Characterization of the putative glyoxysomal protease GX06 in *Arabidopsis thaliana*

Sherry Shen
Honors Thesis

Cellular & Molecular Biology
Department of Molecular, Cellular & Developmental Biology
University of Michigan
April 2, 2012

Sponsor: Laura J. Olsen, Ph.D.
Readers:
Erik Nielsen, Ph.D.
Yanzhuang Wang, Ph.D.

TABLE OF CONTENTS

ABSTRACT

Glyoxysomes are specialized peroxisomes found in plant seedlings that promote the mobilization of lipid stores for energy. Although several glyoxysomal proteins have been identified, many of their functions remain relatively unknown. Here I present a characterization of GX06 (At4g36880), a predicted glyoxysomal protease found in *Arabidopsis thaliana* which has a putative peroxisomal targeting signal type 1 (PTS1) of SSV>. A yellow fluorescent protein-GX06 fusion showed peroxisomal localization by fluorescence microscopy. This suggests the possible identification of a novel PTS1 (SSV>), which must be confirmed with further testing. GX06 exhibited both pH- and temperature-dependent processing of casein in radiolabeled casein substrate assays. An analysis of *GX06* mRNA expression using quantitative real-time PCR revealed high expression in flowers and seedlings, and in particular, dark-grown seedlings. SALK-line mutant plants with *GX06* T-DNA insertions grew larger and developed faster than wild-type *Arabidopsis thaliana* plants, based on measurement of leaf span and number of leaves, suggesting a possible inhibitory function for GX06 in the glyoxysome.

INTRODUCTION

Peroxisomes are single-membrane-bound organelles that perform a wide variety of metabolic functions in eukaryotic organisms. They process very-long-chain fatty acids, which have hydrocarbon tails longer than 22 carbons and cannot be metabolized in the mitochondria. For the detoxification of harmful substances, peroxisomes harbor catalase. Catalase is an enzyme responsible for peroxidation reactions, the reduction of hydrogen peroxide to water. Peroxisomes are also the production site of plasmalogens, ether phospholipids that can be found in a variety of human tissues, most notably in the myelin sheaths of neurons. In humans, various defects in

peroxisome biogenesis or mutations in specific peroxisomal proteins lead to disorders such as Zellweger syndrome, neonatal adrenoleukodystrophy, and rhizomelic chondrodysplasia punctata, with varying degrees of lethality (Braverman et al., 1998; Purdue et al., 1997). Peroxisome biogenesis disorders often involve mutations in the PEX genes, which play an important role in peroxisomal protein import. Consequences of these disorders include the accumulation of very-long-chain fatty acids and deficient plasmalogen synthesis, leading to both physical and mental developmental disorders (reviewed in Steinberg et al., 2006).

Because peroxisomes lack genomic DNA, all peroxisomal proteins, including membrane proteins, must be post-translationally transported across the membrane from the cytosol or endoplasmic reticulum (Lazarow and Fujiki, 1985; Fujiki et al., 2006). The two different pathways for protein import into the peroxisome matrix depend on the peroxisome targeting signal present in the cargo protein, and are mediated by different receptors. A peroxisome targeting signal type 1 (PTS1) is characterized by a carboxyl-terminal tripeptide that has the general consensus sequence [S/A/C/P]-[K/R/H]-[L/M/I] (Lametschwandtner et al., 1998; Hayashi et al., 1997; Reumann et al., 2007). However, it was demonstrated that peroxisomal proteins may have PTS1 sequences that differ greatly from this established scheme. For example, SLM> and SKV> have both been characterized as PTS1s, although the sequences only overlap with two of the three residues specified by the consensus sequence (Reumann et al. 2009). According to the online database AraPerox, peroxisomal proteins may contain ‘rare’ PTS1 tripeptides like SSI> which differ significantly from the general consensus sequence (Reumann et al., 2004). In fact, several confirmed PTS1 proteins listed in the database have residues not even included in the original consensus sequence, suggesting that perhaps the consensus sequence should be modified to [S/A/C/P]-[K/R/H/N/S]-[L/M/I/V]. Novel PTS1 sequences

continue to be identified. A peroxisome targeting signal type 2 (PTS2) is characterized by an amino-terminal nonapeptide sequence that is proteolytically removed in plants and animals, but not in yeast (Helm et al., 2007).

The proteins involved in the import of peroxisomal proteins across the membrane are called peroxins, abbreviated as PEX proteins. PTS1 proteins are recognized by the Pex5 receptor in the cytosol, which physically binds cargo proteins and carries them to the peroxisomal membrane (Urquhart et al., 2000). Upon docking to the membrane, Pex5 interacts with membrane proteins Pex13 and Pex14, which make up the docking complex (Agne et al., 2003). This docking complex also serves to bind Pex7, the cytosolic receptor for PTS2 proteins. There are two main models for the translocation of cargo proteins into the peroxisome lumen, and the mechanism is not well understood. In the simple shuttle model, only the cargo passes the membrane to the peroxisome matrix while the receptors (Pex5 or Pex7) remain on the membrane. The extended shuttle model postulates that the receptor passes through the membrane while still attached to the cargo, releasing it after arrival in the matrix (reviewed in Lanyon-Hogg et al., 2010). After the cargo is released into the peroxisome, the Pex5 receptor recycles back to the cytosol (Dammai and Subramani, 2001).

Plant peroxisomes are involved in the detoxification of photosynthetic byproducts, metabolism of fatty acids, and biosynthesis of hormones (Hayashi and Nishimura, 2003; Nyathi and Baker, 2006; reviewed in Lanyon-Hogg, et al., 2010). Characterization of the various proteins involved in plant peroxisomal functions is crucial to our understanding of this organelle. Different classes of peroxisomes have specialized functions dependent upon the developmental stage of the plant. Glyoxysomes are a type of specialized peroxisome. They are present in developing seedlings, play a vital role in mobilizing lipid stores, and house glyoxylate cycle

enzymes (Beevers, 1979; Olsen and Harada, 1995). Glyoxysomes, which are especially abundant in dark-grown seedlings, are responsible for providing nutrients to growing seedlings by converting fatty acids into useable forms of energy through β-oxidation. The glyoxylate cycle resembles the tricarboxylic acid cycle, but involves the conversion of acetyl-CoA to succinate rather than α-ketogluterate. Succinate is then used for the synthesis of carbohydrates through gluconeogenesis. The four main glyoxysomal enzymes that comprise the glyoxylate cycle are citrate synthase, isocitrate lyase, malate dehydrogenase, and malate synthase (Kato et al., 1995; Turley et al., 1990; Gietl, 1990; Graham et al., 1989). This cycle is especially important for seedlings, which need carbohydrates to germinate and grow but do not yet have the means to obtain and process carbon dioxide through photosynthesis. Once grown in the light, glyoxysomes can become leaf or unspecialized peroxisomes (Titus, 1985; Figure 1).

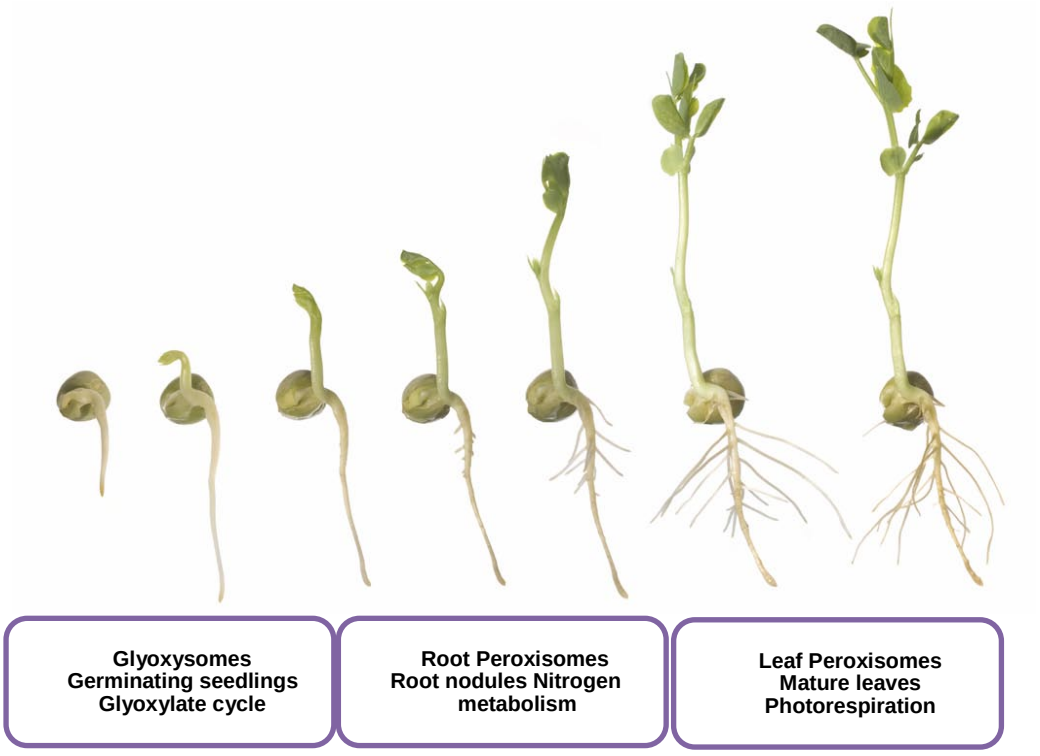

Figure 1. Model of Seedling Development. Glyoxysomes are present at the seedling stage and can become other types of peroxisomes once the plant is exposed to light. Each type of peroxisome has a different function in the plant's physiological processes.

A series of proteins predicted to be targeted to glyoxysomes, each containing a PTS1, was discovered in collaboration with Dr. Jianping Hu's laboratory at Michigan State University. Among these proteins was GX06, the sixth glyoxysomal protein found in the proteomics project (Hu lab, unpublished data). GX06 function has yet to be determined. This 41.6 kD protein has a putative PTS1 of SSV>, which has not been previously described, but resembles predicted rare PTS1 signals like SSL> and SSI> and fits the modified consensus sequence obtained from examination of the confirmed PTS1 sequences annotated on AraPerox (Reumann et al., 2004). If GX06 is indeed a peroxisomal protein and its peroxisomal import is mediated by the SSV> PTS1, this would represent the discovery of a novel PTS1.

GX06 has been declared a putative cysteine protease based on the presence of the D/E-R-F-N-I-N motif, a highly conserved sequence in almost all cysteine proteases (Beers et al., 2004; Figure 2). GX06 is considered a C1A family papain-like cysteine protease. Some cysteine endopeptidases of this family have been implicated in degrading seed storage proteins during the germination stage, a plausible role considering the predicted glyoxysomal localization of GX06 (Yamauchi et al., 1996). Based on homology, the predicted catalytic region of cysteine proteases consists of Cys, His, Asn/Asp, and GX06 has several candidate residues that fit this scheme (Beers et al., 2000; Figure 2).

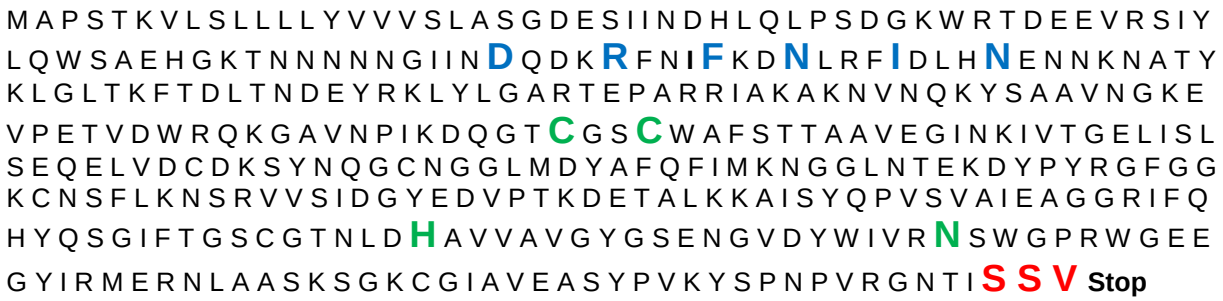

Figure 2. Protein sequence of GX06. GX06 has 376 residues and a predicted molecular weight of 41.6 kD. The putative PTS1 signal consists of the C-terminal tripeptide, SSV>, shown in red. The catalytic region prediction is based on homology to cysteine proteases; residues are shown in green. GX06 contains a D/E-R-I-F-N-I-N motif (shown in blue), a highly conserved sequence present in the propeptide of almost all cysteine proteases (Beers et al., 2004).

An analysis of the *GX06* temporal expression profile on Genevestigator, a gene expression database, indicates that *GX06* mRNA is highly expressed at the earliest stage of development – in seedlings undergoing germination (Figure 3). It has been suggested that once proteins are identified as peroxisomal, they can be categorized by their expression profiles. RNA expressed at high levels in the seedling stage is consistent with glyoxysomal localization (Kamada et al., 2003). Using this reasoning, the expression profile of GX06 suggests that its SSV> C-terminal tripeptide is in fact a true PTS1.

My CMB Honors thesis project was the characterization of GX06, which has putative protease activity in the glyoxysome and may be essential for normal function. Characterizing the function of GX06 in *Arabidopsis thaliana* will assist in the development of a global picture of how glyoxysomal proteins interact to guide the plant through its early developmental stages.

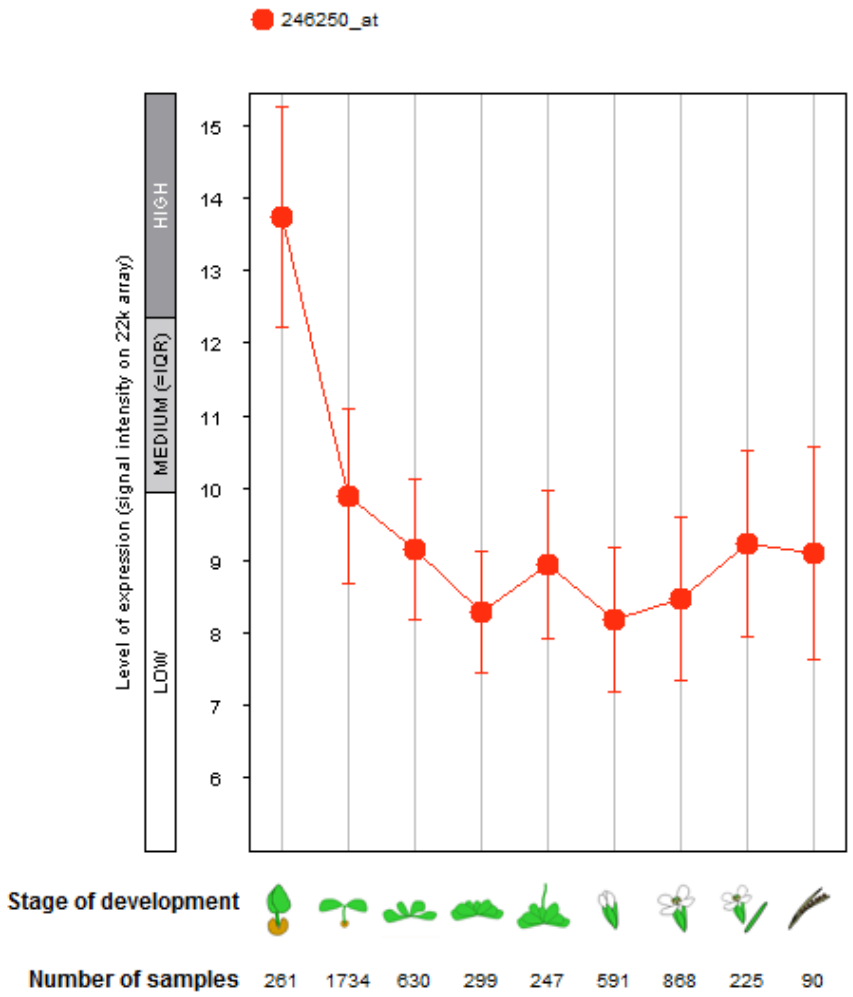

Figure 3. Genevestigator temporal expression profile for *GX06* (At4g36880, https://www.genevestigator.com/gv/plant.jsp). *GX06* mRNA is expressed at relatively high levels at the seedling stage and decreases to a baseline level of expression as the plant matures.

RESULTS

GX06 showed glyoxysomal localization

To test GX06 subcellular localization, I first used an *in vitro* glyoxysomal protein import assay (Behari and Baker 1993, Brickner et al., 1997). The glyoxysome import assay involves the incubation of radiolabeled GX06 with glyoxysomes, Mg-ATP, and buffer at 25ºC, the temperature ideal for import (Figure 4). As a negative control, one sample is incubated at -20°C,

a temperature at which the glyoxysomes will freeze and not allow import. GX06 that has transported across the glyoxysomal membrane would be protected from degradation after the addition of Proteinase K, while any protein that failed to import would be digested. Because the import assay requires the use of the Proteinase K protease, I first generated a protease concentration curve by incubating radiolabeled GX06 with various concentrations of Proteinase K at 37°C to detect GX06 degradation. GX06 appeared to be optimally degraded by Proteinase K at 5.0 µg/mL, a concentration which was then used in the import assay (Figure 5). Glycolate oxidase (GLO), with a major PTS1 of ARL>, was used as the positive control for import. GLO was imported as expected (Figure 6, lane 7). The only GLO protein appearing on the SDS-PAGE gel was that which had been imported into the glyoxysome. GX06 was not imported in this assay; no protein was detected (Figure 6, lane 3). However, even the import of GLO, the positive control, appeared to be minimal, suggesting that GX06 may have been imported, but at levels below detection.

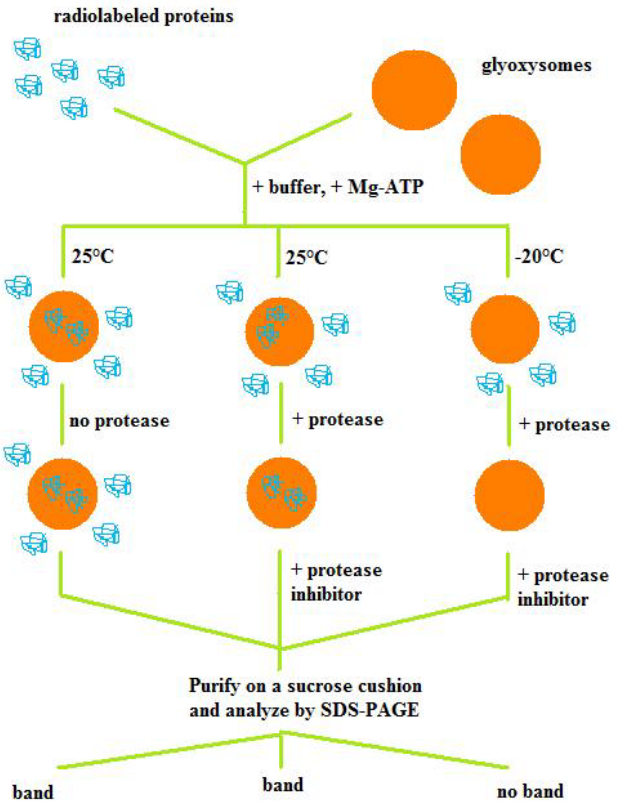

Figure 4. Diagram of glyoxysomal protein import assay. Glyoxysomes were incubated in the presence of radiolabeled GX06 protein at the temperatures specified. Proteinase K and protease inhibitor were added to two of the three samples. Samples were analyzed by SDS-PAGE.

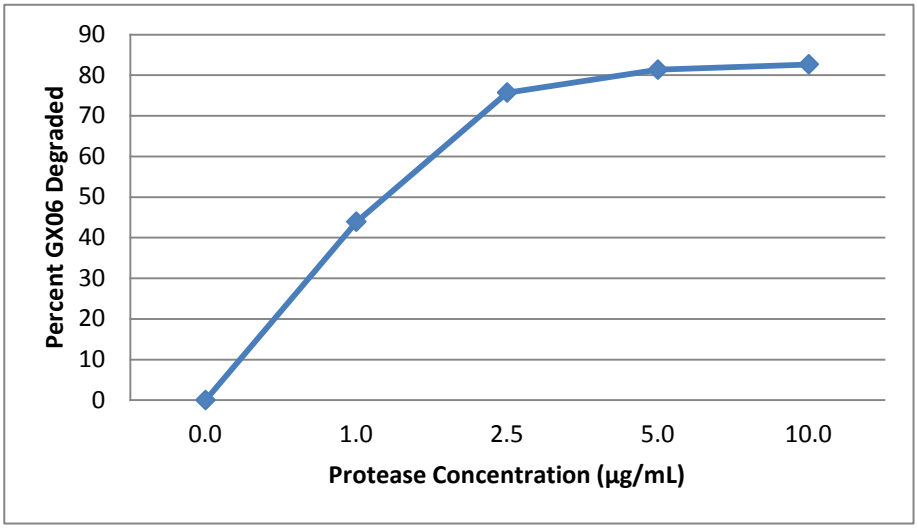

Figure 5. GX06 Protease Concentration Curve. Digestion of GX06 with increasing concentrations of Proteinase K. GX06 degradation increased with protease concentration to a maximum of approximately 80%. For subsequent protein import assays, 5 µg/mL Proteinase K was used to digest unimported GX06 protein.

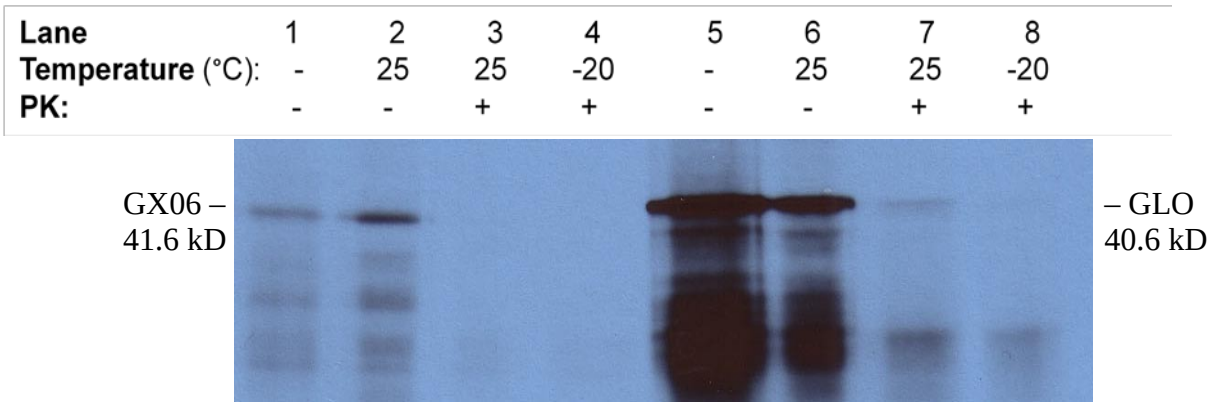

Figure 6. Glyoxysome protein import assay with GX06 and GLO. Lanes 1-4 show GX06, lanes 5-8 show GLO. Lanes 1 and 4 are translated protein only. PK, Proteinase K.

Because the *in vitro* import assay may not be ideal for assessing GX06 localization, an *in vivo* system was also employed. A yellow fluorescent protein-GX06 fusion construct (YFP-GX06) was created for transformation into *Agrobacterium tumefaciens* and subsequent infiltration into *Nicotiana tabacum* leaves. Leaves were given two days post-infiltration to grow and recover, and were then imaged by fluorescence microscopy. The positive control used was CFP-SKL>, a fusion construct between cyan fluorescent protein and three additional amino acids, where SKL> is a major PTS1 known to target protein to all types of peroxisomes (Hayashi et al., 1997). CFP-SKL> and YFP-GX06 appeared to co-localize in the tobacco leaves, suggesting that YFP-GX06 was peroxisomal and its PTS1 was sufficient to mediate transport into the peroxisome (Figure 7). Thus, although the *in vitro* glyoxysome import assay did not detect GX06 import, the data from tobacco infiltration and fluorescence microscopy suggested that GX06 was indeed peroxisomal in localization.

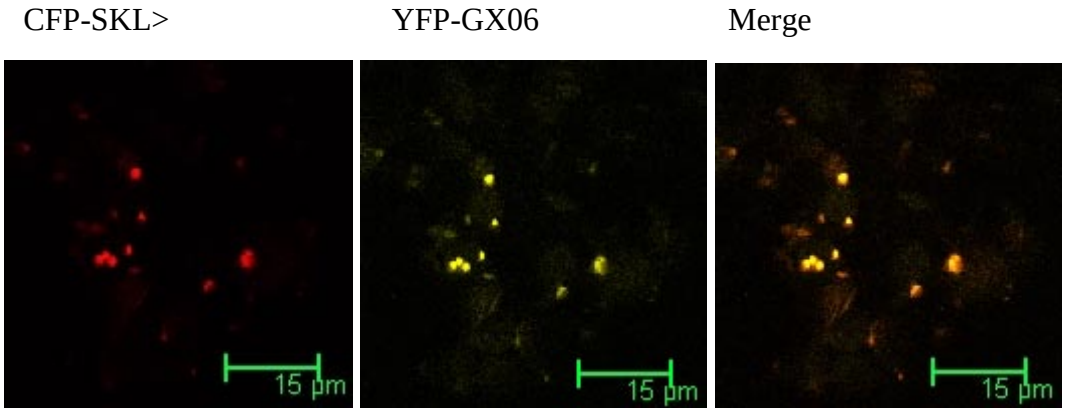

Figure 7. *In vivo* fluorescence microscopy of the YFP-GX06 fusion protein. *A. tumefaciens* strains containing either pCHF3-CFP-SKL or pCAM-eYFP-GX06 were mixed together and infiltrated into tobacco leaves. CFP-SKL> and YFP-GX06 appeared to co-localize, indicating the peroxisomal localization of the YFP-GX06 fusion protein. This experiment was repeated three times; representative results are shown.

GX06 had pH- and temperature-dependent protease activity

To test the optimum conditions for GX06 protease activity, I performed several protease assays using radiolabeled casein as a substrate. Casein, a protein found in mammalian milk, has very unstable tertiary structure due to its lack of disulfide bonds (Mercier et al., 1971; Bhattacharyya and Das, 1999). As a result, it functions as an ideal substrate because it can be cleaved by a variety of proteases. The protease assay consisted of incubating GX06, radiolabeled casein, and buffer, with a (-) protease negative control for 4 hours at 37°C. To determine the optimum pH for GX06 activity, buffers of various pHs were used. Percent casein remaining was determined by quantifying the SDS-PAGE gel with phosphor imaging, and values were adjusted by setting the (-) protease control to 100%. GX06 appeared to be most proteolytically active at acidic pHs, specifically pH 5 (Figure 8). GX06 activity was tested at pH 7-9 as well, but processed casein minimally at basic pHs (data not shown).

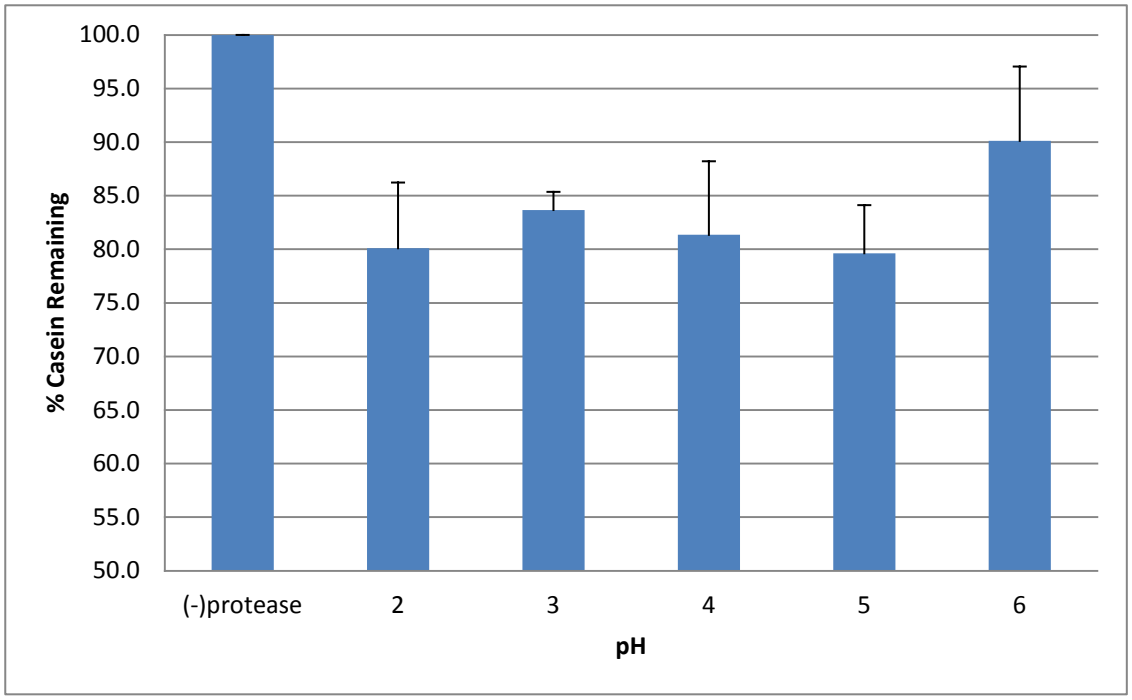

Figure 8. pH-dependent degradation of radiolabeled casein. Data were compiled from 3 trials; error bars indicate standard deviation. GX06 appeared to be more proteolytically active at pH <6, and had maximum activity at pH 5.

To determine the optimum temperature for GX06 activity, the protease assay mixtures were incubated in buffer at pH 5, but at different temperatures. Samples were then run on an SDS-PAGE gel and quantified by phosphor imaging. Again, values for percent casein remaining were adjusted to the (-) protease sample as 100%. GX06 protease activity appeared to increase with temperature, suggested by the decrease in percent casein remaining (Figure 9). Thus, GX06 had pH- and temperature-dependent protease activity.

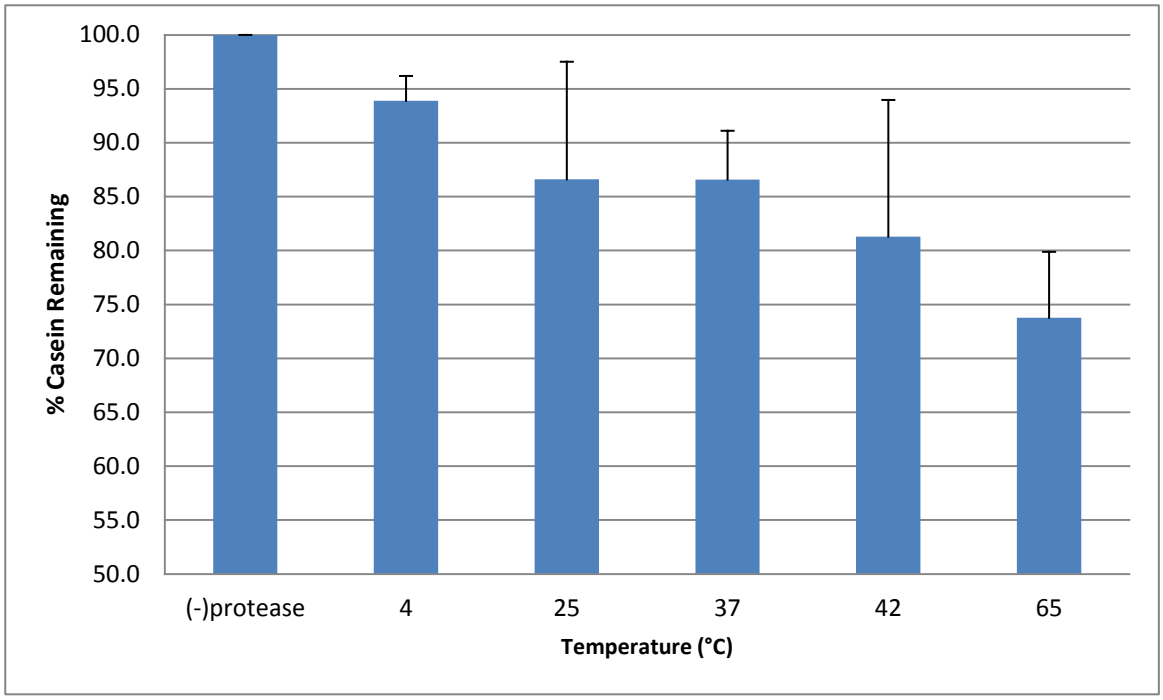

Figure 9. Temperature-dependent degradation of radiolabeled casein. Data were compiled from three trials; error bars indicate standard deviation. GX06 protease activity appeared to increase with temperature.

GX06 mRNA was highly expressed in flowers and dark-grown seedlings

A search for the *GX06* gene (At4g36880) in the gene expression search engine Genevestigator gave a temporal map of *GX06* expression across the developmental stages of *A. thaliana* (Figure 3). To construct a spatial map of *GX06* mRNA expression within different types of plant tissues, RNA was isolated from *Arabidopsis thaliana* plants, converted to cDNA through reverse transcription, and analyzed with quantitative real-time PCR for *GX06*. *GX06* was highly expressed in flowers and seedlings compared to other tissues, as determined by the smaller C_T value (Figure 10). *UB10*, ubiquitin, served as the positive control because it is expressed at a high and steady level across all plant tissues (Callis and Vierstra, 1989).

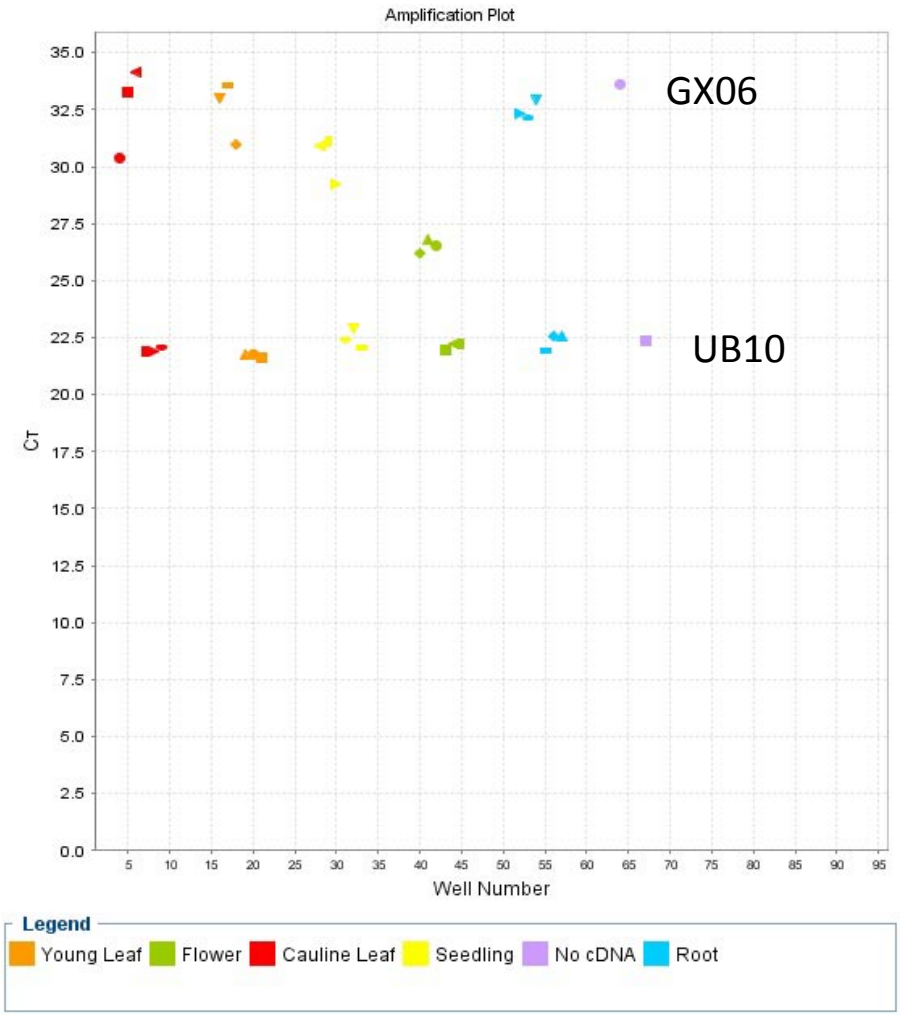

Figure 10. Quantitative real-time PCR of *GX06* mRNA levels in various tissues. Total RNA was isolated from various tissues as indicated and converted to cDNA for qPCR. High expression is indicated by a low C_T, the cycle at which fluorescence is detected above a threshold value. *GX06* was more highly expressed in seedlings and flowers, as shown. *UB10*, ubiquitin, was equally expressed across all tissues.

To further characterize the high expression seen in seedlings, I tested whether *GX06* would be more highly expressed in dark-grown or light-grown seedlings. At the seedling stage, glyoxysomes are important in mobilizing lipid stores for energy and growth. Once seedlings are grown in light, they begin to photosynthesize and no longer have great need for glyoxysomes. RNA was isolated from dark-grown and light-grown seedlings, converted to cDNA, and again analyzed by quantitative real-time PCR. Dark-grown seedlings had higher *GX06* mRNA

expression levels than light-grown seedlings (Figure 11). This provided additional confirmation that GX06 is a glyoxysomal protein by correlating *GX06* expression with its predicted role in the plant.

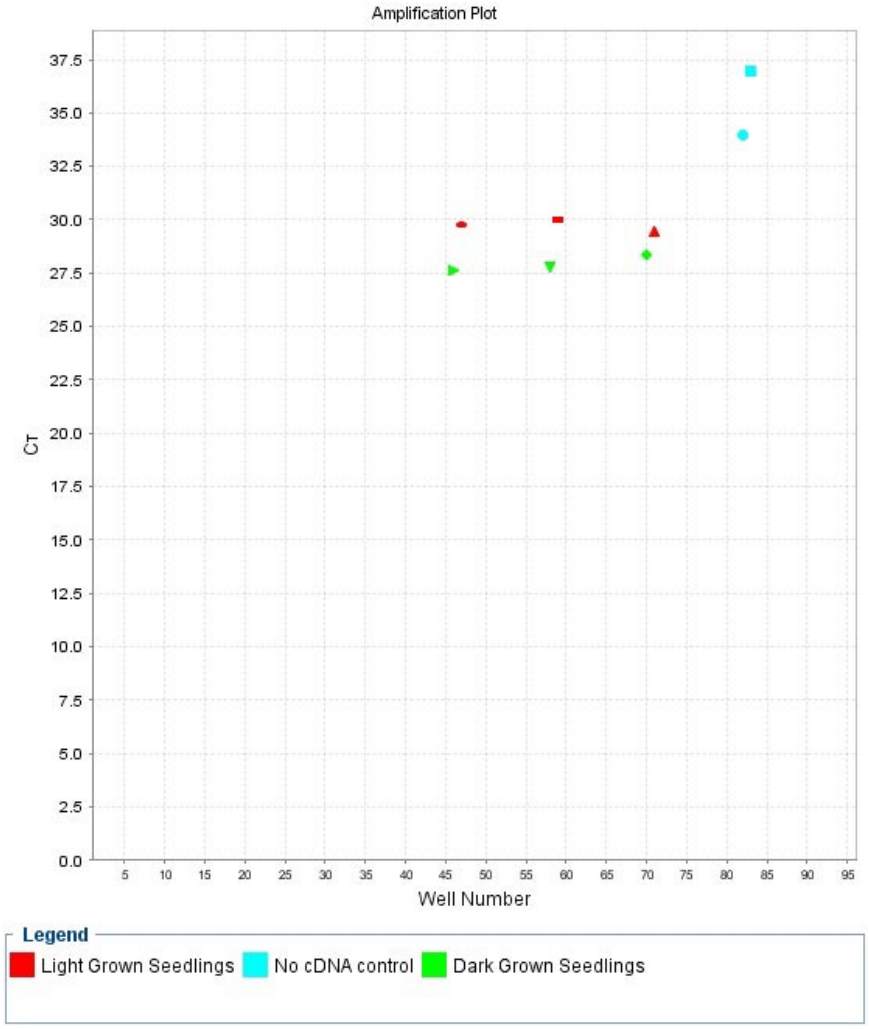

Figure 11. Quantitative real-time PCR of *GX06* mRNA levels in light-grown and dark-grown seedlings. *GX06* was expressed at higher levels in dark-grown seedlings. This corresponds with the higher abundance of glyoxysomes, and therefore glyoxysomal proteins, at early stages in plant development before seedling germination.

gx06 mutant plants developed faster and grew larger than wild-type *Arabidopsis thaliana*

To assess the phenotype of *gx06* mutant plants, seeds from two lines containing T-DNA insertions in the *GX06* gene were obtained from the Arabidopsis Biological Resource Center. SALK 051510C has a T-DNA insertion in intron 1 of *GX06* which runs in the reverse direction

to the gene, while SALK 085378C has a T-DNA insertion in exon 3 of *GX06* which runs in the same direction as the gene (Figure 12A, B). The T-DNA insertions are greater than 4 kb in length, effectively knocking out any potential GX06 function. Both mutations are in Col-0 *Arabidopsis thaliana* background lines. 48 seeds from each 051510C, 085378C, and Col-0 wild-type lines were planted. Throughout the course of four weeks, they were assessed for percent survival, leaf span, and number of leaves. By the first week, the fewest wild-type Col-0 plants had germinated, and the survival rate of these plants continued to decline at a greater rate than both of the SALK mutant lines (Figure 13). Plants that germinated but died afterward were usually much smaller than normal and grew very poorly until death. SALK 085378C plants had the greatest survival rates, with 88% survival after four weeks, compared to wild-type plants whose survival rate was 71% at the end of four weeks.

A

B

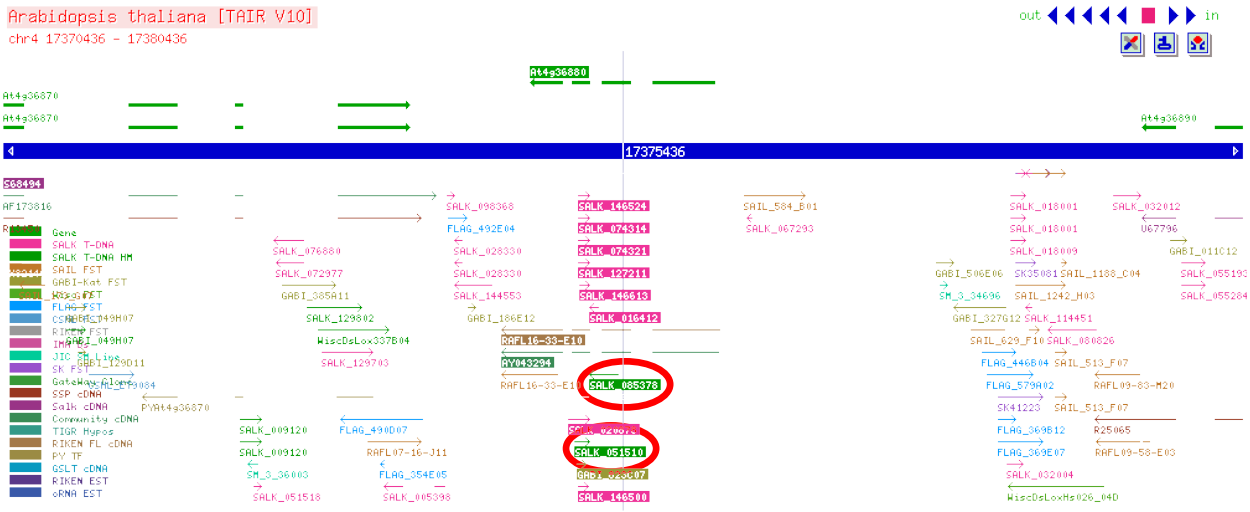

Figure 12. Gene maps of *GX06* with the insertion sites of SALK 051510C and 085378C lines. A, SALK 051510C insertion in intron 1 and SALK 085378C insertion in exon 3 are shown. B, the two SALK lines of interest are circled in red; their relative directionalities are indicated.

Leaf span and number of leaves were the two criteria used to assess growth of the plants. As *Arabidopsis thaliana* plants grow, the leaf spans of the rosette leaves increase; larger leaf spans indicate more robust growth. The number of leaves per plant increases quickly over time, indicating developmental progression. On average, plants of both SALK lines 051510C and 085378C had greater leaf spans than wild-type plants over all four weeks (Figure 14). Plants of SALK line 051510C appeared to have the greatest leaf spans of all three lines.

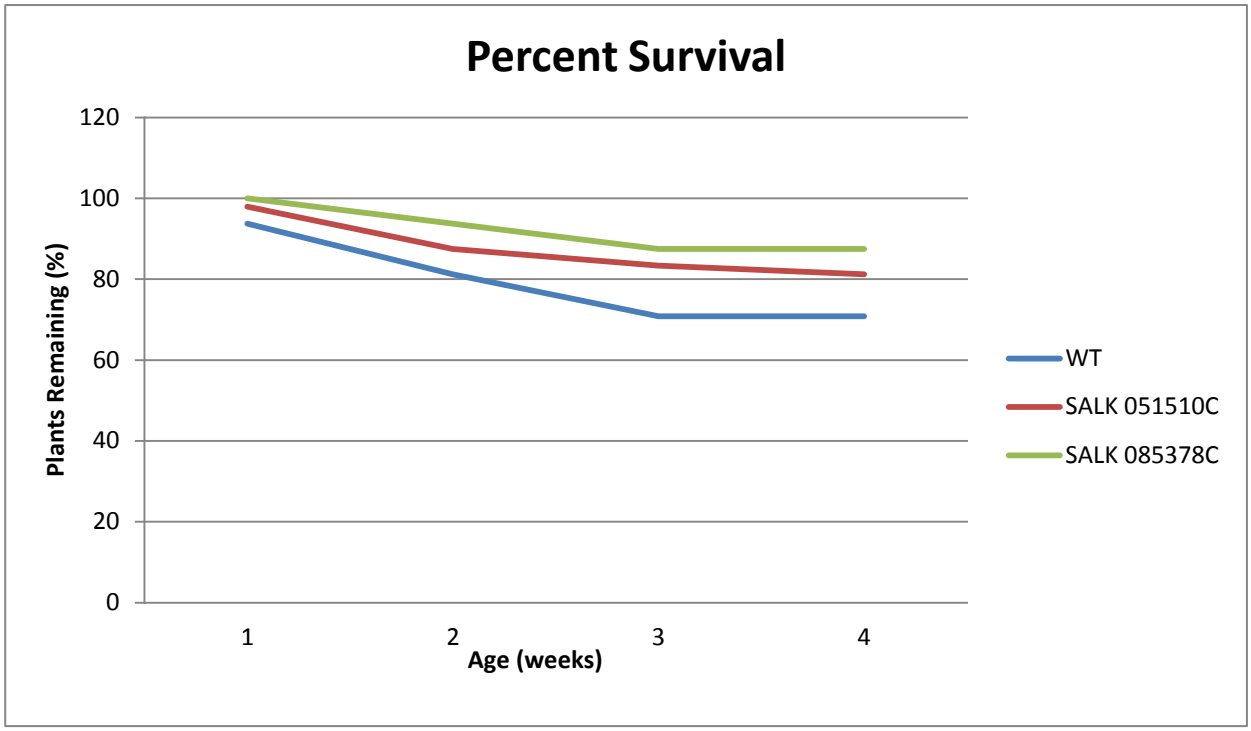

Figure 13. Percent survival of wild-type and mutant plants over four weeks. 48 seeds per line were planted and scored at each week for survival.

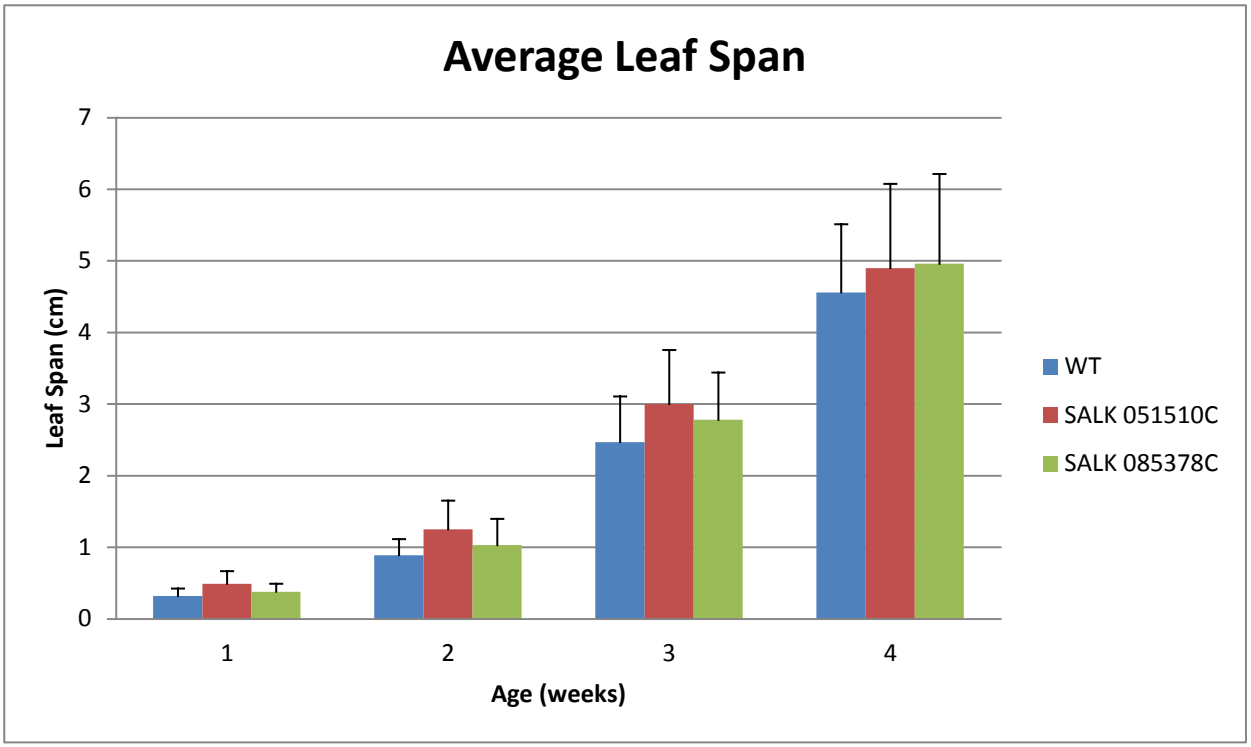

Figure 14. Average leaf spans of wild-type and mutant plants over four weeks. 48 seeds per line were planted. Leaf span was measured from the tips of the longest rosette leaves. Error bars indicate standard deviation.

The number of leaves per plant was also scored in order to measure growth and development in the three plant lines. After the plants grew inflorescences at three weeks, the cauline leaves that grew on the inflorescences were added to the number of rosette leaves for the total number of leaves per plant. Again, the SALK line mutant plants generally grew more leaves than the wild-type plants (Figure 15). The most obvious difference was seen between SALK 051510C and wild-type plants, while the SALK 085378C had very similar numbers of leaves as wild-type plants.

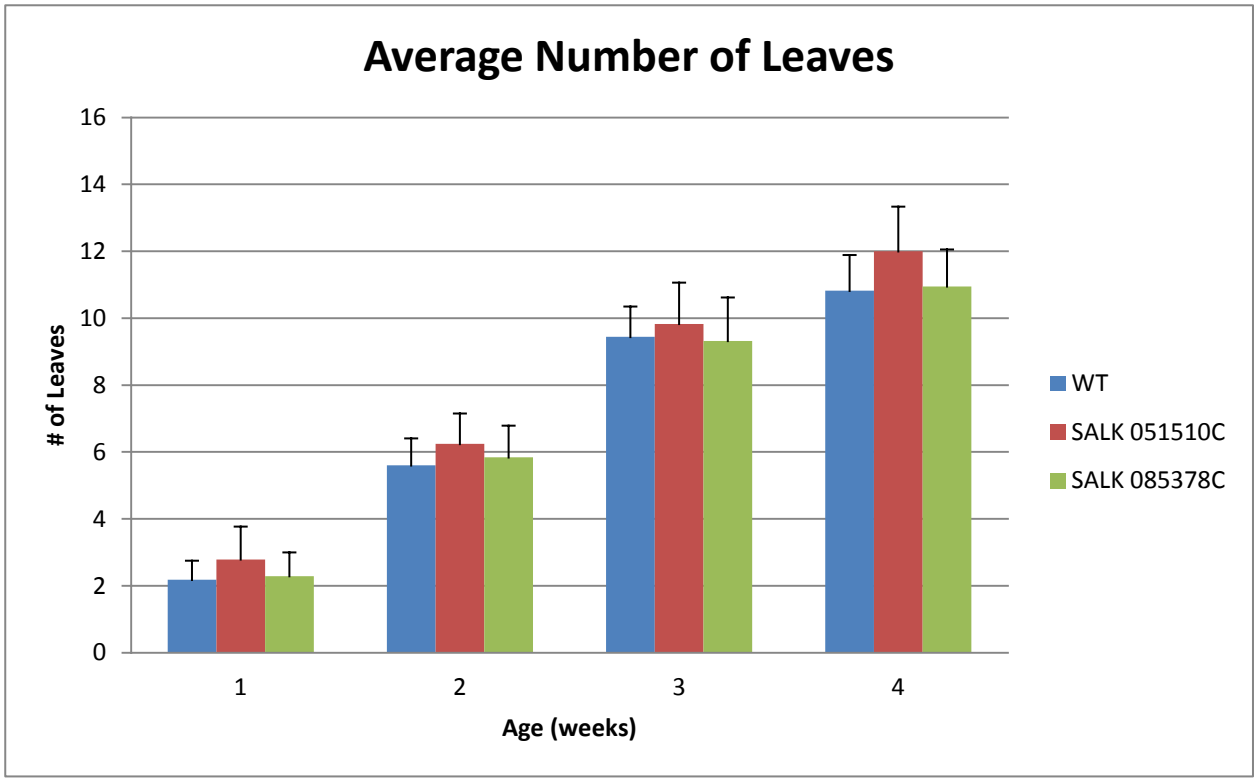

Figure 15. Average number of leaves of wild-type and mutant line plants over four weeks. 48 seeds per line were planted and scored for number of leaves per plant. Total number of leaves was calculated as number of rosette leaves plus number of cauline leaves after inflorescences grew. Error bars indicate standard deviation.

Although the differences in measurements of average leaf span and average number of leaves were not statistically significant, the phenotypic differences were clearly seen. SALK line mutant plants appeared larger in leaf span and had greater numbers of leaves. This was most notable when comparing SALK 051510C plants to wild-type plants (Figure 16). Based on all the phenotypic data describing the differences between wild-type lines and SALK mutants, it was inferred that GX06 is not vital to *Arabidopsis thaliana* because *gx06* plants still survived and grew. However, GX06 may have an inhibitory function in the plant because the absence of normal protein in mutant lines correlates to larger and faster plant growth. Further evaluation is needed to elucidate the role of GX06 in the plant.

Figure 16. Phenotypic comparison of wild-type and SALK 051510C plants. Plants were two weeks old in these photographs.

DISCUSSION AND FUTURE DIRECTIONS

The results presented above show that GX06 is peroxisomal, has pH- and temperature-dependent protease activity, is highly expressed in flowers and dark-grown seedlings, and that plants with *gx06* mutations grow larger and faster than wild-type plants. GX06 peroxisomal localization was first examined through glyoxysome protein import assays, which tested whether GX06 imported into glyoxysomes and was protected from protease digestion by the glyoxysomal membrane. The results obtained from the import assays were not conclusive; GX06 did not appear to import successfully into the glyoxysome (Figure 6). The GLO positive control showed successful import, but the band in the SDS-PAGE gel was very faint, indicating that only a small

fraction of the GLO protein added to the assay mixture actually imported. Since GLO has a major PTS1 and is a known peroxisomal protein, yet it showed weak import, it could be that too few GX06 proteins imported across the glyoxysome membrane to be detectable on the gel. Because this was an *in vitro* approach to assessing localization, the conditions used in the assay may not have been ideal for GLO or GX06 import. Further optimization of this assay may provide more information about GX06 peroxisomal localization.

A yellow fluorescent protein-GX06 fusion protein construct (YFP-GX06) was created to test *in vivo* localization. Because proteins are constantly being imported into leaf peroxisomes naturally, this system is more sensitive to GX06 import than the *in vitro* glyoxysome protein import assays. YFP-GX06 appeared to colocalize with CFP-SKL>, the positive control, suggesting that GX06 is peroxisomal (Figure 7). This represents the discovery of a novel PTS1. Interestingly, the SSV> sequence fits only one of the three residues in the general consensus scheme [S/A/C/P]-[K/R/H]-[L/M/I] for PTS1 proteins (Lametschwandtner et al., 1998 and Hayashi et al., 1997). However, as more novel PTS1 proteins have recently been discovered, it is evident that sequences may deviate significantly from the scheme and yet still import successfully into the peroxisome (Reumann et al., 2009). If the SSV> PTS1 is corroborated in further testing, a modification of the known consensus sequence may be necessary. Although these results are promising, more work needs to be done to understand GX06 targeting to the peroxisome. For example, the SSV> PTS1 signal in GX06 could be mutated through site-directed mutagenesis. The mutated protein should not import into peroxisomes and would therefore not co-localize with CFP-SKL>.

GX06 processed radiolabeled casein in a pH- and temperature-dependent manner. GX06 was most active at acidic pH levels of 2-5 (Figure 8). The acidic optimum pH of GX06 is

interesting in comparison with the different optimum pH requirements of other known glyoxysomal proteins, such as isocitrate lyase at pH 6.75, and malate synthase at pH 8 (Lamb et al, 1978, Bowden and Lord 1978). It has been reported that the matrix of peroxisomes is an acidic environment with pH of approximately 5.8-6.0. This environment is the result of the presence of ATPases on the peroxisomal membrane which establish a proton gradient in yeast (Nicolay et al., 1987; Waterham et al., 1990). ATPase activity has also been detected in rat liver peroxisomes, suggesting that matrix acidity is a general property of peroxisomes in many different species (Del Valle et al., 1988). GX06 optimal activity at acidic pH levels is consistent with acidic peroxisome matrix pH, providing further evidence that GX06 is a glyoxysomal protease.

GX06 protease activity increased with temperature and degraded casein optimally at 65°C (Figure 9). The temperature curve was performed to be sure that 37°C, the temperature of incubation in the pH assay, was suitable for GX06 activity. GX06 processed similar amounts of radiolabeled casein at 25°C and 37°C, suggesting that the activity seen in the pH assay (37°C) could be comparable to physiological conditions in the plant (25°C). Interestingly, GX06 had greatest protease activity at 65°C, a temperature at which many enzymes are degraded. This suggests that GX06 may be a thermostable protease.

These results are preliminary, and the optimal conditions under which GX06 would function may be different *in vivo* because these experiments were performed with a casein substrate. In the future, it would be interesting to determine the true substrate of GX06, possibly by using an α-GX06 antibody for immunoprecipitation and examining the proteins to which GX06 is bound *in vivo*. Once a substrate for GX06 is identified, the catalytic region of GX06 could be determined through protease assays. Based on homology to other cysteine proteases, the

putative catalytic triad of GX06 consists of a Cys, His, and Asn. The DNA sequence coding for these residues could be mutated through site-directed mutagenesis and the effect on protease activity of GX06 could be assessed through degradation of the substrate, similar to the casein-substrate assays used above. Identifying the true substrate of GX06 and testing its catalytic region would further clarify the role of this protease in the glyoxysome.

GX06 mRNA was highly expressed in flowers and seedlings, particularly dark-grown seedlings (Figures 10 and 11). Dark-grown seedlings have never been exposed to light, and would therefore have an abundance of glyoxysomes and glyoxysomal enzymes. *GX06* is still expected to be expressed in light-grown seedlings at higher than basal levels because there may be glyoxysomes remaining in the seedling that have not yet been converted to other types of peroxisomes. Although flowers, the sites of pollen production, do not normally have glyoxysomes, it has been shown that that glyoxysomal function is induced during pollen development (Zhang et al., 1994). The higher *GX06* expression in dark-grown seedlings and in flowers further supports the conclusion that GX06 is a glyoxysomal protease. In the future, *GX06* expression should be examined in the individual flower tissues such as the pistil, sepal, and anthers.

Plants with T-DNA insertions in the *GX06* gene grew larger and developed faster than wild-type *Arabidopsis thaliana* plants, as determined by measurement of leaf span, and number of leaves over four weeks (Figures 13-16). Based on the abundance of *GX06* mRNA in seedlings compared with other tissues, it is likely that the *gx06* mutation exerted an early growth effect on the *A. thaliana* plants (Figure 10). *gx06* mutant plants may have developed normally after the seedling stage, but the differences in phenotypes between *gx06* and wild-type plants that arose at early developmental stages persisted throughout the rest of the plants' life cycle. Since *gx06*

plants were larger, had greater average leaf spans, and a greater number of leaves than wild-type plants, it is possible that GX06 has an inhibitory function in the plant. *gx06* mutant plants still show relatively normal phenotypes; thus, it is clear that GX06 is not critical for plant survival. This suggests that the protease may have a generalized or redundant function.

The work presented in this thesis serves as a preliminary investigation into the role of GX06 in *Arabidopsis thaliana*. It has been shown that GX06 was likely glyoxysomal based on fluorescence microscopy, had pH- and temperature-dependent protease activity, was highly expressed in dark-grown seedlings compared to other plant tissues, and may have inhibited growth in the plant. Future work on defining the PTS1 of GX06, identifying a true substrate, and determining sites of high expression will all serve to further elucidate the exact function of GX06 in *Arabidopsis thaliana.* If the carboxyl-terminal sequence SSV> of GX06 is confirmed through further testing as a true PTS1, it will represent the discovery of a novel signal and may require an extension of the general consensus sequence and a change in the way PTS1 sequences are defined. Identifying a substrate and further refining the expression profile for GX06 will aid in understanding how the role of this protease relates to other proteases in the overall function of glyoxysomes in the developmental progression from seedling to mature plant.

MATERIALS & METHODS

Glyoxysome Isolations and Glyoxysome Protein Import Assays

Glyoxysomes were purified as described previously (Harrison-Lowe and Olsen 2006). Pumpkin seeds (*Cucurbita pepo* var. Connecticut Fields, supplied by Petoseed Co., Inc., Saticoy, CA) were grown in damp vermiculite in complete darkness for 7 d at 25°C. Approximately 50 g cotyledons were harvested and homogenized in grinding buffer (40 mM tetrasodium

pyrophosphate, 2 mM EDTA, 0.6 M D-mannitol) in a blender with five 3 s bursts. The mixture was filtered through Miracloth and centrifuged at 3000 × *g* at 4°C for 10 minutes in a swinging bucket Sorvall HB-6 rotor. The lipid layer was removed and the sample was poured into fresh tubes and spun at 10,000 × *g* at 4°C for 20 min. Pellets were gently resuspended in 1 mL resuspension buffer (20 mM Hepes, 0.6 M D-mannitol). The resulting sample was loaded onto 28% Percoll in resuspension buffer, overlaid on a 2M sucrose cushion, and centrifuged at 18,000 × *g* at 4°C for 30 min in the swinging bucket rotor without braking. Glyoxysomes forming a visible band at the 28% Percoll/Suc interface were collected and diluted 3-fold with resuspension buffer. The sample was then centrifuged at 6800 × *g* at 4°C for 14 min with braking. The pellet was resuspended in resuspension buffer to 25 μg/μL.

The glyoxysome protein import assay was conducted as previously described (Brickner et al., 1997). 10 μL fresh glyoxysomes (250 μg) were incubated with 360,000 cpm GX06 translation product, 100 mM Mg-ATP, and import buffer (50 mM Mes-KOH, pH 6.0, 1M Suc, 20 mM KCl, 2 mM $MgCl_2$, 10 mM NaN_3) in 200 μL final volume. Import reactions were incubated at 25°C for 1 h, except for the negative control, which was incubated at -20°C for 1 h. Samples were then treated with 5 μg/mL Proteinase K for 30 min on ice to digest any protein that had not imported into the glyoxysomes. Reactions were stopped with 1 mM phenyl methane sulphonyl fluoride (PMSF) final concentration for inhibition of Proteinase K. The import samples were loaded onto 0.7 M Suc cushion (in import buffer, 500 μL), and centrifuged at 8500 × *g* at 4°C for 15 min. Pellets were resuspended in 70 μL 1X sample buffer, boiled for 5 min, and stored at -20°C until further analysis. Samples were visualized by SDS-PAGE and autoradiography.

Generation of Constructs

GX06 was amplified from a λ-YES cDNA library with Pfu Turbo using the primers SalI-*GX06*-forward 5'-GTC GAC ATG GCT CCT TCA ACA AAA G-3' and SalI-*GX06*-reverse 5'-GTC GAC TCA AAC ACT GCT GAT AGT AT-3' for the insertion of flanking SalI restriction sites. *GX06* was cloned into the pCRII-TOPO vector (Invitrogen, Grand Island, NY).

Radiolabeled GX06 protein used in glyoxysomes protein import assays was created by linearizing *GX06* in pCRII-TOPO vector with KpnI and *in vitro* transcription with T7 RNA polymerase (Promega, Madison, WI). *In vitro* translation was performed by adding *GX06* mRNA to a cell-free wheat germ extract translation system containing radiolabeled ^{35}S-methionine.

Transient cDNA Expression and Fluorescence Microscopy

The pCAM-35S-eYFP-c1 vector was designed for gene fusion to YFP for fluorescence microscopy (a gift from Nielsen lab, University of Michigan, Ann Arbor, MI). *GX06* was subcloned from pCRII-TOPO into the vector at the 3' end of the e*YFP* gene by digestion with SalI, gel purification with a QIAEXII kit (Qiagen, Valencia, CA), and ligation with T4 DNA ligase (Promega, Madison, WI). *Agrobacterium tumefaciens* GV3101 strains were transformed with the pCAM-eYFP-GX06, pCHF3-CFP-SKL, or p19 helper plasmids. *A. tumefaciens* cells were incubated with 10 μg corresponding plasmid for 5 minutes on ice, 5 min in liquid nitrogen, and 5 min at 37°C. 1 mL LB was added and cells recovered with rotation at room temperature for 2 h. Cells were centrifuged at 2000 × *g* for 2 min and all but 200 μL were removed. The pellet was resuspended and cells were plated onto LB + gent^{20}, riff^{25}, and kan^{100} (pCAM-eYFP-GX06) or amp^{100} (pCHF3-CFP-SKL, p19) media and grown for 2 days at 30°C. For tobacco infiltrations, cells picked from media plates were grown in 5 mL liquid cultures at 30°C overnight and resuspended in infiltration buffer (0.01 M $MgCl_2$, 0.1 μM Acetosyringone) to an

optical density of 0.8 at 600nm. Equal volumes of the strains containing the three plasmids were mixed and infiltrated into 5-week old *N. tabacum* leaves with 1-mL needleless syringes. Leaf tissues were prepared for imaging after 2 days. Confocal images of leaf tissues were acquired using a Leica SP5 laser confocal microscope (excitation 514 nm, emission 525-566 nm YFP, excitation 458 nm, emission 470–498 nm, GFP) with a 40X lens. Images were cropped using Leica Application Suite Advanced Fluorescence software.

Radioactive Casein Substrate Assays

The *CSN2* gene encoding casein in pCMV-SPORT6 vector was linearized using NotI and transcribed *in vitro* with SP6 RNA polymerase (Promega, Madison, WI). *In vitro* translation was performed by adding *CSN2* mRNA to a cell-free wheat germ extract translation system containing radiolabeled ^{35}S-methionine.

The pET-28a(+) vector was designed for the addition of an N-terminal His-Tag for the overexpression and purification of a protein of interest. *GX06* was cloned into the pET-28a(+) vector with the SalI restriction enzyme and transformed into BL21 *Escherichia coli* overexpression cells. 1 L BL21 *E. coli* containing pET-28a(+)-GX06 was grown with 0.1 mg/mL final kanamycin to OD_{600} 0.45, induced with 1 mL of 100 mM IPTG, and grown for 1 h. Cells were harvested at 4000 × *g* for 15 min in a Sorvall GSA rotor. Pellets were resuspended in 50 mL ice-cold binding buffer (50 mM Tris-HCl, pH 7.8, 150 mM NaCl, 20 mM imidazole). The sample was transferred to new tubes for storage at -20°C until purification. For purification, each sample of cells was sonicated for 2 minutes, pooled and sonicated again for 30 s, and centrifuged at 13,000 × *g* for 15 min at 4°C. The supernatant was combined with 2 mL equilibrated nickel bead slurry and rotated at 4°C for 1 h. The solution was centrifuged for 3 min at 1500 × *g* and the supernatant was transferred to a Ni Sepharose column. 15 mL elution buffer was added (10 mM

imidazole in binding buffer). Fractions were collected; the four fractions with highest OD_{280} were pooled and filtered at 4100 × *g* for 5 min and subjected to overnight dialysis at 4°C in dialysis buffer (50 mM Tris-HCl, 150 mM NaCl, 100 mM Suc). Dialyzed GX06 was aliquoted for storage at -70°C. Samples were run on SDS-PAGE gel and Coomassie stained to check for presence of GX06 protein.

In the radiolabeled caseine-substrate protease assays, GX06 was incubated with ^{35}S-methionine-labeled casein, and 5X degradation buffer (125 mM Tris-HCl, 500 mM KCl, 50 mM $MgCl_2$) for 4 h at 37°C. Buffer pH was adjusted to pH 2-6 with HCl or NaOH. In the temperature assays, all samples were incubated in 5X degradation buffer at pH 5 at the specified temperatures for 4 h. Samples were analyzed by SDS-PAGE and casein degradation was quantified by phosphor imaging.

Isolation of RNA and Quantitative Real-Time PCR

For spatial expression analysis of *GX06* mRNA, young leaf, flower, cauline leaf, and root tissues were isolated from 4-week old Col-0 ecotype *Arabidopsis thaliana* plants. Seedlings were grown on MS medium with 2% Suc in 12 h light/12 h dark cycles and harvested after 7 days. For comparison of expression in light- and dark-grown seedlings, *A. thaliana* seeds were grown on MS medium supplemented with 2% Suc either in 12 h light/12 h dark cycles (light-grown seedlings) or in complete darkness (dark-grown seedlings) for 7 days.

RNA was isolated from various *A. thaliana* tissue types with a TRI reagent (Chomczynski and Sacchi 1987). RNA was then converted to cDNA following Promega protocol #TM337 for first-strand cDNA synthesis (Promega, Madison, WI). The cDNA, primers, and SYBR Green mix were combined and analyzed with qPCR. Primers used were *GX06* forward 5'-CGC TAA GGC CAA GAA TGT CAA CCA-3' and *GX06* reverse 5'-TCT GGA ACC TCC

TTG CCG TTT ACA-3’ Real-time PCR thermocycler conditions were: 95°C for 15 min, 40 cycles of 95°C for 15 s and 60°C for 1 min, final hold at 4°C. Quantification and analysis were done using Applied Biosystems StepOnePlus software.

Plant Growth Conditions

Wild-type *Arabidopsis thaliana* plants were of the Col-O ecotype. The *gx06* mutants, SALK lines 051510C and 085378C, in Col-O background, were ordered from the Arabidopsis Biological Resource Center. Plants were grown on soil in a 12-h-light/12-h-dark cycle at 25°C for assessment of mutant phenotypes.

Wild type *Nicotiana tabacum* plants were grown on soil in 12-h-light/12-h-dark cycles at 25°C in a separate chamber.

ACKNOWLEDGEMENTS

I would first like to thank Dr. Laura J. Olsen for her mentorship over the last four years – from high school to the present – and for her help in my pursuit of a career in medicine. I would also like to thank Dr. Nicola Harrison-Lowe for her preliminary work on GX06, Tara Nimkar for her collaboration, and the rest of the Olsen lab. I thank Dr. Erik Nielsen (University of Michigan, Ann Arbor) for providing the pCAM-35S-eYFP-c1 vector. Thank you to both my readers, Dr. Erik Nielsen and Dr. Yanzhuang Wang, for taking the time to evaluate my thesis. I am also grateful to Kitae Chang and Anna Cacciaglia for editing early drafts. Finally, I would like to thank my family for their continued support.

The research for this thesis was supported, in part, by the NSF Research Experience for Undergraduates, Honors Summer Fellowship, the UROP Summer Biomedical and Life Sciences

Fellowship, the Molecular, Cellular & Developmental Biology Summer Research Fellowship, the K.L. Jones Award for Research in Plant Biology, and by the National Science Foundation.

REFERENCES

Agne B, Meindl NM, Niederhoff K, Einwächter H, Rehling P, Sickmann A, Meyer HE, Girzalsky W, Kunau WH (2003) Pex8p: an intraperoxisomal organizer of the peroxisome import machinery. Mol Cell **11**: 635-646

Beers EP, Woffenden BJ, Zhao C (2000) Plant proteolytic enzymes: possible roles during programmed cell death. Plant Mol Biol **44**: 399-415

Beers EP, Jones AM, Dickerman AW (2004) The S8 serine, C1A cysteine and A1 aspartic protease families in *Arabidopsis*. Phytochem **65**: 43-58

Beevers H (1979) Microbodies in higher plants. Annu Rev Plan Physiol **30**: 159-193

Behari R, Baker A (1993) The carboxyl terminus of isocitrate lyase is not essential for import into glyoxysomes in an in vitro system. J Biol Chem **268**: 7315-7322

Bhattacharyya J, Das KP (1999) Molecular chaperone-like properties of an unfolded protein, α_S-Casein. J Biol Chem **274**: 15505–15509

Braverman N, Dodt G, Gould SJ, Valle D (1998) An isoform of Pex5p, the human PTS1 receptor, is required for the import of PTS2 proteins into peroxisomes. Hum Mol Genet **7**: 1195-1205

Brickner DG, Harada JJ, Olsen LJ (1997) Protein transport into higher plant peroxisomes. Plant Physiol **113**: 1213-1221

Bowden L, Lord M (1978) Purification and comparative properties of microsomal and glyoxysomal malate synthase from castor bean endosperm. Plant Physiol **61**: 259-265

Callis J, Vierstra RD (1989) Ubiquitin and ubiquitin genes in higher plants. Oxford Surv Plant Mol Cell Biol. **6**: 1-30

Chomczynski P, Sacchi N (1987) Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem **162**: 156-159

Dammai V, Subramani S (2001) The human peroxisomal targeting signal receptor, Pex5p, is translocated into the peroxisomal matrix and recycled to the cytosol. Cell **105**: 187-196.

Del Valle R, Soto U, Necochea C, Leighton F (1998) Detection of an ATPase activity in rat liver peroxisomes. Biochem Biophys Res Commun **156**: 1353-1359

Fujiki Y, Matsuzono Y, Matsuzaki T, Fransen M (2006) Import of peroxisomal membrane proteins: the interplay of Pex3p- and Pex19p-mediated interactions. Biochim Biophys Acta **1763**: 1639-1646

Fukao Y, Hayashi M, Nishimura M (2002) Proteomic analysis of leaf peroxisomal proteins in greening cotyledons of *Arabidopsis thaliana*. Plant Cell Physiol **43**: 689-696

Gietl, C (1990) Glyoxysomal malate dehydrogenase from watermelon is synthesized with an amino-terminal transit peptide. Proc Natl Acad Sci **87**: 5773–5777

Graham IA, Smith LM, Brown JMS, Leaver CJ, Smith SM (1989) The malate synthase gene of cucumber. Plant Mol Biol **13**: 673–684

Harrison-Lowe N, Olsen LJ (2006) Isolation of glyoxysomes from pumpkin cotyledons. Curr Protoc Cell Biol: 3.19. 11-3. 19. 18

Hayashi M, Aoki M, Kondo M, Nishimura M (1997) Changes in targeting efficiencies of proteins to plant microbodies caused by amino acid substitutions in the carboxy-terminal tripeptide. Plant Cell Physiol. **38**: 759–768

Hayashi M, Nishimura M (2003) Entering a new era of research on plant peroxisomes. Curr Opin Plant Biol **6**: 577-582

Helm M, Lück C, Prestele J, Hierl G, Huesgen PF, Frölich T, Arnold GJ, Adamska I, Görg A, Lottspeich F, Gietl C (2007) Dual specificities of the glyoxysomal/peroxisomal processing protease Deg15 in higher plants. Proc Natl Acad Sci **104**: 11501-11506

Kamada T, Nito K, Hayashi H, Mano S, Hayashi M, Nishimura M (2003) Functional differentiation of peroxisomes revealed by expression profiles of peroxisomal genes in *Arabidopsis thaliana*. Plant Cell Physiol **44**: 1275-1289

Kato A, Hayashi M, Mori H, Nishimura M (1995) Molecular characterization of a glyoxysomal citrate synthase that is synthesized as a precursor of higher molecular mass in pumpkin. Plant Mol Biol **27**: 377–390

Lamb JE, Riezman H, Becker WM (1978) Regulation of glyoxysomal enzymes during germination of cucumber. Plant Physiol **62**: 754-760

Lametschwandtner G, Brocard C, Fransen M, Van Veldhoven P, Berger J, Hartig A (1998) The difference in recognition of terminal tripeptides as peroxisomal targeting signal 1 between yeast and human is due to different affinities of their receptor Pex5p to the cognate signal and to residues adjacent to it. J Biol Chem **273**: 33635-33643

Lanyon-Hogg T, Warriner S, Baker A (2010) Getting a camel through the eye of a needle: the import of folded proteins by peroxisomes. Biol Cell **102**: 245-263

Lazarow PB, Fujiki Y (1985) Biogenesis of peroxisomes. Annu Rev Cell Biol **1**: 489-530

Lingard M, Bartel B (2000) *Arabidopsis* LON2 is necessary for peroxisomal function and sustained matrix protein import. Plant Physiol **151**: 1354-1365

Mercier JC, Grosclaude F, Dumas BR (1971) Structure primaire de la caséine α_{S1}-bovine, séquence complète. Eur J Biochem **23**: 41–51

Nayathi Y, Baker A (2006) Plant peroxisomes as a source of signaling molecules. Biochim Biophys Acta **1763**: 1478-1495

Nicolay K, Veenhuis M, Douma AC, Harder W (1987) A ^{31}P NMR study of the internal pH of yeast peroxisomes. Arch Microbiol **147**: 37-41

Olsen LJ, Harada J (1995) Peroxisomes and their assembly in higher plants. Annu Rev Plant Biol **46**: 123-146

Pool MR, Lopez-Huertas E, Horng JT, Baker A (1998) NADPH is a specific inhibitor of protein import into glyoxysomes. Plant J **15**: 1-14

Purdue PE, Lazarow PB (2001) Peroxisome biogenesis. Annu Rev Cell Dev Biol **17**: 701-752

Purdue PE, Zhang JW, Skoneczny M, Lazarow PB (1997) Rhizomelic chondrodysplasia puntata is caused by deficiency of human PEX7, a homologue of the yeast PTS2 receptor. Nat Genet **15**: 381-384

Reumann S, Ma C, Lemke S, Babujee L (2004) AraPerox. A database of putative *Arabidopsis* proteins from plant peroxisomes. Plant Physiol **136**: 2587-2608

Reumann S, Babujee L, Ma CL, Wienkoop S, Siemsen T, Antonicelli GE, Rasche N, Luder F, Weckwerth W, Jahn O (2007) Proteome analysis of *Arabidopsis* leaf peroxisomes reveals novel targeting peptides, metabolic pathways, and defense mechanisms. Plant Cell **19**: 3170-3193

Reumann S, Quan S, Aung K, Yang P, Mandandhar-Shrestha K, Holbrook D, Linka N, Switzenberg R, Wilkerson CG, Weber APM, Olsen LJ, Hu J (2009) In-depth proteome analysis of *Arabidopsis* leaf peroxisomes combined with in vivo subcellular targeting verification indicates novel metabolic and regulatory functions of peroxisomes. Plant Physiol **150**: 125-143

Steinberg SJ, Dodt G, Raymond GV, Braverman NE, Moser AB, Moser HW (2006) Peroxisome Biogenesis Disorders. Biochim Biophys Acta **1763**: 1733-1748

Titus DE, Becker WM (1985) Investigation of the glyoxysome-peroxisome transition in germinating cucumber cotyledons using double-label immunoelectron microscopy. J Cell Biol **101**: 1288-1299

Turley RB, Choe SM, Trelease RN (1990) Characterization of a cDNA clone encoding the complete amino acid sequence of cotton isocitrate lyase. Biochim Biophys Acta **1049**: 223–226

Twining SS (1984) Fluorescein isothiocyanate-labeled casein assay for proteolytic enzymes. Anal Biochem **143**: 30-34

Urquhart AJ, Kennedy D, Gould SJ, Crane DI (2000) Interaction of Pex5p, the type 1 peroxisome targeting signal receptor, with the peroxisomal membrane proteins Pex14p and Pex13p. J Biol Chem **275**: 4127-4136

Waterham HR, Keizer-Gunnick I, Goodman JM, Harder W, Veenhuis M (1990) Immunocytochemical evidence for the acidic nature of peroxisomes in methylotrophic yeasts. FEBS Lett **262**: 17-19

Yamauchi D, Terasaki Y, Okamoto T, Minamikawa T (1996) Promoter regions of cysteine endopeptidase genes from legumes confer germination-specific expression in transgenic tobacco seeds. Plant Mol Biol **30**: 321-329

Zhang J, Laudencia-Chingcuanco D, Comai L, Li M, Harada J (1994) Isocitrate lyase and malate synthase genes from *Brassica napus* L. are active in pollen. Plant Physiol **104**: 857-864