Title: Molecular cloning and biochemical characterization of three phosphoglycerate kinase isoforms from developing sunflower (*Helianthus annuus* L.) seeds.

Authors: Troncoso-Ponce MA[a], Rivoal J[b], Venegas-Calerón M[a], Dorion S[b], Sánchez R[a], Cejudo FJ[c], Garcés R[a], and Martínez-Force E[a]*.

Authors' affiliation: [a]Instituto de la Grasa, CSIC, Avda. Padre Garcia Tejero 4, 41012 Seville, Spain. [b]Institut de Recherche en Biologie Végétale, Université de Montréal, 4101 Rue Sherbrooke est, Montréal, Que., Canada H1X 2B2. [c]Instituto de Bioquímica Vegetal y Fotosíntesis, Universidad de Sevilla y CSIC, Avda. Américo Vespucio 49, 41092 Seville, Spain.

*Correspondence to: Martínez-Force E.

Email: emforce@cica.es

Tel: 34-954611550

Fax: 34-954616790

Keywords: Phosphoglycerate kinase, Sunflower, Helianthus annuus, Seed, Oil, Fatty acids.

Abbreviations: TAG, triacylglycerides; PGK, Phosphoglycerate kinase; 3-PGA, 3-phosphoglycerate; 1,3-bPGA, 1,3-diphosphoglycerate Ni-NTA, nickel-nitrilotriacetic acid; LB, Luria Bertani; IPTG, Isopropyl β-D-1-thiogalactopyranoside; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis.

Abstract

Three cDNAs encoding different phosphoglycerate kinase (PGK, EC 2.7.2.3) isoforms, two cytosolic (*HacPGK1* and *HacPGK2*) and one plastidic (*HapPGK*), were cloned and characterized from developing sunflower (*Helianthus annuus* L.) seeds. The expression profiles of these genes showed differences in heterotrophic tissues, such as developing seeds and roots, where *HacPGK1* was predominant, while *HapPGK* was highly expressed in photosynthetic tissues. The cDNAs were expressed in *Escherichia coli*, and the corresponding proteins purified to electrophoretic homogeneity, using immobilized metal ion affinity chromatography, and biochemically characterized. Despite the high level of identity between sequences, the *Ha*cPGK1 isoform showed strong differences in terms of specific activity, temperature stability and pH sensitivity in comparison to *Ha*cPGK2 and *Ha*pPGK. A polyclonal immune serum was raised against the purified *Ha*cPGK1 isoform, which showed cross-immunoreactivity with the other PGK isoforms. This serum allowed the localization of high expression levels of PGK isozymes in embryo tissues.

1. Introduction

Phosphoglycerate kinase (PGK, EC 2.7.2.3) is an essential soluble enzyme for the basic metabolism in all living organisms. It catalyses the reversible transfer of phosphate from the C1-position of 1,3-diphosphoglycerate (1,3-DPG) to ADP to form 3-phosphoglycerate (3-PG) and ATP, in the presence of Mg^{2+}. PGK, as a typical kinase, is a monomeric enzyme with two structural domains of about equal size joined by a flexible hinge (Watson et al., 1982). The C-terminal domain binds the nucleotide, Mg-ADP or Mg-ATP, while the N-terminal domain binds 3-PG or 1,3-DPG. During catalysis the domains move toward each other to bring the substrates into the proper vicinity for the transfer of the phosphoryl group (Zerrad et al., 2011). In higher plants there are two classes of phosphoglycerate kinase, one is localized in the cytosol (cPGK), and the other in plastids (pPGK) (Anderson and Advani, 1970). The cytosolic enzyme functions primarily in glycolysis for substrate-level ATP generation, one of the two glycolytic steps that produce ATP (Plaxton, 1996), and is also involved in gluconeogenesis. The plastidial isoform is involved in the photosynthetic carbon reduction cycle and in chloroplast glycolysis. In plants, both isozymes generally show a high degree of conservation (Longstaff et al., 1989). The most significant difference in their sequences is the presence of a chloroplast transit peptide at the N-terminus of the chloroplastic isoform. In tobacco leaves, the genes encoding these isozymes are differentially expressed in a developmental and tissue-specific manner and both genes show light-modulated expression *in vivo* (Bringloe et al., 1996). In addition, both isozymes appear to be involved in the regulation of carbon partitioning between starch and sucrose (Shah and Bradbeer, 1994). Additionally, a sub-population of cytosolic PGK has been localized in the nucleus in pea leaves and shoots, due to the presence of a functional nuclear localization signal near its N-terminus (Brice et al., 2004). Interestingly, it has been hypothesized that PGK has

more than one role (Jeffery, 1999) and it has been regarded as a "moonlighting" protein. Several lines of evidence support this view, as PGK participates in DNA replication and repair in both the mammalian (Jindal and Viswanatha 1990; Kumble and Viswanatha 1991; Viswanatha et al., 1992) and the plant nucleus (Al-Rashdi and Bryant 1994; Burton et al., 1997, Bryant and Anderson 1999, Bryant et al., 2000), and is involved in *in vitro* mRNA synthesis of the Sendai virus (Ogino et al., 1999; Popanda et al., 1998).

During the synthesis of reserve lipids in heterotrophic oilseeds, sucrose is converted into hexose-phosphates and metabolized by the glycolytic pathway to produce carbon skeletons, energy and reducing power necessary for lipid biosynthesis (Schwender and Ohlrogge, 2002; Hill et al., 2003). In developing sunflower (*Helianthus annuus* L.) embryos, one of the glycolytic activities that increases in conjunction with the oil content during the period of active storage lipid synthesis is PGK (Troncoso-Ponce et al., 2009). Moreover, the study of the glycolytic enzymatic activities in developing seeds from standard and low oil content sunflower lines, has pointed to phosphoglycerate kinase and enolase as two of the activities probably implicated in the differences in fat content between these lines (Troncoso-Pone et al., 2010a). In addition, labelling and modelling studies have shown that the major source of carbon for the synthesis of fatty acids in sunflower plastids originates from the metabolism of triose phosphates to phospho*enol*pyruvate (Alonso et al., 2007). PGK activity is involved in this sequence of reactions, which can take place in the cytosolic and plastidial compartments (Plaxton and Podestá, 2006; Alonso et al., 2007). It is also one of the highest glycolytic activities measured in rapeseed (*Brassica napus*) embryo cultures (Junker et al., 2007) and has been identified in the group of the most abundant proteins in castor bean (*Ricinus communis*) developing seeds (Houston et al., 2009), both important crops characterized by high rate of fatty acid synthesis in

their seeds . In transgenic canola seeds with increased oil content, expression of a subset of genes involved in fatty acid biosynthesis and glycolysis is up-regulated in developing seeds, including PGK (Tan et al., 2011). All these observations suggest that this activity may participate in the generation of energy and carbon for the fatty acid synthesis.

In order to better understand the carbon metabolism in sunflower developing seeds, we have focused on the characterization of cytosolic and plastidial PGK isoforms. In this study, we have identified three cDNAs from developing sunflower seeds, which are predicted to encode two cytosolic and one plastidic PGK. We describe their phylogenetic relationship with other homologous genes, and perform model predictions for the corresponding protein structures. The biochemical characterization of the recombinant proteins provides evidence of differences in kinetic properties between PGK isoforms. A survey of the expression pattern of these genes and the immunolocalisation of PGK isoforms in developing embryos supports the view that sunflower PGKs are differentially regulated.

2. Results and discussion

2.1. Isolation and sequence analysis of three sunflower PGK cDNAs

Conserved regions from known phosphoglycerate kinase sequences were used to design degenerated oligonucleotide primers dPGK-F and dPGK-R (Table 1). Using these primers, a 575 bp fragment was PCR amplified from developing sunflower seeds cDNA, which corresponded to an internal region of PGK mRNAs. Subsequently, three full-length cDNA clones were obtained by RACE using the primers shown in the Materials and Methods. These PCR fragments were cloned and sequenced, and their homology to other plant PGKs was confirmed using the Blast software (Altschul et al., 1990). The cDNA sequences were identified as two cytosolic PGK isoforms, *Ha*cPGK1 (1206 bp) and *Ha*cPGK2 (1203 bp), and a plastidial one *Ha*pPGK (1446 bp).

The full-length PGK cDNAs were predicted to encode proteins of 401 amino acids (*Ha*cPGK1, predicted molecular mass of 42.3 kDa and pI of 6.06), 400 amino acids (*Ha*cPGK2, predicted molecular mass of 42.4 kDa and pI of 7.61) and 481 amino acids (*Ha*pPGK, predicted molecular mass of 50.1 kDa and pI of 7.18). The *Ha*pPGK sequence contained a predicted plastid transit peptide of 76 residues at its N-terminus, not present in the two cytoplasmic isoforms, which would produce a 405 amino acids mature polypeptide of 42.2 kDa and pI of 5.16. The alignment of *Ha*PGK proteins from sunflower (Eudicotyledon subclass Asterids family Asteraceae) along with other PGKs from *Arabidopsis thaliana* (Eudicotyledon subclass Rosids family Brassicaceae), cytosolic and chloroplastic forms, and *Bacillus stearothermophilus* (Bacteria phylum Firmicutes family Bacillaceae), showed a high degree of identity (Fig. 1).

Using the sequences reported here and other known PGK sequences, either cytoplasmic or plastidial isoforms, from Viridiplantae, a phylogenetic tree was generated using the PGK

sequence from the bacteria firmicute *Geobacillus stearothermophilus* as an outgroup to root the tree (Fig. 2). This comparison indicated that *Ha*cPGK1 and *Ha*cPGK2 grouped with the other plant cPGKs and were more closely related to PGKs of Solanaceae, like *Nicotiana tabacum*, than to other known PGKs, although PGKs have been highly conserved throughout evolution. Moreover, *Ha*pPGK grouped with other dicot plant PGK plastidial isoforms. The species distribution was essentially consistent with conventional species trees, with Monocots (Poaceae species) and Dicots species differentiating from chlorophytas, mosses and bacteria.

2.2. Tertiary structure prediction of HaPGK proteins

Although crystal structure data for plant PGKs are not yet available, the sequence from *Ha*PGK proteins showed enough identity, around 60%, with the *Bacillus stearothermophilus* PGK protein, the structure of which is well-known (PDB 1PHP). The use of protein modelling software programs based on the homology with proteins of a known crystalline structure (Swiss-Model, Schwede et al., 2003; 3D-JIGSAW, Bates et al., 2001) allowed us to obtain structural models for sunflower PGK proteins. Since the models for *Ha*cPGK1 and *Ha*cPGK2 were very similar, only the models for *Ha*cPGK1 and *Ha*pPGK are shown in Fig. 3. The predicted structures show that both proteins are typical kinases (Watson et al., 1982) with two structural domains joined by a waist region (Fig. 3), each domain binds one of the two substrates, 3-phosphoglycerate (3-PG) and ADP. For the phospho-transfer reaction to take place the substrates must be brought closer by a hinge-bending domain closure. The polypeptide chains are distinctly folded into two domains of unequal size (Fig. 3): the large and small domains separated by a deep cleft containing the residues involved in the enzyme active site. In the secondary structure, 15 α-helices and 17 β-strands are predicted for *Ha*cPGK1 and 14 α-helices and 19 β-strands for

*Ha*pPGK (Fig. 3 A and C). The 3-phosphoglycerate will fit into the cleft between the two PGK domains, interacting with the residues D25, N27, R41, H64 and R122 in the case of *Ha*cPGK1 and with residues D99, N101, R115, H138 and R196 in the case of *Ha*pPGK (white arrow heads in Fig. 1 and residues in green in Fig. 3B and D) forming the substrate binding site. Similarly, it is predicted that ADP will interact with residues G225, G297, N321, P323, G325, V326, F327, E328, G358, D359 and S360 in the case of *Ha*cPGK1 and with residues G299, G371, N395, P397, G399, V400, F401, E402, G432, D433 and S434 in the case of *Ha*pPGK (black arrow heads in Fig. 1 and residues in pink in Fig. 3B and D) those residues forming the ADP binding site in pink and the catalytic site in blue). PGK hinge regions, allowing flexibility of relative domain positions (open and closed state), are formed by residues V187, S188, N189, P190, I376, S377 and T378 in the case of *Ha*cPGK1 and with residues V261, S262, N263, P264, I450, S451 and T452 in the case of *Ha*pPGK (grey arrow heads in Fig. 1). These models identify residues likely to be important in the catalytic function of the enzyme that are potential targets for further studies involving site-directed mutagenesis and kinetic characterisation of the resultant enzymes.

2.3. Expression and purification of recombinant HacPGK

The proteins encoded by the three PGK cDNAs were expressed in *E. coli* with a (6His) tag at the N-terminus and purified by NTA chromatography (Fig. 4). The purified proteins (lanes 2, 4 and 6) were > 95% homogenous as judged from Coomassie staining and image analysis and showed apparent MWs around 45 kDa close to the predicted MWs of the recombinant proteins encoded by the expression plasmid. (6His)*Ha*cPGK1 was purified to a specific activity of 254.4 nkat mg^{-1} protein at pH 8.2 (Table 2). This value is comparable to that reported for recombinant *Plasmodium falciparum* PGK (Pal et al., 2004), but 4-fold lower than

affinity-purified native barley leaf cPGK (MacMorrow and Bradbeer, 1990). While the specific activities obtained with (6His)*Ha*cPGK2 and (6His)*Ha*pPGK were over 5000 nkat mg^{-1} protein (Table 2). Highlighting, especially in the case of the two cytosolic isoform, the fact of how very similar proteins are able to show such striking kinetic differences, although the possibility of a misfolding in the case of (6His)*Ha*cPGK1 cannot be ruled out.

2.4. Thermal stability of recombinant PGKs

To investigate the temperature stability of these enzymes, the purified recombinant PGKs were subjected to 3 min treatments at various temperatures ranging from 20 to 96^{o}C (Fig. 5). Treatments were conducted in the presence of glycerol, a co-solvent that inhibits heat-induced aggregation for yeast PGK (Strucksberg et al., 2007). Near maximum activity was retained by all the enzyme preparations incubated from 22 to 35^{o}C. (6His)*Ha*cPGK1 was clearly the most temperature sensitive of the three enzymes tested, as it underwent a rapid thermal denaturation at temperatures between 40 and 65^{o}C. Activity was completely lost with treatments over 65^{o}C. The two other recombinant enzymes showed similar thermal stability, retaining about 80% activity at temperatures up to 65ºC, then rapidly decreasing between 65 and 80ºC (Fig. 5).

2.5. Kinetic characteristics of recombinant PGKs

The activity of purified recombinant PGKs was monitored in the pH range between 5.75 and 9.25 (Fig. 6). All the enzymes displayed a broad optimum pH between 7.5 and 8.75, similar to that reported for barley leaf and spinach leaf PGKs (Köpke-Secundo et al., 1990; MacMorrow and Bradbeer, 1990). Their activity remained relatively high at alkaline pHs with more than 75% activity retained at pH 9.2 in all cases. At pH values below 7, a sharp reduction of

(6His)*Hac*PGK1 activity was observed, in contrast to the two other enzymes. Interestingly, (6His)*Hac*PGK2 activity was not significantly affected by pH. The sensitivity of (6His)*Hac*PGK1 to pH below 7 would render the enzyme responsive to acidic conditions in the cytoplasm. Cytoplasmic acidosis has been shown to occur in conditions of hypoxia (Roberts et al., 1984; Drew, 1997; Gout et al., 2001), which is common in seed maternal and embryonic tissues (Rolletschek et al., 2002; Borisjuk et al, 2004; Borisjuk and Rolletschek, 2009).

Substrate saturation kinetics for 3-PGA and ATP substrates were studied at three physiologically relevant pH values (7.0 , 7.6 and 8.2) (Table 2). All enzymes displayed Michaelis-Menten kinetics with 3-PGA and ATP at all pH values. ATP saturation kinetics showed that both *V*max and *K*m values were stable across all pHs tested for the three enzymes. Consistent with the specific activity data, *V*max values were always much higher for (6His)*Hac*PGK2 and (6His)*Ha*pPGK. We did not observe biphasic kinetics or inhibition at high ATP concentrations as previously reported for the spinach leave enzymes (Köpke-Secundo et al., 1990). The affinity of (6His)*Hac*PGK1 for 3-PGA was comparable to values previously reported for *Trypanosoma brucei* PGKs (Zomer et al., 1998) but was 2.4 to 3 times lower than the values obtained with purified spinach leave PGKs (Köpke-Secundo et al., 1990) and also with the two other recombinant *H. annuus* enzymes characterized here. The *K*m values obtained using ATP as varying substrate were comparable to previously reported values (Köpke-Secundo et al., 1990; Zomer et al., 1998). Interestingly, (6His)*Hac*PGK1 displayed the higher affinity for this substrate compared to the two other recombinant *H. annuus* PGKs. Under the conditions tested here, the calculated catalytic efficiencies (*V*max/*K*m) of the recombinant enzymes were not strongly affected by pH.

The effects of glycerol-2-P and glycerol-3-P were tested on all three PGK enzymes because these two metabolites are known inhibitors of animal PGKs (Szilagyi and Vas, 1998). However, the activity of three enzymes was unaffected by these compounds, even in the presence of relatively high concentrations (Table 3). Likewise, UDP (10 mM) did not significantly affect the activity of the recombinant enzymes. The inclusion of ADP and AMP in the assay mixture resulted in a significant decrease in activity for all PGKs (Table 3). The effect was more pronounced with ADP, with concentrations as low as 0.1 mM resulting in up to 40% inhibition. These data indicate that sunflower seed PGKs could be sensitive to variations in the ATP/ADP ratio or the adenylate energy charge ([ATP] + 0.5 [ADP])/ ([ATP] + [ADP] + [AMP]). Such behaviour was described before for pig muscle PGK (Molnar and Vas, 1993; Kóvari et al., 2002) and yeast PGK (Larsson-Raznikiewicz and Arvidsson, 1971). This feature could be physiologically relevant for the regulation of the seed enzyme in the gluconeogenic direction because high ADP status is commonly observed under hypoxic conditions known to prevail in seed tissues (Shelp et al., 1995; Borisjuk and Rolletschek, 2009).

2.6. Expression of HacPGK1, HacPGK2 and HapPGK genes

Expression levels of the three PGK genes in different tissues have been studied (Fig. 7). *HacPGK1* showed the highest level of expression during seed development, especially at the initial stages, 12 days after flowering (DAF), and accounted for most (approx. 89%) of the combined expression of the three genes. As seed development progressed, the level of expression for this gene decreased. The temporal expression pattern of *HacPGK1* followed a declining trend similar to other genes involved in carbohydrate and lipid metabolism described in sunflower seed and recently in four different oilseeds (Troncoso-Ponce et al., 2010b, 2011a and 2011b).

Transcripts of the other two genes, *HacPGK2* and *HapPGK*, remained low in all the stages studied though, as *HacPGK1* decreased the proportional participation of *HacPGK2* and *HapPGK* in the combined pool of transcripts increased.

The comparison of *Ha*cPGK1, *Ha*cPGK2 and *Ha*pPGK expression levels in different tissues (Fig. 7A) shows that *Ha*cpgk1 gene is strongly expressed in the initial stages of the seed's development. Meanwhile, highest expression of *Ha*pPGK in photosynthetically active tissues, cotyledons and leaves, might reflect the role of this activity in the Calvin cycle. It is relevant to compare these data to the expression data of *Arabidopsis thaliana* PGK genes (Fig. 7B) taken from published microarrays analyses (Schmid et al., 2005), since both plants are model systems for the accumulation of seed lipid reserves. *A. thaliana* contains one cytosolic PGK (At1g79550) and two PGKs predicted to be targeted to the plastid (At1g56190 and At3g1270). The expression patterns of *H. annuus* and *A. thaliana* cPGKs are broadly similar during seed development, with higher levels at early stages. However, in *A. thaliana*, the difference between the early and the later stages of cPGK expression was not as great. Expression of genes encoding pPGKs was lower than that of cPGK in *A. thaliana*, as was observed in *H. annuus.* Interestingly, in both systems, high levels of PGK expression coincided with the onset of lipid deposition in the developing seed. This process occurs around 10-12 DAF in sunflower (Troncoso-Ponce et al., 2009, 2010b) and around the torpedo stage (approximately 4-5 DAF) in *A. thaliana* (Becerra et al., 2006). Expression of *A. thaliana* genes encoding cPGKs and pPGKs in various tissues (Fig. 7B) was also relatively similar to observations made in sunflower, with cPGK being more predominant in roots and cPGKs more abundantly expressed in leaves.

2.7. PGK protein and activity profile during seed development.

A polyclonal antibody was generated against (6His)*Ha*cPGK1, which efficiently recognized all PGK isoforms as shown by Western blot analysis of the recombinant proteins (Fig. 8A). This immune serum was used to analyse PGK protein profiles in extracts of developing seeds (Fig. 8B). A single band was observed in Western blot analysis of developing seed extracts. The intensity of the band appeared to decline slightly between 15 and 20 DAF, to rise at later developmental stages. Total PGK activity was measured in extracts of developing seed harvested between 10 and 26 DAF (Fig. 8C). PGK activity was low at early stages of seed development (10-14 DAF) and increased steadily to reach a plateau at 18 DAF, at which point PGK activity remained relatively constant up to 26 DAF. This activity profile paralleled the accumulation of fatty acids in seeds (Troncoso-Ponce et al., 2009). The immunoblot profiles were compared to total PGK activity profiles in developing seeds (Fig. 8C). It was apparent from this comparison that the weakest signal observed on immunoblot corresponded to a high level of PGK activity suggesting that the regulation of total PGK activity was complex, and could involve, for example, a differential contribution of the more active isoforms, *Ha*cPGK2 and *Ha*pPGK throughout the developmental process.

2.8. Immunocytochemical localisation of PGK isoforms in developing seeds

Localization of cells with the highest PGK isoforms content has been achieved. Immunohistochemistry was used to study the spatial distribution of PGKs in developing 15DAF sunflower seed tissues (Fig. 9). The polyclonal antiserum generated against purified sunflower 6(His)*Ha*cPGK1 could recognize the three PGK isoforms studied, allowing the visualization of those cells with maximum presence of these enzymes. The spatial distribution of PGK proteins overlaps with the localization of the cytosolic phosphoglucose isomerase (cPGI), a glycolytic

enzyme previously studied in developing sunflower seeds (Troncoso-Ponce et al., 2010b). The fact that both enzymes are found in all cell types is not surprising. However, stronger signals were observed in the procambial ring and especially in the case of PGK at the most external cell layers of the proximal end of the seed.

3. Concluding remarks

In spite of the central role play by the glycolysis pathway during fatty acid synthesis in oilseed and in other plant tissues (Bourgis et al., 2011; Troncoso-Ponce et al., 2011b), little information is available about the characteristics of the individual enzymes. Although the study of *E. coli* plant recombinant expressed enzymes may not have identical kinetics of a plant expressed proteins, we have tackled the characterization of different isoforms of PGK. Our data show that at least two cytosolic and one plastidial PGK are differently expressed during sunflower seed development and within different plant tissues. The three genes encode proteins that have distinct physical properties and analyses of the recombinant proteins in E. coli suggest that these isozymes have different kinetic characteristics. Thus the changes in total PGK activity measured during the development of sunflower seeds represent the complex pattern of variation in both the expression and kinetic properties of these isozymes.

4. Experimental

4.1. Plant culture conditions, biological materials and chemicals

Sunflower (*Helianthus annuus* L.) line CAS-6, with normal fatty acid content, was used in this work. Plants were cultivated in growth chambers at 25/15 ºC (day/night), with a 16-h photoperiod, and a photon flux density of 150 µmol m^{-2} s^{-1}. Fertilization with Bayfolan S (Bayer) was done using fertirrigation lines. Seeds from 10 to 26 DAF from the external seed rings of the capitulum were harvested for analysis. *Escherichia coli* (XL1-Blue strain) was used as plasmid host for cloning and protein expression. All primers were synthesized by MWG Biotech AG (Ebersberg, Germany). Bacteria were grown at 37ºC in LB media (1% Bacto Tryptone, 0.5% Bacto Yeast Extract, 1% NaCl, pH 7). When appropriate, ampicillin (100 µg/mL) was added for plasmid selection. Except when mentioned otherwise, buffers, chemicals and reagents were of analytical grade from Sigma Chemical Co. (St Louis, MO) or Fisher Scientific (Nepean, ON, Canada).

4.2. Cloning of the cDNAs encoding three sunflower PGKs

Approximately 0.4 g of developing sunflower seeds was harvested at 15 DAF. Seeds were ground in liquid N_2 with a precooled sterile mortar and pestle and mRNA was isolated using the MicroFastTrack Kit (Invitrogen, Groningen, The Netherlands). The mRNA pellet was resuspended in 33 µl TE buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8) and the cDNA was obtained using a Ready-To-Go T-Primed First Strand Kit (Amersham Bioscience, Roosendaal, The Netherlands). PGK protein sequences from public databases were aligned using the ClustalX v.2.0 program (Thompson et al. 1997) to identify highly conserved regions. PCR fragments were amplified with two degenerate primers designed from these regions: dPGKF and dPGKR (Table

1). The fragments were cloned into the pGEM-T-Easy® vector (Promega) and several clones were sequenced on both strands by GATC Gmbh (Konstanz, Germany). The identity of the clones was confirmed using the BLAST software (Altschul et al. 1990), identifying three different clones, *HacPGK1*, *HacPGK2* and *HapPGK*. The 5′- ends were obtained using the Smart™- RACE cDNA amplification kit (Clontech) and specific reverse internal oligonucleotides pairs for each cloned sequence: *HacPGK1*-R1 and *HacPGK1*-R2, *HacPGK2*-R1 and *HacPGK2*-R2, and *HapPGK*-R1 and *HapPGK*-R2 (Table 1). The 3′-end of the cDNAs was obtained by PCR using the external oligo FA2Z (Table 1), complementary to the sequences incorporated during the initial cDNA synthesis, and specific internal oligos for each cloned sequence: *HacPGK1*-F, *HacPGK2*-F and *HapPGK*-F (Table 1). The PCR fragments were cloned, sequenced and assembled to obtain DNA sequences coding for two PGK cytosolic isoforms, *HacPGK1* (1206 bp) and *HacPGK2* (1203 bp), and a plastidial one *HapPGK* (1446 bp). These cDNA sequences were deposited in GenBank under accession numbers DQ835564, HM490307 and HM490308, respectively.

4.3. cDNA and protein sequence analyses

Sequences homologous to the predicted sequences of sunflower PGKs were retrieved using the BLASTP program (www.ncbi.nlm.nih.gov). Alignment of the amino acid sequences, including the transit peptides, for PGK proteins deposited at GENBANK was performed using the ClustalX v.2.0 program with the default settings (Thompson et al. 1997). These entire alignments were used to generate a phylogenetic tree based on the neighbor-joining algorithm (Saitou and Nei 1987), and the resulting ‘phenogram’ was drawn using the MEGA 4.0 program (Tamura et al. 2007).

4.4. Modelling of the three-dimensional structure of sunflower PGKs.

Homology modelling studies were performed using Swiss Model server (http://swissmodel.expasy.org/) (Schwede et al. 2003) and JPred prediction server (Cuff and Barton 2000). The sequence used as a template was a bacterial PGK from *Bacillus stearothermophilus*, accession number P18912 (gi129919). The chosen template was the most homologous PGK for which X-ray structure information is available (PDB Entry: 1php; Davies et al. 1994), showing 60.6% sequence identity with 392 residues of *Hac*PGK1. SWISS-MODEL was used in first approach and project (optimise) modes using default parameters. Structures were visualized using Swiss-PDBViewer (Guex and Peitsch 1997).

4.5. Enzyme extraction from plant tissues and protein assays

Typically, 5 peeled sunflower seeds were ground in 1 mL of 50 mM Tris-HCl (pH 8.0), 5 mM DTT using an ice-cooled glass homogenizer. The resulting homogenate was centrifuged for 10 min at 10,000 × *g*. The supernatant was used immediately for enzymatic assays. Protein concentration was determined according to Bradford (1976), using the Bio-Rad protein assay reagent (Bio-Rad Laboratories, Mississauga, ON, Canada) and bovine gamma globulin as standard.

4.6. Constructs for recombinant sunflower PGKs expression in E. coli

Primers with internal *Sph*I and *Pst*I restriction sites were designed, *SphHacPGK1*-F and *PstHacPGK*1-R respectively, to amplify the entire coding sequence of *HacPGK1* by PCR (Table 1). The PCR product obtained was subcloned into the *Sph*I-*Pst*I sites of pQE-80L (Qiagen, Hilden, Germany) to produce a fusion protein with a hexahistidine tag at the N terminus. In a

similar way, full *HacPGK2* cDNA was cloned into pQE-80L as a *Sph*I-*Sal*I fragment using *SphHacPGK2*-F and *SalHacPGK2*-R pair of primers (Table 1) and *HapPGK* cDNA was cloned into pQE-80L as a *Sph*I-*Pst*I fragment missing the signal peptide using *SphHapPGK*-F and *PstHapPGK*-R primers (Table 1). Ligation into the correct reading frame was confirmed by sequencing and the resulting constructs were designated pQEcPGK1, pQEcPGK2 and pQEpPGK, respectively. The recombinant plasmids were introduced and expressed in *E. coli* strain XL1-Blue. The predicted molecular mass for each recombinant protein 6(His)*Hac*PGK1, 6(His)*Hac*PGK2 and 6(His)*Hap*PGK was 43.1, 43.2 and 43.0 kDa, respectively.

4.7. Heterologous protein expression and purification

E. coli cells harbouring recombinant plasmids, pQEcPGK1, pQEcPGK2 and pQEpPGK were grown under continuous shaking at 37ºC in LB broth containing ampicillin. The cells were induced at OD_{600} 0.5 with 0.6 mM IPTG, and grown for an additional 2 h at 37ºC. Cells were harvested by centrifugation (10 min at 10,000 x*g*), and pellets were frozen at -80°C until used. Purification steps were carried out at 4°C as previously described (Dorion et al., 2006). Protein concentration was determined according to Bradford (1976) as described above. Glycerol was added to the purified enzyme preparation at a final concentration of 50% (v/v) and the solution was stored at -20ºC until used.

4.8. PGK activity assay and kinetic analyses

PGK activity assays were conducted according to a protocol modified from Journet et al. (1986). The PGK reaction was coupled to the glyceraldehyde-3-phosphate dehydrogenase (GAPDH, EC 1.2.1.12) reaction and assayed at 25 ºC by monitoring NADH oxidation at 340 nm

using a VersaMax (Molecular Devices, Sunnyvale, CA, USA) microplate reader. The 200 µL standard reaction mixture contained 100 mM Tris-HCl, pH 7.8, 5 mM $MgCl_2$, 0.2 mM NADH, 2 mM ATP, 2 mM DTT, 5 mM 3-phosphoglycerate and 0.5 U/mL GAPDH. Reaction rates were linear with time and proportional to the amount of enzyme added to the assay within a range spanning one order of magnitude. To study the effect of glycerol-3-phosphate, glycerol-2-phosphate UDP, ADP and AMP on PGK activity, these metabolites were added to the standard assay mix described above. For all analyses, presented data are means ± SD of determinations carried out with three independent enzyme preparations and quadruplicate assays unless otherwise mentioned. Statistical analysis of the kinetic data was performed using Student's t test (SigmaPlot 8.0, SPSS, Chicago, IL, USA), with $P<0.05$ considered significant.

The temperature stability of the purified sunflower recombinant PGKs was tested by incubating 20 µl samples of pure enzyme for 3 min at various temperatures using a water bath. After heat treatment, the samples were cooled on ice for 3 min. Residual PGK activity was then immediately assayed as described above. The effect of pH on sunflower recombinant PGKs activity was studied using a three-component buffer to maintain a constant ionic strength throughout the pH range (Ellis and Morrison, 1982). The Tris-HCl buffer used in the standard reaction mixture was therefore replaced with a mixture of 0.05 M acetic acid, 0.05 M 2-(N-morpholino)-ethanesulfonic acid (MES) and 0.1M Tris-HCl adjusted at different pH values with 1 M NaOH or HCl. The pH value assay was measured directly in the reaction mixture using a microelectrode immediately after completion of the spectrophotometric assay. Assays were initiated by addition of enzyme preparation and corrected for background activity by omitting 3-phosphoglycerate from the reaction mixture. Apparent *K*m and *V*max values were calculated

from the Michaelis-Menten equation using a non-linear least-squares regression program (SigmaPlot 8.0, SPSS, Chicago, IL, USA).

4.9. Quantitative real time PCR

cDNAs from developing seeds, roots, stems, green cotyledons and leaves were obtained as previously described. The cDNAs were subjected to quantitative real time PCR (QRTPCR) with specific pairs of primers (Table 1; Q*HacPGK1*-F and Q*HacPGK1*-R for *HacPGK1*; Q*HacPGK2*-F and Q*HacPGK2*-R for *HacPGK2*; and Q*HapPGK*-F and Q*HapPGK*-R for *HapPGK*) and using SYBR Green I (QuantiTect® SYBR® Green PCR Kit, Qiagen, Crawley, UK) in a MiniOpticon system to monitor the resulting fluorescence (Bio-Rad). The reaction mixture was heated to 50ºC for 2min and then to 95 ºC for 15 min before subjecting it to 40 PCR cycles consisting of: 94ºC for 15 s; 60.5ºC for 30 s; and 72ºC for 15 s. Calibration curves were drawn up using sequential dilutions of cDNA. The Livak method (Livak and Schmittgen, 2001) was applied to calculate comparative expression levels between samples and the sunflower actin gene *HaACT1* (GenBank accession number FJ487620) was used as the reference gene using a specific pair of primer (Table 1; Q*HaActin*-F4 and Q*HaActin*-R4).

4.10. HacPGK1 immune serum

To generate antibodies against purified sunflower recombinant *Hac*PGK1, approximately 1 mg of protein was subjected to preparative SDS-PAGE, followed by electroelution (Dorion et al. 2006). After collection of the preimmune serum, a polyclonal antiserum was raised in a New Zealand SPF rabbit at the Centre of Animal Production and Experimentation, University of

Seville (Spain), following standard methods. The crude immune serum was used for immunoblots and immunocytochemical analysis.

4.11. SDS-PAGE and immunoblot analysis

Proteins, previously separated by SDS-PAGE on a gel containing 16.5% (w/v) polyacrylamide (Schagger and Vonjagow 1987), were transferred to a nitrocellulose membrane (Sigma-Aldrich, 0.45 µm pore size) using an XCell II mini-Cell system (Novex) and Tris-Glycine buffer (12 mM Tris base, 96 mM glycine, 20% methanol). The transfer was carried out at 25 V for 2 h. The membrane was stained with a Ponceau solution (Sigma-Aldrich) for 15 min to verify the quality of the transfer and then washed with TBST (0.8% NaCl, 20 mM Tris-HCl pH 7.6, 0.1% Tween 20). For immunodetection, blots were incubated with anti-*Ha*cPGK immune serum (1/5,000 dilution). Polypeptides were detected using an anti-rabbit IgG-Peroxidase (1/10,000 dilution) from Sigma-Aldrich. A Pierce ECL Western Blotting Substrate kit was used for the detection of antigen-antibody complexes, according to the manufacturer's instructions. Alternatively, an anti-rabbit IgG coupled to alkaline phosphatase (Promega, Nepean, ON, Canada) was used as described in Dorion et al. (2005).

4.12. Immunocytochemical analysis

Freshly cut organs from 15 DAF developing seeds were immediately fixed by incubation in FAE (3.7% formaldehyde: 5% acetic acid: 50% ethanol, by vol.) with occasional vacuum, dehydrated in a graded series of aqueous ethanol solutions and embedded in Paraplast Plus (Sigma Chemical Co.) as described in González et al. (1998). 10 µm sections were cut with a Leica RM 2025 microtome and placed on poly-L-lysine coated microscope slides. After

deparaffinizing in xylol and rehydrating in decreasing concentrations of ethanol, sections were blocked for 3 h in TBS buffer containing 1% (w/v) BSA. Anti-*Ha*cPGK1 immune serum or pre-immune serum, (diluted 1:5,000 in TBS) was added to the samples and incubated overnight at 4ºC. Unbound primary antibodies were removed by 3 x 10 min washes in TBS. Tissue sections were then incubated with alkaline phosphatase-conjugated goat anti-rabbit IgG for 2 h at 37ºC. The reaction of alkaline phosphatase was developed with nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl-phosphate.

Acknowledgements

This work was supported by Ministerio de Ciencia e Innovación and FEDER, project AGL2011-23187 and by Advanta Seeds Company. This work was supported in part by a discovery grant from the Natural Sciences and Engineering Research Council of Canada to J. Rivoal.

References

Alonso, A.P., Goffman. F.D., Ohlrogge, J.B., Shachar-Hill, Y., 2007. Carbon conversion efficiency and central metabolic fluxes in developing sunflower (*Helianthus annuus* L.) embryos. Plant J. 52, 296-308.

Al-Rashdi, J., Bryant, J.A., 1994. Purification of a DNA-binding protein from a multi-protein complex associated with DNA polymerase-α in pea (*Pisum sativum*). J. Exp. Bot. 45, 1867-1871.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403-410.

Anderson, L.E., Advani, V.R., 1970. Chloroplast and cytoplasmic enzymes: three distinct isoenzymes associated with the reductive pentose phosphate cycle. Plant Physiol. 45, 583-585.

Anderson, L.E., Carol, A.A., 2005. Enzyme co-localization in the pea leaf cytosol: 3-P-glycerate kinase, glyceraldehyde-3-P dehydrogenase, triose-P isomerase and aldolase. Plant Sci. 169, 620-628.

Bates, P.A., Kelley, L.A., MacCallum, R.M., Sternberg, M.J.E., 2001. Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. Proteins 45, 39-46.

Becerra, C., Puigdomenech, P., Vicient, C.M., 2006. Computational and experimental analysis identifies Arabidopsis genes specifically expressed during early seed development. BMC Genomics 7, 38.

Borisjuk, L. and Rolletschek, H., 2009. The oxygen status of the developing seed. New Phytol. 182, 17-30.

Borisjuk, L., Rolletschek, H., Radchuk, R., Weschke, W., Wobus, U., Weber, H., 2004. Seed development and differentiation: a role for metabolic regulation. Plant Biol. 6, 375-386.

Bourgis, F., Kilaru, A., Cao, X., Ngando-Ebongue, G.F., Drira, N., Ohlrogge, J.B., Arondel, V., 2011. Comparative transcriptome and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning. Proc. Natl. Acad. Sci. U S A. 26, 12527-12532.

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248-254.

Brice, D.C., Bryant, J.A., Dambrauskas, G., Drury, S.C. and Littlechild, J.A., 2004. Cloning and expression of cytosolic phosphoglycerate kinase from pea (*Pisum sativum* L.). J. Exp. Bot. 55, 955-956.

Bringloe, D.H., Rao, S.K., Dyer, T.A., Raines, C.A., Bradbeer, J.W., 1996. Differential gene expression of chloroplast and cytosolic phosphoglycerate kinase in tobacco. Plant Mol. Biol. 30, 637-640.

Bryant, J.A., Anderson, L.E., 1999. What's a nice enzyme like you doing in a place like this? A possible link between glycolysis and DNA replication, in: Bryant, J.A., Burrell, M.M., Kruger, N.J. (Eds.), Plant carbohydrate biochemistry. Bios Scientific Publishers, Oxford, pp. 295-304.

Bryant, J.A., Brice, D.C., Fitchett, P.N., Anderson, L.E., 2000. A novel DNA-binding protein associated with DNA polymerase-a in pea stimulates polymerase activity on frequently primed templates. J. Exp. Bot. 51, 1945-1947.

Burton, S.K., Van't Hof, J., Bryant, J.A., 1997. Novel DNA-binding characteristics of a protein associated with DNA polymerase-α in pea. Plant J. 12, 357-365.

Cuff, J.A., Barton, G.J., 2000. Application of enhanced multiple sequence alignment profiles to improve protein secondary structure prediction. Proteins: Struct. Funct. Genet. 40, 502-511.

Davies, G.J., Gamblin, S.J., Littlechild, J.A., Dauter, Z., Wilson, K.S., Watson, H.C., 1994. Structure of the ADP complex of the 3-phosphoglycerate kinase from *Bacillus stearothermophilus* at 1.65 Å. Acta Crystallogr. D Biol. Crystallogr. 50, 202-209.

Dorion, S., Matton, D.P., Rivoal, J., 2006. Characterization of a cytosolic nucleoside diphosphate kinase associated with cell division and growth in potato. Planta 224, 108–124.

Dorion, S., Parveen, J.J., Matton, D.P., Rivoal, J., 2005. Cloning and characterization of a cytosolic triosephosphate isomerase developmentally regulated in potato leaves. Plant Sci. 168, 183-194.

Drew, M.C., 1997. Oxygen deficiency and root metabolism: Injury and acclimation under hypoxia and anoxia. Annu. Rev. Plant Physiol. Plant Mol. Biol. 48, 223-250.

Ellis, K.J., Morrison, J.F., 1982. Buffers of constant ionic strength for studying pH-dependent processes. Methods Enzymol. 87, 405-426.

González, M.C., Osuna, L., Echevarria, C., Vidal, J., Cejudo, F.J., 1998. Expression and localization of phosphoenolpyruvate carboxylase in developing and germinating wheat grains. Plant Physiol. 116, 1249-1258.

Gout, E., Boisson, A.M., Aubert, S., Bligny, R., 2001. Origin of the cytoplasmic pH changes during anaerobic stress in higher plant cells. Carbon-13 and phosphorus-31 nuclear magnetic resonance studies. Plant Physiol. 125, 912–925.

Guex, R.E.F., Peitsch, M.C., 1997. Swiss Model and Swiss Pdb Viewer: an environment for comparative protein modelling. Electrophoresis 18, 2714-2723.

Hajduch M., Hearne L.B., Miernyk J.A., Casteel J.E., Joshi T., Agrawal G.K., Song Z., Zhou M., Xu D., Thelen J.J., 2010. Systems analysis of seed filling in Arabidopsis: using general linear modeling to assess concordance of transcript and protein expression. Plant Physiol. 152, 2078-2087.

Hill, L.M., Morley-Smith, E.R., Rawsthorne, S., 2003. Metabolism of sugars in the endosperm of developing seeds of oilseed rape. Plant Physiol. 131, 228-236.

Houston, N.L., Hajduch, M., Thelen, J.J., 2009. Quantitative Proteomics of seed filling in castor: comparision with soybean and rapeseed reveals differences between photosynthetic and non-photosynthetic seed metabolism. Plant Physiol. 151, 857-868.

Jeffery, C.J., 1999. Moonlighting proteins. Trends Biochem. Sci. 24, 8-11.

Jindal, H.K., Vishwanatha, J.K., 1990. Functional identity of a primer recognition protein as phosphoglycerate kinase. J. Biol. Chem. 265, 6540-6543.

Journet, E.P., Bligny, R., Douce, R., 1986. Biochemical changes during sucrose deprivation in higher plant cells. J. Biol. Chem. 261, 3193-3199.

Junker, B.H., Lonien, J., Heady, L.E., Rogers, A., Schwender, J., 2007. Parallel determination of enzyme activities an *in vitro* fluxes in *Brassica napus* embryos grown on organic or inorganic nitrogen source. Phytochemistry 68, 2232-2242.

Köpke-Secundo, E., Molnar, I., Schnarrenberger, C., 1990. Isolation and characterization of the cytosolic and chloroplastic 3-phosphoglycerate kinase from spinach leaves. Plant Physiol. 93, 40-47.

Kovári, Z., Flachner, B., Náray-Szabó, G., Vas, M., 2002. Crystallographic and thiol-reactivity studies on the complex of pig muscle phosphoglycerate kinase with ATP and analogues:

correlation between nucleotide binding mode and helix flexibility. Biochemistry 41, 8796-8806.

Kumble, K.D, Vishwanatha, J.K., 1991. Immunoelectron microscopic analysis of the intracellular distribution of primer recognition proteins, annexin 2 and phosphoglycerate kinase, in normal and transformed cells. J. Cell Sci. 99, 751–758.

Larsson-Raznikiewicz, M., Arvidsson, L., 1971. Inhibition of phosphoglycerate kinase by products and product homologues. Eur. J. Biochem. 22, 506-512.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25, 402-408.

Longstaff, M., Raines, C.A., McMorrow, E.M., Bradbeer, J.W., Dyer, T.A., 1989. Wheat phosphoglycerate kinase: evidence for recombination between the genes for the chloroplastic and cytosolic enzymes. Nucleic Acids Res. 17, 6569-6580.

McMorrow, E.M., Bradbeer, J.W., 1990. Separation, purification, and comparative properties of chloroplast and cytoplasmic phosphoglycerate kinase from barley leaves. Plant Physiol. 93, 374-383.

Molnar, M., Vas, M., 1993. Mg^{2+} affects the binding of ADP but not ATP to 3-phosphoglycerate kinase - correlation between equilibrium dialysis binding and enzyme kinetic data. Biochem. J. 293, 595-599.

Ogino, T., Iwama, M., Kinouchi, J., Shibagaki, Y., Tsukamoto, T., Mizumoto, K., 1999. Involvement of a cellular glycolytic enzyme, phosphoglycerate kinase, in Sendai virus transcription. J. Biol. Chem. 274, 35999-36008.

Pal, B., Pybus, B., Muccio, D.D., Chattopadhyay, D., 2004. Biochemical characterization and crystallization of recombinant 3-phosphoglycerate kinase of *Plasmodium falciparum*. Biochim. Biophys. Acta 1699, 277-280.

Plaxton, W.C., 1996. The organization and regulation of plant glycolysis. Annu. Rev. Plant Physiol. Plant Mol. Biol. 47, 185-214.

Plaxton, W.C., Podestá, F.E., 2006. The functional organization and control of plant respiration. Crit. Rev. Plant Sci. 25, 159-198.

Popanda, O., Fox, G., Thielmann, H.W., 1998. Modulation of DNA polymerases alpha, delta and epsilon by lactate dehydrogenase and 3-phosphoglycerate kinase. Biochim. Biophys. Acta 1397, 102–117.

Rivoal, J., Dunford, R., Plaxton, W.C., Turpin, D.H., 1996. Purification and properties of four phosphoenolpyruvate carboxylase isoforms from the green alga *Selenastrum minutum*: evidence that association of the 102-kDa catalytic subunit with unrelated polypeptides may modify the physical and kinetic properties of the enzyme. Arch. Biochem. Biophys. 332: 47-57.

Roberts, J.K., Callis, J., Wemmer, D., Walbot, V., Jardetzky, O., 1984. Mechanisms of cytoplasmic pH regulation in hypoxic maize root tips and its role in survival under hypoxia. Proc. Natl. Acad. Sci. USA. 81, 3379-3383.

Rolletschek, H., Borisjuk, L., Koschorreck, M., Wobus, U., Weber, H., 2002. Legume embryos develop in a hypoxic environment. J. Exp. Bot. 53, 1099-1107.

Saitou, N., Nei, M., 1987. The neighbour-joining method -a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406-425.

Schagger, H., Vonjagow, G., 1987. Tricine sodium dodecyl-sulfate polyacrylamide-gel electrophoresis for the separation of proteins in the range from 1 kDa to 100 kDa. Anal. Biochem. 166, 368-379.

Schwede, T., Kopp, J., Guex, N., Peitsch, M.C., 2003. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. 31, 3381-3385.

Schwender, J., Ohlrogge, J.B., 2002. Probing in vivo metabolism by stable isotope labeling of storage lipids and proteins in developing *Brassica napus* embryos. Plant Physiol. 130, 347-361.

Shah, N., Bradbeer, J.W., 1994. The occurrence of chloroplastic and cytosolic isoenzymes of phosphoglycerate kinase in a range of plant species. Planta 193, 232-237.

Shelp, B.J., Walton, C.S., Snedden, W.A., Tuin, L.G., Oresnik, I.J., and Layzell, D.B., 1995. Gaba shunt in developing soybean seeds is associated with hypoxia. Physiol. Plant. 94, 219-228.

Strucksberg, K.H., Rosenkranz T., Fitter J., 2007. Reversible and irreversible unfolding of multi-domain proteins. Biochim. Biophys. Acta 1774, 1591-1603.

Szilagyi, A.N., Vas, M., 1998. Anion activation of 3-phosphoglycerate kinase requires domain closure. Biochemistry 37, 8551-8563.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596-1599.

Tan, H.L., Yang, X.H., Zhang, F.X., Zheng, X., Qu, C.M., Mu, J.Y., Fu, F.Y., Li, J.A., Guan, R.Z., Zhang, H.S., Wang, G.D., Zuo, J.R., 2011. Enhanced seed oil production in canola by conditional expression of *Brassica napus* LEAFY COTYLEDON1 and LEC1-LIKE in developing seeds. Plant Physiol. 156, 1577-1588.

Thomas, T.M., Scopes, R.K., 1998. The effects of temperature on the kinetics and stability of mesophilic and thermophilic 3-phosphoglycerate kinases. Biochem. J. 330, 1087-1095.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876-4882.

Troncoso-Ponce, M.A., Garcés, R., Martínez-Force, E., 2010a. Glycolytic enzymatic activities in developing seeds involved in the differences between standard and low oil content sunflowers (*Helianthus annuus* L.). Plant Physiol. Biochem. 48, 961-965.

Troncoso-Ponce, M.A., Rivoal, J., Cejudo, F.J., Dorion, S., Garcés, R., Martínez-Force, E., 2010b. Cloning, biochemical characterisation, tissue localization and possible post-translational regulatory mechanism of the cytosolic phosphoglucose isomerase from developing sunflower seeds. Planta 232, 845-859.

Troncoso-Ponce, M.A., Kilaru, A., Cao, X., Durrett, T.P., Fan, J., Jensen, J.K., Thrower, N.A., Pauly, M., Wilkerson, C., Ohlrogge, J.B., 2011a. Comparative deep transcriptional profiling of four developing oilseeds. Plant J. 68, 1014-1027.

Troncoso-Ponce, M.A., Rivoal, J., Dorion, S., Moisan, M.-C., Garcés, R., Martínez-Force, E., 2011b. Cloning, biochemical characterization and expression of a sunflower (*Helianthus annuus* L.) hexokinase associated with seed storage compounds accumulation. J. Plant Physiol. 168, 299-308.

Troncoso-Ponce, M.A., Kruger, N.J., Ratcliffe, R.G., Garcés, R., Martínez-Force, E., 2009. Characterization of glycolytic initial metabolites and enzyme activities in developing sunflower (*Helianthus annuus* L.) seeds. Phytochemistry 70, 1117-1122.

Vishwanatha, J.K., Jindal, H.K., Davis, R.G., 1992. The role of primer recognition proteins in DNA replication: association with nuclear matrix in HeLa cells. J. Cell Sci. 101, 25-34.

Watson H.C., Walker N.P.C., Shaw P.J., Bryant T.N., Wendell P.L., Fothergill L.A., Perkins R.E., Conroy S.C., Dobson M.J., Tuite M.F., Kingsman A.J., Kingsman S.M., 1982. Sequence and structure of yeast phosphoglycerate kinase. EMBO J. 1, 1635-1640.

Zerrad, L., Merli, A., Schroder, G.F., Varga, A., Graczer, E., Pernot, P., Round, A., Vas, M., Bowler, M.W., 2011. A spring-loaded release mechanism regulates domain movement and catalysis in phosphoglycerate kinase. J. Biol. Chem. 286, 14040-14048.

Zomer, A.W.M., Allert, S., Chevalier, N., Callens, M., Opperdoes, F.R., Michels, P.A.M., 1998. Purification and characterisation of the phosphoglycerate kinase isoenzymes of *Trypanosoma brucei* expressed in *Escherichia coli*. Biochim. Biophys. Acta 1386, 179-188.

Table 1

Sequence of PCR primers used in this work.

Primer name	Sequence[a]
dPGK-F	5′ -GCTKVRMTYCCWGADGGBGG-3′
dPGK-R	5′ -GGYTGGATGGGATTWGA-3′
FA2Z	5′ -AACTGGAAGAATTCGCGG-3′
HacPGK1-F	5′ -GCATCCGTGCTGCTGTTC-3′
HacPGK1-R1	5′ -GGTCGACATCCTTGTTCT-3′
HacPGK1-R2	5′ -CCGACCTGTACGTCAACGAT-3′
HacPGK2-F	5′ -GCATACGCGCCGCCGTCC-3′
HacPGK2-R1	5′ -GTCAACATCCTTGTGCTC-3′
HacPGK2-R2	5′ -CAGACTTATACGTCAATGAT-3′
HapPGK-F	5′ -ATGGTGCTAAAGTTATTC-3′
HapPGK-R1	5′ -GTGTGACATTTTGCTTTT-3′
HapPGK-R2	5′ -CAGATCTTTACGTCAATGAT-3′
SphHacPGK1-F	5′ -GACT**GCATGC**ATGGCGACTCTTAATCTCAGACACTGG-3′
PstHacPGK1-R	5′ -CTGA**CTGCAG**TTATATCTTAGGCAATAGAGTTGTGG-3′
SphHacPGK2-F	5′ -CCC**GCATGC**ATGGCGACAAAGAAGAGTGTAAG-3′
SalHacPGK2-R	5′ -TTT**GTCGAC**TCATGCCTCATTCAAGGCA-3′
SphHapPGK-F	5′ -CCC**GCATGC**GCGAAGAAGAGTGTTGGTGAC-3′
PstHapPGK-R	5′ -TTT**CTGCAG**TTACGCAGAGACAGTTGCGA-3′
Q*HacPGK1*-F	5′ -CGGCGGTTGAGAAAGTTGGAC-3′
Q*HacPGK1*-R	5′ -ACCCTTAACCAACAACTTCAG-3′ (3′-UTR)
Q*HacPGK2*-F	5′ -CAGCTGTGGAAAAGGTTGGGT-3′
Q*HacPGK2*-R	5′ -TTCTAGTCTCAACTAAAACCG-3′ (3′-UTR)

Q*HapPGK*-F	5′ -AAGGTTGGAGTGGCGGATGTG-3′	
Q*HapPGK*-R	5′ -ATCACGATTTAGAACTACGTC-3′	(3′-UTR)
Q*HaActin*-F4	5′ -GCTAACAGGGAAAAGATGACT-3′	
Q*HaActin*-R4	5′ -ACTGGCATAAAGAGAAAGCACG-3′	

[a] Restriction sites are indicated in bold.

Table 2. Apparent kinetic parameters for (6His)*Ha*cPGK1, (6His)*Ha*cPGK2 and (6His)*Ha*pPGK recombinant protein at different pH values. V_{max}, maximum velocity (nkat mg^{-1} of protein); K_m, Michaelis-Menten (μM). Data are mean values and standard deviations from three independent experiments.

Substrate	Kinetic parameters	pH 7.0	pH 7.6	pH 8.2
(6His)*Ha*cPGK1				
3-PGA	*V*max	243.7 ±3.8	244.5 ± 5.2	254.4 ± 3.3
	*K*m	1948.6±96.4	1787.9±121.1	1635.1±176.7
	*V*max/*K*m	0.12	0.13	0.15
ATP	*V*max	177.0 ± 1.5	179.2 ± 5.1	168.8 ± 4.6
	*K*m	243.3 ± 5.7	227.6 ± 17.4	215.7 ± 15.9
	*V*max/*K*m	0.72	0.78	0.78
(6His)*Ha*cPGK2				
3-PGA	*V*max	5583.4 ± 58.1	6056.4 ± 61.5	5829.0 ± 123.7
	*K*m	601.7 ± 27.9	711.2 ± 30.6	683.3 ± 62.1
	*V*max/*K*m	9.27	8.51	8.53
ATP	*V*max	5203.5 ± 135.2	5150.2 ± 156.5	5470.2 ± 210.0
	*K*m	357.1 ± 21.8	367.5 ± 26.0	404.7 ± 35.2
	*V*max/*K*m	14.6	14.0	13.5
(6His)*Ha*pPGK				
3-PGA	*V*max	5199.7 ± 58.2	5621.3 ± 88.7	5131.2 ± 56.2
	*K*m	707.6 ± 33.5	719.6 ± 47.8	589.6 ± 29.1
	*V*max/*K*m	7.35	7.81	8.70
ATP	*V*max	4665.6 ± 56.3	5682.8 ± 132.2	4485.4 ± 132.3
	*K*m	300.1 ± 9.0	406.7± 21.4	356.9 ± 24.8
	*V*max/*K*m	15.5	13.9	12.5

Table 3. Effect of various metabolites on (6His)*Ha*cPGK1, (6His)*Ha*cPGK2 and (6His)*Ha*pPGK recombinant protein activities. Values in bold indicate that PGK activity was significantly different from that of the control (no effector).

Effectors	Concentration (mM)	PGK activity (%)		
		(6His)*Ha*cPGK1	(6His)*Ha*cPGK2	(6His)*Ha*pPGK
None		100.0 ± 2.8	100.0 ± 4.3	100.0 ± 2.3
Glycerol-3P	10	92.0 ± 11.1	95.9 ± 1.2	97.2 ± 0.4
Glycerol-2P	10	97.3 ± 3.2	100.5 ± 1.2	101.0 ± 2.6
UDP	10	94.3 ± 4.4	97.9 ± 1.0	104.5 ± 1.7
ADP	0.1	**63.1 ± 2.0**	**86.7 ± 2.9**	**85.6 ± 1.7**
	0.25	**62.4 ± 2.7**	**74.2 ± 2.3**	**69.7 ± 4.4**
	0.5	**62.4 ± 2.7**	**50.8 ± 3.4**	**48.1 ± 3.6**
	1	**32.3 ± 4.7**	**31.4 ± 2.2**	**32.5 ± 2.6**
AMP	0.625	104.3 ± 6.4	95.2 ± 1.5	98.9 ± 1.5
	1.25	99.8 ± 4.8	**88.4 ± 1.1**	**94.4 ± 1.7**
	2.5	96.5 ± 3.4	**77.1 ± 1.8**	**86.1 ± 1.4**
	5	**78.1 ± 6.8**	**62.6 ± 2.3**	**72.4 ± 1.8**
	10	**40.5 ± 5.8**	**44.4 ± 2.6**	**50.9 ± 3.4**

Figure legends

Fig. 1. Alignment of the deduced amino acid sequences of sunflower plastidial, *Ha*pPGK (accession number ADV16381), and cytoplasmic, *Ha*cPGK1 (ABI18157) and *Ha*cPGK2 (ADV16380), PGKs with the closely related sequences from *Arabidopsis thaliana* (*At*pPGK1 NP176015 , *At*pPGK2 NP187884 and *At*cPGK NP178073) and the more phylogenetically distant sequence from *Geobacillus stearothermophilus* (*Gs*PGK, P18912). Asterisks mark identical residues, colon marks conservative changes, and dot marks weakly conservative changes between the sequences. Residues involved in the substrate binding site are indicated with white triangles, those involved in hinge regions with grey triangles and the ones forming the ADP binding site with black triangles.

Fig. 2. Phylogenetic comparison of PGK proteins from plants, green algae and bacteria rooted in the bacteria *Geobacillus stearothermophilus* PGK protein sequence, used as template for *Ha*PGKs three-dimensional structure modelling. GenBank accession numbers follow the species names.

Fig. 3. Proposed structural models for *Ha*cPGK1 and *Ha*pPGK phosphoglycerate kinases. (A and C) Ribbon diagrams for *Ha*cPGK1 and *Ha*pPGK respectively. (B and D) Views of the molecular surfaces for *Ha*cPGK1 and *Ha*pPGK respectively, showing the residues involved in the substrate binding site in green, those forming the ADP binding site in pink and the catalytic site in blue.

Fig. 4. SDS-PAGE showing the (6His)*Ha*cPGK1, (6His)*Ha*cPGK2 and (6His)*Ha*pPGK purification steps. MW, molecular mass standards; lanes 1, 3 and 5 *E. coli* crude extracts after induction of (6His)*Ha*cPGK1, (6His)*Ha*cPGK2 and (6His)*Ha*pPGK respectively; and lanes 2, 4 and 6 purified (6His)*Ha*cPGK1, (6His)*Ha*cPGK2 and (6His)*Ha*pPGK proteins respectively. The running positions of molecular mass standards are indicated on the left.

Fig. 5. Thermal stability of (6His)*Ha*cPGK1, (6His)*Ha*cPGK2 and (6His)*Ha*pPGK activities. Data are mean values and standard deviations from three independent experiments.

Fig. 6. Effect of pH variation on (6His)*Ha*cPGK1, (6His)*Ha*cPGK2 and (6His)*Ha*pPGK activities. Data are mean values and standard deviations from three independent experiments.

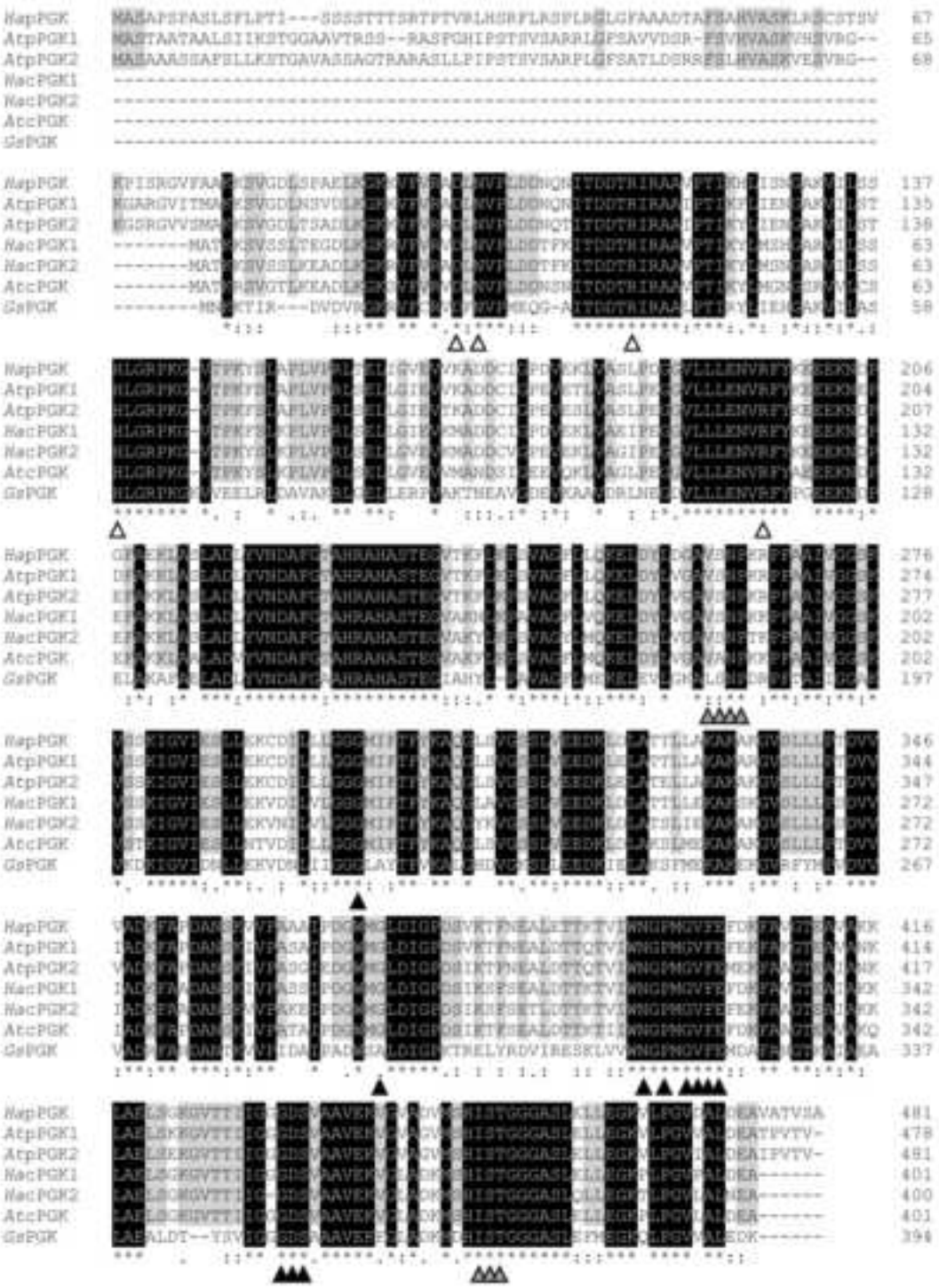
Fig. 7. PGK expression levels from developing seeds and vegetative tissues from *Helianthus annus* (A) and *Arabidopsis thaliana* (B). A, *HacPGK1*, *HacPGK2* and *HapPGK* expression determined by real time PCR form sunflower line CAS-6; B, *AtcPGK (*At1g79550*), AtpPGK1 (*At1g56190) and *AtpPGK2 (*At3g1270) expression estimated from microarrays of Schmid et al. (2005). Sunflower seeds stages counting days after flowering (DAF); Arabidopsis stages: 3, mid globular to early heart embryos; 4, early to late heart embryos; 5, late heart to mid torpedo embryos; 6, mid to late torpedo embryos; 7, late torpedo to early walking-stick embryos; 8, walking-stick to early curled cotyledons embryos; 9, curled cotyledons to early green cotyledons embryos; 10, green cotyledons embryos. Values in panel A represent mean values of three independent samples.

Fig. 8. PGK protein and activity levels in developing sunflower seeds. (A) Recognition of all sunflower PGKs by an immune serum rose against cPGK1. Western blot analysis of 25 ng each of the following purified recombinant proteins: lane 1: (6His)*Ha*cPGK1, lane 2: (6His)*Ha*cPGK2 and lane 3: (6His)*Ha*pPGK. (B) Immunoblot analysis of *Ha*PGKs in developing sunflower seed extracts. Developing seeds were collected at 15, 20, 25 and 30 days after flowering (number of DAF indicated above each lane). Fifteen µg of proteins from each extract were subjected to immunoblot analysis using the immune serum as in panel (A). Purified (6His)*Ha*cPGK1 protein (lane P) was used as a control. (C) Total PGK activity levels measured in developing sunflower kernels from line CAS-6. PGK activity values are the means ± SD from three independent samples.

Fig. 9. PGKs localisation in sunflower seed tissues. Longitudinal section of the proximal end incubated with antibodies raised against purified sunflower 6(His)*Ha*cPGK1 (A) or with pre-immune serum (B). Bar, 500 µm; PR, procambial ring.

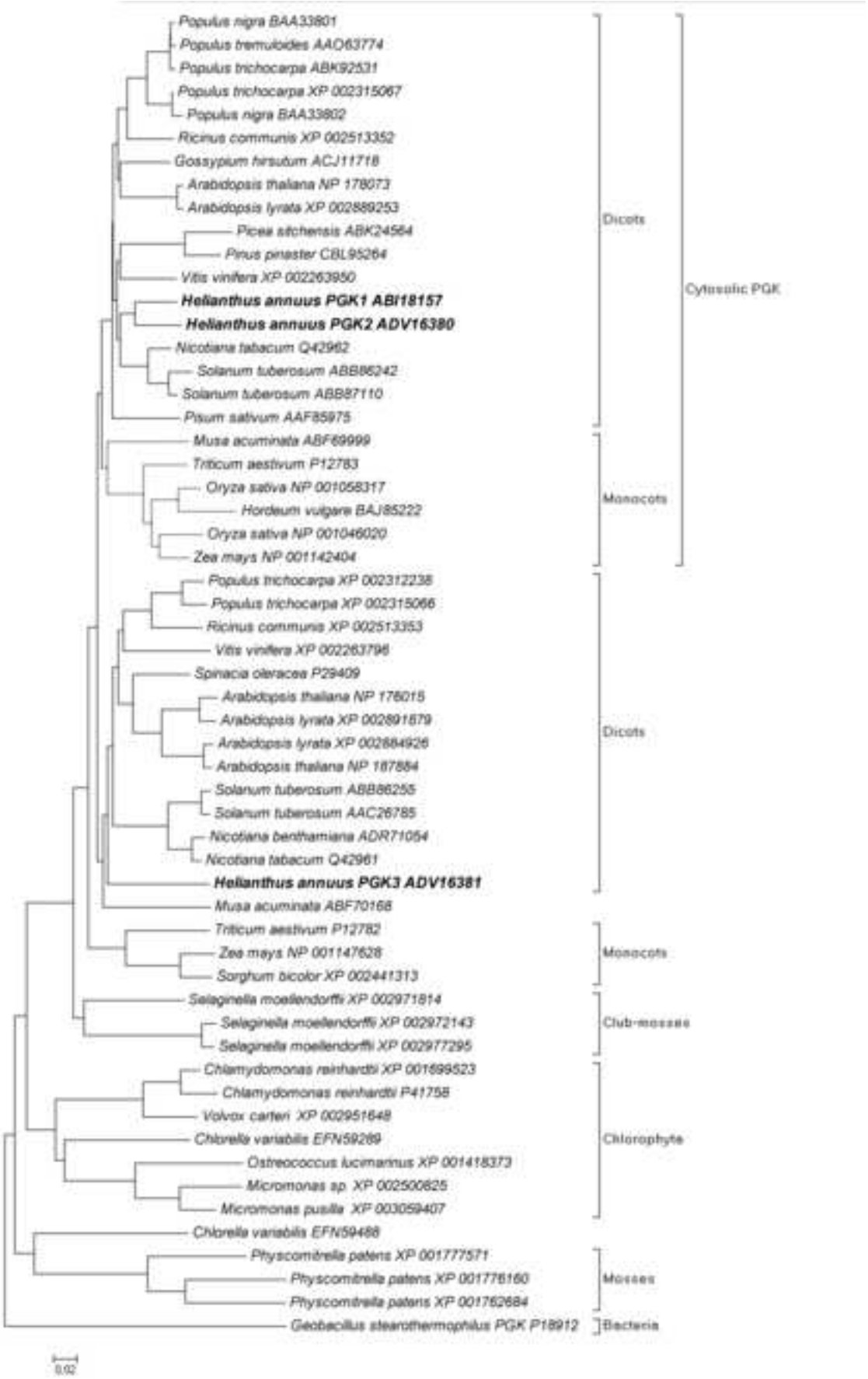
Figure(s)

Click here to download high resolution image



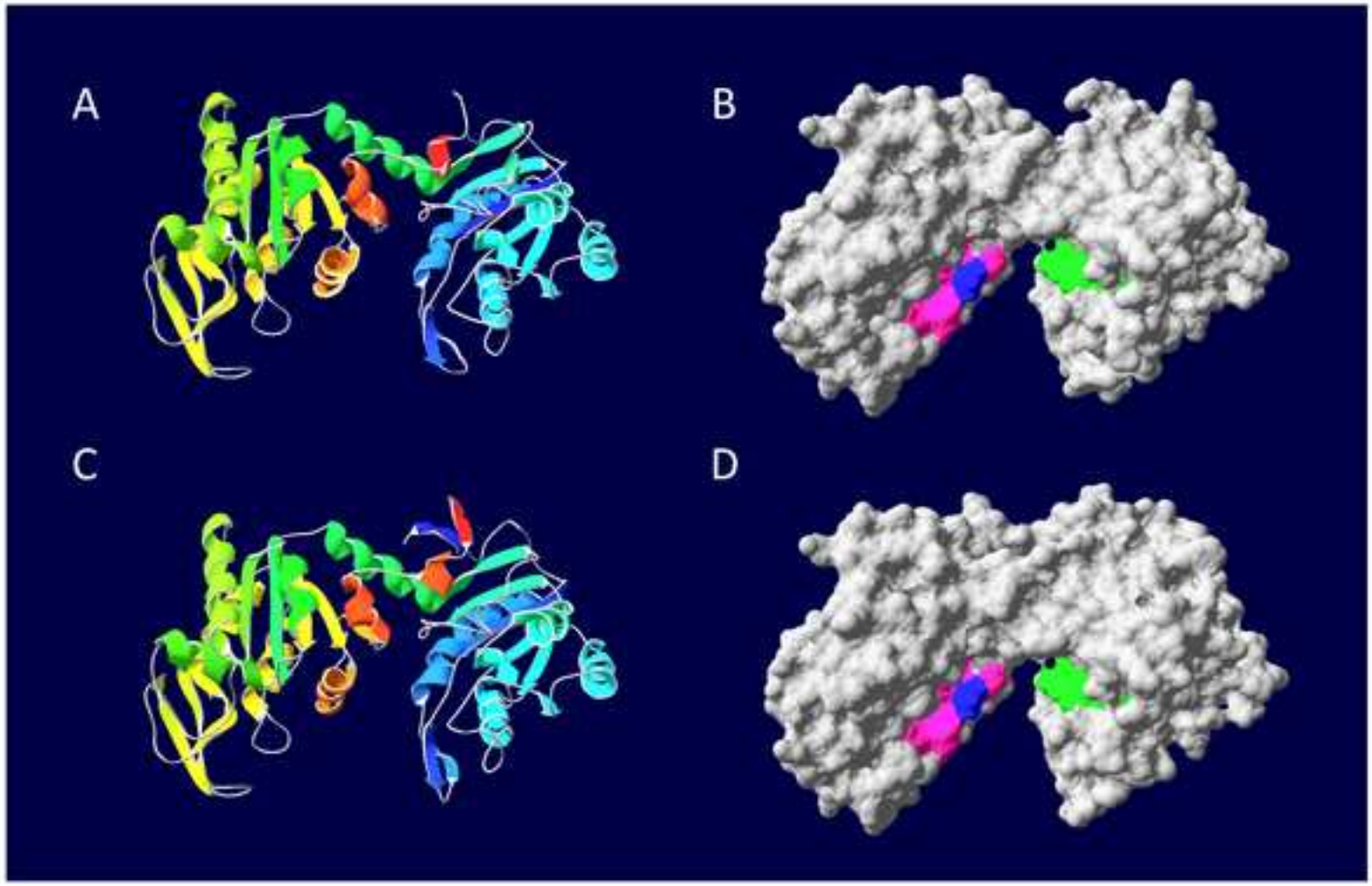
Figure(s)

Click here to download high resolution image



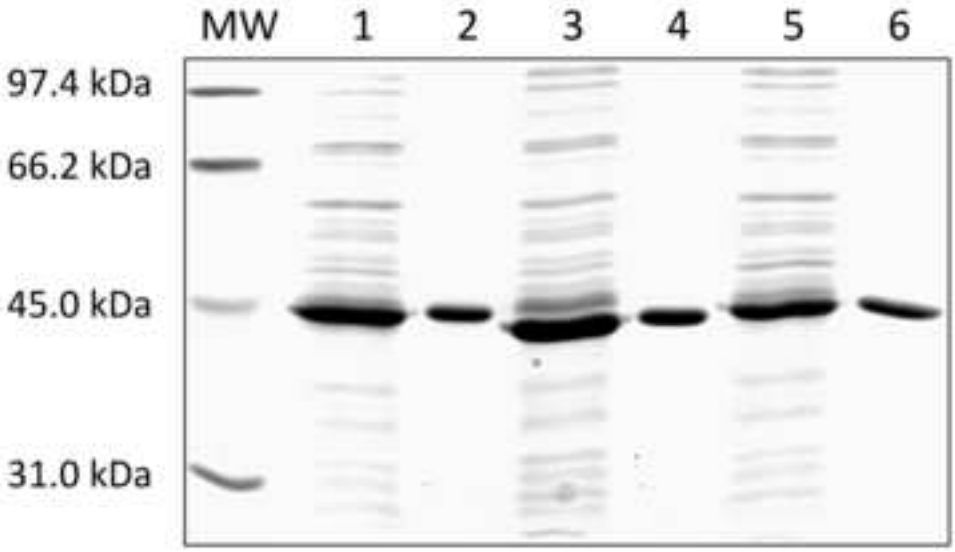
Figure(s)

Click here to download high resolution image



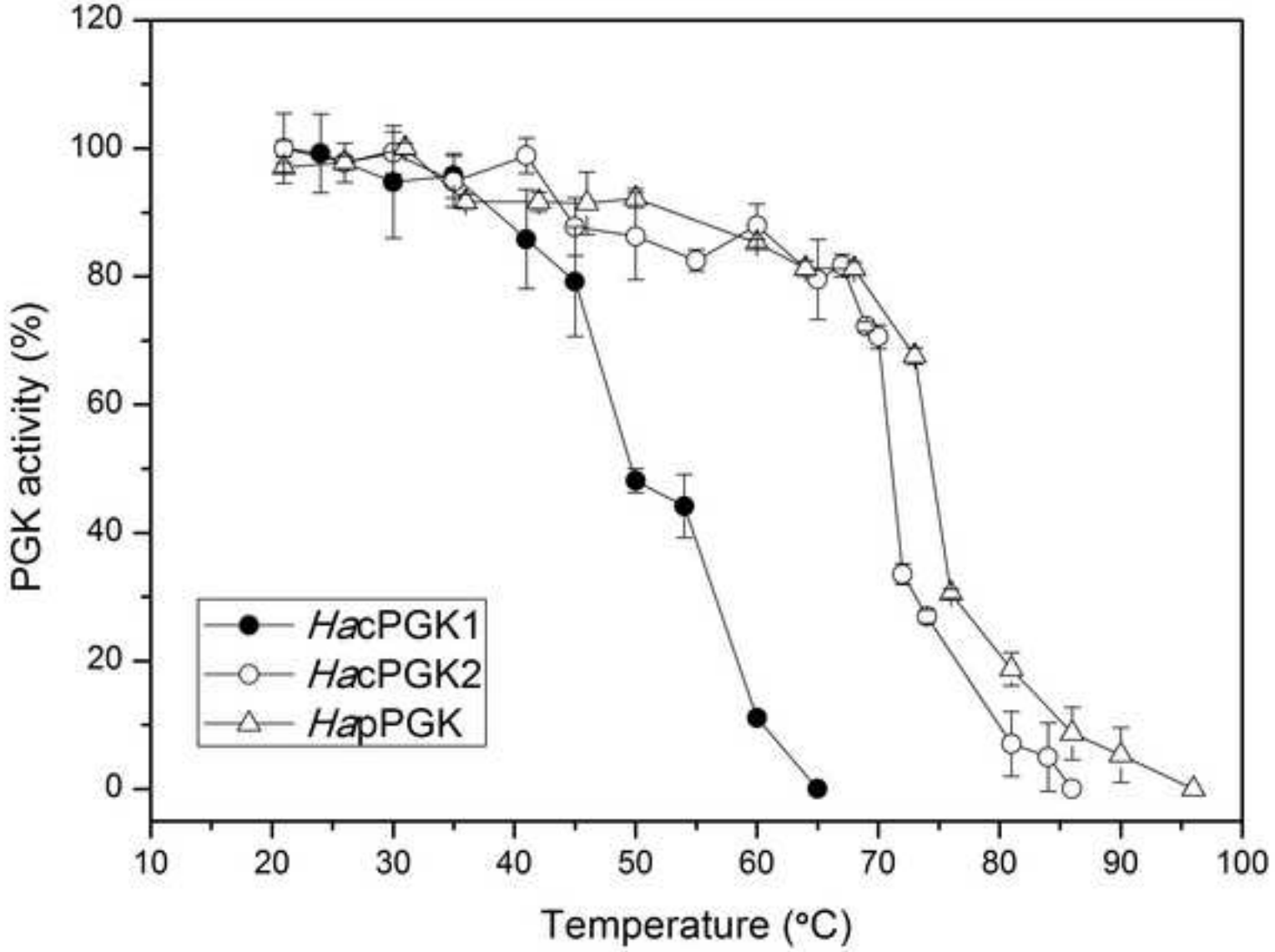
Figure(s)

Click here to download high resolution image



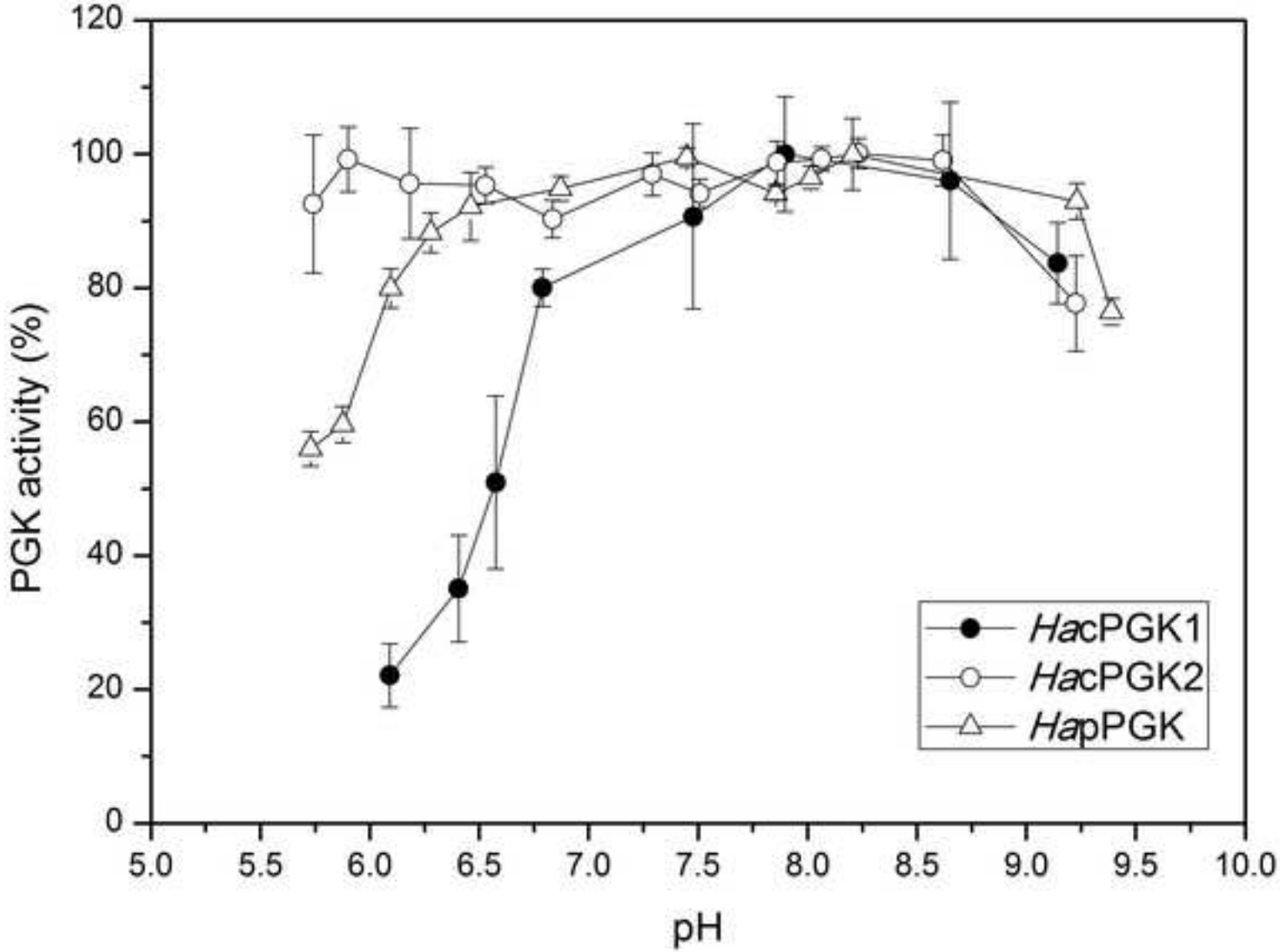
Figure(s)

Click here to download high resolution image



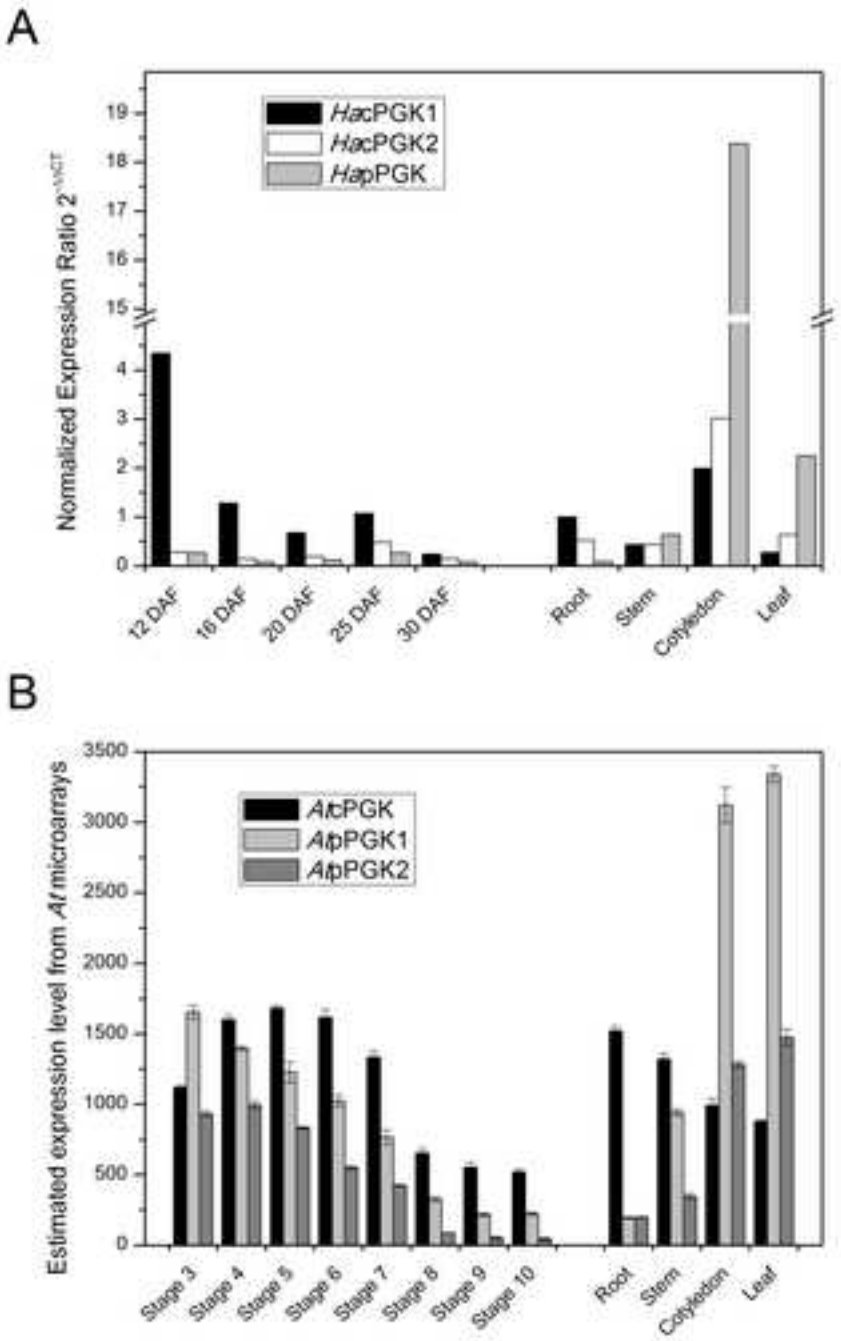
Figure(s)

Click here to download high resolution image



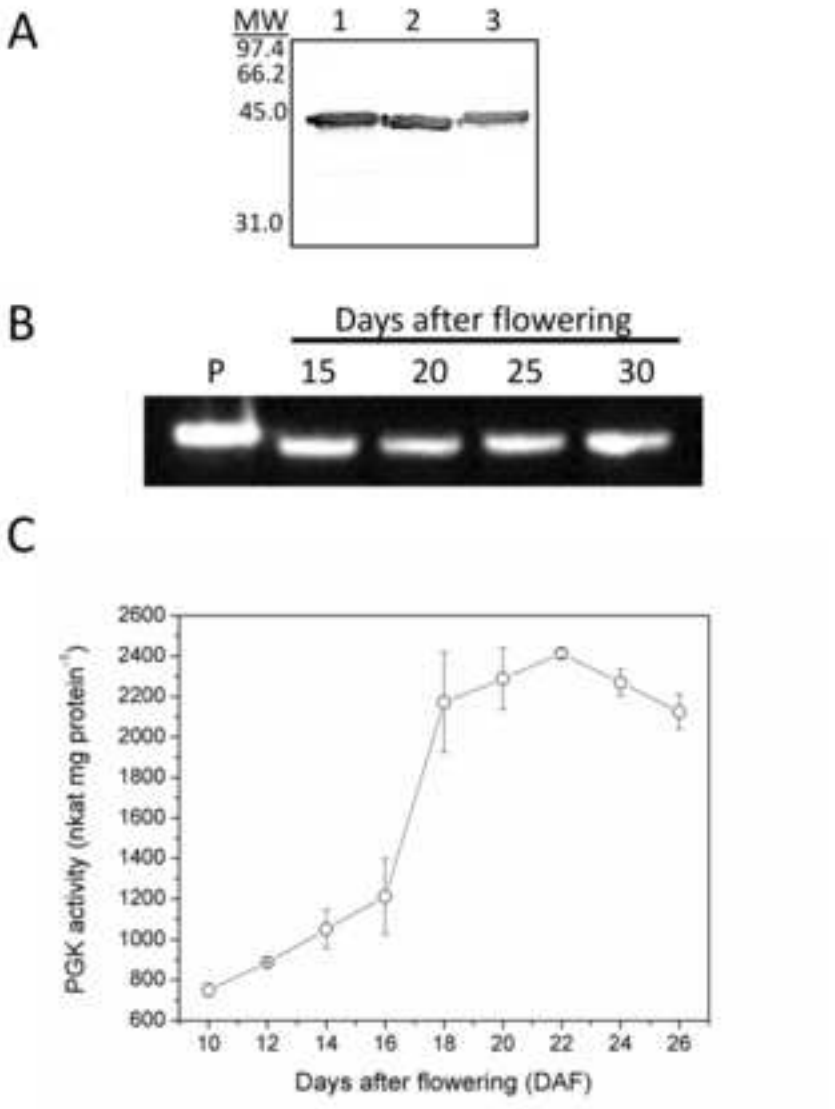
Figure(s)

Click here to download high resolution image



Figure(s)

Click here to download high resolution image



Figure(s)

Click here to download high resolution image