Identification and Characterization of Late Pathway Enzymes in Phytic Acid Biosynthesis in *Glycine max*

Amanda R. Stiles

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Plant Physiology

Elizabeth A. Grabau, Chair
Cynthia Gibas
Glenda Gillaspy
John Jelesko
Saghai Maroof

July 30th, 2007
Blacksburg, VA

Keywords: *Glycine max*, soybean, phytic acid, phytate, *myo*-inositol kinase, pathway

Identification and Characterization of Late Pathway Enzymes in Phytic Acid Biosynthesis in *Glycine max*

Amanda R. Stiles

ABSTRACT

Phytic acid, also known as *myo*-inositol hexa*kis*phosphate or Ins(1,2,3,4,5,6)P_6, is the major storage form of phosphorus in plant seeds. Phytic acid is poorly digested by non-ruminant animals such as swine and poultry, and it chelates mineral cations including calcium, iron, zinc, and potassium, classifying it as an anti-nutrient. The excretion of unutilized phytic acid in manure translates to an excess amount of phosphorus runoff that can lead to eutrophication of lakes and ponds. Understanding the phytic acid biosynthetic pathway will allow for the development of low phytic acid *(lpa*) soybeans by the down-regulation of specific genes. The goal of this research was to elucidate the pathway(s) for phytic acid biosynthesis in soybean (*Glycine max*). We have isolated several *myo*-inositol phosphate kinase genes in soybean as possible candidates for steps in the biosynthetic pathway. We have characterized the genes for four *myo*-inositol(1,3,4)P_3 5/6-kinases (*Gm*Itpk1-4), one *myo*-inositol(1,4,5)P_3 6/3/5-kinase (*Gm*Ipk2), and one *myo*-inositol(1,3,4,5,6)P_5 2-kinase (*Gm*Ipk1). We have examined expression in developing seeds and other tissues by Northern blot analysis and quantitative RT-PCR. We have expressed all six genes as tagged fusion proteins in *E. coli*, and verified enzyme activity on the proposed substrates. For each enzyme, we have conducted biochemical characterization to determine enzyme kinetics and substrate specificities. We have verified *in vivo* activity of *Gm*Ipk2 and *Gm*Ipk1 by complementing yeast mutants in the respective genes. Our studies indicate the likelihood that three of the genes may be involved in phytic acid biosynthesis: *Gm*Itpk3,

*Gm*Ipk2 and *Gm*Ipk1. For future work, to more fully understand the contribution of each kinase gene to phytic acid biosynthesis, an RNA interference approach will be employed. The gene sequences identified in this study will be used to construct silencing vectors for use in future transformation of soybean embryogenic cultures to determine the effects of down-regulation on *myo*-inositol phosphate profiles.

ACKNOWLEDGEMENTS

I am very grateful for all the people that helped me with this project and I would like to thank them. First of all, thanks to my advisor Dr. Beth Grabau. She took a chance taking on a student with very little background in molecular biology, and was very patient throughout the learning curve. I am very grateful for her support and guidance over the years.

I've been very lucky to have such a supportive committee of faculty members helping me along the way: Dr. Glenda Gillaspy, Dr. John Jelesko, Dr. Cynthia Gibas and Dr. M.A. Saghai-Maroof. They answered countless questions and their doors were always open.

I'd like to thank all the people that were in the lab when I started graduate school. Jaime Hampton was the first to show me around, teach me how to use the equipment, and was a listening ear for years. She is one of the nicest people I have ever had the pleasure to meet. Dr. Joe Chiera and Dr. Malcolm Livingstone provided endless assistance and advice, and they kept the lab an entertaining place with their constant good humor. I enjoyed countless conversations with Jenny Jenrette and Piyum Khatibi and appreciate all the assistance they provided. Jennifer Bradford and Joseph Johnson were both very talented undergraduates and I wish them the best in graduate school.

I would like to thank my current lab members Javad Torabinejad, Shanna Chriscoe, Nicole Juba, Amanda Gasser, and Haijie Liu. Javad was always willing to help with lab work, answer questions, and to provide much needed encouragement towards the end of the project. Shanna, Nicole, Amanda and Haijie have made the lab a fun and lively place to be. I'm glad that I have had the opportunity to work with each of them.

I'd like to thank the other graduate students for their friendship, knowledge, and allowing me to come in and borrow all sorts of reagents and equipment. So, thanks to Sherry Hildrith, Stacey Simon, Troy Hoff, Shannon Alford, Elitsa Ananieva-Stoyanova, Ryan Burnette, Ryan Anderson, Pete Bowerman, and everyone else, there are too many to name you all.

I would like to thank Ken Hurley for his friendship, support, and detailed answers to all the biochemistry questions I could throw at him. Aaron Thomas and Kevin Slocum have been good friends and ski buddies, and I've enjoyed the good times. Penelope Eisenbies was a fantastic roommate and friend, and I'd like to thank her for all her advice about graduate school and life in general. There are too many people to name, but I'd like to thank the members of the VPI cave club for their friendship, support, and great caving trips.

Most of all, I'd like to thank my family for their continued love and unwavering support. My brother George and his wife Heather have always had a kind word of encouragement, and I love showing off the pictures they send of my nephews, Alex and Christopher. I am incredibly blessed to have such supportive parents. Mom and Dad, your continuing support and encouragement has been simply amazing, and I can't express how much I appreciate it. And, Mom, I'm finally finished writing.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

Literature Review

SOYBEANS

Soybean History

Soybean (*Glycine max*) is the leading oilseed crop in the world today. Over fifty countries produce soybeans, and the United States is the leading producer. *Glycine max* is the cultivated variety of soybean, and it is an annual that grows 75-125 cm in height with purple, pink or white flowers (Evans, 1975). The *Glycine* genus is divided into two subgenera, *Glycine* (perennials) and *Soja* (annuals). *Glycine max,* is in the subgenus *Soja* along with *Glycine soja*, the wild soybean. The genome of *Glycine max* is approximately 1.1 Mbp/1C (Arumuganathan and Earle, 1991) and 40-60% of the genome can be defined as repetitive (Goldberg, 1978; Gurley et al., 1979). *Glycine max* has a chromosome number of 2n = 40 and is believed to have originally been a tetraploid that over time has become diploidized (Shoemaker et al., 1996).

Soybeans are believed to have been first domesticated in northern China during the Shang dynasty (ca. 1700-1100 B.C.) and to have traveled to central and south China by the first century (Hymowitz, 1990). Over the next fifteen centuries they were disseminated through Japan, Indonesia, the Philippines, Vietnam, Thailand, Malaysia, Burma, Nepal and northern India following sea and land trade routes. The soybean was described by Linnaeus in the *Hortus Cliffortianus*, so it is known to have reached the Netherlands by 1737. The earliest recorded introduction to the United States was by Henry Yonge at the request of a former seaman Samuel Bowen. Bowen used the bean to manufacture soy sauce and soybean noodles and received a patent for the manufacturing tools he invented to produce them. In 1851, a packet of soybean

seeds was received as a gift to Dr. Benjamin Franklin Edwards from a Japanese crew rescued at sea. The seeds were planted and two years later distributed to the New York State Agricultural Society, the Massachusetts Horticultural Society, and the commissioner of patents. The seeds were then sent to dozens of farmers and by the late 1850s farmers were evaluating them for forage potential. Between 1860 and 1899, there were more than 180 publications pertaining to the soybean in the USA. In 1907, Bill Morse joined the USDA and began to test soybean cultivars as his first assignment and subsequently devoted his entire 42 year career to fostering the soybean industry in the United States. In the 1920's, expeditions to China, Japan, and Korea by member of the United States Department of Agriculture (USDA) sent about 4,500 soybean accessions back to the United States and today the collection contains about 13,000 cultivars of soybeans (Hymowitz, 1990). During WWII the use of soybeans expanded, and additional uses for soybean meal and oil were generated. From the 1920's until today, the cultivation of soybeans in the US has grown from about 5 million bushels (Hymowitz, 1990) to over 3 billion in 2004 (American Soybean Association, http://www.soystats.com/2005) and the average yield has grown from approximately 11 bushels per acre (Hymowitz, 1990) to more than 40 (American Soybean Association, http://www.soystats.com/2005).

Soybean Use in Livestock Feed

In the United States, almost 90 percent of the soybean meal produced is used to formulate livestock feed (Iowa Soybean Association, http:\\www.soymeal.org). On average, 51 percent is used by poultry and laying hens, 24 percent for swine, and 19 percent for beef and dairy. The soybean meal is highly useful for livestock feed for a number of reasons. It is available year-round, it is consistent in its nutrient content, and no other common plant protein contains a higher

crude protein content. Soybean meal provides a good balance of all the amino acids with the exception of the amino acid methionine, which may be provided as an inexpensive supplement produced by the chemical industry. In 1948, the discovery of vitamin B_{12} ended the necessity for animal protein in livestock diets.

Soybean meal currently represents 98 percent of the plant protein used in poultry feeds (Iowa Soybean Association, http:\\www.soymeal.org). Poultry feeds generally consist of a corn and soybean mix, a methionine supplement, as well as vitamins and minerals. Starter feeds consist of approximately 30 percent soybean meal, grower feeds contain 24 percent, and finisher diets contain18 percent soybean meal. The consumption of poultry in the United States is increasing. Consumption averaged 52 pounds per capita in 1988 annually and by 1998 it had increased to 65.8 pounds per capita.

The swine industry also uses soybean meal as a major component of livestock feed (Iowa Soybean Association, http:\\www.soymeal.org). Similar to poultry feed, a diet of a corn and soybean meal mixture fortified with vitamins and minerals is a common choice. The most common limiting amino acid in swine diets is lysine, and soybean meal has a percentage of lysine exceeded only by peas, fish, blood and milk proteins. The meal is also high in tryptophan, threonine and isoleucine, all of which are limiting in a grain based diet. Cereal grains tend to be high in methionine, the one limiting amino acid in soybean, and this allows the soybean/corn mixtures to provide a successful complement.

Currently the only potential detriment to soybean meal as a component of poultry feed is the presence of the antinutrient phytate, and the levels of two carbohydrates, raffinose and stachyose (Iowa Soybean Association, http:\\www.soymeal.org). Poultry lack the alpha-galactosidase enzyme necessary to metabolize raffinose and stachyose. When the two

oligosaccharides are excreted, they contribute to litter that is wet and sticky and may contribute to poultry leg disorders.

PHYTIC ACID

Overview

Phytic acid, also known as *myo*-inositol hexa*kis*phosphate, Ins(1,2,3,4,5,6)P_6, or InsP_6, is the major storage form of phosphorous in plant seeds. The structure of phytic acid is an asymmetric cyclitol ring with a phosphate group bound at each carbon (Loewus and Murthy, 2000). The asymmetric ring allows for a large variety of *myo*-inositol phosphates to form, twenty of which are found in eukaryotic cells (Majerus, 1992).

The plant seed accumulates phytic acid in protein storage bodies as mixed salts called phytate that chelate a number of mineral cations. During germination, the stored phosphorous and minerals are released from the *myo*-inositol ring by an enzyme or enzymes known as phytase. This release of phosphorous and other minerals is beneficial to the plant because it occurs at a developmental stage at which the seedling is unable to absorb the nutrients from its environment (Loewus and Murthy, 2000).

Lacking the phytase enzyme in their digestive tracts, non-ruminant animals such as swine and poultry are unable to remove phosphates from the *myo*-inositol ring and therefore are unable to utilize the majority of phosphorous found in legumes and cereals. A typical soybean seed contains over 50% more total phosphorous than a cereal grain and approximately 50% of the total phosphorous content in poultry and swine rations is derived from soybean meal (Wilcox

et al., 2000). Although abundant phosphorous is present in the livestock rations, its incorporation into the indigestible phytic acid molecule renders it unavailable to the livestock digestive system and thus their dietary needs are not fulfilled. In order to meet animal growth requirements for phosphorous in animal feed, a variety of methods are currently employed. Livestock diets may be supplemented with inorganic phosphorous in order to provide the animals with sufficient dietary phosphorous, however the unutilized phytate is then excreted and contributes to phosphorous pollution. Microbial phytase may be added to livestock feed in order to break down the phytate pre-consumption, however this is a costly alternative for farmers and a phytase able to withstand the high temperatures and pressure of feed processing is not commercially available. The phytase must be added post-feed processing, which involves investment in additional equipment for application of the enzyme by spraying.

The six attached phosphorous groups on the *myo*-inositol ring form a high density of negative charge, allowing the molecule to bind a number of mineral cations such as iron, zinc, calcium and potassium. In addition to decreasing phosphorous availability, phytic acid in food binds mineral cations and prevents their absorption. The ability of phytic acid to chelate these cations contributes to its negative effect on both livestock and human nutrition.

Structure and Numbering

The numbering system for *myo*-inositol and phytic acid can be explained using Bernie Agranoff's turtle analogy (Irvine and Schell, 2001). Figure 1-1a shows a Haworth projection of *myo*-inositol, while 1-1b shows the chair conformation for *myo*-inositol. If the turtle's right

flipper (1-1c) is designated hydroxyl "1", following the body counter-clockwise represents the "D" numbering conformation. The D conformation will be used throughout this discussion, and is the conformation typically used in current literature. To describe the "L" conformation, one would label the left "flipper" as hydroxyl number "1", and continue clockwise around the molecule. Therefore, D-*myo*-inositol-3-phosphate, the product of the *myo*-inositol phosphate synthase (MIPS) enzyme, refers to the same molecule as L-*myo*-inositol-1-phosphate (Irvine and Schell, 2001).

Functions of Phytic Acid

Research has implicated several roles for phytic acid in addition to its function as a phosphorus storage molecule in plants. In addition its presence in plants, phytic acid is found in yeast and almost all mammalian cells (Vucenik and Shamsuddin, 2006). In yeast, phytic acid is necessary for mRNA export from the nucleus (York et al., 1999) and in humans it is a potential anti-cancer agent (Vucenik and Shamsuddin, 2006). In addition, it functions as an antioxidant by inhibiting Fe^{3+} catalyzed hydroxyl-radical formation (Hawkins et al., 1993), it may have an affect on K^+ channels in guard cells, and it stimulates protein kinases such as those which phosphorylate a protein involved in synaptic vesicle recycling, and another protein kinase that performs end-joining of DNA (Irvine and Schell, 2001). $InsP_6$ is also used as a substrate for the generation of *myo*-inositol pyrophosphates that may be involved in DNA metabolism, chemotaxis, and environmental stress (Seeds et al., 2004).

Many of the studies on the potential roles of $InsP_6$ have been conducted in yeast. In yeast, there appears to be only one significant pathway for the biosynthesis of phytic acid. Three

separate yeast mutants were identified in conjunction with the yeast *gle1* mutant that showed decreased nuclear export of mRNA and decreased $InsP_6$ synthesis. The mutants were identified as enzymes defective in the $InsP_6$ biosynthetic pathway, specifically the phospholipase C enzyme, an Ins(1,4,5)P_3 6/3-kinase, and an Ins(1,3,4,5,6)P_5 2-kinase. In yeast these mutants are termed *plc1*, *ipk2*, and *ipk1* respectively (York et al., 1999). Studies demonstrated that all three independent mutants showed defects in mRNA export, and the common defect among the three independent mutants is the lack of $InsP_6$ production. The fact that the *ipk2* mutant is still unable to export mRNA from the nucleus despite the functional IPK1 enzyme is further evidence that the necessary molecule is $InsP_6$ (Miller et al., 2005). Later studies by Miller et al. (2005) showed that $InsP_6$ synthesized in the cytoplasm of a *S. cerevisiae* cell by an IPK1 protein anchored in the plasma membrane was able to rescue a *gle1-2 ipk1-4* lethal mutant and restore the Gle1-mRNA export pathway (Miller et al., 2005).

$InsP_6$ is a strong antioxidant, due to the phosphate grouping on the 1, 2, and 3 positions of axial-equatorial-axial. It is able to specifically bind iron and inhibit its ability to form hydroxyl radicals (Vucenik and Shamsuddin, 2003). Data show that diets containing a high level of cereals and legumes show a negative correlation with colon cancer, and *in vitro* experiments with a wide variety of cell lines showed $InsP_6$ to be taken up by the cells, dephosphorylated to lower *myo*-inositol phosphates, and to have an anti-proliferative effect on the cells. $InsP_6$ was shown to be effective in reducing the number of tumors and size of tumors in intestinal cancer (Shamsuddin and Ullah, 1989).

Larsson et al. (1997) found that $InsP_6$ was able to inhibit protein phosphatases in insulin-secreting pancreatic β cells in a concentration dependent manner. This inhibition increased the activity of calcium channels. They found two isomers of the lower *myo*-inositol phosphate $InsP_5$

to inhibit the phosphatases as well, but were less effective. The $InsP_4$ molecules had no affect, which indicated it was not simply the presence of the phosphate groups. They tested *myo*-inositol hexasulfate ($InsS_6$) and found it to function at approximately $1/14^{th}$ the effectiveness of $InsP_6$. There are several types of voltage-gated Ca^{2++} channels in insulin-secreting cells, but the L and T types are dominant, and they found that the L type appeared to be involved in the $InsP_6$ inactivation (Larsson et al., 1997). Poyner et al. (1993) showed that $InsP_6$ tends to be localized to membranes, and thus it is in the right place to regulate membrane ion channels (Poyner et al., 1993).

Lemtiri-Chlieh et al. (2000) conducted experiments on *Vicia faba* and *Solanum tuberosum* guard cell protoplasts to examine the affect of $InsP_6$ in the K^+ conductance. They used a patch electrode and found that submicromolar concentrations of $InsP_6$ blocked the inward K^+ current in a dose dependant manner in the same way that ABA and internal calcium had been shown to do. They attempted a similar experiment with other isomers of $InsP_6$ such as the s*cyllo* or *neo*-inositol, and found they did not have an affect. When EGTA was used together with the $InsP_6$, the inhibition of K^+ inward transfer was inhibited indicating that it is Ca^{++} dependent, and they also found that $InsP_6$ was 100x more powerful than $Ins(1,4,5)P_3$ in affecting the K^+ channels (Lemtiri-Chlieh et al., 2000).

Phytic Acid Biosynthetic Pathway

A number of the biosynthetic steps necessary for the formation of phytic acid have been studied in a variety of organisms including slime mold (Stephens and Irvine, 1990), duckweed (Brearley and Hanke, 1996), yeast (Odom et al., 2000), mung beans (Majumder, 1972), maize

(Shi et al., 2003), Arabidopsis (Stevenson-Paulik et al., 2005), drosophila (Seeds et al., 2004), and human (Verbsky et al., 2005). Two main routes have been hypothesized for the formation of phytate: a phosphatidylinositol (PtdIns) dependent pathway, and a PtdIns independent pathway (Figure 1-2).

The first step in both the lipid-dependent and lipid-independent pathways is the same. D-*myo*-inositol-3-phosphate-synthase (MIPS) is the enzyme that acts in the single entryway into production of any of the *myo*-inositol phosphates or phosphoinositides and it is the rate-limiting enzyme for inositol-containing compounds (Majumder et al., 2003). The isomerase converts D-glucose-6-phosphate to D-*myo*-inositol-3-phosphate. The isomerization involves three partial reactions in which two enzyme-bound intermediates have been implicated (Majumder et al., 2003). The MIPS enzyme was first identified in *S. cerevisiae* (Dean-Johnson and Henry, 1989) and has since been reported from more than 60 organisms including archaea, bacteria, green algae, fungi, parasites, higher plants and animals (Majumder et al., 2003). Due to its diversity and antiquity, the gene has thus been used as a model for studying evolution (Majumder et al., 2003). In the higher plants, MIPS has been identified in many species including sesame (Chun et al., 2003), tobacco (Hara et al., 2000), common ice plant (Ishitani et al., 1996), Arabidopsis (Johnson and Sussex, 1995) and duckweed (Smart and Fleming, 1993). The gene is found in multiple copies in some plant species, thus forming a MIPS gene family. In soybean, at least four different MIPS sequences were identified using an EST database search and Southern hybridization studies (Hegeman et al., 2001; Chappell et al., 2006). In maize, seven sequences hybridized to a MIPS probe (Larson and Raboy, 1999), and in Arabidopsis, at least three MIPS sequences have been identified and published (Johnson, 1994; Johnson and Sussex, 1995; Torabinejad and Gillaspy, 2006), while several others are listed in the NCBI database.

Following the formation of Ins(3)P by MIPS, a lipid-independent pathway involving sequential phosphorylation has been described in both *Spirodela polyrhiza* (Brearley and Hanke, 1996) and *Dictyostelium discoideum* (Stephens and Irvine, 1990). In *S. polyrhiza*, the pathway progresses as follows: Ins(3)P $\rightarrow$ Ins(3,4)P_2 $\rightarrow$ Ins(3,4,6)P_3 $\rightarrow$ Ins(3,4,5,6)$P_4$$\rightarrow$ Ins(1,3,4,5,6)P_5 $\rightarrow$ InsP_6. In *Dictyostelium*, the order is slightly different: Ins(3)P $\rightarrow$ Ins(3,6)P_2 Ins(3,4,6)P_3 $\rightarrow$ Ins(1,3,4,6)P_4 $\rightarrow$ Ins(1,3,4,5,6)P_5 $\rightarrow$ InsP_6, but contains several common intermediates. Providing evidence for a lipid-independent pathway in plants, a *myo*-inositol kinase (MIK) enzyme has been identified in maize (Shi et al., 2005). A mutation in the *myo*-inositol kinase in maize results in an approximate 50% reduction in phytic acid levels in seeds. In addition to a lipid-independent pathway in *Dictyostelium*, evidence has been found for the PtdIns-dependent pathway in a nucleus associated fraction, demonstrating that more than one pathway may be in use by a single organism (van Haastert and van Dijken, 1997).

The PtdIns dependent pathway, also known as the lipid-dependent pathway, proceeds through PtdIns(4,5)P_2. PtdIns(4,5)P_2 is cleaved by phospholipase C to form Ins(1,4,5)P_3 and diacylglycerol. This is a widespread reaction used for signaling in eukaryotic cells and Ins(1,4,5)P_3 has been extensively studied for its ability to release calcium from intracellular stores (Berridge, 1993). From Ins(1,4,5)P_3 in *Schizosaccharomyces pombe* and *Dictyostelium*, the InsP_4 intermediate Ins(1,3,4,5)P_4 is formed, and in *S. pombe*, Ins(1,4,5,6)P_4 is found as well (Ongusaha et al., 1998). In *Saccharomyces cerevisiae* and Arabidopsis, a dual specificity InsP_3/InsP_4 6/3-kinase has been identified that phosphorylates Ins(1,4,5)P_3 to Ins(1,3,4,5,6)P_5. The final step is believed to be a kinase which phosphorylates the InsP_5 molecule at the 2nd position (Stephens and Irvine, 1990; Brearley and Hanke, 1996; York et al., 1999).

In humans, an alternate lipid-dependent pathway has been demonstrated (Verbsky et al., 2005). In this pathway, the Ins(1,4,5)P_3 intermediate is believed to be phosphorylated by an Ins(1,4,5)P_3 3-kinase enzyme to Ins(1,3,4,5)P_4 and subsequently dephosphorylated to Ins(1,3,4)P_3. The Ins(1,3,4)P_3 intermediate is phosphorylated at either the 5th or 6th position by an Ins(1,3,4)P_3 5/6-kinase enzyme. The Ins(1,3,4,6)P_4 product is then phosphorylated to InsP_5 by an Ins(1,3,4,6)P_4 5-kinase and finally to InsP_6 via an Ins(1,3,4,5,6)P_5 2-kinase.

In the human biosynthetic pathway, the rate-limiting enzyme is the Ins(1,3,4)P_3 5/6-kinase. This enzyme has been characterized in several other organisms including Arabidopsis (Stevenson-Paulik et al., 2002), *Entomoeba histolytica* (Field et al., 2000), maize (Shi et al., 2003), and bovine (Wilson and Majerus, 1996). In maize, a mutant lacking this enzyme showed an approximate 35% reduction in InsP_6 production (Shi et al., 2003). The enzyme has not been identified in either yeast or Drosophila (Seeds et al., 2004), suggesting that InsP_6 is synthesized via different biosynthetic pathways in different organisms, or that the enzyme is not conserved at the nucleic acid or amino acid level.

Ins(1,3,4,5,6)P_5 2-kinase is the last step in the biosynthesis of phytic acid in both the PtdIns-dependent and independent pathways, and has been identified and characterized in human brain tissue (Verbsky et al., 2002), Arabidopsis (Stevenson-Paulik et al., 2005; Sweetman et al., 2006), *S. cerevisiae* (Ives et al., 2000), and *S. pombe* (Ives et al., 2000).

Only limited information has been reported for the phytic acid biosynthetic pathway in soybean. In studies of immature soybean seed extracts, Phillippy (1998) partially purified a protein with Ins(1,3,4)P_3 5-kinase and Ins(1,3,4,5)P_4 6-kinase activity. The apparent K_m value for Ins(1,3,4)P_3 kinase was 200 nM while for ATP it was 171 μM and the reaction was inhibited by a variety of *myo*-inositol phosphates (Phillippy, 1998). In a similar study using immature

soybean seeds at five weeks post-anthesis, an enzyme with Ins(1,3,4,5,6)P_5 2-kinase activity was also purified (Phillippy et al., 1994). The K_m for this enzyme was 2.3 μM for Ins(1,3,4,5,6)P_5, and 8.4 μM for ATP, and it was also able to phosphorylate Ins(1,4,5,6)P_4 at the 2 position (Phillippy et al., 1994).

Low Phytic Acid Mutants

Mutants with a low phytic acid phenotype have been recovered in a number of cereal and legume species including maize (Raboy et al., 2000), barley (Dorsch et al., 2003), rice (Larson et al., 2000) and soybean (Wilcox et al., 2000; Hitz et al., 2002). These phenotypes showed a 30% to 95% decrease in seed phytate with an accompanying increase in free phosphate. Total seed phosphorous levels remain similar to wild type seeds (Raboy, 2002). Three different *lpa* phenotypes have been described in maize. In the first, known as *lpa1*, the reduction in phytic acid corresponds with a concomitant increase in free inorganic phosphorus. In seeds containing the *lpa2* mutation, the reduction in phytic acid corresponds to an increase in inorganic phosphorous as well as several *myo*-inositol phosphates including $InsP_5$, $InsP_4$, and $InsP_3$ (Raboy et al., 2000). In seeds containing the *lpa3* mutation, the reduction in $InsP_6$ corresponded to an increase in both inorganic phosphorus and *myo*-inositol (Shi et al., 2005).

Rasmussen et al. (1998) mutagenized barley seeds and recovered two plants with the *lpa1* mutation and seven with the *lpa2* mutation. The *lpa1* mutation affected only one recessive locus in the genome, and showed a moderate change in $InsP_6$ levels, while the *lpa2* mutation affected three recessive alleles and had a greater effect on phytate levels. HPLC and NMR studies by Hatzack et al. (2001) on *lpa* barley mutants showed the *myo*-inositol phosphates formed in *lpa2* mutants were as follows: Ins(1,2,3,4)P_4, Ins(1,2,3,4,5)P_5, Ins(1,2,3,4,6)P_5, Ins(1,2,4,5,6)P_5,

Ins(1,2,5,6)P_4, and Ins(1,4,5,6)P_4 (Hatzack et al., 2001). They also found an unusual accumulation of Ins(1,3,4,5)P_4 in the *lpa1* seeds, which is not a normal degradation product of phytic acid, and suggests it may be an intermediate of phytic acid biosynthesis in barley. Additional studies were conducted by Dorsch et al. (2003) on barley *lpa* mutants which demonstrated the following *myo*-inositol phosphates in *lpa2* seeds: Ins(1,2,3,4,6)P_5, Ins(1,2,4,6)P_4 and/or its enantiomer Ins(2,3,4,6)P_4, Ins(1,2,3,4)P_4 and/or its enantiomer Ins(1,2,3,6)P_4, Ins(1,2,6)P_3 and/or its enantiomer Ins(2,3,4)P_3, Ins(1,5,6)P_3 and/or its enantiomer Ins(3,4,5)P_3.

Lpa1 and *lpa2* mutants were isolated in maize by Raboy et al. (2000) with similar phenotypes to those found in barley. In both mutants, total seed phosphorous remained the same while phytic acid decreased 50 to 66%. In the *lpa2* mutants the *myo*-inositol phosphates formed were Ins(1,2,4,5,6)P_5, Ins(1,4,5,6)P_4, Ins(1,2,6)P_3 and/or their respective enantiomers. In maize, the loci for both mutations *lpa2* and *lpa1* map to chromosome 1S and seeds homozygous for either mutation were found to have a dry weight loss of 4 to 23%. One of the seven MIPS homologous sequences was found to map to chromosome 1S in close proximity to the *lpa1* mutation (Raboy et al., 2000). The close proximity of the 1S MIPS gene and *lpa1* mutation indicated that *lpa1* phenotype could be due to a mutation in the MIPS gene. This hypothesis was called into question when the single copy rice and barley MIPS genes were found on chromosomes 3 and 4H respectively, while their *lpa1* mutations were mapped to chromosome 2L in rice, and 2H in barley (Larson et al., 2000).

Three types of *lpa* maize mutants were studied by the laboratory of Shi (Shi et al., 2003; Shi et al., 2005). They discovered the maize *lpa1* mutant was caused by a mutation in a multidrug resistance-associated protein (United States Patent 20060143728). Shi et al. (2003)

identified the *lpa2* mutation in maize as a single nucleotide substitution in the Ins(1,3,4)P_3 5/6-kinase gene. In the mutant gene, a C was replaced by a T at position 158 creating an early stop codon in the gene's open reading frame and thus truncating the protein to only 34 amino acids. The maize Ins(1,3,4)P_3 5/6-kinase gene has activity on several of the *myo*-inositol phosphates including Ins(1,3,4)P_3, Ins(3,5,6)P_3, Ins(3,4,5,6)P_4, and Ins(1,2,5,6)P_4 and its identification as the mutation in the *lpa2* mutant supports its involvement in phytic acid biosynthesis (Shi et al., 2003). The *lpa3* mutant was identified as a mutation in a *myo*-inositol kinase enzyme (Shi et al., 2005).

In soybean, using EMS as a mutagenic agent, two heritable, non-lethal *lpa* mutants have been identified, M153 and M766 derived from breeding line CX1834 (Wilcox et al., 2000). M2 through M6 generation plants were examined for phytic acid levels, inorganic phosphate levels, and changes in the levels of other *myo*-inositol phosphates. It was shown that the increase in inorganic phosphorous correlated with the decrease in phytic acid and there was no change in the presence of other *myo*-inositol phosphates. The seed phytic acid was reduced by approximately 50% in the mutants and they were phenotypically similar to the *lpa1* mutants in rice, maize and barley. Walker et al. (2006) mapped two loci that contributed to the *lpa* phenotype in soybean mutant M153. One locus explained approximately 11% of the variation in seed P_i levels and the other 41%. The interaction between the two loci contributed to another 8-11% of the seed P_i level variation (Walker et al., 2006).

Studies by Hitz et al. (2002) identified a one nucleotide base change in a soybean MIPS gene that causes decreased levels of phytic acid as well as decreased levels of raffinosaccharides. A mutation in the third base pair of the codon encoding residue 396, a lysine, in the MIPS

protein, converted the amino acid to an asparagine and reduced activity of the enzyme by 90% (Hitz et al., 2002).

Nutritional Studies of Lpa Seeds

Nutritional studies have been conducted by a number of research groups using low phytate crops in the diets of livestock such as swine, poultry and fish, as well as test subjects such as rats and humans. Studies have shown that lower phytic acid corresponds to improved calcium and zinc utilization. Sugiura et al. (1999) conducted feeding studies on rainbow trout using *lpa* maize and barley. They found that the fecal phosphorous content decreased by 42.9%, when the grains were combined with low-ash ingredients, and the *lpa* corn showed a significantly higher apparent availability of calcium, iron, zinc and strontium (Sugiura et al., 1999). Mendoza et al. (1998) used *lpa1* maize to conduct feeding studies on humans (14 non-anemic men), measuring the amount of iron absorption from tortillas made with the *lpa* modified corn compared to its parent, wild-type strain. They found that the iron absorption was 49% higher from the tortillas made from the *lpa* corn (Mendoza et al., 1998). Adams et al. (2002) also performed a study on human test subjects using *lpa* maize and showed the fractional absorption of zinc was almost double in subjects consuming the *lpa* maize compared to those consuming wild-type maize (Adams et al., 2002).

Li et al. (2000) conducted *in vivo* studies on broiler chicks using *lpa1* maize as well as *in vitro* studies in a simulated gastrointestinal digestion system. Five separate feeding treatments were assigned to 30 chicks for 21 days, three using the *lpa* corn, and two using wild type corn. Supplementary KH_2PO_4 was added to two of the wild-type diets to raise levels to that of the *lpa*

corn for one treatment, and to the optimal available phosphorous level the other. They found that the chicks fed the *lpa* diets retained more P, and in the diets in which no supplementary KH_2PO_4 was added, the chicks on the *lpa* diet retained more Ca as well. In the *in vitro* studies conducted in the simulated gastrointestinal environment, it was found that 65% of the total P (1,420 mg/kg) was released from the *lpa* corn, while only 23% (543 mg/kg) was released from the wild-type corn (Li et al., 2000). A similar five treatment study was conducted on growing pigs by Veum et al. (2002) in which thirty-five barrows consumed wild-type or *lpa1* barley diets. They found that before supplementation with inorganic P, the *lpa* diets provided higher bone ash weight, bone breaking strength, P and Ca absorption and retention and a reduced P excretion of 55%. However, in the supplementation diets the bone and absorption tests there were no differences, although the reduction in excreted P was still 16% less in the *lpa* diet. In their *in vitro* simulated digestive systems, the P availability was 52% from the *lpa* barley compared 32% for the wild-type (Veum et al., 2002). Spencer et al. (2000) also conducted feeding studies on pigs. In this study, ten different treatments using *lpa* and wild-type corn were consumed by fifty individually penned pigs. They found that when the *lpa* corn was not supplemented with P, the pigs had increased digestibility and retention of phosphorous, while the excretion of P was lower. Their *in vitro* system calculated the bioavailability to be 57% for the *lpa* corn compared to 11% for the wild-type (Spencer et al., 2000).

These data suggest the high potential value for developing soybeans with reduced phytate content while retaining superior agronomic characteristics. Understanding the phytic acid biosynthetic pathway and its role in plant growth and development will be essential for manipulating this important trait.

LATE PATHWAY *MYO*-INOSITOL KINASES

Three late pathway *myo*-inositol kinases have been identified that contribute to phytic acid biosynthesis in plants. A mutation in an Ins(1,3,4)P_3 5/6-kinase decreased phytic acid levels in maize by approximately 30% (Shi et al., 2003), while T-DNA insertions in an Ins(1,4,5)P_3 6/3/5-kinase and in an Ins(1,3,4,5,6)P_5 kinase in Arabidopsis decreased phytate levels by 35% and 83% respectively (Stevenson-Paulik et al., 2005).

Ins(1,3,4)P_3 5/6-kinase

The Ins(1,3,4)P_3 5/6-kinase enzyme has been purified or partially purified from a variety of mammalian tissues including rat liver (Hansen et al., 1988; Abdullah et al., 1992), calf (Wilson and Majerus, 1995), and human brain (Wilson and Majerus, 1996). It has been identified in the plant species *Zea mays* (Shi et al., 2003) and Arabidopsis thaliana (Wilson and Majerus, 1997), and characterized in *Entamoeba histolytica* (Field et al., 2000). In each organism, the enzyme phosphorylates the Ins(1,3,4)P_3 substrate to generate two different InsP_4 products Ins(1,3,4,6)P_4 and Ins(1,3,4,5)P_4 in varying ratios. In mammals, the InsP_4 products are formed in a 3:1 ratio, while in plants the ratio is 1:3. *E. histolytica* generates the two InsP_4 molecules at an almost equal ratio of 1.5:1, however it is the only organism identified thus far that is able to use another *myo*-inositol phosphate, Ins(1,4,5)P_3, as a substrate to generate the tetra*kis*phosphate Ins(1,3,4,5)P_4. In addition to its 5/6-kinase function, the human kinase enzyme has also been shown to have activity as an Ins(3,4,5,6)P_4 1-kinase (Yang and Shears, 2000) and as an Ins(1,3,4,5,6)P_5 1-phosphatase (Shears et al., 2004). The two substrates had similar K_m

values with an apparent K_m of 300 nM for the Ins(1,3,4)P_3 substrate versus the 100 nM for the Ins(3,4,5,6)P_4 substrate. This similarity in kinetic values may allow for competition for the active site of the enzyme. In addition, the 1-kinase substrate Ins(3,4,5,6)P_4, is a known cellular signal, which inhibits chloride-channel conductance. The regulation of Cl- channels allow for processes such as smooth-muscle contraction and neurological transmission. The interaction between the substrate Ins(1,3,4)P_3, an inhibitor of the Ins(3,4,5,6)P_4 1-kinase activity, and the roles of the enzyme as a Ins(3,4,5,6)P_4 1-kinase/ Ins(1,3,4,5,6)P_5 1-phosphatase may allow for physiological regulation of Ins(3,4,5,6)P_4 signaling (Shears et al., 2004).

The apparent K_m values for the enzyme on the Ins(1,3,4)P_3 substrate had a range of 40 nM for the rat liver enzyme (Abdullah et al., 1992) to 500 nM for the enzyme purified from porcine brain (Hughes et al., 1994). An enzyme with Ins(1,3,4)P_3 kinase activity was partially purified from soybean seed extracts by Phillippy et al. (1998). Kinetics on the partially purified enzyme indicated an apparent K_m of 200 nM with the Ins(1,3,4)P_3 substrate, and it was found to be inhibited more than 80% by various *myo*-inositol phosphates at a concentration of 100 μM (Phillippy, 1998).

Initial studies indicated that the Ins(1,3,4)P_3 5/6-kinase may function as a protein kinase in addition to its role as an *myo*-inositol kinase (Wilson et al., 2001), but these results were later determined to be the result of a contaminating protein (Qian et al., 2005).

Ins(1,4,5)P_3 6/3/5-kinase

The Ins(1,4,5)P_3 6/3-kinase gene is a multifunctional kinase that phosphorylates an *myo*-inositol phosphate, Ins(1,4,5)P_3, at the 6th and the 3rd position. The first phosphorylation reaction

forms one of two tetra*kis*phosphate products- Ins(1,3,4,5)P_4 or Ins(1,4,5,6)P_4. The second phosphorylation step forms the Ins(1,3,4,5,6)P_5 penta*kis*phosphate. The Ins(1,4,5)P_3 substrate is a well characterized signaling molecule in animals that triggers the release of intracellular calcium stores (Berridge, 1989), while the Ins(1,3,4,5)P_4 product is also believed to regulate calcium levels by promoting calcium sequestration (Hill et al., 1988). Kinases that phosphorylate the Ins(1,4,5)P_3 signaling molecule have been characterized in a number of organisms including human (Takazawa et al., 1991; Dewaste et al., 2000), rat (Choi et al., 1990; Thomas et al., 1994), chicken (Bertsch et al., 1999), and Arabidopsis (Stevenson-Paulik et al., 2002).

Two isoforms of the Ins(1,4,5)P_3 6/3-kinase gene were identified in Arabidopsis: AtIpk2α and AtIpkβ. The enzymes were able to phosphorylate Ins(1,4,5)P_3 with a K_m of 14.6 μM, Ins(1,3,4,5)P_4 with a K_m of 15.5 μM, and Ins(1,4,5,6)P_4 with a K_m of 31.7 μM and were found to generate an $InsP_5$ product mainly through the Ins(1,4,5,6)P_4 intermediate. In addition, the enzyme was able to phosphorylate the two substrates Ins(1,3,4,6)P_4 and Ins(1,2,3,4,6)P_5 at the 5 position making the enzyme a polyphosphate 6/3/5-kinase (Stevenson-Paulik et al., 2002).

In yeast the Ins(1,4,5)P_3 6/3-kinase is a crucial step in the $InsP_6$ biosynthetic pathway, and disruption of the gene lowers the production of $InsP_6$ 100-fold (Saiardi et al., 2000). It was also discovered in yeast that the *myo*-inositol kinase (designated IPK2) has another function in addition to the phosphorylation of *myo*-inositol phosphates. Odom et al. (2000) found that the IPK2 protein was identical to ARG82, a previously characterized protein in yeast which is a component of the ARGR-MCM1 transcriptional complex (Odom et al., 2000). The ARGR-MCM1 transcriptional complex is comprised of four proteins, ARG80, ARG81, ARG82 and MCM1. The complex forms on site-specific DNA sequences in response to arginine levels. The

activated complex induces two arginine catabolic genes, CAR1 and CAR2, and represses the arginine anabolic genes ARG1, ARG3, ARG8 and the bifunctional ARG5,6 (Bechet et al., 1970; Dubois et al., 1987). With a functional transcriptional complex, wild-type yeast is able to grow on medium with arginine as it sole nitrogen source. Yeast that contains a mutation in the ARG82 protein is unable to express the arginine catabolic genes, and therefore is unable to grow on the arginine medium. One of the promoter elements to which the ArgR-Mcm1 transcriptional complex binds is the Arg5,6 promoter element. The Arg5,6 protein was identified as one of the first examples of a metabolic protein that functions as a transcriptional regulator (Hall et al., 2004), when in fact the Arg5,6 protein itself may be transcriptionally regulated by a metabolic protein, the Ins(1,4,5)P_3 6/3/5-kinase.

Further research was conducted on the protein to determine whether the kinase properties of the enzyme were necessary for its gene regulation function. Experiments conducted by Odom et al. (2000), imply that the kinase activity of the ARG82 is not necessary for formation of the ARGR-MCM1 transcriptional complex on the DNA binding site. The experiments show that the kinase activity is necessary however, for the induction of the CAR1 and CAR2 arginine catabolic genes. To arrive at these results, they developed a yeast *arg82Δ* mutant strain and found that the transcriptional complex did not form on the promoter of the Arg5,6 gene. When the *arg82Δ* mutant yeast strain was "rescued" with a plasmid encoding an Arg82 gene containing a mutation in the kinase region, the complex was then able to form on the Arg5,6 promoter region. Although the complex was bound to the promoter in the complementation strain, neither strain was able to grow on medium with arginine as its sole nitrogen source, indicating that neither strain was able to induce the arginine catabolic proteins and therefore serve as a functional transcriptional regulator. The lab of Xia et al. (2003) conducted a related complementation study

that showed that one of the two identified Arabidopsis Ins(1,4,5)P_3 6/3/5-kinases was able to rescue the *arg82Δ* mutant phenotype. The ability for an Arabidopsis protein to complement the yeast mutant and allow for the assembly of a functional ARGR-MCM1 transcriptional complex opens the possibility that this protein is likely to have a similar function as a transcriptional regulator in higher plants (Xia et al., 2003).

S. cerevisiae arg82Δ mutant strains showed a variety of phenotypic effects including changes in cellular processes including mating, sporulation and growth at 37˚C (Dubois et al., 1987; Qiu et al., 1990). ARG82 may function in several varying roles. It may function as part of a transcriptional complex, as a regulator of other proteins, or its ability to phosphorylate the Ins(1,4,5)P_3 substrate may produce $InsP_4$ and $InsP_5$ signaling molecules. There are several examples of changes in the *myo*-inositol phosphate levels having wide reaching affects. Saiardi et al. (2000) developed a *S. cerevisiae* Arg82 mutant and found that the $InsP_3$ levels increased over 170-fold whiles the $InsP_6$ levels decreased by 100-fold. Phenotypically, they found impaired nuclear export as well as slowed cell growth. When other metabolic enzymes in the $InsP_6$ pathway were disrupted, such as the enzyme that forms the $InsP_3$ substrate for Arg82 (PLC), or the enzyme that phosphorylates the $InsP_5$ product of ARG82, mRNA export was inhibited. These findings indicated that $InsP_6$, or possibly one of the other *myo*-inositol phosphates with altered levels in each mutant played an important role in mRNA export from the nucleus. A study by Steger et al. (2003) showed a very different use for *myo*-inositol phosphates. ARG82 mutants had been found to be defective in the induction of the phosphate responsive gene PHO5. They found that in the ARG82 mutants, chromatin remodeling at the PHO5 promoter was impaired, and the remodeling complexes SWI/SNF and INO80 were not recruited to the promoter region (Steger et al., 2003).

Ins(1,3,4,5,6)P_5 2-kinase

The Ins(1,3,4,5,6)P_5 2-kinase enzyme has been studied in several organisms including yeast (York et al., 1999), human (Verbsky et al., 2002), Arabidopsis (Stevenson-Paulik et al., 2005; Sweetman et al., 2006) and recently in maize (Sun et al., 2007). The kinase is believed to be the final step in phytic acid biosynthesis.

Phillippy et al. (1994) used a five step purification procedure to purify an enzyme with 2-kinase activity from immature soybean seeds. The enzyme was able to phosphorylate Ins(1,3,4,5,6)P_5 and Ins(1,4,5,6)P_4 at the 2 position, however it was unable to phosphorylate any of the other $InsP_5$ isomers. The 52 kD purified enzyme showed a K_m for the $InsP_5$ substrate of 2.3 μM, while having slightly higher K_m of 8.4 uM for ATP and a V_{max} of 243 nmol/min/mg. They calculated a pH optima of 6.8, and a temperature optima of 42˚C. By examining the *myo*-inositol phosphate concentrations of several $InsP_4$ and $InsP_5$ isomers in both immature and mature seeds, they found that the concentrations favored a forward reaction in developing seeds, while favoring the reverse reaction, or ATP generation in mature seeds. This suggests a potential function of the 2-kinase as a generator of ATP in germinating seeds, and phytate as a phosphate store for this purpose (Phillippy et al., 1994).

The $InsP_6$ biosynthetic pathway was described in *Saccharomyces cerevisiae* (York et al., 1999) and the pathway shared a common final step with that in *Dictyostelium discoideum* of the phosphorylation of Ins(1,3,4,5,6)P_5 to $InsP_6$ by a 2-kinase enzyme designated as IPK1. The *S. cerevisiae ipk1Δ* mutant showed an almost complete inability to synthesize $InsP_6$, and showed a decrease in the ability to export mRNA from the nucleus.

Using several motifs from the three previously identified fungal 2-kinases as well as two partial putative fungal sequences, Verbsky et al. (2002) was able to identify a human sequence containing the same motifs. The sequence encoded an Ins(1,3,4,5,6)P_5 2-kinase that was able to phosphorylate the $InsP_5$ substrate with a K_m of 400 nM while having a K_m of 21 μM for ATP. The gene was found on chromosome nine, and able to complement a yeast *ipk1Δ* mutant, and rescue a lethal *gle1-2 ipk1-4* double mutant. Using northern blots to characterize expression, the enzyme was found to be most highly expressed in the human brain, heart, placenta and testis (Verbsky et al., 2002).

Sweetman et al. (2006) identified an Arabidopsis Ins(1,3,4,5,6)P_5 2-kinase based on its homology with the human enzyme (Sweetman et al., 2006). The Arabidopsis sequence was expressed in *E. coli* and found to phosphorylate Ins(1,3,4,5,6)P_5 to $InsP_6$ *in vitro* with a K_m of 22 μM and a V_{max} of 38 nmol/min/mg. The Arabidopsis 2-kinase was able to complement an *ipk1Δ* strain and allow for $InsP_6$ production, as well as rescue the temperature sensitive growth. *In situ* hybridization studies showed the transcript to be highly expressed in the male and female organs of flower buds, as well as in the seeds and siliques. Semiquantative PCR tests demonstrated the transcript to be present in siliques, leaves and cauline leaves (Sweetman et al., 2006). Studies by Stevenson-Paulik et al. (2005) using an *AtIpk1* T-DNA insertion mutant showed the enzyme to be necessary for phosphorus sensing.

OBJECTIVES

The goal of my research is to contribute to the elucidation of the pathway for phytic acid biosynthesis in soybean (*Glycine max*). Understanding phytic acid biosynthesis will aid in the development of strategies for reducing seed phytate levels. My research will focus on three late pathway kinases that are hypothesized to function in two separate branches of the biosynthetic pathway. Specific objectives are:

1. To identify the cDNA sequences for potential late pathway enzymes in soybean phytate biosynthesis, specifically: Ins(1,3,4)P_3 5/6-kinase, Ins(1,4,5)P_3 6/3/5-kinase and Ins(1,3,4,5,6)P_5 2-kinase genes or gene families from the public soybean EST database and compare them to known genes.

2. To examine RNA expression patterns for each gene during seed development.

3. To clone each cDNA sequence into bacterial expression vectors for protein production and purification.

4. To characterize the purified proteins for biochemical properties, including substrate specificity and kinetic parameters.

FIGURES

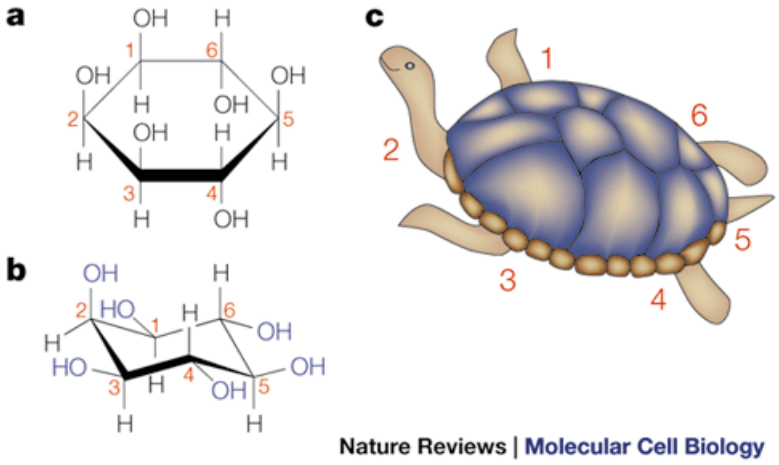

Figure 1-1. An analogy explaining the numbering system for *myo*-inositol in terms of Bernie Agranoff's turtle. a) Haworth projection of D-*myo*-inositol. b) Stair-chair depiction of D-*myo*-inositol c) Turtle representation of D-*myo*-inositol (Taken from Irvine et al., 2001).

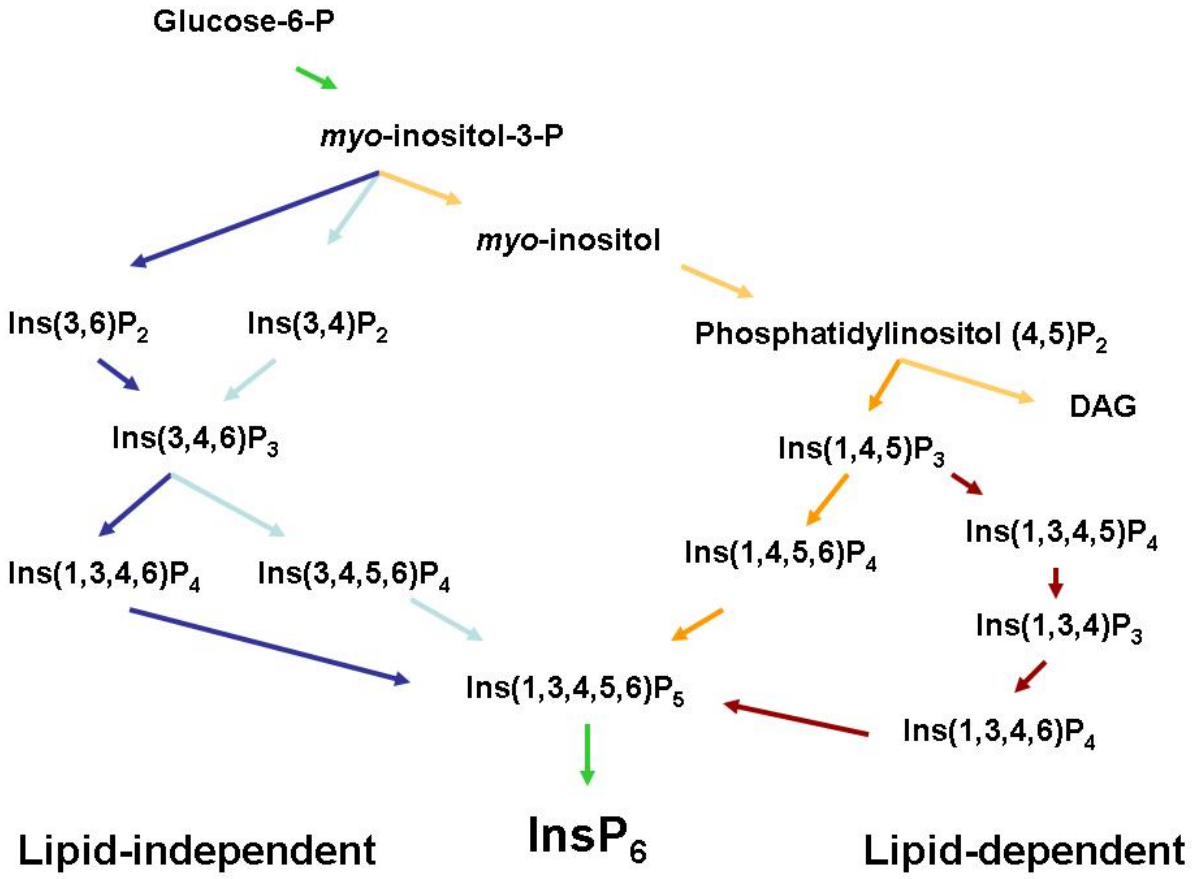

Figure 1-2. Compilation of phytic acid biosynthesis pathways from *D. discoideum*, *S. polyrhiza*, *S. cerevesiae*, and *H. sapiens*. The lipid-independent pathway, depicted on the left, is utilized by *D. discoideum* (dark blue arrows) and *S. polyrhiz*a (light blue arrows). The lipid-dependent pathway, depicted on the right, is utilized by *S. cerevesiae* (orange arrows) and *H. sapiens* (dark red arrows). The initial and final steps are common to all four pathways (green arrows). (Stephens and Irvine, 1990; Brearley and Hanke, 1996; York et al., 1999; Verbsky et al., 2005)

CHAPTER II

Identification and Characterization of an *Myo*-Inositol 3,4,5,6-Tetrakisphosphate 1-Kinase / *Myo*-Inositol 1,3,4-Trisphosphate 5,6-Kinase Gene Family in *Glycine max*

*To be submitted as a manuscript to FEBS Letters by

Stiles AR, Qian X, Shears SB, Grabau EA

INTRODUCTION

Phytate, also known as *myo*-inositol hexa*kis*phosphate or $InsP_6$, is comprised of a six carbon *myo*-inositol ring with a phosphate at each carbon, and is the most abundant *myo*-inositol phosphate found in plants. Soybeans are commonly used in livestock feed, however phytate is poorly digested by non-ruminant animals such as swine and poultry (Reddy et al., 1989; Ravindran et al., 1995). The unavailable phytate phosphorus from animal rations is excreted in manure. Phytate also chelates cations such as calcium, iron, zinc, and potassium, lowering their dietary availability and classifying phytate as an anti-nutrient (Raboy, 2001).

The excretion of unutilized phytate in manure and application of manure to soil as fertilizer can contribute to phosphorous runoff, and the accompanying potential for eutrophication. Eutrophication of lakes and streams results from an influx of excess nutrients into water systems. The presence of excess nutrients, including phosphorus, results in the growth of undesirable algae and aquatic weeds, which is followed by decay and oxygen depletion. In 1996, the EPA identified eutrophication as the main cause of impaired surface water quality, and it has been linked to the outbreaks of *Pfiesteria piscidida,* a dinoflagellate that

has been implicated in fish kills and serious neurological toxicity (Sharpley et al., 2003). Phosphorus is often the limiting element for aquatic growth, and in the United States there is an annual average surplus of 30 lb/acre excess phosphorous (Sharpley et al., 2003).

A long-term goal of this project is to reduce environmental phosphorous pollution by improving phosphorous availability in feed. Selectively modifying key steps in the $InsP_6$ biosynthetic pathway should result in the ability to modulate phytate levels and provide better nutrient availability in plant-based animal diets. A clearer understanding of the plant phytate biosynthetic pathway will be required to achieve this goal. Numerous *myo*-inositol phosphates are found in plant and animal species, many of which are critical for proper cellular function. Pathway modification must avoid detrimental effects on plant growth and development to provide a useful product.

Kinases involved in the phosphorylation of *myo*-inositol phosphates have been studied in several plant species, including mung bean (Majumder, 1972), Arabidopsis (Stevenson-Paulik et al., 2002), and maize (Shi et al., 2003). The $InsP_6$ pathway has also been examined in yeast (Odom et al., 2000) and humans (Verbsky et al., 2005). Two main routes have been hypothesized for the formation of phytate: a phosphatidylinositol (PtdIns)-dependent pathway, and a PtdIns-independent pathway that involves sequential phosphorylation steps.

In both the PtdIns-dependent and -independent pathways, the isomerase D-*myo*-inositol-3-phosphate synthase (MIPS), catalyzes the first step in $InsP_6$ synthesis by converting D-glucose-6-phosphate to D-*myo*-inositol-3-phosphate. This enzyme catalyzes the single entry step into *myo*-inositol phosphate or phosphoinositide production, and it is the rate-limiting enzyme for synthesis of inositol-containing compounds (Majumder et al., 2003).

The synthesis of $InsP_6$ via PtdIns intermediates, also known as the lipid-dependent pathway, proceeds through PtdIns(4,5)P_2, which is cleaved by phospholipase C to form Ins(1,4,5)P_3 and diacylglycerol. Ins(1,4,5)P_3 is a well known signaling molecule that releases calcium from intracellular stores (Berridge, 1993). In certain organisms, including humans, $InsP_3$ may be phosphorylated by a 3-kinase enzyme to form Ins(1,3,4,5)P_4, another known signaling molecule (Majerus, 1992). The Ins(1,4,5)P_3 3-kinase enzyme has not been found in any plant species. In *Saccharomyces cerevisiae* and Arabidopsis, a dual specificity $InsP_3$/$InsP_4$ 6/3-kinase has been identified that sequentially phosphorylates Ins(1,4,5)P_3 to Ins(1,4,5,6)P_4 and then to Ins(1,3,4,5,6)P_5. The final step in phytate biosynthesis in previously characterized systems is the phosphorylation of $InsP_5$ at the 2-position by an Ins(1,3,4,5,6)P_5 2-kinase (Stephens and Irvine, 1990; Brearley and Hanke, 1996; York et al., 1999).

A lipid-independent pathway involving sequential phosphorylation has been described in both *Spirodela polyrhiza* (Brearley and Hanke, 1996) and *Dictyostelium discoideum* (Stephens and Irvine, 1990). In *S. polyrhiza*, the pathway progresses as follows: Ins(3)P → Ins(3,4)P_2 → Ins(3,4,6)P_3 → Ins(3,4,5,6)P_4 → Ins(1,3,4,5,6)P_5 → $InsP_6$. In *Dictyostelium*, the order is slightly different: Ins(3)P → Ins(3,6)P_2 → Ins(3,4,6)P_3 → Ins(1,3,4,6)P_4 → Ins(1,3,4,5,6)P_5 → $InsP_6$, but contains several common intermediates. In addition to this lipid-independent pathway in *Dictyostelium*, evidence has shown a PtdIns-dependent pathway in a nucleus associated fraction, demonstrating that more than one pathway may be in use by a single organism (van Haastert and van Dijken, 1997).

Only limited information has been reported for the phytic acid biosynthetic pathway in soybean. The initial enzyme in the pathway, MIPS, has been identified and characterized as a four member gene family (Hegeman et al., 2001; Chappell et al., 2006). In studies of immature

soybean seeds (Phillippy, 1998), two separate *myo*-inositol kinase activities were partially purified from extracts of developing seeds. One protein demonstrated activity on an Ins(1,3,4)P_3 substrate, and formed two distinct products, Ins(1,3,4,5)P_4 as the major product and Ins(1,3,4,6)P_4 as the minor product. The second protein phosphorylated Ins(1,3,4,5)P_4 to Ins(1,3,4,5,6)P_5 but showed no activity on the Ins(1,3,4)P_3 substrate, indicating the existence of two separate enzymes from these purification procedures. Phillippy et al. (1994) also partially purified a protein from immature soybean seeds that possessed Ins(1,3,4,5,6)P_5 2-kinase activity. Enzymes catalyzing the proposed first and final step in the pathway have thus been identified in soybean, however the intervening steps remain unclear. It has not been determined whether soybeans use a lipid-dependent pathway, a lipid-independent pathway or a combination of both to synthesize phytic acid. In a study of the timing of phytic acid accumulation in soybean, it was shown that trace amounts of phytic acid were detectable in the seeds at 14 days after flowering, followed by a continuous increase from 21 days after flowering until seed maturity (Raboy and Dickinson, 1987).

Mutants with a low phytic acid (*lpa*) phenotype have been recovered in a number of cereal and legume species including maize (Ertl et al., 1998), barley (Rasmussen and Hatzack, 1998; Dorsch et al., 2003), wheat (Guttieri et al., 2003), rice (Larson et al., 2000) and soybean (Wilcox et al., 2000; Hitz et al., 2002). These phenotypes showed a 30% to 95% decrease in seed phytate with an accompanying increase in free phosphate. Total seed phosphorus levels in *lpa* mutants remain similar to levels in wild-type seeds. Three maize *lpa* phenotypes have been described. In the first, known as *lpa1*, the reduction in phytic acid corresponds with a concomitant increase in free inorganic phosphorus (Raboy et al., 2000). In seeds containing the maize *lpa2* mutation, the reduction in phytic acid corresponds to an increase in several *myo*-

inositol phosphates, including $InsP_3$, $InsP_4$, and $InsP_5$, in addition to an increase in inorganic phosphorus (Raboy et al., 2000). Shi et al. (2003) identified the *lpa2* mutation in maize as a single nucleotide substitution in an Ins(1,3,4)P_3 5/6-kinase gene. The mutation created an early stop codon and truncated the coding region to only 34 amino acids. The third maize mutant, *lpa3*, accumulates both *myo*-inositol and free phosphorus, and has been identified as a mutation in a *myo*-inositol kinase enzyme (Shi et al., 2005).

The Ins(1,3,4)P_3 5/6-kinase identified in the maize *lpa2* mutant phosphorylated several *myo*-inositol phosphates including Ins(1,3,4)P_3, Ins(3,5,6)P_3, Ins(3,4,5,6)P_4, and Ins(1,2,5,6)P_4. The maize Ins(1,3,4)P_3 5/6-kinase mutant showed a 30% decrease in seed phytate levels, indicating a role for the enzyme in the biosynthetic pathway. The Ins(1,3,4)P_3 5/6-kinase gene has also been described in Arabidopsis (Wilson and Majerus, 1997), and characterized in *Entamoeba histolytica* (Field et al., 2000). In each organism, the enzyme phosphorylates the substrate Ins(1,3,4)P_3 to generate two $InsP_4$ products, Ins(1,3,4,6)P_4 and Ins(1,3,4,5)P_4, in varying ratios. The human enzyme has been shown to interconvert Ins(1,3,4,6)P_4 and Ins(1,3,4,5)P_4 (Miller et al., 2005).

In addition to its Ins(1,3,4)P_3 5/6-kinase function, the human kinase enzyme functions as an Ins(3,4,5,6)P_4 1-kinase (Yang and Shears, 2000). In human cells, the Ins(3,4,5,6)P_4 molecule functions to inhibit chloride channel conductance and it is inactivated after phosphorylation by the Ins(3,4,5,6)P_4 1-kinase. This finding suggests a role for the kinase in physiological regulation. The K_m values are 300 nM for the Ins(1,3,4)P_3 substrate versus 100 nM for the Ins(3,4,5,6)P_4 substrate. The similarity of the K_m values may allow for competition for the active site of the enzyme (Yang and Shears, 2000).

It has been suggested that the Ins(1,3,4)P_3 5/6-kinase enzyme functions both as a protein kinase that had the ability to phosphorylate transcription factors (Wilson et al., 2001), and was a member of the COP9 signalosome (Sun et al., 2002). However, more recently it appears that the protein kinase activity may have been due to a contaminating protein rather than evidence of a multifunctional enzyme (Qian et al., 2005).

The current study reports the identification of an Ins(1,3,4)P_3 5/6-kinase gene family in *Glycine max* and examines both the expression pattern and activity of each enzyme in developing seeds.

RESULTS

Gene Identification and Comparison

Putative soybean Ins(1,3,4)P_3 5/6-kinase (*Gm*Itpk) sequences were identified from a database search based on their similarity to known genes from *Arabidopsis thaliana* (Wilson and Majerus, 1996) and *Zea mays* (Shi et al., 2003). Thirty-six expressed sequence tag (EST) sequences were identified with homology to the *Arabidopsis* and *Zea mays* genes (Figure 2-1). The EST sequences were arranged into seven contiguous regions, and corresponding soybean cDNA clones were obtained from the Public Soybean Database (Shoemaker et al., 2002) and sequenced. Sequencing results indicated the existence of four complete Ins(1,3,4)P_3 5/6-kinase genes, designated *Gm*Itpk-1 to *Gm*Itpk-4, and one partial sequence (Table 2-1).

The coding region of each complete gene was translated to yield a predicted amino acid sequence using SIXFRAME in Biology Workbench (Subramaniam, 1998). An alignment of the

four predicted soybean protein sequences with those of Arabidopsis and *Zea mays* is shown in Figure 2-2 and shows several highly conserved regions interspersed with regions with less identity (Figure 2-2). The four proteins range in size from 35 kD for *Gm*Itpk2 to 40.1 kD for *Gm*Itpk3 (Table 2-1).

The similarity between amino acid sequences was analyzed using the program ALIGN (Myers and Miller, 1989) in Biology Workbench (Subramaniam, 1998). The sequences show a pattern of pairwise similarity; *Gm*Itpk-1 and *Gm*Itpk-2 (Group A) show a higher sequence similarity to each other than to *Gm*Itpk-3 and *Gm*Itpk-4 (Group B) and vice versa (Figure 2-3). Of the two groups, the Group A proteins have a higher percentage identity to the published Arabidopsis and *Zea mays* protein sequences.

Table 2-1. Four *Gm*Itpk sequences and predicted protein sizes. GenBank accession numbers are as indicated.

Enzyme	Open Reading Frame (bp)	Predicted Protein size (kD)	Accession #
*Gm*Itpk-1	1020	38.5	EU033958
*Gm*Itpk-2	948	35	EU033959
*Gm*Itpk-3	1065	40.1	EU033960
*Gm*Itpk-4	1026	38.4	EU033961

Recombinant Protein Expression and Activity

Each soybean coding sequence was cloned into the protein expression vector pGEX4T-1 (GE Healthcare, Waukesha, WI) for expression in *E. coli.* The expression vector contains an N-terminal GST sequence followed by a thrombin cleavage site. Each of the four fusion proteins

was expressed in bacteria, followed by affinity purification using GST-sepharose beads (GE Healthcare, Waukesha, WI) (Figure 2-4). In the process of purification, it was observed that the Group A proteins (*Gm*Itpk-1, *Gm*Itpk-2) were more soluble than Group B proteins (*Gm*Itpk-3, *Gm*Itpk-4). The hydropathy profiles (data not shown) showed no obvious differences among the four enzymes (Kyte and Doolittle, 1982), however the isoelectric points for *Gm*Itpk-1 and *Gm*Itpk-2 were 5.67 and 5.88, respectively, compared to 7.84 and 6.46 for *Gm*Itpk-3 and *Gm*Itpk-4.

The purified proteins were tested for kinase activity on the substrate [^{3}H]Ins(1,3,4)P_3. [^{3}H]Ins(1,3,4)P_3 was generated using a recombinant 5-phosphatase (see Materials and Methods) to cleave the fifth phosphate from commercially available [^{3}H]Ins(1,3,4,5)P_4. Because the cleavage reaction was incomplete, the substrate contained a mixture of [^{3}H]Ins(1,3,4)P_3 and [^{3}H]Ins(1,3,4,5)P_4. Results of activity assays were analyzed using high performance liquid chromatography (HPLC) as shown in Figure 2-5. Panel A shows a control reaction lacking enzyme illustrating the mixed substrate, while panels B through E demonstrate that all four proteins phosphorylated [^{3}H]Ins(1,3,4)P_3 to the InsP_4 product.

Each protein was also tested for activity on Ins(3,4,5,6)P_4 by collaborators Dr. Stephen Shears and Dr. Xun Qian at the National Institute of Environmental Health Sciences (NIEHS) using [^{3}H]Ins(3,4,5,6)P_4 and analyzing the product using HPLC (see Appendix). Each enzyme was able to phosphorylate the InsP_4 substrate at the 1 position to form Ins(1,3,4,5,6)P_5 (Figure A-3). The enzyme was also tested for its ability to dephosphorylate Ins(1,3,4,5,6)P_5 (data not shown). All four enzymes were able to function as both Ins(3,4,5,6)P_4 1-kinases and Ins(1,3,4,5,6)P_5 1-phosphatases. Previous studies indicated that phosphatase activity of the human Ins(3,4,5,6)P_4 1-phosphatase enzyme was stimulated by Ins(1,3,4)P_3. Tests were

conducted to see if *Gm*Itpk1-4 from soybean functioned in the same manner (Figure A-4). Ins(1,3,4)P_3 did not stimulate the dephosphorylation; on the contrary, a slight inhibition of the dephosphorylation rate was observed in the presence of Ins(1,3,4)P_3, which may indicate competition for the active site of the enzyme (Chamberlain et al., 2007).

Each fusion protein was also tested for kinase activity on the substrate Ins(3,4,6)P_3 (see Appendix). Each enzyme phosphorylated Ins(3,4,6)P_3 to an InsP_4 product (Figure A-5). HPLC analysis confirmed that the InsP_4 molecule was not Ins(3,4,5,6)P_4, and was either Ins(1,3,4,6)P_4 or Ins(2,3,4,6)P_4. It is unlikely to be Ins(2,3,4,6)P_4 because the Ins(1,3,4,5,6)P_5 2-kinase is thus far the only enzyme with the demonstrated ability to phosphorylate at the 2 position, and therefore the product is more likely to be Ins(1,3,4,6)P_4.

Transcript Expression Analysis

Steady-state mRNA levels were examined to determine whether expression levels from the *Gm*Itpk gene family members corresponded with the production of phytic acid in the soybean seed. The first analysis of gene expression during seed development utilized northern blotting to visualize mRNA abundance. It was previously shown that cross-hybridization did not occur between any pairwise combinations of the four cloned genes under the hybridization conditions employed (data not shown). Northern blot results indicated that patterns of expression differed among the four genes in the developing seeds (Figure 2-6). *Gm*Itpk-1 and *Gm*Itpk-2 showed a low level of expression during seed development with a slight decrease in the expression of *Gm*Itpk-1 as the seeds increased in size. *Gm*Itpk-4 was also expressed at a low level for the first three stages of seed development and decreased as the seed approached maturity at the 9-10 mm

size stage. *Gm*Itpk-3 was highly expressed in all four stages of seed development, with the highest level in the 0-4 mm seed size, and a decreased amount in larger seeds. The frequency of *Gm*Itpk-3 sequences found among database entries corroborates the high level of *Gm*Itpk-3 expression. Of the original 35 EST sequences identified for *Gm*Itpk, 15 corresponded to *Gm*Itpk-3 (Figure 2-1).

To more accurately quantify gene expression, RT-PCR was performed for all four family members (Table 2-2 and Figure 2-7). In the real-time PCR experiment, cDNA from each developmental stage was analyzed from three biological replicates and the results were averaged. As in the northern blots, *Gm*Itpk-3 was much more highly expressed (2 to 10 fold at different stages of seed development) compared to the other three genes. The quantitative RNA data also demonstrated a pattern of expression for *Gm*Itpk-3 with an average of 3.13×10^7 transcript copies in 0-4 mm seeds, decreasing by approximately 50% to 1.49×10^7 copies in the 9-10 mm seeds (Table 2-2). As was observed in the northern blots, quantitative PCR results indicated that the other three genes are expressed at lower, less variable levels. The quantitative results showed a higher level of expression for *Gm*Itpk-1 than predicted from the northern blot data. While the northern blot data indicated a low level of *Gm*Itpk-1 expression, the quantitative data showed an average transcript level ranging from approximately one-third to one-half that of *Gm*Itpk-3 as the seeds increased in size. In the real-time PCR experiment, *Gm*Itpk-1 also showed variation in expression among the three biological replicates. This variability of expression could have originated from slight variations of plant growth conditions and may indicate a possible physiological difference such as an inducible response to plant stress.

Table 2-2. Average transcript number and standard deviation for each RT-PCR sample. Seed samples (0-4 mm, 5-6 mm, 7-8 mm, and 9-10 mm) are the averages of three biological replicates, embryogenic tissue culture, leaf, and root samples are the average of two biological replicates.

Average transcript number using 10 ng/ cDNA	Gene	0-4 mm	5-6 mm	7-8 mm	9-10 mm	Embryogenic culture	Leaves	Roots
	*Gm*Itpk1	8.95E+06	9.54E+06	9.48E+06	9.92E+06	4.67E+06	6.36E+06	7.91E+06
	*Gm*Itpk2	1.65E+06	1.82E+06	2.14E+06	1.98E+06	1.49E+06	1.89E+06	2.10E+06
	*Gm*Itpk3	3.13E+07	1.99E+07	1.56E+07	1.49E+07	2.11E+07	2.17E+07	4.46E+07
	*Gm*Itpk4	5.37E+06	6.22E+06	3.34E+06	6.45E+06	7.83E+06	5.61E+06	1.01E+07
Standard deviation for each sample	*Gm*Itpk1	1.03E+06	4.78E+06	4.39E+06	3.67E+06	3.26E+05	1.32E+06	1.23E+06
	*Gm*Itpk2	7.13E+05	5.06E+05	4.32E+05	5.52E+05	1.03E+05	2.75E+05	7.68E+05
	*Gm*Itpk3	1.46E+07	6.47E+06	6.72E+06	7.79E+06	2.15E+07	2.64E+06	5.25E+07
	*Gm*Itpk4	2.77E+06	1.55E+06	2.76E+06	1.52E+06	5.38E+06	2.45E+06	3.10E+06

Enzyme Kinetics

To determine the likely physiological substrate for the kinase enzymes in soybean, we collaborated with Dr. Stephen Shears and Dr. Xun Qian (NIEHS) to perform enzyme kinetics with both Ins(1,3,4)P_3 and Ins(3,4,5,6)P_4 as substrates (Figures A-1, A-2). The V_{max}, K_m and K_{cat} values were determined for each enzyme/substrate pair. The K_m values for Ins(1,3,4)P_3 ranged from 3.7 μM to 46.2 μM while the K_m for Ins(3,4,5,6)P_4 ranged from 0.28 μM to 1.05 μM (Table A-1). All four enzymes had a lower K_m value for the Ins(3,4,5,6)P_4 substrate compared to the Ins(1,3,4)P_3 substrate. Of the four enzymes, *Gm*Itpk-1 had the lowest K_m value with both substrates. *Gm*Itpk-2 and *Gm*Itpk-3 appeared to have similar kinetic values with each substrate, 0.78 μM and 1.05 μM with Ins(3,4,5,6)P_4, and 46.2 μM and 43.4 μM with Ins(1,3,4)P_3 and. *Gm*Itpk-4 showed similar kinetics to *Gm*Itpk-2 and *Gm*Itpk-3 with the Ins(3,4,5,6)P_4 substrate with a K_m value of 0.7μM, however with Ins(1,3,4)P_3, the K_m value was similar to *Gm*Itpk-1 at 4.4 μM.

DISCUSSION

For this report, we identified four genes for the soybean Ins(1,3,4)P_3 5/6-kinase, which represent the first gene family for this enzyme described in plants. Gene expression and biochemical characterizations were performed with the four fusion proteins produced in *E. coli*. The results indicated that one member of the family, *Gm*Itpk-3, was more highly expressed in developing seeds than the other three. The expression pattern was consistent with the timing for accumulation of phytate and indicates a potential role for *Gm*Itpk-3 in seed phytate biosynthesis. However, the other three genes identified in the family were also expressed, but at lower levels. This may indicate functional redundancy of the kinases or additional roles for the Ins(1,3,4)P_3 5/6-kinase enzyme in plants.

A low phytic acid mutant in maize (*lpa2*) showed a 30% reduction in phytate levels due to a mutation in a Ins(1,3,4)P_3 5/6-kinase gene (Shi et al., 2003). The observed reduction of phytate to only 70% of wild type levels resulting from the kinase gene mutation suggests the possible existence of a similar gene family in *Zea mays* and a possible functional redundancy. Alternatively, the remaining phytate in *lpa* mutants may result from an alternative biosynthetic pathway that does not include the Ins(1,3,4)P_3 5/6-kinase enzyme, or from a combination of both possibilities. A BLAST search of the *Zea mays* EST database using the characterized Ins(1,3,4)P_3 5/6-kinase sequence returned 124 EST sequences with an e-value less than one. The majority of these sequences (115) have sequence identity greater than 95% with the originally characterized gene. The remaining nine sequences have a sequence identity ranging from 81% to 94% compared to the published *Zea mays* Ins(1,3,4)P_3 5/6-kinase sequence, differences that may be due to poor sequence data in the database. This suggests the possibility that there is only

one gene in *Zea mays*, or that only one gene is highly expressed. A single gene hypothesis would predict that functional redundancy in maize is unlikely and favors the existence of an alternative pathway.

In Arabidopsis, one Ins(1,3,4)P_3 5/6-kinase enzyme has been characterized and published (Wilson and Majerus, 1997), and a search of the TAIR database shows at least three additional full length sequences that are annotated as Ins(1,3,4)P_3 5/6-kinase family proteins (At2g43980, At4g08170, At4g33770). Although these three Arabidopsis proteins have not yet been characterized to demonstrate function, the presence of an Ins(1,3,4)P_3 5/6-kinase gene family in another plant species seems likely. The four Arabidopsis sequences separate into pairs based on homology, allowing assignment into two groups as in soybean (Group A and Group B). In fact, one pair of predicted proteins in Arabidopsis (Group B) is more similar to the soybean Group B enzymes than to the members of the Arabidopsis enzyme Group A. Using Genevestigator (Zimmermann et al., 2004) to examine expression levels for each gene in Arabidopsis, it was found that two enzymes, one Group A and one Group B enzyme, are most highly expressed. The Group A enzyme has been characterized by Wilson and Majerus (1997), while the sequence of the Group B enzyme (At4g08170) has been analyzed by Chen et al. (2003). Chen et al. (2003) also examined the promoter region of the potential Ins(1,3,4)P_3 5/6-kinase enzyme in Arabidopsis and identified putative dehydration-responsive element/C-repeat (DRE/CRT) cis-acting elements. Concurrent with the promoter analysis, they found the gene to be highly induced by salt and cold, but unresponsive to abcissic acid and drought stress (Chen et al., 2003). Similarly, soybean ESTs corresponding to *Gm*Itpk-3 and *Gm*Itpk-4 originated in libraries of cold stressed germinating shoots, indicating possible roles for the enzymes in stress response.

The *Gm*Itpk enzymes have been examined for their ability to phosphorylate Ins(1,3,4)P_3 at the five or six position of the *myo*-inositol ring. However, the only currently documented pathway for the formation of Ins(1,3,4)P_3 is through a two step route that has not been demonstrated in plants. The first step is the phosphorylation of Ins(1,4,5)P_3, a well studied signaling molecule, at the three position to form Ins(1,3,4,5)P_4. This step is carried out by an Ins(1,4,5)P_3 3-kinase enzyme. The InsP_4 product is then dephosphorylated at the five position by a *myo*-inositol 5-phosphatase. The Ins(1,4,5)P_3 3-kinase has been documented in a number of organisms including both rat (Irvine et al., 1986) and human (Takazawa et al., 1991), but it has not been found in any plant species. When the kinetic data were examined, *Gm*Itpk-3 had a much higher affinity for Ins(3,4,5,6)P_4 than Ins(1,3,4)P_3 as a substrate. Together, these data indicate that Ins(1,3,4)P_3 is unlikely to be the physiological substrate of the *Gm*Itpk enzymes.

Phosphorylation by *Gm*Itpk1-4 at the one position of the Ins(3,4,5,6)P_4 substrate yields the product Ins(1,3,4,5,6)P_5. This reaction forms the substrate for the proposed final step in the pathway, the phosphorylation by an Ins(1,3,4,5,6)P_5 2-kinase. The Ins(1,3,4,5,6)P_5 2-kinase enzyme has been identified in developing soybean seeds (Phillippy et al., 1994), in Arabidopsis (Stevenson-Paulik et al., 2005; Sweetman et al., 2006), and has been characterized in soybean in this study (Chapter III). If the physiological substrate of the *Gm*Itpk-3 enzyme is Ins(3,4,5,6)P_4, the next question to consider is the route for generation of the InsP_4 molecule. The formation of this *myo*-inositol phosphate intermediate has been demonstrated in *Spirodela polyrhiza*, an aquatic plant with the common name duckweed (Brearley and Hanke, 1996) in the previously mentioned lipid-independent pathway (Figure 1-2). It is also possible that this molecule is functional in pathways other than phytic acid biosynthesis. Ins(3,4,5,6)P_4 is an important signaling molecule in human cells (Ismailov et al., 1996), and is likely to function as a signaling

molecule in plants as well. For example, Zonia et al. (2002) identified a role for the molecule in pollen tube growth and Cl^- flux (Zonia et al., 2002). This role for Ins(3,4,5,6)P_4 in signaling suggests a function for the other members of the *Gm*Itpk gene family. The enzyme kinetics comparing the K_m and V_{max}/K_m values for each enzyme indicate a lower apparent K_m value for each of the other enzymes with Ins(3,4,5,6)P_4, and a higher V_{max}/K_m value for *Gm*Itpk1 and *Gm*Itpk2, the group A enzymes. The higher affinity for the Ins(3,4,5,6)P_4 substrate by *Gm*Itpk1, *Gm*Itpk2, and *Gm*Itpk4, and low transcript levels for each in developing seeds may indicate a role for one or more of these three enzymes in signaling pathways.

The most likely substrate of the *Gm*Itpk3 enzyme in the pathway to phytic acid biosynthesis is Ins(3,4,6)P_3. *Gm*Itpk1-4 demonstrated the ability to phosphorylate Ins(3,4,6)P_3 to an InsP_4 product, either Ins(1,3,4,6)P_4 or Ins(2,3,4,6)P_4. As the Ins(1,3,4,5,6)P_5 2-kinase enzyme is the only inositol kinase currently known to phosphorylate at the two position, we believe the product to be Ins(1,3,4,6)P_4. Using this product, two subsequent phosphorylation steps are necessary to synthesize phytic acid. The Ins(1,3,4,6)P_4 product may be phosphorylated at the five position by an Ins(1,3,4,6)P_4 5-kinase to form Ins(1,3,4,5,6)P_5, followed by phosphorylation at the two position by an Ins(1,3,4,5,6)P_5 2-kinase. Both *myo*-inositol kinases have been identified and characterized in soybean (see Chapter III). The Ins(3,4,6)P_3 substrate has been identified as an intermediate in the lipid-independent pathways in both *S. polyrhiza* and *D. discoideum* (Figure 1-2). In addition, the expression pattern of *Gm*Itpk3 in developing seeds supports this hypothesis. Phytic acid is first synthesized early in seed development, and increases in concentration in a linear fashion as the seed matures (Raboy and Dickinson, 1987). If the Ins(3,4,6)P_3 substrate is phosphorylated to Ins(1,3,4,6)P_4 by *Gm*Itpk3, it would be necessary for this enzyme to be transcribed early in seed development to function in the pathway.

The expression pattern is similar to that of the seed-specific MIPS enzyme, also believed to function in the early pathway of phytic acid biosynthesis (Hegeman et al., 2001; Chappell et al., 2006).

This research identifies a previously uncharacterized gene family in soybean and provides evidence for the participation of at least one of the isozymes in the biosynthesis of phytic acid. These data, in combination with previous reports on the expression pattern of the MIPS enzyme in developing soybean seed (Hegeman et al., 2001; Chappell et al., 2006), as well as the recent characterization of an Arabidopsis Ins(1,3,4,5,6)P_5 2-kinase enzyme (Stevenson-Paulik et al., 2005; Sweetman et al., 2006), provide a framework for defining phytic acid biosynthesis in this important feed crop. To fully elucidate the phytate pathway, the enzyme(s) responsible for conversion of Ins(3)P to Ins(3,4,6)P_3 and/or Ins(3,4,5,6)P_4 must be identified. Due to the ubiquitous nature and numerous functions of *myo*-inositol phosphates, the potential for more than one pathway in soybean is likely. Understanding phytic acid biosynthesis is the first step in the long-term goal of targeted pathway modification for development of a feed crop with high phosphorus availability.

MATERIALS AND METHODS

Gene Identification

To identify soybean expressed sequence tag (EST) cDNAs for Ins(1,3,4)P_3 5/6-kinase genes, a BLAST database search was conducted utilizing sequences previously identified in

Arabidopsis thaliana (AF080173) and *Zea mays* (AY172635). The resulting soybean EST clones identified from the National Center for Biotechnology Information (NCBI) database were arranged into sets of overlapping sequences (contigs) using a contig assembly program (CAP3) (Huang and Madan, 1999) to identify potential coding regions. EST clones containing the sequences corresponding to the 5' ends of the putative genes were obtained from the Iowa State University (Iowa State University, Ames, IA) Public Soybean Database (Shoemaker et al., 2002) via Biogenetic Services (Brookings, South Dakota). Plasmid DNA was purified and sent to the University of Chicago Cancer Research Center for sequencing. Open reading frames were predicted for each gene based on homology to previously characterized kinases. Expected protein sizes and isoelectric points were calculated using the Compute pI/MW program (Bjellqvist et al., 1993).

Recombinant Protein Expression

In order to clone the Ins(1,3,4)P_3 5/6-kinase genes for expression in *E. coli*, primers were designed to amplify the complete coding regions and to include the restriction sites *Sma*I or *Eco*RI on the forward primer, and *Xho*I on the reverse primer (Table 2-2). DNA amplification was performed using AccuPrime Pfx SuperMix (Invitrogen, Carlsbad, CA). PCR conditions consisted of denaturation for 30 sec at 95°C, annealing for 30 sec at 55°C and extension for 90 sec at 72°C for 38 cycles. The original public soybean database clones were used for the template DNA. The PCR products were digested with the respective restriction enzymes and ligated (New England Biolabs Inc., Beverly, MA) into the protein expression vector pGEX4T-1 (GE Healthcare, Waukesha, WI). pGEX4T-1 includes an N-terminal GST tag followed by a

thrombin cleavage site upstream of the multiple cloning site. The ligation reaction was transformed into chemically competent TOP10 bacterial cells (Invitrogen, Carlsbad, CA) for replication of the plasmid. The plasmid insert was sequenced at the Virginia Bioinformatics Institute core laboratory facility (Blacksburg, VA) and transformed into chemically competent DH5α cells (Invitrogen, Carlsbad, CA) for protein expression.

The cells were grown to an OD_{600} of 0.5 at 37°C, induced with isopropyl thiogalactoside (IPTG) at a final concentration of 0.4 mM for 4 hours and then harvested by centrifugation at 6000 x g for 15 minutes at 4°C. The cells were resuspended in cold lysis buffer (50 mM Tris-HCl, 100 mM NaCl, 100 μM phenylmethylsulfonyl fluoride, pH 7.4) and passed through a French press (Thermo Fisher Scientific, Waltham, MA) twice at 16,000 psi. Triton-X was added to a concentration of 1% and the lysate incubated on ice for one hour. The lysate was centrifuged at 10,000 x g for 10 minutes to pellet the insoluble material. GST-sepharose beads (GE Healthcare, Waukesha, WI) were added to the soluble fraction using a bed volume of 0.05% of the original culture volume. The beads were incubated on a rotating shaker at 4°C for one hour. The beads were loaded onto a column and washed three times with 10 volumes of cold cell lysis buffer. The proteins were eluted three times with buffers containing increasing concentrations of reduced glutathione (50 mM Tris-HCl, 100 mM NaCl, with 10 mM, 25 mM or 50 mM reduced glutathione, pH 8.0), using an elution volume equal to the bead volume. Protein concentrations were determined using Coommassie Plus Protein Assay Reagent (Pierce Biotechnology, Rockford, IL). The eluted protein fractions were aliquoted and stored in 15% glycerol at -80°C. Purified proteins were analyzed for size and purity on 10% gels by SDS PAGE (Invitrogen, Carlsbad, CA).

Table 2-3. Primer sequences utilized for cloning *Gm*Itpk1-4 sequences into the pGEX4T-1 protein expression vector and for RT-PCR experiments (restriction sites underlined)*.

Primers for cloning into protein expression vector (pGEX4T-1).	*Gm*Itpk-1	FWD	GATCCCCGGGAATGGCGGAGAAGAGATTCG
		REV	CCTACTCGAGTCAAGCTTGAAGAGATTCCTCTT
	*Gm*Itpk-2	FWD	GATCGAATTCATGTCCGAGTCGGAAGTAGCA
		REV	CACACTCGAGCTACGCAGTCTTGGAGCGTA
	*Gm*Itpk-3	FWD	GATCGAATTCATGAGGTTGAGGGAGGAGGTAG
		REV	GATCCTCGAGCTATTTTTTCTTGTACTTCCCCTGC
	*Gm*Itpk-4	FWD	GATCCCCGGGAATGAGGCTAAACGGTGAAATCTC
		REV	GGCCCTCGAGTTAGGCAGCAAGTTTCTTATTA
Quantitative PCR (Each product ~200bp)	*Gm*Itpk-1	FWD	CTGCGAAGTAATGCTCAAGA
		REV	GCAACTCGTGCCAACC
	*Gm*Itpk-2	FWD	TGAGGACGCTGAAATGCC
		REV	AGACAACAGTGTAAATGTGTAATAACATC
	*Gm*Itpk-3	FWD	CCTACTGTTGCTGAGCTTC
		REV	GGGCTTACGTCATGTGGG
	*Gm*Itpk-4	FWD	ATGCCAGGCTATGAGCAC
		REV	ACAGACCCTATTTCCACCTT
	EIFA**	FWD	AGCGTGGTTATGTTGCCTCAAACT
		REV	CTTGATGACTCCCACAGCAACAGT
	G6PDH**	FWD	GAAGAATTGGCCGTTTGGTA
		REV	GCCTTGTCCTTATCGGTGAA

* All primers listed in 5' to 3' orientation. **Housekeeping genes; Elongation factor 1a (EF1A) and Glucose-6-phosphate dehydrogenase (G6PHD).

Myo-inositol Kinase Activity Assays and HPLC Analysis

[^{3}H]Ins(1,4,5)P_3 and [^{3}H]Ins(1,3,4,5)P_4 were purchased from Perkin Elmer (Wellesley, MA). [^{3}H]Ins(1,3,4)P_3 was synthesized using a recombinant 5-phosphatase (generously donated by Dr. Gillaspy at Virginia Tech) to cleave the phosphate at position five from [^{3}H]Ins(1,3,4,5)P_4. The reaction was incubated at 30°C for 16 hours in reaction buffer (50 mM HEPES, 100 mM NaCl, 0.25 mM EDTA, 2 mM $MgCl_2$, pH 7.2) and the product(s) verified using HPLC.

Ins(1,3,4)P_3 5/6-kinase assays were carried out using the method previously described by Wilson and Majerus (1996). Reactions were performed in a volume of 20 µL in reaction buffer (20 mM HEPES pH 7.2, 6 mM $MgCl_2$, 10 mM $LiCl_2$, 100 mM KCl, 10 mM phosphocreatine, 10 units phosphocreatine kinase, 1 mM DTT, 5 mM ATP) and 1 µg of purified enzyme. The reactions were incubated at 37°C for 1.5 hours and terminated by addition of 80 µL of 10 mM $NH_4H_2PO_4$ at pH 3.5, followed immediately by incubation at 100°C for five minutes. The samples were centrifuged at 12,000 x g for two minutes to pellet the denatured proteins prior to HPLC analysis using a Beckman Coulter System Gold HPLC unit (Beckman Coulter Inc., Fullerton, CA). *Myo*-inositol phosphate products were separated using methods previously described (Stevenson-Paulik et al., 2005). The products were separated using a Whatman Partisphere 5 strong anion exchange column (Whatman Inc., Florham Park, NJ) with a linear gradient of 10 mM $NH_4H_2PO_4$ (pH 3.5) to 1.7 M $NH_4H_2PO_4$ (pH 3.5) over 12 minutes followed by 25 minutes of 1.7 M $NH_4H_2PO_4$ (pH 3.5). Tritiated reaction products were detected using an inline β-Ram Model 3 radioisotope detector (IN/US, Tampa, FL). Elution profiles for *myo*-inositol phosphates were determined using [^{3}H]Ins(1,3,4)P_3, [^{3}H]Ins(1,3,4,5)P_4, and [^{3}H]Ins(1,4,5)P_3 as standards.

Activity assays using Ins(3,4,5,6)P_4, Ins(3,4,6)P_3 and Ins(1,3,4,5,6)P_5 were conducted by collaborators Dr. Stephen Shears and Dr. Xun Qian. See Appendix for detailed methods.

After experiments were completed, it was discovered that there was a single base pair change at the end of the cloned *Gm*Itpk-4 sequence compared to the EST cDNA. Nucleotide 1018 is a "T" instead of a "G". This mutation changes the penultimate amino acid in *Gm*Itpk-4 from an alanine to a serine. An analysis of the enzyme crystal structure from ***E. histolytica***

(Miller et al., 2005) indicates this amino acid is not part of the catalytic site, and it is unlikely that it would have an effect on enzyme activity.

Northern Blots

Northern blot analyses were performed using total RNA extracted from soybean seeds at different stages of development as determined by increasing seed size. Soybean seeds were harvested from multiple plants and sorted into four size groups (0-4 mm, 5-6 mm, 7-8 mm, 9-10 mm). Total RNA was extracted using Tri-Reagent according to manufacturer's instructions (MRCgene, Cincinnati, OH) and resuspended in RNAsecure resuspension buffer (Ambion, Austin, TX). Total RNA (10 μg per lane) was separated on a 0.9% glyoxal gel using commercial gel buffer and glyoxal loading dye (Ambion, Austin, TX). Prior to transfer to Nytran SuperCharge membrane (Schleicher and Schuell, Keene, NH), the samples were visualized with ethidium bromide and photographed to document RNA integrity and equal loading of lanes. The blots were transferred using Ambion transfer buffer for two hours at room temperature and UV cross-linked using 120 joules/cm^2. The blots were pre-hybridized for 4 hours at 68°C using modified Church's reagent (0.5 M Na_2HPO_4, 7% SDS, 1 mM EDTA, pH 7.8) (Church and Gilbert, 1984). The blots were hybridized overnight at 68°C in the same buffer with full length DNA probes generated using a Psoralen-Biotin kit (Ambion, Austin, TX). The blots were washed for 10 minutes at room temperature with 100 ml 1xSSC (150 mM NaCl, 15 mM sodium citrate, pH 7.0) , 0.1% sodium dodecyl sulfate (SDS), then washed three times for 10 minutes at 68°C with 0.5xSSC, 0.1% SDS. Chemiluminesence was detected using a Biodetect kit (Ambion, Austin, TX) and film exposure for 16 hours (*Gm*Itpk-1, *Gm*Itpk-2) or 1 hour (*Gm*Itpk-3,

*Gm*Itpk-4).

The soybean plants were grown in the greenhouse using a 16 hour light/8 hour dark photoperiod using high pressure sodium vapor lights for supplemental light. The plants were grown year-round, and the greenhouse temperature was a constant 28°C. Plants were grown in MetroMix 510 (MM510, Scotts Co., Marysville, OH) and fertilized with Miracle-Gro (Stern's Miracle-Gro Products, Port Washington, NY) at one-fourth the manufacturer's recommended rate.

Real-Time PCR Expression Analysis

For RT-PCR analysis, total RNA was extracted from developing soybean seeds using the same procedure as for northern blots. Three biological replicates were analyzed for each gene using tissue samples from multiple plants grown in the greenhouse. Two of the biological replicates were from the soybean cultivar "Jack" and the third from Virginia experimental line V71-370. Total RNA was also extracted from soybean embryogenic tissue culture, soybean leaves, and soybean roots from the soybean cultivar "Jack", and two biological replicates analyzed. Purified RNA was quantified in triplicate using absorbance at 260 nm, separated by gel electrophoresis, and visualized by ethidium bromide. Total RNA was treated with DNase I using a Turbo DNA-free kit per manufacturer's instructions (Ambion, Austin, TX) and tested for residual contaminating DNA by real-time PCR analysis using treated RNA as the template. cDNA first strand synthesis was generated using oligo dT primers (12-18 bp) (Invitrogen, Carlsbad, CA) and the Omniscript cDNA kit (Qiagen, Valencia, CA). Primers were designed using Lightcycler Software (Roche Applied Science, Indianapolis, IN) to amplify fragments of

approximately 200 bp (Table 2-3). For specificity, the reverse primer for each gene was chosen from the 3' untranslated region. Each primer set was tested for amplification of a single product and the predicted product size was verified. Products were separated on a 1% agarose gel prior to use in real-time analysis. To check for equivalent synthesis of cDNA among samples, all twelve cDNA samples were simultaneously analyzed in triplicate using primers for two housekeeping genes, Elongation factor 1a (EF1A) and Glucose-6-phosphate dehydrogenase (G6PHD) (Table 2-3).

Each 50 µL reaction consisted of POWER SYBR green master mix (Applied Biosystems, Foster City, CA), cDNA corresponding to 10 ng of original total RNA, and two primers at a concentration of 200 nM each. Each reaction was performed in triplicate and the results averaged. Analysis of each gene included cDNA samples from each seed development stage, controls lacking template, and a standard curve. The standard curve was generated for each gene using 10^4 to 10^{10} molecules of plasmid DNA containing the gene of interest. PCR conditions began with 2 minutes at 50°C followed by 40 cycles consisting of denaturation for 30 sec at 95°C, annealing for 30 sec at 60°C, and extension for 45 sec at 72°C. Each reaction was followed by a melt curve analysis to ensure single product amplification. Amplification and analysis were conducted using an ABI Prism 7700 Sequence Detection System (Applied Biosystems, Foster City, CA).

Enzyme Kinetics

Enyzme kinetics were conducted for *Gm*Itpk1-4 on substrates Ins(1,3,4)P_3 and Ins(3,4,5,6)P_4 by collaborators Dr. Stephen Shears and Dr. Xun Qian. See Appendix for detailed methods.

ACKNOWLEDGEMENTS

This work was supported by the United States Department of Agriculture National Research Initiative Competitive Grants Program and the United Soybean Board.

FIGURES

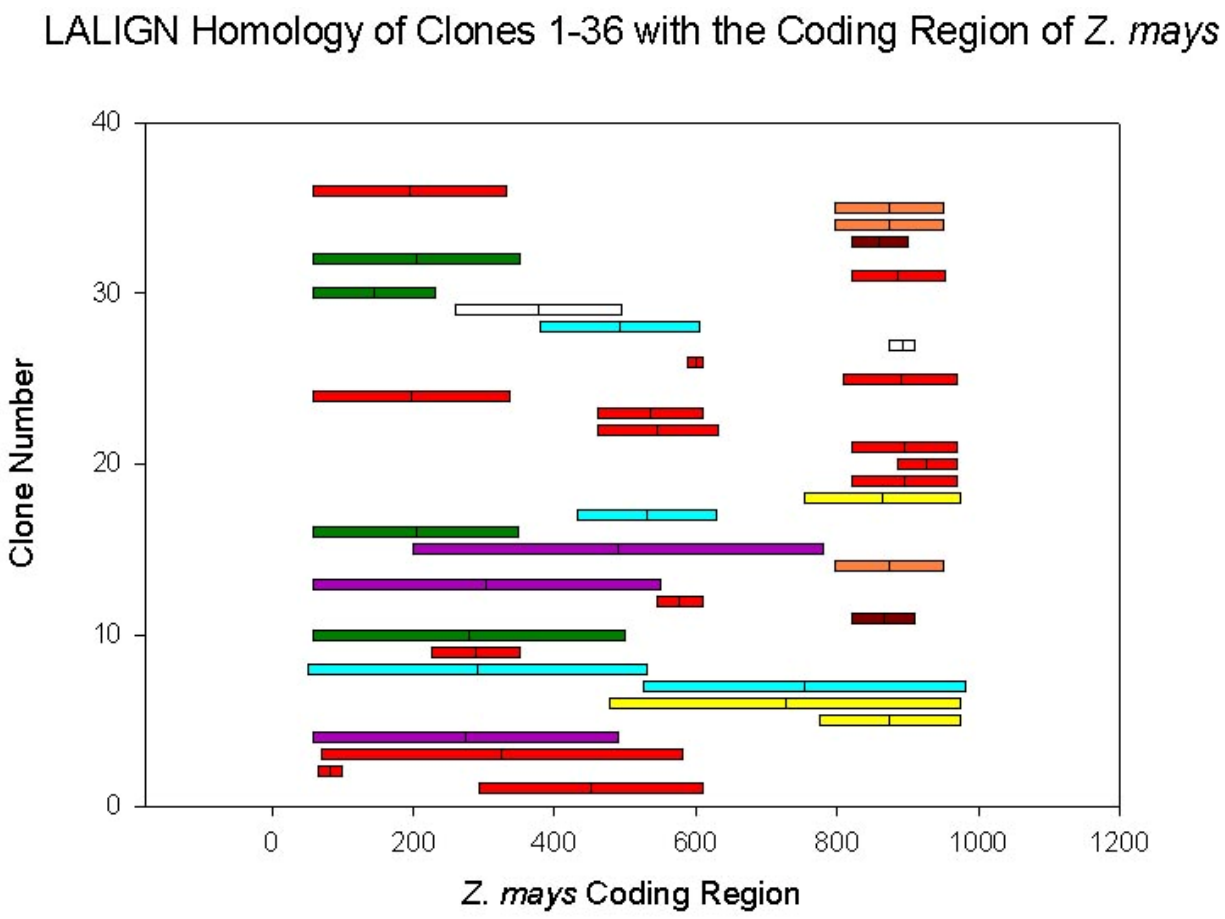

Figure 2-1. Alignment of the 36 identified EST sequences with homology to characterized genes from *Arabidopsis thaliana* (AF080173) (Wilson and Majerus, 1997) and *Zea mays* (AY172635) (Shi et al., 2003). Sequences were grouped into seven contiguous regions and two singletons. An EST clone from each contig was obtained and sequenced. Sequence analysis identified four full genes, and one partial gene sequence. The contigs corresponded to the identified genes as follows: *Gm*Itpk1 = Purple, *Gm*Itpk2 = Lt. Blue, *Gm*Itpk3 = Red, *Gm*Itpk4 = Green and Lt. Brown, *Gm*Itpk5 (partial gene) = Yellow. The two singletons (white) and remaining ESTs (dark brown) did not correspond to an identified gene.

Figure 2-2. Alignment of predicted amino acid sequences for the *Gm*Itpk genes (EU033958 – EU033961). The four soybean sequences were identified based on the homology of EST sequences to previously characterized genes from *Arabidopsis thaliana* (AF080173) (Wilson and Majerus, 1997) and *Zea mays* (AY172635) (Shi et al., 2003).

A.

	*Gm*ltpk-1	*Gm*ltpk-2	*Gm*ltpk-3	*Gm*ltpk-4	*At*ltpk
***Gm*ltpk-1**					
***Gm*ltpk-2**	46.0				
***Gm*ltpk-3**	37.1	38.0			
***Gm*ltpk-4**	39.3	41.9	67.2		
***At*ltpk**	52.0	53.3	36.8	40.2	
***Zm*ltpk**	59.1	47.5	39.8	40.8	52.8

B.

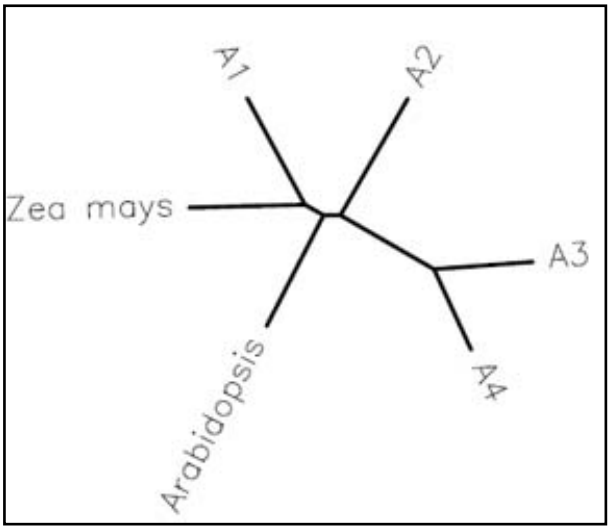

Figure 2-3. Comparison of the identity of the Itpk proteins at the amino acid level. The coding region for each sequence was translated using SIXFRAME in Biology Workbench (Subramaniam, 1998) and compared using ALIGN (Myers and Miller, 1989) in Biology Workbench. Panel A compares the percent identity between each sequence at the amino acid level. Panel B shows an unrooted phylogenetic tree demonstrating the division of the sequences into two separate groups.

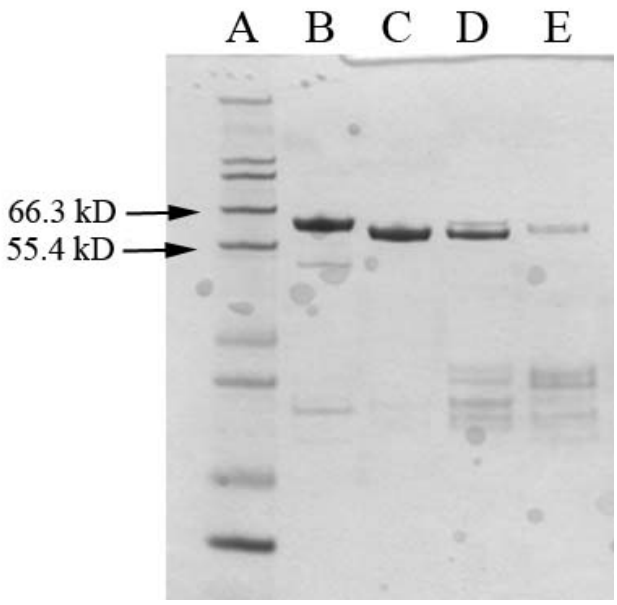

Figure 2-4. SDS-PAGE gel showing purification of the *Gm*Itpk1-4 recombinant proteins. Lane A) Protein size marker (Mark12 prestained protein marker, Invitrogen, Carlsbad, CA). Lanes B-E) GST-tagged *Gm*Itpk1-4, Expected sizes are 64.5 kD, 61 kD, 66.1 kD, and 64.4 kD respectively.

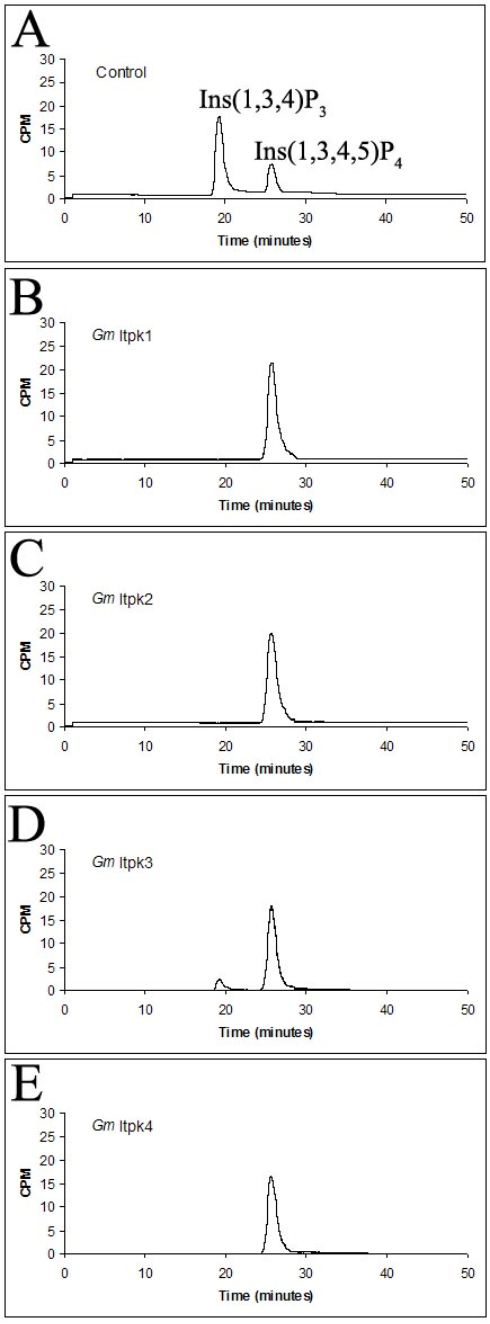

Figure 2-5. HPLC elution profiles of the products of the Ins(1,3,4)P_3 5/6-kinase activity assays. A) Profile of substrate without added enzyme. B – E) Profiles showing the reaction products of the activity assays for enzymes *Gm*Itpk-1 through *Gm*Itpk-4.

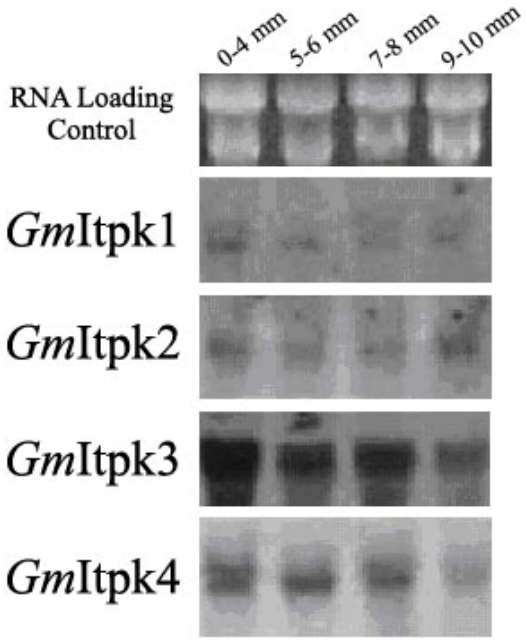

Figure 2-6. Northern blots of total RNA from developing soybean seeds (*Gm*Itpk1-4) using full-length probes. Developing seeds were grouped by size (0-4 mm, 5-6 mm, 7-8 mm, 9-10 mm) for RNA isolation.

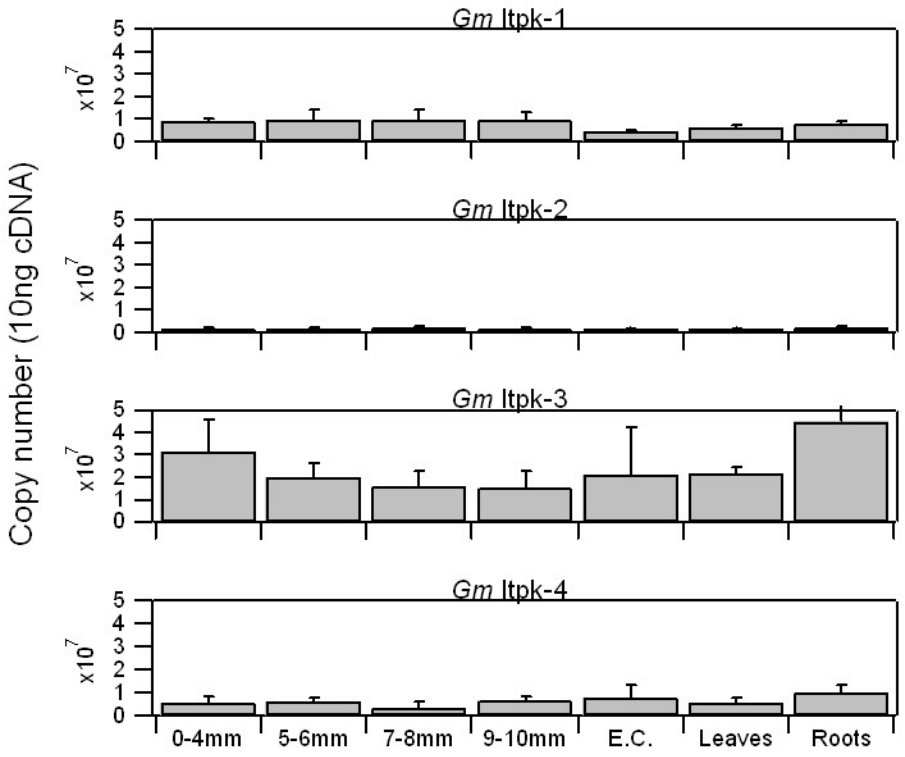

Figure 2-7. RT-PCR analyses quantifying steady state transcript levels of *Gm*Itpk1- *Gm*Itpk4. cDNA was generated from total RNA extracted from developing soybean seeds grouped by size (0-4 mm, 5-6 mm, 7-8 mm, 9-10 mm), embryogenic soybean tissue culture (E.C.), soybean leaves, and soybean roots. Reverse primers were generated from the 3'UTR region of each gene for specificity.

CHAPTER III

Identification and Characterization of *Myo*-Inositol (1,4,5)P_3 6/3/5-Kinase and *Myo*-Inositol (1,3,4,5,6)P_5 2-Kinase in *Glycine max*

*To be submitted as a manuscript to Plant Molecular Biology by

Stiles AR, Torabinejad J, Grabau EA

INTRODUCTION

Phytic acid, also known as *myo*-inositol hexa*kis*phosphate or $InsP_6$, is the major storage form of phosphorus in plant seeds, comprising 1-2% of the total seed dry weight (Raboy, 1997). The negatively-charged six carbon ring containing six phosphate groups forms a mixed salt termed phytate and chelates cations such as iron and zinc. Phytate is stored in protein bodies known as globoids until it is catabolized by the phytase enzyme during seed germination (Raboy, 2001). During germination and seedling development, the stored phosphorus, *myo*-inositol, and mineral cations are utilized until the young plant is able to extract nutrients from the environment.

While beneficial to the young seedling, phytate is considered an anti-nutrient for monogastric animals (Reddy et al., 1989; Ravindran et al., 1995). In the animal digestive system that lacks the phytase enzyme, phytate, along with any chelated mineral cations, travels undigested through the gut and into the environment. Bacteria in the soil break down phytate to its components and the released phosphorus becomes a source of environmental phosphorous pollution. Excess phosphorous in soil can enter waterways through runoff and eventually lead to

eutrophication of lakes and streams (Sharpley et al., 2003). One long-term goal of soybean crop improvement is to develop a low phytic acid (*lpa*) soybean. In order to generate a useful *lpa* soybean that lacks deleterious effects resulting from changes in the *myo*-inositol phosphates, the phytate biosynthetic pathway must first be elucidated.

Two pathways for phytic acid biosynthesis have been proposed, one termed lipid-independent, and the other lipid-dependent. The lipid-dependent pathway proceeds via a PtdIns(4,5)P_2 intermediate, with the cleavage of the lipid to form Ins(1,4,5)P_3 and diacylglycerol. In yeast, the InsP_3 molecule is then phosphorylated to InsP_5 via an Ins(1,4,5)P_3 6/3-kinase and to InsP_6 by a Ins(1,3,4,5,6)P_5 2-kinase (York et al., 1999). In humans the pathway is more complex and proceeds via phosphorylation of Ins(1,4,5)P_3 to Ins(1,3,4,5)P_4 by a 3-kinase, followed by a dephosphorylation step to Ins(1,3,4)P_3, which is then sequentially phosphorylated to InsP_6 (Verbsky et al., 2005).

A lipid-independent pathway has been described in both *Spirodela polyrhiza* (Brearley and Hanke, 1996) and *Dictyostelium discoideum* (Stephens and Irvine, 1990). In *S. polyrhiza*, the pathway progresses as follows: Ins(3)P → Ins(3,4)P_2 → Ins(3,4,6)P_3 Ins(3,4,5,6)P_4 → Ins(1,3,4,5,6)P_5 → InsP_6. In *Dictyostelium*, the order is slightly different: Ins(3)P → Ins(3,6)P_2 → Ins(3,4,6)P_3 → Ins(1,3,4,6)P_4 → Ins(1,3,4,5,6)P_5 → InsP_6, but contains several common intermediates. In addition to this lipid-independent pathway in *Dictyostelium*, evidence has been found for the PtdIns-dependent pathway in a nuclear associated fraction, demonstrating that more than one pathway may be in use by a single organism (van Haastert and van Dijken, 1997).

Little is known about the phytic acid biosynthetic pathway in plants. Current research suggests that rather than a single pathway, multiple pathways may exist in plants involving

enzymes with both kinase and phosphatase activities. Josefsen et al. (2007), found that enzymes with very similar sequences may have different functions in different plant species. In their study, they characterized *myo*-inositol phosphate kinases (IPKs) from both rice and barley and found that although the enzymes shared high sequence homology, the two enzymes were able to phosphorylate and dephosphorylate different *myo*-inositol phosphates (Josefsen et al., 2007). Therefore, the pathway to phytic acid may differ among species.

Several *myo*-inositol kinase enzymes have been identified in plants including *myo*-inositol kinase (Shi et al., 2005), Ins(1,3,4)P_3 5/6-kinase (Wilson and Majerus, 1997), Ins(1,4,5)P_3 6/3/5-kinase (Stevenson-Paulik et al., 2002), and Ins(1,3,4,5,6)P_5 2-kinase (Stevenson-Paulik et al., 2002; Sweetman et al., 2006). However, the role of these kinases is not yet fully understood. Each kinase has the ability to phosphorylate several different *myo*-inositol phosphates *in vitro*, but their functions *in vivo* remain largely unconfirmed.

Stevenson-Paulik et al. (2005) examined the Ins(1,4,5)P_3 6/3/5-kinase gene (*AtIpk2β-1*) and the Ins(1,3,4,5,6)P_5 2-kinase gene (*AtIpk1-1*) using T-DNA insertion mutants in Arabidopsis. Phytate was reduced in the *AtIpk2β-1* mutant by 35%, in the *AtIpk1-1* by 83%, and in the double mutant by greater than 95%.

Myo-inositol-3-phosphate synthase (MIPS) is the first enzyme in the phytic acid biosynthetic pathway. MIPS has been identified and characterized as a four member gene family in soybean (Hegeman et al., 2001; Chappell et al., 2006). In studies of immature soybean seeds, two separate *myo*-inositol kinase activities were partially purified from extracts of developing seeds (Phillippy, 1998). The first protein purified showed activity on an Ins(1,3,4)P_3 substrate, and formed two distinct products, Ins(1,3,4,5)P_4 as the major product and Ins(1,3,4,6)P_4 as the minor product. The second protein activity phosphorylated Ins(1,3,4,5)P_4 to Ins(1,3,4,5,6)P_5 but

showed no activity on the Ins(1,3,4)P_3 substrate, indicating the existence of two separate enzymes from these purification procedures. Phillippy et al. (1994) also partially purified a protein from immature soybean seeds that possessed Ins(1,3,4,5,6)P_5 2-kinase activity. Enzymes catalyzing the proposed first and final step in the pathway have been identified in soybean, but the intermediate steps remain unclear. It has not been determined whether soybeans use a lipid-dependent pathway, a lipid-independent pathway or a combination of both to synthesize phytic acid.

In this study, we identified and characterized an Ins(1,4,5)P_3 6/3/5-kinase (*Gm*Ipk2) and an Ins(1,3,4,5,6)P_5 2-kinase (*Gm*Ipk1) in soybean. To gain greater understanding of their roles in phytic acid biosynthesis, we examined enzyme activity and kinetics using several *myo*-inositol phosphate substrates, determined expression patterns in developing soybean seeds, and confirmed *in vivo* complementation of yeast mutants by the isolated plant gene sequences. If these two enzymes play a vital role in phytic acid biosynthesis in soybean as they do in Arabidopsis, they will be ideal enzymes to target for down-regulation and generation of *lpa* soybeans.

RESULTS

Gene Identification

The soybean Ins(1,4,5)P_3 6/3/5-kinase sequence was identified based on its similarity to two known genes (AY147935, AY147936) from *Arabidopsis thaliana* (Stevenson-Paulik et al., 2002). Using the Arabidopsis sequences, a BLAST search was conducted against the soybean

Expressed Sequence Tag (EST) database at the National Center for Biotechnology Information (NCBI). A cDNA clone corresponding to the putative 5' end of the kinase gene (GenBank Accession BE210763) was obtained from the Public Soybean Database (Shoemaker et al., 2002) and sequenced. Putative start and stop codons were identified based on sequence similarity to the Arabidopsis sequences and the coding region was translated to its predicted amino acid sequence using the program SIXFRAME in Biology Workbench (Subramaniam, 1998). The results indicated the existence of a single complete Ins(1,4,5)P_3 6/3/5-kinase gene (*Gm*Ipk2). The coding region is 840 nucleotides and encodes a protein with a predicted size of 30.9 kD (Table 3-1). An alignment of the amino acid sequence with the two previously characterized Arabidopsis proteins was generated using ClustalW (Figure 3-1) in Biology Workbench. The percentage identity was computed using the program Align in Biology Workbench. The *Gm*Ipk2 protein sequence shares 52.9% identity with AtIpkα and 53.4 % identity with AtIpkβ. AtIpkα and AtIpkβ share 70.4% identity with each other.

A BLAST search was conducted to identify putative plant Ins(1,3,4,5,6)P_5 2-kinase sequences using a partial maize amino acid sequence for Ins(1,3,4,5,6)P_5 2-kinase reported by Verbsky et al. (2002). Several Arabidopsis EST sequences were identified and aligned using the Contig Assembly Program 3 (Huang and Madan, 1999). The resulting sequence was used in a BLAST search of the soybean EST database, and several ESTs identified. An EST corresponding to the proposed 5' region of the soybean gene (AW920567) was ordered from the Public Soybean Database (Shoemaker et al., 2002) and sequenced. Based on alignments with previously characterized genes in other organisms, the coding region for the soybean gene was identified. The coding region is 1371 nucleotides in length and the predicted protein size 51.1 kD (Table 3-1). The EST results indicate the existence of a potential second *Gm*Ipk1 sequence,

but the 5' region was not identified and the sequence not pursued for this study. The predicted amino acid sequence for *Gm*Ipk1 sequence was aligned with the published Arabidopsis and maize sequences (Figure 3-2). *Gm*Ipk1 shares 53.9% identity with the predicted Arabidopsis protein and 50.7% identity with the maize protein. The Arabidopsis and maize proteins share 46.9% amino acid identity. Four Ipk1 motifs were identified by Verbsky et al. (2002), and were also present in the soybean sequence (data not shown).

Table 3-1. *Gm*Ipk2 and *Gm*Ipk1 coding sequences and predicted protein sizes.

Enzyme	Open Reading Frame (bp)	Predicted Protein Size (kD)	Accession #
*Gm*Ipk2	840	30.9	EU033957
*Gm*Ipk1	1371	51.1	EU033956

Recombinant GmIpk2 and GmIpk1 Protein Expression and Activity

The soybean Ins(1,4,5)P_3 6/3/5-kinase sequence was cloned into the protein expression vector pGEX4T-1 for expression in *E. coli* (GE Healthcare, Waukesha, WI). The expression vector encodes a fusion protein containing an N-terminal GST tag followed by a thrombin cleavage site. The GST tag adds 26.7 kD to the predicted protein size for a total fusion protein size of 57.6 kD. The protein was expressed in BL21 *E. coli* cells and purified by affinity purification using GST-sepharose beads (GE Healthcare, Waukesha, WI) (Figure 3-3).

The purified *Gm*Ipk2 protein was tested for kinase activity using [^{3}H]Ins(1,4,5)P_3 as a substrate. Results of activity assays were analyzed using high performance liquid chromatography (HPLC) (Figure 3-4). The enzyme phosphorylated [^{3}H]Ins(1,4,5)P_3 to an InsP_5

product that could be subsequently phosphorylated to $InsP_6$ with the Ins(1,3,4,5,6)P_5 2-kinase enzyme. *Gm*Ipk2 also phosphorylated Ins(1,3,4,5)P_4 and Ins(1,3,4,6)P_4 to $InsP_5$ products.

The soybean Ins(1,3,4,5,6)P_5 2-kinase enzyme was originally cloned into pGEX4T-1, but the resulting protein was highly insoluble in bacterial cells. A small amount of protein was successfully purified, however no enzyme activity could be demonstrated. The coding region was subsequently cloned into the pDEST24 vector (Invitrogen, Carlsbad, CA) that encoded a C-terminal GST tag. The solubility of the expressed protein was not significantly improved but the purified protein successfully phosphorylated Ins(1,3,4,5,6)P_5 in activity assays indicating that the large GST tag at the N-terminus likely interfered with enzyme activity. In order to purify the protein, the coding region was cloned into another protein expression vector, pCOLD1 (Takara Mirus Bio, Madison, WI). This vector contains an N-terminal HIS tag and utilizes a cold inducible promoter to induce protein expression at 16°C. The pCOLD1-*Gm*Ipk1 plasmid was co-expressed in BL21 cells with a chaperone plasmid pTF16 (Takara Mirus Bio, Madison, WI) to aid in solubility. The HIS fusion tag adds approximately 2.1 kD to the size of the predicted protein for a fusion protein of 53.2 kD. The size of the purified protein on SDS-PAGE was consistent with the predicted size (Figure 3-3). The purification showed several contaminating bands, however these bands were also seen in control cultures using an empty vector plasmid.

The purified *Gm*Ipk1 fusion protein was tested in *in vitro* activity assays using Ins(1,3,4,5,6)P_5 as a substrate, and the products were analyzed using HPLC. The recombinant enzyme phosphorylated the Ins(1,3,4,5,6)P_5 substrate to $InsP_6$ (Figure 3-5).

Analysis of RNA Expression for GmIpk2 and GmIpk1

Northern blotting and quantitative PCR were used to determine steady state mRNA levels of the two genes during soybean seed development. The Northern blots showed a very low level of expression of the Ins(1,4,5)P_3 6/3/5-kinase gene during seed fill (Figure 3-6). This low level was observed in multiple experiments, but could be due to probe-specific labeling or hybridization issues. The Ins(1,3,4,5,6)P_5 2-kinase enzyme showed a higher level of mRNA abundance and increasing transcript levels as the seed increased in size. In order to more accurately quantify transcript levels, quantitative RT-PCR was conducted for the two genes (Table 3-2 and Figure 3-7). The quantitative data showed similar results, indicating a low constant level of expression of the Ins(1,4,5)P_3 6/3/5-kinase gene, and a pattern of increasing expression for the Ins(1,3,4,5,6)P_5 2-kinase gene. The quantitative PCR results indicated a higher level of *Gm*Ipk2 expression than observed in the Northern blot data and is considered a more accurate method of analysis.

Table 3-2. Average transcript number and standard deviation for each *Gm*Ipk2 and *Gm*Ipk1 RT-PCR sample. Seed samples (0-4 mm, 5-6 mm, 7-8 mm, and 9-10 mm) are the averages of three biological replicates, embryogenic tissue culture, leaf, and root samples are the average of two biological replicates.

Average transcript number using 10 ng/ cDNA	Gene	0-4 mm	5-6 mm	7-8 mm	9-10 mm	Embryogenic culture	Leaves	Roots
	GmIpk2	4.35E+06	4.71E+06	5.70E+06	3.96E+06	3.46E+06	2.33E+06	5.26E+06
	GmIpk1	1.31E+06	3.77E+06	5.01E+06	9.81E+06	6.14E+05	2.19E+06	1.52E+06
Standard deviation for each sample	GmIpk2	1.25E+06	1.88E+06	3.43E+06	4.42E+05	2.48E+06	1.92E+05	4.10E+06
	GmIpk1	3.10E+05	2.38E+06	2.16E+06	1.54E+06	8.36E+04	6.02E+05	6.11E+04

Enzyme Kinetics

Michaelis-Menten kinetics were performed to compare the activity of each enzyme using various substrates. Each reaction contained excess amounts of ATP and was halted at less than 25% conversion of substrate to product. Velocities were calculated based on nmole converted/sec/mg of recombinant protein and the K_m and V_{max} values were determined using Kaleidagraph software version 4.03 (Synergy Software, Reading, PA). The K_m,V_{max} and catalytic processivity (K_{cat}) values were determined for each enzyme/substrate combination (Tables 3-3 and 3-4).

*Gm*Ipk2 analyses resulted in apparent K_m values for Ins(1,4,5)P_3, Ins(1,3,4,5)P_4 of 5.3 μM and 74.2 μM respectively (Figure 3-8). The K_{cat} values were 1383.9 and 140.9 respectively, indicating that kinetically Ins(1,4,5)P_3 is approximately ten times better as a substrate than Ins(1,3,4,5)P_4. *Gm*Ipk1 was tested with the substrate Ins(1,3,4,5,6)P_5 and had an apparent K_m value of 18.9 μM and a V_{max} value of 350.1 nmol/sec/mg (Figure 3-9).

Table 3-3. Kinetic constants for *Gm*Ipk2 with *myo*-inositol phosphate substrates.

*Gm*Ipk2			
Substrate	K_m (μM)	V_{max} (nmole/sec/mg)	K_{cat} (V_{max} / K_m)
Ins(1,4,5)P_3	5.3	7335.2	1383.9
Ins(1,3,4,5)P_4	74.2	10458	140.9

Table 3-4. Kinetic constants for *Gm*Ipk1 with *myo*-inositol phosphate substrates.

*Gm*Ipk1			
Substrate	K_m (μM)	V_{max} (nmole/sec/mg)	K_{cat} (V_{max} / K_m)
Ins(1,3,4,5,6)P_5	18.9	350.1	18.5

Yeast Complementation

To verify enzyme function *in vivo*, the soybean enzymes were used to complement yeast mutants in the respective genes. Yeast knockout strains in the Ins(1,4,5)P_3 6/3-kinase gene (*ipk2Δ*), and the Ins(1,3,4,5,6)P_5 2-kinase (*ipk1Δ*) were obtained. The *ipk2Δ* was generously provided by Dr. John York (Howard Hughes Medical Institute, Duke University Medical Center Durham, NC) and *ipk1Δ* was purchased from Open Biosystems (Foster City, CA). The coding regions of the soybean *myo*-inositol kinase enzymes were inserted into the yeast expression plasmid pDEST52 and an empty vector control was generated by insertion of a random DNA sequence lacking a start codon.

The yeast cells were grown to late-logarithmic phase and labeled using [^{3}H]*myo*-inositol. The soluble *myo*-inositol phosphates were extracted and analyzed by HPLC (Figure 3-10). The wild-type yeast cells showed a high level of *myo*-inositol and $InsP_2$, and an accumulation of $InsP_6$. The *ipk2Δ* cells showed the same *myo*-inositol phosphate pattern previously reported by other laboratories (Odom et al., 2000), namely an accumulation of $InsP_3$, presumably Ins(1,4,5)P_3, and a high level of $InsP_2$, presumably formed by dephosphorylation of a portion of the accumulated $InsP_3$. Transformation with the pDEST52-Control plasmid had no effect on the mutant phenotype and showed the same *myo*-inositol phosphate pattern as the *ipk2Δ* cells. Transformation of the *ipk2Δ* cells with the pDEST52-*Gm*Ipk2 plasmid restores the production of $InsP_6$, resulting in a *myo*-inositol phosphate pattern similar to the wild-type (Figure 3-10). The pDEST52-GmIpk2 transformed *ipk2Δ* cells showed a peak at $InsP_4$ that did not occur in the wild-type cells. This size of this $InsP_4$ peak correlated to the length of time the cells were allowed to grow. Several incubation times were analyzed (data not shown) and it was noted that

a longer incubation time at late-logarithmic phase allowed for more *myo*-inositol to be incorporated into $InsP_6$, and simultaneously increased the level of unconverted $InsP_4$. This may indicate that the GmIpk2 enzyme has higher enzyme activity on the $InsP_3$ substrate compared to the $InsP_4$ substrate.

The *ipk1Δ* cells showed a high level of accumulation of $InsP_5$ (Figure 3-11). Transformation with the pDEST52-Control plasmid had no effect on the mutant phenotype and showed the same labeled *myo*-inositol phosphate pattern as the *ipk1Δ* cells. Transformation of the *ipk1Δ* cells with the pDEST52-*Gm*Ipk1 plasmid restored the production of $InsP_6$ and produced the same profile as the wild-type.

DISCUSSION

This study identified genes for an Ins(1,4,5)P_3 6/3/5-kinase (*Gm*Ipk2) and an Ins(1,3,4,5,6)P_5 2-kinase (*Gm*Ipk1) in soybean. Based on their similarity to sequences in Arabidopsis and maize, the proteins were hypothesized to phosphorylate specific *myo*-inositol phosphate substrates. To verify their function as *myo*-inositol kinase enzymes, *in vitro* assays were conducted using radiolabeled *myo*-inositol phosphates. The recombinant *Gm*Ipk2 enzyme phosphorylated Ins(1,4,5)P_3, Ins(1,3,4,5)P_4 and Ins(1,3,4,6)P_4 to $InsP_5$ products. The *Gm*Ipk1 enzyme phosphorylated Ins(1,3,4,6)P_4 to an $InsP_5$ product, and Ins(1,3,4,5,6)P_5 to $InsP_6$. In addition to their *in vitro* activities, both enzymes functioned *in vivo*, complementing yeast mutants.

To determine if the two enzymes are expressed during seed development as would be

expected for involvement in phytic acid biosynthesis, steady state mRNA levels were examined during various stages of seed development. The expression pattern of *Gm*Ipk2 showed a low constant level using both Northern blots and quantitative PCR while expression levels of *Gm*Ipk1 increased as the seeds developed. The increasing level of *Gm*Ipk1 mRNA as the seed increased in size corresponds with the timing of phytic acid deposition and provides evidence that the enzyme is part of the phytic acid pathway. The low expression level of *Gm*Ipk2 indicates that it may not be involved in phytic acid biosynthesis in seeds.

The Ins(1,3,4,5,6)P_5 2-kinase enzyme catalyzes the final step in the $InsP_6$ pathway in all organisms for which the pathway has been elucidated, including yeast (York et al., 1999), human (Verbsky et al., 2002; Verbsky et al., 2005), and fruit fly (Seeds et al., 2004) . In Arabidopsis, a T-DNA disruption of the $InsP_5$ 2-kinase results in an 83% decrease in phytic acid levels (Stevenson-Paulik et al., 2005). With this precedent and the observed *Gm*Ipk1 expression pattern in developing soybean seeds, it is likely that *Gm*Ipk1 catalyzes the final step in phytic acid biosynthesis in soybean.

The role of *Gm*Ipk2 in soybean is still unclear. Although the enzyme functions as an *myo*-inositol kinase, it may play only a minor role in $InsP_6$ production in seeds. Several alternate functions have been identified in plants for the IPK2 enzyme. Xu et al. (2005) examined the role of AtIPK2α in Arabidopsis using an antisense approach to down-regulate its expression. They found enhanced pollen germination, pollen tube growth, and root elongation under suboptimal Ca^{++} concentrations. In higher plants Ca^{++} is of special importance during the regulation of pollen tube growth and fertilization (Xu et al., 2005). Disruption of the Ca^{++} gradient results in almost an immediate arrest in pollen tube growth. This supports the hypothesis that AtIpk2α plays a role in the regulation of growth through the regulation of $InsP_3$-mediated calcium

accumulation in plants. More recently it has been shown that the AtIpk2β enzyme plays a role in auxiliary shoot development and is induced by auxin signaling (Zhang et al., 2007).

*Gm*Ipk2 could have multiple roles. In yeast, Ipk2 functions as both a crucial step in the $InsP_6$ pathway (Saiardi et al., 2000) and in an unrelated role as a member of a transcriptional complex (Odom et al., 2000) . Odom et al. (2000) found that the yeast Ipk2 protein was identical to Arg82, a previously characterized protein in yeast that is a component of the ArgR-Mcm1 transcriptional complex (Odom et al., 2000). With a functional transcriptional complex, yeast is able to use arginine as its sole nitrogen source. The Arabidopsis protein AtIpk2β is able to complement yeast lacking the Arg82 gene and create a functional transcriptional complex (Xia et al., 2003).

The enzyme kinetics for each enzyme aided in elucidating their preferred substrates. *Gm*Ipk2 has the lowest K_m value for Ins(1,4,5)P_3, 5.3 μM compared to 74.2 μM for Ins(1,3,4,5)P_4. The K_{cat} values were calculated for each substrate, and were 1383.9 and 140.9 respectively. The K_{cat} value for the Ins(1,4,5)P_3 substrate is approximately 10-fold higher than the K_{cat} value for Ins(1,3,4,5)P_4, indicating that Ins(1,4,5)P_3 is a better substrate. The *Gm*Ipk1 enzyme had an apparent K_m value of 18.9 μM for the Ins(1,3,4,5,6)P_5 and a V_{max} value of 350.1 nmol/sec/mg. The soybean K_m value is similar to that of the Arabidopsis enzyme, determined by two independent labs to be 7.6 μM (Stevenson-Paulik et al., 2005) or 22 μM (Sweetman et al., 2006).

Our previous research identified a four member *Gm*Itpk *myo*-inositol kinase family in soybean (see Chapter II). These enzymes have the ability to phosphorylate Ins(3,4,6)P_3 to Ins(1,3,4,6)P_4, potentially creating a substrate for the *Gm*Ipk2 enzyme, and therefore providing a role for *Gm*Ipk2 in $InsP_6$ biosynthesis. In this case, the later steps of the pathway would be like

those of *D. discoideum* (Stephens and Irvine, 1990); $Ins(3,4,6)P_3 \rightarrow Ins(1,3,4,6)P_4 \rightarrow Ins(1,3,4,5,6)P_5 \rightarrow InsP_6$. The *Gm*Itpk enzymes may alternately function *in vivo* as $Ins(3,4,5,6)P_4$ 1-kinase enzymes. In that case, the last two steps in the pathway would then be identical to those found in *S. polyrhiza* (Brearley and Hanke, 1996), carried out by *Gm*Itpk and *Gm*Ipk1, $Ins(3,4,5,6)P_4 \rightarrow Ins(1,3,4,5,6)P_5 \rightarrow InsP_6$. In the second hypothesized pathway, *Gm*Ipk2 would be unnecessary. It is possible that rather than a single linear pathway to soybean $InsP_6$ production, multiple pathways may be utilized as in *D. discoideum* (van Haastert and van Dijken, 1997).

This study identifies *Gm*Ipk1 as the likely final step in the phytic acid biosynthetic pathway in soybean and identifies and characterizes the *Gm*Ipk2 enzyme. To further investigate the pathway in soybean, down-regulation of individual *myo*-inositol kinase enzymes and analysis of the resulting *myo*-inositol phosphate profile will be necessary.

MATERIALS AND METHODS

Gene Identification

To identify the soybean $Ins(1,4,5)P_3$ 6/3/5-kinase sequence and the soybean $Ins(1,3,4,5,6)P_5$ 2-kinase sequence, BLAST searches were conducted using previously characterized sequences in Arabidopsis and/or maize against the soybean Expressed Sequence Tag (EST) database at the National Center for Biotechnology Information (NCBI). The resulting EST sequences were aligned using Contig Assembly Program 3 (Huang and Madan, 1999). Corresponding soybean cDNA clones were obtained from the Public Soybean Database

(Shoemaker et al., 2002) and sequenced. The coding region of each gene was translated to its predicted amino acid sequence using SIXFRAME in Biology Workbench (Subramaniam, 1998). The similarity of each amino acid sequence to the Arabidopsis and/or maize genes were analyzed using the program ALIGN (Myers and Miller, 1989) in Biology Workbench.

Recombinant Protein Expression of Soybean Ins(1,4,5)P_3 6/3/5-Kinase

To clone the soybean Ins(1,4,5)P_3 6/3/5-kinase gene for expression in *E. coli*, primers were designed to amplify the complete coding regions and include restriction sites *Eco*RI and *Xho*I on the forward and reverse primers, respectively (Table 3-5). Amplification was performed using AccuPrime Pfx SuperMix (Invitrogen, Carlsbad, CA). PCR conditions consisted of 38 cycles consisting of denaturation for 30 sec at 95°C, annealing for 30 sec at 55°C and extension for 90 sec at 72°C. The original public soybean database clones were used for the template DNA. The PCR products were digested with the respective restriction enzymes and ligated (New England Biolabs Inc., Beverly, MA) into the protein expression vector pGEX4T-1 (GE Healthcare, Waukesha, WI). pGEX4T-1 includes an N-terminal GST tag followed by a thrombin cleavage site. The ligation reaction was transformed into chemically competent TOP10 bacterial cells (Invitrogen, Carlsbad, CA) for replication of the plasmid. The purified plasmid was sequenced at the Virginia Bioinformatics Institute (Blacksburg, VA) and transformed into chemically competent BL21 cells for protein expression.

The cells were grown to an OD_{600} of 0.5 at 37°C, induced with 0.4 mM isopropyl thiogalactoside (IPTG) for five hours and then harvested by centrifugation at 6000 x g for 15 minutes at 4°C. The cells were resuspended in 15 ml Bugbuster Protein Extraction Reagent, 15

µl Benzonase, and 4 µl rLysozyme (EMD Biosciencs, Inc., San Diego, CA). The cells were incubated at room temperature for 30 minutes with gentle agitation. The lysate was centrifuged at 16,000 x g for 20 minutes at 4°C to pellet the insoluble fraction. GST-sepharose beads (GE Healthcare, Waukesha, WI) were added to the soluble fraction using a bed volume of 0.05% of the original culture volume. The beads were incubated on a rotating shaker at 4°C for one hour. The beads were then loaded onto a column and washed three times with 10 bed volumes of cold phosphate buffered saline (pH 7.4). The proteins were eluted in buffers containing increasing concentrations of reduced glutathione (50 mM Tris-HCl, 100 mM NaCl, with 10 mM, 25 mM or 50 mM reduced glutathione, pH 8.0), with an elution volume equal to the bead volume. Protein concentrations were determined according to manufacturer's instructions using Coommassie Plus Protein Assay Reagent (Pierce Biotechnology, Rockford, IL). The elutions were aliquoted and stored in 15% glycerol at -80°C. Purified proteins were analyzed for size and purity on 10% gels by SDS-PAGE (Invitrogen, Carlsbad, CA).

Table 3-5. Primer sequences utilized for cloning *Gm*Ipk2 and *Gm*Ipk1 kinase sequences into the pGEX4T-1 and pCOLD1 protein expression vectors and for RT-PCR experiments (restriction sites underlined)*.

Primers for cloning into protein expression vectors*	*Gm*Ipk2	FWD	GACTGGATCC ATGCTCAAGATCCCGGAGC
		REV	ATGATCGCGGCCGCTCAGTCACTTGTGAAGACATGCTAC
	*Gm*Ipk1	FWD	GCGCGACATATGATGGCATTGACTTTGAAA
		REV	CGCAGTCGACTCAATATGCAGCATTAGAT
Quantitative PCR (product size ~200 bp)	*Gm*Ipk2	FWD	CTTGGGTGGCCTTTGT
		REV	AGATTGCCTCTCAGTCCAT
	*Gm*Ipk1	FWD	AGAAGATAGTGAACTGCTACAGAC
		REV	TTCAGGAGGCAACAAATACTAGAT
	EIFA**	FWD	AGCGTGGTTATGTTGCCTCAAACT
		REV	CTTGATGACTCCCACAGCAACAGT
	G6PDH**	FWD	GAAGAATTGGCCGTTTGGTA
		REV	GCCTTGTCCTTATCGGTGAA

* All primers listed in 5' to 3' orientation. **Housekeeping genes; Elongation factor 1a (EF1A) and Glucose-6-phosphate dehydrogenase (G6PHD).

Recombinant Protein Expression of Soybean Ins(1,3,4,5,6)P_5 2-kinase

To clone the soybean Ins(1,3,4,5,6)P_5 2-kinase gene for expression in *E. coli*, primers were designed to amplify the complete coding regions and to include the restriction sites *Nde*I and *Sal*I on the forward and reverse primers, respectively (Table 3-5). The reverse primer was designed to include a C-terminal 6x-HIS tag. Amplification was performed using AccuPrime Pfx SuperMix (Invitrogen, Carlsbad, CA). PCR conditions consisted of 38 cycles of denaturation for 30 sec at 95°C, annealing for 30 sec at 55°C and extension for 90 sec at 72°C. The PCR products were digested with the respective restriction enzymes and ligated (New England Biolabs Inc., Beverly, MA) into the protein expression vector pCOLD1 (Takara Mirus Bio, Madison, WI). The ligation reaction was transformed into chemically competent TOP10 bacterial cells (Invitrogen, Carlsbad, CA) for replication of the plasmid. The purified plasmid was sequenced at the Virginia Bioinformatics Institute (Blacksburg, VA). The plasmid was then transformed into competent BL21 cells containing chaperone plasmid pTF16 (Takara Mirus Bio, Madison, WI) for protein expression.

The cells were grown to an OD_{600} 1.0 at 37°C in LB containing 100 μg/ml ampicillin, 20 μg/ml chloramphenicol, 0.05% arabinose, and 0.01% glucose followed by 30 minutes of cold shock at 15°C. The cells were induced with 20 μM isopropyl thiogalactoside (IPTG) for 16 hours at 16°C and then harvested by centrifugation at 6000 x g for 15 minutes at 4°C. The cells were resuspended in 30 ml binding buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidizole, pH 8) and passed twice through a French Press at 16,000 psi. The lysate was incubated on ice for one hour with 2% Triton-X and centrifuged at 16,000 x g for 15 minutes at 4°C to pellet the insoluble fraction. The soluble fraction was incubated with NiNTA beads (Qiagen, Valencia,

CA) for two hours at 4°C with gentle agitation. The beads were loaded onto a column, and washed with 50 volumes of cold wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidizole, pH 8). The proteins were eluted in buffers containing increasing concentrations of imidizole (50 mM NaH_2PO_4, 300 mM NaCl, with 50 mM or 250 mM imidizole, pH 8), using an elution volume equal to the bead volume. Protein concentrations were determined according to manufacturer's instructions using Coommassie Plus Protein Assay Reagent (Pierce Biotechnology, Rockford, IL). The eluted fractions were aliquoted and stored in 15% glycerol at -80°C. Purified proteins were analyzed for size and purity on 10% gels by SDS-PAGE (Invitrogen, Carlsbad, CA).

Prior to successful purification using the pCOLD1 vector and chaperone system from Takara (Takara Mirus Bio, Madison, WI), several other methods were attempted. This included expression in pGEX4T-1 (GE Healthcare, Waukesha, WI), pET28a (Invitrogen, Carlsbad, CA), and pDEST24 (Invitrogen, Carlsbad, CA). These vectors were transformed into several different *E. coli* cell lines including BL21(DE3), Rosetta™, BL21(DE3)pLysS, and Origami™ (EMD Biosciences, San Diego, CA). Induction parameters were adjusted including OD_{600} at induction (ranging from 0.5-1), concentration of IPTG (ranging from 100 μM to 1 mM), induction time (ranging from four hours to overnight), induction temperature (18°C or 37°C), and the addition of additional media at the time of induction. In addition, six types of media were tested using a media optimization kit from Athena Enzyme Systems (AthenaES, Baltimore, MD) and inclusion body purification was attempted using BugBuster® lysis solution (EMD Biosciences, San Diego, CA). These methods met with limited success and did not allow purification of sufficient protein for kinetic assays.

Myo-Inositol Kinase Activity Assays and HPLC Analysis

Unlabeled *myo*-inositol phosphate substrates were purchased from Inositols.com (Division of AG Scientific, San Diego, CA) and Cayman Chem (Ann Arbor, MI). [^{3}H]Ins(1,4,5)P_3 and [^{3}H]Ins(1,3,4,5)P_4 were purchased from Perkin Elmer (Wellesley, MA). The [^{3}H]Ins(1,3,4,5,6)P_5 and [^{3}H]Ins(1,3,4,6)P_4 substrates were generated using purified recombinant enzymes. To generate [^{3}H]Ins(1,3,4,5,6)P_5, 100 µl reactions were conducted using reaction buffer (50 mM HEPES pH 7.5, 100 mM KCl, 10 mM $MgCl_2$, 5 mM ATP) and an excess of the Ins(1,4,5)P_3 6/3/5-kinase enzyme. The reaction was incubated for 1-2 hours at 37°C, adding additional enzyme as necessary. The reaction was terminated by incubation at 100°C for 2 minutes, and the product identified by HPLC. HPLC analysis was conducted using a Beckman Coulter System Gold HPLC unit (Beckman Coulter Inc., Fullerton, CA) using methods previously described (Stevenson-Paulik et al., 2005).

The products were separated on a Whatman Partisphere strong anion exchange (SAX) column (Whatman Inc., Florham Park, NJ) using a linear gradient of 10 mM ammonium phosphate (pH 3.5) to 1.7 M ammonium phosphate (pH 3.5) over 12 minutes followed by 25 minutes of 1.7 M ammonium phosphate (pH 3.5). Tritiated reaction products were detected using an inline β-Ram Model 3 radioisotope detector (IN/US, Tampa, FL). To generate [^{3}H]Ins(1,3,4,6)P_4, a concerted reaction using a recombinant 5-phosphatase enzyme (generously donated by the laboratory of Dr. Gillaspy, Virginia Tech) and a GST-tagged recombinant soybean Ins(1,3,4)P_3 5/6-kinase was conducted as previously described (Chang et al., 2002). The product was confirmed using a 250 mm x 4.6 mm Adsorbosphere SAX HPLC column (Alltech, Deerfield, IL) using a linear gradient of 0 – 1 M ammonium phosphate (pH 3.5) at 1

ml/min over 120 minutes as previously described (Chang et al., 2002). [^{3}H]Ins(1,3,4,6)P_4 elutes just before the [^{3}H]Ins(1,3,4,5)P_4 standard.

Ins(1,4,5)P_3 6/3/5-kinase assays were conducted using 20 µl reactions consisting of reaction buffer (50 mM HEPES pH 7.5, 100 mM KCl, 10 mM $MgCl_2$,1 mM ATP), a mixture of cold and tritiated substrate, and 100 ng of purified recombinant GST-*Gm*Ipk2 enzyme. The reactions were incubated at 25°C for various times and terminated by addition of 80 µl of 10 mM ammonium phosphate (pH 3.5) followed by incubation at 100°C for 2 minutes. The samples were centrifuged at 12,000 x g for 2 min to pellet the denatured proteins prior to HPLC analysis. Ins(1,3,4,5,6)P_5 kinase assays were conducted using the same method as above using 250 ng of purified recombinant HIS-*Gm*Ipk1 enzyme.

Northern Blots

Northern-blot analyses were performed using total RNA extracted from soybean seeds at increasing seed sizes representing stages of development. Soybean seeds were harvested from multiple plants and sorted into four size groups (0-4 mm, 5-6 mm, 7-8 mm, and 9-10 mm). Total RNA was extracted using Tri-Reagent according to manufacturer's instructions (MRCgene, Cincinnati, OH) and resuspended in RNAsecure resuspension buffer (Ambion, Austin, TX). 10 µg total RNA was separated on a 0.9% glyoxal agarose gel using commercial gel buffer and glyoxal loading dye (Ambion, Austin, TX). Prior to transfer to Nytran SuperCharge membrane (Schleicher and Schuell, Keene, NH) the gels were visualized with ethidium bromide and photographed to document RNA integrity and equal loading of lanes. The blots were transferred using Ambion transfer buffer for two hours at room temperature and UV cross-linked using 120

joules/cm^2. The blots were pre-hybridized for 4 hours at 68°C using modified Church's reagent (0.5 M Na_2HPO_4, 7% SDS, 1 mM EDTA, pH 7.8) (Church and Gilbert, 1984). The blots were hybridized overnight at 68°C in the same buffer with full length DNA probes generated using a Psoralen-Biotin kit (Ambion, Austin, TX). The blots were washed for 10 minutes at room temperature with 100 ml 1xSSC (150 mM NaCl, 15 mM sodium citrate, pH 7.0) , 0.1% sodium dodecyl sulfate (SDS), then washed three times for 10 minutes at 68°C with 0.5xSSC, 0.1% SDS. The blots were visualized by chemiluminesence using a Biodetect kit (Ambion, Austin, TX) and exposed on film.

RT-PCR Expression Analysis

Total RNA for RT-PCR was extracted from developing soybean seeds using the same procedure as for northern blots. Three biological replicates were analyzed for each gene using tissue samples from multiple greenhouse grown plants. Two of the biological replicates were from the soybean cultivar "Jack" and the third from Virginia experimental line V71-370. Total RNA was also extracted from soybean embryogenic tissue culture, soybean leaves, and soybean roots from the soybean cultivar "Jack", and two biological replicates analyzed. Purified RNA was quantified by measuring absorbance at 260 nm, separated by agarose gel electrophoresis and visualized by ethidium bromide staining. Total RNA was treated with DNase I using a Turbo DNA-free kit according to manufacturer's instructions (Ambion, Austin, TX) and tested to verify a lack of residual DNA by RT-PCR analysis using treated RNA as the template. cDNA was generated using oligo dT primers 12-18 nucleotides in length (Invitrogen, Carlsbad, CA) and the Omniscript cDNA kit (Qiagen, Valencia, CA). Primers were designed using Lightcycler

Software (Roche Applied Science, Indianapolis, IN) to amplify fragments of approximately 200 bp (Table 3-5). For specificity, the reverse primer for each gene was chosen from the 3' untranslated region. Each primer set was tested for single product amplification and product size prior to use in real-time analysis. Products were separated on a 1% agarose gel followed by ethidium bromide staining and visualization. To check for equal synthesis of cDNA among samples, all twelve cDNA samples (3 biological repetitions at 4 stages) were simultaneously analyzed in triplicate using primers for two housekeeping genes; Elongation factor 1a (EF1A) and Glucose-6-phosphate dehydrogenase (G6PHD) (Table 3-5).

Each 50 µl RT-PCR reaction consisted of POWER SYBR green master mix (Applied Biosystems, Foster City, CA), cDNA corresponding to 10 ng of original total RNA, and primers at a concentration of 200 nM each. Each reaction was conducted in triplicate and the results averaged. Analysis of each gene included cDNA samples from each seed development stage, controls lacking template, and a standard curve. The standard curve was generated for each gene using 10^4 to 10^{10} molecules of plasmid DNA containing the gene of interest. PCR conditions began with 2 minutes at 50°C followed by denaturation for 30 sec at 95°C, annealing for 30 sec at 60°C, and extension for 45 sec at 72°C for 40 cycles. Each reaction was followed by a melt curve analysis to ensure single product amplification. Amplification and analysis were conducted using an ABI Prism 7700 Sequence Detection System (Applied Biosystems, Foster City, CA).

Enzyme Kinetics

Enzyme kinetic analyses were performed using identical reaction conditions as described in *myo*-inositol kinase activity assays above. To determine reaction velocity, reactions were stopped before 25% of the substrate was converted to product. Each *Gm*Ipk2 assay used 100 ng of recombinant enzyme, and each *Gm*Ipk1 assay used 250 ng. Products were detected and quantified using HPLC as described above.

Yeast Vector Generation

*Gm*Ipk2 and *Gm*Ipk1 were cloned into the yeast expression vector pYES-DEST52 using the Gateway™ system (Invitrogen, Carlsbad, CA). Both sequences were amplified using primers that corresponded to the start and stop of the coding region (Table 3-5). The forward primer contained the CACC sequence for directional insertion into the pENTR/D/SD vector. The amplified sequence was inserted into the pENTR/D/SD according the manufacturer's instructions (Invitrogen, Carlsbad, CA). The clone was transformed into TOP10 cells (Invitrogen, Carlsbad, CA) for plasmid replication. The plasmid was isolated using a miniprep kit (Qiagen, Valencia, CA) and sequenced at the Virginia Bioinformatics Institute Core Facility (Blacksburg, VA). The coding region was recombined into the pYES-DEST52 vector using an LR recombinase reaction (Invitrogen, Carlsbad, CA). Following the recombination reaction, the plasmid was transformed, isolated and sequenced as described above. Due to the presence of the ccdB gene in the pYES-DEST52 vector, the original vector could not be used as an empty vector control. To generate an empty vector control plasmid, a recombination reaction was necessary to

replace the ccdB gene with random DNA sequence. A sequence of DNA lacking a start codon was amplified and inserted into the pENTR/D/SD vector followed by the methods described above for plasmid transformation, replication, isolation and recombination.

Yeast Transformation

The yeast *ipk2Δ* strain was generously provided by Dr. York's laboratory (Duke University Medical Center, Durham, NC) and the *ipk1Δ* and wild-type yeast strain (*ipk1Δ*: Clone ID 3674, wild-type strain: BY4730) were purchased from Open Biosystems (Huntsville, AL). The mutants were transformed with the plasmid of interest, pDEST52-Control, pDEST52-*Gm*Ipk2, or pDEST52-*Gm*Ipk1, using the lithium acetate method (Ito et al., 1983). The yeast transformed with pDEST52-*Gm*Ipk2 or pDEST52-*Gm*Ipk1 were plated on complete minimal media lacking uracil and containing 200 μg/ml Geneticin (Invitrogen, Carlsbad, CA) for selection.

Yeast Complementation Assay and HPLC Analysis

Complemented yeast strains were grown to late logarithmic phase in 3 ml complete minimal media lacking uracil, with 200 μg/ml Geneticin, 20 μCi/ml [^{3}H]*myo*-inositol, and 2% galactose for induction. Wild-type cells were grown in 3 ml complete minimal media with 20 μCi/ml [^{3}H]*myo*-inositol, and 2% galactose. The cells were pelleted and washed three times using ice cold media and stored at -80°C until extraction. To isolate soluble *myo*-inositol phosphates, the pellets were resuspended in 100 μl 0.5 N HCl. 372 μl of chloroform/methanol

(1:2 v/v, freshly prepared) and 200 µl of glass beads were added and agitated three times for 30 seconds each. 125 µl chloroform and 125 µl 2 M KCl were added and followed by agitation three times for 30 seconds each. The lysate was clarified by centrifugation at 13,000 x g for five minutes and the upper aqueous phase recovered. The aqueous phase was dried under nitrogen gas and resuspended in dH_20. An aliquot was counted on the scintillation counter and the samples normalized for equal counts. The *myo*-inositol phosphates were separated by HPLC using a Beckman Coulter System Gold HPLC unit (Beckman Coulter Inc., Fullerton, CA) by methods previously described (Stevenson-Paulik et al., 2005). The products were separated using a Whatman Partisphere strong anion exchange (SAX) column (Whatman Inc., Florham Park, NJ) using a linear gradient of 10 mM ammonium phosphate (pH 3.5) to 1.7 M ammonium phosphate (pH 3.5) over 12 minutes followed by 25 minutes of 1.7 M ammonium phosphate (pH 3.5) at a rate of 1 ml/min. Tritiated reaction products were detected using an inline β-Ram Model 3 radioisotope detector (IN/US, Tampa, FL) and Tru-Count scintillation fluid (IN/US, Tampa, FL) at a rate of 3 ml/min. Peaks were identified by comparison to known standards.

ACKNOWLEDGEMENTS

The authors thank Jennifer Webster for cloning the *Gm*Ipk1 gene into pGEX4T-1, Joseph Johnson and Nicole Juba for assistance with the yeast vector generation and yeast transformation, and Dr. John York for generously providing the *ipk2Δ* yeast strain. We would also like to thank Dr. Peter Kennelly and Dr. Stephen Shears for suggestions for the kinetic assays and Dr. Shao-Chun Chang for advice regarding the *myo*-inositol phosphate separation procedure.

Figure 3-1. Alignment of the predicted amino acid sequence of soybean *Gm*Ipk2 (Genbank Accession: EU033957) with the Arabidopsis sequences *At*Ipk2α and *At*Ipk2β (Genbank Accession: AY147935, AY147936). Black boxes indicate identical amino acids.

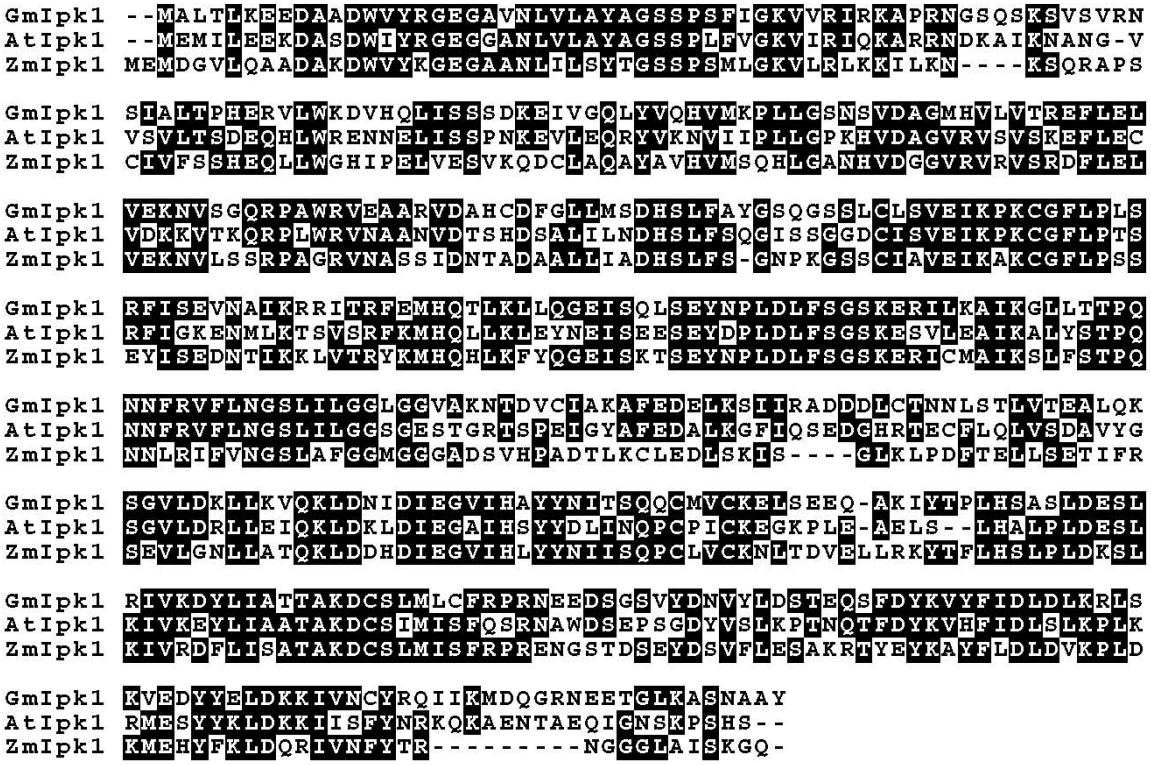

Figure 3-2. Alignment of the predicted amino acid sequence of soybean *Gm*Ipk1 (Genbank Accession: EU033956) with the Arabidopsis *At*Ipk1 and maize *Zm*Ipk1 sequences (Genbank Accession: AAZ99216, ABO33318). Black boxes indicate identical amino acids.

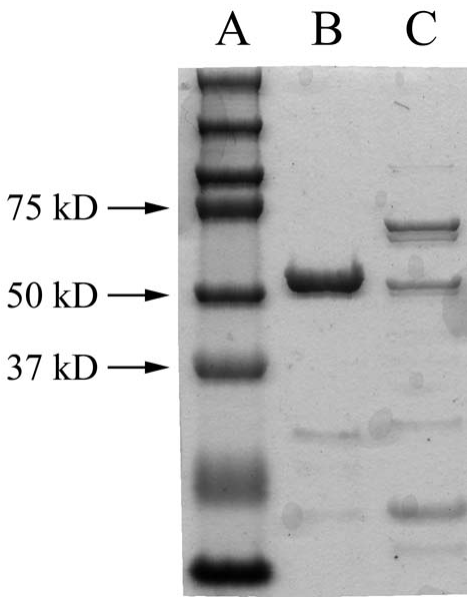

Figure 3-3. SDS-PAGE gel showing purification of the *Gm*Ipk1 and *Gm*Ipk2 recombinant proteins. Lane A) Protein size marker (Precision Plus Kaleidoscope, Biorad, Richmond, CA). Lane B) GST-tagged *Gm*Ipk2, 2.5 μg of total protein. Expected size is 57.6 kD. Lane C) His-tagged *Gm*Ipk1, 2.5 μg of total protein. Expected size is 53.2 kD.

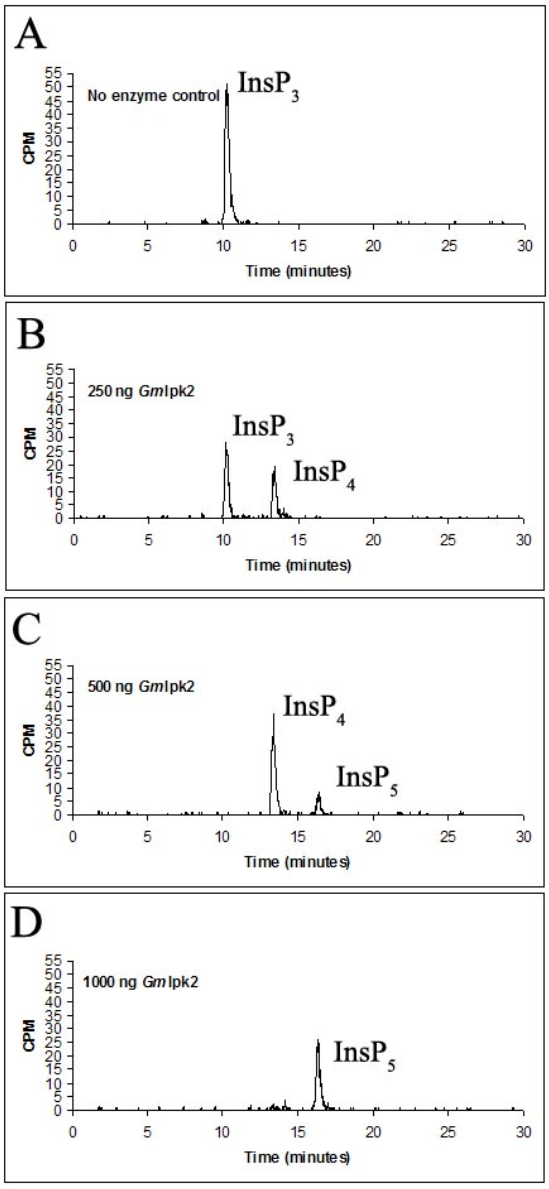

Figure 3-4. HPLC elution profiles of the products of the Ins(1,4,5)P_3 6/3/5-kinase activity assay. [^{3}H]Ins(1,4,5)P_3 was incubated with increasing concentrations of the *Gm*Ipk2 enzyme for a fixed time period. A) Profile of the [^{3}H]Ins(1,4,5)P_3 substrate lacking enzyme. B-E) Profiles showing the products of the *Gm*Ipk2 activity assay with increasing concentrations of recombinant enzyme, 250 ng, 500 ng, and 1000 ng.

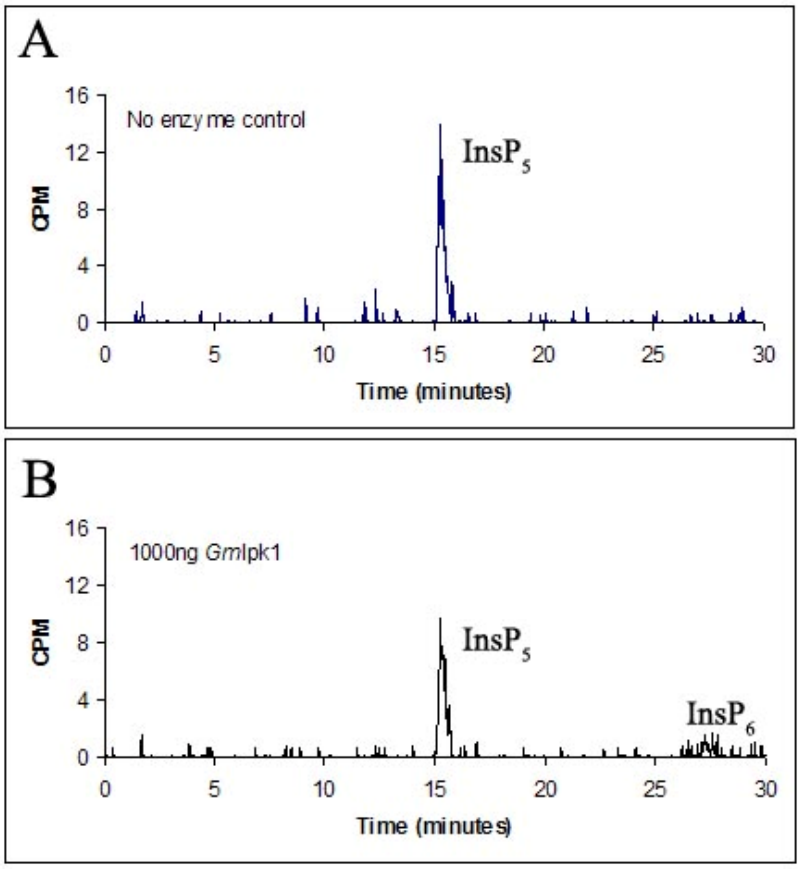

Figure 3-5. HPLC elution profile of the product of the Ins(1,3,4,5,6)P_5 2-kinase activity assay. The [^{3}H]Ins(1,3,4,5,6)P_5 substrate was generated using the recombinant *Gm*Ipk2 enzyme. A) Profile of [^{3}H]Ins(1,3,4,5,6)P_5 substrate lacking enzyme. B) Profile showing the product of the recombinant *Gm*Ipk1 activity assay.

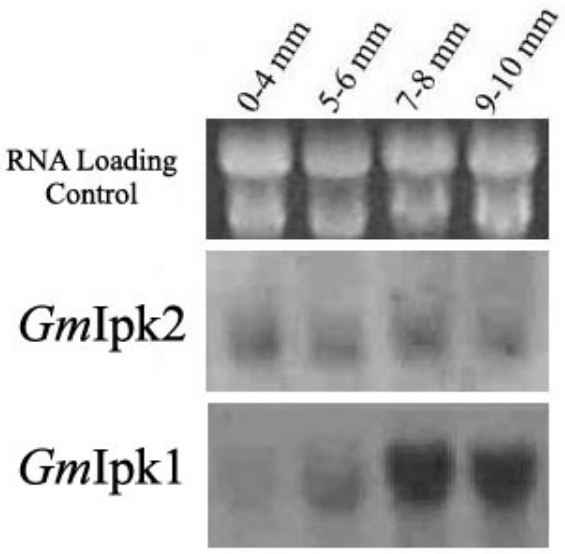

Figure 3-6. Northern blots of total RNA from developing soybean seeds using full-length *Gm*Ipk2 and *Gm*Ipk1 probes. Developing seeds were grouped by size (0-4 mm, 5-6 mm, 7-8 mm, 9-10 mm) for RNA isolation.

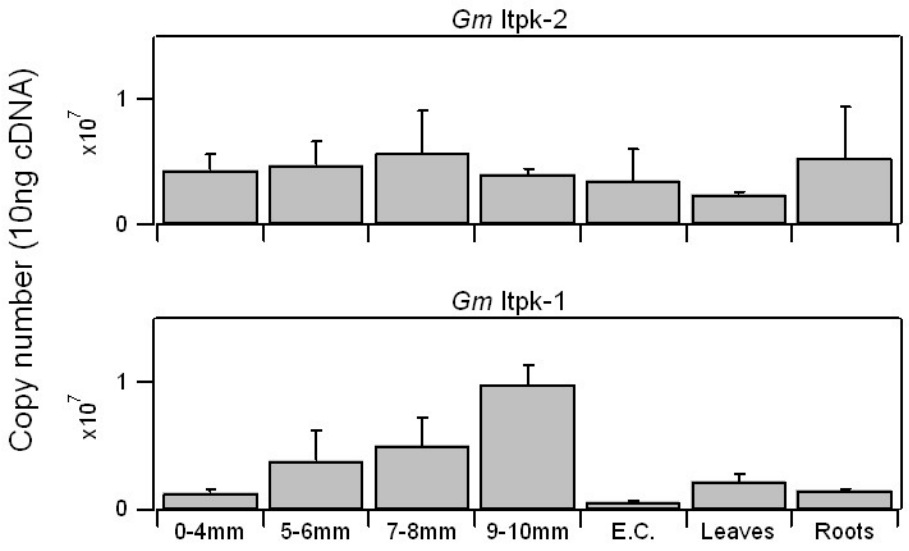

Figure 3-7. RT-PCR analyses quantifying steady state transcript levels of *Gm*Ipk1 and *Gm*Ipk2. cDNA was generated from total RNA extracted from developing soybean seeds grouped by size (0-4 mm, 5-6 mm, 7-8 mm, 9-10 mm), embryogenic soybean tissue culture (E.C.), soybean leaves, and soybean roots. Reverse primers were generated from the 3'UTR region of each gene for specificity.

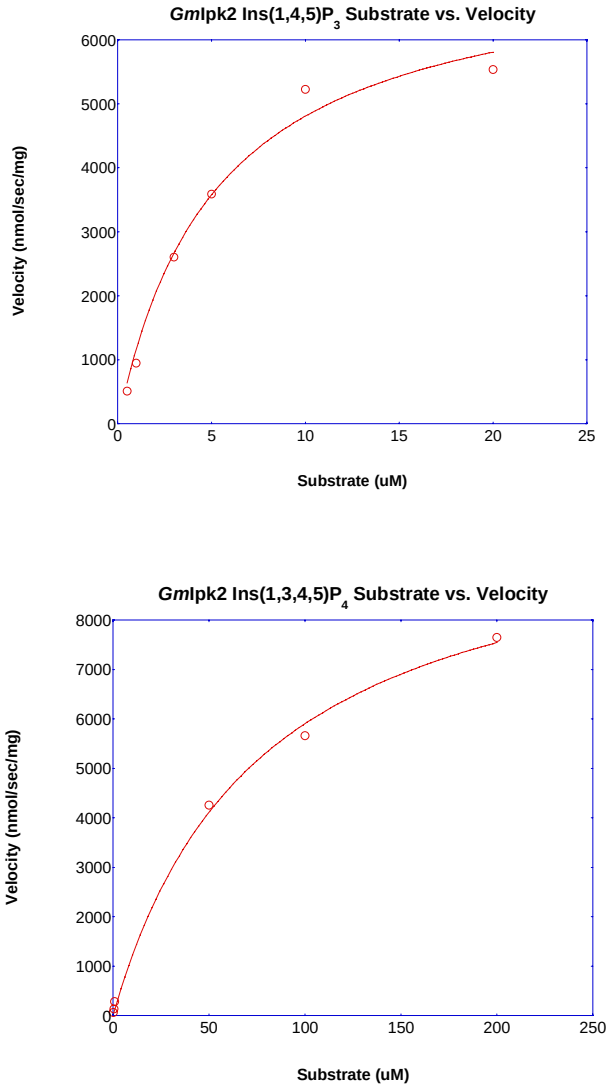

Figure 3-8. Kinetics of the recombinant soybean Ins(1,4,5)P_3 6/3/5-kinase enzyme. Kinetic assays were conducted using [^{3}H]Ins(1,4,5)P_3 and [^{3}H]Ins(1,3,4,5)P_4 as substrates. Reactions were run using a fixed concentration of ATP at 1 mM while varying the concentration of the substrate. The data was analyzed using KaleidaGraph version 4.0.3 (Synergy Software, Reading, PA). Each data point represents the average of three repetitions.

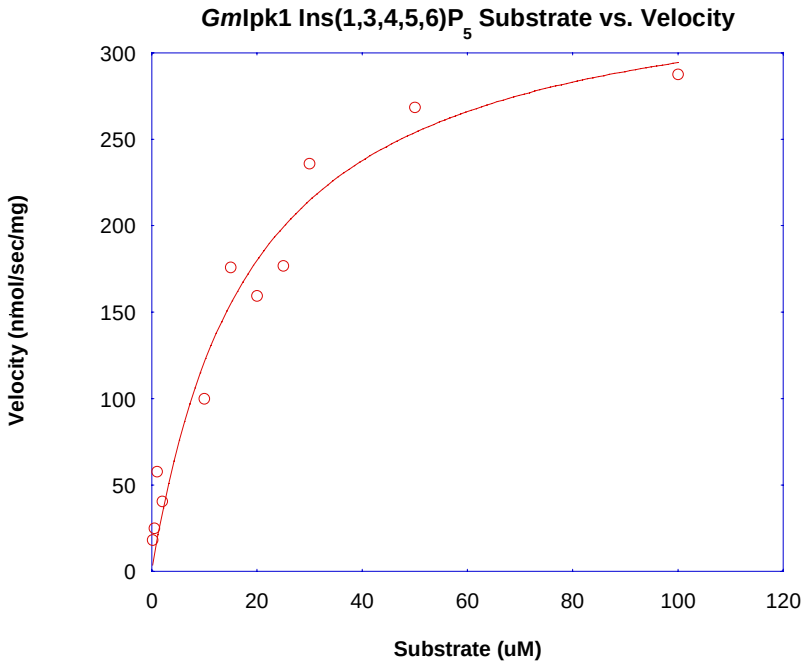

Figure 3-9. Kinetics of the recombinant soybean Ins(1,3,4,5,6)P_5 2-kinase enzyme. Kinetic assays were conducted using [^{3}H]Ins(1,3,4,5,6)P_5 as a substrate. Reactions were run using a fixed concentration of ATP at 1 mM while varying the concentration of the substrate. The data was analyzed using KaleidaGraph version 4.0.3 (Synergy Software, Reading, PA). Each data point represents the average of three repetitions.

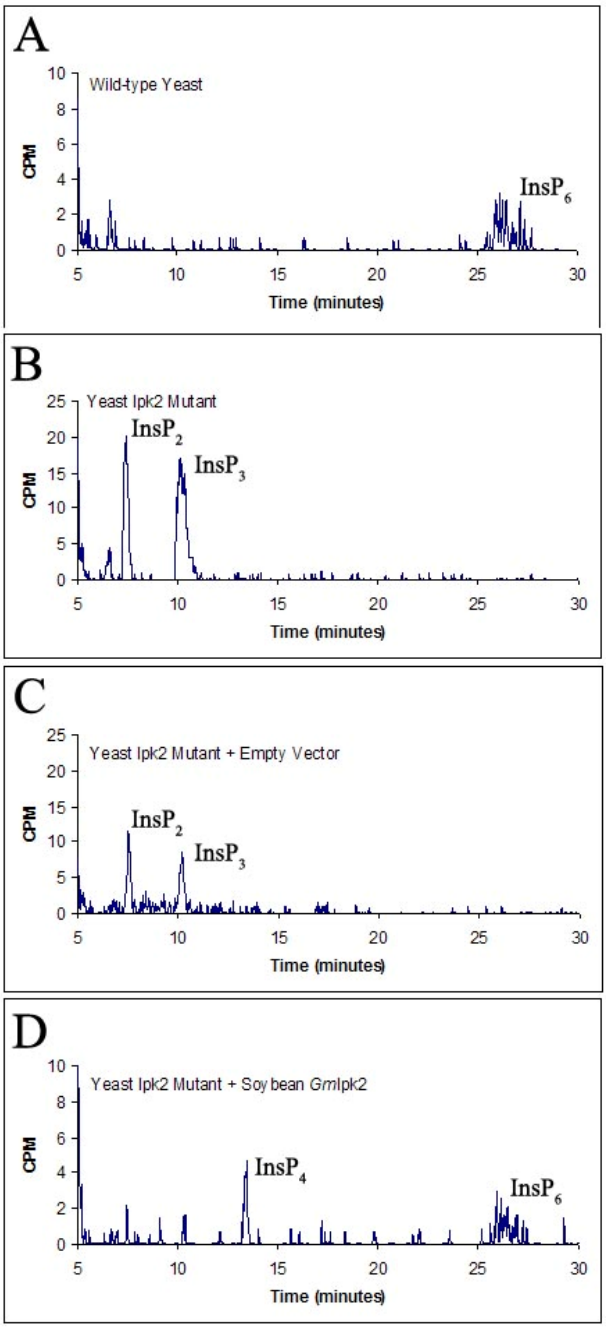

Figure 3-10. Complementation of the yeast *ipk2Δ* mutant by the soybean Ins(1,4,5)P_3 6/3/5-kinase. Yeast strains were grown to late-logarithmic phase with 20 μCi/ml [^{3}H]*myo*-inositol. Soluble inositol phosphates were extracted, normalized for equal counts, and analyzed using HPLC. A) Wild-type yeast strain, B) Yeast *ipk2Δ* strain, C) Yeast *ipk2Δ* strain transformed with the pDEST52-empty vector, D) Yeast *ipk2Δ* strain transformed with pDEST52-*Gm*Ipk2.

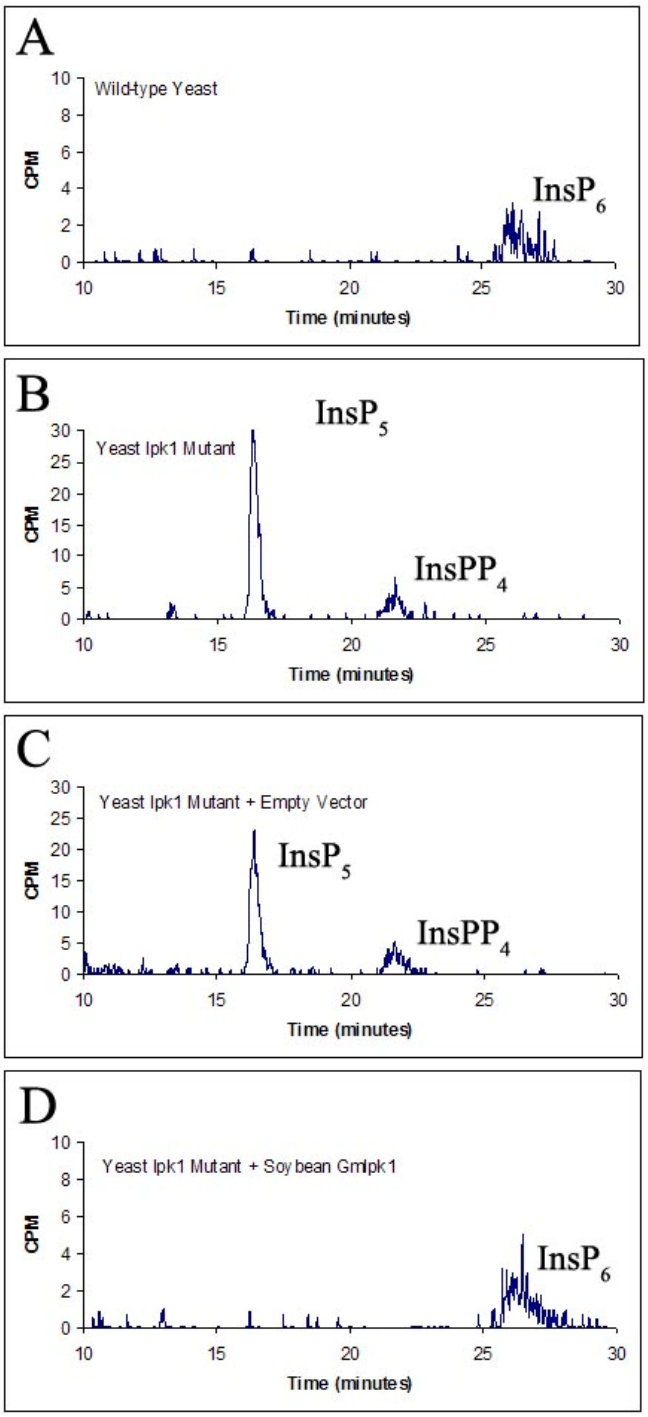

Figure 3-11. Complementation of the yeast *ipk1Δ* mutant by soybean Ins(1,3,4,5,6)P_5 2-kinase. Yeast strains were grown to late-logarithmic phase with 20 μCi/ml [^{3}H]*myo*-inositol. Soluble inositol phosphates were extracted, normalized for equal counts, and analyzed using HPLC. A) Wild-type yeast strain, B) Yeast *ipk1Δ* strain, C) Yeast *ipk1Δ* strain transformed with the pDEST52-empty vector, D) Yeast *ipk1Δ* strain transformed with pDEST52-*Gm*Ipk1.

CHAPTER IV

Summary and Future Directions

Summary

In this study we have isolated and characterized six *myo*-inositol kinase enzymes in soybean. We have examined enzyme activity both *in vitro* and *in vivo* and expression patterns in developing seeds. Our main goal in this study was to identify enzymes involved in phytic acid biosynthesis and to determine other possible functions in which the enzymes may be involved. A long-term goal is to develop *lpa* soybeans by targeting specific genes for down-regulation, rather than using the more common, random EMS mutagenesis approach. The *myo*-inositol pathway plays many important roles in plant growth and development, germination, and potentially in signaling cascades. Perturbations in the pathway have the potential to create a myriad of deleterious effects. In this study we focused on three types of *myo*-inositol kinase enzymes, *Gm*Itpk (Ins(1,3,4)P_3 5/6-kinase / Ins(3,4,5,6)P_4 1-kinase), *Gm*Ipk2 (Ins(1,4,5)P_3 6/3/5-kinase) and *Gm*Ipk1 (Ins(1,3,4,5,6)P_5 2-kinase). All three kinases have shown an *lpa* phenotype when down-regulated in Arabidopsis or maize (Shi et al., 2003; Stevenson-Paulik et al., 2005). A loss-of-function mutation in the *Gm*Itpk homolog in maize resulted in an approximate 30% decrease in phytic acid levels (Shi et al., 2003). In Arabidopsis, T-DNA insertions in *At*Ipk2-β and *At*Ipk1 resulted in a 35% and 83% decrease in $InsP_6$ levels respectively (Stevenson-Paulik et al., 2005). Targeting early steps in the pathway is more likely to disrupt other important functions involving *myo*-inositol and the lower *myo*-inositol phosphates. These three enzymes

were also chosen for analysis because their substrates in other organisms suggest that they may be involved in later steps of phytic acid biosynthesis.

*Gm*Itpk makes up a gene family of at least four members in the soybean genome. One of these members, *Gm*Itpk-3, is highly expressed in seeds with a pattern of increasing transcript abundance during seed development, which indicates a potential role in phytic acid biosynthesis. Each enzyme phosphorylates both Ins(1,3,4)P_3 and Ins(3,4,6)P_3 to an InsP_4 product, and phosphorylates Ins(3,4,5,6)P_4 to Ins(1,3,4,5,6)P_5. Each enzyme also functions as a 1-phosphatase on Ins(1,3,4,5,6)P_5, dephosphorylating it to Ins(3,4,5,6)P_4. Of the three substrates tested, Ins(1,3,4)P_3 is unlikely to be a physiological substrate in plants as it is in humans. However, both Ins(3,4,6)P_3 and Ins(3,4,5,6)P_4 have been found in plants and are potential intermediates in the pathway.

There appears to be a single copy of *Gm*Ipk2 in soybean. This enzyme showed a low but constant level of expression in soybean seeds during seed development. The kinase phosphorylated Ins(1,4,5)P_3, Ins(1,3,4,5)P_4 and Ins(1,3,4,6)P_4 in *in vitro* activity assays, and demonstrated *in vivo* activity by rescuing an *ipk2Δ* yeast mutant. Kinetic assays were conducted to compare the enzyme's relative affinity for the three analyzed substrates and to compare their catalytic efficiency values. The results showed the lowest K_m and K_{cat} value for Ins(1,4,5)P_3. This may indicate a potential step in phytate biosynthesis, or a role for the enzyme in Ins(1,4,5)P_3 signaling. Ins(1,4,5)P_3 is a well-studied signaling molecule in higher organisms for releasing intracellular calcium stores (Berridge, 1993) and may function in a similar manner in plants (Xu et al., 2005).

*Gm*Ipk1 has Ins(1,3,4,5,6)P_5 2-kinase activity *in vitro*. Ins(1,3,4,5,6)P_5 2-kinase activity was also demonstrated *in vivo* by the enzyme's ability to rescue an *ipk1Δ* yeast mutant. Kinetic

assays were conducted to examine the enzyme's affinity for the Ins(1,3,4,5,6)P_5 substrate. Based on the kinetic data and studies in other organisms, it seems most likely that the physiological function of the enzyme is as an Ins(1,3,4,5,6)P_5 2-kinase. The expression data support this hypothesis based on the observations that transcript levels increase in abundance as the seeds develop.

Future Directions

Having identified and characterized several *myo*-inositol kinase enzymes in soybean, we would like to more specifically elucidate their role in phytic acid biosynthesis. To address this question, and RNAi approach will be used to down-regulate individual enzymes. A preliminary approach to down-regulation will be conducted using embryogenic soybean culture. Somatic soybean embryos have been previously shown to be a suitable model for developing seeds, and have been successfully used to study genes expressed in the seed (Cahoon et al., 2001; Dhugga et al., 2004). Each enzyme will be down-regulated using an RNAi construct corresponding to either the coding region of the gene, or the 3'UTR region to ensure specificity. Down-regulation will be analyzed using quantitative PCR, and the *myo*-inositol phosphate profile examined using HPLC. Preliminary data using quantitative PCR indicate that all six *myo*-inositol kinase enzymes characterized in this study are expressed in soybean tissue culture (unpublished data). As shown in developing soybean seeds, *Gm*Itpk-3 is also the most highly expressed in embryogenic culture. The soybean tissue culture will be grown in a liquid medium and labeled using [^{3}H]*myo*-inositol. After several days of labeling, the soluble *myo*-inositol phosphates will

be extracted and identified using HPLC. A preliminary experiment using this method shows that *myo*-inositol phosphates are produced in soybean tissue culture, and the most abundant *myo*-inositol phosphate is phytic acid (unpublished data). Using embryogenic soybean culture will provide a quick method to identify the relative contribution of each enzyme to phytic acid production and will aid in elucidation of the pathway.

The information gained from the RNAi experiments in embryogenic culture will help to identify candidate enzymes for down-regulation. However, a whole plant system will be necessary to examine the effects of down-regulation of each enzyme on other processes in the plant. To study whole-plant effects, transgenic soybean plants will be generated containing RNAi constructs to down-regulate gene expression of the enzymes deemed most important by the results of the embryogenic culture experiments. The *myo*-inositol phosphate profiles and the phenotypes of the transgenic plants will be analyzed. Data will be collected on characteristics such as germination rate, seed weight and yield and seed phytate content.

Our long-term goal is the generation of an *lpa* soybean plant using a targeted method that minimizes the deleterious effects of pathway modification. The first step in achieving this goal is to elucidate the pathway. The second step is to understand the effects that perturbations in the pathway may cause. Once the pathway is understood and additional roles of each enzyme uncovered, the most effective method for *lpa* plant generation can be determined. In order to target phytate biosynthesis specifically, seed-specific promoters will be used to target down-regulation of the genes solely during seed development and phytate deposition. This may prove to be the most efficient method of decreasing phytate levels while allowing the soybean plant to continue normal growth and development.

APPENDIX

INTRODUCTION

Several experiments testing the biochemical characteristics of the *Gm*Itpk enzymes were conducted in collaboration with Dr. Stephen Shears and Dr. Xun Qian at the National Institute of Environmental Health Sciences (NIEHS). In this collaboration, I supplied the sequence data for *Gm*Itpk1-4, purified recombinant proteins *Gm*Itpk1-4, and generated the substrate [^{3}H]Ins(1,3,4)P_3. The Shears laboratory conducted kinetic assays of *Gm*Itpk1-4 on the substrates Ins(1,3,4)P_3 and Ins(3,4,5,6)P_4, and activity assays for each protein on the phosphorylation of Ins(3,4,6)P_3 and the dephosphorylation of Ins(1,3,4,5,6)P_5. The use of the soybean sequence data and recombinant proteins resulted in the publication *Integration of Inositol Phosphate Signaling Pathways Via Human Itpk1* (Chamberlain et al., 2007). A second manuscript is currently in preparation describing the *Gm*Itpk soybean gene family. The materials and methods used by the Shears laboratory are described below, along with the figures describing the results of each experiment.

MATERIALS AND METHODS

GmItpk Enzyme Kinetics on Ins(1,3,4)P_3 and Ins(3,4,5,6)P_4

[^{3}H]Ins(1,3,4)P_3 was synthesized using a recombinant 5-phosphatase (generously donated by Dr. Gillaspy at Virginia Tech) to cleave the phosphate from position five from

[^{3}H]Ins(1,3,4,5)P_4. The reaction was incubated at 30°C for 16 hours in reaction buffer (50 mM HEPES pH 7.2, 100 mM NaCl, 0.25 mM EDTA, 2 mM $MgCl_2$) and product(s) verified using high performance liquid chromatography (HPLC). [^{3}H]Ins(3,4,5,6)P_4 was generated as previously described (Qian et al., 2005). Non-radioactive Ins(1,3,4)P_3 and Ins(3,4,5,6)P_4 were purchased from CellSignals (Columbus, OH).

Enzyme kinetic analyses were conducted using 100 µl reactions consisting of reaction buffer (20 mM HEPES pH 7.2, 100 mM KCl, 6 mM $MgSO_4$, 10 mM phosphocreatine, 2 units/ml creatine phosphokinase, 0.3mg/ml BSA), a mixture of cold and tritiated substrate and a fixed amount of protein. The reactions were halted before 25% of the substrate had been converted to product. The reactions were quenched using 20 µl 2 M perchloric acid, samples left on ice for 10 minutes, then centrifuged to remove protein precipitate. The supernatant was neutralized with 33 µl 1 M sodium carbonate, 1 mM EDTA. The sample was incubated on ice for 10 minutes, centrifuged and the supernatant stored for analysis.

To quantify the *myo*-inositol phosphate products, the samples were diluted in 10 ml of dH_20 and applied to Dowex AG 1-X8 resin (formate form, 200-400 mesh, Bio-Rad, Richmond, CA). To separate $InsP_3$ and $InsP_4$, the sample was applied to 0.7 ml columns, and the column washed with 17.5 ml 0.7 M ammonium formate, 0.1 M formic acid to remove the $InsP_3$ substrate. $InsP_4$ was eluted with 5 ml 2 M ammonium formate, 0.1 M formic acid. Scintillation fluid was added to the elution and activity counted on a scintillation counter. To separate $InsP_4$ and $InsP_5$, the sample was applied to 0.8 ml columns, and the column washed with 22.5 ml 0.9 M ammonium formate, 90 mM formic acid and $InsP_5$ eluted with 5 ml 2 M ammonium formate, 0.1 M formic acid.

GmItpk Activity on Ins(3,4,6)P_3

Non-radioactive Ins(3,4,6)P_3 was purchased from CellSignals (CellSignals, Columbus, OH), ^{32}P-ATP was purchased from Perkin Elmer ((Perkin Elmer, Wellesley, MA)). To analyze activity on Ins(3,4,6)P_3, each isoform was incubated for 1 hour at 37°C in 20 µl assay buffer (20 mM HEPES pH 7.2, 100 mM KCl, 6 mM $MgSO_4$, 0.3 mg/ml BSA) with [^{32}P]ATP. Products were analyzed using a 4.6 x 125 mm Partisphere SAX HPLC column. The gradient consisted of mixing Solution A (1 mM Na_2EDTA) and Solution B (1 mM Na_2EDTA, 1.3 M $(NH_4)_2HPO_4$, pH 3.85) as follows: 0-5 min, 0% B; 5-10 min, 0-30% B; 10-60 min, 30-55% B at a flow rate of 1 ml/min. Radioactivity was assessed inline from the Cerenkov radiation using a Radiometric D515 Flow Scintillation Analyzer radioactivity (Packard Instrument Co., Meriden, CT). Data were exported as an ASCII file into SigmaPlot version 8.02.

To distinguish Ins(1,3,4,6)P_4 and Ins(2,3,4,6)P_4 from Ins(3,4,5,6)P_4, samples were analyzed on an 250 x 4.6 mm Q100 column (Thompson Instruments, Clear Brook, VA). Samples were eluted at a flow-rate of 1 ml/min using a gradient generated from Buffer A (1 mM EDTA) and Buffer B (1 mM EDTA, 2 M $NH_4H_2PO_4$, pH 3.35) as follows: 0-2 min, 0 % B, 2 - 120 min, 0-44 % B. 1 ml fractions were collected, mixed with 3 ml Monoflow-4 scintillant (National Diagnostics, Atlanta, GA) and counted for radioactivity.

TABLES

Table A-1. Kinetic parameters of *Gm*Itpk1-4 on the $Ins(1,3,4)P_3$ and $Ins(3,4,5,6)P_4$ substrates.

Kinase	K_m (μM)		V_{max} (pmol/ng/min)		K_{cat} (V_{max} / K_m)	
	$Ins(1,3,4)P_3$ substrate	$Ins(3,4,5,6)P_4$ substrate	$Ins(1,3,4)P_3$ substrate	$Ins(3,4,5,6)P_4$ substrate	$Ins(1,3,4)P_3$ substrate	$Ins(3,4,5,6)P_4$ substrate
*Gm*Itpk-1	3.7	0.28 +/- 0.004	0.5	0.083 +/- 0.02	0.135	0.296
*Gm*Itpk-2	46.2	0.78 +/- 0.04	1.97	0.76 +/- 0.06	0.043	0.974
*Gm*Itpk-3	43.4	1.05 +/- 0.16	0.86	0.22 +/- 0.1	0.020	0.210
*Gm*Itpk-4	4.4 +/- 0.6	0.7 +/- 0.03	0.007 +/- 0.001	0.004 +/- 0.0003	0.002	0.006

FIGURES

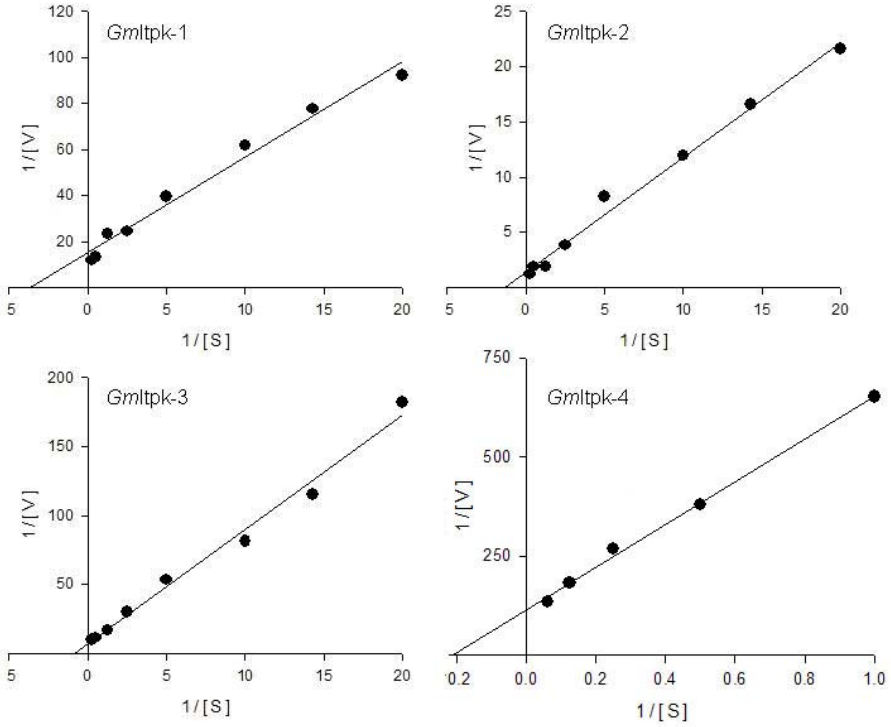

Figure A-1. *Gm*Itpk1-4 kinetics on the Ins(1,3,4)P_3 substrate. Kinetic assays were conducted using [^{3}H]Ins(1,3,4)P_3 as a substrate. Reactions were run using a fixed concentration of ATP while varying the concentration of the substrate. The data is depicted as Lineweaver-Burke plots, each data point represents the average of at least three repetitions.

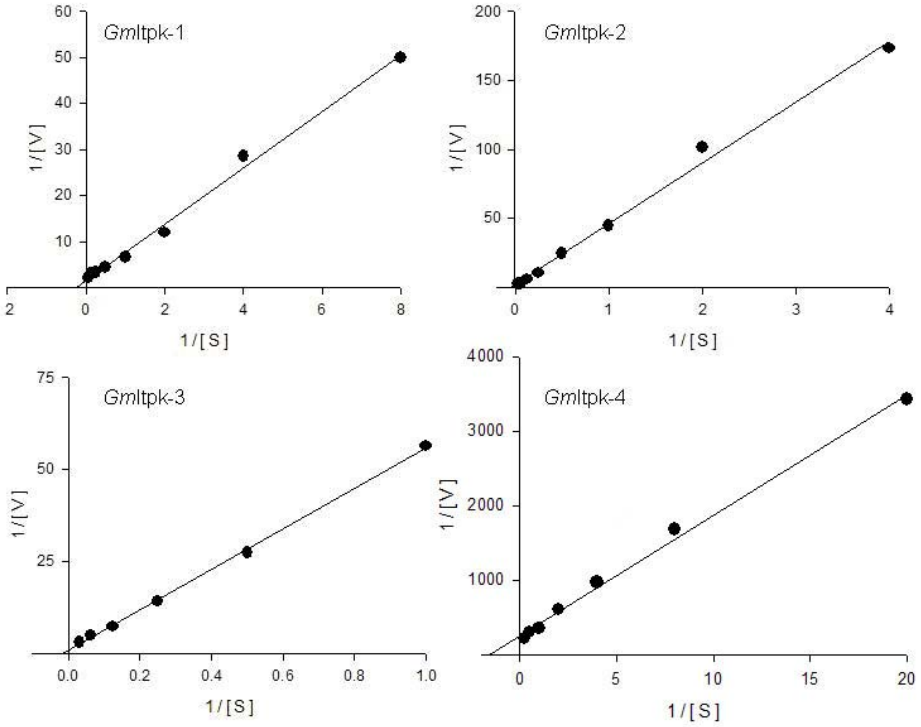

Figure A-2. *Gm*Itpk1-4 kinetics on the Ins(3,4,5,6)P_4 substrate. Kinetic assays were conducted using [^{3}H]Ins(3,4,5,6)P_4 as a substrate. Reactions were run using a fixed concentration of ATP while varying the concentration of the substrate. The data is depicted as Lineweaver-Burke plots, each data point represents the average of at least three repetitions.

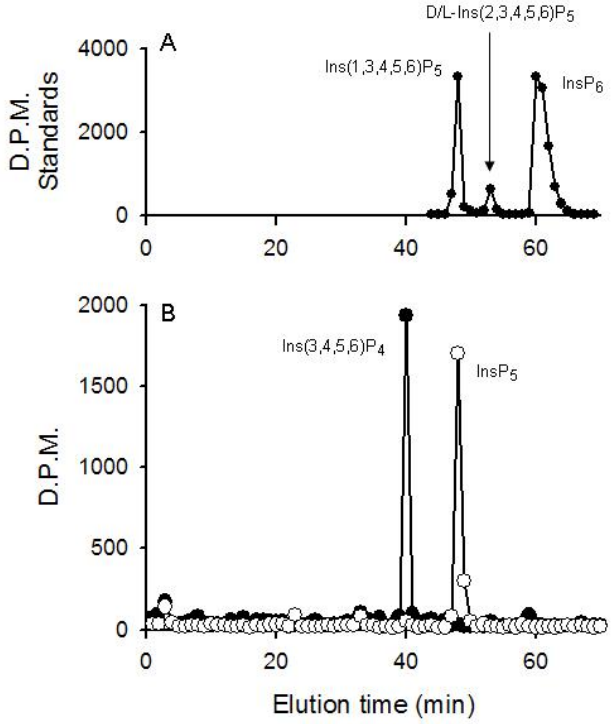

Figure A-3. Phosphorylation of Ins(3,4,5,6)P_4 to InsP_5 product. Panel A shows Ins(1,3,4,5,6)P_5, Ins(2,3,4,5,6)P_5 and InsP_6 standards, Panel B shows a representative activity assay of *Gm*Itpk enzymes on the Ins(3,4,5,6)P_4 substrate. The panel shows traces for both the the Ins(3,4,5,6)P_5 substrate (closed circles) and the InsP_5 product (open circles).

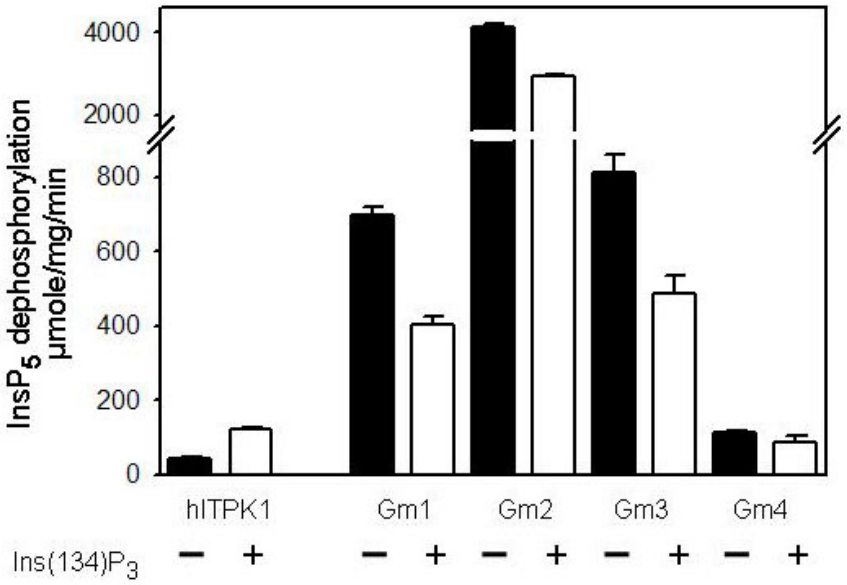

Figure A-4. Test for the stimulation of Ins(1,3,4,5,6)P_5 dephosphorylation by *Gm*Itpk1-4 using Ins(1,3,4)P_3. White bars indicate the addition of Ins(1,3,4)P_3 to the reaction, black bars indicate no Ins(1,3,4)P_3 is present. Human Itpk (hITPK1) results shown in comparison.

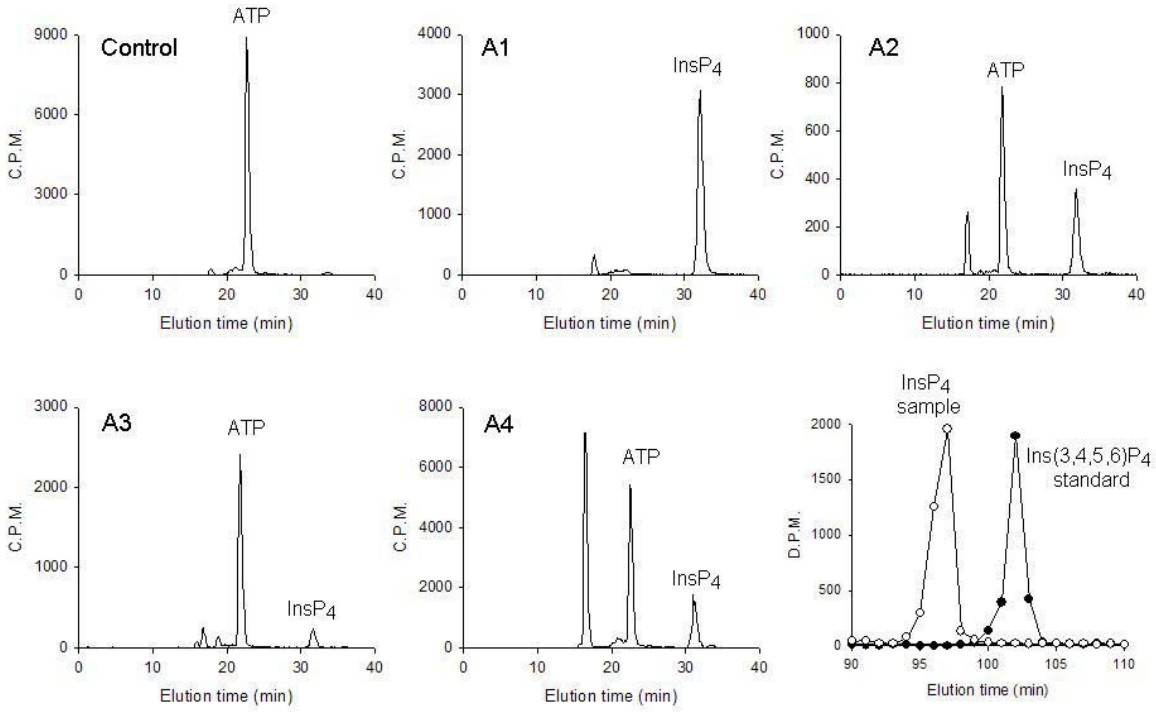

Figure A-5. *Gm*Itpk1-4 phosphorylation of Ins(3,4,6)P_3 to an InsP_4 product. Ins(3,4,6)P_3 was incubated with [^{32}P]ATP. Control panel shows [^{32}P]ATP alone, Panels A1-A4 demonstrate the ability of GmItpk1-4 to form an InsP_4 product. The final panel shows the InsP_4 product (open circles) with an Ins(3,4,5,6)P_4 standard (closed circles).

Abdullah M, Hughes P, Craxton A, Gigg R, Desai T, Marecek J, Prestwich G, Shears S (1992) Purification and characterization of inositol-1,3,4-trisphosphate 5/6- kinase from rat liver using an inositol hexakisphosphate affinity column. Journal of Biological Chemistry **267:** 22340-22345

Adams CL, Hambidge M, Raboy V, Dorsch JA, Sian L, Westcott JL, Krebs NF (2002) Zinc absorption from a low-phytic acid maize. American Journal of Clinical Nutrition **76:** 556-559

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Molecular Biology Rep. **9:** 208-218

Bechet J, Greenson M, Wiame JM (1970) Mutations affecting the repressibility of arginine biosynthetic enzymes in Saccharomyces cerevisiae. European Journal of Biochemistry **12:** 31-39

Berridge MJ (1989) Inositol 1,4,5-trisphosphate-induced calcium mobilization is localized in Xenopus oocytes. Proceedings of the Royal Society of London Series B **238:** 235-243

Berridge MJ (1993) Inositol trisphosphate and calcium signaling. Nature **361:** 315-325

Bertsch U, Haefs M, Moller M, Deschermeier C, Fanick W, Kitzerow A, Ozaki S, Meyer HE, Mayr GW (1999) A novel A-isoform-like inositol 1,4,5-trisphosphate 3-kinase from chicken erythrocytes exhibits alternative splicing and conservation of intron positions between vertebrates and invertebrates. Gene **228:** 61-71

Bjellqvist B, Hughes GJ, Pasquali C, Paquet N, Ravier F, Sanchez J-C, Frutiger S, Hochstrasser DF (1993) The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. Electrophoresis **14:** 1023-1031

Brearley CA, Hanke DE (1996) Inositol phosphates in barley (Hordeum vulgare L.) aleurone tissue are stereochemically similar to the products of breakdown of InsP6 in vitro by wheat-bran phytase. Biochemical Journal **318:** 279-286

Brearley CA, Hanke DE (1996) Metabolic evidence for the order of addition of individual phosphate esters to the myo-inositol moiety of inositol hexakisphosphate in the duckweed Spirodela polyrhiza L. Biochemical Journal **314:** 227-233

Cahoon EB, Ripp KG, Hall SE, Kinney AJ (2001) Formation of conjugated Delta(8),Delta(10)-double bonds by Delta(12)-oleic-acid desaturase-related enzymes - Biosynthetic origin of calendic acid. Journal of Biological Chemistry **276:** 2637-2643

Chamberlain PP, Qian X, Stiles AR, Cho J, Jones DH, Lesley SA, Grabau EA, Shears SB, Spraggon G (2007) Integration of inositol phosphate signaling pathways via human ITPK1. Journal of Biological Chemistry

Chang S-C, Miller AL, Feng Y, Wente SR, Majerus PW (2002) The Human Homolog of the Rat Inositol Phosphate Multikinase Is an Inositol 1,3,4,6-Tetrakisphosphate 5-Kinase. Journal of Biological Chemistry **277:** 43836-43843

Chappell AS, Scaboo AM, Wu X, Nguyen H, Pantalone RV, Bilyeu KD (2006) Characterization of the *MIPS* gene family in *Glycine max*. Plant Breeding **125:** 493-500

Chen QJ, Niu XG, Chai MF, Chen J, Liu Q, Wang XC (2003) Isolation of an Arabidopsis gene encoding ins (1,3,4) P-3 5/6-kinase-like protein and involved in plant response to abiotic stresses. Acta Botanica Sinica **45:** 211-218

Choi KY, Kim HK, Lee SY, Moon KH, Sim SS, Kim JW, Chung HK, Rhee SG (1990) Molecular cloning and expression of a complementary DNA for inositol 1,4,5-trisphosphate 3-kinase. Science **248:** 64-66

Chun JA, Jin UH, Lee JW, Yi YB, Hyung NI, Kang MH, Pyee JH, Suh MC, Kang CW, Seo HY, Lee SW, Chung CH (2003) Isolation and characterization of a myo-inositol 1-phosphate synthase cDNA from developing sesame (Sesamum indicum L.) seeds: functional and differential expression, and salt-induced transcription during germination. Planta **216:** 874-880

Church GM, Gilbert W (1984) Genomic sequencing. Proceedings of the National Academy of Sciences USA **81:** 1991-1995

Dean-Johnson M, Henry SA (1989) Biosynthesis of inositol in yeast. Primary structure of myo-inositol-1-phosphate synthase (EC 5.5.1.4) and functional analysis of its structural gene, the INO1 locus. Journal of Biological Chemistry **264:** 1274-1283

Dewaste V, Pouillon V, Moreau C, Shears S, Takazawa K, Erneux C (2000) Cloning and expression of a cDNA encoding human inositol 1,4,5-trisphosphate 3-kinase C. Biochemical Journal **352 Pt 2:** 343-351

Dhugga KS, Barreiro R, Whitten B, Stecca K, Hazebroek J, Randhawa GS, Dolan M, Kinney AJ, Tomes D, Nichols S, Anderson P (2004) Guar seed beta-mannan synthase is a member of the cellulose synthase super gene family. Science **303:** 363-366

Dorsch JA, Cook A, Young KA, Anderson JM, Bauman AT, Volkmann CJ, Murthy PPN, Raboy V (2003) Seed phosphorus and inositol phosphate phenotype of barley low phytic acid genotypes. Phytochemistry **62:** 691-706

Dubois E, Bercy J, Messenguy F (1987) Characterization of two genes, ARGRI and ARGRIII required for specific regulation of arginine metabolism in yeast. Molecular and General Genetics **207:** 142-148

Ertl DS, Young KA, Raboy V (1998) Plant genetic approaches to phosphorus management in agricultural production. Journal of Environmental Quality **27:** 299-304

Evans LT (1975) Crop physiology : some case histories. Cambridge University Press, London ; New York

Field J, Wilson MP, Mai Z, Majerus PW, Samuelson J (2000) An Entamoeba histolytica inositol 1,3,4-trisphosphate 5/6-kinase has a novel 3-kinase activity. Molecular and Biochemical Parasitology **108:** 119-123

Goldberg RB (1978) DNA sequence organization in the soybean plant. Biochemical Genetics **16:** 45-68

Gurley WB, Hepburn AG, Key JL (1979) Sequence organization of the soybean genome. Biochimica et Biophysica Acta **561:** 167-183

Guttieri M, Bowen D, Dorsch JA, Raboy V, Souza E (2003) Identification and characterization of a low phytic acid wheat. Crop Science **44:** 418-424

Hall DA, Zhu H, Zhu X, Royce T, Gerstein M, Snyder M (2004) Regulation of gene expression by a metabolic enzyme. Science **306:** 482-484

Hansen CA, vom Dahl S, Huddell B, Williamson JR (1988) Characterization of inositol 1,3,4-trisphosphate phosphorylation in rat liver. FEBS Letters **236:** 53-56

Hara K, Yagi M, Koizumi N, Kusano T, Sano H (2000) Screening of wound-responsive genes identifies an immediate-early expressed gene encoding a highly charged protein in mechanically wounded tobacco plants. Plant Cell Physiology **41:** 684-691

Hatzack F, Hubel F, Zhang W, Hansen PE, Rasmussen SK (2001) Inositol phosphates from barley low-phytate grain mutants analysed by metal-dye detection HPLC and NMR. Biochemical Journal **354:** 473-480

Hawkins PT, Poyner DR, Jackson TR, Letcher AJ, Lander DA, Irvine RF (1993) Inhibition of iron-catalysed hydroxyl radical formation by inositol polyphosphates: a possible physiological function for myo-inositol hexakisphosphate. Biochemical Journal **294 (Pt 3):** 929-934

Hegeman CE, Good LL, Grabau EA (2001) Expression of D-myo-inositol-3-phosphate synthase in soybean. Implications for phytic acid biosynthesis. Plant Physiology **125:** 1941-1948

Hill TD, Campos-Gonzalez R, Kindmark H, Boynton AL (1988) Inhibition of inositol trisphosphate-stimulated calcium mobilization by calmodulin antagonists in rat liver epithelial cells. Journal of Biological Chemistry **263:** 16479-16484

Hitz WD, Carlson TJ, Kerr PS, Sebastian SA (2002) Biochemical and molecular characterization of a mutation that confers a decreased raffinosaccharide and phytic acid phenotype on soybean seeds. Plant Physiology **128:** 650-660

Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. Genome Research **9:** 868-877

Hughes PJ, Kirk CJ, Michell RH (1994) Inhibition of porcine brain inositol 1,3,4-trisphosphate kinase by inositol polyphosphates, other polyol phosphates, polyanions and polycations. Biochimica et Biophysica Acta **1223:** 57-70

Hymowitz T (1990) Soybeans: The sucess story. *In* Advances in New Crops: Proceedings of the First National Symposium on New Crops: Research, Development, Economics, held at Indianapolis, Ind., 23-26 October 1998. *Edited by* J. Janick, and J.E. Simon. Timber Press, Portland OR. pp. 159-163

Irvine RF, Letcher AJ, Heslop JP, Berridge MJ (1986) The inositol tris/tetrakisphosphate pathway--demonstration of Ins(1,4,5)P3 3-kinase activity in animal tissues. Nature **320:** 631-634

Irvine RF, Schell MJ (2001) Back in the water: The return of the inositol phosphates. Nature Reviews Molecular Cell Biology **2:** 327-338

Ishitani M, Majumder AL, Bornhouser A, Michalowski CB, Jensen RG, Bohnert HJ (1996) Coordinate transcriptional induction of myo-inositol metabolism during environmental stress. Plant Journal **9:** 537-548

Ismailov, II, Fuller CM, Berdiev BK, Shlyonsky VG, Benos DJ, Barrett KE (1996) A biologic function for an "orphan" messenger: D-myo-inositol 3,4,5,6-tetrakisphosphate selectively blocks epithelial calcium-activated chloride channels. Proceedings of the National Academy of Sciences USA **93:** 10505-10509

Ito H, Fukuda Y, Murata K, Kimura A (1983) Transformation of intact yeast cells treated with alkali cations. Journal of Bacteriology **153:** 163-168

Ives EB, Nichols J, Wente SR, York JD (2000) Biochemical and functional characterization of inositol 1,3,4,5,6-pentakisphosphate 2-kinases. Journal of Biological Chemistry **275:** 36575-36583

Johnson MD (1994) Arabidopsis-Thaliana Myoinositol 1-Phosphate Synthase (Ec-5.5.1.4). Plant Physiology **105:** 1023-1024

Johnson MD, Sussex IM (1995) 1 L-myo-Inositol 1-Phosphate Synthase from Arabidopsis thaliana. Plant Physiology **107:** 613-619

Josefsen L, Bohn L, Sorensen MB, Rasmussen SK (2007) Characterization of a multifunctional inositol phosphate kinase from rice and barley belonging to the ATP-grasp superfamily. Gene **397:** 114-125

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. Journal of Molecular Biology **157:** 105-132

Larson SR, Raboy V (1999) Linkage mapping of maize and barley myo-inositol 1-phosphate synthase DNA sequences: correspondence with a low phytic acid mutation. Theoretical and Applied Genetics **99:** 27-36

Larson SR, Rutger JN, Young KA, Raboy V (2000) Isolation and genetic mapping of a non-lethal rice (Oryza sativa L.) low phytic acid 1 mutation. Crop Science **40:** 1397-1405

Larsson O, Barker CJ, Sjoholm A, Carlqvist H, Michell RH, Bertorello A, Nilsson T, Honkanen RE, Mayr GW, Zwiller J, Berggren PO (1997) Inhibition of phosphatases and increased Ca2+ channel activity by inositol hexakisphosphate. Science **278:** 471-474

Lemtiri-Chlieh F, MacRobbie EAC, Brearley CA (2000) Inositol hexakisphosphate is a physiological signal regulating the K+-inward rectifying conductance in guard cells. Proceedings of the National Academy of Sciences USA **97:** 8687-8692

Li YC, Ledoux DR, Veum TL, Raboy V, Ertl DS (2000) Effects of low phytic acid corn on phosphorus utilization, performance, and bone mineralization in broiler chicks. Poultry Science **79:** 1444-1450

Loewus FA, Murthy PPN (2000) myo-inositol metabolism in plants. Plant Science **150:** 1-19

Majerus PW (1992) Inositol phosphate biochemistry. Annual Review of Biochemistry **61:** 225-250

Majumder AL, Chatterjee A, Ghosh Dastidar K, Majee M (2003) Diversification and evolution of L-myo-inositol 1-phosphate synthase. FEBS Letters **553:** 3-10

Majumder ANL, Mandal, N.C., and Biswas, B.B. (1972) Phosphoinositol kinase from germinating mung bean seeds. Phytochemistry **11:** 503-508

Mendoza C, Viteri F, Lonnerdal B, Young K, Raboy V, Brown K (1998) Effect of genetically modified, low-phytic acid maize on absorption of iron from tortillas. Am. J. Clinical Nutrition **68:** 1123-1127

Miller GJ, Wilson MP, Majerus PW, Hurley JH (2005) Specificity determinants in inositol polyphosphate synthesis: crystal structure of inositol 1,3,4-trisphosphate 5/6-kinase. Molecular Cell **18:** 201-212

Myers W, Miller E (1989) Optimal alignments in linear space. Computer Applications in the Biosciences **4:** 11-17

Odom AR, Stahlberg A, Wente SR, York JD (2000) A Role for Nuclear Inositol 1,4,5-Trisphosphate Kinase in Transcriptional Control. Science **287:** 2026-2029

Ongusaha PP, Hughes PJ, Davey J, Michell RH (1998) Inositol hexakisphosphate in Schizosaccharomyces pombe: synthesis from Ins(1,4,5)P-3 and osmotic regulation. Biochemical Journal **335:** 671-679

Phillippy B, Ullah A, Ehrlich K (1994) Purification and some properties of inositol 1,3,4,5,6-pentakisphosphate 2-kinase from immature soybean seeds [published erratum appears in J Biol Chem 1995 Mar 31;270(13):7782]. Journal of Biological Chemistry **269:** 28393-28399

Phillippy BQ (1998) Identification of inositol 1,3,4-trisphosphate 5-kinase and inositol 1,3,4,5-tetrakisphosphate 6-kinase in immature soybean seeds. Plant Physiology **116:** 291-297

Poyner DR, Cooke F, Hanley MR, Reynolds DJ, Hawkins PT (1993) Characterization of metal ion-induced [3H]inositol hexakisphosphate binding to rat cerebellar membranes. Journal of Biological Chemistry **268:** 1032-1038

Qian X, Mitchell J, Wei SJ, Williams J, Petrovich RM, Shears SB (2005) The Ins(1,3,4)P3 5/6-kinase/Ins(3,4,5,6)P4 1-kinase is not a protein kinase. Biochemical Journal **389:** 389-395

Qiu HF, Dubois E, Broen P, Messenguy F (1990) Functional analysis of ARGRI and ARGRIII regulatory proteins involved in the regulation of arginine metabolism in Saccharomyces cerevisiae. Molecular and General Genetics **222:** 192-200

Raboy V (1997) Accumulation and storage of phosphate and minerals. *In* BA Larkins, and Vasil, I.K., ed, Cellular and Molecular Biology of Plant Seed Development. Kluwer Academic Publishers, Dordrecht, Netherlands, pp 441-477

Raboy V (2001) Seeds for a better future: 'low phytate', grains help to overcome malnutrition and reduce pollution. Trends in Plant Science **6:** 458-462

Raboy V (2002) Progress in breeding low phytate crops. Journal of Nutrition **132:** 503s-505s

Raboy V, Dickinson DB (1987) The Timing and Rate of Phytic Acid Accumulation in Developing Soybean Seeds. Plant Physiology **85:** 841-844

Raboy V, Gerbasi PF, Young KA, Stoneberg SD, Pickett SG, Bauman AT, Murthy PPN, Sheridan WF, Ertl DS (2000) Origin and seed phenotype of maize low phytic acid 1-1 and low phytic acid 2-1. Plant Physiology **124:** 355-368

Rasmussen S, Hatzack F (1998) Identification of two low-phytate barley (Hordeum vulgare L) grain mutants by TLC and genetic analysis. Hereditas **129:** 107-112

Ravindran V, Bryden WL, Kornegay ET (1995) Phytates: Occurrence, bioavailability and implications in poultry nutrition. Poultry and Avian Biology Reviews **6:** 125-143

Reddy NR, Pierson MD, Sathe SK, Salunkhe DK (1989) Phytate in cereals and legumes. CRC Press Inc., Boca Raton, Florida

Saiardi A, Caffrey JJ, Snyder SH, Shears SB (2000) Inositol polyphosphate multikinase (ArgRIII) determines nuclear mRNA export in Saccharomyces cerevisiae. FEBS Letters **468:** 28-32

Seeds AM, Sandquist JC, Spana EP, York JD (2004) A molecular basis for inositol polyphosphate synthesis in drosophila melanogaster. Journal of Biological Chemistry **279:** 47222-47232

Shamsuddin AM, Ullah A (1989) Inositol hexaphosphate inhibits large intestinal cancer in F344 rats 5 months after induction by azoxymethane. Carcinogenesis **10:** 625-626

Sharpley AN, Daniel T, T. S, Lemunyon J, Stevens R, Parry R (2003) Agricultural Phosphorus and Eutrophication. *In* U.S. Department of Agriculture. Agricultural Research Service, ARS 149

Shears S, Yang L, Qian X (2004) Cell signaling by a physiologically reversible inositol phosphate kinase/phosphatase. Advances in Enzyme Regulation **44:** 265-277

Shi J, Wang H, Hazebroek J, Ertl DS, Harp T (2005) The maize low-phytic acid 3 encodes a myo-inositol kinase that plays a role in phytic acid biosynthesis in developing seeds. Plant Journal **42:** 708-719

Shi J, Wang H, Wu Y, Hazebroek J, Meeley RB, Ertl DS (2003) The maize low-phytic acid mutant lpa2 is caused by a mutation in an inositol phosphate kinase gene. Plant Physiology **131:** 507-515

Shoemaker R, Keim P, Vodkin L, Retzel E, Clifton SW, Waterston R, Smoller D, Coryell V, Khanna A, Erpelding J, Gai X, Brendel V, Raph-Schmidt C, Shoop EG, Vielweber CJ, Schmatz M, Pape D, Bowers Y, Theising B, Martin J, Dante M, Wylie T, Granger C (2002) A compilation of soybean ESTs: generation and analysis. Genome **45:** 329-338

Shoemaker RC, Polzin K, Labate J, Specht J, Brummer EC, Olson T, Young N, Concibido V, Wilcox J, Tamulonis JP, Kochert G, Boerma HR (1996) Genome duplication in soybean (Glycine subgenus soja). Genetics **144:** 329-338

Smart CC, Fleming AJ (1993) A plant gene with homology to D-myo-inositol-3-phosphate synthase is rapidly and spatially up-regulated during an abscisic-acid-induced morphogenic response in Spirodela polyrrhiza. Plant Journal **4:** 279-293

Spencer JD, Allee GL, Sauber TE (2000) Phosphorus bioavailability and digestibility of normal and genetically modified low-phytate corn for pigs. Journal of Animal Science **78:** 675-681

Steger DJ, Haswell ES, Miller AL, Wente SR, O'Shea EK (2003) Regulation of chromatin remodeling by inositol polyphosphates. Science **299:** 114-116

Stephens LR, Irvine RF (1990) Stepwise phosphorylation of myoinositol leading to myoinositol hexakisphosphate in Dictyostelium. Nature **346:** 580-583

Stevenson-Paulik J, Bastidas RJ, Chiou S-T, Frye RA, York JD (2005) Generation of phytate-free seeds in Arabidopsis through disruption of inositol polyphosphate kinases. Proceedings of the National Academy of Science USA **102:** 12612-12617

Stevenson-Paulik J, Odom AR, York JD (2002) Molecular and biochemical characterization of two plant inositol polyphosphate 6-/3-/5-kinases. Journal of Biological Chemistry **277:** 42711-42718

Subramaniam S (1998) The Biology Workbench - a seamless database and analysis environment for the biologist. Proteins **32:** 1-2

Sugiura SH, Raboy V, Young KA, Dong FM, Hardy RW (1999) Availability of phosphorus and trace elements in low-phytate varieties of barley and corn for rainbow trout (Oncorhynchus mykiss). Aquaculture **170:** 285-296

Sun Y, Thompson M, Lin G, Butler H, Gao Z, Thornburgh S, Yau K, Smith DA, Shukla VK (2007) Inositol 1,3,4,5,6-pentakisphosphate 2-kinase from maize (Zea mays L.): molecular and biochemical characterization. Plant Physiology **144:** 1278-1291

Sun Y, Wilson MP, Majerus PW (2002) Inositol 1,3,4-trisphosphate 5/6-kinase associates with the COP9 signalosome by binding to CSN1. Journal of Biological Chemistry **277:** 45759-45764

Sweetman D, Johnson S, Caddick SE, Hanke DE, Brearley CA (2006) Characterization of an Arabidopsis inositol 1,3,4,5,6-pentakisphosphate 2-kinase (AtIPK1). Biochemical Journal **394:** 95-103

Takazawa K, Perret J, Dumont JE, Erneux C (1991) Molecular cloning and expression of a new putative inositol 1,4,5-trisphosphate 3-kinase isoenzyme. Biochemical Journal **278 (Pt 3):** 883-886

Thomas S, Brake B, Luzio JP, Stanley K, Banting G (1994) Isolation and sequence of a full length cDNA encoding a novel rat inositol 1,4,5-trisphosphate 3-kinase. Biochimica et Biophysica Acta **1220:** 219-222

Torabinejad J, Gillaspy GE (2006) Functional genomics of inositol metabolism. Subcellular Biochemistry **39:** 47-70

van Haastert PJ, van Dijken P (1997) Biochemistry and genetics of inositol phosphate metabolism in Dictyostelium. FEBS Letters **410:** 39-43

Verbsky JW, Chang S-C, Wilson MP, Mochizuki Y, Majerus PW (2005) The pathway for the production of inositol hexakisphosphate in human cells. Journal of Biological Chemistry **280:** 1911-1920

Verbsky JW, Wilson MP, Kisseleva MV, Majerus PW, Wente SR (2002) The synthesis of inositol hexakisphosphate - Characterization of human inositol 1,3,4,5,6-pentakisphosphate 2-kinase. Journal of Biological Chemistry **277:** 31857-31862

Veum TL, Ledoux DR, Bollinger DW, Raboy V, Cook A (2002) Low-phytic acid barley improves calcium and phosphorus utilization and growth performance in growing pigs. Journal of Animal Science **80:** 2663-2670

Vucenik I, Shamsuddin AM (2003) Cancer inhibition by inositol hexaphosphate (IP6) and inositol: from laboratory to clinic. Journal of Nutrition **133:** 3778S-3784S

Vucenik I, Shamsuddin AM (2006) Protection against cancer by dietary IP6 and inositol. Nutrition and Cancer **55:** 109-125

Walker DR, Scaboo AM, Pantalone VR, Wilcox JR, Boerma HR (2006) Genetic mapping of loci associated with seed phytic acid content in CX1834-1-2 soybean. Crop Breeding, Genetics, and Cytology **46:** 390-397

Wilcox JR, Premachandra GS, Young KA, Raboy V (2000) Isolation of high seed inorganic P, low-phytate soybean mutants. Crop Science **40:** 1601-1605

Wilson ME, Majerus PW (1995) Purification and characterization of calf brain inositol 1,3,4-trisphosphate kinase. Blood **86:** 2723-2723

Wilson MP, Majerus PW (1996) Isolation of inositol 1,3,4-trisphosphate 5/6-kinase, cDNA cloning, and expression of the recombinant enzyme. Journal of Biological Chemistry **271:** 11904-11910

Wilson MP, Majerus PW (1996) Purification, cloning, and expression of inositol 1,3,4 trisphosphate 5/6-kinase. FASEB Journal **10:** 2285-2285

Wilson MP, Majerus PW (1997) Characterization of a cDNA encoding Arabidopsis thaliana inositol 1,3,4-trisphosphate 5/6-kinase. Biochemical and Biophysical Research Communications **232:** 678-681

Wilson MP, Sun Y, Cao L, Majerus PW (2001) Inositol 1,3,4-trisphosphate 5/6-kinase is a protein kinase that phosphorylates the transcription factors c-Jun and ATF-2. Journal of Biological Chemistry **276:** 40998-41004

Xia HJ, Brearley C, Elge S, Kaplan B, Fromm H, Mueller-Roeber B (2003) Arabidopsis inositol polyphosphate 6-/3-kinase is a nuclear protein that complements a yeast mutant lacking a functional ArgR-Mcm1 transcription complex. Plant Cell **15:** 449-463

Xu J, Brearley CA, Lin WH, Wang Y, Ye R, Mueller-Roeber B, Xu ZH, Xue HW (2005) A role of Arabidopsis inositol polyphosphate kinase, AtIPK2alpha, in pollen germination and root growth. Plant Physiology **137:** 94-103

Yang X, Shears SB (2000) Multitasking in signal transduction by a promiscuous human Ins(3,4,5,6)P(4) 1-kinase/Ins(1,3,4)P(3) 5/6-kinase. Biochemical Journal **351:** 551-555

York JD, Odom AR, Murphy R, Ives EB, Wente SR (1999) A phospholipase c-dependent inositol polyphosphate kinase pathway required for efficient messenger RNA export. Science **285:** 96-100

Zhang ZB, Yang G, Arana F, Chen Z, Li Y, Xia HJ (2007) Arabidopsis inositol polyphosphate 6-/3-Kinase (AtIpk2beta) is involved in axillary shoot branching via auxin signaling. Plant Physiology **144:** 942-951

Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W (2004) GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. Plant Physiology **136:** 2621-2632

Zonia L, Cordeiro S, Tupy J, Feijo JA (2002) Oscillatory chloride efflux at the pollen tube apex has a role in growth and cell volume regulation and is targeted by inositol 3,4,5,6-tetrakisphosphate. Plant Cell **14:** 2233-2249

VITAE

Amanda Rose Stiles was born to Vernon Daniel Stiles and Elizabeth Rose Stiles on December 5, 1977 in Fairfax, Virginia. In 1995 she graduated from W.T. Woodson High and attended Virginia Polytechnic Institute and State University in Blacksburg, VA. She graduated in 1999 with a B.S. in Biology and a minor in Chemistry. While at Virginia Tech she pursued undergraduate research in the laboratories of Dr. Bob Jones and Dr. Jack Webster in forest and stream ecology, and spent a semester abroad in Ronda, Spain. She also spent a summer participating in a research experience for undergraduates (REU) project studying nitrogen cycling in streams as part of the five year Lotic Intersite Nitrogen eXperiment (LINX). Following graduation she participated in a year long exchange program in Germany in which she attended the University of Saarbrücken for one semester, and then spent six months working at the State Institute of Health and Environment in Saarbrücken, Germany. Upon returning to the United States she worked as a technician at the Norman M. Cole pollution control plant in Lorton, Virginia. In 2002, she returned to Virginia Tech to pursue graduate studies in the Department of Plant Pathology, Physiology, and Weed Science in the laboratory of Dr. Elizabeth Grabau. Her studies focused on elucidating the biosynthetic pathway of phytate, a molecule found in soybean seeds that contributes to phosphorous pollution in water systems. She will join the laboratory of Dr. Norman Terry at University of California, Berkeley in August 2007 where she will study biofuels and phytoremediation as a Postdoctoral Research Associate.