2016

The Characterization of Alpha- and Beta-Carbonic Anhydrases of Arabidopsis thaliana

Robert John DiMario
Louisiana State University and Agricultural and Mechanical College, rjdimario@gmail.com

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_dissertations

Recommended Citation
https://digitalcommons.lsu.edu/gradschool_dissertations/3758

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
THE CHARACTERIZATION OF α- AND β-CARBONIC ANHYDRASES OF ARABIDOPSIS THALIANA

A Dissertation

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Doctor of Philosophy in

The Department of Biological Sciences

by

Robert John DiMario
B.S., Louisiana State University, 2010
May 2016
ACKNOWLEDGEMENTS

I would like to thank Dr. James Moroney for initially taking me in as an undergraduate to work in his lab. Without this step, I might not have pursued scientific research. I also thank Dr. Moroney for agreeing to be my major professor in graduate school and for being a great advisor and colleague.

I would like to thank Dr. David Longstreth for teaching me how to obtain gas exchange rates from Arabidopsis leaves, for providing indispensable feedback on my projects, and for agreeing to be on my PhD committee. Dr. Longstreth helped and interacted with me so much that he had become a co-major professor.

I would like to thank Dr. John Larkin, Dr. Sue Bartlett, Dr. Aaron Smith, and Dr. Jong Hyun Ham for all agreeing to be a part of my PhD committee. All four professors provided crucial insight towards data analysis, experimental design and troubleshooting, and provided insight for future career options in research. The conversations I have had with all of my PhD committee members have kept up my enthusiasm for research.

I would also like to thank Dr. Bernard Genty for his advice and input into this project. Also, I would like to thank Dr. Maheshi Dassanayake and Dr. Dong-Ha Oh for providing their RNAseq data for this project as well as their input into this project. I want to thank Alice Simmons for teaching me plant techniques and for helping me start my plant project.

I would like to thank all the members of the Moroney lab for making the lab environment such a wonderful place to work. Tiffany Lindsey was the RA of the lab when I joined as an undergraduate and she taught me most of the basic lab techniques. Susan Laborde has kept our lab organized and running and helped perform many experiments. Yunbing Ma, Bratati Mukherjee, Joanna Bajsa-Hirschel, and Mary Machingura all provided guidance while I was a
graduate student. Ananya Mukherjee has been a good friend in the lab and a great person with whom to discuss scientific ideas. Monica Hartman started the Arabidopsis carbonic anhydrase project and was the one who originally convinced me to work with Arabidopsis over Chlamydomonas. Wes Frey provided interesting scientific conversations. Kristen Bice, Michael Sonnier, Dequantarius Speed, and Jennifer Quebedeaux are all past undergraduates who have helped extensively with the Arabidopsis carbonic anhydrase project. Also, I would like to thank the current members of the Arabidopsis project, Madelinn Fink, Amanda Quattrone, Kayla Schwartzenburg, and Lance Pounds, for helping with this project and for being great co-workers. I wish you all the best of luck with this project.

I want to thank my sister, Rachel, my friends, and Annette Hebert. I am indebted to you all for supporting me, for working your schedules around mine, and for the memories we made throughout my graduate career.

I want to thank my parents for all of their loving support and guidance. They taught me to “run to daylight” and through their encouragement for every endeavor I undertook, I have found it.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .................................................................................................................. ii  
LIST OF TABLES .......................................................................................................................... v  
LIST OF FIGURES ........................................................................................................................ vi  
ABBREVIATIONS ......................................................................................................................... ix  
ABSTRACT ..................................................................................................................................... x  

CHAPTER  
1 INTRODUCTION ......................................................................................................................... 1  
2 MATERIALS AND METHODS ...................................................................................................... 31  
3 CYTOSOLIC CARBONIC ANHYDRASE IS REQUIRED FOR OPTIMAL GROWTH IN LOW CO$_2$ .............................................................................................................................. 42  
4 THE CHLOROPLAST CARBONIC ANHYDRASES $\beta$CA1 AND $\beta$CA5 ..................................... 71  
5 THE $\alpha$CAS OF ARABIDOPSIS THALIANA .................................................................................. 92  
6 IMPROVEMENTS TO C$_3$ PHOTOSYNTHESIS .......................................................................... 112  
7 CONCLUSIONS ............................................................................................................................ 123  

REFERENCES ................................................................................................................................... 130  

APPENDIX A. LIST OF PRIMERS FOR GENOMIC PCR AND RT-PCR ........................................ 147  
APPENDIX B. LIST OF CLONING PRIMERS .................................................................................. 148  
APPENDIX C. PERMISSION TO REPRINT .................................................................................... 149  

VITA ................................................................................................................................................. 151
**LIST OF TABLES**

Table 1.1. List of Arabidopsis carbonic anhydrase genes .......................................................... 24

Table 3.1. $\beta$CA2, $\beta$CA3, and $\beta$CA4 are all expressed in roots and leaves .................................. 44

Table 3.2. Predicted subcellular locations of $\beta$CA1, $\beta$CA2.1, and $\beta$CA2.2 using protein subcellular location prediction programs ................................................................. 46

Table 3.3. Predicted locations of $\beta$CA4.1 and $\beta$CA4.2 using four different protein targeting programs ................................................................. 50

Table 3.4. Gas exchange values of wildtype (COL), $\beta$ca2, $\beta$ca4, and $\beta$ca2$\beta$ca4 plants ............. 58

Table 3.5. Levels of free amino acid pools of wildtype (COL), $\beta$ca2, $\beta$ca4, and $\beta$ca2$\beta$ca4 plants grown at low CO$_2$ ........................................................................................................ 61

Table 3.6. Free amino acid levels of wildtype (COL) and $\beta$ca2$\beta$ca4 plants grown in high CO$_2$ (1,000 μL L$^{-1}$ CO$_2$) ........................................................................................................ 63

Table 4.1. Predicted subcellular locations of $\beta$CA1 and $\beta$CA5 using protein subcellular location prediction programs ................................................................. 80

Table 4.2. $\beta$CA1 and $\beta$CA5 are expressed in roots and leaves .................................................. 82

Table 5.1. $\alpha$CA1, $\alpha$CA2, and $\alpha$CA3 are the only $\alpha$CAs expressed in Arabidopsis leaves and roots ........................................................................................................ 94

Table 5.2. Predicted subcellular locations of $\alpha$CA1, $\alpha$CA2, and $\alpha$CA3 using protein subcellular location prediction programs ................................................................. 100

Table 6.1. Predicted subcellular location of CAH3 ............................................................................ 115
LIST OF FIGURES

Figure 1.1. The carboxylation and oxygenation reactions of Rubisco .................................. 10

Figure 1.2. The products of the Rubisco carboxylation reaction are subsequently phosphorylated and reduced in the Calvin Benson Basham cycle ............................................. 10

Figure 1.3. Modeled atmospheric concentrations of CO₂ and O₂ spanning the history of the Earth .................................................................................................................................................. 12

Figure 1.4. A simplified version of the proposed carbon concentrating mechanism model of Chlamydomonas ........................................................................................................... 14

Figure 1.5. A simplified model of the carbon concentrating mechanism of cyanobacteria ...... 16

Figure 1.6. The fatty acid synthesis pathway ........................................................................... 20

Figure 1.7. The carboxylation reaction of phosphoenolpyruvate carboxylase and downstream aminotransferase reaction to generate aspartate ............................................................................................................................ 22

Figure 1.8. A proposed model of C₁ accumulation and CO₂ diffusion based on carbonic anhydrase activity derived from the βCAs of Arabidopsis .............................................................................................................. 29

Figure 3.1. Aligning RNAseq reads from root and shoot RNA samples to the Arabidopsis reference genome ....................................................................................................................... 45

Figure 3.2. Aligning the peptide sequence of the chloroplast carbonic anhydrase, βCA1, with the two βCA2 peptide sequences, βCA2.1 and βCA2.2 using Clustal Omega ................................. 45

Figure 3.3. Leaf and root RNA samples yielded two forms of βCA4 mRNA ...................................... 46

Figure 3.4. Maps of promoter regions used to drive GUS expression ........................................ 47

Figure 3.5. GUS staining of the various pβCA2::GUS, pβCA3::GUS, and pβCA4::GUS plants .............................................................................................................................................................................. 48

Figure 3.6. Clustal Omega alignment of the peptide sequences of βCA4.1 and βCA4.2 ............ 50

Figure 3.7. eGFP fluorescence from leaf tissue of various stably expressing eGFP lines .......... 51

Figure 3.8. eGFP fluorescence of protoplasts generated from stably expressing eGFP Arabidopsis lines .............................................................................................................................................. 52

Figure 3.9. The location of T-DNA insertions in the βCA2 and βCA4 genes of the βca2 and βca4 mutant lines, respectively ................................................................................................................................. 53

Figure 3.10. Genotyping of the βca2, βca4, and βca2βca4 mutants ........................................ 54

Figure 3.11. Two Western blots using an antibody raised against spinach carbonic anhydrase . 55
Figure 3.12. The growth phenotype of the βca2βca4 double mutant ...........................................56

Figure 3.13. The βca2βca4 double mutant shows reduced growth in low CO₂ over a 10 week study at 200 μL L⁻¹ CO₂. ........................................................................................................57

Figure 3.14. The stomatal density of 10 week old carbonic anhydrase single and double mutant plants are unchanged from wildtype values ..........................................................59

Figure 3.15. Free amino acid levels in wildtype (COL), βca2, βca4, and βca2βca4 plants .............60

Figure 3.16. Free amino acid levels in wildtype (COL) and βca2βca4 plants grown in high (1,000 μL L⁻¹ CO₂) ........................................................................................................62

Figure 3.17. Expressing the βCA2 coding region in βca2βca4 plants restores wildtype growth in low CO₂ .........................................................................................................................64

Figure 3.18. Genomic PCR probing for the p35S::βCA2 construct in the βca2βca4 mutant .........65

Figure 4.1. Locations of the T-DNA insertions within the βCA1 and βCA5 genes ....................73

Figure 4.2. The βCA1 and βCA5 genes are disrupted by T-DNA insertions .................................74

Figure 4.3. The βCA1 protein is missing in the βca1 mutant .......................................................75

Figure 4.4. The predicted mature chloroplast carbonic anhydrase, βCA2.1, is nearly identical in size and sequence to βCA1 ........................................................................................................76

Figure 4.5. The T-DNA insertions of the βca1 and βca5 mutant lines have reduced transcription of the βCA1 and βCA5 genes, respectively ............................................................77

Figure 4.6. Growth histograms of wildtype (COL), βca1, and βca5 plants in ambient air (400 μL L⁻¹) and in very high CO₂ (50,000 μL L⁻¹)... .................................................................78

Figure 4.7. Wildtype (COL), βca1, and βca5 plants grown in ambient (400 μL L⁻¹) and very high CO₂ (50,000 μL L⁻¹) ........................................................................................................79

Figure 4.8. The cotyledons of βca5 mutant plants already show a distinct phenotype as compared to wildtype (COL) plants ..................................................................................80

Figure 4.9. Both βCA1 and βCA5 are chloroplast proteins .........................................................81

Figure 4.10. There are two mRNA forms of βCA1 that differ at the 3’ ends .................................82

Figure 4.11. The RNAseq reads match well with the TAIR βCA5 gene model ............................83

Figure 4.12. The promoter regions used to make the pβCA1::GUS and pβCA5::GUS constructs .................................................................................................................................83

Figure 4.13. Three-day old βCA1 and βCA5 seedlings show different expression patterns .......84
Figure 4.14. Three week old pβCA1::GUS and pβCA5::GUS plants still show different expression patterns ................................................................. 85

Figure 4.15. Multiple Independent pβCA1::GUS and pβCA5::GUS lines show different expression patterns in rosettes ................................................................. 86

Figure 4.16. Protein alignments of βCA1 and βCA5 ................................................................. 89

Figure 5.1. The RNAseq data present evidence that the αCA2 gene has an extra 5’ exon and two extra 3’ exons as compared to the TAIR αCA2 reference gene ..................... 95

Figure 5.2. Upstream promoter regions used for the (A) pαCA1::GUS, (B) pαCA2::GUS, and (C) pαCA3::GUS constructs ................................................................. 96

Figure 5.3. GUS expression powered by the various αCA promoter regions ..................... 96

Figure 5.4. GUS staining is prevalent in the trichomes of pαCA2::GUS plants ..................... 97

Figure 5.5. αCA3 has a point mutation that is near the 3’ end of the message ..................... 98

Figure 5.6. The truncated αCA3 may still be active ............................................................... 99

Figure 5.7. Transient eGFP expression assays in tobacco show αCA1-eGFP localizes to the periphery of epidermal cells ............................................................... 100

Figure 5.8. The αCA1 antibody does not pick up the αCA1 protein ..................................... 101

Figure 5.9. Locations of T-DNA insertions in the αCA1 and αCA2 genes ......................... 103

Figure 5.10. The aca1 and aca2 genes are disrupted by the presence of T-DNA insertions..... 103

Figure 5.11. RT-PCR of wildtype (COL), aca1, and aca2 samples ..................................... 104

Figure 5.12. aca1 and aca2 single mutants grow similarly to wildtype (COL) plants in a low CO2 (200 μL L⁻¹) environment ............................................................... 105

Figure 5.13. Rosette areas of the double mutants are smaller than wildtype (COL) at the end of a 10 week growth period at low CO2 ......................................................... 106

Figure 5.14. aca1βca4 and aca2βca4 plants are slightly smaller than wildtype (COL) plants growing in low CO2 (200 μL L⁻¹) at the end of a 10 week growth study ......... 106

Figure 6.1. A plan for implementing a simplified version of the algal carbon concentrating mechanism in plant chloroplasts to improve photosynthesis rates ..................... 114

Figure 6.2. The Chlamydomonas carbonic anhydrase, CAH3, locates to the chloroplasts of Arabidopsis protoplasts ............................................................... 116
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-PGA</td>
<td>3-phosphoglycerate</td>
</tr>
<tr>
<td>ACCase</td>
<td>Acetyl-CoA carboxylase</td>
</tr>
<tr>
<td>AIR carboxylase</td>
<td>phosphoribosylaminoimidazole carboxylase</td>
</tr>
<tr>
<td>Ambient CO₂</td>
<td>400 μL L⁻¹ CO₂</td>
</tr>
<tr>
<td>CA</td>
<td>carbonic anhydrase</td>
</tr>
<tr>
<td>CBB</td>
<td>Calvin Benson Basham cycle</td>
</tr>
<tr>
<td>CCM</td>
<td>carbon concentrating mechanism</td>
</tr>
<tr>
<td>Cᵢ</td>
<td>inorganic carbon</td>
</tr>
<tr>
<td>CO₂</td>
<td>carbon dioxide</td>
</tr>
<tr>
<td>CPS</td>
<td>carbamoyl phosphate synthetase</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>GUS</td>
<td>β-glucuronidase</td>
</tr>
<tr>
<td>HCO₃⁻</td>
<td>bicarbonate</td>
</tr>
<tr>
<td>High CO₂</td>
<td>1,000 μL L⁻¹ CO₂</td>
</tr>
<tr>
<td>Low CO₂</td>
<td>200 μL L⁻¹ CO₂</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PEPC</td>
<td>Phosphoenolpyruvate Carboxylase</td>
</tr>
<tr>
<td>Rubisco</td>
<td>Ribulose-1,5-Bisphosphate Carboxylase Oxygenase</td>
</tr>
<tr>
<td>RuBP</td>
<td>ribulose-1,5-bisphosphate</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>TE</td>
<td>Tris-EDTA</td>
</tr>
</tbody>
</table>
ABSTRACT

Carbonic anhydrases (CAs) are zinc-metalloenzymes that interconvert two inorganic carbon (C\textsubscript{i}) species, CO\textsubscript{2} and HCO\textsubscript{3}\textsuperscript{-}. In *Arabidopsis thaliana*, there are eight αCA genes, six βCA genes, three γCA genes, and two γCA-like genes. The majority of CA research in plants has focused on finding a link between CA activity and photosynthesis rates. Since the CA genes are expressed in different plant tissues and multiple CA isoforms are distributed among various organelles of the plant cell, I hypothesize that CAs facilitate CO\textsubscript{2} diffusion among cell compartments and maintain C\textsubscript{i} pools for carbon-requiring reactions by interconverting CO\textsubscript{2} and HCO\textsubscript{3}\textsuperscript{-}. This thesis focuses on the αCAs and βCAs of Arabidopsis and how they may affect various reactions throughout the plant. CA T-DNA insertion lines were used to determine if removing one or more CAs from Arabidopsis affects the plant growth. The βca5 single mutant and βca2βca4, αca1βca4, and αca2βca4 double mutants show different growth phenotypes. The βca2βca4 plants were smaller in size and chlorotic in their younger leaves under low CO\textsubscript{2} conditions, but showed improved growth in high CO\textsubscript{2} conditions. The growth of the βca5 single mutant was severely stunted in ambient CO\textsubscript{2} conditions and high CO\textsubscript{2} partially rescued wildtype growth in the βca5 plants. The αca1βca4 and αca2βca4 double mutants were slightly smaller than wildtype plants in low CO\textsubscript{2} conditions. Interestingly, it seems the reduced growth of the βca5 single mutant and βca2βca4 double mutant plants was not linked to deficiencies in photosynthesis rates but rather may be required for other carbon requiring reactions. These results suggest that CAs are playing more complex roles in plants than once thought and that the various isoforms are affecting different carbon-requiring pathways.
CHAPTER 1
INTRODUCTION

THE THREE CA TYPES IN PLANTS

Carbonic anhydrases (CAs) are zinc-metalloenzymes that catalyze the interconversion of CO$_2$ and HCO$_3^-$.

CAs are ubiquitous in nature and are an example of convergent evolution, as multiple, structurally and sequentially distinct, families of CA have been discovered (Hewett-Emmett and Tashian, 1996). *Arabidopsis thaliana*, the model organism used in this work, has three types of CA: α-, β-, and γ-CAs. The α-type CA (αCA) was first found in erythrocytes and was the first CA family discovered (Brinkman et al., 1932; Meldrum and Roughton, 1932; Meldrum and Roughton, 1933a; Meldrum and Roughton, 1933b; Stadie and O’Brien, 1933). Most αCAs form monomers, although there are reports of αCA dimers (Whittington et al., 2001; Hilvo et al., 2008; Suzuki et al., 2010; Suzuki et al., 2011; Cuesta-Seijo et al., 2011). The majority of the enzyme is comprised of ten β-strands that create a large central β-sheet which is surrounded by seven α-helices on the periphery of the protein (Liljas et al., 1972). The N-terminus of the mature αCA is composed of two α-helices that are substantially exposed to their surroundings. Interestingly, these two α-helices contain very few interactions with the rest of the protein, although they do play a role in the stability of the enzyme (Aronsson et al., 1995).

Removing the first five amino acids of the mature human αCA2 (CAII) led to a loss of protein stability and a 70% reduction in activity, although the enzyme was still very active with a measured $K_{cat}$ of $10^5$ sec$^{-1}$ (Aronsson et al., 1995). Removing the first 28 amino acids of the mature protein led to a complete loss of enzyme activity (Aronsson et al., 1995). The zinc at the αCA active site is coordinated by three histidine residues and one water molecule organized in a tetrahedral conformation (Liljas et al., 1972; Eriksson, Jones, and Liljas, 1988a; Håkansson et al., 1992) and is located in the central part of the protein at the bottom of a cone-shaped crevice.
Multimeric αCAs have been discovered as well as αCAs containing extra domains. One example is the periplasmic αCA, CAH1, from the single-celled green alga, *Chlamydomonas reinhardtii*. Not only is CAH1 a dimer (Kamo et al., 1990; Ishida, Muto, and Miyachi, 1993; Suzuki et al., 2011), the αCA also contains an extra loop structure assumed to be required for enzyme stability (Kamo et al., 1990; Ishida, Muto, and Miyachi, 1993; Suzuki et al., 2011). To generate the extra loop structure of CAH1, its peptide sequence is cleaved resulting in a long and short chain with sizes of 35 kDa and 4 kDa, respectively (Kamo et al., 1990). After cleavage, the short chain is linked to the long chain via disulfide bonds to create the extra loop structure of CAH1 (Ishida, Muto, and Miyachi, 1993; Suzuki et al., 2011). Another example of an αCA dimer is human carbonic anhydrase IX (CAIX). The multimeric form of CAIX was once thought to be stabilized via a disulfide bond formed between Cys41 residues of the two peptide chains (Hilvo et al., 2008), but a CAIX mutant containing a C41S mutation abolishing that disulfide bond still dimerizes via large networks of van der Waals interactions (Alterio et al., 2009).

The β-class CA (βCA) was discovered much later in plants (Burnell, Gibbs, and Mason, 1989; Fawcett, Volokita, and Bartlett, 1990; Roeske and Ogren, 1990) and its protein sequence and structure are very different from that of αCA. The βCA zinc ion is coordinated by two cysteine residues, one histidine residue, and a water molecule in contrast to three histidines in αCAs (Kimber and Pai, 2000). The structure of a βCA monomer is mostly comprised of α-helices that surround a β-sheet consisting of four parallel β-strands. There is also a fifth, C-terminal β-strand involved in the oligomerization of βCA (Kimber and Pai, 2000). The functional unit of the βCA is a dimer, although the most common βCA oligomerization is a tetramer (Kimber and Pai, 2000; Rowlett, 2010). The βCA dimer is formed via extensive
interactions created by two N-terminal α-helices of one monomer wrapping around the second monomer and by minor hydrogen bonding between the second β-strand of each monomer (Kimber and Pai, 2000). Tetramers are formed by interactions made by primarily by the fifth, C-terminal β-strand (Kimber and Pai, 2000). In pea, βCA forms an octamer. Dicots have a unique C-terminal extension of the fifth β-strand of βCA, whereas monocots do not (Kimber and Pai, 2000; Rowlett, 2010). Octamers are formed via slightly different interactions with the fifth, C-terminal β-strand (Kimber and Pai, 2000; Rowlett, 2010).

The γ-class CA (γCA) has been found in plants (Moroney, Bartlett, and Samuelsson, 2001), photosynthetic bacteria (Price et al., 1993), and archaea (Alber and Ferry, 1994). The first crystal structure of γCA from *Methanosarcina thermophila* was reported by Kisker and colleagues in 1996. Much like the active site of αCA, the active site for γCA also contains a zinc atom coordinated by three histidines and a water molecule (Kisker et al., 1996). Although the coordinating amino acids are similar, the protein structures of αCA and γCA are vastly different. αCAs are reported as predominantly monomeric whereas γCAs form homotrimers where the histidines that coordinate the zinc ion are provided by multiple subunits (Kisker et al., 1996). β-strands dominate the structure of γCA and this β-strand region consists of seven complete turns creating a left-handed β-helix (Kisker et al., 1996). Each full turn contains three β-strands making the β-helix look like an equilateral triangle from the top view (Kisker et al., 1996). Two of the three β-sheets, or sides of the β-helix have seven β-strands whereas the third β-sheet has eight β-strands (Kisker et al., 1996). There are also two α-helices at the C-terminus of the γCA monomer (Kisker et al., 1996). The functional unit of γCA is the trimer, as the active site is formed between two of the monomers. Using *Methanosarcina thermophila* numbering, the three histidines required for zinc coordination are His81 and His122 from one monomer and His117
from the neighboring monomer (Kisker et al., 1996). Hydrogen bonding and hydrophobic interactions between β-sheets of each monomer are required to make the trimer (Kisker et al., 1996). Also, three highly conserved residues, Arg59 and Asp61 of one monomer and Asp76 of a second monomer, are required to form salt bridges that also help to stabilize the trimer (Kisker et al., 1996).

**CATALYTIC MECHANISM OF CA**

CA is a very efficient enzyme and the human αCA, CAII, is one of the fastest enzymes characterized, having a turnover rate of $10^6$ sec$^{-1}$ (Khalifah, 1971). The catalytic mechanism of CA was established from previous work on CA protein structure and CA activity assays performed on various site-directed mutants. The majority of the research field focused on the αCA isozymes when studying the CA mechanism. From the collective work, it was determined that all CA isozymes have similar mechanisms only differing in a few details (Lindskog, 1997). The reaction that CA catalyzes, with a pKa of around 6.4, is represented by equation 1.

$$\text{CO}_2 + \text{H}_2\text{O} \leftrightarrow \text{HCO}_3^- + \text{H}^+ \quad (1)$$

More specifically, the CA catalyzed reaction occurs in two steps represented by equations 2 and 3 (Mikulski and Silverman, 2010).

$$\text{CO}_2 + \text{CA-OH}^- \leftrightarrow \text{CA-HCO}_3^- \leftrightarrow \text{CA-H}_2\text{O} + \text{HCO}_3^- \quad (2)$$

$$\text{CA-H}_2\text{O} + \text{B} \leftrightarrow \text{CA-OH}^- + \text{BH}^+ \quad (\text{B = proton acceptor}) \quad (3)$$

It is thought that CO$_2$ interacts with and moves through a hydrophobic pocket (Val121, Val143, Leu198, Val207, and Trp209 of CAII) that opens to the αCA active site (Domsic and McKenna, 2010). The evidence for CO$_2$ requiring this hydrophobic pocket to gain access to the αCA active site is shown when mutating Val143 to Phe or Tyr, as their side groups block passage through the hydrophobic pocket and extinguishes CA activity (Liang and Lipscomb, 1990; Fierke,
Calderone, and Krebs, 1991). Interestingly, when the CO₂ molecule binds the active site of αCA, there is almost no change in the structure of the enzyme, a characteristic of enzymes with high turnover rates (Domsic and McKenna, 2010). The CO₂ molecule is weakly bound to the hydrophobic pocket of the active site and is oriented in a way via interactions between one of the CO₂ oxygens and the amide nitrogen of Thr199 that allows the carbon of CO₂ to easily undergo a nucleophile attack from the zinc-bound hydroxide (Liang and Lipscomb, 1990; Domsic and McKenna, 2010). The location of the CO₂ molecule within the hydrophobic pocket of the αCA active site is supported by multiple studies. The addition of 3-acetoxymercuri-4-aminobenzenesulphonamide inhibits CA activity by blocking the hydrophobic pocket of αCA (Fisher et al., 2007). The inhibitor also displaces a water molecule from Thr199 much like when CO₂ binds the active site of αCA (Fisher et al., 2007). Also, it was shown that site-directed mutants containing the V143Y alteration lack the ability to bind CO₂ (Fierke, Calderone, and Krebs, 1991; Alexander, Nair, and Christianson, 1991). Lastly, with the addition of cyanate, CO₂ is unable to bind to the active site as cyanate interacts with Thr199 (Lindahl, Svensson, and Liljas, 1993).

Once CO₂ is hydrated to HCO₃⁻, the HCO₃⁻ must be released from the active site. Domsic et al. (2008) found that when CO₂ kicks out the water molecule from Thr199, a new water molecule is formed at Thr200. It is thought that this water molecule dislodges the HCO₃⁻ from the active site, as the water molecule is very close to the HCO₃⁻ (Domsic et al., 2008).

The second half of the CA mechanism involves a H⁺ transfer step allowing for the regeneration of the zinc-bound hydroxide. Steiner et al. (1975) postulated that the rate-limiting step of the CA reaction was the proton shuttling involving His64. To regenerate the zinc-bound hydroxide, a zinc-bound water molecule protonates His64 (Mikulski and Silverman, 2010). In
order to protonate His64, the $\text{H}^+$ has to travel along a network of hydrogen bonded water molecules spanning from the $\alpha$CA active site to the His64 residue (Mikulski and Silverman, 2010). The evidence for His64 being the proton shuttling residue comes from an H64A site-directed mutant whose activity was lowered by a factor of 20 but CA activity could be rescued by immersing the H64A mutant in imidazole buffer (Tu et al., 1989).

The CA mechanism is very similar for $\beta$CAs and $\gamma$CAs. Crystal structures of $\beta$CAs and $\gamma$CAs also show hydrophobic pockets (Kimber and Pai, 2000; Rowlett, 2010; Kisker et al., 1996; Iverson et al., 2000). The formation of the hydrophobic pocket of the $\beta$CA occurs when two monomers assemble into the active dimeric $\beta$CA form (Kimber and Pai, 2000; Rowlett, 2010; Domsic and McKenna, 2010). The residues that form the hydrophobic pocket of $\beta$CA from *Pisum sativum* are Val184 from one monomer and Phe179, Tyr205, and Leu210 from the other monomer (Kimber and Pai, 2000; Domsic and McKenna, 2010). The hydrophobic pocket of the $\gamma$CA from *Methanosarcina thermophila* contains the residues Leu83, Phe138, Phe140, Ile157, and Val172 from one monomer and Phe132 and Met135 from the second monomer (Domsic and McKenna, 2010). In both cases, it is thought that the hydrophobic pockets are close enough to the active site for CO$_2$ to be hydrated by the zinc-bound hydroxide (Domsic and McKenna, 2010). Proton shuttling residues have also been identified for $\beta$CAs and $\gamma$CAs. In Arabidopsis, the proton shuttling residue of $\beta$CA is His216 (Rowlett et al., 2002). In *Methanosarcina thermophila*, the proton shuttling residue of the $\gamma$CA is Glu84 (Iverson et al., 2000).

**CA EXPRESSION IN PHOTOSYNTHETIC ORGANISMS**

CA is a highly expressed protein and can account for 1-2% of soluble protein in the leaf (Tobin, 1970; Okabe et al., 1984; Peltier et al., 2005). Determining a single physiological role of CA based on gene expression is difficult due to the lack of agreement among CA expression and
activity data collected from various plants. Although not all the data agree, \( CO_2 \) may affect CA expression levels in plants hinting at possible roles for CA in photosynthesis, lipid biosynthesis, and nucleotide biosynthesis, pathways that all require either \( CO_2 \) or \( HCO_3^- \) for a carboxylation step. Placing bean plants in elevated \( CO_2 \) conditions reduces CA activity (Porter and Grodzinsky, 1984). In contrast, one study found Arabidopsis CA expression to increase in high \( CO_2 \) conditions (Raines et al., 1992). Another study using Arabidopsis plants found CA expression was mostly unaffected by various \( CO_2 \) concentrations, although the expression of \( \alpha CA2 \) and \( \alpha CA3 \) both increased under low \( CO_2 \) conditions (Fabre et al., 2007). A recent study examining the expression of the \( \beta CA \) genes in Arabidopsis found that moving Arabidopsis plants to high \( CO_2 \) resulted in a permanent reduction of \( \beta CA4 \) expression, whereas the expression profiles of the other \( \beta CAs \) in Arabidopsis either had initial drops in expression that eventually returned to normal levels or had expression patterns that oscillated over time (Wang et al., 2014). When Arabidopsis plants were moved to low \( CO_2 \), most of the \( \beta CA \) expression levels increased except for \( \beta CA6 \), whose expression level remained unchanged (Wang et al., 2014). Placing cotton plants in low \( CO_2 \) also resulted in increased CA expression in leaves (Hoang and Chapman, 2002a). Light can also affect CA expression, which might indicate CA has a role in photosynthesis. When maize plants are incubated in the dark, they have lower CA activity than when maize plants are growing in the light (Burnell, Suzuki, and Sugiyama, 1990). Other studies showed that light can stimulate the production of CA in tobacco (Majeau and Coleman, 1992) and rice (Suzuki and Burnell, 1995). Another study showed that dark-adapted plants exposed to light greatly increased CA expression (Majeau and Coleman, 1994). In Arabidopsis, \( \beta CA1 \) was shown to decrease when plants were placed in the dark, whereas there was no change seen in \( \beta CA2 \) levels (Fett and Coleman, 1994). Other studies showed that altering light levels
does not affect CA expression in cotton (Hoang and Chapman, 2002) nor in Arabidopsis (Fabre et al., 2007). Another CA expression study placing Arabidopsis plants in the dark showed CA expression levels dropped for most βCAs except for βCA2 whose mRNA levels remained unchanged and βCA6 whose expression level increased (Wang et al., 2014). It was shown that CA expression can change due to varying nutrients as well. For example, reducing the nitrogen source for plants will lower CA activity (Burnell, Suzuki, and Sugiyama, 1990). Other studies showed CA activity increased when adding either exogenous glutamine or nitrate in combination with zeatin to detached maize leaves from nitrogen-starved plants (Sugiharto et al., 1992; Sugiharto, Burnell, and Sugiyama, 1992). Interestingly, it seems CA from different plants can respond to various environmental changes differently. Also, one can speculate from previous studies that different CA genes in the same organism can have different expression patterns in various environmental changes. Since there are multiple CA genes in a single organism that purportedly have different expression profiles under varying conditions, it is likely that CA can have multiple physiological roles and is not limited to a single role.

**THE ROLE OF CA IN PHOTOSYNTHESIS**

**Photosynthesis - an overview**

Photosynthesis is the process of converting light energy into chemical energy that is used to synthesize organic compounds from CO₂ captured from the atmosphere. Photosynthesis is split into two sets of reactions named the light reactions and the Calvin-Benson-Basham Cycle (CBB Cycle). The light reactions take place on the thylakoid membranes of the chloroplast and serve to replenish the ATP and NADPH pools used in the CBB Cycle. Four multimeric protein complexes, Photosystem II (PSII), Cytochrome b6/f complex (Cyt b6/f), Photosystem I (PSI), and ATP synthase, are embedded in the thylakoid membranes and comprise the core of the light
reactions (Grossman et al., 1995). Light drives the splitting of water molecules in the oxygen evolving complex (OEC) of PSII which releases electrons into PSII and generates O$_2$ as a byproduct. A series of redox reactions involving PSII, Cyt b6/f, PSI, and the electron carriers: plastoquinone, plastocyanin, and ferredoxin, all drive electron transport along the thylakoid membranes to reduce NADP$^+$ produced in the CBB Cycle to NADPH. The electron transport along the thylakoid membranes also results in the pumping of protons into the chloroplast lumen from the chloroplast stroma, creating a proton gradient. The gradient of protons across the thylakoid membrane is used by ATP synthase to generate ATP for the CBB Cycle.

The CBB cycle consists of three steps. The first step is the carboxylation of a five carbon sugar, ribulose 1,5-bisphosphate (RuBP). The enzyme that performs this reaction is ribulose 1,5-bisphosphate carboxylase oxygenase (Rubisco). Rubisco binds its substrate, RuBP, and carboxylates it using CO$_2$ that has diffused into the chloroplast from the atmosphere. From the five carbon substrate, RuBP, and the carboxylating CO$_2$, two molecules of 3-phosphoglycerate (3-PGA) are created (Figure 1.1). The second step of the CBB cycle is the reduction and phosphorylation of 3-PGA (Figure 1.2). The products of Rubisco are subsequently phosphorylated and reduced in the CBB cycle using the ATP and NADPH pools created in the light reactions (Raines, 2003). The triose phosphates produced in the CBB cycle, are then transported out of the chloroplast to be used to make sugars. The last step of the CBB cycle is to regenerate the RuBP pool in the chloroplast. Five triose phosphates (containing a sum of 15 carbons) undergo multiple aldolase and transketolase reactions to form three RuBP molecules that can be carboxylated by Rubisco again (Raines, 2003).

Rubisco is the main carboxylating enzyme of C$_3$ photosynthesis. There are different forms of Rubisco found in nature. Many photosynthetic organisms have Form I Rubisco, a 560
kDa hexadecamer consisting of eight large subunits that are chloroplast encoded and eight small subunits that are nuclear encoded (Knight, Andersson, and Brändén, 1990; Andersson and Backlund, 2008). Dinoflagellates and certain prokaryotes carry the Form II Rubisco which is a dimer of Rubisco large subunits (Schneider et al., 1986; Whitney, Shaw, and Yellowlees, 1995). Archaebacteria have a Rubisco whose structure is a decamer containing dimers of five Rubisco large subunits (Kitano et al., 2001). Although the chloroplast encoded large subunits have not
diverged much in structure or sequence in various organisms, there is variation seen among the nuclear encoded small subunits of various organisms (Spreitzer and Salvucci, 2002). Even with the variations of subunits and multiple forms of Rubisco, the enzyme is still an infamously inefficient enzyme. Rubisco catalyzes 1-3 reactions per active site per second and has a low affinity for CO₂ (Spreitzer and Salvucci, 2002). Also, in place of CO₂, Rubisco can react with O₂ to form products that cannot be used in the CBB cycle. These products, 2-phosphoglycolate (2-PGA), are recycled in the energy costly photorespiratory pathway so three-fourths of the carbon from the 2-PGA pools can be salvaged and used for the CBB cycle (Moroney et al., 2013).

Rubisco is a poor enzyme now, but it may have been a much more efficient enzyme under the conditions it evolved. Rubisco first appeared around three billion years ago when the atmospheric CO₂ concentration was much higher than present day values (Figure 1.3; Berner and Kothavala, 2001; Sage, 2004). As a result of the evolution of oxygenic photosynthesis, atmospheric CO₂ concentrations slowly dropped as competitive atmospheric O₂ concentrations increased over the next hundreds of millions of years (Figure 1.3; Berner and Kothavala, 2001; Sage, 2004). Since the ratio of CO₂:O₂ in the vicinity of Rubisco directly impacts the rate of carboxylation, one can see that the changing environment could have a negative impact on the enzyme’s efficiency. Researchers have tried altering Rubisco to make it a more efficient enzyme in the current atmospheric conditions. In theory, by having a better Rubisco, photosynthesis rates can increase in photosynthetic organisms. An early study by Whitney and Andrews (2001) showed it was possible to successfully replace tobacco’s Rubisco with another, albeit the transformed tobacco plants containing Rhodospirillum rubrum’s Rubisco grew poorly. More
recently, Lin et al., (2014b) generated tobacco transformants with the cyanobacteria Rubisco replacing the native Rubisco. Their plants grew very poorly, but similarly to the transformed tobacco plants of Whitney and Andrews (2001). Their results again show that using a transgenic Rubisco is possible. Another approach is to recombine Rubisco large subunits from one organism with Rubisco small subunits from another organism. Using hybrid Rubiscos in Chlamydomonas combining the Rubisco large subunit of Chlamydomonas with the Rubisco small subunit of tobacco, sunflower, or Arabidopsis resulted in Chlamydomonas mutants that did not grow as well as wildtype Chlamydomonas cells under ambient CO$_2$ conditions as a pyrenoid, a Rubisco-rich region of the algal chloroplast, was not present in these hybrid Rubisco Chlamydomonas mutants (Genkov et al., 2010). Interestingly, these hybrid Rubiscos also had slightly higher CO$_2$ specificity factors but lower carboxylation rates as compared to the wildtype Chlamydomonas Rubisco (Genkov et al., 2010). There is also the possibility of site-directed
mutagenesis to create a better Rubisco. Changing conserved residues at the C-terminus of the Chlamydomonas Rubisco large subunit to mimic the C-terminus of the large subunit of Rubisco from spinach generated a stable, hybrid Rubisco within Chlamydomonas with altered enzyme kinetics (Satagopan and Spreitzer, 2008). This hybrid Rubisco had an increased CO$_2$ specificity factor, although the CO$_2$ specificity factor was lower than the native spinach Rubisco’s specificity factor and had a higher $K_{cat}$ than the native spinach Rubisco, but lower than the Chlamydomonas Rubisco $K_{cat}$ (Satagopan and Spreitzer, 2008). Besides modified Rubiscos, photosynthetic organisms have made other physiological and biochemical changes to compensate for Rubisco’s poor enzymatic activity.

**Photosynthetic adaptations and CA**

Certain photosynthetic organisms have developed their own ways to compensate for the inefficiencies of Rubisco. These organisms, found mostly in less than optimal photosynthetic environments, evolved to have carbon concentrating mechanisms (CCMs) to enhance their photosynthetic properties. All CCMs have three functional characteristics. One, CCMs need inorganic carbon (C$_i$) transporters to move their carbon source. Two, CCMs require a CA to catalyze the interchange of CO$_2$ and HCO$_3^-$ . Lastly, CCMs need Rubisco to be concentrated to a specific region of the chloroplast. Chlamydomonas is a single-celled aquatic green alga that contains a CCM that fulfills each category (Figure 1.4). This green alga evolved to have a CCM because it grows in water where the diffusion of CO$_2$ is 10,000 times slower than it is in air, limiting its CO$_2$ source. Furthermore, in a neutral or alkaline pH environment, the major C$_i$ species available to Chlamydomonas will be in the form of HCO$_3^-$ which has much greater difficulty diffusing through membranes than CO$_2$. In response to these conditions, the CCM of Chlamydomonas has an external CA (CAH1) and putative HCO$_3^-$ transporters (HLA3 and LCI1)
on its plasma membrane to help bring C\textsubscript{i} into the cell (Im and Grossman, 2001; Duanmu et al., 2009; Ohnishi et al., 2010). Once inside the cell, HCO\textsubscript{3}\textsuperscript{-} can enter the Chlamydomonas chloroplast via the chloroplast envelope HCO\textsubscript{3}\textsuperscript{-} transporter, Nar1.2 (Mariscal et al., 2006; Wang and Spalding, 2014). In the chloroplast, the accumulated HCO\textsubscript{3}\textsuperscript{-} needs to be dehydrated to CO\textsubscript{2}. It is thought that HCO\textsubscript{3}\textsuperscript{-} enters the acidic thylakoid lumen, where the lumen CA, CAH3, dehydrates HCO\textsubscript{3}\textsuperscript{-} to CO\textsubscript{2} (Karlsson et al., 1998; Duanmu, Wang, and Spalding, 2009; Moroney

Figure 1.4. A simplified version of the proposed carbon concentrating mechanism model of Chlamydomonas. Whereas CO\textsubscript{2} can freely diffuse from the environment to the site of carboxylation, HCO\textsubscript{3} needs a set of transporters to get to Rubisco. HCO\textsubscript{3} from the environment can be transported across the plasma membrane by HLA3 and LCI1 (Im and Grossman, 2001; Duanmu et al., 2009; Ohnishi et al., 2010). Once in the cytosol, HCO\textsubscript{3} can cross the chloroplast envelope via the HCO\textsubscript{3} transporter, Nar1.2 (Mariscal et al., 2006; Wang and Spalding, 2014). The accumulated HCO\textsubscript{3} in the chloroplast stroma enters the acidic thylakoid lumen via an assumed unknown HCO\textsubscript{3} transporter on the thylakoid membranes and is then dehydrated to CO\textsubscript{2} in the acidic lumen by CAH3 in the pyrenoid (Karlsson et al., 1998; Duanmu, Wang, and Spalding, 2009; Moroney and Ynalvez, 2007). Figure modified from Moroney and Ynalvez (2007). PM = plasma membrane; CE = chloroplast envelope; T = thylakoid lumen
and Ynalvez, 2007). The morphology of the Chlamydomonas chloroplast changes in low CO$_2$ conditions. The pyrenoid becomes very pronounced (Ramazanov et al., 1994) and the thylakoid tubules prominently enter the pyrenoid (Blanco-Rivero, 2012; Engel et al., 2015). To possibly reduce the amount of accumulated C$_i$ leakage from the pyrenoid, nearly 90% of the chloroplast Rubisco content relocates to the pyrenoid under low CO$_2$ conditions to fix the accumulated CO$_2$ within the pyrenoid (Rawat et al., 1995; Borkhsenious, Mason, and Moroney, 1998), a prominent starch sheath forms around the pyrenoid (Ramazanov et al., 1994), and the protein LciB forms a ring around the pyrenoid (Duanmu, Wang, and Spalding, 2009; Wang and Spalding, 2014). The dehydration of accumulated HCO$_3^-$ in the acidic lumen by CAH3 within the pyrenoid will create an increased CO$_2$ concentration in the pyrenoid for Rubisco to fix, increasing the photosynthetic capabilities of Chlamydomonas (Moroney and Ynalvez, 2007).

Cyanobacteria also have an efficient CCM (Figure 1.5). Transporters on the cyanobacterial plasma membrane transport HCO$_3^-$ and CO$_2$ into the cyanobacterial cell (Shibata et al., 2002; Price et al., 2004; Price and Howitt, 2011; Price et al., 2013). The accumulated HCO$_3^-$ in the cytosol then enters the carboxysome, a 20-sided protein shell made up of carboxysome shell proteins, a CA, and Rubisco (Price, Coleman, and Badger, 1992; So and Espie, 2005; Long et al., 2007; Cot, So, and Espie, 2008; Espie and Kimber, 2011). Inside the carboxysome, a CA dehydrates the accumulated HCO$_3^-$ to CO$_2$ in the presence of Rubisco (Price, Coleman, and Badger, 1992).

Another CO$_2$ concentrating mechanism is active in certain plants that have evolved to perform C$_4$ photosynthesis instead of C$_3$ photosynthesis. It is called C$_4$ photosynthesis because the first product of C$_4$ photosynthesis is the four carbon molecule, oxaloacetate (von Caemmerer
Figure 1.5. A simplified model of the carbon concentrating mechanism of cyanobacteria. HCO$_3^-$ transporters on the plasma membrane such as, BicA, SbtA, and BCT1, can transport HCO$_3^-$ into the cytosol (Shibata et al., 2002; Price et al., 2004; Price and Howitt, 2011; Price et al., 2013). There are two NDH-1 complexes on the thylakoid membranes with special subunits that can interact with CO$_2$ and convert it to HCO$_3^-$ by using electrons from NADPH (Price et al., 2002). This accumulated HCO$_3^-$ in the cytosol moves into the carboxysome where CA dehydrates the HCO$_3^-$ to CO$_2$ for Rubisco to fix. Figure modified from Price, Badger, and von Caemmerer (2011). PM = plasma membrane; TM = thylakoid membrane; C = carboxysome and Furbank, 2003; Sage, Sage, and Kocacinar, 2012). C$_4$ photosynthesis is usually compartmentalized between two cells, mesophyll cells and bundle sheath cells, of the leaf (von Caemmerer and Furbank, 2003; Sage, Sage, and Kocacinar, 2012). The oxaloacetate is converted to malate in C$_4$ NADP-malic enzyme type organisms and then moves into the bundle sheath cell chloroplast. Then, the malate is decarboxylated to pyruvate and CO$_2$. This accumulated CO$_2$ in the bundle sheath cell is fixed by Rubisco. The pyruvate moves back to mesophyll cells to be doubly phosphorylated to phosphoenolpyruvate. Although not as effective
in high CO₂ conditions, C₄ photosynthesis gives a distinct advantage over C₃ photosynthesis under lower CO₂ conditions.

**Photosynthesis and CA in C₄ plants**

The role, if any, CA plays in photosynthesis of terrestrial plants is nebulous. In C₄ plants, it is thought that CA activity is required to provide HCO₃⁻ for phosphoenolpyruvate carboxylase (PEPC) in mesophyll cells (Hatch and Burnell, 1990). Modeling CA activity in C₄ plants show that CA is required for optimal rates of C₄ photosynthesis as CA activity is near limiting for optimal photosynthetic requirements/PEPC carboxylation (Hatch and Burnell, 1990). Modeling CA activity also suggests that removal of mesophyll CA will lower C₄ photosynthesis rates since the majority of Cᵢ passing through photosynthetic reactions initially reacted with a CA (Badger and Price, 1994). Early CA antisense work in *Flaveria bidentis* showed that CA activity must be cut by at least 70% in order to see effects (von Caemmerer et al., 2004). This finding suggests that CA activity is not limiting in C₄ photosynthesis although if CA activity is reduced to a low enough level, reductions in photosynthesis rates can be observed (von Caemmerer et al., 2004). A more recent study that removed ca1 and ca2 of maize showed poor growth of the CA double mutant, but no CA effect on photosynthesis rates, suggesting a possible anaplerotic role for CA such as providing Cᵢ for intermediate reactions in other carbon-requiring pathways that are independent of photosynthesis (Studer et al., 2014).

**The role of CA in C₃ photosynthesis**

The effect of CA activity on C₃ photosynthesis is also unclear. CAs in the chloroplast are thought to facilitate CO₂ diffusion into the chloroplast, since lowering the chloroplast CA activity in tobacco plants reduced the CO₂ partial pressure in the tobacco chloroplasts, although without lowering photosynthesis rates (Price et al., 1994). It is thought that cytosolic CA activity
plays a minimal role in C₃ photosynthesis (Badger and Price, 1994; Terashima et al., 2011; Tholen et al., 2012; Tholen, Ethier, and Genty, 2014). Studies using antisense lines show reducing chloroplast CA levels below 10% of total CA activity in tobacco did not significantly reduce photosynthesis rates (Price et al., 1994; Williams et al., 1996; Majeau, Arnoldo, and Coleman, 1994) although CA antisense studies showed reduced water use efficiency (WUE) and increased stomatal conductance (Majeau, Arnoldo, and Coleman 1994; Kim, 1997). However, these studies were conducted before the numbers of CA genes in plants were known. It is still possible that other CA isoforms could be compensating for the loss of the targeted CA.

The effect of location on CA activity and photosynthesis

The location of CA activity within a photosynthesizing cell is important to the proper functioning of a CCM. It is believed that ectopic CA activity can disrupt the function of CCMs by hindering an organism’s CCM to concentrate Cᵢ around Rubisco. In cyanobacteria, CA activity is restricted to the carboxysome (Price and Badger, 1989). Cyanobacteria producing cytosolic human CAII grow very poorly in low CO₂ conditions (Price and Badger, 1989). The cytosolic human CAII in the cyanobacteria dehydrated the accumulated HCO₃⁻ to CO₂ before it could reach the carboxysome holding the Rubisco of the cell (Price and Badger, 1989). This CO₂ diffused out of the cell which caused the cyanobacteria to grow poorly in low Cᵢ conditions.

In C₄ plants, CA activity is found in the cytosol of mesophyll cells to provide HCO₃⁻ to PEPC to make oxaloacetate (Hatch and Burnell, 1990). Tobacco chloroplast CA was placed in the C₄ plant Flaveria bidentis (Flaveria) and increased the ¹³C discrimination of the plant (Ludwig et al., 1998). The increase in ¹³C discrimination suggests that the tobacco chloroplast CA in Flaveria increased the amount of Cᵢ leaking from the cell which lowers the C₄ CCM efficiency (Ludwig et al., 1998).
THE ROLE OF CA IN LIPID BIOSYNTHESIS

Lipid biosynthesis overview

Fatty acid biosynthesis is a part of primary metabolism in the plant. It occurs in the plastids of plant cells and its pathway is composed of repeated steps to build a fatty acid chain consisting of 16 to 18 carbons. Acetyl-CoA carboxylase (ACCase) is a multi-subunit protein that catalyzes a two-step, ATP-powered reaction that generates malonyl-CoA from acetyl-CoA and HCO$_3^-$ (Sasaki and Nagano, 2004). The four subunits that form ACCase are biotin carboxylase (BC), biotin carboxyl carrier protein (BCCP), and the α- and β-carboxyltransferase (CT) subunits (Sasaki and Nagano, 2004). In plants there are usually two forms of ACCase present, a homomeric ACCase made from a single, large polypeptide containing all four subunits and a heteromeric ACCase that is constructed from four, separately coded, protein subunits (Alban, Baldet, and Douce, 1994; Konishi and Sasaki, 1994; Konishi et al., 1996; Schulte et al., 1997; Ke et al., 2000). The homomeric form comes from a eukaryotic origin and is mostly found in the cytosol of plants (Alban, Baldet, and Douce, 1994; Konishi and Sasaki, 1994; Konishi et al., 1996; Sasaki and Nagano, 2004). The heteromeric form is of prokaryotic origins and the various nuclear-encoded ACCase subunits are all targeted to the chloroplast (Alban, Baldet, and Douce, 1994; Konishi and Sasaki, 1994; Konishi et al., 1996; Sasaki and Nagano, 2004). Not all plants have both eukaryotic and prokaryotic ACCase forms, as the Graminaeae family lacks the prokaryotic form (Konishi et al., 1996). The first step of the ACCase catalyzed reaction involves BC carboxylating biotin carried by BCCP seen in equation 3 (Sasaki and Nagano, 2004). The second step of the reaction is catalyzed by CT which transfers the carboxyl group from the carboxylated biotin carried by BCCP to acetyl-CoA as seen in equation 4 (Sasaki and Nagano, 2004).
BCCP-biotin + HCO$_3^-$ + ATP $\rightarrow$ BCCP-biotin-CO$_2$ + ADP + P$_i$ \hspace{1cm} (3)

BCCP-biotin-CO$_2$ + acetyl-CoA $\rightarrow$ BCCP-biotin + malonyl-CoA \hspace{1cm} (4)

Once malonyl-CoA is made, a series of repeated steps are executed to generate 16C and 18C fatty acid chains (Figure 1.6).

![Diagram of fatty acid synthesis pathway]

Figure 1.6. The fatty acid synthesis pathway. Modified from Ohlrogge and Browse (1995).

**CA and Lipid Biosynthesis**

It is possible that a CA is required to deliver HCO$_3^-$ to ACCase to carboxylate biotin. A few experiments have shown that CA activity can impact lipid biosynthesis rates in plants. Cottonseed embryos incubated with the CA inhibitor, ethoxyzolamide, had decreased amounts of radiolabeled acetate incorporated into its lipid pools (Hoang and Chapman, 2002a). A second experiment using tobacco cell suspensions incubated in ethoxyzolamide also showed reduced
levels of radiolabeled acetate in lipid samples (Hoang and Chapman, 2002a). A third experiment showed that tobacco CA-antisense lines had reduced amounts of radiolabeled acetate incorporated into chloroplast lipid pools (Hoang and Chapman, 2002a). A study using CA RNAi tobacco lines showed that reducing chloroplast CA can lower the rate of C$_i$ entering the chloroplast (Price et al., 1994). Hoang and Chapman (2002) considered the possibility that CA activity traps C$_i$ within compartments in the form of HCO$_3^-$.

Whether it is a matter of facilitating CO$_2$ diffusion into the chloroplast, trapping C$_i$ as HCO$_3^-$, or a combination of the two, the lower C$_i$ levels in the chloroplast due to lower levels of chloroplast CA can possibly affect lipid synthesis rates as the $K_m$ (HCO$_3^-$) of ACCase is already higher than normal plastid HCO$_3^-$ levels.

**CA AND AMINO ACID BIOSYNTHESIS**

It is possible that cytosolic CA activity is required to maintain a steady pool of C$_i$ for amino acid biosynthesis (Raven and Newman, 1994). A major carboxylation reaction in amino acid biosynthesis of plants includes PEPC. In C$_4$ plants, PEPC is mostly known for catalyzing the first step of photosynthesis, but for C$_3$ plants PEPC is associated with amino acid biosynthesis. PEPC uses cytosolic HCO$_3^-$ to carboxylate phosphoenolpyruvate (PEP) to oxaloacetate (OAA). OAA then forms aspartate when aspartate aminotransferase removes an amino group from glutamate forming α-ketoglutarate and adds the amino group to OAA forming aspartate (Figure 1.7). It is estimated that 50 percent of the free aspartate pool in the plant is created by PEPC (Melzer and O’Leary, 1987). C$_4$ Cytosolic PEPC has a $K_m$ (HCO$_3^-$) between 25 and 100 μM (O’Leary, 1982; Bauwe, 1986; Hatch and Burnell, 1990) and the C$_3$ cytosolic PEPC has an even higher $K_m$ (HCO$_3^-$) between 100 and 200 μM (Mukerji and Yang, 1974; Sato, Koizumi, and Yamada, 1988). Since the pKa of CO$_2$ is 6.4 and assuming the cytosol of a C$_3$ cell has a pH of 7.1 and the CO$_2$ concentration of the cytosol is 12 μM, the cytosolic HCO$_3^-$
concentration is near 60 μM. This cytosolic HCO$_3^-$ concentration is much lower than the $K_m$ (HCO$_3^-$) of PEPC. It is possible that CA activity is responsible for trapping C$_i$ in the form of HCO$_3^-$ so C$_i$ pools used for carbon-requiring reactions do not leak out of subcellular compartments or the plant cell (Hoang and Chapman, 2002). If this is the case, CA activity in the cytosol may be required to maintain cytosolic HCO$_3^-$ levels to maintain PEPC activity.

Figure 1.7. The carboxylation reaction of phosphoenolpyruvate carboxylase and downstream aminotransferase reaction to generate aspartate. Figure modified from Heldt and Piechulla (2011).

CA AND NUCLEOTIDE BIOSYNTHESIS

It was suggested that CA activity may play an anaplerotic role in plastid nucleotide biosynthesis (Raven and Newman, 1994; Hoang and Chapman, 2002). There are two plastid nucleotide biosynthesis enzymes that require C$_i$ to carboxylate their substrate. The first enzyme is carbamoyl-phosphate synthetase (CPS) which catalyzes the first step of pyrimidine synthesis and the second enzyme is 5-Amino-imidazole ribonucleotide (AIR) carboxylase involved in
purine biosynthesis (Zrenner et al., 2006). It is assumed that CPS activity does not match the C\textsubscript{i} assimilation levels of PEPC and ACCase (Raven and Newman, 1994), so these reactions may not be affected as much by the presence or absence of CA activity in the plastid, although if their affinities for C\textsubscript{i} are low, they may still be impacted by a loss of substrate in the chloroplast.

**CA IN ARABIDOPSIS**

Arabidopsis is a C\textsubscript{3} dicot and contains eight \(\alpha\)CAs, six \(\beta\)CAs, three \(\gamma\)CAs, and two \(\gamma\)-like CAs (Table 1.1). Although the \(\beta\)CAs and \(\gamma\)CAs have been studied in Arabidopsis, the \(\alpha\)CAs have remained relatively unanalyzed. Of the eight \(\alpha\)CAs, only \(\alpha\)CA1, \(\alpha\)CA2, and \(\alpha\)CA3 are expressed at a moderate level (Fabre et al., 2007; Ferreira, Guo, and Coleman, 2008; Schmid et al., 2005; Winter et al., 2007). \(\alpha\)CA4 through \(\alpha\)CA8 are either pseudogenes genes or expressed at an undetectable level or in very specific cell types. \(\alpha\)CA1 is the only Arabidopsis \(\alpha\)CA that has been studied. Although the function of \(\alpha\)CA1 has not been ascertained, a unique mechanism for the localization of this protein has been reported. \(\alpha\)CA1 has a leader sequence that designates the protein to be directed to the endoplasmic reticulum (ER) in the secretory pathway (Villarejo et al., 2005). Although it was shown that \(\alpha\)CA1 is directed to the ER, surprisingly \(\alpha\)CA1 is not secreted, nor is it located in the plasma membrane (Villarejo et al., 2005; Burén et al., 2011). GFP studies show \(\alpha\)CA1 is located in the chloroplast (Villarejo et al., 2005). Later it was determined that \(\alpha\)CA1 is glycosylated in the ER which redirects the protein to the chloroplast (Burén et al., 2011). Site-directed mutagenesis targeting glycosylation sites on \(\alpha\)CA1 as well as adding glycosylation inhibitors prevented glycosylation of \(\alpha\)CA1 to occur as seen by changes in \(\alpha\)CA1’s molecular weight on Western blots when treated (Burén et al., 2011) as well as changes in the subcellular location of \(\alpha\)CA1 when treated (Villarejo et al., 2005).
Table 1.1. List of Arabidopsis carbonic anhydrase genes. There are eight αCA genes, six βCA genes, three γCA genes, and two γCA-like genes in Arabidopsis.

<table>
<thead>
<tr>
<th>CA</th>
<th>GENE I.D.</th>
<th>SUBCELLULAR LOCATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>αCA1</td>
<td>At3g52720</td>
<td>Chloroplast&lt;sup&gt;1,2&lt;/sup&gt;</td>
</tr>
<tr>
<td>αCA2</td>
<td>At2g28210</td>
<td>Secreted&lt;sup&gt;3,4,5,6,7&lt;/sup&gt;</td>
</tr>
<tr>
<td>αCA3</td>
<td>At5g04180</td>
<td>Secreted&lt;sup&gt;3,4,5,6,7&lt;/sup&gt;</td>
</tr>
<tr>
<td>αCA4</td>
<td>At4g20990</td>
<td>Unknown</td>
</tr>
<tr>
<td>αCA5</td>
<td>At1g08065</td>
<td>Unknown</td>
</tr>
<tr>
<td>αCA6</td>
<td>At4g21000</td>
<td>Unknown</td>
</tr>
<tr>
<td>αCA7</td>
<td>At1g08080</td>
<td>Unknown</td>
</tr>
<tr>
<td>αCA8</td>
<td>At5g56330</td>
<td>Unknown</td>
</tr>
<tr>
<td>βCA1</td>
<td>At3g01500</td>
<td>Chloroplast&lt;sup&gt;8,9&lt;/sup&gt;</td>
</tr>
<tr>
<td>βCA2</td>
<td>At5g14740</td>
<td>Cytosol&lt;sup&gt;8&lt;/sup&gt;</td>
</tr>
<tr>
<td>βCA3</td>
<td>At1g23730</td>
<td>Cytosol&lt;sup&gt;8&lt;/sup&gt;</td>
</tr>
<tr>
<td>βCA4</td>
<td>At1g70410</td>
<td>Plasma Membrane&lt;sup&gt;8,10&lt;/sup&gt;</td>
</tr>
<tr>
<td>βCA5</td>
<td>At4g35580</td>
<td>Chloroplast&lt;sup&gt;8&lt;/sup&gt;</td>
</tr>
<tr>
<td>βCA6</td>
<td>At1g58180</td>
<td>Mitochondria&lt;sup&gt;8,11&lt;/sup&gt;</td>
</tr>
<tr>
<td>γCA1</td>
<td>At1g19580</td>
<td>Mitochondria&lt;sup&gt;12&lt;/sup&gt;</td>
</tr>
<tr>
<td>γCA2</td>
<td>At1g47260</td>
<td>Mitochondria&lt;sup&gt;12,13&lt;/sup&gt;</td>
</tr>
<tr>
<td>γCA3</td>
<td>At5g66510</td>
<td>Mitochondria&lt;sup&gt;12,14&lt;/sup&gt;</td>
</tr>
<tr>
<td>γCAL1</td>
<td>At5g63510</td>
<td>Mitochondria&lt;sup&gt;12&lt;/sup&gt;</td>
</tr>
<tr>
<td>γCAL2</td>
<td>At3g48680</td>
<td>Mitochondria&lt;sup&gt;12&lt;/sup&gt;</td>
</tr>
</tbody>
</table>


βCAs in Arabidopsis are much more highly expressed than the αCAs. βCA1 and βCA5 are both located in the chloroplast (Fabre et al., 2007; Hu et al., 2015), but βCA1 is the highest expressed CA in Arabidopsis leaves whereas βCA5 is only expressed at a level that is 5% of βCA1 (Fabre et al., 2007; Ferreira, Guo, and Coleman, 2008; Schmid et al., 2005; Winter et al., 2007). βCA2 and βCA3 are both cytosolic CAs (Fabre et al., 2007) and βCA2 is the second highest expressed CA in Arabidopsis leaves whereas βCA3 is not expressed at a high level (Fabre et al., 2007; Ferreira, Guo, and Coleman, 2008; Schmid et al., 2005; Winter et al., 2007). βCA4 is located on the plasma membrane and although moderately expressed in Arabidopsis...
leaves, it is the highest expressed CA in roots (Fabre et al., 2007; Wang et al., 2014). It is thought that the active site of βCA4 is on the cytosolic side of the plasma membrane because it was found that βCA4 interacts with the plasma membrane aquaporin, PIP2;1 (Wang et al., 2016). βCA6 is a mitochondrial CA and is expressed at a moderate level in Arabidopsis leaves (Fabre et al., 2007; Jiang et al., 2014).

The three βCAs that have been most extensively researched are βCA1, βCA4, and βCA6. Removing βCA1 from Arabidopsis plants results in no discernible growth phenotype in various CO₂ concentrations (Ferreira, Guo, and Coleman, 2008). Yet seedlings lacking βCA1 are reported to have lower survival rates after germination as well as having a lower CO₂ uptake rate as measured by acid-stable ¹⁴C incorporation (Ferreira, Guo, and Coleman, 2008). It was found that βCA1, βCA4, and βCA6 of Arabidopsis are the three highest expressed CAs in guard cells and have been implicated in controlling guard cell movements via sensing CO₂ concentrations (Hu et al., 2010; Xue et al., 2011; Hu et al., 2015). It is thought that the HCO₃⁻ hydrated from CO₂ via guard cell CAs is a signal mediated by OST1, a protein kinase, to activate SLAC1 dependent S-type anion channels to close stomata (Xue et al., 2011). Studies using double or triple knockout lines lacking βCA1 and βCA4 or βCA1, βCA4, and βCA6 show these mutant lines have a slower stomatal response to changing CO₂ concentrations (Hu et al., 2010; Hu et al., 2015). Arabidopsis plants that lack both βCA1 and βCA4 have a different stomatal density than wildtype plants (Hu et al., 2010; Engineer et al., 2014). Stomatal densities of the CA mutant plants were lower than wildtype plants in low CO₂ environments and had higher stomatal densities than wildtype plants in ambient and higher CO₂ environments (Hu et al., 2010; Engineer et al., 2014).
Plants missing βCA6 show reduced growth in a low CO₂ environment, although plants lacking βCA6 still have normal photosynthesis rates (Jiang et al., 2014). Alternatively, overexpressing βCA6 in Arabidopsis resulted in larger plants with higher reported dry-weights (Jiang et al., 2014). βCA6 was assumed to affect cellular respiration of Arabidopsis because of its mitochondrial location and because βCA6 levels were shown to affect the growth of Arabidopsis and change its respiration rates but not its photosynthesis rates (Jiang et al., 2014).

γCAs and γCA-like (γCAL) proteins interact with Complex I of plant mitochondria (Parisi et al., 2004; Perales et al., 2004). Interestingly, this γCA and γCAL complex that produces an extra structure on Complex I is only seen in photosynthetic organisms (Heazlewood et al., 2003; Perales et al., 2004; Sunderhaus et al., 2006). Unfortunately, enzyme activity has yet to be measured from the γCA and γCAL proteins (Parisi et al., 2004; Perales et al., 2004; Martin et al., 2009). Certain conserved amino acids are missing from the Arabidopsis γCA and γCAL proteins that are needed for catalysis (Parisi et al., 2004; Perales et al., 2004). The three γCAs in Arabidopsis all have the zinc coordinating residues as well as most catalytic amino acid residues, but the γCAs lack the important proton shuttling residues, Glu62 and Glu84 (Parisi et al., 2004). The γCAL proteins lack two of the three zinc coordinating histidines (Perales et al., 2004). Although the γCAs do not show activity, it is possible that the enzymes require an unknown factor for activity supplied only in vivo in the mitochondria. Arabidopsis γCA2 expressed and purified from E. coli showed the ability to bind C_i, although CA activity could not be detected in HCO₃⁻ dehydration assays (Martin et al., 2009). Because γCA2 shows the ability to bind C_i, the idea that γCAs in Arabidopsis could possibly help shuttle C_i from the mitochondria to chloroplasts was put forward (Zabaleta, Martin, and Braun, 2012). Another hypothesis is that these γCAs could affect photorespiration in the mitochondria as the expression
of γCA and γCAL genes were downregulated in elevated CO₂ (Perales et al., 2005; Zabaleta, Martin, and Braun, 2012; Soto et al., 2015). Removing both γCA2 and γCAL2 in Arabidopsis led to increased production of glycine and reduced photosynthesis rates (Soto et al., 2015). Also, it is possible that γCAs affect cellular respiration since plants lacking γCA2 have severely reduced levels of complex I (Zabaleta, Martin, and Braun, 2012; Perales et al., 2005). Interestingly, even though complex I is reduced by 80% in the γca2 knockout, γca2 mutant plants show no discernible phenotype in the conditions tested, although cell cultures lacking γca2 were reported to have lower respiration rates (Perales et al., 2005). Although γCA single mutants do not show a discernible growth phenotype, γca1γca2 double mutant plants cannot be obtained because γca1γca2 mutants are lethal at the embryo stage (Córdoba et al., 2015). Other studies reported sterile Arabidopsis plants ubiquitously overexpressing γCA2 (Villarreal et al., 2009). Sterility in the Arabidopsis plants is thought to be due to an inability of their anthers to release pollen grains because the abnormal expression of γCA2 led to a reduction in reactive oxygen species (ROS) burst needed for anther development (Villarreal et al., 2009).

**THESIS PROJECT**

The purpose of this project was to further characterize the αCA and βCA isoforms of Arabidopsis to add information to the CA collective work as well as to determine the necessity of various plant CA isoforms for application-based experiments to improve photosynthesis. One goal of this project was to determine CA expression levels and expression patterns in tissues of Arabidopsis plants. These experiments confirmed the CA expression levels determined by previous microarray experiments (Schmid et al., 2005; Winter et al., 2007). Surprisingly, the RNAseq experiment which provided the CA expression data also showed that there were multiple mRNA forms for various CA genes, which corroborates the findings of Aubry et al.
Promoter::GUS studies were performed to confirm the findings of the RNAseq experiment as well as determining the CA expression patterns in Arabidopsis tissues.

Another major goal of this project was to confirm the subcellular locations of various CA isoforms in Arabidopsis. I paid particular attention to the genes that produced multiple mRNA forms as they possibly coded for proteins that have different subcellular locations. Indeed, the \( \beta CA4 \) gene produced two proteins, \( \beta CA4.1 \) which localized to the plasma membrane and \( \beta CA4.2 \) which localized to the cytosol. The localization data were of particular importance as they dictated which CA single mutant lines to cross to generate double mutants.

One hypothesis of this work was that the CA isoforms in Arabidopsis have overlapping functions and can compensate for a missing CA. Previous studies with \( \beta \)CAs in maize and \( \gamma \)CAs in Arabidopsis provide evidence that CAs have overlapping functions in plants (Studer et al., 2014; Wang et al., 2012; Soto et al., 2015; Córdoba et al., 2015; Fromm et al., 2016). Indeed, the removal of one CA isoform from the Arabidopsis plant usually led to no discernible growth phenotype. Using the subcellular localization data, Arabidopsis single mutants lacking CA isoforms in the same subcellular region were crossed to produce double mutants. The double mutants created were \( \beta ca2\beta ca4 \), lacking CAs in the cytosol, and \( \alpha ca1\beta ca4 \) and \( \alpha ca2\beta ca4 \), lines lacking secreted and plasma membrane CAs. Making a chloroplast CA double mutant line by crossing the \( \beta ca1 \) and \( \beta ca5 \) lines was not necessary as \( \beta ca5 \) plants already had extremely reduced growth. These CA mutant lines were crucial in setting the groundwork for investigating the main research topic.

The main hypothesis of this work is that Arabidopsis CA isoforms facilitate CO\(_2\) diffusion through mesophyll cells and build C\(_1\) pools as HCO\(_3^-\) to maintain carbon-requiring reactions (Figure 1.8). T-DNA insertion lines were used to identify roles for CA in various
carbon-requiring reactions. eGFP plants were generated to analyze the subcellular locations of CA isoforms in Arabidopsis and to determine which CA T-DNA lines to cross to generate double mutant lines. GUS plants were generated to confirm earlier RNAsseq work and to determine the tissue expression patterns of various CA genes in Arabidopsis. Interestingly, the

![Diagram](image1.png)

Figure 1.8. A proposed model of C\textsubscript{i} accumulation and CO\textsubscript{2} diffusion based on carbonic anhydrase activity derived from the βCAs of Arabidopsis. Since the pKa of CO\textsubscript{2} is 6.4, CA in the plasma membrane and neutral pH cytosol can hydrate CO\textsubscript{2} diffusing across the plasma membrane and into the cytosol to HCO\textsubscript{3}\textsuperscript{-}. This may keep the CO\textsubscript{2} diffusion gradient moving into the cytosol from the intercellular air space (IAS) and effectively traps the carbon source in the cytosol as the C\textsubscript{i} form, HCO\textsubscript{3}\textsuperscript{-}, which can be used for carbon-requiring reactions such as amino acid synthesis and the creation of secondary metabolites. CO\textsubscript{2} can diffuse directly into the chloroplast from the IAS as well as from the cytosol. CO\textsubscript{2} entering the chloroplast can be hydrated to HCO\textsubscript{3}\textsuperscript{-} because of the basic pH of the stroma. This also keeps the CO\textsubscript{2} diffusion gradient moving into the chloroplast and traps a large C\textsubscript{i} source in the stroma in the form of HCO\textsubscript{3}\textsuperscript{-}. The diffusing CO\textsubscript{2} can be used by Rubisco to make 3-phosphoglycerate (3-PGA) which feeds into the CBB cycle. Two prominent, carbon-requiring pathways that can utilize the large pool of HCO\textsubscript{3}\textsuperscript{-} in the stroma include fatty acid synthesis and nucleotide synthesis. PM = plasma membrane, Thy = thylakoid lumen.
cytoplasmic CA double mutant, βca2βca4, shows a CO₂ dependent growth phenotype, yet the βca2βca4 plants show no reduction in photosynthesis. Analysis of the βca2βca4 free amino pools show that aspartate and glutamine are reduced while glycine and serine are elevated compared to wildtype amino acid levels. The chloroplast βca5 mutant has a developmental growth phenotype that is partially corrected when grown in high CO₂ conditions. Double knockout lines, αca1βca4 and αca2βca4, have slightly reduced growth in a low CO₂ environment. The results of this dissertation reveal that CAs may impact multiple carbon-requiring reactions and that the CA network throughout the Arabidopsis plant may be more complex than once thought.
CHAPTER 2
MATERIALS AND METHODS

GROWTH CONDITIONS

All Arabidopsis wildtype and T-DNA plants used in this work are of the Columbia (COL) ecotype. For all transient expression assays, *Nicotiana benthamiana* plants were used. Arabidopsis plants were grown in ambient CO$_2$ at 21° C in 24 hour light at an intensity of 80 μmol photons m$^{-2}$ sec$^{-1}$ unless otherwise specified. All plants used in growth studies were watered bi-weekly, alternating between distilled H$_2$O and a 1:3 dilution of Hoagland’s nutrient solution in distilled H$_2$O (Epstein and Bloom, 2005).

DNA EXTRACTION AND GENOMIC PCR

50 mg of leaf tissue was ground in a sterile 1.5 mL microcentrifuge tube before adding 400 μL of Edward’s Extraction Buffer (200 mM Tris-HCl pH 7.5, 250 mM NaCl, 25 mM EDTA pH 8.0, 0.5% SDS). The samples were vortexed and then spun in a benchtop centrifuge at 14,000 rpm for five minutes. 300 μL of the supernatant was added to a fresh 1.5 mL microcentrifuge tube whereas the pellet was discarded. 100% isopropanol was added to make a final 1:1 dilution of sample and isopropanol. The samples were vortexed briefly and then spun down at 14,000 rpm for ten minutes. The supernatant was discarded before 400 μL of 70% ethanol was added to each DNA pellet. Samples were vortexed briefly before being spun down at 14,000 rpm for five minutes. The supernatant was discarded and DNA pellets were allowed to air dry for at least one hour before being resuspended in 100 μL of 1x TE (10 mM Tris-HCl pH 8.0, 1mM EDTA pH 8.0). DNA samples were quantified using the NanoDrop 1000 Spectrophotometer (Thermo Scientific). 100 ng of DNA was used for a 50 μL PCR reaction. Amplification of DNA was carried out using the standard protocol for One Taq (New England Biolabs).
**RNA EXTRACTION AND RT-PCR**

80 mg of leaf tissue from Arabidopsis plants in the vegetative stage was collected in a sterile, 1.5 mL microcentrifuge tube and frozen in liquid N\(_2\) before being crushed with a small, sterile pestle. RNA was extracted from crushed leaf samples using the Qiagen RNAeasy kit. RNA samples were quantified using the NanoDrop 1000 Spectrophotometer (Thermo Scientific). 3 \(\mu\)g of RNA was used for the reverse transcriptase reaction to generate cDNA. The reverse transcriptase reaction was performed using the SuperScript\textsuperscript{®} First-Strand RT-PCR kit and protocol (Invitrogen). 0.5 \(\mu\)L of cDNA (2 \(\mu\)g \(\mu\)L\(^{-1}\)) was used as the template for a 25 \(\mu\)L PCR reaction using the standard protocol for One Taq (New England BioLabs).

**RNAseq ANALYSIS**

RNAseq reads were generated and processed by Dr. Maheshi Dassanayake and Dr. Dong-Ha Oh to calculate expression counts. The methods performed are described in Oh et al. (2014). Three biological replicates were used to generate an average count for each gene.

**PROTEIN EXTRACTION AND WESTERN BLOTTING**

50 mg of leaf tissue was placed in a sterile 1.5 mL microcentrifuge tube and ground using a sterile plastic pestle. 132 \(\mu\)L of Protein Extraction Buffer (1x TE, 1.2 \%SDS, 2.7\% sucrose, 7.5 \(\mu\)g mL\(^{-1}\) bromophenol blue) was added to the ground leaf tissue. Samples were vortexed and placed in ice for 15 minutes before being centrifuged at 14,000 rpm for five minutes using a benchtop centrifuge. The supernatant was collected and placed in a new sterile 0.5 mL microcentrifuge tube whereas the pellet was discarded. Protein samples were quantified using the Bradford protein assay kit (Pierce). 2-mercaptoethanol was added to a final concentration of 350 mM and then samples were incubated at 100\(^\circ\) C for three minutes. 5 \(\mu\)g of total protein was added to a 12\% polyacrylamide gel and proteins were electrophoresed (Bio-Rad) before being
transferred to a PVDF membrane (Bio-Rad) using the Semi-Dry system. Membranes were blocked in 5% powdered milk dissolved in 1x TBS (50 mM Tris-HCl pH 7.5, 150 mM NaCl) for up to two hours at room temperature before being washed three times with TTBS (0.05% Tween-20, 19 mM Tris base, 500 mM NaCl, pH 7.5). The PVDF membrane was then placed in a TBSB solution (1% BSA, 19 mM Tris base, 500 mM NaCl, pH 7.5) containing either a spinach CA polyclonal antibody (Fawcett et al., 1990) at a final dilution of 1:20,000 or a polyclonal antibody directed against Arabidopsis βCA1 (Proteintech Group) at a final dilution of 1:20,000 and was allowed to incubate overnight at 4°C. The following morning, the membrane was washed five times with TTBS before being placed in a TBSB solution containing a Bio-Rad goat anti-rabbit secondary antibody at a final dilution of 1:20,000. The membrane was allowed to incubate at room temperature for one hour and was then washed five times with TTBS. Membranes were incubated in a 1:1 mixture of peroxide and luminol (Bio-Rad) for at least 30 seconds. Protein bands on the membrane were visualized on X-ray film using chemiluminescence.

VECTOR CONSTRUCTION USING THE GATEWAY SYSTEM

Amplicons for pENTR™ Gateway construction were generated using Phusion polymerase (New England Biolabs). PCR fragments were gel purified using the Qiaquick Gel Extraction kit (Qiagen). 1 to 2 μL of purified PCR product was added to a pENTR™ master mix (1 μL of a 1.2 M NaCl and 0.06 M MgCl₂ mix (Invitrogen), 1 μL pENTR™/dTOPO® vector mix (Invitrogen), and dH₂O (Invitrogen) to final volume of 6 μL) for pENTR™ vector construction. Vectors were transformed into E. coli TOP10 chemically competent cells and plated onto YEP plates (for 1 L: 10 g peptone, 5 g NaCl, 10 g yeast extract, 15 g agar) supplemented with 50 μg mL⁻¹ kanamycin. pENTR™ vectors were subjected to restriction digestion and sequencing to confirm the correct orientation and sequence of the construct. pENTR™ vectors containing
coding regions or genomic regions built for eGFP were recombined into the pDEST™ vector, pB7FWG2 (Karimi, Inzé, and Depicker, 2002). pENTR™ vectors containing coding regions or genomic regions constructed for complementation were recombined into the pDEST™ vector, pMDC32 (Curtis and Grossniklaus, 2003). Promoter regions of CA genes held in pENTR™ vectors were recombined into the pDEST™ vector, pKGWFS7 (Karimi, Inzé, and Depicker, 2002). The correct orientation of the pDEST™ vector was confirmed via restriction digestion.

**TRANSFORMATION OF AGROBACTERIUM TUMEFACIENS**

Agrobacterium were transformed following the Freeze-thaw protocol from Weigel and Glazebrook (2002). 10 μL of an *E. coli* plasmid miniprep containing a pDEST™ vector was added to 250 μL of thawed chemically competent Agrobacterium cells (strain GV3101 or GV3101::pMP90RK). Samples were then placed in ice for five minutes. To re-freeze Agrobacteria after adding 10 μL of plasmid prep DNA, samples were placed in liquid N₂ for five minutes. Samples were then placed in 37⁰ C for five minutes. The transformed bacteria samples were pipetted into a 1.5 mL microcentrifuge tube containing 1 mL of YEP media (for 1 L: 10 g peptone, 5 g NaCl, 10 g yeast extract). Samples were placed on a shaker at were allowed to incubate at room temperature for four hours before being plated on YEP plates supplemented with appropriate antibiotics. Colonies that formed on the selection plates were genotyped via PCR.

**TRANSFORMATION OF ARABIDOPSIS PLANTS**

Arabidopsis plants were transformed using the floral dip method of Weigel and Glazebrook (2002). Agrobacterium cells were pipetted into flasks containing 200 mL of YEP medium supplemented with the appropriate antibiotics. The flasks were then placed on shakers and the Agrobacterium were allowed to grow at 28⁰ C overnight. The following morning,
Agrobacterium cells were pelleted by centrifugation at 6,000 rpm for 10 minutes using the Beckman J2-HS centrifuge and JA-10 rotor. Agrobacterium cells were resuspended in 400 mL of infiltration media (half-strength Murashige and Skoog Salts, half-strength Gamborg’s B5 Vitamins, 5% (w/v) sucrose, 0.044 μM benzylamino purine, 50 μL L⁻¹ Silwet L-77). Plant flowers and inflorescences were dipped in the Agrobacterium infiltration medium for 40 seconds before being placed on their sides on a tray. The trays were covered with a plastic dome and placed in 21 °C with constant light for one day before removing the dome and returning the plants to their upright position. Plants were kept alive to allow siliques to form. Seeds were collected once the plant tissue dried from lack of watering.

**TRANSIENT TRANSFORMATION OF TOBACCO LEAVES FOR eGFP EXPRESSION**

Four to five week old *Nicotiana benthamiana* (tobacco) plant leaves were used for transient eGFP expression. Two Agrobacterium strains were used to generate transient eGFP expression in tobacco leaves. When infiltrating tobacco leaves, one strain, GV3101 containing the αCA1-eGFP construct, was combined with a second Agrobacterium strain, AGL-1 [p19], containing a suppressor to gene silencing construct (Voinnet et al., 2003). Two days before infiltrating tobacco leaves with an Agrobacterium solution, inoculate a 2 mL YEP liquid culture with either a single Agrobacterium colony from an agar plate or from a glycerol stock and were placed in a 28 °C shaker overnight. The GV3101 culture should also contain 100 μg mL⁻¹ Spectinomycin (MP Biomedicals, LLC), 30 μg mL⁻¹ Gentamycin (GOLDBIO), and 10 μg mL⁻¹ Rifampicin (GOLDBIO). The AGL-1 [p19] culture should also contain 50 μg mL⁻¹ Kanamycin (Sigma). The following afternoon, 100 μL of the GV3101 and AGL-1 [p19] cultures were transferred to 5 mL of YEP solution containing the appropriate antibiotics and were placed in a 28 °C shaker overnight. The following morning, the OD₆₀₀ was measured. An OD₆₀₀ of 2.0 was
the target cell concentration as an OD$_{600}$ below 2.0 is not concentrated enough and an OD$_{600}$ well above 2.0 will have cell death. If an OD$_{600}$ of 2.0 was not reached by morning, cultures were allowed to incubate at 28°C until an OD$_{600}$ of 2.0 was reached. If the OD$_{600}$ of the overnight culture was well above 2.0, a new 5 mL culture was inoculated with the overnight culture to give the new culture an OD$_{600}$ of 1.0. The new culture was allowed to incubate in 28°C for at least two hours to allow the OD$_{600}$ reach 2.0. Once an OD$_{600}$ of 2.0 was reached, 0.5 mL of the GV3101 culture and 0.5 mL of the AGL-1 [p19] culture were combined in a 2 mL microcentrifuge tube. The sample was spun at 3,000 RPM for 10 minutes and the supernatant was discarded. The Agrobacterium pellet was resuspended in 1 mL of 10 mM MgCl$_2$ to remove the antibiotics. The Agrobacterium solution was spun again at 3,000 RPM for 10 minutes and the supernatant was discarded. The Agrobacterium pellet was resuspended in 2 mL of an Agrobacterium resuspension solution (1 mL of 100 mM MES; 1 mL of 100 mM MgCl$_2$; 100 μL of 1.5 Acetosyringone (Sigma); 7.9 mL dH$_2$O). Using a needle-less syringe, the 2 mL of Agrobacterium solution was infiltrated into the abaxial side of multiple tobacco leaves. Infiltrated tobacco plants were returned to normal growth conditions and were imaged by confocal microscopy three days later.

**PROTOPLAST GENERATION FROM ARABIDOPSIS LEAVES**

Protoplasts from stably transformed Arabidopsis leaves were generated for better resolution with confocal microscopy. The protocol of Wu et al., (2009) was used to make protoplasts from two g of leaf tissue. Magic Tape was added to the upper epidermis of leaves whereas Time Tape was added to the lower epidermis of the leaves. To remove the lower epidermis, the Time Tape was peeled from the leaf, taking the lower epidermal layer with it. Leaves with exposed mesophyll layers were incubated in 10 mL of enzyme solution (1%
Cellulase from Trichoderma viride (Sigma), 0.25% Pectinase from Rhizopus sp. (Sigma), 0.4 M mannitol, 10 mM CaCl₂, 20 mM KCl, 0.1% BSA, 20 mM MES at pH 5.7) for one hour in light. Protoplasts were then concentrated by spinning the samples at 800 rpm using a Beckman J2-HS centrifuge and JS-13.1 rotor. The supernatant was discarded and the protoplasts were resuspended in a solution containing 0.4 M mannitol, 15 mM MgCl₂, and 4 mM MES at pH 5.7 and were then ready to image.

**eGFP VISUALIZATION USING ARABIDOPSIS LEAVES AND PROTOPLASTS**

At least three independent, stably transformed eGFP Arabidopsis plants or at least three leaves from three independent transiently transfected tobacco plants were used for eGFP localization of a specific CA. On average, a 30 mm² tissue sample from a leaf was submerged in perfluorodecalin (Littlejohn and Love, 2012; Sigma) for up to five minutes in a microscope well slide before a 1.5 mm cover slip was added and the tissue imaged. To image immobilized protoplasts, 30 μL of a protoplast sample was added to a microscope slide and a 1.5 mm cover slip was carefully added so as not to burst the protoplasts. To excite eGFP and chlorophyll, the Kr/Ar laser for the Leica Sp3 confocal microscope was set to 488 nm and the photomultiplier tube was set to detect wavelengths of 505-520 nm to observe eGFP fluorescence and a second photomultiplier tube was set to detect wavelengths of 660-700 nm to observe chlorophyll autofluorescence.

**PROMOTER::GUS ANALYSIS**

GUS staining was visualized following a modified protocol of Jefferson, Kavanagh, and Bevan (1987). At least five stably transformed promoter::GUS Arabidopsis plants were used for GUS expression analysis. Plants were grown in ambient air at 21°C in 24 hour light with a light intensity of 80 μmol photons m⁻² sec⁻¹. Cotyledons three days post germination, three-week old
rosettes, and floral tissues grown in either soil or on ½ MS plates were used for various GUS experiments. In the afternoon, plants were submerged in a GUS staining solution (0.1M NaPO₄ pH 7, 10 mM EDTA, 0.1% Triton X-100, 1 mM K₃Fe(CN)₆, 2 mM 5-bromo, 4-chloro, 3-indol-β-D-glucuronic acid (X-Gluc, from GoldBio) suspended in N,N-Dimethylformamide) in either small scintillation vials or 50 mL beakers, depending upon the size of the plants, and were subjected to vacuum infiltration for five minutes. After vacuum infiltration, plants were placed in a 37°C incubator in the dark overnight. The following morning, plants were taken out of the incubator and the GUS staining solution was aspirated from the beakers or vials. To remove chlorophyll and other plant pigments, 100% methanol was added to the GUS plants after the staining solution was aspirated and then the plants were placed in a 60°C oven for 15 minutes. This step was repeated until all of the pigments were removed from the wildtype, negative control plants. GUS plants were imaged using a Canon EOS Rebel T5i camera and Canon EF 100 mm F/2.8L IS USM macro lens.

**GROWTH STUDIES OF βca2βca4 MUTANTS**

Rosette areas were obtained weekly by photographing plants growing in either 200 μL L⁻¹ CO₂ or 1,000 μL L⁻¹ CO₂ with an eight hour photoperiod at a light intensity of 120 μmol photons m⁻² sec⁻¹ using a Canon Rebel T5i camera with a Canon 15-85mm f/3.5-5.6 IS USM lens. Pictures of plants were analyzed using ImageJ (NIH) by tracing the outline of the projected plant rosette to obtain areas. Rosette areas of at least eight plants were measured per line.

Dry-weights of each plant line growing in 200 μL L⁻¹ CO₂ with an eight hour photoperiod at a light intensity of 120 μmol photons m⁻² sec⁻¹ were measured once a week. Five plants per line were used for each dry-weight measurement. Arabidopsis rosettes were clipped at the crown
of each plant and were placed in 60° C. Dry weight measurements were recorded once a day until plant weights stabilized.

**STOMATA COUNTING**

Detached leaf segments of 10-week old Arabidopsis plants grown in 200 μL L⁻¹ CO₂ and short days were submerged in a 1 mg/mL solution of propidium iodide (Biotium) for five minutes before being imaged with a Leica SP3 confocal microscope. The Kr/Ar laser was set to 488 nm to excite propidium iodide and the photomultiplier tube was set to absorb wavelengths of 620-650 nm. Images from fluorescing epidermal leaf fragments taken by the confocal microscope were analyzed for stomata counts using ImageJ (NIH).

**AMINO ACID ANALYSIS**

When plants reached six weeks of age growing in either 200 μL L⁻¹ or 1,000 μL L⁻¹ CO₂, three plants from each line were chosen for amino acid analysis. 100 mg of leaf tissue from each plant was frozen in liquid N₂ and was ground in a sterile 1.5 mL microcentrifuge tube using a small plastic pestle. 500 μL of 80% methanol was added to the ground leaf tissue and then each sample was vortexed and placed in a 75° C hot water bath for 10 minutes. The samples were then placed in a Beckman centrifuge JA-18.1 rotor and were centrifuged at 20,000 g for five minutes at 4° Celsius. The supernatant was placed in a new sterile 1.5 mL microcentrifuge tube while the pellet was resuspended using 20% methanol and was centrifuged again at 20,000 g for five minutes at 4° Celsius. The two supernatants for each plant sample were pooled and were then pushed through a 0.2 μm filter (VWR International) and collected in a new sterile 2 mL microcentrifuge tube. The leaf samples were quantified using the protocol from Lowry et al. (1951) and were then shipped to the TAMU Protein Chemistry Laboratory at Texas A&M University for amino acid analysis.
GAS EXCHANGE ANALYSIS

Leaf gas exchange rates were measured with a LI-COR 6400XT gas analyzer system. Plants used for gas exchange measurements were grown at 200 μL L⁻¹ CO₂ for 10 weeks and then were shifted to 1,000 μL L⁻¹ CO₂ for two weeks in order to obtain leaves big enough from the βca2βca4 line to fill the LI-COR 6400-40 leaf fluorescence cuvette. Leaves were first allowed to acclimate in the leaf cuvette at 400 μL L⁻¹ CO₂, 1,000 μmol photons m⁻² s⁻¹ (saturating irradiance for these leaves), and 22-24°C for one hour or until a steady-state rate of CO₂ uptake was reached. A/Cᵢ curves were measured on the 16th youngest leaf from four separate plants of each plant line. Each curve started at 400 μL L⁻¹ CO₂ and decreased to 50 μL L⁻¹ CO₂ before returning to 400 μL L⁻¹ CO₂ and subsequently increasing to 2,200 μL L⁻¹ CO₂. For each CO₂ point, individual leaves reached steady state photosynthesis within three minutes on average before measurements were recorded.

GROWTH STUDIES OF βca5 MUTANTS

10-day old βca5 homozygous mutants, 10-day old wildtype (COL) plants, and three-day old wildtype (COL) cotyledons growing in ambient CO₂ and a 24 hour photoperiod with a light intensity of 80 μmol photons m⁻² sec⁻¹ were photographed using a Canon Rebel T5i camera and Canon EF 100 mm F/2.8L IS USM macro lens.

Plants growing in either ambient air and a 24 hour photoperiod with a light intensity of 80 μmol photons m⁻² sec⁻¹ or 50,000 μL L⁻¹ CO₂ and a 24 hour photoperiod with a light intensity of 80 μmol photons m⁻² sec⁻¹ were photographed with a Canon Rebel T5i camera with a Canon 15-85mm f/3.5-5.6 IS USM lens. Pictures of plants were analyzed using ImageJ (NIH) by tracing the outline of the projected plant rosette to obtain areas. The rosette area was measured on at least four plants from each line.
GROWTH STUDIES OF αCA MUTANTS

Wildtype (COL) and αCA mutant plants were grown in low CO₂ (200 μL L⁻¹ CO₂) in an 8 hour light/16 hour dark photoperiod with a light intensity of 120 μmol photons m⁻² sec⁻¹ for a 10 week growth study. Plant rosettes were photographed using a Canon Rebel T5i camera and Canon 15-85mm f/3.5-5.6 IS USM lens. Rosette images were analyzed using ImageJ (NIH) to obtain projected rosette areas of each plant line.
CHAPTER 3
CYTOSOLIC CARBONIC ANHYDRASE IS REQUIRED FOR OPTIMAL GROWTH IN LOW CO₂

INTRODUCTION

The roles of cytosolic CAs vary in different types of photosynthetic organisms. It is hypothesized that cytosolic CAs are important in the functioning of the CCM in C₄ plants (Burnell and Hatch, 1990; Badger and Price, 1994). The majority of the carbon fixed in C₄ plants is expected to pass through a CA catalyzed reaction, as PEPC uses HCO₃⁻ which is generated through the action of CA on CO₂ diffusing into the plant (Badger and Price, 1994). Models have been generated showing the pools of HCO₃⁻ formed in the cytosol of C₄ plants are just above limiting for PEPC, suggesting CA activity is required to maintain optimal rates of C₄ photosynthesis (Burnell and Hatch, 1990). CA-antisense constructs which reduce the activity of cytosolic CA in Flaveria bidentis mesophyll cells by at least 70% leads to diminished rates of photosynthesis, albeit CA levels must be severely reduced in order to see effects on photosynthesis rates due to the high enzymatic activity of CA (von Caemmerer et al., 2004). In maize, insertional mutants of the ca1 and ca2 genes decreased plant growth but led to no significant changes in photosynthesis rates, suggesting possible anaplerotic roles for CA (Studer et al., 2014).

The roles of CAs in C₃ photosynthesis are poorly understood. CAs in the cytosol and chloroplast have been proposed to help facilitate the diffusion of inorganic carbon to the chloroplast, however recent modeling studies indicate the effect of CA activity in the cytoplasm

might be minimal (Badger and Price, 1994; Terashima et al., 2011; Tholen et al., 2012; Tholen, Ethier, and Genty, 2014). Earlier studies using antisense lines show reducing chloroplast CA levels below 10% of total CA activity in tobacco did not significantly reduce photosynthesis rates (Price et al., 1994; Williams et al., 1996; Majeau, Arnoldo, and Coleman, 1994). Other tobacco CA antisense lines have shown reduced water use efficiency (WUE) and increased stomatal conductance (Majeau, Arnoldo, and Coleman 1994; Kim, 1997). All of these studies were conducted before it was known that there were multiple CA genes so it is still possible that other CA isoforms could be compensating for the loss of the targeted CA. Therefore it is unclear which CAs, if any, were contributing to CO$_2$ conductance or fixation in C$_3$ plants.

Based on previous reports and our preliminary studies, I hypothesize that there are multiple forms of CA in different cell compartments and these CAs may have overlapping functions. In this chapter, I report our investigation of the physiological roles of cytoplasmic βCAs. I have found that βCA2 and a previously unknown short form of βCA4 (βCA4.2) are the most abundant cytoplasmic CAs in Arabidopsis leaves. Using a transgenic plant missing βCA2, βCA4.1, and βCA4.2, I found that these cytosolic CAs are required for optimal growth under low CO$_2$ conditions. I put forth the hypothesis that optimal cytosolic PEPC activity requires CA activity.

RESULTS

Expression of the cytosolic CAs

There are three cytosolic CAs in Arabidopsis: βCA2, βCA3, and βCA4. To determine if these three CAs are expressed in leaves, RNA from three biological replicates was extracted from both root and leaf tissue of Arabidopsis and was used for RNAseq analysis (RNAseq data is deposited in the NCBI SRA database as BioSample:SAMN03339724). Root and leaf RNAseq
reads were refined as normalized counts of 100 bp reads that uniquely mapped to a CA gene. Analyzing the normalized RNAseq counts show that all three βCA genes are expressed in both roots and leaves and that the overall cytosolic CA expression pattern agrees with the deposited EST counts found on TAIR (Table 3.1). Examining the RNAseq data shows that βCA2 is the most highly expressed cytosolic CA in the leaf and βCA4 is the most highly expressed CA in the root (Table 3.1). βCA3 is expressed at a very low level, corresponding to only 1% of βCA2 and 5% of βCA4 expression (Table 3.1; Ferreira, Guo, and Coleman, 2008; Schmid et al., 2005; Winter et al., 2007).

Table 3.1. βCA2, βCA3, and βCA4 are all expressed in roots and leaves. Three Arabidopsis plants were used for leaf and root RNA samples. RNAseq reads were normalized to 100 bp counts that uniquely mapped to a CA gene in the Arabidopsis reference genome. RNAseq values are given in RPKM (Reads Per Kilo base per Million mapped). Actin1 was used as a reference to compare CA expression levels.

<table>
<thead>
<tr>
<th>GENE</th>
<th>GENE I.D.</th>
<th>ESTs(^a)</th>
<th>RNAseq Reads(^b) (Shoots)</th>
<th>RNAseq Reads(^c) (Roots)</th>
</tr>
</thead>
<tbody>
<tr>
<td>βCA2</td>
<td>At5g14740</td>
<td>693</td>
<td>34,488</td>
<td>214</td>
</tr>
<tr>
<td>βCA3</td>
<td>At1g23730</td>
<td>51</td>
<td>376</td>
<td>21</td>
</tr>
<tr>
<td>βCA4</td>
<td>At1g70410</td>
<td>279</td>
<td>7,781</td>
<td>9,026</td>
</tr>
<tr>
<td>Actin1</td>
<td>At2g37620</td>
<td>65</td>
<td>1,471</td>
<td>1,056</td>
</tr>
</tbody>
</table>

\(^a\)EST values were extracted from The Arabidopsis Information Resource webpage (Arabidopsis.org)
\(^b\),\(^c\)RNAseq data generated by Oh et al. (2014)

**RNAseq data show alternative forms of βCA2 and βCA4**

The RNAseq data did not match any of the βCA2 gene models from TAIR. The uniquely mapped RNAseq reads from leaf and root samples were aligned to the genomic region of βCA2 in the Arabidopsis reference genome (TAIR10) and visualized using a genome browser (Figure 3.1). Both root and leaf RNAseq reads report a shorter βCA2 transcript as the transcription start site begins in exon 2 of the TAIR10 model (Figure 3.1). This result impacts the predicted subcellular localization of the βCA2 protein. The short form of βCA2 will lose its predicted chloroplast transit peptide and will become a predicted cytosolic protein (Figure 3.2 and Table
which agrees with the findings of Fabre et al. (2007). Due to the predicted chloroplast location and lack of expression of the long form of βCA2, named βCA2.1, the long βCA2 form was not included in this chapter of work. In this chapter, the term, βCA2, will be used in exclusive reference to the short βCA2 form, named βCA2.2.

Figure 3.1. Aligning RNAseq reads from root and shoot RNA samples to the Arabidopsis reference genome. RNAseq reads begin in the second predicted exon of the βCA2 reference gene. This shifts the βCA2 ATG start downstream, causing the loss of the predicted chloroplast transit peptide of βCA2. RNAseq data generated by Oh et al. (2014).

Figure 3.2. Aligning the peptide sequence of the chloroplast carbonic anhydrase, βCA1, with the two βCA2 peptide sequences, βCA2.1 and βCA2.2 using Clustal Omega. βCA1 and βCA2.1 both have putative chloroplast transit peptides indicated in red.
Table 3.2. Predicted subcellular locations of βCA1, βCA2.1, and βCA2.2 using protein subcellular location prediction programs. Predotar, ChloroP, TargetP, and YLoc are four programs that predict the subcellular location of proteins based on amino acid sequence. A value of (1) denotes strong confidence in the prediction. Ch = Chloroplast, O = Other location, Cy = Cytoplasm; (+) = contains a chloroplast transit peptide, (-) does not contain a chloroplast transit peptide.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Predotar</th>
<th>ChloroP</th>
<th>TargetP</th>
<th>YLoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>βCA1</td>
<td>Ch (0.89) (+)</td>
<td>Ch (0.97)</td>
<td>Ch (1)</td>
<td></td>
</tr>
<tr>
<td>βCA2.1</td>
<td>Ch (0.52) (+)</td>
<td>O (0.669)</td>
<td>Ch (0.99)</td>
<td></td>
</tr>
<tr>
<td>βCA2.2</td>
<td>O (0.97) (-)</td>
<td>O (0.903)</td>
<td>Cy (0.98)</td>
<td></td>
</tr>
</tbody>
</table>


The RNAseq data also show two distinct transcripts of βCA4 when the unique reads generated by RNAseq were aligned to the genomic region of βCA4 (Figure 3.3). This finding is in agreement with the earlier findings of Aubry et al. (2014) and Wang et al. (2014). In addition, each transcript seems to be expressed in a tissue specific manner. The longer transcript of βCA4, named βCA4.1, is expressed mainly in leaves and has two 5’ exons unique to its transcript. The shorter transcript, named βCA4.2, is expressed in both roots and leaves and has one 5’ exon unique to its transcript.

Figure 3.3. Leaf and root RNA samples yielded two forms of βCA4 mRNA. The long mRNA form is found primarily in the leaf and contains 10 exons where the first two exons are unique to the long form. The short mRNA form has nine exons where the first exon is unique to the short mRNA form and can be found in both the root and shoot RNA samples. RNAseq data generated by Oh et al. (2014).
Promoter::GUS analysis of βCA2, βCA3 and βCA4

To confirm the RNAseq expression patterns of βCA2, βCA3, and βCA4, various promoter regions upstream of the βCA2, βCA3, and βCA4 ATG start sites (Figure 3.4) were inserted upstream of the GUS reporter gene in the vector, pKGWFS7 (Karimi, Inzé, and Depicker, 2002), and stably introduced into wildtype Arabidopsis plants by transformation. The positive control

![Promoter::GUS analysis of βCA2, βCA3 and βCA4](image)

Figure 3.4. Maps of promoter regions used to drive GUS expression. Promoter regions used to drive GUS expression from (A) βCA2, (B) βCA3, and (C) βCA4 GUS constructs are marked with black bars under each gene model. Blue boxes and blue lines represent exons and introns, respectively. Red boxes and arrows represent the 5’ and 3’ UTRs, respectively. Purple lines represent intergenic regions. Lightly shaded boxes and lines represent alternative gene model forms for the multiple mRNA forms for the gene. Arrows indicate orientation of the genes.
line, pβCA1::GUS, was generated by stably transforming a GUS construct containing 805 bp of the βCA1 promoter region into wildtype Arabidopsis plants. Three week old Arabidopsis plants were used for GUS analysis. Plants containing the pβCA1::GUS construct showed strong GUS expression in rosettes (Figure 3.5), similar to the results reported by Wang et al. (2014). A

![Figure 3.5. GUS staining of the various pβCA2::GUS, pβCA3::GUS, and pβCA4::GUS plants. pβCA2::GUS and pβCA4.1::GUS plants show GUS staining in the rosette leaves, whereas pβCA3::GUS and pβCA4.2::GUS plants primarily show GUS staining in the roots. pβCA4.1::GUS plants also have strong GUS staining in the roots. pβCA4.1v.2::GUS and pβCA4.2v.2::GUS plants do not show GUS staining. As a negative control, wildtype (COL) plants were used as a positive control for their strong GUS activity in the rosette leaves.](image-url)
similar pattern was seen with the plants containing the \( p\beta CA2::GUS \) construct (Figure 3.5). \( p\beta CA3::GUS \) plants showed minimal GUS staining around the edges of rosette leaves and staining in the roots (Figure 3.5). Analyzing expression for \( \beta CA4 \) was more difficult because of the multiple transcription start sites. Multiple promoter regions were used to elucidate the expression patterns of the two \( \beta CA4 \) transcripts (Figure 3.4, C). Plants containing the construct \( p\beta CA4.1::GUS \) showed strong GUS expression in both roots and leaves, whereas the construct \( p\beta CA4.2::GUS \) led to strong GUS expression primarily in the roots of plants (Figure 3.5). The constructs \( p\beta CA4.1v.2::GUS \) and \( p\beta CA4.2v.2::GUS \) led to no GUS expression in plants (Figure 3.5). The results from the \( p\beta CA2::GUS \) and \( p\beta CA4.1::GUS \) plant lines confirm the RNAseq data that \( \beta CA2 \) and \( \beta CA4 \) are expressed in leaves. Furthermore, the promoter::GUS work illustrates the differential expression of the two forms of \( \beta CA4 \), which agrees well with the RNAseq work, as well as depicts the complexity of the genomic region regulating the two \( \beta CA4 \) transcript forms.

**Subcellular localization of \( \beta CA2 \), \( \beta CA4.1 \), and \( \beta CA4.2 \)**

The transcription start site of \( \beta CA4.2 \) is in a genomic region of \( \beta CA4 \) that excludes the predicted leader sequence that directs \( \beta CA4.1 \) through the secretory pathway to the plasma membrane (Figure 3.6 and Table 3.3). Continuing the work of Fabre et al. (2007), Hu et al. (2010), and Hu et al. (2015), a C-terminal eGFP tag was added to \( \beta CA4.2 \) to determine if the subcellular location of \( \beta CA4.2 \) has changed. Arabidopsis plants stably expressing either \( \beta CA2\)-eGFP or \( \beta CA4.1\)-eGFP were generated as controls for cytoplasmic and plasma membrane signals, respectively. Leaves of Arabidopsis plants stably expressing \( \beta CA2\)-eGFP confirmed the cytosolic localization (Figure 3.6; Fabre et al., 2007). Leaves of plants stably expressing \( \beta CA4.1\)- eGFP gave a plasma membrane signal, confirming earlier reports by Fabre et al. (2007),
Figure 3.6. Clustal Omega alignment of the peptide sequences of βCA4.1 and βCA4.2. Marked in red is the secretory pathway leader sequence of βCA4.1, whereas βCA4.2 lacks the secretory pathway leader sequence.

Table 3.3. Predicted locations of βCA4.1 and βCA4.2 using four different protein targeting programs. Predotar, ChloroP, TargetP, and YLoc are four programs that predict the subcellular location of proteins based on amino acid sequence. A value of (1) denotes strong confidence in the prediction. Ch = Chloroplast, O = Other location, Cy = Cytoplasm; (+) = contains a chloroplast transit peptide, (-) does not contain a chloroplast transit peptide.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Predotar</th>
<th>ChloroP</th>
<th>TargetP</th>
<th>YLoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>βCA4.1</td>
<td>O (0.95)</td>
<td>(-)</td>
<td>O (0.79)</td>
<td>Ch (0.99)</td>
</tr>
<tr>
<td>βCA4.2</td>
<td>O (0.99)</td>
<td>(-)</td>
<td>O (0.99)</td>
<td>Cy (0.73)</td>
</tr>
</tbody>
</table>

Small et al. (2004); Emanuelsson et al. (2000); Nielsen H, and Von Heijne G (1999); Briesemeister, Rahnenführer, and Kohlbacher (2010a); Briesemeister, Rahnenführer, and Kohlbacher (2010b).

Hu et al. (2010), and Hu et al. (2015), whereas βCA4.2-eGFP gave a cytosolic signal (Figure 3.7).

To increase the resolution of the localization data, protoplasts were made using leaves that were stably expressing the different eGFP constructs. The cytosolic localization for βCA2 and βCA4.2 was confirmed using protoplasts from the βCA2-eGFP and βCA4.2-eGFP plant lines (Figure 3.7). The protoplasts made from βCA4.1-eGFP plants generated a thin ring of GFP fluorescence, confirming the plasma membrane localization for βCA4.1 (Figure 3.8).
Figure 3.7. eGFP fluorescence from leaf tissue of various stably expressing eGFP lines. βCA2-eGFP and βCA4.2-eGFP samples give cytosolic fluorescence signals whereas βCA4.1-eGFP samples give a plasma membrane fluorescent signal. eGFP fluorescence is represented as yellow and chlorophyll autofluorescence is represented as blue.
Figure 3.8. eGFP fluorescence of protoplasts generated from stably expressing eGFP Arabidopsis lines. The βCA2-eGFP and βCA4.2-eGFP protoplasts confirm the cytosol subcellular localization and the βCA4.1-eGFP protoplasts confirm the plasma membrane localization. eGFP is represented as yellow and chlorophyll autofluorescence is represented as blue.
**Genotyping the βca2 and βca4 T-DNA lines**

Fabre et al. (2007) determined that the cytosolic CAs in Arabidopsis are βCA2 and βCA3. Previous reports and our own RNAseq data show that βCA3 is expressed at only 1% of βCA2 (Table 3.1; Ferreira, Guo, and Coleman, 2008; Schmid et al., 2005; Winter et al., 2007). Also, promoter::GUS experiments show that βCA3 is not strongly expressed in the rosette leaves of Arabidopsis (Figure 3.5), thus βCA3 was not considered for this project. Wang et al. (2016) suggests that βCA4.1 is attached to cytoplasmic side of the plasma membrane via interactions with aquaporins. The eGFP data show that βCA4.2 is located in the cytoplasm. To determine the effects of βCA2 and both forms of βCA4 on plant growth, T-DNA insert lines for both genes were obtained from TAIR. The βca2 line, SALK_145785, has a T-DNA insertion in the fifth intron of the βCA2 gene, whereas the βca4 line, CS859392, contains a T-DNA insertion in the fourth intron of the βCA4 gene (Figure 3.9). Genomic PCR using the gene-specific primers

![Figure 3.9](image-url)  
**Figure 3.9.** The location of T-DNA insertions in the βCA2 and βCA4 genes of the βca2 and βca4 mutant lines, respectively. The βca2 T-DNA insertion (SALK_145785) is located in the fourth intron of the βCA2 gene. The βca4 T-DNA insertion (CS859392) is located in the fourth intron of the βCA4 gene. T-DNA insertions are represented by triangles. Primers used for PCR and reverse transcriptase PCR are indicated by black arrows. F and R represent gene specific primers that span the location of the insert and β2-I and β4-I represent T-DNA specific primers. Blue boxes and blue lines represent exons and introns, respectively. Red boxes and red arrows represent the 5’ and 3’ UTRs, respectively.
showed that the T-DNA insertions disrupted the genomic regions of the $\beta CA2$ and $\beta CA4$ genes (Figure 3.10, A). Primers specific to sequences unique to the T-DNA inserts were combined with gene-specific primers to confirm the location of the inserts within the two genes (Figure 3.10, A). RT-PCR was performed to confirm the T-DNA mutants were knock-out lines. Transcripts for the $\beta CA2$ and $\beta CA4$ genes are absent in the $\beta ca2$ and $\beta ca4$ lines, respectively, whereas both transcripts are present in the wildtype plants (Figure 3.10, B). Transcripts for both genes are missing in the $\beta ca2\beta ca4$ double mutant (Figure 3.10, B).

![Figure 3.10](image.png)

**Figure 3.10.** Genotyping of the $\beta ca2$, $\beta ca4$, and $\beta ca2\beta ca4$ mutants. (A) Genomic PCR probing for the presence of the T-DNA insert in the $\beta CA2$ and $\beta CA4$ genes. Amplicons for the respective T-DNA insertions can be seen in the $\beta ca2$, $\beta ca4$, and $\beta ca2\beta ca4$ lines. (B) RT-PCR illustrating the lack of $\beta CA2$ mRNA in the $\beta ca2$ and $\beta ca2\beta ca4$ lines and lack of $\beta CA4$ mRNA in the $\beta ca4$ and $\beta ca2\beta ca4$ lines.
An antibody raised against spinach chloroplast CA (Fawcett et al., 1990) can detect both βCA1 and βCA2 in Arabidopsis because their peptide sequences are nearly 90% identical (Figure 3.2 and Figure 3.11, A). The mature βCA1 protein, after losing its chloroplast transit peptide, is 233 amino acids long resulting in a protein having a molecular weight of around 25.3 kDa (Fawcett et al., 1990; Fett and Coleman 1994). βCA2 has no transit peptide, so its 259 amino acids gives it a molecular weight of around 28.4 kDa. Using the spinach CA antibody, two bands can be detected around 25 kDa and 28 kDa (Figure 3.11, A). The band at 25 kDa is indicative of the βCA1 protein and shows that βCA1 is present in all four lines. The band around 28 kDa represents the βCA2 protein and that band is missing in both the βca2 single mutant and the βca2βca4 double mutant (Figure 3.11, B), indicating that the βCA2 protein is missing in these lines.

![Figure 3.11. Two Western blots using an antibody raised against spinach carbonic anhydrase. (A) A Western blot demonstrating the anti-CA antibody recognizes both the βCA1 and βCA2 proteins. (B) Western blot demonstrating that the βCA2 protein is absent in the βca2 and βca2βca4 mutants.](image)
Growth of plants lacking cytoplasmic CAs

βca2βca4 plants grown in 200 μL L\(^{-1}\) CO\(_2\) (low CO\(_2\)) in an eight hour photoperiod with a light intensity of 120 μmol of photons m\(^{-2}\) sec\(^{-1}\) grew more slowly compared to wildtype plants and the βca2 and βca4 lines (Figure 3.12, A). βca2βca4 plants also showed chlorosis in their youngest leaves when growing in low CO\(_2\). The growth of βca2βca4 plants improved in 1,000 μL L\(^{-1}\) CO\(_2\) (high CO\(_2\)) in an eight hour photoperiod (Figure 3.12, B). Chlorosis in βca2βca4 plants remained when grown in high CO\(_2\) (Figure 3.12, B). The measured rosette areas and dry-weights of the wildtype, βca2, and βca4 plants were similar, but the rosette areas and dry-weights of the βca2βca4 plants were significantly reduced (Figure 3.13, A and B). The growth differences between the βca2βca4 plants and the wildtype and single mutant plants became apparent after the second or third week of growth (Figure 3.13, inserts).

Figure 3.12. The growth phenotype of the βca2βca4 double mutant. The βca2βca4 mutant displays slow growth in (A) low CO\(_2\) conditions (200 μL L\(^{-1}\) CO\(_2\)) but grows much more like wildtype plants in (B) high CO\(_2\) conditions (1,000 μL L\(^{-1}\) CO\(_2\)).
Figure 3.13. The βca2βca4 double mutant shows reduced growth in low CO₂ over a 10 week study at 200 μL L⁻¹ CO₂. (A) The average dry weights and (B) average projected rosette areas of the βca2βca4 double mutant plants are already reduced by the second to third week of the growth study (inserts). βca2 and βca4 single mutants grow similarly to wildtype (COL) plants in low CO₂. Each dry weight symbol represents the mean ± SD of five independent plants. Each symbol for projected rosette area represents the mean ± SD of nine independent plants.

**Photosynthetic properties of the double mutant**

Photosynthetic properties of 10-week old wildtype, βca2, βca4, and βca2βca4 plants grown at low CO₂ in an eight hour photoperiod with a light intensity of 120 μmol of photons m⁻² sec⁻¹ were measured to determine if a reduction in carbon fixation of the double mutants led to their slow growth in low CO₂. The CO₂ compensation points of double mutant plants were similar to those of the single mutant plants and wildtype plants (Table 3.4). The CO₂ assimilation rate (A) under various CO₂ concentrations was similar among the βca2βca4 plants, wildtype plants, and both single mutants (Table 3.4).

**Stomate density of the double mutant**

Removing βCA1 and βCA4, two CAs expressed highly in guard cells, changes the stomate density of Arabidopsis leaves (Hu et al., 2010; Engineer et al., 2014). Double mutant plants were checked for compensatory changes in stomate density that could account for the lack
Table 3.4. Gas exchange values of wildtype (COL), βca2, βca4, and βca2βca4 plants. The CO₂ compensation points were generated by finding the slope of the initial linear portion of the A/Cᵢ curve and solving for the X-intercept. CO₂ assimilation rate (A), stomatal conductance (gₛ), and water use efficiency (WUE) values are listed for low (200 μL L⁻¹), ambient (400 μL L⁻¹), and high (1,000 μL L⁻¹) CO₂. Measurements were made with a LI-COR 6400XT by using the LI-COR 6400-40 leaf fluorescence cuvette. Values are taken from A/Cᵢ curves performed on the 16th youngest leaf of four independent 10-week old plants from each plant line.

<table>
<thead>
<tr>
<th></th>
<th>WT (COL)</th>
<th>βca2</th>
<th>βca4</th>
<th>βca2βca4</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO₂ comp. pt.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200 μL L⁻¹ CO₂</td>
<td>55.9 ± 2.0</td>
<td>53.1 ± 1.4</td>
<td>59.9 ± 4.1</td>
<td>56.4 ± 4.1</td>
</tr>
<tr>
<td>A</td>
<td>4.9 ± 0.32</td>
<td>5.3 ± 0.71</td>
<td>6.3 ± 0.38</td>
<td>5.3 ± 0.90</td>
</tr>
<tr>
<td>gₛ</td>
<td>0.28 ± 0.04</td>
<td>0.25 ± 0.04</td>
<td>0.37 ± 0.03</td>
<td>0.35 ± 0.04</td>
</tr>
<tr>
<td>WUE</td>
<td>0.16 ± 0.02</td>
<td>0.18 ± 0.01</td>
<td>0.15 ± 0.01</td>
<td>0.14 ± 0.01</td>
</tr>
<tr>
<td>400 μL L⁻¹ CO₂</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>11.2 ± 1.1</td>
<td>12.1 ± 0.85</td>
<td>13.6 ± 0.6</td>
<td>11.4 ± 1.7</td>
</tr>
<tr>
<td>gₛ</td>
<td>0.32 ± 0.03</td>
<td>0.29 ± 0.04</td>
<td>0.38 ± 0.03</td>
<td>0.35 ± 0.04</td>
</tr>
<tr>
<td>WUE</td>
<td>0.33 ± 0.04</td>
<td>0.35 ± 0.02</td>
<td>0.33 ± 0.01</td>
<td>0.30 ± 0.02</td>
</tr>
<tr>
<td>1,000 μL L⁻¹ CO₂</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>18.0 ± 2.1</td>
<td>19.7 ± 1.4</td>
<td>20.8 ± 1.1</td>
<td>18.7 ± 2.2</td>
</tr>
<tr>
<td>gₛ</td>
<td>0.29 ± 0.04</td>
<td>0.21 ± 0.04</td>
<td>0.38 ± 0.03</td>
<td>0.35 ± 0.04</td>
</tr>
<tr>
<td>WUE</td>
<td>0.57 ± 0.12</td>
<td>0.77 ± 0.09</td>
<td>0.51 ± 0.03</td>
<td>0.49 ± 0.02</td>
</tr>
</tbody>
</table>

of change in photosynthesis rate of the double mutant. Leaf sections from various 10-week old plant lines grown in low CO₂ were incubated in a propidium iodide solution for five minutes before being imaged by confocal microscopy. Confocal images revealed no change in stomate density in the double mutant plants when compared to wildtype, single mutant, or complemented plants (Figure 3.14, A, B, and C).

**Amino acid analysis of the βca2βca4 double mutant**

PEP carboxylase in the cytoplasm uses HCO₃⁻ to generate 50% of the aspartate in leaf cells of *Nicotiana tabacum* (Melzer and O’Leary, 1987). Leaf samples from wildtype, βca2, βca4, and βca2βca4 plants grown in 200 μL L⁻¹ CO₂ were analyzed for amino acid content to determine if aspartate levels as well as other amino acid levels are altered in the double mutant. The amino acid levels of wildtype plants are comparable to levels in the single mutants, βca2 and βca4 (Figure 3.15 and Table 3.5). In leaf samples of the double mutant, the aspartate
Figure 3.14. The stomatal density of 10 week old carbonic anhydrase single and double mutant plants are unchanged from wildtype values. (A) Average stomatal densities of wildtype (COL), βca2, βca4, βca2βca4, and the complementation line, βca2βca4 OE βCA2. Each bar represents the mean ± SD of eight independent plants. (B) COL and (C) βca2βca4 leaf samples stained with propidium iodide imaged using confocal microscopy.
Figure 3.15. Free amino acid levels in wildtype (COL), βca2, βca4, and βca2βca4 plants. Aspartate and glutamine pools are significantly lower in the double mutant plants whereas the glycine and serine pools of the double mutant are elevated compared to wildtype (COL) and single mutant plants. Samples were sent to TAMU Protein Chemistry Laboratory at Texas A&M for amino acid analysis. Each bar represents the mean ± SD of three independent plants. Differences between the plant lines were tested using the Student-Newman-Keuls ANOVA. Means that are significantly different (P ≤ 0.05) are represented by different letters.

concentration is only 87 ± 28 μg g⁻¹ of leaf tissue, well below the aspartate levels of wildtype and single mutant plants (Figure 3.15 and Table 3.5). Interestingly, glutamate and glutamine levels are also lower in the double mutant, whereas glycine and serine are higher in the double-mutant (Figure 3.15 and Table 3.5). Since the double mutant plants grew better in high CO₂ (1,000 μL L⁻¹ CO₂) conditions, the amino acid content was analyzed in wildtype (COL) and double mutant plants grown in high CO₂ (Figure 3.16 and Table 3.6). The amino acid levels returned to near wildtype levels in double mutant plants grown in high CO₂, suggesting amino acid biosynthesis may play a role in the CO₂ dependent phenotype seen in the double mutant.
Table 3.5. Levels of free amino acid pools of wildtype (COL), βca2, βca4, and βca2βca4 plants grown at low CO₂. Samples were sent to TAMU Protein Chemistry Laboratory at Texas A&M for amino acid analysis. Each value represents the mean ± SD of three independent plants. Differences between the plant lines were tested using the Student-Newman-Keuls ANOVA. Means that are significantly different (P ≤ 0.05) are represented by different letters.

<table>
<thead>
<tr>
<th>Amino Acids</th>
<th>Amino Acid Composition (μg g⁻¹ Fresh wt.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>COL</td>
</tr>
<tr>
<td>ASP</td>
<td>185.3 ± 12.6ᵃ</td>
</tr>
<tr>
<td>GLU</td>
<td>339.1 ± 13.0ᵃ</td>
</tr>
<tr>
<td>ASN</td>
<td>23.5 ± 2.2ᵃ</td>
</tr>
<tr>
<td>SER</td>
<td>83.6 ± 9.6ᵃ</td>
</tr>
<tr>
<td>GLN</td>
<td>246.5 ± 19.5ᵃ</td>
</tr>
<tr>
<td>HIS</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>GLY</td>
<td>76.5 ± 18.3ᵃ</td>
</tr>
<tr>
<td>THR</td>
<td>59.0 ± 11.4ᵃ</td>
</tr>
<tr>
<td>ALA</td>
<td>45.5 ± 7.2ᵃ</td>
</tr>
<tr>
<td>ARG</td>
<td>9.3 ± 5.0ᵃ</td>
</tr>
<tr>
<td>TYR</td>
<td>1.2 ± 0.2ᵃ</td>
</tr>
<tr>
<td>VAL</td>
<td>8.2 ± 1.5ᵃ</td>
</tr>
<tr>
<td>MET</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>TRP</td>
<td>6.8 ± 8.7ᵃ</td>
</tr>
<tr>
<td>PHE</td>
<td>9.7 ± 2.3ᵃ</td>
</tr>
<tr>
<td>ILE</td>
<td>4.3 ± 2.9ᵃ</td>
</tr>
<tr>
<td>LEU</td>
<td>1.9 ± 0.5ᵃ</td>
</tr>
<tr>
<td>LYS</td>
<td>1.9 ± 0.3ᵃ</td>
</tr>
<tr>
<td>PRO</td>
<td>24.5 ± 5.0ᵃ</td>
</tr>
</tbody>
</table>

Complementing the βca2βca4 double mutant

To confirm that T-DNA inserts in the βCA2 and βCA4 genes are responsible for the reduced growth of the βca2βca4 line, complementation lines expressing the wildtype βCA2 coding region powered by a 2x35S promoter were generated. Reestablishing the wildtype βCA2 coding region in the double mutant restored wildtype growth at low CO₂ (Figure 3.17, A and B). Using the spinach CA antibody, the protein band for βca2 was present in the βca2βca4 p35S::βCA2 plant (Figure 3.17, C). Genomic PCR was also performed on the βca2βca4 p35S::βCA2 plant to ensure both βCA2 and βCA4 genes were still disrupted by T-DNA inserts.
Figure 3.16. Free amino acid levels in wildtype (COL) and βca2βca4 plants grown in high (1,000 μL L\(^{-1}\) CO\(_2\)). When grown in high CO\(_2\) levels, the amino acid levels of the βca2βca4 plants are similar to levels seen in wildtype (COL) plants. Samples were sent to TAMU Protein Chemistry Laboratory at Texas A&M for amino acid analysis. Each bar represents the mean ± SD of three independent plants. Differences between the plant lines were tested using the unpaired T-test with Welch correction. Means that are significantly different (P ≤ 0.05) are represented by (*).

(Figure 3.18). Adding βCA2 back to βca2βca4 plants restores normal growth in low CO\(_2\) conditions indicating this is a CA-facilitated problem.

DISCUSSION

In this chapter, I present evidence that βCA2 and βCA4 are the two most abundant cytosolic CAs in Arabidopsis leaves. Eliminating either βCA2 or βCA4 produces plants that are indistinguishable from wildtype plants (Figure 3.12, A). However, disrupting both βCA2 and βCA4 together resulted in a plant that exhibited slow growth and chlorosis at 200 μL L\(^{-1}\) CO\(_2\) and an eight-hour photoperiod. High CO\(_2\) alleviated the poor growth phenotype of the double mutant (Figure 3.12, B). In high CO\(_2\) conditions, the double mutant more closely resembled
Table 3.6. Free amino acid levels of wildtype (COL) and βca2βca4 plants grown in high CO₂ (1,000 μL L⁻¹ CO₂). Samples were sent to TAMU Protein Chemistry Laboratory at Texas A&M for amino acid analysis. Each value represents the mean ± SD of three independent plants. Differences between the plant lines were tested using the unpaired T-test with Welch correction. Means that are significantly different (P ≤ 0.05) are represented by (*).

<table>
<thead>
<tr>
<th>Amino Acids</th>
<th>Amino Acid Composition (μg g⁻¹ Fresh wt.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>COL</td>
</tr>
<tr>
<td>ASP</td>
<td>227.7 ± 12.4</td>
</tr>
<tr>
<td>GLU</td>
<td>423.9 ± 9.6</td>
</tr>
<tr>
<td>ASN</td>
<td>88.5 ± 39.6</td>
</tr>
<tr>
<td>SER</td>
<td>48.6 ± 8.1</td>
</tr>
<tr>
<td>GLN</td>
<td>422.0 ± 17.1</td>
</tr>
<tr>
<td>HIS</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>GLY</td>
<td>17.2 ± 10.1</td>
</tr>
<tr>
<td>THR</td>
<td>50.4 ± 6.5</td>
</tr>
<tr>
<td>ALA</td>
<td>37.2 ± 9.6</td>
</tr>
<tr>
<td>ARG</td>
<td>59.1 ± 24.1</td>
</tr>
<tr>
<td>TYR</td>
<td>1.5 ± 0.2</td>
</tr>
<tr>
<td>VAL</td>
<td>9.4 ± 0.5</td>
</tr>
<tr>
<td>MET</td>
<td>48.1 ± 6.0</td>
</tr>
<tr>
<td>TRP</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>PHE</td>
<td>5.9 ± 0.9</td>
</tr>
<tr>
<td>ILE</td>
<td>5.0 ± 2.9</td>
</tr>
<tr>
<td>LEU</td>
<td>1.9 ± 1.6</td>
</tr>
<tr>
<td>LYS</td>
<td>3.3 ± 0.8</td>
</tr>
<tr>
<td>PRO</td>
<td>28.4 ± 4.2</td>
</tr>
</tbody>
</table>

wildtype plants in size although minor chlorosis was still evident (Figure 3.12, B). Surprisingly, photosynthesis was not impaired in the double mutant even though a CO₂ effect on growth was observed (Table 3.4). However, analyzing the free amino acid content in leaves of the double mutant showed that aspartate was significantly lower in the double mutant compared to wildtype leaves (Figure 3.15). Since 50% of the aspartate in the plant is made as a result of PEPC activity (Melzer and O’Leary, 1987) we hypothesize the loss of βCA2 and βCA4 lowers PEPC activity in the double mutant. Our results are consistent with the amino acid concentrations seen in PEPC knockout plants (Shi et al., 2015) and support the hypothesis that CA activity is required for optimal PEPC activity in the cytoplasm.
Figure 3.17. Expressing the βCA2 coding region in βca2βca4 plants restores wildtype growth in low CO₂. (A) Normal growth was restored in βca2βca4 p35S::βCA2 plants growing at 200 μL L⁻¹ CO₂. All plants were grown under an 8 hour light (22° C)/16 hour dark (18° C) regime with a light intensity of 120 μmol photons m⁻² s⁻¹. (B) Weekly average projected rosette areas for each line showed normal growth was restored in the double mutant by adding p35S::βCA2. Each symbol represents the mean ± SD of nine individual plants per line. (C) Western blotting shows the presence of the βCA2 protein restored in the βca2βca4 p35S::βCA2 plants. Each lane contains 5 μg μL⁻¹ of total protein extract from leaves wildtype (COL), βca2βca4, and βca2βca4 p35S::βCA2 plants.
Figure 3.18. Genomic PCR probing for the p35S::βCA2 construct in the βca2βca4 mutant. A βca2βca4 mutant transformed with the p35S::βCA2 construct displays a band for the βCA2 CDS whereas the wildtype (COL) and double-mutant samples do not display this band. The complemented line still does not display the genomic βCA2 nor βCA4 bands but still shows the PCR products for the T-DNA inserts in the βCA2 and βCA4 genes. Actin was used as a loading control.

The poor growth of the double mutant but not the single mutants indicates that βCA2 and βCA4 have overlapping functions. One hypothesis examined in this study was that the double mutant grew slowly on low CO₂ because cytoplasmic CA activity is required to facilitate the diffusion of inorganic carbon to the chloroplast for photosynthesis. In this scenario, photosynthesis in the double mutant would be reduced because the CO₂ concentration at Rubisco would be reduced. However, CO₂ assimilation rates in all the mutant lines were similar to values in 10-week old wildtype plants (Table 3.4). From these measurements, it was concluded that the cytoplasmic CAs are not playing an important role in photosynthesis. These observations are in
agreement with the models of Badger and Price (1994), Terashima et al., (2011), Tholen et al. (2012), and Tholen, Ethier, and Genty (2014). They argued that the cytoplasm only offers minimal resistance to CO₂ diffusion because the chloroplasts are often close to the plasma membrane in mesophyll cells.

Another possible role of the cytoplasmic CAs would be to provide HCO₃⁻ for cytoplasmic PEP carboxylase. While PEPC is normally thought to have a very high affinity for inorganic carbon compared to Rubisco, the PEPC Kₘ (HCO₃⁻) has been reported to be between 25-100 μM for C₄ plants (O’Leary, 1982; Bauwe, 1986; Hatch and Burnell, 1990) and between 100 and 200 μM for C₃ plants (Mukerji and Yang, 1974; Sato, Koizumi, and Yamada, 1988). Since the dissolved CO₂ concentration in the cytoplasm of C₃ plants is expected to be about 12 μM at 400 μL L⁻¹ CO₂ and 25°C, the HCO₃⁻ concentration at equilibrium would be close to 60 μM assuming a cytoplasmic pH of 7.0 and a CO₂ to HCO₃⁻ pKa of 6.4. This estimated HCO₃⁻ concentration in the cytosol is less than the 100-200 μM PEPC Kₘ (HCO₃⁻) for C₃ plants. A drop in CO₂ concentration to 200 μL L⁻¹ in air would therefore drop the HCO₃⁻ to about 25-30 μM, again well under the reported Kₘ(HCO₃⁻) for PEPC. It is estimated that 50% of carbon at position four in aspartate can be attributed to PEPC activity (Melzer and O’Leary, 1987); therefore we measured aspartate levels in wildtype plants, the single mutants, and the double mutant (Figure 3.14 and Table 3.5). The wildtype plants and the single mutants had similar amino acid profiles with aspartate levels above 170 μg g⁻¹ of leaf tissue. In contrast, the aspartate level in the double mutant was only 87 ± 28 μg g⁻¹ of leaf tissue. The double mutant also showed elevated levels of glycine and serine (Figure 3.14). Plants lacking PEPC (Shi et al., 2015) also have low levels of aspartate and very high levels of glycine and serine. The PEPC double mutant, containing T-DNA insertions in the two highest expressed PEPC genes in leaves, exhibited a reduced growth
rate and was chlorotic (Shi et al., 2015). While the βca2βca4 double mutant also showed reduced growth and some initial chlorosis (Figure 3.11), the βca2βca4 plants appeared healthier than the PEPC double mutant. The healthier appearance of the βca2βca4 double mutant would be consistent with the fact that the βca2βca4 double mutant still has PEPC activity.

The results reported here also clarify the gene models for βCA2 and βCA4. Previously, only βCA2 (Fett and Coleman, 1994; Fabre et al., 2007) and βCA3 (Fabre et al., 2007) were reported to be in the cytoplasm. The only reports for βCA4 localization indicated the protein was associated with the plasma membrane (Fabre et al., 2007; Hu et al., 2010; Hu et al., 2015).

In addition, we found the gene models presented in TAIR are incomplete for βCA4 and inaccurate for βCA2. For βCA2, the gene models all show an exon/intron pattern very similar to βCA1 and predict that βCA2 should localize in the chloroplast. However, RNAseq data (Figure 3.1) and deposited ESTs show that >95% of the βCA2 transcripts begin in the middle of the second exon in the TAIR model. This results in an mRNA that encodes a cytoplasmic protein which was shown in the work of Fett and Coleman (1994) and Fabre et al. (2007). The fact that the mature βCA2 protein (28kDa) is larger than the mature βCA1 protein (25kDa) (Figure 3.2 and Figure 3.10, A) further supports a cytoplasmic localization for βCA2. Abundant GUS staining is seen in the leaf when GUS is linked to a promoter made immediately upstream of the ATG start site for the predicted cytoplasmic βCA2 (Figure 3.5). This GUS expression pattern for βCA2 coincides with the leaf CA activity levels reported in pea (Majeau and Coleman, 1994) and bean (Porter and Grodzinski, 1984) and also fits with observed EST abundance in TAIR, microarray data (Ferreira, Guo, and Coleman, 2008; Schmid et al., 2005; Winter et al., 2007), as well as RNAseq data (Table 3.1). Our results are in contrast to those of Wang et al. (2014) who observed little or no GUS staining in the leaf with their βCA2 promoter. However, they used a
sequence upstream of the first exon of the TAIR gene model and our data and the EST data indicate that the first exon in the TAIR model is transcribed at a very low level, if at all.

The gene model for βCA4 is somewhat complex. There are two different and abundant transcripts made from the βCA4 gene in the leaf, βCA4.1 and βCA4.2 (Table 3.1 and Figure 3.3). The longer leaf transcript, βCA4.1, encodes a protein that is targeted to the plasma membrane as shown by our localization studies and the published work of Fabre et al. (2007), Hu et al., (2010), and Hu et al., (2015). Here, we also determined that βCA4.2, the shorter transcript, encodes a cytoplasmic CA (Figure 3.6 and Figure 3.7). The RNAseq data also show that only βCA4.2 is expressed in the roots while both βCA4.1 and βCA4.2 are expressed in leaf tissue. GUS expression studies support the RNAseq data. When a βCA4.2 promoter (Figure 3.4, C) was linked to GUS, only root expression was observed (Figure 3.5). A more complete promoter linked to GUS showed both root and leaf expression (Figure 3.5). However, we were unable to find a βCA4.1 promoter sequence that showed only expression in leaves (data not shown). Other studies have found multiple transcripts of βCA4 (Aubry et al., 2014; Wang et al., 2014) but this is the first report showing that the different transcripts encode proteins with different destinations in the cell.

βCA2 and βCA4 are highly expressed in leaves (Figure 3.5) The only CA gene transcribed at higher levels in the leaf is βCA1. The only other CA reported to be in the cytoplasm, βCA3, is expressed at less than 5% of the level of either βCA2 or βCA4 (Table 3.1; Fabre et al., 2007; Ferreira, Guo, and Coleman, 2008; Schmid et al., 2005; Winter et al., 2007). In addition, Hu et al. (2010) presented evidence showing high expression of βCA2 and βCA4 in mesophyll cells, whereas βCA3 had very low expression in mesophyll cells. Using promoter::GUS analysis, Wang et al. (2014) showed very low expression of βCA3. Since the
other CAs that show significant expression are either in the chloroplast (Fett and Coleman, 1994; Villarejo et al., 2005; Fabre et al., 2007; Burén et al., 2011; Hu et al., 2015) or mitochondria (Fabre et al., 2007; Jiang et al., 2014), we conclude that βCA2 and βCA4 are the most abundant CA isoforms in the cytoplasm. This contention is supported by leaf RNAseq data (Table 3.1), GUS staining (Figure 3.5), CA microarray analysis (Ferreira, Guo, and Coleman, 2008), as well as publically available EST and microarray data (Schmid et al., 2005; Winter et al., 2007).

Previously, researchers lowered the expression of the chloroplastic CA, βCA1 (Majeau, Arnoldo, and Coleman, 1994; Price et al., 1994; Ferreira, Guo, and Coleman, 2008) and found normal growth and carbon assimilation rates in plants with reduced βCA1. More recently, Jiang et al. (2014) reported that plants lacking the mitochondrial βCA6 grew slowly on low CO₂. There have been few studies of other CA isoforms. A notable exception has been the construction of double and triple mutant lines of the most abundant CAs in leaf guard cells including βCA1, βCA4, and βCA6 (Hu et al. 2010, Xue et al. 2011, and Hu et al. 2015). These CAs are localized to different organelles in the guard cell with βCA1 located in the chloroplast, βCA4 in the plasma membrane, and βCA6 in the mitochondria. Eliminating the expression of these guard cell CAs caused changes in stomatal density and temporal changes in stomatal conductance in response to changes in CO₂ level or humidity (Hu et al., 2010; Engineer et al., 2014) leading to the hypothesis that CA activity is an important factor in how plant guard cells sense CO₂ concentration (Xue et al., 2011). In the present study, stomatal density was unaffected by knocking out cytosolic βCA2 and/or βCA4 (Figure 3.13).

In conclusion, this chapter presents evidence that βCA2 and βCA4 are the most abundant cytoplasmic CAs in Arabidopsis leaves. Loss of both of these proteins causes reduced growth on short days and low CO₂ concentrations although photosynthetic properties appear unaffected.
Our results are similar to those of Studer et al., (2014) who found that knocking out CA1 and CA2 in maize led to reduced growth but no significant effect on photosynthesis. These cytosolic CAs still appear important in plant carbon metabolism as judged by the free amino acid profiles of the double mutant in our study. We hypothesize that βCA2 and βCA4 are necessary for the proper function of cytosolic PEPC needed for production of amino acids. Our growth and amino acid profiles are consistent with the recent work of Shi et al. (2015) who studied an Arabidopsis double mutant missing the two most abundant isoforms of PEPC. They also observed low levels of ASP and GLN as well as stunted growth and chlorosis. It is also likely that the large number of CA genes and isoforms in plants indicates that CA may be needed for a number of metabolic pathways in different tissues.
CHAPTER 4
THE CHLOROPLAST CARBONIC ANHYDRASES βCA1 AND βCA5

INTRODUCTION

The Chloroplast became an organelle of eukaryotic photosynthetic organisms as a result of an endosymbiotic event between a cyanobacteria-like cell and a host cell over 600 million years ago (Mereschkowsky, 1905 in Martin and Kowallik, 1999; McFadden and van Dooren, 2004). Chloroplasts still retain a partial genome containing 100-200 genes depending on the organism (Sugiura, 1994; Shinozaki et al., 1986; Maier et al., 1995) as well as the machinery to generate proteins from those genes. Over time, most of the chloroplast genome has been transferred to the nucleus. To function in the chloroplast, nuclear-encoded proteins require a chloroplast transit peptide (cTP) (Dobberstein, Blobel, and Chua, 1977; Chua and Schmidt, 1978; Bruce, 2000) and cells need a chloroplast protein targeting system (Highfield and Ellis, 1978; Soll and Schleiff, 2004). cTPs are peptide sequences found at the N-terminus of a protein sequence that are recognized by chaperones in the cytosol and these chaperones guide chloroplast proteins to the outer membrane of the chloroplast (Soll and Schleiff, 2004; Lee, Jung, and Hwang, 2012). cTPs are difficult to detect, as their sequence lengths are variable and amino acid content resemble mitochondrial transit peptides (mTPs). Generally, a cTP is between 40-100 amino acids in length (Bruce, 2000; Bionda et al., 2010). Also, the cTP region is enriched in small, hydroxylated amino acids and small hydrophobic amino acids while excluding acidic amino acid residues, much like mTPs (Bhushan et al., 2006).

Eukaryotic photosynthetic organisms have multiple chloroplast carbonic anhydrases that are nuclear encoded. These nuclear encoded chloroplast CAs have been linked to both photosynthesis and lipid biosynthesis. In Chlamydomonas, the chloroplast lumen CA, CAH3, is an important component of the algal CCM (Karlsson et al., 1998; Moroney and Ynalvez, 2007;
Duanmu, Wang, and Spalding, 2009). Located in the pyrenoid under low CO₂ conditions (Blanco-Rivero et al., 2012), it is thought that CAH3 dehydrates HCO₃⁻ to CO₂ to supply Rubisco in the pyrenoid (Moroney and Ynalvez, 2007; Duanmu, Wang, and Spalding, 2009). Studies removing CAH3 from Chlamydomonas resulted in cells having poor growth in low CO₂ (Karlsson et al., 1998; Duanmu, Wang, and Spalding, 2009) as well as accumulating high level of Cᵢ as the accumulated HCO₃⁻ cannot be used by Rubisco (Duanmu, Wang, and Spalding, 2009).

It is much harder to associate chloroplast CA activity with photosynthesis in plants. Researchers have tried linking chloroplast CA activity to C₃ photosynthesis. Mathematical models by Badger and Price (1994) indicate that CA activity may have a minimal effect on C₃ photosynthesis. Other models predict a minimal effect of CA activity on C₃ photosynthesis because chloroplasts are pushed up against the plasma membranes of mesophyll cells, reducing the CO₂ diffusion pathway length (Terashima et al., 2011; Tholen et al., 2012; Tholen, Ethier, and Genty, 2014). Early RNAi studies reducing the amount of chloroplast CA activity in tobacco did not reduce photosynthesis rates (Price et al., 1994), although changes in water use efficiency (WUE) and ¹³C isotopic discrimination of the RNAi plants were reported (Price et al., 1994; Majeau, Armoldo, and Coleman, 1994).

Although changes in chloroplast CA activity might not affect photosynthesis, CA activity has been shown to affect plastid lipid synthesis. Reducing chloroplast CA activity using ethoxyzolamide reduces lipid synthesis in cotton embryos (Hoang and Chapman, 2002). In a follow-up experiment, the addition of ethoxyzolamide to tobacco cell suspensions also reduces lipid synthesis (Hoang and Chapman, 2002). Tobacco CA-antisense lines also show reductions in total lipid synthesis (Hoang and Chapman, 2002).
In Arabidopsis, I found that plants lacking the chloroplast CA, βCA5, grow very poorly under ambient CO₂ conditions whereas plants lacking the chloroplast CA, βCA1, grow normally. I hypothesize that βCA1 and βCA5 play two different physiological roles in plant cells and that the complete removal of βCA5 negatively impacts lipid synthesis and plant development after germination. In this chapter, I examine why lacking one chloroplast CA, βCA5, results in a drastic change in phenotype, whereas lacking another chloroplast CA, βCA1, does not change the plant phenotype.

RESULTS

βca1 and βca5 mutants have drastically different phenotypes

The βca1 mutant line (SALK_106570) has a T-DNA insertion in the last exon of the βCA1 gene whereas the βca5 mutant line (SALK_121932) has a T-DNA insertion in the eighth exon of the βCA5 gene (Figure 4.1). Genomic PCR confirmed that each T-DNA insert is

Figure 4.1. Locations of the T-DNA insertions within the βCA1 and βCA5 genes. The βCA1 T-DNA insertion (SALK_106570) is in the ninth exon of the βCA1 gene. The βCA5 T-DNA insert (SALK_121932) is in the eighth exon of the βCA5 gene. Triangles represent T-DNA insertions and black arrows represent locations of gene specific (F and R) primers and insert (I) primers. Blue boxes and blue lines represent exons and introns, respectively. Red boxes and red arrows represent the 5’ and 3’ UTRs, respectively.
disrupting their respective gene. Amplifying from the insert located in either the βCA1 or βCA5 gene generated PCR bands from βca1 or βca5 mutant plants only (Figure 4.2). Using genespecific primers that span the putative location of the insert generated bands only from wildtype samples indicating that the βCA1 and βCA5 genes are disrupted by the T-DNA insert in the βca1 and βca5 lines, respectively (Figure 4.2).

![Figure 4.2](image.png)

Figure 4.2. The βCA1 and βCA5 genes are disrupted by T-DNA insertions. Genomic PCR of wildtype, βca1 and βca5 plants show that the T-DNA insert, SALK_106570, is present in the βca1 line and the T-DNA insert, SALK_121932, is present in the βca5 line. Both insert disrupt their corresponding genes in the βca1 and βca5 lines.

To confirm that βca1 plants are knockouts, Western blots were performed using antibodies directed against Arabidopsis βCA1 and spinach CA. (Figure 4.3). Using the antibody directed against Arabidopsis βCA1 produced a protein band around 25 kDa in the wildtype lane only (Figure 4.3, A). The antibody directed against spinach CA produced three bands in the
wildtype lane (Figure 4.3, B). The band near 25 kDa represents the βCA1 protein as this protein band runs at the same position as the protein band obtained with the Arabidopsis βCA1 antibody. This 25 kDa band obtained with the spinach CA antibody is seen in both the wildtype and mutant βca2 protein samples (Figure 4.3, B). There is a faint band near 25 kDa in the mutant βca1 protein sample (Figure 4.3, B). The second protein band around 28 kDa represents the βCA2 protein, as bands are found in the wildtype and βca1 lanes, but not the βca2 lane (Figure 4.3, B).

![Figure 4.3](image)

**Figure 4.3.** The βCA1 protein is missing in the βca1 mutant. (A) A Western blot using a polyclonal antibody raised against a peptide sequence of the Arabidopsis βCA1 protein. The band near 25 kDa, which is the predicted size of βCA1, is missing in the βca1 mutant. (B) A Western blot using a polyclonal antibody raised against extracted spinach CA. This antibody reacts with βCA1 seen near 25 kDa and βCA2 seen near 28 kDa. The 25 kDa band is missing in the βca1 mutant and the 28 kDa band is absent in the βca2 mutant. A very thin band near 25 kDa can be seen in the βca1 mutant.

The third band near 55 kDa in all three lanes is assumed to be the large subunit of Rubisco. The faint 25 kDa band picked up by the spinach CA antibody in the βca1 sample could be the chloroplastic long form of βCA2, as the sequence identity between βCA1 and βCA2 is nearly
90% (Figure 4.4). Also, when the cTP of the chloroplastic βCA2 is cleaved, its new molecular weight is comparable to the mature βCA1 protein (Figure 4.4).

| βCA1 (At3g01500) | MSTAFLSGFLTSLSQYQSLKLSLKTSSTVACLPAASSSSSSEESSEERVF----
| βCA2.1 (At5g14740) | ----MVFPWTVSRRGSSDSGETLQASA------TKQKYPLSPHLHSLILFF
| βCA2.2 (At5g14740) | 

| βCA1 (At3g01500) | -TLIRNEFVAAPAPIIAPYWSEEMTEAYDEAIAlanlllleKEELKTVAAKVEQITA
| βCA2.1 (At5g14740) | PHLSANGACFRCTCFSHPKLELRMENGSIEDAELKALKLKELGDADVAAAKKEITA
| βCA2.2 (At5g14740) | **:***:***:***:***:***:***:***:***:***:***

| βCA1 (At3g01500) | ALQTSTSSDSKAFDPVETIKQGFIKFEEKETNPALYGELAKQSPKYNFPACSDS
| βCA2.1 (At5g14740) | ELQAASSDSKAFPVERSEKQGFIKFEEKETNPALYGELAKQSPKYNFPACSDS
| βCA2.2 (At5g14740) | ELQAASSDSKAFPVERSEKQGFIKFEEKETNPALYGELAKQSPKYNFPACSDS

**Figure 4.4.** The predicted mature chloroplast carbonic anhydrase, βCA2.1, is nearly identical in size and sequence to βCA1. Removing the chloroplast transit peptides, marked in red, upon entry into the chloroplast generates mature βCA1 and βCA2.1 proteins that are predicted to be 232 and 231 amino acids in length with a mature peptide sequence identity of 87 percent.

We were unable to generate an antibody directed against βCA5, so RT-PCR was performed to see if transcription was blocked by the T-DNA insertion in the βca5 plants. Using gene specific primers to detect the βCA5 cDNA, a band was present in the wildtype sample, but not the βca5 sample (Figure 4.5) indicating transcription of the βCA5 gene has been stopped by the T-DNA insertion. Since RT-PCR was performed on the βca5 mutant, I did the same for the βca1 mutant to confirm the results of the Western blot. Using βCA1 specific primers, the βCA1 message could be detected in the wildtype cDNA sample, but not in the βca1 mutant sample.
(Figure 4.5) suggesting the T-DNA insertion has disrupted transcription which results in the lack of βCA1 protein as seen in the protein sample from βca1 mutant used in the Western blot.

![Image of RT-PCR results for βCA1 and βCA5 genes](image)

Figure 4.5. The T-DNA insertions of the βca1 and βca5 mutant lines have reduced transcription of the βCA1 and βCA5 genes, respectively. Using gene-specific primers for RT-PCR show that βCA1 messages cannot be detected in the βca1 mutant and βCA5 messages cannot be detected in the βca5 mutant. 3 μg of RNA was used from each sample to carry out RT-PCR and Actin (At2g37620) was used as a loading control.

**Growth of the βca1 and βca5 mutant lines**

Knocking out the βCA1 gene results in no apparent phenotype, but knocking out the βca5 gene results in plants having severe growth defects. In ambient CO₂, βca1 plants closely resemble wildtype (COL) plants whereas homozygous βca5 plants exhibit severe growth impediment (Figure 4.6 and Figure 4.7). Interestingly, growth of the homozygous βca5 plants is partially restored when growing in very high (50,000 μL L⁻¹) CO₂ (Figure 4.6 and Figure 4.7). Finding homozygous βca5 mutant plants was difficult. A growth phenotype was not expected for the βca5 mutant since most CA single mutant lines, especially the βca1 mutant line, do not
Figure 4.6. Growth histograms of wildtype (COL), βca1, and βca5 plants in ambient air (400 μL L⁻¹) and in very high CO₂ (50,000 μL L⁻¹). Wildtype (COL) and βca1 plants grow similarly in ambient and very high CO₂ conditions. The βca5 mutant plants grow very poorly in ambient air, but the growth of the βca5 plants improves in very high CO₂ conditions.

show a growth phenotype. The difficulty of finding a homozygous βca5 plant became apparent as homozygous βca5 seedlings show an immediate developmental phenotype (Figure 4.8). Additionally, the severity of the βca5 phenotype was surprising as earlier CA expression studies showed βCA5 expression levels are only 5% of the expression level of βCA1 (Fabre et al., 2007; Ferreira, Guo, and Coleman, 2008; Schmid et al., 2005; Winter et al., 2007), yet βca1 plants are indistinguishable from wildtype plants phenotypically. This raises the question, what is different between βCA1 and βCA5?

**βCA1 and βCA5 are both chloroplast proteins**

Both βCA1 and βCA5 are predicted chloroplast proteins (Table 4.1) and proteomics data show both βCA1 and βCA5 are found in the chloroplast (Sun et al., 2009). Studies using GFP fusions have shown that βCA1 and βCA5 are localized to the chloroplast (Fabre et al., 2007; Hu et al., 2015). Transient expression assays, however, show βCA1 present in the plasma...
Figure 4.7. Wildtype (COL), βca1, and βca5 plants grown in ambient (400 μL L\(^{-1}\)) and very high CO\(_2\) (50,000 μL L\(^{-1}\)). Wildtype and mutant plants 24 days post germination grown in either ambient air or very high CO\(_2\) conditions in 24 hour light at a light intensity of 80 μmol of photons m\(^{-2}\) sec\(^{-1}\).

membrane (Hu et al., 2010). To confirm that βCA1 and βCA5 are both chloroplast proteins, the eGFP coding region was fused to the 3’ end of the βCA1 and βCA5 coding regions. Protoplasts generated from stably transformed Arabidopsis plants expressing either 35S::βCA1-eGFP or 35S::βCA5-eGFP (Figure 4.9) confirm the chloroplast localization of βCA1 (Fabre et al., 2007; Hu et al., 2015) and βCA5 (Fabre et al., 2007).
Figure 4.8. The cotyledons of βca5 mutant plants already show a distinct phenotype as compared to wildtype (COL) plants. The lag in development of βca5 mutant plants is apparent as βca5 plants do not put out new leaves as quickly as wildtype (COL) plants. Wildtype (COL) plants 10 days post germination (DPG) have already put out their first two leaves and are starting to put out the next two leaves. βca5 plants 10 DPG are still in the cotyledon stage. Also, the cotyledons of the βca5 plants are darker in color as compared to wildtype (COL) cotyledons.

Table 4.1. Predicted subcellular locations of βCA1 and βCA5 using protein subcellular location prediction programs. Predotar, ChloroP, TargetP, and YLoc are four programs that predict the subcellular location of proteins based on amino acid sequence. A value of (1) denotes strong confidence in the prediction. Ch = Chloroplast, O = Other location; (+) = contains a chloroplast transit peptide, (-) does not contain a chloroplast transit peptide.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Predotar</th>
<th>ChloroP</th>
<th>TargetP</th>
<th>YLoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>βCA1</td>
<td>Ch (0.89)</td>
<td>(+)</td>
<td>Ch (0.97)</td>
<td>Ch (1)</td>
</tr>
<tr>
<td>βCA5</td>
<td>O (0.87)</td>
<td>(+)</td>
<td>Ch (0.92)</td>
<td>Ch (1)</td>
</tr>
</tbody>
</table>


Expression patterns of βCA1 and βCA5 vary

Deposited ESTs on TAIR show that βCA1 is the most highly expressed CA in Arabidopsis leaves (Table 4.2). The number of TAIR ESTs deposited for βCA5 suggests that it is expressed at a much lower level than βCA1 in Arabidopsis (Table 4.2). RNAseq analysis was performed on root and shoot RNA from three biological replicates to confirm the CA expression.
levels observed in TAIR (RNAseq data is deposited in the NCBI SRA database as BioSample:SAMN03339724). RNAseq reads were refined as normalized counts of 100 base-pair reads that uniquely mapped to a CA gene. The RNAseq data showed βCA1 as the predominant chloroplast CA in leaves, whereas βCA5 is only expressed at around 5% the level of βCA1 in leaves (Table 4.2). In roots, the predominant plastid CA is βCA5, whereas βCA1 expression is much lower (Table 4.2). The overall pattern of RNAseq values for βCA1 and βCA5 matches the pattern of EST values for βCA1 and βCA5 from TAIR confirming that βCA1 is expressed more than βCA5 in leaves (Table 4.2). Interestingly, when matching the RNAseq reads to the βCA1 gene, there seem to be two mRNA forms differing at the 3’ end, indicating another CA has multiple mRNA forms (Figure 4.10). ESTs from TAIR indeed confirm that
Table 4.2. \( \beta CA1 \) and \( \beta CA5 \) are expressed in roots and leaves. Three Arabidopsis plants were used for leaf and root RNA samples. RNAseq reads were normalized to 100 bp counts that uniquely mapped to a CA gene in the Arabidopsis reference genome. RNAseq values are given in RPKM (Reads Per Kilo base per Million mapped). Actin1 was used as a reference to compare CA expression levels.

<table>
<thead>
<tr>
<th>GENE</th>
<th>GENE I.D.</th>
<th>ESTs(^a)</th>
<th>RNAseq Reads(^b) (Shoots)</th>
<th>RNAseq Reads(^c) (Roots)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta CA1 )</td>
<td>At5g14740</td>
<td>1,538</td>
<td>109,333</td>
<td>662</td>
</tr>
<tr>
<td>( \beta CA5 )</td>
<td>At1g70410</td>
<td>116</td>
<td>2,262</td>
<td>2,443</td>
</tr>
<tr>
<td>Actin1</td>
<td>At2g37620</td>
<td>65</td>
<td>1,471</td>
<td>1,056</td>
</tr>
</tbody>
</table>

\(^a\)EST values were extracted from The Arabidopsis Information Resource webpage (TAIR).

\(^b,c\)RNAseq data generated by Oh et al. (2014)

Figure 4.10. There are two mRNA forms of \( \beta CA1 \) that differ at the 3’ ends. Shoot RNAseq reads that uniquely map to the \( \beta CA1 \) gene show the majority of \( \beta CA1 \) transcripts contain an extended ninth exon whereas only a few \( \beta CA1 \) transcripts contain the ninth intron and tenth exon represented in the \( \beta CA1 \) TAIR model. Blue lines and blue boxes represent introns and exons, respectively. The red boxes and red arrows represent the 5’ and 3’ UTRs, respectively. RNAseq data generated by Oh et al. (2014).

there is a second mRNA form (The Arabidopsis Information Resource). The majority of ESTs show an extended ninth exon, whereas a few ESTs match the predicted length of the ninth exon and contain the tenth exon from the TAIR model (Figure 4.10; The Arabidopsis Information Resource). Unlike \( \beta CA1 \), the RNAseq reads matched the \( \beta CA5 \) gene from the reference genome (Figure 4.11).
Figure 4.11. The RNAseq reads match well with the TAIR βCA5 gene model. RNAseq reads for βCA5 nicely align to the exons of the βCA5 reference gene. Blue boxes and lines represent exons and introns, respectively. Red boxes and the red arrow represent the 5’ and 3’ UTRs, respectively. RNAseq data generated by Oh et al. (2014).

To confirm the RNAseq results, the promoter regions and 5’ UTRs of βCA1 and βCA5 were cloned in front of the GUS gene to drive GUS expression (Figure 4.12). Analyzing the patterns of GUS staining in pβCA1::GUS and pβCA5::GUS plants show that βCA1 and βCA5 have almost complementary expression patterns (Figure 4.13). Stably transformed Arabidopsis seedlings expressing the pβCA1::GUS construct showed strong blue staining in the cotyledons and hypocotyls of the seedlings but not roots. However, pβCA5::GUS seedlings have GUS staining in roots and strong GUS staining in the root and shoot meristems (Figure 4.13). A closer look shows that pβCA1::GUS plants have very little GUS staining in their meristematic

Figure 4.12. The promoter regions used to make the pβCA1::GUS and pβCA5::GUS constructs. Blue boxes and blue lines represent exons and introns, respectively. Black lines represent the promoter regions used to drive expression of the GUS constructs. Red boxes represent the 5’ UTRs. Red arrows represent the 3’ UTRs and gene orientation. Purple lines represent intergenic regions.
regions (Figure 4.13, inserts). Three week old plants stably expressing the pβCA1::GUS construct showed strong staining in the rosette, but not the roots (Figure 4.14). The three week old plants expressing the pβCA5::GUS construct showed GUS staining in the youngest leaves of the Arabidopsis rosettes and in the meristematic region of the roots. Interestingly, complementary to pβCA5::GUS plants, mature rosettes of pβCA1::GUS plants contain little to

![Image](image_url)

**Figure 4.13.** Three-day old βCA1 and βCA5 seedlings show different expression patterns. Three day old pβCA1::GUS plants show strong GUS staining in the cotyledons. Three day old pβCA5::GUS plants show GUS staining primarily in the root and shoot meristems.
no GUS staining in their shoot meristem and leaf primordia (Figure 4.15). Since βCA5 is expressed in meristematic regions of cotyledons and in the youngest leaves of Arabidopsis rosettes, it is possible that βCA5 is needed for managing C\textsubscript{i} levels for lipid synthesis or nucleotide biosynthesis for cell division since those pathways occur in plastids and require a carboxylation step.

Figure 4.14. Three week old pβCA1::GUS and pβCA5::GUS plants still show different expression patterns. pβCA1::GUS plants show GUS expression only in the rosette leaves whereas pβCA5::GUS plants show GUS expression in root and shoot meristematic regions and younger leaves.
Figure 4.15. Multiple Independent pβCA1::GUS and pβCA5::GUS lines show different expression patterns in rosettes. Three independent pβCA1::GUS lines show lower GUS expression in shoot meristematic regions whereas three independent pβCA5::GUS lines show the strongest GUS expression in the shoot meristematic regions.
DISCUSSION

In this chapter, I present evidence that βCA1 and βCA5 are chloroplast proteins (Table 4.1 and Figure 4.9) and that their different expression patterns hint that these chloroplast CAs most likely have different functions in the plant. The most striking result of this chapter is the poor growth of the βca5 mutant when grown at ambient CO₂ (Figure 4.6 and Figure 4.7). The reduced growth of the βca5 plants at ambient CO₂ was at least partially restored to normal by high CO₂ (Figure 4.6 and Figure 4.7). Surprisingly, the βcal mutant plants grow just like wildtype (COL) plants in ambient air (Figure 4.6 and Figure 4.7). This is interesting because βCA1 comprises nearly 1% of all soluble protein in the leaf (Tobin, 1970; Okabe et al., 1984; Peltier et al., 2006) and βCA5 is expressed at only 5% of the expression level as βCA1 (Table 4.2). Despite such high levels of βCA1 in the leaf, it is the βca5 mutant that shows such a drastic reduction in growth in ambient air. The results of the growth study provide two questions to answer. One, why is the growth of the βca5 plants so different from the βca1 plants if βCA1 and βCA5 are both located in the chloroplast? And two, why does deleting βCA1, a highly abundant protein, not have an effect on the plant?

There are two possibilities as to why there is a growth difference between the βca1 and βca5 mutant plants. The first possibility, which is the less likely of the two, is that βCA1 and βCA5 are both involved in photosynthesis but have different sub-chloroplast localizations. The chloroplast stroma has a basic pH due to protons pumped into the lumen from the light reactions. Because of the basic pH, the predominant C₅ molecule in the stroma is HCO₃⁻, a carbon source that Rubisco cannot use to carboxylate RuBP. In Chlamydomonas, it is hypothesized that the accumulated HCO₃⁻ in the stroma enters the acidic lumen of the chloroplast where CAH₃, the thylakoid lumen CA, generates CO₂ from HCO₃⁻ to provide a usable carbon source for Rubisco
It is possible that βCA5 could be playing the role of CAH3 in Arabidopsis chloroplasts since removing βCA5 will result in poor growth just as Chlamydomonas mutants lacking CAH3 grow poorly in low CO₂ conditions. Proteins that enter the chloroplast lumen have a bipartite transit peptide where half of the cTP is cleaved entering the chloroplast stroma and the other signal is cleaved after entering the lumen (Albiniak, Baglieri, and Robinson, 2012). About 50 percent of chloroplast lumen proteins traverse one of two characterized pathways to enter the lumen and proteins for each pathway have conserved peptide motifs (Albiniak, Baglieri, and Robinson, 2012). Proteins that enter the lumen through the Twin-arginine translocation (TAT) pathway have a transit peptide that contains a conserved twin-RR motif, followed by a hydrophobic region, and ending with a conserved A-(X)-A terminal processing motif (Albiniak, Baglieri, and Robinson, 2012). Proteins that enter the lumen using the Sec pathway have a transit peptide containing a conserved lysine residue followed by a hydrophobic region (Albiniak, Baglieri, and Robinson, 2012). βCA1 and βCA5 do not have the conserved amino acid residues to be transported by the TAT or Sec pathways suggesting they do not enter the lumen via these pathways (Figure 4.16). Curiously, a report presented evidence of CA activity in isolated thylakoids from the Arabidopsis βca1 mutant background (Fedorchuk et al., 2014). More work is needed to confirm those observations. Other works, including entire Arabidopsis chloroplast proteome studies and Arabidopsis chloroplast lumen proteome studies indicate that there is no chloroplast lumen CA present in Arabidopsis (Friso et al., 2004; Sun et al., 2009). It is unlikely that βCA5 has a different sub-chloroplast localization than βCA1 which allows it to play the role of a thylakoid CA in Arabidopsis.
Figure 4.16. Protein alignments of βCA1 and βCA5. The protein sequences of βCA1 and βCA5 were aligned using Clustal Omega (Sievers et al., 2011). The predicted chloroplast transit peptides of βCA1 and βCA5 indicated in red lack the conserved peptide sequences needed to enter the thylakoid lumen through either the TAT pathway or SEC pathway.

The more likely possibility is that βCA1 and βCA5 have different functions in the plant. This idea is supported by the different tissue expression patterns of βCA1 and βCA5 as seen with the GUS plants (Figure 4.13, Figure 4.14, and Figure 4.15). Since βCA5 is primarily expressed in meristematic regions and young developing tissues, βCA5 could possibly be linked to lipid synthesis and not photosynthesis (Figure 4.13). Other CAs have been shown to be linked to lipid synthesis (Hoang and Chapman, 2002). If lipid synthesis is reduced, there may not be enough lipids for the membranes of dividing cells, slowing the development of plant tissues. The immediate phenotype seen with the βca5 cotyledons (Figure 4.8) and slow development and growth (Figure 4.6 and Figure 4.7) suggest that the βca5 plants have a developmental problem and not a photosynthesis problem. ACCase, the enzyme that catalyzes the first committed step of lipid biosynthesis, requires HCO₃⁻ to carboxylate acetyl-CoA. It is possible that a CA
interacts with ACCase to provide HCO$_3^-$ for carboxylation. The $K_m$ (HCO$_3^-$) of ACCase is near 1.4 mM (Herbert et al., 1996). Assuming the chloroplast stroma CO$_2$ concentration is 12 μM and assuming a leaf chloroplast stroma pH of 8.2, the stromal HCO$_3^-$ concentration would be near 750 μM optimally, assuming equilibrium conditions in the chloroplast. This concentration of HCO$_3^-$ is well below the $K_m$ (HCO$_3^-$) of ACCase, so the presence of a CA near ACCase may help with ACCase activity. The disparity between plastid HCO$_3^-$ concentration and the ACCase $K_m$ (HCO$_3^-$) is even greater in the roots. Since roots are not exposed to light, the plastids in root cells do not perform photosynthesis and the pH of the plastid stroma is expected to be closer to neutral. Assuming equilibrium in the root-plastid stroma, the HCO$_3^-$ concentration will be much lower and the lack of a CA interacting with ACCase will be even more problematic. Indeed, the root systems of βca5 plants are much smaller than the root systems of wildtype plants (data not shown). Growing βca5 plants with supplemented malonate, which bypasses the ACCase step of lipid synthesis, may improve the growth of the βca5 plants if it is a lipid synthesis problem. It is speculated that malonate can be converted to Malonyl-CoA by a homolog of Malonyl-CoA synthetase found in Arabidopsis (Baud et al., 2004). This possible bypass of the ACCase step was demonstrated by Baud et al. (2004) where they found that exogenously added malonate can partially compensate for the loss of cytosolic ACCase in Arabidopsis plants.

It was also posited that CAs may affect nucleotide biosynthesis (Raven and Newman, 1994; Hoang and Chapman, 2002). Carbamoyl-phosphate synthetase and AIR carboxylase both require C$_i$ to catalyze reactions for pyrimidine and purine biosynthesis, respectively (Zrenner et al., 2006). Lowering chloroplast CA activity reduces the CO$_2$ partial pressure in the chloroplast (Price et al., 1994). It is possible that removing the chloroplast CAs in Arabidopsis will lower the C$_i$ substrates in the chloroplast for these two reactions, possibly slowing nucleotide
biosynthesis. If nucleotide biosynthesis is reduced enough, it may affect the rate of cell division, slowing down the growth and development of the plant.

The lack of a growth phenotype when removing the highly abundant βCA1 protein from the plant is baffling. It is possible that the βca1 mutant still has some leaky expression although Western blots and RT-PCR show that βca1 is eliminated from the plant (Figure 4.3 and Figure 4.5). Interestingly, when a leaf protein extract is probed with an anti-CA antibody, there is a faint band near the size of the βCA1 protein that is present when the large βCA1 band is missing (Figure 4.3). βCA2 has a second, less expressed, mRNA form that codes for a putative chloroplast protein (Table 4.1). When the cTP of the long form of βCA2 is cleaved, the mature βCA2 protein will have the same molecular weight as βCA1 (Figure 4.4). Since the peptide sequence of βCA2 has an 89% sequence identity to βCA1, it is not surprising that the βCA1 polyclonal antibody cross-reacts with βCA2 as well. This cross-reaction with βCA2 can be seen in the Western blot of Figure 4.3B. It is possible that when the primary stromal CA, βCA1, is removed from the plant, the expression of the chloroplast form of βCA2 increases to compensate for the loss of βCA1. This can explain why the removal of βCA1 has no noticeable effect on the plant. Crossing the βca1 mutant with the βca2 mutant should confirm if the faint protein band seen in the βca1 lane is the chloroplastic form of βCA2 and reveal if a βca1βca2 double mutant shows a reduced growth phenotype. The loss of the faint 25 kDa band seen in the βca1 protein sample coupled with the reduced growth of a βca1βca2 double mutant plant may hint to a chloroplastic βCA2 isoform that may be compensating for the loss of βCA1. This chapter brings to light the complex network of CA activity in the plant and how it is possible for various CA isoforms to impact multiple biochemical reactions. It is becoming clear that deciphering their role(s) in Arabidopsis will be challenging.
CHAPTER 5
THE αCAs OF ARABIDOPSIS THALIANA

INTRODUCTION

In Chlamydomonas, there are two αCA isoforms that contribute to the functioning of the alga’s CCM. The first algal αCA, CAH1, is a periplasmic CA whose gene expression is highly upregulated when Chlamydomonas is introduced to a low CO₂ environment (Fukuzawa et al., 1990; Fujiwara et al., 1990). CAH1 is thought to help facilitate Cᵢ movement into the cell from the periplasmic space (Moroney and Ynalvez, 2007). The evidence to support this comes from using the CA inhibitor, acetazolamide, during photosynthesis measurements of Chlamydomonas cultures under various pH conditions (Moroney and Tolbert, 1985). Under high pH conditions where the predominant Cᵢ molecule is HCO₃⁻, the photosynthesis rate of Chlamydomonas is decreased when acetazolamide inhibits CAH1 whereas the effect on photosynthesis is much less under acidic conditions where CO₂ is the predominant Cᵢ molecule and can freely diffuse into the cell (Moroney, Husic, and Tolbert, 1985). The growth also seems to be inhibited slightly due to the lower photosynthesis rates although this growth observation has been challenged by Van and Spalding (1999). The other algal αCA, CAH3, is a thylakoid lumen CA that relocates to the pyrenoid of Chlamydomonas when it is phosphorylated (Moroney and Ynalvez, 2007; Blanco-Rivero et al., 2012). CAH3 is thought to dehydrate HCO₃⁻ in the acidic thylakoid lumen to CO₂ for Rubisco to fix to RuBP. Chlamydomonas mutants lacking CAH3 grow very poorly in a low CO₂ environment, although grow normally in the presence of high CO₂ (Karlsson et al., 1998; Duanmu, Wang and Spalding, 2009).

There are eight αCA genes found in the Arabidopsis genome. Microarray analysis of CA expression shows that of the eight αCAs, only αCA1, αCA2, and αCA3 are expressed at a detectable level in the leaves (Fabre et al., 2007; Ferreira, Guo, and Coleman, 2008; Schmid et
Expression of $\alpha CA1$ is constant under varying concentrations of CO$_2$ (Fabre et al., 2007). Interestingly, expression of $\alpha CA2$ and $\alpha CA3$ is upregulated under low CO$_2$ conditions (Fabre et al., 2007). Of the eight $\alpha$CAs, only $\alpha CA1$ has been investigated to a degree. $\alpha CA1$ was found to move through a novel pathway leading to the chloroplast (Villarejo et al., 2005; Burén et al., 2011). $\alpha CA1$ contains a secretory pathway transit peptide that brings the protein to the ER where it is glycosylated and transported to the chloroplast (Villarejo et al., 2005; Burén et al., 2011).

In this chapter, I characterize the $\alpha$CA isoforms of Arabidopsis and attempt to find similarities between the $\alpha$CAs of Chlamydomonas and the $\alpha$CAs of Arabidopsis. RNAseq results were used to confirm the available Arabidopsis $\alpha$CA microarray data. Using promoter::GUS studies, I found that although $\alpha CA1$, $\alpha CA2$, and $\alpha CA3$ are expressed in leaves, they are not primarily expressed in mesophyll cells. Interestingly, eGFP data and protein localization prediction software indicate that $\alpha CA1$, $\alpha CA2$, and $\alpha CA3$ are found either in the plasma membrane or are secreted from the cell into the cell wall. Growth studies show mutant lines lacking either $\alpha CA1$ or $\alpha CA2$ resemble wildtype plants, so these T-DNA mutations were crossed into the $\beta$ca4 mutant lacking the plasma membrane CA. The $\alpha$ca1$\beta$ca4 and $\alpha$ca2$\beta$ca4 double mutants showed slight reductions in their overall rosette size in a low CO$_2$ environment.

RESULTS

Expression of $\alpha CA4$ genes in Arabidopsis

To confirm the EST data gathered from TAIR, RNA was extracted from root and leaf tissue of Arabidopsis plants for RNAseq analysis. Comparing the RNAseq data to the EST counts from TAIR confirm that $\alpha CA1$, $\alpha CA2$, and $\alpha CA3$ are the only three $\alpha$CAs expressed in Arabidopsis. From the RNAseq data, $\alpha CA1$ was the highest expressed $\alpha$CA in both roots and
shoots, albeit only showing moderate expression levels compared to \( \beta CA1 \) (Table 5.1). \( \alpha CA2 \) and \( \alpha CA3 \) show low expression in both roots and shoots (Table 5.1). RNA levels for \( \alpha CA4 \) through \( \alpha CA8 \) could not be quantified because of expression levels for \( \alpha CA4 \) through \( \alpha CA8 \) are too low to detect (Table 5.1).

Table 5.1. \( \alpha CA1, \alpha CA2, \) and \( \alpha CA3 \) are the only \( \alpha CAs \) expressed in Arabidopsis leaves and roots. Three Arabidopsis plants were used for leaf and root RNA samples. RNAseq reads were normalized to 100 bp counts that uniquely mapped to a CA gene in the Arabidopsis reference genome. RNAseq values are given in RPKM (Reads Per Kilo base per Million mapped). Actin1 was used as a reference to compare CA expression levels.

<table>
<thead>
<tr>
<th>GENE</th>
<th>GENE I.D.</th>
<th>ESTs(^a)</th>
<th>RNAseq Reads(^b) (Shoots)</th>
<th>RNAseq Reads(^c) (Roots)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha CA1 )</td>
<td>At3g52720</td>
<td>36</td>
<td>1,109</td>
<td>18</td>
</tr>
<tr>
<td>( \alpha CA2 )</td>
<td>At2g28210</td>
<td>2</td>
<td>248</td>
<td>531</td>
</tr>
<tr>
<td>( \alpha CA3 )</td>
<td>At5g04180</td>
<td>31</td>
<td>134</td>
<td>231</td>
</tr>
<tr>
<td>( \alpha CA4 )</td>
<td>At4g20990</td>
<td>0</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>( \alpha CA5 )</td>
<td>At1g08065</td>
<td>2</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>( \alpha CA6 )</td>
<td>At4g21000</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \alpha CA7 )</td>
<td>At1g08080</td>
<td>0</td>
<td>6</td>
<td>68</td>
</tr>
<tr>
<td>( \alpha CA8 )</td>
<td>At5g56330</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \beta CA1 )</td>
<td>At3g01500</td>
<td>1,538</td>
<td>109,333</td>
<td>662</td>
</tr>
<tr>
<td>Actin1</td>
<td>At2g37620</td>
<td>65</td>
<td>1,471</td>
<td>1,056</td>
</tr>
</tbody>
</table>

\(^a\)EST values were extracted from The Arabidopsis Information Resource webpage (TAIR).

\(^b,c\)RNAseq data generated by Oh et al. (2014)

Interestingly, the RNAseq data show that \( \alpha CA2 \) may be expressed at a higher level than the deposited EST counts for \( \alpha CA2 \) suggest. Matching the uniquely mapped RNAseq reads to the \( \alpha CA2 \) reference gene (TAIR10) suggest that the annotated \( \alpha CA2 \) gene provided by TAIR is incorrect (Figure 5.1), which corroborates an earlier study of the \( \alpha CA2 \) gene structure by Matt Brown (Brown, 2005). The gene model from TAIR indicates that \( \alpha CA2 \) contains four exons and three introns (Figure 5.1). The RNAseq data show that there is an extra exon and intron upstream of the first exon in the TAIR model as well as two extra exons and introns downstream of the 3’ exon of the TAIR model (Figure 5.1).
Figure 5.1. The RNAseq data present evidence that the \(\alpha CA2\) gene has an extra 5’ exon and two extra 3’ exons as compared to the TAIR \(\alpha CA2\) reference gene. The \(\alpha CA2\) reference gene from TAIR contains four exons and three introns represented as blue boxes and lines, respectively. The \(\alpha CA2\) reference gene also contains a 5’ and 3’ UTR represented as a red box and red arrow, respectively. The RNAseq reads mapping to the \(\alpha CA2\) genomic region, represented in red beneath the TAIR gene model, show there are seven exons in the region instead of four. RNAseq data generated by Oh et al. (2014).

**Tissue expression patterns of \(\alpha CA1\), \(\alpha CA2\), and \(\alpha CA3\)**

Promoter regions upstream of the \(\alpha CA1\), \(\alpha CA2\), and \(\alpha CA3\) genes were used to generate promoter GUS constructs to look at \(\alpha CA\) expression patterns in Arabidopsis (Figure 5.2). Arabidopsis plants stably expressing the \(p\alpha CA1::GUS\) construct show strong GUS staining in the petioles and vasculature of the rosette leaves (Figure 5.3). Expression of the \(p\alpha CA2::GUS\) construct generated a unique GUS staining pattern. There is GUS staining present in the roots and at the base of the young, developing leaves of the rosette, but the trichomes of the rosette are also stained blue (Figure 5.3 and Figure 5.4). The GUS expression in the roots and shoots of \(p\alpha CA2::GUS\) plants nicely correlates with the RNAseq data (Table 5.1 and Figure 5.3). Also, the root and trichome staining of the \(p\alpha CA2::GUS\) plants coincides with an earlier enhancer trap experiment that found \(\alpha CA2\) expression in trichomes and roots of Arabidopsis plants (Brown, 2005). Like the \(p\alpha CA1::GUS\) plants, Arabidopsis plants expressing the \(p\alpha CA3::GUS\) construct show GUS staining in the petioles which may extend into the midvein of the leaves, but they did not show prominent staining in the leaf vasculature like the \(p\alpha CA1::GUS\) plants (Figure 5.3).
Figure 5.2. Upstream promoter regions used for the (A) αCA1::GUS, (B) αCA2::GUS, and (C) αCA3::GUS constructs. Black lines represent the upstream promoter regions used for each GUS construct. Blue boxes and blue lines represent exons and introns, respectively. Red boxes represent 5’ UTRs. Red arrows represent 3’ UTRs and the orientation of genes. Purple lines represent intergenic regions.

Figure 5.3. GUS expression powered by the various αCA promoter regions. αCA1::GUS plants display GUS expression in petioles and rosette leaf vasculature. αCA2::GUS plants show GUS expression in roots and trichomes. αCA3::GUS plants display GUS expression in the petioles.
Peptide sequence of αCA3

Cloning the coding region of αCA3 into a pENTR™ vector and subsequently sequencing the cloned sequence led to the discovery of a nonsense mutation near the 3’ end of the αCA3 coding region. To confirm this point mutation, two sources of αCA3 mRNA were used for sequencing the αCA3 coding region. Both the coding region of αCA3 obtained from TAIR in the pENTR™ vector, DQ446916, as well as mRNA obtained from wildtype (COL) Arabidopsis leaves contained the 3’ transversion from a guanine to a thymine (Figure 5.5). This 3’ nonsense mutation will shorten the C-terminal end of the αCA3 peptide sequence by 19 amino acids. It is possible for the truncated αCA3 to still be a viable protein as the N-terminus of the protein is noted for its role in the stability of the protein and not its C-terminus (Aronsson et al., 1995).
The truncated αCA3 peptide sequence still contains all of its zinc coordinating residues as well as all catalytic amino acids (Figure 5.6). Also, the peptide sequences of human CA2 and αCA2 are only two and five amino acid residues longer at the C-terminus than the truncated αCA3 protein (Figure 5.6).

<p>| αCA3TAIR | AGCATCGGAAAGATAGATCCAGAGATAATTGGAATGCTACGATTACATTGGAATAC |
| DQ446916 | AGCATCGGAAAGATAGATCCAGAGATAATTGGAATGCTACGATTACATTGGAATAC |</p>
<table>
<thead>
<tr>
<th>COL</th>
<th>AGCATCGGAAAGATAGATCCAGAGATAATTGGAATGCTACGATTACATTGGAATAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>αCA3TAIR</td>
<td>AGAGGTCCTCTCAGACTCTCCTTGCACGGAGATGCTATGGACCACTCATCAAGAG</td>
</tr>
<tr>
<td>DQ446916</td>
<td>AGAGGTCCTCTCAGACTCTCCTTGCACGGAGATGCTATGGACCACTCATCAAGAG</td>
</tr>
<tr>
<td>COL</td>
<td>AGAGGTCCTCTCAGACTCTCCTTGCACGGAGATGCTATGGACCACTCATCAAGAG</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------------------------------</td>
</tr>
<tr>
<td>αCA3TAIR</td>
<td>GTGGGGACTGTTTACACGGAGCAGGAAATTGTGATTACGAGATGCTCGGTGTTAT</td>
</tr>
<tr>
<td>DQ446916</td>
<td>GTGGGGACTGTTTACACGGAGCAGGAAATTGTGATTACGAGATGCTCGGTGTTAT</td>
</tr>
<tr>
<td>COL</td>
<td>GTGGGGACTGTTTACACGGAGCAGGAAATTGTGATTACGAGATGCTCGGTGTTAT</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------------------------------</td>
</tr>
<tr>
<td>αCA3TAIR</td>
<td>GAGAAGAGCGCGAGACGACGACTGCAACCTGCTGACGAGCCTGCTGTTATTTAAACGACAG</td>
</tr>
<tr>
<td>DQ446916</td>
<td>GAGAAGAGCGCGAGACGACGACTGCAACCTGCTGACGAGCCTGCTGTTATTTAAACGACAG</td>
</tr>
<tr>
<td>COL</td>
<td>GAGAAGAGCGCGAGACGACGACTGCAACCTGCTGACGAGCCTGCTGTTATTTAAACGACAG</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------------------------------</td>
</tr>
<tr>
<td>αCA3TAIR</td>
<td>TCACGTCCAGTCTCCACTCCACGCGGTAAGATACGACGAGTCTCGGTCTTG</td>
</tr>
<tr>
<td>DQ446916</td>
<td>TCACGTCCAGTCTCCACTCCACGCGGTAAGATACGACGAGTCTCGGTCTTG</td>
</tr>
<tr>
<td>COL</td>
<td>TCACGTCCAGTCTCCACTCCACGCGGTAAGATACGACGAGTCTCGGTCTTG</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------------------------------</td>
</tr>
</tbody>
</table>

Figure 5.5. αCA3 has a point mutation that is near the 3’ end of the message. The point mutation is a transversion from a guanine to a thymine that creates an early stop codon.

Subcellular localization of αCA1, αCA2, and αCA3

αCA1, αCA2, and αCA3 are the only αCAs in Arabidopsis that are expressed at an appreciable level (Table 5.1). All three αCAs are predicted to move through the secretory pathway of Arabidopsis (Table 5.2). To confirm these predictions, the coding regions of the three αCAs were cloned upstream of a C-terminal eGFP tag to generate the 35S::αCA1-eGFP, 35S::αCA2-eGFP, and 35S::αCA3-eGFP constructs. For αCA3, the coding region upstream of
Figure 5.6. The truncated αCA3 may still be active. Black highlighted residues mark the zinc coordinating residues. Blue highlighted residues are conserved catalytic amino acids. Yellow highlighted residues are conserved proton shuttling residues.

the nonsense mutation was used. Stably transformed Arabidopsis lines were unable to produce GFP fusion proteins so transiently expressing the αCA-eGFP constructs in tobacco leaves was implemented instead. Of the three constructs, only the 35S::αCA1-eGFP construct provided a good signal in the transiently transfected tobacco leaves. In the epidermal cell layer of tobacco leaves expressing the 35S::αCA1-eGFP construct, an eGFP signal can be detected outlining each cell (Figure 5.7). Because epidermal cells have large vacuoles that push the cell contents to the
Table 5.2. Predicted subcellular locations of $\alpha$CA1, $\alpha$CA2, and $\alpha$CA3 using protein subcellular location prediction programs. Predotar, ChloroP, TargetP, and YLoc are four programs that predict the subcellular location of proteins based on amino acid sequence. A value of (1) denotes strong confidence in the prediction. ER = Endoplasmic Reticulum, SP = Secretory Pathway; (+) = contains a chloroplast transit peptide, (-) does not contain a chloroplast transit peptide.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Predotar$^1$</th>
<th>ChloroP$^2$</th>
<th>TargetP$^3$</th>
<th>YLoc$^4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$CA1</td>
<td>ER (0.99)</td>
<td>(-)</td>
<td>SP (0.99)</td>
<td>SP (0.83)</td>
</tr>
<tr>
<td>$\alpha$CA2</td>
<td>ER (0.99)</td>
<td>(-)</td>
<td>SP (0.99)</td>
<td>SP (0.49)</td>
</tr>
<tr>
<td>$\alpha$CA3</td>
<td>ER (0.99)</td>
<td>(-)</td>
<td>SP (0.68)</td>
<td>SP (0.85)</td>
</tr>
</tbody>
</table>


Figure 5.7. Transient eGFP expression assays in tobacco show $\alpha$CA1-eGFP localizes to the periphery of epidermal cells. Yellow denotes eGFP fluorescence and red denotes chlorophyll autofluorescence.

edges of cells, determining if the $\alpha$CA1-eGFP fusion protein is located in the cytosol, the plasma membrane, or is excreted into the apoplast is difficult. Further studies are required to resolve the $\alpha$CA1-eGFP localization within the plant cell.

Although preliminary transient eGFP expression analysis shows $\alpha$CA1 has a cytosol/plasma membrane/apoplast-type signal in tobacco epidermal cells (Figure 5.7) which agrees with the $\alpha$CA1 secretory transit peptide (Table 5.2), previous studies using $\alpha$CA1-eGFP fusions and immunolocalization techniques have shown $\alpha$CA1 moving through the ER en route to its final destination in the chloroplast (Villarejo et al., 2005; Burén et al., 2011). The anti-
αCA1 antibody used in Villarejo et al. (2005) was generously given to our lab to genotype our αca1 T-DNA lines and to confirm the antibody recognizes the αCA1 protein. A Western blot using the donated anti-αCA1 antibody was performed to confirm the knockout of the αCA1 gene and to confirm the reliability of the antibody. The mature Arabidopsis αCA1 peptide sequence is predicted to have a molecular weight around 30 kDa. According to a previous report, this anti-αCA1 antibody detects a glycosylated αCA1 protein around 37 kDa (Villarejo et al., 2005). When using this anti-αCA1 antibody, a band around 37 kDa can be seen in both the wildtype and αca1 mutant protein lanes (Figure 5.8). This raises a question that either the αca1 mutant lines are leaky or their antibody does not detect the αCA1 protein.

![Western blot image](image)

Figure 5.8. The αCA1 antibody does not pick up the αCA1 protein. Two prominent bands are seen with sizes near 38 kDa and 30 kDa. The mature αCA1 protein is predicted to have a weight of 30 kDa. Both bands are seen in the wildtype (COL) and αca1 mutant samples.
Genomic characterization of the αCA mutants

Mutant Arabidopsis lines containing T-DNA insertions in the αCA1 and αCA2 genes were obtained to see if the highest expressed αCAs have an effect on plant growth. An αca1 single mutant line (GK-235E05-014338) and an αca1βca4 (GK-235E05-014338; CS859392) double mutant line were graciously given to us by the laboratory of Bernard Genty. An αca2 T-DNA line (SALK_080341) was obtained from TAIR. Unfortunately, there is no T-DNA line available for αCA3. The αca1 T-DNA insertion is located near the end of the third exon of the αCA1 gene whereas the αca2 T-DNA insertion is present in the second intron of the αCA2 gene (Figure 5.9). The βca4 T-DNA insertion is the same line used from Chapter 3 (Figure 3.8 and Figure 3.9, A and B). Genomic PCR was performed to confirm the location of the T-DNA insertions in their respective genes. Gene specific primers for αCA1 and αCA2 that span the location of the T-DNA insert generate bands for the wildtype (COL) samples but not for the αca1 and αca2 DNA samples (Figure 5.10). Conversely, PCR bands are present in the αca1 and αca2 DNA samples, but not the wildtype (COL) sample when using a T-DNA specific primer with an αCA1 or αCA2 gene specific primer (Figure 5.10). cDNA samples were generated from RNA extracts of the αca1 and αca2 mutants to confirm that the T-DNA insertions were indeed disrupting transcription of the αCA1 and αCA2 genes. Using gene specific primers for αCA1 and αCA2, PCR bands representing the αCA1 and αCA2 transcripts were seen with the wildtype (COL) sample, but not with the αca1 and αca2 cDNA samples (Figure 5.11) suggesting that the T-DNA insertions are preventing transcription of the αCA1 and αCA2 genes.

Growth of the αCA knockout mutants

Much like the growth phenotype of the cytoplasmic βCA single mutants, the αCA single mutants also lack a discernible growth phenotype when grown in 200 μL L⁻¹ CO₂ in an 8 hour
Figure 5.9. Locations of T-DNA insertions in the αCA1 and αCA2 genes. The αca1 T-DNA insertion (GK235E05-014338) is located in the third exon of the αCA1 gene. The αca2 T-DNA insertion (SALK_080341) is located in the second intron of the αCA2 gene. Blue boxes and blue lines represent exons and introns, specifically. Red boxes represent 5’ UTRs and red arrows represent 3’ UTRs and the directionality of the genes. Black triangles represent T-DNA insertions. Black arrows represent locations of gene-specific primers. F and R are the forward and reverse primers used for genomic PCR and RT-PCR. I = insert primer used for genomic PCR.

Figure 5.10. The αca1 and αca2 genes are disrupted by the presence of T-DNA insertions. A genomic PCR using gene-specific primers and an insert primer shows that the αCA1 and αCA2 genes are disrupted in the αca1 and αca2 mutant lines due to the presence of T-DNA insertions in the two genes.
Figure 5.1. RT-PCR of wildtype (COL), αca1, and αca2 samples. A signal for αCA1 cDNA cannot be detected in the αca1 mutant sample. The αca2 mutant does not produce transcripts for the αCA2 gene. 3 μg of RNA was used for RT-PCR.

light/16 hour dark photoperiod with a light intensity of 120 μmol of photons m⁻² sec⁻¹ (Figure 5.12). Also, the growth of the αCA single mutants was similar to wildtype plants while growing in 200 μL L⁻¹ CO₂ for 10 weeks (Figure 5.12). Because of the predicted secretory localization for αCA1 and αCA2 and the preliminary localization data for αCA1, the αca1 and αca2 single mutants were crossed with the plasma membrane βca4 single mutant to produce the αca1βca4 and αca2βca4 double mutants. The double mutant lines grew normally in 200 μL L⁻¹ CO₂ for the first seven weeks of a growth study (Figure 5.13). As the double mutants aged, their growth
rates waned in the final three weeks of the growth study (Figure 5.14). This resulted in smaller rosettes of the double mutants by the end of the 10 week growth study (Figure 5.13 and Figure 5.14).

Figure 5.12. αca1 and αca2 single mutants grow similarly to wildtype (COL) plants in a low CO₂ (200 μL L⁻¹) environment. After 10 weeks of growth, wildtype (COL), αca1, and αca2 plants are indistinguishable from one another.

**DISCUSSION**

Unlike the βCAs and γCAs of Arabidopsis, little is known about the αCAs. The reason for the lack of interest in the αCAs is that earlier Arabidopsis expression studies show low expression profiles for the αCAs (Fabre et al., 2007; Ferreira, Guo, and Coleman, 2008; Schmid et al., 2005; Winter et al., 2007). Of the eight αCAs in Arabidopsis, only αCA1, αCA2, and αCA3 are expressed at an appreciable level (Table 5.1; Fabre et al., 2007; Ferreira, Guo, and Coleman, 2008; Schmid et al., 2005; Winter et al., 2007). Fabre et al. (2007) found that the expression profile of αCA1 is different from αCA2 and αCA3. Although αCA1 is the highest expressed αCA in Arabidopsis (Table 5.1; Fabre et al., 2007; Ferreira, Guo, and Coleman, 2008; Schmid et al., 2005; Winter et al., 2007), its expression remains constant under varying CO₂ concentrations whereas the expression levels for both αCA2 and αCA3 increase under low CO₂
Figure 5.13. Rosette areas of the double mutants are smaller than wildtype (COL) at the end of a 10 week growth period at low CO$_2$. Plants were grown in a low CO$_2$ environment in an 8 hour light/16 hour dark photoperiod at a light intensity of 120 μmol of photons m$^{-2}$ sec$^{-1}$.

Figure 5.14. αca1βca4 and αca2βca4 plants are slightly smaller than wildtype (COL) plants growing in low CO$_2$ (200 μL L$^{-1}$) at the end of a 10 week growth study.
conditions (Fabre et al., 2007). Two papers have looked at the post-translational modification and the localization process of αCA1 (Villarejo et al., 2005; Burén et al., 2011).

The aim of this chapter was to broaden what is known about the αCAs in Arabidopsis. Obtaining T-DNA mutants from the laboratory of Bernard Genty and TAIR provided a way to examine the role of the αCAs on plant growth. Much like the growth phenotype of the βCA single mutants in low CO₂ (200 µL L⁻¹ CO₂) in Chapter 3, the αca1 and αca2 single mutants were indistinguishable from wildtype (COL) plants in low CO₂ with an 8 hour day/16 hour night photoperiod at a light intensity of 120 µmol photons m⁻² sec⁻¹ (Figure 5.12). Since crossing βCA single mutants to generate βCA double mutants lacking multiple CAs in the same subcellular region led to a growth phenotype, obtaining the αca1βca4 double mutant from the lab of Bernard Genty as well as crossing the αca2 mutant with the βca4 mutant to produce the αca2βca4 double mutant was required for this study. The βca4 mutant was crossed with the αca1 and αca2 single mutants because βCA4 is found in the plasma membrane (Fabre et al., 2007; Hu et al., 2010; Hu et al., 2015) and would be close to the putative subcellular locations of αCA1 and αCA2 (Table 5.2; Figure 5.7). Also, βCA4 is expressed at a much higher level than αCA1 and αCA2, so crossing the αca1 and αca2 mutants with the βca4 mutant would give the best possible chance at seeing a growth phenotype. αca1βca4 and αca2βca4 double mutant plants grown in a short day cycle and in low CO₂ were slightly smaller than single mutant and wildtype plants at the end of a 10-week study (Figure 5.13 and Figure 5.14).

Seeing a growth phenotype with the αca1βca4 and αca2βca4 plants in low CO₂ was somewhat surprising given the low expression levels and unique tissue locations. Looking at the expression levels of the αCAs compared to βCA1 (Table 5.1), one would suspect that the αCAs are not impacting the plant much, although von Caemmerer et al. (2004) found that the majority
of CA activity must be eliminated to see an effect due to the high $K_{cat}$ of the enzyme. So it is possible that low abundance CA can have an impact on the plant. From the RNAseq data, $\alpha CA1$, $\alpha CA2$, and $\alpha CA3$ were the only $\alpha$CAs to show an appreciable level of expression in Arabidopsis leaves (Table 5.1). Interestingly, the number of RNAseq reads for $\alpha CA2$ did not match up with the $\alpha CA2$ EST counts from TAIR. Looking at the RNAseq data more closely, it was found that three exons of the $\alpha CA2$ gene were not present in the TAIR model (Figure 5.1). Since the first exon and the last two exons of the $\alpha CA2$ gene were not a part of the TAIR10 reference gene, it is possible that ESTs for those regions were not being registered for $\alpha CA2$ which could account for the contrast between the RNAseq reads and EST counts from TAIR.

Looking at the tissue expression patterns of $\alpha CA1$, $\alpha CA2$, and $\alpha CA3$, it is possible the $\alpha$CAs cannot compensate for $\beta CA4$ or vice versa. As seen in Chapter 3, $\beta CA4$ is expressed in the majority of the leaf whereas $paCA1::GUS$ plants are expressed in the leaf vasculature or cells surrounding the vasculature and in the petioles (Figure 5.3). Reports using woody plants show CO$_2$ transported in xylem can diffuse into leaf mesophyll cells and be utilized for photosynthesis (Levy et al., 1998). If $\alpha CA1$ is present in the plasma membrane (Table 5.2 and Figure 5.7) and is found in the leaf vasculature or in the surrounding cells, it is possible to propose that CA is present to help CO$_2$ in the xylem enter leaf cells.

$paCA2::GUS$ plants showed GUS staining in roots and in leaf trichomes (Figure 5.3 and Figure 5.4). Interestingly, this confirms the enhancer trap data Matt Brown previously obtained from Scott Poethig’s laboratory which found that $\alpha CA2$ was expressed in trichomes and roots (Haseloff, 1999; Brown, 2005). If $\alpha CA2$ is present in the plasma membrane, seeing GUS staining in the trichomes offers an idea that CA could be helping CO$_2$ diffuse into trichomes from the atmosphere. Trichomes contain an array of secondary metabolites including modified
fatty acids (Schilmiller, Last, and Pichersky, 2008), so the HCO$_3^-$ pool of the trichome could possibly be utilized by cytoplasmic ACCase to produce those metabolites.

In the future, a second $paCA3::GUS$ construct should be made to confirm the GUS staining seen in the petioles of $paCA3::GUS$ plants (Figure 5.3). The promoter used for the $paCA3::GUS$ construct most likely includes the promoter region of the upstream gene, At5g04190, especially since the upstream gene is oriented in the opposite direction (Figure 5.2). At5g04190 codes for Phytochrome Kinase Substrate 4 (PKS4), a protein that participates in phytochrome and phototropin signaling to regulate phototropism (Lariguet et al., 2006; Schepens et al., 2008). It would make sense that the protein would be found in the petioles of Arabidopsis which is what is seen with these $paCA3::GUS$ plants. Using a smaller promoter region to drive GUS expression should help to clarify the true tissue expression pattern of $aCA3$.

A nonsense mutation was discovered when sequencing the $aCA3$ coding region provided by TAIR in the vector, DQ446916. Both $aCA3$ sources sequenced contained the same transversion mutation from guanine to thymine (Figure 5.5). This mutation codes for a truncated protein that has 19 fewer amino acids at its C-terminal end than the predicted $aCA3$ protein (Figure 5.6). Although the truncated $aCA3$ protein still has all of the catalytic amino acids and zinc-coordinating residues as well as having a similar C-terminal length as compared to other $a$CAs (Figure 5.6), obtaining eGFP fluorescence from the truncated $aCA3$-eGFP construct was problematic. It is possible that the truncation creates an $aCA3$ protein that is degraded.

Although the protein localization programs predict a secretory pathway prediction for $aCA1$ (Table 5.2; Emanuelsson O, Nielsen H, and Von Heijne G, 1999; Emanuelsson et al., 2000; Small et al., 2004; Briesemeister, Rahnenführer, and Kohlbacher, 2010a; Briesemeister, Rahnenführer, and Kohlbacher, 2010b), it was found that $aCA1$ moves through a novel protein
localization mechanism that directs proteins to the chloroplast from the ER (Villarejo et al., 2005). αCA1-eGFP fluorescent signals were imaged in Arabidopsis chloroplasts, whereas an αCA1-eGFP construct containing the ER peptide retention signal, KDEL, remained in the ER (Villarejo et al., 2005). Using an αCA1 antibody, a protein band could be seen in the chloroplast fractions of a Western blot confirming the eGFP data (Villarejo et al., 2005). Using an anti-HA antibody confirmed the HA-tagged αCA1 localizes to the chloroplast via immuno-gold labeling as well as showing the HA-tagged αCA1 protein is post-translationally modified via Western blotting (Burén et al., 2011). To confirm αCA1 is located in chloroplasts, I attached a C-terminal eGFP tag to αCA1. Transiently expressing the αCA1-eGFP construct under the control of the 35S promoter in tobacco leaves showed a fluorescent signal near the plasma membrane of epidermal cells and not in the chloroplasts (Figure 5.7). It is possible that the periphery of the epidermal cells is fluorescing because they lack chloroplasts to which the αCA1-eGFP protein can move. But, when looking at guard cells which contain chloroplasts, αCA1-eGFP can be seen around the periphery of the guard cells and not in the chloroplasts (Figure 5.7). It is interesting to note that it is possible for the same protein to have multiple subcellular locations (Xu et al., 2013).

Since we have an αca1 mutant, the laboratory of Flor Martinez graciously donated their antibody that recognizes αCA1. Unfortunately, the protein bands from the wildtype sample matched up with the protein bands from the αca1 mutant (Figure 5.8). It is possible that the αca1 mutant has leaky expression, although the RT-PCR experiment shows that the T-DNA insertion in the αCA1 gene disrupts transcription in the αca1 mutant line (Figure 5.11). It is also possible that the protein that was detected in the αCA1 papers is not αCA1, reducing the strength of the αCA1 localization studies. Two chloroplast proteome studies favor the predicted secretory
localization of αCA1 as both proteomes failed to find αCA1 in the chloroplast (Sun et al., 2009; Ferro et al., 2010). A plasma membrane localization for the αCAs is more likely than the αCAs being secreted into the apoplast as a cell wall and apoplastic proteomics study performed on Arabidopsis rosettes did not find any of the αCAs that may have been secreted from the cell (Boudart et al., 2005). The αCAs are not highly expressed and the study only analyzed 93 cell wall and apoplastic proteins, so it is possible that the study missed the αCA proteins. An apoplast-proteomics study analyzed over 200 proteins in poplar, but also did not find an apoplastic CA (Pechanova et al., 2010).

Even though αCA1 and αCA2 supposedly move through the secretory pathway like βCA4 (Table 5.2 and Figure 5.7), αCA1 and αCA2 do not seem to be expressed in the same cell types as βCA4 (Figure 5.3). The differences in tissue expression make it less likely that the αCAs would be compensating for βCA4 or vice versa. The low CA expression of the αCAs as compared to the βCAs suggests that the αCAs play a minimal role in the plant. Strangely, even with these observations, a small reduction in growth is seen with the αca1βca4 and αca2βca4 double mutants. The growth study at low CO2 needs to be repeated, although this growth phenotype has been observed twice now for the double mutants in low CO2. Growing the double mutants, single mutants, and wildtype (COL) at high CO2 (1,000 μL L⁻¹ CO2) should be done to confirm that the CO2 concentration affects how these mutants grow. The growth of the double mutants may allude to CAs having multiple unknown functions and more work is needed to understand the possibly many roles of CAs in plants.
CHAPTER 6
IMPROVEMENTS TO C₃ PHOTOSYNTHESIS

INTRODUCTION

A recent major focus in plant biotechnology has been to make biochemical and physiological improvements to C₃ food crops. This focus has ignited a second ‘Green Revolution’ as for the second time in recent history human population will surpass global food products. In 2011, the world’s population surpassed seven billion people and mathematical models predict that by 2050 the world will be accommodating nine billion people (FAO). In order to feed the growing population, global food production must increase by 70 to 120% by 2050 (Ray et al., 2013). This is a daunting task as the allocation of sources for biofuel production, global warming, and increases in demand for poultry, fish, and dairy all strain the global food supply (Godfray et al., 2010).

Increasing the amount of arable land is one way to increase food production that had worked in the past. Currently, about 38% of the terrestrial surface of Earth is designated for agriculture (Foley et al., 2011). Unfortunately, this statistic represents the best land suited for agriculture as the rest of terrestrial Earth is comprised of deserts, mountains, tundra, cities, and other land not suitable for growing crops (Foley et al., 2011). Worldwide increases in arable land availability have decreased, causing agricultural expansion into tropical rainforests that will negatively impact the environment (Godfray et al., 2010; Foley et al., 2011; Tilman et al., 2011). Also, unsustainable land practices, global warming, and urbanization have led to a decrease in arable land in many regions (Godfray et al., 2010; Foley et al., 2011; Tilman et al., 2011). With these factors considered, it is assumed that more food will be needed to be produced from the same amount of land or less than what is currently available (Godfray et al., 2010).
Instead of cultivating new land to increase crop production, various groups have looked to improve the output of crop plants by improving their photosynthetic efficiencies. Free Air CO$_2$ Enrichment (FACE) experiments that increase the atmospheric CO$_2$ concentration around a selected area of crop plants have shown that the increased CO$_2$ concentration led to higher photosynthesis rates, which led to a 9% increase in crop biomass and food production (Leakey et al., 2009). Stephen Long and colleagues have shown by modeling that the greatest limitation to photosynthesis in plants is the conversion of CO$_2$ and H$_2$O into carbohydrates (Zhu, Long, and Ort, 2008; Zhu, Long, and Ort, 2010). To improve the photosynthetic output of crop plants, various groups are working on implementing cyanobacterial and algal CCMs into C$_3$ crops (Price, Badger, and von Caemmerer, 2011; Pengelly et al., 2014; Lin et al., 2014) as well as changing the biochemical and physiological traits of C$_3$ crops so they may be able to perform C$_4$ photosynthesis (von Caemmerer, Quick, and Furbank, 2012).

We have proposed a way to improve C$_3$ photosynthesis by adding components of an algal CCM into higher plants (Figure 6.1). This model provides the simplest method to increase the CO$_2$ concentration in the chloroplast using algal CCM components. This model includes C$_i$ transporters to increase C$_i$ concentration in the chloroplast and to bring HCO$_3^-$ into the acidic thylakoid lumen to be dehydrated to CO$_2$. Also, this model requires that CA activity be lowered in the chloroplast stroma to reduce CO$_2$ leakage out of the stroma. Lastly, the algal thylakoid lumen CA, CAH3, is needed to catalyze the dehydration of HCO$_3^-$ to CO$_2$, as the uncatalyzed dehydration of HCO$_3^-$ to CO$_2$ is too slow to support cellular functions (Raven, 1997).

The first step to implementing an algal CCM in higher plants is to determine if the selected algal CCM components transferred to higher plants can produce a viable protein that localizes to the correct sub-cellular region of the plant cell. I chose the Chlamydomonas
Figure 6.1. A plan for implementing a simplified version of the algal carbon concentrating mechanism in plant chloroplasts to improve photosynthesis rates. (1) Step one includes implementing the algal chloroplast envelope \( \text{HCO}_3^- \) transporter, Nar1.2, into the plant chloroplast envelope. This step will allow cytoplasmic \( \text{HCO}_3^- \) direct access to the chloroplast stroma, increasing the \( \text{C}_1 \) pool of the chloroplast. This step should increase the photosynthetic rate of the plant according to McGrath et al. (2014). If not, step 2 will be implemented. (2) Because the stroma is basic, \( \text{HCO}_3^- \) is the major \( \text{C}_1 \) form in the stroma. Since Rubisco fixes \( \text{CO}_2 \) and not \( \text{HCO}_3^- \) to RuBP, moving \( \text{HCO}_3^- \) to the acidic lumen will allow it to be dehydrated to \( \text{CO}_2 \) and diffuse back into the stroma for Rubisco to fix. Moving \( \text{HCO}_3^- \) to the lumen will require another \( \text{HCO}_3^- \) transporter. So far, there are no known thylakoid membrane \( \text{HCO}_3^- \) transporters. It may be possible to exchange the chloroplast membrane transit peptide of Nar1.2 with an Arabidopsis thylakoid membrane transit peptide. Also, the algal lumen CA, CAH3, will need to be added to the Arabidopsis chloroplast lumen since uncatalyzed dehydration of \( \text{HCO}_3^- \) is very slow. (3) It is possible that the stromal CA, \( \beta \text{CA}1 \), needs to be removed. If only a chloroplast membrane \( \text{HCO}_3^- \) transporter is added and \( \text{CO}_2 \) and \( \text{HCO}_3^- \) are allowed to equilibrate in the stroma, since the Rubisco active site concentration in the stroma is 200 \( \mu \text{M} \), it is possible that Rubisco can fix the increased stromal \( \text{CO}_2 \) pool before \( \text{CO}_2 \) leaks out of the chloroplast. If step 2 was implemented or if Rubisco is relocated to an added pyrenoid/carboxysome structure in the chloroplast, stromal CA activity may disrupt this simplified algal CCM in the plant chloroplast and will need to be removed. More additions to this model may be needed such as adding an algal pyrenoid or cyanobacterial carboxysome. Also, it is possible that ion transporters are needed to balance the \( \text{HCO}_3^- \) charge across various membranes.
carbonic anhydrase, CAH3, to express in Arabidopsis plants. In this chapter, I show that the native coding region of CAH3 can be expressed in Arabidopsis and that Arabidopsis recognizes the cTP of CAH3 for chloroplast localization.

RESULTS

CAH3 is a chloroplast protein in Chlamydomonas, but that does not necessarily mean the protein will localize to the chloroplast of Arabidopsis. Analyzing the peptide sequence of CAH3 using various protein localization programs show CAH3 is predicted to move to the mitochondria of plants (Table 6.1). To confirm the location of CAH3 in Arabidopsis, CAH3 was GFP-tagged. The native coding region, minus the stop codon, of CAH3 from Chlamydomonas was cloned into a pENTR vector where the CAH3 coding region was recombined into the pB7FWG2 vector upstream of the eGFP coding region. The resulting p35S::CAH3-eGFP construct was generated and stably introduced into Arabidopsis via the Agrobacterium floral dip technique. Protoplasts were generated from leaf tissue of a transformed Arabidopsis plant and were analyzed via confocal microscopy. Weak GFP fluorescence was detected in chloroplasts of Arabidopsis protoplasts (Figure 6.2), confirming that a Chlamydomonas construct can be expressed in a plant and produce a protein that localizes to the correct organelle of a plant.

Table 6.1. Predicted subcellular location of CAH3. Predotar, ChloroP, TargetP, and YLoc are four programs that predict the subcellular location of proteins based on amino acid sequence. A value of (1) denotes strong confidence in the prediction. M = Mitochondria, O = Other location; (+) = contains a chloroplast transit peptide, (-) does not contain a chloroplast transit peptide.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Predotar$^1$</th>
<th>ChloroP$^2$</th>
<th>TargetP$^3$</th>
<th>YLoc$^4,5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAH3</td>
<td>O (0.62) / M (0.35)</td>
<td>(+)</td>
<td>M (0.80)</td>
<td>M (0.99)</td>
</tr>
</tbody>
</table>

Figure 6.2. The Chlamydomonas carbonic anhydrase, CAH3, locates to the chloroplasts of Arabidopsis protoplasts. Fluorescence from the CAH3-eGFP fusion protein can be detected in chloroplasts. eGFP fluorescence is depicted as green and chlorophyll autofluorescence is depicted as red.

DISCUSSION

We have proposed a way to improve C₃ photosynthesis by adding components of an algal CCM into higher plants (Figure 6.1). Assuming the pH of the cytosol of a plant mesophyll cell is around 7.1, and the pKa of CO₂ is 6.4, under optimal conditions the cytosolic HCO₃⁻ concentration will be around 5 times higher than the cytosolic CO₂ concentration. Unfortunately, HCO₃⁻ cannot easily pass through membranes like CO₂. By adding the algal chloroplast inner envelope HCO₃⁻ transporter, Nar1.2, the HCO₃⁻ pool in the cytosol can have a direct route to the chloroplast, increasing the Cᵢ concentration of the chloroplast (Figure 6.1).

Although a larger Cᵢ concentration will be present in the chloroplast stroma as a result of the HCO₃⁻ transporter, the major Cᵢ species of the chloroplast stroma will be HCO₃⁻, a molecule Rubisco does not fix. In order to dehydrate HCO₃⁻ to CO₂ for Rubisco fixation, the algal thylakoid lumen CA, CAH3, must be present in the plant thylakoid lumen. In the presence of a CA, HCO₃⁻ in the acidic thylakoid lumen will be dehydrated to CO₂ and that CO₂ can diffuse out of the thylakoid lumen and into the stroma for Rubisco to fix to RuBP (Figure 6.1). Without the
thylakoid lumen CA, the dehydration of HCO$_3^-$ to CO$_2$ would be very slow (Raven, 1997). Lastly, the mechanism for HCO$_3^-$ entering the thylakoid lumen is poorly understood. Since HCO$_3^-$ does not easily move through membranes, it is thought that there is a HCO$_3^-$ transporter on the thylakoid membranes. A thylakoid membrane C$_i$ transporter has yet to be discovered in Chlamydomonas or any other alga with a CCM. It would be difficult to modify the transit peptide of Nar1.2 so it is redirected to the thylakoid membranes, but if no HCO$_3^-$ transporter is found, this method may be the only way to move HCO$_3^-$ into the lumen.

Stroma CA activity may possibly be problematic with the functioning of an algal CCM in plant cells. With the addition of only the algal chloroplast envelope HCO$_3^-$ transporter, Nar1.2, the removal of βCA1 may not be needed. With a working Nar1.2 present in the plant chloroplast envelope, the chloroplast C$_i$ concentration will increase. Since the Rubisco active site concentration in the stroma is 200 μM, if CO$_2$ and HCO$_3^-$ are allowed to equilibrate in the chloroplast stroma, the increased CO$_2$ concentration resulting from the increased HCO$_3^-$ concentration can still be fixed by Rubisco before leaking back out of the chloroplast.

Price and Badger (1989) found that the addition of human CA to the cytosol of cyanobacteria disrupted the working CCM of cyanobacteria. The accumulated HCO$_3^-$ within the cyanobacteria cells was dehydrated to CO$_2$ by the cytosolic human CA and leaked back out of the cyanobacteria cell before the HCO$_3^-$ could move into the carboxysome where Rubisco is located (Price and Badger, 1989). This leakage of CO$_2$ from the cyanobacteria cell led to a cell line that grew poorly in low C$_i$ conditions (Price and Badger, 1989). If a thylakoid lumen HCO$_3^-$ transporter and lumen CA are added, βCA1 may need to be removed. Adding both a thylakoid lumen HCO$_3^-$ transporter and lumen CA without a pyrenoid/carboxysome-like structure will be adding CO$_2$ to the stroma from the thylakoid lumen. CO$_2$ diffusing into the stroma from the
lumen that is not immediately fixed by Rubisco can be easily hydrated to $\text{HCO}_3^-$ by the stromal CA. The removal of $\beta\text{CA}1$ can increase the chances of Rubisco fixing CO$_2$ from the acidic lumen. If a pyrenoid/carboxysome-like structure is added to plant chloroplasts containing a $\text{HCO}_3^-$ transporter on the inner chloroplast membrane and Rubisco is relocated to this structure, $\beta\text{CA}1$ should be removed. With the addition of a pyrenoid/carboxysome-like structure to plant chloroplasts, the activity from $\beta\text{CA}1$ may cause the accumulated $\text{HCO}_3^-$ in the stroma from the added $\text{HCO}_3^-$ transporter to be dehydrated back to CO$_2$ and leak back out of the chloroplast before the accumulated C$_i$ can reach the pyrenoid/carboxysome-like structure to be fixed by Rubisco. Removal of the other major plant stromal CA, $\beta\text{CA}5$, could be detrimental to the plant, as seen in Chapter 4. Hence, with or without a pyrenoid/carboxysome structure, a small portion of CA activity must remain in the chloroplast stroma to maintain other possible anaplerotic reactions.

The Chlamydomonas thylakoid lumen CA, CAH3, is an essential component of the algal CCM (Karlsson et al., 1998). CAH3 does not have the typical expression profile of an algal CCM component as its expression level and its protein level remain fairly constant in high and low CO$_2$ conditions (Karlsson et al., 1998; Blanco-Rivero et al., 2012). Although CAH3 does not have the typical CCM expression pattern, the protein is phosphorylated when Chlamydomonas cells are introduced to a low CO$_2$ environment (Blanco-Rivero et al., 2012). This post-translational modification of CAH3 appears to partially regulate CA activity as the addition of Staurosporine, a SER/THR kinase inhibitor, reduced CA activity in the thylakoid fractions of Chlamydomonas (Blanco-Rivero et al., 2012). In addition, the post-translational modification of CAH3 in response to low CO$_2$ conditions causes the enzyme to concentrate in the pyrenoid thylakoids of Chlamydomonas to possibly dehydrate concentrated $\text{HCO}_3^-$ in the
pyrenoid for Rubisco to fix to RuBP (Blanco-Rivero et al., 2012). The Chlamydomonas mutant lacking CAH3, cia3, shows a typical CCM deficient phenotype as cia3 cells grow very poorly in a low CO₂ environment, but grow well in high CO₂ (Karlsson et al., 1998; Duanmu, Wang, and Spalding, 2009).

Since CAH3 is an important component of the algal CCM, and because of my background in CA research, I saw it fitting to use CAH3 as the starting point to begin implementing an algal CCM into a C₃ plant. A construct carrying the coding region of CAH3 fused to the 5’ end of the eGFP gene was stably introduced into the Arabidopsis genome via the Agrobacterium floral dip method. This was performed for multiple reasons. This construct would confirm that an algal CCM component could be expressed and ultimately produce a viable protein in a plant. Also, the CAH3-eGFP fusion protein would confirm if CAH3 localizes to the correct subcellular location of a plant cell. Earlier work shows that CAH3 is located to the thylakoid lumen of Chlamydomonas chloroplasts (Karlsson et al., 1998; Blanco-Rivero et al., 2012). Programs analyzing the transit peptide of CAH3 predict that it will locate to the mitochondria of plants (Table 6.1). When examining protoplasts of transformed CAH3-eGFP Arabidopsis leaves, a GFP signal could be seen coming from the chloroplasts (Figure 6.2). Although the majority of the GFP signal resonates from the Arabidopsis chloroplasts, there seems to be some GFP fluorescence outside of the chloroplasts, too (Figure 6.2). This could be due to the plant’s protein translocation machinery not fully recognizing the cTP of CAH3 correctly or CAH3 having problems entering the chloroplasts of Arabidopsis. This aberrant CAH3 localization may be problematic as foreign CA activity introduced to incorrect locations of cells has had negative effects in the past (Badger and Price, 1994). It is possible to achieve nearly 100% chloroplast localization of a foreign protein by replacing the native protein’s transit
peptide with a transit peptide of a protein in the host organism that localizes to the chloroplast (Bionda et al., 2010). Another problem with the localization of CAH3 in Arabidopsis is that we do not know for certain that CAH3 is present in the thylakoid lumen of Arabidopsis chloroplasts. Access to better instrumentation to analyze CAH3-eGFP fluorescence in the chloroplasts will provide better resolution to determine the sub-chloroplast localization of CAH3. Using transmission electron microscopy (TEM) to image immune-gold labeled samples incubated with an antibody directed against CAH3 might provide the resolution needed to determine CAH3’s sub-chloroplast localization.

The protein level of CAH3 in Arabidopsis could be higher. To increase the protein content of CAH3 in Arabidopsis, optimizing the codons of the CAH3 coding region to better suit the translation machinery of Arabidopsis might improve protein levels. The genome of Chlamydomonas is GC rich and it is possible that the codon usage for Chlamydomonas does not match the codon usage of Arabidopsis. Codons used in CAH3 may be rare in Arabidopsis, slowing down the production of CAH3 in Arabidopsis. Also, adding an artificial intron to the coding region or using the CAH3 genomic region from Chlamydomonas may improve CAH3 protein levels in Arabidopsis. Introns in pre-processed mRNA have been shown to help mRNA exit the nucleus (Reed and Hurt, 2002).

Although CAH3 was present in Arabidopsis, the growth of the plants containing CAH3 was not different from wildtype plants, which was to be expected (data not shown). There are multiple possibilities for the lack of a growth phenotype in the plants containing CAH3. One reason is that CAH3 is localizing to the chloroplast, but not to the thylakoid lumen. Because the pKa of HCO$_3^-$ is 10.3, CAH3 needs to localize to the acidic lumen in order to dehydrate HCO$_3^-$ to CO$_2$ to provide a C$_i$ source in the chloroplast Rubisco can fix to RuBP. Also, CAH3 might not
be active in Arabidopsis as post-translational modification is needed for optimal levels of CA activity in thylakoids of Chlamydomonas (Blanco-Rivero et al., 2012). It is not guaranteed that a chloroplast kinase will phosphorylate CAH3 in Arabidopsis.

The main reason there were no expectations for improved growth in the CAH3 Arabidopsis plants is because a CCM needs all three components (CA requirements, C_i transporter, pyrenoid/carboxysome) to work. Our model provides two of these components as the removal of stromal βCA1 and the addition of CAH3 satisfies the CA requirements and the addition of Nar1.2 will satisfy the C_i transporter requirement. Although those additions should improve the photosynthesis rate of Arabidopsis, other additions could be added to further increase the photosynthesis rate (McGrath and Long, 2014). It is possible that the CO_2 produced from CAH3 and the increased chloroplast C_i pool will be fixed by Rubisco since the concentration of Rubisco active sites is near 200 μM, a concentration much higher than the CO_2 concentration. But, it is also possible that increasing the CO_2 concentration in the stroma will cause increased CO_2 leakage from the chloroplast as well. The addition of a pyrenoid-like or carboxysome-like structure that can concentrate Rubisco and the increased C_i pool within a small volume will increase the chance of interaction between Rubisco and CO_2 and decrease the chances of CO_2 diffusing out of the chloroplast before Rubisco can fix it to RuBP. Unfortunately, the assembly of the algal pyrenoid is poorly understood. There is a much better understanding of the cyanobacterial carboxysome and recently it was shown that tobacco plants expressing carboxysome components can assemble carboxysome-like structures in their chloroplasts (Lin et al., 2014). One other addition to consider is to add other ion transporters to balance the negative charge of HCO_3^- entering the chloroplast stroma. A charge imbalance in the
chloroplast stroma could affect other processes and without the major stromal CA, βCA1, the pH of the stroma may also be affected.

Adding a working CCM to plants is a daunting challenge, but is a way to improve crop production to alleviate world hunger. Although transforming CAH3 into Arabidopsis plants did not improve plant growth, it is a promising start to implementing algal CCMs into plants. This work shows that plants can express algal genes and that the plant translational machinery can generate a viable protein from the algal gene. Also, this work shows that the cTP of CAH3 can be recognized by a plant and that an algal protein will localize to the correct organelle of a plant cell. More work will be needed to improve foreign gene expression as well as protein targeting for other algal CCM components. This is the first step in building an algal CCM in higher plants.
CHAPTER 7
CONCLUSIONS

CA research in plants has predominantly focused on trying to link CAs with photosynthesis. Early CA studies on C₄ plants suggest that CA activity is required for photosynthesis. Early modeling of CA activity in C₄ plants show that a CA isoform is required in the cytosol of C₄ mesophyll cells (Hatch and Burnell, 1990; Badger and Price, 1994). CA antisense lines of the C₄ plant, *Flaveria bidentis*, show that although wildtype levels of CA activity are in excess, lowering CA activity to below 30 percent of wildtype will reduce photosynthesis rates in the plant (von Caemmerer et al., 2004). Recently, CA research showed that removing CA1 and CA2 of the C₄ plant, maize, had no effect on photosynthesis, although the mutant lines did not grow as well as the wildtype line in a low CO₂ environment (Studer et al., 2014).

Early studies and mathematical modeling of CA activity in C₃ plants have shown that CA is not required for C₃ photosynthesis. The architecture of plant mesophyll cells creates a diffusion pathway for CO₂ to directly enter the mesophyll chloroplasts from the intercellular air spaces of the leaf, reducing the need of CA activity (Badger and Price, 1994; Terashima et al., 2011; Tholen et al., 2012; Tholen, Ethier, and Genty, 2014). Tobacco CA RNAi lines show that chloroplast CA activity can be greatly reduced, yet photosynthesis rates remain constant (Price et al., 1994; Williams et al., 1996; Majeau, Arnoldo, and Coleman, 1994). Interestingly, although photosynthesis rates are unchanged in the tobacco CA RNAi lines, the RNAi plants are reported to be slightly smaller than wildtype plants and have lower WUEs than wildtype plants (Price et al., 1994; Majeau, Arnoldo, and Coleman, 1994). From these studies, it seems CA activity can have other effects on the plant and not just affect photosynthesis.
After many of the early CA studies had been performed, it was discovered that plants had more CA isoforms than once thought (Moroney, Bartlett, and Samuelsson, 2001; Grigoriev et al., 2012; Kawahara et al., 2013). It is possible that multiple CA isoforms can affect photosynthesis rates in plants and the removal of just a few CAs might not be enough to affect photosynthesis since other CA isoforms can compensate for the missing CAs. With the knowledge of multiple CA isoforms in multiple subcellular and tissue locations, it is possible to consider CAs affecting multiple metabolic and physiological aspects of the plant. Since CAs catalyze the interconversion of two C\textsubscript{i} species, CO\textsubscript{2} and HCO\textsubscript{3}\textsuperscript{-}, it is possible that CAs provide C\textsubscript{i} sources to many carbon-requiring reactions within the plant. The work described in this dissertation was performed to determine the effect of CAs on multiple carbon-requiring reactions such as photosynthesis, lipid biosynthesis, amino acid biosynthesis, and nucleotide biosynthesis.

I hypothesized that cytosolic CAs may affect photosynthesis and amino acid biosynthesis. Although not playing a major role in photosynthesis, CA may play a role in lowering the amount of C\textsubscript{i} leakage from the cell by trapping C\textsubscript{i} in the cytosol as HCO\textsubscript{3}\textsuperscript{-}. CA may affect amino acid biosynthesis as PEPC in the cytosol has a much higher $K_m$ (HCO\textsubscript{3}\textsuperscript{-}) than the [HCO\textsubscript{3}\textsuperscript{-}] in the cytoplasm of a C\textsubscript{3} plant (Mukerji and Yang, 1974; Sato, Koizumi, and Yamada, 1988). From earlier work, it was known that the two cytosolic CAs in Arabidopsis are $\beta$CA2 and $\beta$CA3 (Fabre et al., 2007). I found that two mRNA forms can be made from the $\beta$CA4 gene. More importantly, I found that the two mRNA forms of $\beta$CA4 code for proteins that have two different subcellular localizations. This directly impacts the number of cytosolic CA proteins since the short $\beta$CA4 protein, $\beta$CA4.2, lost its secretory pathway leader sequence. Now, instead of $\beta$CA2 and $\beta$CA3 being the major CA isoforms in the cytosol, it is now $\beta$CA2 and $\beta$CA4.2 that are the major CA isoforms in the cytosol since $\beta$CA3 is expressed at a very low level. Although the
RNAseq data showed two distinct mRNA forms for \( \beta CA4 \) that are expressed in a tissue-specific manner, the GUS data could not fully corroborate this finding. It is possible that by trying to split up promoter regions to drive GUS expression, important enhancers and other DNA elements are removed from the various GUS construct assemblies. Databases mapping various DNA elements show that the genome is more complicated than once thought and elements that regulate a certain gene’s expression can be found in various locations upstream of a gene, downstream of a gene, or even within the gene itself or other genes (Yilmaz et al., 2011). More work is needed to understand the regulation of gene expression and pattern of expression of the two mRNA forms of \( \beta CA4 \).

Another major discovery of Chapter 3 is that a CA isoform in the same subcellular location can compensate for a missing CA isoform in the same subcellular location. An example of this is that both the \( \beta \text{ca}2 \) and \( \beta \text{ca}4 \) single mutants grow like wildtype plants in all growth conditions tested, but the \( \beta \text{ca}2\beta \text{ca}4 \) double mutant grows poorly compared to wildtype and the single mutant plants in a low CO\(_2\) environment. The atmospheric CO\(_2\) concentration has an effect on the growth of the \( \beta \text{ca}2\beta \text{ca}4 \) double mutant as the double mutant plants grow poorly in low CO\(_2\) (200 \( \mu \text{L L}^-1 \)) but grow much more like wildtype plants in high CO\(_2\) (1,000 \( \mu \text{L L}^-1 \)). Even though a CO\(_2\) effect can be seen with the double mutants, the photosynthesis rates of the \( \beta \text{ca}2\beta \text{ca}4 \) plants remained unchanged. A recent study analyzing an Arabidopsis PEPC double mutant revealed that plants lacking PEPC had a similar growth phenotype to our \( \beta \text{ca}2\beta \text{ca}4 \) double mutant (Shi et al., 2015). The PEPC double mutants are severely chlorotic with drastic reductions in growth and aspartate pools compared to wildtype (Shi et al., 2015). The PEPC study prompted us to analyze the free amino acid pools of the \( \beta \text{ca}2\beta \text{ca}4 \) double mutants. Interestingly, ASP levels in the double mutant were lower than wildtype levels while GLY and
SER levels were higher in the double mutant plants as compared to wildtype plants, similar to
the results reported by Shi et al. (2015). Analysis of the free amino acid levels of βca2βca4
plants grown at high CO₂ showed that the double mutant plants have amino acid levels similar to
wildtype plants. Although altered amino acid biosynthesis rates is one explanation for the
growth phenotype seen with the βca2βca4 plants, future metabolomics studies using the
βca2βca4 plants may help discern other roles for the cytosolic CAs in Arabidopsis.

Chloroplast CAs are suggested to play a role in photosynthesis. In Arabidopsis βCA1
and βCA5 are the chloroplast CA isoforms (Fabre et al., 2007). βCA1 comprises nearly 1% of
all leaf soluble protein (Tobin, 1970; Okabe et al., 1984; Peltier et al., 2006). βCA1 and βCA5
are both chloroplast proteins, but βCA5 is expressed at only 5% the level of βCA1. Interestingly,
it is the βca5 mutant line that possesses a very sick growth phenotype. I suggest that the growth
phenotype of the βca5 mutant is caused by a developmental problem and is not a photosynthetic
issue. This is evident by the altered phenotype of the βca5 cotyledons and lack of βCA5
expression beyond the meristematic regions of the plant. Since βCA5 is in the chloroplast, it is
possible that lipid biosynthesis is crippled due to the lack of CA activity. Growing plants on
plates supplemented with malonate should alleviate the drastic growth reduction of the βca5
mutant plants as the added malonate should bypass the ACCase carboxylation step (Baud et al.,
2004). Another experiment that should be performed is looking at the uptake of 14C acetate into
lipids of wildtype, βca1 plants, and βca5 plants. Also, examining seed and embryo development
of the βca5 heterozygous plants in ambient CO₂ conditions and βca5 homozygous plants in
elevated CO₂ conditions can hint at lipid biosynthesis problems as lipid mutants show defects in
the embryos (Baud et al., 2003). If this turns out to be a lipid synthesis mutant, the βca5 mutant
line could be a powerful tool to examine lipid synthesis in plants, as lipid synthesis mutants are
typically embryo lethal (Ohlrogge and Browse, 1995). Also, since many groups are trying to increase crop production, it would be interesting to see if overexpressing βCA5 in the wildtype background could lead to larger plants that would produce more food.

The βca5 mutant may be deficient in nucleotide biosynthesis. It is possible that reduced nucleotide levels in the plant can lead to reduced cell division and slower growth. Two enzymes in the chloroplast require HCO$_3^-$ for nucleotide biosynthesis, CPS in the pyrimidine synthesis pathway and AIR carboxylase in the purine synthesis pathway (Zrenner et al., 2006). If either CPS activity or AIR carboxylase activity is reduced due to a reduction in HCO$_3^-$ substrate, nucleotide biosynthesis may be reduced which could cause the growth phenotype of βca5. If nucleotide biosynthesis is reduced, cell division should be noticeably affected in the plant tissues. Adding propidium iodide that stains plant cell walls and imaging wildtype, βca1, and βca5 plant tissues can help determine if cells are not elongating, which would point towards a lipid synthesis problem, or if cells are not dividing, which could be either a lipid problem or a nucleotide source problem. Growing wildtype, βca1, and βca5 plants in a high CO$_2$ environment and then moving the plants to a low CO$_2$ environment to subsequently collect samples in a time dependent manner for metabalomic analysis could determine what is affected by the lack of βCA5 as the pathway most affected should show the earliest signs from movement into low CO$_2$. If CA is linked to nucleotide biosynthesis, the βca5 mutant line could be another tool used to study the cell cycle and possibly be used for cancer research as there are many CA cancer studies found in the literature.

αCAs have not been studied much due to their low expression levels in plants. I studied αCA1, αCA2, and αCA3 because these are the highest expressed αCAs in Arabidopsis. Since αCA3 does not have a T-DNA mutant line, the αCA3 isoform was not as high of a priority as
αCA1 and αCA2. αCA1 is thought to be a chloroplast protein (Villarejo et al., 2005; Burén et al., 2011). From my preliminary αCA1-eGFP transient expression studies in tobacco, αCA1 was not localized to chloroplasts, but to the perimeter of epidermal cells and guard cells. Our lab was graciously given the anti-αCA1 antibody used in the Villarejo et al. (2005) study to determine if our αca1 mutant lacked the αCA1 protein. Unfortunately, the protein banding pattern from the Western blot using the anti-αca1 antibody was the same in both the wildtype and αca1 protein lanes. It is possible that the αca1 mutant has leaky expression of αCA1, but no αCA1 signal can be detected from the αca1 RNA samples in RT-PCR experiments. Also, multiple chloroplast proteomics studies could not find the αCA1 protein in their chloroplast preps (Sun et al., 2009; Ferro et al., 2010). On the other hand, an apoplast proteome study also could not detect the αCA1 protein either (Boudart et al., 2005). But, it is possible that the protein is found in the plasma membrane of cells. Generating stable Arabidopsis αCA1-eGFP lines will help determine where the αCA1 protein is localizing.

Interestingly, αCA2 was found to be expressed in trichomes and roots, confirming earlier work by Matt Brown (Brown, 2005). It is possible that a CA is needed in trichomes to maintain a C1 source for lipid synthesis or nucleotide synthesis, as trichomes greatly expand as single cells as well as undergo endoreplication (Kasili et al., 2010).

Why do plants need so many CAs and what are they all doing? We now know that many of the CA isoforms are expressed in a tissue specific manner in Arabidopsis and that many of the CA isoforms go to different locations within the cell. With that knowledge, it is easy to see that the various CA isoforms can have different roles throughout the plant and that all the isoforms are not present to regulate only a single plant function. The work presented in this dissertation provides the foundation in trying to understand the many roles of CA in plants and also provides
an abundant collection of CA knockout lines to test if CAs affect other carbon-requiring reactions in the plant. There is a reason CAs account for one to two percent of total soluble protein in the leaf (Tobin, 1970; Okabe et al., 1984; Peltier et al., 2005), the roles and network of CAs throughout the plant are just more complex than initially thought before.
REFERENCES


135


Konishi T, Shinohara K, Yamada K, Sasaki Y (1996) Acetyl-CoA carboxylase in higher plants: most plants other than gramineae have both the prokaryotic and the eukaryotic forms of this enzyme. Plant and Cell Physiology 37: 117-122.


Meldrum N, Roughton F (1932) Some properties of carbonic anhydrase, the CO2 enzyme present in blood. J Physiol 75: 15.


### APPENDIX A

**LIST OF PRIMERS FOR GENOMIC PCR AND RT-PCR**

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>αCA1 F</td>
<td>5’- TTG CTG TGG TGG CAA GTC TC - 3’</td>
</tr>
<tr>
<td>αCA1 R</td>
<td>5’- GTT GAC CAC ACG GTC TTG AAT TGT TCT TG - 3’</td>
</tr>
<tr>
<td>αCA1 (8-4) F</td>
<td>5’- ACA CAG CAC AAG TGG AAG TAG GAA - 3’</td>
</tr>
<tr>
<td>αCA1 (8-4) R</td>
<td>5’- AAT TGT TCT TGA AAG AAG TGT CCA A - 3’</td>
</tr>
<tr>
<td>(α1-F) αCA1 GK F</td>
<td>5’- GAA GGA GTA GTG TTT GGA TAT AAA GG - 3’</td>
</tr>
<tr>
<td>(α1-R) αCA1 GK R</td>
<td>5’- GTG CCG ATT TTG AAG AGA CTT G - 3’</td>
</tr>
<tr>
<td>αCA1 GK F</td>
<td>5’- GAC ACT TAA ACC CTC ACT TCA CCA CAT - 3’</td>
</tr>
<tr>
<td>αCA1 GK R</td>
<td>5’- CTT CCT TTA GCT TCA CCA ATT TCT CCT TC - 3’</td>
</tr>
<tr>
<td>(α2-F) MH αCA2 F</td>
<td>5’- TCC ATT GAC CGT AAT ACT CCC TGA C - 3’</td>
</tr>
<tr>
<td>(α2-R) MH αCA2 R</td>
<td>5’- AAA TTC AAG TCC CCA AAG CTC AAA A - 3’</td>
</tr>
<tr>
<td>(β1-F) βCA1 (9-4) F</td>
<td>5’- AAA TAC GGT GGC GTT GGA GC - 3’</td>
</tr>
<tr>
<td>(β1-R) βCA1 (9-4) R</td>
<td>5’- AAA TTC AAG TCC CCA AAG CTC AAA A - 3’</td>
</tr>
<tr>
<td>(β2-F) βCA2 F new</td>
<td>5’- TGT GCC CAT CAC ACG TAC ACT - 3’</td>
</tr>
<tr>
<td>(β2-R) βCA2 (9-6) R</td>
<td>5’- GTC AAG AGG AAA TGA CAT AAG CCC C - 3’</td>
</tr>
<tr>
<td>(β4-F) 17</td>
<td>5’- GCG AAG ATC AAA GCG TTG ACG - 3’</td>
</tr>
<tr>
<td>(β4-R) 18</td>
<td>5’- CGC CAA CTC CAG AGT GTC T - 3’</td>
</tr>
<tr>
<td>(β4-I) 19</td>
<td>5’- AAC GTC CGC AAT GTG TTA TTA AGT TGT C - 3’</td>
</tr>
<tr>
<td>(β5-F) βCA5 (10-8) F</td>
<td>5’- AAC ACA TCT TAG TCA TTG GTC ATA GCC - 3’</td>
</tr>
<tr>
<td>(β5-R) βCA5 (10-8) R</td>
<td>5’- AAC ACA TCT TAG TCA TTG GTC ATA GCC - 3’</td>
</tr>
<tr>
<td>LB-OURS a</td>
<td>5’- CCA AAC TGG AAC AAC ACT CAA CCC - 3’</td>
</tr>
<tr>
<td>RB-OURS b</td>
<td>5’- CTT AGC TCA TTA AAC TCC AGA AAC CC - 3’</td>
</tr>
<tr>
<td>GK INSERT LB a</td>
<td>5’- GAT GAA ATG GGT ATC TGG GAA TGG - 3’</td>
</tr>
<tr>
<td>GK INSERT LB-R</td>
<td>5’- CTA CTT TAC CCA TAG ACC CTT ACC - 3’</td>
</tr>
<tr>
<td>GK INSERT RB</td>
<td>5’- CAC TTC TAT CAC CTT TTC CCA CCA C - 3’</td>
</tr>
<tr>
<td>GK INSERT RB-R</td>
<td>5’- GTG AAG ATA GTG GAA AAG GAA GGT GG - 3’</td>
</tr>
<tr>
<td>GFP F</td>
<td>5’- ATG GTG AGC AAG GGC GAG G - 3’</td>
</tr>
<tr>
<td>GFP R</td>
<td>5’- ATT TGG TAC AGC TCG TCC ATG CC - 3’</td>
</tr>
<tr>
<td>GUS F</td>
<td>5’- ATG TTA CGT CCT GTA GAA ACC CCA ACC C - 3’</td>
</tr>
<tr>
<td>GUS R</td>
<td>5’- TCA TGG TTT GCC TCC CTG CTG C - 3’</td>
</tr>
<tr>
<td>2x35S F</td>
<td>5’- GGT CAA CAT GGT GGA CGA CGA C - 3’</td>
</tr>
<tr>
<td>2x35S R</td>
<td>5’- GGG ATC CTC TAG AGT CGA GGT C - 3’</td>
</tr>
<tr>
<td>Actin F</td>
<td>5’- CAT CTA TGA TGG GTA TGG AGA ATG CTG G - 3’</td>
</tr>
<tr>
<td>Actin R</td>
<td>5’- CCT TTG CGA TCC ACA TCT GCT - 3’</td>
</tr>
</tbody>
</table>

*a Insert primer used for the SALK T-DNA insert in the αCA2, βCA1, βCA2, and βCA5 genes.
*b Insert primer used for the GK T-DNA insert in the αCA1 gene.
## APPENDIX B

### LIST OF CLONING PRIMERS

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAH3-GFP-F</td>
<td>5'-CAC CAT GCG TCT AGC CGT TCT ACA A - 3'</td>
</tr>
<tr>
<td>CAH3-GFP-R</td>
<td>5'-CAG CTC GTA TTC GAC CAG CGG - 3'</td>
</tr>
<tr>
<td>α1-GFP-F</td>
<td>5'-CAC CCA GGA ACT AAA ACA CAA GAT GAA GAT T - 3'</td>
</tr>
<tr>
<td>α1-GFP-R</td>
<td>5'-ATT GGG TTT TTT CTT TTT GGT TTC TTT TTT ATC - 3'</td>
</tr>
<tr>
<td>pα1-GUS-F</td>
<td>5'-CAC CGA AAG ACT TCC ACT AAG CTG TG - 3'</td>
</tr>
<tr>
<td>pα1-GUS-R</td>
<td>5'-CTT GTG TTT TAG TCC TGT TGT TGG AAA TCT CGT - 3'</td>
</tr>
<tr>
<td>α2-GFP-F</td>
<td>5'-CAC CCA AAC ATT TAT CTC CAA TCA TCA CCT CTC T - 3'</td>
</tr>
<tr>
<td>α2-GFP-R</td>
<td>5'-TAG TGA TTT TGG TTT GTA TAA CTT TAC CAC - 3'</td>
</tr>
<tr>
<td>pα2-GUS-F</td>
<td>5'-CAC CTA TTT CTA TTG AAT GAT TTT TGG ACA C - 3'</td>
</tr>
<tr>
<td>pα2-GUS-R</td>
<td>5'-GTG TTT CTA TTA CAA TTA CAA GAG CGG - 3'</td>
</tr>
<tr>
<td>α3-GFP-F</td>
<td>5'-CAC CAT GAA AAC CAT TAT CCT TTT TGT TGT AAC ATT TCT T - 3'</td>
</tr>
<tr>
<td>α3-GFP-R</td>
<td>5'-GTT TAA ATA AAC CAG ACG TCC GTT CAG AGG - 3'</td>
</tr>
<tr>
<td>pα3-GUS-F</td>
<td>5'-CAC CTT TGG TCT ATA GTC TCT CTA TGA - 3'</td>
</tr>
<tr>
<td>pα3-GUS-R</td>
<td>5'-CTT TGG TTT TTT ATT TTT TTT TTT TTT TTT ATC - 3'</td>
</tr>
<tr>
<td>β1-GFP-F</td>
<td>5'-CAC CAT GTC GAC CGC TCC TCT C - 3'</td>
</tr>
<tr>
<td>β1-GFP-R</td>
<td>5'-TAC AGA GCT AGT TTC GGA GAG GCC AAA - 3'</td>
</tr>
<tr>
<td>βp1-GUS-F</td>
<td>5'-CAC CGC GTG TGC CCA CTT TAC ATA TAA A - 3'</td>
</tr>
<tr>
<td>βp1-GUS-R</td>
<td>5'-TTG TGG CGC TCC AAG AGA AGG CGG GAT - 3'</td>
</tr>
<tr>
<td>β2-GFP-F</td>
<td>5'-CAC CAT GGG AAA CGA ATC ATA T - 3'</td>
</tr>
<tr>
<td>β2-GFP-R</td>
<td>5'-TAT AGA ATG AAC GGG GGA - 3'</td>
</tr>
<tr>
<td>β2-OE-R</td>
<td>5'-TCA TAT AGA ATG AAC GGG GGA - 3'</td>
</tr>
<tr>
<td>βp2-GUS-F</td>
<td>5'-CAC CGA AAT GGC TCA TCA GAC TCA GAG A - 3'</td>
</tr>
<tr>
<td>βp2-GUS-R</td>
<td>5'-TTC TCA GAT CTG TGC AAG TCA TGG - 3'</td>
</tr>
<tr>
<td>βp3-GUS-F</td>
<td>5'-CAC CCG TCA CTA TGT CAA TAA ATC TGT AAT AAT A - 3'</td>
</tr>
<tr>
<td>βp3-GUS-R</td>
<td>5'-ATC TTT TGG TCT CAA TAA CAC AAT AAC AC - 3'</td>
</tr>
<tr>
<td>β4.1-GFP-F</td>
<td>5'-CAC CAT GGC TCC TGC ATT CGG AAA A - 3'</td>
</tr>
<tr>
<td>β4.1-GFP-R</td>
<td>5'-AGA GAA GGC AAA AGA AGG AGT GGT C - 3'</td>
</tr>
<tr>
<td>βp4.1-GUS-F</td>
<td>5'-CAC CGA ATC AGA ACG GAA CTT ATA AAG ATG ATG A - 3'</td>
</tr>
<tr>
<td>βp4.1-GUS-R</td>
<td>5'-TAT AAA TGG TCT TCT TGT TCC TCT ATC TCA CCA ACC - 3'</td>
</tr>
<tr>
<td>βp4.1(v-2)-GUS-R</td>
<td>5'-AGC AAT ATG GGG GGT TCT GGT AGG GGA TAG - 3'</td>
</tr>
<tr>
<td>βp4.2-GUS-F</td>
<td>5'-CAC CTG TGG CAA CGG GAA GGG CAA CTG AAT - 3'</td>
</tr>
<tr>
<td>βp4.2-GUS-R</td>
<td>5'-TGT GTC CCA AAT ATG ACC ACG TGG CTA TAC AAT ACT A - 3'</td>
</tr>
<tr>
<td>βp4.2(v-2)-GUS-F</td>
<td>5'-CAC CTT TGT GCC TTA TTG ATC CAG ATT C - 3'</td>
</tr>
<tr>
<td>β5-GFP-F</td>
<td>5'-AGA CCA AAC TGA CGG GTC TTT AAC AAG AGT TAC - 3'</td>
</tr>
<tr>
<td>β5-GFP-R</td>
<td>5'-CAC CGA AAT ACC ATG GAA ATT TTA ATT ATG TTT - 3'</td>
</tr>
<tr>
<td>βp5-GUS-F</td>
<td>5'-CAC CGA AAT ACC ATG GAA ATT TTA ATT ATG TTT - 3'</td>
</tr>
<tr>
<td>βp5-GUS-R</td>
<td>5'-GAA AAA GGT CTA GAG AGT AAA GGT GGG TAC T - 3'</td>
</tr>
</tbody>
</table>
APPENDIX C
PERMISSION TO REPRINT

Dear Robert:

ASPB allows authors of works published in Plant Physiology to freely reuse their content for any purpose without written permission from the Society/publisher.

However, per your request, I am sending this message granting you permission to freely reuse your images and text in your dissertation. I understand that your institution will be making your dissertation openly accessible. That's fine.

Congratulations on completing your dissertation and best of luck. Thank you for publishing with us.

Nancy Winchester
Director of Publications

Plant Physiology permission information: http://www.aspb.org/publications/permission.cfm

Permission to Use Content from Plant Physiology® and The Plant Cell

Permission to make digital or hard copies of part or all of a work published in Plant Physiology® or The Plant Cell is granted without fee for personal or classroom use, provided that copies are not made or distributed for profit or commercial advantage and that copies bear the full citation, the journal URL (www.plantphysiol.org or www.plantcell.org), and the following notice: "Copyright American Society of Plant Biologists." Please request permission in writing to reproduce material if the use is commercial or if you wish to make multiple copies other than for educational purposes. The Copyright Clearance Center is the authorized agent of ASPB for permission requests. Contact the Copyright Clearance Center at

Copyright Clearance Center Inc.
Re: Plant Physiology®/The Plant Cell
222 Rosewood Drive
Danvers, MA 01923 USA
Voice: 978-750-8400
Fax: 978-750-4470
Internet: http://www.copyright.com/

To Our Authors:

ASPB grants to authors whose work has been published in Plant Physiology® or The Plant Cell the royalty-free right to reuse images, portions of an article, or full articles in any book, book chapter, or journal article of which the author is the author or editor. Reproductions must bear the full citation, the journal URL (www.plantphysiol.org or www.plantcell.org), and the following notice: "Copyright American Society of Plant Biologists." ASPB further grants to authors the permission to make digital or hard copies of part or all of a work published in Plant Physiology® or The Plant Cell without fee for personal or classroom use.
To Whom It May Concern,

I hereby grant Robert DiMario permission from his published first-authored work as part of his Ph.D. thesis.

Michael E. Hellberg
Director of Graduate Studies

Mr. DiMario,

As you know our Department requires our doctoral students to have a publication from their dissertation. You certainly have permission to use the publication as part of your dissertation.

Joseph F. Siebenaller
VITA

Robert DiMario attended Louisiana State University for his B.S. degree. While in college, he asked to join Dr. James Moroney’s laboratory to earn research credit. His project was to genotype various Arabidopsis carbonic anhydrase T-DNA mutant lines. Robert earned his B.S. in spring of 2010. Robert continued his work in Dr. Moroney’s lab when he entered the graduate program at Louisiana State University. His graduate work included characterizing the roles of α- and β-carbonic anhydrases in Arabidopsis. Robert is a candidate to receive his PhD in Biological Sciences in spring of 2016.