Redesigning Crop Products for Biotechnology: Starch and Fructan

Society/Host Colloquium Organized and Edited by P. John (University of Reading). 638th Meeting held at Reading University, 10–12 April, 1991

Starch biosynthesis and its regulation
Jack Preiss, K. Ball, B. Smith-White, A. Iglesias, G. Kakefuda and L. Li
Department of Biochemistry, Michigan State University, East Lansing, MI 48824, U.S.A.

Introduction

The biosynthetic reactions of starch synthesis

There is strong evidence to indicate that the ADP-glucose (ADPGlc) pathway, consisting of the enzymatic reactions catalysed respectively by ADPGlc synthetase (or ADPGlc pyrophosphorylase, EC 2.7.7.27; reaction 1), starch synthase (EC 2.4.1.21; reaction 2) and branching enzyme (EC 2.4.1.18; reaction 3), is the predominant if not the sole pathway towards starch synthesis in leaf as well as in non-photosynthetic tissue such as endosperm or tubers.

\[
\begin{align*}
\alpha\text{-Glucosyl-1-P} + \text{ATP} & \leftrightarrow \text{ADPGlc} + \text{PP}_i \\
\text{ADPGlc} + (\text{glucosyl})_n & \rightarrow \text{ADP} + (\text{glucosyl})_{n+1}
\end{align*}
\]

(1) (2)

Linear glucosyl chain of α-glucan → branched chain glucan with α1–6-linkage branch point (3)

The data supporting this view are from studies of isolated mutants of maize endosperm [1, 2] having deficient levels of ADPGlc synthetase, which are also deficient in starch levels. Recently, it has been shown that the \(rb\) locus controls the level of ADPGlc synthetase levels in developing pea embryos [3]. One pea line having recessive \(rb\) genes contained 38–72% of the starch found in the pea line having the \(Rb\) loci and about 3–5% of the ADPGlc synthetase activity. Other convincing evidence that the ADPGlc pathway is the dominant pathway, if not the sole pathway, in non-photosynthetic or reserve tissue has been obtained from Dr L. Willmitzer’s group in Berlin where they have shown that application of anti-sense RNA to either the gene of the small subunit or the large subunit of the potato tuber ADPGlc synthetase resulted in the production of minitubers containing only 3% of the starch observed in the normal or untreated minituber (L. Willmitzer, private communication). These data in four diverse species of plants strongly indicate that the ADPGlc pathway must play a major role in starch synthesis. Other data demonstrating a direct relationship between increase in activity of the above starch biosynthetic enzymes and increased starch accumulation in various plants have been previously reviewed [5–7].

Regulation of the ADPGlc pathway

There is evidence \emph{in vitro} suggesting that ADPGlc synthesis is regulated by activation of the plant ADPGlc synthetase by 3-phosphoglycerate (3PGA) and inhibition by inorganic phosphate (\(P_i\)). \emph{In vivo} or \emph{in situ} evidence showing a correlation between the concentrations of 3PGA and starch and inverse correlations between \(P_i\) and starch levels have been obtained and were reviewed [5, 6]. Recently, Pettersson & Ryde-Pettersson [8] have applied modern control theory as developed by Kacser & Burns [9, 10] to develop a kinetic model to determine the extent that stromal metabolites, known to affect leaf ADPGlc synthetase activity \emph{in vitro}, controlled the rate of photosynthetic starch production under conditions of light and \(CO_2\) saturation. The model consists of the 13 enzyme-catalysed steps of

Abbreviations used: ADPGlc, ADP-glucose; 3PGA, 3-phosphoglycerate; Fru2,6-P₂, fructose-2,6-bisphosphate; PGO, phenylglyoxal, 8-N₃-ATP, 8-azido-ATP.
the reductive pentose phosphate pathway and starch synthesis and photosynthetic export (P3 triose-P translocator) from the chloroplast. Using this model, the steady state concentrations of various stromal metabolites, the corresponding rates of CO2 fixation and starch production were defined as a function of the P3 concentration external to the chloroplast. The model essentially agrees with reported experimental data with respect to metabolite concentrations and to CO2 fixation and starch synthetic rates. It showed that ATP and various stromal metabolites, the corresponding contributions to starch synthetic rate changes induced by increasing P3 concentrations external to the chloroplast. At low P3 concentrations, the most significant contribution to increased starch synthetic rate was ATP and to a smaller extent 3PGA. At P3 concentrations higher than 0.12 mm, 3PGA becomes the predominant regulator of starch synthesis with glucose-1-P and fructose-6-P contributing to regulation to a smaller extent. Thus, using modern mathematical control analysis [9, 10], they reached the conclusion that 3PGA and P3 play an important role in regulating starch synthesis with contributions made by ATP, glucose-1-P and fructose-6-P. Since these metabolites are either substrates or effectors of the ADPGlc synthetase, the analysis is consistent with the view that the 3PGA/P3 ratio regulates starch synthesis by regulation of ADPGlc synthetase.

The availability of chloroplast mutants of phosphoglucoisomerase of *Clarkia xantiana* [11, 12], phosphoglucomutase [13], and ADPGlc synthetase of *Arabidopsis thaliana* [14, 15] have allowed the analysis of the extent of control that these enzymes exert on chloroplast starch synthesis. The plastid P3-glucoisomerase exerted very little control over starch or sucrose synthesis in low light but did exert some control of starch synthesis in saturating light. Lowering the cytosolic enzyme activity had little effect on either starch or sucrose synthesis in saturating light but did increase starch synthetic rate in low light and decreased sucrose synthetic rate. Partitioning of carbon between sucrose and starch was thus affected by variation of the cytosolic phosphoglucoisomerase levels. Metabolite levels were also affected in these mutants. In the mutant containing only 18% of the wild-type cytosolic phosphoglucoisomerase activity both fructose 2,6-bisphosphate (Fru-2,6-P2) and 3PGA levels increase about 100%. Neuhaus et al. [12] indicated that the reduced rate of sucrose synthesis was owing to the increased Fru-2,6-P2 level which caused increased inhibition of cytosolic fructose 1,6-bisphosphate, which is on the pathway towards sucrose synthesis. The data also strongly support the view that increased starch synthesis in the mutants with reduced levels of phosphoglucoisomerase is due to activation of the ADPGlc synthetase by the increased 3PGA concentration and 3PGA/P3 ratio.

These experiments have been extended to the null chloroplast phosphoglucomutase [13] and the low activity (7% of wild type) ADPGlc synthetase mutants [14] of *Arabidopsis thaliana* [15]. In low light, a 50% decrease in phosphoglucomutase activity had no significant effect on the above fluxes. However, a 50% and 93% decrease of ADPGlc synthetase activity resulted in a 23% and 74% decrease in flux of starch synthesis with a concomitant increase of 17% and 42% respectively in sucrose synthetic rate. Thus diminution of ADPGlc synthesis activity not only significantly affected starch synthesis but also affected the partitioning of photosynthetic carbon causing more to be directed towards sucrose biosynthesis. In high light a 50% decrease in phosphoglucomutase activity resulted in a 20% decrease in starch synthesis rate with little effect on sucrose synthesis rate. However, reduction of the ADPGlc synthesizing activity to 50% and 93% resulted in a 39% and 90% decrease respectively, in starch synthesis flux. The flux control coefficients [10] for the enzymes for starch synthesis were calculated to determine the distribution of control, and compared with previous results obtained with the *Clarkia xantiana* phosphoglucoisomerase. Of the enzymes under study the flux to starch synthesis is only regulated by ADPGlc synthetase in low light. In high light and CO2, ADPGlc synthetase activity exerts major control and other enzymes such as plastid phosphoglucomutase and phosphoglucoisomerase exert a small (but significant) control.

In summary, various analyses of the starch biosynthetic system of a number of plants utilizing data obtained *in vivo* from different plants and applying the Kacser & Burns control analysis method [9, 10] point out that the major site of regulation of starch synthesis is at ADPGlc synthetase and that 3PGA and P3 are important regulatory metabolites of that enzyme.

**Properties of ADPGlc synthetase**

*Summary of regulatory and structural properties*

*Regulation by 3-P-glycerate and by P3.* The properties
of the plant and bacterial ADPGlc synthetases have been extensively reviewed and for detailed information on the plant systems the reader is referred to [5–7, 18]. For the bacterial systems the reader is referred to [19, 20]. As indicated earlier, and in the above reviews on plant systems, every plant ADPGlc synthetase studied has as its major activator, 3PGA and its inhibitor, P_i. For every leaf system studied, whether the leaf source is from a plant utilizing C_3 or C_4 pathways or Crassulacean metabolism, the major activator is still 3PGA and the inhibitor is P_i. In non-photosynthetic systems studied (e.g. maize endosperm, potato tuber) the ADPGlc synthetase activity is highly dependent on the presence of 3PGA and is inhibited by P_i. Activation of ADPGlc synthetase of carrot roots, avocado mesocarp [21], of cassava tubers [22] and of lilly pollen [23] by 3PGA have been reported.

**Structural properties of ADPGlc synthetase**

The most studied ADPGlc synthetase with respect to structural properties is the spinach leaf enzyme [24–27]. This enzyme has a molecular mass of 206 kDa and is composed of two different subunits with molecular masses of 51 and 54 kDa. These subunits are distinguished not only by differences in their molecular mass but also by different amino acid compositions, N-terminal sequences, peptide patterns on h.p.l.c. of tryptic digests and antigenic properties. Thus, the two subunits are distinct and probably are the products of two genes. In contrast, the bacterial ADPGlc synthetase is composed of only one subunit, 50–55 kDa in molecular mass, depending upon the source, and the native enzymes are homotetrameric [19].

Recently, other plant enzymes have been shown to be composed of two dissimilar subunits. The maize endosperm ADPGlc synthetase, which has a molecular mass of 230 kDa, is composed of two subunits of 55 and 60 kDa. Of interest, was the analysis of the maize endosperm mutants, *shrunken* 2 (*sh2*) and *brittle* 2 (*bt2*). In Western blotting experiments and using the native and subunit antisera towards native spinach leaf ADPGlc synthetase the presence of two immunoreactive bands in gel electrophoresis of extracts of wheat, rice and maize leaves was observed [33]. Similar data were obtained with *Arabidopsis thaliana* ADPGlc synthetase [4, 14]. It is composed of two subunits, molecular masses of 48 and 54 kDa. Indeed one of the *Arabidopsis thaliana* ADPGlc synthetase mutants, TL46, lacks the larger, 54 kDa subunit [14]. The mutant enzyme has been partially purified and its kinetic properties have been studied. It was activated by 3PGA and inhibited by P_i, as was the normal enzyme. However, there was about a 29-fold lower affinity for 3PGA and a 4–5-fold higher affinity for P_i, for the larger-subunit-deficient mutant enzyme. Moreover, in the presence of saturating concentrations of 3PGA, the mutant enzyme had *K_m* values for ATP, glucose-1-P and Mg^{2+}, 6-, 5-, and 2-fold higher, respectively, than those of the normal enzyme. Thus, changes in the kinetics of the mutant enzyme may be caused by the absence of the large subunit. This mutation provides further evidence that the larger subunit is a necessary component of the native ADPGlc synthetase for optimal activity as the mutant enzyme has only 7% of the wild-type activity. The mutant synthesizes starch at 9% of the rate seen for the normal strain in high light and only at 26% of the rate at low light [15].

The potato tuber ADPGlc synthetase has been highly purified, and by two-dimensional PAGE two different proteins could be distinguished by their slight differences in molecular mass, 50 and 51 kDa, and in net charge [30, 31]. Thus, as found with the spinach leaf and maize endosperm enzyme, the potato tuber enzyme is composed of two distinct subunits and not one as previously thought [32]. It appears that other plant ADPGlc synthetases may also be composed of two subunits on the basis of Western blotting experiments. Using antisera towards native spinach leaf ADPGlc synthetase the presence of two immunoreactive bands in gel electrophoresis of extracts of wheat, rice and maize leaves was observed [33].
The allosteric 3PGA activator binding site of ADPGlc synthetase: determination of the amino acid sequence at the 3PGA allosteric binding site

It is of interest to know why the two subunits are required for optimal catalytic activity. The bacterial ADPGlc synthetases from at least seven organisms contain only one subunit and the native structure is homotetrameric [19, 20].

Initial studies were designed to determine the nature and location of the 3PGA binding site. These studies were facilitated by the finding that pyridoxal 5-phosphate is an activator of spinach leaf ADPGlc synthetase [24, 25]. Pyridoxal-P is not as effective as 3PGA with regard to activation of velocity; only a 6-fold stimulation is observed compared with a 25-fold stimulation caused by 3PGA. Three other ADPGlc synthetases, (maize endosperm, Arabidopsis thaliana and from the cyanobacterium, Synechocystis 6803) are also stimulated by pyridoxal-5-P (A. Iglesias & J. Preiss, unpublished work). The apparent affinity, or $A_{50}$ value (concentration of activator giving 50% of maximal stimulation), for the spinach leaf enzyme for pyridoxal-5-P is about 15 $\mu$M and is somewhat lower than the $A_{50}$ of 45 $\mu$M for 3PGA. Thus, there is a higher apparent affinity for pyridoxal-5-P even though there is a lesser stimulation of maximal velocity. Pyridoxal-5-P can be covalently bound to the spinach leaf ADPGlc pyrophosphorylase by reduction with sodium borohydride [24, 25, 27, 34], producing a modified enzyme that is highly active in the absence of activator. The activator, 3PGA, as well as the inhibitor, P, prevents or inhibits the reductive covalent binding of the pyridoxal-5-P and reduces the increase in activity caused by the reductive phosphopyridoxylation [24, 25, 27, 34]. The modified enzyme is also quite insensitive to P inhibition [24, 25, 34]. All the above results are consistent with the view that pyridoxal-5-P binds to the activator site and that the covalent modification with pyridoxal-5-P places the enzyme in a conformational state with higher catalytic activity and where it is resistant to P inhibition.

Incorporation of $[^3H]$pyridoxal-5-P into the spinach leaf enzyme ADPGlc synthetase has been observed and there is a relationship between the amount of pyridoxal-5-P incorporated and the increase in activity caused by the modification [25, 34]. Many metabolites and substrates were tested and only P, and 3PGA inhibited incorporation of the labelled pyridoxal-5-P into the enzyme. Labelled pyridoxal-5-P is incorporated about equally into both the 54 and 51 kDa subunits [34]. The activator, 3PGA, prevented incorporation of the pyridoxal-5-P into both subunits to the same extent [34]. The labelled 51 kDa subunit has been subjected to trypsin digestion. The labelled peptide isolated using reverse-phase h.p.l.c. was sequenced [27, 34]. The sequence obtained is SGIVTVIKDALL with Lys covalently linked to the pyridoxal-P. The amino acid sequence found is different to that corresponding to the fructose 1,6-bisphosphate activator site in the Escherichia coli ADPGlc synthetase [35]. There are many more basic amino acid residues in the E. coli sequence than in the spinach leaf 51 kDa activator sequence. The amino acid sequence of the 54 kDa subunit activator site of the spinach leaf has not been elucidated as yet. At present, up to two or three radioactive peptides have been isolated from tryptic digests and these appear to be different from the peptide isolated from the 51 kDa peptide [34]. It would be of interest to know if the subunit activator sites are independent in causing the enzyme to go into an active conformation upon binding of 3PGA or if both peptides interact to provide the active conformation upon 3PGA binding.

Recent experiments have shown that phenylglyoxal (PGO), an arginine-specific reagent, can also affect the allosteric properties of the spinach leaf, Anabaena and Synechocystis ADPGlc synthetases. Fig. 1 shows that the allosteric effectors 3PGA and P, can protect the spinach leaf enzyme from inactivation by PGO whereas the substrate glucose-1-P gives minimum protection. About 50% protection occurred with 30 $\mu$M-P, and with 300
μM-3PGA. ATP or Mg2+ did not protect the enzyme from PGO inactivation. Similar results were also noted for the ADPGlc synthetase from *Synechocystis*. The apparent affinity for 3PGA and the Vₘₐₓ at saturating 3PGA concentrations were decreased by PGO inactivation both with the spinach leaf and *Synechocystis* enzymes. Moreover, PGO treatment also causes a decreased sensitivity to Pᵢ inhibition for the cyanobacterial enzyme but increased the ability of Pᵢ to inhibit the spinach leaf activated ADPGlc synthetase activity. These results suggest that PGO interferes primarily with the enzyme's allosteric properties and that arginine residues are involved in maintaining the regulatory properties of the ADPGlc synthetase. When > 90% of the enzyme activity was inhibited about 1.5 mol of [14C]phenylglyoxal were bound per mole of subunit and both subunits of the spinach leaf enzyme were labelled.

**Studies on the catalytic site of the spinach leaf ADPGlc synthetase using the photoaffinity substrate analogue, 8-azido-ATP**

8-Azido-ATP(8-N₃-ATP) has been shown to be a substrate for the spinach leaf ADPGlc synthetase with a Kₘ larger than that of the natural substrate, ATP [36]. The Kₘ is 0.81 μM for 8-N₃-ATP and 0.12 μM for ATP. Moreover, the Vₘₐₓ with 8-N₃-ATP is only 1% of that observed with ATP. The product of the reaction, 8-N₃-ADPGlc, has a Kₘ for the spinach leaf ADPGlc synthetase of 80 μM whereas the Kₘ for ADPGlc is 225 μM. The maximal velocity rate for 8-N₃-ADPGlc is only 0.3% of that for ADPGlc. When u.v. light (257 nm) irradiates 8-azido compounds, a nitrile radical is formed which can react with electron-rich residues. Indeed it is so reactive that it can form a secondary amine with a C–H linkage. Preliminary experiments [36] have shown that labelled 8-N₃-ADPGlc covalently link to the spinach leaf ADPGlc synthetase and inactivate it upon u.v. irradiation. The substrate, ADPGlc, was very effective in inhibiting the chemical modification by the azido compound while UDPglucose, a non-substrate, did not prevent the inactivation. The labelled 8-azido-ADPGlc is incorporated mainly, if not solely, in the 54 kDa subunit [36]. ADPGlc inhibits the incorporation by at least 67%. It thus appears that the substrate binding site is on the 54 kDa subunit. It is probable, however, that the substrate/catalytic site is shared between the two subunits and that the nitrene radical reacts with a strong nucleophilic residue on the larger subunit causing the incorporation to be mainly on the larger subunit.

**Isolation of the ADPGlc synthetase gene**

Cloning of ADPGlc synthetase genes and their sequence comparisons: cDNA or genomic clones for the small subunit ADPGlc synthetase gene of rice endosperm [33, 37, 38], maize endosperm [29], spinach leaf [36], *Arabidopsis thaliana* (A. Lonnewbo, private communication), maize endosperm [39, 40] and potato tuber [30, 41] have been isolated. In addition, cDNA clones for the maize endosperm ADPGlc synthetase larger molecular mass subunit [29, 39] (shrunken 2 locus) and potato tuber have also been isolated [41]. cDNA clones have been isolated from wheat leaf and wheat endosperm [42]. The deduced amino acid sequence derived from the wheat endosperm cDNA would suggest that it is more homologous with the deduced amino acid sequence of the maize endosperm cDNA of the larger (shrunken 2) subunit than the rice seed cDNA of the smaller synthetase subunit gene suggesting that they are representative of the large subunit peptide. While no cDNA clone has been isolated representing the spinach leaf large subunit, the major portion of the spinach leaf large molecular mass (54 kDa) subunit has been sequenced by Edman degradation technique (K. Ball, J. Hutny, J. Leykam & J. Preiss, unpublished work).

At the DNA level the isolated genes are quite dissimilar. For wheat leaf and wheat endosperm there is only 56% identity [42] and on the basis of Southern hybridization analyses and restriction enzyme mapping, it is concluded that there are at least two distinct gene families in wheat. For spinach leaf and rice endosperm there is only about an 86% identity between the genes (B. Smith-White & J. Preiss, unpublished work). The deduced amino acid sequence of the clones from rice and wheat endosperm, spinach, wheat and *Arabidopsis thaliana* leaf and potato tuber have been compared. Two endosperm cDNA clones were isolated [42] and they represented two closely related gene subfamilies in wheat endosperm. One, AGA.3, had a DNA insert, 1272 bp in length, with an open reading frame of 888 bp encoding 296 amino acids. The other, AGA.7, is 1798 bp in length comprising an open reading frame of 1500 bp encoding 500 amino acids. The clones were considered incomplete but it was estimated that the AGA.7 insert was nearly complete, possibly lacking some of the transit peptide portion. These wheat endosperm clones are 96% identical in shared regions of their open reading frames and differ in only 10 of the 296 amino acid residues. In contrast, there is only a 54% identity between the wheat leaf, AGA.1 DNA-deduced amino acid sequence and that from AGA.7. AGA.1
is 947 bp in size and encodes for 301 amino acids. Comparisons of the various clones and their proteins were made with the deduced amino acid sequence derived from the rice endosperm cDNA clone which was a full length clone of 1647 bp and coded for 483 amino acids [34]. AGA.1, AGA.3 and AGA.7 showed only about 43%, 43% and 44% identity respectively, with respect to amino acids with the rice seed amino acid sequence. If similar amino acids were considered, the identity between the sequences and the rice sequence is about 62, 64 and 65%, respectively. Greater identity was seen with the potato, spinach and Arabidopsis leaf cDNA clones with respect to amino acid sequences. The cDNA clone of potato tuber is almost a full length clone which codes for 444 amino acids [30, 41], while the spinach leaf clones are incomplete, with two of them encoding for a total of 450 amino acids [36]. The Arabidopsis genomic clone is incomplete and its sequence encodes 240 amino acids (A. Lonneborg, private communication). Of interest, is that for the potato tuber, spinach leaf 51 kDa subunit and Arabidopsis leaf deduced amino acid sequences there is about 90%, 86% and 88% identity, respectively, with the rice seed amino acid sequence. If similar amino acids are considered, the homology is then, respectively, 94%, 93% and 93%. All of these three last mentioned clones have been shown to be representative genes of the smaller subunit of the ADPGlc synthetase. In contrast, less identity and similarity is seen with the derived amino acid sequences of the cDNA clone of the maize shrunken 2 gene protein (higher molecular mass) subunit of the ADPGlc synthetase with the rice seed and spinach leaf cDNA clones derived amino acid sequences. The shrunken 2 sequence [39] codes for 542 amino acids. Included in the comparison were the sequences of the 54 kDa, high molecular mass subunit of the spinach leaf enzyme obtained by automated Edman degradation procedures (K. Ball, J. Leykam, T. Jones & J. Preiss, unpublished work) and then aligning the peptide sequences to the derived sequences of the rice seed and spinach leaf lower molecular mass subunit. There is only a 44% and 46% identity seen in the amino acid sequences of shrunken 2 protein with those of rice seed and spinach leaf, respectively. If similar amino acids are considered, then the similarity is 67% and 68%, respectively. Similarly, about 42% identity is observed for the spinach leaf peptides (total of 227 amino acids) of the 54 kDa subunit compared with the amino acid sequences of rice seed and spinach leaf. If homologous amino acids are considered then the similarity is about 50 to 54%. Comparison of the shrunken 2 sequence with the spinach leaf 54 kDa subunit shows a higher identity. About 288 amino acids can be compared and 52% identity is seen. When similar amino acids are considered, 63% similarity is found. Also of interest is the comparison of the amino acid sequences seen for the wheat endosperm cDNA clone, AGA.7 and the shrunken 2 cDNA clone. There is a 66% identity in sequence and similarity is 78% when similar amino acids are considered. Thus, very good identity is observed in comparing similar subunits of the ADPGlc synthetase from the different plants, and this is expected as the spinach leaf lower molecular mass subunit antibody reacts very well with the equivalent subunits of maize endosperm [28], rice seed [33, 37], Arabidopsis [4, 14] and potato tuber [31] enzymes. The lower molecular mass subunit antibody does not react well with the higher molecular mass subunit of the ADPGlc synthetase of these various plants. Therefore it was not expected to see much homology between the lower and higher molecular mass subunits. However, there appears to be identity of about 40% between the 54 and 51 kDa subunits of the spinach leaf ADPGlc synthetase. The sequence analyses point out a greater relationship in identity between the spinach leaf 54 kDa subunit and the maize shrunken 2 subunit and most certainly between the wheat endosperm subunit encoded by the cDNA insert, AGA.7. Most probably, the isolated cDNA clones of wheat endosperm are representative of the large molecular mass subunit of the wheat endosperm ADPGlc synthetase. Because of the low but certain similarity between the two subunits of the ADPGlc synthetase it can be speculated that they may have arisen originally from the same gene. The bacterial ADPGlc synthetase has been purified to homogeneity from seven different and diverse species and in each case has been shown to be a homotetramer composed of only one subunit [19, 20]. Recent results with cyanobacterial ADPGlc synthetase, which has 3PGA as an allosteric activator and P, as an inhibitor, similar to the higher plant enzyme [43], suggests that it is homotetrameric in structure unlike the higher plant enzymes (A. Iglesias, G. Kakefuda & J. Preiss, unpublished work). Thus it is quite possible that during evolution, there was gene duplication in the higher plant photosynthetic systems of the synthetase gene and then divergence of the genes to produce two different peptides for the native ADPGlc synthetase. Yet, both subunits are required for optimal activity of the native plant enzyme. The reason for this gene duplication is
presently unknown but conservation of amino acid sequences may point out the particular importance of those sequences.

**Comparison of the plant and bacterial ADPGlc synthetase amino acid sequences**

The rice seed ADPGlc synthetase cDNA clone corresponding to the small molecular mass, 51 kDa subunit has been compared with the deduced amino acid sequence of the *E. coli* *glg* C gene, the structural gene for ADPGlc synthetase [37]. There is an overall identity of about 29% at the amino acid level between the plant and bacterial subunits suggesting a common origin for these two genes. Divergence appears to have occurred in a non-random manner for when the peptides are matched in residues 48 to 227 of rice seed to 19 to 196 of *E. coli* and rice seed residues 297 to 394 to *E. coli* amino-acid residues 264 to 368, the degree of identity is higher, about 43%. Regions of largest divergence are at the N-terminus, amino acid residues 29 to 47 of rice seed, a central portion between rice seed residues 228 and 296 and at the C-terminus beginning with Met 395. Where significant identity occurs in region 48 to 227 for the rice seed is of great interest in that, for the to 47 of rice seed, a central portion between rice and bacterial enzyme, Lys 39, where 17 of the 24 amino acids are the same in the equivalent rice seed sequence, residues 54-77. The potato tuber and spinach leaf subunits have the same identical sequences as the rice seed subunit. Fructose 1,6-bis-P is an activator of the various higher plant ADPGlc synthetases [22, 44, 45]. However, it is not as potent an activator as 3PGA. The tyrosine residue 114 in *E. coli* has been shown to be involved in binding of the substrates and activator, as changing it to Phe via site-directed mutagenesis caused a lowering of the apparent affinities for substrates and activator [46]. Although there is a change to Phe for the rice seed, spinach leaf and potato tuber subunits in their equivalent sequences, 9 of the 12 amino acids in these sequences are identical with the *E. coli* sequence (Fig. 2). The conservation of amino acids in this region as well as in the region equivalent to rice seed residues 65-74 in the plant enzymes suggest that these amino acids play a major role in maintenance of protein conformation and of the regulatory and catalytic functions of the plant enzyme having 3PGA as the major activator and P as the inhibitor. Whereas the activator site for the *E. coli* enzyme is at the N-terminus, the activator site for spinach leaf enzyme is at the C-terminus [27, 34]. The reactive residue is Lys 447 in the 51 kDa subunit of the spinach leaf enzyme and the equivalent lysine in the rice seed enzyme is residue 469 (Fig. 2). The sequence around the reactive lysine is also observed in the derived amino acid sequences from the cDNA and genomic clones of either the high or low molecular mass subunits. Thus the extended portion of the plant peptide is not observed for the bacterial enzyme. The change in activator specificity may then be owing to an extension of the C-terminal end and possibly other amino acid replacements in other regions of the plant enzyme. This possibility can be tested via construction of a recombinant enzyme where the DNA fragment transcribing the higher plant C-terminus is fused to the mutant and wild-type *E. coli* enzyme gene and analysing the regulatory kinetics of the hybrid enzyme.

The above comparisons of the primary structures of 11 ADPGlc pyrophosphorylases thus indicate that many domains of the protein are conserved from both the enteric bacteria and angiosperm plants because of catalytic activity as well as
Comparison of amino acid sequences of regulatory and catalytic sites of *Escherichia coli* and plant ADP-Glc synthetases

*E. coli* Lys 39 (*) is the amino acid shown to bind to the activator, fructose 1,6-bis-P. The equivalent lysine in the spinach leaf enzyme is Lys 38. Lys 47 in the spinach leaf enzyme has been shown to be involved in the binding of the activator, 3PGA. Tyr 114 and Lys 195 of the *E. coli* ADP-Glc synthetase have been shown to be involved in the binding of the substrates ATP, ADP-Glc and glucose-1-P.

**E. coli activator site**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Spinach leaf</th>
<th>Rice/potato</th>
<th>Spinach leaf/A.t.</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>30</td>
<td>39</td>
<td>*</td>
</tr>
</tbody>
</table>

**Rice endosperm 51 kDa subunit**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Spinach leaf</th>
<th>Rice/potato</th>
<th>Spinach leaf/A.t.</th>
</tr>
</thead>
<tbody>
<tr>
<td>53</td>
<td>*</td>
<td>24</td>
<td></td>
</tr>
</tbody>
</table>

**Plant ADP-Glc synthetase allosteric activator sites**

<table>
<thead>
<tr>
<th>Site</th>
<th>Spinach leaf</th>
<th>Rice endosperm</th>
<th>Potato tuber</th>
<th>Maize endosperm</th>
</tr>
</thead>
<tbody>
<tr>
<td>51</td>
<td>SGGVTVYETALPSOTVI</td>
<td>SGGVTVYETALL</td>
<td>SGGVTVYETPSGIII</td>
<td>SGGVTVYETALPSOTVI</td>
</tr>
<tr>
<td>55</td>
<td>SGGVTVYETALL</td>
<td>SGGVTVYETPSGIII</td>
<td>SGGVTVYETALPSOTVI</td>
<td></td>
</tr>
</tbody>
</table>

**Wheat endosperm subunit**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Spinach leaf</th>
<th>Wheat leaf subunit</th>
<th>Maize endosperm 60 kDa subunit</th>
</tr>
</thead>
<tbody>
<tr>
<td>440</td>
<td>SGGVTVYETALPSOTVI</td>
<td>SGGVTVYETMAKEIY</td>
<td>SGGVTVYETNATHELCLVI</td>
</tr>
<tr>
<td>462</td>
<td>SGGVTVYETALL</td>
<td>SGGVTVYETMAKEIY</td>
<td></td>
</tr>
</tbody>
</table>

**E. coli catalytic sites**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Spinach leaf</th>
<th>Rice</th>
<th>Potato</th>
<th>Spinach-51</th>
<th>Arabidopsis</th>
<th>Brittle 2</th>
<th>Spinach-54</th>
<th>Wheat-E</th>
<th>Wheat-L</th>
<th>Shrunken 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>112</td>
<td>NYKROTADEFQP</td>
<td>119</td>
<td>139</td>
<td>143</td>
<td>143</td>
<td>143</td>
<td>143</td>
<td>143</td>
<td>143</td>
<td>143</td>
</tr>
</tbody>
</table>

**Fig. 2**

In addition, the nucleotide sequences of the genes of eleven of these proteins were compared for relationships between them and the analysis suggests that the protein for the small subunit has been subjected to greater selective pressure to retain a particular primary structure.

16. Reference deleted
17. Reference deleted
24. Preiss, J., Bloom, M., Morell, M., Knowles, V., Plaxton, W. C., Okita, T. W., Larsen, R., Harmon,
Abbreviation used: SBE, starch-branching enzyme.

The pathway of starch synthesis in developing pea embryos

Alison M. Smith, Cathie Martin and Kay Denyer
John Innes Institute, Colney Lane, Norwich NR4 7UH, U.K.

Introduction
The aim of our work is to discover what regulates the amount and structure of starch that accumulates in the pea embryo during its development. This organ provides an excellent model system for the study of starch synthesis. At maturity half of its dry weight is starch, and high rates of synthesis are maintained over a long developmental period [1]. Many aspects of its development have been characterized in detail, and well-defined stages can be easily recognized [2–4]. The tissue is biochemically very amenable, in that interference by proteases and secondary metabolites during protein extraction and purification is minimal. Activities of the enzymes of starch degradation are generally low during development, facilitating the assay of the enzymes of starch synthesis. Analysis of the pathway of starch synthesis is also facilitated by the occurrence of two mutations – at the r and rb loci – which affect directly the activity of enzymes in the pathway [1, 5]. Wild-type and mutant lines of peas that are essentially isogenic except at one or both of these loci have recently been developed at the John Innes Institute.


Received 12 April 1991