THE STARCH GRANULE SURFACE: TECHNOLOGICAL AND BIOLOGICAL IMPLICATIONS OF PUROINDOLINE AND HOST-PATHOGEN INTERACTIONS

by

Michael Louis Wall

A thesis submitted to the Faculty of Graduate and Postdoctoral Studies

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy - Biochemistry

Department of Biochemistry, Microbiology and Immunology

Faculty of Medicine

University of Ottawa

ABSTRACT

The sun is the primary source of all chemical energy on the planet. Starch granules have evolved as storage deposits for captured light energy. Many complex biological functions take place at the starch granule surface, including starch granule metabolism and defense. The starch granule-associated protein puroindoline is a known antimicrobial with unique functional and biological properties, attributed to the presence of a unique tryptophan-rich domain. To test puroindoline's tight association, puroindoline removed from the starch granule surface during water-washing was assessed. Washing more than eight times failed to further reduce puroindoline content of starch granules, suggesting a strong association of puroindoline with the starch granule surface. To identify the tryptophan-rich domain tightly associated with the starch granule surface, we used a combination of *in situ* tryptic digestion and mass spectrometry. We identified the tryptophan-rich domain of puroindoline directly bound to the starch granule surface of wheat. This is the first instance of the tryptophan-rich domain directly observed at the starch granule surface. In addition, using mass spectrometry, we determined that during development and maturation, wheat seeds appear to have resisted infection and lysed the pathogens where, upon desiccation, the molecular evidence remained fixed at the starch granule surface. Proteins with known antimicrobial activity were identified, as well as several proteins from the plant pathogens *Agrobacterium tumefaciens*, *Pectobacterium carotovorum*, *Fusarium graminearum*, *Magnaporthe grisea*, *Xanthomonas axonopodis*, and *X. oryzae*. Future characterization may reveal previously unknown host-pathogen interactions. Finally, we have demonstrated that puroindoline, when expressed in the seeds of transgenic corn, will localize and associate with the starch granule surface in a pattern similar to the puroindoline expression pattern observed in wheat. Surprisingly, puroindoline expression in transgenic corn is correlated with an increase in total seed oil content.

Acknowledgements

This thesis is dedicated to the members of the Altosaar lab, both past and present. Though students come and go, the camaraderie and collegiality of the Altosaar lab has been a constant during my time there. I would like to thank my graduate supervisor, Dr. Illimar Altosaar. He has always been supportive of my ideas and creativity and I am constantly in awe of his incredible and infectious passion for science and life.

This thesis would not have been possible without the encouragement of my family. I am ever indebted to my parents, John and Vicki Wall, for their love and support and for making me who I am today. I need to thank my sister Julie and nieces Kaylee and McKenzie. All of you are such great sources of encouragement and fun. To my brother-in-arms, Mark: You're the greatest brother a guy could have, thanks for always being there. Finally, I need to thank Natalie, whom I love and cherish and for being the greatest person in my life.

Table of Contents

List of Tables and Figures

LIST OF ABBREVIATIONS

ASK - Alexho single kernel
BHT - Butylated hydroxy toluene
BLAST - Basic Local Alignment Search Tool
BSA - Bovine serum albumin
Ct - Threshold cycle
DAP - Days after pollination
DGAT - diacylglycerol acyltransferase
DGDG - digalactosyl diacylglycerol
ESI - Nanoelectrospray ionization
EST - Expressed sequence tag
FAME - Fatty acid methyl esters
FITC - Fluorescein isothiocyanate
GBSSI - Granule bound starch synthase I
GC - Gas chromatography
GIMP - GNU Image Manipulation Program
GSP - Grain softness protein
HPLC - High pressure liquid chromatography
HRP - Horseradish peroxidase
IHO - Illinois high-oil
LC - Liquid chromatography
LTP - Lipid transfer protein
MAP - Mitogen-activated protein
MS - Mass spectrometry
NCBI - National Center for Biotechnology Information
nr - Non-redundant
nsLTP - Non-specific lipid transfer protein
ORF - Open reading frame
PAGE - Polyacrylamide gel electrophoresis
PBST - Phosphate buffered saline Tween 20
PE - Phosphatidylethanolamine
PIN - Puroindoline
PR - Pathogenesis-related
QqTOF - Quadrupole/quadrupole time-of-flight
qRT-PCR - Quantitative real-time polymerase chain reaction
QTL - Quantitative trait loci
RT - Room temperature
SBD - Starch-binding domain
SDS - Sodium dodecyl sulphate
SGAP - Starch granule associated protein
TAG - Triacylglycerol
TOF - Time-of-flight
TMB - 3,3',5,5'-tetramethylbenzidine
TRD - Tryptophan-rich domain
TX-114 - Triton X-114
XIP - Xylanase-inhibiting protein

Chapter 1

General Introduction

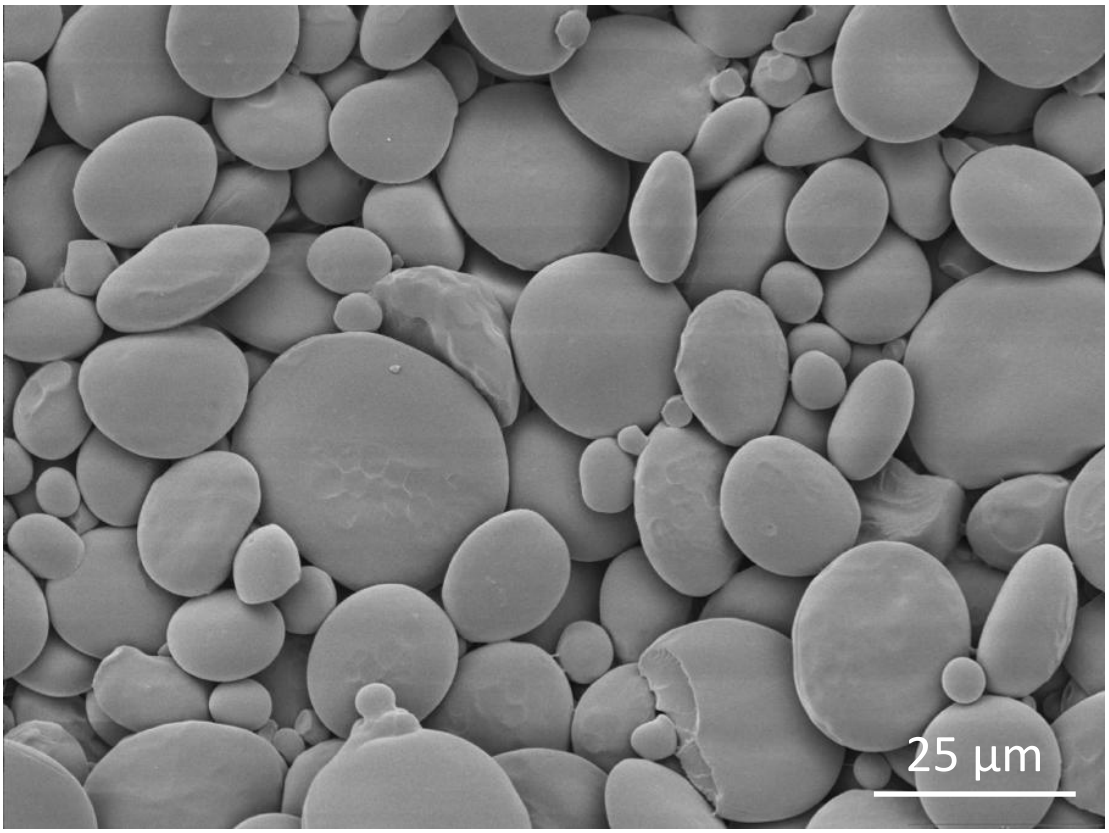
1.1 Starch granules

1.1.1 Starch granule function

The sun is the primary source of all chemical energy on the planet (Krause and Weis 1991). Green plants use the energy of incident photons to split water and use the resulting constituent elements and air-derived carbon dioxide to create carbohydrates (Voncaemmerer and Farquhar 1981). This chemically-captured sunlight is stored as carbohydrate polymers into high-density granular deposits called starch granules for either short-term (day/night) consumption or long term reserve in the form of dry endosperm, typified by the grain of cereal plants (**Figure 1.1**) (Schnyder 1993).

Starch granule evolution in the grasses has resulted in a diverse array of functions for these insoluble sub-cellular organelles (Zeeman et al. 2007). Functions include gravisensing (Moulia and Fournier 2009), diurnal energy reserve in green leaf tissue (Zeeman et al. 2007) and, most characteristic of cereals, long term reserve energy storage in the form of starchy endosperm for the developing embryo (Smith et al. 2005).

Figure 1.1. Scanning electron micrograph of isolated, water-washed starch granules from wheat. Starch granules are high-density semi-crystalline deposits of amylose and amylopectin found in a wide variety of organisms and are the primary component of cereal endosperm. (Unpublished data).


25 µm

1.1.2 Starch granule structure

The starch granules of photosynthetic organisms represent high-density energy reserves of captured sunlight. The synthesis of its constituent carbohydrate polymers, amylose and amylopectin, and the mechanism for the accumulation of these sugars into semi-crystalline aggregates has been the subject of intense research for over fifty years (Buleon et al. 1998). Amylose is a linear molecule composed of α-1,4-linked glucose subunits with rare α-1,6-branching and, generally, the major structural component of starch granules is not amylose, but amylopectin (Ball and Morell 2003). Similar to amylose, amylopectin is a polymer composed of α-1,4-linked glucose subunits, but differs in that it is highly branched, with α-1,6-branching occurring every 20 to 24 glucose subunits (Ball and Morell 2003). There are high-amylose as well low-amylose "waxy" cereals and potatoes that have either zero or very low amylose content (Sasaki et al. 2000). Though all starch granules are composed of amylopectin and amylose, the ratio of these polymers result in subtle structural differences and widely different physiological properties (Sasaki et al. 2000). In addition to species and variety-specific differences in starch granule properties, tissue-specific starch granules differ in composition and physical properties. For example, the diurnal transient starch granules of leaf tissue are primarily composed of amylopectin (Zeeman et al. 2007), whereas the starch granules found in endosperm cells for long-term energy storage are significantly larger and typically have an amylose content of ~30% (Buleon et al. 1998; Smith 2001).

1.1.3 Starch granule surface structure

The dense nature of the starch granule necessitates that proteins involved in defense of the starch granule or starch granule metabolism be situated within the intracellular protein matrix, cytoplasm or at the granule surface. The granule surface itself then becomes an interesting subcellular compartment distinct in functionality from that of the rest of the granule.

The surface structure of starch granules is marked by protrusions, channels, pores and other features, the result of the individual starch granule growth patterns (Fannon et al. 1992; Fannon et al. 2003; Huber and BeMiller 1997, 2000). Atomic force microscopy was used by Baldwin et al. (1998) to reveal spherical 'blocklet' structures, assumed to be composed of amylopectin, protruding 30-300 nm from the surface of potato starch granules (Baldwin et al. 1998). The wheat starch granule surface did not have these features. This demonstrated that, aside from relatively large-scale differences in starch granule structure, starch granules can differ in subtle surface features. The presence, size and frequency of channels and pores depend on the organism, the tissue type and the starch granule type and variations of such features have been observed in wheat (Han et al. 2005; Huber and BeMiller 2000). The importance or effect of surface features on host-pathogen interaction or host metabolism is unknown.

1.1.4 Role of starch granules in biotechnology

Recently, the study of starch granule metabolism has intensified (Streb et al. 2008), a consequence of the emerging paradigm of 'alternative fuels' and the potential for high-energy

plants, such as the grasses, to potentially play a role (Somerville 2006). It is clear that a detailed knowledge of plastid, amyloplast and starch granule molecular biology, biochemistry and metabolism will be necessary before the directed manipulation of these pathways is possible (Morell and Myers 2005).

Genetically modified plants for the production of useful recombinant proteins is still in an early stage of development with issues of risk, purity and containment still needing to be addressed (Murphy 2007; Ramessar et al. 2008; Shama and Peterson 2008; Spoek 2007). In the meantime, research to improve or overcome technical challenges presented by the need for high-level production of recombinant enzymes, vaccines (Streatfield 2007) and antibodies (De Muynck et al. 2010; Pogue et al. 2010), control of post-translation modifications (Gomord et al. 2004; Saint-Jore-Dupas et al. 2007) and effective delivery (Menassa et al. 2007; Takagi et al. 2005) are making steady progress. Various end-use quality and value-added enhancements of genetically modified plants are also underway (Oksman-Caldentey 2007).

Technological modifications, such as increased or modified starch in potato and cereals is desirable; starch and starch products are some of the most highly used natural materials in industry (Blennow et al. 2005; Kossmann and Lloyd 2000). Starch is a renewable resource of chemically-captured sunlight that is utilized in many industrial applications. Though most starch is derived from corn and potato, increasing proportions are being derived from cassava, rice, and wheat (Blennow et al. 2005). Technological modifications can include breeding for higher starch yields or direct modification using genetic engineering techniques. For example, BASF Plant Science's Amflora genetically

engineered potato variety set to be grown commercially in Europe starting summer, 2010 (Benner 2010).

1.1.5 Starch granule lipids and membranes

Lipids are components of all living organisms and are involved in a wide variety of biological processes. Aside from their role in compartmentalization and reaction localization, they can act as important signalling molecules and hormones (Munnik and Testerink 2009; Testerink and Munnik 2005; Xue et al. 2009). The lipids in the membranes of plant cells and organelles vary in their composition (Jouhet et al. 2007). Like most successful biological mechanisms, the exact distribution of lipids is well conserved among many plants species and tissues (Jouhet et al. 2007).

Plastid membranes are primarily composed of sulfolipid and galactolipids whereas phospholipids are the main components of extraplastidic membranes (Jouhet et al. 2007)**.** Phosphatidylethanolamine (PE) is completely absent from plastidic membranes (Jouhet et al. 2007). Sterols and glycosphingolipids are present in the plasma membrane where they are involved in membrane domain organization (Jouhet et al. 2007).

Upon phosphate deprivation, digalactosyl diacylglycerol (DGDG), a polar plastid membrane-specific lipid in standard conditions, has been found in the plasma membrane (Debet and Gidley 2006). Although phospholipids are rare in plastidic membranes, they are one of the primary components of starch granule-associated lipids (Bechtel and Wilson 2003) and it is polar lipids with which the wheat starch granule binding protein, puroindoline (PIN), is associated *in vivo* (Feiz et al. 2009).

1.2 Plant proteomics

1.2.1 Mass spectrometry

Mass spectrometry refers to a category of analytical techniques used for the analysis of molecules based on their mass to charge ratio and is integral to modern proteomics. There are several types of mass spectrometers available along with a variety of sample delivery methods. Two of the most popular sample delivery methods include electrospray ionization (ESI; Fenn et al. 1989) and matrix-assisted laser desorption ionization (MALDI; Karas and Bahr 1990). ESI has become routine for compounds which have been first separated by liquid chromatography and remains one of the most common sample delivery and sample ionization techniques used.

Components used for mass selection and detection include ion traps (3D and linear), quadrupoles and time of flight mass analyzers (TOF). For the purpose of this thesis, only ESI, quadrupoles and TOF will be discussed with emphases on the QStar QqTOF triple quadrupole MS/MS.

There are three steps in electrospray ionization: (1) droplet formation, (2) droplet shrinkage and (3) gaseous ion formation. In step one, electrostatic force on the extruded liquid causes it to emerge from the capillary tip as a cone where the liquid breaks into fine droplets. As these droplets evaporate, the gaseous ions pass into the mass spectrometer. A key advantage of ESI is the formation of multiply charged ions for large proteins. Since a mass spectrometer analyzes ions on the basis of their mass to charge ratios rather than just their masses, multiple charges reduce the mass to charge ratio of the protein, bringing high mass proteins within the usable mass range of the mass spectrometer. The ESI system then

delivers the charged ions to the rest of the mass spectrometer for analysis. For mass analysis, the QStar hybrid QqTOF (**Figure 1.2**) consists of three quadrupoles coupled to a TOF mass analyzer. The triple quadrupoles are arranged such that the first quadrupole is responsible for initial ion selection, allowing the researcher to focus on a specific mass range. This is followed by a collision chamber where selected ions can be broken down into smaller ions through collision-induced dissociation (CID), usually with an inert gas such as helium. The third quadrupole allows for a second round of ion selection before feeding the selected ions into the TOF mass analyzer. Quadrupoles are useful for selection and retention of ions whereas TOF allows for very high precision mass analysis.

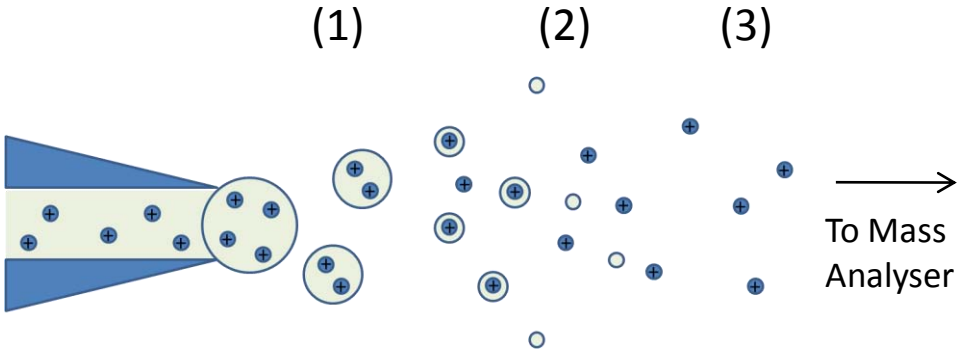
TOF mass spectrometry is based on measuring the time required for ions to travel a set distance to an ion detector. As ions enter the TOF chamber, the ions are accelerated by a magnetic field of predetermined strength. The velocity of the ions moving through the chamber, and thus the time taken, is determined by the ions' mass to charge ratio.

1.2.2 *Peptide Sequencing*

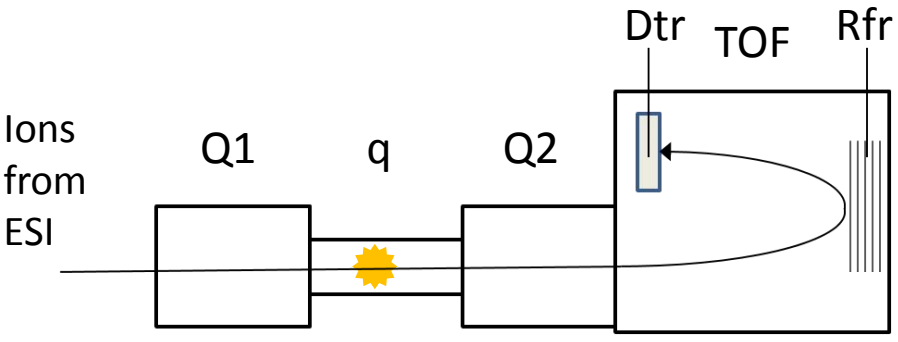
Peptide sequencing is one popular method for determining the sequences of peptides passed through a mass spectrometer. MS/MS of positively charged ions results in a predictable fragmentation pattern produced from cleavage along the peptide's amide bond. The naming convention for these ions refers to b ions as the fragment containing the N-terminus, and y ions for the fragments containing the C-terminus (Roepstorff and Fohlman 1984). The mass difference between fragmented ions compared with the parent mass allows for determination of the amino acid sequence of the peptide in a *de novo* fashion. Based on this amino acid sequence, it is possible to identify the protein that it was generated from, even in the presence

Figure 1.2. Diagrams of ESI and TOF mass spectrometry. ESI (A) involves three basic steps: (1) droplet formation, (2) droplet shrinkage and (3) gaseous ion formation. QqTOF mass spectrometry involves ion selection in the first quadrupole (Q1), collision induced dissociation in the collision cell (q), a non-selecting quadrupole. Selection of fragmented ions is possible with a second quadrupole (Q2). The *m/*z ratio of ions is calculated based on their travel time in the TOF chamber. A reflector (Rfr) increases the path length and improves precision. A detector (Dtr) is used to detect incident ions that have passed through the TOF chamber.

A



B



of many other peptides. Peptide sequencing depends upon available knowledge of the proteome of the organism studied. It is here where the reliance upon genomic databases becomes exceedingly valuable. Using genomic and experimental information, a searchable proteome for an organism can be constructed. Current proteomics databases have thousands of non-redundant proteins from many organisms, many of which are derived from genomic predicted open reading frame (ORF) information. The development of search algorithms which can quickly analyze MS/MS peptide spectra generated from one LC-MS/MS analysis have been developed, with MASCOT being one of the most popular search engines available (Perkins et al. 1999). The search algorithms used by MASCOT can compare experimentally obtained fragmentation spectra to computed theoretical fragmentation spectra derived from the proteomic database. Various scoring methods and 'decoy' database searches are used to ensure quality and reliability of the data. Decoy database searches use protein databases containing only nonsensical proteins, such as a reversed proteome sequence and have become a requirement to properly assess false discovery rates in an experiment (Elias et al. 2005).

1.2.3 Proteomic analysis of grasses

Generally, proteomics of plants has yet to reach the same level of maturity as proteomics for model eukaryotes and prokaryotes. This can be attributed to the fact that unicellular organisms, especially prokaryotes, have relatively simpler proteomes, a consequence of their relatively smaller genomes. Total genomic sequencing of plants has not matched the pace of sequencing projects for animal genomes. Although mass spectrometry is most useful when genomic DNA or EST data of the organism being studied is available, it is still possible to obtain information on large genomes having little or no annotation

(Porubleva et al. 2001). Despite the slower pace, the complete genomes of rice (*Oryza sativa*) (Goff et al. 2002; Yu et al. 2002), soybean (http://www.phytozome.net/soybean) and maize (Schnable et al. 2009) as well as the model plants *Arabidopsis* (Kaul et al. 2000) and *Brachypodium* (Vogel et al. 2010) are now available. The successful completion of the *Brachypodium* genome, for example, will assist in proteomics analysis of not only this model plant, but of proteomics analysis of related organisms with significantly more complex genomes, such as wheat.

1.2.4 The amyloplast proteome

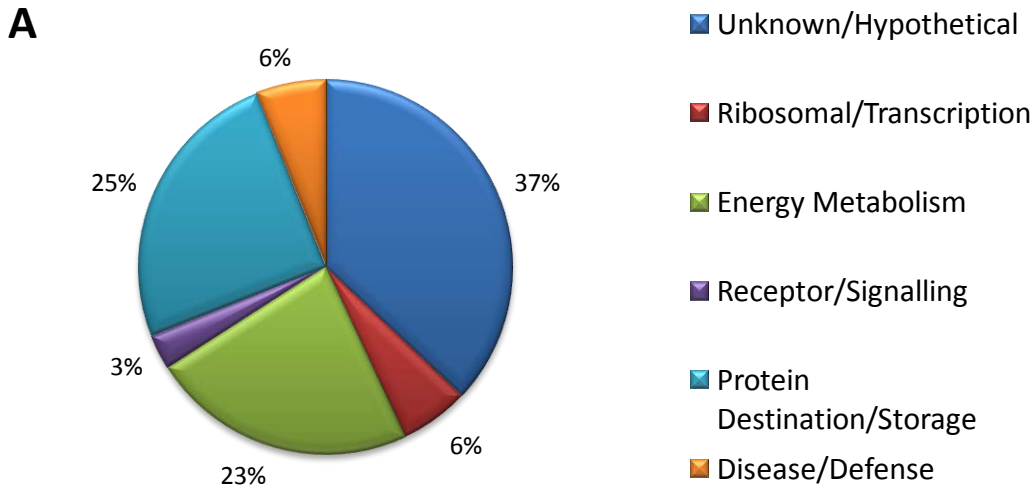
Proteomic studies of specific plant tissues and organelles, such as the chloroplast (Kleffmann et al. 2004; Peltier et al. 2000; Zybailov et al. 2009) and chloroplast membrane (Ferro et al. 2003; Rolland et al. 2003) have been a success. The chloroplast has been studied in significantly more detail than the amyloplast, not only because it is more common in the plant kingdom but also because amyloplasts are difficult to isolate in the laboratory due to fragility of the amyloplast membrane in contrast with the density of the resident starch granules (Denyer and Pike 2008). Though the isolation of whole amyloplasts is challenging, isolation of starch granules is relatively easy due to their density (1.5 g/mL) and insolubility (Dengate 1978). The wheat amyloplast has been looked at (Andon et al. 2002). They discovered that a large fraction of proteins in the amyloplast membrane and whole amyloplast are unknown and still need to be characterized and classified and a large fraction of the amyloplast proteome was dedicated to metabolism (**Figure 1.3**). Though the amyloplast is the center of starch biosynthesis, it was recently discovered that the wheat amyloplast plays a role in at least 18 other metabolic processes of the plant (Dupont 2008). In

addition to whole amyloplast proteomic analysis, proteomics can also be applied to the study of isolated starch granules themselves. During wheat seed maturation, cell division and growth is replaced by a stage characterized by rapid starch synthesis and accumulation inside the amyloplast (Ball and Morell 2003), referred to as grain filling. Amyloplasts are specially developed leucoplasts (Ball and Morell 2003), and it is only within these organelles that the starch granules develop. Several proteomic studies of starch have already been undertaken. Kasarda et al. (2008) analyzed the proteome of commercially-available starch from several manufacturers to test the suitability of the starch for celiac patients, who are sensitive to the seed protein α-gliadin. Andon et al. (2002) explored only the whole amyloplast and amyloplast membrane proteome of immature winter wheat endosperm. No studies have so far analyzed the total accumulated proteome of the reserve starch granules from the desiccated endosperm of cereal plants to assess endogenous and exogenous proteins of incidental and host-targeted pathogens.

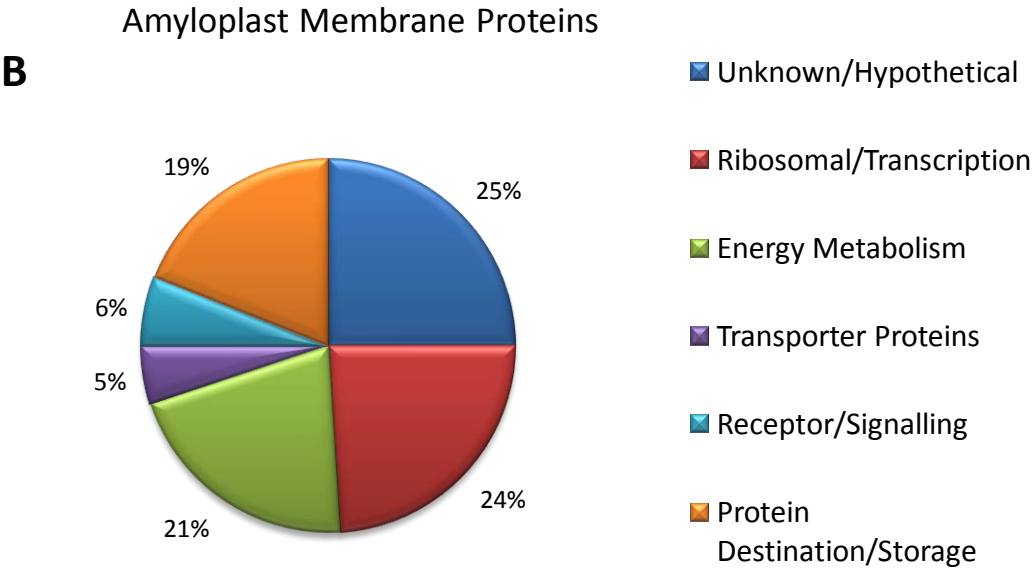
Analysis of the potato tuber amyloplast proteome also shows diversity of metabolic function (Stensballe et al. 2008). Andon et al. (2002) further showed that a significant proportion of the amyloplast proteome is comprised of defense-related proteins. It is clear that characterizing the amyloplast proteome is important for understanding starch granule metabolism as well as antimicrobial defense to preserve bioenergy reserves intact for the subsequent sporophytic generation.

Figure 1.3. Mass spectrometric analysis of the wheat amyloplast membrane and whole proteome. Pie charts show relative abundance of broad protein classes identified by 2D SDS-PAGE and mass spectrometry from the whole amyloplast of developing wheat endosperm (A) or isolated from the amyloplast membrane of developing wheat endosperm (B). A significant proportion of the protein complement is unknown with 37% of whole amyloplast proteins (A) and 25% of amyloplast membrane proteins (B) unnamed and uncharacterized. Reproduced with permission from Andon et al. (2002).

A


Whole Amyloplast Proteins
6%
37%
25%
3%
6%
23%
Unknown/Hypothetical
Ribosomal/Transcription
Energy Metabolism
Receptor/Signalling
Protein Destination/Storage
Disease/Defense

B


Amyloplast Membrane Proteins
19%
25%
6%
5%
24%
21%
Unknown/Hypothetical
Ribosomal/Transcription
Energy Metabolism
Transporter Proteins
Receptor/Signalling
Protein Destination/Storage

1.3 Plant-pathogen interactions

1.3.1 Fungal and bacterial infection of grasses

The high energy density of the starch granule glucose reserves makes it a valuable target for invading pathogens. Plant-pathogen interactions have been increasingly compared to an open warfare, the protein armaments of which comprise only a fraction of the total arsenal available to either side in the conflict (Ferreira et al. 2007). Plants are well equipped with physical, chemical and enzymatic mechanisms for defending against infection and attack from bacteria, fungi (Ferreira et al. 2007) and herbivorous insects (Cheng et al. 1998; Chen 2008). Constitutive and inducible resistance mechanisms exist and the outcome of each battle results in the establishment of either resistance or pathogenesis. Structural mechanisms, such as protective layering of tissues, waxy coatings and cuticle (Russin et al. 1997) and cell wall composition (Kikot et al. 2009) are passive, acting as insurmountable barriers to some pathogens or a temporary hurdle for others. In addition, the caryopsis (grain) of cereals has physical protection in the form of thick and multi-layered bran tissue comprised of pericarp layers, cross cells, tube cells and aleurone layers (Hueckelhoven 2007). As a case in point, the composition and concentration of wax in the outer pericarp layers of grain are factors of *Aspergillus flavus* infection in maize and affect aflatoxin accumulation (Russin et al. 1997). Because of physical defenses, fungal pathogens spend significant time on the grain surface before penetration into the aleurone and subaleurone layers and the energy-rich endosperm tissue (Skadsen and Hohn 2004) (**Figure 1.4**). Though *Fusarium graminearum* infection is associated with shrunken kernels, suggesting that these infectious organisms target the rich energy reserves of the starchy endosperm, little study has been performed in this area. It is

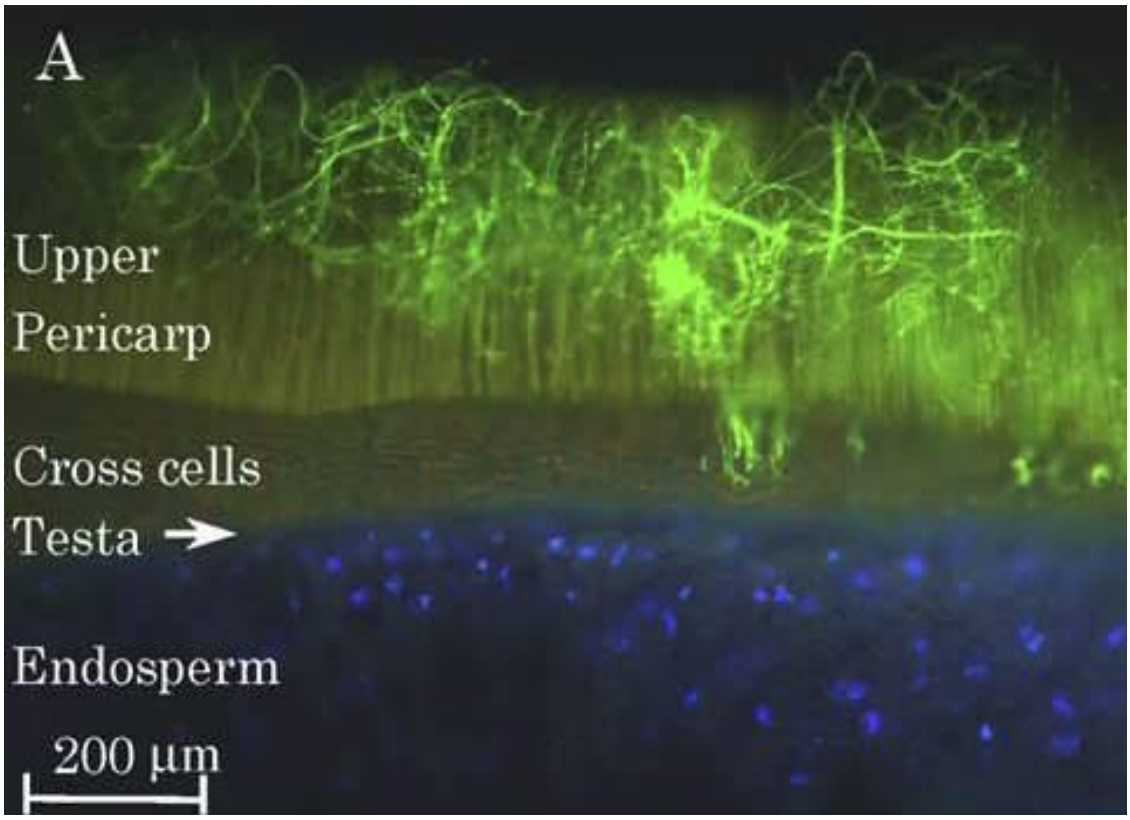
not clear if infection and digestion of endosperm tissue is needed to gain high fungal biomass in situations of widespread infection in the field.

1.3.2 Pathogen infection strategies

A component of early successful fungal pathogen infection in wheat is the secretion of enzymes for the degradation of the host physical barriers, such as the waxy cuticle and polysaccharide-rich cell wall (Phalip et al. 2005; Samuels et al. 2008). In the case of a typical fungal pathogen, such as *Fusarium* spp., secreted enzymes can include lipases (Feng et al. 2005; Jenczmionka and Schafer 2005) cutinases (Kang and Buchenauer 2000; Kang et al. 2005) and pectinases (ten Have et al. 1998; Valette-Collet et al. 2003). In additon, *Fusarium* spp. produce a large amount of cell wall degrading enzymes and toxins (Bottalico and Perrone 2002). Lipases and cutinases are involved in the digestion of the waxy cuticle whereas pectinases modify cell walls to facilitate access of other cell wall-degrading enzymes (e.g. polygalacturonases, endoxylanases and cellulases); (Juge 2006). Studies suggest that cell wall-degrading enzymes produced by fungal pathogens facilitate rapid colonization of wheat spikes before infection of the grain (Wanjiru et al. 2002). Nutrition for the invading pathogen is mostly derived from degraded cytoplasmic proteins, lipids and portions of digested cell wall components (Kikot et al. 2009).

After combating with the physical and chemical barriers of the pericarp and protective cell layers, progression of infection in grain eventually leads to entry of the infecting pathogen into the endosperm tissue (Kikot et al. 2009). Here, proteinases and amylases serve to digest the nutrient-rich stored protein and starch (Schwarz et al. 2001, 2002). Electron microscopy of the endosperm tissue of wheat has shown that after prolonged

Figure 1.4. Fluorescent micrograph of GFP-expressing *Fusarium graminearum* actively infecting wheat pericarp and cross cells. Fungal hyphae become highly concentrated on the grain surface before penetration into the sub pericarp layers or starch granule-rich endosperm. Reproduced with permission from Skadsen and Hohn (2004).


A
Upper
Pericarp
Cross cells
Testa
Endosperm
200 µm

infection with *F. graminearum*, starch granules were extensively degraded and the intracellular protein matrix intercalating the starch granules was completely digested (Jackowiak et al. 2005; Nightingale et al. 1999; Pekkarinen et al. 2000).

The secretion of many pathogenesis-related enzymes by invading pathogens necessary for infection is carefully regulated by a common signalling pathway. For example, the pathogenic stages of filamentous fungi are controlled by mitogen-activated protein (MAP) kinases and it has been shown that disruption of various MAP kinases in *F. graminearum* can reduce extracellular lipolytic, glucolytic, xylanolytic or proteolytic activity (Di Pietro et al. 2001; Hou et al. 2002; Jenczmionka and Schafer 2005; Ruiz-Roldan et al. 2001). In addition to MAP kinase control of secreted digestive enzymes, MAP kinases are involved in adaptation to environmental changes such as pH and osmotic stress (Xu 2000). The MAP kinase Chk1 from the leaf pathogen *Cochliobolus heterostrophus* controls the induction of the cellobiohydrolase CBH7 and the endoglucanase EG6 (Lev and Horwitz 2003) whereas the MAP kinase Hog1 is responsible for adaptation of *F. graminearum* to hyperosmotic stress (Ramamoorthy et al. 2007; Xu 2000).

1.3.3 Pathogenesis-related (PR) proteins in host defense

In conjunction with the physical defenses discussed above, molecular interactions of pathogens with their host are important. Pathogenesis-related (PR) proteins in plants are proteins expressed either locally or systemically in response to pathogens and are necessary for defense against infection (reviewed by Van Loon and Van Strien 1999). Historically, PR proteins were identified by their appearance on protein gels from plant material which had been exposed to a pathogen (van Loon 1985). Now with the development of mass

spectrometry and microarray analysis, there is a greater range of methods available to discover novel PR proteins. The PR proteins are currently categorized into 17 families and proper classification still requires protein expression profiling in addition to any nucleotide sequence information (van Loon and Van Strien 1999). Many PR proteins are directly involved in defense, and can be antipathogenic. For example, the pathogenesis-related protein families PR-2 to PR-5 are largely involved with the degradation of chitin and other glucan polymers of invading pathogens (Nawrath and Metraux 1999). There is significant overlap among the PR proteins and other classes of plant proteins; several classes of known antifungal proteins are also classified as pathogenesis-related or PR-like (van Loon et al. 1994). The PR-like proteins are those homologous to the PR proteins, but are expressed by the host plant under normal physiological conditions (van Loon et al. 1994). Cross-classed PR proteins include members of the defensins (family PR-12), thionins (family PR-13) and some lipid transfer proteins (LTPs; family PR-14). For example, the antimicrobial protein PIN is classified as a PR-like protein due to its constitutive expression during seed maturation even though PIN-a expression can be up-regulated by wounding in developing seeds (Evrard et al. 2007).

1.4 Puroindoline - antimicrobial starch granule surface-associated protein of wheat

1.4.1 Overview of puroindoline biology

The starch granule-associated protein, PIN, is the primary molecular determinant of grain 'texture' (also known as grain hardness) and mutation of the wild-type protein is solely responsible for the emergence of wheat grain useable for baked goods (for a full review, see

Bhave and Morris 2008a; Bhave and Morris 2008b). As a consequence, PIN genetics is partially responsible for the growth and development of human civilization; the technological advancement of bread-making would not have been possible without the discovery and propagation of 'hard' wheat possessing mutant *PIN* alleles, making it millable into more manipulable flour and stable baked forms.

1.4.2 History of puroindoline

The amino acid sequence of PIN was first characterized by Blochet et al. (1993). It was identified from a wheat endosperm protein fraction isolated by triton X-100 phase partitioning, a method developed to isolate integral membrane-bound proteins (Bordier 1981) and since has been used to isolate lipid transfer proteins (Neumann et al. 1994). Primary amino acid sequence analysis revealed the presence of a unique tryptophan-rich domain (TRD) in two isoforms of PIN; PIN-a with the sequence 66-WKWWKWWK-73 and PIN-b with the sequence 68-WPTKWWK-74. PIN-a and PIN-b share 85% cDNA sequence identity and 55% amino acid sequence identity (**Figure 1.5**) (Gautier et al. 1994). Both genes are present at the so-called Hardness (*Ha*) locus of the distal end of chromosome 5D in hexaploid wheat (Doekes and Belderok 1976). The *Ha* locus in wheat includes the *PIN-a, PIN-b*, and *Gsp-1* genes and these genes encode the proteins PIN-a and -b and grain softness protein-1 (GSP-1), respectively. In addition to PIN-a, PIN-b and Gsp-1, the *Ha* locus contains a PseudoPinb, a Pinb-relic, two predicted genes (Gene3 and Gene5) and several transposable elements (Chantret et al. 2005). Tetraploid and diploid wheat do not possess the D genome, though PIN genes are conserved in the A and B genomes of these varieties (Chantret et al. 2005; Gautier et al. 2000).

1.4.3 Puroindoline localization

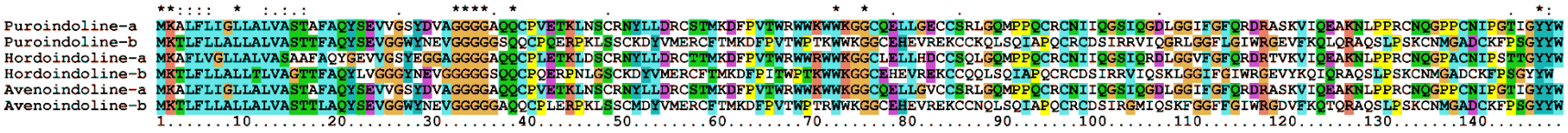
Localization of PIN in wheat indicated that the PIN isoforms, PIN-a and PIN-b, always colocalize in endosperm tissue (Capparelli et al. 2005). Furthermore, the total amount of both PIN-a and PIN-b bound to starch granules *in vivo* is limited by the quantity of PIN-b (Swan et al. 2006) but not PIN-a; PIN-b is the limiting factor for total PIN binding. Ziemann et al. (2008) demonstrated, using a yeast two-hybrid system, that PIN-b interacts with itself *in vitro* and weakly interacts with PIN-a whereas PIN-a did not strongly interact with itself. The impact of PIN protein-protein interactions and aggregations with itself on its antimicrobial activity is unknown.

Though earlier experiments suggested that PIN is expressed in both the endosperm and aleurone layers (Capparelli et al. 2005), further study indicated that PIN is expressed exclusively in the endosperm tissue of developing wheat grain (Wiley et al. 2007). Expression of PIN is generally constitutive, though PIN-a expression can be up-regulated by wounding (Evrard et al. 2007), which may be a response to possible pathogen infection. The structural similarity of PIN to other classes of endosperm proteins, such as the antimicrobial purothionins (Bihan et al. 1996; Petit et al. 1994) and PIN's lipid-binding characteristics, were factors which originally led to the hypothesis that PIN plays a role in antimicrobial defense (Dubreil et al. 1998).

1.4.4 Puroindoline homologs

There are no known homologs of PIN in maize, sorghum or rice (Fabijanski et al. 1988; Tanchak et al., 1998; Gautier et al. 2000; Darlington et al. 2001; Morris 2002).

Figure 1.5. Sequence alignment of puroindoline and its homologs. Isoforms PIN-a and -b from wheat are aligned with the PIN homologs hordoindoline-a and -b from barley (*Hordeum vulgare* L.) and avenoindoline-a and -b from oat (*Avena sativa* L.). Sequences are represented by single letter amino acid code. Amino acids annotated with an asterisk are identical across all six sequences. Amino acids annotated with ':' are highly conserved and amino acids annotated with '.' are weakly conserved. Colours: Orange = GPST, Red = HKR, Blue = FWY, Green = ILMV. Amino acid position is indicated at the bottom of the alignment.


Puroindoline-a
Puroindoline-b
Hordoindoline-a
Hordoindoline-b
Avenoindoline-a
Avenoindoline-b

However, a genomic sequence of 150 bp with similarity to the *Gsp-1* gene (67% nucleotide identity) is present in rice and known as the Ha-rice-relic. Comparative genome analysis shows that a short genomic sequence of 105 bp, with 67% amino acid similarity to *Gsp*-1 gene, is present in an otherwise orthologous rice locus (Caldwell et al. 2004; Chantret et al. 2004, 2005). Homologs of PIN in other cereals include hordoindoline in barley (Rouves et al. 1996) and avenoindoline (also known as tryptophanin) in oat (Tanchak et al. 1998). Although no full PIN homologs are present in the Panicoideae (sorghum and maize) and Ehrhartoideae (rice), they do possess non-functional *Ha*-like gene relics (Charles et al. 2009). It has been hypothesized that the current PIN homologs appeared in the *Ha* locus during or after the divergence of the Pooideae and Ehrhartoideae (Charles et al. 2009). PIN-like relics in related species are not known to localize at starch granule surfaces or affect grain hardness. In addition, they are not known to interact with lipids or create stable protein-lipid foams.

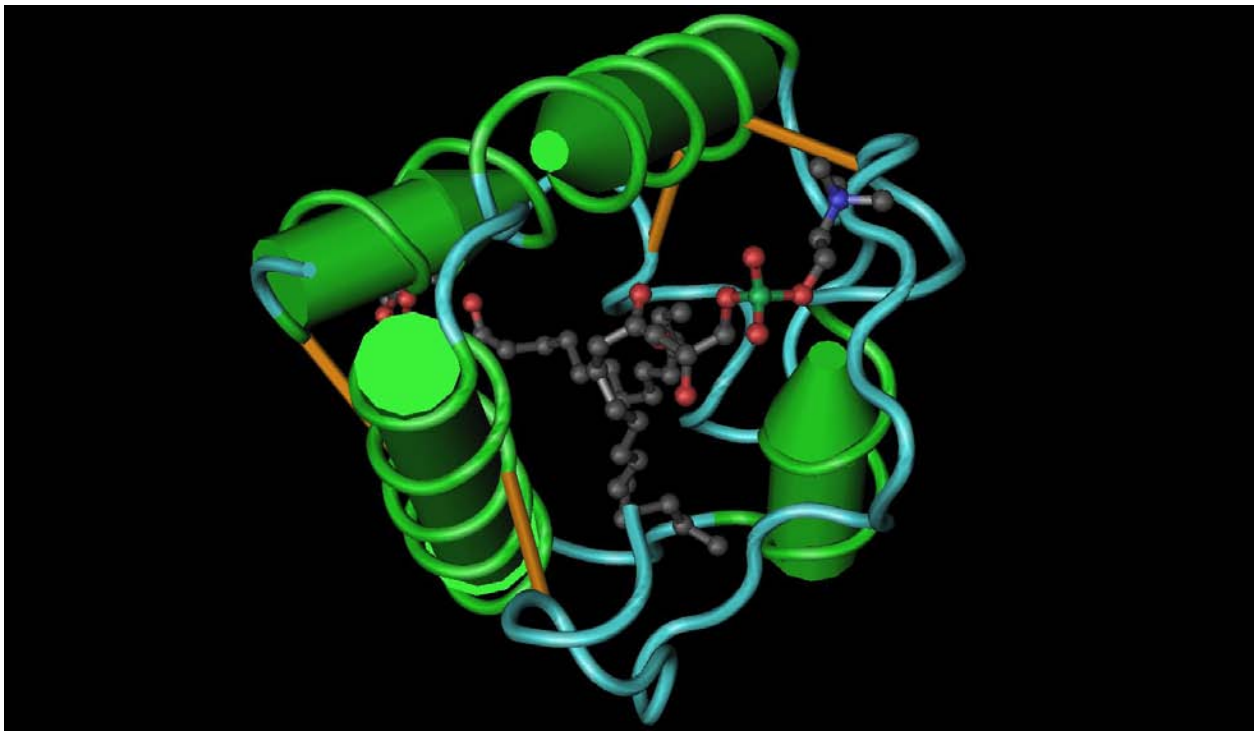
1.4.5 Relationship with lipid binding proteins

Similar to non-specific lipid transfer proteins, PIN binds polar lipids, as demonstrated by its *in vitro* binding to Lysopalmitoylphosphatidylcholine (Wilde et al. 1993), and can participate in the formation of highly stable lipid foams (Dubreil et al. 1997; Le Guerneve et al. 1998; Wilde et al. 1993). It has been suggested that PIN's lipid-binding characteristics were dependent upon PIN's tryptophan-rich domain (Dubreil et al. 1997). Rationale for this came from the fact that a synthetic protein, also possessing a region similar to the tryptophan-rich domain was shown several years previous to have strong foaming properties (Enser et al. 1990). These characteristics lead to the classification of PIN as a lipid binding protein. Lipid binding proteins are found in all organisms and play a role in a wide variety of biological

functions and can be generally placed into two groups based on lipid-interaction characteristics (Marion et al. 2007). Some function as lipid transporters, binding individual lipid molecules inserted into cavities within the protein itself. The nsLTPs function in this way (Douliez et al. 2000). The second category of lipid-binding proteins function by inserting into lipid membranes or by associating with the surface of a lipid membrane (Banaszak et al. 1994). The second functional category applies to PIN, which displays a variety of lipid-binding and lipid foam-stabilizing behaviour (Biswas et al. 2001; Clifton et al. 2007; Dubreil et al. 1997, 2003; Le Guerneve et al. 1998; Sandras et al. 2009), differing it from the nsLTPs. However, the physiochemical characteristics of PIN, including basic pI (pH 10.5), low molecular weight (~13kDa), secondary structure and predicted tertiary structure suggest homology with the nsLTPs (**Figure 1.6**) (Bihan et al. 1996; Marion et al. 2007). It is suggested that the primary role of plant LTPs is the biosynthesis of the cutin layers and waxy surfaces through the transport of acyl monomers (fatty acids, fatty alcohols and hydroxy-fatty acids) (Kolattukudy 1981; Thoma et al. 1993). nsLTPs also likely play a role in defense, supported by the observed *in vitro* antimicrobial properties of wheat nsLTP1 (Cheng et al. 2004; Douliez et al. 2000; Molina et al. 1993).

The first LTP to have its primary amino acid sequence discovered was a phospholipid transfer protein from spinach leaf (Bouillon et al. 1987). Most konwn lipid transfer proteins fall into two structurally related families, LTP1 and LTP2, which have low sequence homology but similar folding properties (Douliez et al. 2000), though recent genomic analysis of rice by Boutrot et al. (2008) has revealed novel lipid transfer proteins which have been placed into new families, bringing the total number of LTP families to ten. The large hydrophobic cavity for binding of lipids and hydrophobic molecules is formed in LTPs by a

Figure 1.6. Three dimensional structure of wheat non-specific lipid transfer protein complexed with two molecules of phospholipid. Analysis of PIN secondary structure suggested that its tertiary structure is similar to the non-specific lipid transfer proteins. Image based on NMR solution structure submitted to the Protein Databank by Charvolin et al. (1999).



helix fold stabilized by four disulphide bonds. The topology of this fold and the topology of the cavity have an influence on the lipid binding properties of members of the LTPs. LTPs can enhance the intermembrane transfer and exchange of bound lipids (Douliez et al. 2000). Though PIN is structurally similar to nsLTP1, PIN, unlike LTPs, interacts directly with lipid aggregates to form highly-stable lipid foams (Wilde et al. 1993). PIN aggregates in the presence of membrane lipids and the exact organization of these aggregates is modified by the structure of the lipids. In the absence of lipids, PIN will self-aggregate (Ziemann et al. 2008).

The antimicrobial activity of lipid binding proteins is generally associated with lipid membrane binding and insertion rather than with binding to individual lipids (Chan et al. 2006). There are several mechanisms of action for antimicrobial proteins and peptides, ranging from lipid membrane destabilization and disintegration to pore formation and cytoplasmic leakage (Shai 2002). The strong affinity of PIN with polar lipids (Dubreil et al. 1997) suggests that it may act as a membranotoxin similar to the thionins.

1.5 Antimicrobial proteins and peptides

1.5.1 Similarity of the puroindoline TRD to antimicrobial peptides

As mentioned, PIN is structurally similar to the non-specific lipid transfer proteins and is antifungal and antibacterial, most likely due to its tryptophan-rich domain. It is also classified as a PR-like protein. Interestingly, the tryptophan-rich domain itself is similar to antimicrobial peptides from non-plant sources, such as indolicidin, an antimicrobial peptide from bovine neutrophils (Selsted et al. 1992). Jing et al. (2003) demonstrated that a 13-

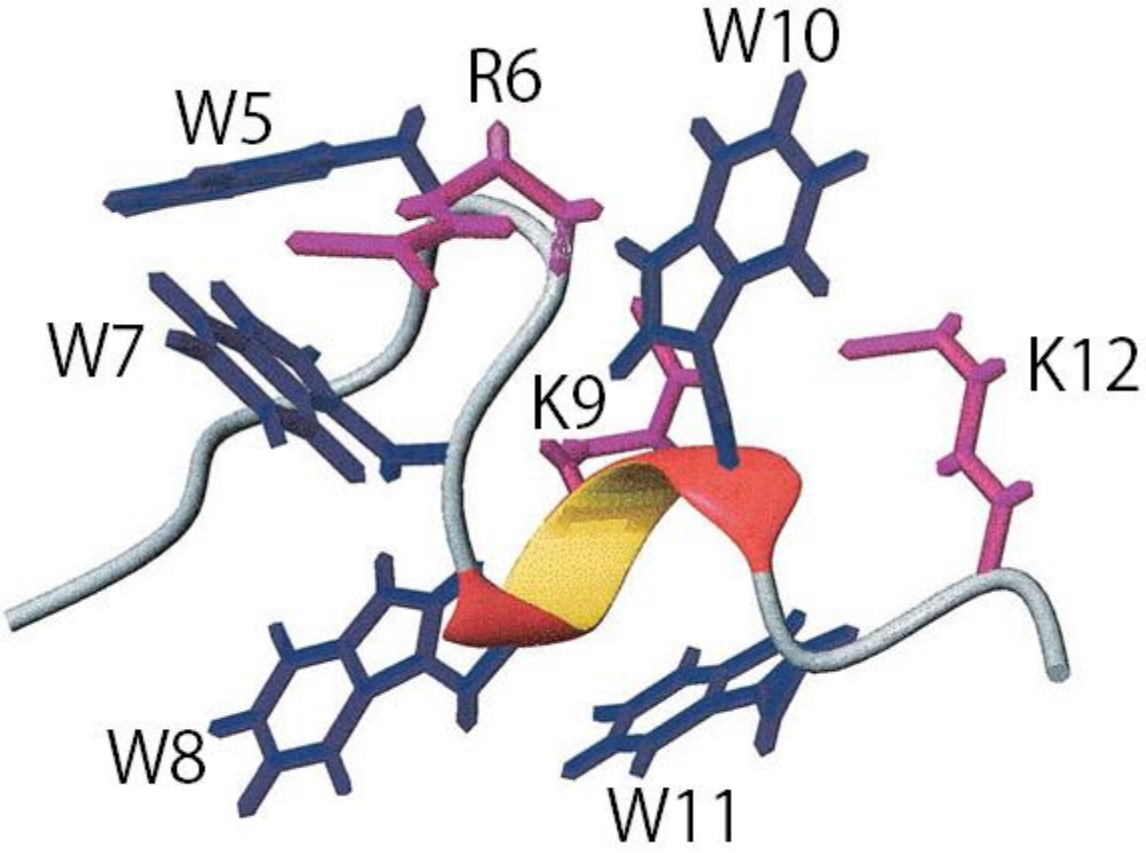
residue peptide named PuroA, corresponding to the tryptophan-rich domain of PIN-a (FPVTWRWWKWWKG-NH2), exhibited bactericidal activity against both *Escherichia coli* and *Staphylococcus aureus* bacteria (**Figure 1.7**). It is notable that some antimicrobial peptides (e.g. indolicidin) and membrane binding proteins (e.g. PIN) have a high tryptophan content that may plays a role in their activity, generally due to tryptophan's stability in interfacial regions of lipid bilayers (Killian et al. 1996; Persson et al. 1998; Yau et al. 1998).

1.5.2 Antimicrobial peptides

The first characterized antibacterial peptides, magainin 1 and 2, were discovered in skin extracts of the African clawed frog *Xenopus laevis* (Zasloff 1987). Since then, many classes of antimicrobial peptides have been found and these peptides show a variety of activities, including antiviral (Andersen et al. 2004), antifungal (De Lucca 2000; Theis and Stahl 2004), antitumor (Papo and Shai 2005), and immunomodulatory (Jerala and Porro 2004; McPhee et al. 2005). The beta-defensins, for example, are involved in all of these functions (Theis and Stahl 2004; Thevissen et al. 1997). Though the exact mechanism of action of antimicrobial peptides is not fully understood for any given peptide, some antimicrobial peptides achieve activity by interacting with lipid membranes or interact with membranes only to gain entry into the target cell (Chan et al. 2006; Shai 2002; Theis and Stahl 2004). Exactly how PIN interacts with lipid membranes or if this interaction is solely responsible for its antimicrobial activity is unknown.

There are various models to describe the interaction of antimicrobial peptides in contact with lipid membranes (Chan et al. 2006). In three models the peptide begins bound to the surface of the lipid membrane due to hydrophobic interactions (Brogden 2005). From this

Figure 1.7. NMR solution structure of PuroA, a 13 amino acid peptide corresponding to the tryptophan-rich domain (TRD) of puroindoline-a. The TRD of PIN and PIN-like proteins is similar to tryptophan-rich antimicrobial proteins and is hypothesized to be the domain responsible for PIN's lipid-binding and antimicrobial properties. Tryptophan (W), arginine (R) and lysine (K) residues are numbered based on their position in the PuroA peptide. Reproduced with permission from Jing et al. (2003).


W5
R6
W10
W7
K9
K12
W8
W11

position, three transitions are possible, depending on the specific properties of the antimicrobial peptide. In one model, the peptides induce micelle formation, leading to the disintegration of the lipid membrane (Oren and Shai 1998). In another model, the peptides aggregate and align perpendicular to the membrane surface, forming stable pores (Oren and Shai 1998). The third model is similar to the second except that transient pore formation is achieved instead of stable pore formation and the lipids of the membrane intercalate with the peptides and participate in inter-peptide interactions (Huang 2004). The transient pore formation resulting from these interactions allows antimicrobial peptides to transfer to the other side of the lipid leaflet, effectively transporting the peptides to the cytoplasm. Two models of antimicrobial peptide action exist which differ from the three models mentioned above. In the first model, cationic peptides associate with the lipid membrane and generate a voltage across the membrane. It is hypothesized that as the voltage increases, the cationic peptides will eventually push through the membrane, resulting in transient pore formation (Miteva et al. 1999; Tieleman 2004). The last model involves binding of amphyphatic peptides to a lipid membrane which than self-associate and cause local increases in membrane curvature. As the association increases, transient pore formation occurs, resulting in transfer of the peptides to the interior lipid surface and overall destabilization of the lipid leaflet (Pokorny and Almeida 2005).

1.5.3 Puroindoline antimicrobial activity

Several studies show that antimicrobial peptides synergistically act with commonly used antibiotics (Giacometti et al. 2005a, 2005b). Cooperation can also been seen among different classes of antimicrobial proteins and peptides and has been observed for the

synergistic antifungal action of PIN with purothionin (Dubreil et al. 1998). In fact, synergistic antimicrobial action between non-specific lipid transfer proteins, of which PIN is closely related, and thionins is known (Molina et al. 1993). PIN-a will cooperatively form pores in giant liposomes with purothionin (Llanos et al. 2006). The similar spacial and temporal localization of PIN, nsLTP1 and purothionin in wheat starchy endosperm suggests that the synergistic effects of these antimicrobial proteins observed *in vitro* may also occur *in vivo* (Dubreil et al. 1998). PIN and its homologs in other cereals is localized at the surface of starch granules where, for wild-type PIN, it strongly binds and can protect the starch granule from pathogen attack (Dubreil et al. 1998).

PIN's binding to the starch granule surface *in vivo* is mainly supported by immunolocalization and its lipid and lipid-binding properties supported by *in vitro* binding to lipids, membranes and stability in lipid foams (Biswas et al. 2001; Clifton et al. 2007; Dubreil et al. 1997; Dubreil et al. 2003; Le Guerneve et al. 1998; Sandras et al. 2009). However, direct evidence of PIN's tryptophan-rich domain binding to the starch granule surface is lacking. In addition, though the effects of PIN expression in transgenic rice increases microbial resistance and grain softness (Krishnamurthy and Giroux 2001; Krishnamurthy et al. 2001), the localization of PIN in these transgenic rice plants has not been published.

1.6 Hypothesis and Objectives

1.6.1 Hypothesis

I hypothesize that the starch granule surface from the desiccated endosperm of mature wheat would contain the protein remnants of failed pathogen infection. Furthermore, I hypothesize that the tryptophan-rich domain of PIN is directly responsible for its binding to starch granules *in vivo* and that PIN is capable of localizing to the surface of starch granules of transgenic plants which do not normally express PIN.

1.6.2 Objectives

1. Assess the starch granule surface proteome of water-washed wheat starch granules from wild-type *PIN* and mutant *PIN* genotypes and identify peptides from pathogen proteins.

2. Assess the binding strength of PIN in wild-type *PIN* wheat and mutant *PIN* wheat and determine the domain of PIN directly bound to the starch granule surface of water-washed starch granules.

3. Localize PIN in transgenic corn expressing PIN to determine whether the effects of PIN expression are due to its localization at the starch granule surface in these plants.

1.7 References

Andersen, J. H., Jenssen, H., Sandvik, K., Gutteberg, T. J. 2004. Anti-HSV activity of lactoferrin and lactoferricin is dependent on the presence of heparan sulphate at the cell surface. J. Med. Virol. 74:262-271.

Andon, N. L., Hollingworth, S., Koller, A., Greenland, A. J., Yates 3rd, J. R., Haynes, P. A. 2002. Proteomic characterization of wheat amyloplasts using identification of proteins by tandem mass spectrometry. Proteomics 2:1156-1168.

Baldwin, P. M., Adler, J., Davies, M. C., Melia, C. D. 1998. High resolution imaging of starch granule surfaces by atomic force microscopy. J. Cereal Sci. 27:255-265.

Ball, S. G., Morell, M. K. 2003. From bacterial glycogen to starch: Understanding the biogenesis of the plant starch granule. Annu. Rev. Plant Biol. 54:207-233.

Banaszak, L., Winter, N., Xu, Z. H., Bernlohr, D. A., Cowan, S., Jones, T. A. 1994. Lipid-binding proteins - a family of fatty-acid and retinoid transport proteins. Adv. Protein Chem. 45:89-151.

Bechtel, D. B., Wilson, J. D. 2003. Amyloplast formation and starch granule development in hard red winter wheat. Cereal Chem. 80:175-183.

Benner, S. 2010. European Commission approves Amflora starch potato. http://www.basf.com/group/corporate/en/function/conversions:/publish/content/news-and-media-relations/news-releases/downloads/2010/P179-Amflora-e.pdf (last viewed 08.07.10).

Bhave, M. and Morris, C. F. 2008a. Molecular genetics of puroindolines and related genes: allelic diversity in wheat and other grasses. Plant Mol. Biol. 66:205-219.

Bhave, M. and Morris, C. F. 2008b. Molecular genetics of puroindolines and related genes: regulation of expression, membrane binding properties and applications. Plant Mol. Biol. 66:221-231.

Bihan, T. L., Blochet, J. E., Desormeaux, A., Marion, D., Pezolet, M. 1996. Determination of the secondary structure and conformation of puroindolines by infrared and Raman spectroscopy. Biochemistry 35:12712-12722.

Biswas, S. C., Dubreil, L., Marion, D. 2001. Interfacial behaviour of wheat puroindolines: monolayers of puroindolines at the air-water interface. Colloid Polym. Sci. 279:607-614.

Blennow, A., Wischmann, B., Houborg, K., Ahmt, T., Jorgensen, K., Engelsen, S. B., Bandsholm, O., Poulsen, P. 2005. Structure function relationships of transgenic starches with engineered phosphate substitution and starch branching. Int. J. Biol. Macromol. 36:159-168.

Blochet, J. E., Chevalier, C., Forest, E., Pebaypeyroula, E., Gautier, M. F., Joudrier, P., Pezolet, M., Marion, D. 1993. Complete amino-acid-sequence of puroindoline, a new basic and cystine-rich protein with a unique tryptophan-rich domain, isolated from wheat endosperm by triton x-114 phase partitioning. FEBS Lett. 329:336-340.

Bordier, C. 1981. Phase-separation of integral membrane-proteins in triton X-114 solution. J. Biol. Chem. 256:1604-1607.

Bottalico, A., Perrone, G. 2002. Toxigenic *Fusarium* species and mycotoxins associated with head blight in small-grain cereals in Europe. Eur. J. Plant Pathol. 108:611-624.

Boutrot, F., Chantret, N., Gautier, M.F. 2008. Genome-wide analysis of the rice and arabidopsis non-specific lipid transfer protein (nsLtp) gene families and identification of wheat nsLtp genes by EST data mining. BMC Genomics 9:86.

Brogden, K. A. 2005. Antimicrobial peptides: Pore formers or metabolic inhibitors in bacteria? Nat. Rev. Microbiol. 3:238-250.

Buleon, A., Colonna, P., Planchot, V., Ball, S. 1998. Starch granules: Structure and biosynthesis. Int. J. Biol. Macromol. 23:85-112.

Caldwell, K. S., Langridge, P., Powell, W. 2004. Comparative sequence analysis of the region harboring the hardness locus in barley and its colinear region in rice. Plant Physiol. 136:3177-3190.

Capparelli, R., Amoroso, M. G., Palumbo, D., Iannaccone, M., Faleri, C., Cresti, M. 2005. Two plant puroindolines colocalize in wheat seed and *in vitro* synergistically fight against pathogens. Plant Mol. Biol. 58:857-867.

Chan, D. I., Prenner, E. J., Vogel, H. J. 2006. Tryptophan- and arginine-rich antimicrobial peptides: Structures and mechanisms of action. Biochim. Biophys. Acta-Biomembr. 1758:1184-1202.

Chantret, N., Cenci, A., Sabot, F., Anderson, O., Dubcovsky, J. 2004. Sequencing of the *Triticum monococcum* Hardness locus reveals good microcolinearity with rice. Mol. Genet. Genomics 271:377-386.

Chantret, N., Salse, J., Sabot, F., Rahman, S., Bellec, A., Laubin, B., Dubois, I., Dossat, C., Sourdille, P., Joudrier, P., Gautier, M. F., Cattolico, L., Beckert, M., Aubourg, S., Weissenbach, J., Caboche, M., Bernard, M., Leroy, P., Chalhoub, B. 2005. Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (*Triticum* and *aegilops*). Plant Cell 17:1033-1045.

Charles, M., Tang, H., Belcram, H., Paterson, A., Gornicki, P., Chalhoub, B. 2009. Sixty million years in evolution of soft grain trait in grasses: Emergence of the softness locus in the

common ancestor of Pooideae and Ehrhartoideae, after their divergence from panicoideae. Mol. Biol. Evol. 26:1651-1661.

Charvolin, D., Douliez, J. P., Marion, D., Cohen-Addad, C., Pebay-Peyroula, E. 1999. The crystal structure of a wheat nonspecific lipid transfer protein (ns-LTP1) complexed with two molecules of phospholipid at 2.1 Angstrom resolution. Eur. J. Biochem. 264:562-568.

Chen, M. 2008. Inducible direct plant defense against insect herbivores: A review. Insect Sci. 15:101-114.

Cheng, C. S., Samuel, D., Liu, Y. J., Shyu, J. C., Lai, S. M., Lin, K. F., Lyu, P. C. 2004. Binding mechanism of nonspecific lipid transfer proteins and their role in plant defense. Biochemistry 43:13628-13636.

Cheng, X. Y., Sardana, R., Kaplan, H., Altosaar, I. 1998. *Agrobacterium*-transformed rice plants expressing synthetic *cryIA(b)* and *cryIA(c)* genes are highly toxic to striped stem borer and yellow stem borer. Proc. Natl. Acad. Sci. U. S. A. 95:2767-2772.

Clifton, L. A., Green, R. J., Frazier, R. A. 2007. Puroindoline-b mutations control the lipid binding interactions in mixed puroindoline-a : puroindoline-b systems. Biochemistry 46:13929-13937.

De Lucca, A. J. 2000. Antifungal peptides: potential candidates for the treatment of fungal infections. Expert Opin. Investig. Drugs 9:273-299.

De Muynck, B., Navarre, C., Boutry, M. 2010. Production of antibodies in plants: status after twenty years. Plant Biotechnol. J. 8:529-563.

Debet, M. R., Gidley, M. J. 2006. Three classes of starch granule swelling: Influence of surface proteins and lipids. Carbohydr. Polym. 64:452-465.

Denyer K., and Pike M. 2008. Isolation of Amyloplasts. in: Current Protocols in Cell Biology 3.28.1-3.28.15. in: Bonifacino J. S., Dasso M., Harford J. B., Lippincott-Schwartz J., andYamada K. M., eds.John Wiley & Sons, Inc., New York.

Di Pietro, A., Garcia-Maceira, F. I., Meglecz, E., Roncero, M. I. G. 2001. A MAP kinase of the vascular wilt fungus *Fusarium oxysporum* is essential for root penetration and pathogenesis. Mol. Microbiol. 39:1140-1152.

Doekes, G. J., Belderok, B. 1976. Kernel hardness and baking quality of wheat - genetic analysis using chromosome substitution lines. Euphytica 25:565-576.

Douliez, J. P., Michon, T., Elmorjani, K., Marion, D. 2000. Structure, biological and technological functions of lipid transfer proteins and indolines, the major lipid binding proteins from cereal kernels. J. Cereal Sci. 32:1-20.

Dubreil, L., Compoint, J. P., Marion, D. 1997. Interaction of puroindolines with wheat flour polar lipids determines their foaming properties. J. Agric. Food Chem. 45:108-116.

Dubreil, L., Gaborit, T., Bouchet, B., Gallant, D. J., Broekaert, W. F., Quillien, L., Marion, D. 1998. Spatial and temporal distribution of the major isoforms of puroindolines (puroindoline-a and puroindoline-b) and non specific lipid transfer protein (ns-LTPle(1)) of *Triticum aestivum* seeds. Relationships with their in vitro antifungal properties. Plant Sci. 138:121-135.

Dubreil, L., Vie, V., Beaufils, S., Marion, D., Renault, A. 2003. Aggregation of puroindoline in phospholipid monolayers spread at the air-liquid interface. Biophys. J. 85:2650-2660.

Dupont, F. M. 2008. Metabolic pathways of the wheat (*Triticum aestivum*) endosperm amyloplast revealed by proteomics. BMC Plant Biology 8:39.

Kikot, G. E., Hours, R. A., Alconada, T. M. 2009. Contribution of cell wall degrading enzymes to pathogenesis of *Fusarium graminearum*: A review. J. Basic Microbiol. 49:231-241.

Elias, J. E., Haas, W., Faherty, B. K., Gygi, S. P. 2005. Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. Nature Methods 2:667-675.

Enser, M., Bloomberg, G. B., Brock, C., Clark, D. C. 1990. *De novo* design and structure-activity-relationships of peptide emulsifiers and foaming agents. Int. J. Biol. Macromol. 12:118-124.

Evrard, A., Meynard, D., Guiderdoni, E., Joudrier, P., Gautier, M. 2007. The promoter of the wheat puroindoline-a gene (PinA) exhibits a more complex pattern of activity than that of the PinB gene and is induced by wounding and pathogen attack in rice. Planta 225:287-300.

Fannon, J. E., Gray, J. A., Gunawan, N., Huber, K. C., BeMiller, J. N. 2003. The channels of starch granules. Food Sci. Biotechnol. 12:700-704.

Fannon, J. E., Hauber, R. J., Bemiller, J. N. 1992. Surface pores of starch granules. Cereal Chem. 69:284-288.

Feiz, L., Wanjugi, H. W., Melnyk, C. W., Altosaar, I., Martin, J. M., Giroux, M. J. 2009. Puroindolines co-localize to the starch granule surface and increase seed bound polar lipid content. J. Cereal Sci. 50:91-98.

Feng, J., Liu, G. S., Selvaraj, G., Hughes, G. R., Wei, Y. D. 2005. A secreted lipase encoded by LIP1 is necessary for efficient use of saturated triglyceride lipids in *Fusarium graminearum*. Microbiology 151:3911-3921.

Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., Whitehouse, C.M. 1989. Electrospray ionization for mass spectrometry of large biomolecules. Science 246:64–71.

Ferreira, R. B., Monteiro, S., Freitas, R., Santos, C. N., Chen, Z., Batista, L. M., Duarte, J., Borges, A., Teixeira, A. R. 2007. The role of plant defence proteins in fungal pathogenesis. Mol. Plant Pathol. 8:677-700.

Ferro, M., Salvi, D., Brugiere, S., Miras, S., Kowalski, S., Louwagie, M., Garin, J., Joyard, J., Rolland, N. 2003. Proteomics of the chloroplast envelope membranes from *Arabidopsis thaliana*. Mol. Cell. Proteomics 2:325-345.

Gautier, M. F., Aleman, M. E., Guirao, A., Marion, D., Joudrier, P. 1994. *Triticum aestivum* puroindolines, 2 basic cystine-rich seed proteins - cDNA sequence-analysis and developmental gene-expression. Plant Mol. Biol. 25:43-57.

Gautier, M. F., Cosson, P., Guirao, A., Alary, R., Joudrier, P. 2000. Puroindoline genes are highly conserved in diploid ancestor wheats and related species but absent in tetraploid *Triticum* species. Plant Sci. 153:81-91.

Giacometti, A., Cirioni, O., Kamysz, W., D'Amato, G., Silvestri, C., Licci, A., Nadolski, P., Riva, A., Lukasiak, J., Scalise, G. 2005. *In vitro* activity of MSI-78 alone and in combination with antibiotics against bacteria responsible for bloodstream infections in neutropenic patients. Int. J. Antimicrob. Agents 26:235-240.

Giacometti, A., Cirioni, O., Kamysz, W., Silvestri, C., Del Prete, M. S., Licci, A., D'Amato, G., Lukasiak, J., Scalise, G. 2005. *In vitro* activity of citropin 1.1 alone and in combination with clinically used antimicrobial agents against *Rhodococcus equi*. J. Antimicrob. Chemother. 56:410-412.

Goff, S. A., Ricke, D., Lan, T. H., Presting, G., Wang, R. L., Dunn, M., et al. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp japonica). Science 296:92-100.

Gomord, W., Sourrouille, C., Fitchette, A. C., Bardor, M., Pagny, S., Lerouge, P., Faye, L. 2004. Production and glycosylation of plant-made pharmaceuticals: the antibodies as a challenge. Plant Biotech. J. 2:83-100.

Han, X., Benmoussa, M., Gray, J. A., BeMiller, J. N., Hamaker, B. R. 2005. Detection of proteins in starch granule channels. Cereal Chem. 82:351-355.

Hou, Z. M., Xue, C. Y., Peng, Y. L., Katan, T., Kistler, H. C., Xu, J. R. 2002. A mitogen-activated protein kinase gene (MGV1) in *Fusarium graminearum* is required for female fertility, heterokaryon formation, and plant infection. Mol. Plant-Microbe Interact. 15:1119-1127.

Huang, H. W., Chen, F. Y., Lee, M. T. 2004. Molecular mechanism of peptide-induced pores in membranes. Phys. Rev. Lett. 92:198304.

Huber, K. C., BeMiller, J. N. 2000. Channels of maize and sorghum starch granules. Carbohydr. Polym. 41:269-276.

Huber, K. C., BeMiller, J. N. 1997. Visualization of channels and cavities of corn and sorghum starch granules. Cereal Chem. 74:537-541.

Hueckelhoven, R. 2007. Cell wall - Associated mechanisms of disease resistance and susceptibility. Annu. Rev. Phytopathol. 45:101-127.

Jackowiak, H., Packa, D., Wiwart, M., Perkowski, J. 2005. Scanning electron microscopy of *Fusarium* damaged kernels of spring wheat. Int. J. Food Microbiol. 98:113-123.

Jenczmionka, N. J., Schafer, W. 2005. The Gpmk1 MAP kinase of *Fusarium graminearum* regulates the induction of specific secreted enzymes. Curr. Genet. 47:29-36.
Jerala, R., Porro, M. 2004. Endotoxin neutralizing peptides. Curr. Top. Med. Chem. 4:1173-1184.

Jing, W. G., Demcoe, A. R., Vogel, H. J. 2003. Conformation of a bactericidal domain of puroindoline a: Structure and mechanism of action of a 13-residue antimicrobial peptide. J. Bacteriol. 185:4938-4947.

Jouhet, J., Marechal, E., Block, M. A. 2007. Glycerolipid transfer for the building of membranes in plant cells. Prog. Lipid Res. 46:37-55.

Juge, N. 2006. Plant protein inhibitors of cell wall degrading enzymes. Trends Plant Sci. 11:359-367.

Kang, Z. S., Buchenauer, H. 2000. Cytology and ultrastructure of the infection of wheat spikes by *Fusarium culmorum*. Mycol. Res. 104:1083-1093.

Kang, Z. S., Zingen-Sell, I., Buchenauer, H. 2005. Infection of wheat spikes by *Fusarium avenaceum* and alterations of cell wall components in the infected tissue. Eur. J. Plant Pathol. 111:19-28.

Karas, M., Bahr, U. 1990. Laser desorption ionization mass-spectrometry of large biomolecules. Trends Analytical Chem. 9:321-325.

Kasarda, D. D., Dupont, F. M., Vensel, W. H., Altenbach, S. B., Lopez, R., Tanaka, C. K., Hurkman, W. J. 2008. Surface-associated proteins of wheat starch granules: Suitability of wheat starch for celiac patients. J. Agric. Food Chem. 56:10292-10302.

Kaul, S., Koo, H. L., Jenkins, J., Rizzo, M., Rooney, T., Tallon, L. J., et al. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796-815.

Killian, J. A., Salemink, I., dePlanque, M. R. R., Lindblom, G., Koeppe, R. E., Greathouse, D. V. 1996. Induction of nonbilayer structures in diacylphosphatidylcholine model membranes by transmembrane alpha-helical peptides: Importance of hydrophobic mismatch and proposed role of tryptophans. Biochemistry 35:1037-1045.

Kleffmann, T., Russenberger, D., von Zychlinski, A., Christopher, W., Sjolander, K., Gruissem, W., Baginsky, S. 2004. The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. Curr. Biol. 14:354-362.

Kolattukudy, P. E. 1981. Structure, biosynthesis, and biodegradation of cutin and suberin. Annu. Rev. Plant Physiol. Plant Mol. Biol. 32:539-567.

Kossmann, J., Lloyd, J. 2000. Understanding and influencing starch biochemistry. Crit. Rev. Plant Sci. 19:171-226.

Krause, G. H., Weis, E. 1991. Chlorophyll fluorescence and photosynthesis - the basics. Annu. Rev. Plant Physiol. Plant Mol. Biol. 42:313-349.

Krishnamurthy, K., Balconi, C., Sherwood, J. E., Giroux, M. J. 2001. Wheat puroindolines enhance fungal disease resistance in transgenic rice. Mol. Plant-Microbe Interact. 14:1255-1260.

Krishnamurthy, K., Giroux, M. J. 2001. Expression of wheat puroindoline genes in transgenic rice enhances grain softness. Nat. Biotechnol. 19:162-166.

Le Guerneve, C., Seigneuret, M., Marion, D. 1998. Interaction of the wheat endosperm lipid-binding protein puroindoline-a with phospholipids. Arch. Biochem. Biophys. 360:179-186.

Lev, S., Horwitz, B. A. 2003. A mitogen-activated protein kinase pathway modulates the expression of two cellulase genes in *Cochliobolus heterostrophus* during plant infection. Plant Cell 15:835-844.

Llanos, P., Henriquez, M., Minic, J., Elmorjani, K., Marion, D., Riquelme, G., Molgo, J., Benoit, E. 2006. Puroindoline-a and alpha 1-purothionin form ion channels in giant liposomes but exert different toxic actions on murine cells. FEBS J. 273:1710-1722.

Marion, D., Bakan, B., Elmorjani, K. 2007. Plant lipid binding proteins: Properties and applications. Biotechnol. Adv. 25:195-197.

McPhee, J. B., Scott, M. G., Hancock, R. E. W. 2005. Design of host defence peptides for antimicrobial and immunity enhancing activities. Comb. Chem. High Throughput Screen. 8:257-272.

Menassa, R., Du, C., Yin, Z. Q., Ma, S., Poussier, P., Brandle, J., Jevnikar, A. M. 2007. Therapeutic effectiveness of orally administered transgenic low-alkaloid tobacco expressing human interleukin-10 in a mouse model of colitis. Plant Biotechnol. J. 5:50-59.

Miteva, M., Andersson, M., Karshikoff, A., Otting, G. 1999. Molecular electroporation: a unifying concept for the description of membrane pore formation by antibacterial peptides, exemplified with NK-lysin. FEBS Lett. 462:155-158.

Molina, A., Segura, A., Garciaolmedo, F. 1993. Lipid transfer proteins (nsLTPs) from barley and maize leaves are potent inhibitors of bacterial and fungal plant-pathogens. FEBS Lett. 316:119-122.

Morell, M. K., Myers, A. M. 2005. Towards the rational design of cereal starches. Curr. Opin. Plant Biol. 8:204-210.

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

Moulia, B., Fournier, M. 2009. The power and control of gravitropic movements in plants: a biomechanical and systems biology view. J. Exp. Bot. 60:461-486.

Munnik, T., Testerink, C. 2009. Plant phospholipid signaling: "in a nutshell". J. Lipid Res. 50:S260-S265.

Murphy, D. J. 2007. Improving containment strategies in biopharming. Plant Biotechnol. J. 5:555-569.

Nawrath, C., Metraux, J. P. 1999. Salicylic acid induction-deficient mutants of *Arabidopsis* express PR-2 and PR-5 and accumulate high levels of camalexin after pathogen inoculation. Plant Cell 11:1393-1404.

Neumann, G. M., Condron, R., Thomas, I., Polya, G. M. 1994. Purification and sequencing of a family of wheat lipid transfer protein homologs phosphorylated by plant calcium-dependent protein-kinase. Biochim. Biophys. Acta - Prot. Struct. Mol. Enzymol. 1209:183-190.

Nightingale, M. J., Marchylo, B. A., Clear, R. M., Dexter, J. E., Preston, K. R. 1999. *Fusarium* head blight: Effect of fungal proteases on wheat storage proteins. Cereal Chem. 76:150-158.

Oksman-Caldentey, K. 2007. Tropane and nicotine alkaloid biosynthesis-novel approaches towards biotechnological production of plant-derived pharmaceuticals. Curr. Pharm. Biotechnol. 8:203-210.

Oren, Z., Shai, Y. 1998. Mode of action of linear amphipathic alpha-helical antimicrobial peptides. Biopolymers 47:451-463.

Papo, N., Shai, Y. 2005. Host defense peptides as new weapons in cancer treatment. Cell Mol. Life Sci. 62:784-790.

Pekkarinen, A., Mannonen, L., Jones, B. L., Niku-Paavola, M. L. 2000. Production of proteases by *Fusarium* species grown on barley grains and in media containing cereal proteins. J. Cereal Sci. 31:253-261.

Peltier, J. B., Friso, G., Kalume, D. E., Roepstorff, P., Nilsson, F., Adamska, I., van Wijk, K. J. 2000. Proteomics of the chloroplast: Systematic identification and targeting analysis of lumenal and peripheral thylakoid proteins. Plant Cell 12:319-341.

Perkins, D. N., Pappin, D. J. C., Creasy, D. M., Cottrell, J. S. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551-3567.

Persson, S., Killian, J. A., Lindblom, G. 1998. Molecular ordering of interfacially localized tryptophan analogs in ester- and ether-lipid bilayers studied by H-2-NMR. Biophys. J. 75:1365-1371.

Petit, M. C., Sodano, P., Marion, D., Ptak, M. 1994. 2-Dimensional H-1-NMR studies of maize lipid-transfer protein - sequence-specific assignment and secondary structure. Eur. J. Biochem. 222:1047-1054.

Phalip, V., Delalande, F., Carapito, C., Goubet, F., Hatsch, D., Leize-Wagner, E., Dupree, P., Van Dorsselaer, A., Jeltsch, J. M. 2005. Diversity of the exoproteome of *Fusarium graminearum* grown on plant cell wall. Curr. Genet. 48:366-379.

Pogue, G. P., Vojdani, F., Palmer, K. E., Hiatt, E., Hume, S., Phelps, J., Long, L., Bohorova, N., Kim, D., Pauly, M., Velasco, J., Whaley, K., Zeitlin, L., Garger, S. J., White, E., Bai, Y., Haydon, H., Bratcher, B. 2010. Production of pharmaceutical-grade recombinant aprotinin and a monoclonal antibody product using plant-based transient expression systems. Plant Biotechnol. J. 8:638-654.

Pokorny, A., Almeida, P. F. F. 2005. Permeabilization of raft-containing lipid vesicles by delta-lysin: A mechanism for cell sensitivity to cytotoxic peptides. Biochemistry 44:9538-9544.

Porubleva, L., Vander Velden, K., Kothari, S., Oliver, D. J., Chitnis, P. R. 2001. The proteome of maize leaves: Use of gene sequences and expressed sequence tag data for identification of proteins with peptide mass fingerprints. Electrophoresis 22:1724-1738.

Ramamoorthy, V., Zhao, X., Snyder, A. K., Xu, J., Shah, D. M. 2007. Two mitogen-activated protein kinase signalling cascades mediate basal resistance to antifungal plant defensins in *Fusarium graminearum*. Cell. Microbiol. 9:1491-1506.

Ramessar, K., Sabalza, M., Capell, T., Christou, P. 2008. Maize plants: An ideal production platform for effective and safe molecular pharming. Plant Sci. 174:409-419.

Roepstorff, P., Fohlman, J. 1984. Proposal for a common nomenclature for sequence ions in mass-spectra of peptides. Biomed. Mass Spectrom. 11:601-601.

Rolland, N., Ferro, M., Seigneurin-Berny, D., Garin, J., Douce, R., Joyard, J. 2003. Proteomics of chloroplast envelope membranes. Photosynth. Res. 78:205-230.

Rouves, S., Boeuf, C., ZwickertMenteur, S., Gautier, M. F., Joudrier, P., Bernard, M., Jestin, L. 1996. Locating supplementary RFLP markers on barley chromosome 7 and synteny with homoeologous wheat group 5. Plant Breeding 115:511-513.

Ruiz-Roldan, M. C., Maier, F. J., Schafer, W. 2001. PTK1, a mitogen-activated-protein kinase gene, is required for conidiation, appressorium formation, and pathogenicity of *Pyrenophora teres* on barley. Mol. Plant-Microbe Interact. 14:116-125.

Russin, J. S., Guo, B. Z., Tubajika, K. M., Brown, R. L., Cleveland, T. E., Widstrom, N. W. 1997. Comparison of kernel wax from corn genotypes resistant or susceptible to *Aspergillus flavus*. Phytopathology 87:529-533.

Saint-Jore-Dupas, C., Faye, L., Gomord, V. 2007. From planta to pharma with glycosylation in the toolbox. Trends Biotechnol. 25:317-323.

Sandras, F., Pezolet, M., Marion, D., Grauby-Heywang, C. 2009. Raman study of the Puroindoline-a/Lysopamitoylphosphatidylcholine interaction in free standing black films. Langmuir 25:8181-8186.

Sasaki, T., Yasui, T., Matsuki, J. 2000. Effect of amylose content on gelatinization, retrogradation, and pasting properties of starches from waxy and nonwaxy wheat and their F1 seeds. Cereal Chem. 77:58-63.

Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., et al. 2009. The B73 maize genome: complexity, diversity, and dynamics. Science 326:1112-1115.

Schnyder, H. 1993. The role of carbohydrate storage and redistribution in the source-sink relations of wheat and barley during grain filling - a review. New Phytol. 123:233-245.

Schwarz, P. B., Jones, B. L., Steffenson, B. J. 2002. Enzymes associated with *Fusarium* infection of barley. J. Am. Soc. Brew. Chem. 60:130-134.

Schwarz, P. B., Schwarz, J. G., Zhou, A., Prom, L. K., Steffenson, B. J. 2001. Effect of *Fusarium graminearum* and *F. poae* infection on barley and malt quality. Mon. schr. Brauwiss. 54:55-63.

Selsted, M. E., Novotny, M. J., Morris, W. L., Tang, Y. Q., Smith, W., Cullor, J. S. 1992. Indolicidin, a novel bactericidal tridecapeptide amide from neutrophils. J. Biol. Chem. 267:4292-4295.

Shai, Y. 2002. Mode of action of membrane active antimicrobial peptides. Biopolymers 66:236-248.

Shama, L. M., Peterson, R. K. D. 2008. Assessing risks of plant-based pharmaceuticals: I. Human dietary exposure. Hum. Ecol. Risk Assess. 14:179-193.

Skadsen, R. W., Hohn, T. A. 2004. Use of *Fusarium graminearum* transformed with *gfp* to follow infection patterns in barley and *Arabidopsis*. Physiol. Mol. Plant Pathol. 64:45-53.

Smith, A. M., Zeeman, S. C., Smith, S. M. 2005. Starch degradation. Annu. Rev. Plant Biol. 56:73-98.

Smith, A. M. 2001. The biosynthesis of starch granules. Biomacromolecules 2:335-341.

Somerville, C. 2006. The billion-ton biofuels vision. Science 312:1277-1277.

Spoek, A. 2007. Molecular farming on the rise - GMO regulators still walking a tightrope. Trends Biotechnol. 25:74-82.

Stensballe, A., Hald, S., Bauw, G., Blennow, A., Welinder, K. G. 2008. The amyloplast proteome of potato tuber. FEBS J. 275:1723-1741.

Streatfield, S. J. 2007. Approaches to achieve high-level heterologous protein production in plants. Plant Biotechnol. J. 5:2-15.

Streb, S., Delatte, T., Umhang, M., Eicke, S., Schorderet, M., Reinhardt, D., Zeeman, S. C. 2008. Starch granule biosynthesis in *Arabidopsis* is abolished by removal of all debranching enzymes but restored by the subsequent removal of an endoamylase. Plant Cell 20:3448-3466.

Swan, C. G., Meyer, F. D., Hogg, A. C., Martin, J. M., Giroux, M. J. 2006. Puroindoline b limits binding of Puroindoline a to starch and grain softness. Crop Sci. 46:1656-1665.

Takagi, H., Saito, S., Yang, L. J., Nagasaka, S., Nishizawa, N., Takaiwa, F. 2005. Oral immunotherapy against a pollen allergy using a seed-based peptide vaccine. Plant Biotechnol. J. 3:521-533.

Tanchak, M. A., Schernthaner, J. P., Giband, M., Altosaar, I. 1998. Tryptophanins: Isolation and molecular characterization of oat cDNA clones encoding proteins structurally related to puroindoline and wheat grain softness proteins. Plant Sci. 137:173-184.

ten Have, A., Mulder, W., Visser, J., van Kan, J. A. L. 1998. The endopolygalacturonase gene Bcpg1 is required for full virulence of *Botrytis cinerea*. Mol. Plant-Microbe Interact. 11:1009-1016.

Testerink, C., Munnik, T. 2005. Phosphatidic acid: a multifunctional stress signaling lipid in plants. Trends Plant Sci. 10:368-375.

Theis, T., Stahl, U. 2004. Antifungal proteins: targets, mechanisms and prospective applications. Cell Mol. Life Sci. 61:437-455.

Thevissen, K., Osborn, R. W., Acland, D. P., Broekaert, W. F. 1997. Specific, high affinity binding sites for an antifungal plant defensin on *Neurospora crassa* hyphae and microsomal membranes. J. Biol. Chem. 272:32176-32181.

Thoma, S., Kaneko, Y., Somerville, C. 1993. A nonspecific lipid transfer protein from *Arabidopsis* is a cell-wall protein. Plant Journal 3:427-436.

Tieleman, D. P. 2004. The molecular basis of electroporation. Biophys. J. 86:371A-372A.

Valette-Collet, O., Cimerman, A., Reignault, P., Levis, C., Boccara, M. 2003. Disruption of *Botrytis cinerea* pectin methylesterase gene *Bcpme1* reduces virulence on several host plants. Mol. Plant-Microbe Interact. 16:360-367.

van Loon, L. C. 1985. Pathogenesis-related proteins. Plant Mol. Biol. 4:111-116.

van Loon, L. C., Pierpoint, W. S., Boller, T., Conejero, V. 1994. Recommendations for naming plant pathogenesis-related proteins. Plant Mol. Biol. Rep. 12:245-264.

Van Loon, L. C., Van Strien, E. A. 1999. The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. Physiol. Mol. Plant Pathol. 55:85-97.

Vogel, J. P., Garvin, D. F., Mockler, T. C., Schmutz, J., Rokhsar, D., Bevan, M. W., International *Brachypodium* Initiative. 2010. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature 463:763-768.

Voncaemmerer, S., Farquhar, G. D. 1981. Some relationships between the biochemistry of photosynthesis and the gas-exchange of leaves. Planta 153:376-387.

Wanjiru, W. M., Kang, Z. S., Buchenauer, H. 2002. Importance of cell wall degrading enzymes produced by *Fusarium graminearum* during infection of wheat heads. Eur. J. Plant Pathol. 108:803-810.

Wilde, P. J., Clark, D. C., Marion, D. 1993. Influence of competitive adsorption of a lysopalmitoylphosphatidylcholine on the functional-properties of puroindoline, a lipid-binding protein isolated from wheat-flour. J. Agric. Food Chem. 41:1570-1576.

Wiley, P. R., Tosi, P., Evrard, A., Lovegrove, A., Jones, H. D., Shewry, P. R. 2007. Promoter analysis and immunolocalisation show that puroindoline genes are exclusively expressed in starchy endosperm cells of wheat grain. Plant Mol. Biol. 64:125-136.

Xu, J. R. 2000. MAP kinases in fungal pathogens. Fungal Genet. Biol. 31:137-152.

Xue, H., Chen, X., Me, Y. 2009. Function and regulation of phospholipid signalling in plants. Biochem. J. 421:145-156.

Yau, W. M., Wimley, W. C., Gawrisch, K., White, S. H. 1998. The preference of tryptophan for membrane interfaces. Biochemistry 37:14713-14718.

Yu, J., Hu, S. N., Wang, J., Wong, G. K. S., Li, S. G., Liu, B., et al. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp indica). Science 296:79-92.

Zasloff, M. 1987. Magainins, a class of antimicrobial peptides from *Xenopus* skin - isolation, characterization of 2 active forms, and partial cDNA sequence of a precursor. Proc. Natl. Acad. Sci. U. S. A. 84:5449-5453.

Zeeman, S. C., Smith, S. M., Smith, A. M. 2007. The diurnal metabolism of leaf starch. Biochem. J. 401:13-28.

Ziemann, M., Ramalingam, A., Bhave, M. 2008. Evidence of physical interactions of puroindoline proteins using the yeast two-hybrid system. Plant Sci. 175:307-311.

Zybailov, B., Friso, G., Kim, J., Rudella, A., Rodriguez, V. R., Asakura, Y., Sun, Q., van Wijk, K. J. 2009. Large scale comparative proteomics of a chloroplast Clp protease mutant reveals folding stress, altered protein homeostasis, and feedback regulation of metabolism. Mol. Cell. Proteomics 8:1789-1810.

Chapter 2

The tryptophan-rich domain of puroindoline is directly associated with the starch granule surface as judged by tryptic shaving and mass spectrometry

Michael L. Wall, Heather L. Wheeler, Matthew P. Huebsch, Jeffrey C. Smith, Daniel Figeys, Illimar Altosaar

2.1 Author contributions

Michael L. Wall was the primary author, wrote and revised the manuscript with the assistance of his supervisor (Dr. Illimar Altosaar), designed and directed all of the experimental work, and performed the experimental work with the assistance of Heather L. Wheeler, a former undergraduate honours project student. Mass spectrometric analysis was performed by Dr. Jeffrey Smith, a postdoctoral fellow in Dr. Figeys's lab at the time, and his Masters student Matthew P. Huebsch now a MSc student in The Smith Lab at Carelton University's Chemistry Department.

2.2 Significance of manuscript

This manuscript is the first example of using a combination of "tryptic shaving" and mass spectrometry to probe the orientation of a protein at the surface of the wheat starch granule. The tryptophan-rich domain of puroindoline was hypothesized to be the region of puroindoline responsible for its direct binding to the starch granule surface and the data presented in this manuscript provides direct evidence to support this hypothesis. The

conclusion that puroindoline was oriented such that the tryptophan-rich domain was buried into the starch granule surface came from the observation that tryptic digestion released a peptide into the aqueous supernatant that corresponded to the tryptophan-rich domain as well as two peptides which were upstream and downstream of the tryptophan-rich domain. Mass spectrometric analysis of the peptides left behind at the starch granule surface following this tryptic digestion revealed only the tryptophan-rich domain but not peptides corresponding to other regions.

2.3 Abstract

The starch granule surface is a frontline of microbial attack and defence, operating in the background of normal starch granule metabolism. Puroindoline, a wheat protein which binds starch granule surfaces, contains a unique tryptophan-rich domain likely responsible for this property, though direct evidence is lacking. To test puroindoline's tight association, prime starch granule extracts were water-washed 8 or 20 times and residual puroindoline removed using a solution of 50% isopropanol/50 mM NaCl. We found that this solvent was consistent in the amount of protein extracted from wheat flour and washed starch, regardless of initial protein content. Relative quantification of puroindoline following water-washing was performed using dot blot. Washing more than 8 times did not further reduce puroindoline content of starch granules suggesting a strong association with the starch granule surface. To identify the tryptophan-rich domain tightly associated with the starch granule surface, a combination of *in situ* tryptic digestion and mass spectrometry was used. Following digestion and water-washing, 50% isopropanol/ 50 mM NaCl was used to remove tightly-associated peptides for identification by mass spectrometry. Using this method, we identified the tryptophan-rich domain of puroindoline directly bound to the starch granule surface of wheat.

2.4 Introduction

The starch granule is a semi-crystalline carbohydrate storage body deposited by a wide variety of photosynthetic organisms, the internal structure and molecular composition of which has been extensively studied (for review see Ball and Morell 2003; Tester et al. 2004). Only photosynthetic eukaryotes and their non- photosynthetic derivatives possess starch, with green algae and the higher plants forming starch granules within plastids. The complex metabolic relationship of the starch granule with its host necessitates a variety of proteins for biosynthesis, degradation, and defence of this important energy deposit (Ball and Morell 2003). Protection of the starch granule is of primary importance to the host organism and research in this area has characterized amylase inhibitors and anti-microbial proteins associated with the starch granule surface as well as the endogenous and exogenous amylases which attempt to degrade it (Smith et al. 2005; Wall et al. 2010; Zeeman et al. 2004). As a consequence of the starch granule's physical characteristics, the surface of the starch granule can be thought of as a frontline of attack and defence, operating in the background of normal starch granule metabolism. Some organisms, including plant pathogens, seek the starch granule for nourishment and have adapted molecular mechanisms to circumvent plant innate defence mechanisms and penetrate these energy reserves. It therefore may prove informative to characterize the starch granule surface and the biologically relevant proteins associated with it, referred to here as starch granule associated proteins (SGAPs; Baldwin 2001).

To facilitate degradation of starch polysaccharides, amylolytic enzymes possess a starch-binding domain (SBD; Machovic and Janecek 2006). The SBD facilitates the association of the enzyme with granular starch, as well as soluble amylose. In addition to direct binding of protein to starch amylose helices, some lipid binding proteins have been

found to associate with the starch granule surface. Puroindoline from wheat and its homologues in other cereals (avenoindolines in oat, hordoindolines in barley) are 13 kDa cysteine-rich lipid associated antimicrobial proteins found to accumulate at the starch granule surface during seed maturation (Bhave and Morris 2008; Blochet et al. 1993; Feiz et al. 2009). In addition to being the primary determinant of grain hardness, both isoforms of puroindoline, puroindoline-a and puroindoline-b (PIN- a and PIN-b respectively), possess antimicrobial activity both *in vitro* and *in vivo* (Dubreil et al. 1998). Genetic experiments with wheat have revealed that expression of wild-type puroindoline recuperates a soft phenotype and that puroindoline-b is a limiting factor in the association of both puroindoline-a and puroindoline- b to starch granules (Beecher et al. 2002; Swan et al. 2006). The structure-function relations responsible for the remarkable antimicrobial activity displayed by puroindoline-like proteins has been the focus of much research ever since tryptophan-rich sequences were first observed in cereal storage protein-encoding cDNA sequences (Fabijanski et al. 1988). *In vitro* experiments using mutants of the puroindoline isoforms, both natural and generated in lab, have aided in identifying the tryptophan-rich domain (TRD) as the region most likely responsible for its lipid binding properties and antimicrobial activity. Puroindoline's anti-microbial and lipid binding characteristics have subsequently been probed *in vitro* via the binding of two synthetic polypeptides representing the puroindoline-a and puroindoline-b tryptophan-rich domains to micelles of sodium dodecyl sulphate (Jing et al. 2003). The tryptophan-rich domain alone behaved independently as an antimicrobial peptide with characteristics similar to those of other well characterized antimicrobial peptides (Andrushchenko et al. 2008). In addition, recombinantly expressed entire PIN-a and PIN-b, with mutations in the tryptophan-rich domain, insert differently into

the intracellular membranes of host yeast cells (Evrard et al. 2008). However, direct evidence linking the tryptophan-rich domain of puroindoline with the starch granule surface is lacking. Therefore, we sought to test whether the TRD of puroindoline is primarily responsible for its binding to the starch granule surface. To do this, we have quantified puroindoline's relative affinity for the starch granule surface from water-washed wheat starch granules and directly probed the starch granule surface for the presence of the TRD using a combination of tryptic shaving and mass spectrometry.

2.5 Materials and Methods

2.5.1 Starch extraction and washing

The kernels of *Triticum turgidum* (L.) subspecies durum and *Triticum aestivum* (L.) cultivars AC Andrew (Pina-D1a/Pinb-D1a), AC Sadash (SWS 349), Bhisaj, Hoffman and AC Barrie (Pina-D1a/Pinb- D1b) were obtained from Agriculture and Agri-Food Canada. Wheat samples were ground to form whole grain meal in a small coffee mill and 2 g aliquots were mixed with small amounts of distilled water to form a dough ball. Starch was extracted from the samples by passing deionized water through the dough while gently squeezing. The flow-through from the dough ball was collected in three aliquots for each sample and centrifuged at 1600 x g for 2 min. The supernatants were discarded and the starch pellets were washed by suspension in 8 ml of distilled deionized water and vortexed for 20 s. The samples were centrifuged as described above and the washes were repeated to a total of 8 or 20 times. One aliquot of each wheat variety was left unwashed.

2.5.2 Extraction of isopropanol/NaCl-soluble protein fraction

Starch pellets were suspended in 12 ml of protein extraction buffer (50% isopropanol (v/v), 50 mM NaCl) per milligram wet weight. Wheat kernel protein extractions were prepared by suspending wheat flour in 24 ml of protein extraction buffer per milligram of dry flour. The samples were mixed at room temperature (23 °C) for 45 min and centrifuged at 2500 x g for 5 min. Insoluble pellets were discarded. Protein concentration of the supernatant was determined by Bradford assay. Aliquots of 200 ml were removed from each supernatant and dried in a speedvac. The dried samples were resuspended in 100 ml of 1x Laemmli SDS-PAGE loading buffer, heated at 100 °C for 3 min and analyzed by SDS-PAGE and Western blot. The remainder of each supernatant was decanted and stored at -20 °C for puroindoline semi-quantification by dot blot analysis.

2.5.3 SDS-PAGE

Samples prepared in Laemmli buffer were loaded onto 1 mm thick 15% SDS-polyacrylamide mini-gels. For silver stained gels, 5 ml of a 1/16 dilution Benchmark Protein Ladder (Invitrogen) was loaded. For gels intended for Western blotting, 5 ml of a biotinylated protein ladder (Cell Signalling Technology) was used. All gels were electrophoresed for 180 V hours.

2.5.4 Western blotting

Proteins were transferred onto nitrocellulose membranes (Bio-Rad) by electrophoresis for 120 V hours using a BioRad semi-dry transblotting system. The membranes were blocked

using 5% (w/v) skim milk powder (Carnation) in PBST (phosphate buffered saline (PBS), 0.05% (v/v) Tween 20 (BioRad)) for 1 h at 23 °C or overnight at 4 °C. Blots were incubated with primary antibody for 1 h (polyclonal rabbit anti-PIN-a 1:1000 dilution in 5% (w/v) skim milk powder/PBST). The membranes were washed three times for 10 min in PBST, followed by incubation in secondary antibody for 1 h (goat anti-rabbit horse radish peroxidase-conjugated, 1:2000 dilution and goat anti-biotin, 1:2000 dilution, in PBST). The washes were repeated and the blots developed using an enhanced chemiluminescent detection system (GE Healthcare) and exposed to HyBlot CL auto-radiography film (Denville Scientific).

2.5.5 Relative quantification of isopropanol/NaCl-soluble PINs

Protein extracts from flour and washed starch samples were dried in a speedvac to a tenth their original volume and aliquots of 6 ml applied directly to a nitrocellulose membrane. The membrane was dried at RT for 1 h and blocked using 5% (w/v) skim milk powder in PBST. Membranes were incubated with Durotest (R-Biopharm Rhone Ltd.) primary antibody for 1 h (mouse anti-friabilin, 1:20 000 dilution in 5% (w/v) skim milk powder/PBST) followed by three washes for 10min each in PBST. Membranes were incubated in secondary antibody for 1 h (goat anti-mouse HRP- conjugated, 1:5000 dilution in PBST). The washes were repeated and the dot-blots developed using an enhanced chemiluminescent detection system (GE Healthcare) and exposed to HyBlot CL auto-radiography film (Denville Scientific). The resulting spots were quantified using the MultiImage™ Light Cabinet gel documentation system and AlphaEase Version 5.00 (Alpha Innotech Corp.).

2.5.6 Identification of granule binding protein domains

Kernels of *T. aestivum* cultivar AC Andrew and *T. turgidum* subspecies durum were rinsed several times with deionized water. Kernels were dried overnight in a laminar flow hood and ground using a separate, clean mortar and pestle for each wheat variety. Starch was extracted and washed 20 times as described in Section 2.5.1. Aliquots of approximately 150 mg were removed from the starch pellets for tryptic shaving. Samples were either set aside for proteolysis or treated with BSA (1.5 mg in deionized water, 37 °C, 1 h). BSA treated samples were centrifuged (18 000 x g, 1 min) and the supernatants discarded. All samples were subjected to enzymatic shaving with trypsin (Promega; 5 mg in 50 mM sodium bicarbonate, 37 °C overnight). Samples were centrifuged (18 000 x g, 1 min) and the supernatant transferred to a fresh tube. The remaining starch pellet was washed three times in deionized water. Peptides remaining on the granule were extracted from the starch pellets as described in Section 2.5.2. All peptides were denatured by heating (70 °C, 5 min) and treated with dithiothreitol (5 mM, 56 °C, 1 h). Denatured proteins were carbamylated by modification with iodoacetamide in the dark (10mM, 23 °C, 1 h). All samples were dried in a speedvac (30 °C, 4 h), resuspended in 40 ml of deionized water, purified using Millipore ZipTips to remove salt and residual starch, and dried again in a speedvac (30 °C, 1 h). Peptides were resuspended through sonication in 40 ml of 0.1% formic acid (J.T. Baker Chemical Co.) for 15 min prior to mass spectrometric analysis.

2.5.7 Chromatography and mass spectrometry

Peptides were injected onto a 7 cm x 200 mm inner diameter trap column, fritted, and packed in-house with 5 cm of 5 mm YMC ODS A reversed phase packing material (Waters, Milford, MA) using an Agilent 1100 HPLC (Agilent, Santa Clara, CA). The peptides were washed for 4 min at 20 mL/min with an aqueous solution containing 5% acetonitrile and 0.1% formic acid. The trap column was connected in series to a 6 cm x 75 mm Picofrit analytical column with a tip opening of 10 mm (New Objective, Woburn, MA), packed with 5 cm of YMC ODS A reversed phase packing material, for mass spectrometric analysis. Starch peptides were eluted from the trap and analytical columns at a flow rate of e 250 nL/min, ionized by nanoelectrospray ionization (ESI) and analyzed using a QSTAR Pulsar QqTOF mass spectrometer (Applied Biosystems, Foster City, CA) operating in information-dependent acquisition mode. Mass analysis included a 1-s survey scan followed by four 3-s tandem mass spectrometric scans on the most intense peaks in the spectrum; masses could be sequenced twice before being added to an exclusion list for 90 s. The HPLC pumped 0.1% formic acid in water with the following percentage gradient of acetonitrile: 0 min: 5%, 3min: 15%, 50min: 40%, 55min: 50%, and 60min: 80%. MS/MS data were searched against a custom database consisting of all wheat proteins from the NCBI nr database using Mascot (Matrixscience Ltd, Boston, MA). Mass tolerances were set to ±100 ppm and ±0.2 Da for the peptide and fragment ion spectra, respectively; up to 2 missed cleavages were considered. Oxidation of methionine residues was selected as a variable modification; no fixed modifications were selected as the 'tryptic shaving' did not involve an alkylation step. Peptides identified by Mascot were manually examined and verified for accuracy.

2.6 Results

2.6.1 Water-washing does not reduce the isopropanol/salt-extractable protein content of starch granules from non-durum wheat

To test the affinity of puroindoline for the starch granule surface, primary starch granule extracts from soft and hard wheat cultivars were water-washed and puroindoline selectively extracted using a solution of 50% isopropanol, 0.05 M NaCl. The protein content of each puroindoline-enriched extract was determined by Bradford assay relative to a BSA standard (**Table 2.1**). For comparison, the same protein fractions from whole flour using the isopropanol/salt solution was also determined. The isopropanol/salt-extractable protein content of whole flour for the soft wheat varieties, reported as percentage dry weight, was 1.28±0.30%, 1.28±0.50% and 1.01±0.04% (AC Andrew, Bhisaj and AC Sadash, respectively). The protein content of similar protein extracts from whole flour of the hard wheat varieties Hoffman and AC Barrie was determined to be 0.74±0.10% and 1.26±0.20%, respectively. In contrast, the isopropanol/salt-extractable protein content of prime starch granules from soft wheat was 0.10±0.01% for AC Andrew, 0.18±0.02% for AC Sadash, and 0.10±0.02% for Bhisaj. The amounts of isopropanol/salt-extractable protein from prime starch granules of hard wheat varieties were 0.10±0.03% for Hoffman and 0.12±0.04% for AC Barrie. Extended water-washing of any non-durum starch granules did not greatly reduce the isopropanol/salt-soluble protein content of starch granules. The largest change was found

Table 2.1. The affinity of isopropanol/salt-soluble proteome for starch granule surfaces is independent of genotype. Percentage of protein from whole flour, prime starch, 8x water-washed starch and 20x water-washed starch, respectively, which is soluble in 50% isopropanol, 50mMNaCl, relative to whole flour protein content as determined by Bradford assay. All samples were conducted in triplicate.

Genotype	Flour	Prime starch	8× washed	20× washed	Literature value	Reference
Durum	0.90 ± 0.05	0.12 ± 0.04	0.06 ± 0.01	0.05 ± 0.01	14.5%	Clarke et al. (2009)
AC Barrie	1.26 ± 0.20	0.10 ± 0.01	0.09 ± 0.01	0.08 ± 0.01	14.2%	McCaig et al. (1996)
Hoffman	0.74 ± 0.10	0.10 ± 0.03	0.08 ± 0.00	0.08 ± 0.03	13.4%	Day (2008)
Bhisaj	1.28 ± 0.50	0.10 ± 0.02	0.07 ± 0.01	0.06 ± 0.02	11.8%	SeCan Technical Bulletin (2009)
AC Andrew	1.28 ± 0.30	0.10 ± 0.01	0.07 ± 0.01	0.05 ± 0.01	11.2%	Sadasivaiah et al. (2004)
AC Sadash	1.01 ± 0.04	0.18 ± 0.02	0.11 ± 0.01	0.07 ± 0.02	10.8%	SeCan Technical Bulletin (2009)

for AC Andrew, the granules of which exhibited a final isopropanol/salt-extractable protein content of 0.05±0.01% dry weight after 20 water-washes.

2.6.2 Immuno-quantification of puroindoline bound to starch granules after water-washing

Isopropanol/salt-soluble proteins from the flour, extracted prime starch granules, and water-washed starch granules of three soft wheat cultivars, two hard wheat cultivars, and durum were separated by SDS-PAGE and the presence of puroindoline verified by Western blotting (**Figure 2.1**). A band at the relative mobility position corresponding to that of puroindoline (~15 kDa) was present for all cultivars except durum. In addition, each cultivar was tested for the presence of puroindoline by immuno-detection (**Figure 2.1, B**). Puroindoline was present in all isopropanol/salt-soluble protein extracts from every cultivar except durum. Therefore, this isopropanol/salt extraction method recovers puroindoline levels that authentically represent the expression of puroindoline among genetic lines. Adjacent lanes were loaded with isopropanol/salt extracts from prime starch granules (**Figure 2.1, Lane 1**), starch granules water-washed 8x (**Lane 2**), or washed 20x (**Lane 3**). To quantify the amount of puroindoline in samples of prime starch, water-washed 8x and water-washed 20x starch granules, a dot-blot method was used (**Figure 2.2**). For the three soft cultivars AC Andrew, AC Sardash and Bhisaj, puroindoline concentration was roughly halved after eight washes but was not reduced further after 20 washes. The amount of puroindoline from starch granules of the hard wheat cultivars Hoffman and AC Barrie were unchanged after water washing and the control puroindoline-null wheat cultivar durum did not result in a signal greater than background levels.

Figure 2.1. Puroindoline remains bound to starch granules after extensive water-washing of granules from soft wheats (var. AC Andrew, AC Sadash [SWS 349], Bhisaj) and non- durum hard wheats (Hoffman, AC Barrie). Durum wheat extracts were included as PIN-null control. Puroindoline-enriched isopropanol/salt extracts of whole flour (A) and starch granules (B) were separated by SDS-PAGE and silver stained. An arrow indicates the 15 kDa molecular weight region, where puroindoline (PIN) is expected to migrate. For each variety, Lane 1 was loaded with the isopropanol/salt extract from the initial batch of sedimented starch granules (prime starch). Lanes 2 and 3 were loaded with isopropanol/salt extract from starch granules washed 8 or 20 times respectively. Corresponding gels were Western blotted and challenged with anti-PIN-a antibody. The PIN signal at 15 kDa is displayed at the bottom of each lane.

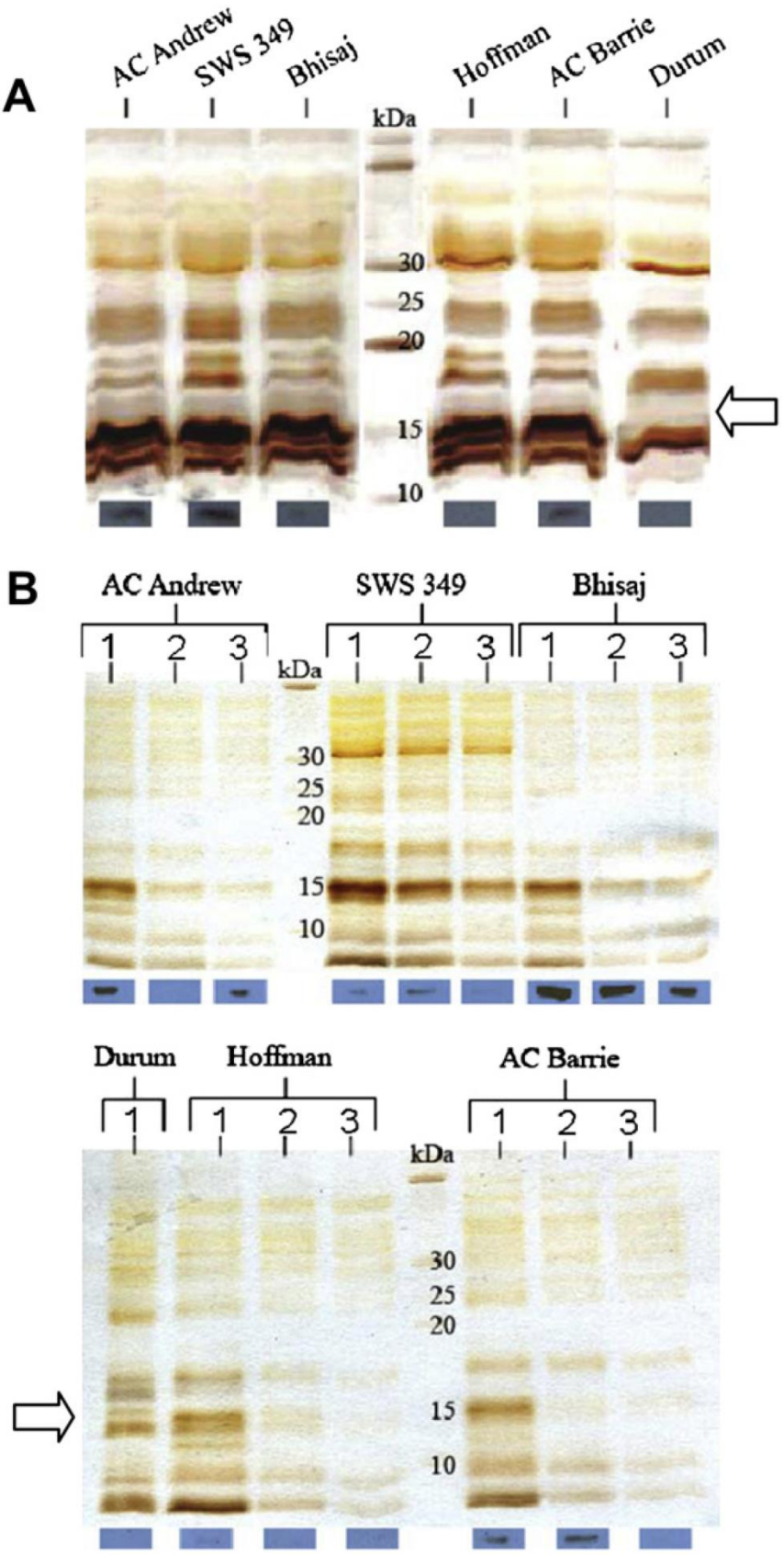
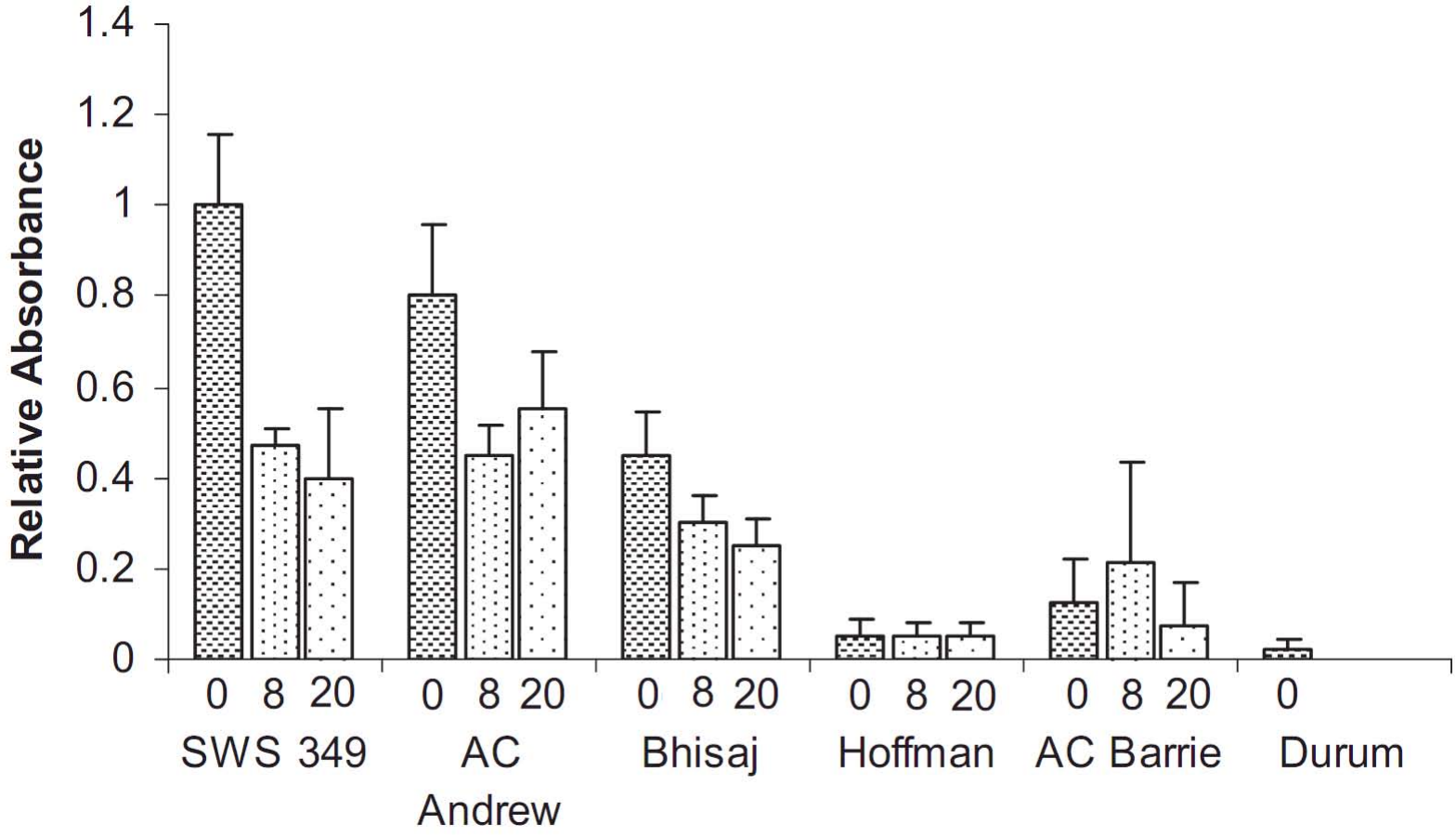

A
AC Andrew
SWS 349
Bhisaj
Hoffman
AC Barrie
Durum
kDa
30
25
20
15
10
B
AC Andrew
SWS 349
Bhisaj
1 2 3
kDa
30
25
20
15
10
Durum
Hoffman
AC Barrie
1
1 2 3
1 2 3
kDa
30
25
20
15
10

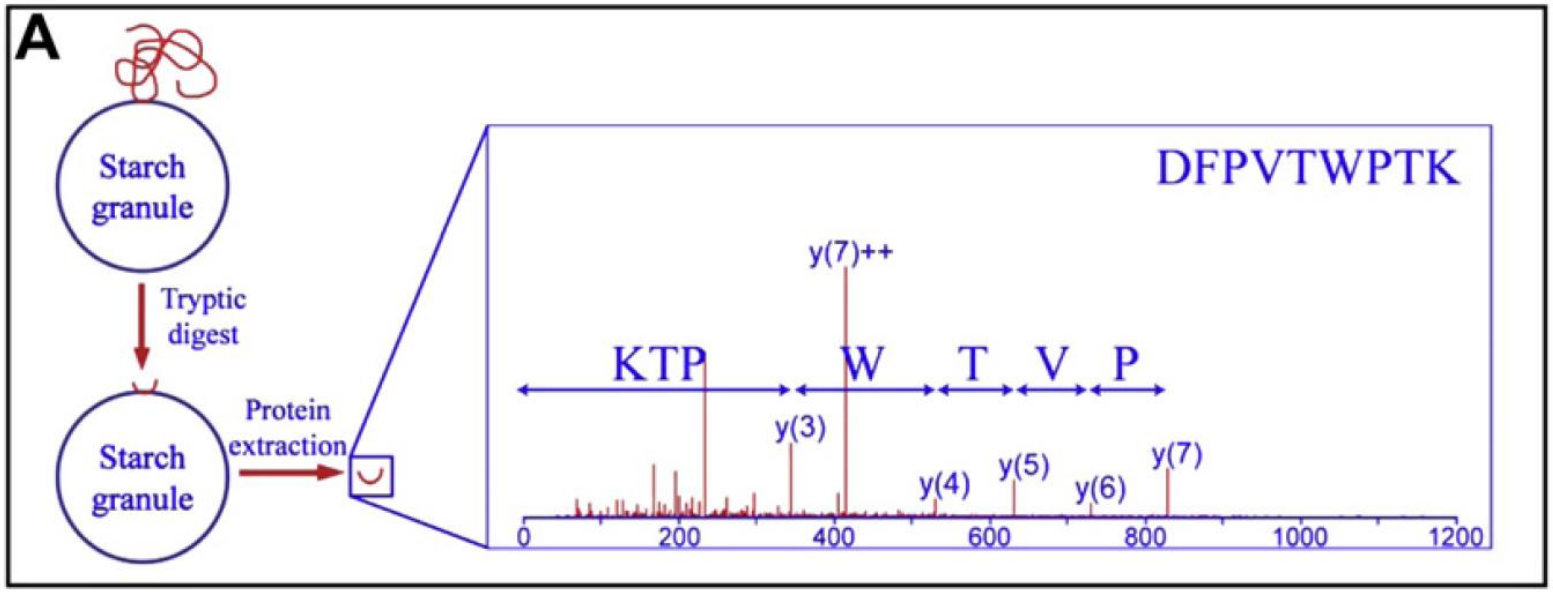
Figure 2.2. Relative abundance of isopropanol/salt-soluble puroindoline at the starch granule surface of six wheat varieties as determined by dot blotting. Prime starch granules (labelled 0), granules washed 8 times, and granules washed 20 times are indicated. Soft wheats: AC Sadash (SWS 349), AC Andrew, Bhisaj; Hard varieties: Hoffman, AC Barrie; Control: unwashed prime granules of durum wheat were extracted for puroindoline quantification.


Relative Absorbance
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0 8 20
SWS 349
0 8 20
AC Andrew
0 8 20
Bhisaj
0 8 20
Hoffman
0 8 20
AC Barrie
0
Durum

2.6.3 The tryptophan-rich domain of puroindoline is present at the starch granule surface

The starch granule-associative properties of puroindoline have been attributed to the presence of a tryptophan-rich domain which may allow puroindoline to bind polar lipids at the starch granule surface. However, evidence identifying those domains of puroindoline which may be responsible for direct binding to the starch granule surface has not been shown. We were therefore interested in identifying whether particular domains of puroindoline had high affinity for the starch granule surface *in vivo*. To probe the starch granule surface and determine whether a particular domain of puroindoline is associated with it, a combination of tryptic proteolysis and mass spectrometry was used. Suspensions of 20x water-washed starch granules of AC Andrew were treated with trypsin to generate peptides for analysis by mass spectrometry and to create conditions in which starch granule associative peptides may remain bound to the granule surface. The supernatant was kept for mass spectrometric analysis and the remaining starch granule suspension saved. The starch granule suspension was washed in water and the isopropanol/salt-extractable protein was obtained. This extract was analysed by mass spectrometry. The analysis revealed that a peptide corresponding to the tryptophan- rich domain of puroindoline-b, residues 65-73, remained on the starch granule surface after the initial tryptic shaving (**Figure 2.3, Panels A, B**). This same peptide of puroindoline-b was also identified in the supernatant after treatment with trypsin. The tryptophan-rich region, 64-KDFPVTWPTKWWKG-77 (PIN-b) contains three trypsin-susceptible residues that permitted such direct detection of this specific peptide fragment by mass spectrometry. In addition, two peptides of puroindoline-a were identified in supernatants after tryptic shaving (RLGQMPPQCRC and DVAGGGGAQQCPVETKL). These peptides were not identified in fractions of starch granule surface peptides.

Figure 2.3. The tryptophan-rich domain of puroindoline remains bound to starch granule surface after proteolysis with trypsin. Panel A. Schematic diagram indicates workflow of the tryptic shaving technique and mass spectrum of a representative peptide from puroindoline-b. Panel B. The identified peptide DFPVTWPTK from (A) corresponded to the tryptophan-rich domain of puroindoline-b (highlighted in red) and is indicated by a boxed region. The homologous region in puroindoline-a is also highlighted in red. The partial amino acid sequences portrayed (residues 49-111, 51-113) are arbitrarily depicted to display the TRD-containing region of puroindolines.



B

```
PIN-a 49 RNYLLDRCSTMKDFPVTWR-WWKWWKGGCQ-ELLGECCSRLGQMPPQCRCNIIQGSIQGDLGGIF 111
PIN-b 51 KDYVMERCFTMKDFPVTW--PTKWWKGGCEHEVREKCCKQLSQIAPQCRCDSIRRVIQGRLGGFL 113
```

2.7 Discussion

2.7.1 Water-washing does not reduce the isopropanol/salt- extractable protein content of starch granules from non-durum wheat

There is no single solvent which can be used to extract and purify all proteins from cereal seeds. The classical Osborne solvents consisting of water, alcohol, salt, and weak acid have been used in varying forms for over 100 years to obtain protein fractions that together constitute the whole range of seed storage proteins (Osborne 1907). More recently, Morris et al. (1994) systematically tested non-Osborne solvent conditions to obtain extracts of puroindoline proteins from wheat starch granule preparations. The combination of 50% isopropanol and 50 mM NaCl was found to be the most effective solvent to extract puroindoline and proved to be more effective than either salt or isopropanol alone. Morris et al. (1994) demonstrated the complete removal of puroindoline proteins from starch granule using this solvent by subsequently extracting with 1% SDS and separating the extracts by SDS-PAGE. Regardless of initial protein content, the 50% isopropanol, 50 mM NaCl solvent consistently extracted ~1% protein, whole flour dry weight, independent of flour total protein content or wheat variety. This was surprising considering that total protein content previously reported among wheat varieties ranges from 8% for soft wheat to >14% for very hard wheats (**Table 2.1**). From this we infer that either the 50% isopropanol, 50 mM NaCl solvent enriches for puroindoline from water-washed wheat starch granules or that puroindoline is the major protein component of water-washed starch granule surfaces and thus makes up a larger portion of the 50% isopropanol, 50 mM NaCl fraction. This would be consistent with

the hypothesis that extraction with 50% isopropanol, 50 mM NaCl enriches for puroindoline and water-washing removes nonpuroindoline proteins from the starch granule surface.

2.7.2 Immuno-quantification of puroindoline bound to starch granules after water-washing

To test the degree of association of puroindoline for the starch granule surface more directly, extensive water washing was employed as in Section 2.7.1 and a dot-blotting technique was used to detect and quantify puroindoline (**Figure 2.2**). The Durotest antibody was selected for its ability to recognize both a- and b-isoforms of puroindoline, allowing for relative quantitation based upon total puroindoline content. After eight washes, the detectable level of puroindoline in isopropanol/salt extracts was half of that determined in extracts from unwashed, initial prime starch granules. After twenty washes, the puroindoline content was not significantly reduced relative to the 8x washed samples. In this case, the amount of puroindoline detected represents starch granule-associated puroindoline but not total puroindoline content. Some puroindoline at the starch granule surface *in situ* may be associated with the protein matrix or free polar lipids which are not covalently bonded to starch granule carbohydrate and are significantly removed during the water-washes, designed to reduce any intracellular proteins or protein bodies. While direct *in vivo* measurement of puroindoline's association with the starch granule surface has yet to be reported, puroindoline's affinity for the polar lipids (Clifton et al. 2007, 2008; Le Guerneve et al. 1998; Wilde et al. 1993) at the starch granule surface (Baldwin 2001; Morrison 1995) may be responsible for its associative properties as well as for the inadequacy of water-washing to fully remove puroindoline from the starch granule surface. In contrast, a previous study used 1 M sodium chloride to extract puroindoline from water-washed starch granules from dry-

sieved flours of wheat varieties Mercia and Riband (Darlington et al. 2000). In that case, granules washed 8 or 20 times did not show any detectable puroindoline but the 1 M sodium chloride may have failed to remove puroindoline from the starch granule surface. That is, 1 M NaCl may only remove the PIN fraction that is associated with the protein matrix or free polar lipids and this PIN fraction is easily removed by water washing. Alternatively, 1 M NaCl could have interfered with the binding of proteins to the nitrocellulose dot membrane. The current results, however, indicate that even in low-puroindoline containing hard wheat granules (**Figure 2.2, Hoffman, AC Barrie**), the 50% isopropanol/50 mM NaCl solution used here was effective in removing proteins in sufficient quantities to permit reproducible immuno-quantification.

2.7.3 The tryptophan-rich domain of puroindoline is present at the starch granule surface

Using SDS-PAGE and dot blotting on prime starch isopropanol/salt extracts, puroindoline was observed to be tightly associated with the starch granule surface despite 20x water washing. We therefore set out to identify which particular domains or regions of puroindoline have higher affinity for the starch granule surface. Tryptic 'shaving' of the prepared granules and subsequent mass spectrometric analysis was performed to identify and compare released versus bound peptides of puroindoline. The isopropanol/salt extracts of trypsin-treated granules theoretically contained peptides of targeting or stabilization domains associated with starch granule lipids, or peptides embedded in the granule surface. Trypsin's cleavage selectivity after arginine and lysine residues was convenient for the study of the tryptophan-rich domain in puroindoline because the TRD is rich in these amino acids (Tanchak et al. 1998). Several putative peptides containing at least some portion of the TRD

would be possible. Our approach for analyzing puroindoline-binding domains did reveal other starch granule surface-associated proteins. New proteins associated with the starch granule are also being identified by other laboratories (Kasarda et al. 2008) and can differ from proteins associated with amyloplast membranes (Andon et al. 2002; Balmer et al. 2006). The positive identification of a peptide specifically matching the TRD of puroindoline-b (**Figure 2.3, A and B**) suggests that this peptide remained at the starch granule surface following tryptic shaving and subsequent washing. The specific interaction of the TRD for the starch granule surface is supported by the finding that the peptides RLGQMPPQCRC and DVAGGGGAQQCPVETKL of puroindoline were identified in supernatants but not in granule-bound extracts. The fact that the TRD of puroindoline-b is clearly hydrophobic and remained with the granule surface after shaving and washing but was extractable with isopropanol/salt confirms that the TRD is responsible for binding of PINs to the starch granule surface. The binding may also be due to puroindoline's affinity to polar lipids at the starch granule surface or its insertional behaviour into membranes during seed development and subsequent association with the granule surface. Supporting this, puroindoline is known to co-localize to the starch granule surface and increase bound lipid content (Feiz et al. 2009). Similarly, in other puroindoline-expressing cereals such as oat. Mohammadi et al. (2007) used fluorescein isothiocyanate (FITC)-labelled anti-tryptophanin to show that the puroindoline homologue, tryptophanin, colocalizes to the starch granule surface during seed development. The precise molecular basis for this transfer of puroindoline from the protein matrix to the starch granule surface during seed maturation and the mechanisms how this occurs to a greater extent in soft varieties requires further study.

2.8 References

Andon, N. L., Hollingworth, S., Koller, A., Greenland, A. J., Yates 3rd, J. R., Haynes, P. A. 2002. Proteomic characterization of wheat amyloplasts using identification of proteins by tandem mass spectrometry. Proteomics 2:1156-1168.

Andrushchenko, V. V., Aarabi, M. H., Nguyen, L. T., Vogel, H. J. 2008. Thermodynamics of the interactions of tryptophan-rich cathelicidin antimicrobial peptides with model and natural membranes. Biochim. Biophys. Acta - Biomembr. 4:1004-1014.

Baldwin, P. M. 2001. Starch granule-associated proteins and polypeptides: a review. Starch-Starke 53:475-503. Ball, S. G., Morell, M. K. 2003. From bacterial glycogen to starch: understanding the biogenesis of the plant starch granule. Annual Rev. Plant Biol. 54:207-233.

Balmer, Y., Vensel, W. H., DuPont, F. M., Buchanan, B. B., Hurkman, W. J. 2006. Proteome of amyloplasts isolated from developing wheat endosperm presents evidence of broad metabolic capability. J. Exp. Bot. 57:1591-1602.

Beecher, B., Bettge, A., Smidansky, E., Giroux, M. J. 2002. Expression of wild-type pinB sequence in transgenic wheat complements a hard phenotype. Theor. Appl. Genet. 105:870-877.

Bhave, M., Morris, C. F., 2008. Molecular genetics of puroindolines and related genes: allelic diversity in wheat and other grasses. Plant Mol. Biol. 66:205-219.

Blochet, J. E., Chevalier, C., Forest, E., Pebaypeyroula, E., Gautier, M. F., Joudrier, P. Pezolet, M., Marion, D. 1993. Complete amino-acid-sequence of puroindoline, a new basic and cystine-rich protein with a unique tryptophan-rich domain, isolated from wheat endosperm by triton X-114 phase partitioning. FEBS Let. 329:336-340.

Clarke, F. R., Clarke, J. M., Pozniak, C. J., Knox, R. E., McCaig, T. N. 2009. Protein concentration inheritance and selection in durum wheat. Can. J. Plant Sci. 89:601-612.

Clifton, L. A., Green, R. J., Frazier, R. A. 2007. Puroindoline-b mutations control the lipid binding interactions in mixed puroindoline-a: puroindoline-b systems. Biochemistry 46:13929-13937.

Clifton, L. A., Green, R. J., Hughes, A. V., Frazier, R. A. 2008. Interfacial structure of wild-type and mutant forms of puroindoline-b bound to DPPG monolayers. J. Phys. Chem. B 112:15907-15913.

Darlington, H. F., Tecsi, L., Harris, N., Griggs, D. L., Cantrell, I. C., Shewry, P. R. 2000. Starch granule associated proteins in barley and wheat. J. Cereal Sci. 32:21-29.

Day, S. 2008. The crops that will power biofuels (ethanol in particular). http://www.umanitoba.ca/afs/agronomists_conf/proceedings/2007/Scott_Day.pdf (last viewed 27.08.09).

Dubreil, L., Gaborit, T., Bouchet, B., Gallant, D. J., Broekaert, W. F., Quillien, L., Marion, D. 1998. Spatial and temporal distribution of the major isoforms of puroindolines (puroindoline-a and puroindoline-b) and non specific lipid transfer protein (ns-LTPle(1)) of Triticum aestivum seeds. Relationships with their in vitro antifungal properties. Plant Sci. 138:121-135.

Evrard, A., Lagarde, V., Joudrier, P., Gautier, M. 2008. Puroindoline-a and puroindoline-b interact with the *Saccharomyces cerevisiae* plasma membrane through different amino acids present in their tryptophan-rich domain. J. Cereal Sci. 48:379-386.

Fabijanski, S., Chang, S. C., Dukiandjiev, S., Bahramian, M. B., Ferrara, P., Altosaar, I. 1988. The nucleotide-sequence of a cDNA for a major prolamin (avenin) in oat (*Avena sativa* L. cultivar Hinoat) which reveals homology with oat globulin. Biochemie und Physiologie der Pflanzen 183:143-152.

Feiz, L., Wanjugi, H. W., Melnyk, C. W., Altosaar, I., Martin, J. M., Giroux, M. J. 2009. Puroindolines co-localize to the starch granule surface and increase seed bound polar lipid content. J. Cereal Sci. 50:91-98.

Jing, W. G., Demcoe, A. R., Vogel, H. J. 2003. Conformation of a bactericidal domain of puroindoline a: structure and mechanism of action of a 13-residue antimicrobial peptide. J. Bacteriol. 185:4938-4947.

Kasarda, D. D., Dupont, F. M., Vensel, W. H., Altenbach, S. B., Lopez, R., Tanaka, C. K., Hurkman, W. J. 2008. Surface-associated proteins of wheat starch granules: suitability of wheat starch for celiac patients. J. Ag. Food Chem. 56:10292-10302.

Le Guerneve, C., Seigneuret, M., Marion, D. 1998. Interaction of the wheat endosperm lipid-binding protein puroindoline-a with phospholipids. Arch. Biochem. Biophys. 360:179-186.

Machovic, M., Janecek, S. 2006. Starch-binding domains in the post-genome era. Cell. Mol. Life Sci. 63:2710-2724.

McCaig, T. N., DePauw, R. M., Clarke, J. M., McLeod, J. G., Fernandez, M. R., Knox, R. E. 1996. AC Barrie hard red spring wheat. Can. J.Plant Sci. 76:337-339.

Mohammadi, M., Zaidi, M. A., Ochalski, A., Tanchak, M., Altosaar, I. 2007. Immunodetection and immunolocalization of tryptophanins in oat (*Avena sativa* L.) seeds. Plant Sci. 172:579-587.

Morris, C. F., Greenblatt, G. A., Bettge, A. D., Malkawi, H. I. 1994. Isolation and characterization of multiple forms of friabilin. J. Cereal Sci. 20:167-174.

Morrison, W. R. 1995. Starch lipids and how they relate to starch granule structure and functionality. Cereal Foods World 40:437-446.

Osborne, T.B. 1907. The proteins of the wheat kernel. Carnegie Institution of Washington.

Sadasivaiah, R. S., Perkovic, S. M., Pearson, D. C., Postman, B., Beres, B. L. 2004. Registration of 'AC Andrew' wheat. Crop Sci. 44:696-697.

SeCan Technical Bulletin. AC® Sadash (SWS 349) Soft White Spring Wheat. http://www.secan.com/en/data/files/_pages/Varieties/files/varieties/876/AC%20Sadash%20Feb%202008.pdf (last viewed 27.08.09).

Smith, A. M., Zeeman, S. C., Smith, S. M. 2005. Starch degradation. Annual Rev. Plant Biol. 56:73-98.

Swan, C. G., Meyer, F. D., Hogg, A. C., Martin, J. M., Giroux, M. J. 2006. Puroindoline b limits binding of puroindoline a to starch and grain softness. Crop Sci. 46:1656-1665.

Tanchak, M. A., Schernthaner, J. P., Giband, M., Altosaar, I. 1998. Tryptophanins: isolation and molecular characterization of oat cDNA clones encoding proteins structurally related to puroindoline and wheat grain softness proteins. Plant Sci. 137:173-184.

Tester, R. F., Karkalas, J., Qi, X. 2004. Starch - composition, fine structure and architecture. J. Cereal Sci. 39:151-165.

Wall, M. L., Wheeler, H. L., Smith, J., Figeys, D., Altosaar, I. 2010. Mass spectrometric analysis reveals remnants of host-pathogen molecular interactions at the starch granule surface in wheat endosperm. Phytopathology 100:848-854.

Wilde, P. J., Clark, D. C., Marion, D. 1993. Influence of competitive adsorption of a lysopalmitoylphosphatidylcholine on the functional-properties of puroindoline, a lipid-binding protein isolated from wheat-flour. J. Ag. Food Chem. 41:1570-1576.

Zeeman, S. C., Smith, S. M., Smith, A. M. 2004. The breakdown of starch in leaves. New Phytol. 163:247-261.

Chapter 3

Mass Spectrometric Analysis Reveals Remnants of Host-Pathogen Molecular Interactions at the Starch Granule Surface in Wheat Endosperm

Michael L. Wall, Heather L. Wheeler, Jeffrey C. Smith, Daniel Figeys, Illimar Altosaar

3.1 Author contributions

Michael L. Wall was the primary author, wrote and revised the manuscript with the assistance of his supervisor (Dr. Illimar Altosaar), designed and directed all of the experimental work, and performed the experimental work with the assistance of Heather L. Wheeler, a former undergraduate honours project student. Mass spectrometric analysis was performed by Dr. Jeffrey Smith, a postdoctoral fellow in Dr. Figeys's lab at the time.

3.2 Significance of manuscript

This manuscript is the first example of mass spectrometry used to probe the starch granule surface for proteins derived from plant pathogens. The starch granule surface is the final line of defense against pathogens seeking the nutrient-rich starch granule interior. By probing the starch granule surface, we have revealed many host proteins which have only been hypothesized from cDNA library information. In addition, most of the proteins identified from the pathogens were unknown. The specificity of this study in analysing proteins at the starch granule surface only increased the chance that identified proteins from both the host and any pathogens are involved in known or novel host-pathogen interactions.

3.3 Abstract

The starch granules of wheat seed are solar energy-driven deposits of fixed carbon and, as such, present themselves as targets of pathogen attack. The seed's array of antimicrobial proteins, peptides, and small molecules comprises a molecular defense against penetrating pathogens. In turn, pathogens exhibit an arsenal of enzymes to facilitate the degradation of the host's endosperm. In this context, the starch granule surface is a relatively unexplored domain in which unique molecular barriers may be deployed to defend against and inhibit the late stages of infection. Therefore, it was compelling to explore the starch granule surface in mature wheat seed, which revealed evidence of host–pathogen molecular interactions that may have occurred during grain development. In this study, starch granules from the soft wheat *Triticum aestivum* cv. AC Andrew and hard wheat *T. turgidum durum* were isolated and water washed 20 times, and their surface proteins were digested *in situ* with trypsin. The peptides liberated into the supernatant and the peptides remaining at the starch granule surface were separately examined. In this way, we demonstrated that the identified proteins have a strong affinity for the starch granule surface. Proteins with known antimicrobial activity were identified, as well as several proteins from the plant pathogens *Agrobacterium tumefaciens*, *Pectobacterium carotovorum*, *Fusarium graminearum*, *Magnaporthe grisea*, *Xanthomonas axonopodis*, and *X. oryzae*. Although most of these peptides corresponded to uncharacterized hypothetical proteins of fungal pathogens, several peptide fragments were identical to cytosolic and membrane proteins of specific microbial pathogens. During development and maturation, wheat seed appeared to have resisted

infection and lysed the pathogens where, upon desiccation, the molecular evidence remained fixed at the starch granule surface.

3.4 Introduction

Wheat, like all plants, captures energy from sunlight and converts it into vitreous amyloplast-derived starch granule deposits composed of amylose and amylopectin (Buleon et al. 1998; Tester et al. 2004). These granules can be transitory, as found in green tissues; but, importantly, starch granules are also synthesized in developing grain endosperm for long-term storage (Buleon et al. 1998). Fungal and bacterial pathogens that attack the developing kernel have adapted to seeking and obtaining vital nutrition from the starchy endosperm, an energy-rich storage tissue meant to fuel germination of the cognate embryo (Pritsch et al. 2000; Zhou et al. 2006). There are constant metabolic interactions between environmental pathogens and the cereal host over the energy stored in the endosperm (Dangl and Jones 2001). Evolution by adaptation suggests that novel, localized defense mechanisms are always in various stages of development and are constantly refined to combat the ever-changing strategies of attacking pathogens. Host resistance includes organelle-specific innate defense mechanisms, and antimicrobial small molecules, peptides, and proteins at starch granule surfaces can be considered the final line of defense against the invading organisms (Dangl and Jones 2001). Therefore, the starch granule surface in mature wheat grain represents a subcellular compartment where one may observe not only the direct interactions of host molecular defense mechanisms but also the pathogen molecular attack mechanisms. Studying such host-pathogen/protein- protein interactions along the “walls of the energy vault” may identify novel resistance mechanisms. The infection success of cereal pathogens depends on

many factors. For example, wheat heads are most susceptible to pathogen attack during anthesis (Bottalico and Perrone 2002; Mentewab et al. 2000). Environmental factors are also important. Although temperature influences the rate and severity of disease caused by *Fusarium* spp. (Brennan et al. 2005; Miller et al. 2004), moisture is the overriding factor determining successful infection by *Fusarium* spp. (Cowger et al. 2009; Lacey et al. 1999). The infection pattern of fungal pathogens of wheat seed usually begins with germination on the inner surfaces of the lemma and palea as well as on the ovary within 12 to 24 h (Zange et al. 2005). The fungus then forms a network of mycelia after 2 h. Following this, the hyphae penetrate inter- and intracellularly, causing severe damage to the infected host tissues. This can be viewed using microscopy techniques (Schwarz 2003). Using scanning electron microscopy, fungal hyphae have been observed within the seed coat as well as within the endosperm (Jackowiak et al. 2005). Bacterial pathogens, such as *Pectobacterium carotovorum* (also known as *Erwinia carotovora*), rely on evasion of plant defense mechanisms and secretion of degradative enzymes to digest plant cell walls (Hugouvieux-Cotte-Pattat et al. 1996; Keen and Tamaki 1986). In the last decade, mass spectrometry has enabled the disentanglement of proteins and factors involved in host response to pathogen infection (Mehta et al. 2008). Paper et al. (2007) used a comparative approach to identify proteins expressed by *Fusarium graminearum* during infection of whole wheat seed. An advantage of mass spectrometry over other popular techniques, such as microarray analysis that only measures mRNA transcript levels, is that mass spectrometry can obtain information on extracellular and organellar protein localization. In addition, unlike microarray analysis, mass spectrometry deals with proteins directly, which is critical in that they are the primary determinants of phenotype. Mass spectroscopic analysis of wheat amyloplasts revealed a

high percentage of proteins involved in both storage and metabolism (Andon et al. 2002; Balmer et al. 2006). This is also true of proteomic studies in barley (Finnie et al. 2002). Considering what is known about pathogen infection of wheat kernels, this area of research as it pertains to the starch granule surface proteome remains largely unexplored (Kasarda et al. 2008). Therefore, the present study aimed to use a proteomics approach to identify proteins tightly associated with the starch granule surface.

3.5 Materials and Methods

3.5.1 Wheat samples

Mature seed of two model genotypes (Wall et al. 2010; Zange et al. 2005) were chosen for this study, one durum with no puroindoline and the other a soft wheat with wild-type PIN alleles. *Triticum turgidum durum* and *T. aestivum* cv. AC Andrew were obtained from the Eastern Cereal and Oilseed Research Centre (Agriculture and Agri-Food Canada, Ottawa), summer harvest year 2006. Both seed lots were from regular bulk plots on the same field with no fungicide treatment. Reagents were purchased from Sigma- Aldrich (St. Louis) unless otherwise noted. The preparation of wheat starch granules for mass spectrometric analysis was performed in triplicate.

3.5.2 Starch granule extraction

Wheat seed samples were sterilized in 0.25% sodium hypochlorite (4% domestic bleach solution) for 10 min, rinsed with distilled water 20 to 30 times, and dried overnight in a laminar flowhood. Seed were ground into a meal using a separate, clean mortar and pestle

for each wheat species. The following procedure was performed in triplicate for each wheat sample: an aliquot of meal (2 g) was mixed with distilled water to form a dough ball. Distilled water was passed through the dough ball and, by gently squeezing the dough ball, a starch-granule-rich flow-through was obtained. To collect the granules, this flow-through was centrifuged at 1,600 × g for 2 min; the supernatant was discarded and the starch pellet was resuspended in distilled water. The resuspension and centrifugation wash step was repeated 20 times for each wheat sample.

3.5.3 Sampling the starch granule surface proteome

For "tryptic shaving", aliquots of water-washed starch granules (~150 mg) were subjected to enzymatic shaving with trypsin (Promega Corp., Madison, WI) (5 µg of trypsin in 50 mM sodium bicarbonate, 37°C overnight). Trypsin-treated granules were centrifuged (18,000 × g, 1 min) and the aqueous supernatant was transferred to a fresh tube. The insoluble starch pellet was washed three times in distilled water. Peptides remaining on the insoluble granules were extracted by suspending the granules in 12 µl of 50% isopropanol (vol/vol), 50 mM NaCl per milligram of wet weight. The granule suspensions were mixed at room temperature for 45 min and centrifuged at 2,500 × g for 5 min to obtain the isopropanol supernatant. Insoluble pellets were discarded. The aqueous and isopropanol supernatants were both dried in a speedvac at 30°C for 4 h (Speed Vac Concentrator model number SVC100H; Savant Instruments, Inc. Hicksville, NY) Peptide pellets were resuspended in 40 µl of distilled water, purified using ZipTips (Millipore, Bedford, MA) to remove salt and residual starch, and dried again in a speedvac (30°C, 1 h). Peptides were resuspended in 40 µl

of 0.1% formic acid, separated by high-performance liquid chromatography (HPLC), and analyzed by electrospray ionization mass spectrometry, as detailed below.

3.5.4 Chromatography and mass spectrometry

Peptides were injected onto a 7-cm by 200-μm inner diameter trap column, fritted, and packed in-house with 5 cm of 5 μm YMC ODS A reversed phase packing material (Waters, Milford, MA) and separated using an Agilent 1100 HPLC (Agilent, Santa Clara, CA). The peptides were washed for 4 min at 20 μl/min with an aqueous solution containing 5% acetonitrile and 0.1% formic acid. The trap column was connected in series to a 6-cm by 75-μm Picofrit analytical column with a tip opening of 10 μm (New Objective, Woburn, MA) and packed with 5 cm of YMC ODS A for mass spectrometric analysis as well as a diverter valve to split the solvent flow prior to the columns. Peptides were eluted from the trap and analytical columns at a flow rate of ~250 nl/min, ionized by nanoelectrospray ionization, and analyzed using a QSTAR Pulsar QqTOF mass spectrometer (Applied Biosystems, Foster City, CA) operating in information-dependent acquisition mode. Mass analysis included a 1-s survey scan followed by four 3-s tandem mass spectrometric scans on the most intense peaks in the spectrum; masses could be sequenced twice before being added to an exclusion list for 90 s. The HPLC pumped 0.1% formic acid in water with the following percentage gradient of acetonitrile: 0 min, 5%; 3 min, 15%; 50 min, 40%; 55 min, 50%; and 60 min, 80%. Tandem mass spectrometric data were searched against a custom database consisting of all proteins from the comprehensive National Center for Biotechnology Information nonredundant (NCBInr) protein database using Mascot (Matrixscience Ltd., Boston, MA). Mass tolerances were set to ±100 ppm and ±0.2 Da for the peptide and fragment ion spectra, respectively; up

to two missed cleavages were considered. Oxidation of methionine residues was selected as a variable modification; no fixed modifications were selected because the tryptic shaving did not involve an alkylation step. Mass spectral significance was determined by searching the data against a "decoy" database containing a copy of the NCBInr database in reverse order and adjusting the ion score cutoff to ≥15 such that the false-positive rate (proportional to the number of peptides matching to reversed sequences in the database) was ≤1% (Higdon et al. 2005). Each match from the database was verified manually; data were deemed acceptable if at least three successive y- or b-ions were present (or y–++ or b–++ ions if the charge state of the peptide was >2+).

3.6 Results

3.6.1 Organization of the data

Wheat starch granules from hard and soft wheat were isolated and water washed 20 times. Trypsin was added directly to the granule suspensions. The peptides liberated from the granule surfaces were collected as aqueous supernatants. To recover the peptide fragments remaining on the surface, the trypsin-treated starch granules were also extracted using 50% isopropanol and 50 mM NaCl, and these second supernatants were collected by centrifugation. All four peptide samples (aqueous supernatant and isopropanol supernatant for both the hard and soft wheat lines) were analyzed by mass spectrometry. The results of the mass spectrometry analysis and Mascot protein identification are presented in five tables based on several criteria. Compiling fragment ion spectra from all four peptide samples (aqueous and isopropanol supernatants of hard and soft wheat starch granule peptides),

Table 3.1 lists all peptides identical to regions of several known plant pathogen proteins, as annotated in the NCBI protein database. To assess the nature of the starch granule surface proteome that may be either distantly tethered or, conversely, tightly bound, the peptides found in the two different supernatants are presented separately in **Tables 3.2 and 3.3**, respectively. Peptides in **Table 3.2** are listed based upon two criteria: (i) those peptides found distal to the starch granule surface and, thus, released into the aqueous supernatant following tryptic digestion; and (ii) reported as being antimicrobial, as annotated in the NCBI protein database. Similar to **Table 3.2**, **Table 3.3** lists peptides annotated by the NCBI protein database as antimicrobial in function but is restricted to peptides found proximal to the starch granule surface. Proximal peptides are defined as those peptides which remain bound to the starch granule surface following tryptic digestion and collected by extraction with 50% isopropanol and 50 mM NaCl. To highlight the fact that there were, from all four samples, peptides identical to regions of hypothetical or uncharacterized proteins, or proteins of unknown function, we have compiled these peptides into **Table 3.4**. Finally, peptides from all four samples identical to regions of plant proteins which are not antimicrobial are listed in **Table 3.5**.

3.6.2 The proteome of the starch granule surface

Several peptides identical to regions of proteins from common plant pathogens were identified on the surfaces of starch granules (**Table 3.1**). The species to which peptide identifications were made include *F. graminearum*, *P. carotovorum*, *Xanthomonas oryzae, Agrobacterium tumefaciens*, *Magnaporthe grisea*, and *X. axonopodis*. The hard wheat contained peptides identical to regions of proteins from all of the above-mentioned microbial

Table 3.1. List of peptides identical to regions of proteins from the plant pathogens *Agrobacterium tumefaciens*, *Fusarium graminearum*, *Pectobacterium carotovorum*, *Magnaporthe grisea*, *Xanthomonas axonopodis*, and *X. oryzae*, generated from mass spectrometric analysis of aqueous supernatants and isopropanol extracts of trypsin-treated hard and soft wheat starch granules.

Reference[a]	Protein[b]	Peptide sequence[c]	
		Hard wheat	Soft wheat
gi|50123425	Bifunctional N-acetylglucosamine-1-phosphate uridyltransferase [*P. carotovorum*]	(K)LDDPAGYGRIVR(E)	…
gi|39941602	Hypothetical protein MG5712.4 [*M. grisea*]	(R)DRVADAWPVFCK(Y)	…
gi|46121515	Hypothetical protein FG05136.1 [*F. graminearum*]	(R)ALGKSKNAR(T)	…
gi|84623867	Hypothetical protein XOO_2210 [*X. oryzae*]	(R)ATLANRKR(V)	…
gi|39977703	Hypothetical protein MG06736.4 [*M. grisea*]	(R)VRAEMVAKMDEK(G)	…
gi|50121895	Putative branched-chain amino acid aminotransferase [*P. carotovorum*]	(K)VVDTAAKKR(G)	…
gi|39942636	Hypothetical protein MG03398.4 [*M. grisea*]	(K)ASRVRDSQRASR(A)	…
gi|39940010	Hypothetical protein MG05235.4 [*M. grisea*]	(K)IAMSELAR(R)	…
gi|15891657	Hypothetical protein AGR_L_3085 [*A. tumefaciens*]	(K)ACPTAEMRGVALGR(F)	…
gi|15891318	Hypothetical protein AGR_L_2410 [*A. tumefaciens*]	(K)KPTAQEIAITR(G)	…
gi|15891555	Hypothetical protein AGR_L_2889 [*A. tumefaciens*]	(R)VVVMQDGR(I)	…
gi|16119407	Hypothetical protein AGR_pAT_256 [*A. tumefaciens*]	(R)LSITPAAVSRNVAMLER(N)	…
gi|21243478	D-amino acid dehydrogenase subunit [*X. axonopodis*]	(R)AREVVIATGPWSPALAAQLGLR(L)	…
gi|21242862	Hypothetical protein XAC2127 [*X. axonopodis*]	…	(R)MRKTLLAR(H)
gi|21245076	Oxidoreductase [*X. axonopodis*]	…	(M)NIPVPVVR(L)

[a] GenBank GenInfo Identifier (gi) number for the most likely protein candidate for which peptides were identified.

[b] Species from which each protein derives is indicated in square brackets.

[c] Where peptides were detected as having an oxidized methionine, the residues have been underlined. Amino acids in parentheses indicate residues expected to flank each identified peptide as determined from the corresponding protein's amino acid sequence. N-terminal peptides are indicated by (–) at the N-terminus.

species, whereas the soft wheat contained peptides identical to regions of proteins of only one plant pathogen, *X. axonopodis*. The majority of the proteins listed in **Table 3.1** are hypothetical proteins derived from genomic open reading frame (ORF) sequences in the NCBI nucleotide database. However, several proteins listed in **Table 3.1** have been previously characterized and are known to localize to the cytosol, as annotated in the NCBI protein database. For example, one 14-mer peptide was identical to a region of a bifunctional N-acetylglucosamine-1-phosphate uridyltransferase, an enzyme involved in bacterial cell wall synthesis (GenBank GenInfo Identifier no. gi|50123425).

Among the peptides identical to regions of plant proteins, several proteins belong to classes that have antimicrobial activity, as annotated in the NCBI protein database. These proteins are organized into **Table 3.2** (aqueous supernatant distal peptides) and **Table 3.3** (isopropanol supernatant proximal peptides). Seven unique antimicrobial proteins were identified in the aqueous supernatant of trypsin-treated hard wheat starch granules as opposed to nine in the aqueous supernatant of trypsin-treated soft wheat starch granules (**Table 3.2**). Although most were identified by a single peptide, a few peptides corresponded to regions of the same protein. For example, 10 peptides, each identical to a region of the ribosome-inactivating protein tritin (gi|391929), were found in both the aqueous supernatant (**Table 3.2**) and isopropanol supernatant (**Table 3.3**) of the hard wheat sample. From soft wheat, only one peptide of tritin was found in the aqueous supernatant whereas four peptides of tritin were found in the isopropanol supernatant.

From all four samples, several peptides were identical to regions of hypothetical plant proteins (genomic ORF DNA sequence only) or plant proteins of unknown function (protein

Table 3.2. List of peptides identical to regions of known antimicrobial proteins, as annotated in the National Center for Biotechnology Information database. Generated from mass spectrometric analysis of aqueous supernatants of trypsin-treated hard and soft wheat starch granules.

Reference[a]	Protein[b]	Peptide sequence[c]	
		Hard wheat	Soft wheat
gi|391929	Tritin [*Triticum aestivum*]	(K)LTNVALGR(Q)	…
	…	(R)ADNLYWEGFK(S)	…
	…	(K)TSPASTGLTLATR(A)	…
	…	(R)FQTVSGFVAGVLEPK(E)	…
	…	(K)AQVNGWQDLSEALLK(T)	…
	…	(R)EAVTTLLLMVHEATR(F)	…
	…	(R)DLLGDTDKLTNVALGR(Q)	…
	…	(K)SSDGTWWELTPGLIPGATHVGFGGTYR(D)	(R)DLLGDTDKLTNVALGR(Q)
gi|5669008	Secretory protein [*T. aestivum*]	(K)VLSDASAFIWK(T)	…
	…	(K)SGYTDDFFAQILGK(N)	(K)SGYTDDFFAQILGK(N)
gi|62465514	Class II chitinase [*T. aestivum*]	(R)AIGVDLLSNPDLVATDPTVSFK(T)	(R)AIGVDLLSNPDLVATDPTVSFK(T)
gi|22001285	Peroxidase 1 [*T. aestivum*]	(K)DIGLAAGLLR(L)	(K)DIGLAAGLLR(L)
	…	(R)DSVVVSGGPDYR(V)	…
	…	(R)EGLFVSDQDLFTNDITRPIVER(F)	(R)EGLFVSDQDLFTNDITRPIVER(F)
gi|132577	Protein synthesis inhibitor I (Ribosome-inactivating protein I) [*Hordeum vulgare*]	(K)LTNVALGR(Q)	…
	…	(R)FQTVSGFVAGLLHPK(A)	…
		(R)DLLGDTDKLTNVALGR(Q)	…
gi|55669876	Chain X, Crystal structure of xylanase inhibitor I [*T. aestivum*]	(–)LPVLAPVTK(D)	…
gi|20804336	Xylanase inhibitor protein I [*T. aestivum*]	…	(R)ALATGIFER(A)
gi|116344	Endochitinase precursor (MF1 antigen) [*Brugia malayi*]	…	(K)SAQKTER(F)
gi|2894148	Monomeric α-amylase inhibitor [*T. aestivum*]	…	(K)VPIPNPSGDR(A)
gi|1709917	Puroindoline-B precursor [*T. aestivum*]	…	(K)DFPVTWPTK(W)

[a] GenBank GenInfo Identifier (gi) number for the most likely protein candidate for which peptides were identified.
[b] Species from which each protein derives is indicated in square brackets.
[c] Amino acids in parentheses indicate residues expected to flank each identified peptide as determined from the corresponding protein's amino acid sequence. N-terminal peptides are indicated by (–) at the N-terminus.

Table 3.3. List of peptides identical to regions of known wheat antimicrobial proteins, generated from mass spectrometric analysis of isopropanol extracts of trypsin-treated hard and soft wheat starch granules.

Reference[a]	Protein[b]	Peptide sequence[c]	
		Hard wheat	Soft wheat
gi|391929	Tritin [*Triticum aestivum*]	(R)ADNLYWEGFK(S)	…
	…	(K)TSPASTGLTLATR(A)	(K)TSPASTGLTLATR(A)
	…	(R)QQMADAVTALYGR(T)	(R)QQMADAVTALYGR(T)
	…	(R)FQTVSGFVAGVLHPK(E)	…
	…	(K)AQVNGWQDLSEALLK(T)	(K)AQVNGWQDLSEALLK(T)
	…	(R)EAVTTLLLMVHEATR(F)	…
	…	(R)DLLGDTDKLTNVALGR(Q)	(R)DLLGDTDKLTNVALGR(Q)
	…	(R)TAEQAAATLGILLFVQVPGGMTVAQALELFHK(S)	…
gi|5669008	Secretory protein [*T. aestivum*]	(K)VLSDASAFIWK(T)	…
	…	(K)SGYTDDFFAQILGK(N)	(K)SGYTDDFFAQILGK(N)
gi|62465514	Class II chitinase [*T. aestivum*]	(R)AIGVDLLSNPDLVATDPTVSFK(T)	(R)GPIQLSHNYNYGPAGR(A)
	…	(R)AIGVDLLSNPDLVATDPTVSFK(T)	…
gi|22001285	Peroxidase 1 [*T. aestivum*]	(K)DIGLAAGLLR(L)	…
gi|20804336	Xylanase inhibitor protein I [*T. aestivum*]	…	(R)ALATGIFER(A)

[a] GenBank GenInfo Identifier (gi) number for the most likely protein candidate for which peptides were identified.
[b] Species from which each protein derives is indicated in square brackets.
[c] Amino acids in parentheses indicate residues expected to flank each identified peptide as determined from the corresponding protein's amino acid sequence.

sequence only). These peptides have been compiled into **Table 3.4** and represent a source of novel starch-granule-associated proteins.

Peptides from all four samples that were found to match regions of plant proteins are listed in **Table 3.5**. They compose a variety of protein classes, including starch granule synthesis (gi|4760582 starch synthase, GBSSI), storage (gi|228310 globulin 2), and transcription (gi|18401374 transcription factor). Proteins such as GBSSI and a grain softness protein (gi|607198) are known to be strongly associated with the starch granule surface (Dubreil et al. 1998; Feiz et al. 2009). Overall, more peptides were identified in the hard wheat (52 peptides) than the soft wheat samples (30 peptides). Finally, several peptides corresponded to proteins of related organisms with no homologous proteins sequenced in wheat as well as numerous peptides that have not been sequenced in any organism.

3.7 Discussion

3.7.1 Proteins of pathogen origin

Peptides identical to regions of proteins of plant pathogens were identified at the starch granule surface (**Table 3.1**). More peptides (13 peptides from six pathogen species) were identified from hard wheat starch granules than from soft wheat starch granules (2 peptides from one pathogen species). It may be that fewer molecular "footprints" were left behind by pathogen intruders in soft wheat, implying that soft wheat has a more robust antimicrobial armament than hard wheat. Many factors may be at play in such a scenario; anatomical and biochemical characteristics may include differences in the waxy cuticle

Table 3.4. List of peptides identical to regions of proteins identified by Mascot as being hypothetical or of unknown function. Generated from mass spectrometric analysis of aqueous supernatants and isopropanol extracts of trypsin-treated hard and soft wheat starch granules.

Reference[a]	Protein[b]	Peptide sequence[c]	
		Hard wheat	Soft wheat
gi|1323750	Unknown protein [*Triticum aestivum*]	(R)ANGGLIEGIADYVR(L)	(R)ANGGLIEGIADYVR(L)
gi|22238	Unnamed protein product [*Zea mays*]	(R)VPTVDVSVVDLTVR(I)	…
gi|4662644	Unknown protein [*Arabidopsis thaliana*]	(R)VLIVVNTEER(K)	…
gi|92874420	At1g27150/T7N9_21-related [*Medicago truncatula*]	(M)LAFPLLELGQMK(E)	…
gi|38345325	OSJNBa0004N05.22 [*Oryza sativa*]	(R)NTAGEEKQQTLALR(V)	…
gi|115485083	Os11g0276000 [*O. sativa*]	(K)SCLMPLRK(S)	…
gi|115440299	Os01g0778800 [*O. sativa*]	(K)FDRVESEVALLR(E)	…
gi|125572766	Hypothetical protein OsJ_004106 [*O. sativa*]	(R)SAKPQHR(I)	…
gi|79561588	Unknown protein [*A. thaliana*]	(M)VEMMRSENHLR(Q)	…
gi|50428631	Protein of unknown function [*O. sativa*]	(K)CDNMASQYFVCCR(W)	…
gi|12321294	Unknown protein [*A. thaliana*]	(K)TPPPPPPPPPPLPSR(S)	…
gi|18087666	Unknown protein [*O. sativa*]	(R)IVVADPPYLSK(E)	…
gi|15223729	Unknown protein [*A. thaliana*]	(K)QVEEMVLVKWVEWKK(E)	…
gi|51090857	Hypothetical protein [*O. sativa*]	(R)IPRWRRR(L)	…
gi|115434166	Os01g0116100 [*O. sativa*]	(K)CSIAKDKPKR(N)	…
gi|18972	Unnamed protein product [*Hordeum vulgare*]	…	(R)AIGVDLLR(N)
	…	…	(R)NPDLVATDPTVSFK(T)
	…	…	(R)GPIQLSHNYNYGPAGR(A)
gi|37718800	Hypothetical protein [*O. sativa*]	…	(R)SQAKAVAPTSRAR(S)
gi|115434624	Os01g0157200 [*O. sativa*]	…	(R)EAASLAEAAK(D)
gi|15225665	Unknown protein [*A. thaliana*]	…	(R)VVESMRSK(S)
gi|32400760	Unknown protein [*T. aestivum*]	…	(R)QLVQIPEQAR(C)
gi|38346495	OSJNBa0020I02.9 [*O. sativa*]	…	(R)CSAKYLK(S)
gi|24414171	Hypothetical protein [*O. sativa*]	…	(K)VVTVSARPK(–)
gi|9755378	F17F8.11 [*A. thaliana*]	…	(R)ASKSGLHR(V)
gi|125533309	Hypothetical protein OsI_033816 [*O. sativa*]	…	(R)VQKASISLVMK(L)
gi|6692258	Unknown protein [*A. thaliana*]	…	(K)NPLAIVNVKGR(A)

[a] GenBank GenInfo Identifier (gi) number for the most likely protein candidate for which peptides were identified.

[b] Species from which each protein derives is indicated in square brackets.

[c] Amino acids in parentheses indicate residues expected to flank each identified peptide as determined from the corresponding protein's amino acid sequence. C-terminal peptides are indicated by (–) at the C-terminus. Where peptides were detected as having an oxidized methionine, the residues have been underlined.

Table 3.5. List of peptides identical to regions of proteins from the plant species *Oryza sativa*, *Hordeum vulgare*, *Zea mays*, *Arabidopsis thaliana* and *Medicago truncatula*. Generated from mass spectrometric analysis of aqueous supernatants and isopropanol extracts of trypsin-treated hard and soft wheat starch granules.

Reference[a]	Protein[b]	Peptide sequence[c]	
		Hard wheat	Soft wheat
gi|4760582	Starch synthase (GBSSI) [*Triticum aestivum*]	(K)DAWDTSVVSEIK(V)	…
	…	(K)EALQAEVGLPVDR(K)	…
	…	(K)FLAANYDVTTALEGK(A)	…
	…	(K)SSFDFIDGYDKPVEGR(K)	…
	…	(K)VLTVSPYYAEELISGEAR(G)	…
gi|32400764	β Amylase [*T. aestivum*]	(R)YDPTAYNTILR(N)	…
	…	(R)FFVDNGTYLTEQGR(F)	…
	…	(K)AAAAMVGHPEWEFPR(D)	…
gi|169777	β Amylase [*Oryza sativa*]	(K)SAPEELVQQVLSAGWR(E)	…
	…	(R)NIEYLTLGVDDQPLFHGR(T)	…
gi|1323750	Unknown protein [*T. aestivum*]	(R)ANGGLIEGIADYVR(L)	(R)ANGGLIEGIADYVR(L)
gi|167004	Embryo globulin [*Hordeum vulgare*]	(R)VAIMEVNPR(A)	…
	…	(R)VAIMEVNPR(A)	…
	…	(R)DTFNLLEQRPK(I)	…
gi|228310	Globulin 2 [*Zea mays*]	(K)EGEGVIVLLR(G)	(K)EGEGVIVLLR(G)
gi|17425168	Low-molecular-weight glutenin subunit group 3 type II [*T. aestivum*]	(R)TTTRVPFGVGTGVGGY(–)	…
gi|15220239	RNA binding protein [*Arabidopsis thaliana*]	(K)IVAILGSMRVPN(–)	…
gi|15148385	Gamma-gliadin [*T. aestivum*]	(K)APFASIVADIGGQ(–)	…
gi|607198	15-kDa grain softness protein [*T. aestivum*]	(K)AIWTSIQGDLSGFK(G)	…
gi|22238	Unnamed protein product [*Z. mays*]	(R)VPTVDVSVVDLTVR(I)	…
gi|15241221	Structural molecule [*A. thaliana*]	(K)VLRSRSRR(L)	…
gi|4662644	Unknown protein [*A. thaliana*]	(R)VLIVVNTEER(K)	…
gi|12323783	Putative exportin, tRNA [*A. thaliana*]	(R)SKVTSFIHR(M)	…
gi|92874420	At1g27150/T7N9_21-related [*Medicago truncatula*]	(M)LAFPLLELGQMK(E)	…
gi|886965	Low molecular weight glutenin [*T. aestivum*]	(R)QLPQIPEQSR(Y)	…
gi|38345325	OSJNBa0004N05.22 [*O. sativa*]	(R)NTAGEEKQQTLALR(V)	…
gi|115485083	Os11g0276000 [*O. sativa*]	(K)SCLMPLRK(S)	…
gi|115440299	Os01g0778800 [*O. sativa*]	(K)FDRVESEVALLR(E)	…
gi|125572766	Hypothetical protein OsJ_004106 [*O. sativa*]	(R)SAKPQHR(I)	…
gi|79561588	Unknown protein [*A. thaliana*]	(M)VEMMRSENHLR(Q)	…
gi|50428631	Protein of unknown function [*O. sativa*]	(K)CDNMASQYFVCCR(W)	…
gi|12321294	Unknown protein [*A. thaliana*]	(K)TPPPPPPPPPPLPSR(S)	…
gi|18401374	Transcription factor [*A. thaliana*]	(K)TLWSRAILSKAVK(L)	…
gi|18087666	Unknown protein [*O. sativa*]	(R)IVVADPPYLSK(E)	…
gi|15223729	Unknown protein [*A. thaliana*]	(K)QVEEMVLVKWVEWKK(E)	…
gi|51090857	Hypothetical protein [*O. sativa*]	(R)IPRWRRR(L)	…
gi|15232073	Lyase/pectate lyase [*A. thaliana*]	(R)IYVVTSPR(D)	…
gi|115434166	Os01g0116100 [*O. sativa*]	(K)CSIAKDKPKR(N)	…
gi|75282488	Protein H2A.5 [*T. aestivum*]	(R)LLAGVTIAHGGVIPNINSVLLPK(K)	…
gi|871551	Serpin (protease inhibitor) [*T. aestivum*]	(K)ISFGIEASDLLK(C)	…
gi|18972	Unnamed protein product [*H. vulgare*]	…	(R)AIGVDLLR(N)
	…	…	(R)NPDLVATDPTVSFK(T)
	…	…	(R)GPIQLSHNYNYGPAGR(A)
gi|37718800	Hypothetical protein [*O. sativa*]	…	(R)SQAKAVAPTSRAR(S)
gi|92875830	Trp repressor/replication initiator [*M. truncatula*]	…	(K)VFSDEVIR(F)
gi|92893314	Ribonuclease H; Glutathione S-transferase, C-terminal-like [*M. truncatula*]	…	(K)KIWSKDIPPSK(S)
gi|115434624	Os01g0157200 [*O. sativa*]	…	(R)EAASLAEAAK(D)
gi|15225665	Unknown protein [*A. thaliana*]	…	(R)VVESMRSK(S)
gi|32400760	Unknown protein [*T. aestivum*]	…	(R)QLVQIPEQAR(C)
gi|54291437	Pr1-like protein [*O. sativa*]	…	(R)RELESGKEEGR(R)
gi|14587294	Putative permease 1 [*O. sativa*]	…	(R)GMPWWVPFR(S)
	…	…	(R)GMPWWVPFR(S)
gi|38346495	OSJNBa0020I02.9 [*O. sativa*]	…	(R)CSAKYLK(S)
gi|24414171	Hypothetical protein [*O. sativa*]	…	(K)VVTVSARPK(–)
gi|9755378	F17F8.11 [*A. thaliana*]	…	(R)ASKSGLHR(V)
gi|125533309	Hypothetical protein OsI_033816 [*O. sativa*]	…	(R)VQKASISLVMK(L)
gi|50726319	Putative BRCA1 interacting protein C-terminal helicase 1 [*O. sativa*]	…	(K)DPVPVNTRRSK(F)
gi|6692258	Unknown protein [*A. thaliana*]	…	(K)NPLAIVNVKGR(A)
gi|15229913	CYP71B38; heme binding/iron ion binding/monooxygenase /oxygen binding [*A. thaliana*]	…	(K)KVPLVLIQTSHR(W)
gi|30688003	Phosphoinositide 5-phosphatase [*A. thaliana*]	…	(R)NFKLTLIAR(R)

[a] GenBank GenInfo Identifier (gi) number for the most likely protein candidate for which peptides were identified.

[b] Species from which each protein derives is indicated in square brackets.

[c] Amino acids in parentheses indicate residues expected to flank each identified peptide as determined from the corresponding protein's amino acid sequence. C-terminal peptides are indicated by (–) at the C-terminus. Where peptides were detected as having an oxidized methionine, the residues have been underlined.

between hard and soft wheat grains, numbers of pericarp cell layers, suberization, and even enzyme inhibitors in the aleurone or subaleurone cell layers. Pathogen proteins closely associated with the starch granules of wheat may be the molecular remnants of a failed pathogen attack against the developing or mature granules of the wheat kernel. Using a green fluorescent protein-expressing *Fusarium* sp., Miller et al. (2004) showed that the fungus was able to infect the majority of a wheat seed coat 16 days after inoculation, though no infection of the endosperm was observed using this method. However, microscopy of *F. graminearum*-infected wheat and barley endosperm have shown pitted starch granules and fungal hyphae (Miller et al. 1984; Pritsch et al. 2000). Therefore, it is not surprising to find remnant proteins of fungal pathogens within the endosperm of mature wheat kernels. We propose that such molecular footprints of grain shipments may be used to trace the cumulative historical record of its incurred biotic stresses or pathogen load. Of the pathogen proteins identified (**Table 3.1**), several are cytosolic and serve no known direct role in pathogenicity. These include an oxidoreductase (gi|21245076), a D-amino acid dehydrogenase subunit (gi|21243478), a putative branched-chain amino acid aminotransferase (gi|50121895), and a bifunctional Nacetylglucosamine- 1-phosphate uridyltransferase (gi|50123425). This suggests that bacterial pathogens and fungal hyphae were lysed during seed maturation and fixed in place within the desiccated endosperm. The in planta role of these hypothetical plant pathogen proteins in **Table 3.1** may be in some way associated with starch granule degradation or they may be related to the recognition and degradation of carbohydrate polymers in general. However, like the cytosolic pathogen proteins, the hypothetical proteins may not be directly involved in infection. Further study is needed to characterize and elucidate the function of these hypothetical proteins to determine the role they may play in host-pathogen interactions.

The presence of peptides identical to regions of plant pathogen approach of probing the starch granule surface goes beyond providing insight into the starch granule proteome only: it provides a view of the "late-game" interaction between plant and pathogen. Further experiments may determine the effect of pathogen infection on the wheat starch granule proteome or, conversely, the effect of variation in the wheat starch granule proteome on pathogen infection.

3.7.2 Host defense proteins

Because defense-related proteins are known to compose an important fraction of the cereal proteome (Koller et al. 2002), we set out to discover proteins defending the seed's energy stores. Peptides identical to regions of several antimicrobial proteins, as annotated in the NCBI protein database, were identified in the supernatant and granule-bound isolates of tryptically shaved hard and soft wheat starch granules (**Tables 3.2 and 3.3**). If present at the starch granule surface, these proteins may play a role in protecting the starch granule from invading pathogens. This would ensure that the starch granules are preserved exclusively for the developing wheat embryo. Interestingly, a peptide identical to a region of an amylase inhibitor was identified in the supernatant (gi|2894148) (**Table 3.2**). α-Amylase and other glycosylase inhibitors may play a role in preventing the breakdown of starch and cell wall components by invading pathogens or may play a role in regulating growth and degradation of the starch granule by the plant itself (Howell et al. 2005; Lionetti et al. 2007). Another interesting result is the detection of a xylanase inhibitor (gi|20804336) (**Tables 3.2 and 3.3**), specifically a xylanase-inhibiting protein (XIP)-type inhibitor associated with the starch granules of soft wheat. A proteomic analysis of four commercial starches by Kasarda et al.

(2008) identified xylanase inhibitors, suggesting that they may play a role in antimicrobial activity. XIP is exclusively expressed in wheat endosperm and has been suggested to play a role in plant defense (Croes et al. 2009). Although XIP-type xylanase inhibitors are secreted into the intercellular matrix (Juge et al. 2006), it may be possible that, during seed maturation and desiccation, some of the highly expressed xylanase inhibitors become associated with the starch granule surface. It is important to note, when comparing these tables, that several peptides in **Table 3.3** are identical to those in **Table 3.2**, indicating that there were some proteins found in both the aqueous and isopropanol extracts. Given that the latter is 50% aqueous, it is not surprising that the two solvents were not mutually exclusive to a degree high enough as to effect a discriminatory or differential solubility of granule-bound proteins.

3.7.3 Hypothetical proteins

The proteome of any organism is highly diverse and the identification and characterization of most proteomes is largely incomplete. The motivation to characterize novel proteins generally depends on their candidacy for connection or interaction with a characterized system. By limiting our analysis to the starch granule proteome, novel identifications may be related to starch granule defense or metabolism, pending future experiments. The use of mass spectrometric techniques to discover unknown and hypothetical proteins localized to specific organellar compartments, such as the starch granule surface, may aid in expanding the body of knowledge of host–pathogen interactions and, perhaps, result in the discovery of novel resistance mechanisms. For example, the unknown protein with gi|1323750 (**Table 3.4**) is the protein product of a cDNA sequence (gi|1323750). This cDNA sequence from wheat was entered into the database in 1996 and

contains an ORF annotated as “unknown”. Subsequently, this protein of unknown function was found to match the superfamily of basic secretory protein class of proteins thought to play a role in host defense (Kuwabara et al. 1999). Other hypothetical proteins, such as OsJ_004106 (gi|125572766), did not match with any known proteins, nor do they contain any known conserved domains when analyzed using the NCBI BLAST protein search tools. Therefore, these hypothetical proteins can no longer be considered hypothetical because the present peptide extraction actually recovered them from granules as bona fide wheat proteins. This novel cohort of wheat proteins presents an opportunity to explore potentially new classes of antimicrobial or starch-metabolism-related proteins not yet characterized.

3.7.4 Host proteins related to storage and metabolism

Similar to other proteomic studies of wheat and potato amyloplasts (Andon et al. 2002; Balmer et al. 2006; Stensballe et al. 2008), a number of proteins related to defense (**Tables 3.2 and 3.3**) and starch granule metabolism and storage (**Table 3.5**) were identified in the present study. These include the starch metabolism enzymes granule-bound starch synthase I (GBSSI) and β-amylase, as well as the storage protein globulin, identified here as the barley protein, embryo globulin. The identification of proteins already known to associate with the starch granule, such as GBSSI and grain softness protein, validates the efficacy of this current tryptic shaving technique to probe the starch granule surface proteome and extend this organellar approach to the investigation of host–pathogen interactions. Peptides identical to regions of proteins from the cereals rice, barley, and corn as well as the model plants *Arabidopsis* and alfalfa were identified. For these protein hits, it is most likely that the reason they are annotated as non-wheat is that they share sequence identity with homologs

from wheat which have not yet had DNA, RNA, or protein sequence information reported and, therefore, are annotated only under related organisms which have had their genomes sequenced. It is interesting to note that our Mascot identification did not proffer any host stress-related proteins. Other proteomic studies, such as that by Kasarda et al. (2008), detected peptides identical to regions of stress-related proteins. In conclusion, this tryptic peptide shaving of hard and soft wheat starch granules was a useful means to characterize host-pathogen interactions at the gateway to energy reserves. Isolating tightly bound peptides from the starch granule surface identified peptide sequences from six common field pathogens of wheat. Many other recovered peptides were unique in that they were the first occurrence of previous putative proteins indicated by genomic data. Recovering this novel community of both host and pathogen proteins clustered on the granule landscape now permits a more dynamic tracking of the penetrative power of pathogens and their proteins.

3.8 References

Andon, N. L., Hollingworth, S., Koller, A., Greenland, A. J., Yates, J. R., and Haynes, P. A. 2002. Proteomic characterization of wheat amyloplasts using identification of proteins by tandem mass spectrometry. Proteomics 2:1156-1168.

Balmer, Y., Vensel, W. H., DuPont, F. M., Buchanan, B. B., and Hurkman, W. J. 2006. Proteome of amyloplasts isolated from developing wheat endosperm presents evidence of broad metabolic capability. J. Exp. Bot. 57:1591-1602.

Bottalico, A., and Perrone, G. 2002. Toxigenic *Fusarium* species and mycotoxins associated with head blight in small-grain cereals in Europe. Eur. J. Plant Pathol. 108:611-24.

Brennan, J. M., Egan, D., Cooke, B. M., and Doohan, F. M. 2005. Effect of temperature on head blight of wheat caused by *Fusarium culmorum* and *F. graminearum*. Plant Pathol. 54:156-160.

Buleon, A., Colonna, P., Planchot, V., and Ball, S. 1998. Starch granules: Structure and biosynthesis. Int. J. Biol. Macromol. 23:85-112.

Cowger, C., Patton-Ozkurt, J., Brown-Guedira, G., and Perugini, L. 2009. Post-anthesis moisture increased *Fusarium* head blight and deoxynivalenol levels in North Carolina winter wheat. Phytopathology 99:320- 327.

Croes, E., Gebruers, K., Luyten, N., Delcour, J. A., and Courtin, C. M. 2009. The three classes of wheat xylanase-inhibiting proteins accumulate in an analogous way during wheat ear development and germination. J. Plant Physiol. 166:1253-1262.

Dangl, J. L., and Jones, J. D. G. 2001. Plant pathogens and integrated defence responses to infection. Nature 411:826-833.

Dubreil, L., Gaborit, T., Bouchet, B., Gallant, D. J., Broekaert, W. F., Quillien, L., and Marion, D. 1998. Spatial and temporal distribution of the major isoforms of puroindolines (puroindoline-a and puroindoline-b) and non specific lipid transfer protein (ns-LTPle(1)) of *Triticum aestivum* seeds. Relationships with their in vitro antifungal properties. Plant Sci. 138:121-135.

Feiz, L., Wanjugi, H. W., Melnyk, C. W., Altosaar, I., Martin, J. M., and Giroux, M. J. 2009. Puroindolines co-localize to the starch granule surface and increase seed bound polar lipid content. J. Cereal Sci. 50:91-98.

Finnie, C., Melchior, S., Roepstorff, P., and Svensson, B. 2002. Proteome analysis of grain filling and seed maturation in barley. Plant Physiol. 129:1308-1319.

Higdon, R., Hogan, J. M., Van Belle, G., and Kolker, E. 2005. Randomized sequence databases for tandem mass spectrometry peptide and protein identification. Omics 9:364-79.

Howell, J. T., and Davis, M. R. 2005. Plant defense mechanisms against fungal pathogens: Polygalacturonase inhibitor proteins. Can. J. Plant Pathol. 24:5-15.

Hugouvieux-Cotte-Pattat, N., Condemine, G., Nasser, W., and Reverchon, S. 1996. Regulation of pectinolysis in *Erwinia chrysanthemi*. Annu. Rev. Microbiol. 50:213-257.

Jackowiak, H., Packa, D., Wiwart, M., and Perkowski., J. 2005. Scanning electron microscopy of *Fusarium* damaged kernels of spring wheat. Int. J. Food Microbiol. 98:113-123.

Juge, N., and Svensson, B. 2006. Proteinaceous inhibitors of carbohydrate- active enzymes in cereals: Implication in agriculture. J. Sci. Food Agric. 86:1573-1586.

Kasarda, D. D., Dupont, F. M., Vensel, W. H., Altenbach, S. B., Lopez, R., Tanaka, C. K., and Hurkman, W. J. 2008. Surface-associated proteins of wheat starch granules: Suitability of wheat starch for celiac patients. J. Agric. Food Chem. 56:10292-10302.

Keen, N. T., and Tamaki, S. 1986. Structure of two pectate lyase genes from *Erwinia chrysanthemi* EC16 and their high-level expression in *Escherichia coli*. J. Bacteriol. 168:595-606.

Koller, A., Washburn, M. P., Lange, B. M., Andon, N. L., Deciu, C., Haynes, P. A., Hays, L., Schieltz, D., Ulaszek, R., Wei, J., Wolters, D., and Yates, J. R. 2002. Proteomic survey of metabolic pathways in rice. Plant Biol. 99:11969-11974.

Kuwabara, C., Arakawa, K., and Yoshida, S. 1999. Abscisic acid-induced secretory proteins in suspension-cultured cells of winter wheat. Plant Cell Physiol. 40:184-191.

Lacey, J., Bateman, G. L., and Mirocha, C. J. 1999. Effects of infection time and moisture on development of ear blight and deoxynivalenol production by *Fusarium* spp. in wheat. Ann. Appl. Biol. 134:277-283.

Lionetti, V., Raiola, A., Camardella, L., Giovane, A., Obel, N., Pauly, M., Favaron, F., Cervone, F., and Bellincampi, D. 2007. Overexpression of pectin methylesterase inhibitors in *Arabidopsis* restricts fungal infection by *Botrytis cinerea*. Plant Physiol. 143:1871-1880.

Mehta, A., Brasileiro, A. C. M., Souza, D. S. L., Romano, E., Campos, M. A., Grossi-de-Sá, M. F., Silva, M. S., Franco, O. L., Fragoso, R. R., Bevitori, R., and Rocha, T. L. 2008. Plant–pathogen interactions: What is proteomics telling us? FEBS J. 275:3731-3746.

Mentewab, A., Rezanoor, H. N., Gosman, N., Worland, A. J., and Nicholson, P. 2000. Chromosomal location of *Fusarium* head blight resistance genes and analysis of the relationship between resistance to head blight and brown foot rot. Plant Breed. 119:15-20.

Miller, S. S., Chabot, D. M. P., Ouellet, T., Harris, L. J., and Fedak, G. 2004. Use of a *Fusarium graminearum* strain transformed with green fluorescent protein to study infection in wheat (*Triticum aestivum*). Can. J. Plant Pathol. 26:453-463.

Miller, S. S., Yiu, S. H., Fulcher, R. G., and Altosaar, I. 1984. Preliminary evaluation of lectins as fluorescent-probes of seed structure and composition. Food Microstruct. 3:133-139.

Paper, J. M., Scott-Craig, J. S., Adhikari, N. D., Cuomo, C.A., and Walton, J. D. 2007. Comparative proteomics of extracellular proteins in vitro and in planta from the pathogenic fungus *Fusarium graminearum*. Proteomics 7:3171-3183.

Pritsch, C., Muehlbauer, G. J., Bushnell, W. R., Somers, D. A., and Vance, C. P. 2000. Fungal development and induction of defense response genes during early infection of wheat spikes by *Fusarium graminearum*. Mol. Plant-Microbe Interact. 13:159-169.

Schwarz, P. B. 2003. Impact of Fusarium head blight on malting and brewing quality of barley. Pages 395-419 in: Fusarium head blight of wheat and barley. K. J. Leonard and W. R. Bushnell, eds. American Phytopathological Society Press, St. Paul, MN.

Stensballe, A., Hald, S., Bauw, G., Blennow, A., and Welinder, K. G. 2008. The amyloplast proteome of potato tuber. FEBS J. 275:1723-1741.

Tester, R. F., Karkalas, J., and Qi, X. 2004. Starch - composition, fine structure and architecture. J. Cereal Sci. 39:151-165. 32.

Wall, M. L., Wheeler, H., Huebsch, M. P., Smith, J. C., Figeys, D., and Altosaar, I. The tryptophan rich domain of puroindoline is directly associated with the starch granule surface as judged by tryptic shaving and mass spectrometry. J. Cereal Sci. 52:115-120.

Zange, B. J., Kang, Z., and Buchenauer, H. 2005. Effect of Folicur in infection process of *Fusarium culmorum* in wheat spikes. Plant Dis. Prot. 112:52-64.

Zhou, W., Eudes, F., and Laroche, A. 2006. Identification of differentially regulated proteins in response to a compatible interaction between the pathogen *Fusarium graminearum* and its host, *Triticum aestivum*. Proteomics 6:4599-4609.

Chapter 4

The Wheat Puroindoline Genes Increase Germ Size and Seed Oil Content in Transgenic Corn

Jinrui Zhang, John M. Martin, Brian Beecher, Chaofu Lu, L. Curtis Hannah, Michael L. Wall, Illimar Altosaar, Michael J. Giroux

4.1 Author contributions

Michael L. Wall contributed bright field and fluorescent microscopy of wild-type corn B73 and transgenic line UP8 and also contributed to the interpretation of the integrative plant anatomy of the transgenic maize seeds, targeting of puroindoline in maize kernel tissues, to visualization of unintended effects in transgenic organisms, as well as to the editing of the manuscript with the assistance of his supervisor, Dr. Illimar Altosaar. Primary writing of the manuscript and preparation of most figures and tables was conducted by Dr. Jinrui Zhang in the lab of Dr. Michael. J. Giroux (Montana State University, Bozeman, MO), John M. Martin (Montana State University, Bozeman, MO) and Brian Beecher (USDA, Pullman, WA). Seed oil and lipid analysis was performed with the assistance of Dr. Chaofu Lu (Montanta State University, Bozeman, MO). Corn transformation and field planting was performed with the assistance of Dr. L. Curtis Hannah (University of Florida, Gainesville, FL).

4.2 Significance of manuscript

Increasing the nutritional content of grain is one of the greatest goals of crop breeding. Achieving increased oil content in a grain has been one of the most difficult challenges. We demonstrate one of the highest increases in total lipid content ever achieved by recombinant DNA technology in a crop plant, achieved by expressing the wheat protein puroindoline in the embryo and endosperm tissue of transgenic corn. The expression of puroindoline in the endosperm of these transgenic corn plants demonstrated that puroindoline, a wheat protein which localizes at the starch granule surface, will also localize to the surface of corn starch granules. This localization has significant implications in controlling grain texture and milling properties, antimicrobial activity and possible implications for plant biotechnology.

4.3 Abstract

Plant oil content and composition improvement is a major goal of plant breeding and biotechnology. Puroindoline a and b (PIN-a and PIN-b) together control whether wheat seeds are soft or hard textured. They share a similar structure to that of plant non-specific lipid-transfer proteins. Here we transformed corn (*Zea mays* L.) with the wheat (*Triticum aestivum* L.) puroindoline genes (*Pina* and *Pinb*) to assess their effects upon seed oil content and quality. *Pina* and *Pinb* were introduced into corn under the control of a corn Ubiquitin promoter. Three *Pina*/*Pinb* expression positive transgenic events were evaluated over two growing seasons. The results showed that Pin expression increased germ size significantly without negatively impacting seed size. Germ yield increased 33.8% while total seed oil

content was increased by 25.23%. Seed oil content increases were primarily the result of increased germ size. This work indicates that higher oil content corn hybrids having increased oil content or increased feed value could be produced via puroindoline over-expression.

4.4 Introduction

Plant oil is an important renewable resource for both edible and industrial uses. An increase in seed oil content is a major goal of plant breeding and biotechnology (reviewed in Jaworski and Cahoon 2003). Maize or corn (*Zea mays* L.) oil is the main co-product of corn wet and dry milling processes, and contains a high concentration of essential, polyunsaturated fatty acids, which make it an excellent source of energy and essential fatty acids for food and feed uses (Weber 1987). High oil corn is particularly attractive as a feed, due to its high metabolizable energy (Alexander 1988a). Increases in corn seed oil content are possible by traditional breeding. The Illinois high-oil (IHO) and the Alexho single kernel (ASK) synthetic are examples of high-oil corn developed by breeders through recurrent selection (Laurie et al. 2004; Alexander 1988b). However, commercial high-oil corn hybrids are associated with reduced seed size and agronomic yield (Dudley and Lambert 2004). Many researchers have attempted to identify quantitative trait loci (QTLs) for oil content or quality to increase or modify plant oil content via manipulation of various oil biosynthetic pathway genes (Poneleit and Alexander 1965; Alrefai et al. 1995; Song et al. 2004; Beló et al. 2008; Zheng et al. 2008). For example, Zheng et al. (2008) identified a major oil QTL and demonstrated that increasing acyl-CoA:diacylglycerol acyltransferase (DGAT) during corn seed development can achieve relative increases in oil of approximately 40%. The

synthesis of storage oils requires the coordinated activity of many enzymes in the lipid biosynthetic pathway. Despite a good understanding of the genes involved, there has been little success at transgenically increasing seed oil content (Clemente and Cahoon 2009). Genes whose products physically interact with lipids may be able to increase seed lipid content by decreasing the level of lipid breakdown.

Puroindoline-a and -b genes (*Pina* and *Pinb*) together make up the functional components of the wheat grain hardness locus (*Ha*) and control whether wheat seeds are soft or hard textured (Giroux and Morris 1998). They share a similar structure to that of non-specific lipid-transfer proteins (nsLTPs) (Giroux and Morris 1997; Marion et al. 1994; Le et al. 1996; Douliez et al. 2000). Both Puroindoline (PIN) isoforms, and their homologs in barley and oat, contain a backbone of 10 cysteine residues and likely form a tertiary structure comprised of four α-helices separated by loops and stabilized by five disulphide bridges (reviewed in Bhave and Morris 2008). The unique tryptophan-rich domain found in PINs has been demonstrated to have lipid-binding properties (Kooijman et al. 1997). This motif is believed to be a non-stick agent that enables PINs to bind to starch granule surface lipids preventing adhesion between starch granules and the surrounding protein matrix during seed maturation (Marion et al. 1994; Giroux and Morris 1998). Significant quantitative differences exist in seed polar lipid (glyco- and phospho-lipids) content between wheats varying in PIN content (Greenblatt et al. 1995; Feiz et al. 2009).

Plant lipids are generally stored as triacylglycerols (TAGs) in oil bodies (Huang 1992). Only a limited number of proteins are specifically associated with plant seed lipid bodies (reviewed in Purkrtova et al. 2008). Oleosins are the major proteins associated with oil bodies and maintain oil bodies as small single units preventing their coalescence during

seed desiccation (Leprince et al. 1997). Siloto et al. (2006) observed that oleosins are important factors in determining seed oil body size in *Arabidopsis*. Both PIN and oleosins have a central hydrophobic region which confers a lipid-binding property. However, while oleosins are integral to oil bodies, PINs are normally only associated with the surface of starch granules (Capparelli et al. 2005; Feiz et al. 2009; Wall et al. 2010).

Here we tested whether the increased expression of Pins can improve seed oil content. Puroindolines are endosperm specific and are found only in the *Triticeae* (Wiley et al. 2007). Corn does not express PINs or even contain PIN homologues (Gautier et al. 2000). Therefore, both *Pina* and *Pinb* were introduced into corn under the control of a corn Ubiquitin gene (*Ubi1*) promoter, and the effect of PINs on seed composition and oil content was studied.

4.5 Materials and Methods

4.5.1 Plasmid constructs

The *Pina* and *Pinb* expression vectors, pUbiPina and pUbiPinb, were described in Krishnamurthy and Giroux (2001). They carry the *Pina-D1a*/*Pinb-D1a* alleles reported by Gautier et al. (1994) that are found in all soft hexaploid wheats (Morris et al. 2001). In short, PCR products containing the coding sequences of *Pina* or *Pinb* were digested with BamHI and ligated into BamHI-digested and phosphatase treated pAHC17 plasmid (Christensen and Quail 1996), downstream of the corn *Ubi1* promoter and intron. All resultant plasmids were sequenced to ensure correct orientation and fidelity of ligation junctions. The construct pBAR184 used for selection (Frame et al. 2000) contains the corn *Ubi1* promoter-*Bar* gene

cassette as a selectable marker conferring resistance to the herbicides bialaphos (4-[hydroxy(methyl-)phosphinoyl]-L-homoalanyl-L-alanyl-L-alanine) and glufosinate (DL-homoalanin-4-yl(methyl) phosphinic acid).

4.5.2 Corn transformation

Corn transformation was conducted in the Iowa State University Plant Transformation Facility according to their standard protocols (Frame et al. 2000). Highly embryogenic corn line Hi-II was used in biolistic transformation. Embryogenic calli generated from immature embryos of the Hi-II genotype were co-bombarded with the plasmids using a 1:2.5:2.5 molar ratio (pBAR184:pUbiPina:pUbiPinb). Bialaphos-resistant calli were screened by PCR for the *Pina* and *Pinb* coding sequences using gene specific primers (**Table 4.1**). The method for screening of the transgenic positive and negative lines of each event was described in Zhang et al. (2009). Herbicide-resistant *Pina*/*Pinb* PCR positive T0 lines were pollinated using inbred line B73. Herbicide-resistance of T0 and all subsequent progeny plants was tested by painting an individual leaf with 0.1% glufosinate ammonium (Zhang et al. 2009). Herbicide-resistant T1 plants were heterozygous for the transgene locus and were self-pollinated to produce T2 progeny. Individual T2 plants from each transformation event were tested for both *Pin* genes as well as for the presence of *Bar*. T2 plants giving progeny (T3 plants) in which more than 12 consecutive sprayed plants were herbicide resistant and *Pina*/*Pinb* PCR positive were considered homozygous positive for *Bar* and *Pina*/*Pinb*, and T2 plants having progeny where more than 5 consecutive sprayed plants were herbicide susceptible and *Pina*/*Pinb* PCR negative were considered homozygous

Table 4.1. The sequence of primers used for PCR screening and qRT-PCR.

No*	Gene*	Database accession	Primer sequences		Size of PCR product
			Forward(5'3')	Reverse(5'-3')	
1	*Pina*	DQ363911	GGTGTGGCCTCATCTCATCT	TCACCAGTAATAGCCAATAGTG	501(795-1296)
2	*Pinb*	DQ363913	AATAAAGGGGAGCCTCAACC	TCACCAGTAATAGCCACTAGGGAA	507(213-720)
3	*Pina*	X69913	TAGCGAAGTTGTTGGCAGTT	TTGAGCATCGATCTAGCAGG	109(87-196)
4	*Pinb*	X69912	GAAGTTGGCGGAGGAGGT	TTTTGTGGGCCAGGTGAC	105(104-229)
5	*Actin*	J01238.1	TCCTGACACTGAAGTACCCGATTG	CGTTGTAGAAGGTGTGATGCCAGTT	83(498-581)

*1-2 PCR primers; 3-5 qRT-PCR primers. The position of the amplified region within the accession sequence is shown in parentheses

negative for *Bar* and *Pina*/*Pinb*. T0 and T1 plants were planted in a Montana State University greenhouse. T2 and T3 plants were planted both in the greenhouse at Montana State University and in the field at the University of Florida in 2007. T4 plants were planted in the field at the University of Florida in 2008. Each line was planted in two rows and each row had 15 plants with between row spacing of 90 cm and within row spacing of 30 cm. Ears from individual selfed plants were harvested, dried for three days at 37 °C in a forced air incubator and then maintained at room temperature and ambient humidity (30%). The moisture content of all seeds at the time analyses were performed was ~10%. The same seed source was used for all the seed quality analysis including kernel weight, starch and protein content, germ yield and oil content. All values represent three independently-derived homozygous positive or negative T3 lines in 2007 and T4 lines in 2008 for each event. Seeds from three ears within the same line were pooled.

4.5.3 Southern blot and Northern blot analysis

Southern-blot analysis was performed using DNA extracted from young leaf tissue of homozygous transgene positive T3 transgenic lines as previously described (Zhang et al. 2009). Genomic DNA was isolated from young leaves according to Riede and Anderson (1996). The blot, hybridization and film development were the same with Zhang et al. 2009. Probes were made from the coding sequence of wheat *Pina* or *Pinb*, amplified as previously described (Gautier et al., 1994). Northern-blot analysis was performed on the RNA from 21 days after pollination (DAP) developing corn seeds (Zhang et al. 2009). RNA was prepared from 21 DAP developing corn seeds using a Trizol protocol (Invitrogen, Carlsbad, CA) as

described in detail in Zhang et al. 2009. The blot, hybridization and film development methods were also the same with Zhang et al. 2009.

4.5.4 Quantitative real-time PCR (qRT-PCR)

Expression of the genes in the different tissues of transgenic lines was quantified by qRT-PCR. The measurements were performed using a Dynamo SYBR green 2-step qRT-PCR kit (Finnzymes, Espoo, Finland) and a MiniOpticon RT-PCR machine (Bio-Rad, Hercules, CA), according to the manufacturer's instructions. Total RNA was prepared from 21 DAP corn embryos and endosperms as well as leaves of plants isolated three weeks after germination. RNA from embryo and endosperm tissue was extracted as described above and corn and wheat control leaf RNA was extracted using a standard Trizol protocol (Invitrogen, Carlsbad, CA). The RNA samples were treated with RQ1 DNase (Promega, Madison, WI) and cDNA was generated using random hexamers and M-MuLV RNase H+ reverse transcriptase (Finnzymes, Espoo, Finland). The primer sequences used to detect *Pina*, *Pinb*, and the control gene *Actin 1* are listed in **Table 4.1**. Thermocycling conditions were as follows: 15 min at 95 ºC, then 40 cycles of 10 s at 94 ºC, 25 s at 56 ºC and 30 s at 72 ºC. Threshold cycle (Ct) values were automatically detected and compared between samples. The specificity and length of PCR products were verified by agarose gel electrophoresis and homogeneity was assessed via melting curve analysis using a temperature increment of 0.2 ºC and a hold time of 2 s between 72 and 90 ºC. The expression of *Pina* and *Pinb* was normalized using *Actin 1* expression level for each RNA sample. The comparative Ct method ($2^{-\Delta\Delta Ct}$) was used for relative quantification. The fold differences presented are comparisons

of homozygous positive transgenic lines and untransformed inbred line B73. Data were represented by two replications.

4.5.5 PIN polyclonal antibody production and ELISA determination of PIN-a and PIN-b content

PIN-a and PIN-b specific polyclonal antibodies (#3825 for PIN-a and #3826 for PIN-b) were prepared by purifying intact PIN-a or PIN-b from wheat seeds by Triton X-114 (TX-114) fractionation followed by SDS-PAGE gel purification. The TX-114 fractionation, SDS-PAGE gel and antibody production protocols have each been described previously (Krishnamurthy and Giroux 2001; Giroux et al. 2003). The PIN-a or PIN-b specific genotypes used as the source material to purify PIN-a or PIN-b were chosen from the transgenic genotypes described by Wanjugi et al. (2007) and contained only PIN-a or PIN-b. For enzyme-linked immunosorbent assay (ELISA), total puroindoline was extracted from 100 mg of mature whole seed meal from three transgenic corn lines and the soft wheat control (Heron) using TX-114 detergent by modifying the previously described method (Giroux et al. 2003) to include an additional round of TX-114 fractionation. Whole meal powder for each assay was created by grinding a random sample of seeds for each genotype using a Perten 3303 Laboratory Mill (Perten Instruments, Stockholm, Sweden).

Extracted proteins were diluted 1/200 in coating buffer (15 mM Na2CO3, 35 mM NaHCO3, pH 9.6). Each well (Immulon 4HBX 96-well plate; Thermo Electron Corporation, Byron Medical Waltham, MA) was coated with 100 µL of the diluted protein samples and the plate was left overnight at 4 ºC. The plate was washed three times at RT with a phosphate buffered saline solution (PBS) containing Tween-20 (137 mM NaCl, 1.5 mM KH2PO4, 8

mM Na2HPO4, 2.7 mM KCl, pH 7.4, 0.1% (v/v) Tween 20). PBS (300 µL) containing 5% (w/v) bovine serum albumin (PBS-BSA) was added and plates were incubated for 2 hrs at RT. After three washes with PBST, 100 µL of a solution containing polyclonal PIN-a or PIN-b specific antiserum (diluted 1:250 in PBS-BSA) was added and plates were incubated overnight at 4 ºC. After three RT washes with PBST, the PIN linked antibodies were revealed by adding 100 µL of a solution containing goat anti-rabbit IgG conjugated with horseradish peroxidase (Millipore, Billerica, MA) (dilution 1:100,000 in PBS-BSA). After 2 hrs at RT, wells were washed eight times with PBST and then 100 µL TMB (3,3',5,5'-tetramethylbenzidine) (Pierce Chemical Co, Rockford, IL) was added and plates were incubated for 15 min at RT. Reactions were stopped by addition of 100 µL 0.18 M H2SO4 and absorbance was measured at 450 nm using a SpectraMax Plus384 spectrophotometer (Molecular Devices, Sunnyvale, CA). Sample wells were blanked against transgene negative corn protein extracts, and all measurements were performed in duplicate.

4.5.6 Germ recovery and total oil content

The oil content of whole seeds and germs of transgenic lines was tested by a procedure similar to that of Li et al. (2006). Three T3 lines in 2007 and T4 lines in 2008 of each event were tested. Data were averages of three lines. Germ yield was measured by a small-scale wet milling method adapted from a previously published procedure (Vignaux et al. 2006). Briefly, 10 g of kernels were steeped in 20 mL H2O containing 0.2% (w/v) SO2 and 0.5% (v/v) lactic acid, pH 3.2, 50 ºC for 48 hr. The pericarp and germ were manually removed with forceps. Germs were dried for 2 days at 37 ºC using a forced air incubator and weighed. Germ yield was calculated as the ratio of the weight of germ recovered to the total

weight of the sample. For oil content tests, about 25 mg of whole seed or germ was put in glass tubes (1×10 cm) fitted with Teflon-lined screw caps. The germ was milled in the glass tubes using a Teflon® Pestle. Then 1 mL of 5% (v/v) sulfuric acid in methanol, 25 µL of BHT (0.2% (v/v) butylated hydroxy toluene in methanol), and 300 µl of toluene were added to each tube. One hundred µg glyceryl triheptadecanoate (Sigma-Aldrich, St. Louis, MO) in 20 µg/µLhexane solution was added as an internal TAG standard for each sample to generate 17:0 fatty acid methyl esters (FAME). The mixture was vortexed for 30 s then heated at 90 ºC for 90 mins. After cooling to RT, 1.5 mL of 0.9 % NaCl (w/v) aqueous solution and 2 mL of hexane were added and samples were mixed. After being centrifuged at 3,000 rpm for 4 min (Medilite, Thermo Electron Corporation, Byron Medical Waltham, MA), the top organic phase containing FAMEs (fatty acid methyl esters) was transferred to a new glass tube. The hexane extraction step was repeated two more times to completely extract FAMEs. Pooled extracts were evaporated under a nitrogen stream and dissolved in 400 µl hexane and analyzed using a GC-2010 dual-FID gas chromatograph (GC) (Shimadzu, Columbia, MD). One µl of each sample (1 µl) was injected into the GC using a Twin-PAL dual rail robot (LEAP Technologies, Carrboro, NC) employing two DB-23 10 m × 0.18 mm i.d. columns (Agilent, Palo Alto, CA) using helium as the carrier gas. The GC was programmed for an initial temperature of 160 ºC for 1.5 min, followed by an increase of 20 ºC/min to 200 ºC and then maintained for a further 1.5 min. Fatty acid composition was analyzed using Shimadzu's GC Solution software, and oil content was calculated based on the internal 17:0 standard.

4.5.7 Polar lipid content

The total lipid was further analyzed for UP6 and UP8 events by quantifying total seed polar lipids. Ten seeds of each event were bulked and ground with a mortar and pestle in liquid N2. Total polar lipids were extracted from the whole meal samples as previously described (Welti et al. 2002). Lipid samples were analyzed on an API 4000 electrospray ionization triple quadrupole mass spectrometer (Applied Biosystems, Foster City, CA) at the Kansas Lipidomics Research Center (Kansas State University, Manhattan, KS).

4.5.8 Immunofluorescent localization of PINs

Immunofluorescent localization of PIN-a and PIN-b was conducted according to the basic methodologies of Feiz et al. (2009), with the following modifications. Corn seeds from both parental and transgenic lines were cut with a scalpel longitudinally through the embryo. One half of each seed sample was fixed in 2% formaldehyde, 2% glutaraldehyde in 100 mM phosphate buffer, pH 7.4 for 48 h. Fixed half seeds were dehydrated with sequential 2 h treatments in 70%, 85% and 100% (v/v) ethanol/water. Dehydrated seed halves were imbibed in paraffin wax, 4 μm sections were made using a microtome and the sections mounted on microscope slides. Seed sections on slides were deparaffinized using three washes (3 min each) in 100% xylene, and one final wash in 1:1 xylene:ethanol (3 min). Xylene was removed by successive 3 min washes in 100%, 90% and 70% ethanol, followed by gentle rinsing of the slides under tap water for 10 min (Hayat, 2002). Seed sections were blocked in 5% skim milk powder in phosphate-buffered saline/0.1% Tween (PBST) overnight at 4°C. The next morning, seed sections were incubated in a 1/100 dilution of Durotest® mouse anti-

puroindoline primary antibody (R-Biopharm AG, Darmstadt, Germany) for 1 h in 2% skim milk powder in PBST (Mohammadi et al. 2007). Seed sections, still adsorbed on the surface of the glass slide, were washed for 10 min x 3 in PBST. Slide sections were incubated in a 1/2000 dilution of goat anti-mouse Alexa Fluor® 488-conjugated secondary antibody (Invitrogen, Burlington, ON) for 1 h in PBST followed by three 10-minute washes in PBST. Finally, the slide sections were overlaid with Prolong Gold® (Invitrogen) anti-fade reagent and cured at room temperature in the dark for 48 h. Bright field and fluorescent microscopy images were viewed with an Axiophot (Zeiss, Oberkochen, Germany) epifluorescent microscope and captured using a Go-3 CMOS digital camera (QImaging, Surrey, B.C.). Overlay images were generated using the 'color to alpha' function of the GNU Image Manipulation Program (GIMP), version 2.6.7.

4.5.9 Starch and protein content

Starch content of whole seed meal was measured via a total starch assay method (AACC Method 76-13, 2003) using a Megazyme kit (Megazyme International, Wicklow, Ireland). Protein content (N × 6.25) of whole seed meal was measured using a Leco FP-2000 (Leco Corp, St. Joseph, MI).

4.5.10 Statistical analysis

All mean values represent testing of three independently derived T1 homozygous positive or negative lines for each event. The means of a transgene positive event was

compared with its respective transgene negative events using a t-test. All statistical tests were conducted using GraphPad InStat software (v. 3.06, San Diego, CA).

4.6 Results

4.6.1 Development of transgenic corn lines containing Pina and Pinb genes

The transgenic plants were created in Hi-II via biolistic transformation at Iowa State University. Herbicide-resistant T_0 lines having both *Pina* and *Pinb* were pollinated with pollen from inbred line B73. Resistant T_1 plants were selfed to produce segregating T_2 progeny. The *Bar* marker gene segregated in a 3:1 ratio in each of the three independent lines presented here (**Table 4.2**). T_3 plants were tested using both PCR for the *Pin* genes as well as herbicide resistance to identify T_2 derived lines that were homozygous positive or negative for the transgene locus. There were seven independent transgenic events generated, and three of them were selected for analysis on the basis of having good vigor, fertility and being expression positive for both PIN-a and PIN-b (results not shown). The three selected homozygous T_3 events carrying *pBAR184*, *pUbiPina* and *pUbiPinb* were designated UP6, UP8 and UP9. All the seeds analyzed were taken from the central cob region of the ears, and all seeds were equilibrated to the same moisture content.

Southern-blot analysis was used to confirm that each line resulted from an independent transformation event. Total genomic DNA was isolated from seedlings and restriction digested with *Hin*dIII which cuts once within the *pUbiPina* and *pUbiPinb* constructs. Each transgenic event had multiple copies of the transgenes, ranging from one to

Table 4.2. *Pina* and *Pinb* PCR and glufosinate herbicide resistance test results for PIN transgenic corn.

Event[a]	T_0 PCR (*Pina*/*Pinb*)[b]	T_2 segregation [c] (resistant/susceptible)	Chi-square (3:1)[d]	T_3 PCR for *Pins* (+/-)[e]	T_3 herbicide for *Bar* (+/-)[e]
B73	-/-	0/5	15	0/5	0/12
UP6	+/+	7/2	0.037	16/0	24/0
UP8	+/+	10/5	0.556	18/0	12/0
UP9	+/+	6/2	0	16/0	10/0

[a]All PCR positive T_0 plants were crossed with untransformed plants to produce T_1 seeds; T_1 hemizygous lines were self-pollinated to produce T_2 seeds. B73 is an inbred line used here as a negative control.
[b]PCR screening was performed using *Pina* or *Pinb* specific primer pairs on samples of genomic DNA from each T_0 plant [c]T_2 progenies were leaf-painted with 0.1% glufosinate. Seedlings which showed minimal adverse effects were scored as resistant.
[e]T_3 seedlings were tested with *Pina* and *Pinb* PCR screening and herbicide tests to obtain T_3 seed pools co-segregating as homozygous positive or negative for the *Pina*, *Pinb* and *Bar* genes. PCR screening was tested on *Bar* positive and negative plants.
[d]Chi-square value tests the fit of resistant/susceptible of T_2 plants to a 3:1 ratio.

likely more than 10 copies (**Figure 4.1**). The banding pattern was unique for each transgenic line, indicating that each arose from an independent integration event.

4.6.2 Pina *and* Pinb *tissue-specific expression levels*

Total RNA was extracted from immature seeds of transgenic lines at 21 DAP and analyzed by probing RNA gel blots with *Pina* and *Pinb* specific probes (**Figure 4.2**). RNA extracted from developing seeds of the soft wheat control variety Heron was used for quantification of signal intensity. No *Pina* or *Pinb* transcripts were detected in control corn inbred B73. The RNA expression levels of *Pina* and *Pinb* in transgenic lines were much less than that of the soft wheat control with *Pin* expression levels in UP8 and UP9 approximately 1-2% that of the soft wheat control variety Heron. UP6 was omitted here, but was analyzed by quantitative real-time PCR (qRT-PCR) along with UP8 and UP9.

The *Pina* and *Pinb* expression in immature embryo, endosperm and leaves of transgenic lines was determined by qRT-PCR (**Figure 4.3**). PCR efficiencies of *Actin 1* and *Pin* genes were approximately equal, which ranged from 1.0 to 1.6. Coefficients of correlation (r^2) to evaluate the quality of the standard curve for reference and target genes were between 0.96 and 0.99. There were no transcripts of either *Pin* detected in untransformed inbred line B73. The highest expression levels observed among the *Pin*-expressing transgenic tissues was for *Pina* in UP8 and UP9 endosperms. The lowest expression level was for *Pina* in endosperm of UP6. The *Pinb* expression pattern was similar to *Pina* for UP8 and UP9 but in UP6 *Pinb* was expressed at a much high level than *Pina*. Expression levels in germ, endosperm and leaf were compared. Endosperm had the highest

Figure 4.1. Southern-blot analysis for three transgenic lines. Genomic DNA isolated from homozygous T_3 seedlings was digested with *Hin*dIII. Untransformed B73 was used as a wildtype control. The ^{32}P-labeled DNA coding sequence of *Pina* or *Pinb* was used as a probe for hybridization. The position of DNA molecular size markers is indicated.

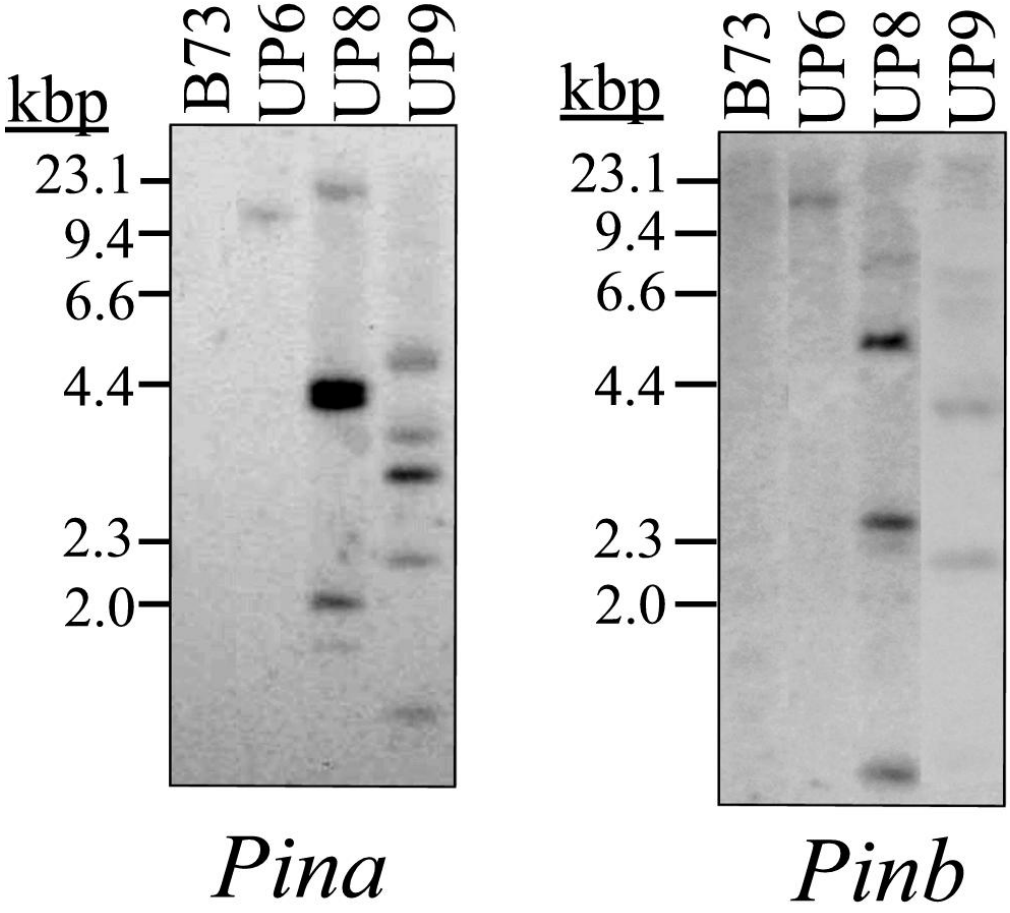

kbp
B73
UP6
UP8
UP9
23.1
9.4
6.6
4.4
2.3
2.0
Pina
kbp
B73
UP6
UP8
UP9
23.1
9.4
6.6
4.4
2.3
2.0
Pinb

Figure 4.2. Northern analysis of *Pina* and *Pinb* expression in transgenic corn seeds. B73 is a negative control. Soft wheat cultivar Heron was used as a positive control and RNA of Heron was loaded in a series of increasing amounts to account for varying signal intensities among transformed lines. The control was exposed 4 hours and the transgenic lines were exposed for one week. A duplicate ethidium bromide stained agarose gel shows discrete bands of rRNA fractionated, indicating a similar loading from lane to lane and a lack of RNA degradation.

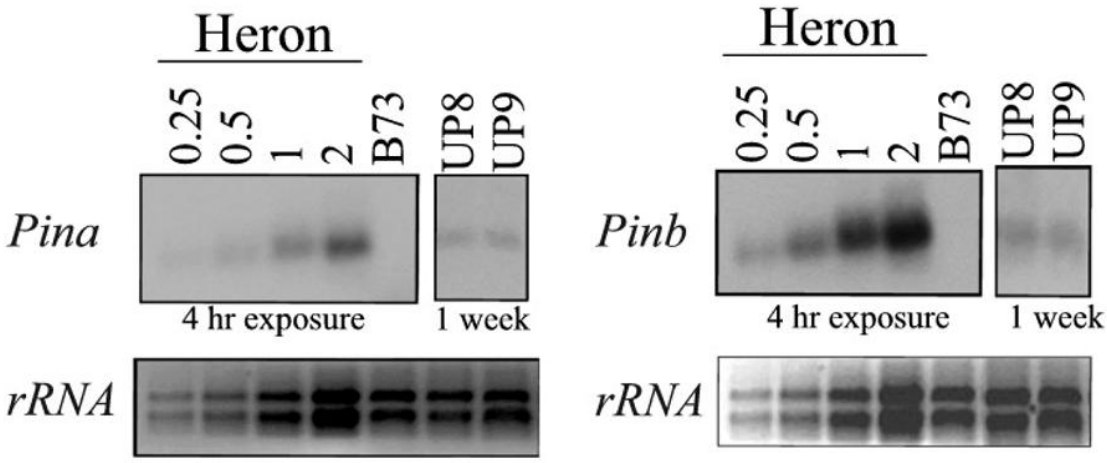
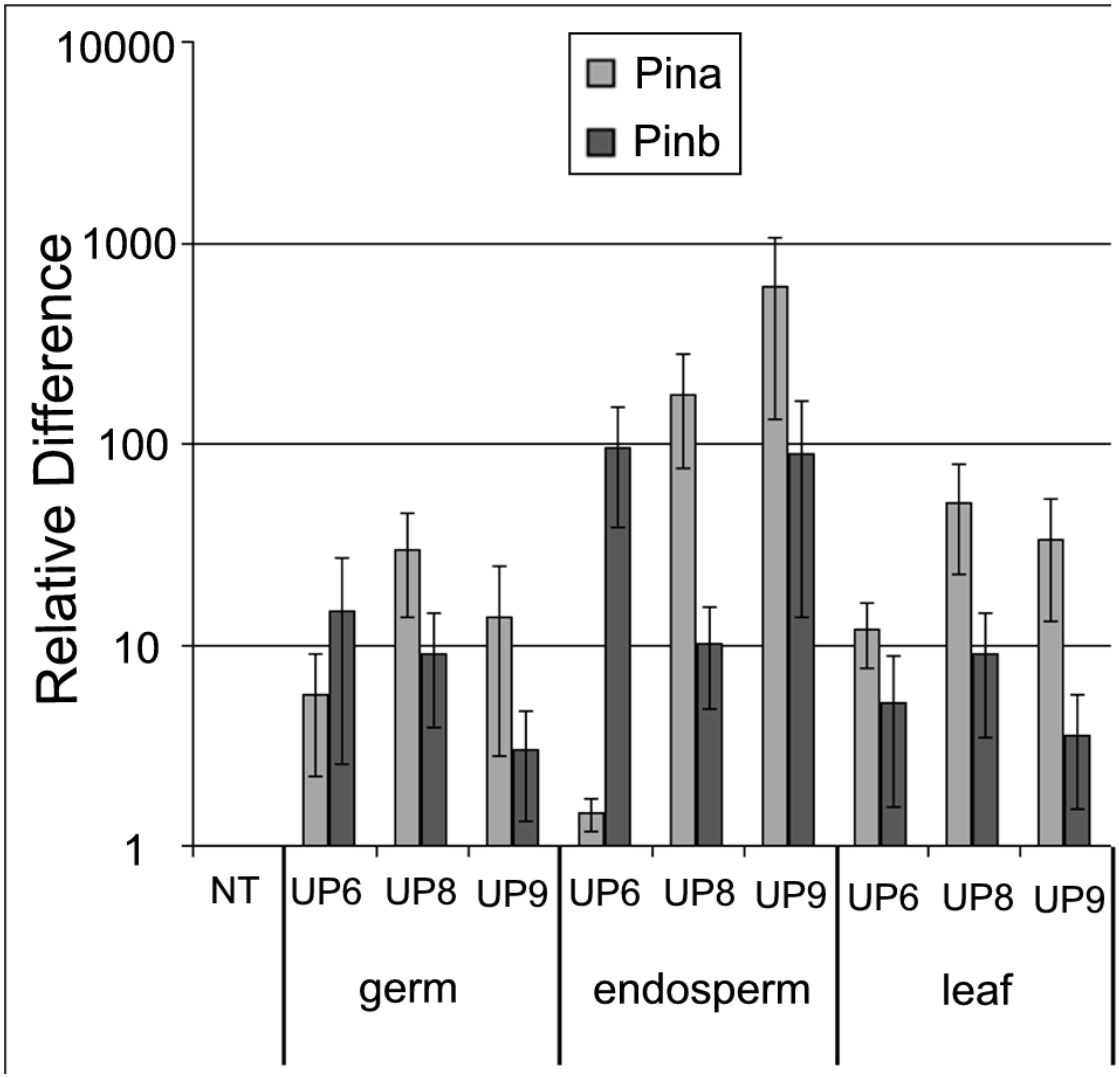

Heron
0.25
0.5
1
2
B73
UP8
UP9
Pina
4 hr exposure
1 week
rRNA
Heron
0.25
0.5
1
2
B73
UP8
UP9
Pinb
4 hr exposure
1 week
rRNA

Figure 4.3. qRT-PCR analysis of *Pina* and *Pinb* expression in transgenic lines. The endogenous corn *Actin* 1 gene was used as a reference. RNA was prepared from 21 DAP embryos, and endosperms as well as leaves of three week post germination seedlings. Untransformed B73 was used as a negative control. The relative difference is between the samples and the B73 control for either *Pina* or *Pinb*. Data shown are on a logarithmic scale of the relative expression level difference. Error bars denote standard deviations.


10000
1000
100
10
1
Relative Difference
Pina
Pinb
NT
UP6
UP8
UP9
UP6
UP8
UP9
UP6
UP8
UP9
germ
endosperm
leaf

expression levels compared to germ and leaves. The *Pina* expression levels in germs were the lowest among the different tissues, and *Pinb* expression levels were similar in germ and leaves. Since the comparative C_t method ($2^{-\Delta\Delta Ct}$) was used for relative quantification, the relative differences are comparisons of homozygous positive transgenic lines and untransformed inbred line B73. Heron does not have the corn *Actin* gene. However, the expression level of Pina and Pinb in Heron was 50 to 1000 times greater than that in the transgenic corn seeds (data not shown).

ELISA was performed to quantify PIN-a and PIN-b in the transgenic corn seeds. Mature homozygous transgene positive or negative control seeds along with seeds of the soft wheat cultivar Heron were analyzed by ELISA for the presence of PIN-a and PIN-b (**Figure 4.4**). The soft wheat Heron, which contains both *Pina* and *Pinb*, consistently gave strong ELISA reactions for both PINs. All three transgenic events UP6, UP8, and UP9 were ELISA positive for both PIN-a and PIN-b (**Figure 4.4**). Seeds from UP8 gave the strongest reactions among the three events for both PIN-a and PIN-b while the PIN-a level in UP6 was lowest among all events. The PIN-a and PIN-b levels in the transgenic corn lines were roughly 1/100 to 1/1000 of Heron PIN-a and PIN-b levels.

4.6.3 Puroindolines increase germ recovery and oil content without modifying seed size

Data for kernel weight, protein, and starch content are presented in **Table 4.3**. Kernel weight, protein and starch content among the transgene positive lines and their corresponding negative control were similar (P>0.05) for UP8 and UP6. However, UP9 had reduced kernel

Figure 4.4. ELISA of T_3 homozygous transgenic corn seed proteins. Each reading represents an ELISA assay in which a well was loaded with 200 times dilution of 100 mg equivalents of total corn or wheat seed flour meal extracts prepared by TX-114 phase partitioning. Readings are representative of duplicate measurements. Sample wells were blanked against homozygous negative corn protein extracts. Soft wheat cultivar Heron was used as a positive control while the negative control was untransformed B73. Error bars denote standard deviations.

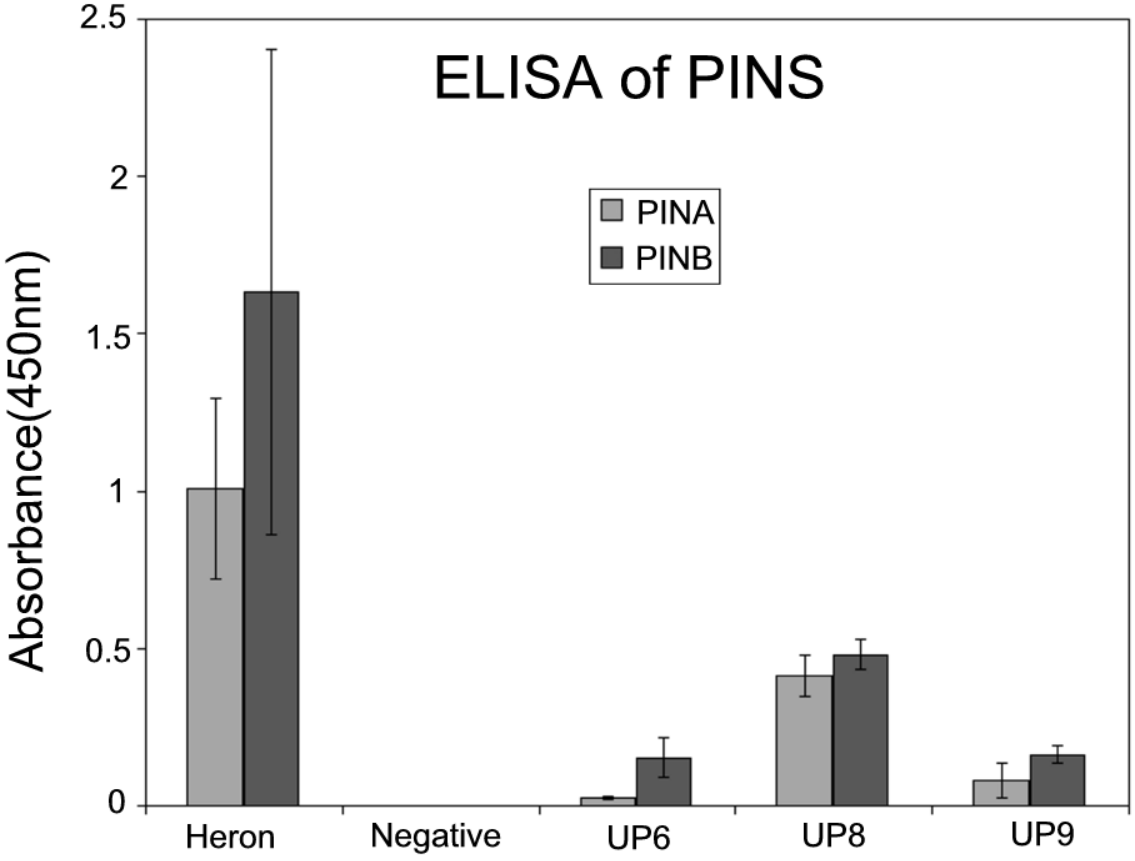

ELISA of PINS
PINA
PINB
Absorbance(450nm)
2.5
2
1.5
1
0.5
0
Heron
Negative
UP6
UP8
UP9

Table 4.3. Means for kernel weight, protein content and starch content of PIN transgenic corn seeds.

Event[a]	Source[b]	Kernel weight (g)	Protein content (%)[c]	Starch content (%)[d]
UP6	+/+	0.248± 0.010(ns)	10.40±0.20 (ns)	68.0±5.66 (ns)
	-/-	0.258±0.004	10.68±0.65	68.9±5.16
UP8	+/+	0.256±0.016 (ns)	11.93± 1.02(ns)	69.1± 2.12(ns)
	-/-	0.250±0.026	11.00±0.52	68.1±2.12
UP9	+/+	0.236±0.046*	12.09±0.04 (ns)	66.1±1.41(ns)
	-/-	0.275±0.054	11.02±1.43	68.5±4.03
B73	wt	0.265±0.025	10.35±0.21	68.2±5.23

*Denotes significance at $P<0.05$ in comparisons of homozygous positive (+/+) transgenic lines and the corresponding homozygous negative (-/-) lines. ns, not significant. wt, wild-type.

[a]The values presented are the average +/- standard deviation of three lines of each event over two growing seasons.

[b]T_2 derived T_3 seeds with co-segregation homozygous positive(+/+) or homozygous negative(-/-) for the transgene locus and *Bar* gene.

[c]Determined by a Leco FP-2000 (N × 6.25)

[d]Determined by a Megazyme total starch assay kit.

weight relative to its control lines (0.236 vs 0.275 g, $P < 0.05$) and tended to have higher protein and reduced starch content.

Germ recovery (germ weight as a percent of seed weight) along with germ and whole seed oil content for the three events is shown in **Table 4.4**. Germ recovery was increased significantly in each of the three transgenic lines relative to their respective controls ($P<0.01$). The relative increase in germ recovery ranged from 22.62% (UP6, 9.54 vs 7.78 %) to 42.26% (UP9, 12.22 vs 8.59%) with the three events averaging a 33.77% increase. Germ oil content was not significantly different for both UP8 and UP9 events ($P>0.05$) but was for UP6 in which the transgenic lines had higher germ oil content than their negative controls (33.62 vs 30.65%). The oil content of seeds was significantly higher for all three events ($P<0.001$). The oil content of inbred B73 was 3.54% and was similar to the transgene negative controls (3.79%). UP8 had the greatest increase in total oil content (4.97 vs 3.96%). The relative increases in oil content for UP6. UP8, UP9 were 25.70%, 25.51%, and 24.47%, respectively. We quantified the germ fatty acid composition of transgenic and control lines by gas chromatography (GC) (**Table 4.5**). Stearic, and oleic acids increased while linoleic acid decreased in UP6 (stearic, and oleic acids increased by 41.06% and 31.31%, respectively, while linoleic acid decreased by 11.86%). The same trend was apparent in UP8 in 2007 (data not shown), but did not reach statistical significance over two years while the fatty acid composition of UP9 was relatively unaltered. The lipid profile of the mature seeds of two events (UP6 and UP8) was analyzed and is shown in **Table 4.6**. Only total LysoPC was different between the transgenic positive and negative lines (47.27% increase for UP6 and 24.63% increase for UP8).

Table 4.4. Means for seed oil traits of PIN transgenic corn.

Event[a]	Source	Germ yield (%)[b]	Germ oil content (%)[c]	Germ oil increase (%)	Seed oil content (%)[d]	Seed oil increase (%)
UP6	+/+	9.54±1.43**	33.62±0.28*	9.69	4.55±0.35***	25.70
	-/-	7.78±0.74	30.65±2.51		3.62±0.29	
UP8	+/+	10.86±1.99***	34.29± 3.25(ns)	3.40	4.97±0.88***	25.51
	-/-	7.96±1.02	33.16±4.49		3.96±0.52	
UP9	+/+	12.22±2.04***	27.47± 0.76(ns)	0.81	4.73±0.23***	24.47
	-/-	8.59± 0.86	27.25±3.95		3.80±0.09	
B73	-/-	7.03±0.04	30.23±1.45		3.54±0.08	

*, **, *** denote significance at $P<0.05$, 0.01, and 0.001, respectively, in comparisons of homozygous positive (+/+) transgenic lines and the corresponding homozygous negative (-/-) lines. ns, not significant.

[a]The values +/- standard deviations presented are the average of three lines of each event over two growing seasons.

[b]Germ yield was determined by a wet milling process and was calculated as the ratio of the weight of germ recovered to the total weight of the sample (Vignaux et al. 2006).

[c]Germ oil content was determined by GC analysis using 17:0 TAG as an internal standard for each sample. It was calculated as the ratio of the total oil content to the weight of germ.

[d]Seed oil content was calculated as the ratio of the oil content to the weight of whole seed meal.

Table 4.5. Fatty acid composition of PIN transgenic corn embryo as determined by GC analysis. The values presented are the average of the percentage of the fatty acid of each event over two years.

Event[a]	Source	Palmitic acid C16:0(%)	Stearic acid C18:0(%)	Oleic acid C18:1(%)	Linoleic acid C18:2(%)	Linolenic acid C18:3(%)
UP6	+/+	10.95 ±0.26	2.13±0.39*	27.97±0.23**	58.04±0.89**	0.93 ±0.01
	-/-	10.46 ±1.27	1.51 ±0.86	21.30 ±0.70	65.85 ±2.43	0.89 ±0.39
UP8	+/+	11.59 ±1.12	1.84 ±0.74*	21.75 ±4.20	64.24 ±6.03	0.60 ±0.01
	-/-	11.03 ±1.33	1.43 ±0.57	21.38 ±5.64	65.07 ±7.69	1.10 ±0.14
UP9	+/+	10.36 ±0.69	1.45 ±0.69	23.54±4.12	63.58 ±5.59	1.08 ±0.07
	-/-	11.04 ±1.53	1.67 ±1.03	24.54 ±3.09	61.60 ±6.02	1.16 ±0.37
B73	-/-	10.01 ±0.64	1.41 ±0.71	23.25 ±3.46	63.97 ±5.09	1.03 ±0.21

*, ** denote significance at $P<0.05$, 0.01, respectively, in comparisons of homozygous positive (+/+) transgenic lines and the corresponding homozygous negative (-/-) lines.

[a]The values +/- standard deviations presented are the average of the percentage of the total fatty acid of each event over two growing seasons.

Table 4.6. Polar lipid composition of PIN transgenic corn seeds.

Description[a]	UP6(+/+)[b] (nmol/mg)	UP6(-/-)[b] (nmol/mg)	UP8(+/+)[b] (nmol/mg)	UP8(-/-)[b] (nmol/mg)
Total DGDG	0.078±0.002	0.087±0.018	0.087±0.003	0.090±0.001
Total MGDG	0.011±0.002	0.014±0.003	0.009±0.003	0.009±0.000
Total PG	0.024±0.002	0.024±0.004	0.034±0.009	0.037±0.001
Total lysoPG	0.000±0.000	0.000±0.000	0.001±0.001	0.000±0.000
Total LysoPE	0.015±0.000	0.012±0.003	0.020±0.002	0.017±0.001
Total LysoPC	0.081±0.002*	0.055±0.015	0.110±0.001*	0.087±0.004
Total PC	1.692±0.204	1.888±0.228	1.762±0.159	1.921±0.009
Total PE	0.406±0.036	0.448±0.139	0.469±0.052	0.508±0.005
Total PI	0.128±0.034	0.069±0.028	0.159±0.030	0.165±0.011
Total PS	0.013±0.001	0.013±0.003	0.013±0.001	0.013±0.001
Total PA	0.011±0.002	0.017±0.004	0.015±0.008	0.014±0.000
Total	2.458±0.277	2.628±0.446	2.679±0.263	2.861±0.007

*, **, *** denote significance at $P<0.05$, 0.01, and 0.001, respectively, in comparisons of homozygous positive (+/+) transgenic lines and the corresponding homozygous negative (-/-) lines. SD, standard deviation.

The bright field micrographs of embryos show that the overall cellular shape of oil bodies remained unaltered between the transgenic line and B73 (**Figure 4.5 A, 4.6 A**). The overall size of the oil bodies between the UP8 and B73 embryos was similar. However, the numbers of oil bodies increased in the transgenic lines (**Figure 4.5, 4.6 C**), and they occupied most of the cell in the transgenic embryos. The corresponding immunofluorescent micrographs show the localization of PINs in endosperm for UP8 line (**Figure 4.6 D, F**). PINs localized to the surface of starch granules in endosperm (**Figure 4.6 F**). However, no PIN was detected in the embryo using immunofluorescent analysis though mRNA expression *Pina* and *Pinb* in the embryo were confirmed by qRT-PCR (**Figure 4.3**).

4.7 Discussion

Approximately 84% of corn seed oil is located in the embryo (Watson 1984) and therefore corn seed oil content is primarily determined by embryo oil content and size. Our experiments indicate that both germ size and oil content were increased via the constitutive expression of puroindolines (**Table 4.4**). No increase in germ size or oil content was seen in transgenic corn where puroindolines were expressed under the control of a γ-Zein endosperm-specific promoter (data not shown, Zhang et al. 2009). Therefore, the germ size increase observed here was likely due to PIN expression in germ, not endosperm specific gene expression. The embryo oil concentration of the transgenic lines was similar to that of the negative controls ($P>0.05$), except in UP6, which had higher germ oil content than the negative control (33.62% vs 30.65%, $P<0.05$) (**Table 4.4**). These results suggest that total oil content in transgenic corn was increased by increasing germ size, not by increasing embryo oil content. In contrast, the key corn oil QTL identified by Zheng et al. (2008) controls

Figure 4.5. Immunolocalization of puroindoline in untransformed B73 mature corn seed sections incubated with Durotest® mouse anti-puroindoline primary antibody and goat anti-mouse Alexa Fluor® 488 secondary antibody. (**a**) Bright field micrograph and (**b**) corresponding immunofluorescent micrograph of embryo tissue. (**c**) Bright field micrograph and (**d**) corresponding immunofluorescent micrograph of endosperm tissue. (**e**) Bright field micrograph of untransformed B73 mature corn (**a**) overlaid with its corresponding immunofluorescent micrograph (**b**). (**f**) Bright field micrograph of untransformed B73 mature corn endosperm tissue (**c**) overlaid with its corresponding immunofluorescent micrograph (**d**). Immunofluorescent images were exposed for 1s. Scale bar = 25 μm.

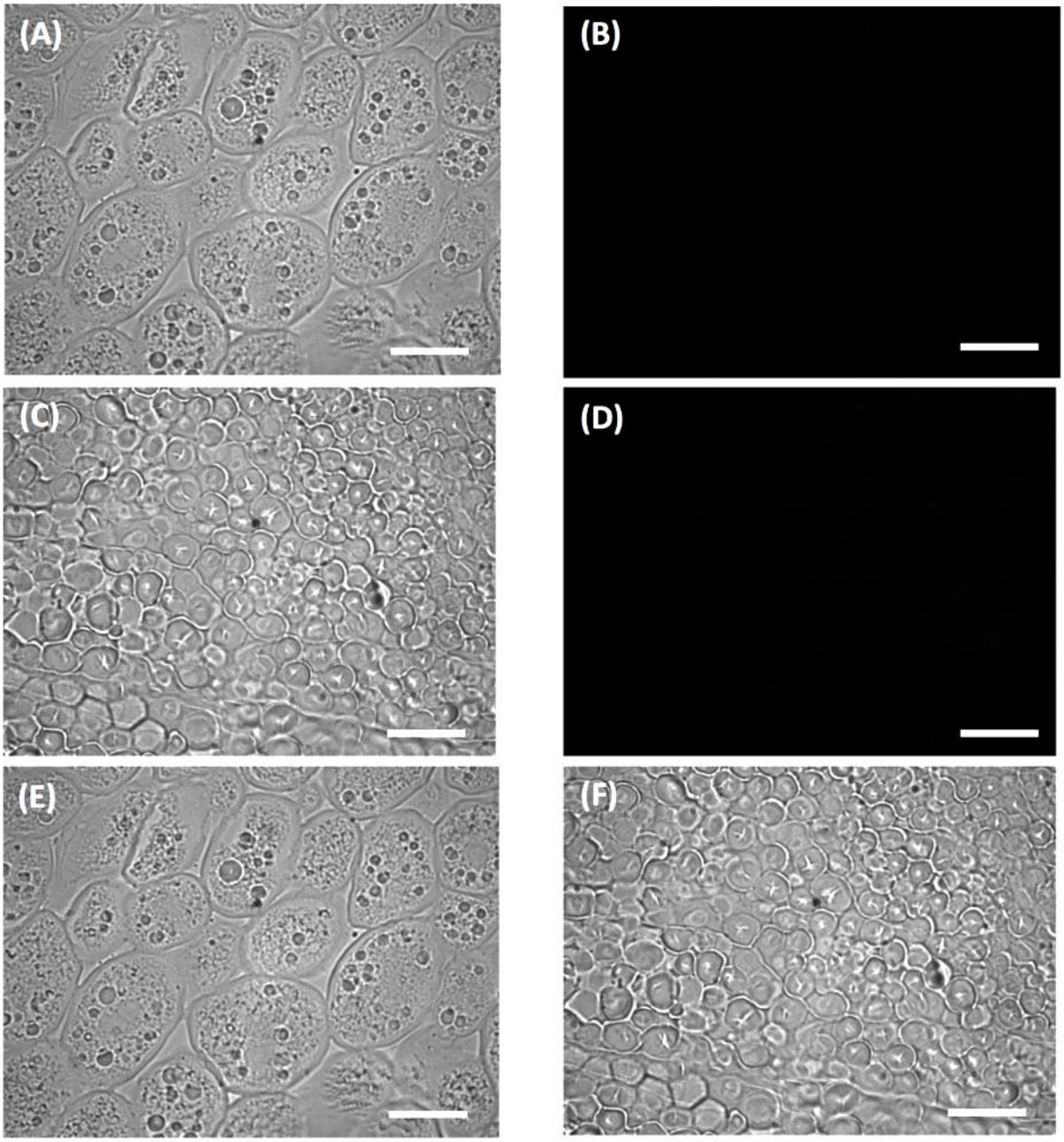
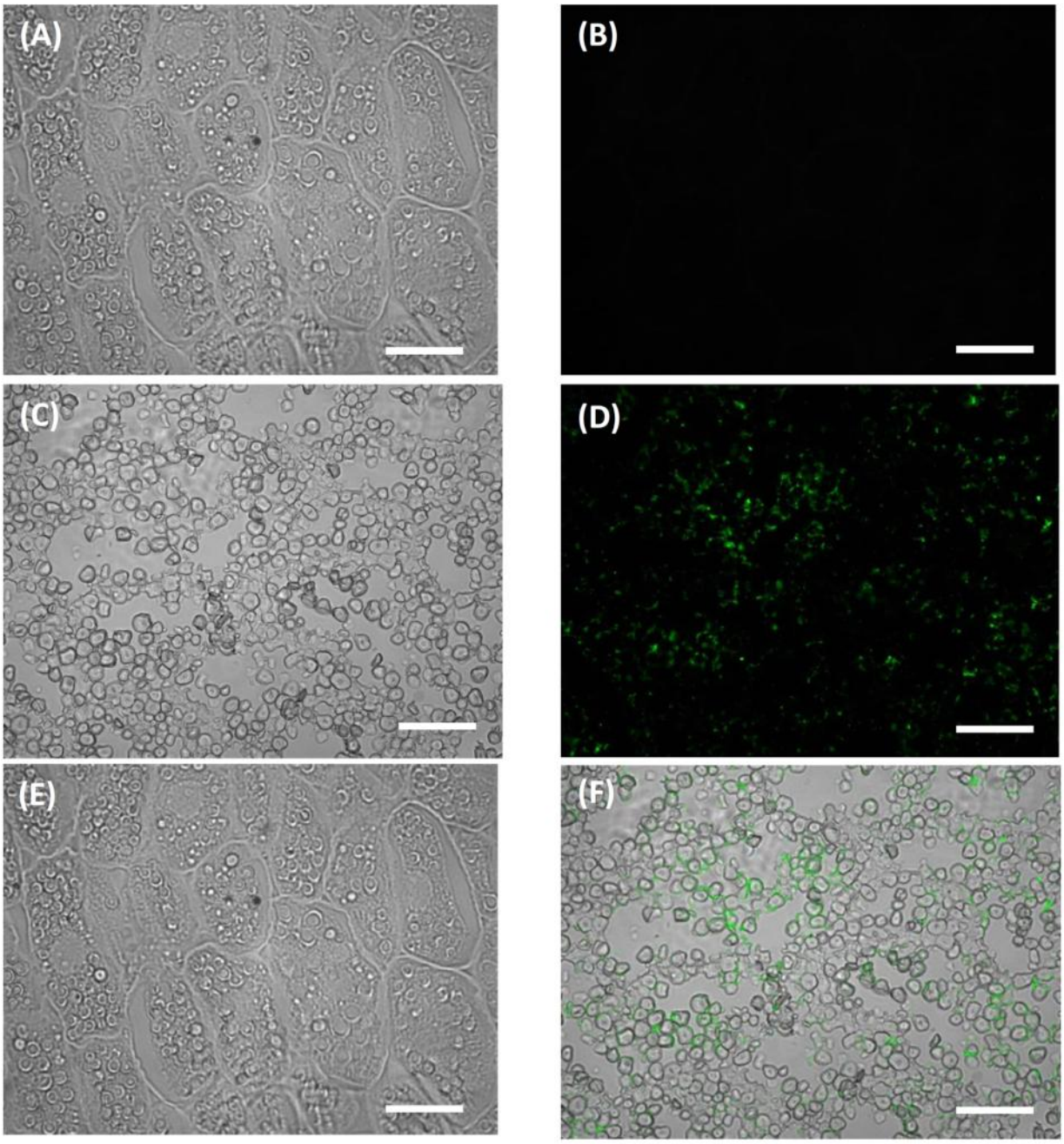

(A)
(B)
(C)
(D)
(E)
(F)

Figure 4.6. Immunolocalization of puroindoline in UP8 mature corn seed sections incubated with Durotest® mouse anti-puroindoline primary antibody and goat anti-mouse Alexa Fluor® 488 secondary antibody. (**a**) Bright field micrograph and (**b**) corresponding immunofluorescent micrograph of embryo tissue. (**c**) Bright field micrograph and (**d**) corresponding immunofluorescent micrograph of endosperm tissue. (**e**) Bright field micrograph of UP8 mature corn embryo tissue (**a**) overlaid with its corresponding immunofluorescent micrograph (**b**). (**f**) Bright field micrograph of UP8 mature corn endosperm tissue (**c**) overlaid with its corresponding immunofluorescent micrograph (**d**). Immunofluorescent images were exposed for 1s. Scale bar = 25 μm.


(A)
(B)
(C)
(D)
(E)
(F)

embryo oil concentration and seed oil content, but not embryo size. Recurrent selection for increased oil content is typically associated with reduced seed size and agronomic yield whereas we did not observe consistently reduced seed size in the PIN transgenic corn in the current study. The typical fatty acid composition of dent corn oil includes 11.0% palmitic acid, 2.0% stearic acid, 24.1% oleic acid, 61.9% linoleic acid, 0.7% linolenic acid and other fatty acids (White et al. 2007). Oil content increases in corn are associated with decreased linoleic acid and increased oleic and saturated acid content (Sniegowski and Baldwin 1954). Increased oleic acid and decreased linoleic acid was also found in the IHO × B73 (Wassom et al. 2008) and the ASKC28IB1 (ASK cycle 28 inbred 1) × PH09B (normal-oil inbred line) populations (Zheng et al. 2008). Increased stearic, oleic, and decreased linoleic acid was indeed observed in this UP6 event. However, that trend was not seen in either UP8 or UP9 (**Table 4.5**).

Our results indicate that cereal seed oil content can be increased by *Pin* expression. However, the mechanism of the increase is not clear. Fatty acids are generally stored in oil bodies and oleosins are the major proteins associated with oil bodies. The importance of the central hydrophobic domain of oleosins for their correct insertion into lipid bodies was addressed by Van Rooijen and Moloney 1995. The tryptophan-rich domain of PINs has a high affinity for binding to lipids (Kooijman et al. 1997; Wall et al. 2010). Ting et al. (1996) found oleosins are constitutively expressed independent of oil content in the IHO and ILO populations. High oil corn had larger and spherical oil bodies and low oil corn has smaller and irregularly shaped oil bodies. In the present transgenic lines, the size of the oil bodies was not significantly different between the seeds expressing PIN and the controls while oil body number was increased (**Figure 4.5, 4.6**), suggesting that constitutive expression of the

puroindoline gene in germ functions to increase oil content differently than effected by oleosins.

Another possible explanation of this unusual effect on germ size lies in the apparent tissue specificity of the two promoters used in our recent transgenic programs, the γ-Zein promoter and the ubiquitin promoter, both from maize (Zhang et al. 2009, present study, respectively). Taken together with our previous transgenic rice line expressing *Pin* (Krishnamurthy et al. 2001) it would be interesting to study the genomic landscapes, flanking the sites of integration, in these various events to better understand whether any genes determining embryo or oil-content characteristics may have been influenced or enhanced (Zaidi et al. 2007), as well as perform global microarray profiling to check for unintended effects. The use of such techniques to identify unintended effects in crops has aided in the determination of the "equivalence" of a transgenic plant to its non-transgenic counterpart, and may illuminate the genetic factors in these UP6, 8, and 9 endosperms versus embryos (Abdeen et al. 2010).

Compared with many previous studies where the goal was to increase or modify oil content via manipulation of the oil biosynthetic pathway (Sharma et al. 2008, Zheng et al. 2008), our work may yield numerous advantages and make it possible to develop a new method to increase the oil content of plant seeds. We predict that over-expression of PIN may be useful in increasing oil content of cereal crops via increasing germ size. Oil content increases in oilseed crops might also be enhanced since the potential for oil storage in oilseed crops is likely greater than in monocots.

4.8 References

Abdeen A., Schnell J., Miki B. 2010. Transcriptome analysis reveals absence of unintended effects in drought-tolerant transgenic plants overexpressing the transcription factor ABF3. BMC Genomics 11:69.

Alexander, D. E. 1988a. High oil corn: breeding and nutritional properties, 43rd Annual Corn and Sorghum Research Conference, pp 97-105.

Alexander, D. E. 1988b. Breeding special nutritional and industrial types. In Corn and Corn Improvement. Sprague GF, Dudley JW (eds). American Society of Agronomy, pp 869-880.

Alrefai, R., Berke, T. G., Rocheford, T. R. 1995. Quantitative trait locus analysis of fatty acid concentrations in maize. Genome 38:894-901.

American Association of Cereal Chemists. 2003. Approved Methods of the American Association of Cereal Chemists. 10th ed. AACC.

Arondel, V., Vergnolle, C., Tchang, F., Kader, J. C. 1990. *Bifunctional lipid-transfer* fatty acid-binding proteins in plants. Mol. Cell. Biochem. 98:49-56.

Beló, A., Zheng, P., Luck, S., Shen, B., Meyer, D. J., Li, B., Tingey, S., Rafalski, A. 2008. Whole genome scan detects an allelic variant of fad2 associated with increased oleic acid levels in maize. Mol. Genet. Genomics 279:1-10.

Bhave, M., Morris, C. F. 2008. Molecular genetics of puroindolines and related genes: Regulation of expression, membrane binding properties and applications. Plant Mol. Biol. 66:221-231.

Capparelli, R., Amoroso, M. G., Palumbo, D., Iannaccone, M., Faleri, C., Cresti, M. 2005. Two plant puroindolines colocalize in wheat seed and in vitro synergistically fight against pathogens. Plant Mol. Biol. 58:857-867.

Christensen, A. H., Quail, P. H. 1996. Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. Trans. Res. 5:213-218.

Clemente, T. E. and Cahoon, E. B. 2009. Fatty acid compositions of corn oils in relation to oil contents of the kernels. J. Am. Oil Chem. Soc. 31:414-416.

Douliez, J. P., Michon, T., Elmorjani, K., Marion, D. 2000. Structure, biological and technological functions of lipid transfer proteins and indolines, the major lipid binding proteins from cereal kernels. J. Cereal Sci. 32:1-20.

Dudley, J. W., Lambert, R. J. 2004. 100 Generations of selection for oil and protein in corn. Plant Breed. Rev. 24. 1:79-110

Feiz, L., Wanjugi, H. W., Melnyk, C. W., Altosaar, I., Martin, J. M., Giroux, M. J. 2009. Puroindolines co-localize to the starch granule surface and increase seed bound polar lipid content. J. Cereal Sci. 50:91-98.

Frame, B. R., Zhang, H. Y., Cocciolone, S. M., Sidorenko, L. V., Dietrich, C. R., Pegg, S. E., Zhen, S. F., Schnable, P. S., Wang, K. 2000. Production of transgenic maize from bombarded type II callus: Effect of gold particle size and callus morphology on transformation efficiency. In Vitro Cell Dev. Biol. Plant 36:21-29

Gautier, M. F., Aleman, M. E., Guirao, A., Marion, D., Joudrier, P. 1994. *Triticum aestivum* puroindolines, 2 basic cystine-rich seed proteins - cDNA sequence-analysis and developmental gene-expression. Plant Mol. Biol. 25:43-57.

Gautier, M. F., Cosson, P., Guirao, A., Alary, R., Joudrier, P. 2000. Puroindoline genes are highly conserved in diploid ancestor wheats and related species but absent in tetraploid *Triticum* species. Plant Sci. 153:81-91.

Giroux, M. J., Morris, C. F. 1997. A glycine to serine change in puroindoline b is associated with wheat grain hardness and low levels of starch-surface friabilin. Theor. Appl. Genet. 95:857-864.

Giroux, M. J., Morris, C. F. 1998. Wheat grain hardness results from highly conserved mutations in the friabilin components puroindoline a and b. Proc. Natl. Acad. Sci. U. S. A. 95:6262-6266.

Giroux, M. J., Sripo, T., Gerhardt, S., Sherwood, J. E. 2003. Puroindolines: Their role in grain hardness and plant defense. In Biotechnology and Genetic Engineering Reviews. Stephen, E. H. (ed). Intercept, Hampshire, pp 277-290.

Greenblatt, G. A., Bettge, A. D., Morris, C. F. 1995. Relationship between endosperm texture and the occurrence of friabilin and bound polar lipids on wheat starch. Cereal Chem. 72:172-176.

Hayat, M. A. 2002. Microscopy, immunohistochemistry, and antigen retrieval methods: For light and electron microscopy. Kluwer Academic Publishers, New York.

Huang, A. H. C. 1992. Oil bodies and oleosins in seeds. Annu. Rev. Plant Physiol. Plant Mol. Biol. 43:177-200.

Jaworski, J., Cahoon, E. B. 2003. Industrial oils from transgenic plants. Curr. Opin. Plant Bio. 6:178-184.

Kooijman, M., Orsel, R., Hessing, M., Hamer, R. J., Bekkers, A. C. A. P. A. 1997. Spectroscopic characterisation of the lipid-binding properties of wheat puroindolines. J. Cereal Sci. 26:145-159.

Krishnamurthy, K., Giroux, M. J. 2001. Expression of wheat puroindoline genes in transgenic rice enhances grain softness. Nat. Biotechnol. 19:162-166.

Krishnamurthy, K., Balconi, C., Sherwood, J. E., Giroux, M. J. 2001. Increased tolerance to fungal diseases of rice plants transformed with puroindoline genes. Mol. Plant Microbe. Interact. 14:1255-1260.

Laurie, C. C., Chasalow, S. D., LeDeaux, J. R., McCarroll, R., Bush, D., Hauge, B., Lai, C.,

Clark, D., Rocheford, T. R., Dudley, J. W. 2004. The genetic architecture of response to long-term artificial selection for oil concentration in the maize kernel. Genetics 168:2141-2155.

Le, B., Blochet, J. E., Désormeaux A, Marion D, Pézolet M. 1996. Determination of the secondary structure and conformation of puroindolines by infrared and Raman spectroscopy. Biochemistry 35:12712-12722.

Leprince, O., van Aelst, A. C., Pritchard, H. W., Murphy, D. J. 1997. Oleosins prevent oil-body coalescence during seed imbibition as suggested by a low-temperature scanning electron microscope study of desiccation-tolerant and -sensitive oilseeds. Planta 204:109-119.

Li, Y., Beisson, F., Pollard, M., Ohlrogge, J. 2006. Oil content of arabidopsis seeds: The influence of seed anatomy, light and plant-to-plant variation. Phytochemistry 67:904-915.

Marion, D., Gautie, M. F., Joudrier, P., Ptak, M., Pezolet, M., Forest, E., Clark, D. C.,

Broekaert, W. 1994. Structure and function of wheat lipid binding proteins. In Wheat kernel proteins: Molecular and functional aspects. al Cimino, S. M. (ed). Proc. Int. Mtg., Universita Degli Studi della Tuscia, Viterbo, pp 175-180.

Marion, D., Bakan, B., Elmorjani, K. 2007. Plant lipid binding proteins: Properties and applications. *Biotechnol. Adv.* 25:195-197.

Mohammadi, M., Zaidi, M. A., Ochalski, P. A., Tanchak, M. A., Altosaar, I. 2007. Immunodetection and immunolocalization of tryptophanins in oat seeds. Plant Sci. 172:579-587.

Morris, C. F., Lillemo, M., Simeone, M. C., Giroux, M. J., Babb, S. L., Kidwell, K. K. 2001. Prevalence of puroindoline grain hardness genotypes among historically significant North American spring and winter wheats. Crop Sci. 41:218-228.

Poneleit, C. G., Alexander, D. E. 1965. Inheritance of linoleic and oleic acids in maize. Science 147:1585-1586.

Purkrtova, Z., Jolivet, P., Miquel, M., Chardot, T. 2008. Structure and function of seed lipid body-associated proteins. Comptes Rendus Biologies 331:746-754.

Riede, C. R. and Anderson, J. A. 1996. Linkage of RFLP markers to an aluminum tolerance gene in wheat. Crop Sci., 36, 905-909.

Sharma, N., Anderson, M., Kumar, A., Zhang, Y., Giblin, E. M., Abrams, S. R., Zaharia, L. I., Taylor, D. C., Fobert, P. R. 2008. Transgenic increases in seed oil content are associated with the differential expression of novel *Brassica*-specific transcripts. BMC Genomics 9:619.

Siloto, R. M. P., Findlay, K., Lopez-Villalobos, A., Yeung, E. C., Nykiforuk, C. L., Moloney, M. M. 2006. The accumulation of oleosins determines the size of seed oilbodies in Arabidopsis. Plant Cell. 18:1961-1974.

Sniegowski, M. S., Baldwin, A. R. 1954. Fatty acid compositions of corn oils in relation to oil contents of the kernels. J. Am. Oil Chem. Soc. 31:414-416.

Song, X. F., Song, T. M., Dai, J. R., Rocheford, T. R. 2004. QTL mapping of kernel oil concentration with high-oil maize by SSR markers. Maydica 49:41-48.

Ting, J. T. L., Lee, K., Ratnayake, C., Platt, K. A., Balsamo, R. A., Huang, A. H. C. 1996. Oleosin genes in maize kernels having diverse oil contents are constitutively expressed independent of oil contents. Planta 199:158-165.

Van Rooijen, G. J., Moloney, M. M. 1995. Plant seed oil-bodies as carriers for foreign proteins. Bio/Technology 13:72-77.

Vignaux, N., Fox, S. R., Johnson, L. A. 2006. A 10-g laboratory wet-milling procedure for maize and comparison with larger scale laboratory procedures. Cereal Chem. 83:482-490.

Wall, M. L., Wheeler, H., Huebsch, M. P., Smith, J. C., Figeys, D. Altosaar I. 2010. The tryptophan rich domain of puroindoline is directly associated with the starch granule surface as judged by tryptic shaving and mass spectrometry. J. Cereal Sci. 52:115-120.

Wanjugi, H. W., Hogg, A. C., Martin, J. M., Giroux, M. J. 2007. The role of puroindoline A and B individually and in combination on grain hardness and starch association. Crop Sci. 47:67-76.

Wassom, J. J., Mikkelineni, V., Bohn, M. O., Rocheford, T. R. 2008. QTL for fatty acid composition of maize kernel oil in Illinois high oil × B73 backcross-derived lines. Crop Sci. 48:69-78.

Watson, S. A. 1984. Corn and sorghum starches: Production. In Starch: Chemistry and technology. Whistler, R. L, BeMiller, J. N, Paschall, E. F. (eds). Academic Press, Orlando, pp 417-418.

Weber, E. J. 1987. Lipids of the kernel. In Corn: Chemistry and technology. Watson, S. A., Ramstad, P. E. (eds). AACC, St. Paul, MN, pp 311-350.

Welti, R., Li, W., Li, M., Sang, Y., Biesiada, H., Zhou, H., Rajashekar, C. B., Williams, T. D., Wang, X. 2002. Profiling membrane lipids in plant stress responses: Role of phospholipase D{alpha} in freezing-induced lipid changes in Arabidopsis. J. Biol. Chem. 277:31994-32002.

White, P. J., Pollak, L. M., Duvick, S. 2007. Improving the fatty acid composition of corn oil by using germplasm introgression. Lipid Technol. 19:35-38.

Wiley, P. R., Tosi, P., Evrard, A., Lovegrove, A., Jones, H. D., Shewry, P. R. 2007. Promoter analysis and immunolocalisation show that puroindoline genes are exclusively expressed in starchy endosperm cells of wheat grain. Plant Mol. Biol. 64:125-136.

Zaidi, M. A., Cheng, X. Y., Altosaar, I. 2007. Characterization of left-border flanking sequences of T-DNA integration in transgenic rice (*Oryza sativa* L.) expressing cry1Ab. Cereal Res. Comm. 35:1375-1383.

Zhang, J., Martin, J. M., Beecher, B., Morris, C. F., Hannah, L. C., Giroux, M. J. 2009. Seed-specific expression of the wheat puroindoline genes improves maize wet milling yields. Plant Biotech. J. 7:733-743.

Zheng, P. Z., Allen, W. B., Keith, R., Williams, M. E., Zhang, S., Li, J., Glassman, K., Ranch, J., Nubel, D., Solawetz, W., Bhattramakki, D., Llaca, V., Deschamps, S., Zhong, G., Tarczynski, M. C., Shen, B. 2008. A phenylalanine in DGAT is a key determinant of oil content and composition in maize. Nat. Genet. 40:367-372.

Chapter 5

General Discussion

5.1 Summary of results

The starch granule surface is one of the last lines of protection for the host plant against invading bacterial and fungal pathogens. PIN is an antimicrobial protein which strongly binds the starch granule surface in the endosperm of some cereal plants (**Figure 2.2**). This thesis provides the first direct evidence that wheat PIN strongly binds to wheat starch granule surfaces due to its tryptophan-rich domain (**Figure 2.3**).

Also present at the starch granule surface are other wheat proteins, some of which are also involved in antimicrobial activity (**Tables 3.2 and 3.3**). These proteins are further accompanied by the protein remnants of known pathogens (**Table 3.1**). It is hypothesized that these pathogen remnants are the result of a failed pathogen attack(s) on the wheat grain which become 'fixed' in the desiccated tissue as a protein 'fingerprint' of the intruders.

Finally, we show that PIN will localize to the surface of starch granules in transgenic corn (**Figure 4.5**), a cereal which normally does not express PIN. In addition, the expression of PIN in these plants increases the seed oil content (**Table 4.4**).

5.2 Extraction of puroindoline from starch granule surfaces

PIN is strongly bound to the starch granule surface (**Figure 2.2**). To extract PIN from wheat starch granules, we employed the use of a solvent previous shown to be effective at extracting PIN (Hogg et al. 2004). This solvent has advantages over non-ionic detergents for

PIN extraction. Primarily, the use of isopropanol instead of a detergent allows for evaporation of the alcohol, as opposed to requiring precipitation of extracted proteins and additional polishing (i.e. downstream processing). Not only was the 50% isopropanol, 50mM NaCl extraction solution effective at removing PIN from the surface of starch granules, it effectively removed all protein from the starch granule surface. Following protein extraction with 50% isopropanol / 50 mM NaCl, the decanted starch granules were incubated in SDS for 1h. Bradford assays of this extract demonstrated that no protein remained at the starch granule surface following the initial isopropanol/salt extraction (data not shown).

We have shown that PIN-null (durum) and mutant-PIN (hard) varieties of wheat had similar total protein extractable with 50% isopropanol, 50 mM NaCl compared to wilt-type PIN varieties (**Table 2.1**). An explanation for this may be that for durum and hard wheat, the reduced adhesion between the starch granule surface and surrounding intracellular matrix increases the proportion of 'contaminating' protein from the protein matrix during starch granule extraction. This is in contrast to the case for soft wheat, where the reduced adhesion between the starch granule surface and surrounding intracellular matrix due to the presence of PIN may result in reduction of 'contaminating' protein from the protein matrix. The reduced protein matrix contamination on soft wheat starch granules is compensated for by the increased starch granule surface PIN.

5.3 Puroindoline - starch granule binding properties

Recently, the study of PIN and its tryptophan-rich domain has greatly increased. However, the process by which PIN becomes associated with the starch granule surface during seed desiccation and whether it is bound to surface lipids or carbohydrate is not

entirely clear. Based on current evidence, the proposed model of PIN deposition is as follows: the PIN proprotein signal peptide directs PIN to the endoplasmic reticulum where it eventually becomes deposited into developing intracellular protein bodies. Through a mechanism not yet understood, during grain dessication, PIN becomes associated with the starch granule surface (Dubreil et al. 1998). It may be that PIN is associated with lipids during its initial deposition within the intracellular protein matrix and, as the seed desiccates, the other proteins in the intracellular matrix coalesce, forcing PIN to the periphery where it has the opportunity to contact the starch granule surface lipids and, through interaction mediated by its TRD, becomes strongly bound.

There is little evidence to suggest that PIN can bind carbohydrate polymers, nor is there a plausible biochemical mechanism based upon the PIN sequence. PIN's ability to bind lipids, lipid membranes and stabilize lipid foams, as well as the lipid membrane-binding and antimicrobial properties of the TRD all suggest that PIN binds associates with the surface of starch granules by binding lipid remnants. Though the lipid content at the starch granule surface has been studied in some detail (Panozzo et al. 1993; Konopka et al. 2005), more study is needed to assess the nature of the lipid remnants at the starch granule surface. The starch granule surface supports a 'loosely' bound lipid fraction, the presence of which contributes to grain texture and is influenced by PIN . However, it is difficult to differentiate between lipids which are 'tightly' bound to the starch granule surface versus those which are embedded within the starch granule interior and do not actually play a role in PIN binding. Milling conditions and solvent extraction choice can greatly influence measured lipid content (Konopka et al. 2005). The result of water-washing on starch granules PIN content (**Figure**

2.1), seen as a roughly one-half decrease in PIN content after x8 and up to x20 washes, may be due to PIN's indiscriminate association with both loosely-bound and tightly-bound lipids.

5.4 Significance of puroindoline for biotechnology

PIN and PIN-like proteins are involved in biological functions and phenotypic effects with potential for use in biotechnology applications. Several patents already exist regarding its use or expression in a variety of plant systems. For example, such patents include those on expression of PIN for improvement of grain texture (Freeman 2003; Morris 1998) and antimicrobial resistance (Giroux 2000; Giroux 2003). One potential for puroindoline may be as an affinity tag for efficient recombinant fusion protein expression in cereals. The starch granules could act as natural affinity beads, to which the puroindoline fusion protein would bind. Recovery of the recombinant protein would involve isolation of the starch granules and separation of the fusion partner from the puroindoline tag (Altosaar 2006). Overall, puroindoline's physical characteristics, antimicrobial activity, association with starch granules and influence on grain texture make it an interesting target of research for many aspects of biotechnological and industrial applications.

5.5 Analysis of mass spectrometric data of host-pathogen interactions

I have demonstrated the first use of mass spectrometry to identify pathogen proteins at the starch granule surface in wheat. However, the study reported here was not the first to perform mass spectrometric analysis of endosperm tissue. Other studies have looked at the proteome of commercial wheat starch (Kasarda et al. 2008), the proteome of wheat and

potato amyloplast and amyloplast membranes (Andon et al. 2002; Stensballe et al. 2008), and the proteome of starch granule surface protein extracts from wheat (Skylas et al. 2005). Interestingly, these studies did not identify proteins from microorganisms. Reason for this comes from the data collection and analysis. Instead of searching the mass spectra data against a protein database from all organisms the mass spectra data in these other studies was searched against a database of proteins from green plants only. By expanding the search to proteins from all organisms, as we have done, it was possible to identify proteins from microorganisms that would normally have been precluded from the results.

Mass spectrometry, in general, has the potential to become one of the most important techniques for probing host-pathogen interactions. Currently, the main challenge to overcome is the functional assignment of identified proteins. Transcriptional changes alone are not sufficient to inform on cellular regulatory mechanisms, since post-transcriptional regulatory processes that can alter the quantity of active protein. Therefore, a complementary approach including, proteome-based expression profiling is needed to obtain a full picture of all regulatory elements involved. This is apparant with PR proteins: they are traditionally identified solely by their upregulation during pathogenesis, regardless of their fuction or whether they play a direct or indirect role in host defense. With the expansion of known PR proteins and the lagging classification of these proteins based upon function or localization, the PR desigation may become obsolete.

5.6 Identification of pathogens versus endophytes

Endogenous microorganisms can coexist with plants intercellularly within various tissues (Clay 1990; Hallmann et al. 1997) and without inducing a hypersensitive response

(Hallmann et al. 1997). Generally, endophytic microorganisms are defined as those that are not killed or removed during surface sterilization procedures. Based upon this definition, there is a case for defining the microorganisms identified in **Table 3.1** as endophytic. However, there are two considerations which limit this conclusion. The first is that no microorganisms were cultured from the wheat starch granule samples, nor is there any data to suggest that the microorganisms are viable. Unless the organism is capable of forming endospores, it can be reasonably assumed that they are not viable. However, none of the bacteria identified (*Agrobacterium tumefaciens, Pectobacterium carotovorum, Xanthomonas axonopodis and X. oryzae*) are known to form endospores. The second argument is that none of these microorganisms have ever been shown to be endophytic of wheat endosperm. Only *A. tumefaciens* has been shown to be endophytic. *A. tumefaciens* is a common soil bacterium and has been shown to endophytically colonize the roots of *Melilotus dentatus*, a species of sweet clover (Wang et al. 2006). It is not known whether *A. tumefaciens* can endophytically colonize the roots of wheat or any other wheat tissue. In conclusion, it may be reasonable to consider the microorganisms from which peptides were identified (**Table 3.1**), as pathogens only.

5.7 Mass spectrometry for assessment of 'pathogen load' in grain

The finding that the desiccated endoperm of wheat, and maybe other cereals, can 'fix' the protein remnants of failed pathogen attack (**Table 3.1**) opens the possibility of using mass spectrometry to assess the 'pathogen load' that the plant material had acquired during growth, development, storage and transport. Identification of pathogens usually involves culturing the invasive pathogens in media and identifying them morphologically (Lievens and Thomma

2005). However, these methods would be unable to identify very low quantities of pathogen or unviable pathogens. Since the proteins are highly stable in the dry grain, this technique may be useful for analyzing the pathogen load of crops grown in different regions or climates. Although it has already been argued that the pathogens identified in the present study are not endophytic, there is no limitation on this method to identify endophytic organisms in endosperm tissue. In cases of assessing 'pathogen load', it is not strictly necessary to assess the starch granule surface only. Any tissue may be assessed, though the quantity of proteins from whole tissues may significantly overshadow the low quantity of proteins remaining, post pathogen infection. The significance of the study is still related to the fact that by restricting analysis to the starch granule surface only, it is possible to focus on those proteins which may be involved in starch granule host defense or pathogen attack.

5.8 Expression levels of PIN mRNA and protein in transgenic corn

Northern analysis of RNA extracted from whole corn 21 DAP (**Figure 4.2**) showed the presence of *Pina* and *Pinb* transcripts in the seeds of transgenic corn lines UP8 and UP9 (no data for UP6). To support the northern analysis, qRT-PCR was performed on RNA extractions from three separate tissues: leaf, embryo and endosperm (**Figure 4.3**) for all three transgenic corn lines (UP6, UP8 and UP9). The analysis confirmed that the *Pina* and *Pinb* transcripts were present in all three tissues analyzed from all three transgenic corn lines. In addition, it showed that the levels of *Pina* and *Pinb* transcript not only varied among the tissues but the levels of *Pina* and *Pinb* transcript varied within each tissue. For example, the mean level of *Pinb* transcript in the developing endosperm of UP6 was 100-fold, relative to negative control, which is 60 times more than the level of *Pina* transcript in this same tissue.

This difference in *Pina* versus *Pinb* transcript levels in UP6 endosperm is in stark contrast to the levels of *Pina* and *Pinb* transcript in UP8 endosperm, where the *Pina* transcript level was 250-fold, approximately 25 times higher than the level of *Pinb* transcript. If transcript level is an indicator for protein expression, then the endosperm tissue of UP6 should have had higher levels of PIN-b protein than PIN-a and vice versa for line UP8. In fact, the protein levels of PIN-a and PIN-b from UP6 whole seed (total from embryo and endosperm), determined by ELISA (**Figure 3.4**), roughly correspond to the total transcript levels found in the embryo and endosperm. However, the association of PIN protein levels to *Pin* transcript levels observed for UP6 was not seen for UP8 and UP9. Other studies attempting to determine whether there is a correlation between mRNA levels and protein levels are generally inconclusive (Greenbaum et al. 2003).

The difference in *Pina* and *Pinb* transcript levels found in all three transgenic lines was interesting considering that expression of both *Pin* genes was under the control of the ubiquitin promoter (**Figure 4.3**). The ubiquitin protein plays a vital house-keeping function (Hershko and Ciechanover 1998) and the promoter for the *Ubi* gene was originally thought to be constitutive (Christensen and Quail 1996). Constitutive expression is desirable for high-level expression in all tissues of a transgenic organism and the maize ubiquitin promoter has been used to drive expression of transgenes in monocotyledonous plants for nearly 15 years (Christensen and Quail 1996; Peremarti et al. 2010). However, more recent studies on tissue-specificity of the maize ubiquitin promoter suggests that tissue-specific expression of transgenes under the control of the maize ubiquitin promoter may vary (Rooke et al. 2000).

5.9 Expression pattern of puroindoline in trangenic cereals

To localize PIN in the grain of transgenic corn lines UP6, UP8 and UP9, microscopy sections of these samples were probed with an anti-PIN antibody and AlexaFluor^{488} secondary antibody. Immunofluorescent imaging of the endosperm tissue of UP8 (**Figure 4.5**) corroborates the ELISA data (**Figure 4.4**) for this sample and confirms that the PIN protein is expressed in the grain of this transgenic line. However, no fluorescence was observed in the embryo of UP8 or any grain-specific tissue of UP6 or UP9. The reason for this may be that the higher expression of PIN in UP8 and its localization at the starch granule surface surpassed the sensitivity limits of the microscopy technique. Protein levels in the endosperm and embryo of UP6 and UP9 may have been below the detection limit. Since the ELISA quantified PIN for whole grain, it was not certain if PIN was expressed in the embryo. In addition to confirmation of PIN expression, the immunofluorescent microscopy can reveal the pattern of PIN localization in the seed endosperm. The high fluorescence around the starch granules in the UP8 endosperm is similar to the pattern of PIN localization in the mature endosperm of wheat (Feiz et al. 2009). The fluorescence creates a 'halo' pattern, resulting from high PIN concentration at the starch granule surface and no PIN within the starch granule interior. There are no cases of PIN or a PIN homologues localizing to the starch granule interior, but fluorescence may be seen in the center of the 'halo' if the intact starch granule is viewed using fluorescence techniques. However, preparation of endosperm samples generally results in sectioning of the granule granule itself, exposing the interior portion only. This pattern has been observed for PIN in wheat (Feiz et al. 2009) and tryptophanin in oat (Mohammadi et al. 2007).

5.10 Future work

An issue with searching for members of molecular systems of any kind is setting a selection criteria that will reduce the overall pool of potential positive 'hits' against the background of any biological system. Selection of the starch granule surface was performed under the hypothesis that any discovery would have a higher likelihood of being related to starch granule metabolism, defense or, in the case of pathogens, attack and degradation.

The tryptophan-rich domain of PIN is similar to known antimicrobial peptides (Jing et al. 2003). Research by Jing et al. (2004) and others (Evrard et al. 2008) demonstrated that single amino acid differences in the tryptophan-rich motif of these peptides can have an effect on their affinity for lipid membranes, affecting their antimicrobial activity (Clifton et al. 2007, 2008). The properties of PIN may be significantly altered by modifying residues at and around the tryptophan-rich domain to improve its lipid binding, texture-altering and antimicrobial effects. For example, it has been shown that substitution of the lysine following tryptophan doublets with arginine will increase the antimicrobial activity of the tryptophan-rich domain peptide from PIN (Jing et al. 2004; Chan et al. 2006). The PIN homolog avenoindoline (tryptophanin) from oat is very similar to PIN but possess a single lysine to arginine substitution within the tryptophan-rich domain (Tanchak et al. 1998). Whether this change significantly alters the properties and phenotypic effects of this protein relative to its homologs in other cereals is not known and mutational analysis of the tryptophan-rich domain in wheat and other cereals will be very informative.

5.11 References

Altosaar, I. 2006. Puroindolines as fusion-protein carriers in molecular pharming. Canadian Patent 2555137.

Andon, N. L., Hollingworth, S., Koller, A., Greenland, A. J., Yates, J. R.,3rd, Haynes, P. A. 2002. Proteomic characterization of wheat amyloplasts using identification of proteins by tandem mass spectrometry. Proteomics 2:1156-1168.

Chan, D. I., Prenner, E. J., Vogel, H. J. 2006. Tryptophan- and arginine-rich antimicrobial peptides: Structures and mechanisms of action. Biochim. Biophys. Acta-Biomembr. 1758:1184-1202.

Christensen, A. H., Quail, P. H. 1996. Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. Transgenic Res. 5:213-218.

Clay, K. 1990. Fungal Endophytes of Grasses. Annu. Rev. Ecol. Syst. 21:275-297.

Clifton, L. A., Green, R. J., Hughes, A. V., Frazier, R. A. 2008. Interfacial structure of wild-type and mutant forms of puroindoline-b bound to DPPG monolayers. J. Phys. Chem. B 112:15907-15913.

Clifton, L. A., Lad, M. D., Green, R. J., Frazier, R. A. 2007. Single amino acid substitutions in puroindoline-b mutants influence lipid binding properties. Biochemistry 46:2260-2266.

Evrard, A., Lagarde, V., Joudrier, P., Gautier, M. 2008. Puroindoline-a and puroindoline-b interact with the *Saccharomyces cerevisiae* plasma membrane through different amino acids present in their tryptophan-rich domain. J. Cereal Sci. 48:379-386.

Feiz, L., Wanjugi, H. W., Melnyk, C. W., Altosaar, I., Martin, J. M., Giroux, M. J. 2009. Puroindolines co-localize to the starch granule surface and increase seed bound polar lipid content. J. Cereal Sci. 50:91-98.

Freeman, J. P. 2003. Method for improving corn starch extractability. US Patent 7,279,616.

Giroux, M. J., Sherwood, J. E., Krishnamurthy, K., Morris, C. F. 2000. Transgenic plants expressing puroindolines and methods for producing such plants. US Patent 6,600,090.

Giroux, M. J., Sherwood, J. E., Krishnamurthy, K., Morris, C. F. 2003. Transgenic plants expressing puroindolines and methods for producing such plants. US Patent 7,425,669.

Greenbaum, D., Colangelo, C., Williams, K., Gerstein, M. 2003. Comparing protein abundance and mRNA expression levels on a genomic scale. Genome Biol. 4:117.

Hallmann, J., QuadtHallmann, A., Mahaffee, W. F., Kloepper, J. W. 1997. Bacterial endophytes in agricultural crops. Can. J. Microbiol. 43:895-914.

Hershko, A., Ciechanover, A. 1998. The ubiquitin system. Annu. Rev. Biochem. 67:425-479.

Jing, W. G., Demcoe, A. R., Vogel, H. J. 2003. Conformation of a bactericidal domain of puroindoline a: Structure and mechanism of action of a 13-residue antimicrobial peptide. J. Bacteriol. 185:4938-4947.

Jing, W., Schibli, D. J., Nguyen, L. T., Vogel, H. J. 2004. Tryptophan-rich antimicrobial peptides: The mechanism of action of tritrpticin and puroindoline-derived peptides. J. Pept. Sci. 10:184-184.

Kasarda, D. D., Dupont, F. M., Vensel, W. H., Altenbach, S. B., Lopez, R., Tanaka, C. K., Hurkman, W. J. 2008. Surface-associated proteins of wheat starch granules: Suitability of wheat starch for celiac patients. J. Agric. Food Chem. 56:10292-10302.

Konopka, W., Rotkiewicz, D., Tanska, M. 2005. Wheat endosperm hardness. Part II. Relationships to content and composition of flour lipids. Eur. Food Res. Technol. 220:20-24.

Li, X. X., Franceschi, V. R., Okita, T. W. 1993. Segregation of storage protein messenger RNAs on the rough endoplasmic reticulum membranes of rice endosperm cells. Cell 72: 869–879.

Lievens, B., Thomma, B. P. H. J. 2005. Recent developments in pathogen detection arrays: Implications for fungal plant pathogens and use in practice. Phytopathology 95:1374-1380.

Mohammadi, M., Zaidi, M. A., Ochalski, A., Tanchak, M., Altosaar, I. 2007. Immunodetection and immunolocalization of tryptophanins in oat (*Avena sativa* L.) seeds. Plant Sci. 172:579-587.

Morris, C. F., Giroux, M. J. 1998. Modification of cereal grain hardness via expression of puroindoline protein. US Patent 6,596,930.

Panozzo, J. F., Hannah, M. C., Obrien, L., Bekes, F. 1993. The relationship of free lipids and flour protein to breadmaking quality. J. Cereal Sci. 17:47-62.

Peremarti, A., Twyman, R. M., Gomez-Galera, S., Naqvi, S., Farre, G., Sabalza, M., Miralpeix, B., Dashevskaya, S., Yuan, D. W., Ramessar, K., Christou, P., Zhu, C. F., Bassie, L., Capell, T. 2010. Promoter diversity in multigene transformation. Plant Mol. Biol. 73:363-378.

Rooke, L., Byrne, D., Salgueiro, S. 2000. Marker gene expression driven by the maize ubiquitin promoter in transgenic wheat. Ann. Appl. Biol. 136:167-172.

Skylas, D. J., Cordwell, S. J., Craft, G., McInerney, B., Wu, M. J., Chin, J., Wrigley, C. W.,

Sharp, P. J. 2005. Proteome analysis of two soft-grained wheats of different processing quality: cultivar-specific proteins. Aust. J. Agric. Res. 56:145-155.

Stensballe, A., Hald, S., Bauw, G., Blennow, A., Welinder, K. G. 2008. The amyloplast proteome of potato tuber. FEBS J. 275:1723-1741.

Tanchak, M. A., Schernthaner, J. P., Giband, M., Altosaar, I. 1998. Tryptophanins: Isolation and molecular characterization of oat cDNA clones encoding proteins structurally related to puroindoline and wheat grain softness proteins. Plant Sci. 137:173-184.

Wang, L. L., Wang, E. T., Liu, J., Li, Y., Chen, W. X. 2006. Endophytic occupation of root nodules and roots of *Melilotus dentatus* by *Agrobacterium tumefaciens*. Microb. Ecol. 52:436-443.