

THE TWO AGPase SUBUNITS EVOLVE AT DIFFERENT RATES IN ANGIOSPERMS,
YET THEY ARE EQUALLY SENSITIVE TO ACTIVITY-ALTERING AMINO ACID
CHANGES WHEN EXPRESSED IN BACTERIA

By

NIKOLAOS GEORGELIS

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2007

© 2007 Nikolaos Georgelis

To my parents, Panayiotis and Kathryn

ACKNOWLEDGMENTS

I wholeheartedly thank my supervisory committee (Curt Hannah, Don McCarty, Ken Cline, Rob Ferl, and Jon Stewart) for their advice and patience. I especially thank my advisor, Curt Hannah, for his tremendous support and the freedom he gave me to develop as an independent scientist. I also acknowledge Hannah Lab members (Sue Boehlein, Janine Shaw, John Baier, and Maureen Clancy) for cooperating and putting up with me all these years. Special thanks go to Janine Shaw and Sue Boehlein who contributed to my research in different ways. I also extend special thanks to Ed Braun for his invaluable help with evolution and Regina Shaw and William Farmerie for their help with sequencing. Finally, I would like to express my gratitude to my parents and my fiancée for their support and patience.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS	4
LIST OF TABLES	7
LIST OF FIGURES	8
ABSTRACT	9
CHAPTER	
1 INTRODUCTION	11
Starch Importance	11
Starch Structure	11
Starch Synthesis	12
Models for Starch Debranching Enzyme (DBE) Action	14
Role of AGPase	14
Evolution of AGPase	15
Allosteric Regulation and Heat Stability of AGPase	16
AGPase Structure	17
Sub-Cellular Localization of AGPase	17
Research Objectives	18
2 EVOLUTION OF ADP-GLUCOSE PYROPHOSPHORYLASE ¹	20
Introduction	20
Materials and Methods	23
Random Mutagenesis	23
Glycogen Detection	24
Glycogen Quantitation	24
DNA Sequencing	25
Classification of Amino Acid Changes (Conservative, Non-conservative)	25
Sequence Retrieval and Alignment	25
Phylogenetic Analysis	26
Purification of BT2	26
Results	27
Alignment of SH2 and BT2 with Homologs	27
Evolutionary Constraints of the Large and Small Subunit of AGPase by Estimating the ω Value	27
Evolutionary Constraints of the Large and Small Subunit of AGPase by Estimating the Amino Acid Substitution Rate	29
Distinguishing Temporary Constraint Relaxation from Permanently Relaxed Constraints in the Large Subunit	30

Probability that a Non-Synonymous Mutation in <i>Sh2</i> or <i>Bt2</i> Affects Enzyme Activity	31
Glycogen Quantitation.....	33
Estimating the Ratio of Non-Synonymous to Synonymous Mutations in Functional Clones	34
Proteins Interacting with BT2	35
Tissue Specificity and Gene Number as a Cause of the Greater Evolutionary Constraints on AGPase Small Subunit	36
Discussion.....	40
3 ISOLATION OF AGRONOMICALLY IMPORTANT ADP-GLUCOSE PYROPHOSPHORYLASE VARIANTS.....	69
Introduction.....	69
Importance of Allosteric Properties of AGPase for Plants and Humans.....	69
Agronomic Importance of AGPase Heat Stability	70
Goal and Approach.....	73
Materials and Methods	73
Random mutagenesis.....	73
Bacterial Expression System	74
Glycogen Detection	74
Glycogen Quantitation.....	74
DNA Sequencing.....	75
Purification of Maize Endosperm AGPase from AC70R1-504 <i>E.coli</i> Cells.....	75
Kinetic Characterization of AGPase	76
Measuring AGPase Specific Activity from Crude or Partially Purified Protein	
Extracts	76
Determining Heat Stability of AGPase	77
Qualitative Determination of the AGPase Purity	77
Protein Blot Analysis of Crude Extracts	77
3D Modeling.....	78
Yeast Two Hybrid	78
Results.....	79
Discussion.....	81
4 DISCUSSION AND FUTURE DIRECTIONS.....	94
APPENDIX SMALL AND LARGE AGPase SUBUNIT ACCESSION NUMBERS	99
LIST OF REFERENCES	103
BIOGRAPHICAL SKETCH	118

LIST OF TABLES

<u>Table</u>	<u>page</u>
2-1. Amino acid changes not tolerated in BT2 and SH2.	48
2-2. Glycogen production from randomly-chosen <i>sh2</i> and <i>bt2</i> mutants expressed in <i>E.coli</i>	49
2-3. Distribution of observed and expected positive clones if multiple intragenic mutations function independently.....	50
2-4. X_f of <i>Sh2</i> and <i>Bt2</i> using different stringencies to classify mutants as non-functional.....	51
2-5. Ratio of non-synonymous to synonymous mutations in functional <i>Sh2</i> and <i>Bt2</i> clones.....	52
2-6. Distribution of positive clones with different numbers of non-synonymous mutations from the heavily mutagenized 3 rd and 4 th libraries.	53
2-7. Amino acid changes of BT2 variants that were expressed in <i>E.coli</i> with SH2 and AGPLEMZM.	54
2-8. Comparison of AGPase small and large subunits from unicellular green algae.	55
3-1. Kinetic properties of purified recombinant AGPase variants.....	85
A-1. Cyanobacterial and small AGPase subunit accession numbers.....	99
A-2. Large AGPase subunit accession numbers.	101

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
2-1. Amino acid conservation patterns of BT2 and SH2. BT2 and SH2 were aligned with 30 and 42 homologs respectively.....	56
2-2. Estimates of AGPase Subunit Gene Trees and ω Values.....	57
2-3. Species tree. Times of divergence are indicated in million of years (MY) at nodes.....	58
2-4. Estimates of the number of amino acid substitutions per site per million years (MY) for AGPase subunits.....	59
2-5. AGPase mutants exhibit different phenotypes following iodine staining.....	61
2-6. Placement of point mutations created by error-prone PCR in <i>Sh2</i> and <i>Bt2</i> clones.....	62
2-7. The distribution of <i>Bt2</i> and <i>Sh2</i> clones with different numbers of missense mutations.....	63
2-8. Glycogen contents of <i>sh2</i> and <i>bt2</i> mutants compared to wildtype AGPase.....	64
2-9. Purification of BT2 using a monoclonal antibody column.....	65
2-10. Comparison of glycogen produced by <i>Bt2</i> mutants expressed with wildtype <i>Sh2</i> and wildtype <i>Agplemzm</i> in <i>E. coli</i>	66
2-11. Phylogenetic tree of AGPase isoforms from plants and unicellular algae.....	67
2-12. Differential rates of evolution between <i>Bt2</i> and <i>Zea mays 2</i> after gene duplication.....	68
3-2. Dot blots of crude extracts from <i>E.coli</i> cells expressing BT2, BT2-TI, MP, MP-TI with the complementary subunit SH2.....	87
3-3. Specific activity of AGPase variants in crude and partially purified protein extracts from non-induced <i>E.coli</i> cells.....	88
3-4. Purification of AGPase. SDS-PAGE of purified recombinant BT2/SH2, TI/SH2, MP/SH2, and MP-TI/SH2.....	89
3-5. Heat stability of purified BT2/SH2, BT2-TI/SH2, and MP/SH2.....	90
3-6. Heat stability of purified MP/SH2 and MP-TI/SH2.....	91
3-7. 3D modeling of BT2 and TI.....	92
3-8. Strength of AGPase subunit-subunit interactions.....	93

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

THE TWO AGPase SUBUNITS EVOLVE AT DIFFERENT RATES IN ANGIOSPERMS,
YET THEY ARE EQUALLY SENSITIVE TO ACTIVITY-ALTERING AMINO ACID
CHANGES WHEN EXPRESSED IN BACTERIA

By

Nikolaos Georgelis

December 2007

Chair: Larkin Curtis Hannah

Major: Plant Molecular and Cellular Biology

ADP-glucose pyrophosphorylase (AGPase) catalyzes a rate-limiting step in starch synthesis. AGPase is a heterotetramer in angiosperms consisting of two identical large and two identical small subunits. The large and the small subunit are encoded by *Shrunken-2* (*Sh2*) and *Brittle-2* (*Bt2*), respectively, in the maize endosperm. Random mutagenesis of *Sh2* and *Bt2* and expression of the mutants in a heterologous bacterial expression system were used to address the cause of the differential sequence divergence between the two AGPase subunits as well as to isolate agronomically promising AGPase variants.

It has been hypothesized that the AGPase small subunit is more conserved compared with the large subunit because a higher proportion of the amino acids of the small subunit are required for enzyme activity compared with the large subunit. Evolutionary analysis indicates that the AGPase small subunit has been subjected to more intense purifying selection than has the large subunit in the angiosperms. However, random mutagenesis and expression of *Sh2* and *Bt2* in bacteria show that the two AGPase subunits are equally predisposed to enzyme activity-altering amino acid changes when expressed in one environment with a single complementary subunit. As an alternative hypothesis, I suggest that the small subunit exhibits more evolutionary constraints

in planta than does the large subunit because it is less tissue-specific and thus must form functional enzyme complexes with different large subunits. Independent approaches provide data consistent with this alternative hypothesis.

Random mutagenesis of *Bt2* and expression in *Escherichia coli* with a functional *Sh2* led to the isolation of a heat stable AGPase variant (BT2-TI). BT2-TI differs from BT2 by one amino acid change. Introduction of this amino acid change into an existing, agronomically promising AGPase variant (MP) led to a variant (MP-TI) with increased heat stability compared to MP. MP-TI may lead to an increase in starch yield when expressed in cereal endosperm.

CHAPTER 1 INTRODUCTION

Starch Importance

Starch serves as a means of energy storage and as a mechanism to transfer energy from parents to progeny. Part of the fixed carbon gives rise to starch in leaves. This is converted back to monosaccharides at night when plants become heterotrophic (Zeeman et al., 2007). Starch is also formed in large quantities in seed and more specifically the monocot endosperm where it is mobilized by the young seedling during its early heterotrophic growth. Starch from the monocot endosperm and potato tubers is also important to animals, including humans, since it provides the vast majority of the daily needs in carbohydrates (James et al., 2003). U.S. is the major producer of maize while Europe relies more on potato and wheat. Finally, starch has several uses in industrial applications. It can be used as a source for alcohol production while its gelling capacity makes it useful in food and drug industry. Its structural properties are crucial for these uses (Jobling, 2004).

Starch Structure

Starch in maize endosperm consists of ~25% of amylose and ~75% amylopectin. Amylose is virtually a linear glucan, with glycosyl groups connected by α -1,4 linkages. In amylopectin, approximately 5% of the linkages are α -1,6 bonds that create glucan branches. α -1,6 bonds are not randomly distributed but are concentrated in distinct areas that are known as amorphous lamellae (Myers et al., 2000; James et al., 2003). In areas where there is lack of α -1,6 bonds, linear glucans are intertwined in double helices creating a crystalline structure known as crystalline lamellae. Hence, amylopectin has a semi-crystalline structure consisting of concentric areas of amorphous and crystalline lamellae. Specific primers for starch synthesis remain elusive. Amylopectin shows similarity to glycogen in that glucans are connected by α -1,4 and α -1,6

bonds (Ball et al., 1996). However, glycogen is more branched than amylopectin (10% vs 5%), its α -1,6 linkages are evenly distributed and it is water soluble. By evolving a semi-crystalline water insoluble means of storage, plants have been able to store more energy without changing the osmotic pressure inside the cells.

Starch Synthesis

The first committed step in starch synthesis is catalyzed by ADP-glucose pyrophosphorylase (AGPase) (Hannah, 2005). It involves the conversion of glucose-1-P (G-1-P) and ATP to ADP-glucose and pyrophosphate (PPi). ADP-glucose is the prevalent if not sole substrate of starch synthases (SS) that extend linear glucan chains by creating α -1,4 bonds between adjacent glycosyl groups. The several isoforms of SS can be classified into two groups depending on their ability to bind to the starch granule (James et al., 2003; Ball and Morell, 2003; Tetlow et al., 2004; Hannah, 2005). Mutant analysis indicated that granule-bound starch synthase I (GBSSI) is important in amylose synthesis. GBSSI is encoded by the *Waxy* gene and mutants in that locus resulted in reduced amounts of amylose (Nelson and Rines, 1962). This provides evidence that amylopectin synthesis does not require prior amylose formation. GBSSI also participates in the extension of long glucans within amylopectin. The second group of SSs (SSI, SSII, SSIII, SSIV/V) can be soluble in the stroma or granule bound depending on the tissue, species and developmental stage. James et al. (2003) proposed that SSI is important at the early stages of glucan chain elongation while SSII and SSIII are important at late stages. SSIII has a longer N terminus than other SSs and contains a 14-3-3 binding motif. Sehnke et al. (2001) showed that starch was increased 2- to 4- fold in Arabidopsis leaves when the epsilon subgroup of 14-3-3 proteins was reduced by antisense technology. This indicated that 14-3-3 proteins have an important role in starch synthesis perhaps through SSIII. However, the exact roles of different

SSs are not clear due to pleiotropic effects and subtle phenotypes of single mutants (James et al., 2003).

Starch branching enzymes (SBE) hydrolyze an α -1,4 bond within α -1,4-linked glycosyl groups and create α -1,6 bonds. SBE isoforms can be classified into two groups (I, II). In monocots, group II can be further divided into two classes (a, b) (Yao et al., 2004). Different isoforms are expressed in different tissues and developmental stages depending on plant species. SBEs have different substrate specificities (Guan and Preiss, 1993; Misuno et al., 2001) and chain length of the products (Takeda et al., 1993). However, the exact role of each isoform remains unresolved. Increased amylose/amylopectin has been observed only in SBEII mutants including a SBEIIb mutant known as *amylose extender (ae)* in maize. More severe phenotypes are observed in double mutants involving SBEI and SBEII. A maize endosperm SBEI mutant lacks an obvious mutant phenotype indicating that its function can be at least partially complemented by SBEII (Yao et al., 2004).

Starch debranching enzymes (DBE) hydrolyze α -1,6 bonds and can be classified into two groups, isoamylases and pullulanases, based on substrate specificity. Mutant analysis of both classes has indicated that they are both important in synthesis and degradation of starch. *Sugary1*, an isoamylase mutant, results in reduced starch and increased amounts of sugars and water soluble polyglucans (WSP) also termed phytyglycogen (Pan and Nelson, 1984). Also long-chained amylopectin increases at the expense of short-chains in the mutant. Different DBEs appear to have overlapping roles but they are not completely redundant (James et al., 2003; Hannah, 2005).

Bacteria and plants share a similar pathway for synthesizing glycogen and starch respectively. In bacteria, the first committed step is also catalyzed by AGPase. However, the rest

of the glycogen synthesis pathway is simpler than in plants. Indeed, there are only single isoforms of a glycogen synthase, glycogen branching enzyme and glycogen debranching enzyme with the latter participating exclusively in glycogen degradation (Ball and Morrell, 2003).

Models for Starch Debranching Enzyme (DBE) Action

It is now well accepted that DBEs participate in both starch synthesis and degradation. Three models for the role of DBEs in starch synthesis have been proposed (Myers et al., 2000; James et al., 2003; Ball and Morell, 2003; Tetlow et al., 2004; Hannah, 2005). The glucan trimming model proposes that DBEs directly participate in amylopectin synthesis by removing α -1,6 bonds non-randomly (Ball et al., 1996). This allows for the formation of highly branched areas and areas with packed linear α -1,4 chains and leads to the semicrystalline structure of amylopectin. Recently, it was shown that DBEs exist, in vivo, in complexes of high molecular weight supporting the notion that they may not hydrolyze closely spaced branches due to steric hindrance. Hence their action maybe limited to loosely and medium branched areas. The WSP clearing model proposes that DBEs catabolize WSP providing the substrates for SSs and BEs. Finally, from studies of DBE mutants in barley, Burton et al. (2002) suggested that DBEs may participate in the degradation of an amylopectin primer reducing the number of starch granules per plastid. The above models are not necessarily mutually exclusive.

Role of AGPase

AGPase catalyzes the first committed step in starch synthesis. It converts G-1-P and ATP to ADP-glucose and PPi. AGPase is a homotetramer in bacteria including cyanobacteria. However, it is a heterotetramer in algae and higher plants. It consists of two identical small and two identical large subunits. The large and the small subunit are encoded by *Shrunken-2* (*Sh2*) and *Brittle-2* (*Bt2*) respectively in the maize endosperm (Hannah and Nelson, 1976; Bae et al., 1990; Bhave et al., 1990).

The importance of maize endosperm AGPase in starch synthesis has been shown by the kernel phenotype of mutants in either subunit of the enzyme. Indeed, such mutants result in shrunken kernels and a large reduction in endosperm starch content (Tsai and Nelson, 1966; Hannah and Nelson, 1976). Also, knockout mutants of the structural genes of AGPase in *Arabidopsis* led to severely reduced amounts of starch in leaves (Lin et al., 1988a;1988b).

Evolution of AGPase

The two subunits of plant AGPase show identity and similarity pointing to gene duplication. Each subunit apparently has undergone several duplication events, giving rise to different tissue specific isoforms. Even though the two subunits are not functionally interchangeable the small subunit is more conserved among tissues and species than is the large subunit. Smith-White and Preiss (1992) suggested that the small subunit might have been under more selection pressure than the large subunit. Ballicora et al. (2003) and Greene et al. (1996) working with potato tuber AGPase have proposed that the two subunits have different roles in AGPase function. It has been proposed that the small subunit is catalytic while the large subunit modifies the response to allosteric regulators. Being catalytic, the small subunit has more constraints in evolution than does the large subunit. However, higher selective constraints for the small subunit than the large subunits have never been shown definitively. Also, the relative roles of AGPase subunits in enzyme function are far from clear. Recent evidence shows that both subunits control kinetic and allosteric properties (Hannah and Nelson, 1976; Cross et al., 2004; Cross et al., 2005; Hwang et al., 2006; 2007).

It is possible that different evolutionary histories or mutation rates for the small and large subunit may explain the higher sequence conservation of the small subunit in plants. It is also possible that different evolutionary constraints, if they exist, maybe explained by reasons other than relative roles in enzyme activity.

Allosteric Regulation and Heat Stability of AGPase

One feature of AGPase isoforms is that they are allosterically regulated by small effector molecules that indicate the cellular energy status. In bacteria, AGPase is activated by fructose 6-P, fructose 1,6-bisphosphate, or pyruvate, and deactivated by AMP (Ballicora et al., 2003). In cyanobacteria, green algae and angiosperms, AGPase is activated by 3-PGA, the first carbon assimilatory product and deactivated by Pi. The importance of AGPase regulation by 3-PGA and Pi is obvious in source leaves, where sucrose and starch synthesis need to be coordinated to meet the needs of the whole plant (Hannah, 2000). However, in sink tissues, the physiological importance of such regulation is not clear. AGPase regulation in sink tissues maybe a by-product of evolution (Hannah, 2000). Alternatively, the allosteric regulator levels maybe an indicator of the amount of energy imported into the sink tissue. These levels may coordinate the amount of energy to be used by the cell, i.e. glycolysis, and the amount to be stored as starch.

AGPase catalyses a rate-limiting step in glycogen and starch synthesis. The first genetic evidence of the rate-limiting role of AGPase and the importance of its allosteric regulation came from *E.coli*, where an AGPase variant (*glgC-16*) that had altered allosteric properties resulted in glycogen overproduction (Preiss and Romero, 1989). Expression of *glgC-16* in the potato tuber resulted in 35% starch increase (Stark et al., 1992). Expression of an allosteric variant of AGPase (*Rev6*) in maize endosperm resulted in 15% and 40% starch increase in a single and triple dose respectively (Giroux et al., 1996). Starch increase was due to increased kernel size.

Heat stability of maize endosperm AGPase is another feature that is agronomically important (Greene and Hannah, 1998a; 1998b). A heat stable variant of maize endosperm AGPase, *HS33*, along with *Rev6*, described above, were placed in a construct that increased starch production in wheat, rice, and maize (Smidansky et al., 2002; 2003; Greene and Hannah,

in preparation). Unexpectedly, starch increase was due to increased kernel number. Clearly, allosteric properties and heat stability of AGPase are agronomically important.

AGPase Structure

The crystal structure of an AGPase heterotetramer has not been resolved. The only available crystal structure is a potato tuber small subunit homotetramer in its inactivated form (Jin et al., 2005). Each monomer can be divided into two modules, the N and the C terminal domain. The active site of the enzyme is in the N terminus and it resembles a Rossman fold (Rao and Rossman, 1973). Rossman fold is an a/b topology that consists of 3 β -sheets and 2 α -helices (β -sheet – α -helix – β -sheet – α -helix – β -sheet). It is often present in nucleotide-binding proteins. The C terminus forms a β -helix that participates in allosteric regulation and oligomerization. The link between the N and C terminal domain is heavily involved in subunit-subunit interactions. The N terminus is also involved in subunit-subunit interactions including a cysteine in the small subunit of some isoforms that forms a disulfide bridge between the small subunits of the enzyme complex under oxidative conditions.

Sub-Cellular Localization of AGPase

AGPase is localized in the plastids, in most plant tissues. However, this is not true for monocot endosperms where AGPase is localized in the cytosol. Cell fractionation experiments in maize endosperm resulted in detection of 90-95% of AGPase activity in the cytosol (Denyer et al., 1996). SH2 and BT2 precursor and mature proteins are of the same size pointing to a lack of transit peptides (Giroux and Hannah, 1994). Additionally, maize amyloplasts possess a plastidic translocator BT1 that shows similarity to mitochondrial adenylate translocators (Shannon et al., 1998). BT1 maybe involved in transport of ADP-glucose into the amyloplast since loss-of-function mutations in *bt1* resulted in an accumulation of ADP-glucose in the cells (Shannon et al., 1996; Kirchberger et al., 2007). If AGPase were plastidic, ADP-glucose would not

accumulate, rather it would immediately serve as a glucose donor for starch formation by starch synthases. Finally, metabolic profiling studies support the existence of a cytosolic form of AGPase in monocot endosperms (Beckles et al., 2001). Here, the authors used a high ratio ADP-glucose/UDP-glucose as an indication for the presence of a cytosolic AGP. That ratio was much higher in monocot endosperms compared to dicot or other monocot tissues. On the other hand, Brangeon et al. (1997) showed that AGP was confined to the amyloplasts by using in situ immunolocalization and light microscopy. Their data and data from the literature cited above may not be mutually exclusive. At that level of resolution, the authors admitted that it was not possible to locate AGP inside or outside of the plastids. Taking into account that most literature supports a cytosolic localization, perhaps AGPase is bound to the outside of the amyloplasts. For instance, Shannon et al. (1998) suggested a cytosolic localization from nonaqueous cell fractionation studies, but they also acknowledged the possibility that AGP may be confined to the outside or the intermembrane domain of the amyloplasts.

A feature of plastidic small subunit isoforms is the presence of a cysteine in the N terminus that is not found in the cytosolic small subunits. It has been shown that the cysteine of the two small subunits of AGPase can form a disulfide bridge between the small subunits (Fu et al., 1998). Reduction of the disulfide bridge, under light or after feeding with sucrose, upregulates AGPase (Tiessen et al., 2002). This disulfide bridge maybe a form of post-translational regulation of the plastidic isoforms not shared by the cytosolic isoforms.

Research Objectives

The first goal of this research was the elucidation of the differential sequence divergence between the small and the large AGPase subunits. More specifically, phylogenetic analysis for the two subunits was performed to compare their relative evolutionary constraints. Additionally, the premise that the large subunit can tolerate more amino acid changes than the small subunit

with respect to AGPase activity was tested. Additional reasons that could explain the higher sequence conservation of the AGPase small subunit were explored.

The second goal was to isolate agronomically important variants of maize endosperm by using a bacterial expression system. Mutants with altered allosteric properties and increased heat stability would be particularly promising for future agronomic gain after expression in plants.

CHAPTER 2 EVOLUTION OF ADP-GLUCOSE PYROPHOSPHORYLASE¹

Introduction

ADP-glucose pyrophosphorylase (AGPase) is an allosteric enzyme that catalyzes a rate-limiting step in starch synthesis, the conversion of glucose-1-P (G-1-P) and ATP to ADP-glucose and pyrophosphate (PPi). ADP-glucose then serves as the main if not sole precursor for starch synthesis by starch synthases (Hannah, 2005). Angiosperm and green algal AGPases are heterotetramers consisting of two identical large and two identical small subunits. In the case of the maize endosperm AGPase, the subject of this research, the large subunit is encoded by *Shrunken-2* (*Sh2*) and the small subunit by *Brittle-2* (*Bt2*) (Bae et al., 1990; Bhave et al. 1990). SH2 and BT2 show considerable amino acid identity (43.2%) and similarity (61%). They also exhibit identity to the sole subunit of the homotetrameric *Escherichia coli* (*E.coli*) AGPase (SH2: 24.4%, BT2: 29.0%). Gene duplication early in the evolution of the plant lineage followed by sequence divergence is the most parsimonious explanation. The two subunits are not functionally interchangeable as shown by analysis of mutants. Loss of *Sh2* or *Bt2* function abolishes more than 90% of endosperm AGPase activity (Hannah and Nelson, 1976). Expression of each of the maize endosperm AGPase subunits separately in *E.coli* showed that SH2 or BT2 alone gives only 3.5 % and 2.5 % activity, respectively, of the heterotetramer (Burger et al., 2003).

Following the initial duplication, each subunit underwent several duplication events, giving rise to different tissue specific isoforms. In general, angiosperm species contain fewer small subunit genes compared to large subunit genes. *Arabidopsis* (*Arabidopsis thaliana*)

¹Reprinted with permission from Georgelis, N., Braun, E.L., Shaw, J.R., and Hannah, L.C. (2007). The two AGPase subunits evolve at different rates in angiosperms, yet they are equally sensitive to activity altering amino acid changes when expressed in bacteria, *Plant Cell* 19: 1458–1472.

contains one functional small AGPase subunit and four tissue-specific large subunit genes (Crevillen et al., 2003; 2005). Rice (*Oryza sativa*) contains two small subunit genes and four large subunit genes (Akihiro et al., 2005; Ohdan et al., 2005). Populus (*Populus trichocarpa*) contains one small subunit and six large subunit genes (Tuskan et al., 2006). Even though three small subunit and three large subunit genes have been cloned in maize (Bae et al., 1990; Bhave et al., 1990; Giroux et al., 1995; Shaw and Hannah unpublished) the exact number of maize AGPase genes is unknown since the complete maize genome sequence is presently unavailable.

Small subunits show striking conservation among species while the large subunits are less conserved (Smith-White and Preiss, 1992; see Results). Smith-White and Preiss (1992) suggested that the small subunit has more selective constraints than the large subunit. The idea that the small subunit has been subject to greater constraint than has the large is supported by the higher percentage of identity between cyanobacterial AGPase and the small subunit compared to the large subunit (Greene et al., 1996). While the exact role of each subunit remains unclear, Ballicora et al. (2003; 2005), Frueauf et al. (2003) and Greene et al. (1996) working with potato tuber AGPase suggested that the two subunits have different roles in AGPase function. The small subunit has been proposed to be catalytic while the large subunit presumably modifies the response to allosteric regulators. According to this hypothesis, this presumed specificity in roles has placed different evolutionary constraints on the two proteins. The large subunit, then, can accumulate more non-deleterious non-synonymous mutations for AGPase activity than can the small subunit. However, the roles of each AGPase subunit in enzyme function are not clear, since each subunit has activity when expressed alone in *E.coli* (Iglesias et al., 1993; Burger et al., 2003) and mutations in either subunit affect both catalytic and allosteric properties (Hannah and Nelson, 1976; Cross et al., 2004; Cross et al., 2005; Hwang et al. 2006; 2007). It should also be

noted that the regulatory subunits of the Archaeal type H⁺-ATPase (A-ATPase) and vacuolar type H⁺-ATPase (V-ATPase) evolved at a slower rate compared to the catalytic subunits (Marin et al., 2001).

Herein, I show that the higher degree of sequence divergence in the large subunit is, in fact, due to increased evolutionary constraints on the small subunit compared to the large subunit. The different evolutionary constraints were shown by estimating the overall ratio of non-synonymous substitutions per non-synonymous site to synonymous substitutions per synonymous site (ω) for the large and small subunit coding genes as well as the number of amino acid substitutions in the large and in the small subunits per site per million years (aass MY⁻¹).

I performed two independent tests to determine whether the difference in evolutionary rates of the two subunits reflected different sensitivity of the subunits to activity altering amino acid changes. First, random mutations were introduced separately into *Sh2* and *Bt2* subunits and AGPase activity was monitored in an *E.coli* expression system. Mutant genes were sequenced. The probability that a non-synonymous mutation alters activity was calculated using methodology and a formula recently described by Guo et al. (2004). These experiments showed that the probability of obtaining activity-altering non-synonymous mutations was equal for the two subunits in *E.coli*. Second, functional clones from heavily mutagenized libraries were sequenced and the ratio of non-synonymous to synonymous mutations was calculated. This ratio was the same for both subunits. Our results indicate then that SH2 and BT2 are equally predisposed to activity-altering non-synonymous mutations. The subunits are equally robust to random amino acid changes affecting AGPase activity when expressed in one common environment (*E.coli*) with one complementary subunit.

I examined other hypotheses to account for the different evolutionary constraints of the two subunits. First I asked whether the BT2 protein interacted with other proteins. The only protein found to interact with BT2 was SH2. This result suggests that the AGPase small subunit does not have additional functions mediated by protein-protein interactions. I did, however, find that the effect of some *Bt2* mutants on AGPase activity in the *E. coli* expression system was dependent on whether the interacting partner was *Sh2* or the maize embryo large subunit (*Agplemzm*) (Giroux and Hannah, 1994; Giroux et al., 1995; Georgelis et al., 2007). This is significant since plants generally have fewer small subunits genes than large subunit genes, and therefore individual small subunits must be expressed in several tissues and interact with multiple large subunit partners. Accordingly, I suggest that the greater evolutionary constraints on the angiosperm small subunit reflect, at least in part, the fact that the angiosperm AGPase small subunit is less tissue-specific than is the large subunit and it must interact with different large subunits whereas each large subunit interacts with only one small subunit and is expressed in fewer tissues. As a test of this hypothesis, I examined the rate of sequence divergence of the small and large subunits in unicellular green algae. These organisms contain only one functional small and one functional large AGPase subunit. The rate of sequence divergence is identical for the two subunits. Also, several large subunit genes are expressed in at least some tissues; with the resulting redundancy possibly reducing selection pressure on individual large subunit genes.

Materials and Methods

Random Mutagenesis

Mutations were introduced into *Sh2* and *Bt2* by PCR random mutagenesis (GeneMorph II EZClone Domain Mutagenesis Kit, Stratagene). A mixture of non-biased, error-prone DNA polymerases was used to introduce point mutations. Wildtype *Sh2* and *Bt2* coding sequences in pMONc*Sh2* and pMONc*Bt2* (Giroux et al., 1996) respectively were used as templates for PCR.

Two pairs of primers (*Sh2*: 5'GAAGGAGATATATCCATGG 3', 5' GGATCCCCGGGTACCGAGCTC 3' *Bt2*: 5'GAAGGAGATATATCCATGG 3', 5' GTTGATATCTGAATTCGAGCTC 3') flanking *Sh2* and *Bt2* were used for error-prone PCR. Mutant *sh2* clones produced by PCR were subcloned into vector pMONc*Sh2* according to Stratagene protocols. pMONc*Sh2* was then used to transform *E. coli* strain AC70R1-504 that contained wildtype *Bt2* in the compatible vector pMONc*Bt2*. Mutant *bt2* clones produced by PCR were subcloned into vector pMONc*Bt2*. pMONc*Bt2* was then used to transform *E. coli* strain AC70R1-504 that contained wildtype *Sh2* in the compatible vector pMONc*Sh2*.

Glycogen Detection

Glycogen synthesis was detected by production of brown staining colonies following a two-minute exposure to iodine vapors. *E. coli* cells were grown on Luria broth (LB) media in the presence of 75 µg/mL spectinomycin, 50 µg/mL kanamycin and 2% w/v glucose. Iodine staining provided a fast and convenient means to identify active and inactive AGPases. Colonies with inactive AGPase produced no color following contact with iodine vapors while active AGPase produced glycogen and, in turn, brown staining with iodine.

Glycogen Quantitation

Glycogen quantitation of AGPase mutants was performed by phenol reaction (Hanson and Phillips, 1981). In brief, glycogen was extracted from *E. coli* cells (OD₆₀₀=2.0) grown in LB containing 2% w/v glucose by boiling for 3 hours in 50% w/v KOH. Glycogen was then precipitated by adding ethanol to 70% v/v and centrifuging at 10000 x g for 10 min. After pellet drying, 200 µl deionized water, 200 µl of 5% w/v phenol and 1 mL of concentrated sulfuric acid were added. Glycogen was estimated by the absorbance at 488 nm.

DNA Sequencing

Five hundred and seventy six *sh2* and *bt2* mutants were double-pass sequenced by the Genome Sequencing Services Laboratory (GSSL) of the Interdisciplinary Center for Biotechnology Research at the University of Florida. Data analysis was performed by Bioedit software (Hall, 1999).

Classification of Amino Acid Changes (Conservative, Non-conservative)

BLOSUM62 matrix was used for determining conservative and non-conservative amino acid changes (Henikoff and Henikoff, 1992). Amino acid changes that were given a positive or neutral score by the BLOSUM62 matrix were considered conservative (Cargill et al., 1999). Amino acid changes that were given a negative score were considered non-conservative.

Sequence Retrieval and Alignment

I retrieved all full-length plant large and small subunits from the National Center for Biotechnology Information (NCBI) database. The AGPase isoforms of *Populus trichocarpa* were obtained from US DOE Joint Genome Institute (JGI) (http://genome.jgi-psf.org/Poptr1_1/Popt_r1_1.home.html) by a tblastn search using BT2 and SH2 as queries. The AGPase isoforms of *Chlamydomonas reinhardtii*, *Ostreococcus tauri* and *Ostreococcus lucimarinus* were obtained from JGI (<http://genome.jgi-psf.org/Chlre3/Chlre3.home.html>, <http://genome.jgi-psf.org/Ostta4/Ostta4.home.html>, http://genome.jgi-psf.org/Ost9901_3/Ost9901_3.home.html) either as annotated sequences or by tblastn searches. Alignments of 31 full length small subunits sequences and 43 full length large subunits sequences were obtained by using the BIOEDIT software (Hall, 1999) with BLOSUM62 matrix followed by manual inspection of the resulting alignments. The poorly aligned N' termini (~50 amino acids) of both the large and the small subunits were excluded from alignment. Pairwise identity was estimated using BIOEDIT and the IDENTITY matrix.

Phylogenetic Analysis

Maximum likelihood (ML) estimates of gene trees were obtained using nucleotide data and the GARLI (Genetic Algorithm for Rapid Likelihood Inference) software (Zwickl, 2006). These analyses used the default General Time Reversible (GTR) model of sequence evolution (Yang, 1994) and accommodating among-sites rate heterogeneity by assuming some sites are invariant and the rates at the remaining sites are G (gamma) distributed (Gu et al., 1995). Bootstrap values (Felsenstein, 1985) were also calculated by GARLI using 100 replicates. Branch lengths based upon amino acid changes were estimated using AAML, which is part of the PAML (Phylogenetic Analysis by Maximum Likelihood) package (Yang, 1997), using Dayhoff (PAM) empirical model (Dayhoff et al., 1978) with G distributed rates. ω for the large and small subunit coding sequences were estimated using CODEML, distributed with PAML package. Reconciled tree analysis, which allows the differentiation between divergences likely to represent gene duplications and those likely to represent speciation events given a specific species tree (Goodman et al., 1979), was performed using GeneTree 1.3.0 (Page, 1998). ML estimates of phylogenetic trees using amino acid sequences were obtained by using the rapid search algorithm implemented in PhyML (Guindon and Gascuel, 2003) and the Dayhoff model with G distributed rates, and the same program was used to estimate bootstrap support using 100 replicates.

Purification of BT2

Formaldehyde cross-linking of proteins associated in maize endosperm cells was done according to a modified protocol described by Rohila et al. (2004). In brief, developing maize endosperm harvested 18 days after pollination and stored at -80°C was cut into small pieces and treated with 2% w/v formaldehyde in ice-cold PBS buffer for 4 hours. Cross-linking was quenched by addition of 2M glycine for 1 hour. Total proteins were then extracted as described by Boehlein et al. (2005). A BT2 monoclonal antibody column described in Boehlein et al.

(2005) was used to purify BT2 along with associated proteins. The purification products were run on a 7.5% SDS polyacrylamide gel (+DTT) and visualized by staining with Coomassie blue (Laemmli, 1970).

Results

Alignment of SH2 and BT2 with Homologs

A previous study showed that the AGPase small subunit exhibits greater sequence conservation than does the large subunit (Smith-White and Preiss, 1992). To explore the much larger collection of AGPase sequences and update the comparison of the two AGPase subunits, I aligned all available angiosperm AGPase large and small subunits (Appendix A). The hypervariable amino termini of the two subunits, shown in Figure 2-1, were excluded from the alignments. These alignments showed that the average pairwise identity is 70.8% for large subunits and 91.3% for small subunits. More strikingly, 21.6% of SH2 residues and 64.1% of BT2 residues are invariant within the respective family (Figure 2-1). It is also apparent that the large subunit is more diverse than is the small subunit throughout their sequences (Figure 2-1).

Evolutionary Constraints of the Large and Small Subunit of AGPase by Estimating the ω Value

To determine whether the difference in amino acid identity is due to evolutionary constraints or to the timing and pattern of gene duplication events, the overall ω for the large and the small subunit genes was calculated. To check for changes in evolutionary constraints within a gene tree that includes the large and the small AGPase subunits from angiosperms, I applied two nested likelihood models to the gene tree in Figure 2-2A. A model that assumes two different ω values, one for the large and one for the small subunits, was favored over a model that assumes one ω for both subunits by a likelihood ratio test (LRT; $P < 0.001$) (Felsenstein, 1981; Goldman, 1993; Yang et al., 1995; Huelsenbeck and Rannala, 1997; Hileman and Baum, 2003). The overall

ω estimates were 0.087 for the large subunit and 0.032 for the small subunit (Figure 2-2A). This result indicates that the large subunit evolved under reduced purifying selection compared with the small subunit. To monitor for saturation of the rapidly accumulating synonymous substitutions that might cause bias, ω values were calculated for sets of closely related isoforms within each subunit. The small subunit genes were clustered into two major groups (Figure 2-2B). Groups 1 and 2 consist of genes expressed in monocots and eudicots, respectively. The large subunit genes were clustered into five major groups (Figure 2-2C). Group 1 comprises genes expressed in monocot and eudicot leaf tissue. Group 2 includes genes expressed in eudicot source and sink tissues. Groups 3 and 4 contain genes expressed in monocot seed and eudicot sink tissues, respectively. Group 5 is composed of only two sequences whose tissue specificity has not been studied in detail. A model that assumes a different ω value for each group of the small subunit was favored over a model that assumes a single ω for all small subunit groups (LRT; $P < 0.001$). On the other hand, groups 1, 2, 5 of the large subunit share a common ω . Group 4 has a significantly higher ω (LRT; $p=0.002$) than groups 1, 2, 5, while group 3 evolves significantly faster than group 4 (LRT; $p<0.001$). The ω values of the large subunit groups ranged from 0.073 to 0.132 (Figure 2-2C), while small subunit groups varied from 0.027 to 0.054 (Figure 2-2B). By using the gene tree shown in Figure 2-2A, I performed an LRT test that showed that the large subunit groups with the lowest ω (groups 3, 4, 5) evolved significantly faster than the small subunit group with the highest ω (group 1) ($p=0.003$). These results indicate that mutational saturation is unlikely to have had a major impact on calculated ω values and further suggest that the large subunit has been subjected to fewer constraints than has the small subunit.

Evolutionary Constraints of the Large and Small Subunit of AGPase by Estimating the Amino Acid Substitution Rate

Although ω provides a convenient method to examine evolutionary constraints, it has typically been used to examine sequences that have diverged relatively recently. There is evidence that the number of synonymous substitutions per site (K_S) is often saturated even between recent evolutionary divergences (~100 MY ago)(Rabinowicz et al., 1999). Although such saturation is likely to represent a worst-case scenario, the accuracy of ω when applied to ancient divergences is an open question. To complement and validate the previous method, I estimated the absolute rate of sequence evolution (aass MY^{-1}) for all the large and the small subunits. Estimates of the absolute rate for each subunit were obtained by comparing estimates of the number of amino acid substitutions per site to the most recent speciation event that was reliably dated. Paths from each extant sequence to a dated speciation event were chosen to be unique so the rate estimates did not represent pseudoreplication.

The estimates of phylogeny for both subunits reflect both speciation and successful gene duplication events. If the estimate of phylogeny for the large and small subunit gene trees is correct, it is possible to distinguish between speciation and successful duplication events by reconciling the gene trees with the probable species tree (Figure 2-3). Reconciled tree analysis is a procedure that minimizes the number of duplication events needed to explain a specific gene tree given the species tree (Goodman et al., 1979). However, estimates of phylogeny for individual genes are not without error, and accommodating error in gene trees is difficult in reconciled tree analyses (Page and Cotton, 2000). I assessed the impact of uncertainty in the gene tree using a procedure described by Chen et al. (2000), who rearranged nodes with low bootstrap support to increase congruence with the species trees. This was done by manually evaluating all gene tree rearrangements that do not alter well-supported nodes (bootstrap value $\geq 70\%$) that

might increase congruence with the species tree. This yielded two estimates of the numbers of successful duplications, one based on the optimal gene tree and a conservative estimate based on well-supported nodes. Both reconciled tree analyses showed that the large subunit genes underwent more successful duplications than did the small subunit genes (16 versus 7 given the optimal trees and 11 versus 4 given the conservative criterion) (Figure 2-4A, B).

The absolute rate of amino acid substitution (aass MY^{-1}) can be estimated by examining branch lengths for those divergences in the gene tree that reflect speciation events with divergence time for species from literature (de Sa and Drouin, 1996; Grant et al., 2000; Ku et al., 2000; Wikström et al., 2001; Gaut, 2002; Choi et al., 2004; Schlueter et al., 2004; Soltis et al., 2005; Bar-Or et al., 2006; Yang et al., 2006). All divergence times were examined for consistency with the fossil record (Crepet et al., 2004) and attention was restricted to the tips in the phylogeny. The average rate was 0.001261 aass MY^{-1} for the large subunit and 0.000462 aass MY^{-1} for small subunit (Figure 2-4C), a 2.7-fold difference that was both significant ($P=0.0006$ by Student's unpaired t-test) and in agreement with the results of the ω analysis. Limiting the analysis to more strongly supported duplication events (those that were not eliminated when the bootstrap was considered) does not affect the estimate of the average rate of evolution of the small or large subunit. Overall, these results indicate that the small subunit evolved more slowly than did the large subunit, and suggest that the small subunit had more evolutionary constraints *in planta* than did the large subunit.

Distinguishing Temporary Constraint Relaxation from Permanently Relaxed Constraints in the Large Subunit.

Lynch and Conery (2000) showed that recent duplicates (defined as those with $K_S < 0.1$) accumulated more non-synonymous mutations per site (K_A) than did older duplicates ($K_S > 0.1$). The large subunit appears to have undergone more successful gene duplications than has the

small subunit (Figure 2-4A,B), so it is possible that the elevated ω observed for the large subunit relative to the small subunit reflects more frequent episodes of temporary relaxation and/or positive selection rather than permanently relaxed constraints. To distinguish between these possibilities, branches of the large subunit gene tree (Figure 2-4A) were broken into two groups. One group was composed of those that follow a strongly supported duplication event and have a K_S approximately 0.1 or lower. The other group was composed of all other branches. I used ML to estimate ω for each group revealing that ω for short branches following duplications was 0.125 while ω for all the other branches was 0.086. This difference is statistically different (LRT; $P=0.005$). However, the branches of the large subunits, excluding those that follow a strongly supported duplication event and have a K_S approximately 0.1 or lower, had an ω significantly higher than the branches of the small subunit in the gene tree shown in Figure 2-2A (LRT; $P<0.001$). These results indicate that temporary relaxation or positive selection alone cannot explain the differential sequence divergence of the two AGPase subunits, since the class with the low ω value still had a value of ω that was 2.6-fold greater than that estimated for the small subunit. These results are consistent with the conclusion that the small subunit is under more purifying selection than is the large subunit.

To directly determine whether the different rates of sequence divergence for the two subunits (and the correlated differences in the degree of sequence conservation) reflect differences in the relative impact of changes in the amino acids of the two subunits for enzyme activity, as suggested by other investigators, I used the mutagenesis approach described below.

Probability that a Non-Synonymous Mutation in *Sh2* or *Bt2* Affects Enzyme Activity

Two *Sh2* and two *Bt2* libraries were created by error-prone PCR. The resulting *sh2* and *bt2* clones were expressed in *E.coli* strain AC70R1-504 and 96 colonies from each of the four libraries were chosen at random. Colonies were rated as functional or non-functional by the

formation of brown-staining glycogen following exposure to iodine vapors (Figure 2-5). I expect that mutations that modify catalytic and regulatory properties, enzyme stability and enzyme assembly would affect enzyme activity. DNA sequencing revealed the nature and position of nucleotide changes. Mutations that result in loss of activity are shown in Table 2-1. Mutants listed are those that come from non-functional clones containing only a single missense mutation. Non-functional clones with multiple missense mutations were excluded from this table because the causal mutation could not be determined.

The distribution of all nucleotide substitutions within *Sh2* and within *Bt2* is uniform for all libraries (Figure 2-6). This is further supported by the fact that the numbers of mutants with different number of non-synonymous substitutions within each library follow a Poisson distribution ($0.889 < r^2 < 0.967$) (Figure 2-7). The percentage of conservative missense mutations (Material and Methods) was $54.8 \pm 6.9\%$ and $51.3 \pm 6.6\%$ (Mean \pm 2xSE) in the *Bt2* and *Sh2* libraries respectively. These percentages are not significantly different indicating that the comparison of the robustness of *Sh2* and *Bt2* to missense mutations is unlikely to be biased by introduction of more conservative changes in one of the two subunits. Clones containing indels, stop codons or no non-synonymous mutations were excluded from further analysis.

The probability that a non-synonymous mutation abolishes gene function was estimated by the formula recently published by Guo et al., (2004):

$$S = f_1(100-X_f)^1 + f_2(100-X_f)^2 + \dots + f_n(100-X_f)^n$$

S: fraction of functional clones

f_n : fraction of clones having n non-synonymous mutations

X_f : probability that a non-synonymous mutation in *Sh2* or *Bt2* reduces AGPase activity leading to no obvious production of glycogen

Results from the two *Sh2* and *Bt2* libraries are summarized in Table 2-2. The X_f for *Bt2* is $34.02 \pm 0.82\%$ and $33.36 \pm 2.27\%$ for *Sh2*. These values are not statistically different. I conclude

that *Sh2* and *Bt2* show little to no difference in robustness to non-synonymous mutations with respect to AGPase activity expressed in *E. coli*.

In these analyses, I assumed that multiple non-synonymous mutations in the same gene act independently. To test the validity of this assumption, I compared the fractions of clones having 1, 2, 3 ... non-synonymous mutations not altering AGPase function to the expected fractions if multiple intragenic mutations act independently (Table 2-3). Chi-square goodness-of-fit test indicates that the observed fractions are not significantly different from the expected fractions. This indicates that the assumption is valid for *Sh2* and *Bt2* for the mutational loads of this experiment.

Glycogen Quantitation

Since iodine staining is largely a qualitative method, the glycogen levels of *Sh2* and *Bt2* colonies of the second library were measured by a phenol reaction. This allowed us to determine the relationship between iodine staining and glycogen content. Glycogen contents were compared to the level produced by wildtype *Sh2* and *Bt2*. Iodine-staining is a reliable method for classifying clones for AGPase function (Figure 2-8). Clones rated as non-functional by iodine-staining produced less than 4% wildtype glycogen levels. There was little overlap in the glycogen content values among functional and non-functional clones as judged by iodine-staining. The small overlap was a result of slight inaccuracies in either iodine-staining or glycogen quantification. Only three ambiguous clones, producing around 3-5% wildtype glycogen levels, were identified. The calculated X_f was not significantly affected whether they were classified as functional or non-functional.

I asked if a relationship exists between the presence of synonymous mutations and glycogen production. Pearson Correlation Coefficient r is 0.039 for *Bt2* and 0.043 for *Sh2* ($p > 0.05$ for both) indicating little to no correlation between the number of synonymous

mutations per clone and glycogen production. This result indicates that synonymous mutations are unlikely to affect the comparison of tolerance to non-synonymous mutations between *Bt2* and *Sh2* in this experiment.

Glycogen quantitation also provides the opportunity to change the stringency of classification of functionality. Changing the stringency of classification from 0% glycogen to 99% of wildtype glycogen yields almost identical X_f values for *Sh2* and *Bt2* (Table 2-4). This result shows that stringency of classification does not affect the conclusion that *Sh2* and *Bt2* are equally prone to the reduction or elimination of function after accumulating non-synonymous mutations.

Estimating the Ratio of Non-Synonymous to Synonymous Mutations in Functional Clones

As an independent estimate of the robustness of *Sh2* and *Bt2* to non-synonymous mutations affecting AGPase activity, I isolated functional clones from heavily mutagenized libraries and measured the ratio of non-synonymous to synonymous mutations. In our previous studies described above, the ratios of non-synonymous to synonymous mutations of randomly-selected *Sh2* (2.936 ± 0.1037) and of randomly-selected *Bt2* (3.0585 ± 0.063) clones are not significantly different. Hence, synonymous mutations in functional clones can be used as a comparison of mutational load between *Sh2* and *Bt2* in this experiment.

Accordingly, two additional *Sh2* and two additional *Bt2* heavily mutagenized libraries were created by error-prone PCR. Resulting clones were expressed in *E. coli* AC70R1-504 and 48 functional clones from each of the four libraries were sequenced. The number of synonymous mutations per 1000 nucleotides of *Sh2* and *Bt2* cDNA is 0.951 ± 0.107 and 0.989 ± 0.204 respectively (Table 2-5). This indicates that the average mutational frequency is the same for both genes.

The ratio of non-synonymous to synonymous mutations is 1.188 ± 0.036 for *Sh2* and 1.176 ± 0.006 for *Bt2* (Table 2-5). These are not significantly different. This indicates that *Sh2* and *Bt2* are equally robust to non-synonymous mutations with respect to AGPase activity.

Additionally, chi-square goodness of fit tests indicates that the distribution of functional *Bt2* clones with different numbers (e.g., 0, 1, 2, 3...) of non-synonymous mutations was not significantly different from the distribution of functional clones found in *Sh2* (Table 2-6). This result is in agreement with the conclusion that *Sh2* and *Bt2* do not differ in their ability to tolerate non-synonymous mutations.

To summarize, two independent tests of protein robustness of the SH2 and the BT2 protein were performed. Both tests lead to the conclusion that the genes are equally predisposed to non/less-functional non-synonymous mutations at least when expressed in one environment (*E.coli*) and in the presence of a single functional complementary subunit.

Proteins Interacting with BT2

I tested the hypothesis that relative to the large subunit the small subunit is more conserved because it has additional roles that are mediated through protein-protein interactions. To investigate this possibility, I searched for proteins interacting with BT2 in maize endosperm. Formaldehyde was used to cross-link associated proteins in maize endosperm cells using methodology described elsewhere (Hall and Struhl, 2002; Rohila et al., 2004). Crosslinking with formaldehyde traps the BT2 protein into a complex of approximately 250 kDa (Figure 2-9A). Purification of BT2 from maize endosperm by a monoclonal BT2 antibody column shows that SH2 is the major protein that interacts with BT2 (Figure 2-9B) as verified by Western blot using a polyclonal SH2 antibody (data not shown). The extra proteins that are co-purified with BT2 either interacted directly with BT2 or AGPase or they were cross-reaction products with the monoclonal BT2 antibody column (Figure 2-9B). To distinguish among these possibilities, the

same monoclonal BT2 antibody column was used to purify cross-linked proteins from the developing endosperm of *bt2-7410*, a *bt2* knock-out mutant that has undetectable amounts of BT2 (Giroux and Hannah, 1994). Six of the eight extra proteins purified from wildtype were also purified from *bt2-7410* (Figure 2-9B). This indicates that these extra proteins are a cross-reaction product with the BT2 antibody column rather than BT2 or AGPase interacting proteins. The remaining two extra proteins purified from wildtype, but not *bt2-7410*, are SH2/BT2 polymers since they hybridize to both SH2 and BT2 antibodies (data not shown). These results show that SH2 is the only protein interacting with BT2 in endosperm. They also suggest that the AGPase small subunit is unlikely to have additional functions mediated by protein-protein interactions.

Tissue Specificity and Gene Number as a Cause of the Greater Evolutionary Constraints on AGPase Small Subunit

The timing of gene duplications giving rise to tissue-specific isoforms differed dramatically for the two subunits (Hannah et al., 2001 and above). At least one duplication of the large subunit occurred before the monocot-eudicot separation, while small subunit duplications occurred after this divergence. The original small subunit duplications appear to have occurred independently in monocots and eudicots and after their evolutionary separation.

The number of successful duplication events also appears to differ for the two subunits. Perusal of the rice, *Populus*, and *Arabidopsis* genomes suggest that the large subunit genes underwent more successful duplications than have the small subunit genes, a result consistent with our reconciled tree analysis (Figure 2-4A,B). *Arabidopsis* genome contains a single functional gene for the AGPase small subunit but four genes for the large subunit (Crevillen et al., 2003; 2005); rice contains two small subunit genes and four large subunit genes (Akihiro et al., 2005; Ohdan et al., 2005), while *Populus* apparently has one small and six large subunit genes (Tuscan et al., 2006). Three large subunit and one small subunit genes have been isolated

from tomato and from potato (La Cognata et al., 1995; Park and Chung, 1998). The major small subunits in barley endosperm and leaves (Thorbjørnsen et al., 1996; Johnson et al., 2003; Rosti et al., 2006) are encoded by a single gene that is alternatively spliced, while the interacting large subunits are encoded by different genes. Expression patterns and functional characterization of large and small subunits from tomato, potato, barley, rice and *Arabidopsis* indicate that large subunits are more tissue specific than are small subunits (La Cognata et al., 1995; Park and Chung, 1998; Akihiro et al., 2005; Crevillen et al., 2005; Ohdan et al., 2005; Rosti et al. 2006). Taken as a whole, these results strongly suggest that the large subunit genes, in comparison to small subunit genes, underwent more successful duplications that were then followed in some cases by subfunctionalization.

There is ample evidence that broadly-expressed genes are more conserved than are tissue specific genes in plants and mammals (Duret and Mouchiroud, 2000; Akashi, 2001; Wright et al., 2004). Broadly-expressed genes, like the small subunit AGPase genes, may have additional constraints because they have a greater impact on fitness and/or because they must satisfy multiple cellular and subcellular environments (Akashi, 2001).

As a result of broad expression, the small AGPase subunit must form an enzyme complex with multiple large subunits in multiple cellular environments. This is not the case for the large subunits. If this hypothesis is true, the effect of small subunit non-synonymous mutations on AGPase activity should depend on the identity of the large subunit. Mutations in the small subunit that are tolerated in a complex with one large subunit may not be tolerated in a complex with a different large subunit.

To test one aspect of this hypothesis, I randomly selected 20 *Bt2* variants that function with the SH2 protein and expressed them in *E. coli* with the wildtype maize embryo large subunit

(*Agplemzm*). Resulting glycogen was measured and compared to the glycogen produced by cells expressing wildtype *Bt2* with *Sh2* and wildtype *Bt2* with *Agplemzm*, respectively. Cells expressing wildtype *Bt2* with *Sh2* and wildtype *Bt2* with *Agplemzm* produce approximately equal amounts of glycogen (data not shown). DNA sequencing indicated there were 7 unique *Bt2* clones out of 20 excluding clones without any non-synonymous mutations (Table 2-7). Approximately 42.8% of *Bt2* mutants result in significantly different amounts of glycogen depending on the large subunit partner (Figure 2-10). These results indicate that the effect of some amino acid changes in the small subunit on AGPase activity depends on the interacting large subunit.

To further test the hypothesis that the small subunit is more constrained because of broad tissue expression, I perused the evolutionary relationships of AGPase subunits from the sequenced unicellular green algae, *Chlamydomonas reinhardtii*, *Ostreococcus tauri* and *Ostreococcus lucimarinus*. There is one major large and one small AGPase subunit in *C. reinhardtii* (Zabawinski et al., 2001). This is most likely the case in *O. tauri* (Ral et al., 2004) and *O. lucimarinus* since a third gene homologous to an AGPase subunit in these genomes shows very low identity (~31%) to the large and small subunit from the respective organisms. Additionally, alignment of the divergent *Ostreococcus* AGPase homologs to AGPase subunits from plants, algae and bacteria shows that residues conserved in all AGPase subunits are not conserved in these extra subunits (data not shown).

In support of our hypothesis that the greater conservation reflects the existence of multiple large subunit genes and/or broad patterns of expression, the large and the small subunits in these organisms show similar levels of sequence conservation (Table 2-8). Indeed, the small subunit in these unicellular algae evolved more rapidly than did the small subunit in plants and at the same

rate as the large subunit of the unicellular algae (Figure 2-11). It should be noted that the rate of evolution of nuclear genes in green algae and in plants is comparable based on small and large subunit ribosomal RNA (Van de Peer et al., 1996; Ben Ali et al., 2001) and proteins such as the elongation factor 1 α , actin, α -tubulin and β -tubulin (Baldauf et al., 2000). Hence, the most parsimonious explanation for the faster evolution of the small subunit in algae is that it has fewer constraints in algae than it does in plants.

It is also possible that different tissues require different levels of AGPase function. This may place different evolutionary constraints on the large subunit depending on the tissue(s) of expression. Indeed, expression of each of various large subunits with the common small subunit in *Arabidopsis* yields AGPases with different kinetic and allosteric properties (Crevillen et al., 2003). This suggests that different tissues may require different levels of AGPase function. Interestingly, large subunits expressed primarily in leaves have lower ω (Figure 2-3A) than do large subunits expressed in monocot endosperm (Figure 2-3B). Further support for more evolutionary pressure on subsets of AGPase genes comes from analysis of a recent duplication of a small subunit maize gene. While rice, barley and probably wheat have one small subunit gene expressed in both leaves and endosperm, maize has one small subunit gene specific for endosperm and one specific for leaves (Prioul et al., 1994; Rosti and Denyer, 2007). These two maize isoforms were created by recent gene duplication. Following duplication, the maize leaf isoform (*Zea mays 2*) evolved slightly faster (LRT; $p=0.1$) than before duplication while the maize endosperm isoform (*Bt2*) evolved significantly faster (LRT; $p=0.007$) (Figure 2-4A, Figure 2-12). This suggests that tissue identity influences the rate of evolution of both the large and the small subunit of AGPase. Expression in multiple tissues then would compound the various selection pressures of the various tissues.

Discussion

The focus of these investigations was an understanding of the cause of the differential rate of sequence conservation exhibited by the two subunits of angiosperm AGPase. Substantially greater sequence divergence is found among large subunits than is found among the small subunits. Here, I definitively show that the different rate of sequence divergence distinguishing the large and the small subunit of AGPase is due to different evolutionary constraints rather than the different paths the genes have taken during evolution. Estimation of both ω and the rates of amino acid substitution indicate that the angiosperm small subunit is under more evolutionary pressure than is the angiosperm large subunit. This result is especially compelling since estimates of ω and the absolute rate of amino acid substitution provide congruent results, indicating that factors that confound each approach (saturation of K_S for ω ; inaccurate divergence time estimates for absolute rates) do not confound our results.

One explanation in the literature for the observed differences in the degree of sequence conservation is the conjecture that the two subunits play vastly different roles in enzyme catalysis. It has been proposed that the small subunit is catalytic and, accordingly, more constrained in amino acid substitutions compared to the largely regulatory large subunit. This model predicts that random non-synonymous mutations in the small subunit will have a greater probability of disrupting AGPase activity than random non-synonymous mutations in the large subunit.

A bacterial expression system (Iglesias et al., 1993) allowed us to randomly mutagenize maize endosperm AGPase genes and score AGPase activity in a fast and efficient way. The *E. coli* system is ideal for studying plant AGPases for a number of reasons. First, compared to the enzyme from the host tissue, *E. coli*-expressed plant AGPases exhibit virtually identical kinetic and allosteric properties (Iglesias et al., 1993; Boehlein et al., 2005). Second, mutations affecting

both kinetic and allosteric properties of plant AGPases have been isolated in *E. coli*. A potato large subunit mutant (P52L) with reduced affinity for 3-PGA exhibited no staining in *E. coli* (Greene et al., 1996). A mosaic small subunit between potato tuber and maize endosperm (MP) having wildtype kinetic properties but reduced affinity for Pi showed elevated staining compared to wildtype maize endosperm AGPase in a complex with maize endosperm large subunit (Cross et al., 2004; Boehlein et al., 2005; Georgelis and Hannah, unpublished data). And, potato tuber AGPase mutants with altered kinetic properties, such as decreased affinity for substrates, have also been detected. A potato tuber large subunit mutant (M143) with decreased affinity for G-1-P had reduced staining in *E. coli* (Kavakli et al., 2001). A potato tuber small subunit mutant with decreased affinity for ATP (A106T) also showed reduced staining in *E. coli* (Laughlin et al., 1998). It should be noted that potato tuber and maize endosperm AGPases have identical affinities for substrates (Fu et al., 1998; Boehlein et al., 2005). This indicates that our bacterial system is sensitive enough to detect kinetic mutants in maize endosperm AGPase. Finally, the system allows for the detection of mutants with increased or decreased enzyme stability. A heat-stable *Sh2* variant, *HS33*, was isolated by intense iodine-staining compared to wildtype *Sh2* (Greene and Hannah, 1998a) following *E. coli* growth at elevated temperatures. This variant enhances heat stability when expressed in the wheat endosperm (Meyer et al., 2004). Hence, the *E. coli* expression system is a sensitive and valid means for screening AGPase activity and comparing the relative robustness of the large and small AGPase subunits to activity altering mutations.

Two independent mutagenesis studies were used to test the hypothesis that the large subunit is more robust to mutations than is the small subunit in terms of enzyme activity. In the first approach, random non-synonymous mutations were introduced into *Sh2* and *Bt2*. Resulting

clones were picked at random and gene function and gene sequence were determined to calculate the probability that a non-synonymous mutation would abolish or alter gene function (X_f). In the second approach, functional clones were chosen from new, heavily mutagenized libraries and the frequency of non-synonymous and synonymous mutations was determined.

The first series of experiments revealed that random, non-synonymous mutations in either subunit have the same probability of altering AGPase activity. Approximately 1/3 of the missense mutation reduced glycogen content to approximately 4% wildtype levels. While the frequency of mutants having less drastic effects on glycogen content occurred more frequently, their frequency was identical for both genes.

In our second series of experiments, only functional clones as judged by iodine staining were selected and the frequencies of non-synonymous and synonymous mutations were determined. The frequency of synonymous mutations was used to measure mutation frequency. If random, non-synonymous mutations were tolerated at a higher frequency in *Sh2*, then I would expect a higher non-synonymous/synonymous ratio for *Sh2* in comparison to the ratio from *Bt2*. In contrast, I observed the same ratio for both genes. Furthermore, the numbers of non-synonymous mutations in equally mutagenized *Sh2* and *Bt2* libraries show identical distributions, further corroborating the same conclusion.

The conclusion I draw from both sets of experiments is that the genes are equally robust to mutation. The probability that a non-synonymous mutation affects AGPase activity is the same for both the large subunit and the small subunit of AGPase when expressed in one common environment with one complementary partner. These results clearly indicate that different roles of the large and small subunit in AGPase activity, if they exist, cannot account for the increased evolutionary constraints placed on the plant small subunit. This is also supported by the fact that

in green unicellular algae the large subunit and small subunit evolve at similar rates. It has been proposed that the small subunit in green alga *C.reinhardtii* is catalytic and the large is regulatory (Zabawinski et al., 2001). If true, the proposed difference in roles does not result in more conservation in the small subunit compared to the large subunit.

The proposition that in angiosperms, the small subunit is catalytic and the large subunit is somehow involved in regulation is based on three observations. The first is that mutations of some small subunit residues presumably involved in catalysis had greater effect on activity compared to mutants in cognate sites in the large subunit. (Fu et al., 1998; Frueauf et al., 2003). The second observation is that the small subunit has activity as a homotetramer while the large subunit did not (Ballicora et al., 1995). Finally, the third observation is that the small subunit is more conserved than is the large subunit.

The first observation has been challenged by biochemical studies (Hannah and Nelson, 1976; Cross et al., 2004; Cross et al., 2005; Hwang et al., 2006; 2007). These studies show that mutations in the large subunit affect catalytic properties as well as allosteric ones. With regard to the second observation, it should be noted that activity arising from a homotetramer composed of only the large subunit can be seen in the data of Burger et al. 2003 and Iglesias et al. 1993. While the presence of activity arising from a large subunit homotetramer is not universally accepted, the lack of activity of a large subunit homotetramer does not show that the large subunit is not catalytic in a heterotetramer.

The third observation, that the small subunit is more evolutionarily conserved than is the large subunit because the small subunit participates in catalysis whereas the large subunit simply modulates regulation, predicts that I should have found more missense mutations in the small subunit (compared to the large subunit) affecting activity. In contrast, in two separate and

independent series of experiments, I found that both proteins were equally predisposed to activity altering missense mutations. While our experiments do not completely dismiss the possibility that the large and the small subunit have different roles in enzyme activity, these different roles in enzyme activity, if they exist, cannot explain the higher evolutionary constraints placed on the small subunit. Taking into account the biochemical studies, mentioned above, that show that both subunits are important for catalytic and allosteric properties I believe the simplistic separation of subunit roles to catalytic and allosteric modules is not longer tenable. Such a separation is not sufficient to explain the accumulating data.

Because both AGPase subunits exhibit equal constraints on amino acid sequence when both subunits are expressed in one common environment with a common complementary subunit, other explanations must be explored to account for the observed differences in amino acid substitution rates.

I considered the possibility that the small subunit has biological roles not shared with the large subunit and separate from its function as an AGPase subunit. Here I show that SH2 was the only protein found to interact with the BT2 protein in the maize endosperm. Thus, there is no evidence from this study or previous studies that suggest BT2 has functions separate from interacting with SH2 to form an active AGPase.

Several lines of evidence reviewed above suggest that there are fewer copies of the genes encoding the small subunit than genes encoding the large subunit, and that small subunit genes have a broader expression profile compared to large subunit genes. Accordingly, evidence that broadly-expressed genes evolve at a slower rate than do tissue specific genes (Duret and Mouchiroud, 2000; Akashi, 2001; Wright et al., 2004) is relevant. Studies focused on specific gene families have shown that genes encoding broadly expressed protein isoforms evolve more

slowly than those with a narrow expression profile (Hastings, 1996; Schmidt et al., 1997). This may reflect a larger fitness effect for mutations that impair the function of genes expressed in many different tissues or it may reflect the existence of more constraints upon proteins expressed in multiple biochemical environments (Akashi, 2001). All of these studies are consistent with the small subunit having more evolutionary constraints because it must function in multiple cellular environments, polymerize with different large subunits and provide AGPase activity in multiple tissues.

Although the small subunit does not have interaction partners other than the large subunit, the existence of multiple large subunits that must interact with a single small subunit makes the correlation between numbers of interaction partners and conservation (Fraser et al., 2002) relevant. The relationship between conservation and numbers of interaction partners is complex (Jordan et al., 2003), making it difficult to predict *a priori* whether one would expect the need to interact with multiple paralogous large subunits to provide additional constraints. However, when I asked whether BT2 variants known to be functional with the SH2 protein were also functional when expressed with another wildtype large subunit isoform I found that 42.8% of randomly selected BT2 variants which interact with the wildtype SH2 protein to produce a functional AGPase interact differently when expressed with the maize embryo large subunit. This provides empirical evidence that large subunit-small subunit interactions differ based upon the large subunit involved, and indicates that the X_F estimate derived from *E. coli* did not consider all mutations that are relevant *in planta*. This reflects the fact that mutations that are deleterious for even one complex of the small subunit and a large subunit could reduce the fitness of the organism and be eliminated by natural selection.

The general hypothesis that tissue specificity and the need to interact with multiple large subunits constrains the evolution of the small subunit is supported by the fact that the small subunit evolved as rapidly as the large subunit in unicellular algae *C. reinhardtii*, *O. tauri* and *O. lucimarinus* and that the evolution of the small subunit was considerably faster than it was in angiosperms. Since there is a single cell type in these algae (although it can undergo developmental changes) and there is likely to be a single large subunit, one would predict that the evolutionary rates for the large and small subunits would be identical in these organisms, just as I observed. If different constraints in evolution of the small and the large subunit were due to different roles in AGPase catalysis I would expect to see different conservation between the small and the large subunit in algae. The fact that the small subunit in unicellular green algae evolved considerably faster than in multicellular angiosperms suggests that multiple environments including multiple large subunit partners in angiosperms constrain the evolution of the small subunit.

Finally, an additional reason that may allow the large subunit to evolve faster than the small subunit is partial redundancy. Even though genetic studies have shown that one large subunit (APL1) is of major importance in Arabidopsis leaves, the possibility that other large subunits contribute to AGPase activity in the same tissue is likely. Indeed, a knockout of the large subunit gene APL1 reduced leaf starch levels to only 40% of wildtype (Lin et al., 1988b) whereas a small subunit knockout resulted in starchless leaves (Lin et al., 1988a). Partial redundancy may then reduce the impact of the large subunit mutation on plant fitness and accelerate its evolution rate.

Overall, this study shows that the higher conservation in the small subunit than the large subunit of AGPase is due to greater evolutionary constraints on the small subunit. It also shows

that the large and the small subunit of AGPase are equally tolerant to non-synonymous mutations with respect to enzyme activity when expressed with a single partner in a common cellular environment. This implies that both subunits are equally important for the structure and function of AGPase. The high degree of sequence identity between the large and small subunits of AGPase indicates that their structures are likely to be virtually identical, and the estimate of X_f indicates that both subunits are equally sensitive to non-synonymous mutation when expressed in one common environment. While I cannot formally exclude the possibility that subtle (and as yet unknown) differences distinguish the mechanism AGPase catalysis of the maize enzyme in *E.coli* versus in the maize endosperm, I suggest that the difference in evolutionary rates of sequence divergence seen in angiosperms is caused by the different number of genes encoding the small and large subunits of AGPase. This, in turn, results in differences in their patterns of gene expression, differences in the number of interaction partners and differences in functional redundancy. The detailed analyses possible in focused studies like this one complement large-scale bioinformatic studies (Wolf et al., 2006) and suggest that detailed examination of other paralogs exhibiting different degrees of sequence conservation will prove useful.

Table 2-1. Amino acid changes not tolerated in BT2 and SH2.

BT2			SH2		
Position	From	To	Position	From	To
65*	R	P	102*	T	R
112*	A	P	127	F	S
188	A	V	172	Q	H
193*	A	T	298	M	L
222	K	I	412	H	Y
343*	P	Q	416*	G	E
447	A	D	474	I	S
			510	T	S

The *E.coli* colonies expressing these mutants along with the wt complementary AGPase subunit produce less than 3-4 % of glycogen compared to wildtype. * denotes that the original amino acid residue is absolutely conserved in the small or large subunit homologs

Table 2-2. Glycogen production from randomly-chosen *sh2* and *bt2* mutants expressed in *E.coli*.

<i>Bt2</i> (1 st library)			<i>Bt2</i> (2 nd library)		
# of non-synonymous mutations per clone	# of clones	# of positive clones	# of non-synonymous mutations per clone	# of clones	# of positive clones
1	13	10	1	12	9
2	12	4	2	14	7
3	8	1	3	7	1
4	3	2	4	4	0
5	1	0	5	2	0

$$X_{f_{Bt2(1)}} = 33.67 \%$$

$$X_{f_{Bt2(2)}} = 34.35 \%$$

$$X_{f_{Bt2(Ave)}} = 34.02 \pm 0.82 \%$$

<i>Sh2</i> (1 st library)			<i>Sh2</i> (2 nd library)		
# of non-synonymous mutations per clone	# of clones	# of positive clones	# of non-synonymous mutations per clone	# of clones	# of positive clones
1	10	5	1	16	13
2	17	7	2	8	2
3	12	6	3	5	1
4	4	0	4	3	1
5	1	0	5	0	0

$$X_{f_{Sh2(1)}} = 34.75 \%$$

$$X_{f_{Sh2(2)}} = 31.45 \%$$

$$X_{f_{Sh2(Ave)}} = 33.36 \pm 2.27 \%$$

X_f : probability that a non-synonymous mutation abolishes AGPase activity. Positive clones exhibit some glycogen production as judged by iodine staining. Average X_f = Mean \pm 2 x standard error. Clones with insertions or deletions, stop codons or lacking non-synonymous mutations were excluded.

Table 2-3. Distribution of observed and expected positive clones if multiple intragenic mutations function independently.

<i>Sh2</i> (1 st and 2 nd library)				
# of non-synonymous mutations per clone	# of clones (positive and negative)	# of observed positive clones	# of expected positive clones	X ² goodness of fit test
1	26	17	17.38	p=0.844
2	25	9	11.17	
3	17	7	5.08	
4	7	1	1.4	
5	1	0	0.133	
<i>Bt2</i> (1 st and 2 nd library)				
# of non-synonymous mutations per clone	# of clones (positive and negative)	# of observed positive clones	# of expected positive clones	X ² goodness of fit test
1	25	18	16.37	p=0.798
2	26	11	11.15	
3	14	2	3.93	
4	8	2	1.472	
5	3	0	0.361	

Positive clones exhibit some glycogen production as judged by iodine staining.

Table 2-4. X_f of *Sh2* and *Bt2* using different stringencies to classify mutants as non-functional.

Threshold of functionality (% of wildtype)	$X_f(Bt2)$ (Mean \pm SE)	$X_f(Sh2)$ (Mean \pm SE)
0%	11.76 \pm 2.01%	12.90 \pm 0.65%
Staining upon iodine exposure (4 to 5%)		
15%	44.26 \pm 4.58%	49.97 \pm 7.71%
30%	58.56 \pm 3.21%	60.77 \pm 4.87%
60%	65.76 \pm 12.11%	72.72 \pm 7.27%
99%	79.55 \pm 8.15%	87.95 \pm 1.25%

Two samples (approximate sample size: 16-20 clones) from the second *Sh2* and *Bt2* libraries were used for calculating X_f . X_f : Probability that a non-synonymous mutation reduces AGPase activity.

Table 2-5. Ratio of non-synonymous to synonymous mutations in functional *Sh2* and *Bt2* clones.

	# of synonymous mutations per 1Kb of <i>Bt2</i>	# of synonymous mutations per 1Kb of <i>Sh2</i>	Student's t-test
3 rd library	1.091	1.004	p=0.768
4 th library	0.888	0.897	
Average \pm 2 x SE	0.989 \pm 0.102	0.951 \pm 0.053	
	# of non-synonymous/ # of synonymous mutations in <i>Bt2</i>	# of non-synonymous/ # of synonymous mutations in <i>Sh2</i>	Student's t-test
3 rd library	1.179	1.171	p=0.574
4 th library	1.173	1.207	
Average \pm 2 x SE	1.176 \pm 0.006	1.188 \pm 0.036	

Mutational load is expressed as number of synonymous mutations/kilobase. The ratio of non-synonymous to synonymous mutations in functional *Sh2* and *Bt2* clones is determined by sequencing two samples of 24 random functional clones from each of two heavily mutagenized libraries (labeled 3rd and 4th). Positive clones exhibit some glycogen production as judged by iodine staining.

Table 2-6. Distribution of positive clones with different numbers of non-synonymous mutations from the heavily mutagenized 3rd and 4th libraries.

# of non-synonymous mutations per clone	# of positive clones		X ² goodness of fit test
	<i>Bt2</i>	<i>Sh2</i>	
0	8	10	p=0.608
1	33	29	
2	23	23	
3	15	7	
4	4	7	
5	0	3	
Total	83	79	

Positive clones exhibit some glycogen production as judged by iodine staining. Chi-square goodness of fit test was used to compare the distribution of positive clones with different numbers of non-synonymous mutations between *Sh2* and *Bt2*.

Table 2-7. Amino acid changes of BT2 variants that were expressed in *E.coli* with SH2 and AGPLEMZM.

Mutant	# of non-synonymous substitutions	aa changes									
1	2	118	G-E	442	D-G						
2	2	27	D-G	72	L-V						
3	1	439	L-F								
4	5	5	L-M	6	R-P	177	E-V	203	T-I	299	D-E
5	1	241	D-E								
6	3	14	W-R	313	N-S	402	K-Q				
7	3	43	D-V	359	E-G	374	L-I				

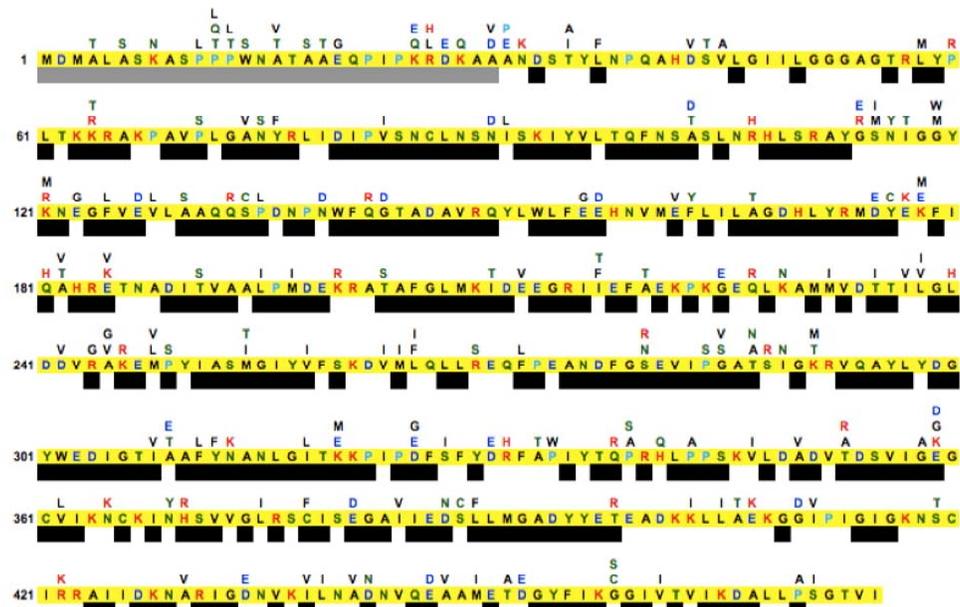
The effects of some of these BT2 variants on AGPase activity depended on the identity of the complementary large subunit as judged by *E.coli* glycogen production.

Table 2-8. Comparison of AGPase small and large subunits from unicellular green algae.

Pairwise comparisons	Percent Identity and Similarity	
	Small subunits	Large subunits
<i>O.lucimarinus-O.tauri</i>	84.3-91.1	91.0-95.8
<i>O.lucimarinus-C.reinhardtii</i>	64.3-79.3	63.1-78.1
<i>C.reinhardtii-O.tauri</i>	63.4-79.8	64.0-78.1

Alignment of amino acid sequences was done by BIOEDIT using BLOSUM62 matrix. The hyper-variable amino terminus was excluded.

BT2



SH2

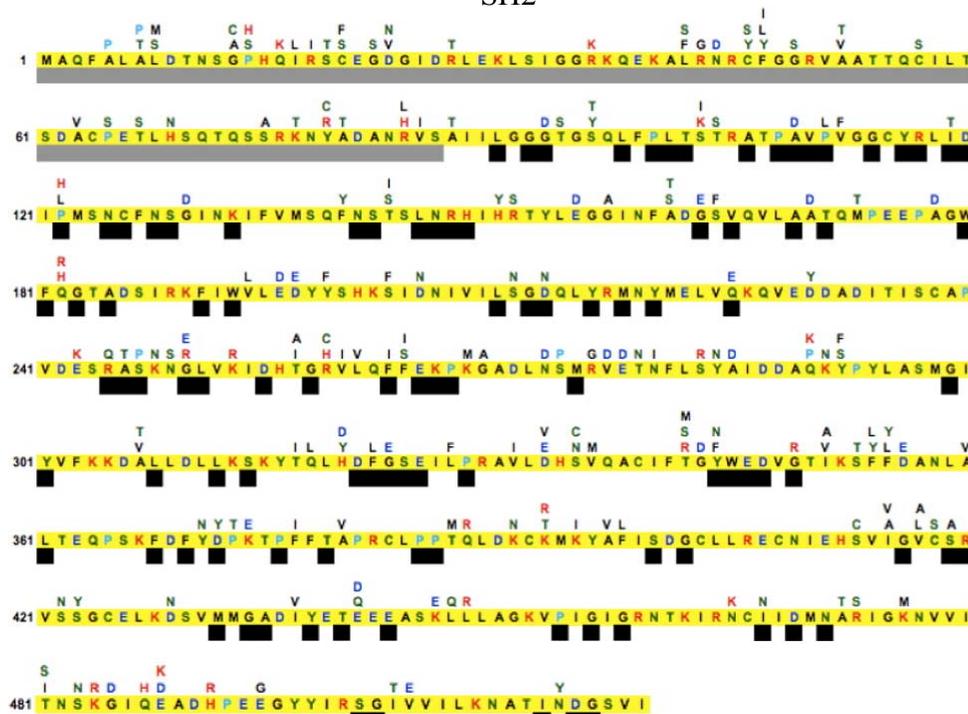
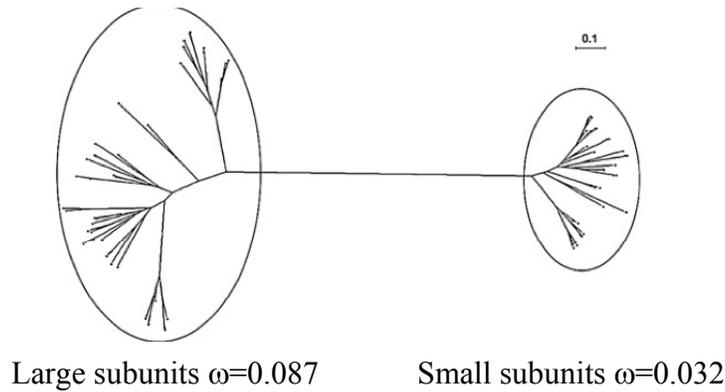
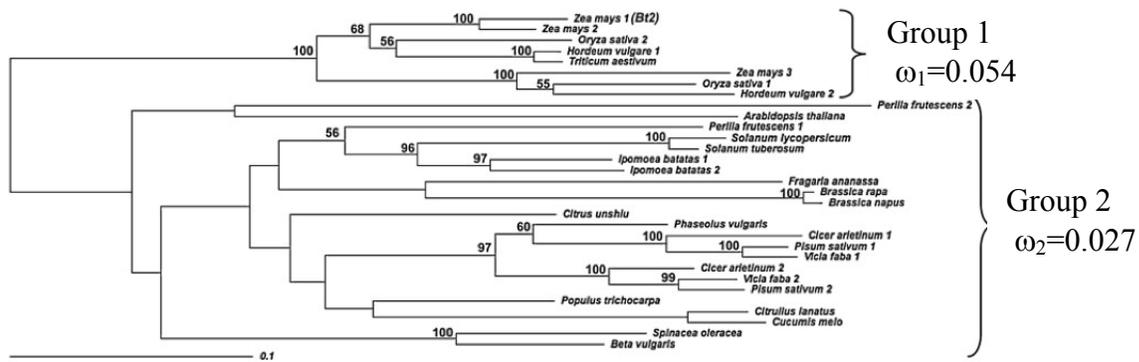


Figure 2-1. Amino acid conservation patterns of BT2 and SH2. BT2 and SH2 were aligned with 30 and 42 homologs respectively. Tolerated amino acid changes created by error-prone PCR in this experiment are shown above the amino acid sequence of BT2 and SH2. Black: non-polar, Green: uncharged polar, Red: basic, Blue: acidic, Turquoise: proline, Gray box: Poor alignment with homologous sequences, Black box: Absolutely conserved residues among homologous sequences.

A Gene tree of the angiosperm large and small AGPase subunits



B Gene tree of the angiosperm small AGPase subunits



C Gene tree of the angiosperm large AGPase subunits

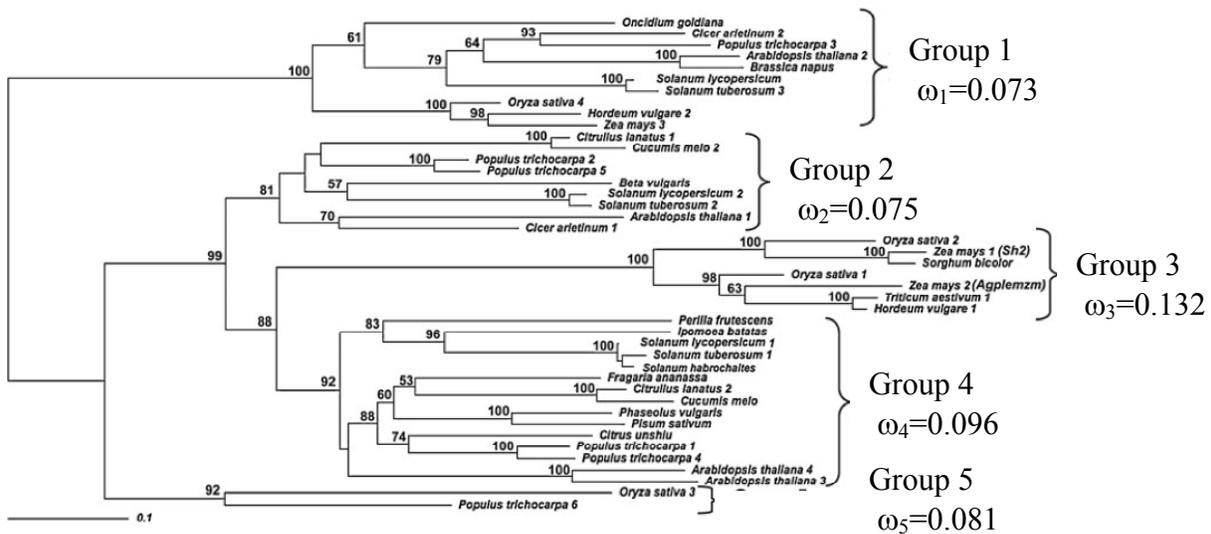


Figure 2-2. Estimates of AGPase Subunit Gene Trees and ω Values. The overall estimate of ω for each subunit and groups within each subunit were calculated by CODEML. Gene trees, used to estimate ω , were constructed using GARLI. Bootstrap values 50% are shown, and the bars indicate the number of nucleotide substitutions per site. A) Small and Large subunit gene tree. B) Small subunit gene tree. C) Large subunit gene tree.

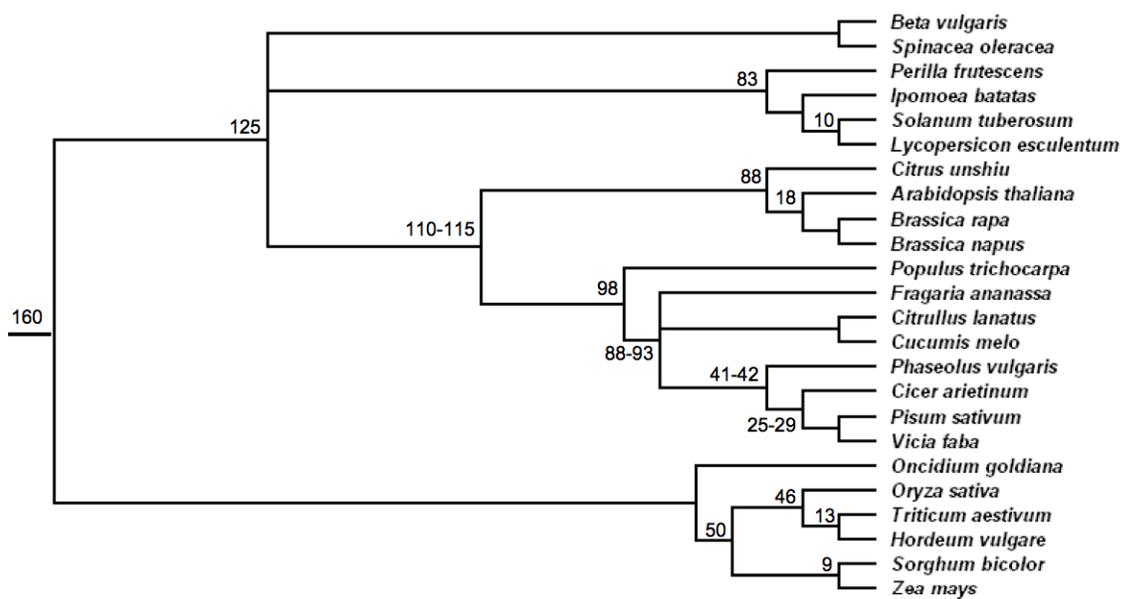


Figure 2-3. Species tree. Times of divergence are indicated in million of years (MY) at nodes.

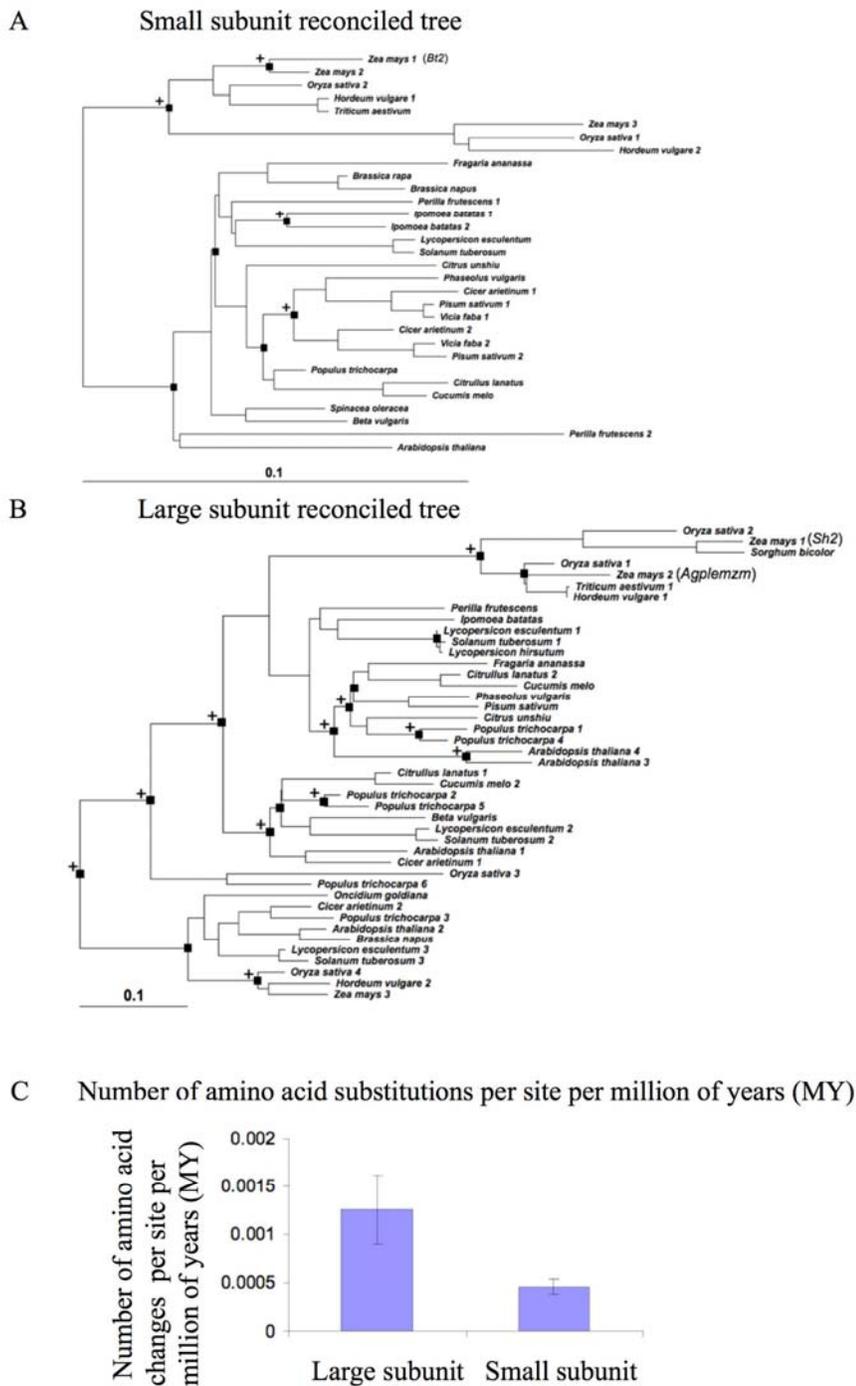


Figure 2-4. Estimates of the number of amino acid substitutions per site per million years (MY) for AGPase subunits. The topologies of phylogenetic trees A) and B) correspond to the small and large subunit gene trees, respectively. Branch lengths reflect the number of amino acid substitutions per site, estimated using AAML. Black boxes at branches indicate gene duplication events established by reconciling the gene trees with a species tree (Figure 2-3) using GENETREE. The numbers of amino acid substitutions

per site are indicated using the scale bars. C) Overall numbers of amino acid substitutions per site per MY for the small and the large subunit, calculated from the most recent speciation event for any set of taxa that has an associated date. The error bars correspond to 2 x standard error. + denotes well-supported duplication events (duplication events supported by gene tree nodes with bootstrap value >70%).



Figure 2-5. AGPase mutants exhibit different phenotypes following iodine staining. Iodine staining of *E. coli* expressing wildtype *Sh2* and mutant *Bt2* clones. Brown staining indicates AGPase activity.

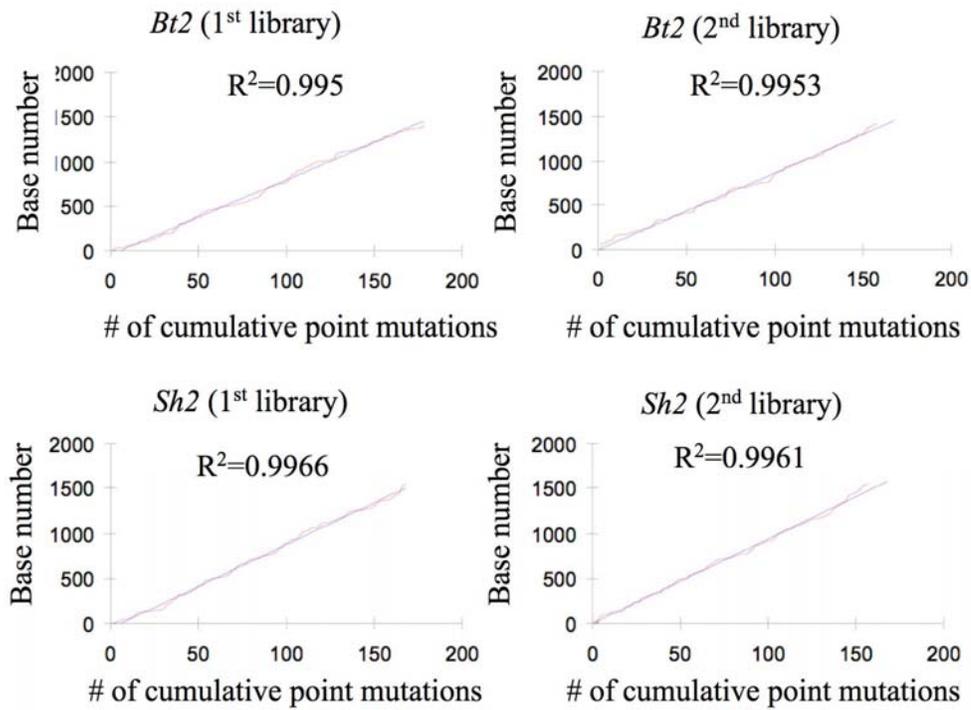


Figure 2-6. Placement of point mutations created by error-prone PCR in *Sh2* and *Bt2* clones. Blue curve: expected distribution of random mutations. Red curve: observed distribution of mutations

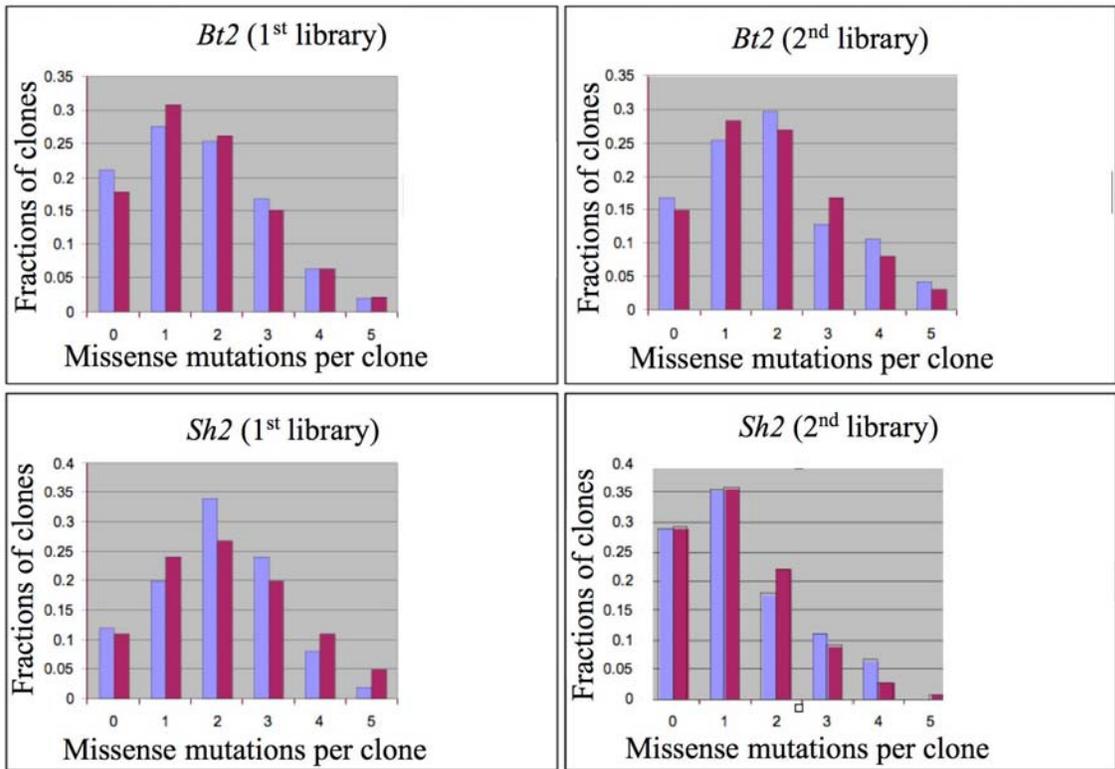


Figure 2-7. The distribution of *Bt2* and *Sh2* clones with different numbers of missense mutations. Blue: Actual distribution. Magenta: Fitted Poisson distribution.

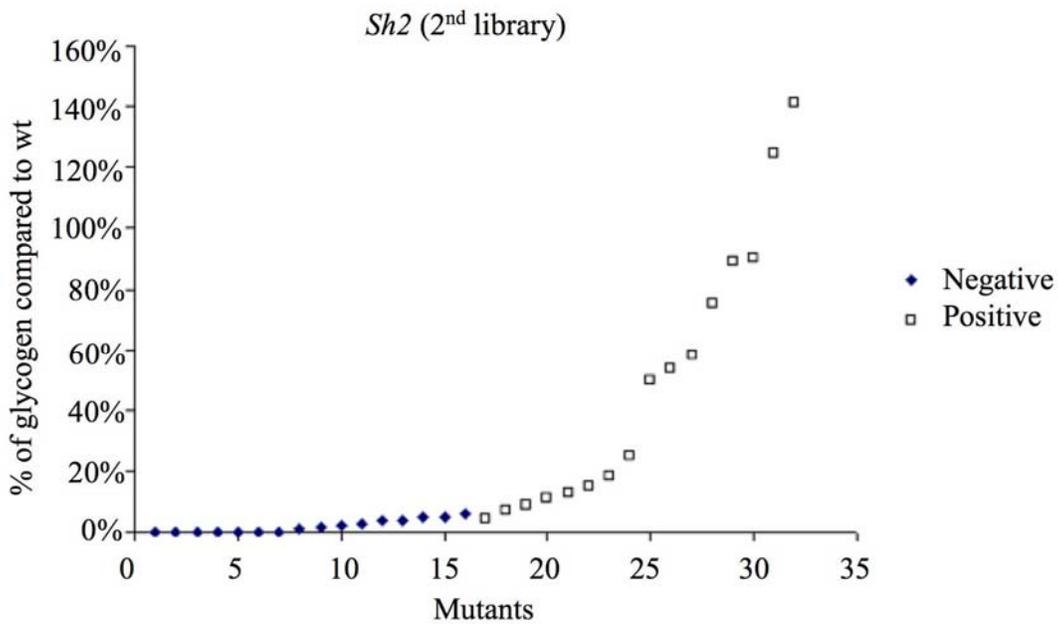
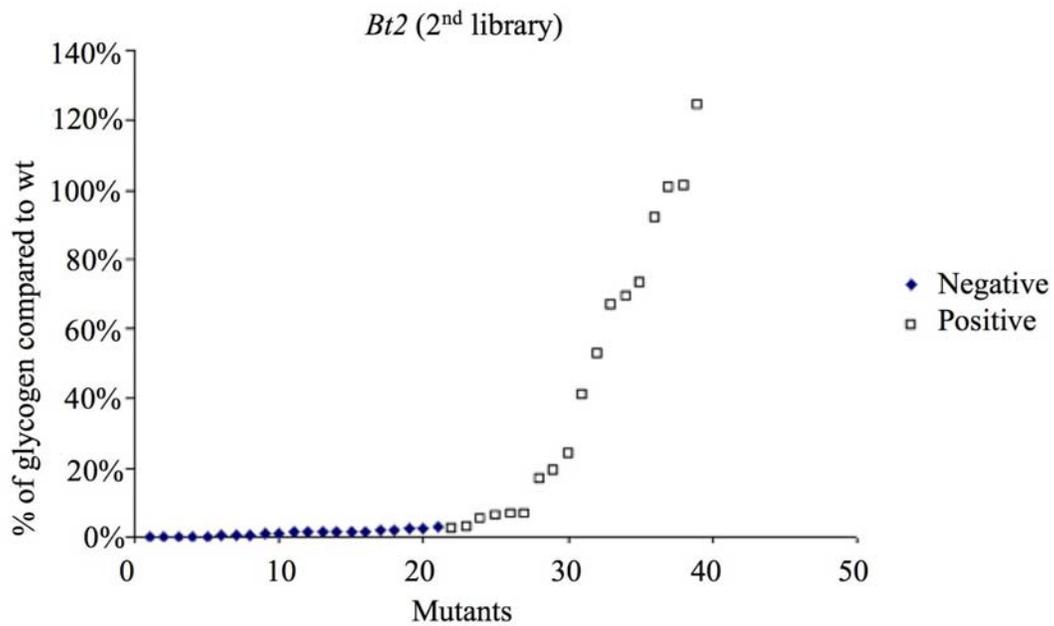


Figure 2-8. Glycogen contents of *sh2* and *bt2* mutants compared to wildtype AGPase. Positive clones exhibit some glycogen production as judged by iodine staining. Negative clones do not exhibit detectable amounts of glycogen as judged by iodine staining. Negative clones produced 4 - 5% or less glycogen compared to wildtype. Partially staining clones, classified as positive in this experiment, produced 5 - 40% glycogen compared to wildtype.

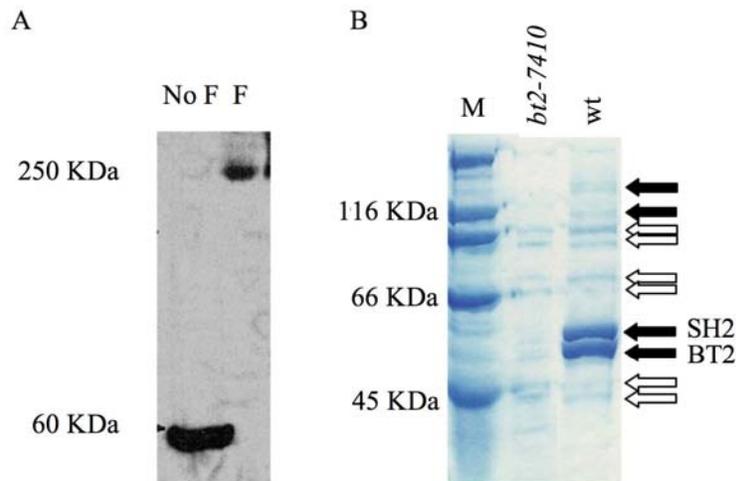


Figure 2-9. Purification of BT2 using a monoclonal antibody column. A) Formaldehyde cross-linking of BT2. Total proteins were extracted from maize endosperm with or without prior formaldehyde treatment. Extracted proteins were probed with a monoclonal BT2 antibody following SDS-PAGE. B) SDS-PAGE of proteins from *bt2-7410* (a *bt2* loss-of-protein mutant) and wildtype maize endosperm bound to a monoclonal BT2 antibody column. Formaldehyde cross-linking was performed on maize endosperm before protein extraction. Formaldehyde cross-linking was reversed by heating (95°C for 20 min) after column purification and before loading the purified proteins on the gel. M: Marker F: formaldehyde treatment before protein extraction. White arrows point to proteins purified from both *wildtype* and *bt2-7410*. Black arrows point to proteins purified only from wildtype.

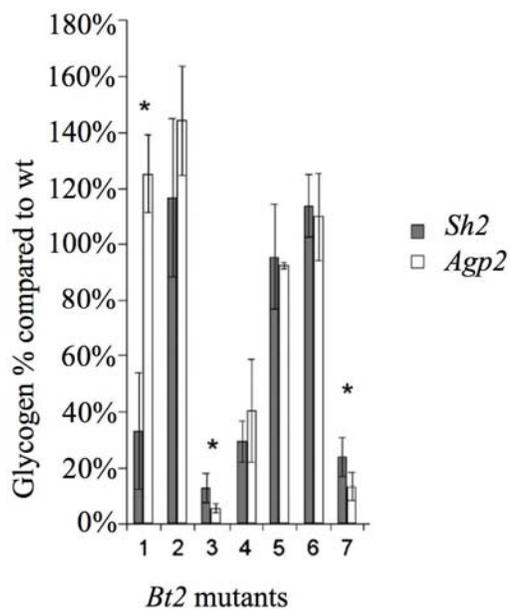


Figure 2-10. Comparison of glycogen produced by *Bt2* mutants expressed with wildtype *Sh2* and wildtype *Agp2* in *E. coli*. Glycogen produced by *E. coli* cells expressing *Bt2* mutants with wildtype *Sh2* or wildtype *Agp2*. Values are the percentage of glycogen produced compared to glycogen produced by cells expressing wildtype *Bt2* with wildtype *Sh2* or *Agp2*. Three replicates were used for determining glycogen production for each mutant. Error bars: 2 x standard error.

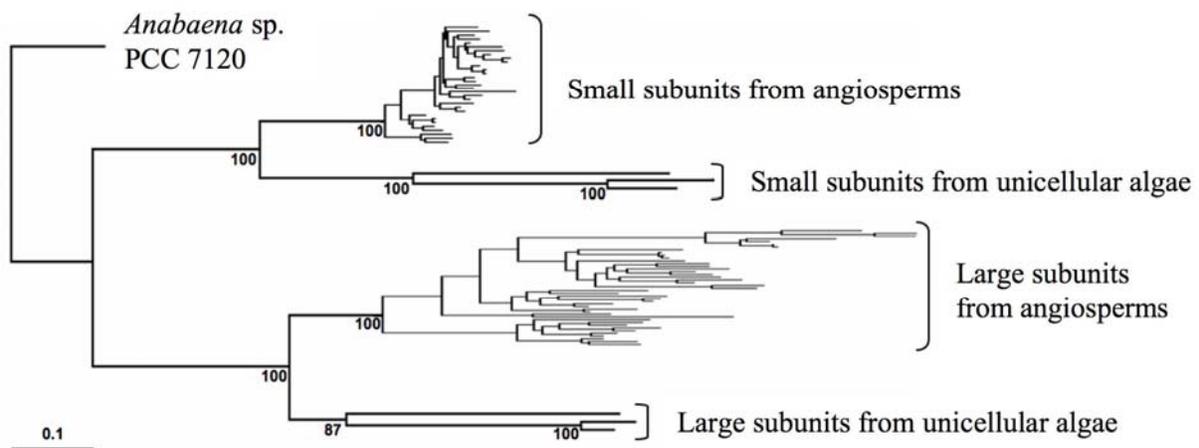


Figure 2-11. Phylogenetic tree of AGPase isoforms from plants and unicellular algae. The topology of the tree was determined by ML using aligned amino acid sequences analyzed by PhyML. The length of the branches reflects numbers of amino acid substitutions per site estimated using the Dayhoff+ Γ model of sequence evolution. ML bootstrap values are indicated below branches and the scale bar shows the number of amino acid substitutions per site.

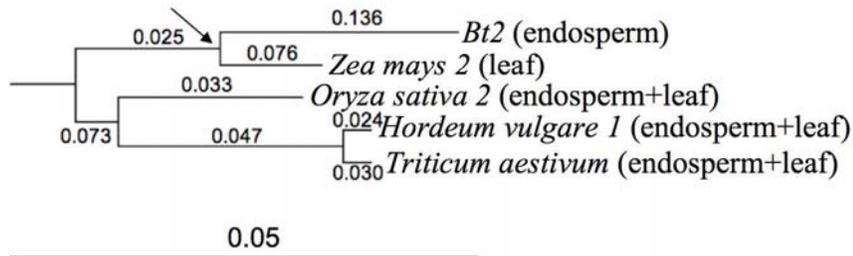


Figure 2-12. Differential rates of evolution between *Bt2* and *Zea mays 2* after gene duplication. The topology of the phylogenetic tree corresponds to the small subunit gene tree. Branch lengths reflect the number of amino acid substitutions per site, estimated using AAML. The value above or below each branch correspond to the ω value of the branch estimated by CODEML. The arrow points to the duplication between *Bt2* and *Zea mays 2*. Tissue(s) of major expression of each isoform are indicated in parentheses. The numbers of amino acid substitutions per site are indicated using the scale bars.

CHAPTER 3
ISOLATION OF AGRONOMICALLY IMPORTANT ADP-GLUCOSE
PYROPHOSPHORYLASE VARIANTS

Introduction

Importance of Allosteric Properties of AGPase for Plants and Humans

ADP-glucose pyrophosphorylase (AGPase) catalyzes the first committed step in starch (plants) and glycogen (bacteria) synthesis. It involves the conversion of glucose-1-P (G-1-P) and ATP to ADP-glucose and pyrophosphate (PPi). AGPase is a heterotetramer in plants consisting of two identical small and two identical large subunits. The large and the small subunits are encoded by *Shrunken-2* (*Sh2*) and *Brittle-2* (*Bt2*) respectively in maize endosperm. The importance of maize endosperm AGPase in starch synthesis has been shown by the kernel phenotype of mutants in either subunit of the enzyme. Indeed, such mutants result in shrunken kernels and a large reduction in endosperm starch content (Tsai and Nelson, 1966; Hannah and Nelson, 1976).

AGPase is allosterically regulated by small effector molecules that indicate the energy status of the cell. In bacteria, AGPase is activated by fructose 6-P, fructose 1,6-bisphosphate, or pyruvate, and deactivated by AMP (Ballicora et al, 2003). All these effectors are glycolytic intermediates. AGPase is activated by 3-PGA, the first carbon assimilatory product, and deactivated by inorganic phosphate (Pi) in cyanobacteria, green algae and angiosperms. The importance of AGPase regulation by 3-PGA and Pi is obvious in source leaves, where sucrose and starch synthesis need to be coordinated to meet the needs of the whole plant (Hannah, 2000). However, in sink tissues, the physiological importance of such regulation is not clear.

Evidence of the rate limiting role of AGPase and the importance of its allosteric properties came from an *Escherichia coli* (*E.coli*) AGPase variant, *glgC-16*, that was relatively insensitive to inhibitors AMP and Pi (Preiss and Romero, 1989). *E.coli* cells expressing *glgC-16* produced

more glycogen than wildtype. However, starch content in maize endosperm was not increased by changing the number of wildtype *Sh2* or *Bt2* alleles from 2 to 3 doses (Singletary et al., 1997; Hannah and Greene, 1998). This suggested that overexpression of either wildtype gene would have little effect on starch content. Expression of wildtype *E.coli* AGP in potato tuber had no effect on starch content (Stark et al., 1992). However, expression of *glgC-16* that was less dependent on allosteric regulators for activity resulted in 35% starch increase in potato tuber. This points to the physiological importance of the allosteric properties of AGPase. More recently, Giroux et al. (1996) isolated a *Sh2* allele that resulted in a less Pi sensitive AGPase (*Rev6*) compared to wildtype. *Rev6* increased maize seed weight by 11-18% in single dose, confirming the agronomic importance of the allosteric properties of AGPase.

Agronomic Importance of AGPase Heat Stability

Heat stress leads to decreased maize yield (Peters et al., 1971; Thompson, 1975; Chang, 1981; Christy and Williamson, 1985). This can be attributed to reduced photosynthate availability and transportation from source to sink tissues, poor pollination, reduced cell and granule size and number, early seed abortion and reduced grain filling period. Growth of endosperm starts with a lag phase in which cells actively divide and continues with a linear phase where cells increase in size and starch synthesis occurs. Elevated temperature during lag phase resulted in reduced yield (Jones et al., 1984). They suggested that reduced yield was due to reduced cell and granule number and size as well as seed abortion. Additionally, elevated temperatures during the linear phase resulted in shorter grain filling period and subsequently smaller kernels (Jones et al., 1984). Similar results were found by Hunter et al. (1977) and Tollenaar and Bruulsema (1988).

Records for five states that traditionally produce more than 50% of the US corn showed that average daily temperature was 23.6°C, around 2°C higher than optimum during grain filling

(Singletary et al., 1994). Photosynthate availability during grain filling is not reduced at high temperatures, at least in barley and wheat. Indeed, sucrose content in barley and wheat seeds was either unchanged or elevated at high temperatures (Bhullar and Jenner, 1986; Wallwork et al., 1998). Also photosynthesis in maize increases up to 32°C (Duncan and Hesketh, 1968; Hofstra and Hesketh, 1969; Christy et al., 1985). Moreover, Cheikh and Jones (1995) studied the ability of maize kernels to fix ¹⁴C sucrose and hexoses at different temperatures. They found that these types of sugars increased in the seed at elevated temperatures. The evidence above suggests that limited sugar availability and transport into the kernel during grain filling are not the reason for yield decrease at elevated temperatures.

There have been extensive efforts to identify biochemical pathways that impact grain filling during elevated temperatures. Singletary et al. (1993; 1994) assayed enzyme activities in maize kernels grown *in vitro* at elevated temperatures (22°C to 36°C) and found that AGPase and soluble starch synthase (SSS) were more heat labile compared to other enzymes participating in starch synthesis. They suggested that heat lability of AGPase and SSS contributes to grain filling cessation. Duke and Doehlert (1996) found that transcripts of several genes encoding enzymes of the starch synthesis pathway, including *Sh2* and *Bt2*, were decreased at 35°C compared to 25°C. However, enzyme assays showed that only AGPase activity was strikingly lower. They suggested that this could be due to a higher turnover rate of AGPase compared to other enzymes. Finally, Wilhelm et al. (1999), through Q₁₀ analysis, showed that AGPase had the most pronounced reduction in activity compared to several other enzymes. Maize AGPase indeed lost 96% of its activity when heated at 57°C for 5 min (Hannah et al., 1980).

Greene and Hannah (1998a) isolated a mutant form of maize AGP with a single amino acid change in the large subunit termed *HS33*. They showed that the altered enzyme was more heat

stable and that stability was due to stronger subunit-subunit interactions. When wheat and rice were transformed with a *Sh2* variant that contains the *Rev6* and *HS33* changes together, yield was increased by 38% and 23% respectively (Smidansky et al., 2002; 2003). Remarkably, the increase was due to an increase in seed number rather than individual seed weight. The phenotype cannot be explained by *Rev6* since, when expressed alone in maize, it increases seed weight but not seed number. Transformation of maize with the *Sh2* variant containing *Rev6* and *HS33* changes gives similar results. A detailed characterization of the transformants is under way (Greene and Hannah, personal communication). The above studies show the importance of AGPase heat stability in cereal yield.

Cross et al. (2004) generated a mosaic small subunit (MP) that consisted of the first 200 amino acids of BT2 and the last 275 amino acids of the potato tuber small subunit. MP in a complex with SH2 had several features that could lead to agronomic gain (Cross et al., 2004; Boehlein et al., 2005). Some of those features were increased activity in the absence of the activator 3-PGA, increased affinity for 3-PGA and elevated heat stability compared to wildtype maize endosperm AGPase (BT2/SH2). MP and BT2 were separately expressed in maize endosperm under an endosperm specific promoter. Preliminary data show that plants with transgenic MP but no functional endogenous small subunit had a 2-3 fold increase in seed number compared to plants that had only one dose of endogenous BT2 (Hannah, unpublished data). Plants with transgenic BT2 but no functional endogenous small subunit were indistinguishable from plants that had only one dose of endogenous BT2. These results indicate that MP leads to agronomic gain in maize endosperm due to its kinetic and allosteric properties and/or heat stability rather than overexpression.

Goal and Approach

As discussed above, allosteric and heat-stable variants of AGPase can lead to agronomic gain. Our goal was to isolate AGPase variants that may lead to agronomic gain when expressed in plants and more specifically in the cereal endosperm. Promising variants would include AGPase mutants with increased k_{cat} , affinity for substrates, increased sensitivity to the activator 3-PGA or decreased sensitivity to the inhibitor Pi and mutants with increased heat stability.

Sh2 and *Bt2* were randomly mutagenized by error-prone PCR. The mutants were then expressed with the complementary wildtype subunit in *E.coli*. Their activity was scored by glycogen production of *E.coli* as judged by iodine staining. Mutants staining darker than wildtype were selected and sequenced. Using this approach, a dark staining small subunit variant that harbored a single amino acid change compared to wildtype was isolated. A detailed study of the properties of the mutant showed that its dark staining was due to its high heat stability. The agronomic potential of the variant is discussed.

Materials and Methods

Random mutagenesis

Mutations were introduced into *Sh2* and *Bt2* by PCR random mutagenesis (GeneMorph II EZClone Domain Mutagenesis Kit, Stratagene). A mixture of non-biased, error-prone DNA polymerases was used to introduce point mutations. Wildtype *Sh2* and *Bt2* coding sequences in pMONc*Sh2* and pMONc*Bt2* (Giroux et al., 1996) respectively were used as templates for PCR. Two pairs of primers (*Sh2*: 5'GAAGGAGATATATCCATGG 3', 5' GGATCCCCGGGTACCGAGCTC 3' *Bt2*: 5'GAAGGAGATATATCCATGG 3', 5' GTTGATATCTGAATTCGAGCTC 3') flanking *Sh2* and *Bt2* were used for error-prone PCR. Mutant *Sh2* clones produced by PCR were subcloned into vector pMONc*Sh2* according to Stratagene protocols. pMONc*Sh2* was then used to transform *E. coli* strain AC70R1-504 that

contained wildtype *Bt2* in the compatible vector pMONc*Bt2*. Mutant *Bt2* clones produced by PCR were subcloned into vector pMONc*Bt2*. pMONc*Bt2* was then used to transform *E.coli* strain AC70R1-504 that contained wildtype *Sh2* in the compatible vector pMONc*Sh2*.

Bacterial Expression System

A bacterial expression system (Iglesias et al., 1993) allowed us to randomly mutagenize maize endosperm AGPase genes and score AGPase activity in a fast and efficient way. The *E. coli* system is ideal for studying plant AGPases for a number of reasons discussed in Georgelis et al. (2007).

Glycogen Detection

Glycogen synthesis was detected by production of brown staining colonies following exposure to iodine vapors. *E.coli* cells were grown on Kornberg media in the presence of 75 µg/mL spectinomycin, 50 µg/mL kanamycin and 1% w/v glucose for 16 h at 37°C (Govons et al., 1969). The colonies were exposed to iodine vapors for 1 min. Iodine staining provided a fast and convenient means to identify active and inactive AGPases. Colonies with inactive AGPase produced no color following contact with iodine vapors while active AGPase produced glycogen and, in turn, brown staining with iodine. AGPase variants staining darker than wildtype were selected for further study.

Glycogen Quantitation

Glycogen quantitation of AGPase mutants was performed by phenol reaction (Hanson and Phillips, 1981). In brief, glycogen was extracted from *E.coli* cells (OD₆₀₀=2.0) grown in LB containing 2% w/v glucose by boiling for 3 hours in 50% w/v KOH. Glycogen was then precipitated by adding ethanol to 70% v/v and centrifuging at 10000 x g for 10 min. After pellet drying, 200 µl deionized water, 200 µl of 5% w/v phenol and 1 mL of concentrated sulfuric acid were added. Glycogen was estimated by the absorbance at 488 nm.

DNA Sequencing

Sh2 and *Bt2* mutants that produced more glycogen than wildtype were double-pass sequenced by the Genome Sequencing Services Laboratory (GSSL) of the Interdisciplinary Center for Biotechnology Research at the University of Florida. Data analysis was performed by Bioedit software (Hall, 1999).

Purification of Maize Endosperm AGPase from AC70R1-504 *E.coli* Cells

AC70R1-504 *E.coli* cells expressing maize endosperm AGPase were grown in 2 L of Luria-Broth (LB) medium in the presence of 75 µg/mL spectinomycin, 50 µg/mL kanamycin and 2% w/v glucose for 16h at 37°C with shaking. At OD₆₀₀=0.6, 0.2mM isopropyl-beta-D-thiogalactoside (IPTG) and 0.02mg/mL nalidixic acid were added to induce protein expression. The cultures were immediately moved to room temperature and grown for 4h with shaking. The following steps were conducted at 4°C. Cells were harvested by centrifuging at 3000 x g and the pellet was resuspended in 16mL of buffer A (50mM KH₂PO₄ pH 7.0, 5mM MgCl₂, 0.5mM EDTA) and protease inhibitors (1µg/mL pepstatin, 0.1mM PMSF, 10 µg/mL chymostatin, and 1mM benzamidine). The cells were lysed with a French press and centrifuged at 26000 x g. The protein concentration of the supernatant was adjusted to 30mg/mL by adding buffer A. Three tenths of volume of 1% protamine sulfate were added and the mixture stirred on ice for 20 min and then centrifuged at 26000 x g for 20min. The supernatant was brought to 45% saturation with ammonium sulfate, stirred on ice for 20min and centrifuged at 26000 x g for 20min. The pellet was resuspended in 2-2.5mL of buffer A. The mixture went through a strong anion exchange column (macro-prep High Q support, Biorad), and an Econo-pac hydroxyapatite cartridge (Biorad) as described by Boehlein et al. (2005). AGPase was desalted by using Zeba Micro Desalt Spin Columns (Pierce) before assaying. AGPase was exchanged into 50mM HEPES, 5mM MgCl₂, 0.5mM EDTA and 0.5mg/mL BSA (for stability).

Kinetic Characterization of AGPase

The forward direction of the reaction was used ($G-1-P + ATP \rightarrow ADP\text{-glucose} + PPi$) for estimating k_{cat} , K_m for ATP and G-1-P, and affinities for 3-PGA and Pi. More specifically, 0.04-0.06 μ g of purified AGPase was assayed for specific activity in a total volume of 200 μ l of 50mM HEPES pH 7.4, 15mM $MgCl_2$, 1.0mM ATP, and 2.0mM G-1-P at 37°C for 10min. For determining K_m s for ATP and G-1-P and K_a for 3-PGA, varying amounts of ATP, G-1-P and 3-PGA respectively. K_i for Pi was estimated by adding various amounts of Pi, 1mM ATP, 2mM G-1-P, and 2.5mM 3-PGA were used. The reaction was stopped by boiling for 2min. PPi was estimated by coupling the reaction to a reduction in NADH concentration using 300 μ l of coupling reagent. The coupling reagent contained 25mM imidazole pH 7.4, 4mM $MgCl_2$, 1mM EDTA, 0.2mM NADH, 0.725U aldolase, 0.4U triose phosphate isomerase, 0.6U glycerophosphate dehydrogenase, 1mM fructose 6-phosphate and 0.8 μ g of pyrophosphate dependent phosphofructokinase (PPi-PFK). All the enzymes were purchased from Sigma except for PPi-PFK which was produced and purified according to Deng et al. 1999 with some modifications (Boehlein and Hannah, unpublished data). NADH concentration was estimated by absorbance at 340nm. PPi concentration was calculated by a standard curve developed by using various amounts of PPi instead of AGPase. The amount of PPi produced by AGPase was linear with time and enzyme concentration. The kinetic constants of AGPase were calculated by Prism 4.0 (Graph Pad, San Diego CA).

Measuring AGPase Specific Activity from Crude or Partially Purified Protein Extracts

AC70R1-504 *E.coli* cells expressing maize endosperm AGPase were grown in 2 L of Luria-Broth (LB) medium in the presence of 75 μ g/mL spectinomycin, 50 μ g/mL kanamycin and 2% w/v glucose at 37°C with shaking until $OD_{600}=2.0$. Gene expression was not induced. The following steps were conducted at 4°C. Cells were harvested by centrifuging at 3000 x g and the

pellet was resuspended in 16mL of buffer A (50mM KH_2PO_4 pH 7.0, 5mM MgCl_2 , 0.5mM EDTA) and protease inhibitors (1 $\mu\text{g}/\text{mL}$ pepstatin, 0.1mM PMSF, 10 $\mu\text{g}/\text{mL}$ chymostatin, and 1mM benzamidine). The cells were lysed with a French press and centrifuged at 26000 x g. Part of the supernatant was stored at -80°C for determining AGPase activity of the crude extract. The rest of the supernatant went through the protamine sulfate and ammonium sulfate steps described above. All protein extracts were desalted as described by Boehlein et al. (2005) before assaying and exchanged into 50mM HEPES, 5mM MgCl_2 , and 0.5mM EDTA. AGPase specific activity was monitored in the reverse direction ($\text{ADP-glucose} + \text{PPi} \rightarrow \text{G-1-P} + \text{ATP}$) using saturating amounts of substrates and activator as described by Boehlein et al. (2005).

Determining Heat Stability of AGPase

AGPase was purified as described above. AGPase was further diluted 1/100 (v/v) in 50mM HEPES, 5mM MgCl_2 , 0.5mM EDTA and 0.5mg/mL BSA and heat treated at 42 or 53°C for various times, and then cooled on ice. The activity remaining after heat treatment was monitored in the forward and reverse direction by using saturating amounts of ATP, G-1-P and 3-PGA. The data were plotted as log of percentage of remaining activity versus time of heat treatment. The inactivation constant $t_{1/2}$ was calculated from the formula $t_{1/2} = 0.693/(-2.3 \cdot \text{slope})$.

Qualitative Determination of the AGPase Purity

The purity of AGPase was monitored in the following way. Six μg of AGPase were diluted 1:1 in denaturing solution (100 mM Tris-Cl pH 6.8, 4% SDS, and 8mM DTT), heated at 95°C for 5min, electrophoresed on a 5% SDS polyacrylamide gel at 150V for 1h and visualized by staining with Coomassie Brilliant Blue (Laemmli, 1970).

Protein Blot Analysis of Crude Extracts

Samples were vacuum blotted onto a PVDF membrane (Biorad) by using Hybri-Dot blot apparatus (Life Technologies). The PVDF membrane had been pre-soaked in methanol for 5min

and then in transfer buffer [20% (v/v) methanol, 0.303% (w/v) Tris and 1.44% (w/v) glycine] for 10min. The membrane was incubated with blocking buffer [0.8% (w/v) NaCl, 0.02% (w/v) KCl, 0.144% (w/v) Na₂HPO₄, 0.024% (w/v) KH₂PO₄, 5% (w/v) bovine serum albumin (BSA), and 0.05% (v/v) Tween-20] for 1h with constant shaking. The blot was incubated with blocking buffer containing 1:10000 (v/v) of monoclonal antibody against BT2 (kindly provided by Sue Boehlein) for 1h with shaking. Then, the blot was washed 3 x 10 min with washing buffer (blocking buffer –BSA) with constant shaking. The blot was then incubated with a 1:60000 dilution of goat anti–mouse secondary antibody conjugated with Horseradish peroxidase (Pierce) for 45min. Finally the blot was washed 3 x 10 min. Proteins were visualized using an enhanced chemiluminescent substrate kit (Pierce).

3D Modeling

BT2 monomer structure was modeled after the potato tuber AGPase small subunit known 3D structure (RCSB Protein Data Bank #:1YP2). Homology modeling was done by using SWISS MODEL (Peitsch, 1995; Guex and Peitsch, 1997; Schwede et al. 2003; Kopp and Schwede, 2004; Arnold et al, 2006). Amino acid 462 (Thr or Ile) contacting residues were determined by using Jmol, an open-source Java viewer for chemical structures in 3D (<http://www.jmol.org/>).

Yeast Two Hybrid

Yeast transformations and a β -galactosidase assay were conducted as indicated in Greene and Hannah (1998b). The only modifications were the use of pGBKT7 and pGADT7 as vectors for the bait and the prey respectively. pGBKT7-53 and pGADT7-T plasmids were used as a positive control.

Results

One library of *Bt2* and one library of *Sh2* mutants were created by error-prone PCR. The mutational load of both libraries was ~2 non-synonymous mutations per clone. The mutants were expressed in *E.coli* along with a wildtype complementary subunit gene. Approximately 50000 colonies from the *Bt2* and 20000 colonies from the *Sh2* library were screened for glycogen production. Ten and three colonies from the *Bt2* and *Sh2* library respectively were picked as dark stainers. To verify that the dark staining was gene- and not cell-specific, plasmid DNA extracted from these colonies was used to re-transform *E.coli*. Indeed, the freshly transformed cells stained darker than cells transformed with wildtype *Bt2* and *Sh2*. The two darkest staining *Bt2* mutants were sequenced. Both had the same non-synonymous mutation resulting in a change of BT2 amino acid 462 from threonine to isoleucine (TI). The amino acid in that position (Thr) is absolutely conserved among the higher plant small subunits (data not shown). Glycogen quantitation provided support that BT2-TI/SH2 produced more glycogen than did BT2/SH2 (Figure 3-1). Cells expressing BT2-TI and BT2, as homotetramers, did not produce detectable amounts of glycogen (Figure 3-1). This indicates that the amount of glycogen, produced by *E.coli*, depends exclusively on the complex of BT2-TI or BT2 with SH2.

A dot-blot of the crude extracts from cells expressing BT2/SH2 and TI/SH2 indicated that TI is found in higher amounts in *E.coli* (Figure 3-2). The AGPase activity of the crude extract from non-induced cells expressing BT2/SH2 and BT2-TI/SH2 was too low to detect. However, the partially purified extract from BT2-TI/SH2 had 20 times more activity than the partially purified extract from BT2/SH2 (Figure 3-3). The possibility that BT2-TI/SH2 produced more protein and activity because of more efficient transcription/translation is unlikely since the codons ACA (T) to ATA (I) are used with the same frequency in *E.coli* (6.1 and 5.0%

respectively)(Nakamura et al., 2000). I suggest that the higher amount of protein and activity in BT2-TI/SH2 cells is due to increased stability of the AGPase.

To determine the kinetic properties and heat stability of BT2-TI/SH2 and understand why it produces more glycogen in *E.coli*, recombinant BT2-TI/SH2 along with BT2/SH2 were purified (Figure 3-4). The kinetic properties of BT2-TI/SH2, as summarized in Table 3-1, show that the K_m for G-1-P and ATP, K_a for 3-PGA and K_i for Pi were indistinguishable from BT2/SH2. The k_{cat} of BT2-TI is 30% lower than the k_{cat} of BT2/SH2. These kinetic properties cannot account for the darker staining of BT2-TI/SH2 in *E.coli*. However, BT2-TI/SH2 is clearly more heat stable than BT2/SH2 (Figure 3-5A,B). These results strongly suggest that the high heat stability of BT2-TI/SH2 accounts for the high amount of glycogen in *E.coli*.

As mentioned before, MP is a small subunit variant that can lead to agronomic gain. Some of its features were increased activity in the absence of the activator 3-PGA, increased affinity for 3-PGA, decreased affinity for Pi (Table 3-1) and elevated heat stability compared to BT2/SH2 (Figure 3-3)(Cross et al., 2004; Boehlein et al., 2005). Since BT2-TI/SH2 was not as heat stable as MP/SH2 (Figure 3-3), the amino acid change of TI was introduced into MP in an effort to further increase the heat stability of MP.

Cells expressing MP-TI/SH2 produced the same amount of glycogen as cells expressing MP/SH2 (Figure 3-1). However, greater amounts of the MP-TI protein relative to BT2 were found in crude extracts of *E.coli* expressing the two proteins with SH2 (Figure 3-2). The activity of the crude extracts and the partially purified extracts from MP-TI/SH2 was 2-3 fold higher than from MP/SH2 (Figure 3-3). MP-TI/SH2, in its pure form (Figure 3-4), maintained the favorable kinetic properties of MP/SH2 (Table 3-1) except that its k_{cat} was reduced ~30% compared to MP/SH2. Additionally, MP-TI/SH2 had greater heat stability than MP/SH2 (Figure 3-6).

The crystal structure of maize endosperm AGPase has not been resolved. The only relevant structure is a potato tuber small subunit homotetramer (Jin et al., 2005). The potato tuber small subunit shows 88% identity and 96% similarity to BT2. BT2 monomer structure was modeled after the resolved structure of the potato tuber small subunit (Figure 3-7A). The residue mutated in TI is part of a β -helix and it makes hydrophobic contact with two residues (Pro, Leu) of the N-terminus of the small subunit (Figure 3-7B). The amino acid change from Thr to Ile in TI shortens the distance from the Pro and Leu mentioned above (Figure 3-7C). It is tempting to speculate that the TI mutation strengthens the hydrophobic interaction between the C- and the N-terminus of the small subunit and results in greater stability. Unlike MP, whose heat stability is attributed to residues at/near the subunit-subunit interfaces, TI may not directly affect subunit-subunit interactions since it is far from the subunit-subunit interfaces (Figure 3-7A).

To determine whether TI affects the strength of subunit-subunit interactions SH2 was used as a bait and BT2, TI, and MP were used as a prey in a yeast two-hybrid system. A quantitative β -galactosidase assay indicated that, in contrast to MP, TI did not increase the strength of subunit-subunit interactions (Figure 3-8).

Discussion

Our goal was to isolate agronomically promising maize endosperm AGPase variants by using random mutagenesis and a heterologous bacterial expression system. BT2-TI was isolated as a small subunit variant that increased the amount of glycogen produced by *E.coli* cells when expressed along with SH2. Cells expressing BT2-TI/SH2 had 20-fold higher AGPase activity than cells expressing BT2/SH2. A dot-blot indicated that the crude protein extract from cells expressing BT2-TI/SH2 had more detectable BT2 protein compared to cells expressing BT2/SH2. This result could be attributed to more efficient transcription/ translation or to greater AGPase stability and/or solubility. As mentioned previously, a more efficient

transcription/translation is unlikely based on codon usage. On the other hand, it was showed that the purified form of BT2-TI/SH2 was significantly more heat-stable than the purified form of BT2/SH2. This may render the BT2-TI/SH2 complex less prone to proteolysis and/or aggregation compared to BT2/SH2 in *E.coli*.

The kinetic and allosteric properties of BT2-TI/SH2 were indistinguishable from BT2/SH2 except for a 30% lower k_{cat} . These results point out that the only feature of BT2-TI/SH2 that may lead to agronomic gain when expressed in plants is its higher heat stability compared to BT2/SH2. However, heat stability seems to be a crucial factor for AGPase activity in *E.coli* cells grown at 37°C. I interpret the 20-fold increase in AGPase activity of BT2-TI/SH2 expressing cells as a higher number of active AGPase molecules compared to BT2/SH2 expressing cells. This also means that less than 5% of the potential AGPase molecules actually function in BT2/SH2.

It has been reported that the potato tuber small subunit can form a homotetramer that has significant activity when given extremely high amounts of the activator 3-PGA (Ballicora et al., 1995). Here I show that *E.coli* cells expressing BT2 and BT2-TI as homotetramers do not produce detectable amounts of glycogen. Hence, the increased amounts of glycogen observed in cells expressing BT2-TI/SH2 compared to cells expressing BT2/SH2 is due to the complex of BT2-TI with SH2 rather than the BT2-TI homotetramer.

Another small subunit variant that results in increased heat stability in a complex with SH2 is MP. The heat stability conferred by MP has been mapped to residues near or at the subunit-subunit interaction interfaces (Boehlein, Shaw, Stewart, and Hannah, unpublished data). In contrast, the amino acid change of BT2-TI is far from these interfaces. Structure modeling suggests that the TI change strengthens the intra-subunit hydrophobic interactions between the

C- and the N-terminus. The results from Y2H support the idea that, in contrast to MP, TI does not strengthen the subunit-subunit interactions. However, the possibility that TI indirectly affects subunit-subunit interactions through a conformational change cannot be dismissed. It is possible that Y2H may not be sensitive enough to reveal a difference in subunit-subunit interactions between TI/SH2 and BT2/SH2 at the yeast growth temperature of 30°C.

A comparison of the heat stability of pure BT2-TI/SH2 and MP/SH2 indicated that BT2-TI/SH2 was not as heat-stable as MP/SH2. Additionally, MP/SH2 has several advantages not shared by BT2-TI/SH2, such as activity in the absence of the activator 3-PGA, a higher affinity for 3-PGA, a lower affinity for the inhibitor Pi and a higher k_{cat} compared to BT2/SH2. Clearly, MP/SH2 is a more promising AGPase variant for agronomic gain. It was investigated whether the heat stability of MP/SH2 could be further improved by introducing the TI change into MP. The resulting variant, MP-TI, when expressed with SH2 in *E.coli*, yielded an equal amount of glycogen as MP/SH2. This could mean that either MP-TI/SH2 was not more heat stable than MP/SH2 or that the production of ADP-glucose catalyzed by AGPase was not limiting in *E.coli* anymore. I heavily favor the latter interpretation since AGPase activity in crude and partially purified extracts of cells expressing MP-TI/SH2 was 2-3 fold higher than cells expressing MP/SH2. MP-TI/SH2 maintained all the kinetic and allosteric properties of MP/SH2 with the exception a 30% lower k_{cat} . Most importantly, MP-TI/SH2 was more heat stable than MP/SH2. Two phases of heat stability in both MP-TI/SH2 and MP/SH2 were observed. These phases were probably a result of different states of AGPase. This biphasic mode of heat stability has been observed before and it is not specific to MP/SH2 or MP-TI (Boehlein, Shaw, Stewart and Hannah, submitted). The first phase shows lower heat stability than does the second one. MP-TI/SH2 is more heat-stable in both phases compared to MP/SH2. However, what exactly the

state of AGPase in each phase remains enigmatic. The biphasic mode of heat stability was not observed in BT2/SH2 and BT2-TI/SH2 because the samples were not heated enough time to reach the second phase and they were heated at lower temperature than MP/SH2 and MP-TI/SH2 (42°C instead of 53°C).

Is MP-TI a variant worth expressing in the endosperm of grasses instead of MP? As mentioned above, the TI mutation increased heat stability but reduced the k_{cat} of the resulting AGPase. Whether MP-TI leads to higher starch yield in plants than MP remains an open question. In *E.coli*, MP-TI/SH2 resulted in higher AGPase activity than did MP/SH2 at 37°C. Under the growing conditions of *E.coli*, heat stability of AGPase is apparently of paramount importance. On the other hand, maize is commercially grown in areas with average temperature lower than 37°C during the grain filling period. However, it is possible that MP-TI is more stable than MP at lower temperatures too. Additionally, plants are subject to variable periods of temperatures of 37°C or higher in some growth areas and MP-TI may confer an advantage to plants growing under these conditions.

Table 3-1. Kinetic properties of purified recombinant AGPase variants.

	BT2/SH2	BT2-TI/SH2	MP/SH2	MP-TI/SH2
K_m G-1-P (mM)	0.050 (±0.008)	0.040 (±0.006)	0.079 (±0.007)	0.059 (±0.005)
$k_{cat}(s^{-1})$ (G-1-P)	38.170 (±1.323)	26.200 (±1.401)	62.655 (±2.569)	42.880 (±1.880)
K_m ATP (mM)	0.102 (±0.020)	0.146 (±0.050)	0.133 (±0.021)	0.112 (±0.013)
$k_{cat}(s^{-1})$ (ATP)	43.321 (±1.554)	29.112 (±1.030)	69.337 (±3.276)	49.031 (±1.903)
K_a 3-PGA (mM)	0.480 (±0.137)	0.330 (±0.060)	0.100 (±0.010)	0.068 (±0.010)
K_i Pi(mM)	2.320 (0.530,4.100)	4.070 (2.120,6.020)	6.610 (4.530,8.690)	5.870 (4.410,7.330)

$V_{max} - 3\text{-PGA} / V_{max} + 3\text{-PGA}$ (nmol/min/mg)			
BT2/SH2	BT2-TI/SH2	MP/SH2	MP-TI/SH2
280/4000 (7.2±3.2%)	134/2737 (4.9±1.5%)	1856/6584 (28.2±3.1%)	1101/4513 (24.4±5.3%)

The kinetic and allosteric properties of AGPase variants were determined in the forward direction. $k_{cat}(s^{-1})(G-1-P)$ was estimated by varying the amount of G-1-P and keeping ATP at saturating amounts (1mM). $k_{cat}(s^{-1})(ATP)$ was estimated by varying the amount of ATP and keeping G-1-P at saturating amounts (2mM). K_i 's are expressed as mean (95% confidence interval). All other values are expressed as mean ± standard deviation. The specific activity of AGPase in the absense of 3-PGA is expressed as a percentage of the specific activity in the presense of 10mM of 3-PGA (mean ± standard error).

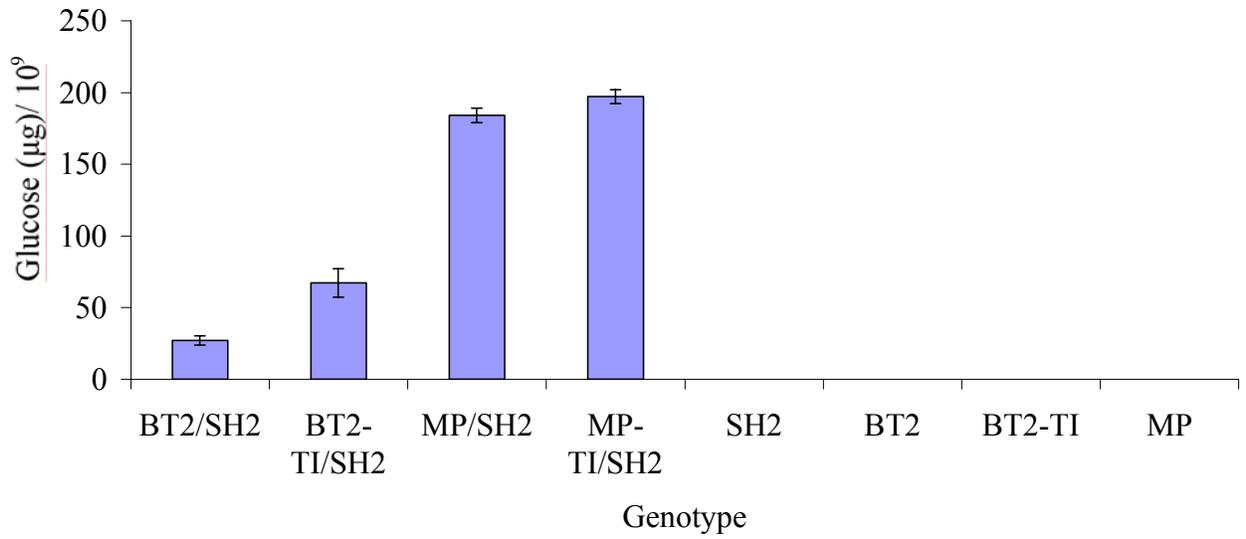


Figure 3-1. Glycogen produced by *E.coli* cells expressing BT2, BT2-TI, MP, MP-TI along with SH2. Glycogen from cells expressing only SH2, BT2, BT2-TI, and MP alone. Glycogen is measured in glucose units. The error bars indicate standard deviation (N=3).

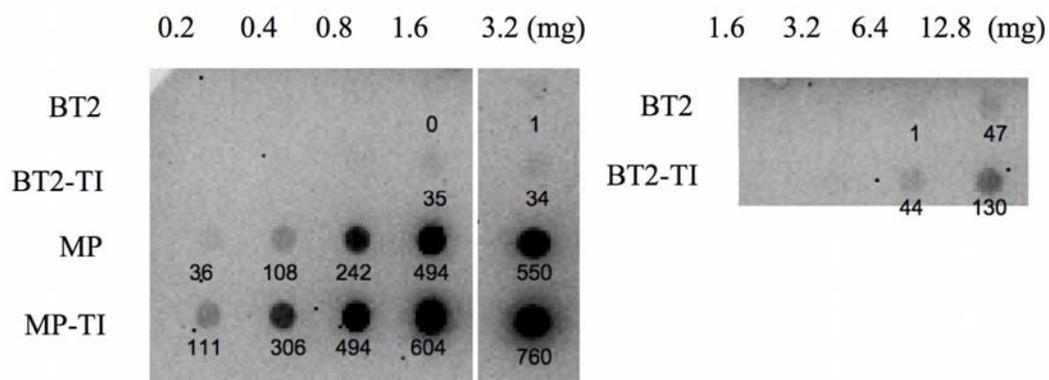


Figure 3-2. Dot blots of crude extracts from *E. coli* cells expressing BT2, BT2-TI, MP, MP-TI with the complementary subunit SH2. AGPase was visualized by using a monoclonal antibody against BT2. The density of the spots was estimated by using ImageJ (<http://rsb.info.nih.gov/ij/>).

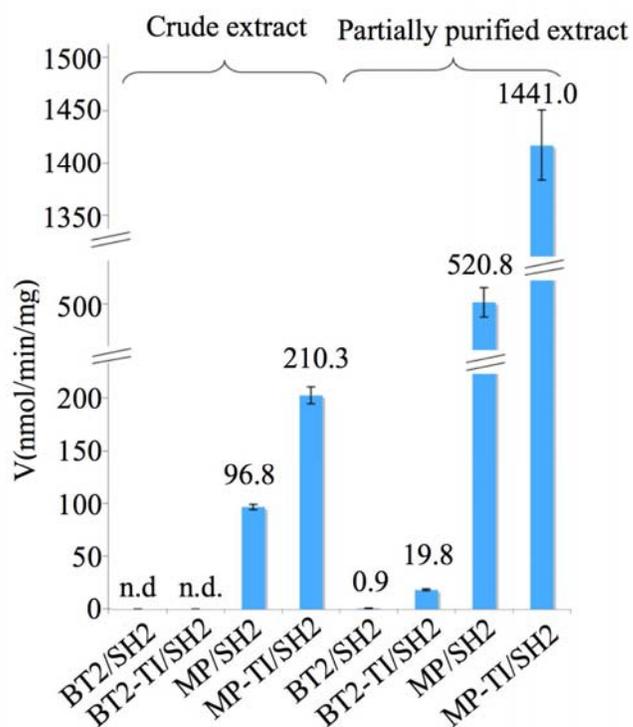


Figure 3-3. Specific activity of AGPase variants in crude and partially purified protein extracts from non-induced *E.coli* cells. Activity was measured in the reverse direction. n.d.:not detectable. The error bars indicate standard deviation (N=3).

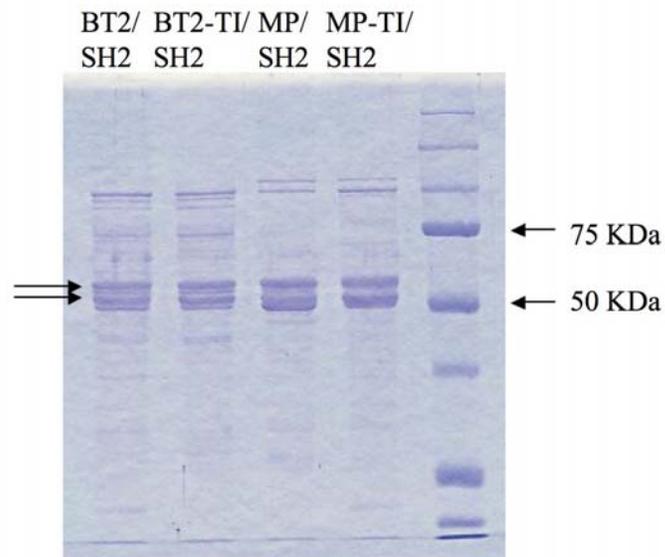
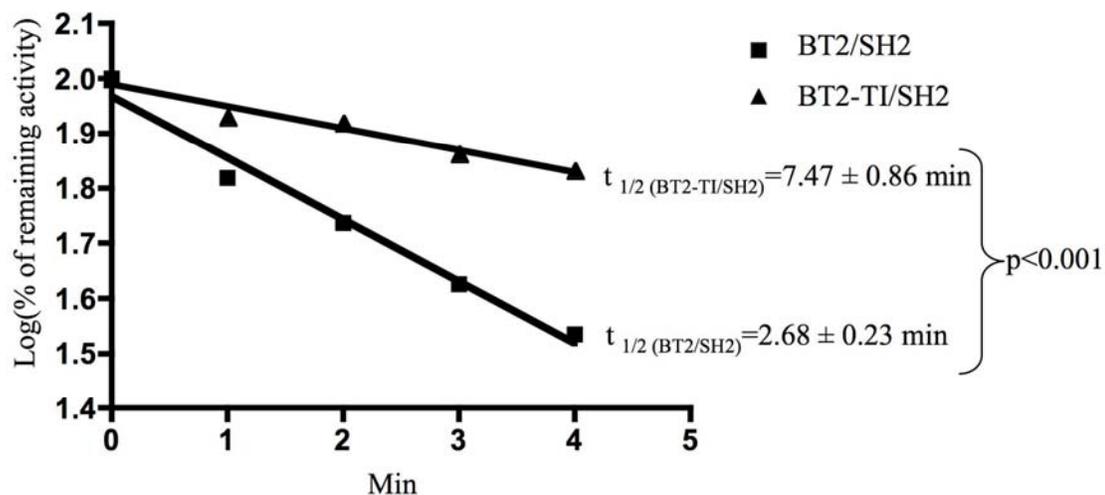


Figure 3-4. Purification of AGPase. SDS-PAGE of purified recombinant BT2/SH2, TI/SH2, MP/SH2, and MP-TI/SH2. Precision Plus Protein All Blue Standard from Biorad was used as a marker. The upper arrow on the left points to the large subunit. The lower arrow on the left points to the small subunit.

A. Forward direction 42°C



B. Reverse direction 42°C

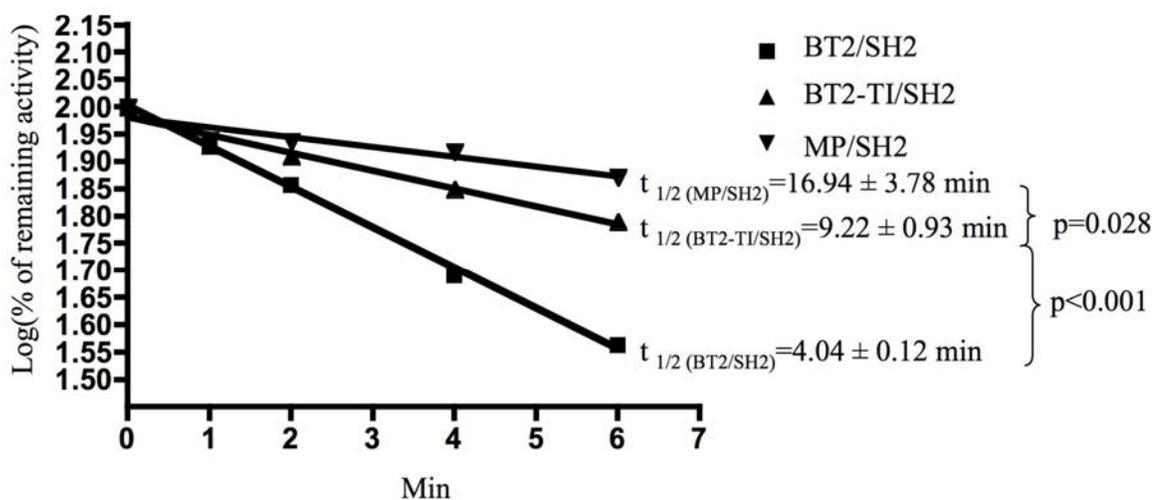


Figure 3-5. Heat stability of purified BT2/SH2, BT2-TI/SH2, and MP/SH2. The half-life ($t_{1/2}$) of each AGPase is expressed as mean \pm standard error. The p-values are estimated by an F-test implemented by Prizm (Graph pad, San Diego CA). A) The assay was conducted in the forward direction. B) The assay was conducted in the reverse direction.

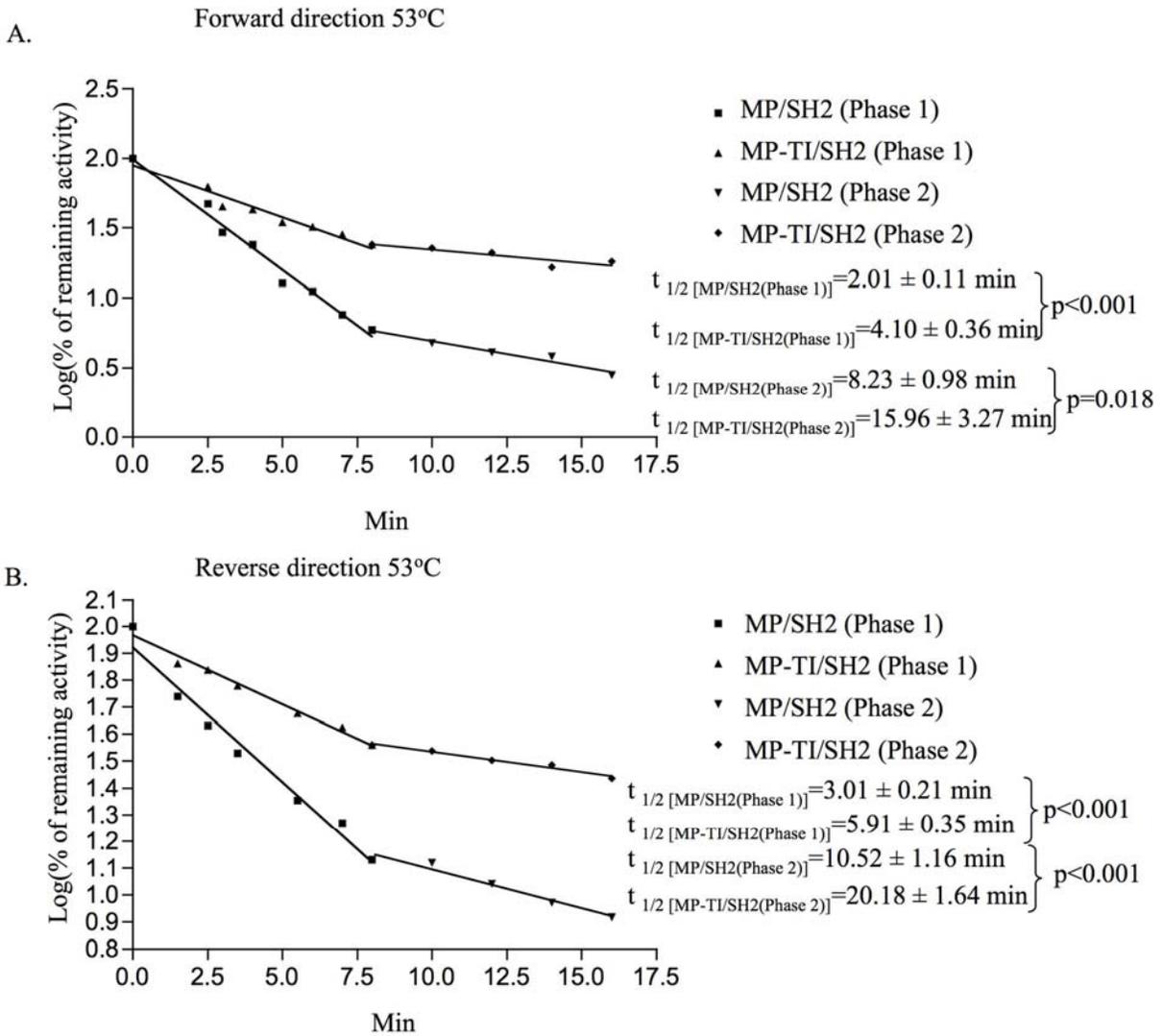


Figure 3-6. Heat stability of purified MP/SH2 and MP-TI/SH2. The half-life ($t_{1/2}$) of each AGPase is expressed as mean \pm standard error. The p-values are estimated by an F-test implemented by Prizm (Graph pad, San Diego CA). A) The assay was conducted in the forward direction. B) The assay was conducted in the reverse direction.

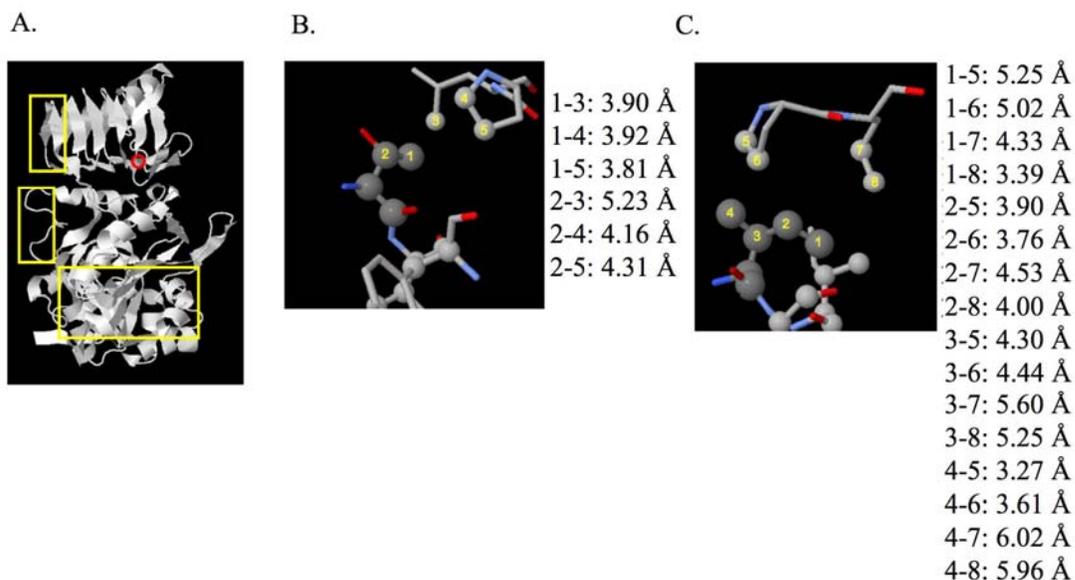


Figure 3-7. 3D modeling of BT2 and TI. A) Predicted 3D structure of BT2 monomer. The TI change is marked by a red circle. The areas of BT2 that are directly involved in subunit-subunit interactions are highlighted by yellow boxes. B) Distances of carbon atoms of Thr462 (1,2) from those of Pro60 (4,5) and Leu61 (3). C) Distances of carbon atoms of Ile462 (1,2,3,4) from those of Pro60 (5,6) and Leu61 (7,8). The Thr462 and Ile462 contacting residues were determined by using FirstGlance Jmol. Dark gray spheres indicate carbon atoms of Thr462 and Ile462. Light gray spheres indicate carbon atoms of contacting residues. Oxygen and nitrogen atoms are indicated by red and blue color respectively.

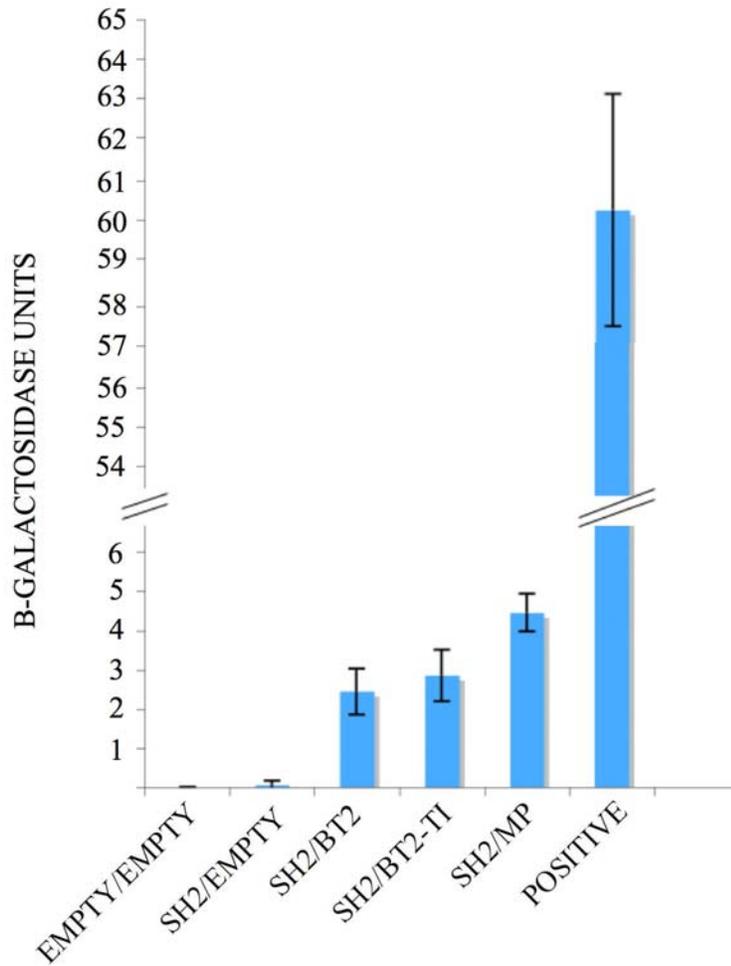


Figure 3-8. Strength of AGPase subunit-subunit interactions. SH2 was used as a bait and BT2, TI, and MP as a prey in a yeast two hybrid system. A quantitative β -galactosidase assay was used to quantify the interactions between the bait and the prey. The error bars indicate 2 x standard error (N=4).

CHAPTER 4 DISCUSSION AND FUTURE DIRECTIONS

Plants are the exclusive producers of starch, the major source of carbohydrates for animals, including humans. The first committed step in starch biosynthesis is catalysed by ADP-glucose pyrophosphorylase (AGPase). Evidence from maize, wheat, and rice endosperm and potato tuber indicates that AGPase catalyses a rate-limiting step in starch synthesis (Stark et al., 1992; Giroux et al., 1996; Smidansky et al., 2002; 2003). This is the main reason that AGPase has received considerable attention.

A bacterial expression system was developed by Iglesias et al. (1993) that allowed the expression of plant AGPase in *E. coli*. This accelerates the phenotypic evaluation of various plant AGPase variants and facilitates their purification and study of their kinetic properties and stability. Here we use this system in combination with random mutagenesis to generate numerous mutants in the small and the large subunit of the major maize endosperm AGPase and evaluate their effects on AGPase activity. By using that methodology, we attempted to understand the evolutionary differences between the large and the small subunit of AGPase in angiosperms and isolate AGPase variants that can potentially lead to agronomic gain when expressed in plants.

AGPase is a heterotetramer consisting of two identical small and two identical large subunits in algae and angiosperms. The large and the small subunit are encoded by *Shrunken-2* (*Sh2*) and *Brittle-2* (*Bt2*), respectively, in the maize endosperm. It is an allosteric enzyme that is upregulated by 3-PGA and downregulated by inorganic phosphate (Pi). Even though the two subunits are clearly paralogs, they are not functionally interchangeable. Site-directed mutagenesis has been used to delineate the roles of each subunit in AGPase activity. It has been proposed that the small subunit is catalytic while the large simply modulates the response to

allosteric regulators. One observation for this model is the fact that the large subunit is considerably less conserved in angiosperms than is the small subunit. It has been speculated that the small subunit, being the catalytic subunit, has more evolutionary constraints than does the merely regulatory large subunit. In this research, we showed that the small subunit has indeed more evolutionary constraints than does the large subunit. We tested the above hypothesis by expressing a large series of *Sh2* and *Bt2* mutants in a heterologous bacterial expression system. We found that the two subunits were equally susceptible to activity-altering amino acid changes when expressed in one environment. This is not consistent with the idea that the small subunit is the catalytic subunit while the large simply fine-tunes the response of the small subunit to allosteric regulators. The above hypothesis was further disputed by studies that showed that mutations in the large subunit specifically affected the catalytic properties of AGPase (Hannah and Nelson, 1976; Cross et al., 2004; Cross et al., 2005; Hwang et al., 2006; 2007). It is becoming clear that both subunits are important in catalytic and allosteric properties.

In an effort to explain the higher constraints of the small subunit in angiosperms, we suggest that the difference in constraints between the subunits is due to tissue-specificity and we provided evidence consistent with our hypothesis. More specifically, we suggested that the small subunit evolves slower because it has broad tissue specificity and it has to interact with different large subunits. On the other hand, the large subunit is more tissue-specific and it usually interacts with only one small subunit. Our study complements large-scale bioinformatics studies, which have also indicated that broadly expressed genes evolve slower than do narrowly expressed genes (Duret and Mouchiroud, 2000; Akashi, 2001; Wright et al., 2004). In these studies, thousands of genes encoding unrelated proteins were used. Further support of our hypothesis comes from studies of several protein families, which have members with varying degrees of

tissue-specificity (Hastings, 1996; Schmidt et al., 1997). More specifically, members with broad tissue specificity evolved slower than members with narrow tissue specificity within 15 out of 16 protein families. These results along with our data indicate that two proteins that have almost identical structures and perform the same function can evolve with different rates because of different expression patterns. Hence, tissue specificity should be taken into account when the evolutionary rates of different proteins are compared.

The large subunits can be classified into different groups by expression pattern. Amino acids that are invariant in one pattern of expression but differing from large subunits expressed in different tissues can be identified and then modified by site-directed mutagenesis to determine whether their identity is essential for expression in the identified tissues.

An interesting observation we noted was that the large subunit had more successful duplications than did the small subunit. Even though gene duplicability is poorly understood it has been correlated with the complexity of the encoded protein (number of protein modules) and the number of cis-elements of the promoter (He and Zhang, 2005; Lin et al., 2007). It has been speculated that a high number of protein modules and cis-elements make the gene more “splittable”, increasing the chances of subfunctionalization and retainment of both genes after a gene duplication. The large and the small subunit should not differ in the number of protein modules since they are clearly paralogs. However, there is evidence that the transcription of the large subunit genes is more responsive to stimuli, such as sucrose and ABA, than is the transcription of the small subunit genes (Müller-Röber et al., 1990; Prioul et al., 1994; Weber et al., 1995; Burgess et al., 1997; Choi et al., 2001; Akihiro et al., 2005; Crevillén et al., 2005). This suggests that the large subunit genes have more cis-elements in their promoters, which may explain why the large subunit genes duplicate more than the small subunit genes. Perusal of the

promoters of the large and small subunit genes in completely sequenced genomes may offer some insight on that matter. Sequencing of more plant genomes in addition to Arabidopsis, rice and Populus will be helpful in that respect.

As stated above, AGPase catalyzes a rate-limiting step in starch synthesis. The allosteric properties and heat stability of AGPase are of paramount agronomic importance. Even though AGPase variants with desirable allosteric properties and increased heat stability are already available there is always room for isolating additional AGPase variants or improving the existing ones. The methodology of random mutagenesis along with expression of the variants in a heterologous bacterial system, that was used to address the first objective of this research, was used to isolate agronomically promising AGPase variants. Isolation and subsequent characterization of variants with greater glycogen production in bacteria led to the isolation of a heat stable AGPase variant (TI). Its heat stability was transferable to an existing promising AGPase variant, called MP.

The heat stability of TI was conferred by a T to I change at position 462 of the small subunit. Exploration of more changes in that position may lead to an even greater heat stability. Additionally, different agronomically promising AGPase variants can be isolated through our current methodology. Introduction of such mutations into existing promising variants such as MP and MP-TI can lead to further improvement of the existing variants.

Whether MP-TI leads to higher starch yield in cereal endosperm than does MP remains an open question. It is possible that MP-TI leads to higher starch yield only under high temperatures. It is also possible that downstream enzymes such as starch synthases become limiting in plants expressing MP-TI. Isolation of more active or heat stable variants of other

enzymes in starch synthesis and expression along with MP-TI in cereal endosperm may lead to higher agronomic gain than expressing MP-TI alone.

APPENDIX A
SMALL AND LARGE AGPase SUBUNIT ACCESSION NUMBERS

Table A-1. Cyanobacterial and small AGPase subunit accession numbers.

Accession	Name	Tissue(s) of expression	Reference(s)
S24991	<i>Anabaena</i> sp.PCC 7120	-	Charng et al. 1992
Small subunits			
Accession	Name	Tissue(s) of expression	Reference(s)
P30523	<i>Triticum aestivum</i>	seed endosperm	Burton et al. 2002
AAQ14870	<i>Zea mays 1</i>	seed endosperm	Bae et al. 1990
AAK69628	<i>Zea mays 2</i>	leaf	Hannah et al. 2001
AAK39640	<i>Zea mays 3</i>	seed embryo	Hannah et al. 2001
CAA88449	<i>Hordeum vulgare 1</i>	seed endosperm	Johnson et al. 2003; Thorbjornsen et al. 1996
AAO16183	<i>Hordeum vulgare 2</i>	seed endosperm	Johnson et al. 2003
XP_481806	<i>Oryza sativa 2</i>	seed endosperm	Ohdan et al. 2005; Akihiro et al. 2005
AAK27313	<i>Oryza sativa 1</i>	seed endosperm, leaf	Ohdan et al. 2005; Akihiro et al. 2005; Sikka et al. 2001
AAK27721	<i>Cicer arietinum 1</i>	seed, weak in leaf	Singh et al. 2002
AAK27720	<i>Cicer arietinum 2</i>	seed, weak in leaf	Singh et al. 2002
CAB89863	<i>Brassica napus</i>	leaf, seed	Vigeolas et al. 2004
AAK27684	<i>Brassica rapa</i>	leaf	No publication
CAA65540	<i>Pisum sativum 1</i>	seed, weak in leaf,stem,root	Burgess et al. 1997
CAA65539	<i>Pisum sativum 2</i>	seed, leaf	Burgess et al. 1997
CAA54259	<i>Vicia faba 1</i>	seed	Weber et al. 1995
CAA54260	<i>Vicia faba 2</i>	leaf,seed	Weber et al. 1995
BAC66693	<i>Phaseolus vulgaris</i>	seed	Omoto et al. 2003
P23509	<i>Solanum tuberosum</i>	tuber,leaf	Nakata and Okita 1996 Nakata et al. 1994
CAA58473	<i>Ipomoea batatas 1</i>	tuber,leaf	No publication
CAB01912	<i>Ipomoea batatas 2</i>	tuber,stem,leaf	Bae and Liu 1997
AAB00482	<i>Solanum lycopersicum</i>	leaf,fruit	Chen et al. 1998
CAA58475	<i>Spinacea oleracia</i>	leaf	No publication
AAF66434	<i>Perilla frutescens 1</i>	stem, root, seed, leaf	Choi et al. 2001
AAF66435	<i>Perilla frutescens 2</i>	stem, root, seed, leaf	Choi et al. 2001
AAB09585	<i>Arabidopsis thaliana</i>	leaf,stem, inflorescence	Villand et al. 1993; Kim et al. 1998; Wang et al. 1998
AAB91466	<i>Citrullus lanatus</i>	fruit	Kim et al. 1998
AAS00541	<i>Fragaria ananassa</i>	Unknown	No publication

Table A-1. Continued.

AAD56041	<i>Citrus unshiu</i>	fruit, leaf	Kim et al. 2001
CAA55515	<i>Beta vulgaris</i>	leaves, weak in hypocotyl, tap root, root	Müller-Röber et al. 1995
AAB91462	<i>Cucumis melo</i>	leaf, early fruit, weak in root and stem	Park et al. 1998
573143 (Protein ID in JGI)	<i>Populus trichocarpa</i>	Unknown	No publication
AAF75832	<i>Chlamydomonas reinhardtii</i>	-	Zabawinski et al. 2001
AAS88879	<i>Ostreococcus tauri</i>	-	Ral et al. 2004
32753 (Protein ID in JGI)	<i>Ostreococcus lucimarinus</i>	-	No publication

Table A-2. Large AGPase subunit accession numbers.

Accession	Name	Large subunits	
		Tissue of expression	Reference
CAA79980	<i>Triticum aestivum</i>	seed endosperm	Burton et al. 2002; Ainsworth et al. 1995
P55241	<i>Zea mays 1</i>	seed endosperm	Bhave et al. 1990
CAA86227	<i>Zea mays 2</i>	embryo	Giroux et al. 1995
Shaw and Hannah unpublished	<i>Zea mays 3</i>	leaf	No publication
AAB94012	<i>Sorghum bicolor</i>	seed endosperm	Chen et al. 1998
CAA47626	<i>Hordeum vulgare 1</i>	seed endosperm	Eimert et al. 1997; Villand et al. 1992
AAC49729	<i>Hordeum vulgare 2</i>	leaf, weak in endosperm	Eimert et al. 1997
BAA23490	<i>Oryza sativa 1</i>	seed endosperm	Ohdan et al. 2005; Akihiro et al. 2005
AAB38781	<i>Oryza sativa 2</i>	seed endosperm	Ohdan et al. 2005; Akihiro et al. 2005
AAT78793	<i>Oryza sativa 4</i>	leaf	Ohdan et al. 2005; Akihiro et al. 2005
NP_911710	<i>Oryza sativa 3</i>	stem,leaf,seed	Ohdan et al. 2005; Akihiro et al. 2005
AAK27718	<i>Cicer arietinum 2</i>	leaf	Singh et al. 2002
AAK27719	<i>Cicer arietinum 1</i>	seeds,weak in leaf,stem and root	Singh et al. 2002
AAK27685	<i>Brassica rapa</i>	leaf	No publication
CAA65541	<i>Pisum sativum</i>	sink tissues: seed,pod,seed coat	Burgess et al. 1997
BAC66692	<i>Phaseolus vulgaris</i>	seed	Omoto et al. 2003
CAA53741	<i>Solanum tuberosum 3</i>	tuber	La Cognata et al. 1995
CAA52917	<i>Solanum tuberosum 2</i>	leaf, tuber	La Cognata et al. 1995
CAA43490	<i>Solanum tuberosum 1</i>	tuber,weak in leaf	Nakata and Okita 1995
AAC21562	<i>Ipomoea batatas</i>	tuber, stem	Harn et al. 2000
AAC49941	<i>Solanum lycopersicum</i> 2	stems,roots,early fruit	Park and Chung 1998
AAC49943	<i>Solanum lycopersicum</i> 3	leaf	Park and Chung 1998
AAC49942	<i>Solanum lycopersicum</i> 1	fruit, root	Park and Chung 1998
AAD56405	<i>Solanum habrochaites</i>	fruit	Schaffer et al. 2000
AAF66436	<i>Perilla frutescens</i>	stem,cotyledon	Choi et al. 2001
AAM14190	<i>Arabidopsis thaliana 1</i>	low levels of expression	Crevillen et al. 2005
AAM20291	<i>Arabidopsis thaliana 3</i>	sink tissues	Crevillen et al. 2005
CAA77173	<i>Arabidopsis thaliana 4</i>	sink tissues	Rook et al. 2001

Table A-2. Continued.

BAA76362	<i>Arabidopsis thaliana</i> 2	leaf	Kavakli et al. 2002
AAB91468	<i>Citrullus lanatus</i> 1	fruit	Kim et al. 1998
AAB91467	<i>Citrullus lanatus</i> 2	fruit	Kim et al. 1998
AAS00542	<i>Fragaria ananassa</i>	Unknown	No publication
AAD56042	<i>Citrus unshiu</i>	fruit, leaf	Kim et al. 2001
CAA55516	<i>Beta vulgaris</i>	leaf, weak in hypoc, root	Müller-Röber et al. 1995
AAB91463	<i>Cucumis melo</i> 1	stem, root, early fruit, weak in leaf	Park et al. 1998
AAB91464	<i>Cucumis melo</i> 2	stem, root, early fruit, weak in leaf	Park et al. 1998
AAM95945	<i>Oncidium goldiana</i>	leaf, flower	Li et al. 2003
720391 (Protein ID in JGI)	<i>Populus trichocarpa</i> 1	Unknown	No publication
346791 (Protein ID in JGI)	<i>Populus trichocarpa</i> 2	Unknown	No publication
721986 (Protein ID in JGI)	<i>Populus trichocarpa</i> 3	Unknown	No publication
546089 (Protein ID in JGI)	<i>Populus trichocarpa</i> 4	Unknown	No publication
685406 (Protein ID in JGI)	<i>Populus trichocarpa</i> 5	Unknown	No publication
57133 (Protein ID in JGI)	<i>Populus trichocarpa</i> 6	Unknown	No publication
187891 (Protein ID in JGI)	<i>Chlamydomonas reinhardtii</i>	Unknown	Zabawinski et al. 2001
AAS88891	<i>Ostreococcus tauri</i>	Unknown	Ral et al. 2004
42209 (Protein ID in JGI)	<i>Ostreococcus lucimarinus</i>	Unknown	No publication

LIST OF REFERENCES

- Ainsworth, C., Hosein, F., Tarvis, M., Weir, F., Burrell, M., Devos, K. M., and Gale, M.D.** (1995). Adenosine diphosphate glucose pyrophosphorylase genes in wheat: differential expression and gene mapping. *Planta* **197**: 1-10.
- Akashi, H.** (2001). Gene expression and molecular evolution. *Curr. Opin. Genet. Dev.* **11**: 660-666.
- Akihiro, T., Mizuno, K., and Fujimura, T.** (2005). Gene Expression of ADP-glucose Pyrophosphorylase and Starch Contents in Rice Cultured Cells are Cooperatively Regulated by Sucrose and ABA. *Plant Cell Physiol.* **46**: 937-946.
- Arnold, K., Bordoli, L., Kopp, J., and Schwede, T.** (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* **22**: 195-201.
- Bae, J.M., Giroux, M.J., and Hannah, L.C.** (1990). Cloning and molecular characterization of the *brittle-2* gene of maize. *Maydica* **35**: 317-322.
- Bae, J.M., and Liu, J.R.** (1997). Molecular cloning and characterization of two novel isoforms of the small subunit of ADP-glucose pyrophosphorylase from sweetpotato. *Mol. Gen. Genet.* **254**: 179-185.
- Baldauf, S.L., Roger, A.J., Wenk-Siefert, I., and Doolittle, W.F.** (2000). A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science* **290**: 972-977.
- Ball S., Guan, H., James, M., Myers, A., Keeling, P., Mouille, G., Buleon, A., Colonna, P., and Preiss, J.** (1996). From glycogen to amylopectin: a model for the biogenesis of the plant starch granule. *Cell* **86**: 349-352.
- Ball, S.G. and Morell, M.K.** (2003). From bacterial glycogen to starch: understanding the biogenesis of the plant starch granule. *Annu. Rev. Plant Biol.* **54**: 207-233.
- Ballicora, M.A., Dubay, J.R., Devillers, C.H., and Preiss, J.** (2005). Resurrecting the ancestral enzymatic role of a modulatory subunit. *J. Biol. Chem.* **280**: 10189-10195.
- Ballicora, M.A., Iglesias, A.A., Preiss, J.** (2003). ADP-glucose pyrophosphorylase, a regulatory enzyme for bacterial glycogen synthesis. *Microbiol. Mol. Biol. Rev.* **67**: 213-225.
- Ballicora, M.A., Laughlin, M.J., Fu, Y., Okita, T.W., Barry, G.F. and Preiss, J.** (1995). Adenosine 5'-diphosphate-glucose pyrophosphorylase from potato tuber. Significance of the N-terminus of the small subunit for catalytic properties and heat stability. *Plant Physiol.* **109**: 245-251.

- Bar-Or, C., Bar-Eyal, M., Gal, T.Z., Kapulnik, Y., Czosnek, H., and Koltai, H.** (2006). Derivation of species-specific hybridization-like knowledge out of cross-species hybridization results. *BMC Genomics* **7**: 110.
- Beckles, D.M., Smith, A.M., and apRees, T.** (2001). A cytosolic ADP-glucose pyrophosphorylase is a feature of graminaceous endosperms, but not of other starch-storing organs. *Plant Physiol.* **125**: 818-827.
- Ben Ali, A., De Baere, R., Van der Auwera, G., De Wachter, R., and Van de Peer, Y.** (2001). Phylogenetic relationships among algae based on complete large-subunit rRNA sequences. *Int. J. Syst. Evol. Microbiol.* **51**: 737–749.
- Bhave, M.R., Lawrence, S., Barton, C., and Hannah, L.C.** (1990). Identification and molecular characterization of *shrunk-2* cDNA clones of maize. *Plant Cell* **2**: 581-588.
- Bhullar, S.S., and Jenner, C.F.** (1985). Differential responses to high temperatures of starch and nitrogen accumulation in the grain of four cultivars of wheat. *Aust. J. Plant Physiol.* **12**: 363-375.
- Boehlein, S.K., Sewell, A.K., Cross, J., Stewart, J.D., and Hannah, L.C.** (2005). Purification and characterization of adenosine diphosphate glucose pyrophosphorylase from maize/potato mosaics. *Plant Physiol.* **138**: 1552–1562.
- Brangeon, J., Reyss, A., and Prioul, J.L.** (1997). In situ detection of ADPglucose pyrophosphorylase expression during maize endosperm development. *Plant Physiol. Biochem.* **35**: 847-858.
- Burger, B.T., Cross, J., Shaw, J.R., Caren, J., Greene, T.W., Okita, T.W., and Hannah, L.C.** (2003). Relative turnover numbers of maize endosperm and potato tuber ADP-glucose pyrophosphorylases in the absence and presence of 3-PGA. *Planta* **217**: 449-456.
- Burgess, D., Penton, A., Dunsmuir, P. and Dooner, H.** (1997). Molecular cloning and characterization of ADP-glucose pyrophosphorylase cDNAs isolated from pea cotyledons. *Plant Mol. Biol.* **33**: 431–444.
- Burton, R.A., Johnson, P.E., Beckles, D.M., Fincher, G.B., Jenner, H.L., Naldrett, M.J., and Denyer, K.** (2002). Characterization of the genes encoding the cytosolic and plastidial forms of ADP-Glucose pyrophosphorylase in wheat endosperm. *Plant Physiol.* **130**: 1464-1475.
- Cargill, M., Altshuler, D., Ireland, J., Sklar, P., Ardlie, K., Patil, N., Shaw, N., Lane, C.R., Lim, E.P., Kalyanaraman, N., Nemesh, J., Ziaugra, L., Friedland, L., Rolfe, A., Warrington, J., Lipshutz, R., Daley, G.Q., and Lander, E.S.** (1999). Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat. Genet.* **22**: 231-238.
- Chang, J.** (1981). Corn yield in relation to photoperiod, night temperature, and solar radiation. *Agricul. Metero.* **24**: 253-262.

- Chang, Y.Y., Kakefuda, G., Iglesias, A.A., Buikema, W.J., and Preiss, J.** (1992). Molecular cloning and expression of the gene encoding ADP-glucose pyrophosphorylase from the cyanobacterium *Anabaena* sp. strain PCC 7120. *Plant Mol. Biol.* **20**: 37-47.
- Cheikn, N., and Jones, R.** (1995). Heat stress effects on sink activity of developing maize kernels grown in vitro. *Physiol. Plant.* **95**: 59-66.
- Chen, B.Y., Janes, H.W., and Gianfagna, T.** (1998). PCR cloning and characterization of multiple ADP-glucose pyrophosphorylase cDNAs from tomato. *Plant Sci.* **136**: 59-67.
- Chen, K, Durand, D, and Farach-Colton, M.** (2000). NOTUNG: A program for dating gene duplications and optimizing gene family trees. *Journal of Computational Biology* **7**: 429-447.
- Choi, H.K., Mun, J.H., Kim, D.J., Zhu, H., Baek, J.M., Mudge, J., Roe, B., Ellis, N., Doyle, J., Kiss, G.B., Young, N.D., and Cook, D.R.** (2004). Estimating genome conservation between crop and model legume species. *Proc. Natl. Acad. Sci. USA* **101**: 15289–15294.
- Choi, S.B., Kim, K.H., Kavakli, I.H., Lee, S.K., and Okita, T.W.** (2001). Transcriptional expression characteristics and subcellular localization of ADP-glucose pyrophosphorylase in the oil plant *Perilla frutescens*. *Plant and Cell Physiology* **42**: 146–153.
- Christy, A.L., and Williamson, D.R.** (1985). Characteristics of CO₂ fixation and productivity of corn and soybeans. Pages 379-387 in P.W. Luden and J.E. Burris eds. *Nitrogen Fixation and CO₂ Metabolism*. Elsevier Science Publishing Co., New York.
- Christy, A.L., Williamson, D.R., and Wideman, A.S.** (1985). Maize source development and activity. In 'Regulation of Carbon and Nitrogen Reduction and Utilization in Maize'. (Eds J.C. Shannon and C.D. Boyer.) pp 11-20. (American Society of Plant Physiologists: Rockville).
- Crepet, W.L., Nixon, K.C., and Gandolfo, M.A.** (2004). Fossil evidence and phylogeny: the age of major angiosperm clades based on mesofossil and macrofossil evidence from Cretaceous deposits. *Am. J. Botany* **91**: 1666 – 1682.
- Crevillen, P., Ballicora, M.A., Merida, A., Preiss, J., and Romero, J.M.** (2003). The different large subunit isoforms of *Arabidopsis thaliana* ADP-glucose pyrophosphorylase confer distinct kinetic and regulatory properties to the heterotetrameric enzyme. *J. Biol. Chem.* **278**: 28508-28515.
- Crevillen, P., Ventriglia, T., Pinto, F., Orea, A., Merida, A., and Romero, J.M.** (2005). Differential Pattern of Expression and Sugar Regulation of *Arabidopsis thaliana* ADP-glucose Pyrophosphorylase-encoding Genes. *J. Biol. Chem.* **280**: 8143-8149.
- Cross, J. M., Clancy, M., Shaw, J., Boehlein, S., Greene, T. Schmidt, R. Okita, T. and Hannah L.C.** (2005). A polymorphic motif in the small subunit of ADP-glucose pyrophosphorylase modulates interactions between the small and large subunits. *Plant J.* **41**, 501-511.

Cross, J. M., Clancy, M., Shaw, J., Greene, T. W., Schmidt, R. R. Okita, T. W. and Hannah, L. C. (2004). Both subunits of ADP-glucose pyrophosphorylase are regulatory. *Plant Physiol.* **135**: 137-140.

Dayhoff, M.O., Schwartz, R.M., and Orcutt, B.C. (1978). A model of evolutionary change in proteins. In *Atlas of Protein Sequence and Structure, Vol 5, Suppl. 3* (ed M.O. Dayhoff), National Biomedical Research Foundation, Silver Springs, MD pp. 345-352.

de Sa, M., and Drouin, G. (1996). Phylogeny and substitution rates of angiosperm actin genes. *Mol. Biol. Evol.* **13**: 1198-1212.

Denyer, K., Dunlap, F., Thorbjornsen, T., Keeling, P., and Smith, A. M. (1996). The major form of ADP-glucose pyrophosphorylase in maize endosperm is extra-plastidal. *Plant Physiol.* **112**: 779-785.

Duke, E., and Doehlert, D. (1996). Effects of heat stress on enzyme activities and transcript levels in developing maize kernels grown in culture. *Environ. Exp. Botany* **36**: 199-208.

Duncan, W.G., and Hesketh, J.D. (1968). Net photosynthetic rates, relative leaf growth rates, and leaf numbers of 22 races of maize grown at eight temperatures. *Crop Science* **8**: 670-674.

Duret, L., and Mouchiroud, D. (2000). Determinants of substitution rates in mammalian genes, expression pattern affects selection intensity but not mutation rate. *Mol. Biol. Evol.* **17**: 68-74.

Eimert, K., Luo, C., Dejardin, A., Villand, P., Thorbjornsen, T., and Kleczkowski, L.A. (1997). Molecular cloning and expression of the large subunit of ADP-glucose pyrophosphorylase from barley (*Hordeum vulgare*) leaves. *Gene* **189**: 79-82.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol.* **17**: 368-376.

Felsenstein, J. (1985). Confidence-limits on phylogenies: an approach using the bootstrap. *Evolution* **39**: 783-791.

Fraser, H.B., Hirsh, A.E., Steinmetz, L.M., Scharfe, C., and Feldman, M.W. (2002). Evolutionary rate in the protein interaction network. *Science* **296**: 750-752.

Frueauf, J.B., Ballicora, M.A., and Preiss, J. (2003). ADP-glucose pyrophosphorylase from potato tuber, site-directed mutagenesis of homologous aspartic acid residues in the small and large subunits. *Plant J.* **33**: 503-511.

Fu, Y., Ballicora, M.A., Leykam, J.F., and Preiss, J. (1998). Mechanism of reductive activation of potato tuber ADP-glucose pyrophosphorylase. *J. Biol. Chem.* **273**: 25045-25052.

- Fu, Y., Ballicora, M.A., and Preiss, J.** (1998). Mutagenesis of the glucose-1-phosphate-binding site of potato tuber ADP-glucose pyrophosphorylase. *Plant Physiol.* **117**: 989–996.
- Gaut, B.S.** (2002). Evolutionary dynamics of grass genomes. *New phytologist* **154**: 15-28.
- Georgelis, N., Braun, E.L., Shaw, J.R., and Hannah, L.C.** (2007). The two AGPase subunits evolve at different rates in angiosperms, yet they are equally sensitive to activity altering amino acid changes when expressed in bacteria, *Plant Cell* **19**: 1458–1472.
- Giroux, M.J., and Hannah, L.C.** (1994). ADP-glucose pyrophosphorylase in *shrunk-2* and *brittle-2* mutants of maize. *Mol. Gen. Genet.* **243**: 400-408.
- Giroux, M.J., Shaw, J., Barry, G., Cobb, B.G., Greene, T.W., Okita, T.W., and Hannah, L.C.** (1996). A single mutation that increases maize seed weight. *Proc. Natl. Acad. Sci. USA* **93**: 5824-5829.
- Giroux, M., Smith-White, B., Gilmore, V., Hannah, L.C., and Preiss, J.** (1995). The large subunit of the embryo isoform of ADP glucose pyrophosphorylase from maize. *Plant Physiol.* **108**: 1333-1334.
- Goldman, N.** (1993). Statistical tests of models of DNA substitution. *J. Mol. Evol.* **36**: 182–198.
- Goodman, M., Czelusniak, J., Moore, G.W., Romero-Herrera, A.E., and Matsuda, G.** (1979). Fitting the gene lineage into its species lineage: a parsimony strategy illustrated by cladograms constructed from globin sequences. *Syst. Zool.* **28**: 132-168.
- Govons, S., Vinopal, R., Ingraham, J., and Preiss J.** (1969). Isolation of mutants of *Escherichia coli* B altered in their ability to synthesize glycogen. *J. Bacteriol.* **97**: 970-972.
- Grant, D., Cregan, P., and Shoemaker, R.C.** (2000). Genome organization in dicots: genome duplication in *Arabidopsis* and synteny between soybean and *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **97**: 4168–4173.
- Greene, T.W., Chantler, S.E., Kahn, M.L., Barry, G.F., Preiss, J., and Okita, T.W.** (1996). Mutagenesis of the potato ADPglucose pyrophosphorylase and characterization of an allosteric mutant defective in 3-phosphoglycerate activation. *Proc. Natl. Acad. Sci. USA* **93**: 1509-1513.
- Greene, T.W., and Hannah, L.C.** (1998a). Enhanced stability of maize endosperm ADP-glucose pyrophosphorylase is gained through mutants that alter subunit interactions. *Proc. Natl. Acad. Sci. USA* **95**: 13342-13347.
- Greene, T.W., and Hannah, L.C.** (1998b). Assembly of maize endosperm ADP-glucose pyrophosphorylase requires motifs located throughout the large and small subunit units. *Plant Cell* **10**: 1295-1306.

- Gu, X., Fu, Y.X., and Li, W.H.** (1995). Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. *Mol. Biol. Evol.* **12**: 546-557.
- Guan, H.P., and Preiss, J.** (1993). Differentiation of the Properties of the Branching Isozymes from Maize (*Zea mays*). *Plant Physiol.* **102**: 1269-1273.
- Guex, N., and Peitsch, M.C.** (1997). SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modelling. *Electrophoresis* **18**: 2714-2723.
- Guindon, S., and Gascuel, O.** (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**: 696-704.
- Guo, H.H., Choe, J., and Loeb, L.A.** (2004). Protein tolerance to random amino acid change. *Proc. Natl. Acad. Sci. USA* **101**: 9205-9210.
- Hall, D.B., and Struhl, K.** (2002). The VP16 activation domain interacts with multiple transcriptional components as determined by protein-protein cross-linking in vivo. *J. Biol. Chem.* **277**: 46043-46050.
- Hall, T.A.** (1999). BioEdit, a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp. Ser.* **41**: 95-98.
- Hannah, L.C.** (2000). Starch Biosynthesis and Genetic Potential. In: "Designing Crops for Added Value" pp. 181-200. Ed. by C. F. Murphy and D. M. Peterson, American Society of Agronomy, Inc. Madison Wisconsin, USA.
- Hannah, L.C.** (2005). Starch synthesis in the maize endosperm. *Maydica* **50**: 497-506.
- Hannah, L.C., and Greene, T.W.** (1998). Maize seed weight is dependent on the amount of endosperm ADP-glucose pyrophosphorylase. *J. Plant Physiol.* **152**: 649-652.
- Hannah, L.C., and Nelson, O.E., Jr.** (1976). Characterization of ADP-glucose pyrophosphorylase from *shrunk-2* and *brittle-2* mutants of maize. *Biochem. Genet.* **14**: 547-560.
- Hannah, L.C., Shaw, J.R., Giroux, M.J., Reys, A., Prioul, J.L., Bae, J.M., and Lee, J.Y.** (2001). Maize genes encoding the small subunit of ADP-glucose pyrophosphorylase. *Plant Physiol.* **127**: 173-183.
- Hannah, L.C., Tuschall, D., and Mans, R.** (1980). Multiple forms of maize endosperm ADP-glucose pyrophosphorylase and their control by *Shrunk-2* and *Brittle-2*. *Genetics* **95**: 961-970.
- Hanson, R.S., and Phillips, J.A.** (1981). Chemical composition, p. 328-364. In P. Gerhardt, *et al.* (ed.), *Manual of methods for general bacteriology*. American Society for Microbiology, Washington, D.C.

- Harn, C.H., Bae, J.M., Lee, S.S., Min, S.R., and Liu, J.R.** (2000). Presence of multiple cDNAs encoding an isoform of ADP-glucose pyrophosphorylase large subunit from sweet potato and characterization of expression levels. *Plant Cell Physiol.* **41**: 1235-1242.
- Hastings, K.E.M.** (1996). Strong evolutionary conservation of broadly expressed protein isoforms in the troponin I gene family and other vertebrate gene families. *J. Mol. Evol.* **42**: 631–640.
- He, X., and Zhang, J.** (2006). Higher duplicability of less important genes in yeast genomes. *Mol. Biol. Evol.* **23**: 144-151.
- Henikoff, S., and Henikoff, J.G.** (1992). Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. USA* **89**: 10915-10919.
- Hileman, L.C., and Baum, D.A.** (2003). Why do paralogs persist? Molecular evolution of *CYCLOIDEA* and related floral symmetry genes in Antirrhineae (Veronicaceae). *Mol. Biol. Evol.* **20**: 591–600.
- Hofstra, G., and Hesketh, J.D.** (1969). Effects of temperature on the gas exchange of leaves in the light and dark. *Planta* **85**: 228-237.
- Huelsenbeck, J.P., and Rannala, B.** (1997). Phylogenetic methods come of age: Testing hypotheses in an evolutionary context. *Science* **276**: 227–232.
- Hunter, R., Tollenaar, M., and Breuer, C.** (1977). Effects of photoperiod and temperature on vegetative and reproductive growth of maize (*Zea mays*) hybrid. *Can. J. Plant Sci.* **57**: 1127-1133.
- Hwang, S.K., Hamada, S., and Okita, T.W.** (2006). ATP binding site in the plant ADP-glucose pyrophosphorylase large subunit. *FEBS Lett.* **580**: 6741-6748.
- Hwang, S.K., Hamada, S., and Okita, T.W.** (2007). Catalytic implications of the higher plant ADP-glucose pyrophosphorylase large subunit. *Phytochemistry* **68**: 464-477.
- Iglesias, A., Barry, G.F., Meyer, C., Bloksberg, L., Nakata, P., Greene, T., Laughlin M.J., Okita T.W., Kishore G.M., and Preiss, J.** (1993). Expression of the potato tuber ADP-glucose pyrophosphorylase in *Escherichia coli*. *J. Biol. Chem.* **268**: 1081-1086.
- James, M.G., Denyer, K., and Myers, A.** (2003). Starch synthesis in the cereal endosperm. *Curr. Opin. Plant Biol.* **6**: 215-222.
- Jin, X., Ballicora, M.A., Preiss, J., and Geiger, J.H.** (2005). Crystal structure of potato tuber ADP-glucose pyrophosphorylase. *EMBO J.* **24**: 694-704.
- Jobling, S.** (2004). Improving starch for food and industrial applications. *Curr. Opin. Plant Biol.* **7**: 210-218.

- Johnson, P.E., Patron, N.J., Bottrill, A.R., Dinges, J.R., Fahy, B.F., Parker, M.L., Waite, D.N., and Denyer, K.** (2003). A low-starch barley mutant, riso 16, lacking the cytosolic small subunit of ADP-glucose pyrophosphorylase, reveals the importance of the cytosolic isoform and the identity of the plastidial small subunit. *Plant Physiol.* **131**: 684–696.
- Jones, R., Ouattar, S., and Crookston, R.** (1984). Thermal environment during endosperm cell division and grain filling in maize: effects on kernel growth and development in vitro. *Crop Science* **24**: 133-137.
- Jordan, I.K., Wolf, Y.I., and Koonin, E.V.** (2003). No simple dependence between protein evolution rate and the number of protein-protein interactions: Only the most prolific interactors tend to evolve slowly. *BMC Evol. Biol.* **3**: 1.
- Kavakli, I.H., Greene, T.W., Salamone, P.R., Choi, S.B., and Okita, T.W.** (2001). Investigation of subunit function in ADP-glucose pyrophosphorylase. *Biochem. Biophys. Res. Commun.* **281**: 783–787.
- Kavakli, I.H., Kato, C., Choi, S.B., Kim, K.H., Salamone, P.R., Ito, H., and Okita, T.W.** (2002). Generation, characterization, and heterologous expression of wild-type and up-regulated forms of *Arabidopsis thaliana* leaf ADP-glucose pyrophosphorylase. *Planta* **215**: 430-439.
- Kim, I.J., Kahng, H.Y., and Chung, W.I.** (1998). Characterization of cDNAs encoding small and large subunits of ADP-glucose pyrophosphorylases from watermelon (*Citrullus vulgaris* S.). *Biosci. Biotechnol. Biochem.* **62**: 550-555.
- Kim, I.J., Noh, S.J., Lee, B.H., Jo, J., Kim, Y.S., and Chung, W.I.** (2001). Molecular characterization of cDNA clones for ADP-glucose pyrophosphorylase from Citrus. *Biochim. Biophys. Acta* **1518**: 324-328.
- Kirchberger, S., Leroch, M., Huynen, M.A., Wahl, M., Neuhaus, H.E., and Tjaden, J.** (2007). Molecular and biochemical analysis of the plastidic ADP-glucose transporter (ZmBT1) from *Zea mays*. *J. Biol. Chem.* **282**: 22481-22491.
- Kopp, J., and Schwede, T.** (2004) The SWISS-MODEL Repository of annotated three-dimensional protein structure homology models *Nucleic Acids Research* **32**: D230-D234.
- Ku, H.M., Vision, T., Liu, J., and Tanksley, S.D.** (2000). Comparing sequenced segments of the tomato and *Arabidopsis* genomes: Large-scale duplication followed by selective gene loss creates a network of synteny. *Proc. Natl. Acad. Sci. USA* **97**: 9121-9126.
- La Cognata, U., Willmitzer, L., and Müller-Röber, B.** (1995). Molecular cloning and characterization of novel isoforms of potato ADP-glucose pyrophosphorylase. *Mol. Gen. Genet.* **246**: 538–548.

- Laemmli, U.K.** (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680-685.
- Laughlin, M.J., Payne, J.W., and Okita, T.W.** (1998). Substrate binding mutants of the higher plant ADP-glucose pyrophosphorylase. *Phytochemistry* **47**: 621–629.
- Li, C.R., Zhang, X.B., and Hew, C.S.** (2003). Molecular cloning of ADP-glucose pyrophosphorylase large subunit cDNA from *Oncidium*. *Biol. Plantarum* **47**: 613-615.
- Lin, T.P., Caspar, T., Somerville, C.R., and Preiss, J.** (1988a). Isolation and Characterization of a Starchless Mutant of *Arabidopsis thaliana* (L.) Heynh Lacking ADPglucose Pyrophosphorylase Activity. *Plant Physiol.* **86**: 1131-1135.
- Lin, T.P., Caspar, T., Somerville, C.R., and Preiss, J.** (1988b). A Starch Deficient Mutant of *Arabidopsis thaliana* with Low ADPglucose Pyrophosphorylase Activity Lacks One of the Two Subunits of the Enzyme. *Plant Physiol.* **88**: 1175-1181.
- Lin, Y.S., Hwang, J.K., and Li, W.H.** (2007). Protein complexity, gene duplicability and gene dispensability in the yeast genome. *Gene* **387**: 109-117.
- Lynch, M., and Conery, J.S.** (2000). The evolutionary fate and consequences of duplicate genes. *Science* **290**: 1151–1155.
- Marin, I., Fares, M.A., Gonzalez-Candelas, F., Barrio, E., and Moya, A.** (2001). Detecting changes in the functional constraints of paralogous genes. *J. Mol. Evol.* **52**: 17-28.
- Meyer, F.D., Smidansky, E.D., Beecher, B., Greene, T.W., and Giroux, M.J.** (2004). The maize Sh2r6hs ADP-glucose pyrophosphorylase (AGP) large subunit confers enhanced AGP properties in transgenic wheat (*Triticum aestivum*). *Plant Sci.* **167**: 899–911.
- Mizuno, K., Kobayashi, E., Tachibana, M., Kawasaki, T., Fujimura, T., Funane, K., Kobayashi, M., and Baba, T.** (2001). Characterization of an isoform of rice starch branching enzyme, RBE4, in developing seeds. *Plant Cell Physiol.* **42**: 349-357.
- Müller-Rober, B., Kossmann, J., Hannah, L.C., Willmitzer, L., and Sonnewald, U.** (1990). Only one of two different ADPglucose pyrophosphorylase genes from potato responds strongly to elevated levels of sucrose. *Molecular and General Genetics* **224**: 136–146.
- Myers, A.M., Morell, M.K., James, M.G., and Ball, S.G.** (2000). Recent progress toward understanding biosynthesis of the amylopectin crystal. *Plant Physiol.* **122**: 989-997.
- Nakamura, Y., Gojobori, T., and Ikemura, T.** (2000). Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res.* **28**: 292.
- Nakata, P.A., Anderson, J.M., and Okita, T.W.** (1994). Structure and expression of the potato ADP-glucose pyrophosphorylase small subunit. *J. Biol. Chem.* **269**: 30798-30807.

- Nakata, P.A., and Okita, T.W.** (1995). Differential regulation of ADP-glucose pyrophosphorylase in the sink and source tissues of potato. *Plant Physiol.* **108**: 361-368.
- Nakata, P.A., and Okita, T.W.** (1996). *Cis*-elements important for the expression of the ADP-glucose pyrophosphorylase small-subunit are located both upstream and downstream from its structural gene. *Mol. Gen. Genet.* **250**: 581-592.
- Nelson, O.E., and Rines, H.W.** (1962). The enzymatic deficiency in the waxy mutant of maize. *Biochem. Biophys. Res. Commun.* **9**: 297-300.
- Ohdan, T., Francisco, P.B. Jr, Sawada, T., Hirose, T., Terao, T., Satoh, H., and Nakamura, Y.** (2005). Expression profiling of genes involved in starch synthesis in sink and source organs of rice. *J. Exp. Bot.* **56**: 3229-3244.
- Omoto, D., Ito, H., Obana, Y., Matsumoto, K., Isono, N., Hamada, S., Okita, T.W., and Matsui, H.** (2003). Isolation and characterization of two cDNAs for large and small subunits of ADP-glucose pyrophosphorylase. *J. Appl. Glycosci.* **50**: 475-479.
- Page, R.D.M.** (1998). GeneTree: comparing gene and species phylogenies using reconciled trees. *Bioinformatics* **14**: 819-820.
- Page, R.D.M. and Cotton, J.A.** (2000). GeneTree: a tool for exploring gene family evolution. In D. Sankoff, and J. Nadeau, (eds.), *Comparative Genomics: Empirical and Analytical Approaches to Gene Order Dynamics, Map Alignment, and the Evolution of Gene Families*. Kluwer Academic Press, Dordrecht, pp. 525-536.
- Pan, D., and Nelson, O.E.** (1984). A debranching enzyme deficiency in endosperms of the *sugary-1* mutants of maize. *Plant Physiol.* **74**: 324-328.
- Park, S.W., and Chung, W.I.** (1998). Molecular cloning and organ-specific expression of three isoforms of tomato ADP-glucose pyrophosphorylase gene. *Gene* **206**: 215-221.
- Park, S.W., Kahng, H.Y., Kim, I.J., Park, J.O., and Chung, W.I.** (1998). Molecular cloning and characterization of small and large subunits of ADP-glucose pyrophosphorylase from oriental melon. *J. Plant Res.* **111**: 59-63.
- Peitsch, M.C.** (1995). Protein modeling by E-mail Bio/Technology **13**: 658-660.
- Peters, D.B., Pendleton, J.W., Hageman, R.H., and Brown, C.M.** (1971). Effect of night air temperature on grain yield of corn, wheat and soybeans. *Agron. J.* **63**: 809.
- Preiss, J., and Romero, T.** (1989). Physiology, biochemistry and genetics of bacterial glycogenesis. *Adv. Microbiol. Phys.* **30**: 183-238.

Prioul, J.L., Jeannette, E., Reyss, A., Gregory, N., Giroux, M., Hannah, L.C., and Causse, M. (1994). Expression of ADP-glucose pyrophosphorylase in maize (*Zea mays* L.) grain and source leaf during grain filling. *Plant Physiol.* **104**: 179-187.

Rabinowicz, P.D., Braun, E.L., Wolfe, A.D., Bowen, B., and Grotewold, E. (1999). Maize R2R3 Myb genes: Sequence analysis reveals amplification in the higher plants. *Genetics* **153**: 427-444.

Ral, J.P., Derelle, E., Ferraz, C., Wattebled, F., Farinas, B., Corellou, F., Buléon, A., Slomianny, M.C., Delvalle, D., d'Hulst, C., Rombauts, S., Moreau, H., and Ball, S. (2004). Starch division and partitioning. A mechanism for granule propagation and maintenance in the picophytoplanktonic green alga *Ostreococcus tauri*. *Plant Physiol.* **136**: 3333–3340.

Rao, S., and Rossmann, M. (1973). Comparison of super-secondary structures in proteins. *J. Mol. Biol.* **76**: 241-256.

Rohila, J.S., Chen, M., Cerny, R., and Fromm, M.E. (2004). Improved tandem affinity purification tag and methods for isolation of protein heterocomplexes from plants. *Plant J.* **38**:172-181.

Rook, F., Corke, F., Card, R., Munz, G., Smith, C., and Bevan, M.W. (2001). Impaired sucrose-induction mutants reveal the modulation of sugar-induced starch biosynthetic gene expression by abscisic acid signalling. *Plant J.* **26**: 421-433.

Rosti, S., and Denyer, K. (2007). Two paralogous genes encoding small subunits of ADP-glucose pyrophosphorylase in maize, *Bt2* and *L2*, replace the single alternatively spliced gene found in other cereal species. *J. Mol. Evol.* PMID: 17846820.

Rosti, S., Rudi, H., Rudi, K., Opsahl-Sorteberg, H.G., Fahy, B., and Denyer, K. (2006). The gene encoding the cytosolic small subunit of ADP-glucose pyrophosphorylase in barley endosperm also encodes the major plastidial small subunit in the leaves. *J. Exp. Bot.* **57**: 3619–3626.

Schaffer, A.A., Levin, I., Oguz, I., Petreikov, M., Cincarevsky, F., Yeselson, Y., Shen, S., Gilboa, N., and Bar, M. (2000). ADP-glucose pyrophosphorylase activity and starch accumulation in immature tomato fruit: the effect of a *Lycopersicon hirsutum*-derived introgression encoding for the large subunit. *Plant Sci.* **152**: 135-144.

Schmidt, T.R., Jaradat, S.A., Goodman, M., Lomax, M.I., and Grossman, L.I. (1997). Molecular evolution of cytochrome c oxidase: rate variation among subunit VIa isoforms. *Mol. Biol. Evol.* **14**: 595-601.

Schlueter, J.A., Dixon, P., Granger, C., Grant, D., Clark, L., Doyle, J.J., and Shoemaker, R.C. (2004). Mining EST databases to resolve evolutionary events in major crop species. *Genome* **47**: 868-876.

- Schwede, T., Kopp, J., Guex, N., and Peitsch, M.C.** (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* **31**: 3381-3385.
- Sehnke, P.C., Chung, H.J., Wu, K., and Ferl, R.J.** (2001). Regulation of starch accumulation by granule associated plant 14-3-3 proteins. *Proc. Natl. Acad. Sci. USA* **98**: 765-770.
- Shannon, J.C., Pien, F.M., Cao, H., and Liu, K.C.** (1998). *Brittle-1*, an adenylate translocator, facilitates transfer of extraplasmidial synthesized ADP-Glucose into amyloplasts of maize endosperms. *Plant Physiol.* **117**: 1235–1252.
- Shannon J.C., Pien, F.M., and Liu, K.C.** (1996). Nucleotides and Nucleotide Sugars in Developing Maize Endosperms: Synthesis of ADP-Glucose in *brittle-1*. *Plant Physiol.* **110**: 835-843.
- Singh, S., Choi, S.B., Modi, M.K., and Okita, T.W.** (2002). Isolation and characterization of cDNA clones encoding ADP-glucose pyrophosphorylase large and small subunits from chickpea (*Cicer arietinum* L.). *Phytochem.* **59**: 261-268.
- Sikka, V.K., Choi, S.B., Kavakli, I.H., Sakulsingharoj, C., Gupta, S., Ito, H., and Okita, T.W.** (2001). Subcellular compartmentation and allosteric regulation of the rice endosperm ADPglucose pyrophosphorylase. *Plant Sci.* **161**: 461-468.
- Singletary, G., Banisadr, R., and Keeling, P.** (1993). Decreased starch synthesis in heat stressed maize kernels results from reduced ADPG-pyrophosphorylase and starch synthase activities. *Plant Physiol. Suppl.* **102**: 6.
- Singletary, G., Banisadr, R., and Keeling, P.** (1994). Heat stress during grain filling in maize: effects of carbohydrate storage and metabolism. *Aust. J. Plant Physiol.* **21**: 829-841.
- Singletary, G.W., Banisadr, R., and Keeling, P.L.** (1997). Influence of Gene Dosage on Carbohydrate Synthesis and Enzymatic Activities in Endosperm of Starch-Deficient Mutants of Maize. *Plant Physiol.* **113**: 293-304.
- Stark, D., Timmerman, K., Barry, G., Preiss, J., and Kishore, G.** (1992). Regulation of the amount of starch in plant tissues by ADP glucose pyrophosphorylase. *Science* **258**: 287-291.
- Smidansky, E.D., Clancy, M., Meyer, F.D., Lanning, S.P., Blake, N.K., Talbert, L.E., and Giroux, M.J.** (2002). Enhanced ADP-glucose pyrophosphorylase activity in wheat endosperm increases seed yield. *Proc. Natl. Acad. Sci.* **99**: 1724-1729.
- Smidansky, E.D., Martin, J.M., Hannah, L.C., Fischer, A.M., and Giroux, M.J.** (2003). Seed yield and plant biomass increases in rice are conferred by deregulation of endosperm ADP-glucose pyrophosphorylase. *Planta* **216**: 656-664.
- Smith-White, B.J., and Preiss, J.** (1992). Comparison of proteins of ADP-glucose pyrophosphorylase from diverse sources. *J. Mol. Evol.* **34**: 449-464.

- Soltis, P., Soltis, D., Edwards, C.** (2005). Angiosperms, Flowering Plants, Version 03 June 2005. <http://tolweb.org/Angiosperms/20646/2005.06.03> in The Tree of Life Web Project, <http://tolweb.org/>
- Takeda, Y., Guan, H.P., and Preiss, J.** (1993). Branching of amylose by the branching isoenzymes of maize endosperm. *Carbohydr. Res.* **240**: 253–263.
- Tetlow, I.J., Morell, M.K., and Emes, M.J.** (2004). Recent developments in understanding the regulation of starch metabolism in higher plants. *J. Exp. Bot.* **55**: 2131-2145.
- Thompson, L.** (1975). Weather variability, climatic change and grain production. *Science* **188**: 535-541.
- Thorbjørnsen, T., Villand, P., Denyer, K., Olsen, O., and Smith, A.M.** (1996). Distinct isoforms of ADPglucose pyrophosphorylase occur inside and outside the amyloplast in barley endosperm. *Plant J.* **10**: 243–250.
- Tiessen, A., Hendriks, J.H.M., Stitt, M., Branscheid, A., Gibon, Y., Farre, E.M., and Geigenberger, P.** (2002). Starch synthesis in potato tubers is regulated by post-translational redox modification of ADP-glucose pyrophosphorylase: A novel regulatory mechanism linking starch synthesis to the sucrose supply. *Plant Cell* **14**: 2191 – 2213.
- Tollenaar, M., and Bruulsema, T.** (1988). Effects of temperature on rate and duration of kernel dry matter accumulation of maize. *Can. J. Plant Sci.* **68**: 935-940.
- Tsai, C.Y., and Nelson, O.E.** (1966). Starch deficient maize mutants lacking adenosine diphosphate glucose pyrophosphorylase activity. *Science* **151**: 341-343.
- Tuskan, G.A., et al.** (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**: 1596-1604.
- Van de Peer, Y., Rensing, S.A., Maier, U.G., and De Wachter, R.** (1996). Substitution rate calibration of small subunit ribosomal RNA identifies chlorarachniophyte endosymbionts as remnants of green algae. *Proc. Natl. Acad. Sci. USA* **93**: 7732–7736.
- Vigeolas, H., Mohlmann, T., Martini, N., Neuhaus, H.E., and Geigenberger, P.** (2004). Embryo-specific reduction of ADP-Glc pyrophosphorylase leads to an inhibition of starch synthesis and a delay in oil accumulation in developing seeds of oilseed rape. *Plant Physiol.* **136**: 2676-2686.
- Villand, P., Aalen, R., Olsen, O. A., Luthi, E., Lonneborg, A., and Kleczkowski, L.A.** (1992). PCR amplification and sequences of cDNA clones for the small and large subunits of ADP-glucose pyrophosphorylase from barley tissues. *Plant Mol. Biol.* **19**: 381-389.

- Villand, P., Olsen, O.A., and Kleczkowski, L.A.** (1993). Molecular characterization of multiple cDNA clones for ADP-glucose pyrophosphorylase from *Arabidopsis thaliana*. *Plant Mol. Biol.* **23**: 1279-1284.
- Wallwork, M.A.B., Logue, S.J., MacLeod, L.C., and Jenner, C.F.** (1998) Effect of high temperature during grain filling on starch synthesis in the developing barley grain. *Aust. J. Plant Physiol.* **25**: 173-181.
- Wang, S.M., Lue, W.L., Yu, T.S., Long, J.H., Wang, C.N., Eimert, K., and Chen, J.** (1998). *adg2-1* represents a missense mutation in the ADPG pyrophosphorylase large subunit gene of *Arabidopsis thaliana*. *Plant J.* **13**: 63-70.
- Weber, H., Heim, U., Borisjuk, L., and Wobus, U.** (1995). Cell-type specific, co-ordinate expression of two ADPglucose pyrophosphorylase genes in relation to starch biosynthesis during seed development in *Vicia faba* L. *Planta* **195**: 352–361.
- Wikstrom, N., Savolainen, V., and Chase, M.W.** (2001). Evolution of the angiosperms: calibrating the family tree. *Proc. Biol. Sci.* **268**: 2211-2220.
- Wilhelm, E., Mullen, R., Keeling, P., and Singletary, G.** (1999). Heat stress during grain filling in maize: Effects on kernel growth and metabolism. *Crop Science* **39**: 1733-1741.
- Wolf, Y.I., Carmel, L., and Koonin, E.V.** (2006). Unifying measures of gene function and evolution. *Proc. R. Soc. Lond. B Biol. Sci.* **273**: 1507–1515.
- Wright, S.I., Yau, C.B., Looseley, M., and Meyers, B.C.** (2004). Effects of gene expression on molecular evolution in *Arabidopsis thaliana* and *Arabidopsis lyrata*. *Mol. Biol. Evol.* **21**: 1719-1726.
- Yang, T.J., Kim, J.S., Kwon, S.J., Lim, K.B., Choi, B.S., Kim, J.A., Jin, M., Park, J.Y., Lim, M.H., Kim, H.I., Lim, Y.P., Kang, J.J., Hong, J.H., Kim, C.B., Bhak, J., Bancroft, I., and Park, B.S.** (2006). Sequence-Level Analysis of the Diploidization Process in the Triplicated FLOWERING LOCUS C Region of *Brassica rapa*. *Plant Cell* **18**: 1339-1347.
- Yang, Z.** (1994). Estimating the pattern of nucleotide substitution. *J. Mol. Evol.* **39**: 105-111.
- Yang, Z.** (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. *Computer Applications in BioSciences* **13**: 555-556.
- Yang, Z., Goldman, N., and Friday, A.** (1995). Maximum likelihood trees from DNA sequences: a peculiar statistical estimation problem. *Syst. Biol.* **44**: 384–399.
- Yao Y., Thompson, D., and Gultinan, M.** (2004). Maize starch-branching enzyme isoforms and amylopectin structure. In the absence of starch-branching enzyme IIb, the further absence of starch-branching enzyme Ia leads to increased branching. *Plant Physiol.* **136**: 3515–3523.

Zabawinski, C., Van den Koornhuise, N., D'Hulst, C., Schlichting, R., Giersch, C., Delrue, B., Lacroix, J.M., Preiss, J., and Ball, S. (2001). Starchless mutants of *Chlamydomonas reinhardtii* lack the small subunit of a heterotetrameric ADP-glucose pyrophosphorylase. *J. Bacteriol.* **183**: 1069–1077.

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

Zwickl, D.J. (2006). Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The University of Texas at Austin.

BIOGRAPHICAL SKETCH

I was born in 1977, to Panayiotis and Kathryn. I grew up in the Greek island of Lesbos which I left in 1995 to pursue undergraduate studies at Aristotle University in Thessaloniki. I graduated in 2000 with a major in horticultural sciences. Then, I pursued an M.S. degree in the Department of Horticultural Sciences at the University of Florida. In 2002, I completed my degree, under the supervision of Dr. Jay W. Scott, specializing in molecular breeding, That same year I joined the Ph.D. program of Plant Molecular and Cellular Biology at the University of Florida, under the supervision of Dr. L. Curtis Hannah. Ever since, I have been working in the area of biochemical genetics.