TALE activation and functional analysis of endogenous genes in Chlamydomonas reinhardtii

Han Gao
Iowa State University

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

Recommended Citation
Gao, Han, "TALE activation and functional analysis of endogenous genes in Chlamydomonas reinhardtii" (2014). Graduate Theses and Dissertations. 14142.
https://lib.dr.iastate.edu/etd/14142
TALE activation and functional analysis of endogenous genes in *Chlamydomonas reinhardtii*

by

Han Gao

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
Diane Bashham
Jeffrey Essner
Aragula Rao
Yanhai Yin

Iowa State University

Ames, Iowa

2014

Copyright © Han Gao, 2014. All rights reserved.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>ACKNOWLEDGEMENTS</strong></td>
<td>iv</td>
</tr>
<tr>
<td></td>
<td><strong>CHAPTER 1  GENERAL INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>General Introduction</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Literature Cited</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td><strong>CHAPTER 2  TALE ACTIVATION OF ENDOGENOUS GENES IN</strong></td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>CHLAMYDOMONAS REINHARDTII</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abstract</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Material and Methods</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Results and Discussion</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>Concluding Remarks</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>Literature Cited</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>Figure Legends</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td><strong>CHAPTER 3  EXPRESSION ACTIVATION AND FUNCTIONAL</strong></td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>ANALYSIS OF HLA3, A PUTATIVE INORGANIC CARBON TRANSPORTER IN**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CHLAMYDOMONAS REINHARDTII</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abstract</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>Results</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>Discussion</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>Material and Methods</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>Literature Cited</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>Figure Legends</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td><strong>CHAPTER 4  dTALE MEDIATED ACTIVATION OF β-CAROTENE KETOLASE AND</strong></td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>β-CAROTENE HYDROXYLASE IN ATTEMPT TO INCREASE THE AMOUNT OF**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ASTAXANTHIN IN THE GREEN ALGA IN CHLAMYDOMONAS REINHARDTII</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abstract</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>Results and Discussion</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>Material and Methods</td>
<td>109</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>Literature Cited</td>
<td>111</td>
<td></td>
</tr>
<tr>
<td>Figure Legends</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>CHAPTER 5  GENERAL SUMMARY</td>
<td>120</td>
<td></td>
</tr>
<tr>
<td>General Summary</td>
<td>120</td>
<td></td>
</tr>
<tr>
<td>Figure Legends</td>
<td>124</td>
<td></td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my major professor Dr. Martin Spalding, for not only guiding me through this research, but 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 have learned and grew tremendously in this lab, which would have been impossible without the kindness of Dr. Spalding.

My sincere appreciation also goes to my committee members Dr. Diane Bassham, Dr. Jeffrey Essner, Dr. Aragula Rao, and Dr. Yanhai Yin. 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. Bing Yang and Dr. Ting Li for being incredibly helpful in conducting this research together.

I am very grateful for the lab environment created by Dr. Yingjun Wang, Dr. David Wright, Dr. Dan Stessman, Wei Fang, Soujanya Akella, Bo Chen, Alfredo Kono, and many other former Spalding lab members. Spalding lab is the most supportive and friendly lab I have ever experienced or heard of. Our lab members, and especially Dr. Yingjun Wang and Dr. David Wright 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 years.
CHAPTER 1

GENERAL INTRODUCTION

General Introduction

For decades, the unicellular green alga *Chlamydomonas reinhardtii* (hereinafter referred to as Chlamydomonas) has been known as an important model organism in both fundamental biological studies and bio-product development. With the complete genomic sequence information and a wide variety of mutants of Chlamydomonas, researchers were able to make substantial breakthroughs in photosynthesis, flagella assembly, evolutionary relationships among plants, animals, and human beings, biomedicine and biofuel production, and many other fields (Grossman et al. 2004; Snell et al. 2004; Merchant et al. 2007; Ruppecht 2009; Takahashi et al. 2011; Rosales-Mendoza et al. 2012).

One particular focus of research in Chlamydomonas in our lab and of this thesis is to understand the unique CO\textsubscript{2} concentrating mechanism of this green alga.

Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) is the initial enzyme catalyzing photosynthetic carbon fixation in most known photosynthetic organisms. As is indicated by its name, in addition to RuBP, Rubisco uses both CO\textsubscript{2} and O\textsubscript{2} as substrates, and has low apparent affinity for CO\textsubscript{2} (Laing et al. 1975; Lilley and Walker 1975). The concentration of CO\textsubscript{2} in the vicinity of Rubisco therefore has a large impact on carbon assimilation, and RuBP oxygenation occurs when local CO\textsubscript{2} concentration is low, causing a wasteful side reaction. Multiple strategies have evolved to increase the CO\textsubscript{2} supply to Rubisco, including the C\textsubscript{4} and CAM (Crassulacean acid metabolism) photosynthetic pathways in terrestrial plants, both of which rely on subcellular compartments and an intermediate CO\textsubscript{2} fixation step resulting in a four-carbon product for final supply to Rubisco (Ranson and Thomas, 1960; Slack and
Hatch, 1967). A different type of CO₂ concentrating mechanism (CCM) exists in many aquatic photosynthetic microorganisms, including cyanobacteria and photosynthetic microalgae. While the CCM in cyanobacteria has been extensively studied and revealed (Badger and Price 2003; Price 2011; Kupriyanova et al. 2013), the functions of components in the eukaryotic CCMs remain largely unknown, partly due to the complicated subcellular structure of eukaryotic cells. One model microalgae, *Chlamydomonas reinhardtii* (hereafter, Chlamydomonas), provides researchers an ideal platform to study the CCM. Compared to many other green algae species, Chlamydomonas cells are easily grown under most laboratory conditions. The Chlamydomonas genome has been fully sequenced, and a series of mutants are also available, thanks to years of previous work. Despite all these advantages, a big picture of the Chlamydomonas CCM still calls for more in-depth investigation.

Genetics and molecular techniques such as mutagenesis, gene knockdown, targeted overexpression and gene activation have long been used in revealing the mechanism of the Chlamydomonas CO₂ concentrating system, and continues to facilitated breakthroughs in this field of research.

Transcription activator-like effectors (TAL effectors, or TALEs) are disease related proteins secreted by plant pathogenic genus *Xanthomonas* when infecting host plants (Zhu et al. 2000; Buttner and Bonas, 2002, 2006). These proteins function like transcription factors by recognizing and binding the upstream regions of host susceptible genes or resistance genes, causing different host responses (White and Yang 2009; Boch and Bonas 2010). TALE proteins are typically composed of three domains, with the central repeats mainly responsible for DNA recognition and the C-terminal domain responsible for gene activation. The sequence specific binding pattern between TALE central repeats and target DNA has been
successfully decoded in recent years, both computationally and structurally. These 34 amino acid residue repeats are nearly identical, with only the 12\textsuperscript{th} and 13\textsuperscript{th} residues being polymorphic, known as repeat variable diresidues (RVDs). Each RVD specifically interacts with one single nucleotide, and a combination of TALE repeats with different RVDs hence recognizes a specific DNA sequence (Boch et al. 2009; Moscou and Bogdanove, 2009; Deng et al. 2012). The true modular recognition of TALE repeats makes them promising tools for expression manipulation and genome editing. By simply replacing the original repeats with artificially designed ones, designed TALEs (dTALEs) have been used to activate target genes in rice, human cells, and other mammalian cells (Zhang et al. 2011; Li et al. 2013; Maeder et al. 2013). Fusion proteins of dTALE and endonuclease FokI have also been demonstrated to effectively interrupt target genomes at expected sites in plants, mammalian cells, zebra fish, and fruit flies (Li et al. 2011; Wood et al. 2011; Sander et al. 2012; Liu et al. 2012). Successful uses of these techniques in Chlamydomons, especially on genes involved in the CCM, will doubtlessly provide information for a better understanding of gene functions and CCM pathways.

The model organism \textit{Chlamydomonas reinhardtii}

As tiny as it might seem, the single-celled green algae \textit{Chlamydomonas} has played essential roles in fundamental biological studies as well as bio-product production in past decades (Grossman et al. 2004; Merchant et al. 2007; Corchero et al. 2013). Under easily achievable lab conditions, the doubling time of Chlamydomonas liquid cultures is less than ten hours, and the cells can be quickly synchronized and behave homogenously (Dent et al. 2001). The vegetative cells of Chlamydomonas are haploids, while the sexual cycle can be readily
induced for genetic analysis (Levine and Ebersold, 1960). As the whole genome sequence as well as other information became available, hundreds of insightful works have been conducted to reveal the coding secrets of these tiny cells (Merchant et al. 2007).

A unique feature of Chlamydomonas is its function to acclimate to various levels of inorganic carbon sources, via the system known as the CO₂ concentrating mechanism (CCM). Unlike terrestrial plants, aquatic photosynthetic species face environmental CO₂ concentrations that are not only often limiting but also changing drastically. The Chlamydomonas CCM relies on active uptake of more than one inorganic carbon (Ci) species, and allows for fast response and acclimation to different environmental Ci concentrations. Previous studies using insertional mutagenesis, suppressor screening, and RNAi technology have identified many important components of the CCM, including potential Ci transporters, carbonic anhydrases that catalyze rapid equilibrium between two major Ci species, and regulatory proteins that control the acclimation states of the CCM (Fukuzawa et al. 2001; Wang and Spalding, 2006; Spalding 2008; Duanmu et al. 2009; Wang et al. 2011). Different models have been suggested to integrate pathways and proteins involved in the CCM, while a thorough understanding of the system will benefit from more specific confirmation of the potential transporters and detailed examination into different acclimation states (Wang et al. 2011; Fang et al. 2012).

As the very basis of life on the earth, photosynthetic carbon assimilation has long been the focus of plant biology and agronomy research. With the ever growing population and the unprecedented need for food and fuel, improving the efficiency of Rubisco has become a more tempting topic. While many have focused on engineered Rubisco (Evans 2013),
maintaining sufficient supply to at the site of the enzyme might provide a more feasible and effective means to achieve optimized Rubisco performance.

**CCM in Chlamydomonas**

In the past, CCMs have been studied extensively in aquatic photosynthetic microorganisms, especially in cyanobacteria (reviewed in Price and Badger 2003; Price 2011; Kupriyanova et al. 2013). In recent years, since the completion of its three genomes as well as development of new genetic approaches, the unicellular green alga Chlamydomonas has facilitated major progress in understanding eukaryotic CCMs (Merchant et al. 2007; Spalding 2008; Wang et al. 2011).

Three species carbon sources co-exist in most aquatic environments, the ratio of which depending largely on the pH. The two major species of Ci in most naturally occurring environments are CO$_2$ and HCO$_3^-$, while CO$_3^{2-}$ is relatively rare, except in very alkaline environments. Evidence based on physiology experiments showed that direct uptake of both CO$_2$ and HCO$_3^-$ occurs at the plasma membrane as well as the chloroplast envelope (Sultemeyer et al., 1989; Badger et al., 1994; Palmqvist et al., 1994; Amoroso et al. 1998). The internal Ci pool is mainly composed of HCO$_3^-$, which is ultimately converted to CO$_2$ for the supply to Rubisco (Spalding et al. 1983a, Spalding and Ogren 1985, Spalding and Portis 1985, Moroney et al. 1985, 1987; Duanmu et al. 2009). Also, the ability of Chlamydomonas cells to quickly acclimate to limiting Ci conditions and to effectively accumulate Ci internally suggests rapid and tight regulation mechanisms underlying the system. Therefore, at least three essential components are required for the CCM: Ci transporters, carbonic anhydrases (CAs) that enables fast equilibrium between the two Ci species, and regulatory
proteins that control the on/off switch of the CCM. Through mutagenesis based screening and transcript profiling analysis, crucial components have been identified in the Chlamydomonas CCM, including potential Ci transporters, CAs, and regulatory proteins.

**Potential Carbon Transporters**

Two major physical boundaries need to be passed during the process of Ci accumulation: the plasma membrane and the chloroplast envelope. It has been reported that both CO₂ and HCO₃⁻ are actively taken up under limiting CO₂ conditions, with CO₂ being the preferred species at the plasma membrane (Sultemeyer et al. 1989). Given the low permeability of biological membranes to HCO₃⁻, it is reasonable to believe that HCO₃⁻ transport proteins are needed at both the plasma membrane and the chloroplast envelope. Although no specific protein has been identified as directly responsible for Ci uptake, many have been proven involved in Ci transport at either the plasma membrane or the chloroplast envelope, discussed here as potential Ci transporters.

Two proteins have been identified as possible Ci transporters across the plasma membrane, namely LCI1 and HLA3/MRP1, and four chloroplast envelope proteins have been identified as candidate Ci transporters, ycf10, CCP1, CCP2 and LCIA/Nar1.2.

LCI1 is a membrane protein with four transmembrane helices. Under limiting CO₂ conditions, the expression of *LCI1* transcript is significantly upregulated, indicating its potential role in the CCM (Burow et al. 1996; Yoshioka et al. 2004; Fang et al. 2012). Two regulatory proteins, CIA5 and LCR1, control the expression of LCI1, indicating its potential role in acclimation to limiting Ci. Impaired photosynthetic Ci affinity was observed in the *lcr1* mutant, which lacks a functional LCR1. LCI1 is located at the plasma membrane, and
when LCII was overexpressed in the lcr1 mutant, the cells showed increased Ci accumulation and photosynthetic affinity (Ohnishi et al. 2010). A recent study using dark-light cycles also showed that LCII is up-regulated in light, suggesting its function involved in photosynthesis (Tirumani et al. 2014). Therefore, LCII is considered potentially responsible for Ci uptake, but its precise role requires much further investigation (Burow et al., 1996; Yoshioka et al., 2004; Ohnishi et al. 2010).

HLA3/MRP1 is a putative ATP-binding cassette (ABC) type transporter that belongs to the multi-drug-resistance-related proteins subfamily (Gaedeke et al. 2001; Kruh and Belinsky, 2003; Hanikenne et al., 2005; Klein et al. 2006). The expression of HLA3 is also controlled by the master regulator gene, CIA5, and is drastically induced by limiting CO₂ conditions.

Membrane fractionation followed by immuno-blotting using HLA3 specific antibodies suggested that HLA3 is located in the plasma membrane of Chlamydomonas (Fei and Spalding, unpublished). When expressed in onion epidermal cells, HLA3-GFP chimeric protein is localized on the plasma membrane, agreeing with the membrane fractionation experiment (Fei and Spalding, unpublished). Co-knockdown of HLA3 and LCIB, a low CO₂ induced gene, results in severely decreased Ci uptake efficiency and growth rate, especially at alkaline pH where HCO₃⁻ is the dominant Ci species (Duamnu et al. 2009b). It is generally accepted that HLA3 is involved in the active transport of Ci, mainly responsible for HCO₃⁻ uptake.

With regard to potential plastid envelope Ci transport, ycf10 is a plastid gene whose open reading frame encodes a chloroplast envelope protein that shares sequence similarity with the plastid-encoded CemA product from plants and the pxcA gene from cyanobacteria. Knockdown of ycf10 causes decreased Ci uptake in Chlamydomonas, arguing for its possible
position in the CCM directly or indirectly, as its involvement in light-induced Na⁺ dependent proton extrusion similar to that of the pxcA gene product in cyanobacteria is likely (Rolland et al., 1997). CCP1 and CCP2 encode two almost identical chloroplast envelope proteins that share sequence similarity with the mitochondrial carrier protein superfamily, a family of small proteins involved in transport across mitochondria inner membrane of various metabolites (Spalding and Jeffery, 1989; Ramazanov et al., 1993; Chen et al., 1997). Knockdown of CCP1 and CCP2 did not affect Ci uptake significantly but hampered growth of Chlamydomonas cells in low CO₂ concentrations, indicating at least partial participation of these two proteins in the CCM (Pollock et al., 2005).

LCIA (NAR1.2) encodes a membrane protein that belongs to the Formate/Nitrite transporter (FNT) family that is involved in nitrite transport (Rexach et al., 2000; Galvan et al., 2002). Despite the common function of the NAR subfamily, LCIA expression is remarkably elevated with limiting CO₂, irrespective of the nitrogen source (Miura et al. 2004). The 366 amino acid gene product, LCIA, has six transmembrane domains and a predicted chloroplast transit peptide. Immunolocalization of LCIA located it at the chloroplast periphery, agreeing with the sequence-based prediction (Wang and Spalding, unpublished).

Photosynthetic Ci affinity in lcia mutant cells is very mildly affected (Wang and Spalding, unpublished). However, co-knockout of LCIA and LCIB, another low CO₂ induced protein, nearly abolishes Ci uptake and photosynthetic Ci affinity. In mutants lacking a functional LCIA, the photosynthetic affinity appeared most severely impacted at alkaline pH and under very low CO₂ (≤100ppm or 7µM), suggesting that LCIA is mainly involved in the active uptake of HCO₃⁻ under very low CO₂. Interestingly, when transformed into rice or Arabidopsis, LCIA overexpression leads to decreased size of the plants in general (Dong and
Spalding, unpublished). It was also reported that when expressed in *Xenopus* oocytes, LCIA showed low affinity for HCO$_3^-$ and high affinity for nitrite, leaving its actual role in Ci transport mysterious (Mariscal et al., 2006).

**Carbonic Anhydrases**

At least 12 CA isoforms or predicted CAs were identified in Chlamydomonas, including all three major lineages in evolution: 3 alpha CAs, 6 beta CAs, and 3 gamma (or gamma-like) type CAs (Reviewed in 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. 

*CAH1* was one of the first genes identified as limiting-CO$_2$ induced. Under limiting CO$_2$ conditions, the transcript level of *CAH1* increases drastically. The *CAH1* gene encodes a periplasmic CA, CAH1, and is proposed to convert HCO$_3^-$ to CO$_2$ at the cell surface to facilitate the CO$_2$ uptake across the plasma membrane (Moroney et al., 1985; Fukuzawa et al., 1990). Despite its low-CO$_2$ induced expression and periplasmic location, *CAH1* does not appear to be essential for a functional CCM. A *CAH1* deletion mutant showed no significant growth difference from wild type cells under limited CO$_2$ (Van and Spalding, 1999). Interestingly, another periplasmic CA, CAH2, shares high amino acid sequence similarity with CAH1. The expression of *CAH2* is repressed under limiting-CO$_2$ conditions (Rawat and Moroney, 1991), and its possible function(s) is not clear yet.

CAH4 and CAH5 are two mitochondrial CAs that share very high sequence similarities. Although their expressions are also elevated under limiting-CO$_2$ conditions, whether these two CAs directly participate in the CCM is still unclear (Eriksson et al., 1998). A few
hypotheses have been proposed to explain their possible functions, including converting CO$_2$ from respiration to HCO$_3^-$ which is then transported into chloroplast, buffering the mitochondrial matrix during photorespiration, and facilitating N assimilation via PEP carboxylase (Raven 2001; Giordano et al., 2003). However, the exact roles of these two CAs are still unclear.

CAH6 is reportedly a chloroplast CA whose expression is also up-regulated under limiting-CO$_2$ conditions. Its reported cellular location suggests its possible role in trapping unused CO$_2$ by Rubisco that diffuses from the pyrenoid and converting it to HCO$_3^-$ to prevent loss of Ci from the cell (Mitra et al., 2004, 2005).

CAH7 and CAH8 are also a pair of CAs that share sequence similarity. Both genes are constitutively expressed at the transcript and the protein levels, but are targeted to different cellular locations (Moroney and Ynalvez, 2007). CAH7 is predicted to be a chloroplast CA, while CAH8 was localized in the periplasmic space (Ynalvez et al. 2008). The location and constitutive expression of CAH8 provides a possible explanation for the seemingly non-essential role of CAH1, in that the defect in the $cah1$ mutant may at least be partially complemented by CAH8. CAH9 was predicted to be located in the cytoplasmic space. Unlike other CAs in *Chlamydomonas*, its sequence is more closely related with bacterial CAs. Its possible function has not been revealed yet (Mitra et al. 2005).

The gamma CAs, CAG1, CAG2, and CAG3, were identified as putative CAs in Chlamydomonas (Cardol et al. 2005). Gamma-CAs in higher plants are found in the complex I in mitochondria (Klodmann et al 2010). Although the function(s) of these gamma CAs in Chlamydomonas remain unclear, transcriptome analysis indicate that they are well-expressed (http:// genomes.mcdb.ucla.edu/Cre454/).
Among all CAs in *Chlamydomonas*, CAH3 is the only one proved essential for the CCM. Mutants lacking active CAH3 accumulate internal $\text{HCO}_3^-$ to high levels but are CO$_2$ limited in growth (Spalding et al., 1983b). The subcellular localization of CAH3 to the thylakoid lumen suggests its possible role in rapid conversion of $\text{HCO}_3^-$ to CO$_2$ at the site of Rubisco (Karlsson et al., 1998).

CAH3 is the only CA localized to the thylakoid lumen side in Chlamydomonas (Karlsson et al. 1998). Mutant lacking an active CAH3 (*cia3*) accumulate a high level of internal $\text{HCO}_3^-$, but is CO$_2$ limited in growth and photosynthesis (Spalding et al. 1983a). It was later demonstrated that during acclimation to limited CO$_2$ conditions, a significant portion of CAH3 protein rapidly become phosphorylated and relocated to the thylakoids that intrude into the pyrenoid, where Rubisco is concentrated (Blanco-Rivero et al. 2012). It is therefore reasonable to infer that CAH3 functions at the luminal side of thylakoids by catalyzing rapid conversion of $\text{HCO}_3^-$ to CO$_2$ for the supply to Rubisco.

A low CO$_2$ induced protein, LCIB, has long been known as crucial for the survival of Chlamydomonas cells under air level of CO$_2$ (Spalding et al. 1983b). LCIB mutants shows similar growth in high and very low CO$_2$ (<100 ppm), and die in ambient CO$_2$ (~400ppm, Wang and Spalding 2006). It has been proposed that LCIB, possibly together with another low CO$_2$ induced gene product, LCIC, in the form of a LCIB/LCIC complex, captures CO$_2$ or helps to maintain the CO$_2$ pool in the chloroplast stroma (Yamano et al. 2010; Wang and Spalding, 2006; Wang and Spalding, unpublished) Knockout of *CAH3* in the *LCIB* mutant partially rescues the air-die phenotype, suggesting an upstream position of *CAH3* relative to *LCIB* in the Ci transport pathway (Duanmu et al. 2009a), and also adds evidence to the interpretation of CAH3 as a CO$_2$ provider to Rubisco.
Multiple acclimation states and tight regulation of the CCM

One major regulator of the *Chlamydomonas* CCM is CIA5 (CCM1), the “master controller” of Ci acclimation (Fukuzawa et al., 2001; Xiang et al. 2001). CIA5 is a zinc-dependent transcription regulator that directly or indirectly regulates many CCM genes in *Chlamydomonas*. Mutants lacking CIA5 activity are unable to form typical pyrenoid structure in the chloroplast, and are deficient in CCM induction, with low Ci affinity and low Ci accumulation (Fukuzawa et al., 1998, 2001). Expression profiling and transcriptome data showed that CIA5 regulates the expression of hundreds of genes in Chlamydomonas, many of which are involved in the CCM, including *LCIA, HLA3, CAH3, LCI1*, and *CCP2* (Miura et al., 2004; Fang et al. 2013). Among these CIA5 controlled genes, a Myb transcription factor, *LCRI1*, was also shown to directly regulate the expression of limiting-CO$_2$ responsive genes *CAH1, LCI1* and *LCI6* (Yoshioka S et al., 2004).

*LCIB* and the *LCIB* gene family (*LCIC, LCID, and LCI6*) are limiting-CO$_2$ responsive genes in *Chlamydomonas* (Miura et al., 2004). Studies on air-dier mutants, *pmp1* and *ad1*, showed defects on these mutants in *LCIB* and revealed multiple acclimation states in *Chlamydomonas: pmp1* and *ad1* mutants are only lethal and deficient in Ci accumulation in air level CO$_2$ (0.4%-0.03%), but not in very low (less than 0.01%) CO$_2$ (Van et al. 2001; Spalding et al., 2002; Vance and Spalding, 2005; Wang and Spalding, 2006). LCIB is a soluble protein, thus it is unlikely that the protein is a Ci transporter at either the plasma membrane or chloroplast envelope. As was mentioned above, a second mutation in *CAH3* rescues the air-dier phenotype of *LCIB* mutants and allows them to grow in air level CO$_2$ (Duanmu et al. 2009). The epistatic relationship between *CAH3* and *LCIB* suggests that
LCIB functions downstream of CAH3, possibly acting as a trap of CO$_2$ that is converted by CAH3 to prevent Ci leakage out of chloroplast. Recent research has added more information to the mysterious roles played by LCIB. Under very low CO$_2$, the LCIB protein forms a ring shaped structure surrounding the pyrenoid, while under low CO$_2$, LCIB is localized in the chloroplast stroma in a randomly spread manner. Under air level CO$_2$, LCIB mutants exhibit a lethal phenotype, and the inhibition of photosynthesis due to limiting Ci is the most apparent when CO$_2$ is the dominant Ci species. LCIB/LCIA double mutants are lethal under very low CO$_2$, although the LCIA single mutation seems to have very mild impact on growth under very low CO$_2$ (Wang and Spalding, unpublished). Therefore, the main function of LCIB appears to be maintaining CO$_2$ uptake and supply, and it contributes to the CCM under very low CO$_2$, but is only essential under air level CO$_2$. These findings strongly suggest a system responding to multiple states of environmental Ci levels, with the sensor being either the CO$_2$ concentration or sensor protein(s), and rapid on/off transitions of limiting CO$_2$ induced genes. The fact that transcriptome analysis did not detect notable difference of the expression of these genes at the mRNA level (Fang et al. 2012), as well as the immediate inhibition of LCIA when the environmental CO$_2$ is elevated from very low to air level, suggests post-translation regulation might play an important role in the CCM.

**The role of pyrenoid**

A pyrenoid is the structure containing the majority of Rubisco that is observed in many photosynthetic microalgae species including Chlamydomonas. It has been considered an essential part of the CCM that shares functional similarity with the carboxysome of photosynthetic cyanobacteria (Moroney and Ynalvez, 2007, Schwarz et al. 1995, Price and Badger 2003). By compartmenting intracellular space and elevating the local concentration...
of Rubisco, pyrenoids allow the possibility of efficient CO2 supply to Rubisco and therefore carboxylation. The internal Ci pools examined in microalgae strains lacking a pyrenoid are around 10-fold lower compared to those have pyrenoids, despite their ability to induce the rest of the CCM (Morita et al. 1998). Although it is not known how pyrenoids are formed or structurally maintained, it has been reported that Rubisco, especially specific motifs in its small subunits, play an important role in pyrenoid assembly (Meyer et al. 2013).

The big picture

A promising model of the CCM system has been proposed based on identification of potential Ci transporters and epistatic relationships of limiting Ci responsive genes (Fig.1, Wang et al. 2011). While possibly interconverted by periplasmic CAs, CO2 and HCO3- pass through plasma membrane via diffusion and active transport (possibly by HLA3/LC11), respectively. Similarly, upon arrival in the cytosol, CO2 is able to diffuse across chloroplast envelope, while HCO3- has to be transported in to chloroplast stroma (possibly by LCIA), where it accumulates in the stroma 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 HCO3- permease into the thylakoid lumen to be quickly converted by CAH3 to CO2, which diffuses back out to the stroma as the substrate for Rubisco. The CO2 leaked from Rubisco/the pyrenoid is proposed to be trapped, or one-directionally converted to HCO3-, by the LCIB/LCIC complex and/or CAH6, and returns to the circulation via CAH3. As was briefly mentioned before, this suggested model is supported by the epistatic relationship between LCIB and CAH3, and by the observations
that during acclimation to limited CO2, part of the CAH3 protein is quickly phosphorylated and transferred to the thylakoids that intrudes into the pyrenoid, where most of the Rubisco molecules are found (Duanmu et al. 2009; Blanco-Rivero 2012).

**TAL Effectors and designed TALE mediated transcriptome/genome modification**

Transcription activator like effectors (TALEs) are virulence factors secreted by plant pathogen genus *Xanthomonas* during host infection. These proteins are delivered to host plant cells through type III secretion system. Upon injection into plant cells, TALs either activate the expression of host genes that make the plant more susceptible to pathogen infection (S genes), or, in fewer cases, genes that trigger plant resistance (R genes) (White and Yang, 2009; Bogdanove et al. 2010; Boch and Bonas 2010). This gene specific activation relies on the functions of each modular motif of TAL effectors, and especially the DNA recognition domain in the middle of the protein.

A typical TALE can be divided into three functional domains: the N-terminal domain that contains a translocation signal needed for infection, the C-terminal domain that contains a nuclear localization signal and an acidic activation motif, and the central repeats (also known as “canonical repeats”) that are the major contributors to TALE-DNA recognition specificity (Figure 2, Boch et al. 2009). A lot of effort has been invested in investigating the code of the recognition, and it was successfully deciphered (Boch et al. 2009; Moscou and Bogdanove 2009). The central region is composed of an array of nearly identical repeats of 34 amino acid residues, with the 12th and the 13th residues being polymorphic ones among repeats, known as repeat variable diresidues (RVDs). Each RVD specifically recognizes one nucleotide of the target DNA, thus an array of RVD combination is able to recognize a
specific DNA sequence, specifically referred to as effector binding elements (EBEs). The four most predominant RVD types in nature are NG, HD, NI, and NN, each recognizing T, C, A, and G (or A). Other types of RVDs also exist, including aberrant repeats where the 13th aa is missing and are often designated as N*; the N* RVD repeats are reported to recognize any nucleotide (Boch et al. 2009; Moscou and Bogdanove 2009).

The specificity and affinity of TALEs have also been extensively studied. Recent works on crystal structures have revealed more details in the binding pattern of TALEs. The canonical repeats wind around the sense strand of the target DNA as a right-handed helix, and each repeat forms a hairpin/loop structure, with the RVD residing in the inter loop position. The 12th aa residue interacts with the 8th residue within the loop and stabilizes the interloop position of the RVD, while the 13th protrudes into the major groove of the DNA strand, and directly interacts with the nucleotide (Mak et al. 2012; Deng et al. 2012; Murakami et al. 2010). This finding, together with the fact that RVDs HD and ND as well as RVDs NI and HI recognize the same nucleotides (Cong et al. 2012), indicates that it is the 13th, and not the 12th aa residue that determines the specificity of nucleotide recognition, which also explains the general tolerance of RVD N* for any nucleotide (Streubel et al. 2012).

The affinity of TALE binding, however, is not necessarily related to its specificity. Experiments with long stretches of synthesized TALE containing the same RVDs revealed that certain RVDs, such as HD, NN, and HN, make major contribution to the binding affinity, while others such as NI appear to have lower affinity contribution (Streubel et al. 2012). The N-terminal region immediately preceding the canonical repeats has also been proved crucial to TAL affinity. Despite the distinct aa compositions, structures similar to canonical repeats have been identified in this region, designated repeats -3, -2, -1, and 0.
These non canonical repeats, especially repeats -1 and 0 are required for effective DNA binding; a tryptophan within repeat -1 specifically interacts with nucleotide T immediately preceding the EBE, which is strictly conserved in all naturally found TALEs (Miller et al. 2011; Mussolino et al. 2011; Gao et al. 2012).

The true modularity of TAL effector interactions with individual nucleotides makes them ideal tools for DNA targeting, compared to other DNA targeting tools such as zinc finger proteins. Successful uses of TALEs or fusion proteins of TALE and other functional domains have been reported in targeted gene activation, genome editing, and recently epigenomic modifications (Li et al. 2012, 2013; Mahfouz et al. 2012; Maeder et al. 2013a).

TALE mediated transcriptome modification takes advantage of the original function of TAL effectors and allows activation of any desired target gene, only requiring substitution of the canonical repeat region of naturally occurring TALEs. Besides general concerns such as codon bias and chromatin structure in the target organism, it is reasonable to take into consideration factors that might affect the efficiency of such activation, including the position of the EBE relative to transcription initiation, the composition of the RVD array, and the length of the EBE. Given the importance of interaction between nucleotide T immediately preceding the EBE and the tryptophan in the N-terminal non-canonical repeat, an EBE should always be chosen with a thymine at position 0. In previous studies, artificially constructed TALEs without the N-terminal non-canonical repeats showed no measurable binding to the target EBE (Gao et al. 2012). It has been reported that EBEs surrounding the TATA box of target genes exhibited ideal activation efficiency (Li et al. 2011, 2012; Liu et al. 2013). While all naturally occurring TALEs include RVDs with high and low affinities and non-optimal RVDs (possibly to tolerate potential mismatches and mutations), stretches
of RVDs with lower affinity should be avoided when designing artificial TALEs. Although the increasing length of the EBE would doubtlessly increase the specificity of TALE across the genome, it is also noteworthy that the affinity of each RVD decreases from the N terminus to the C terminus within the TAL array (Garg et al. 2012; Meckler et al. 2013; Perez Quintero et al. 2013). Generally, TALE mediated activation has been demonstrated in a few organisms, including plants, human cells, and other mammalian cells. Over 10,000-fold increased target transcript level was reported in rice (Li et al. 2012). Highest activation reported was shown in human cells when the original C terminal domain was replaced with the VP16 activation domain from herpes simplex virus (Zhang et al. 2011; Bultmann et al. 2012). It was also reported that when a series of TALEs were targeted simultaneously to different EBEs of the same target gene in human cells, activation was largely increased in a synergistic pattern (Meader et al. 2013a).

TALEs fused with other functional proteins have been used in both transcriptome and genome modification. One such application has been in gene repression. The TAL-EBE binding near transcription initiation site apparently occupies the position, resulting in possible repression of genes downstream because of the physical blocking at the transcription initiation sites. When the C terminal domain of the TALE is fused with a repression domain or replaced with a repression domain, expression of the target genes has been shown to be notably reduced (Mahfouz et al. 2012). Recently, fusion proteins of dTALEs and demethylase or methytransferase have also been successfully used to repress the expression of target genes in human cells (Maeder et al. 2013b).

TALE-nuclease, or TALEN, is probably the most widely applied technique since the decoding of TALE and recognition of the potential of synthetic RVD arrays. A fusion protein
of TALE and the endonuclease FokI was used in many model organisms including rice, zebrafish, drosophila, C. elegans, and mammalian cells (Li et al. 2012; Sander et al. 2011; Liu et al. 2012; Wood et al. 2011). Guided by two TAL domains recognizing two EBEs separated by a short spacer, the two FokI nucleases dimerize and thus introduce double strand breaks, which are subsequently repaired by nonhomologous end joining or homologous recombination to mediate specific gene disruptions or specific gene modifications, respectively.

TAL effectors are unique proteins created by nature and have been proved as powerful tools in genetic and molecular studies. Although the prokaryotic system CRSPR/cas9 has gained a lot of research attention regarding genome editing in recent years given its simplicity, TALE-based genome modification tools still have some advantages, because of their specificity and reportedly low off-targeting effect, as well as their utility in gene activation and/or repression.

**Transgenes in Chlamydomonas: Achievements and Challenges**

Because of its simple structure and high growth potential, Chlamydomonas has been used as a model organism and a potential host for bioproduct generation. Years of study have revealed profound amounts of information on gene functions and metabolic pathways in Chlamydomonas. In recent years such studies have mainly utilized reverse genetic approaches, such as insertional mutagenesis, and forward genetic approaches, such as overexpression and/or repression tools. These classical molecular methods, however, rely on transgene expression in Chlamydomonas cells and can often be labor intensive. While transformation serves an effective method to generate transgenic clones, the expression level
of transgenes in Chlamydomonas has often been disappointing. Transgenes introduced via transformation integrate into the Chlamydomonas genome in a random manner (Zhang et al. 2014). Because of the unknown nature of constraints on transgene expression, it is difficult to obtain clones that express transgenes at a high level (Fuhrmann et al., 1999; Schroda et al. 2000). Various strategies have been used to overcome these expression limits, such as modifying the transgene sequences to match the high GC content and the codon bias of Chlamydomonas genome (Fuhrmann et al. 2004; Shao and Bock 2008), introducing multiple ORFs relying on the T2A peptide (Rasala et al. 2012), and selecting for strains that can potentially maximize transgene expression and minimize the effect of transgene insertion position (Neupert et al. 2009). However, no general solution has been established yet. In addition, mutant screening based on random mutagenesis can take years of work, and gene repression using methods such as RNAi tends to be highly variable, relatively transient and decreasingly effective over generations, despite the integration of transgenes into the genome. Therefore, both dTALE mediated gene activation and genome editing with TALENs provide promising potential means for research of gene functions in Chlamydomonas given their specificity and low-level of required transgene expression.

Dissertation Organization

This thesis is written in the format to include a published paper (chapter 2) that demonstrates robust activation of two endogenous genes, ARSI and ARS2, mediated by dTALEs in Chlamydomonas. The activation of a potential Ci transporter, HLA3, and its possible functions in the CCM is described as a manuscript ready for submission in Chapter 3. Chapter 4 discusses TALE induced activation of BKT and CHYB, genes that can potentially lead to elevated expression of astaxanthin and provide biomedical and commercial uses, as
well as advantages, drawbacks and future directions of the method. The dissertation is then concluded with a general summary in Chapter 5.


Buttner D, Bonas U (2002). Getting across--bacterial type III effector proteins on their way to the plant cell. EMBO J., 21, 5313-5322.


Cardol P, Gonzalez Halphen D, Reyes-Prieto A, Baurain D, Matagne RF, Remacle C (2005). The mitochondrial oxidative phosphorylation proteome of *Chlamydomonas*


Pollock SV, Prout DL, Godfrey A, Lemaire SD, Moroney JV (2004). The *Chlamydomonas reinhardtii* proteins CCP1 and CCP2 are required for long-term growth, but are not necessary for efficient photosynthesis, in a low-CO₂ environment. Plant Molecular Biology 91: 505-513.


Throughput genotyping of green algal mutants reveals random distribution of mutagenic insertion sites and endonucleolytic cleavage of transforming DNA. Plant Cell http://dx.doi.org/10.1105/tpc.114.124099.


Figure Legends

Figure 1. Wang and Spalding, 2011. Hypothetical schematic model of the Chlamydomonas CCM with emphasis on inorganic carbon (Ci, CO₂, and HCO₃⁻) transport. The proposed transport of Ci from the extracellular space to the thylakoid lumen via the cytosol and stroma is illustrated. The carbonic anhydrases (red ovals) catalyze the interconversion of CO₂ and HCO₃⁻ in respective compartments. Non-membrane-bound subcellular structure pyrenoid (boundary labeled as dashed line) is within the stroma and is the site of Rubisco localization. Blue ovals indicate HCO₃⁻ transporters and these include HLA3, LCI1, and LCIA. Additional potential HCO₃⁻ transporters (light blue circles) or HCO₃⁻ facilitated diffusion channels (light blue barrels) are indicated with question marks. Putative CO₂ gas channel and proton extrusion from stroma are operated by RHP1 (purple barrel) and ycf10/CemA (purple oval), respectively. Stroma localized LCIB/LCIC complex is proposed to be involved in rehydration of CO₂ released from thylakoid lumen into the HCO₃⁻ pool in the stroma, possibly including involvement of the putative stromal CA, CAH6.

Figure 2. Adapted from Boch et al. 2009. Model for DNA-target specificity of TAL effectors. (A) TAL effectors contain central tandem repeats, NLSs, and an AD. Repeat
variable di-residues (RVD) 12 and 13 are shaded in gray. (B) RVD amino acid residues at position 12 and 13 of the 17.5 repeat TALE AvrBs3 are aligned to the target gene EBE.

Figure 1.
Figure 2.
CHAPTER 2 TALE ACTIVATION OF ENDOGENOUS GENES IN CHLAMYDOMONAS REINHARDTII

A paper published in Algal Research 5: 52-60
Han Gao, David A. Wright, Ting Li, Yingjun Wang, Kempton Horken, Donald P. Weeks, Bing Yang, Martin H. Spalding

The plasmid library used to assemble artificially designed TALE repeats in this work was a generous gift from Dr. Bing Yang’s lab, contributed by Dr. Ting Li. The Chlamydomonas expression vector backbone was constructed by Dr. David Wright.

Abstract

Background
Transcription activator-like effectors (TALEs) are effector proteins secreted by certain plant pathogenic bacteria when infecting their hosts. Upon translocation, TALEs bind via a well-deciphered recognition code to specific sequences in the promoter region of targeted host genes, thereby activating expression of those targeted genes. Gene activation induced by artificially designed TALEs (dTALEs) has been studied in multiple organisms, but is yet to be demonstrated in green algae, such as Chlamydomonas reinhardtii, a well-known model organism for fundamental biological studies, as well as for biofuel production.

Results
We chose two endogenous Chlamydomonas genes, ARS1 and ARS2, as the targets of dTALE induced activation. Both genes encode a periplasmic arylsulfatase (ARS), and are located in an apparent tail-to-tail inverted duplication on chromosome 16. dTALEs independently
targeting $ARS1$ and $ARS2$ promoters were generated and successfully expressed in Chlamydomonas. Both target genes exhibited noticeably increased expression induced by their respective dTALEs at the transcript level as well as the protein level, which was confirmed at the protein activity level by ARS colorimetric assays. The level of induced target gene activation was closely correlated with the dTALE transcript abundance.

**Conclusions**

Our work demonstrates robust gene-specific activation induced by artificially designed TALEs in the green alga Chlamydomonas. The frequency and efficiency of the induced expression demonstrate dTALEs as powerful tools for targeted gene activation in Chlamydomonas. The pattern of activation shown here provides insights into the mechanism of TALE induced expression, and also confirms the activity of the activation domain of naturally occurring TALEs in another organism. Our success with dTALE-induced activation in Chlamydomonas may open a new avenue to fast, high throughput gene manipulation in other related organisms, such as green algae and crop plants.

**Keywords**

Chlamydomonas; TAL effector; targeted gene activation; arylsulfatase

**Introduction**

Transcription activator-like effectors (TALEs) are disease-related proteins secreted by certain plant pathogenic bacteria, e.g., those of the genus *Xanthomonas*. Through their type III secretion system, *Xanthomonas* cells deliver TALEs to host plant cells during the infection (Zhu et al. 2000; Buttner et al. 2002, 2006). These proteins function as transcription factors
by binding to the promoter sequences of specific host target genes. This results in activation of host resistance genes or susceptibility genes and, therefore, triggers either resistance responses or disease susceptibility in the host (White et al. 2009; Boch and Bonas 2010). All known TALEs typically contain three major domains: the N-terminal secretion signal, the C-terminal nuclear localization domain and trans-activation domain, and the central region of 1.5 to 33.5 repeats of 34 amino acid residues that is essential to DNA recognition and binding (Boch and Bonas 2010). The multiple repeats of this region are nearly identical except for the 12th and 13th amino acid residues, named the repeat variable di-amino-acids (RVDs). Each RVD recognizes one nucleotide of the target DNA. The order in which the coding regions for the repeats and their associated RVDs are arranged in the TALE gene dictates the specific gene sequence to which the mature TALE protein binds. Four predominant types of RVDs are found in most native TALEs, each recognizing an A, G, C, or T nucleotide in the target DNA sequences, specifically referred to as effector binding elements (EBEs) (Boch et al. 2009, Moscou and Bogdanove 2009).

The simple modular character of the TALE recognition code suggested its potential as a tool for targeted gene manipulation. Synthetic designer TALEs (dTALEs) with repeat units and RVDs arranged in a predetermined order have been used successfully to target specific genes for activation in plant and mammalian cells (e.g., Li et al. 2013; Maeder et al. 2013). Designed TALE nucleotide recognition domains also have been paired with FokI nuclease domains to build TALE nucleases (TALENs)(e.g., Christian et al. 2010; Li et al. 2011). TALENs can create double-strand DNA breaks in precise locations that are then repaired by either non-homologous-end-joining or homologous recombination to either disrupt or precisely edit target genes (Li et al. 2012; Cade et al. 2012).
Neither TALE-mediated activation nor TALEN-mediated genome editing has yet to be demonstrated in aquatic algae species such as *Chlamydomonas reinhardtii* (hereafter, Chlamydomonas), a model organism of long known importance in fundamental biological studies and biotechnology research. Gene activation and overexpression in Chlamydomonas have frequently been attempted in functional analysis studies, but often fall short of ideal because of difficulties in cloning Chlamydomonas genes containing high GC levels and the inability to gain adequate expression of even introduced copies of Chlamydomonas endogenous genes (Fuhrmann et al., 1999; Schroda et al. 2000).

Here we demonstrate the feasibility of targeted gene activation triggered by artificially tailored dTALEs, using two endogenous Chlamydomonas genes, *ARS1* and *ARS2*, as our targets. *ARS1* and *ARS2* are located on Chlamydomonas chromosome 16 in an inverse repeat pattern (de Hostos et al. 1988; Davies et al 1992). Both genes encode a periplasmic arylsulfatase (ARS), which cleaves the sulfate group of aromatic substrates (de Hostos et al. 1989). The activity of ARS can be easily assayed with 5-bromo-4-chloro-3-indolyl sulfate (XSO₄), a chromogenic compound that turns blue upon cleavage of its sulfate group (Davies et al. 1992). Transcripts of either *ARS1* or *ARS2* are barely detectable when Chlamydomonas cells are grown under sufficient sulfur conditions. Over 6 hours of sulfur deprivation, the abundance of *ARS1* transcript increases more than 1000-fold, while the *ARS2* transcript also increases several fold (Gonzalez-Ballester et al. 2010). By using dTALEs targeting the endogenous genes *ARS1* and *ARS2*, we show that target gene transcript abundance and target protein activity can be considerably elevated in dTALE transformed Chlamydomonas cells grown under replete sulfur. The results detailed here indicate that dTALEs can be used in
targeted gene activation and, therefore, for functional analysis of genes of interest in Chlamydomonas.

Material and Methods

Cell strains and culture conditions

Chlamydomonas reinhardtii strain cw10 (cc849, mating type minus) was obtained from the Chlamydomonas Stock Center, Duke University, Durham, NC. All cells were maintained on agar plates containing minimal medium (Geraghty et al. 1990) and were kept in high-CO₂ (~5% CO₂ in air) growth chambers at room temperature. Liquid cultures were grown on a gyratory shaker at 175 RPM in TAP medium (Gorman and Levine 1965) under 60 μmols photons m⁻² s⁻¹ illumination.

Chlamydomonas dTALE construction

The dTALE construction was performed as previously described using a library of repeat modules (Li and Yang 2013). “Core” repeats from avrXa7 encoding RVDs of NI, NG, NN and HD from this library were ligated in the receptor vector pTL-n to generate 8-mers of repeats in the desired order (see Fig. 1). A stop codon and XbaI digestion site were inserted into pSKavrXa7 (Li et al. 2010), resulting in pSKavrXa7delta. Three 8-mers were then cloned into pSKAvrXa7delta to create pSKavrXa7delta-ARS1 and pSKavrXa7delta-ARS2, each of which contained a full length dTALE with the original repeat region substituted with a synthetic repeat region targeting the designated ARS1 EBE or ARS2 EBE. A 9531bp plasmid, pDW2177, which consists of Chlamydomonas optimized Ble gene with introns, Chlamydomonas Rbcs2/Hsp70A chimeric promoter (Schroda et al. 2000), and
Chlamydomonas RbcS2 terminator (Wright and Spalding, unpublished data), was used as the backbone for TALE expression in Chlamydomonas. dTALE repeats from pSKavrXa7delta-ARS1 and pSKavrXa7delta-ARS2 were subsequently cloned into pDW2177 with BglII and XbaI to created full-length dTALEs driven by Chlamydomonas promoter Rbcs2/Hsp70A. The final constructs, pARS1-2177 and pARS2-2177 were used for Chlamydomonas transformation as described previously (Shimogawara et al. 1998).

**Colorimetric assay**

ARS activity in transformants was assayed using 5-bromo-4-chloro-3-indolyl sulfate (X-SO₄) according to previous protocols (Davies et al. 1992). During the screening process, individual transformants were grown in 300µl of TAP medium in 96-well microtiter wells for 60 hours until the cultures reached 5-10 x 10⁶ cells/ml (Davies et al. 1992), with XSO₄ added to the medium to a final concentration of 0.6mM. The microtiter plates were then centrifuged at 5000rpm (4472 x g), and the resulting supernatants from each well were screened visually for blue-green color to identify strains with potential ARS1 or ARS2 gene activation. Strains exhibiting apparent blue-green color were further assayed using consistent culture densities of 0.5 x 10⁶ cells/ml (measured with Beckman Coulter Z1 Coulter cell particle counter) grown in sulfur sufficient TAP medium for 60 hours and with 0.6mM XSO₄ (Davies et al. 1992). Cells were centrifuged at 5000rpm (4919 x g) and the OD₆₁₀nm of the resulting supernatant was measured with Cary 50 Bio Spectrophotometer (Varian) to quantify the apparent ARS activity.
Gene Expression Analysis

RNA was extracted with an RNAeasy extraction kit (Qiagen), and the RNA concentration was measured using an ND-1000 Nanodrop spectrophotometer (Nanodrop Technologies). The cDNA derived from 50ng of total RNA (Super Script III First-strand synthesis system; Invitrogen) 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 an iCycler iQ real-time PCR detection system (Biorad) using a SYBR green one-step quantitative PCR system (Quanta Biosciences). Amplicons from qRT-PCR were sequenced to ensure specificity of amplifications. 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 \( \Delta C_t = C_t_{\text{target gene}} - C_t_{\text{CBLP}} \) to represent the difference between the transcript abundance of genes examined and the transcript abundance of CBLP. After normalization, \( \Delta C_t \) values of each transformant were compared with that of wild type cells, and the dTALE induced expression is represented by \( \Delta \Delta C_t = \Delta C_t_{\text{transformant}} - \Delta C_t_{\text{wildtype}} \). (Li et al., 2013). For direct comparisons, fold change is also used to represent the relative abundance of target genes, which equals \( 2^{\Delta \Delta C_t} \).

DNA and protein blot analysis

Southern blot analysis was performed as described by Van and Spalding (1999). For western immunoblotting, total protein was obtained as described previously (Duanmu et al. 2008). Proteins were separated by SDS-PAGE on 12% polyacrylamide gels, and immunoblotting using an anti-FLAG-tag antibody (Santa Cruz Biotechnology; catalog no. sc-51590) was performed as described in the protocol from Bio-Rad Laboratories (catalog no. 500-0006). Protein immunoblot results were quantified with ImageJ Histogram analysis using the
Results and Discussion

Design of dTALEs for gene activation in Chlamydomonas

ARS1 and ARS2 are located on chromosome 16 in a tail-to-tail inverted repeat, and have different and distantly separated promoter sequences (Fig. 1A). We selected target EBEs for ARS1 and ARS2 (Fig. 1B & Fig.1C) based on criteria demonstrated by previous studies: 1) target EBEs should be located between 100bp to 35bp upstream of transcription initiation site; 2) the nucleotide preceding the EBEs should be a “T”; 3) each EBE will be 23bp long, and 4) there should be no identical off-target sequences in promoter regions across the Chlamydomonas genome (Boch et al. 2009; Boch and Bonas 2010; Li et al. 2010). To ensure specificity of dTALE recognition, the Chlamydomonas genome was searched by BLAST for additional sequences matching the candidate EBE sequences (Altschul et al. 1990). The dTALE constructs for ARS1 and ARS2, p2177-ARS1 and p2177-ARS2, respectively, were generated using an assembly method developed previously (Li and Yang 2013). The assembled 24-mers were used to replace the variable repeat region of the TALE AvrXa7. The resulting plasmids were subsequently cloned into Chlamydomonas expression vector pDW2177 (Wright and Spalding, unpublished) to contain a Chlamydomonas Rbcs2/Hsp70A chimeric promoter, antibiotic selective marker, and full-length dTALEs with artificially designed repeats (Fig. 1D).
dTALEs were successfully delivered and expressed in Chlamydomonas cells

We transformed wild type Chlamydomonas strain cw10 (cc849) with the dTALe constructs p2177-ARS1 and p2177-ARS2 and selected transformants based on resistance to zeocin. We confirmed the insertion of intact dTALe constructs into the Chlamydomonas genome using colony PCR to amplify both 3’ and 5’ fragments of the inserted dTALe construct in zeocin resistant transformants (data not shown). Of the 149 p2177-ARS1 transformant colonies tested, 52 (35%) showed apparently intact constructs inserted into the genome. Similarly, of 460 tested zeocin resistant colonies from the p2177-ARS2 transformation, 227 (49%) showed apparently intact dTALe constructs (Table 1).

Southern blot analysis with a probe specific for the p2177-ARS construct backbone (Fig. 1D) in selected transformants confirmed the presence of intact dTALe constructs and indicated a varied number of inserts among the transformants (Fig.1E). The combination of XhoI/SwaI restriction endonucleases cuts outside the dTALe construct, releasing the same fragment size from all intact dTALe inserts. The restriction endonuclease PvuII cuts inside the dTALe construct and in the flanking DNA, yielding different sized fragments for non-identical inserts. Single inserts were shown by PvuII digestion in most transformed lines, while multiple inserts were observed for lines ARS1-2177-201, 206, and 163, and possibly in lines ARS1-2177-167 and ARS2-2177-134. The intensity of bands when the DNA was cut with the XhoI/SwaI combination also was higher for lines ARS1-2177-201, 206, and 163, and ARS1-2177-167 and ARS2-2177-134, consistent with the apparent presence of multiple inserts in these lines.
Successful expression of dTALE constructs at the transcript level was confirmed with RT-PCR. We tested 20 transformant lines of p2177-ARS1, using primers specific for the 5’ region of dTALE constructs for p2177-ARS1 transformants and 3’ region of dTALE constructs for p2177-ARS2 transformants. 18 lines (2177-ARS1-121,149,163,167, 182, 201, 206, 211, 217, 222, 225, 228, 246, 264, 142, 157, 184) with intact p2177-ARS1 construct, based on colony PCR results, showed detectable expression of dTALE-ARS1 transcripts, while 2 lines (2177-ARS1-101,102) without an intact construct did not (Fig. 2). Similarly, of the 22 transformant lines tested, 18 lines (2177-ARS2-160, 560, 510, 177, 124, 207, 120, 121, 134, 205, 213, 502, 179, 102, 287, 512, 534, 575) with intact p2177-ARS2 construct exhibited detectable expression of dTALE-ARS2 transcripts, and 2 lines (2177-ARS2-125, 110) without an intact construct did not (Fig. 2). We also used western blot analysis to detect dTALE expression at the protein level. Of 14 tested transformant lines exhibiting RT-PCR-detectable expression of dTALE-ARS1 or dTALE-ARS2, dTALE protein abundance was detected with anti-FLAG tag antibody (supplemental Fig. 1).

**dTALEs induced expression of ARS1 and ARS2 in Chlamydomonas**

We used a colorimetric assay under sulfur-replete growth conditions to screen for ARS activity, which should indicate activation of either ARS1 or ARS2 in the absence of sulfur stress (Supplemental Fig. 2). Of the 149 p2177-ARS1 transformants screened, including lines with or without the intact p2177-ARS1 dTALE, 14 lines (9.4% of all transformants screened), all of which contained the intact dTALE construct, appeared blue-green in sulfur-replete culture medium after incubation with XSO$_4$. None of the transformant colonies lacking a demonstrated intact TALE construct showed detectable color change under the same growth conditions. Of 460 p2177-ARS2 transformants screened by the same method,
43 lines (9.3% of all transformants screened) containing the intact p2177-ARS2 dTALE showed visible ARS activity, while none of the p2177-ARS2 transformants lacking an intact dTALE construct showed detectable ARS activity (Table 1).

Activation of both target genes, \textit{ARS1} and \textit{ARS2}, was demonstrated at the transcript level with RT-PCR, using primers specific for the \textit{ARS1} 5’UTR or the \textit{ARS2} 5’UTR. When grown under sulfur starvation, elevated expression levels of both \textit{ARS1} and \textit{ARS2} were observed in wild type cells (Fig.3). dTALE-induced expression of \textit{ARS1} transcripts under sulfur-replete conditions was observed in 14 lines (2177-ARS1-121,149,163,167, 182, 201, 206, 211, 217, 222, 225, 228, 246) showing an intact 2177-ARS1 dTALE construct and colorimetric ARS activity, as well as in 4 lines (2177-ARS1-264, 142, 157, 184) showing an intact 2177-ARS2 dTALE construct but no detectable colorimetric ARS activity, while 2 transformants (2177-ARS1-101,102) lacking an intact 2177-ARS1 dTALE construct showed no detectable increase in \textit{ARS2} mRNA abundance under sufficient sulfur conditions (Fig.3). Similarly, 12 lines (2177-ARS2-160, 560, 510, 177, 124, 207, 120, 121, 134, 205, 213, 502) showing an intact 2177-ARS2 dTALE construct and colorimetric ARS activity, and 6 lines (2177-ARS2-179, 102, 287, 512, 534, 575) showing an intact 2177-ARS2 dTALE construct yet no detectable colorimetric ARS activity exhibited induced \textit{ARS2} transcript expression under sulfur-replete conditions, while 2 transformants (2177-ARS2-125, 110) lacking an intact dTALE construct showed no detectable increase in \textit{ARS2} transcript abundance under the same conditions (Fig. 3).

3.4 Activation of target genes is positively correlated to the expression of dTALEs

We used ΔΔCt values normalized to the internal control gene, \textit{CBLP}, to interpret data from qRT-PCR assays (Li et al., 2013). In all 18 lines containing an apparently intact p2177-ARS1
dTALE construct, qRT-PCR quantification indicated that induced ARS1 transcript abundance was positively related to dTALE-ARS1 transcript abundance (Fig.4A). In the 2 lines lacking an intact p2177-ARS1 dTALE construct, no expression above the baseline level was detected, and the non-targeted ARS2 transcript abundance showed no detectable increase in these lines (Fig. 4A, showing two lines tested). In the 18 lines containing an intact p2177-ARS2 dTALE construct, the transcript abundance of ARS2 also correlated strongly with the transcript abundance of dTALE-ARS2 (Fig.4B). Similar to the p2177-ARS1 transformants, no off-target ARS1 transcript abundance was detected in the 2 dTALE-ARS2 transformants tested, and the 2 lines lacking an intact dTALE construct did not show any increased ARS2 expression (Fig.4B). The two transformant lines with the most abundant target gene transcripts were ARS1-2177-206 and ARS2-2177-560, where the ΔΔCt of dTALE activated expression was 1.4 and 2.6 lower, respectively, than that of sulfur depletion induced ARS1 expression. Compared with uninduced expression of ARS genes in wild type cells, the ΔΔCt values of dTALE mediated expression of target genes ARS1 and ARS2 in ARS1-2177-206 and ARS2-2177-560 were 19.7 and 21.2 higher, respectively, and the fold change of the target genes in these lines are more than 100,000 times higher. These data indicate effective activation induced by dTALEs in Chlamydomonas.

The expression of dTALE proteins was detected in western immunoblots with anti-FLAG tag antibody (supplemental Fig. 1). The abundance of ARS1 and ARS2 transcripts in all transformed lines examined appeared to be positively related to the amount of expressed dTALEs (Fig. 5), although the relative protein density measured seemed to saturate at a relative protein abundance of about 230, as indicated in Fig. 5A. In all other lines, the transcript abundance of ARS1 and ARS2 correlated with dTALE expression in a near linear
manner. This demonstrates clearly that the dTALE constructs were being expressed at the protein level and more closely correlates the activation of the target genes *ARS1* and *ARS2* with the functional abundance of dTALE expression.

The relationship between qRT-PCR based *ARS1* or *ARS2* transcript abundance and colorimetric ARS activity assays performed on both p2177-ARS1 and p2177-ARS2 transformed lines in sulfur-replete medium showed that colorimetric ARS activity was detectable in only those selected lines with transcript abundance yielding a ΔΔCt=12.3 or higher (Fig 6). Also evident in Fig. 6, there was a positive correlation between ΔΔCt and increased *ARS1* or *ARS2* transcript abundance for those lines with a ΔΔCt greater than 12.3.

Concluding Remarks

TAL effectors, especially when fused with nucleases, have risen as promising tools in molecular biology studies and bioengineering in recent years. Breakthroughs in deciphering the TALE-DNA recognition code have made possible manipulations of gene expression mediated by artificially designed TALEs or TALENs targeting endogenous or exogenous genes in many organisms (Li et al. 2010; Mobitzer et al. 2010; Zhang et al 2011; Li et al 2013). In this study, we expanded the use of TALE activation to another key model organism, the aquatic green alga *Chlamydomonas reinhardtii*. We generated dTALEs targeting the endogenous Chlamydomonas genes *ARS1* and *ARS2*, and delivered dTALE constructs to wild type Chlamydomonas cells. Despite their adjacent locations and the resemblance of their coding sequences, *ARS1* and *ARS2* are two distinct genes possibly resulting from a single duplication event (Hallmann and Sumper, 1994). The tail-to-tail pattern of the duplication determines the relatively long distance between the *ARS1* and *ARS2* promoters, which share no obvious sequence similarity. Gene expression analysis revealed
marked and highly selective dTALE activation of either ARS1 or ARS2 (Fig.2, Fig.3), with no off-target activation of the non-targeted ARS gene. This indicates that specific activation of most if not any desired gene might be achieved with this dTALE-based system in Chlamydomonas.

The frequency of activation as determined by the colorimetric assay of both target genes appeared to be ~9% of all transformants screened. However, in all examined transformant lines where the inserted dTALE construct was intact, elevated transcript abundance of the specific target genes was demonstrated by both RT-PCR and qRT-PCR results (Fig.2, Fig.4). Therefore, the frequency of successful dTALE-based activation of the target gene above baseline level appears to be much higher than 9%.

The lack of detectable colorimetric ARS activity in lines with induced but relatively lower transcript abundance, as indicated by a ΔΔCt less than 12.3, suggests that the colorimetric assay has a much lower sensitivity than qRT-PCR, which is not surprising. It also suggests that the expression level of either ARS1 or ARS2 must exceed some threshold level before a detectable color develops, possibly because the sensitivity of the ARS colorimetric assay cannot detect low levels of activity or because low level transcript abundance did not yield ARS protein and thus ARS activity. The usefulness of the colorimetric assay, aside from its obvious ease relative to RT-PCR or qRT-PCR, may lie in its identification of relatively higher-level activation lines. Therefore, this colorimetric assay clearly suggests that relatively high level activation of these two ARS genes occurs in about 9% of the transformants.

Chlamydomonas cells are able to acclimate to different nutritional conditions by altering the expression level of various genes beneficial in the scavenging of the limiting nutrients. Under
sulfur depletion, the expression of ARSI is up-regulated more than 1000 times within 6 hours. Based on a previous study, the RPKM values of ARS1 under sulfur depletion and repletion are 2717 and 2, respectively. (Gonzalez-Ballester et al. 2010). In dTALE transformant lines with the highest target genes activation levels, the transcript abundance of ARSI and ARS2 were comparable to that of sulfur-starvation induced ARSI. Among the lines generated and analyzed in this study, the maximum expression of dTALE-activated ARSI and ARS2 were only 3-4-fold lower than sulfur-starvation induced ARSI expression. The significant activation of both target genes, especially of ARS2, a gene that is almost completely unexpressed in all known nutritional conditions (Davies et al. 1996; Ravina et al. 2002; Gonzalez-Ballester et al. 2010), demonstrates the power of dTALE activation in Chlamydomonas.

The highest level of expression attained for ARSI and ARS2, which was observed with ≤10% of the transformants, should be sufficient for many potential TALE activator applications, such as perturbations of metabolic enzymes to examine their functional roles or to shift partitioning to a specific metabolic product. As discussed above, lower levels of activation appear to occur at a much higher frequency. This could be an important distinction depending on what type of gene one is seeking to activate. Genes encoding metabolic or biosynthetic enzymes generally will require relatively high expression to have the desired impact. On the other hand, targeted activation of regulatory genes, such as those encoding transcription factors generally would require only fairly low level activation for impact and, thus, should occur at a much higher frequency.

Because of its simple structure and high growth potential, Chlamydomonas has been used as a model organism and a potential host for bioproduct generation. The expression level of
transgenes in Chlamydomonas, however, has often been disappointing. While transformation serves an effective method to generate transgenic clones, transgene introduced via transformation integrate into the Chlamydomonas genome in a random manner (Zhang et al. 2014). Because the unknown nature of constraints on transgene expression, it is difficult to obtain clones that express transgenes at a high level (Fuhrmann et al., 1999; Schroda et al. 2000). Various strategies have been used to overcome limits, such as the high GC content and the codon bias of Chlamydomonas genome (Fuhrmann et al. 2004; Shao and Bock 2008), and the effect of transgene integration position (Rasala et al. 2012; Bock et al. 2009), but no general solution has been established yet. The dTALE-mediated activation demonstrated here provides a promising strategy in that it does not rely on the expression level of the transgene itself, but only on its activation potential to induce endogenous gene expression.

Despite the effort invested in studying TALEs, a complete understanding of the TALE activation mechanism is still lacking. It has been suggested that TALEs function in a similar way as transcription factors, with the DNA recognition domain binding DNA sequences and an acidic activation domain facilitating transcription initiation (Gu et al. 2005; Boch and Bonas 2010). While marked activation has been achieved in certain studies, in other cases the activation seemed modest (Zhang et al. 2011; Garg et al. 2012). Previous studies have shown significantly increased levels of targeted activation when multiple dTALE constructs, rather than one, were introduced into mammalian cells (Meader et al. 2013; Perez-Pinera et al. 2013). Our results agree with this synergistic or dosage effect of dTALEs, in that the transcript abundance of activated genes appeared to be positively and closely correlated with the abundance of dTALE transcripts (Fig. 3). Such a correlation suggests that at the level of
activation obtained, the dTALE abundance limits the activation of the target gene. To achieve higher levels of activation of target genes, it might prove useful to elevate the abundance of the artificially delivered dTALEs by increasing the copies of dTALE genes or including a strong constitutive promoter.

For decades, *Chlamydomonas reinhardtii* has served as an important model system for both fundamental molecular and genetics research and biotechnology advances. The simplicity and availability of this unicellular green alga has facilitated studies on photosynthesis, chloroplast biogenesis, flagellar assembly and mobility, as well as most if not all basic metabolic pathways (Grossman et al. 2004; Merchant et al. 2007; Takahashi et al 2011; Urzica et al. 2012). Studies with regard to the CO₂ concentrating mechanism of Chlamydomonas have revealed this alga’s unique ability to acclimate to multiple states of environmental CO₂ concentrations and therefore achieve high photosynthetic efficiency, which provides promising insights to carbon assimilation and potentially into crop production (Wang and Spalding, 2006; Spalding 2008; Wang et al. 2011). In recent years, Chlamydomonas has also drawn attention as a system for production of biopharmaceuticals, as well as a potential clean source for hydrogen and biodiesel production (Ruppecht 2009; Rosales-Mendoza et al. 2012; James et al. 2013; Pourmir et al. 2013). Our study has for the first time demonstrated robust gene activation in this model system, precisely targeting any desired genes. Unlike the traditional promoter-transgene method, this approach provides a powerful experimental system of endogenous gene activation that is not limited by the requirement for complex cloning strategies. Given its robustness, apparent high efficiency and specificity, this dTALE-based system may have great potency as a tool that helps to open
the gate to fast and efficient functional studies as well as biotechnology advances in Chlamydomonas.

Similar to Chlamydomonas, other microalgae species, such as diatoms, chlorella, and Haematococcus, have long been used to study basic biological pathways, acclimation to environment changes, biomedicine and food supplement production, as well as renewable energy sources (Tretacoste et al. 2013, Romano et al. 2013, Cao et al. 2013, Raman et al. 2012, Lorentz and Cysewski, 2000.) Besides the advantage brought to Chlamydomonas and carbon metabolism research, our success in targeted gene activation in Chlamydomonas suggests dTALE activation as a potentially valuable tool for both basic and applied research in other algae.
Literature Cited


Buttner D, Bonas U (2002). Getting across--bacterial type III effector proteins on their way to the plant cell. EMBO J., 21, 5313-5322.


Figure Legends

**Figure 1.** Structure of the designed TAL effectors (dTALEs) for Chlamydomonas. A, Relative positions of *ARS1* and *ARS2* genes and effector binding elements (EBEs) on chromosome 16. Red stars, positions of EBEs. Angled arrows, transcription initiation sites. B, DNA sequence of *ARS1* EBE and corresponding dTALE repeat, only showing repeat variable residues (RVDs). C, DNA sequence of *ARS2* EBE and corresponding dTALE repeats. D, Schematic drawing of Chlamydomonas dTALE constructs. The dTALE constructs contain full length sequences of the N-terminal and C-terminal domains of *avrXa7*, but with the *avrXa7* DNA recognition domain replaced by 23 repeats designed for *ARS1* or *ARS2* EBEs. Arrows indicate restriction sites of *XhoI*, *PvuII*, and *SwaI*. Yellow bar indicates binding position of the probe used for Southern blot analysis. NLS, nuclear localization signal. AD, activation domain. E, Southern blot analysis of transformed lines. Genomic DNA concentrations were measured on electrophoresis gel by comparing band strength with ladder bands and then digested with indicated restriction enzymes, with wild type cw10 cells as a control. 1-, 2177-ARS1; 2-, 2177-ARS2; wt, wild type.

**Figure 2.** RT-PCR analysis of dTALE transgene transcript presence in Chlamydomonas, using CBLP transcript as an internal control. A, dTALE-*ARS1* transcript presence in lines with intact dTALE-*ARS1* transgene (left panel) and absence in lines without intact dTALE-*ARS1* (right panel). B, dTALE-*ARS2* transcript presence in lines with intact dTALE-*ARS2* transgene (left panel) and absence in lines without the intact transgene (right panel). Transformed lines were grown under sulfur sufficient conditions, while cw10 cells were
grown under either sulfur-sufficient or sulfur-deficient conditions, as indicated, for 60 h. numbers underneath each panel indicate transformant line numbers.

**Figure 3.** RT-PCR analysis of *ARS1* and *ARS2* transcript presence in transformed lines using *CBLP* transcript as an internal control. A, presence of *ARS1* transcript in lines with intact *dTALE-ARS1* transgene in sulfur replete medium (upper panel), absence of *ARS1* transcript in those without the intact transgene (lower right panel) and absence of *ARS2* transcript in *dTALE-ARS1* transformed lines grown in sulfur replete medium (lower left panel). B, presence of *ARS2* transcript in lines with intact *dTALE-ARS2* transgene in sulfur replete medium (upper panel), absence of *ARS2* transcript in those without the intact transgene (lower right panel) and absence of *ARS1* transcript in *dTALE-ARS2* transformed lines grown in sulfur replete medium (lower left panel). Transformed lines were grown under sulfur sufficient conditions, while cw10 cells were grown under either sulfur-sufficient or sulfur-deficient conditions, as indicated, for 60 hrs.

**Figure 4.** Analysis of *ARS1* and *ARS2* target gene transcript abundance by qRT-PCR as a function of the corresponding dTALE transcript abundance. A, *ARS1* transcript abundance as a function of *dTALE-ARS1* transcript abundance, showing transformed lines with intact *dTALE-ARS1* and visible ARS activity (diamonds), transformed lines with intact *dTALE-ARS1* and no visible ARS activity (triangles), transformed lines without an intact *dTALE-ARS1* (crosses), and wild type (round dot). The absence of *ARS2* transcript in two lines with activated *ARS1* is also shown (squares). B, *ARS2* transcript abundance as a function of *dTALE-ARS2* transcript abundance, showing transformed lines with intact *dTALE-ARS2* and visible ARS activity (diamonds), transformed lines with intact *dTALE-ARS2* and no visible ARS activity (triangles), transformed lines without an intact *dTALE-ARS2* (crosses), and wild
type (round dot). The absence of ARSI transcript in two lines with activated ARS2 is also shown (squares).

$$\Delta \Delta Ct = \Delta Ct_{\text{transformant}} - \Delta Ct_{\text{wt}} = (Ct_{\text{targetgene-transformant}} - Ct_{\text{CBLP-transformant}}) - (Ct_{\text{targetgene-wt}} - Ct_{\text{CBLP-wt}}),$$

where $\Delta Ct$ equals the difference between target transcript abundance and the transcript abundance of $CBLP$, the internal control gene. Three repeats were performed for each transformants, and the maximum standard error was 0.81.

**Figure 5.** ARSI (A) and ARS2 (B) target gene transcript abundance by qRT-PCR as a function of the responding dTALE protein abundance. Round dots, wild type controls. Diamonds, transformant lines. The relative transcript abundance is represented by fold change. Western blot results were quantified with ImageJ histogram analysis. Relative protein abundance was determined by the ratio of pixels from each sample lane on western blot over pixels from the control lane, as was described in material and methods.

**Figure 6.** ARS activity resulting from ARSI or ARS2 dTALE activation as a function of the transcript abundance of the corresponding ARS gene. A, OD$_{610\text{nm}}$ of p2177-ARS1 transformants and cw10 cells grown under sufficient sulfur for 60h as a function of ARSI transcript abundance, showing transformed lines with intact dTALE-ARS1 and visible ARS activity (diamonds), transformed lines with intact dTALE-ARS1 and no visible ARS activity (triangles), and wild type (round dots). B, OD$_{610\text{nm}}$ of p2177-ARS2 transformants and cw10 cells grown under sufficient sulfur for 60hrs as a function of ARS2 transcript abundance.

$$\Delta \Delta Ct = \Delta Ct_{\text{transformant}} - \Delta Ct_{\text{wt}} = (Ct_{\text{targetgene-transformant}} - Ct_{\text{CBLP-transformant}}) - (Ct_{\text{targetgene-wt}} - Ct_{\text{CBLP-wt}}).$$

Three repeats were performed for qRT-PCR of each transformed lines, and the maximum standard error was 0.35. Three technical repeats were used for absorbance at 610nm.
Table 1. Numbers of transformants showing efficiencies of transformation and activation

<table>
<thead>
<tr>
<th></th>
<th>Zeocin$^a$</th>
<th>Intact dTALE$^a$</th>
<th>X-SO$_4$$^b$</th>
<th>ARS activation$^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>pARS1-2177</td>
<td>149</td>
<td>52/149</td>
<td>14/149</td>
<td>18/18</td>
</tr>
<tr>
<td>pARS2-2177</td>
<td>460</td>
<td>227/460</td>
<td>43/460</td>
<td>18/18</td>
</tr>
</tbody>
</table>

$^a$Intact dTALE: transformants containing intact dTALE construct as a fraction of total colonies screened.

$^b$X-SO$_4$: transformants showing visible color change in colorimetric assay as a fraction of total colonies screened.

$^c$ARS activation: transformants show elevated ARS1 or ARS2 expression detectable with RT-PCR. 18 transformants with intact construct were tested for both dTALE-ARS1 and dTALE-ARS2, including some with visible ARS activity and some without.
Figure 1. Structure of the designed TAL effectors (dTALEs) for Chlamydomonas. A, Relative positions of ARS1 and ARS2 genes and effector binding elements (EBEs) on chromosome 16. Red stars, positions of EBEs. Angled arrows, transcription initiation sites. B, DNA sequence of ARS1 EBE and corresponding dTALE repeat, only showing repeat variable residues (RVDs). C, DNA sequence of ARS2 EBE and corresponding dTALE repeats. D, Schematic drawing of Chlamydomonas dTALE constructs. The dTALE constructs contain full length sequences of the N-terminal and C-terminal domains of avrXa7, but with the avrXa7 DNA recognition domain replaced by 23 repeats designed for ARS1 or ARS2 EBEs. Arrows indicate restriction sites of Xhol, PvuII, and Swai. Yellow bar indicates binding position of the probe used for Southern blot analysis. NLS, nuclear localization signal. AD, activation domain. E, Southern blot analysis of transformed lines. Genomic DNA concentrations were measured on electrophoresis gel by comparing band strength with ladder bands and then digested with indicated restriction enzymes, with wild type cw10 cells as a control. 1-, 2177-ARS1; 2-, 2177-ARS2; wt, wild type.
**Figure 2.** RT-PCR analysis of dTALE transgene transcript presence in Chlamydomonas, using CBLP transcript as an internal control. A, dTALE-ARS1 transcript presence in lines with intact dTALE-ARS1 transgene (left panel) and absence in lines without intact dTALE-ARS1 (right panel). B, dTALE-ARS2 transcript presence in lines with intact dTALE-ARS2 transgene (left panel) and absence in lines without the intact transgene (right panel). Transformed lines were grown under sulfur sufficient conditions, while cw10 cells were grown under either sulfur-sufficient or sulfur-deficient conditions, as indicated, for 60 h. Numbers underneath each panel indicate transformant line numbers.
**Figure 3.** RT-PCR analysis of *ARS1* and *ARS2* transcript presence in transformed lines using *CBLP* transcript as an internal control. A, presence of *ARS1* transcript in lines with intact *dTALE-ARS1* transgene in sulfur replete medium (upper panel), absence of *ARS1* transcript in those without the intact transgene (lower right panel) and absence of *ARS2* transcript in *dTALE-ARS1* transformed lines grown in sulfur replete medium (lower left panel). B, presence of *ARS2* transcript in lines with intact *dTALE-ARS2* transgene in sulfur replete medium (upper panel), absence of *ARS2* transcript in those without the intact transgene (lower right panel) and absence of *ARS1* transcript in *dTALE-ARS2* transformed lines grown in sulfur replete medium (lower left panel). Transformed lines were grown under sulfur sufficient conditions, while cw10 cells were grown under either sulfur-sufficient or sulfur-deficient conditions, as indicated, for 60 hrs.
Figure 4. Analysis of ARS1 and ARS2 target gene transcript abundance by qRT-PCR as a function of the corresponding dTALE transcript abundance. A, ARS1 transcript abundance as a function of dTALE-ARS1 transcript abundance, showing transformed lines with intact dTALE-ARS1 and visible ARS activity (diamonds), transformed lines with intact dTALE-ARS1 and no visible ARS activity (triangles), transformed lines without an intact dTALE-ARS1 (short bars), and wild type (round dot). The absence of ARS2 transcript in two lines with activated ARS1 is also shown (squares). B, ARS2 transcript abundance as a function of dTALE-ARS2 transcript abundance, showing transformed lines with intact dTALE-ARS2 and visible ARS activity (diamonds), transformed lines with intact dTALE-ARS2 and no visible ARS activity (triangles), transformed lines without an intact dTALE-ARS2 (short bars), and wild type (round dot). The absence of ARS1 transcript in two lines with activated ARS2 is also shown (squares).

$\Delta\Delta Ct = \Delta Ct_{\text{transformant}} - \Delta Ct_{\text{wt}} = (Ct_{\text{target gene-transformant}} - Ct_{\text{CBLP-transformant}}) - (Ct_{\text{target gene-wt}} - Ct_{\text{CBLP-wt}}),$ where $\Delta Ct$ equals the difference between target transcript abundance and the transcript abundance of CBLP, the internal control gene. Three repeats were performed for each transformants, and the maximum standard error was 0.81.
Figure 5. *ARS1* (A) and *ARS2* (B) target gene transcript abundance by qRT-PCR as a function of the responding dTALE protein abundance. Round dots, wild type controls. Diamonds, transformant lines. The relative transcript abundance is represented by fold change. Western blot results were quantified with ImageJ histogram analysis. Relative protein abundance was determined by the ratio of pixels from each sample lane on western blot over pixels from the control lane, as was described in material and methods. In panel A, the 30,000 to 850,000 range on the y axis is not shown for a clear illustration of the result.
Figure 6. ARS activity resulting from ARS1 or ARS2 dTALE activation as a function of the transcript abundance of the corresponding ARS gene. A, OD_{610nm} of p2177-ARS1 transformants and cw10 cells grown under sufficient sulfur for 60h as a function of ARS1 transcript abundance, showing transformed lines with intact dTALE-ARS1 and visible ARS activity (diamonds), transformed lines with intact dTALE-ARS1 and no visible ARS activity (triangles), and wild type (round dots). B, OD_{610nm} of p2177-ARS2 transformants and cw10 cells grown under sufficient sulfur for 60hrs as a function of ARS2 transcript abundance.

\[ \Delta \Delta Ct = \Delta Ct_{\text{transformant}} - \Delta Ct_{\text{wt}} = (Ct_{\text{target gene-transformant}} - Ct_{\text{CBSA-transformant}}) - (Ct_{\text{target gene-wt}} - Ct_{\text{CBSA-wt}}) \]

Three repeats were performed for qRT-PCR of each transformed lines, and the maximum standard error was 0.35. Three technical repeats were used for absorbance at 610nm.
CHAPTER 3. EXPRESSION ACTIVATION AND FUNCTIONAL ANALYSIS OF HLA3, A PUTATIVE INORGANIC CARBON TRANSPORTER IN CHLAMYDOMONAS REINHARDTII

A manuscript to be submitted to Plant Journal or Photosynthesis Research before 08.11. 2014.

The library used to assemble artificially designed TALE in this chapter was a generous gift from Dr. Bing Yang’s lab, contributed by Dr. Ting Li. The subcellular localization of HLA3 protein was contributed by Dr. Xiaowen Fei. The Chlamydomonas expression vector backbone was constructed by Dr. David Wright. The LCIA overexpression strain LA43 was created thanks to the work of Dr. Yingjun Wang.

Abstract

The CO₂ concentrating mechanism (CCM) is a key component of the carbon assimilation strategy of aquatic microalgae. Induced by limiting CO₂ and tightly regulated, the CCM enables these microalgae to rapidly respond to varying environmental CO₂ supplies and to perform photosynthetic CO₂ assimilation in a cost effective way. A functional CCM in eukaryotic algae requires Rubisco sequestration, rapid interconversion between CO₂ and HCO₃⁻ catalyzed by carbonic anhydrases (CAs), and active inorganic carbon (Ci) uptake. In the model microalga Chlamydomonas reinhardtii, a membrane protein HLA3 is proposed to be involved in active Ci uptake across the plasma membrane. In this study, we use an artificially designed transcription activator-like effector (dTALE) to activate the expression of HLA3. The successful activation of HLA3 expression demonstrates dTALE as a promising tool for gene specific activation and investigation of gene function in Chlamydomonas. Activation of HLA3 expression under in high CO₂ acclimated cells, where HLA3 is not expressed, resulted in increased Ci accumulation and Ci-dependent photosynthetic O₂ evolution specifically in very low CO₂ concentrations, which confirms that HLA3 is indeed involved in Ci uptake, and suggests it is mainly associated with HCO₃⁻.
transport in very low CO₂ concentrations, conditions where active CO₂ uptake is highly limited.

Introduction

As the substrate of photosynthesis, CO₂ has provided the basis of the organic world on earth since the origin of life. Interestingly, the central enzyme of photosynthetic CO₂ assimilation, Rubilose-1,5-bisphosphate carboxylase oxygenase (Rubisco) has a notoriously low turnover rate, and the carboxylation reaction is competitively inhibited by O₂, as reflected by the name of the enzyme. The efficiency of photosynthesis therefore heavily relies on the concentration of CO₂ in the vicinity of Rubisco, and different strategies have evolved in various plant species to ensure a high local concentration of CO₂. For example, in species harboring Crassulacean acid metabolism (CAM) and C4 metabolism (so called “C4 plants”), CO₂ is first sequestered in a 4-carbon metabolic product and then enriched by decarboxylation of the 4-carbon metabolite in the vicinity of Rubisco.

Unlike terrestrial plants, aquatic photosynthetic species face an even more challenging environment regarding inorganic carbon (Ci) resources. Given the $10^4$ times slower diffusion rate of CO₂ in water relative to air and the impact of pH, sediment, soil respiration, and the Ci source, CO₂ for aquatic species is oftentimes scarce and can be drastically changing over a short period of time. Cyanobacteria and microalgae species have developed limiting CO₂ inducible systems generally known as CO₂ concentrating mechanisms (CCMs) to help scavenge Ci when CO₂ concentration is low. Under limiting CO₂, the CCMs enable these organisms to quickly respond and acclimate to the ever-changing environmental CO₂ supply, drastically increase the CO₂ concentration supplied to Rubisco, and utilize metabolic energy to boost the rate of carbon assimilation.

The CCMs in cyanobacteria have been extensively studied, and major components of the system have been revealed (Badger and Price 2003; Price 2011; Kupriyanova et al. 2013). Compared to the cyanobacterial CCMs, the CCMs in eukaryotic microalgae are more complex. One species of green alga, *Chlamydomonas reinhardtii* (hereafter,
Chlamydomonas), has been known as a model organism for its fast growth, facile genetics and complete genome sequence. Decades of research have identified key components involved in the Chlamydomonas CCM, including carbonic anhydrases (CAs) and potential HCO$_3^-$ transporters and active CO$_2$ uptake mechanisms, as well as regulatory genes that control the energy/Ci dependent response of the CCM. However, a complete picture of the CCM remains unclear, and much needs to be revealed regarding the exact functions of these elements. In most natural environments, two species of Ci exist as the major source of carbon assimilation for Chlamydomonas: CO$_2$ and HCO$_3^-$. As the Ci concentration, the pH, and the temperature vary, the availability of both Ci species also vary greatly, as does the acclimation status of the CCM. At least three acclimation states have been defined based on physiological responses to varying CO$_2$ concentrations (Vance and Spalding 2005) and the phenotype caused by mutations in one gene, LCIB, which exhibits an air dier phenotype characterized by lethality under air level (~0.035%) of CO$_2$ but survival under either high (>0.4%) or very low (<0.02%) CO$_2$ (Spalding et al. 1983; Wang and Spalding, 2006).

To better understand the function of the CCM, it is important to investigate the specific roles of the proteins involved, especially the proposed Ci transporters. Transcriptome analyses have recognized hundreds of genes (Miura et al 2004; Fang et al. 2012; Brueggeman et al. 2012) whose expressions are induced by limiting CO$_2$ and are controlled by the “master regulator” of the CCM, CIA5. In the absence of a functional CIA5, the cells are unable to induce the CCM and require high CO$_2$ to grow (Moroney et al. 1989, Fukuzawa et al. 2001; Xiang et al. 2001). Many of these genes encode proteins that are potential Ci transporters, including CCP1/CCP2 and LCIA on the chloroplast envelope (Pollock et al. 2004, Mariscal et al. 2006), as well as LCI1 and HLA3 on the plasma membrane (Fukuzawa et al. 2001; Onishi et al 2010; Im and Grossman, 2002; Miura et al. 2004; Burow et al. 1996).

Originally identified as a high-light-activated protein (Im and Grossman, 2002), HLA3 is a putative ATP-binding cassette (ABC) type transporter that belongs to the multidrug resistance related protein (MRP) family. The MRP proteins are widely found in animals and plants as transporters of variant anions involved in pathways including drug extrusion, detoxification, and guard cell regulation (Gaedeke et al, 2001; Kruh et al., 2003; Hanikenne
et al., 2005; Klein et al., 2003, 2006). While the substrate for other potential Ci transporters remains unclear, recent work has provided insight regarding the Ci species specificity of HLA3. Knockdown of HLA3 alone resulted in decreased photosynthetic Ci uptake and assimilation under alkaline pH and very low CO₂ conditions, and co-knockdown of HLA3 and LCIA severely hampered growth, photosynthesis and Ci uptake in Chlamydomonas, especially in very low CO₂ at alkaline pH where HCO₃⁻ is the dominant Ci species (Duanmu et al. 2009a). It is therefore reasonable to infer that HLA3 and LCIA both function in HCO₃⁻ transport, with HLA3 facilitating its transport across the plasma membrane and LCIA across the chloroplast envelope, as the next gateway to the vicinity of Rubisco. To further understand the roles of these genes, overexpression of HLA3, especially when combined with different genetic backgrounds, can provide evidence for a more detailed picture of the CCM function.

In recent years, transcription activator-like effectors (TAL effectors, or TALEs) have gained wide attention as an effective tool for targeted gene activation and genome editing. TALEs are effector proteins secreted by plant pathogenic bacteria species Xanthomonas when infecting its hosts. Delivered through type III secretion systems, TALEs function like transcription activators by recognizing and binding onto the promoter region of and activating the expression of certain host genes, causing disease susceptibility or resistance responses of the host plants (Zhu et al. 2000; Buttner et al. 2002, 2006). The recognition pattern between naturally occurring TALEs and their target genes has been well deciphered. Each TALE is composed of a N-terminal domain containing a translocation signal, a C-terminal domain containing an activation motif and a nuclear localization signal, and central repeats that are essential for DNA recognition (Bogdanove et al. 2010; Boch and Bonas 2010; White and Yang, 2009; Wright et al. 2014). This region typically contains an average of 17.5 repeats, although TALEs ranging from 1.5 to 33.5 repeats have been observed. The repeats contain 33 amino acid (aa) residues and are all nearly identical except for the 12th and 13th aa residues, known as repeat variable diresidues (RVDs). The RVDs have been proved crucial to DNA recognition and binding, with the 12th aa residue intruding into the major groove of target DNA strands and the 13th aa residue interacting with other repeats to ensure stable binding. Each pair of RVDs specifically binds one single nucleotide, and the four most
The predominant type of RVDs observed in nature are HD, NG, NN, and NI, recognizing C, T, G (or A), and A, respectively (Boch et al. 2009, Moscou and Bogdanove 2009). Other types of RVDs have also been identified, with different levels of specificity (Streubel et al. 2012; Christian et al. 2012). Compared with earlier editing tools, such as zinc finger proteins and mega nucleases, the truly modular recognition pattern of TALEs have made them superior tools for targeted gene activation and genome editing. By substituting the original central repeats with artificially designed repeats targeting desired gene promoters, designed TALEs (dTALEs) have been created and successfully used in plant, algal, human, and other mammalian cells to mediate gene specific activation (Li et al. 2013; Maeder et al. 2013; Gao et al. 2014). The combination of dTALEs and endonuclease FokI has also achieved great success in site directed mutagenesis and genome editing in plants, zebra fish, C. elegans, human cells, and other mammalian cells (Cade et al. 2010, Cermak et al. 2011; Li et al. 2011, 2012).

In this study, we use a dTALE targeting the 5’ promoter region of HLA3 to induce constitutive expression of HLA3 in Chlamydomonas. The activation of HLA3 expression in different genetic backgrounds reveals a clear role of HLA3 in the Chlamydomonas CCM. In agreement with previous studies, we found that HLA3 is probably responsible for active uptake of HCO$_3^-$, specifically under very low CO$_2$ concentrations.

Results

HLA3 Localization.
HLA3 is predicted to be located on the plasma membrane based on its deduced aa sequence (Duanmu et al. 2009a). To confirm its subcellular localization, we performed protein immunoblots of purified plasma membrane fractions from Chlamydomonas cells with anti-HLA3 antibodies. As indicated in Figure 1, HLA3 is highly enriched in the plasma membrane fraction, which is distinguished by enrichment of the plasma membrane H$^+$-ATPase and depletion of the thylakoid membrane protein D1 and the chloroplast envelope proteins CCP1/CCP2. Accordingly, HLA3 is depleted from the membrane fraction.
containing the chloroplast envelope and thylakoid membranes, which are distinguished by enrichment of CCP1/CCP2 and D1, respectively (Figure 1).

**dTALE-Mediated Activation of HLA3 Expression.**

Using criteria described previously (Gao et al. 2014), a 24 bp sequence located 60 – 36 base pairs upstream of the transcription initiation site of *HLA3* was selected as the EBE, with a nucleotide T preceding the sequence (Figure 2). To avoid obvious off-target effects, we searched the Chlamydomonas genome by BLAST for additional sequence similarities to the candidate EBE, and found no additional matching sequence of comparable length (Altschul et al. 1990). The dTALE construct pHLA3-2177 was then introduced into plasmids described previously to contain Zeocin resistance (*Zeo^R*) as a selection marker, Chlamydomonas chimeric promoter RbcS2/Hsp70, and full-length TALE with an artificially designed repeat region (Gao et al. 2014; Figure 2B). We first transformed wild type wall-less strain cw10 (CC829, mt -) with pHLA3-2177. After initial selection for *Zeo^R*, insertion of the complete dTALE cassette into the Chlamydomonas genome was confirmed in the selected transformants with colony PCR (data not shown; primers indicated in Table 1).

Of the 60 *Zeo^R* transformants with a potentially intact pHLA3-2177 construct, 5 strains, namely dTHLA3-4, dTHLA3-5, dTHLA3-19, dTHLA3-43, and dTHLA3-60, exhibited elevated transcript abundance both of *dTALE-HLA3* and of *HLA3*, which normally is repressed when wild type cells are grown under high CO₂ (Figure 3A). HLA3 antibody also detected dTALE-mediated HLA3 expression under high CO₂ (Figure 3B). To quantify the expression levels of *HLA3* in pHLA3-2177 transformed strains as well as in wild type under different CO₂ levels, we used qRT-PCR analysis with primers specific for HLA3 cDNA (Table 1). Using the expression of *CBLP* as an internal control, the transcript abundance of *HLA3* in transgenic and wild type lines were compared by ΔΔCt as defined previously (Gao et al. 2014): ΔCt=Ct_{HLA3}-Ct_{CBLP}, and ΔΔCt=ΔCt_{transgenicline}−ΔCt_{wildtype}. Compared to high CO₂ acclimated wild type (cw10) cells, significant increase of *HLA3* transcript abundance was achieved in all five dTALE-HLA3 transformed lines (Figure 3C, ANOVA, p<0.005).

**HLA3 Expression Impact on Ci-dependent O₂ Evolution.** To examine the effect of activated HLA3 expression on Ci uptake and photosynthesis, we measured Ci-dependent photosynthetic O₂ evolution of dTALE-HLA3 transformant lines in comparison with the cw10 progenitor line. All cell lines were acclimated to high CO₂, under which conditions the
CCM is repressed, allowing the contribution of HLA3 to Ci uptake to be tested with a nearly null baseline level.

At pH 7.3, all five transformants (dTHLA3-4, 5, 19, 43, 60) showed significantly (MANOVA, p<=0.05) higher Ci dependent O2 evolution rates compared to the wild type cw10 cells in Ci concentrations below 100 µM, although the increase appears fairly low when compared with the photosynthetic O2 evolution of very low CO2 acclimated wild type cells (Figure 4). The most marked increase occurs at 50 µM, and diminishes gradually between 50 µM and 100 µM (Figure 4 A-F). In addition, the calculated K50Ci of O2 evolution for each of the transformant strains is significantly lower than that of the wild type strain cw10 (Figure 4G).

To further investigate the function of HLA3 in different genetic backgrounds, we chose two mutant strains, LA43 and cia5, to cross with dTALE-HLA3 activated strains. LA43 constitutively expresses LCIA (Fig. 5B), a putative HCO3- transporter predicted to be located on the chloroplast envelope (Miura et al. 2004; Mariscal et al. 2006; Duanmu et al. 2009a), from an introduced transgene. The cia5 mutant contains a point mutation in CIA5, which encodes the master regulator of Chlamydomonas CCM. As a result, cia5 is unable to induce a CCM and is not viable under very low CO2. Two dTALE-HLA3 activation strains, dTHLA3-19 and dTHLA3-60, were used as parents in these crosses. Resulting progeny were collected, and their genetic background as well as the intactness of dTALE-HLA3 were examined. In crosses with LA43, the presence of the LCIA overexpression cassette was tested with colony PCR (data not shown), and the constitutive expression of both HLA3 and LCIA was confirmed with RT-PCR, qRT-PCR, and/or protein immunoblots. In crosses with cia5, the mutation in CIA5 was reconfirmed by PCR and amplicon sequencing, and the expression of HLA3 was also confirmed both at the transcript and the protein levels. As indicated in Figure 5, the dTALE expression and activated HLA3 expression are inherited and maintained in the progeny generated from these crosses.

The Ci dependent photosynthetic O2 evolution rate of LA43 and cia5, progeny from the above-mentioned crosses, and wild type strains 21gr and cw10 are shown in Figure 6. The two wild-type strains exhibited very similar Ci response curves, and, interestingly, no
obvious increase in photosynthetic rate was observed in high CO$_2$ acclimated LA43 cells, where $\text{LCIA}$ alone is overexpressed (Figure 6A and B). We selected two progeny each from crosses between LA43 and both dTHLA3-19 and dTHLA3-60, namely HLLA19-1, HLLA19-18, HLLA60-16, and HLLA60-26. Photosynthetic O$_2$ evolution rates suggested that in all four strains, a combination of overexpression of both $\text{LCIA}$ and $\text{HLA3}$ leads to increased photosynthesis in high CO$_2$ acclimated cells (Figure 6C-F). The increase was similar but more significant compared with $\text{HLA3}$ overexpression alone, and exhibited a similar pattern, with the highest increase observed at 50 $\mu$M Ci. The calculated photosynthetic K$_{5}$Ci of all four strains based on Ci response from 10 $\mu$M to 100 $\mu$M Ci are significantly lower compared to the K$_{5}$Ci of wild type, as well as to those of both parental strains (Figure 6G; ANOVA, p≤0.05). To further investigate the impact of the combined $\text{HLA3}$ and $\text{LCIA}$ overexpression on Ci uptake, we compared the [$^{14}$C]-Ci uptake in HLLA progeny with those of the wild type 21gr (Fig.7). The 4 HLLA progeny tested each exhibited an increased Ci uptake into the internal Ci pool compared to the wild type (Figure 7; ANOVA, p≤0.05).

We also selected four strains that exhibited activated $\text{HLA3}$ expression from the crosses between $\text{cia5}$ and either dTHLA3-19 or dTHLA3-60, namely HLCI19-1, HLCI19-28, HLCI60-14, and HLCI60-28, for Ci-dependent O$_2$ evolution experiments. Despite the obvious expression of $\text{HLA3}$ at both the transcript and the protein levels (Figure 5), Ci-dependent O$_2$ evolution in these strains did not show a significant difference when compared to the $\text{cia5}$ mutant (Figure 8).

Discussion

TALE-mediated, gene-specific activation has been successfully applied in different organisms, including plants and mammals (Morbitzer et al. 2010; Zhang et al. 2011; Li et al. 2013; Maeder et al. 2013), as well as in Chlamydomonas (Gao et al. 2014). In this study, activation was successful for ~8% of the transformants generated by introduction of the dTALE, which is similar to previous observations (Gao et al. 2014). In crosses between successful dTALE activation lines and LA43 or $\text{cia5}$, about 20%-25% of progeny possessing the antibiotic selective marker exhibited dTALE activation detectable by RT-PCR and
protein immunoblot. This demonstrates clearly that the dTALE activation can be transmitted genetically to progeny, although the level of genetic transmission is lower than the 100% that would be expected if all dTALE copies were linked to the antibiotic resistance and fully expressed when transmitted to progeny from the dTALE activated parental strains. Of the Zeo\textsuperscript{R} progeny examined, only \(~80\%\) retained \textit{dTALE-HLA3} expression detectable by RT-PCR, suggesting that gene silencing may be occurring in about 20\% of the dTALE-HLA3 containing lines. Even considering a 20\% silencing rate, it is not clear why the transmission rate is only 20-25\%, but it is possible that, in addition to the apparent gene silencing, the dTALE transgenes may simply be less effective in the progeny. In support of this possibility, the transcript abundance of \textit{HLA3} is \(~2^{4}\) fold lower in the progeny with activated \textit{HLA3} expression than that in the parental dTALE-HLA3 transgenic lines (Figure 3C, Figure 5C and F), which suggests that the effectiveness of the dTALE activation was diminished even when transmitted.

HLA3 has long been considered a possible Ci transporter in the Chlamydomonas CCM, because of its limiting CO\textsubscript{2} induction, its placement in the ABC transporter gene family and its predicted plasma membrane location (Im and Grossman 2002; Duanmu et al. 2009). However, neither the function of HLA3 in Ci transport nor its plasma membrane location has been confirmed before now. The alkaline-pH-dependent inhibition of photosynthesis in \textit{HLA3} knockdown strains and the retarded growth phenotype of \textit{HLA3-LCIA} co-knockdown strains specifically under alkaline pH suggest that HLA3 is mainly involved in the active uptake of HCO\textsubscript{3}\textsuperscript{-} rather than of CO\textsubscript{2} (Duanmu et al. 2009). The clear impact of pH on the phenotype of the \textit{HLA3} knockdown strains is consistent with a plasma membrane location, and the strong enrichment of HLA3 in the plasma membrane fraction from purified microsome membranes (Fig. 1) confirms the predicted plasma membrane localization. The exact role of HLA3 in active Ci uptake, however, remains unclear, as direct evidence for a transport function has been lacking.

The increased Ci dependent photosynthesis rate in high CO\textsubscript{2} acclimated, \textit{HLA3}-overexpressing cw10 lines shown here supports the notion that HLA3 contributes to active Ci uptake in Chlamydomonas. However, it is clear that \textit{HLA3} expression under high CO\textsubscript{2} conditions, where other CCM components are not induced or up-regulated, provides only a
small stimulation of the photosynthetic $O_2$ evolution rate, suggesting that HLA3 alone is not very effective in driving up the $CO_2$ concentration available to Rubisco in the absence of other induced CCM components (Figure 4). In the HLLA progeny, when activated HLA3 expression was combined with overexpression of LCIA, more evident increases were observed in Ci-dependent $O_2$ evolution. Compared to the $O_2$ evolution rate when the CCM is induced, however, the increase caused by activation/overexpression of HLA3 and LCIA is still relatively low, suggesting that induction or expression of other CCM components is required to achieve substantial increases in Ci uptake and photosynthesis (Figure 6).

In agreement with the observed increase in $O_2$ evolution, $[^{14}C]$-Ci uptake was also increased in the HLLA progeny. The observed 0.3-0.4 mM internal Ci in HLLA progeny when 50$\mu$M Ci was provided, which represents a 40-85% increase in $[^{14}C]$-Ci accumulation, relative to wild type, indicates the ability of HLA3, in combination with LCIA, to generate an internal $HCO_3^-$ pool. However, the observed internal Ci accumulation still is relatively low compared to the 2-3 mM Ci routinely observed with low-$CO_2$ or very low-$CO_2$ acclimated wild type Chlamydomonas (Spalding et al. 1983; Duanmu et al. 2009a, 2009b), which argues that additional CCM components not induced in high $CO_2$ acclimated cells are required for full function of $HCO_3^-$ transport. In addition, a lower relative rate of photosynthetic Ci affinity demonstrated by $O_2$ evolution suggests that additional CCM components not induced in high $CO_2$ acclimated cells also are required for the cells to effectively utilize the accumulated Ci.

The ratio of the two physiologically relevant species of Ci, $CO_2$ and $HCO_3^-$, varies greatly as the pH of the medium changes. Chlamydomonas responds to changes in Ci sources with at least three acclimation states, previously defined as very low $CO_2$, low $CO_2$ or air level of $CO_2$, and high $CO_2$ (Wang and Spalding, 2006). Previous work on the air dier LCIB mutants has demonstrated that LCIB or the LCIB/LCIC complex is primarily responsible for Ci uptake in the defined low $CO_2$ concentration range (Wang and Spalding 2006; Duanmu et al. 2009b), and that LCIB probably is responsible for active $CO_2$ uptake rather than $HCO_3^-$ transport (Duanmu et al. 2009b; Wang and Spalding 2014). Since LCIB mutants grow relatively normally in very low $CO_2$, some CCM component or components other than LCIB must facilitate Ci uptake in the very low $CO_2$ concentration range. At $pH$ 7.3, when 10 - 100 $\mu$M Ci is provided, $HCO_3^-$ is the dominant species, and $CO_2$ availability is very limited (1 - 9
μM), and corresponds with the previously defined very low CO₂ acclimation range (Vance and Spalding 2005; Wang and Spalding, 2006), suggesting that Ci assimilation at pH 7.3 in very low CO₂ concentrations is likely dependent on HCO₃⁻ transport rather than active CO₂ uptake.

The observation that the O₂ evolution rate is increased in Chlamydomonas lines with HLA3 overexpression alone or with combined overexpression of both HLA3 and LCIA within the very low CO₂ concentration range (10-100µM Ci or 1-9µM CO₂), but not at higher Ci concentrations, therefore, strongly implies that HLA3 is mainly involved in active uptake of HCO₃⁻ rather than of CO₂.

The impact of HLA3 overexpression and HLA3 plus LCIA overexpression on photosynthetic Ci assimilation under very low CO₂ conditions agrees with the previously proposed role of HLA3 in HCO₃⁻ uptake (Duanmu et al. 2009a), and also suggests that HLA3 might function in tandem with LCIA. LCIA is proposed to be responsible for HCO₃⁻ uptake across the chloroplast envelope, especially under very low CO₂ conditions, as supported by compromised photosynthesis rates observed under very low CO₂ concentrations and alkaline pH in LCIA/HLA3 RNAi knockdown strains (Duanmu et al. 2009a). Similar to LCIA, HLA3 appears to function mainly in HCO₃⁻ transport and is active when CO₂ concentration is very low. It is likely that HLA3 and LCIA work sequentially and synergistically to maintain the Ci pool by actively transporting HCO₃⁻ when CO₂ concentration is very low. Overexpression of LCIA alone did not result in detectable photosynthetic O₂ evolution rate increase. This should not be surprising, since any LCIA-mediated HCO₃⁻ transport at the chloroplast envelope would rely on cytoplasmic HCO₃⁻ being provided by active Ci uptake across the plasma membrane, which is likely provided by HLA3 or other potential limiting CO₂ inducible transporters not present in the high CO₂ acclimated cells used here.

Activation of HLA3 expression in the cia5 mutant did not seem to increase the photosynthetic O₂ evolution rate, in spite of the apparently elevated expression of HLA3 on both the transcript and the protein levels. CIA5 is a zinc finger protein that regulates the expression of thousands of genes, including CO₂ regulated genes, as well as CO₂-insensitive genes that involved in other pathways such as terpenoid metabolism (Fukuzawa et al. 2001; Xiang et al.
The expression profiles of the cia5 mutant and high CO₂ acclimated wild type cells are notably different at the transcript level, and the expression levels of most CO₂ regulated genes are even lower than their repressed expression levels in high CO₂ acclimated wild type cells. The observation that HLA3 expression in the cia5 mutant failed to increase photosynthesis rate implies that the CCM function in cia5 differs from that in high CO₂ acclimated wild type, even though the CCM is not induced in either, and that the impact of CIA5 lies beyond induction of the CCM alone. It is possible that, via some mechanism that is not yet clear, components either directly needed for HLA3 activity or indirectly needed to utilize the HCO₃⁻ transported by HLA3 are regulated by CIA5 and require its presence even in high CO₂.

Following the original proof-of-concept study (Gao et al. 2014), the successful heritable activation of HLA3 expression in Chlamydomonas demonstrates again that artificially designed TALEs can be used as effective tools for targeted gene activation in this model microalga. In agreement with previous reports (Duanmu et al. 2009a), the confirmed subcellular localization of HLA3 and increased Ci dependent photosynthetic O₂ evolution in HLA3 activation lines indicate that HLA3 is indeed involved in active Ci uptake across the plasma membrane. Also in agreement with previous studies regarding the role of HLA3 and the acclimation states of the Chlamydomonas CCM (Duanmu et al. 2009a), the observation that HLA3 mediated increase of O₂ evolution occurs in the range of very low CO₂ implies that HLA3 is mainly involved in HCO₃⁻ uptake, and that HLA3 likely functions sequentially and synergistically with LCIA. The suggested transition from HLA3/LCIA involved HCO₃⁻ uptake in very low CO₂ to LCIB/LCIC involved CO₂ uptake in air level of CO₂ might represent an effective strategy utilized by the Chlamydomonas CCM to respond to the environmental carbon sources in a rapid, cost efficient manner.

Material and Methods

Cell Strains and Culture Conditions.

Chlamydomonas reinhardtii strains cw10 (CC 849, mt -), CC125, and 21gr (CC 1690, mt +) were obtained from the Chlamydomonas culture center, Duke University, Durham, NC. The LCIA overexpression line LA43 was generated by introducing a LCIA expression cassette
into 21gr (see below). The *cia5* mutant was a gift from Dr. Donald Weeks (University of Nebraska, Lincoln).

All cells were maintained on agar plates containing minimal medium (Geraghty et al. 1990) and were kept in high-CO$_2$ (\(~5\%\) CO$_2$ in air) growth chambers at room temperature. Liquid cultures were grown on a gyratory shaker at 175 RPM in TAP medium under 60 μmol photons m$^{-2}$ s$^{-1}$ illumination (Gorman and Levine, 1965). For very low CO$_2$ acclimation, cells grown in high CO$_2$ to late log phase were switched to very low CO$_2$ (0.005%-0.01%) by bubbling with a controlled ratio of CO$_2$-free air (air filtered through saturated sodium hydroxide) and normal air (Wang and Spalding, 2006).

**Membrane fractionation.**

Membrane fractionation and separation of the plasma membrane procedure followed the method described by Norling et al. in 1996. 10L of cw10 cells culture were collected and broken with a N2 bomb. The crude was first centrifuged at 1,200 g for 2.5 min, resulting in a pellet containing unbroken cells, chloroplasts, and cell debris. The supernatant was centrifuged at 13,000 g for 5 min, resulting in a pellet containing mitochondria and thylakoid. The second supernatant was then centrifuged at 125,000 g for 1 hour to obtain a pellet containing total microsome. This pellet was then applied to a two-phase system of 6.5% Dextran, 6.5% polyethylene glycol 3350, 0.25 M sucrose, 5mM KH2PO4, and 4mM KCl. After three times of repartitioning, the lower phase was collected by centrifugation at 125,000 g for 1 hour.

**Construction of the dTALE and LA43 LCIA Overexpression Cassettes.**

The dTALE construction was performed using a library of repeat modules (Li et al. 2013). The dTALE-HLA3 was constructed as previously described using an altered TAL plasmid pSKavrXa7delta and a Chlamydomonas expression vector pDW2177 (Gao et al. 2014) to create final construct pHLA3-2177, which was subsequently transformed to cw10 (CC849, mt -) cells for dTALE expression in Chlamydomonas.
For the LCIA overexpression cassette, the Hsp70/Rbcs2 chimeric Chlamydomonas promoter was amplified from pSI103Δ plasmid (Sizova et al. 2001) to introduce an NdeI restriction site at the 3’ end. 2kb of genomic DNA containing the LCIA gene was PCR amplified with an NdeI restriction site at the 5’ end of the start codona EcoRI restriction site after at the 3’ end after the stop codon. The Hsp70/Rbcs2 chimeric promoter and the LCIA PCR fragment were then ligated to each other by NdeI and subsequently ligated into pGenD plasmid (Fischer and Rochaix 2001) with NotI/EcoRI. The AphIIIIV gene from pSI103Δ was then inserted downstream of the LCIA overexpression cassette for paromomycin resistance. The resulting plasmid was used to transform Chlamydomonas cells, and transformants resistant to paromomycin were selected to test for LCIA overexpression.

**Gene Expression Analysis.**

RNA was extracted with an RNAeasy extraction kit (Qiagen), and the RNA concentration was measured using an ND-1000 Nanodrop spectrophotometer (Nanodrop Technologies). The cDNA derived from 50 ng of total RNA (Super Script III First-Strand Synthesis System; Invitrogen) was used for reverse transcription polymerase chain reaction (RT-PCR) and for quantitative RT-PCR (qRT-PCR). Primer sequences are provided in Table 1. The qRT-PCR was performed on an iCycler iQ real-time PCR detection system (Biorad) using a SYBR green one-step quantitative PCR system (Quanta Biosciences). Amplicons from qRT-PCR were sequenced to ensure specificity of amplifications. 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_targetgene - Ct_CBLP to represent the difference between the transcript abundance of genes examined and the transcript abundance of CBLP. After normalization, ΔCt values of each transformant were compared with that of wild type cells, and the dTALE induced expression is represented by ΔΔCt = ΔCt_transformant - ΔCt_wildtype. (Li et al., 2013). For direct comparisons, fold change is also used to represent the relative abundance of target genes, which equals 2^{ΔΔCt}.

For western immunoblotting, total protein was obtained by harvesting cells grown to late-log phase and directly dissolving the pellets in 1x SDS-PAGE buffer followed by boiling for 5 minutes. A rabbit HLA3 antibody was generated against a synthetic peptide
EANASLQRLEAYLLEEC at the protein facility at Iowa State University. Proteins were separated by SDS-PAGE on 12% polyacrylamide gels, and immunoblotting using antibody described above was performed at 1:750 dilution as described in the protocol from Bio-Rad Laboratories (catalog no. 500-0006).

**HLA3 Overexpression Strain Selection and Genetic Crosses.**

Overexpression transformants were first selected with 10µg/ml zeocin on TAP medium agar plates, and then confirmed with colony PCR for presence of the complete dTALE construct (Primers listed in Table 1). Expression analysis was performed as described above.

Crosses between HLA3 overexpression lines and CC125, LA43, 21gr, and cia5 were performed as described by Harris et al (2009). Progeny carrying both Ble and Aph8 were selected on TAP-agar plates with 10 µg/ml zeocin and 15 µg/ml paromomycin. Progeny carrying only Ble were selected on plates supplemented with 10 µg/ml zeocin. Progeny carrying the cia5 mutation were identified by PCR amplification (Table 1).

Photosynthetic O₂ Evolution and Ci Uptake Measurements and Data Analysis

Photosynthetic O₂ evolution was measured as previously described (Duanmu et al. 2009a). At 25 °C, the O₂ evolution rate was obtained with a Clark-type O₂ electrode controlled by the software Oxy-lab (Hanstech). Cells grown to early log phase were collected by centrifuging at 1,000g for 7 minutes and resuspended in 25 mM MOPS-Tris (pH7.3) for a final chlorophyll concentration of 20 µg/ml. The cell suspension was placed in the oxygen electrode chamber under illumination of 500 µmol photons. m⁻². s⁻¹, and frequently flushed with N₂ gas. The depletion of internal and external Ci in the chamber was determined by the cessation of O₂ production. A series of measurements were obtained by adding different concentrations of NaHCO₃ to the chamber. For each measurement, every dose of NaHCO₃ was added to the chamber only once the O₂ evolution rate from the previous dose reached nearly zero, except for the last addition of NaHCO₃ of 16mM for the measurement of the Vₘₐₓ. The K₁/₂ Ci estimates, Ci concentration at which ½ of the Vₘₐₓ was reached, were
calculated by fitting the Ci dependent O2 evolution plots to rectangular hyperbola enzyme kinetic curves with simulation provided by the software Graphpad.

Ci uptake at 50 µM was measured by the silicone oil filtration method (Badger et al. 1980). The cell volume was estimated using ¹⁴C-sorbitol and ³H₂O, as was described previously (Heldt et al. 1980).

O₂ evolution data of each strain examined was obtained from three biological replicates. Two-way ANOVA was performed between each pair of lines examined with the software StatPlus, and the difference is considered significant when p ≤ 0.05. Ci accumulation data was collected from 3 replicates at each time point, and ANOVA was performed for Ci accumulation at each time point as well as for the K₅/₂ estimates. Standard errors were in each figure to indicate the range of variance.
Literature Cited


Buttner D, Bonas U (2002). Getting across--bacterial type III effector proteins on their way to the plant cell. EMBO J., 21, 5313-5322.


Figure Legends

**Figure 1.** Subcellular localization of HLA3. Enrichment of HLA3 in the chloroplast membrane is shown by wild type (cw10) cells grown under very low CO₂. Chloroplast protein CCP2 and D1, and plasma membrane protein H⁺-ATPase are used to demonstrate effective separation of different membrane fractions. H, high CO₂. VL, very low CO₂.

**Figure 2.** Structure of pHLA3-2177 and the targeting position of dTALE-HLA3. Upper panel: schematic drawing of pHLA3-2177. The construct contains full length sequences of the N-terminus and C-terminus of avrXa7, but with the avrXa7 DNA recognition repeats replaced with 23 repeats designed for the HLA3 effector binding element (EBE). BleRK, selective marker for zeocin resistance. Yellow and green lines across the drawing indicate positions of the nuclear localization signal and the activation domain, respectively, of avrXa7. Middle panel: EBE sequence of HLA3 and the corresponding dTALE repeat RVDs. Lower panel: relative position of the HLA3 EBE on chromosome 2 of Chlamydomonas.

**Figure 3.** Expression analysis of dTALE mediated HLA3 gene activation with intact dTALE constructs. Number and letters below each figure refer to different cell strains. 4, 5, 19, 43, 60: transgenic strains dTHLA3-4, dTHLA3-5, dTHLA3-19, dTHLA3-43, and dTHLA3-60. cw10: wall-less wild type strain cw10 (CC829). H CO₂: high CO₂. VL CO₂: very low CO₂. A, RT-PCR showing presence of dTALE-HLA3 and HLA3 transcripts in p2177-HLA3 transformed lines. CBLP is a constitutively expressed gene and is used as the internal control. B, Protein immunoblot of p2177-HLA3 transformed lines, using rabbit anti-HLA3 antibody. C, Quantification of HLA3 expression on the transcript level by qRT-PCR. The relative transcript abundance, ΔΔCt, is defined using CBLP as an internal control, where ΔCt=CtHLA3-CtCBLP, and ΔΔCt=ΔCttransformant-ΔCtwild type. Three repeats were performed for each strain, and HLA3 transcript abundance in all 5 dTHLA3 transgenic strains were significantly higher than that in high CO₂ acclimated cw10 cells (ANOVA, p<0.005).
**Figure 4.** Photosynthetic O$_2$ evolution (A-F) and K$_{1/2}$ Ci estimates (G) of wild type cells (cw10) and five p2177-HLA3 transformed lines exhibiting activated HLA3 expression: cw10 (A), dTHLA3-5 (B), dTHLA3-43 (C), dTHLA3-19 (D), dTHLA3-60 (E), dTHLA3-4 (F). Open circles in A: very low CO$_2$ acclimated cw10 cells. Closed circles in A: high CO$_2$ acclimated cw10 cells. Open circles in B-F: high CO$_2$ acclimated dTHLA3 transgenic lines. Closed circles in B-F: high CO$_2$ acclimated cw10. Solid black bar in G: high CO$_2$ acclimated wild type cells (cw10). Solid grey bar in G: very low CO$_2$ acclimated wild type cells. Open bars in G: high CO$_2$ acclimated dTHLA3 transgenic strains. Three replicates of each strain were used to obtain each figure, and MANOVA was performed to estimate the significance of the difference. All five dTHLA3 transgenic strains examined exhibited significant increase in O$_2$ evolution between 10-100µM Ci provided (p ≤ 0.05).

**Figure 5.** Expression analysis of dTALE mediated HLA3 gene activation heritable in progeny resulting from crosses between two dTHLA3 lines and LA43 or cia5 on the transcript level (A, C, D, F) and the protein level (B and E). Numbers and letters under A, B, and C refer to HLLA progeny and wild type. LA43: LCIA overexpression strain. 21gr: wild type (CC1690). Numbers and letters under D, E, and F refer to HLCI progeny, cia5 mutant, and wild type 21gr (CC1690). H CO$_2$: high CO$_2$. VL CO$_2$: very low CO$_2$. The relative transcript abundance in C and F, ΔΔCt, is defined using CBLP as an internal control, where ΔCt=Ct$_{HLA3}$-Ct$_{CBLP}$, and ΔΔCt=ΔCt$_{transmant}$-ΔCt$_{wild}$ type. All progeny in C and F were high CO$_2$ acclimated. Three repeats were performed for each strain, and the transcript abundance of HLA3 in both HHLA and HLCI progeny were significantly higher than that of wild type (21gr) tested by ANOVA (p<0.005).

**Figure 6.** Photosynthetic O$_2$ evolution of wild type strains, high CO$_2$ acclimated LA43 (A and B) and HLLA progeny (C-F), as well as K$_{1/2}$ estimates calculated from these strains (G). A: O$_2$ evolution of high CO$_2$ (closed squares) and low CO$_2$ (open squares) acclimated wild type strain 21 gr. B: Photosynthetic O$_2$ evolution of high CO$_2$ acclimated wild type strains cw10 (closed circles) and 21gr (closed squares), and LA43 (closed triangles). C-F: Photosynthetic O$_2$ evolution of high CO2 acclimated HLLA60-16, HLLA60-26, HLLA19-1, and HLLA19-18 (open triangles in C-F, respectively) in comparison to high CO$_2$ acclimated
wild type strain 21gr (closed triangles). G: $K_{1/2}$ Ci estimates of each strain. Three repeats were included to generate each figure. High $CO_2$ acclimated HLLA $K_{1/2}$ estimates were significantly lower than that of high $CO_2$ acclimated 21gr (ANOVA, $p \leq 0.05$).

**Figure 7.** $^{14}$C Ci accumulation of high $CO_2$ acclimated HLLA progeny and wild type strain 21gr. $^{14}$C uptake was performed at pH7.3 with 50µM NaHCO$_3$ provided initially. Solid bars represent Ci accumulation in 21gr. Open bars represent Ci accumulation HLLA progeny, each strain indicated by numbers inside the bars. 60-16: HLLA60-16. 60-26: HLLA60-26. 19-18: HLLA19-18. 19-1: HLLA 19-1. 3 repeats were included for each strain. Ci accumulation of HLLA progeny significantly different from that of 21gr are indicated with asterisks (ANOVA, $p \leq 0.05$).

**Figure 8.** Photosynthetic $O_2$ evolution of wild type strain 21gr and high $CO_2$ acclimated HLCI progeny (A-D), and $K_{1/2}$ estimates calculated (F). A-D: Photosynthetic $O_2$ evolution of high $CO_2$ acclimated HLCI60-14, HLCI60-28, HLCI19-1, HLCI19-19 (open diamonds in A-D, respectively) in comparison to high $CO_2$ acclimated 21gr (closed diamonds). F: $K_{1/2}$ estimates of each strain calculated based on A-D. $O_2$ evolution of the HLCI progeny and 21gr are not significantly different.
<table>
<thead>
<tr>
<th>Primer</th>
<th>Gene</th>
<th>Sequence</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>94790PC</td>
<td>dTALE-HLA3-5'</td>
<td>GGCCTACCCGAATTCGAAGC</td>
<td>Colony/RT-PCR</td>
</tr>
<tr>
<td>94792PC</td>
<td>dTALE-HLA3-3'</td>
<td>AACAGATCTGCCACCATGGA</td>
<td>Colony/RT-PCR</td>
</tr>
<tr>
<td>114944PC</td>
<td>LA43 Cassette</td>
<td>TGATCTTCTCTTGCTGCTGC</td>
<td>Colony/RT-PCR</td>
</tr>
<tr>
<td>111532PC</td>
<td>HLA3</td>
<td>CTACCTCCCCCTGATCTCCTCT</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>20106PC</td>
<td>HLA3</td>
<td>CTCCGAGCGTTCGTTTGTTC</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>72092PC</td>
<td>LCIA</td>
<td>GCGCTCATCGAGAAGAAGG</td>
<td>qRT-PCR</td>
</tr>
</tbody>
</table>

Table 1. Primers used in this work
Figure 1.
Figure 2.
Figure 3.
Figure 5.
Figure 6.
Figure 7.
Figure 8.
CHAPTER 4

dTALE MEDIATED ACTIVATION OF β-CAROTENE KETOLASE AND β-CAROTENE HYDROXYLASE IN ATTEMPT TO INCREASE ASTAXANTHIN AMONG IN THE GREEN ALGA CHLAMYDOMONAS REINHARDTII

All dTALE constructs used in this chapter were contributed by Dr. David Wright. All transgenic strains of the dTALE constructs were also generated by Dr. David Wright.

Abstract

Astaxanthin is a carotenoid that is widely appreciated for its potential health benefits and its value in the aquaculture industry. It is naturally accumulated in very few organisms, including some microalgae species that grow slowly and are hard to maintain. Synthesis of astaxanthin in the microalga Chlamydomonas would be of great potential, in that the model organism grows rapidly and is easy to manipulate genetically. We attempted to induce astaxanthin production in Chlamydomonas by specifically activating the expression of two key enzymes involved in the biosynthesis of astaxanthin, β-carotene ketolase (BKT) and β-carotene hydroxylase (CHYB), with artificially designed transcription activator like effectors (dTALEs). While the expressions of the target genes were activated on the transcript level, no accumulation of astaxanthin was observed in Chlamydomonas transformants. However, this preliminary work provides insights in possible improvements in dTALE mediated gene specific activation for future directions.

Introduction

Carotenoids, including xanthophylls and carotenes, are isoprenoids produced by plants and other photosynthetic species. It is commonly believed that they are involved light harvesting and in protection against photo-oxidative stress (Armstrong and Hearst1996; Ruban et al. 2012; Ramel et al. 2013). The keto-carotenoid astaxanthin has attracted worldwide interest because of its recognized importance in aquaculture, as well as its potential health benefit in food supplements. As a strong anti-oxidant, astaxanthin has been suggested to be beneficial
in the treatment of human diseases (Guerin et al. 2003; Fassett et al 2009; Lennikov et al. 2012). Unlike many other carotenoids, the natural sources of astaxanthin are very limited. It is only synthesized in a number of bacteria, cyanobacteria, yeast, and a few microalgal species including *Haematococcus pluvialis* and *Chlorella zofingiensis* (Misawa et al. 1995; Fernandez-Gonzalez 1997; Boussiba et al. 1999; Visser et al. 2003; Huang 2005). Currently, the most dominant source of natural astaxanthin is the green alga *H. pluvialis*, whose doubling time is two weeks, and only accumulates astaxanthin when exposed to photooxidative stress, making the scale-up production of astaxanthin costly and inefficient (Boussiba et al. 2000).

The green alga *Chlamydomonas reinhardtii* (hereafter, Chlamydomonas), has long been used as a model organism in both fundamental biological studies as well as biomedicine and biofuel production (Grossman et al. 2004; Merchant et al. 2007; Corchero et al. 2013). Under easily achievable lab conditions, the doubling time of Chlamydomonas liquid cultures is less than ten hours, and the cells can be quickly synchronized and behave homogenously (Dent et al. 2001). All three genomes of Chlamydomonas have been fully sequenced, and the molecular techniques for genetic manipulation in Chlamydomonas are mature (Merchant et al. 2007). Biosynthesis of astaxanthin in Chlamydomonas, if successful, would be of tremendous potential economically and commercially.

The biosynthesis of astaxanthin, initiated by the synthesis of the first intermediate carotenoid phytoene, requires the involvement of many enzymes. \( \beta \)-carotene ketolase (BKT) and \( \beta \)-carotene hydroxylase (CHYB) are the two key enzymes catalyzing the final steps of astaxanthin synthesis (Figure 1, Cordero et al. 2012). Because most photosynthetic species do not posses BKT activity, it is generally considered the limiting factor in astaxanthin biosynthesis. Multiple attempts have been made to achieve synthesis of astaxanthin in Chlamydomonas, mostly by introducing foreign genes from algal species that naturally synthesize astaxanthin to Chlamydomonas cells. However, no successful accumulation of astaxanthin has been achieved yet. Expression of the BKT encoding gene from *H. pluvialis*, *bkt1*, resulted in accumulation of 4-keto-lutein or ketozeaxanthin, but not zeaxanthin or astaxanthin (Leon et al. 2007). Introduction of a phytoene synthase gene from *C. zofingiensis*
in Chlamydomonas increased the content of carotenoids violaxanthin and lutein, but not keto-carotenoids such as astaxanthin (Cordero et al. 2011).

The Chlamydomonas genome does contain genes annotated as encoding β-carotene ketolase (BKT) and β-carotene hydroxylase (CHYB), but under favorable growth conditions, the expression of both BKT and CHYB in wild type Chlamydomonas cells are nearly undetectable by RNA-Seq (Fang et al. 2012). When cells are exposed to photooxidative stress, the expression of genes involved in carotenoid biosynthesis are upregulated, resulting in accumulation of zeaxanthin as a way to protect the photosynthetic machinery from photooxidative damage (Baroli et al. 2003; Grossman et al. 2005). In the cia5 mutant, a zinc finger protein that regulates the expressions of hundreds of other genes, CIA5, is not functional, and the expression of BKT is shown to be 4-5 folds higher than that in the wild type (Fukuzawa et al. 2001; Fang et al. 2012). When the BKT gene from various sources was cloned and transferred to E.coli, the Chlamydomonas BKT exhibited high conversion rate of zeaxanthin to astaxanthin (85%), much higher than its homologs isolated from many other green algae species, including H. pluvialis and S. oocystiformis (Huang et al. 2012). Given the undesired results of attempts to synthesize astaxanthin by introducing foreign genes to Chlamydomonas, and given the expression patterns of CHYB and BKT as well as the enzymatic activity of BKT in E.coli, it is a reasonable strategy to induce the expression of endogenous genes that encode astaxanthin synthesis enzymes.

Transcription activator-like effectors (TAL effectors, or TALEs) are effector proteins secreted by plant pathogenic bacteria species Xanthomonas when infecting its hosts. Delivered through type III secretion systems, TALEs function like transcription activators by recognizing and binding onto the promoter region of and activating the expression of certain host genes, causing disease susceptibility or resistance responses of the host plants (Zhu et al. 2000; Buttner et al. 2002, 2006). The recognition pattern between naturally occurring TALEs and their target genes has been well deciphered. Each TALE is composed of a N-terminal domain containing a translocation signal, a C-terminal domain containing an activation motif and a nuclear localization signal, and central repeats that are essential for DNA recognition (Bogdanove et al. 2010; Boch and Bonas 2010; White and Yang, 2009; Wright et al. 2014). The repeats contain 33 amino acid (aa) residues and are all nearly identical except for the 12th
and 13th aa residues, known as repeat variable diresidues (RVDs). Each pair of RVDs specifically binds one single nucleotide, and the four most predominant type of RVDs observed in nature are HD, NG, NN, and NI, recognizing C, T, G (or A), and A, respectively (Boch et al. 2009, Moscou and Bogdanove 2009). Compared with earlier editing tools, such as zinc finger proteins and mega nucleases, the truly modular recognition pattern of TALEs have made them superior tools for targeted gene activation and genome editing. By substituting the original central repeats with artificially designed repeats targeting desired gene promoters, designed TALEs (dTALEs) have been created and successfully used in plant, algal, human, and other mammalian cells to mediate gene specific activation (Li et al. 2013; Maeder et al. 2013; Gao et al. 2014). The combination of dTALEs and endonuclease FokI has also achieved great success in site directed mutagenesis and genome editing in plants, zebra fish, C. elegans, human cells, and other mammalian cells (Cermak et al. 2011, Cade et al. 2010, Li et al. 2011, 2012).

In this study, we use dTALEs targeting the 5’ promoter region of BKT and CHYB in Chlamydomonas, in an attempt to induce the synthesis of astaxanthin. We used CC4033, a yellow-in-the-dark mutant that is deficient in chlorophyll synthesis and turns light yellow when grown in the dark (Meinecke et al. 2010), to facilitate screening for possible accumulation of astaxanthin. Despite the activation of BKT and CHYB expression on the transcript level, no accumulation of astaxanthin was observed. However, analysis into possible reasons why no detectable astaxanthin was accumulated provides preliminary data for further investigation in production of astaxanthin and further application of dTALEs in Chlamydomonas.

Results and Discussion
Sequences 20 to 24 base pairs long from upstream of the transcription initiation site of BKT and CHYB genes were selected as dTALE effector binding elements (EBEs), using criteria described previously (Gao et al. 2014). Sequences from 27 to 7 base pairs (BKT EBE1) and 55 to 35 base pairs (BKT EBE2) upstream of the transcription initiation site were used as EBEs for BKT. Sequences from 60 to 36 base pairs (CHYB EBE1) and 84 to 60 base pairs (CHYB EBE2) upstream of transcription initiation were used as EBEs for CHYB. Two sets of
dTALEs were constructed: the first set included a Hygromycin resistance selective marker (HygR), the Chlamydomonas Hsp70A/Rbcs2 chimeric promoter, a dTALE activation cassette derived from the natural TALE pthXo1, and the Chlamydomonas PSAD terminator, and are named pDW2969 and pDW2970, targeting BKT EBE1 and BKT EBE2, respectively (Figure 2A). The second set was constructed to include the same HygR selective marker and the promoter/terminators as the first set, but included two tandemly positioned dTALE cassettes targeting BKT EBE1 and CHYB EBE1, BKT EBE1 and CHYB EBE2, BKT EBE2 and CHYB EBE1, and BKT EBE2 and CHYB EBE2, respectively, and are named pDW3321, pDW3322, pDW3323, and pDW3324 (Figure 2B). The dTALE cassettes derived from the natural TALE pthXo1 was edited to remove the codons for its first 152 amino acid residues at the N-terminus (Miller et al. 2011), and codon-optimized the remaining N-terminus to fit the Chlamydomonas codon usage preference.

Potentially complete dTALE-BKT constructs were confirmed by colony PCR, with sequence specific primers at positions indicated in Figure 2A. 45 transformant colonies, about 50% of the ZeoR transformant colonies tested, appeared to harbor both fragments at the 3’ end and the 5’ end of the dTALE-BKT construct. We examined the expression of the target gene, BKT, on the transcript level, with RT-PCR and qRT-PCR. 5 strains containing a potentially intact dTALE construct showed increased expression of the BKT gene at the transcript level. Based on qRT-PCR results, the activated transcript expression levels are 32 to 64 fold higher than that of the wild type, and are comparable to that of cia5 mutant (Figure 3A and B). None of the dTALE-BKT activation transformants exhibited noticeable color change, despite the yellow-in-the-dark phenotype of the host strain. When grown on TAP medium, no color change was observed in colonies of these transformants. We performed thin-layer-chromatography (TLC) assay on these 5 strains, as well as on wild type strain CC125 and the host strain CC4033. In both transgenic and non-transgenic lines, β-carotene was accumulated to a high level, but no obvious accumulation of astaxanthin was observed (data not shown).

Two short paths exist in the final steps of the biosynthesis of astaxanthin from β-carotene: one is through zeaxanthin, the other through canthaxanthin (Figure 1). Given the apparent abundance of the substrate of astaxanthin synthesis, β-carotene, we used the second set of
dTALE plasmids, which contains dTALE cassettes targeting both BKT and CBLP EBEs, to activate the expression of both genes. To screen for potentially intact dTALE cassettes, Hyg\textsuperscript{R} transformants were tested by colony PCR with four primer pairs indicated in Figure 2B. Primer pair 1532/1533 was used for the first round of colony PCR, and the resulting positive transformants were subsequently tested with primer pair 1529/1531, followed by the other two primer pairs towards the 5’ end of the dTALE-BKT-CHYB construct. Interestingly, the number of colony PCR-positive transformants decreased in each round of colony PCR. 550 colonies were originally tested, 93 appeared to have the DNA fragment amplified by primer pair 1532/1533, out of which only 39 were amplified by 1529/1531, and the number of positive transformants decreased to 33 and eventually 22, as the other two primer pairs towards the 5’ end were used. This is different from what was expected, since the 5’ end of the construct carries the antibiotic resistance selective marker. Aside from the possibility that colony PCR is not 100% effective, it appears that foreign DNA introduced to Chlamydomonas cells tend to be randomly sliced (and not chewed down from one end) by some unknown mechanism, and the possibility of this digestion or degradation increases with the size of the foreign DNA. Similar DNA digestion on a much larger scale and a possible mechanism was observed and proposed by Zhang et al. in 2014, by which the Chlamydomonas cells deploy endonuclease(s) to digest foreign DNA fragments introduced. This is especially noteworthy when future DNA transformations with Chlamydomonas are performed.

The 22 transformant strains potentially containing an intact construct were examined for the activation of BKT and/or CHYB expression on the transcript level. Compared to that in the wild type, 8 showed increased expression of BKT, and 9 showed increased expression of CHYB. Out of these transformants, 3 showed increased expression of both BKT and CHYB (Figure 3C). qRT-PCR results indicate that the expression of the target genes on the transcript level are generally 16 to 32 fold higher than that in the wild type cells (Figure 3D). However, neither transformants exhibiting activated expression of either BKT or CHYB, nor those that exhibited activated expression of both genes exhibited any apparent color change. Astaxanthin is rarely accumulated in Chlamydomonas or most other photosynthetic species, and the regulation of its biosynthesis is not completely understood. The Chlamydomonas
BKT, the essential enzyme in synthesizing astaxanthin, showed high efficiency in converting zeaxanthin to astaxanthin when expressed in *E.coli* (Huang et al. 2012). It is possible that, despite the possible expression of BKT, was not effectively delivered to the subcellular location near the vicinity of β-carotene. However, given the presence and abundance of β-carotene, it is also likely that the expressions of *BKT* and/or *CHYB* mediated by dTALEs are too low to elicit enough functional enzymes to catalyze the reactions. dTALE mediated gene activation efficiency observed here is much lower than previously reported in Chlamydomonas when more than 10,000 fold increase was achieved with different genes (Gao et al. 2014).

The reason why the activation efficiency is so low is unclear. One possible cause may be a property of the target genes. For example, their positions on the chromosome, or the characteristics of the upstream regions of these genes might restrict their expression level. However, it seems unlikely that both genes would have the same or similar characteristics.

In previous successful dTALE activation of endogenous genes in Chlamydomonas, the complete N-terminus and C-terminus of a native TALE, AvrXa7, was used. It is possible the N-terminal truncation reduces the effectiveness of the native TALEs, although the first 152 amino acids only contains host cell translocation signal, which presumably is dispensable for the activation function. Others have found that similarly truncated TALE nucleases (TALENs), which also include N-terminal truncations, appear not to be as effective as full-length bacterial TALENs in plant genome editing experiments (Wright and Spalding unpublished), so there may be additional features associated with the N-terminal and C-terminal parts of TALEs than we currently understand. dTALE activation using a full-length, native TALE backbone targeting the same genes of Chlamydomonas should provide insights into answering these questions, and may possibly achieve higher expression of the target genes and accumulation of astaxanthin.

**Material and Methods**

**Cell strains and culture conditions.** *Chlamydomonas reinhardtii* strain CC125 and the yellow-in-the-dark mutant strain CC4033 was obtained from the Chlamydomonas culture center, (University of Minnesota, St. Paul, MN). All cells were maintained on agar plates
containing minimal medium (Geraghty et al. 1990) and were kept in high-CO$_2$ (~ 5% CO$_2$ in air) growth chambers at room temperature. Liquid cultures were grown on a gyratory shaker at 175 RPM in TAP medium under 60 μmol photons m$^{-2}$ s$^{-1}$ illumination (Gorman and Levine, 1965).

**Plasmids and primers used.** The dTALE repeats were assembled using the Golden Gate method as described previously (Gao et al. 2014). The repeats were ligated into plasmid pDW2871 (Wright and Spalding, unpublished) with restriction enzymes SphI and AatII, resulting in a dTALE cassette including Chlamydomonas codon optimized, 5’ truncated N-terminus (Δ152, Miller et al. 2011) and native C-terminus of pthXo1, and the assembled repeats, called pDW2965. This was subsequently digested with BglII and SpeI, and ligated into the Chlamydomonas expression vector pDW2178 (Wright and Spalding, unpublished) so that the dTALE cassette is under the control of Chlamydomonas chimeric promoter Hsp70A/Rbcs2 and selectable by hygromycin resistance.

**Gene expression analysis.** RNA was extracted with an RNAeasy extraction kit (Qiagen), and the RNA concentration was measured using an ND-1000 Nanodrop spectrophotometer (Nanodrop Technologies). The cDNA derived from 50 ng of total RNA (Super Script III First-Strand Synthesis System; Invitrogen) was used for reverse transcription polymerase chain reaction (RT-PCR) and for quantitative RT-PCR (qRT-PCR). The qRT-PCR was performed on an iCycler iQ real-time PCR detection system (Biorad) using a SYBR green one-step quantitative PCR system (Quanta Biosciences). Amplicons from qRT-PCR were sequenced to ensure specificity of amplifications. 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$_{\text{target gene}}$ - Ct$_{\text{CBLP}}$ to represent the difference between the transcript abundance of genes examined and the transcript abundance of CBLP. After normalization, ΔCt values of each transformant were compared with that of wild type cells, and the dTALE induced expression is represented by ΔΔCt = ΔCt$_{\text{transformant}}$ - ΔCt$_{\text{wildtype}}$ (Li et al., 2013). For direct comparisons, fold change is also used to represent the relative abundance of target genes, which equals $2^{\Delta\Delta C t}$. The primers used for RT-PCR and qRT-PCR are shown in Table 1.
Literature Cited


Buttner D, Bonas U (2002). Getting across--bacterial type III effector proteins on their way to the plant cell. EMBO J., 21, 5313-5322.


Figure Legends

**Figure 1.** Adapted from Leon et al. (2006). Biosynthesis pathways of ketocarotenoids including astaxanthin. BKT, β-carotene ketolase. CHYB, β-carotene hydroxylase.

**Figure 2.** Schematic drawings of dTALE constructs used to activate *BKT* and/or *CHYB* expression. Hyg<sup>R</sup>: hygromycin selective marker. hsp70/Rbcs2P: Chlamydomonas hsp7-A/Rbcs2 chimeric promoter. PSADT: Chlamydomonas *PSAD* terminator. Pink arrows and letters above each drawing indicate the positions and directions of primers used for colony PCR or RT-PCR. EBE: effector binding element. Angled arrows and “+1”: transcription initiation sites. A: dTALE constructs used to activate *BKT* expression. B: dTALE constructs used to activate the expression of both *BKT* and *CHYB*. Note that the length of the bars do not correspond with the actual length of DNA sequences of the constructs.

**Figure 3.** Expression analysis of dTALE mediated expression activation of BKT and/or CHYB. A: RT-PCR showing the presence of dTALE-BKT and *BKT* transcripts, in comparison with wild type and the *cia5* mutant. B: qRT-PCR result showing the transcript abundance of *BKT* induced by dTALE-BKT, using *CBLP* as an internal control. The transcript abundance is defined as ΔCt=Ct<sub>targetgene</sub>-Ct<sub>CBLP</sub>, and ΔΔCt=ΔCt<sub>transformant</sub>-ΔCt<sub>wildtype</sub>. C: RT-PCR showing the presence of dTALE-*BKT/CHYB* and *BKT* or *CHYB* transcript. D and E: qRT-PCR result showing the transcript abundance of *BKT* and *CHYB*, respectively, induced by dTALE-BKT/CHYB. 4033: yellow-in-the-dark mutant CC4033.
<table>
<thead>
<tr>
<th>Primer</th>
<th>Gene</th>
<th>Sequence</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEP1528</td>
<td>dTALE</td>
<td>AGGTTCATCACGCCGTTTGT</td>
<td>Colony PCR</td>
</tr>
<tr>
<td>AEP1529</td>
<td>dTALE</td>
<td>CGGTCTGTTTATAGGTGACATTGAG</td>
<td>Colony PCR</td>
</tr>
<tr>
<td>AEP1530</td>
<td>dTALE</td>
<td>GCCCTACGCCTTTGTGACCA</td>
<td>Colony PCR</td>
</tr>
<tr>
<td>AEP1531</td>
<td>dTALE</td>
<td>GTTCGAATGAAAGAAGACCTCGAG</td>
<td>Colony PCR</td>
</tr>
<tr>
<td>AEP1532</td>
<td>dTALE</td>
<td>CCAGCACCCTGATCTCTGGGAA</td>
<td>Colony/RT-PCR</td>
</tr>
<tr>
<td>AEP1533</td>
<td>dTALE</td>
<td>GTGCCACCTGACCCTTAAGGT</td>
<td>Colony/RT-PCR</td>
</tr>
<tr>
<td>108552</td>
<td>BKT</td>
<td>GTATCCCGCACCAGCTCCGC</td>
<td>RT-PCR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TTGCCTCTACACACCTCAAG</td>
<td></td>
</tr>
<tr>
<td>109067</td>
<td>BKT</td>
<td>AGTAGGGAAATAAGGGACAT</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AACGATACCGTTCAATTCC</td>
<td></td>
</tr>
<tr>
<td>111539</td>
<td>CHYB</td>
<td>CGCCGCCTTTGCTTTTACAT</td>
<td>RT-PCR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TCGGATGTTTTCTGGGCGTG</td>
<td></td>
</tr>
<tr>
<td>111534</td>
<td>CHYB</td>
<td>CACAGACTCACGCCACCGTG</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GATAGCTTCTCTCCTCGGG</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1.
Figure 2.
Figure 3.
CHAPTER 5
GENERAL SUMMARY

General Summary

Since the breakthroughs in deciphering the DNA recognition code transcription activator-like effectors (TALEs) in 2009, great achievements have been made in genome editing and gene specific activation in multiple organisms. The true modular characteristic of TALEs allows for easy design and assembly of artificially designed TALEs (dTALEs), which are created by replacing the recognition repeat domain of native TALEs with artificially assembled repeats targeting the promoter regions of genes of interest.

In Chlamydomonas, dTALEs targeting endogenous genes ARS1, ARS2, and HLA3 using complete N-terminus and C-terminus of the native TALE AvrXa7 mediated effective activation of the target genes, where the transcript abundance increased more than 10,000 fold compared to that in wild type cells. This suggests that dTALEs can be used as a powerful tool for gene specific activation in Chlamydomonas. dTALEs constructed with truncated N-terminus and complete C-terminus of pthXo1 targeting BKT and CHYB, however, did not appear as effective in gene activation. Only 32-64 fold increase in transcript abundance was achieved in dTALE transgenic lines compared to that in wild type, in clear contrast with the dTALEs included complete amino acid (aa) sequences of both N-terminus and C-terminus. While the characteristics of the target genes could have caused the difference in the level of activation achieved, the difference in the N-terminus might also explain the loss of activation ability of the BKT and/or CHYB targeting dTALEs. Despite its dispensability demonstrated by studies in human and yeast cells, the first 152 aa of the N-terminus of native pthXo1 that is truncated in the BKT/CHYB targeting dTALEs might be required for effective activation in Chlamydomonas, although it is only known to contain the translocation signal related to the host infection process of Xanthomonas. Experiments with dTALEs containing full length C-terminus and N-terminus of native TALEs targeting BKT
and CHYB could help to understand whether possible unknown functions of the N terminus of TALEs are needed for effective dTALE-mediated gene activation in Chlamydomonas. It is also noteworthy that when dTALE harboring Chlamydomonas strains are crossed with other strains, the resulting progeny tend to exhibit lower level and lower frequency of dTALE mediated activation compared to the expected situation, in which full activation capacity is linked to the antibiotic selective marker and transmitted through crosses to all progeny containing the transgenes. In the dTALE-HLA3 activation study, when the dTHLA3 strains were crossed with either LA43 or cia5, the expression of the dTALE appeared to have been silenced in some progeny, since there was no transcript of the dTALE detectable by RT-PCR. In some other progeny, the dTALEs were expressed on the transcript level, but appeared to be inefficiently translated, or have lost the ability to activate the target gene HLA3. In the HLLA progeny where the expression and activation function were both retained, the level of dTALE mediated expression of the target gene were generally 16 – 32 fold lower than that in the parental dTHLA3 strains. It is possible that multiple copies of dTALEs were inserted into the dTHLA3 strains during transformation, and worked synergistically in activating the expression of HLA3. Some of the copies might have been lost or silenced through crosses, and the remaining copy (ies) were not able to produce enough amount of dTALEs to induce a detectable activation of the target gene. Examining the copy numbers in the parental dTHLA3 as well as the progeny strains will help in testing this hypothesis.

Chlamydomonas provides researchers an excellent system to study the eukaryotic CO₂ concentrating mechanism (CCM). Although many important genes involved in the Chlamydomonas CCM have been identified, the specific functions of these genes, and especially the transition between the three acclimation states (high CO₂, low (air-level) CO₂, and very low CO₂) have long been puzzling. The increase in inorganic carbon (Ci) dependent photosynthetic O₂ evolution in dTHLA3 compared to that in wild type is limited to the Ci range of 0-100 µM (0-10 µM CO₂) at pH7.3. This CO₂ range corresponds well with the previously defined very low CO₂ state, suggesting that HLA3 is mainly involved in the uptake of HCO₃⁻ across the plasma membrane, which is in agreement with previous discoveries.
The combination of overexpression/activation of both HLA3 and LCIA resulted in more evidently increased Ci dependent O$_2$ evolution, and the increase was also limited to the Ci range of 0-100 $\mu$M at pH7.3. It has been suggested previously that the chloroplast membrane protein LCIA is also involved in the uptake of HCO$_3^-$.

Recent work on the LCIB mutant, the LCIA mutant, and LCIB/LCIA double mutant revealed presumptive roles of LCIB and LCIA (and possibly other Ci transporters) in the CCM, and suggested a model that explains the transition between different acclimation states (Wang and Spalding, unpublished). In the very low CO$_2$ range, HCO$_3^-$ uptake appears to play a dominant Ci uptake role. In this CO$_2$ range, LCIA presumably functions in HCO$_3^-$ uptake, either as a transporter or a channel. When CO$_2$ concentrations reach air level (>7 $\mu$M), HCO$_3^-$ uptake mediated by LCIA was shown to be clearly inhibited, although the mechanism of the inhibition is not yet clear. At air level of CO$_2$, presumably, the LCIB/LCIC complex becomes the major contributor to total Ci uptake in the form of CO$_2$, in agreement with the lethality of the LCIB mutant in air. The Ci range in which increased O$_2$ evolution was observed here suggests that HLA3 might function in a pattern similar to that of LCIA. It is possible that, similar to LCIA, HLA3 is mainly involved in HCO$_3^-$ uptake and is active when CO$_2$ concentration is very low, and that as more CO$_2$ becomes available, the active uptake of Ci by HLA3 is inhibited by CO$_2$, and Ci uptake is taken over by other components of the CCM responsible for CO$_2$ uptake, namely the LCIB/LCIC complex.

Based on the above-mentioned findings, a more up-to-date model of the Chlamydomonas CCM has been suggested (Wang and Spalding unpublished). In this model, Chlamydomonas cells acclimate to varying Ci concentrations by varying which components in the CCM are active. Under very low CO$_2$, HLA3, LCIA, and possibly other components involved in HCO$_3^-$ transport are largely responsible for the Ci uptake, although CO$_2$ uptake via the LCIB/LCIC complex also contributes (Figure 1A). As the CO$_2$ concentration reaches air level, the components involved in HCO$_3^-$ uptake are inhibited through an unknown mechanism, and the LCIB/LCIC complex functions in the chloroplast stroma as the main contributor to the internal Ci pool, possibly by unidirectionally converting CO$_2$ to HCO$_3^-$ (Figure 1B). When the CO$_2$ concentration is above air level, CO$_2$ diffusion might be sufficient to support survival and growth of the Chlamydomonas cells. The regulation mechanism behind this fast transition of roles played by different CCM components between
different environmental CO2 concentrations is unclear. However, compared to regulation on the transcript level, which transitions between gene induction (at higher CO2) and gene repression (at lower CO2) at CO2 concentrations of about 0.5%, this regulation may allow the cells to promptly respond to the environmental CO2 availability in a cost-efficient way.
Figure Legends

Figure 1. Adapted from unpublished work Wang and Spalding (2014). Working model illustrating possible roles played by HLA3- and LCIA-involved HCO$_3^-$ uptake system and the LCIB-associated CO$_2$ uptake system in very low CO$_2$ acclimated cells (A) and low CO$_2$ acclimated cells (B). PM: plasma membrane. ChE: chloroplast envelope.
Figure 1.