Pointillism in plant systems biology:
I. Proteomic analysis of plant exosome-like particles
II. Amyloplast-binding puroindoline fusion proteins for recombinant protein expression.

Trevor Greenham

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements
for the Doctorate of Philosophy degree in Biochemistry

Department of Biochemistry, Microbiology and Immunology
Faculty of Medicine
University of Ottawa

ABSTRACT

Expanding upon our understanding of plant defense is critical, particularly with the perilous threats of climate change and overpopulation to our food security, health and well-being. In this study, we focused on plant defense using two distinct approaches. First, we performed a proteomic analysis of plant exosome-like nanoparticles in order to elucidate their defense related protein cargo. Secondly, we used a wheat antimicrobial protein, puroindoline, as a fusion partner for the expression of recombinant proteins in rice endosperm.

Plant exosome-like nanoparticles (ELP) were isolated from fresh tomato and subjected to mass spectrometry (MS) analysis. The ELPs were compared to fresh pressed tomato juice, and the proteins that were significantly upregulated in the ELPs were analyzed for their defensive properties. Bioinformatic analysis identified 30 proteins upregulated in the ELPs, with a majority of these being involved in plant defense.

Puroindoline is a protein found in soft wheat varieties. A unique feature of this protein is the presence of a tryptophan-rich domain, which causes it to localize and tether onto starch granule surfaces; a property we are seeking to exploit for recombinant protein isolation. We hypothesized that when expressed in a pin-null crop, such as rice, puroindoline along with its fusion partner will localize and adhere to starch granule surfaces. PIN fusions were expressed in rice, and their subcellular localization was determined by immunolocalization. It was observed that PIN localizes to rice starch

granules *in vitro* and *in planta*, and retains its starch granule binding abilities as a fusion partner.

To identify other possible starch granule binding fusion partners, an anhydrous cleavage method was developed that can scan dry biological materials for associated proteins, in this case the starch granule surface. Incubation of our cleavage reagent with isolated rice starch granules yielded several cleavage products as determined through SDS-PAGE. These cleavage products were compared with previous proteomic data of trypsin digested rice starch granules.

ACKNOWLEDGMENTS

I would like to thank my supervisor Professor Illimar Altosaar for providing me with a great environment full of exciting opportunities. Discovering the beauty of plant biochemistry is something I will always appreciate. I am grateful for the mentorship and friendship you have provided. I would also like to thank my thesis advisory committee members Dr. Danielle Carrier, Dr. Anne-Gaelle Rolland-Lagan, Dr. Ashok Kumar, Professor Linda Bonen and Professor Mary-Ellen Harper for their invaluable feedback, guidance and support.

I owe a large amount of gratitude to my many mentors in the lab over the years Dr. Adam Koziol, Dr. Shen Wan, Dr. Tonya Ward, Dr. George Styles and Ms. Jamie Kraft. I am very grateful that you were always available for insight and guidance which helped me immensely throughout my studies.

Finally I would like to thank my family and friends for their support.

TABLE OF CONTENTS

Contents

ABBREVIATIONS

7-HFBA: 7-heptafluorobutyric acid

97-1: Puroindoline^{+} rice line

A:
AP: Alkaline phosphatase

B:
BCIP: 5-bromo,4-chloro,3-indolylphosphate

C:
CHO: Chinese hamster ovary
CV: cultivar

D:
DAMP: Damage-associated molecular pattern
DC: Detergent compatible
db: database
DGDG: Digalactosyl diacylglyceride
DP: Aspartyl-Prolyl bond
DTT: Dithiothreitol

E:
ELP: Exosome-like nanoparticle
ESCRT: Endosomal sorting complex required for transport
EV: Extracellular vesicles

F:
FDR: False discovery rate
FITC: Fluorescein isothiocyanate

G:
GO: Gene ontology

I:
IGF-1: Insulin-like growth factor-1

L:
LFQ: Label free quantification

M:
MAMPs: Molecular-associated molecular pattern
MGDG: Monogalactosyl diacylglyceride
MS: Mass spectrometry

N:
NTA: Nanoparticle tracking analysis

O:
OMV: Outer membrane vesicle

P:
PAMPs : Pathogen associated molecular pattern
PBS: Phosphate buffered saline
PBST: Phosphate buffered saline tween-20
PCR: Polymerase chain reaction
PIN: Puroindoline seed specific protein

S:
SEM: Scanning electron microscopy

T:
TRD: Tryptophan rich domain

U:
UTR: Untranslated region

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION

1.1 General Introduction

With food shortages being seen around the world, crop yield has become a primary focus for plant scientists. Increasing yield and improving crop defense, from both an abiotic and biotic point of view, are the two main areas of focus when tackling the food shortage issue. The estimated global losses of crops to pests are estimated to be 50% for wheat, 26-29% for soybean, 31% for maize, 37% for rice and 40% for potatoes **(Oerke et al., 2006)**. These percentages are expected to increase 10-15% with each degree increase in global temperature. To date the primary strategy for crop defense is through the use of environmentally harmful pesticides. More recently, plant biotechnology techniques have been employed to help curb crop loss and increase yield.

Plant pathogens use a variety of strategies for plant infection. Pathogenic bacteria can enter through a wound in a plant, or through a plant's gas and/or water pores (stomata and hydathodes, respect.). Pathogenic fungi can infect a plant by entering through its epidermal cells, or by penetration with a haustoria structure, allowing it to feed off the plants' nutrients. Nematodes and aphids use a stylet, a piercing structure, inserting it directly into plant cells. Each of these pathogens has virulent effector molecules, which improves their pathogenic virulence. Seeing as plants are immobile, and the wide array of plant pathogens that exist, it has caused what is referred to as an evolutionary arms race. Unable to achieve the flight response upon predatory or pathogenic invasion, plants have one choice, to stand their ground and fight.

Physically immobile, plants also lack mobile defender cells and an adaptive immune system, but rather rely on innate immunity of each cell and systemic signaling responses triggered from sites of infection **(Jones, and Dangl, 2006)**. This immunity is broken down into two components: microbial/pathogen associated molecular patterns (MAMPs and PAMPs) and the expression of resistance genes (R –genes). MAMPs and PAMPs are based around cell surface receptors, which recognize molecules released from pathogens, causing a signaling cascade thus inducing downstream defense responses. Common molecular triggers include flagellin, Ef-Tu, lipoproteins and chitin **(Boller, and Felix, 2009)**.

In addition to the biotic stresses, plants must also deal with abiotic assaults caused by environmental stresses such as drought or salinity. Upon abiotic stress, the plant can release proteins into the extracellular space activating further the components of the plant's innate immune system. These molecules are termed damage-associated molecular patterns (DAMPs). DAMPs can be released from damaged or stressed cells. MAMPs and PAMPs elicit an immune reaction in response to pathogenic signals, DAMPs elicit an immune response through the plant's own protein secretion network.

In the past decade, considerable attention has been given to extracellular vesicles, and most notably plant exosome-like particles (ELP), as a new avenue of exploration for plant defense and plant stress responses.

1.2 Extracellular vesicles

Extracellular vesicles (EVs) are encapsulation structures composed of a lipid bilayer that are released from cells and involved in intercellular communication **(Thery et al., 2019)**. EVs have been shown to package a broad range of components such as protein, RNA and DNA. Different subtypes of EVs across different kingdoms have been identified and classified into exosomes, apoptotic bodies, shed microvesicles, bacterial outer membrane vesicles, bacterial membrane vesicles and plant exosome-like nanoparticles (ELPs). These vesicles differ in their biogenesis, size and their content **(Thery et al., 2019)**.

1.2.1 Mammalian exosomes

In the last decade, considerable attention has been given to elucidating the cargo, the role and the function of mammalian exosomes **(Figure 1.1)**. Initially thought to be involved in cell waste disposal **(Johnstone et al., 1987)**, recent findings have shown a broad range of functions spanning from intercellular communication to cross-kingdom regulation **(Monisha et al., 2015).** The wide array of function partly comes from their ability to transport DNA, mRNA, miRNA, lipids and proteins between cells and species **(Robbins et al., 2014).** Exosomes are thought to form through the fusion of multivesicular bodies with the plasma membrane. The hypothesis on biogenesis begins with the plasma membrane getting endocytosed into the cytoplasm, thereby generating small internal vesicles termed intraluminal vesicles.

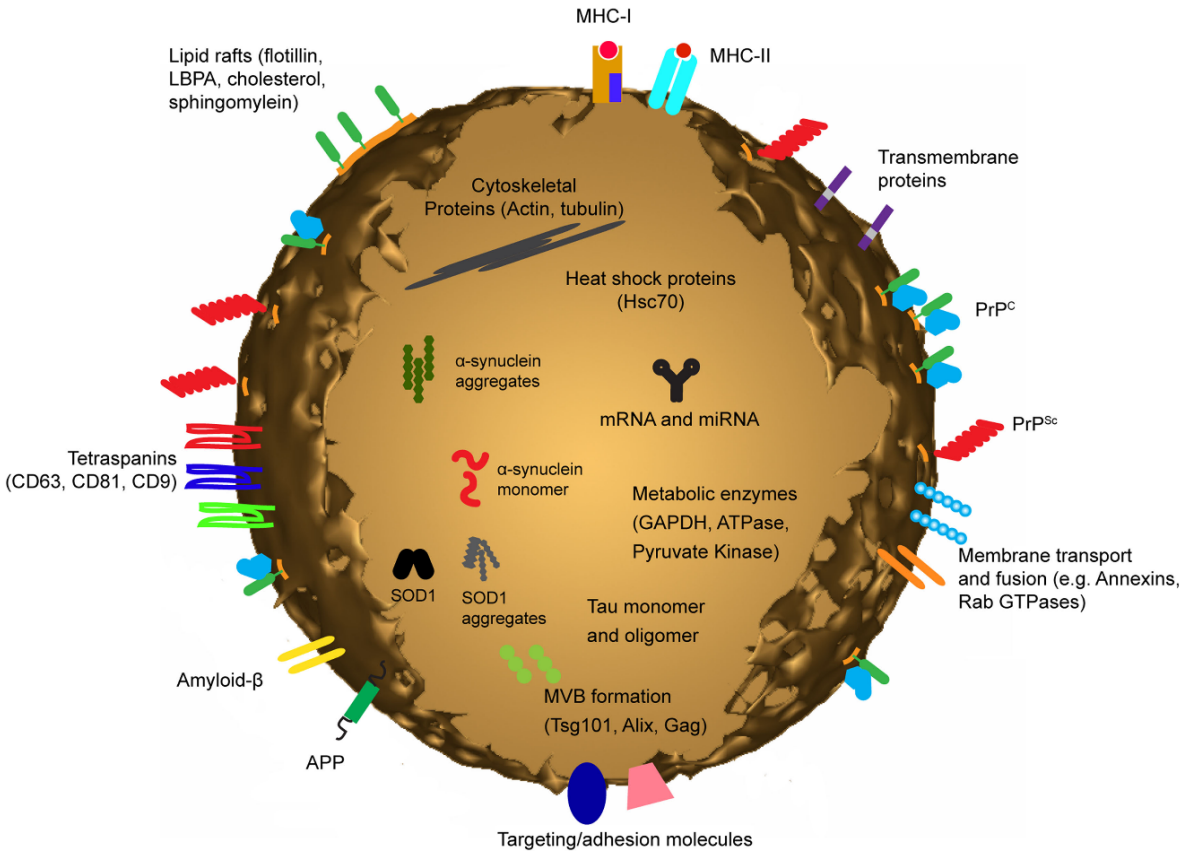

MHC-I
MHC-II
Lipid rafts (flotillin, LBPA, cholesterol, sphingomylein)
Transmembrane proteins
Cytoskeletal Proteins (Actin, tubulin)
Heat shock proteins (Hsc70)
PrPC
α-synuclein aggregates
mRNA and miRNA
PrPSc
Tetraspanins (CD63, CD81, CD9)
α-synuclein monomer
Metabolic enzymes (GAPDH, ATPase, Pyruvate Kinase)
Membrane transport and fusion (e.g. Annexins, Rab GTPases)
SOD1
SOD1 aggregates
Tau monomer and oligomer
Amyloid-β
MVB formation (Tsg101, Alix, Gag)
APP
Targeting/adhesion molecules

Figure 1.1. A schematic of a mammalian exosome.

Image courtesy of Bellingham, S et al., 2012.

Bellingham, S., et al. (2012). Exosomes: vehicles for the transfer of toxic proteins associated with neurodegenerative diseases? Frontiers in Physiology 3, 1-12.

These multivesicular bodies then fuse with the plasma membrane and release the exosomes into the environment **(Miyado et al., 2017).** Despite considerable advances in the field, the exact mechanism of biogenesis remains largely unknown **(Xu et al., 2016)**. The process of biogenesis, protein markers and size differentiate exosomes from other types of extracellular vesicles **(Thery et al., 2002; Thery et al., 2019).** The endosomal sorting complex required for transport (ESCRT) has been deemed necessary in the formation of multivesicular bodies and intraluminal vesicles and several ESCRT components are considered as mammalian exosomal biomarkers **(Kowal et al., 2016)**. Shed microvesicles are also linked to the ESCRT machinery; however, their production entails the outward budding of the plasma membrane **(Greening et al., 2018)**. Apoptotic bodies are produced during cell apoptosis where the entire cell is fragmented. Unlike other EVs, the process of biogenesis allows apoptotic bodies to encase entire organelles **(Elmore et al., 2007)**.

Exosomes have been isolated from a wide variety of biological fluids some of which include breast milk, blood, urine and saliva **(De Palma et al., 2016, Karlsson et al., 2016; Zheng et al., 2017)**. These nanovesicles have recently been explored for their delivery properties **(Gyorgy et al., 2015)**. Wolfers et al., demonstrated their ability to be packaged and used as a means of vaccination **(Wolfers et al., 2001)**.

1.2.2 Plant exosome-like nanoparticles (ELPs)

To date, most exosome studies have been performed using mammalian derived samples. Plant ELPs, though discovered approximately 15 years previous to mammalian exosomes **(Halperin and Jensen, 1967)**, have received little attention. A main reason for their neglect was due to the preconception that plant ELPs were simply artifacts

generated upon cell disruption during apoplastic fluid extraction. It was also thought that plant cell walls would inhibit ELP secretion, providing additional reasoning for discounting their existence. Recently, ELPs have since been found in various plant species such as sunflower, Arabidopsis, grapes, ginger, carrots, coconut, grapefruit and other plant species **(Mu et al., 2014),** and do not appear to be limited by the cell wall. Although they share similar size and apparent function to their mammalian counterparts, their biogenesis *in planta* is unknown. It has been hypothesized that these plant made nanoparticles are derived from protoplast plasma membranes, however, it has yet to be confirmed. One of the first attempts at isolating ELPs was performed on water-imbibed sunflower seeds, where vesicles ranging from 50-200nm were observed using transmission electron microscopy **(Regente et al., 2009)**. A similar finding was observed in olive pollen-derived samples where the presence of nanovesicles was detected in the stigmata exudates and pollen germination medium **(Prado et al., 2014)**. Further evidence of plant ELPs was recently reported in leaf apoplastic fluids of Arabidopsis **(Rutter and Innes, 2017)**. Since these findings, plant ELPs are gaining traction, and more resources are being invested into elucidating their properties and function.

1.2.3 Plant exosome-like nanoparticle proteomics

Proteomics, the large-scale study of proteins, is a leading tool for quantifying the relative expression of up to thousands of proteins within a single sample. Not only will the identification of proteins provide insight, but their relative expression will also help elucidate the biological functions and processes being impacted, most notably for plant defense and crop yield purposes. There are two approaches for protein identification: bottom-up and top-down proteomics. Bottom-up approaches use digestion of proteins

prior to MS analysis whereas top-down uses intact proteins. The bottom-up approach is applied for a complex mixture of proteins whereas top-down requires simple protein samples. There are multiple approaches for quantitative proteomics including chemical labeling, metabolic labeling and label free quantification (LFQ). LFQ is a cost effective method, as it does not require isotopically heavy peptides which are necessary for metabolic labeling. The accuracy of LFQ is, however, limited by sample variability **(Xie et al., 2011)**. In bottom-up proteomics, the comparison of empirical spectra to theoretical spectra allows for peptides to be assigned to their source protein **(Eng et al., 1994)**.

Proteomics has been used to identify the protein cargo of plant ELPs. To our knowledge there have been two proteomic studies on plant ELPs. The first was on *Arabidopsis thaliana* leaf apoplast ELPs **(Rutter and Innes, 2017)**. In their study they identified 170 proteins associated with their ELPs. Comparing the Gene Ontology (GO) of the ELPs to the whole proteome, it was observed that the ELPs were enriched for proteins involved in biotic and abiotic stress. From the proteins found in the ELPs, 26% were involved directly with responses from stress stimuli whereas this percentage drops to 11% for the entire proteome. These results suggest that plant ELPs are specialized for stress response and host defense purposes. Furthermore, this group found that the abundance of ELPs more than doubled in the presence of a plant pathogen leading to the hypothesis that plant ELP secretion is a general immune response.

The second proteomic study was performed on ELPs isolated from sunflower seedlings **(Regente et al., 2009)**. This study identified 278 proteins in isolated ELPs. Similar to ELPs from Arabidopsis, sunflower ELPs yielded a significant amount of cell wall remodeling and defense related proteins. It is interesting to note that exosomes

isolated from fungi and gram-positive bacteria, both with cell walls, also contain cell wall modifying proteins **(Albuquerque et al., 2008)**. It has been hypothesized that for ELPs to cross the cell wall remodeling may be required, though it is still unknown. Leaving leaves alone for the time being, I chose a more fruitful target to explore, the fruit of the tomato plant. I turned my attention to the tomato fruit itself, because the 'Flavr Savr' tomato was the first commercially grown genetically engineered food licensed for human consumption **(Morin et al., 2008)**.

1.3 Molecular pharming

Molecular pharming refers to the production of recombinant proteins in genetically modified plant hosts. The ability to express transgenes in plants dates back to 1986, where human growth hormone was expressed in transgenic tobacco **(Barta et al., 1986)**. Plant recombinant protein production has proven its advantages over traditional expression systems such as mammalian Chinese hamster ovary (CHO) cells, yeast, insect, microbial and transgenic animal hosts. Although these platforms are widely used, the main downside of these systems is the high capital requirements. These capital requirements are usually in the form of bioreactors and downstream purification, usually heavily dependent on chromatographic techniques, of the target protein. It is estimated that the production costs associated with plant recombinant protein expression is 2-10% that of microbial expression systems and 0.1% the cost of using mammalian cell cultures **(Chen et al., 2014)**. Butyrylcholinesterase when produced recombinantly in *Nicotiana benthamiana* (tobacco) is approximately 20 times less costly than when obtained from blood **(Dirisala et al., 2017)**. Cellulase enzymes for ethanol production when produced in plant-based systems are on average 30% the cost of fungal produced cellulases **(Tuse**

et al. 2014). Since the capital advantages are quite clear, several recombinant proteins are currently being produced for pharmaceutical use. Plant molecular pharming companies are seeing interest from big pharma, further proving their disruptive nature in this field. Protalix Bioscience, an Israeli based molecular pharming company, has received funding from Pfizer to produce taliglucerase alpha for the treatment of Gaucher disease in carrot cell cultures. Medicago, a Canadian biotechnology start-up received considerable funding from tobacco giant Phillip Morris for the production of the hemaglutinin protein of the H5N1 influenza vaccine in tobacco plants **(Maxmen et al. 2012)** and was acquired by Tanabe Mitsubishi in 2013.

1.3.1 Puroindoline

One biotechnology approach for plant defense was achieved through the genetic engineering of puroindoline, a wheat antimicrobial protein, into rice. The transgenic rice lines exhibited a higher resistance to microbial infection when compared to their non-GMO counterpart. Upon further analysis of these lines, it was discovered that the antimicrobial protein, puroindoline, elicited a defense property, partly due to its unique deposition site in plant kernels, on the starch granule surfaces. This led us to ask the following question. Can molecular farming be designed to reduce the cost of expensive biologics even farther? A difficult question to pose, but the answer may lie in soft wheat. In hexaploid wheat, grain hardness is determined by a hardness locus found on chromosome arm 5D. This locus contains the friabilin genes, and more specifically the antimicrobial puroindoline (PIN) genes. In *Triticum turgidum* L. cv. Langdon, a durum wheat variety, this section of the genome is not present and thus the grain softening PIN proteins are not present. The absence of PINs results in a hard grain texture. PINs are

being researched extensively due to their influence in altering grain textures, which ultimately dictates the end usage of wheat **(Ali et al., 2015)**. It has been observed that mutations within PIN proteins can cause change in the overall grain hardness and texture **(Bloch et al., 2001)**. Transgenic studies expressing PIN in hard wheat and PIN-null rice led to a significant softening of grain texture. Furthermore, RNAi silencing of PIN in soft wheat led to an increase in grain hardness **(Gasparis et al., 2011)**.

PIN proteins are basic, cysteine rich with a molecular weight of 13 kDa, initially isolated from soft wheat kernels using Triton X-114 phase partitioning **(Gautier et al., 2000)**. One of the most unique features of PIN proteins is this presence of a tryptophan rich domain (TRD) **(Figure 1.2)**.

PIN-a contains the domain (-WRWWKWWK-) at positions 38 to 45, and PIN-b contains a slightly truncated version (-WPTKWWK-) at positions 39 to 45 **(Day et al., 2006)**. PIN proteins usually accumulate in the wheat endosperm, and more specifically, accumulate on starch granule surfaces. Upon maturation, the wheat endosperm contains lipid remnants from amyloplast membranes. During seed desiccation these membrane remnants collapse and deposit themselves on the surface of starch, resulting in lipid coated granules **(Rosicka-Kaczmarek et al., 2015)**. This lipid coat interacts with the surrounding protein matrix, and is thought to be the main area influencing grain hardness **(Biswas and Marion, 2006)**. The most abundant lipids found on starch granule surfaces in wheat are monogalactosyldiacylglycerols (MGDG) and digalactosyldiacylglycerols (DGDG), both of which are found in high quantities in amyloplast membranes. It has been shown that PINs have a high affinity for these lipids **(Pauly et al., 2013)**. Due to the lipid binding properties of tryptophans, it is hypothesized that PIN interacts with

MKALFLIGLLALVASTAFAQYSEIVGSYDVAGGGGAQQCPLETKLNSCRNYLLD
RCSTMKFPVTWRWWKWWKGGCLELLGECCSQLGQMPPQCRCNIIQGSIQGDLSG
IFGFQRDRASKVIQEAKNLPPRCNQGPPCNIPGTIGYYW

MKTLFLLALLALVASTTFAQYSEVGGWYNEVGAGGGSQQCPLERPKLS
SCKDYVMERCFTMKDFPVTWPTKWWKGGCEHEVREKCCQQLSQIAPQC
LCDSIRGMIQGKLGGFFGIWRGDVFKQIQRAQSLPSKCNMGADCKLPS
GYYW

Figure 1.2. The amino acid sequence of Puroindoline

Puroindoline-a accession CAB89541.1 (top) and Puroindoline-b accession CAC33799.1 (bottom) from the species *Triticum monococcum*. The Tryptophan Rich Domain is highlighted in yellow.

starch granule surfaces by binding their lipid coats. This hypothesis was verified through tryptic shaving of wheat starch granule surfaces. Briefly, starch granules were water washed to remove any loosely associated proteins followed by incubation with trypsin. Mass spectrometry analysis confirmed that the TRD of PIN remains bound to the starch granule surfaces, while the N- and C- terminal portions of the protein did not **(Wall et al., 2010)**. This study demonstrated that the tethering point of PINs to the starch granule surface is achieved using the TRD. It is this tethering ability we aim to exploit as a fusion partner for the production of recombinant proteins in plants.

1.3.2 Rice as a recombinant protein expression host

Rice is one of the most important foods for over half the world's population. The rice seed has been used as an efficient location for the production of recombinant proteins due to its high biomass yield, its self-pollinating properties, it is hypoallergenic and can be conveniently scaled up. The rice kernel at the mature stage consists of four different tissues; the starchy endosperm, basal endosperm transfer cells, aleurone layer, and the embryo. The endosperm comprises approximately 90% of the total seed weight, in which 80-90% is attributed to starch granules. Seed proteins account for 7-15% and are usually accumulated in the endosperm in the form of protein bodies **(Takaiwa et al., 2017)**. The major rice storage proteins are glutelins, accounting for 60-80% of total seed protein. To date, several high value recombinant products have been expressed in rice endosperm, perhaps the most notable being lactoferrin, produced by Ventria Bioscience.

One advantage of seed based protein expression systems and starch granule surface recombinant protein deposition is that the starch granules can be isolated from the endogenous dry host cellular components using aqueous-free methods such as milling

and air-classification. Together, these technologies can isolate starch granules from the other cellular components based on their size and density. These technologies are a fraction of the cost of aqueous based methods, and provide a dry environment where the recombinant fusion protein remains stable. There are several dry and wet milling techniques, but only a select few reduce particles to the required size of <10um diam. The preferred technique preserves the rice powder in the dry state (i.e. <14% total moisture). Hammer milling pulls particles into a milling chamber through vacuum suction. Inside this chamber, high-speed hammers fracture the rice particles. These particles are continuously hammered until they reach the desired size, being able to pass through the pores of the exit sieve and then collected. This milling technique is effective at reducing rice to particles 100um wide, but readily clogs finer exit sieves. Jet milling uses air pressure to feed the powder into the milling chamber similar to the hammer mill, but instead of using hammers to induce particle dissociation, it relies on rapid directional changes of air-flow. Air is injected tangentially to the wall of the milling chamber. This jet of air causes the particles in the chamber to rapidly circulate around the chamber. The air is drawn at a selected radial degree from the center of the chamber to an exit port. The particles in the chamber experience a drag force proportional to their cross-sectional area. If the particles do not contain sufficient momentum to continue on the circular path around the chamber, they will be drawn out through the exit port and collected. Jet-mills do not impact the powder with machinery; rather they depend on high-energy particle-particle collisions to induce milling and particle size reduction. As a result, jet-milling can readily mill rice to small particles, 3-5 micron in diameters **(Jeong et al., 2008)**.

Air-Classifiers function in a similar manner to the jet-mill, where the particles exhibit a drag force along with a small moment of inertia. Particles with a large drag force and a small moment of inertia (e.g. protein bodies) are selectively drawn out in the fines stream, separating them from the larger components (starch granules). A second stream of air is added in air-classifiers to remove the larger, heavier particles. This prevents particle particle collision as would occur in the jet mill. Three factors control the particle separation: particle density, shape, and size ('aspect'). The particle separation is improved as the difference in particle diameter between fine particles and coarse particles is increased. Experimenters commonly use laser diffraction particle size distribution, scanning electron microscopy, and combustion protein assays (e.g., ELEMENTAR instrument) to characterize the physical properties of cereal starches **(Pande et al., 2017)**. Our starches were analyzed following milling and air-classification by Malvern Mastersizer 2000 particle size distribution.

1.4 Aqueous purification of recombinant proteins

Aqueous purification of recombinant proteins is the capital and technical bottleneck of production. Not only is aqueous purification expensive, there are functional limitations such as contamination, product degradation via proteases, and large amounts of hazardous waste byproducts. These shortfalls present themselves due to the aqueous nature of the process. Given the unique deposition site of the antimicrobial PIN protein, and the nature of the dry rice kernel, we set out in search of an anhydrous purification process, one completely void of water. After an extensive literature search, no anhydrous purification methods were found. Since the rice seed is dry, and that is the deposition

environment to be tested for recombinant proteins, we reasoned that the purification scheme should remain dry, avoiding the aqueous purification steps.

The search for an anhydrous strategy led us into the field of gas chromatography and gas-phase protein sequenators **(Edman 1967)**. These technologies rely heavily on gas-phase chemical cleavage of amino acid residues in proteins and peptides for further downstream analysis. Selective chemical cleavage is a useful way to identify proteins through the observation of subsequent cleavage patterns. In 1968, there was a report of selective bond cleavages for peptides that contained serine, threonine, and glycine residues when exposed to hydrochloric acid at room temperature **(Kaneko et al., 1968)**. The cleavages at the N-terminal of the serine and threonine followed a mechanism involving a N→O shift of hydroxyl groups. The first selective cleavage at an aspartic residues peptide bond was first observed in 1950 upon heating and incubation of a protein in a weak acid solution **(Partridge and Davis 1950)**. This resulted in cleavage at aspartic and asparagine residues. In 1962, selective cleavage was observed following methionine residues using a bovine pancreatic ribonuclease protein and cyanogen bromide as the cleavage reagent **(Gross and Witkop, 1962)**. Under acidic conditions, cyanogen bromide only interacts with the methylmercapto group of methionine. These studies led us to the hypothesis that specific gas-phase cleavable linker sequences could be employed to release starch granule tethered recombinant proteins, using an aqueous-free, gas phase cleavage.

In 1993, a very facile and specific cleavable peptide bond was observed in the gas phase **(Wen et al., 1993)**. This bond was the Asp-Pro peptide bond (D-P), and it is much

more unstable than any other peptide bonds formed between other amino acids. This peptide bond is facile due to the presence of a labile proton found on the aspartic acid residue side chain, coupled with the basicity of the upstream proline. The esterification of the aspartyl group inhibits this mechanism of cleavage, demonstrating that the labile proton found in the side chain of the aspartic acid is necessary. Proximity of this proton to the peptide backbone is also important for cleavage, as observed through the absence of Glu-Pro peptide cleavages **(Wen et al., 1993)**. An important property of the Asp-Pro bond is that cleavage can occur under conditions where all other peptide bonds are stable. Furthermore, the Asp-Pro pairing is amongst the rarest of all amino acid pairs found in nature. These unique features of the Asp-Pro bond make it an ideal selective, gas-phase cleavable linker. The proposed method aims to overcome a number of disadvantages found in conventional protein purification strategies i.e., those that rely on aqueous buffers/reagents and chromatography. A critical advantage of PIN and dry purification is that it prevents product loss that usually occurs from proteolysis and/or contamination. Most known proteases are of aqueous cytoplasmic origin, and thus are only active in an aqueous environment, and inactive in low moisture environments, such as rice flour. The anhydrous strategy can eliminate the requirement for conventional bioprocessing protease inhibitors and hundreds of thousands of liters of bio-waste, produced from the large-scale aqueous purification technologies.

1.5 Rationale, Hypothesis and Objectives

Plant exosome-like nanoparticles (ELPs) are just starting to receive attention as important mediators in plant stress and defense responses. We aimed to provide further research in this new field through proteomic analysis on healthy mature tomatoes.

Proteomic studies completed to date have all been performed on ELPs isolated from plants still in their growth phase. To our knowledge, this is the first proteomic analysis performed on plant ELPs from a mature plant specimen. Given tomato's position as a global fruit consumed broadly, and its world production of 177 million tons, I set out to probe its 95% water content of the fruit. Are ELPs only present during growth and development of the plant? If they are found in mature specimens, do they still contain defensive and stress based proteins? Could the gene expression controlling ELPs be engineered to improve yield, transport, processing and nutritional impact?

We hypothesize that

1- Despite having been harvested, packaged and shipped, the tomato will contain ELPs that contain defensive and stress response proteins.

My objectives were to:

1-Isolate tomato ELPs using differential centrifugation

2-Perform proteomic analysis to elucidate their protein cargo.

Plant biotechnology approaches involving defense proteins are usually directed towards the generation of transgenic lines that confer some sort of resistance to pests or environmental conditions. In this study, we took a novel approach on plant defense proteins, and used the wheat antimicrobial puroindoline protein as a tool to express and purify recombinant proteins. The need for more affordable and widespread use of recombinant proteins is evident. We aimed to express recombinant proteins in rice endosperm using PIN as a fusion partner. PIN elicits its antimicrobial effects from the surface of starch granules, a targeting and localization property we aim to exploit. Rice is

an ideal host to test our fusion technology as its endosperm consists primarily of starch granules and it is a PIN-null crop, allowing our PIN fusions to tether freely to its starch granule surfaces.

I hypothesized that

1- PIN will interact and tether to rice starch granules *in vitro*, and that this interaction will occur *in vivo* when expressed as a recombinant fusion protein accompanied by our protein of interest insulin-like growth factor-1 (IGF-1).

My objectives were to

1-Investigate the tethering ability of PIN to rice starch granules *in vitro*

2-Identify the subcellular location of PIN in transgenic rice seeds

3-Express PIN-IGF-1 fusions in rice endosperm and identify their subcellular localization.

An anhydrous method would solve the limitations of current aqueous recombinant protein purification methods by first pinning or tethering the recombinant protein onto the surface of a cellular particle such as starch granule, a particle that is then isolated in a dry state through standard milling and air-classification techniques, followed by gas-phase cleavage of the fusion protein, employing an anhydrous purification method.

Our hypothesis is that

1-The Asp-Pro bond can be used as a selective gas-phase cleavable linker sequence, allowing liberation of starch granule associated proteins anhydrously.

My objectives were to:

1-Assess labileness by verifying the ability of the gas-phase to cleave Asp-Pro bonds in model proteins

2-Isolate starch granules using milling and air-classification and scan the starch granule surface associated proteome for Asp-Pro containing proteins.

CHAPTER 2. MATERIALS AND METHODS

2.1.1 Isolation, purification and identification of exosome-like nanoparticles from tomato

Tomatoes were washed three times with distilled water. After washing, tomatoes were cut in half, gloved-hand squeezed and filtered through a 0.45um Mylar filter. Half the juice was set aside for proteomic analysis. The remaining juice was differentially centrifuged. An Eppendorf bench top centrifuge 5417R was used at 500g for 10 min, 2,000g for 20 min, 5,000g for 30 min and 20,000g for 60 min. The resulting supernatant was collected and centrifuged at 100,000g for 120 min in a Beckman Coulter ultracentrifuge optima XPN-100 **(Regente et al., 2017)**. The final ELP pellet was re-suspended in PBS and retained for proteomic analysis. To determine the particle size distribution, the ELP fraction was analyzed by nanoparticle tracking analysis (Nanosite LM10).

2.1.2 Preparation of tomato juice and exosome-like nanoparticles for mass spectrometry analysis

A DC protein assay (Bio-Rad) was performed to determine the protein concentration as per manufacturer's instructions. Protein (25ug) was digested by in solution trypsin digest. Briefly, proteins were solubilized in 200uL 8M urea and 50mM ammonium bicarbonate (pH 8.0). Insoluble material was spun down at 5000g for 10 min and discarded. A reduction buffer (1M dithiothreitol in 200uL) was added to a final concentration of 10mM and incubated for 60 min at room temperature. An alkylation buffer (4uL of 1M iodoacetemide in 200uL) was added to a final concentration of 20mM and incubated for 40 min at room temperature in the dark. Samples were diluted with 5 volumes of 50mM ammonium bicarbonate. Trypsin was added (1ug/20ug sample) and

incubated overnight at room temperature with gentle agitation. The resulting peptides were desalted using 20 uL C18 desalting tips as per the manufacturer's protocol and resuspended in 0.5% (vol/vol) formic acid. Samples were run in triplicate.

2.1.3 Mass spectrometry analysis and Bioinformatic analysis of tomato exosome-like particles

The mass spectrometry analysis and database search were performed by the Ottawa Institute of Systems Biology. Peptides resuspended in 0.5% formic acid were analyzed by high resolution mass spectrometry on a Q Exactive (Thermo Scientific) using a 2 hour gradient of increasing (5-25%) acetonitrile concentration (in 0.1% formic acid) and a 250nL/min flow rate. Using a data-dependent acquisition mode the top 12 most intense ions were selected for MS/MS analysis. Peptide identification and quantification were determined using the MS-acquired data with MaxQuant version 1.5.3.30 **(Cox et al., 2009)** using the *Solanum lycopersicum* (downloaded from Uniprot on 2018/03/11) database. The number of maximum missed cleavages was fixed at two, oxidation on methionine and protein N-terminal acetylation were set as variable modifications. The Label-free quantification was enabled and carbamidomethyl on cysteine was selected as a fixed modification. Normalization was performed in MaxQuant using a technology termed MaxLFQ, which relies on a set of algorithms for label-free quantification (**Jurgen et al., 2014**). The normalization is based upon peptide ion intensity and does not rely on housekeeping proteins.

Column correlation (Pearson correlation), two-sample test (Student's T-test) with Benjamini-Hochberg FDR adjustment were performed in Perseus 1.5.1.6 in addition to the heat map of differentially expressed proteins between the juice and exosome-like nanoparticles. Relative protein expression was plotted in GraphPad Prism 7.

2.2.1 Plant material

Seeds of soft wheat *Triticum aestivum* L. cv. Augusta were obtained from Dr Fregeau-Reid, Agriculture and Agri-Food Canada. Rice seeds, *Oryza sativa* L. cv. M202, as supplied by Karen Moldenhauer, Arkansas Rice Research and Extension Center. Durum wheat *Triticum turgidum* L. cv. Langdon was obtained from Sylvie Cloutier, Eastern Cereal and Oilseed Research Center, Agriculture and Agri-Food Canada, Ottawa. Transgenic rice, *Oryza sativa* L. cv. M202 containing wheat PINs was obtained from Drs. Giroux and Beecher (Montana State University). The rice was transformed with PIN-a and PIN-b genes, and is referred to as line 97-1 **(Krishnamurthy and Giroux, 2001)**.

2.2.2 Seed protein extraction and SDS-PAGE

Seeds were husked and ground in liquid nitrogen using a mortar and pestle. Non-reducing Laemmli buffer was added to the dry seed in a ratio of 1.5g seed: 5mL buffer. The sample was boiled at 100 °C and centrifuged at 2000 x g (Beckman Model J-6B) for 10 minutes, after which the pellet was discarded. To reduce samples, dithiothreitol (DTT) was added at a concentration of 5 to 100 mM in Laemmli buffer. One milligram of whole kernel equivalents was loaded per lane. SDS-PAGE gels were prepared per standard protocol **(Wall et al., 2010)**. Protein extracts were loaded in 75mm thick, 5% acrylamide stacking and 15% acrylamide resolving gels. Seven microliters of benchmark ladder was added. Electrophoresis was performed at 20 mA for 2 hours. Gels were stained with coomassie blue for two hours followed by an overnight destaining treatment. Gel doc system (MultiImage Light Cabinet, Alpha Inno-tech Corporation) was used for imaging.

2.2.3 Immunoblotting

For the detection of PINs, the monoclonal antibody Durotest®, an antibody raised in mice, was obtained from R-Biopharm Rhone Ltd. A polyclonal anti-mouse fluorescein isothiocyanate (FITC) conjugated secondary antibody raised in donkey was obtained from Jackson ImmunoResearch Laboratories Inc. The polyclonal anti-mouse alkaline-phosphatase (AP) conjugated secondary antibody raised in goat was obtained from Sigma-Aldrich. Following electrophoresis, the proteins were transferred to a nitrocellulose transblot membrane using a semi-dry set up. Gels were electroblotted onto nitrocellulose membranes for 1.5 hours at 11V. The membrane was rinsed with Ponceau S stain to confirm protein transfer. Once the membrane was destained through water washing, it was incubated in phosphate buffered saline Tween 20 solution (PBST) for 5 minutes. The membrane was blocked overnight using a 5% skim milk solution. Following three, 10 minute water washes, primary antibody Durotest® was added at a 1:20,000 dilution in PBST and incubated for 1 hour at 4 °C. The membrane was washed three times with PBST and incubated with a 1:5000 dilution of the secondary anti-mouse conjugated AP antibody. The membrane was washed three additional times with PBST. Developing reagents (5-bromo,4-chloro,3-indolylphosphate (BCIP) and nitroblue tetrazolium (Invitrogen, Burlington, Ontario) were mixed in 5 mL AP buffer and incubated with the membrane. Colorimetric reactions took approximately 10 minutes to develop at room temperature and were stopped with the addition of water. Membranes were scanned with a Hewlett Packard PSC7000 and imaged with HP PSC Director software.

2.2.4 Cereal Seed sectioning and fixation

Fixing of seeds was performed in 2.5% formaldehyde, 0.5% glutaraldehyde, and 100 mM NaH_2PO_4 (pH 7.4) for two days. Seeds were transversely cut and transferred into 60% ethanol solution for 2 hours. After two 1-hour washes with toluene, the seeds were infiltrated with paraffin. Seeds were embedded in paraffin blocks by the tissue-embedding center (Leica EG 1160, Richmond Hill, Ontario). Paraffin blocks were cross-sectioned into 4-micrometer thick portions using a microtome. The resulting sections were mounted on Fischer Superfrost Plus slides, dried and toluene deparaffinized, ethanol gradient rehydrated, and water washed.

2.2.5 Wheat lysate incubation

Mounted seed sections were circled with a toluene pen to create a catchment vessel or pool basin. To determine whether PINs adhere to rice starch granules, soft and hard wheat seed protein extract were incubated on the rice seed cross-sections for one hour. The slides were thoroughly washed three times in PBS for ten minutes each prior to antibody addition.

2.2.6 Primary and secondary immunostaining of wheat lysates

After incubation with wheat seed protein extracts, slides were rinsed three times for 10 minutes in PBS. Twenty microliters of primary antibody (1:5000 dilution) was added to the hydrophobic pool basin and incubated at room temperature for 1 hour. After incubation of the primary antibody, and three 10-minute rinses with PBS, the secondary antibody was added at a 1:100 dilution and incubated at room temperature for 1 hour. Slides were rinsed three times with PBS and dried. After drying, a drop of Prolong Gold antifade reagent (Invitrogen) was added, and 0.2 mm coverslips were placed over the sections and incubated overnight at room temperature in the dark. Fluorescent and phase

contrast images were visualized using a Zeiss Axiophot microscope (Oberkochen, Germany) equipped with an Olympus DP70 CCD camera. Imaging software, MetaMorph (MetaMorph Imaging Systems, Sunnyvale, California), was used to capture images.

2.2.7 Puroindoline extraction from transgenic 97-1 rice lines

PINs were isolated using the Triton X-114 phase partitioning protocol previously described **(Gautier et al., 2000)**. Briefly, whole kernels were crushed using a mortar and pestle in liquid nitrogen and added to 1% (vol/vol) Triton X-114 in Tris-buffered saline solution (10 mM Tris/150 mM NaCl, pH7.5) at 4 °C for 30 minutes. The solution was transferred to 25 °C for an additional 30 minutes. After incubation, the lower detergent phase was collected and subjected to an additional round of phase partitioning. Following the second phase portioning step, proteins were precipitated with 80% (vol/vol) acetone. The pellet was washed with ether and dried at room temperature. Non-reducing SDS sample buffer was added to the protein extract and loaded to a final concentration of 1 milligram whole kernel equivalents/lane. SDS-PAGE and western blotting was performed as described in sections 2.2.2 and 2.2.3.

2.2.8 Design and synthesis of a PIN-IGF-1 sequence for expression in rice endosperm tissue

To ensure successful expression of our transgene in rice endosperm, codon optimization of the IGF-1 sequence was performed. The IGF-1 sequence used was obtained from the Genbank Database accession number NM_001111285. From this sequence, the mature peptide coding sequence was codon optimized for expression in *Oryza sativa* according to the codon usage database Kazusa (www.kazusa.or.jp/codon/). The optimized IGF-1

sequence was assembled using the oligo primer walking method **(Young et al., 2004)** and cloned into a pUC19 vector using the KpnI and BamHI restriction sites. To achieve expression of the recombinant protein of interest in transgenic endosperm tissue of the novel host plant, the rice GlutelinC promoter was used to drive expression. This promoter was accompanied by the rice Globulin 5' and 3' untranslated regions (UTR) and a Glutelin C terminator **(Qu et al., 2008)**. The Glutelin C promoter, Globulin 5' UTR, PIN, 3' UTR and Glutelin C terminator were synthesized and cloned into a pUC19 vector by Blue Heron Biotechnology. The codon optimized IGF-1 was cloned into the expression vector upstream from PIN and downstream from the 3' globulin UTR using KpnI and BamHI restriction sites **(Figure 2.1)**. Successful cloning was verified by DNA sequencing (Illumina). The construct was cloned into the plant expression vector pCAMBIA1305.1.


GlutC promoter
Glob 5'UTR
Pin-IGF-1
Glob 3'UTR
GlutC terminator

Figure 2.1. Rice Expression construct

DNA construct design for expression in *Oryza sativa*, cloned into a pCambia1305.2 vector (Not shown to scale). The GlutelinC and Globulin elements are from rice endosperm specific proteins to ensure endosperm specific expression of the recombinant Pin-IGF-1 protein. GlutC represents rice GlutelinC (EU264107.1), Glob UTR represents rice globulin untranslated regions (D50643.1), Pin represents puroindoline-a (CAB89541.1) and human IGF-1 (A29117.1).

2.2.9 Transformation of *Agrobacterium-tumefaciens* with pCambia1305.1 PIN-IGF-1 construct

Agrobacterium-tumefaciens strain LBA4404 was transformed with the pCambia1305.1 PIN-IGF-1 vector according to the freeze thaw method **(Hofgen et al., 1988)**. Briefly, a 500μL aliquot of *A. tumefaciens* was incubated with 0.5-1.0μg of plasmid DNA. The cells and vector were incubated successively 5 min on ice and 5 min at 37. After dilution in 1mL YEP media the mixture was shaken at room temperature for 2-4 hours. Aliquots of 50μL were plated out on YEP plates containing the antibiotic kanamycin (50μg/mL) and hygromycin (50μg/mL) and incubated for 2 days at 28 °C. Single colonies were picked and subjected to DNA isolation and PCR analysis by agarose gel electrophoresis **(Figure 2.2)**.

2.2.10 Agrobacterium-mediated transformation of rice callus tissue

Transformation of rice calli took place as previously described **(Hiei and Komari, 2008)**. Mature rice seeds cv. Nipponbare are dehusked and sterilized in 70% ethanol and 2% sodium hypochlorite with a drop of Tween-20 for 30 min with stirring. The seeds are rinsed and placed on 2N6 plates and incubated at 30 °C in the dark for seven days. During this time the seeds produce calli tissue that are harvested and propagated further on 2N6 plates for ten days. After this calli induction step, they are inoculated with the transformed *A. tumefaciens* strain LBA4404 OD=0.1 at 660nm for 2 min at 25 °C. Following inoculation the calli are plated on 2N6-Acetylsyringone and incubated in the dark at 25 °C for three days. After this co-cultivation step, the calli underwent two rounds of antibiotic selection and the surviving calli were used for plant regeneration.

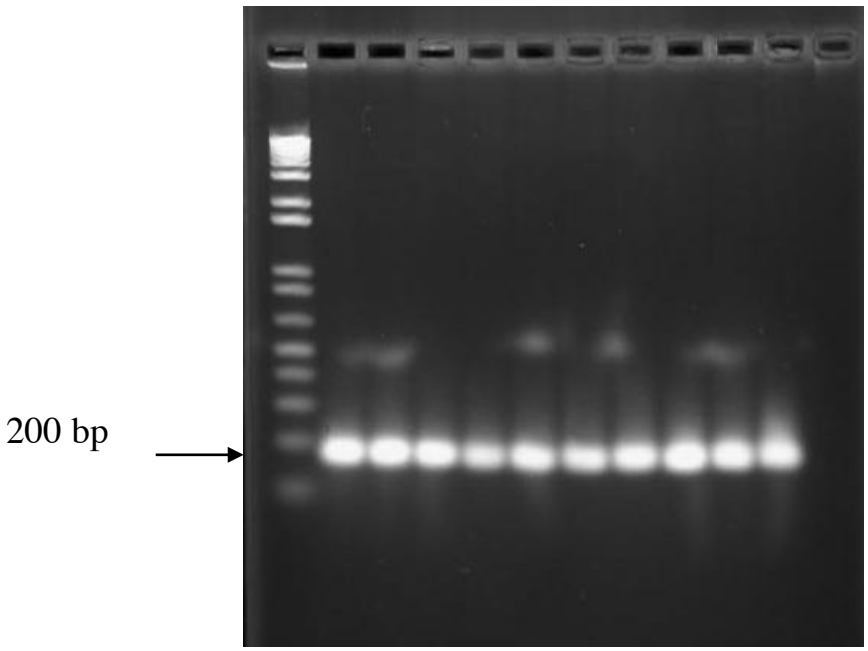

200 bp

Figure 2.2. Genomic PCR of *Agrobacterium-tumefaciens*

Genomic PCR of *Agrobacterium-tumefaciens* DNA to confirm insertion of the transgene construct from Figure 8, using primers for the N- and C-terminal portions of the human IGF-1 protein (210bp).

2.2.11 Rice callus regeneration

The calli remaining after the two selection stages were plated on N6RH50 Petri plates containing 1-Naphthaleneacetic acid and 6-Benzyladenine and incubated for 28 days under continuous illumination at 32 °C. This regeneration step produces plantlets that were transferred to rooting induction plates for an additional 14 days.

2.2.12 Cultivation of transgenic rice plants

Transgenic plants were transferred to 17 cm of potting soil and grown in a greenhouse for 95 days at 18-24 °C (Centre for Advanced Research in Environmental Genomics, Biology Dept., University of Ottawa). After this period, leaf tissue and seeds were collected for further analysis.

2.2.13 Rice genomic DNA isolation and transgene analysis

Leafy tissue was collected two months after planting. Using a mortar and pestle, the leafy tissue was ground in liquid nitrogen and subjected to a DNeasy mini kit from Qiagen. PCR analysis was performed on the extracted genomic DNA to confirm transgene integration.

2.2.14 PIN-IGF-1 immunolocalization

Transgenic rice seeds were cross-sectioned as described in section 2.2.5. Immunolocalization was performed as described in section 2.2.7 using an anti-IGF-1 primary antibody.

2.2.15 PIN-IGF-1 SDS-PAGE and immunoblotting

SDS-PAGE and western blot of rice PIN-IGF-1 extracts were performed as described in sections 2.2.3 and 2.2.4 using a primary anti-IGF-1 primary antibody.

2.3.1 Gas-phase reagents

The model proteins catalase and tyrosinase were obtained from Sigma-Aldrich (St. Louis, Mo., cat. nos. 120E-7160 and 29C-9640, respectively). The cleavage reagent, 7-Heptafluorobutyric acid (7-HFBA) was also purchased from Sigma. Rice seeds, cultivar Nipponbare (Arkansas Rice Research Center, Stuttgart, AR, USA) were grown in a greenhouse and collected. Milling was performed using the Perten Hammer and Sturtevant Jet mill and air-classification was carried out using a Matsubo elbow jet air-classifier (AAAmachines, Arlington, IL).

2.3.2 Gas-phase cleavage of model proteins

Catalase (2 mg) and tyrosinase (2 mg) were weighed and placed into separate microcentrifuge tubes. The gas-cleavage apparatus developed in this study was a novel modification of the gas-phase amino acid sequenators, comprising of a filter packed in the bottom of a glass reaction vessel. This reaction vessel was vacuum-sealed during the cleavage reaction. Cleavage tests were performed under vacuum at 60°C for 16 hours using a concentration of 0.01% 7-HFBA. Heptafluorobutyric acid (0.01%) exposed samples from various incubation periods were placed in 96 well Terasaki Plate (Alpha Biotech Ltd London, UK), in duplicates. BSA standards were also prepared with concentrations ranging from 0.2-2 mg/ml using the Albumin Standard (Pierce, Ill., USA, #23209). Absorbance were read at 595nm with the PowerWave Microplate Spectrophotometer (BioTek Instruments Inc. Winooski, VT). Model proteins, tyrosinase (2 mg) and catalase (2 mg) were exposed to the optimal cleavage conditions using the novel, gas-phase cleavage apparatus to observe anhydrous peptide cleavage. Each sample was incubated with 7-HFBA (0.01%) in the cleavage apparatus for 16 hours at 60°C to

permit cleavage of peptides, specifically at any labile Asp-Pro bonds in the model proteins. Following incubation, samples were transferred from the filter paper directly into a microcentrifuge tube (1.5 ml). The remaining 7-HFBA was removed from the samples by vacuum, and the proteins were solubilized in 1 ml of Tris buffer and kept for further analysis.

2.3.3 Isolation of rice starch granules and gas-phase cleavage of surface associated proteins

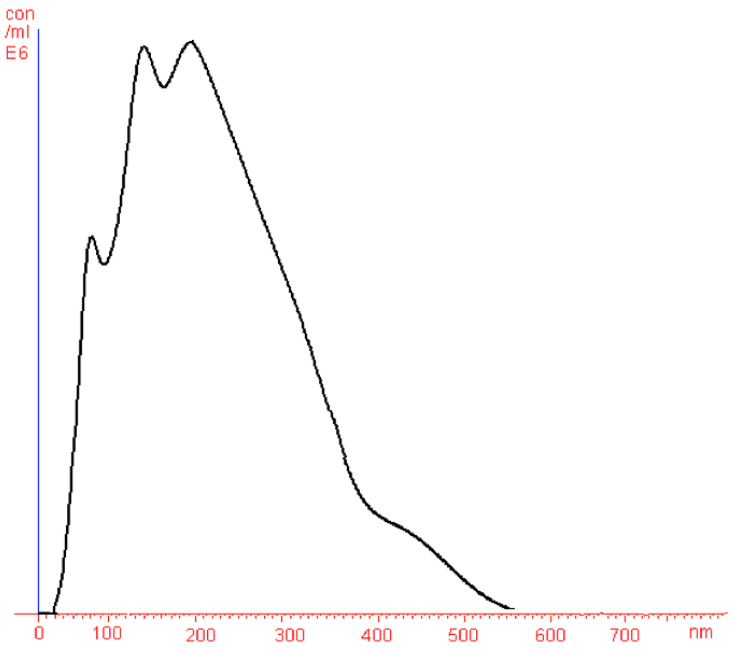
Rice cultivar Nipponbare was grown in a greenhouse for 90 days under 18 hour light periods at a temperature of 28°C (CAREG, Biology Dept., uOttawa). Following maturation, rice seeds were harvested and husked. Seeds were sent to Sturtevant (Hanover, MA) for Jet milling. The resulting flour was air-classified at AAAmachines (Chicago, IL). Particle size distribution was performed on the resulting starch granule isolate. This starch granule fraction was incubated under vacuum with 0.1% heptafluorobutyric acid for 16 hours at 60°C to elicit Asp-Pro cleavage of the starch granule associated proteins.

CHAPTER 3. Proteomic analysis of exosome-like nanoparticles from *Solanum lycopersicum* (tomato)

3.1 Results

3.1.1 Identification of exosome-like nanoparticles in mature tomato

Exosomes have been routinely isolated and analyzed from mammalian cell cultures and biological fluids while plant ELPs are just starting to garner attention. The primary isolation method for exosomes is differential centrifugation **(Regente et al., 2017)**, while isolation using columns is less common. We employed a differential centrifugation method **(Regente et al., 2009)** for isolation of tomato ELPs. Briefly, tomatoes were gently pressed and vacuum filtered through a 0.22um diameter Mylar filter. It is important to note that the tomatoes were not ground or blended, which would cause release of intercellular materials, which could cause product contamination. The obtained juice was differentially centrifuged at 2,000g, 5,000g, 10,000g and 20,000g then ultracentrifuged at 100,000g. The ELP pellet was collected and resuspended in PBS. To determine the particle size distribution, samples were subjected to nanoparticle tracking analysis (NTA). The sample contained spherical structures with an average diameter of 125-150nm, consistent with previous reports **(Figure 3.1) (Regente et al., 2009).** The protein yield obtained in each biological replicate as determined by protein assay (DC Bio-Rad) for the juice was 2.75, 2.45 and 2.00ug/uL respectively. The protein yields obtained from the ELPs were 1.11, 0.76 and 1.14ug/uL respectively **(Figure 3.2)**. These quantities were similar to those found in sunflower and Arabidopsis ELPs.



Particle Size / Concentration

Figure 3.1. Particle size distribution of tomato ELPs

Particle size distribution of tomato ELPs following differential centrifugation as determined by nanoparticle tracking analysis (Nanosite LM10). X-axis units in nm to indicate ELP particle diameter. Y-axis units in con/ml.

	Protein concentration ug/uL	Volume for 25ug protein
Tomato juice 1	2.752231	9.083541
Tomato juice 2	2.450734	10.20103
Tomato juice 3	2.005667	12.46468
EV tomato 1	1.112456	22.47279
EV tomato 2	0.756608	33.04222
EV tomato 3	1.148349	21.77039

Figure 3.2. Protein concentration of tomato juice and tomato ELPs

The purpose of this experiment was to investigate whether fresh pressed exudate from the fruit differs from pasteurized protein concentration of tomato juice and tomato ELPs as determined by Detergent Compatible assay.

3.1.2 Bioinformatic analysis of tomato exosome-like nanoparticles

Proteomic analysis of the tomato juice detected 1242, 1365 and 1793 proteins from the Uniprot tomato database. Analysis of the ELPs identified 155, 141 and 179 proteins **(Figure 3.3)**. The number of proteins identified is consistent with other plant ELP proteome studies, characterized from sunflower and Arabidopsis **(Prado et al., 2014; Rutter and Innes, 2017).** Biological replicates yielded high Pearson correlation values for both the juice (0.969 to 0.978) and ELP samples (0.91 to 0.945) **(Figures 3.4 and 3.5).** There were 84 proteins common to all ELP and juice biological replicates of which 54 were deemed to have significantly different expression levels between the ELPs and juice (FDR 0.05) **(Figure 3.6)**. Of these 54 proteins, 30 were upregulated in the ELPs of which 15 remain uncharacterized. Out of those 30 ELP specific proteins, a majority of them are related to plant defense and stress response.

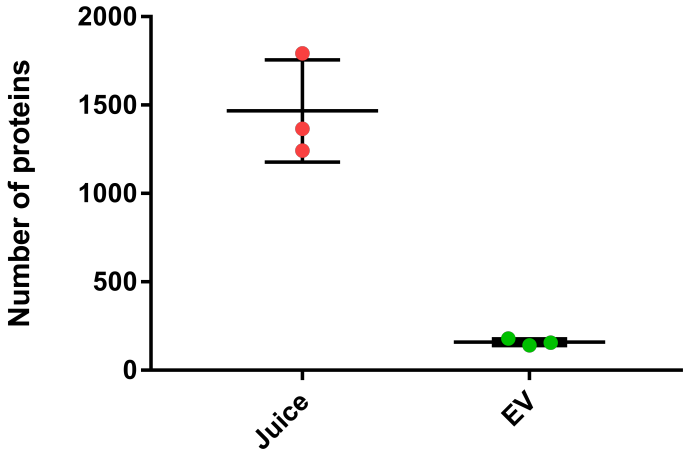

Number of proteins
2000
1500
1000
500
0
Juice
EV

Figure 3.3. Total number of proteins found in tomato juice and tomato ELPs

Total number of proteins identified by MS from the pressed juice and the differentially centrifuged ELP isolated samples respectively. Image was generated using GraphPad prism 7.

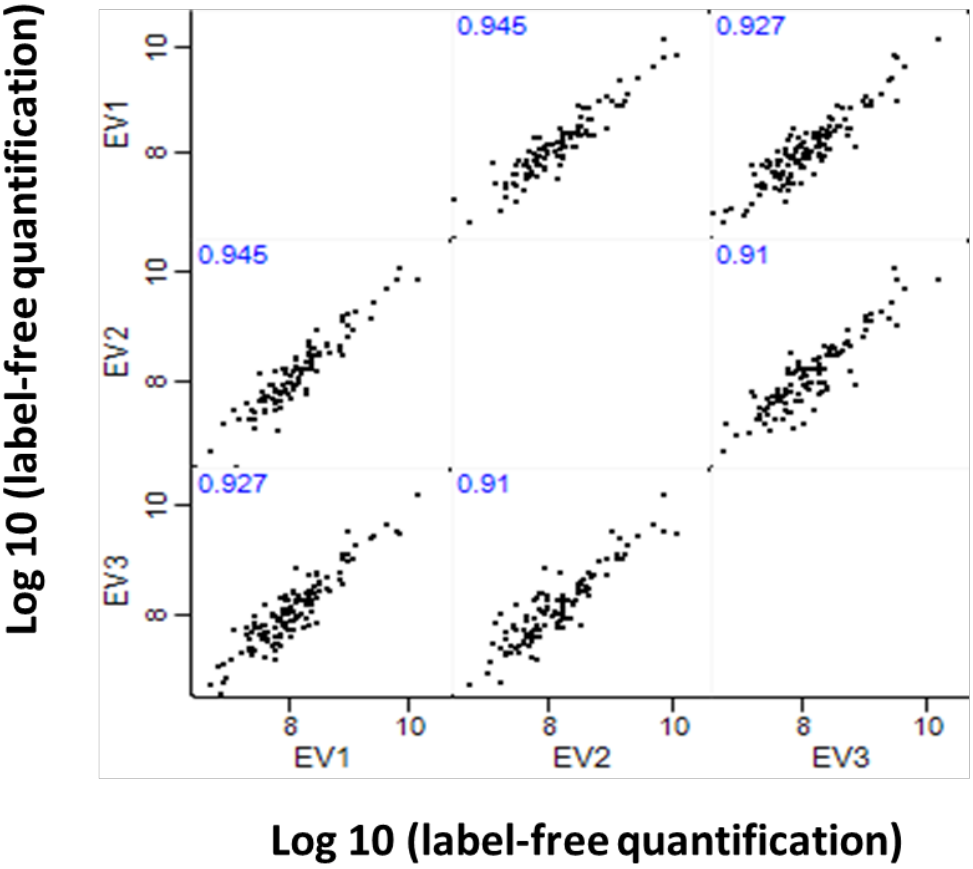

Log 10 (label-free quantification)
0.945
0.927
0.945
0.91
0.927
0.91
EV1
EV2
EV3
8
10
Log 10 (label-free quantification)

Figure 3.4. Pearson correlation values of tomato ELP samples

Pearson correlation values (numeric value indicated in blue) created in Perseus 1.5.1.6 for the three ELP samples from tomato, EV1, EV2, and EV3. Pearson correlation indicates the linear correlation between two variables.

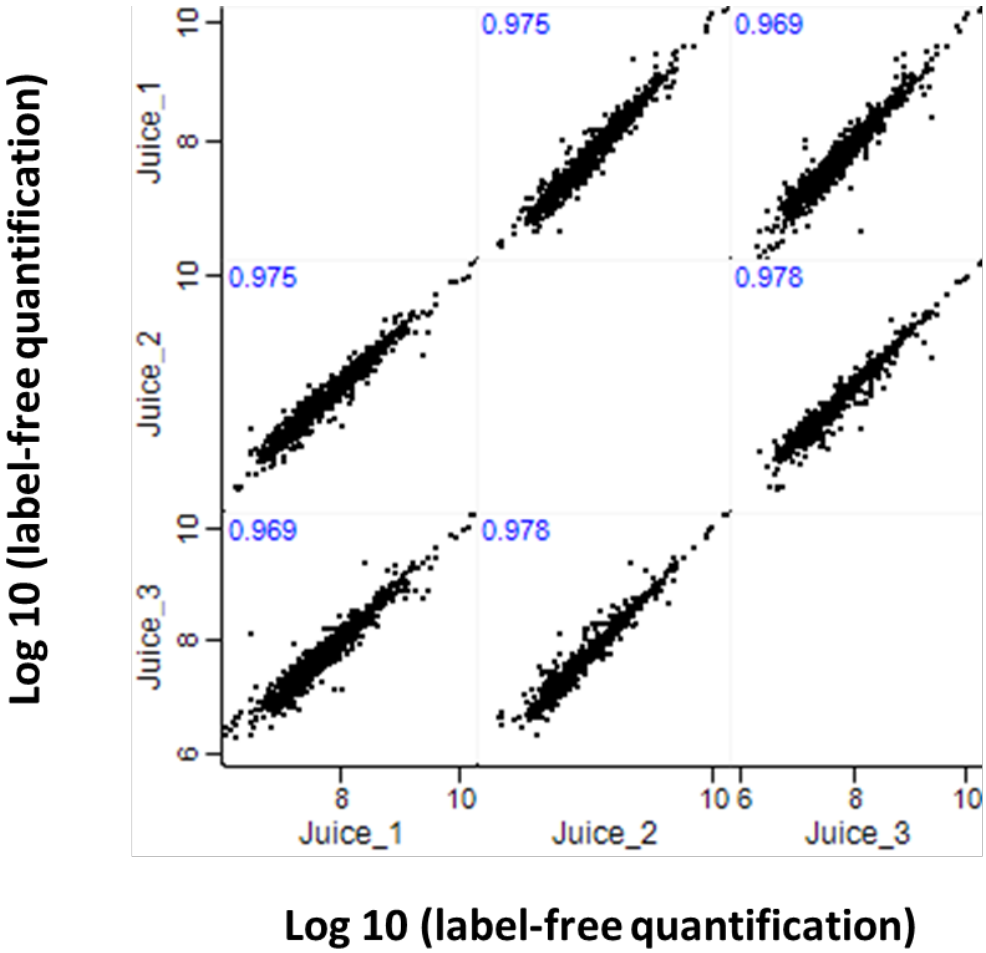

Log 10 (label-free quantification)
0.975
0.969
0.975
0.978
0.969
0.978
Juice_1
Juice_2
Juice_3
10
8
6
Log 10 (label-free quantification)

Figure 3.5. Pearson correlation values of tomato juice samples

Pearson correlation values (numeric value indicated in blue) created in Perseus 1.5.1.6 for the three tomato juice samples. Pearson correlation indicates the linear correlation between two variables.

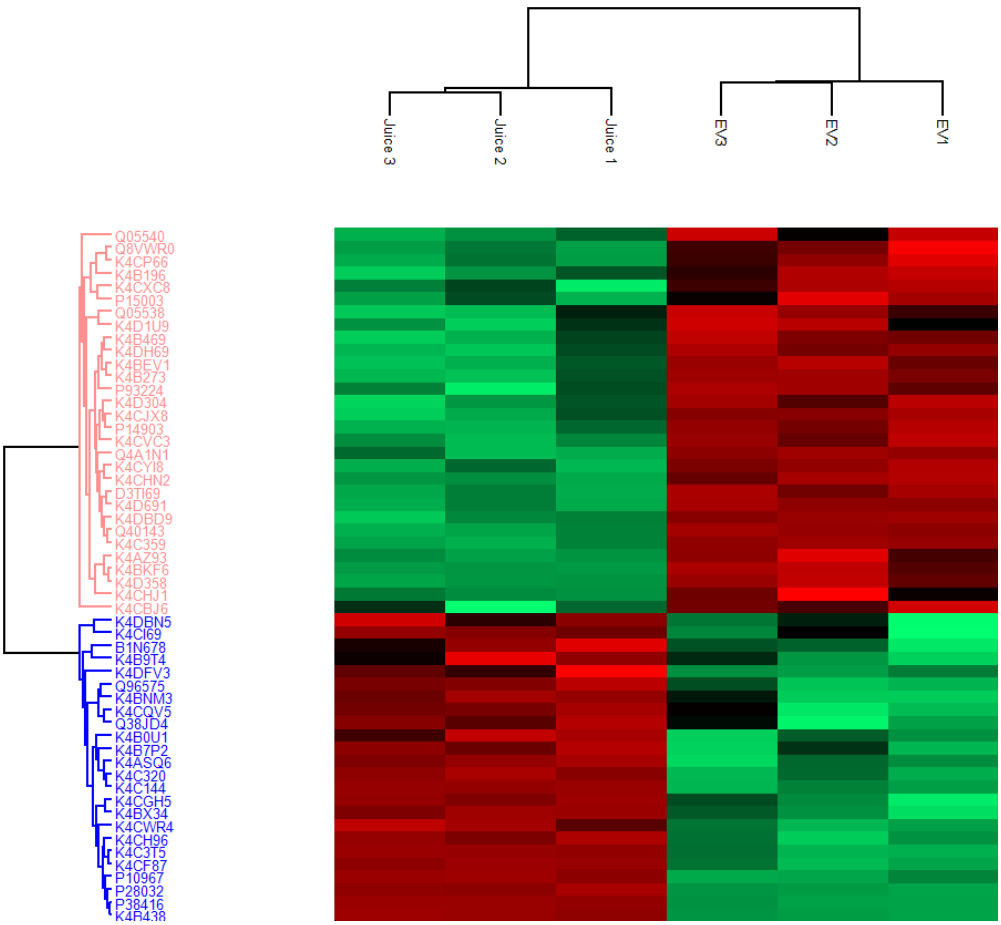

Juice 3
Juice 2
Juice 1
EV3
EV2
EV1
Q05540
Q8VWR0
K4CP66
K4B196
K4CXC8
P15003
Q05538
K4D1U9
K4B469
K4DH69
K4BEV1
K4B273
P93224
K4D304
K4CJX8
P14903
K4CVC3
Q4A1N1
K4CYI8
K4CHN2
D3TI69
K4D691
K4DBD9
Q40143
K4C359
K4AZ93
K4BKF6
K4D358
K4CHJ1
K4CBJ6
K4DBN5
K4CI69
B1N678
K4B9T4
K4DFV3
Q96575
K4BNM3
K4CQV5
Q38JD4
K4B0U1
K4B7P2
K4ASQ6
K4C320
K4C144
K4CGH5
K4BX34
K4CWR4
K4CH96
K4C3T5
K4CF87
P10967
P28032
P38416
K4B438

Figure 3.6. Heat map showing the proteins that are expressed at significantly different levels between the juice and ELPs

Left panel shows the Uniprot accession numbers, right panel shows the relative expression levels following Z score normalization. Green represents an increase in expression; red represents a decrease in expression. Heat map was plotted in Perseus 1.5.1.6.

We chose to focus on the proteins that were either significantly upregulated in the ELPs or proteins found solely in the ELP samples. It has been demonstrated in recent studies that ELPs isolated from leafy tissue are packaged with defense related proteins, and that they do indeed have antifungal properties **(Regente et al., 2017).** Our studies further support the hypothesis that plant ELPs play a role in plant defense. In contrast with ELPs isolated from leafy tissue of living plants, our ELPs were isolated from a mature tomato. Despite undergoing harvesting, packaging and transport, the tomato cytoplasm and apoplast fluids contained ELPs that lay dormant, waiting to spring into action in the event of stresses or pathogen attacks. Of the significantly upregulated proteins found in the ELPs, a considerable proportion are related to plant defense and stress **(Table 3.1).** Furthermore, the defense related proteins are significantly upregulated in the ELPs when compared to the whole tomato juice. There were defense specific proteins found solely in the ELPs. Two notable defense proteins, pathogenesis related protein P2 and *Phytophthora* inhibited protease (water molds) were specific to the ELP fraction.

Table 3.1. . **List of defense proteins found in tomato ELPs:** comparison of stress and defense related proteins identified in tomato fruit exosome-like nano-particle (ELP) and juice samples

Protein	ELP (Log10) LFQ group average	JUICE (Log10) LFQ group average
Suberization-associated anionic peroxidase	9.875	9.154
Profiling	8.585	8.208
Non-specific lipid transfer protein	8.448	7.605
Basic 30 kDa endochitinase	7.922	7.344
Acidic 30 kDa endochitinase	8.198	7.425
Pectin esterase A	8.654	8.075
Cysteine protease 3	9.028	8.472
Polygalacturonase	9.028	N/A
Pathogenesis related protein P2	7.677	N/A
Phytophthora inhibited protease 1 (water molds)	7.359	N/A
Low temperature induced cysteine proteinase	7.634	N/A
Chitinase activity	7.649	N/A

*N/A indicates there were no hits in those samples for the corresponding proteins

The ELPs contained cell wall modifying enzymes. These results are in accordance with previous exosome proteomic studies of plant samples **(Rutter and Innes, 2017)**. Plant defense is heavily based on cell wall modifications, further supporting the hypothesized properties of plant ELPs **(Park et al., 2017)**. Upon pathogen attack, the plant's primary response is in the form of encasements **(Hansen et al., 2017)**. Plant pathogens' main form of attack is through haustorium structures **(Hansen et al., 2017)**. Haustoria are slender cone like projections that can penetrate plant tissue and allow for the absorption of the host's nutrients. Cell wall modifying proteins are used by plants to restrict or encase haustorium structures **(Both et al., 2005)**. Furthermore, proteins that promote the deposition of cutin, a waxy substance, can be used by the plant to restrict pathogen attack efficiency **(Savatin et al., 2014)**. The major constituent of plant parasitic fungi is in the form of chitin **(Sharma et al., 2011)**. When I analyzed these plant ELPs, I found a significant increase in endochitinases **(Figure 3.7)**. These results further support the hypothesis that plant ELPs have an important role in plant mediated defense.

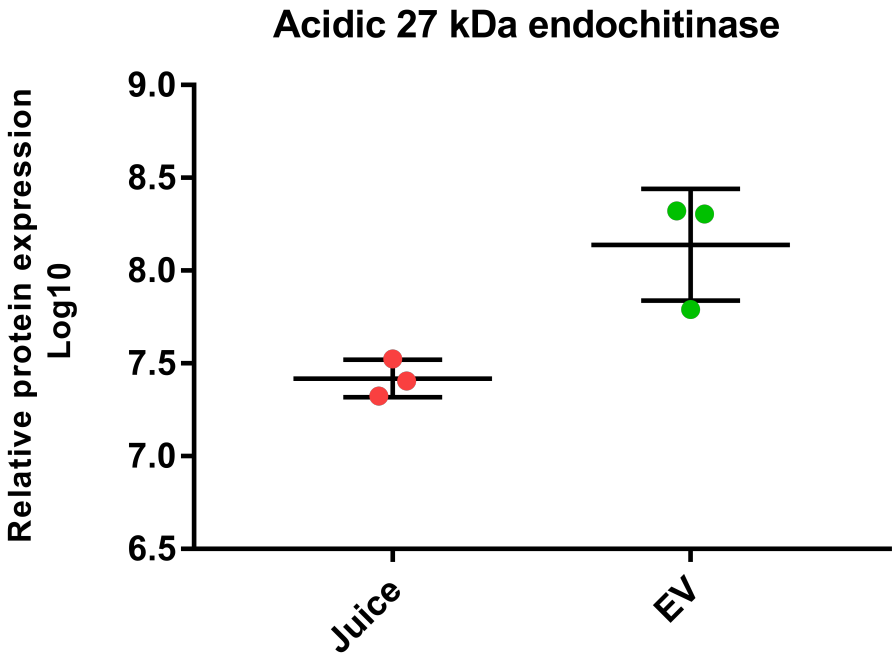

Acidic 27 kDa endochitinase
Relative protein expression
Log10
9.0
8.5
8.0
7.5
7.0
6.5
Juice
EV

Figure 3.7. Relative protein expression levels of acidic endochitinase

Exosome protein content differs from the hand-pressed juice exudate. Comparison of the relative protein expression of the defense protein acidic 27 kDa endochitinase (Gene ID: 101266570) between the tomato juice and ELP samples. Image was generated using GraphPad prism 7.

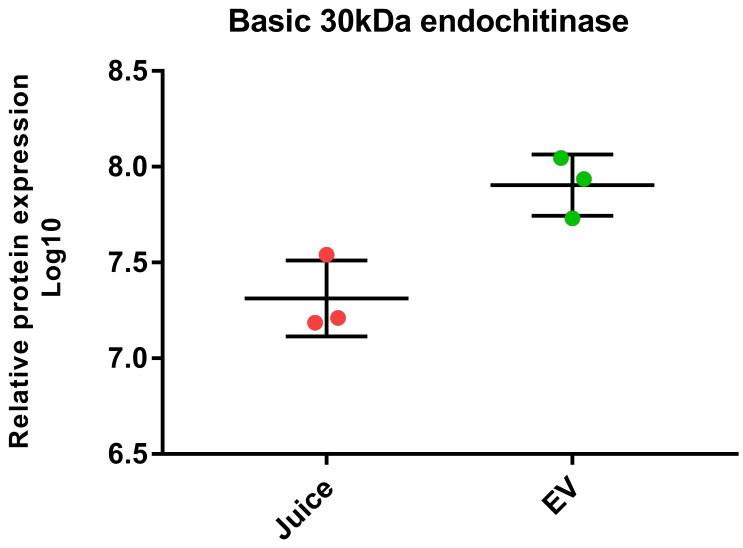

Basic 30kDa endochitinase
Relative protein expression
Log10
8.5
8.0
7.5
7.0
6.5
Juice
EV

Figure 3.8. Relative protein expression levels of basic endochitinase.

Exosome protein content differs from the hand-pressed juice exudate. Comparison of the relative protein expression of the defense protein acidic 27 kDa endochitinase (Gene ID: 544146) between the tomato juice and ELP samples. Image was generated using GraphPad prism 7.

CHAPTER 4. Puroindoline as a fusion partner for recombinant proteins expressed in rice

4.1 Results

4.1.2 In situ localization of endogenous PIN proteins

The natural starch granule tethering ability of PINs was verified using immunofluorescent staining of PINs in PIN rich soft wheat, PIN-null hard wheat and PIN-null rice seed cross-sections. A strong signal was observed on and surrounding the starch granules in the soft wheat cross-section **(Figure 4.1 A, B)**. A weak localization was noticed in the surrounding protein matrix, with a majority of the fluorescence occurring on the starch granule surfaces. Phase contrast images are included to provide orientation. Hard wheat cross-sections showed no fluorescent signals, confirming the absence of PINs **(Figure 4.1 C, D)**. To confirm that the Durotest antibody does not interact with rice seed cross-sections in the absence of PIN, immunostaining on the parental rice 'M202' was performed. No fluorescence was observed, demonstrating the specificity of the Durotest antibody for PIN **(Figure 4.1 E, F)**.

4.1.3 In vitro localization of exogenous proteins

To determine whether PIN will bind rice starch granules in a similar manner as it interacts with wheat starch granules, rice seed cross-sections were treated with soft wheat seed homogenate.

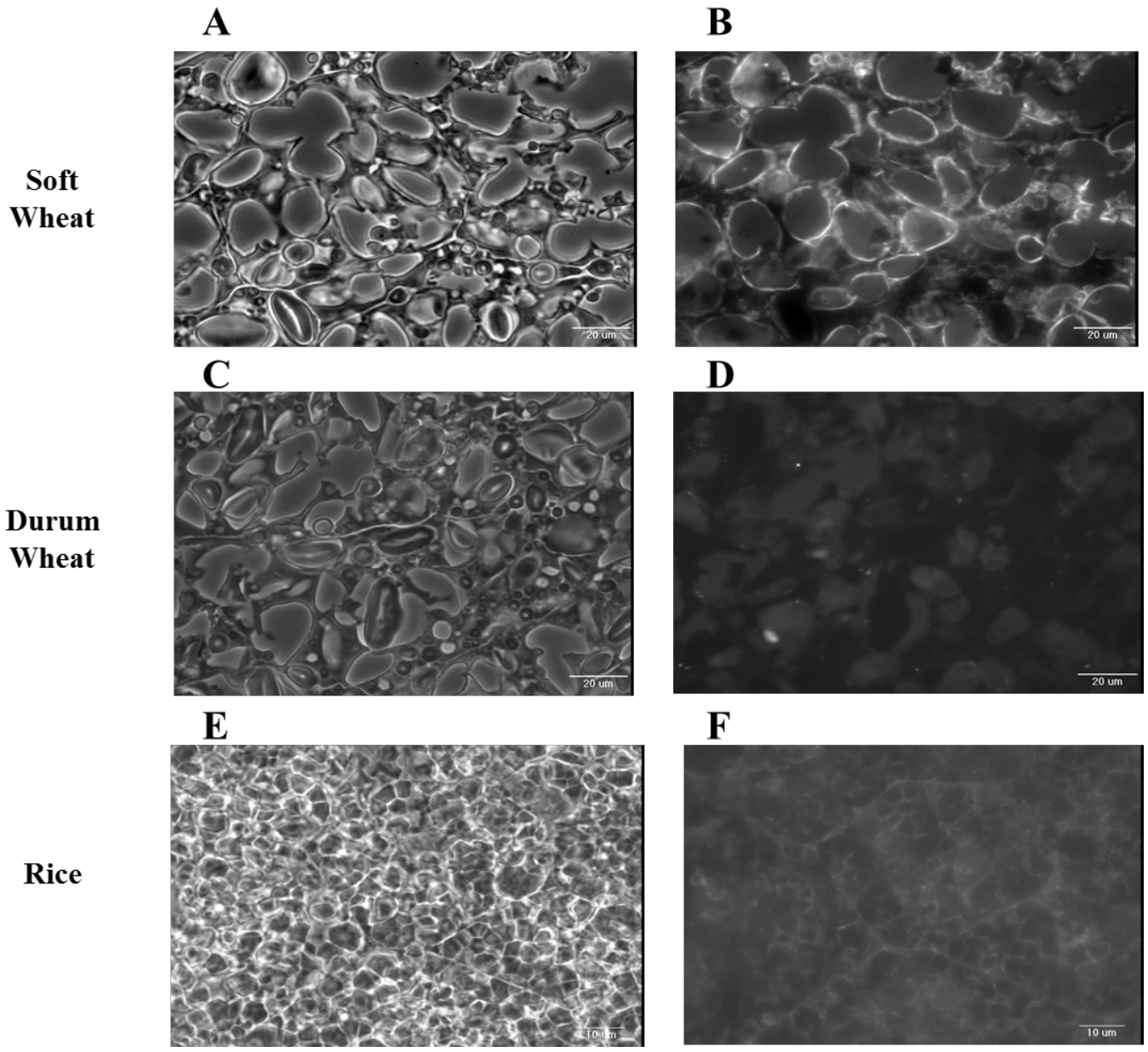

A
B
Soft Wheat
20 um
20 um
C
D
Durum Wheat
20 um
20 um
E
F
Rice
10 um
10 um

Figure 4.1. Immunolocalization of PIN on starch granule surfaces in wheat

Immunolocalization of puroindolines showed localization on the starch granule surfaces in the seed endosperm of hexaploid wheat *T. aestivum*. Phase contrast (left) and immunofluorescent microscopy images (right), on soft wheat *T. aestivum* L. cv. Augusta (A,B), durum wheat *T. turgidum* L. cv. Langdon (pin-null) (C,D), and wild type rice *O. sativa* L. cv. M202 (pin-null) (E,F).
These micrographs were produced by an undergraduate student in Dr Altosaar's lab named Charles Melnyk.

Despite the fact that rice contains a cryptic PIN-like sequence (Chantret et al., 2005), no transcripts encoding PINs have been observed in rice **(Morris, 2002; Tanchak et al., 1998)**. Rice has a hard kernel texture and lacks the genes homologous to PIN and grain softness protein-1. The presence of PIN-like genes in the *Triticeae* and *Aveneae* tribes and the lack of PINs in rice is consistent with the phylogeny of the *Gramineae*. Given the known interactions between PIN and polar lipids **(Biswas and Marion 2006)**, we hypothesized that native PINs *ex situ* from soft wheat seed extracts, when incubated with rice starch granules, would transfer from the suspension of wheat granules to the *in situ* PIN-null rice starch granules. Following immunostaining of the rice seed-cross sections, localization was observed to be specific for rice starch granule surfaces **(Figure 4.2 C)**. The experiment was repeated using PIN-null durum wheat seed homogenate as a negative control, and to ensure no cross-reactivity of the antibodies **(Figure 4.2 D)**. It was important to note that no wheat starch granules remained fixed to the cross-sections as they can be distinguished by their larger shape and ellipsoidal lenticular shape.

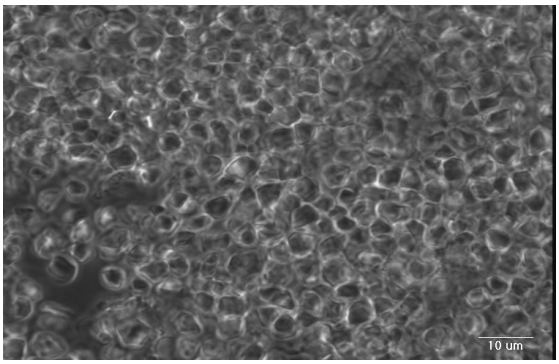
4.1.4 *In vivo* localization of exogenous PIN in transgenic 97-1 rice

To determine whether PIN interacts with rice starch granules *in planta* in addition to the rice seed cross-sections, transgenic rice transformed with the wheat PIN genes were analyzed. The Durotest antibody showed a strong fluorescent localization surrounding the starch granules in 97-1 rice **(Figure 4.3).** To confirm the secondary antibody was not interacting in a non-specific manner, rice cross-sections were stained with the secondary antibody alone. No fluorescence was observed.

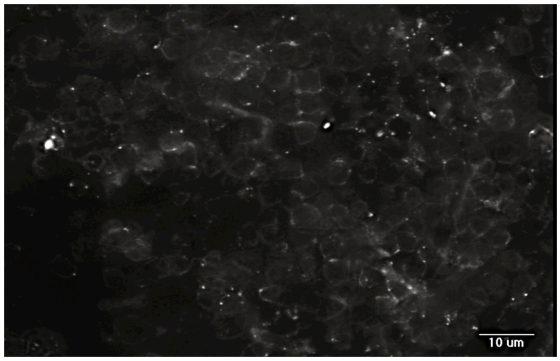
Phase Contrast

Durotest+

A

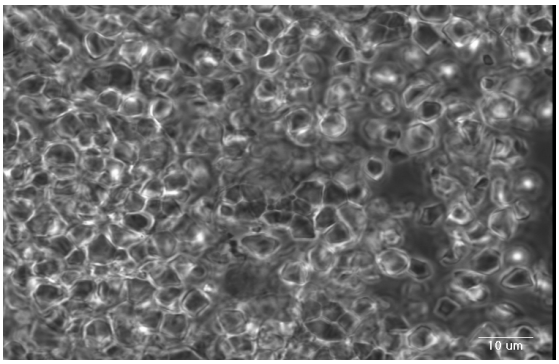


B

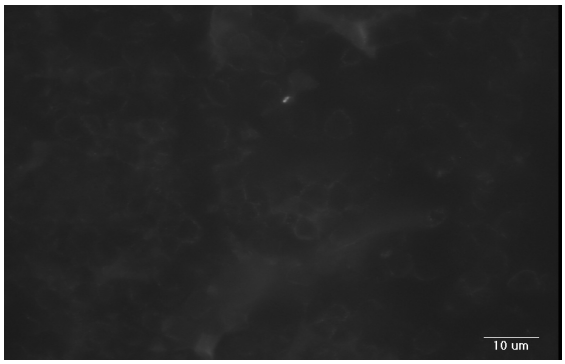


Rice + Soft Wheat Extract

C



D



Rice + Durum Wheat Extract

Figure 4.2. 'Roll-over' immunolocalization of PIN on rice starch granule surfaces

'Roll-over' of puroindolines from preparations of starch granules from soft wheat specifically localize onto the starch granule surfaces of rice. Phase contrast (left) and immunofluorescent microscopy (right), of rice, *O. sativa* L. cv. M202 cross-sections incubated with soft wheat, *T. aestivum* L. cv. Augusta (A,B) and durum wheat, *T. turgidum* L. cv. Langdon (C,D), seed protein extracts.

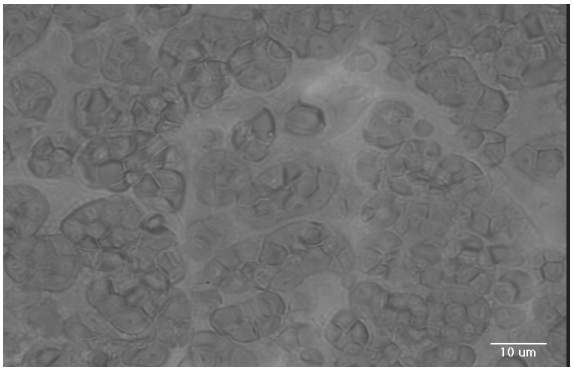
Phase Contrast

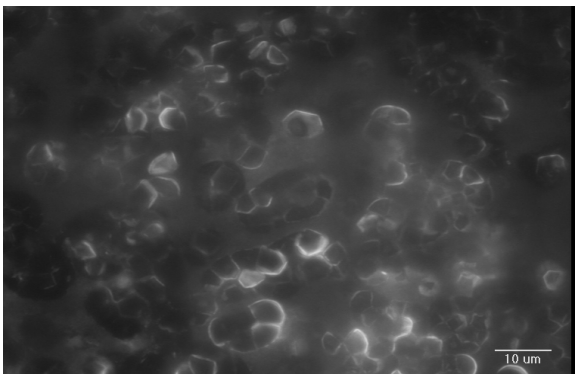
Durotest+

A

B

PIN+ rice

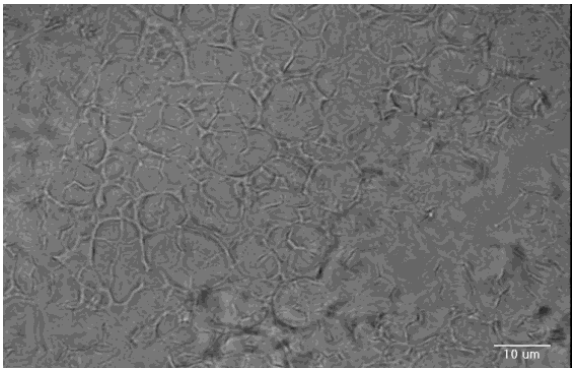




C

D

Rice



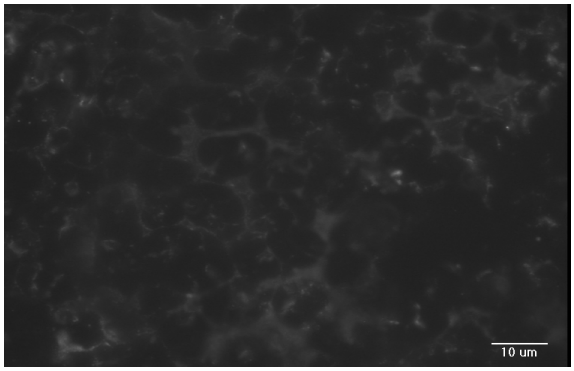


Figure 4.3. Immunolocalization of PIN in transgenic rice seed

Puroindolines, when expressed in a pin-null rice endosperm compartment, localize and deposits themselves onto the starch granule surfaces. Phase contrast (left) and immunofluorescent microscopy images (right), of transgenic PIN^{+} rice (97-1) parental *O. sativa* L. cv. M202 (A,B) and wild type rice *O. sativa* L. cv. M202 (C,D) seed cross-sections.

These micrographs were taken by an undergraduate student in Dr Altosaar's lab named Charles Melnyk,

4.1.5 Synthesis of codon optimized IGF-1 for rice expression

To optimize the expression of IGF-1 in transgenic rice seed, a codon-optimized version was designed. The sequence was codon optimized using high codon frequencies tabulated by kazuza.co.jp for *Oryza sativa*. The codons replaced those from the mature human sequence. The optimized sequence was ligated into to plant expression vector pCambia1305.1 downstream of the puroindoline gene. The cartoon of the construct design can be seen in **(Figure 2.1)**.

4.1.6 Agrobacterium-tumefaciens transformation

The strain LBA4404 was transformed with the pCambia1305.1 PIN-IGF-1 codon optimized construct using the freeze thaw method. To confirm successful transformation, *A. tumefaciens* were grown under the influence of three antibiotic selection markers: Kanamycin, hygromycin and rifampicin. The kanamycin and hygromycin resistant markers are found on the pCambia1305.1 vector, and the rifampicin is found with the LBA4404 strain. Colonies surviving the antibiotics were collected and subjected to a PCR to further verify successful transformation. Primers for the human codon optimized IGF-1 were used, with an expected product of 210bp. The PCR products were fractionated on an agarose gel to determine fragment length. As expected, an observed product of 210bp was observed **(Figure 2.2)**. *A. tumefaciens* colonies were collected for rice calli transformations.

4.1.7 *Agrobacterium*-mediated transformation of rice

Transformation of mature rice calli was carried out using the established transformation protocol **(Hiei and Komari, 2008)**. Similar to other *A. tumefaciens* transformation methods, the mature calli technique has extremely low transformation

efficiencies, with 1 in 150 calli surviving the first antibiotic selection, and only 1 in 400 surviving the second selection. The co-cultivation step seemed to be the bottleneck, as the *A. tumefaciens* were overly virulent causing necrosis of calli tissue. There was frequent bacterial overgrowth on co-cultivated calli when inoculated with a culture having an OD of 0.3-0.5. Selection and regeneration of transgenic lines was accomplished by incubating calli on root and shoot regeneration media containing the antibiotic selection marker hygromycin. Resulting plantlets were grown in a greenhouse at 30°C and 16hr day periods for four months. During plant growth, leafy tissue was isolated and genomic DNA was extracted using the DNeasy plant mini kit from Qiagen. The genomic DNA was subjected to a PCR using primers for the promoter and terminator and subjected to a restriction digest to map and to confirm the integration into the rice genome **(Figure 4.4)**.

4.1.8 Immunolocalization of PIN-IGF-1 in transgenic rice

Following maturation, transgenic rice seeds were cross-sectioned and immunolocalization of PIN-IGF-1 was performed as described in section 2.2.15. Similar to the in vitro roll-over results **(Figure 4.2 B)** the pin-IGF-1 fusions appear to solely localize surrounding the starch granule surfaces **(Fig. 4.5)**. These results demonstrate that PIN retains its starch granule binding ability despite having a C-terminal fusion partner. It also confirms the ability of the rice GlutelinC promoter to drive transgene expression in the endosperm tissue. SDS-PAGE and Western Blot analysis demonstrated the presence of PIN-IGF-1, and IGF-1 following protease cleavage at the proper molecular weights **(Figures 4.6 and 4.7)**.

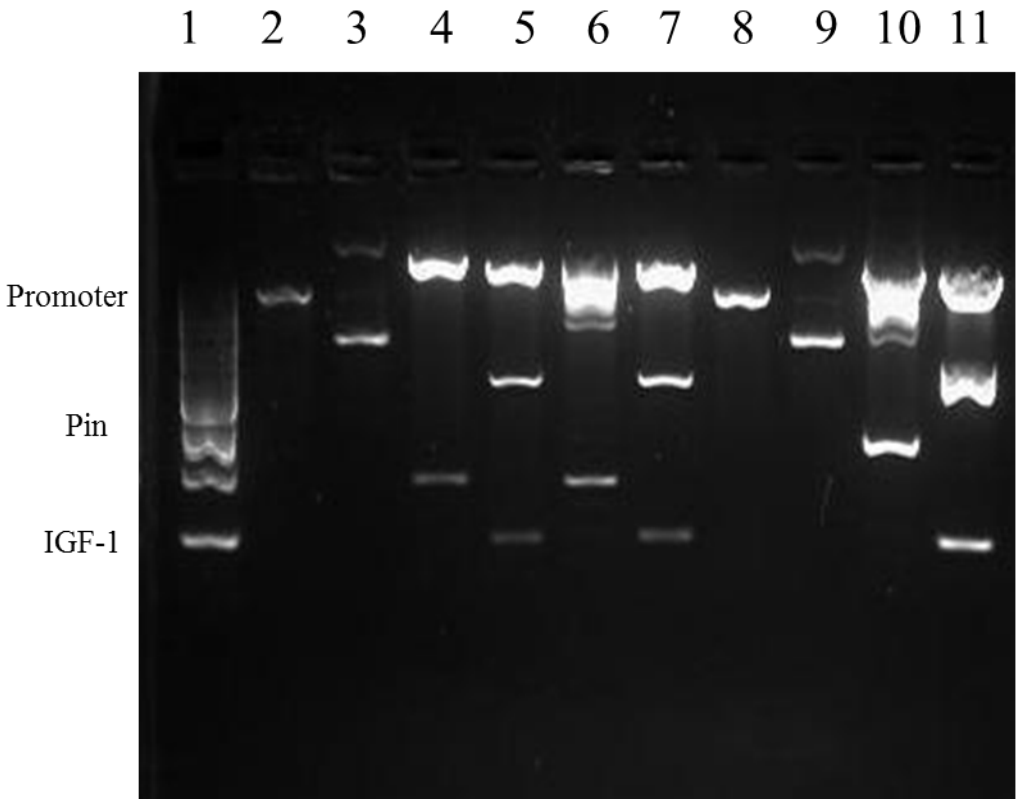
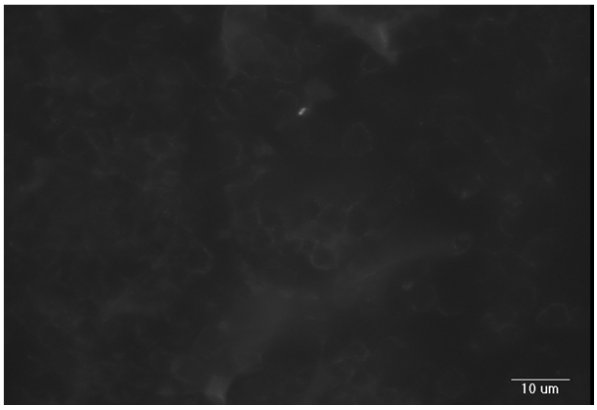

1 2 3 4 5 6 7 8 9 10 11
Promoter
Pin
IGF-1

Figure 4.4. Genomic PCR of transgenic rice

Genomic PCR of transgenic rice DNA, *O. sativa* L. cv. Nipponbare, using primers for the GlutelinC promoter and GlutelinC terminator to confirm *Agrobacterium*-mediated transformation. PCR products were cleaved using restriction enzymes EcoRI, HindIII, BamHI and KpnI to obtain a digest map of the construct.

A



B

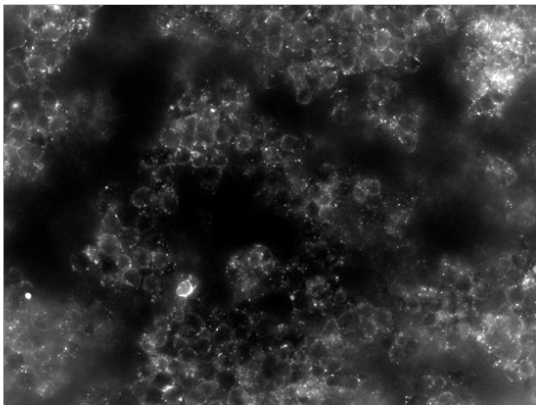


Figure. 4.5. Immunolocalization of PIN fusions in transgenic rice

When expressed in a PIN-null endosperm, puroindoline-IGF-1 fusions localized and deposited themselves on starch granule surfaces. Phase contrast (left) and immunofluorescent (right) microscopy images using IGF-1 primary antibody of transgenic PIN-IGF1 rice, parental *O. sativa* L. cv. Nipponbare seed cross-sections.

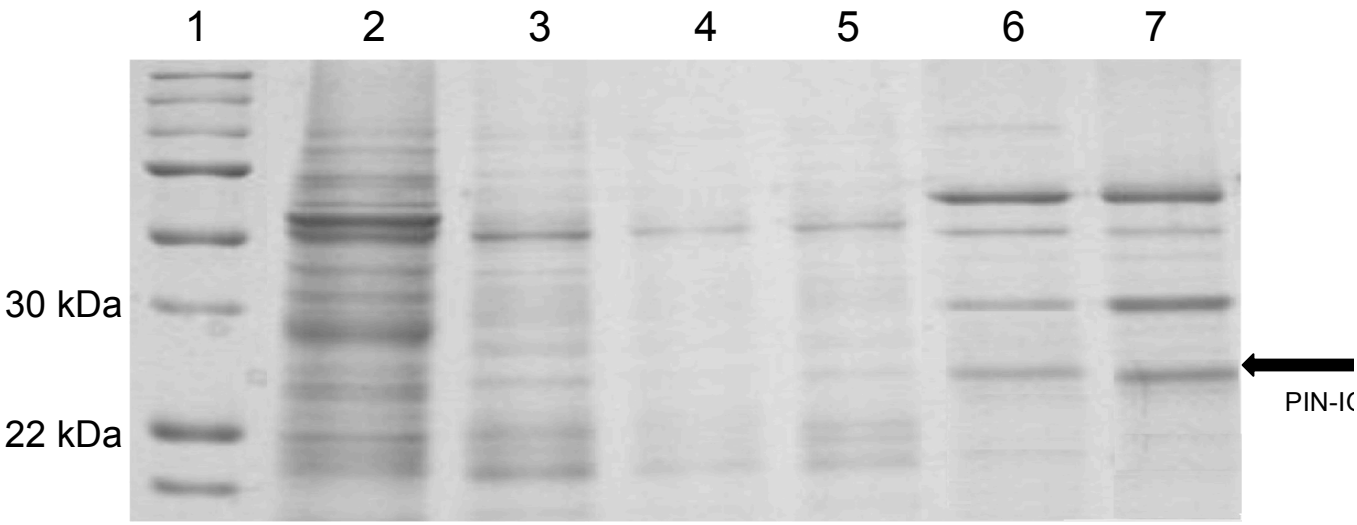

1
2
3
4
5
6
7
30 kDa
22 kDa
PIN-I

Figure 4.6. Detection of PIN-IGF-1 in transgenic rice through SDS-PAGE

Detection of Pin-IGF-1 fusions in transgenic rice seed through SDS-PAGE. Lane 1: Benchmark protein ladder (Bio-Rad). Lane 2: Wild type Nipponbare seed protein extracts. Lanes 3-5 transgenic Nipponbare seed protein extracts. Lanes 7-8: Transgenic starch granule protein extracts.

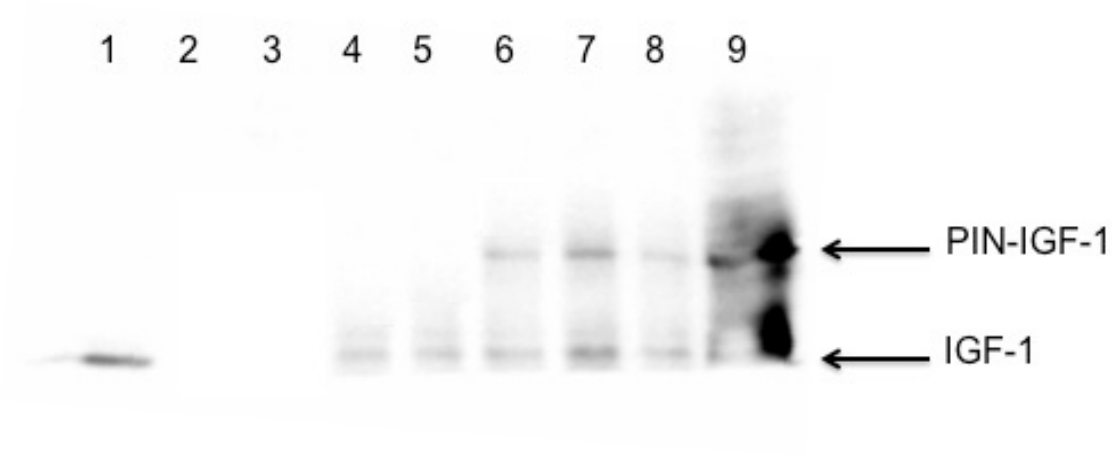

1
2
3
4
5
6
7
8
9
PIN-IGF-1
IGF-1

Figure 4.7. Detection of PIN-IGF-1 and IGF-1 from transgenic rice seeds through western blot

Detection of PIN-IGF-1 and enterokinase cleaved IGF-1 from transgenic rice seed. Lane 1: purified hIGF-1 from *E.coli*. Lane 2: wild type rice *O. sativa* L. cv. Nipponbare. Lane 3: total protein extracts from leafy tissue. Lanes 4-5: purified IGF-1 by enterokinase cleavage of starch granule surfaces. Lane 6-8: purified IGF-1 using enterokinase along with Triton X-100 based purification of PIN-IGF-1 fusions. Lane 9: Total seed protein extraction.

CHAPTER 5. Dry Phission of fusion proteins: Gas-phase cleavage of proteins at aspartic-proline bonds

5.1 Results

5.1.1 Gas-phase cleavage of model proteins

The advantage of rice flour as a high-throughput food-grade platform includes convenient and inexpensive features such as the ease of determining optimal anhydrous peptide cleavage conditions. This is achieved through the adjustment of varying factors such as concentration and volume of gaseous reagents, length and temperature of incubation, and the moisture content and mass of the rice flour or other biological medium. Anhydrous 7-HFBA should cleave Asp-Pro bonds within proteins associated with starch granules isolated from rice. Anhydrous cleavage can be observed readily by exposing model proteins containing the labile Asp-Pro bond to a small amount of gaseous 7-HFBA. Samples of our model proteins were loaded onto a filter, where the cleavage gas passage could flow through. The reaction vessel was connected to a line of tubing enabling a flow path for the 7-HFBA gas vapors. The samples are treated with inert nitrogen gas prior to exposure of gaseous 7-HFBA. The nitrogen gas flow removed loosely associated starch granule associated proteins. Samples were then exposed to gaseous 7-HFBA at the optimal cleavage conditions of 16 hour exposure at 60°C **(Wen et al., 1993)**.

Our first model protein, Tyrosinase, isolated from *Bos taurus* is an 80 kDa protein containing two Asp-Pro cleavage sites. With the presence of two Asp-Pro cleavage sites, three cleavage products are predicted following 7-HFBA exposures. Based on the amino

acid sequence (GenBank: AAL02331.2), the expected peptide sizes are 4.4, 17, and 47 kDa **(Figure 5.1 A)**. Upon 16-hour exposure to 0.1% 7-HFBA using the novel anhydrous cleavage apparatus, cleavage of tyrosinase was observed, and indicated by arrows **(Figure 5.2 A).** The right hand lane in the silver stained SDS-PAGE gel showed the presence of three cleavage products, as three faint bands were observed. The middle lane represents a control sample of tyrosinase, unexposed to gaseous 7-HFBA. It is therefore expected that cleavage products be absent in the center lane. The multiple faint bands of higher Mr in the center and right hand lanes may indicate low purity of the tyrosinase protein sample or different possible isoforms. It is expected that one band at 80 kDa can be identified as tyrosinase using western blotting or mass spectrometry of the gel slice. Left hand lane represents 5 µl of the PageRuler protein Ladder. The migration distance of the cleavage products on the gel corresponds to the predicted peptide sizes of 11, 20 and 49 kDa.

Anhydrous cleavage of Catalase, our second model protein, exemplifies the power of particle tethering of fusion proteins followed by liberation via gas cleavage of the recombinant protein. Catalase was isolated from *Bos taurus* and is a 72 kDa protein containing at least one Asp-Pro cleavage site. The amino acid sequence (NCBI Reference Sequence: NP 001030463.1) shows that three Asp-Pro peptide bonds occur with to being at approximately third intervals in the protein, with an Asp-Pro site at the N-terminal region. Due to the proximity of this last DP site to the terminal, no cleavage product is expected to be observed. Therefore, three cleavage products of roughly similar sizes are predicted from that scissile bond **(Figure 5.1 B)**.

A)
MLLAALYCLLWSFRTSAGHFPRACASSKSLTEKECCPPWAGDGSPCGRLSGRGSCQDVILSTAPLGPQFP
FTGVDDRESWPSIFYNRTCQCFSNFMGFNCGSCKFGFRGPRCTERRLLVRRNIFDLSVPEKNKFLAYLTL
AKHTTSPDYVIPTGTYGQMNHGTTPLFNDVSVYDLFVWMHYYVSRDTLLGDSEVWRDIDFAHEAPGFLPW
HRLFLLLWEQEIQKLTGDENFTIPYWDWRDAENCDVCTDEYMGGRNPANPNLLSPASFFSSWQIVCSRLE
EYNSRQALCNGTSEGPLLRNPGNHDKARTPRLPSSADVEFCLSLTQYESGSMDKAANFSFRNTLEGFA
DPVTGIADASQSSMHNALHIYMNGTMSQVPGSAN**DP**IFLLHHAFVDSIFEQWLRKYHPLQDVYPEANAPIG
HNRESYMVPFIPLYRNGDFFISSKDXGYDYSYLQDSEPDIFQDYIKPYLEQAQRIWPWLIGAAVVGSVLTA
VLGGLTSLLCRRKRNQLPEEKQPLLMEKEDYHNLMYQSHL

MLLAALYCLLWSFRTSAGHFPRACASSKSLTEKECCPPWAGDGSPCGRLSGRGSCQDVILSTAPLGPQFP
FTGVDDRESWPSIFYNRTCQCFSNFMGFNCGSCKFGFRGPRCTERRLLVRRNIFDLSVPEKNKFLAYLTL
AKHTTSPDYVIPTGTYGQMNHGTTPLFNDVSVYDLFVWMHYYVSRDTLLGDSEVWRDIDFAHEAPGFLPW
HRLFLLLWEQEIQKLTGDENFTIPYWDWRDAENCDVCTDEYMGGRNPANPNLLSPASFFSSWQIVCSRLE
EYNSRQALCNGTSEGPLLRNPGNHDKARTPRLPSSADVEFCLSLTQYESGSMDKAANFSFRNTLEGFA
-47 kDa

PVTGIADASQSSMHNALHIYMNGTMSQVPGSAN**D**
— 4.4 kDa

IFLLHHAFVDSIFEQWLRKYHPLQDVYPEANAPIGHNRESYMVPFIPLYRNGDFFISSKDXGYDYSYLQDS
EPDIFQDYIKPYLEQAQRIWPWLIGAAVVGSVLTAVLGGLTSLLCRRKRNQLPEEKQPLLMEKEDYHNLMY
QSHL
- 17 kDa

B)
MNAMTNKTLTTAAGAPVADNNNTMTAGPRGPALLQDVWFLEKLAHFDRERIPERVVHAKGSGAYGTFTVT-
HDISRYTRARIFAEVGKQTPLFLRFSTVAGERGAADAERDVRGFAIKFYTDEGNWDLVGNNTPVFFIR**DP**L
KFPDFIHTQKRKTNLRNATAAWDFWSLNPESLHQVTILMSDRGLPQNYRQQHGFGSHTYSFVNDAGERFYV
KFHFKSQQGIACYTDGEAAELVGRDRESAQRDLFQNIEQGQFPRWTLKVQVMPEAEAATYHINPFDLTKVW
PHAD**DP**YPLIEVGVLELNKNPENYFAEVEQAAFTPANVVPGIGFSPDKMLQGRLFSYGDTHRYRLGINHHQ
IPVNAPRCPFHSFHRDGMGRVDGNGGATLNYEPNSFGEWREAKHAAEPPLALDGQAADRWNHRVDEDYYSQ
PGALFRLMNDDQKQQLFGNIGRH**DP**AYAAGVAKALGLK

MNAMTNKTLTTAAGAPVADNNNTMTAGPRGPALLQDVWFLEKLAHFDRERIPERVVHAKGSGAYGTFTVT-
HDISRYTRARIFAEVGKQTPLFLRFSTVAGERGAADAERDVRGFAIKFYTDEGNWDLVGNNTPVFFIR**D**
-15.6 kDa

PLKFPDFIHTQKRKTNLRNATAAWDFWSLNPESLHQVTILMSDRGLPQNYRQQHGFGSHTYSFVNDAGERF
YVKFHFKSQQGIACYTDGEAAELVGRDRESAQRDLFQNIEQGQFPRWTLKVQVMPEAEAATYHINPFDLTK
VWPHAD**D**
-17.2kDa

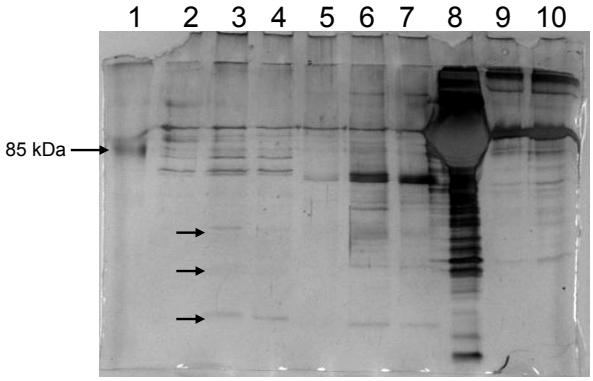
PYPLIEVGVLELNKNPENYFAEVEQAAFTPANVVPGIGFSPDKMLQGRLFSYGDTHRYRLGINHHQIPVNA
PRCPFHSFHRDGMGRVDGNGGATLNYEPNSFGEWREAKHAAEPPLALDGQAADRWNHRVDEDYYSQPGALF
RLMNDDQKQQLFGNIGRH**D**
-17.8

Figure 5.1. Amino acid sequences and expected cleavage products of tyrosinase and catalase

Amino acid sequence of tyrosinase (a) and catalase (b) and their corresponding DP cleavage sites (yellow) and resulting fragments below.

a)

1-Ladder
2-Tyrosinase (4ug)
3-Tyrosinase + gas (16hr) (4ug)
4-Tyrosinase +gas (8hr) (4ug)
5-Tyrosinase (4ug)
6-Tyrosinase + gas (16hr) (8ug)
7-Tyrosinase + gas (8hr) (8ug)
8-hAlbumin (4ug)
9-hAlbumin + gas (16hr) (4ug)
10-hAlbumin + gas (8hr)(4ug)



b)

1-Catalase (4ug)
2-Catalase + gas (4ug)
3-Empty
4-hAlbumin
5-hAlbumin + gas
6-Ladder

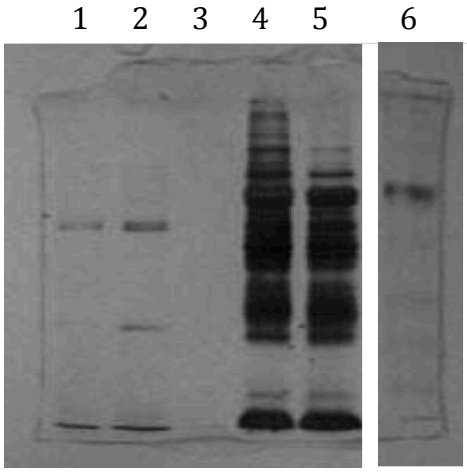


Figure 5.2. Cleavage of model proteins using 7-HFBA

Cleavage of model proteins using anhydrous 7-heptafluorobutyric acid vapors. A represents tyrosinase, here cleavage products are indicated by the arrows. B represents catalase. hAlbumin was used as a negative control, lacking DP sites.

The identification of the faint 35 kDa band, however, suggests that the band corresponding to this cleavage product may represent a protein that has not undergone complete cleavage, thus two cleavage products are being analyzed as one. Similar to the results obtained for tyrosinase, a 16-hour exposure to gaseous heptafluorobutyric acid in our anhydrous cleavage apparatus resulted in cleavage of catalase **(Figure 5.2 B)**. As opposed to tyrosinase, catalase was represented by a single band in the silver stain gel, indicating a higher level of purity. One distinct band was observed as a cleavage product in lane 2. This single band is expected to represent two cleavage products of similar masses, which is expected upon anhydrous cleavage of catalase at the Asp-Pro bonds. Lane 1 represents a control sample of catalase, unexposed to gaseous 7-HFBA. It is therefore expected that cleavage products be absent in lane 1.

5.1.2 Rice starch granule isolation and cleavage of Asp-Pro bonds found in the starch granule associated proteome

Milling of rice seeds is a common technique used to produce flours. In this study we experimented with two different types of milling, hammer and jet. Hammer mills rely on rapidly moving hammer arms within the mill to induce particle dissociation, whereas jet milling relies on particle-particle collisions. Jet milling resulted in a significant reduction of median particle size as compared to hammer milling as determined through scanning electron microscopy **(Figure 5.3)**. To separate the rice starch granules from the endogenous protein bodies, agglomerates and other seed constituents, air-classification using a Matsubo Elbow Jet EJ-L3 air classifier was used (AAAmachines, Arlington, IL) on the jet-milled flour. Using a Malvern Mastersizer, particle size distribution of the rice flour sample was obtained, showing two prominent peaks, one at ~1 μm (protein body

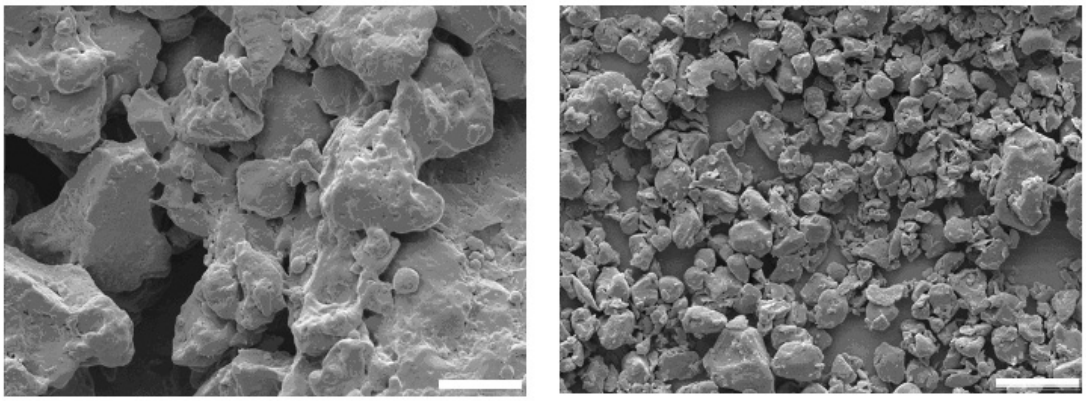

10μm
10μm

Figure 5.3. SEM of rice flour after hammer and jet milling

Scanning electron microscopy of rice flour after hammer milling in Perten Lab Mill 3100 before jet-milling (left), and following jet-milling in a Sturtevant jet mill (right). Scale bar is10μm.
These SEM images were produced by an undergraduate student in Dr Altosaar's lab named Mark Horsman.

fraction) and another around 3-5 μm (starch granule fraction) **(Figure 5.4)**. The 3-5 μm fraction was collected and used for our gas-phase cleavage experiments.

Following the cleavage of model proteins at Asp-Pro sites, we wanted to scan the starch granule associated proteome for proteins containing Asp-Pro bonds. Transgenic studies utilizing PIN and the Asp-Pro bond as a cleavable linker will need to compete with the endogenous rice starch granule associated proteins upon anhydrous classification. A previous student in our lab conducted proteomic experiments on rice starch granule associated proteins **(Koziol et al., 2012)**. Rice starch from Sigma was incubated with trypsin and the cleavage products were subjected to mass spec analysis. From the analysis, they identified rice starch granule associated proteins that contain DP sites **(Table 5.1)**. This provided us with a list of expected cleavage products, allowing us to evaluate the efficacy of anhydrous cleavage through SDS-PAGE gels. Sigma rice starch and our isolated rice starch granules were subjected to anhydrous cleavage as per tyrosinase and catalase referenced earlier in this chapter. Cleavage products were separated on an SDS-PAGE and silver stained **(Figure 5.5).** As expected, the Sigma rice starch contained higher protein content than our air-classified rice starch following anhydrous cleavage. It was also determined through previous MS studies that the Sigma starch samples contained proteins that are not endogenous to rice. Whether these proteins are a result of their starch isolation process, or contaminants introduced through another source is unknown **(Koziol et al., 2012)**. Our air-classified starch showed four prominent bands following anhydrous cleavage. These bands are found at 27, 22, 18 and 16 kDa. Upon comparison with our lab's previous MS analysis of the rice starch granule

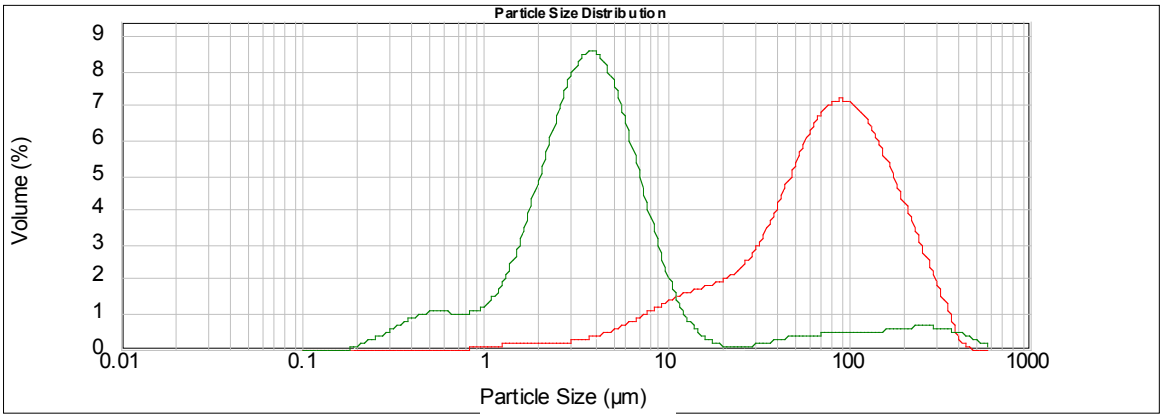

Particle Size Distribution
Volume (%)
9
8
7
6
5
4
3
2
1
0
0.01
0.1
1
10
100
1000
Particle Size (µm)

Figure 5.4. Particle size distribution of rice flour after Matsubo air-classification.

Red line represent particle size distribution of milled rice flour and the green line represents the particle size distribution of milled and air-classified rice flour. There was no separation of particles after milling. Air-classification showed a sharp peak at the starch granule size (4-6μm).

Table 5.1. List of starch granule associated proteins as determined from mass spectrometric analysis of isopropanol extracts of trypsin-treated *Oryza sativa* starch granules

Accession	Protein	#D-P Cleavable bonds	Source	Size of Fragments (kDa)
CAA34756	Elongation factor-1 alpha	1	Human	36 14
BAA22419	Orthophosphate dikinase	4	Rice	17 12 46 19 8
Q42968	Granule-bound starch synthase 1, chloroplastic/amyloplastic	1	Rice	40 27
AAD50279	Branching enzyme	2	Grass	10 16 66
EEC73106	Hypothetical protein OsI_07091:glutelin	2	Rice	20 6 27
EAY75915	Hypothetical protein OsI_03835:glutelin	1	Rice	22 33
CAA32565	Preprolamin	1	Rice	2 15

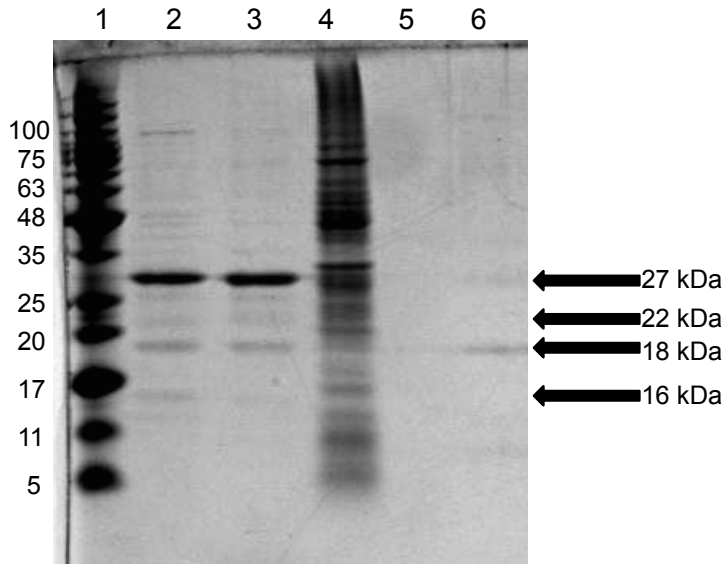

1
2
3
4
5
6
100
75
63
48
35
25
20
17
11
5
27 kDa
22 kDa
18 kDa
16 kDa

Figure 5.5. Cleavage products from rice starch granules obtained through DryPhission.

Lane 1: Benchmark protein ladder. Lane 2: Sigma rice starch cleavage products. Lane 3: Sigma rice starch cleavage products. Lane 4: Total rice seed protein extraction. Lane 5: Cleavage products from endogenous rice starch granule surfaces following milling and air-classification. Lane 6: Cleavage products from endogenous rice starch granule surfaces following milling and air-classification.

proteome **(Table 5.1)**, and our analysis of Asp-Pro sites found in these proteins, we could predict which proteins remained following our isolation process. The 27 kDa band can correspond to a fragment from either granule-bound starch synthase 1 or a rice glutelin protein. The 22 kDa band could correspond to a rice glutelin protein; the 18 kDa band could represent orthophosphate dikinase, branching enzyme or preprolamin. The 16 kDa band could represent preprolamin, branching enzyme or elongation factor-1. These fragments are found in both the Sigma rice starch and the air-classified rice starch samples. These protein identifications are just a prediction, but they can be confirmed through MS analysis of the gel bands.

CHAPTER 6. DISCUSSION

6.1. Tomato ELPs contain defense proteins

Protein content of tomato ELPs reported herein is consistent with those identified from Sunflower and Arabidopsis. Pearson correlation showed > 0.9 similarity between our three ELP samples. This demonstrated the consistency in our ELP isolation method, and MS preparation. Clustering of proteins identified in the ELPs according to their gene ontology (GO) terms was unsatisfactory since a majority of the proteins had no predicted GO term for biological process, highlighting the need for more fruit proteomic studies. Interestingly, many proteins identified in the tomato ELPs have also been identified in mammalian exosomes as determined by the Extracellular Vesicle Database (EVpedia www.evpedia.info). This was the case for proteins involved in glycolysis, cytoskeleton modeling and membrane trafficking. There were also some plant specific proteins identified including proteins involved in photosynthesis and cellular signaling. Other protein families that were enriched in ELPs included proteases, lipid transfer proteins and expansins. It is difficult to identify all the proteins found in ELPs that are involved in defense or stress, as a majority of the tomato proteins are uncharacterized. It is also unclear if the cell wall modifying enzymes detected are all involved in plant defense or if they have an alternative function. What is clear is that plant ELPs are packaged with defense and stress related proteins. Interesting candidates included chitinases. Tomato does not contain any endogenous substrates for chitinases, thus it has been hypothesized that their role is strictly in defense **(Harikrishna et al., 1996)**. The acidic endochitinase is an extracellular protein and is considered to be a class II chitinase, lacking the cysteine-rich N-terminal portion presents in class I chitinases. The basic endochitinase is an

intracellular protein class I chitinase that has a cysteine rich N-terminal, a valine-rich signal peptide and vacuolar localization **(Danhash et al., 1993)**. The ELPs analyzed in this study showed an increase in the 27 kDa acidic endochitinase and the 30 kDa basic endochitinase compared to the fresh pressed tomato juice **(Figure 3.7 and 3.8)**. Comparing the expression of acidic to the basic endochitinase, it was found that the expression of the basic endochitinase increased more in the ELPs than did the acidic. The total expression levels were higher for the acidic endochitinase, but the expression increase compared to the juice was greater for the basic protein. This is an interesting result because the basic endochitinase is an intracellular protein yet has a higher increase in expression in the ELPs when compared to its acidic extracellular counterpart. This helps confirm the ability of plant ELPs to package and distribute proteins, whether intracellular or extracellular in nature. These results are the first to show the durability and longevity of plant ELPs. ELPs are present in plants long after growth, development, maturation and harvesting. This provides a new avenue for plant genetic engineering techniques that aim to improve plant defense/stress **(Albani et al., 1991; Krishnamurthy et al., 2001)**.

Future work could pursue questions like, does Heinz tomato juice contain ELPs also? Do the ELPs in V-8 Juice contain common ELPs from all eight fruit/vegetable sap/fluid exudates? Do such tomato ELPs help preserve freshness of this global beverage and nutritional supplement by means of its antifungal signals borne on the surface of, or in the interior of *Lycopersicum* exosomes? Can these ELPs be used to carry other active ingredients? This exploratory research lays the groundwork for the role of exosomal

proteins, miRNA and lipidomics in the burgeoning field of nutraceuticals and cosmeceuticals **(Fernando et al., 2019;Yahya et al., 2018)**.

6.2. PIN fusion system appears to localize to the starch granule surfaces in rice

To determine the subcellular localization of PINs, immunostaining of mature seed cross-sections were performed in concordance with the previous results showing PIN localization in mature wheat endosperm **(Feiz et al., 2009)**. The results help solidify the hypothesis that grain hardness is altered by PINs due to the interactions taking place at the starch granule-protein matrix interface **(Fujiwara et al., 2014)**. The immunolocalization technique is a viable assay to predict the tethering abilities of PIN, its homologues and of PIN mutants in mature endosperm, and their potential to alter grain hardness. It may also be useful in studying other starch granule interacting proteins **(Rosicka-Kaczmarek et al., 2015)**. Rice was used as our model cereal crop because it does not contain PIN or any of its homologues; it has been previously transformed with PIN (97-1), and the 97-1 rice lines showed a reduction in grain hardness **(Krishnamurthy and Giroux, 2001)**. As shown by the 97-1 experiments, it is possible to determine the grain altering effects through transgenic studies, however, these methods are extremely laborious and can take months to generate a transgenic line. The seed cross-section immunoassay developed in this study provides a quick and easy 'roll-over' assessment technique to analyze and predict the grain softening effects of PINs, their homologues and PIN mutants **(Kammeraad et al., 2016)**. Upon incubation with soft wheat seed protein extracts, rice seed cross-sections attracted PINs to their surfaces, causing them to translocate from the aqueous medium to the embedded specimen.

Since the localization of PIN is involved in dictating grain hardness, it becomes possible to predict its influence based on the quantity that interacts with the starch granule surfaces. To confirm this interaction takes place *in planta* and is not an artifact from the *in vitro* methodology, we compared the immunofluorescence to the transgenic PIN rice line 97-1. As observed in the 'roll-over' experiments, transgenic PIN had a strong affinity for starch granule surfaces **(Figure 4.2)**. These results support the idea that the 'roll-over' technique can be used to help predict the softening potential of a protein expressed *in planta*.

In addition to predicting the influence on grain hardness, the 'roll-over' study also provides insight into how PIN fusion proteins may interact *in planta*. Due to the nature of the starch granule; 5-8μm in diameter and a density of 1.5, we reasoned they could be easily purified from endogenous protein bodies (1-2μm in diameter) and based on the observations seen in the 'roll-over' experiment, we hypothesized that PIN fusions would tether onto rice starch granule surfaces.

Our fusion protein comprised PIN-a fused to a codon optimized hIGF-1. PIN-a has been shown to have a higher affinity for the lipids found on starch granule surfaces compared to PIN-b as determined by Langmuir trough experiments **(Clifton et al., 2008)**. It is interesting to note that PIN-a and PIN-b have a cooperative binding property, leading to an increased association with starch granule surfaces. If the PIN-a fusion protein does not show an affinity towards starch granule surfaces, it may be beneficial to attempt a double transformation, comprising PIN-a and PIN-b in bimodal, chimeric fusions. *Agrobacterium*-mediated transformation of rice had a very low efficiency. It was determined that the co-cultivation phase, where the mature calli are incubated with the

bacterium, was the limiting step. This was apparent due to the amount of necrosis and bacterial overgrowth taking place during and following this stage. To overcome this shortfall, we attempted to reduce the interaction between the rice calli and the *Agrobacterium*. At first we simply tried to reduced the OD_{600} of the inoculum from 0.3 to as low as 0.05. This seemed to have little to no effect on the efficiency. To overcome this problem, we implemented a new technique that involved the addition of a physical barrier between the calli and bacterium during the co-cultivation step. Rather than the traditional liquid co-cultivation step where the calli soaks in the inoculum, we developed a semi-dry inoculation step where the inoculum is placed underneath two filter papers, and the calli are placed on top. This semi-dry method seemed to limit the exposure of the bacterium to the calli, resulting in a more delicate transformation process and drastically increased yield. The resulting plantlets were grown in a greenhouse at 30 °C and 16hr day periods for four 4 months. Two months into plant growth, leafy tissue was collected and the DNA was extracted using the DNeasy plant mini kit from Qiagen. The genomic DNA was subjected to PCR analysis to confirm integration of the transgene. A restriction digest map was performed on the PCR fragment to ensure proper size of the promoter (1222bp), PIN-a (450bp) and IGF-1 (210bp) sequences **(Figure 4.4)**. Agarose gel electrophoresis of the digested products had the corresponding sizes confirming the correct assembly and integration of our construct into the rice genome. Upon rice plant maturation, seeds were collected and cross-sectioned as described in section 2.8 except the cross-sections were probed with an anti-IGF-1 primary antibody. Immunofluorescence of our fusion protein showed localization surrounding starch granule surfaces **(Figure 4.5)**. These results are in accordance with the hypothesis

generated from the 'roll-over' experiments, that PIN-fusions will also have an affinity for starch granule surfaces. To further verify the presence of our fusion protein in rice seeds, SDS-PAGE and Western Blot analyses were performed **(Figures 4.6 and 4.7)**. SDS-PAGE showed a band at 24 kDa, the expected size of our PIN-IGF-1 protein. This band was confirmed to be our transgene through western blotting using the anti-hIGF-1 primary antibody. It is important to note that no signal was observed for the leafy tissue protein extraction (lane 3 **Figure 4.7)**, demonstrating the seed localization capabilities of our GlutelinC promoter.

In conclusion, the PIN fusion system appears to localize to the starch granule surfaces in rice. These granules can be subsequently isolated using standard milling and air classification technologies **(Gomez and Martinez 2016),** alleviating some of the workload required by chromatography.

6.3. Dry Phission of fusion proteins

Smaller particle size of starch granules observed with jet milling as compared to hammer milling can be due to the heat associated with hammer milling. The gelatinization temperature of rice starch is 62°C **(Tsutsui et al., 2005)**, and while the temperature measured at the top of the comminution chamber was under this threshold, it does not account for the slight increase of temperature occurring at the tips of the hammers, where maximal impact friction is occurring. An increase in heat at these points could have caused micro-gelatinization, resulting in larger particle sizes. Jet milling allows for a higher degree of control through feeding and pressure alterations. A decrease in feed rate and an increase in pressure results in smaller particle diameter.

These are consistent with previous jet milling reports **(Angelidis et al., 2016)**. To determine whether milling had any effect on the flour structure, scanning electron microscopy was performed. According to the micrographs, a higher degree of disaggregation was observed in jet-milled samples as compared to the hammer milled samples **(Fig. 5.3)**. Some starch granules appeared deformed with rounder shape, whereas other granules, freed from the surrounding protein matrix, appeared to be in smaller aggregates or standing alone with their standard polygonal shape. These starch granules, released from the protein matrix are consistent with previous reports, suggesting jet milling may be an efficient process for the separation of starch granules from proteins **(Sanguansri et al., 2006)**.

Cleavage of the model proteins tyrosinase and catalase yielded fragments of the anticipated size as determined through electrophoresis. The location of the DP sites found within the proteins had no impact on gas-phase non-protease 'proteolytic' fragmentation. This is in accordance with previous results that used cyanogen bromide as the cleavage reagent on bovine pancreatic ribonuclease **(Gross et al., 1962)**. It was also determined that cleavage using 7-HFBA required a minimum reaction time of 16 hours, and a temperature of at least 60°C to obtain complete fragmentation. Similar findings were observed through cleavage of an *E. coli* derived thioredoxin **(Wen et al., 1993)**. Anhydrous cleavage products of rice starch granule proteins **(Figure 5.5)** were in accordance with the aqueous trypsin digested starch granule associated proteins **(Table 5.1)**. There was a lower amount of protein observed in the air-classified samples as compared to the Sigma starch samples. Based on the MS results from the tryptic digested Sigma starch, it was observed that several contaminating proteins were present. In the

Koziol study it was hypothesized that these contaminants were introduced mostly through the aqueous process **(Koziol et al., 2012)**. These present results support the hypothesis that an anhydrous method will reduce contaminants when compared to its aqueous counterpart. Comparisons between the anhydrous and aqueous cleavage products through electrophoresis showed several similarities (**Figure 5.5** lanes 2/3 and lanes 5/6), which could correspond to some well known and expected rice seed proteins such as glutelin, prolamin, starch granule bound synthase, amongst others **(Koller et al., 2002)**. A fragment for the 27kDa glutelin protein was observed in all samples, and appeared to be the most abundant starch granule associated protein. Glutelins are the most abundant rice endosperm based proteins **(Xing et al., 2016)**, thus these results are understandable. However, glutelins are found predominantly within protein bodies (**Li et al., 1993**), so their high abundance may allude to the inefficiencies of the air-classification process. The most likely explanation is that following jet-milling, some protein bodies may piggyback along with individual or within compound starch granule structures. A reduction in wait time of sample handling between jet milling and air-classification may reduce the agglomeration of particles and thus limit the adherence of contaminants with the starch granules. It would also be interesting to perform the milling and air-classification in an inert gas stream that is chilled, which would reduce the electrostatic and van der Waals forces present in the powder, providing a more pure starch granule fraction **(Li et al., 2004)**.

Future experiments involving MS analysis of the anhydrous cleavage products would further confirm our technique, and may also provide for new starch granule fusion partners, for example novel carbohydrate binding domains to serve as carrier tethers in

plant protein farming of vaccines, hormones, cytokines and biocatalysts **(Greenham and Altosaar 2013)**.

CHAPTER 7. CONCLUSION

Plant molecular pharming presents an efficient alternative to standard cell culture protein production methods. The dry rice seed represents an ideal deposition site for recombinant proteins. Fusion strategies such as the one implemented in this thesis can help ease the purification process and reduce the environmentally harmful byproducts associated with chromatographic purification strategies. By tethering recombinant proteins to starch granule surfaces, purification can be accomplished anhydrously. These anhydrous methods are a fraction of the costs of the standard aqueous methods (solubilization, fractionation, precipitation, size exclusion, ion exchange, affinity chromatography, iontophoresis, etc.). Furthermore, the anhydrous environment ensures a pathogen and protease free area of deposition. Another advantage of starch granule tethering is that biologics that do not require high purity, for example industrial enzymes such as transglutaminase, can simply be purified using milling and air-classification strategies. The cost effective scalability of plants provides a new avenue for production of more commodity-based biologics. Future studies will involve the addition of the anhydrous cleavable linker sequence that will link puroindoline to the protein of interest in order to liberate and purify the recombinant protein with a completely anhydrous strategy.

Plant-made pharmaceuticals have a promising future. This was highlighted by the FDA approval of a taliglucerase alfa for Gauchers disease produced in carrot cell-culture,

and more recently through the success of ZMAPP, a tri-monoclonal antibody product for Ebola treatment produced in Tobacco. As the regulatory framework becomes more streamlined, we should expect to see several more plant made biologics to hit the market, following the earlier success of Sigma-Aldrich's made-in-plant chicken egg white avidin (cat.no. A8706) and bovine trypsin (cat.no. T3568, TrypZean).

Plant ELPs are also gaining considerable attention in plant research and present an exciting new field for biotechnology applications. In this study we characterized the tomato ELP proteome. In conjunction with previous reports **(Regente et al., 2017; Rutter and Innes, 2017)**, our ELPs were enriched in plant defense and stress related proteins. This study demonstrated that ELPs remain intact and loaded with proteins long after fruit ripening and harvest. This has led to the hypothesis that ELPs help extend the shelf life of plant products and offer an attractive target for biotechnological engineering. Future studies can be directed at characterizing the DNA and RNA content of tomato ELPs. Understanding how exosomes are packaged *in planta* will be a key component in exploiting them for plant defense and stress response purposes. Isolated ELPs are being used as nanocarriers for different compounds and proteins **(Akuma et al., 2019)**. Perhaps the most intriguing use shown thus far has been the incorporation of doxorubicin, a chemotherapy drug, in ginger derived ELPs through electrostatic interaction **(Zhang et al., 2016b)**. The method of incorporation allowed for a pH dependent release following intravenous delivery.

This thesis explored protein farming in plant hosts, because the plant system is being focused on globally as a more cost-effective and safe protein production platform

for various biosimilars, much needed yet still awaiting affordable scale-up **(Jin et al., 2019)**.

REFERENCES

Akers, J.C., et al. (2013) MiR-21 in the extracellular vesicles (EVs) of cerebrospinal fluid (CSF): a platform for glioblastoma biomarker development. PLoS one. *8*, e78115.

Akuma, P., et al. (2019) Naturally occurring exosome vesicles as potential delivery vehicle for bioactive compounds. Frontiers in sustainable food systems.

Albani, D., et al., (1991) A gene showing sequence similarity to pectin esterase is specifically expressed in developing pollen of brassica-napus sequences in its 5' flanking region are conserved in other pollen-specific promoters. Plant molecular biology. 16, 501-513.

Albuquerque, PC., et al. (2008) Vesicular transport in *Histoplasma capsulatum*: an effective mechanism for trans-cell wall transfer of proteins and lipids in ascomycetes. Cellular microbiology. 10**,** 1695–1710.

Ali, I., et al. (2015) Molecular characterization of the puroindoline-a and b alleles in synthetic hexaploid wheats and in silico functional and structural insights into Pina-D1. Jouranl of theoretical biology. 376, 1-7.

Angelidis, G., et al. (2016) Jet milling effect on wheat flour characteristics and starch hydrolysis. Journal of food science and technology. 53, 784-791.

Basu, S et. al. (2018) Estimation of global insulin use for type 2 diabetes, 2018-30: a microsimulation analysis. The Lancet. 7, 25-33.

Barta, A., et al. (1986) The expression of a nopaline synthase human growth hormone chimaeric gene in transformed tobacco and sunflower callus tissue. Plant molecular biology, 6, 347–57.

Biswas, S.C., and Marion, D., (2006) Interaction between puroindolines and the major polar lipids of wheat seed endosperm at the air-water interface. Colloids surf. B-biointer. 53, 167-174.

Bloch, H.A., et al. (2001) In vitro binding of puroindolines to wheat starch granules. Cereal chemistry. 78, 74-78.

Boller, T., and Felix, G. (2009) A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. Annual Reviews in Plant Biology. 60, 379-406.

Both, M., et al, (2005) Gene expression profiles of *Blumeria graminis* indicate dynamic changes to primary metabolism during development of an obligate biotrophic pathogen. The Plant cell. 17, 2107–2122.

Chantret, N., et al. (2005) Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploidy wheat species (Triticum and aegilops). Plant cell. 17, 1033-1045.

Chen, Q., et al. (2014) Plant made biologics. Biomedical research 2004, 1-3.

Clifton, L.A., et al. (2008) Interfacial structure of wild-type and mutant forms of puroindoline-b bound to DPPG monolayers. Journal of physical chemistry B. 112, 15907-15913.

Dangl, J., and Jones, DG. (2001) Plant pathogens and integrated defense responses to infection. Nature. 411, 826-833.

Danhash, N., et al., (1993) Molecular characterization of four chitinase cDNAs obtained from fulvum-infected tomato. Plant molecular biology. 6, 1017-1029.

Day, L., et al. (2006) Characterization of wheat puroindoline proteins. FEBS J. 273:,5358-5373.

De palma, G., et al. (2016) Clinical application of human urinary extracellular vesicles in kidney and urologic diseases. International journal of molecular science. *17*.

Dirisala, M., et al. (2017) Recombinant pharmaceutical protein production in plants: unraveling the therapeutic potential of molecular pharming. Acta physiologiae plantarum 39,18.

Edman, P., (1967) A Protein Sequenator. European journal of biochemistry. 1, 80-91.

Elmore, S., (2007) Apoptosis: a review of programmed cell death. Toxicology and Pathology. *35*, 495-516.

Eng, JK., et al. (1994) An approach to correlate tandem mass-spectral data of peptides with amino acid sequences in a protein database. Journal of the American society for mass spectrometry. 11, 976-989.

Fernando, I., et al. (2019) Algal polysaccharides: potential bioactive substances for cosmeceutical applications. Critical reviews in biotechnology. 39, 99-113.

Feiz, L., et al. (2009) Puroindolines co-localize to the starch granule surface and increase seed bound polar lipid content. Journal of cereal science. 50, 91-98.

Fujiwara, M., et al. (2014) Localization of transgene-derived friabilins in rice endosperm cells. Plant biotechnology. 31, 67-70.

Gasparis, S., et al. (2011) The RNA-mediated silencing of one of the Pin genes in allohexaploid wheat simultaneously decreases the expression of the other, and increases grain hardness. Journal of experimental botany. 62, 4025-4036.

Gautier, M.F., et al. (2000) Puroindoline genes are highly conserved in diploid ancestor wheats and related species but absent in tetraploid *Triticum* species. Plant science. 153, 81-91.

Gomez, M., et al. (2016) Changing flour functionality through physical treatments for the production of gluten-free baking goods. Journal of cereal science. 67, 68-74.

Greenham, T., and Altosaar, I., (2013) Molecular strategies to engineer transgenic rice seed compartments for large-scale production of plant-made pharmaceuticals. Methods in molecular biology. 956, 311-326.

Greening, D., et al. (2018) Understanding extracellular vesicle diversity – current status. Expert review of proteomics. 15, 887-910.

Gross, E., and Witkop, B., (1962) Nonenzymatic cleavage of peptide bonds: The methionine residues in bovine pancreatic ribonuclease. Journal of biological chemistry 237, 1856-1860.

Gyorgy, B., et al. (2015) Therapeutic applications of extracellular vesicles: clinical promise and open questions. Annual review in pharmacology and toxicology. *55*, 439-464.

Halperin, W and Jensen, WA (1967) Ultrastructural changes during growth and embryogenesis in carrot cell cultures. Journal of ultrastructure research. 18, 428–443.

Hansen, L., et al, (2017) Plant exosomes: using an unconventional exit to prevent pathogen entry? Journal of experimental botany. 69, 59-68.

Harikrishna, K., et al., (1996) N endochitinase gene expressed at high levels in the stylar transmitting tissue of tomatoes. Plant molecular biology. 30, 899-911.

Hiei, Y., and Komari, T., (2008) Agrobacterium-mediated transformation of rice using immature embryos or calli induced from mature seed. Nature protocols. 3**,** 824-834.

Hofgen, J., et al. (1988) Storage of competent cells for *Agrobacterium* transformation. Nucleic Acids research 16, 20-25.

Jeong, E., et al. (2008) Effect of particle size on the solubility and dispersibility of endosperm bran and husk powders of rice. Food science and biotechnology 4,833-838.

Jin, N., et al. (2019) Low binding affinity and reduced complement-dependant cell death efficacy of ofatumumab produced using a plant system (Nicotiana benthamiana L.). Protien expression and purification. 159, 34-41.

Johnstone, RM., et al. (1987) Vesicle formation during reticulocyte maturation – association of plasma-membrane activities with released vesicles (exosomes). Journal of biological chemistry 19, 9412-9420.

Jurgen Cox, Marco Y Hein, Christian A Luber, Igor Paron, Nagarjuna Nagaraj, and Matthias Mann. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed maxlfq. Molecular & Cellular Proteomics, 13(9):2513–2526, 2014.

Kammeraad, J.D., et al. (2016) Mutagenesis-derived puroindoline alleles in *Triticum aestivum* and their impacts on milling and bread quality. Cereal chemistry. 93, 201-208.

Kaneko, T., (1968) Chemical cleavage of Threonine peptides. Bulletin of the chemical society of Japan. 41, 2155-2161.

Karlsson, O., et al. (2016) Detection of long non-coding RNAs in human breastmilk extracellular vesicles: Implications for early child development. Epigenetics. 11, 721-729.

Koziol, A., et al. (2012) The starch granule associated proteomes of commercially purified starch reference materials from rice and maize. Journal of proteomics 3, 993-1003.

Krishnamurthy, K., et al. (2001) Wheat puroindolines enhance fungal disease resistance in transgenic rice. Molecular plant-microbe interactions. 14, 1255-1260.

Krishnamurthy, K., and Giroux, M.J., (2001) Expression of wheat puroindoline genes in transgenic rice enhances grain softness. Nature biotechnology. 19, 162-166.

Kowal, J., et al. (2016) Proteomic comparison defines novel markers to characterize heterogeneous populations of extracellular vesicle subtypes. PNAS. *113*, 968-977.

Li, X., et al. (1993) Segregation of storage protein messenger-RNAs on the rough endoplasmic reticulum membranes of rice endosperm cells. Cell. 6, 869-879.

Li, Q., et al. (2004) Interparticle Van Der Waals force in powder flowability and compactibility. International Journal of pharmaceuticals. 280, 77-93.

Maxmen, A., (2012) Drug-making plant blooms. Nature 485,160.

Miyado, K., et al. (2017) Exosomes versus microexosomes: Shared components but distinct functions. Journal of plant research. 130, 479-483.

Monisha, S., et al. (2015) Extracellular vesicles including exosomes in cross kingdom regulation: a viewpoint from plant-fungal interactions. Frontiers in plant science. 2, 13-24.

Morin, XK., (2008) Genetically modified food from crops: progress, pawns, and possibilities. Analytical and bioanylitical chemistry. 392, 333-340.

Morris, C.F., (2002) Puroindolines: the molecular genetic basis of wheat grain hardness. Plant molecular biology 48, 633-647.

Mu, J., et al. (2014) Interspecies communication between plant and mouse gut host cells through edible plant derived exosomes-like nanoparticles. Molecular nutrition and food research. 58, 1561-1573.

Oerke, EC (2006) Crop losses to pests. The journal of agricultural science. 144, 31-43.

Pande, S., et al. (2017) *Atta* (whole wheat flour) with multi-wholegrains: flour characterization, nutritional profiling and evaluation of *chapatti* making quality. Journal of food science technology 11, 3451-3458.

Park, E., et al. (2017) Plant-microbe interactions: organelles and the cytoskeleton in action. Tansley review. 217, 1012-1028.

Pauly, A., et al. (2013) Wheat (*Triticum aestivum L*. and *T. turgidum L.* ssp durum) kernel hardness: II. Implications for end-product quality and role of puroindolines therein. Comprehensive reviews in food science. 12, 427-438.

Partridge, S., and Davis, H., (1950) Preferential release of aspartic acid during the hydrolysis of proteins. Nature 4185, 62-63.

Prado, N., et. al. (2014) Nanovesicles are secreted during pollen germination and pollen tube growth: a possible role in fertilization. Molecular plant. 7**,** 573–577.

Qu, L., et al. (2008) Expression pattern and activity of six glutelin gene promoters in transgenic rice. Journal of experimental botany. 9, 2417-2424.

Regente, M., et. al. (2009) Vesicular fractions of sunflower apoplastic fluids are associated with potential exosome marker proteins. FEBS Lett. 583, 3363– 3366.

Regente, M., et al. (2017) Plant extracellular vesicles are incorporated by a fungal pathogen and inhibit its growth. Journal of experimental biology. 68, 5485-5495.

Robbins, P. D and Morelli, A. E., (2014) Regulation of immune responses by extracellular vesicles. Nature Reviews in immunology. *14*, 195–208.

Rosicka-Kaczmarek, J., et al. (2015) Fluorimetric studies of the interactions of wheat puroindolines with polar lipids on the surface of starch granules. Journal of cereal science. 66, 53-58.

Rutter, B and Innes, R (2017) Extracellular Vesicles Isolated from the Leaf Apoplast Carry Stress-Response Proteins. Plant physiology. 173, 728-741.

Sanguansri, H., and Augustin, M., (2006) Nanoscale materials development – a food industry perspective. TIFST 17, 547-556.

Savatin, D., et al. (2014) Wounding in the plant tissue: the defense of a dangerous passage. Frontiers in plant science. 5, 470.

Sharma, N., et al. (2011) Role of chitinase in plant defense. Asian journal of biochemistry. 6, 29-37.

Takaiwa, F., et al. (2017) An overview on the strategies to exploit rice endosperm as production platform for biopharmaceuticals. Plant science 263, 201-209.

Tanchak, M., et al. (1998) Tryptophanins: isolation and molecular characterization of oat cDNA clones encoding proteins structurally related to puroindoline and wheat grain softness proteins. Plant science. 137, 173-184.

Thery, C., et. al. (2002) Exosomes: composition, biogenesis and function. Nature reviews. 2, 569-579.

Thery, C., et al. (2019) Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. Journal of extracellular vesicles 8, 7-42.

Tsutsui, K., et al. (2005) Effect of Annealing Temperature on Gelatinization of Rice Starch Suspension As Studied by Rheological and Thermal Measurements Journal of agricultural and food chemistry. 53, 9056-9063

Tuse, D., et al. (2014) Manufacturing economics of plant-made biologics: case studies in therapeutic and industrial enzymes. Biomedical research 2014, 1-16.

Wall, M.L., et al. (2010) The tryptophan-rich domain of puroindoline is directly associated with the starch granule surface as judged by tryptic shaving and mass spectrometry. Journal of cereal science. 52, 115-120.

Wen, Y., et al. (1993) Identification of the facile gas-phase cleavage of the Asp-Pro and Asp-Xxx peptide bonds in matrix-assisted laser desorption time-of-flight mass spectrometry. Analytical chemistry 65, 3015-3023.

Wolfers, J., et al. (2001) Tumor-derived exosomes are a source of shared tumor rejection antigens for CTL cross-priming. Nature medicine. 7, 297–303.

Xie, F., et al. (2011) Liquid chromatography-mass spectrometry-based quantitative proteomics. Journal of biological chemistry. *286*, 25443-25449.

Xing, S., et al. (2016) Proteome Profile of Starch Granules Purified from Rice (*Oryza sativa*) Endosperm. Plos one 11, 168-179.

Xu, A.T., et al. (2016) Exosome in intestinal mucosal immunity. Journal of Gastroenterology and Hepatology *31*, 1694-1699.

Yahya, N., et al. (2018) An overview of cosmeceutically relevant plant extracts and strategies for extraction of plant-based bioactive compounds. Food and bioproducts processing. 112, 69-85.

Young, L., et al. (2001) Two-step total gene synthesis. Nucleic acid research 32, 59-65.

Zernecke, A., et al. (2009) Delivery of microRNA-126 by apoptotic bodies induces CXCL12-dependent vascular protection. Science signal *2*, 180-186.

Zhang, Y., et al. (2013) Protein analysis by shotgun/bottom-up proteomics. Chemical reviews. *113*, 2343-2394.

Zhang, M., et al. (2016b) Edible ginger-derived nano-lipids loaded with doxorubicin as a novel drug –delivery approach for colon cancer therapy. Molecular therapy. 24:10.

Zheng, X., et al. (2017) Salivary exosomal PSMA7: a promising biomarker of inflammatory bowel disease. Protein cell. *8*, 686-695.

CONTRIBUTION OF COLLABORATORS

1. Dr Illimar Altosaar (University of Ottawa)
Dr. Altosaar contributed to the study design and interpretation of results and provided critical content to my thesis.

2. Charles Melnyk (University of Ottawa)
Charles Melnyk performed the Microscopy images on the soft wheat, hard wheat, rice and 97-1 rice micrographs.

3. Mark Horsman (University of Ottawa)
Mark Horseman performed the SEM of the rice flour particles after hammer and jet milling.

4. Ottawa Institute of Systems Biology

The mass spectrometry analysis and database search was performed by the Ottawa Institute of Systems Biology.

Trevor J. Greenham

Education

University of Ottawa, Ottawa, ON

Ph.D. Candidate in Biochemistry, Winter 2013 to Present

- Recipient of a NSERC Industrial Postgraduate Scholarship 2 (April 2012 – March 2015)
- Recipient of a NRC IRAP
- Recipient of MITACS grant
- Recipient of Excellence Scholarship (July 2012 and July 2015)

Graduate courses:

- RNA class: Dr Pelchat (A)
- Protein form and function: Dr Figeys (A)
- Molecular biology lab course: Dr Ekker (A+)
- Cholesterol and miRNA: Dr Rayner (A+)

University of Ottawa, Ottawa, ON

M.Sc. Candidate in Biochemistry, September 2010 to December 31, 2012

University of Ottawa, Ottawa, ON

B.Sc. Specialization Biochemistry & Minor Business Administration, 2005-2009

- Graduated Cum Laude, 2010

Experience

Cosmeceutica Biotechnology (2016-Present)

- Founder, CEO
- Recipient of the start-up garage entrepreneurial competition ($20,000)
- Production of active ingredients for the skin care markets

Proteins Easy Corp

Business Development Manager, July 1 2010 – June 30 2011

- Recipient of a NRC-IRAP YES grant
- Developed a business plan, learned the marketing, administration, operation and financial management aspects of the company
- Competed in Carleton University's Ontario Talent First Network's Lead-To-Win entrepreneurs program and achieved Top-five status for PEC's team, out of a total of forty companies, gaining ecosystem acceptance.

- Improved my public speaking skills by preparing and presenting a 5 minute business pitch in front of two Venture Capitalists, one Angel Investor and a third year University Entrepreneurial class

Protein Biotechnology Laboratories

PhD, Msc, and BSc Hon. Thesis experimentations, Sept 2009 to present

- Investigated the effects of Nitrous Oxide Reductase and the *nos* gene clusters from *Pseudomonas stutzeri* when expressed in transgenic tobacco
- Expression of recombinant fusion proteins in Chlamydomonas reinhardtii
- Molecular pharming of human proteins within *Arabidopsis thaliana* and *Oryza sativa* using a puroindoline fusion system.
- Anhydrous cleavage or proteins
- Plant exosome isolation and packaging

Entrepreneur Competitions http://techvc.org/

Tech Venture Challenge, Business plan competition Ottawa, ON

- Placed top four out of 32 teams in 2009
- Placed top five out of 38 teams in 2010

Peer Help Center – University of Ottawa, Sept 2007 to Jan 2012

- Tutoring Biochemistry, Metabolism, Physics, Organic chemistry, Analytical Chemistry, Physical Chemistry, Biochemistry Lab 2 and Molecular Biology Lab.

Residence Study Group Leader – University of Ottawa, 2008 and 2009

- Lead weekly study groups for first year undergraduates in physics 1101/1102.

Patent Applications

Co-Inventor: USPTO Application #61887986. Title of Invention: Methods for separating and purifying endogenous, exogenous and recombinant proteins/peptides from plants and animals using aqueous-free, anhydrous strategies.

Invention Disclosures

Co-inventor: Isolation and packaging of plant exosome-like vesicles and/or plant extracellular vesicles for use in the skin care markets

Publications

Greenham T and Altosaar, I (2018) Wheat puroindolines tether to starch granule surfaces in puroindoline-null (Pin-null) plants. Journal of Cereal Science.

Greenham T, and Altosaar I, (2012) Engineering transgenic rice seed compartments for large-scale biological production of plant-made pharmaceuticals. In: "Rice Protocols Handbook" Ed. Yang Y., Humana Press, Springer, Totawa NJ.

Zaidi M, El Bilali J, Koziol A, Ward T, Styles G, **Greenham T,** Faiella W, Son H, Wan S, Taga I, Altosaar I. (2012) Gene technology in agriculture, environment and biopharming: Beyond Bt-rice and building better breeding budgets for crops. J Plant Biochem & Biotechnology

Shen Wan, **Trevor Greenham,** Kagami Goto, Yaseen Mottiar, Amanda M. Johnson, Julianne M. Staebler, Mohsin A. Zaidi, Qingyao Shu and Illimar Altosaar. (2012) A novel nitrous oxide mitigation strategy: expressing nitrous oxide reductase from *Pseudomonas stutzeri* in transgenic plants. Can J Plant Sci.

RIGHT AND PERMISSIONS

For Greenham, T., and Altosaar, I., (2018) Wheat puroindoline tether to starch granule surfaces in puroindoline-null (Pin-null) plants. Journal of Cereal Science. 79, 286-293.

https://service.elsevier.com/app/answers/detail/a_id/14290/kw/rights%20and%20permission%20/supporthub/publishing/
Rights and permission: As a journal author, you retain *rights* for a large number of author uses, including use by your employing institute or company. These *rights* are retained and permitted without the need to obtain specific *permission* from Elsevier.

Licences (Permissions) For Frontiers in physiology Bellingham, S., (2012) Exosomes: vehicles for the transfer of toxic proteins associated with neurodegenerative diseases?. 124, 1-12.

When you submit an article to Frontiers, you grant to Frontiers and to the world at large a permanent, non-cancellable, free-of-charge, worldwide licence (permission) to publish, display, store, copy and re-use that article – including any third-party materials – and to create derivative works from it. You can not terminate that licence. You must ensure that you have all necessary permissions from third parties. Ownership by the third party of the copyright can still be notified on the relevant materials, and attribution must be made in accordance with usual scholarly practices.