Identification of the DEG15 processing site in Aspartate Aminotransferase 3 (ASP3) in *Arabidopsis thaliana.*

Pallavi S. Abraham
Cell and Molecular Biology Honors Thesis
Department of Molecular, Cellular, and Developmental Biology
The University of Michigan
April 1st, 2013

Sponsor: Laura J. Olsen, Ph.D.

Readers:
Matthew Chapman, Ph.D.
Yanzhuang Wang, Ph.D

Undergraduate Senior Honors Thesis

Identification of the DEG15 processing site in Aspartate Aminotransferase 3 (ASP3) in *Arabidopsis thaliana.*

Pallavi S. Abraham

Table of Contents

ABSTRACT

The peroxisome is a highly specialized organelle found in eukaryotic cells. It is responsible for the catabolism of very-long-chain fatty acids and branched-chain fatty acids through β-oxidation. In order for the peroxisome to perform these specific functions, proteins must be imported from the cytosol, as the peroxisome does not have its own genome. The proteins are synthesized in the cytosol, and then, using a targeting signal, are transported to the peroxisome. There are two known peroxisomal targeting signals (PTSs): a carboxyl terminal PTS1 and an amino-terminal PTS2. Once in the peroxisome, PTS2-containing proteins undergo cleavage by the protease DEG15, to remove the PTS2. The DEG15 recognition site is predicted to be a cysteine residue conserved in many PTS2 proteins. Interestingly, Aspartate Aminotransferase 3 (ASP3), a known peroxisomal protein that contains a PTS2, does not include this conserved cysteine. There is little currently known about the DEG15 recognition site in ASP3 and the purpose of this study was to identify this processing site. Using site-directed mutagenesis, I found that a region of four amino acids, residues #38-41, are critical to DEG15 recognition of ASP3 and processing of the PTS2.

INTRODUCTION

Peroxisomes are ubiquitous organelles found in eukaryotic cells. They are responsible for the catabolism of very-long-chain fatty acids and branched-chain fatty acids through β-oxidation, and the biosynthesis of plasmalogens. Peroxisomes in plants have distinctive features compared to the organelle in other organisms, including the remarkable plasticity of their functions based upon the cell type they are found in. But peroxisomes in all cell types carry out the detoxification of hydrogen peroxide (H_2O_2) through the action of catalase. There are three different classes of plant peroxisomes: leaf peroxisomes, glyoxysomes, and unspecialized peroxisomes. Glyoxysomes are a specialized type of peroxisome, which are found in germinating seeds and are responsible for converting lipids into sugars via β-oxidation and the glyoxylate cycle. Leaf peroxisomes are involved in photorespiratory glycolate metabolism (Hayashi 2003). Glyoxysomes transition into leaf peroxisomes as the plants become photosynthetic. Therefore, during this transition the glyoxysomes begin to adopt leaf peroxisome functions such as photorespiratory glycolate metabolism (Hayashi et al., 2000).

There are two processes involved in the assembly of peroxisomes. First, the peroxisomal membranes are assembled from phospholipids from the endoplasmic reticulum. Next, peroxisomal matrix proteins are targeted from the cytosol and then imported to the organelle. The process of peroxisomal protein import depends on specialized proteins termed peroxins, which are encoded by *PEX* genes. Therefore the expression of multiple *PEX* genes is critical to peroxisomal biogenesis. There are 16 *PEX* genes that have been identified in humans (Scott 2011). Defects in these PEX

proteins or the non-PEX enzymes found within the peroxisomal matrix, can result in a range of disorders including the severe neonatal Zellweger syndrome and Type 1 Rhizomelic chondrodysplasia punctata (RCDP1). Symptoms of Zellweger syndrome include dysmorphic features, neurological abnormalities, hepatorenal and gastrointestinal dysfunction due to peroxisomal dysfunction (Pavelka 2010). RCDP1 is characterized by systemic shortening of the proximal bones, seizures, congenital cataracts and frequent respiratory infections (Braverman 1993).

Unlike organelles such as mitochondria and chloroplasts, peroxisomes do not contain their own DNA or their own protein synthesis machinery. Peroxisomal proteins are nuclear-encoded and must be synthesized in the cytosol and then post-translationally transported into the peroxisome (Lanyon-Hogg et al., 2010). Protein transport within a cell is controlled by amino acid-encoded targeting signals at the termini of cytosolically translated proteins. The targeting signal then directs the newly synthesized protein to the correct organelle. There have been two different peroxisomal targeting signals (PTSs) characterized thus far; PTS1 and PTS2. The PTS1 is a three-residue sequence located at the carboxyl terminus of the protein and the PTS2 is a nine-residue sequence usually located 20-30 residues from the amino terminus.

The PTS1 sequence was initially indentified as the three-residue sequence SKL at the extreme carboxyl terminus, but has since been expanded to the current consensus sequence [S/A/C]-[K/R/H]-[L/M] (Lanyon-Hogg el al., 2010, Reumann 2004). PEX5, a cytosolic receptor, recognizes proteins with a PST1 sequence, after synthesis in the cytosol. PEX5 brings the PTS1 protein (cargo protein) to the docking

complex (PEX14, PEX17, PEX13) on the peroxisomal membrane, and the cargo protein and PEX5 are translocated across the bilayer into the lumen. (Figure 1; Nito et. al). Once the cargo protein has been translocated across the membrane, PEX5 is recycled back to the cytosol to aid in further rounds of PTS1 protein trafficking (Purdue and Lazarow, 2001).

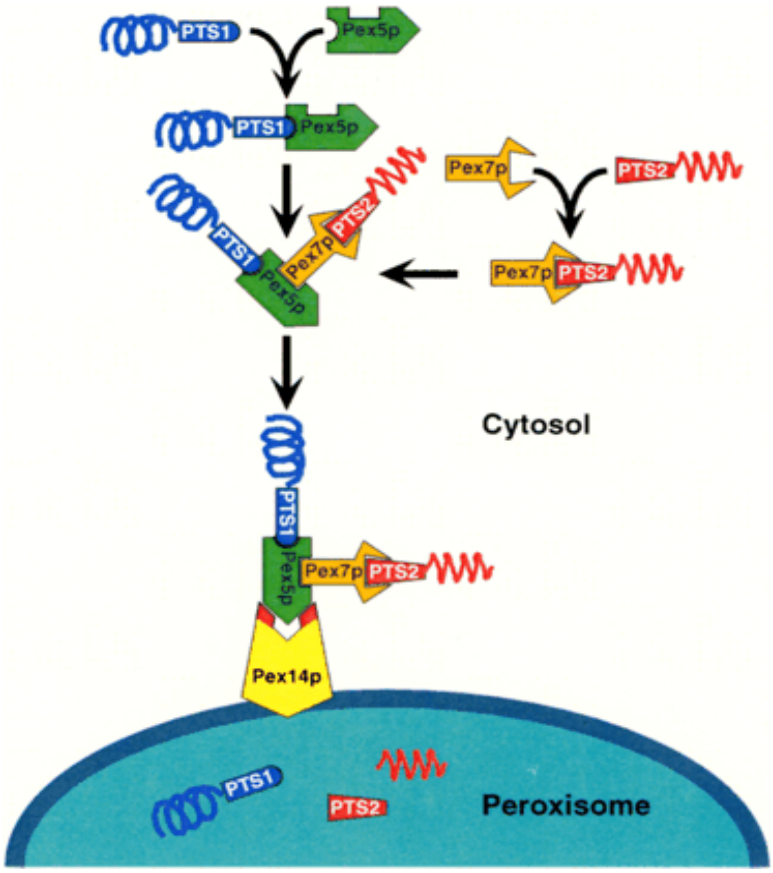

Figure 1. A schematic of peroxisomal protein import in *Arabidopsis thaliana*. All peroxisomally targeted matrix proteins are tagged with either a peroxisomal targeting signal located at the carboxyl terminus (PTS1) or at the amino terminus (PTS2). PEX5 and PEX7 are cytosolic PTS1 and PTS2 receptors, respectively. The receptor-cargo complex then interacts with PEX14, a peroxisomal membrane protein, which facilitates protein import into the peroxisomal lumen. Taken from Nito et. al. 2002.

The PTS2 consensus sequence was initially know to be [R/K]-[L/V/I]-x_5-[H/Q]-[L/A], where x denotes any amino acid (Kato, et al. 1998, Reumann 2004). But, similar to PTS1 proteins, this preliminary consensus sequence did not encompass all of the possible configurations of amino acids. A more complete consensus sequence was proposed by Petriv et al in 2004; [R/K]-[L/V/I/Q]-xx-[L/V/I/H/Q]-[L/S/G/A/K]-x-[H/Q]-[L/A/F] (Figure 2). Proteins with a PST2 sequence are recognized by PEX7, after synthesis in the cytosol. PEX7 brings the PTS2 protein to the receptor complex on the peroxisomal membrane, and the protein is translocated across the bilayer into the peroxisomal matrix. (Figure 1; Nito et. al). Once the cargo protein has been translocated across the membrane, PEX7 is recycled back to the cytosol to aid in further rounds of PTS2 protein trafficking (Purdue and Lazarow, 2001).

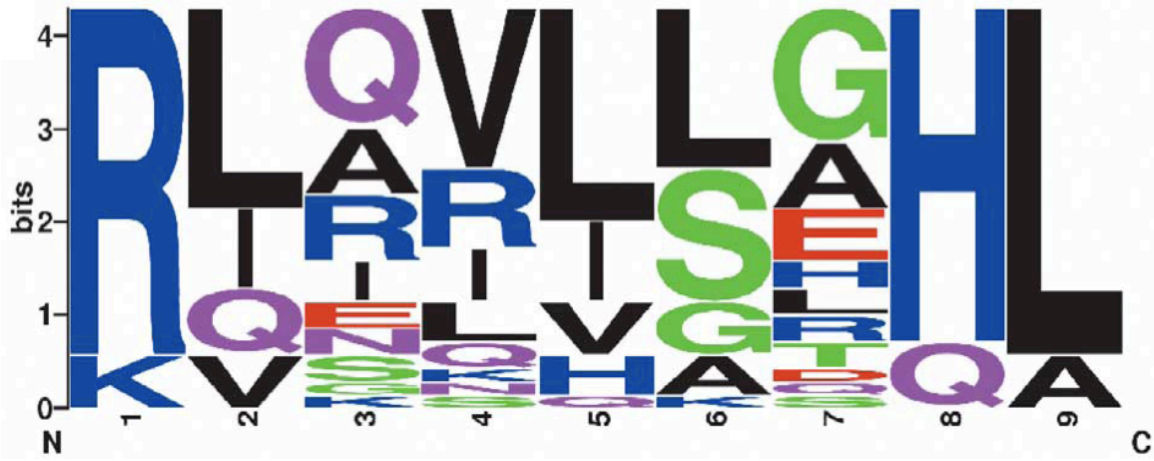

Figure 2. Schematic of currently recognized PTS2s. Data were collected from 30 distinct PTS2 motifs known to target proteins to peroxisomes. The x-axis signifies the position in the PTS2. The size of the letter represents the relative frequency of the residue in that particular position. Taken from Petriv, et al. (2004).

After protein translocation across the membrane, the targeting signal is cleaved off of PTS2, but not PTS1 proteins. This cleavage is performed by DEG15 in plants and the homolog TYSND1 in mammals (Helm et al., 2007). *A. thaliana* DEG15 mutants will not cleave the targeting signal of PTS2-containing proteins and therefore show phenotypes similar to mildly peroxisome-deficient plants (Lingard and Bartel, 2009). The cleavage site for DEG15 on PTS2 proteins is predicted to be a conserved cysteine residue located 15-20 amino acids downstream of the PTS2 targeting signal (Helm et al., 2007; Figure 4). Though found in many, this cysteine cleavage site is not found in all PTS2 proteins, leaving the cleavage site to be completely defined.

DEG15 belongs to the family of Deg/HtrA proteases and is found in two forms: a 72 kDa monomer and a 144 kDa dimer. Ca^{2+} is responsible for the shift between the monomer and dimer form, with Ca^{2+} presence supporting the dimer, which is responsible for removing the PTS2 pre-sequence. The monomer functions as a degrading protease that acts on denatured, non-functional proteins (Helm et al, 2007). Interestingly, the relationship between Ca^{2+} presence and CSY3 processing was tested by Kyle Helzer in 2012 but did not show an effect of Ca^{2+} on PTS2 cleavage (Helzer 2012). DEG15 is a serine protease with a catalytic triad formed by residues His-392, Asp-491, and Ser-580 (Schuhmann et al., 2008). Serine proteases contain a nucleophilic serine residue at their active site, where cleavage of the PTS2 occurs.

Aspartate aminotransferases catalyze the reversible transamination reaction between aspartate and α-ketoglutarate to give glutamate and oxaloacetate, with

pyridoxal 5'-phosphate (PLP) as an essential cofactor (Figure 3). The enzyme plays a critical role in the regulation of carbon and nitrogen flux in all organisms. In eukaryotes, aspartate aminotransferases along with malate dehydrogenase are part of a system for transporting reducing equivalents across organelle membranes called the malate-aspartate shuttle (Wilke 1996).

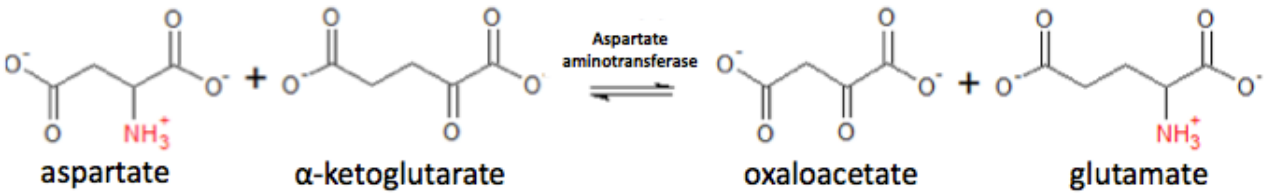

Figure 3. The reversible reaction catalyzed by aspartate aminotransferase. The enzyme transfers an amino group (highlighted in red) from aspartate to glutamate using α-ketoglutarate as a secondary substrate and producing oxaloacetate as a by-product. The enzyme functions as a homodimer, requiring pyridoxal 5'-phosphate (PLP) as a cofactor. The enzyme plays a crucial role in the moderation of carbon and nitrogen fluctuation in all organisms. Taken from Schutz and Corruzzi (1995).

In plants, at least five different isoenzymatic forms of aspartate aminotransferase have been found (Figure 4), each being associated with a particular subcellular compartment (e.g. cytosol, chloroplast, mitochondrion). The genes for each ASP2 and ASP4 encode cytosolic isoenzymes, with ASP2 being the most highly expressed (Schultz and Coruzzi 1995). The ASP1 gene is predicted to encode a mitochondrial isoenzyme, while ASP3 is predicted to encode either a plastid or peroxisomal enzyme based on transit peptide sequence analysis (Schultz and Coruzzi 1995). A fifth ASP gene (ASP5) was identified as a chloroplastic isoenzyme based upon in vitro chloroplastic protein import experiments (Wilke et.

al 1995, Murooka 2002). ASP3 shares the most sequence identity with ASP2, the cytosolic isozyme (Table 1). The percent sequence identity is based upon the full-length proteins, including the targeting signals. The lower percent homology could be due to the differences between the targeting signals.

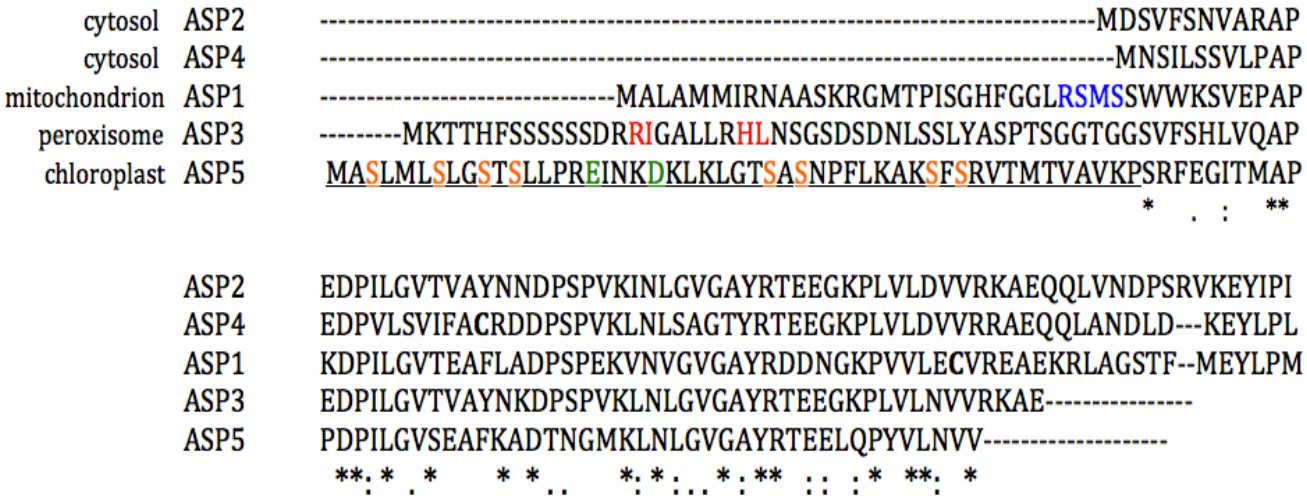

Figure 4. A DNA sequence alignment using ClustalW for the first 100 amino acids of the five aspartate aminotransferase isozymes found in Arabidopsis. They are all distinct nuclear-encoded genes with differential expression in various types of tissues. The order of alignment is based upon the homology of the first 100 amino acids. ASP2 and ASP4 encode cytosolic aminotranserases, and are therefore lacking the amino-terminal pre-protein sequences found in the organelle targeted isozymes. ASP1 shows the consensus sequence for a mitochondrial targeting signal, indicated in blue. ASP3 is peroxisomally targeted with the PTS2 highlighted in red. The chloroplastic targeting signal in ASP5 has no defined consensus sequence but is based upon the high percentage of serines (orange) and the two acidic amino acids in the first 50 residues (green).

Table 1. The percentage of identical amino acids between ASP3 and each of its isozymes. ASP3 shows the highest percent identity with the two cytosolically targeted isozymes.

Isozyme	Percent Sequence Identity between the ASP Isozymes and Peroxisomal ASP3
ASP1 (mitochondrion)	48.16%
ASP2 (cytosol)	81.17%
ASP4 (cytosol)	71.74%
ASP5 (chloroplast)	48.34%

The various aspartate aminotransferase isoenzymes play specific roles in: (a) converting newly formed organic nitrogen to the nitrogen carriers, Glu and Asp; (b) the formation of Asp used to generate several essential amino acids such as Asn, Met, Thr, and Ile; (c) the regeneration of carbon skeletons (α-ketoglutarate) for further primary nitrogen assimilation (Ryan and Fottrell 1974); (d) the assimilation of organic nitrogen compounds formed from photorespiration (Hatch and Osmond 1973); and (e) the shuttling of reducing equivalents between cells (Given 1980).

The chloroplastic targeting signal usually contains a high percentage of serine and threonine residues and a low percentage of acidic residues. ASP3 has 37.5% serines and threonine residues and only 6.7% acidic residues in its first 100 amino acids. Therefore, ASP3 was originally characterized as a chloroplastic enzyme based upon the amino acids in the presequence, but no import analysis was performed to confirm the prediction (Schulz and Coruzzi 1995). ASP3 is also a

putative peroxisomal protein, as it contains a strong PTS2 targeting signal (RI-X_5-HL). Based upon import analysis, ASP3 translocates to the peroxisome (Olsen Lab unpublished data).

Interestingly, the 48.9 KDa protein does not contain the conserved cysteine found in many other PTS2 proteins (Figure 5). Therefore, further analysis of the processing site is required, as most PTS2 proteins are cleaved by DEG15 after peroxisomal import at the conserved cysteine region. At present, little work as been done to concretely identify the critical residues of the ASP3 DEG15 processing site. The purpose of this study is to test previous processing site mutant studies and develop new ASP3 processing site mutants to definitively determine the location of the DEG15 processing site.

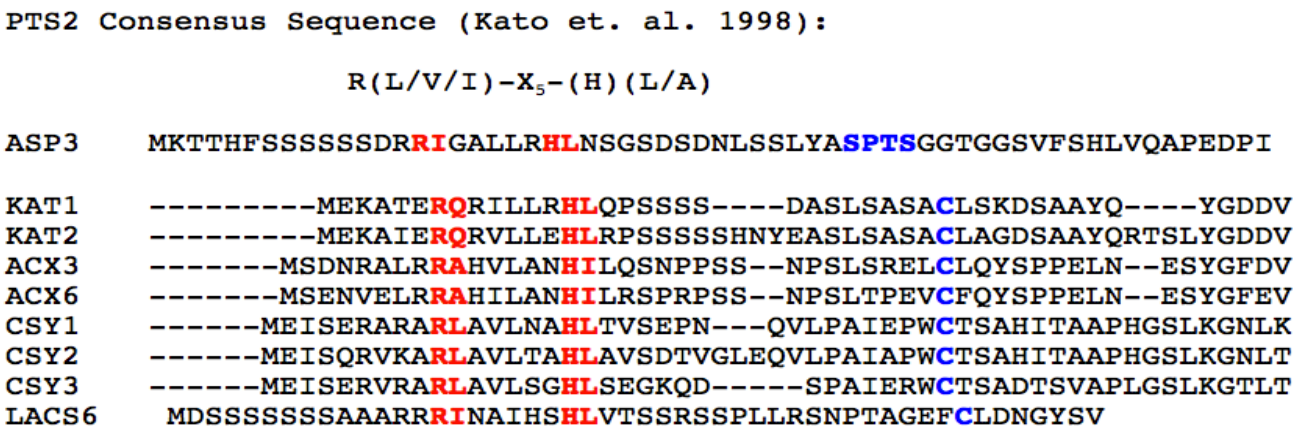

Figure 5. Alignment of ASP3 sequence surrounding the PTS2 signal and putative cleavage site using ClustalW. Other PTS2 containing proteins show are: KAT1, 3-keto-acyl-CoA thiolase 1; KAT2, 3-keto-acyl-CoA thiolase 2; ACX3, acyl-CoA oxidase 3; ACX6, acyl-CoA oxidase 6; CSY1, citrate synthase 1; CsY2, citrate synthase 2; CSY3, citrate synthase 3; LACS6, long-chain acyl-CoA synthetase. Residues highlighted in red match the consensus sequence (Kato, et al. 1998; Petriv et al. 2004) for PTS2 proteins, which is simpler than the consensus sequence found in Figure 2. The residues highlighted in blue are the putative cleavage site for DEG15 activity after peroxisomal import. ASP3 does not contain the conserved cysteine found in many other PTS2 proteins.

RESULTS

After import into the peroxisomal matrix, PTS2 proteins in *Arabidopsis thaliana* undergo proteolytic cleavage by DEG15 to remove the PTS2 presequence. This processing creates a mature protein, as DEG15 removes the segment from the amino terminus containing the targeting signal. In order for cleavage to occur, there must be DEG15 recognition of a particular set of residues, or a structural domain on the amino terminus of a PTS2 protein. The goal of this thesis project was to identify the DEG15 recognition site in ASP3.

Based upon sequence analysis, ASP3 is different from many other proteins that contain the conserved PTS2 (Figure 5). Approximately 15-20 amino acids after the PST2 signal, many other proteins contain a conserved cysteine where DEG15 cleaves off the targeting signal to form the mature form of the protein (Schuhmann et al. 2008). ASP3, unlike many of the other PTS2 proteins lacks this conserved residue, and therefore further mutant analysis was necessary to identify the processing site.

<u>Identification of ASP3 processing site</u>

Antoinette Williams and Meiyan Jin started studies of this processing site in 2009. Since ASP3 did not contain the cysteine residue expected to be involved in processing, mutations were made using SDM to change serine residues in the same region as where the cysteine residue would have been. Mutants created for this study follow the nomenclature system of: original amino acid, residue position in protein, mutant amino acid *or* Δ indicating a deletion of the original residue (e.g.

S38G signifies that the serine residue at position 38 has been changed to a glycine or or S38Δ signifies that the serine has been deleted)

The first few mutants created (S34A, S47G and S50A) did not show an effect on the percent processing by DEG15 (Appendix Table 1). This work was continued by Nicole Buller and Humphrey Fang to further identify which residues were critical in DEG15 processing. After several years of mutant analysis, 22 mutants had been created to alter the putative DEG15 processing site (Appendix Table 1).

Examination of initial mutation analysis

The most relevant mutant constructs were determined (Table 2A), based upon the work conducted by Nicole Buller, Antoinette Williams and Humphrey Fang (Appendix, Table 1). This determination was based upon the effect of the mutation measured by the average corrected percent processing (percent processing with DEG15 minus percent processing without DEG15) of each of the constructs as compared to the wild type. These results were then recalculated to be the average percent processing as a percentage of the wild type, to account for differences in wild type processing between different assays. A "Score" of "+ processing" or "- processing" was assigned to mutants whose results were close to the wild type, or clearly abolished processing, respectively (Appendix, Table 1). Based upon these results, Table 2A, indicates which constructs would be further analyzed. It was determined that mutation to residues 38-41 (SPTS) showed the greatest effect on DEG15 processing. This indicates that this region contains the putative DEG15 processing site. This thesis study began with a preliminary set of mutants to analyze this putative site; S38-41Δ: deletion of residues #38-41, P39A: change of

residue #39 from a proline to an alanine, T40A: change of residue #40 from a threonine to an alanine (Table 2B). The subsequent experiment used a set of serine mutants; S38G: change of residue #38 from a serine to an glycine, S41G: change of residue #41 from a serine to an glycine, S38/41G: change of both residues #38 and #41 from serines to glycines, S38/41G P39A: change of both residues #38 and #41 from serines to glycines and residue #39 from a proline to an alanine (Table 2C). The final experiment used deletion mutants; S38-41Δ: deletion of residues #38-41, S38Δ: deletion of residue #38, P39Δ: deletion of residue #39, T40Δ: deletion of residue #40, S41Δ: deletion of residue #41 (Table 2D).

Table 2A. List of most relevant constructs based upon original studies conducted by Humphrey Fang (HF) and Nicole Buller (NEB). The S38Δ and S41Δ constructs were designed by me (PSA). The initials found in column 1 indicate the person who originally designed the construct and conducted the preliminary studies. The name of construct, which indicates the mutated residues are found in column 2. The "average percent processing" is based upon at least an n=3 and "N.D." indicates that the study has not been performed yet. The "average processing as a % of wild type" sets the processing of each mutant as a percentage of the wild type in that particular assay to correct for difference between wild type processing between assays.

designer	name of construct	Residue number													average % processing	average processing as a % of wt
		34	35	36	37	38	39	40	41	42	43	44	45	46		
na	wt ASP3	S	L	Y	A	**S**	**P**	**T**	**S**	G	G	T	G	G	11.4	100.0
HF	S38G	S	L	Y	A	**G**	P	T	S	G	G	T	G	G	6.1	68
HF	P39A	S	L	Y	A	S	**A**	T	S	G	G	T	G	G	0.0	0.1
NEB	T40A	S	L	Y	A	S	P	**A**	S	G	G	T	G	G	8.9	94
HF	S41G	S	L	Y	A	S	P	T	**G**	G	G	T	G	G	2.5	38
NEB	S38G S41G	S	L	Y	A	**G**	P	T	**G**	G	G	T	G	G	4.3	35.6
HF	S38G S41G P39A	S	L	Y	A	**G**	**A**	T	**G**	G	G	T	G	G	0.8	6.8
NEB	S38-41Δ	S	L	Y	A	Δ	Δ	Δ	Δ	G	G	T	G	G	0.6	4.1
PSA	S38Δ	S	L	Y	A	Δ	P	T	S	G	G	T	G	G	N.D.	N.D.
HF	P39Δ	S	L	Y	A	C	Δ	T	S	G	G	T	G	G	3.0	19.0
NEB	T40Δ	S	L	Y	A	S	P	Δ	C	G	G	T	G	G	1.1	11.7
PSA	S41Δ	S	L	Y	A	S	P	T	Δ	G	G	T	G	G	N.D.	N.D.

Table 2B. Preliminary mutant studies from the list of most relevant constructs.

designer	name of construct	Residue number													average % processing	average processing as a % of wt
		34	35	36	37	38	39	40	41	42	43	44	45	46		
na	wt ASP3	S	L	Y	A	**S**	**P**	**T**	**S**	G	G	T	G	G	11.4	100.0
Preliminary Significant Mutations																
NEB	S38-41Δ	S	L	Y	A	Δ	Δ	Δ	Δ	G	G	T	G	G	0.6	4.1
HF	P39A	S	L	Y	A	S	**A**	T	S	G	G	T	G	G	0.0	0.1
NEB	T40A	S	L	Y	A	S	P	**A**	S	G	G	T	G	G	8.9	94

Table 2C. Serine mutant studies from the list of most relevant constructs.

designer	name of construct	Residue number													average % processing	average processing as a % of wt
		34	35	36	37	38	39	40	41	42	43	44	45	46		
na	wt ASP3	S	L	Y	A	**S**	**P**	**T**	**S**	G	G	T	G	G	11.4	100.0
Serine Mutations																
HF	S38G	S	L	Y	A	**G**	P	T	S	G	G	T	G	G	6.1	68
HF	S41G	S	L	Y	A	S	P	T	**G**	G	G	T	G	G	2.5	38
NEB	S38G S41G	S	L	Y	A	**G**	P	T	**G**	G	G	T	G	G	4.3	35.6
HF	S38G S41G P39A	S	L	Y	A	**G**	**A**	T	**G**	G	G	T	G	G	0.8	6.8

Table 2D. Deletion mutant studies from the list of most relevant constructs.

designer	name of construct	Residue number													average % processing	average processing as a % of wt
		34	35	36	37	38	39	40	41	42	43	44	45	46		
na	wt ASP3	S	L	Y	A	**S**	**P**	**T**	**S**	G	G	T	G	G	11.4	100.0
Deletion Mutations																
PSA	S38Δ	S	L	Y	A	**Δ**	P	T	S	G	G	T	G	G	N.D.	N.D.
HF	P39Δ	S	L	Y	A	C	**Δ**	T	S	G	G	T	G	G	3.0	19.0
NEB	T40Δ	S	L	Y	A	S	P	**Δ**	C	G	G	T	G	G	1.1	11.7
PSA	S41Δ	S	L	Y	A	S	P	T	**Δ**	G	G	T	G	G	N.D.	N.D.

Processing of ASP3 by DEG15

To investigate the processing of ASP3 by the peroxisomal protease DEG15, radioactively labeled ASP3 was incubated with DEG15 *in vitro* at 37°C for 4 hours, separated by SDS-PAGE, and visualized by autoradiography (Figure 6). Processing is indicated on the exposed gel by a shorter band (the mature protein) directly beneath the full-length protein, indicating that a portion (the amino terminus with the PTS2) has been cleaved off. In Figure 6, three experimental results are shown. Lane 2 shows 100% cleavage with no full-length band visible, lane 3 shows no

cleavage with no processed band visible and finally, lane 4 shows partial cleavage with a full-length and processed band visible. Partial cleavage of ASP3 is expected in this study.

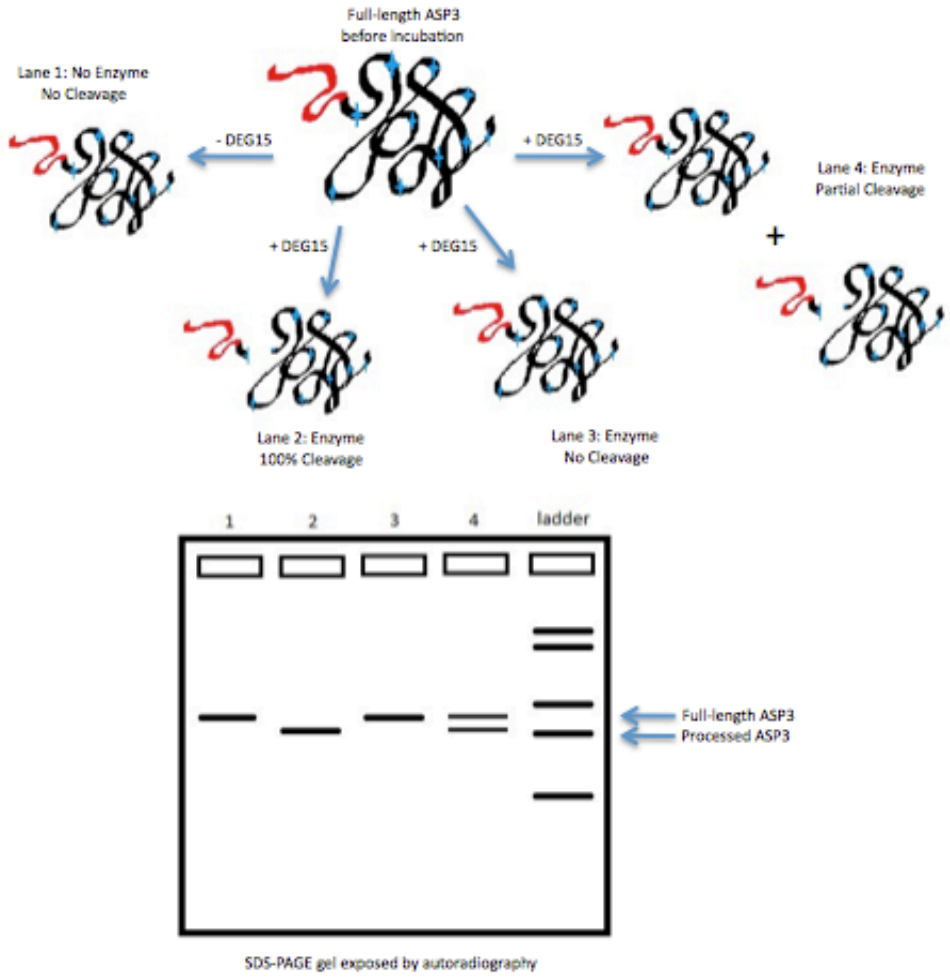

Figure 6. Diagram of a DEG15 protease assay. The ASP3 lysate was translated with radioactively labeled ^{35}S-Methionine using a wheat germ system to produce a radioactive protein. The blue stars represent the radiolabelled methionine residues in the newly translated protein. The PTS2 is indicated in red at the amino terminus. The labeled protein was then incubated in either the presence (Lanes 2-4) or absence (Lane 1) of DEG15 for 4 hours at 37°C. The results were analyzed by SDS-PAGE and visualized by autoradiography. The predicted experimental results show three possible outcomes after ASP3 is incubated with DEG15. Lane 2 shows 100% cleavage by DEG15, Lane 3 shows no cleavage in the presence of DEG15 and Lane 4 shows partial cleavage by DEG15. The full-length, unprocessed band appears longer than the mature, processed band.

Processing of Preliminary Mutants

In this preliminary study, processing was seen in the wild-type ASP3 as well as the T40A mutant (Figure 7). There was minimal processing seen in the S38-41Δ and some seen in the P39A. The processed form of ASP3 appeared at the expected size after cleavage. The quantification of processing was done using a phospho-screen, which when analyzed by a molecular imager, calculated the amount of radioactivity per band in the gel. This was then used to compare the amount of "mature" protein and the amount of unprocessed protein. The quantification of the mutant processing by DEG15 showed that S38-41Δ was processed at 10% the level of the wild type. The substitution mutants, P39A and T40A showed 51% and 83% processing, respectively. These results showed that the deletion of residues 38-41 (S38-41Δ) abolished processing and these two single substitution mutations in the SPTS processing site also had negative effects on overall processing.

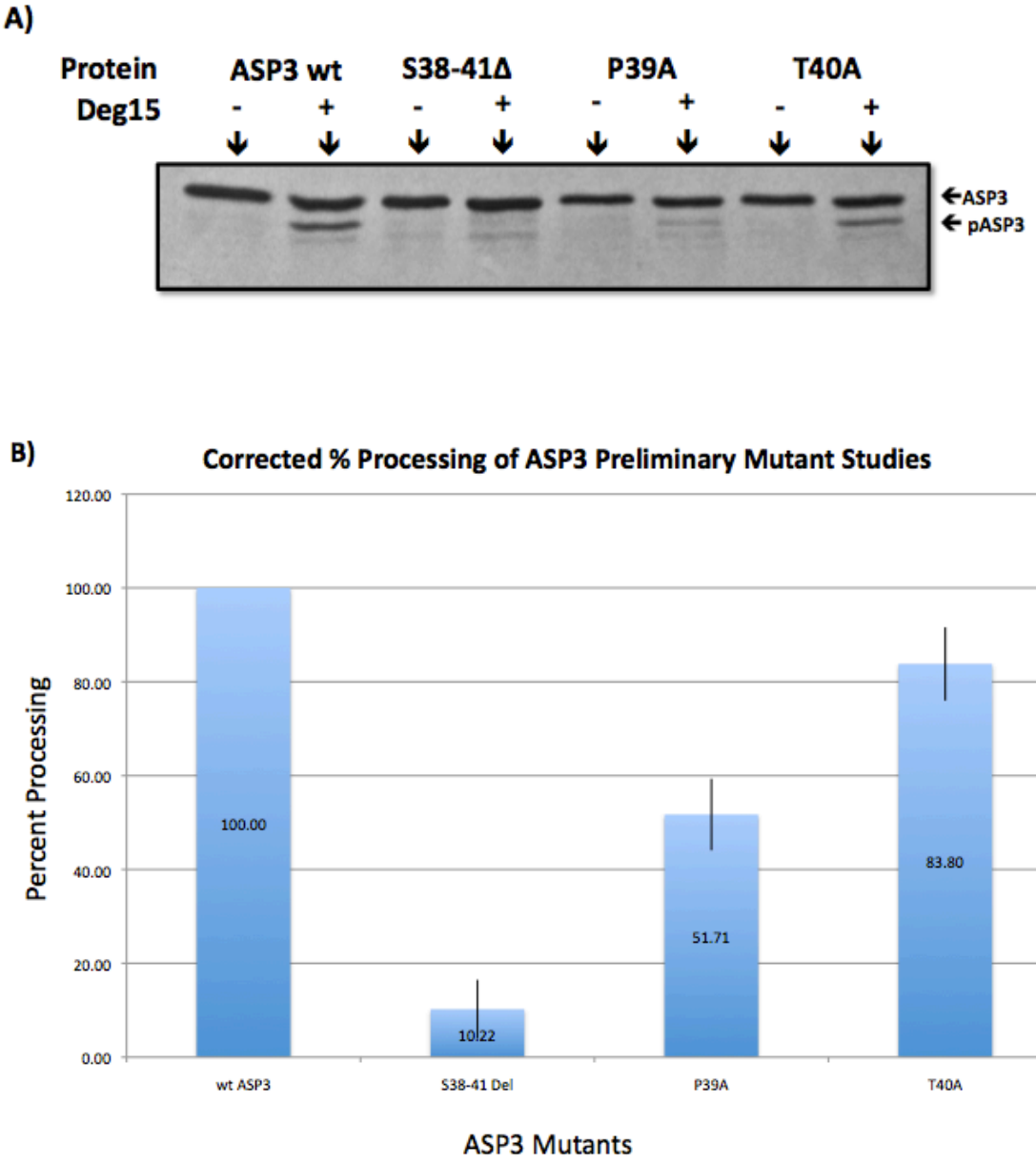

Figure 7. Processing of preliminary mutant constructs by Deg15. 10 µg DEG15 was incubated with 50,000 cpm radioactively labeled ASP3 for 4 hours at 37°C. Constructs studied: ASP3 wild-type, S38-41Δ: deletion of residues #38-41, P39A: substitution of residue #39 from a proline to an alanine, T40A: substitution of residue #40 from a threonine to an alanine. A) 10,000 cpm loaded per lane on SDS-PAGE gel with a 3-day exposure on film. (+) indicates 10 µg DEG15; (-) indicates no DEG15. "p" indicates processed form of the protein, after DEG15 cleavage while "ASP3" designates uncleaved wild-type or mutant protein. **B)** Quantification of Protease Assay. SDS-PAGE gel was exposed overnight on a phospho-screen and quantified. The percent processing was calculated using the ratio of processed protein to the total amount of protein (processed + unprocessed). The percent processing was corrected for background on the phospho-screen and set as a percentage of the wildtype. The background on the phosopho-screen was defined as the processing of the –DEG15 samples. Error bars represent one standard deviation; n=3.

Processing of Constructs Containing Serine Mutations

Processing was clearly seen in the wild-type ASP3 as well as thiolase, which was used as a positive control. Thiolase (KAT2) is a known PTS2 protein, which is consistently processed by DEG15 (Lingard and Bartel 2009). Based upon the results seen in the gel (Figure 8A), an intermediate amount of processing was seen for both the S38G and S41G mutants as compared to the wild type. The double (S38/41G) and triple (S38/41G P39A) mutants showed little to no processing as compared to the wild type. The quantification of the mutant processing by DEG15, using a phospo-screen (Figure 8B), showed that S38G and S41 yielded 41% and 47% processing, respectively, as compared to the wild type. This showed that both residues individually have an equal effect on processing by DEG15. The double mutant, S38/41G yielded 20.7% processing and surprisingly the triple mutant, S38/41G P39A did not have a significantly lower average percent processing, at 19.2% as compared to the wild type. This result was inconsistent with the results from the preliminary mutant study, which showed that the P39A mutation had a significant effect on processing. When the P39A mutation was added to the double mutant (S38/41G) it did not have its own effect, by further decreasing the percent processing. Instead we saw almost no change in the percent processing between the double and triple mutant.

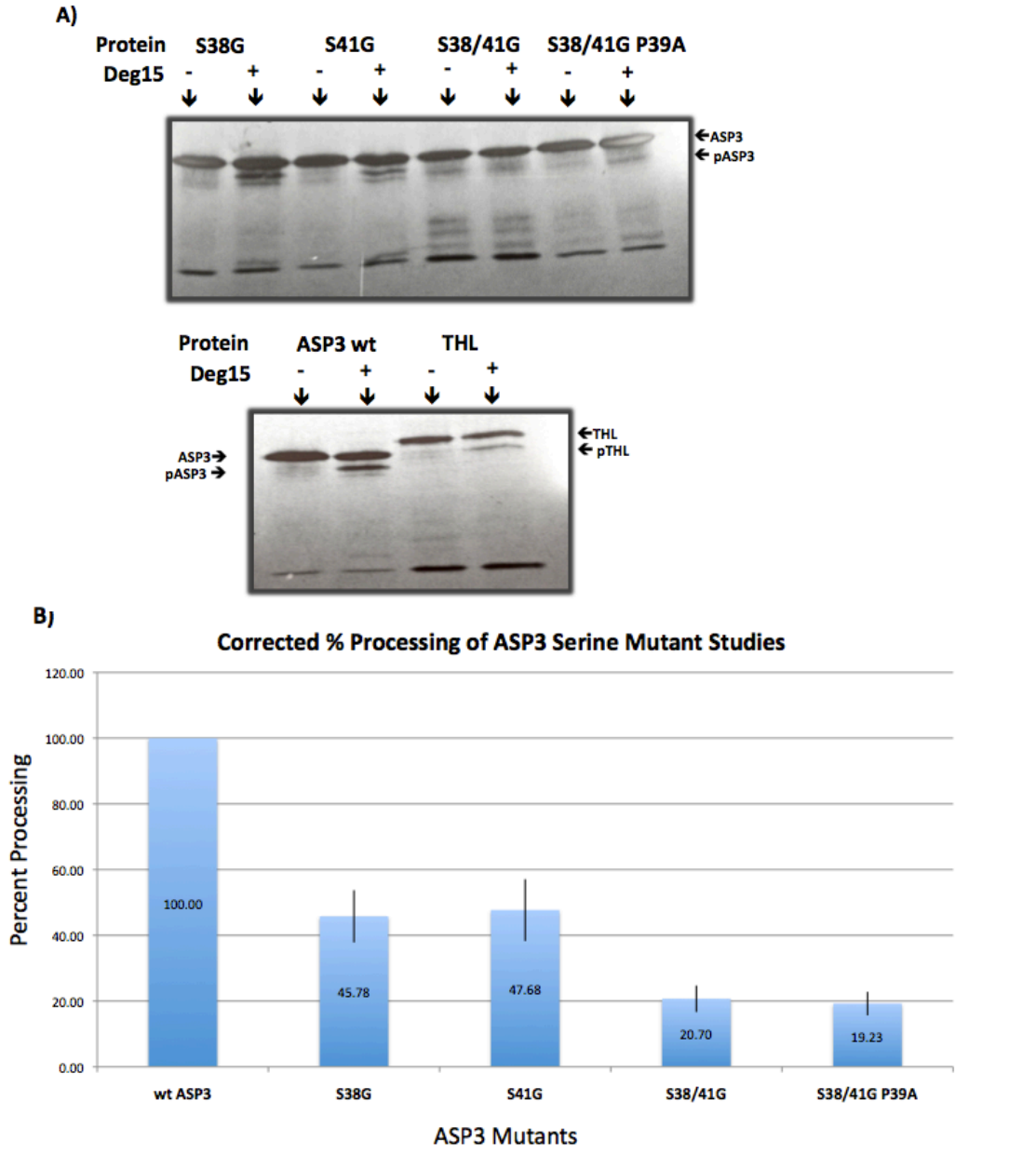

Figure 8. Processing of serine mutants by Deg15. 10 µg DEG15 was incubated with 50,000 cpm radioactively labeled ASP3 for 4 hours at 37°C. Constructs studied: S38G: substitution of residue #38 from a serine to an glycine, S41G: substitution of residue #41 from a serine to an glycine, S38/41G: substitution of both residues #38 and #41 from serines to glycines, S38/41G P39A: substitution of both residues #38 and #41 from serines to glycines and residue #39 from a proline to an alanine. ASP3 wild-type and thiolase (THL) were used as positive controls. **A)** 10,000 cpm loaded per lane on SDS-PAGE gel with a 3-day exposure on film. (+) indicates 10 µg DEG15; (-) indicates no DEG15. "p" indicates processed form of the protein, after DEG15 cleavage while "ASP3/THL" designates uncleaved wild-type or mutant protein. **B)** Quantification of Protease Assay. SDS-PAGE gel was exposed overnight on a phospho-screen and quantified. The percent processing was calculated as described in Figure 7B; n=3.

Processing of Constructs Containing Deletion Mutations

Processing was clearly seen in the wild-type ASP3 as well as thiolase, both of which were used as a positive control. Based upon the results qualitatively seen in the gel (Figure 9A), the deletion mutations showed large reduction in the amount of processing seen as compared to the wild type. The quantification of the mutant processing by DEG15 (Figure 9B), using a phospho-screen showed that S38Δ and S41Δ yielded 26% and 22% processing respectively. This reconfirms that both residues have an equal impact on DEG15 recognition and processing. The P39Δ and T40Δ mutations showed 11.3% and 16.1% respectively. This indicates that the proline residue is slightly more important to DEG15 processing than the threonine residue, as seen in the substitution mutations, P39A and T40A. Finally, the S38-41Δ mutation yielded results of 10.3% processing which reconfirmed the data from the preliminary mutant studies. In total, all of the deletion mutations had severe detrimental effects on the percent processing by DEG15.

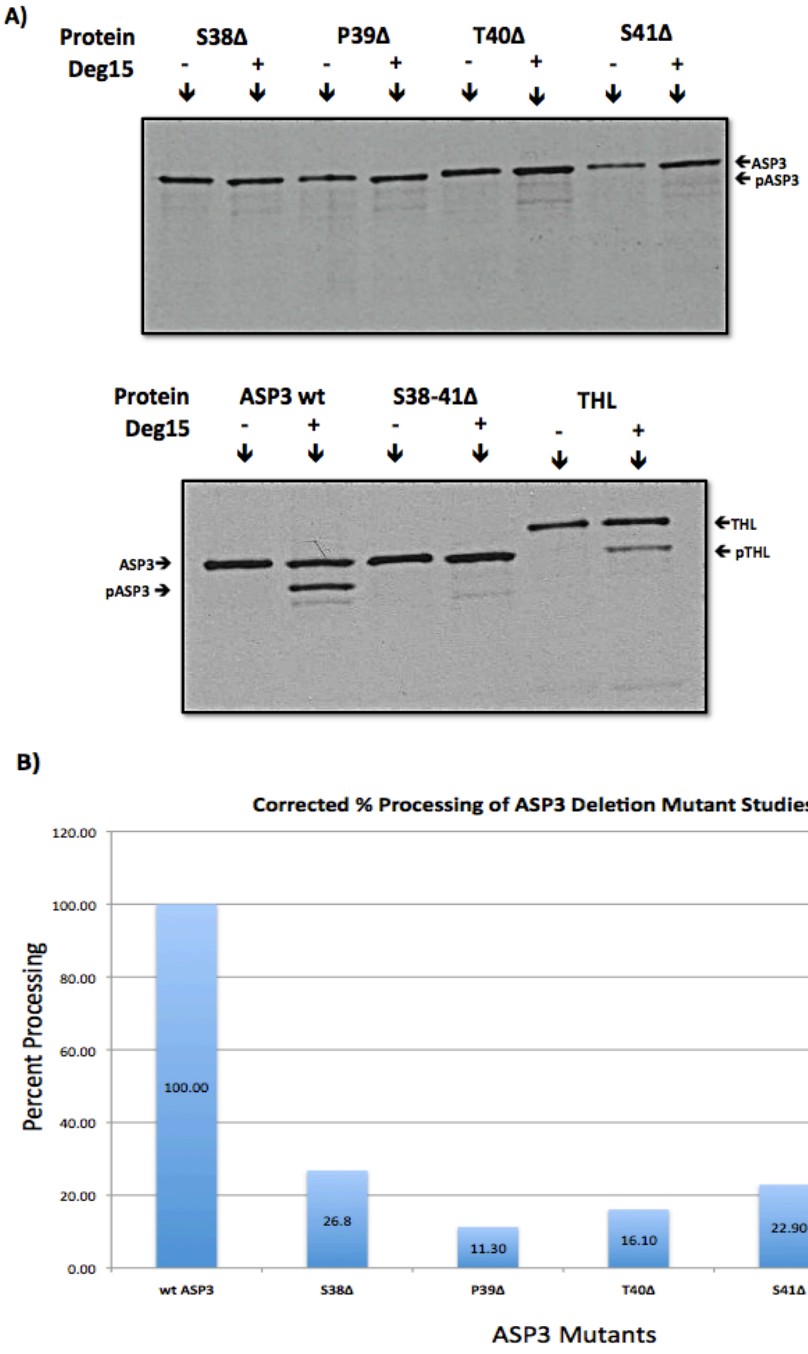

Figure 9. Processing of deletion mutants by Deg15. 10 µg DEG15 was incubated with 50,000 cpm radioactively labeled ASP3 for 4 hours at 37°C. Constructs studied: S38Δ: deletion of residue #38, P39Δ: deletion of residue #39, T40Δ: deletion of residue #40, S41Δ: deletion of residue #41, S38-4141Δ: deletion of residues #38-41. ASP3 wild-type and thiolase (THL) were used as positive controls. **A)** 10,000 cpm loaded per lane on SDS-PAGE gel with a 2-day exposure on film. (+) indicates 10 µg DEG15; (-) indicates no DEG15. "p" indicates processed form of the protein, after DEG15 cleavage while "ASP3/THL" designates uncleaved wild-type or mutant protein. **B)** Quantification of Protease Assay. SDS-PAGE gel was exposed overnight on a Phospho-screen and quantified. The percent processing was calculated as described in Figure 7B. The results shown are from a representative experiment, which was repeated three times.

MATERIALS AND METHODS

Site-Directed Mutagenesis

Primers (F: 5' CTGAGCTCTCTCTTTATGCACCGACATCGGGAGGC, R: 5' GCCTCCCGATGTCGGTGCATAAAGAGAGCTCAG) for *ASP3* S38Δ were designed using the strategy for the QuikChange® Site-Directed Mutagenesis Kit (Stratgene). This technique created a circular PCR product with a deleted codon corresponding to serine residue #38. The conditions for PCR were initial denature: 30 s at 95°C, denature: 30 s at 95°C, anneal: 60 s at 59°C (primer melting temperature (T_m)-9°C); extension: 6 min, 68°C; final extension, 1 min, 68°C. The amplification steps were cycled 16 times, using the high-fidelity *PfuTurbo*® DNA polymerase. After the mutagenic PCR amplification, the final product was treated with *Dpn1. Dpn1* digests the original template strand, which is methylated, unlike the newly made PCR product, leaving only the desired mutated DNA.

Primers for ASP3 S41Δ were designed using the Primer X protocol (F: 5' GGAGGCACCGGTGGTTCTGTTTTCTCTCATCTT, R: 5' TGTCGGAGATGCATAAAGAGAGCTCAGGTTGTC . This protocol (Olsen Lab Standard protocol) created a linear PCR product with a deleted codon corresponding to serine residue #41. The primers were initially phosphorylated with T4 polynucleotide kinase (PNK) to allow for subsequent ligation post-PCR. The conditions for PCR were initial denature: 5 min-95°C, denature: 30 s- 95°C, anneal: 30s-62°C (primer T_m-9); extension: 6 min, 72°C; final extension, 1 min, 72°C. The amplification steps were cycled 25 times, using the high-fidelity *PfuTurbo*® DNA polymerase. The final PCR product was *Dpn1* treated. The *Dpn1* digested PCR product was run on a 1%

agarose gel and the final product was extracted from the gel. This extraction removes the components of the digestion reaction, such as enzyme and bovine albumin serum, leaving only pure linear DNA. Finally, the ends of the linear mutant DNA were ligated using T4 DNA ligase to form a mutant plasmid (circular).

ASP3 expression in *E. coli*

The products from the site-direct mutagenesis were transformed into Top 10 F' competent *E. coli* cells. Using 10 μL PCR product, the cells were heat shocked for 60s at 42°C. After recovery in 250μL LB, the transformations were plated on agarose plates containing ampicillin (100 μg/mL). Colonies were selected at random and grown overnight at 37°C in 3mL LB. The DNA was extracted from the bacterial cells by lysing the bacterial cells and purifying the DNA (Olsen Lab Standard Protocol). The resulting DNA was sequenced using the standard M13 Reverse primers for the PCR-II-TOPO plasmid. Bacteria containing the plasmid with the correct mutation were chosen for growth in large-scale bacterial cultures (250 mL) and purified by cell-lysis and DNA precipitation with PEG (Olsen Lab Standard protocol).

In vitro Transcriptions

All transcription protocols were performed under RNase-free conditions. Before transcription, 10 μg ASP3 DNA (1μg/mL), was linearized with *Xba1* for 2 hours at 37°C . Linearization is necessary because transcription proceeds to the end of the DNA template, and linearization ensures that RNA transcripts of a defined length and sequence are generated. The linearization reactions were purified to remove the components of the linearization reaction, resulting in pure linear DNA. Transcriptions of *ASP3* and its mutants were performed by combining 2.5 μg

linearized DNA with 20 units of SP6 or T7 RNA polymerase in Transcription 5X Buffer, 2.5 μg BSA, 10mM DTT, 0.5mM rNTPs, 0.5mM GTP, 0.5mM GpppG analog, and 0.05mM RNasin. The 50μL transcription reactions were incubated at 37°C for 90 minutes. The final mRNA was extracted from the other reaction components and resuspended in 50μL 10mM DTT + 0.1U/μL RNasin (Olsen Lab Standard protocol).

In vitro Translations

Radiolabelled ASP3 was synthesized using a cell-free wheat germ extract system (Olsen Lab Standard protocol). Each tube contained 250μL wheat germ extract, 1M potassium acetate, 1mM amino acids without methionine, and radioactive ^{35}S-methionine (Specific Activity: 1175 Ci/mmol, Concentration: 10.0 mCi/mL). The translation reaction was incubated for 90 minutes at room temperature. 2μL nonradioactive methionine was added at the end of the translation to stop the reaction. To quantify the amount of ^{35}S-Methionine in the newly translated protein, 2 μL translation product was spotted onto Whatman glass microfiber filters (Cat. No. 1827021). The filters were washed twice for 5 minutes each with 10% TCA, 5% TCA and 95% ethanol and dried for 5 minutes under a heat lamp. Using a LS 6500 multi-purpose scintillation counter (Beckman) the cpm (counts per minute)/μL were calculated for each translation. Duplicates were counted for each translation reaction, and the results averaged (cpm/μL).

ASP3 Processing Assays

10μL radiolabelled proteins (5000 cpm/μL) were incubated with 10μg DEG15 in 25μL TYSND1 buffer (50 mM HEPES, 115 mM NaCl, 2 mM DTT; Olsen Lab Standard protocol) at 37°C for 4 hours. After the incubation period, the assay was

stopped with 50μL 2X SDS-PAGE sample buffer (20% glycerol, 10% β-mercapto-ethanol, 4% SDS), and the samples were boiled at 95°C for 10 min. To view the protein cleavage by DEG15, the reaction products were separated on a 10% SDS-PAGE gel and visualized by autoradiography.

The extent of cleavage by DEG15 was quantified using a phospho-screen and a Personal Molecular Imager (Bio-Rad). The percent processing was calculated using the ratio of processed protein to the total amount of protein (processed + unprocessed). The percent processing was corrected for background on the phospho-screen and set as a percentage of the wild-type level. The background on the phosorpho-screen was defined as the processing of the samples incubated without DEG15.

DISCUSSION

In *Arabidopsis thaliana,* PTS2-containing proteins are imported into the peroxisome via recognition by PEX7 and the translocation complex located on the peroxisomal membrane (Rehling et al., 1996; review by Lanyon-Hogg et al., 2010 and see Figure 1). After import, the PTS2 signal is cleaved off by the processing protease, DEG15, leaving a mature peroxisomal protein (Helm et al., 2007). The goal of this thesis project was to identify the DEG15 processing site in the PTS2 protein Asparate Aminotransferase 3 (ASP3).

Based upon the work of previous lab members (Table 1, Supplementary Material), eleven mutations that showed the most significant impact on DEG15 processing of ASP3 were selected for further studies. The S38-41Δ mutant construct showed an average of 0.6% processing of ASP3 as compared to the wild type, which showed an average of 11.4% processing. Compared to the S33-50Δ and S38-50Δ mutants, the S38-41Δ mutant was the smallest and most specific deletion and showed the lowest percent processing, indicating that the critical residues of DEG15 action were located in this region.

It is important to mention that most PTS2-containing proteins have a conserved cysteine residue involved in DEG15 recognition. The first cysteine residue found in ASP3 is located at residue 148, well into the mature protein. In proteins containing this conserved cysteine, the cysteine is usually located shortly after the PTS2. In the initial mutant analysis performed by Antionette Williams, this cysteine residue was mutated to a glycine (C148G- Appendix, Table 1) and did not

show an effect on ASP3 processing. This indicates that this cysteine is not involved in DEG15 recognition, and returned our attention to the residue 38-41 region.

Based upon the results of the first few mutants analyzed, a contrast between the deletion mutant S38-41Δ and the single substitution mutants P39A and T40A was apparent (Figure 6). The two substitution mutants each decreased processing, but to a lesser extent than the deletion mutant. This indicates that the proline and threonine could be involved in recognition of ASP3 by DEG15 but are not individually decisive for processing. The T40A substitution mutation showed less of an effect on processing than the P39A mutation, which may indicate that the threonine residue is less important to DEG15 recognition of the processing site than the proline. The proline residue structure could have a greater effect on the conformation of that particular region of the protein due to its cyclic structure, which can create kinks or bends in the protein that threonine cannot.

The next set of residues examined were the serines at positions 38 and 41 (Figure 7). Both individual substitution mutations S38G and S41G showed approximately equal reduction in processing, roughly 50%. This shows that neither was a greater contributor than the other. The double substitution mutant S38/41G showed a dramatic decrease in processing, down to only 20%, which indicates that the presence of at least one serine is necessary for partial DEG15 recognition. The S38/41G P39A mutation also yielded interesting results, as there was no significant decrease in processing as compared to the S38/41G mutation. This indicates that the proline residue is not as critical as the serine residues, as its substitution mutant did not produce a dramatic decrease as compared to the serine mutations. These

results were surprising, as the individual substitution of the proline to an alanine previously caused a dramatic decrease in processing (Figure 6).

When comparing the substitution mutants S38G, P39A, T40A, and S41G, we can see that the two serine mutations have the greatest effect on processing. The proline and the threonine mutations both have negative effects, but are less detrimental than the serine mutations.

The deletion mutant analysis showed that all of the deletion mutations (S38Δ, P39Δ, T40Δ, S41Δ and S38-41Δ) had negative effects on processing. The serine deletion mutations yielded similar results, but, unlike the substitution mutants, the serine deletion mutations were *less* detrimental to processing than the P39Δ and T40Δ mutations. As expected, the proline deletion mutation showed a significant decrease in processing due to the unique structural characteristics of proline residues. Interestingly, this significant decrease in processing is not seen with the construct containing the proline substitution mutation alone(P39A).

A concern when studying deletion mutants is a potential change the protein's tertiary structure. A removal of residues creates steric changes in a particular region that could affect the overall folding of the protein. Reduction in the percent processing of mutants could therefore be a result of misfolded proteins and not the result of reduced recognition of processing sites. Another possibility is that a change in tertiary structure could affect the conformation at the recognition site, which could be critical to protease binding and activity.

Another concern when analyzing data from the Bio-Rad phospho-screen is the proximity of the ASP3 full-length and processed bands. This close proximity

makes it difficult to accurately calculate protein concentrations of each band. This is a source of errors for all data points calculated in this study and potentially affected the comparative nature of the residue analysis.

In conclusion, the location of the DEG15 processing site has been narrowed to residues 38-41, SPTS. Substitution and deletion mutation analysis has shown that all four residues are critical to DEG15 recognition, as substitution or deletion of any of these residues has a negative effect on processing. The serine residues at either end of this region are equally critical, and the proline and threonine residues yielded conflicting results between the substitution and deletion mutation analysis.

This study is critical to our knowledge of PTS2-containing proteins as our results diverge from the commonly accepted theory that the targeting signal is cleaved at a conserved cysteine residue. This study provides a basis for potentially identifying the DEG15 cleavage region in the minority of PTS2 proteins that do not contain the conserved cysteine residue. An example of such a protein is long-chain acyl-CoA synthetase 7 (LACS7), which does not contain the conserved cysteine residue but does contain the tetrapeptide sequence SPTN soon after the PTS2 signal.

Potential future directions include peroxisomal protein import assays to test the PTS2 signal in ASP3. Import assays could be used to test which residues in the PTS2 signal are critical for import. Also, it would be interesting to test for localization to the peroxisome using fluorescence microscopy. Future directions for PTS2 proteins in general, include studying other PTS2-containing proteins that lack this conserved cysteine. It would be interesting to examine if there is another set of conserved residues that are involved in DEG15 cleavage.

APPENDIX

Table 1. Processing site mutants worked on from 2009-Summer 2011. The initials found in column 1 indicate the person who originally designed the construct; Nicole Buller (NEB), Humphrey Fang (HF), (MJ), Antoinette (ALW). The "score" designated in column 5 qualitatively indicates if processing occurs "+" or does not occur "-"

Primer Designer	ASP3 Mutant	Average Corrected % Processing	Partial Amino Acid Sequence Residue 29-50	Score
na	wt ASP3	11.4	S D N L S S L Y A S P T S G G T G G S V F S	+
MJ	S29-34Δ	8.3	- - - - - - L Y A S P T S G G T G G S V F S	
MJ	S29-41Δ	1.3	- - - - - - - - - - - - - - G G T G G S V F S	
MJ	S33/34Δ	9.1	S D N L - - L Y A S P T S G G T G G S V F S	
MJ	S33-50Δ	0.2	S D N L - - - - - - - - - - - - - - - - - - -	-
NEB	S38-41Δ	0.6	S D N L S S L Y A - - - - G G T G G S V F S	-
NEB	S38-50Δ	0.8	S D N L S S L Y A - - - - - - - - - - - - -	-
HF	P39Δ	3.0	S D N L S S L Y A S - T S G G T G G S V F S	
HF	P39Δ S38G S41G	5.9	S D N L S S L Y A - - T - G G T G G S V F S	
NEB	T40Δ	1.1	S D N L S S L Y A S P - S G G T G G S V F S	-
MJ	S34A	9.8	S D N L S A L Y A S P T S G G T G G S V F S	
NEB	S38C	11.2	S D N L S S L Y A C P T S G G T G G S V F S	+
NEB	S38G	6.1	S D N L S S L Y A G P T S G G T G G S V F S	
NEB	S38G S41G	4.3	S D N L S S L Y A G P T G G G T G G S V F S	
HF	S38G S41G P39A	0.8	S D N L S S L Y A G A T G G G T G G S V F S	-
HF	S38G S50A	4.8	S D N L S S L Y A G P T S G G T G G S V F A	
HF	P39A	0.0	S D N L S S L Y A S A T S G G T G G S V F S	-
NEB	T40A	8.9	S D N L S S L Y A S P A S G G T G G S V F S	
ALW	S41G	2.5	S D N L S S L Y A S P T G G G T G G S V F S	
ALW	S47G	11.4	S D N L S S L Y A S P T S G G T G G G V F S	+
MJ	S50A	9.1	S D N L S S L Y A S P T S G G T G G S V F A	
ALW	C148G	13.4	S D N L S S L Y A S P T S G G T G G S V F S	+
NEB	S38C P39L T40A S41G	3.9	S D N L S S L Y A C L A G G G T G G S V F S	

ACKNOWLEDGEMENTS

First, I would like to thank Dr. Laura Olsen for giving me the opportunity to work in the Olsen Lab. Without her words of wisdom and constant willingness to help, this senior thesis would not have been possible. Laura has not only been an amazing person to work for, but has shaped both my undergraduate education and my life goals. As a future educator, I hope to be able to inspire students as Laura has inspired me.

Secondly, I would like to thank the people who helped make the Olsen Lab such a fun and rewarding part of my undergraduate career. The previous lab members; Nicola Harrison-Lowe, Nicole Buller, Katie Baughman, Humphrey Fang & Sherry Shen, all of whom kindly answered my many questions and taught me most of what I know about laboratory science. The current lab members; Tara Nimkar, Nirali Shah, Ryan Cummings and Ivanna Murskyj who have made these last few years in the lab so enjoyable, filled with giggles and good times. And finally, Ali Dorchak and Kyle Helzer, two people who guided me through my thesis project with total patience and a sense of humor about my many lab and life disasters.

Next, I would like to thank my friends who have made my undergraduate career the amazing experience it was; my roommates from both 3E and pyramidhouse, friends from Michigan Manzil and LSA Student Government, and of course, my forever friends from grade school.

Finally, I would like to thank my family for the constant love and support. My parents have always been accepting of my goals and pushed me to challenge myself to be the best person I can be. For that, I am forever grateful. And last, my sister, my favorite loon, without whom my life would be extremely dull. I am so glad to have shared my last two years in Ann Arbor with you.

REFERENCES

Braverman, Nancy E., Ann B. Moser, and Steven J. Steinberg. "Rhizomelic chondrodysplasia punctata type 1." *et al., GeneReviews [Internet], University of Washington, Seattle, Seattle (WA)(1993–2001 Nov 16)[updated 2010 Mar 2]*(2010).

Givan, Curtis V. "Aminotransferases in higher plants." *The biochemistry of plants* 5 (1980): 329-357.

Hatch, M. D., and C. B. Osmond. "Compartmentation and transport in C4 photosynthesis." *Transport in Plants III*. Springer Berlin Heidelberg, (1976): 144-184.

Hayashi, Makoto, and Mikio Nishimura. "Entering a new era of research on plant peroxisomes." *Current opinion in plant biology* 6.6 (2003): 577-582.

Hayashi, Makoto, et al. "Functional transformation of plant peroxisomes." *Cell biochemistry and biophysics* 32.1-3 (2000): 295-304.

Helm, Michael, et al. "Dual specificities of the glyoxysomal/peroxisomal processing protease Deg15 in higher plants." *Proceedings of the National Academy of Sciences* 104.27 (2007): 11501-11506.

Helzer, K.T. (2012). Processing of Citrate Synthase by the Peroxisomal Protease DEG15 in *Arabidopsis thaliana*. B.S. Thesis. The University of Michigan: U.S.A.

Kato, Akira, et al. "Glyoxysomal malate dehydrogenase in pumpkin: cloning of a cDNA and functional analysis of its presequence." *Plant and cell physiology*39.2 (1998): 186-195.

Lanyon-Hogg, Thomas, Stuart L. Warriner, and Alison Baker. "Getting a camel through the eye of a needle: the import of folded proteins by peroxisomes."*Biology of the Cell* 102.4 (2010): 245-263.

Lingard, Matthew J., and Bonnie Bartel. "Arabidopsis LON2 is necessary for peroxisomal function and sustained matrix protein import." *Plant physiology* 151.3 (2009): 1354-1365.

Murooka, Yoshikatsu, Yukiko Mori, and Makoto Hayashi. "Variation of the amino acid content of." Arabidopsis seeds by expressing soybean aspartate aminotransferase gene." *Journal of bioscience and bioengineering* 94.3 (2002): 225-230.

Nito, Kazumasa, Makoto Hayashi, and Mikio Nishimura. "Direct interaction and determination of binding domains among peroxisomal import factors in Arabidopsis thaliana." *Plant and cell physiology* 43.4 (2002): 355-366.

Pavelka, Margit, and Jürgen Roth. "Peroxisomal Disorders." *Functional Ultrastructure* (2010): 138-139.

Petriv, Oleh I., et al. "A new definition for the consensus sequence of the peroxisome targeting signal type 2." *Journal of molecular biology* 341.1 (2004): 119-134.

Purdue, P. Edward, and Paul B. Lazarow. "Peroxisome biogenesis." *Annual review of cell and developmental biology* 17.1 (2001): 701-752.

Reumann, Sigrun. "Specification of the peroxisome targeting signals type 1 and type 2 of plant peroxisomes by bioinformatics analyses." *Plant physiology* 135.2 (2004): 783-800.

Ryan, Eamon, and Patrick F. Fottrell. "Subcellular localization of enzymes involved in the assimilation of ammonia by soybean root nodules."*Phytochemistry* 13.12 (1974): 2647-2652.

Schultz, C and G. Coruzzi. "The aspartate aminotransferase gene family of Arabidopsis encodes isoenzymes localized to three distinct subcellular compartments." *The Plant Journal*. 7 (1995): 61–75.

Scott, Camilla, and Simon Olpin . "Peroxisomal disorders." *Paediatrics and Child Health* 21.2 (2011): 71-75.

Wilkie, Susan E., Richard Lambert, and Martin J. Warren. "Chloroplastic aspartate aminotransferase from Arabidopsis thaliana: an examination of the relationship between the structure of the gene and the spatial structure of the protein." *Biochemical Journal* 319.Pt 3 (1996): 969.