Orchestral manoeuvres in the light: Crosstalk needed for regulation of the Chlamydomonas carbon concentration mechanism

Indu Santhanagopalan¹, Rachel Wong, Tanya Mathur and Howard Griffiths¹

Department of Plant Sciences, Downing Street
University of Cambridge
CB2 3EA UK

¹for correspondence
Abstract

The inducible carbon concentrating mechanism in *Chlamydomonas reinhardtii* has been well defined from a molecular and ultrastructural perspective. Inorganic carbon transport proteins, and strategically located carbonic anhydrases deliver CO$_2$ within the chloroplast pyrenoid matrix where Rubisco is packaged. However, there is little understanding of the fundamental signalling and sensing processes leading to CCM induction. While external CO$_2$ limitation has been believed to be the primary cue, the coupling between energetic supply and inorganic carbon demand through regulatory feedback from light harvesting and photorespiration signals, could provide the original CCM trigger. Key questions regarding the integration of these processes are addressed in this review. We consider how the chloroplast functions as a crucible for photosynthesis, importing and integrating nuclear-encoded components from the cytoplasm, and sending retrograde signals to the nucleus to regulate CCM induction. We hypothesise that induction of the CCM is associated with retrograde signals associated with photorespiration and/or light stress. We have also examined the significance of common evolutionary pressures for origins of two co-regulated processes- CCM and photorespiration, in addition to identifying genes of interest involved in transcription, protein folding and regulatory processes which are needed to fully understand the processes leading to CCM induction.
Introduction

The carbon concentrating mechanism (CCM) traits found in algae (and cyanobacteria) have evolved to improve the operating efficiency of Rubisco, which is normally packaged within a specific microcompartment: in algae, this is the chloroplast pyrenoid. Inorganic carbon, in the form of bicarbonate, is delivered to the chloroplast stroma using a series of membrane transporters. Saturating internal CO$_2$ concentrations (Ci), about x40 above ambient (Badger et al., 1980), are generated within the pyrenoid by strategically placed transporters of inorganic carbon and carbonic anhydrases (CAs) (Hennacy and Jonikas, 2020; Meyer and Griffiths, 2013; Moroney and Ynalvez, 2007). The availability of a sequenced genome (Merchant et al., 2007), transcriptomic studies for synchronised cells across 24h light/dark cycles (Strenkert et al., 2019; Zones et al., 2015) and extensive mutant libraries (Li et al., 2019; Li et al., 2016; Vilarrasa-Blasi et al., 2020) for Chlamydomonas have provided additional opportunities for CCM characterisation.

Substantial molecular and mechanistic advances in our understanding of algal CCM have been recently reviewed (Goudet et al., 2020; Hennacy and Jonikas, 2020; Meyer and Griffiths, 2013; Meyer et al., 2017). CCM induction is associated with enhancement of aggregation of Rubisco with specific linker proteins (He et al., 2020; Mackinder et al., 2016; Meyer et al., 2020) in the pyrenoid surrounded by a starch sheath, with an existing network of knotted tubules making connections with thylakoid stacks (Engel et al., 2015; Meyer et al., 2017). Establishment of the CCM must be co-ordinated between chloroplast and nucleus in sensing induction stimuli, triggering CCM genes-expression, translation, and intracellular transport and assembly of CCM proteins.

The aim of this review is to characterize the various novel aspects of molecular mechanisms leading to mechanisms of CCM induction and establishment. We explore the regulatory interplay between environmental sensing, photosynthesis and CCM induction that is critical for Chlamydomonas and identify future avenues for investigation. First, we consider the control of nuclear gene expression and the need to identify transcription factors (TFs) associated with sensing the different environmental stimuli- light and CO$_2$; second, regulation of export of translated proteins and folding within the chloroplast, and associated chaperone systems; third, formation of the pyrenoid matrix, starch sheath and intrapyrenoidal tubule network; fourth, the role of retrograde signalling in delivering signals to alter nuclear gene expression; finally, we discuss the potential evolution of CCM induction from existing photorespiration regulatory mechanisms, through the master regulator CIA5.
I. CCM induction and control of gene expression

Changes in [CO₂] in the external medium have traditionally been thought to be sensed through Ci of the photosynthesizing algal cell and conveyed to the nucleus to change the expression of genes that turn on/off the CCM. Studies carried out with asynchronous cells grown under continuous light revealed [CO₂]-dependent expression for over 5000 genes at the transcription level (Brueggeman et al., 2012; Fang et al., 2012). Over 600 of these differentially expressed genes identified in these genome-wide studies have been implicated in CCM (Mackinder et al., 2017). With several hundred genes orchestrating CCM, the following questions come to the fore: (a) What are the regulators in the nucleus that respond to [CO₂]/Ci changes, and (b) how do they bring about changes at the transcriptional level? In this section, we discuss regulatory mechanisms operating at the transcriptional level to modulate the inducible CCM in Chlamydomonas.

CIA5- The ‘master-regulator’ of CCM

One of the first interesting candidates to be identified as a ‘CCM master regulator’ was CIA5/CCM1. CIA5 was identified in 1989 as essential for growth in limiting CO₂ conditions through studies on a UV-generated mutant, cia5 (Moroney et al., 1989). Studies in the early 2000s established CIA5 as being essential for induction of expression of several CCM genes inorganic carbon transporters HLA3, LCI1, CAS -CAH3, CAH1, alanine α-ketoglutarate aminotransferase ATT1, pyrenoid protein EPYC1, transcription factor LCR1, and mitochondrial membrane proteins - CCP1, CCP2 (Fukuzawa et al., 2001; Miura et al., 2004; Xiang et al., 2001). It must be noted that genes denoted as CCM genes in this review are based on previous findings (Mackinder et al., 2017; Strenkert et al., 2019). Genome-wide studies (Fang et al., 2012) showed CIA5-dependent-expression for 15% genes, but only around half of these CIA5-dependent genes responded to changes in [CO₂]. Furthermore, the mechanism of CIA5 in helping cells acclimate to [CO₂] changes is not understood. CIA5 is proposed to be a TF based on the presence of two Zn-finger domains (Fukuzawa et al., 2001; Xiang et al., 2001). The DNA activating region (Chen, 2016) and [CO₂]-dependent domain that triggers the expression of CIA5-dependent-CCM genes (Xiang et al., 2001), lie in the C-terminal end sequences of 130 and 54 residues respectively. Primarily, the expression of CIA5 seems to be [CO₂]-independent (Fang et al., 2012), and associated [CO₂]-dependent expression changes in the genome are believed to be mediated by post-translational modifications of CIA5 (Chen, 2016). This is supported by CIA5 having several putative sites.
for phosphorylation, glycosylation and myristoylation (Fukuzawa et al., 2001), and anomalous electrophoretic mobility (Chen, 2016).

Absence of evidence for DNA-CIA5 complexes suggests that CIA5 might act indirectly through other proteins, although no such proteins have been identified (Kohinata et al., 2008). Recombinantly expressed full-length CIA5 showed very weak affinity in vitro for the 9-bp sequence (GGGGCGGGG), identified from analysis of upstream sequences of select CIA5-dependent genes (Chen, 2016). However, no motif-dependent binding for CIA5 could be established in vivo when genes with upstream mutated motifs showed similar expression patterns to those of non-mutated ones (Chen, 2016).

Understanding CIA5 mechanism will require identification of the different post-translationally modified forms of CIA5 and the corresponding cis-regulatory elements of CCM genes. The roles played by CIA5 are revisited in the context of chaperone expression (Section II), and the implications of strikingly similar CIA5-dependent expression profiles of photorespiratory and CCM genes are explored below (Sections I and V).

**Other transcriptional and post-transcriptional regulators of CCM**

A search for other CCM TFs and transcription regulators (TRs) has been made over the years. LCR1 (Yoshioka et al., 2004) is a Myb TF that plays a crucial role in CCM by regulating expression of CAH1, LCII, and LC16. LCR1 expression is [CO₂]- and CIA5-dependent. Absence of LCR1 leads to a reduction in affinity for Ci (Yoshioka et al., 2004).

Sequence analysis identified 234 genes as potential TFs and TRs in *Chlamydomonas* (Riano-Pachon et al., 2008), and they need to be checked for their activity in CCM regulation. A recent proteomic analysis (Arias et al., 2020) on nuclei obtained from *Chlamydomonas* grown in 5% CO₂/0.04% CO₂, revealed the presence of 117 proteins which were potential TFs/TRs. Of these 117 nuclear proteins, 35 were of differential abundance dependent on [CO₂], which have been listed in Table 1. It is worth noting that neither CIA5 nor LCR1 were detected in the nuclear proteome in this study. However, the candidates identified in this study might act as a good starting point for investigating other TFs and TRs regulating CCM.

What also remains to be determined are the cis-DNA elements that respond to these TFs and TRs. *Chlamydomonas CAH1* is the only CCM gene for which regulatory elements have been systematically investigated and identified. The 5' upstream 543 bp region of the *CAH1* gene was shown to contain a silencer region and an enhancer region with enhancer
elements EE-1 (AGATTTTCACCGGTGGAAGGAGGT) and EE-2 (CGACTTACGAA) (Kucho et al., 1999; Kucho et al., 2003). The upstream regions of duplicated genes CAH4 and CAH5, which confer CO2 dependence in the presence of light was narrowed to 194 bp. No similarities were seen between the upstream regions of CAH4/5 and CAH1, and no shorter segments in this 194 bp have been identified as the elements responsible for the [CO2]-dependent transcription (Villand et al., 1997). Potential genome-wide cis-DNA regulatory elements in Chlamydomonas that had been shifted from high to low [CO2] were identified by FAIRE-seq (Winck et al., 2013). The potential regulatory regions in these genes require further validation.

While the above studies help identify specific transcription regulatory elements, Chlamydomonas also carries a post-transcriptional regulatory machinery of an extensive system of small RNAs (sRNAs), with 3 argonaute and 3 dicer-like proteins (Chung et al., 2019; Molnar et al., 2007; Valli et al., 2016; Zhao et al., 2007). This gene expression modulatory system with 6164 loci predicted to give rise to sRNAs (Muller et al., 2020) needs potential targets identified. Whether this extensive system of sRNAs regulates CCM requires investigation.

**Mechanisms of regulation of CCM induction: [CO2] is not the only cue**

The CCM genes that appear to be responsive to [CO2] changes may be responding to photosynthetic activity or carbohydrate metabolism as an indicator of Ci, and not directly to [CO2] changes. A process allied to photosynthetic carbon reduction that is of particular interest is photorespiration (PR). PR in photosynthetic organisms removes the toxic metabolite: 2-phosphoglycolate(2-PG) arising from Rubisco’s oxygenase activity (Fig 1) (Table 2). A series of PR reactions occurring in chloroplast and mitochondria regenerates the CBB intermediate 3-phosphoglycerate(3-PGA) from 2-PG, with the associated loss of 25% fixed carbon as CO2. It must be noted that the Chlamydomonas PR differs from that in higher plants in two aspects: (i) Glycolate is converted to glyoxylate in the mitochondria rather than peroxisomes, (ii) Glyoxylate formation is catalysed by algal glycolate dehydrogenase (GDH), as opposed to glycolate oxidase. PR acts as a sink for energy and reducing power from photosynthetic electron transfer (PET) when CBB activity is limiting (Kozaki and Takeba, 1996; Moroney et al., 2013). PR helps prevent blockades in reduced PET chains, which otherwise would result in formation of reactive oxygen species (ROS) resulting in deleterious effects. ROS resulting from overreduction of PET include \( ^1 \text{O}_2 \) (singlet oxygen), \( \text{H}_2\text{O}_2 \) (hydrogen peroxide), \( \text{O}_2^- \) (superoxide), and \( \cdot \text{OH} \) (hydroxyl radical). While \( ^1 \text{O}_2 \) is generated...
predominantly at the reaction center of PSII and is the main ROS responsible for photo-oxidative damage, the other three are formed at the acceptor side of PSI (Erickson et al., 2015). $^{1}$O$_2$ is formed primarily by energy transfer from triplet state of photosensitizers such as chlorophyll, tetrapyrroles and flavins. Chlamydomonas exhibits acclimation to $^{1}$O$_2$ (Ledford et al., 2007), and this acclimation response is an indicator of $^{1}$O$_2$ signaling mechanism (Erickson et al., 2015). PR is thus intricately connected with photosynthesis and CCM and could help resolve imbalances in PET and CBB (Fig 2) (Caspari et al., 2017), by altering ROS levels. This interconnectedness suggests that metabolites resulting from PR and CBB, or ROS from light absorption by saturated PET chains, could be signals leading to CCM induction.

The impact of PET rates was demonstrated by downregulation of CCM genes CAH1, HLA3, HLA1, HLA2, and HLA4 by 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU) - an inhibitor of photosystem II (Im and Grossman, 2002). DCMU also affects the intrachloroplastic localization of a retrograde-signalling-CCM protein called CAS, which in turn affects expression of 13 CCM genes including HLA3 and LCI1 (as discussed in Section IV). The use of DCMU leads to $^{1}$O$_2$ accumulation (Fufezan et al., 2002). Whether it is the downstream effects of $^{1}$O$_2$, or the reduced photosynthesis that impacts gene expression, requires further systematic investigations using modulators of photosynthesis and intracellular ROS. This study also showed that the above CCM genes require high intensity light, in addition to low [CO$_2$], to upregulate their expression, again hinting towards PET or allied CBB activity as potential influencers of CCM induction.

Like CCM, PR activity is observed due to low [CO$_2$]/[O$_2$] in actively photosynthesizing organisms. Mutations of both PR and CCM genes have reduced ability to survive, unless in a high [CO$_2$] environment (Moroney et al., 2013). This led researchers to suggest a metabolite of PR as the signalling molecule for CCM induction in Chlamydomonas over three decades ago (Spalding et al., 1985). A Chlamydomonas PR mutant lacking the phosphoglycolate phosphatase1(PGP1) that converts 2-PG to glycolate, had affinity for inorganic carbon comparable to wild type (wt) with an induced CCM (Suzuki et al., 1990) leading authors to suggest 2-PG as a CCM signalling molecule.

Further evidence for CCM-PR crosstalk comes from differential processing of glycolate in Chlamydomonas with or without CCM (Moroney et al., 1986). While glycolate is excreted under high CO$_2$ conditions (CCM uninduced and PR downregulated), glycolate excretion is minimal ($1/80^{th}$ of that in high CO$_2$ conditions) in low CO$_2$, CCM induced
conditions. Experiments with labelled CO$_2$, labelled glycolate and inhibitors of PR showed that this is not owing to lowered Rubisco oxygenase activity, but due to high rates of processing of glycolate by PR enzymes. This crosstalk between algal CCM and PR (Fig 1) is further strengthened by co-regulation of several CCM and PR genes by similar environmental stimuli of CO$_2$ and the master regulator CIA5 (Fang et al., 2012). The PR genes are shown in Table 2, with those highlighted co-expressing with CCM genes. Co-regulation of expression of PR and CCM genes is discussed more in section V where we propose a hypothesis for evolution of CCM regulatory mechanisms. In addition to the transfer of metabolites between chloroplast and mitochondria in *Chlamydomonas* during PR, light and CO$_2$ induce changes in mitochondrial arrangement within the cells (Geraghty and Spalding, 1996; Polukhina *et al.*, 2016). The mitochondria lying between chloroplast and cell membrane in CCM-active cells are thought to capture the glycolate exiting chloroplast (Geraghty and Spalding, 1996), and the implications for mitochondrial CAs and inorganic carbon transporters for the CCM are discussed in Section IV. Whether the PR metabolites moving between organelles with light- and CO$_2$- dependent intracellular location impact expression of nuclear CCM genes needs further investigation.

There have only been a few studies to explore the flux between PR, CCM and photosynthesis. Caspari *et al.*, studied pyrenoid-less CCM-defective mutant expressing a higher plant version of Rubisco small subunit (Caspari *et al.*, 2017). This mutant exhibited lower PET rate without affecting the intra-pyrenoidal thylakoid morphologies to compensate for the limited CO$_2$ supply arising from lack of CCM. The rate of PET was restored to levels comparable to wt cells when exposed to high CO$_2$ and the mutant also had increased non-photochemical quenching (NPQ) to dissipate the energy reaching the photosystems.

A large-scale experiment studying metabolite flux and gene expression in photoautotrophic *Chlamydomonas* cells grown under low CO$_2$ (0.04%) and high CO$_2$ (10%) showed that the mitochondrial processes of glycolysis, gluconeogenesis, the glyoxylate pathway, dicarboxylate and metabolism rates of amino acids in PR were responsive to CO$_2$ concentration (Winck *et al.*, 2016). A more systematic analysis of PET, PR, CBB reactions and inorganic carbon uptake under different conditions of [CO$_2$] is needed. Engineering of the biophysical (algal/cyanobacterial) CCM into higher plants might require tinkering with PR processes, in addition to introducing various CCM components (Atkinson *et al.*, 2016; Atkinson *et al.*, 2020), for optimising fluxes between light and dark reactions of photosynthesis.
While CCM is active in the presence of light; and the transcription of several of the CCM genes is light- and \([\text{CO}_2]\)-dependent (Brueggeman et al., 2012; Fang et al., 2012; Tirumani et al., 2014), there have been observations of transcription of \(\text{CAH1, CAH3, CAH6}\) and \(\text{LCIB}\) in low \([\text{CO}_2]\) conditions even in the dark (Mitchell et al., 2014; Rawat and Moroney, 1995; Tirumani et al., 2019; Tirumani et al., 2014). However, the increase of protein levels of these dark-low \([\text{CO}_2]\) transcribed genes, and localization of \(\text{CAH3}\) to the pyrenoid, does not happen until the cells are exposed to light (Mitchell et al., 2014; Tirumani et al., 2014). These observations indicate towards regulation of expression at not just transcription, but also during translation and transport.

In conclusion, much ground remains to be covered for identification of key regulatory elements of the inducible CCM transcription machinery. Efforts are needed to utilize the extensive temporal transcription data from \(\text{Chlamydomonas}\) (Strenkert et al., 2019; Zones et al., 2015) to identify potential CCM TFs and TRs by developing gene regulatory networks (Emmert-Streib et al., 2014). Experiments are needed to understand CIA5 mechanism, validate and characterise potential CCM TFs and TRs identified computationally or via organism-wide experiments discussed in this section. Understanding of CCM regulation requires a general comparison with other physiological mechanisms associated with sensing light and signalling photosynthetic, photorespiratory and phototinhibitory responses. This interconnectedness of PR and CCM led us to consider their evolutionary origins in Section V.

II. Chaperones and the Import and assembly of chloroplastic CCM proteins

Fewer than 100 genes encoding proteins are found in the chloroplast, with the remaining plastidic proteins, including those involved in the CCM, encoded in the nucleus (Martin et al., 2002). The spatial segregation between nuclear gene expression and localization of photosynthetic apparatus in the chloroplast means that many of the photosynthetic and CCM components need to be translated on cytoplasmic ribosomes and transported to across the chloroplast membranes. This implies the need for chaperones for transport and folding of CCM proteins. In this section, we evaluate experimental evidence for chaperones and cellular transport machinery having roles in the assembly and functioning of the CCM.
Folding and transport of CCM proteins

Based on existing data (Brueggeman et al., 2012; Fang et al., 2012; Mackinder et al., 2016; Wang et al., 2015) and proteomics studies, Mackinder et al. (Mackinder et al., 2017) collated 624 nuclear genes involved in the CCM. Analysis of the encoded protein sequences suggests that a significant proportion are significantly disordered (140 seq > 70% disorder, 201 seq > 50% disorder, 291 seq > 30% disorder). These included 76 chloroplast-localized CCM proteins of which 29, 21 and 16 sequences have disorderliness greater than 30%, 50% and 70%, respectively (Mackinder et al., 2017). The occurrence of disorderliness is often an indicator of presence of scaffolding modules for interaction with other proteins. Such disordered regions also destabilise proteins, and very often chaperones are required to prevent their irreversible misfolding (Pechmann and Frydman, 2014). Analysis of the 624 sequences also showed that 142 nuclear and 21 chloroplastic sequences carry at least one transmembrane (TM) domain (Krogh et al., 2001). The insertion of TM regions into membranes requires the assistance of chaperones within the cell (Guna and Hegde, 2018; Jarvis and Kessler, 2014). The disorderliness and the presence of TM domains suggests a requirement for assisted folding in several of the CCM proteins that are translated on cytosolic ribosomes, as the proteins need to be transported across the chloroplast membranes (envelope and/or thylakoid) or embedded within them. Chloroplast TM transport is achieved through channels formed by TOC and TIC (translocon on the outer/inner chloroplast membrane) complexes. This transport is a complex process involving recognition of a transit peptide, protein unfolding and threading through TOC-TIC complexes, cleavage of the transit peptide in the stroma, and finally refolding of the protein. The chaperones Hsp70 (cytosolic and stromal), cytosolic Hsp90, stromal Hsp93 and Cpn60 are involved in chloroplast import of proteins (Flores-Perez and Jarvis, 2013).

Chaperones involved with CCM

The roles of chaperones in organellar transport and folding mean that they could be vital for establishing the CCM, although little work has been carried out in this area. Here, using previously published protein interaction data we have identified the CCM proteins that interact with chaperones HSP90, HSP70 and sHSPs (22E and 22F) (Table 3) (Mackinder et al., 2017; Rutgers et al., 2017). The list indicates that several key CCM proteins including EPYC1, rubisco small subunits, CAs and transporters might need chaperone assistance.

Despite the expectation that chaperones are essential for CCM, the only chaperone which has been suggested to be essential for photosynthesis in a large-scale mutant screen,
CDJ2- a chloroplastic DnaJ protein (Li et al., 2019). Previously, expression of DNJ12 (DnaJ protein) was shown to be dependent on both CO₂ and CIA5 (Fang et al., 2012). Two other DnaJ chaperones- DNJ15 and DNJ31 have CIA5 dependent expression (Fang et al., 2012), and hence feature as CCM proteins in a list compiled recently (Mackinder et al., 2017). The expression of DNJ31 is also dependent on the retrograde signalling mediated by CCM protein CAS (Wang et al., 2016) (Section IV discusses role of CAS), further hinting towards a role in the CCM. DnaJ or Hsp40 proteins are chaperones that work in conjunction with Hsp70, to help fold nascent proteins. Certain DnaJ proteins are known to confer substrate specificity. Whether the algal DnaJ proteins are indeed chaperones and are specifically interacting with CCM proteins needs further characterization. An upregulation of HSF1 in response to low CO₂ conditions in Chlamydomonas has also been shown (Winck et al., 2013). HSF1 is a TF that is known to bind to promoter elements of HSP22F and HSP70A (Strenkert et al., 2011), suggesting that CCM responses include an upregulation of at least two chaperones. Overexpression of chaperones important for CCM expression might be worth considering in a strategy for engineering the CCM.

The limited evidence for chaperones associated with gene expression and protein import essential for the CCM should not negate their importance. Chaperones which are CCM-specific might be few, and there might be several like Hsp70 and Hsp90 members, which cater for varied substrates, including CCM proteins. The varied nature of substrates for several chaperones makes it difficult to identify CCM associated chaperones. The chaperones discussed in this section which have featured in genome wide studies, and in interactomes, are good candidates to further explore this area.

III. Dynamics of starch sheath, pyrenoid matrix and thylakoid tubules

While the response to extracellular environment changes rests primarily with the nuclear genes, the vital process of carbon capture occurs in the pyrenoids. The discovery of the role of EPYC1 as a linker protein (Mackinder et al., 2016; Wunder et al., 2018), the visualization of Rubisco-EPYC1 dynamics during pyrenoid division (Freeman Rosenzweig et al., 2017), proteomics studies (Mackinder et al., 2016; Mackinder et al., 2017; Zhan et al., 2018) and the identification of a motif linking key elements (Meyer et al., 2020) have given the CCM community much needed information about pyrenoidal composition and dynamics. In this section, we focus on specific aspects of the pyrenoid matrix: the extra-pyrenoidal starch sheath and intra-pyrenoidal tubule network, as well as the linkages and interactions that promote their assembly.
Role of the extra-pyrenoidal starch sheath

An extra-pyrenoidal starch sheath in *Chlamydomonas* was first clearly defined in 1957 (Sager and Palade, 1957). While the formation of a starch sheath under low [CO$_2$] suggested that starch prevents CO$_2$ leakage (Ramazanov et al., 1994), there have been contradictory findings showing CCM induction in starchless algal mutants with >10-fold increase in affinity for inorganic carbon (Plumed et al., 1996; Villarejo et al., 1996). However, recent studies in *Chlamydomonas* are indicative of a more fundamental role for the starch sheath in the CCM.

An important CCM protein associated with the starch sheath is LCIB that has low [CO$_2$] and light dependent localization around the pyrenoid in a complex with LCIC. The localization of the LCIB/LCIC complex close to thylakoid tubule emergence and starch plate convergence (Yamano et al., 2014; Yamano et al., 2010), together with a preferential role in uptake of CO$_2$ over HCO$_3^-$ during pH dependent photosynthetic activity measurements (Wang and Spalding, 2014) are consistent with a role for the complex in capturing CO$_2$ retro-diffusing from the pyrenoid. Whether LCIB/LCIC complex acts as a physical barrier or functions as an inducible CA, needs examination. Structures of both LCIB and LCIB/LCIC have attributes of CAs, but no detectable CA activity (Jin et al., 2016). LCIB localization in extra-pyrenoidal starch affects pyrenoid size and number (Yamano et al., 2014), and inorganic carbon affinity (Toyokawa et al., 2020). LCID and LCIE are two more members of the same predicted CA family as LCIB occurring in *Chlamydomonas*, needing characterization (Jin et al., 2016; Wang and Spalding, 2014).

The importance of starch in affecting pyrenoid number and orientation around the tubule network was further supported by studies with *Chlamydomonas saga1* mutants with lowered photosynthetic efficiency, and containing many pyrenoid-like structures, in contrast to the single pyrenoid in wt (Itakura et al., 2019). SAGA1 is believed to link starch and Rubisco. Pyrenoids harbour a thylakoid tubule network that acts as a conduit between thylakoid lumen and pyrenoid. In *saga1* mutants, while the number of pyrenoid-like structures is increased, there was only one tubule network which was often displaced to the periphery of a pyrenoid. Both formation of an entire starch sheath enclosing a single pyrenoid in the canonical position, and a central thylakoid tubule network, appear important for maintaining photosynthetic efficiency of the *saga1* mutants. Recent findings show that SAGA2, 30% similar in sequence to SAGA1, with a starch binding domain localizes to the interface of the pyrenoid and starch sheath (Meyer et al., 2020). Both SAGA1 and SAGA2...
also contain a Rubisco binding motif (RbM), the role of which is discussed later in this section.

**Thylakoid membrane proteins involved in CCM**

The function of the thylakoid tubule network is also thought to be critical for CCM induction. An important protein in the thylakoid lumen is the carbonic anhydrase CAH3, that becomes phosphorylated and localises to the tubule network. Though interactions of CAH3 with thylakoid associated kinases (Depege *et al.*, 2003; Lemeille *et al.*, 2010; Mackinder *et al.*, 2017) have been observed, the exact kinase responsible for its phosphorylation remains unknown. CAH3 is responsible for converting HCO$_3^-$ to CO$_2$ for release into the heart of the pyrenoid (Blanco-Rivero *et al.*, 2012). Such a mechanism will require a HCO$_3^-$ transporter to be located in the thylakoid membrane, and *Chlamydomonas* cells lacking CAH3 grow slower compared to wt cells, particularly in low [CO$_2$], highlighting its role in the upkeep of the CCM (Sinetova *et al.*, 2012). Recently, three genes coding for Bestrophin-like proteins in *Chlamydomonas* called BST1-3 were identified (Mukherjee *et al.*, 2019). These genes are under the regulation of the CCM master TF CIA5 (Fang *et al.*, 2012), and interact with CCM components such as LCIB (Mackinder *et al.*, 2017). BST1-3 localise to the thylakoid membrane and are thought to transport HCO$_3^-$ ions to CAH3, with downregulation of these proteins hampering cell growth at low [CO$_2$] (Mukherjee *et al.*, 2019).

**Linkages involved in the formation of the pyrenoid matrix and starch sheath**

Previous studies have shown that EPYC1, an intrinsically disordered repeat linker protein, is necessary to form a Rubisco-EPYC1 pyrenoidal matrix (Mackinder *et al.*, 2016). Recently, five Rubisco binding regions in EPYC1 were identified using structural data for complexes of Rubisco and peptides representing different regions of EPYC1 (He *et al.*, 2020). EPYC1 acts as a “molecular glue” that tethers multiple Rubisco molecules together to give rise to the pyrenoid matrix.

The presence of a RbM similar to that found in EPYC1 were identified in other pyrenoid localised Rubisco-interacting proteins (Meyer *et al.*, 2020). Identified first in EPYC1, RbM [D/N]W[R/K]XX[L/I/V/A] has been found in a putative chloroplast epimerase CSP41A, SAGA1, SAGA2, and in the thylakoid localising proteins RBMP1 and RBMP2. Disruption of the motif caused these pyrenoid targeted proteins to diffuse homogeneously across the chloroplast, while introduction of the motif into non-pyrenoidal proteins lead to their accumulation in the pyrenoid matrix (Meyer *et al.*, 2020). It must be noted that the
presence of this motif does not always lead to pyrenoidal localization. As discussed previously, the presence of the motif in EPYC1 allows for its interaction with Rubisco to give rise to the EPYC1-rubisco condensate forming the pyrenoid matrix (Atkinson et al., 2020; Wunder et al., 2018). Similarly, the presence of the RbM along with the starch binding regions in photosynthesis is accepted and SAGA2 might help the starch sheath envelope the pyrenoid. Two newly identified thylakoid tubule network localising proteins RBMP1(Cre06.g261750) and RBMP2 (Cre09.g416850) might help tether the pyrenoid to the tubules (Fig 3). Among them, RBMP1 is predicted to be a member of the bestrophin family and may play a role in transporting HCO\textsubscript{3}\textsuperscript{-} to the pyrenoid directly through the tubule network, in contrast to the other bestrophin-like proteins which are found outside the pyrenoid. How these tubule network proteins are localised to the tubule network, which is formed even in cases where there is no pyrenoid matrix formation (Caspari et al., 2017) needs investigation.

**Membrane bending and plasticity of the tubule network for CCM maintenance**

A key area requiring study is the mechanism regulating the formation of the thylakoid tubule network for effective CCM operation. Insights for these processes may arise from studies of cell division in *Chlamydomonas*, when each of the four daughter cells normally contain a nascent pyrenoid as identified from fluorophore-tagged components (Freeman Rosenzweig et al., 2017). Although a small proportion of cells synthesise a pyrenoid de novo, the extent to which thylakoid tubule network is partitioned is not known. The intricate tubule network is formed as thylakoid membranes coalesce near entry points into the pyrenoid, but an added complexity arises from the internal mini-tubules which provide connectivity and allow CBB cycle intermediates to exchange between pyrenoid matrix and chloroplast stroma (Engel et al., 2015). How this complex array of thylakoid membrane remodelling is regulated needs examination.

Membrane remodelling proteins have been discovered in cyanobacteria and chloroplasts of algae and plants. But there is limited understanding of their interactions during thylakoid tubule network formation and potential to mediate CCM development. A set of proteins called CURT1 (Curvature Thylakoid 1) have emerged as important modulators of thylakoid membrane bending and plasticity. First identified in *Arabidopsis thaliana* (CURT1A, B, C, and D), these proteins are conserved across photosynthetic organisms, including three homologs in *Chlamydomonas* (Armbruster et al., 2013). *Arabidopsis* CURT1 proteins concentrate around granal margins and oligomerize to induce membrane tubulation,
with their inhibition negatively affecting photosynthetic efficiency (Armbruster et al., 2013; Pribil et al., 2018). The role of CURT in *Chlamydomonas* is pending inquiry, but it probably contributes to similar membrane dynamics, and potentially plays a role in co-ordinating the assembly of CCM components, and the possible exclusion of PSII from thylakoid tubules within the pyrenoid matrix (McKay and Gibbs, 1991).

Another protein implicated in membrane remodelling is VIPP1, a member of the ESCRT-III family found in eukaryotes (Liu et al., 2020). Disrupting VIPP1 in vascular plants altered thylakoid structure and decreased the volume of thylakoid membranes, suggesting roles in thylakoid membrane upkeep, biogenesis and remodelling (Gutu et al., 2018; Heidrich et al., 2017; Hennig et al., 2015; Hennig et al., 2017; Kroll et al., 2001; Westphal et al., 2001). VIPP1 localizes to *Chlamydomonas* pyrenoid (Zhan et al., 2018). *vipp1* mutant is sensitive to high light and heat stress, and shows altered thylakoid membrane structures close to the pyrenoid, suggesting a role in tubule biogenesis (Nordhues et al., 2012). *Chlamydomonas* also harbours a paralog of VIPP1 called VIPP2, that is not found in many land plants. VIPP1 and VIPP2 are both upregulated under high light or H$_2$O$_2$ induced stress. Both oligomerize to form rod-like structures (Theis et al., 2020). VIPP2 is expressed only in high light conditions in contrast to the constitutive expression of VIPP1, and forms a complex with VIPP1 and HSP22E/F. The lack of upregulation of *HSP22E/F* in a *vipp2* mutant led authors to suggest a role for VIPP2 in conveying chloroplastic stress to the nucleus. This recurring pattern of thylakoid membrane remodeling proteins being upregulated during high light stress in *Chlamydomonas*, is consistent with CCM inductive stimuli (Im and Grossman, 2002). Notably, HSP22E/F interacts with various starch synthesis proteins (Table I), with role of remobilization and starch plate formation being important for pyrenoid assembly and CCM induction (Itakura et al., 2019; Ramazanov et al., 1994). Another candidate involved in shaping thylakoid morphology is FzI, which is a dynamin-like protein involved in grana organization in *Arabidopsis* (Gao et al., 2006). A GTP-binding, oligomerizing homolog of this protein in *Chlamydomonas* called crFZL was recently found important for coping with high light stress (Findinier et al., 2019).

While the recent identifications of EPYC1 as a linker protein and RbM (Meyer et al., 2020) are major developments, many intriguing questions related to thylakoid organization during CCM induction, whether during transfer from high to low CO$_2$, or in synchronized cells during cell division and development, remain unanswered. Structural and mechanistic insights about the thylakoid membrane organisational proteins discussed in this section, in
conjunction with life-cycle and environmental regulation of the pyrenoid starch and tubule network, are needed to further our understanding of the CCM.

IV. Retrograde signalling in CCM regulation

In previous sections, we have navigated from the nucleus to the pyrenoid examining different algal mechanisms to establish and regulate CCM. The chloroplast effectively operating as the hub orchestrating photosynthetic reactions, must be sensitive to environmental factors that affect photosynthesis and/or CCM. In this section, we see how changes in the chloroplast are communicated to the nucleus via signalling molecules (Rea et al., 2018) for CCM regulation. In *Chlamydomonas*, molecular transducers such as tetrapyrrole intermediates, ROS, as well as Ca\(^{2+}\) ions help relay signals to the nucleus, with potential changes to the nuclear transcriptome. This section describes how changes within the chloroplast affect nuclear gene expression with implications for photosynthesis and CCM.

Redox status signalling in photosynthesis

Photosynthesis is a redox-centred metabolic process, subject to fluctuations in environmental conditions (Dietz et al., 2016). Maintaining a functional PET requires balanced excitation of the two photosystems (Rea et al., 2018) without which over-reduced PET components might generate harmful ROS. As discussed in Section I, PET rates need to be tuned for CCM, CBB and PR (Caspari et al., 2017). The influence of changes in PET affecting CCM gene expression (Im and Grossman, 2002), discussed in Section I, suggests the redox status of PET components acts as a regulator of CCM gene expression. It is therefore essential to have inter and intra-organellar redox status communication (Pfannschmidt et al., 2020) so that both CCM and photosynthesis rates are optimal. Here, we explore the *gun4* mutant, which demonstrates a link between ROS generation and retrograde signalling.

Retrograde signalling by GUN4

The biosynthesis of tetrapyrroles for chlorophyll generation is a process that is sensitive to oxidative stress and needs to be tightly regulated. Tetrapyrroles, as well as many of their biosynthetic intermediates interact with oxygen in their triplet state to generate ROS, mainly singlet oxygen (Rea et al., 2018; Tanaka and Tanaka, 2007). Tetrapyrrole metabolism occurs in the chloroplast making use of nuclear-encoded enzymes, necessitating communication between chloroplast and nucleus. The role of tetrapyrrole intermediates in ‘retrograde signalling’ to modulate expression of certain nuclear genes has been studied using
‘GUN mutants’ in plants and algae (Larkin, 2016). In wt cells, disrupted chlorophyll biosynthesis is known to activate a specific retrograde signal, which results in downregulation of photosynthetic nuclear genes, such as Lhcb2. The gun mutant alleles are defective in this particular signal, instead elevating levels of these usually suppressed gene transcripts (Larkin, 2016). One such mutant- GUN4 has been identified in *Chlamydomonas* with half as much chlorophyll as wt cells (Formighieri et al., 2012). While GUN4 is not essential for chlorophyll synthesis, gun4 mutants in *Arabidopsis* and *Chlamydomonas* exhibit impaired chlorophyll accumulation, suggesting a role for GUN4 in regulation of the enzyme Mg$^{2+}$ chelatase (MgCh) (Fig 4). RNA-Seq analysis has shown that the expression of 803 nuclear-encoded genes in *Chlamydomonas* is disrupted in the gun4 mutants as compared to the wt (Formighieri et al., 2012).

While the role of retrograde signalling in regulating photosynthesis is accepted, we checked if it also plays a part in modulating the CCM. We compared the 803 differentially expressed genes in the gun4 mutant (Formighieri et al., 2012) and those identified as being important for CCM (Mackinder et al., 2017; Strenkert et al., 2019). Our examination of these datasets shows that expression of nuclear genes encoding 52 CCM genes are affected >8-fold by GUN4 and, by extension, retrograde signalling (Table 4). This list of 52 genes includes a redox proteins (peroxiredoxin), a PR pathway proteins (serine hydroxymethyl transferase2), three photosystem components (PSAK, PSAH, PSIIPbs27), bestrophin-3, three CAs (mitochondrial-CAH4, CAH5, and periplasmic-CAH8), and the chaperone DNJ15. That PR, mitochondrial CAs and PET proteins are influenced by GUN4 further strengthens the cross-talk between these processes which has previously been highlighted in Section I. GUN4-influenced CCM genes also include DNJ15, implying protein homeostasis regulation by retrograde signalling. Seven of the CCM genes in Table 4 are uncharacterised. These proteins of unknown function which are influenced by GUN4 make a case for further exploring of the CCM and retrograde signalling connections.

The mechanistic action of GUN4 in relaying a signal to the nucleus remains undetermined. GUN4 orthologues are present only in species that carry out oxygenic photosynthesis (Formighieri et al., 2012), suggesting a role in photo-oxidative acclimation strategies. Based on the structure of GUN4 in unliganded form (Verdecia et al., 2005), it was hypothesized that GUN4 protects protochlorophyllide intermediates from O$_2$, leading to degradation products of GUN4 going on to act as retrograde signals to the nucleus. A contrasting hypothesis arising from a GUN4-PPIX structure solved by Chen et al., suggested that O$_2$ could access bound tetrapyrroles (Chen et al., 2015) leading to singlet oxygen (O$_2^{{\cdot}}$).
(Tabrizi et al., 2016), which acts as the retrograde signal. The proposed models for the GUN4 mode of action are summarized in Fig 4. Such signalling responses in Arabidopsis may be mediated through the chloroplast-localised EXECUTER 1 and 2 systems (Singh et al., 2015). It is proposed that a similar sensing system is required in Chlamydomonas. However, no corresponding homolog has been found, with the most similarity shared with EXECUTER 1 and 2 was 10.2% and 11.2% by the Chlamydomonas protein Cre03.g163500. Mechanistic regulation of CCM gene expression by retrograde signalling in Chlamydomonas, as discussed in the Section 1 for nuclear regulators, needs to be explored in detail with systematic analysis for signalling molecules, TFs and cis-elements.

Another CCM protein that has been characterized in the last five years is CAS. It is the only CCM protein that has been studied in a systematic manner for its potential role in retrograde signalling independent of GUN 4 and is described in the following paragraphs.

**Retrograde signalling by CAS**

*Chlamydomonas* Ca$^{2+}$- binding protein (CAS) is a chloroplastic thylakoid membrane protein that mediates signalling, as part of acclimation to high light and low carbon (LC) conditions (Wang et al., 2016). CAS was initially studied in Arabidopsis, where it acts as a Ca$^{2+}$-binding protein, regulating stomatal closure (Nomura and Shiina, 2014). Although no catalytic activity has been demonstrated for CAS, *A. thaliana* and, Chlamydomonas CAS have Ca$^{2+}$ binding ability in their N-terminus, and a rhodanese domain of unknown function (Wang, 2017). A comparison of the transcriptomes of a cas mutant strain with wt and complemented strains demonstrated that absence of CAS leads to greater than 4-fold decrease in transcript levels of 13 genes (Wang, 2017; Wang et al., 2016) (Table 5). All 13 genes except one encoding for a predicted phosphatase have previously been identified as CCM genes. Of particular interest was reduced gene transcription and accumulation of transporters HLA3 and LCIA for the uptake of inorganic carbon into the cell (Yamano et al., 2015) in the cas mutant. Three of the genes (LC1 D, Cre12.g541550, Cre26.g756747) in Table 5 have been identified as CCM genes based on previous large-scale expression studies, but their cellular functions are unknown. The expression of CIA5-dependent DNJ31, encoding a DnaJ chaperone discussed in Section II, is also influenced by CAS. These expression features of DNJ31 make it an interesting candidate requiring its functional characterization and identification of its substrates. Two mitochondrial carbonic anhydrases (CAH4, CAH5- also regulated by GUN4) and two mitochondrial envelope proteins (CCP1, CCP2) of unknown function, (Atkinson et al., 2016; Mackinder et al., 2017) are among those encoded by the 13
genes with CAS-dependent expression. The chloroplastic CCM protein CAS influencing the expression of 4 mitochondrial CCM proteins represents another feature of the mitochondria-chloroplast crosstalk needed to establish the CCM. Considering the importance of PR-CCM connections as described in Section I, it is probable that CAH4 and CAH5 help recapture photorespired and respired CO₂ in mitochondria as bicarbonate ions prior to export by unidentified transporters. CAS also influences PET by affecting expression of genes encoding two proteins involved in NPQ- LHCSR2 and LHCSR3. The over-reduction of the PET chain caused by high light intensity could be a potential trigger for CAS activation, and its modulation of NPQ (Caspari et al., 2017). These observations further reiterate the need for tuning the rates of PET, PR, CBB and inorganic carbon uptake with CCM induction, as discussed in Section I.

Similar to Rubisco, CAS was also revealed to move into the pyrenoid upon transition from high to low [CO₂], with light being a prerequisite for this relocalisation to occur (Yamano et al., 2018). The importance of Ca²⁺ binding is demonstrated by use of the chelator BAPTA that lessens CAS-mediated accumulation of LCIA and HLA3 (Yamano et al., 2018). The Ca²⁺-rich environment in the pyrenoid leads to Ca²⁺-CAS binding, which is thought to trigger a conformational change and mediate a signal to the nucleus, regulating CCM gene expression. The precise manner in which this protein signals for LC acclimation in *Chlamydomonas* is unknown, but a broad model is depicted in Fig 5. The PET inhibitor DCMU prevents CAS relocalisation under low [CO₂] (Wang, 2017). Upon activation, CAS appears to relocate to a focal region within the pyrenoid. Though unclear, the signal propagation is affected by intracellular Ca²⁺ levels, [CO₂] and light intensity. While the CAS mode of action is still unknown, these results suggest that the protein mediates a Ca²⁺-dependent retrograde signal to the nucleus, as part of the Chlamydomonas high light/LC acclimation. High light acclimation and CCM induction appear as important factors in the case of CIA5 mediated CCM gene expressions (Section I) and thylakoid membrane structuring proteins (Section III), and also in retrograde signalling *Chlamydomonas*. CAS mediated retrograde signalling further strengthens the ideas presented in Section I, integrating the various physical and biological factors regulating CCM.

### V. Insights for evolution of CCM regulation

The cues regulating CCM genes, which have been frequently re-iterated throughout this review, are not low ambient [CO₂] alone, but also light intensity. The response to CO₂ and light intensity changes, as discussed in Sections I and IV could be indirectly mediated by
metabolites or ROS that result from photosynthesis and/or PR. The physiological connections between PR and CCM along with the uncanny similarity in gene expression regulation of several CCM and PR genes, led us to hypothesize on the evolutionary origins of CCM regulatory mechanisms.

The regulation of PR in *Chlamydomonas* bears several similarities to that of CCM. The genome wide expression study with *cia5* showed that several key PR enzymes (Fig 1, Table 2) (alanine aminotransferase1 (AAT1), glycerate kinase (GLYK), glycolate dehydrogenase (GDH), hydroxypyruvate reductase1 (HPR1), serine glyoxylate aminotransferase1 (SGAT1), glycine decarboxylase (GDC) complex enzymes) were dependent on the master-regulator CIA5 and CO₂ in a manner similar to several CCM genes (Fang *et al.*, 2012). The expression profiles of the PR and CCM genes that clustered in ‘CCM clusters’ (Fang *et al.*, 2012) are shown in Fig 6. It is worth noting that not all PR genes display this expression pattern. Notable exceptions are *PGP* genes, serine hydroxymethyltransferase (*SHMT*) genes and *GDC-H* gene. GDC-H is in a GDC complex with two other subunits (GDC-T and GDC-P) that are encoded by genes part of ‘CCM clusters’.*PGP1* was discussed in Section I. Although the PGP genes- *PGP1*, *PGP2*, and *PGP3* did not show CIA5-dependent differential expression (Fang *et al.*, 2012), the expression of these PR genes was upregulated under low [CO₂] (Tural and Moroney, 2005). [CO₂]-independent expression of *PGP* genes, in contrast to most genes that encode enzymes downstream of PGP in PR, lends more weight to PGP substrate (2-PG) or product (glycolate) as CCM induction signalling molecules. 2-PG was suggested as a potential signalling molecule in a *pgp1* algal mutant study (Section I, (Suzuki *et al.*, 1990)). *SHMT* genes which also have expression profiles different from that of other PR genes, encode proteins that catalyze conversion of glycine to serine. If mitochondrial glycine:serine ratio impacts CCM induction requires investigation.

A recent study (Tirumani *et al.*, 2019) showed that five of the PR enzymes- AAT1, GDH, HPR1, PGP1 and GDC-H were not only dependent on low [CO₂] for enhanced expression, but also were affected by light intensity. These genes also displayed diel regulation of expression without any circadian rhythmicity. These expression characteristics displayed by the 5 PR genes were also seen for 4 CCM genes (*CAH3, LCIB, LCI1, CCP1*) in the same study. This similarity in expression profiles dependent on [CO₂] changes and the presence of a common TF led us to look at the evolutionary origins of these processes. These studies provide compelling evidence that the stimulus for the CCM is likely determined by interactions between light intensity and PR activity when external CO₂ is limiting.
The origin of oxygenic photosynthesis in Cyanobacteria is placed around 2.4 billion years ago (bya) (Kopp et al., 2005). The endosymbiotic event conferring eukaryotic cells with a photosynthetic chloroplast through engulfment of a cyanobacteria is believed to have occurred 1.8 bya, and the origin of Chlorophytes (the branch carrying green alga such as Chlamydomonas) from a common ancestor for all green plants and algae is dated around 700 million years ago (mya) (Becker, 2013). The endosymbiotic events of internalising cyanobacteria and proteobacteria are believed to have not only conveyed the photosynthetic machinery via chloroplast (cyanobacterial endosymbiosis) establishment, but also the PR enzymes from both cyano- and proteobacteria (mitochondrial evolution) (Bauwe et al., 2012; Eisenhut et al., 2008). PR had to develop in response to the oxygenase activity of Rubisco. PR is seen in all photosynthetic organisms— with and without biophysical (cyanobacteria, algae, hornworts) or biochemical CCMs (C4 and CAM plants) (Hagemann et al., 2016). It is an ancillary pathway to photosynthesis that evolved over 1.8 bya, before the evolution of the algal CCM, which was perhaps as early as 500 mya, and C4 and CAM in land plants evolved less than 100 mya. The driver for evolution of PR, biophysical CCM, C4 and CAM was the changing atmospheric composition with increasing \([\text{O}_2]:[\text{CO}_2]\) ratios (Griffiths et al., 2017). This similarity in evolutionary drivers and co-regulation by CIA5 of both PR and CCM led us to hypothesize that the cellular machinery adopted a pre-existing PR regulatory tool to establish the CCM.

The observation that Rubisco occurring in organisms with CCM/C4/CAM have greater affinity for O\(_2\) (i.e., lower specificity factor) than those lacking concentration mechanisms (Griffiths et al., 2017) hints at PR being an essential physiologically linked process for concentration mechanisms to function. Whether there are common TFs between PR and C4/CAM in higher plants, such as Chlamydomonas’ CIA5 requires further analysis. While the existence of common TFs like CIA5 as a co-regulatory tool for concentration mechanisms and PR may not exist due to homoplastic origins of eukaryotic CCM, C4 and CAM pathways (Sage et al., 2011; Sage, 2016; Raven et al., 2017; Edwards, 2019), there is a possibility of evolution of other regulatory modes for co-ordinating the processes. However, the identification of common TFs regulating PR responses, as well as co-ordinating expression of the algal CCM and C4/CAM pathways in higher plants, is a promising line of investigation (I. Santhanagopalan, P. Singh, M. Chomthong, H. Griffiths and G. Yadav, unpublished data).
Conclusions

We began this review by considering the role of the TF CIA5, and were faced with the confounding observation that this key regulatory molecule not only activates some CCM genes, it is also involved in activating PR. Working from the hypothesis that [CO₂] changes are sensed through indirect cues from photosynthetic activity or carbohydrate metabolism, we propose that signals associated with imbalance of PET and CBB rates in response to [CO₂] and light intensity changes, such as photorespiratory metabolites, or associated redox signalling, mediated by specific chloroplastic retrograde signalling mediators such as CAS or GUN4, may be effectors in CCM induction. Whilst CCM is a response to CO₂ limitation, the associated signalling has been co-opted secondarily from the requirement to regulate genes processing photorespiratory intermediates, or cope with associated light stress when the PET is over-energised and NPQ is upregulated (Caspari et al., 2017). From an evolutionary perspective, this signalling would be consistent with the theory that CCM evolved in Chlorophyceae at the point of dissolved [CO₂]:[O₂] ratios being equivalent, some 400-500 mya (Griffiths et al., 2017), probably building on the evolution of PR machinery that had evolved 1.8 bya (Becker, 2013). The close co-operation between mitochondria and chloroplast metabolite exchanges, PR activity and inducible CCM components strengthens this notion. We also propose that the interconnections between PET, CBB, CCM and PR might be further strengthened by recapturing of photorespired CO₂ by CCM machinery and suggest that optimisation of an engineered algal CCM in higher plants requires consideration of the PR apparatus.

Whilst building on the tremendous progress that has been made in recent years in identifying novel CCM components and assembly of a functional pyrenoid (Meyer et al., 2012; Mackinder et al., 2016; Mackinder et al., 2017; Atkinson et al., 2020; Meyer et al., 2020), future studies should seek to identify the key effectors for CCM activation. The promise of identifying additional TFs (Arias et al., 2020) and potential cis-activational motifs (Winck et al., 2013), and post-transcriptional regulations through sRNA systems (Muller et al., 2020) will underpin efforts to understand CCM regulation. The analyses of large transcriptional data sets (Zones et al., 2015; Strenkert et al., 2019) using gene regulatory networks (Emmert-Streib et al., 2014), in conjunction with cis-element analysis (Winck et al., 2013) can help identify key TF and TR for future focus. Equally, by revisiting earlier studies which associated photorespiratory activity and CCM induction (Spalding et al., 1985; Moroney et al., 1986; Im and Grossman, 2002), we have highlighted additional regulatory processes which perhaps lead to CCM induction. Subsequent investigations with relevant
knockout mutants from the large repositories generated (Li et al., 2016; Li et al., 2019; Vilarrasa-Blasi et al., 2020) will help test the above hypotheses and obtain further insights about regulation.

Major questions still remain, such as the origins and regulatory processes leading to the knotted thylakoid tubule network at the heart of the pyrenoid, the formation of associated connective minitubules, and spatial segregation between PSI and PSII. The energetic balance between CO₂ reduction in the pyrenoid matrix, associated metabolite exchange and regulation of the CBB operating in the stroma, may also account for the regulatory signalling complexities outlined above. The answers to questions about the complex regulatory processes leading to CCM induction will be revealed by molecular and physiological analyses and will require critical decisions on appropriate growth conditions (light intensity, photoperiods, synchronisation of cells, [CO₂]) to identify the signalling mechanisms which orchestrate CCM.

**Acknowledgements:** The authors thank Dr Gitanjali Yadav for helpful discussions, and Dr Jessica Royles for meticulous proofreading of this document. We are grateful for financial support from the GCRF Collective Call: BBSRC BB/P027970/1 (TIGR²ESS) programme.
Arias C, Obudulub O, Zhaoa X, Ansolia P, Zhang X, Paula S, Bygdell J, Pirmoradiane M, Zubareve R, Samuelsson G, Wingsle G, Bajhaiya AK. 2020. Nuclear proteome analysis of Chlamydomonas with response to CO2 limitation. Algal Research 46.

Armbruster U, Labs M, Pribil M, Viola S, Xu W, Scharfenberg M, Hertle AP, Rojahn U, Jensen PE, Rappaport F, Joliot P, Dörmann P, Wanner G, Leister D. 2013. Arabidopsis CURVATURE THYLAKOID1 proteins modify thylakoid architecture by inducing membrane curvature. Plant Cell 25, 2661-2678.

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, Mao Y, Chan KK, McCormick AJ. 2020. Condensation of Rubisco into a protop鄱ynoid in higher plant chloroplasts. Nature Communications 11, 6303.

Badger MR, Kaplan A, Berry JA. 1980. Internal Inorganic Carbon Pool of Chlamydomonas reinhardtii. EVIDENCE FOR A CARBON DIOXIDE-CONCENTRATING MECHANISM. Plant Physiology 66, 407-413.

Bauwe H, Hagemann M, Kern R, Timm S. 2012. Photosynthesis has a dual origin and manifold links to central metabolism. Current Opinion in Plant Biology 15, 269-275.

Becker B. 2013. Snow ball earth and the split of Streptophyta and Chlorophyta. Trends in Plant Science 18, 180-183.

Blanco-Rivero A, Shutova T, Roman MJ, Villarejo A, Martinez F. 2012. Phosphorylation controls the localization and activation of the luminal carbonic anhydrase in Chlamydomonas reinhardtii. PLoS One 7, e49063.

Brueggeman AJ, Gangadharaiyah DS, Cserhati MF, Casero D, Weeks DP, Ladunga I. 2012. Activation of the carbon concentrating mechanism by CO2 deprivation coincides with massive transcriptional restructuring in Chlamydomonas reinhardtii. Plant Cell 24, 1860-1875.

Caspari OD, Meyer MT, Tolleret D, Wittkopp TM, Cunniffe NJ, Lawson T, Grossman AR, Griffiths H. 2017. Pyrenoid loss in Chlamydomonas reinhardtii causes limitations in CO2 supply, but not thylakoid operating efficiency. Journal of Experimental Botany 68, 3903-3913.

Chen B. 2016. The function and regulation of CIA5/CCM1 in Chlamydomonas reinhardtii. Ph.D thesis, Iowa State University.

Chen X, Pu H, Fang Y, Wang X, Zhao S, Lin Y, Zhang M, Dai HE, Gong W, Liu L. 2015. Crystal structure of the catalytic subunit of magnesium chelatase. Nature Plants 1, 15125.

Chung BY, Valli A, Deery MJ, Navarro FJ, Brown K, Hnatova S, Howard J, Molnar A, Baulcombe DC. 2019. Distinct roles of Argonaute in the green alga Chlamydomonas reveal evolutionary conserved mode of miRNA-mediated gene expression. Scientific Reports 9, 11091.

Depege N, Bellafiore S, Rochaix JD. 2003. Role of chloroplast protein kinase Stt7 in LHCII phosphorylation and state transition in Chlamydomonas. Science 299, 1572-1575.

Dietz KJ, Turkcan I, Krieger-Liszkay A. 2016. Redox- and Reactive Oxygen Species-Dependent Signaling into and out of the Photosynthesizing Chloroplast. Plant Physiology 171, 1541-1550.

Eisenhut M, Ruth W, Haímovich M, Bauwe H, Kaplan A, Hagemann M. 2008. The photorespiratory glycolate metabolism is essential for cyanobacteria and might have been conveyed endosymbiontically to plants. Proceedings of the National Academy of Sciences, USA 105, 17199-17204.

Emmert-Streib F, Dehmer M, Haibe-Kains B. 2014. Gene regulatory networks and their applications: understanding biological and medical problems in terms of networks. Frontiers in Cell and Developmental Biology 2, 38.

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, doi:10.7554/elife.04889.

Erickson E, Wakao S, Niyogi KK. 2015. Light stress and photoprotection in Chlamydomonas reinhardtii. Plant Journal 82, 449-465.

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 CO2-concentrating mechanism regulator CIA5/CCM1. Plant Cell 24, 1876-1893.

Findier J, Delevoye C, Cohen MM. 2019. The dynamin-like protein Fzl promotes thylakoid fusion and resistance to light stress in Chlamydomonas reinhardtii. PLoS Genetics 15, e1008047.
Proceedings of the National Academy of Sciences, USA

Structural insights into the LCIB protein family reveals a new group of beta-Jin S, Sun J, Wunder T, Tang D, Cousins AB, Sze SK, Mueller FA. (eds) Plastid Biology. Advances in Plant Biology, vol 113. Springer, New York, NY., 241-270.

Fuentes C, Rutherford AW, Krieger-Liszkay A. 2002. Singlet oxygen production in herbicide-treated photosystem II. FEBS Letters 532, 407-410.

Fukuzawa H, Miura K, Ishizaki K, Kucho KI, Saito T, Kohinata T, Ohyama K. 2001. Ccm1, a regulatory gene controlling the induction of a carbon-concentrating mechanism in Chlamydomonas reinhardtii by sensing CO2 availability. Proceedings of the National Academy of Sciences, USA 98, 5347-5352.

Gao H, Sage TL, Osteryoung KW. 2006. FZL, an FZO-like protein in plants, is a determinant of thylakoid and chloroplast morphology. Proceedings of the National Academy of Sciences, USA 103, 6759-6764.

Geraghty AM, Spalding MH. 1996. Molecular and Structural Changes in Chlamydomonas under Limiting CO2 (A Possible Mitochondral Role in Adaptation). Plant Physiology 111, 1339-1347.

Goudet MMM, Orr DJ, Melkonian M, Muller KH, Meyer MT, Carmo-Silva F, Griffiths H. 2020. Rubisco and carbon-concentrating mechanism co-evolution across chlorophyte and streptophyte green algae. New Phytologist 227, 810-823.

Griffiths H, Meyer MT, Rickaby REM. 2017. Overcoming adversity through diversity: aquatic carbon concentrating mechanisms. Journal of Experimental Botany 68, 3689-3695.

Guna A, Hegde RS. 2018. Transmembrane Domain Recognition during Membrane Protein Biogenesis and Quality Control. Current Biology 28, R498-R511.

Gutu A, Chang F, O'Shea EK. 2018. Dynamical localization of a thylakoid membrane binding protein is required for acquisition of photosynthetic competency. Molecular Microbiology 108, 16-31.

Hagemann M, Kern R, Maurino VG, Hanson DT, Weber AP, Sage RF, Bauwe H. 2016. Evolution of photorespiration from cyanobacteria to land plants, considering protein phylogenies and acquisition of carbon concentrating mechanisms. Journal of Experimental Botany 67, 2963-2976.

He S, Chou HT, Matthies D, Wunder T, Meyer MT, Atkinson N, Martinez-Sanchez A, Jeffrey PD, Port SA, Patena W, He G, Chen VK, Hughson FM, McCormick AJ, Mueller-Cajar O, Engel BD, Yu Z, Jonikas MC. 2020. The structural basis of Rubisco phase separation in the pyrenoid. Nature Plants 6, 1480-1490.

Heidrich J, Thurotte A, Schneider D. 2017. Specific interaction of IM30/Vipp1 with cyanobacterial and chloroplast membranes results in membrane remodeling and eventually in membrane fusion. Biochimica et Biophysica Acta - Bioenergetics 1858, 126-136.

Im CS, Grossman AR. 2002. Identification and regulation of high light-induced genes in Chlamydomonas reinhardtii. Plant Journal 30, 301-313.

Itakura AK, Chan KK, Atkinson N, Pallesen L, Wang L, Reeves G, Patena W, Caspari O, Roth R, Goodenough U, McCormick AJ, Griffiths H, Jonikas MC. 2019. A Rubisco-binding protein is required for normal pyrenoid number and starch sheath morphology in Chlamydomonas reinhardtii. Proceedings of the National Academy of Sciences, USA 116, 18445-18454.

Jarvis P, Kessler F. 2014. Mechanisms of chloroplast protein import in plants. In: Theg S., Wollman FA. (eds) Plastid Biology. Advances in Plant Biology, vol 5. Springer, New York, NY., 241-270.

Jin S, Sun J, Wunder T, Tang D, Cousins AB, Sze SK, Mueller-Cajar O, Gao YG. 2016. Structural insights into the LCIB protein family reveals a new group of beta-carbonic anhydrases. Proceedings of the National Academy of Sciences, USA 113, 14716-14721.
Kohinata T, Nishino H, Fukuzawa H. 2008. Significance of zinc in a regulatory protein, CCM1, which regulates the carbon-concentrating mechanism in Chlamydomonas reinhardtii. Plant and Cell Physiology 49, 273-283.

Kopp RE, Kirschvink JL, Hilburn IA, Nash CZ. 2005. The Paleoproterozoic snowball Earth: a climate disaster triggered by the evolution of oxygenic photosynthesis. Proceedings of the National Academy of Sciences, USA 102, 11131-11136.

Kozaki A, Takeba G. 1996. Photorespiration protects C3 plants from photooxidation. Nature 384, 557-560.

Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. Journal of Molecular Biology 305, 567-580.

Kroll D, Meierhoff K, Bechtold N, Kinoshita M, Westphal S, Voithknecht UC, Soll J, Westhoff P. 2001. VIPP1, a nuclear gene of Arabidopsis thaliana essential for thylakoid membrane formation. Proceedings of the National Academy of Sciences, USA 98, 4238-4242.

Kucho K, Ohyama K, Fukuzawa H. 1999. CO2-responsive transcriptional regulation of CAH1 encoding carboxic anhydrase is mediated by enhancer and silencer regions in Chlamydomonas reinhardtii. Plant Physiology 121, 1329-1338.

Kucho K, Yoshioka S, Taniguchi F, Ohyama K, Fukuzawa H. 2003. Cis-acting elements and DNA-binding proteins involved in CO2-responsive transcriptional activation of CAH1 encoding a periplasmic carboxic anhydrase in Chlamydomonas reinhardtii. Plant Physiology 133, 783-793.

Larkin RM. 2016. Tetrapyrrole Signaling in Plants. Frontiers in Plant Science 7, 1586.

Ledford HK, Chin BL, Niyogi KK. 2007. Acclimation to singlet oxygen stress in Chlamydomonas reinhardtii. Eukaryotic Cell 6, 919-930.

Lemeille S, Turkina MV, Vener AV, Rochaix JD. 2010. Stt7-dependent phosphorylation during state transitions in the green alga Chlamydomonas reinhardtii. Molecular and Cellular Proteomics 9, 1281-1295.

Li X, Patena W, Fauser F, Jinkerson RE, Saroussi S, Meyer MT, Ivanova N, Robertson JM, Yue R, Zhang R, Vilarasssa-Blasi J, Wittkopp TM, Ramundo S, Blum SR, Goh A, Laudon M, Srikumar T, Lefebvre PA, Grossman AR, Jonikas MC. 2019. A genome-wide algal mutant library and functional screen identifies genes required for eukaryotic photosynthesis. Nature Genetics 51, 627-635.

Li X, Zhang R, Patena W, Gang SS, Blum SR, Ivanova N, Yue R, Robertson JM, Lefebvre PA, Fitz-Gibbon ST, Grossman AR, Jonikas MC. 2016. An Indexed, Mapped Mutant Library Enables Reverse Genetics Studies of Biological Processes in Chlamydomonas reinhardtii. Plant Cell 28, 367-387.

Liu J, Tassinari M, Souza DP, Naskar S, Noel JK, Bohuszewicz O, Buck M, Williams TA, Baum B, Low HH. 2020. Bacterial Vipp1 and PspA are members of the ancient ESCRT-III membrane-remodelling superfamily. BioRxiv, doi.org/10.1101/2020.1108.1113.249979 [Preprint].

Mackinder LC, Meyer MT, Mettler-Altmann T, Chen VK, Mitchell MC, Caspari O, Freeman Rosenzweig ES, Pallesen L, Reeves G, Itakura A, Roth R, Sommer F, Geimer S, Muhlhaus T, Schroda M, Goodenough U, Stitt M, Griffiths H, Jonikas MC. 2016. A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle. Proceedings of the National Academy of Sciences, USA 113, 5958-5963.

Mackinder LCM, Chen C, Leib RD, Patena W, Blum SR, Rodman M, Ramundo S, Adams CM, Jonikas MC. 2017. A Spatial Interactome Reveals the Protein Organization of the Algal CO2-Concentrating Mechanism. Cell 171, 133-147.e114.

Martin W, Rajan T, Richly E, Hansen A, Cornelisen S, Lins T, Leister D, Stoeb B, Hasegawa M, Penny D. 2002. Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proceedings of the National Academy of Sciences, USA 99, 12246-12251.

McKay RML, Gibbs SP. 1991. Composition and function of pyrenoids: cytochemical and immunocytochemical approaches. Canadian Journal of Botany 69, 1040-1052.

Merchant SS, Prochnik SE, Vallon O, Harris EH, Karpowicz SJ, Witman GB, Terry A, Salamov A, Fritz-Laylin LK, Marechal-Drouard L, Marshall WF, Qu LH, Nelson DR, Sanderfoot AA, Spalding MH, Kapitonov VV, Ren Q, Ferris P, Lindquist E, Shapiro H, Lucas SM, Grimwood J, Schmutz J, Cardol P, Cerutti H, Chanfreau G, Chen CL, Cognat V, Croft MT, Dent R, Dutcher S, Fernandez E, Fukuzawa H, Gonzalez-Ballester D, Gonzalez-Halphen
D, Hallmann A, Hanikenne M, Hippler M, Inwood W, Jabbari K, Kalanon M, Kuras R, Lefebvre PA, Lemaire SD, Löbanov AV, Lohr M, Manuell A, Meier I, Mets L, Mittag M, Mittelmeier T, Moroney JV, Moseley J, Napoli C, Nedelcu AM, Niyogi K, Novoselov SV, Paulsen IT, Pazour G, Purton S, Rai JP, Riano-Pachon DM, Riekhof W, Rymarquis L, Schroda M, Stern D, Umen J, Willows R, Wilson N, Zimmer SL, Allmer J, Balk J, Bisoka V, Chen CJ, Elias M, Gendler K, Hauser C, Lamb MR, Ledford H, Long JC, Minagawa J, Page MD, Pan J, Pootakham W, Roje S, Rose A, Stahlberg E, Terauchi AM, Yang P, Ball S, Bowler C, Dieckmann CL, Gladyshev VN, Green P, Jorgensen R, Mayfield S, Mueller-Roeber B, Rajamani S, Sayre RT, Broekstein P, Dubchak I, Goodstein D, Hornick L, Huang YW, Jhaveri J, Luo Y, Martínez D, Ngu WC, Otillar B, Poliakov A, Porter A, Szajkowski L, Werner G, Zhou K, Grigoriev IV, Rokhsar DS, Grossman AR. 2007. The Chlamydomonas genome reveals the evolution of key animal and plant functions. Science 318, 245-250.

Meyer M, Griffiths H. 2013. Origins and diversity of eukaryotic CO2-concentrating mechanisms: lessons for the future. Journal of Experimental Botany 64, 769-786.

Meyer MT, Itakura AK, Patena W, Wang L, He S, Emrich-Mills T, Lau CS, Yates G, Mackinder LCM, Jonikas MC. 2020. Assembly of the algal CO2-fixing organelle, the pyrenoid, is guided by a Rubisco-binding motif. Science Advances 6, eaad2408.

Meyer MT, Whittaker C, Griffiths H. 2017. The algal pyrenoid: key unanswered questions. Journal of Experimental Botany 68, 3739-3749.

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, Kohinata T, Inoue Y, Taniguchi F, Asamizu E, Nakamura Y, Tabata S, Yamato KT, Ohyama K, Fukuzawa H. 2004. Expression profiling-based identification of CO2-responsive genes regulated by CCM1 controlling a carbon-concentrating mechanism in Chlamydomonas reinhardtii. Plant Physiology 135, 1595-1607.

Molnar A, Schwach F, Studholme DJ, Thuenemann EC, Baulcombe DC. 2007. miRNAs control gene expression in the single-cell alga Chlamydomonas reinhardtii. Nature 447, 1126-1129.

Moroney JV, Husic HD, Tolbert NE, Kitayama M, Manuel LJ, Togasaki RK. 1989. Isolation and Characterization of a Mutant of Chlamydomonas reinhardtii Deficient in the CO(2) Concentrating Mechanism. Plant Physiology 89, 897-903.

Moroney JV, Jungnick N, Dimario RJ, Longstreth DJ. 2013. Photorespiration and carbon concentrating mechanisms: two adaptations to high O2, low CO2 conditions. Photosynthesis Research 117, 121-131.

Moroney JV, Wilson BJ, Tolbert NE. 1986. Glycolate Metabolism and Excretion by Chlamydomonas reinhardtii. Plant Physiology 82, 821-826.

Moroney JV, Ynalvez RA. 2007. Proposed carbon dioxide concentrating mechanism in Chlamydomonas reinhardtii. Eukaryotic Cell 6, 1251-1259.

Mukherjee A, Lau CS, Walker CE, Rai AK, Prejean CI, Yates G, Emrich-Mills T, Lemoine SG, Vinyard DJ, Mackinder LCM, Moroney JV. 2019. Thylakoid localized bestrophin-like proteins are essential for the CO2 concentrating mechanism of Chlamydomonas reinhardtii. Proceedings of the National Academy of Sciences, USA 116, 16915-16920.

Muller SY, Matthews NE, Valli AA, Baulcombe DC. 2020. The small RNA locus map for Chlamydomonas reinhardtii. PLoS One 15, e0242516.

Nordhues A, Schöttler MA, Unger AK, Geimer S, Schönfelder S, Schmollinger S, Rütgers M, Finazzi G, Soppa B, Sommer F, Mühlhaus T, Roach T, Krieger-Liszakay A, Lokstein H, Crespo JL, Schroda M. 2012. Evidence for a role of VIPP1 in the structural organization of the photosynthetic apparatus in Chlamydomonas. Plant Cell 24, 637-659.

Pechmann S, Frydman J. 2014. Interplay between chaperones and protein disorder promotes the evolution of protein networks. PLoS Computational Biology 10, e1003674.

Pfannschmidt T, Terry MJ, Van Aken O, Quiros PM. 2020. Retrograde signals from endosymbiotic organelles: a common control principle in eukaryotic cells. Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences 375, 20190396.

Plumed MP, Villarejo A, Roos A, García-Reina G, Ramazanov Z. 1996. The CO2-concentrating mechanism in a starchless mutant of the green unicellular alga Chlorella pyrenoidosa. Planta 200, 28-31.
Polukhina I, Fristedt R, Dinc E, Cardol P, Croce R. 2016. Carbon Supply and Photoacclimation Cross Talk in the Green Alga Chlamydomonas reinhardtii. Plant Physiology 172, 1494-1505.

Pribil M, Sandoval-Ibanez O, Xu W, Sharma A, Labs M, Liu Q, Galgenmuller C, Schneider T, Wessels M, Matsubara S, Jansson S, Winner G, Leister D. 2018. Fine-Tuning of Photosynthesis Requires CURVATURE THYLAKOID1-Mediated Thylakoid Plasticity. Plant Physiology 176, 2351-2364.

Ramazanov Z, Rawat M, Henk MC, Mason CB, Matthews SW, Moroney JV. 1994. The induction of the CO2-concentrating mechanism is correlated with the formation of the starch sheath around the pyrenoid of Chlamydomonas reinhardtii. Planta 195, 210-216.

Rawat M, Moroney JV. 1995. The Regulation of Carbonic Anhydrase and Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase Activase by Light and CO2 in Chlamydomonas reinhardtii. Plant Physiology 109, 937-944.

Rea G, Antonacci A, Lambreva MD, Mattio AK. 2018. Features of cues and processes during chloroplast-mediated retrograde transcription in the algae Chlamydomonas. Plant Science 272, 193-206.

Riano-Pachon DM, Correa LG, Trejos-Espinosa R, Mueller-Roeber B. 2008. Green transcription factors: a chlamydomonas overview. Genetics 179, 31-39.

Rutgers M, Muranaka LS, Muhlhaus T, Sommer F, Thoms S, Schurig J, Willmund F, Schulz-Raffelt M, Schröda M. 2017. Substrates of the chloroplast small heat shock proteins 22E/F point to thermolability as a regulative switch for heat acclimation in Chlamydomonas reinhardtii. Plant Molecular Biology 95, 579-591.

Sager R, Palade GE. 1957. Structure and development of the chloroplast in Chlamydomonas. I. The normal green cell. The Journal of Biophysical and Biochemical Cytology. 3, 463-488.

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 Biophysica Acta 1817, 1248-1255.

Singh R, Singh S, Parihar P, Singh VP, Prasad SM. 2015. Retrograde signaling between plastid and nucleus: A review. Journal of Plant Physiology 181, 55-66.

Spalding MH, Spritzer RJ, Ogren WL. 1985. Use of mutants in analysis of the CO2-concentrating pathway of Chlamydomonas Reinhardtii. Botany Publication and Papers 72, http://lib.dr.iastate.edu/bot_pubs/72.

Strenkert D, Schmollinger S, Gallaher SD, Salome PA, Purvine SO, Nicora CD, Mettler-Altmann T, Soubyrand E, Weber APM, Lipton MS, Basset GJ, Merchant SS. 2019. Multiomics resolution of molecular events during a day in the life of Chlamydomonas. Proceedings of the National Academy of Sciences, USA 116, 2374-2383.

Strenkert D, Schmollinger S, Sommer F, Schulz-Raffelt M, Schröda M. 2011. Transcription factor-dependent chromatin remodeling at heat shock and copper-responsive promoters in Chlamydomonas reinhardtii. Plant Cell 23, 2285-2301.

Suzuki K, Marek LF, Spalding MH. 1990. A Photorespiratory Mutant of Chlamydomonas reinhardtii. Plant Physiology 93, 231-237.

Tabrizi TS, Sawicki A, Zhou S, Luo M, Willows RD. 2016. GUN4-Protoporphyrin IX Is a Singlet Oxygen Generator with Consequences for Plastid Retrograde Signaling. Journal of Biological Chemistry 291, 8978-8984.

Tanaka R, Tanaka A. 2007. Tetrapyrrole biosynthesis in higher plants. Annual Review of Plant Biology 58, 321-346.

Theis J, Niemeyer J, Schmollinger S, Ries F, Rütgers M, Gupta TK, Sommer F, Muranaka LS, Venn B, Schulz-Raffelt M, Willmund F, Engel BD, Schröda M. 2020. VIPP2 interacts with VIPP1 and HSP22E/F at chloroplast membranes and modulates a retrograde signal for HSP22E/F gene expression. Plant Cell and Environment 43, 1212-1229.

Tirumani S, Gothandam KM, B JR. 2019. Coordination between photorespiration and carbon concentrating mechanism in Chlamydomonas reinhardtii: transcript and protein changes during light-dark diurnal cycles and mixotrophy conditions. Protoplasma 256, 117-130.

Tirumani S, Kokkanti M, Chaudhari V, Shukla M, Rao BJ. 2014. Regulation of CCM genes in Chlamydomonas reinhardtii during conditions of light-dark cycles in synchronous cultures. Plant Molecular Biology 85, 277-286.

Toyokawa C, Yamano T, Fukuzawa H. 2020. Pyrenoid Starch Sheath Is Required for LCIB Localization and the CO2-Concentrating Mechanism in Green Algae. Plant Physiology 182, 1883-1893.
Tural B, Moroney JV. 2005. Regulation of the expression of photorespiratory genes in Chlamydomonas reinhardtii. Canadian Journal of Botany 83, 810-819.

Valli AA, Santos BA, Hnatova S, Bassett AR, Molnar A, Chung BY, Baulcombe DC. 2016. Most microRNAs in the single-cell alga Chlamydomonas reinhardtii are produced by Dicer-like 3-mediated cleavage of introns and untranslated regions of coding RNAs. Genome Research 26, 519-529.

Verdecia MA, Larkin RM, Ferrer JL, Riek R, Chory J, Noel JP. 2005. Structure of the Mg-chelatase cofactor GUN4 reveals a novel hand-shaped fold for porphyrin binding. PLoS Biology 3, e151.

Villarasa-Blasi J, Fauser F, Onishi M, Ramundo S, Patena W, Millican M, Osaki J, Philp C, Nemeth M, Salomé PA, Li X, Wakao S, Kim RG, Kaye Y, Grossman AR, Niyogi KK, Merchant S, Cutler S, Walter P, Dinneny JR, Jonikas MC, Jinkerson RE. 2020. Systematic characterization of gene function in a photosynthetic organism. BioRxiv doi.org/10.1101/2020.12.11.420950, [Preprint]

Villand P, Eriksson M, Samuelsson G. 1997. Carbon dioxide and light regulation of promoters controlling the expression of mitochondrial carbonic anhydrase in Chlamydomonas reinhardtii. Biochemical Journal 327, 51-57.

Villarejo A, Martínez F, Plumed MP, Ramazanov Z. 1996. The induction of the CO₂ concentrating mechanism in a starch-less mutant of Chlamydomonas reinhardtii. Physiologia Plantarum 98.

Wang L. 2017. A calcium-binding protein CAS regulates the CO₂-concentrating mechanism in the green alga Chlamydomonas reinhardtii. Ph.D thesis, Kyoto University.

Wang L, Yamano T, Takane S, Niikawa Y, Toyokawa C, Ozawa SI, Tokutsu R, Takahashi Y, Minagawa J, Kanesaki Y, Yosihikawa H, Fukuzawa H. 2016. Chloroplast-mediated regulation of CO₂-concentrating mechanism by Ca2+-binding protein CAS in the green alga Chlamydomonas reinhardtii. Proceedings of the National Academy of Sciences, USA 113, 12586-12591.

Wang Y, Spalding MH. 2014. Acclimation to very low CO₂: contribution of limiting CO₂ inducible proteins, LCIB and LCIA, to inorganic carbon uptake in Chlamydomonas reinhardtii. Plant Physiology 166, 2040-2050.

Wang Y, Steissman DJ, Spalding MH. 2015. The CO₂ concentrating mechanism and photosynthetic carbon assimilation in limiting CO₂: how Chlamydomonas works against the gradient. Plant Journal 82, 429-448.

Westphal S, Heins L, Soll J, Vothknecht UC. 2001. Vipp1 deletion mutant of synechocystis: A connection between bacterial phage shock and thylakoid biogenesis? Proceedings of the National Academy of Sciences, USA 98, 4243-4248.

Winck FV, Arvidsson S, Riano-Pachon DM, Hempel S, Koseska A, Nikoloski Z, Urbina Gomez DA, Rupprecht J, Mueller-Roeber B. 2013. Genome-wide identification of regulatory elements and reconstruction of gene regulatory networks of the green alga Chlamydomonas reinhardtii under carbon deprivation. PloS One 8, e79909.

Winck FV, Melo DO, Riano-Pachon DM, Martins MC, Caldana C, Barrios AF. 2016. Analysis of Sensitive CO₂ Pathways and Genes Related to Carbon Uptake and Accumulation in Chlamydomonas reinhardtii through Genomic Scale Modeling and Experimental Validation. Frontiers in Plant Science 7, 43.

Wunder T, Cheng SLH, Lai SK, Li HY, Mueller-Cajar O. 2018. The phase separation underlying the pyrenoid-based microalgal Rubisco supercharger. Nature Communications 9, 5076.

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, Asada A, Sato E, Fukuzawa H. 2014. Isolation and characterization of mutants defective in the localization of LCIB, an essential factor for the carbon-concentrating mechanism in Chlamydomonas reinhardtii. Proceedings of the National Academy of Sciences, USA 98, 5341-5346.

Yamano T, Sato E, Iwuchi H, Fukuda Y, Fukuzawa H. 2015. Characterization of cooperative bicarbonate uptake into chloroplast stroma in the green alga Chlamydomonas reinhardtii. Photosynthesis Research 121, 193-200.

Yamano T, Sato E, Iwuchi H, Fukuda Y, Fukuzawa H. 2018. 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.

Yamano T, Toyokawa C, Fukuzawa H. 2018. High-resolution suborganellar localization of Ca(2+)-binding protein CAS, a novel regulator of CO₂-concentrating mechanism. Protoplasma 255, 1015-1022.
Yamano T, Tsujikawa T, Hatano K, Ozawa S, Takahashi Y, Fukuzawa H. 2010. Light and low-
CO2-dependent LCIB-LCIC complex localization in the chloroplast supports the carbon-
concentrating mechanism in Chlamydomonas reinhardtii. Plant and Cell Physiology 51, 1453-1468.
Yoshioka S, Taniguchi F, Miura K, Inoue T, Yamano T, Fukuzawa H. 2004. The novel Myb
transcription factor LCR1 regulates the CO2-responsive gene Cah1, encoding a periplasmic carbonic
anhydrase in Chlamydomonas reinhardtii. Plant Cell 16, 1466-1477.
Zhan Y, Marchand CH, Maes A, Mauries A, Sun Y, Dhaliwal JS, Uniacke J, Arragain S, Jiang
H, Gold ND, Martin VJJ, Lemaire SD, Zerges W. 2018. Pyrenoid functions revealed by
proteomics in Chlamydomonas reinhardtii. PLoS One 13, e0185039.
Zhao T, Li G, Mi S, Li S, Hannon GJ, Wang XJ, Qi Y. 2007. A complex system of small RNAs in
the unicellular green alga Chlamydomonas reinhardtii. Genes and Development 21, 1190-1203.
Zones JM, Blaby IK, Merchant SS, Umen JG. 2015. High-Resolution Profiling of a Synchronized
Diurnal Transcriptome from Chlamydomonas reinhardtii Reveals Continuous Cell and Metabolic
Differentiation. Plant Cell 27, 2743-2769.
Table 1: Transcription factors and regulators occurring with different relative abundances in *Chlamydomonas* grown in low (0.04%) and high (5%) CO\textsubscript{2} conditions (Arias et al., 2020)

| Protein                     | Description                                      | TF family\(^{3}\) | Fold change |
|-----------------------------|--------------------------------------------------|-------------------|-------------|
| **Transcription factors**   |                                                  |                   |             |
| Cre01.g000050.t1.1          | RWP-RK Transcription Factor                      | RWP-RK            | 5.5         |
| Cre10.g444450.t1.1          | Predicted Protein                                | C3H               | 3.1         |
| Cre14.g625802.t1.1          | Ring Finger Protein-Related                      | FHA               | 2.9         |
| Cre16.g656250.t1.1          | U1 Small Nuclear Ribonucleoprotein               | CSD               | 2.5         |
| Cre17.g714500.t1.2          | Histone H2A                                       | CCAAT             | 2.5         |
| Cre06.g288750.t1.2          | Nuclear Rna Cap-Binding Protein                  | CSD               | 2.4         |
| Cre06.g254650.t1.2          | Zinc Finger Protein 183                          | C3H               | 1.9         |
| Cre14.g632050.t1.2          | RPGR-Interacting Protein 1 Related               | VARL              | 1.9         |
| Cre01.g035150.t1.1          | Zinc Finger (CCCH-Type) Family Protein           | C3H               | 1.9         |
| Cre10.g446900.t1.2          | WD40 Repeat Proteinprl1/PRLl2-Related            | Orphans           | 1.7         |
| Cre02.g115250.t1.1          | Centriole Proteome Protein                       | Orphans           | 1.6         |
| Cre12.g523200.t1.1          | Nucleosome Remodeling Factor                     | Orphans           | 1.6         |
| Cre09.g389550.t1.1          | Dnaj-Like Protein                                | MYB-related       | 1.6         |
| Cre03.g197350.t1.2          | Cell Division Cycle 5-Like Protein               | MYB-related       | 1.5         |
| Cre17.g713900.t1.2          | Tor Kinase Binding Protein                       | Orphans           | 1.5         |
| Cre01.g020400.t1.2          | WD40 Repeat Protein                              | Orphans           | 2.3         |
| Cre09.g392350.t1.2          | Rna Recognition Motif. (Rnp Domain) (Rrm_1)      | CSD               | 1.9         |
| Cre01.g035000.t1.2          | Wd Repeat Protein                                | Orphans           | 1.8         |
| Cre17.g729150.t1.2          | Rna Recognition Motif. (Rnp Domain) (Rrm_1)      | CSD               | 1.7         |
| Cre16.g662800.t1.2          | Splicing Factor, Component Of The U4/U6-U5 Snrnp Complex | Orphans           | 1.6         |
| Cre06.g275100.t1.2          | Splicing Factor 3B, Subunit 4                    | CSD               | 1.5         |
| Cre06.g274200.t1.2          | Histone H2A                                       | CCAAT             | 1.6         |
| Cre12.g507650.t1.2          | Chloroplast Dnaj-Like Protein                    | MYB-related       | 1.6         |

| **Transcription regulators**|                                                  |                   |             |
| Cre16.g668200.t1.1          | Chromatin Remodeling Protein, Contains Phd Zn-Finger | ARID              | 3.9         |
| Cre16.g672300.t1.2          | Swi/Snf-Related Chromatin Binding Protein        | HMG               | 2.8         |
| Cre01.g015050.t1.1          | Unknown                                          | SNF2              | 2.2         |
| Cre07.g322450.t1.1          | Pwwp Domain (Pwwp) // Set Domain (Set)           | PHD               | 1.7         |
| Cre08.g380151.t1.1          | Phd-Finger (Phd) // Wstf, Hb1, Itc1P, Mbd9 Motif  | PHD               | 1.7         |
| Cre07.g334200.t1.2          | Atp-Dependent Rna Helicase Ddx41-Related          | SNF2              | 1.6         |
| Cre02.g078700.t1.1          | Lysine-Specific Demethylase 4A-Related            | JUMONJI           | 1.5         |
| Cre06.g261450.t1.2          | Swi/Snf-Related Chromatin Binding Protein        | HMG               | 1.6         |
| Cre17.g709550.t1.2          | Histone Demethylase 1A                           | JUMONJI           | 1.8         |
| Cre01.g029450.t1.1          | Non-Histone Protein 10                           | HMG               | 1.5         |
| Cre08.g367300.t1.1          | Bromodomain Extra-Terminal - Bet                 | DDT               | 1.5         |
| Cre08.g358532.t1.1          | Gata Zinc Finger (Gata) // Bah Domain (Bah)      | PHD               | 1.5         |

\(^{3}\)Transcription factor(TF) family has been determined from Plant transcription factor database. Only proteins with fold change \(\geq 1.5\) are shown in table. Proteins more abundant in low CO\textsubscript{2} conditions are indicated in bold.

\(^{*}\)Proteins have been detected only in one of the two conditions- low/high CO\textsubscript{2}
Table 2: List of photorespiration genes in *Chlamydomonas*

| Gene ID     | Short name | Brief description                                                                 |
|-------------|------------|----------------------------------------------------------------------------------|
| Cre03.g168700 | PGLP1      | Phosphoglycolate phosphatase/4-nitrophenylphosphatase                              |
| Cre10.g438100 | PGLP2      | Phosphoglycolate phosphatase/4-nitrophenylphosphatase                              |
| Cre03.g162601 | PGLP3      | CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase                 |
| Cre06.g288700 | GDH        | Glycolate dehydrogenase                                                           |
| Cre10.g451950 | AAT        | Alanine aminotransferase                                                          |
| Cre06.g294650 | AGT1       | Alanine-glyoxylate transaminase                                                   |
| Cre03.g182800 | AGT2       | Alanine-glyoxylate transaminase                                                   |
| Cre12.g534800 | GDC-P      | Glycine cleavage system, P protein                                                |
| Cre06.g253350 | GDC-H      | Glycine cleavage system, H-protein                                                |
| Cre03.g193750 | GDC-T      | Glycine cleavage system, T protein                                                |
| Cre18.g749847 | DLDH       | Dihydrolipoyl dehydrogenase                                                       |
| Cre16.g664550 | SHMT1      | Serine hydroxymethyltransferase                                                    |
| Cre06.g293950 | SHMT2      | Serine hydroxymethyltransferase 2                                                 |
| Cre09.g411900 | SHMT3      | Serine hydroxymethyltransferase 3                                                 |
| Cre01.g005150 | SGAT       | Serine glyoxylate aminotransfer                                                   |
| Cre06.g295450 | HPR1       | Hydroxyppyrurate reductase                                                         |
| Cre12.g542300 | GLYK       | Glycerate kinase                                                                  |

The genes highlighted in bold were identified to be regulated by both CIA5 and CO₂, and were classified having expression patterns similar to CCM clusters (Fang et al., 2012)
Table 3: CCM proteins that are found to interact with chaperones from proteomics studies (Mackinder et al., 2017, Rutgers et al., 2017)

| Protein ID | Description                           | Protein ID | Description                           |
|------------|---------------------------------------|------------|---------------------------------------|
| Cre06.g295450 | HRP1- Hydroxypyruvate reductase      | Cre03.g162800 | LC11- Low-CO2-inducible membrane protein |
| Cre05.g248450 | CAH5- Mitochondrial carbonic anhydrase | Cre16.g6652800 | Unannotated                           |
| Cre10.g444700 | SBE3- Starch binding enzyme          | Cre04.g229300 | RCA1- rubisco activase                |
| Cre03.g151650 | Unannotated                          | Cre12.g509050 | PSBP3- OEE2-like protein of thylakoid lumen |
| Cre16.g651050 | CYC6- Cytochrome c6                  | Cre13.g577100 | ACP2- acyl-carrier protein            |
| Cre10.g444700 | SBE3- Starch binding enzyme          | Cre16.g651050 | CYC6- Cytochrome c6                  |
| Cre03.g179800 | LCI24- low-CO2-inducible membrane protein | Cre09.g394473 | LCI9- low-CO2-inducible protein      |
| Cre16.g663450 | LCI11- Low-CO2-inducible membrane protein | Cre12.g560950 | PSAG- photosystem I reaction center subunit V |
| Cre16.g662600 | Unannotated                          | Cre10.g436550 | EPYC1/ LCI5- low-CO2-inducible protein |
| Cre10.g444700 | SBE3- Starch binding enzyme          | Cre02.g120150 | Unannotated                           |
| Cre03.g179800 | LCI24- low-CO2-inducible membrane protein | Cre02.g120100 | RBCS1-Rubisco small subunit 1         |
| Cre16.g663450 | LCI11- Low-CO2-inducible membrane protein | Cre07.g330250 | PSAG-subunit H of photosystem I       |
| Cre06.g307500 | LCI1- CO2-inducible protein          | Cre14.g626700 | PETF- apoferredoxin                    |
| Cre01.g054850 | Unannotated                          | Cre12.g507300 | LCI30- low-CO2-inducible protein      |
| Cre16.g663450 | LCI11- Low-CO2-inducible membrane protein | Cre12.g519300 | TEF9- Unannotated                     |
| Cre06.g307500 | LCI1- CO2-inducible protein          | Cre16.g662600 | Unannotated                           |
| Cre01.g054850 | Unannotated                          | Cre16.g663450 | LCI11- Low-CO2-inducible membrane protein |
| Cre16.g663450 | LCI11- Low-CO2-inducible membrane protein | Cre06.g283750 | HST1- homogentisate solanesyltransferase |
| Cre08.g372450 | PSBQ- oxygen evolving enhancer protein 3 | Cre10.g444700 | SBE3- Starch binding enzyme            |
| Cre02.g097800 | HL3A- ABC transporter                | Cre03.g151650 | Unannotated                           |
| Cre17.g724300 | PSAK-photosystem I reaction center subunit | Cre16.g6652800 | Unannotated                           |
| Cre09.g415700 | CAH3- carboxy anhydrase 3            | Cre17.g724300 | PSAK-photosystem I reaction center subunit |
| Cre01.g151650 | Unannotated                          | Cre09.g394473 | LCI9- low-CO2-inducible protein      |
| Cre06.g309000 | LCI1- CO2-inducible protein          | Cre10.g436550 | EPYC1/ LCI5- low-CO2-inducible protein |
| Cre01.g054850 | Unannotated                          | Cre06.g307500 | LCI1- CO2-inducible protein            |
| Cre03.g151650 | Unannotated                          | Cre06.g309000 | NAR1.2- anion transporter              |
| Cre06.g295450 | HRP1- Putative hydroxypyruvate reductase | Cre12.g519300 | TEF9- Unannotated                     |
| Cre16.g663450 | LCI11- CO2-inducible protein         | Cre03.g191250 | LCI34- Low-CO2-inducible protein      |
| Cre06.g309000 | LCI1- CO2-inducible protein          | Cre01.g051500 | ULPI- uncharacterized luminal polypeptide |
| Cre03.g151650 | Unannotated                          | Cre09.g415700 | CAH3- carboxy anhydrase 3            |
| Cre12.g485050 | CAH6- carboxy anhydrase 6            | Cre02.g097800 | HL3A- ABC transporter                |
| Cre01.g051500 | ULPI- uncharacterized luminal polypeptide | Cre10.g452800 | LCI1- CO2-inducible protein            |
| Cre03.g179800 | LCI1- CO2-inducible protein          | Cre04.g229300 | RCA1- rubisco activase                |
| Cre06.g283750 | HST1- homogentisate solanesyltransferase | Cre12.g509050 | PSBP3- OEE2-like protein of thylakoid lumen |
| Cre04.g223050 | CAH2- carboxy anhydrase, alpha type, periplasmic | Cre02.g120100 | RBCS1-Rubisco small subunit 1         |
| Cre03.g191250 | LCI34- Low-CO2-inducible protein     | Cre12.g560950 | PSAG- photosystem I reaction center subunit V |
| Protein ID          | Description                                                                 |
|--------------------|------------------------------------------------------------------------------|
|Cre01.g054850       | Unannotated                                                                 |
|Cre02.g120150       | Unannotated                                                                 |
|Cre03.g179800       | LC124- low-CO2-inducible membrane protein                                     |
|Cre04.g223300       | CCP1- low-CO2-inducible chloroplast envelope protein                         |
|Cre16.g651050       | CYC6- Cytochrome c6                                                           |
|Cre08.g372450       | PSBQ- oxygen evolving enhancer protein 3                                     |
|Cre10.g436550       | EPYC1L/LC15- low-CO2-inducible protein                                        |
|Cre08.g362900       | PSBP4- luminal PsbP-like protein                                              |
|Cre04.g223050       | CAH2- carbonic anhydrase, alpha type, periplasmic                             |
|Cre12.g485050       | CAH6- carbonic anhydrase 6                                                   |
|Cre03.g162800       | LC11- Low-CO2-inducible membrane protein                                      |
|Cre05.g248450       | CAH5- Mitochondrial carbonic anhydrase                                       |
|Cre07.g330250       | PSAH-subunit H of photosystem I                                              |
|Cre12.g507300       | LC130- low-CO2-inducible protein                                             |
|Cre06.g295450       | HRP1- Putative hydroxypropyruvate reductase                                  |
|Cre14.g626700       | PETF- apoferritin                                                             |
|Cre13.g577100       | ACP2- acyl-carrier protein                                                   |
|Cre17.g721500       | STA2- granule-bound starch synthase I                                         |
|              |                                                                                   |
|Protein ID          | Description                                                                 |
|--------------------|------------------------------------------------------------------------------|
|Cre05.g248435       | CAH5- Mitochondrial carbonic anhydrase                                       |
|Cre16.g662600       | Unannotated                                                                 |
|Cre10.g436550       | EPYC1L/LC15- low-CO2-inducible protein                                        |
|Cre09.g394473       | LC19- low-CO2-inducible protein                                              |
|Cre06.g295450       | HRP1- Putative hydroxypropyruvate reductase                                  |
|Cre16.g663450       | LC11- Low-CO2-inducible membrane protein                                      |
|Cre03.g151650       | Unannotated                                                                 |
|Cre06.g283750       | HST1- homogentisate solanesyltransferase                                     |
|Cre04.g229300       | RCA1- rubisco activase                                                        |
|Cre16.g652800       | Unannotated                                                                 |
|Cre10.g444700       | SBE3- Starch binding enzyme                                                  |
|Cre06.g309000       | NAR1.2- anion transporter                                                     |
|Cre02.g097800       | HLA3- ABC transporter                                                        |
|Cre17.g724300       | PSAK-photosystem I reaction center subunit                                    |
|Cre16.g651050       | CYC6- Cytochrome c6                                                           |
|Cre09.g415700       | CAH3- carbonic anhydrase 3                                                   |
|Cre04.g223300       | CCP1- low-CO2-inducible mitochondrial envelope protein                       |
Table 4: Nuclear encoded CCM genes\* with greater than 8-fold change in expression in gun4 mutant with respect to wt *Chlamydomonas*

| Gene ID       | Short name | Short description                        | Gene ID       | Short name | Short description                        |
|--------------|------------|------------------------------------------|--------------|------------|------------------------------------------|
| Cre01.g014350| PRX5       | Type II peroxiredoxin                    | Cre08.g360200| DUR3       | Urea active transporter                  |
| Cre01.g015350|            | Light-dependent protochlorophyllide reductase | Cre09.g405750| CAH8       | Carbonic anhydrase                       |
| Cre01.g029250| Amino acid hydroxylase-like protein     | Cobalamin-5'-phosphate synthase         | Cre10.g426050| CTPA1      | C-terminal processing peptidase          |
| Cre01.g036950|            | Cobalamin-5'-phosphate synthase         | Cre10.g455700|            | Non-canonical poly(A) polymerase         |
| Cre01.g053950| MOX       | Monoxygenase                             | Cre12.g485150| GAP1       | Glyceraldehyde 3-phosphate dehydrogenase |
| Cre02.g078507| PSII Pbs27| Photosystem II Pbs27                    | Cre12.g519300| TEF9       | Predicted protein                        |
| Cre02.g085500|           | Putative transposase DNA-binding domain | Cre12.g53250  |            | RNA polymerase s factor                  |
| Cre02.g107000|           | Cyclin dependent kinase-2               | Cre12.g541550|            | Unknown function                         |
| Cre02.g143450|           | Unknown function                         | Cre12.g555700| DNJ15      | DnaJ-like protein                        |
| Cre02.g144800| LC18      | Acetylglutamate kinase                   | Cre13.g569000|            | Antibiotic biosynthesis monoxygenase     |
| Cre03.g149050| Cyt b561  | Cytochrome b561                          | Cre14.g625450|            | Methyltransferase                        |
| Cre03.g158000|           | Glu-1-semialdehyde aminotransferase     | Cre14.g630350|            | Unknown function                         |
| Cre03.g171350| SEC61A    | SEC61-α subunit                         | Cre16.g652800|            | Unknown function                         |
| Cre03.g200350| Methyltransferase                      | Cre16.g658400| FDX2       | Ferredoxin                               |
| Cre04.g223250|           | LCIB-like gene                           | Cre16.g659800|            | Unknown function                         |
| Cre05.g236650| CYG63     | Guanilate cyclase                        | Cre16.g663450| BST-3      | Bestrophin-3                             |
| Cre05.g248400| CAH4      | Mitochondrial carbonic anhydrase         | Cre16.g685100|            | Cobalamin synthesis protein              |
| Cre05.g248450| CAH5      | Mitochondrial carbonic anhydrase         | Cre17.g700950| FDX5       | Apoferredoxin                            |
| Cre06.g258850|           | Stage V sporulation protein S            | Cre17.g713700|            | Tryptophan pyrrolase                     |
| Cre06.g266450| Protein kinase (MEC-15)               | Cre17.g720900|            | Unknown function                         |
| Cre06.g284150| RHP2      | Ammonium transport protein               | Cre01.g038400|            | Calreticulin 2                           |
| Cre06.g303050| Nitrate reductase                      | Cre02.g111450|            | Rhodanese-like protein                   |
| Cre06.g310950| Sarcosine dehydrogenase               | Cre03.g198950|            | PSBP domain carrying protein             |
| Cre07.g315050| Gamma-glutamyl hydrolase              | Cre06.g293950| SHMT2      | Serine hydroxymethyltransferase 2        |
| Cre07.g330250| PSAH      | Photosystem I- subunit H                 | Cre07.g321400| FAP113     | Flagellar associated protein             |
| Cre07.g337100|           | Unknown function                         | Cre17.g724300| PSAK       | Photosystem I subunit Psak               |

\*CCM genes were compiled from Mackinder et al., 2017, and Strenkert et al., 2019. All genes upregulated in gun4 mutant are highlighted in bold.
Table 5: Nuclear encoded genes with upregulation greater than 4-fold in wt *Chlamydomonas* with respect to *cas* mutant

| Protein id       | Short name | Short description                                      |
|------------------|------------|--------------------------------------------------------|
| Cre02.g097800    | HLA3       | ABC transporter                                        |
| Cre06.g309000    | LCIA       | Anion transporter                                      |
| Cre03.g204577    | DNJ31      | DnaJ-like protein                                      |
| Cre05.g248400    | CAH4       | Mitochondrial carbonic anhydrase, beta type            |
| Cre05.g248450    | CAH5       | Mitochondrial carbonic anhydrase                       |
| Cre07.g334750*   | PPP30      | Protein phosphatase 2C                                  |
| Cre04.g223300    | CCP1       | Low-CO2-inducible mitochondrial protein                |
| Cre04.g222750    | CCP2       | Low-CO2-inducible mitochondrial protein                |
| Cre04.g222800    | LCID       | Low-CO2 inducible protein                              |
| Cre08.g367500    | LHCSR3.1   | Stress-related chlorophyll a/b binding protein 2       |
| Cre08.g367400    | LHCSR3.2   | Stress-related chlorophyll a/b binding protein 3       |
| Cre12.g541550    | --         | Uncharacterised                                        |
| Cre26.g756747    | --         | Uncharacterised                                        |

*PPP30 is an uncharacterised gene, which has not been identified as a CCM gene in any previous study.*
**Figure 1** Photorespiratory cycle in *Chlamydomonas*: The enzymes Rubisco, PGLP (phosphoglycolate phosphatase), GDH (glycolate dehydrogenase), GGT (glutamate glyoxalate aminotransferase), GDC (glycine decarboxylase complex), SHMT (serine hydroxymethyl transferase), SGAT (serine/alanine glyoxalate aminotransferase), HPR1 (hydroxypyruvate reductase), and GLYK (glycerate kinase) are in red. Other abbreviations used: Glu: Glutamate, 2-OG: 2-Oxoglutarate, Ala: Alanine, Pyr: Pyruvate. The enzymes highlighted in bold have expression dependent on both [CO₂] and CIA5, similar to several CCM genes (Fang et al., 2012).

**Figure 2**: Schematic representation of crosstalk between photosynthetic electron transport (PET), Calvin Benson Basham cycle (CBB), photorespiration (PR) and carbon concentration mechanism (CCM) in *Chlamydomonas*. CCM components- inorganic carbon transporters and carbonic anhydrases, occurring in various parts of the cell are highlighted in grey. The role of mitochondrial proteins CCP1, CCP2, CAH4 and CAH5 is hypothesized, and remains to be explored. Reactive oxygen species (ROS) generated during PET, and PR metabolites are hypothesized to act as signalling molecules for CCM.

**Figure 3**: Assembly of the pyrenoid- The Rubisco binding Motif (RbM) mediates the formation of three regions of the pyrenoid. The RbM bearing protein EPYC1 binds to multiple Rubisco holoenzymes and creates a Rubisco-EPYC1 condensate that forms the pyrenoid matrix. The interaction of RbM bearing thylakoid anchored proteins RBMP1 and RBMP2 with Rubisco tethers the pyrenoid matrix to the tubule network. The starch sheath is moulded around the pyrenoid matrix through the action of SAGA1 and SAGA2, which bind to Rubisco through their RbM domain and bind to the starch sheath through their starch binding domain.

**Figure 4**: (a) GUN4 retrograde signalling model (i) GUN4 is proposed to be an activator of MgCh activity, interacting with the ChlH subunit to promote the catalytic integration of Mg²⁺ with ProtoIX to form the chlorophyll biosynthesis pathway intermediate Mg-ProtoIX. (ii) The accumulation of excess tetrapyrrole intermediates, such as ProtoIX, in the chloroplast can lead to generation of ROS. GUN4 is proposed to bind ProtoIX, shielding its reaction with ROS. In shielding ProtoIX, GUN4 may be progressively modified or degraded with degradation products hypothesised to act as the retrograde signals. (b) A contrasting model for GUN4 - In stead of having a ‘shielding’ effect when bound to ProtoIX, the GUN4-ProtoIX complex appeared to escalate ^1^O₂ generation. The elevated ^1^O₂ produced by GUN4-ProtoIX may be sensed by a ^1^O₂ sensing system (like the *Arabidopsis*...
EXECUTER1/EXECUTER2 or EX1/EX2 system) yet to be discovered, that relays a signal to the nucleus.

**Figure 5:** A schematic of a tentative mechanism for CAS activity in Chlamydomonas retrograde signalling. A Under low light/high CO$_2$ conditions, CAS is dispersed throughout the chloroplast. B Under high light/low CO$_2$ the ETC proteins become over-reduced, triggering the movement of CAS into the pyrenoid along the pyrenoid tubules. In the Ca$^{2+}$-rich pyrenoid, CAS binds to Ca$^{2+}$ and becomes activated. This form of CAS signals back to the nucleus to modulate target genes. It also induces an increase in intracellular Ca$^{2+}$.

**Figure 6:** Relative expression of PR (top) and CCM (bottom) genes in *wt* and *cia5* Chlamydomonas grown in different [CO$_2$]-<0.02% (V-very low), 0.03-0.05% (L-Low) and 5% (H-High). The expression of only genes classified as being in CCM clusters (Fang et al., 2012) are shown here.
Figure 1 Photorespiratory cycle in *Chlamydomonas*: The enzymes Rubisco, PGLP (phosphoglycolate phosphatase), GDH (glycolate dehydrogenase), GGT (glutamate glyoxalate aminotransferase), GDC (glycine decarboxylase complex), SHMT (serine hydroxymethyltransferase), SGAT (serine/alanine glyoxalate aminotransferase), HPR1 (hydroxypyruvate reductase), and GLYK (glycerate kinase) are in red. Other abbreviations used: Glu: Glutamate, 2-OG: 2-Oxoglutarate, Ala: Alanine, Pyr: Pyruvate. The enzymes highlighted in bold have expression dependent on both [CO$_2$] and CIA5, similar to several CCM genes (Fang et al., 2012).
Figure 2: Schematic representation of crosstalk between photosynthetic electron transport (PET), Calvin Benson Basham cycle (CBB), photorespiration (PR) and carbon concentration mechanism (CCM) in *Chlamydomonas*. CCM components- inorganic carbon transporters and carbonic anhydrases, occurring in various parts of the cell are highlighted in grey. *The role of mitochondrial proteins CCP1, CCP2, CAH4 and CAH5 is hypothesized, and remains to be explored. Reactive oxygen species (ROS) generated during PET, and PR metabolites are hypothesized to act as signalling molecules for CCM.*
Figure 3: Assembly of the pyrenoid- The Rubisco binding Motif (RbM) mediates the formation of three regions of the pyrenoid. The RbM bearing protein EPYC1 binds to multiple Rubisco holoenzymes and creates a Rubisco-EPYC1 condensate that forms the pyrenoid matrix. The interaction of RbM bearing thylakoid anchored proteins RBMP1 and RBMP2 with Rubisco tethers the pyrenoid matrix to the tubule network. The starch sheath is moulded around the pyrenoid matrix through the action of SAGA1 and SAGA2, which bind to Rubisco through their RbM domain and bind to the starch sheath through their starch binding domain.
Figure 4: (a) GUN4 retrograde signalling model (i) GUN4 is proposed to be an activator of MgCh activity, interacting with the ChlH subunit to promote the catalytic integration of Mg$^{2+}$ with ProtoIX to form the chlorophyll biosynthesis pathway intermediate Mg-ProtoIX. (ii) The accumulation of excess tetrapyrrole intermediates, such as ProtoIX, in the chloroplast can lead to generation of ROS. GUN4 is proposed to bind ProtoIX, shielding its reaction with ROS. In shielding ProtoIX, GUN4 may be progressively modified or degraded with degradation products hypothesised to act as the retrograde signals. (b) A contrasting model for GUN4 - In stead of having a ‘shielding’ effect when bound to ProtoIX, the GUN4-ProtoIX complex appeared to escalate $^1$O$_2$ generation. The elevated $^1$O$_2$ produced by GUN4-ProtoIX may be sensed by a $^1$O$_2$ sensing system (like the Arabidopsis EXECUTER1/EXECUTER2 or EX1/EX2 system) yet to be discovered, that relays a signal to the nucleus.
**Figure 5:** A schematic of a tentative mechanism for CAS activity in *Chlamydomonas* retrograde signalling. A Under low light/high CO$_2$ conditions, CAS is dispersed throughout the chloroplast. B Under high light/low CO$_2$ the ETC proteins become over-reduced, triggering the movement of CAS into the pyrenoid along the pyrenoid tubules. In the Ca$^{2+}$ rich pyrenoid, CAS binds to Ca$^{2+}$ and becomes activated. This form of CAS signals back to the nucleus to modulate target genes. It also induces an increase in intracellular Ca$^{2+}$. 
Figure 6: Relative expression of PR (top) and CCM (bottom) genes in wt and cia5 Chlamydomonas grown in different [CO₂]- <0.02% (V-very low), 0.03-0.05% (L-Low) and 5% (H-High). The expression of only genes classified as being in CCM clusters (Fang et al., 2012) are shown here.