Plant Physiol. (1995) 108: 1333–1334

Plant Gene Register

The Large Subunit of the Embryo Isoform of ADP Glucose Pyrophosphorylase from Maize[1]

Michael Giroux[2], Brian Smith-White, Virginia Gilmore[3], L. Curtis Hannah, and Jack Preiss*

Program in Genetics and Department of Biochemistry, Michigan State University, East Lansing, Michigan 48824-1319 (B.S.-W., V.G., J.P.); and Horticultural Sciences and Program in Plant Molecular and Cellular Biology, University of Florida, Gainesville, Florida 32611 (M.G., L.C.H.)

ADP Glc pyrophosphorylase (EC 2.7.7.27) is the enzyme catalyzing the first committed step of starch anabolism (reviewed by Sivak and Preiss, 1994). The enzyme in higher plants is a tetramer of two dissimilar subunits (reviewed in Sivak and Preiss, 1994). In *Zea mays*, the enzyme has been shown to be present as kinetically distinguishable isozymes in different tissues (Preiss et al., 1971; reviewed by Sivak and Preiss, 1994). For maize, two genetic loci conditioning the presence of enzyme activity solely in the endosperm have been reported (reviewed by Sivak and Preiss, 1994), and clones corresponding to these loci have been isolated and sequenced (Bae et al., 1990; Shaw and Hannah, 1992). The molecular basis for the observed isozymes is not completely known, since (a) genetic loci conditioning the presence of enzyme activity in other tissues are not known but are presumed to exist because *bt2* and *sh2* influence the enzyme activity exclusively in the endosperm, and (b) the cDNAs or the genomic clones corresponding to these presumptive loci have been reported only for part of the cDNA encoding a small subunit of one of the leaf isoforms (Prioul et al., 1994). This report concerns the determined nucleotide sequence of the cDNA and deduced primary structure corresponding to the large subunit of the embryo isoform of ADP Glc pyrophosphorylase (Table I). A companion report will do the same for the small subunit of the embryo isoform of ADP Glc pyrophosphorylase.

In conjunction with experiments published elsewhere (Giroux and Hannah, 1994), we conclude that this insert corresponds to almost, if not all, of the mRNA encoding the large subunit of the embryo ADP Glc pyrophosphorylase for three reasons. First, this insert recognizes an abundant mRNA in only the embryo (Giroux and Hannah, 1994). Second, the size of the embryo-specific mRNA is 2.1 kb (Giroux and Hannah, 1994), which agrees closely with that determined for the cDNA insert. Third, an identity matrix shows that the encoded protein is most like the large subunit of cereal endosperm enzymes. If the 46 residue positions where the cereal endosperm proteins show unique residues, relative to the other large subunit proteins, are examined in pairwise comparisons, the protein encoded by this cDNA shows a ratio of specific identity to specific difference between 2.1 and 2.3. We conclude that this protein is a member of the Agp2 family of ADP Glc pyrophosphorylase proteins (Smith-White and Preiss, 1994).

Received December 19, 1994; accepted February 10, 1995.
Copyright Clearance Center: 0032-0889/95/108/1333/02.
The GenBank accession number for the sequence reported in this article is Z38111.

LITERATURE CITED

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

Ball K, Preiss J (1994) Allosteric sites of the large subunit of the spinach leaf ADPglucose pyrophosphorylase. J Biol Chem **269:** 24706–24711

Burr B, Burr FA, Matz EC (1991) Database for loci mapping in T × Cm and CoxT × RI families. Maize Genet Coop Newslett **65:** 105–110

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

Joshi CP (1987) Putative polyadenylation signals in nuclear genes of higher plants: a compilation and analysis. Nucleic Acids Res **15:** 9627–9640

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

Prioul J-L, Jeanette 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:** 179–187

Shaw JR, Hannah LC (1992) Genomic nucleotide sequence of a wild type *Shrunken-2* allele of *Zea mays*. Plant Physiol **98:** 1214–1216

Sivak MN, Preiss J (1994) Starch synthesis in seeds. *In* M Negbi, J Kigel, G Galili, eds, Seed Development and Germination. Marcel Dekker, New York, pp 139–168

Smith-White B, Preiss J (1994) Suggested mnemonics for cloned DNA corresponding to enzymes involved in starch metabolism. Plant Mol Biol Rep **12:** S67–S71

[1] Supported in part by Department of Energy grant No. DEFG02-936R20121, by State of Michigan Research Excellence funds to J.P., and National Science Foundation grant No. 93–052818 and U.S. Department of Agriculture grant 94–37300–0453 to L.C.H. This manuscript is the Florida Agriculture Experiment Station Journal Series No. R-04317.

[2] Present address: U.S. Department of Agriculture Wheat Starch Laboratory, Washington State University, Pullman, WA 99164.

[3] Present address: School of Optometry, Ferris State University, Big Rapids, MI 49307.

* Corresponding author; e-mail preiss@pilot.msu.edu; fax 1–517–353–9334.

Table I. *Characteristics of the cDNA and encoded protein of the large subunit of the maize embryo ADP Glc pyrophosphorylase*

Organism:
Maize (*Zea mays*, cv W22).

Function:
Encodes a large subunit of ADP Glc pyrophosphorylase (EC 2.7.7.27).

Gene Location:
Nuclear genome, located on 6L (Burr et al., 1991).

Method of Identification:
Nucleotide and protein sequences were analyzed with algorithms in the Genetics Computer Group suite running on a Silicon Graphics server. Pairwise comparison with all other ADP Glc pyrophosphorylase proteins from plant sources showed that the deduced protein is 89 to 91% identical with members of the Agp2 family and 40 to 70% identical with members of all other ADP Glc pyrophosphorylase gene families (Smith-White and Preiss, 1994).

Sequencing Strategy:
Dideoxynucleotide chain termination protocols with Sequenase (United States Biochemical) using both single-strand templates produced from phagemids and double-strand templates of CsCl-purified DNA. Internal regions were accessed by deletion and subcloning of specific portions. Both strands have been seen across all joints.

G + C Content:
48.9% [excluding poly(A) tract], 48.2% in coding region, 55.6% in untranslated region [excluding poly(A) tract].

Structural Features of Nucleotides:
The 5′ untranslated material has no regions of dyad symmetry other than restriction enzyme sites. This region has 19 possible stem-loop structures with only three or four structures that do not obstruct each other. This region contains a 7-nucleotide block at position 89 that is directly repeated at position 99 and a 4-nucleotide block that is directly repeated at position 85, position 96, and position 137. All five blocks would be located in loops of an optimal RNA secondary structure. The deduced protein is encoded from a translation initiation codon at position 178 to a translation termination codon at position 1758. There is a 7-nucleotide block at position 1885 that is directly repeated at position 1929. A perfect match of the poly(A) addition signal (Joshi, 1987) is found beginning at position 2079. Three single-base mismatches of the consensus poly(A) addition signal are found beginning at positions 1978, 2013, and 2098. The cDNA is terminated with 56 adenine bases.

Structural Features of Encoded Protein:
The protein is composed of 521 residues with a computed molecular mass of 57,987 D and a computed pI of 7.97. The 3 N-terminal residues are identical with those found in the *sh2* protein (Shaw and Hannah, 1992). The isoforms of the large subunit from wheat seed (GenBank accession No. Z21969 from C. Ainsworth) and barley endosperm (GenBank accession No. X67151 from P. Villand) have 3 residues removed from the deduced N terminus of the initial translation product, a 14-residue block that is identical with the 14 N-terminal residues of the encoded protein of this cDNA. The significance of a domain of identity this close to the N terminus of the initial translation product is not clear in light of the essentially random sequence within broad boundary conditions exhibited by other identified transit peptides. Within the body of the mature protein, the Glc-1-P-binding site (reviewed by Sivak and Preiss, 1994), the domain thought to interact with the adenine portion of ATP and ADP Glc (reviewed by Sivak and Preiss, 1994), and the three portions of the allosteric ligand-binding site (Ball and Preiss, 1994) identified previously in other ADP Glc pyrophosphorylase proteins are conserved.

Expression Characteristics:
Most abundant in embryo, present in endosperm (Giroux and Hannah, 1994). Location within the embryo not determined.

Subcellular Location:
Not determined.