Maize Genes Encoding the Small Subunit of ADP-Glucose Pyrophosphorylase[1]

L. Curtis Hannah*, Janine R. Shaw, Michael J. Giroux[2], Agnes Reyss, Jean-Louis Prioul, Jung-Myung Bae[3], and Jung-Youn Lee[4]

Program in Plant Molecular and Cellular Biology, Horticultural Sciences, University of Florida, P.O. Box 110690, 2211 Fifield Hall, Gainesville, Florida, 32611 (L.C.H., J.R.S., M.J.G., J.-M.B., J.-Y.L.); and Institut de Biotechnologie des Plantes, UMR8618, Associé au Centre National de la Recherche Scientifique, Structure et Métabolisme des Plantes, Bât. 630, Université de Paris-Sud, 91405 Orsay cedex, France (A.R., J.-L.P.)

Plant ADP-glucose pyrophosphorylase (AGP) is a heterotetrameric enzyme composed of two large and two small subunits. Here, we report the structures of the maize (*Zea mays*) genes encoding AGP small subunits of leaf and endosperm. Excluding exon 1, protein-encoding sequences of the two genes are nearly identical. Exon 1 coding sequences, however, possess no similarity. Introns are placed in identical positions and exhibit obvious sequence similarity. Size differences are primarily due to insertions and duplications, hallmarks of transposable element visitation. Comparison of the maize genes with other plant AGP small subunit genes leads to a number of noteworthy inferences concerning the evolution of these genes. The small subunit gene can be divided into two modules. One module, encompassing all coding information except that derived from exon 1, displays striking similarity among all genes. It is surprising that members from eudicots form one group, whereas those from cereals form a second group. This implies that the duplications giving rise to family members occurred at least twice and after the separation of eudicots and monocot cereals. One intron within this module may have had a transposon origin. A different evolutionary history is suggested for exon 1. These sequences define three distinct groups, two of which come from cereal seeds. This distinction likely has functional significance because cereal endosperm AGPs are cytosolic, whereas all other forms appear to be plastid localized. Finally, whereas barley (*Hordeum vulgare*) reportedly employs only one gene to encode the small subunit of the seed and leaf, maize utilizes the two genes described here.

The formation of ADP-Glc from Glc-1-phosphate and ATP with the release of pyrophosphate is considered the first committed step in the starch biosynthetic pathway. This key, rate-limiting reaction is catalyzed by the enzyme ADP-Glc pyrophosphorylase (AGP; EC 2.7.7.27). AGP isoforms occur in organisms and tissues ranging from *Escherichia coli* to the potato (*Solanum tuberosum*) tuber and the maize (*Zea mays*) endosperm (for review, see Preiss and Romeo, 1994; Preiss and Sivak, 1996; Hannah, 1997). Conventional mutant analysis and antisense studies have highlighted the important role AGP plays in starch synthesis both in photosynthetic and nonphotosynthetic plant tissues (Tsai and Nelson, 1966; Hannah and Nelson, 1976; Lin et al., 1988a, 1988b; Smith et al., 1989; Hylton and Smith, 1992; Muller-Rober et al., 1992; Singletary et al., 1997).

Cloning and sequencing of *Sh2* and *Bt2* and analysis of informative mutants (Bae et al., 1990, Bhave et al., 1990) of maize endosperm transcripts and genes showed that both loci encode subunits of AGP. *Bt2* encodes the small subunit, whereas the large subunit is derived from *Sh2*. The two proteins are similar to each other and also to the subunit of the homotetrameric *E. coli* AGP. This suggests that the two complementary plant genes arose from an early gene duplication event (Bhave et al., 1990). Although similar in sequence, the subunits are not interchangeable. Active maize endosperm AGP formation requires both subunits (for review, see Hannah 1997).

Sequence analysis of AGP structural genes from other plants confirmed the universal nature of heterotetrameric plant AGPs. Plant AGPs are composed of subunits of dissimilar size and genetic origin (for review, see Smith-White and Preiss, 1992), whereas a somewhat active enzyme containing only the small subunit can be formed when the small subunit gene of the potato tuber or barley (*Hordeum vulgare*) endosperm is expressed in *E. coli* or in insect cells (Ballicora et al., 1995; Doan et al., 1999) or following mutagenesis of the potato small subunit (Greene et

[2] Present address: P.O. Box 173150, Department of Plant Sciences and Plant Pathology, Montana State University, Bozeman, MT 59717–3150.

[3] Present address: Graduate School of Biotechnology, Korea University, 5–1, Anam-dong, Sungbuk-ku, Seoul 136–701, South Korea.

[4] Present address: One Shields Avenue, Life Sciences Additions, Section of Plant Biology, University of California, Davis, CA 95616.

* Corresponding author; e-mail Hannah@gnv.ifas.ufl.edu; fax 352–392–5653.

al., 1998). Conditions required for homotetrameric activity, however, are likely non-physiological.

Following the duplication forming what are now complementary AGP structural genes, other duplications likely occurred giving rise to the genes expressed in the various plant tissues. For example, early enzymological and genetic studies pointed to expression of different batteries of AGP structural genes in the maize endosperm, embryo, and leaf (Preiss et al., 1971; Hannah and Nelson, 1976; Fuchs, 1977; Bryce and Nelson, 1979). The tissue-specific nature of AGP expression was confirmed by subsequent molecular investigations (Giroux and Hannah, 1994; Prioul et al., 1994; Giroux et al., 1995). These latter studies also detected a second set of AGP structural genes expressed in the endosperm. This second endosperm AGP is expressed at low levels and definitive proof for its presence came only through analysis of null *sh2* and *bt2* mutants (Giroux and Hannah, 1994). As in the case of maize, multiple forms of AGP have been found in many if not all plants examined in sufficient detail; for example, wheat (*Triticum aestivum*) (Olive et al., 1989), potato (Okita et al., 1990; La Cognata et al., 1995), tomato (*Lycopersicon esculentum*; Chen and Janes, 1997), barley (Villand et al., 1992), rice (*Oryza sativa*; Nakamura and Kawaguchi, 1992), Arabidopsis (Villand et al., 1993), *Vicia faba* (Weber et al., 1995), pea (*Pisum sativum*; Burgess et al., 1997), and sweet potato (*Ipomoea batatas*; Bae and Liu, 1997).

Although a multitude of genes encode the small subunit and the large subunit of AGP, early sequence data (compiled by Smith-White and Preiss, 1992) clearly showed less sequence heterogeneity within members of the small subunit gene family compared with those of the large subunit class. Interpretation of this result is not obvious because mutation in either subunit can affect both the kinetic and the allosteric properties of AGP (for review, see Hannah, 1997). Hence, the genes encoding both subunits should be under the same evolutionary constraints.

One explanation for the lack of sequence heterogeneity within at least two members of the small subunit family was recently suggested by Thorbjornsen et al. (1996a). They reported that a single barley gene encodes both the endosperm and leaf AGP small subunit. Two different sequences serve as the first exon of the respective genes and the use of two promoters or alternative splicing then provides the two different mature transcripts.

Another observation resulting from the barley sequence analysis was the absence of signal peptide. This observation raises the issue of enzyme localization (Thorbjornsen et al., 1996b). As reviewed below, genetic, biochemical, molecular, and localization studies point to extraplastidal location in maize and barley kernels.

Here, we initially asked whether maize employed one gene to encode both the leaf and endosperm AGP small subunit. In contrast to the single gene reported in barley, we detected two separate genes in maize for the leaf and endosperm AGP small subunits. Except for exon 1, the maize genes exhibit obvious sequence similarity in exons and introns. Sequence similarity in introns may point to a fairly recent duplication giving rise to these two genes. Much of the sequence divergence within introns likely was caused by transposable element visitation. In contrast to the rest of the gene, exon 1 sequences are more divergent and appear to have a different origin compared to the rest of the gene. Within cereal endosperms, the first exon likely has had two independent origins. Finally, from comparison of sequenced plant genes, the duplications giving rise to the rest of the gene likely did not occur until after the evolutionary separation of monocots from eudicots.

RESULTS

Gene Nomenclature

The *brittle-2* (*Bt2*) gene of maize was named for the phenotype resulting from loss of its activity. Subsequent work showed that it encoded the smaller subunit of the major, endosperm-specific AGP. Because the gene encoding the smaller subunit of the leaf AGP is not associated with a mutant phenotype, we term this gene *Agpslzm* (*AGP, small, leaf, Zea mays*). This abbreviation replaces L2 we used previously (Prioul et al., 1994), encapsulates nomenclature of AGP, and renders the necessary plant and tissue specificity. Likewise, we replace AGP2, the term we (Giroux and Hannah, 1994) used previously for the cloned embryo small subunit, with *Agpsemzm*.

Determination of *Bt2* and *Agpslzm* Structures

Fragments from partial cDNA clones of *Bt2* and *Agpslzm* reported previously (Bae et al., 1990; Prioul et al., 1994) were used in primer extension experiments with endosperm or leaf RNA to ascertain 5′ sequences of corresponding transcripts. Corresponding genomic clones isolated by conventional hybridization, harbor sequences matching those of the mature transcripts. Two *Agpslzm* genomic clones from separate maize lines were sequenced. Resulting sequences were identical in a 5,026-bp overlap that encompasses all exons and introns.

Introns of *Bt2* and *Agpslzm*

Both *Bt2* and *Agpslzm* consist of nine introns and 10 exons (Table I). Introns occur in identical positions in both genes. Moreover, corresponding introns of approximately the same size exhibit much sequence similarity (Table I).

Those introns differing in size exhibit alterations that are hallmarks of transposable elements. Straight-

Table I. *Comparison of exons and introns of Bt2 and Agpslzm*

Sequence	Bt2	Agpslzm	Identity
	bp		%
Exon 1	153	>255	ND[a]
Intron 1	1,361	2,001	ND
Exon 2	297	297	97
Intron 2	1,213	300	ND
Exon 3	270	270	95
Intron 3	396	123	25[b]
Exon 4	180	180	96
Intron 4	291	317	~73
Exon 5	104	104	94
Intron 5	117	121	85
Exon 6	112	112	96
Intron 6	120	123	83
Exon 7	99	99	97
Intron 7	76	72	82
Exon 8	120	120	90
Intron 8	589	666	ND
Exon 9	118	118	97
Intron 9	156	175	ND
Exon 10	264	230	<79

[a] ND, Not determined. [b] 25% comparing all 396 bp, 80% at 5' and 3' termini compared with the 123 bp.

forward examples of this are shown in Figure 1. Introns 5 are 85% identical and, with one exception, differ only by single base substitutions principally in the central portion of the intron. The extra bases in intron 5 of *Agpslzm* likely were derived by an internal duplication involving an inversion. Introns 6 (83% identical) and 7 (82% identical) of *Bt2* and *Agpslzm* differ in size by incorporation of relatively short internal repeats. More complex footprints occur as well (Fig. 2). The second introns of *Bt2* and *Agpslzm* are nearly identical except for the central portion. *Bt2* contains a duplicated, partially overlapping 28-bp sequence. A repeating AG motif of 180 bp follows the duplication. Finally, four relatively short direct repeats terminate the sequence unique to *Bt2*.

A more complicated series of rearrangements is found in intron 3 (Fig. 3.) A fragment of 274 bp, unique to *Bt2*, contains a series of direct and inverted duplications. In one case, labeled duplication 6, each duplicate is actually composed of two tandem inverted repeats. Searches of GenBank (October 2000) with the sequence delimited by and including the left member of duplication 1 and the right member of duplication 9 showed that sequences of greatest similarity were located on several rice chromosomes. Arabidopsis sequences were not detected. This suggests that this sequence represents a repetitive element possibly unique to monocots or cereals.

Splicing of *Bt2* and *Agpslzm* Introns Uses Consensus Splice Sites

All introns begin and end with the consensus dinucleotide pairs, GT and AG. This is relevant because Anderson et al. (1991) reported that intron 2 of a rice

AGP small subunit gene began with CA and ended with CC. Intron 2 and adjoining exon sequences of the rice gene, *Bt2, Agpslzm*, the potato small subunit gene, and the Arabidopsis gene for the small subunit are shown in Figure 4. Aside from a T missing in the rice gene, the three cereal genes are virtually identical. We also note that the sequence TCAGG is duplicated at the termini of the cereal introns. This duplication precludes definitive identification of the exact splice sites, but we note that an intron beginning with GT and ending with AG can be identified in the *Bt2* and *Agpslzm* genomic sequences as well as in the potato sequence. The duplication, however, does allow for a number of non-consensus termini in the rice intron. At present, it is unknown whether splicing of this rice intron follows nonconventional patterns of splicing and, if so, which sequences actually terminate the intron or whether the sequence reported is in error. It is interesting that one member of this gene family, from Arabidopsis, lacks this intron as well as one member of the duplication.

Coding Regions of *Bt2* and *Agpslzm*

Exons 2 through 9 of *brittle-2* and *Agpslzm* are of the same size and are identical at 96% of the corresponding bp. Exons 10, which lack protein coding information, exhibit approximately 79% similarity. Both genes initiate translation in exon 1 and termination is near the 3' terminus of exon 9. As discussed below, the high GC content of *Agpslzm* exon 1 prevented unequivocal determination of the start of transcription and translation. Hence, the provisional start of translation is inferred from comparison with the translation start site of the barley leaf small subunit

Figure 1. *Agpslzm* (abbreviated here and below as Asl to conserve space) and *Bt2* introns though consistent in placement vary in length. Comparison of the three most conserved introns indicates the size differences are due to insertions of duplicated sequence, a hallmark of transposable element visitation. The duplicated sequences are in bold and orientation is shown by arrows.
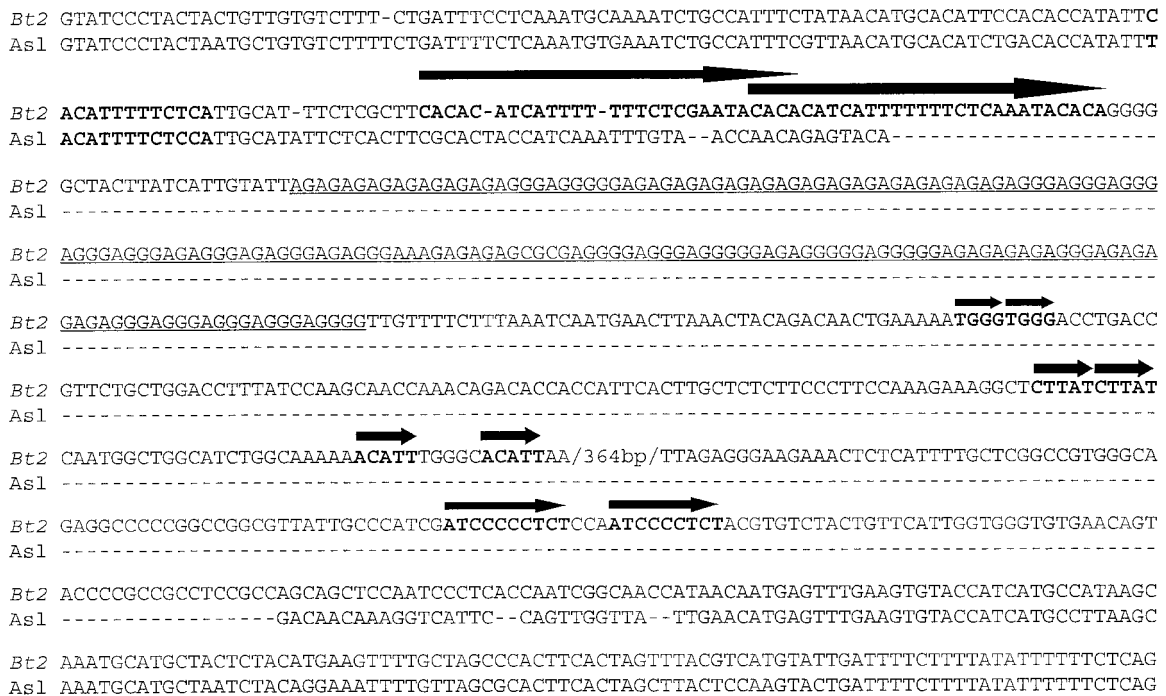
```
Bt2 GTATCCCTACTACTGTTGTGTCTTT-CTGATTTCCTCAAATGCAAAATCTGCCATTTCTATAACATGCACATTCCACACCATATTC
Asl GTATCCCTACTAATGCTGTGTCTTTTCTGATTTTCTCAAATGTGAAATCTGCCATTTCGTTAACATGCACATCTGACACCATATTT

Bt2 ACATTTTTCTCATTGCAT-TTCTCGCTTCACAC-ATCATTTT-TTTCTCGAATACACACATCATTTTTTTCTCAAATACACAGGGG
Asl ACATTTTCTCCATTGCATATTCTCACTTCGCACTACCATCAAATTTGTA--ACCAACAGAGTACA--------------------

Bt2 GCTACTTATCATTGTATTAGAGAGAGAGAGAGAGAGAGGGAGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGGAGGGAGGG
Asl ------------------------------------------------------------------------------------------

Bt2 AGGGAGGGAGAGGGAGAGGGAGAGGGAAAGAGAGAGCGCGAGGGGGAGGGAGGGGGAGAGGGGGGAGGGGGAGAGAGAGAGGGAGAGA
Asl ------------------------------------------------------------------------------------------

Bt2 GAGAGGGAGGGAGGGAGGGAGGGGTTGTTTTCTTTAAATCAATGAACTTAAACTACAGACAACTGAAAAATGGGTGGGACCTGACC
Asl ------------------------------------------------------------------------------------------

Bt2 GTTCTGCTGGACCTTTATCCAAGCAACCAAACAGACACCACCATTCACTTGCTCTCTTCCCTTCCAAAGAAAGGCTCTTATCTTAT
Asl ------------------------------------------------------------------------------------------

Bt2 CAATGGCTGGCATCTGGCAAAAAACATTTGGGCACATTAA/364bp/TTAGAGGGAAGAAACTCTCATTTTGCTCGGCCGTGGGCA
Asl ------------------------------------------------------------------------------------------

Bt2 GAGGCCCCCGGCCGGCGTTATTGCCCATCGATCCCCCTCTCCAATCCCCTCTACGTGTCTACTGTTCATTGGTGGGTGTGAACAGT
Asl ------------------------------------------------------------------------------------------

Bt2 ACCCCGCCGCCTCCGCCAGCAGCTCCAATCCCTCACCAATCGGCAACCATAACAATGAGTTTGAAGTGTACCATCATGCCATAAGC
Asl ----------------GACAACAAAGGTCATTC--CAGTTGGTTA--TTGAACATGAGTTTGAAGTGTACCATCATGCCTTAAGC

Bt2 AAATGCATGCTACTCTACATGAAGTTTTGCTAGCCCACTTCACTAGTTTACGTCATGTATTGATTTTCTTTTATATTTTTTCTCAG
Asl AAATGCATGCTAATCTACAGGAAATTTTGTTAGCGCACTTCACTAGCTTACTCCAAGTACTGATTTTCTTTTATATTTTTTCTCAG
```

Figure 2. Intron 2 of *Agpslzm* and *Bt2* has the greatest variation in size with *Bt2* 913 bp larger than *Agpslzm*. Duplicated sequences are denoted by arrows and bold type. The long overlapping duplication (long arrows) in *Bt2* bears some similarity to a sequence present in both genes (indicated by bold type with no arrow). *Bt2* also contains a long AG repeat, underlined.

gene. We do not know the full extent of *Agpslzm* exon 1; however, the location of the presumed ATG is 255 bp 5′ of the 3′ terminus of exon 1 and the closest possible TATA sequence is 283 bp upstream from the ATG. *Bt2* exon 1 is 153 bp in length with a TATA box 34 bp upstream of the start of transcription.

Exons 1 of AGP Small Subunit Genes Exhibit Striking Sequence Dissimilarity

Exons 1 of *Bt2* and *Agpslzm* show little to no similarity at either the DNA or amino acid level whereas they differ at only 14 of the 434 amino acids encoded by exons 2 to 9 (97% identity). This pattern of sequence similarity is consistent across all AGP small subunits in GenBank as well as the maize embryo form, now termed *Agpsemzm* (accession no. AY032604). The peptides encoded by the first exons are highly variable, whereas the rest of the proteins are virtually identical. Amino acids encoded by exon 1 and the 5′ portion of exon 2 of a number of AGP small subunit genes are shown in Figure 5.

Some similarities can be found in exon 1. *Agpslzm*, *Agpsemzm*, and the other monocot leaf and eudicot genes display discernible conserved motifs. These sequences likely reflect coding information for transit peptides. In contrast, *Bt2* and the seed sequences from the other cereals have few similarities and, in all likelihood, lack transit peptides (see below). Finally,

we note that barley, wheat, and rice seed small subunits share similarities not found in *Bt2*.

Exons 1 Define a Phylogenetic Tree Different from the Rest of the Gene

Because of the demarcation in similarity exhibited by the two portions of the small subunit genes, phylogenetic relationships of each section were determined separately. A phylogenetic tree resulting from the comparison of exon 1-encoded peptides consists of three main branches (Fig. 6). The *Bt2* peptide is the only representative of one of these branches. The other monocot seed/endosperm sequences lie on the second branch, whereas all other sequences cluster at the end of the third branch. The percent accepted mutation distance analysis found no similarity between *Bt2* and any of the other exon 1 encoded peptides. There was also no detectable similarity between the monocot seed small subunit and the other sequences.

In contrast, the tree generated using the remaining coding sequences (exons 2–9) contains only two main branches. Overall, the percent accepted mutation distances in these comparisons are much smaller than those in the monocot leaf, *Agpsemzm*, eudicot cluster in the exon 1 tree. An interesting feature of this tree is the separation of monocots and eudicots. This tree is nearly identical to one resulting from the whole
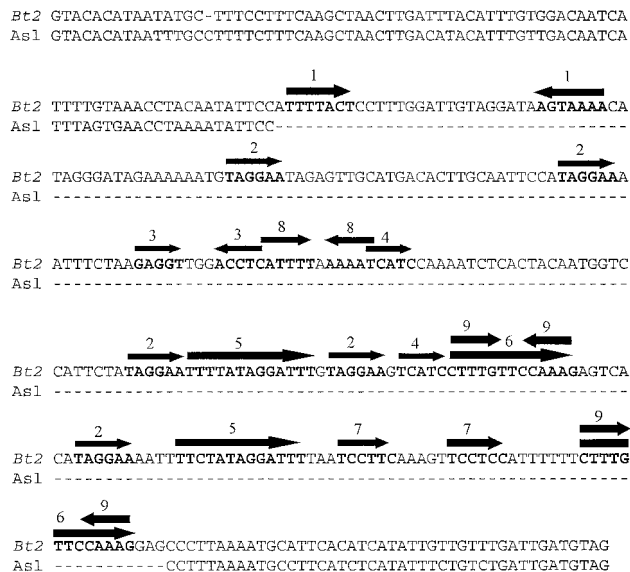
```
Bt2 GTACACATAATATGC-TTTCCTTTCAAGCTAACTTGATTTACATTTGTGGACAATCA
As1 GTACACATAATTTGCCTTTTCTTTCAAGCTAACTTGACATACATTTGTTGACAATCA

                      1                              1
Bt2 TTTTGTAAACCTACAATATTCCATTTTACTCCTTTGGATTGTAGGATAAGTAAAACA
As1 TTTAGTGAACCTAAAATATTCC--------------------------------

              2                                         2
Bt2 TAGGGATAGAAAAAATGTAGGAATAGAGTTGCATGACACTTGCAATTCCATAGGAAA
As1 --------------------------------------------------------

             3       3      8       8     4
Bt2 ATTTCTAAGAGGTTGGACCTCATTTTAAAAATCATCCAAAATCTCACTACAATGGTC
As1 --------------------------------------------------------

             2         5         2       4        9      9
Bt2 CATTCTATAGGAATTTTATAGGATTTGTAGGAAGTCATCCTTTGTTCCAAAGAGTCA
As1 --------------------------------------------------------

             2          5          7         7            9
Bt2 CATAGGAAAATTTTCTATAGGATTTAATCCTTCAAAGTTCCTCCATTTTTTCTTTG
As1 --------------------------------------------------------

        6      9
Bt2 TTCCAAAGGAGCCCTTAAAATGCATTCACATCATATTGTTGTTTGATTGATGTAG
As1 ----------CCTTTAAAATGCCTTCATCTCATATTTCTGTCTGATTGATGTAG
```

Figure 3. Though *Agpslzm* and *Bt2* introns vary in length, the shorter intron sequence can be found in the terminal portions of the longer intron. The internal sequence of the longer intron contains duplications such as the complex sequence shown here in intron 3. These duplications may be due to successive insertions and excisions of transposable elements. Arrows with identical numbers indicate duplicated sequences. Duplicated sequences are in bold.

protein (data not shown) as well as trees developed via Phylip programs.

DISCUSSION

We initially asked whether the closely related plants, maize and barley, exhibit the same form of genetic control of the leaf and endosperm small subunit AGPs. Whereas one gene in barley has been shown to encode subunits expressed in both the leaf and seed, we report here that two different genes are employed in maize. Whether the barley gene is also contained within the maize genome cannot be ascertained at this time; however, screening of approximately 560,000 expressed sequence tags in the Pioneer/DuPont database (T. Helentjaris and C. Zinselmeier, personal communication) failed to detect any similarity to the first exons of the barley gene. Based on extant data, it appears that maize and barley exhibit a fundamental difference in their genetic control of two of the AGP small subunits.

A number of interesting inferences can be gleaned about the possible evolution of the AGP small sub-

unit genes from perusal of the genes encoding this protein.

First, sequence variation within the genes is not randomly distributed. The vast majority of sequence diversity lies in the 5′ portion of the genes. For example, exon 1 and intron 1 of *Agpslzm* bear no similarity to their counterparts in *Bt2*, but beginning in exon 2, 97% of corresponding amino acids and 95.5% of corresponding coding sequences are identical. Identity of third positions of codons is also quite high, 90%. This similarity extends into the introns as well. Of the three introns not disturbed by insertions/deletions, 83.5% of the sequences are identical.

Some, but not all, of the variation at the N terminus likely is associated with the fundamental dichotomy in cell localization among the AGPs. The data presented here show that exon 1 coding sequences clearly separate endosperm and nonendosperm AGP small subunits.

Although AGP is in the chloroplasts of spinach (*Spinacia oleracea*) leaves (Okita et al., 1979) and in the amyloplast of potato tubers (Kim et al., 1989), recent protein (Giroux and Hannah, 1994; Villand and Kleczkowski, 1994), genetic (Cao et al., 1995; Shannon et al., 1996), and cell fractionation studies (Denyer et al., 1996; Thorbjornsen et al., 1996b) provide strong arguments for a cytosolic localization of AGP activity in cereal endosperms. This localization is consistent with immunocytological studies in maize endosperm (Brangeon et al., 1997). Furthermore, Choi et al. (2001) recently transformed tobacco with constructs expressing the N terminus of BT2 fused to the green flourescent protein. The BT2 N terminus did not target the green flourescent protein to the plastid. Finally, Beckles et al. (2001) have concluded that a cytosolic location for AGP is a feature of graminaceous endosperms but not of other starch-storing organs. The clear distinction in exon 1 coding sequences reported above most likely reflects differences in intracellular targeting.

Second, although the sequences of all plant AGP small subunit genes are not available, the present data suggest that this gene appears to have evolved as two separate modules. Exon 1 represents one module, whereas exons 2 through 10 represent the second module. Cereal endosperm genes define exclusively two of the three exon 1 branches. This strongly points to two independent origins of exon 1 in the genes encoding the endosperm AGP small subunit.

Particularly interesting in this regard are the three maize genes encoding the small subunit. Exon 1 sim-

```
Rice     AAT CCT AAC TGG TT  caggtatcatc...tttatattctgaccT CAG GGT ACT GCA GAT GCT GTA AGA
Bt2      AAT CCA AAC TGG TTT CAGgtatccct...tttttatatttttct cag GGT ACT GCA GAT GCT GTA AGG
Agpslzm  AAT CCA AAC TGG TTT CAGgtatcccт...tttttatatttttct cag GGT ACT GCA GAT GCT GTA AGG
Potato   AAC CCC GAT TGG TTC CAGgtaacttc...gctcatgcttgtctg tag GGC ACG GCT GAT GCT GTC AGA
Arabid   AAC CCC AAC TGG TTC CAG                          GGG ACA GCT GAT GCC GTC AGG
```

Figure 4. Intron 2 splice junctions and flanking exon sequences from rice, *Bt2*, *Agpslzm*, potato, and Arabidopsis. Spaces have been placed between codon triplets in the exons. Bolded letters indicate the duplicated sequence.

ilarities place the embryo and leaf genes in one group, whereas the similarities in the remaining coding region place the leaf and endosperm genes into one category. Separate evolutionary histories are clearly suggested.

Third, although the coding region from exon 2 through 9 exhibits strong conservation, the variation within this sequence suggests that duplications giving rise to the small subunit family members expressed in the various tissues occurred at least twice during plant evolution. Further, these duplications most likely occurred after the evolutionary separation of eudicots and cereals. For example, note that the gene most closely related to the maize form of the leaf small subunit, *Agpslzm*, is the maize endosperm gene. The same relationship holds for the two cloned

sweet potato genes. In contrast, the only case for a gene duplication occurring before plant speciation can be made for the genes of pea and fava bean. The extant data strongly point to at least two if not more independent sets of duplications of the progenitor gene. At least some of these duplications apparently occurred after the evolutionary separation of monocots and eudicots.

Some gene features, however, suggest an alternative relationship. We note the pronounced similarity of genes isolated from individual plant species as well as the high level of sequence identity exhibited by *Bt2* and *Agpslzm* in internal regions of some introns and in third positions of codons—genic regions not normally thought to be under evolutionary pressure. These features suggest that sequences of small
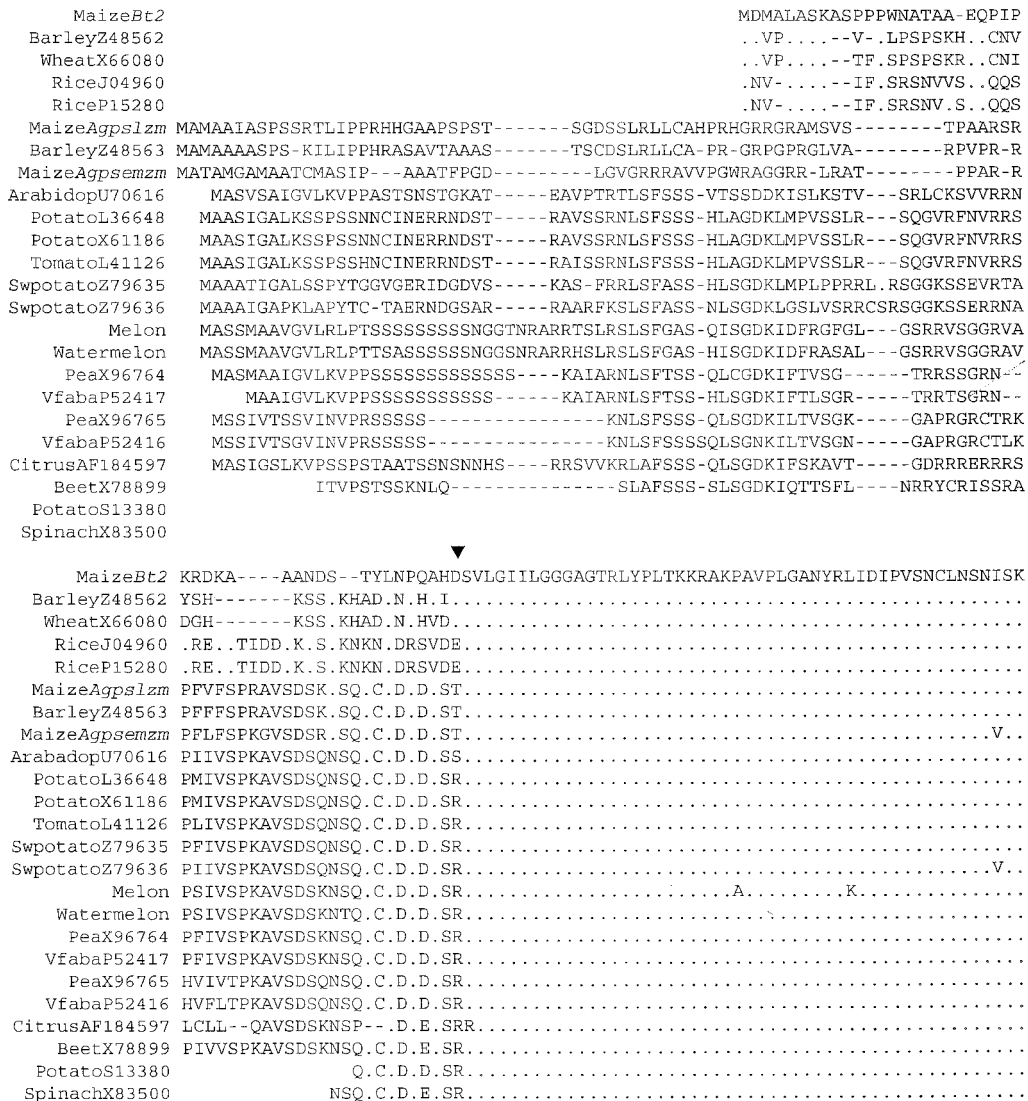
```
          MaizeBt2                                                              MDMALASKASPPPWNATAA-EQPIP
       BarleyZ48562                                                             ..VP.....--V-.LPSPSKH..CNV
        WheatX66080                                                             ..VP.....--TF.SPSPSKR..CNI
          RiceJ04960                                                            .NV-.....--IF.SRSNVVS..QQS
          RiceP15280                                                            .NV-.....--IF.SRSNV.S..QQS
      MaizeAgpslzm MAMAAIASPSSRTLIPPRRHHGAAPSPST-------SGDSSLRLLCAHPRHGRRGRAMSVS--------TPAARSR
      BarleyZ48563 MAMAAAASPS-KILIPPHRASAVTAAAS-------TSCDSLRLLCA-PR-GRPGPRGLVA--------RPVPR-R
     MaizeAgpsemzm MATAMGAMAATCMASIP---AAATFPGD---------LGVGRRRAVVPGWRAGGRR-LRAT--------PPAR-R
     ArabidopU70616    MASVSAIGVLKVPPASTSNSTGKAT-----EAVPTRTLSFSSS-VTSSDDKISLKSTV---SRLCKSVVRRN
      PotatoL36648    MAASIGALKSSPSSNNCINERRNDST-----RAVSSRNLSFSSS-HLAGDKLMPVSSLR---SQGVRFNVRRS
      PotatoX61186    MAASIGALKSSPSSNNCINERRNDST-----RAVSSRNLSFSSS-HLAGDKLMPVSSLR---SQGVRFNVRRS
      TomatoL41126    MAASIGALKSSPSSHNCINERRNDST-----RAISSRNLSFSSS-HLAGDKLMPVSSLR---SQGVRFNVRRS
    SwpotatoZ79635    MAAATIGALSSPYTGGVGERIDGDVS-----KAS-FRRLSFASS-HLSGDKLMPLPPRRL.RSGGKSSEVRTA
    SwpotatoZ79636    MAAAIGAPKLAPYTC-TAERNDGSAR-----RAARFKSLSFASS-NLSGDKLGSLVSRRCSRSGGKSSERRNA
             Melon MASSMAAVGVLRLPTSSSSSSSSSSNGGTNRARRTSLRSLSFGAS-QISGDKIDFRGFGL---GSRRVSGGRVA
         Watermelon MASSMAAVGVLRLPTTSASSSSSSNRARRHSLRSLSFGAS-HISGDKIDFRASAL---GSRRVSGGRAV
           PeaX96764   MASMAAIGVLKVPPSSSSSSSSSSSSSS----KAIARNLSFTSS-QLCGDKIFTVSG------TRRSSGRN-ᵕ
         VfabaP52417      MAAIGVLKVPPSSSSSSSSSSS-------KAIARNLSFTSS-HLSGDKIFTLSGR-----TRRTSGRN--
           PeaX96765 MSSIVTSSVINVPRSSSSS---------------KNLSFSSS-QLSGDKILTVSGK-----GAPRGRCTRK
         VfabaP52416 MSSIVTSGVINVPRSSSSS---------------KNLSFSSSSQLSGNKILTVSGN-----GAPRGRCTLK
      CitrusAF184597 MASIGSLKVPSSPSTAATSSNSNNHS----RRSVVKRLAFSSS-QLSGDKIFSKAVT-----GDRRRERRRS
           BeetX78899          ITVPSTSSKNLQ--------------SLAFSSS-SLSGDKIQTTSFL----NRRYCRISSRA
        PotatoS13380
       SpinachX83500
                                         ▼
          MaizeBt2 KRDKA----AANDS--TYLNPQAHDSVLGIILGGGAGTRLYPLTKKRAKPAVPLGANYRLIDIPVSNCLNSNISK
       BarleyZ48562 YSH-------KSS.KHAD.N.H.I......................................
        WheatX66080 DGH-------KSS.KHAD.N.HVD.....................................
          RiceJ04960 .RE..TIDD.K.S.KNKN.DRSVDE....................................
          RiceP15280 .RE..TIDD.K.S.KNKN.DRSVDE....................................
      MaizeAgpslzm PFVFSPRAVSDSK.SQ.C.D.D.ST....................................
      BarleyZ48563 PFFFSPRAVSDSK.SQ.C.D.D.ST....................................
     MaizeAgpsemzm PFLFSPKGVSDSR.SQ.C.D.D.ST.................................V..
     ArabadopU70616 PIIVSPKAVSDSQNSQ.C.D.D.SS....................................
      PotatoL36648 PMIVSPKAVSDSQNSQ.C.D.D.SR....................................
      PotatoX61186 PMIVSPKAVSDSQNSQ.C.D.D.SR....................................
      TomatoL41126 PLIVSPKAVSDSQNSQ.C.D.D.SR....................................
    SwpotatoZ79635 PFIVSPKAVSDSQNSQ.C.D.D.SR....................................
    SwpotatoZ79636 PIIVSPKAVSDSQNSQ.C.D.D.SR.................................V..
             Melon PSIVSPKAVSDSKNSQ.C.D.D.SR...............A.........K..........
         Watermelon PSIVSPKAVSDSKNTQ.C.D.D.SR....................................
           PeaX96764 PFIVSPKAVSDSKNSQ.C.D.D.SR....................................
         VfabaP52417 PFIVSPKAVSDSKNSQ.C.D.D.SR....................................
           PeaX96765 HVIVTPKAVSDSQNSQ.C.D.D.SR....................................
         VfabaP52416 HVFLTPKAVSDSQNSQ.C.D.D.SR....................................
      CitrusAF184597 LCLL--QAVSDSKNSP--.D.E.SRR...................................
           BeetX78899 PIVVSPKAVSDSKNSQ.C.D.E.SR....................................
        PotatoS13380          Q.C.D.D.SR....................................
       SpinachX83500         NSQ.C.D.E.SR....................................
```

Figure 5. AGP small subunit sequences aligned relative to BT2 demonstrate the vast dissimilarity in variation of sequences encoded by exon 1 and exon 2. Sequences from GenBank are identified by plant name and accession no. − indicates the placement of intron 1. Hyphens indicate gaps in the sequence. Dots identify amino acids identical to those of BT2. All of exon 1-encoded sequences and one-half of exon 2 are shown. The remainder of the proteins exhibits the degree of similarity found in exon 2. Sequences were aligned using the program for multiple sequence alignment at http://www.toulouse.inra.fr/multalin.html.
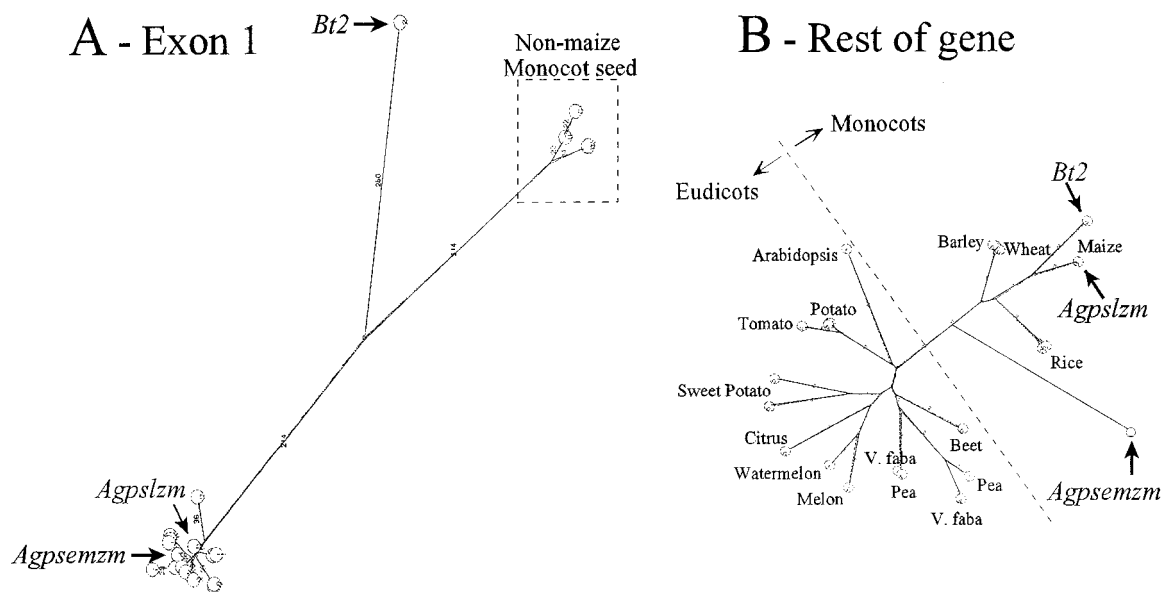
Figure 6. Unrooted phylogenetic trees of comparison of peptides encoded by the first exons only (A) and the remainder of the genes (B). Trees were prepared using all against all comparisons of the following proteins: maize *Bt2* (endosperm), maize *Agpslzm* (leaf), maize *Agpsemzm* (embryo), barley Z48563 (endosperm), barley Z48562 (leaf), potato X61186 (tuber), Arabidopsis U70616, beet (*Beta vulgaris*) X78899 (tap root), citrus (*Citrus unshiu*) AF184597, melon (*Cucumis melo*) AF030382 (mature fruit), pea X96764 (cotyledon), pea X96765 (cotyledon), potato L36648 (leaf), rice J04960 (seed), rice P15280 (endosperm specific), sweet potato Z79635 (tuberous root and leaf), sweet potato Z79636 (tuberous root and leaf), tomato L41126 (fruit), *V. faba* P52416 (cotyledon), *V. faba* P524167 (cotyledon), watermelon (*Citrullus lanatus*) AF032471 (fruit), and wheat X66080 (developing grain). The Internet site for the all against all related peptide sequence comparison was from the Computational Biochemistry Research Group Server of ETHZ at http://cbrg.inf.ethz.ch/subsection311.html. Resulting PostScript files were visualized using Aladdin Ghostscript 5.50 found at http://www.cs.wisc.edu/~ghost/aladdin/get550.html. Scales of A and B differ.

subunit genes may be homogenized, perhaps by gene conversion or related mechanisms involving nonallelic genes. However, this model does not explain the conspicuous separation of eudicots from cereals, would have to be limited to only certain regions of the gene, and does not account for the diversity separating *Agpsemzm* from *Bt2* and *Agpslzm*. Rather, we favor the hypothesis that the close sequence identity detailed here reflects relatively recent duplications.

The pattern of gene duplications inferred for the small subunit of AGP differs markedly from that of other genes. For example, duplications of the genes encoding Suc synthase, an enzyme involved in the same pathway and expressed in the same tissues, appear to be much more ancient (Shaw et al., 1994). Furthermore, genes encoding the large subunit of AGP also exhibit much more sequence diversity compared with those of the small subunit (Smith-White and Preiss, 1992).

Fourth, we note that intron 2 is not present in all gene family members. This polymorphism, coupled with the sequences at its exon-intron borders, suggests an interesting origin for this intron as described below.

Two theories explain the relative ages of exons and introns (for review, see Giroux et al., 1994). In the intron early hypothesis, introns came before genes, are ancient, border exons encoding functional domains of the resulting proteins, and were used in evolution to shuffle common exons into different genes. These introns were then lost in the vast majority of prokaryotic genes in the streamlining of their rapidly dividing genomes. An alternative explanation, "introns late," is that intron creation is a relatively late event occurring after the formation of functional genes. Transposons are usually considered the probable cause. Giroux et al. (1994) expanded on the "introns late" theory and noted that transposon insertions that duplicate certain host sequences can give rise to introns. The duplicate origin of intron termini explains at least some of the consensus sequences in the intron/exon borders. It is intriguing that those AGP genes having intron 2 also contain the duplication TCAGG. The last base of 5′ duplication is the terminal base of the intron donor site and the AG of the 3′ duplication terminates the intron. Splicing then would remove the insertion and effectively one copy of the duplication.

Finally, we note a possible relationship between AGP sequence and function. Endosperm AGPs generally exhibit less activation by the sugar, 3-phosphoglyceric acid, compared with AGPs from leaves, tubers, and embryos (for review, see Hannah, 1997). Sequence data summarized here show that AGPs with increased 3-PGA sensitivity contain more similar exon 1 sequences compared with other AGPs.

These data point to the possible importance of the small subunit N terminus in controlling the allosteric properties of AGP. This is presently under test.

MATERIALS AND METHODS

Structure of the *Brittle2* Mature Transcript

The sequence of the partial 1.7-kb *Bt2* cDNA clone published earlier (Bae et al., 1990) was completed using methods described therein. The 5′ end of the *Bt2* transcript was determined by direct sequencing of the reverse transcriptase products synthesized with gene-specific primers. The primer complementary to the cDNA sequence, TTG GCT GCT CGG CGG CGG TGG CAT TCC ATG GCG G, was 5′ end labeled with [γ-^{32}P]ATP (specific activity 5,000 Ci mmol^{-1}, NEN, Boston) and T4 polynucleotide kinase (Life Tech, Rockville, MD) basically by the methods of Sambrook et al. (1989) and purified via two rounds of ethanol precipitation. Poly(A^{+}) RNA was extracted from 22-d-old *sh2-R* endosperms by the method of McCarty (1986). The *Bt2* transcript is enhanced in *sh2-R* endosperms.

Following annealing, primer extension was performed with Moloney-Murine Leukemia Virus Reverse Transcriptase (Life Tech) in the presence of RNasin (Promega, Madison, WI) in conventional dideoxy sequence reactions. Sequences of resulting DNA fragments were read on 8% (w/v) polyacrylamide gels.

Structure of the *Bt2* Gene

Bt2 genomic clones were isolated from a CLONTECH (Palo Alto, CA)-prepared library of B73 DNA isolated at the two-leaf stage. A partial *Mbo*I digest cloned into the *Bam*HI site of λEMBL3 was probed with the most 5′ 168 bp of *Bt2* cDNA. The probe was prepared by PCR using the *Bt2* cDNA clone as template with the pUC19 T7/T3 primer (5′ AAC AGC TATGAC CAT G 3′) and a gene-specific primer located 148 to 168 bp from the 5′ terminus (5′ GAT TCC AAG AAC ACT ATC ATG 3′). The resulting fragment, corresponding to 153 bp of exon 1 and 15 bp of exon 2, hybridizes to a single fragment on maize (*Zea mays*) genomic Southerns and exhibits little if any sequence similarity to other plant AGP small subunits.

Five genomic clones were isolated from screening 150,000 plaques. Partial restriction mapping of these clones at the 5′ terminus of *Bt2* was performed by use of a probe from the 5′-most 327 bp of *Bt2* cDNA. This probe was isolated via PCR from the *Bt2* cDNA clone using the pUC19 T7/T3 primer with a gene-specific primer (5′ GTT AAA TTG CGT TAG CAC ATA 3′) located 307 to 327 bp from the transcription start site. Two clones, overlapping in the region encompassing exon 1, were chosen for subsequent sequence analysis. Fragments were subcloned into pUC19 and pSPORT vectors (Life Tech) and were sequenced as described below. The *Bt2* gene has been deposited in GenBank (accession no. AF334959).

Structure of the *Agpslzm* Gene

Maize cv Black Mexican Sweet genomic DNA from a single plant was partially digested with *Sau*3A and fractionated on a 5% to 24% (w/v) NaCl gradient. Three fractions containing DNA of 10 to 20 kb were pooled and ligated to *Bam*HI-cut λEMBL3 vector and packaged as in McCarty et al. (1986). The resulting library was probed with *Bt2* cDNA to extract the *Agpslzm* genomic clone. This was subsequently cloned into pSPORT using the *Sal*I sites in the lambda's polylinker regions.

A second genomic clone was obtained by screening a genomic library (CLONTECH FL1030 D) prepared with DNA of 7-d-old leaves of maize W22. This library was cloned into the *Bam*HI site of λEMBL3 after complete digestion with *Mbo*I. Both bacterial strains *Escherichia coli* NM539 and NM 538 were used as hosts following CLONTECH instructions. The partial (0.7 kb) cDNA clone described in Prioul et al. (1994) was used as a probe. The 0.2-kb leaf specific region at its 3′ terminus was used to distinguish *Bt2* and *Agpslzm* genomic clones. A 10-kb *Agpslzm* genomic clone was subcloned into the *Sal*I site of pUC19 and propagated in DH5α *E. coli* cells. *Pst*I and *Hin*dIII fragments were subcloned into pUC18 and sequenced.

Sequencing

Some of the DNA sequencing was done at the University of Florida, Interdisciplinary Center for Biotechnology Research (ICBR) DNA Sequencing Core Laboratory (Gainesville), using ABI Prism Dye Terminator sequencing protocol developed by Applied Biosystems (Foster City, CA) and gene-specific probes synthesized by the ICBR DNA Synthesis Core or Life Technologies (Rockville, MD). Double-stranded sequencing employing pUC19 or pSPORT clones used the dideoxy method and Sequenase (USB, Cleveland). In some cases, radiolabeled dCTP and autoradiography were utilized. Some sequencing employed amplification in a GeneAmp PCR System 9600 (Perkin Elmer, Foster City, CA) and a 373 DNA Sequencer Stretch (Applied Biosystems).

Structure of the *Agpslzm* Mature Transcript

cDNA Analysis (RT-PCR)

The sequence of the *Agpslzm* transcript was determined from total RNA isolated from W64A × 182E maize leaves using Trizol reagent (Life Technologies) following the manufacturer's protocol. Exons 2 through 7 were amplified using 4 μg of total RNA, a 3′ gene-specific primer (L2UDGRTR, CAU CAU CAU CAU GAT ATG CAA GAA CGG AGT CC) derived from the cDNA sequence previously published (Prioul et al., 1994) and Life Technologies SuperScript Preamplification System to prepare first strand cDNA. This was used subsequently as template for PCR amplification with the same 3′ primer, a 5′ gene-specific primer of sequence from exon 2 (L2UDGRTF, CUA CUA CUA CUA CTC GGA GGT GGT GCT GGG AC) and *Taq*

DNA polymerase. The resulting product was purified for sequencing using Ultrafree-MC 30,000 NMWL filter units (Millipore, Bedford, MA).

Amplification of the GC-rich exon 1 was attempted using Thermostable rTth Reverse Transcriptase RNA PCR Kit (PEBiosystems, Foster City, CA). First strand cDNA was prepared using 360 ng W64A × 182E leaf total RNA and either 3′ primer L2GSP2 (CCC GTA GGC TCT TGA GAG GTG ACG G) or L2GSP3UDG (CAU CAU CAU CAU GCG TCG GGG TCG AGG CAG GTC TGG G). The hot start technique of the kit protocol was followed: template DNA, primer, and buffer were incubated at 70°C for 5 min before addition of nucleotides, enzyme, and MnCl₂. The 70°C incubation was then continued for 15 min. PCR amplification following the kit protocol was performed using the same 3′ primer and a 5′ primer designed around the hypothesized transcription start site (L2ATGUDGGSP, CUA CUA CUA CUA GAG CAA TGG CGA TGG CAG CC). Products from the L2GSP3UDG/L2ATGUDGGSP amplification reaction were purified using Millipore's Ultrafree-MC 30,000 NMWL filter units and cloned into pAMP1 using Life Technologies CloneAMP PCR Cloning System.

5′ RACE

In initial 5′ RACE experiments, first strand cDNA was prepared from 400 ng of W22 poly(A⁺) RNA using the gene-specific primer L2GSP1 (CCATCCGGTACAGGTGATCGCCAGC). This is complementary to a sequence in exon 3. First strand cDNA was prepared, purified, and dC tailed as per Life Technologies 5′ RACE System for Rapid Amplification of cDNA Ends protocol. The product was amplified via PCR with a nested 3′ primer complementary to exon 2 sequence (L2GSP2, CCCGTAGGCTCTTGAGAGGTGACGG) and the anchor primer designed for the kit. The resulting products were purified using Millipore's Ultrafree-MC 30,000 NMWL filter units, digested with SalI (a site in the anchor primer) and PstI (a site in L2 5′ to L2GSP2), and cloned into pUC19.

L2GSP2 was used subsequently to initiate first strand synthesis. The nested primers L2GSP3 (GCGTCGGGGTCGAGGCAGGTCTGGG), from exon 1, and L2GSP4 (GAGGAAGGACGAAGGAGGCGAGGCG), 5′ to known exon 1 sequences, were used with the anchor primer for PCR amplification. The resulting products were electrophoresed on a 1% (w/v) agarose gel for size selection. Bands of 100 to 300 bp and 300 to 700 bp were extracted from the gel and the DNA was eluted from the agarose using Millipore's Ultrafree-MC 0.45-μm filter units. The products were digested with SstI or SalI and SstI for cloning into pUC19. A third primer, L2GSP3UDG (CAUCAUCAUCAUGCGTCGGGGTCGAGGCAGGTCTGGG) was used for amplification in order to allow the products to be cloned directly into pAMP1 using Life Technologies CloneAMP PCR Cloning System. These were first purified using Millipore's Ultrafree-MC 30,000 NMWL filter units.

Finally, the Thermostable rTth Reverse Transcriptase RNA PCR Kit (PEBiosystems) was used in conjunction with Life Technologies 5′ RACE System. First strand synthesis was performed using 360 ng of W64A × 182E leaf total RNA and L2GSP2 3′ primer following the rTth-RT protocol for hot start (70°C for 15 min) or cycling the temperature (95°C, 30 s to 70°C, 1 min for 20 cycles). The reaction was then treated with RNase, purified, and dC tailed with components of the 5′ RACE kit. PCR amplification was performed using the anchor primer and L2GSP3UDG and either Taq DNA polymerase or rTth DNA polymerase. The resulting products were purified using Millipore's Ultrafree-MC 30,000 NMWL filter units and cloned directly into pAMP1 using Life Technologies CloneAMP PCR Cloning System.

Agpslzm has been deposited in GenBank (accession no. AF334960).

Sequence Comparisons (Alignment and Phylogenetic Trees)

Sequences other than maize were obtained from GenBank. Unique protein sequences were aligned using Florence Corpet's multiple sequence alignment (Corpet, 1988) at http://www.toulouse.inra.fr/multalin.html. Only full-length or nearly full-length sequences were used to prepare phylogenetic trees using the AllAll: Related peptide sequences comparison from the Computational Biochemistry Research Group Server of ETHZ at http://cbrg.inf.ethz.ch/subsection3 1 1.html. The resulting PostScript files were visualized using Aladdin Ghostscript 5.50 downloaded from http://www.cs.wisc.edu/~ghost/aladdin/get550.html. The Phylip program was obtained from http://evolution.genetics.washington.edu/phylip.html.

LITERATURE CITED

Anderson JM, Larsen R, Laudencia D, Kim WT, Morrow D, Okita TW, Preiss J (1991) Molecular characterization of the gene encoding a rice endosperm- specific ADPglucose pyrophosphorylase subunit and its developmental pattern of transcription. Gene **97:** 199–205

Bae JM, Giroux M, Hannah LC (1990) Cloning and characterization of the *brittle-2* gene of maize. Maydica **35:** 317–322

Bae JM, Liu JR (1997) Molecular cloning and characterization of two novel isoforms of the small subunit of ADP-

glucose pyrophosphorylase from sweet potato. Mol Gen Genet **254:** 179–185

Ballicora MA, Laughlin MJ, Fu Y, Okita TW, Barry GF, 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

Beckles DM, Smith AM, apRees T (2001) A cytosolic ADPglucose pyrophosphorylase is a feature of Graminaceous endosperms but not of other starch-storing organs. Plant Physiol **125:** 818–827

Bhave MR, Lawrence S, Barton C, Hannah LC (1990) Identification and molecular characterization of shrunken-2 cDNA clones of maize. Plant Cell **2:** 581–588

Brangeon J, Reyss A, Prioul JL (1997) In situ detection of ADPglucose pyrophosphorylase expression during maize endosperm development. Plant Physiol Biochem **35:** 847–858

Bryce WH, Nelson OE Jr (1979) Starch-synthesizing enzymes in the endosperm and pollen of maize. Plant Physiol **63:** 312–317

Burgess D, Penton A, Dunsmuir P, Dooner H (1997) Molecular cloning and characterization of ADP-glucose pyrophosphorylase cDNA clones isolated from pea cotyledons. Plant Mol Biol **33:** 431–444

Cao H, Sullivan TD, Boyer CD, Shannon JC (1995) *Bt1*, a structural gene for the major 39–44 kD amyloplast membrane polypeptides. Physiol Plant **95:** 176–186

Chen BY, Janes HW (1997) Multiple forms of ADP-glucose pyrophosphorylase from tomato fruit. Plant Physiol **113:** 235–241

Choi SB, Kim KH, Kavakli IH, Lee SK, Okita TW (2001) Transcriptional expression characteristics and subcellular localization of ADP-glucose pyrophosphorylase in the oil plant *Perilla frutescens*. Plant Cell Physiol **42:** 146–153

Corpet F (1988) Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res **16:** 10881–10890

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

Doan DN, Rudi H, Olsen OA (1999) The allosterically unregulated isoform of ADP-glucose pyrophosphorylase from barley endosperm is the most likely source of ADP-glucose incorporated into endosperm starch. Plant Physiol **121:** 965–975

Fuchs RL (1977) Purification and characterization of ADP-glucose pyrophosphorylase A from maize endosperm. PhD thesis. Texas A&M University, College Station

Giroux MJ, Clancy M, Baier J, Ingham L, McCarty D, Hannah LC (1994) De novo synthesis of an intron by the maize transposable element Dissociation. Proc Natl Acad Sci USA **91:** 12150–12154

Giroux MJ, Hannah LC (1994) ADP-glucose pyrophosphorylase in shrunken-2 and brittle-2 mutants of maize. Mol Gen Genet **243:** 400–408

Giroux M, Smith-White B, Gilmore V, Hannah LC, Preiss J (1995) The large subunit of the embryo isoform of ADP

glucose pyrophosphorylase from maize. Plant Physiol **108:** 1333–1334

Greene TW, Kavakli IH, Kahn ML, Okita TW (1998) Generation of up-regulated allosteric variants of potato ADP-glucose pyrophosphorylase by reversion genetics. Proc Natl Acad Sci USA **95:** 10322–10327

Hannah LC (1997) Starch synthesis in the maize endosperm. *In* BA Larkins, IK Vasil, eds, Advances in Cellular and Molecular Biology of Plants, Vol 4. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 375–405

Hannah LC, Nelson OE Jr (1976) Characterization of ADP-glucose pyrophosphorylase from shrunken-2 and brittle-2 mutants of maize. Biochem Genet **14:** 547–560

Hylton C, Smith AM (1992) The *rb* mutation of peas causes structural and regulatory changes in ADP-glucose pyrophosphorylase from developing embryos. Plant Physiol **99:** 1626–1634

Kim WT, Franceschi VR, Okita TW, Robinson NL, Morell M, Preiss J (1989) Immunocytochemical localization of ADPglucose pyrophosphorylase in developing potato tuber cells. Plant Physiol **91:** 217–220

La Cognata U, Willmitzer L, Muller-Rober B (1995) Molecular cloning and characterization of novel isoforms of potato ADP-glucose pyrophosphorylase. Mol Gen Genet **246:** 538–548

Lin TP, Caspar T, Somerville C, 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 TP, Caspar T, Somerville CR, Preiss J (1988b) A starch deficient mutant of *Arabidopsis thaliana* with low ADP-glucose pyrophosphorylase activity lacks one of the two subunits of the enzyme. Plant Physiol **88:** 1175–1181

McCarty DR (1986) A simple method for extraction of RNA from maize tissue. Maize Genet Coop Newsl **60:** 61

McCarty DR, Shaw JR, Hannah LC (1986) The cloning, genetic mapping, and expression of the constitutive sucrose synthase locus of maize. Proc Natl Acad Sci USA **83:** 9099–9103

Muller-Rober B, Sonnewald U, Willmitzer L (1992) Inhibition of the ADP-glucose pyrophosphorylase in transgenic potatoes leads to sugar-storing tubers and influences tuber formation and expression of tuber storage protein genes. Embo J **11:** 1229–1238

Nakamura Y, Kawaguchi K (1992) Multiple forms of ADP-glucose pyrophorylase of rice endosperm. Plant Physiol **84:** 336–342

Okita TW, Greenberg E, Kuhn DN, Preiss J (1979) Subcellular localization of the starch degradative and biosynthetic enzymes of spinach leaves. Plant Physiol **64:** 187–192

Okita TW, Nakata PA, Anderson JM, Sowokinos JMM, Preiss J (1990) The subunit structure of potato tuber ADP-glucose pyrophosphorylase. Plant Physiol **93:** 785–790

Olive MR, Ellis RJ, Schuch WW (1989) Isolation and nucleotide sequences of cDNA clones encoding ADP-glucose pyrophosphorylase polypeptides from wheat leaves and endosperm. Plant Mol Biol **12:** 525–538

Preiss J, Lammel C, Sabraw A (1971) A unique adenosine diphosphoglucose pyrophosphorylase associated with maize embryo tissue. Plant Physiol **47:** 104–108

Preiss J, Romeo T (1994) Molecular biology and regulatory aspects of glycogen biosynthesis in bacteria. Prog Nucleic Acid Res Mol Biol **47:** 299–329

Preiss J, Sivak M (1996) Starch synthesis in sinks and sources. *In* E Zamski, ed, Photoassimilate Distribution in Plants and Crops: Source-Sink Relationships. Marcel Dekker Inc., New York, pp 139–168

Prioul JL, Jeannette E, Reyss A, Gregory N, Giroux M, Hannah LC, Causse M (1994) Expression of ADP-glucose pyrophosphorylase in maize (*Zea mays* L.) grain and source leaf during grain filling. Plant Physiol **104:** 179–87; erratum **Prioul JL, Jeannette E, Reyss A, Gregory N, Giroux M, Hannah LC, Causse M** (1994) Plant Physiol **105:** 1465; erratum **Prioul JL, Jeannette E, Reyss A, Gregory N, Giroux M, Hannah LC, Causse M** (1994) Plant Physiol **106:** 405

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning, Ed 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Shannon JC, Fang-Mei P, Kang-Chein L (1996) Nucleotides and nucleotide sugars in developing maize endosperms. Plant Physiol **110:** 835–843

Shaw JR, Ferl RJ, Baier J, St. Clair D, Carson C, McCarty DR, Hannah LC (1994) Structural features of the maize *sus1* gene and protein. Plant Physiol **106:** 1659–1665

Singletary GW, Banisadr R, Keeling PL (1997) Influence of gene dosage on carbohydrate synthesis and enzymatic activities in endosperm of starch-deficient mutants of maize. Plant Physiol **113:** 293–304

Smith AM, Bettey M, Bedford ID (1989) Evidence that the *rb* locus alters starch content of developing pea embryos through an effect on ADP-glucose pyrophosphorylase. Plant Physiol **89:** 1279–1284

Smith-White BJ, Preiss J (1992) Comparison of proteins of ADP-glucose pyrophosphorylase from diverse sources. J Mol Evol **34:** 449–464

Thorbjornsen T, Villand P, Denyer K, Olsen O, Smith AM (1996b) Distinct isoforms of ADPglucose pyrophosphorylase occur inside and outside the amyloplast in barley endosperm. Plant J **10:** 243–250

Thorbjornsen T, Villand P, Kleczkowski LA, Olsen OA (1996a) A single gene encodes two different transcripts for the ADP-glucose pyrophosphorylase small subunit from barley (*Hordeum vulgare*). Biochem J **313:** 149–154

Tsai CY, Nelson OE (1966) Starch-deficient maize mutant lacking adenosine dephosphate glucose pyrophosphorylase activity. Science **151:** 341–343

Villand P, Aalen R, Olsen OA, Luthi E, Lonneborg A, Kleczkowski LA (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, Kleczkowski LA (1994) Is there an alternative pathway for starch biosynthesis in cereal seeds? Z Naturforsch **49:** 215–219

Villand P, Olsen OA, Kleczkowski LA (1993) Molecular characterization of multiple cDNA clones for ADP-glucose pyrophosphorylase from *Arabidopsis thaliana*. Plant Mol Biol **23:** 1279–1284

Weber H, Heim U, Borisjuk L, Wobus U (1995) Cell-type specific, coordinate expression of two ADP-glucose pyrophosphorylase genes in relation to starch biosynthesis during seed development of *Vicia faba* L. Planta **195:** 352–361