Molecular Cloning and Spatial Expression of an ApL1 cDNA for the Large Subunit of ADP-Glucose Pyrophosphorylase from Arabidopsis thaliana

Leszek A. Kleczkowski, Lubomir N. Sokolov, Cheng Luo* and Per Villand**
Department of Plant Physiology, Umeå University, 901-87 Umeå, Sweden
Z. Naturforsch. 54c, 353–358 (1999); received January 15/February 5, 1999
Carbon Partitioning, 3′-RACE, Starch Synthesis, Transit Peptide

A cDNA, ApL1a, corresponding to a homologue of the large subunit of ADP-glucose pyrophosphorylase (AGPase), has been isolated/characterised by screening a cDNA library prepared from leaves of Arabidopsis thaliana, followed by rapid amplification of cDNA 3′-ends (3′-RACE). Within the 1685 nucleotide-long sequence (excluding polyA tail), an open reading frame encodes a protein of 522 amino acids (aa), with a calculated molecular weight of 57.7 kDa. The derived aa sequence does not contain any discernible transit peptide cleavage site motif, similarly to two other recently sequenced full-length Arabidopsis homologues for AGPase. The aa sequence shows 58–78% identity to homologous proteins from other plants/tissues. The corresponding gene was found to be expressed in all tissues examined (root and stem leaves, stems, flowers and fruits). The ubiquitous expression of the gene is consistent with its critical role in starch synthesis in Arabidopsis.

Introduction

ADP-glucose pyrophosphorylase (AGPase) is the key enzyme of starch biosynthesis in all plants [reviewed in (Preiss, 1978; Kleczkowski et al., 1991)]. In plants, the enzyme is composed of two subunit types, encoded by different genes (Olive et al., 1988; Lin et al., 1988; Nelson and Pan, 1995). Studies with AGPase mutants (Lin et al., 1988; Nelson and Pan, 1995) and using “antisense” technology (Müller-Röber et al., 1992) have established that AGPase plays a pivotal role in providing carbon skeletons for starch synthesis in plants. AGPase from several tissues (e.g. leaves) has been shown to be potently regulated by metabolic effectors (Preiss, 1978; Kleczkowski et al., 1993b), and this in turn may control the rate of starch synthesis. Effector regulation of AGPase has frequently been considered as a model in studies on control of carbon metabolism in plants (Kleczkowski, 1994). Apart from its possible regulation by metabolic effectors, AGPase is also regulated at the transcriptional level, most notably by sugars (Müller-Röber et al., 1990; Nakata and Okita, 1995; Sokolov et al., 1998) and/or osmotica (Sokolov et al., 1998), suggesting the involvement of complex signal transduction mechanisms that mediate gene expression. In several plants, at least two isozymes of AGPase have been described that may differ in their molecular mass (Mr), kinetic/ regulatory properties and/or intra- and inter-cellular compartmentation (Kleczkowski et al., 1993a, b; Villand and Kleczkowski, 1994; Nelson and Pan, 1995; Weber et al., 1996; Denyer et al., 1996; Kleczkowski, 1996; Chen and Janes, 1998).

Previously, we have used a polymerase chain reaction (PCR) technique to amplify several partial [ca. 540 nucleotide (nt)-long] cDNAs corresponding to small and large subunits of AGPase from various plant tissues (Villand et al., 1992a, 1993). These PCR products have proven to be very useful as gene-specific probes to isolate full length cDNAs, genomic clones and for expression studies (Villand et al., 1992a, b, 1993; Thorbjørnsen et al., 1996; Eimert et al., 1997). In Arabidopsis, a model plant species, at least three genes for the large subunit and one gene for the small subunit have been identified (Villand et al., 1993). One of the large subunit genes (ApL1) corresponds to the large subunit of AGPase that is considered to be critical for starch synthesis, based on mutant studies (Lin
et al., 1988; Wang et al., 1997). A mutant lacking ApL1 protein had greatly reduced AGPase activity and its starch levels were at 40% when compared to wild-type plants (Lin et al., 1988). Expression of the ApL1 gene is strongly dependent on osmotic pressure, as found by exposing excised Arabidopsis leaves to specific sugars and general osmotics (Sokolov et al., 1998), suggesting the involvement of an osmoticum-mediated signal transduction pathway.

In the present study, isolation of a cDNA encoding full open reading frame (ORF) of Arabidopsis ApL1 protein is reported, along with characterisation of its expression in Arabidopsis tissues. The nucleotide sequence of ApL1a has been deposited in GenBank under accession number AF117570.

Materials and Methods

**Isolation and characterisation of ApL1a**

The characterisation of ApL1a cDNA was carried out in two steps. First, a 1669 nt-long ApL1-1 clone was isolated from an Arabidopsis cDNA library by probing with an earlier isolated (Villand et al., 1993) 540-nt-long ApL1 cDNA. The ApL1-1 clone encoded a full ORF for the large subunit of AGPase, but lacked polyA tail, indicating that it corresponded to a 3'-end truncated mRNA. Subsequently, a putative composite full-length cDNA sequence (ApL1a) was obtained by rapid amplification of cDNA 3'-ends (3'-RACE) using specific primers. For this purpose, an aliquot of 5 μg of total RNA was isolated from leaves of one-month-old Arabidopsis (see below) and reversed to cDNA with oligo dNTPs by avian myeloblastosis virus (AMV) reverse transcriptase, following manufacturers’ instruction (Promega). A specific forward primer (5'-CTCGGAGGGTATACAGAAGCAGAT-3') and an adapter primer (5'-CTGATCAAGGTTAGG[T]15-3') were used for reverse transcriptase PCR; the adapter primer contained 15 oligo dT plus 14 nt random nucleotides to balance the antisense primer’s annealing temperature. The specific PCR product was cloned by TA-cloning (Invitrogen). A 196 nt-long clone isolated by this approach fully aligned with the corresponding sequences with the ApL1-1 cDNA and allowed for determination of the complete 3'-end nt sequence of the transcript. All cDNAs were sequenced on both strands using custom-made primers (fmol DNA Sequencing, Promega, Madison, USA). The nt and aa sequence comparisons were performed against GenBank non-redundant databases (www.ncbi.nlm.nih.gov), using the basic local alignment search tool (BLAST).

**Northern blot analyses**

Arabidopsis thaliana (L.) Heynh., ecotype Columbia, plants were grown in a chamber with 10 h white light (125 μE m⁻² s⁻¹, 22 °C) and 14 h darkness (18 °C) photoperiod regime. Total RNA was isolated from various tissues of 6–7 week-old mature plants in the end of the 14 h dark period, using a modification of the guanidine thiocyanate-based method, as described in (Sokolov et al., 1998), and aliquots of 15 μg of total RNA were electrophoretically resolved on 1.2% agarose-formaldehyde gels. Following Northern blot transfer, Hybond-N⁺ (Amersham) membranes with the blotted/immobilised RNA were hybridised with ApL1-1 probe. Hybridisations and subsequent washes were performed at 65 °C, following standard procedures (Sambrook et al., 1989).

**Results and Discussion**

**Nucleotide and deduced aa sequence**

The nt sequence of ApL1a, a cDNA encoding large subunit (ApL1) of AGPase, was obtained from ApL1-1 clone isolated after screening an Arabidopsis cDNA library with an earlier characterised (Villand et al., 1993) 540 nt-long ApL1 clone, followed by 3'-RACE. The 1685 nt-long (excluding polyA tail) ApL1a encodes an ORF of 522 aa (Fig. 1). Based on nt sequence, ApL1a is the most homologous (78–81% identity) to large subunits of AGPase from leaves of tomato (GenBank accession # U85497) (Chen et al., 1998), and aliquots of 15 pg of total RNA were electrophoretically resolved on 1.2% agarose-formaldehyde gels. Following Northern blot transfer, Hybond-N⁺ (Amersham) membranes with the blotted/immobilised RNA were hybridised with ApL1-1 probe. Hybridisations and subsequent washes were performed at 65 °C, following standard procedures (Sambrook et al., 1989).
CTTTCCTGCGAAAATGGTGGTCTCTGCTGACTGCAGAATCTCCCTCTCTGCCCCTAGCTGCATACGTAGTAGCTCCACGGGATTGACTAGGCACATTAAGCTTGGCAGCTTCTGCAATGGTGAGCTCATGGGGAAGAAGCTCAACTTG

Fig. 1. Nucleotide and derived aa sequences of ApL1a (GenBank acc. # AF117570) from Arabidopsis. A nt sequence consensus motif to the translation initiation site is underlined by thin line, and that defining the termination signal for polymerase II (polyadenylation signal) is underlined by thick line. An arrow denotes the 3'-end of ApL1-1 clone, whereas the dotted line underlines nt sequence that was obtained by 3'-RACE amplification. The original earlier reported ApL1 cDNA clone (Villand et al., 1993) is denoted by nt sequence # 920-1428. Lysine residues (Lys184, Lys475, Lys512) which are homologous to those previously shown for spinach AGPase to bind pyridoxal-P (Ball and Preiss, 1994), an analog of activators of the enzyme, are circled.
Sokolov et al., 1998). The ApL1 protein shares 77–78% identity with the tomato, barley and potato large subunits of AGPase. The identity to a host of homologous sequences from other plants and/or tissues is at ca. 58–67%, including a full-length homologue from Arabidopsis (acc. # Y18432) [corresponding to ApL3 (Villand et al., 1993)]. The derived aa sequence of ApL1a contains three highly conserved lysine (Lys) residues (Lys184, Lys475, Lys512) (Fig. 1), which are thought to be critical for binding of pyridoxal-P, an analog of AGPase activators (Ball and Preiss, 1994). The two Lys residues located close to C-terminus are positioned near or at an allosteric site of the leaf enzyme. A similar positioning of Lys residues occurs in a host of other homologous large subunit sequences of AGPase [e.g. (Eimert et al., 1997)], but not in barley seed AGPase large subunit which lacks the lysines in positions corresponding to Lys184 and Lys475 (Villand et al., 1992b). The latter enzyme has been shown to be remarkably insensitive to effector regulation (Kleczkowski et al. 1993a, b).

**Motifs**

In the vicinity of the first ATG codon from the 5’-end of the ApL1a cDNA clone, a sequence homology to the translation initiation region [(A/G)CCATGG] (Kozak motif) has been identified (Fig. 1). In this sequence, the positions of cytidine bases are less conserved than those of other nucleotides. In addition, the 3’-untranslated region of ApL1a contains a motif characteristic of the termination signal for polymerase II (polyadenylation signal) (AATAAA).

Concerning derived aa sequence, we have failed to locate any transit peptide motif characteristic of a plastid-targeted nuclear-encoded proteins. Leaf AGPases are supposed to be located exclusively in plastids (Preiss, 1978) and their immature aa sequences should contain a peptide that is cleaved upon transport to chloroplast stroma. Based on empirical rules summarised in (Bairoch, 1992), the criteria for such a transit peptide in ApL1 protein are not met due to the shortage of G, A, S, T and C residues in the first 15 aa of the sequence, as well as the lack of a consensus cleavage site (I/V−X−A/C−↓−A) within first 80 aa of the ApL1a-derived protein. Interestingly, derived full length aa sequences for Arabidopsis ApL3 (acc. # Y18432) and ApS (acc. # U70616) proteins also do not have such an obvious motif for a transit peptide cleavage site. The significance of this is not entirely clear; whereas most nucleus-encoded chloroplast-located proteins do have such a discernible motif, ca. 30% proteins do not conform to this rule (Bairoch, 1992). Thus, the evidence derived from aa sequence alone is not sufficient to assign an either chloroplastic or cytosolic location for a given protein. Were Arabidopsis AGPase cytosol-located, this would imply that ADP-glucose formed in the cytosol would have to be exported to plastids for starch synthesis, similarly to the situation in cereal seeds (Villand and Kleczkowski, 1994; Denyer et al. 1996; Kleczkowski, 1996). In contrast to the Arabidopsis homologues, barley leaf AGPase subunits do contain discernible transit peptide cleavage sites, as found both by analyses of derived cDNA sequences (Thorbjørnsen et al., 1996; Eimert et al., 1997) and by expression in a heterologous system of full-length and specifically truncated cDNAs (Luo et al., 1997; Luo and Kleczkowski, 1998).

**Spatial expression and the role of ApL1 protein**

Expression of the ApL1 gene was studied using RNA fractions collected from various tissues of intact Arabidopsis (Fig. 2). The gene was expressed predominantly in rosette and stem leaves.
as well as in flowers. Some expression was also detected in stems and in fruits (Fig. 2). In the latter case, we were limited by the low level of total RNA due to contamination with polysaccharides (see rRNA levels in Fig. 2), so the actual comparative expression of ApL1 there is certainly higher than shown. The expression in these various tissues was observed both during dark and light conditions (data not shown), suggesting that the gene is ubiquitously expressed in intact Arabidopsis, plants, albeit its expression level may vary depending on conditions and/or organ studied. All the tissues that were investigated in the present study contain a considerable pool of internal starch, as found by iodine staining tests (K. Eimert, unpublished).

The ubiquitous expression of the ApL1a transcript in Arabidopsis is consistent with the role of the corresponding gene as one of the major determinants of starch production in this species. The ApL1 gene encodes a protein that is absent in leaves of a starch-deficient adg2 mutant of Arabidopsis (Wang et al., 1997). This mutant does express the mutated ApL1 transcript, but its translation is impaired resulting in a greatly reduced AGPase activity and in starch deficiency. The data, apart from underlying the major role for ApL1 with respect to starch synthesis, suggest also that ApL2 and ApL3 genes can not functionally compensate for the lack of ApL1 protein in adg2. One possibility is that these genes encode proteins that confer regulatory properties for native AGPase that are distinct from those where ApL1 protein is present. Large subunits represent critical determinants of the regulatory response of the enzyme (Ball and Preiss, 1994) and, when substituting for ApL1, they could confer a degree of effector (in-)sensitivity to the native AGPase, limiting starch production in the mutant. Other possibility is that these genes are perhaps expressed in some non-mesophyll leaf cells only (e.g. guard cells) (Sokolov et al., 1998). For further evaluation of the roles of individual AGPase genes in Arabidopsis, in situ expression approaches and studies using transgenic plants with “antisense” inhibition of either ApL2 and ApL3 will be necessary.

Acknowledgement

This research was supported, in part, by the Umeå University biotechnology fund, Swedish Natural Science Research Council, and Swedish Foundation For Strategic Research.

Müller-Röber B., Sonnewald U. and 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.


