4-2001

THE EXAMINATION OF THE EXPRESSION PATTERNS OF A GAMMA CARBONIC ANHYDRASE HOMOLOGUE IN ARABIDOPSIS THALIANA

Eric Thomas Dane Nielsen III

THE EXAMINATION OF THE EXPRESSION PATTERNS OF A GAMMA CARBONIC ANHYDRASE HOMOLOGUE IN *ARABIDOPSIS THALIANA*

A Thesis

Submitted to the Honors College of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for
Upper Level Honors

in

Biochemistry

by
Eric Thomas Dane Nielsen III
April 2001

List of Figures

Abstract

A distinct cDNA encoding a gamma carbonic anhydrase (γ-CA2) was isolated from an *Arabidopsis thaliana* total leaf RNA extract by RT-PCR. γ-CA2 is the second gamma carbonic anhydrase isloated in *Arabidopsis.* γ-CA2 encodes a 278 amino acid polypeptide that is very similar to the previously isolated *Arabidopsis* γ-CA (Xi, 1999). Their derived amino acid sequences exhibit 75% identity.

A promoter-GUS construct containing the first 21 amino acids of γ-CA2 was transformed into *Arabidopsis* for studying gene expression of γ-CA2. The pattern of expression was centralized at the immature leaves with limited expression radiating out a short distance into the petioles of the mature leaves and cotyledons and down the hypocotyl. This expression pattern was unlike that of the γ-CA deduced by Xi (1999) using the same method. γ-CA2 appears to be a cytosolic protein. The subcellular areas of expression were visualized as concentrated granules in the cytoplasm. However, microscopic analysis of the transgenic plants revealed that the GUS product is associated with organelles, possibly plastids. Possibly, the amino termini of plant γ-CAs mediate interactions between these proteins and organelles or proteins associated with organelles.

Literature Review

Carbonic anhydrase (CA), which was first discovered in red blood cells, is a zinc metalloenzyme catalyzing the reversible hydration of CO_2. CA is ubiquintous among living organisms and is important in a variety of functions, including facilitated diffusion of CO_2, interconversion of CO_2 and HCO_3^- during photosynthesis and biosynthetic processes, pH homeostasis, and ion transport (Alber *et al.*, 1994). There are three families of CA's with no primary sequence similarity: α-, β-, and γ-. The evolution of the independent families appears to have converged to form zinc metalloenzymes, and all appear to share a similar catalytic mechanism (Lindskog, 1997).

Originally, the families were also thought to be distinct in phylogenetic origin. α-CAs were expressed only in the animal kingdom, while β–CAs were expressed in plants and eubacteria, and γ–CA expression was limited to archaebacteria and eubacteria. However, further studies have shown that CA families are found in organisms across these boundaries. Plants appear to have all three types of CAs. In the *Arabidopsis thaliana* genome at least 14 potential genes encoding CAs, with all three families represented, have been found (Moroney *et al.*, 2001). However, the roles of α-CAs and γ-CAs in higher plants are still not clear. Cyanobacteria have both α-CA and β-CA and a γ-CA-like protein CcmM. In the green alga, *Chlamydomonas reinhardtii*, examples of α-CA and β-CA are represented (Samuelsson and Karlsson, 2000). However, in animals only α-CAs have been found, although there are multiple isoforms of α-CAs in humans.

α-CAs are represented in animals, plants, eubacteria, and viruses making it the most widely distributed family of CAs. The structure of most α-CAs is an active monomer of about 30 kDa.

The active site is a funnel shaped crater with the zinc ion near the bottom. X-ray crystallography shows that the side chains of three histidines and a hydroxyl ion at the active site coordinate the zinc ligand. The structure of the active site is highly conserved among α-CAs.

Although most research involving α-CAs is done in higher animals, such as the mitochondrial α-CA (CAVI) of liver cells, recently researchers have become aware of α-CAs in eubacteria, plants, and algae. Only a few α-CA proteins have been identified in photosynthetic plants. One α-CA in *Arabidopsis* (ACAH1) was identified by Susanne Larsson. Localization by immunogold experiments suggests an association of ACAH1 with the chloroplast stroma. The lysine rich C-terminus may bind to the membrane or a protein of the stroma. In the cyanobacteria *Anabaena* and *Synechococcus* one α-CA has been identified and localized to the periplasmic space. It may function as part of the CCM (CO_2 concentrating mechanism) to replenish the external CO_2 from the HCO_3^- supply as CO_2 enters the cell or it may facilitate the diffusion of CO_2 across the plasma membrane (Soltes-Rak *et al.*, 1997).

In *Chlamydonas reinhardtii*, a photosynthetic green alga, three α-CAs were identified by Karlsson (1998). Two of the α-CAs, Cah1 and Cah2, were localized in the periplasmic space based on physiological experiments using CA inhibitors. Both were very similar in sequence and structure, suggesting that the gene encoding Cah2 arose from duplication of the gene encoding Cah1. However, the two genes are differentially regulated by varying levels of CO_2 in the growth environment of *C. reihardtii.* Cah1 is expressed under low CO_2 but not elevated CO_2 conditions, while Cah2 is poorly expressed under low levels of CO_2 that only slightly increase at elevated levels of CO_2 (Fujiwara, 1990). Cah1 may have similar function to the periplasmic α-CAs of cyanobacteria. However, the function of Cah2 is unresolved. Some biochemical and sequence evidence shows that the third α-CA in *C. reinhardtii*, Cah3, is localized to the thylakoid membrane

(Karlsson 1998). Cah3 is needed for cell growth in air levels of CO_2. Interestingly, no other photosynthetic plants have been shown to contain CAs associated with the thylakoid membrane, suggesting that *C. reihardtii* may be unique in this respect.

β-CA has not been found in animals but has been identified in most photosynthetic organisms, some bacteria, and yeast. The enzyme from higher plants is an active octamer made of dimers that bind to form two tetramers that associate to form the active octamer. EXAFS studies of the structure of a β-CA in spinach describe the conserved residues of the active site (Bracey *et al.*, 1994). One histidine and two cysteines coordinate the zinc ion in the hydrophobic active sites located at the subunit interfaces. Studies have also shown that the amino acid sequence of the β-CA active site is a mirror image of α-CA's (Kimber & Pai, 2000).

The highest activity of β-CA is found in the chloroplast stroma of C3 plants. β-CA comprises two percent of the total leaf protein in mesophyll cells. At least five β-CA encoding sequences are in the *Arabidopsis* genome database. There are an assortment of roles suggested for β-CA in the chloroplast stroma: modulate the pH of the stroma, facilitate diffusion of CO_2 across the chloroplast membrane, or replenish the CO_2 supply in the chloroplast stroma from HCO_3^-.

It would appear that β-CA plays an important role in CO_2 fixation when CO_2 is limited. Price *et al.* (1994) made transgenic plants containing antisense CA constructs. The leaves of tobacco plants containing the antisense β-CA gene have only about 10% or less CA activity than the wild type plants. Yet, no deleterious effects on the transgenic plants resulted from the reduction in CA activity. However, in another experiment by Kim (1997) on transgenic *Arabidopsis* containing a β-CA antisense construct, plants with less than 10% activity could not grow on media lacking sucrose unless they were provided with high levels of CO_2. These results

suggest β-CA does have a role in CO_2 fixation, and bring to question the inconsistence of results between Price *et al.* (1994) and Kim (1997).

Along with photosynthetic organisms, β-CA has been found in eubacteria, archaebacteria and yeast. A β-CA has been identified and localized in the carboxysome of the cyanobacterium *Synechocystis* (Fukuzawa, 1992). Carboxysomes house the abundance of Rubisco in these organisms. When activity of the carboxysomal CA was lost, a high CO_2 requiring phenotype resulted (Price & Badger, 1989). Cells without β-CA still accumulated HCO_3^- but could not rapidly convert the HCO_3^- to CO_2 for fixation.

Eriksson *et al.* (1996) discovered two β-CAs in the mitochondria of *C. reinhardtii*. They found that the two β-CAs differed by only one amino acid in the coding region, and the polypeptides were not synthesized under high CO_2 levels but was strongly induced at air (lower) levels of CO_2. They suggested that the β-CAs functions to buffer the mitochondrial matrix and protect it from adverse effects of NH_3 produced in photorespiration or to decrease the leakage of inorganic carbon back out of the matrix.

In corn, a C4 plant, Ku, Kano-Murakami & Matsuoka (1996) found a β-CA localized in the cytosol of mesophyll cells. This β-CA provides HCO_3^- for PEP carboxylase. This process is essential for photosynthesis in C4 plants, and the presence of β-CA in the cytoplasm of mesophyll cells accelerates the rate of photosynthesis by an estimated 10^4 times (Badger and Price, 1994).

The third family of CAs, the γ-CAs, were discovered when Alber and Ferry (1994) characterized a CA from the archaebacterium *Methanosarcina thermophila*. It was not homologous with either the α-CA or β-CA. The structure of γ-CA, which was solved by Kisker *et al.* (1996) by crystallography, was much different from the α- and β-CAs. The active form is made of a trimer of three identical monomers. Each monomer is dominated by a left-handed β-helix

consisting of three parallel sheets. The active trimer contains three zinc ligands, one at each subunit interface. Like α-CA, three histidines and a water molecule coodinate the zinc into the active site. However the histidines are supplied by two subunits. In *M. thermophilia*, His 81 and His 122 from one subunit and His 117 from the adjacent subunit act as the three ligands on the γ-CA protein. Although the active sites are at the subunit interfaces, their architecture is similar to that of the α-CA monomer.

The coding sequence of the γ-CA characterized from *M. thermophilia* contained an N-terminal sequence of 34 amino acids similar to a signal peptide in secretory proteins. This evidence suggested that the γ-CA is expressed in the periplasmic space. Studies of *M. thermophilia* γ-CA activity by Albers and Ferry (1994) showed that cells grown on a acetate media had higher activity than cells grown on a methanol, suggesting that γ-CA could facilitate the removal of excess CO_2 produced by metabolism of acetate. However the precise role of this γ-CA is not known.

Genes encoding proteins similar to γ-CA have been found in eubacteria and plants. Searches of *Arabidopsis* genome database for coding sequences homologous to *M. thermophilia* γ-CA have produced three matching genomic sequences, and a search of the EST database reveals that all three genes are expressed. Two γ-CAs genes are located on chromosome I, and the third is located on chromosome V (Moroney *et al.*, 2001). A distinct characteristic of the *Arabidopsis* γ-CAs is that they have extensions at the N- and C-terminus not required for catalytic activity while the *M. thermophilia* γ-Ca does not. Analysis of the first 50 amino acids at the N-terminus in γ-CA1 of *Arabidopsis* (Xi, 1999) revealed that its secondary structure is largely α-helix and could be amphipathic with a slight positive charge. These properties are characteristic of transit proteins that target an organelle in the cytosol. The C-terminal extension of γ-CA1 possesses an overall negative charge. However the functions of the N- and C-terminal extensions are unknown. The *Arabidopsis*

γ-CA which is profiled in this research, identified as γ-CA2, is located on chromosome I (protein ID# AAF98404.1) and has a N-terminal extension similar to γ-CA1. Homologues to the *Arabidobsis* γ-CAs have been found in monocots, dicots, and green algae. However, enzymatic activity of γ-CAs has not been directly measured in the higher plants nor has function been pinpointed.

Methods and Materials

Isolation of total DNA from *Arabidopsis thaliana* leaf.

About 10mg of leaf tissue from *Arabidopsis thaliana* ecotype Columbia were excised, immediately frozen in N_2, and stored at –80C in a microfuge tube. The frozen tissue was ground for 15 sec in the tube and 400 ul of Extraction Buffer (200mM Tris pH 7.5, 250mM NaCl, 25mM EDTA, 0.5% SDS) was added. The mixture was vortexed for 5 sec and sat for 5 min at room temperature (RT) before centrifugation at 12,000 rpm for 3 min at RT. 300 ul of supernatant was collected and mixed with 300 ul of isopropanol in a new tube. After sitting for 2 min at RT and centrifugation for 5 min at 12,000rpm, the supernatant was removed and the pellet was washed with 70% ethanol. The clean pellet was air dried and resuspended in 50 ul of TE. Approximately 40ng of total leaf DNA was used for each PCR.

Amplification of the γ–CA2 promoter by PCR.

The total leaf DNA prepared above was used as the template for amplification of the γ-CA2 promoter. The 5' primer (AGAGAGAGACCTTGAACGGCATTA) and 3' primer (GCTTGACCAGTCTCAC GGATC) were synthesized by Genelab at Louisiana State University. These primers flanked about 1.6 kb of the DNA region upstream of the AUG start codon as well as the first 21 codons of the open reading frame. The PCR reaction contained 1X PCR buffer, approximately 40 ng of template DNA, 200 uM dNTPs, 2 uM primers, and 2 units of Vent polymerase in 25 ul total volume. The γ-CA2 promoter was amplified for 35 cycles using the following parameters: 94C for 30 sec, 66C for 30 sec, and 72C for 1 min for the polymerization cycle. The DNA yield was assessed by electrophoresis through a 1% agarose gel in 1X TBE (90 mM Tris-borate, 2 mM EDTA, pH 8.0). A promoter stock of the identical PCR reaction mixture was made for use in later experiments.

Amplification of the γ–CA2 cDNA by RT-PCR.

A sample of total leaf RNA was used as the template for RT-PCR. The 5' primer (TCAAACCAGAGACCGAAACATA) and 3' primer (GCTCGAAGAAAGTGAC GATCAAT) were synthesized by Genelab at Louisiana State University. These primers flanked the 823 bp open reading frame. The PCR reaction contained 1X Taq Plus buffer, 1 ul template RNA, 200 uM dNTPs, 1 uM primers, 1 ul MMLV-RT, and 5 units of Taq polymerase in 50 ul total volume. The DNA yield was assessed by electrophoresis through a 1% agarose gel in 1X TBE. The RT-PCR product was cloned into pGEM-T Easy vector (Promega). The construct was transformed into *Escherichia coli* DH5α cells by electroporation. Presence of the insert was selected for on LB agar plates with ampicillin (30 ug/ml) and X-gal and confirmed by restriction analysis of the minipreparations of plasmid DNA using EcoR I. DNA fragments were separated by electrophoresis through 1 % agarose gel in 1X TBE. The γ–CA2 cDNA was sequenced by Genelab at LSU.

A-Tailing γ–CA2 promoter PCR product.

The promoter PCR product was purified and concentrated by the phenol/chloroform method and ethanol precipitation. Specifically, 100 ul DNA were added to 100 ul (1X vol) phenol/ choloroform and mixed by inversion (20X) before centrifugation for 6 min at 12,000 rpm. The upper layer (approx. 100 ul) was transferred to a new tube and 10 ul(1/10X vol) of 3M Na acetate and 220 ul (2X vol) of 100% ethanol were added. After 10 min at –20C, the mixture was centrifuged at 12,000 rpm. The supernatant was discarded and the pellet was washed with 100 ul of 70% ethanol before centrifugation at 12,000 rpm for 12min. The supernatant was discarded and the pellet was air dried for 20 min and resuspended in 10 ul of ddH_2O. The A-tailing reaction was prepared in a microfuge tube using the 10 ul of DNA purified previously, 1X PCR reaction buffer,

2.5 mM $MgCl_2$, 0.2mM dATP, and 5 units of Taq Polymerase in 100 ul total volume. The tube containing the A-tailing reaction mixture was incubated in a 70C hot water bath for 30 min. After the A-tailing reaction was complete, the ethanol precipitation method was used to concentrate the A-tailed DNA in a total volume of 10 ul of TE. The product was run out on a low melt 1% agarose gel, and the 1.6 kb band was cut out for use in the cloning experiment.

Cloning of γ–CA2 promoter into pGEM-T easy vector.

The A-tailed PCR product was heated at 65C in a water bath for 10 min to melt the agarose gel and cloned into pGEM-T Easy vector (Promega) according to the manufacturer's protocol (Promega). The construct was transformed into *Escherichia coli* DH5α cells by electroporation. Presence of the insert was selected for on LB agar plates with ampicillin (30 ug/ml) and X-gal and confirmed by restriction analysis of the minipreparations of plasmid DNA using EcoR I. The pGEM-T Easy plasmid containing the promoter insert was also digested with EcoR I and Xho I to verify targets. DNA fragments were separated by electrophoresis through 1 % agarose gel in 1X TBE. The γ–CA2 promoter DNA was sequenced by Genelab at LSU.

Cloning of the γ–CA2 promoter into pBI101-2 (Promoter-GUS fusion).

The pGEM-T Easy – γ–CA2 promoter construct was sequentially double digested with Sal I followed by Avr II, and the insert was ligated into the binary vector pBI101-2 which had been sequentially double digested with Sal I followed by Xba I and treated with calf intestinal phosphatase. The ligation was transformed into *E. coli* DH5α cells by electroporation, and the cells were plated on LB agar plates with kanamycin (30 ug/ml). Presence of the insert was confirmed by restriction analysis of minipreparations of plasmid DNA using Hind III. The constructs containing the γ–CA2 promoter inserted in front of the β-glucuronidase (GUS) gene of pBI101-2 are designated as pBI-γ–CA2 promoter.

Transformation of pBI- γ–CA2 promoter into *Agrobacterium tumefaciens*.

The pBI- γ–CA2 promoter construct was transformed into *Agrobacterium tumefaciens* GV3101 pMP90 (Koncz and Schell, 1996) by electroporation and transformants were selected by growth on LB media containing gentamycin (50 ug/ml) and kananycin (30 ug/ml). Isolated colonies were inoculated into 3 ml overnights of LB with antibiotics.

Transformation of pBI- γ–CA2 promoter into *Arabidopsis thaliana* by vacuum infiltration.

The protocol for transformation of *Arabidopsis thaliana* by vacuum infiltration was provided by Andrew Bent (1994) and based on the work of Bechtold *et al.* (1993). To obtain plants for transformation, seeds of *Arabidopsis thaliana* ecotype Columbia seeds were suspended in 0.15% agar and directly spread over soil in a 4" pot covered by nylon screen. The pots were covered with a plastic lid and placed in the dark for two days at 4C, and then transferred to continuous light at 23C. After seeds sprouted the lid was removed. Plants were watered three times per week and were fertilized one time per week. Plants were thinned to nine per pot. After about one month, the plants were trimmed back to induce more shoots. This produced more flowers per plant which increased the chances of obtaining transformants. After one more week the plants were ready to be transformed. Plants were ready when many bolts 1-5 cm long and some longer but only a few dozen open flowers were present. Plants were watered very well the day before infiltration.

When plants were ready to transform, 50 ml of LB media containing gentamycin (50ug/ml) and kanamycin (30ug/ml) were inoculated with cells from a 3 ml overnight starter culture of *Agrobacterium* containing the pBI- γ–CA2 promoter construct. Cultures grew for 36 hours at 28C in a shaker at 300 rpm. The next day 50 ml of culture were transferred into 500 ml LB containing kanamycin and gentamycin to grow for two more days at 28C to an A600 greater than 2.00.

The *Agrobacterium* cultures were harvested and resuspended in two volumes of infiltration media. One liter of infiltration medium contains 2.2 g Murashige & Skoog salts with vitamins, 50 g sucrose, 0.5 g MES, pH to 5.7 with KOH, 44 nM benzylaminopurine (10 ul per liter of a 1 mg/ml stock in DSMO), 200 ul Stilwet L-77 ('Vac-In-Stuff' Lehle Seeds). 400 ml of infiltration medium were used for infiltration of four pots of *Arabidopsis*. The resuspended infiltration medium and cell mixture, about 400 ml, was poured into a container inside a vacuum dessicator. The pot containing the plants was inverted into the solution, so that all aboveground portions of the plants were submerged. A vacuum was established for 5 min by suction from an aspirator. The water pressure was raised until air bubbles were exiting the plants leaves. After the vacuum was released, each pot was placed on its side in a tray. The tray was covered with a plastic cover to maintain humidity. The pots were set upright and uncovered the next day. The transformed plants were grown under continuous light for about three weeks at 23C, before watering was halted to allow the plant to produce seed and die. Seeds were then harvested and collected separately from each plant.

Selection of transgenic plants.

Selection was performed using selection media consisting of 1X MS salts with vitamins, 0.5 g/l MES, pH to 5.7 with KOH, 0.8% agar, kanamycin (50 ug/ml) and gentamycin (30 ug/ml) to select for transgentic plants, and vancomycin to kill any agrobacterium that survived the sterilization without harming the plants.

After the harvested seeds were completely dry, they were collected into separate vials. Transgenic plants which had the highest yields of seed were selected for further study, so that a greater number of seeds could be plated per plant and there would be enough left over to repeat. 1000 seeds from each transgenic plant were weighed out and put in a 15 ml tube to be sterilized: 1 min in 70% ethanol, 7 min rocking in 50% bleach/ 0.02% Triton X-100, 3 rinses in sterile water.

Then they were resuspended in 0.1% agarose (2 ml). The agarose containing about 1000 seeds was evenly distributed on six selection plates (150 mm) using a technique, developed by Linda Shaffer, similar to using a potter's wheel. While the plate on a turntable was spun, one hand steadily the other hand to slowly release the fluid from a P-1000 pipette-man starting at the outside edge and moving towards the center making sure not to gouge the surface. This made a spiral pattern of agarose and seeds on the surface of the plate. The plates were sealed with paper tape, labeled, and placed in the dark at 4C for 2 days, then transferred to continuous light at 23C for ten days. The transformants (transgenic seedlings) were distinguished from non-transformants by the formation of dark green first true leaves and a healthy long root system after ten days. The non-transformants were small, white, short rooted, and lacked first true leaves as compared with transformants. The transformants were used for analysis of GUS activity or were carefully moved to individual pots of soil to grow under normal conditions until the seeds were harvested. The transgenic plants were named γ–CA2 promoter.

Analysis of transgenic *Arabidopsis thaliana* for GUS activity.

In order to examine the pattern of expression of γ–CA2 in *Arabidopsis thaliana,* one to two γ–CA2 promoter transgenic seedlings from each selection plate were assayed for β-glucoronidase (GUS) activity following the method of Gallagher (1992) using 5-bromo-4chloro-3-indolyl-β-D glucoronide (X-Gluc) as a substrate.

The transgenic seedlings were carefully transferred from each plate with fine forceps into microfuge tubes containing 0.5 ml of GUS staining solution (100 mM sodium phosphate pH 7, 10 mM EDTA, 0.1% Triron X-100, 1 mg/ml X-Gluc, and 100 ug/ml Chloramphenicol). After the seedlings were transferred to the staining solution, the tubes were wrapped in aluminum foil and placed in a vacuum dessicator. The vacuum was drawn for 10 min. The tubes were then incubated

at 37C overnight, after which the GUS staining solution was removed and replaced with 70% ethanol at 37C for several hours to destain the seedlings. The seedlings were then visually observed for staining, and the seedlings with the more remarkable staining were selected to be examined by microscopy.

Results and Discussion

Expression of the γ-CA2 gene

From the *Arabidopsis* genome database a gamma carbonic anhydrase homologue, γ-CA2, was chosen for study. γ-CA2 is located on Chromosome I (BAC F14P1; BAC accession: AC024609; Protein ID# AAF98404.1). The 2145 bp gene is expressed as a reverse compliment and contains four introns and five exons (Figure 1a): 124 bp, 143 bp, 125 bp, 155 bp, and 290 bp encoding a protein containing 278 amino acids (Figure 1b). Using the GeneStream align tool, the amino acid sequence of γ-CA2 showed 75% identity to the sequence of an *Arabidopsis* γ-CA analyzed by Xi (1999) (Figure 2). The 0.8 kb RT-PCR product (Figure 3-lane C) and its subsequent restriction analysis with EcoR I (Figure 4) confirmed that γ-CA2 is expresses in *Arabidopsis.*

From the aligning results, the restriction analysis, and the RT-PCR results, the γ-CA2 gene appears to code for a carbonic anhydrase and appears to be expressed in *Arabidopsis*. This does not prove that the γ-CA2 protein is enzymatically active, but further research on γ-CA2 expression patterns may tell more about its activity and importance to *Arabidopsis.*

Confirmation of identity of γ-CA2 promoter

The promoter PCR product was approximately 1.6 kb (Figure 3-lane D) as expected. A restriction digest with EcoR I and Xho I verified the presence of the promoter (Figure 5). From the sequencing results, the promoter was 95% identical to gene sequence from AC024609 (Figure 6), which is enough to assume identity.

A.

AGTGTTGAGTCCTCGTTTACCTGTGAGCTCGAAGAAAGTGACGATCA**ATGGGAACCCTAGGCAGAGCATTTTACTCGGT**
CGGTTTTTGGATCCGTGAGACTGGTCAAGCTCTTGATCGCCTCGGTTGTCGCCTTCAAGGCAAAAATTACTTCCGAGAA
CAACGTAAGATTCTGCGTAACTCTATTTTTCACTCCGATTTCGAATTCGTAAAGAGTGAGGAGTTTCGATTGGGTAATT
CAGCTCACCTATTCTGGTTAGGTTAGAGATCTGGAATTGATGTCTTTGTAGGTATAGCGATTCTGGATTCGAATGGCAC
AAGGGATTAATTGATGTGGAATGTGTTTTGCTGCTTAGTTTTTTTTTTGTTTCCTGGTCACGAAATTTGCTTATATGTA
ATTGTGGTTTTCTCTGTGTAG**TGTCAAGGCATCGGACACTGATGAATGTATTTGATAAGGCTCCGATTGTGGACAAGGA**
AGCTTTTGTGGCACCAAGCGCCTCAGTTATTGGGGACGTTCACATTGGAAGAGGATCGTCCATTTGGTATGGATGCGTA
TTACGAGGTGAGCTAGCTACTAAGCCTCAATGTTTGGTGTTTGGTGGTTGACTATTTTAGTGACTGAAGTTTTGTTTGA
AATAGTATCTACTTCTGTATCCTTAGCTGTGTCAACTCTTACGATAGAAAATAGTGAACGAATATCTGTTTTTTACTCT
TCTGGTTAAGCGATGAACAAGACTACTACTCAAAATGGAGCTTTCTAGCTTTATTGAAGGATCTTAAACAGATCATTAC
TACATATCGTATCTTGATGATACTGGCAGTTGCTCTAATGGGTGATTGGGTATATGATATCCCCAATCATATACAGAAT
GCTTTCATGTCATAACACCAGAATTAGTTGAACGATATCATATACAGATTATTGAGTTTGAGTTGAATCCTTATTGCGA
GTATGAGTTGGCTAGTCTGGATTTTTATTTAATGGTGATGAAATAGGGTTGTCCTTTTTACAGAATGCTTTCATGTCAT
AACACATTGGTACTGTATGGCTAAGCCTGTTAGTACTTTACCAGGAAACTGGGATAACATACAATCTAAACTTGCTTTC
TCTCTTTTTTTCTGCGCAG**GCGATGTGAACACCGTAAGTGTTGGGTCAGGAACTAATATTCAGGACAACTCACTTGTGC**
ATGTGGCAAAATCAAACTTAAGCGGGAAGGTGCACCCAACCATAATTGGAGACAATGTAACCATTGGTTAGTGGGAATC
TCTTACTCACTGTTTATTTTGAAAATGTTTAATATGCATACATCTCCAAGAAGCATTTGGGATATAGTCTTAAATTAAA
TGTGTAACAG**GTCATAGTGCTGTTTTACATGGATGTACTGTTGAGGATGAGACCTTTATTGGGATGGGTGCGACACTTC**
TTGATGGGGTCGTTGTTGAAAAGCATGGGATGGTTGCTGCTGGTGCACTTGTACGACAAAACACCAGAATTC**CTTCTGG**
AGAGGTATGATTTTGGGTTTAACATTTCTGATTACCTGAAAACCCACAAGTAACCATTGGTGCGGGTATACTTCTATTG
CTACTTAGACAGCATAGGCTATATTATTGACATTATTAGTGTCTTTCTTTCATTGGCTGAAAGTATTGGAGGTTGGTGT
GAAATTGAAATACAGTTTCTAGTCTTGAGAATTCTTCTCCTCTATCTGGATAGAACGTATCTTATCCACTTTTCAGATG
CCCTAACAGATTATAAATGTTAGTTAGATGAAAGTGGAAATGTTCTTCTAGCTATGGTTTATCTGCATATGAATGCCAA
AGGGTGTTAGTCTCTCATCATGGCCTTAAAATTTACTTCATCCGCCATATTGACTGACAGAGTTAAAATGTTCGCATTG
GTCTATAATGACTCA**GGTATGGGGAGGAAACCCAGCAAGGTTCCTCAGGAAGCTCACTGATGAGGAAATTGCTTTTATC**
TCTCAGTCAGCAACAAACTACTCAAACCTCGCACAGGCTCACGCTGCAGAGAATGCAAAGCCATTAAATGTGATTGAGT
TCGAGAAGGTTCTACGCAAGAAGCATGCTCTAAAGGACGAGGAGTATGACTCAATGCTCGGAATAGTGAGAGAAACTCC
ACCAGAGCTTAACCTCCCTAACAACATACTGCCTGATAAAGAAACCAAGCGTCCTTCTAATGTGAACTGATTTTTCAGG
GGTATGTTTTCTGGCCGAAGCCCTACAGGGTGAGATACTCAAGGGGATTATGTTTCGGTCTCTGGTTTGAATATGGCAG
GTAGAGTAC

B.

MGTLGRAFYSVGFWIRETGQALDRLGCRLQGKNYFREQLSRHRTLMNVFDKAPIVDKEAFVAPSASVIGDVHIGRGSSI
WYGCVLRGDVNTVSVGSGTNIQDNSLVHVAKSNLSGKVHPTIIGDNVTIGHSAVLHGCTVEDETFIGMGATLLDGVVVE
KHGMVAAGALVRQNTRIPSGEVWGGNPARFLRKLTDEEIAFISQSATNYSNLAQAHAAENKPLNVIEFEKVLRKKHALK
DEEYDSMLGIVRETPPELNLPNNILPDKETKRPSNVN

Figure 1. Nucleotide sequence of the γ-CA2 gene (A) and deduced amino acid sequence of the cDNA encoding γ-CA2 from *Arabidopsis* (B). The primers for RT-PCR are underlined. The EcoR I sites are highlighted. The predicted exons (A) and N-terminus amino acids in the promoter-GUS construct (B) are shown in bold.

```
CLUSTAL W (1.81) multiple sequence alignment

CA1             MGTLGRAFYSVGFWIRETGQALDRLGCRLQGKNYFREQLSRHRTLMNVFDKAPIVDKEAF 60
CA3             ---MGKAFYSVGFWIRETGQALDRLGCRLQGKNHFREQLSRHRTLMNVFDKTPNVDKGAF 57
CA2             MGTLGRAIYTVGNWIRGTGQALDRVGSLLQGSHRIEEHLSRHRTLMNVFDKSPLVDKDVF 60
                   :*:*:*:** *** *******:*. ***.: :.*:**************:* *** .*

CA1             VAPSASVIGDVHIGRGSSIWYGCVLRGDVNTVSVGSGTNIQDNSLVHVAKSNLSGKVHPT 120
CA3             VAPNASLSGDVHVGRGSSIWYGCVLRGDANSISVGAGTNIQDNALVHVAKTNLSGKVLPT 117
CA2             VAPSASVIGDVQIGKGSSIWYGCVLRGDVNNISVGSGTNIQDNTLVHVAKTNISGKVLPT 120
                ***.**: ***::*:************.*.:***:*******:******:*:**** **

CA1             IIGDNVTIGHSAVLHGCTVEDETFIGMGATLLDGVVVEKHGMVAAGALVRQNTRIPSGEV 180
CA3             VIGDNVTIGHSAVLHGCTVEDEAYIGTSATVLDGAHVEKHAMVASGALVRQNTRIPSGEV 177
CA2             LIGDNVTVGHSAVIHGCTVEDDAFVGMGATLLDGVVVEKHAMVAAGSLVKQNTRIPSGEV 180
                :******:*****:********::::* .**:***. ****.***:*:**:**********

CA1             WGGNPARFLRKLTDEEIAFISQSATNYSNLAQAHAAENAKPLNVIEFEKVLRKKHALKDE 240
CA3             WGGNPAKFLRKVTEEERVFFSSSAVEYSNLAQAHATENAKNLDEAEFKKLLNKKNAR-DT 236
CA2             WGGNPAKFMRKLTDEEIVYISQSAKNYINLAQIHASENSKSFEQIEVERALRKKYARKDE 240
                ******:*:**:*:** .::*.** :* **** **:**:* ::  *.:: *.** *  *

CA1             EYDSMLGIVRETPPELNLPNNILPDKETKRPSNVN--- 275
CA3             EYDSVL-------DDLTLPENVPKAA------------ 255
CA2             DYDSMLGITRETPPELILPDNVLPGGKPVAKVPSTQYF 278
                :***:*        :* **:*:

Sequence 1: CA2              278 aa
Sequence 2: CA1              275 aa
Sequence 3: CA3              255 aa

Sequences (2:3) Aligned. Score:  77
Sequences (1:2) Aligned. Score:  75
Sequences (1:3) Aligned. Score:  69
```

Figure 2. Aligning results of γ-CA2 (CA2) versus Xi's γ-CA (CA1) and γ-CA3 (CA3).

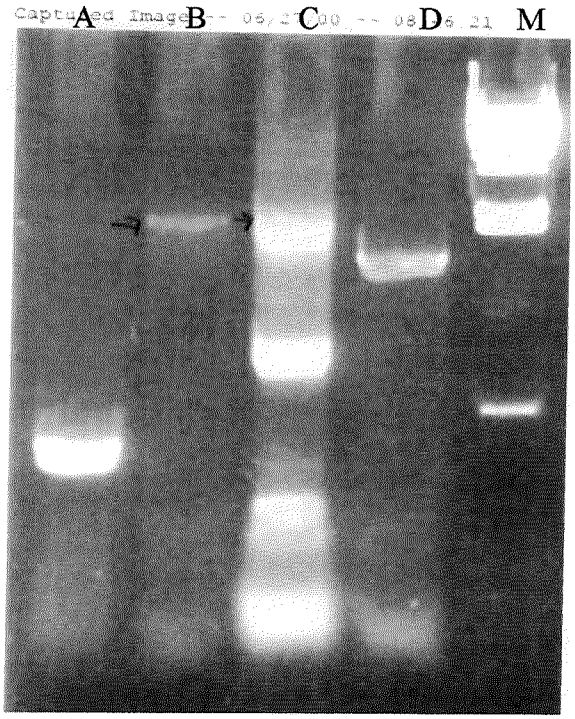


Figure 3. Agarose gel of RT-PCR positive control (lane A), RT-PCR negative control (lane B), RT-PCR product of cDNA from Katrina's 21% leaf total RNA (lane C), PCR product of promoter from *Arabidopsis* leaf DNA (lane D), and λ Hind III marker (lane M)

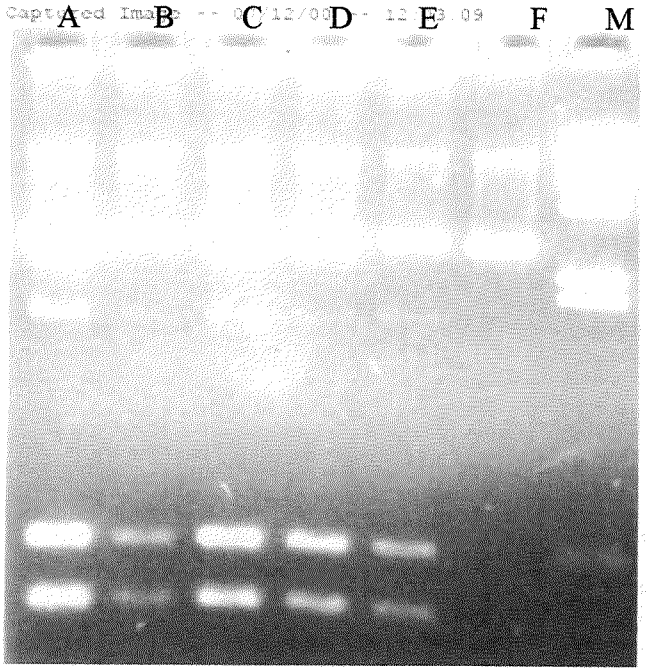


Figure 4. Agarose gel of restriction analysis of pGEM-T Easy-cDNA minipreps (lanes A-F) with EcoR I and the λ Hind III marker (lane M).

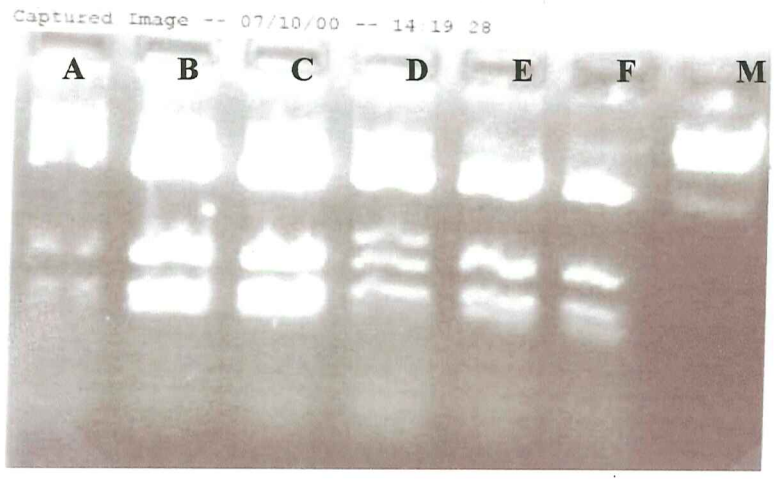


Figure 5. Agarose gel of restriction analysis of pGEM-T Easy-promoter minipreps (lanes A-F) with EcoR I and Xho I, and λ Hind III marker (lane M).

GATATAAGAGAGAGACCTTGAACAACATTACCACACCAACGATCTCATAACGGTGCATAGTTTCAAGGGTTAACCTTGG
GAGCCCATCCAGAGATCACATACTTTTCCTCGGTCTTAGTTTTCTCAAAAGAGTCGAGCGAGAACAGAGTCTTCATGTA
GTTATGGACATGGGGAAAGCTCTCAGGGACAGACCAGCTTTTGAAATGGCCAAGAGCAACTTGAAGGTGGTAAAGCTTT
GGTGCTAAGCTTAGATCCACTGCGGAAACTCTTTCTCCGGCGATAAAAGGGCCATCGTGACTCTTAAGATGGTTTTCCA
AAGCTTCTAGCTCAACAAGCAAGGCATGTTCAGATCCGTCATTGGAGTCTTTGCTCTTCAAGAATGTCCCAAAAGTACC
AAAAATGTTGGATCCAACAGAGGCAAATTCAGCAGGAGTCTTGAGTGGTGGATCAGGATACTTCTCCTCGAGTATACCA
ACGATGACGTCGGAATCAGTCACCCACTTGTCGTCGATCTTAAGCACTGGTACTTTCCCTTGAGGACTAATGTCCAAGA
ACCTGCATCATATATAAAACCAAAGCCACCAATTTCAAGTTCCGGTTAACACAATGCTCGCACATGAAGCCAATCCAAT
TTCCGGTTAACACAAAGCTCGTACCAGTTACAAGTAAAATTTCAGAGATTATGGCTAAACAAGTTCAAACTAGACCGAT
CTTAAGGGTTTAGGCGTGAGAGAAACTGACCACTGGGGTTTGTCAGAGAGGTTAATCAGATGGATTTTGTAGGTAAGAC
TCTTCTCCTCGAGTGTGAGAAGAGCCCGTTGGCTGAACGGACCTAAAACGACAAACAAACACTTAAATTAGTAAATTTA
CCAACGATCAAAGCTCAGAAATTAAGAATTGCAAAATCAAAAGCTTAATGCAAAACAAAAACAGAGAGAGAAAGAGAAG
ATGGTTACAGTCGCCGAGATGATCAGGAGCACCAACAGCAGCTTTCACACAGATTTCCAGAGCCATCTTGATTCTTCTT
CCTTTGTTTGCAGAGACGGATGAATGATTGAGTTAATCAGTGAGAAGGCTTAGCTTTTAAACTGAATTATTACCATTCT
CCACCCACCAGTGGGCCATTTTTACCCGATGTTTATCTTTTCTCACGTGTTATGCGTTGCAAGTGACGACGATACGAAA
CGACATTGTTACTGATAGGGCCATTTGATTTTTGGTCGGCAGGTTTTGGACCGTTCAATTTAGTGAAAATATAGTAATA
AATTCGTACCGAATAATATAATCGGTTTAATGAAACTACTACCTGAAATATCGAATTAGATTTGGTTCATATTTCGTAT
TGTTTTGGTTCGATTAAACCATATTAAACCACTGTAACTGTAATTTATTGTTTCGCCGTCCCGGAATGTTCCTGTTGAA
ATCCATTTTCGCTGATTTTTTTTCTTCCGTCTCTTCTTCAGCTTCGACCATTTTCGTCTTCTTCATTCAGTGTTGAGTC
CTCGTTTACCTGTGAGCTCGAAGAAAGTGACGATCA**ATGGGAACCCTAGGCAGAGCATTTTACTCGGTCGGTTTTTGGA**
TCCGTGAGACTGGTCAAGC

Figure 6. Nucleotide sequence of the γ-CA2 promoter which has been cloned into the pGEM-T Easy and pBI101-2 vectors. The primers for PCR are underlined. The four Hind III sites are highlighted. The nucleotides in the open reading frame are in bold.

Formation of the γ-CA2 promoter – GUS construct

The promoter was inserted in-frame 5' to the β-glucoronidase gene (GUS). The construct was verified by restriction analysis of pBI101-2-promoter minipreps of transformed *E. coli* with Hind III (Figure 7).

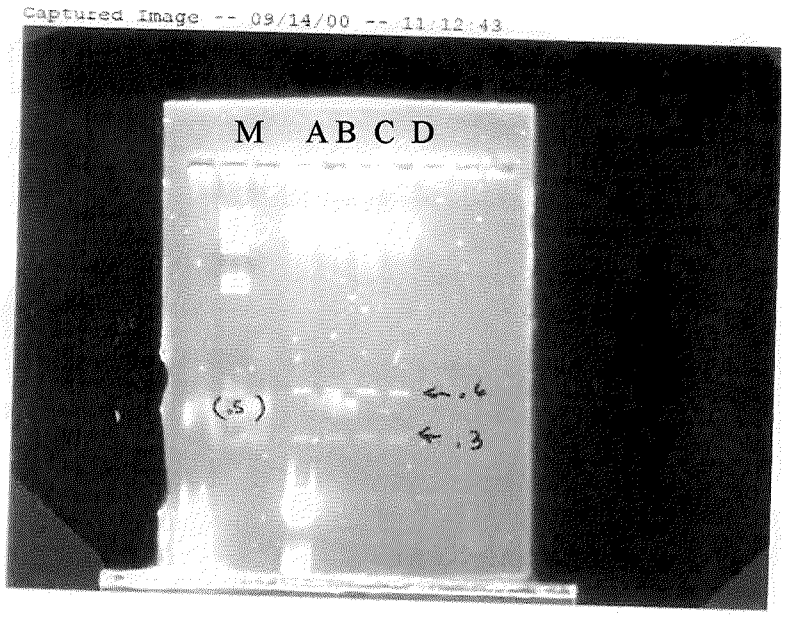


Figure 7. Agarose gel of restricion analysis of pBI101-2-promoter minipreps (lanes A-D) with Hind III, and λ Hind III marker (lane M).

Selection of transgenic *Arabidopsis*

Most of the seeds that germinated on selective media produced white seedlings with poor roots and no true first leaves. However, approximately 5-15 seeds per plate grew into green seedlings with strong root systems and true leaves. The plants were able to thrive because they were antibiotic resistant and contained the γ-CA2 promoter – GUS construct.

These hemizygous first generation transgenic seedlings were used to study γ-CA2 gene expression. However, homozygously transgenic plants could be selected for based on Mendalian genetics. The second generation of plants would include 25% homozygously transgenic plants. Statistically within the third or fourth generation the progeny would all be homozygously transgenic. In the case of homozygous transgenics, nearly all seeds would grow on antibiotic selection plates. Studies using the homozygous transgenics may produce different results than this research did.

GUS Activity Assay and Microscopy of transgenic *Arabidopsis* seedlings

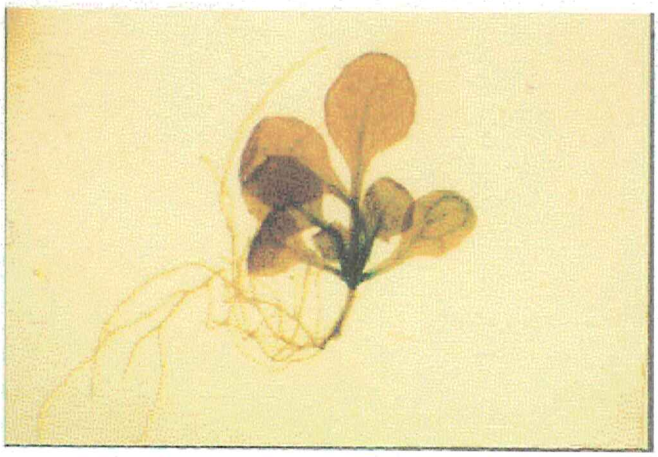
The production of GUS in the transgenic seedlings was visually observed by treating the seedlings with X-Gluc. The enzymatic activity of GUS catalyzed the reaction to cleave X-Gluc into two products one of which is indole. Indole has a blue color, so wherever GUS was produced the blue color was visualized.

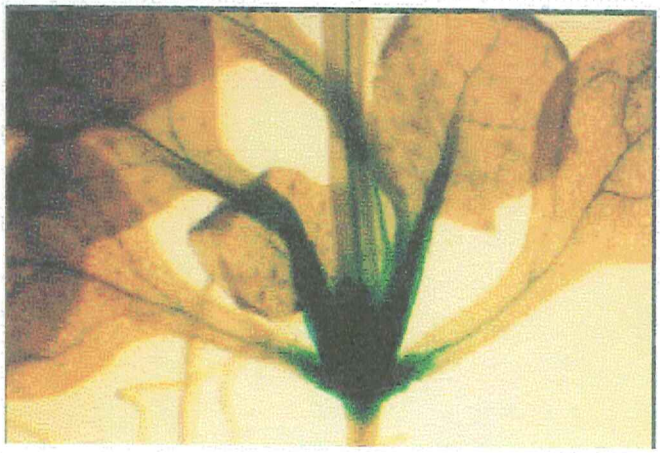
Of the twelve seedlings assayed, eight showed blue staining and four did not. Plant to plant variability in the level of transgene expression is often observed since the T-DNA inserts randomly in the genome. The plants assayed were hemizygous, and homozygous plants may well exhibit detectable GUS activity. The four seedlings showing the darkest staining were sent to the

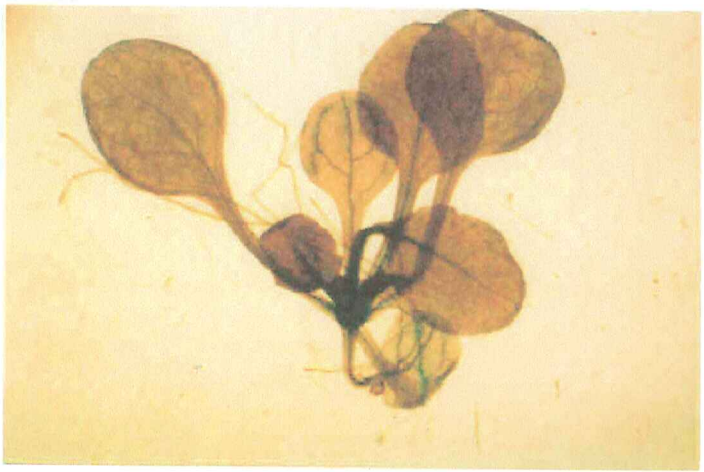
microscopy lab in the Department of Biological Sciences to be photographed (Figure 8) and sectioned (Figure 9). The GUS staining was strongest at the immature leaves, near the apical meristem, and radiated out a short distance into the petiole of the leaves and cotyledons and down the hypocotyl (Figure 8).

A previous study by Xi (1999) examining the expression patterns of another γ-CA in *Arabidopsis* resulted in an expression pattern which was different. That γ-CA was expressed in meristematic tissues namely in leaf primodia, root tips, and cotyledons. Xi's γ-CA was only present at very early stages of development and disappeared quickly. The new pattern for γ-CA2 suggests that the expression patterns of CAs in the γ-family do not overlap. Each γ-CA may be localized and specific to areas of the plant or to a certain time of expression. The γ-CAs may phase in and out during development of the plant. A temporal study of expression of both genes may bring to light the relationship between Xi's γ-CA and γ-CA2.

From sections of the transgenic plants, some information about the subcelluar localization of γ-CA2 was gathered. γ-CA2 appeared to be a cytosolic protein. Expression seems to be absent from the vacuoles and intercellular space which do not show any blue color (Figure 9). Also, the subcellular expression is not generalized throughout the cytoplasm. The areas of expression appear as dark granules in the cytoplasm, suggesting that the 21 N-terminus amino acids encoded by the promoter –GUS construct may be part of an amino terminal signal sequence. γ-CA2 may be linked to a specific organelle or protein complex in the cytoplasm, possibly a plastid because of the large size of the granules and the plastids role in photosynthesis. However, the GUS product could have precipitated out of the cyosol and was aggregating on nonspecifically on organelles or the plasma membrane.







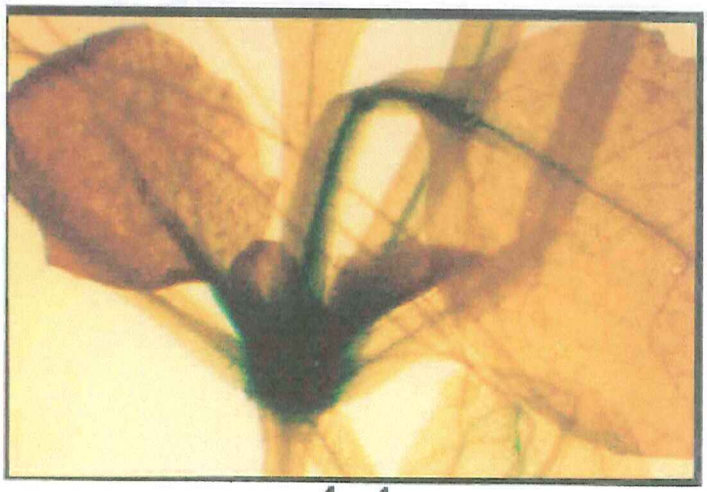
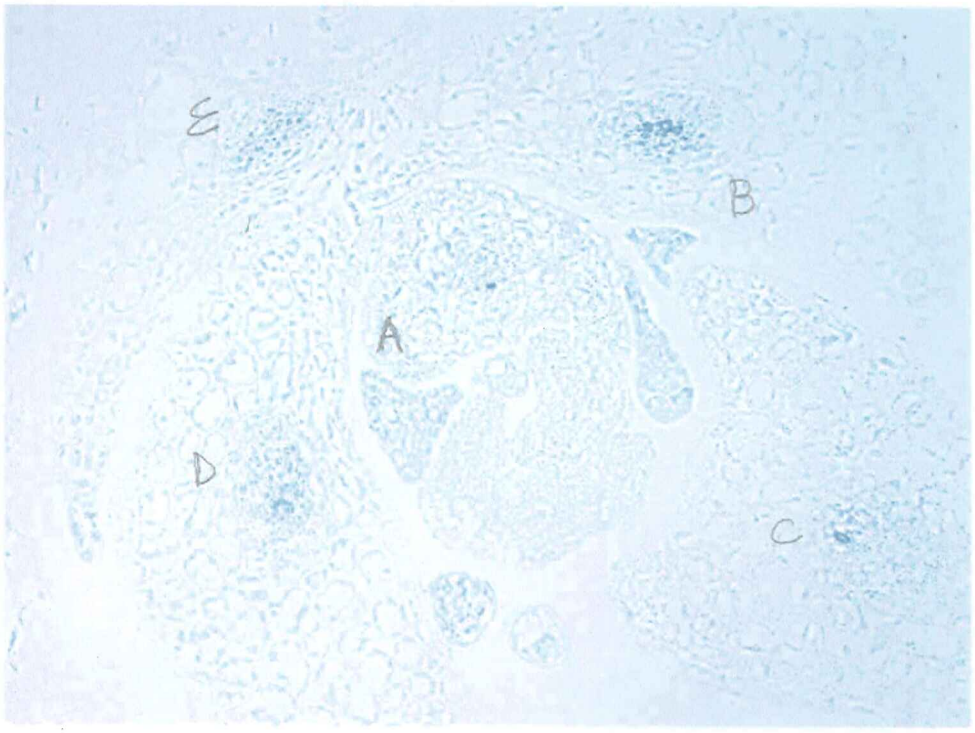
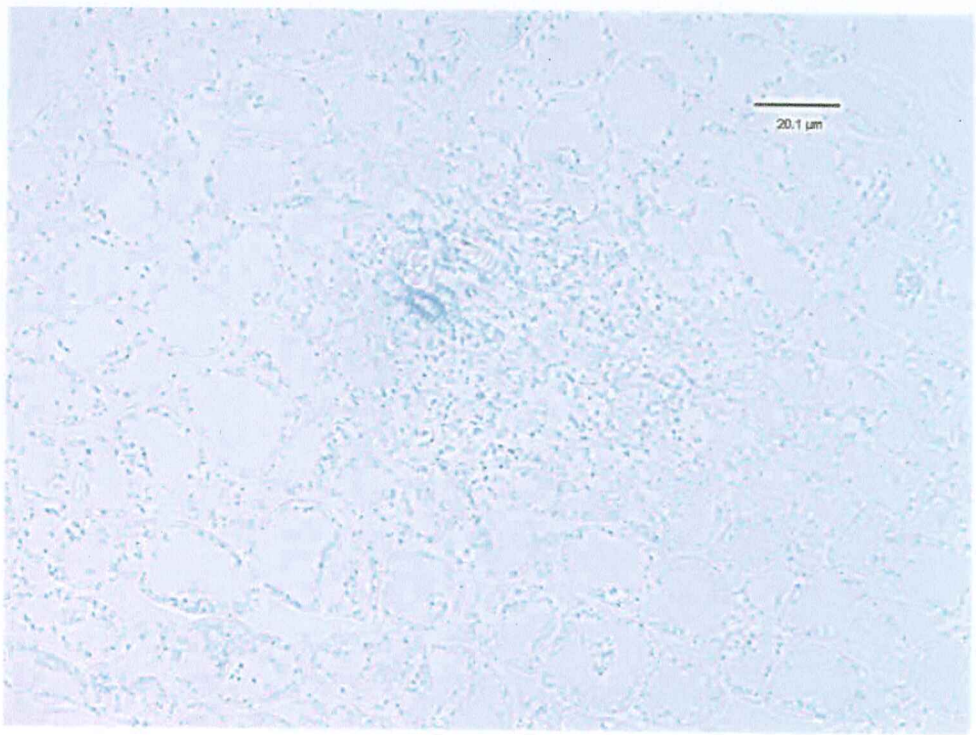


Figure 8. Light microscope pictures of the selected GUS stained *Arabidopsis* seedlings. Note that the dark blue green areas are where γ-CA2 is expressed.

A.



B.



C.

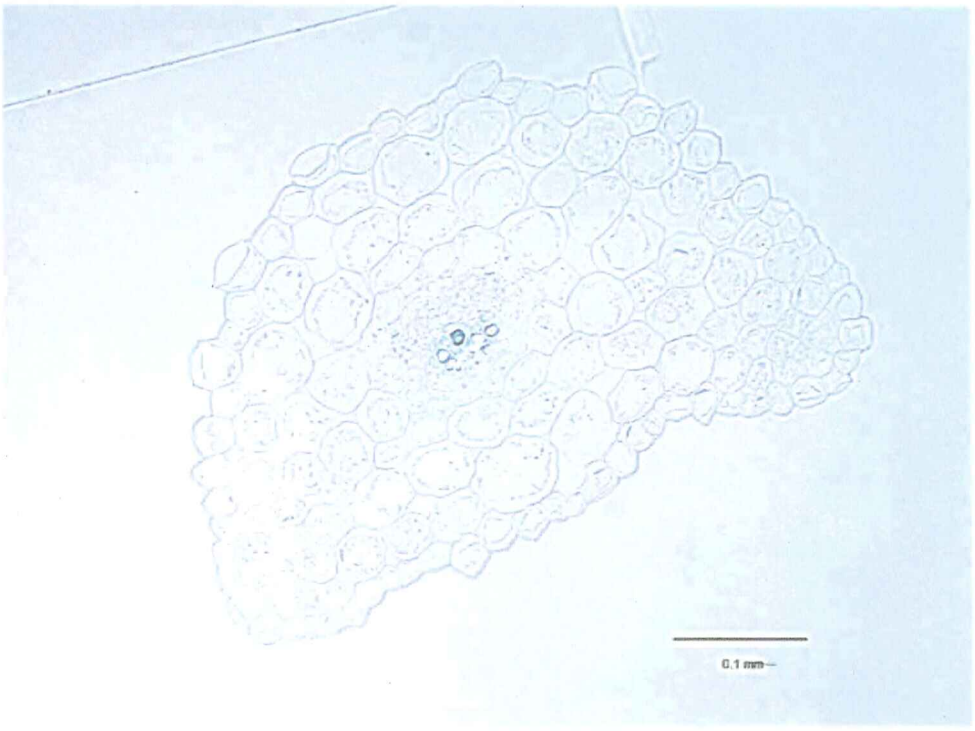


Figure 9A,B,C. Light microscope pictures of cross-sections of GUS stained *Arabidopsis* seedlings. (A) Apical vegetation, (B) Cotyledon, and (C) Petiole.

Conclusions and Future Research

γ-CA2 is the second gamma carbonic anhydrase gene to show expression in *Arabidopsis.* The deduced amino acid sequence showed high similarity to Xi's γ-CA. Also, the amplification of the cDNA by RT-PCR demonstrated that γ-CA2 is expressed in *Arabidopsis.*

The expression pattern of γ-CA2 was different than the expression pattern for the γ-CA characterized by Xi (1999). In the work described here, the difference is likely due to temporal shifts of expression of the γ-CAs. However, specific localization of γ-CAs in *Arabidopsis* or a combination of temporal and spatial regulation cannot be rule out.

The subcellular expression patterns revealed that γ-CA2 is expressed in the cytoplasm, and is likely associated with a complex or organelle, possibly a plastid, in the cytoplasm. The immature amino acid sequence could include an amino terminal signaling sequence as many other enzymes that are localized do. Signaling sequences are specific to organelles, so a better understanding of the significance of the amino terminal of γ-CA2 and sequences that target specific organelles in *Arabidopsis* could reveal more about γ-CA2's subcellular localization. This may lead to better understanding of γ-CA2's possible enzymatic function.

Further studies of the temporal pattern of γ-CA2 gene expression in developing and adult *Arabidopsis* should help us to understand the relationships between γ-CAs. Also, studying changes in the plant phenotype with an antisense construct could shed light on the importance of γ-CA2 to the survival and growth of the plant.

Works Cited

Alber B.E. & Ferry J.G. (1994) A carbonic anhydrase from the archaeon *Methanosarcina thermophila*. Proceedings of the National Academy of Sciences, USA **91,** 6909-6913.

Badger M.R. & Price G.D. (1994) The role of carbonic anhydrase in photosynthesis. Annual Review of Plant Physiology and Plant Molecular Biology, **45**, 369-392.

Bechtold N., Ellis J. & Pelletier G. (1993) In planta *Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. C.R. Acad. Sci. Paris, Life Sciences, **316,** 1194-1199.

Bent A.F., Kunckel B.N., Dahlbeck D., Brown K.L., Schmidt R., Giraudat J., Leung J. &Staskawicz B.J. (1994) RPS2 of *Arabidopsis thaliana*: A leucine-rich repeat class of plant disease resistance genes. Science, **265,** 1856-1860.

Bracey M.H., Christiansen J., Tovar P., Cramer S.P. & Bartlett S.G. (1994) Spinach carbonic anhydrase: Investigation of the zinc-binding ligands by site-directed mutagenesis, elemental analysis and EXAFS. Biochemistry, **33,** 13126-13131.

Eriksson M., Karlsson J., Ramazanov Z., Gardeström P. & Samuelsson G. (1996) Discovery of an algal mitochondrial carbonic anhydrase: molecular cloning and characterization of a low-CO_2-induced polypeptide in *Chlamydomonas reinhardtii*. Proceedings of the National Academy of Sciences, USA **93,** 12031-12034.

Eriksson M., Villand P., Gardeström P. & Samuelsson G. (1998) Induction and regulation of expression of a low-CO_2-induced mitochondrial carbonic anhydrase in *Chlamydomonas reinhardtii*, Plant Physiology, **116,** 637-641.

Fujiwara S., Fukuzawa H., Tachiki A., & Miyachi S. (1990) Structure and differential expression of two genes encoding carbonic anhydrase in *Chlamydomonas reinhardtii*. Proceedings of the National Academy of Sciences USA, **87,** 9779-9783.

Fukuzawa H., Suzuki E., Komukai Y. & Miyachi S. (1992) A gene homologous to chloroplast carbonic anhydrase (*icfA)* is essential to photosynthetic carbon dioxide fixation by *Synechococcus* PCC 7942. Proceedings of the National Academy of Sciences USA, **89,** 4437-4441.

Gallagher S.R. (1992) GUS protocols: using the GUS gene as a repoter of gene expression. Acadeemic Press Inc., San Diego.

Karlsson J., Clarke A.K., Chen Z.-Y., Hugghins .SY., Park Y.I., Husic H.D., Moroney J.V. & Samuelsson G (1998) A novel alpha-type carbonic anhydrase associated with the thylakoid membrane in *Chlamydomonas reinhardtii* is required for growth at ambient CO_2. EMBO Journal **17,** 1208-1216.

Kim H.J. (1997) Identification of the physiological role of carbonic anhydrase using the antisense technique and investigation of its transcriptional regulation in *Arabidopsis thaliana* (L.) Heynh. Ph.D. dissertation, Louisiana State University, Baton Rouge, Louisiana.

Kimber M.S., & Pai E.F. (2000) The active site architecture of *Pisum sativum* β-carbonic anhydrase is a mirror image of that of α-carbonic anhydrases. EMBO Journal, **19**, 1407-1418.

Kisker C., Schindelin H., Alber B.E., Ferry J.G. & Rees D.C. (1996) A left-handed beta-helix revealed by the crystal structure of a carbonic anhydrase from the archaeon *Methanosarcina thermophila*, EMBO Journal, **15,** 2323-2330.

Koncz C. & Schell J. (1986) The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. Mol. Gen. Genet., **294,** 383-396.

Ku M.S.B., Kano-Murakami Y. & Matsuoka M. (1996) Evolution and expression of C_4 photosynthesis genes. Plant Physiology, **111,** 949-957.

Lindskog S. (1997) Structure and mechanism of carbonic anhydrase. Pharmacological Therapeutics, **74,** 1-20.

Moroney J.V., Bartlett S.G., & Samuelsson G. (2001) Cabonic anhydrases in plants and algae (in press)

Price G.D. & Badger M.R. (1989) Expression of human carbonic-anhydrase in the cyanobacterium *Synechococcus* PCC7942 creates a high CO_2-requiring phenotype – evidence for a central role for carboxysomes in the CO_2 concentrating mechanism. Plant Physiology, **91,** 505-513.

Price G.D., von Caemmerer S., Evans J.R., Yu J.-W., Lloyd J., Oja V., Kell P., Harrison K., Gallagher A. & Badger M.R. (1994) Specific reduction of chloroplast carbonic anhydrase activity by antisense RNA in transgenic tobacco plants has a minor effect on photosynthetic CO_2 assimilation. Planta, **193,** 331-340.

Samuelsson G & Karlsson J (2000) Chloroplast carbonic anhydrases. In *Recent Advances on Photosynthesis* (in press)

Soltes-Rak E., Mulligan M.E. & Coleman J.R. (1997) Identification and characterization of a gene encoding a vertebrate-type carbonic anhydrase in cyanobacteria. Journal of Bacteriology, **179,** 769-774.

Xi, X. (1999) Structural and functional analysis of a gamma carbonic anhydrase from *Arabidopsis thaliana*. Masters thesis, Louisiana State University, Baton Rouge, Louisiana.