October 2007

Cloning and characterization of two putative glucuronolactonases from Arabidopsis thaliana involved in ascorbate degradation

Melinda Belisle
Worcester Polytechnic Institute

Cloning and characterization of two putative glucuronolactonases from *Arabidopsis thaliana* involved in ascorbate degradation

A Major Qualifying Project Proposal

Submitted to the Faculty of the

WORCESTER POLYTECHNIC INSTITUTE

in fulfillment of the requirements for the

Degree of Bachelor of Science

in

Biology and Biotechnology

by

_____
Melinda Belisle

October 16th, 2007

APPROVED:

_____	_____
Argelia Lorence, Ph.D.	Pamela Weathers, Ph.D.
Assistant Professor in Metabolic Engineering	Professor of Biology and
Arkansas Biosciences Institute	Biotechnology
Major Advisor	WPI Project Advisor

Abstract

Vitamin C (ascorbic acid, AsA) is the most abundant water-soluble antioxidant in plants and animals. There are four known pathways by which plants synthesize AsA: L-galactose, and D-galacturonate, -gulose, and *myo*-inositol. The inositol route involves four enzymes. The first two have been already characterized at the Lorence Laboratory. Using bioinformatic analysis the the third enzyme was identified from 18 candidate genes. Putative glucuronolactonases (GNLs) were identified in the *Arabidopsis thaliana* genome by comparing them to GNLs from rat and bacteria. Eighteen genes were selected as they shared conserved amino acid regions with the GNLs from rat and bacteria. The corresponding knockout lines were ordered from the Arabidopsis Biological Resource Center, and were grown in selective media along with wild type and β-glucuronidaseover-expressers as controls. After 5 weeks, rosette leaves were collected and the foliar AsA levels were measured via a spectrophotometric-based method. Unexpectedly two mutants were discovered which contained high foliar AsA levels compared to controls. We hypothesized that these two genes are involved in AsA degradation instead of synthesis and we have performed further studies to confirm their participation in AsA metabolism. This work focused on cloning of ORFs, making constructs to allow expression in bacteria and *in planta,* and transformation of wild type and Arabidopsis homozygous knockouts with appropriate constructs.

Table of Contents

List of Abbreviations

AsA- Ascorbic acid
DFP- Diisopropyl phosphorofluoridate
GlcUR- Glucuronolactonase reductase
GLOase- L-Gulono-1,4-lactone oxidase
GNL-Glucuronolactonase
H –Hours
L -Liters
LB- Luria-Bertani media
MIOX- *myo*-nositol oxygenase
Min - Minutes
ORF- Open reading frame
SMP30- Senescence marker protein 30
SOC- Bacterial growth media

List of Figures

Introduction

Vitamin C (ascorbate, AsA; Figure 1**)** functions as an antioxidant, a redox buffer and plays an important role in both plant and animal metabolisms and response to stress (Arrigone and DeTuillio, 2002). The AsA biosynthetic pathway in animals was described in the 1950s. Despite the fact that fruits and vegetables are the most important sources of this essential vitamin for humans and other primates, little is known about its biosynthesis in plants. Since 1998 four routes leading to AsA formation have been proposed: the D-mannose/L-galactose (Wheeler et al., 1998), L-gulose (Wolucka and Van Montagu 2003), D-galacturonate (Agius et al., 2003), and *myo*-inositol (Lorence et al., 2004) pathways (Figure 2). Four enzymes participate in the inositol route to vitamin C: *myo*-inositol oxygenase, glucuronate reductase, glucuronolactonase and L-gulono-1,4-lactone oxidase (Figure 2). The first two enzymes have been already characterized at the Lorence laboratory. My work will focus on the cloning and characterization of two putative glucuronolactonases involved in AsA degradation in *Arabidopsis thaliana*.

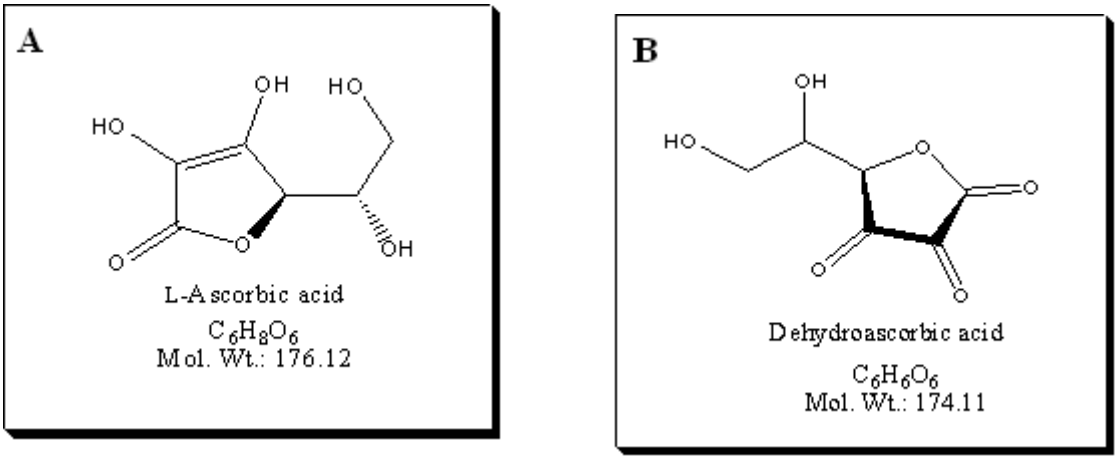

Figure 1: Structures of the reduced (L-ascorbic acid) and oxidized (dehydroascorbic acid) forms of ascorbate (AsA).

Background

Ascorbic acid biosynthesis in higher plants

L-ascorbic acid (AsA; Figure 1) has multiple roles in plant metabolism and plant stress responses (Ishikawa et al., 2006). Ascorbate is characterized as an antioxidant, a redox buffer, an enzyme cofactor, and is known to hydrolyze toxic compounds (Arrigone and De Tuillio, 2002; Kondo et al., 2004). Ascorbate performs its antioxidant function by scavenging reactive oxygen species which can have damaging repercussions to cells (Radzio et al., 2003). AsA has also been known to act in hormone responses, programmed cell death, senescence, pathogen responses, photosynthesis and respiration (Arrigoni and De Tullio, 2002). The synthesis of AsA has become especially important to primates because they have lost the ability to synthesize this multifunctional enzyme cofactor and antioxidant. The pathways to AsA production in plants are, therefore, important because fruits and vegetables are the chief supply of this vitamin for primates and many other animals.

The main substrate for the AsA biosynthetic pathway in mammals is UDP-D-glucose derived from glycogen. The main intermediates that participate in this route include: D-glucuronate, L-gulonate, and L-gulono-1,4-lactone (reviewed in Ishikawa et al., 2006; Figure 2).

Wheeler et al. (1998) proposed one of the main pathways of AsA production in plants. The so called Smirnoff-Wheeler pathway proceeds via GDP-mannose, GDP-L-galactose, L-galactose, and L-galactono 1,4-lactone (Figure 2). Although the Smirnoff-Wheeler pathway is the best characterized, several other routes for AsA biosynthesis have been proposed; these include: L-gulose (Wolucka and Van Montagu 2003), D-galacturonate (Agius et al., 2003), and *myo*-inositol (Lorence et al., 2004) routes (Figure 2).

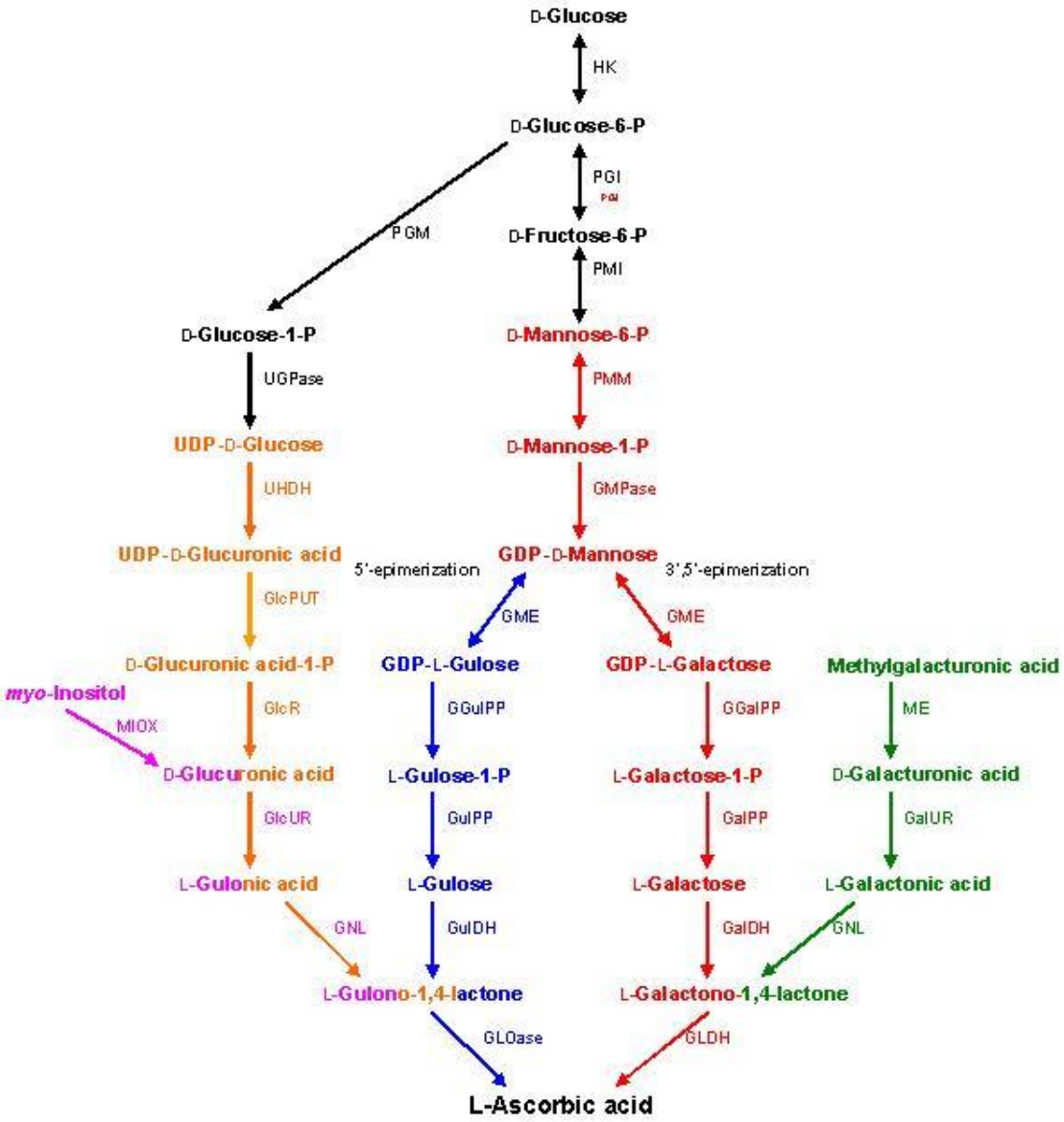

Figure 2: Ascorbic acid biosynthesis in plants and animals. Each pathway is presented in a different color. L-Galactose Pathway, Wheeler et al., 1998.D-Galacturonic Acid Pathway, Agius et al,. 2003. L-Gulose Shunt, Wolucka and Van Montagu, 2003. *Myo*-Inositol Pathway, Lorence et al., 2004. The AsA biosynthetic pathway in animals. Reactions are catalyzed by: (HK) hexokinase, (PGI) phosphoglucose isomerase, (PMI) phosphomannose isomerase, (PMM) phosphomannose mutase, (GMPase) GDP-D-mannose pyrophosphorylase, (GME) GDP-D-mannose-3,5-epimerase, (GGalPP) GDP-L-galactose pyrophosphatase, (GalPP)L-galactose-1-phosphate phosphatase, (GalDH) L-galactose dehydrogenase, (GLDH) L-galactono-1,4-lactone dehydrogenase, (GGulPP) GDP-L-gulose pyrophosphatase, (GulPP) L-gulose-1-phosphate phosphatase, (GuLDH) L-gulose dehydrogenase, (GLOase) L-gulono-1,4-lactone dehydrogenase, PGM phosphoglucose murase, (UGPase) UDP-glucose pyrophosphorylase, (UHDH) UDP-glucose dehydrogenase, (GlcPUT) glucuronate-1-phosphate, (GlcR) glucurono kinase, (MIOX) *myo*-inositol oxidase, (GlcUR) glucuronic acid reductase, (GNL) glucuronolactonase, (ME) methylesterase, (GalUR) galacturonic acid reductase.

The *myo*-inositol pathway

The *myo*-inositol pathway was recently proposed as a possible biosynthetic route to AsA in plants. Four enzymes participate in this pathway: *myo*-inositol oxygenase (MIOX), glucuronate reductase (GlcUR), glucuronolactonase (GNL), and L-gulono-1,4-lactone oxidase (GLOase) (Lorence et al., 2004; Figure 2). Taking advantage of a fully sequenced and annotated genome, Dr. Lorence and her collaborators have performed bioinformatic analyses that allowed them to identify four MIOXs, four GlcUARs, eighteen putative GNLs and seven putative GLOase's in *Arabidopsis thaliana*. All four MIOXs, and one of the GlcUARs have been already characterized (Radzio et al., 2003; Lorence et al., 2004; Lorence and Nessler, 2007).

Animal glucuronolactonase

A protein called senescence marker protein 30 (SMP30) was recently discovered in rats. It was determined that this protein is actually glucuronolactonase (GNL), one of the enzymes involved in AsA biosynthesis (Kondo et al., 2006; Figure 3). It was observed that the expression of *SMP30* decreases in an androgen-independent manner with aging (Ishigami et al., 2002; Kondo et al., 2004, 2006). It is because of this characteristic that SMP30 has been used as a marker of the aging (senescence) process in animals. Recently, Kondo et al. (2006) showed that SMP30 exhibits GNL activity and that *SMP30* knockout mice have a shorter life span than wild type mice and are prone to scurvy. In addition, GNL was recognized by an antibody directed to rat SMP30, and GNL purified from mice liver was identical to that of the reported sequence of rat SMP30 (Kondo et al., 2006). Absence of GNL leads to decreased degradation of AsA because GNL hydrolyzes the lactone ring of dehydroascorbic acid (Figure 1B), the oxidation

product of AsA (Kondo et al., 2006). These data paved the way for further research into the functions of GNL in AsA plant metabolism.

The fact that GNL may be involved in the degradation of AsA to oxalic and tartaric acids is of particular interest. Nakata and McConn (2007) and Debolt et al. (2006) reviewed the evidence indicating that AsA is a precursor to oxalic and tartaric acids in some plants. Debolt et al. (2006) found that a relatively high amount of AsA is converted to oxalic and tartaric acids which indicates that AsA catabolism can be high when necessary. Their data indicated that over 70% of radiolabelled AsA was recovered as either oxalic or tartaric acids. We hypothesize that similar to the rat GNL (Kondo et al., 2006), some of the Arabidopsis GNLs may catalyze one of the reactions involved in this process. Through radiolabelling Kondo et al. also discovered that when AsA is cleaved between carbon atoms 2 and 3, oxalic acid is formed; if AsA is cleaved between carbon atoms 4 and 5, tartaric acid is produced.

Nakata and McConn (2007) worked with Arabidopsis mutants that over-accumulated crystals of calcium oxalate (oxalic acid). In accordance with the data of Debolt et al. (2006), Nakata and McConn discovered that mutants that over-accumulated oxalate had lower AsA content compared to wild type plants indicating that the more oxalate formed, the less AsA accumulates. Together these results strongly suggested that AsA is a biosynthetic precursor to oxalic and tartaric acid (Figure 3). These data support our hypothesis that GNL is involved in the degradation of AsA into oxalic and tartaric acid.

Another interesting fact about Arabidopsis GNLs is that they may hydrolyze diisopropyl phosphorofluoridate. Ishigami et al. (2002) and Kondo et al. (2004) found that *SMP30* knockout mice were more prone to TNF-α and Fas-mediated apoptosis than wild type mice. TNF-α and Fas-mediated apoptosis are two examples of direct initiation of programmed cell death because

the TNF and Fas receptors directly activate apoptic pathways. *SMP30* knockout mice also had reduced ability hydrolyzing diisopropyl phosphorofluoridate (DFP) in the liver (Kondo et al 2004). DFP is a compound resembling chemical warfare nerve agents such as sarine, soman and tabun which are organophosphorous cholinesterase inhibitors. These compounds inhibit tissue enzymes by permanently binding to the enzyme and are soluble in fat and water. The effects of these nerve agents include loss of consciousness, seizures, muscular flaccidity and respiratory failure (Holstege et al., 1997; Dunn et al., 1987).

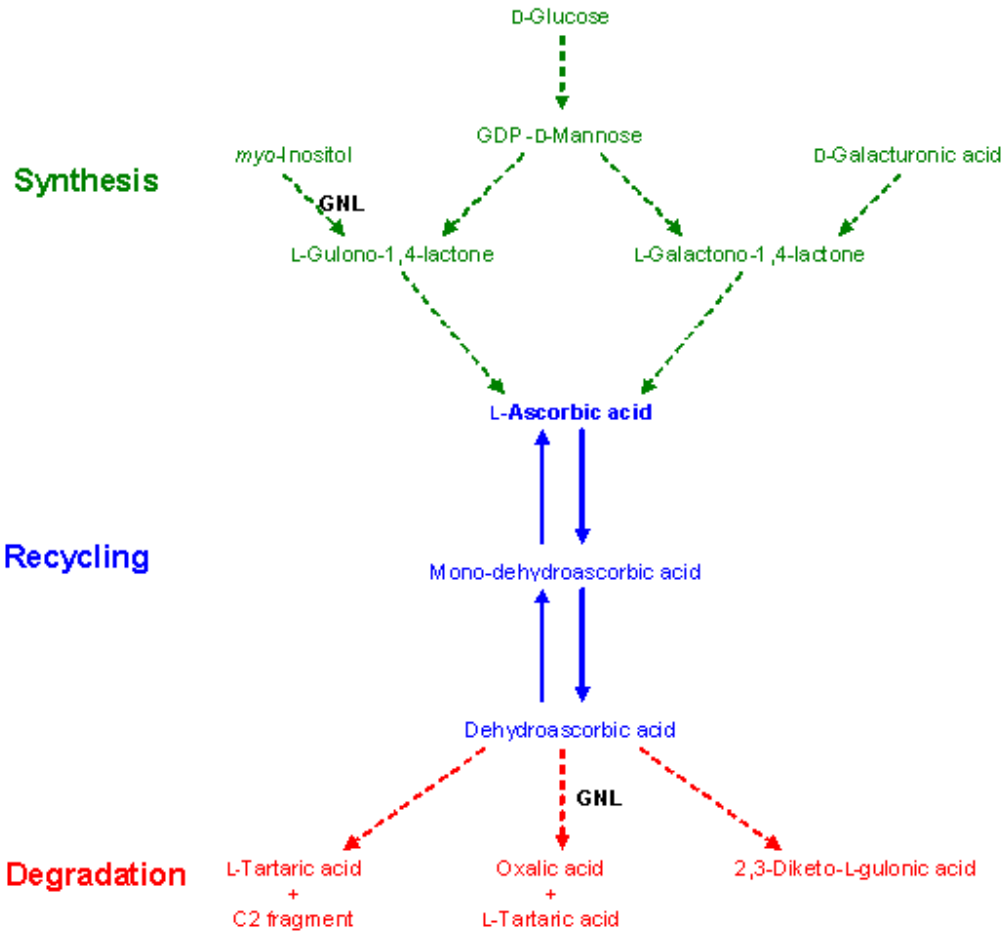

Figure 3: Ascorbic acid biosynthesis and catabolism in plants. Similar to animal models, various Arabidopsis glucuronolactonases (GNLs) may be involved in both the synthesis and degradation of AsA. Solid lines represent single reactions; dashed lines represent reactions catalyzed by multiple enzymes.

Preliminary results on the identification of Arabidopsis GNLs

GNLs have been studied in bacteria (Hucho and Wallenfels, 1972; Zachariou and Scopes, 1986; Kanagasundaram and Scopes, 1992), fungi (Brodie and Lipmann, 1955; Witteven et al., 1993), and animal models (Roberts et al., 1978; Bailey et al., 1979; Kondo et al., 2006 and references therein), however, there are no reports of its enzymatic activity in plants. Bioinformatic analyses, therefore, were conducted to identify this enzyme among candidate genes. By using sequences of GNLs from rat, *Zymomonas* and *Nostoc*, eighteen genes were selected as they shared conserved amino acid regions (motifs) with the homologs from rat and bacteria (Lorence, unpublished). These genes were identified out of 28,607 in the *Arabidopsis thaliana* genome.

Researchers at the Salk Institute have developed a large collection of Arabidopsis mutants in which individual genes are interrupted by a T-DNA. These mutants are called 'knockouts' due to the lack of function of the interrupted gene and are available to the scientific community via the Arabidopsis Biological Resource Center (Alonso et al., 2003). Seeds of the 18 mutants from the Salk collection corresponding to the candidate genes for GNL were ordered, planted, and grown under controlled conditions (23ºC, 16/8 h of light/dark, 150 µmol m^{-2} $s^{-1}$and 65% humidity). Seeds were planted on selective MS medium (Murashige and Skoog, 1962) supplemented with kanamycin (50 mg/) and subsequently transferred to soil. After five weeks of growth, the healthy rosette leaves were collected and their AsA content was measured by two methods, a spectrophotometric-based method and an HPLC-based method. This screening strategy allowed the Lorence laboratory to identify five knockouts with significantly lower AsA

content (~50%) compared to wild type plants: Salk lines 001853, 025037, 009989, 124438 and 091299).

Surprisingly, they also found two knockouts with elevated AsA (Salk lines 026466 and 057855). We hypothesized that these knockouts are defective in an enzyme involved in AsA degradation (Figure 3). My research focused on testing this hypothesis. The cDNAs corresponding to these two genes (At3g51420 and At3g51440) were amplified. I then cloned them into a TA-cloning vector (pDrive) and sequenced both DNA strands, and made the genetic constructs to allow production of recombinant proteins in *Escherichia coli* (in pET41) and to allow for expression in plants (in pCAMBIA1300). I also made transgenic Arabidopsis to test the potential of these plants to over-express enzymes which can be used as potential resources for phytoremediation of soil contaminated with warfare agents such as sarine, soman, and tabun.

Materials and Methods

cDNA amplification from total RNA

The cDNAs corresponding to the putative GNLs At3g51420 and At3g51440 were amplified by PCR from stem cDNA. A 25 µL PCR reaction containing 3 µL of cDNA as a template was performed using illustra PuReTaq Ready-To-Go PCR beads (GE Healthcare, Piscataway, NJ). The first 20-25 bases of the gene of interest were copied and an *NcoI* site was added to ATG. The last 20-25 bases of the gene of interest were also copied, the reverse compliment was

obtained and a *Bam*HI restriction site was added after the stop codon. The primers are as follows: At3g51440Nco CCCATGGCTGTAT TCCTCTCTTC, At3g51440Bam CGGATCC TCAAAGCTTCTTGTTTTG, 51420Nco CCCATGGTTCTCT TCTTCTCCAC, and 51420Bam GGATCC TTAAAGCTTCTTTTGTGCGG. Samples were denatured at 94ºC for 4 min. Amplification was as follows: 30 cycles of 94ºC for 4 min, an annealing temperature of 60ºC for 1 min for At3g51420, and 56ºC for 1 min for the gene At3g51440, 72 ºC for 90 s, and lastly 72 ºC for 10 min. Samples were then run through a 1% agarose gel for confirmation of PCR amplification.

Cloning of At3g51420 and At3g51440 ORFs

Bands of the PCR products were cut out of the 1% agarose gel and cleaned using the QIAEX II agarose gel extraction protocol (Qiagen Sciences, Montgomery County, Maryland). All gels contained ethidium bromide and were visualized using a transilluminator (ultraviolet light box). Resulting products were ligated into pDrive cloning vector (Qiagen), following the protocol provided by the manufacturer. Thunderbolt GC10 electrocompetent *E. coli* cells (Sigma, St. Louis, MO) were transformed with the ligation reaction products and plated on LB medium plus kanamycin (50 mg/mL).

Positive colonies were grown overnight in 3 mL cultures of LB and kanamycin selective media. A positive colony is a colony which had acquired the insert and in turn acquired resistance to kanamycin; only positive colonies grew on the kanamycin selective media. Minipreps were performed on the cultures using the Qiaprep spin kit (Qiagen).A miniprep is a

procedure used to extract plasmid DNA from bacterial cell suspensions A 3 µL aliquot of the DNA obtained was digested with *Eco*RI and run through a 1% agarose gel to verify that the insert was in place in the pDrive vector. The constructs, At3g51420: pDrive and At3g51440:pDrive were verified by sequence analysis done by the University of Chicago Cancer Research Center DNA Sequencing and Genotyping Facility and results were analyzed using the Sequencher software. Both strands were sequenced using Sp6 (5'GATTTAGGTGACACTATAG 3'), and T7 (5'TAATACGACTCACTATAGGG3') primers.

Construct for expression of recombinant protein in *E.coli*

The cDNAs were subcloned from pDrive into pET41 (Novagen, Madison, WI). To do this, all the samples and the empty pET41 vector were digested with *NcoI* and *Bam*HI overnight. These samples were then run on a 1% agarose gel. The insert and pET41 bands were cut and extracted using the QIAEX II Agarose Gel Extraction protocol (Qiagen). The inserts were then ligated to the empty pET41 vector overnight.

Thunderbolt GC10 and DH5α electrocompetent *E.coli* cells (Sigma) were transformed with inserts that had been ligated into pET41. A 3 µL aliquot of DNA was added to 40 µL of electrocompetent *E. coli* cells and loaded into a VWR electroporation cuvette. The cells were shocked and then placed in 500 µL of SOC medium to recover for 1.5 h at 37ºC. Recipe for 1000 mL of SOC media is as follows: 900 mL of dH_2O, 20 g bacto Tryptone, 5 g bacto yeast extract, 2 mL of 5M NaCL, 2.5 mL of 1M KCL, 10ml of 1M $MgCl_2$, 10mL of 1M $MgSO_4$, 20mL of 1M

glucose adjust to 1000 mL with dH_2O and autoclave. The samples were then plated on LB + kanamycin (50 mg/mL) medium.

Qiaprep spin kit (Qiagen) minipreps were performed the next day using the protocol provided by the manufacturer. A double digestion of *NcoI* and *Bam*HI was done to confirm that the vector and insert were present. The DNA saved from the minipreps was used to transform BL21 chemically competent cells (Novagen) which are very efficient at making recombinant proteins.

Aliquots (40 µL) of BL21 chemically competent *E.coli* cells were added to 3 µL of DNA from each sample. The cells were incubated on ice for 30-40 min and then at 37ºC for 5 min. The cells were then returned to ice and 500 µL of SOC medium was added. The cells were left to recover and incubated for at least one hour at 37ºC in a shaker at 200 rpm. Cells were then plated on LB+ kanamyacin media (50 mg/mL).

Constructs for expression in plants

Constructs of At3g51420 and At3g51440 were made in the binary vector pCAMBIA1300 driven by the double-enhanced 35S promoter as described previously (Lorence et al., 2004). These constructs were introduced into *Agrobacterium tumefaciens* strain GV-3101 and *Arabidopsis thaliana* var. Columbia wild type plants (CS6000 stock from ABRC) were transformed using the floral dipping method (Clough and Bent, 1998).

Results and Discussion

Initially At3g51420 and At3g51440 ORFs were amplified using *A. thaliana* stem cDNA as a template. Results on an agarose gel showed expression of both constructs (Figure 4). After PCR amplification, constructs of ORFs At3g51420 and At3g51440 were made in pDrive (Figure 5) and sent to the Core Facility at the University of Chicago for DNA sequencing.

Amplification of cDNA

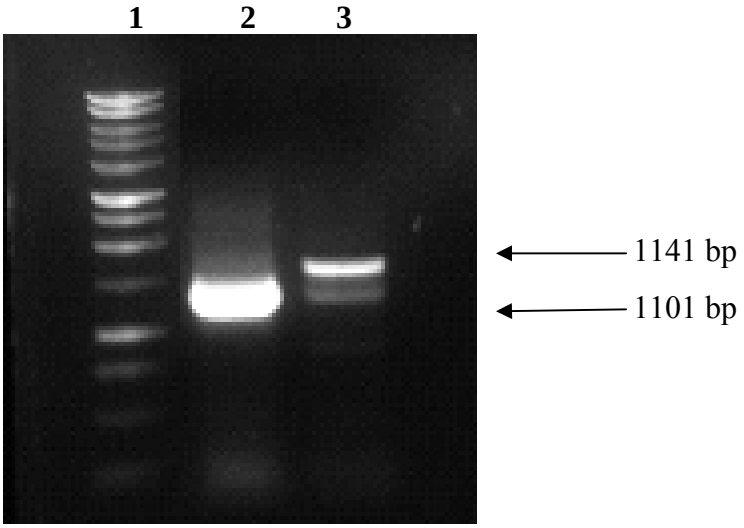

Figure 4: PCR products obtained for ORFs using cDNA from *A. thaliana* stems as a template. Lane1-MWM 1 kb ladder(Promega), Lane 2-At3g51420**,** Lane 3-At3g51440.

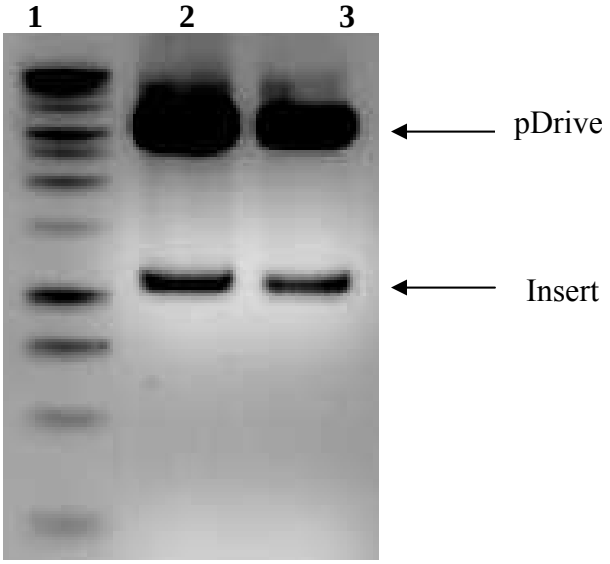

Figure 5: pDrive constructs. A 20 µL *EcoR1* digestion reaction was performed. Lane1-MWM 1 kb ladder (Promega), Lane 2-At3g51440B:pDrive, Lane 3-At3g51420:pDrive

Sequencing results revealed that the At3g51440B ORF was a splicing variant and that there was one intron which had not been out spliced. This intron yielded a truncated version of the gene. The first 278 amino acid residues of the full length protein (371 amino acids) were amplified. However we will test if this truncated gene we amplified contains the active domain of the protein and is still functional.

Another sample of At3g51440 which we called At3g51440A had several introns which were not spliced out. We attribute this result to genomic DNA left over in the total RNA that was used to make cDNA and had contaminated the sample. Since At3g51440A had several introns we chose to work instead with At3g51440B. To correct the problem of introns, a transient expression system which uses tobacco plants to make the protein of interest will be used.

Sequence analysis

```
Query  49    PCLISITLYQLDSFEPASLPADSLITSPTSIPPLLNDRFLTGAEFIGVGLLNNPEDIAYH  228
             PCLISITLYQLDSFEPASLPADSLITSPTSIPPLLNDRFLTGAEFIGVGLLNNPEDIAYH
Sbjct  17    PCLISITLYQLDSFEPASLPADSLITSPTSIPPLLNDRFLTGAEFIGVGLLNNPEDIAYH  76

Query  229   KDSNLTYTGCVDGWVKRVSVHDSANDSIVEDWVNTGGRPLGIAFGLHGEVIVADANKGLL  408
             KDSNL YTGCVDGWVKRVSVHDSANDSIVEDWVNTGGRPLGIAFGLHGEVIVADANKGLL
Sbjct  77    KDSNLIYTGCVDGWVKRVSVHDSANDSIVEDWVNTGGRPLGIAFGLHGEVIVADANKGLL  136

Query  409   SISDGGKKTELLTDEADGVRFKLTDAVTVADNGVLYFTDASSKYDFYQFIFDFLEGKPHG  588
             SISDGGKKTELLTDEADGVRFKLTDAVTVADNGVLYFTDASSKYDFYQFIFDFLEGKPHG
Sbjct  137   SISDGGKKTELLTDEADGVRFKLTDAVTVADNGVLYFTDASSKYDFYQFIFDFLEGKPHG  196

Query  589   RVMSFDPTTRATRVLLKDLYFANGISMSPDQTHFVFCETIMRRCSKYYISEERVEVFIQG  768
             RVMSFDPTTRATRVLLKDLYFANGISMSPDQTHFVFCETIMRRCSKYYISEERVEVFIQG
Sbjct  197   RVMSFDPTTRATRVLLKDLYFANGISMSPDQTHFVFCETIMRRCSKYYISEERVEVFIQG  256

Query  769   LPGYPDNIRYDGDGHYWIALISEVTTSWKLSMKYLFLRKLIYMAAKYGVELLSIKNAAVL  948
             LPGYPDNIRYDGDGHYWIALISEVTTSWKLSMKYLFLRKLIYMAAKYGVELLSIKNAAVL
Sbjct  257   LPGYPDNIRYDGDGHYWIALISEVTTSWKLSMKYLFLRKLIYMAAKYGVELLSIKNAAVL  316

Query  949   QVDLDGNPIAMYHDHPSSHITSGVKIGNHLYFGSLLHSYITRLDLLKYPAQKKL  1110
             QVDLDGNPIAMYHDHP SHITSGVKIGNHLYFGSLLHSYITRLDLLKYPAQKKL
Sbjct  317   QVDLDGNPIAMYHDHPFSHITSGVKIGNHLYFGSLLHSYITRLDLLKYPAQKKL  370
```

Figure 6: At3g51420 ORF cloned from stem cDNA. Query indicates both strands (forward and reverse) of DNA of At3g51420 ORF which were sequenced. The results were analyzed using Sequencher software. Differences in amino acids of the ORF I cloned were compared to the known At3g51420 sequence (GenBank Accession Numbers: NM_115001 and BT024894) indicated by the word Sbjct and are highlighted in yellow.

Two differences were observed At3g51420 ORF when compared to the published sequence. At amino acid Sbjct 83 the published sequence indicates an isoleucine where we detected threonine. These amino acids have similar structures but threonine is hydrophilic while isoleucine is hydrophobic so this may cause disruption of the enzyme activity. At Sbjct 334 there is conversion of phenylalanine to a serine which is similar in size but there is once again a change in hydrophobicity, but instead a hydrophobic amino acid is converted to a hydrophilic one.

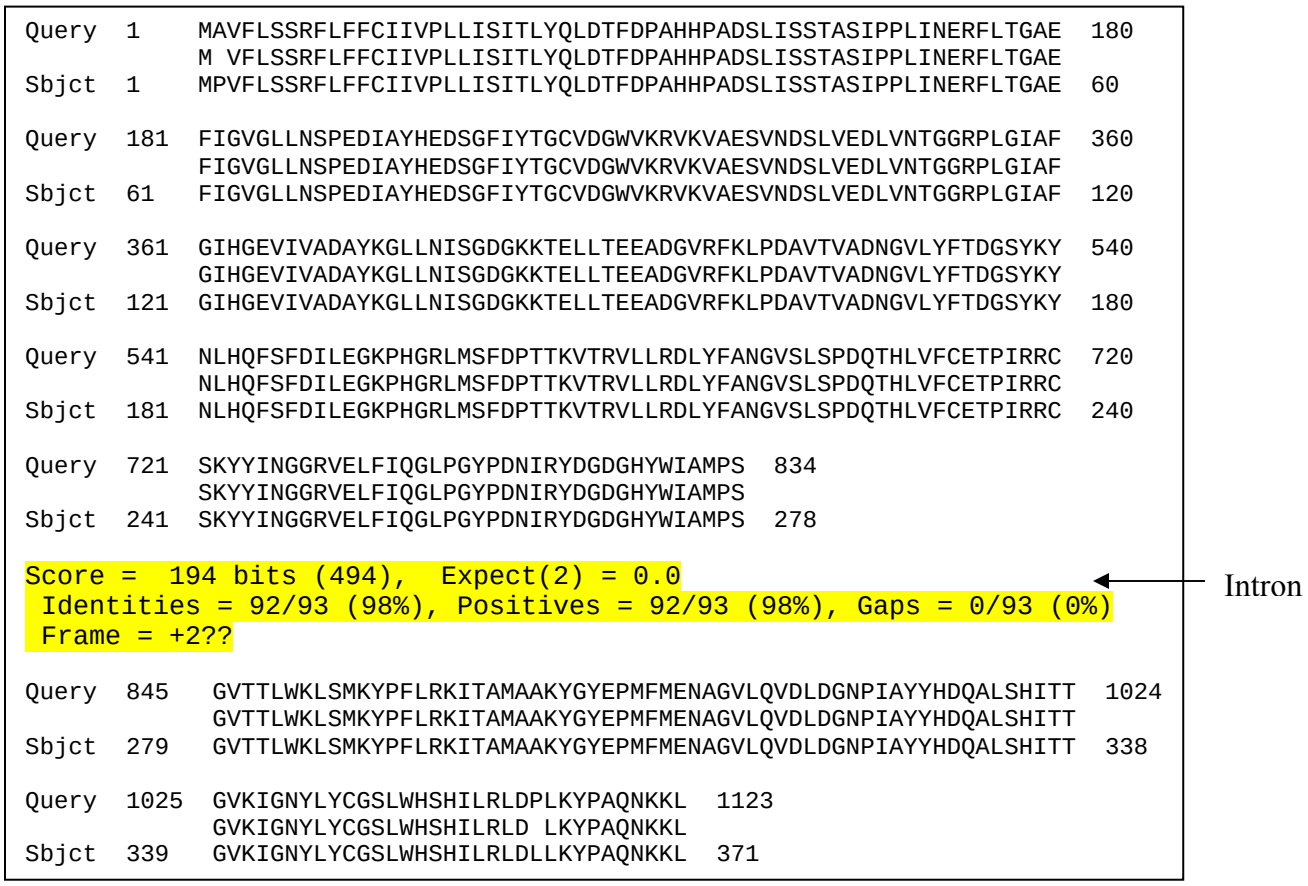

Figure 7: At3g51440 ORF cloned from stem cDNA. Both strands of DNA of At3g51440 ORF were sequenced and the results were analyzed using Sequencher software. Differences in amino acids of the ORF we cloned were compared to the known At3g51440 sequence (Genbank accession number: NM_115001) and are highlighted in yellow. Query indicates forward and reverse strands of ORF while Sbjct indicates published sequence.

The At3g51440 ORF showed no differences from the published sequence with the exception of an intron which has been discussed previously. To obtain the possible sequence of plant GNL comparisons were made to the amino acid sequences of rat SMP30, and the GNLs of *Nostoc* and *Zymomonas* . Comparison revealed several homologous regions in these published sequences (Kondo et al., 2006) to our hypothetical plant GNL. Comparison with published GNL sequence in Genbank supports the argument that these two genes are putative GNLs.

DNA sequence results confirmed that we had amplified the correct genes so we moved forward with the cloning strategy (Figure 6). The cloning strategy involved sub-cloning from pDrive into pRTL2 and then into pCAMBIA1300 for expression in plants (Figures 7A and 7B), as well as subcloning from pDrive into pET41 for protein expression (Figure 8).

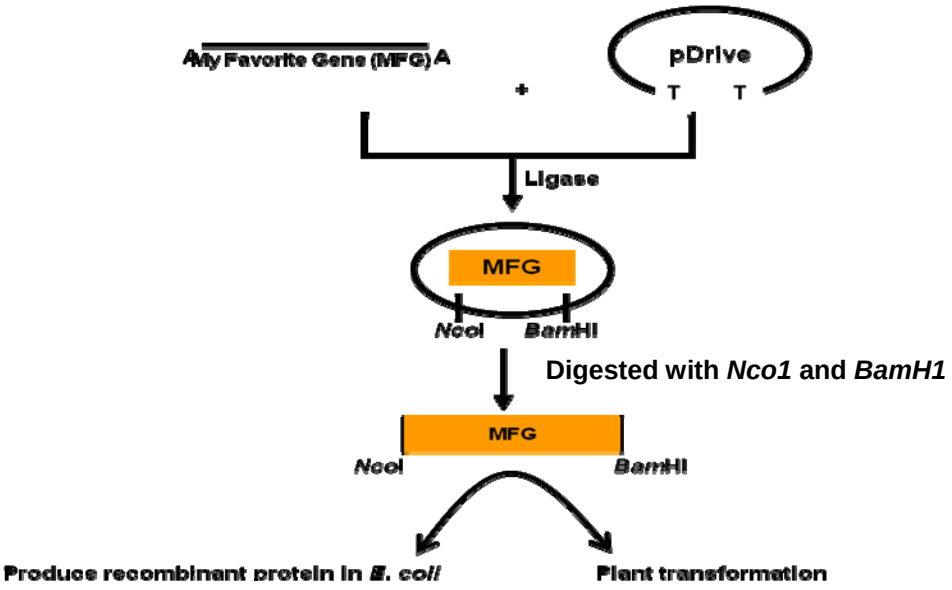

Figure 8: Cloning strategy

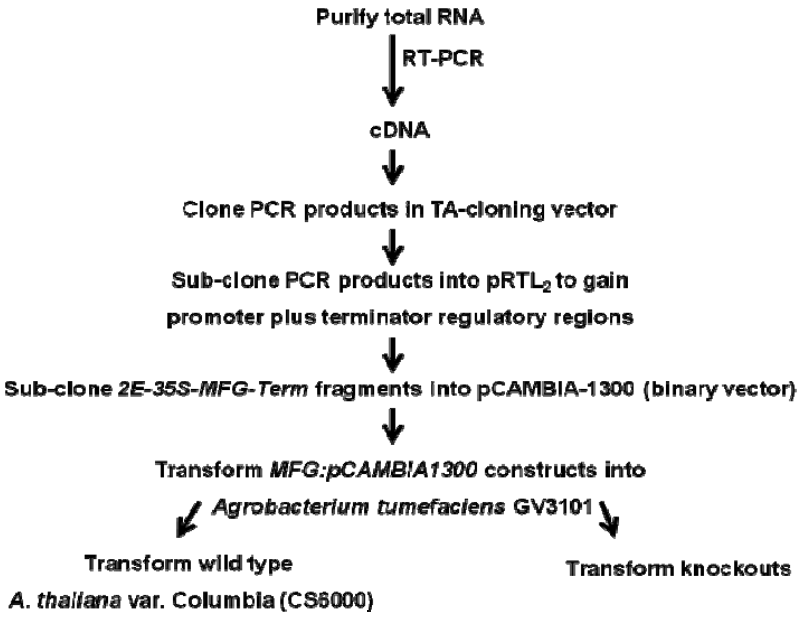

Figure 9A: Cloning strategy for expression in plants

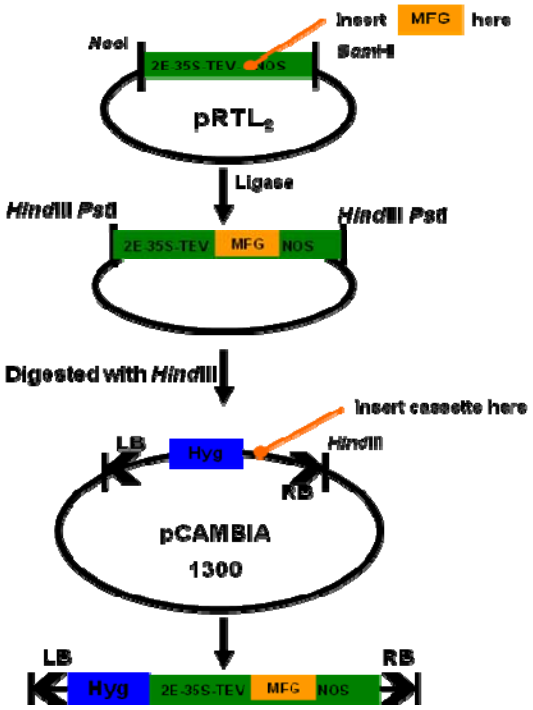

Figure 9B: Detailed cloning strategy for expression in plants

To transform plants we first sub-cloned our genes of interest into the pRTL2 vector. To do this we cut both the vector and the inserts with *NcoI* and *BamHI* restriction enzymes. After extracting the samples from an agarose gel, we ligated the inserts into the vector using ligase. Inserts were ligated into pRTL2 to obtain the standard cassette containing the 35S promoter and NOS terminator.

The pRTL2 constructs were digested with *PstI* restriction enzyme which allows for extraction of the cassette 35S-TEV:At3g51420:NOS. The pRTL2 vector is now 2.3 kb from the original 3.3 kb losing the 2E-35S-TEV promoter and NOS terminator. The inserts went from approximately 1 kb to 2 kb with the addition of the cassette. The pCAMBIA1300 binary vector was also digested with *PstI* and ligated to each one of the inserts. *Agrobacterium tumefaciens* strain GV3101 was transformed with the pCAMBIA1300 constructs and in turn plants were transformed with *Agrobacterium* via the floral dipping method (Clough and Bent, 1998).

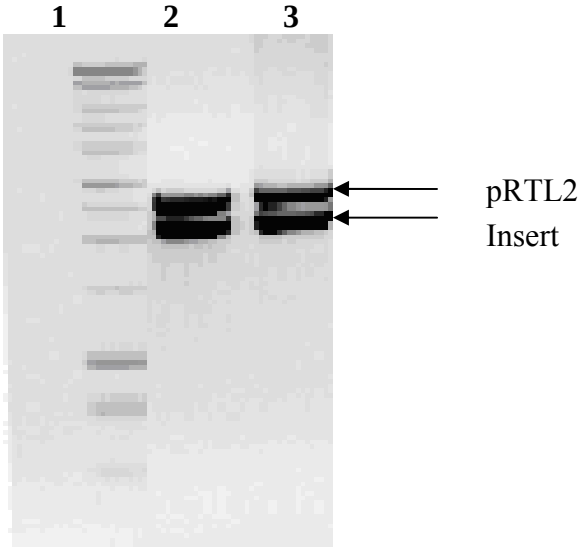

Figure 10: pRTL2 constructs. These samples were digested overnight at 37ºC using *PstI* restriction enzyme. Lane 1-MWM1kb ladder (Promega), Lane 2-At3g51440B:pRTL2, Lane 3-At3g51420:pRTL2

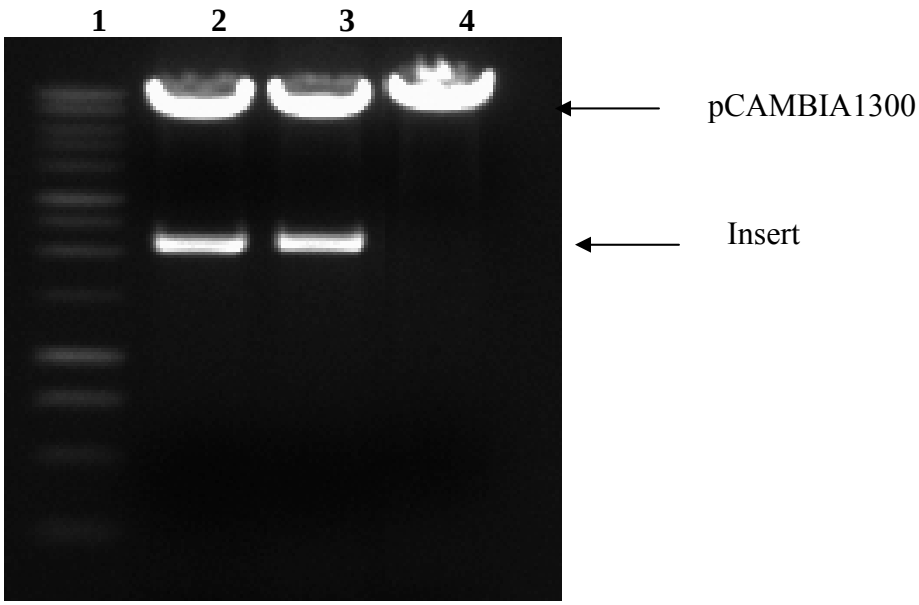

Figure 11: pCAMBIA constructs. A 20 µL digestion reaction using *Pst1* was performed. Lane 1-MWM 1 kb ladder (Promega), Lane 2-At3g51420:pCAMBIA1300, Lane 3-At3g51440B:pCAMBIA1300, Lane 4-pCAMBIA1300 empty vector

Lastly we made constructs for protein expression in the pET41 vector.

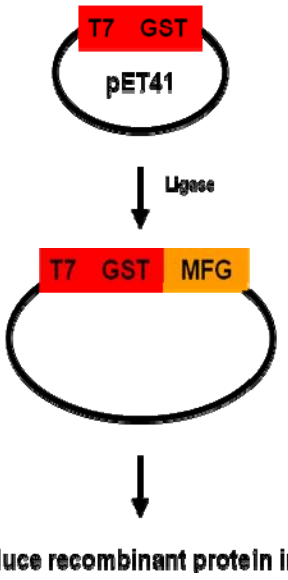

Figure 12: Cloning strategy for protein expression

Constructs in pDrive as well as the pET41 empty vector were digested with *Nco*I and *BamHI* restriction enzymes. The inserts were then ligated into the empty vector.

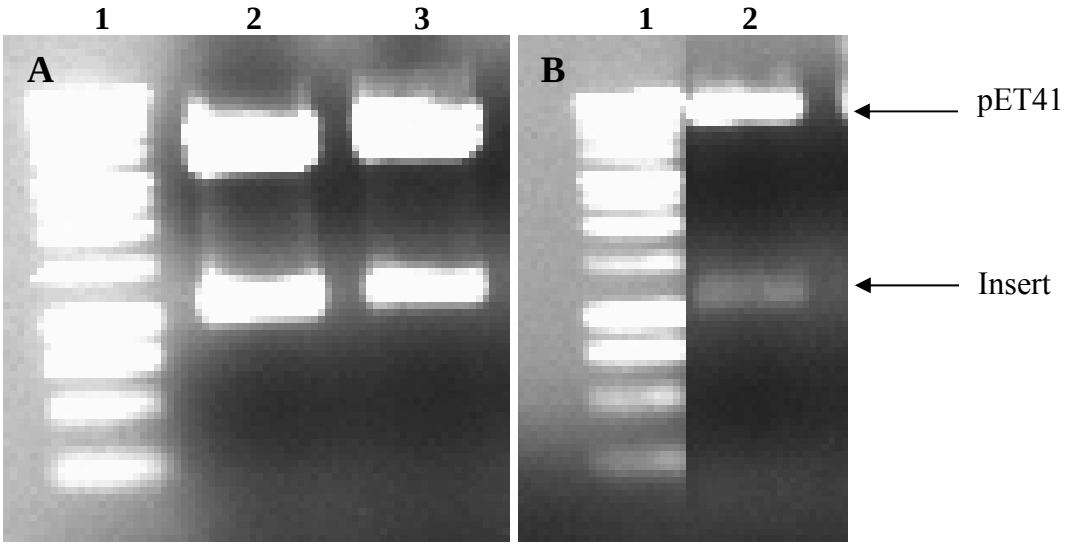

Figure 13: pET41 constructs. An *Nco1/BamH1* digestion was performed on these samples. (**A)** Lane 1-MWM 1 kb ladder (Promega), Lane 2-At3g51440B:pET41, Lane 3-At3g51440B:pET41 (**B**) Lane 1-MWM, Lane 2-At3g51420:pET41

All constructs for production of recombinant proteins *in vitro* and for gene expression *in planta* were successfully made (Figures 5, 8, 9 and 11). These constructs are currently being used for the production of recombinant proteins in *E. coli.* The Lorence lab will then explore a transient expression system and test the activity of these enzymes *in planta* using L-gulonate, dehydroascorbate, and DFP as substrates. This project has paved the way for further investigation of putative GNLs.

I also successfully used the pCAMBIA1300 constructs to transform *Agrobacterium tumefaciens* strain GV3101 and in turn transformed Arabidopsis knockout and wild-type plants with the *Agrobacterium* (Clough and Bent, 1998). We expect the knockouts to have restored phenotypes and the wild-type plants should now over-express the gene of interest. In addition, to test the involvement of GNLs in vitamin C degradation, we will measure AsA, oxalate and

tartarate levels in tissues of Arabidopsis wild type and knockout plants via HPLC. These plants will be used to test the activity of proposed substrates *in planta.* These transgenic Arabidopsis can then be used to test the potential of plants over-expressing these enzymes as resources for phytoremediation of soil contaminated with warfare agents such as sarine, soman, and tabun which are analogous to DFP (Kondo et al., 2004).

Acknowledgements

I would like to thank my Major Advisor, Dr. Argelia Lorence for allowing me to work in her lab at the Arkansas Bioscience Research Institute, and for all of her help throughout the entire project. I would also like to thank the students in Dr. Lorence lab, especially Gwendolyn Wilson my lab partner who worked alongside me throughout the project and who helped create some of the figures. Guillermo Tujullio who collaborated with us on the project and lastly, I am grateful for the guidance of my WPI Project Advisor, Dr. Pam Weathers who helped me initiate the project and prepare my final MQP report.

Literature Cited

Agius, F., Gonzalez-Lamothe, R., Caballero, J.L., Munoz-Blanco, J., Botella, M.A., Valpuesta, V. (2003) Engineering increased vitamin C levels in plants by overexpression of a D-galacturonic acid reductase. *Nat. Biotechnol.* 21: 177-181.

Alonso, J., Stepanova, A., Leissee, T., Kim, C., Chen, H., Shinn, P., Stevenson, D., Zimmerman, C., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H.., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W., Berry, C., Ecker, J. (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301: 653-657.

Arrigoni, O., De Tillio, M.C. (2002) Ascorbic acid: much more than just an antioxidant. *Biochem. Biophys. Acta.* 1569: 1-9.

Bailey, G.D., Roberts, B.D., Buess, C.M., and Carper, W.R. (1979). Purification and partial characterization of beef liver gluconolactonase. *Arch. Biochem. Biophys.* 192: 482-488.

Brodie, A. and Lipmann, F. (1955) Identification of a gluconolactonase. *J. Biol. Chem.* 212: 677-685.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana. Plant J.* 16: 735-743.

Debolt, S., Melino, V., Ford, C.M. (2006) Ascorbate as a biosynthetic precursor in plants. *Annals Botany* 99:3-8.

Dunn, M.A., Hackley, B.E., Sidell, F.R. (1987) Pretreatment for nerve agent exposure. *Medical Aspects of Chemical and Biological Warfare*: 181-196

Holstege, C.P., Kirk, M., Sidell, F.R. (1997) Chemical warfare. Nerve agent poisoning. *Crit Care Clin* 13(4):932-42

Hucho, F., Wallenfels, K. (1972) Gluconolactonase from *Escherichia coli*. *Biochim. Biophys. Acta* 276: 176-179.

Ishigami, A., Fujita, T., Handa, S., Shirasawa, T., Koseki, H., Kitamura, T., Enomoto, N., Sato, N., Shimosawa, T., Naoki, M. (2002) Senescence marker protein-30 knockout mouse liver is highly susceptible to tumor necrosis factor-α- and Fas-mediated apoptosis. *Am. J. Pathol.* 161: 1273-1281.

Ishikawa, T., Dowdle, J., Smirnoff, N. (2006) Progress in manipulating ascorbic acid biosynthesis and accumulation in plants. *Physiol. Plant.* 126: 343-355.

Kanagasundaram,V, and Scopes, R. (1992) Isolation and characterization of the gene encoding gluconolactonase from *Zymomonas mobilis. Biochem. Biophys. Acta* 1171: 198-200.

Kondo, Y., Ishigami, A., Kubo, S., Handa, S., Gomi, K., Hirokawa, K., Kajiyama, N., Chiba, T., Shimokado, K., Maruyama, N. (2004) Senescence marker protein 30 is a unique enzyme that hydrolyzes diisopropyl phophorofluoridate in the liver. *FEBS* 570: 57-62.

Kondo, Y., Inai, Y., Sato, Y., Handa, S., Kubo, S., Shimokado, K., Goto, S., Nishikimi, M., Maruyama, N, Ishigami, A. (2006) Senescence marker protein 30 functions as gluconolactonase in L-ascorbic acid biosynthesis, and its knockout mice are prone to scurvy. *Proc. Natl. Acad. Sci. USA* 103: 5723-5728.

Lorence, A., Chevone, B.I., Mendes, P., Nessler, C.L. (2004) *myo*-Inositol oxygenase offers a possible entry point into plant ascorbate biosynthesis. *Plant Physiol.* 134: 1200-1205.

Lorence, A. and Nessler, C.L. (2007). Pathway engineering of the plant vitamin C metabolic network. In "Applications of Plant Metabolic Engineering" R Verpoorte, AW Alfermann and TS Johnson (eds). Springer, chapter 8, 189-209.

Murashige, T., Skoog, F.A. (1962) A revised medium for rapid growth and bioassays with tobacco culture. *Plant Physiol.* 15: 473-497

Nakata, P.A., McConn, M. (2007) Isolated *Medicago truncatula* mutants with increased calcium oxalate crystal accumulation have decreased ascorbic acid levels. *Plant Physiol. and Biochem*. 45: 216-220.

Radzio, J.A., Lorence, A., Chevone, B.I., Nessler C.L. (2003) L-Gulono-1,4-lactone oxidase expression rescues vitamin C- deficient *Arabidopsis (vtc)* mutants. *Plant Mol. Biol.* 53: 837-844.

Roberts B.D., Bailey, G.D., Buess, C.M., and Carper, W.R. (1978) Purification and characterization of hepatic porcine gluconolactonase. *Biochem. Biophys. Res. Commun.* 84: 322-327.

Wheeler, G.L., Jones M., Smirnoff N. (1998) The biosynthetic pathway of vitamin C in higher plants. *Nature* 393: 365-369.

Witteven, C., van de Voondervoort, P., van den Broeck H., van Engelenburg, F., de Graaff, L., Hillebrand, M., Schaap, P., Visser, J. (1993) Induction of glucose oxidase, catalase, and lactonase in *Aspergillus niger*. *Curr. Genet.* 24: 406-416.

Wolucka, B.A., Van Montagu, M. (2003) GDP-Mannose 3',5'-epimerase forms GDP-L-gulose, a putative intermediate for the *de novo* biosynthesis of vitamin C in plants. *J. Biol. Chem.* 278: 47483-47490.

Zacharias, M., and Scopes, R.K. (1986) Glucose-fructose oxidoreductase, a new enzyme isolated from *Zymomonas mobilis* that is responsible for sorbitol production. *J. Bacteriol.* 167: 863-869.