Application of direct-sequencing peptide proteomics to the characterization of antagonistic (endogenous and exogenous) proteins in cereal grains

Adam Koziol

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

Department of Biochemistry, Microbiology and Immunology

Faculty of Medicine

University of Ottawa

In a Wheat Field, Evening Shadows
James Edward Hervey MacDonald
Canada, 1929

Abstract

The cereal seed plays a crucial role in society – both in the “food as medicine” paradigm, but also in food security. It is the starch and proteins present in the seed that lend it importance in these dissimilar anthropomorphic activities. This thesis investigation first characterized the post-translational processing of the potential diabetogen, wheat globulin-3. Globulin-3-like peptides were observed primarily in the embryo. These peptides varied significantly in their molecular masses and isoelectric points, as determined by two dimensional electrophoresis and immunoblotting. Five major polypeptide spots were sequenced by mass spectrometry, allowing for the development of a model of the post-translational events contributing to the globulin-3 processing profile. Three separate investigations of starch granules from different cereal species were performed. In the first series of experiments, pathogen-susceptible maize kernels were injected with either conidia of the fungal pathogen *Fusarium graminearum* or sterile water controls. Proteins in the desiccated fungal remnants on the surface of the kernels as well as in the endosperm and embryo tissues of the control and infected kernels were isolated and these proteomes were sequenced using tandem mass spectrometry. Approximately 250 maize proteins were identified. These proteins were classified into functional categories. There was an increased representation of defense proteins in the both the embryo and endosperm tissues of infected maize samples. The proteome of the fungal remnants was composed of 18 proteins. Several of these proteins were categorized as being involved in the metabolism of plant-sourced molecules, or in stress response. The second series of experiments detail the investigation of commercially prepared rice and maize starches using tandem mass spectrometry. The majority of identified proteins, in both rice and maize samples, were involved in either carbohydrate metabolism or storage. Markers for seed maturity and for starch mobilization were also documented. Finally, the third series of experiments investigated the non-host proteomes present in commercially-prepared starches. Non-host proteins from a variety of species, including *Homarus americanus* were found in the starch samples. This documentation of *H. americanus* proteins in these starch samples may have food safety implications with regards to shellfish allergies.

Acknowledgements

I would like to thank my family for always being there for me. I could not have done this without them.

I am grateful to my supervisor, Illimar Altosaar. He gave me this opportunity, and supported me every step of the way. A supervisor alone does not make a lab, and the companionship and camaraderie of the various Altosaar lab members over the years has been very much appreciated.

The assistance both in terms of materials and resources I received from the employees of Agriculture and Agri-Food Canada was invaluable.

My TAC members have changed as my doctoral training project has developed over time. I am thankful for the assistance and guidance from every TAC member, and other faculty members over the course of my studies.

Finally, I cherish the friendships, both new and existing, that have made this experience fantastic at times, and just bearable at others, but have always made my life better.

Table of Contents

List of Abbreviations

1D: one-dimensional
2D: two-dimensional
3D: three-dimensional
ABA: abscisic acid
ADP: adenine-diphosphate glucose
AGPase: ADP-glucose pyrophosphorylase
Beg: barley embryo globulin
BCA: bicinchoninic acid
BLAST: basic local alignment search tool
Bt1: brittle 1
CBB: Coomassie Brilliant Blue
CHAPS: 3-([3-cholamidopropyl]dimethylamino)-1-propanesulfonate
CID: collision-induced dissociation
DAP: days after pollination
DON: deoxynivalenol
DTT: dithiothreitol
GBSS: granule bound starch synthase
Glo-3: globulin-3
ESI: electrospray ionization
FDR: false discovery rate
G1P: glucose-1-phosphate
G6P: glucose-6-phosphate
HR: hypersensitive response
HRP: horseradish peroxidase
IEF: isoelectric focusing
IgE: immunoglobulin E
IgG: immunoglobulin G
IPG: immobilized pH gradient
KEGG: Kyoto Encyclopedia of Genes and Genomes
LC: liquid chromatography
LDS: linoleate diol synthase
LEA: late embryogenesis protein
Ma: million years ago
MS: mass spectrometry
MS/MS: tandem mass spectrometry
NCBI: National Center for Biotechnology Information
NCBInr: National Center for Biotechnology Information non-redundant
OYE: old yellow enzyme
OISB: Ottawa Institute of Systems Biology
PGM: phosphoglucomutase
PI: proteinase inhibitor
pI: isoelectric point
PPDK: orthophosphate dikinase
PR: pathogenesis-related

PSV: protein storage vacuole
Q: mass-resolving quadrupole
q: radio frequency quadrupole
Reg: rice embryo globulin
RIP: ribosome inactivating protein
ROS: reactive oxygen species
S: Svedberg unit
SBE: starch branching enzyme
SDBE: starch debranching enzyme
SDH: sorbitol dehydrogenase
SGAP: starch granule associated protein
SGS: starch granule surface
SOD: superoxide dismutase
SS: starch synthase
SuSy: sucrose synthase
T1D: type 1 diabetes
TBP: tributylphosphine
TCA: trichloroacetic acid
TMV: Tobacco mosaic virus
TOF: time of flight
UDP: uracil diphosphate
UDPG: uracil diphosphate glucose
UGPase: UTP-glucose-1-phosphate uridylyltransferase
USD: United States dollars
UTMB: University of Texas Medical Branch
WDEIA: wheat-dependent exercise-induced anaphylaxis

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Preamble – Research outline

Written in manuscript format, the experiments detailed in this thesis use mass spectrometry techniques (MS) to investigate two separate tissues within the cereal seed: the protein matrix and starch granules. To provide a biological overview for these experiments, the introduction is divided into four distinct sections. The first section provides a summary of seed evolution and physiology, the second section details the components of the protein matrix, with an emphasis on the wheat storage globulins. Starch granule development, morphology, and previous proteomic studies are presented in the third section, while the fourth section provides insight into the mass spectrometry techniques used in the experiments performed for this thesis.

1.2 The Seed

1.2.1 Evolution

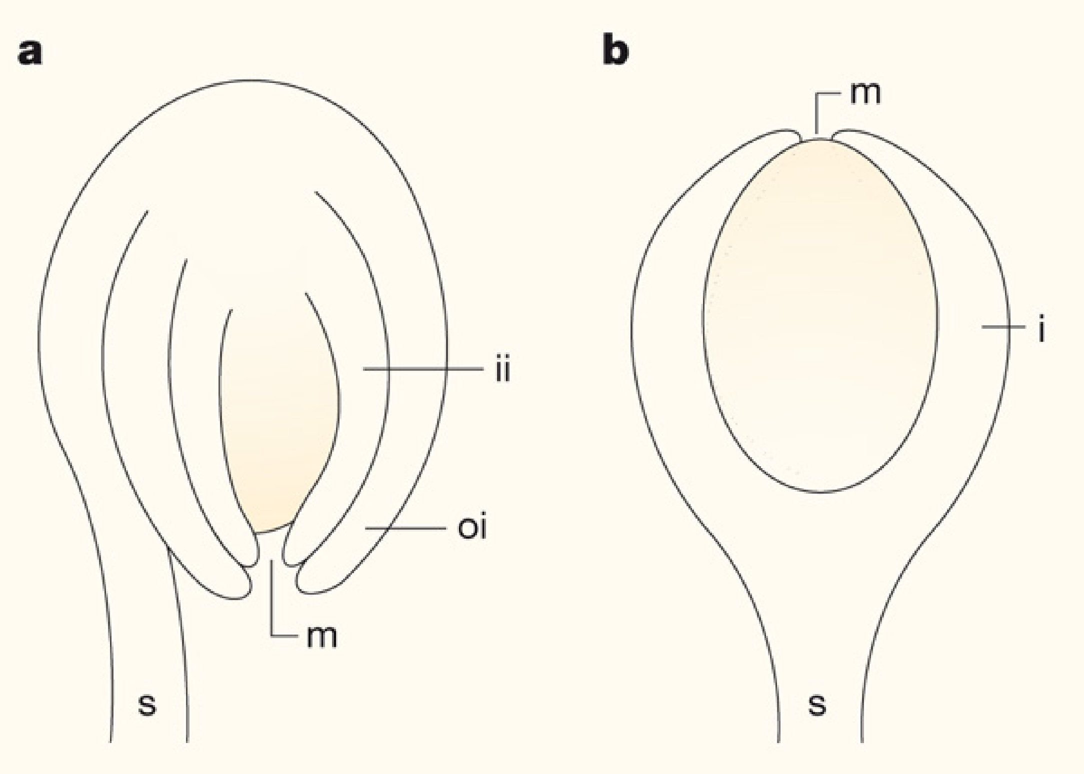
Elksinia polymorpha, the first known plant to bear seeds, evolved in the Late Devonian (Famennian), 365 million years ago (Ma) (Gerrienne et al., 2004). It has been hypothesized, however, that seeds could have evolved as early as between 385 and 365 Ma in vascular plants that belonged to the paraphyletic group termed Pteridospermae, or seed ferns. The seed ferns evolved from progymnosperms, which employed pteridophytic reproduction (spores), but also featured certain vegetative features common to seed ferns (Linkies et al., 2010). The transition from spores to seeds as a reproduction mechanism required the evolution of three individual traits: heterospory, integuments, and pollen-receiving structures (Taylor and Taylor, 2009). Heterospory is the specialization of spores into haploid male-like microspores and female-like megaspores, and integuments are protective layers around the nucellus (Esau, 1977).

1.2.2 Physiology of fruit and seed structures

The common features of all seeds are that, independent of plant lineage, seeds are composed of an embryo, a protective seed coat, and a nutrient source (Linkies et al., 2010). In the case of gymnosperms, seeds are not covered by an ovary, and are attached to the cone scales (< Greek *gymnospermos* “naked seed”). In contrast, angiosperm seeds are enclosed by the ovary (< Greek *angeion* “vessel” and *sperma* “seed”) (Figure 1.1).

The angiosperm cereal *Triticum aestivum* (wheat) is one of the most economically important crop plants, as approximately 60% of the proteins and calories consumed globally by humans are derived from wheat. As the United States Department of Agriculture estimates that more than 356 million metric tons of wheat are consumed annually, wheat is a good model seed system to study. The majority of calories, for both the embryo, and for human consumption, from harvested wheat are derived from the endosperm, which accounts for approximately 83% of total seed weight (Pomeranz et al., 1970). Endosperm tissue arises from double fertilization events that are unique to angiosperms. During the fertilization of the egg cell, a fertilization of a second cell of the megagametophyte by two different sperm cells occurs, leading to the formation of the triploid, nutritive, and supportive tissue that supplies energy to the embryo post-wintering (Floyd and Friedman, 2000). Through an initial cell enlargement and endopolyploidy phase, followed by a cellularization phase, the endosperm tissue is able to grow at very

Figure 1.1: Ovule morphology.
a, Angiosperm ovule. **b**, Gymnosperm ovule. Legend: i, integument (covering); ii, inner integument; m, micropyle (opening); oi, outer integument; s, stalk. Reproduced from Frohlich and Chase, 2007 with permission Macmillan Publishers Ltd: Nature.


a
ii
oi
m
s
b
m
i
s

rapid rates (Olsen, 2004). The endosperm cells are composed of 50-90% compact, high density starch granules, and seed storage proteins.

1.3 Storage proteins

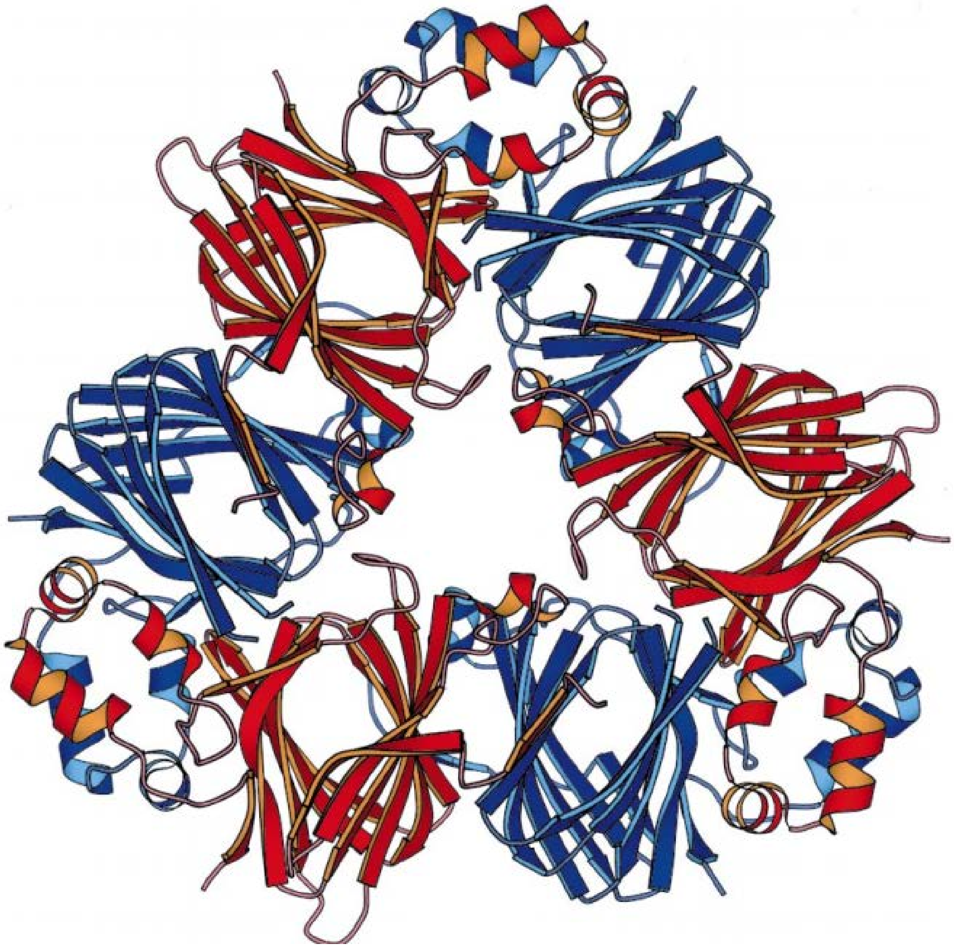
Seed storage proteins serve as sources of carbon, nitrogen, and sulphur source for germinating embryos (Shewry and Tatham, 1990). Additionally, the nutritional value of the seeds for human consumption is dependent on the amino acid balance of the storage proteins. For example, storage proteins high in the amino acids lysine, tryptophan and methionine, are considered to have a very good amino acid balance (Borlaug, 1983). Seed proteins are sequestered in an apoptotic, anhydrous protein matrix or in protein bodies called protein storage vesicles (Hara-Nishimura et al., 1998). Intended to be digested and subsequently taken up by the sprouting grain, they are now anthropomorphically classified based on their solubility in certain solvents. Albumins are soluble in water and dilute buffers, globulins are soluble in salt solutions, prolamins are soluble in 70% alcohol, and glutelins are soluble in dilute acids or bases (Osborne, 1908). Certain storage proteins are soluble in more than one solvent or have sequences and/or structures closely related to proteins of another solubility class. For example, the α-globulins of wheat, maize and rice are 18-25 kDa proteins that are soluble in saline solutions, but share significant sequence identity with a high molecular weight glutenin from wheat (Gu et al., 2004; Shorrosh et al., 1992; Woo et al., 2001). Additionally, rice glutelins are soluble in NaOH, but have related 3D structures to the 11S globulins of oat (Robert et al., 1985; Robert et al., 1985).

Depending on the plant species, the relative proportions of the seed storage proteins can be significantly different. In legumes such as chick pea, the majority of the storage proteins are globulins, with distributions of total protein as follows: albumin (8-12%), globulin (53-60%), prolamin (3-6%) and glutelin (19-24%) (Dhawan et al., 1991). This is very different than the distribution of storage proteins in wheat, where albumins and globulins represent 15% of total endosperm proteins, and prolamins (53%), and glutelins (32%) are the most predominant proteins (Shewry and Halford, 2002).

1.3.1 Globulins

The globulin family of seed storage proteins is of particular interest, as dietary exposure to certain members of this protein family have been implicated in the development of type 1 diabetes in certain mammals (MacFarlane et al., 2003). There are two categories of globulins based on their sedimentation coefficients: 7-9S and 11-13S (Danielsson, 1949). Orthologs to the 7S and 11S globulin protein families, vicilins and legumins, respectively, were first identified in the Leguminoseae (Derbyshire et al., 1976). Globulins are encoded by multigene families, and have characteristic β-barrel cupin domains (Dunwell, 1998) (Figure 1.2). The 7-9S globulins are arranged as homo- or hetero-trimers (Lawrence et al., 1994), while the 11-13S globulins are assembled into hexameric configurations (Adachi et al., 2001). Both the globulins and the legumins undergo significant post-translational modifications, including glycosylation and endoproteolytic cleavage prior to their storage in seed protein storage vesicles in the embryo (Hara-Nishimura et al., 1998).

Figure 1.2: Three-dimensional structure of jack bean 7S globulin trimers.
The N-terminal and the C-terminal domains, coloured red and blue, respectively, are arranged about the threefold axis. From Ko et al., 2000, reproduced with permission of the International Union of Crystallography.



In cereal crops, there have been several studies characterizing the 7S globulins in barley and maize, (Heck et al., 1993; Kriz, 1989), and more recently in *Brachypodium distachyon* (Larre et al., 2010). However, these proteins are still poorly characterized in wheat and oats. The lack of information on oat globulins is unusual, as oats are the first foodstuff to be allowed a medical health claim by the U.S. Food and Drug Administration stating that consumption of soluble fiber from whole oats reduces the risk of heart disease. The expression of 7S globulins is primarily in the embryo and aleurone layer (Sun et al., 1996), with little to no evidence of globulin expression in the endosperm (Loit et al., 2009). Expression of globulins is transcriptionally regulated by the phytohormone abscisic acid (ABA) (Rivin and Grudt, 1991). As globulins are expressed primarily in the embryo and contain increased levels of nitrogen-rich amino acids arginine, lysine, aspartic acid and glutamic acid (Zhu et al., 2006), it is hypothesized that the globulins have evolved to be used as an easily exploitable source of amino groups and energy for germinating embryos (Shewry and Tatham, 1990).

There are two members of the 7S globulin protein family present in multiple cereal species: globulin 1 (Glbl) and globulin 2 (Glb2) in maize, barley embryo globulins (Begl and Beg2) in barley and rice embryo globulins (Regl and Reg2) in rice (Sun et al., 1996; Yupsanis et al., 1990). Previous studies on Glb1, Beg1 and Reg1 have shown that these proteins are all encoded by single copy genes (Belanger and Kriz, 1989; Heck et al., 1993; Sun et al., 1996). However, there is evidence to suggest that *Beg1* may belong to a small multigene family (Zhang et al., 2004). The expression patterns of the members of the protein family can be significantly different, as the rice embryo globulins as well as

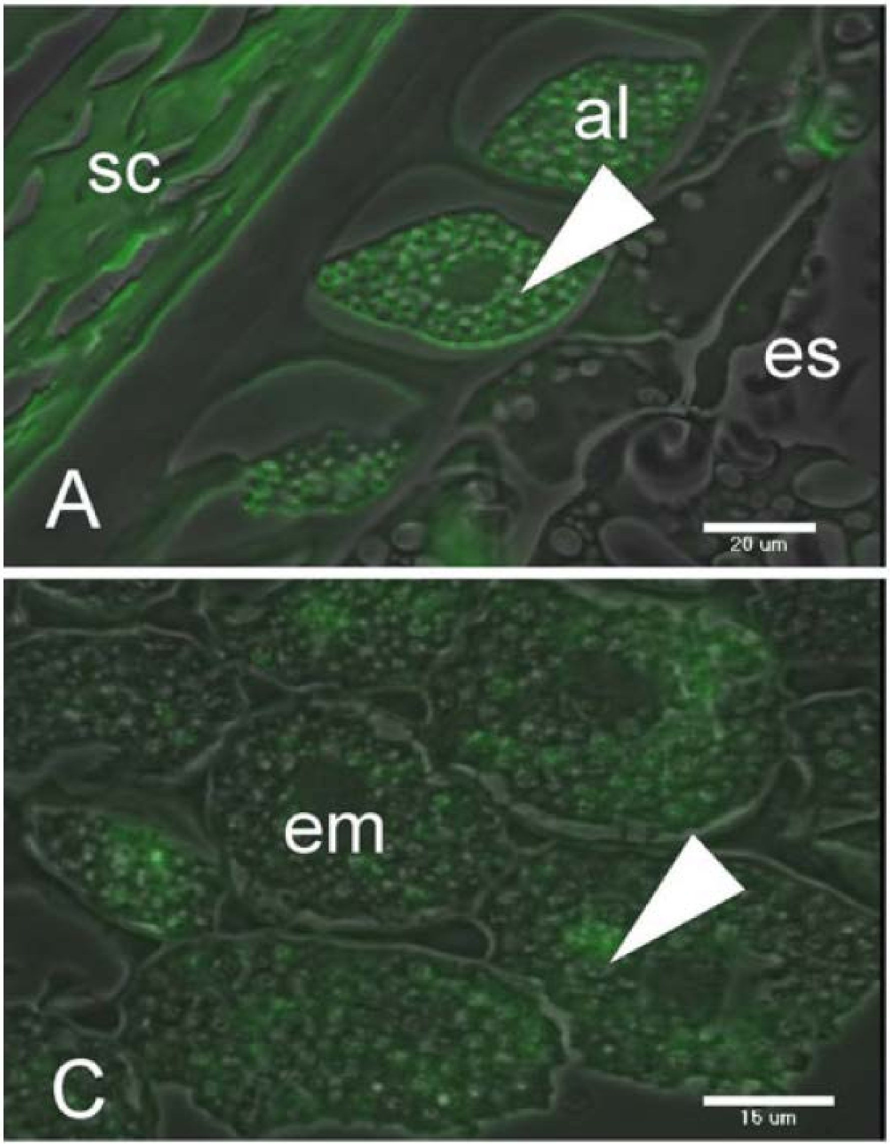
Begl transcripts are highly expressed during early germination, whereas *Beg2* transcript levels decrease during the germination process (Zhang et al., 2004).

In wheat, there are three members of the globulin protein family: globulin-1, -2 and -3, as well as triticin (Gomez et al., 1988; Singh et al., 1991). Globulin-3A is a potential food allergen (Larre et al., 2011), and has been associated with an increased risk for the development of type 1 diabetes (MacFarlane et al., 2003) and celiac disease (Taplin et al., 2011) in susceptible individuals. The globulin-3 gene family in wheat cultivar Glenlea contains three individual globulin-3 genes: *Glo-3A, Glo-3B*, and *Glo-3C* (Loit et al., 2009). These genes share a moderate to high degree of nucleotide sequence identity (73-93%) and are expressed exclusively in the embryo and aleurone layer (Figure 1.3). Previous investigations into the post-translational endoproteolytic cleavage of the globulin-3 family have shown that these proteins can be processed at internal cleavage sites to yield polypeptides with a range of molecular weights and pI values (Dupont et al., 2011; Singh et al., 2001). To gain a more precise understanding of the processing events that orchestrate our dietary proteomes, a targeted study of these processed proteins is required to understand their role in the development of human diseases.

1.3.2 Globulins and human health

Wheat seed storage proteins, including globulin-3, and the γ-gliadins have been associated with the development of wheat allergies, wheat sensitivity and celiac disease (Tatham and Shewry, 2008). Certain occupations have higher incidences of wheat allergies; baker's asthma and rhinitis can affect as many as 30% of professional bakers

Figure 1.3: Immunolocalization of globulin-3 in developing wheat embryos. Panel A: wheat aleurone (al) cells were positively stained while endosperm (es) cells were unstained when sections were stained with Glo-3A antibodies. Panel (C): wheat embryonic (em) tissue stained with Glo-3A antibodies. Green staining in seed coat (sc) shows unspecific staining. Green (examples indicated with arrowheads) indicates positive staining. Reproduced with permission from Loit et al., 2009.


sc
al
es
A
20 um
em
C
15 um

(Bittner et al., 2008). Though these allergies were recognized by the Romans, it was not until the 1970s that wheat albumins were found to be reactive with immunoglobulin E (IgE) antibodies from individuals suffering from baker's asthma (Baldo and Wrigley, 1978). The allergies in these individuals are responses to the inhalation of wheat flour and dust particles.

Type 1 diabetes (T1D) is an autoimmune disease characterized by the destruction of the insulin producing β-cells in the pancreatic islets of Langerhans and the loss of glucose homeostasis (Jahromi and Eisenbarth, 2007). It has been hypothesized that a dysfunctional gut barrier may be involved in disease onset in certain individuals (Maurano et al., 2005). While the incidence of T1D in developed countries has risen by 3% annually since 1960 (Onkamo et al., 1999), there is still no definitive cause of this disease. Identified risk genes impart genetic susceptibility to T1D, but there is under 50% concordance in monozygotic twins (Todd, 2010). Therefore, the development of T1D likely relies on environmental triggers, such as diet. Alternatively, the hygiene hypothesis states that the reduced exposure to infectious agents during childhood is responsible for the increased prevalence of T1D (Yazdanbakhsh et al., 2002). In current societies, with heightened standards for sanitation, healthcare and diet, the natural development of the immune system can be repressed by the lack of exposure to T1D protective factors, such as symbiotic prokaryotes and certain parasites (Todd, 1991).

Celiac disease, also known as gluten enteropathy, or sprue, is a hereditary immune system disorder (Fasano, 2001). The intestinal epithelium and the tight intercellular junctions in healthy individuals are intact and do not permit dietary antigens to pass the gut barrier (Heyman et al., 2012). However, the tight junction system becomes damaged in subjects with celiac disease (Shan et al., 2002). In these individuals, the immune systems can be stimulated by dietary antigens, such as gliadins that pass through the tight junctions (Tjon et al., 2010). An increased expression of the tight junction modulating protein zonulin may have a role in the degradation of the tight junctions (Lammers et al., 2008). Celiac disease is one of the most common genetic diseases, with worldwide prevalence of approximately 0.3-1% (Catassi et al., 1994).

1.3.3 Globulin summary

As the mechanisms of the development of diseases such as type 1 diabetes and celiac disease are still poorly understood, the implication of dietary antagonists is an important avenue to explore. Further research is necessary to discover how globulin-3 is involved with disease development, and whether the post-translation processing of globulin-3 is a factor in the development of these diseases.

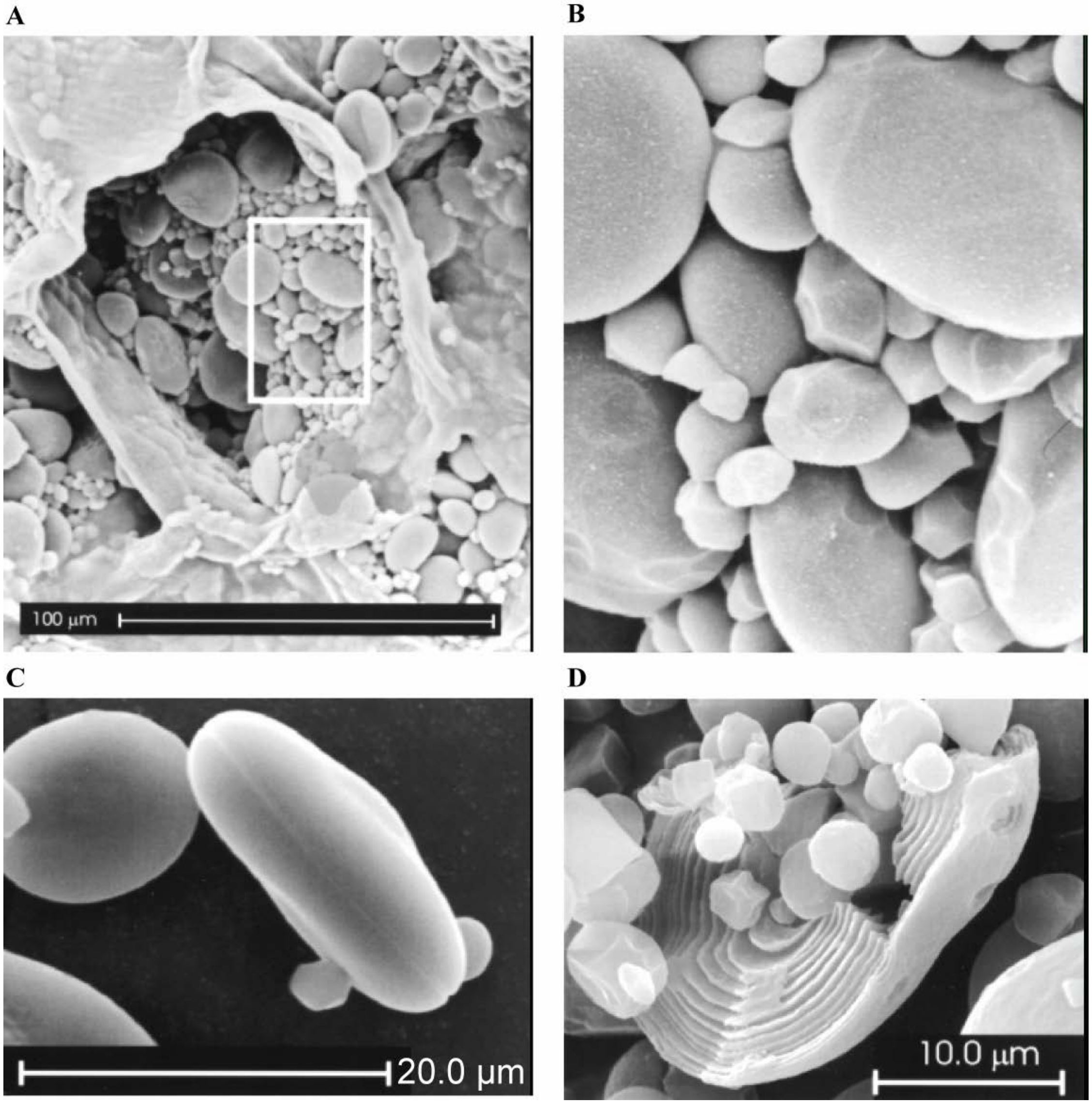
1.4 Starch granules

Photosynthesis fixes carbon dioxide using the reductant generated through the light reactions of photosynthesis. The reductant can be subsequently oxidized to support the energy needs of plant growth and development. Energy generation and consumption needs vary depending on light levels and the developmental stage of plants. Therefore the capacity to store energy as inert starch granules in the form of complex polysaccharides composed of monomeric glucose molecules linked by glycosidic bonds is a necessity.

Starch typically accumulates in photosynthetically active chloroplasts or in amyloplasts in non-photosynthetic tissues, which specialize in starch storage (Tetlow, 2011).

Starch formation in wheat amyloplasts is associated with stromules (stroma-filled tubules) during the development of the endosperm in cereals (Natesan et al., 2005). It was determined that A-type granules (>15 μM) are initiated in undifferentiated plastids in the coenocytic endosperm (Bechtel and Wilson, 2003). Following cellularization of the endosperm, B-type granules (5-15 μM) form in stromules protruding from plastids containing A-type granules in the subaleurone layer. C-type granules (< 5 μM) are initiated in branched stromules that form cytosolic networks (Bechtel and Wilson, 2003) (Figure 1.4). Similar development of starch granules and stromules was observed in cells of the central endosperm (Bechtel and Wilson, 2003). Starch granules are composed of 98-99% amylose and amylopectin molecules (Buléon et al., 1998).These polysaccharide chains are composed of glucose monomers linked by α-1,4 or α-1,6 glycosidic bonds (Ball and Morell, 2003). Amylose contains 99% α-1,4 bonds and 1% α-1,6 linkages and therefore is essentially a long, linear molecule compared to amylopectin, which contains 95% α-1,4 linkages and 5% α-1,6 linkages (Tester et al., 2004). The crystalline structure of starch granules is attributed to the dense packing of the double helices of amylopectin (Perez and Bertoft, 2010) (Figure 1.5). The ratio of amylose to amylopectin can range significantly between starch granules: from under 1% amylose in waxy starches to greater than 70% amylose in high amylose starches (Buléon et al., 1998). The composition of most starch granules ranges from 20-35% amylose (Tester et al., 2004). The size, shape, and composition of starch granules can vary between species. Wheat starch granules are composed of 36% amylopectin (Morrison et al., 1994), and have a bimodal size

Figure 1.4: Scanning electron micrographs of wheat starch granules.
Panels A and B are mature wheat endosperm tissue (Large A-type granules (>15 µm), medium B-type granules (5–15 µm) and small C-type granules (<5 µm) are present in panels A and B. B is a magnified image of the boxed portion in A. C shows starch granules isolated from mature endosperm. D shows a fractured starch granule, revealing the layers of starch. Reproduced from Rahman et al., 2000 with permission from Elsevier.


A
B
C
D
100 µm
20.0 µm
10.0 µm

distribution: type A granules, range from 15 to 35 µm in diameter with a lenticular shape and the smaller type B granules range from 2 to 10 µm with a spherical beach ball shape (Tester et al., 2004). In contrast, rice starch granules have an unimodal size distribution, are approximately 3 to 8 µm in size (Cledat et al., 2004) and are composed of 47-51% amylopectin (Qi et al., 2003). During development, polyhedral compound granules are generated through the interactions of the growing granules (Jane et al., 1994). The shape, size and composition of starch granules all have impacts on the bioutilization, the end use properties as well as on the nutritional benefits of starch for both embryo and predators. The nature of the starch granule surface, and particularly the presence of surface lipids and proteins, has many significant effects on the properties of the starch.

1.4.1 Starch granule surface

The starch granule surface is characterized by the presence of certain features, including pores, protrusions, and channels (Fannon et al., 2003; Fannon et al., 2004) (Figure 1.6). The occurrence, size and number of these topological features is dependent on the species, tissue, and starch granule type (Huber and BeMiller, 2000). In addition to these topological features, proteins have been observed on the starch granule surface using atomic force microscopy (Baker et al., 2001) and other techniques, including mass spectrometry (Wall et al., 2010) and immunofluorescence (Lauriere et al., 1986). These starch granule associated proteins (SGAPs) include a number of proteins of varying molecular weight, including several isoforms of the starch granule bound starch synthases (GBSS), starch synthases (SS) and starch branching enzymes (SBE)

Figure 1.5: Organization of amylopectin and starch granules.
(A) Schematic view of a starch granule with its succession of amorphous and crystalline growth rings. (B) Relation between a section of a crystalline growth ring of the granule and the molecular organization of amylopectin. (C) The succession of ten alternating lamellae in context to the primary structure of a portion of an amylopectin molecule. Each line represents an α-1,4 linked glucan chain. The chains are hooked together by α-1,6 branches. The dotted line delimits the sections appearing in the crystalline and amorphous lamellae. (D) Relation of a part of the primary structure depicted in (C) to the secondary structure of a single cluster displaying the double helical structures. Reproduced from Perez et al., 2010 with permission from John Wiley and Sons.

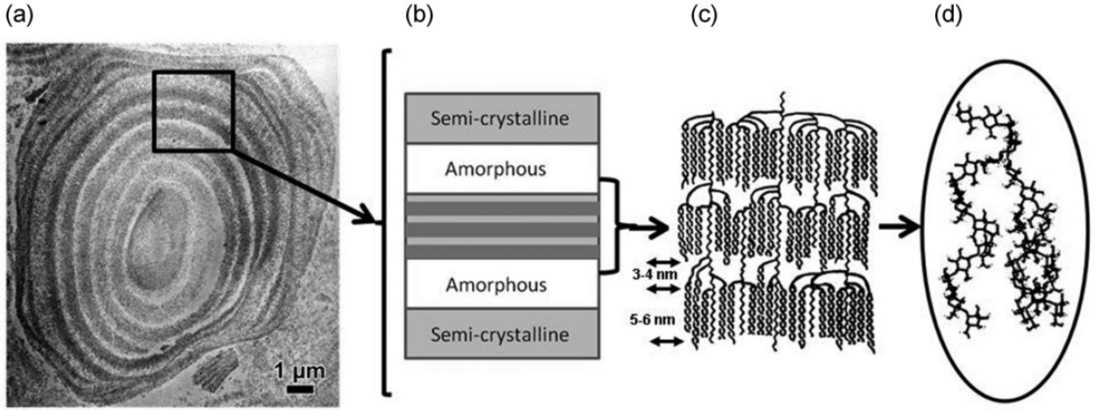
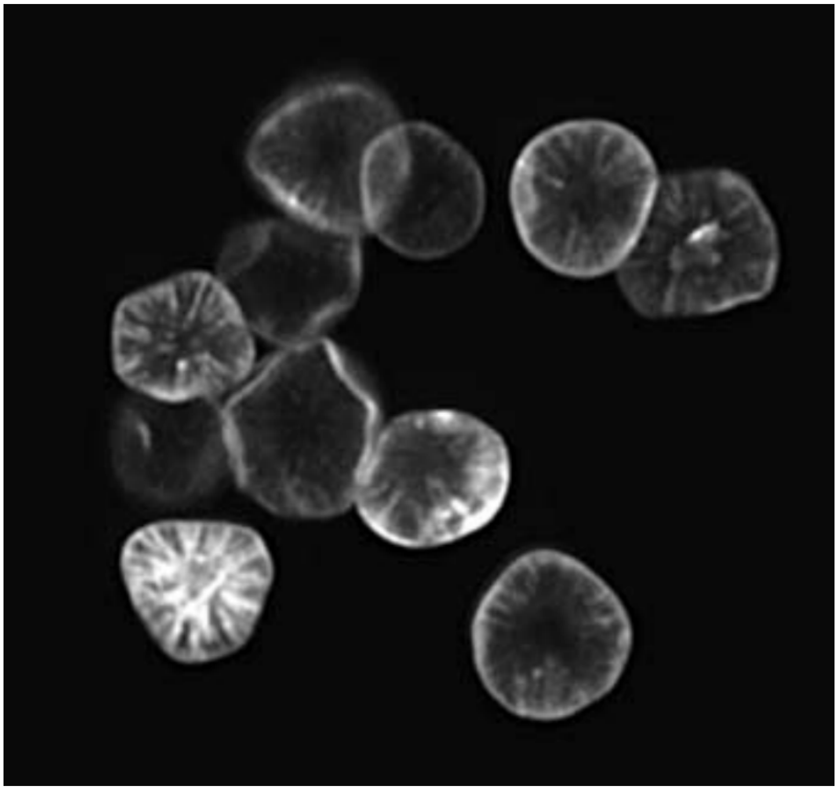

(a)
1 µm
(b)
Semi-crystalline
Amorphous
Amorphous
Semi-crystalline
(c)
3-4 nm
5-6 nm
(d)

Figure 1.6: Channels in starch granules may contain proteins. Compositional backscattered electron image of sorghum granules treated with a solution of merbromin in methanol to highlight granule topography. Channels and cavities are visible only when they are present within the optical plane. Reproduced from Fannon et al., 2004 with kind permission from Springer Science and Business Media.



(Rahman et al., 2007) as well as several other proteins, many of which are still poorly characterized. For instance, an abundant and poorly characterized unnamed wheat SGAP with a molecular weight of approximately 30 kDa is present both on the surface and within the granule (Baldwin, 2001). The function of the protein is unknown, but it has been hypothesized that it may be metabolized by the germinating embryo. The well documented SGAP puroindoline is the major determinant of wheat grain softness, an important property in determining the end-use quality of wheat (Morris, 2002). Puroindolines have also been shown to have antimicrobial properties and have likely evolved as a defence mechanism (Capparelli et al., 2005; Krishnamurthy et al., 2001).

1.4.2 Starch granule summary

Starch granules can therefore be regarded as unique intracellular compartments of endosperm cells that consist of the metabolic and regulatory proteins for starch synthesis and degradation. The proteins localized to this compartment either contain a specific binding domain for starch or directly interact with a protein or group of proteins bound to granules. Since these proteins and the starch granules store large amounts of energy in their chemical bonds, there are important safeguards, such as the innate immune system, in place to protect them from predation.

1.5 Plant/pathogen interactions

To protect their energy reserves (both starch granules and storage proteins), and, similarly, the embryo from predators and pathogens, plants have evolved layers of resistance mechanisms. These mechanisms include physical barriers, such as the cuticle, and chemical barriers, including antimicrobial molecules, such as phytoanticipins (Osbourn, 1996). In addition to these “first-line” defenses, plants have inducible resistance mechanisms that are activated upon the sensing of pathogen infection (Durrant and Dong, 2004). These inducible mechanisms include the initiation of the hypersensitive response (HR), the cross-linking of cell walls, and the accumulation of reactive oxygen species (ROS), secondary metabolites, and pathogenesis-related (PR) proteins (Sels et al., 2008). The PR proteins are proteins that are either undetectable or present at low concentrations in healthy tissues, but are induced following pathogen infection (van Loon et al., 2006).

The PR proteins were originally described in *Nicotiana tabacum* (tobacco) as proteins upregulated in the hypersensitive response to Tobacco mosaic virus (TMV) (van Loon and van Strien, 1999). There are now 17 recognized families of PR proteins (Table 1.1). The PR families are numbered based on the order of their discovery (van Loon et al., 2006). While not all PR families have been demonstrated to have direct roles in defense, many of the families have well documented antimicrobial activities. PR families such as PR-3, -4, -8, and -11 (chitinases), PR-5 (thaumatin), PR-12 (defensin), PR-13 (thionin), and PR-14 (lipid-transfer protein) catalyze the degradation of specific microbial macromolecules. Chitinases catalyze the cleavage of the β-1,4-glycosidic bonds of the fungal cell wall structural molecules chitin and chitosan (Jitonnom et al., 2011). Collectively, PR-5, -12, -13, -14 catalyze the degradation of microbial cell walls, though these families have different target molecules (Dubreil et al., 1998; Lay and Anderson, 2005).

Table 1.1: Families of PR proteins.

Family	Properties
PR-1	Unknown
PR-2	β-1,3-glucanase
PR-3	Chitinase type I, II, IV, V, VI, VII
PR-4	Chitinase type I, II
PR-5	Thaumatin-like
PR-6	Proteinase-inhibitor
PR-7	Endoproteinase
PR-8	Chitinase type III
PR-9	Peroxidase
PR-10	Ribonuclease-like
PR-11	Chitinase, type 1
PR-12	Defensin
PR-13	Thionin
PR-14	Lipid-transfer protein
PR-15	Oxalate oxidase
PR-16	Oxalate-oxidase-like
PR-17	Unknown

Adapted from van Loon et al., 2006.

The functions of the PR proteins are not limited solely to defense. The PR-2 proteins have antimicrobial function through their ability to catalyze the cleavage of 1,3-β-D-glucosidic linkages of the glucans that are prevalent in fungal cell walls (Leubner-Metzger and Meins, 1999). However, a diverse assortment of roles for this family in uninfected plants have been established, including, but not limited to embryogenesis (Helleboid et al., 1998), mobilization of energy reserves (Fincher and Stone, 1993), fruit ripening (Hinton and Pressey, 1980), and cold tolerance (Hincha et al., 1997). Other PR families also have diverse roles unrelated to plant defense, therefore careful study is required to document all the functions of the PR families.

Successful fungal pathogen infection is characterized by the secretion of fungal enzymes that catalyze the degradation of the plant's physical barriers, such as the cuticle and cell walls (Phalip et al., 2005). Enzymes secreted by *Fusarium graminearum*, the fungal pathogen responsible for causing Gibberella ear rot in maize, include lipases and cutinases that degrade the cuticle (Feng et al., 2005; Jenczmionka and Schafer, 2005; Kang and Buchenauer, 2000) and pectinases that modify plant cell walls (Kikot et al., 2009), allowing easier access of other fungal enzymes, such as cellulases and xylanases, to plant cell components (Reignault et al., 2008). These fungal cell wall-degrading enzymes facilitate the colonization of wheat spikes prior to fungal grain infection (Wanjiru et al., 2002).

Following the penetration of a cereal plant seed's physical barriers, the pathogen will target the energy-rich endosperm tissue (Kikot et al., 2009). Fungal enzymes such as proteinases and amylases catalyze the degradation of the seed's storage proteins and starch granules, respectively (Schwarz et al., 2001). For example, following infection of endosperm tissue with *F. graminearum*, starch granules were digested to a great degree and the intracellular protein matrix was fully degraded (Jackowiak et al., 2005) (Figure 1.7).

Summary of plant/pathogen interactions

Plants and pathogens are continuously evolving new mechanisms for the defense or infection of plant tissues, respectively. This ongoing evolutionary arms race between pathogen and plant can have very large economic impacts. Over $1 billion USD of crops are lost to *F. graminearum* infection every year (Goswami and Kistler, 2004), indicating the study of these plant/pathogen interactions using state of the art tools, such as mass spectrometry, can have both economic impacts and benefits for human health by increasing food security.

1.6 Mass spectrometry

The proteomes of wheat endosperm (Vensel et al., 2005), wheat and potato amyloplasts (Andon et al., 2002; Stensballe et al., 2008) and maize starch-granules (Grimaud et al., 2008) have been catalogued using high-throughput mass spectrometry techniques. The proteins identified in these studies were classified based upon their biochemical roles in the cell.

In wheat amyloplasts, many of the identified proteins have roles in metabolism, as well as protein destination and storage (Balmer et al., 2006) (Figure 1.8). It is interesting to note

Figure 1.7: Damaged starch granules of spring wheat infected with *Fusarium*. Reproduced from Jackowiak et al., 2005 with permission Elsevier.

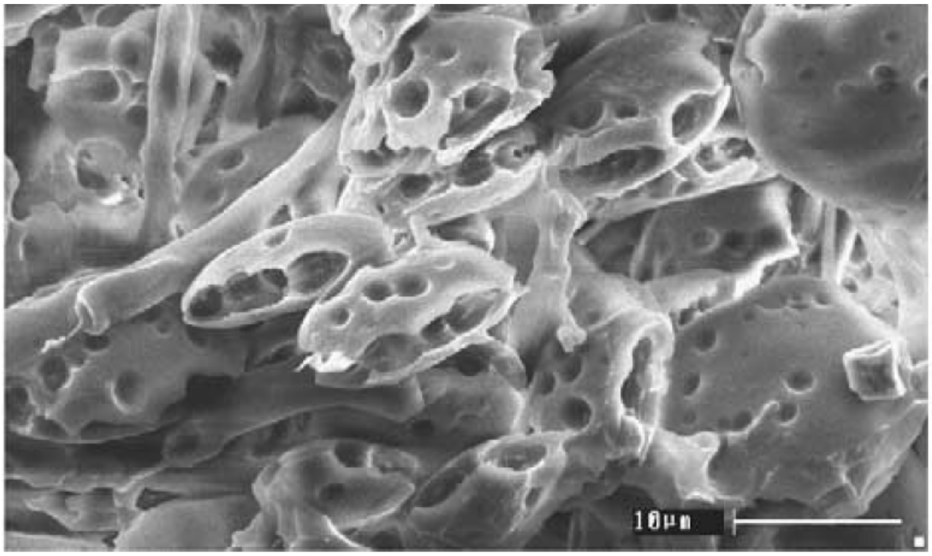
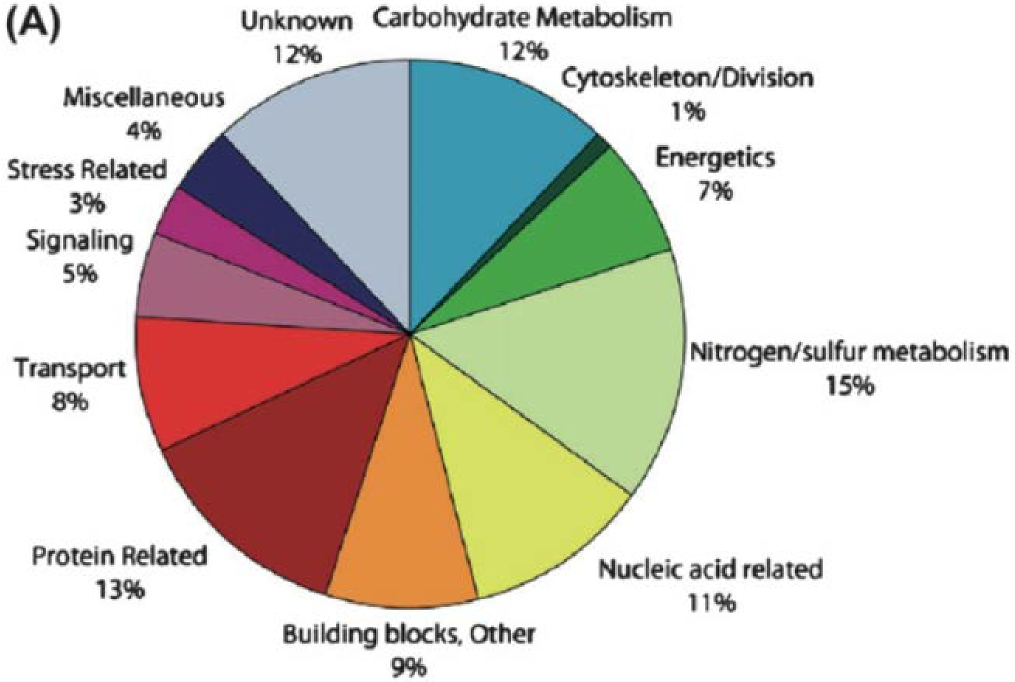

10μm

Figure 1.8: Functional distribution of proteins identified in amyloplasts from wheat endosperm. Reproduced from Balmer et al., 2006 with permission from Oxford University Press.

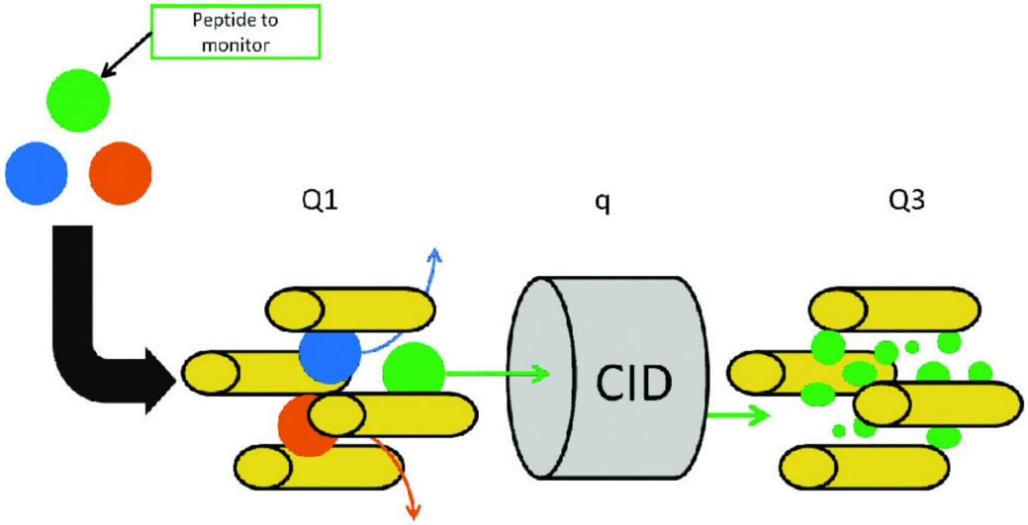

(A)
Unknown 12%
Carbohydrate Metabolism 12%
Cytoskeleton/Division 1%
Energetics 7%
Nitrogen/sulfur metabolism 15%
Nucleic acid related 11%
Building blocks, Other 9%
Protein Related 13%
Transport 8%
Signaling 5%
Stress Related 3%
Miscellaneous 4%

that over one third of the proteins identified in the study were unknown or hypothetical proteins. This underscores the necessity of gene discovery and functional analysis studies for large-scale proteomic analysis. The starch granule associated proteome of water washed soft (pathogen resistant) and hard wheat (pathogen susceptible) cultivars have previously been investigated (Wall et al., 2010). In addition to the many unknown and hypothetical proteins, there were multiple proteins with antimicrobial activities identified. As these samples were collected from field grown plants that were deemed fit for harvest, a natural progression of this work is to study grains intentionally infected with pathogens and investigating the subsequent plant/pathogen interactions at the proteomic level. Furthermore, the proteomes of commercially-available starch has been studied to test the suitability of the starch for patients with celiac disease (Kasarda et al., 2008). The authors discovered many proteins that had not previously been believed to be associated with the starch granule surface, showing that careful study is required when documenting proteins in foods destined for human consumption.

The term mass spectrometry refers to a broad range of analytical techniques used for the analysis of molecules based on their mass to charge ratio. As mass spectrometry covers such a broad range of techniques, only the techniques used in the experiments in this thesis will be covered. These techniques include the electrospray ionization (ESI) delivery system, the time of flight (TOF) mass analyzer and mass selection and detection components, and the QStar QqTOF triple quadrupole tandem mass spectrometer. For this latter instrument, Q refers to a mass-resolving quadrupole, and lower-case q refers to a radio frequency-only quadrupole or hexapole collision cell (Chernushevich et al., 2001). This instrument is a tandem mass spectrometer, as there are two mass selection quadrupoles (Q1 and Q3) separated by ion fragmentation in a collision chamber (q2). When performing peptide sequencing experiments, mixtures of peptides are first separated by liquid chromatography and ionized using electrospray ionization. This technique involves three separate stages: droplet formation, droplet shrinkage and gaseous ion formation (Smith et al., 1990). To promote droplet shrinkage, peptides are suspended in volatile solvents such as acetonitrile. During droplet formation, liquid is extruded from a capillary tip in a "Taylor cone" shape due to electrostatic forces and is subsequently aerosolized (Figure 1.9). The droplet size of this aerosolized liquid is controlled by the inclusion of certain compounds, such as acetic or formic acid, that increase the conductivity of the peptide mixture as well as the applied potential, the flow rate of the solvent, and the diameter of the capillary. As the droplets shrink, gaseous ions pass into the mass spectrometer (Cech and Enke, 2001).

A characteristic of ESI is that the gaseous ions can each have multiple charged residues. Multiple charges reduce the mass to charge ratio of the protein, allowing for the analysis of high molecular weight proteins that are normally outside the usable mass range of the mass spectrometer (Dongre et al., 1996). For mass analysis, the QStar hybrid QqTOF MS/MS consists of three quadrupoles linked to a TOF mass analyzer (Figure 1.9). The first quadrupole (Q1) is responsible for initial ion selection, allowing for the selection of a specific mass range. The second quadrupole (q2) is a non-mass filtering radio frequency quadrupole that serves as a collision cell where ions selected in Q1 can be broken down into smaller ions through collision-induced dissociation (CID) with an inert gas such as argon, nitrogen or helium. The third quadrupole (Q3) allows for further ion selection before the selected ions are delivered to the TOF mass analyzer (Aebersold and Mann,

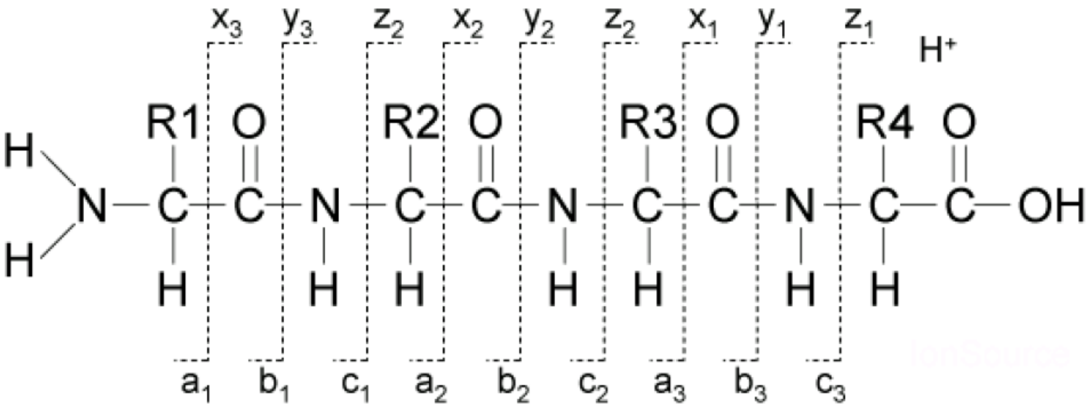
Figure 1.9: Schematic of a triple quadrupole mass spectrometer.
Legend: Q1, quadrupole 1, q, quadrupole 2, Q3 quadrupole 3, CID, collision-induced dissociation. From broadinstitute.org/files/shared/proteomics/tq.jpg. Last viewed September 29, 2012.


Peptide to monitor
Q1
q
CID
Q3

2003). As ions enter the TOF chamber, the ions are accelerated by an electric field of known strength. Due to this acceleration, all ions with identical charge will have the same kinetic energy. As lighter particles are able to accelerate to higher speeds than heavy particles, an ion's velocity is dependent on the mass-to-charge ratio. As the time an ion takes to reach an ion detector at a known distance can be measured, this time and the experimental parameters are used to calculate the mass-to-charge ratio of the ion (Takats et al., 2004).

When analyzing peptide samples using a tandem mass spectrometer, the fragmentation of positively charged ions occurs in a predictable pattern due to the tendency for the amide bond of peptides to be cleaved during ion fragmentation (Kelleher et al., 1999). The fragment containing the N-terminus is referred to as a b-ion, while the fragment containing the C-terminus is referred to as a y-ion (Roepstorff and Fohlman, 1984) (Figure 1.10). The differences in mass between the fragmented ions compared to the mass of the ion prior to fragmentation allows for calculation of the amino acid sequence of the peptide. Based on this deduced amino acid sequence, it is possible to identify the protein from which this ion was generated. The MS/MS peptide spectra generated from liquid chromatography (LC) MS/MS analyses are analyzed using search engines such as Mascot (Perkins et al., 1999). The search algorithms used by Mascot compare experimentally obtained fragmentation spectra to theoretical fragmentation spectra calculated from a database of protein sequences, such as the National Center for Biotechnology Information (NCBI) non-redundant protein database. The scoring system used by Mascot is probability-based. The total score is the probability that the observed match is a random event, with the scores reported as $-10*LOG_{10}(P)$, where P is the absolute probability. To

Figure 1.10: Types of peptide fragment ions observed in an MS/MS spectrum. Adapted from Roepstroff and Fohlmann, 1984.


x_3
y_3
z_2
x_2
y_2
z_2
x_1
y_1
z_1
H^+
R1
R2
R3
R4
O
O
O
O
H
H
N
C
C
N
C
C
N
C
C
N
C
C
OH
H
H
H
H
H
H
H
a_1
b_1
c_1
a_2
b_2
c_2
a_3
b_3
c_3

determine whether the calculated protein score is significant, Mascot computes an expectation value for every hit. The expectation value is defined as the number of matches with equal or better scores that are expected to occur by chance alone. This threshold is commonly set at $p<0.05$, therefore any expectation values below 0.05 are considered significant. Another technique to ensure the significance of the Mascot matches are decoy database searches. These searches use the same protein databases as the standard query, but the sequences have been randomized or reversed. This allows for the calculation of the false discovery rates (FDR) in an experiment (Elias et al., 2005). The FDR is calculated as follows: $FDR = FP / (FP + TP)$, where FP is the number of matches in the decoy database and TP is the number of matches in the target database.

1.6 Research Hypotheses and Objectives

Rationale 1

Previously, globulin-3A in wheat cultivar AC Barrie was shown to have implications in the development of type 1 diabetes in genetically susceptible individuals. As AC Barrie and Glenlea are closely related hexaploid cultivars, it is feasible that the globulin-3 family in AC Barrie is similar to the previously documented globulin-3 proteins in Glenlea. Additionally, the post-translational endoproteolytic processing of globulins have been shown to be similar to the post-translational processing observed in barley.

Hypothesis 1

Globulin-3 in wheat undergoes post-translational processing, including glycosylation and endoproteolytic processing, similar to 7S proteins in other species such as soy and barley.

Objectives

1. Determine the expression pattern of globulin-3 in wheat embryo and endosperm tissues.
2. Sequence the proteolytically-cleaved proteins separated on two-dimensional SDS-PAGE gels.
3. Create a model for proteolytic cleavage of the globulin-3 preproprotein.

Rationale 2

Use of a pathogen-susceptible maize genotype and the mechanical kernel injection technique both promote a fuller expression of the fungal pathogen proteome. Under these optimal pathogen growth conditions, the fungus will take up nutrients from the plant and will express enzymes to metabolize the nutrients. Additionally, the inducible plant defence proteins as well as fungal proteins will be present and interacting within both the endosperm and embryo of the host tissues.

Hypothesis 2

Proteins sequenced from the desiccated fungal remnants on the surface of maize kernels can provide insight into the biochemical processes occurring in the fungus prior to post-infection desiccation. Host defense-related proteins will be present in both the endosperm and embryo tissues, allowing for a better understanding of plant/pathogen interactions.

Objectives

1. Isolate and purify proteins from desiccated fungal remnants on the surface of infected maize kernels as well as proteins from the endosperm and embryo tissues of infected and control kernels.
2. Sequence the proteomes of these tissues.
3. Classify proteins based on their biochemical roles.

Rationale 3

The certificates of analysis for commercially available starch samples give basic measurements of the non-starch components of the samples, such as pH and percentages of moisture, protein, and lipid, but not the full proteome of the starch samples. There is 1% protein in these starches that is undocumented, thus requiring deeper inspection using mass spectrometry techniques.

Hypothesis 3

The protein component of commercially purified starch samples could affect the outcome of other experiments in which the starches will serve as defined baseline substrates. The investigation of these trace proteomes will allow for the determination of protein-sourced contributions to the macroscopic characteristics of these starches.

Objectives

1. Purify the starch granule-associated proteins in commercially-purified starch.
2. Sequence these proteins using tandem mass spectrometry.
3. Classify the proteins based on biochemical characteristics.

Rationale 4

The water used to process the starch during commercial purification can contaminate starch samples with non-host organisms, and can affect the end quality of the starch. For example, proteins from microbial pathogens have been documented on the surface of starch granules in lab-scale purified wheat starch.

Hypothesis 4

The source of any non-host proteins present in the commercial starch preparations is the water used during the extraction of the starch granules.

Objectives

1. Isolate and sequence the proteins present in commercially prepared rice and maize starches.
2. Document any non-host protein present in the starch samples.

Chapter 2

Seed storage proteins of the globulin family are cleaved post-translationally in wheat embryos

Koziol AG, Loit E, M^{c}Nulty M, MacFarlane AJ, Scott FW, Altosaar, I

Published in: *BMC Research Notes* 2012, **5**:385 (28 July 2012)

Contribution of Authors:

AGK performed bioinformatic analyses, analyzed data, assembled all figures and tables, drafted and edited the manuscript. EL helped draft the first manuscript, aided in editing, and helped with experiments. MM performed the electrophoresis experiments, contributed to the initial writing and subsequent editing. AJM, FWS provided the globulin-3 antibodies and edited the manuscript. FWS participated in the study design. IA designed the study, provided all reagents and funding, and edited the manuscript.

2.1 Abstract

Background: The 7S globulins are plant seed storage proteins that have been associated with the development of a number of human diseases, including peanut allergy. Immune reactivity to the wheat seed storage protein globulin-3 (Glo-3) has been associated with the development of the autoimmune disease type 1 diabetes in diabetes-prone rats and mice, as well as in a subset of human patients.

Findings: The present study characterized native wheat Glo-3 in salt-soluble wheat seed protein extracts. Glo-3-like peptides were observed primarily in the wheat embryo. Glo-3-like proteins varied significantly in their molecular masses and isoelectric points, as determined by two dimensional electrophoresis and immunoblotting with anti-Glo-3 antibodies. Five major polypeptide spots were identified by mass spectrometry and N-terminal sequencing as belonging to the Glo-3 family.

Conclusions: These results in combination with our previous findings have allowed for the development of a hypothetical model of the post-translational events contributing to the wheat 7S globulin profile in mature wheat kernels.

2.2 Findings

The 7S globulins, orthologs of the vicilins of the Leguminoseae, are salt-soluble storage proteins that accumulate during seed development (Debiton et al., 2011; Jerkovic et al., 2010). Vicilins were first described by Osborn and Campbell in 1898 as a class of seed storage proteins in *Vicia faba* (horse bean) (Osborne and Campbell, 1898). Both the vicilins and the legumins, distinguishable by their sedimentation coefficients of 7-9S and 11-13S, respectively (Danielsson, 1949), contain characteristic β-barrel cupin domains (Dunwell, 1998). The 7S globulins are translated as preproproteins that, following the co-translational cleavage of the signal peptide, assemble into homo- or heterotrimers (Lawrence et al., 1994) within the lumen of the endoplasmic reticulum (Chrispeels, 1991). Prior to accumulation in seed protein storage vesicles, the trimers undergo post-translational processing, which includes glycosylation and partial endoproteolytic cleavage (Herman and Larkins, 1999, Heck et al., 1993).

Exposure to a number of wheat seed proteins can induce a number of immune-mediated diseases including gluten sensitive enteropathy (celiac disease) (Di Sabatino and Corazza, 2009), Baker's asthma and wheat-dependent exercise-induced anaphylaxis (WDEIA) in predisposed individuals (Tatham and Shewry, 2008). The *Triticum aestivum* (wheat) storage protein WP5212, later named globulin-3 (Glo-3), has been demonstrated to be a potential food allergen (Larre et al., 2011), identified as the first candidate wheat protein associated with the development of type 1 diabetes (T1D) (MacFarlane et al., 2003), and now celiac disease (Taplin et al., 2011). We recently identified the genomic origins of three *Glo-3* genes, *Glo-3A, B* and *C* in the wheat cultivar Glenlea (Loit et al., 2009).

Immunofluorescence studies have localized the Glo-3 gene products to the developing wheat seed embryo and aleurone layer (Loit et al., 2009).

Few studies have sought to characterize wheat 7S globulins because they were thought to be minor storage proteins with little contribution to the bread-making properties of wheat flour (Fabijanski et al., 1985; Robert et al., 1985). However, 7S proteins, based on their sedimentation coefficient, have been characterized in barley and maize, and more recently, two *Glo-3*-like sequences have been identified in the model cereal *Brachypodium distachyon* (Larre et al., 2010). In addition to cultivar Glenlea, Glo-3 proteins have been observed in cultivars Butte 86 and Recital (Dupont et al., 2011; Tasleem-Tahir et al., 2011), indicating that Glo-3 is well-conserved in wheat, thus deserving more attention.

Due to its documented role in the development of T1D, we initiated the present study to characterize the Glo-3-related proteins and peptides in wheat cultivar AC Barrie, the original source of WP5212 (MacFarlane et al., 2003). We hypothesized that Glo-3 undergoes post-translational processing, including glycosylation and endoproteolytic processing, similar to 7S proteins in other species. Therefore, we sought to characterize the expression and the distribution of Glo-3 antigenically related proteins by M_r and pI in the embryo and endosperm of AC Barrie, and to link observed protein fragments with their corresponding endoproteolytic cleavage events.

2.3 Results

2.3.1 Glo-3 antigenically-related proteins co-isolate with wheat globulins

To characterize the Glo-3 antigenically-related proteins in whole AC Barrie seeds, globulins were extracted, following the classical method (Khavkin et al., 1978; Robert et al., 1985). The globulin-enriched fraction was separated by 1D SDS-PAGE and immunoblots were probed with polyclonal rabbit antibodies specific for Glo-3A (Figure 2.1) (Loit et al., 2009). The four most intense protein bands, as resolved by SDS-PAGE, had relative mobilities of 33-36, 47-53 and 64-65 and 66-68 (doublet) kDa. The Glo-3 antigenically-related proteins had comparable M_r to these intense bands (33-37, 47-53, 64-68 kDa). Pre-immune serum and secondary/tertiary antibody controls were negative for immunoreactivity with the Glo-3-related proteins (Figure 2.1).

2.3.2 The Glo-3-related proteins are primarily located in the embryo

Protein expression levels of 7S globulins have been shown to be highest in the embryo and aleurone layers, while almost absent in the endosperm (Burgess and Shewry, 1986; Sun et al., 1996; Thijssen et al., 1996). To study the expression of Glo-3 proteins, AC Barrie endosperm and embryo salt-soluble protein fractions were compared by two-dimensional (2D) electrophoresis according to pI and M_r, followed by immunoblotting using anti Glo-3 antibodies (Figure 2.2). The embryo protein fraction was noticeably more complex than the endosperm fraction, with 287 spots detected by GE Healthcare ImageQuant TL Colony Version 7.0 in the CBB R-250-stained 2D polyacrylamide gel of the embryo protein fraction compared to the 122 spots detected in the 2D gel of the

Figure 2.1: SDS-PAGE and immunoblot analysis of AC Barrie salt-soluble proteins. The salt-soluble fraction from AC Barrie seeds was separated under reducing conditions by SDS-PAGE (12% polyacrylamide) and stained with CBB R-250. Standard lane (M) is Precision Plus Protein (Bio-Rad). Proteins were immunoblotted with polyclonal anti-Glo-3 antiserum at a 1:10,000 dilution, with pre-immune serum (1:10,000), or with secondary and tertiary antibodies alone.

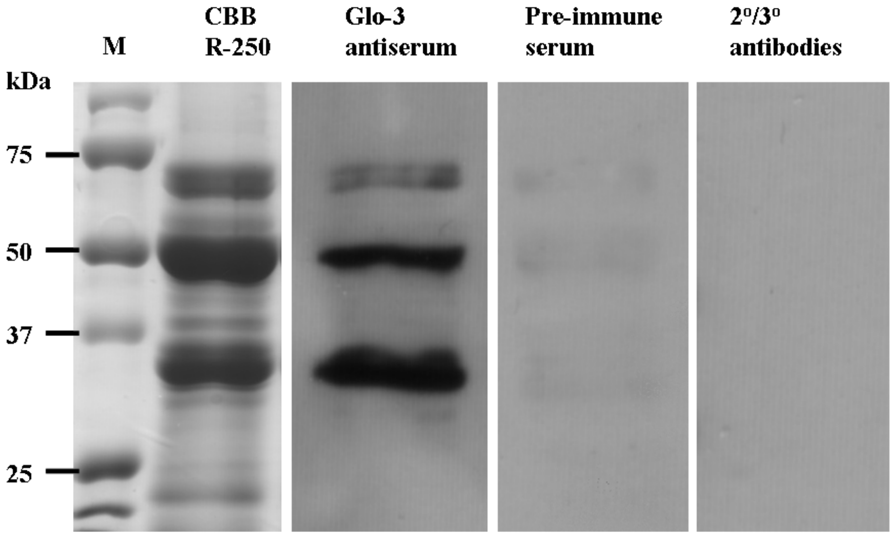
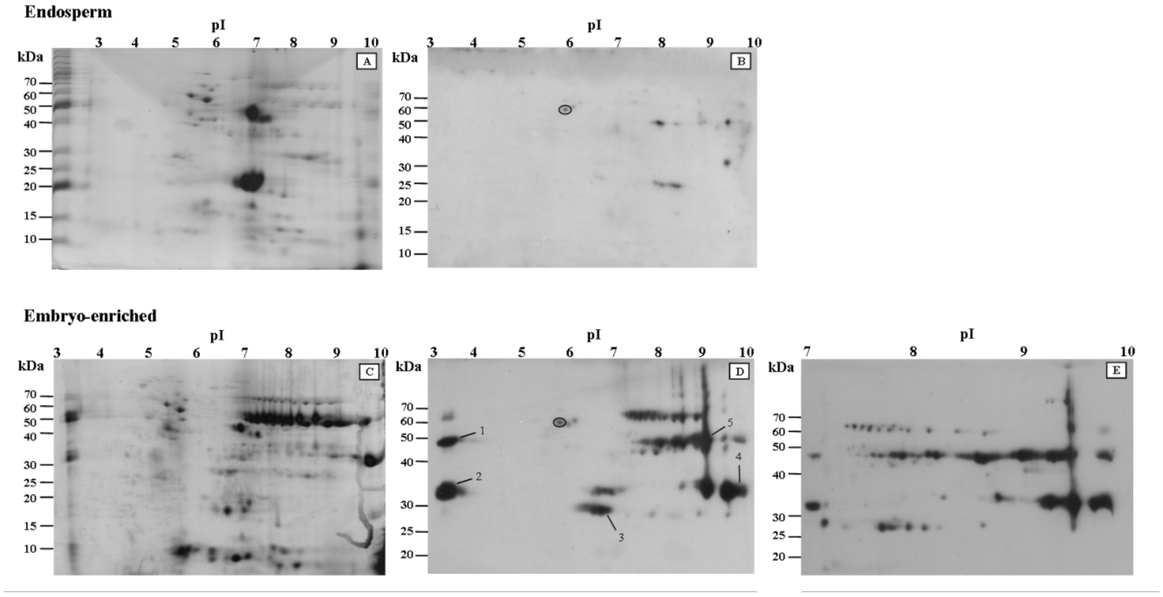

M
CBB
R-250
Glo-3
antiserum
Pre-immune
serum
2°/3°
antibodies
kDa
75
50
37
25

Figure 2.2: 2D gels and immunoblot analysis of AC Barrie salt-soluble proteins. Salt-soluble globulins were extracted from AC Barrie wheat seed endosperm (panels A, B) and embryo-enriched (panels C, D, E) fractions and separated by 2-DE. Proteins were stained with CBB R-250 (panels A, C) or transferred to nitrocellulose and probed with polyclonal rabbit anti-Glo-3 antiserum (panels B, D, E). Marker lanes (M) are Pre-stained Benchmark (Invitrogen). Molecular masses shown on immunoblots are approximations. Spots chosen for mass spectrometry are labelled 1-5 and marked with arrows (panel D), and represent a sampling of the major observed molecular masses (~30 kDa and ~50 kDa) with isoelectric points in the acidic (pH 3), neutral (pH 6-7) and basic (pH 9-10) regions. Circled spots are non-specific spots common to blots probed with pre-immune serum and anti-Glo-3 antibodies.


Endosperm
pI
3 4 5 6 7 8 9 10
kDa
70
60
50
40
30
25
20
15
10
A
B
Embryo-enriched
C
D
E
1
2
3
4
5
7 8 9 10

endosperm protein fraction (Figure 2.2, panels A, C). Analysis of the immunoblots revealed 91 spots corresponding to antigenically-related Glo-3-related proteins in the embryo protein fraction, and 46 spots in the endosperm protein fraction (Figure 2.2, panels B, D). On the basis of the increased anti-Glo-3 immunoreactivity with the salt-soluble embryo protein fraction, further studies focused on the AC Barrie embryo. One immunoreactive spot, with M_r 57 kDa and pI 5.8, was common between blots probed with anti-Glo-3 specific antibodies (circled in Figure 2.2, panels B and D) and with pre-immune serum (data not shown). This spot was considered non-specific for Glo-3 immunoreactivity.

Of the 91 anti-Glo-3 immunoreactive spots in the salt-soluble embryo protein fraction, 59 spots corresponded to the four dominant bands identified in the 1D immunoblot (M_r 33-37, 47-53, 64-65, and 66-68 kDa) (Figure 2.1 and Figure 2.2, panel D). Twelve spots in the M_r 33-37 kDa range and 23 spots in the M_r 47-53 kDa range were observed with 20 spots having pI values between 7.5 and 9.5. Twenty-four spots were in the M_r range of the 64-65 and 66-68 kDa doublet, with 21 spots having pI values between 7.5 and 9.5. As 71 of the 91 spots had pI values between 7 and 10, the salt-soluble embryo protein faction was resolved on a 2D gel with a pH range of 7-10 (Figure 2.2, panel E). There were 103 Glo-3 immunoreactive spots identified within this narrower pH range, with 70 of the 103 spots with M_r of 33-37, 47-53, 64-65, or 66-68 kDa.

2.3.3 Identification of selected Glo-3-related polypeptides

To confirm that the antigenic epitopes detected by the anti-Glo-3 antibodies were specific to Glo-3, five anti-Glo-3-immunoreactive spots from the salt-soluble embryo protein fraction were excised from the 2D gel, and analyzed by mass spectrometry (LC- MS/MS) (Figure 2.2, panel D; numbers indicate location of spots). The spots excised were chosen as they had a wide range of M_r and pI values, and they fell outside the predicted M_r and pI values of proglobulin-3 (GenBank Accession JQ945759) (M_r 66.6 kDa and pI 8.5, as calculated using the Expasy Compute pI/Mw tool). Mass spectrometry results are summarized in Table 1. All five spots were identified as Glo-3A by interrogating the non-redundant NCBI database.

2.3.4 Characterization of selected Glo-3-related polypeptides

To study the post-translational processing of Glo-3, three spots were analyzed with N-terminal sequencing. One sequence (Spot 1) may be N-terminally blocked because no information could be obtained, despite protein visualization after amido black staining. The N-terminal sequence of Spot 3 was determined to be SRDTFNLL, which matched the Glo-3 (GenBank Accession JQ945759) sequence starting at amino acid residue 337. Spot 4 was difficult to visualize following protein transfer and staining. As determined by N-terminal sequencing, the first two residues could not be resolved (X) and the last residue was reported as either arginine (R) or glutamic acid (E). Using arginine as the last residue, the resulting sequence XXHGDSRR matched the findings of Singh et al., and the Glo-3B sequence (GenBank accession FJ439136) starting at residue 117.

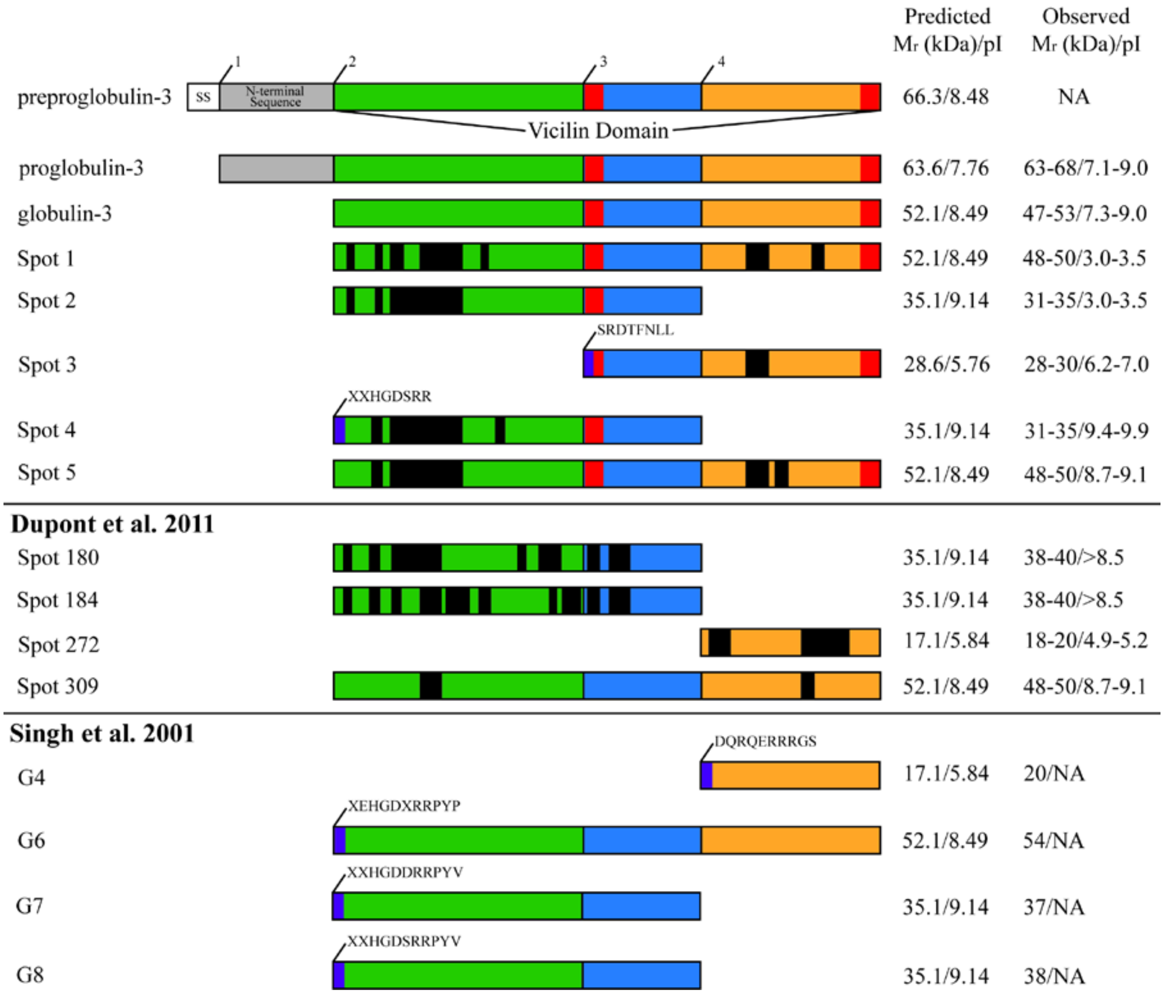
The post-translational endoproteolytic cleavage events of preproglobulin-3 that would be required to yield polypeptides with M_r and pI corresponding to the sequenced spots are

Table 2.1: MS/MS sequencing results of selected gel spots of wheat 7S globulins

Spot ID	Protein (GenBank accession)	Mascot Score	% coverage	Peptides
1	globulin-3A (JQ945759)	210	17.5	VFGPRSF; DEVSRLL; HTISVPGKF; GRPAREVQEVF; RVAIMEVNPRAF; TVRQGDVIVAPAGSIMHL; VVPPGHPVVEIASSRGSSNL; VAQGEGVLTVIENGEKRSY
2	globulin-3A (JQ945759)	131	12.9	VFGPRSF; DEVSRLL; VVPGLTDADGVGY; RVAIMEVNPRAF; TVRQGDVIVAPAGSIMHL; VAQGEGVLTVIENGEKRSY
3	globulin-3A (JQ945759	38	3.4	VVPPGHPVVEIASSRGSSNL
4	globulin-3A (JQ945759)	135	13.8	RPFDEVSRLL; RVAIMEVNPRAF; VAQGEGVLTVIENGEKRSY; SAKPLLASL; TVRQGDVIVAPAGSIMHL; VVPGLTDADGVGY
5	globulin-3A (JQ945759)	144	17.7	DEVSRLL; RPFDEVSRL;VVPGLTDADGVGY; RVAIMEVNPRAF; EINAERNERVWL; TVRQGDVIVAPAGSIMHL; VVPPGHPVVEIASSRGSSNL; VAQGEGVLTVIENGEKRSY

summarized in Figure 2.3. In addition to the M_r and pI of the spots in Figure 2.2, panel D, the location of the MS sequenced peptides within globulin-3 (black bars), N-terminal sequence data (purple bars - when available), as well as the location of the epitopes used to generate the polyclonal anti-Glo-3 antibodies (red bars) were considered. The size and location of the signal sequence was predicted by TargetP 1.1 (Emanuelsson et al., 2000; Nielsen et al., 1997), (http://www.cbs.dtu.dk/services/TargetP/). Proteins from previous studies (Dupont et al., 2011; Singh et al., 2001) were included to demonstrate that the methods used for the determination of the processing events in the current study are applicable to previously published findings and that our observed processing corresponds to the processing inferred in the literature. The processing of Spot 3 (Figure 2.3) is supported by the N-terminal sequence that matched cleavage site 3 (SRDTFNLL), as well as the MS sequenced peptide in the C-terminal vicilin domain segment, and an observed M_r 28-30 kDa and pI 6.2-7.0. Spot 3 corresponds to the cleavage of proglobulin-3A at cleavage site 3, with predicted M_r 28.6 kDa and pI 5.76. The processing of Spot 4 is supported by N-terminal sequence data that match the previously documented cleavage site 2 (XXHGDSRR) (Singh et al., 2001). Spot 4 corresponds to cleavage of proglobulin-3A at cleavage sites 2 and 4, with predicted M_r 35.1 kDa and pI 9.14. Spot 2 lacked N-terminal sequence data, but as with Spot 4, contained MS sequenced peptides present in the N-terminal vicilin domain segment and an observed M_r 31-35 kDa. Cleavage at both sites 1 and 3, or 2 and 4 yield products with M_r of approximately 35 kDa. The epitopes used for the generation of the antiglobulin-3 antibody are present in the middle and C-terminal segments. Neither of those epitope sites is present in products created from cleavage sites 1 and 3, so therefore Spot 2 corresponds to cleavage at sites 2

Figure 2.3: Model of Glo-3 endoproteolytic processing.
The observed M_r, pI, MS/MS sequence data, and N-terminal sequence data from this study and previous studies (Dupont et al., 2011) were reconciled with theoretical peptides created by the endoproteolytic cleavage of preproglobulin-3 (GenBank Accession JQ954759). A linear representation of preproglobulin-3 is shown with approximate locations of potential cleavage sites, labelled 1-4. The protein domains are represented as follows: signal sequence (SS) (white); N-terminal sequence (grey); vicilin domain (divided into three segments by cleavage sites – N-terminal segment (green), middle segment (blue), C-terminal segment (orange). Red boxes correspond to the location of the linear epitopes used when creating the anti-Glo-3 polyclonal antibodies. Black boxes correspond to the location of MS/MS sequenced peptides. Purple boxes correspond to sequence obtained from N-terminal sequencing. The observed and expected molecular masses (kDa) and isoelectric points (pI) of the resulting polypeptides following endoproteolytic cleavage is indicated on the right. Additionally, the expected processing of proteins inferred by previous studies (Dupont et al., 2011) and (Singh et al., 2001) has been demonstrated.


Predicted Mr (kDa)/pI
Observed Mr (kDa)/pI
1
2
3
4
preproglobulin-3
SS
N-terminal Sequence
Vicilin Domain
66.3/8.48
NA
proglobulin-3
63.6/7.76
63-68/7.1-9.0
globulin-3
52.1/8.49
47-53/7.3-9.0
Spot 1
52.1/8.49
48-50/3.0-3.5
Spot 2
35.1/9.14
31-35/3.0-3.5
SRDTFNLL
Spot 3
28.6/5.76
28-30/6.2-7.0
XXHGDSRR
Spot 4
35.1/9.14
31-35/9.4-9.9
Spot 5
52.1/8.49
48-50/8.7-9.1
Dupont et al. 2011
Spot 180
35.1/9.14
38-40/>8.5
Spot 184
35.1/9.14
38-40/>8.5
Spot 272
17.1/5.84
18-20/4.9-5.2
Spot 309
52.1/8.49
48-50/8.7-9.1
Singh et al. 2001
DQRQERRRGS
G4
17.1/5.84
20/NA
XEHGDXRRPYP
G6
52.1/8.49
54/NA
XXHGDDRRPYV
G7
35.1/9.14
37/NA
XXHGDSRRPYV
G8
35.1/9.14
38/NA

and 4. Spots 1 and 5 lack N-terminal sequence data, but had MS sequenced peptides that localized to the N-terminal and C-terminal segments of the vicilin domain, and exhibited relative mobilities of 48-50 kDa (Figure 2.3). The pI of Spot 1 was 3.0-3.5, while the pI of Spot 5 was 8.7-9.1 Both Spot 1 and Spot 5 correspond to the theoretical cleavage of preproglobulin-3 at cleavage site 2 with predicted M_r 52.1 kDa and pI 8.49.

2.4 Discussion

2.4.1 Type 1 diabetes

Inflammation associated with the T-cell-mediated autoimmune disease T1D results in the loss of the insulin-producing β cells in the pancreatic islets of Langerhans (Eizirik et al., 2009). While the incidence of T1D has steadily increased in developed countries over the past 60 years, a definitive cause of T1D has yet to be elucidated (Todd, 2010). While many risk genes for T1D have been identified, it has been proposed that most T1D-related genes are not highly penetrant and that T1D is actually a complex disease requiring both genetic susceptibility and the exposure to environmental risk factors (MacFarlane et al., 2009). Serum IgG antibodies from human patients with T1D, but not from matched control groups, are able to bind Glo-3 *in vitro*, indicating that Glo-3 may be a candidate risk factor for the development of T1D in certain individuals (MacFarlane et al., 2003). We therefore chose to characterize the Glo-3-related proteins in the wheat cultivar AC Barrie with respect to protein maturation in developed seeds.

2.4.2 Characterization of Glo-3

When the amino acid sequences of preproglobulin-3A from the wheat cultivars AC Barrie and Glenlea are compared, there are five amino acid substitutions out of 588 residues (Loit et al., 2009). However, the theoretical isoelectric points of these Glo-3 proteins are 8.48, and 7.78 in AC Barrie and Glenlea, respectively. The difference in pI of 0.70 is due to the substitution of two arginine residues (R43Q and R102H) in the AC Barrie Glo-3. The AC Barrie salt-soluble Glo-3-related proteins demonstrate a range of isoelectric points concentrated in the basic range (pH 7-9) when separated by 2DE (Figure 2.2, panel E). As there are three *Glo-3* genes in the wheat cultivar Glenlea (Loit et al., 2009), it is feasible that AC Barrie would also have three *Glo-3* genes, as both AC Barrie and Glenlea are closely related hexaploid wheat cultivars (McCallum and DePauw, 2008). The charge trains of gel spots with similar M_r and a range of pI values at the major size groups (Figure 2.2) are likely due to post-translational processing and modifications, as previously discussed (Robert et al., 1985), as well as amino acid substitutions between the multiple immunologically related Glo-3 proteins in AC Barrie, or from artifacts generated during the execution of the extraction and examination protocols (Deng et al., 2012) .

When the wheat cultivar AC Barrie salt-soluble globulin fraction was resolved by 1D SDS-PAGE under reducing conditions, bands of three major size ranges of approximately 64-70 kDa, 47-53 kDa and 33-37 kDa were visible (Figure 2.1); referred to as 65 kDa, 50 kDa and 35 kDa, respectively. These three major wheat globulin groups are characteristic of the 7S globulins described in other studies (Burgess and Shewry, 1986; Robert et al., 1985). Intriguingly, while Spot 3 had an observed M_r of 28-30 kDa (Figure 2.2D), no distinct band is visible in this region in Figure 2.1, though a faint band is visible just below the 35 kDa band. The 28-30 kDa band is likely less intense than the major bands, as Spot 3 is the only significant immunoreactive spot at 28-30 kDa, while the doublet at 65 kDa comprises several spots and is still markedly less intense than the 50 and 35 kDa bands, which both include several intense spots.

The 7S globulins are enriched in the embryo and aleurone layer of wheat and other cereals (Heck et al., 1993; Khavkin et al., 1978; Kriz, 1999; Yupsanis et al., 1990). Consistent with our previous findings (Loit et al., 2009), immunoblots probed with anti-Glo-3 polyclonal serum revealed that 7S globulins are expressed at low levels in the salt-soluble fraction of wheat endosperm (Figure 2.1). In the embryo-enriched fraction, Glo-3 related proteins are restricted to each of the major size groups, suggesting that these Glo-3-related proteins have similar post-translational processing patterns as the previously characterized 7S globulins.

2.4.3 Post-translational processing of Glo-3

The 7S storage proteins in other plants undergo a series of post-translational modifications which include limited endoproteolytic events (Gatehouse et al., 1983; Sharma et al., 2010; Spencer et al., 1983). Glycosylation is frequently observed in 7S globulins as N-linked complex glycans (Sturm et al., 1987), but is not required for proper folding or export to the protein storage vacuole (PSV) (Chrispeels et al., 1982). A previous study has shown that wheat 7S globulins bind the lectin concanavalin A, although the exact nature and extent of this binding is unknown (Robert et al., 1985). Additionally, other cleavage events may occur at the N-terminus of Glo-3 as observed for maize Glb1 (Schwartz, 1979). These post-translational modifications, among others, likely contribute to the heterogeneity of the observed isoelectric points as well as the wide range of observed molecular masses among the Glo-3-related proteins.

The MS/MS sequencing of the numbered spots in Figure 2.2, panel D, all returned Glo-3A sequences when queried against the non-redundant NCBI protein database, indicating that, similar to orthologous 7S storage proteins in other species (Heck et al., 1993), the peptide spots in Figure 2.2, panel D were derived from a common precursor protein. The Glo-3-related proteins found at the acidic end of the pH spectrum, including the proteins in spots 1 and 2, exhibited a pI of approximately 3.0 (Figure 2.2, panel D). The low pI of these proteins could be due to protein precipitation during the rehydration of the IPG strips (Sanchez et al., 1997), as similar spots are visible at pH 7.0 in the immunoblot of the 2D gel of pH 7-10 (Figure 2.2, panel E).

Considering features such as the location of the MS/MS sequenced peptides within Glo-3, N-terminal sequence data (when available), as well as the M_r and pI of the spots, we were able to deduce the likely cleavage pattern of the preproglobulin-3 precursor protein (Figure 2.3) that resulted in the globulin-3 polypeptide spectrum recognized by anti-Glo-3 antibodies (Figure 2.2, panel D). It must be noted that the antibodies used to probe the globulin fraction were raised against two linear epitopes from the vicilin domain of WP5212: SRDTFNLLEQRPKIAN and RGDEAVEAFLRMATA (MacFarlane et al., 2003; Scott, 2007)**.** These epitopes are located in the middle (blue) and C-terminal (orange) vicilin domain segments (Figure 2.3). Therefore, N-terminal polypeptides created by the endoproteolytic cleavage at site 3 may not be recognized by the anti-Glo-3 antibodies used in this study.

The model of endoproteolytic cleavage presented in Figure 2.3 suggests that the Glo-3-related proteins of varying molecular masses and isoelectric points originate from 7S globulin precursors possessing a signal peptide, an N-terminal segment and a vicilin domain (Figure 2.3). These precursors belong to a multigene family whose members differ slightly in sequence, with varying length N-terminal segments (Lawrence et al., 1994). We propose that the Glo-3 protein, prior to co-translational removal of the signal sequence, is a ~66 kDa monomer (Figure 2.3). After signal sequence removal (cleavage site 1), the protein becomes a ~64 kDa holoprotein. Three internal cleavage sites have been identified that, when processed, can yield polypeptides with a range of M_r and pI values (Figure 2.3). Spot 3 (Figures 2, panel D, Figure 2.3) was observed at M_r 28-30 kDa and pI 6.2-7.0, and Spot 4, which has a similar M_r (31-35 kDa), a more basic pI (9.4-9.9), and a different N-terminus than Spot 3. The N-terminal sequence determined for Spot 3 SRDTFNLL represents a novel cleavage site for globulin-3. The location of this cleavage site is consistent with processing by vacuolar processing enzymes (Ariizumi et al., 2011), as an aspartic acid residue precedes the N-terminal sequence of Spot 3. These findings suggest that these ~30 kDa globulin-3-related polypeptides could arise from different processing events of the same precursor, as was observed for barley Beg1 (Heck et al., 1993; Yupsanis et al., 1990). Curiously, there were no proteins identified in the immunoblotting experiments that corresponded to the product of processing at cleavage site 4, as was documented in other studies that sequenced globulin-3 (Figure 2.3 and (Dupont et al., 2011; Singh et al., 2001; Tasleem-Tahir et al., 2011)). Potentially, as the antiglobulin-3 antibodies are polyclonal and were prepared by the co-immunization of rabbits with a mixture of two separate peptides (Sancho et al., 2008), it is possible that the antibodies recognize only one of the two peptides. However, further study on the binding of these antibodies to processed globulin-3 proteins is still required in order to achieve a more definitive understanding.

When compared to the Glo-3A, Glo-3B, and Glo-3C sequences previously published by our group (Loit et al., 2009), each of the spots contained at least one peptide in Table 1 that mapped solely to the Glo-3A coding sequence (GenBank Accession JQ945759 - data not shown). As the Glo-3A protein coding sequence is 99% identical between wheat cultivars Glenlea (GenBank accession ACJ65514) and AC Barrie (JQ945759), the Glo-3B and Glo-3C proteins may share a high percentage identity between these two cultivars. Therefore, we attributed all spots to Glo-3A. However, on close inspection of the Glo-3B (FJ439136) genomic sequence, there are several insertion/deletion events (data not shown) that would yield a protein of similar molecular weight as Glo-3A (66.3 kDa). Further sequencing of Glo-3 cDNA clones from both AC Barrie and Glenlea cultivars is necessary to determine the transcribed sequence of Glo-3.

The endoproteolytic processing events outlined here are likely part of a series of post-translational events that lead to the maturation of the Glo-3-like proteins, also similar to those observed for barley Beg1 (Heck et al., 1993). The identification of wheat Glo-3 holoproteins (~65-70 kDa) by immunoblot analysis suggests that the endoproteolytic modifications described are partial and may not be prerequisites for proper folding, transport, targeting, and storage of the Glo-3 proteins. Alternatively, as the wheat genome has not been sequenced, the spots with M_r 65-70 kDa may correspond to proteins in the globulin-3 family that encode proteins recognizable by the antiglobulin-3 antibodies, but are processed in a different manner than the proteins observed in this study. Further sequencing and study are required in order to fully catalogue the globulin-3 family. The presence of polypeptides of ~50 kDa and ~30-38 kDa (Singh et al., 2001 and this study) with the same N-termini reinforce the idea that not every hypothesized set of endoproteolytic processing events occurs, and some could occur with varying degrees of processing. Such maturation of vicilins has not only medical implications, but also implications for the quality of foods rich in vicilins following processing and production (e.g. wheat, soybean (Kaviani et al., 2011; Sancho et al., 2008)).

2.5 Conclusion

With a greater understanding of the endoproteolytic processing events that lead to the maturation of the Glo-3 family of proteins observed in the salt-soluble embryo protein fraction, we can now refine the research of Glo-3 to certain domains present only in the polypeptides that are associated with T1D or celiac disease following endoproteolytic processing. In addition, the specific breeding or genetic modification of wheat could be performed to minimize any potential disease- or food quality-related protein or peptide content compared to that of existing wheat cultivars, as is underway for the deletion of the conglycinin α' subunit from the soybean proteome (Rayhan et al., 2011; Stanojevic et al., 2011).

2.6 Methods

2.6.1 Wheat seed protein extraction and sample preparation

For each extraction, 4 g of whole *Triticum aestivum* AC Barrie seeds or 50-100 embryos or endosperm, which were dissected as in (Walker-Simmons, 1987) were employed. Seeds were initially ground in a domestic coffee grinder. Finer powder was obtained by hand milling using a mortar and pestle under liquid nitrogen. Wheat powder was mixed with 10 volumes of fresh ice-cold acid denaturing solution (10% (w/v) trichloroacetic acid (TCA) and 0.05% (w/v) dithiothreitol (DTT) in acetone). Samples were stirred at $4^{o}C$ for 1 h and left at $-20^{o}C$ overnight. The resulting suspension was centrifuged for 30 min at 35,000 x *g* at $4^{o}C$ (Beckman Avanti J-25, rotor JA 25.50). The supernatant was decanted and the pellet resuspended in ice-cold acetone containing 0.05% (w/v) DTT. The mixture was extracted by incubation for 1 h at $-20^{o}C$. The suspension was centrifuged for 20 min at 35,000 x *g* at $4^{o}C$, the supernatant decanted and the pellet dried on ice. Dried powder was extracted with 15 ml of 1 M NaCl solution (1.0 M NaCl; 0.05 M Tris; pH 8.0). The mixture was stirred for 1 h at room temperature and centrifuged (27,000 x *g* for 30 min at $22^{o}C$). Supernatants were collected and the extraction with 1 M NaCl solution was repeated. The supernatants were pooled and dialysed against 5 changes of ddH_2O for 24 h at $4^{o}C$. Precipitates were collected by centrifugation at 35,000 x *g* at $4^{o}C$ for 45 min. Pellets were drained and extracted for either one-dimensional (1D) or two-dimensional protein separation.

2.6.2 Sample preparation for 1D separation

Pellets were resuspended in 1 M NaCl solution. Samples were centrifuged at 20,000 x *g* for 1 h at 22^{o}C and quantified by the bicinchoninic acid (BCA) assay (Smith et al., 1985). Supernatants were aliquoted and stored at -80^{o}C until use.

2.6.3 Sample preparation for 2D separation

Proteins from pellets were extracted with 1.5 - 3 ml rehydration buffer (8M/2M deionized urea/thiourea, 2% 3-([3-cholamidopropyl]dimethylamino)-1-propanesulfonate (CHAPS), 50 mM DTT, 0.0005% bromophenol blue) containing pH 3-10 ampholytes (Bio-Rad). Following sonication, samples were centrifuged at 200,000 x *g* for 1 h at 22^{o}C in the Beckman TL-100 ultracentrifuge. Supernatants were collected, quantified by Bradford method (Bradford, 1976), aliquoted and stored at -80^{o}C until use. Throughout the 2D separation process, care was taken not to heat the urea/thiourea-containing solutions above 30^{o}C to avoid carbamylation of amino groups which can lead to artifactual spot heterogeneity (Shaw and Riederer, 2003).

2.6.4 1D SDS-PAGE protein fractionation

Protein extracts were fractionated under SDS-PAGE reducing conditions. Protein samples were combined in a 1:1 (v/v) ratio with 2X sample buffer (4% SDS, 20% glycerol, 0.12 M Tris (pH 6.8), 10% (v/v) β-mercaptoethanol, 0.01% bromophenol blue). Samples were boiled for 5-10 min and centrifuged at 20,000 x *g* for 10 min. All samples were loaded on discontinuous (5% stacking, 10%-12% resolving) 1.5 mm SDS polyacrylamide gels (Laemmli, 1970). The SDS running buffer, adjusted to pH 8.3 with NaOH, consisted of 25 mM Tris base, 0.19 M glycine and 0.1% SDS. Electrophoresis was performed in the Mini-PROTEAN 3 System (Bio-Rad) at 100-150 V (400 mA) until dye front reached the bottom of the gel. Gels were either stained with Coomassie Brilliant Blue (CBB) R-250 or were used for immunoblot analysis.

2.6.5 Immunoblot analysis

Proteins from SDS-PAGE were transferred under semi-dry conditions by means of the Trans-Blot SD Semi-Dry Transfer Cell (Bio-Rad). Gels were rinsed 5-10 min in semi-dry transfer buffer (24 mM Tris, 192 mM glycine and 15% methanol) and transferred to nitrocellulose membranes (Bio-Rad) also soaked in transfer buffer. Electroblotting proceeded for 1-1.5 h at 11 V (400 mA). Once complete, the transfer was verified by staining membranes with Ponceau S (0.2% (w/v) in 1% glacial acetic acid) followed by several washes of ddH$_2$O. Prior to immunoblotting, membranes were destained with TBST buffer (10 mM Tris-HCl (pH 7.3), 0.1 M NaCl and 0.5% Tween-20) and were incubated 30 min with shaking, in 5% skim milk powder in TBST buffer at room temperature. Membranes were incubated with polyclonal rabbit anti-Glo-3 antibody (MacFarlane et al., 2003) (diluted 1:10,000 in 5% skim milk powder in TBST and 0.05% NaN$_3$) and left overnight at 4^{o}C with gentle rocking. Next day, membranes were washed 4 x 5 min in TBST, and incubated 1 h in biotin-SP-conjugated AffiniPure goat anti-rabbit IgG (H+L) secondary antibody (Jackson ImmunoResearch Laboratories, Inc., West Grove, PA) with a dilution of 1:100,000. This incubation was followed by another set of washes in TBST and a final incubation in affinity purified goat anti-biotin-horseradish peroxidase (HRP)-conjugated tertiary antibody (Cell Signaling Technology, Inc. Danvers, MA) for 1 h. Following 4 x 5 min washes, membranes were treated with ECL™ Western Blotting Detection Reagents (Amersham Biosciences, Piscataway, NJ) for 1 min, exposed to Kodak BioMax Light Film (Fisher Scientific, Nepean, ON). Film was developed by the Kodak X-OMAT 2000A Processor.

2.6.6 Two-dimensional gel electrophoresis (2DE)

Proteomic analysis of the AC Barrie salt-soluble proteins was conducted by 2DE. The proteins were separated in the first dimension by isoelectric focusing (IEF) according to their isoelectric points (pI). Extracts were applied to linear immobilized pH gradient (IPG) 7 cm strips with pH ranges 3-10 and 7-10 (Bio-Rad, Mississauga, ON). A protein load of 100-150 μg was applied to strips for staining and immunoblot applications; 250 μg for mass spectrometry and N-terminal sequencing applications. Strips were covered in mineral oil and rehydrated overnight in a reswelling tray in 125-150 μl volumes of sample diluted in rehydration buffer. For strips in the pH ranges 3-10, DTT was used as a reducing agent; tributylphosphine (TBP) was used for strips in the pH range of 7-10. Rehydrated strips were focused in the PROTEAN IEF Cell (Bio-Rad) according to the following program: step1, 250 V (30 min), linear ramp; step2, 4000 V (2 h), linear ramp; step3, 4000 V (10,000 Vh), rapid ramp. Immediately following isoelectric focusing, the strips were either stored at -80^{o}C for future use or equilibrated prior to running the second dimension. Strips were thawed, if necessary, and reduced for 15 min in 2 ml of equilibration buffer (6 M urea, 2% SDS, 20% glycerol, 0.375 M Tris-HCl, pH 8.8

containing 2% (wt/vol) DTT or 5 mM TBP). A second step, one of alkylation, was performed for 15 min in the same equilibration buffer with the exception that 2.5% (w/v) iodoacetamide was substituted for reducing agents DTT or TBP. Overlay agarose (Bio-Rad) was used to seal the equilibrated strips at the top of 1 mm vertical resolving gels (10% or 12% polyacrylamide). Second dimension SDS-PAGE was performed using the Bio-Rad MiniProtean 3 at 100 V (400 mA); runs were terminated when the bromophenol dye front had reached the end of the gel. Protein spots were counted using ImageQuant TL Colony Version 7.0 (GE Healthcare).

2.6.7 Liquid chromatography tandem mass spectrometry (LC-MS/MS)

Gels were carefully manipulated under a laminar flow hood at all times to reduce keratin contamination. Following separation in the second dimension, proteins were detected with Bio-Safe Coomassie (Bio-Rad) according to manufacturer's instructions and destained with multiple changes of ddH_2O. Selected spots were excised from the gel by means of a clean scalpel blade and subjected to mass spectrometric analysis (Ottawa Institute of Systems Biology, OISB). Proteins were digested in-gel with chymotrypsin as previously described (Wilm et al., 1996). Trypsin digests were avoided because of the high frequency of arginine and lysine residues in the reported wheat globulin sequences (Kriz, 1999; Loit et al., 2009). Peptides were separated by liquid chromatography on an Agilent 1100 Series HPLC System (Agilent Technologies, Palo Alto, CA) and applied by electrospray to a QSTAR Pulsar quadrupole-TOF mass spectrometer (ABI/MDS Sciex, Concord, ON) as described in (Vasilescu et al., 2005). Resulting peptide masses were used to interrogate the non-repetitive NCBI protein database (06/10/2011; 1,4,324,397 sequences; 4,906,523,086 residues) using Mascot software (version 2.3) (Matrixscience Ltd.) as in our previous studies (Koziol et al., 2012; Wall et al., 2010). Fixed modifications were set for carbamidomethyl (C) and variable modifications for oxidation (M). One missed cleavage was allowed. Peptide and MS/MS mass tolerances permitted were ±100 ppm and 0.2 Da, respectively.

2.6.8 N-terminal sequencing

Gels used for N-terminal sequencing were treated for wet transfer to PVDF membranes. Briefly, the gels were soaked in CAPS (3-(cyclohexylamino)-1-propanesulfonic acid) electroblotting buffer (10 mM CAPS/NaOH pH 11 and 10% methanol) for 5 min. Sequiblot PVDF (Bio-Rad) membranes were wet in 100% methanol, soaked in CAPS electroblotting buffer along with Whatman 3MM paper sheets and transblot sponges. The blotting sandwich was assembled and run in the Mini Trans-Blot Cell (Bio-Rad) at 50 V (170 mA) for 45 min. After disassembly, the membranes were rinsed thoroughly with ddH$_2$O, saturated in 100% methanol and stained for a minimum of 1 min in amido black stain (0.1% amido black, 1% acetic acid and 40% methanol). Membranes were rinsed in multiple changes of ddH$_2$O and air dried before excision of spots. Spots of interest were subjected to Edman degradation using the 494 cLC PROCISE Sequencing System (Applied Biosystems; Foster City, CA) for high sensitivity (femtomole quantities) N-terminal protein sequencing (University of Texas Medical Branch, UTMB, Biomolecular Resource Facility Core).

Chapter 3

Taking stock of the protein remnants on the battlefield between host and pathogen: Maize CL30-*Fusarium* interactome

Koziol AG, Reid L, Altosaar I

Submitted to: *Journal of Proteomics*

Contribution of authors:

AGK contributed to the study design, performed all experiments, analyzed data, contributed all figures, and wrote and edited the manuscript. LR contributed to study design, provided field samples, and edited the manuscript. IA contributed to the study design, data analysis and editing.

3.1 Abstract

The economic losses associated with gibberella ear rot make necessary a better understanding of both the pathogen, *Fusarium graminearum*, and its host, *Zea mays*. The *F. graminearum* susceptible inbred maize line CL30 was exposed to water controls or conidial suspensions of *F. graminearum* through kernel inoculations, and the host plants were allowed to grow to maturity. Proteins were extracted separately from the maize embryo and endosperm tissues, as well as from the fungal remnants on the surface of the maize kernels. The proteins were sequenced using direct-sequencing LC-MS/MS. We identified 250 maize proteins with high confidence. These proteins were classified into several functional categories, including: late embryogenesis abundant, seed storage, metabolism, and defense. There was a significantly increased representation of defense proteins in both the embryo and endosperm tissues of infected maize samples. The proteome of the fungal remnants on the surface of the maize kernels was composed of 18 proteins that were present after desiccation. Several of these proteins were categorized as being involved in the metabolism of plant-sourced molecules, or in stress response. Taken together, these data should help dissect the interactive system between host and pathogen during the important time period of grain maturation and harvest.

3.2 Introduction

Gibberella ear rot and stalk rot in maize, as well as Fusarium head blight in wheat, barley, and oats are caused through the infection of these crops by the fungal pathogen *Fusarium graminearum* (Holbert et al., 1923; Miller et al., 1985; Reid et al., 1992; Takeda and Heta, 1989; Trail, 2009). Not only do these diseases reduce grain yields (Kazan et al., 2012), but *F. graminearum* also synthesizes mycotoxins, such as deoxynivalenol (DON), fusarin C, autofusarin, and zearalenone (Logrieco et al., 2002; Taylor et al., 2008). One of the best studied mycotoxins, DON, is a virulence factor that aids in fungal colonization, and also poses serious human health risks (Harris et al., 1999). For instance, DON inhibits eukaryotic protein biosynthesis, and upon ingestion, causes nausea, vomiting, and convulsions (Goswami and Kistler, 2004; Vesonder et al., 1981), with long-term exposure to DON being linked to immune system suppression and certain neurological disorders (Bennett and Klich, 2003; Pestka and Bondy, 1994). Consequently, the concentration of DON in food products is tightly regulated, with limits of 1 µg DON per gram of food for human consumption (Trail, 2009).

The sequence of the four chromosomes of the *F. graminearum* strain PH-1 (teleomorph *Gibberella zeae*) genome was published in 2007 (Cuomo et al., 2007). The genome is 36.45 Mb and codes for 13 321 predicted proteins (www.broadinstitute.org/annotation/genome/fusarium_group/GenomeStats.html; last viewed 7/09/2012). Compared to other sequenced fungal genomes, *F. graminearum* has an overrepresentation of certain protein categories, including hydrolytic enzymes (Cuomo et al., 2007), which may have implications in pathogenesis. The infection process is initiated for gibberella ear rot when ascospores or conidia are introduced to exposed silks or punctured kernels (Koehler, 1942; Miller et al., 2007; Sutton, 1982). Following germination of the spores, the fungus can then colonize the host plant (Mesterhazy et al., 2012). However, in response to fungal infection, plants have several inducible defense strategies, including the strengthening of cell walls and the production of antimicrobial proteins (Campo et al., 2004; Reid et al., 1992; Reid et al., 1994).

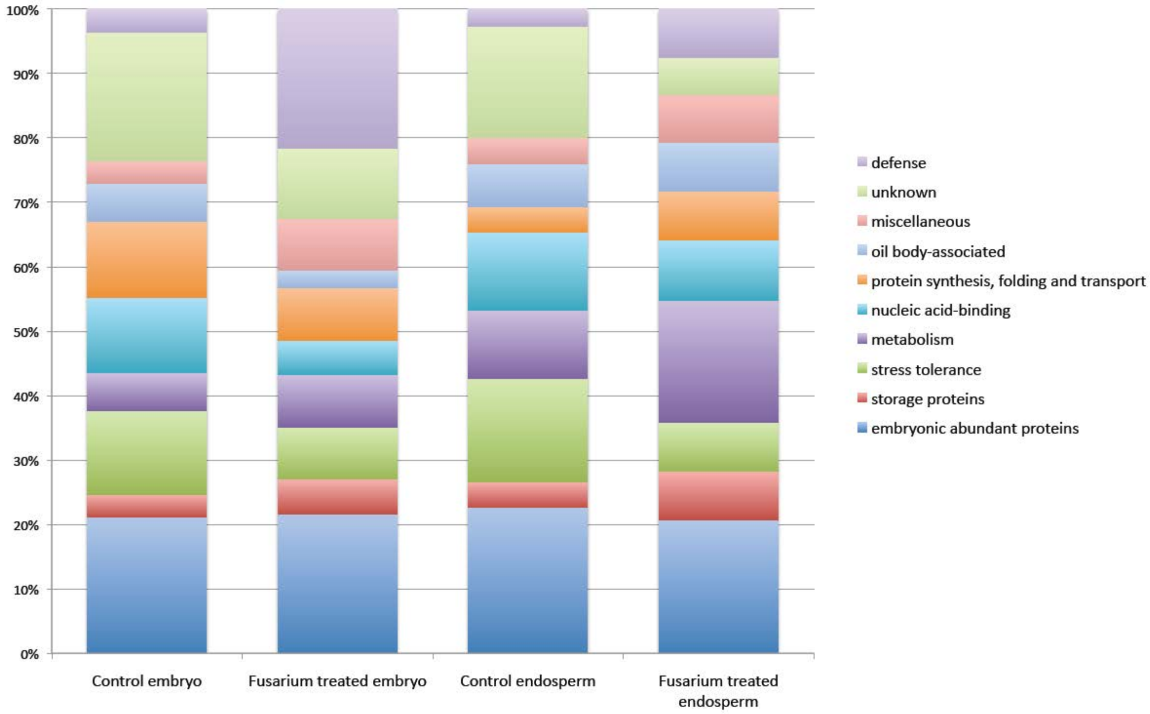
We have previously examined the starch granule-associated proteomes of wheat (Wall et al., 2010), and of commercially prepared rice and maize starches (Koziol et al., 2012; Koziol et al., 2012). In these experiments, it was presumed that the grains used in the preparation of starches were healthy and free of obvious fungal contamination. The use of CL30, a susceptible host genotype of maize, and the mechanical kernel injection technique promotes a fuller expression of the foreign proteome. While there have been proteomic sequencing studies of *F. graminearum* within the host following infection, to our knowledge, the proteome of the fungal remnants on the surface of maize seeds at maturity has not previously been determined. We hypothesized that the proteins sequenced within these fungal remnants could provide insight into the biochemical processes occurring in the fungus prior to desiccation, and that even though CL30 is susceptible to infection by *F. graminearum*, we would be able to identify host defense-related proteins present in both the endosperm and embryo tissues. To that end, we initiated a gel-free protein sequence project using LC-MS/MS to investigate the interplay of the proteomes of host and pathogen.

3.3 Results

Kernels of the inbred maize line, CL30, were injected at 10-15 days post-silk emergence with either water controls or conidial suspensions of *F. graminearum* and allowed to develop to maturity. Following harvest, kernels from both treatments were collected, and the proteomes of discrete tissues from the kernels, as well as the remnants of *F. graminearum* on the surface of the maize seeds were sequenced using LC-MS/MS. Sequences were analyzed and classified by biochemical function.

Cumulatively, 250 plant-sourced proteins were identified in the water control or *Fusarium* treated embryo and endosperm samples. These proteins were classified based upon their molecular function, with the classes as follows: maturation; storage; stress tolerance; metabolism; nucleic acid binding; protein folding, synthesis and transport; oil body-associated; miscellaneous; unknown; and defense. This classification is shown in Figure 3.1. When comparing the percent representation between the control and infected samples, the largest differences in percent representation include nucleic acid binding (11.8% in control and 5.4% in the infected) and oil body associated (5.9% in the control and 2.7% in the infected) in the embryo samples. The storage (4% in the control and 7.5% in the infected), stress tolerance (16% in the control and 7.5% in the infected) and metabolism (10.7% in the control and 18.9% in the infected) proteins in the endosperm also showed differences in percent representation. Additionally, the defense proteins had different representations in both the embryo and endosperm samples.

Figure 3.1: Percent representation of maize functional protein categories. Proteins were isolated from both the *Fusarium* infected and control embryo and endosperm tissues and sequenced by LC-MS/MS. Functional categories are based on KEGG classifications (Kanehisa and Goto, 2000).


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Control embryo
Fusarium treated embryo
Control endosperm
Fusarium treated endosperm
defense
unknown
miscellaneous
oil body-associated
protein synthesis, folding and transport
nucleic acid-binding
metabolism
stress tolerance
storage proteins
embryonic abundant proteins

That both the embryo and endosperm samples had different representation of defense proteins was of particular interest. The percentage representation of defense proteins in the control embryo, infected embryo, control endosperm, and infected endosperm samples was 3.5%, 21.6%, 2.7%, and 7.5%, respectively. The defense proteins identified in each sample are shown in Table 3.1, which includes GenBank accession number, protein name, species, Mascot score, percent coverage, and the number of peptide hits (and unique hits). The percent coverage was calculated by adding together the total number of amino acids residues present in all the peptide hits attributed to a particular protein, and dividing this sum by the total length of the protein in GenBank. There were five classes of defense proteins identified: chitinases; ribosome inhibiting proteins; proteinase inhibitors; proteases; and peroxidases.

We discovered 18 proteins that were not degraded upon the desiccation of the fungal remnants on the surface of the maize kernels (Table 3.2). Of the 18 proteins identified, 10 were attributed to *F. graminearum*, and four proteins were attributed to *F. oxysporum*. A putative linoleate diol synthase from *F. graminearum* had the highest Mascot score (349). Other enzymes involved in the processing of raw nutrients include glycoside hydrolase and xylanase. Additionally, we identified ubiquitous proteins, proteins involved with stress tolerance, metabolism, as well as proteins with unknown functions.

3.3 Discussion

The economic losses from all *F. graminearum* diseases are in excess of $1 billion annually through the reduction in grain yield and certain traits such as grain quality

Table 3.1: Defense proteins in maize samples.

Sample	Accession	Protein	Mascot Score	% Coverage	Peptides, Uniques
H_2O treated maize embryo	NP_001147325	subtilisin-chymotrypsin inhibitor CI-1B	139	38.6	2,2
	ACJ62806	ribosome-inactivating protein	106	12.3	3,3
	AAT39992	chitinase, class GH19	101	7.2	2,1
***Fusarium* treated maize embryo**	ACG36543	peroxidase 12	209	13.6	4,4
	NP_001140805	uncharacterized protein LOC100272880 - pepsin retropepsin-like	202	16.1	4,4
	AAT39991	chitinase, class GH19	159	14.2	4,3
	NP_001105541	chitinase chem5 precursor - contains GH18 hevamine XIP I class III domain	150	8.9	2,2
	NP_001142312	uncharacterized protein LOC100274481 - contains GH18 hevamine XIP I class III domain	147	9.9	3,3
	NP_001232812	trypsin/factor XIIA inhibitor precursor	99	47.8	3,3
	NP_001146946	cysteine proteinase inhibitor B	79	14.0	2,2
	CAN72184	hypothetical protein VITISV_042598 - retropepsin like	78	3.1	4,3
H_2O treated maize endosperm	CAA49723	protein b-32 – ribosome inhibiting protein	237	21.5	5,2
	NP_001149472	endochitinase A precursor, class GH19	120	11.8	2,2
***Fusarium* treated maize endosperm**	ACJ62815	ribosome-inactivating protein	206	20.6	4,2
	ABA34115	maize protease inhibitor	111	38.4	2,2
	NP_001142312	uncharacterized protein LOC100274481 precursor - putative xylanase inhibitor XIP-I, or class III plant chitinase	92	7.1	2,2
	NP_001142211	uncharacterized protein – contains ricin-type β-trefoil domains	77	5.8	2,2

Table 3.2: Proteome of desiccated fungi on the surface of infected maize kernels.

Accession	Protein	Species	Mascot Score	% Coverage	Peptides, Uniques
XP_382844	hypothetical protein FG02668.1 – putative linoleate diol synthase	*Gibberella zeae* PH-1	349	10.1	10,8
XP_381869	hypothetical protein FG01693.1	*Gibberella zeae* PH-1	220	5.4	5,5
XP_389762	hypothetical protein FG09586.1 - putative phosphatidylinositol/phosphatidylglycerol transfer protein	*Gibberella zeae* PH-1	194	26.7	5,5
AAR02398	glycoside hydrolase 7	*Gibberella zeae*	183	7.8	4,3
XP_388897	SODC_NEUCR superoxide dismutase	*Gibberella zeae* PH-1	171	15.8	2,1
EGU80493	hypothetical protein FOXB_08953 – putative Cu-Zn superoxide dismutase	*Fusarium oxysporum* Fo5176	137	16.9	2,1
XP_381699	hypothetical protein FG01523.1 – putative aldo-keto reductase	*Gibberella zeae* PH-1	132	17.9	4,4
XP_381422	hypothetical protein FG01246.1 – putative old yellow enzyme	*Gibberella zeae* PH-1	131	2.6	3,3
ABU48779	translation elongation factor 1-alpha	*Wickerhamomyces rabaulensis*	126	15.9	2,2
XP_001538961	ubiquitin/79aa fusion protein	*Ajellomyces capsulatus* NAm1	121	37.0	3,1
XP_391289	hypothetical protein FG11113.1 – putative bicupin, oxalate decarboxylase	*Gibberella zeae* PH-1	121	9.6	3,3
CAD92004	calmodulin	*Fusarium oxysporum*	117	32.2	3,2
EGU72061	hypothetical protein FOXB_17429	*Fusarium oxysporum* Fo5176	110	36.4	2,1
XP_003044867	predicted protein - ribosomal protein S11	*Nectria haematococca* mpVI 77-13-4	103	22.7	3,3
AAV98254	putative xylanase 2	*Gibberella zeae*	99	8.2	2,2
XP_390596	hypothetical protein FG10420.1	*Gibberella zeae* PH-1	95	21.4	3,3
EGU83490	hypothetical protein FOXB_05994 - contains stress-induced bacterial acidophilic repeat motif	*Fusarium oxysporum* Fo5176	86	33.8	2,1
XP_003007386	60S ribosomal protein L26-1	*Verticillium albo-atrum* VaMs.102	82	32.3	3,2

(Goswami and Kistler, 2004). Previous studies have investigated the proteomic profiles of both pathogen and host at various stages of infection (Geddes et al., 2008; Mohammadi et al., 2011; Shin et al., 2011; Zhou et al., 2006), but we were interested in the proteomes of both the field desiccated fungal remnants on the surface of the seed at harvest and in the tissues of the mature seeds themselves.

3.3.1 Percent representation

The sessile nature of plants forces their pathogen defense systems to be particularly robust (The Arabidopsis Genome Initiative, 2000). It has previously been shown that certain defense proteins are upregulated following pathogen infection (Walter et al., 2010). Our data confirms these findings, as there was an increase in the number of sequenced proteins associated with defense in both the infected embryo and infected endosperm tissues compared to the control samples (Figure 1). The number of identified defense proteins increased 617% in the infected embryo, and 278% in the infected endosperm. The lower percentage increase of defense proteins in the endosperm is due to the apoptotic (and therefore less responsive) nature of the endosperm (Young and Gallie, 2000).

3.3.2 Proteinase inhibitors

Proteinase inhibitors (PIs) are abundantly expressed in storage tissues; comprising up to 10% total protein content (Habib and Fazili, 2007). PIs act by forming complexes with proteolytic enzymes (Sels et al., 2008). These complexes reduce the activity of the proteolytic enzymes, and therefore are important in regulating a number of important biological functions influenced by proteolytic enzymes, including signal transmission, seed dormancy, and apoptosis (Haq et al., 2004). In addition to regulating host biochemical processes, PIs have a proposed role in plant defense. Upon infection, fungi secrete proteases that degrade cell walls to facilitate colonization, as well as digest plant proteins down to the amino acids required for fungal metabolism (Dunaevskii et al., 2005). In response to fungal infection, plant PIs are upregulated, and bind and reduce fungal protease activity (Ryan, 1990). We identified four PIs (Table 3.2). Both the infected embryo and endosperm tissues had more represented PIs compared to the water injected control samples. One PI identified in *Fusarium*-treated embryo was a trypsin/factor XIIA inhibitor (GenBank accession NP_001232812). This PI, in addition to inhibiting proteases, also demonstrates α-amylase inhibitor activity. These dual function inhibitors are particularly effective in defending against fungal pathogens, as the degradation of both proteins and starch granules is inhibited – severely limiting the sources of energy to the growing fungal hyphae.

3.3.3 Chitinases

Chitinases hydrolytically cleave the β-1,4-glycosidic bonds in both chitin and chitosan, which are important structural components of fungal cell walls (Jitonnom et al., 2011). The majority of identified plant chitinases are endochitinases that hydrolyze bonds within the chitin, chitosan, or other molecules, as opposed to exochitinases, which catalyze the cleavage of bonds at the non-reducing end of chitin polymers (Huang et al., 2012). Chitinases play an important role in plant defense against fungal pathogens. Following infection, chitinases are upregulated (van Loon et al., 2006), and hydrolyze exposed chitin polymers in fungal hyphal tips, inhibiting fungal growth (Grover, 2012). Plant chitinases are grouped into two separate families, based on features of their catalytic domains: GH18, which includes chitinase Classes III and V, and GH19, which includes chitinase Classes I, II, and IV (Santos et al., 2008). The GH18 and GH19 families likely do not share a common ancestral protein, as these families have significantly different amino acid sequences and three-dimensional structures (Ubhayasekera, 2011). Class IV chitinases (GenBank accessions AAT39991, AAT39992, NP_001149472) are distinguished by the presence of an N-terminal carbohydrate-binding domain (Ubhayasekera et al., 2009). Class IV chitinases were found in both infected and control samples.

3.3.4 Xylanase inhibitors

Xylan is composed of linked β-1,4-linked D-xylose groups that can be substituted by different residues, including L-arabinose (Juge, 2006). Xylans and their substituted derivative molecules are the most abundant hemicellulose molecule in cereals, representing up to 25% of the dry weight of maize (Pordesimo et al., 2005). Secreted fungal endoxylanases catalyze the degradation of xylans into poly- and oligosaccharides by hydrolyzing the β-1,4-linkages between xylopyranoside residues in xylans (Prade, 1996). There are two classes of fungal xylanase: GH10 and GH11 that differ in size, 3D structure and catalytic properties (Biely et al., 1997). We identified a putative xylanase in our sequencing of the *Fusarium* remnants on the seed surface (Table 3.2). This 40 kDa GH10 xylanase is characterized by broad substrate specificity and a resistance to certain xylanase inhibitor proteins (XIPs) (Goesaert et al., 2004).The two XIP-like proteins we identified (Table 3.2 - GenBank accessions NP_0011005541 and NP_001142312) have been shown to able to inhibit both GH10 and GH11 xylanase classes. Class GH10 is inhibited through XIPs binding to the xylanase substrate-binding pocket, while class GH11 is inhibited through XIPs binding to the active site (Payan et al., 2004). In addition to sharing identity with XIPs, these XIP-like proteins we identified have sequence identity to class III chitinases, though these proteins have modified active sites that lack the ability to catalyze the degradation of chitin (Payan et al., 2003). The discovery of XIPs in only the *Fusarium* treated samples agrees with previous studies that demonstrated that XIPs in maize were upregulated following *F. graminearum* infection (Mohammadi et al., 2011).

3.3.5 Ribosome inactivating proteins

Cereal ribosome inactivating proteins (RIPs) are rRNA N-glycosidases that inactivate ribosomes by catalyzing the irreversible depurination of a single adenine residue in the conserved elongation factor-binding α-sarcin loop of 28S rRNA molecules (Barbieri et al., 1993). Accordingly, protein synthesis is arrested in these depurinated ribosomes. There are three types of RIPs. The maize RIP1 protein (b-32) is a type 3 RIP, as it is expressed as inactive precursors that, following proteolytic removal of N-terminal, internal, and C-terminal domains, yield two chains, which interact non-covalently to form an active form that is 10 000 times more active than the precursor (Walsh et al., 1991). Unlike many of the other pathogen resistance gene products, b-32 is not upregulated in response to pathogen infection; it is expressed in the endosperm concurrently with the deposition of storage proteins (Balconi et al., 2010). This is reflected in Table 3.1, as the RIP proteins are found in both the control and *Fusarium* infected samples.

3.3.6 Peroxidases

Oxidative burst is the rapid and transient production of reactive oxygen species (ROS) in response to pathogen infection (Angel Torres, 2010). There are three major proposed functions of the ROS during oxidative burst: 1) generating a toxic environment that inhibits pathogen growth, 2) reinforcing cell walls to prevent pathogen spreading, and 3) activating signaling cascades to induce pathogen response genes (Passardi et al., 2005). There are several pathways that work synergistically to produce ROS during oxidative burst. From these intertwined pathways, we identified a single peroxidase in our survey. This class III peroxidase (GenBank accession ACG36543) was present in the *Fusarium* treated embryo sample. Class III peroxidases are thermostable glycoproteins with broad substrate specificity that are localized to vacuoles and cell walls (Almagro et al., 2009). The importance of class III peroxidases in response to pathogen attack was demonstrated in *Arabidopsis*, where transgenic lines with reduced transcript levels of class III peroxidases had a reduced oxidative burst and had an increased susceptibility to pathogen infections (Bindschedler et al., 2006). Our discovery of ACG36543 in only the *Fusarium* treated embryo reinforces previous findings that this particular peroxidase is upregulated in response to pathogen infection (Mohammadi et al., 2011) and may be sufficient to convey protection due to its multifunctional properties (Passardi et al., 2005).

Fungal pathogens can mitigate the effects of the ROS expressed during an infected host's oxidative burst through the action of several different proteins. These protein families include peroxidases, catalases and superoxide dismutases (Mayer et al., 2001). We identified two copper-zinc (Cu-Zn) superoxide dismutases (SOD) amongst the *Fusarium* remnants (Table 3.2). There are four classes of SOD, which can be differentiated by their metal co-factors: 1) Cu-Zn, 2) Mn, 3) Fe, and 4) Ni (Mayer et al., 2001). All classes of SOD catalyze the dismutation of $^{\bullet}O_2$ and $^{\bullet}OH$ to O_2 and H_2O_2 (Hwang et al., 2002). Previously, *F. graminearum* Cu-Zn SOD proteins have been identified in infected *Triticum aestivum* (Paper et al., 2007; Zhou et al., 2006), indicating that this protein is highly expressed by *F. graminearum* during infection.

3.3.7 *Fusarium* on the surface

The proteins identified in the fungal remnants on the surface of the maize kernels could be divided into four main categories depending on their biological functions: 1) Proteins responsible for the metabolism of plant molecules (linoleate diol synthase, glycoside hydrolase), 2) ROS resistance (superoxide dismutase), 3) constitutive expressed proteins (elongation factors, ribosomal proteins, calmodulin, and others), and 4) hypothetical proteins with no assigned function. As we identified individual 18 proteins in our desiccated *Fusarium* samples, these proteins may still be present as they are either very highly expressed or resistant to degradation.

3.3.8 *Fusarium* metabolism proteins

Up to 70% of the total oil content from maize seeds is linoleic acid (Jellum, 1970). This unsaturated n-6 fatty acid can be metabolized by *Fusarium* by the putative linoleate diol synthase (LDS) (Table 3.2) in the following reaction: linoleic acid is dioxygenated to (8R)-hydroperoxylinoleate and subsequently isomerized to (7S,8S)-dihydroxylinoleate (Hornsten et al., 1999). Proteins that share sequence identity with LDS have been found in several filamentous fungi (Soanes et al., 2008) and have been hypothesized to be involved with the mobilization of storage lipids (Huber et al., 2002). Gene deletion studies in the rice-blast disease causing *Magnaporthe oryzae* demonstrated that LDS is not required for successful infection or sporulation, possibly due to the presence of redundant linoleic acid metabolizing pathways (Jerneren et al., 2010).

The degradation of different glycosides is performed by the family of glycoside hydrolases that catalyze the hydrolysis of glycosidic linkages, freeing sugar moieties. The glycoside cellulose is a chain of β-(1→4) linked D-glucose units. There are two classes of fungal glycoside hydrolases that degrade cellulose: 1) endoglucanases that catalyze the cleavage of cellulose polymers at random locations, and 2) cellobiohydrolases that hydrolyze the non-reducing ends of cellulose chains (Munoz et al., 2001). The glycoside hydrolase 7 we identified is a cellobiohydrolase. As the most predominant glycoside in maize stalks (approximately 50% of dry weight (Appenzeller et al., 2004)) is cellulose, glycoside hydrolases are necessary for *Fusarium* to degrade plant cell walls to facilitate colonization and nutrient uptake (Yoshizawa et al., 2002).

3.3.9 *Fusarium* stress response

Old Yellow Enzyme (OYE) is a flavin-dependent NADPH-oxidoreductase that was first described in 1932 by Warburg & Christian in 1932 from brewers' bottom yeast. Proteins with sequence similarity to OYE have been identified in a broad phylogenetic range of organisms, including bacteria, fungi, plants, and mammals (Williams and Bruce, 2002). While the physiological role of OYE is, as of yet, unknown, a number of hypothesized functions has been proposed, including oxidative stress response (Lee et al., 1999). OYE is upregulated in the presence of hydrogen peroxide (Fitzpatrick et al., 2003) and can be oxidized by many different compounds, including ketones and aldehydes (Trotter et al., 2006). These data together with our discovery of OYE in the fungal remnants indicates that it may play a protective role during plant colonization.

Plant late embryogenesis proteins (LEA) are a diverse group of hydrophilic proteins associated with abiotic stress tolerance - most notably desiccation and cold stress (Tolleter et al., 2010). There are three main groups of LEA proteins in plants that are distinguishable by the presence of sequence motifs (Tunnacliffe and Wise, 2007). The majority of proteins within each group have only minimal secondary structure when in solution, though they have increased folding when in a dry state (Tolleter et al., 2007). It has been shown that LEA proteins accumulate during the final stages of embryo maturation (Galau et al., 1986). This accumulation of LEA proteins in mature seeds is reflected in this study, because in each of the maize samples investigated the percent representation was greater than 20% (Figure 1). While the function of LEA proteins in stress tolerance is unknown, it has been hypothesized that they have several different functions, including stabilizing membranes (Steponkus et al., 1998) or desiccation-sensitive proteins (Goyal et al., 2005), preventing cellular water loss (Garay-Arroyo et al., 2000), or by binding ions or ROS (Tunnacliffe and Wise, 2007).

In addition to plants, LEA proteins have been discovered in prokaryotes, protists, invertebrates, and certain invertebrates (Tolleter et al., 2007). All fungal LEA-like proteins are group 3 LEAs, which are characterized by the presence of multiple repeats of an 11 amino acid residue motif (Tunnacliffe and Wise, 2007). We identified a single LEA-like protein in the fungal remnants (Table 3.2). This protein also shares 96% sequence identity with a conidiation-specific protein. Conidiation-specific proteins are expressed during conidia (spore) formation, which is triggered by low nutrient or water levels (White and Yanofsky, 1993). It is therefore likely that this LEA-like protein was expressed in the *Fusarium* examined during seed maturation and desiccation in order to form the spores necessary for the biological dispersal of the parental *Fusarium* genomic sequence. As other fungal species propagate in a similar manner, the proteins on the surface of the kernels sourced to fungal species other than *F. graminearum* (Table 3.2) may be present due to the damage from the inoculation of the maize kernels. This damage potentially allowed for the limited colonization by opportunistic fungi that would otherwise not have been able to colonize maize.

3.4 Conclusions

The approximately 250 maize proteins identified were classified into functional categories, including: late embryogenesis abundant, seed storage, metabolism, and defense, with an increased representation of defense proteins in both the embryo and endosperm tissues of infected samples. Most of the proteome of the fungal remnants was categorized as being involved in the metabolism of plant-sourced molecules, or in stress response. Taken together, these data should help dissect the interactive system between host and pathogen during the important time period of grain maturation and harvest.

Understanding the interplay of pathogen and host is important in developing pathogen resistant breeding lines. Greater knowledge of host defense systems will allow for the fine-tuning of the polygenic disease resistance present in maize. Finally, the proteins present in the fungal remnants may be targets for fungicides or for rapid biological screening of grain shipments for presence of pathogen infection.

3.5 Methods

F. graminearum susceptible *Zea mays* inbred line CL30 was planted on the Central Experimental Farm in Ottawa, ON, Canada for the 2010 growing season. Conidial suspensions of *F. graminearum* strain DAOM 180378, obtained from the Canadian Collection of Fungal Cultures, Ottawa, ON were prepared to 5×10^5 conidia/mL as described in (Reid et al., 1996). Maize kernels were inoculated with either the conidial suspensions or sterile water controls using a four-pin automatic kernel inoculator as per Reid, et al., (, 1997). This inoculation was timed to the blister to early milk stages of kernel development; approximately 10-15 days post-silk emergence. Three ears of both the fungal and water treated maize were harvested at maturity.

Kernels were harvested and suspended in protein extraction buffer (10mM Tris-HCl pH 7.8, 0.5 mM EDTA, 10 mM KCl, 1mM PMSF). Kernels were agitated for 30 seconds, allowed to settle. The supernatant was decanted into a fresh tube, and sonicated on ice at 20 amplitude microns 20 times in three second bursts. Cellular debris was pelleted by centrifugation at 14 000 *g* for one min at 4 C, and supernatant transferred to a fresh tube and stored at -80 C. The kernels were surface sterilized in a 10% solution of sodium hypochlorite for 5 min, washed 5 times in ddH$_2$O and dried in a laminar flowhood. Kernels were incubated overnight in 0.3% (w/v) sodium metabisulphite and 85% (v/v) lactic acid at 37 C. Embryo and endosperm tissues were manually dissected from the pericarp using a razor blade. Isolated tissues were ground in a mortar and pestle and suspended in protein extraction buffer. The solution was filtered through a 63 µ mesh and centrifuged at 14 000 *g* for 10 min at 4 C. The pellet was washed two times with protein extraction buffer, two times with ice cold 95% ethanol, and two times with ice cold acetone. The pellets were allowed to dry in a laminar flowhood. Protein extraction buffer was added to the pellets and allowed to incubate for 30 min at room temperature. The solution was centrifuged at 14 000 *g* for ten min, and the supernatant was collected. Along with the protein samples previously stored at -80 C, the samples were supplemented with four volumes of ice cold acetone and stored at -20 C overnight to precipitate proteins. The samples were centrifuged at 14 000 *g* for twenty min and washed twice with ice cold acetone and dried in a laminar flowhood.

Pellets were resuspended in 100 µL ddH$_2$O. Protein concentration was determined using the Bradford assay (Bradford, 1976). Two micrograms of each protein sample was subjected to mass spectrometric analysis at the Ottawa Institute of Systems Biology, OISB. Protein samples were digested with trypsin, and digested peptides were separated by liquid chromatography on an Agilent 1100 Series HPLC System (Agilent

Technologies, Palo Alto, CA). Separated peptides were ionized by electrospray to a QSTAR Pulsar quadrupole-TOF mass spectrometer (ABI/MDS Sciex, Concord, ON) as described in (Vasilescu et al., 2005). Mascot software (version 2.3) (Matrixscience Ltd.) was used to interrogate the non-repetitive NCBI protein database (12/20/2011; 16 392 747 sequences; 5 641 810 382 residues) with the taxonomy filter set to either 'other fungi' (3 787 334 sequences) or 'other green plants' (3 614 282 sequences). Carbamidomethyl (C) and oxidation (M) were set to fixed variable modifications, respectively. One missed cleavage was permitted. Tolerances were set for both peptide (±100 ppm) and MS/MS mass tolerances (0.2 Da). Decoy searches were performed, and false positive rates were approximately 1%.

Molecular function was assigned to proteins using the Kyoto Encyclopedia of Genes and Genomes (KEGG) at http://www.genome.jp/kegg/kegg2.html (last viewed 7/09/2012) (Kanehisa and Goto, 2000; Kanehisa et al., 2006; Kanehisa et al., 2010). Peptides corresponding to unknown or hypothetical proteins in the NCBI database were classified based on sequence similarity by BLAST searches (Altschul et al., 1990).

Chapter 4

The starch granule associated proteomes of commercially purified starch reference materials from rice and maize

Koziol AG, Marquez BK, Huebsch MP, Smith JC, Altosaar, I

Published in: *Journal of Proteomics* 75: 993-1003 (2012)

Contribution of authors:

AGK contributed to the study design, performed experiments except LC-MS/MS, analyzed data, contributed all figures, wrote the manuscript, and participated in editing the manuscript. BKM aided in experiments, analyzing data, and editing. MPH performed LC-MS/MS. JS participated in analyzing data and editing. IA participated in study design, data analysis, and editing.

4.1 Abstract

Commercially available reference materials are integral components of many experimental protocols, as it is critical to compare one's results to those derived from well-characterized standards. Most reference materials are well defined, with all their components being catalogued. However, certain reference materials, such as commercially prepared starch samples, can have undefined components, potentially limiting their usefulness as standards. The proteome of commercially prepared starch has not been documented, and to that end, we initiated a mass spectrometry-based survey of the proteins associated with starch granules in commercially prepared rice and maize starch samples. We performed direct trypsin treatments of starch samples and sequenced both the water-soluble peptides liberated into the aqueous supernatant and the peptides released from the starch granule surface by isopropanol solvent washing. We discovered that the majority of proteins, in both rice and maize samples, were involved in either carbohydrate metabolism or storage. We also documented proteins that are markers for seed maturity and for starch mobilization.

4.2 Introduction

Starch is a versatile and renewable commodity that is exploited as food and biofuel, as well as in industry during the production of chemicals, paper, binders, and adhesives (Roper, 2002). Worldwide, almost 300 billion tons of starch is manufactured per year, with 90% originating from cereal crops (Halford et al., 2011; Tetlow, 2011). These crops, which include rice and maize, produce seeds that contain up to 70% of their dry weight as starch (Dale and Housley, 1986). Having such concentrated reserves of starch so readily available, it is therefore not surprising that as an extractable component, rice starch was used as far back as ancient Egypt for the manufacture of papyrus (Bostock and Riley, 1855). Cereal starch is synthesized in amyloplasts, and stored in the endosperm as semi-crystalline granules consisting of amylose and amylopectin (Zeeman et al., 2010). Both these polymers contain α-(1$\rightarrow$4) glycosidic linkages, however, amylopectin which makes up 75% of the mass of most starch granules, contains approximately 5-6% of randomly distributed α-(1$\rightarrow$6) glycosidic linkages (Grimaud et al., 2008). These linkages introduce branch points in the amylopectin chains that allow for the clustering of chains into tightly packed arrays, and, consequently, the semi-crystalline nature of starch (Perez and Bertoft, 2010).

The origins of starch synthesis can be traced to the common ancestor of the Archaeplastida (Patron and Keeling, 2005). Following an endosymbiotic event between a heterotrophic eukaryotic host and a cyanobacterium, which gave rise to the Archaeplastida, the genomic reorganization in the new, photosynthetic eukaryotic organism merged polyglucan synthesis pathways of bacterial and eukaryotic origins (Deschamps et al., 2008). Cytosolic starch synthesis pathways evolved from these reorganization events, and it was with the evolution of the light-harvesting complexes, which are responsible for light capture and photoprotection (Koziol et al., 2007) that starch synthesis was relocated to plastids (Tetlow, 2011). It has been proposed that this relocation was prompted by the requirements for ATP, generated via starch degradation, to overcome protoporphyrin IX-induced oxidative stress (Deschamps et al., 2008).

In addition to starch, the endosperm of cereal seeds contains proteins, lipids, carbohydrates, and other organic and inorganic compounds (Lopes and Larkins, 1993). Most of the proteins found in seeds are storage proteins, which act as sinks of amino acids for germinating embryos (Shewry and Halford, 2002). However, structural and metabolic proteins, and starch granule associated proteins (SGAPs) are present as well (Lowy et al., 1981). Many of these proteins are enzymes involved in starch biosynthesis, and include adenosine 5′ diphosphate-glucose pyrophosphorylase, granule-bound starch synthase (GBSS), soluble starch synthase, starch branching enzyme, starch debranching enzyme, and plastidial starch phosphorylase (Jeon et al., 2010). Other proteins are involved in defense, transport, signaling, protein folding, and stress response (Balmer et al., 2006; Wall et al., 2010).

Despite SGAP being a term in the biochemical lexicon since 1986, SGAPs have only been investigated on a case-by-case approach (Baldwin, 2001; Greenwell and Schofield, 1986). Regardless of this lack of scientific precision, commercial starch samples have nevertheless still been frequently used as reference materials in experiments. The rice and maize starch samples we examined here have been used in a broad range of experiments that, for example, have studied starch granule structure using light and electron microscopy (Baldwin et al., 1994), the enzymatic digestibility of starch granules by amylases and glucosidases (Erban et al., 2009), the production of ceramic materials with differing levels of porosity through the use of starch additives (Diaz and Hampshire, 2004), the pasting properties of floridean starch in algae (Yu et al., 2002), and the use of sonication treatments during starch separation (Zhang et al., 2005), among others.

The certificates of analysis for these commercially available starch samples give basic measurements of the non-starch components of the samples, such as pH and percentages of moisture, protein, and lipid, but in this current age of systems biology, it is important to be able to catalogue all the detectable components of a sample. It has previously been stated that the presence of proteins and lipids within purified starch samples have an effect on the rate of granule hydrolysis by amylases and glucoamylases (Williamson et al., 1992). Therefore, it is possible that the protein component of commercially purified starch samples could affect the outcome of other experiments in which the starches will serve as defined baseline substrates. It is for this reason that we initiated a tandem mass-spectrometry-based survey of commercially available starch reference materials to catalogue their trace proteomes. We have previously investigated the SGAPs in wheat (Wall et al., 2010; Wall et al., 2010) using strict washing protocols at a laboratory scale. Here, we have investigated the proteomes of rice and maize starch granules following commercial starch purification.

4.3 Experimental procedures

4.3.1 Sampling

Rice and maize starch samples were purchased from Sigma-Aldrich (St. Louis, MO). In order to confirm random sampling, 1 kg rice starch (Cat. No. S7260, Lot No. 107K0052) and 2 kg maize starch (Cat. No. S4126, Lot No. 015K0144) reagent bottles were gently agitated, and four separate 50 mg aliquots from each bottle were removed, with gentle agitation between each sampling step.

4.3.2 Starch granule preparation

To disperse the granules and allow for granule swelling, 350 μL of ddH_2O was added to 200 mg of starch. The samples were incubated for 30 min at room temperature with gentle agitation. To prepare samples for trypsin treatment, ammonium bicarbonate was added to the samples to a final concentration of 100 mM. Trypsin (T6567, Sigma-Aldrich) was suspended in 1 mM HCl to a concentration of 1 μg/μL according to the manufacturer's directions. Starch samples were incubated with 5 μg of trypsin overnight at 37 °C with gentle mixing.

4.3.3 Peptide preparation

The peptides associated with the starch granule surface were collected according to our modified protocol (Wall et al., 2010). Briefly, the trypsin-treated granules were centrifuged at 18,000 x *g* for 1 min. This aqueous supernatant was transferred to a fresh tube. The pellet, which consisted of trypsin-treated starch granules and peptides, was washed five times with a ten-fold excess of ddH_2O to remove any residual water-soluble proteins. Following water washing, proteins remaining on the starch granule surface were extracted by adding 350 µL 50% (v/v) isopropanol, 50 mM NaCl and gently mixing for 45 min at room temperature. The samples were centrifuged at 18,000 x *g* for 1 min, and the supernatant was collected. The peptides from both this isopropanol fraction and the reserved aqueous supernatant were dried in a Speed Vac (Speed Vac Concentrator model number SVC 100H; Savant Instruments, Inc. Hicksville, NY). Peptide pellets were resuspended in 40 µL of ddH_2O, purified using ZipTips with C18 resin (Millipore, Bedford, MA) to remove salt and residual starch, and dried again in a Speed Vac. Peptides were resuspended in 40 µL of 0.1% formic acid.

4.3.4 Chromatography and mass spectrometry

To determine the number and nature of proteins in the samples, peptides were loaded into 7 cm by 200 µm inner diameter trap columns, fritted, and packed in-house with 5 cm of 5 µm Magic C18AQ reversed-phase packing material (Michrom Bioresources, Auburn, CA) using custom pressure vessels constructed in-house and washed with 100 uL of 0.1% formic acid. The column was then connected to a 75 um inner diameter Picofrit tip (New Objective, Woburn, MA) also packed with 5 cm of Magic C18AQ reversed phase material. Peptides were eluted at a flow rate of approximately 250 nL/min using an HP 1090 HPLC (Agilent Technologies, Santa Clara, CA) and the following gradient of acetonitrile containing 0.1% formic acid over 90 min: 0 min 2%, 3 min 5%, 8 min 10%, 70 min 35%, 75 min 80%, 83.5 min 100% (wash step), 84 min 2%, 90 min 2% (re-equilibration steps).

Eluted peptides were ionized by nanoelectrospray ionization, and analyzed using a QSTAR XL QqTOF mass spectrometer (AB Sciex, Foster City, CA) operating in information-dependent acquisition mode. Mass analysis included a 1 s survey scan followed by two 2 s and two 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.

4.3.5 Protein identification

Tandem mass spectrometric data were interpreted using Mascot software (version 2.3) (Matrixscience Ltd., Boston, MA) using a custom database consisting of all proteins from the 20100704 comprehensive National Center for Biotechnology Information non-redundant (NCBInr) protein database (11 365 658 sequences; 3 876 535 693 residues). Mass tolerances were set to ±50 ppm and ±0.2 Da for the peptide and fragment ion spectra, respectively; one missed cleavage was 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. An automated decoy search was performed, and false positive rates were approximately 1%. 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 $> 2^{+}$). Only proteins with two or more peptide-based identifications with Mascot scores ≥ 55

were included to minimize the chance of matches occurring due to random events. Peptides corresponding to unannotated sequences in the database were classified based upon sequence homology by BLAST searches (Altschul et al., 1990). Molecular function of the proteins matched from the database was inferred using the Kyoto Encyclopedia of Genes and Genomes (KEGG) at http://www.genome.jp/kegg/kegg2.html (Kanehisa and Goto, 2000; Kanehisa et al., 2006; Kanehisa et al., 2010). For proteins that corresponded to multiple pathways, the most suitable pathway was chosen based upon the context of the protein being found on the starch granule surface.

4.4 Results

Commercial rice and maize starch samples were treated with trypsin. Peptides generated from the proteolytic digestion of proteins associated with the starch granule surface (SGS) were isolated with aqueous or isopropanol/salt solvents. The peptides were sequenced by tandem mass spectrometry, and MS/MS data were interpreted using Mascot software. The molecular functions of proteins identified by Mascot were determined using KEGG.

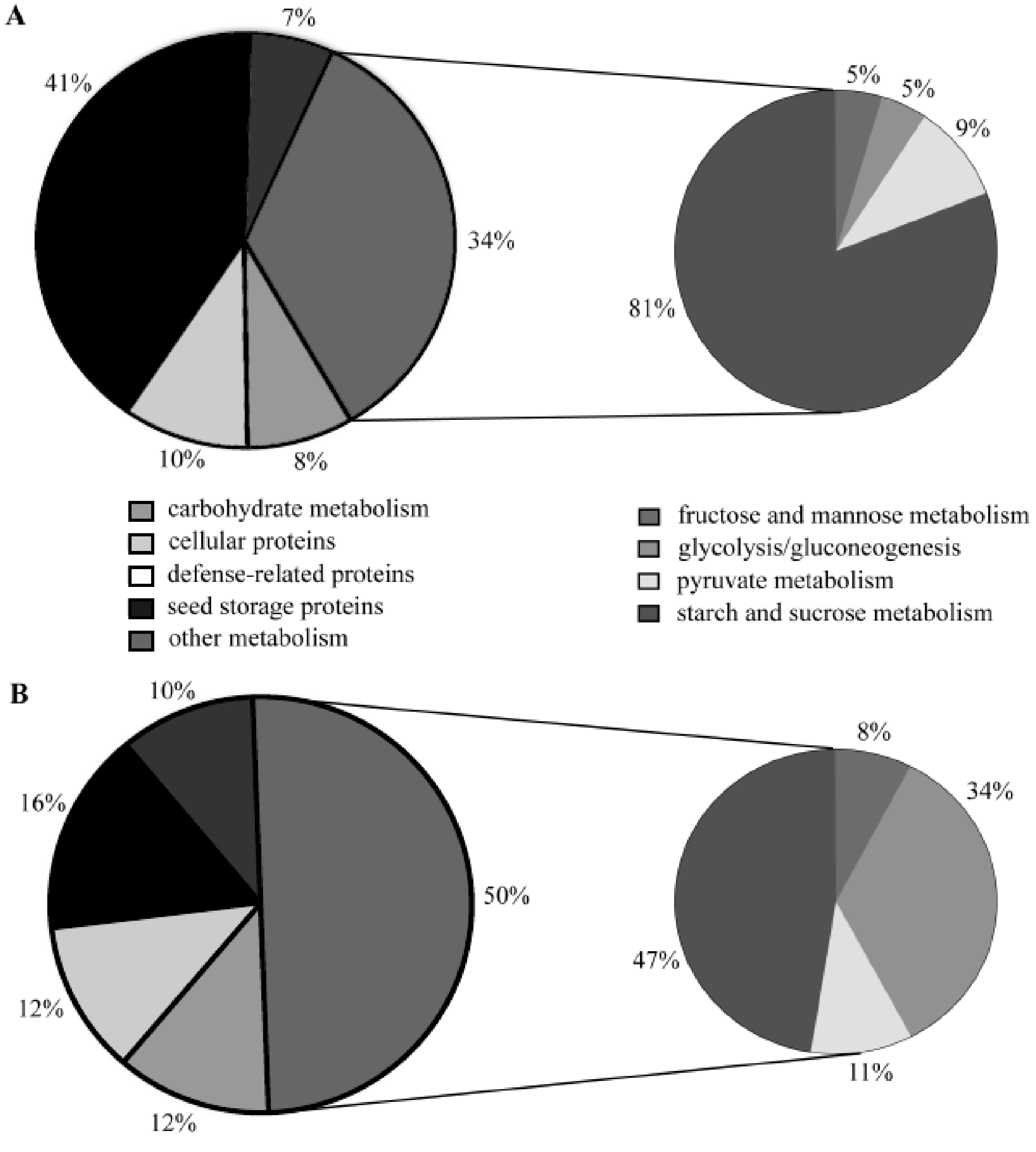
4.4.1 Starch granule associated proteins in rice

Using Mascot, peptides from 107 proteins were identified in the rice starch sample. We determined that 90 of these proteins were plant-sourced with 60 proteins identified in the aqueous fraction and 30 identified in the isopropanol fraction. The remaining 17 non-plant proteins were from a phylogenetically diverse range of prokaryotes and eukaryotes and may be present in the starch samples as remnants from pathogen attacks (Wall et al.,

2010) or contaminants from the commercial purification of the starch granules. The percentage representation of the molecular functions of the proteins determined using KEGG is shown in Figure 4.1A and are as follows: seed storage proteins (41%), carbohydrate metabolism (34%), cellular proteins (8%), defence proteins (10%), and other metabolism proteins (7%). The carbohydrate metabolism-related proteins can be subdivided into starch and sucrose metabolism (81%), glycolysis/gluconeogenesis (11%), pyruvate metabolism (7%), fructose and mannose metabolism (4%), and glyoxylate and dicarboxylate metabolism (4%).

We identified 21 proteins classified as carbohydrate metabolism proteins in the rice starch samples. Of these 21 proteins, 17 were found in the aqueous fraction and four were found in the isopropanol fraction. These proteins are presented in Table 4.1, which includes GenBank accession number, protein name, Mascot score, number of peptide hits (and unique peptide hits), percent protein coverage of the peptide hits, and the carbohydrate metabolism subcategories. Percent protein coverage of the peptide hits was calculated by dividing the number of amino acid residues spanned by the assigned peptides by the total number of amino acid residues of the full-length protein. Proteins generated from peptides in the isopropanol fraction are presented in **bold** font. Within carbohydrate metabolism subcategories, we found 17 proteins related to starch and sucrose metabolism, including sucrose synthase, granule-bound starch synthase, branching enzyme, ADP-glucose pyrophosphorylase, pullulanase, α-glucosidase, and α-1,4-glucan phosphorylase. Other carbohydrate metabolism-related proteins identified included sorbitol

Figure 4.1: Functional classification of identified proteins.
Panel A) rice and panel B) maize. The functional categories were determined using KEGG. In the pie charts on the left include carbohydrate metabolism, cellular proteins, defence-related proteins, seed storage proteins, and other metabolism proteins. The carbohydrate metabolism category is subdivided into fructose and mannose metabolism, glycolysis/gluconeogenesis, pyruvate metabolism, and starch and sucrose metabolism in the pie charts on the right.



A
7%
41%
34%
10%
8%
5%
5%
9%
81%
carbohydrate metabolism
cellular proteins
defense-related proteins
seed storage proteins
other metabolism
fructose and mannose metabolism
glycolysis/gluconeogenesis
pyruvate metabolism
starch and sucrose metabolism
B
10%
16%
50%
12%
12%
8%
34%
47%
11%

Table 4.1: Proteins involved in carbohydrate metabolism from *Oryza sativa* starch.

Accession	Protein	Mascot Score	Number of Peptides[a]	Coverage (%)[b]
Fructose and Mannose Metabolism				
NP_001062412	Os08g0545200:sorbitol dehydrogenase	147	3(3)	12.2
Glycolysis/Gluconeogenesis				
NP_001053139	Os04g0486600:G3P dehydrogenase	142	4(3)	11.0
Pyruvate Metabolism				
NP_001055507	Os05g0405000:orthophosphate dikinase	620	15(11)	18.0
BAA22419[c]	**orthophosphate dikinase**	**90**	**2(2)**	**2.3**
Starch and Sucrose Metabolism				
Q42968	granule-bound starch synthase 1	1047	18(2)	38.3
AAF72561	granule-bound starch synthase	976	17(1)	38.3
ACU82450	granule-bound starch synthase	965	17(1)	38.3
ADI48504	granule-bound starch synthase	802	14(2)	30.9
Q42968	**granule-bound starch synthase 1, chloroplastic/amyloplastic**	**705**	**15(3)**	**36.0**
BAA01584	branching enzyme	471	10(6)	13.4
NP_001047009	Os02g0528200: rice starch branching enzyme	355	8(5)	10.2
EEC81171	hypothetical protein OsI_24144: alpha-glucosidase	346	8(8)	9.8
ACJ71344	ADP-glucose pyrophosphorylase large subunit	320	8(6)	15.3
AAA33890	ADP-glucose pyrophosphorylase 51 kD subunit	298	7(7)	13.9
AAF85966	sucrose synthase-2	246	7(1)	10.0
EEC76742	hypothetical protein OsI_14800: pullulanase	240	6(6)	6.3
NP_001051330	Os03g0758100: plastidic alpha 1,4-glucan phosphorylase	220	6(4)	12.5
EAZ04713	hypothetical protein OsI_26874: sucrose synthase	212	6(1)	8.5
AAD50279	**branching enzyme**	**126**	**4(1)**	**4.0**
AAA33891	**ADP-glucose pyrophosphorylase**	**124**	**2(2)**	**5.4**
ABD57308	UDP-glucose pyrophosphorylase	122	2(2)	5.1

[a] Number of total peptides (number of unique peptides). [b] Percent protein sequence coverage by the total peptides. [c] Entries **in bold** were sequenced from the isopropanol extract.

dehydrogenase, glyceraldehyde phosphate, fructose-bisphosphate aldolase, and orthophosphate dikinase.

There were five proteins identified in rice starch with a broad range of functions. These proteins were termed cellular proteins, of which, three were in the aqueous fraction and two were in the isopropanol fraction (Table 4.2). There were four proteins classified as folding, sorting and degradation proteins, including the endosperm lumenal binding protein, which was found in both aqueous and isopropanol fractions, heat shock protein 70, and protein disulfide-isomerase. There was one protein (elongation factor-1 alpha) associated with translation.

The most abundant class of proteins in rice starch was the seed storage proteins. There were 25 seed storage proteins identified in the rice starch samples (Table 4.3). Of these, there were 21 glutelin, one globulin, and three prolamin proteins. There were 14 seed storage proteins found in the aqueous fraction and 11 seed storage proteins in the isopropanol fraction.

4.4.2 Starch granule associated proteins in maize

In the maize starch samples, peptides from 96 proteins were identified in the aqueous fraction and 41 proteins were identified in the isopropanol fraction, yielding 137 proteins in total. Of these 137 proteins, 109 were identified as plant-sourced, with the remaining

Table 4.2: Proteins involved in cellular processes from *Oryza sativa* starch.

Accession	Protein	Mascot Score	Number of Peptides[a]	Coverage (%)[b]	Molecular Functions
AAB63469	endosperm lumenal binding protein	707	12(7)	21.3	Folding, sorting and degradation
NP_001066486	Os12g0244100: heat shock 70 protein	178	4(2)	7.2	Folding, sorting and degradation
AAX85991	protein disulfide isomerase	142	4(4)	8.2	Folding, sorting and degradation
CAA34756[c]	**elongation factor-1 alpha**	**104**	**2(1)**	**5.0**	**Translation**
AAA62325	**endosperm lumenal binding protein**	**97**	**2(2)**	**3.8**	**Folding, sorting and degradation**

[a] Number of total peptides (number of unique peptides). [b] Percent protein sequence coverage by the total peptides. [c] Entries **in bold** were sequenced from the isopropanol extract.

Table 4.3: Seed storage proteins from *Oryza sativa* starch.

Accession	Protein	Mascot Score	Number of Peptides[a]	Coverage (%)[b]
EEC73106	hypothetical protein Osl_07091:glutelin	689	15(14)	42.3
NP_001046512	Os02g0268100:glutelin	665	12(7)	31.8
EEC72839[c]	**hypothetical protein OsI_06572:glutelin**	**625**	**13(1)**	**35.0**
CAA33838	**unnamed protein product:glutelin**	**567**	**13(1)**	**33.5**
EAY75915	hypothetical protein Osl_03835:glutelin	565	12(4)	32.9
EEC72839	hypothetical protein OsI_06572:glutelin	556	12(7)	28.1
NP_001046512	**Os02g0268100:glutelin**	**554**	**11(7)**	**24.6**
EEC73106	**hypothetical protein OsI_07091:glutelin**	**519**	**13(12)**	**38.7**
BAA00462	pre-proglutelin	479	11(0)	26.7
EEC75527	hypothetical protein OsI_12139:glutelin	470	10(1)	19.0
EAY85201	hypothetical protein OsI_06564:glutelin	469	9(8)	19.0
EEC75527	**hypothetical protein OsI_12139:glutelin**	**468**	**12(9)**	**24.2**
1311273A	hypothetical protein OsI_0656:glutelin	463	11(0)	26.7
CAA38211	glutelin	457	10(1)	25.4
BAA00462	**prepro-glutelin**	**413**	**12(0)**	**33.1**
1311273A	**glutelin**	**411**	**12(1)**	**33.1**
AAO22140	glutelin precursor	382	7(5)	15.8
EAY85201	**hypothetical protein OsI_06564:glutelin**	**339**	**9(2)**	**20.2**
EAY75915	**hypothetical protein OsI_03835:glutelin**	**337**	**10(3)**	**28.6**
AAO22140	**glutelin precursor**	**292**	**6(4)**	**15.4**
CAA29507	glutelin	306	7(1)	39.4
BAA11129	prolamin	291	4(4)	49.0
BAA11129	**prolamin**	**278**	**5(5)**	**53.0**
CAA52764	11S globulin	138	3(1)	6.3
AAV43828	putative prolamin 7	123	2(2)	18.9

[a] Number of total peptides (number of unique peptides). [b] Percent protein sequence coverage by the total peptides. [c] Entries **in bold** were sequenced from the isopropanol extract.

28 proteins being attributed to a phylogenetically diverse range of organisms. The molecular functions of the plant-sourced proteins were determined using KEGG, and the percentage representation of the major classifications is shown in Figure 4.1B. The major classifications include: carbohydrate metabolism-related proteins (50%), seed storage proteins (16%), cellular proteins (12%), defence-related proteins (12%), and other metabolism-related proteins (10%). Due to the high percentage representation of carbohydrate metabolism-related proteins, this category was further divided into: starch and sucrose metabolism (47%), glycolysis/gluconeogenesis (34%), pyruvate metabolism (11%), and fructose and mannose metabolism (8%).

In contrast to the 34% in the rice starch samples, we identified 50% of the proteins in the maize starch samples as being involved in carbohydrate metabolism. Maize starch contained 18 proteins associated with starch and sucrose metabolism, including sucrose synthase, granule-bound starch synthase, branching enzyme, ADP-glucose pyrophosphorylase, UTP-glucose-1-phosphate uridylyltransferase, and glucose-1-phosphate adenylyltransferase large subunit 1 (Table 4.4). While the maize starch samples had only one protein, with three isoforms, involved in fructose and mannose metabolism (sorbitol dehydrogenase), there were 13 proteins corresponding to glycolysis/gluconeogenesis, including fructose-bisphosphate aldolase, phosphoglucomutase, triosephosphate isomerase, glyceraldehyde-3-phosphate dehydrogenase, 3-phosphoglycerate kinase, and brittle 1. Additionally, there were four proteins found in the maize starch samples involved with pyruvate metabolism. These

Table 4.4: Proteins involved in carbohydrate metabolism from *Zea mays* starch.

Accession	Protein	Mascot Score	Number of Peptides[a]	Coverage (%)[b]
Fructose and Mannose Metabolism				
ABA70761	sorbitol dehydrogenase	164	3(1)	8.5
NP_001149440	sorbitol dehydrogenase homolog 1	157	3(1)	8.5
NP_001149440[c]	**sorbitol dehydrogenase homolog 1**	**145**	**4(1)**	**8.7**
Glycolysis/Gluconeogenesis				
NP_001105336	fructose-bisphosphate aldolase, cytoplasmic isozyme	475	10(7)	32.4
ACG24648	triosephosphate isomerase, cytosolic	305	5(1)	21.7
ACG31393	triosephosphate isomerase, cytosolic	221	4(1)	19.2
ACF78275	unknown:putative brittle-1 protein	210	5(5)	12.1
P48495	triosephosphate isomerase	195	4(0)	17.7
CAB39974	**glyceraldehyde-3-phosphate dehydrogenase**	**188**	**4(3)**	**12.2**
ACU23111	unknown: putative fructose-1,6-bisphosphate aldolase	175	4(1)	12.8
NP_001105951	cytosolic glyceraldehyde-3-phosphate dehydrogenase GAPC4	167	4(3)	16.6
AAO32643	cytosolic 3-phosphoglycerate kinase	136	3(2)	14.0
CAN70587	hypothetical protein: putative triosephosphate isomerase	111	3(2)	12.6
NP_001105889	**brittle-1, chloroplastic/amyloplastic precursor**	**90**	**3(3)**	**7.1**
CAA27681	alcohol dehydrogenase 1	74	2(1)	5.0
NP_001105405	phosphoglucomutase, cytoplasmic 2	59	2(2)	3.1
Pyruvate Metabolism				
AAP34174	C4-specific pyruvate orthophosphate dikinase	272	6(2)	8.6
AAV58858	pyruvate orthophosphate dikinase	224	5(1)	5.1
NP_001146891	2-isopropylmalate synthase B	84	2(2)	3.0
ACG36184	malate dehydrogenase	79	2(2)	6.8
Starch and Sucrose Metabolism				
ACG43170	sucrose synthase 1	784	18(7)	24.7
NP_001105001	**granule-bound starch synthase 1, chloroplastic/amyloplastic**	**592**	**13(1)**	**28.8**
AAQ06291	granule-bound starch synthase precursor	558	11(1)	23.5
AAQ06291	**granule-bound starch synthase precursor**	**500**	**12(1)**	**27.1**
NP_001105001	granule-bound starch synthase 1,	550	11(1)	24.8
NP_001105411	**sucrose synthase 1**	**405**	**11(3)**	**14.1**
NP_001130742	LOC100191846: UTP–glucose-1-phosphate uridylyltransferase	332	7(2)	15.4
ABK80479	putative granule bound starch synthase	330	6(0)	12.8
AAB33385	branching enzyme II	258	5(5)	7.1
NP_001121104	glucose-1-phosphate adenylyltransferase large subunit 1	257	7(4)	12.3
AAZ82467	ADP-glucose pyrophosphorylase small subunit	181	3(3)	6.0
AAD02955	granule-bound starch synthase	165	4(1)	16.9
ACA25696	granule-bound starch synthase	160	3(2)	30.9
AAC33764	**starch branching enzyme IIb**	**157**	**5(5)**	**7.1**
ACA25696	**granule-bound starch synthase**	**147**	**3(2)**	**30.1**
ACA25661	granule-bound starch synthase	131	3(3)	13.7
AAY42161	ADP-glucose pyrophosphorylase endosperm large subunit	107	4(1)	13.0
ACA25679	granule-bound starch synthase	77	2(2)	23.8

[a] Number of total peptides (number of unique peptides). [b] Percent protein sequence coverage by the total peptides. [c] **Bold** entries were found in the isopropanol extract.

included pyruvate orthophosphate dikinase, 2-isopropylmalate synthase B, and malate dehydrogenase.

We found nine cellular proteins in the maize starch sample (Table 4.5). There were three proteins identified that were involved with folding, sorting and degradation. These proteins included heat shock protein 70, protein disulfide-isomerase, and thioredoxin H-type. We found one elongation factor, elongation factor 2-gamma 3, which was classified as a translation-related protein. Additionally, we identified histone H4 and two proteins with fasciclin domains associated with cell adhesion. The highest scoring cellular protein in maize starch was the 14-3-3 protein associated with the cell cycle.

There were 12 seed storage proteins identified in the maize starch samples (Table 4.6). Of these, there were four zein, three legumin, two globulin, one vicilin, and two glycinin G1 proteins. Of these 12 proteins, six were found in the aqueous fraction and six were found in the isopropanol fraction.

4.5 Discussion

We have previously examined the starch granule proteome in hard and soft wheats following very strict washing procedures (Wall et al., 2010a; Wall et al., 2010b). Here, we have examined two different species, rice, with starch granules composed of 23.6% amylose (Tan, 2003) and maize with starch granules composed of 23.1% amylose (Uthumporn et al., 2010), whose granules have been processed in large-scale, commercial environments. The protocols for commercial starch processing are optimized for high-throughput processing at low cost, rather than the absolute removal of all non-starch

Table 4.5: Proteins involved in cellular processes from *Zea mays* starch.

Accession	Protein	Mascot Score	Number of Peptides[a]	Coverage (%)[b]	Molecular function
Q01526	14-3-3-like protein GF14-12	158	2(2)	11.1	Cell cycle
AAC36132	heat shock protein 70	104	2(0)	3.8	Folding, sorting and degradation
NP_001105408	endosperm specific protein 1 – contains Fasciclin domain	97	2(2)	7.0	Cell adhesion
1101277A	histone H4	81	2(0)	17.6	Replication and repair
ACG29437	thioredoxin H-type	77	2(2)	13.9	Folding, sorting and degradation
NP_001151465	LOC100285098 – elongation factor 2	66	3(3)	4.2	Translation
P52588	protein disulfide-isomerase	63	3(3)	8.2	Folding, sorting and degradation
NP_001151356	fasciclin-like arabinogalactan protein 10	62	2(2)	4.7	Cell adhesion
ACG36145	elongation factor 1-gamma 3	58	2(2)	4.8	Translation

[a] Number of total peptides (number of unique peptides). [b] Percent protein sequence coverage by the total peptides.

Table 4.6: Seed storage proteins from *Zea mays* starch.

Accession	Protein	Mascot Score	Number of Peptides[a]	Coverage (%)[b]
AAL16994	legumin 1	213	4(1)	10.6
1802402A	globulin 2	209	4(4)	13.8
CAA41809	vicilin-like embryo storage protein	200	4(2)	9.1
P04776	glycinin G1	190	4(4)	9.1
NP_001104865	legumin 1	180	4(1)	10.6
ABV71978[c]	**z1D alpha zein protein**	**179**	**3(1)**	**21.0**
ABV71937	**z1A alpha zein protein**	**155**	**3(1)**	**23.8**
AAL16994	**legumin 1**	**128**	**3(3)**	**52.6**
ABV71937	z1A alpha zein protein	110	2(2)	10.9
1107201E	**zein M36M361**	**108**	**2(2)**	**6.4**
AAC31468	**globulin 1**	**80**	**2(2)**	**7.4**
P04776	**glycinin G1**	**64**	**2(2)**	**3.8**

[a] Number of total peptides (number of unique peptides). [b] Percent protein sequence coverage by the total peptides. [c] Entries **in bold** were sequenced from the isopropanol extract.

fractions (Tester et al., 2007). We were therefore interested in determining whether direct trypsinolysis could be used to access and identify the proteins that remain associated with the starch granules following commercial-scale processing (workflow shown in Figure 4.2A). In order to identify hydrophobic peptides on the starch granule surface, we used an isopropanol extraction protocol adapted from Morris et al. (1994). While this protocol was optimized for the isolation of the wheat SGAP puroindoline, an amphipathic protein, we were able to identify certain peptides that remained associated to the SGS following commercial starch processing and subsequent in-lab washing. While it is possible that these associations are artifactual in nature, it is feasible that they are physiologically important and are therefore worthy of further study.

4.5.1 Starch granule associated proteins

Thoroughly water washing starch granules from cereal seeds in a laboratory setting can reduce the total protein content to 0.25% (Skerritt et al., 1990), whereas commercial starches have protein contents of approximately 0.4% (Appelqvist and Debet, 1997). The endosperm-sourced proteins of cereal grains can be divided into three distinct categories: 1) seed storage proteins, 2) cellular or amyloplastic proteins, or 3) starch granule associated proteins (SGAPs). The SGAPs are defined as proteins that are located at the surface and/or interior of starch granules (Baldwin, 2001). The most abundant SGAP family, granule-bound starch synthase (GBSS), comprises up to 85% of the total internal SGAPs (Echt and Schwartz, 1981), and is involved in starch synthesis through the addition of uridine diphosphate glucose or adenine diphosphate glucose to growing glucan polymers (Baba et al., 1987; Denyer et al., 1996; Hanashiro et al., 2008). Due to the abundance of GBSS, it is appropriate that we found five different GBSS proteins in rice starch and an additional ten GBSS proteins in maize starch.

4.5.2 Starch biosynthetic enzymes

The enzymes involved in starch metabolism are well documented. For a comprehensive review, see Tetlow, 2011. Both rice and maize starch samples had a high percentage representation (34% and 50%, respectively) of proteins involved with carbohydrate metabolism (Figures 4.1A and 4.1B). This is consistent with previous studies that examined amyloplast and starch granule-associated proteomes in *Triticum aestivum* (Andon et al., 2002) and *Solanum tuberosum* (Denyer et al., 1996). Many of the proteins in the starch biosynthetic pathway (Keeling and Myers, 2010) were identified in our study (Tables 4.1 and 4.4). These proteins have been assembled in Figure 4.2B to demonstrate how much of the biosynthetic machinery of proteinaceous origin is still present in commercially processed starches. In the endosperm of monocots, other than GBSS, soluble starch synthase, and the starch branching and de-branching enzymes, and a plastidial form of ADP-glucose pyrophosphorylase (AGPase), starch biosynthetic enzymes are located in the cytoplasm (Tetlow, 2011). It is likely that during the dessication of the endosperm, these cytoplasmic enzymes became associated with the remnants of the amyloplast membrane or to the hydrophobic surface of the starch granules (Wall et al., 2010a; Wall et al., 2010b).

4.5.3 Orthophosphate dikinase

It has been proposed that the composition of the storage protein fractions and the starch-protein balance is controlled, in part, by orthophosphate dikinase (PPDK) (Mechin et al., 2007). Given that starch turnover and carbon allocation are integral in coordinating metabolism with growth and that starch content is negatively correlated with biomass (Sulpice et al., 2009), PPDK may also be involved in starch-biomass balance. The classical role of PPDK in both C3 and C4 plants is in catalyzing the reversible reaction of pyruvate, ATP and phosphate to phosphoenolpyruvate, AMP and diphosphate (Chastain and Chollet, 2003). In maize, the expression of PPDK is upregulated at 21 days after pollination (DAP) (Mechin et al., 2007), while the expression of PPDK in rice is highest from 5 – 15 DAP, after which, PPDK is rapidly degraded or inactivated through phosphorylation. This pool of inactivated PPDK is present in mature seeds, and may play a role in developmental processes during seed germination (Chastain et al., 2006). Future analysis of the starch granule-associated phosphoproteome will allow for the determination of phosphorylation levels of PPDK in these starch samples.

4.5.4 Starch mobilization

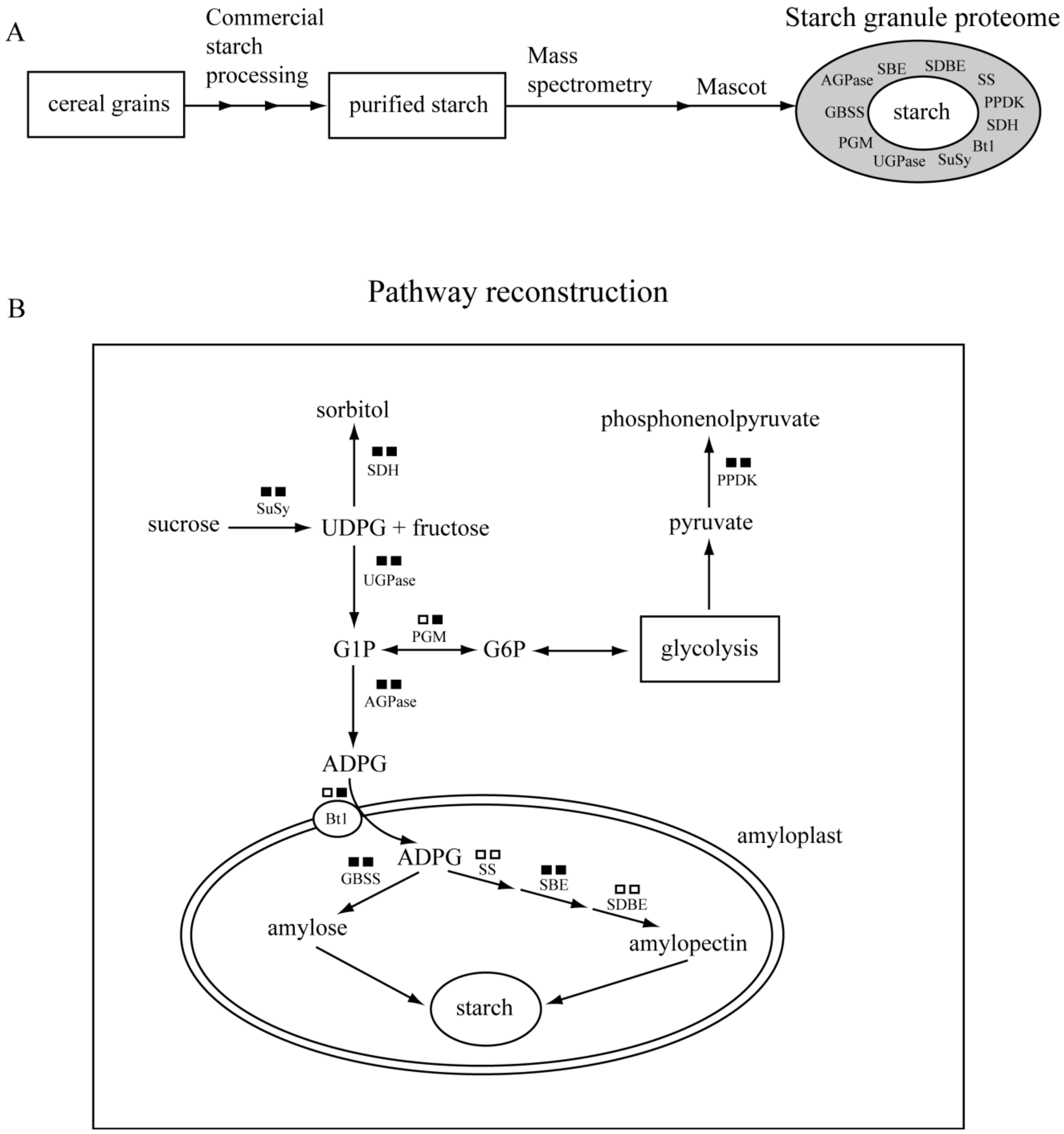
To utilize their stored carbon reserves, plants must be able to degrade their starch granules to oligosaccharides and monosaccharides. However, of the 59 proteins identified in our survey as being involved with carbohydrate metabolism, only two were involved in the mobilization of starch (Table 4.1). Those enzymes, plastidic α-1,4-glucan α-phosphorylase and α-glucosidase were both found in the rice starch samples.

Glucosidase is involved in the hydrolysis of α-1,4 glucosidic linkages (Smith et al., 2005), while plastidic α-1,4-glucan phosphorylase acts by converting the terminal glycosyl units of the non-reducing ends of linear oligosaccharide glucan chains to glucose 1-phosphate (Duwenig et al., 1997). While the well-studied α-amylase is involved in the direct degradation of the polysaccharides that make up starch granules, α-glucosidase is generally believed to preferentially hydrolyze oligosaccharides (Hanes, 1932). There is, however, evidence that α-glucosidase can act synergistically with α-amylase to enhance granule degradation in barley (Sun and Henson, 1990). While α-glucosidase expression is high during seed development in barley, the expression decreases to basal levels seeds upon maturity (Fincher, 1989). Our ability to detect these starch mobilizing enzymes in only the rice samples may indicate that some of the rice kernels were harvested prior to maturity.

4.5.5 14-3-3 proteins

The 14-3-3 protein family consists of 28–33-kDa acidic polypeptides that are ubiquitously expressed in eukaryotes (Morrison, 2009). The 14-3-3 proteins form homo- or hetero-dimers that are able to bind a broad range of ligands due to the specific phospho-serine/phospho-threonine binding activity of these 14-3-3 dimers (Fu et al., 2000). However, not all 14-3-3 binding is phosphorylation dependent (Aitken, 2006). The ligands with which the dimers of 14-3-3 have been demonstrated to interact include transcription factors, cytoskeletal proteins, signaling molecules, and biosynthetic enzymes, among others (Morrison, 2009). While it has been hypothesized that the

Figure 4.2: Workflow and validation of proteome characterization.
Panel A) For both rice and maize, the purified starch samples came from cereal grains that underwent numerous commercial starch processing steps. Direct tryptic digests and mass spectrometry identified a community of starch granule associated proteins that could be clustered around the granule as a specific proteome. Panel B) Starch biosynthetic pathway reconstruction. To validate the workflow, and characterize the defined proteome, the individual proteins were overlaid on the standard metabolic pathway of starch synthesis (Tetlow, 2011). The congruency validates both the workflow and the nature of the specific proteomes in this case. Presence/absence of protein in the proteomic survey is indicated by black (presence) or white (absence) squares; left-side squares correspond to rice, right-side squares correspond to maize. The starting point of starch biosynthesis in endosperm cells is the conversion of sucrose and uracil-diphosphate (UDP) into uracil-diphosphate glucose (UDPG) and fructose by sucrose synthase (SuSy) (Baroja-Fernandez et al., 2003). The fructose can be converted into sorbitol by sorbitol dehydrogenase (SDH) (de Sousa et al., 2008). The UPDG created by sucrose synthase along with diphosphate is converted to glucose 1-phosphate (G1P) by UTP-glucose-1-phosphate uridylyltransferase (UGPase). The G1P is either converted to adenine-diphosphate glucose (ADPG) by ADP-glucose pyrophosphorylase (AGPase), or converted to glucose 6-phosphate (G6P) by phosphoglucomutase (PGM) and enters glycolysis, where, following glycolysis, the pyruvate can be converted to phosphoenolpyruvate by pyruvate orthophosphate dikinase (PPDK). The ADPG, created from the G1P, is transported into amyloplasts by brittle 1 (Bt1), and can be used to form amylose via granule-bound starch synthase (GBSS) or amylopectin via starch synthase (SS), starch branching enzymes (SBE) and starch debranching enzymes (SDBE) (Tetlow, 2011).


A
cereal grains
Commercial starch processing
purified starch
Mass spectrometry
Mascot
Starch granule proteome
starch
AGPase
SBE
SDBE
SS
PPDK
SDH
Bt1
SuSy
UGPase
PGM
GBSS
B
Pathway reconstruction
sorbitol
SDH
sucrose
SuSy
UDPG + fructose
UGPase
G1P
PGM
G6P
glycolysis
pyruvate
PPDK
phosphonenolpyruvate
AGPase
ADPG
Bt1
amyloplast
ADPG
GBSS
SS
SBE
SDBE
amylose
amylopectin
starch

assembly of the enzyme complexes responsible for starch biosynthesis is mediated by the phosphorylation-dependent binding of 14-3-3 (Tetlow, 2011), there is no record of 14-3-3 being found in mature seeds. Previously, 14-3-3 isoforms have been found in maize starch granules isolated from immature pollen (Datta et al., 2002), as well as in Arabidopsis (Sehnke et al., 2001) and barley (Alexander and Morris, 2006). Isoforms of GBSS, soluble starch synthase, and starch branching enzyme, as well as the starch-degrading enzyme α-amylase can be bound by barley grain 14-3-3 proteins (Alexander and Morris, 2006). The particular isoform of 14-3-3 identified in this study is GF14-12 (Table 4.5), which was first cloned in 1994 and has been shown to form DNA/protein complexes (de Vetten and Ferl, 1994). This discovery of a 14-3-3 protein in mature seeds supports the role of 14-3-3 proteins in the assembly of starch biosynthetic complexes.

4.5.6 Storage proteins

The protein content of cereal endosperm can be divided into four categories based on their solubility in solvents. These Osborne fractions include albumins, soluble in water; globulins, soluble in dilute salt solutions; prolamins, soluble in alcohol/water solutions; and glutelins, soluble in weak acids/bases (Osborne, 1908). Rice endosperm is 7-8% protein (Puchongkavarin et al., 2005), which can be divided into 3.8–8.8% albumin, 9.6–10.8% globulin, 2.6–3.3% prolamin and 66–78% glutelin (Cagampan et al., 1966). In maize, the prolamins, represented by the zein family, are more prevalent, while the glutelins are less predominant, depending on the variety of maize used in the production of the maize starch samples. The percentage representation of maize endosperm proteins is usually 4-8% albumin, 3-4% globulin, 47-55% prolamin, and 38-45% glutelin (Alais and Linden, 1991). The protein bodies in rice endosperm are tightly associated with the starch granule surface (Tanaka et al., 1980), and are difficult to remove during starch granule purification, even with extended soaking in solutions of dilute base (Wang and Wang, 2001). The distribution of the seed storage proteins in the rice starch sample favored glutelin, with 84% of the seed storage proteins being identified as glutelins (Table 4.3). Moreover, we found 21 different glutelins associated with the commercial rice starch granule preparations. Given that the current genome sequence database (GenBank) of the rice cultivar Nipponbare contains 22 members of the rice glutelin gene family, it is remarkable that such a high proportion of these are clustered close to the rice granule surface and had not been removed by commercial starch processing and purification. On the other hand, the maize starch samples had an intriguingly high representation of globulins (25%) compared to the prolamins (zeins) (33%) (Table 4.6). This divergence from the expected ratios may be an artefact of the commercial starch purification conditions employed. Future experiments may allow for the determination of the sequence of storage protein isoforms with strong starch-adhering properties. These isoforms could potentially be used to breed varieties of cereals with “customizable” starch adherence that impacts bread-making characteristics or other industrial attributes.

4.6 Conclusions

We were able to look at molecular snapshots of the endosperm proteins that remained associated to starch granules following the commercial processing of starch from two different species, both with 23% amylose content. We were, consequently, able to glean insight into the physico-chemical properties of the protein remnants that make up the intra-granular packing or protein matrix of the mature seeds of two commercially important plants: rice and maize. These proteinaceous remnants could affect the yield of pure starch granules during wet milling and subsequent purification in addition to affecting the experiments in which commercially prepared starches are used as reference materials. Furthermore, the differences inferred between the intracellular architecture and molecular dynamics of rice (C3 plant) and maize (C4 plant) should be instructive to those engineering the maize C4 pathway into rice (Hibberd et al., 2008). Finally, we also found markers of seed maturity in the rice starch samples. These markers indicate that both immature and mature seeds were used in the production of the rice starch samples.

Chapter 5

Commercially produced rice and maize starches contain non-host proteins, as shown by mass spectrometry

Koziol AG, Marquez BK, Huebsch MP, Smith JC, Altosaar, I

Published in: *Journal of Cereal Science*, DOI: 10.1094/CCHEM-04-12-0043-N

Contribution of authors:

AGK contributed to the study design, performed experiments except LC-MS/MS, analyzed data, wrote the manuscript, and participated in editing the manuscript. BKM aided in experiments, analyzing data, writing the manuscript, and editing. MPH performed LC-MS/MS. JS participated in analyzing data and editing. IA participated in study design, data analysis, and editing.

5.1 Abstract

To evaluate the presence of contaminating, non-host proteins in commercially prepared rice and maize starch samples, we initiated a direct sequencing mass spectrometric proteomics survey. We discovered non-host proteins from a variety of species, including *Phytophthora cinnamomi*, *Homarus americanus,* and *Ovis aries*. Our documentation of *H. americanus* proteins in these starch samples may have food safety implications with regards to shellfish allergies. We hypothesize that these proteins were introduced to the starch samples via process wash water used in the milling and deproteination steps in the commercial preparation of the starches analyzed. The introduction of non-host proteins during commercial processing of starch samples that are used routinely in analytical studies indicates that these studies are using impure materials. Therefore, further study and documentation of the starch samples is required to ensure that all components of the samples are properly catalogued.

5.2 Introduction

In plants, glucose molecules are linked for many reasons, including reducing the osmotic potential of free sugars, sequestering and storing glucose units for diurnal/season metabolic needs, and providing dense granules for gravitropism of statoliths (Ball and Morell, 2003; Caspar and Pickard, 1989; Tetlow, 2011). The linked chains of glucose occur as two distinct polymers: amylose, which is primarily α(1→4) linked glucosyl units, with occasional α(1→6) branches, and amylopectin, while still predominantly α(1→4) linked glucosyl units, contains a much higher proportion of α(1→6) branches compared to amylose (Grimaud et al., 2008). Together, amylose and amylopectin are assembled into starch granules (SGs). The SGs in storage organs, such as the endosperm of grains or the storage parenchyma cells of tubers, act as reservoirs for the carbohydrates synthesized during photosynthesis, and have internal semi-crystalline growth rings that are the result of periodic differences in the rate of glucose polymerization during starch synthesis (Perez and Bertoft, 2010).

While the biochemical dissection of SGs *in planta* has progressed at a fine molecular scale (Graf and Smith, 2011; Yandeau-Nelson et al., 2011), the concomitant analytics of the molecular toolkit for furthering starch-based biofuels or foods has lagged (Shi and Gao, 2011; Shu et al., 2006). Starch extraction on an industrial scale, usually by wet milling, involves several steps, including the removal of the pericarp, germ, fibres, and protein matrix, as well as the purification and drying of the starch granules (Sun, 2005). Generally, wet milling includes a steeping step to allow for easy separation of the germ and to loosen the protein matrix within the endosperm. The steeping process involves fermentation by lactic acid bacteria and yeast (Cabrales et al., 2006; Singh et al., 1999).

The seeds are ground gently and the germ is separated from the endosperm based on differences in density in hydrocyclones. The endosperm slurry is ground again, and passed through a series of fine screens to remove fibres and the protein matrix. Starch is purified from any remaining protein based on density in a series of cyclones. The isolated starch granules are then dried prior to being packaged as dry powder and transported (BeMiller and Whistler, 2009).

Commercially extracted starch is ubiquitously present in both industrial and culinary applications. The industrial applications include papermaking, adhesives, bioplastics, biofuel (Kaur et al., 2012; Pitak and Rakshit, 2011; Tasic and Veljkovic, 2011), while modified starches and starch sugars are important food additives (BeMiller, 2011). We have previously documented the remnants of pathogenic microbial proteins on the surface of starch granules in lab-scale purified wheat starch; however, we did not observe proteins from organisms from soil- or water-dwelling organisms (Wall et al., 2010). In our recent study to characterize the proteomes of commercially prepared reference starches from leading biochemical reagent suppliers, we were able to document many host-sourced proteins that remain in starch samples following industrial processing (Koziol et al., 2012). In addition to these host proteins, we discovered proteins from a broad taxonomic range of organisms. It has previously been stated that the purity of the water used to process the starch during commercial purification can contaminate starch samples, and affects the end quality of the starch (Sun, 2005). These findings serve as a caveat for all aspects of biomass research as well as food science and food processing where such biochemical reagents have been assumed to be pure starch. We therefore hypothesized that the non-host proteins present in the commercial starch preparations were sourced from the water used during the extraction of the starch granules.

5.3 Methods

Methods were as published by Koziol et al (2012). Briefly, rice (cat. no.S7260-1KG, lot no. 107K0052) and maize (cat. no.S4126-2KG, lot no. 015K0144) starches were obtained from Sigma-Aldrich (St. Louis, MO). Prior to sampling, reagent bottles were gently agitated, and four separate 50 mg aliquots from each bottle were removed, with gentle shaking between each sampling step.

5.3.1 Protein extraction

The sampled starches were suspended in 350 µL of ddH2O and incubated for 30 min with gentle agitation. Ammonium bicarbonate was added to the samples to a final concentration of 100 mM. To proteolytically digest the proteins present in the samples, 5 µg of 1 µg/µL trypsin (T6567, Sigma-Aldrich) in 1 mM HCl was added. Samples were incubated overnight at 37ºC with gentle agitation. To collect the peptides in the supernatants, samples were centrifuged (Heraeus Sepatech Biofuge A 1302; 18 000 x *g*, 1 min) and the supernatants were transferred into fresh tubes. Peptides in the supernatants were vacuum-dried (Speed Vac Concentrator SVC 100H; Savant Instruments, Inc. Hicksville, NY) for 3 hours. Starch pellets were washed five times in distilled water, were suspended in protein extraction buffer (50% isopropanol (v/v), 50 mM NaCl), and were gently mixed for 45 min. Samples were centrifuged (2 500 x *g*, 5 min). Supernatants were transferred to new tubes. Peptides in the supernatants were vacuum-dried as described above. Starch pellets were discarded. To purify and concentrate the protein fractions, vacuum-dried peptides were resuspended in 40 µL ddH_2O. Peptide suspensions were purified using ZipTips (Millipore, Bedford MA, cat. no. ZTC18M096). The manufacturer's protocol was repeated eight times per sample. Resulting peptide solutions were vacuum-dried as described above.

5.3.2 Protein separation

Peptide separation was accomplished using a 7 cm by 200 µm inner diameter trap column coupled to a 5 cm by 75 µm Picofrit nanoelectrospray tip (New Objective, Woburn, MA) packed with 5 µm Magic C18AQ resin (Michrom Bioresources, Auburn, CA). Using a tee union prior to the columns, the flow from an HP 1090 HPLC (Agilent Technologies, Santa Clara, CA) was split to pass the following gradient of acetonitrile containing 0.1% formic acid at a flow rate of approximately 250 nL/min: 0 min 2%, 3 min 5%, 8 min 10%, 70 min 35%, 75 min 80%, 83.5 min 100% (wash step), 84 min 2%, 90 min 2% (re-equilibration steps). Eluted peptides were analyzed using a QSTAR XL QqTOF mass spectrometer (AB Sciex, Foster City, CA) operating in information-dependent acquisition mode. Mass analysis included a 1 s survey scan followed by two 2 s and two 3 s tandem mass spectrometric scans on the most intense peaks in the spectrum for each cycle. Masses could be sequenced twice before being added to an exclusion list for 90 s.

Tandem mass spectrometric data were interpreted using Mascot software (version 2.3) (Matrixscience Ltd., Boston, MA) using a custom database containing all proteins from the 20100704 comprehensive National Center for Biotechnology Information non-redundant (NCBInr) protein database (11,365,658 sequences; 3,876,535,693 residues). One missed cleavage was allowed, and mass tolerances were set to ±50 ppm for peptides and ±0.2 Da for fragment ion spectra. Oxidation of methionine residues was selected as a variable modification, while no fixed modifications were selected. An automated decoy search was performed, and the rate of false positives was approximately 1%. Each match from the database was verified manually and spectra were only used if at least three successive y- or b-ions were present. Only proteins with two or more peptide-based identifications with Mascot scores ≥ 55 ($p < 0.05$) were included to decrease the chance of matches occurring due to random events.

5.4 Results and Discussion

To better understand the source and nature of the proteins present in commercial starch samples, our approach was applied to both commercially prepared rice and maize starches. These commercial starches have been recently characterized at the level of the host plant proteome (Koziol et al., 2012), but proteins of non-plant origin had not been investigated. Surprisingly, three non-host proteins were identified in the rice samples, and three different non-host proteins were identified in the maize samples (Tables 5.1 and 5.2). Mascot and manual NCBI BLAST searches revealed that these non-host proteins were of microbial and animal origin. The prevalence of the non-host proteins in the samples was approximately 5% calculated by occurrence in Mascot hit tables (data not shown) (Koziol et al 2012). Of the non-host proteins identified, only one (glucose regulated protein/BiP - heat shock protein 70-like) contains a domain that binds carbohydrates (Di Luccio et al., 2007). It is therefore presumed that the other proteins bound non-specifically to the starch granules (Nelson Jr. and Lebrun, 1956).

Microbial remnants from algae and fungi were identified in the starch samples (Table 2).

While the common plant and tree fungus, *Phytophthora cinnamomi*, is not known to

Table 5.1: Non-host protein contaminants associated to rice starch granule.

Accession	Protein[a]	Species	Mascot Score[b]	Percent Coverage (%)[b]	Peptide Sequences
671058A	glyceraldehydephosphate dehydrogenase	*Homarus americanus*	89	4.5	LTGMAFR IGIDGFGR
ADB93318	elongation factor 1-alpha	*Lymanopoda dietzi*	64	8.4	STTTGHLIYK GSFDYAWVLDK
P68251	protein kinase C inhibitor protein 1	*Ovis aries*	63	14.0	NLLSVAYK KQQMGKEYK DSTLIMQLLR

[a] Protein names produced in Mascot database search. Sequences of unknown or hypothetical proteins were searched in protein BLAST and the protein match with the greatest score is shown.

[b] Percent coverage is the percentage of the full-length protein that is represented by the peptide sequences.

target maize, it causes root rot, stem cankers and crown rot in ornamental and forest species such as cranberry, pear and pine (dos Santos et al., 2011; Hardham, 2005). Being a water mould, it is common for *Phytophthora* species to be spread through cereal crops during irrigation (Hong and Moorman, 2005), allowing for infection of many plants in the same field. Alternatively, *P. cinnamomi* may have been introduced to the starch samples during starch processing. Common steps in wet-milling of cereal grains such as sodium dioxide steeping and lactic acid fermentation may also provide optimal growing conditions for thiophilic and acidophilic bacteria that could potentially contaminate starch samples (Sun, 2005). Such contaminants in wheat starch preparations are known to include microbial enzymes (Wall et al., 2010), so these current results underscore the need to be cautious when using such starch preparations in analytical studies, especially in biomass or biomedical studies, where starch-protein interactions are being queried.

Non-microbial contaminants were also found in the commercial starch samples of rice and maize. These proteins are listed among the other contaminants in Tables 1 and 2. Using Mascot and BLAST searches of the non-redundant NCBI database, the contaminating proteins were found to originate from a variety of species, including *Homarus americanus, Lymanopoda dietzi,* and *Ovis aries.* The proteins found from these species are conserved across the animal kingdom; elongation factor 1-alpha, glyceraldehyde phosphate dehydrogenase, protein kinase C inhibitor, and skeletal actin. The purity of commercial starch is highly dependent on the quality of water used in the processing streams removing the bran and storage components of the starting flour feedstock materials. The unexpected presence of proteins from non-host organisms in the starch samples suggests that cross-contamination between water sources occurred in the

Table 5.2: Non-host proteins identified in maize starch samples.

Accession	Protein[a]	Species	Mascot Score	Percent Coverage (%)[b]	Peptide Sequences
CAA53369	glucose regulated protein /BiP - heat shock protein 70-like	*Phytophthora cinnamomi*	153	5.3	LSQEEIDR VEIIANDQGNR IINEPTAAAIAYGIDK AVFPSIVGR
ACI23566	skeletal muscle actin 3	*Homarus americanus*	75	9.3	AGFAGDDAPR EEYDESGPGIVHRKCF
BAE48223	heat shock protein 70	*Chlorella pyrenoidosa*	55	4.2	YKAEDETHRAR IINEPTAAAIAYGLDK

[a] Protein names produced in Mascot database search. Sequences of unknown or hypothetical proteins were searched in protein BLAST and the protein match with the greatest score is shown.

[b] Percent coverage is the percentage of the full-length protein that is represented by the peptide sequences.

starch mills. As *Chlorella pyrenoidosa* is a fresh water alga (Koziol et al., 2007), it was likely introduced to the starch samples by the water used in the starch processing.

The ease of starch contamination during commercial processing is disconcerting, especially since starch is a consumable product. It is therefore possible that a subset of the contaminants could induce allergic reactions in susceptible individuals. It is the protein component of foodstuffs that is commonly responsible for the development of food allergies (Ramesh, 2008). Recently, the prevalence of shellfish allergies in American children was reported as 17.2% (Gupta et al., 2011). Allergies to shellfish are caused by the protein tropomyosin (Ramesh, 2008). While we did not observe tropomyosin in the starch samples, we documented proteins attributed to *Homarus americanus* (American lobster) in both the rice and maize starch samples. With this discovery of *H. americanus* proteins in commercial starch samples, a risk of triggering shellfish allergies has been identified. It may, therefore, be prudent to screen industrially-extracted starches more rigorously for the presence of any potential allergens.

5.5 Conclusion

We have analyzed commercially prepared starches of two agriculturally important crop species: rice and maize by sampling the starch granule associated proteins via water washing and isopropanol stripping followed by direct sequencing tandem mass spectrometry. Unexpectedly, we identified non-host proteins, which included microbial remnants and non-microbial contaminants on the surface of starch granules of both species. Six non-host proteins were identified in the starch samples. More importantly, the results demonstrated that this methodology can be used to ensure that potential allergenic proteins are not present in starch samples to be consumed by animals.

Chapter 6

General Discussion of Major Findings

6.1 General Discussion of Chapter 2

The work in this thesis combined data from 2D-electrophoresis, immunoblotting , N-terminal sequencing and MS/MS sequencing of peptides resulting from the endoproteolytic cleavage of salt-soluble globulin-3 in wheat flour preparation enriched in embryo tissue. The results show that the globulin-3 family is subjected to post-translational processing and accumulates in the form of a set of cleavage products. Previous work that described this protein family was performed on the hexaploid wheat cultivar Glenlea (Loit et al., 2009), while the work performed here was on cultivar AC Barrie, the cultivar that was used in the experiments that discovered the putative link between dietary intake of Glo-3A and the development of type 1 diabetes (MacFarlane et al., 2003). Immunoblot analysis of proteins from both embryo and endosperm tissues revealed that the Glo-3 protein family is expressed predominantly in the embryo with very low expression levels in the endosperm (Figure 2.2). Sequencing protein spots extracted from 2D PAGE gels of salt-soluble embryo proteins using mass spectrometry allowed for the development of a model describing the post-translational partial endoproteolytic cleavage of the preproglobulin-3 family (Figure 2.3).

6.1.1 Glo-3 expression pattern and food safety

Our findings that demonstrated that the expression of Glo-3 was predominantly in the embryo confirms immunolocalization data of Glo-3 in Glenlea (Loit et al., 2009) and therefore demonstrates that Glo-3 shares similar expression patterns to 7S globulins in rice, maize and other cereals (Burgess and Shewry, 1986; Sun et al., 1996; Thijssen et al., 1996). Wheat gluten, first isolated and named by Bartolomeo Beccari in Bologna in 1745 (Beccari, 1745), is today still a generic term that describes a mixture of gliadins and glutelins; seed storage proteins expressed in the endosperm, but not the embryo (van Herpen et al., 2008). However, Glo-3 can be present at very low levels in wheat gluten (MacFarlane et al., 2003); as commercial gluten is often made from whole wheat, which consists of the endosperm and embryo, as well as the aleurone layer. To study the potential diabetogenic effects of Glo-3, an investigation is required that compares wheat gluten that is prepared exclusively from endosperm and therefore does not contain Glo-3 to gluten that is prepared from whole wheat. Stripping or milling off the outer layers of wheat seeds has been demonstrated to reduce the sensitivity of predisposed individuals to wheat flour (Handoyo et al., 2008), as these polished seeds contained lower levels of allergenic proteins. Alternatively, wheat could be genetically transformed to express lower levels of the globulin-3 protein family. A Glo-3 knockdown/knockout line of wheat would allow for increased precision in studying the diabetogenic effects of Glo-3. Using gluten prepared solely from endosperm compared to the entire wheat seed will not only exclude Glo-3, but also other embryo/aleurone layer specific proteins, making it more difficult to definitively prove a role for Glo-3 in the development in T1D. Knockout lines should be identical to the parental lines, with the exception of Glo-3 expression. Maize lines that are *Glb1* and *Glb2* null have been shown to be able to germinate normally (Kriz, 1989), indicating that knocking out the globulin family is not a lethal mutation.

6.1.2 Post-translational endoproteolytic cleavage of Glo-3

The endoproteolytic processing of some wheat flour protein families has been investigated using N-terminal sequencing (Singh et al., 2001) and mass spectrometry (Dupont et al., 2011), but a concentrated study of the processing of Glo-3 specifically has not been reported before. This present thesis used multiple characteristics of protein spots to determine the location of processing events that occurred. The pI and M_r of the spots, the location of the sequenced peptides within the protein sequence, and the N-terminal sequence data (when available) were all taken into account when assembling the processing model (Figure 2.3). As was demonstrated in Figures 2.2 and 2.3, and in agreement with the literature, the endoproteolytic processing of the Glo-3 protein is partial. It is still unknown how the proteins are prevented from not being completely processed at every protease recognition site; though it has been determined that cleavage is not required for proper folding, transport and storage of these proteins (Heck et al., 1993).

With a better understanding of the processing of the Glo-3 protein family, it is now possible to study the diabetogenic effects of the individual polypeptides generated from the cleavage of proglobulin. The genetic engineering of wheat is now common practice, although the 200 odd traits modified to date are mostly agronomic. Wheat can therefore also be bred or engineered to lack certain protease recognition sites or domains within Glo-3. The wheat lacking either certain domains or cleavage sites can then be used in animal feeding trials in an attempt to locate a specific region within the protein that displays increased diabetogenic activity. Given that this type of engineering already promises to benefit celiac patients (Gil-Humanes et al., 2010), the same effort should now be expended for a much more expensive disease, T1D.

6.2 General Discussion of Chapter 3

This work has increased the understanding of the protein expression by pathogen and host in water-treated or *Fusarium graminearum*-infected maize kernels. Previous studies have investigated host-*F. graminearum* interactions in actively growing kernels hours or days after infection (Mohammadi et al., 2011; Zhou et al., 2006), but I was interested in the proteins present in the host and pathogen proteomes following the desiccation of mature kernels under field conditions. Our group has previously documented fungal proteins present in healthy, mature wheat seeds (Wall et al., 2010), I therefore decided to extend these investigations by sequencing the proteome of healthy, mature maize seeds, as well as maize seeds with a heavy pathogen load to determine the number and nature of pathogen and host proteins present under these two conditions. In healthy plants, we found no evidence of fungal proteins within the seeds. However, we determined that fungal proteins are present in both the desiccated fungal remnants on the surface of infected seeds as well as within the seeds' tissues.

6.2.1 Proteins in desiccated fungal remnants

It was intriguing that we only discovered 18 proteins in the fungal remnants despite the fact that silver-stained PAGE gels showed a complex banding pattern (Figure S1), and that there were many fungal proteins sequenced in the endosperm and embryo tissues, it is possible that increased sampling will reveal an increased number of proteins in the fungal remnants. Alternatively, as the seed contains chaperone proteins that aid in maintaining protein folding and integrity, these chaperone(s) may have allowed for the preservation of the fungal proteins within the seed, while the proteins in the fungal remnants were degraded.

6.2.2 Maize proteome response to fungal infection

This study used a gel-free direct peptide sequencing protocol to investigate the host and pathogen proteomes present in mature kernels. Previous studies have used 2D electrophoresis to separate protein spots, with individual spots being excised from the gel and sequenced with mass spectrometry. These protocols are time- and labour-consuming, though they allow for quantitation of protein levels and the sequencing of a larger portion of the proteome when compared to the gel-free methods used in this study. Nevertheless, the gel-free methods were used here as they are rapid to perform and provide sufficient proteomic sequence data for analysis.

The study of infected and control maize endosperm and embryo tissues independently has not previously been performed. The separation of the tissues allowed for the investigation at a scale smaller than the whole seed. As expected, the endosperm tissue had a small increase in the number of defense-related proteins present in the infected samples compared to the controls, while the embryo had a much larger increase in the number of proteins when the two conditions were compared (Figure 3.1). It would be interesting to compare rates of appearance of host defense proteins in the two respective tissues, diploid embryo and triploid endosperm now that we know more about how apoptosis is governed in developing cereal seeds (Chaban et al., 2011; Wang et al., 2012).

6.3 General Discussion of Chapter 4

A previous study investigated the proteomes of wheat starch purified by several different commercial companies to investigate the suitability of these starches for celiac patients (Kasarda et al., 2008). In that study, non-seed storage proteins such as α-amylase inhibitors as well as enolases were identified. We were therefore interested in determining whether commercially prepared rice and maize starches also contained these non-seed storage proteins. Commercial preparation of a bulk commodity such as starch implies that removal of protein is not absolutely complete. This work demonstrated that the contaminating 1% protein content of commercially prepared rice and maize starches includes several biochemical classes of proteins, including many proteins involved with starch metabolism (Figure 4.1A,B).

Additionally, in an unpublished analysis of commercially purified wheat starch with results shown in the Supplemental Materials of this thesis, I discovered that Glo-3 proteins were present in these starch samples (Table S1).The wheat starch was prepared and analyzed using the same protocols as the rice and maize starches in Chapter 3 of this thesis. Although these particular commercially-prepared starches are not intended for human consumption, the presence of Glo-3 could invalidate certain studies if these starches are considered “pure starch controls” in animal feeding trials to study type 1 diabetes or celiac disease. Furthermore, starches that are intended for animal feed, such as

Peking duck force-feedings, are known to be watched by federal surveillance authorities for presence of *Fusarium* toxins, both small molecules like vomitoxin as well as proteins (Danicke et al., 2004).

6.4 General Discussion of Chapter 5

Using mass spectrometry, I show that there are non-host proteins present in commercially-prepared rice and maize starch samples. When starches from healthy or infected wheat or maize seeds are prepared in the lab I discovered that there are certain non-host proteins present in the wheat starch (Wall et al., 2010), but not in maize starch (This thesis, chapter 3). This is potentially due to the manner in which the starches are prepared. The wheat was stored in paper bags after it had been hulled, a procedure that removes the protective palea and lemma from the seed, allowing for organisms to colonize the seed surface. During the starch extraction, wheat seeds were sterilized in hypochlorite and rinsed with distilled water prior to grinding and starch extraction (Wall et al., 2010). While adequate to remove microorganisms and large debris, it is unlikely that this protocol would strip away all foreign proteins present on the surface of the seeds. Conversely, the protocol for extracting maize starch in the laboratory involves a long steeping step in lactic acid, followed by the careful dissection of the endosperm from the pericarp and embryo tissues. The removal of the pericarp, and therefore any non-host surface proteins attached to this tissue, should reduce the number of non-host proteins in the starch samples. When comparing the commercially prepared starches to these control starches prepared in the laboratory, the presence of proteins from *Homarus americanus*, *Chlorella pyrenoidosa*, and *Ovis aries* can therefore be attributed to the commercial starch preparation protocols. While these proteins are present in low concentrations in the samples, their presence, like Glo-3 in wheat starch, must be taken into account when interpreting studies using these starches as controls. Low levels of wheat proteins have been detected in the urine of schizophrenia patients so it is important to develop more such refined techniques to monitor the presence, biotransformation and migration of food-derived peptides (Samaroo et al., 2010).

6.5 Translational medicine

With the current high-throughput technologies available today, the phrase “you are what you eat” is more relevant than ever. Cereal proteins and their degradation products in the diet are being shown to be involved with an ever increasing number of biochemical pathways. For instance, studies on celiac disease are able to show, with great precision, that in predisposed individuals, the partial proteolytic degradation of gliadins by gastrointestinal enzymes can produce the so-called toxic peptides and the immunodominant peptides that damage epithelial cells, or activate mucosal T-cell-mediated adaptive immunity, respectively (Lindfors et al., 2012). Additionally, “nutropioids”, the food-derived opioid oligopeptides, have been studied for many years e.g. (Samaroo et al., 2010; Zioudrou et al., 1979). These oligopeptides with μ-opioid activity are recognized by the μ-opioid receptors (MOR) that are expressed exclusively in the brain and the small intestine, where they are involved with controlling food intake levels (Pfluger et al., 2012) and gut motility and bowel movements (Moughan et al., 2007), respectively. Recently, satiety effects of protein rich diets have been associated with the antagonism of MORs (Duraffourd et al., 2012).

Given that dietary cereal proteins can have effects that range from the development of diseases and allergies to cravings and satiety, the use of mass spectrometry to study these proteins is imperative. In this thesis, I used this tool to gain an understanding of the post-translational endoproteolytic cleavage of Glo-3 and the proteomes of cereals grown under normal or pathogen-infected conditions. The data generated in this thesis not only allowed me to test my hypotheses, but will ultimately allow for a greater understanding of the proteins that are ingested on a daily basis worldwide.

References

Adachi, M., Takenaka, Y., Gidamis, A., Mikami, B., and Utsumi, S. (2001). Crystal structure of soybean proglycinin alaB1b homotrimer. J. Mol. Biol. *305,* 291-305.

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. Nature *422,* 198-207.

Aitken, A. (2006). 14-3-3 proteins: A historic overview. Semin. Cancer Biol. *16,* 162-172.

Alais, C., and Linden, G. (1991). Food Biochemistry (Nancy, France: Ellis Horwood).

Alexander, R.D., and Morris, P.C. (2006). A proteomic analysis of 14-3-3 binding proteins from developing barley grains. Proteomics *6,* 1886-1896.

Almagro, L., Ros, L.V.G., Belchi-Navarro, S., Bru, R., Barcelo, A.R., and Pedreno, M.A. (2009). Class III peroxidases in plant defence reactions. J. Exp. Bot. *60,* 377-390.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215,* 403-410.

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.

Angel Torres, M. (2010). ROS in biotic interactions. Physiol. Plantarum *138,* 414-429.

Appelqvist, I.A.M., and Debet, M.R.M. (1997). Starch-biopolymer interactions - A review. Food Rev. Int. *13,* 163-224.

Appenzeller, L., Doblin, M., Barreiro, R., Wang, H., Niu, X., Kollipara, K., Carrigan, L., Tomes, D., Chapman, M., and Dhugga, K. (2004). Cellulose synthesis in maize: isolation and expression analysis of the cellulose synthase (*CesA*) gene family. Cellulose *11,* 287-299.

Ariizumi, T., Higuchi, K., Arakaki, S., Sano, T., Asamizu, E., and Ezura, H. (2011). Genetic suppression analysis in novel vacuolar processing enzymes reveals their roles in controlling sugar accumulation in tomato fruits. J. Exp. Bot. *62,* 2773-2786.

Baba, T., Yoshii, M., and Kainuma, K. (1987). Acceptor molecule of antigranulocytes-bound starch synthase from sweet-potato roots. Starch-Starke *39,* 52-56.

Baker, A.A., Miles, M.J., and Helbert, W. (2001). Internal structure of the starch granule revealed by AFM. Carbohydr. Res. *330,* 249-256.

Balconi, C., Lanzanova, C., and Motto, M. (2010). Ribosome-Inactivating Proteins in Cereals. In Toxic Plant Proteins, Lord, J. M., and Hartley, M. R. eds., (Berlin: Springer) pp. 149-166.

Baldo, B.A., and Wrigley, C.W. (1978). IgE antibodies to wheat-flour components - studies with sera from subjects with baker's asthma or celiac condition. Clin. Allergy *8,* 109-124.

Baldwin, P.M. (2001). Starch granule-associated proteins and polypeptides: A review. Starch-Starke *53,* 475-503.

Baldwin, P.M., Adler, J., Davies, M.C., and Melia, C.D. (1994). Holes in starch granules - confocal, SEM and light-microscopy studies of starch granule structure. Starch-Starke *46,* 341-346.

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

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.

Barbieri, L., Battelli, M., and Stirpe, F. (1993). Ribosome-inactivating proteins from plants. Biochim. Biophys. Acta *1154,* 237-282.

Baroja-Fernandez, E., Munoz, F.J., Saikusa, T., Rodriguez-Lopez, M., Akazawa, T., and Pozueta-Romero, J. (2003). Sucrose synthase catalyzes the de novo production of ADP-glucose linked to starch biosynthesis in heterotrophic tissues of plants. Plant and Cell Physiology *44,* 500-509.

Beccari, J.B. (1745). De frumento. De Bononiensi Scientiarum et Artium Instituto Atque Academia Commentarii *2,* 122-127.

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

Belanger, F.C., and Kriz, A.L. (1989). Molecular characterization of the major maize embryo globulin encoded by the *glb1* gene. Plant Physiol. *91,* 636-643.

BeMiller, J.N., and Whistler, R.L. (2009). Starch: Chemistry and Technology (Burlington, MA: Academic Press).

BeMiller, J.N. (2011). Pasting, paste, and gel properties of starch-hydrocolloid combinations. Carbohydr. Polym. *86,* 386-423.

Bennett, J., and Klich, M. (2003). Mycotoxins. Clin. Microbiol. Rev. *16,* 497-516.

Biely, P., Vrsanska, M., Tenkanen, M., and Kluepfel, D. (1997). Endo-beta-1,4-xylanase families: differences in catalytic properties. J. Biotechnol. *57,* 151-166.

Bindschedler, L.V., Dewdney, J., Blee, K.A., Stone, J.M., Asai, T., Plotnikov, J., Denoux, C., Hayes, T., Gerrish, C., Davies, D.R., Ausubel, F.M., and Bolwell, G.P. (2006). Peroxidase-dependent apoplastic oxidative burst in *Arabidopsis* required for pathogen resistance. Plant Journal *47,* 851-863.

Bittner, C., Grassau, B., Frenzel, K., and Baur, X. (2008). Identification of wheat gliadins as an allergen family related to baker's asthma. J. Allergy Clin. Immunol. *121,* 744-749.

Borlaug, N. (1983). Contributions of conventional plant-breeding to food-production. Science *219,* 689-693.

Bostock, J., and Riley, H.T. (1855). Pliny the Elder, the Natural History (London: Taylor and Francis).

Bradford, M.M. (1976). Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding. Anal. Biochem. *72,* 248-254.

Buléon, A., Colonna, P., Planchot, V., and Ball, S. (1998). Starch granules: structure and biosynthesis. Int. J. Biol. Macromol. *23,* 85-112.

Burgess, S.R., and Shewry, P.R. (1986). Identification of homologous globulins from embryos of wheat, barley, rye and oats. J. Exp. Bot. *37,* 1863-1871.

Cabrales, L., Niu, Y., Buriak, P., and Eckhoff, S. (2006). Effect of laboratory batch steeping pH on starch yield and pasting properties of selected corn hybrids. Cereal Chem. *83,* 22-24.

Cagampan, G.B., Cruz, L.J., Espiritu, S.G., Santiago, R.G., and Juliano, B.O. (1966). Studies on extraction and composition of rice proteins. Cereal Chem. *43,* 145-155.

Campo, S., Carrascal, M., Coca, M., Abian, J., and San Segundo, B. (2004). The defense response of germinating maize embryos against fungal infection: a proteomics approach. Proteomics *4,* 383-396.

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

Caspar, T., and Pickard, B. (1989). Gravitropism in a starchless mutant of *Arabidopsis* - implications for the starch-statolith theory of gravity sensing. Planta *177,* 185-197.

Catassi, C., Ratsch, I.M., Fabiani, E., Rossini, M., Bordicchia, F., Candela, F., Coppa, G.V., and Giorgi, P.L. (1994). Celiac-disease in the year 2000 - exploring the iceberg. Lancet *343,* 200-203.

Cech, N.B., and Enke, C.G. (2001). Practical implications of some recent studies in electrospray ionization fundamentals. Mass Spectrom. Rev. *20,* 362-387.

Chaban, I.A., Lazareva, E.M., Kononenko, N.V., and Polyakov, V.Y. (2011). Antipodal complex development in the embryo sac of wheat. Russian Journal of Developmental Biology *42,* 79-91.

Chastain, C.J., and Chollet, R. (2003). Regulation of pyruvate, orthophosphate dikinase by ADP-/Pi-dependent reversible phosphorylation in C-3 and C-4 plants. Plant Physiology and Biochemistry *41,* 523-532.

Chastain, C.J., Heck, J.W., Colquhoun, T.A., Voge, D.G., and Gu, X.Y. (2006). Posttranslational regulation of pyruvate, orthophosphate dikinase in developing rice (*Oryza sativa*) seeds. Planta *224,* 924-934.

Chernushevich, I.V., Loboda, A.V., and Thomson, B.A. (2001). An introduction to quadrupole-time-of-flight mass spectrometry. Journal of Mass Spectrometry *36,* 849-865.

Chrispeels, M.J. (1991). Sorting of proteins in the secretory system. Annu. Rev. Plant Physiol. Plant Mol. Biol. *42,* 21-53.

Chrispeels, M.J., Higgins, T.J.V., and Spencer, D. (1982). Assembly of storage protein oligomers in the endoplasmic reticulum and processing of the polypeptides in the protein bodies of developing pea cotyledons. J. Cell Biol. *93,* 306-313.

Cledat, D., Battu, S., Mokrini, R., and Cardot, P.J. (2004). Rice starch granule characterization by flow cytometry scattering techniques hyphenated with sedimentation field-flow fractionation. J. Chromatogr. A *1049,* 131-138.

Cuomo, C.A., Gueldener, U., Xu, J., Trail, F., Turgeon, B.G., Di Pietro, A., Walton, J.D., Ma, L., Baker, S.E., Rep, M.*, et al.* (2007). The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. Science *317,* 1400-1402.

Dale, E.M., and Housley, T.L. (1986). Sucrose synthase activity in developing wheat endosperms differing in maximum weight. Plant Physiol. *82,* 7-10.

Danicke, S., Matthaus, K., Valenta, H., and Halle, I. (2004). Effects of *Fusarium*-toxin contaminated wheat grains and non-starch-polysaccharide (NSP) hydrolyzing enzyme preparation on Pekin duck performance. Archiv Fur Geflugelkunde *68,* 199-205.

Danielsson, C.E. (1949). Seed globulins of the gramineae and leguminosae. Biochem. J. *44,* 387-400.

Datta, R., Chamusco, K.C., and Chourey, P.S. (2002). Starch biosynthesis during pollen maturation is associated with altered patterns of gene expression in maize. Plant Physiol. *130,* pp. 1645-1656.

de Sousa, S.M., Paniago, M.D., Arruda, P., and Yunes, J.A. (2008). Sugar levels modulate sorbitol dehydrogenase expression in maize. Plant Mol. Biol. *68,* 203-213.

de Vetten, N.C., and Ferl, R.J. (1994). Two genes encoding GF14 (14-3-3) proteins in *Zea mays*. Structure, expression, and potential regulation by the G-box binding complex. Plant Physiol. *106,* 1593-1604.

Debiton, C., Merlino, M., Chambon, C., Bancel, E., Decourteix, M., Planchot, V., and Branlard, G. (2011). Analyses of albumins, globulins and amphiphilic proteins by proteomic approach give new insights on waxy wheat starch metabolism. J. Cereal Sci. *53,* 160-169.

Deng, X., Hahne, T., Schroeder, S., Redweik, S., Nebija, D., Schmidt, H., Janssen, O., Lachmann, B., and Waetzig, H. (2012). The challenge to quantify proteins with charge trains due to isoforms or conformers. Electrophoresis *33,* 263-269.

Denyer, K., Clarke, B., Hylton, C., Tatge, H., and Smith, A.M. (1996). The elongation of amylose and amylopectin chains in isolated starch granules. Plant Journal *10,* 1135-1143.

Derbyshire, E., Wright, D., and Boulter, D. (1976). Legumin and vicilin, storage proteins of legume seeds. Phytochemistry *15,* 3-24.

Deschamps, P., Colleoni, C., Nakamura, Y., Suzuki, E., Putaux, J., Buleon, A., Haebel, S., Ritte, G., Steup, M., Falcon, L.I., *et al.* (2008). Metabolic symbiosis and the birth of the plant kingdom. Mol. Biol. Evol. *25,* 536-548.

Deschamps, P., Haferkamp, I., d'Hulst, C., Neuhaus, H.E., and Ball, S.G. (2008). The relocation of starch metabolism to chloroplasts: when, why and how. Trends Plant Sci. *13,* 574-582.

Dhawan, K., Malhotra, S., Dahiya, B., and Singh, D. (1991). Seed protein-fractions and amino-acid-composition in gram (*Cicer-arietinum*). Plant Foods for Human Nutrition *41,* 225-232.

Di Luccio, E., Petschacher, B., Voegtli, J., Chou, H., Stahlberg, H., Nidetzky, B., and Wilson, D.K. (2007). Structural and kinetic studies of induced fit in xylulose kinase from *Escherichia coli*. J. Mol. Biol. *365,* 783-798.

Di Sabatino, A., and Corazza, G.R. (2009). Coeliac disease. Lancet *373,* 1480-1493.

Diaz, A., and Hampshire, S. (2004). Characterisation of porous silicon nitride materials produced with starch. Journal of the European Ceramic Society *24,* 413-419.

Dongre, A.R., Jones, J.L., Somogyi, A., and Wysocki, V.H. (1996). Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: Evidence for the mobile proton model. J. Am. Chem. Soc. *118,* 8365-8374.

dos Santos, A.F., Tessmann, D.J., Alves, T.C.A., Vida, J.B., and Harakava, R. (2011). Root and crown rot of Brazilian pine (*Araucaria angustifolia*) caused by *Phytophthora cinnamomi*. J. Phytopathol. *159,* 194-196.

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 Science *138,* 121-135.

Dunaevskii, Y., Tsybina, T., Belyakova, G., Domash, V., Sharpio, T., Zabreiko, S., and Belozerskii, M. (2005). Proteinase inhibitors as antistress proteins in higher plants. Appl. Biochem. Microbiol. *41,* 344-348.

Dunwell, J.M. (1998). Cupins: a new superfamily of functionally diverse proteins that include germins and plant storage proteins. Biotechnology & Genetic Engineering Reviews *15,* 1-32.

Dupont, F.M., Vensel, W.H., Tanaka, C.K., Hurkman, W.J., and Altenbach, S.B. (2011). Deciphering the complexities of the wheat flour proteome using quantitative two-dimensional electrophoresis, three proteases and tandem mass spectrometry. Proteome Sci. *9,* 10.

Duraffourd, C., De Vadder, F., Goncalves, D., Delaere, F., Penhoat, A., Brusset, B., Rajas, F., Chassard, D., Duchampt, A., Stefanutti, A., Gautier-Stein, A., and Mithieux, G. (2012). Mu-opioid receptors and dietary protein stimulate a gut-brain neural circuitry limiting food intake. Cell *150,* 377-388.

Durrant, W.E., and Dong, X. (2004). Systemic acquired resistance. Annu. Rev. Phytopathol. *42,* 185-209.

Duwenig, E., Steup, M., Willmitzer, L., and Kossmann, J. (1997). Antisense inhibition of cytosolic phosphorylase in potato plants (*Solanum tuberosum* L.) affects tuber sprouting and flower formation with only little impact on carbohydrate metabolism. Plant Journal *12,* 323-333.

Echt, C.S., and Schwartz, D. (1981). Evidence for the inclusion of controlling elements within the structural gene at the waxy locus in maize. Genetics *99,* 275-284.

Eizirik, D.L., Colli, M.L., and Ortis, F. (2009). The role of inflammation in insulitis and beta-cell loss in type 1 diabetes. Nature Reviews Endocrinology *5,* 219-226.

Elias, J., Haas, W., Faherty, B., and Gygi, S. (2005). Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. Nature Methods *2,* 667-675.

Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J. Mol. Biol. *300,* 1005-1016.

Erban, T., Erbanova, M., Nesvorna, M., and Hubert, J. (2009). The importance of starch and sucrose digestion in nutritive biology of synanthropic acaridid mites: alpha-amylases and alpha-glucosidases are suitable targets for inhibitor-based strategies of mite control. Arch. Insect Biochem. Physiol. *71,* 139-158.

Esau, K. (1977). Anatomy of seed plants (London: John Wiley & Sons).

Fabijanski, S., Altosaar, I., Lauriere, M., Pernollet, J.C., and Mosse, J. (1985). Antigenic homologies between oat and wheat globulins. FEBS Lett. *182,* 465-469.

Fannon, J.E., Gray, J.A., Gunawan, N., Huber, K.C., and BeMiller, J.N. (2004). Heterogeneity of starch granules and the effect of granule channelization on starch modification. Cellulose *11,* 247-254.

Fannon, J.E., Gray, J.A., Gunawan, N., Huber, K.C., and BeMiller, J.N. (2003). The channels of starch granules. Food Science and Biotechnology *12,* 700-704.

Fasano, A. (2001). Celiac disease: The past, the present, the future. Pediatrics *107,* 768-770.

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

Fincher, G.B. (1989). Molecular and cellular biology associated with endosperm mobilization in germinating cereal-grains. Annu. Rev. Plant Physiol. Plant Mol. Biol. *40,* 305-346.

Fincher, G.B., and Stone, B.A. (1993). Physiology and biochemistry of germination in barley. In Barley: chemistry and technology, MacGregor, A. W., and Bhatty, R. S. eds., (St. Paul, MN: AACC, American Association of Cereal Chemists) pp. 247-295.

Fitzpatrick, T., Amrhein, N., and Macheroux, P. (2003). Characterization of YqjM, an old yellow enzyme homolog from*Bacillus subtilis*involved in the oxidative stress response. J. Biol. Chem. *278,* 19891-19897.

Floyd, S., and Friedman, W. (2000). Evolution of endosperm developmental patterns among basal flowering plants. Int. J. Plant Sci. *161,* S57-S81.

Frohlich, M.W., and Chase, M.W. (2007). After a dozen years of progress the origin of angiosperms is still a great mystery. Nature *450,* 1184-1189.

Fu, H.A., Subramanian, R.R., and Masters, S.C. (2000). 14-3-3 proteins: Structure, function, and regulation. Annu. Rev. Pharmacol. Toxicol. *40,* 617-647.

Galau, G., Hughes, D., and Dure, L. (1986). Abscisic-acid induction of cloned cotton late embryogenesis-abundant (LEA) messenger-RNAs. Plant Mol. Biol. *7,* 155-170.

Garay-Arroyo, A., Colmenero-Flores, J., Garciarrubio, A., and Covarrubias, A. (2000). Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit. J. Biol. Chem. *275,* 5668-5674.

Gatehouse, J.A., Lycett, G.W., Delauney, A.J., Croy, R.R.D., and Boulter, D. (1983). Sequence specificity of the post-translational proteolytic cleavage of vicilin, a seed storage protein of pea (*Pisum sativum* L). Biochem. J. *212,* 427-432.

Geddes, J., Eudes, F., Laroche, A., and Selinger, L.B. (2008). Differential expression of proteins in response to the interaction between the pathogen *Fusarium graminearum* and its host, *Hordeum vulgare*. Proteomics *8,* 545-554.

Gerrienne, P., Meyer-Berthaud, B., Fairon-Demaret, M., Streel, M., and Steemans, P. (2004). *Runcaria*, a middle Devonian seed plant precursor. Science *306,* 856-858.

Gil-Humanes, J., Piston, F., Tollefsen, S., Sollid, L.M., and Barro, F. (2010). Effective shutdown in the expression of celiac disease-related wheat gliadin T-cell epitopes by RNA interference. Proc. Natl. Acad. Sci. U. S. A. *107,* 17023-17028.

Goesaert, H., Elliott, G., Kroon, P., Gebruers, K., Courtin, C., Robben, J., Delcour, J., and Juge, N. (2004). Occurrence of proteinaceous endoxylanase inhibitors in cereals. Biochimica Et Biophysica Acta - Proteins and Proteomics *1696,* 193-202.

Gomez, L., Sanchezmonge, R., and Salcedo, G. (1988). A family of endosperm globulins encoded by genes located in group-1 chromosomes of wheat and related species. Molecular & General Genetics *214,* 541-546.

Goswami, R., and Kistler, H. (2004). Heading for disaster: *Fusarium graminearum* on cereal crops. Molecular Plant Pathology *5,* 515-525.

Goyal, K., Pinelli, C., Maslen, S., Rastogi, R., Stephens, E., and Tunnacliffe, A. (2005). Dehydration-regulated processing of late embryogenesis abundant protein in a desiccation-tolerant nematode. FEBS Lett. *579,* 4093-4098.

Graf, A., and Smith, A.M. (2011). Starch and the clock: the dark side of plant productivity. Trends Plant Sci. *16,* 169-175.

Greenwell, P., and Schofield, J.D. (1986). A starch granule protein associated with endosperm softness in wheat. Cereal Chem. *63,* 379-380.

Grimaud, F., Rogniaux, H., James, M.G., Myers, A.M., and Planchot, V. (2008). Proteome and phosphoproteome analysis of starch granule-associated proteins from normal maize and mutants affected in starch biosynthesis. J. Exp. Bot. *59,* 3395-3406.

Grover, A. (2012). Plant chitinases: genetic diversity and physiological roles. Crit. Rev. Plant Sci. *31,* 57-73.

Gu, Y., Coleman-Derr, D., Kong, X., and Anderson, O. (2004). Rapid genome evolution revealed by comparative sequence analysis of orthologous regions from four triticeae genomes. Plant Physiol. *135,* 459-470.

Gupta, R.S., Springston, E.E., Warrier, M.R., Smith, B., Kumar, R., Pongracic, J., and Holl, J.L. (2011). The prevalence, severity, and distribution of childhood food allergy in the United States. Pediatrics *128,* E9-E17.

Habib, H., and Fazili, K.M. (2007). Plant protease inhibitors: a defense strategy in plants. Biotech. Mol. Biol. Rev. *2,* 68-85.

Halford, N.G., Curtis, T.Y., Muttucumaru, N., Postles, J., and Mottram, D.S. (2011). Sugars in crop plants. Ann. Appl. Biol. *158,* 1-25.

Hanashiro, I., Itoh, K., Kuratomi, Y., Yamazaki, M., Igarashi, T., Matsugasako, J., and Takeda, Y. (2008). Granule-bound starch synthase I is responsible for biosynthesis of extra-long unit chains of amylopectin in rice. Plant and Cell Physiology *49,* 925-933.

Handoyo, T., Akagawa, M., Morita, N., Maeda, T., and Mitsunaga, T. (2008). Hypoallergenic characteristics of wheat flour produced by stepwise polishing. Int. J. Food Prop. *11,* 243-252.

Hanes, C.S. (1932). Studies on plant amylases. I. The effect of starch concentration upon the velocity of hydrolysis by the amylase of germinated barley. Biochem. J. *26,* 1406-1421.

Haq, S., Atif, S., and Khan, R. (2004). Protein proteinase inhibitor genes in combat against insects, pests, and pathogens: natural and engineered phytoprotection. Arch. Biochem. Biophys. *431,* 145-159.

Hara-Nishimura, I., Shimada, T., Hatano, K., Takeuchi, Y., and Nishimura, M. (1998). Transport of storage proteins to protein storage vacuoles is mediated by large precursor-accumulating vesicles. Plant Cell *10,* 825-836.

Hardham, A. (2005). *Phytophthora cinnamomi*. Molecular Plant Pathology *6,* 589-604.

Harris, L., Desjardins, A., Plattner, R., Nicholson, P., Butler, G., Young, J., Weston, G., Proctor, R., and Hohn, T. (1999). Possible role of trichothecene mycotoxins in virulence of *Fusarium graminearum* on maize. Plant Dis. *83,* 954-960.

Heck, G.R., Chamberlain, A.K., and Ho, T.H.D. (1993). Barley embryo globulin-1 gene, *beg1*- characterization of cDNA, chromosome mapping and regulation of expression. Molecular & General Genetics *239,* 209-218.

Helleboid, S., Bauw, G., Belingheri, L., Vasseur, J., and Hilbert, J.L. (1998). Extracellular beta-1,3-glucanases are induced during early somatic embryogenesis in *Cichorium*. Planta *205,* 56-63.

Herman, E., and Larkins, B. (1999). Protein storage bodies and vacuoles. Plant Cell *11,* 601-613.

Heyman, M., Abed, J., Lebreton, C., and Cerf-Bensussan, N. (2012). Intestinal permeability in coeliac disease: insight into mechanisms and relevance to pathogenesis. Gut *61,* 1355-1364.

Hibberd, J.M., Sheehy, J.E., and Langdale, J.A. (2008). Using C-4 photosynthesis to increase the yield of rice - rationale and feasibility. Curr. Opin. Plant Biol. *11,* 228-231.

Hincha, D.K., Meins, F., and Schmitt, J.M. (1997). Beta-1,3-glucanase is cryoprotective *in vitro* and is accumulated in leaves during cold acclimation. Plant Physiol. *114,* 1077-1083.

Hinton, D.M., and Pressey, R. (1980). Glucanases in fruits and vegetables. J. Am. Soc. Hort. Sci. *105,* 499-502.

Holbert, J., Burlison, W., Biggar, H., Koehler, B., Dungan, G., and Jenkins, M. (1923). Early vigor of maize plants and yield of grain as influenced by the corn root, stalk, and ear rot diseases. Journal of Agricultural Research *23,* 0583-0630.

Hong, C., and Moorman, G. (2005). Plant pathogens in irrigation water: challenges and opportunities. Crit. Rev. Plant Sci. *24,* 189-208.

Hornsten, L., Su, C., Osbourn, A., Garosi, P., Hellman, U., Wernstedt, C., and Oliw, E. (1999). Cloning of linoleate diol synthase reveals homology with prostaglandin H synthases. J. Biol. Chem. *274,* 28219-28224.

Huang, Q., Xie, X., Liang, G., Gong, F., Wang, Y., Wei, X., Wang, Q., Ji, Z., and Chen, Q. (2012). The GH18 family of chitinases: their domain architectures, functions and evolutions. Glycobiology *22,* 23-34.

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

Huber, S., Lottspeich, F., and Kamper, J. (2002). A gene that encodes a product with similarity to dioxygenases is highly expressed in teliospores of *Ustilago maydis*. Molecular Genetics and Genomics *267,* 757-771.

Hwang, C., Rhie, G., Oh, J., Huh, W., Yim, H., and Kang, S. (2002). Copper- and zinc-containing superoxide dismutase (Cu/ZnSOD) is required for the protection of *Candida albicans* against oxidative stresses and the expression of its full virulence. Microbiology *148,* 3705-3713.

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.

Jahromi, M.M., and Eisenbarth, G.S. (2007). Cellular and molecular pathogenesis of type 1A diabetes. Cellular and Molecular Life Sciences *64,* 865-872.

Jane, J.L., Kasemsuwan, T., Leas, S., Zobel, H., and Robyt, J.F. (1994). Anthology of starch granule morphology by scanning electron-microscopy. Starch-Starke *46,* 121-129.

Jellum, M. (1970). Plant introductions of maize as a source of oil with unusual fatty acid composition. J. Agric. Food Chem. *18,* 365-370.

Jenczmionka, N.J., and Schafer, W. (2005). The Gpmk1 MAP kinase of *Fusarium graminearum* regulates the induction of specific secreted enzymes. Curr. Genet. *47,* 29-36.

Jeon, J., Ryoo, N., Hahn, T., Walia, H., and Nakamura, Y. (2010). Starch biosynthesis in cereal endosperm. Plant Physiology and Biochemistry *48,* 383-392.

Jerkovic, A., Kriegel, A.M., Bradner, J.R., Atwell, B.J., Roberts, T.H., and Willows, R.D. (2010). Strategic distribution of protective proteins within bran layers of wheat protects the nutrient-rich endosperm. Plant Physiol. *152,* 1459-1470.

Jerneren, F., Sesma, A., Francheschetti, M., Hamberg, M., and Oliw, E.H. (2010). Gene deletion of 7,8-linoleate diol synthase of the rice blast fungus. Studies on pathogenicity, stereochemistry, and oxygenation mechanisms. J. Biol. Chem. *285,* 5308-5316.

Jitonnom, J., Lee, V.S., Nimmanpipug, P., Rowlands, H.A., and Mulholland, A.J. (2011). Quantum mechanics/molecular mechanics modeling of substrate-assisted catalysis in family 18 chitinases: conformational changes and the role of Asp142 in catalysis in ChiB. Biochemistry *50,* 4697-4711.

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

Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. *28,* 27-30.

Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M., and Hirakawa, M. (2010). KEGG for representation and analysis of molecular networks involving diseases and drugs. Nucleic Acids Res. *38,* D355-D360.

Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K.F., Itoh, M., Kawashima, S., Katayama, T., Araki, M., and Hirakawa, M. (2006). From genomics to chemical genomics: new developments in KEGG. Nucleic Acids Res. *34,* D354-D357.

Kang, Z., and Buchenauer, H. (2000). Ultrastructural and immunocytochemical investigation of pathogen development and host responses in resistant and susceptible wheat spikes infected by *Fusarium culmorum*. Physiol. Mol. Plant Pathol. *57,* 255-268.

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.

Kaur, B., Ariffin, F., Bhat, R., and Karim, A.A. (2012). Progress in starch modification in the last decade. Food Hydrocoll. *26,* 398-404.

Kaviani, B., Pourkhalili, S.T., Sajedi, R.H., and Mosadegh, B. (2011). Salt treatment can change composition of glycinin and beta-conglycinin proteins in soybean seed. Plant Omics *4,* 228-235.

Kazan, K., Gardiner, D.M., and Manners, J.M. (2012). On the trail of a cereal killer: recent advances in *Fusarium graminearum* pathogenomics and host resistance. Molecular Plant Pathology *13,* 399-413.

Keeling, P.L., and Myers, A.M. (2010). Biochemistry and genetics of starch synthesis. Annual Review of Food Science and Technology *1,* 271-303.

Kelleher, N.L., Lin, H.Y., Valaskovic, G.A., Aaserud, D.J., Fridriksson, E.K., and McLafferty, F.W. (1999). Top down versus bottom up protein characterization by tandem high-resolution mass spectrometry. J. Am. Chem. Soc. *121,* 806-812.

Khavkin, E.E., Misharin, S.I., Markov, Y.Y., and Peshkova, A.A. (1978). Identification of embryonal antigens of maize: globulins as primary reserve proteins of embryo. Planta *143,* 11-20.

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

Ko, T.P., Day, J., and McPherson, A. (2000). The refined structure of canavalin from jack bean in two crystal forms at 2.1 and 2.0 Ångstrom resolution. Acta Crystallographica Section D-Biological Crystallography *56,* 411-420.

Koehler, B. (1942). Natural mode of entrance of fungi into corn ears and some symptoms that indicate infection. Journal of Agricultural Research *64,* 0421-0442.

Koziol, A.G., Borza, T., Ishida, K., Keeling, P., Lee, R.W., and Durnford, D.G. (2007). Tracing the evolution of the light-harvesting antennae in chlorophyll *a*/*b*- containing organisms. Plant Physiol. *143,* 1802-1816.

Koziol, A.G., Marquez, B.K., Huebsch, M.P., Smith, J.C., and Altosaar, I. (2012). Commercially produced rice and maize starches contain non-host proteins, as shown by mass spectrometry. Cereal Chemistry *89,* 262-267.

Koziol, A.G., Marquez, B.K., Huebsch, M.P., Smith, J.C., and Altosaar, I. (2012). The starch granule associated proteomes of commercially purified starch reference materials from rice and maize. Journal of Proteomics *75,* 993-1003.

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

Kriz, A.L. (1999). 7S globulins of cereal. In Seed Proteins, Shewry, P. R., and Casey, R. eds., (Dordrecht: Kluwer Academic Publishers) pp. 477-498.

Kriz, A.L. (1989). Characterization of embryo globulins encoded by the maize *glb* genes. Biochem. Genet. *27,* 239-251.

Laemmli, U.K. (1970). Cleavage of structural proteins during assembly of head of bacteriophage-T4. Nature *227,* 680-685.

Lammers, K.M., Lu, R., Brownley, J., Lu, B., Gerard, C., Thomas, K., Rallabhandi, P., Shea-Donohue, T., Tamiz, A., Alkan, S., *et al.* (2008). Gliadin induces an increase in intestinal permeability and zonulin release by binding to the chemokine receptor CXCR3. Gastroenterology *135,* 194-204.

Larre, C., Lupi, R., Gombaud, G., Brossard, C., Branlard, G., Moneret-Vautrin, D.A., Rogniaux, H., and Denery-Papini, S. (2011). Assessment of allergenicity of diploid and hexaploid wheat genotypes: Identification of allergens in the albumin/globulin fraction. Journal of Proteomics *74,* 1279-1289.

Larre, C., Penninck, S., Bouchet, B., Lollier, V., Tranquet, O., Denery-Papini, S., Guillon, F., and Rogniaux, H. (2010). *Brachypodium distachyon*grain: identification and subcellular localization of storage proteins. J. Exp. Bot. *61,* 1771-1783.

Lauriere, C., Lauriere, M., and Daussant, J. (1986). Immunohistochemical localization of beta-amylase in resting barley-seeds. Physiol. Plantarum *67,* 383-388.

Lawrence, M.C., Izard, T., Beuchat, M., Blagrove, R.J., and Colman, P.M. (1994). Structure of phaseolin at 2·2 Å resolution: implications for a common vicilin/legumin structure and the genetic-engineering of seed storage proteins. J. Mol. Biol. *238,* 748-776.

Lay, F.T., and Anderson, M.A. (2005). Defensins - components of the innate immune system in plants. Curr. Protein Peptide Sci. *6,* 85-101.

Lee, J., Godon, C., Lagniel, G., Spector, D., Garin, J., Labarre, J., and Toledano, M. (1999). Yap1 and Skn7 control two specialized oxidative stress response regulons in yeast. J. Biol. Chem. *274,* 16040-16046.

Leubner-Metzger, G., and Meins, F. (1999). Functions and regulation of plant β-1,3-glucanases (PR-2). In Pathogenesis-related proteins in plants, Datta, S. K., and Muthukrishnan, S. eds., (Boca Raton, Florida: CRC Press LLC) pp. 49-76.

Lindfors, K., Rauhavirta, T., Stenman, S., Maki, M., and Kaukinen, K. (2012). *In vitro* models for gluten toxicity: relevance for celiac disease pathogenesis and development of novel treatment options. Exp. Biol. Med. *237,*

Linkies, A., Graeber, K., Knight, C., and Leubner-Metzger, G. (2010). The evolution of seeds. New Phytol. *186,* 817-831.

Logrieco, A., Mule, G., Moretti, A., and Bottalico, A. (2002). Toxigenic *Fusarium* species and mycotoxins associated with maize ear rot in Europe. Eur. J. Plant Pathol. *108,* 597-609.

Loit, E., Melnyk, C.W., MacFarlane, A.J., Scott, F.W., and Altosaar, I. (2009). Identification of three wheat globulin genes by screening a *Triticum aestivum* BAC genomic library with cDNA from a diabetes-associated globulin. BMC Plant Biology *9,* 93.

Lopes, M.A., and Larkins, B.A. (1993). Endosperm origin, development, and function. Plant Cell *5,* 1383-1399.

Lowy, G.D.A., Sargeant, J.G., and Schofield, J.D. (1981). Wheat-starch granule protein - the isolation and characterization of a salt-extractable protein from starch granules. J. Sci. Food Agric. *32,* 371-377.

MacFarlane, A.J., Burghardt, K.M., Kelly, J., Simell, T., Simell, O., Altosaar, I., and Scott, F.W. (2003). A type 1 diabetes-related protein from wheat (*Triticum aestivum*) - cDNA clone of a wheat storage globulin, Glb1, linked to islet damage. J. Biol. Chem. *278,* 54-63.

MacFarlane, A.J., Strom, A., and Scott, F.W. (2009). Epigenetics: deciphering how environmental factors may modify autoimmune type 1 diabetes. Mammalian Genome *20,* 624-632.

Maurano, F., Mazzarella, G., Luongo, D., Stefanile, R., D'Arienzo, R., Rossi, M., Auricchio, S., and Troncone, R. (2005). Small intestinal enteropathy in non-obese diabetic mice fed a diet containing wheat. Diabetologia *48,* 931-937.

Mayer, A., Staples, R., and Gil-ad, N. (2001). Mechanisms of survival of necrotrophic fungal plant pathogens in hosts expressing the hypersensitive response. Phytochemistry *58,* 33-41.

McCallum, B.D., and DePauw, R.M. (2008). A review of wheat cultivars grown in the Canadian prairies. Can. J. Plant Sci. *88,* 649-677.

Mechin, V., Thevenot, C., Le Guilloux, M., Prioul, J.L., and Damerval, C. (2007). Developmental analysis of maize endosperm proteome suggests a pivotal role for pyruvate orthophosphate dikinase. Plant Physiol. *143,* 1203-1219.

Mesterhazy, A., Lemmens, M., and Reid, L.M. (2012). Breeding for resistance to ear rots caused by*Fusarium*spp. in maize - a review. Plant Breeding *131,* 1-19.

Miller, J., Young, J., and Sampson, D. (1985). Deoxynivalenol and *Fusarium* head blight resistance in spring cereals. Phytopathologische Zeitschrift-Journal of Phytopathology *113,* 359-367.

Miller, S.S., Reid, L.M., and Harris, L.J. (2007). Colonization of maize silks by *Fusarium graminearum*, the causative organism of gibberella ear rot. Canadian Journal of Botany *85,* 369-376.

Mohammadi, M., Anoop, V., Gleddie, S., and Harris, L.J. (2011). Proteomic profiling of two maize inbreds during early gibberella ear rot infection. Proteomics *11,* 3675-3684.

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

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

Morrison, D.K. (2009). The 14-3-3 proteins: integrators of diverse signaling cues that impact cell fate and cancer development. Trends Cell Biol. *19,* 16-23.

Morrison, W.R., Tester, R.F., and Gidley, M.J. (1994). Properties of damaged starch granules .2. Crystallinity, molecular order and gelatinization of ball-milled starches. J. Cereal Sci. *19,* 209-217.

Moughan, P.J., Fuller, M.F., Han, K., Kies, A.K., and Miner-Williams, W. (2007). Food-derived bioactive peptides influence gut function. Int. J. Sport Nutr. Exerc. Metab. *17,* S5-S22.

Munoz, I., Ubhayasekera, W., Henriksson, H., Szabo, I., Pettersson, G., Johansson, G., Mowbray, S., and Stahlberg, J. (2001). Family 7 cellobiohydrolases from *Phanerochaete chrysosporium*: crystal structure of the catalytic module of Cel7D (CBH58) at 1.32 Ångstrom resolution and homology models of the isozymes. J. Mol. Biol. *314,* 1097-1111.

Natesan, S.K.A., Sullivan, J.A., and Gray, J.C. (2005). Stromules: a characteristic cell-specific feature of plastid morphology. J. Exp. Bot. *56,* 787-797.

Nelson Jr., R.A., and Lebrun, J. (1956). The requirement for antibody and complement for in vitro phagocytosis of starch granules. J. Hyg. *54,* 8-19.

Nielsen, H., Engelbrecht, J., Brunak, S., and vonHeijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. *10,* 1-6.

Olsen, O. (2004). Nuclear endosperm development in cereals and *Arabidopsis thaliana*. Plant Cell *16,* S214-S227.

Onkamo, P., Vaananen, S., Karvonen, M., and Tuomilehto, J. (1999). Worldwide increase in incidence of Type I diabetes - the analysis of the data on published incidence trends. Diabetologia *42,* 1395-1403.

Osborne, T. (1908). Our present knowledge of plant proteins. Science *28,* 417-427.

Osborne, T.B., and Campbell, G.F. (1898). Proteids of the pea. Journal of the American Chemical Society *20,* 348-362.

Osbourn, A.E. (1996). Preformed antimicrobial compounds and plant defense against fungal attack. Plant Cell *8,* 1821-1831.

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.

Passardi, F., Cosio, C., Penel, C., and Dunand, C. (2005). Peroxidases have more functions than a Swiss army knife. Plant Cell Rep. *24,* 255-265.

Patron, N.J., and Keeling, P.J. (2005). Common evolutionary origin of starch biosynthetic enzymes in green and red algae. J. Phycol. *41,* 1131-1141.

Payan, F., Flatman, R., Porciero, S., Williamson, G., Juge, N., and Roussel, A. (2003). Structural analysis of xylanase inhibitor protein I (XIP-I), a proteinaceous xylanase inhibitor from wheat (*Triticum aestivum*, var. Soisson). Biochem. J. *372,* 399-405.

Payan, F., Leone, P., Porciero, S., Furniss, C., Tahir, T., Williamson, G., Durand, A., Manzanares, P., Gilbert, H., Juge, N., and Roussel, A. (2004). The dual nature of the wheat xylanase protein inhibitor XIP-I - Structural basis for the inhibition of family 10 and family 11 xylanases. J. Biol. Chem. *279,* 36029-36037.

Perez, S., and Bertoft, E. (2010). The molecular structures of starch components and their contribution to the architecture of starch granules: A comprehensive review. Starch-Starke *62,* 389-420.

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

Pestka, J.J., and Bondy, G.S. (1994). Immunotoxic effects of mycotoxins. In Mycotoxins in Grain: Compounds other than Aflatoxin, Miller, J. D., and Trenholm, H. L. eds., (St. Paul, MN: The American Phytopathological Society) pp. 339-358.

Pfluger, P.T., Schriever, S.C., and Tschoep, M.H. (2012). Nutropioids, hedonism in the gut? Cell Metabolism *16,* 137-139.

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

Pitak, N., and Rakshit, S.K. (2011). Physical and antimicrobial properties of banana flour/chitosan biodegradable and self sealing films used for preserving fresh-cut vegetables. LWT-Food Science and Technology *44,* 2310-2315.

Pomeranz, Y., Carvajal, M., Hoseney, R., and Ward, A. (1970). Wheat germ in breadmaking 1. Composition of germ lipids and germ protein fractions. Cereal Chem. *47,* 373-380.

Pordesimo, L., Hames, B., Sokhansanj, S., and Edens, W. (2005). Variation in corn stover composition and energy content with crop maturity. Biomass & Bioenergy *28,* 366-374.

Prade, R. (1996). Xylanases: from biology to biotechnology. Biotechnology and Genetic Engineering Reviews *13,* 101-131.

Puchongkavarin, H., Varavinit, S., and Bergthaller, W. (2005). Comparative study of pilot scale rice starch production by an alkaline and an enzymatic process. Starch-Starke *57,* 134-144.

Qi, X., Tester, R.F., Snape, C.E., and Ansell, R. (2003). Molecular basis of the gelatinisation and swelling characteristics of waxy rice starches grown in the same location during the same season. J. Cereal Sci. *37,* 363-376.

Rahman, S., Bird, A., Regina, A., Li, Z., Ral, J.P., McMaugh, S., Topping, D., and Morell, M. (2007). Resistant starch in cereals: Exploiting genetic engineering and genetic variation. J. Cereal Sci. *46,* 251-260.

Rahman, S., Li, Z., Batey, I., Cochrane, M.P., Appels, R., and Morell, M. (2000). Genetic alteration of starch functionality in wheat. J. Cereal Sci. *31,* 91-110.

Ramesh, S. (2008). Food allergy overview in children. Clin. Rev. Allergy Immunol. *34,* 217-230.

Rayhan, M.U., Van, K., Kim, D.H., Il Kim, S., Kim, M.Y., Lee, Y., and Lee, S. (2011). Identification of Gy4 nulls and development of multiplex PCR-based co-dominant marker for Gy4 and alpha' subunit of beta-conglycinin in soybean. Genes & Genomics *33,* 383-390.

Reid, L.M., Hamilton, R.I., and Mather, D.E. (1996). Screening Maize for Resistance to Gibberella Ear Rot (Ottawa, ON AAFC Technical Bulletin 1996-5E: Eastern Cereal and Oilseed Research Centre).

Reid, L., Hamilton, R., and Schaafsma, A. (1997). Automatic pipettor for inoculation of maize kernels for evaluation of resistance to *Fusarium graminearum*. Canadian Journal of Plant Pathology-Revue Canadienne De Phytopathologie *19,* 185-187.

Reid, L., Mather, D., Bolton, A., and Hamilton, R. (1994). Evidence for a gene for silk resistance to *Fusarium graminearum* Schw. ear rot of maize. J. Hered. *85,* 118-121.

Reid, L., Mather, D., Hamilton, R., and Bolton, A. (1992). Diallel analysis of resistance in maize to *Fusarium graminearum* infection via the silk. Can. J. Plant Sci. *72,* 915-923.

Reid, L., Mather, D., Hamilton, R., and Bolton, A. (1992). Genotypic differences in the resistance of maize silk to *Fusarium graminearum*. Canadian Journal of Plant Pathology-Revue Canadienne De Phytopathologie *14,* 211-214.

Reignault, P., Valette-Collet, O., and Boccara, M. (2008). The importance of fungal pectinolytic enzymes in plant invasion, host adaptability and symptom type. Eur. J. Plant Pathol. *120,* 1-11.

Rivin, C., and Grudt, T. (1991). Abscisic-acid and the developmental regulation of embryo storage proteins in maize. Plant Physiol. *95,* 358-365.

Robert, L.S., Adeli, K., and Altosaar, I. (1985). Homology among 3S and 7S globulins from cereals and pea. Plant Physiol. *78,* 812-816.

Robert, L.S., Nozzolillo, C., and Altosaar, I. (1985). Homology between legumin-like polypeptides from cereals and pea. Biochem. J. *226,* 847-852.

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

Roper, H. (2002). Renewable raw materials in Europe - Industrial utilisation of starch and sugar. Starch-Starke *54,* 89-99.

Ryan, C. (1990). Protease inhibitors in plants - genes for improving defenses against insects and pathogens. Annu. Rev. Phytopathol. *28,* 425-449.

Samaroo, D., Dickerson, F., Kasarda, D.D., Green, P.H.R., Briani, C., Yolken, R.H., and Alaedini, A. (2010). Novel immune response to gluten in individuals with schizophrenia. Schizophr. Res. *118,* 248-255.

Sanchez, J.C., Rouge, V., Pisteur, M., Ravier, F., Tonella, L., Moosmayer, M., Wilkins, M.R., and Hochstrasser, D.F. (1997). Improved and simplified in-gel sample application using reswelling of dry immobilized pH gradients. Electrophoresis *18,* 324-327.

Sancho, A.I., Gillabert, M., Tapp, H., Shewry, P.R., Skeggs, P.K., and Mills, E.N.C. (2008). Effect of environmental stress during grain filling on the soluble proteome of wheat (*Triticum aestivum*) dough liquor. J. Agric. Food Chem. *56,* 5386-5393.

Santos, P., Fortunato, A., Ribeiro, A., and Pawlowski, K. (2008). Chitinases in root nodules. Plant Biotechnology *25,* 299-307.

Schwartz, D. (1979). Analysis of the size alleles of the *pro* gene in maize - evidence for a mutant protein processor. Molecular & General Genetics *174,* 233-240.

Schwarz, P.B., Schwarz, J.G., Zhou, A., Prom, L.K., and Steffenson, B.J. (2001). Effect of *Fusarium graminearum* and *F-poae* infection on barley and malt quality. Monatsschrift Fur Brauwissenschaft *54,* 55-63.

Scott, F.W. (2007). Diabetogenic epitopes. US Patent 20070185021.

Sehnke, P.C., Chung, H.J., Wu, K., and Ferl, R.J. (2001). Regulation of starch accumulation by granule-associated plant 14-3-3 proteins. Proc. Natl. Acad. Sci. U. S. A. *98,* 765-770.

Sels, J., Mathys, J., De Coninck, B.M.A., Cammue, B.P.A., and De Bolle, M.F.C. (2008). Plant pathogenesis-related (PR) proteins: a focus on PR peptides. Plant Physiol. Biochem. *46,* 941-950.

Shan, L., Molberg, O., Parrot, I., Hausch, F., Filiz, F., Gray, G., Sollid, L., and Khosla, C. (2002). Structural basis for gluten intolerance in Celiac sprue. Science *297,* 2275-2279.

Sharma, G.M., Mundoma, C., Seavy, M., Roux, K.H., and Sathe, S.K. (2010). Purification and biochemical characterization of Brazil nut (*Bertholletia excelsa* L.) seed storage proteins. J. Agric. Food Chem. *58,* 5714-5723.

Shaw, M.M., and Riederer, B.M. (2003). Sample preparation for two-dimensional gel electrophoresis. Proteomics *3,* 1408-1417.

Shewry, P.R., and Halford, N.G. (2002). Cereal seed storage proteins: structures, properties and role in grain utilization. J. Exp. Bot. *53,* 947-958.

Shewry, P., and Tatham, A. (1990). The prolamin storage proteins of cereal seeds - structure and evolution. Biochem. J. *267,* 1-12.

Shi, M., and Gao, Q. (2011). Physicochemical properties, structure and in vitro digestion of resistant starch from waxy rice starch. Carbohydr. Polym. *84,* 1151-1157.

Shin, K., Kamal, A.H.M., Cho, K., Choi, J., Yu, J., Paek, N., Lee, Y.W., Lee, J.K., Park, J., Kim, H., and Woo, S.H. (2011). Defense proteins are induced in wheat spikes exposed to *Fusarium graminearum*. Plant Omics *4,* 270-277.

Shorrosh, B., Wen, L., Zen, K., Huang, J., Pan, J., Hermodson, M., Tanaka, K., Muthukrishnan, S., and Reeck, G. (1992). A novel cereal storage protein - molecular-genetics of the 19 kDa globulin of rice. Plant Mol. Biol. *18,* 151-154.

Shu, X., Jiao, G., Fitzgerald, M.A., Yang, C., Shu, Q., and Wu, D. (2006). Starch structure and digestibility of rice high in resistant starch. Starch-Starke *58,* 411-417.

Singh, J., Blundell, M., Tanner, G., and Skerritt, J.H. (2001). Albumin and globulin proteins of wheat flour: immunological and N-terminal sequence characterisation. J. Cereal Sci. *34,* 85-103.

Singh, N.K., Shepherd, K.W., Langridge, P., and Gruen, L.C. (1991). Purification and biochemical-characterization of triticin, a legumin-like protein in wheat endosperm. J. Cereal Sci. *13,* 207-219.

Singh, V., Haken, A., Dowd, M., Niu, Y., Zou, S., and Eckhoff, S. (1999). Batch steeping of corn: Effects of adding lactic acid and sulfur dioxide at different times on starch yields, protein contents, and starch pasting properties. Cereal Chem. *76,* 600-605.

Skerritt, J.H., Frend, A.J., Robson, L.G., and Greenwell, P. (1990). Immunological homologies between wheat gluten and starch granule proteins. J. Cereal Sci. *12,* 123-136.

Smith, A.M., Zeeman, S.C., and Smith, S.M. (2005). Starch degradation. Annual Review of Plant Biology *56,* 73-98.

Smith, P.K., Krohn, R.I., Hermanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., Fujimoto, E.K., Goeke, N.M., Olson, B.J., and Klenk, D.C. (1985). Measurement of protein using bicinchoninic acid. Anal. Biochem. *150,* 76-85.

Smith, R.D., Loo, J.A., Edmonds, C.G., Barinaga, C.J., and Udseth, H.R. (1990). New developments in biochemical mass-spectrometry - electrospray ionization. Anal. Chem. *62,* 882-899.

Soanes, D.M., Alam, I., Cornell, M., Wong, H.M., Hedeler, C., Paton, N.W., Rattray, M., Hubbard, S.J., Oliver, S.G., and Talbot, N.J. (2008). Comparative genome analysis of filamentous fungi reveals gene family expansions associated with fungal pathogenesis. PloS One *3,* e2300.

Spencer, D., Chandler, P.M., Higgins, T.J.V., Inglis, A.S., and Rubira, M. (1983). Sequence interrelationships of the subunits of vicilin from pea seeds. Plant Mol. Biol. *2,* 259-267.

Stanojevic, S.P., Barac, M.B., Pesic, M.B., and Vucelic-Radovic, B.V. (2011). Assessment of soy genotype and processing method on quality of soybean tofu. J. Agric. Food Chem. *59,* 7368-7376.

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

Steponkus, P., Uemura, M., Joseph, R., Gilmour, S., and Thomashow, M. (1998). Mode of action of the *COR15a* gene on the freezing tolerance of *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. U. S. A. *95,* 14570-14575.

Sturm, A., Vankuik, J.A., Vliegenthart, J.F.G., and Chrispeels, M.J. (1987). Structure, position, and biosynthesis of the high mannose and the complex oligosaccharide side-chains of the bean storage protein phaseolin. J. Biol. Chem. *262,* 13392-13403.

Sulpice, R., Pyl, E., Ishihara, H., Trenkamp, S., Steinfath, M., Witucka-Wall, H., Gibon, Y., Usadel, B., Poree, F., Piques, M.C., *et al.* (2009). Starch as a major integrator in the regulation of plant growth. Proc. Natl. Acad. Sci. U. S. A. *106,* 10348-10353.

Sun, J.L., Nakagawa, H., Karita, S., Ohmiya, K., and Hattori, T. (1996). Rice embryo globulins: amino-terminal amino acid sequences, cDNA cloning and expression. Plant and Cell Physiology *37,* 612-620.

Sun, X.S. (2005). Isolation and Processing of Plant Materials. In Bio-based Polymers and Composites, Wool, R., and Sun, X. S. eds., (San Diego, CA: Elsevier Academic Press) pp. 33-55.

Sun, Z., and Henson, C.A. (1990). Degradation of native starch granules by barley alpha-glucosidases. Plant Physiol. *94,* 320-327.

Sutton, J.C. (1982). Epidemiology of wheat head blight and maize ear rot caused by *Fusarium graminearum*. Canadian Journal of Plant Pathology *4,* 195-209.

Takats, Z., Wiseman, J.M., Gologan, B., and Cooks, R.G. (2004). Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. Science *306,* 471-473.

Takeda, K., and Heta, H. (1989). Establishing the testing method and a search for the resistant varieties to Fusarium head blight in barley. Japanese Journal of Breeding *39,* 203-216.

Tan, S.Y. (2003). Resistant Rice Starch Development. M.Sc. Thesis, Louisiana State University.

Tanaka, K., Sugimoto, T., Ogawa, M., and Kasai, Z. (1980). Isolation and characterization of 2-types of protein bodies in the rice endosperm. Agric. Biol. Chem. *44,* 1633-1639.

Taplin, C.E., Mojibian, M., Simpson, M., Taki, I., Liu, E., Hoffenberg, E.J., Norris, J.M., Scott, F.W., and Rewers, M. (2011). Antibodies to the wheat storage globulin Glo-3A in children before and at diagnosis of celiac disease. J. Pediatr. Gastroenterol. Nutr. *52,* 21-25.

Tasic, M.B., and Veljkovic, V.B. (2011). Simulation of fuel ethanol production from potato tubers. Comput. Chem. Eng. *35,* 2284-2293.

Tasleem-Tahir, A., Nadaud, I., Girousse, C., Martre, P., Marion, D., and Branlard, G. (2011). Proteomic analysis of peripheral layers during wheat (*Triticum aestivum* L.) grain development. Proteomics *11,* 371-379.

Tatham, A.S., and Shewry, P.R. (2008). Allergens to wheat and related cereals. Clin. Exp. Allergy *38,* 1712-1726.

Taylor, E.L., and Taylor, T.N. (2009). Seed ferns from the late paleozoic and mesozoic: any angiosperm ancestors lurking there? Am. J. Bot. *96,* 237-251.

Taylor, R.D., Saparno, A., Blackwell, B., Anoop, V., Gleddie, S., Tinker, N.A., and Harris, L.J. (2008). Proteomic analyses of *Fusarium graminearum* grown under mycotoxin-inducing conditions. Proteomics *8,* 2256-2265.

Tester, R.F., Karkalas, J., and Qi, X. (2004). Starch structure and digestibility enzyme-substrate relationship. Worlds Poultry Science Journal *60,* 186-195.

Tester, R.F., Yousuf, R., Kettlitz, B., and Röper, H. (2007). Use of commercial protease preparations to reduce protein and lipid content of maize starch. Food Chem. *105,* 926-931.

Tetlow, I.J. (2011). Starch biosynthesis in developing seeds. Seed Science Research *21,* 5-32.

The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature *408,* 796-815.

Thijssen, M.H., Spoelstra, P., and Emons, A.M.C. (1996). Immunodetection and immunolocalization of globulin storage proteins during zygotic and somatic embryo development in *Zea mays*. Physiol. Plantarum *98,* 539-549.

Tjon, J.M., van Bergen, J., and Koning, F. (2010). Celiac disease: how complicated can it get? Immunogenetics *62,* 641-651.

Todd, J.A. (2010). Etiology of type 1 diabetes. Immunity *32,* 457-467.

Todd, J.A. (1991). A protective role of the environment in the development of type-1 diabetes. Diabetic Med. *8,* 906-910.

Tolleter, D., Hincha, D.K., and Macherel, D. (2010). A mitochondrial late embryogenesis abundant protein stabilizes model membranes in the dry state. Biochimica et Biophysica Acta - Biomembranes *1798,* 1926-1933.

Tolleter, D., Jaquinod, M., Mangavel, C., Passirani, C., Saulnier, P., Manon, S., Teyssier, E., Payet, N., Avelange-Macherel, M., and Macherel, D. (2007). Structure and function of a mitochondrial late embryogenesis abundant protein are revealed by desiccation. Plant Cell *19,* 1580-1589.

Trail, F. (2009). For blighted waves of grain: *Fusarium graminearum* in the postgenomics era. Plant Physiol. *149,* 103-110.

Trotter, E.W., Collinson, E.J., Dawes, I.W., and Grant, C.M. (2006). Old yellow enzymes protect against acrolein toxicity in the yeast *Saccharomyces cerevisiae*. Appl. Environ. Microbiol. *72,* 4885-4892.

Tunnacliffe, A., and Wise, M.J. (2007). The continuing conundrum of the LEA proteins. Naturwissenschaften *94,* 791-812.

U.S. Food and Drug Administration. Food Labeling: Health Claims; Oats and Coronary Heart Disease; Final Rule, 62 Federal Register (23 January, 1997), pp. 3583-3601.

Ubhayasekera, W. (2011). Structure and function of chitinases from glycoside hydrolase family 19. Polym. Int. *60,* 890-896.

Ubhayasekera, W., Rawat, R., Ho, S.W.T., Wiweger, M., Von Arnold, S., Chye, M., and Mowbray, S.L. (2009). The first crystal structures of a family 19 class IV chitinase: the enzyme from Norway spruce. Plant Mol. Biol. *71,* 277-289.

United States Department of Agriculture. Grains > Wheat consumption by country. www.NationMaster.com/graph/agr_gra_whe_con-agriculture-grains-wheat-consumption.*Last Accessed: September 8, 2012.*

Uthumporn, U., Zaidul, I.S.M., and Karim, A.A. (2010). Hydrolysis of granular starch at sub-gelatinization temperature using a mixture of amylolytic enzymes. Food Bioprod. Process. *88,* 47-54.

van Herpen, T.W.J.M., Riley, M., Sparks, C., Jones, H.D., Gritsch, C., Dekking, E.H., Hamer, R.J., Bosch, D., Salentijn, E.M.J., Smulders, M.J.M., Shewry, P.R., and Gilissen, L.J.W.J. (2008). Detailed analysis of the expression of an alpha-gliadin promoter and the deposition of alpha-gliadin protein during wheat grain development. Annals of Botany *102,* 331-342.

van Loon, L.C., Rep, M., and Pieterse, C.M.J. (2006). Significance of inducible defense-related proteins in infected plants. Annu. Rev. Phytopathol. *44,* 135-162.

van Loon, L.C., and 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.

Vasilescu, J., Smith, J.C., Ethier, M., and Figeys, D. (2005). Proteomic analysis of ubiquitinated proteins from human MCF-7 breast cancer cells by immunoaffinity purification and mass spectrometry. Journal of Proteome Research *4,* 2192-2200.

Vensel, W.H., Tanaka, C.K., Cai, N., Wong, J.H., Buchanan, B.B., and Hurkman, W.J. (2005). Developmental changes in the metabolic protein profiles of wheat endosperm. Proteomics *5,* 1594-1611.

Vesonder, R., Ellis, J., and Rohwedder, W. (1981). Elaboration of vomitoxin and zearalenone by *Fusarium*-isolates and the biological-activity of *Fusarium*-produced toxins. Appl. Environ. Microbiol. *42,* 1132-1134.

Walker-Simmons, M. (1987). ABA levels and sensitivity in developing wheat embryos of sprouting resistant and susceptible cultivars. Plant Physiol. *84,* 61-66.

Wall, M.L., Wheeler, H.L., Huebsch, M.P., Smith, J.C., Figeys, D., and 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.

Wall, M.L., Wheeler, H.L., Smith, J., Figeys, D., and 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.

Walsh, T., Morgan, A., and Hey, T. (1991). Characterization and molecular-cloning of a proenzyme form of a ribosome-inactivating protein from maize - novel mechanism of proenzyme activation by proteolytic removal of a 2.8-kiloDalton internal peptide segment. J. Biol. Chem. *266,* 23422-23427.

Walter, S., Nicholson, P., and Doohan, F.M. (2010). Action and reaction of host and pathogen during *Fusarium* head blight disease. New Phytol. *185,* 54-66.

Wang, H., Wang, Z., Wang, F., Gu, Y., and Liu, Z. (2012). Development of basal endosperm transfer cells in *Sorghum bicolor* (L.) Moench and its relationship with caryopsis growth. Protoplasma *249,* 309-321.

Wang, L.F., and Wang, Y.J. (2001). Comparison of protease digestion at neutral pH with alkaline steeping method for rice starch isolation. Cereal Chem. *78,* 690-692.

Wanjiru, W.M., Kang, Z.S., and 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.

Warburg, O., and Christian, W. (1932). On a new oxidation enzyme and its absorption spectrum. Biochem. Z. *254,* 438-458.

White, B., and Yanofsky, C. (1993). Structural characterization and expression analysis of the *Neurospora* conidiation gene *con*-6. Dev. Biol. *160,* 254-264.

Williams, R., and Bruce, N. (2002). 'New uses for an old enzyme' - the old yellow enzyme family of flavoenzymes. Microbiology *148,* 1607-1614.

Williamson, G., Belshaw, N.J., Self, D.J., Noel, T.R., Ring, S.G., Cairns, P., Morris, V.J., Clark, S.A., and Parker, M.L. (1992). Hydrolysis of A-type and B-type crystalline polymorphs of starch by alpha-amylase, beta-amylase and glucoamylase-1. Carbohydr. Polym. *18,* 179-187.

Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., and Mann, M. (1996). Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature *379,* 466-469.

Woo, Y., Hu, D., Larkins, B., and Jung, R. (2001). Genomics analysis of genes expressed in maize endosperm identifies novel seed proteins and clarifies patterns of zein gene expression. Plant Cell *13,* 2297-2317.

Yandeau-Nelson, M.D., Laurens, L., Shi, Z., Xia, H., Smith, A.M., and Guiltinan, M.J. (2011). Starch-branching enzyme IIa is required for proper diurnal cycling of starch in leaves of maize. Plant Physiol. *156,* 479-490.

Yazdanbakhsh, M., Kremsner, P., and van Ree, R. (2002). Immunology - Allergy, parasites, and the hygiene hypothesis. Science *296,* 490-494.

Yoshizawa, T., Shimizu, T., Hirano, H., Sato, M., and Hashimoto, H. (2002). Structural basis for inhibition of xyloglucan-specific endo-beta-1,4-glucanase (XEG) by XEG-protein inhibitor. J Biol Chem *287,* 18710-18716.

Young, T., and Gallie, D. (2000). Programmed cell death during endosperm development. Plant Mol. Biol. *44,* 283-301.

Yu, S.K., Blennow, A., Bojko, M., Madsen, F., Olsen, C.E., and Engelsen, S.B. (2002). Physico-chemical characterization of floridean starch of red algae. Starch-Starke *54,* 66-74.

Yupsanis, T., Burgess, S.R., Jackson, P.J., and Shewry, P.R. (1990). Characterization of the major protein component from aleurone cells of barley (*Hordeum vulgare* L.). J. Exp. Bot. *41,* 385-392.

Zeeman, S.C., Kossmann, J., and Smith, A.M. (2010). Starch: its metabolism, evolution, and biotechnological modification in plants. Annual Review of Plant Biology *61,* 209-234.

Zhang, H.N., Sreenivasulu, N., Weschke, W., Stein, N., Rudd, S., Radchuk, V., Potokina, E., Scholz, U., Schweizer, P., Zierold, U., *et al.* (2004). Large-scale analysis of the barley transcriptome based on expressed sequence tags. Plant Journal *40,* 276-290.

Zhang, Z.T., Niu, Y.X., Eckhoff, S.R., and Feng, H. (2005). Sonication enhanced cornstarch separation. Starch-Starke *57,* 240-245.

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.

Zhu, K., Zhou, H., and Qian, H. (2006). Proteins extracted from defatted wheat germ: nutritional and structural properties. Cereal Chem. *83, 69-75.*

Zioudrou, C., Streaty, R.A., and Klee, W.A. (1979). Opioid peptides derived from food proteins - exorphins. J. Biol. Chem. *254,* 2446-2449.

Appendix A: Supplemental Tables and Figures

Figure S1: Silver stained SDS-PAGE of protein extracts.
Lane 1: Bio-Rad molecular weight standards. Lane 2: proteins extracted from desiccated *Fusarium* on the surface of CL-30 maize kernels. Lane 3: water washed endosperm proteins from *Fusarium*-infected maize. Lane 4: water washed embryo proteins from *Fusarium*-infected maize. Lane 5: crude endosperm proteins from *Fusarium*-infected maize. Lane 6: crude embryo proteins from *Fusarium*-infected maize. Lane 7: crude endosperm proteins from control maize. Lane 8: water washed endosperm proteins from control maize. Lane 9: crude embryo proteins from control maize. Lane 10: water washed embryo proteins from maize.

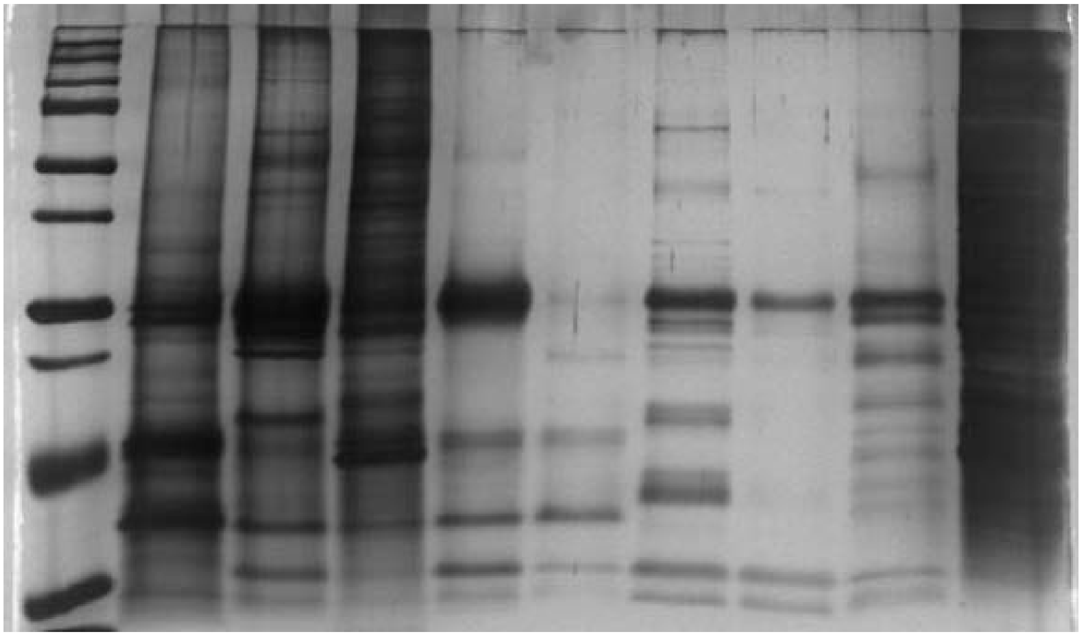


Table 6.1: Proteins isolated from commercially purified wheat starch.

Genbank Accession	Protein	Species	Mascot Score
BAE48800	starch synthase II-D	*Triticum aestivum*	742
AAD53263	starch synthase IIA	*Triticum aestivum*	686
CAB96627	starch synthase IIa-3	*Triticum aestivum*	679
BAA77351	starch synthase (GBSSI)	*Triticum aestivum*	555
AAD26156	granule-bound starch synthase precursor	*Triticum aestivum*	516
P27736	granule-bound starch synthase 1	*Triticum aestivum*	454
AAB17086	1,4-alpha-D-glucan 6-alpha-D-(1,4-alpha-D-glucanotransferase)	*Triticum aestivum*	233
ACN97412	granule-bound starch synthase I	*Elymus repens*	219
ABK80483	putative granule bound starch synthase	*Sorghum bicolor*	214
AAW80631	starch branching enzyme IIb	*Triticum aestivum*	194
BAC06488	granule bound starch synthase	*Setaria italica*	175
ACZ74552	granule-bound starch synthase	*Zea mays*	160
ACJ65514	globulin-3	*Triticum aestivum*	142
ACS36512	waxy	*Arundo donax*	133
AAB02197	soluble starch synthase	*Triticum aestivum*	128
2007234A	chitinase a	*Secale cereale*	106

Appendix B: *Curriculum Vitae*

Adam Koziol

Academic History

Ph.D. Biochemistry and Systems Biology. University of Ottawa. 2007 - 2013.

M.Sc. Biology. University of New Brunswick. Fredericton, New Brunswick, Canada. 2004-2007.

B.Sc. Biology and Biotechnology (High Honours). Carleton University. Ottawa, Ontario, Canada. 1999-2003.

Theses and Supervisors

Koziol AG. Application of direct-sequencing peptide proteomics to the characterization of antagonistic (endogenous and exogenous) proteins in cereal grains. University of Ottawa. Supervisor: Illimar Altosaar.

Koziol AG. 2007. The evolution of the chlorophyll *a*/*b* light-harvesting complex superfamily. M.Sc. Thesis. University of New Brunswick, Fredericton, New Brunswick, Canada, 100 pp. Supervisor: Dion G. Durnford.

Koziol AG. 2003. The role of auxin in phase II secondary xylem development. B.Sc. Honours Thesis. Carleton University, Ottawa, Ontario, Canada. 25 pp. Supervisor: Sharon Regan.

Areas of Academic Interest

Food safety, mass spectrometry, bioinformatics, molecular biology, molecular evolution, antimicrobial peptides, recombinant protein expression.

a. publications in refereed journals

Koziol AG, Marquez BK, Huebsch MP, Smith JC and Altosaar I. (2012). The starch granule associated proteomes of commercially purified starch reference materials from rice and maize. **Journal of Proteomics**. 75: 993-1003. (PhD work)

Koziol AG, Loit E, McNulty M, MacFarlane A, Scott FW and Altosaar I. (2012). Seed storage proteins of the globulin family are cleaved post-translationally in wheat embryos. Accepted July 5, 2012. 19 pages. **BMC Research Notes**. 5: 385, 10 pages, DOI: 10.1186/1756-0500-5-385. (PhD work)

Koziol AG, Marquez BK, Huebsch MP, Smith JC and Altosaar I. (2012). Commercially produced rice and maize starches contain non-host proteins, as shown by mass spectrometry. 14 pages. **Cereal Chemistry**. 89(5): 262-264, DOI: 10.1094/CCHEM-04-12-0043-N. (PhD work)

Zaidi MA, El Bilali JE, **Koziol AG,** Styles G, Ward TL, Greenham TJ, Faiella WM, Son HH, Wan S, Taga I, and Altosaar I. (2012). Gene technology in agriculture, environment and biopharming: Beyond Bt-rice and building better breeding budgets for crops. 20 pages. **Journal of Plant Biochemistry and Biotechnology,** 21 (Suppl 1): S2-S9, DOI: 10.1007/s13562-012-0128-z.

Koziol AG and Durnford DG. (2008). *Euglena* light-harvesting complexes are encoded by multifarious polyprotein mRNAs that evolve in concert. **Molecular Biology and Evolution**. 25: 92–100. (MSc work)

Koziol AG, Borza T, Ishida KI, Keeling P, Lee RW and Durnford DG. (2007). Tracing the evolution of the light-harvesting antennae in chlorophyll *a*/*b*-containing organisms. **Plant Physiology**. 143: 1802-1816. (MSc work)

b. Articles submitted to refereed journals

Koziol AG, Reid LM, Altosaar I. (2012). Taking stock of the protein remnants on the battlefield between host and pathogen: Maize CL30-*Fusarium* interactome. 18 pages. Submitted to **Journal of Proteomics** July, 2012 JPROT-S-12-00689. (PhD work)

c. Non-refereed contributions

Koziol AG and Altosaar I. 2012. Molecular bio-pharming. BMI Poster Day, Ottawa, ON, May 19. Poster. (PhD work)

Koziol AG. 2011. The proteome of a standard reference material: Starch granule associated proteins in rice. BMI Work in progress series. Ottawa, ON, March 2.

Koziol AG and Altosaar I. 2011. The development of a puroindoline protein tethering model. BMI Seminar Symposium (1st annual), Ottawa, ON, February 18.

Koziol AG Wheeler H and Altosaar I. 2008. Solaromics. Meeting of the Canadian Society of Plant Scientists (50th annual). Ottawa, ON. June 14-17. Poster. (PhD work)

Koziol AG and Durnford DG. 2007. Fusion, Duplication, and Deletion: Evolution of *Euglena gracilis* LHC Polyprotein-coding genes. Northeast Algal Society (46th Annual), Narragansett, RI, USA, April 20 – 22.

Koziol AG and Durnford DG. 2006. An Examination of the Evolution and Diversification of Light-Harvesting Complexes in *Euglena gracilis*. Eastern Regional Photosynthesis Conference (23rd annual), Woods Hole, MA, USA, April 21 - 23. Poster.

Appendix C: Reuse permissions

Figure 1.1

NATURE PUBLISHING GROUP LICENSE
TERMS AND CONDITIONS

Feb 26, 2013

This is a License Agreement between Adam G Koziol ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3078201221583
License date	Jan 29, 2013
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature
Licensed content title	After a dozen years of progress the origin of angiosperms is still a great mystery
Licensed content author	Michael W. Frohlich and Mark W. Chase
Licensed content date	Dec 20, 2007
Volume number	450
Issue number	7173
Type of Use	reuse in a thesis/dissertation
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/ illustrations	1
High-res required	no
Figures	Box 1 Figure: Diagrammatic ovules
Author of this NPG article	no
Your reference number	
Title of your thesis / dissertation	Application of direct-sequencing peptide proteomics to the characterization of antagonistic (endogenous and exogenous) proteins in cereal grains
Expected completion date	Jan 2013
Estimated size (number of pages)	171
Total	0.00 USD
Terms and Conditions	

Terms and Conditions for Permissions

Nature Publishing Group hereby grants you a non-exclusive license to reproduce this material for this purpose, and for no other use, subject to the conditions below:

1. NPG warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to Nature Publishing Group and does not carry the copyright of another entity (as credited in the published version). If the credit line on any part of the material you have requested indicates that it was reprinted or adapted by NPG with permission from

another source, then you should also seek permission from that source to reuse the material.

2. Permission granted free of charge for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to the work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version. Where print permission has been granted for a fee, separate permission must be obtained for any additional, electronic re-use (unless, as in the case of a full paper, this has already been accounted for during your initial request in the calculation of a print run). NB: In all cases, web-based use of full-text articles must be authorized separately through the 'Use on a Web Site' option when requesting permission.

3. Permission granted for a first edition does not apply to second and subsequent editions and for editions in other languages (except for signatories to the STM Permissions Guidelines, or where the first edition permission was granted for free).

4. Nature Publishing Group's permission must be acknowledged next to the figure, table or abstract in print. In electronic form, this acknowledgement must be visible at the same time as the figure/table/abstract, and must be hyperlinked to the journal's homepage.

5. The credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)
   For AOP papers, the credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

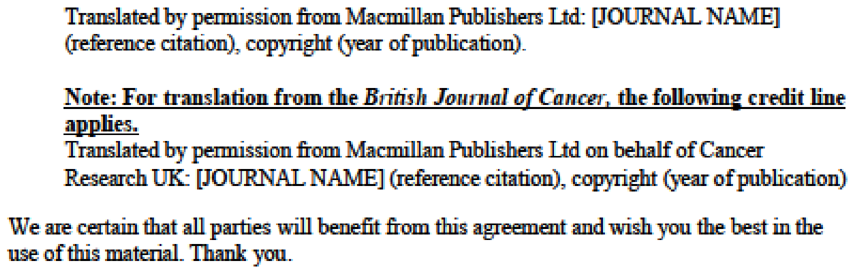
   <u>Note: For republication from the *British Journal of Cancer*, the following credit lines apply.</u>
   Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication) For AOP papers, the credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/ sj.[JOURNAL ACRONYM].XXXXX)

6. Adaptations of single figures do not require NPG approval. However, the adaptation should be credited as follows:

   Adapted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)

   <u>Note: For adaptation from the *British Journal of Cancer*, the following credit line applies.</u>
   Adapted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

7. Translations of 401 words up to a whole article require NPG approval. Please visit http://www.macmillanmedicalcommunications.com for more information.
   Translations of up to a 400 words do not require NPG approval. The translation should be credited as follows:

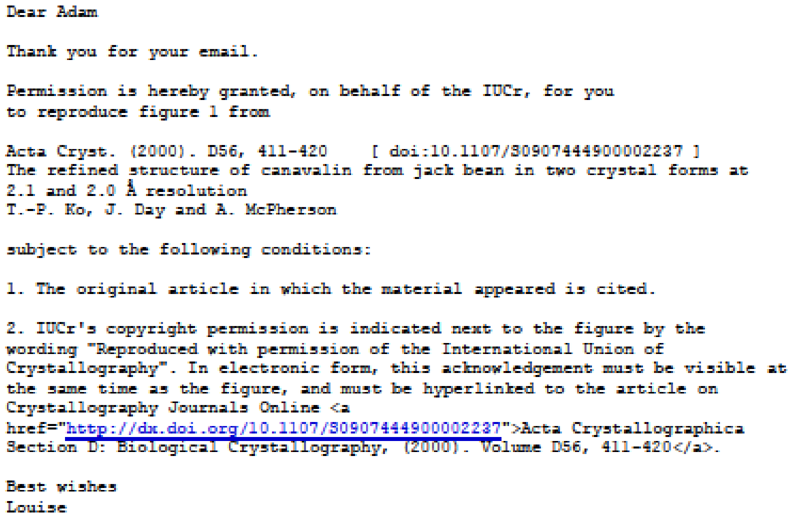


Translated by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication).

Note: For translation from the *British Journal of Cancer*, the following credit line applies.
Translated by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

We are certain that all parties will benefit from this agreement and wish you the best in the use of this material. Thank you.

Figure 1.2



Dear Adam

Thank you for your email.

Permission is hereby granted, on behalf of the IUCr, for you
to reproduce figure 1 from

Acta Cryst. (2000). D56, 411-420 [doi:10.1107/S0907444900002237]
The refined structure of canavalin from jack bean in two crystal forms at
2.1 and 2.0 Å resolution
T.-P. Ko, J. Day and A. McPherson

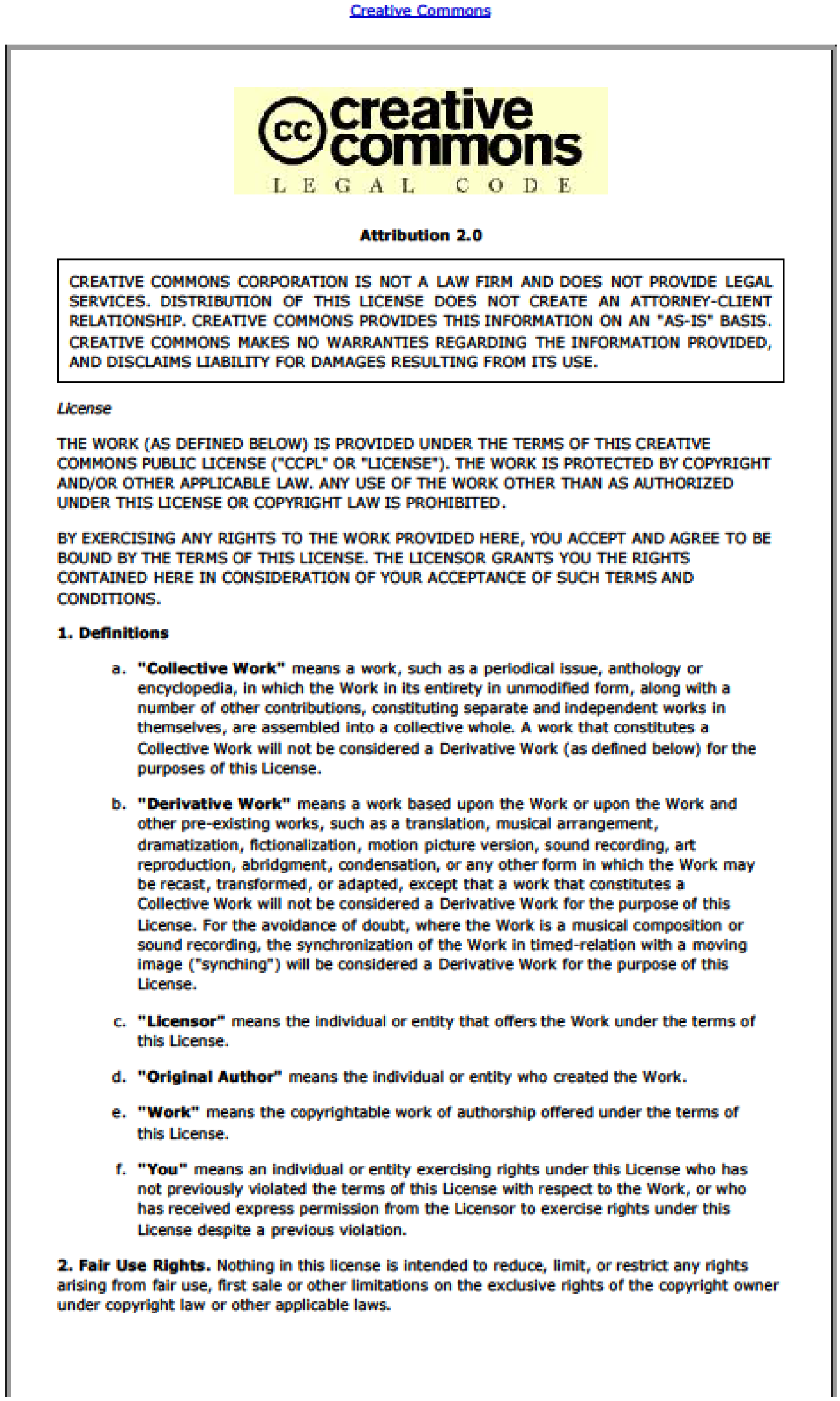
subject to the following conditions:

1. The original article in which the material appeared is cited.

2. IUCr's copyright permission is indicated next to the figure by the
wording "Reproduced with permission of the International Union of
Crystallography". In electronic form, this acknowledgement must be visible at
the same time as the figure, and must be hyperlinked to the article on
Crystallography Journals Online <a
href="http://dx.doi.org/10.1107/S0907444900002237">Acta Crystallographica
Section D: Biological Crystallography, (2000). Volume D56, 411-420.

Best wishes
Louise

Figure 1.3



Creative Commons

creative commons

LEGAL CODE

Attribution 2.0

CREATIVE COMMONS CORPORATION IS NOT A LAW FIRM AND DOES NOT PROVIDE LEGAL SERVICES. DISTRIBUTION OF THIS LICENSE DOES NOT CREATE AN ATTORNEY-CLIENT RELATIONSHIP. CREATIVE COMMONS PROVIDES THIS INFORMATION ON AN "AS-IS" BASIS. CREATIVE COMMONS MAKES NO WARRANTIES REGARDING THE INFORMATION PROVIDED, AND DISCLAIMS LIABILITY FOR DAMAGES RESULTING FROM ITS USE.

License

THE WORK (AS DEFINED BELOW) IS PROVIDED UNDER THE TERMS OF THIS CREATIVE COMMONS PUBLIC LICENSE ("CCPL" OR "LICENSE"). THE WORK IS PROTECTED BY COPYRIGHT AND/OR OTHER APPLICABLE LAW. ANY USE OF THE WORK OTHER THAN AS AUTHORIZED UNDER THIS LICENSE OR COPYRIGHT LAW IS PROHIBITED.

BY EXERCISING ANY RIGHTS TO THE WORK PROVIDED HERE, YOU ACCEPT AND AGREE TO BE BOUND BY THE TERMS OF THIS LICENSE. THE LICENSOR GRANTS YOU THE RIGHTS CONTAINED HERE IN CONSIDERATION OF YOUR ACCEPTANCE OF SUCH TERMS AND CONDITIONS.

1. Definitions

a. **"Collective Work"** means a work, such as a periodical issue, anthology or encyclopedia, in which the Work in its entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole. A work that constitutes a Collective Work will not be considered a Derivative Work (as defined below) for the purposes of this License.

b. **"Derivative Work"** means a work based upon the Work or upon the Work and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Work may be recast, transformed, or adapted, except that a work that constitutes a Collective Work will not be considered a Derivative Work for the purpose of this License. For the avoidance of doubt, where the Work is a musical composition or sound recording, the synchronization of the Work in timed-relation with a moving image ("synching") will be considered a Derivative Work for the purpose of this License.

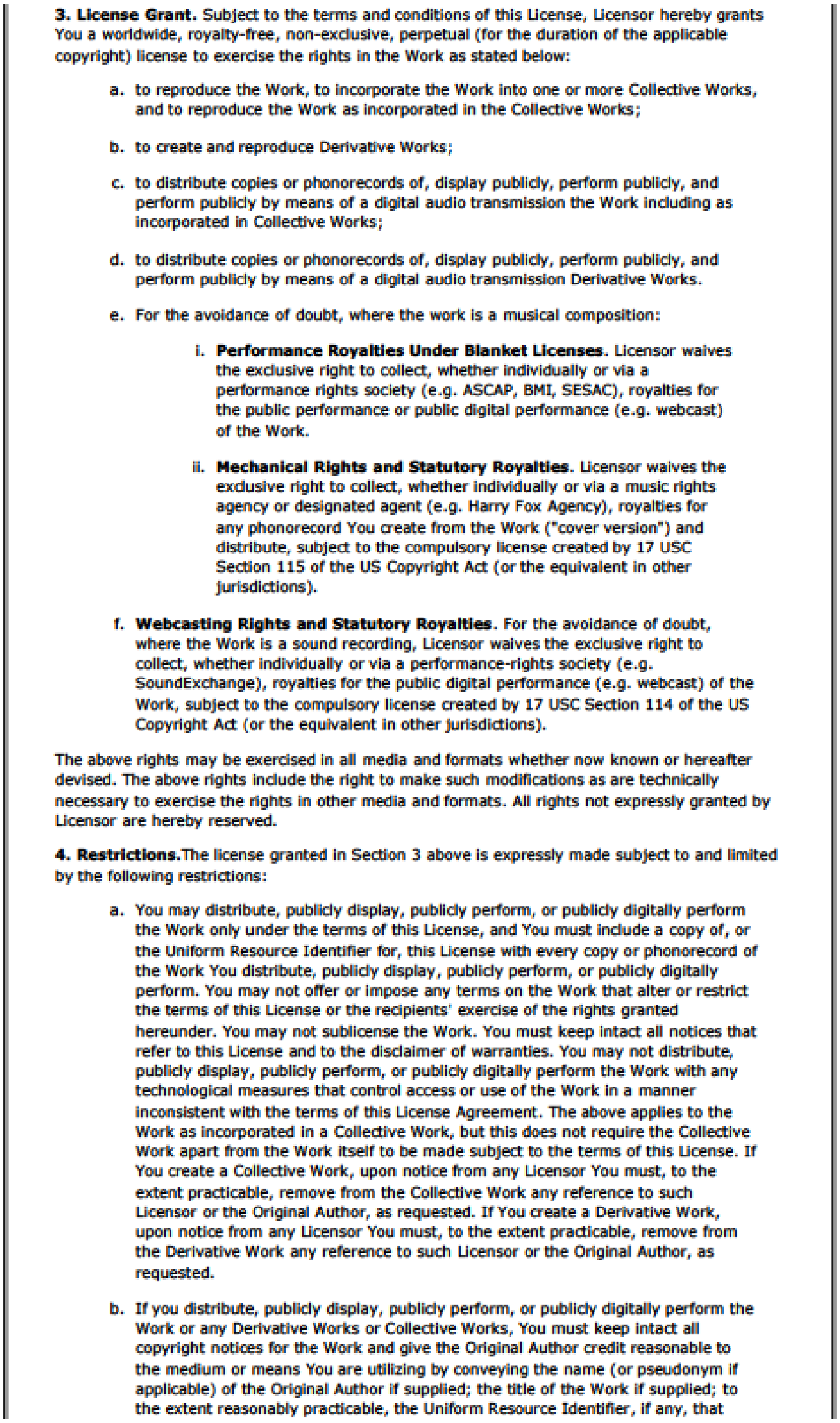
c. **"Licensor"** means the individual or entity that offers the Work under the terms of this License.

d. **"Original Author"** means the individual or entity who created the Work.

e. **"Work"** means the copyrightable work of authorship offered under the terms of this License.

f. **"You"** means an individual or entity exercising rights under this License who has not previously violated the terms of this License with respect to the Work, or who has received express permission from the Licensor to exercise rights under this License despite a previous violation.

2. Fair Use Rights. Nothing in this license is intended to reduce, limit, or restrict any rights arising from fair use, first sale or other limitations on the exclusive rights of the copyright owner under copyright law or other applicable laws.



3. License Grant. Subject to the terms and conditions of this License, Licensor hereby grants You a worldwide, royalty-free, non-exclusive, perpetual (for the duration of the applicable copyright) license to exercise the rights in the Work as stated below:

a. to reproduce the Work, to incorporate the Work into one or more Collective Works, and to reproduce the Work as incorporated in the Collective Works;

b. to create and reproduce Derivative Works;

c. to distribute copies or phonorecords of, display publicly, perform publicly, and perform publicly by means of a digital audio transmission the Work including as incorporated in Collective Works;

d. to distribute copies or phonorecords of, display publicly, perform publicly, and perform publicly by means of a digital audio transmission Derivative Works.

e. For the avoidance of doubt, where the work is a musical composition:

    i. **Performance Royalties Under Blanket Licenses.** Licensor waives the exclusive right to collect, whether individually or via a performance rights society (e.g. ASCAP, BMI, SESAC), royalties for the public performance or public digital performance (e.g. webcast) of the Work.

    ii. **Mechanical Rights and Statutory Royalties.** Licensor waives the exclusive right to collect, whether individually or via a music rights agency or designated agent (e.g. Harry Fox Agency), royalties for any phonorecord You create from the Work ("cover version") and distribute, subject to the compulsory license created by 17 USC Section 115 of the US Copyright Act (or the equivalent in other jurisdictions).

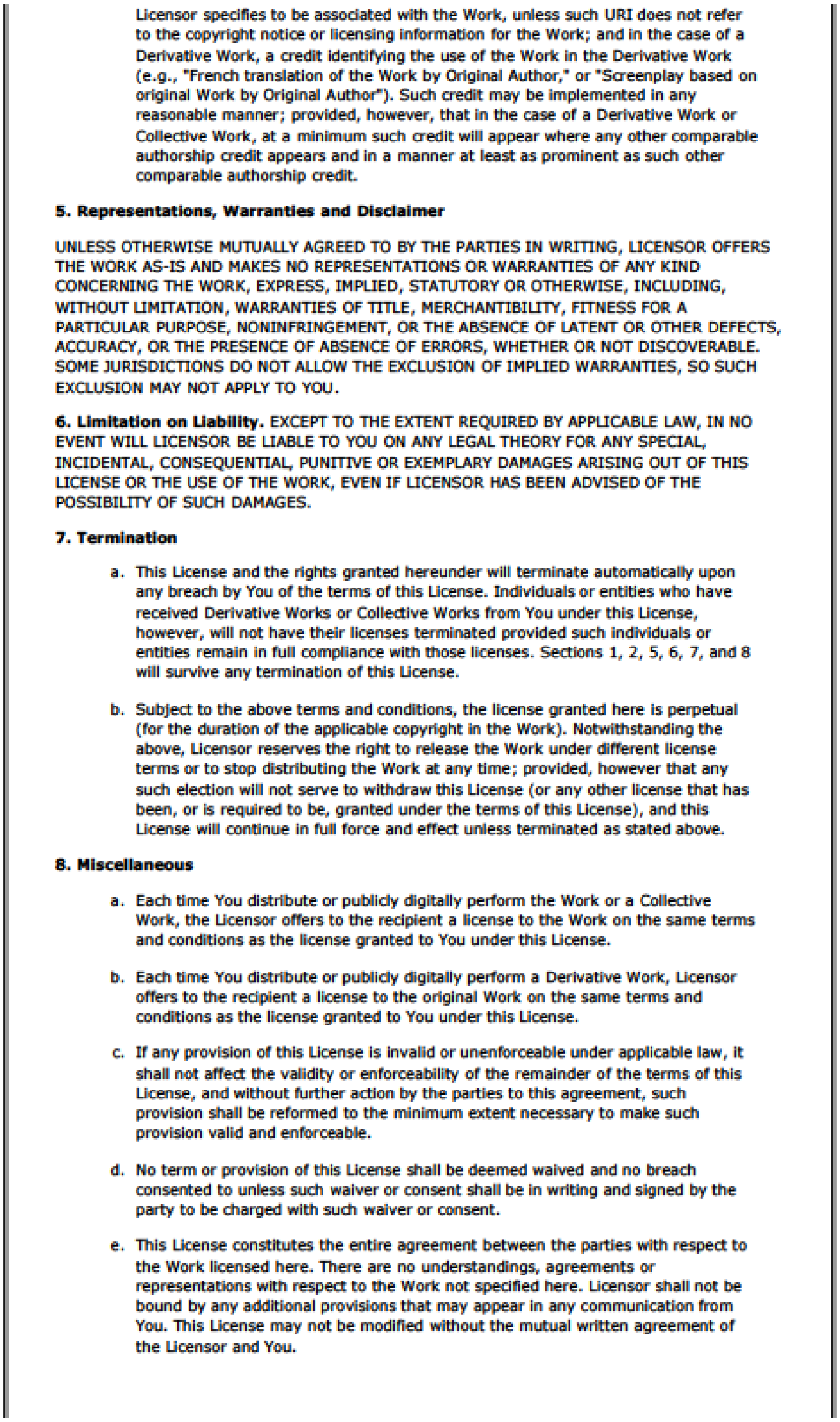
f. **Webcasting Rights and Statutory Royalties.** For the avoidance of doubt, where the Work is a sound recording, Licensor waives the exclusive right to collect, whether individually or via a performance-rights society (e.g. SoundExchange), royalties for the public digital performance (e.g. webcast) of the Work, subject to the compulsory license created by 17 USC Section 114 of the US Copyright Act (or the equivalent in other jurisdictions).

The above rights may be exercised in all media and formats whether now known or hereafter devised. The above rights include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. All rights not expressly granted by Licensor are hereby reserved.

4. Restrictions. The license granted in Section 3 above is expressly made subject to and limited by the following restrictions:

a. You may distribute, publicly display, publicly perform, or publicly digitally perform the Work only under the terms of this License, and You must include a copy of, or the Uniform Resource Identifier for, this License with every copy or phonorecord of the Work You distribute, publicly display, publicly perform, or publicly digitally perform. You may not offer or impose any terms on the Work that alter or restrict the terms of this License or the recipients' exercise of the rights granted hereunder. You may not sublicense the Work. You must keep intact all notices that refer to this License and to the disclaimer of warranties. You may not distribute, publicly display, publicly perform, or publicly digitally perform the Work with any technological measures that control access or use of the Work in a manner inconsistent with the terms of this License Agreement. The above applies to the Work as incorporated in a Collective Work, but this does not require the Collective Work apart from the Work itself to be made subject to the terms of this License. If You create a Collective Work, upon notice from any Licensor You must, to the extent practicable, remove from the Collective Work any reference to such Licensor or the Original Author, as requested. If You create a Derivative Work, upon notice from any Licensor You must, to the extent practicable, remove from the Derivative Work any reference to such Licensor or the Original Author, as requested.

b. If you distribute, publicly display, publicly perform, or publicly digitally perform the Work or any Derivative Works or Collective Works, You must keep intact all copyright notices for the Work and give the Original Author credit reasonable to the medium or means You are utilizing by conveying the name (or pseudonym if applicable) of the Original Author if supplied; the title of the Work if supplied; to the extent reasonably practicable, the Uniform Resource Identifier, if any, that



Licensor specifies to be associated with the Work, unless such URI does not refer to the copyright notice or licensing information for the Work; and in the case of a Derivative Work, a credit identifying the use of the Work in the Derivative Work (e.g., "French translation of the Work by Original Author," or "Screenplay based on original Work by Original Author"). Such credit may be implemented in any reasonable manner; provided, however, that in the case of a Derivative Work or Collective Work, at a minimum such credit will appear where any other comparable authorship credit appears and in a manner at least as prominent as such other comparable authorship credit.

5. Representations, Warranties and Disclaimer

UNLESS OTHERWISE MUTUALLY AGREED TO BY THE PARTIES IN WRITING, LICENSOR OFFERS THE WORK AS-IS AND MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND CONCERNING THE WORK, EXPRESS, IMPLIED, STATUTORY OR OTHERWISE, INCLUDING, WITHOUT LIMITATION, WARRANTIES OF TITLE, MERCHANTIBILITY, FITNESS FOR A PARTICULAR PURPOSE, NONINFRINGEMENT, OR THE ABSENCE OF LATENT OR OTHER DEFECTS, ACCURACY, OR THE PRESENCE OF ABSENCE OF ERRORS, WHETHER OR NOT DISCOVERABLE. SOME JURISDICTIONS DO NOT ALLOW THE EXCLUSION OF IMPLIED WARRANTIES, SO SUCH EXCLUSION MAY NOT APPLY TO YOU.

6. Limitation on Liability. EXCEPT TO THE EXTENT REQUIRED BY APPLICABLE LAW, IN NO EVENT WILL LICENSOR BE LIABLE TO YOU ON ANY LEGAL THEORY FOR ANY SPECIAL, INCIDENTAL, CONSEQUENTIAL, PUNITIVE OR EXEMPLARY DAMAGES ARISING OUT OF THIS LICENSE OR THE USE OF THE WORK, EVEN IF LICENSOR HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

7. Termination

a. This License and the rights granted hereunder will terminate automatically upon any breach by You of the terms of this License. Individuals or entities who have received Derivative Works or Collective Works from You under this License, however, will not have their licenses terminated provided such individuals or entities remain in full compliance with those licenses. Sections 1, 2, 5, 6, 7, and 8 will survive any termination of this License.

b. Subject to the above terms and conditions, the license granted here is perpetual (for the duration of the applicable copyright in the Work). Notwithstanding the above, Licensor reserves the right to release the Work under different license terms or to stop distributing the Work at any time; provided, however that any such election will not serve to withdraw this License (or any other license that has been, or is required to be, granted under the terms of this License), and this License will continue in full force and effect unless terminated as stated above.

8. Miscellaneous

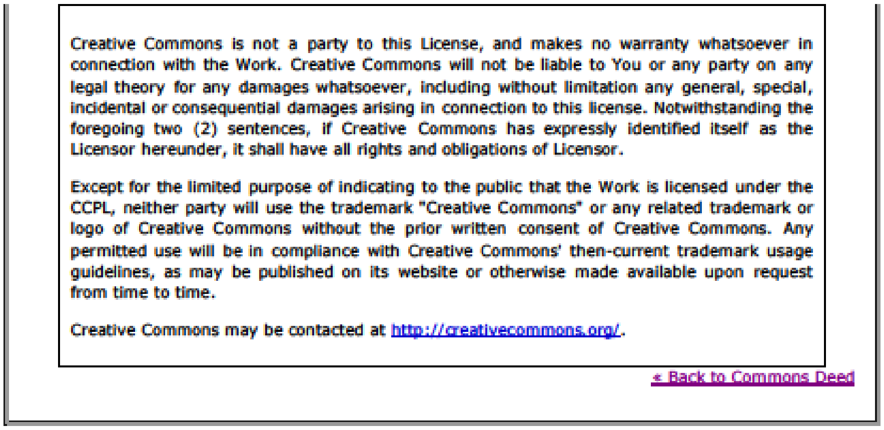
a. Each time You distribute or publicly digitally perform the Work or a Collective Work, the Licensor offers to the recipient a license to the Work on the same terms and conditions as the license granted to You under this License.

b. Each time You distribute or publicly digitally perform a Derivative Work, Licensor offers to the recipient a license to the original Work on the same terms and conditions as the license granted to You under this License.

c. If any provision of this License is invalid or unenforceable under applicable law, it shall not affect the validity or enforceability of the remainder of the terms of this License, and without further action by the parties to this agreement, such provision shall be reformed to the minimum extent necessary to make such provision valid and enforceable.

d. No term or provision of this License shall be deemed waived and no breach consented to unless such waiver or consent shall be in writing and signed by the party to be charged with such waiver or consent.

e. This License constitutes the entire agreement between the parties with respect to the Work licensed here. There are no understandings, agreements or representations with respect to the Work not specified here. Licensor shall not be bound by any additional provisions that may appear in any communication from You. This License may not be modified without the mutual written agreement of the Licensor and You.



Creative Commons is not a party to this License, and makes no warranty whatsoever in connection with the Work. Creative Commons will not be liable to You or any party on any legal theory for any damages whatsoever, including without limitation any general, special, incidental or consequential damages arising in connection to this license. Notwithstanding the foregoing two (2) sentences, if Creative Commons has expressly identified itself as the Licensor hereunder, it shall have all rights and obligations of Licensor.

Except for the limited purpose of indicating to the public that the Work is licensed under the CCPL, neither party will use the trademark "Creative Commons" or any related trademark or logo of Creative Commons without the prior written consent of Creative Commons. Any permitted use will be in compliance with Creative Commons' then-current trademark usage guidelines, as may be published on its website or otherwise made available upon request from time to time.

Creative Commons may be contacted at http://creativecommons.org/.

« Back to Commons Deed

Figure 1.4

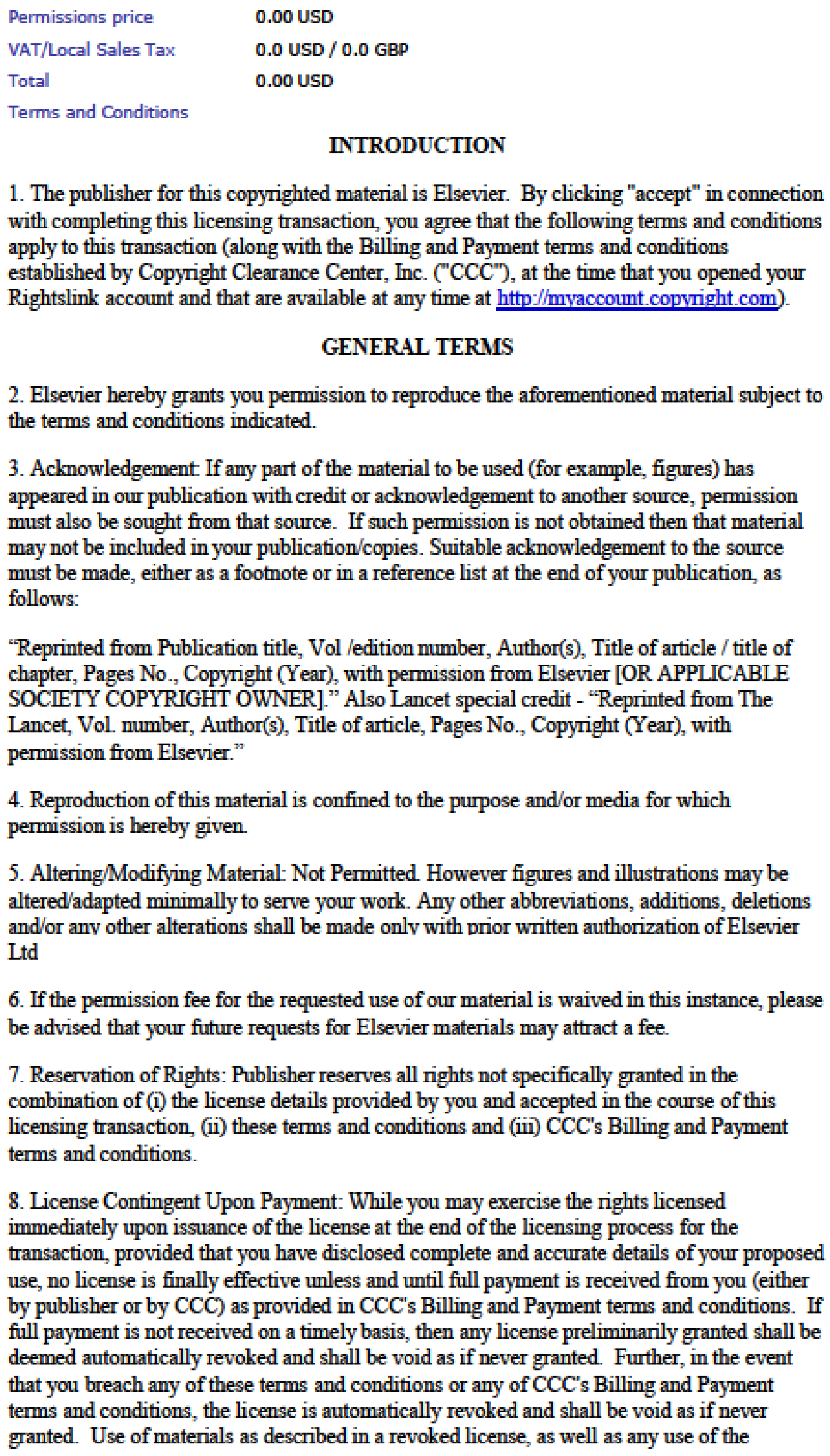
ELSEVIER LICENSE
TERMS AND CONDITIONS

Feb 26, 2013

This is a License Agreement between Adam G Koziol ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited
Registered Company Number	1982084
Customer name	Adam G Koziol
Customer address	
License number	3078230092519
License date	Jan 29, 2013
Licensed content publisher	Elsevier
Licensed content publication	Journal of Cereal Science
Licensed content title	Genetic Alteration of Starch Functionality in Wheat
Licensed content author	S. Rahman,Z. Li,I. Batey,M.P. Cochrane,R. Appels,M. Morell
Licensed content date	January 2000
Licensed content volume number	31
Licensed content issue number	1
Number of pages	20
Start Page	91
End Page	110
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	other
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Order reference number	
Title of your thesis/dissertation	Application of direct-sequencing peptide proteomics to the characterization of antagonistic (endogenous and exogenous) proteins in cereal grains
Expected completion date	Jan 2013
Estimated size (number of pages)	171
Elsevier VAT number	GB 494 6272 12



Permissions price	0.00 USD
VAT/Local Sales Tax	0.0 USD / 0.0 GBP
Total	0.00 USD
Terms and Conditions	

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the

materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article. If this license is to re-use 1 or 2 figures then permission is granted for non-exclusive world rights in all languages.

16. **Website**: The following terms and conditions apply to electronic reserve and author websites:
Electronic reserve: If licensed material is to be posted to website, the web site is to be password-protected and made available only to bona fide students registered on a relevant course if:
This license was made in connection with a course,
This permission is granted for 1 year only. You may obtain a license for future website posting,
All content posted to the web site must maintain the copyright information line on the

bottom of each image,
A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com , and
Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

17. **Author website** for journals with the following additional clauses:

All content posted to the web site must maintain the copyright information line on the bottom of each image, and the permission granted is limited to the personal version of your paper. You are not allowed to download and post the published electronic version of your article (whether PDF or HTML, proof or final version), nor may you scan the printed edition to create an electronic version. A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx . As part of our normal production process, you will receive an e-mail notice when your article appears on Elsevier's online service ScienceDirect (www.sciencedirect.com). That e-mail will include the article's Digital Object Identifier (DOI). This number provides the electronic link to the published article and should be included in the posting of your personal version. We ask that you wait until you receive this e-mail and have the DOI to do any posting.

Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

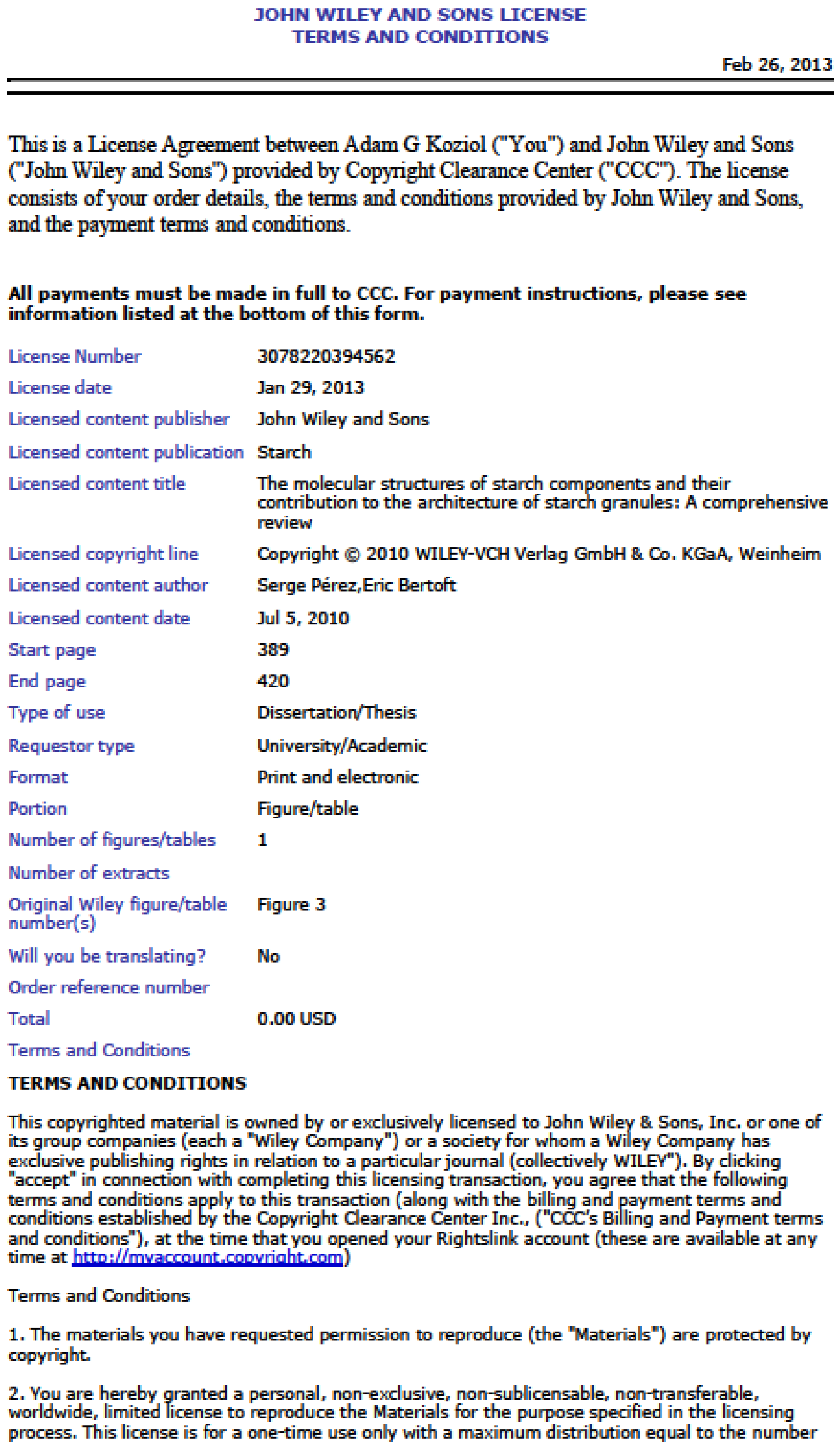
18. **Author website** for books with the following additional clauses:
Authors are permitted to place a brief summary of their work online only.
A hyper-text must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version.

Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

19. **Website** (regular and for author): A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx. or for books to the Elsevier homepage at http://www.elsevier.com

20. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission.

Figure 1.5



JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

Feb 26, 2013

This is a License Agreement between Adam G Koziol ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3078220394562
License date	Jan 29, 2013
Licensed content publisher	John Wiley and Sons
Licensed content publication	Starch
Licensed content title	The molecular structures of starch components and their contribution to the architecture of starch granules: A comprehensive review
Licensed copyright line	Copyright © 2010 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
Licensed content author	Serge Pérez,Eric Bertoft
Licensed content date	Jul 5, 2010
Start page	389
End page	420
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Number of extracts	
Original Wiley figure/table number(s)	Figure 3
Will you be translating?	No
Order reference number	
Total	0.00 USD
Terms and Conditions	

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or a society for whom a Wiley Company has exclusive publishing rights in relation to a particular journal (collectively WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your Rightslink account (these are available at any time at http://myaccount.copyright.com)

Terms and Conditions

1. The materials you have requested permission to reproduce (the "Materials") are protected by copyright.

2. You are hereby granted a personal, non-exclusive, non-sublicensable, non-transferable, worldwide, limited license to reproduce the Materials for the purpose specified in the licensing process. This license is for a one-time use only with a maximum distribution equal to the number

that you identified in the licensing process. Any form of republication granted by this licence must be completed within two years of the date of the grant of this licence (although copies prepared before may be distributed thereafter). The Materials shall not be used in any other manner or for any other purpose. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Material. Any third party material is expressly excluded from this permission.

3. With respect to the Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Materials without the prior permission of the respective copyright owner. You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Materials, or any of the rights granted to you hereunder to any other person.

4. The Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc or one of its related companies (WILEY) or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto.

5. NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

6. WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

7. You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

8. IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

9. Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

10. The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

11. This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

12. Any fee required for this permission shall be non-refundable after thirty (30) days from receipt.

13. These terms and conditions together with CCC's Billing and Payment terms and conditions

(which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

14. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

15. WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

16. This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

17. This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

Wiley Open Access Terms and Conditions

All research articles published in Wiley Open Access journals are fully open access: immediately freely available to read, download and share. Articles are published under the terms of the Creative Commons Attribution Non Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes. The license is subject to the Wiley Open Access terms and conditions:
Wiley Open Access articles are protected by copyright and are posted to repositories and websites in accordance with the terms of the Creative Commons Attribution Non Commercial License. At the time of deposit, Wiley Open Access articles include all changes made during peer review, copyediting, and publishing. Repositories and websites that host the article are responsible for incorporating any publisher-supplied amendments or retractions issued subsequently.
Wiley Open Access articles are also available without charge on Wiley's publishing platform, **Wiley Online Library** or any successor sites.

Use by non-commercial users

For non-commercial and non-promotional purposes individual users may access, download, copy, display and redistribute to colleagues Wiley Open Access articles, as well as adapt, translate, text- and data-mine the content subject to the following conditions:

- The authors' moral rights are not compromised. These rights include the right of "paternity" (also known as "attribution" - the right for the author to be identified as such) and "integrity" (the right for the author not to have the work altered in such a way that the author's reputation or integrity may be impugned).
- Where content in the article is identified as belonging to a third party, it is the obligation of the user to ensure that any reuse complies with the copyright policies of the owner of that content.
- If article content is copied, downloaded or otherwise reused for non-commercial research and education purposes, a link to the appropriate bibliographic citation (authors, journal, article title, volume, issue, page numbers, DOI and the link to the definitive published version on Wiley Online Library) should be maintained. Copyright notices and disclaimers must not be deleted.
- Any translations, for which a prior translation agreement with Wiley has not been agreed, must prominently display the statement: "This is an unofficial translation of an article that appeared in a Wiley publication. The publisher has not endorsed this translation."

Use by commercial "for-profit" organisations

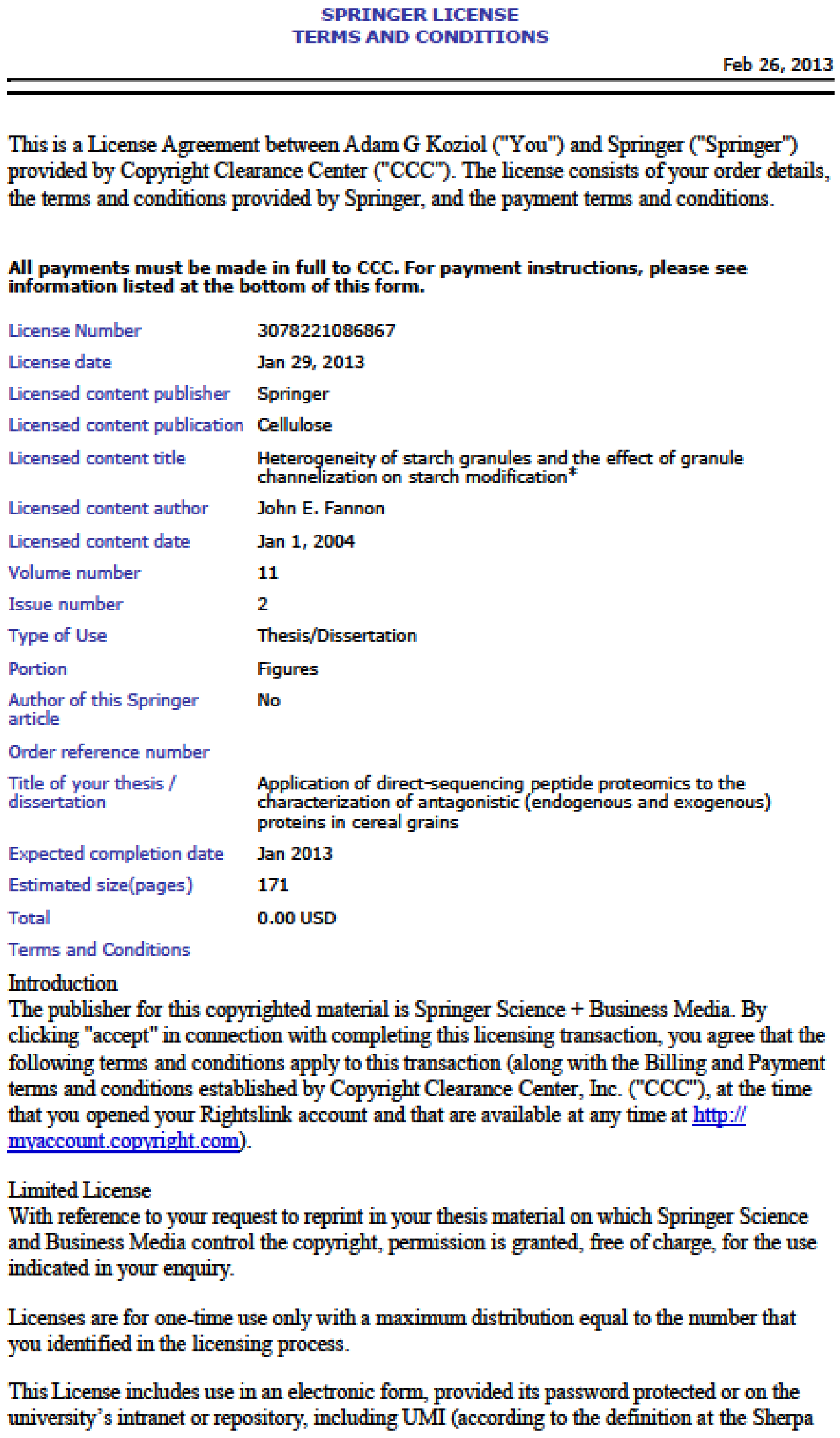
Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee. Commercial purposes include:

- Copying or downloading of articles, or linking to such articles for further redistribution, sale or licensing;
- Copying, downloading or posting by a site or service that incorporates advertising with such content;
- The inclusion or incorporation of article content in other works or services (other than normal quotations with an appropriate citation) that is then available for sale or licensing, for a fee (for example, a compilation produced for marketing purposes, inclusion in a sales pack)
- Use of article content (other than normal quotations with appropriate citation) by for-profit organisations for promotional purposes
- Linking to article content in e-mails redistributed for promotional, marketing or educational purposes;
- Use for the purposes of monetary reward by means of sale, resale, licence, loan, transfer or other form of commercial exploitation such as marketing products
- Print reprints of Wiley Open Access articles can be purchased from:

Other Terms and Conditions:

BY CLICKING ON THE "I AGREE..." BOX, YOU ACKNOWLEDGE THAT YOU HAVE READ AND FULLY UNDERSTAND EACH OF THE SECTIONS OF AND PROVISIONS SET FORTH IN THIS AGREEMENT AND THAT YOU ARE IN AGREEMENT WITH AND ARE WILLING TO ACCEPT ALL OF YOUR OBLIGATIONS AS SET FORTH IN THIS AGREEMENT.

Figure 1.6



SPRINGER LICENSE
TERMS AND CONDITIONS

Feb 26, 2013

This is a License Agreement between Adam G Koziol ("You") and Springer ("Springer") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3078221086867
License date	Jan 29, 2013
Licensed content publisher	Springer
Licensed content publication	Cellulose
Licensed content title	Heterogeneity of starch granules and the effect of granule channelization on starch modification*
Licensed content author	John E. Fannon
Licensed content date	Jan 1, 2004
Volume number	11
Issue number	2
Type of Use	Thesis/Dissertation
Portion	Figures
Author of this Springer article	No
Order reference number	
Title of your thesis / dissertation	Application of direct-sequencing peptide proteomics to the characterization of antagonistic (endogenous and exogenous) proteins in cereal grains
Expected completion date	Jan 2013
Estimated size(pages)	171
Total	0.00 USD
Terms and Conditions	

Introduction
The publisher for this copyrighted material is Springer Science + Business Media. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

Limited License
With reference to your request to reprint in your thesis material on which Springer Science and Business Media control the copyright, permission is granted, free of charge, for the use indicated in your enquiry.

Licenses are for one-time use only with a maximum distribution equal to the number that you identified in the licensing process.

This License includes use in an electronic form, provided its password protected or on the university's intranet or repository, including UMI (according to the definition at the Sherpa website: http://www.sherpa.ac.uk/romeo/). For any other electronic use, please contact Springer

The material can only be used for the purpose of defending your thesis, and with a maximum of 100 extra copies in paper.

Although Springer holds copyright to the material and is entitled to negotiate on rights, this license is only valid, provided permission is also obtained from the (co) author (address is given with the article/chapter) and provided it concerns original material which does not carry references to other sources (if material in question appears with credit to another source, authorization from that source is required as well).

Permission free of charge on this occasion does not prejudice any rights we might have to charge for reproduction of our copyrighted material in the future.

Altering/Modifying Material: Not Permitted
You may not alter or modify the material in any manner. Abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of the author (s) and/or Springer Science + Business Media. (Please contact Springer at (permissions.dordrecht@springer.com or permissions.heidelberg@springer.com)

Reservation of Rights
Springer Science + Business Media reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

Copyright Notice:Disclaimer
You must include the following copyright and permission notice in connection with any reproduction of the licensed material: "Springer and the original publisher /journal title, volume, year of publication, page, chapter/article title, name(s) of author(s), figure number(s) , original copyright notice) is given to the publication in which the material was originally published, by adding; with kind permission from Springer Science and Business Media"

Warranties: None

Example 1: Springer Science + Business Media makes no representations or warranties with respect to the licensed material.

Example 2: Springer Science + Business Media makes no representations or warranties with respect to the licensed material and adopts on its own behalf the limitations and disclaimers established by CCC on its behalf in its Billing and Payment terms and conditions for this licensing transaction.

Indemnity
You hereby indemnify and agree to hold harmless Springer Science + Business Media and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

No Transfer of License
This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without Springer Science + Business Media's written permission.

No Amendment Except in Writing
This license may not be amended except in a writing signed by both parties (or, in the case of

Springer Science + Business Media, by CCC on Springer Science + Business Media's behalf).

Objection to Contrary Terms
Springer Science + Business Media hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and Springer Science + Business Media (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

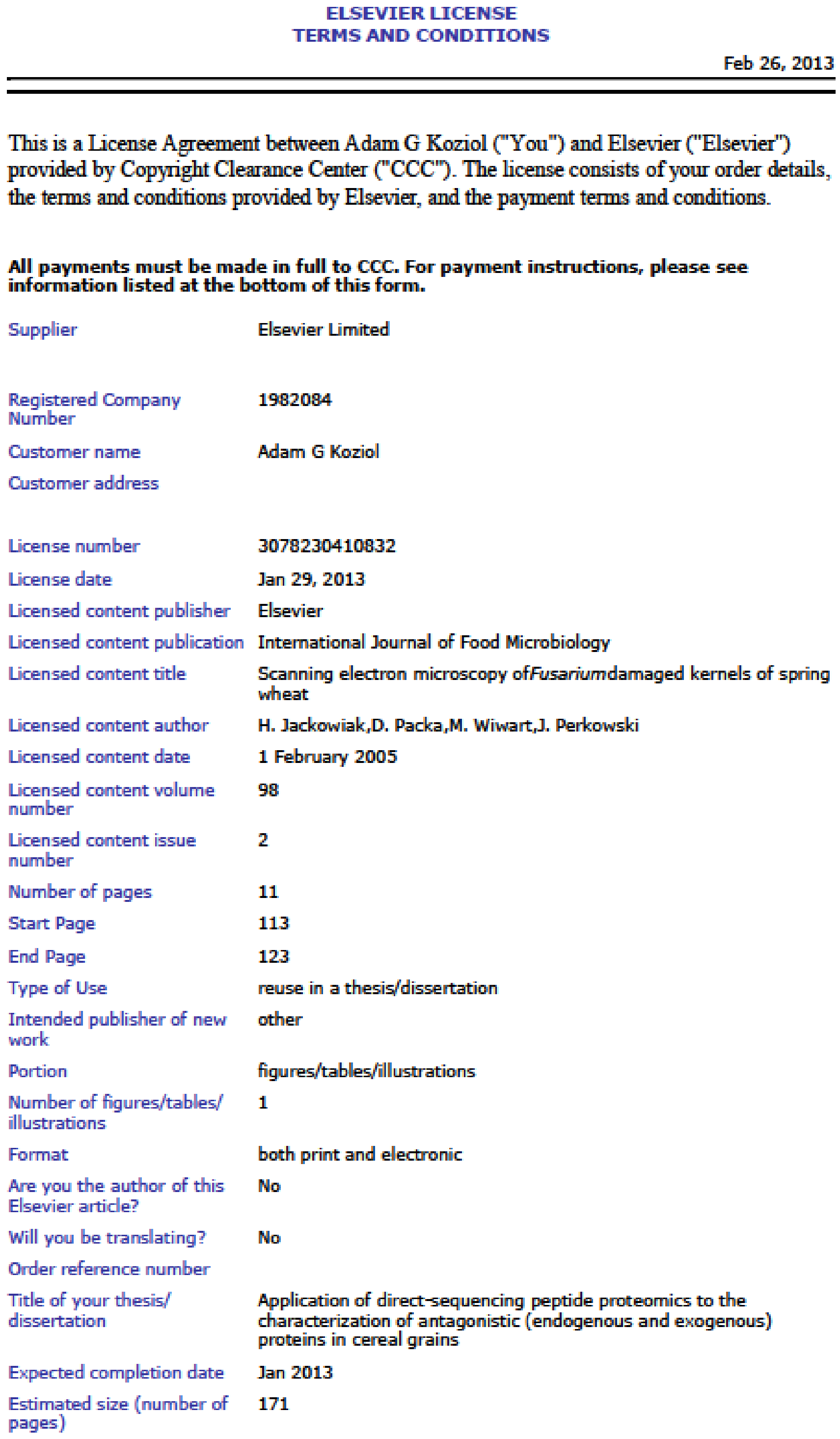
Jurisdiction
All disputes that may arise in connection with this present License, or the breach thereof, shall be settled exclusively by arbitration, to be held in The Netherlands, in accordance with Dutch law, and to be conducted under the Rules of the 'Netherlands Arbitrage Instituut' (Netherlands Institute of Arbitration).***OR:***

All disputes that may arise in connection with this present License, or the breach thereof, shall be settled exclusively by arbitration, to be held in the Federal Republic of Germany, in accordance with German law.

Other terms and conditions:

v1.3

Figure 1.7



ELSEVIER LICENSE
TERMS AND CONDITIONS

Feb 26, 2013

This is a License Agreement between Adam G Koziol ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited
Registered Company Number	1982084
Customer name	Adam G Koziol
Customer address	
License number	3078230410832
License date	Jan 29, 2013
Licensed content publisher	Elsevier
Licensed content publication	International Journal of Food Microbiology
Licensed content title	Scanning electron microscopy of*Fusarium*damaged kernels of spring wheat
Licensed content author	H. Jackowiak,D. Packa,M. Wiwart,J. Perkowski
Licensed content date	1 February 2005
Licensed content volume number	98
Licensed content issue number	2
Number of pages	11
Start Page	113
End Page	123
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	other
Portion	figures/tables/illustrations
Number of figures/tables/ illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Order reference number	
Title of your thesis/ dissertation	Application of direct-sequencing peptide proteomics to the characterization of antagonistic (endogenous and exogenous) proteins in cereal grains
Expected completion date	Jan 2013
Estimated size (number of pages)	171

Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD
VAT/Local Sales Tax	0.0 USD / 0.0 GBP
Total	0.00 USD
Terms and Conditions	

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never

granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article. If this license is to re-use 1 or 2 figures then permission is granted for non-exclusive world rights in all languages.

16. **Website**: The following terms and conditions apply to electronic reserve and author websites:
Electronic reserve: If licensed material is to be posted to website, the web site is to be password-protected and made available only to bona fide students registered on a relevant course if:
This license was made in connection with a course,
This permission is granted for 1 year only. You may obtain a license for future website posting,

All content posted to the web site must maintain the copyright information line on the bottom of each image,
A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com , and
Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

17. **Author website** for journals with the following additional clauses:

All content posted to the web site must maintain the copyright information line on the bottom of each image, and the permission granted is limited to the personal version of your paper. You are not allowed to download and post the published electronic version of your article (whether PDF or HTML, proof or final version), nor may you scan the printed edition to create an electronic version. A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx . As part of our normal production process, you will receive an e-mail notice when your article appears on Elsevier's online service ScienceDirect (www.sciencedirect.com). That e-mail will include the article's Digital Object Identifier (DOI). This number provides the electronic link to the published article and should be included in the posting of your personal version. We ask that you wait until you receive this e-mail and have the DOI to do any posting.

Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

18. **Author website** for books with the following additional clauses:
Authors are permitted to place a brief summary of their work online only.
A hyper-text must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version.

Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

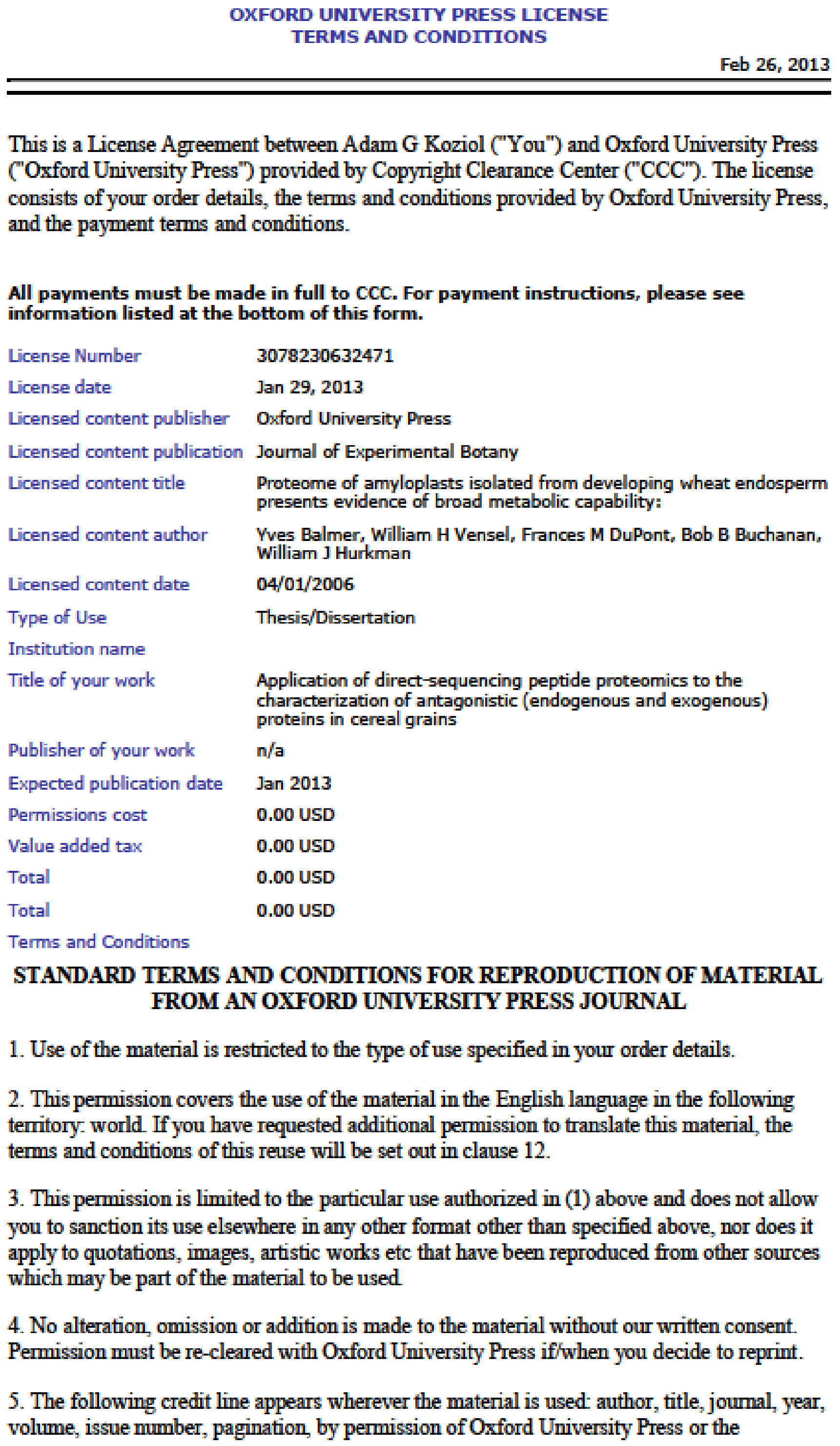
19. **Website** (regular and for author): A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx. or for books to the Elsevier homepage at http://www.elsevier.com

20. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission.

21. **Other Conditions**:

v1.6

Figure 1.8



OXFORD UNIVERSITY PRESS LICENSE
TERMS AND CONDITIONS

Feb 26, 2013

This is a License Agreement between Adam G Koziol ("You") and Oxford University Press ("Oxford University Press") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Oxford University Press, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3078230632471
License date	Jan 29, 2013
Licensed content publisher	Oxford University Press
Licensed content publication	Journal of Experimental Botany
Licensed content title	Proteome of amyloplasts isolated from developing wheat endosperm presents evidence of broad metabolic capability:
Licensed content author	Yves Balmer, William H Vensel, Frances M DuPont, Bob B Buchanan, William J Hurkman
Licensed content date	04/01/2006
Type of Use	Thesis/Dissertation
Institution name	
Title of your work	Application of direct-sequencing peptide proteomics to the characterization of antagonistic (endogenous and exogenous) proteins in cereal grains
Publisher of your work	n/a
Expected publication date	Jan 2013
Permissions cost	0.00 USD
Value added tax	0.00 USD
Total	0.00 USD
Total	0.00 USD
Terms and Conditions	

STANDARD TERMS AND CONDITIONS FOR REPRODUCTION OF MATERIAL FROM AN OXFORD UNIVERSITY PRESS JOURNAL

1. Use of the material is restricted to the type of use specified in your order details.

2. This permission covers the use of the material in the English language in the following territory: world. If you have requested additional permission to translate this material, the terms and conditions of this reuse will be set out in clause 12.

3. This permission is limited to the particular use authorized in (1) above and does not allow you to sanction its use elsewhere in any other format other than specified above, nor does it apply to quotations, images, artistic works etc that have been reproduced from other sources which may be part of the material to be used.

4. No alteration, omission or addition is made to the material without our written consent. Permission must be re-cleared with Oxford University Press if/when you decide to reprint.

5. The following credit line appears wherever the material is used: author, title, journal, year, volume, issue number, pagination, by permission of Oxford University Press or the

sponsoring society if the journal is a society journal. Where a journal is being published on behalf of a learned society, the details of that society must be included in the credit line.

6. For the reproduction of a full article from an Oxford University Press journal for whatever purpose, the corresponding author of the material concerned should be informed of the proposed use. Contact details for the corresponding authors of all Oxford University Press journal contact can be found alongside either the abstract or full text of the article concerned, accessible from www.oxfordjournals.org Should there be a problem clearing these rights, please contact

7. If the credit line or acknowledgement in our publication indicates that any of the figures, images or photos was reproduced, drawn or modified from an earlier source it will be necessary for you to clear this permission with the original publisher as well. If this permission has not been obtained, please note that this material cannot be included in your publication/photocopies.

8. While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by Oxford University Press or by Copyright Clearance Center (CCC)) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and Oxford University Press reserves the right to take any and all action to protect its copyright in the materials.

9. This license is personal to you and may not be sublicensed, assigned or transferred by you to any other person without Oxford University Press's written permission.

10. Oxford University Press reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

11. You hereby indemnify and agree to hold harmless Oxford University Press and CCC, and their respective officers, directors, employs and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

12. Other Terms and Conditions:

v1.4