The function and regulation of CIA5/CCM1 in Chlamydomonas reinhardtii

Bo Chen
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Part of the Genetics Commons

Recommended Citation
Chen, Bo, "The function and regulation of CIA5/CCM1 in Chlamydomonas reinhardtii" (2016). Graduate Theses and Dissertations. 15101.
https://lib.dr.iastate.edu/etd/15101
The function and regulation of CIA5/CCM1 in *Chlamydomonas reinhardtii*

by

Bo Chen

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Genetics

Program of Study Committee:
Martin H. Spalding, Major Professor
Steven R. Rodermel
Gururaj A. Rao
Bing Yang
Yanhai Yin

Iowa State University
Ames, Iowa
2016

Copyright © Bo Chen, 2016. All rights reserved.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .................................................................................................................. v

ABBREVIATIONS .......................................................................................................................... vi

CHAPTER 1. GENERAL INTRODUCTION AND LITERATURE REVIEW ...........................................1

1. General Introduction ..................................................................................................................1

2. Literature Review ......................................................................................................................2

2.1 The model organism *Chlamydomonas reinhardtii* .................................................................2

2.2 CO₂ concentrating mechanism (CCM) ....................................................................................3

2.3 Important components of CCMs ............................................................................................5

2.4 The schematic model for *Chlamydomonas* CCM ................................................................12

2.5 *Chlamydomonas* CIA5 ........................................................................................................13

3. Dissertation Organization .........................................................................................................15

References .....................................................................................................................................17

Figure Legend ...............................................................................................................................24

CHAPTER 2. CIA5 ACTIVATION DOMAIN ....................................................................................26

Abstract .......................................................................................................................................27

1. Introduction ...............................................................................................................................27

2. Material and Methods ...............................................................................................................30

2.1 Cell strains and culture conditions .........................................................................................30

2.2 Modified yeast two-hybrid and LacZ assay ..........................................................................30

2.3 *Chlamydomonas* dTALE construction and colorimetric assay ........................................30

2.4 Gene expression analysis ......................................................................................................31

2.5 Western immunoblot analysis ...............................................................................................32

2.6 Expression of protein fragments in *E. coli* .........................................................................32

3. Results .......................................................................................................................................32

3.1 CIA5/CCM1 auto-activation domain identification in yeast ..................................................32

3.2 Activation domain confirmation in *Chlamydomonas* ...........................................................33

3.3 Region responsible for abnormal migration of CIA5 during SDS-PAGE ...............................36

3.4 Region indispensable for complementation of *cia5* mutant .................................................38

4. Conclusion and Discussion ......................................................................................................39

References .....................................................................................................................................41
2.5 Western immunoblot analysis .................................................................105
2.6 Chlorophyll measurement .....................................................................106
3. Results ........................................................................................................106
  3.1 Recombinant mini-CIA5 construct can complement cia5 ..................106
  3.2 Downstream gene expression in mini-CIA5 transformants ...............107
  3.3 Liquid growth experiments .....................................................................108
  3.4 Screening/selection for cia5 suppressors .............................................109
4. Conclusions and Discussion .....................................................................110
References ......................................................................................................113
Figure and Table Legends ............................................................................116
CHAPTER 5. GENERAL SUMMARY .................................................................129
ACKNOWLEDGEMENTS

I would never have been able to finish my dissertation without the guidance of my committee members, help from friends, and support from my family and wife.

I would like to express my deepest gratitude to my advisor Dr. Martin Spalding, for his excellent guidance, caring, patience, and providing me with an excellent atmosphere for doing research, and also providing me the opportunity to achieve significant growth academically and personally. Dr. Spalding’s patience, perseverance, and insightful thoughts have always inspired me as well as the whole Spalding lab.

I also deeply appreciate all the help I got from my committee members Dr. Steven Rodermel, Dr. Aragula Rao, Dr. Bing Yang, and Dr. Yanhai Yin. I would like to express my sincere gratitude for their guidance and critical review of my thesis. Throughout my years of study at ISU, they have always supported me with encouragement and constructive suggestions. I would also like to specifically thank Dr. Yanhai Yin for being incredibly helpful in conducting experiments, and also kindly providing materials and technique supports.

I am very grateful for the whole lab groups, Dr. Yingjun Wang, Dr. David Wright, and Dr. Deqiang Duanmu, Dr. Dan Stessman, Dr. Wei Fang, Soujanya Akella, Alfredo Kono, and many other former Spalding lab members, and especially Dr. David Wright and former members Dr. Yingjun Wang, Dr. Deqiang Duanmu have always been there to kindly answer my millions of questions since I was a new graduate student.

Finally, thanks to my friends and family for their encouragement, support, and their unconditional love as I walked through these long year.
### Abbreviations

Abbreviated terminology used in present work

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>Activation Domain</td>
</tr>
<tr>
<td>BD</td>
<td>Binding Domain</td>
</tr>
<tr>
<td>C3</td>
<td>Three carbon</td>
</tr>
<tr>
<td>C4</td>
<td>Four carbon</td>
</tr>
<tr>
<td>CA</td>
<td>Carbonic Anhydrase</td>
</tr>
<tr>
<td>CAM</td>
<td>Crassulacean Acid Metabolism</td>
</tr>
<tr>
<td>CC-</td>
<td>prefix for strains in the <em>Chlamydomonas</em> Resource Center collection</td>
</tr>
<tr>
<td>CCM</td>
<td>CO₂ (or Carbon) Concentrating Mechanism</td>
</tr>
<tr>
<td>Chl</td>
<td>Chlorophyll</td>
</tr>
<tr>
<td>CO₂</td>
<td>Carbon Dioxide</td>
</tr>
<tr>
<td>Ci</td>
<td>Inorganic carbon</td>
</tr>
<tr>
<td>Cr</td>
<td><em>Chlamydomonas reinhardtii</em></td>
</tr>
<tr>
<td>DIC</td>
<td>Dissolved inorganic carbon</td>
</tr>
<tr>
<td>dTALE</td>
<td>designed Transcription Activator-Like Effector</td>
</tr>
<tr>
<td>EMS</td>
<td>the mutagen ethyl methanesulfonate</td>
</tr>
<tr>
<td>GPD3</td>
<td>Glycerol-3-Phosphate Dehydrogenase-3-Phosphate Reductase</td>
</tr>
<tr>
<td>H-CO₂</td>
<td>High CO₂</td>
</tr>
<tr>
<td>HCO₃⁻</td>
<td>Bicarbonate</td>
</tr>
<tr>
<td>HCR</td>
<td>High CO₂ Requiring</td>
</tr>
<tr>
<td>GMSA</td>
<td>Gel Mobility Shift Assay</td>
</tr>
<tr>
<td>LCI</td>
<td>Limiting-CO₂-inducible</td>
</tr>
<tr>
<td>L-CO₂</td>
<td>Low CO₂ or air level CO₂</td>
</tr>
<tr>
<td>O₂</td>
<td>Molecular Oxygen</td>
</tr>
<tr>
<td>PPM</td>
<td>Parts Per Million</td>
</tr>
<tr>
<td>RBSS</td>
<td>Random Binding Site Selection</td>
</tr>
<tr>
<td>Rubisco</td>
<td>Ribulose-1, 5-bisphosphate carboxylase/oxygenase</td>
</tr>
<tr>
<td>RuBP</td>
<td>Ribulose-1, 5-bisphosphate</td>
</tr>
<tr>
<td>TAP</td>
<td>Tris-acetate-phosphate culture medium</td>
</tr>
<tr>
<td>VL-CO₂</td>
<td>Very low CO₂</td>
</tr>
</tbody>
</table>
CHAPTER 1. GENERAL INTRODUCTION AND LITERATURE REVIEW

1. General Introduction

Carbon dioxide (CO₂) has dual roles in plants and other photosynthetic organisms, serving both as the substrate for photosynthesis and as an important signal to regulate growth and development, so variable CO₂ concentrations affect photosynthesis, growth and productivity of plants and photosynthetic microbes. Although atmospheric CO₂ has increased from a preindustrial concentration of about 280 ppm to a globally averaged concentration of approximately 400 ppm at present (http://CO₂now.org/), this level of CO₂ concentration is still limiting to photoautotrophic growth of most plants and photosynthetic microorganisms, because of the inefficiency of the enzyme Ribulose-1, 5-bisphosphate carboxylase/oxygenase (Rubisco), which has both an oxygenase and a carboxylase function. Rubisco has a low affinity for CO₂, and the high O₂/CO₂ ratio in the atmosphere increases its oxygenase activity, which consumes energy and ultimately releases fixed CO₂ through the wasteful photorespiration pathway (LONG et al. 2004; DUANMU and SPALDING 2011; WANG et al. 2015).

Rubisco is the central enzyme for photosynthetic carbon assimilation in plants, catalyzing the first and most important carboxylation reaction by fixation of CO₂ via the carboxylation of ribulose-1, 5-bisphosphate (RuBP). As an ancient enzyme that evolved in high-CO₂, low-O₂ environments, Rubisco has a slow catalytic rate, low affinity for CO₂ compared with many other enzymes and the oxygenation of RuBP, which is competitive with the carboxylation reaction, make it much less efficient (FANG et al. 2012; WANG et al. 2015).

Many photosynthetic organisms have developed strategies to increase the level of CO₂ at the location site of Rubisco, including the well-known C4 photosynthesis and Crassulacean Acid Metabolism (CAM) pathways in terrestrial vascular plants (NIYOGI et al. 2015). These systems are believed to have evolved several times as mechanisms that adapt organisms to unfavorable growth conditions. Based on recent experimental evidence, another more general CCM system was proposed to present in all plants, a basal CO₂ concentrating mechanism (bCCM) is
hypothesized to be composed of mitochondrial carbonic anhydrases (a β-type carbonic anhydrase and the γ-type carbonic anhydrase) and probably further unknown components. The bCCM is proposed to reduce leakage of CO₂ from plant cells and allow efficient recycling of mitochondrial CO₂ for carbon fixation in chloroplasts (ZABALETA et al. 2012).

In contrast, some C3-type aquatic photosynthetic microorganisms, such as cyanobacteria and microalgae, which also face a challenge of slower diffusion rate for dissolved CO₂ relative to the rate of diffusion in air, induce a different type of CO₂-concentrating mechanism (CCM) when the available inorganic carbon (Ci: CO₂, and HCO₃⁻) for photosynthesis is limited. Much knowledge of the mechanisms of the cyanobacterial and microalgal CCMs has been gained from well-investigated prokaryotic and eukaryotic model organisms such as Synechocystis, Synechococcus (BADGER and PRICE 2003; PRICE et al. 2008; PRICE 2011; PRICE et al. 2013) and Chlamydomonas reinhardtii (herein referred to as Chlamydomonas) (MORONEY and YNALVEZ 2007; SPALDING 2008; SPALDING 2009). When the availability of Ci is limiting, the induced CCM acts via active Ci uptake systems across both the plasma membrane and the chloroplast envelope to increase internal Ci accumulation, and various carbonic anhydrases (CAs) function to dehydrate accumulated HCO₃⁻ and provide elevated internal CO₂ concentrations to the pyrenoid, a Rubisco-containing micro-compartment of the chloroplast (MORONEY and YNALVEZ 2007; SPALDING 2008; WANG et al. 2011). Interestingly, no homologs of cyanobacterial CCM genes have been identified in Chlamydomonas, even though it has a functionally similar CCM, indicating the independent evolution of CCM in these microalgae and cyanobacteria.

2. Literature Review

2.1 The model organism Chlamydomonas reinhardtii

The genus Chlamydomonas (Greek: chlamys, a cloak or mantle; monas: solitary, now used as a generic term for certain unicellular flagellates) was named by C.G. Ehrenberg in 1833 and Chlamydomonas species C. reinhardtii was described in 1888 by Dangeard, and named for Ludwig Reinhard(t) in 1876 (HARRIS et al. 2009).
This single-celled green algae became a more prominent model organism for many fundamental cellular processes, including flagella assembly, motility, DNA methylation, photosynthesis, chloroplast biogenesis, mitochondria characterization, metabolism, sex determination, and recently attracted attention for their utility in industrial applications including biofuel production, because it is unicellular, vegetatively haploid, amenable to tetrad analysis, and has a short life cycle, well-characterized genetics, and a large, available collection of various mutants. In addition, its three genomes are subject to specific transformation, it can grow either photoautotrophically or heterotrophically on acetate, and its genome is sequenced and many molecular tools have been developed for this organism (HARRIS 2001; FUNES et al. 2007; GROSSMAN et al. 2007; MERCHANT et al. 2007; HARRIS et al. 2009; BRENNAN and OWENDE 2010; FLOWERS et al. 2015; WANG et al. 2015).

2.2 CO₂ concentrating mechanism (CCM)

As an aquatic photosynthetic organism, *Chlamydomonas* has to overcome not only low atmospheric CO₂ concentrations but also about 10,000-fold slower diffusion of CO₂ in water relative to air. Rubisco also has a poor apparent affinity for CO₂, with a $K_m$(CO₂) greater than 25 µmol L⁻¹, while in air-equilibrated water, the CO₂ concentration is around 10 µmol L⁻¹, so Rubisco functions at less than 20% of its capacity (MORONEY 2001; MORONEY and YNALVEZ 2007; MUKHERJEE and MORONEY 2011). The efficiency of the CCM may be assessed from the ratio between the intrinsic $K_m$(CO₂) of the particular Rubisco and the apparent photosynthetic $K_{\frac{1}{2}}$ for extracellular dissolved CO₂, and this ratio varies from several thousands in cyanobacteria to close to one in organisms lacking a CCM (KAPLAN and REINHOLD 1999). It has been reported that *Chlamydomonas* cells grown in limiting CO₂ have an apparent photosynthetic $K_{\frac{1}{2}}$ for DIC 10-100 fold lower than that in high CO₂ grown cells (WANG et al. 2015). Such high efficiency can be attributed to its capacity for accumulating a large intracellular Ci pool. Indeed, *Chlamydomonas* has been demonstrated to accumulate intracellular Ci at least 40 to 100-fold over ambient CO₂ levels (DUANMU et al. 2009b; WANG et al. 2015).
Ci active transport and accumulation plays a critical role in the *Chlamydomonas* CCM (Moroney and Ynalvez 2007; Spalding 2008). Internal accumulation of Ci occurs against a large concentration gradient, so accumulation must occur as HCO$_3^-$ because its permeability across lipid membranes is 1000-fold lower than that of CO$_2$ (Wang and Spalding 2014a). However, Rubisco uses CO$_2$ as substrate, so, along with Ci transporters, CAs, which catalyze interconversion of CO$_2$ and HCO$_3^-$, also play important roles in the CCM (Spalding *et al.* 1983a; Coleman and Grossman 1984; Moroney *et al.* 2011).

The *Chlamydomonas* CCM is induced when the CO$_2$ concentration is limited, and one strategy for discovery of CCM-related genes has been based on identifying genes with elevated expression under limiting CO$_2$ (lower than 0.05%) compared with high CO$_2$ (1 to 5% CO$_2$) (Spalding and Jeffrey 1989; Chen *et al.* 1997; Somanchi and Moroney 1999; Miura *et al.* 2004; Yamano and Fukuzawa 2009). Many CAs and putative transporters or other LCI (low CO$_2$ inducible) genes have been discovered using this criterion and have been hypothesized to relate to the CCM of *Chlamydomonas* (Miura *et al.* 2004; Yamano and Fukuzawa 2009).

Recently, at least three distinguishable CO$_2$ concentration-dependent acclimation states were identified in *Chlamydomonas*, including high CO$_2$ (H-CO$_2$, 5-0.5% CO$_2$), low CO$_2$ (L-CO$_2$, 0.4-0.03% CO$_2$), and very low CO$_2$ (VL-CO$_2$, 0.01-0.005% CO$_2$) (Vance and Spalding 2005; Spalding 2008; Dujuanmu *et al.* 2009b). The defining features of these states were determined to be: H-CO$_2$, lack of limiting-CO$_2$-inducible gene expression, photosynthetic $K_{\text{m}}$(CO$_2$) similar to that of the $K_m$(CO$_2$) of Rubisco; L-CO$_2$, induction of limiting-CO$_2$-regulated genes and marked decrease in photosynthetic $K_{\text{m}}$(CO$_2$); VL-CO$_2$, a decreased photosynthetic Vmax and further decreased $K_{\text{m}}$(CO$_2$). The abundance of specific transcripts associated with CO$_2$ limitation showed an increase in both L-CO$_2$ and VL-CO$_2$, suggesting that differences between these two states may arise from a quantitative difference in transcript levels (Vance and Spalding 2005; Spalding 2008; Spalding 2009). Besides the largely decreased growth rate and lower photosynthetic efficiency in VL-CO$_2$ compared with L-CO$_2$ (Spalding 2009), no statistically distinguishable differentially expressed genes were identified in VL-CO$_2$ from L-CO$_2$ by transcriptome analysis (Fang *et al.* 2012).
2.3 Important components of CCMs

In recent years, many molecular components involved in the eukaryotic microalgal CCM have been revealed by genetic/molecular studies and genome-wide gene expression analysis in the model organism *Chlamydomonas* (Spalding 2009). However, a comprehensive picture of how the CCM operates in eukaryotic microalgae is still not well-known due to the lack of detailed functional analysis of many proteins involved in the CCM. Furthermore, many proteins that apparently are essential for the CCM or limiting CO₂ acclimation are novel proteins that lack known functional domains or motifs, making it difficult to precisely decipher their functions (Wang and Spalding 2014b).

The essential components of a CCM in the aquatic organisms include active Ci uptake systems, which transport outside Ci for intracellular HCO₃⁻ accumulation, internal CAs which catalyze rapid interconversion of CO₂ and HCO₃⁻, and the pyrenoid, a specific internal micro-compartment to provide a localized elevated CO₂ concentration for Rubisco utilization (Morney and Ynalvez 2007; Spalding 2008; Spalding 2009; Wang et al. 2011).

a. Putative Ci transporters

Regarding the mechanism of Ci transport and accumulation in the CCM, the first barrier to Ci uptake is the plasma membrane. Two CIA5-regulated genes encoding candidate transporters have been implicated in Ci transport across the plasma membrane: HLA3 and LC11.

High light-induced gene3 (*HLA3*) encodes a putative ATP binding cassette type transporter induced under low CO₂ conditions, and the involvement of HLA3 in active Ci uptake was experimentally confirmed by either efficient RNAi-mediated knockdown in both wild type and mutant *Chlamydomonas* strains (Duanmu et al. 2009a; Wang et al. 2011; Fang et al. 2012) or insertional mutants of *HLA3* and/or *LCIA* (Yamano et al. 2015), When *HLA3* expression was reduced alone, substantial decreases in both photosynthetic Ci affinity and Ci uptake were observed at alkaline pH, at which the predominant form of Ci is HCO₃⁻. When combined with either *LCIB* mutations or
simultaneous knockdown of LCIA mRNA, the cells represented more severe growth defects, together with further decreased photosynthetic Ci affinity and Ci uptake. These results strongly support the hypothesis that HLA3 is directly or indirectly required for HCO$_3^-$ transport into the cells (DUANMU et al. 2009a; YAMANO et al. 2015). HLA3 homologs also were found in several other microalgae, such as Volvox carteri, Chlorella sp. NC64A, and Ostreococcus RCC809 (YAMANO and FUKUZAWA 2009). Individually overexpressed HLA3 or LCIA in high CO$_2$ conditions did not affect Ci affinity, simultaneous overexpression of HLA3 with LCIA significantly increased Ci affinity and accumulation in Chlamydomonas (YAMANO et al. 2015), the similar results were also observed when overexpression HLA3 in Chlamydomonas by using dTALE in high CO$_2$ resulted in increased Ci accumulation specifically under very low CO$_2$ conditions, which confirmed that HLA3 indeed involved in Ci uptake, and it might associate with HCO$_3^-$ transport under very low CO$_2$ conditions in which active Ci uptake was highly limited (GAO et al. 2015). But further biochemical characterization of HLA3 is still needed to elucidate the mechanism of this putative plasma membrane HCO$_3^-$ transporter.

LCI1 was first identified as a limiting CO$_2$ inducible gene (BUROW et al. 1996), it was reported to increase Ci uptake and mainly localized to the plasma membrane when expressed transgenically in lcr1 mutant (OHNISHI et al. 2010). The expression of LCI1 is under control of both CIA5/CCM1 and a MYB type transcription factor LCR1 (MIURA et al. 2004; YOSHIOKA et al. 2004). LCI1 also has no identifiable homologues or recognizable domains, other than predicted transmembrane domains, so it appears to represent a unique membrane protein. Because it is CIA5-regulated and mainly plasma membrane localization, it is a possible Ci transport candidate, though little else is known about this gene or its product (SPALDING 2009; OHNISHI et al. 2010; WANG et al. 2015). The LCI1-RNA interference (LCI1-RNAi) strains with reduced expression of LCI1 showed reduced growth on low CO$_2$ (Mason and Moroney, unpublished observations), but the physiological role of LCI1 remains to be determined (MORONEY and YNALVEZ 2007).

To be effective in the CCM, all Ci uptake must ultimately concentrate in chloroplasts, and must include active Ci uptake across the chloroplast envelope.
Some chloroplast envelope proteins are candidates to transport Ci into the chloroplast stroma, including LCIA (NAR1.2), CCP1, and CCP2 (Spalding 2008; Spalding 2009; Wang et al. 2011; Wang et al. 2015). The low CO₂-induced gene A (LCIA [NAR1.2]), which encodes a Formate/Nitrite transporter family protein targeted to the chloroplast envelope, is induced in limiting CO₂ and requires CIA5 for expression (Miura et al. 2004). LCIA has been reported to increase HCO₃⁻ transport when transfected into Xenopus laevis oocytes (Mariscal et al. 2006), and it has also been shown that LCIA and HLA3 co-knockdown in wild type Chlamydomonas cells gave more severe phenotypic impact than HLA3 knockdown alone, and simultaneous overexpression of HLA3 with LCIA significantly increased Ci affinity and accumulation, suggesting that LCIA and HLA3 are key cooperative or complementary components of the active Ci transport pathway in limiting Ci acclimated cells (Duanmu et al. 2009a; Wang et al. 2011; Wang et al. 2015; Yamano et al. 2015).

RNAi knockdown of chloroplast carrier protein 1 (CCP1) and CCP2 (Pollock et al. 2004), which encode nearly identical, LCI membrane proteins (Spalding and Jeffrey 1989; Chen et al. 1997) reportedly located in the chloroplast envelope (Ramazanov et al. 1993), resulted in poor growth under low CO₂ conditions, although no direct evidence for a defect in Ci transport or photosynthesis was demonstrated. The low CO₂ condition which was used to screen CCP1/CCP2 RNAi transformants was 35 ppm CO₂ (Pollock et al. 2004), such low CO₂ concentration is actually very low CO₂ conditions instead of low CO₂, and high-CO₂ requiring CCP1/CCP2 RNAi transformants showed dramatically decreased not only in growth but also in the message abundance of both transcripts and proteins under such very low CO₂ conditions. While RNAi silenced strains showed normal rates of photosynthesis at all tested concentrations of DIC, and over a range of pHs which was measured under air-level CO₂ conditions (low CO₂) (Pollock et al. 2004). We may argue that CCP1/CCP2 might function under very low CO₂ instead of air-level CO₂ conditions. Also, although long considered to be chloroplast envelope proteins, CCP1 and CCP2 were recently reported to be located in the mitochondrial membrane instead (Atkinson et al. 2015).

It was also suggested that these two putative transporters may be involved with transport of metabolic intermediates important in limiting CO₂ acclimation, either in the
chloroplast or the mitochondrion. It is also possible that, as was observed with the HLA3 RNAi knockdown or HLA3 insertional mutant, which exhibited no clear phenotype unless paired with decreased function of another transporter or of a Ci accumulation system (DUANMU et al. 2009a; YAMANO et al. 2015), any Ci transporter functions of CCP1/CCP2 may have been masked by some other Ci transport or accumulation systems under the conditions in which the Ci affinity was tested (WANG et al. 2011; WANG et al. 2015).

b. Carbonic anhydrase (CA)

The combined transport of HCO$_3^-$ across both the plasma membrane and the chloroplast envelope results in the accumulation of HCO$_3^-$ in the chloroplast stroma. Rubisco, which is located in the pyrenoid, uses CO$_2$ as its substrate. CO$_2$ and HCO$_3^-$ are the two major Ci species in aquatic environments, and the ratio of their concentrations in solution is strongly dependent on pH. *Chlamydomonas* maintains an average cytosolic pH of 7.1 (MESSERLI et al. 2005), while, in the light, the pH of the chloroplast stroma and the thylakoid lumen are about 8 and between 4 and 5, respectively (MORONEY and YNALVEZ 2007). This pH differential is significant because CO$_2$ is a weak acid with a pKa of about 6.3 (HCO$_3^-$ + H$^+$ ⇌ CO$_2$ + H$_2$O). At equilibrium under these conditions, HCO$_3^-$ would be the predominant Ci species in the chloroplast stroma, while CO$_2$ would be the most abundant form of Ci in the thylakoid lumen (MORONEY and YNALVEZ 2007). HCO$_3^-$ becomes more predominant over CO$_2$ at neutral to alkaline pHs, although spontaneous conversion between Ci species is very slow on a physiological time scale. Both supplying substrate CO$_2$ to Rubisco from an accumulated, stromal HCO$_3^-$ pool and preventing depletion of HCO$_3^-$ or CO$_2$ at the cell surface during rapid Ci influx driven by active HCO$_3^-$ transport or active CO$_2$ uptake, respectively, arguably may require a CAs to maintain the supply of the relevant Ci species via rapid equilibration between CO$_2$ and HCO$_3^-$.

So far at least 12 CA isoforms or predicted CAs have been identified in *Chlamydomonas*, including all three, distinct evolutionary classes: 3 α-CAs, 6 β-CAs, and 3 γ- (or γ-like) CAs (SPALDING 2009; MORONEY et al. 2011). Some of these proteins
are proposed to have possible roles in the CCM, but the exact functions of most of them remain unknown.

The first essential CA of the *Chlamydomonas* CCM identified was CAH3, a thylakoid lumen CA responsible for dehydration of HCO$_3^-$ actively accumulated internally, probably in the stroma (FUNKE et al. 1997; KARLSSON et al. 1998). CAH3 mutants show leaky or poor growth under air level CO$_2$, cannot survive under very low CO$_2$ conditions and over-accumulate internal Ci in a form of HCO$_3^-$, which is unavailable for photosynthesis (SPALDING et al. 1983a; FUNKE et al. 1997; KARLSSON et al. 1998; HANSON et al. 2003). This α-CA is constitutively expressed, but modestly up-regulated in limiting CO$_2$ conditions, and may be associated with PS II particles (KARLSSON et al. 1998; PARK et al. 1999; BENLOCH et al. 2015). The localization of CAH3 in the thylakoid lumen has been confirmed (SHUTOVA et al. 2008; SINETOVA et al. 2012; BENLOCH et al. 2015) where it can take advantage of the acidic lumen environment to drive nearly complete conversion of HCO$_3^-$ to CO$_2$ (SPALDING 2008; DUANMU et al. 2009b; MORONEY et al. 2011; FANG et al. 2012), but its location and the possible association with PS II have led to two alternative hypotheses with regard to its primary function. One is that CAH3 supplies Rubisco with CO$_2$ from dehydration of the accumulated HCO$_3^-$, and the other involves PS II function or proton gradient formation across the thylakoid membranes. However, a recent report appears to have resolved these contradictory functions by demonstrating that CAH3 is phosphorylated and re-localizes from PSII particles to pyrenoidal thylakoid tubules in low CO$_2$ conditions (BLANCO-RIVERO et al. 2012). Mutants defective in the thylakoid lumen CA, CAH3, suppress the LCIB mutation phenotype (DUANMU et al. 2009b), suggesting a role for LCIB and LCIC, in which LCIB forms a heteromeric complex (YAMANO et al. 2010; WANG et al. 2011), and functions downstream of CAH3 in preventing the leakage of CO$_2$ from the stroma (DUANMU et al. 2009b).

CAH6, a putative chloroplast stromal CA, may be involved in CO$_2$-to-HCO$_3^-$ conversion in the alkaline stroma, maintaining a high Ci concentration or helping the pyrenoid surrounded LCIB/LCIC complex to recapture leaked CO$_2$ in *Chlamydomonas* (MITRA et al. 2005; MORONEY and YNALVEZ 2007; WANG et al. 2015).
A periplasmic CA, CAH1 has been extensively studied as a limiting-CO₂-inducible protein (Coleman and Grossman 1984; Van and Spalding 1999; Yoshioka et al. 2004; Wang et al. 2015), and might function in maintaining HCO₃⁻ for plasma membrane HCO₃⁻ transporters or CO₂ for active CO₂ uptake. However, no growth defects were observed in a CAH1 mutant (cah1), other than slightly decreased photosynthetic Ci affinity in very low CO₂ conditions (Van and Spalding 1999). It is possible that there are multiple complementary Ci uptake systems that function in conditions of limiting CO₂ environment (Wang and Spalding 2014a; Wang et al. 2015) and CAH1 function might be unmasked if combined with other defective Ci uptake systems.

Highly induced mitochondrial CAs CAH4 and CAH5 might help CO₂-to-HCO₃⁻ conversion to reduce leakage of CO₂ from plant cells and allow efficient recycling of mitochondrial CO₂ for carbon fixation in chloroplasts (Zabaleta et al. 2012). Because of the high-level limiting-CO₂ induction of these two mitochondrial CAs and the recently reported mitochondrial localization of the two strongly limiting-CO₂ induced CCP1 and CCP2 membrane transporters (Atkinson et al. 2015), mitochondria may play a significant role in limiting CO₂ acclimation responses.

c. The Pyrenoid

The pyrenoid of Chlamydomonas is penetrated by a network of thylakoid tubules where one important CAs, CAH3, has been localized to the thylakoid lumen, especially in the thylakoid tubules inside the pyrenoid in cells acclimated to low CO₂ and very low CO₂ conditions (Karlsson et al. 1998; Mitra et al. 2005; Shutova et al. 2008; Duanmu et al. 2009b; Blanco-Rivero et al. 2012; Sinetova et al. 2012; Benlloch et al. 2015; Engel et al. 2015). With the help of CAH3, the accumulated HCO₃⁻ in the thylakoid lumen is rapidly dehydrated and release CO₂ into the pyrenoid, a Rubisco-containing internal compartment of the chloroplast, for assimilation by Rubisco (Borkhensive et al. 1998; Price et al. 2002; Spalding et al. 2002).

The pyrenoid is a very large, electron-dense, proteinaceous structure within the chloroplasts of most algae and only found in lower plants such as hornworts (Ramazanov et al. 1994; Rawat et al. 1996; Morita et al. 1997; Ma et al. 2011;
SINETOVA et al. 2012; ENGEL et al. 2015). In the pyrenoid of Chlamydomonas, Rubisco is the predominant protein (MORITA et al. 1997; BORKHSENIOUS et al. 1998), and nonsense mutations in the Rubisco large subunit totally abolish the formation of the pyrenoid (RAWAT et al. 1996). Immuno-localization studies with Chlamydomonas show that the pyrenoid undergoes a dramatic morphological change when cells are switched from high- to low- CO₂ conditions (RAWAT et al. 1996; MORONEY and YNALVEZ 2007). When the CCM is functional, development of the starch sheath surrounding the chloroplast pyrenoid is increased and almost all of Rubisco is present in the pyrenoid (RAMAZANOV et al. 1994; MORITA et al. 1997; BORKHSENIOUS et al. 1998; MORONEY and YNALVEZ 2007; MUKHERJEE and MORONEY 2011). The observed increase in pyrenoid starch is consistent with limiting CO₂ up-regulated expression of STA2 and STA3 (WATTEBLED et al. 2002; MIURA et al. 2004; HARRIS et al. 2009), a general correlation that also was reported in Chlorella kessleri (OYAMA et al. 2006). Once proposed to be functionally important as a diffusion barrier to minimize loss of CO₂ from the Rubisco-rich pyrenoid (BADGER and PRICE 1994), the absence of any detectable decrease in CCM functionality in a starchless mutant argues against this hypothesis (VILLAREJO et al. 1996).

A Chlamydomonas mutant, cia6, was identified as unable to grow in low CO₂ and having a highly disorganized pyrenoid under electron microscopy. Complementation of the mutant with a protein methylase restored not only the pyrenoid but also the ability to grow under low-CO₂ conditions, and the ability to actively concentrate Ci. These results indicate that CIA6 is required for formation of the pyrenoid and further supported the pyrenoid as required for a functional CCM in Chlamydomonas (MA et al. 2011). Together we may conclude that the possible role of the pyrenoid is to provide a micro-compartment for CO₂ to be generated, providing a localized elevated CO₂ concentration for Rubisco utilization.

d. Other important components

In addition, a low CO₂ induced protein, LCIB, together with another low CO₂ induced protein LCIC, forming the LCIB-LCIC complex, is localized around the pyrenoid (YAMANO et al. 2010; WANG and SPALDING 2014b; WANG and SPALDING
2014a) when the cells were acclimated to very low CO2 conditions. Mutants with defects in LCIB expression result in an unusual air-dier phenotype, which fails to accumulate internal Ci and dies at air level of CO2 but survives under very low CO2 (Spalding et al. 1983b; Wang and Spalding 2006; Duanmu et al. 2009b). Mutants that lack CAH3 suppress the LCIB mutant air-dier phenotype, which indicates that LCIB functions downstream of CAH3 in the CCM (Duanmu et al. 2009b), LCIB may participate in trapping stromal CO2 released from the thylakoid into the stromal HCO3- pool, and LCIB might also function to capture the external CO2 diffusing into chloroplasts and maintain the intracellular Ci accumulation (Duanmu et al. 2009b; Wang and Spalding 2014b).

The identification of an LCIA-LCIB double mutant based on its inability to survive in very low CO2 suggests that both LCIA (also named NAR1.2) and LCIB are critical for survival in very low CO2. The contrasting effects of individual mutations in LCIB and LCIA compared with the effects of LCIB-LCIA double mutations on growth and Ci-dependent photosynthetic O2 evolution reveal distinct roles of LCIA and LCIB in the CCM (Wang and Spalding 2014a).

2.4 The schematic model for Chlamydomonas CCM

As shown in Fig.1, the hypothetical schematic model illustrated here for acclimation to limiting CO2 in Chlamydomonas with emphasis on Ci, transport, and relationships of limiting Ci responsive genes (Mukherjee and Moroney 2011; Wang et al. 2011; Zabaleta et al. 2012; Wang and Spalding 2014a; Wang et al. 2015). The proposed model illustrates the transport of Ci from the extracellular periplasmic space to the thylakoid lumen via the cytosol and chloroplast stroma. While possibly interconverted by periplasmic CA (CAH1), CO2 and HCO3- pass through plasma membrane via diffusion and active transport (possibly by HLA3/LCI1), respectively. Similarly, upon arrival in the cytosol, CO2 is able to diffuse across the chloroplast envelope, while HCO3- has to be transported, probably with the help of LCIA, into chloroplast stroma, where it accumulates as HCO3-. CO2 diffusing across the chloroplast envelope also accumulates as HCO3-, possibly after hydration catalyzed by CAH6 and/or the LCIB/LCIC complex. HCO3- accumulated in the chloroplast
stroma then must cross the thylakoid membrane via a predicted but not identified HCO$_3^-$ permease or possible transporter into the pyrenoidal tubule thylakoid lumen to be quickly converted by CAH3 to CO$_2$, which then diffuses into pyrenoid as the substrate for Rubisco utilization.

Any CO$_2$ not consumed by Rubisco and thus leaked from the pyrenoid is proposed to be trapped, or uni-directionally converted to HCO$_3^-$, by the stroma-localized LCIB/LCIC complex, possibly including involvement of the putative stromal CA, CAH6, and returned to the stromal HCO$_3^-$ pool for re-circulation via CAH3 within the lumen of the pyrenoid-associated thylakoid tubules for dehydration to CO$_2$.

Mitochondrial CAs (CAH4/CAH5) are proposed to reduce leakage of CO$_2$ from plant cells and allow efficient recycling of mitochondrial CO$_2$ for carbon fixation in chloroplasts (ZABAleta et al. 2012). This suggested model is supported by the re-localization of Rubisco, LCIB/LCIC complex and CAH3 upon acclimation from high CO$_2$ to limited CO$_2$ conditions, and part of the CAH3 protein is quickly phosphorylated and transferred to the thylakoids that intrude into the pyrenoid, where most of the Rubisco is found.

2.5 Chlamydomonas CIA5

Even though the Chlamydomonas CCM has been extensively studied in recent years, we still know little about the limiting CO$_2$ acclimation process, the acclimation to limiting CO$_2$ and induction of the CCM. In Chlamydomonas, CCM induction appears to be regulated by CIA5 (or CCM1) (Miura et al. 2004; Wang et al. 2005; Fang et al. 2012). A UV- induced mutant, cia5, was first identified as a slow growing mutant in limiting CO$_2$ conditions (Moroney et al. 1989). This mutant can grow under high CO$_2$ and grows more slowly than the wild type in low CO$_2$, but it cannot grow in very low CO$_2$ and appears to completely lack the changes in gene expression associated with acclimation to low CO$_2$ (Spalding 2009). Further studies show that almost all identified Low CO$_2$ Inducible (LCI) genes remain uninduced, fail to be up-regulated or decreased when cia5 mutant is exposed to limiting CO$_2$ (Moroney et al. 1989; Spalding et al. 2002; Miura et al. 2004; Fang et al. 2012). The defective gene in cia5, CIA5, also known as CCM1, yields two mRNAs, CCM1-A and CCM1-B, from a single
gene by an alternative splicing between the 3rd and 4th exons (only 2 aa difference, 183rd GR substituted to 183rd E. CCM1-A: 699 aa; CCM1-B also named CIA5: 698 aa) (FUKUZAWA et al. 2001; XIANG et al. 2001). Expression of CIA5 in Chlamydomonas is constitutive and stable but only at a very low abundance and independent of CO2 concentration levels, so some type of posttranslational activation of CIA5 in low CO2 apparently is required for CIA5 to regulate downstream gene expression (XIANG et al. 2001; MIURA et al. 2004; WANG et al. 2005; FANG et al. 2012).

Most Ci transporter candidate genes and CA genes are under control of CIA5/CCM1, a transcription regulator coordinating expression of almost all low-CO2 induced genes (MORONEY et al. 1989; FUKUZAWA et al. 2001; XIANG et al. 2001; MIURA et al. 2004; FANG et al. 2012). Because of the extensive connection of CIA5 to the regulation of most CCM-related gene expression, including regulation of all candidate Ci transporters, CIA5 is often called the “master regulator” of the CCM (FANG et al. 2012). The 76 kD CIA5 protein migrates as an approximately 90-100 kD protein during SDS-PAGE (WANG et al. 2005; KOHINATA et al. 2008). This suggests either the aberrant migration or potential posttranslational modification of CIA5.

Aside from being a critical upstream regulator of the CCM, the details of CIA5 function remain undiscovered. We know very little about the genes, if any, that it directly regulates downstream (FANG et al. 2012), let alone anything about sequences that might be recognized by its putative DNA binding domain. The transcriptome comparison by Fang et al. (FANG et al. 2012), identified a massive impact of CIA5/CCM1 and CO2 on the transcriptome and revealed an array of gene clusters with distinctive expression patterns that provide insight into the regulatory interaction between CIA5 and CO2. Individual gene clusters responded primarily to CIA5, to CO2, or to an interaction between the two (FANG et al. 2012).

CIA5/CCM1 has been proposed to be a transcription factor. It has two non-typical zinc finger motifs within an N-terminal domain (SOMANCHI and MORONEY 1999), which may represent a DNA-binding motif typical of transcription factors. It was confirmed that 2 mol zinc can bind to the N-terminal zinc-finger domain, and if this protein is mutated, H54Y, C77V and C80V, the zinc binding ability is lost (KOHINATA et
al. 2008). CIA5 also contains a Gln-rich repeat, which might participate in regulation of other eukaryotic transcription factors, and a Gly-rich region with unknown function.

Generally, zinc-binding domains, including zinc finger motifs have been shown to function as either sequence-specific nucleic acid recognition motifs or as recognition motifs for protein-protein interactions (KOHINATA et al. 2008). It was assumed that the CCM1 protein could be a DNA-binding protein, since CCM1 impacts expression of so many genes and is located in the nucleus (KOHINATA et al. 2008). However, the zinc finger motif was tested in southwestern blot analyses by using the anti-CCM1 antibody, but no signals were detected on the DNA-protein complexes, suggesting that CCM1 may bind to some unidentified proteins rather than to DNA in the nucleus (KOHINATA et al. 2008).

As a master regulator of the CCM, the detailed function and regulation of CIA5/CCM1 still remain unclear. The research described in this dissertation was undertaken to better characterize this CCM master regulator and to answer the following questions: Does CIA5/CCM1 have activation activity? Are any genes directly regulated by CIA5/CCM1? Do any DNA fragments interact with this regulator? Which CIA5 domains are needed to make a fully functional CIA5?

3. Dissertation Organization

This dissertation is written in a format that includes currently unpublished work to demonstrate that CIA5, the master CCM transcription regulator of the Chlamydomonas CCM, has transcription activation activity either in yeast or in Chlamydomonas, and has an acidic activation domain that also is responsible for abnormal SDS-PAGE migration (chapter 2). Chapter 3 describes the first successful overexpression and purification of CIA5 protein from E. coli., which was used to demonstrate DNA binding ability of CIA5/CCM1 in vitro using random binding site selection and gel mobility shift assays, and Chapter 4 reports the development of a mini-CIA5 construct, which only contains a partial CIA5 CDS (nt 1 - 330 plus nt 1282 - 1674). This mini-CIA5, when transformed into cia5, was able to complement the cia5 lethal phenotype in very low CO2. However, detailed assessment of gene expression in the mini-CIA5 complemented cia5 suggested that other CIA5 sequences missing
from mini-CIA5 may also be important for a fully functional CIA5. Screening or selection for cia5 suppressors by insertional mutagenesis, UV mutagenesis or EMS mutagenesis identified no candidate colonies other than cia5 revertants that converted the point mutation back to its wild-type form. The inability to identify second-site suppressors further supports and strengthens the significant role of CIA5 in the Chlamydomonas CCM. The dissertation is then concluded with a general summary of these new contributions to our understanding of CIA5 function in Chapter 5.

The data from the chapter 2 Fig.8 performed by K.L. and chapter 2 Fig. 9 performed by T.P., D.D. and Y.W. performed the initial yeast two-hybrid cDNA library screening by using the full length CIA5 CDS. All other experiments from the dissertation were designed by B.C. and M.H.S., B.C. performed all the researches, analyzed data; and wrote the dissertation. M.H.S. edited the whole dissertation and D.P.W. kindly helped to edit the chapter 2.
References


Duanmu, D., A. R. Miller, K. M. Horken, D. P. Weeks and M. H. Spalding, 2009a Knockdown of limiting-CO₂-induced gene HLA3 decreases HCO₃⁻ transport
and photosynthetic Ci affinity in *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA **106**: 5990-5995.


Funke, R. P., J. L. K Kovar and D. P. Weeks, 1997 Intracellular carbonic anhydrase is essential to photosynthesis in *Chlamydomonas reinhardtii* at atmospheric levels of CO2. Demonstration via genomic complementation of the high-CO2-requiring mutant ca-1. Plant Physiol **114**: 237-244.


Hanson, D. T., L. A. Franklin, G. Samuelsson and M. R. Badger, 2003 The *Chlamydomonas reinhardtii cia3* mutant lacking a thylakoid lumen-localized
carbonic anhydrase is limited by CO₂ supply to Rubisco and not photosystem II function in vivo. Plant Physiol 132: 2267-2275.


SOMANCHI, A., and J. V. MORONEY, 1999 As *Chlamydomonas reinhardtii* acclimates to low-CO2 conditions there is an increase in cyclophilin expression. Plant Mol Biol 40: 1055-1062.


WANG, Y. J., and M. H. SPALDING, 2006 An inorganic carbon transport system responsible for acclimation specific to air levels of CO₂ in *Chlamydomonas reinhardii*. Proc Natl Acad Sci USA 103: 10110-10115.


YOSHIOKA, S., F. TANIGUCHI, K. MIURA, T. INOUE, T. YAMANO et al., 2004 The novel Myb transcription factor LCR1 regulates the CO$_2$-responsive gene Cah1, encoding a periplasmic carbonic anhydrase in *Chlamydomonas reinhardtii*. Plant Cell 16: 1466-1477.

Figure Legend

**Fig. 1.** Schematic, hypothetical model for CCM function in *Chlamydomonas* with emphasis on active Ci uptake and transport adapted from Mukherjee and Moroney (Mukherjee and Moroney 2011; Wang et al. 2011; Zabaleta et al. 2012; Wang and Spalding 2014a; Wang et al. 2015). The proposed transport of Ci from the extracellular space to the thylakoid lumen via the cytosol and stroma is illustrated. CAs catalyze interconversion of CO$_2$ and HCO$_3^-$ in respective compartments. The non-membrane-bound subcellular structure, the pyrenoid, (surrounded by starch sheath) is within the stroma and is the site of Rubisco localization. Blue and Green ovals indicate confirmed and putative HCO$_3^-$ transporters, including HLA3, LCI1 (plasma membrane, PM) and LCIA, CCP1/CCP2 (Chloroplast Envelope, ChE). Additional potential HCO$_3^-$ transporters or HCO$_3^-$ facilitated diffusion channels are indicated with question marks. Stroma localized LCIB/LCIC complex is proposed to function in hydration of incoming CO$_2$ and rehydration of CO$_2$ released from thylakoid lumen into the HCO$_3^-$ pool in the stroma, possibly including involvement of the putative stromal CA, CAH6. Green dashed lines represent the hypothesized function of the LCIB-LCIC complex in conversion of CO$_2$ to HCO$_3^-$, which is then delivered to CAH3 within the lumen of the pyrenoid-associated thylakoid tubules for dehydration to CO$_2$. Mitochondrial CAs (CAH4/CAH5) are proposed to reduce leakage of CO$_2$ from cells and allow efficient recycling of mitochondrial CO$_2$ for carbon fixation in chloroplasts.
Fig. 1
CHAPTER 2. CIA5 ACTIVATION DOMAIN

The activation domain of CIA5/CCM1 in *Chlamydomonas reinhardtii*

Bo Chen¹, Kwanghong Lee², Tom Plucinak², Deqiang Duanmu¹, Yingjun Wang¹, Donald P. Weeks² and Martin H. Spalding¹a

1. Department of Genetics, Development, and Cell Biology, Iowa State University, Ames, IA 50011
2. Department of Biochemistry, University of Nebraska, Lincoln, NE 68588-0664

a: Corresponding Author:

Martin H. Spalding
Telephone: 515 294-1749
FAX: 515 294-1303
Email: mspaldin@iastate.edu
Abstract

Acclimation to limiting CO$_2$ conditions and induction of the CO$_2$ concentrating mechanism (CCM) in *Chlamydomonas reinhardtii* appear to be regulated by the transcription regulator, CIA5 (or CCM1), the so-called “master regulator” of the CCM. Mutants lacking CIA5 function, including the mutant *cia5*, completely lack CCM induction, although they are viable under high CO$_2$ conditions and grow more slowly than wild type in air level CO$_2$. In *cia5*, most identified LCI (low CO$_2$ inducible) genes remain either uninduced by or fail to be up-regulated by limiting CO$_2$. Using a yeast two hybrid system, we identified a 109aa region (436aa - 544aa) of CIA5 and a similar, highly conserved, 130aa region (428aa - 557aa) in the *Volvox* CCM1 protein that both exhibit auto-activation in yeast. We further demonstrated that a region overlapping this highly conserved auto-activation domain causes the CIA5 protein to exhibit abnormal migration in SDS-PAGE. When divided in half, either half of the 130aa region supported auto-activation in yeast. However, deletion of the first half of this conserved domain rendered CIA5 unable to complement the *cia5* mutant. Therefore, this 130aa region may represent two independent activation domains (428aa - 487aa and 488aa - 557aa), both of which are required for CIA5-mediated gene activation. These putative CIA5/CCM1 activation domains were demonstrated to activate expression of a targeted gene in *Chlamydomonas* when used to replace the activation domain of a designed transcription activator-like element (dTALE). This is the first demonstration that CIA5, the master CCM transcription regulator of *Chlamydomonas*, has transcription activation activity both in yeast and in *Chlamydomonas*, and confirms that this acidic activation domain also is responsible for abnormal migration of CIA5 during SDS-PAGE.

**Keywords:** *Chlamydomonas*; CO$_2$ concentrating mechanism; CIA5/CCM1; activation domain; dTALE; targeted gene activation

1. Introduction

Carbon dioxide (CO$_2$) serves both as the substrate for photosynthesis and as an important signal to regulate plant growth and development. Variable CO$_2$
concentrations have effects on photosynthesis, growth and productivity of plants and other photosynthetic organisms. Although atmospheric CO$_2$ has increased from a preindustrial concentration of about 280 ppm to a globally averaged concentration of approximately 400 ppm at present (http://co2now.org/), this CO$_2$ concentration is still limiting to photoautotrophic growth of most plants and photosynthetic microorganisms, because of the inefficiency of the enzyme Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco), which has both an oxygenase as well as a carboxylase function, as well as a low affinity for CO$_2$ (LONG et al. 2004; DUANMU and SPALDING 2011).

A number of photosynthetic organisms have developed strategies to increase the level of CO$_2$ around Rubisco to improve its function, including the well-known C4 photosynthesis and Crassulacean Acid Metabolism (CAM) in terrestrial vascular plants (NIYOGI et al. 2015). In contrast, some aquatic photosynthetic microorganisms accomplish this through induction of a different type of CO$_2$-concentrating mechanism (CCM) when photosynthesis is limited by dissolved inorganic carbon (Ci; CO$_2$, and HCO$_3^-$). Much knowledge of the mechanisms of the microalgal CCM has been gained from well-investigated eukaryotic the microalga, Chlamydomonas reinhardtii (MORONEY and YNALVEZ 2007; SPALDING 2008) (hereafter, Chlamydomonas). When the Ci is limiting, the induced CCM acts via active Ci uptake systems on both the plasma membrane and the chloroplast envelope to increase internal Ci accumulation, and various carbonic anhydrases function to dehydrate accumulated HCO$_3^-$ and provide elevated internal CO$_2$ concentrations to Rubisco (MORONEY and YNALVEZ 2007; SPALDING 2008; WANG et al. 2011).

Even though the Chlamydomonas CCM has been extensively studied in recent years, we still know little about the limiting CO$_2$ acclimation process, the acclimation to limiting CO$_2$ and induction of the CCM. In Chlamydomonas, CCM induction appears to be regulated by the master regulator, CIA5 (or CCM1) (MIURA et al. 2004; WANG et al. 2005). A UV- induced mutant, cia5, was first identified as a slow growing mutant in limiting CO$_2$ (MORONEY et al. 1989). This mutant can grow under high CO$_2$ and grows more slowly than the wild type in low CO$_2$, but it cannot grow in very low CO$_2$ and appears to completely lack the changes in gene expression associated with
acclimation to low CO$_2$ (SPALDING 2009). Further studies show that almost all identified Low CO$_2$ Inducible (LCI) genes remain uninduced, fail to be up-regulated or decreased when $cia5$ mutant is exposed to limiting CO$_2$ (MORONEY et al. 1989; SPALDING et al. 2002; SPALDING 2009; FANG et al. 2012). The defective gene in $cia5$, CIA5 (also known as $CCM1$), was identified independently by two research groups as encoding a predicted 698 amino acid hydrophilic protein (FUKUZAWA et al. 2001; XIANG et al. 2001).

CIA5/CCM1 has been proposed to be a putative transcription factor or transcription co-activator; it has two non-typical zinc finger domains (SOMANCHI and MORONEY 1999), which may comprise a DNA-binding motif typical of transcription factors. It was confirmed that 2 mol zincs can bind to the N-terminal domain, and that, if mutated to perturb putative zinc binding sites, the zinc binding ability was lost (KOHINATA et al. 2008). CIA5 also contains a Gln-rich repeat that may allow interactions with and/or activation of other transcription factors, and a Gly-rich region, but the function of these regions has not been experimentally determined. The 76 kD CIA5 protein that regulates the Chlamydomonas CCM migrates as an approximately 90-100 kD during SDS PAGE (WANG et al. 2005; KOHINATA et al. 2008). This suggested either aberrant migration of CIA5 or potential posttranslational modification of CIA5.

Aside from being a critical upstream regulator of the CCM and other low CO$_2$ acclimation responses (and likely requiring posttranslational modifications during high and low CO$_2$ concentration transitions), the details of CIA5 function remain undiscovered. We know very little about sequences recognized by its putative DNA binding domain or the genes it directly regulates downstream (FANG et al. 2012). An experiment exploring the impact of the $cia5$ mutation on the transcriptome identified a massive impact of CIA5/CCM1 and CO$_2$ on the transcriptome and revealed an array of gene clusters with distinctive expression patterns that provide insight into the regulatory interaction between CIA5 and CO$_2$ (BRUEGGE et al. 2012; FANG et al. 2012). Individual gene clusters responded primarily to CIA5, to CO$_2$, or to an interaction between the two (FANG et al. 2012). In this study, we found that CIA5 has transcription activation activity not only in yeast but also in Chlamydomonas, and
confirmed that this acidic activation domain also is responsible for abnormal SDS-PAGE migration of CIA5.

2. Material and Methods

2.1 Cell strains and culture conditions

*Chlamydomonas* strain CW10 (cc849) was obtained from the *Chlamydomonas* Stock Resource (University of Minnesota, St. Paul, MN). Media and growth conditions for *Chlamydomonas* strains have been previously described (Wang and Spalding 2006). All strains were maintained on CO$_2$ minimal plates and kept in high CO$_2$ (air enriched with 5% vol/vol CO$_2$) chambers at room temperature, under continuous illumination (50 μmol photons m$^{-2}$ s$^{-1}$). Liquid cultures for experimental use were grown on a gyratory shaker (180 rpm) under aeration in white light (approximately 100 μmol photons m$^{-2}$ s$^{-1}$). For experiments in which cells were shifted from high to limiting CO$_2$ (low CO$_2$, 350-400 ppm or very low CO$_2$, 50-100 ppm) conditions, cells were cultured in CO$_2$ minimal medium aerated with 5% CO$_2$ to a density of $\approx 2 \times 10^6$ cells/ml and then shifted to aeration with the appropriate limiting CO$_2$ for various times. Very low CO$_2$ was obtained by mixing normal air with compressed, CO$_2$-free air.

2.2 Modified yeast two-hybrid and LacZ assay

The yeast two-hybrid procedure was performed in accordance with the protocol for the HybriZAP-2.1 XR library construction kit (Stratagene). The *EcoRI*-SalI fragments containing the full length coding DNA sequence (CDS) or partial CDS of *CIA5* were subcloned into pGBKT7 (Clontech) to make bait constructs. Yeast (AH109) cells transformed with indicated constructs were grown in SD media lacking Trp or Trp and His. Putative positive clones were used for LacZ assays using the X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside; Sigma) colony-lift filter assay performed following the manufacturer’s instructions (Clontech).

2.3 *Chlamydomonas* dTALE construction and colorimetric assay
The dTALE was constructed as previously described using an altered TAL plasmid pSKavrXa7delta and a *Chlamydomonas* expression vector pDW2177 (GAO *et al.* 2014). Only the pDW2177ARS1 construct was used for this study. This dTALE construct was cut with *Bbv*CI and *Xba*I to remove the activation domain (ZHU *et al.* 1998), and CIA5 activation domain (AD) candidates (WF: whole fragment; LH: left half and RH: right half as indicated on Fig. 1, green colored fragments) were integrated into the dTALE AD domain adjacent portion by fusion PCR (the procedure is shown on Fig. S1), then subcloned into dTALE *Bbv*CI and *Xba*I site to assure that only the dTALE AD domain was replaced and to create the final construct, pDW2177ARS–CIA5. For expression in *Chlamydomonas*, this cTALE construct was used to transform cw10 (CC849, mt-* *) cells by the glass bead transformation technique (VAN *et al.* 2001). ARS activity in transformants was assayed as previously described (GAO *et al.* 2014).

2.4 Gene expression analysis

RNA was extracted using an RNeasy Plant Mini Kit (Qiagen, Cat. No. 74904), and the RNA concentration was measured using an ND-2000 Nanodrop spectrophotometer (Nanodrop Technologies). 1-5 μg total RNA was treated with TURBO DNA-Free™ kit (Life Technologies), and cDNA was generated from 0.5 μg - 1 μg of treated total RNA (Verso cDNA Synthesis Kit; Thermo Scientific). The resulting cDNA (25 - 50ng) was used for reverse transcription polymerase chain reaction (RT-PCR) and for quantitative RT-PCR (qRT-PCR). Primer sequences and efficiencies are provided in the supplemental information (Table S1). The qRT-PCR was performed on a CFX Connect™ Real-Time PCR Detection System (Biorad) using a SYBR green two-step quantitative PCR system (Quanta Biosciences). For quantitative analyses, the CBLP gene was used as an internal control for normalization of qRT-PCR data (FANG *et al.* 2012). The relative transcript abundance in each sample is defined as \( \Delta Ct = Ct_{\text{target gene}} - Ct_{\text{CBLP}} \) to represent the difference between the transcript abundance of genes examined and the transcript abundance of CBLP. After normalization, \( \Delta Ct \) values of each transformant were compared with that of wild type cells, and the dTALE induced expression is represented by \( \Delta \Delta Ct = \Delta Ct_{\text{transformant}} - \Delta Ct_{\text{wild type}} \) (GAO *et al.*
2014). For direct comparisons, fold change, which equals $2^{\Delta\Delta Ct}$, is also used to represent the relative abundance of target genes.

2.5 Western immunoblot analysis

For western immunoblotting, total protein was obtained as described previously (DUANMU et al. 2009b). Proteins were separated by SDS-PAGE on 4-12% polyacrylamide gels. Immunoblotting utilized an anti-FLAG-tag antibody (Santa Cruz Biotechnology; catalog no. sc-51590) and was performed as described in the protocol from Bio-Rad Laboratories (product catalog no. 500-0006).

2.6 Expression of protein fragments in E. coli

A fragment of the CIA5 gene from Eagl (1214) to NotI (1930) sites and all other fragments used in this study were subcloned into pET30 vector and then transformed into E. coli. Cells were IPTG induced at 37°C. Proteins were separated by SDS-PAGE on 8% or 10% polyacrylamide gels, and detected by immunoblotting using an anti-His-tag antibody.

3. Results

3.1 CIA5/CCM1 auto-activation domain identification in yeast

Yeast two-hybrid screening is a technique used to discover protein-protein interactions by testing for physical interactions (such as binding) between two proteins. The most common screening approach is the yeast two-hybrid assay. (VAN CRIEKINGE and BEYAERT 1999). Using a yeast two hybrid GAL4 system to detect CIA5/CCM1 interacting components, we found too many positive yeast colonies to count when we used full length CIA5 cDNA (fragment 1 in Fig. 1) as bait to screen a cDNA library (WANG and SPALDING 2014b). We hypothesized that CIA5/CCM1 itself can auto-activate in yeast, so we introduced the full-length CIA5 bait plasmid alone, without the prey plasmid, and again observed too many positive yeast colonies to count. Thus, the CIA5/CCM1 protein itself appears to have auto-activation activity, at least in yeast.
Taking advantage of this auto-activation effect in yeast to narrow down which CIA5/CCM1 regions might act as putative activation motifs, we tested whether fragments of CIA5/CCM1 can replace the yeast activation domain in activating gene expression. In testing several other fragments (Fig. 1, fragments 2 - 28), we found that the GAL4 DNA-binding domain fused to any CIA5 fragments that included aa 436 to 544 (Red box region as shown on Fig. 1), including fragments 17, 18, 21 or 22, could auto-activate the GAL4 yeast two hybrid system in the absence of the GAL4 activation domain. This 109-aa delineated auto-activation region agrees very well with the 130-aa C-terminal region highly conserved between Volvox carteri (hereafter, Volvox olvox) CCM1 and Chlamydomonas CIA5/CCM1 (aa 428 to 557) (Golden color labeled region on Fig. 1) (YAMANO et al. 2011).

Therefore, using a GAL4 yeast two hybrid system, we identified a 109-aa region of CIA5/CCM1 (aa 436 - 544) that is contained within the 130-aa region conserved between Volvox and Chlamydomonas and that has auto-activation activity in yeast. We further investigated the CIA5 fragments capable of auto-activation in yeast and discovered that either half of the 109-aa or the 130-aa conserved C-terminus domain (Fig. 1, fragments 16, 20, 24, & 25) could activate the GAL4 system in yeast. Results from experiments using either the SD dropout selective medium assay or an X-gal test (Fig. 2 and Table S1) confirmed that either the 109-aa and 130-aa region or either half of these regions behave as true activation domains, at least in yeast.

### 3.2 Activation domain confirmation in Chlamydomonas

Although the efficiency for activating target gene expression is relatively low, a dTALE system can be used to activate gene expression in Chlamydomonas (GAO et al. 2014; GAO et al. 2015). We reasoned that we could test the activation function of putative activation domains in Chlamydomonas cells by using them to replace the TALE activation domain (ZHU et al. 1998).
A. dTALEs were successfully delivered and expressed in 

*Chlamydomonas* cells

We used a two-step PCR reaction to create constructs pDW2177ARS–CIA5 containing either the whole 130-aa (aa 428 - 557) putative activation domain (pDW2177ARS–CIA5–WF) or either half of this domain (either pDW2177ARS–CIA5–LH or pDW2177ARS–CIA5–RH) (Fig S1). We then transformed wild-type *Chlamydomonas* strain cw10 (cc849) with these dTALE constructs and selected transformants based on resistance to zeocin. The original dTALE construct, pDW2177ARS (Gao et al. 2014), and a modified dTALE lacking any activation domain, pDW2177ARS-AD, were used as positive and negative controls, respectively (Fig 3). We confirmed the insertion into the *Chlamydomonas* genome of intact dTALE constructs in about 20-40% of the 192 zeocin resistant transformants screened for each of the five constructs, using colony PCR (Fig. S2).

B. dTALEs induced expression of ARS1 in *Chlamydomonas*

A colorimetric assay was used under sulfur-replete growth conditions to screen for ARS activity, which should indicate any activation of *ARS1* in the absence of sulfur stress. Of the 192 transfantants screened for the positive control (pDW2177ARS) and the three CIA5-fragment-containing pDW2177ARS-derived constructs (pDW2177ARS–CIA5–WF, pDW2177ARS–CIA5–LH or pDW2177ARS–CIA5–RH) about 5-20% show XSO₄⁻ based color changes in sulfur-replete culture medium (representative data in Fig. 4; not all data shown), all of which contained a PCR-detected, intact dTALE construct. None of the transformant colonies lacking a PCR-detected intact dTALE construct showed detectable color change under the same growth conditions (data not shown). Of 192 negative control transformants (i.e., those transformed with pDW2177ARS-AD which lacks an activation domain), 42 lines (21.9 % of all transformants screened) contained a PCR-detected intact dTALE, but none showed visible ARS activity (representative data for four transformants (-AD1, -AD2, -AD3 and -AD4) are provided in Fig. 4; not all data shown).
Successful expression of dTALE constructs at the transcript level was confirmed using RT-PCR (Fig. 5). We tested 4 candidate transformant lines for each of the 5 pDW2177ARS-derived constructs, using primers specific for the 5' region of the dTALE constructs. All showed detectable expression of dTALE-ARS1 transcripts. We also used western immunoblot analysis to detect dTALE expression at the protein level. Of all tested transformant lines exhibiting RT-PCR-detectable expression of dTALEs, protein expression was confirmed using anti-FLAG tag antibody (Fig. 6). The expression of dTALE also was detected in pDW2177ARS-AD transformants lines (lacking an activation domain) at both the transcript and protein levels (Figs. 5 & 6).

Activation of target gene, ARS1, was demonstrated at the transcript level with RT-PCR, using primers specific for the ARS1 3'UTR. When grown under sulfur starvation, elevated expression of ARS1 was observed in wild type cells (Fig. 5). dTALE-induced expression of ARS1 transcripts under sulfur-replete conditions was observed in all tested lines with PCR-detected, intact pDW2177ARS-derived dTALE construct and colorimetric ARS activity (positive control and three CIA5-fragment-containing constructs), but not in pDW2177ARS-AD lines without activation domain showing an intact pDW2177-ARS1 dTALE construct but no detectable colorimetric ARS activity (Figs. 4 & 5).

C. Each half of the CIA5 activation domain can function as an independent activation domain

All three dTALE constructs containing the putative CIA5 activation domain or its left and right fragments (i.e., pDW2177ARS-CIA5-WF, pDW2177ARS-CIA5-LH, and pDW2177ARS-CIA5-RH, respectively) activate expression of the target gene, ARS1, under sulfur-replete conditions (Fig. 5). To gain a more quantitative appraisal of gene transcription enhancement supported by the p2177ARS-CIA5-based dTALE constructs, we used DNA from the same transformed lines employed above to conduct qRT-PCR analyses. The abundance of ARS1 transcripts in all examined lines from the positive control (i.e., CW10 -S) and the three CIA5-fragment-containing constructs (dTALE containing an activation domain) (Fig. 7) were all positively correlated to the levels of ARS activity observed in the in vivo (in Chlamydomonas)
measurements of ARS activity present in cultures of transformed cells (Fig. 4). Transformants lacking an activation domain (i.e., pDW2177ARS-AD) showed only lower, basal level ARS expression, even though the dTALE abundance was similar to that in all other transformants (Fig. 7). The relationship between qRT-PCR based ARS1 transcript abundance and colorimetric ARS activity assays performed on transformants expressing pDW2177ARS-based dTALEs in sulfur-replete medium showed that colorimetric ARS activity was detectable in only those selected lines with substantial transcript abundance. Interestingly, the levels of ARS gene transcripts in transformants containing constructs encoding the left or right fragments of the CIA5 activation domain were somewhat higher than in the constructs carrying the intact activation domain, but this did not translate to markedly increased responses in the ARS colorimetric analyses (Figure 4).

3.3 Region responsible for abnormal migration of CIA5 during SDS-PAGE

CIA5 always shows abnormal SDS-PAGE migration, whether detected in Chlamydomonas cell extracts with anti-CIA5 antibody (Wang et al. 2005; Kohinata et al. 2008) or in E. coli extracts overexpressing His-tagged CIA5 and detected with anti-His antibody (Fig. S3) or anti-CIA5 antibody (chapter 3). To investigate what region of CIA5 is responsible for the abnormal SDS-PAGE migration, we expressed in E. coli different constructs bearing either full length or partial CIA5 coding regions (containing C-terminal extensions encoding a 6xHis-tag) (Fig. 8) and then detected their expression using anti-His-tag antibody to probe western immunoblots (Fig. S3). As shown both in Fig. 8 and Fig. S3, a constructs containing the full-length CIA5 cDNA (Frag 31, Fig. 8 and Fig. S3) produced a protein migrating as a ~95 kD protein instead of the expected 72.1 kD (i.e., an approximately 23 kD difference between observed and predicted molecular mass). Indeed, any partial cDNA containing the coding region for amino acids from residue 402 to residue 642 [or the Ascl-NotI region (aa 399-642)] showed aberrant migration rates during SDS-PAGE (i.e., Fragments 32, 34, 35, 36, 37, 38, 39, and Frag 40 in Fig. 8 and Fig. S3). Those constructs containing Frag 36 (aa 404 to 642) show the greatest abnormal SDS-PAGE migration (about 22kD difference from the expected). The constructs containing only a small part of this
region (Frag 37, 38, 39 and Frag 40 in Fig. 8 and Fig. S3) show less abnormal migration, and those including no part of this region (e.g., Frag 33 in Fig. 8 and Fig. S3) exhibit no abnormal migration. This apparent size aberrancy is observed both with CIA5 isolated from both *Chlamydomonas* and CIA5 produced from *E. coli* - and, thus, not likely due to post-translational modification of the protein. In particular, the aberrant migration also does not result from post-translational SUMOylation of CIA5 because the protein still ran aberrantly and could still complement the *cia5* mutant when every potential CIA5 SUMOylation site was mutated (Wang et al. 2008).

The region apparently responsible for abnormal SDS-PAGE migration (aa 399 - 642) overlaps extensively with the 130-aa C-terminal CIA5/CCM1 domain conserved between *Chlamydomonas* and *Volvox* that we demonstrated above to behave as an activation domain in both yeast and *Chlamydomonas*. The aa sequence of this region contains an abundance of acidic aa residues (E and D), which suggests that it may function as an acidic activation domain (Zhu et al. 1998).

Increased mobility during SDS-PAGE migration has been reported to occur when amino acid residues are substituted that decrease the net negative charge on a protein such as H80R, G85R (Shi et al. 2012). Further studies demonstrated that aa substitutions such as H80R, G85R, D90A, G93R, E100G, E100K, D101G, and D101N can increase migration (by ~2 - 3 kDa) during SDS-PAGE by promoting the binding of three to four additional SDS molecules, without significantly altering the secondary structure, while another substitution, G93D, decreases migration (by ~2 - 3 kDa) (Shi et al. 2012). A slight discrepancy between the theoretical MW and the MW observed during SDS-PAGE is not uncommon, incomplete unfolding might influence electrophoretic mobility, also, the observed MW of a protein might differ from the theoretical value when the protein binds more or less SDS than it should on average due to amino acid composition, thus appearing to be larger or smaller than expected. It was reported that basic proteins bind more SDS and acidic proteins bind less detergent than the average (Bronner and Kleinzeller 1979). The pI for CIA5 aa 1 - 698 is 5.58, while for the 130-aa conserved C-terminal region is 4.14 (predicted by Online Analysis Tool-Protein Chemistry, http://web.expasy.org/compute-pi/). We may reason that it is possible that high density of acidic aa residues such as CIA5 130-aa
region should decrease SDS binding, thus decreasing the net negative charge of the acidic region and causing abnormal migration.

3.4 Region indispensable for complementation of \textit{cia5} mutant

The CIA5 130-aa C-terminal region conserved between \textit{Chlamydomonas} and \textit{Volvox} that we confirmed to act as an activation domain in yeast and \textit{Chlamydomonas}, also appears essential for complementation of the \textit{cia5} mutant. This conclusion was reached from results of experiments in which a number of restriction enzyme-based deletions in the CIA5 gene were created and tested for their ability to complement the \textit{cia5} mutant. In all cases, constructs were designed (and confirmed by DNA sequencing) to ensure that a proper reading frame for the coding sequence was created. Deletion of either \textit{BglII}-\textit{XbaI} or of \textit{NotI} resulted in deletion of 3'-UTR (Fig. 9, 1) or near the C-terminal encoding aa 642 - 698 region (Fig. 9, 2) shows no effect on the ability to complement \textit{cia5} mutant. Deletion of the large \textit{Ascl} restriction fragment (encoding aa 153 - 399) (Fig. 9, 3) or deletion of a fragment encoding aa 261 - 420 (Fig. 9, 7) from the CIA5 WT gDNA does not affect the ability of this DNA molecule to complement the \textit{cia5} mutant, but deletion of the larger \textit{XmnI}-\textit{NotI} restriction fragment encoding aa 192 - 642 (Fig.9, 4), which contains the C-terminal conserved domain, prevents complementation. This large coding region overlaps with the \textit{EagI}-\textit{NotI} region (data not shown) responsible for the aberrant SDS-PAGE migration (see above) and the conserved 130-aa putative activation domain associated with gene activation in yeast and \textit{Chlamydomonas}. Deletion of the \textit{EagI}-\textit{NotI} restriction fragment (encoding aa 403 - 642) also eliminates CIA5 function. Deletion the approximate 130-aa region aa 430 - 550 (Fig9, 5) also eliminates the \textit{CIA5} function. Even deletion of only nucleotides 1288 to 1458 (encoding aa 430 - 487), approximately the first half of the 130-aa C-terminal putative activation region conserved between \textit{Chlamydomonas} and \textit{Volvox}, also eliminates the ability of the resulting CIA5 fragment to complement the \textit{cia5} (Fig. 9, 6).
4. Conclusion and Discussion

As a master regulator of the CCM in *Chlamydomonas*, CIA5 (also known as CCM1) may function alone or work with other factors to control downstream gene expression. *CIA5* was identified independently by two research groups (*Fukuzawa et al.* 2001; *Xiang et al.* 2001) more than two decades ago, but little progress has been made in understanding how CIA5 regulates such a large network of genes (*Fang et al.* 2012). The regulation of thousands of genes, including induction or up-regulation of almost all limiting-CO₂ inducible genes (including most putative Ci transporters and some CAs) is impaired in the *cia5* mutant (*Miura et al.* 2004; *Wang et al.* 2005; *Fang et al.* 2012).

CIA5/CCM1 contains two N-terminal zinc finger domains. If one of these zinc finger domains is mutated, as in the *cia5* or *ccm1* mutants (*Xiang et al.*, 2001; *Fukuzawa et al.*, 2001), CIA5 activity is lost and the regulation of most CO₂-responsive genes is disrupted. Aside from this role as the regulator of the CCM and other low CO₂ acclimation responses, the details of CIA5 function remain largely undefined. Transcriptome analyses of cells shifted from high to low CO₂ conditions identified a massive impact of CIA5/CCM1 and CO₂ on the transcription of several thousand genes and revealed an array of gene clusters with distinctive expression patterns (*Brueggeman et al.* 2012; *Fang et al.* 2012). These observations provide important insights into the regulatory interaction between CIA5 and CO₂ and the knowledge that individual gene clusters responded primarily to CO₂-regulated changes in CIA5, activity (*Fang et al.* 2012).

The yeast two hybrid system helped us to identify a 109aa region (436aa - 544aa) of CIA5/CCM1 that corresponds to a 130-aa region (428aa - 557aa) very highly conserved between *Chlamydomonas* CIA5 and *Volvox* CCM1 - a region that also exhibits auto-activation in yeast. Replacement of the normal TALE activation domain in a dTALE with this 130-aa putative activation domain demonstrated that this region also behaves as an activation domain in *Chlamydomonas*, suggesting it also may function as an activation domain in CIA5. Deletion of the *XmnI-NotI* region, which contains the 130-aa C-terminal conserved, putative activation domain, totally
eliminates the ability of CIA5 to complement the \textit{cia5} mutant - an observation which is consistent with this region having an obligate function in CIA5.

The 130-aa, conserved C-terminal putative activation domain contains an abundance of acidic amino acid residues, and is included within the \textsc{Acl} and \textsc{Nof}l region (aa 399-642) responsible for abnormal SDS-PAGE migration of CIA5. Both observations are consistent with the identified 130-aa conserved C-terminal region acting as an acidic activation domain.

This 130-aa, putative activation region may represent two independent activation domains (aa 428 - 487 and aa 488 - 557), since each half of the 130-aa domain could function in auto-activation in yeast as well as in the dTALE system to activate gene expression in \textit{Chlamydomonas}. However, deletion of only aa 430-487, approximately the first half of the 130-aa, C-terminal, conserved domain, totally eliminates the ability of CIA5 to complement the \textit{cia5} mutant. This observation argues that the second putative activation domain within this 130 aa region is not sufficient by itself to allow CIA5 function.

The present work represents the first demonstration that a conserved, acidic domain from the CCM master regulator, CIA5, can function as an activation domain both in \textit{Chlamydomonas} and in yeast. This conserved acidic activation domain also appears indispensable for CIA5 function, as judged from complementation experiments with the \textit{cia5} mutant. Together, these observations provide strong suggestive evidence that CIA5 functions, at least in part, through gene direct activation of gene expression. Whether CO$_2$-responsive genes are specifically targeted for activation by CIA5 alone or by collaboration of CIA5 with specific transcription factors remains a prime question to be answered in future studies.
References


SHI, Y., R. A. MOWERY, J. ASHLEY, M. HENTZ, A. J. RAMIREZ et al., 2012 Abnormal SDS-PAGE migration of cytosolic proteins can identify domains and mechanisms that control surfactant binding. Protein Sci **21**: 1197-1209.

SOMANCHI, A., and J. V. MORONEY, 1999 *As Chlamydomonas reinhardtii* acclimates to low-CO$_2$ conditions there is an increase in cyclophilin expression. Plant Mol Biol **40**: 1055-1062.


VAN, K., Y. WANG, Y. NAKAMURA and M. H. SPALDING, 2001 Insertional mutants of *Chlamydomonas reinhardtii* that require elevated CO$_2$ for survival. Plant Physiol **127**: 607-614.


WANG, Y., I. LADUNGA, A. R. MILLER, K. M. HORKEN, T. PLUCINAK et al., 2008 The small ubiquitin-like modifier (SUMO) and SUMO-conjugating system of *Chlamydomonas reinhardtii*. Genetics **179**: 177-192.

WANG, Y., and M. H. SPALDING, 2014 LCIB in the *Chlamydomonas* CO$_2$-concentrating mechanism. Photosynth Res **121**: 185-192.

WANG, Y. J., and M. H. SPALDING, 2006 An inorganic carbon transport system responsible for acclimation specific to air levels of CO2 in Chlamydomonas reinhardtii. Proc Natl Acad Sci USA 103: 10110-10115.


Figure and Table Legends

**Fig. 1.** CIA5 full coding DNA sequences (CDS) or DNA fragments containing regions of the CDS used for yeast two hybrid experiments. The uppermost bar represents the full CIA5 coding sequence. The position of each fragment in the CIA5 coding sequence is indicated by its position relative to the full coding sequence and by the numbers representing its span of nucleic acid residues in the coding DNA sequence. Fragments colored either blue (●) or green (■) showed auto activation in the yeast two hybrid assay. Fragments lacking color were negative when tested using the two hybrid assay. The red box (□) indicates the C-terminal region conserved between *Chlamydomonas* and *Volvox* and that overlaps with the auto activation region identified using the yeast two hybrid assay, and the grey dashed line (┆) splits this region into two halves. The golden color region (●) is the region which was identified as promoting auto activation in the two hybrid assay. The green colored fragments (16, 18 and 25) were used for replacing dTALE activation domain and designated as LH (16: 1282 - 1461, left half), WF (18: 1282 - 1671, whole fragment) and RH (25: 1462 - 1671, right half) respectively. The grey colored fragments (29 and 30) are those containing a His^{54} to Tyr^{54} point mutation.

**Fig. 2.** Yeast two hybrid experiment using CIA5 CDS and CIA5 CDS fragments, but in the absence of activation-domain partners to identify auto-activation domains. pGBKT7 constructs containing either full length or partial CIA5 cDNAs (numbered as indicated in Figure 1) were transformed into yeast AH109, and selected using SD dropout Trp (SD-Trp) selective plates. Transformants were then transferred to SD dropout Trp and His (SD-Trp-His) selective plates to screen for possible auto-activation. Any growth on SD-Trp-His medium indicates potential auto-activation candidates. X-gal induction was used as a secondary screen to confirm auto-activation. Because α-galactosidase is a secreted enzyme, it can be assayed directly using X-α-Gal indicator plates.

**Fig. 3.** dTALE constructs used for testing CIA5 putative activation domains in *Chlamydomonas* cells. The original activation domain of pDW2177ARS construct was
replaced by either the whole CIA5 130 aa putative activation domain (whole fragment - WF) or by one half of that domain (left half, LH or right half, RH).

**Fig. 4.** Representative examples of colorimetric screens for ARS activity induced by dTALEs. CW10 cells grown either with (+S) or without (-S) sulfur serve as positive and negative controls, respectively. Other transformants were all grown with sufficient sulfur for 60 hours, then incubated with XSO4. The blue-green color indicates ARS activity. Microtiter wells -AD1 to 4 contain cells transformed with the pDW2177ARS-AD construct lacking an activation domain; wells +AD1 to 4 contain cells transformed with construct pDW2177ARS that contains the native TALE activation domain; wells LH1 to 4 contain cells transformed with pDW2177ARS-CIA5-LH that contains the left half of the CIA5 activation domain; wells WF1 to 4 contain cells transformed with pDW2177ARS-CIA5-WF that contains the entire CIA5 activation domain; wells RH1 to 4 contain cells transformed with pDW2177ARS-CIA5-RH that contains the right half of the CIA5 activation domain.

**Fig. 5.** RT-PCR analysis of *ARS1* and *dTALE* transgene transcription in *Chlamydomonas* cells transformed with various pDW2177ARS-derived constructs. Detection of transcripts from the *CBLP* gene (upper panel) in all lines tested served as a positive internal control. The middle panel shows the presence or absence of *ARS1* transcripts in the transformants tested when grown in sulfur replete medium. The lower panel shows the presence of *dTALE* transcripts in all lines tested. *ARS1* transcripts are evident in lines with an inserted activation domain (middle panel, lines 1 through 4 of +AD, LH, WF and RH), but not in those without activation domain transgene (middle panel, lines 1 through 4 of -AD). Transformed lines were grown under sulfur sufficient conditions, while CW10 cells (right two lanes) were grown under either sulfur-sufficient or sulfur-deficient conditions, as indicated, for 60 hours. Numbers under each panel indicate transformant line.

**Fig. 6.** Western immunoblots of *Chlamydomonas* cells transformed with various pDW2177ARS-derived constructs showing the relative abundance of the dTALE proteins. Anti-Flag-tag antibody was used as described in Material and Methods. Cells were grown under sulfur replete conditions for 60 hours. CW10 cells grown without sulfur (-S) were used as a control.
Fig. 7. Analysis of ARS1 target gene and dTALE transcript abundance by qRT-PCR. ARS1 and dTALE transcript abundance (Relative log2 Fold Change), showing transformed lines with intact dTALE-ARS1 and visible ARS activity. Transformants without an inserted activation domain (-AD 1-4) showed only basal level ARS expression. The ARS transcript abundance is much higher in those with an inserted activation domain, but not as high as the sulfur starvation induced control (CW10-S).

$$\Delta \Delta C_t = \Delta C_t^{\text{transf}} - \Delta C_t^{\text{WT}} = (C_t^{\text{target gene-transf}} - C_t^{\text{CBLP-transf}}) - (C_t^{\text{target gene-WT}} - C_t^{\text{CBLP-WT}})$$, where $\Delta C_t$ equals the difference between target transcript abundance and the transcript abundance of CBLP, the internal control gene. Three technical replicates were performed for each transformant and error bar is standard error. Transformed lines were grown under sulfur replete conditions, while CW10 cells were grown under either sulfur-replete or sulfur-deficient conditions, as indicated, for 60 hours. Numbers underneath each panel indicate transformant lines.

Fig. 8. Migration of CIA5 and CIA5 fragments during SDS-PAGE. CIA5 fragments were used for testing the effect of partial deletions on protein migration during SDS-PAGE. Beside each construct is indicated the predicted mass of the CIA5 fragment, the mass of the fragment deduced from migration of protein standards during SDS-PAGE and the difference between the two - a measure of migration aberrancy. The red box (□) indicates the approximately 130-aa region.

Fig. 9. Regions of the CIA5 gene or cDNA that were deleted in experiments to test the ability of truncated CIA5 molecules to complement the cia5 mutant. Seven different deletions (1-7) were created to determine which regions were essential for CIA5 activity as deduced from the ability or inability of each deletion to complement the CIA5 defect in the Chlamydomonas cia5 mutant.
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1-2097</td>
</tr>
<tr>
<td>10</td>
<td>922-1281</td>
</tr>
<tr>
<td>11</td>
<td>922-1305</td>
</tr>
<tr>
<td>12</td>
<td>922-1461</td>
</tr>
<tr>
<td>13</td>
<td>922-1632</td>
</tr>
<tr>
<td>14</td>
<td>922-1671</td>
</tr>
<tr>
<td>15</td>
<td>922-2097</td>
</tr>
<tr>
<td>16</td>
<td>1282-1461</td>
</tr>
<tr>
<td>17</td>
<td>1282-1632</td>
</tr>
<tr>
<td>18</td>
<td>1282-1671</td>
</tr>
<tr>
<td>19</td>
<td>1282-2097</td>
</tr>
<tr>
<td>20</td>
<td>1306-1461</td>
</tr>
<tr>
<td>21</td>
<td>1306-1632</td>
</tr>
<tr>
<td>22</td>
<td>1306-1671</td>
</tr>
<tr>
<td>23</td>
<td>1306-2097</td>
</tr>
<tr>
<td>24</td>
<td>1462-1632</td>
</tr>
<tr>
<td>25</td>
<td>1462-1671</td>
</tr>
<tr>
<td>26</td>
<td>1462-2097</td>
</tr>
<tr>
<td>27</td>
<td>1633-2097</td>
</tr>
<tr>
<td>28</td>
<td>1672-2097</td>
</tr>
</tbody>
</table>

**Fig. 1**
Fig. 2
Fig. 3
Transcripts Relative Abundance (ΔΔCt)

Fig. 6

Fig. 7
Fig. 8
A: Results from CIA5 deletion complementation

Summary of CIA5 deletions evaluated:
1. CIA5 gDNA Δ BglII-XbaI (3’-UTR)
2. CIA5 gDNA Δ NotI (aa 642-698)
3. CIA5 gDNA Δ Ascl (aa ~153-399)
4. CIA5 gDNA Δ XmnI-NotI (aa ~192-642)
5. CIA5 gDNA Δ aa 430-550
6. CIA5 gDNA Δ aa 430-487
7. CIA5 gDNA Δ aa 261-420
8. CIA5 WT gDNA

Effect on cia5 (H54Y) complementation:
1. None
2. None
3. Minimal
4. Complete
5. Severe
6. Severe
7. None
8. None

B: Summary of CIA5 deletion complementation

Fig. 9
Supplemental Information

Table S1. A summary of all the primers used for this work. A: Primers used for amplifying CIA5 fragments in yeast two-hybrid (inserts into pGBKT7 through EcoRI/SalI site). CIA5-Tyr primer pairs (primer 17 and 18) are used to amplify the fragment containing His<sup>54</sup> to Tyr<sup>54</sup> point mutation. B: Primers used for constructing CIA5 putative activation domain (either half LH/RH or whole fragment, WF) by fusion PCR (as shown on Fig. S1) to replace the dTALE activation domain through BbvCI/XbaI multicloning site. C: Primers for checking dTALE constructs intactness. D: Primers used for the RT-PCR and qRT-PCR to check dTALE and ARS1 expression.

Table S2. CIA5 Fragments used for yeast two-hybrid auto-activation identification, and the summary for the results from yeast two-hybrid and X-gal screening. The fragments of CIA5 full or partial CDS were used for Yeast two-hybrid experiments to identify auto-activation domains, and the results from both yeast two hybrid and X-gal test are indicated as either positive (+) or negative (−). Any constructs which are able to auto-activate can grow on SD-Trp-His selective plates and show blue color on X-gal screening.

Fig. S1. Schematic of a two-step PCR (fusion PCR) approach for synthesis and insertion of CIA5 activation domain (AD) into pDW2177ARS construct. In a first round PCR reaction, pairs of primers (sequences provided in Table S2) are used to amplify two (soon-to-be overlapping) fragments. Primers R1 and F2 are complementary to each other (red), 20 nucleotide (nt) from the pDW2177ARS vector, and 19 nt from CIA5 AD sequence. Following the first PCR reaction, the two products of the reaction are purified and mixed into a second PCR reaction to become overlapping PCR mega-primers as a result of the pairing of their complementary 3' ends. Simultaneously, they also become PCR reaction templates for the PCR primers, F1 and R2.

Fig. S2. Colony PCR determination of whether transformants have intact dTALE construct. Two primer pairs F1/R1 and F2/R2 (Sequence provided on Table S2) were pooled together then used to perform colony PCR on individual Chlamydomonas transformants. Those containing AD domains clearly showed two bands. Construct pDW2177ARS-AD, which lacks an inserted AD domain and is
missing the R1 primer, only shows a single size PCR product. The AD adjacent flanking regions were amplified and ligated to AD candidates (CIA5-WF, RH or LH) by fusion PCR (as shown on Fig. S1) to assure only the original AD sequence was replaced, then the fragments were inserted into pDW2177ARS plasmid through BbvCl/XbaI site. The recombinants plasmids were transformed into wall-less Chlamydomonas strain CW10 and selected against Zeocion resistance. Primer pairs F1/R1 and F2/R2 were used to check the dTALE construct intactness. All constructs contain the same 5’dTALE fragment (primer pair F2/R2 will amplify 520bp in all 5 constructs), each contains different length of 3’ dTALE activation domain (primer pair F1/R1 will amplify the sequence including dTALE AD domain, for pDW2177ARS-CIA5-WF: 858bp, pDW2177ARS-CIA5-RH: 678bp, pDW2177ARS-LH: 648bp, and pDW2177ARS: 600bp. For pDW2177ARS–AD, the one without AD domain disrupted the R1 reverse primer, so there is no PCR product if using primer pair F1/R1. If pooled primers F1/R1 and F2/R2 together when we perform colony PCR to check constructs dTALE intactness, for the ones containing AD domain will have two different size of PCR product, the one which lack AD domain only shows one band.

Fig. S3. Western immunoblot detecting recombinant CIA5 fragment migration on SDS-PAGE.
Table S1. Primers used to generate CIA5 fragments

<table>
<thead>
<tr>
<th></th>
<th>5’-EcoRI</th>
<th>CCGGAAATTCATGGAGACCTTAGACGCGCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3’-303-Sal</td>
<td>ACCGCCTCACTCTACGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>3</td>
<td>3’-330-Sal</td>
<td>ACCGCAGTCTACCTACGGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>4</td>
<td>3’-588-Sal</td>
<td>ATTGTGAGGCTACGGGAAGGGGTTC</td>
</tr>
<tr>
<td>5</td>
<td>3’-1281-Sal</td>
<td>ACCGCAGTCTACCTACGGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>6</td>
<td>3’-1305-Sal</td>
<td>ACCGCAGTCTACCTACGGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>7</td>
<td>3’-1461-Sal</td>
<td>ACCGCAGTCTACCTACGGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>8</td>
<td>3’-1632-Sal</td>
<td>GATGAGGCTACGGGAAGGGGTTC</td>
</tr>
<tr>
<td>9</td>
<td>3’-1671-Sal</td>
<td>AGGCTCTAGCTAGCCCTCTCTCCGG</td>
</tr>
<tr>
<td>10</td>
<td>3’-2097-Sal</td>
<td>ACCGCAGTCTACCTACGGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>11</td>
<td>5’-922-EcoRI</td>
<td>CCAGCAGTCTACCTACGGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>12</td>
<td>5’-1282-EcoRI</td>
<td>CCAGCAGTCTACCTACGGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>13</td>
<td>5’-1306-EcoRI</td>
<td>CCAGCAGTCTACCTACGGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>14</td>
<td>5’-1462-EcoRI</td>
<td>CCAGCAGTCTACCTACGGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>15</td>
<td>5’-1633-EcoRI</td>
<td>CCAGCAGTCTACCTACGGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>16</td>
<td>5’-1672-EcoRI</td>
<td>CCAGCAGTCTACCTACGGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>17</td>
<td>CIA5-Tyr F</td>
<td>CGATGAGGCTACGGGAAGGGGTTC</td>
</tr>
<tr>
<td>18</td>
<td>CIA5-Tyr R</td>
<td>AGGCTCTAGCTAGCCCTCTCTCCGG</td>
</tr>
<tr>
<td>19</td>
<td>5’-BbvCI</td>
<td>ATTTGCCCTCAGC</td>
</tr>
<tr>
<td>20</td>
<td>5’-1282</td>
<td>CCAGCAGTCTACCTACGGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>21</td>
<td>3’-1282</td>
<td>CCAGCAGTCTACCTACGGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>22</td>
<td>2’-1462</td>
<td>CCAGCAGTCTACCTACGGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>23</td>
<td>3’-1462</td>
<td>CCAGCAGTCTACCTACGGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>24</td>
<td>3’-1461-Xbal</td>
<td>GGTGCTAAGCTAGCCCTCTCTCCGG</td>
</tr>
<tr>
<td>25</td>
<td>3’-1671-Xbal</td>
<td>GGTGCTAAGCTAGCCCTCTCTCCGG</td>
</tr>
<tr>
<td>26</td>
<td>Colony F1</td>
<td>GGCCTACGGGCTACGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>27</td>
<td>Colony R1</td>
<td>CATTTACACGGAGCGGGGTC</td>
</tr>
<tr>
<td>28</td>
<td>Colony F2</td>
<td>AACAGATCTGGCACCATTGGGATC</td>
</tr>
<tr>
<td>29</td>
<td>Colony R2</td>
<td>TAGATCTTCTCTAGTGGGATC</td>
</tr>
<tr>
<td>30</td>
<td>ARS1_QF1</td>
<td>GATCTAGCTACGGGCTACGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>31</td>
<td>ARS1_QR1</td>
<td>CCTAAACATTTGGGTCTCTAGCCCTCTCTCCGG</td>
</tr>
<tr>
<td>32</td>
<td>dTALE_QF4</td>
<td>TCGCAGATCTTACGAGGGGTTC</td>
</tr>
<tr>
<td>33</td>
<td>dTALE_QR4</td>
<td>CCTGTGCTCAGCTAGCCCTCTCTCCGG</td>
</tr>
<tr>
<td>34</td>
<td>CBLP_QF1</td>
<td>GACCGTACGGGCTACGGGCGCAAGGGGTTC</td>
</tr>
<tr>
<td>35</td>
<td>CBLP_QR1</td>
<td>GACCGTACGGGCTACGGGCGCAAGGGGTTC</td>
</tr>
</tbody>
</table>

A: Primers (primer 1-18) for amplifying CIA5 fragments used in yeast two hybrid (inserts into pGBK7 through EcoRI/SalI site). CIA5-Tyr primer pairs are used to amplify the fragment containing His^{54} to Tyr^{54} point mutation.

B: 19-25 used for constructing CIA5 putative activation domain (either LH/RH or whole WF) to replace the dTALE AD.

C: 26-29 for colony PCR to check dTALE constructs intactness.

D: Primers (30-35) used for the RT-PCR and qRT-PCR to check dTALE and ARS1 expression.
<table>
<thead>
<tr>
<th>Constructs</th>
<th>2:</th>
<th>3:</th>
<th>4:</th>
<th>5:</th>
<th>6:</th>
<th>7:</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD-Trp-His</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>X-Gal</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>response</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constructs</td>
<td>8:</td>
<td>9:</td>
<td>1:</td>
<td>10:</td>
<td>11:</td>
<td>12:</td>
</tr>
<tr>
<td>SD-Trp-His</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>X-Gal</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>response</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constructs</td>
<td>13:</td>
<td>14:</td>
<td>15:</td>
<td>16:</td>
<td>17:</td>
<td>18:</td>
</tr>
<tr>
<td>SD-Trp-His</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>X-Gal</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>response</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD-Trp-His</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>X-Gal</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>response</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constructs</td>
<td>25:</td>
<td>26:</td>
<td>27:</td>
<td>28:</td>
<td>29:</td>
<td>30:</td>
</tr>
<tr>
<td>SD-Trp-His</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>X-Gal</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>response</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* The constructs with the point mutation H54Y.

Table S2
Fig. S1
Fig. S2
Fig. S3
CHAPTER 3. CIA5 DNA BINDING DOMAIN

Investigation the DNA binding ability of CIA5/CCM1 in *Chlamydomonas reinhardtii*

Abstract

The acclimation to limiting CO$_2$ conditions and induction of the CCM in *Chlamydomonas reinhardtii* (herein referred to as *Chlamydomonas*) appears to be regulated by a critical transcription regulator, CIA5 (or CCM1), often called the master regulator of the CCM. The *cia5* mutant completely lacks induction of the CCM and is lethal in very low CO$_2$ conditions, although it is viable under high CO$_2$ conditions and also grows more slowly than the wild type in air levels of CO$_2$. Also, most identified LCI (low CO$_2$ inducible) genes remain either uninduced, fail to be up-regulated or dramatically decreased in expression when *cia5* is exposed to limiting CO$_2$. By using I-TASSER, an on-line platform for protein structure and function predictions, the single amino acid mutation (54$^{th}$ His to Tyr) in *cia5* is predicted to change the conformation and zinc coordination of CIA5/CCM1 3-D structure, which may explain why zinc binding ability was lost in the mutated CIA5. We report the first to successful over-expression of full length of CIA5/CCM1 proteins in *E. coli*, confirmed by SDS-PAGE and Western immunoblots. We also used purified, full length CIA5/CCM1 proteins from *E. coli* to identify potential specific DNA-binding sequences using random binding site selection (RBSS), which, after 5-8 rounds of selection, identified a particular motif with core sequence GGTT that was enriched by this selection. Thus, the RBSS experiment showed direct evidence that CIA5/CCM1 protein can bind specific DNA sequences *in vitro*, which was confirmed using a gel mobility shift assay (GMSA) to demonstrate highly specific protein-DNA interaction with purified, full-length CIA5. In addition, using online software BioProspector from Stanford University, we identified a 9-bp GC rich (GGGGCGGGGG) motif from the promoters of a selection of CIA5 dependent genes, and demonstrated using GMSA that promoter fragments from genes *LCI5*, *LCI1* and *LCIB*, which contain the candidate sequence, showed highly specific protein-DNA interaction with CIA5, although the interactions were somewhat weaker than with the RBSS-identified sequence. We also attempted *in vivo*
experiments using constructs containing LCIB genomic DNA combined with LCI5, LCI1 and LCIB promoters, both wild-type or with a mutated GC rich motif, to complement ad1. Unfortunately, these were inconclusive, so failed to demonstrate CIA5 protein-DNA interaction in vivo, but also did not refute the in vivo DNA binding activity of CIA5/CCM1. Therefore, it is not clear whether the GC rich motif represents an essential binding site in vivo for CIA5 to regulate target gene expression. Nonetheless, this work clearly provides the first direct evidence that CIA5 can bind specific DNA sequences in vitro and thus open the way for more extensive in vivo experiments to determine whether the specific DNA-binding of CIA5 has any biological relevance in vivo.

**Keywords:** Chlamydomonas; CIA5/CCM1; DNA binding domain; RBSS; GMSA;

1. Introduction

Many photosynthetic organisms have developed strategies to increase the level of CO₂ around Rubisco, including the well-known C4 photosynthesis and Crassulacean Acid Metabolism (CAM) in terrestrial vascular plants (NiyoGi et al. 2015). The normal atmospheric CO₂ concentration is limiting to photoautotrophic growth of most plants and photosynthetic microorganisms, in part because of the inefficiency of the enzyme Ribulose-1, 5-bisphosphate carboxylase/oxygenase (Rubisco), which has both an oxygenase as well as a carboxylase function, as well as a low affinity for CO₂ (Long et al. 2004; Duanmu and Spalding 2011). Some C3-type aquatic photosynthetic microorganisms, including cyanobacteria and microalgae, which face an additional challenge of 10,000-fold slower diffusion rate for dissolved CO₂ relative to diffusion in air, have developed a very different mechanism to increase the CO₂ around Rubisco.

A different type of CO₂-concentrating mechanism (CCM) from C4 photosynthesis and CAM evolved in microalgae and cyanobacteria that is induced when the supply of dissolved inorganic carbon (Ci) for photosynthesis is limited. Much knowledge regarding the microalgal CCM has been gained from well-investigated eukaryotic organisms such as Chlamydomonas (Moroney and Ynalvez 2007;
When the availability of Ci (inorganic carbon, CO$_2$, and HCO$_3^-$) is limiting, the induced CCM acts via active Ci uptake systems across both the plasma membrane and the chloroplast envelope to increase internal Ci accumulation, and various carbonic anhydrases (CAs) function to dehydrate accumulated HCO$_3^-$ and provide elevated internal CO$_2$ concentrations to the pyrenoid, a Rubisco-containing micro-compartment of the chloroplast (MORONEY and YNALVEZ 2007; SPALDING 2008; WANG et al. 2011). The essential components of a CCM in aquatic organisms include active Ci uptake systems that transport outside Ci for intracellular HCO$_3^-$ accumulation, internal CAs that catalyze rapid interconversion of CO$_2$ and HCO$_3^-$, and specific internal compartments to provide a localized elevated CO$_2$ concentration for Rubisco (MORONEY and YNALVEZ 2007; SPALDING 2008; WANG et al. 2011). Even though the *Chlamydomonas* CCM has been extensively studied for many years, we still know little about the limiting CO$_2$ acclimation process.

A key regulatory mutant, *cia5*, was first identified as a High-CO$_2$-Requiring mutant (MORONEY et al. 1989); it can grow under high (5%) CO$_2$ and grows more slowly than the wild type in low (0.03-0.04%) CO$_2$, but it is unable to grow in very low (0.01%) CO$_2$ and appears to completely lack induction of the CCM. The mutated gene, *CIA5*, also named *CCM1*, yields two mRNAs CCM1-A and CCM1-B from a single gene by an alternative splicing between the 3rd and 4th exons (only 2aa difference, 183$^{rd}$ GR substituted to E. CCM1-A: 699aa; CCM1-B also named CIA5: 698 aa) (FUKUZAWA et al. 2001; XIANG et al. 2001). CIA5/CCM1 has been proposed to be a putting transcription factor; it has two non-typical zinc finger motifs (SOMANCHI and MORONEY 1999), which may represent a DNA-binding motif typical of transcription factors. It was confirmed that 2 mol zinc can bind to the N-terminal zinc-finger domain, and that if this protein is mutated, H$^{54}$Y, C$^{77}$V and C$^{80}$V, the zinc binding ability is lost (KOHINATA et al. 2008). CIA5 also contains a Gln-rich repeat which might participate in regulation of other eukaryotic transcription factors, and a Gly-rich region with unknown function.

Aside from being a critical upstream regulator of the CCM and other low CO$_2$ acclimation responses and likely requiring posttranslational activation in low CO$_2$, the details of CIA5 function remain undiscovered. We know very little about the genes it
directly regulates downstream (FANG et al. 2012), let alone anything about sequences that might be recognized by its putative DNA binding domain. A transcriptome comparison by Fang et al. (FANG et al. 2012) identified a massive impact of CIA5/CCM1 and CO$_2$ on the transcriptome and revealed an array of gene clusters with distinctive expression patterns that provide insight into the regulatory interaction between CIA5 and CO$_2$. Individual gene clusters responded primarily to CIA5, to CO$_2$, or to an interaction between the two (FANG et al. 2012).

Most transporter candidate genes and CA genes are under control of CIA5/CCM1, a transcription regulator coordinating expression of almost all low-CO$_2$ induced genes (MORONEY et al. 1989; FUKUZAWA et al. 2001; XIANG et al. 2001; MIURA et al. 2004; FANG et al. 2012). Because of the extensive connection of CIA5 to the regulation of most CCM-related gene expression, including regulation of all candidate Ci transporters, CIA5 is often called the “master regulator” of the CCM (FANG et al. 2012). The 76 kD CIA5 protein migrates as an approximately 90-100 kD protein during SDS-PAGE (WANG et al. 2005; KOHINATA et al. 2008), which suggests either aberrant migration or potential posttranslational modification of CIA5.

Generally, zinc-binding domains including zinc finger motifs have been shown to function as either sequence-specific nucleic acid recognition motifs or recognition motifs for protein-protein interactions (KOHINATA et al. 2008). It was assumed that the CIA5/CCM1 protein could be a DNA-binding protein, since CIA5/CCM1 impacts expression of so many genes and is located in the nucleus (KOHINATA et al. 2008). However, when only the zinc-binding domain (N-terminal region) was investigated in southwestern blot analyses using an anti-CCM1 antibody, no signals were detected for DNA-protein complexes, suggesting that CIA5/CCM1 may bind to some unidentified proteins rather than to DNA in the nucleus (KOHINATA et al. 2008). The research described here investigated the DNA binding ability of CIA5/CCM1 either in vitro or in vivo.

2. Materials and Methods

2.1 Cell strains and culture conditions
*Chlamydomonas* strain CW10 (cc849) was obtained from the *Chlamydomonas* Stock Center (University of Minnesota, St. Paul, MN), and LCIB-defective mutant *ad1* was generated by insertional mutagenesis (WANG and SPALDING 2006). Media and growth conditions for *Chlamydomonas* strains have been previously described (WANG and SPALDING 2006). All strains were maintained on CO₂ minimal plates and kept in high CO₂ (air enriched with 5% vol/vol CO₂) chambers at room temperature, under continuous illumination (50 μmol photons m⁻² s⁻¹). Liquid cultures for experimental use were grown on a gyratory shaker (180 rpm) under aeration in white light (approximately 100 μmol photons m⁻² s⁻¹). For experiments in which cells were shifted from high to limiting CO₂ (low CO₂, 350-400 ppm or very low CO₂, 50-100 ppm) conditions, cells were cultured in CO₂ minimal medium aerated with 5% CO₂ to a density of ≈2×10⁶ cells/ml and then shifted to aeration with the appropriate limiting CO₂ for various times. Very low CO₂ was obtained by mixing normal air with compressed, CO₂-free air.

### 2.2 Preparation of competent *E.coli* and transformation

Several 2-3mm diameter *E. coli* colonies were picked from a freshly streaked Super Optimal Broth (SOB) plate then dispersed in 3-5mL SOB medium and grown overnight. 50-100 μL of overnight culture was inoculated into 1L SOB medium (medium volume: flask volume should be from 1:10 to 1:30) and incubated at 37°C with moderate aeration (275rpm), until cell density reached 5-9×10⁷ viable cells per mL. The *E. coli* suspension was collected into centrifuge tubes and chilled on ice for 10-15 min, then the cells were pelleted by centrifugation for 15 min at 750-1000g, 4°C (SS34 rotator, 5min, 5000rpm). The resulting supernatant was decanted and any remaining supernatant was removed by micropipette. The cell pellet was resuspended in 1/3 volume of ice-cold buffer1 with moderate vortexing, and incubated on ice for 15 min. Cells were pelleted by centrifugation as above (SS34 rotator, 5 min, 5000 rpm), drained thoroughly, and resuspended in 1/12.5 original volume of ice-cold buffer2, then incubated on ice for 15min. The resulting cell suspension was aliquoted into chilled screw cap tubes or microcentrifuge tubes, frozen in liquid nitrogen, and stored at -80°C (recipes for buffer1, buffer2 and SOB are shown in Table S1).
For transformation, 1-10 ng DNA was added to 50-100 μL of competent cells, which were incubated on ice for 30 min, then heat shocked at 42°C for 60-90 sec, and returned to ice for 2 min. After sitting on ice, 900 μL SOB was added and the cells were incubated at 37°C with moderate aeration for 30-60 min. A 100 μL aliquot was plated on selective medium, the remainder was plated on another selective plate, and both were incubated at 37°C overnight.

2.3 CIA5 constructs for overexpression in *E. coli*

Full length *CIA5* coding DNA sequence (CDS) was amplified from a subcloned cosmid containing the full-length *CIA5* gene (Xiang et al. 2001), a gift from Dr. Donald P. Weeks (University of Nebraska, Lincoln), then inserted into an *EcoRI/SalI* cloning site in either pET28a or pET42a vector, kindly provided by Dr. Yanhai Yin (Iowa State University, Ames IA 50011), to generate final constructs pET28a_CIA and pET42a_CIA. These constructs were transformed (as described on 2.2 above) into *E. coli* strain BL21-CodonPlus (DE3) for overexpression. For affinity purification of fusion proteins, we used epitope purification of protein from plasmid pET28a, which adds a His-tag, and plasmid pET42a, which adds both a His-tag and an additional 220aa GST-tag that adds an additional 25kD to the molecular weight (MW) of the fusion protein. But even with IPTG induction, either none of the expected CIA5/CCM1 protein was seen (pET28a_CIA on Fig. S2) or it was expressed at the wrong size (pET42a_CIA, about 42kD with red star marked on Fig. S2). After carefully re-checking the sequence of the constructs we generated, we identified a “G” missing around 185nt-189nt (Fig. S3, upper panel) of the *CIA5* CDS we amplified from the cosmid. This error resulted in an earlier termination around 486nt and generated, together with the GST-tag (220aa), a shorter peptide (161aa), which explains why we detect a fusion protein with the incorrect size induced in pET42a_CIA (red star marked on Fig. S2).

To overcome the 1nt deletion (there is one “G” missing in the full length *CIA5* CDS which we amplified from the cosmid) described above, a primer 5’-TCC GGA CAT TCG GGG GAG T-3’, which contains unique restriction enzyme site *BasWI* (highlight as Italic red) near the mutation position was chosen to correct GGGG to
GGGGG, and another primer 5’-GTA **CAC GAG** CGT GCC GGG CAA-3’, containing a unique restriction enzyme site BssSI (highlight as Italic dark red), was used to make a shorter CDS, to minimize PCR errors. Then we performed 4-piece ligation (3 fragments for CIA5 CDS plus the vector itself) to regenerate final constructs pET28a_CIA5 and pET42a_CIA5.

For total protein analyses, cells were harvested and resuspended in a buffer containing 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 10 mM NaCl, 1 mM PMSF, 1 mM benzamidine and 5 mM DTT. Fusion proteins were purified by His-tag affinity purification according to the protocol from Novagen (Cat. No. 70666-3) by batch purification under native/denatured conditions. His-tag purified proteins were re-purified using Mag-Strep beads (IBA) by affinity purification against Strep-tag. Protein concentrations were measured using Bio-Rad protein assay kit (Bio-Rad Catalog Number 500-0006). Proteins were separated on 10% polyacrylamide gels as described previously (LAEMMLI 1970). Both anti his-tag antibody (Pierce, Cat. No. MA1-21315) and anti CIA5 antibody were used to detect overexpressed target CIA5 protein. Two CIA5 antibodies were used in this work: CIA5 antibody 1 was a gift from Dr. Donald P. Weeks (University of Nebraska, Lincoln NE), and was generated against a synthetic peptide containing a 19-aa sequence found in the C-terminus of CIA5 (**C-EHHRDHLLLDAETFRLLQSC**) (WANG et al. 2005), and CIA5 antibody 2 was generated against a synthetic peptide containing a 14-aa sequence near the N-terminus of CIA5 (**SAPRQTASKRNRTG-C**) (GenScript USA Inc. Piscataway, NJ), in which an extra "C" (high-lighted as bold red) was added to the C-terminus (or N-terminus) to facilitate conjugation.

**2.4 Random binding site selection**

Random binding site selection (RBSS) was performed as described (LIN et al. 2000) with slight modifications. The 16-nt random, double-stranded DNA oligonucleotides, which were kindly provided by Dr. Yanhai Yin (Iowa State University, Ames IA 50011), were generated using a random oligomer (5’- TGG AGA AGA GGA GAG TGG GC[N]16CTT TGC ATT CTT CTT CGA TTC CGG G-3’) and forward
and reverse primers with the sequences 5’-TGG AGA AGA GGA GAG TGG GC-3’ and 5’-CCC GGA ATC GAA GAA GAA TGC AAA AGA G-3’, respectively.

The amplification reaction was carried out using 5 µg of random oligomer, 10 µg of forward primer, and 10 µg of reverse primer for three cycles, with each cycle consisting of 1.5 min at 95°C, 2 min at 55°C, and 2 min at 72°C. The binding mixture (25 µl) contained 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 50 mM NaCl, 2 mM dithiothreitol [DTT], 5% glycerol, 0.5% Nonidet P-40 [NP-40], 10 µg of bovine serum albumin (BSA) per µl, 62.5 µg of poly (dl-dC) per ml, and 100 ng of purified CIA5/CCM1 protein. After incubation for 15 min, 10 µl of Mag-strep beads (IBA) were added to immobilize the protein-DNA complex by affinity against Strep-tag, and the mixture was incubated with constant rotation for 15 min at room temperature. Bound protein-DNA complexes were washed twice with 500 µl of cold binding buffer, without BSA but containing poly (dl-dC), by placing reaction tube in a magnetic separator and removing supernatant. The bound DNA was eluted at 50°C in 120 µl of elution buffer 1 (5 mM EDTA, 0.5% sodium dodecyl sulfate [SDS], 100 mM sodium acetate, 50 mM Tris-HCl [pH 7.6]), and the DNA was recovered by ethanol precipitation. The recovered DNA was amplified by PCR using 100 pmol of forward primer and 100 pmol of reverse primer for 15 cycles under the conditions described above. The purified PCR products were incubated with protein for another round of PCR-enriched random site binding selection. After five to eight rounds of selection, the protein-DNA complexes were separated by electrophoresis with a 5% polyacrylamide gel in 0.5 × Tris-borate-EDTA (TBE). The bound DNA was excised from the dry gel and eluted in 400 ml of elution buffer 2 (0.5 M ammonium acetate, 1 mM EDTA [pH 8.0]). Bound DNA was recovered by ethanol precipitation and amplified by PCR. The products were then cloned into pGEM®-T Vector (Promega, Cat. No. A3600) and subjected to sequence analysis (Lin et al. 2000).

2.5 Gel mobility shift assay

Protein-DNA gel mobility shift assays (GMSAs) were performed as described (Yin et al. 2005), and the work-flow is shown in Table S3. Oligonucleotide probes (the probe sequences for each candidate are listed in Table S4 and S5) were synthesized,
annealed, and labeled with $^{32}$P-$\gamma$-ATP using T4 nucleotide kinase. The binding reactions were performed in 20 µl binding buffer (25 mM HEPES-KOH [pH 8.0], 50 mM KCL, 1 mM DTT, and 10% glycerol) with about 0.1 ng labeled probe (20,000 cpm) and the indicated amount of proteins purified from *E. coli*, plus 0.1 mg/ml dl-dC or dA-dT. After 30 min incubation on ice, the reactions were resolved by 5% native polyacrylamide gels with 1X TGE buffer (6.6g/l Tris, 28.6g/l glycine, 0.78 g/l EDTA [pH 8.7]) for 2-3 hours (YIN et al. 2005). After electrophoresis, gels were dried on Whatman® 3MM filter paper after covering with plastic wrap and dried on a gel dryer. The dried gels were exposed to and analyzed using phosphorimaging instrumentation (Molecular Imager™ System, Bio-Rad).

2.6 Constructs for *ad1* complementation

To test whether CIA5 can bind the GGGGCGGG DNA motif *in vivo*, modified pBluescript II KS_LCIB plasmid containing *LCIB* genomic DNA (Fig. S1) was used to complement the *ad1* mutant. *LCIB* genomic DNA was inserted into this plasmid through *Sfu*I/*BamH*I site, and *LCI1*, *LCIB* and *LCI5* promoters (around 800bp ~ 1000bp upstream of +1) with either wild-type or mutated GGGGCGGGG motif (amplifying primers in Table S6) were inserted in place of the original promoter. The 5P and 3P primer pairs were used to amplify wild-type promoters from *Chlamydomonas* genomic DNA, and the 5P/MR or MF/3P primer pairs were applied to amplify fragments with mutations and used to generate the mutated promoter by fusion PCR. All promoters were inserted into pBluescript II KS_LCIB through *Not*I/*Sfu*I site to replace the original promoter and to make 8 final constructs named LCI1PB, LCI1PBM, LCIBPB, LCIBPBM, LCI5PB and LCI5PBM1/M2/M3 (M stands for mutation), which were all transformed into *Chlamydomonas ad1* cells by glass bead transformation performed as described previously (VAN and SPALDING 1999).

2.7 Gene expression analysis

RNA was extracted with an RNeasy Plant Mini Kit (Qiagen, Cat. No. 74904), and the RNA concentration was measured using an ND-2000 Nanodrop
spectrophotometer (Nanodrop Technologies). 1-5 μg total RNA was treated with TURBO DNA-Free™ kit (Life Technologies), and cDNA was generated from 0.5 μg - 1 μg of treated total RNA (Verso cDNA Synthesis Kit; Thermo Scientific). cDNA (25-50 ng) was used for reverse transcription polymerase chain reaction (RT-PCR) and for quantitative RT-PCR (qRT-PCR) with a CFX Connect™ Real-Time PCR Detection System (Biorad) using a SYBR green two-step quantitative PCR system (Quanta Biosciences). Primer sequences and efficiencies are provided as supplemental information (Table S6). For quantitative analyses, the CBLP gene was used as an internal control for normalization of qRT-PCR data (FANG et al. 2012). The relative transcript abundance in each sample is defined as ΔCt = Ct(target gene) - Ct(CBLP) to represent the difference between the transcript abundance of genes examined and the transcript abundance of CBLP.

2.8 Western immunoblot analysis

For western immunoblotting, total protein was obtained as described previously (DUANMU et al. 2009b). Proteins were separated by SDS-PAGE on 4-12% gradient polyacrylamide gels, and immunoblotting using an anti-His-tag antibody (Pierce, Cat. No. MA1-21315) or anti-CIA5 antibody 1 and anti-CIA5 antibody 2 was performed as described in the protocol from Bio-Rad Laboratories (catalog no. 500-0006).

3. Results

3.1 CIA5/CCM1 protein 3-D structure prediction

The I-TASSER server is an on-line platform for protein structure and function predictions. The server is in active development with the goal to provide the most accurate structural and functional predictions using state-of-the-art algorithms (ZHANG 2008; ROY et al. 2010; ROY et al. 2012). Both the wild-type CIA5 and cia5-point-mutation (H54Y) protein sequences were submitted to I-TASSER, which yielded predicted 3-D structures of the CIA5/CCM1 protein as shown in Fig. 1 (left is the wild-type protein, right is the cia5 protein, yellow balls structures indicate the putative C2H2
zinc binding motif). In cia5, the point mutation causes His-54 to be mutated into Tyr in exon2, which makes cia5 mutant unable to grow under very low CO₂ conditions and impairs the induction of almost limiting-CO₂ inducible genes in cia5 (including most putative Ci transporters and some CAs) (Miura et al. 2004; Fang et al. 2012). In the I-TASSER predicted CIA5/CCM1 3-D structure, this single amino acid difference H^{54}Y and loss of a coordinating Histidine is predicted to cause a CIA5 protein confirmation change, which could cause the mutant CIA5/CCM1 protein to lose zinc binding ability.

3.2 CIA5 overexpression in E. coli

Although the cia5 mutant has been studied for more than two decades, as far as we know, no one has successfully expressed the full length CIA5/CCM1 protein and used it to perform in vitro experiments. CIA5 expression of in Chlamydomonas is constitutive and stable but only at a very low abundance; so it is extremely difficult to purify enough CIA5/CCM1 protein from Chlamydomonas cells for in vitro experiments. We reasoned that if we could synthesize and purify large quantities of CIA5/CCM1 proteins in another system, we would be able to perform in vitro DNA binding experiments such as Random Binding Site Selection (RBSS) to test its DNA binding ability, or Gel Mobility Shift Assays (GMSAs) to test whether it can bind to the promoter regions of putative CIA5 target genes, etc.

The full length CIA5 coding sequence (CDS), when overexpressed from vectors pET-28a_CIA5 and pET42a_CIA5 with IPTG-induction in E. coli, yielded overexpressed CIA5 protein as detected by both SDS-PAGE (Fig. 2) and western immunoblot by using antibody against His-tag (Fig. 3). However, these constructs yielded only rare CIA5 protein products when purified from E. coli cells either by His-tag (pET28a_CIA5), or by His-tag or GST-tag (pET42a_CIA5) affinity purification, and included lots of either nonspecific or degraded product (Fig. S5), especially from pET42a_CIA5 (Fig. 4), and most over-expressed proteins were in the form of inclusion bodies (data not shown).

To increase the yield and purity of CIA5 protein, another plasmid pETMal was also used to regenerate CIA5 overexpression constructs. With these new constructs, which all contain both N-terminal His-tag and C-terminal Strep-tag (Fig. 5A), (His-tag
to be used for initial His-tag affinity protein purification, and Strep-tag for further purification), IPTG induction (0.3 mM IPTG, 4 h, 37°C at 250 rpm continuous agitation) yielded overexpressed CIA5/CCM1 protein in *E. coli* cells (Fig. 5B), which was confirmed by western immunoblotting with anti-CIA5/CCM1 specific antibody 1 and antibody 2. As shown in Fig. 5B, the overexpressed CIA5/CCM1 protein is easily degraded even in *E. coli* cells (indicated by the lower yellow arrows). Based on its size, the degraded partial CIA5/CCM1 (15-20 kD) is very close to the size of the N-terminal conserved CIA5 region (Kohinata *et al.* 2008). We modified the induction conditions to lower temperature, lower agitation speed and a longer induction (0.3 mM IPTG induction overnight at 16°C with 180 rpm continuous agitation), under which conditions most of the over-expressed protein was soluble and the yield of high purity CIA5, which was similar to that from pET28a-CIA5 (Fig. 4) from *E. coli*, was sufficient for *in vitro* experiments.

### 3.3 Random binding site selection (RBSS)

The master regulator of the CCM, CIA5/CCM1, was predicted to have a putative zinc biding motif in the highly conserved N-terminal domain (Fukuzawa *et al.* 2001; Xiang *et al.* 2001; Yamano *et al.* 2011; Blanc *et al.* 2012). It was also confirmed that 2 mol zinc can bind to the N-terminal domain, and that, if this protein was mutated to H54Y, C77V and C80V, the zinc binding ability was lost (Kohinata *et al.* 2008). However, so far there is no evidence showing that this zinc binding protein can interact with a specific DNA sequence. After successfully over expressing and purifying CIA5/CCM1 protein from *E. coli* by affinity His-tag and Strep-tag, we performed PCR-assisted DNA binding site selection using random oligonucleotides (Random Binding Site Selection: RBSS).

After 5-8 rounds of RBSS, performed as described in material and methods, specific sequences were enriched that bind to CIA5. The enriched binding sequences have a motif with a core sequence GGT or GGTT (Fig. 6, Table S2), two sequences containing GGTT occurred more than once (ACTAGGCAGGGTTACA occurred 6 times and GTGGCTTAGGTTGCAA occurred 4 times). These two showing high occurrence frequency in our results were later used to generate sequence logo.
(weblog from http://weblogo.berkeley.edu/logo.cgi), and also used for further experiments, including GMSA, which was further examined in gel mobility shift assays discussed below to confirm CIA5 binding.

### 3.4 Gel mobility shift assay

As demonstrated either in microarray or RNA-seq experiments, the expression of thousands of genes is regulated directly or indirectly by CIA5/CCM1 (Miura et al. 2004; Wang et al. 2005; Fang et al. 2012). However there is still no direct evidence showing that CIA5/CCM1 binds directly to any of the promoter regions from these identified genes to regulate their expression.

BioProspector is useful software to discover conserved DNA motifs in upstream regulatory regions of co-expressed genes (Liu et al. 2001). We chose fifteen candidate genes whose expressions are highly regulated by CIA5/CCM1 based on microarray data and RNA-Seq (Miura et al. 2004; Wang et al. 2005; Fang et al. 2012), selected promoter regions 100-2000-bp upstream from transcription start sites for each, and entered these 15 sequences into BioProspector, an online free motif finding software developed by Stanford University (http://ai.stanford.edu/~xsliu/BioProspector/). It was reported that some C2H2 type zinc finger proteins such as transcription factor Sp1 (specificity protein 1) strongly binds to G-rich elements such as GC (GGGGCGGGG) and GT/CACC boxes (GGTGTCGGG) (Kamiuchi et al. 1998; Philipsen and Suske 1999). We identified 8 genes (LCI1, LCI5, LCIB, LCIC, CCP1, CCP2, LHCSR1 and CAH3) (Table S3), whose promoter regions contain a highly conserved 9-bp, GC rich region (GGGGCGGGG). We chose 29-bp sequences (9-bp consensus sequence plus 10-bp sequences flanking each side) from each promoter as a set of DNA motifs on which to perform protein-DNA binding experiments using a gel mobility shift assays to detect protein-DNA interactions in vitro.

Protein-DNA gel mobility shift assays (GMSA) were performed as described (Yin et al. 2005) in material and methods using a work-flow outlined in Table S4. Only 3 of the 8 candidate genes showed a positive response in the GMSA indicative of a protein-DNA interaction (Data not shown), although even these three gave relatively weak responses. As indicated in Fig. 7, the selected DNA motifs from LCI1, LCI5 and
LCIB promoters exhibited a GMSA response indicating they bound to CIA5/CCM1. Increasing the concentration of each unlabeled DNA motif oligo itself as a specific competitive inhibitor decreased the GMSA signal (compared the panel “1”, “10” and “100” with panel “-“), suggesting the CIA5/CCM1 binding is specific for each tested motif. Also, when mutated LCI5 sequences (Table S5) were used as probes, no Protein-DNA interaction was detected (Fig. 7, M1 to M5).

We also used the DNA motif generated by RBSS in GMSA and found that the Protein-DNA binding signal was much stronger than for the BioProspector-defined motifs (Fig. 7, lower panel). In this case also, increasing the concentration of the specific, unlabeled DNA motif oligo as competitor decreased the GMSA signal, confirming that the CIA5/CCM1 binding is specific for the tested motifs.

Based on GMSA results, it appears that CIA5/CCM1 protein binds nucleotide sequences from the promoters of some low-CO2 inducible genes such as, LCI1, LCI5 and LCIB. Although the interactions appear rather weak, the impact of both unlabeled competitors and single nucleotide mutations suggested that the binding is quite specific. Nonetheless, CIA5 has a much stronger a still very specific interaction with the motif identified from the RBSS experiments.

3.5 The effect of replaced promoter on complementation with ad1 mutant

To test whether there is protein-DNA interaction in vivo between CIA5 and the GC rich motif identified in the BioProspector analysis (GGGGCGGGGG), modified constructs containing LCIB genomic DNA (Fig. S1) and LCI1, LCI5 and LCIB promoters with either wild-type or mutated motifs (Table S6) were used to complement the air-dier phenotype of the LCIB ad1 mutant. If the GC rich motif is vital and essential for CIA5 binding in vivo to regulate target gene expression, the mutated promoters should not allow CIA5 binding, as was seen for the LCI5-derived motif in vitro (Fig. 7). This predicted decrease in CIA5/CCM1 binding might be expected to prevent the constructs from complementing ad1, and might be expected to either decrease or increase LCIB expression in the complemented lines. The 8 constructs containing LCIB genomic DNA combined with either wild-type or mutated LCI1, LCIB and LCI5 promoters (the sequences used to amplify the promoter as shown in Table S6) were
transformed into \textit{ad1}, and all 8 constructs were able to complement \textit{ad1}, that is, all enabled \textit{ad1} transformants to grow under air-level \textit{CO}_2. \textit{LCIB} expression was assessed at both the transcriptional and translational levels in all transformants. Western immunoblot analysis using specific anti-\textit{LCIB} antibody (Fig. 8), showed \textit{LCIB} was expressed in all transformants with either the wild-type or mutated promoters but not in the \textit{ad1} mutant, which indicated that all constructs with alternative promoters (either wild type or mutated) successfully complemented the mutant. However, when compared with wild type CW10, the expressed \textit{LCIB} protein levels were lower than wild type in all transformant lines either under high \textit{CO}_2 or when induced with air-level \textit{CO}_2. \textit{LCIB} transcript expression results from both RT-PCR (Fig. 9) and qRT-PCR (Fig. 10) yielded very similar results to the western immunoblot results in confirming that all three promoters enabled expression of the \textit{LCIB} gene. In addition, both western immunoblots and qRT-PCR results indicated that all three promoters provided higher expression in air-level \textit{CO}_2 than in high-\textit{CO}_2, as would be expected if they functioned normally. However, there was no clear indication that any of the mutated promoters had a negative impact on \textit{LCIB} expression.

4. Conclusions and Discussion

Although the zinc-finger type transcription regulator, CIA5 (or CCM1), was identified more than two decades ago both by complementation of the \textit{cia5} mutant (MORONEY \textit{et al.} 1989), and by cloning from a tagged allele of \textit{cia5}, \textit{ccm1} (FUKUZAWA \textit{et al.} 2001; XIANG \textit{et al.} 2001), we still know little about the detailed function of CIA5. CIA5/CCM1 has been proposed to be a putative transcription factor; it has two non-typical zinc finger domains (SOMANCHI and MORONEY 1999), which may be a DNA-binding motif typical of transcription factors. Aside from being a critical upstream regulator of the CCM, the details of CIA5 function remain undiscovered. We know very little about whether it actually binds DNA or, if so, what sequences are recognized by its putative DNA binding domain or what genes it directly regulates downstream (FANG \textit{et al.} 2012).

Kohinata \textit{et al.} reported no detectable signals of DNA-protein complexes in southwestern blot analyses using anti-CCM1 antibody, which argues that CIA5/CCM1
Zn-finger domain might bind protein rather than DNA (KOHINATA et al. 2008). In this current research, we report the first successful overexpression (Fig. 2, 3 and Fig. 5) and purification (Fig. 4 and Fig. S5) of full-length CIA5 protein from *E. coli*. Using random binding site selection and gel mobility shift assay, we confirmed that CIA5 protein indeed can bind specific DNA sequences (Fig. 6 and 7), with the impact of specific competitors and mutation effects providing very strong evidence to support the protein-DNA specific binding.

Although our experiments clearly demonstrated *in vitro* DNA binding activity of CIA5, the southwestern immunoblot analysis performed by Kohinata *et al.* (KOHINATA *et al.* 2008) failed to detect any sign of specific DNA-protein binding. We might argue that one possibility to explain this apparent contradiction would be the higher sensitivity of our detection methods, since the sensitivity for western immunoblots is usually down to ng or pg, depending on different specific antibodies, while for $^{32}\text{P}$-based GMSA, the detection sensitivity can be as low as fg (OSBORN 2000). Another possible explanation for the different results is the difference in the CIA5 proteins used, since we used full-length CIA5 protein for both RBSS and GMSA, while Kohinata *et al.* used only the N-terminal Zn-finger domain for southwestern immunoblot analysis. It is possible, therefore, that the other domains of CIA5 might also be needed to form stable, specific DNA-protein complexes.

The *in vivo* experiments were performed and aimed to test whether *in vitro* results were biologically relevant. Although the *in vivo* results (Fig. 8, 9 and Fig. 10) failed to confirm expectations based on *in vitro* results, they also were inconclusive and thus cannot rule out *in vivo* DNA binding activity of CIA5/CCM1. Nonetheless, while it is possible that the *in vivo* process is more complicated than we imagined, it is certainly possible that the DNA motif we characterized might not represent vital and essential CIA5 binding sites important *in vivo* to regulate target gene expression.
References


SOMANCHI, A., and J. V. MORONEY, 1999 As Chlamydomonas reinhardtii acclimates to low-CO2 conditions there is an increase in cyclophilin expression. Plant Mol Biol 40: 1055-1062.


WANG, Y. J., and M. H. SPALDING, 2006 An inorganic carbon transport system responsible for acclimation specific to air levels of CO$_2$ in *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA **103**: 10110-10115.


Figure and Table Legends

**Fig. 1.** Predicted 3-D protein structures of CIA5/CCM1 by online software I-TASSER. Left is the wild type CIA5, right is CIA5 protein with the His$^{54}$ to Tyr$^{54}$ point mutation of cia5. Yellow marked atoms indicate the non-typical C2H2 zinc finger C$^{36}$C$^{42}$H$^{54}$D$^{60}$. In the predicted CIA5/CCM1 3-D structure as shown above, this single amino acid difference H$^{54}$Y and loss of a coordinating histidine is predicted to cause a CIA5 protein confirmation change, which may cause the mutated CIA5/CCM1 protein to lose zinc binding ability.

**Fig. 2.** SDS-PAGE detection of CIA5 overexpression in *E. coli*. Following IPTG induction (0.5 mM IPTG, 3hrs), we detected successfully overexpressed CIA5 in *E. coli*. The red arrows indicate the expected fusion protein appearing at the correct size, and the yellow arrow shows the partially degraded protein.

**Fig. 3.** Western immunoblot confirming the overexpressed CIA5 fusion protein in *E. coli*. Following IPTG induction (0.5 mM IPTG for 3hrs), we could detect a fusion protein with the expected size using antibody against the His-tag.

**Fig. 4.** SDS-PAGE showing CIA5 fusion protein affinity purified using the His-tag. Lanes A and B show IPTG induction before purification, and lanes E1, E2 and E3 show different elution fraction portion of the fusion protein after purification. Although pET42a_CIA5 yielded more fusion protein than pET28a_CIA5 when detected by SDS-PAGE (Fig. 2), it appeared to yield lower purity fusion protein either by His-tag affinity purification or GST-tag (data not shown).

**Fig. 5.** A: New His-Strep-tagged constructs for CIA5 overexpression in *E. coli*. New pET28a: His-Strep-tagged vector. New pET42a: His-GST-Strep-tagged vector. New pETMal: His-MBP-Strep-tagged vector. **B:** Detection of over-expressed CIA5 fusion protein by both SDS-PAGE and Western immunoblot. **Left panel:** SDS-PAGE indicating induced CIA5 proteins. **Middle panel:** western-immunoblot against CIA5 antibody 1 (from Dr. Weeks lab), which recognizes 19aa at the CIA5 C-terminus is detecting only full CIA5. **Right:** western-immunoblot against CIA5 antibody 2 (from Dr. Spalding’s Lab), which recognizes 14aa near the CIA5 N-terminus is detecting both full length and degraded CIA5. **Red arrows** indicate the overexpressed CIA5 protein
(full length), and **yellow arrows** indicate partially degraded CIA5 protein. (-, no IPTG) (+, 0.3 mM IPTG for 4 hours).

**Fig. 6.** Schematic diagram showing the process used for random binding site selection (RBSS), as described in material and methods. The random double-stranded DNA oligonucleotides library uses a 16 nt random oligomer flanked by two known sequences (5'- TGG AGA AGA GGA GAG TGG GC [N]16 CTC TTT TGC ATT CTT CTT CGA TTC CGG G-3'), which can be used as forward and reverse primers for PCR amplification with the sequences 5'-TGG AGA AGA GGA GAG TGG GC-3' and 5'-CCC GGA ATC GAA GAA GAA TGC AAA AGA G-3', respectively. After incubating the CIA5 protein-DNA mixture, CIA5 protein-DNA complexes were immobilized to Mat-strep beads (IBA) using the CIA5 Strep-tag, and washed twice. Bound DNA was eluted, precipitated, PCR amplified, and subjected to another round of RBSS. After five to eight rounds of RBSS, CIA5 protein-DNA complexes were electrophoretically separated and bound DNA precipitated and PCR amplified, cloned and sequenced. Of 88 recovered sequences, 46 contained the core sequences either GGT or GGTT. Two sequences containing GGTT occurred more than once (sequence ACTAGGCAGGGTTACA occurred 6 times and GTGGCTTAGGTTGCAA occurred 4 times) by RBSS and were used to generate sequence logo (weblog from http://weblogo.berkeley.edu/logo.cgi), which were used for further experiments, including GMSA.

**Fig. 7.** CIA5 gel mobility shift assays (GMSAs). The upper panel shows GMSA using oligonucleotide probes composed of 29-nt DNA sequences centered on the GGTT core motif identified in RBSS experiments. The lower panel shows GMSA using oligonucleotide probes composed of 29-nt DNA sequences from the LCI1, LCI2 or LCI3 promoters centered on the GGGGCGGGGG motif from the Bioprospector analysis. Either poly dI-dC or dA-dT was added as a non-specific competitor in all reactions, and various concentrations of the un-labeled probes themselves were used as specific competitors in individual reactions. Increased specific competitor concentration, as indicated, decreased the signal for CIA5 protein-DNA binding.

**Fig. 8.** Western immunoblot analysis of LCI2 expression among all ad1 transformants complemented with LCI2 under regulation by LCI1, LCI2 or LCI3
promoter sequences in either high CO$_2$ (H) or air-level CO$_2$ (A). LCI1PB/M, LCIBPB/M and LCI5PB/M1/M2/M3 indicate various ad1 transformants complemented with LCIB containing either LCI1, LCIB or LCI5 promoter (M stands for mutation).

**Fig. 9.** RT-PCR analysis of LCIB gene expression as transcript abundance among all ad1 transformants complemented with LCIB under regulation by LCI1, LCIB or LCI5 promoter sequences in either high CO$_2$ (H) or air-level CO$_2$ (A). LCI1PB/M, LCIBPB/M and LCI5PB/M1/M2/M3 indicate various ad1 transformants complemented with LCIB containing either LCI1, LCIB or LCI5 promoter (M stands for mutation).

**Fig. 10.** qRT-PCR analysis of LCIB gene expression as transcript abundance among all ad1 transformants complemented with LCIB under regulation by LCI1, LCIB or LCI5 promoter sequences in either high CO$_2$ (H) or air-level CO$_2$ (A). LCI1PB/M, LCIBPB/M and LCI5PB/M1/M2/M3 indicate various ad1 transformants complemented with LCIB containing either LCI1, LCIB or LCI5 promoter (M stands for mutation).
Fig. 1

Fig. 2
Fig. 3

Fig. 4
A. His-Strep-tagged CIA5 constructs.

B. His-Strep-tagged CIA5 overexpression.

Fig. 5
Conclusion:
After 5-8 rounds of selection, some particular sequences which can bind to CIA5 proteins and have the motif with core sequence GGTT enriched by this PCR assisted RBSS selection.

Fig. 6
Fig. 7

Fig. 8
Fig. 9

Fig. 10
Supplemental Information

**Table S1.** The recipe for making *E. coli* competent cells, including solutions used.

**Table S2.** CIA5 binding results from random binding site selection (RBSS). RBSS was performed as described in Material and Methods. After five to eight rounds of selection, protein-DNA complexes were separated by electrophoresis with a 5% polyacrylamide gel and bound DNA was recovered by ethanol precipitation and amplified by PCR. The PCR products cloned into pGEM®-T Vector yielded 88 sequences, about 46 of which contained the core sequences either GGT or GGTT and sequence log was also generated based on all 46 sequences.

**Table S3.** Sequences from the promoters of *LCIB, LCI1* and *LCI5* that were identified by Bioprospector and used for GMSA. Bold red text indicates the 9-nt GC rich motif identified in Bioprospector, which was flanked on each side by 10 nucleotide acids from each promoter. Also shown are mutations within the GGGGCGGGG motif indicated with blue text.

**Table S4.** Gel mobility shift assay (GMSA) working sheet. Each reagent was added to the reaction according to the order on the working sheet.

**Table S5.** The wild-type and mutant *LCI5* promoter sequences used for GMSA. The mutated nucleotide within the standard motif (red text) is indicated in purple text.

**Table S6.** Primers used to amplify *LCIB, LCI1* and *LCI5* promoters for *in vivo* ad1 mutant complementation experiments. Primer pairs 5P and 3P were used to amplify each of the wild type promoter sequences (about 800bp ~ 1000bp upstream of +1), the primer pairs 5P and MR or MF and 3P were used to amplify each partial promoter containing the desired mutation, then 5P/3P primer pairs were used to ligate each partial fragments together by fusion PCR to generate mutated promoters. The promoters were inserted into pBluescript II KS_LCIB through NotI/Sful to generate final constructs LCI1PB/M, LCIBPB/M and LCI5PB/M1/M2/M3 (M stands for mutation). NotI/Sful sites are highlighted as italic bold red, the conserved GC motifs are marked as bold red font, and mutated nucleotides are indicated in bold purple font.
Fig. S1. Map of pBluescript II KS_Ble LCIB, which was used for cloning the LCIB, LCI1 and LCI5 promoters in place of the HSP70/Rbcs2 promoter in the plasmid.

Fig. S2. Overexpression of CIA5 in E. coli. Full length CIA5 coding DNA sequence (CDS) was amplified from a subcloned cosmid containing the full length CIA5 CDS, and sub-cloned into E. coli expression vector pET28a (His-tag) and pET42a (His- and GST-tag) through EcoRI/SalI site to generate the final constructs pET28a_CIA and pET42a_CIA, then transformed into E. coli BL21-CodonPlus (DE3) (see material and methods). The expected fusion protein size should be more than 70kD and 100kD in pET28a_CIA and pET42a_CIA (220aa GST-tag), respectively. None of 8 transformants for each construct with or without IPTG induction, showed expected protein size. Only pET42a_CIA constructs shows an IPTG inducible response, but with the incorrect protein size (marked with red star, around 42 kD).

Fig. S3. Schematic showing reconstruction of the CIA5 to repair error. Upper panel indicates one G missing around 185 to 189 nt. Middle panel indicates the position and use of two unique restriction enzyme site to overcome the missing “G” error. Bottom panel shows the primers used to correct the sequence error.

Fig. S4. Three PCR fragments containing CIA5 CDS to be used for constructing full-length CIA5 CDS. These three fragments were purified (Qiagen Gel Extraction Kit, Cat. No. 28704) and digested with EcoRI/BsaWI, BsaWI/BssSI and BssSI/SalI, then ligated with EcoRI/SalI digested vector in a 4-fragments ligation to regenerate the pET28a_CIA5 and pET42a_CIA5 constructs.

Fig. S5. His-CIA5 purification from pET28a_CIA5. Lane 1: total extraction; Lane 2: flowthrough; Lane 3-4: elution; Lane 5-6: wash.
Table S1. The receipt for making *E. coli* competent cells

<table>
<thead>
<tr>
<th>Buffer 1</th>
<th>100mM RbCl</th>
<th>6.05g</th>
<th>RbCl</th>
<th>TOTAL: 500mL, adjust pH to 5.8 with acetic acid (Better to use diluted acetic acid). Filter sterilize through a prerinsed 0.22μm membrane</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50mM MnCl₂</td>
<td>4.95g</td>
<td>MnCl₂.4H₂O</td>
<td></td>
</tr>
<tr>
<td></td>
<td>30mM KAc</td>
<td>1.47g</td>
<td>KAc</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10mM CaCl₂</td>
<td>0.735g</td>
<td>CaCl₂.2H₂O</td>
<td></td>
</tr>
<tr>
<td>15% Glycerol</td>
<td>75mL</td>
<td>Glycerol</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Buffer 2</th>
<th>10mM MOPS</th>
<th>209mg</th>
<th>MOPS</th>
<th>TOTAL: 100mL, adjust pH to 6.8 with NaOH. Filter sterilize through a prerinsed 0.22μm membrane</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10mM RbCl</td>
<td>121mg</td>
<td>RbCl</td>
<td></td>
</tr>
<tr>
<td></td>
<td>75mM CaCl₂</td>
<td>1.1g</td>
<td>CaCl₂.2H₂O</td>
<td></td>
</tr>
<tr>
<td>15% Glycerol</td>
<td>15mL</td>
<td>Glycerol</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SOB</th>
<th>2%</th>
<th>Bactotryptone</th>
<th>Autoclave and then add, filter sterilized, 10mM MgSO₄ and 10mM MgCl₂.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.50%</td>
<td>Bacto Yeast Extract</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10mM</td>
<td>NaCl</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.5mM</td>
<td>KCl</td>
<td></td>
</tr>
</tbody>
</table>
Table S2. Sequence from RBSS

<table>
<thead>
<tr>
<th>Sequence 1</th>
<th>Sequence 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACAGGCAAGGTTACA</td>
<td>TCTCCCGGGTCAAT</td>
</tr>
<tr>
<td>ACAGGCAAGGTTACA</td>
<td>ACGGGATGGGCTGCA</td>
</tr>
<tr>
<td>ACAGGCAAGGTTACA</td>
<td>GCAGTCTGCTTGGG</td>
</tr>
<tr>
<td>ACAGGCAAGGTTACA</td>
<td>CCTGGTAGTTACGTTA</td>
</tr>
<tr>
<td>ACTAGGCAGGTTACA</td>
<td>ACCAGTCTGCTGGTA</td>
</tr>
<tr>
<td>ACTAGGCAGGTTACA</td>
<td>AGTGTGCTCAGTTA</td>
</tr>
<tr>
<td>GTGACTAAGGTTGCAA</td>
<td>GGTCAACAGGTATA</td>
</tr>
<tr>
<td>GTGACTAAGGTTGCAA</td>
<td>ACCAGTCTGCTGGTA</td>
</tr>
<tr>
<td>GTGACTAAGGTTGCAA</td>
<td>ATGGGCGAGAATA</td>
</tr>
<tr>
<td>GTGACTAAGGTTGCAA</td>
<td>TCAACGGTGATTACAA</td>
</tr>
<tr>
<td>GTGACTAAGGTTGCAA</td>
<td>TACCGGCGAGCACG</td>
</tr>
<tr>
<td>GTGACTAAGGTTGCAA</td>
<td>TAGAAGGTCTGTTG</td>
</tr>
<tr>
<td>GTGACTAAGGTTGCAA</td>
<td>TAGAAGGTCTGTTG</td>
</tr>
<tr>
<td>GTGACTAAGGTTGCAA</td>
<td>TCGCTGCTGTTG</td>
</tr>
<tr>
<td>GTGACTAAGGTTGCAA</td>
<td>CGTCTCAGCTTCG</td>
</tr>
<tr>
<td>GTGACTAAGGTTGCAA</td>
<td>AGTGGCGAGAAG</td>
</tr>
<tr>
<td>GTGACTAAGGTTGCAA</td>
<td>TAGAAGGTCTGTTG</td>
</tr>
<tr>
<td>GTGACTAAGGTTGCAA</td>
<td>AACAGTCTGCTGGG</td>
</tr>
<tr>
<td>GTGACTAAGGTTGCAA</td>
<td>TCAGTGGGCTGTA</td>
</tr>
<tr>
<td>GTGACTAAGGTTGCAA</td>
<td>AAGGTGTAGGGATAC</td>
</tr>
<tr>
<td>GTGACTAAGGTTGCAA</td>
<td>TTCTCGATCCAAAGT</td>
</tr>
<tr>
<td>GTGACTAAGGTTGCAA</td>
<td>GTGCAAGGTAGGATA</td>
</tr>
<tr>
<td>GTGACTAAGGTTGCAA</td>
<td>CATGGGCGAGAATAG</td>
</tr>
</tbody>
</table>

Table S2
<table>
<thead>
<tr>
<th>No.</th>
<th>Sequence Name</th>
<th>Bases</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LCI1 motif F</td>
<td>29</td>
<td>TGGCGAGTTGAGGGCGGGGTTGACATCGCA</td>
</tr>
<tr>
<td></td>
<td>LCI1 motif R</td>
<td>29</td>
<td>TGGCATGCAATCCCCGGCGCAACACTCGCA</td>
</tr>
<tr>
<td>2</td>
<td>LCIIB motif F</td>
<td>29</td>
<td>CTTGGGTCGCGGGGGCCTGTTGGCGGT</td>
</tr>
<tr>
<td></td>
<td>LCIIB motif R</td>
<td>29</td>
<td>ACGCCACGGCGCGCCCGGCCACCAG</td>
</tr>
<tr>
<td>3</td>
<td>LCIC motif F</td>
<td>29</td>
<td>TCAGGGGGCGTGGGGGGGGGGGCTAGAAGGG</td>
</tr>
<tr>
<td></td>
<td>LCIC motif R</td>
<td>29</td>
<td>CCGTTCTAGCCCCGGGCGGACGCCCTGA</td>
</tr>
<tr>
<td>4</td>
<td>CCP1 motif F</td>
<td>29</td>
<td>TGGACATCATGGCGGGGCGAGGCGGTTC</td>
</tr>
<tr>
<td></td>
<td>CCP1 motif R</td>
<td>29</td>
<td>GAATGCCGCTGGCCCCGGCATGTATTTCCA</td>
</tr>
<tr>
<td>5</td>
<td>CCP2 motif F</td>
<td>29</td>
<td>TGGACATCATGGCGGGGCGAGGCGGTTC</td>
</tr>
<tr>
<td></td>
<td>CCP2 motif R</td>
<td>29</td>
<td>GAACGCGCCTGGCCCCGGCATGTATTTCCA</td>
</tr>
<tr>
<td>6</td>
<td>LCI5 motif F</td>
<td>29</td>
<td>TTGAGGGAGCGGGCGGGGGCGATTCCTCC</td>
</tr>
<tr>
<td></td>
<td>LCI5 motif R</td>
<td>29</td>
<td>GGAGGAATGGCGCGGCCCCGGCTGCTCAGAA</td>
</tr>
<tr>
<td>7</td>
<td>Lhcsr1 motif F</td>
<td>29</td>
<td>TGCTGTGCAAGGGGGCAGGGGGGGGGAGGGGGGGG</td>
</tr>
<tr>
<td></td>
<td>Lhcsr1 motif R</td>
<td>29</td>
<td>CCCCCCCCCCTCCCCGGCGGCTGCCAGAN</td>
</tr>
<tr>
<td>8</td>
<td>CAH3 motif F</td>
<td>29</td>
<td>ACTGCCGGGCGGGGGGGAGGGGGGGGAGGGGGG</td>
</tr>
<tr>
<td></td>
<td>CAH3 motif R</td>
<td>29</td>
<td>GGGGGCCTCTCCCCGGCGGCTCCCCGCAAGT</td>
</tr>
</tbody>
</table>
### Table S4. Gel Mobility Shift Assay (GMSA) working sheet Date:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₂O</td>
<td>μl</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4×Buffer</td>
<td>μl</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>dl-dC or dA-dT</td>
<td>μl</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Protein (unlabeled)</td>
<td>μl</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antibody</td>
<td>μl</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incubate on ice for 10 min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probe (labeled)</td>
<td>μl</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>μl</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incubate on ice for 30 min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table S5. LCI5 motif mutant

<table>
<thead>
<tr>
<th></th>
<th>Forward sequence</th>
<th>Reverse sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>LCI5</td>
<td>TTTGAGGAGCGGGGCGGGGCTTTTCC</td>
<td>GGAGGAATGGCCCCGCTCTCAAA</td>
</tr>
<tr>
<td>M1</td>
<td>TTTGAGGAGCGGGGCTTTTCC</td>
<td>GGAGGAATGGCCCCGCTCTCAAA</td>
</tr>
<tr>
<td>M2</td>
<td>TTTGAGGAGCGGGGCTTTTCC</td>
<td>GGAGGAATGGCCCCGCTCTCAAA</td>
</tr>
<tr>
<td>M3</td>
<td>TTTGAGGAGCGGGGCTTTTCC</td>
<td>GGAGGAATGGCCCCGCTCTCAAA</td>
</tr>
<tr>
<td>M4</td>
<td>TTTGAGGAGCGGGGCTTTTCC</td>
<td>GGAGGAATGGCCCCGCTCTCAAA</td>
</tr>
<tr>
<td>M5</td>
<td>TTTGAGGAGCGGGGCTTTTCC</td>
<td>GGAGGAATGGCCCCGCTCTCAAA</td>
</tr>
</tbody>
</table>

### Table S5
### Table S6. Primers used to amplify promoters

<table>
<thead>
<tr>
<th>Primer ID</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>LCI1_5P</td>
<td>AAGGAAAAAA<strong>GGGCCGCC</strong> AACATCCGCTGCATTCCAGG</td>
</tr>
<tr>
<td>LCI1_3P</td>
<td>AAGGAAAAAA<strong>TTGGAATGCAAACTAGTTAATCCAC</strong></td>
</tr>
<tr>
<td>LCIB_5P</td>
<td>AAGGAAAAAA<strong>GGGCCGC</strong>TCGATGATGGGCAAGTTC</td>
</tr>
<tr>
<td>LCIB_3P</td>
<td>AAGGAAAAAA<strong>TTGGAATGCAAACTAGTTAATCCAC</strong></td>
</tr>
<tr>
<td>LCI5_5P</td>
<td>AAGGAAAAAA<strong>GGGCCGC</strong>TTACTCGCACTGCCAGC</td>
</tr>
<tr>
<td>LCI5_3P</td>
<td>AAGGAAAAAA<strong>TTGGAATGCAAACTAGTTAATCCAC</strong></td>
</tr>
<tr>
<td>LCI1_MF</td>
<td>GGGGGGGGGG TTGTTGTTGTTGTTGTTGTTGTTG</td>
</tr>
<tr>
<td>LCI1_MR</td>
<td>CAAATCCCTCCCCCAAC</td>
</tr>
<tr>
<td>LCIB_MF</td>
<td>TGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG</td>
</tr>
<tr>
<td>LCIB_MR</td>
<td>GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG</td>
</tr>
<tr>
<td>LCI5_MF1</td>
<td>AGCGGGGTTGGGGGCA</td>
</tr>
<tr>
<td>LCI5_MR1</td>
<td>TGGGGGACGGGGGCA</td>
</tr>
<tr>
<td>LCI5_MF2</td>
<td>AGCGGGGTTGGGGGCA</td>
</tr>
<tr>
<td>LCI5_MR2</td>
<td>TGGGGGACGGGGGCA</td>
</tr>
<tr>
<td>LCI5_MF3</td>
<td>AGCGGGGTTGGGGGCA</td>
</tr>
<tr>
<td>LCI5_MR3</td>
<td>TGGGGGACGGGGGCA</td>
</tr>
</tbody>
</table>
Fig. S1
Fig. S2
Fig. S3
Fig. S4

Fig. S5
CHAPTER 4. MINI RECOMBINANT CIA5

Investigating the function of recombinant mini-CIA5 and suppressor screening

Abstract

Acclimation to limiting CO\textsubscript{2} conditions and induction of the CO\textsubscript{2} concentrating mechanism (CCM) in \textit{Chlamydomonas reinhardtii} (herein referred to as \textit{Chlamydomonas}) appear to be regulated by the transcription regulator, CIA5 (or CCM1), the so-called “master regulator” of the CCM, which reportedly affects expression of over 2,000 genes. Mutants lacking CIA5 function, including the \textit{cia5} mutant, completely lack CCM induction. Most identified LCI (low CO\textsubscript{2} inducible) genes in \textit{cia5} remain either uninduced or fail to be up-regulated by limiting CO\textsubscript{2}. Spot tests indicated that a mini-CIA5 construct that combines highly conserved CIA5 N- and C-terminal domains from CIA5 could complement the \textit{cia5} growth phenotype, and a selection of CIA5-regulated, downstream genes in the mini-CIA5 complemented transformants exhibited the same patterns of expression, including CO\textsubscript{2} regulation, as seen in the wild type or in \textit{cia5} complemented with full-length CIA5. Since \textit{cia5} suppressor screening/selection by either insertional mutagenesis or UV and EMS mutagenesis failed to identify any putative suppressors other than revertants, and because multiple mini-CIA5 transformants were identified as capable of complementing \textit{cia5}, the apparent complementation of \textit{cia5} with mini-CIA5 is very unlikely to have resulted from phenotype suppression caused by additional introduced mutations or by insertion-site gene disruptions.

Although gene expression patterns were restored, the amplitudes of the gene expression changes in mini-CIA5 complemented transformants were intermediate between the wild type and \textit{cia5} mutant when evaluated by RT-PCR and qRT-PCR analysis. Also, the observed liquid-culture growth rate for mini-CIA5 complemented lines under either high or air-level CO\textsubscript{2} was similar to wild-type and full-length CIA5 complemented lines, except for slightly slower growth of the mini-CIA5 lines late in the growth curve. However, under very low CO\textsubscript{2} conditions the mini-CIA5 complemented lines exhibited growth rates intermediate between \textit{cia5} and the wild-type or full-length...
CIA5 complemented lines. Therefore, although the recombinant mini-CIA5 clearly can complement cia5 mutant and properly mediate CIA5-regulated and CO2-regulated gene expression similar to full-length CIA5, it apparently lacks some additional functionality of the full-length CIA5 that affects the magnitude of transcript abundance, the amplitude of gene expression changes, and the full growth potential, especially in very low CO2.

Key words: mini-CIA5, CO2 concentrating mechanisms, spot test, growth rate, RT-PCR, Chlamydomonas

1. Introduction

As an aquatic photosynthetic organism, Chlamydomonas uses a CO2-concentrating mechanism (CCM) to overcome not only the low atmospheric CO2 concentration but also the about 10,000-fold slower diffusion of CO2 in water relative to air (MORONEY and YNALVEZ 2007). Internal accumulation of Ci (CO2 or as HCO3−) in the CCM occurs against a large concentration gradient, so accumulation must occur as HCO3− because its permeability across lipid membranes is 1000-fold lower than that of CO2. Active Ci transport and accumulation plays a critical role in the Chlamydomonas CCM (MORONEY and YNALVEZ 2007; SPALDING 2008). However, Rubisco uses CO2 as substrate, so, along with Ci transporters, carbonic anhydrases (CAs) which catalyze interconversion of CO2 and HCO3−, also play important roles in CCMs (SPALDING et al. 1983a; COLEMAN and GROSSMAN 1984; MORONEY et al. 2011).

The Chlamydomonas CCM is induced by limiting CO2 concentrations, and the discovery of CCM-related genes has been based on identifying genes with elevated expression under limiting CO2 (lower than 0.05%) compared with high CO2 (1 to 5% CO2) (SPALDING and JEFFREY 1989; CHEN et al. 1997; SOMANCHI and MORONEY 1999; MIURA et al. 2004; YAMANO and FUKUZAWA 2009). The essential components of a CCM in the aquatic organisms include active Ci uptake and transport systems that transport external Ci for intracellular HCO3− accumulation, internal CAs that catalyze rapid interconversion of CO2 and HCO3−, and specific internal compartments to provide a localized elevated CO2 concentration for Rubisco (MORONEY and YNALVEZ 2007; SPALDING 2008; WANG et al. 2011).
Even though the *Chlamydomonas* CCM has been extensively studied, we still know little about the limiting CO\(_2\) acclimation process. Acclimation to limiting CO\(_2\) and induction of the CCM in *Chlamydomonas* appear to be regulated by the so-called master regulator, CIA5 (or CCM1) (MIURA et al. 2004; WANG et al. 2005). The mutant *cia5* was first identified as a High-CO\(_2\)-Requiring (HCR) mutant (MORONEY et al. 1989). The *cia5* mutant can grow under high (5%) CO\(_2\) and grows more slowly than the wild type in low (0.03-0.04%) CO\(_2\), but is unable to grow in very low (0.01%) CO\(_2\) and also appears to completely lack induction of the CCM (MIURA et al. 2004; WANG et al. 2005).

Also, induction of almost all limiting-CO\(_2\) inducible genes is impaired in the *cia5* mutant (including most putative Ci transporters and some CAs) when it is exposed to low CO\(_2\) (MORONEY et al. 1989; SPALDING et al. 2002; MIURA et al. 2004; WANG et al. 2005; FANG et al. 2012). The gene defective in the *cia5* mutant, CIA5, also named CCM1, yields two mRNAs *CCM1*-A and *CCM1*-B from a single gene by an alternative splicing between the 3rd and 4th exons (only 2aa difference, 183\(^{rd}\) GR substituted to E. CCM1-A: 699aa; CCM1-B also named CIA5: 698 aa) (FUKUZAWA et al. 2001; XIANG et al. 2001). CIA5/CCM1 has been proposed to be a putative transcription factor; it has two non-typical zinc finger motifs (SOMANCHI and MORONEY 1999), which may represent a DNA-binding motif typical of transcription factors. It was confirmed that 2 mol zinc can bind to the N-terminal zinc-finger domain, and, if this protein is mutated, H\(^{54}\)Y, C\(^{77}\)V and C\(^{80}\)V, the zinc binding ability is lost (KOHNATA et al. 2008). CIA5 also contains a Gln-rich repeat which might participate in protein-protein interaction involved in regulation of other eukaryotic transcription factors, as well as a Gly-rich region with unknown functions.

Aside from being a critical upstream regulator of the CCM and other low CO\(_2\) acclimation responses in limiting CO\(_2\), the details of CIA5 function remain undiscovered. We know very little about any sequences recognized by its putative DNA binding domain or the genes it may directly regulate downstream (FANG et al. 2012). It is reported that the presence or absence of CIA5 affects over 1,000 CO\(_2\)-responsive genes (FANG et al. 2012) and also the degree of Rubisco packing inside the pyrenoid (SINGH et al. 2014). However, CIA5 expression is constitutive and independent of CO\(_2\) concentration (WANG et al. 2005), and is reportedly incapable of
binding DNA directly, so may require additional activation systems and/or transcription factors, such as LCR1, to impact expression of the genes it regulates (YoshioKA et al. 2004). CIA5 seems to be conserved only among green algae (Volvox, Chlorella and Coccomyxa) (YAMANO et al. 2011; SINGH et al. 2014). The work reported here used a recombinant mini-CIA5 to complement the mutant cia5, and used RT-PCR and qRT-PCR to investigate the downstream gene expression in transformants complemented with this mini-CIA5. Additional suppressor screening experiments were also applied to study CIA5 function.

2. Material and Methods

2.1 Cell strains, culture conditions and growth

*Chlamydomonas* strain cw10 (cc849) was obtained from the *Chlamydomonas* Stock Center (University of Minnesota, St. Paul, MN) and the mutant cia5 was a gift from Dr. Donald P. Weeks (University of Nebraska, Lincoln), and *HCR209* (CIA5 disrupted mutant) generated by insertional mutagenesis (Van et al. 2001). Media and growth conditions for *Chlamydomonas* strains have been previously described (Wang and Spalding 2006). All strains were maintained on CO2 minimal plates and kept in high CO2 (air enriched with 5% vol/vol CO2) chambers at room temperature, under continuous illumination (50 μmol photons m⁻² s⁻¹). Liquid cultures were grown on a gyratory shaker (180 rpm) under aeration in white light (approximately 100 μmol photons m⁻² s⁻¹). For experiments in which cells were shifted from high to limiting CO2 (low CO2, 350-400 ppm or very low CO2, 50-100 ppm) conditions, cells were cultured in CO2 minimal medium aerated with 5% CO2 to a density of ≈2×10⁶ cells/ml and then shifted to aeration with the indicated limiting CO2 for various times. Very low CO2 was obtained by mixing normal air with compressed, CO2-free air.

For liquid-culture growth experiments, cells inoculated from agar plates were cultured in CO2 minimal medium aerated with 5% CO2 for about 2 days, then centrifuged at 1300 rpm for 5 min, and an appropriate aliquot used to inoculate experimental cultures at a starting density of 1× 10⁵/mL in a fresh 50 ml culture. Liquid cultures were grown in 250 ml flasks on a gyratory shaker (180 rpm) under aeration
with 3 CO₂ conditions: high CO₂, air-level CO₂ and very low CO₂, and were sampled daily for both cell density and chlorophyll concentration. Optical density at 750 nm (OD₇₅₀) was used to monitor algal cell density, since Sager and Granick found OD₇₅₀ to be linear over a range of 2×10⁵ to 1×10⁷ cells/ml (SAGER and GRANICK 1953), and that at this wavelength, chlorophyll absorbance does not interfere. Growth assays were performed in clear, flat-bottom 96-well microtiter plates monitored at 750 nm using a Synergy 2 Multi-Mode Plate Reader (BioTek Instruments, Inc) as described by Bernd and Cook (BERND and COOK 2002).

To investigate the function of CIA5, two conserved regions of the CIA5 CDS, the putative binding domain, nt 1-330, and the putative activation domain, nt 1282-1671 (Chapter 2), were fused to generate a recombinant mini-CIA5. This mini-CIA5 was then transformed into cia5 by glass bead transformation (VAN and SPALDING 1999) to test whether it can complement the lethal phenotype of cia5 in very low CO₂ and/or affect downstream gene expression normally regulated by CIA5. Full-length CIA5 CDS (F-CIA5) also was transformed into cia5 as a positive control. The primers used to construct mini-CIA5 and F-CIA5 are shown in Table S1. The CIA5 fragments generated by PCR using these primers were then subcloned through the NdeI/EcoRI multi-cloning site into pGenD_AphVIII plasmid, which is a modified version of the nuclear expression vector pGenD described by Fischer and Rochaix (FISCHER and ROCHAIX 2001).

2.3 Selection for suppressors and growth spot-tests

Glass bead transformations were performed as described previously (VAN and SPALDING 1999). Either the original mutant cia5 (point mutation) or a cia5 allele with a deletion, HCR209 (VAN et al. 2001), was used for suppressor screening. For insertional mutagenesis and suppressor screening, the cells were transformed with linearized plasmid pSI103 (SIZOVA et al. 2001) and spread onto CO₂ minimal plates containing 10 μg/mL paromomycin (Par). Plates either were first selected under very low CO₂ conditions followed by a screen for antibiotic resistance (Par⁰) or were directly grown under high CO₂ conditions followed by transfer of transformant colonies to very low CO₂ for suppressor selection.
For UV mutagenesis, mutants cia5 and HCR209 were treated with UV radiation in a UV DNA/RNA crosslinker (Stratalinker™ 1800, Stratagene) at 4x10^6 cells per dish for 0, 1, 2 or 5 minutes. Ethyl MethaneSulfonate (EMS), which produces random, chemically-induced mutations in genetic material by nucleotide substitution, was also used to screen for suppressors of cia5. Cell suspensions (cia5 and HCR209) at 10^7 cells/ml were incubated with EMS (Sigma-Aldrich, USA) at various concentrations, i.e. 20, 30 and 40 µl/ml, for 2 hours with gentle agitation. The UV-radiation or EMS treated samples were washed three times with CO_2 minimal medium and resuspended in 20 ml of fresh CO_2 minimal medium for 2 days, then spread uniformly on CO_2 minimal plates and kept under very low CO_2 for suppressor screening.

For spot tests of growth, actively growing cells were 5-fold serially diluted to similar cell densities in minimal medium, spotted (5 µL/spot) onto minimal agar plates, and grown in various CO_2 concentrations for around 9 days.

2.4 Gene expression analysis

RNA was extracted with an RNeasy Plant Mini Kit (Qiagen, Cat. No. 74904), and the RNA concentration was measured using an ND-2000 Nanodrop spectrophotometer (Nanodrop Technologies). 1-5 µg total RNA was treated with TURBO DNA-Free™ kit (Life Technologies) and 0.5 µg -1 µg of the treated total RNA was used to generate cDNA (Verso cDNA Synthesis Kit; Thermo Scientific). 25-50 ng of this cDNA was used for reverse transcription polymerase chain reaction (RT-PCR) and for quantitative RT-PCR (qRT-PCR). Primer sequences and efficiencies are provided in the supplemental information. The qRT-PCR was performed on a CFX Connect™ Real-Time PCR Detection System (Biorad) using a SYBR green two-step quantitative PCR system (Quanta Biosciences). For quantitative analyses, the CBLP gene was used as an internal control for normalization of qRT-PCR data (FANG et al. 2012). The relative transcript abundance in each sample is defined as \( \Delta C_t = C_t(\text{target gene}) - C_t(CBLP) \) to represent the difference between the transcript abundance of genes examined and the transcript abundance of CBLP.

2.5 Western immunoblot analysis
For western immunoblotting, total protein was obtained as described previously (Duanmu et al. 2009b). Proteins were separated by SDS-PAGE on 4-12% polyacrylamide gels, and detected by immunoblotting using an anti-FLAG-tag antibody (Santa Cruz Biotechnology; catalog no. sc-51590) as described in the protocol from Bio-Rad Laboratories (catalog no. 500-0006).

2.6 Chlorophyll measurement

The chlorophyll content was measured as described by Glaesener et al. (Glaesener et al. 2013) and the calculations were based on Porra et al. (Porra et al. 1989). Workflow for the spectrophotometric determination of the chlorophyll content of Chlamydomonas cells is shown in Fig. S1.

3. Results

3.1 Recombinant mini-CIA5 construct can complement cia5

Both mini-CIA5 and full length CIA5 CDS were subcloned into pGenD_AphVIII vector through Ndel/EcoRI (Fig. 1), then transformed into cia5 (=1×10⁸ cells) by glass bead transformation (Van and Spalding 1999). Cells were plated on either CO₂ minimal plates or CO₂ minimal with 10 μg/L Par, the plates were placed into either high CO₂ (aerated with 5% CO₂) growth chamber (plates with Par) or very low CO₂ (50-100 ppm CO₂) chamber (plates without Par). The Par⁺ transformants from the high CO₂ growth plates were transferred to very low CO₂ for complementation screening. The majority of Par⁺ transformant colonies obtained from high CO₂ either died or were unable to continue growth after transfer to very low CO₂ for about one month. After 4-5 weeks selection under very low CO₂, we obtained 5 Par⁺ transformants from the mini-CIA5 (named mini-CIA5) construct and 11 Par⁺ transformants from the full length CIA5 CDS construct (thereafter named F-CIA5). Colony-PCR was used to confirm the presence of the transgene, and spot-tests were applied to confirm the complementation by growth. Primer pair P1/P6 was used for colony PCR (Cao et al. 2009), and could only amplify a 435-bp genomic DNA fragment...
from the control lines (CW10, \textit{ad1} and \textit{cia5}), but could amplify both a 330-bp CDS fragment and the 435-bp genomic DNA fragment for transformants complemented with either mini-CIA5 or F-CIA5 (Fig. 2). Sequencing of both these bands from colony PCR in the complemented lines revealed that the genomic DNA band still retained the point mutation (data not shown), confirming that the complemented lines were not revertants. Unlike the \textit{cia5} mutant itself, \textit{cia5} transformants complemented with either mini-CIA5 or F-CIA5 were able to grow under very low CO\textsubscript{2}, just like the controls strains \textit{ad1} and CW10 did (Fig. 3). Unfortunately, Par\textsuperscript{R} was lost (presumably silenced) in all the complemented \textit{cia5} transformants after long selection under very low CO\textsubscript{2} conditions in the absence of Par.

\textbf{3.2 Downstream gene expression in mini-CIA5 transformants}

Although the mini-CIA5 construct can complement \textit{cia5}, whether it functions fully or partially with regard to all types of CIA5-regulated gene expression (\textit{FANG \textit{et al.} 2012}) is in question. Based on transcriptome data, over 1,000 genes appear to be affected mainly by CIA5 but only minimally by CO\textsubscript{2} in “CIA5 clusters” (\textit{FANG \textit{et al.} 2012}). A few genes from typical “CIA5 clusters”, especially cluster 1 genes showing high expression in WT but low expression in \textit{cia5}, and cluster 2 genes having the opposite expression pattern, as well as a few representative, well-characterized CCM-related genes, were selected to test the mini-CIA5 function by assessing the impact of mini-CIA5 complementation on their expression. Ten candidates from each of these 3 categories were chosen for initial screening (detailed information, such as gene names or transcript IDs and primers used for RT-PCR and qRT-PCR, are indicated in Table 1). All primers were tested by RT-PCR for 30 cycles in CW10 or \textit{cia5} (primers from cluster 2), and the amplification efficiency for each primer pair is shown in Fig. 4A. The primers for candidate genes with good amplification efficiency were used to perform both RT-PCR and qRT-PCR in cell lines, including CW10, \textit{cia5}, \textit{cia5} transformants complemented with mini-CIA5, and \textit{cia5} transformants complemented with F-CIA5, under high CO\textsubscript{2} (H) and air level CO\textsubscript{2} (A) conditions.

F-CIA5 transformants complemented with CIA5 full length CDS enable the complemented transformants to function very similarly to the CW10 WT strain with
regard to expression patterns and amplitudes of all candidate genes investigated (see F-CIA5 in Fig. 4B, 4C and 4D, and Fig. 5A, 5B and 5C). On the other hand, mini-CIA5 transformants complemented with mini-CIA5 carrying only a partial CIA5 CDS (nt 1-330 plus nt 1282-1671), although clearly complementing cia5 and showing similar expression patterns, did not function quite the same as either CW10 or F-CIA5-complemented lines, with regard to expression amplitude for any of the candidate genes for which CIA5 functions as either an apparent activator (cluster 1 plus some CCM-related genes) or an apparent repressor (cluster 2). When complemented with mini-CIA5, the expression of each of the candidate genes behaved as intermediate between cia5 and CW10 (see mini-CIA5 in Fig. 4B, 4C and 4D, and Fig. 5A, 5B and 5C). These results demonstrate that, although mini-CIA5 might complement cia5 and enable the transformants complemented with mini-CIA5 to grow under very low CO₂, mini-CIA5 apparently did not fully activate or fully repress downstream gene expression to the same extent that CIA5 did, and mini-CIA5 also did not enable the same impact amplitude as CIA5 did when both CIA5 and CO₂ affected expression of CCM related genes (see CCM related genes in Fig. 4B, 4C and 4D, and Fig. 5A, 5B and 5C).

3.3 Liquid growth experiments

We compared the growth rate and chlorophyll content of the complemented cia5 transformants (F-CIA5 and mini-CIA5) with those of CW10, ad1 and cia5 under high CO₂, air-level CO₂ and very low CO₂ conditions. Growth experiments showed patterns of photoautotrophic growth for cia5 transformants consistent with those observed in spot tests (Fig. 2). Active, 2-d-old air-acclimated cells were inoculated into liquid minimal medium with similar starting cell densities (1×10⁵ cells mL⁻¹), grown with aeration, and the cell densities measured by OD₇₅₀ daily at the same time of day for 10 d. Not much difference was noted among all strains grown under high CO₂ (Fig. 6A), except the maximum density of mini-CIA5 from day 8 forward was lower than the others. Also, in air-level CO₂, the cell densities for CW10, F-CIA5, mini-CIA5 and cia5 were all lower than those in high CO₂ but almost uniformly much higher than that of ad1 (Fig. 6B), except that the mini-CIA5 complemented line again was slightly lower
than the others from day 10 forward. However, in very low CO₂, all strains exhibited a sharply decreased growth rate and a much lower final cell density than in high or air-level CO₂, with CW10, \textit{ad1} and F-CIA5 showing similarly higher growth rates and final cell densities than \textit{cia5}, while mini-CIA5 exhibited a growth rate and final cell density intermediate between CW10 (or F-CIA5 or \textit{ad1}) and \textit{cia5} (Fig. 6C). Chlorophyll concentration also was measured in these cultures over the same time course, with the growth curves based on chlorophyll exhibiting a pattern similar to those for OD₇₅₀-based cell density (Fig. 6D-6F).

### 3.4 Screening/selection for \textit{cia5} suppressors

Screening and/or selection for \textit{cia5} suppressors was also performed following either insertional mutagenesis, UV mutagenesis or chemical mutagenesis. For insertional mutagenesis, the \textit{cia5} (≈1×10⁸ cells) was transformed with linearized plasmid pSI103 (SZOVA \textit{et al.} 2001) by either glass beads transformation (VAN and SPALDING 1999) or electroporation (SHIMOGAWARA \textit{et al.} 1998), and spread onto either CO₂ minimal plates or plates containing 10 μg/mL Par. Plates grown without Par were first subjected to selection under very low CO₂ conditions, then secondarily screened for Par\textsuperscript{R}, while plates initially subjected to Par\textsuperscript{R} selection directly were grown under high CO₂ conditions then secondarily subjected to very low CO₂ selection. Unlike other strains, the transformation efficiency was pretty low in \textit{cia5}. Beginning with ≈1×10⁸ cells, we found only about 25 Par\textsuperscript{R} transformant colonies by glass-bead transformation. Even the normally much higher electroporation transformation efficiency, which can yield more than 10⁵ transformants (SHIMOGAWARA \textit{et al.} 1998), only yielded less than 100 Par\textsuperscript{R} transformants from ≈1×10⁸ cells. None of the Par\textsuperscript{R} transformants obtained from high CO₂ by either glass-bead or electroporation transformation were able to grow after transfer to very low CO₂ for about one month. Plates initially subjected to very-low CO₂ selection, with or without Par, yielded no apparent suppressors. After about 6-8 weeks’ selection no colonies were found on plates with Par, and less than 10 colonies grew on plates without Par, but none of those colonies survived after transfer to Par-containing plates even under high CO₂.
For UV mutagenesis, cia5 and HCR209 treated with UV light for 0, 1, 2 or 5 minutes and plated on CO$_2$ minimal plates under very low CO$_2$ for suppressor screening, yielded no putative suppressors from the CIA5 disrupted mutant HCR209, and yielded only about 5 colonies from cia5 (data not shown). For EMS treated cia5 or HCR209, no putative suppressor colonies grew for mutant HCR209, and 13 putative suppressor colonies grew on cia5 plates (Fig. S2). However, all suppressor candidates from either UV or EMS mutagenesis of cia5 turned out to be revertants, in which the cia5 C$\rightarrow$T point mutation at position nt 431 was reverted to the wild-type C after either UV or EMS treatment (Fig. S3, not all data shown).

4. Conclusions and Discussion

The cia5 mutant was first identified by its slow growth in limiting CO$_2$ (MORONEY et al. 1989). This UV-induced mutant cannot grow under very low CO$_2$ and is defective in only a single gene, CIA5 or CCM1 (FUKUZAWA et al. 2001; XIANG et al. 2001). Based on transcriptome data, CIA5 appears to serve much broader and more extensive roles than indicated by the phenotype of cia5 (FANG et al. 2012). CIA5 is very likely involved in the upstream regulation of multiple physiological processes, and, although its own transcript abundance does not change, the presence/absence of CIA5 (and its potential activation/inactivation) appears to have a major impact on the expression of genes encoding many secondary regulatory genes, including those encoding transcription factors and signal transduction components involved in regulation of a number of processes (FANG et al. 2012). Within the CIA5 clusters described from transcriptome studies, such as cluster 1, reported to have increased expression in wild type, the presence of CIA5 increases transcript abundance for specific genes such that it appears to behave as an upstream activator of positive signaling pathways and/or other gene expression activators. On the other hand, in clusters, such as cluster 2, where the presence of CIA5 results in decreased expression of specific genes (decreased expression in wild type), CIA5 appears to act as an upstream repressor of specific genes and metabolic functions (FANG et al. 2012). The mini-CIA5 construct, which carries only a partial CIA5 CDS (nt 1-330 plus nt 1282-1674) was
transformed into cia5 to learn more about the domains of CIA5 required for regulation of downstream genes.

Based on selection of transformants in very-low CO₂, mini-CIA5 appeared able to complement the cia5 lethal phenotype in very low CO₂. Both spot tests (Fig. 2) and liquid growth rate experiments (Fig. 6) corroborated the observation that mini-CIA5 transformants, which contained only two small parts of the CIA5 CDS, were able to complement cia5 and enable growth under very low CO₂. All downstream candidate genes investigated here by either RT-PCR (Fig. 4) or qRT-PCR (Fig. 5) responded to mini-CIA5 in the same expression patterns as they did to F-CIA5, including clear CO₂-regulated expression patterns for those genes normally regulated by both CIA5 and CO₂. However, mini-CIA5 complemented transformants exhibited an apparently intermediate response between CW10 and cia5, in that mini-CIA5 apparently did not either fully activate or fully repress downstream gene expression, and exhibited somewhat intermediate growth phenotypes in liquid culture. These observations together demonstrate that the N-terminal, conserved, Zn-finger-containing, putative binding domain (nt 1-330 or aa 1-110) combined with the C-terminal, conserved, putative activation domain (nt 1282-1671 or aa 428-557) (Chapter 2, this dissertation) are sufficient to provide CIA5 function, including responsiveness to CO₂ concentration, enable mini-CIA5 to complement the cia5 mutant phenotype, and enable growth under very low CO₂, but that other CIA5 sequences missing from this mini-CIA5 may also be important to make a fully functional CIA5.

As previously reported, low CO₂ acclimation state of Chlamydomonas can be distinguished from the high CO₂ acclimation state by an acclimation related delay in culture growth, induction of limiting CO₂-regulated genes, smaller cell size and a markedly decreased photosynthetic $K_{\frac{1}{2}}$ (CO₂) (SPALDING 2009). This difference in growth rate and total chlorophyll content between high CO₂ and low CO₂ acclimated cells could be easily distinguished in this work at earlier growth stages of liquid culture growth studies, such as from day 3 to day 7, but with continued growth to about 10 days, this difference decreased, as expected (Fig. 6 A, B and D, E). In addition, the very low CO₂ acclimation state was easily distinguishable in this work from the low
CO₂ acclimation state by a much decreased liquid culture growth rate and much lower chlorophyll concentration (Fig. 6 C and F), as expected.

Screening or selection for cia5 suppressors by insertional mutagenesis, UV mutagenesis or EMS mutagenesis resulted in either no candidate colonies or in cia5 revertants that converted the point mutation back to its wild-type form, which further supports and strengthens the significant role of CIA5 in the CCM pathways. The functional importance of CIA5 may be one explanation to account for the lack of identification of any bona fide suppressors in this screen/selection. However, other contributing reasons, such as the cia5 mutant is difficult to transform, so the numbers of putative transformants screened by insertional mutagenesis was very low, and also that the conditions needed for selection under very low CO₂ for at least 7 weeks to two months, are so harsh that they may eliminate even some bona fide suppressors. Nonetheless, the complete lack of identification of any putative suppressors argues strongly that the observed complementation of cia5 by mini-CIA5 represents true complementation and not suppression caused by the transgene insertion site.
References


SOMANCHI, A., and J. V. MORONEY, 1999 As Chlamydomonas reinhardtii acclimates to low-CO₂ conditions there is an increase in cyclophilin expression. Plant Mol Biol 40: 1055-1062.


VAN, K., Y. WANG, Y. NAKAMURA and M. H. SPALDING, 2001 Insertional mutants of *Chlamydomonas reinhardtii* that require elevated CO$_2$ for survival. Plant Physiol **127**: 607-614.


WANG, Y. J., and M. H. SPALDING, 2006 An inorganic carbon transport system responsible for acclimation specific to air levels of CO$_2$ in *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA **103**: 10110-10115.


YOSHIOKA, S., F. TANIGUCHI, K. MIURA, T. INOUE, T. YAMANO et al., 2004 The novel Myb transcription factor LCR1 regulates the CO$_2$-responsive gene Cah1, encoding a periplasmic carbonic anhydrase in *Chlamydomonas reinhardtii*. Plant Cell **16**: 1466-1477.
Figure and Table Legends

**Fig. 1** Schematic diagram of mini-CIA5 construct. A: diagram of full length CIA5 CDS (F-CIA5) and also indicating the nt 1-330 and nt 1282-1671 regions. B: diagram of mini-CIA5 which combines both the conserved N-terminus (nt 1-330) and C-terminus (nt 1282-1671) through fusion PCR. Primer pairs P1/P3 and P4/P2 were used to amplify nt 1-330 and nt 1282-1671 respectively, then primer P1/P2 was used to join this two fragments together by fusion PCR. Primer P3 is complementary to P4. Primer pair P1/P5 was used to amplify CIA5 full length CDS and P1/P6 pair used for colony PCR (all primer sequences were shown in Table S1). C: detailed map of pGenD_AphVIII, which is derived from a modified version of the nuclear expression vector pGenD described by Fischer and Rochaix (FISCHER and ROCHAIX 2001), in which mini-CIA5 fragment was subcloned into the vector through NdeI/EcoRI multi-cloning site.

**Fig. 2.** Colony PCR assessing the presence of CIA5 CDS in transformants. The primer pair P1/P6 (Table S1) which can amplify either a 330 bp (CIA5 CDS as the template) or a 435 bp (CIA5 genomic DNA as the template) fragment was used to perform colony PCR. The transformants that carried either CIA5 full CDS (F-CIA5) or partial CDS (mini-CIA5) should have both the 330 bp and 435 bp PCR products, while the control lines CW10, ad1 and cia5 only contain genomic DNA template, so should amplify only the single 435-bp PCR product.

**Fig. 3.** Spot test growth of control lines (CW10, ad1 and cia5) and cia5 transformants (F-CIA5 and mini-CIA5) on minimal plates in high CO₂, low (air-level) CO₂ or VL-CO₂. Cells grown to logarithmic phase were diluted to similar concentrations, spotted (5 μl) on plates and incubated for 9 days under different CO₂ concentration conditions.

**Fig. 4.** RT-PCR analysis of candidate gene expression. A: PCR amplification efficiency for primers (as shown in Table 1) from different categories was tested by RT-PCR. Primers from each category with good amplification were used for RT-PCR (Fig. 4B, C and D) and qRT-PCR (Fig. 5A, B and C). RT-PCR analysis of candidate gene expression from different categories as transcript abundance among CW0, cia5
and cia5 transformants complemented with either CIA5 full length CDS (F-CIA5) or CIA5 partial CDS (mini-CIA5) under either high CO2 (H) or air-level CO2 (A) conditions. The names or assigned numbers for candidate genes from each category are listed in Table 1.

**Fig. 5.** qRT-PCR analysis of candidate gene expression from different categories as transcript abundance among CW10 (blue color), cia5 (red color) and cia5 transformants complemented with either CIA5 full length CDS (F-CIA5) (green color) or CIA5 partial CDS (mini-CIA5) (purple color) under either high CO2 (H) (solid fill) or air-level CO2 (A) (pattern fill) conditions. The names or assigned numbers for candidate genes from each category are listed in Table 1. Two replicates were performed for each strain.

**Fig. 6.** Cell density and total chlorophyll (Chl) concentration. Cell density was measured and monitored as the absorbance at 750 nm using clear, flat-bottom 96-well plates. Chlorophyll concentrations were measured by monitoring the absorbance at 750 nm, 663.6 nm and 646.6 nm using clear, flat-bottom 96-well plates, using the flow-chart and calculations shown in Fig 1S. Two replicates were performed for each measurement.

**Table 1.** Candidate genes and primer lists used to analyze gene expression in cia5 transformants complemented with either CIA5 full length CDS (F-CIA5) or CIA5 partial CDS (mini-CIA5). Candidate genes are from three different expression patterns: cluster 1, high in WT vs low in cia5 mutant, where CIA5 appears to behave as an activator; cluster 2, low in WT vs high in cia5 mutant, where CIA5 appears to behave as a repressor; all others are well-characterized CCM-related genes whose expression is both CO2 and CIA5 regulated.

**Fig. S1.** Flowchart for measurement the chlorophyll (Chl) concentration in Chlamydomonas cells including calculations based on Porra et al. (PORRA et al. 1989).

**Table S1.** Primers used in the present work to generate either F-CIA5 or mini-CIA5 constructs to complement cia5 as described in Fig. 1 legend. Primers P1/P5 were used to clone the full length CIA5 construct, and primers P1/P6 were used for colony PCR to detect the presence or intactness of each construct in the transformants.
**Fig. S2.** Putative suppressor colonies obtained from EMS treated *cia5*.

**Fig. S3.** Sequence information for the putative suppressors obtained from UV and EMS treated *cia5*. For *cia5*, which has a point mutation C→T at 431, all putative suppressors have a “G” (complement strand) instead of A.
Fig. 1
**Fig. 2**

![DNA gel electrophoresis result](image1)

**Fig. 3**

![Petri dish cultures](image2)

**High CO₂**

**Low CO₂**

**Very Low CO₂**
Fig. 4

A: Test the primer by RT-PCR

B: RT-PCR results from cluster 1

C: RT-PCR results from cluster 2

D: RT-PCR results from CCM related genes
Fig. 5
Fig. 6
<table>
<thead>
<tr>
<th>Cluster</th>
<th>Candidate genes name or ID</th>
<th>Primers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>517586 (unknown transmembrane )</td>
<td>F: AAAAGAAATCAAATGGCGGGG&lt;br&gt;R: GTGATAGAGGCTCAAAGG</td>
</tr>
<tr>
<td>2</td>
<td>520461 (unknown)</td>
<td>F: ACAAGTACACGGAGGCAAG&lt;br&gt;R: AGTCCGAGATGAAAGTTG</td>
</tr>
<tr>
<td>3</td>
<td>516299 (RNA binding protein )</td>
<td>F: TCGGCTTTGACCACTTACCC&lt;br&gt;R: TGCTTTGAGTGTCCTTGG</td>
</tr>
<tr>
<td>4</td>
<td>520777 (hypothetical protein )</td>
<td>F: GTGAACCTGTGAACAGGAGAG&lt;br&gt;R: CACCGGCAAAATCATTAATG</td>
</tr>
<tr>
<td>5</td>
<td>523829 (HYDG, Hydrogenase assembly factor)</td>
<td>F: AAAAGAAATCAAATGGCGGGG&lt;br&gt;R: GTGATAGAGGCTCAAAGG</td>
</tr>
<tr>
<td>6</td>
<td>517585 (SULTR2, sulfate transporter)</td>
<td>F: CTGGGAGAAGGTTAAAGAGGAC&lt;br&gt;R: ACAGCTCACAACTCAAG</td>
</tr>
<tr>
<td>7</td>
<td>526302 (calcium transporting ATPase)</td>
<td>F: GGGGTGGAATTTAAAAGAGG&lt;br&gt;R: GTCAATCTCCCTAAACG</td>
</tr>
<tr>
<td>8</td>
<td>514370 (Ca-dependent protein kinase)</td>
<td>F: AAGGATGCAATTTAAGGCAAG&lt;br&gt;R: CAAGGCTCCCTCAACTTAC</td>
</tr>
<tr>
<td>9</td>
<td>511727 (GPD3)</td>
<td>F: AAGGATGCAATTTAAGGCAAG&lt;br&gt;R: CAAGGCTCCCTCAACTTAC</td>
</tr>
<tr>
<td>Cluster 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>516705 (plasma membrane calcium-transporting ATPase)</td>
<td>F: AAAAGAAATCAAATGGCGGGG&lt;br&gt;R: GTGATAGAGGCTCAAAGG</td>
</tr>
<tr>
<td>2</td>
<td>520263 (PTB2, sodium/phosphate symporter)</td>
<td>F: ACAAGTACACGGAGGCAAG&lt;br&gt;R: AGTCCGAGATGAAAGTTG</td>
</tr>
<tr>
<td>3</td>
<td>519946 (PTB4, sodium/phosphate symporter)</td>
<td>F: ACAAGTACACGGAGGCAAG&lt;br&gt;R: AGTCCGAGATGAAAGTTG</td>
</tr>
<tr>
<td>4</td>
<td>521890 (Chl a/b binding domain)</td>
<td>F: AAAAGAAATCAAATGGCGGGG&lt;br&gt;R: GTGATAGAGGCTCAAAGG</td>
</tr>
<tr>
<td>5</td>
<td>512596 (glycoside-hydrolase-like protein)</td>
<td>F: AAAAGAAATCAAATGGCGGGG&lt;br&gt;R: GTGATAGAGGCTCAAAGG</td>
</tr>
<tr>
<td>6</td>
<td>517768 (serine protease inhibitor)</td>
<td>F: AAAAGAAATCAAATGGCGGGG&lt;br&gt;R: GTGATAGAGGCTCAAAGG</td>
</tr>
<tr>
<td>7</td>
<td>526006 (ABC transporter-like domain)</td>
<td>F: AAAAGAAATCAAATGGCGGGG&lt;br&gt;R: GTGATAGAGGCTCAAAGG</td>
</tr>
<tr>
<td>8</td>
<td>524684 (NAR1.4, formate nitrite transporter)</td>
<td>F: AAAAGAAATCAAATGGCGGGG&lt;br&gt;R: GTGATAGAGGCTCAAAGG</td>
</tr>
<tr>
<td>9</td>
<td>515108 (TST1, thiosulfate sulfurtransferase)</td>
<td>F: AAAAGAAATCAAATGGCGGGG&lt;br&gt;R: GTGATAGAGGCTCAAAGG</td>
</tr>
<tr>
<td>10</td>
<td>520668 (MTP3, CDF transporter, membrane protein)</td>
<td>F: AAAAGAAATCAAATGGCGGGG&lt;br&gt;R: GTGATAGAGGCTCAAAGG</td>
</tr>
<tr>
<td>CCM related genes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Cluster 8</td>
<td>LCI5</td>
<td>F: AAAAGAAATCAAATGGCGGGG&lt;br&gt;R: GTGATAGAGGCTCAAAGG</td>
</tr>
<tr>
<td>2 Cluster 14</td>
<td>LCI3</td>
<td>F: AAAAGAAATCAAATGGCGGGG&lt;br&gt;R: GTGATAGAGGCTCAAAGG</td>
</tr>
<tr>
<td>3 Cluster 14</td>
<td>LCI15</td>
<td>F: AAAAGAAATCAAATGGCGGGG&lt;br&gt;R: GTGATAGAGGCTCAAAGG</td>
</tr>
<tr>
<td>4 Cluster 15</td>
<td>LCIA</td>
<td>F: AAAAGAAATCAAATGGCGGGG&lt;br&gt;R: GTGATAGAGGCTCAAAGG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>7</td>
<td>LCI1</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>LCR1</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>CAH1</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>HLA3</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>CIA5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CBLP</td>
<td></td>
</tr>
</tbody>
</table>

Table 1
1. Pellet $1.5 \times 10^6$ cells

2. Resuspend the pellet in 1mL 80% acetone 20% methanol

3. Mix thoroughly by vortexing

4. Centrifuge 5 min at maximum speed, 4°C

5a. If the pellet is white, then proceed to step 6
5b. If the pellet is still green, go back to step 3

6. Measure reading at $A_{750\text{mm}}$, $A_{663.6\text{mm}}$ and $A_{646.6\text{mm}}$

7. Subtract $A_{750\text{mm}}$ from $A_{663.6\text{mm}}$ and $A_{646.6\text{mm}}$ readings

8. Calculate chlorophyll (Chl) content ($\mu$g/ml):
   - Chl a: $12.25 \times A_{663.6\text{mm}} - 2.55 \times A_{646.6\text{mm}}$
   - Chl b: $20.31 \times A_{646.6\text{mm}} - 4.91 \times A_{663.6\text{mm}}$
   - Total Chl: $17.76 \times A_{646.6\text{mm}} + 7.34 \times A_{663.6\text{mm}}$

Fig. S1
Table S1. Primers used for construct mini-CIA5 and colony PCR

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1 CIA5_5’-NdeI</td>
<td>GGAATTCATATGGAAGCCTTAGACGCgCA</td>
</tr>
<tr>
<td>P2 1671_3’-EcoRI</td>
<td>ACGCGAATTCCTAGCCGCCTCCTCCgG</td>
</tr>
<tr>
<td>P3 Fusion_R</td>
<td>AACACGGTGCAGGTTGTCGTTCTATTTGCTTGTTG</td>
</tr>
<tr>
<td>P4 Fusion_F</td>
<td>CCAGAACCGCAATAGAACGGAACCTGGGCAACGTTT</td>
</tr>
<tr>
<td>P5 CIA5_3’-EcoRI</td>
<td>ACGCGAATTCCTAATCGCAGGACTGCAAGCAGT</td>
</tr>
<tr>
<td>P6 CIA5_3’-330</td>
<td>CGTTCTATTTGCTTGCTGG</td>
</tr>
</tbody>
</table>

Table S1

Fig. S2
Fig. S3
CHAPTER 5. GENERAL SUMMARY

The *cia5* mutant was first identified by its slow growth in limiting CO$_2$. This UV-induced mutant is defective in only a single gene, *CIA5* or *CCM1*. This mutation is caused by a single nucleotide C to T transition resulting in the zinc finger motif His$^{54}$ mutated to Tyr$^{54}$, which disrupts the zinc binding ability. Mutants lacking CIA5 function, including *cia5*, completely lack CCM induction, although they are viable under high CO$_2$ conditions and only grow more slowly than wild type in air-level CO$_2$. In *cia5*, most identified LCI (low CO$_2$ inducible) genes remain either uninduced, fail to be up-regulated or dramatically decreased by limiting CO$_2$. It was reported that the expression of more than 1000 genes was regulated, either directly or indirectly, by CIA5. Based on transcriptome data (FANG *et al.* 2012), CIA5 appears to serve much broader and more extensive roles than indicated by the limiting-CO$_2$-sensitive phenotype of *cia5*. CIA5 is very likely involved in upstream regulation of multiple physiological processes, and, although its own transcript abundance does not change, the presence/absence of CIA5 (and its potential activation/inactivation) appears to have a major impact on the expression of genes encoding many secondary regulatory genes, including those encoding transcription factors and signal transduction components involved in regulation of a number of processes. Within the CIA5 clusters, such as cluster1, reported to have increased expression in wild type, the presence of CIA5 increases transcript abundance for specific genes such that it appears to behave as an upstream activator of positive signaling pathways and/or other gene expression activators. On the other hand, in clusters, such as cluster 2, where the presence of CIA5 results in decreased expression of specific genes (decreased expression in wild type), CIA5 appears to act as an upstream repressor of specific genes and metabolic functions. Although it is very important for induction of a functional CCM, so far only one CIA5/CCM1 ortholog was identified in a closely-related, multicellular green alga, *Volvox carteri*. These two CCM1s share a highly conserved N-terminal region containing zinc-binding amino acid residues (CIA5, aa 1 - 110) and a highly-conserved C-terminal region rich in acidic amino acid residues (CIA5, aa 428 - 557).
Although CIA5 clearly is a very critical transcription regulator that impacts expression of thousands of genes, including functional regulation of the very important CCM, very little was known prior to the research described in this dissertation with regard to how CIA5 might function to regulate all these genes. The research described here was therefore undertaken to better characterize this CCM master regulator and to address questions about whether CIA5 functions in direct gene activation, whether CIA5 binds DNA, and what domains are essential for CIA5 function.

In Chapter 2, the yeast two hybrid system helped us to identify a 109-aa region (aa 436 - 544) of CIA5 that exhibits auto-activation in yeast and overlaps with the highly-conserved 130-aa region (aa 428 - 557) in Volvox CCM1. We further focused on this conserved 130-aa C-terminal region and demonstrated that a region overlapping this conserved, auto-activation domain causes the CIA5 protein to exhibit abnormal migration in SDS-PAGE.

Auto-activation in yeast was demonstrated for either half of the identified 130-aa region, suggesting that this 130-aa region may represent two independent activation domains (aa 428 - 487 and aa 488 - 557). These putative CIA5/CCM1 activation domains were demonstrated in Chapter 2 to activate gene expression in Chlamydomonas when used to replace the activation domain of a designed transcription activator-like element (dTALE). The ability of this 130-aa putative activation domain or either half of this domain to function as an activation domain in Chlamydomonas suggests that this conserved C-terminal domain may function as an activation domain in CIA5.

Deletion of a CIA5 region containing the 130-aa C-terminal conserved putative activation domain, totally eliminated the ability of CIA5 to complement a cia5 mutant, which is consistent with this region having an obligate function in CIA5. This 130-aa, conserved C-terminal putative activation domain contains an abundance of acidic aa residues, and is included within the Ascl and NotI region (aa 399 - 642) responsible for abnormal SDS-PAGE migration of CIA5. Both observations are consistent with the identified 130-aa conserved C-terminal region acting as an acidic activation domain. This 130-aa, putative activation region may represent two independent activation domains (aa 428 - 487 and aa 488 - 557), since each half of the 130-aa domain could
function in auto-activation in yeast as well as in the dTALE system to activate gene expression in *Chlamydomonas*. However, deletion of only aa 430 - 487, approximately the first half of the 130-aa C-terminal, conserved domain, totally eliminates the ability of CIA5 to complement the *cia5* mutant, which argues that the second putative activation domain within this 130-aa region is not sufficient by itself to allow CIA5 function. Therefore, even though either half apparently can function in both yeast auto-activation and *Chlamydomonas in vivo* activation, the full domain is required for CIA5 function.

This present work represents the first demonstration that a conserved, acidic domain from the CCM master regulator, CIA5, can function as a transcription activation domain both in *Chlamydomonas* and in yeast. This conserved acidic activation domain also appears indispensable for CIA5 function, as judged by complementation of a *cia5* mutant, suggesting that CIA5 functions, at least in part, through gene direct activation of gene expression. Further research could explore the detailed sequence constraints within this activation domain, as well as to confirm unequivocally its requirement for CIA5 function.

Chapter 3 examines the protein-binding activity of CIA5. CIA5/CCM1 contains an N-terminal zinc finger domain that binds 2 mol zinc, and, when the zinc finger domain was mutated, both zinc binding ability and regulatory activity were lost. By making both His- and Strep- tagged CIA5/CCM1 constructs, I was the first to successfully over-express full length CIA5/CCM1 proteins in *E. coli*, as confirmed by SDS-PAGE and Western immunoblot. Purified full length CIA5/CCM1 proteins from *E. coli* were used to identify candidate nucleotide binding sequences by performing random binding site selection (RBSS), which yielded enrichment for a specific motif with the core sequence GGTT. Thus, the RBSS experiment demonstrated that CIA5/CCM1 can bind specific DNA sequences *in vitro*. I also used online software BioProspector (Stanford University) to identify a common 9-bp GC-rich (GGGGCGGGG) motif from the promoter regions of a selection of 8 *CIA5* dependent genes.

Analysis by gel mobility shift assay (GMSA) confirmed strong, specific interaction of CIA5 with the RBSS-identified core sequence, but demonstrated only 3
of 8 candidate CIA5-dependent gene promoters as showing a relatively weak, but specific, CIA5-DNA interaction. In summary, I have determined that CIA5/CCM1 can specifically bind nucleotide sequences from the promoters of low-CO$_2$ inducible genes such as, $LCI5$, $LCI1$ and $LCIB$, although the interactions are weak, and that CIA5/CCM1 also binds another, specific DNA core sequence identified in RBSS experiments.

Previous work using southwestern blot analyses to probe interaction of the CIA5/CCM1 zinc-finger domain with DNA reported no detectable signals, which was interpreted as indicating that the CIA5/CCM1 Zn-finger domain might bind protein rather than DNA. That work used only the N-terminal, zinc-binding domain expressed in *E. coli*, since they failed in attempts to express full-length CIA5/CCM1. In my current research, I report the first successful overexpression and purification of CIA5 protein from *E. coli*. This full-length CIA5 exhibited highly-specific DNA-binding activity for sequences identified both by RBSS and BioProspector experiments, confirming that CIA5 indeed can bind specific DNA sequences and providing very strong evidence to support protein-DNA specific binding by CIA5. Although the previously reported southwestern immunoblot analysis using the CIA5/CCM1 zinc-finger domain and anti-CCM1 antibody failed to detect evidence for DNA-protein interactions, that the success in our work might result from our use of full-length CIA5 and/or from the higher sensitivity of my GMSA detection methods relative to the southwestern blot analysis. Regardless of the reason for this apparent contradiction, the experiments described in this dissertation provide the first direct evidence that CIA5/CCM1 can bind specific DNA sequences *in vitro*.

I also performed *in vivo* experiments to test whether *in vitro* results were biologically relevant. Although the *in vivo* results failed to confirm expectations based on *in vitro* results, they also were inconclusive and thus also do not refute the *in vivo* DNA binding activity of CIA5/CCM1. Nonetheless, while the *in vivo* experiments were inconclusive, it is also possible that the DNA motifs I characterized might not represent vital and essential CIA5 binding sites important *in vivo* to regulate target gene expression. However, the *in vitro* results reported here do open the way for more
extensive *in vivo* experiments to determine whether the specific DNA-binding of CIA5 has any biological relevance *in vivo*.

In Chapter 4, the mini-CIA5 construct, which only carries a partial *CIA5* CDS, was transformed into *cia5* to learn more about the parts of CIA5 required for regulation of downstream genes. Previous reports had demonstrated that the N-terminal zinc-finger domain was required for CIA5 function, and results presented in Chapter 2 of this dissertation demonstrated that the conserved, acidic C-terminal activation domain was apparently required for CIA5 function. Therefore, I constructed a mini-CIA5 construct containing only the conserved N-terminal zinc-finger domain (nt 1-330) and the conserved C-terminal activation domain (nt 1282 - 1674) to test whether these two conserved domains alone could provide sufficient CIA5 function to complement the *cia5* mutant phenotype.

Based on transformant selection in very-low CO₂, mini-CIA5 appeared able to complement the *cia5* lethal phenotype in very low CO₂. Both spot tests and liquid growth rate experiments corroborated the observation that mini-CIA5 transformants, which contained only two small parts of the CIA5 CDS, were able to complement *cia5* and enable growth under very low CO₂. All downstream candidate genes investigated here by either RT-PCR or qRT-PCR responded to mini-CIA5 in the same way as they did to F-CIA5, although mini-CIA5 apparently exhibited an intermediate response between CW10 and *cia5* so may not fully either activate or repress downstream gene expression. These observations together demonstrate that the N-terminal, conserved, Zn-finger-containing, putative DNA-binding domain (nt 1 - 330 or aa 1 - 110) and the C-terminal, conserved, putative activation domain (nt 1282 - 1671 or aa 428 - 557) are sufficient to provide CIA5 function. However, although this mini-CIA5 clearly can complement *cia5* and enable growth under very low CO₂, other CIA5 sequences missing from this mini-CIA5 may also be important to make a fully functional CIA5. Further experiments assessing the *in vivo* function of other mini-CIA5 constructs containing an iterative combination of domains and sequences might enable identification of additional, important domains.

To summarize, I have 1) demonstrated that CIA5/CCM1 contains a C-terminal, acidic activation domain that functions in both yeast and *Chlamydomonas* and
appears to be required for CIA5 function; 2) confirmed that CIA5/CCM1 can bind specific DNA sequences \textit{in vitro}, suggesting that the N-terminal, zinc-finger domain required for CIA5 function might bind DNA \textit{in vivo}; and 3) demonstrated that a mini-CIA5 construct containing only the putative DNA-binding, N-terminal zinc-finger domain and the conserved C-terminal acidic activation domain is sufficient to provide CIA5 function. These data taken together argue that CIA5 functions as a transcription regulator able to directly activate target genes, and that CIA5 may directly bind DNA in the promoter regions of at least some of the genes regulated by this “master regulator” of the CCM. The tools generated in this work, including the successful \textit{E. coli} expression and purification of full-length CIA5 and the construction of a functional mini-CIA5, together with the identification of a functional CIA5 activation domain and specific CIA5-binding DNA sequences as candidates for \textit{cis}-elements, provide a very solid foundation for further work to define the function of CIA5 even better.