

REVIEW PAPER

# Progress and challenges of engineering a biophysical CO<sub>2</sub>-concentrating mechanism into higher plants

Benjamin D. Rae<sup>1,2</sup>, Benedict M. Long<sup>2</sup>, Britta Förster<sup>2</sup>, Nghiem D. Nguyen<sup>1,2</sup>, Christos N. Velanis<sup>3</sup>, Nicky Atkinson<sup>3</sup>, Wei Yih Hee<sup>2</sup>, Bratati Mukherjee<sup>1,2</sup>, G. Dean Price<sup>1,2</sup> and Alistair J. McCormick<sup>3,\*</sup>

<sup>1</sup> Australian Research Council Centre of Excellence for Translational Photosynthesis

<sup>2</sup> Research School of Biology, The Australian National University, 134 Linnaeus Way, Acton ACT 2601, Australia

<sup>3</sup> SynthSys and Institute of Molecular Plant Sciences, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3BF, UK

\* Correspondence: [alistair.mccormick@ed.ac.uk](mailto:alistair.mccormick@ed.ac.uk)

Received 1 February 2017; Editorial decision 29 March 2017; Accepted 28 March 2017

Editor: Howard Griffiths, University of Cambridge

## Abstract

Growth and productivity in important crop plants is limited by the inefficiencies of the C<sub>3</sub> photosynthetic pathway. Introducing CO<sub>2</sub>-concentrating mechanisms (CCMs) into C<sub>3</sub> plants could overcome these limitations and lead to increased yields. Many unicellular microautotrophs, such as cyanobacteria and green algae, possess highly efficient biophysical CCMs that increase CO<sub>2</sub> concentrations around the primary carboxylase enzyme, Rubisco, to enhance CO<sub>2</sub> assimilation rates. Algal and cyanobacterial CCMs utilize distinct molecular components, but share several functional commonalities. Here we outline the recent progress and current challenges of engineering biophysical CCMs into C<sub>3</sub> plants. We review the predicted requirements for a functional biophysical CCM based on current knowledge of cyanobacterial and algal CCMs, the molecular engineering tools and research pipelines required to translate our theoretical knowledge into practice, and the current challenges to achieving these goals.

**Key words:** Algae, carboxysome, cyanobacteria, photosynthesis, pyrenoid, Rubisco, transporter.

## Introduction

Generating new C<sub>3</sub> crop varieties with increased yields is critical for safeguarding future food security (Long *et al.*, 2015). To achieve this goal, much research has focused on improving the efficiency of photosynthetic CO<sub>2</sub> uptake by the primary carboxylase, Rubisco. Rubisco might be the most important limiting component that restricts photosynthetic productivity (Parry *et al.*, 2013). Compared with other enzymes, Rubisco has a slow turnover rate ( $k_{\text{cat}}^{\text{c}}$ ) and a relatively low affinity for CO<sub>2</sub>. Nonetheless, Rubisco in higher plants has a  $K_{\text{m}}$  for CO<sub>2</sub> at ambient O<sub>2</sub> ( $K_{\text{c}}^{\text{air}}$ ) close to the CO<sub>2</sub> concentration in a C<sub>3</sub> leaf mesophyll cell (Galmés *et al.*, 2014), indicating functional adaptation to its operating environment. Rubisco

also catalyses a competitive reaction with O<sub>2</sub> (~30% of activity is apportioned to oxygenation in C<sub>3</sub> plants), leading to substantial metabolic costs associated with photorespiration (Sharkey, 2001). Thus, increasing the carboxylation efficiency of Rubisco and reducing photorespiration are important strategies for improving yields in C<sub>3</sub> crops (Betti *et al.*, 2016; Sharwood, 2017).

Photosynthetic organisms have evolved to compensate for the slow and promiscuous nature of Rubisco. Different strategies have emerged through evolutionary history as organisms adapted to changes in both atmospheric gas composition and growth environment (Shih, 2015). For example,

C<sub>3</sub>, C<sub>4</sub>, and CAM (Crassulacean acid metabolism) modes of photosynthesis have evolved to occupy specialist niches in terrestrial environments. C<sub>3</sub> plants predominate in temperate ecosystems (Woodward *et al.*, 2004) and have compensated for the limitations of Rubisco by maintaining high stromal Rubisco concentrations. Thus, C<sub>3</sub> plants maximize CO<sub>2</sub> capture by increasing the likelihood of carboxylation events within the leaf at the cost of higher rates of photorespiration and water use. C<sub>4</sub> plants dominate tropical/subtropical grasslands (Woodward *et al.*, 2004; Edwards *et al.*, 2010) and have adapted a biochemical CO<sub>2</sub>-concentrating mechanism (CCM), typically shared between two adjacent leaf cell types (mesophyll and bundle sheath). C<sub>4</sub> plants utilize cytosolic phosphoenolpyruvate carboxylase in the mesophyll to perform biochemical fixation of inorganic carbon (C<sub>i</sub>) as the bicarbonate ion (HCO<sub>3</sub><sup>-</sup>) into four-carbon organic acids, which are transported to bundle sheath chloroplasts where Rubisco is primarily localized for subsequent decarboxylation and CO<sub>2</sub> fixation. The biochemical CCMs found in CAM species utilize temporal separation to capture CO<sub>2</sub> as malic acid within the mesophyll vacuole during the night followed by decarboxylation and CO<sub>2</sub> fixation by Rubisco during the day. As an adaptation to drought stress, CAM species can facilitate CO<sub>2</sub> fixation while stomata are closed, ensuring minimal water loss (Keeley and Rundel, 2003).

In aquatic environments, additional limitations on photosynthesis are imposed by the relatively low solubility and diffusion of CO<sub>2</sub> through water. Furthermore, the effects of temperature and pH/ionic strength on CO<sub>2</sub> solubility and the equilibrium ratios of carbonate species (H<sub>2</sub>CO<sub>3</sub>, HCO<sub>3</sub><sup>-</sup>, and CO<sub>3</sub><sup>2-</sup>) are such that supply of C<sub>i</sub> is both limited and sometimes highly variable (Mangan *et al.*, 2016). Capture of C<sub>i</sub> in the form of the less membrane permeable, yet usually more predominant, bicarbonate ion is a major driving force in CCMs found in aquatic photosynthetic organisms. Algae and cyanobacteria have evolved highly efficient biophysical CCMs that can enhance CO<sub>2</sub> levels around Rubisco by up to 1000-fold, making these the most effective biological C<sub>i</sub> uptake systems as yet identified in nature (Wang *et al.*, 2015). Their success is evident in that no less than 80% of photosynthetic C<sub>i</sub> uptake in aquatic environments is driven by biophysical CCMs, which accounts for roughly half of global primary productivity (Raven and Beardall, 2015).

#### *Increased ambient CO<sub>2</sub> improves crop yield*

Several decades of experimental data clearly show that C<sub>3</sub> crops grown under elevated CO<sub>2</sub> concentrations consistently have higher yields, demonstrating that productivity and growth efficiencies are improved by increasing CO<sub>2</sub> levels at the Rubisco active site (Leakey *et al.*, 2009). These findings are complemented by the numerous examples in nature where CCMs have evolved to enhance the operating efficiency of Rubisco. Thus, introduction of CCM components into C<sub>3</sub> crops using new synthetic biology-based approaches is considered a promising strategy to increase photosynthetic productivity and environmental resilience (Furbank, 2016; Long *et al.*, 2016; Meyer *et al.*, 2016; Schuler *et al.*, 2016; Wang

*et al.*, 2017). The biophysical CCMs found in algae and cyanobacteria are particularly attractive due to their high levels of efficiency and establishment in a single cell.

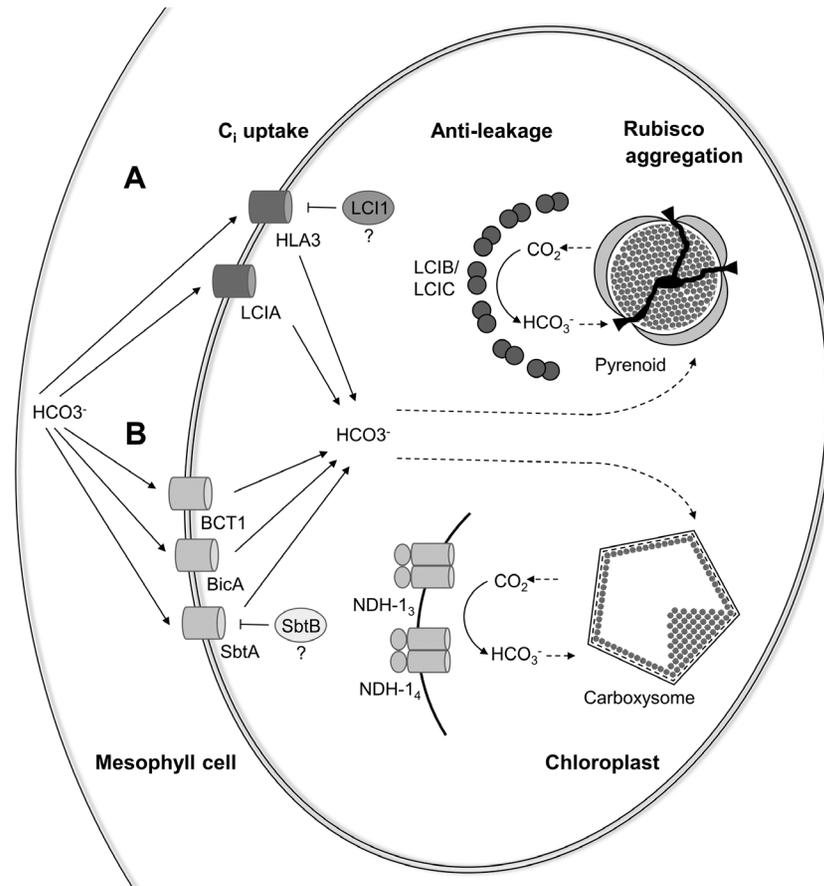
#### *Common characteristics of biophysical CCMs*

Biophysical CCMs have four common characteristics: (i) energy-dependent uptake of C<sub>i</sub>; (ii) dehydration of bicarbonate to CO<sub>2</sub> via a carbonic anhydrase (CA) located close to Rubisco; (iii) assembly of Rubisco into subcellular microcompartments (cyanobacterial carboxysomes) or regions (algal pyrenoids); and (iv) systems to minimize CO<sub>2</sub> diffusion from the site of carboxylation (Fig. 1). Carboxysomes are essential CCM components in photosynthetic cyanobacteria, whereas pyrenoids are found in most but not all eukaryotic microalgae (Badger *et al.*, 1998; Giordano *et al.*, 2005). Furthermore, carboxysomes are permanent subcellular features, whereas pyrenoid assembly/disassembly is circadian and C<sub>i</sub> dependent (Mitchell *et al.*, 2014). Accordingly, key components of the cyanobacterial CCM are well characterized, and detailed models exist for predicting the best routes for introduction into C<sub>3</sub> chloroplasts (Price *et al.*, 2011a, 2013; McGrath and Long, 2014). However, the factors that regulate activity and assembly are still unclear, and the focus of current investigation. In contrast, eukaryotic algal CCMs are relatively less well understood overall and probably more complex due to the nested-compartmentalized structure of the eukaryotic cell (Wang *et al.*, 2015).

In this review, we highlight recent progress and current challenges of engineering biophysical CCMs into C<sub>3</sub> plants. Specifically, we outline progress towards the incorporation of inorganic carbon transporters into the chloroplastic inner membrane, utilizing both algal and cyanobacterial genes as sources, and emphasizing the need for further knowledge of their activation and energization. We also draw attention to current knowledge of the Rubisco microbodies (pyrenoids) and microcompartments (carboxysomes) and how this informs us of their potential reconstruction within C<sub>3</sub> plant chloroplast hosts. The review also touches on the costs of microcompartment construction within the chloroplast and synthetic approaches to generating a chloroplastic CCM.

### **Model-guided strategies for improvement of C<sub>3</sub> photosynthesis**

Mathematical modelling reveals that modest photosynthetic benefits might be achievable by incorporating bicarbonate transporters and their ancillary systems into the chloroplast envelope of C<sub>3</sub> plants (Price *et al.*, 2011a, 2013; McGrath and Long, 2014). Bicarbonate transporters could be harnessed independently or in concert from different species (e.g. algae and cyanobacteria) to boost C<sub>i</sub> levels in the chloroplast. However, native chloroplastic CA activities, which are highly abundant in C<sub>3</sub> plants (Meyer *et al.*, 2016), would need to be removed or significantly re-regulated to avoid dissipation of any C<sub>i</sub> gradient generated by active transport (Price *et al.*, 2013; McGrath and Long, 2014). These features are



**Fig. 1.** Introduction of key algal and cyanobacterial CCM components into  $C_3$  plant chloroplasts. Algal (A, dark grey) and cyanobacterial (B, light grey) components are involved in three unique features: inorganic carbon ( $C_i$ ) uptake (HLA3, LCIA, BCT1, BicA, and SbtA), Rubisco aggregation into subcellular organelles (shown as the pyrenoid and carboxysome from algae and cyanobacteria, respectively) and anti-leakage systems (LCIB/LCIC, NDH-1<sub>3</sub>, and NDH-1<sub>4</sub>) to minimize loss of  $CO_2$ . Putative regulators of HLA3 (LCI1) and SbtA (SbtB) activity are indicated. Together, components of the cyanobacterial and algal CCMs could constitute a functional synthetic biophysical CCM. We intentionally present components of the algal and cyanobacterial CCMs together as constituents of both systems could be used in concert to achieve the intended goal of enhanced  $CO_2$  concentration at the site of carbon fixation. For detailed recent reviews of these CCMs, see Wang *et al.* (2015), Long *et al.* (2016), Meyer *et al.* (2016), and Rae *et al.* (2013a, b).

a necessary pre-condition for downstream introduction of Rubisco-aggregating factors, such as carboxysome and pyrenoid-like structures. Additional aggregation of Rubisco into carboxysomes is predicted to enhance the photosynthetic potential of higher plants by ~60% (McGrath and Long, 2014). Earlier models have also indicated that the presence of a pyrenoid significantly improves the maintenance of elevated  $CO_2$  concentrations around Rubisco (Badger *et al.*, 1998), suggesting that pyrenoid-like structures would likewise be beneficial in  $C_3$  plants in the presence of a functional chloroplastic  $C_i$  uptake system.

McGrath and Long (2014) predicted that the further addition of complex thylakoid membrane-localized systems to reduce  $CO_2$  leakage (multisubunit NDH-1 complexes, NDH-1<sub>3</sub> and NDH-1<sub>4</sub>) would only marginally improve photosynthesis. Thus far, the technical challenges associated with localizing multigene complexes on the thylakoid membrane have precluded introduction of NDH-1<sub>3</sub> and NDH-1<sub>4</sub> as a major engineering focus. The increasing availability of component libraries for gene expression and development of rapid assembly approaches for multigene constructs may simplify this goal (Patron, 2016). An alternative  $CO_2$  recapture system is found in algae, which is conceivably simpler, and

consists of a stromal soluble complex of two proteins (LCIB and LCIC in *Chlamydomonas reinhardtii*) (Jin *et al.*, 2016). Expression of LCIB/LCIC may represent the simplest means of achieving the  $CO_2$  recapture step of the CCM process outlined by McGrath and Long (2014).

## Introducing bicarbonate transporters into the chloroplast envelope

### Cyanobacterial bicarbonate transporters

To date, five  $C_i$  uptake systems have been identified in cyanobacteria that differ in expression patterns (constitutive versus inducible) as well as substrate affinity and uptake rates (Fig. 1) (Price *et al.*, 2008, 2011a). Many cyanobacterial genomes possess only a subset (Rae *et al.*, 2011), suggesting that the full complement of  $C_i$  uptake systems is not required to accumulate sufficient intracellular  $C_i$  in their natural environments. Common to all systems is that  $C_i$  uptake is an active process relying on various forms of energization.

While models suggest that the greatest improvement in  $C_3$  photosynthetic efficiency would require the introduction of a complete CCM into higher plant chloroplasts, they also

indicate that the introduction of one or two active bicarbonate transporters alone can increase photosynthetic efficiency significantly (Price *et al.*, 2011a; McGrath and Long, 2014). Light-saturated CO<sub>2</sub> assimilation rates could be increased by 9% by using a single cyanobacterial bicarbonate transporter [Bicarbonate Transporter A (BicA)], while utilizing all cyanobacterial bicarbonate transporters together could improve photosynthetic rates by as much as 16% (McGrath and Long, 2014). Notably, McGrath and Long (2014) found that introducing bicarbonate transporters into the plasma membrane would result in no additional benefits for photosynthetic efficiency. Observations in field-based CO<sub>2</sub> enrichment studies for C<sub>3</sub> crops show, on average, a 17% increase in biomass accumulation based on improvements in photosynthetic efficiency (Long *et al.*, 2006). Modelled photosynthetic improvements using cyanobacterial bicarbonate transporters alone (i.e. a partial biophysical CCM) indicate that active bicarbonate transporters could translate into a 5–10% increase in biomass on a daily basis. However, the compounding interest effect could easily lead to a larger increase in biomass accumulated at harvest time. This effect has already been observed with recent strategies for crop plant photosynthetic improvement (Simkin *et al.*, 2015; Kromdijk *et al.*, 2016), where relatively small changes in photosynthetic efficiency led to significant increases in endpoint biomass production. Furthermore, downstream addition of carboxysomes or pyrenoids, and pluralist approaches drawing from all photosynthetic improvement strategies will provide additive improvement.

Three different transporters located in the cyanobacterial plasma membrane actively take up bicarbonate. The Bicarbonate Transporter 1 (BCT1) complex, encoded by the *cmpABCD* operon, is an ATP-binding cassette (ABC) type transporter [traffic ATPase family (Omata *et al.*, 1999)]. Even though fewer subunits constitute a functional BCT1, the same multigene engineering challenges exist as for NDH-1 complexes. Therefore, the two single-gene bicarbonate transporters, BicA and Sodium-dependent Bicarbonate Transporter A (SbtA), encoded by the *bicA* and *sbtA* genes, respectively, are considered the most suitable candidates for expression in higher plant chloroplast envelopes (Shibata *et al.*, 2002; Price *et al.*, 2004). BicA belongs to the large SLC26A/SulP type protein family of anion carriers and channels, present in many prokaryotic and eukaryotic organisms. SbtA is widely distributed amongst cyanobacteria and found in some bacteria, but is absent from eukaryotes. Structurally, SbtA has no resemblance to other known transport proteins (Price and Howitt, 2014). The activity of both transporters depends on a cell inward-directed sodium ion (Na<sup>+</sup>) gradient, and most probably both function as HCO<sub>3</sub><sup>-</sup>/Na<sup>+</sup> symporters. However, direct evidence is lacking and the stoichiometry of transported ions needs to be determined. Based on partial homology to low-resolution structures of bacterial SulP proteins (Compton *et al.*, 2011), it is likely that BicA forms homodimers, while there is some proteomic evidence for tetramers of SbtA (Zhang *et al.*, 2004). Nevertheless, the active configuration and assembly mechanisms of BicA and SbtA have not yet been elucidated. Important structural information about the expression of the mature protein in the plasma membrane

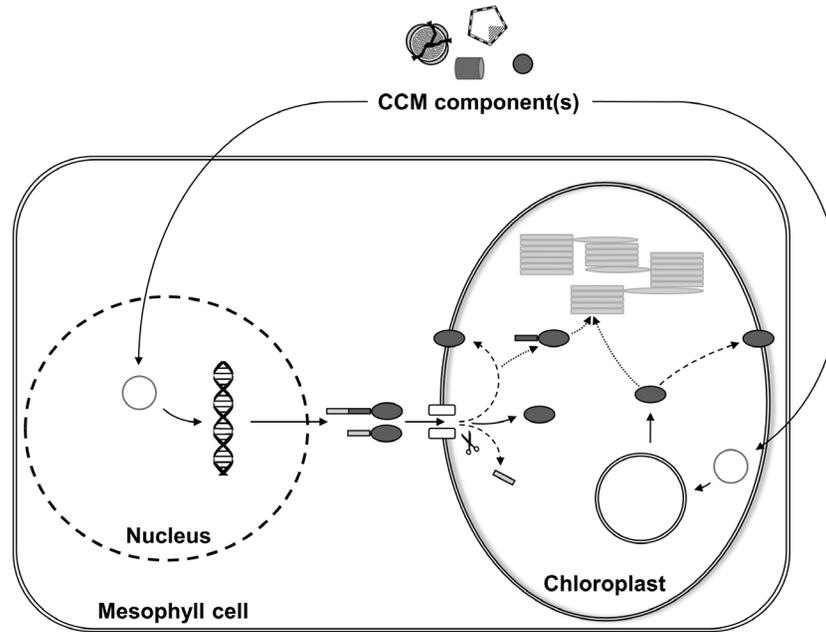
has been gained from topology mapping and superimposition of known structures of similar proteins. The N- and C-termini of BicA have been shown to be located in the cytoplasm (Shelden *et al.*, 2010), whereas both N- and C-termini of SbtA are unusual in that they protrude into the periplasmic space (Price *et al.*, 2011b).

#### Expression of cyanobacterial membrane transporters in planta

Both BicA and SbtA have been successfully introduced into C<sub>3</sub> plants using chloroplast and nuclear transformation approaches (Pengelly *et al.*, 2014; Rolland *et al.*, 2016; Uehara *et al.*, 2016). Introduction of transgenes directly into the chloroplast genome may negate the need for a chloroplast targeting signal for insertion into the chloroplast inner envelope membrane (IEM) (Fig. 2). On the other hand, the vast majority of IEM-targeted proteins are encoded in the nuclear rather than the chloroplast genome, with only a few known exceptions: Ycf1/Tic214 identified in *Arabidopsis thaliana* (Kikuchi *et al.*, 2013), CemA in pea (Sasaki *et al.*, 1993), and its orthologue Ycf10 in *Chlamydomonas* (Rolland *et al.*, 1997).

Insertion of chloroplast- and nuclear-encoded proteins into the thylakoid membrane is thought to occur via four pathways. The first three are analogous to bacterial systems, and require additional factors [i.e. the signal recognition particle (SRP), secretory (Sec), and twin-arginine translocase (Tat)/ΔpH-dependent pathways (Aldridge *et al.*, 2009)]. The fourth pathway is less common and relies on spontaneous insertion. For example, several membrane-spanning proteins are able to insert unassisted into the thylakoid membrane [e.g. photosystem subunits PsbY, PsaG, and PsaK (Zygodlo *et al.*, 2006)]. It is important to note though that those proteins contain at most four membrane spans, and therefore require significantly less threading of protein domains across the lipid bilayer than insertion of BicA and SbtA, which contain 14 and 10 transmembrane/membrane integral domains, respectively. Detailed analysis of insertion requirements of PsaG suggests that the positively charged stromal loop region between the membrane spans is essential. This is consistent with the ‘positive inside rule’ proposed for spontaneous insertion of membrane-integral proteins in bacteria (Dalbey *et al.*, 2011).

It is not known how membrane-integral proteins such as BicA and SbtA are sorted into the plasma membrane in cyanobacteria, or in the special cases of Ycf1/Tic214 and CemA from the stroma to the IEM in chloroplasts. Based on topology mapping, BicA and SbtA may follow the positive inside rule, as positive charges are present on their cytoplasm-exposed domains (Shelden *et al.*, 2010; Price *et al.*, 2011b). Hence, the chloroplast transformation approach relies on either spontaneous insertion of BicA or SbtA into the IEM, or sufficient compatibility of these proteins with other chloroplastic insertion pathways. Given the prokaryotic origin of chloroplasts, it is conceivable that the membrane protein insertion pathways have not completely diverged from the ancestral cyanobacterial system. Mature



**Fig. 2.** Transformation strategies for engineering a biophysical CCM into higher plant chloroplasts. Stable nuclear and chloroplast genome transformation are both feasible strategies for introducing CCM components. Chloroplast transformation (right) is advantageous for expression of stromal CCM components and gene operons (e.g. [Lin et al., 2014b](#)), but is less predictable for delivery to the envelope (dashed arrow) and thylakoid membrane (dotted arrow) ([Pengelly et al., 2014](#)). Robust chloroplast transformation techniques are currently established in only a very limited number of plant species ([Bock, 2014](#)). Nuclear transformation (left) requires that CCM components be modified to carry an appropriate chloroplast transit signal sequence. The latter approach has successfully delivered CCM components to the stroma and inner envelope membrane (IEM) ([Atkinson et al., 2016, 2017](#); [Rolland et al., 2016](#); [Uehara et al., 2016](#)) and could be used also for targeting to the thylakoid membrane or lumen (as indicated), although additional processing would be required ([Aldridge et al., 2009](#)). Multigene assembly tools offer the opportunity to deliver several CCM components to different locations following a single transformation step ([Patron, 2016](#)). A disadvantage of nuclear transformation is the current limited understanding of the transit signal sequences required for correct orientation of transmembrane CCM components (e.g. bicarbonate transporters) in the IEM.

BicA protein of the expected size was detected throughout the life of the plant both in the thylakoid and in the IEM fraction in *Nicotiana tabacum* (tobacco hereafter) chloroplast transformants ([Pengelly et al., 2014](#)), whereas this approach has been unsuccessful for SbtA, as protein did not accumulate (BF, unpublished data). In both cases, the original cyanobacterial coding sequences were transferred to the same region of the inverted repeat in the chloroplast genome and expressed from the same promoter. Evidently, there are unknown factors that differentially affect synthesis, stability, and targeting of BicA and SbtA. Physiological characterization of mature BicA-expressing plants suggested the BicA protein was inactive ([Pengelly et al., 2014](#)). It remains unclear whether BicA was inserted into the IEM incorrectly or whether other BicA-specific regulatory factors (e.g. post-translational modifications and/or interacting proteins or metabolites) are missing in chloroplasts. The latter seems more likely, as BicA inserts correctly into the plasma membrane when expressed in *Escherichia coli*, evidenced by successful topology mapping ([Shelden et al., 2010](#)). Active uptake of radioactively labelled bicarbonate mediated by BicA could not be detected using *E. coli* as a heterologous expression system ([Du et al., 2014](#)). Likewise, BicA showed no bicarbonate uptake activity when expressed in *Xenopus laevis* oocytes (BF and M.D. Parker, unpublished data). Taken together, this emphasizes the absolute necessity to understand post-translational regulation/activation mechanisms of BicA to make it a useful candidate for  $C_i$  uptake.

#### *Nuclear expression and chloroplast targeting of cyanobacterial bicarbonate transporters*

The alternative to transgene expression in the chloroplast is nuclear transformation. This approach requires synthesis of precursors of BicA and SbtA that have peptide additions that facilitate correct targeting and insertion into the chloroplast IEM, via a stop-transfer or post-import mechanism ([Viana et al., 2010](#)). Even though our understanding of these components and processes is growing, we are far from being able to predict the right signal peptide for each protein. Using a trial and error approach, a cleavable transit peptide (cTP) followed by a potentially uncleaved membrane protein leader sequence (MPL) of known chloroplast IEM-targeted proteins was fused to BicA or SbtA, resulting in successful insertions in the IEM of *Nicotiana benthamiana* and *Arabidopsis* ([Rolland et al., 2016](#); [Uehara et al., 2016](#)). [Rolland et al. \(2016\)](#) identified two *Arabidopsis* cTP+MPL sequences, *AtPLGG1* and *AtHP59*, which effectively direct the green fluorescent protein (GFP)-tagged bicarbonate transporters to the IEM in *N. benthamiana*. The *AtPLGG1* leader can correctly target both SbtA and BicA, while the *AtHP59* leader is only effective in directing SbtA. This demonstrates that the effectivity of a leader sequence is highly dependent on the protein to be delivered to the IEM, and must be determined for each foreign membrane protein separately.

There is some evidence that information for protein insertion into the IEM may be contained elsewhere in the mature protein sequence ([Theg and Scott, 1993](#); [Okawa et al., 2014](#);

Oh and Hwang, 2015; Endow *et al.*, 2016). Uehara *et al.* (2016) were able to direct a *Staphylococcus* protein A-tagged BicA and SbtA chimera to the IEM in Arabidopsis using the cTP and the mature IEM-located Arabidopsis Cor413im1 protein. One possibility is that an MPL may be contained in the Cor413im1 sequence leading up to the first transmembrane domain, suggesting that a similar targeting structure (cTP+MPL) was generated as in Rolland *et al.* (2016). Interestingly, BicA was also targeted to the IEM when Cor413im1 was fused to the C-terminus of BicA. All IEM-inserted chimeras tested by Rolland *et al.* (2016) had an N-terminal cTP+MPL fusion, and domain swapping experiments showed that MPL sequences had to be in the correct orientation and matched with a compatible cTP.

If we assume that the IEM targeting signal is not, or only partially, removed from the transporter upon IEM insertion, the positioning of the IEM signal may have important implications with respect to the functionality of the transporters, particularly for SbtA. In *E. coli* (Du *et al.*, 2014) and in cyanobacteria (BF and GDP, unpublished data), we have shown previously that SbtA activity is abolished by adding epitope tags, GFP or yellow fluorescent protein (YFP). Furthermore, bicarbonate transporters can only function effectively on the chloroplast envelope if they are inserted in the correct orientation. From the work of Uehara *et al.* (2016) using detergent and protease treatments, and Rolland *et al.* (2016) using visualization of BicA::GFP and SbtA::GFP fusions by confocal microscopy, we know that both transporters were integrated into the IEM but not their orientation or exact N-terminal amino acid sequence.

#### Characterization of cyanobacterial bicarbonate transporter activity in transgenic plants

Achieving expression of BicA and SbtA in the IEM was a significant milestone. Future research now faces two key challenges. First, bicarbonate transporter activity on the IEM cannot be readily measured directly. Researchers currently rely on measurements of leaf photosynthetic traits [e.g. the CO<sub>2</sub> compensation point and carbon isotope discrimination ( $\Delta^{13}\text{C}$ )] and plant biomass accumulation as indicators of a change in C<sub>i</sub> uptake capacity (Pengelly *et al.*, 2014). These measurements are largely indirect and assume that bicarbonate uptake will lead immediately to an improvement in photosynthetic efficiency, as suggested by modelling (Price *et al.*, 2011a; McGrath and Long, 2014). However, there are several confounding factors that may mask or counteract phenotypic effects of bicarbonate uptake, including (i) the expression levels required for an active bicarbonate transporter to elevate C<sub>i</sub> in a plant chloroplast are not yet clear; (ii) photosynthetic improvement may not translate into an increase in growth due to other limiting factors (e.g. light, temperature, or source–sink dynamics) under the given growth conditions; and (iii) transporter activity may have detrimental side effects on native metabolism. The latter is a potential issue for BicA and SbtA, as both transporters are HCO<sub>3</sub><sup>-</sup>/Na<sup>+</sup> symporters. An increased influx of Na<sup>+</sup> into the chloroplast needs to be matched with an Na<sup>+</sup>/H<sup>+</sup> antiporter activity on the IEM to

maintain ion homeostasis and prevent a build-up of Na<sup>+</sup> from the stroma. Higher plants have endogenous HCO<sub>3</sub><sup>-</sup>/Na<sup>+</sup> symporters in the IEM, such as the well-characterized NHD1 transporter in Arabidopsis (Müller *et al.*, 2014), but whether native transporters provide sufficient Na<sup>+</sup> extrusion or additional capacity is required in the presence of BicA and SbtA remains unclear (Price, 2011). Direct measurements of bicarbonate/CO<sub>2</sub> uptake can be performed on isolated chloroplasts, either by monitoring changes in photosynthetic gas exchange using membrane-inlet mass spectrometry (MIMS), or by measuring uptake of radioactively labelled [<sup>14</sup>C]bicarbonate into chloroplast using the silicon oil centrifugation method. Neither technique is simple and requires isolation of intact and functional chloroplasts, which is difficult for many plant species. One solution may be to develop improved MIMS techniques to detect C<sub>i</sub> uptake activity at the leaf level.

#### Transport regulation of cyanobacterial bicarbonate transporters *ex situ*

A further key challenge is to overcome the apparent inactivity of the bicarbonate transporters *ex situ* and, potentially, to regulate activity when expressed in plants. The observed lack of activity in plants, particularly of BicA, is not yet understood. Requirements for a sufficient Na<sup>+</sup> gradient to drive activity are met—cytosolic levels of Na<sup>+</sup> are estimated at 1–3 mM, which is in the range of Na<sup>+</sup> concentration required for half-maximal activity of bicarbonate uptake by BicA and SbtA (Price *et al.*, 2013). Elevated Na<sup>+</sup> in the chloroplast stroma is unlikely to be problematic due to the presence of Na<sup>+</sup>/H<sup>+</sup> antiporters, and Na<sup>+</sup>-coupled transporters, in the chloroplast membrane (Price *et al.*, 2013). Furthermore, no serious effect of increased cytoplasmic bicarbonate is expected (Price *et al.*, 2011a, 2013). One possible cause for a lack of function is incorrect insertion in the IEM. A bimolecular fluorescence complementation (BiFC) system could be suitable for determining the orientation of BicA and SbtA *in planta* (Waadt *et al.*, 2008). Previously, BiFC-based approaches have been successfully used to determine the orientation of TOC75 in the outer chloroplast envelope membrane (OEM) in Arabidopsis (Chen *et al.*, 2016).

In cyanobacteria, BicA and SbtA activity are regulated by environmental cues. One important feature is the rapid inactivation of bicarbonate transport in the dark (BF, unpublished data), which is likely to avoid futile C<sub>i</sub> uptake and energy wastage when photosynthetic CO<sub>2</sub> fixation does not occur. The underlying molecular mechanisms that activate/deactivate BicA and SbtA are unknown. Heterologous expression of SbtA in *E. coli* has shown that SbtA is constitutively active in the absence of photosynthetic CO<sub>2</sub> fixation (Du *et al.*, 2014) However, when co-expressed with a companion protein of unknown function (SbtB), SbtA is inactive. The genome of every sequenced cyanobacterial species containing *sbtA* inextricably has a *sbtB* gene in the same operon or in close proximity. The SbtB protein shows 21% identity to the cyanobacterial PII/GlnB proteins that are regulators of nitrogen and carbon metabolism in bacteria. Structurally, SbtB trimers (unpublished crystal structure: ncbi.nlm.nih.gov

structure 3DFE) fold in a similar way to trimeric GlnB and GlnK (a PII homologue). In *E. coli*, the ammonium channel AmtB is inactivated upon binding of GlnK (Conroy *et al.*, 2007). Therefore, we postulate a similar 'lock-down' function for SbtB when bound to SbtA, which is further corroborated by *in vitro* interaction of SbtA with SbtB in *E. coli* protein extracts (Du *et al.*, 2014).

SbtA proteins from at least two different cyanobacterial species bind efficiently to their cognate, immobilized SbtB orthologues during immobilized metal affinity chromatography (IMAC). However, analysis of cyanobacterial mutants lacking SbtB suggests that additional post-translational mechanisms are in place to regulate SbtA activity (BF and GDP, unpublished results). Bioinformatics revealed that highly conserved, putative phosphorylation sites are present in BicA, SbtA, and SbtB, which suggests that phosphorylation could be involved in regulation. Furthermore, high molecular weight protein complexes containing SbtA or BicA can be detected in native protein extractions from cyanobacteria, raising the question of whether interactions with other proteins and/or metabolic intermediates determine BicA or SbtA activity (BF and GDP, unpublished results). Clearly, we need to improve our understanding of post-translational regulation of C<sub>i</sub> uptake in cyanobacteria in order to engineer active forms of BicA and SbtA into higher plant chloroplasts.

#### Algal bicarbonate transporters

The algal CCM is most well understood in the model green alga *Chlamydomonas*. The *Chlamydomonas* CCM is dynamic and only induced in the light and under limiting CO<sub>2</sub>, which results in the expression of many hundreds of CCM-associated genes to sustain photosynthetic growth (Miura *et al.*, 2004; Brueggeman *et al.*, 2012; Fang *et al.*, 2012). Several of these genes have been confirmed as essential or contributing to the functionality of the algal CCM through molecular, genetic, and physiological studies of CCM mutants (Wang *et al.*, 2015).

Two putative C<sub>i</sub> transporters have been identified as primary targets for engineering C<sub>i</sub> uptake into higher plants using the algal CCM (Fig. 1). The currently accepted CCM model based on *Chlamydomonas* indicates that active C<sub>i</sub> uptake under limiting CO<sub>2</sub> is driven by High Light-Activated 3 (HLA3) and Low Carbon-inducible A (LCIA), located on the plasma membrane and chloroplast envelope, respectively. The model predicts that a third C<sub>i</sub> transporter may be located on the thylakoid membrane to facilitate transport of C<sub>i</sub> into the pyrenoid, but a candidate has yet to be identified. HLA3 is a putative ATP-dependent ABC-type transporter belonging to the multidrug resistance-related protein family that pumps bicarbonate from the periplasm to the cytosol (Im and Grossman, 2002). LCIA is a putative anion transporter and a homologue of the bacterial formate transporter FocA, a symmetric pentamer that closely resembles an aquaporin channel (Wang *et al.*, 2009). Similar to FocA, LCIA forms a protein complex under native conditions (Yamano *et al.*, 2015). Although a C<sub>i</sub> uptake mechanism has yet to be confirmed, LCIA is believed to shuttle cytosolic bicarbonate

into the stroma (Wang and Spalding, 2014). Both HLA3 and LCIA have been localized *in vivo* (Yamano *et al.*, 2015), and capacity for C<sub>i</sub> uptake has been observed *in vitro* in *Xenopus* oocytes (Mariscal *et al.*, 2006; Atkinson *et al.*, 2016).

An additional component of interest is Low Carbon-inducible 1 (LCI1), a small protein localized primarily at the plasma membrane (Ohnishi *et al.*, 2010). LCI1 is highly induced under low CO<sub>2</sub>, and LCI1 overexpression studies have shown enhanced rates of C<sub>i</sub> uptake in *Chlamydomonas*. Nevertheless, topology predictions have indicated that the hydrophobic domains of LCI1 are too short to support a transmembrane channel, at least not in the monomeric form (Meyer and Griffiths, 2013). In addition, C<sub>i</sub> uptake was not observed *in vitro* in *Xenopus* oocytes (AJM, unpublished data). These data suggest that LCI1 is not directly involved in C<sub>i</sub> uptake, but may act as an ancillary component to enhance the activity of other transporters, such as HLA3 (Fig. 1). Notably, no close homologues of LCI1 are present in any other available algal proteome.

Engineering approaches for the introduction of algal CCM components have thus far focused on nuclear transformation, as the key CCM components identified in *Chlamydomonas* are encoded in the nuclear genome (Atkinson *et al.*, 2016). Highlighting one of the potential advantages of a eukaryotic CCM, the majority of algal CCM components appear to localize to the appropriate locations in higher plants without modification of the coding sequence, indicating that higher plants are able to recognize cTPs from *Chlamydomonas*. In particular, LCIA and LCIB/LCIC were able to localize to the chloroplast envelope and stroma, respectively. Atkinson *et al.* (2016) also demonstrated that LCI1 could be re-localized to the chloroplast stroma using a higher plant cTP, but not the envelope, suggesting again that LCI1 may not be capable of traversing membranes. Although initial attempts to re-localize HLA3 to the chloroplast envelope were ineffective, recent attempts have yielded more success (AJM, unpublished data). This again highlights efforts by Rolland *et al.* (2016), where selection of an appropriate transit signal sequence is required for targeting particular proteins to the chloroplast envelope (Fig. 2).

Stable expression of either HLA3 or LCIA has not yet resulted in a significant growth advantage or a measurable change in photosynthetic capacity in higher plants (Atkinson *et al.*, 2016). It is possible that algal transporters may require additional regulation for activity *in planta*, similar to BicA and SbtA (Du *et al.*, 2014; Pengelly *et al.*, 2014). Yamano *et al.* (2015) have recently shown that HLA3 and LCIA act co-operatively when the CCM is active in *Chlamydomonas*, suggesting that both transporters may be required for efficient C<sub>i</sub> uptake. If LCIA is a passive channel for bicarbonate anions rather than an active transporter, C<sub>i</sub> uptake into the stroma would need to proceed against an electrical gradient, as the stromal side of the chloroplast envelope carries a negative charge (Wang and Spalding, 2014). Active C<sub>i</sub> transport at the plasma membrane would be required (e.g. by HLA3) to facilitate a sufficiently high bicarbonate concentration gradient in the cytosol to overcome this limitation. Likewise, HLA3 would be unable to deliver bicarbonate to the chloroplast

effectively in the absence of LCIA. In fact, double insertion mutants for LCIA and HLA3 have a lower affinity for  $C_i$  and slower growth rates compared with HLA3 or LCIA single-insertion mutants in *Chlamydomonas* (Duanmu *et al.*, 2009a; Yamano *et al.*, 2015).

#### *Ancillary requirements to algal bicarbonate transporter function*

It is apparent that additional features are required for a biophysical CCM to function efficiently, in both algae and cyanobacteria, including structures for Rubisco aggregation (discussed below), components to reduce  $C_i$  leakage, and removal of inhibitory CA activities (Meyer *et al.*, 2016). For example, expression of a heterologous CA in the cytosolic compartment of cyanobacteria or mistargeting of the native carboxysomal CA is sufficient to ‘short-circuit’ the CCM, resulting in loss of the ability to accumulate internal  $C_i$  (Price and Badger, 1989a, b). Likewise, in *Chlamydomonas*, removal of putative CA enzymes LCIB and/or CAH3 [an  $\alpha$ -form CA located in the thylakoid lumen and inside the pyrenoid when the CCM is active (Blanco-Rivero *et al.*, 2012; Sinetova *et al.*, 2012)] disrupts the capacity of the algal CCM to function under subambient  $CO_2$  conditions (Wang and Spalding, 2006; Duanmu *et al.*, 2009b; Jin *et al.*, 2016). Notably, disruption of LCIA compromises  $C_i$  uptake and growth in *Chlamydomonas* under very limiting  $CO_2$  levels, where photosynthesis is thought to be driven primarily by active  $HCO_3^-$  uptake, but disruption of both LCIA and LCIB is lethal under the same conditions (Wang and Spalding, 2014; Yamano *et al.*, 2015), suggesting that  $C_i$  transporters may require specific stromal CA activities for a functional CCM.

Engineering and localization strategies for algal  $C_i$  transporters and other CCM components have been largely guided by models for introducing the cyanobacterial CCM into plants (Price *et al.*, 2013; McGrath and Long, 2014). The model of McGrath and Long (2014) predicts that improvements in photosynthesis from localization of cyanobacterial  $C_i$  transporters in the chloroplast envelope are not enhanced by additional  $C_i$  transporter activity on the plasma membrane. However, growing knowledge of the apparent co-dependencies and redundancies in the eukaryote biophysical CCM illustrate the requirement for predictive models specifically focused on the multilayered algal CCM system (Wang *et al.*, 2015). The algal CCM (at least in *Chlamydomonas*) appears to employ different CCM components under different limiting  $C_i$  conditions. Improved understanding of the molecular mechanisms and catalytic characteristics of algal CCM components, such as LCIA and HLA3, will aid the development of models to test whether retention of algal CCM components in the plasma membrane and/or chloroplast envelope is a tractable strategy in plants. Such models should also include the critical role of pH that has been demonstrated for the cyanobacterial CCM (Mangan *et al.*, 2016).

The generation of an elevated bicarbonate pool within the chloroplast by way of active bicarbonate transporters is the first step toward introducing a complete biophysical CCM into higher plant chloroplasts. While these transporters

alone theoretically lead to enhanced photosynthetic performance (Price *et al.*, 2011a), the addition of further CCM components, such as a Rubisco microcompartment (i.e. a carboxysome) or microbody (i.e. a pyrenoid), will allow for progression towards the maximum theoretical performance (McGrath and Long, 2014).

## **Carboxysome expression in higher plants: challenges and current status**

### *Two convergently evolved carboxysome subtypes with different biogenesis*

Carboxysomes are protein-bound prokaryotic organelles that function to separate the  $CO_2$  fixation step of photosynthesis from the cytoplasmic milieu (Rae *et al.*, 2013a, b). It is thought that enhancement of  $CO_2$  fixation is conferred by a selectively permeable shell which provides some level of diffusional resistance to  $CO_2$ , while allowing free transit of other Rubisco substrates and products (Menon *et al.*, 2010; Mangan *et al.*, 2016). These microbodies have evolved at least twice, once in diverse proteobacterial species, and incorporating Form-1A Rubisco, and again in cyanobacteria, incorporating Form-1B Rubisco—the higher plant Rubisco lineage (Tabita, 1999; Badger and Bek, 2008). These convergent forms of carboxysome are named for the phyletic affinity of their Rubisco enzymes:  $\alpha$ -carboxysomes are the proteobacterial type, containing the proteobacterial Rubisco Form-1A, and  $\beta$ -carboxysomes are the cyanobacterial type, and encapsulate Rubisco Form-1B (Badger and Bek, 2008). To complicate matters somewhat, the marine/oligotrophic cyanobacterial clade encompassing the genera *Cyanobium*, *Synechococcus*, and *Prochlorococcus* gained the  $\alpha$ -carboxysome by horizontal gene transfer (Rae *et al.*, 2011). Accordingly, this lineage of cyanobacteria is referred to as  $\alpha$ -cyanobacteria, and the remainder are  $\beta$ -cyanobacteria (Badger *et al.*, 2002).

Bioinformatic and structural comparison of the proteins of  $\alpha$ - and  $\beta$ -carboxysomes has shown that both contain phylogenetically related, albeit divergent, protein shell structures (Heinhorst *et al.*, 2014). However, the internal enzymes of each carboxysome form are quite different. Each has a unique type of CA, and the principal structural proteins of each form are not evolutionarily related. The fine detail of carboxysome structure is not our focus, and we divert the reader’s attention to recent comprehensive reviews of carboxysome structure, function, and evolution (Heinhorst *et al.*, 2014; Kimber, 2014; Rae *et al.*, 2013a, b). To summarize briefly:  $\alpha$ -carboxysomes are assembled from proteins with the Cso nomenclature (Carboxysome), where the CsoS2 protein is responsible for packaging the  $\alpha$ -carboxysomal CA (CsoSCA) and Form-1A Rubisco within shells composed of multiple paralogues of CsoS1, and CsoS4 structural proteins (Cai *et al.*, 2015).  $\beta$ -Carboxysomes are assembled from proteins with the Ccm nomenclature ( $CO_2$  concentrating mechanism) whereby at least two proteins derived from the *ccmM* gene assemble the  $\beta$ -carboxysomal CA (CcmM, CcaA, or both, depending on the species) along with Form-1B Rubisco (Long *et al.*, 2007; Cot *et al.*, 2008), and these arrayed enzymes are

bridged by CcmN to shell structures formed by paralogues of CcmK, CcmO, CcmP, and CcmL structural proteins (Kinney *et al.*, 2012; Rae *et al.*, 2012; Cai *et al.*, 2013). These represent two fundamentally different biogenesis pathways whereby  $\beta$ -carboxysomes primarily rely on the formation of a luminal Rubisco complex (Cameron *et al.*, 2013; Chen *et al.*, 2013) while empty  $\alpha$ -carboxysome structures can form in the absence of its Rubisco cargo (Menon *et al.*, 2008; Rae *et al.*, 2013a).

### Carboxysome-mediated metabolism

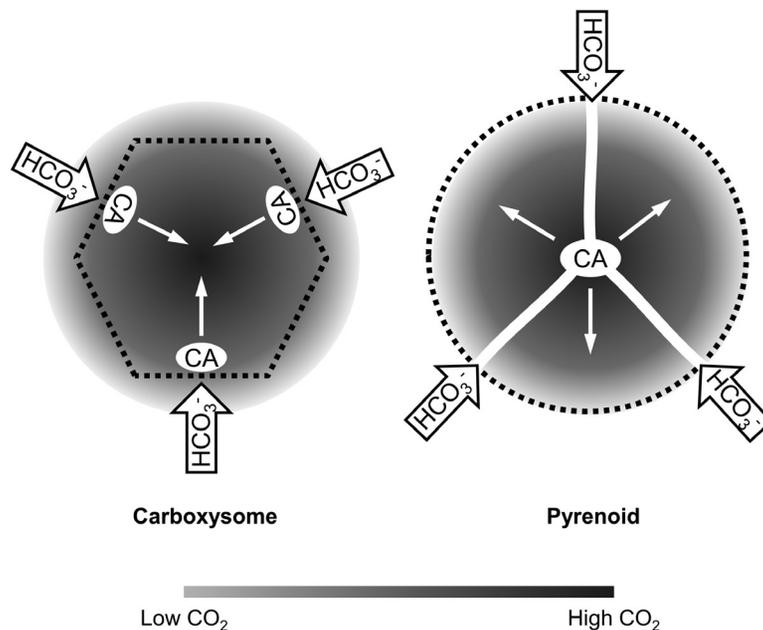
In theoretical structure–function models, the cyanobacterial carboxysome operates as a discrete physico-metabolic unit, and the two types of carboxysome operate in essentially the same way (Fig. 3) (Whitehead *et al.*, 2014). It is helpful to view the steady-state kinetics of carboxysomes as sequential metabolic steps, whereby intracellular  $C_i$ , which exists primarily as the bicarbonate ion in model cyanobacteria, transits the outer protein shell, is interconverted with  $CO_2$  by the carboxysomal CA(s) (CsoSCA in  $\alpha$ -carboxysomes, and CcmM and/or CcaA in  $\beta$ -carboxysomes), and fixed into 3-phosphoglycerate (3-PGA) by Rubisco before release into the cytosol. By virtue of the outer shell functioning as a  $CO_2$  barrier, the carboxysome lumen is comparatively  $CO_2$  rich, and this environment favours Rubisco carboxylation to the extent that photorespiration is negligible in some cyanobacterial species (Cheng and Colman, 1974; Birmingham *et al.*, 1982).

Adaptive evolution of carboxysomal Rubiscos has resulted in these enzymes occupying different catalytic optima from

‘free’ Rubisco. Carboxysomal Rubiscos are comparatively fast enzymes, but have lower  $CO_2/O_2$  specificity ( $S_{C/O}$ ) (Tcherkez *et al.*, 2006). These characteristics allow the cyanobacterial cell to make a comparatively small investment in Rubisco biosynthesis, especially in comparison with  $C_3$  plants. In theoretical models, carboxysomal Rubisco operates more effectively because of its localization with CA at the top of the  $CO_2$  concentration gradient. However, recent work reiterates that merely co-locating Rubisco with CA is not enough to enhance photosynthesis appreciably (Mangan and Brenner, 2014; Gonzalez-Esquer *et al.*, 2015; Mangan *et al.*, 2016). Rather, the important function is the resistance to  $CO_2$  diffusion, within an elevated pool of  $HCO_3^-$ , which allows carboxysomes to operate effectively. This contrasts with pyrenoids, whose structure superficially resembles that of carboxysomes, but apparently lack this clearly defined barrier (Fig. 3).

### Folding, assembly, and activation of cyanobacterial Rubisco in higher plants

The two subunits of Form-I Rubisco enzymes are translated and assembled via a complex system of post-translational modifications, and folding and assembly chaperone proteins to form hexadecameric holoenzymes of eight large and eight small subunits (Houtz *et al.*, 2008; Whitney *et al.*, 2011; Wilson and Whitney, 2015). Many of the steps in this pathway remain obscure and are poorly understood. However, it is clear that phylogenetically distant Rubiscos have co-evolved with distinct types of folding and activation chaperone proteins (Whitney *et al.*, 2015). For example, the two



**Fig. 3.** Concentration of inorganic carbon in the carboxysome and pyrenoid Rubisco microcompartments. Carbon fixation proceeds by ‘outside-in’ and ‘inside-out’ concentrating mechanisms in carboxysomes and pyrenoids, respectively. Bicarbonate ( $HCO_3^-$ ) diffuses into the cyanobacterial carboxysome and conversion to  $CO_2$  proceeds by the activity of an internal carbonic anhydrase (CA). High internal  $CO_2$  concentrations are maintained by a selectively permeable protein shell (dotted line, left). In the algal pyrenoid, bicarbonate diffuses down a concentration gradient within the lumen of thylakoid membranes traversing the pyrenoid. Luminal CA activity is localized within the pyrenoid. Here, bicarbonate is rapidly converted to  $CO_2$ , which diffuses into the surrounding Rubisco aggregate. A starch sheath localized around the pyrenoid (dotted line, right) may function as a barrier to maintain high internal  $CO_2$  concentrations.

Rubisco forms found in cyanobacteria [Form-1A Rubisco in  $\alpha$ -cyanobacteria and Form-1B Rubisco in  $\beta$ -cyanobacteria (Rae *et al.*, 2013a, b)] have complex and differing chaperone requirements. In some species with Form-1A Rubisco, the  $\alpha$ -carboxysome Rubisco assembly factor (acRAF) is required for efficient Rubisco folding (Wheatley *et al.*, 2014), while Rubisco activation proceeds by CbbQ–CbbO and CbbX (the Cbb nomenclature stands for the Calvin–Benson–Basham cycle) (Sutter *et al.*, 2015; Tsai *et al.*, 2015). Species with Form-1B Rubisco require chaperonin proteins, Rubisco protein X (RbcX) and Rubisco accumulation factor 1 (RAF1), which interact with the large subunit (LSU) of Rubisco to assist in LSU folding (Emlyn-Jones *et al.*, 2006; Kolesinski *et al.*, 2014; Hauser *et al.*, 2015), while Rubisco activase (Rca) is required for Rubisco activation (Li *et al.*, 1999).

The roles of Rubisco activation chaperones within carboxysomes remains somewhat unclear. Evidence suggests that carboxysomes are persistent through cell division (Chen *et al.*, 2013), highlighting that, once incorporated into the carboxysome, any requirement for Rubisco activation would probably need to occur within the structure. The activase module CbbQ appears to be a minor component of the  $\alpha$ -carboxysome shell in *Halothiobacillus neapolitanus* carboxysomes (Sutter *et al.*, 2015). How it might act on inhibited Rubisco active sites within the carboxysome remains unclear. Indeed, the requirement for re-activation may be low within  $\beta$ -carboxysomes due to the relatively high rate of inhibitor release from Form-1B Rubisco active sites when catalytic misfiring occurs (Pearce, 2006). This detail is currently unknown for Form-1A Rubiscos. The  $\beta$ -cyanobacterial activase, Rca, was not associated with carboxysomes in cells of *Synechococcus elongatus* PCC7942 (*Synechococcus* hereafter) (Friedberg *et al.*, 1993), while the  $\alpha$ -cyanobacterial activase, CbbX, has not been reported in  $\alpha$ -carboxysomes. However, inactivation of the *rca* gene in *Anabaena* led to light-dependent effects on Rubisco activation and cell growth (Li *et al.*, 1999), highlighting its importance. Notably, a proportion of Rubisco is always found exterior to the carboxysome in both  $\alpha$ - and  $\beta$ -cyanobacteria (McKay *et al.*, 1993), which would be more accessible to activases. Within green algae, activase enzymes and Rubisco both inhabit the pyrenoid (McKay *et al.*, 1991) and, while Rca is required for optimal photosynthesis in *Chlamydomonas*, its absence is partially compensated for by an active CCM (Pollock *et al.*, 2003).

Understanding the importance of chaperones for engineering heterologous Rubisco enzymes into higher plants is currently an active area of research. Whitney *et al.* (2015) showed that the content and assembly rate of Arabidopsis Rubisco (Form-1B) in transgenic tobacco was improved 2- to 3-fold when co-transformed with Arabidopsis RAF1 (*AtRAF1*). This work demonstrated that cognate chaperones can significantly improve the biogenesis of heterologous Rubiscos, even in closely related higher plant species (Whitney *et al.*, 2015). In contrast, Rubisco from the  $\beta$ -cyanobacterium *Synechococcus* introduced into tobacco did not require or greatly benefit from the presence of RbcX (Lin *et al.*, 2014b; Occhialini *et al.*, 2016). Although the cyanobacterial Rubisco assembled and was active, the Rubisco content of these transgenic

lines was very low, comparable with the Rubisco content achieved by expression of archaeal Rubisco enzymes (Wilson *et al.*, 2016). Furthermore, the catalytic characteristics of the cyanobacterial Rubisco were not sufficient to support plant growth under ambient CO<sub>2</sub> levels, highlighting the expectation that carboxysomes alone, or their cognate Rubiscos, are not expected to improve C<sub>3</sub> photosynthetic performance (McGrath and Long, 2014). Lines co-transformed with the  $\beta$ -carboxysomal Rubisco-interlinking protein, CcmM-35, had a higher Rubisco content, suggesting that CcmM-35 might act to stabilize the cyanobacterial Rubisco *in planta* (Occhialini *et al.*, 2016). Whether co-transformation with *Synechococcus* RAF1 would result in further improvements remains to be tested.

#### *Appropriate carboxysome operons for chloroplast engineering*

Current efforts to express  $\beta$ -carboxysomes in higher plant chloroplasts have utilized genes from the well-characterized carboxysome of *Synechococcus* (Lin *et al.*, 2014a, b; Occhialini *et al.*, 2016). Lin, Occhialini, and co-workers have been successful in transiently expressing shell (CcmK, L, and O) and lumen (CcmM and N) proteins, which have been imported into the chloroplasts of *N. benthamiana* to form carboxysome-related structures (Lin *et al.*, 2014a). They have also generated crucial luminal Rubisco–CcmM35 complexes within tobacco chloroplasts (Lin *et al.*, 2014b; Occhialini *et al.*, 2016). While these breakthrough results are highly informative, the carboxysome gene set of *Synechococcus* is complicated by the existence of additional genes and translational controls that are crucial for the formation of functional carboxysomes (Long *et al.*, 2016). For example, there are several critical isoforms of CcmK (Rae *et al.*, 2012), required in unknown quantities within the shell. The Rubisco-binding protein CcmM is complicated by expression of two protein isoforms (Long *et al.*, 2010), with a shorter form arising from a ribosomal entry site within the reading frame of *ccmM* (Ludwig *et al.*, 2000; Long *et al.*, 2005, 2007). The carboxysomes of *Synechococcus* and related species also contain the additional carbonic anhydrase CcaA (So and Espie, 1998; Long *et al.*, 2011), whereas other  $\beta$ -carboxysomes make use of an active  $\gamma$ -CA at the N-terminus of the CcmM protein (Peña *et al.*, 2010; de Araujo *et al.*, 2014). There is also the potential requirement for an additional shell protein, CcmP, which may play a role in the passage of large metabolites across the shell (Cai *et al.*, 2013). The requirement for Rubisco ancillary proteins (mentioned above) is still not completely clear, and their presence within the  $\beta$ -carboxysome is completely unknown. The complexity of the  $\beta$ -carboxysome is underlined by the arrangement of carboxysome-related genes into clusters, compared with more tightly regulated operons for  $\alpha$ -carboxysomes (Cai *et al.*, 2008; Heinhorst *et al.*, 2014).

Even if expression of all the components required for carboxysome assembly is achieved, we see carboxysomes with smaller genetic footprints as the most logically appealing for C<sub>3</sub> chloroplast transformation (Long *et al.*, 2016). Not only is the genetic load lower, and the potential for recombination

between homologous genes significantly reduced, but self-assembly of comparatively simpler carboxysomes might be more efficient in the chloroplast. [Cai et al. \(2016\)](#) were successful in engineering relatively simple synthetic carboxysome shells and cargo in *E. coli* using genes from the  $\beta$ -cyanobacterium *Halotheca* sp. PCC7418, demonstrating that both simple gene units and synthetic approaches to the challenge of chloroplastic carboxysome expression have significant merit. Self-assembly of the carboxysome is a topic of ongoing investigation, but current evidence suggests that carboxysomes can assemble from their component parts with remarkable ease ([Cameron et al., 2013](#); [Chen et al., 2013](#)).

Several simple candidate operons exist for both  $\alpha$ - and  $\beta$ -carboxysomes. *Gloeobacter violaceus* PCC7421 has the smallest known  $\beta$ -carboxysome gene content ([Nakamura et al., 2003](#)), and is also likely to encode a CA-competent form of CcmM, obviating the need for a separate carboxysomal CA gene. Moreover, *G. violaceus* carboxysomes appear to require only five outer shell proteins (CcmK1, CcmK2, CcmO, CcmL, and CcmP). However, *G. violaceus* probably still expresses two isoforms of CcmM. The high-light-adapted *Prochlorococcus* clade have the simplest known  $\alpha$ -carboxysomes in terms of gene content, with as few as nine structural genes ([Roberts et al., 2012](#)), and with some expressing a single form of CsoS2 (the  $\alpha$ -carboxysomal equivalent of CcmM) ([Roberts et al., 2012](#); [Chaijarasphong et al., 2016](#)). Thus, based on gene number and apparent ease of expression,  $\alpha$ -carboxysome operons are currently the most attractive candidates for chloroplastic transformation.

Rubisco fixation rates are assumed to approach their theoretical maximum for a fully functional biophysical CCM in a  $C_3$  chloroplast ([McGrath and Long, 2014](#)). Thus, the catalytic properties of extant carboxysomal Rubiscos must be considered when choosing appropriate sources for engineering. Based on measurements thus far, Rubiscos from  $\alpha$ -carboxysomes appear catalytically inferior to those from  $\beta$ -carboxysomes. *Prochlorococcus* and *Cyanobium* Form-1A Rubiscos have  $k_{\text{cat}}^c$  values of  $6.6 \text{ s}^{-1}$  and  $9.0 \text{ s}^{-1}$ , respectively ([Whitehead et al., 2014](#); [Shih et al., 2016](#)), compared with reports of up to  $14.4 \text{ s}^{-1}$  for Form-1B Rubisco from *Synechococcus* ([Whitehead et al., 2014](#); [Occhialini et al., 2016](#)). The current lack of catalytic data for  $\alpha$ -carboxysomal Form-1A Rubiscos warrants greater analysis of this diverse family of enzymes to identify potential high-performance candidates. However, some  $\alpha$ -carboxysomes are amenable to sequestration of heterologous Rubiscos ([Menon et al., 2008](#)), which opens up the possibility of engineering structurally simple carboxysomes containing kinetically superior Rubiscos into  $C_3$  chloroplasts.

The generation of a carboxysome-based biophysical CCM in the chloroplast will be a multistep process. In the absence of functional transporters, much work has focused on optimizing carboxysome expression in the chloroplast ([Price et al., 2013](#); [Long et al., 2016](#)). The latter is challenging, as the functional performance of expressed carboxysomes is difficult to determine *in planta*. The loss of single protein components from the carboxysome (due either to genetic lesion or to stoichiometric perturbations) can lead to a loss or decrease of

function, even though morphologically the carboxysome may not appear altered ([Cai et al., 2009](#); [Bonacci et al., 2012](#); [Rae et al., 2013a](#)). Thus, the ability to isolate carboxysome structures expressed in the chloroplast, and to determine their function, will be essential. While synthetic  $\beta$ -carboxysomes have been expressed and purified from *E. coli* ([Cai et al., 2016](#)), successful purification of intact, functional  $\beta$ -carboxysomes from wild-type (WT) cyanobacteria has been elusive ([Long et al., 2005, 2011](#)). Conversely,  $\alpha$ -carboxysomes can be purified intact ([So et al., 2004](#); [Gonzales et al., 2005](#); [Roberts et al., 2012](#)), and some findings indicating function have been reported ([Cai et al., 2009](#); [Menon et al., 2010](#)). The development of further assays for carboxysome functionality *in vitro* will be important to guide future engineering strategies.

### *Costs of carboxysome expression in the chloroplast—how many do we need?*

Conservative predictions estimate between 8 and 13 separate polypeptides will be required to generate a functional  $\alpha$ - or  $\beta$ -carboxysome in a  $C_3$  chloroplast ([Long et al., 2016](#)). While there are still unknowns regarding the absolute protein structure of both  $\alpha$ - and  $\beta$ -carboxysomes, existing knowledge of protein stoichiometries for each type ([Roberts et al., 2012](#); [Rae et al., 2013a](#)) makes it possible to estimate the nutritional and spatial costs of biosynthesis in plants. In [Table 1](#), we have calculated the costs of carboxysome synthesis in a  $C_3$  leaf under two possible scenarios (in a fully functional CCM where  $C_i$  accumulation allows for the operation of Rubisco at their catalytic maximum). First, the predicted costs of both  $\alpha$ - and  $\beta$ -carboxysomes are estimated assuming the replacement of the WT  $C_3$  enzyme with an equivalent number of carboxysomal Rubiscos ([Table 1](#), upper rows). Under this scenario, the additional structural and functional components of carboxysomes (i.e. shell, CA, and Rubisco-binding proteins) mean that there is additional nitrogen, carbon, and energy investment for the construction of carboxysomes. However, these costs are offset by the expected enhancement in carbon fixation rates within a fully functional chloroplastic CCM when we model a scenario in which the theoretical *in vitro* maximum carboxylation rates ( $V_{c,\text{max}}$ ) are matched with those expected in WT leaves ([Table 1](#), bottom rows). Given that the  $k_{\text{cat}}^c$  values for carboxysomal Rubiscos are 3- to 4-fold higher than  $C_3$  counterparts, and assuming that a chloroplastic CCM can deliver  $\text{CO}_2$  concentrations to carboxysomal Rubiscos in excess of their relatively high  $K_c$  (and minimizing the adverse effects of low  $S_{C(O)}$ ), our calculations actually predict significant overall savings in carbon, nitrogen, total protein, and energy investments under this scenario. One caveat here is that the costs of the bicarbonate transporters required to elevate the stromal  $C_i$  concentration are not included. Nonetheless, a 44–63% saving in leaf protein and 40–65% saving in stromal volume dedicated to Rubisco could be expected in the scenario presented, and is likely to offset the costs of transporter synthesis significantly. Notably, predicted savings in the energy cost of peptide synthesis, under conditions which support the same theoretical  $V_{c,\text{max}}$  as WT plants ([Table 1](#)), may also contribute to the cost

**Table 1.** Estimated carboxysome requirements in tobacco leaves in a fully functional chloroplastic CCM

Rubisco source	Rubisco $k_{cat}^c$ ( $s^{-1}$ )	Rubisco active sites ( $\mu\text{mol m}^{-2}$ leaf)	Rubisco holoenzymes or carboxysomes ( $\text{m}^{-2}$ leaf)	Energy cost of peptide synthesis ( $\text{kJ m}^{-2}$ )	% Stromal volume	Nitrogen cost ( $\text{mg N m}^{-2}$ leaf)	Carbon cost ( $\text{mg m}^{-2}$ leaf)	Protein mass ( $\text{mg m}^{-2}$ leaf)	Protein cost/saving ( $\text{mg m}^{-2}$ leaf)
To achieve the same Rubisco active site density as WT tobacco ( $25 \mu\text{mol m}^{-2}$ )									
WT <i>N. tabacum</i>	3.2	25	$1.88 \times 10^{18}$	2.62	13	288	915	1691	
$\alpha$ -Carboxysome	9.0	25	$1.24 \times 10^{16}$	3.97	22	468	1440	2681	-990
$\beta$ -Carboxysome	14.4	25	$1.68 \times 10^{15}$	4.07	21	492	1497	2851	-1160
To achieve the same theoretical <i>in vitro</i> maximum carboxylation rates as WT tobacco ( $V_{c,max}^c$ , $80 \mu\text{mol m}^{-2} s^{-1}$ )									
WT <i>N. tabacum</i>	3.2	25	$1.88 \times 10^{18}$	2.62	13	288	915	1691	
$\alpha$ -Carboxysome	9.0	8.9	$4.40 \times 10^{15}$	1.41	7.8	166	512	953	738
$\beta$ -Carboxysome	14.4	5.6	$3.73 \times 10^{14}$	0.91	4.6	109	333	634	1057

Tobacco (*Nicotiana tabacum*) Rubisco  $k_{cat}^c$  ( $3.2 s^{-1}$ ) and active site density ( $25 \mu\text{mol m}^{-2}$ ) values are taken from Sharwood (2017) and are used to calculate a theoretical *in vitro* maximum carboxylation rate ( $V_{c,max}^c$ ) of  $80 \mu\text{mol m}^{-2} s^{-1}$ . Values for  $\alpha$ - and  $\beta$ -carboxysome Rubisco  $k_{cat}^c$  are those of *Cyanobium* sp. PCC7001 and *Synechococcus elongatus* PCC7942, respectively (Whitehead et al., 2014). Carboxysome numbers are calculated in terms of Rubisco active site numbers per  $\text{m}^2$  leaf area for carboxysomes of 90 nm diameter [representing *Prochlorococcus marinus* MED4  $\alpha$ -carboxysomes for which protein stoichiometries have been estimated (Roberts et al., 2012)] and 172 nm diameter [representing *S. elongatus* PCC7942  $\beta$ -carboxysomes for which protein stoichiometries have been estimated (Long et al., 2011)]. Carboxysomes are assumed to be idealized icosahedrons containing 152 and 1120 Rubisco holoenzymes in a Kepler packing arrangement for both  $\alpha$ - and  $\beta$ -carboxysomes, respectively, where the internal and total volumes are calculated allowing for shell thicknesses of 4 nm for  $\alpha$ -carboxysomes (Lancu et al., 2007) and 5.5 nm for  $\beta$ -carboxysomes (Kaneko et al., 2006). Protein, carbon, and nitrogen masses of each carboxysome type were determined using published stoichiometries and amino acid sequences from both *P. marinus* MED4 and *S. elongatus* PCC7942. Protein masses for  $\alpha$ - and  $\beta$ -carboxysomes were determined to be 130.39 MDa and 1021.91 MDa, respectively. The contributions of proteins CcmP and CcmN to  $\beta$ -carboxysome masses, along with any Rubisco activation and folding chaperones, have not been included in the calculations due to their unknown contribution, which is likely to be relatively small. The amino acid sequence of *N. tabacum* RbcL and RbcS were used to calculate carbon and nitrogen masses and a molecular mass of 541 kDa. Numbers of amino acids required for the synthesis of tobacco Rubisco and the entire carboxysome of each type were used to estimate the energy cost of peptide synthesis at a rate of  $160 \text{ kJ mol}^{-1}$  for each peptide bond. The cost of the tobacco RbcS chloroplast targeting sequence is included in its energy cost. Rubisco holoenzymes from each species were assumed to occupy a volume of  $905 \text{ nm}^3$  (based on a spherical diameter of 12 nm). Stromal volumes were calculated from mesophyll surface area per unit leaf area ( $15 \text{ m}^2 \text{ m}^{-2}$ ) and stromal volume per mesophyll surface area ( $0.87 \times 10^{-6} \text{ m}^3 \text{ m}^{-2}$ ) as used by Tholen and Zhu (2011). Protein cost/savings are calculated from total protein masses for carboxysomes compared with wild-type (WT) tobacco plants with free Rubisco, and assume the total replacement of free Rubisco with carboxysomes. A negative value represents a net cost and a positive value represents a net saving.

of energy-driven bicarbonate transport uptake. These predicted gains are also facilitated in part by the extremely high concentration of Rubisco active sites within both carboxysome types [10.9–14.1 mM (Whitehead *et al.*, 2014)], which approach or exceed the maximum concentrations predicted in  $C_3$  stromae (Pickersgill, 1986). Differential costs between the two carboxysome types are primarily due to their intrinsic Rubisco kinetic maxima.

In order to achieve the equivalent number of Rubisco active sites in a chloroplastic carboxysome as in a WT plant, there are significant costs in both protein requirement and stromal volume (Table 1). However, this represents a substantial overinvestment and is probably not required with a fully functional chloroplastic CCM. The estimated carboxysome numbers required to achieve the same  $V_{c,max}$  as WT plants suggest that leaf carboxysomal Rubisco active site concentrations need only be 2.8- to 4.5-fold lower than those of WT plants. While this estimate is heavily dependent on the relative improvement in  $k_{cat}^c$  between the WT Rubisco and its replacement carboxysomal counterpart, our calculations indicate that even low levels of carboxysomal Rubisco expression may provide a benefit within the context of a chloroplastic CCM. For example, the level of *Synechococcus* Rubisco expression reported by Occhialini *et al.* (2016) was low, but those values do approach the relative levels required in a fully functional chloroplastic CCM. While the data presented in Table 1 present a simplified approach to estimating carboxysome requirements, this analysis highlights the value in modelling plant performance to understand if there is an over- or underinvestment in carboxysome synthesis.

## Structure and dynamics of the algal pyrenoid

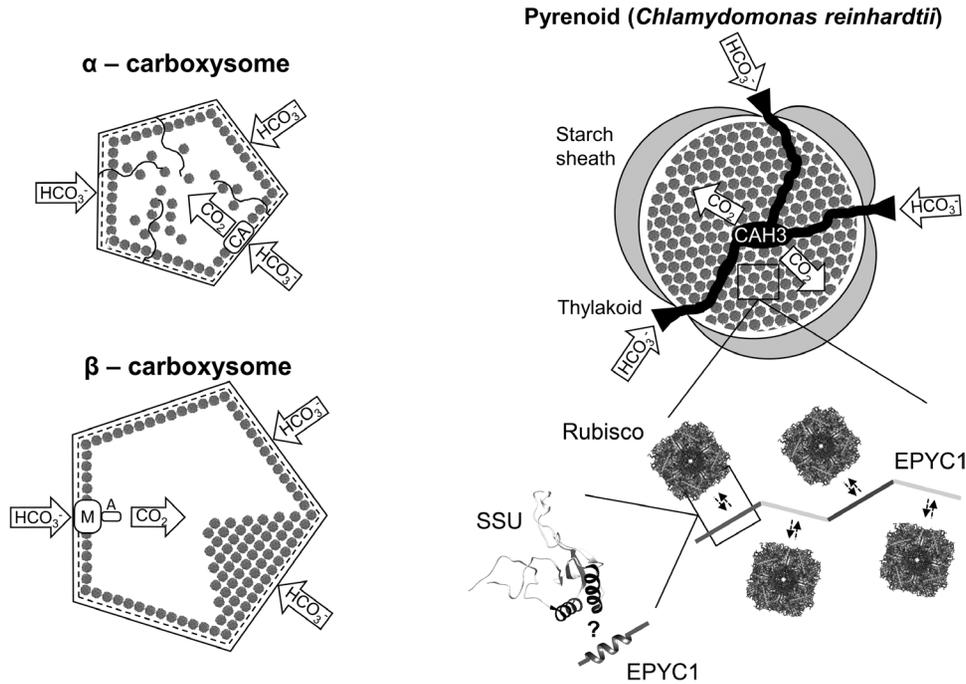
Pyrenoids are subcellular organelles located in the plastid stroma of almost all eukaryotic microalgae, as well as diatoms, dinoflagellates, coccolithophores, and a single group of land plants, the hornworts (Vaughn *et al.*, 1990; Schmid, 2001; Hanson *et al.*, 2014; Hopkinson *et al.*, 2016). Species may contain a single pyrenoid (as in *Chlamydomonas*) or several. Like the carboxysome, the fine structure of the pyrenoid is well known to be formed of a dense protein aggregate (Kowallik, 1969; Bertagnolli and Nadakavukaren, 1970; Leadbeater and Manton, 1971; Markowitz and Hoffman, 1974), which we now interpret to be primarily Rubisco (Engel *et al.*, 2015). Although pyrenoids are absent in some chlorophytes that appear to retain a capacity for active  $C_i$  uptake (Morita *et al.*, 1998), pyrenoids probably play a key role in the efficiency of biophysical CCMs in eukaryotes where they are present, and are associated with species with high internal  $C_i$  levels (Fig. 3) (Badger *et al.*, 1998; Morita *et al.*, 1999). While the catalytic properties of microalgal Rubiscos enable photosynthesis at ambient  $CO_2$  (Jordan and Ogren, 1981), disruption of the pyrenoid inhibits induction of the algal CCM in species where pyrenoids are present (Genkov *et al.*, 2010; Ma *et al.*, 2011; Meyer *et al.*, 2012; Mackinder *et al.*, 2016). Thus, pyrenoids are thought to be of similar importance to

the carboxysomes (Giordano *et al.*, 2005), and their structures do share several key features: a dense aggregation of spatially arranged Rubiscos and an intrinsic internal CA activity (CAH3 in *Chlamydomonas*) (Fig. 4). However, there are several marked structural distinctions. For example, pyrenoids do not have a defined shell structure, but are typically (but not always) surrounded by a sheath of starch plates. It remains unclear if the starch layer inhibits  $CO_2$  leakage in a manner akin to the carboxysome shell. In *Chlamydomonas* under  $C_i$ -limiting conditions, the enigmatic LCIB/LCIC complex is concentrated in a discrete region around the starch layer, but can also be found in the surrounding stroma (Wang and Spalding, 2014). LCIB/LCIC may be involved in unidirectional re-capture of  $CO_2$  released from the pyrenoid and thylakoid membranes (Jin *et al.*, 2016). Pyrenoids are also traversed by a dense network of specialized thylakoid membranes that fenestrate the starch sheath and fuse to form internal minitubules contiguous with the chloroplast stroma (Engel *et al.*, 2015). The internal minitubules are thought to provide a channel for the substrates and products of Rubisco, while the surrounding larger tubules facilitate bicarbonate diffusion to CAH3 localized within the trans-pyrenoid thylakoids (Meyer *et al.*, 2016). As the thylakoid membranes traversing the pyrenoid are PSI enriched, Smith and Griffiths (2000) have suggested that the pyrenoid could also provide a means of spatially separating the oxygenic PSII complex from Rubisco to reduce the occurrence of competitive photorespiratory reactions. Nevertheless, pyrenoid Rubiscos do have greater  $S_{C/O}$  values than those from carboxysomes (Badger *et al.*, 1998).

### Protein determinants of pyrenoid assembly and ultrastructure

Identifying factors that regulate assembly of the pyrenoid is currently an active area of research. Pyrenoids appear considerably more dynamic than carboxysomes, and undergo changes in size, composition, and density depending on  $C_i$  availability and light (Ramazanov *et al.*, 1994; Mitchell *et al.*, 2014). Under photoautotrophic growth at high  $CO_2$  levels (1% or higher), less than half of Rubisco is located in the pyrenoid, and starch accumulates in the stroma rather than around the pyrenoid (Borkhsenius *et al.*, 1998). Following CCM induction at air levels of  $CO_2$  (0.04%) or below, >90% of Rubisco aggregates in the pyrenoid, and the pyrenoid starch sheath is formed consequently. Similarly, pyrenoid Rubisco content increases following diel transitions from dark to light, which coincides with light-dependent (or dawn) induction of the CCM (Mitchell *et al.*, 2014).

The structure of the Rubisco small subunit (SSU) plays a key role in pyrenoid assembly. Hybrid Rubiscos that contain native *Chlamydomonas* LSUs and a higher plant SSU cannot form pyrenoids (Genkov *et al.*, 2010). Meyer *et al.* (2012) elegantly demonstrated that two  $\alpha$ -helices located on the SSU periphery are more hydrophobic in species with pyrenoids, and are a critical feature for assembly in *Chlamydomonas*. In contrast, the LSU does not appear to be important. Pyrenoid morphology remained unaffected by mutations in



**Fig. 4.** Substructural organization of the cyanobacterial  $\alpha$ - and  $\beta$ -carboxysomes, and the pyrenoid from *Chlamydomonas*. In both  $\alpha$ - and  $\beta$ -carboxysomes (left), pore structures on the outer shell barrier facilitate bicarbonate ( $\text{HCO}_3^-$ ) uptake. Bicarbonate is converted to  $\text{CO}_2$  by one of several potential carbonic anhydrases [CsoSCA (CA) in the case of  $\alpha$ -carboxysomes and CcaA (A) or CcmM (M) in the case of  $\beta$ -carboxysomes], probably located at the subsurface of the each carboxysome shell. In  $\alpha$ -carboxysomes, Form-1A Rubisco (filled grey circles) is anchored internally via the highly disordered CsoS2 protein (black line), whereas in  $\beta$ -carboxysomes Form-1B Rubiscos are interlinked via the full-length form of CcmM to interact with the internal surface of the shell, and via a short isoform of CcmM to generate a regular, ordered array of enzymes (only a portion of this internal structure is shown). Carboxysome figures are adapted from Long et al. (2011) and Rae et al. (2013a, b). The *Chlamydomonas* pyrenoid (right) consists of a dense semi-organized protein aggregation, primarily of Rubisco, traversed by a network of specialized thylakoid membranes (minitubules not shown here), and surrounded by a starch sheath. Bicarbonate is taken up into the thylakoid lumen and converted into  $\text{CO}_2$  by a luminal  $\alpha$ -type carbonic anhydrase (CAH3) localized within the pyrenoid. EPYC1 is abundant in the pyrenoid when the CCM is active, and is shown as a linker chaperone that stabilizes Rubisco within the pyrenoid structure. In *Chlamydomonas*, EPYC1 has four nearly identical repeat regions (shown in dark grey and light grey), and each repeat is suggested as a site for RuBisCO binding. Within each repeat, a less disordered domain containing a predicted  $\alpha$ -helix may interact with the two outward-facing  $\alpha$ -helices of the small subunit of Rubisco (SSU).

the LSU, even those resulting in an inactive Rubisco (Rawat et al., 1996; Meyer, 2010). The pyrenoid is known to contain two additional non-thylakoid-specific proteins, Rca and the recently characterized Essential Pyrenoid Component 1 [EPYC1; previously LCI5 (McKay et al., 1991; Mackinder et al., 2016)]. Null mutants for Rca show an expected decrease in the amount of active Rubisco, but no effect on pyrenoid morphology or  $\text{C}_i$  affinity (Pollock et al., 2003). In contrast, EPYC1 is required for pyrenoid assembly and CCM function, and is believed to be a linker chaperone that facilitates aggregation of Rubisco within the pyrenoid structure during CCM induction (Fig. 4). EPYC1 is an intrinsically disordered protein with four nearly identical repeat regions, and each repeat is suggested as a site for Rubisco binding, possibly by interactions with the SSU. Thus, EPYC1 may play a similar role to the  $\alpha$ -carboxysomal Rubisco-binding protein CsoS2 which is also highly disordered and binds SSUs (Cai et al., 2015).

Successful engineering of the pyrenoid into higher plants will require a better understanding of pyrenoid biogenesis, regulation, and integration with other CCM components. Fortunately, recent pioneering research has already started to provide important details. For example, CAH3 and EPYC1 undergo multiple phosphorylations when  $\text{CO}_2$  is limiting (Blanco-Rivero et al., 2012; Mackinder et al., 2016). Similarly,

models have implicated phosphorylation in activity regulation of the LCIB/LCIC complex (Jin et al., 2016). Identifying components involved in post-translational modification and signalling is ongoing. Recently, the thylakoid-localized  $\text{Ca}_2^+$ -binding protein (CAS) was shown to re-localize to minitubules in the pyrenoid during CCM induction in *Chlamydomonas* (Wang et al., 2016). CAS may operate within a  $\text{Ca}^{2+}$ - and  $\text{CO}_2$ -mediated chloroplast retrograde signalling pathway to regulate expression of nuclear-encoded CCM components, such as HLA3 and LCIA. Although it remains unclear how thylakoids come to traverse pyrenoids, newly identified CCM components may shed light on the relationship between Rubisco, thylakoids, and starch surrounding the pyrenoid (Powell, 2016), and known key regulators involved in pyrenoid assembly and CCM function, such as CIA5 and CIA6 (Fukuzawa et al., 2001; Xiang et al., 2001; Ma et al., 2011).

## Towards building a pyrenoid in higher plants

As with the cyanobacterial CCM, developing a chloroplastic pyrenoid-based CCM in  $\text{C}_3$  plants will probably require a step-wise incorporation of pyrenoidal components, and

modification of native Rubisco (Meyer *et al.*, 2012). In this respect, the algal CCM may have a tractable advantage over the cyanobacterial CCM. Although cyanobacterial Rubiscos have high  $k_{cat}$  rates, their catalytic efficiencies in higher plants are hampered by low values for  $K_c$  and  $S_{CO}$  (Lin *et al.*, 2014b; Occhialini *et al.*, 2016). This reflects their co-evolution within the carboxysome structure, and adaptation to higher  $CO_2$  concentrations compared with Rubiscos prevailing in either angiosperm chloroplasts or algal pyrenoids (Meyer and Griffiths, 2013). In contrast, the catalytic characteristics of algal Rubiscos are close to those of Rubiscos in  $C_3$  plants (Meyer *et al.*, 2012; Galmés *et al.*, 2014; Prins *et al.*, 2016). Thus, in a development pipeline for engineering the pyrenoid-based CCM, initial introduction of an algal Rubisco is likely to be neutral compared with the expected drop in photosynthetic performance and growth during the development of a cyanobacterial CCM (McGrath and Long, 2014).

If the algal LSU is not required for pyrenoid assembly, as evidence suggests [see the previous section (Rawat *et al.*, 1996; Meyer, 2010)], engineering a plant with Rubisco capable of pyrenoid assembly may require replacement only of the higher plant SSU. Thus, requirements for co-transformation of additional Rubisco assembly chaperones, as seen for the LSU (Bracher *et al.*, 2015; Whitney *et al.*, 2015), may be limited to those needed for the SSU. Currently, there is little evidence to suggest that the SSU requires assembly chaperones in the chloroplast (Wilson and Whitney, 2015). Recent work in *Arabidopsis* has shown that modified SSUs competent for pyrenoid assembly can be incorporated into  $C_3$  plant Rubisco with minimal effects on Rubisco catalysis, content, and plant growth at ambient  $CO_2$  (Atkinson *et al.*, 2017). These *Arabidopsis* lines are proving to be useful backgrounds in which to (i) examine pyrenoid assembly chaperones, such as EPYC1, and other candidates as they emerge; (ii) clarify the nature of SSU-associated interactions, and requirements for Rubisco aggregation *in planta*; and (iii) integrate other essential CCM components involved in active  $C_i$  uptake.

## Synthetic approaches to constructing a biophysical CCM in higher plants

### Algal CCMs

Due to initiatives such as the 1KP project (onekp.com), the increasing availability of sequencing data has made it possible to begin exploring how ubiquitous the CCM components of the *Chlamydomonas* CCM are in other eukaryotes with biophysical CCMs. For example, some components, such as LCIA and HLA3, appear to be common, while others, such as LCI1, are not (Meyer and Griffiths, 2013). Currently the majority of sequenced genomes remain poorly annotated, but this growing resource could offer a useful tool to mine alternative components for building a synthetic biophysical CCM.

Although the *Chlamydomonas* pyrenoid is clearly the best understood, it is also arguably one of the most structurally complex. Pyrenoids do vary significantly in terms of shape, size, and apparent complexity between different species. Although some algae and bryophytes possess highly complex

pyrenoids (i.e. with traversing minitubules and/or thylakoid invaginations, sometimes with and without a starch sheath present), the simplest pyrenoids, as viewed in transmission electron micrographs, appear to be homogenous bodies with no starch sheaths and often without any clear localization near the thylakoids, as seen in *Laminaria* spp. (Chi, 1971) and many haptophytes (Griffiths, 1970; Bendif *et al.*, 2011). These types of pyrenoid might be the simplest in terms of structural protein components, and thus may be the most feasible models for assembly in higher plant chloroplasts. Recent work has shown that a wide diversity of haptophyte algae with ‘simple’ pyrenoids, or in some cases no pyrenoid at all, are capable of high photosynthetic rates (BDR and R.E.M. Rickaby, unpublished results). However, it is clear we still know very little about the molecular make up of pyrenoids, and assembly components, in other eukaryotes. Growing evidence for the contribution of single-cell  $C_4$ -like CCMs in some algae and diatoms suggests that additional complexities may also exist in certain species (Derelle *et al.*, 2006; Hopkinson *et al.*, 2016). Thus, improving our current molecular and physiological understanding of biophysical CCMs in other globally significant eukaryotes will be important for optimizing the chances of successful transfer to higher plants. The hornworts are particularly attractive candidates for further study, as they are phylogenetically the closest group to angiosperms that retain pyrenoids and biophysical CCMs (Li *et al.*, 2017). The pyrenoid-bearing hornwort, *Anthoceros agrestis*, was recently established as a tractable model system and may prove useful as a resource for comparative studies (Szövényi *et al.*, 2015).

### Cyanobacterial CCMs

Efforts to increase the existing genomic data for the cyanobacteria phylum have demonstrated the huge variety of potential CCM component homologues available for exploitation (Shih *et al.*, 2013). Gaudana *et al.* (2015) showed that genes with homology to the key transporters BicA and SbtA (and the putative regulator SbtB) were widely conserved in 128 different sequenced genomes. Further research may reveal the malleability of the cyanobacterial CCM and minimal requirements for functionality. Nonetheless, several good examples of minimal CCMs are evident in nature. While the minimal carboxysome gene sets of *G. violaceus* and *Prochlorococcus* spp. were discussed earlier, the endosymbiotic cyanobacteria *Richellia intracellularis*, *Anabaena azollae*, *Synechococcus spongarium*, and the endosymbiont of *Paulinella chromatophora* converge on a minimal gene set comprising one identifiable bicarbonate transporter, one  $CO_2$  uptake/scavenging system, and minimal carboxysome operons. The free-living  $\beta$ -cyanobacterium *Trichodesmium erythraeum* and the  $\alpha$ -cyanobacterium *Prochlorococcus marinus* also have a small set of identifiable CCM components (Kranz *et al.*, 2011; Hopkinson *et al.*, 2014).

### Rational designs for a chimeric biophysical CCM

Recent progress highlights the possibility of introducing CCM components from either the nuclear genome, the plastid genome, or both (Lin *et al.*, 2014a, b; Atkinson *et al.*,

2016, 2017). These findings are exciting as they open up several opportunities, including parallel development of transformant lines and their subsequent conjunction through crossing, or transformation of crop species currently recalcitrant to chloroplastic transformation. However, nuclear expression and import into the chloroplast will require a better understanding of the effect of any transit peptide scars on functional CCM components in the chloroplast. Previous successes in appropriate localization of algal CCM components in the stroma and chloroplast envelope, and formation of carboxysome structural intermediates from small proteins when fused with native transit peptides are promising (Lin *et al.*, 2014a; Atkinson *et al.*, 2016).

With such a wide availability of tools, the structural and functional concepts of biophysical CCM components from eukaryotes and cyanobacteria could be potentially recapitulated via synthetic biology approaches. For example, a feasible solution for maximizing  $C_i$  uptake may be a chimeric approach, such as combining  $C_i$  uptake systems from cyanobacteria and algae. In addition, orthogonal regulators could be used to control transporter activity. Efforts are underway to engineer rhodopsin-driven bicarbonate transporters, which would enable activity to be regulated by light and photosynthetic activity (Gaudana *et al.*, 2015). The co-encapsulation of a Rubisco and CA into a simplified (single protein) capsid shell was recently demonstrated by Frey *et al.* (2016). While this 'primordial carboxysome' system showed marginal changes in Rubisco catalysis over the naked enzyme, it established a logical foundation for further development. For example, such systems could utilize directed evolution to advance the functionality of synthetic carboxysomes. While the functional role of the carboxysome shell is still under research scrutiny, designing a simplified microcompartment that approaches the functionality of evolved carboxysomes with a minimal set of genetic components (e.g. Gonzalez-Esquer *et al.*, 2015; Cai *et al.*, 2016) may be a useful strategy. Alternatively, a combinatorial approach may be more beneficial (e.g. Fuentes *et al.*, 2016) where multiple genes could be expressed in the chloroplast, covering every possible requirement of an active carboxysome. Given recent progress in using mutant libraries and high-throughput screening approaches in *Chlamydomonas* to detect novel putative eukaryotic CCM components (Li *et al.*, 2016; Powell, 2016), the latter approach may also be useful for studying pyrenoid assembly. Our understanding of pyrenoid biogenesis and regulation still lingers far behind that of carboxysomes, but recent advances in high-fidelity image analyses of pyrenoid architecture (Engel *et al.*, 2015) and high-throughput genetics may help to provide a rapid catch up.

## Future work

Much progress has been made in defining the requirements for introducing biophysical CCMs, and engineering different CCM components into  $C_3$  plants. Future work requires focus in key areas. In particular, characterization of the processes involved in regulating bicarbonate transporter activities is

critical to facilitate elevated  $C_i$  levels in  $C_3$  chloroplasts. Price *et al.* (2011a) noted that an important research focus toward a  $C_3$  chloroplastic CCM is an understanding of the activation and energization processes for bicarbonate transporters. Initial modelling indicates that the cost of  $C_i$  import is minimal and is offset by reductions in photorespiration (Price *et al.*, 2011a). Nonetheless, further accurate modelling of the energy costs is needed to confirm that a  $C_3$  chloroplast can meet the challenge. A related issue that has received little focus but will need to be addressed is the burden and potential limitations of protein expression on the  $C_3$  chloroplast envelope (Breuers *et al.*, 2012). This is an exciting time for algal CCM research, as many new details of the structure and physiology of the eukaryotic pyrenoid are currently emerging. These should be used to inform algal-specific models for engineering  $C_3$  plants. While our understanding of the components required for carboxysomes is significantly more mature than that of pyrenoids, improved techniques to study carboxysome assembly and functionality *in planta* (including the chaperone requirements for carboxysomal Rubiscos) will greatly assist efforts to engineer a functional biophysical CCM into higher plants.

## Acknowledgements

Work highlighted in this review was funded by the UK Biotechnology and Biological Sciences Research Council and US National Science Foundation as part of the Combining Algal and Plant Photosynthesis (CAPP) consortium (AJM, NA, and CNV supported through grants BB/1024453/1 and BB/M006468/1). BML, BF, WYH, and GDP are supported through a sub-award from the University of Illinois as part of the Realizing Increased Photosynthetic Efficiency (RIPE) project, funded by the Bill & Melinda Gates Foundation. BM, BDR, NDN, and GDP are supported by the Australian Research Council, Centre of Excellence grant for 'Translational Photosynthesis' (CE140100015). The authors thank Niall Mangan (University of Washington, USA) for comments on Fig. 3, and Murray Badger and Rob Sharwood (Australian National University, Australia) for comments on Table 1.

## References

- Aldridge C, Cain P, Robinson C. 2009. Protein transport in organelles: protein transport into and across the thylakoid membrane. *FEBS Journal* **276**, 1177–1186.
- Atkinson N, Feike D, Mackinder LC, Meyer MT, Griffiths H, Jonikas MC, Smith AM, McCormick AJ. 2016. Introducing an algal carbon-concentrating mechanism into higher plants: location and incorporation of key components. *Plant Biotechnology Journal* **14**, 1302–1315.
- Atkinson N, Leitão N, Orr DJ, Meyer MT, Carmo-Silva E, Griffiths H, Smith AM, McCormick AJ. 2017. Rubisco small subunits from the unicellular green alga *Chlamydomonas* complement Rubisco-deficient mutants of *Arabidopsis*. *New Phytologist* **214**, 655–667.
- Badger MR, Andrews TJ, Whitney SM, Ludwig M, Yellowlees DC, Leggat W, Price GD. 1998. The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based  $CO_2$ -concentrating mechanisms in algae. *Canadian Journal of Botany* **76**, 1052–1071.
- Badger MR, Bek EJ. 2008. Multiple Rubisco forms in proteobacteria: their functional significance in relation to  $CO_2$  acquisition by the CBB cycle. *Journal of Experimental Botany* **59**, 1525–1541.
- Badger MR, Hanson D, Price GD. 2002. Evolution and diversity of  $CO_2$  concentrating mechanisms in cyanobacteria. *Functional Plant Biology* **29**, 161–173.
- Bendif EI M, Probert I, Hervé A, Billard C, Goux D, Lelong C, Cadoret JP, Véron B. 2011. Integrative taxonomy of the Pavlovophyceae (Haptophyta): a reassessment. *Protist* **162**, 738–761.

- Bertagnolli BL, Nadakavukaren MJ.** 1970. An ultrastructural study of pyrenoids from *Chlorella pyrenoidosa*. *Journal of Cell Science* **7**, 623–630.
- Betti M, Bauwe H, Busch FA, et al.** 2016. Manipulating photorespiration to increase plant productivity: recent advances and perspectives for crop improvement. *Journal of Experimental Botany* **67**, 2977–2988.
- Birmingham BC, Coleman JR, Colman B.** 1982. Measurement of photorespiration in algae. *Plant Physiology* **69**, 259–262.
- Blanco-Rivero A, Shutova T, Román MJ, Villarejo A, Martínez F.** 2012. Phosphorylation controls the localization and activation of the luminal carbonic anhydrase in *Chlamydomonas reinhardtii*. *PLoS One* **7**, e49063.
- Bock R.** 2014. Genetic engineering of the chloroplast: novel tools and new applications. *Current Opinion in Biotechnology* **26**, 7–13.
- Bonacci W, Teng PK, Afonso B, Niederholtmeyer H, Grob P, Silver PA, Savage DF.** 2012. Modularity of a carbon-fixing protein organelle. *Proceedings of the National Academy of Sciences, USA* **109**, 478–483.
- Borkhsenius ON, Mason CB, Moroney JV.** 1998. The intracellular localization of ribulose-1,5-bisphosphate carboxylase/oxygenase in *Chlamydomonas reinhardtii*. *Plant Physiology* **116**, 1585–1591.
- Bracher A, Hauser T, Liu C, Hartl FU, Hayer-Hartl M.** 2015. Structural analysis of the rubisco-assembly chaperone RbcX-II from *Chlamydomonas reinhardtii*. *PLoS One* **10**, e0135448.
- Breuers F, Braeutigam A, Geimer S, Welzel UY, Stefano G, Renna L, Brandizzi F, Weber AP.** 2012. Dynamic remodeling of the plastid envelope membranes—a tool for chloroplast envelope in vivo localizations. *Frontiers in Plant Science* **3**, 7.
- Brueggeman AJ, Gangadharaiyah DS, Cserhati MF, Casero D, Weeks DP, Ladunga I.** 2012. Activation of the carbon concentrating mechanism by CO<sub>2</sub> deprivation coincides with massive transcriptional restructuring in *Chlamydomonas reinhardtii*. *The Plant Cell* **24**, 1860–1875.
- Cai F, Bernstein SL, Wilson SC, Kerfeld CA.** 2016. Production and characterization of synthetic carboxysome shells with incorporated luminal proteins. *Plant Physiology* **170**, 1868–1877.
- Cai F, Dou Z, Bernstein SL, Leverenz R, Williams EB, Heinhorst S, Shively J, Cannon GC, Kerfeld CA.** 2015. Advances in understanding carboxysome assembly in *Prochlorococcus* and *Synechococcus* implicate CsoS2 as a critical component. *Life* **5**, 1141–1171.
- Cai F, Heinhorst S, Shively JM, Cannon GC.** 2008. Transcript analysis of the *Halothiobacillus neapolitanus* cso operon. *Archives of Microbiology* **189**, 141–150.
- Cai F, Menon BB, Cannon GC, Curry KJ, Shively JM, Heinhorst S.** 2009. The pentameric vertex proteins are necessary for the icosahedral carboxysome shell to function as a CO<sub>2</sub> leakage barrier. *PLoS One* **4**, e7521.
- Cai F, Sutter M, Cameron JC, Stanley DN, Kinney JN, Kerfeld CA.** 2013. The structure of CcmP, a tandem bacterial microcompartment domain protein from the β-carboxysome, forms a subcompartment within a microcompartment. *Journal of Biological Chemistry* **288**, 16055–16063.
- Cameron JC, Wilson SC, Bernstein SL, Kerfeld CA.** 2013. Biogenesis of a bacterial organelle: the carboxysome assembly pathway. *Cell* **155**, 1131–1140.
- Chaijarasphong T, Nichols RJ, Kortright KE, Nixon CF, Teng PK, Oltrogge LM, Savage DF.** 2016. Programmed ribosomal frameshifting mediates expression of the α-carboxysome. *Journal of Molecular Biology* **428**, 153–164.
- Chen AH, Robinson-Mosher A, Savage DF, Silver PA, Polka JK.** 2013. The bacterial carbon-fixing organelle is formed by shell envelopment of preassembled cargo. *PLoS One* **8**, e76127.
- Chen YL, Chen LJ, Li HM.** 2016. Polypeptide transport-associated domains of the Toc75 channel protein are located in the intermembrane space of chloroplasts. *Plant Physiology* **172**, 235–243.
- Cheng KH, Colman B.** 1974. Measurements of photorespiration in some microscopic algae. *Planta* **115**, 207–212.
- Chi EY.** 1971. Brown algal pyrenoids. *Protoplasma* **72**, 101–104.
- Compton EL, Karinou E, Naismith JH, Gabel F, Javelle A.** 2011. Low resolution structure of a bacterial SLC26 transporter reveals dimeric stoichiometry and mobile intracellular domains. *Journal of Biological Chemistry* **286**, 27058–27067.
- Conroy MJ, Durand A, Lupo D, Li X-D, Bullough PA, Winkler FK, Merrick M.** 2007. The crystal structure of the Escherichia coli AmtB–GlnK complex reveals how GlnK regulates the ammonia channel. *Proceedings of the National Academy of Sciences, USA* **104**, 1213–1218.
- Cot SS, So AK, Espie GS.** 2008. A multiprotein bicarbonate dehydration complex essential to carboxysome function in cyanobacteria. *Journal of Bacteriology* **190**, 936–945.
- Dalbey RE, Wang P, Kuhn A.** 2011. Assembly of bacterial inner membrane proteins. *Annual Review of Biochemistry* **80**, 161–187.
- de Araujo C, Arefeen D, Tadesse Y, Long BM, Price GD, Rowlett RS, Kimber MS, Espie GS.** 2014. Identification and characterization of a carboxysomal γ-carbonic anhydrase from the cyanobacterium *Nostoc* sp. PCC 7120. *Photosynthesis Research* **121**, 135–150.
- Derelle E, Ferraz C, Rombauts S, et al.** 2006. Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proceedings of the National Academy of Sciences, USA* **103**, 11647–11652.
- Du J, Förster B, Rourke L, Howitt SM, Price GD.** 2014. Characterisation of cyanobacterial bicarbonate transporters in *E. coli* shows that SbtA homologs are functional in this heterologous expression system. *PLoS One* **9**, e115905.
- Duanmu D, Miller AR, Horken KM, Weeks DP, Spalding MH.** 2009a. Knockdown of limiting-CO<sub>2</sub>-induced gene HLA3 decreases HCO<sub>3</sub><sup>−</sup> transport and photosynthetic Ci affinity in *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences, USA* **106**, 5990–5995.
- Duanmu D, Wang Y, Spalding MH.** 2009b. Thylakoid lumen carbonic anhydrase (CAH3) mutation suppresses air-Dier phenotype of LCIB mutant in *Chlamydomonas reinhardtii*. *Plant Physiology* **47**, 929–937.
- Edwards EJ, Osborne CP, Strömberg CA, et al.** 2010. The origins of C<sub>4</sub> grasslands: integrating evolutionary and ecosystem science. *Science* **328**, 587–591.
- Emlyn-Jones D, Woodger FJ, Price GD, Whitney SM.** 2006. RbcX can function as a Rubisco chaperonin, but is non-essential in *Synechococcus* PCC7942. *Plant and Cell Physiology* **47**, 1630–1640.
- Endow JK, Rocha AG, Baldwin AJ, Roston RL, Yamaguchi T, Kamikubo H, Inoue K.** 2016. Polyglycine acts as a rejection signal for protein transport at the chloroplast envelope. *PLoS One* **11**, e0167802.
- Engel BD, Schaffer M, Kuhn Cuellar L, Villa E, Plitzko JM, Baumeister W.** 2015. Native architecture of the *Chlamydomonas* chloroplast revealed by *in situ* cryo-electron tomography. *Elife* **4**, e04889.
- Fang W, Si Y, Douglass S, Casero D, Merchant SS, Pellegrini M, Ladunga I, Liu P, Spalding MH.** 2012. Transcriptome-wide changes in *Chlamydomonas reinhardtii* gene expression regulated by carbon dioxide and the CO<sub>2</sub>-concentrating mechanism regulator CIA5/CCM1. *The Plant Cell* **24**, 1876–1893.
- Frey R, Mantri S, Rocca M, Hilvert D.** 2016. Bottom-up construction of a primordial carboxysome mimic. *Journal of the American Chemical Society* **138**, 10072–10075.
- Friedberg D, Jager KM, Kessel M, Silman NJ, Bergman B.** 1993. Rubisco but not Rubisco activase is clustered in the carboxysomes of the cyanobacterium *Synechococcus* sp. PCC 7942: Mud-induced carboxysomeless mutants. *Molecular Microbiology* **9**, 1193–1201.
- Fuentes P, Zhou F, Erban A, Karcher D, Kopka J, Bock R.** 2016. A new synthetic biology approach allows transfer of an entire metabolic pathway from a medicinal plant to a biomass crop. *Elife* **5**, e13664.
- Fukuzawa H, Miura K, Ishizaki K, Kucho K-I, Saito T, Kohinata T, Ohyama K.** 2001. *Ccm1*, a regulatory gene controlling the induction of a carbon-concentrating mechanism in *Chlamydomonas reinhardtii* by sensing CO<sub>2</sub> availability. *Proceedings of the National Academy of Sciences, USA* **98**, 5347–5352.
- Furbank RT.** 2016. Walking the C<sub>4</sub> pathway: past, present, and future. *Journal of Experimental Botany* **67**, 4057–4066.
- Galmés J, Kapralov MV, Andralojc PJ, Conesa MÀ, Keys AJ, Parry MA, Flexas J.** 2014. Expanding knowledge of the Rubisco kinetics variability in plant species: environmental and evolutionary trends. *Plant, Cell and Environment* **37**, 1989–2001.
- Gaudana SB, Zarzycki J, Moparthy VK, Kerfeld CA.** 2015. Bioinformatic analysis of the distribution of inorganic carbon transporters and prospective targets for bioengineering to increase C<sub>i</sub> uptake by cyanobacteria. *Photosynthesis Research* **126**, 99–109.

- Genkov T, Meyer M, Griffiths H, Spreitzer RJ.** 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**, 19833–19841.
- Giordano M, Beardall J, Raven JA.** 2005. CO<sub>2</sub> concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution. *Annual Review of Plant Biology* **56**, 99–131.
- Gonzales AD, Light YK, Zhang Z, Iqbal T, Lane TW, Martino A.** 2005. Proteomic analysis of the CO<sub>2</sub>-concentrating mechanism in the open-ocean cyanobacterium *Synechococcus* WH8102. *Canadian Journal of Botany* **83**, 735–745.
- Gonzalez-Esquer CR, Shubitowski TB, Kerfeld CA.** 2015. Streamlined construction of the cyanobacterial CO<sub>2</sub>-fixing organelle via protein domain fusions for use in plant synthetic biology. *The Plant Cell* **27**, 2637–2644.
- Griffiths DJ.** 1970. The pyrenoid. *Botanical Review* **36**, 29–58.
- Hanson DT, Renzaglia K, Villarreal JC.** 2014. Diffusion limitation and CO<sub>2</sub> concentrating mechanisms in bryophytes. In: Hanson DT, Rice SK, eds. *Photosynthesis in bryophytes and early land plants*. Dordrecht: Springer Netherlands, 95–111.
- Hauser T, Bhat JY, Miličić G, Wendler P, Hartl FU, Bracher A, Hayer-Hartl M.** 2015. Structure and mechanism of the Rubisco-assembly chaperone Raf1. *Nature Structural and Molecular Biology* **22**, 720–728.
- Heinhorst S, Cannon GC, Shively JM.** 2014. Carboxysomes and their structural organization in prokaryotes. In: Barton LL, Bazylinski DA, Xu H, eds. *Nanomicrobiology: physiological and environmental characteristics*. New York: Springer New York, 75–101.
- Hopkinson BM, Dupont CL, Matsuda Y.** 2016. The physiology and genetics of CO<sub>2</sub> concentrating mechanisms in model diatoms. *Current Opinion in Plant Biology* **31**, 51–57.
- Hopkinson BM, Young JN, Tansik AL, Binder BJ.** 2014. The minimal CO<sub>2</sub>-concentrating mechanism of *Prochlorococcus* spp. MED4 is effective and efficient. *Plant Physiology* **166**, 2205–2217.
- Houtz RL, Magnani R, Nayak NR, Dirk LM.** 2008. Co- and post-translational modifications in Rubisco: unanswered questions. *Journal of Experimental Botany* **59**, 1635–1645.
- Iancu CV, Ding HJ, Morris DM, Dias DP, Gonzales AD, Martino A, Jensen GJ.** 2007. The structure of isolated *Synechococcus* strain WH8102 carboxysomes as revealed by electron cryotomography. *Journal of Molecular Biology* **372**, 764–773.
- Im CS, Grossman AR.** 2002. Identification and regulation of high light-induced genes in *Chlamydomonas reinhardtii*. *The Plant Journal* **30**, 301–313.
- Jin S, Sun J, Wunder T, Tang D, Cousins AB, Sze SK, Mueller-Cajal O, Gao YG.** 2016. Structural insights into the LCIB protein family reveals a new group of β-carbonic anhydrases. *Proceedings of the National Academy of Sciences, USA* **113**, 14716–14721.
- Jordan DB, Ogren WL.** 1981. Species variation in the specificity of ribulose biphosphate carboxylase/oxygenase. *Nature* **291**, 513–515.
- Kaneko Y, Danev R, Nagayama K, Nakamoto H.** 2006. Intact carboxysomes in a cyanobacterial cell visualized by Hilbert differential contrast transmission electron microscopy. *Journal of Bacteriology* **188**, 805–808.
- Keeley JE, Rundel PW.** 2003. Evolution of CAM and C<sub>4</sub> carbon-concentrating mechanisms. *International Journal of Plant Sciences* **164**, S55–S77.
- Kikuchi S, Bédard J, Hirano M, Hirabayashi Y, Oishi M, Imai M, Takase M, Ide T, Nakai M.** 2013. Uncovering the protein translocon at the chloroplast inner envelope membrane. *Science* **339**, 571–574.
- Kimber MS.** 2014. Carboxysomes—sequestering RubisCO for efficient carbon fixation. In: Hohmann-Marriott MF, ed. *The structural basis of biological energy generation*. Dordrecht: Springer Netherlands, 133–148.
- Kinney JN, Salmeen A, Cai F, Kerfeld CA.** 2012. Elucidating essential role of conserved carboxysomal protein CcmN reveals common feature of bacterial microcompartment assembly. *Journal of Biological Chemistry* **287**, 17729–17736.
- Kolesinski P, Belusiak I, Czarnocki-Cieciura M, Szczepaniak A.** 2014. Rubisco accumulation factor 1 from *Thermosynechococcus elongatus* participates in the final stages of ribulose-1,5-bisphosphate carboxylase/oxygenase assembly in *Escherichia coli* cells and *in vitro*. *FEBS Journal* **281**, 3920–3932.
- Kowallik K.** 1969. The crystal lattice of the pyrenoid matrix of *Procoentrum micans*. *Journal of Cell Science* **5**, 251–269.
- Kranz SA, Eichner M, Rost B.** 2011. Interactions between CCM and N<sub>2</sub> fixation in *Trichodesmium*. *Photosynthesis Research* **109**, 73–84.
- Kromdijk J, Glowacka K, Leonelli L, Gabilly ST, Iwai M, Niyogi KK, Long SP.** 2016. Improving photosynthesis and crop productivity by accelerating recovery from photoprotection. *Science* **354**, 857–861.
- Leadbeater BSC, Manton I.** 1971. Fine structure and light microscopy of a new species of *Chrysochromulina* (*C. acantha*). *Archiv für Mikrobiologie* **78**, 58–69.
- Leakey AD, Ainsworth EA, Bernacchi CJ, Rogers A, Long SP, Ort DR.** 2009. Elevated CO<sub>2</sub> effects on plant carbon, nitrogen, and water relations: six important lessons from FACE. *Journal of Experimental Botany* **60**, 2859–2876.
- Li FW, Villarreal JC, Szovenyi P.** 2017. Hornworts: an overlooked window into carbon-concentrating mechanisms. *Trends in Plant Science* **22**, 275–277.
- Li LA, Zianni MR, Tabita FR.** 1999. Inactivation of the monocistronic *rca* gene in *Anabaena variabilis* suggests a physiological ribulose biphosphate carboxylase/oxygenase activase-like function in heterocystous cyanobacteria. *Plant Molecular Biology* **40**, 467–478.
- Li X, Zhang R, Patena W, et al.** 2016. An indexed, mapped mutant window enables reverse genetics studies of biological processes in *Chlamydomonas reinhardtii*. *The Plant Cell* **28**, 367–387.
- Lin MT, Occhialini A, Andralojc PJ, Devonshire J, Hines KM, Parry MA, Hanson MR.** 2014a. β-Carboxysomal proteins assemble into highly organized structures in *Nicotiana* chloroplasts. *The Plant Journal* **79**, 1–12.
- Lin MT, Occhialini A, Andralojc PJ, Parry MA, Hanson MR.** 2014b. A faster Rubisco with potential to increase photosynthesis in crops. *Nature* **513**, 547–550.
- Long BM, Badger MR, Whitney SM, Price GD.** 2007. Analysis of carboxysomes from *Synechococcus* PCC7942 reveals multiple Rubisco complexes with carboxysomal proteins CcmM and CcaA. *Journal of Biological Chemistry* **282**, 29323–29335.
- Long BM, Price GD, Badger MR.** 2005. Proteomic assessment of an established technique for carboxysome enrichment from *Synechococcus* PCC7942. *Canadian Journal of Botany* **83**, 746–757.
- Long BM, Rae BD, Badger MR, Price GD.** 2011. Over-expression of the β-carboxysomal CcmM protein in *Synechococcus* PCC7942 reveals a tight co-regulation of carboxysomal carbonic anhydrase (CcaA) and M58 content. *Photosynthesis Research* **109**, 33–45.
- Long BM, Rae BD, Rolland V, Förster B, Price GD.** 2016. Cyanobacterial CO<sub>2</sub>-concentrating mechanism components: function and prospects for plant metabolic engineering. *Current Opinion in Plant Biology* **31**, 1–8.
- Long BM, Tucker L, Badger MR, Price GD.** 2010. Functional cyanobacterial beta-carboxysomes have an absolute requirement for both long and short forms of the CcmM protein. *Plant Physiology* **153**, 285–293.
- Long SP, Ainsworth EA, Leakey AD, Nösberger J, Ort DR.** 2006. Food for thought: lower-than-expected crop yield stimulation with rising CO<sub>2</sub> concentrations. *Science* **312**, 1918–1921.
- Long SP, Marshall-Colon A, Zhu XG.** 2015. Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. *Cell* **161**, 56–66.
- Ludwig M, Sültemeyer D, Price GD.** 2000. Isolation of *ccmKLMN* genes from the marine cyanobacterium *Synechococcus* sp. PCC7002 (Cyanophyceae), and evidence that CcmM is essential for carboxysome assembly. *Journal of Phycology* **36**, 1109–1119.
- Ma Y, Pollock SV, Xiao Y, Cunnusamy K, Moroney JV.** 2011. Identification of a novel gene, *CIA6*, required for normal pyrenoid formation in *Chlamydomonas reinhardtii*. *Plant Physiology* **156**, 884–896.
- Mackinder LCM, Meyer MT, Mettler-Altmann T, et al.** 2016. A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle. *Proceedings of the National Academy of Sciences, USA* **113**, 5958–5963.
- Mangan N, Brenner M.** 2014. Systems analysis of the CO<sub>2</sub> concentrating mechanism in cyanobacteria. *Elife* **3**, e02043.
- Mangan NM, Flammholz A, Hood RD, Milo R, Savage DF.** 2016. pH determines the energetic efficiency of the cyanobacterial CO<sub>2</sub>

concentrating mechanism. Proceedings of the National Academy of Sciences, USA **113**, E5354–E5362.

**Mariscal V, Moulin P, Orsel M, Miller AJ, Fernández E, Galván A.** 2006. Differential regulation of the *Chlamydomonas* *Nar1* gene family by carbon and nitrogen. *Protist* **157**, 421–433.

**Markowitz MM, Hoffman LR.** 1974. Chloroplast inclusions in zoospores of *Oedocladium*. *Journal of Phycology* **10**, 308–315.

**McGrath JM, Long SP.** 2014. Can the cyanobacterial carbon-concentrating mechanism increase photosynthesis in crop species? A theoretical analysis. *Plant Physiology* **164**, 2247–2261.

**McKay RML, Gibbs SP, Espie GS.** 1993. Effect of dissolved inorganic carbon on the expression of carboxysomes, localization of Rubisco and the mode of inorganic carbon transport in cells of the cyanobacterium *Synechococcus* UTEX 625. *Archives of Microbiology* **159**, 21–29.

**McKay RML, Gibbs SP, Vaughn KC.** 1991. RuBisCo activase is present in the pyrenoid of green algae. *Protoplasma* **162**, 38–45.

**Menon BB, Dou Z, Heinhorst S, Shively JM, Cannon GC.** 2008. *Halothiobacillus neapolitanus* carboxysomes sequester heterologous and chimeric RubisCO species. *PLoS One* **3**, e3570.

**Menon BB, Heinhorst S, Shively JM, Cannon GC.** 2010. The carboxysome shell is permeable to protons. *Journal of Bacteriology* **192**, 5881–5886.

**Meyer M, Griffiths H.** 2013. Origins and diversity of eukaryotic CO<sub>2</sub>-concentrating mechanisms: lessons for the future. *Journal of Experimental Botany* **64**, 769–786.

**Meyer MT.** 2010. Physiological and molecular determinants of the *Chlamydomonas reinhardtii* pyrenoid. Doctoral thesis, Cambridge University.

**Meyer MT, Genkov T, Skepper JN, Jouhet J, Mitchell MC, Spreitzer RJ, Griffiths H.** 2012. Rubisco small-subunit  $\alpha$ -helices control pyrenoid formation in *Chlamydomonas*. Proceedings of the National Academy of Sciences, USA **109**, 19474–19479.

**Meyer MT, McCormick AJ, Griffiths H.** 2016. Will an algal CO<sub>2</sub>-concentrating mechanism work in higher plants? *Current Opinion in Plant Biology* **31**, 181–188.

**Mitchell MC, Meyer MT, Griffiths H.** 2014. Dynamics of carbon-concentrating mechanism induction and protein relocalization during the dark-to-light transition in synchronized *Chlamydomonas reinhardtii*. *Plant Physiology* **166**, 1073–1082.

**Miura K, Yamano T, Yoshioka S, et al.** 2004. Expression profiling-based identification of CO<sub>2</sub>-responsive genes regulated by CCM1 controlling a carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. *Plant Physiology* **135**, 1595–1607.

**Morita E, Abe T, Tsuzuki M, Fujiwara S, Sato N, Hirata A, Sonoike K, Nozaki H.** 1998. Presence of the CO<sub>2</sub>-concentrating mechanism in some species of the pyrenoid-less free-living algal genus *Chloromonas* (Volvocales, Chlorophyta). *Planta* **204**, 269–276.

**Morita E, Abe T, Tsuzuki M, Fujiwara S, Sato N, Hirata A, Sonoike K, Nozaki H.** 1999. Role of pyrenoids in the CO<sub>2</sub>-concentrating mechanism: comparative morphology, physiology and molecular phylogenetic analysis of closely related strains of *Chlamydomonas* and *Chloromonas* (Volvocales). *Planta* **208**, 365–372.

**Müller M, Kunz HH, Schroeder JI, Kemp G, Young HS, Neuhaus HE.** 2014. Decreased capacity for sodium export out of Arabidopsis chloroplasts impairs salt tolerance, photosynthesis and plant performance. *The Plant Journal* **78**, 646–658.

**Nakamura Y, Kaneko T, Sato S, et al.** 2003. Complete genome structure of *Gloeobacter violaceus* PCC 7421, a cyanobacterium that lacks thylakoids (supplement). *DNA Research* **10**, 181–201.

**Occhialini A, Lin MT, Andralojc PJ, Hanson MR, Parry MA.** 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<sub>2</sub>. *The Plant Journal* **85**, 148–160.

**Oh YJ, Hwang I.** 2015. Targeting and biogenesis of transporters and channels in chloroplast envelope membranes: unsolved questions. *Cell Calcium* **58**, 122–130.

**Ohnishi N, Mukherjee B, Tsujikawa T, Yanase M, Nakano H, Moroney JV, Fukuzawa H.** 2010. Expression of a low CO<sub>2</sub>-inducible protein, LCI1, increases inorganic carbon uptake in the green alga *Chlamydomonas reinhardtii*. *The Plant Cell* **22**, 3105–3117.

**Okawa K, Inoue H, Adachi F, Nakayama K, Ito-Inaba Y, Schnell DJ, Uehara S, Inaba T.** 2014. Targeting of a polytopic membrane protein to the inner envelope membrane of chloroplasts *in vivo* involves multiple transmembrane segments. *Journal of Experimental Botany* **65**, 5257–5265.

**Omata T, Price GD, Badger MR, Okamura M, Gohta S, Ogawa T.** 1999. Identification of an ATP-binding cassette transporter involved in bicarbonate uptake in the cyanobacterium *Synechococcus* sp. strain PCC 7942. Proceedings of the National Academy of Sciences, USA **96**, 13571–13576.

**Parry MA, Andralojc PJ, Scales JC, Salvucci ME, Carmo-Silva AE, Alonso H, Whitney SM.** 2013. Rubisco activity and regulation as targets for crop improvement. *Journal of Experimental Botany* **64**, 717–730.

**Patron NJ.** 2016. Blueprints for green biotech: development and application of standards for plant synthetic biology. *Biochemical Society Transactions* **44**, 702–708.

**Pearce FG.** 2006. Catalytic by-product formation and ligand binding by ribulose biphosphate carboxylases from different phylogenies. *Biochemical Journal* **399**, 525–534.

**Peña KL, Castel SE, de Araujo C, Espie GS, Kimber MS.** 2010. Structural basis of the oxidative activation of the carboxysomal  $\gamma$ -carbonic anhydrase, CcmM. Proceedings of the National Academy of Sciences, USA **107**, 2455–2460.

**Pengelly JJ, Förster B, von Caemmerer S, Badger MR, Price GD, Whitney SM.** 2014. Transplastomic integration of a cyanobacterial bicarbonate transporter into tobacco chloroplasts. *Journal of Experimental Botany* **65**, 3071–3080.

**Pickersgill RW.** 1986. An upper limit to the active site concentration of Rubisco in chloroplasts. *Biochemical Journal* **236**, 311.

**Pollock SV, Colombo SL, Prout DL Jr, Godfrey AC, Moroney JV.** 2003. Rubisco activase is required for optimal photosynthesis in the green alga *Chlamydomonas reinhardtii* in a low-CO<sub>2</sub> atmosphere. *Plant Physiology* **133**, 1854–1861.

**Powell K.** 2016. Martin Jonikas: bringing high-throughput genetics to photosynthesis. *Journal of Cell Biology* **213**, 139–140.

**Price GD.** 2011. Inorganic carbon transporters of the cyanobacterial CO<sub>2</sub> concentrating mechanism. *Photosynthesis Research* **109**, 47–57.

**Price GD, Badger MR.** 1989a. Expression of human carbonic anhydrase in the cyanobacterium *Synechococcus* PCC7942 creates a high CO<sub>2</sub>-requiring phenotype: evidence for a central role for carboxysomes in the CO<sub>2</sub> concentrating mechanism. *Plant Physiology* **91**, 505–513.

**Price GD, Badger MR.** 1989b. Isolation and characterization of high CO<sub>2</sub>-requiring-mutants of the cyanobacterium *Synechococcus* PCC7942: two phenotypes that accumulate inorganic carbon but are apparently unable to generate CO<sub>2</sub> within the carboxysome. *Plant Physiology* **91**, 514–525.

**Price GD, Badger MR, von Caemmerer S.** 2011a. The prospect of using cyanobacterial bicarbonate transporters to improve leaf photosynthesis in C<sub>3</sub> crop plants. *Plant Physiology* **155**, 20–26.

**Price GD, Badger MR, Woodger FJ, Long BM.** 2008. Advances in understanding the cyanobacterial CO<sub>2</sub>-concentrating-mechanism (CCM): functional components, C<sub>i</sub> transporters, diversity, genetic regulation and prospects for engineering into plants. *Journal of Experimental Botany* **59**, 1441–1461.

**Price GD, Howitt SM.** 2014. Topology mapping to characterize cyanobacterial bicarbonate transporters: BicA (SulP/SLC26 family) and SbtA. *Molecular Membrane Biology* **31**, 177–182.

**Price GD, Pengelly JJ, Forster B, Du J, Whitney SM, von Caemmerer S, Badger MR, Howitt SM, Evans JR.** 2013. The cyanobacterial CCM as a source of genes for improving photosynthetic CO<sub>2</sub> fixation in crop species. *Journal of Experimental Botany* **64**, 753–768.

**Price GD, Shelden MC, Howitt SM.** 2011b. Membrane topology of the cyanobacterial bicarbonate transporter, SbtA, and identification of potential regulatory loops. *Molecular Membrane Biology* **28**, 265–275.

**Price GD, Woodger FJ, Badger MR, Howitt SM, Tucker L.** 2004. Identification of a SulP-type bicarbonate transporter in marine cyanobacteria. Proceedings of the National Academy of Sciences, USA **101**, 18228–18233.

**Prins A, Orr DJ, Andralojc PJ, Reynolds MP, Carmo-Silva E, Parry MA.** 2016. Rubisco catalytic properties of wild and domesticated

relatives provide scope for improving wheat photosynthesis. *Journal of Experimental Botany* **67**, 1827–1838.

**Rae BD, Förster B, Badger MR, Price GD.** 2011. The CO<sub>2</sub>-concentrating mechanism of *Synechococcus* WH5701 is composed of native and horizontally-acquired components. *Photosynthesis Research* **109**, 59–72.

**Rae BD, Long BM, Badger MR, Price GD.** 2012. Structural determinants of the outer shell of β-carboxysomes in *Synechococcus elongatus* PCC 7942: roles for CcmK2, K3-K4, CcmO, and CcmL. *PLoS One* **7**, e43871.

**Rae BD, Long BM, Badger MR, Price GD.** 2013a. Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO<sub>2</sub> fixation in cyanobacteria and some proteobacteria. *Microbiology and Molecular Biology Reviews* **77**, 357–379.

**Rae BD, Long BM, Whitehead LF, Förster B, Badger MR, Price GD.** 2013b. Cyanobacterial carboxysomes: microcompartments that facilitate CO<sub>2</sub> fixation. *Journal of Molecular Microbiology and Biotechnology* **23**, 300–307.

**Ramazanov Z, Rawat M, Henk MC, Mason CB, Matthews SW, Moroney JV.** 1994. The induction of the CO<sub>2</sub>-concentrating mechanism is correlated with the formation of the starch sheath around the pyrenoid of *Chlamydomonas reinhardtii*. *Planta* **195**, 210–216.

**Raven JA, Beardall J.** 2016. The ins and outs of CO<sub>2</sub>. *Journal of Experimental Botany* **67**, 1–13.

**Rawat M, Henk MC, Lavigne LL, Moroney JV.** 1996. *Chlamydomonas reinhardtii* mutants without ribulose-1,5-bisphosphate carboxylase-oxygenase lack a detectable pyrenoid. *Planta* **198**, 263–270.

**Roberts EW, Cai F, Kerfeld CA, Cannon GC, Heinhorst S.** 2012. Isolation and characterization of the Prochlorococcus carboxysome reveal the presence of the novel shell protein CsoS1D. *Journal of Bacteriology* **194**, 787–795.

**Rolland N, Dorne AJ, Amoroso G, Sültemeyer DF, Joyard J, Rochaix JD.** 1997. Disruption of the plastid ycf10 open reading frame affects uptake of inorganic carbon in the chloroplast of *Chlamydomonas*. *EMBO Journal* **16**, 6713–6726.

**Rolland V, Badger MR, Price GD.** 2016. Redirecting the cyanobacterial bicarbonate transporters BicA and SbtA to the chloroplast envelope: soluble and membrane cargos need different chloroplast targeting signals in plants. *Frontiers in Plant Science* **7**, 185.

**Sasaki Y, Sekiguchi K, Nagano Y, Matsuno R.** 1993. Chloroplast envelope protein encoded by chloroplast genome. *FEBS Letters* **316**, 93–98.

**Schmid AMM.** 2001. Value of pyrenoids in the systematics of diatoms: their morphology and ultrastructure. In: Economou-Amilli A, ed. 16th International Diatom Symposium. Athens: University of Athens Press.

**Schuler ML, Mantegazza O, Weber AP.** 2016. Engineering C<sub>4</sub> photosynthesis into C<sub>3</sub> chassis in the synthetic biology age. *The Plant Journal* **87**, 51–65.

**Sharkey TD.** 2001. Photorespiration. eLS.

**Sharwood RE.** 2017. Engineering chloroplasts to improve Rubisco catalysis: prospects for translating improvements into food and fiber crops. *New Phytologist* **213**, 494–510.

**Shelden MC, Howitt SM, Price GD.** 2010. Membrane topology of the cyanobacterial bicarbonate transporter, BicA, a member of the SulP (SLC26A) family. *Molecular Membrane Biology* **27**, 12–22.

**Shibata M, Katoh H, Sonoda M, Ohkawa H, Shimoyama M, Fukuzawa H, Kaplan A, Ogawa T.** 2002. Genes essential to sodium-dependent bicarbonate transport in cyanobacteria: function and phylogenetic analysis. *Journal of Biological Chemistry* **277**, 18658–18664.

**Shih PM.** 2015. Photosynthesis and early Earth. *Current Biology* **25**, R855–R859.

**Shih PM, Occhialini A, Cameron JC, Andralojc PJ, Parry MA, Kerfeld CA.** 2016. Biochemical characterization of predicted Precambrian RuBisCO. *Nature Communications* **7**, 10382.

**Shih PM, Wu D, Latifi A, et al.** 2013. Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proceedings of the National Academy of Sciences, USA* **110**, 1053–1058.

**Simkin AJ, McAusland L, Headland LR, Lawson T, Raines CA.** 2015. Multigene manipulation of photosynthetic carbon assimilation increases

CO<sub>2</sub> fixation and biomass yield in tobacco. *Journal of Experimental Botany* **66**, 4075–4090.

**Sinetova MA, Kupriyanova EV, Markelova AG, Allakhverdiev SI, Pronina NA.** 2012. Identification and functional role of the carbonic anhydrase Cah3 in thylakoid membranes of pyrenoid of *Chlamydomonas reinhardtii*. *Biochimica et Biophysica Acta* **1817**, 1248–1255.

**Smith EC, Griffiths H.** 2000. The role of carbonic anhydrase in photosynthesis and the activity of the carbon-concentrating-mechanism in bryophytes of the class Anthocerotae. *New Phytologist* **145**, 29–37.

**So AK, Espie GS.** 1998. Cloning, characterization and expression of carbonic anhydrase from the cyanobacterium *Synechocystis* PCC6803. *Plant Molecular Biology* **37**, 205–215.

**So AK, Espie GS, Williams EB, Shively JM, Heinhorst S, Cannon GC.** 2004. A novel evolutionary lineage of carbonic anhydrase (epsilon class) is a component of the carboxysome shell. *Journal of Bacteriology* **186**, 623–630.

**Sutter M, Roberts EW, Gonzalez RC, Bates C, Dawoud S, Landry K, Cannon GC, Heinhorst S, Kerfeld CA.** 2015. Structural characterization of a newly identified component of α-carboxysomes: the AAA+ domain protein CsoCbbQ. *Scientific Reports* **5**, 16243.

**Szövényi P, Frangedakis E, Ricca M, Quandt D, Wicke S, Langdale JA.** 2015. Establishment of *Anthoceros agrestis* as a model species for studying the biology of hornworts. *BMC Plant Biology* **15**, 98.

**Tabita FR.** 1999. Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: a different perspective. *Photosynthesis Research* **60**, 1–28.

**Tcherkez GG, Farquhar GD, Andrews TJ.** 2006. Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proceedings of the National Academy of Sciences, USA* **103**, 7246–7251.

**Theg SM, Scott SV.** 1993. Protein import into chloroplasts. *Trends in Cell Biology* **3**, 186–190.

**Tholen D, Zhu XG.** 2011. The mechanistic basis of internal conductance: a theoretical analysis of mesophyll cell photosynthesis and CO<sub>2</sub> diffusion. *Plant Physiology* **156**, 90–105.

**Tsai YC, Lapina MC, Bhushan S, Mueller-Cajar O.** 2015. Identification and characterization of multiple Rubisco activases in chemoautotrophic bacteria. *Nature Communications* **6**, 8883.

**Uehara S, Adachi F, Ito-Inaba Y, Inaba T.** 2016. Specific and efficient targeting of cyanobacterial bicarbonate transporters to the inner envelope membrane of chloroplasts in *Arabidopsis*. *Frontiers in Plant Science* **7**, 16.

**Vaughn KC, Campbell EO, Hasegawa J, Owen HA, Renzaglia KS.** 1990. The pyrenoid is the site of ribulose 1,5-bisphosphate carboxylase/oxygenase accumulation in the hornwort (Bryophyta: Anthocerotae) chloroplast. *Protoplasma* **156**, 117–129.

**Viana AAB, Li M, Schnell DJ.** 2010. Determinants for stop-transfer and post-import pathways for protein targeting to the chloroplast inner envelope membrane. *Journal of Biological Chemistry* **285**, 12948–12960.

**Waadt R, Schmidt LK, Lohse M, Hashimoto K, Bock R, Kudla J.** 2008. Multicolor bimolecular fluorescence complementation reveals simultaneous formation of alternative CBL/CIPK complexes in planta. *The Plant Journal* **56**, 505–516.

**Wang L, Yamano T, Takane S, et al.** 2016. Chloroplast-mediated regulation of CO<sub>2</sub>-concentrating mechanism by Ca<sup>2+</sup>-binding protein CAS in the green alga *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences, USA* **113**, 12586–12591.

**Wang S, Tholen D, Zhu XG.** 2017. C<sub>4</sub> photosynthesis in C<sub>3</sub> rice: a theoretical analysis of biochemical and anatomical factors. *Plant, Cell and Environment* **40**, 80–94.

**Wang Y, Huang Y, Wang J, et al.** 2009. Structure of the formate transporter FocA reveals a pentameric aquaporin-like channel. *Nature* **462**, 467–472.

**Wang Y, Spalding MH.** 2006. An inorganic carbon transport system responsible for acclimation specific to air levels of CO<sub>2</sub> in *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences, USA* **103**, 10110–10115.

**Wang Y, Spalding MH.** 2014. Acclimation to very low CO<sub>2</sub>: contribution of limiting CO<sub>2</sub> inducible proteins, LCIB and LCIA, to inorganic carbon uptake in *Chlamydomonas reinhardtii*. *Plant Physiology* **166**, 2040–2050.

**Wang Y, Stessman DJ, Spalding MH.** 2015. The CO<sub>2</sub> concentrating mechanism and photosynthetic carbon assimilation in limiting CO<sub>2</sub>: how *Chlamydomonas* works against the gradient. *The Plant Journal* **82**, 429–448.

- Wheatley NM, Sundberg CD, Gidaniyan SD, Cascio D, Yeates TO.** 2014. Structure and identification of a pterin dehydratase-like protein as a ribulose-bisphosphate carboxylase/oxygenase (RuBisCO) assembly factor in the  $\alpha$ -carboxysome. *Journal of Biological Chemistry* **289**, 7973–7981.
- Whitehead L, Long BM, Price GD, Badger MR.** 2014. Comparing the in vivo function of  $\alpha$ -carboxysomes and  $\beta$ -carboxysomes in two model cyanobacteria. *Plant Physiology* **165**, 398–411.
- Whitney SM, Birch R, Kelso C, Beck JL, Kapralov MV.** 2015. Improving recombinant Rubisco biogenesis, plant photosynthesis and growth by coexpressing its ancillary RAF1 chaperone. *Proceedings of the National Academy of Sciences, USA* **112**, 3564–3569.
- Whitney SM, Houtz RL, Alonso H.** 2011. Advancing our understanding and capacity to engineer nature's CO<sub>2</sub>-sequestering enzyme, Rubisco. *Plant Physiology* **155**, 27–35.
- Wilson R, Whitney S.** 2015. Photosynthesis: getting it together for CO<sub>2</sub> fixation. *Nature Plants* **1**, 15130.
- Wilson RH, Alonso H, Whitney SM.** 2016. Evolving *Methanococcoides burtonii* archaeal Rubisco for improved photosynthesis and plant growth. *Scientific Reports* **6**, 22284.
- Woodward FI, Lomas MR, Kelly CK.** 2004. Global climate and the distribution of plant biomes. *Philosophical Transactions of the Royal Society B: Biological Sciences* **359**, 1465–1476.
- Xiang Y, Zhang J, Weeks DP.** 2001. The Cia5 gene controls formation of the carbon concentrating mechanism in *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences, USA* **98**, 5341–5346.
- Yamano T, Sato E, Iguchi H, Fukuda Y, Fukuzawa H.** 2015. Characterization of cooperative bicarbonate uptake into chloroplast stroma in the green alga *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences, USA* **112**, 7315–7320.
- Zhang P, Battchikova N, Jansen T, Appel J, Ogawa T, Aro EM.** 2004. Expression and functional roles of the two distinct NDH-1 complexes and the carbon acquisition complex NdhD3/NdhF3/CupA/SII1735 in *Synechocystis* sp PCC 6803. *The Plant Cell* **16**, 3326–3340.
- Zygadlo A, Robinson C, Scheller HV, Mant A, Jensen PE.** 2006. The properties of the positively charged loop region in PSI-G are essential for its 'spontaneous' insertion into thylakoids and rapid assembly into the photosystem I complex. *Journal of Biological Chemistry* **281**, 10548–10554.