Phosphoenolpyruvate Carboxylase Kinase in Higher Plants

by

Justin Marsh

Thesis submitted for the degree of doctor of philosophy

Division of Biochemistry and Molecular Biology

Institute of Biomedical and Life Sciences

University of Glasgow

September, 2002

Contents of Thesis

List of Figures and Tables

Chapter 4

Chapter 5

Chapter 6

Appendices

Abbreviations

The abbreviations used in this thesis are described in full below:

ADP, adenosine diphosphate; AMP, adenosine monophosphate; AMPK, AMP-activated protein kinase; ATP, adenosine triphosphate, [α-^{32}P]ATP, radioactively labelled adenosine triphosphate; ATPase, adenosine triphosphatase; CAM, Crassulacean acid metabolism; cAMP, cyclic adenosine monophosphate; cDNA, complementary DNA; CDPK, calcium-dependent protein kinase (or calmodulin-like domain protein kinase); DMSO, dimethyl sulphoxide; DNA, deoxyribonucleic acid; EtOH, ethanol; est, expressed sequence tag; GA, gibberellic acid; Gln, glutamine; Glu, glutamate; GOGAT, glutamine 2-oxoglutarate aminotransferase (glutamate synthase); GS, glutamine synthetase; HMG-CoA reductase, 3-hydroxy-3-methylglutaryl-coenzyme A reductase; KAPP, kinase associated protein phosphatase; MAPK, MAPKK or MAPKKK, mitogen activated protein kinase, kinase kinase or kinase kinase kinase; ME, malic enzyme; MDH, malate dehydrogenase; MeOH, methanol; [^{35}S]Met, radioactively labelled methionine; NAD^+/NADH, nicotinamide adenine dinucleotide oxidised/ reduced form; $NADP^+$/NADPH, nicotinamide adenine dinucleotide phosphate oxidised/reduced form; NR, nitrate reductase; OAA, oxaloacetate; PDC, pyruvate dehydrogenase complex; PEP, phosphoenolpyruvate; PEPc, phosphoenolpyruvate carboxylase; PEPCK, phosphoenolpyruvate carboxykinase; Pi, inorganic phosphate; PKA, PKG and PKC, protein kinase A, G and C; PPDK, pyruvate, phosphate dikinase; PPi inorganic pyrophosphate; PP1, protein phosphatase type-1; PP2A, 2B, 2C, protein phosphatase type-2A, -2B or -2C; RLK, receptor-like protein kinase; RNA, ribonucleic acid (mRNA, messenger ribonucleic acid); RNase, ribonuclease; Rubisco, ribulose bisphosphate carboxylase/oxygenase; SA, salycylic acid; SDS-PAGE, sodium dodecyl sulphate polyacrylamide gel electrophoresis; SNF-1, sucrose non-fermenting protein kinase-1; SPS, sucrose phosphate synthase; SRK, S-locus receptor kinase; SS or SuSy, sucrose synthase; TP, triose phosphate; UDP, uridine diphosphate.

Throughout this thesis I will use italicised abbreviations for gene names. For example, the *Kalanchoe fedtschenkoi* Phosphoenolpyruvate carboxylase kinase gene will be shortened to *KfPPCK*. I have also abbreviated the protein (and transcript) names in this thesis. This will be non- italicised. For example, KfPPCK.

Acknowledgements

I would like to acknowledge my supervisor, Prof. Hugh. G. Nimmo for guidance and support throughout my research. This was particularly helpful when positive experimental results appeared myth-like.

I gratefully acknowledge the financial support of the BBSRC.

I would also like to thank members of the Bond lab, both past and present, for ensuring a happy and 'relaxed' workplace and their generous help and ideas. I would particularly like to thank Jim Jardine, Stuart Sullivan, Dr Matt Shenton, Dr Veronique Fontaine, Dr Gill Nimmo and Dr James Hartwell for their experimental suggestions and demonstrations.

Abstract

Phosphoenolpyruvate carboxylase (PEPc) is a ubiquitous enzyme in higher plants. Several isoforms of PEPc exist in plants. They carry out a variety of roles including the anapleurotic role of re-supplying the tricarboxylic acid cycle with four-carbon intermediates that are depleted by aminoacid biosynthesis. PEPc is also involved in processes requiring four-carbon acid accumulation, such as ripening. In C_4 and CAM plants, PEPc has a further photosynthetic role in performing the primary fixation of CO_2. PEPc is regulated by reversible phosphorylation, which modulates the allosteric properties of the enzyme. The phosphorylated form is less sensitive to negative feedback inhibition by malate. The phosphorylation state of PEPc is primarily controlled by the activity of its kinase, phosphoenolpyruvate carboxylase kinase (PPCK). This Ca^{2+}-independent kinase, with minimal N- or C-terminal extensions is in turn dependent upon *de novo* protein synthesis for its activity. This activity is controlled by a circadian oscillator in CAM plants and induction by light in C_4 and C_3 plants. The aim of this work was to identify and characterise PPCK in C_3 and C_4 plants.

A partial maize PPCK was identified by est database mining. A full-length clone was obtained by RACE PCR and the cDNA clone was sequenced, revealing high similarity to other known PPCKs. To prove function, the cDNA clone was *in vitro* transcribed and translated and shown to phosphorylate PEPc. The gene expression was then studied in mature leaves by RT-PCR. Analysis revealed that this PPCK was not upregulated in the light. It is therefore thought not to be the photosynthetic relevant PPCK but a separate isoform with a possible house-keeping role *in vivo*. Several other ests from the maize database show similarity to PPCK and are in the process of being investigated by members of our lab.

A similar database mining approach revealed two putative PPCK in the tomato database. Full-length genomic and cDNA clones were generated and sequenced. The cDNA clones were tested for PPCK activity as the maize PPCK was. The first of these, *LePPCK1*, showed PPCK activity. It bears high similarity to other PPCKs genes with a small 3' end intron. The second, *LePPCK2*, appears to be a novel PPCK gene. RT-PCR analysis and sequencing show that a second intron exists in this gene that is subject to alternative splicing. This intron, near the middle of the gene, allows the introduction of a premature stop codon into the open reading frame, truncating the kinase, thereby preventing PPCK activity. Correct splicing of this intron though,

allows PPCK activity. This is demonstrated in a similar fashion to the aforementioned maize PPCK. Expression analysis of these genes was undertaken by semi-quantitative RT-PCR analysis over a range of tissues and compartments of the tomato plant. This revealed that *LePPCK1* is expressed at similar rates across tissues suggesting that its role is anapleurotic. The expression of *LePPCK2* is induced in post-breaker fruit. In combination, the alternative splicing of the pre-mRNA appears to be dependent upon tissue type and compartment type as suggested by semi-quantitative RT-PCR transcript ratios and transcripts in red-ripe fruit compartments. Furthermore, these types of PPCK genes appear to exist in other *Solanaceae* plants, suggesting a PPCK sub-family.

A phylogenetic analysis of all known PPCKs (and ones recently identified by myself and members of our lab) is also undertaken. This reveals a number of interesting traits. Perhaps the most interesting is the above average GC content of monocot PPCKs of ~70%. Collectively these results indicate that monocots and dicots probably hold multigene PPCK families.

Chapter 1

Introduction

This thesis deals with several aspects of the control of phosphoenolpyruvate carboxylase kinase (PPCK) in higher plants. The activity of PPCK is of fundamental importance to the regulation of several metabolic pathways within higher plants. In this chapter I describe the potential roles of this enzyme by its modification of the several phosphoenolpyruvate carboxylase (PEPc) isoforms found in C_3, C_4 and CAM plants. Since many of the recent advances in this area concern the control of photosynthetic isoforms of PEPc in C_4 and CAM, I start by reviewing photosynthesis in the different types of higher plant (1.1 - 1.3). I then discuss recent advances in the understanding of PEPc and PPCK (1.4 - 1.6). Some of the data presented in this thesis demonstrate alternative splicing of PPCK transcripts in tomato. I therefore include a short section (1.7) reviewing alternative splicing in higher plants.

1.1 C_3 pathway of photosynthesis

1.1.1 Biochemistry of the C_3 pathway

In the majority of higher plants the first stable product of photosynthesis following the primary fixation of CO_2 is a three carbon compound. These plants are therefore called C_3 plants. The reactions comprising photosynthesis take place within the mesophyll cells of the leaf (see Figure 1.1). The reactions are further compartmentalised within the mesophyll cell chloroplasts. This specialised organelle separates the photochemical [light] reactions, which occur in the thylokoid and grama, from the biochemical [dark] reactions that reduce CO_2 to carbohydrate in a reductive pentose phosphate cycle (RPP), in the stroma of the chloroplast (Macdonald and Buchanan, 1990).

The RPP cycle begins with the covalent attachment of CO_2 to ribulose-1, 5-bisphosphate (RubP). This reaction is catalysed by the enzyme ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco). The unstable six-carbon compound formed by Rubisco breaks down into two molecules of 3-phosphoglycerate (3-PGA), the first stable product of photosynthesis. The energy captured by the photochemical reactions is used to generate reduced nicotinamide adenine dinucleotide phosphate (NADPH) and adenine 5'-triphosphate (ATP). NADPH is used in the reduction of 3-PGA to glyceraldehyde-3-phosphate (G-3-P). In net terms, one in six molecules of G-3-P is converted to dihydroxyacetone phosphate (DHAP) and represents the net gain

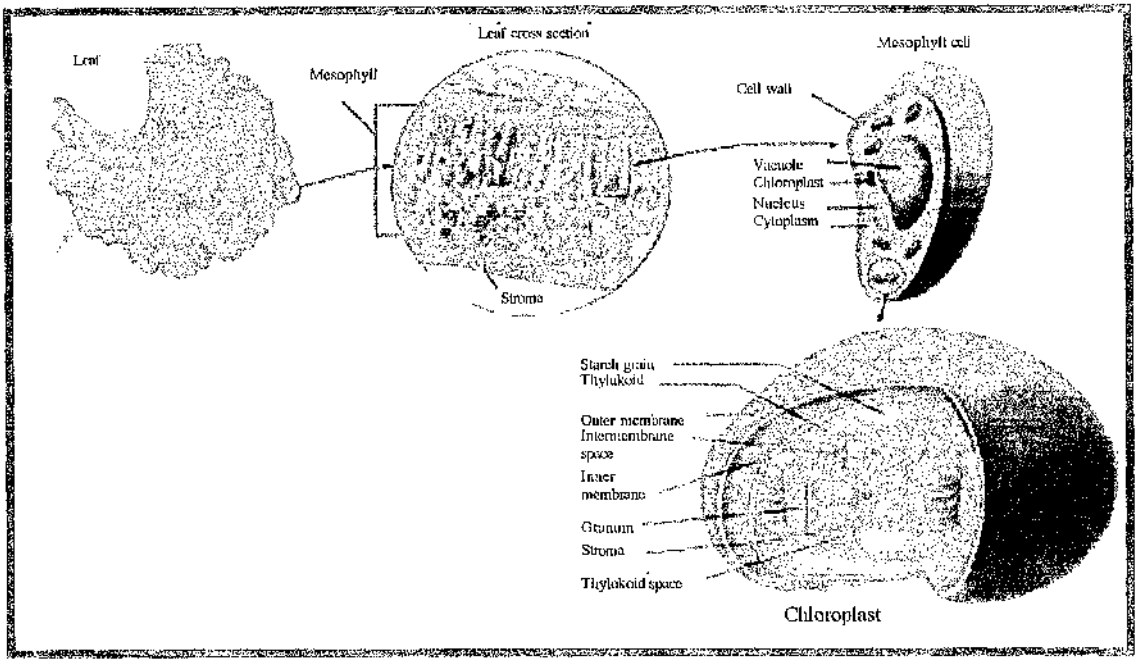
Figure 1.1 Summary of photosynthesis within the mesophyll cell chloroplast

Panel A summarises the anatomy of a leaf, a mesophyll cell and a chloroplast.

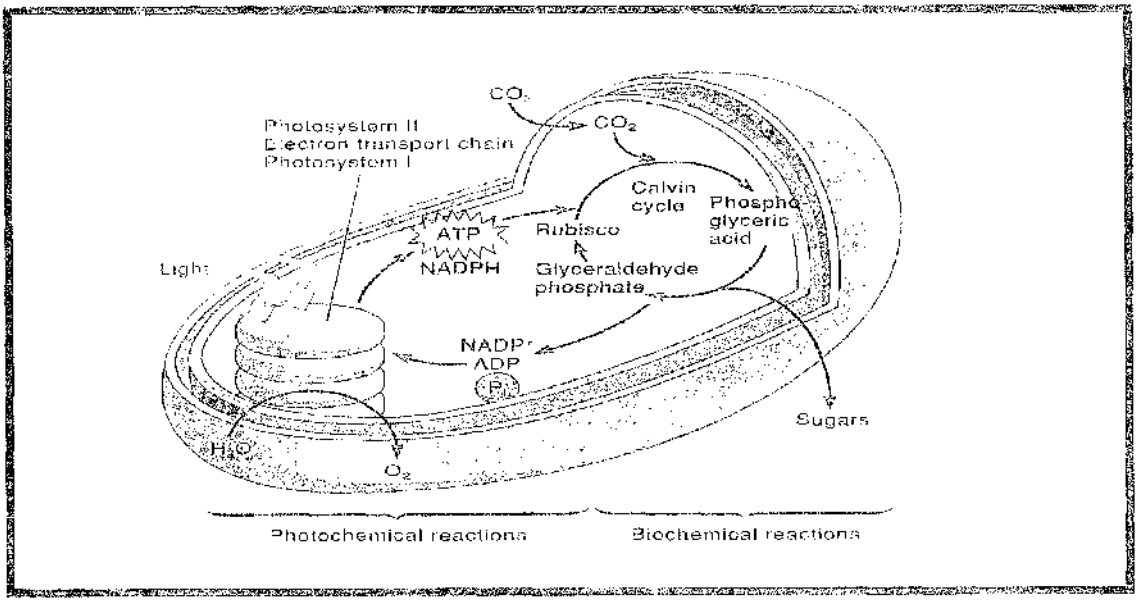
Panel B shows a summary of photosynthesis demonstrating the separation of the dark (biochemical) and light (photochemical) reactions within the chloroplast.

This figure was adapted from Moore *et al.*, 1998.

A


Leaf
Mesophyll
Leaf cross section
Stroma
Mesophyll cell
Cell wall
Vacuole
Chloroplast
Nucleus
Cytoplasm
Starch grain
Thylakoid
Outer membrane
Intermembrane space
Inner membrane
Granum
Stroma
Thylakoid space
Chloroplast

B


CO2
CO2
Photosystem II
Electron transport chain
Photosystem I
Calvin cycle
Phospho glyceric acid
ATP
NADPH
Rubisco
Glyceraldehyde phosphate
Light
NADP+
ADP
Pi
Sugars
H2O
O2
Photochemical reactions
Biochemical reactions

of fixed carbon from the RPP cycle. The remaining G-3-P molecules are used to regenerate the RubP (Macdonald and Buchanan, 1990). A summary of this pathway is shown in figure 1.2.

1.1.2 Photorespiration

In 1920, Warburg showed that photosynthesis is inhibited in high concentrations of oxygen (O_2). This was early evidence of the consequences of Rubisco's oxygenase activity. Rubisco can also catalyse the oxygenation of RubP, forming one molecule of 3-PGA and one molecule of 2-phosphoglycolate and leading to photorespiration (Canvin, 1990). At a fixed temperature and normal concentrations of atmospheric O_2, the carboxylase activity of Rubisco is less than the oxygenase activity at concentrations of CO_2 less than about 50 ppm. Furthermore, Rubisco's affinity for O_2 relative to its affinity for CO_2 increases with high temperature. Hence low atmospheric CO_2 and high temperature favour photorespiration over the RPP cycle. Thus the leaf will achieve a steady state in which the proportion of oxygenation to carboxylation is determined by temperature and internal CO_2. This in turn depends on factors such as stomatal aperture and atmospheric CO_2.

In photorespiration RubP is oxygenated, forming 3-PGA and 2-phosphoglycolate acid (2-PG). 3-PGA is used to regenerate RubP, but 2-PG is oxidised via the photosynthetic carbon oxidation (PCO) system. The PCO system involves the chloroplasts, peroxisomes and mitochondria (see Figure 1.3). 2-PG is converted to glycine in a series of steps within the peroxisome before transport to the mitochondria. Within the mitochondria two molecules of glycine combine to form serine, CO_2 and NH_3. Serine is then converted to 3-PGA at the expense of a molecule of ATP and is used to replenish the RPP cycle. Although some of the NH_3 produced in photorespiration is reassimilated by glutamine synthesase to provide glutamine for continuation of the PCO cycle, much is lost in this action (Canvin, 1990).

The oxygenase activity of Rubisco means that the majority of plants are not well suited to a hot or dry environment. However a number of plants have adapted to these conditions by evolving a specialised photosynthetic biochemistry, physiology and/or anatomy to overcome Rubisco's failings and ensure that the ratio of CO_2 to O_2 around Rubisco is high. These are the Crassulacean Acid Metabolism (CAM) and C_4 plants.

Figure 1.2. The C_3 pathway of photosynthesis (the reductive pentose phosphate pathway)

Abbreviations: RubP, ribulose 1,5-bisphosphate; PGA, 3-phosphoglycerate; bis-PGA, 1,3-bisphosphoglycerate; Ru5P, ribulose 5-phosphate; G3P, glyceraldehyde 3-phosphate; DHAP, dihydroxyacetone phosphate; FbP, fructose 1,6-bisphosphate; F6P, fructose 6-phosphate; Xu5P, xylulose 5-phosphate; E4P, erythrose 4-phosphate; SbP, sedoheptulose 1,7-bisphosphate; S7P, sedoheptulose 7-phosphate; R5P, ribose 5-phosphate; G6P, glucose 6-phosphate; G1P, glucose 1-phosphate; 2PG, 2-phosphoglycolate.

Here and in the following figures oxidised cofactors are written as NAD etc rather than NAD^+.

The enzymes which catalyse each step are as follows:

1. ribulose 1,5-bisphosphate carboxylase/oxygenase
2. phosphoglycerate kinase
3. glyceraldehyde 3-phosphate dehydrogenase
4. triose phosphate isomerase
5. aldolase
6. fructose 1,6-bisphosphatase
7. transketolase
8. sedoheptulose 1,7-bisphosphatase
9. phosphopentoepimerase
10. phosphoriboisomerase
11. phosphoribulokinase
12. hexose phosphate isomerase
13. phosphoglucomutase
14. ADPglucose pyrophosphorylase and starch synthase

This figure was adapted from Moore *et al.*, 1998.

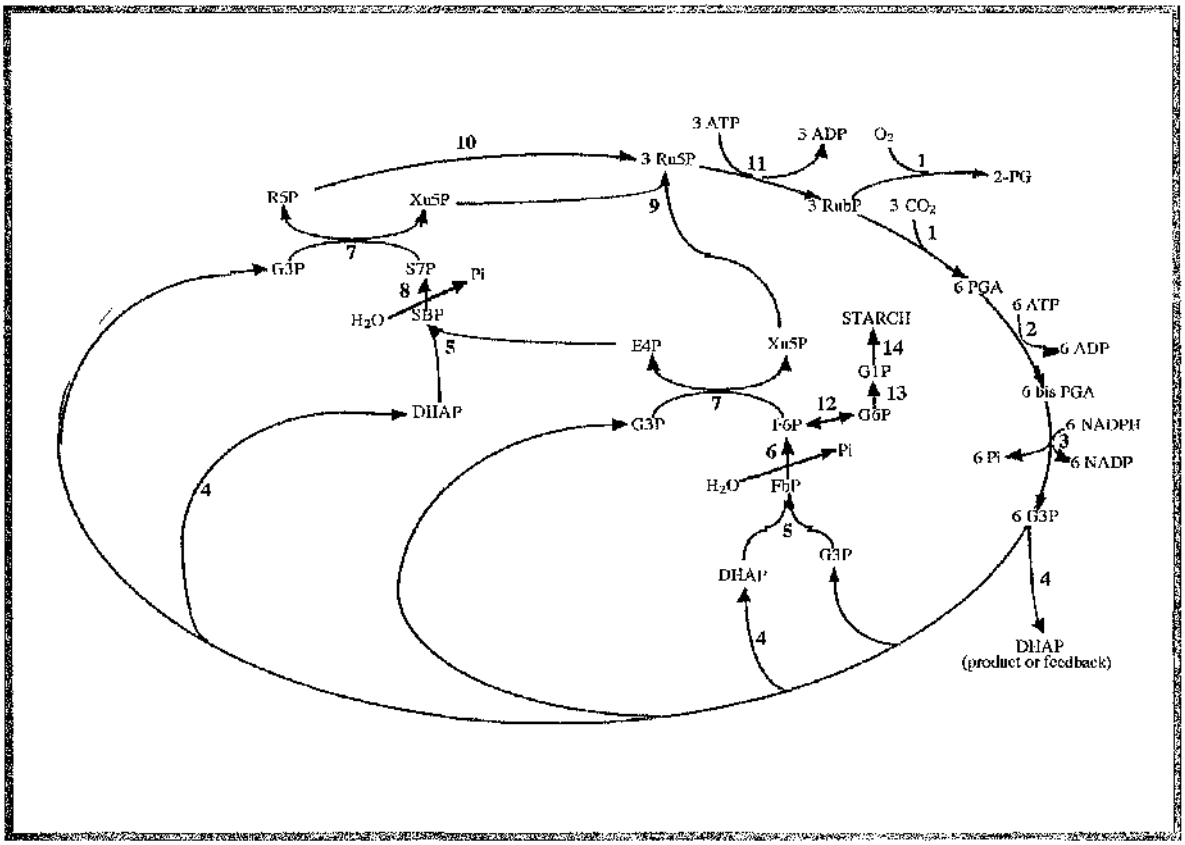
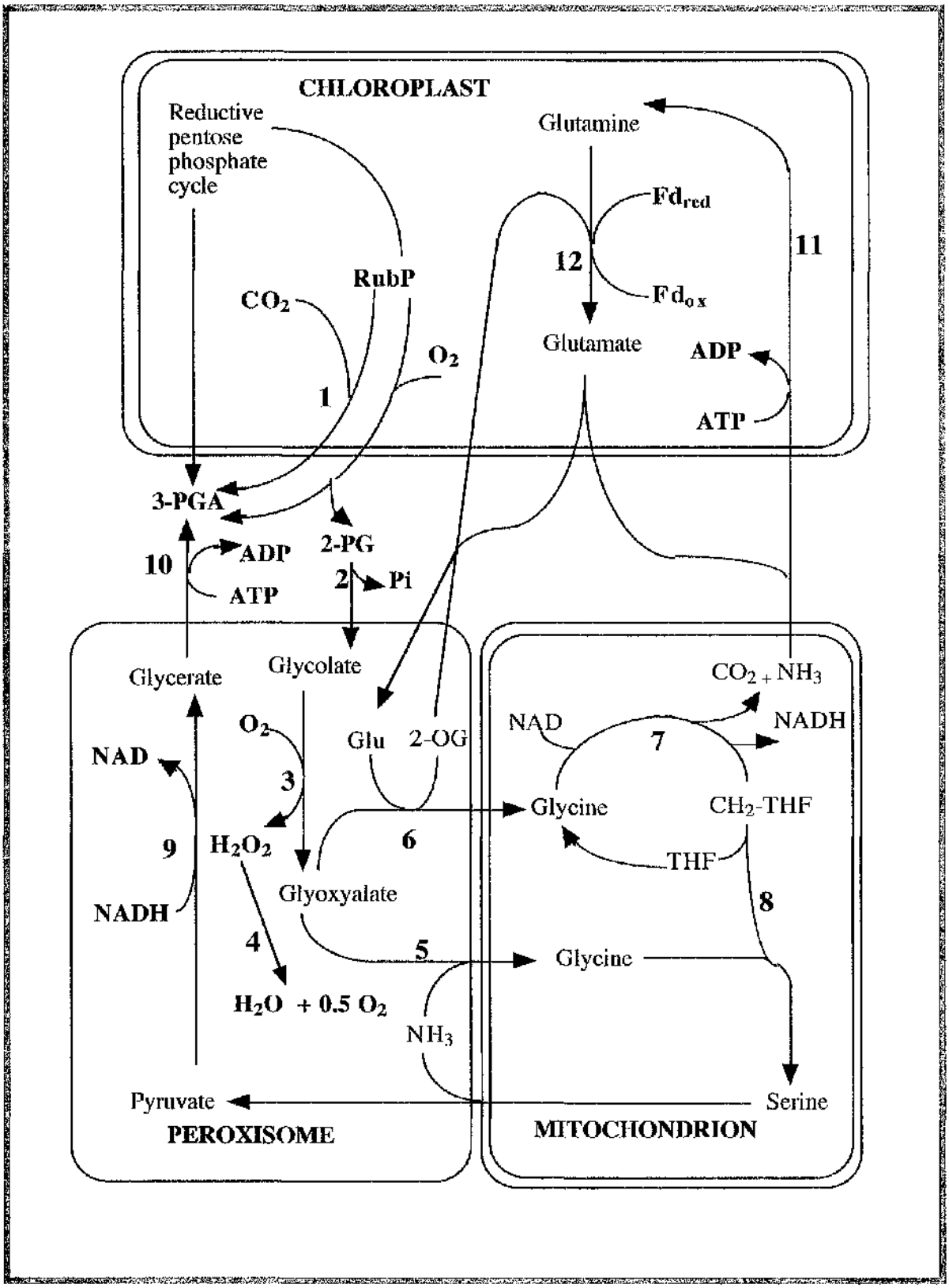

3 ATP
3 ADP
O_2
10
3 Ru5P
11
1
2-PG
R5P
Xu5P
9
3 RuBP
3 CO_2
1
7
G3P
S7P
Pi
8
6 PGA
6 ATP
2
H_2O
SBP
STARCH
6 ADP
5
E4P
Xu5P
14
G1P
6 Bis PGA
DHAP
7
12
13
G6P
6 NADPH
G3P
F6P
3
6
Pi
6 Pi
6 NADP
H_2O
FbP
6 G3P
5
DHAP
G3P
4
4
4
DHAP
(product or feedback)

Figure 1.3. The photosynthetic carbon oxidation (PCO) cycle

Abbreviations: RubP, ribulose 1,5-bisphosphate; 2PG, 2-phosphoglycolate; PGA, 3-phosphoglycerate; THF, tetrahydrofolic acid; Fd(red or ox), ferredoxin (reduced or oxidised). The figure was adapted from Canvin (1990).

The enzymes which catalyse each step are as follows:

1. ribulose 1,5-bisphosphate carboxylase/oxygenase
2. phosphoglycolate phosphatase
3. glycolate oxidase
4. catalase
5. serine: glyoxylate aminotransferase
6. glutamate: glyoxylate aminotransferase
7. glycine decarboxylase
8. serine hydroxymethyltransferase
9. hydroxypyruvate reductase
10. glycerate kinase
11. glutamine synthetase
12. glutamate synthase


CHLOROPLAST
Reductive pentose phosphate cycle
Glutamine
Fd_{red}
11
12
RubP
CO_2
Fd_{ox}
O_2
Glutamate
ADP
1
ATP
3-PGA
2-PG
10
ADP
2
Pi
ATP
Glycerate
Glycolate
$CO_2 + NH_3$
O_2
Glu
2-OG
NAD
NADH
7
NAD
3
Glycine
CH_2-THF
9
H_2O_2
6
THF
NADH
Glyoxylate
8
4
5
Glycine
H_2O + 0.5 O_2
NH_3
Pyruvate
Serine
PEROXISOME
MITOCHONDRION

1.2 C_4 pathway of photosynthesis

1.2.1 Physiology of the leaves of C_4 plants

Hatch and Slack (1970) found that the primary product of photosynthesis in sugarcane was a C_4 compound, not the expected C_3 compound. Since then many other plants have been found to show C_4 metabolism. Almost all C_4 plants possess a specialised leaf anatomy termed 'Kranz' anatomy. In Kranz anatomy a set of cells with thickened cell walls and prominent starch filled chloroplasts, known as bundle-sheath cells, are arranged in a ring around the vascular strands. Beyond these cells lie the mesophyll cells, which are characterised by thin cell walls and small chloroplasts. These features can be understood in terms of the biochemistry of C_4 plants (McDonald and Buchanan, 1990). Perhaps surprisingly C_4 metabolism can be achieved in a single-cell system. In the family *Chenopodiaceae* the terrestrial plants *Borszczowia aralocaspica* and *Bienertia cycloptera* have a leaf anatomy of single elongated cells, arranged radially around the central water storage tissue. Chloroplasts are positioned at the centripetal end of the cell and are large, granal and hold high amounts of starch and Rubisco. Here there is no intercellular space. At the centrifugal end the intercellular space is exposed (Voznesenskaya *et al.*, 2001, 2002). Immunolocalization studies have shown that PEPc is distributed throughout the cytosol, while NAD-malic enzyme is specifically located in the mitochondria near the centripetal ends (Voznesenskaya *et al.*, 2001, 2002). The single cell appears to carry out C_4 photosynthesis by conducting the 'bundle-sheath' reactions in the centripetal end and the mesophyll reactions at the centrifugal end of the cell. C_4 metabolism has also been shown to occur in the single aquatic cell - *Hydrilla* (Bowes *et al.*, 2002). The absence of Kranz anatomy could be due to the submersed nature of the plant as the high diffusion resistance of the aqueous boundary reduces CO_2 leakage and thus the need for the bundle-sheath cell.

The C_4 pathway is essentially the RPP cycle with an addition of a CO_2 pumping system that generates a high concentration of CO_2 at the site of action of Rubisco. The unique leaf anatomy and biochemistry of C_4 plants allows the spatial separation of the primary fixation of carbon, within the mesophyll cells, from an operating RPP cycle, within the bundle-sheath cells (McDonald and Buchanan, 1990).

After CO_2 has entered the mesophyll cells, carbonic anhydrase (CA) catalyses its hydration to bicarbonate (HCO_3^-). The primary fixation of HCO_3^- occurs in the cytoplasm of the mesophyll cells. Phosphoenolpyruvate carboxylase (PEPc) catalyses

this fixation of HCO_3^-, using phosphoenolpyruvate (PEP) and forming oxoloacetate (OAA). OAA is rapidly converted to malate or aspartate, depending upon the C_4 subtype. This 4-carbon product is moved through the plasmodesmata, which links the mesophyll and bundle sheath cells. The malate (or aspartate) is rapidly decarboxylated in the bundle sheath cells, liberating CO_2, which is fixed by Rubisco in the RPP cycle. The plasmadesmata pores pierce a suberin layer, allowing only small metabolites to pass through. This allows a number of diffusive fluxes across the barrier. The decarboxylation step differs between C_4 subtypes, being catalysed by either NAD-dependent malic enzyme (NAD-ME), NADP-dependent malic enzyme (NADP-ME) or phosphoenolpyruvate carboxykinase (PEPCK). The pyruvate is transported across the plasmodesmata where it is converted to PEP by mesophyll pyruvate phosphate dikinase (PPDK), within the chloroplast to complete the cycle. Figure 1.4 shows an overview of this process (McDonald and Buchanan, 1990).

1.2.3 Regulation of the C_4 pathway

1.2.3.1 Transcriptional regulation in the C_4 pathway

The transcriptional regulation of C_4 cycle genes plays an important role in ensuring the spatial separation of the primary fixation of CO_2 from the RPP pathway. Both promoter control of the genes and processing of the ribonucleic acid (RNA) transcripts are utilised as control mechanisms.

Two CA genes have been identified in the C_4 plants *Flaveria bidentis* (a dicot), maize and *Sorghum bicolor* (both monocots). For one gene, expression in mature leaves is enhanced in the light and is mesophyll cell-specific (Sheen, 1999). Similarly, maize and *Sorghum* have been shown to have two malate dehydrogenase (MDH) genes, one of which shows specific expression in mesophyll cells in a light-inducible fashion (Luchetta *et al.*, 1990, Metzer *et al.*, 1989). However, *Flaveria* species apparently have only a single MDH gene (McGonigle *et al.*, 1995). The general assumption that the photosynthetically important CA and MDH in maize and *Sorghum* are the light-inducible and spatially regulated (mesophyll cell-specific) forms implies that transcriptional regulation is an important factor in separating the primary CO_2 fixation from the RPP cycle (Burnell *et al.*, 1997 Luchetta *et al.*, 1990, Metzer *et al.*, 1989). The second MDH gene in maize and *Sorghum* is assumed to play an anapleurotic role. Since this gene is not found in *Flaveria* species, the control of expression of their single MDH gene is presumably different from that of the specialised maize and *Sorghum* genes (Ludwig *et al.*, 1995, Sheen, 1999).

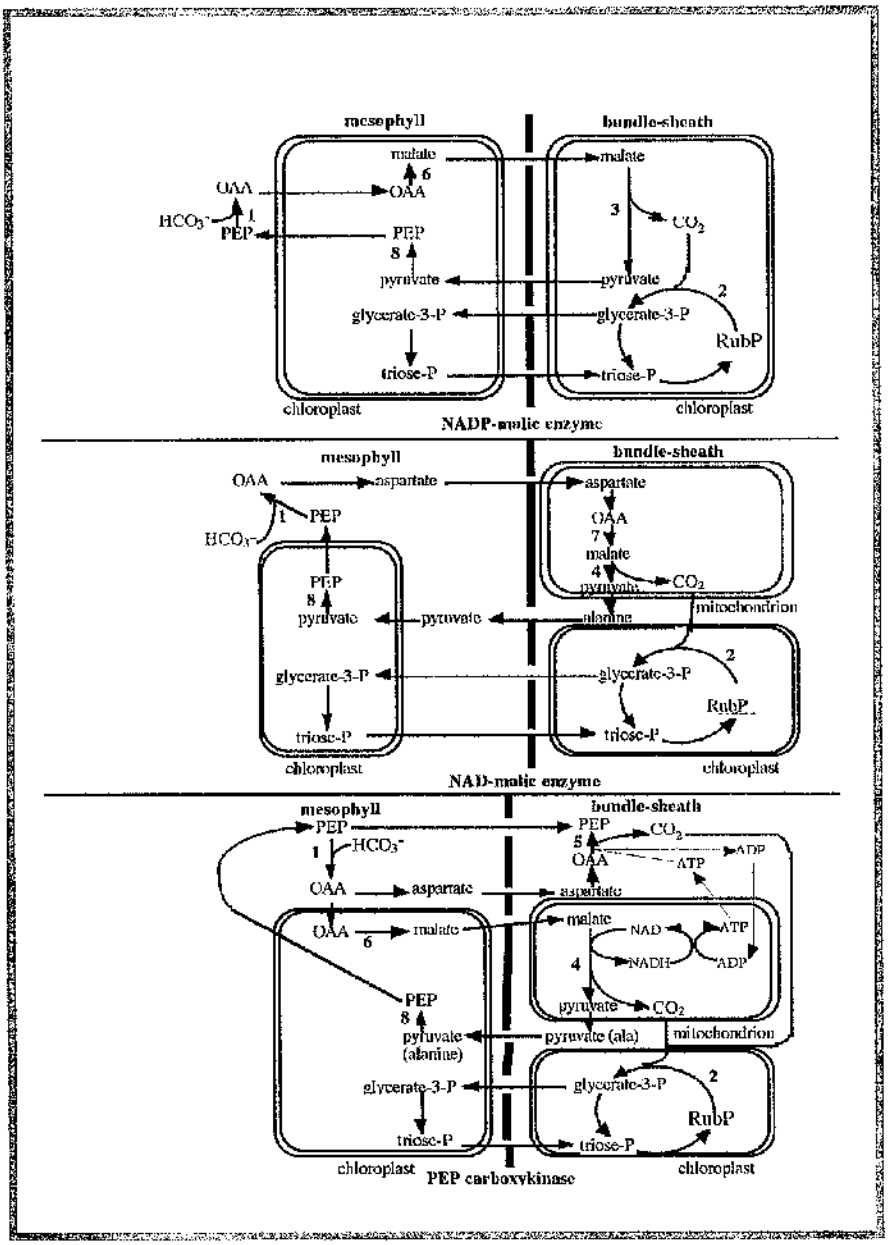
Figure 1.4. The C_4 pathway of photosynthesis

Abbreviations: RubP, ribulose 1,5-bisphosphate; OAA, oxaloacetate; Ala, alanine; PEP, phosphoenolpyruvate.

The diagram gives an overview of the pathway and its intracellular compartmentation for each of the three subgroups of C_4 plants. For most reactions, cofactors, transaminations etc. have been omitted for clarity. In PEP carboxykinase type C_4 plants NAD-malic enzyme also performs malate decarboxylation. The diagram was adapted from Leegood (1997).

The enzymes which catalyse key steps in each subgroup are as follows:

1. phosphoenolpyruvate carboxylase
2. ribulose 1,5-bisphosphate carboxylase/oxygenase
3. NADP-malic enzyme
4. NAD-malic enzyme
5. phosphoenolpyruvate carboxykinase
6. NADP-malate dehydrogenase
7. NAD-malate dehydrogenase
8. pyruvate, phosphate dikinase


mesophyll
bundle-sheath
malate
OAA
HCO3-
PEP
pyruvate
glycerate-3-P
triose-P
CO2
RubP
chloroplast
NADP-malic enzyme
aspartate
alanine
mitochondrion
NAD-malic enzyme
ATP
ADP
NAD
NADH
pyruvate (ala)
PEP carboxykinase

Some C_4 cycle enzymes must be targeted to the chloroplasts. For example, two NADP-ME genes in maize and *Flaveria bidentis* encode chloroplast transit peptide sequences at the N-terminus of the proteins (Marshall *et al.*, 1996, Rothermel *et al.*, 1989). The two genes show different expression patterns, with one showing bundle-sheath specificity and light induction in mature leaves suggesting that it is important in C_4 photosynthesis. Maize PPDK also contains a chloroplast transit peptide (Glackin *et al.*, 1990). Molecular biology evidence suggests that this gene may also encode the non-photosynthetic cytosolic PPDK form by the utilisation of two alternative transcriptional start sites. Transcription from the distal start site allows the translation of the transit peptide. The proximal start site is found in an intron between the PPDK coding sequence and the transit peptide coding sequence. This promoter drives ubiquitous expression of a cytosolic enzyme, rather than leaf-specific, light inducible transcription in the mesophyll cells. Interestingly, the PPDK gene for the C_3 species rice also contains two promoters, but the apparent major difference is the strength of the promoters (Imaizumi *et al.*, 1997). In C_4 plants the PPDK transcript encoding the chloroplast form is high, while that encoding the cytosolic PPDK is low (Glackin *et al.*, 1990). In C_3 rice this difference is reversed. Transcriptional control of PEPc genes will be discussed in section 1.4.

1.2.3.2 Post-translational regulation in the C_4 pathway

A number of C_4 enzymes have been shown to be controlled post-translationally, aiding the co-ordination between CO_2 fixation and the RPP cycle in response to environmental changes (such as light and temperature) and the metabolic flux within the cell. Three features contribute to the fine-tuning of the C_4 enzymes. They include control by metabolites, the thioredoxin system and post-translational modification of key enzymes by phosphorylation.

PPDK, which is localised to the chloroplast stroma, is strongly regulated by light (Leegood, 1997). During catalysis by PPDK a phosphate group becomes transiently attached to the catalytic histidine (position 458). This phosphate group (P) is thought to be very labile and can be lost non-ezymatically, resulting in E-His^{458}/Thr^{456} (Rocske *et al.*, 1988). The regulation of this enzyme is by phosphorylation. Unusually, the phosphoryl donor is ADP. This ADP-dependent phosphorylation inactivates PPDK. The phosphate group becomes attached to a threonine residue (position 456) adjacent to the catalytic histidine (Carrol *et al.*, 1990, Huber *et al.*, 1994). The phosphorlytic cleavage of phosphate from His^{458} forms

pyrophosphate and active E-His/Thr (Burnell, 1984). Addition of the phosphate group to Thr^{456} and the removal of the phosphate group from His^{458} are catalysed by a single bifunctional regulator protein (RP), which posses two distinct active sites (Roeske and Chollet, 1987). It has been proposed that an increase in the concentration of pyruvate, brought about by light dependent transport into the chloroplast, mediates light activation by affecting RP activity (Leegood, 1997). This allows the enzyme-phosphoryl intermediate (E-His^{458}-P) to donate its phosphate group to pyruvate generating PEP. This mechanism is outlined in figure 5.

Several enzymes involved in photosynthesis are regulated by thiol-disulphide interchange. One example relevant to C_4 is NADP-malate dehydrogenase, located in the mesophyll cell chloroplast, which is light activated via thiol-disulphide interchange (Edwards *et al.*, 1985). The oxidation of this enzyme forms two disulphide bridges between four of the six cysteine residues within each subunit and causes loss of activity (Jenkins *et al.*, 1986). Activation of NADP-malate dehydrogenase occurs upon reduction of these two bridges causing four free thiol groups (Kagawa and Bruno, 1988). The redox state of NADP-malate dehydrogenase is controlled by thioredoxin regulation that determines activity of the enzyme (Ashton and Hatch, 1983, Usada, 1988). This allows sensing of light-intensity due to the redox state of ferredoxin, a component of the photosynthetic electron transport chain, in turn controlling the redox state of thioredoxin.

A third enzyme of the C_4 cycle is post-translationally controlled. PEPc is regulated by the phosphorylation of a single serine residue by a protein kinase termed PPCK. This will be discussed further in section 1.4.

1.3 Crassulacean acid metabolism (CAM)

1.3.1. The physiology and biochemistry of CAM

Ranson and Thomas (1960) showed a unique form of photosynthesis in the family Crassulaceae. This form of photosynthesis was therefore termed Crassulacean acid metabolism - CAM. Since then, CAM has been identified in a number of families within both monocotyledon and dicotyledon angiosperms (Dittrich *et al.*, 1973; Winter and Smith, 1996).

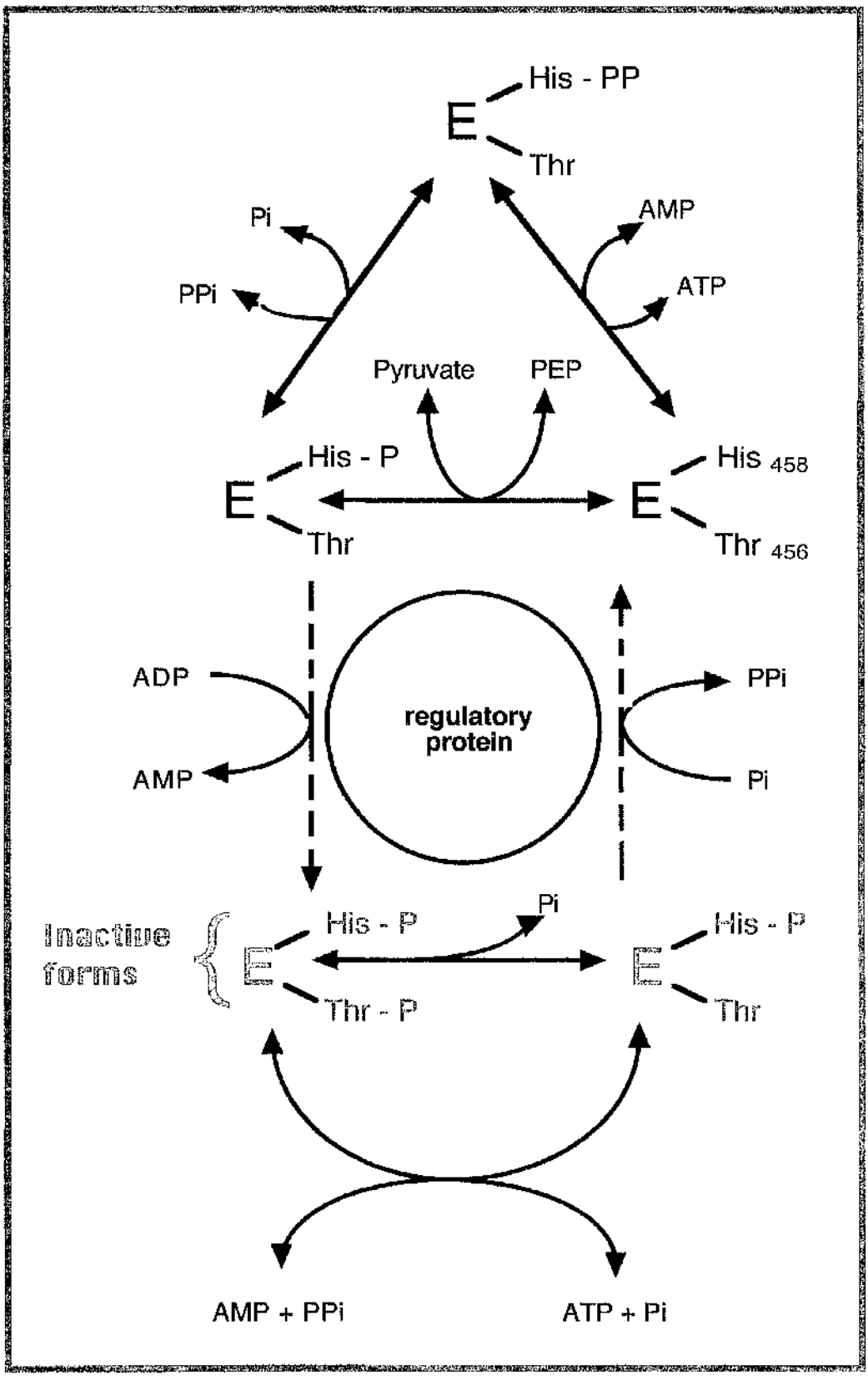
Obligate CAM can be considered in four phases (Osmond, 1978). During the night, in Phase I, the stomatal pores of the leaf are open. CO_2 enters freely and is converted to bicarbonate by CA. HCO_3^- is fixed by PEPc forming OAA. Malate is produced by MDH and then pumped into the vacuole, where it accumulates as malic

Figure 1.5. The regulation of PPDK

Abbreviations: E, PPDK enzyme; PEP, phosphoenolpyruvate

This diagram represents the three-step catalytic mechanism (bold arrows) and the ADP-dependent phosphorylation (inactivation) of Thr^{456} catalysed by the bifunctional regulatory protein (dotted arrows).

Adapted from Leegood, 1997.


His - PP
E
Thr
Pi
PPi
AMP
ATP
Pyruvate
PEP
His - P
E
Thr
His 458
E
Thr 456
ADP
AMP
regulatory protein
PPi
Pi
Inactive forms
His - P
Pi
E
Thr - P
His - P
E
Thr
AMP + PPi
ATP + Pi

acid. Towards the end of Phase I, primary CO_2 fixation by PEPc slows. This is caused by a decrease in PEPc activity thought to be due to the increased sensitivity of PEPc to malate inhibition (Nimmo *et al.*, 1984; Winter, 1980; Winter, 1982). Phase II is a further short burst of CO_2 fixation by both Rubisco and PEPc starting at dawn. Malic acid levels peak during Phase II as decarboxylation starts. During Phase III, the stomatal pores close due to increased intracellular partial pressure of CO_2, hence gas exchange falls to zero. Malate released from the vacuole is decarboxylated and the released CO_2 is fixed in the RPP cycle behind closed stomata. Phase IV is a period of CO_2 fixation by Rubisco late in the light period after decarboxylation has ceased and the stomata have reopened. This pathway prevents photorespiration and greatly increases water efficiency. An overview of CAM is shown in figure 1.6.

This strategy of photosynthesis is a selective advantage in hot and arid environments. Some plants perform CAM constitutively, in others it is induced by drought or salt stress (Osmond, 1984).

1.3.2 Regulation of the CAM pathway

CAM photosynthesis exhibits a range of control inputs, allowing a great degree of metabolic plasticity. These include environmental, developmental, tissue-specificity, hormonal and circadian control pathways (Dodd et al., 2002, Cushman and Bohnert, 1999).

CAM genes (those specifically induced in the CAM pathway) show specific spatial and temporal induction as defined by their cell type and night-dark expression pattern, respectively. Post-translational regulation via phosphorylation of PEPc was discovered in this system, with the dephosphorylated 'day' form more sensitive to malate inhibition than the phosphorylated 'night' form. This observation was made in a study of circadian rhythmicity in *Kalanchoe (Bryophyllum) fedtschenkoi* (Nimmo *et al.*, 1984). Detached leaves of CAM species, kept in constant conditions, exhibit circadian rhythms of CO_2 metabolism due to circadian changes in the flux through PEPc (see Wilkins, 1992 for a review). Following the observation of a day/night difference in the malate sensitivity of PEPc (Winter, 1981) Nimmo *et al.* (1984, 1986, 1987) showed that this was caused by reversible phosphorylation and contributed to the circadian rhythms of CO_2 metabolism. Control of CAM PPCK will be discussed further in 1.6.

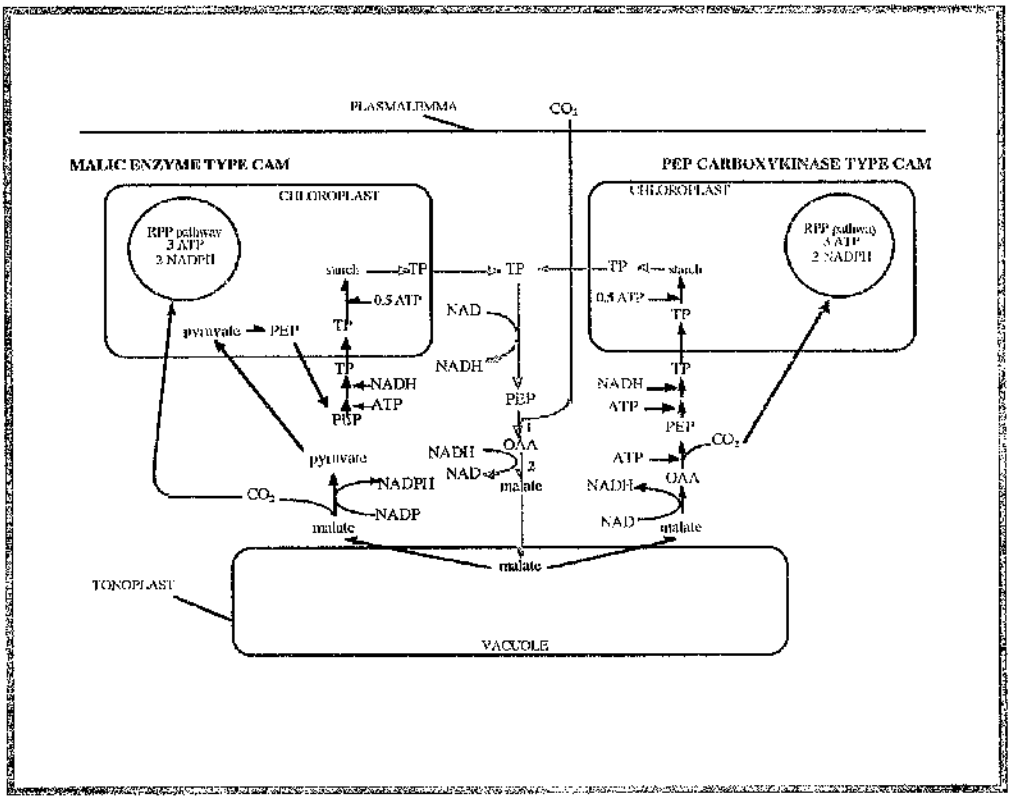
Figure 1.6. Crassulacean acid metabolism (CAM)

The arrows in blue represent steps which occur during the dark period, whilst the arrows in black and red represent steps which occur during periods of illumination in malic enzyme type CAM and PEP carboxykinase type CAM repectively.

Abbreviations: RPP, reductive pentose phosphate pathway; PEP; phosphoenolpyruvate; OAA, oxaloacetate; TP, triose phosphate.

The key enzymes of the CAM pathway are as follows:

1. carbonic anhydrase
2. phosphoenolpyruvate carboxylase
3. malate dehydrogenase
4. phosphoenolpyruvate carboxykinase
5. NADP-malic enzyme (cytosol) and/or NAD-malic enzyme (mitochondria)
6. pyruvate, phosphate dikinase


PLASMALEMMA
CO_2
MALIC ENZYME TYPE CAM
PEP CARBOXYKINASE TYPE CAM
CHLOROPLAST
CHLOROPLAST
RPP pathway
3 ATP
2 NADPH
RPP pathway
3 ATP
2 NADPH
starch
TP
TP
TP
starch
0.5 ATP
0.5 ATP
TP
TP
pyruvate
PEP
NAD
NADH
TP
NADH
ATP
PEP
NADH
ATP
PEP
CO_2
pyruvate
NADH
NAD
OAA
2
malate
ATP
NADH
OAA
NAD
malate
NADPH
CO_2
NADP
malate
malate
TONOPLAST
VACUOLE

1.4 Phosphoenolpyruvate carboxylase in higher plants

1.4.1 Introduction to the catalytic and regulatory properties of phosphoenolpyruvate carboxylase

The photosynthetic role of PEPc in primary CO_2 assimilation in C_4 and CAM pathways is generally well known. This role is carried out by specific photosynthetic isoforms. As a ubiquitous enzyme in higher plants, PEPc is also involved in a number of housekeeping functions. These functions tend to be referred to as C_3-type roles, and are carried out by 'C_3' isoforms of PEPc, differentiating them from the photosynthetic isoforms. The 'C_3' isoforms are known to have specific roles in C_3 tissue and non-photosynthetic tissue in CAM and C_4. It is therefore not surprising that small multi-gene families of PEPc genes exist in each plant species investigated.

1.4.2 Transcriptional regulation of PEPc

Distinct PEPc isoforms are expressed in photosynthetic and non-photosynthetic tissues. In C_3 plants PEPc gene families have been shown to exist in *Brassica napus*, tobacco, tomato, alfalfa, rice, *Flaveria pringlei*, wheat and *Arabidopsis thaliana*. The suggestion that these genes may encode specific housekeeping, root and ripening isoforms is reinforced by studies of C_4 plants. In the C_4 plant *Sorghum*, three PEPc genes have been characterised (Lepiniec *et al.*, 1993). Two C_3 type genes encode a housekeeping and root specific isoform, while the third is a C_4 type photosynthetic form that is expressed in greening, mediated by a phytochrome response (Toh *et al.*1995). Maize contains three PEPc genes which exhibit different expression patterns (Dong *et al.*, 1998, Hudsperth *et el.*, 1989, Kawamura *et al.*, 1992). These patterns suggest that they are C_4 photosynthetic, root and anapleurotic specific isoforms. Interestingly the PEPc genes of the C_3 plant *F. pringlei* are very similar to those of the C_4 plant, *F. trinervia*. However, in the *F. pringlei* the PEPc gene most closely related to the C_4-type gene of *F. trinervia* is only weakly expressed and is not specifically accumulated within the leaf tissue (Ernst and Westhoff, 1997). In CAM plants, the facultative CAM plant *M. crystallinum* is thought to have two or three PEPc genes, with one showing considerable enhancement of expression at C_3 to CAM switching.

Promoter elements are thought to play a role in PEPc gene expression. One piece of evidence for this is the presence of CpG islands at sites within C_4 and C_3 PEPc genes, potentially causing changes in DNA methylation states and thus enhancement or repression (Schaffner *et al.*, 1992). Motifs for the light responsive

element AT-1, nod- and G-boxes have also been identified within promoter sequences of *Sorghum* PEPc genes (Lepiniec *et al.*, 1994).

Although no clear consensus has emerged linking regulatory mechanisms to transcriptional control of the different PEPc genes in plants, multi-gene families with isogenes exhibiting differing expression patterns could indicate that transcriptional control of PEPc is as important as the better studied post translational control of PEPc.

1.4.3 Post translational regulation of PEPc

PEPc is potently controlled by metabolic effectors. Glucose-6-phosphate (G-6-P) is an allosteric activator of PEPc (O'Leary, 1982). The extent of the activation of PEPc by G-6-P is determined by PEPc's phosphorylation state. Malate, produced from the rapid reduction of OAA by NADP-MDH, is a negative feedback effector of PEPc, and shows complex, often sigmoidal, kinetics. The potency of this inhibitor is dependent upon a number of factors. These include whether PEPc has undergone proteolysis of the N-terminal, the assay pH and the phosphorylation status of the enzyme (see section 1.5) (Auschus and O'Leary, 1992; Duff *et al.*, 1995; Jiao and Chollet, 1991; McNaughton *et al.*, 1991; Wang *et al.*, 1992). The effect of assay pH is illustrated by the K_i for malate of the recombinant *Sorghum* C_4-type PEPc. At pH 7.3, the K_i was 25 times less than at pH 8.0. This indicates that cytosolic pH could contribute to PEPc regulation *in vivo*, in the absence of changes to cytosolic malate (Leegood *et al.*, 1990).

The effect of phosphorylation of PEPc on inhibition by malate is illustrated again by experiments utilising recombinant *Sorghum* PEPc. At pH 7.3, the K_a (G-6-P) of the PEPc at 1 mM PEP was 1.3 mM and 0.28 mM for dephospho- and phospho-forms of PEPc, while the K_i (L-malate) of the PEPc was 0.17 mM and 1.2 mM for the dephospho- and phospho- forms of PEPc (Duff *et al.*, 1995). This will be expanded upon in section 1.6.

Other effectors have been shown to affect PEPc *in vitro*, and these may contribute to the fine-tuning of its activity. They included organic acids, flavonoids and shikimic acid (Colombo *et al.*, 1996; Paroba *et al.*, 1996). Colombo *et al.* (1996) suggested that this may allow PEP to be diverted to the shikimate pathway and subsequent flavonoid biosynthesis under conditions causing excess glycolysis. However the significance of these observations *in vivo* is unknown.

The effect of phosphorylation upon PEPc will be discussed in section 1.5.

1.4.4 The structure and catalytic mechanism of phosphoenolpyruvate carboxylase

The enzyme phosphoenolpyruvate carboxylase is ubiquitous in higher plants and is also found in bacteria, cyanobacteria and green algae (Toh *et al.*, 1994; Lepiniec *et al.*, 1994). PEPc catalyses the β-carboxylation of PEP, using HCO_3^-. This releases OAA and P_i (Andreo *et al.*, 1987). The enzyme requires a divalent cation as a cofactor. *In vitro*, the most effective cofactor is Mg^{2+}, with Mn^{2+} and Co^{2+} also showing activity, although Ca^{2+} is inhibitory (Andreo *et al.*, 1997).

The mechanism of PEPc catalysis is relatively well understood. The ligands bind in a preferred order, the cation first, followed by the substrates PEP and HCO_3^- (Chollet *et al.*, 1996). The phosphate group is transferred to form carboxyphosphate and the enolate of pyruvate (Walsh, 1979). Stereochemical studies suggest that the carbonyl carbon in the carboxyphosphate intermediate immediately after transfer is positioned too far away from carbon-3 of the enolate to interact (Ausenhus and O'Leary, 1992). Therefore a conformational change in the enzyme is required to bring the two carbons close enough to interact. The suggested mechanism for C-C bond formation is that a base in the active site of PEPc attacks the carboxyl group of the carboxyphosphate intermediate, deprotonating the carboxyl group. The carboxyphosphate now decomposes to form enzyme bound P_i and CO_2 in a reversible fashion (Ausenhus and O'Leary, 1992). The final reaction step is that enzyme bound CO_2 combines with the metal-stabilised enolate. This part of the reaction is thought to be irreversible (Chollet *et al.*, 1996). In the final step OAA and Pi are released. An outline of this mechanism is shown in figure 1.7.

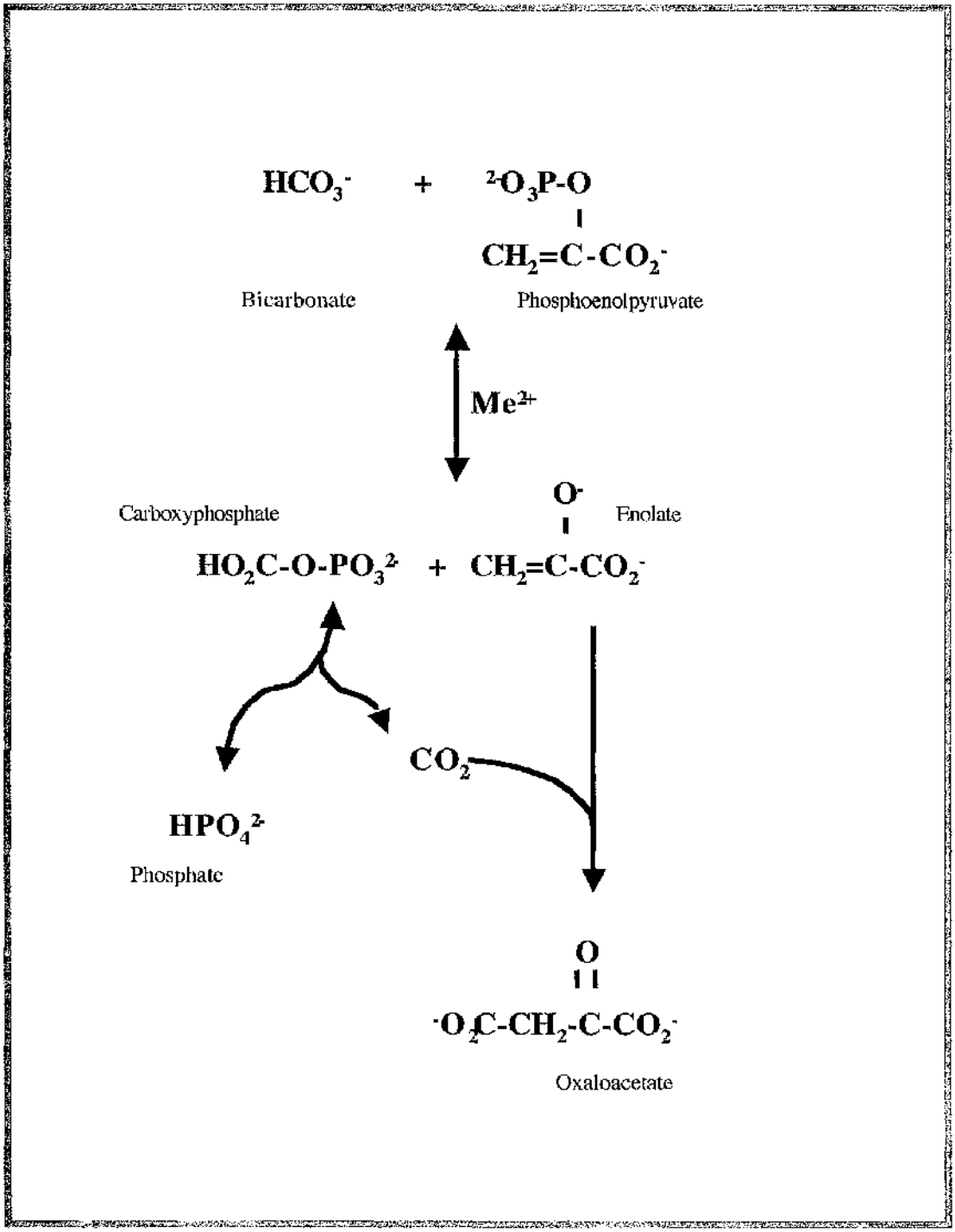
Early indicators about residues involved in catalysis by PEPc came from chemical modification and site-directed mutagenesis allied to sequence comparisons. These implicated lysine^{773}, histidine^{138}, histidine^{579} and arginine^{587} of *E. coli* PEPc. The first structures for PEPc became available in 1999 when Matsumura *et al.* and Kai *et al.* solved the structures of *E. coli* bound to Mn^{2+} (Matsumura *et al.*, 1999) and an enzyme-aspartate complex (Kai *et al.*, 1999). PEPc forms a homo-tetramer structure in a dimer of dimers formation. The allosteric inhibitor for *E. coli* PEPc, L-aspartate, was shown to interact with four aminoacid residues of *E. coli* PEPc. These are the residues lysine^{773}, asparagine^{881}, arginine^{587} and arginine^{832}. The lysine and the two arginines bind to L-aspartate via salt bridges to the carboxyl group. Arginine^{587} is

Figure 1.7 The proposed reaction mechanism of phosphoenolpyruvate carboxylase

Mechanism of the biotin-independent carboxylation of PEP by PEPc

Abbreviation: Me^{2+}, divalent cation.

This figure was adapted from Vidal and Chollet (1997)


HCO3-
+
2-O3P-O
CH2=C-CO2-
Bicarbonate
Phosphoenolpyruvate
Me2+
Carboxyphosphate
O-
Enolate
HO2C-O-PO32-
+
CH2=C-CO2-
CO2
HPO42-
Phosphate
O
-O2C-CH2-C-CO2-
Oxaloacetate

in a glycine-rich loop adjacent to histidine[579], and both of these residues are thought to be involved in catalysis. Hence the current view is that this region of the active site is flexible. The form of the enzyme in which the inhibitor L-aspartate is bound to several residues including arginine[587] is inactive. It is thought that in the active form of the enzyme, the flexibility of the active site region allows substrates to bind arginine[581], arginine[587] and other residues and for a 'lid' to close over the active site to prevent reaction of intermediates with water. However no structures with substrate analogues bound have yet been reported.

Alignments of plant PEPc sequences reveal a conserved consensus motif within the N-terminal region. This motif, E/D R/K X X **S** I D A Q L/M R indicates the phosphorylatable serine which is indicated in bold (Vidal and Chollet, 1997). The presence of this phosphorylation motif in the deduced amino acid sequences of all the higher plant PEPc indicates that they are all capable of undergoing phosphorylation on this N-terminal serine residue with the consequences for activity that have already been noted (discussed further in section 1.5 - 1.6).

1.5 Roles and regulation of phosphoenolpyruvate carboxylase in higher plants

1.5.1 Regulation of phosphoenolpyruvate carboxylase in CAM plants

The first demonstration that PEPc is regulated via phosphorylation came from a study of the CAM plant *Kalanchoë fedtschenkoi* (Nimmo *et al.*, 1986), which showed that PEPc was tenfold more sensitive to malate inhibition during the light period than in the middle of the dark period. This correlated with the presence of phosphorus in the enzyme at night but not in the day. Subsequent work examining purified PEPc showed that the' night' form was phosphorylated but the 'day' was not (Nimmo *et al.* 1986). This is now known to be due to the phosphorylation of a single serine residue near the N-terminus of the protein (Chollet, Vidal *et al.*, 1997). In the early 1980s several groups found that light reduced the malate sensitivity of PEPc in C_4 species (see references cited in Nimmo *et al.*, 1987b). Conversions between the two forms of the enzyme are also observed in detached leaves kept in constant environmental conditions (Nimmo *et al.*, 1987), confirming that they are controlled by a circadian rhythm rather than by light-dark transitions. (Nimmo *et al.*, 1984). Phosphorylation of PEPc has also been shown in several other CAM species (Brulfert *et al.*, 1986; Kluge *et al.*, 1988; Baur *et al.*, 1992; Weigend, 1994). In a detailed study in which *Kalanchoë daigremontiana* (a constitutive CAM plant) and *Clusia minor* (C_3-CAM intermediate) adapted to an 11 h photoperiod were assessed by gas

exchange and instantaneous isotope discrimination analysis (Borland and Griffiths, 1997). This allowed the *in vivo* flux through PEPc to be estimated; in both species, there was a good correlation between the estimated flux and the phosphorylation state of the enzyme, as judged by its malate sensitivity. The control of phosphorylation of PEPc will be discussed in further depth in section 1.6.

1.5.2 Regulation of phosphoenolpyruvate carboxylase in C_4 plants

Initial work by Budde and Chollet (1986) showed that maize leaf PEPc could be phosphorylated *in vitro* when ATP is supplied. The following year, it was found that illumination caused phosphorylation of maize leaf PEPc and that this was accompanied by increases in the K_i of PEPc for L-malate during (Nimmo *et al.*, 1987b). In these experiments the K_i of PEPc for L-malate was assayed in rapidly desalted extracts of illuminated and darkened leaves. The extract from illuminated tissue was shown to have a K_i of PEPc for L-malate of 2-3 times higher than the extract from the darkened tissue. *In vivo* labelling experiments revealed that PEPc from the illuminated leaf extract was phosphorylated with a K_i for L-malate of 1.2 mM and in darkened extracts was dephosphorylated with a Ki for L-malate of 0.5 mM. These results were confirmed by Jiao and Chollet (1989), who showed the association of the K_i changes of *in vitro* maize leaf PEPc to a seryl-phosphorylation, caused by a soluble protein kinase

The sequence of a $^{32}P_i$-labelled regulatory site phosphopeptide purified from a tryptic digest of the *in vitro* phosphorylated, purified dark-form maize leaf PEPc was determined (Jiao and Chollet, 1990; Terada *et al.*, 1990). The sequence of the phosphopeptide was His-His-Ser (P)-Ile-Asp-Ala-Gln-Leu-Arg. Both mammalian cyclic AMP-dependent protein kinase and a protein kinase from illuminated maize leaves phosphorylated this site, although it should be noted that PEPc is a very poor substrate for mammalian PKA (Jiao and Chollet, 1990; Terada *et al.*, 1990). This sequence corresponds exactly to residues 13-21 of the deduced amino acid sequence of maize leaf PEPc (Jiao and Chollet, 1990; Terada *et al.*, 1990). Thus it is Ser^{15} in the maize PEPc sequence that is the target for phosphorylation. This *in vitro* phosphorylation site was confirmed to be identical to the *in vivo* phosphorylation site by performing automated Edman sequencing on *in vivo* phosphorylated PEPc from both maize and *Sorghum* (Jiao *et al.*, 1991b). Only a single N-terminal serine residue (Ser^{15} for maize and Ser^{8} for *Sorghum*) was phosphorylated *in vivo* (Jiao *et al.*, 1991b). Examination of the deduced N-terminal sequences of other PEPcs,

encompassing C_4 and CAM species indicated a conserved N-terminal motif, Lys/Arg-X-X-Ser, which corresponds to this phosphorylation site (Jiao and Chollet, 1990). Using specific antibodies to phosphorylated maize PEPc, Ueno *et al.*, (2000) later showed that there is a light/dark reversible phosphorylation and that there seems to be no other sites of phosphorylation other than this Ser close to the N terminus, hereafter termed for convenience as the N-terminal serine. The phosphorylation of the N-terminal serine of C_4 PEPc in the light was shown to be catalysed by a soluble protein kinase (Echevarria *et al.*, 1990, Jiao and Chollet, 1989, McNaughton *et al.*, 1991). This will be further discussed in section 1.6.

Three PEPc isoforms have been reported in maize and *Sorghum*, both C_4 plants (Yannagisawa *et al.*, 1988; Kawamura *et al.*, 1992; Cretin *et al.*, 1991; Lepiniec *et al.*, 1993; Dong *et al.*, 1998). The first is the C_4-specific isoform involved in the primary fixation of carbon, the second appears to be a housekeeping isoform and the third is a root expressed-form. Investigations into non-C_4 roles of PEPc in C_4 plants reflect the research conducted on PEPc in C_3 plants. Work conducted upon the *Sorghum* PEPc within seeds, revealed two PEPc proteins of 108 and 110 kDa that both accumulate and are phosphorylated in germinating seeds (Nhiri *et al.*, 2000). The work by this group also revealed that these isoforms of PEPc had a higher affinity for their substrate, PEP, and sensitivity for the inhibitor, L-malate, in phospho- and non-phospho forms than the photosynthetic PEPc form of the mesophyll. Thus the regulatory phosphorylation of plant PEPc is universal and will be discussed further in section 1.6.

1.5.3 Regulation of phosphoenolpyruvate carboxylase in C_3 plants

The C_3 isoforms of PEPc have been shown to have specific roles within C_3 tissues. These include the reported roles within roots, root nodules, guard cells, fruit ripening, seeds and leaves (Deroche and Carrayol, 1989; Law and Plaxton, 1995; Melzer and O'Leary, 1987; Osuna *et al.*, 1996; Pacquit *et al.*, 1993; Sangwan *et al.*, 1992; Zhang *et al.*, 1994). The major function of C_3 PEPc is to resupply a depleted TCA cycle with four-carbon skeletons as the TCA cycle intermediates are re-directed for use in aminoacid biosynthesis.

Response of C_3 PEPc to nitrogen

The *in vivo* phosphorylation of C_3 PEPc was demonstrated in 1991 by Van Quy *et al.* Wheat leaves, grown in a NO_3^- depleted environment and then detached, were supplied with $^{32}P_i$ during the dark and light period. The labelling of PEPc was

shown to be 4.16 fold higher in the illuminated sample compared to the control dark sample. Labelling was further increased when the detached leaves were from an illuminated, high NO_3^- environment (6.65 fold higher than the low NO_3^-, dark treated leaf control). The labelling changes were accompanied by changes in the L-malate sensitivity to PEPc, with $I_{0.5\,(malate)}$ values of 1.45 mM (dark control), 1.75 mM (illuminated, low NO_3^-) and 3.4 mM (illuminated, high NO_3^-). Duff and Chollet (1995) clarified these results by conducting *in vitro* PPCK assays that showed the phosphorylation of PEPc was Ca^{2+}-independent and light inducible. They further showed that the kinase was specific to the serine found in the N-terminus of PEPc because it failed to phosphorylate a *Sorghum* PEPc phosphorylation site mutant (serine to asparagine substitution). This was preliminary evidence that 'anapleurotic' PEPc may be controlled via post-translational modification in conditions where the rate of aminoacid biosynthesis is high.

The signal transduction causing phosphorylation of wheat leaf PEPc in high NO_3^- and light was further investigated by inhibitor studies. Phosphorylation was shown to be completely dependent upon an active photosynthetic electron transport chain from H_20 to NADP and to a lesser extent *de novo* protein synthesis as shown by addition of methyl viologen and cycloheximide respectively (Duff and Chollet, 1995).

Further dissection of this pathway was achieved by Manh *et al.* (1993), who fed detached, nitrogen-depleted illuminated wheat leaves with a number of aminoacids (glutamine, glutamate and aspartate), and nitrogen sources (NO_3^- and NH_4^+). Only glutamate decreased PEPc activity while the others increased the activity of PEPc. Similar experiments were also conducted with inhibitors of nitrate reductase (NR)(sodium tungstate), glutamine synthase (GS)(methionine sulphoximine), glutamate synthase (GOGAT)(azaserine) and amino transferases (amino oxyacetate). Only treatment with azaserine did not block the effect of NO_3^-. Together these results suggest that in nitrogen-depleted C_3 leaves, the light activation of PEPc is due to an increase in the *in vivo* glutamine concentration, and is not directly due to NO_3^-. Similar results were reported by Diaz *et al.* (1996) with nitrogen-depleted barley leaves. Manh *et al.* (1993) also suggested, from *in vitro* assays on partially purified PPCK, that glutamine activation and glutamate inhibition of PEPc phosphorylation was due to effects on PPCK activity. However, Duff and Chollet (1995) repeated this experiment but assayed PPCK in the presence of the calcium chelater EGTA. They

therefore detected only the Ca^{2+}-independent PPCK. They showed no direct effect of glutamine in the *in vitro* activity of the Ca^{2+}-independent PPCK.

Activation of C_3 PEPc by illumination

The response of PEPc to light in C_3 plants has been investigated in several plants. Smith *et al.* (1996) used barley protoplasts to characterise the properties of PEPc and PPCK activity in response to light. They showed that the specific activity of PEPc in the protoplasts remained constant in light/dark changes, but illumination caused a loss of sensitivity to inhibition by L-malate. This suggested that PPCK activity was up regulated on illumination in C_3 plants. This light induction of C_3 PPCK was confirmed in experiments which showed that barley protoplasts illuminated for 1-2 h contained higher PPCK activity then those that had been incubated in the dark. Inhibitor studies suggested that this light induction does not require photosynthesis but does require protein synthesis. They also showed that prolonged cycloheximide treatment only removes about 50 % of PPCK activity in the protoplasts. The intracellular Ca^{2+} of protoplasts was reduced by incubation with EGTA and the calcium ionophore-A23187. These experiments increased the tolerance of PEPc to the inhibitor L-malate in the light, but appeared to have no effect on PPCK activity. Together these experiments suggested that two types of PPCK are present in barley protoplasts, one that turned over rapidly and is responsible for the light enhancement, and a second a that is slowly turned over and is present in the dark.

Li *et al.* (1996) investigated the properties of the tobacco leaf PEPc, another C_3 plant. PEPc was phosphorylated only in the light phase. As in barley protoplasts, this light induction was inhibited by cycloheximide, but unlike them, it required photosynthesis. Li *et al.* also provided evidence that agrees with Manh *et al.* (1993) (see above) showing that methionine sulphoximine inhibits the light induction of PEPc phosphorylation, with a partial, but specific reversal of this inhibition by feeding with glutamine.

A number of other C_3 plants have also been investigated, these results will be discussed in section 1.6.

Specialised roles of C_3 PEPc

C_3 PEPc in guard cells

Guard cells provide the main control of gas exchange. These cells are known to respond to a number of factors including light, humidity, CO_2 concentration, temperature and phytochromes. The stomatal pore, a pair of guard cells, widens as a

plasma membrane H^+-ATPase hyperpolarizes the plasma membrane. This allows an uptake of potassium from the apoplast, which increases the osmotic potential of the cell. To balance this potassium (K^+) electrically, the guard cell accumulates anions, mainly $malate^{2-}$ (Cottele *et al.*, 1999). This accumulation of anions and potassium causes an influx of water leading to a turgor rise, swelling of the cells and stomatal opening.

The cytosolic synthesis of malate is catalysed by PEPc and MDH, and PEPc is potentially an important enzyme in stomatal control. In *Vicia faba*, two unique PEPc isoforms have been identified within guard cells in comparison to mesophyll cells (Schulz *et al.*, 1992). Recently, a guard cell PEPc cDNA was cloned from potato (Muller-Rober *et al.*, 1998). This isoform contains the highly conserved serine residue within the N-terminal region, and could therefore be phosphorylated by PPCK allowing activation of PEPc during stomatal opening. Interestingly, this PEPc cDNA hybridised to two different mRNAs, which were derived from the same gene. The smaller transcript was found exclusively in guard cells, while the larger was found in other tissues. Promoter-GUS fusions of this gene showed staining exclusively to guard cells, thus raising questions of what regulatory elements were absent in the promoter-GUS fusion system to prevent the ubiquitous larger transcript.

There is now both direct and indirect evidence showing that guard cell PEPc can be phosphorylated. Biochemical analysis of guard cell PEPc showed it was inhibited by 400 μM malate that during the dark but not in the light, this correlated with its phosphorylation status (Zhang *et al.*, 1994; Du *et al.*, 1997). Meinhard and Schnabl (2001) also presented data that guard cell PEPc is regulated by reversible phosphorylation. Illumination in the presence of fusicoccin caused a decrease in malate sensitivity, which was reversed by abscisic acid (causing stomatal closure). Increasing evidence that fusicoccin stabilises a 14-3-3 protein interacting with plasma membrane H^+-ATPase suggest that fusicoccin directly modulates the proton pump, combined with the increased specific activity of PEPc at suboptimal pH's (7.0-7.3) imply a role for $[H^+]$ in the up regulation of the regulatory kinase (Muller-Rober *et al.*, 1998).

C_3 PEPc in root nodules

The most characterised symbiotic N_2-fixing system is the symbiosis between legumes and rhizobia. The rhizobia induce the formation of nodules on the roots of

their host-plants. The nodule formation occurs as the rhizobia enter the plant cell and become enclosed by a plasmalemma-derived membrane, the peribacteroid membrane, and form an organelle-like structure, the symbiosome (Vance, 1983).

During symbiosis the bacteroids fix atmospheric N_2 to ammonia, which is then assimilated by the plant via the GS/GOGAT pathway. In exchange the plant provides the bacteroid with reduced carbon in the form of malate which is used as fuel for bacteroid metabolism, supporting the high energy demands of nitrogen fixation. To assist this the plant produces PEPc in nodule-specific or nodule-abundant forms (Michelis and Vanderleyden, 1994).

Nodule PEPc has been suggested to carry out several (mainly anapleurotic) functions. These include the provision of dicarboxylic acids (malate, succinate) for metabolism by the bacteroids in support of nitrogenase activity; the synthesis of carbon skeletons for the assimilation of NH_4^+ into amino acids (Christeller, Laing *et al.*, 1977; Coker and Schubert, 1981) and the synthesis of organic acids for export in the xylem to counterbalance cation transport and regulate pH (Israel and Jackson, 1982).

Four PEPc isoforms have been identified in soybean (Sugimoto, Kawasaki *et al.*, 1992; Vazquez-Tello, Whittier *et al.*, 1993; Hata, Izui *et al.*, 1998) and at least three are expressed in soybean root nodules (Hata, Izui *et al.*, 1998). *In situ* hybridisation with the nodule-enhanced isoform supported the results of immunogold and immunochemical studies, showing that transcripts are present in all cell types in nodules, including both infected and uninfected cells (Vidal, Nguyen *et al.*, 1986; Robinson, Pathirana *et al.*, 1996; Hata, Izui *et al.*, 1998).

Since nitrogen assimilation is dependent on a recent supply of photosynthate (sucrose) from the host plant's source leaves (Vance and Heichel, 1991), activation of PEPc by hexose phosphates could coordinate sucrose availability with synthesis of dicarboxylic acid to the bacteroids (Schuller, Turpin *et al.* 1990). G 6-P affects nodule PEPc activity in a number of ways. These include the broadening of the pH optimum, therefore limiting the effect of small fluctuations in pH on PEPc activity *in vivo*; stimulating PEPc activity at low (physiological) PEP concentrations by increasing the enzyme's affinity for PEP and protecting the enzyme against inhibition by malate (Woo and Xu, 1996).

Schuller and Werner (1993) demonstrated that soybean nodule PEPc is phosphorylated *in vitro* by an endogenous protein kinase present in crude nodule

extracts. They reported that the phosphorylated form of nodule PEPc was more resistant to malate inhibition, recording I_{50} (malate) values of 0.35 and 1.24 mM before and after phosphorylation, and that phosphorylation increased its specific activity when assayed at sub-optimal conditions (pH 7, 1.25 mM PEP). Zhang *et al.* (1995) were able to show phosphorylation of PEPc *in planta* and investigated the effects of stem girdling and/or darkness on the apparent phosphorylation status of PEPc as indicated by its malate sensitivity. They found that stem-girdling reduced the phosphorylation state of PEPc, as did extended periods of darkness. Subsequent illumination reversed the effect of the dark pre-treatment but this reversal was prevented by concomitant stem girdling. This suggested that the phosphorylation state of soybean nodule PEPc is modulated by photosynthate transported from the shoots, due to either upregulation of PPCK by sucrose (Zhang, Li *et al.* 1995) or due to reduced ATP levels in the cytosol in the absence of sucrose (Wadham, Winter *et al.*, 1996).

C_3 PEPc in ripening

In fleshy fruits, the predominant organic acids are malic and citric acids (Etienne *et al.*, 2002). The accumulation of these acids occurs during the process of fruit expansion. This process involves several pathways, with PEPc implicated in both citric and malic acid synthesis within the cytosol (Chollet *et al.*, 1996; Monselise *et al.*, 1986). Fleshy fruits accumulate organic acids during the early stages of fruit development. In the model fruit, tomato, the onset of accumulation is begun at the end of the cell division phase, with the organic acid levels peaking at the end of cell expansion. As the fruit continues to ripen, levels of the acids gradually decrease (Varga and Bruinsma, 1986). The range of accumulation of organic acids differs from species to species as shown by a comparison of concentrations of organic acids at maturity of tomato (75 mM) and *Citrus* (250 mM).

Biochemical studies showed that in tomato, PEPc activity rose and then fell in parallel with organic acid levels. Lavel-Martin *et al.* (1977) showed that PEPc activity declines before the onset of ripening. In other fruit though, as in banana, the sudden decline is not seen, but a stable level of PEPc activity was observed (Paquit *et al.*, 1993).

PEPc is also thought to play a role in other aspects of fruit development. The onset of ripening is accompanied by a respiratory-crisis that occurs in all climacteric fruit. This leads to great losses of CO_2. PEPc would be an ideal candidate to re-fix this

lost CO_2 into malate (Latzko and Kelly, 1983; Blanke and Lenz, 1989; Chollet *et al.* 1996). However, there is no good experimental evidence for this role of PEPc. A more recent hypothesis concerns a putative role of PEPc in metabolising recently delivered assimilates to the developing fruit. Famiani *et al.* (2000) showed that in grapes, PEPc was closely associated to the vasculature of the fruit, via immunohistochemical studies.

Recently, Guillet *et al.* (2001), have reported the cloning and analysed the expression of two PEPc's from tomato (*LePPc1* and *LePPc2*). RT-PCR experiments showed that *LePPc1* is expressed in a low and constant fashion across tomato tissues, with the exception of a transient increase during early fruit development. They suggested that LePPc1 is anapleurotic in nature. Conversely, the expression of LePPc2 was specifically and strongly up-regulated during the fruit expansion phase before the breaker phase. LePPc2 transcripts were detected also at a far lower levels in stems and cotyledons.

Localisation studies of LePPc2 transcripts in fruit in the early ripening phase showed strong signals in the vacuolated cells in the pericarp, enlarging cells in the gel, near the vicinity of the seeds and in the periphery of the vascular bundles. This supports the localisation of grape PEPc noted by Famiani *et al.* (2000). In late ripening, *LePPc2* expression drops considerably. Guillet *et al.* (2001) suggested that LePPc2 is probably involved in malate accumulation, balancing potassium uptake and aiding cellular expansion.

C_3 PEPc in developing seeds

The role of PEPc in developing seeds is thought to concern the biosynthesis of amino acids for the incorporation into storage proteins. The evidence for this is the correlation between wheat seed protein accumulation and wheat PEPc concentration, as detected by immunolocalization (Arraus *et al.*, 1993). Similar results were obtained in soybean and *Vicia faba* (Sugimato *et al.*, 1989; Golombek *et al.*, 1999).

Two PEPc genes have been identified in *Vicia faba*, *VfPEPc1* and *VfPEPc2*, (Golombek *et al.*, 1999). The expression patterns of the genes were strikingly different. *VfPEPc1* is thought to represent a "sink-specific" form with predominant expression in roots and cotyledons, while *VfPEPc2* was expressed in green tissue, maternal tissue and young cotyledons, and is thought to be the "housekeeping" form. *VfPEPc1* is thought to be sink specific as its expression peaks just before storage proteins begin to accumulate. In addition, it decreased in expression and activity in the

absence of glutamine and sucrose. Furthermore the effect of glutamine was only obvious in the absence of sucrose. Golombek *et al.* (1999) suggested sucrose may override the modulating effect of glutamine on *VfPEPc1* expression, as sucrose is thought to be a prerequisite for storage activity in sink organs (Martin *et al.*, 1997; Weber *et al.*, 1997). While the transcript level of PEPc 1 decreases as storage proteins begin to accumulate in seeds, the enzyme activity continues to rise. This would indicate that PEPc turnover is low during seed development as PEPc 2 transcripts do not fluctuate.

The expression and localisation of PEPc in wheat grain was also studied by Gonzalez *et al.* (1998). Antibodies to PEPc recognised two proteins of 103 and 108 kDa. The larger form appeared to be inducible, disappearing 15 days post anthesis, while the 103 kDa form, the majority in grains, appears to be a stable housekeeping form. This seems to be determined by mRNA levels. Localisation studies showed that both forms are found in seed tissues exhibiting high metabolic activity such as the aleurone layer leading to malate accumulation in the endosperm. Interestingly, PEPc was also seen at high amounts around and in the vascular tissue of seeds. On imbibition, PEPc was detected across the grain suggesting a role in re-fixing CO_2 released from respiring tissue and supplying 4-carbon skeletons for aminoacid biosynthesis. The larger form was found at a high level in the scutellar epithelium of the seed after imbibition. This may link to the synthesis and transport of nutrient from the endosperm (Drozdowicz and Jones, 1995).

1.6 The phosphorylation of phosphoenolpyruvate carboxylase in higher plants

1.6.1 PPCK in higher plants

The phosphorylation state of PEPc represents the steady-state balance of phosphatase and kinase activities. The identification of the soluble protein kinase in plants responsible for the phosphorylation of PEPc on its target serine residue became a core research question in the 1990's. Several groups (Bakrim *et al.*, 1992; Echevarria *et al.*, 1988; Jiao and Chollet, 1989; Li and Chollet, 1993; McNaughton *et al.*, 1991; Wang and Chollet, 1993) attempted to purify the enzyme, with limited success, partially purifying PPCK from illuminated maize and *Sorghum* leaves by ammonium sulphate precipitation and affinity chromatography. Using an *in vitro* PPCK assay, some groups identified both Ca^{2+}-dependent and Ca^{2+}-independent kinases capable of phosphorylating PEPc. The Ca^{2+}-dependent kinase was characterised in maize (Ogawa *et al.*, 1992, 1998) using specific inhibitors of myosin light chain kinase, a

Ca^{2+}-calmodulin dependent kinase. The calmodulin antagonist, W7, and the presence of EGTA inhibited the kinase activity, confirming that their kinase was indeed Ca^{2+}-dependent. The enzyme was able to phosphorylate the target serine of PEPc, but was unable to cause any significant changes in the malate sensitivity of PEPc possibly because only a low phosphorylation stoichiometry was achieved. It was also shown by SDS-PAGE to have a subunit size of 50 kDa. Gel permeation chromatography suggested that it existed as a homodimer or heterodimer. In *Sorghum* a Ca^{2+}-dependent PPCK was shown to bind to photosynthetic PEPc forming a stable complex *in vitro* (Bakrim *et al.*, 1992, Nhiri *et al.*, 1998). Giglioli-Guivarch *et al.* (1996) had suggested that photosynthetic PEPc phosphorylation may be mediated by an increase in cytosolic pH and calcium in mesophyll protoplasts. Nhiri *et al.* (1998) investigated whether induction of the Ca^{2+}-dependent PPCK could be caused by incubation of illuminated *Sorghum* protoplasts with weak base. After incubation, soluble protein was extracted and *in vitro* PEPc assays were conducted. The results indicated that the Ca^{2+}-dependent kinase in the mesophyll hardly phosphorylates photosynthetic PEPc, while a Ca^{2+}-independent PPCK that is up regulated in the light is responsible for the phosphorylation *in vivo*.

A Ca^{2+}-independent PPCK was partially purified from illuminated maize leaves. This kinase was purified 4000-fold and shown to be a 30 kDa monomer (Wang and Chollet, 1993b). Despite applying rapid purification protocols in the presence of a complete arsenal of protease inhibitors (as suggested by Ogawa *et al.*, 1992) Wang and Chollet found no traces of a Ca^{2+}-dependent PPCK in their purified PPCK preparation. Their Ca^{2+}-independent PPCK was shown to cause an increase in the $I_{0.5}$ (L-malate) of recombinant dephospho-*Sorghum* PEPc in parallel with the phosphorylation of Ser-8, the target serine.

As had been reported previously , the Ca^{2+}-independent PPCK studied by Wang and Chollet (1993b) was inhibited by L-malate (Jiao and Chollet, 1991; McNaughton *et al.*, 1991; Wang and Chollet, 1993b). This was most probably due to an effect of L-malate on the structure of the kinase substrate (PEPc) because the phosphorylation of PEPc by mammalian cAMP-dependent protein kinase (PKA) was also inhibited by L-malate, whilst L-malate had no effect on the phosphorylation of casein by PKA (Wang and Chollet, 1993b).

Chollet's group made significant progress using an "in gel" kinase assay. This involved separation of the kinase sample on denaturing gels containing dephospho-

PEPc. After *in situ* renaturation of the proteins in the gel, segments were incubated in the presence of [γ-^{32}P] ATP. Using this approach Li and Chollet (1993) resolved one Ca^{2+}-dependent (57 kDa) and two Ca^{2+}-independent (37 and 30 kDa) polypeptides capable of phosphorylating PEPc *in vitro* exclusively at the target serine residue. The 57 kDa and 37 kDa polypeptides were capable of autophosphorylation whilst the 30 kDa polypeptide was not, and was therefore assumed to correspond to the 30 kDa PPCK partially purified by Wang and Chollet (1993).

In order to resolve which of these polypeptides was the physiologically relevant PEPc kinase, Li and Chollet (1993) performed experiments in which maize leaves were supplied with inhibitors known to prevent the light-induction of PPCK activity (cycloheximide and methyl viologen). Subsequent analysis of extracts from these leaves using the *in situ* renaturation approach revealed that the 57 kDa polypeptide was present in darkness and unaffected by the inhibitor treatments. The 37 and 30 kDa polypeptides however were not present in darkness and their upregulation by light was blocked by inhibitors of protein synthesis and photosynthetic electron transport. The authors argued that the physiologically relevant PEPc kinase, which undergoes light activation that is blocked by cycloheximide and methyl viologen, is Ca^{2+}-independent and is either one or both of the 37 kDa and 30 kDa polypeptides (Li and Chollet, 1993).

The appearance of the C_4 PPCK activity in response to light requires protein synthesis. The simplest explanation for this observation was that the component synthesised was the PPCK itself. Hartwell *et al.* (1996) developed an *in vitro* method of measuring the translatable mRNA message for the kinase based upon the *in vitro* translation of mRNA, followed by direct assay of PPCK activity on purified dephospho-PEPc. They showed that in maize, illumination of mature leaves caused a seven-fold increase in translatable PPCK mRNA, compared to dark leaf control. This indicates that the light induction of maize PEPc phosphorylation is due at least in part to light-induced increases in maize PPCK mRNA levels.

In C_3 plants, PPCK has been investigated by a number of groups in a range of species. Tobacco leaf PPCK was found to be Ca^{2+}-independent confirming early findings for both the tobacco leaf enzyme and the wheat leaf enzyme (Duff and Chollet, 1995; Li *et al.*, 1996; Wang and Chollet, 1993a). "In gel" kinase assays on the partially purified kinase from tobacco demonstrated that the light-induced, Ca^{2+}-

independent PPCK polypeptides in tobacco leaves have molecular weights of 30 and 37 kDa similar to maize and *Sorghum* (Li *et al.*, 1996). However, the tobacco PPCK appears to be influenced by the glutamine status of the leaf (Li *et al.*, 1996) as the light induction of PPCK is blocked by incubation with methionine sulphoxine (a GS inhibitor). This is reversed, partially, by feeding glutamine to the leaves.

Zhang and Chollet (1997) partially purified PEPc kinase from soybean nodules and revealed two Ca^{2+}-independent PPCK polypeptides with approximate molecular weights of 32 and 37 kDa. They showed that the effects of photosynthate supply from the shoots on PEPc phosphorylation was mediated through parallel changes in PPCK activity. However, given the high malate levels that could occur in nodules (Streeter 1987; Schuller, Turpin *et al.*, 1990), Woo and Xu (1996) proposed that protein phosphorylation and metabolite activation by G-6-P could act synergistically.

Osuna *et al.* (1999) partially purified a Ca^{2+}-independent PPCK from aleurone cell protoplasts and aleurone endosperm tissue of dry and germinating barley seeds. They showed that the activity of this PPCK appears to be present in dry seed. In germinating seeds, the *in situ* phosphorylation state of PEPc increases following imbibition. Unlike other PPCK activities, this phosphorylation is not affected by treatments preventing protein synthesis (Osuna *et al.*, 1999). Hartwell *et al.* (1996) conducted similar PPCK assays, as described before for maize mRNA, upon barley leaf total RNA. The results showed that the translatable kinase mRNA was up regulated (doubled) in a light induced fashion over the dark leaf control. This increase suggested that the translatable mRNA of PPCK in C_3 plants is controlled in a similar fashion as C_4 plants. Several other studies in the leaves of C_4 and CAM plants have also shown the importance of protein synthesis in the up-regulation of Ca^{2+}-independent PPCK with addition of protein synthesis inhibitors preventing PPCK induction in the light (in C_4 plants) and the dark (in CAM plants) (Echevarria *et al.*, 1990, Carter *et al.*, 1991, Jiao *et al.*, 1991a, McNaughton *et al.*, 1991, Vidal and Chollet, 1997). These results would agree with the findings presented by Hartwell *et al.* (1996).

In the CAM plant *K. fedtschenkoi*, a Ca^{2+}-independent kinase from the leaves was shown to be capable of phosphorylating PEPc (Carter *et al.*, 1991). The plants were adapted to short days (illumination of 0800-1600 h). The kinase appeared

between 2000-2200 h and persisted throughout the night, becoming undetectable by 0600 h. The presence of this activity correlated with the period during which the K_i of PEPc was high. Hence it was concluded that this kinase was probably responsible for the *in vivo* phosphorylation of PEPc. Furthermore, the kinase activity oscillated in constant conditions, suggesting that it caused the circadian changes in the phosphorylation state of PEPc previously reported by Nimmo *et al.* (1987a). It has also been shown (Carter *et al.*, 1991, Nimmo *et al.* 1993) that the nocturnal appearance of PPCK in *K. fedtschenkoi* is blocked by inhibitors of protein synthesis (cycloheximide and puromycin) or RNA synthesis (cordycepin and actinomycin D). The rapid decline in PPCK activity both following exposure of leaves to inhibitors and towards the end of the dark period in untreated leaves led to the conclusion that this kinase protein probably has a short half-life. Brulfert *et al.* (1996) showed similar results in PEPc phosphorylation from *K. diagremontiana* leaves, suggesting that the conclusions are valid for other types of CAM species.

The level of PPCK translatable mRNA in *K. fedtschenkoi* was shown to vary 20 fold over the diurnal cycle, oscillated in constant environmental conditions and paralleled the phosphorylation state of the leaf PEPc (Hartwell *et al.*, 1996). Inhibitor experiments, similar to those described above, showed that the appearance of kinase translatable mRNA was blocked by inhibitors of RNA synthesis.

In the facultative CAM plant *Mesembryanthemum crystallinum*, salt stress or water stress induces CAM photosynthesis (Edwards *et al.*, 1996). A range of genes is induced by this stress, including PEPc 1 (cytosol) with maximal induction after 10 days. These stresses also greatly increased the level of a Ca^{2+}-independent PPCK activity which was much more active at night then during the day (Li and Chollet., 1994; Cushman and Bohnert, 1997). This PPCK activity was partially purified and shown to have a molecular weight of 39 and 32 kDa by "in gel" assays.

A common feature of PPCK in all systems is that it is a very low abundance protein, probably less than 1 in 10^6 of soluble protein (Ueno *et al.*, 2000). Consequently no antibodies have been raised against PPCKs purified from plants.

1.6.2 The dephosphorylation of PEPc

Incubation experiments on *in vitro* labelled PEPc suggested that the catalytic subunit of protein phosphatase type 2A (PP2A) could dephosphorylate PEPc. In contrast protein phosphatase type 1 (PP1) could not do so (Carter *et al.*, 1990, McNaughton *et al.*, 1991). This dephosphorylation by PP2A caused a rise in the

sensitivity of PEPc to L-malate. Investigations into the level of PP2A activity in maize leaf extracts were undertaken, using phosphorylase-a as a substrate. This showed that that in illuminated and darkened extracts, the specific activity of PP2A remained constant (McNaughton *et al.*, 1991). PP2A activity was also investigated in the leaves of the CAM plant K. fedtschenkoi (Carter *et al.*, 1991). Examination of the specific activity of PP2A over the 24 h period showed no diurnal variation in its activity. These results from a C_4 and CAM plant indicate that PPCK activity is the main regulator of PEPc phosphorylation state. PP2A was recently partially purified from *Zea mays* leaves (Dong *et al.*, 2001). It was shown to be analogous to other eukaryotic PP2As in being a approximate to 170-kDa heteromer. This heterotrimer lacks any strict substrate specificity in that it dephosphorylates C_4 PEPC, mammalian phosphorylase a, and casein *in vitro*. Also there were limited light/dark effects *in vivo*, suggesting that PPCK is the main regulator of PEPc phosphorylation state (Dong *et al.*, 2001).

1.6.3 Identification of PPCK

The first PPCK cDNA was cloned by adapting an assay for translatable PPCK mRNA. Hartwell *et al.* (1999) were able to detect clones expressing PPCK activity by transcribing linearized plasmid, and then translating and assaying the transcription products. They isolated a single clone from a cDNA library from the CAM plant *K. fedtschenkoi*, by successively subdividing the library into pools and selecting the population whose plasmids encoded most PPCK translatable mRNA. The expressed protein was shown to have high Ca^{2+}-independent PPCK activity and a molecular weight of 30 kDa - properties that were expected from previous biochemical work. Furthermore, peptide mapping showed that the expressed kinase phosphorylated PEPc at the same site as becomes labelled *in vivo* (the N-terminal serine). Interestingly, Northern analysis showed that the PPCK transcript level varies over the diurnal cycle in the same fashion as the PPCK translatable mRNA. Also the PPCK transcript level in detached leaves in constant environmental conditions maintained a circadian oscillation, suggesting that the PPCK gene is controlled by a circadian clock. Thus this PPCK gene clearly encodes the kinase responsible for controlling the phosphorylation state of CAM photosynthetic PEPc.

The sequence of PPCK shows it is a novel member of the calmodulin-dependent protein kinase (CaMK) group. It is closely related to plant Ca^{2+}-dependent protein kinases (CDPKs), but lacks the autoinhibitory region and the Ca^{2+}-binding EF

hands of CDPKs (Hartwell *et al.*, 1999). This explains its Ca^{2+}-independence, and suggests PPCK is essentially unregulated when expressed (Nimmo, 2000). This will be discussed further in section 1.6.4.

Database searches revealed two PPCK genes (*AtPPCK1* and *AtPPCK2*) in the *A. thaliana* genome (Hartwell *et al.*, 1999; Fontaine *et al.*, 2002), both encoding proteins with a molecular weight of 31 kDa. Both corresponding cDNAs were cloned by RT-PCR, transcribed and translated *in vitro*. Both translation products showed high Ca^{2+}-independent PPCK activity. Northern analysis of these genes failed to detect any transcripts, suggesting that they are transcribed at low levels. Semi-quantitative RT-PCR analysis show that the genes have different tissue expression patterns and are both light induced (Fontaine *et al.*, 2002). *AtPPCK1* is most highly expressed in rosette leaves, whereas *AtPPCK2* is low in rosette leaves but highly expressed in flowers and roots. The effect of protein synthesis inhibitors was tested by adding cycloheximide to a 3 day old *A. thaliana* cell culture. This caused increases in expression of both genes, in both light and dark. One possible explanation is that preventing synthesis of kinase proteins leads to an increased expression of PPCK genes via an effect on metabolism, because PEPc activation cannot be achieved in the absence of PPCK.

A fourth PPCK cDNA clone was identified by Taybi *et al.* (2000) from salinity-stressed leaves of the facultative CAM plant, *M. crystallinum*. This was achieved by using a protein kinase-targeted differential display reverse transcriptase polymerase chain reaction (DDrtPCR) approach. They used an anchored dT-primer and a degenerate primer designed to subdomain VIb of the serine/threonine protein kinase catalytic domain (Hanks and Hunter, 1995). PCR reaction products from cDNA of the salinity-stressed leaf mRNA (CAM induced) pool and the unstressed leaf mRNA (C_3) pool were labelled using [α-^{33}P]dATP. Following autoradiography, selected salinity-stressed induced bands were purified, cloned and sequenced. Putative PPCK sequences were analysed for specific expression by semi-quantitative RT-PCR. Clones showing a difference in day/night expression were subjected to RACE PCR to obtain a full length cDNA clone. The resulting clone, *McPPCK1*, encodes a 31.8 kDa Ca^{2+}-independent PPCK. The transcript levels of this gene in stressed leaves was controlled by a circadian oscillator in a similar manner as *KfPPCK*, as suggested by

RT-PCR experiments. Southern blot analysis provided some evidence for a second PPCK isogene in *M. crystallinum*.

The cloned kinase was over-expressed as a recombinant His-tagged PPCK in *E. coli* and was purified by His·Bind affinity chromatography on iminodiacetic acid-agarose. The purified recombinant form was shown to phosphorylate PEPc from *Sorghum*, maize and *M. crystallinum in vitro*, in a Ca^{2+}-independent fashion with maximal activity at pH 8.0, agreeing with the activity of the partially purified Mc PPCK from salinity stressed leaves (Li and Chollet, 1995). The protein was unable to phosphorylate a phosphorylation-site mutant of *Sorghum* PEPc, or sucrose synthase, a target enzyme for CDPK from soybean nodules. These results indicate that Mc PPCK 1 targets the N-terminal region of PEPc in a specific manner.

The L-malate inhibition of dephospho-PEPc from illuminated *M. crystallinum* leaves was shown to be reversed by *in vitro* phosphorylation upon incubation with recombinant Mc PPCK 1. Similar experiments with the phospho-PEPc (night form) had little effect upon PEPc K_i (L-malate). Thus the biochemical properties were in agreement with those reported for purified leaf PPCK (Carter *et al.*, 1991; Li *et al.*, 1996, 1997; Li and Chollet, 1993, 1994; Wang and Chollet, 1993a, 1993b; Zhang and Chollet, 1995).

Recently, a fifth PPCK cDNA clone, was identified from the C_4 dicot plant, *F. trinervia* and then expressed in *E. coli* (Tsuchida *et al.*, 2001). Degenerate primers to subdomain I and VII (Hanks and Hunter, 1995) designed from conserved regions of *KfPPCK* and *McPPCK*, were used in a RT-PCR reaction. This generated a 520 bp product from *F. trinervia* leaf cDNA. The RT-PCR product was used to screen a *F. trinervia* illuminated leaf cDNA library. One of the plaques that hybridised to the 520 bp product was shown after sequencing to be a full length Ft PPCK.

The recombinant (Nus-tagged) clone was expressed in *E. coli* and purified. Biochemical characterisation showed that the recombinant kinase phosphorylated dephospho-PEPc from maize, but was unable to phosphorylate a maize PEPc phosphorylation site mutant, nor did it phosphorylate common substrates of protein kinases such as casein and histone IIIS. Thus the enzyme seemed very specific for the N-terminal serine of PEPc. Furthermore, the kinase was Ca^{2+}-independent as seen by other groups (Fontaine *et al.*, 2002; Hartwell *et al.*, 1999, Taybi *et al.*, 2000). Interestingly, addition of oxidant and reductant to recombinant Ft PPCK caused inhibition and activation respectively. Northern analysis suggest that Ft PPCK is the

C_4 photosynthetic PPCK, with transcript levels high at 1200 h and low at 2400 h. Southern blot analysis of *F. trinervia* (C_4) and *F. pringlei* (C_3) appeared to show two genes, suggesting the presence of another PPCK gene, perhaps a housekeeping form.

Finally several other full length PPCK cDNAs have been putatively identified by our lab and others. These include one from sugar beet (C_3) (Herl *et al.*, unpublished) [GenBank number: AJ309171], *Sorghum* (C_4) (Fontaine *et al.*, unpublished) [GenBank number: AF399915], three from soybean (C_3) (Sullivan *et al.*, unpublished) [AY144181, AY144183 and AY144165] and one from tomato (C_3) (Hartwell *et al.*, unpublished) [GenBank number: AF203481].

1.6.4 Structure of PPCK

The sequencing of the known PPCK genes (including *AtPPCK1, AtPPCK2, FtPPCK, KfPPCK, McPPCK*) showed that PPCK comprises a typical protein kinase catalytic domain with minimal N- and C-terminal extensions (Fontaine *et al.*, 2002; Hartwell *et al.*, 1999; Taybi *et al.*, 2000; Tsuchida *et al.*, 2001). The PPCKs bear similarity to CaMK group of kinases which hold members that are regulated by Ca^{2+}/calmodulin (Hanks and Hunter, 1995). Unlike these kinases the PPCKs lack the EF hands and autoinhibitory regions common to plant CDPKs. The genomic sequences of *A. thaliana* and *K. fedtschenkoi* have also been published (Fontaine *et al.*, 2002, Hartwell *et al.*, 1999) revealing that a 3'-end intron is conserved positionally in all the kinase sequences. Our lab has shown that this trend is continued in the genomic sequences of *Sorghum* (C_4) (Fontaine *et al.*, unpublished) and all three from soybean (C_3) (Sullivan *et al.*, unpublished).

Several lines of evidence exist to suggest a small family of *PPCK* genes exist in each plant species. *A. thaliana, Sorghum* and soybean have all been shown to have multiple forms of PPCK (Fontaine *et al.*, 2002, unpublished; Hartwell *et al.*, 1999; Sullivan *et al.*, unpublished), while Southern analysis suggests that *M. crystallinum* and *F trinervia* and *F. pringlie* also hold multiple forms (Taybi *et al.*, 2000; Tsuchida *et al.*, 2001). Of the published sequences so far, only the Ca^{2+}-independent ~30 kDa forms has been cloned and identified. The absence of a 37 kDa form would suggest that this form may be post-translationally modified rather than the product of a separate gene. As more plant genomes and ests are sequenced it is reasonable to expect further additions to the PPCK sequences.

1.6.5 Control of PPCK

The light induction of C_4 maize PPCK activity was found to require a photosynthetic photon flux density (PPFD) of at least 300 μmol m^{-2} s^{-1} as inferred from the malate sensitivity of PEPc at a range of PPFDs (Bakrim *et al.*, 1992; Nimmo *et al.*, 1987a). Light quality had no effect on the light-induced decrease in the L-malate sensitivity of PEPc, with white, red and blue light proving equally effective (McNaughton *et al.*, 1991), suggesting that the induction of PPCK was not controlled by phytochrome or the blue-light photoreceptor. The requirement for at least 300 μmol m^{-2} s^{-1} of light for a change in the malate sensitivity of PEPc suggested that flux through photosynthesis is involved in the light-induction of C_4 PPCK.

To test whether various segments of the overall pathway of photosynthesis are required for induction of PPCK, detached maize leaves were allowed to take up inhibitor prior to illumination. Photosystem II (DCMU, isocil), photosystem I (methyl viologen) and RPP cycle (DL-glyceraldehyde) inhibitors all blocked the subsequent light activation of PPCK and halted the change in the L-malate sensitivity of PEPc (Jiao and Chollet, 1992; McNaughton *et al.*, 1991). Similar data were obtained using *Sorghum* leaves, suggesting that the signal transduction pathway for the regulation of PPCK and PEPc in C_4 plants is conserved (Bakrim *et al.*, 1992). Also, Bakrim *et al.* (1992) supplied an ATP synthesis inhibitor (gramicidin) to detached *Sorghum* leaves and demonstrated that ATP synthesis is also required for the light-induction of C_4-PEPc kinase.

The elucidation of the upstream signal transduction cascade leading to PEPc phosphorylation in C_4 plants has recently been aided by studies using mesophyll cell protoplasts from *Digitalis sanguinalis* (Giglioli-Guivarc'h *et al.*, 1996; Coursol *et al.*, 2000). The studies showed that increased PPCK activity and phosphorylation of PEPc could be achieved in illuminated mesophyll cell protoplasts by incubating them in NH_4Cl or methylamine. This treatment of protoplasts with weak bases caused an increase in pH of the cytosol. Giglioli-Guivarc'h *et al.* (1996) reasoned that illumination might cause a similar increase in cytosolic pH. They confirmed this using pH indicator dyes. They also suggested that the cause of the pH increase might be uptake of 3-PGA into mesophyll cell chloroplasts, in the C_4 pathway. Incubating illuminated mesophyll protoplasts with 3-PGA mimicked the weak acid effect. The inhibitors TMB-8 and W7, an inositol trisphosphate ($InsP_3$)-gated Ca^{2+} channel blocker and a calmodulin antagonist respectively, were shown to block this weak

base/illumination PPCK activation (Giglioli-Guivarc'h *et al.*, 1996). Further research by Coursol *et al.* (2000) showed that the light and weak base treatment of mesophyll protoplasts promotes a rapid transient increase of $InsP_3$. Treatment with U-73122, an inhibitor of phosphoinositide-specific phospholipase C activity, prevented the $InsP_3$ increase and the phosphorylation of PEPc, and reduced the activation of PPCK (Coursol *et al.*, 2000).

An additional factor controlling PPCK activity was suggested by Saze *et al.* (2001). They purified maize leaf PPCK to homogenity, revealing it to be a 30 kDa protein. During purification it was noted that the maize PPCK only retained its activity in the presence of a mild reducing agent in the purification buffer, and was inactivated by oxidising agents. Experiments showed that thioredoxins from *E. coli* or plants were capable of reactivating PPCK, suggesting a role for redox regulation of maize PPCK. Testing of this hypothesis would be facilitated by the sequencing of maize PPCK cDNA which might reveal Cys residues involved in potential disulphide bridges.

A third component that has recently been postulated in the control of PEPc phosphorylation is the presence of a protein inhibitor of PPCK. The inhibitor was detected within maize leaf (C_4) and *K. fedtschenkoi* leaf (CAM) (Nimmo *et al.*, 2001). Preliminary dilution experiments showed that leaf extracts contained some inhibitory material. The inhibitor was separated from PPCK by hydrophobic chromatography and its nature was investigated. This work showed that it was a protein which inhibits PPCK reversibly. It was not a protease, an ATPase or a protein phosphatase. It was present at sufficient concentrations to inhibit the basal activity of PPCK during the day in CAM plants and during the night in C_4 plants. Thus the inhibitor protein may play a physiologically important role in the regulation of photosynthetic PEPc in ensuring that phosphorylation only occurs when the primary fixation of CO_2 is required. The suggested model is outlined in figure 1.8.

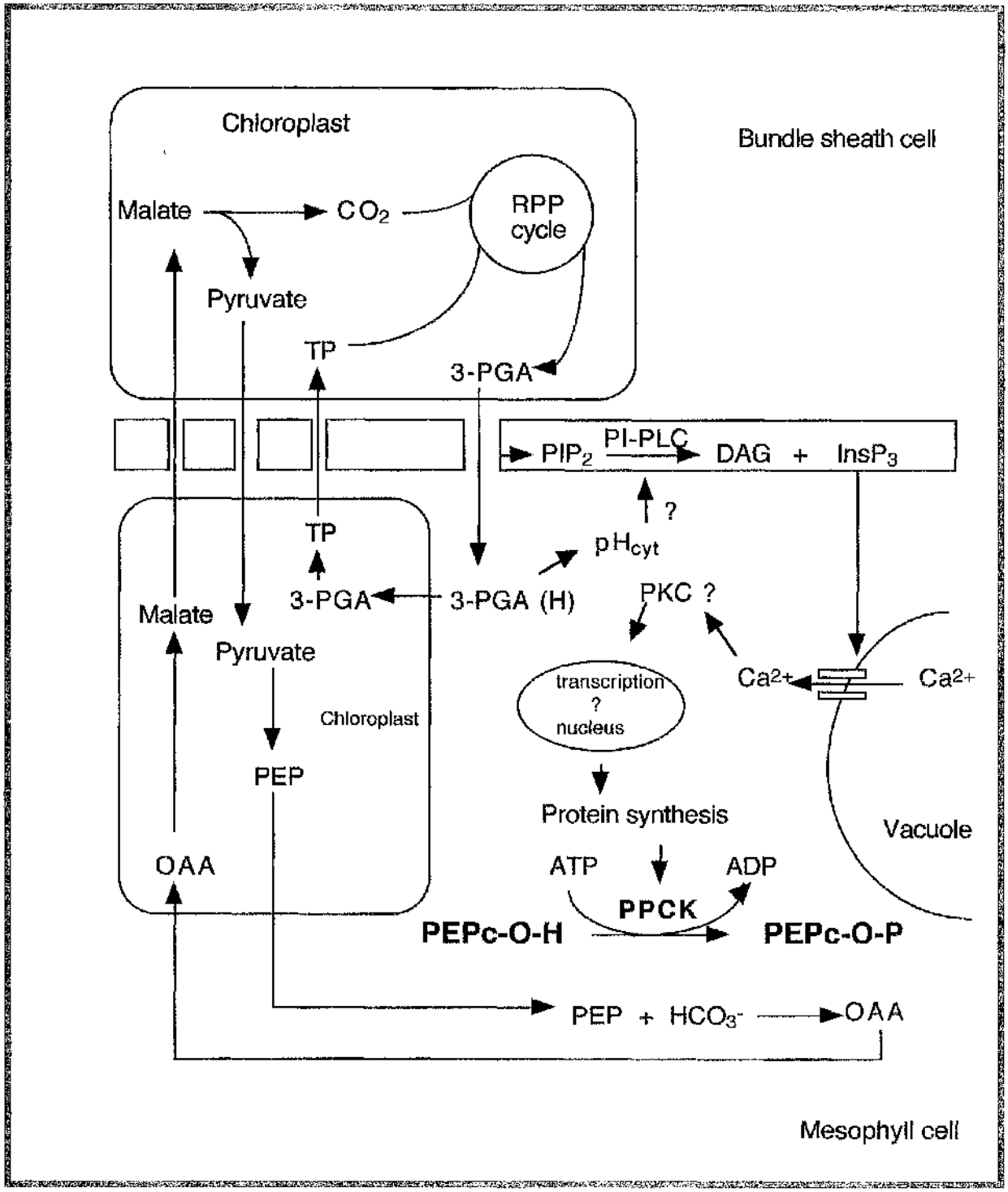
As noted earlier, detached CAM leaves show persistent circadian rhythms of CO_2 fixation in constant conditions. These are paralleled by rhythms in the transcript level and activity of PPCK and the phosphorylation state of PEPc (Wilkins, 1992, Nimmo *et al.*, 1987, Hartwell *et al.*, 1996, 1999). In common with many other circadian rhythms, these CO_2 rhythms in LL (continuous light) or DD (continuous dark) can be re-set by exposure to low or high temperature. This prompted Carter *et al.* (1995a and b) to examine the effect of temperature on both the activity and the

Figure 1.8 The proposed signalling pathway which mediates the light induction of PPCK in C_4 plants

Abbreviations: AOA, oxaloacetate; PiP_2,

Schematic model for the spatio-temporal organization of the transduction cascade in C_4 leaves. An increase in 3-phosphoglyceric acid (3-PGA) ensures the intercellular coupling via pH_{cyt} changes in the mesophyll cell cytosol. This is followed by activation of a mesophyll cell phospholipase C (PI-PLC) and transient production of the second messengers inositol-1, 4, 5-trisphosphate ($InsP_3$) and diacylglycerol (DAG). $InsP_3$ causes tonoplast calcium channels to open and efflux of calcium into the cytosol while DAG and calcium activate a PKC-like (?) activity. The rapid synthesis of the Ca^{2+}-independent PPCK leads to the phosphorylation of C_4 PEPc. Question marks indicate the steps that remain to be elucidated.

This figure was adapted from Jeanneau *et al.* (2002)


Chloroplast
Bundle sheath cell
Malate
CO_2
RPP cycle
Pyruvate
TP
3-PGA
PIP_2
PI-PLC
DAG + $InsP_3$
?
pH_{cyt}
TP
3-PGA
3-PGA (H)
PKC ?
Malate
Pyruvate
transcription ? nucleus
Ca^{2+}
Ca^{2+}
Chloroplast
PEP
Protein synthesis
Vacuole
OAA
ATP
ADP
PPCK
PEPc-O-H
PEPc-O-P
PEP + HCO_3^-
OAA
Mesophyll cell

phosphorylation state of PEPc. They found that reducing the temperature, which of course reduces the V_{max} of PEPc, also reduced its malate sensitivity. This would help to maintain flux through PEPc during cold nights in CAM plants.

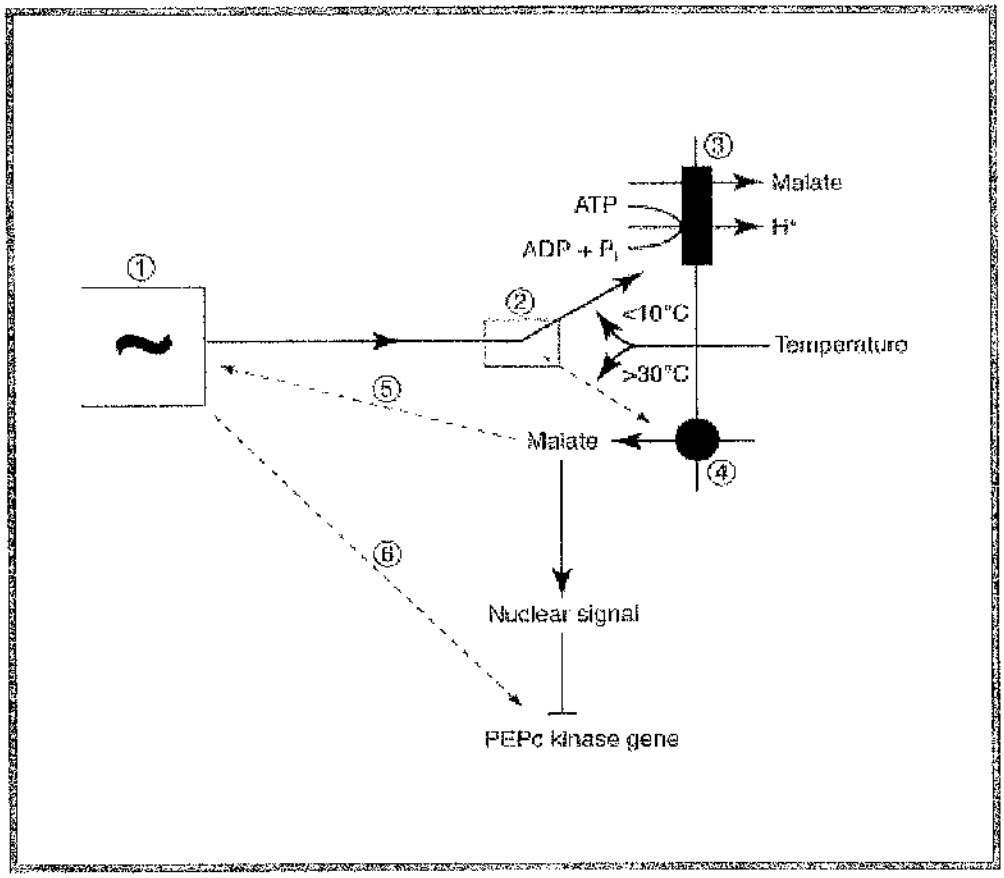
More importantly they studied the effect of temperature in constant darkness. In CO_2-free air at 15°C in DD, leaves show persistent circadian rhythms of CO_2 output. In normal air at 15°C in DD, leaves show one big phase of CO_2 uptake corresponding to a normal night. Then PEPc becomes dephosphorylated, the leaf maintains a high level of malate, and slow CO_2 output is observed. However, if the temperature is reduced to 3°C, PEPc remains phosphorylated and CO_2 is fixed continuously for much longer - up to 60 h. In contrast, at temperature greater than 25°C, PEPc remains dephosphorylated and a rhythm of CO_2 output may be observed. Hence the phosphorylation of PEPc does not drive the circadian rhythms in CO_2 fixation. Rather, its role is probably to modulate the amplitude of CO_2 uptake (Nimmo, 2000). This work also shows that the circadian control of PPCK can be disrupted by temperature changes.

Borland *et al.* (1999) showed that circadian control of PPCK could also be disrupted by preventing nocturnal accumulation of malate. This was achieved by enclosing leaves in an atmosphere of N_2 overnight. When the leaves were released into normal air at dawn, they fixed CO_2 via PEPc to a much greater extent and for much longer than control leaves. This was because the levels of PPCK translatable mRNA and activity, and the phosphorylation state of PEPc, remained high for much longer into the day in the N_2-treated leaves than the control leaves (Borland *et al.* 1999)

Recently a model for the circadian control of PPCK expression has been proposed, see Figure 1.9 (Nimmo, 2000). In this model, the primary output from the circadian oscillator sets a switch that controls malate transport across the tonoplast. This would result in circadian oscillation in the level of cytosolic malate (and/or other metabolites). Cytosolic malate (or a related metabolite) acts as a feedback inhibitor of PPCK expression by an unknown mechanism. There may also be a direct connection between the oscillator and PPCK gene expression, perhaps via a homologue of the

Figure 1.9 Hypothetical model for the circadian control of PEP carboxylase by metabolites

The primary oscillator (1) drives a switch (2) that affects the tonoplast, controlling the rates of malate uptake energized by an H^+-ATPase (3) and malate release (4). Temperature extremes affect the switch operation as shown. The nuclear signal inhibits the expression of PEP carboxylase kinase. In addition, there might be feedback from cytosolic malate to the oscillator (5) and some direct connection between the oscillator and the expression of PEP carboxylase kinase (6).
Broken arrows represent connections that might occur.
This figure was adapted from Nimmo *et al.* (2001).


③
Malate
ATP
H⁺
ADP + P_i
①
②
<10°C
Temperature
>30°C
⑤
Malate
④
⑥
Nuclear signal
PEPc kinase gene

Arabidopsis transcription factor CCA1 whose expression is controlled by the circadian clock. However the recent observation (Hartwell *et al.*, 2002) that the effect of temperature on the expression of PPCK in *K. fedtschenkoi* is "gated" by the clock argues against this possibility and in favour of indirect control such as via metabolites.

1.7 Alternate splicing in higher plants

The majority of plant mRNAs (transcripts) are synthesised as precursors containing intervening sequences (introns) between coding sequences (exons) (Lorkovic *et al.*, 2000). To splice these introns from the primary transcripts, prior to translation, the spliceosome is assembled after recruitment to the primary mRNA splice recognition sites, resulting in the excision of the intron. The spliceosome is a large ribonucleoprotein (RNP) complex, which is assembled at each intron from small nuclear RNP particles and numerous protein factors. The first step involves cleavage at the 5' splice site, followed by the formation of an intron lariat at an adenosine nucleotide (the branchpoint). The second step involves the cleavage of the 3' splice site and the ligation of the 3' end of the first exon to the 5' end of the second exon. This dynamic but ordered reaction is outlined in Figure 1.10. The recognition of 5' and 3' splice sites by the spliceosome is multifaceted. Like other eukaryotes, plant genes define their exon-intron borders by consensus sequences. The 5' and 3' spice site consensus sequences are AG/**GT**AAGT and TGC**AG**/C respectively. Consensus sequences are recognition motifs that snRNA bind to prior to the formation of the spliceosome. A third loose consensus sequence is the intron internal branch site, CTR[A]Y, usually 18-40 nucleotides downstream of the 5' splice site. The bracketed A indicates the adenosine nucleotide that the 5' splice site is linked to. Plant introns though show an important difference to other eukaryotic introns by having a strong composition bias for AT-rich sequences, compared to the exon. These AT-rich introns have been shown to be necessary for efficient intron processing and splice site selection (Ko *et al.*, 1998; Simpson *et al.*, 1996). It is thought that AT-rich sequences minimise the secondary structure of introns because the introduction of hairpins into introns appears to have a strong negative effect upon splicing efficiency in dicot plant cells (Lui *et al.*, 1995). A second role may be to aid recognition of hnRNP-like proteins involved in the early formation of the spliceosome as suggested by the identification of the intron-binding protein, UBP-1 (Gniadkowski *et al.*, 1996). The mechanism of exon and intron definition is still not fully understood in any eukaryotic system, but it seems likely that trans acting elements may also be involved with the

Figure 1.10 The spliceosome cycle in mammals

The splicing of a single intron from a pre-mRNA is shown.

Base pairings of the U1 and U6 small nuclear RNAs (snRNAs) with 5' splice sites and the U2 snRNA with the branch point are indicated based on conservation of these sequences in plant pre-mRNAs and snRNAs.

Filled black circle at the ends of snRNAs represents a 5' terminal cap structure.

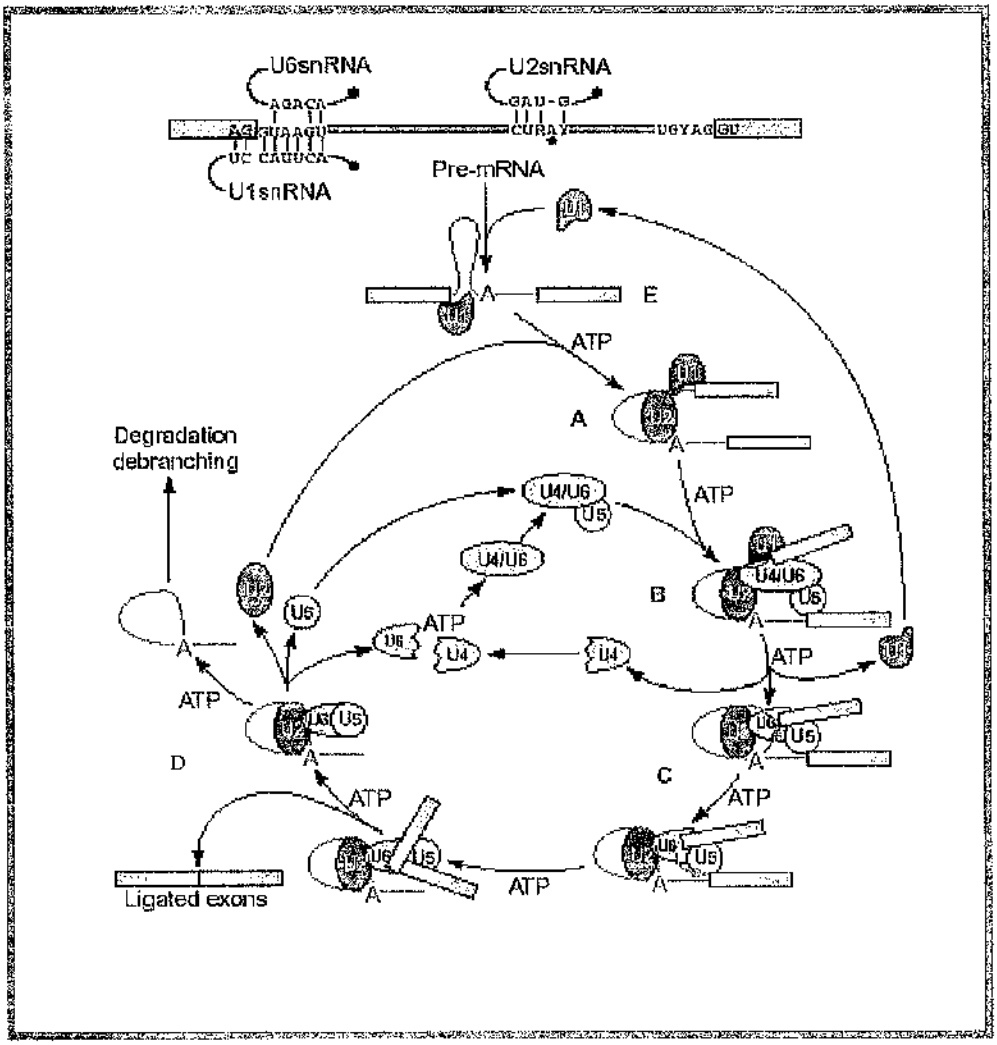
The involvement of five snRNPs and their interactions are indicated at distinct steps in spliceosome formation and catalysis.

All steps, bar the formation of complex E, require ATP.

Arrows indicate transitions between complexes and recycling of snRNPs.

After two transesterfications reactions (red arrows), two ligated exons are released as mRNA and the intron lariat is released and subsequently debranched and degraded.

This figure was adapted from Lorkovic *et al.* (2000).


U6snRNA
U2snRNA
Pre-mRNA
U1snRNA
E
ATP
A
Degradation
debranching
U4/U6
U5
U6
U4
B
C
D
Ligated exons

cis acting recognition sequences within the intron. These elements may include serine-arginine rich (SR) proteins that probably assist in interactions in mammalian exon-intron definition by binding to exonic splicing enhancers, protecting the exon from spliceosome recruitment and assembly. However exonic splicing enhancers have not been identified in plants (Kramer, 1996; Valcarcel *et al.*, 1996). SR proteins are composed of one or two N-terminally placed RNA binding domains that interact with specific sequences in the pre-mRNA and a domain rich in Ser-Arg dipeptides that are involved in protein-protein interactions (Lorkovic *et al.*, 2000).

Alternative splicing allows a gene to produce primary transcripts that may be spliced differently to produce several different mRNAs that encode proteins with different functions. In plants only a few examples of alternative splicing have emerged. Many alternative splicing events in higher plants are constitutive, with similar ratios of variant mRNAs in different cells, whereas others are subject to tissue specific or developmental regulation (Lorkovic *et al.*, 2000). The most common form of alternative splicing is intron retention (Brown and Simpson, 1998; Lorkovic *et al.*, 2000). This is thought to reflect the low efficiency of normal splicing rather than a regulated process of any biological significance. However one exception, where intron retention is thought to be regulated, is in alternative splicing of *FCA* transcripts. The *FCA* floral promotion gene contains 20 introns, of which alternative splicing of introns 3 and 13 produces four different transcripts α, β, δ and γ. Only transcript γ, where all introns are accurately spliced, encodes a protein that functions to control flowering time (Macknight *et al.*, 2002). Transcript β is produced by termination and polyadenylation within intron 3, encoding a protein that is little more than the N-terminal of the full length protein. Transcripts α and γ are alternatively spliced in intron 3 and intron 13 respectively, introducing in-frame stop codons, and thus producing truncated proteins. Upon overexpression of the gene, flowering time was accelerated, possibly due to a small increase in functional transcript γ. The level of polyadenylated transcript β, though, increased by 100 fold in these transgenic plants. This suggested to Macknight *et al.*(2002) that the polyadenylation event controls the formation of FCA to keep flowering time constant.

Selection of alternative 5' and 3' splice sites takes place in a small number of known plant introns. In some cases the ratio of alternative spliced variants is

constitutive, as in rubisco activase (Werneke *et al.*, 1989). In others, splice site selection appears to be a tissue or cell specific event. Examples of these include the H-protein of the glycine cleavage system in *F. trinervia*. There are two possible 3' splice sites in the first, with the distal site selected within floral and root tissue and the proximal site selected in leaf, stem and cotyledons (Kopriva *et al.*, 1995). Alternative 5' splice selection occurs in the hydroxypyruvate reductase gene in pumpkin at intron 12, producing two transcripts, one with a premature stop codon. The level of the truncated protein appears to be greatly enhanced by light, suggesting that alternative splicing may be light regulated in plants (Mano *et al.*, 1999).

Only a few alternatively spliced plant products have been shown to have a role, including separation of signal sequences (Brown and Simpson, 1998). Functional roles for alternative spliced products such as truncated proteins due to in frame stop codons, are not greatly characterised. In some cases the truncated peptides have been shown to accumulate, suggesting a functional role. Such a role may be to act as trans-dominant inhibitors (Macknight *et al.*,1997).

Understanding plant regulated and constitutive alternative splicing will require better characterisation of trans and cis acting factors of splicing. The dissection of this machinery will be intensively studied in the next decade.

1.8 Objectives

The major aim of this work was to identify and characterise PPCK genes from C_3 and C_4 plants. At the outset of my work only one PPCK gene had been identified, a PPCK from the CAM plant *Kalanchoe fedtschenkoi*. My initial objective was to use this sequence as the basis of a homology-based cloning approach in maize. The subsequent identification of a maize PPCK would allow testing for C_4 photosynthetic expression. A second aim arose from several sequences in the est database that bore predicted amino-acid similarity to the *K. fedtschenkoi* PPCK. Two of these are from the C_3 plant - tomato. This would allow cloning of full length cDNA sequences and testing for PPCK activity using a transcription-translation PPCK assay (as described in Hartwell *et al.*, 1996). In combination, the tomato plant allows us to investigate ripening. Therefore a third objective was to characterise the expression of these putative tomato PPCKs in fruit as well as 'standard' C_3 plant tissue. A final objective was to assemble a full list of putative PPCK sequences from the ever growing est databases. This would allow phylogenetic analysis to be conducted, possibly allowing us to assign function to the PPCKs.

Chapter 2

Materials and methods

2.1 Materials

Acetic acid (glacial), acetone, boric acid, chloroform, disodium hydrogen phosphate, ethylene diamine tetra-acetic acid (EDTA, disodium salt), glycine, Hepes, hydrochloric acid (HCl), magnesium chloride ($MgCl_2$), methanol, potassium chloride (KCl), sodium acetate, sodium chloride (NaCl), sodium hydrogen carbonate ($NaHCO_3$) and sodium hydroxide (NaOH) were from Fisher Scientific (U.K.) Ltd., Loughborough.

Actinomycin D, ammonium peroxodisulphate, antipain hydrochloride, benzamidine hydrochloride, Bis-Tris, bovine serum albumin (BSA), Bromophenol Blue, cantharidin, casein enzymatic hydrolysate, chymostatin, Coomassie Brilliant Blue G250, cordycepin, cyclohexamide, Denhardt's solution, diethyl pyrocarbonate (DEPC), dimethyl sulphoxide (DMSO), ethanol (100 % v/v), formaldehyde, formamide, glucose-6-phosphate, (G6P, monosodium salt), hexadecyltrimethyl ammonium bromide (CTAB), hydrogen peroxide, leupeptin (hemisulphate salt), lithium chloride, L-malic acid, 2-mercaptoethanol, octanol, phosphoenolpyruvate (PEP, monosodium salt), polyethylene glycol (PEG, MW 15000-20000), polyvinylpolypyrrolidone (PVPP or insoluble PVP), potassium dihydrogen orthophosphate, puromycin, Ribonuclease A, sodium azide, sodium dodecyl sulphate (SDS), spemidine hydrochloride and N,N,N',N'-tetramethylethylenrdiamine (TEMED) and trichloroacetic acid were from Sigma (London) Chemical Co., Poole, Dorset, U.K.

Adenosine 5'-triphosphate (ATP, disodium salt), nicotinamide adenine dinucleotide reduced form (NADH, disodium salt) and pig heart malate dehydrogenase (MDH) were from Roche Diagnostics Biochemicals, Mannheim, Germany.

[γ-^{32}P]ATP (triethylammonium salt, 3000 Ci mmol^{-1}), [α-^{32}P]dCTP (3000 Ci mmol^{-1}), Hybond N blotting membrane, rabbit reticulocyte *in vitro* translation kit and Redivue™ ^{35}S-methionine (1000 Ci mmol^{-1}) were from Amersham International, Bucks., U.K.

Agar, tryptone (peptone from casein) and yeast extract were from Merck, Darmstedt, Germany.

Agarose and 1 Kb DNA ladder were from Gibco BRL Life Technologies Ltd., Paisley, Scotland, U.K.

AMV Reverse transcriptase, C600 cells, Poly Atract mRNA isolation system, RNA size markers (0.28 – 6.6 Kb), RQ1 RNase-free Dnase, all DNA restriction enzymes and Wizard lambda preps DNA purification system were from Promega (U.K) Ltd., Southhampton, U.K.

Dithiothreitol (DTT) was from Alexis Corporation (U.K.) Ltd., Nottingham, U.K.

DNA customised synthesis and sequencing was carried out by MWG-Biotech (U.K.) Ltd., Milton Keynes, U.K.

DNA-*free*™ kit, mMessage mMachine T3 and T7 *in vitro* transcription kit was from Ambion, the RNA Company, Ambion Inc., U.S.A.

Hydroxyapatite (Bio-gel HTP) and 37.5:1 acrylamide:bis-acrylamide solution were from Bio-Rad Laboratories (England) Ltd., Bramley, Kent, U.K.

Oligo (dT)-cellulose redi column, Sephadex G25 (medium), Sephadex G50 and FPLC Superose 6 pre-packed column were from Pharmacia, Milton Keynes, Bucks., U.K.

Okadaic acid was from Moana Bioproducts.

QIA-quick gel extraction kit, Wizard DNA clean up kit and QIA-prep spin plasmid miniprep kit were from Qiagen Ltd., Crawley, West Sussex, U.K.

The maize seedling cDNA library was kindly supplied by Prof. Katsumura Izui. University of Kyoto, Japan.

Tris was from ICN Biomedicals, Ohio, U.S.A.

Topo TA Cloning Kit was from Invitrogen Life Sciences, UK.

2.2 Plant material

2.2.1 *Z. mays*

Maize (*Zea mays*, cv. Jubilee F1) was grown from seed in potting and bedding compost (William Sinclair Horticulture Ltd., Lincoln, U.K.) in a greenhouse under a 16 h photoperiod supplemented with mercury-vapour lamps. Three week-old plants were re-potted and transferred to a controlled environment chamber (Fitotron TM 600 growth chamber, Fisons, Loughborough, U.K.) and grown under a 12 h photoperiod using white fluorescent tubes, tungsten lamps, giving a light intensity of

700 μmol m^{-2} s^{-1}, and at a constant temperature (20°C). Mature leaves were harvested after 6 weeks, after either 10 h of darkness or 3 h of light, and used to isolate RNA.

2.2.2 *L. esculentum*

Tomato plants (*Lycopersicon esculentum* cv Alicante) were grown from seed on bedding compost (William Sinclair Horticulture Ltd., Lincoln, U.K.) in a greenhouse under a 16 h photoperiod supplemented with mercury-vapour lamps (300 μmol m^{-2} s^{-1}). Plants were re-potted at 4 weeks and then watered every 2 days.

Tomato fruit were allowed to ripen on the vine. Stages of ripening were assigned according to the colour classification by the U.S Department of Agriculture (Standards for Grade of Fresh Tomatoes) and harvested. Quartered segments were used to isolate RNA. Young leaves were harvested as the primary leaf after 10 days sowing, while mature leaves were harvested after 6 weeks. Tissue was collected during the dark (after 10 h) and light (after 3 h) before RNA purification.

2.2.3 Aubergine

Aubergine fruit were obtained from a local fruit shop and used to isolate DNA.

2.2.4 *S. tuberosum*

Potato (*Solanum tuberosum* cv. Desiree) plants were grown from tuber on bedding compost (William Sinclair Horticulture Ltd., Lincoln, U.K.) in a greenhouse under a 16 h photoperiod supplemented with mercury-vapour lamps (300 μmol m^{-2} s^{-1}). Plants were re-potted at six weeks and watered every two days. Leaves were harvested and used to isolate RNA and DNA.

2.2.5 *K. fedtchenkoi*

Kalanchoe fedtschenkoi Hamet et Perrier was propagated by cuttings from the original stock used in previous studies (Wilkins, 1959; Wilkins, 1960). Cuttings were grown in a greenhouse under a 16 h photoperiod supplemented with mercury-vapour lamps. Plants were watered every 5 days. Leaf harvesting was taken from between nodes six and ten.

2.3 General biochemical methods

2.3.1 pH calibrations were carried out using a Russel pH probe connected to an EDT instruments microprocessor pH meter. The pH of all solutions and buffers was adjusted at their working temperature (i.e. room temperature).

2.3.2 Glassware and plastics were cleaned in Haemo-sol solution (Alfred Cox (Surgical) Ltd., UK), rinsed in distilled water and oven dried.

2.3.3 Chromatographic materials. Sephadex G-25M, DEAE-cellulose and hydroxylapatite were packed according to the manufacturer's protocols. Sephadex G-25M was stored in 0.02 % (w/v) sodium azide. DEAE-cellulose and hydroxylapatite bound plant pigments tightly and therefore were discarded after use.

2.3.4 Concentration of protein samples was carried out using Centricon 30 microconcentrators (Amicon Ltd., Stonehouse, Glos., U.K.) according to the manufacturer's instructions.

2.3.5 Spectrophotometric assays were carried out in semi-micro quartz cuvettes (1 cm path length, 1 ml volume). A UV/Vis scanning spectrophotometer (Philips PU 8700 series) was used for the determination of the concentration of DNA/RNA solutions.

2.3.6 Centrifugation was performed in a microfuge, using Eppendorf tubes, for samples less than 1.5 ml. Larger samples were centrifuged in a Beckmann model J2-21 with a JA20 rotor.

2.3.7 Micropipetting was performed using adjustable Finnpipettes (Labsystems, Helsinki, Finland) or Gilson pipettes.

2.4 Enzyme extraction and purification

2.4.1 Purification of phosphoenolpyruvate carboxylase from *Kalanchoe fedtschenkoi*

The purification was carried out at 4°C except for the Superose 6 (FPLC) column, which was carried out at room temperature. Buffers used in the process are described in 2.4.2.

Dephosphorylated PEPc was extracted from 40 g of *K. fedtschenkoi* leaves that were collected between 09:00 and 11:00 h. The leaves were homogenised for 30 s at low speed, in a Waring blender, in the presence of 30 ml of Buffer 1, 1.5 g of insoluble polyvinyl pyrrolidone, 2.0 g of sodium bicarbonate and several drops of octanol. The homogenate was passed through two layers of muslin and centrifuged at 15000 rpm, for 15 min. The supernatant was passed through two fresh layers of muslin before loading upon a Sephadex G-25 M column (4.5 cm x 15 cm bed volume), which was pre-equilibrated with buffer 2. The sample was passed through the column in the presence of buffer 2, and 5 ml fractions were collected at a flow rate

of 5 ml min^{-1}. The $A_{280\ nm}$ of each fraction was monitored and the peak fractions were pooled. The resulting volume was added to 10 ml of hydroxyapatite (HAP), pre-equilibrated with buffer 2, and mixed gently by stirring, for 10 min. The HAP was pelleted out of the mixture by a brief centrifugation, at 5000 rpm, and the supernatant was discarded. The HAP pellet was re-suspended and re-poured in a column. The initial eluate was discarded and the column was washed with a further 30 mls of buffer 2. The column was then washed with Buffer 3 until the $A_{280\ nm}$ returned to baseline. Finally PEPc was eluted by changing to Buffer 4 at 2 ml min^{-1}. 2 ml fractions were collected, and the peak fractions were pooled. This volume was desalted on a Sephadex G-25 M column (2.2 cm x 55 cm) pre-equilibrated with Buffer 5, at a flow rate of 7 ml min^{-1}. Fractions (3.5 ml) were collected and the protein peak fractions were pooled to a maximum of 30 ml. These fractions were slowly loaded onto a DEAE-cellulose column (1 cm x 1.2 cm) pre-equilibrated with Buffer 5. Buffer 5 was washed through the column until the $A_{280\ nm}$ returned to baseline. The PEPc was eluted from the DEAE-cellulose column by Buffer 6. Fractions of 500 µl were collected and assayed for PEPc activity. The PEPc activity peak fractions were pooled, filtered and concentrated in Centricon-30 tubes at 5000 rpm using a 8 x 50 ml angled-rotor (JA-20) in a Beckman centrifuge (Model J2-21) until the volume was less than 500 µl. The concentrated sample was then loaded upon a Pharmacia (FPLC) Superose 6 column (30 cm x 1cm) pre-equilibrated with Buffer 7, and passed through the column at a flow rate of 0.3 ml min^{-1}. Three fractions of 3 ml were collected followed by 0.3 ml fractions. PEPc was eluted in fractions 20-25. The fractions were checked for PEPc activity and the peak fractions were pooled and dialysed overnight at 4 °C, against dialysis buffer. The purified PEPc was then aliquoted into separate tubes at 0.3 units per tube and stored at –70°C.

2.4.2 Buffers used in the purification of PEPc

Buffer 1

100 mM Tris/HCl (pH 8.0) containing 2 mM EDTA, 10 mM L-malate, 2% (w/v) polyethylene glycol 20000 and 1 mM benzamidine hydrochloride.

Buffer 2

100 mM Tris/HCl (pH 7.5) containing 2 mM EDTA, 10 mM L-malate, 1 mM DTT and 1 mM benzamidine hydrochloride.

Buffer 3

100 mM Tris/HCl (pH 7.5) containing 0.1 mM EDTA, 10 mM L-malate, 40 mM potassium phosphate, 1 mM DTT and 1 mM benzamidine hydrochloride.

Buffer 4

100 mM Tris/HCl (pH 7.5) containing 0.1 mM EDTA, 10 mM L-malate, 150 mM potassium phosphate, 1 mM DTT and 1 mM benzamidine hydrochloride.

Buffer 5

50 mM Tris/HCl (pH 7.5) containing 0.1 mM EDTA, 1 mM DTT and 1 mM benzamidine hydrochloride.

Buffer 6

50 mM Tris/HCl (pH 7.5) containing 0.1 mM EDTA, 100 mM NaCl, 1 mM DTT and 1 mM benzamidine hydrochloride.

Buffer 7

50mM Bis-tris Propane (pH 7.5) containing 0.1 mM EDTA, 50 mM potassium chloride, 1 mM DTT and 1 mM benzamidine hydrochloride

Dialysis buffer

50mM Bis-tris Propane (pH 7.5) containing 0.1 mM EDTA, 1 mM DTT, 1 mM benzamidine hydrochloride and 50% (v/v) glycerol.

2.5 Enzyme assays

2.5.1 Estimation of PEPc activity and the apparent K_i of PEPc for L-malate

The standard 1 ml assay cocktail contained 50 mM Tris/HCl, pH 7.8, 5 mM $MgCl_2$, 2 mM PEP, 0.2 mM NADH, 10 mM $NaHCO_3^-$, 5 units of MDH and the enzyme sample. The activity of PEPc was determined via spectrophotometric assay. The rate of decrease in A_{340nm} is dependent upon PEPc concentration (Nimmo *et al.* 1984). One unit of enzyme activity is equivalent to the amount of enzyme required to catalyze the formation of 1 μmol of product min^{-1}.

The apparent K_i of PEPc for L-malate was determined by adding a range of concentrations of L-malate to the assay cocktail. Apparent K_i (the concentration giving 50% of inhibition of initial PEPc activity) was determined using a plot of percentage inhibition of the rate minus L-malate versus L-malate concentration (Nimmo *et al.*, 1984).

2.6 DNA isolation

2.6.1 Isolation of DNA from plant leaf tissue

Plant leaf tissue was harvested and flash-frozen in liquid nitrogen and stored at –70°C. Frozen plant tissue (2-3 g) was ground to fine powder using a mortar and pestle. The fine powder was transferred to a plastic centrifuge tube (Sarstedt, U.K) containing 15 ml of DNA extraction buffer, which was pre-warmed in a 65°C water bath. The DNA extraction buffer was composed of 2 % (w/v) hexadecyltrimethyl ammonium bromide, 1.4 M sodium chloride, 20 mM EDTA, 100 mM Tris/HCl, pH 8.0 and 0.2% (v/v) β-mercaptoethanol (added after autoclaving). The tube was inverted several times, mixing the ground tissue and DNA extraction buffer, and returned to the 65°C water bath for a minimum of 30 min. Tubes were inverted during the incubation period to ensure thorough mixing. Next, 15 ml of Sevag solution (96 % (v/v) chloroform, 4 % (v/v) isoamylalcohol) was added to each tube and the tubes were shaken for 15 min, mixing both phases. Phases were separated by centrifuging for 10 min at 9500 rpm at 4°C using a 6 x 50 ml swing-out rotor (JS 13.1) in a Beckman centrifuge (Model J2-21). The upper aqueous phase was removed to a fresh 30 ml tube containing 0.67 volumes of ice-cold isopropanol. Samples were then placed upright, at 4 °C in a fridge for 20 min, to allow precipitation of DNA. The tubes were centrifuged for 10 min at 9500 rpm and at 4 °C to pellet the DNA. The supernatant was removed and discarded by using a sterile disposable plastic pipette. The pellets were dried for 5 min, by inverting the tubes onto clean, dry tissues. DNA pellets were re-suspended in 500 µl of ice-cold 10 mM Tris/HCl, pH 7.5 containing 1 mM EDTA (TE). The samples were transferred to a 1.5 ml microfuge tube containing 10 µl of RNase A (5 mg/ml). The samples were incubated for 30 min at 37°C. Equal volumes of phenol and Sevag solution were added and mixed to the samples and then the tubes were centrifuged (13000 rpm for 5 min at room temperature) using a 6 x 50 ml swing-out rotor (JS 13.1) in a Beckman centrifuge (Model J2-21). The upper aqueous phase was transferred to a fresh, sterile microfuge tube and 10 % (v/v) 3 M sodium acetate, pH 5.2 and 2.5 volumes of ice-cold ethanol were added and mixed. The DNA was allowed to precipitate for at least 60 min at –70°C. Tubes were then

centrifuged for 10 min at 13000 rpm, room temperature. The supernatant was removed and discarded using a sterile micropipette tip, and 250 µl of 70 % ethanol was added. The tubes were re-centrifuged for 10 min at 13000 rpm, room temperature. The supernatant was removed and discarded using a sterile micropipette tip and the DNA pellet was allowed to dry on ice. The dried pellet was resuspended in 50 µl of sterile TE and stored at –70 °C. This method was adapted from Sambrook *et al* (1989).

The quantity and purity of the DNA was established spectrophotometrically by monitoring the absorbance of the sample at 260 nm and 280 nm as described in Sambrook *et al* (1989). For DNA an A_{260nm} of 1 is equivalent to a concentration of 50 µg/ml. An $A_{260/280}$ ratio of between 1.8 and 2.0 indicates "clean" RNA, relatively free from protein contamination. An $A_{260/280}$ ratio below 1.8 indicates significant protein contamination.

2.6.2 Isolation of plasmid DNA

Plasmid DNA was isolated from 10 ml overnight cultures of the desired strain of *E. coli* containing the plasmid of interest using QIA-prep spin plasmid miniprep kit (Qiagen, U.K.) according to the manufacturer's protocols.

2.6.3 Isolation of lambda DNA

Lambda DNA was isolated from liquid lysate using a Wizard Lambda Preps DNA Purification System (Promega, U.K.) according to the manufacturer's protocols.

2.7 Polymerase chain reaction techniques

2.7.1 Design of primers

Degenerate oligonucleotide primers were designed for several conserved regions among previously cloned PEPc kinases of C_3 and CAM plants. The nucleotide sequences are shown in Table 1. The forward primers correspond to the sequences:

1. CEEIG, subdomain I.
2. FACKSIDK, subdomain II.
3. CHRDVKPEN, subdomain V.

Reverse primers 1 and 2 correspond to the sequences:

1. PYYVAPE, subdomain VIII.
2. PFFYGE, subdomain IX.

Reverse primer 3 was an anchored oligo dT primer.

2.7.2 Polymerase chain reaction

Polymerase chain reaction (PCR) was carried out upon DNA and cDNA using a hot start Hybaid PCR machine, Reddy-Mix (dNTP and polymerase mix) from Abgene and primers that were synthesised by MWG-Biotech. The amplification was carried out for a range of cycles (25-40) (a 95°C denaturing cycle for 30 s, an annealing cycle for 45 s (5°C below the predicted hybridisation temperature of the primers), and a 72°C extension cycle for 1 min per kb of DNA. This method was adapted from Sambrook *et al* (1989)

2.7.3 DDrtPCR

PCR was carried out upon cDNA from mature maize leaves that had been harvested in the light and dark. The amplification reaction was carried out over 35 cycles (a 95°C denaturing cycle for 30 s, a 50°C annealing cycle for 45 s, and a 72°C extension cycle for 1 min per kb of DNA.). The primers used were a degenerate forward primer, corresponding to the motif in subdomain V (CHRDIKPEN) in *KfPPCK* and a reverse dT oligonucleotide primer. The amplified fragments were labelled with [α-^{32}P]dCTP by addition to the PCR master mix and run out upon a sequencing gel to aid differentiation. . This method was adapted from Sambrook *et al* (1989)

2.7.4 RACE-PCR

RACE-PCR was carried out using a RACE-PCR kit from Ambion according to the manufacturer's instructions.

2.7.5 RT-PCR

The total RNA samples (100 ng) were mixed with 0.25 µM oligo dT for 5 minutes at 70°C and cooled at 4 °C for 5 minutes. Reverse transcription was carried out in a reaction mixture (50 µl) containing AMV reverse transcriptase buffer (Promega), 1 mM dNTP's (Promega), 1 U. µl^{-1} RNase inhibitor (Promega) and 0.4 U. µl^{-1} AMV reverse transcriptase (Promega). The reaction was performed at 48°C for 45 min. The enzyme was then heat-inactivated at 95°C for 5 min and the samples used directly for PCR.

PCR reactions were performed using 2.5 µl of each cDNA sample in a reaction mixture (25µl) containing 12.5 µl of 2 x Reddy Mix (Abgene) and 0.5 µM of each 5' and 3' primer. The primer sequences and the predicted product sizes are indicated in Table 1. PCR reactions were conducted in a programmable thermocycler (PCR Sprint, Hybaid). The cycles were carried out with an initial denaturation step followed by a hybridisation step and a final extension step. After amplification, the reaction was resolved by electrophoresis on a 1 % agarose gel stained with ethidium bromide. This method was adapted from Sambrook *et al* (1989)

Table 1. Primers used in RT-PCR analyses and cloning

Primer name and specificity	Sequence
Degenerate PPCR	
Universal PPCK-forward 1	TGCGAGGAGATCGGCCGKG
Universal PPCK-forward 2	GGAGGAGTTCGCCGTAAAGTC
Universal PPCK-forward 3	5'-GGCCACCGGGACATCAARCCNGANAA-3'
Universal PPCK-reverse 1	5'-ACCTCCGGCGCCACGTARTAC-3'
Universal PPCK-reverse 2	5'-CTCSCCGTAAAASGGTGG-3'
Anchored poly dT primer	5' TTTTTTTTTTTTTTTTTTVN-3'
RT-PCR	
Le PPCK 1, 5'	5'-TCAAATTTGCGAAGAAATCG-3'
Le PPCK 1, 3'	5' CCTTCTCCTTCTCTCTTCCACA 3'
Le PPCK 2, 5'	5' GAGCTACTCGCCGTCAAGTC 3'
Le PPCK 2, 3'	5'-AAATCAGCCAATTTCAGTTCG-3'
Le Actin 52, 5'	5'-GATGCCCATGTTGGTGACGA-3'
Le Actin 52, 3'	5'-ATCCTCCGATCCAGACACTG-3'
Le Ubiquitin, 5'	5'-GACAACGTCAAGGCTAAGATCC-3'
Le Ubiquitin, 3'	5'-TTACCAGCGAAAATCAACCTCT-3'
St PPCK 2, 5'	5'-GAGCCATTCGCCGTCAAGTC-3'
St PPCK 2, 3'	5'-AAATCAGCCAATTTCAGCTCG-3'
Zm PPCK 2, 5'	5' GCCCAGGCTCTTCGGCGGGGTGT 3'
Zm PPCK 2, 3'	5'-CGTCCCCAAAAACCCGGGGGCAAAGGAA-3'
Zm Actin 1, 5'	5'-TCACACCTTCTACAACGAGC-3'
Zm Actin 1, 3'	5'-AACGTTACCATACAAATCCTTCC-3'
Cloning	
LePPCK2, 5' FL	5'-AGGAATTCGGCACGAGAAA-3'
LePPCK2, 3' FL	5'-GACCAGAGTATACTGGTGCCTGT-3'
StPPCK1, 5'	5' ACGTTCCAAACCTTGAAAAA-3'
StPPCK1, 3'	5'-TCAGTTTAGATCAGCCATTG-3'
pGAD10, 5'	5'-TACCACTACAATGGATG-3'
pGAD10, 3'	5'-CTTGCGGGGTTTTTCAG-3'
GeneRACE, 5'	5'-CGACTGGAGCACGAGGACACTGACATGGACTGAAGGAGTAGAAA-3'

Zm PPCK 2, 5'	5'-GCCCAGGCTCTTCGGCGGGGTGT
Zm PPCK 2, nested, 3'	5'-CTGCAGCTGGCAGCAGCTACCGTACAC T-3'
Zm PPCK 2, 3'	5'-CGTCCCCAAAAACCCGGGGGCAAAGGAA-3'

PCR product sizes with these primers: LePPCK1: 1154 bp from cDNA and 1242 bp from genomic DNA; LePPCK2: 595 bp, 500 bp and 362 bp from cDNA and 595 bp from genomic DNA; LeActin52: 1000 bp from cDNA; LeUbiquitin: 530 bp from cDNA, StPPCK2: 595 bp, 500 bp and 362 bp from cDNA; LePPCK2FL: 1191 bp, 1096 bp and 958 bp from cDNA and 1360 bp from genomic DNA; StPPCK1: 953 bp from genomic DNA; pGAD10: primers specific to vector on either side of multiple cloning site.; ZmActin1: 631 bp from cDNA; ZmPPCK2: 474 bp from genomic DNA and 351 bp from cDNA.

2.8 RNA isolation

2.8.1 Isolation of total RNA

Plant tissue was flash-frozen in liquid nitrogen and stored at –70°C. During total RNA isolation, standard precautions to reduce the risk of RNase contamination were carried out. These precautions included: the use of sterile/autoclaved plasticware; wearing of gloves throughout the protocol; storage of samples on ice to minimise risk of RNA degradation and the treatment of solutions (bar ones containing Tris) with 0.5 % (v/v) diethylpyrocarbonate (DEPC) overnight followed by autoclaving for 20 min to destroy the DEPC. Tris-containing solutions were prepared using autoclaved DEPC-treated distilled water and then autoclaved.

Frozen plant tissue (2-3 g) was ground to a fine powder using an autoclaved mortar and pestle. The fine powder was quickly transferred to a 30 ml plastic centrifuge tube (Sardstedt, U.K.) that held 10 ml of RNA extraction buffer (2 % (w/v) hexadecyltrimethylammonium bromide, 2 % (w/v) polyvinylpyrrollidone 40, 100 mM Tris/HCl, pH 8.0, 25 mM EDTA, 2 M NaCl, 0.5 g/l spermidine-HCl and 2 % (v/v) β-mercaptoethanol (added after autoclaving)), which was pre-warmed in a 65°C water bath. The tube was inverted several times, mixing the ground tissue and RNA extraction buffer, and returned to the 65°C water bath for a minimum of 15 min. Tubes were inverted during the incubation period to ensure thorough mixing. At the end of this period, 10 ml of chloroform was added to the tube which was inverted to mix both phases. The tube was then centrifuged (9500 rpm) for 10 min at 4 °C using a

6 x 50 ml swing-out rotor (JS 13.1) in a Beckman centrifuge (Model J2-21). The upper aqueous phase was removed to a fresh 30 ml centrifuge tube on ice, using a sterile disposable plastic pipette. An equal volume of chloroform was added and the contents of the tube were mixed by inversion. The tube was centrifuged (5000 rpm) for 10 min at 4°C. The upper aqueous phase was removed to a fresh 30 ml centrifuge tube on ice as before. 30 % (v/v) 10 M lithium chloride, 4°C was added to the tube and shaken to mix. The samples was then placed upright at 4°C in a fridge overnight to allow precipitation of RNA. The tube was then centrifuged (9500 rpm) for 20 min at 4°C to pellet the RNA. The supernatant was removed and discarded by using a sterile disposable plastic pipette. The pellet was then dried for 5 min, by inverting the tube onto clean, dry tissues. The RNA pellet was then re-suspended in 500 µl of ice-cold TE and transferred to a 1.5 ml microfuge tube. An equal volume of chloroform was added and mixed, and then the tubes were centrifuged (13000 rpm) for 20 min at 4°C using a 6 x 50 ml swing-out rotor (JS 13.1) in a Beckman centrifuge (Model J2-21). The upper aqueous phase was transferred to a fresh, sterile microfuge tube and 10 % (v/v) 3 M sodium acetate, pH 5.2 and 2.5 volumes of ice-cold ethanol were added and mixed. The RNA was allowed to percipitate for at least 60 min at –70 °C. The tubes was then centrifuged (13000 rpm) for 20 min at 4°C. The supernatant was removed and discarded using a sterile micropipette tip. The RNA pellet was allowed to dry on ice. The dried pellet was resuspended in 50 µl of sterile, DEPC-treated, distilled water and stored at –70°C. This method was adapted from Sambrook *et al* (1989).

The quantity and purity of the RNA was established spectrophotometrically by monitoring the absorbance of the sample at 260 nm and 280 nm. For RNA an A_{260nm} of 1 is equivalent to a concentration of 40 µg/ml. An $A_{260/280}$ ratio of between 1.8 and 2.0 indicates "clean" RNA, relatively free from protein contamination. An $A_{260/280}$ ratio below 1.8 indicates significant protein contamination. Intactness of the RNA was determined by detection of ribosomal RNA bands via agarose gel electrophoresis.

2.8.2 Isolation of poly (A)+ RNA

A poly Atract system (Promega, U.K.) was used to isolate poly (A)+ RNA from total RNA according to the manufacturer's protocols.

2.8.3 Preparation of RNA for RT-PCR

A DNA-free™ system kit (Ambion, U.S.A) was used to ensure elimination of contaminating DNA from RNA preparations, to levels below limits of detection by PCR, according to the manufacturer's protocols.

2.9 *In vitro* translation

A rabbit reticulocyte lysate system (Amersham, U.K.) was used for *in vitro* translation of isolated RNA according to the manufacturer's protocols and Hartwell *et al.* (1996), using Redivue™ [^{35}S]-methionine as the labelled amino acid. Incubations were for 45 min at 30 °C. The incorporation of [^{35}S]-Met into protein was measured by precipitation with trichloroacetic acid according to the manufacturer's protocols to allow standardisation of the subsequent PPCK assays.

2.10 PEPc kinase activity in *in vitro* translation products

The activity of PPCK in the *in vitro* translated products was assayed (Hartwell *et al.*, 1996) by incubating 5 µl of translation products, in a total volume of 25 µl with 50 mM Tris/HCl, pH 7.8 containing 3 mM $MgCl_2$, 1 mM benzamidine hydrochloride, antipain (10 µg ml$^{-1)}$, leupeptin (10 µg ml^{-1}), 5 nM okadaic acid, 0.03 units of purified dephosphorylated *K. fedtschenkoi* PEPc and 10 µCi [γ-^{32}P]ATP. After a 30 min incubation at 30°C, the reaction was halted by adding 10 µl of polyclonal rabbit anti-*K. fedtschenkoi* PEPc antibody (Nimmo *et al.* 1986) and incubating, upon ice, for 1 hour. The immunoprecipitated PEPc was pelleted by centrifugation at 13000 rpm in a microfuge for 10 min. The supernatant was discarded and the pellet was washed in 200 µl of 1.3 M NaCl, 2 mM EDTA (pH 7.0). The PEPc was re-pelleted by centrifugation at 13000 rpm in a bench top microfuge for 10 min, and re-suspended in 30 µl of 5X SDS sample buffer, followed by a 4 min incubation at 100 °C. ^{32}P incorporation into PEPc was measured by SDS-PAGE and autoradiography or phosphorimaging.

The kinase activities derived from different samples were compared by calculating 'relative intensities' of ^{32}P incorporation corrected for translational

efficiency of each sample. ^{32}P incorporation for a sample (corrected for the blank value obtained from a control with no RNA) was divided by the ^{35}S incorporation for that sample (also corrected for the no RNA blank).

2.11 Gel electrophoresis techniques

2.11.1 SDS-polyacrylamide gel electrophoresis (SDS-PAGE) of proteins

Proteins were separated by discontinuous SDS-PAGE according to the method outlined by Laemmli (1970). All SDS-PAGE gels used a 3 % polyacrylamide stacking gel and a separating gel of either 15 % or 8 %. Prior to loading on the gel, samples were denatured by the addition of 1/5 volume 5 x concentration sample buffer (125 mM Tris/HCl, pH 6.8, 2.5% (w/v) SDS, 25 % (v/v) glycerol, 0.05 % bromophenol blue and 2.5 % (v/v) 2-mercaptoethanol). Samples were then heated in a boiling water bath for 4 min. Samples were then loaded onto the gel and electrophoresed at 60 mA for 180 min or until the tracking dye had left the bottom of the separating gel.

2.11.2 Staining SDS-PAGE gels

After separation, the stacking gel was removed and discarded. The separation gel was then transferred to a pre-warmed (37°C) staining solution (0.1 % Coomassie brilliant blue G250, 50 % (v/v) methanol, 10 % (v/v) glacial acetic acid), for 30 min. Gels were then washed in several changes of destaining solution (10 % (v/v) methanol, 10 %(v/v) glacial acetic acid) at 37°C.

2.11.3 Drying, autoradiography and phosphorimaging of SDS-PAGE gels

Processed gels were dried onto Whatman 3MM chromatography paper using a BioRad Laboratories Gel Drier (model 1125) connected to an Aquavac Junior multi-purpose vacuum unit (Uniscience Ltd, U.K.), for 90 min at 65°C.

Dried radioactive gels and blots were analysed by exposing the gel/blot onto X-ray film (Fuji RX) using two intensifying screens at –70 °C for 1-3 days. The film was developed by using a Kodak X-OMAT Processor (model ME-3).

Radioactive gels/blots were also analysed by phosphorimaging using a Fuji Bio-imaging Analyser (Fuji Photo Film Co. Ltd., Japan). Pre-blanked plates were exposed to the gel or blot for 1-24 h in a cassette before imaging. The images were analysed using Mac-Bas software (Fuji Photo Film Co. Ltd., Japan).

2.11.4 Agarose gel electrophoresis of RNA and DNA

Samples of DNA from PCR reactions and digests were analysed for integrity and molecular weight distribution using horizontal agarose gel electrophoresis. 1 % (w/v) agarose gels were prepared and run in 0.5 x TBE (45 mM Tris-borate, pH 8.0, 1 mM EDTA) buffer containing 0.25 μg/ml ethidium bromide. Samples of DNA were mixed with 4 x loading buffer (0.25 % (w/v) bromophenol blue, 0.25 % (w/v) xylene cyanol and 30 % glycerol (v/v) in water). This causes the sample to be dyed and allows the visualisation of loading and running. The gel and loaded samples were electophoresed at 5 V/cm for 1-2 h depending upon the width of the gel. For estimation of the molecular weight of the sample DNA, a DNA ladder (1 kb ladder (Gibco BRL, U.K.) or 100bp ladder (Promega, U.K.)) was run alongside the samples. This method was adapted from Sambrook *et al* (1989).

2.11.5 Denaturing agarose gel electrophoresis of RNA

RNA samples were routinely analysed for integrity and molecular weight distribution by electrophoresis though a denaturing formaldehyde/Mops agarose gel thereby removing secondary structure. This gel was composed of 0.8-1.3 % (w/v) agarose, 1 X MOPS buffer, pH 8.0 (10 X MOPS buffer: 200 mM MOPS, pH 8.0, 50 mM sodium acetate and 10 mM EDTA) and 10% (v/v) formaldehyde. To prepare the gel agarose was dissolved in water, by heating. Upon cooling to ~ 60 °C, the formaldehyde and 10 x MOPS was added and mixed in a fume hood. The gel was poured and allowed to set. It was electophoresed at 5 V/cm for 2-4 hours, in 1 x MOPS, pH 7.0 after loading samples and ladder. This method was adapted from Sambrook *et al* (1989).

RNA samples were prepared by combining 8 μl of RNA (5-20 μg) with 6 μl of sample buffer (72 μg/ml ethidium bromide, 2 x MOPS, pH 8.0, 6 % (v/v) formaldehyde and 70 % (v/v) formamide). The sample was then heated to 65°C for 2.5 min and snap cooled on ice for 5 min. 1-2 μg of RNA size markers (Promega, U.K.) were treated similarly and run alongside the samples to aid molecular weight estimation of bands.

2.12 Northern blotting

2.12.1 Transfer of RNA from denaturing agarose gels onto nitrocellulose membranes

Upon completion of RNA separation on a formaldehyde/MOPS gel, the RNA was transferred onto Hybond N nitro-cellulose membrane. Transfer buffer, (20 x SSC 3 M NaCl and 0.3 M sodium citrate, pH 7.0) was poured into a shallow tray. A platform, large enough to support the gel, was placed within the transfer buffer, and covered with a "wick" made from Whatman 3MM filter paper. On saturation and removal of air bubbles from the wick, the gel was placed face down. A piece of Hybond N nitrocellulose membrane, cut to the size of the gel and pre-wet in distilled water, was positioned on top of the gel. Air bubbles were removed from the membrane and 3 sheets of Whatman 3MM filter paper, pre-wet in transfer buffer and cut similarly to the membrane, were placed on top of the membrane. On the filter paper a stack of paper towels was placed and secured by a 1 kg weight. Finally, cling film was cut to surround the gel, preventing the transfer buffer from "short circuiting" around the gel. Transfer was then allowed to proceed overnight.

After transfer, the blot was dismantled. The membrane was briefly dipped in distilled water to remove any excess salt deposits left by the transfer buffer. The blot was then allowed to dry in the air before crosslinking the RNA to the membrane. This was carried out using a UV crosslinker (UVP Ultraviolet Crosslinker) by exposing the membrane to a set dose (120000 $\mu J\ cm^{-2}$) of UV irradiation. The blot was stored between two pieces of Whatman 3MM paper, wrapped in aluminium foil. This method was adapted from Sambrook *et al* (1989).

2.12.2 Hybridisation of radiolabelled cDNA probes to Northern blots

All hybridisation steps were performed in a hybridisation oven (Hybaid). Before addition of a cDNA probe, the blots were prehybridised at 55 °C for 2 h in Northern pre-hybridisation buffer (0.5 M Na_2HPO_4, pH 7.2, 7 % (w/v) SDS, 10 mg/ml BSA). A radioactive probe was then added to the pre-hybridisation buffer, and allowed to hybridise overnight at 55°C. The buffer and probe was then poured from the blot and retained for further use. The blot was then washed three times in fresh 2xSSC/1 % SDS at 55°C for 10 min. Further washes, of increasing stringency, were used, depending on the "signal to background" ratio. In increasing stringency order

these washes were: 2 x SSC/1 % SDS at 65°C for 10 min, 1 x SSC/1 % SDS at 65°C for 10 min, 0.5 x SSC/1 % SDS at 65°C, 0.5 x SSC/0.1 % SDS at 65°C for 10 min. The blot was then heat sealed into plastic and exposed to X-ray film or phosphorimaged. This method was adapted from Sambrook *et al* (1989).

2.13 Southern blotting

2.13.1 Transfer of DNA from agarose gels onto nitro-cellulose membranes

Upon completion of DNA separation on an agarose gel, the gel was denatured and neutralized to facilitate DNA transfer. This was carried out by placing the agarose gel in denaturation buffer (1.5 M NaCl, 0.5 M NaOH) for 30 min with gentle agitation at room temperature. This was followed by a second incubation in neutralization buffer (1 M Tris/HCl, 1.5 M NaCl, pH 7.4) for 30 min with gentle agitation at room temperature. A final incubation in 10 x SSC for 30 min with gentle agitation at room temperature removed any neutralization buffer. The DNA, from the gel, was transferred onto Hybond N nitrocellulose membrane, via blotting. with the transfer buffer (20 X SSC, 3 M NaCl and 0.3 M sodium citrate, pH 7.0) which was poured into a shallow tray. A platform, large enough to support the gel, was placed within the transfer buffer, and covered with a "wick", made from Whatman 3MM filter paper. Upon saturation and removal of air bubbles from the wick, the gel was placed face down on the paper. A piece of Hybond N nitrocellulose membrane, cut to the size of the gel and pre-wet in distilled water, was positioned on top of the gel. Air bubbles were removed from the membrane and 3 sheets of Whatman 3MM filter paper, pre-wet from transfer buffer and cut similarly to the membrane, were placed on top of the membrane. On the filter paper a stack of paper towels was placed and secured by a 1 kg weight. Finally, cling film was cut to surround the gel, preventing the transfer buffer from "short circuiting" around the gel. Transfer was then allowed to proceed overnight.

After transfer, the blot was dismantled. The membrane was briefly dipped in distilled water to remove any excess salt deposits left by the transfer buffer. The blot was then allowed to dry in the air before crosslinking the DNA to the membrane. This was carried out using a UV crosslinker (UVP Ultraviolet Crosslinker) by exposing the membrane to a set dose (120000 $\mu J\ cm^{-2}$) of UV irradiation. The blot was stored between two pieces of Whatman 3MM paper, wrapped in aluminium foil. This method was adapted from Sambrook *et al* (1989).

2.13.2 Transfer of phage DNA from LB-agar plates onto nitrocellulose membranes

Nitrocellulose filters were gently placed upon the top of plaque holding LB-agar/TB top agar plates, and the orientation of the filter to the plate was recorded. After 10 min at room temperature, the filters were gently removed and allowed to dry at room temperature, for 20 min, while the plates were stored at 4°C. 3 trays containing Whatman 3MM paper were saturated with one of three buffers:

Buffer A (0.2M NaOH, 1.5M NaCl)

Buffer B (0.4M Tris/HCl, pH 7.6, 2 x SSC)

Buffer C (2 x SSC).

Filters were placed (plaque side up) on each saturated Whatman 3MM paper for 3 min, in sequence. The filter was then briefly dipped in distilled water to remove any excess salt deposits left by the transfer buffer. The blot was then allowed to dry in the air before crosslinking the phage DNA to the membrane. This was carried out using a UV crosslinker (UVP Ultraviolet Crosslinker) by exposing the membrane to a set dose (120000 $\mu J\ cm^{-2}$) of UV irradiation. The filter was stored between two pieces of Whatman 3MM paper, wrapped in aluminium foil. This method was adapted from Sambrook *et al* (1989).

2.13.3 Hybridisation of radiolabelled DNA probes to Southern blots

All hybridisation steps were performed in a hybridisation oven (Hybaid). Before addition of a gDNA/cDNA probe, the blots were prehybridised at 55°C for 2 h in Southern pre-hybridisation buffer (5 x Denhardts reagent, 6 x SSC, 0.5 % SDS, 100 $\mu g\ ml^{-1}$ of denatured fragmented salmon sperm DNA, 50 % formamide). A radioactive probe was then added to the pre-hybridisation buffer, and allowed to hybridise overnight at 42°C. The buffer and probe was then poured from the blot and retained for further use. The blot was then washed three times in fresh 2 x SSC/1 % SDS at 55°C for 10 min. Further washes of increasing stringency were used depending on the "signal to background" ratio. In increasing stringency order these washes were: 2 x SSC/1 % SDS at 65°C for 10 min, 1 x SSC/1 % SDS at 65°C for 10 min, 0.5 x SSC/1 % SDS at 65°C for 10 min. The blot was then heat sealed into

plastic and exposed to X-ray film or phosphorimaged. This method was adapted from Sambrook *et al* (1989).

2.13 Cloning DNA into plasmids and culture and transformation of *Eschericia coli*

2.14.1 DNA cloning

PCR amplified DNA inserts were cloned into pCR4-TOPO vectors and transformed into *E. coli* cells, by using a Topo TA Cloning Kit (Invitrogen Life Sciences), according to the manufacturer's protocols.

2.14.2 Culture of *E. coli* strains possessing plasmids containing cloned cDNA/gDNA inserts

As outlined in Sambrook *et al.* (1989), transformed cells (10 µl and 90 µl) were plated out onto LB broth plates (1% (w/v) Bactotryptone, 0.5% (w/v) Bactoyeast extract, 1% (w/v) NaCl, 1.5% (w/v) Bactoagar, pH 7.5) containing 100 g/ml ampicillin or/and 100 g/ml kanamycin (added after broth had cooled to ~ 45°C), and then incubated overnight at 37°C.

Single colonies were "picked" from the selection plates, after overnight growth, by using a sterile yellow micropipette tip. The colony holding tip was ejected into 10 ml of LB broth (1% (w/v) Bactotryptone, 0.5% (w/v) Bactoyeast extract, 1% (w/v) NaCL, pH 7.5) containing 100g/ml ampicillin or/and 100g/ml kanamycin. The culture was allowed to grow overnight in a shaking incubator (at 37 °C and 200 rpm). Glycerol stocks were made by mixing 700 µl of the overnight culture with 300 µl of sterile glycerol. The mixture was vortexed and flash-frozen in liquid nitrogen before storage at –70°C. This method was adapted from Sambrook *et al* (1989).

2.14.3 Sequencing

DNA vector inserts were sequenced by MWG-biotech from the T7 and T3 priming sites, after purification.

2.15 DNA fragment isolation and generation of radiolabelled cDNA probes for Northern and Southern blotting

2.15.1 Excision of DNA inserts from isolated plasmids

Restriction digests of purified plasmids were carried out to excise DNA fragments. Restriction endonucleases were chosen to cut a unique site flanking either side of the insert. Digests were carried out according to the restriction endonuclease

manufacturer's protocols. Digested DNA was then run out upon a 1 % (w/v) agarose gel, separating the insert from linearized plasmid. A sterile scalpel was then used to excise the desired DNA fragment from the gel, under UV visualisation. The DNA was extracted from the gel-slice by using a QIA-quick gel extraction kit according to the manufacturer's protocols.

2.15.2 Random priming of DNA to synthesize radioactive probes

DNA inserts were radiolabelled for use as probes on Northern and Southern blots by random priming, using the Klenow fragment of DNA polymerase I. DNA (20 ng) was labelled with the Rediprime DNA labelling system (Amersham, U.K.) in the presence of 50 μCi of Redivue [α-^{32}P]dCTP (Amersham, U.K), according to the manufacturer's protocols. After synthesis the probe was denatured by boiling for 5 min followed by chilling on ice, prior to addition to the hybridisation solution.

2.16 *In vitro* transcription

In vitro transcription was performed using a T3 or T7 mMessage mMachine *in vitro* transcription kit, allowing synthesis of capped RNAs according to the manufacturer's protocol. The template DNA, a plasmid, was linearized by performing a restriction endonuclese digestion in the multiple cloning site after the 3' end of the insert. The freshly digested DNA was purified using the Wizard DNA clean up kit (Promega, U.K.) before use as template in the *in vitro* transcription reaction.

Chapter 3

The identification of maize and wheat phosphoenolpyruvate carboxylase kinase cDNAs

3.1 Introduction

PEPc catalyses the fixation of CO_2 into OAA. In the leaves of C_4 plants, this step is carried out in the mesophyll cell and is central to the accumulation of atmospheric CO_2. PEPc plays a range of functions in other tissues. To carry out these roles, higher plants contain a number of PEPc genes characterised as 'C_3' or 'C_4' forms. The 'C_4' forms of PEPc are involved in photosynthesis and may be distinguished from the anapleurotic and specialised forms ('C_3') of PEPc on the basis of sequence and expression. All known forms are regulated by reversible phosphorylation of a single N-terminal serine residue. The phosphorylation state of PEPc is largely controlled by the activity of Ca^{2+}-independent PPCK (see Chapter 1). Several Ca^{2+}-independent PPCKs have been identified recently, and it has been suggested that most plants contain a small family of PPCK genes. All these forms appear to be of the smaller Ca^{2+}-independent isoform (~ 30 kDa).

At the time I started this work (October 1998) no PPCK gene or cDNA had been identified in a C_4 plant. My objective was to examine PPCK genes in maize, and identify and characterise the gene(s) corresponding to the light-induced PPCK activity that had already been studied by biochemical techniques (for a review see Vidal and Chollet, 1997). Several direct approaches proved unsuccessful, so I reverted to the indirect database-mining approach. In the course of this I also assessed PPCK ests from other monocots including wheat.

3.2 Results

3.2.1 The amplification of maize gDNA and cDNA using primers designed to conserved domains of *K. fedtschenkoi PPCK* failed to produce an identifiable PPCK sequence

Following the sequencing of the first full-length PPCK from *K. fedtschenkoi* (Hartwell *et al.*, 1999) a number of primers were designed to conserved domains. Preliminary PCR experiments on gDNA from a number of plant species using these primers amplified a number of products that showed the expected size of ~500 bp. The cloning and sequencing of several of these products revealed that they bore high similarity to *K. fedtschenkoi PPCK* (*KfPPCK*), they were therefore thought to be

partial PPCK sequences. These included a rice partial PPCK sequence [GenBank accession number: AF203484]; a banana partial PPCK sequence [GenBank accession number: AF203483], a oilseed rape partial PPCK [GenBank accession number: AF203482] and a tomato partial PPCK sequence [GenBank accession number: AF203481, now termed *LePPCK1*, see Chapter 4]. A similar approach was used in an attempt to identifying a maize PPCK sequence. Primers were designed to the conserved domains of *KfPPCK*, see Figure 3.1, and were used in PCR experiments on gDNA from maize. Products of the expected size were cloned and sequenced. These PCR products showed no similarity to *KfPPCK*, and were thought to be intergenic gDNA, which is present in the maize genome in large amounts. An example of an isolated partial maize gDNA sequence is shown in Appendix A. This sequence has no significant homology to any known sequence.

To enrich for PPCK sequences, I then used RT-PCR. Previous work had shown that light increases PPCK activity and translatable mRNA. I therefore isolated total RNA from maize leaves in the light and the dark. Figure 3.2 shows that the abundance of PPCK translatable mRNA was much greater in the "light" RNA than the "dark" RNA. Several cDNA products from the "light" RNA were cloned and sequenced. These sequences have no significant homology to KfPPCK. An example of an isolated maize light induced cDNA sequence is shown in Appendix B. This sequence is identical to *Zea mays* calcium-dependent kinase (CDPK) 1. Given the similarity between PPCKs and CDPKs at the aminoacid sequence level, this result is perhaps, not surprising.

A further refinement of this approach involved the use of differential display reverse transcriptase PCR (DDrtPCR). Light and dark cDNA's were labelled with [α-^{32}P]dCTP, run out on a sequencing gel and visualized by phosphorimaging (Figure 3.3). Light induced products with the expected size were extracted, re-amplified, cloned and sequenced. The sequences showed that the products were not similar to *KfPPCK*. An example of an isolated maize light induced cDNA sequence is shown in Appendix C. This sequence bears similarity to plant serine/threonine protein kinases, specifically *Glycine max* SPK-3, and it is therefore thought to be a maize homologue of this gene.

3.2.2 The heterologous probing of maize mRNA with a partial rice PPCK sequence yielded a light-induced 2 kb band

In another approach, a Northern blot of maize RNA was probed with the partial PPCK clone from rice mentioned above. Figure 3.4 shows that light increases the abundance of a 2 kb band that is recognised by the probe. Full length cDNA sequences of *A. thaliana* and *K. fedtschenkoi* had been previously identified within our lab (Hartwell *et al.*, 1999; Fontaine *et al.*, 2000). These sequences, when transcribed, produce a transcript of ~2000 bp. This suggested that the partial rice PPCK sequence may hybridise to a light-induced maize PPCK message. In an attempt to identify a maize PPCK clone, a maize seedling cDNA library was screened using this rice sequence.

3.2.3 The heterologous probing of a maize cDNA library with a partial rice PPCK sequence fails to identify a maize PPCK clone

The maize cDNA seedling phage library was screened with the rice PPCK sequence at low stringency. Six plaques hybridised to this sequence. These plaques were isolated and the lambda vectors, containing maize cDNA inserts were purified (as described in section 2.6.3). The cDNA inserts were digested out of the vector, resolved by electrophoresis and blotted onto nitrocellulose membranes. The Southern blot (Figure 3.5) shows that a maize cDNA insert of 2 kb hybridised to the rice PPCK sequence. The sequenced clones showed no similarity to *KfPPCK* or *AtPPCK1* or *AtPPCK2* sequences. This suggests that the stringency of the screening was too low to discriminate PPCK cDNA's. Alternatively it may mean that transcript levels of maize PPCK in seedlings is very low. An example of a sequence of an isolated clone is shown in Appendix E. This sequence is similar to the 6.1 kDa polypeptide of photosystem II. Clustal analysis reveals a 25 % degree of similarity between the two nucleotide sequences, overall. However, a 100 bp region has much higher similarity (65 %). As the probe was randomly labelled, rather than end-labelled, this could indicate why this cDNA was selected.

3.2.4 The identification and characterisation of a putative maize PPCK est

During 2000, Dr Veronique Fontaine in our lab obtained and sequenced a PPCK clone from *Sorghum* [GenBank number: , unpublished data]. In early 2001, an est from the maize database was identified during a routine database search in our lab as having high similarity to the 3'-end of this *Sorghum PPCK*. The maize est [GenBank accession number: BG3321123], appears to retain the intron that has been

found close to the 3' end of the coding sequences in other PPCK genomic sequences (Fontaine *et al.*, 2002, Hartwell *et al.*, 1999). Figure 3.6 shows the translation map of est BG3321123 with the putative PPCK coding regions shown in blue. One frame shift has to be introduced to position the completely conserved DFG sequence correctly.

I therefore investigated whether this sequence was present in the cDNA of maize. Preliminary experiments showed that primers designed to the est amplified a cDNA product that was approximately 100 bp smaller than the corresponding gDNA product. This suggested that the putative intron could be correctly spliced but that the est represented an unspliced PPCK transcript. A full-length cDNA clone was obtained by RACE-PCR as described in section 2.10, and fully sequenced. From this sequence a 5'-end primer was designed allowing the amplification of a full-length gDNA clone, revealing a single intron matching the intron within BG321123 and the cDNA sequence. Figure 3.7 shows the translation map of the full genomic sequence of this maize PPCK and a schematic of the gene structure. This sequence has now been named *ZmPPCK2* (see section 3.3). Unusually this sequence has a 35 bp insertion in the middle of the coding region (from position 454 to position 489), encoding highly acidic aminoacids. A similar insertion is seen in the *Sorghum PPCK* gene (*SbPPCK*) mentioned above.

A sequence alignment of *ZmPPCK2* to several known PPCKs is shown in Figure 3.8. This alignment reveals a strong similarity between SbPPCK and ZmPPCK2 exists. A 88 % identity exists between these two aminoacid sequences while a 89% identity exists between the nucleotide sequences. As noted above, ZmPPCK2 and SbPPCK both hold a unique highly acidic extension within the middle of the sequence, giving the predicted molecular weight of ZmPPCK2 at 31.5 kDa in comparison to AtPPCK1 31.8 kDa, although the length of ZmPPCK2 is 298 residues compared to AtPPCK1 of 284 residues. This reflects the fact that ZmPPCK2 has a much higher content of glycine and alanine, but a lower content of phenylalanine and tyrosine , than AtPPCK1.

3.2.5 Semi-quantitative RT-PCR analysis of a maize PPCK reveals it is not induced by light treatment

The level of PPCK translatable mRNA is significantly higher in illuminated then darkened maize leaves (Hartwell *et al.*, 1996, see also Figure 3.4). Furthermore a *F. trinervia PPCK* (*FtPPCK*) has recently been shown to be induced by light (Saze *et*

al., 2001). RT-PCR was therefore used to investigate whether the level of ZmPPCK2 transcripts was increased in the light. Figure 3.9 shows the results of these experiments. The data shows maize PPCK2 is not light induced, as similar band intensities for the light and dark cDNA are observed. Interestingly, a second larger transcript can be seen, approximately 100 bp larger than the correctly spliced variant. Since the primers spanned the intron, this form is probably a transcript variant that retains the intron, similar to the est BG321123. At lower cycle numbers, no products were detectable from either the light or dark sample.

To test whether the maize *PPCK2* gene does indeed encode PPCK activity, the cDNA clone was transcribed and translated and the translation products were assayed for PPCK activity as described in 2.9 and 2.10 (see also Figure 4.3 for a similar experiment in tomato PPCK clones). This experiment confirmed that the clone encodes PPCK activity (not shown). Unfortunately the data was lost in the Bower Fire on 24/10/01. This fire resulted in the loss of all our equipment and most of our anti-PEPc antibodies, which are required for this experiment. Hence it has not been possible to repeat the experiment as yet.

3.2.6 The identification of a putative wheat PPCK est

A wheat est [GenBank accession number: BE517315] was identified as a putative PPCK sequence in a database search. I obtained the est and sub-cloned the insert into an expression vector (pCR-TOPO). Sequencing analysis (see Figure 3.10) and PPCK alignment comparisons, see figure 3.8, reveal that it is a full-length cDNA clone of PPCK, similar to a group of maize ests, termed ZmPPCK1 (see section 3.3). We have thus termed this wheat kinase *Triticum aestivum PPCK1* (*TaPPCK1*). The amino acid sequences of ZmPPCK1 and TaPPCK1 have a 73 % identity, while their nucleotide identity is 70 %.

3.3 Discussion

One of my original objectives was to clone and analyse maize PPCK genes. Several attempts failed. However the approaches used were not flawed, similar approaches used on *Flaveria* and *Mesembryantheum* were successful in identifying PPCK (Saze *et al.*, 2001, Taybi *et al.*, 2000). At a late stage in the project I did succeed in getting one maize clone by database mining. This is the first PPCK gene characterised from maize. However like the *Sorghum* homologue it is not light up-regulated. Presumably it has a different function from FtPPCK and KfPPCK even though it is expressed in leaves.

In late 2001, I identified further ests and genomic sequences in the maize database that resembled PPCK. While several of the ests did not overlap with any others, further analysis of the maize gene family by Dr M. Shenton in our lab has clarified the position. Table 2 shows the assignment of ests to three separate genes termed *PPCK1*, *2* and *3*. Our lab has begun to characterise the nature of all three genes by studying their tissue specificity and control of expression. Full length cDNA cloning and validation of the PPCK activity of ZmPPCK1 have already been accomplished, although ZmPPCK3 is not yet fully sequenced. Recent results from expression analysis by real time RT-PCR of the two other putative maize PPCKs suggest that ZmPPCK1 transcript levels are upregulated in light in mature leaves by a factor of 16, in comparison to mature leaves from the dark (Shenton *et al.*, unpublished). This level of increase agrees with the maize data published by Hartwell *et al.* (1999) that showed translated illuminated leaf mRNA produced a seven fold increase of PPCK activity to that of translated dark leaf mRNA. There are few clues to the function of either of these maize kinases as yet. However one unexplained observation is that *Sorghum* PPCK, a homologue of *ZmPPCK2*, is strongly upregulated by cycloheximide. Another point of interest is the absence of cysteine residues in the positions suggested by Saze *et al.* (2001) in the predicted aminoacid sequences of maize PPCK1 and 2. However, four cysteine residues are present in the PPCK sequences, allowing the possibility of thioredoxin regulation.

Table 2 Maize ests and genomic fragments corresponding to putative PPCKs

Identity of maize PPCK	GenBank number	Source
Zm PPCK 1	BM073675	Seedling and silk
	BM379295	Unpollinated first ears
	BQ034770	Immature ear
	BQ048400	Immature ear
	BQ529100	Immature ear
Zm PPCK 2	BG321132	Leaf, crown
Zm PPCK 3	BH414995	Genomic DNA
	BH414996	Genomic DNA
	BH623642	Genomic DNA
	BH629101	Genomic DNA
	BH630546	Genomic DNA
	BH630622	Genomic DNA
	BH631926	Genomic DNA
	BH631990	Genomic DNA

	BH631991	Genomic DNA
	BH63215	Genomic DNA

ZmPPCK2 and SbPPCK show a unique extension within the middle of their coding region The role for this highly acidic region is unknown, but it could be involved in solubility or recognition by unknown factors. This region is not likely to affect catalysis. Structure prediction suggests it is located at the surface on the opposite side of the molecule to the active site.

A further interesting observation that this data reveals is the highly biased GC-rich nature of the maize, wheat and *Sorghum* coding sequences. Analysis of known PPCK coding regions show that these sequences are non-biased (see Chapter 5). Following the sequencing of the rice genome, Yu *et al.* (2002) noted that coding regions are nearly all less than 70 % in GC content. The coding regions of *ZmPPCK2*, *ZmPPCK1*, *TaPPCK1* and *SbPPCK* show a GC content of 70 % or more, which is much greater than the normal population of coding regions of monocots. This data also explains the failure of the PCR approach - presumably the nucleotide sequence similarity was not enough for this to work. The heterologous probing should have worked but the cDNA library may have been harvested from seedlings illuminated at a too low light intensity to contain much PPCK mRNA. The data obtained from wheat reveal that TaPPCK and ZmPPCK are highly similar. This will be further discussed in chapter 5.. My early experiments that focused on identifying a maize PPCK by a direct approach failed. Similar approaches used on *Flaveria* and *Mesembryantheum* were successful in identifying PPCK (Saze *et al.*, 2001, Taybi *et al.*, 2000).

Figure 3.1 The amino acid sequence of KfPPCK and universal PPCK primers designed to conserved regions

Arrows indicate relative position of primers and orientation.

This sequence is taken from Hartwell *et al.* (1999).

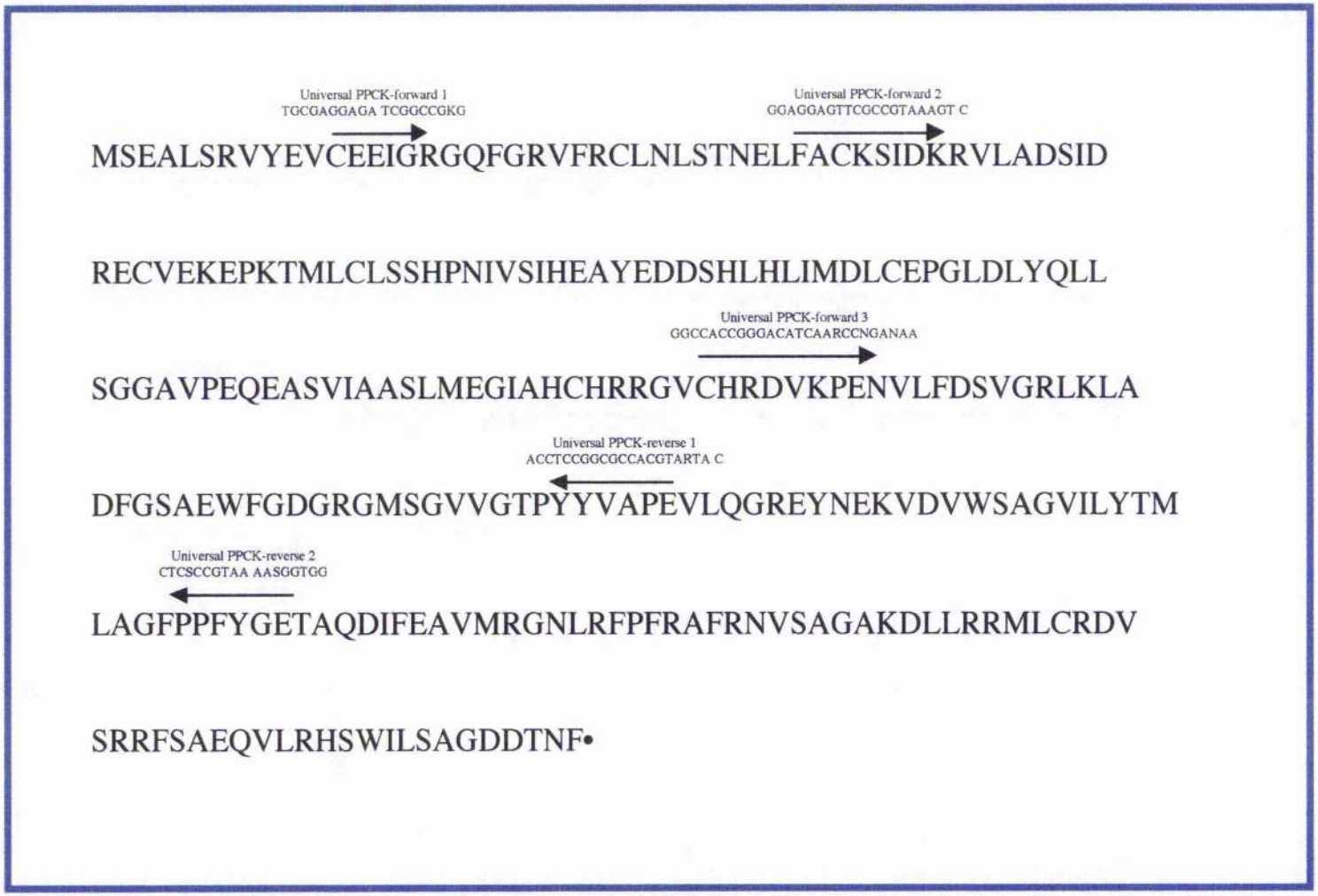

Universal PPCK-forward 1
TGCGAGGAGA TCGGCCGKG
Universal PPCK-forward 2
GGAGGAGTTCGCCGTAAAGT C
MSEALSRVYEVCEEIGRGQFGRVFRCLNLSTNELFACKSIDKRVLADSID
RECVEKEPKTMLCLSSHPNIVSIHEAYEDDSHLHLIMDLCEPGLDLYQLL
Universal PPCK-forward 3
GGCCACCGGGACATCAARCCNGANAA
SGGAVPEQEASVIAASLMEGIAHCHRRGVCHRDVKPENVLFDSVGRLKLA
Universal PPCK-reverse 1
ACCTCCGGCGCCACGTARTA C
DFGSAEWFGDGRGMSGVVGTPYYVAPEVLQGREYNEKVDVWSAGVILYTM
Universal PPCK-reverse 2
CTCSCCGTAA AASGGTGG
LAGFPPFYGETAQDIFEAVMRGNLRFPFRAFRNVSAGAKDLLRRMLCRDV
SRRFSAEQVLRHSWILSAGDDTNF•

Figure 3.2 Effect of light on the amount of PPCK translatable mRNA in maize leaves

Poly A^+ RNA was isolated and translated as described in 2.8.2 - 2.9.

The panel shows incorporation of ^{32}P into PEPc in assays of the translation products

Lane (1) 300 ng of poly A^+ RNA from illuminated maize leaves.

Lane (2) 300 ng of poly A^+ RNA from darkened maize leaves.

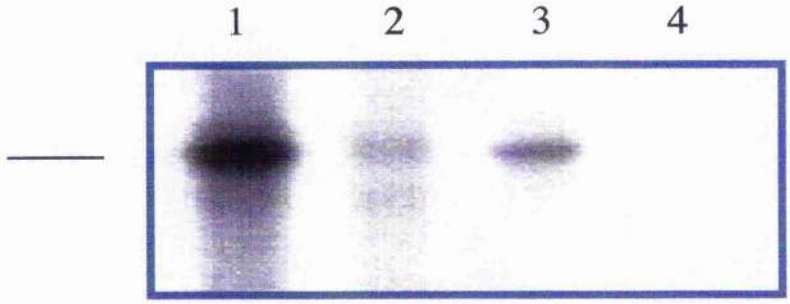
Lane (3) 300 ng of poly A^+ RNA from darkened *K. fedtschenkoi* leaves harvested at 24:00 h (positive control).

Lane (4) no RNA blank

The figures below each lane indicate relative intensity as a percentage of lane 3.

The bar shows immunoprecipitated ^{32}P-labelled PEPc.

This experiments repeats one shown in Hartwell *et al.* (1999).



	1	2	3	4
Relative intensity (%)	210	48	100	0

Figure 3.3 The identification of differentially displayed RT-PCR products from maize leaf mRNA

The RT-PCR products were labelled with [α-^{32}P]dCTP and resolved on a sequencing gel. The figure shows a phosphorimage of the gel. The products were obtained from:

Lane (1): 500 ng of poly A^+ RNA from illuminated maize leaves.

Lane (2): 500 ng of poly A^+ RNA from darkened maize leaves.

Lane (3): 250 ng of poly A^+ RNA from illuminated maize leaves.

Lane (4): 250 ng of poly A^+ RNA from darkened maize leaves.

Lane (5): Water control.

Arrows show extracted bands.

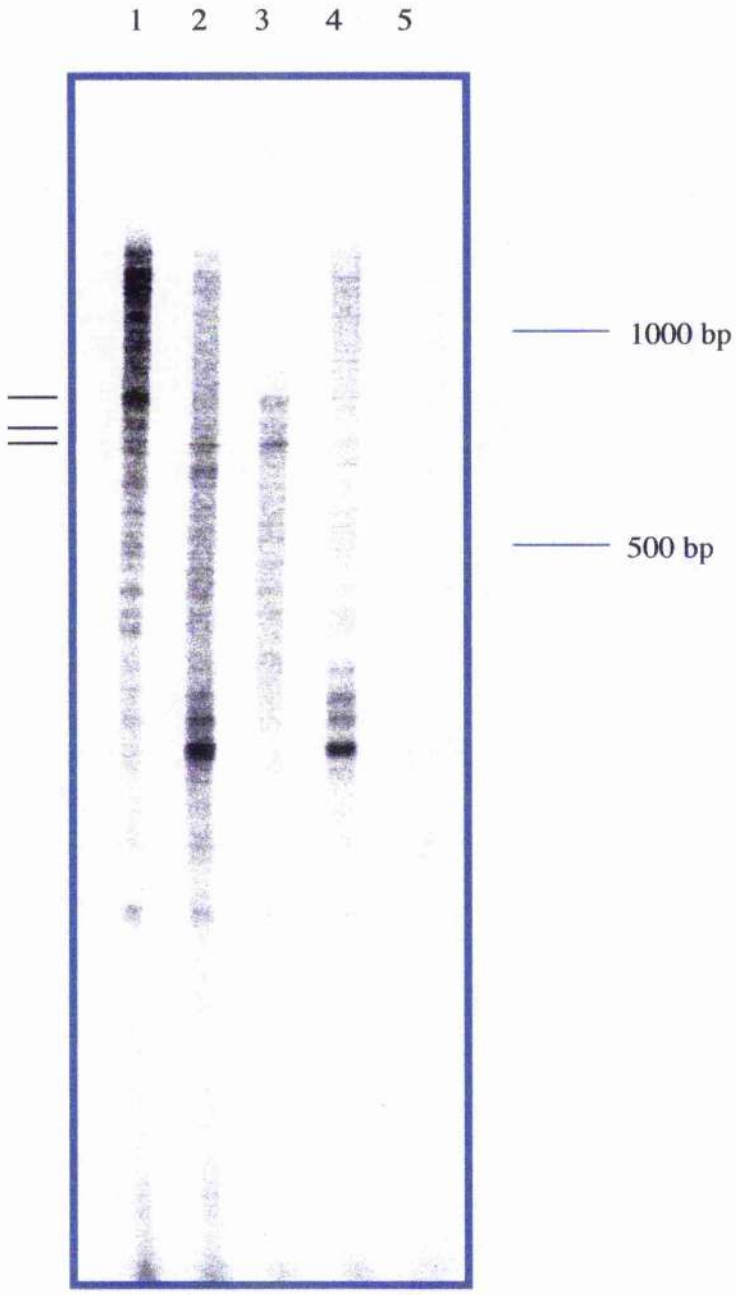

1
2
3
4
5
1000 bp
500 bp

Figure 3.4 Northern analysis of maize RNA

The blot was probed with a partial PPCK cDNA from rice. Methods are described in 2.12. The indicated band is approximately 2000 bp.

Lane (1): 300 ng of poly A^+ RNA from illuminated maize leaves.

Lane (2): 600 ng of poly A^+ RNA from illuminated maize leaves.

Lane (3): 300 ng of poly A^+ RNA from darkened maize leaves.

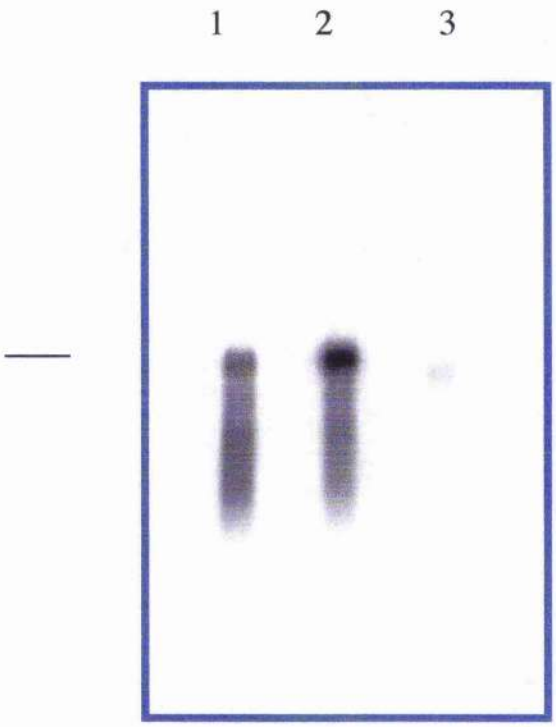

1
2
3

Figure 3.5 Southern analysis of putative maize PPCK cDNA clones

Six plaque colonies were detected by library screening. The maize inserts from the lambda phage of the plaques were digested and resolved upon a 1 % agarose gel. This gel was blotted onto nitrocellulose paper and probed with a partial PPCK cDNA from rice. Methods are described in 2.6.3, 2.11, 2.13 and 2.15. Lane (5) shows a cDNA insert that hybridised to the probe. The indicated band is approximately 700 bp.

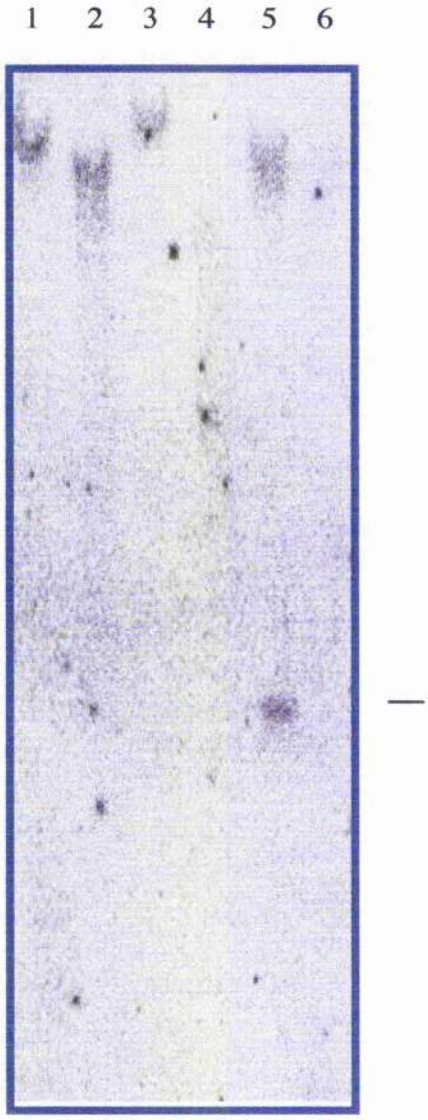
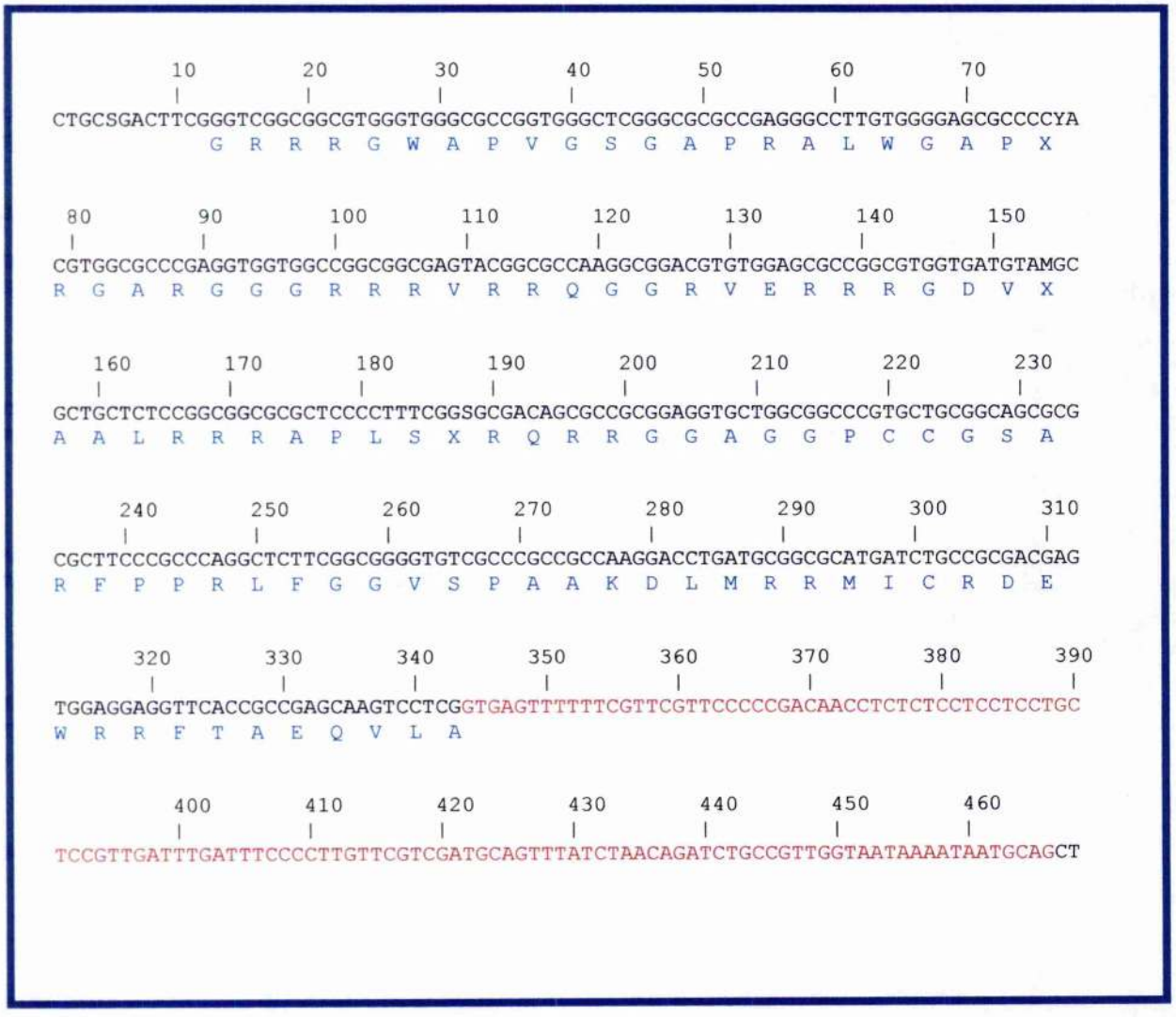
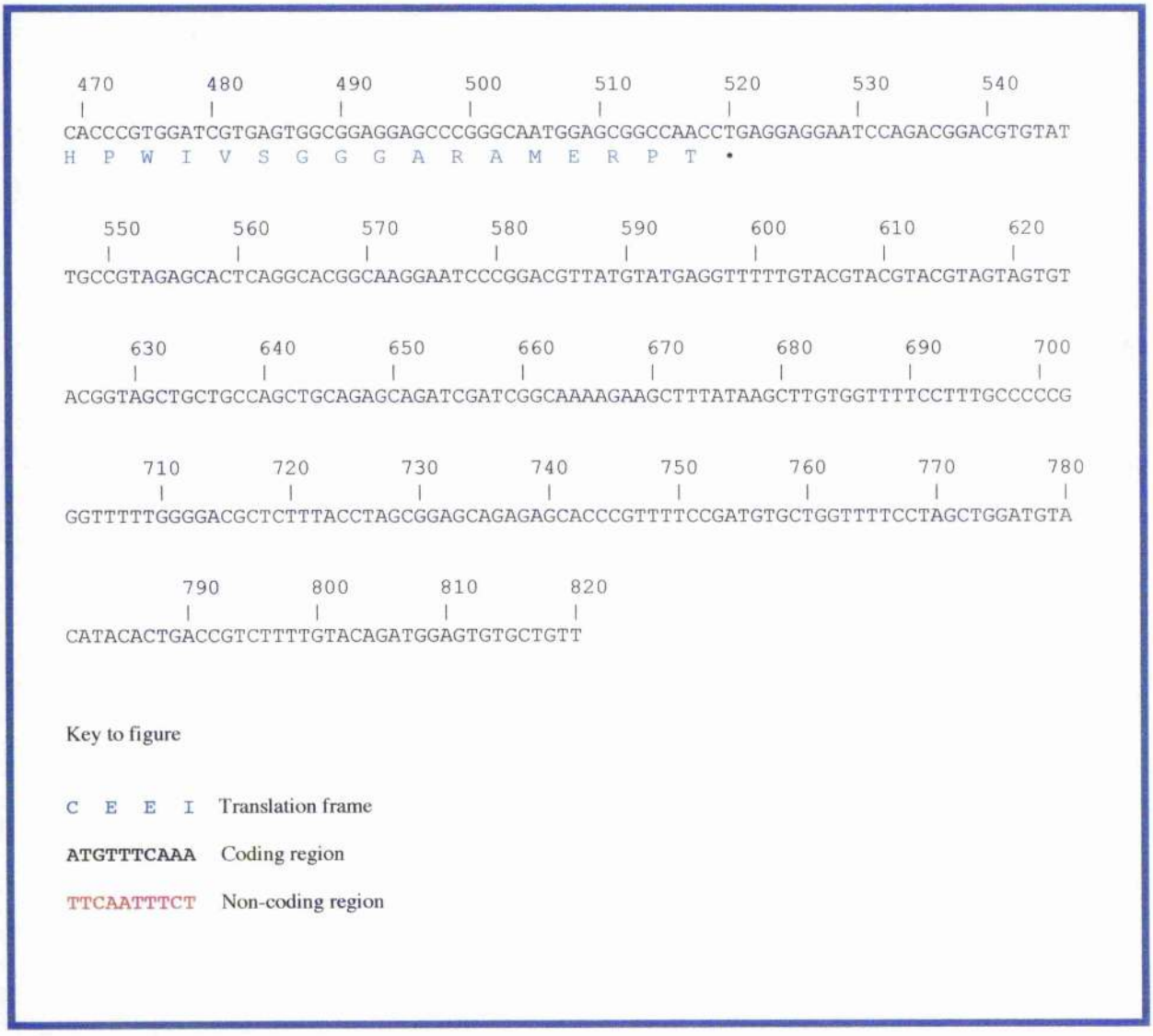

1 2 3 4 5 6

Figure 3.6 Translation map of maize est BG321123

The nucleotide sequence in red indicates the putative intron. The reading frame of the putative PPCK aminoacid sequence is in blue. The red font indicates the intron/non coding region.


10 20 30 40 50 60 70
CTGCSGACTTCGGGTCGGCGGCGTGGGTGGGCGCCGGTGGGCTCGGGCGCGCCGAGGGCCTTGTGGGGAGCGCCCCYA
G R R R G W A P V G S G A P R A L W G A P X
80 90 100 110 120 130 140 150
CGTGGCGCCCGAGGTGGTGGCCGGCGGCGAGTACGGCGCCAAGGCGGACGTGTGGAGCGCCGGCGTGGTGATGTAMGC
R G A R G G G R R R V R R Q G G R V E R R R G D V X
160 170 180 190 200 210 220 230
GCTGCTCTCCGGCGGCGCGCTCCCCTTTCGGSGCGACAGCGCCGCGGAGGTGCTGGCGGCCCGTGCTGCGGCAGCGCG
A A L R R R A P L S X R Q R R G G A G G P C C G S A
240 250 260 270 280 290 300 310
CGCTTCCCGCCCAGGCTCTTCGGCGGGGTGTCGCCCGCCGCCAAGGACCTGATGCGGCGCATGATCTGCCGCGACGAG
R F P P R L F G G V S P A A K D L M R R M I C R D E
320 330 340 350 360 370 380 390
TGGAGGAGGTTCACCGCCGAGCAAGTCCTCGGTGAGTTTTTTCGTTCGTTCCCCCGACAACCTCTCTCCTCCTCCTGC
W R R F T A E Q V L A
400 410 420 430 440 450 460
TCCGTTGATTTGATTTCCCCTTGTTCGTCGATGCAGTTTATCTAACAGATCTGCCGTTGGTAATAAAATAATGCAGCT



Key to figure

C E E I Translation frame

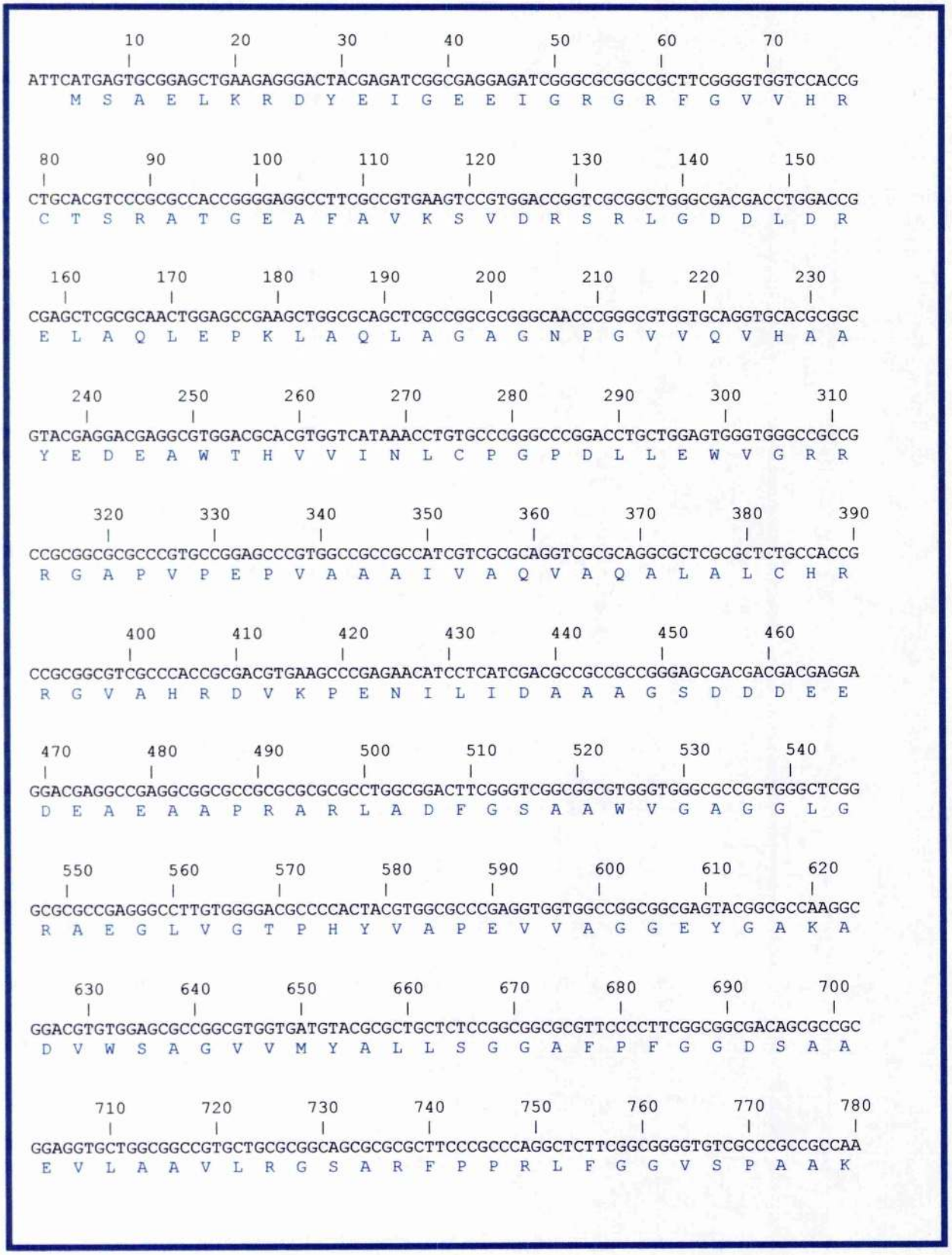
ATGTTTCAAA Coding region

TTCAATTTCT Non-coding region

Figure 3.7 Translation map of the genomic sequence of *ZmPPCK2* and a schematic of the gene structure

Panel A. shows the coding and non-coding region of *ZmPPCK2* and the correct reading frame of ZmPPCK2.

Panel B. shows a schematic of the gene structure of *ZmPPCK2*.


10 20 30 40 50 60 70
ATTCATGAGTGCGGAGCTGAAGAGGGACTACGAGATCGGCGAGGAGATCGGGCGCGGCCGCTTCGGGGTGGTCCACCG
M S A E L K R D Y E I G E E I G R G R F G V V H R
80 90 100 110 120 130 140 150
CTGCACGTCCCGCGCCACCGGGGAGGCCTTCGCCGTGAAGTCCGTGGACCGGTCGCGGCTGGGCGACGACCTGGACCG
C T S R A T G E A F A V K S V D R S R L G D D L D R
160 170 180 190 200 210 220 230
CGAGCTCGCGCAACTGGAGCCGAAGCTGGCGCAGCTCGCCGGCGCGGGCAACCCGGGCGTGGTGCAGGTGCACGCGGC
E L A Q L E P K L A Q L A G A G N P G V V Q V H A A
240 250 260 270 280 290 300 310
GTACGAGGACGAGGCGTGGACGCACGTGGTCATAAACCTGTGCCCGGGCCCGGACCTGCTGGAGTGGGTGGGCCGCCG
Y E D E A W T H V V I N L C P G P D L L E W V G R R
320 330 340 350 360 370 380 390
CCGCGGCGCGCCCGTGCCGGAGCCCGTGGCCGCCGCCATCGTCGCGCAGGTCGCGCAGGCGCTCGCGCTCTGCCACCG
R G A P V P E P V A A A I V A Q V A Q A L A L C H R
400 410 420 430 440 450 460
CCGCGGCGTCGCCCACCGCGACGTGAAGCCCGAGAACATCCTCATCGACGCCGCCGCCGGGAGCGACGACGACGAGGA
R G V A H R D V K P E N I L I D A A A G S D D D E E
470 480 490 500 510 520 530 540
GGACGAGGCCGAGGCGGCGCCGCGCGCGCGCCTGGCGGACTTCGGGTCGGCGGCGTGGGTGGGCGCCGGTGGGCTCGG
D E A E A A P R A R L A D F G S A A W V G A G G L G
550 560 570 580 590 600 610 620
GCGCGCCGAGGGCCTTGTGGGGACGCCCCACTACGTGGCGCCCGAGGTGGTGGCCGGCGGCGAGTACGGCGCCAAGGC
R A E G L V G T P H Y V A P E V V A G G E Y G A K A
630 640 650 660 670 680 690 700
GGACGTGTGGAGCGCCGGCGTGGTGATGTACGCGCTGCTCTCCGGCGGCGCGTTCCCCTTCGGCGGCGACAGCGCCGC
D V W S A G V V M Y A L L S G G A F P F G G D S A A
710 720 730 740 750 760 770 780
GGAGGTGCTGGCGGCCGTGCTGCGCGGCAGCGCGCGCTTCCCGCCCAGGCTCTTCGGCGGGGTGTCGCCCGCCGCCAA
E V L A A V L R G S A R F P P R L F G G V S P A A K

```
          790         800         810         820         830         840         850
           |           |           |           |           |           |           |
GGACCTGATGCGGCGCATGATCTGCCGCGACGAGTGGAGGAGGTTCACCGCCGAGCAAGTCCTCGGTGAGTTTTTTCG
 D  L  M  R  R  M  I  C  R  D  E  W  R  R  F  T  A  E  Q  V  L  A

 860         870         880         890         900         910         920         930
 |           |           |           |           |           |           |           |
TTCGTTCCCCCGACAACCTCTCTCCTCCTCCTGCTCCGTTGATTTGATTTCCCCTTGTTCGTCGATGCAGTTTATCTA

   940         950         960         970         980         990        1000        1010
   |           |           |           |           |           |           |           |
ACAGATCTGCCGTTGGTAATAAAATAATGCAGCTCACCCGTGGATCGTGAGTGGCGGAGGAGCCCGGGCAATGGAGCG
                                    H  P  W  I  V  S  G  G  G  A  R  A  M  E  R

     1020        1030        1040        1050        1060        1070        1080        1090
      |           |           |           |           |           |           |           |
GCCAACCTGAGGAGGAATCCAGACGGACGTGTATTGCCGTAGAGCACTCAGGCACGGCAAGGAATCCCGGACGTTATG
 P  T  •

       1100        1110        1120        1130        1140
        |           |           |           |           |
TATGAGGTTTTTGTACGTACGTAGTAGTGTACGGTAGCTGCTGCCAGCTGCAG
```

Key to figure

C E E I	Translation frame
ATGTTTCAAA	Coding region
TTCAATTTCT	Non-coding region
CGGTAGCTGC	Primer site

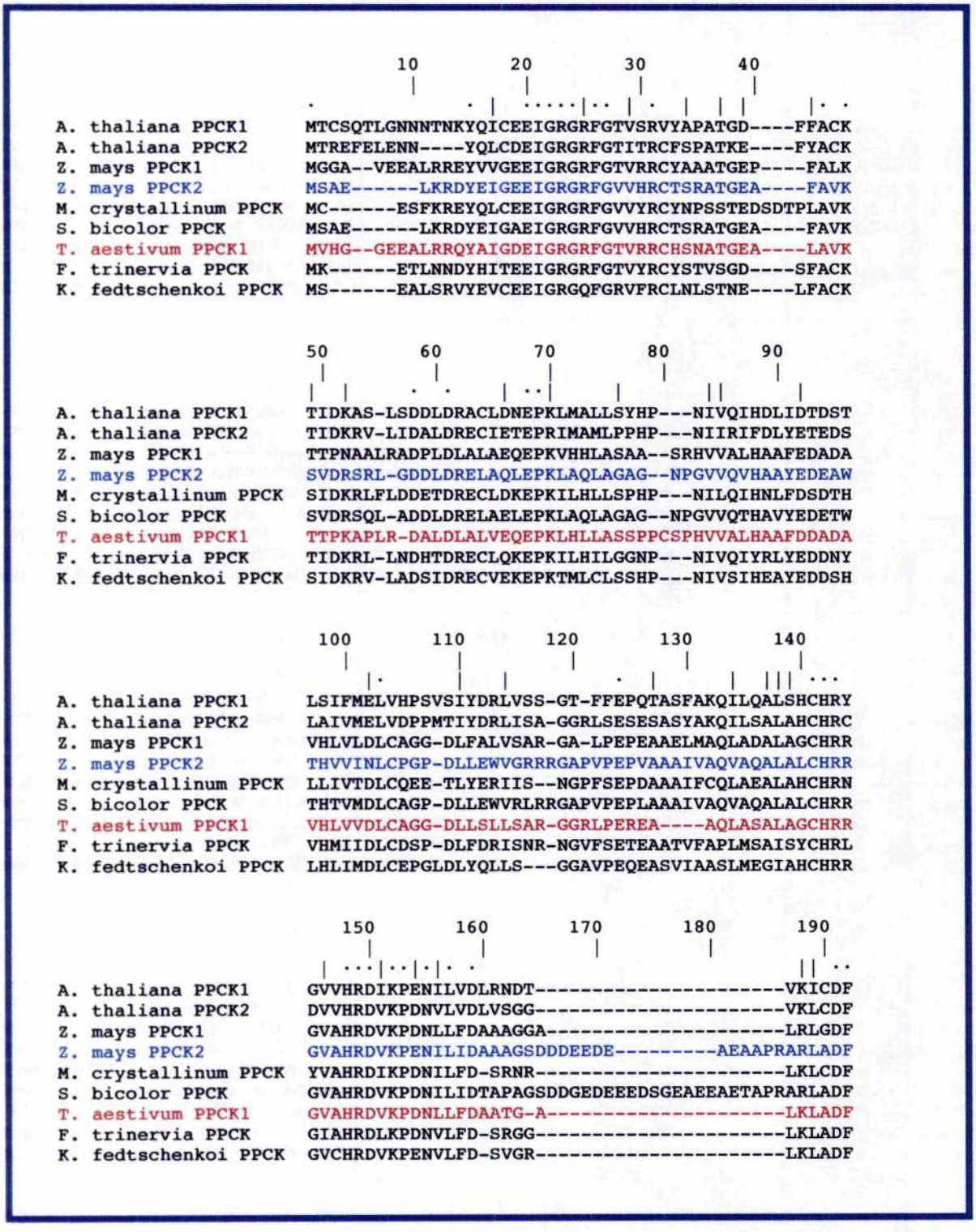
Key to figure

Translational start site

Intron and splice sites

Inframe stop codon

Figure 3.8 Alignment of the aminoacid sequences of *A. thaliana* PPCK1 and 2, *F. trinervia* PPCK, *M. crystallinum* PPCK, *K. fedtschenkoi* PPCK, *S. bicolor* PPCK, *Z. mays* PPCK2 and *T. aestivum* PPCK


10 20 30 40
A. thaliana PPCK1 MTCSQTLGNNNTNKYQICEEIGRGRFGTVSRVYAPATGD----FFACK
A. thaliana PPCK2 MTREFELENN----YQLCDEIGRGRFGTITRCFSPATKE----FYACK
Z. mays PPCK1 MGGA--VEEALRREYVVGEEIGRGRFGTVRRCYAAATGEP----FALK
Z. mays PPCK2 MSAE------LKRDYEIGEEIGRGRFGVVHRCTSRATGEA----FAVK
M. crystallinum PPCK MC------ESFKREYQLCEEIGRGRFGVVYRCYNPSSTEDSDTPLAVK
S. bicolor PPCK MSAE------LKRDYEIGAEIGRGRFGVVHRCTSRATGEA----FAVK
T. aestivum PPCK1 MVHG--GEEALRRQYAIGDEIGRGRFGTVRRCHSNATGEA----LAVK
F. trinervia PPCK MK------ETLNNDYHITEEIGRGRFGTVYRCYSTVSGD----SFACK
K. fedtschenkoi PPCK MS------EALSRVYEVCEEIGRGQFGRVFRCLNLSTNE----LFACK
50 60 70 80 90
A. thaliana PPCK1 TIDKAS-LSDDLDRACLDNEPKLMALLSYHP---NIVQIHDLIDTDST
A. thaliana PPCK2 TIDKRV-LIDALDRECIETEPRIMAMLPPHP---NIIRIFDLYETEDS
Z. mays PPCK1 TTPNAALRADPLDLALAEQEPKVHHLASAA--SRHVVALHAAFEDADA
Z. mays PPCK2 SVDRSRL-GDDLDRELAQLEPKLAQLAGAG--NPGVVQVHAAYEDEAW
M. crystallinum PPCK SIDKRLFLDDETDRECLDKEPKILHLLSPHP---NILQIHNLFDSDTH
S. bicolor PPCK SVDRSQL-ADDLDRELAELEPKLAQLAGAG--NPGVVQTHAVYEDETW
T. aestivum PPCK1 TTPKAPLR-DALDLALVEQEPKLHLLASSPPCSPHVVALHAAFDDADA
F. trinervia PPCK TIDKRL-LNDHTDRECLQKEPKILHILGGNP---NIVQIYRLYEDDNY
K. fedtschenkoi PPCK SIDKRV-LADSIDRECVEKEPKTMLCLSSHP---NIVSIHEAYEDDSH
100 110 120 130 140
A. thaliana PPCK1 LSIFMELVHPSVSIYDRLVSS-GT-FFEPQTASFAKQILQALSHCHRY
A. thaliana PPCK2 LAIVMELVDPPMTIYDRLISA-GGRLSESESASYAKQILSALAHCHRC
Z. mays PPCK1 VHLVLDLCAGG-DLFALVSAR-GA-LPEPEAAELMAQLADALAGCHRR
Z. mays PPCK2 THVVINLCPGP-DLLEWVGRRRGAPVPEPVAAAIVAQVAQALALCHRR
M. crystallinum PPCK LLIVTDLCQEE-TLYERIIS--NGPFSEPDAAAIFCQLAEALAHCHRN
S. bicolor PPCK THTVMDLCAGP-DLLEWVRLRRGAPVPEPLAAAIVAQVAQALALCHRR
T. aestivum PPCK1 VHLVVDLCAGG-DLLSLLSAR-GGRLPEREA----AQLASALAGCHRR
F. trinervia PPCK VHMIIDLCDSP-DLFDRISNR-NGVFSETEAATVFAPLMSAISYCHRL
K. fedtschenkoi PPCK LHLIMDLCEPGLDLYQLLS---GGAVPEQEASVIAASLMEGIAHCHRR
150 160 170 180 190
A. thaliana PPCK1 GVVHRDIKPENILVDLRNDT----------VKICDF
A. thaliana PPCK2 DVVHRDVKPDNVLVDLVSGG----------VKLCDF
Z. mays PPCK1 GVAHRDVKPDNLLFDAAAGGA----------LRLGDF
Z. mays PPCK2 GVAHRDVKPENILIDAAAGSDDDEEDE---------AEAAPRARLADF
M. crystallinum PPCK YVAHRDIKPDNILFD-SRNR----------LKLCDF
S. bicolor PPCK GVAHRDVKPDNILIDTAPAGSDDGEDEEEDSGEAEEAETAPRARLADF
T. aestivum PPCK1 GVAHRDVKPDNLLFDAATG-A----------LKLADF
F. trinervia PPCK GIAHRDLKPDNVLFD-SRGG----------LKLADF
K. fedtschenkoi PPCK GVCHRDVKPENVLFD-SVGR----------LKLADF

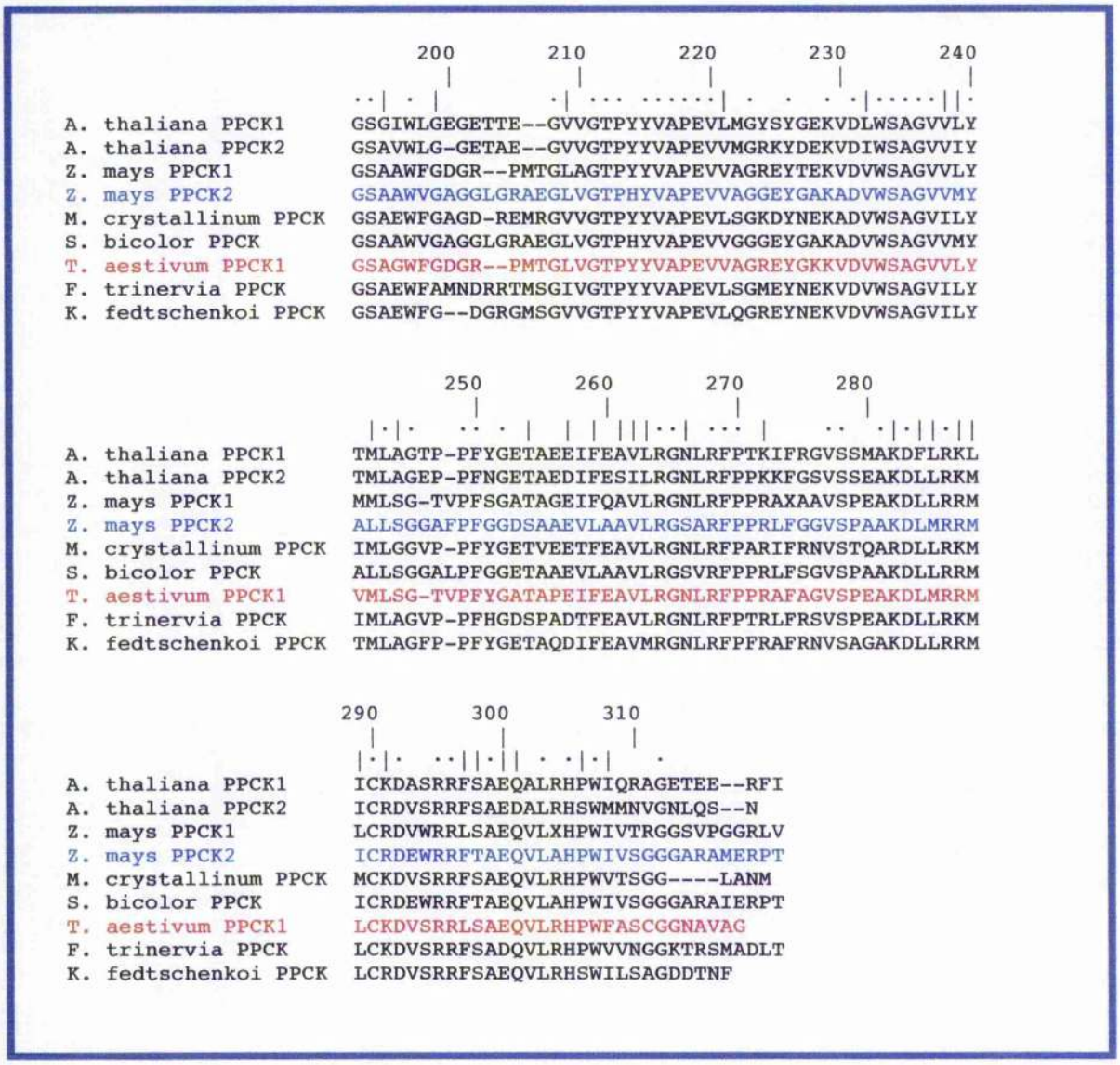

200 210 220 230 240
A. thaliana PPCK1 GSGIWLGEGETTE--GVVGTPYYVAPEVLMGYSYGEKVDLWSAGVVLY
A. thaliana PPCK2 GSAVWLG-GETAE--GVVGTPYYVAPEVVMGRKYDEKVDIWSAGVVIY
Z. mays PPCK1 GSAAWFGDGR--PMTGLAGTPYYVAPEVVAGREYTEKVDVWSAGVVLY
Z. mays PPCK2 GSAAWVGAGGLGRAEGLVGTPHYVAPEVVAGGEYGAKADVWSAGVVMY
M. crystallinum PPCK GSAEWFGAGD-REMRGVVGTPYYVAPEVLSGKDYNEKADVWSAGVILY
S. bicolor PPCK GSAAWVGAGGLGRAEGLVGTPHYVAPEVVGGGEYGAKADVWSAGVVMY
T. aestivum PPCK1 GSAGWFGDGR--PMTGLVGTPYYVAPEVVAGREYGKKVDVWSAGVVLY
F. trinervia PPCK GSAEWFAMNDRRTMSGIVGTPYYVAPEVLSGMEYNEKVDVWSAGVILY
K. fedtschenkoi PPCK GSAEWFG--DGRGMSGVVGTPYYVAPEVLQGREYNEKVDVWSAGVILY
250 260 270 280
A. thaliana PPCK1 TMLAGTP-PFYGETAEEIFEAVLRGNLRFPTKIFRGVSSMAKDFLRKL
A. thaliana PPCK2 TMLAGEP-PFNGETAEDIFESILRGNLRFPPKKFGSVSSEAKDLLRKM
Z. mays PPCK1 MMLSG-TVPFSGATAGEIFQAVLRGNLRFPPRAXAAVSPEAKDLLRRM
Z. mays PPCK2 ALLSGGAFPFGGDSAAEVLAAVLRGSARFPPRLFGGVSPAAKDLMRRM
M. crystallinum PPCK IMLGGVP-PFYGETVEETFEAVLRGNLRFPARIFRNVSTQARDLLRKM
S. bicolor PPCK ALLSGGALPFGGETAAEVLAAVLRGSVRFPPRLFSGVSPAAKDLLRRM
T. aestivum PPCK1 VMLSG-TVPFYGATAPEIFEAVLRGNLRFPPRAFAGVSPEAKDLMRRM
F. trinervia PPCK IMLAGVP-PFHGDSPADTFEAVLRGNLRFPTRLFRSVSPEAKDLLRKM
K. fedtschenkoi PPCK TMLAGFP-PFYGETAQDIFEAVMRGNLRFPFRAFRNVSAGAKDLLRRM
290 300 310
A. thaliana PPCK1 ICKDASRRFSAEQALRHPWIQRAGETEE--RFI
A. thaliana PPCK2 ICRDVSRRFSAEDALRHSWMMNVGNLQS--N
Z. mays PPCK1 LCRDVWRRLSAEQVLXHPWIVTRGGSVPGGRLV
Z. mays PPCK2 ICRDEWRRFTAEQVLAHPWIVSGGGARAMERPT
M. crystallinum PPCK MCKDVSRRFSAEQVLRHPWVTSGG----LANM
S. bicolor PPCK ICRDEWRRFTAEQVLAHPWIVSGGGARAIERPT
T. aestivum PPCK1 LCKDVSRRLSAEQVLRHPWFASCGGNAVAG
F. trinervia PPCK LCKDVSRRFSADQVLRHPWVVNGGKTRSMADLT
K. fedtschenkoi PPCK LCRDVSRRFSAEQVLRHSWILSAGDDTNF

Figure 3.9 RT-PCR analysis of maize PPCK2 transcripts and maize Actin1 transcripts from illuminated and darkened mature maize leaves

RT-PCR was carried out for 28 cycles

Lane (1) ZmPPCK2 transcripts in illuminated mature maize leaves.

Lane (2) ZmPPCK2 transcripts in darkened mature maize leaves.

Lane (3) ZmActin1 transcripts in illuminated mature maize leaves.

Lane (4) ZmActin1 transcripts in darkened mature maize leaves.

The lower bar indicates correctly processed transcripts. The upper bar indicates transcripts that have retained the intron.

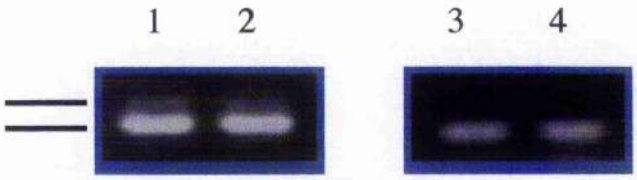
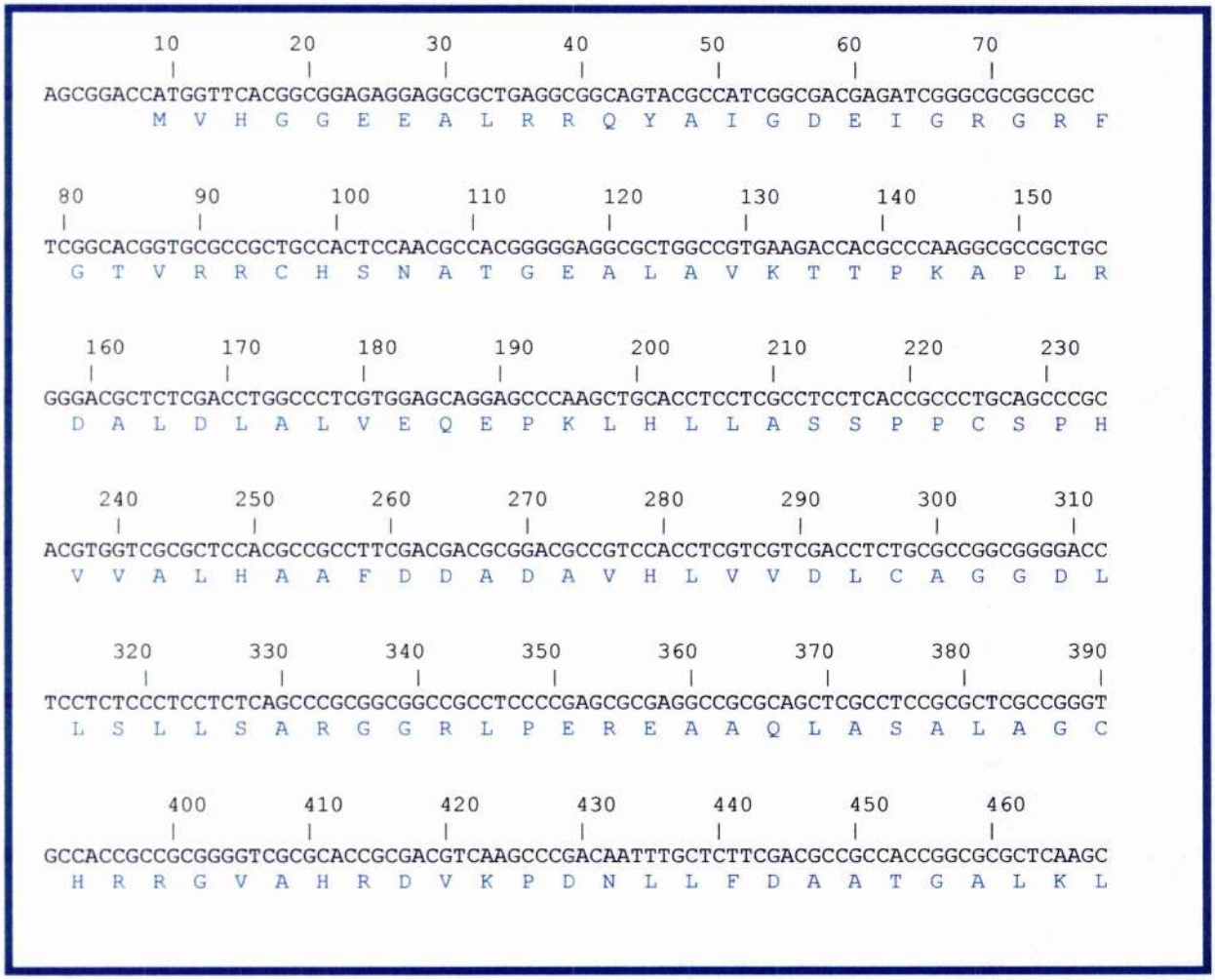
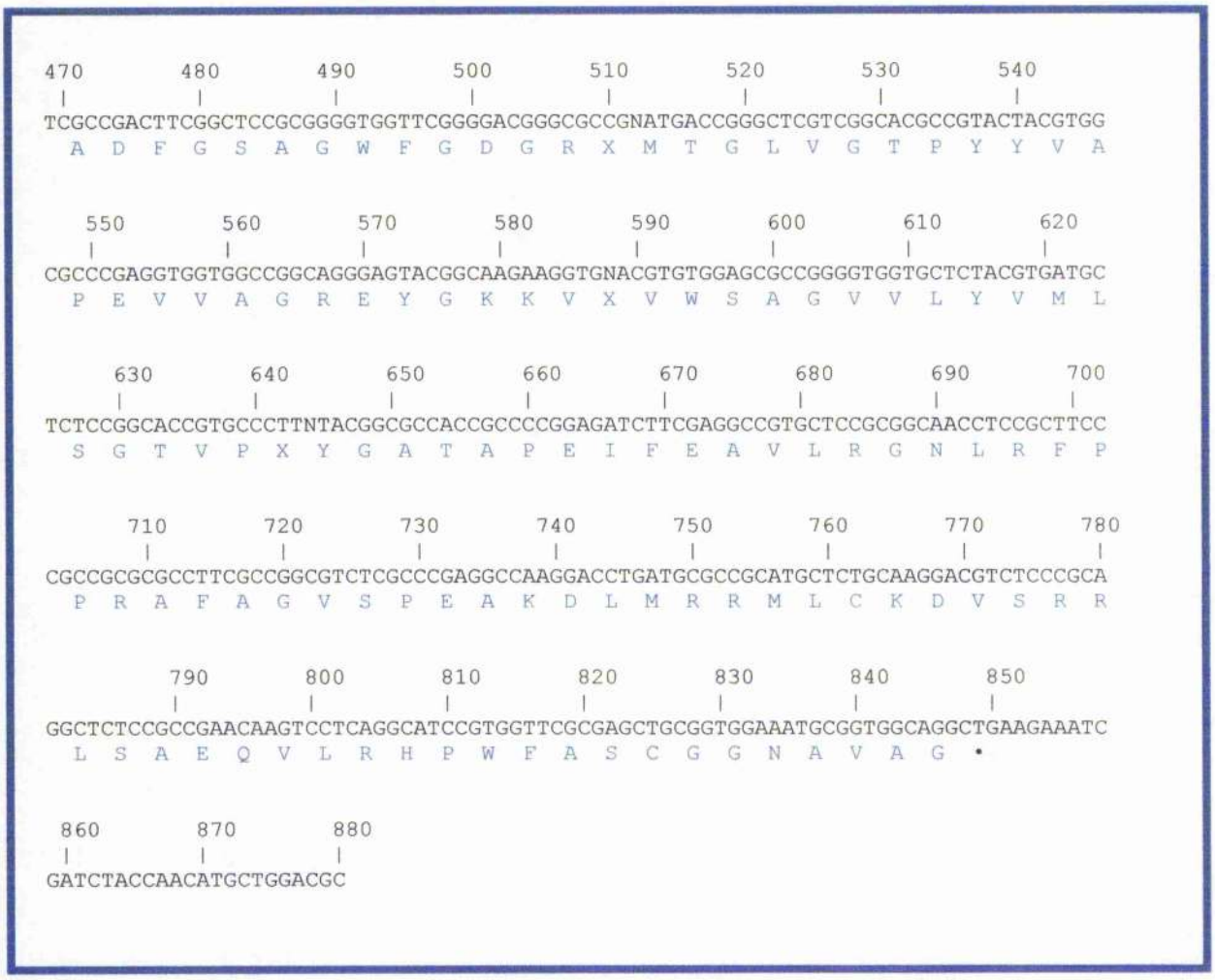

1
2
3
4

Figure 3.10 Translation map of the cDNA sequence of a *TaPPCK1*

This shows the predicted coding of *TaPPCK1* and the correct reading frame of TaPPCK2.


10 20 30 40 50 60 70
AGCGGACCATGGTTCACGGCGGAGAGGAGGCGCTGAGGCGGCAGTACGCCATCGGCGACGAGATCGGGCGCGGCCGC
M V H G G E E A L R R Q Y A I G D E I G R G R F
80 90 100 110 120 130 140 150
TCGGCACGGTGCGCCGCTGCCACTCCAACGCCACGGGGGAGGCGCTGGCCGTGAAGACCACGCCCAAGGCGCCGCTGC
G T V R R C H S N A T G E A L A V K T T P K A P L R
160 170 180 190 200 210 220 230
GGGACGCTCTCGACCTGGCCCTCGTGGAGCAGGAGCCCAAGCTGCACCTCCTCGCCTCCTCACCGCCCTGCAGCCCGC
D A L D L A L V E Q E P K L H L L A S S P P C S P H
240 250 260 270 280 290 300 310
ACGTGGTCGCGCTCCACGCCGCCTTCGACGACGCGGACGCCGTCCACCTCGTCGTCGACCTCTGCGCCGGCGGGGACC
V V A L H A A F D D A D A V H L V V D L C A G G D L
320 330 340 350 360 370 380 390
TCCTCTCCCTCCTCTCAGCCCGCGGCGGCCGCCTCCCCGAGCGCGAGGCCGCGCAGCTCGCCTCCGCGCTCGCCGGGT
L S L L S A R G G R L P E R E A A Q L A S A L A G C
400 410 420 430 440 450 460
GCCACCGCCGCGGGGTCGCGCACCGCGACGTCAAGCCCGACAATTTGCTCTTCGACGCCGCCACCGGCGCGCTCAAGC
H R R G V A H R D V K P D N L L F D A A T G A L K L


470 480 490 500 510 520 530 540
TCGCCGACTTCGGCTCCGCGGGGTGGTTCGGGGACGGGCGCCGNATGACCGGGCTCGTCGGCACGCCGTACTACGTGG
A D F G S A G W F G D G R X M T G L V G T P Y Y V A
550 560 570 580 590 600 610 620
CGCCCGAGGTGGTGGCCGGCAGGGAGTACGGCAAGAAGGTGNACGTGTGGAGCGCCGGGGTGGTGCTCTACGTGATGC
P E V V A G R E Y G K K V X V W S A G V V L Y V M L
630 640 650 660 670 680 690 700
TCTCCGGCACCGTGCCCTTNTACGGCGCCACCGCCCCGGAGATCTTCGAGGCCGTGCTCCGCGGCAACCTCCGCTTCC
S G T V P X Y G A T A P E I F E A V L R G N L R F P
710 720 730 740 750 760 770 780
CGCCGCGCGCCTTCGCCGGCGTCTCGCCCGAGGCCAAGGACCTGATGCGCCGCATGCTCTGCAAGGACGTCTCCCGCA
P R A F A G V S P E A K D L M R R M L C K D V S R R
790 800 810 820 830 840 850
GGCTCTCCGCCGAACAAGTCCTCAGGCATCCGTGGTTCGCGAGCTGCGGTGGAAATGCGGTGGCAGGCTGAAGAAATC
L S A E Q V L R H P W F A S C G G N A V A G •
860 870 880
GATCTACCAACATGCTGGACGC

Chapter 4

The identification and characterisation of two *L. esculentum* PPCK genes

4.1 Introduction

Several isoforms of PEPc in C_3 dicots exhibit expression patterns suggestive of specialised roles. These include root nodulation (Hata *et al.*, 1998), stomatal opening (Kopka *et al.*, 1997), development and germination of seeds (Gonzalez *et al.*, 1998, Golombek *et al.*, 1999), extension of cotton fibres (Smart *et al.*, 1998) and fruit ripening (Law and Plaxton, 1995, Guillet *et al.*, 2001).

The tomato plant is a C_3 dicot that is recognised as the model plant for investigating fleshy fruit development. Two tomato PEPc isoforms have been identified, with one showing expression and activity in a range of tissues (Ppc1) and the second being expressed in fruit (Ppc2) (Guillet *et al.*, 2001). The development of fleshy fruit is thought to involve PEPc in the synthesis of the organic acids, malic and citric acid. These accumulate during fruit development, rising at the end of the cell division phase, peaking towards the end of the cell expansion phase and decreasing thereafter. The malic acid synthesis, which of course involves PEPc, is thought to aid cellular expansion (Gehlen *et al.*, 1996, Guillet *et al.*, 2001, Latzko and Kelly, 1983, Macnicol and Raymond. 1998, Scheible *et al.*, 1997). As might be expected, PEPc expression and activity closely correspond to the changes in malic acid concentration of tomato fruit (Guillet *et al.*, 2001).

The expression and activity of LePpc2 seems to be tightly correlated to the development stages of fruit ripening before the breaker stage (Guillet *et al.*, 2001). At 8 days post anthesis (DPA), preferential localisation of LePpc2 transcripts was apparent in the thin layer of cytoplasm of the large vacuolated cells of the pericarp. The expression was strongest in the inner pericarp, facing the locule and in the expanding cells in the vicinity of the seeds (Guillet *et al.*, 2001) where cells undergo the highest enlargement rate during cell expansion (Mohr and Stein 1969). A second signal from LePpc2 transcripts were also noted around the vascular bundles (Guillet *et al.*, 2001). It is worth noting that in grapes, PEPc is expressed near the vascular bundles (Famiani *et al.*, 2000). This may suggest that PEPc provides an anapleurotic role in metabolising assimilates during their delivery to fruit via fruit vascular bundles. The main areas of localisation for LePpc2 appear to imply a close association

to cell expansion. LePpc2 expression peaks in the tissues, temporally and spatially, where tomato fruit cell diameter increases up to 40-fold leading to a fruit increase from 18 % (at 6 DPA) to 80% (at 20 DPA) of its final size (Mohr and Stein, 1969). As in guard cells and cotton fibre extension, the main osmoticum utilised to provide osmotic pressure driving expansion is organic acids along with sugars in the vacuoles of the cells (Cottele *et al.*, 1999, Smart *et al.*, 1998).

A further role for PEPc during ripening may be the recapture of respiratory CO_2. This can account for up to 25 % of the carbon imported into fruit during the cell expansion phase. Respiratory CO_2 increases in the ripening phases following expansion, suggesting a change in role for PEPc (Blanke and Lenz, 1989, Ho, 1996, Latzko and Kelly, 1983, Laval-Martin *et al.*, 1977, Law and Plaxton, 1997).

To further understand the role of PEPc in tomato plants it seemed necessary to identify and characterise PPCK genes and their expression patterns within a variety of tissues and organs.

4.2 Results

4.2.1 The identification of *LePPCK1*

A putative tomato PEPc kinase gene was identified from the est database. This cDNA clone [GenBank accession number: AF20348], here termed *LePPCK1*, was completely sequenced and the deduced aminoacid sequence was similar to those of previously identified PPCKs (Fontaine *et al.*, 2002, Hartwell *et al.*, 1999, Taybi *et al.*, 2000, Tsuchida *et al.*, 2001). To examine the structure of the gene, I designed PCR primers (see Table 1) to amplify genomic DNA. A single band of 1292 bp was generated, cloned and sequenced. Inspection of this sequence showed that the gene comprises two exons with an 89 bp intron from base 807 to base 896 (see Figure 4.1). The same PCR primers were used to amplify cDNA generating a band of 1203 bp. This was cloned and sequenced. Inspection revealed that the cDNA sequence matched the genomic sequence, bar the presence of the intron. A sequence alignment of the translated product with several known PPCKs is shown in Figure 4.2. This alignment reveals close similarity to the photosynthetic McPPCK, FtPPCK and KfPPCK.

LePPCK1 was cloned behind a T3 promoter in pBluescipt. To test whether the full-length cDNA clone was functional, I linearized the clone with *Not1* and transcribed and translated the insert. The translation product was then assayed for PPCK activity. Figure 4.3 shows that the translation product was of the expected size

(31 kDa) and was able to phosphorylate PEPc, confirming the identity of the gene product.

4.2.2 The identification of *LePPCK2*

I also identified and sequenced another putative tomato PPCK est [GenBank accession number: AW223421], here termed *LePPCK2*. Inspection of this sequence showed that the cDNA contains an insertion of 138 bases relative to other PPCKs. The insertion also included an inframe stop codon. Several hypothesis concerning the origin and role of this insertion were considered. Database comparisons suggest that the insertion is not similar to anything else in the database. However, perhaps the most obvious hypothesis, that it is an intron, did not appear to fit the data because the est sequence did not contain an obvious 3' splice end point (AG nucleotide signal). Another suggestion, that the stop codon actually encodes selenocysteine, is described further in Chapter 6.

The solution became evident when I examined the structure of the gene by designing PCR primers (see Table 1) to amplify cDNA (from tomato fruit) and genomic DNA. A single band of 1360 bp was generated from the genomic PCR, cloned and sequenced, while three smaller bands of 1192 bp, 1104 bp and 959 bp were generated from RT-PCR, cloned and sequenced (see Figure 4.5). Analysis of the sequences revealed that unlike other PPCKs, *LePPCK2* contains two introns. One 170 bp intron is found close to the 3' end of the coding sequence in the same position as the single intron in other PPCK genes; a second 233 bp intron is located close to the middle of the coding region (see Figure 4.4).

4.2.3 Alternate splicing occurs in *LePPCK2*

Unusually, the extra intron has two alternative 5' splice start sites, thus allowing alternative splicing. The region between these sites contains an inframe stop codon. Hence, splicing can give rise to a transcript which encodes a functional PPCK (transcript-α), an incorrectly spliced transcript with a premature stop codon (transcript-β) and an unspliced transcript with a premature stop codon (transcript-γ). The latter two both encode a truncated, non-functional PPCK (see Figure 4.5). The sequence alignment of the predicted aminoacid sequence from LePPCK2 transcript-α with several known PPCKs is shown in figure 4.2. This alignment reveals close similarity to other PPCKs.

LePPCK2 transcript-α and LePPCK2 transcript-γ were cloned behind a T7 promoter in pCR TOPO4, LePPCK2 transcript-β was cloned behind a T3 promoter in pBluescript. All the clones were linearized, transcribed and translated. The translation products were then assayed for PPCK activity. Figure 4.3 shows that the translation products were of the expected size - 31 kDa for transcript-α and of 16 kDa for transcripts-β and -γ due to the premature inframe stop codon. Only the products of transcript-α were able to phosphorylate PEPc.

4.2.4 Semi-quantitative RT-PCR analysis of tomato PPCKs

The expression of both tomato PPCKs was studied. I chose to investigate transcript levels by RT-PCR, rather than Northern blotting, in light of the alternative splicing of *LePPCK2* and the low levels of expression of PPCK genes in C_3 plants (e.g. Fontaine *et al.*, 2002). Primers for tomato *Actin52* and *Ubiquitin* (see Table 1) were used as a constitutive control in semi-quantitative RT-PCR, while specific primers (see Table 1) to the two tomato PPCK genes were used to amplify PPCK transcripts. A range of tissues were studied, including several stages of fruit development (see figure 4.6); young leaves; mature leaves; flowers; roots; seedlings in the light and dark.

LePPCK1 transcripts are present at a relatively low level in the tissues studied (see figure 4.7), as evidenced by the high number of cycles needed to amplify a visible signal. Relative to *LeActin52*, the level is highest in mature leaves and lowest in expansion phase 1 fruit. This suggests that LePPCK1 may be involved in a housekeeping function, perhaps phosphorylation of PEPc to provide metabolic intermediates for amino acid synthesis. Ests corresponding to LePPCK1 have been found in cDNA libraries from a range of tissues. This is consistent with the conclusion reached from the RT-PCR data (see Table 3).

The expression of *LePPCK2* shows three interesting features (see figure 4.7). First, it is strongly induced on fruit ripening. Secondly, the relative abundance of the three transcripts appears to depend on both tissue and conditions. In ripening fruit and illuminated seedlings, transcripts-β and -γ are more abundant than transcript-α. Whereas in roots and seedlings in the dark, the opposite is the case. Thirdly, as for *LePPCK1* expression, transcript abundance is light-enhanced. Again this pattern of expression is supported by the est database (see Table 3).

Table 3 Analysis of tomato ests

	GenBank number	Source of cDNA library
LePPCK1	AI774158	Pseudomonas resistant.
	AW033195	callus
	AW933544	mature green fruit
	BE431605	breaker fruit
	BE459112	developing/immature green fruit
	BF096819	nutrient deficient roots
	BG129025	shoot and meristem
	BI205159	suspension cultures
LePPCK2	AW223421	red ripe fruit
	AW222608	red ripe fruit
	AW441584	red ripe fruit
	AW442172	red ripe fruit
	AW738217	tomato flower buds
	BE462009	breaker fruit
	BF112946	breaker fruit
	BG127024	shoot and meristem
	BI205375	suspension cultures
	BI934398	flower, anthesis
	BM409042	breaker fruit
	BM409651	breaker fruit
	BM411455	breaker fruit
	BM536300	breaker fruit

This analysis includes ests up to 1/8/02.

The expression of both genes was investigated in compartments of red-ripe fruit. Five compartments were chosen - skin, outer pericarp, inner pericarp, locule and seeds. RNA was purified from these compartments and was analysed by RT-PCR. Figure 4.8 shows the results of these experiments. *LePPCK1* was most heavily expressed in the inner pericarp, with tissues on either side of this compartment showing lower transcript levels. No transcript was detected in the seed cDNA.

The expression data for *LePPCK2* shows two interesting and related characteristics. The most obvious of these is the absence of transcript-β and -γ within the locule and seeds. This is similar to the expression data obtained for "dark" seedlings (see Figure 4.7) suggesting that "darkened" tissues (i.e. those to which light does not penetrate) are unable to accumulate these transcripts. In this hypothesis, it is

assumed that light can penetrate to the inner pericarp, hence the expression of *LePPCK2* is similar in skin, outer and inner pericarp tissue. One might at first expect *LePPCK1* to be expressed similarly in skin, inner pericarp and outer pericarp, since it is also light-induced (see Figure 4.8 lanes 10 and 11). However LePPCK1 is clearly under some additional spatial control, since it is expressed only poorly in skin and is undetectable in seeds. Hence the expression pattern of *LePPCK1* is not incompatible with this hypothesis. Secondly, transcript-β and -γ are present in the skin and pericarp, but the level of transcript-α is about the same in all of the tissues. This suggests that there might be a similar, limiting amount of a splicing factor in each compartment.

4.2.5 PPCK genes in other *Solanaceous* plants

Further analysis of the est database revealed two putative potato PPCK cDNA clones. One is similar to *LePPCK1* while the other is a full length est similar to LePPCK2 transcript-γ [GenBank accession numbers: BG888336 and BG594065 respectively]. Subsequently other PPCK ests have been identified, as listed in Table 4. To examine the structure of these genes, now termed *StPPCK1* and *StPPCK2*, we designed PCR primers (see Table 1) to amplify cDNA (from mature leaves) and genomic DNA. Using StPPCK1 primers, a single band of 953 bp was generated from genomic PCR, cloned and sequenced. An intron of 89 bp was identified at the 3' end of the sequence. Primers designed to the region flanking the suspected middle intron of *StPPCK2* gave amplification of three bands from potato cDNA, of 597 bp, 502 bp and 363 bp, indicating that alternative splicing occurs in StPPCK2 transcripts (see figure 4.10). Comparison of the PPCK sequences show that StPPCK1 and LePPCK1 share 98 % identity while StPPCK2 and LePPCK2 share 92 % identity (see Figure 4.11).

Further evidence for a PPCK sub group in the *Solanaceae* that is alternatively spliced is provided by the amplification of a 600 bp product from gDNA of aubergine using PPCK2 primers as shown in Figure 3.10. This sequence, see Figure 4.11, is very similar to the LePPCK2 (and StPPCK2) genomic sequences, with conserved 5' and 3' splice sites in the putative intron, indicating that the intron is very conserved amongst the family *Solanaceae*.

Table 4 Analysis of potato ests

	GenBank number	Source of cDNA library
St PPCK 1	BG887336	dormant tuber
	BG887353	dormant tuber
	BQ045511	leaf
St PPCK 2	BG594065	sprouting eyes from tubers
	BG594668	sprouting eyes from tubers
	BQ505603	mixed tissues
	BQ505604	mixed tissues
	BQ505884	mixed tissues
	BQ505885	mixed tissues

This analysis includes ests up to 1/8/02.

4.3 Discussion

The results demonstrate that the tomato gene *LePPCK1* encodes a functional PPCK. Structurally the gene is similar to other reported PPCK genes (Fontaine *et al.*, 2002, Hartwell *et al.*, 1999, Taybi *et al.*, 2000, Tsuchida *et al.*, 2001) in that it encodes a protein kinase catalytic domain with no N or C-terminal extensions, and contains one small intron in the 3' region. However tomato contains a second gene *LePPCK2* which shows a novel feature - it contains an additional intron subject to alternative splicing. Of the three transcripts detected *in vivo*, two would lead to the production of an inactive truncated protein. It is not yet clear whether this protein accumulates *in vivo*. Only the third transcript (α) produces a functional PPCK protein. Further experiments show the existence of a homologue of LePPCK2 in other members of the *Solanaceae*. Indeed, transcripts from the potato homologue *StPPCK2* are also subject to alternative splicing. This suggests that the *Solanaceae* may contain a unique PPCK gene exhibiting alternative splicing. This leads to the question of whether alternative splicing is functionally significant. The observation that the relative ratios of the three transcripts depends on cell type and conditions suggest that this may be the case. At least three explanations seem possible.

The first possibility is that the truncated protein encoded by transcripts-β and -γ play a functional role, perhaps an inhibitor of PPCK. Recognition of PEPc by PPCK

is assumed to be at a tertiary structure level, rather than a primary structure level. This conclusion has been reached because PPCK, unlike most protein kinases, can only phosphorylate peptides corresponding to the sequence round the phosphorylation site very poorly (Li *et al.*, 1997). The putative truncated PPCK includes the entire N-terminal domain of the protein. Although this is the ATP-binding domain of protein kinases, it is possible that the N-terminal domain of PPCK includes a region that can bind to PEPc. It seems unlikely that this could be functionally important unless the truncated kinase is expressed at a similar level to PEPc in molar terms. It is also possible that the truncated protein binds ATP and has some function based on this property. It is conceivable that the truncated protein could bind to full-length PPCK and affect its function; however this seems unlikely because PPCK is thought to behave as a monomer. It is worth noting that the PDE4 system might provide an analogy, whereby the alternatively spliced, catalytically inactive PDE4A isoform 2EL has a putative role as a competitive inhibitor to its substrate (Sullivan *et al.*, 1998).

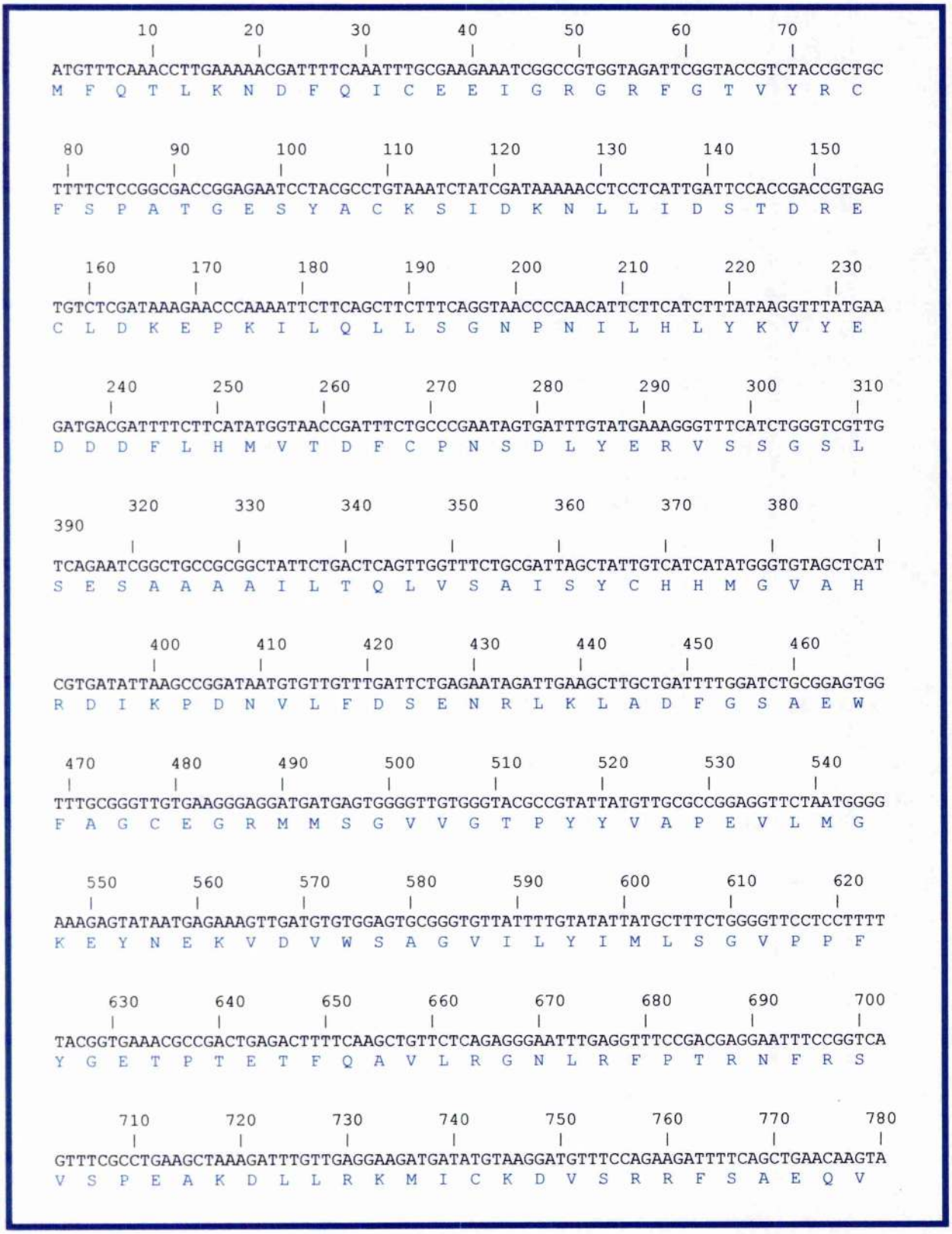
Another explanation of the alternative splicing is that it represents a second level of control ensuring a precise regulation of the level of abundance of LePPCK2 transcript-α in the cell. It is clear that *LePPCK2* gene expression is up regulated in fruit, post breaker phase. It seems unlikely that every cell in the post breaker fruit would need high amounts of PPCK activity, especially since PEPc activity actually falls after the expansion phases in fruit (Guillet *et al.*, 2001). Therefore, the alternatively spliced intron in the PPCK2 gene may act as a potential 'braking' mechanism by which high levels of gene expression may be nullified in a cell. This would allow the cell to produce either large amounts of active PPCK or very little. If this is the case, it seems likely that a regulatory splicing factor is involved. A third explanation is suggested by the uniqueness of the *Solanaceae* PPCK2 genes. It is possible that the alternatively spliced intron was gained by an ancestor of this family. The *Solanaceae* plants may have adapted to this chance event and learnt to 'cope with' the extra splicing requirement with no great cost. Thus the alternative splicing and ratio differences observed may have no physiological significance.

I have also shown expression studies of both tomato PPCKs, across a range of tissues. Encouragingly, my RT-PCR data agrees with the est distribution across tissues, showing that LePPCK1 is found across most types of tissue, while LePPCK2 is predominantly found in ripening tissue, post breaker (see Table 3). This illustrates

the use of est libraries in 'virtual expression' experiments. The case of LePpc2 validates this approach. Some 20 ests correspond to this gene, and all come from developing (**not** ripening) fruit libraries. As noted earlier Guillet *et al.* (2001) showed directly that this gene is expressed almost exclusively in green fruit. This would suggest that LePPCK1 plays an anapleurotic role, mainly involved in ensuring the replenishment of the TCA cycle, while the role of LePPCK2 role is limited to late ripening. An accumulation of PPCK in late ripening could be explained by three possible hypotheses, or some combination of these. One putative function is that PPCK2 is up regulated in post breaker fruit to ensure that the limited amount of fruit PEPc is phosphorylated thereby aiding the re-fixation of respiratory CO_2. Respiratory CO_2 is produced in high amounts in ripening fruits. A second hypothesis is that LePPCK2 could phosphorylate another substrate beside PEPc. This seems unlikely as PPCK has been shown to be extremely specific to PEPc. A third hypothesis is that LePPCK2 transcript-α is localised to specialised cells in ripening fruit. One suspect tissue is around vascular tissue as grapes and tomato have been shown to have active PEPc localised here (Famini *et al.*, 2000, Guillet *et al.*, 2001). It is possible that LePPCK2 pre-mRNA is correctly spliced only in these cells. It is therefore clear that more work is necessary to clarify the role and function of LePPCK2.

Figure 4.1 Translation map of the genomic sequence of a *LePPCK1* and a schematic of the gene structure

Panel A shows the translation map of the genomic sequence of *LePPCK1* with annotations showing correct translation frame, primer sites and non-coding regions. Panel B shows a schematic of *LePPCK1* gene structure.


10 20 30 40 50 60 70
ATGTTTCAAACCTTGAAAAACGATTTTCAAATTTGCGAAGAAATCGGCCGTGGTAGATTCGGTACCGTCTACCGCTGC
M F Q T L K N D F Q I C E E I G R G R F G T V Y R C
80 90 100 110 120 130 140 150
TTTTCTCCGGCGACCGGAGAATCCTACGCCTGTAAATCTATCGATAAAAACCTCCTCATTGATTCCACCGACCGTGAG
F S P A T G E S Y A C K S I D K N L L I D S T D R E
160 170 180 190 200 210 220 230
TGTCTCGATAAAGAACCCAAAATTCTTCAGCTTCTTTCAGGTAACCCCAACATTCTTCATCTTTATAAGGTTTATGAA
C L D K E P K I L Q L L S G N P N I L H L Y K V Y E
240 250 260 270 280 290 300 310
GATGACGATTTTCTTCATATGGTAACCGATTTCTGCCCGAATAGTGATTTGTATGAAAGGGTTTCATCTGGGTCGTTG
D D D F L H M V T D F C P N S D L Y E R V S S G S L
320 330 340 350 360 370 380 390
TCAGAATCGGCTGCCGCGGCTATTCTGACTCAGTTGGTTTCTGCGATTAGCTATTGTCATCATATGGGTGTAGCTCAT
S E S A A A A I L T Q L V S A I S Y C H H M G V A H
400 410 420 430 440 450 460
CGTGATATTAAGCCGGATAATGTGTTGTTTGATTCTGAGAATAGATTGAAGCTTGCTGATTTTGGATCTGCGGAGTGG
R D I K P D N V L F D S E N R L K L A D F G S A E W
470 480 490 500 510 520 530 540
TTTGCGGGTTGTGAAGGGAGGATGATGAGTGGGGTTGTGGGTACGCCGTATTATGTTGCGCCGGAGGTTCTAATGGGG
F A G C E G R M M S G V V G T P Y Y V A P E V L M G
550 560 570 580 590 600 610 620
AAAGAGTATAATGAGAAAGTTGATGTGTGGAGTGCGGGTGTTATTTTGTATATTATGCTTTCTGGGGTTCCTCCTTTT
K E Y N E K V D V W S A G V I L Y I M L S G V P P F
630 640 650 660 670 680 690 700
TACGGTGAAACGCCGACTGAGACTTTTCAAGCTGTTCTCAGAGGGAATTTGAGGTTTCCGACGAGGAATTTCCGGTCA
Y G E T P T E T F Q A V L R G N L R F P T R N F R S
710 720 730 740 750 760 770 780
GTTTCGCCTGAAGCTAAAGATTTGTTGAGGAAGATGATATGTAAGGATGTTTCCAGAAGATTTTCAGCTGAACAAGTA
V S P E A K D L L R K M I C K D V S R R F S A E Q V

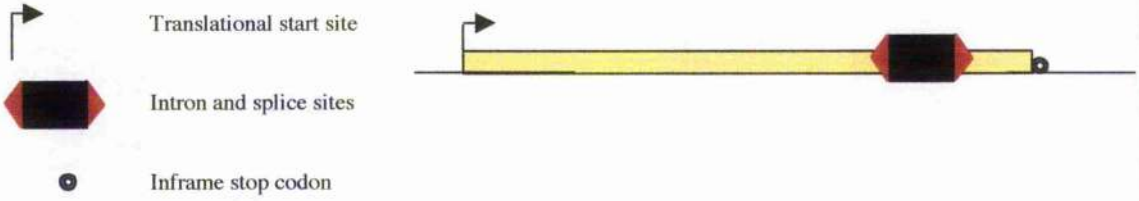
```
          790         800         810         820         830         840         850
          |           |           |           |           |           |           |
CTGAGTAAGTTCAATTTCTCTTATTTTCACTTAATTTTGCTTTCCCCTGTTTTTTTCGAGTTGCTGATTTTGATGTCC
L  R


 860         870         880         890         900         910         920
 |           |           |           |           |           |           |
TTTGTTATATTACAGGACACCCATGGGTGATTAATGGAGGTGAAACAAGATCAATGGCTGATCTAAACTGA
                H  P  W  V  I  N  G  G  E  T  R  S  M  A  D  L  N  •
```

Key to figure

C E E I	Translation frame
ATGTTTCAAA	Coding region
TTCAATTTCT	Non-coding region

Key to figure



Translational start site

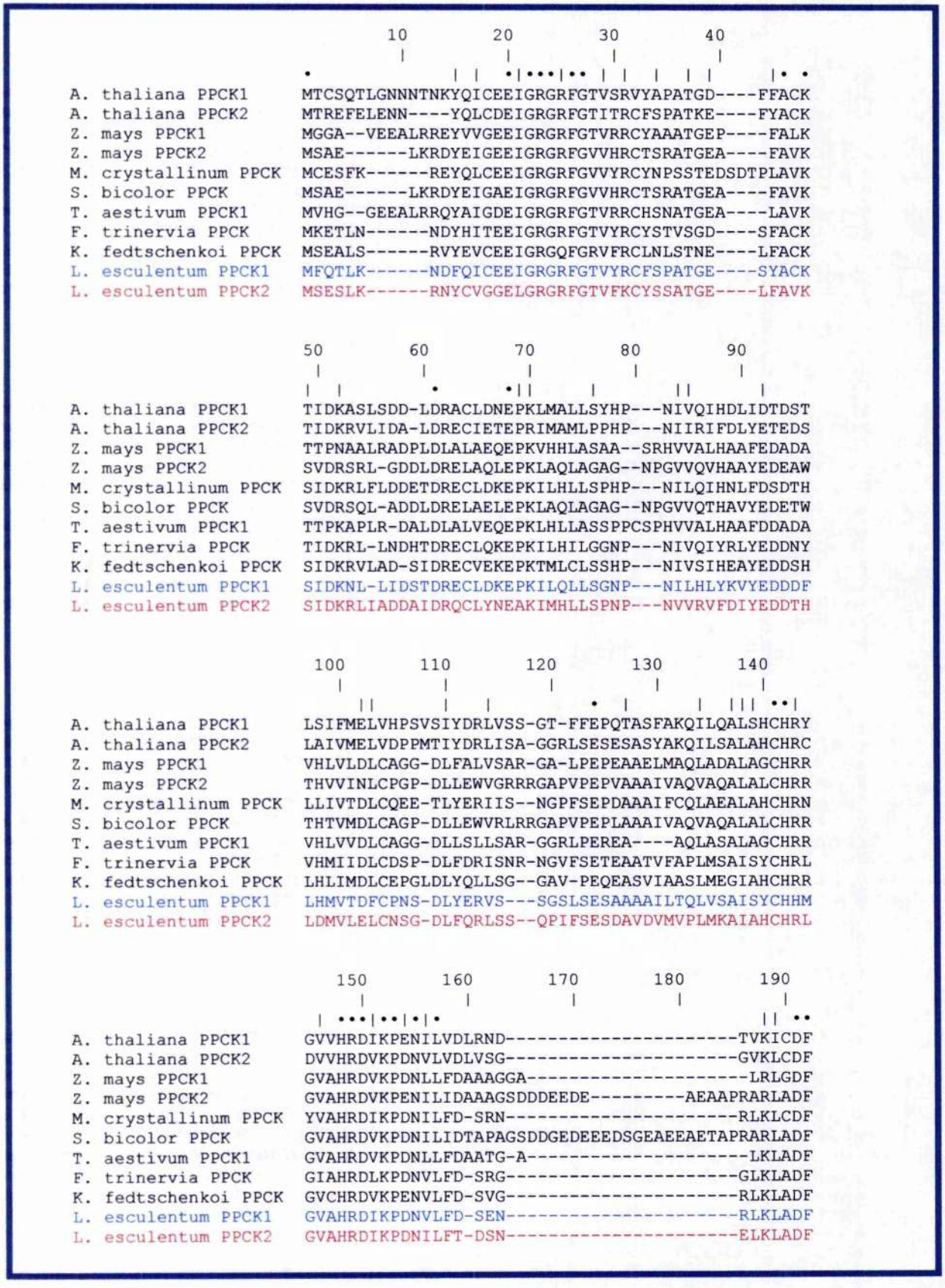
Intron and splice sites

Inframe stop codon

Figure 4.2 Alignment of the aminoacid sequences of *A thaliana* PPCK1 and 2, F trinervia PPCK, *M. crystallinum* PPCK, *K. fedtschenkoi* PPCK, *S. bicolor* PPCK, *Z. mays* PPCK1 and 2, *T. aestivum* PPCK and *L. esculentum* PPCK1 and 2

The green font indicates the predicted aminoacid sequence of LePPCK1.

The blue font indicates the predicted aminoacid sequence of LePPCK2.


10 20 30 40
A. thaliana PPCK1 MTCSQTLGNNNTNKYQICEEIGRGRFGTVSRVYAPATGD----FFACK
A. thaliana PPCK2 MTREFELENN----YQLCDEIGRGRFGTITRCFSPATKE----FYACK
Z. mays PPCK1 MGGA--VEEALRREYVVGEEIGRGRFGTVRRCYAAATGEP----FALK
Z. mays PPCK2 MSAE------LKRDYEIGEEIGRGRFGVVHRCTSRATGEA----FAVK
M. crystallinum PPCK MCESFK------REYQLCEEIGRGRFGVVYRCYNPSSTEDSDTPLAVK
S. bicolor PPCK MSAE------LKRDYEIGAEIGRGRFGVVHRCTSRATGEA----FAVK
T. aestivum PPCK1 MVHG--GEEALRRQYAIGDEIGRGRFGTVRRCHSNATGEA----LAVK
F. trinervia PPCK MKETLN------NDYHITEEIGRGRFGTVYRCYSTVSGD----SFACK
K. fedtschenkoi PPCK MSEALS------RVYEVCEEIGRGQFGRVFRCLNLSTNE----LFACK
L. esculentum PPCK1 MFQTLK------NDFQICEEIGRGRFGTVYRCFSPATGE----SYACK
L. esculentum PPCK2 MSESLK------RNYCVGGELGRGRFGTVFKCYSSATGE----LFAVK
50 60 70 80 90
A. thaliana PPCK1 TIDKASLSDD-LDRACLDNEPKLMALLSYHP---NIVQIHDLIDTDST
A. thaliana PPCK2 TIDKRVLIDA-LDRECIETEPRIMAMLPPHP---NIIRIFDLYETEDS
Z. mays PPCK1 TTPNAALRADPLDLALAEQEPKVHHLASAA--SRHVVALHAAFEDADA
Z. mays PPCK2 SVDRSRL-GDDLDRELAQLEPKLAQLAGAG--NPGVVQVHAAYEDEAW
M. crystallinum PPCK SIDKRLFLDDETDRECLDKEPKILHLLSPHP---NILQIHNLFDSDTH
S. bicolor PPCK SVDRSQL-ADDLDRELAELEPKLAQLAGAG--NPGVVQTHAVYEDETW
T. aestivum PPCK1 TTPKAPLR-DALDLALVEQEPKLHLLASSPPCSPHVVALHAAFDDADA
F. trinervia PPCK TIDKRL-LNDHTDRECLQKEPKILHILGGNP---NIVQIYRLYEDDNY
K. fedtschenkoi PPCK SIDKRVLAD-SIDRECVEKEPKTMLCLSSHP---NIVSIHEAYEDDSH
L. esculentum PPCK1 SIDKNL-LIDSTDRECLDKEPKILQLLSGNP---NILHLYKVYEDDDF
L. esculentum PPCK2 SIDKRLIADDAIDRQCLYNEAKIMHLLSPNP---NVVRVFDIYEDDTH
100 110 120 130 140
A. thaliana PPCK1 LSIFMELVHPSVSIYDRLVSS-GT-FFEPQTASFAKQILQALSHCHRY
A. thaliana PPCK2 LAIVMELVDPPMTIYDRLISA-GGRLSESESASYAKQILSALAHCHRC
Z. mays PPCK1 VHLVLDLCAGG-DLFALVSAR-GA-LPEPEAAELMAQLADALAGCHRR
Z. mays PPCK2 THVVINLCPGP-DLLEWVGRRRGAPVPEPVAAAIVAQVAQALALCHRR
M. crystallinum PPCK LLIVTDLCQEE-TLYERIIS--NGPFSEPDAAAIFCQLAEALAHCHRN
S. bicolor PPCK THTVMDLCAGP-DLLEWVRLRRGAPVPEPLAAAIVAQVAQALALCHRR
T. aestivum PPCK1 VHLVVDLCAGG-DLLSLLSAR-GGRLPEREA----AQLASALAGCHRR
F. trinervia PPCK VHMIIDLCDSP-DLFDRISNR-NGVFSETEAATVFAPLMSAISYCHRL
K. fedtschenkoi PPCK LHLIMDLCEPGLDLYQLLSG--GAV-PEQEASVIAASLMEGIAHCHRR
L. esculentum PPCK1 LHMVTDFCPNS-DLYERVS---SGSLSESAAAAILTQLVSAISYCHHM
L. esculentum PPCK2 LDMVLELCNSG-DLFQRLSS--QPIFSESDAVDVMVPLMKAIAHCHRL
150 160 170 180 190
A. thaliana PPCK1 GVVHRDIKPENILVDLRND----------TVKICDF
A. thaliana PPCK2 DVVHRDVKPDNVLVDLVSG----------GVKLCDF
Z. mays PPCK1 GVAHRDVKPDNLLFDAAAGGA----------LRLGDF
Z. mays PPCK2 GVAHRDVKPENILIDAAAGSDDDEEDE---------AEAAPRARLADF
M. crystallinum PPCK YVAHRDIKPDNILFD-SRN----------RLKLCDF
S. bicolor PPCK GVAHRDVKPDNILIDTAPAGSDDGEDEEEDSGEAEEAETAPRARLADF
T. aestivum PPCK1 GVAHRDVKPDNLLFDAATG-A----------LKLADF
F. trinervia PPCK GIAHRDLKPDNVLFD-SRG----------GLKLADF
K. fedtschenkoi PPCK GVCHRDVKPENVLFD-SVG----------RLKLADF
L. esculentum PPCK1 GVAHRDIKPDNVLFD-SEN----------RLKLADF
L. esculentum PPCK2 GVAHRDIKPDNILFT-DSN----------ELKLADF

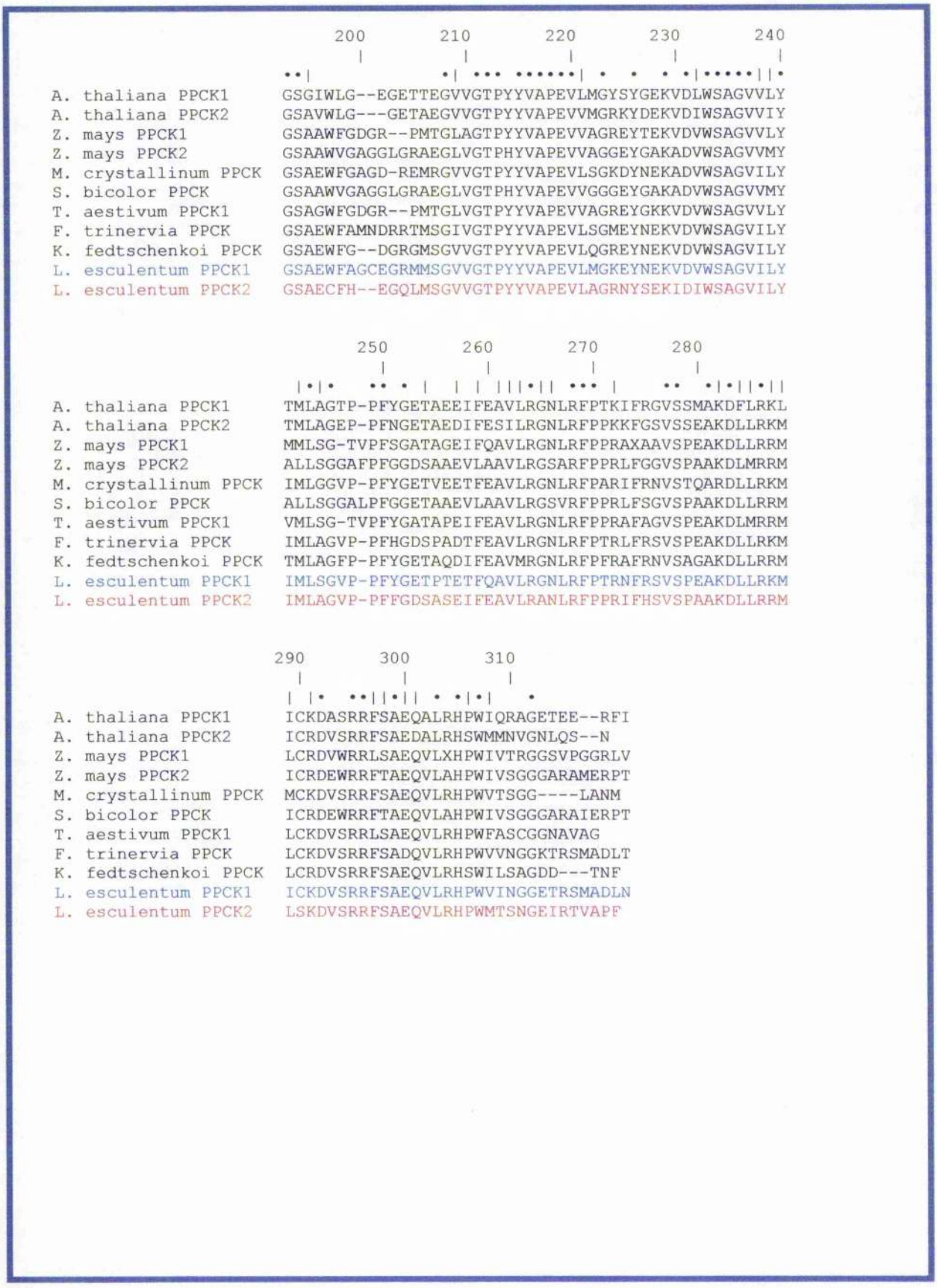

200 210 220 230 240
A. thaliana PPCK1 GSGIWLG--EGETTEGVVGTPYYVAPEVLMGYSYGEKVDLWSAGVVLY
A. thaliana PPCK2 GSAVWLG---GETAEGVVGTPYYVAPEVVMGRKYDEKVDIWSAGVVIY
Z. mays PPCK1 GSAAWFGDGR--PMTGLAGTPYYVAPEVVAGREYTEKVDVWSAGVVLY
Z. mays PPCK2 GSAAWVGAGGLGRAEGLVGTPHYVAPEVVAGGEYGAKADVWSAGVVMY
M. crystallinum PPCK GSAEWFGAGD-REMRGVVGTPYYVAPEVLSGKDYNEKADVWSAGVILY
S. bicolor PPCK GSAAWVGAGGLGRAEGLVGTPHYVAPEVVGGGEYGAKADVWSAGVVMY
T. aestivum PPCK1 GSAGWFGDGR--PMTGLVGTPYYVAPEVVAGREYGKKVDVWSAGVVLY
F. trinervia PPCK GSAEWFAMNDRRTMSGIVGTPYYVAPEVLSGMEYNEKVDVWSAGVILY
K. fedtschenkoi PPCK GSAEWFG--DGRGMSGVVGTPYYVAPEVLQGREYNEKVDVWSAGVILY
L. esculentum PPCK1 GSAEWFAGCEGRMMSGVVGTPYYVAPEVLMGKEYNEKVDVWSAGVILY
L. esculentum PPCK2 GSAECFH--EGQLMSGVVGTPYYVAPEVLAGRNYSEKIDIWSAGVILY
250 260 270 280
A. thaliana PPCK1 TMLAGTP-PFYGETAEEIFEAVLRGNLRFPTKIFRGVSSMAKDFLRKL
A. thaliana PPCK2 TMLAGEP-PFNGETAEDIFESILRGNLRFPPKKFGSVSSEAKDLLRKM
Z. mays PPCK1 MMLSG-TVPFSGATAGEIFQAVLRGNLRFPPRAXAAVSPEAKDLLRRM
Z. mays PPCK2 ALLSGGAFPFGGDSAAEVLAAVLRGSARFPPRLFGGVSPAAKDLMRRM
M. crystallinum PPCK IMLGGVP-PFYGETVEETFEAVLRGNLRFPARIFRNVSTQARDLLRKM
S. bicolor PPCK ALLSGGALPFGGETAAEVLAAVLRGSVRFPPRLFSGVSPAAKDLLRRM
T. aestivum PPCK1 VMLSG-TVPFYGATAPEIFEAVLRGNLRFPPRAFAGVSPEAKDLMRRM
F. trinervia PPCK IMLAGVP-PFHGDSPADTFEAVLRGNLRFPTRLFRSVSPEAKDLLRKM
K. fedtschenkoi PPCK TMLAGFP-PFYGETAQDIFEAVMRGNLRFPFRAFRNVSAGAKDLLRRM
L. esculentum PPCK1 IMLSGVP-PFYGETPTETFQAVLRGNLRFPTRNFRSVSPEAKDLLRKM
L. esculentum PPCK2 IMLAGVP-PFFGDSASEIFEAVLRANLRFPPRIFHSVSPAAKDLLRRM
290 300 310
A. thaliana PPCK1 ICKDASRRFSAEQALRHPWIQRAGETEE--RFI
A. thaliana PPCK2 ICRDVSRRFSAEDALRHSWMMNVGNLQS--N
Z. mays PPCK1 LCRDVWRRLSAEQVLXHPWIVTRGGSVPGGRLV
Z. mays PPCK2 ICRDEWRRFTAEQVLAHPWIVSGGGARAMERPT
M. crystallinum PPCK MCKDVSRRFSAEQVLRHPWVTSGG----LANM
S. bicolor PPCK ICRDEWRRFTAEQVLAHPWIVSGGGARAIERPT
T. aestivum PPCK1 LCKDVSRRLSAEQVLRHPWFASCGGNAVAG
F. trinervia PPCK LCKDVSRRFSADQVLRHPWVVNGGKTRSMADLT
K. fedtschenkoi PPCK LCRDVSRRFSAEQVLRHSWILSAGDD---TNF
L. esculentum PPCK1 ICKDVSRRFSAEQVLRHPWVINGGETRSMADLN
L. esculentum PPCK2 LSKDVSRRFSAEQVLRHPWMTSNGEIRTVAPF

Figure 4.3 PPCK activity encoded by *L. esculentum* PPCK1 and *L. esculentum* PPCK2

Panel A shows phosphorimages of ^{35}S Met-labelled products from the *in vitro* translation of RNA transcribed from the full length tomato cDNAs, separated on a 12.5 % SDS polyacrylamide gel. Clones were transcribed and translated, and translation products were assayed for PPCK activity as described in section 2.16, 2.9 and 2.10 respectively.

Lane (1) RNA transcribed from the full length LePPCK2 transcript-β

Lane (2) RNA transcribed from the full length LePPCK1

Lane (3) no RNA control

Lane (4) RNA transcribed from the full length LePPCK2 transcript-α

Lane (5) RNA transcribed from the full length LePPCK2 transcript-γ

Lane (6) RNA transcribed from the full length KfPPCK.

Lane (7) no RNA control

Panel B shows the immunoprecipitated ^{32}P-labelled PEPc from assays of the PPCK activity of translation products. Lanes are loaded as in panel A. The bar indicates PEPc. The figures below each lane indicate relative intensity as a percentage of lane 6 as a control.

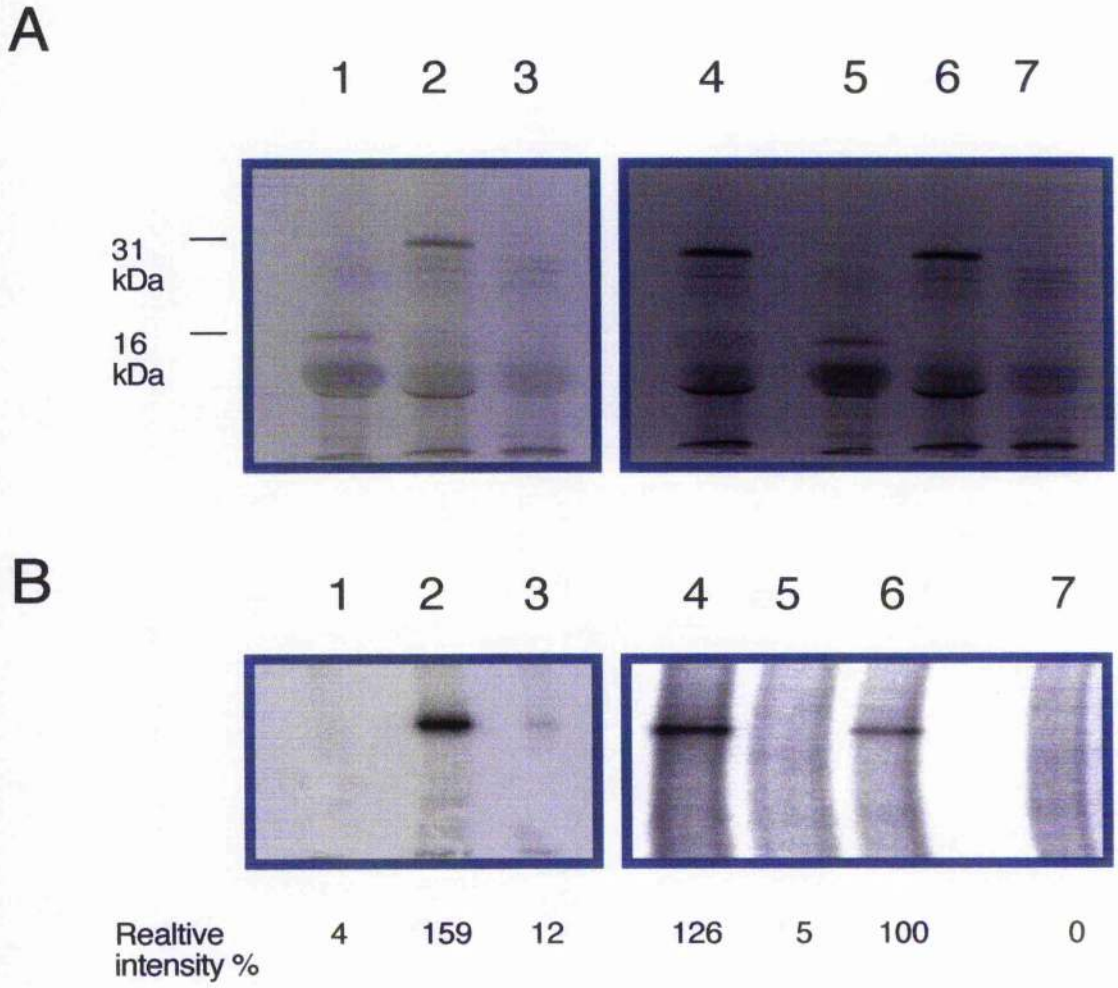
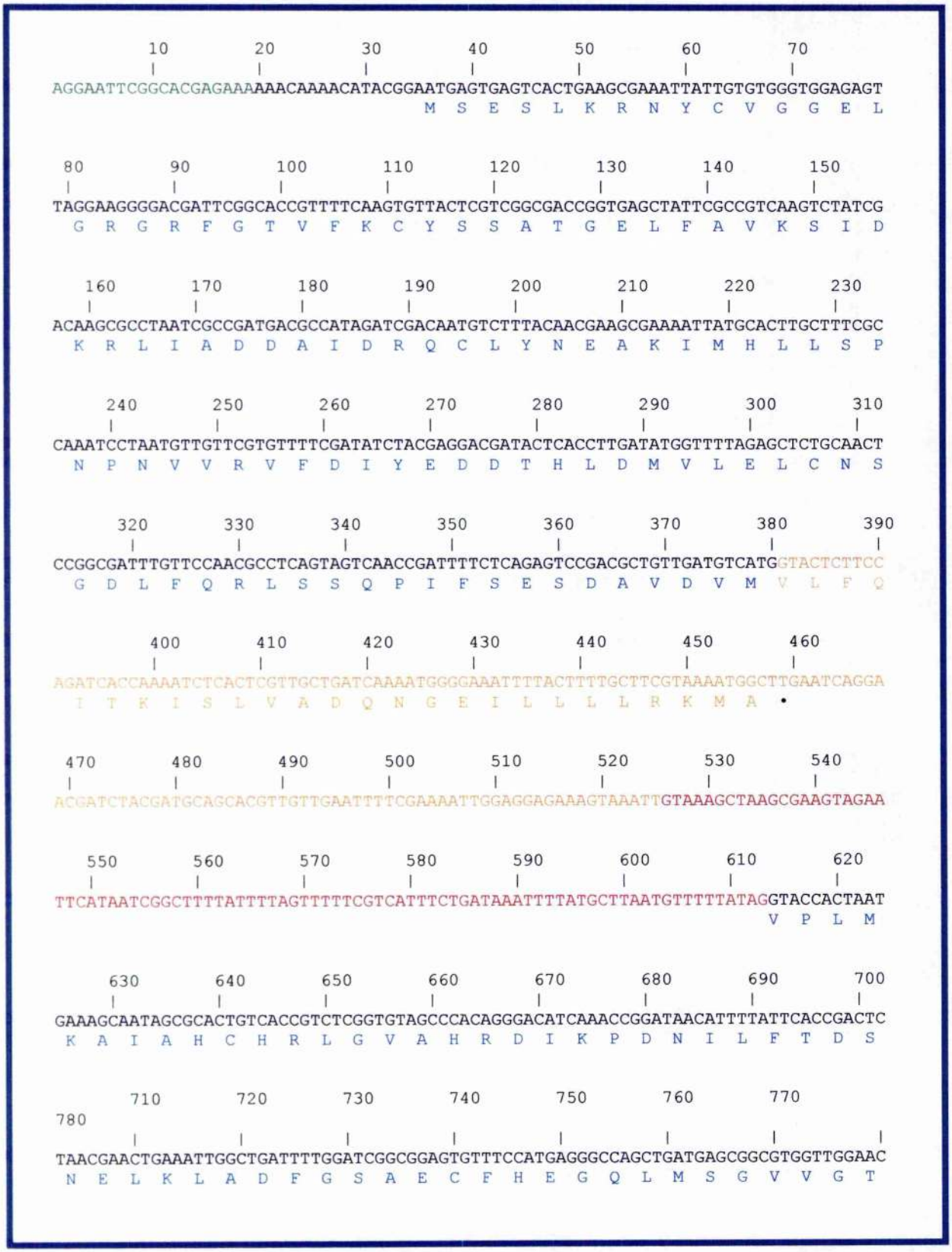
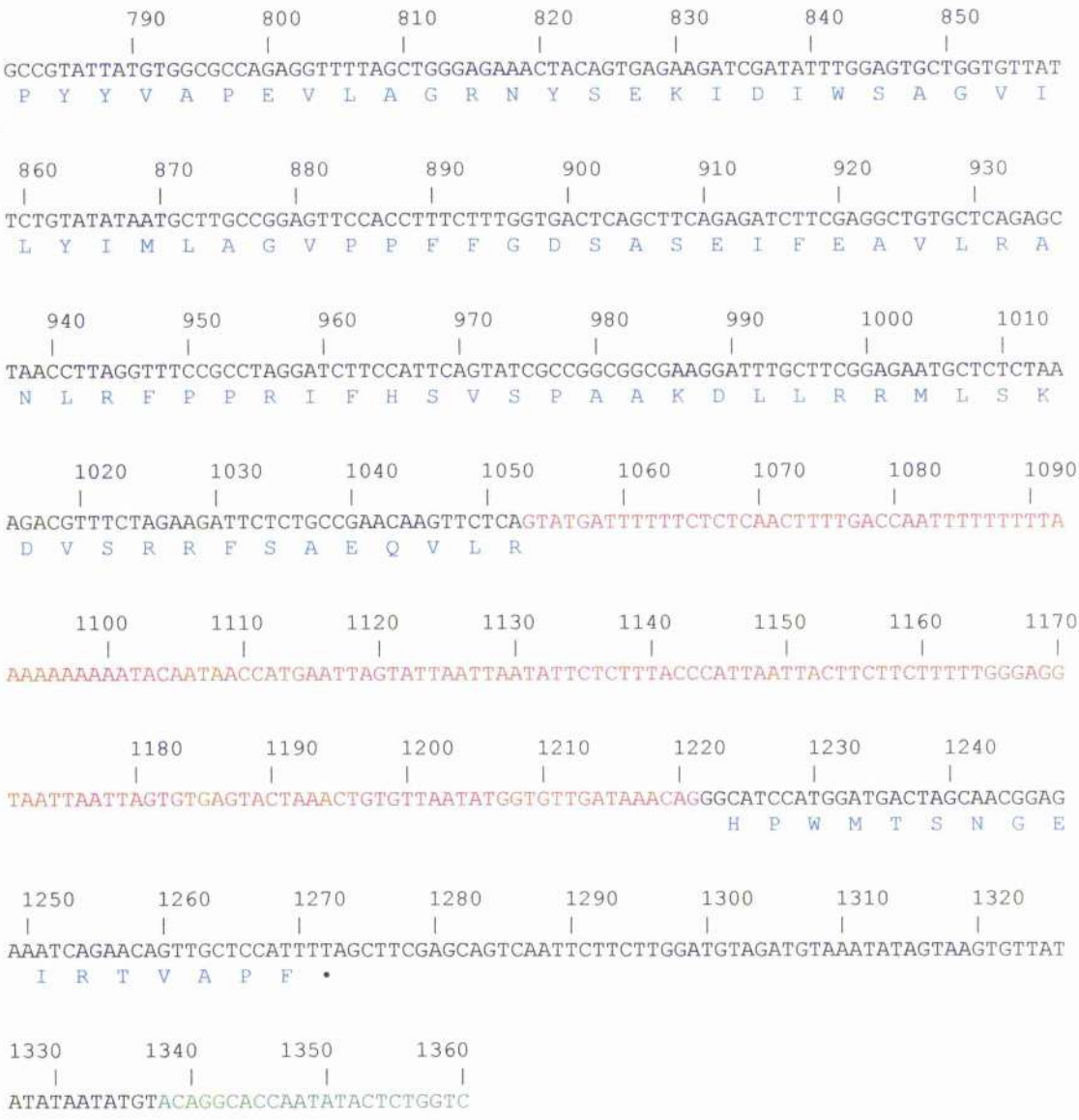

A
1 2 3 4 5 6 7
31 kDa
16 kDa
B
1 2 3 4 5 6 7
Realtive intensity %
4 159 12 126 5 100 0

Figure 4.4 Translation map of the genomic sequence of *L. esculentum PPCK2*

This shows the translation map of the genomic sequence of *LePPCK2* with annotations showing correct translation frame, primer sites, alternatively spliced regions and non-coding regions.


10 20 30 40 50 60 70
AGGAATTCGGCACGAGAAAAAACAAAACATACGGAATGAGTGAGTCACTGAAGCGAAATTATTGTGTGGGTGGAGAGT
M S E S L K R N Y C V G G E L
80 90 100 110 120 130 140 150
TAGGAAGGGGACGATTCGGCACCGTTTTCAAGTGTTACTCGTCGGCGACCGGTGAGCTATTCGCCGTCAAGTCTATCG
G R G R F G T V F K C Y S S A T G E L F A V K S I D
160 170 180 190 200 210 220 230
ACAAGCGCCTAATCGCCGATGACGCCATAGATCGACAATGTCTTTACAACGAAGCGAAAATTATGCACTTGCTTTCGC
K R L I A D D A I D R Q C L Y N E A K I M H L L S P
240 250 260 270 280 290 300 310
CAAATCCTAATGTTGTTCGTGTTTTCGATATCTACGAGGACGATACTCACCTTGATATGGTTTTAGAGCTCTGCAACT
N P N V V R V F D I Y E D D T H L D M V L E L C N S
320 330 340 350 360 370 380 390
CCGGCGATTTGTTCCAACGCCTCAGTAGTCAACCGATTTTCTCAGAGTCCGACGCTGTTGATGTCATGGTACTCTTCC
G D L F Q R L S S Q P I F S E S D A V D V M V L F Q
400 410 420 430 440 450 460
AGATCACCAAAATCTCACTCGTTGCTGATCAAAATGGGGAAATTTTACTTTTGCTTCGTAAAATGGCTTGAATCAGGA
I T K I S L V A D Q N G E I L L L L R K M A •
470 480 490 500 510 520 530 540
ACGATCTACGATGCAGCACGTTGTTGAATTTTCGAAAATTGGAGGAGAAAGTAAATTGTAAAGCTAAGCGAAGTAGAA
550 560 570 580 590 600 610 620
TTCATAATCGGCTTTTATTTTAGTTTTTCGTCATTTCTGATAAATTTTATGCTTAATGTTTTTTATAGGTACCACTAAT
V P L M
630 640 650 660 670 680 690 700
GAAAGCAATAGCGCACTGTCACCGTCTCGGTGTAGCCCACAGGGACATCAAACCGGATAACATTTTATTCACCGACTC
K A I A H C H R L G V A H R D I K P D N I L F T D S
710 720 730 740 750 760 770
780
TAACGAACTGAAATTGGCTGATTTTGGATCGGCGGAGTGTTTCCATGAGGGCCAGCTGATGAGCGGCGTGGTTGGAAC
N E L K L A D F G S A E C F H E G Q L M S G V V G T



Key to figure

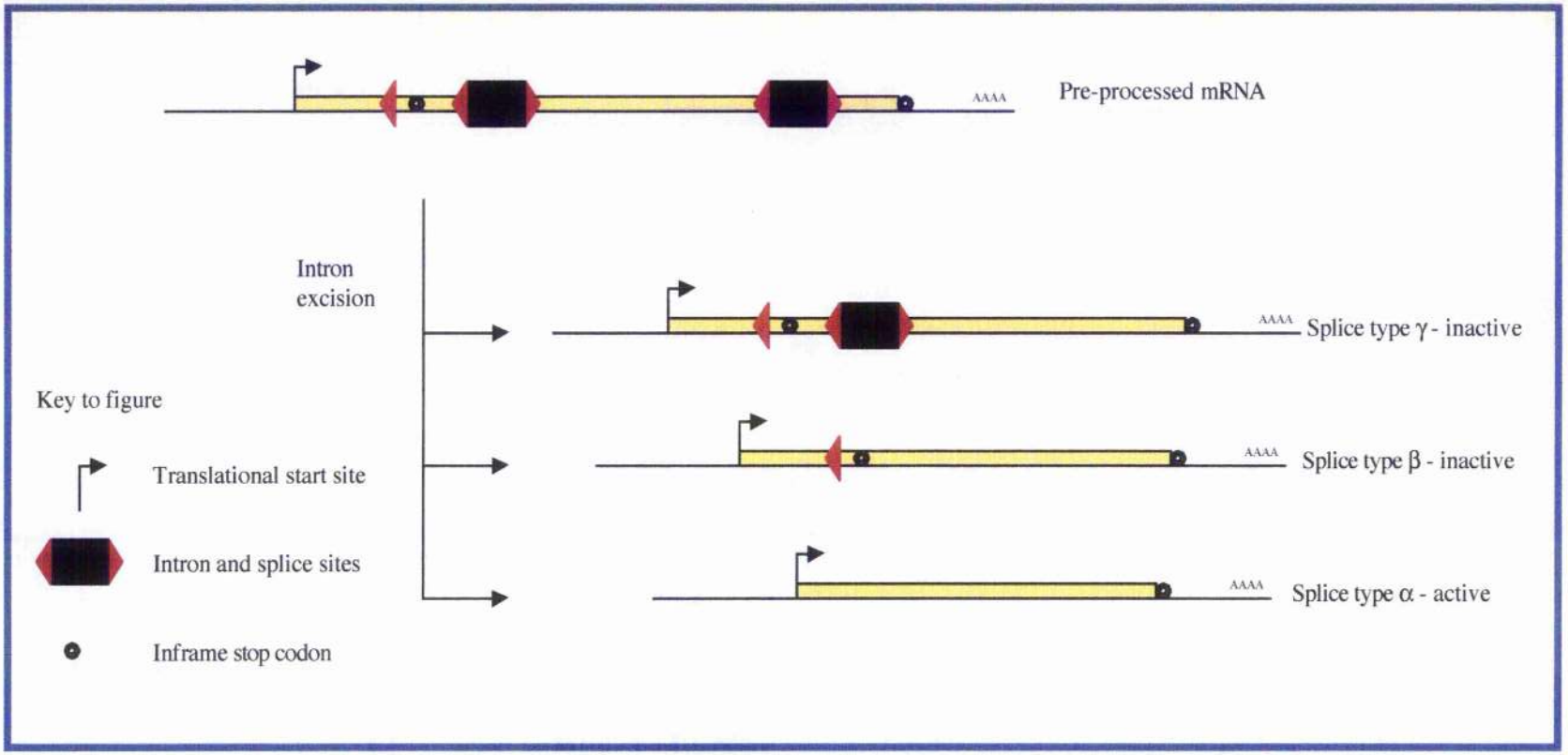
C E E I	Translation frame
ATGTTTCAAA	Coding region
CAAAATGGG	Alternate coding region
TTCAATTTCT	Non-coding region
TTCGGCACTT	Primer sites

Figure 4.5 Cartoon of the *Le PPCK 2* gene and transcript products

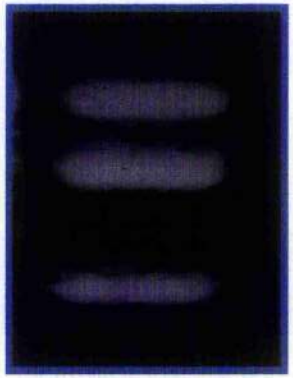
Panel A. shows a schematic of *LePPCK2* pre-processed mRNA and the three observed splice variants.

Panel B. shows a PCR analysis of full length LePPCK2 transcripts from fruit cDNA.

A



B



LePPCK2 transcript - γ

LePPCK2 transcript - β

LePPCK2 transcript - α

Figure 4.6 Classification of stages in tomato fruit development and ripening

Arrow indicates start of ethylene induced ripening

Early development

Expansion phase I:
0.5-2cm diameter

Expansion phase II:
2-5cm diameter

Expansion phase III
5-8cm diameter

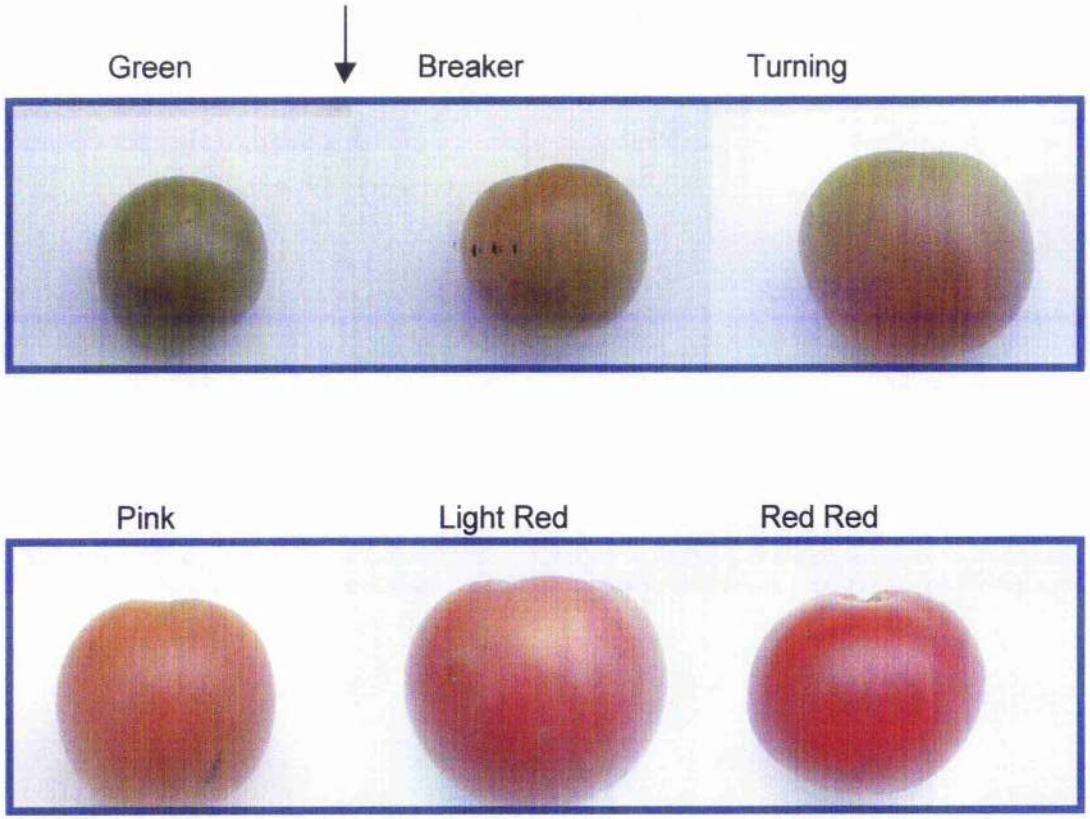


Figure 4.7 RT-PCR of LePPCK1, LePPCK2 and LeActin52 transcripts from tomato tissue

Panel A shows RT-PCR (35 cycles) of LePPCK1 transcripts from a number of tissues.

Lane (1) expansion phase 1 fruit

Lane (2) expansion phase 2 fruit

Lane (3) green fruit

Lane (4) breaker fruit

Lane (5) pink fruit

Lane (6) red-ripe fruit

Lane (7) young leaves

Lane (8) mature leaves

Lane (9) flowers

Lane (10) seedlings in light

Lane (11) seedlings in dark

Lane (12) roots

Panel B shows RT-PCR (35 cycles) of LePPCK2 transcripts. Lanes are loaded as in panel A.

Panel C shows RT-PCR (35 cycles) of LeActin52 transcripts. Lanes are loaded as in panel A.

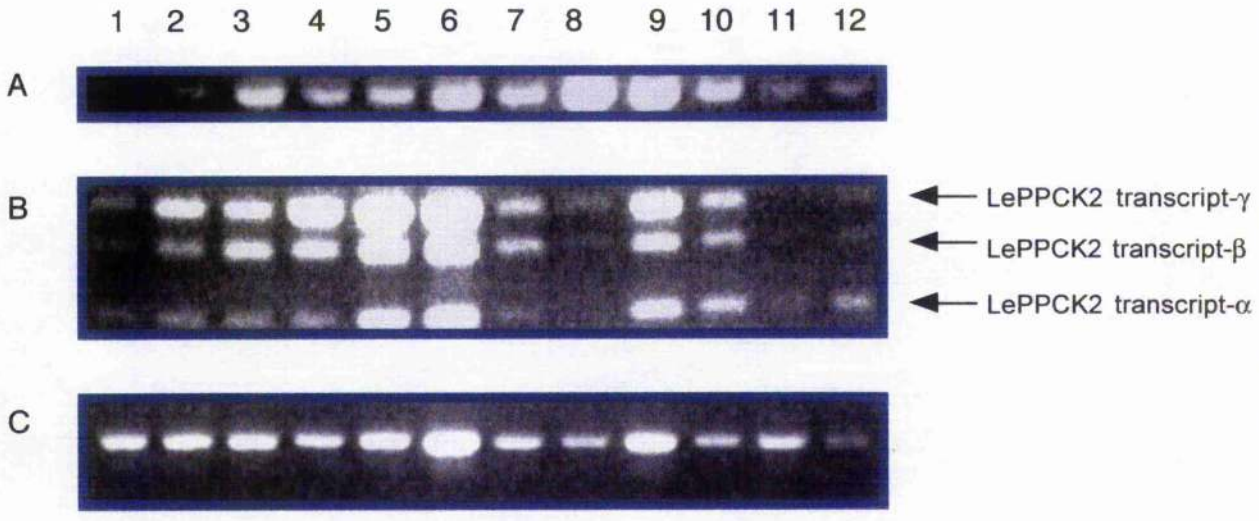

1 2 3 4 5 6 7 8 9 10 11 12
A
B
C
LePPCK2 transcript-γ
LePPCK2 transcript-β
LePPCK2 transcript-α

Figure 4.8 RT-PCR of LePPCK1, LePPCK2 and LeUbiquitin transcripts from fruit compartments

Panel A shows the RT-PCR of Le PPCK1 (40 cycles) from separate compartments of a red-ripe tomato fruit.

Lane (1) tomato skin

Lane (2) tomato outer pericarp

Lane (3) tomato inner pericarp

Lane (4) tomato locule

Lane (5) tomato developing seeds

Panel B shows the RT-PCR of Le PPCK 2 transcripts (40 cycles). Lanes are loaded as in Panel A.

Panel C shows the RT-PCR of LeUbiquitin transcripts (30 cycles). Lanes are loaded as in Panel A.

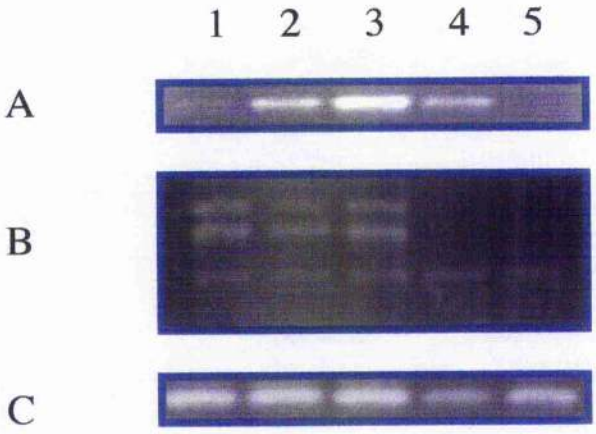

1 2 3 4 5
A
B
C

Figure 4.9 Alignment of the aminoacid sequences of potato and tomato PPCKs

Panel A shows the alignment of potato and tomato PPCK1.

Panel B shows the alignment of potato and tomato PPCK2.

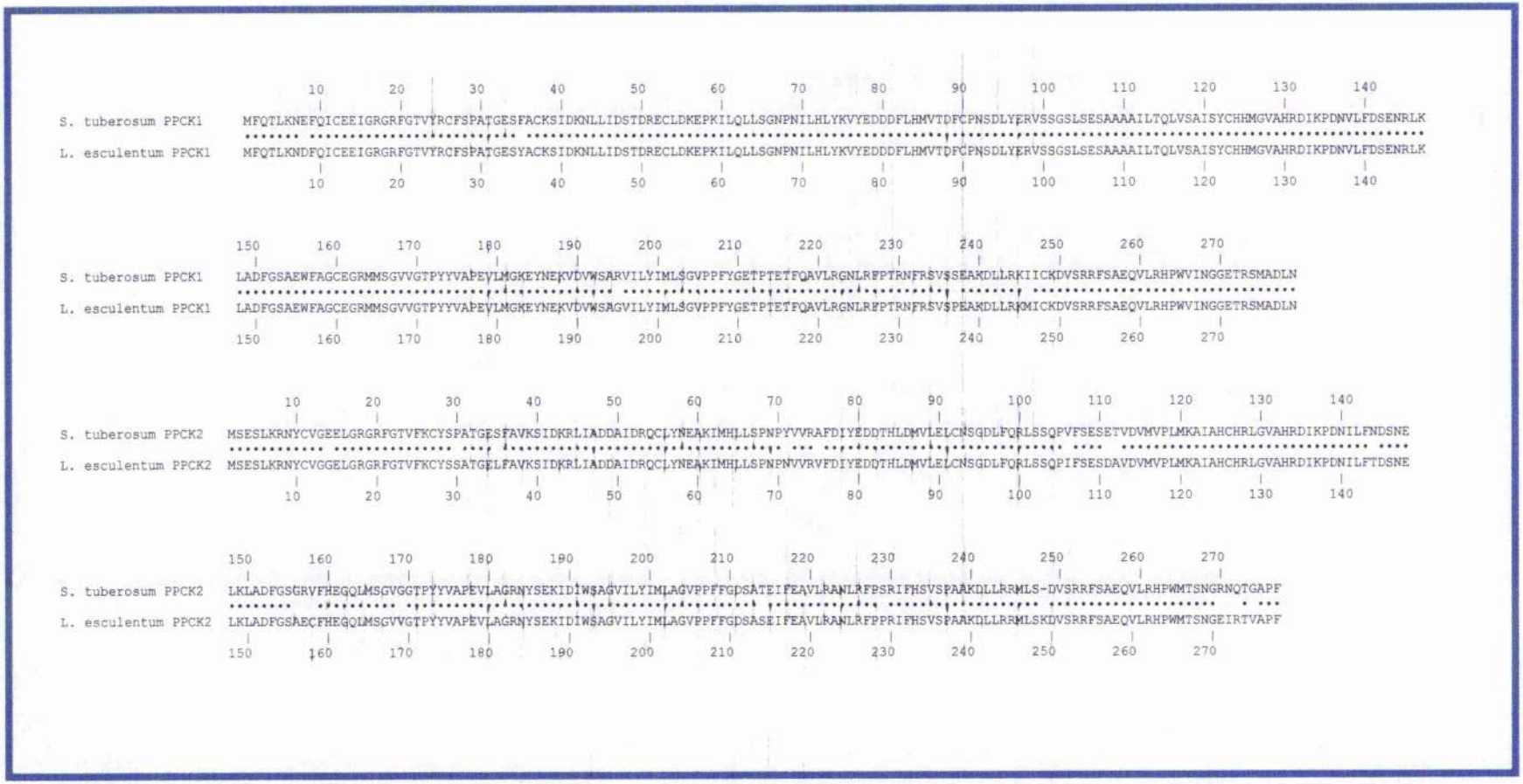

S. tuberosum PPCK1
L. esculentum PPCK1
S. tuberosum PPCK2
L. esculentum PPCK2

Figure 4.10 PCR analysis of potato PPCK2

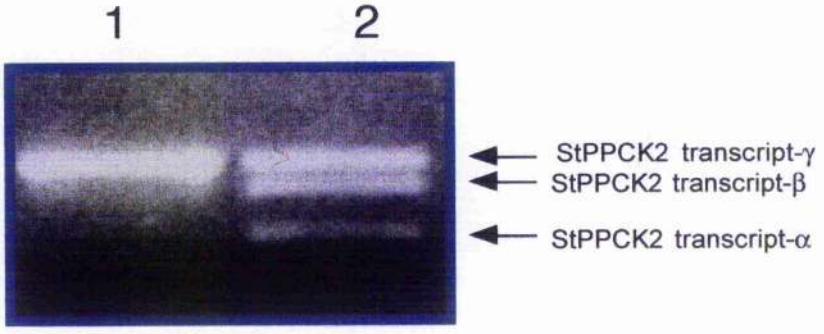
Panel A shows a PCR analysis of potato PPCK DNA and cDNA.

Lane (1) genomic DNA

Lane(2) cDNA

Panel B shows a cartoon of *StPPCK2* and primer sites.

A


1
2
StPPCK2 transcript-γ
StPPCK2 transcript-β
StPPCK2 transcript-α

B

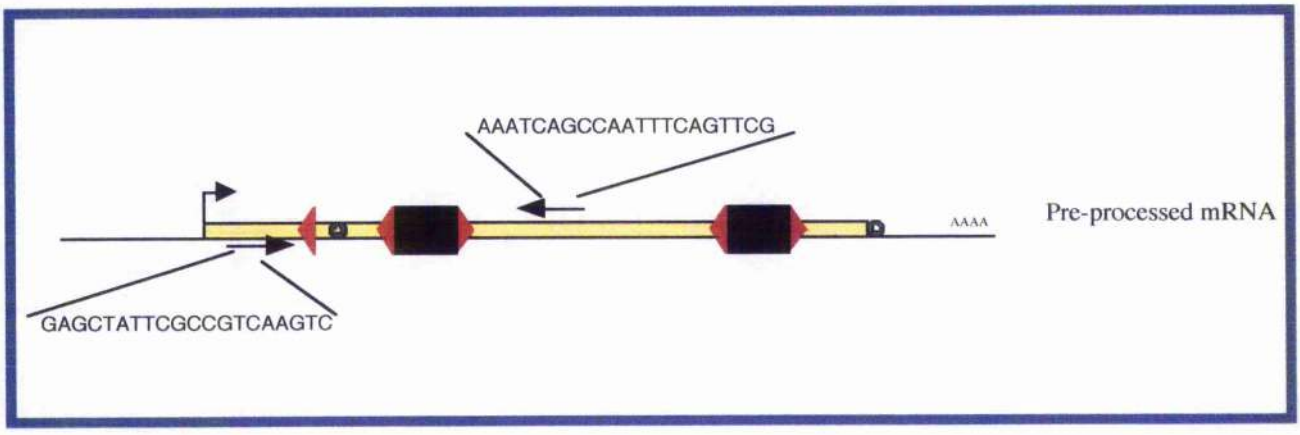
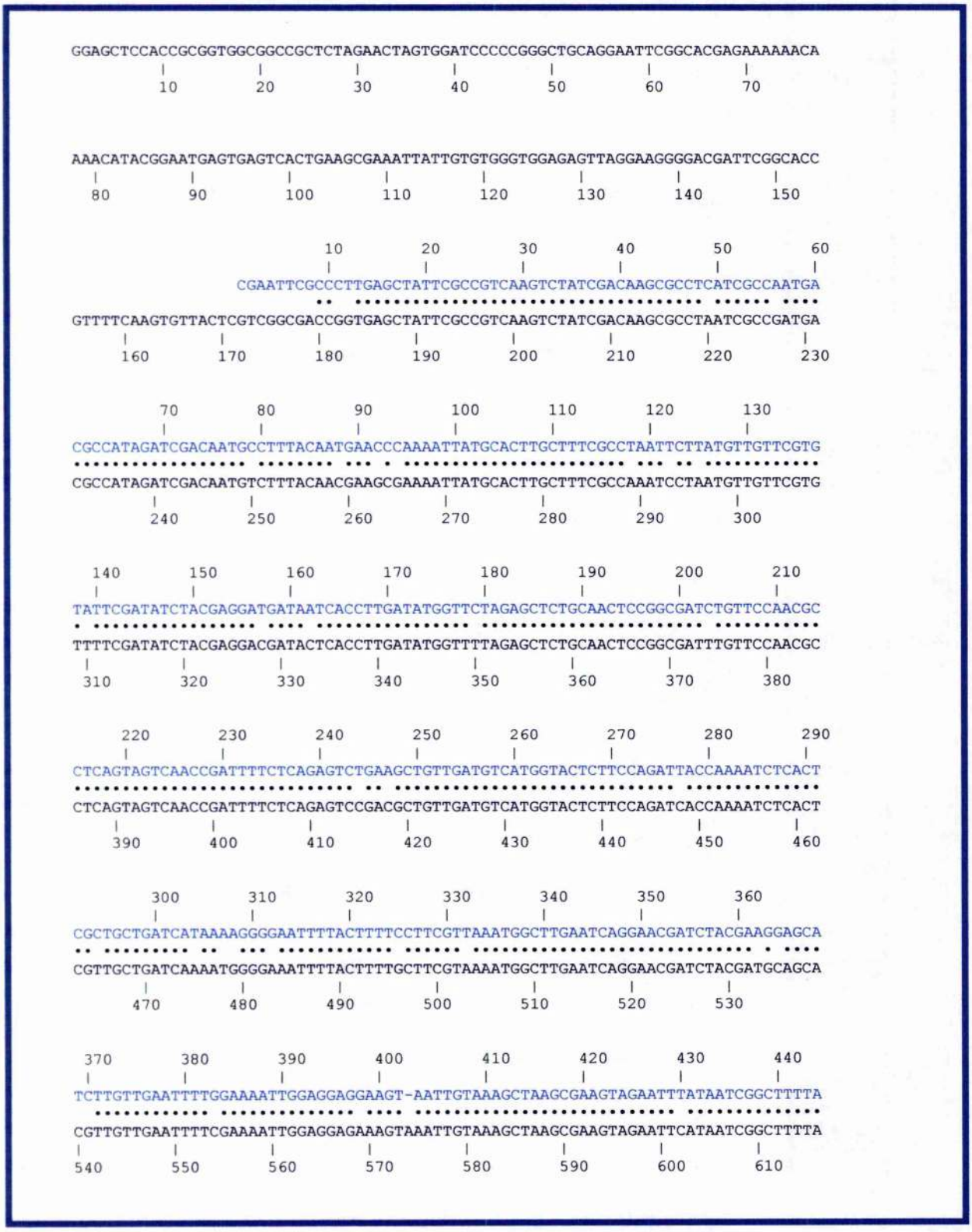
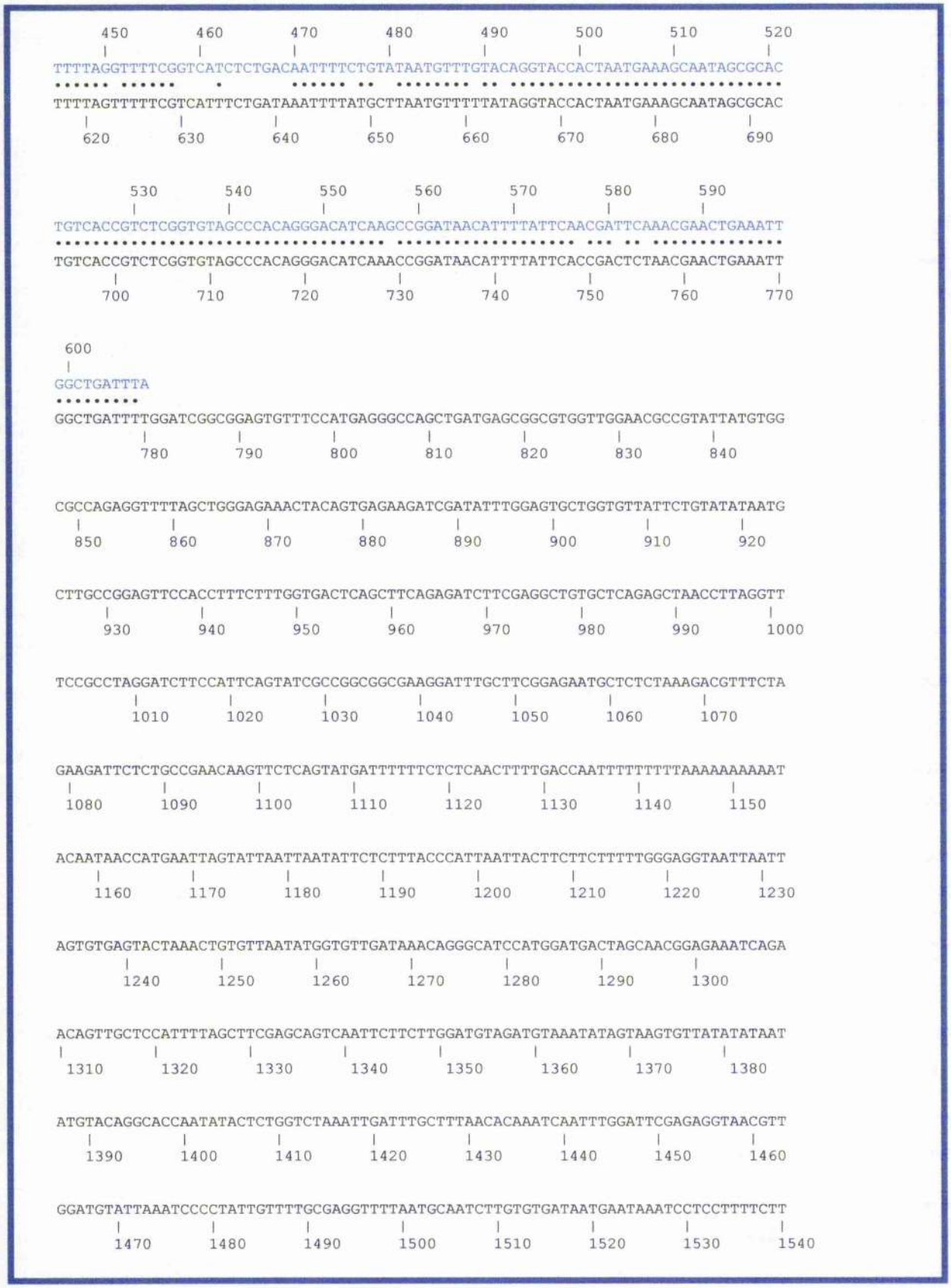

AAATCAGCCAATTTCAGTTCG
AAAA
Pre-processed mRNA
GAGCTATTCGCCGTCAAGTC

Figure 4.11 Alignment of the genomic nucleotide sequences of aubergine and tomato PPCK2

The blue font shows the partial nucleotide sequence of aubergine PPCK2. the black font the full-length nucleotide sequence of tomato PPCK2. The red font indicates the 5' and 3' splice sites.


GGAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGGCACGAGAAAAAACA
10 20 30 40 50 60 70
AAACATACGGAATGAGTGAGTCACTGAAGCGAAATTATTGTGTGGGTGGAGAGTTAGGAAGGGGACGATTCGGCACC
80 90 100 110 120 130 140 150
10 20 30 40 50 60
CGAATTCGCCCTTGAGCTATTCGCCGTCAAGTCTATCGACAAGCGCCTCATCGCCAATGA
GTTTTCAAGTGTTACTCGTCGGCGACCGGTGAGCTATTCGCCGTCAAGTCTATCGACAAGCGCCTAATCGCCGATGA
160 170 180 190 200 210 220 230
70 80 90 100 110 120 130
CGCCATAGATCGACAATGCCTTTACAATGAACCCAAAATTATGCACTTGCTTTCGCCTAATTCTTATGTTGTTCGTG
CGCCATAGATCGACAATGTCTTTACAACGAAGCGAAAATTATGCACTTGCTTTCGCCAAATCCTAATGTTGTTCGTG
240 250 260 270 280 290 300
140 150 160 170 180 190 200 210
TATTCGATATCTACGAGGATGATAATCACCTTGATATGGTTCTAGAGCTCTGCAACTCCGGCGATCTGTTCCAACGC
TTTTCGATATCTACGAGGACGATACTCACCTTGATATGGTTTTAGAGCTCTGCAACTCCGGCGATTTGTTCCAACGC
310 320 330 340 350 360 370 380
220 230 240 250 260 270 280 290
CTCAGTAGTCAACCGATTTTCTCAGAGTCTGAAGCTGTTGATGTCATGGTACTCTTCCAGATTACCAAAATCTCACT
CTCAGTAGTCAACCGATTTTCTCAGAGTCCGACGCTGTTGATGTCATGGTACTCTTCCAGATCACCAAAATCTCACT
390 400 410 420 430 440 450 460
300 310 320 330 340 350 360
CGCTGCTGATCATAAAAGGGGAATTTTACTTTTCCTTCGTTAAATGGCTTGAATCAGGAACGATCTACGAAGGAGCA
CGTTGCTGATCAAAATGGGGAAATTTTACTTTTGCTTCGTAAAATGGCTTGAATCAGGAACGATCTACGATGCAGCA
470 480 490 500 510 520 530
370 380 390 400 410 420 430 440
TCTTGTTGAATTTTGGAAAATTGGAGGAGGAAGT-AATTGTAAAGCTAAGCGAAGTAGAATTTATAATCGGCTTTTA
CGTTGTTGAATTTTCGAAAATTGGAGGAGAAAGTAAATTGTAAAGCTAAGCGAAGTAGAATTCATAATCGGCTTTTA
540 550 560 570 580 590 600 610


450 460 470 480 490 500 510 520
TTTTAGGTTTTCGGTCATCTCTGACAATTTTCTGTATAATGTTTGTACAGGTACCACTAATGAAAGCAATAGCGCAC
TTTTAGTTTTTCGTCATTTCTGATAAATTTTATGCTTAATGTTTTTATAGGTACCACTAATGAAAGCAATAGCGCAC
620 630 640 650 660 670 680 690
530 540 550 560 570 580 590
TGTCACCGTCTCGGTGTAGCCCACAGGGACATCAAGCCGGATAACATTTTATTCAACGATTCAAACGAACTGAAATT
TGTCACCGTCTCGGTGTAGCCCACAGGGACATCAAACCGGATAACATTTTATTCACCGACTCTAACGAACTGAAATT
700 710 720 730 740 750 760 770
600
GGCTGATTTA
GGCTGATTTTGGATCGGCGGAGTGTTTCCATGAGGGCCAGCTGATGAGCGGCGTGGTTGGAACGCCGTATTATGTGG
780 790 800 810 820 830 840
CGCCAGAGGTTTTAGCTGGGAGAAACTACAGTGAGAAGATCGATATTTGGAGTGCTGGTGTTATTCTGTATATAATG
850 860 870 880 890 900 910 920
CTTGCCGGAGTTCCACCTTTCTTTGGTGACTCAGCTTCAGAGATCTTCGAGGCTGTGCTCAGAGCTAACCTTAGGTT
930 940 950 960 970 980 990 1000
TCCGCCTAGGATCTTCCATTCAGTATCGCCGGCGGCGAAGGATTTGCTTCGGAGAATGCTCTCTAAAGACGTTTCTA
1010 1020 1030 1040 1050 1060 1070
GAAGATTCTCTGCCGAACAAGTTCTCAGTATGATTTTTTCTCTCAACTTTTGACCAATTTTTTTTTAAAAAAAAAAT
1080 1090 1100 1110 1120 1130 1140 1150
ACAATAACCATGAATTAGTATTAATTAATATTCTCTTTACCCATTAATTACTTCTTCTTTTTGGGAGGTAATTAATT
1160 1170 1180 1190 1200 1210 1220 1230
AGTGTGAGTACTAAACTGTGTTAATATGGTGTTGATAAACAGGGCATCCATGGATGACTAGCAACGGAGAAATCAGA
1240 1250 1260 1270 1280 1290 1300
ACAGTTGCTCCATTTTAGCTTCGAGCAGTCAATTCTTCTTGGATGTAGATGTAAATATAGTAAGTGTTATATATAAT
1310 1320 1330 1340 1350 1360 1370 1380
ATGTACAGGCACCAATATACTCTGGTCTAAATTGATTTGCTTTAACACAAATCAATTTGGATTCGAGAGGTAACGTT
1390 1400 1410 1420 1430 1440 1450 1460
GGATGTATTAAATCCCCTATTGTTTTGCGAGGTTTTAATGCAATCTTGTGTGATAATGAATAAATCCTCCTTTTCTT
1470 1480 1490 1500 1510 1520 1530 1540

Chapter 5

The phylogenetic analysis of plant PPCK

5.1 Introduction

Since the start of my research on PPCK in higher plants a great deal of information has been obtained from the sequencing of plant genomes, namely those of the C_3 dicot *A. thaliana* and the C_3 monocot *O. sativa* (rice). In tandem with this many est sequencing projects have generated formidable amounts of data. This allowed our lab to identify several putative partial and full-length PPCK genes and cDNAs. The PPCK genes characterised in the literature to date are *AtPPCK1* and *2*, *Beta vulgaris PPCK* (sugar beet PPCK), *FtPPCK*, *KfPPCK* and *McPPCK* (Fontaine *et al.*, 2002, Hartwell *et al.*, 1999, Herl *et al.*, GenBank accession, Taybi *et al.*, 2000, Tsuchida *et al.*, 2001). Hence only limited phylogenetic analysis has been possible. My research has allowed the identification of several more PPCKs including *LePPCK1* and *2*, *StPPCK1* and *2*, *TaPPCK1* and *ZmPPCK2*. In this short chapter I show the relationships between the known PPCKs and many putative PPCKs that have been identified in our lab by myself and others.

5.2 Results

5.2.1 The identification of plant PPCKs from est databases

Database searches revealed the presence of a number of ests whose deduced aminoacid sequences bear a high similarity to those of already characterised PPCKs. Using a similar approach that I undertook for identifying *LePPCK1, 2* and *ZmPPCK2*, our lab has successfully sequenced and in some cases characterised a number of these PPCKs. These include three soybean (*Glycine max*) genes (*GmPPCK1, 2* and *3*) (Sullivan *et al.*, unpublished), one *Sorghum bicolor* gene (*SbPPCK*) (Fontaine *et al.*, unpublished) and one maize gene (*ZmPPCK1*) (Shenton *et al.*, unpublished). In combination, Stuart Sullivan and I have also identified full-length PPCK genes from rice (*OsPPCK 1, 2* and *3*) and *Brassica olerace* (*BoPPCK*) from genome sequence databases. Table 5 lists the GenBank numbers of these sequences. All these PPCKs bar *LePPCK2* and *StPPCK2* have a similar gene structure, namely one small intron close to 3' end of the coding sequence. The predicted molecular weights of these PPCKs are in the range 30-31 kDa. However *OsPPCK2* contains three putative translation start sites and may encode proteins of 31 kDa, 32.4 kDa or 36.5 kDa.

A number of partial sequences of PPCKs have also been identified by assembly of tentative est and/or genomic contigs. These include *Euphorbia esula PPCK* (*EePPCK*), *Gossypium arboreum PPCK* (*GaPPCK*), *Hordeum vulgare PPCK1* and *PPCK2* (*HvPPCK1* and *2*), *Lotus japonicus PPCK* (*LjPPCK*), *Medicago truncatula PPCK* (*MtPPCK*), *Musa acuminata PPCK* (*MaPPCK*), *Phaleolus vulgaris PPCK* (*PvPPCK*), *Triticum aestivum PPCK2* (*TaPPCK2*) and *Zea mays PPCK3* (*ZmPPCK3*). The sequences that make up these tentative contigs are listed in Table 5. An alignment of these putative PPCKs is shown in Figure 5.1.

5.2.2 The phylogenetic analysis of plant PPCK

This large set of full and partial length PPCKs allowed me to analyse the phylogenetic relationship of the plant PPCKs to other protein kinases. A phylogenetic tree was assembled using Clustal, see Figure 5.2. Several points of interest can be seen in this tree. The most striking is the separation of the monocot and dicot PPCKs. The monocots PPCKs themselves fall into two distinct groups. Monocot PPCK1 appears to form a separate branch to that of monocot PPCK2 and 3. This suggests that the monocot PPCK2 and 3 may result from a gene duplication. This clear separation suggests a different function and/or expression pattern for monocot PPCK1 compared to monocot PPCK2 and 3. However, it is difficult to assign function as of yet. A second observation is the unsurprising grouping of the nodule forming plant PPCKs, including GmPPCK1, 2 and 3, LjPPCK, MtPPCK and PvPPCK. The LePPCKs and StPPCKs have aligned, as expected with each other, with the *Solanaceae* PPCK1 showing high similarity to the photosynthetic FtPPCK. One can speculate that *Solanaceae* PPCK1 may be a leaf enhanced form. This would agree with the data obtained from RT-PCR as discussed in section 4.2.

5.2.3 The moncot plant PPCKs show high GC content

The analysis of plant PPCKs showed that monocot PPCK open reading frames are much more GC-rich than the open reading frames of the dicot PPCK genes. This trait has been noticed before with monocot genes averaging about 51 % GC in contrast to the 41 % normally found in dicot open reading frames (Yu *et al.*, 2002). It has been suggested that this contributes to the hypothesis for recognition of exons and introns. Introns are AT-rich and exons are GC-rich and this might aid intron recognition and splicing in monocot plants. The recent sequencing of the monocot rice genome has allowed a study of this trait with the sequencing revealing that exons of rice tend to have a range of GC content, from 50 to 60 %. Higher GC content is not

normally seen. The GC content of PPCK and PEPc was analysed with figure 5.2 showing a graphical representation of these results. The graphs show that monocot PPCK is 67-75 % GC-rich while monocot PEPc is 50-64 % GC-rich. This suggests that PPCK genes in monocots are unusual and possibly unique in nature.

5.3 Discussion

The phylogentic tree allows us to make a number of predictions. The first is that the monocots contain three PPCK genes. This is true in rice, for which essentially the whole genome sequence is now available, and is probably true for maize, given that our analysis of the extensive maize est database has reached three distinct putative PPCK genes. Hence we would expect, for example, barley to contain homologues of *PPCK2* and *PPCK3*, and wheat to contain a homologue of *PPCK2* (note that the est termed *TaPPCK2* is actually in the *PPCK3* subfamily). In terms of gene numbers, the dicots fall into two classes. *Arabidopsis* contains only two PPCK genes, while soybean contains at least three. Two of these, *GmPPCK2* and *3*, are both very similar to each other and most heavily expressed in nodules. Given the number of nodule ests now available from the model legumes *Lotus japonicus* and *Medicago truncatula,* it is interesting that the ests available for these species seem to derive from a single gene in each case. It is possible that there has been a gene duplication in soybean that has not occurred in other legumes. Hence it is too early to predict whether all legumes contain three PPCK genes.

For several species only one PPCK gene has been reported to date. For example, although Taybi *et al.* (2000) presented very preliminary evidence for the existence of a second PPCK gene in *M. crystallinum*, this has resisted all attempts at cloning (J. Hartwell, personal communication). Given the similarity between PPCKs and CDPKs, it is possible that the data of Taybi *et al.* (2000) were not interpreted correctly. Given the est sequencing programme and the prospects for a genome sequencing programme in *M. crystallinum,* it may be possible to resolve the number of PPCK genes in this species in the near future. In any event, it is important to ask why there is a difference in PPCK gene number between some plants. This will require careful analysis of the tissue expression pattern, control of expression and function of the different genes.

The GC-richness that has been observed in monocot plant PPCK is unusual and at the moment unexplainable. This trait may have no importance at all, other than to aid exon/intron definition. Alternatively GC-richness may aid secondary structure

formation in the RNA. This could in theory be involved in recognition and control by an as yet unidentified protein involved in the control of four carbon acid synthesis, such as the amino acid biosynthesis. This theory could be investigated by plant genetic modification. For example, by replacing the non GC-rich PPCK gene from *Arabidopsis* for a GC-rich PPCK gene from rice with the *Arabidopsis* PPCK promoter, expression studies could then determine whether this causes any differences in comparison to the wild type. This experiment though would face a number of difficulties unless a homologous recombination technique was developed for plants.

Table 5: Nomenclature and origin of known and putative PPCK sequences

Table 5a: Completed sequences

Species name	Gene nomenclature	GenBank accession of the full-length sequence	Status of proof	Authors
A.thaliana	*AtPPCK1*	AF162660 BACF22013	Full gDNA and cDNA sequences and PPCK activity shown	Fontaine *et al.*, 2002, Hartwell *et al.*, 1999
A.thaliana	*AtPPCK2*	AF358915 BACT27C4	Full gDNA and cDNA sequences and PPCK activity shown	Fontaine *et al.*, 2002
B.vulgaris	*BvPPCK*	AJ309171	Full cDNA sequence	Kloos *et al.*, (2002)
F.trinervia	*FtPPCK*	AB065100	Full cDNA sequence and PPCK activity shown	Tsuchida *et al.*, (2001)
G.max	*GmPPCK1*	AY144181	Full gDNA and cDNA sequences	Sullivan *et al.*, unpublished
G.max	*GmPPCK2*	AY144183	Full gDNA and cDNA sequences and PPCK activity shown	Sullivan *et al.*, unpublished
G.max	*GmPPCK3*	AY144185	Full gDNA and cDNA sequences	Sullivan *et al.*, unpublished
K.fedtschenkoi	*KfPPCK*	AF162661 AF162662	Full gDNA and cDNA sequences and PPCK activity shown	Hartwell *et al.*, 1999
L.esculentum	*LePPCK1*	See table 3	Full gDNA and cDNA sequences and PPCK activity shown	Marsh *et al.*, this thesis
L.esculentum	*LePPCK2*	See table 3	Full gDNA and cDNA sequences and PPCK activity shown	Marsh *et al.*, this thesis
M.crystallinum	*McPPCK*	AF158091	Full cDNA sequence and PPCK activity shown	Taybi *et al.*, 2000
O.sativa	*OsPPCK1*	AAAA01001629 (Genomic clone)	Full gDNA sequence identified on a single genomic clone	
O.sativa	*OsPPCK2*	AAAA01019193 (Genomic clone)	Full gDNA sequence identified on a single genomic clone	
S.bicolor	*SbPPCK*	AF399915	Full gDNA and cDNA sequences and PPCK activity shown	Fontaine *et al.*, unpublished

S.tuberosum	*StPPCK1*	See table 4	Full gDNA sequence identified	Marsh *et al.*, this thesis
S.tuberosum	*StPPCK2*	See table 4	Full cDNA sequence identified on a single est	
T. aestivum	*TaPPCK1*	BE515532	Full cDNA sequence	Marsh *et al.*, this thesis
		BF484446		
Z.mays	*ZmPPCK1*	See table 2	Full cDNA sequence	Shenton *et al.*, unpublished
Z.mays	*ZmPPCK2*	See table 2	Full gDNA and cDNA sequences and PPCK activity shown	Marsh *et al.*, this thesis

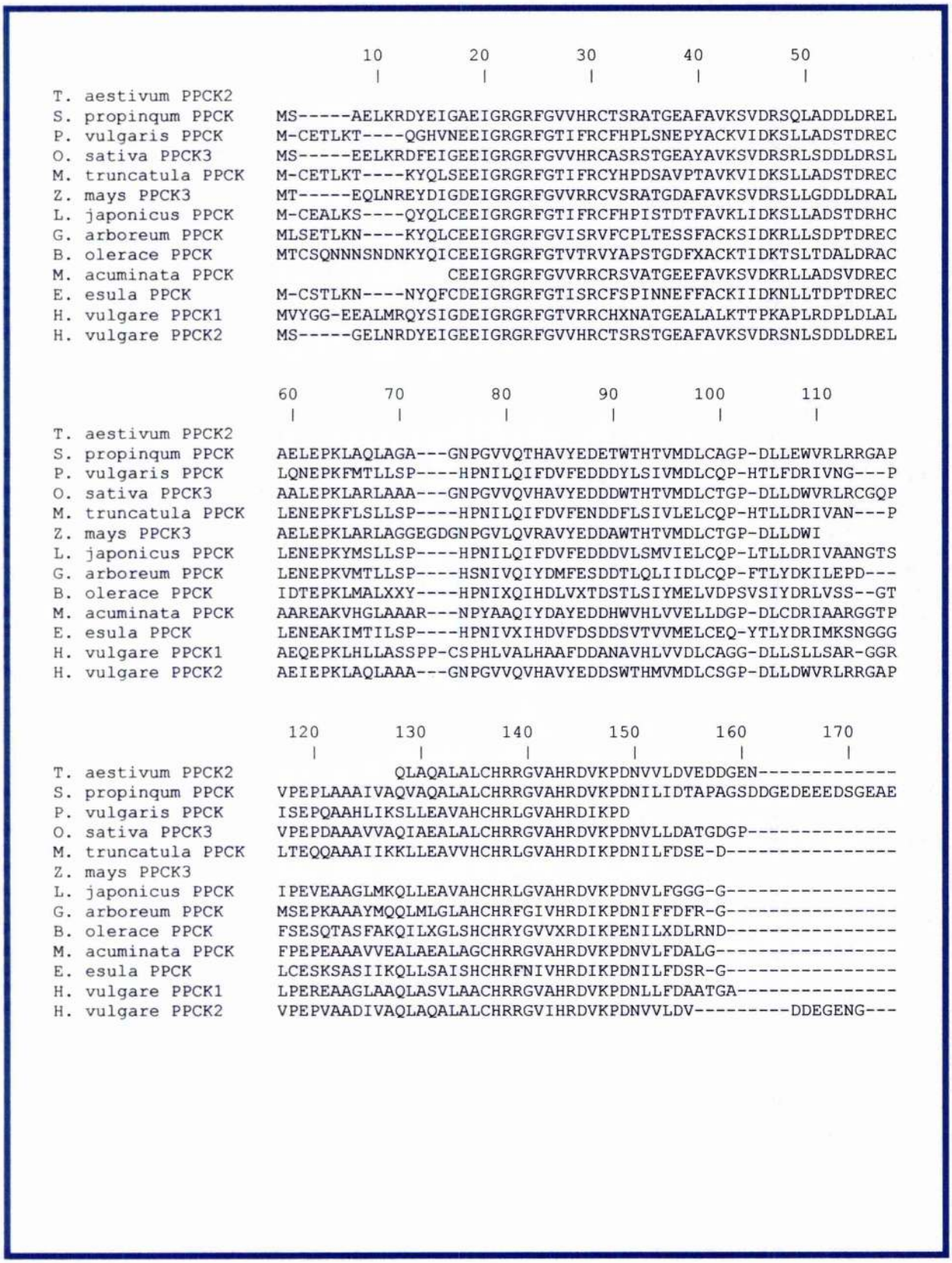
Table 5b: Incomplete sequences derived from est and genomic sequencing programs

Species name	Gene nomenclature	GenBank accession of the full-length sequence	Notes
B. oleracea	*BoPPCK*	BH688180	Contig of 5 overlapping genomic fragments gives a full length sequence of 1920 bp, including 5' and 3' untranslated regions and an intron close to the 3' end. Putative PPCK molecular weight is 32.5 kDa.
		BH445482	
		BH455517	
		BH531334	
		BH607126	
E. esula	*EePPCK*	BG354958	Single est of 565 bp of partial sequence.
G. arboreum	*GaPPCK*	BE053633	Contig of 2 overlapping ests giving 'convincing' sequence of 650 bp. Additional 3' sequence seemed to contain many errors.
		BF277462	
H. vulgare	*HvPPCK1*	BE454588	BE454588 contains 'convincing' sequence up to 776 bp. The sequence seems to contain errors beyond this point. The other ests are very similar to BE454588 but have large deletions and have therefore been ignored in prediction of aminoacid sequence.
		BE231198	
		BM368735	
		AI503765	
		BE413415	
		BE602544	
H. vulgare	*HvPPCK2*	BJ465402	Contig of 5 overlapping ests gives 723 bp of 'convincing' sequence.
		BJ468646	
		BI779702	
		BI780918	
		BM374766	
L. japonicus	*LjPPCK*	AV425263	Contig of 9 overlapping ests gives 535 bp of 'convincing' sequence.
		AW719293	
		AW720549	
		BI416343	
		BI418379	
		BI419338	
		BI419762	
		BI419889	
		BI420586	
M. acuminata	*MaPPCK*	AF203483	Single est of 500 bp of 'convincing' sequence.
M. truncatula	*MtPPCK*	AW574075	Contig of 5 overlapping ests gives 821 bp of 'convincing' sequence.

		AW688391	
		BF637934	
		BG583672	
		AW070003	
O. sativa	*OsPPCK3*	AAAA01027319 (Genomic clone)	Contig of 3 genomic fragments and an est gives a full sequence of 4016 bp including 5' and 3' untranslated regions and one intron close to the 3' end of the coding sequence. Putative PPCK molecular weight is 30.6 kDa.
		AAAA01028858 (Genomic clone)	
		AAAA01088245 (Genomic clone)	
		AU095289	
P. vulgaris	*PvPPCK*	BQ481874	Single est of 432 bp of 'convincing' sequence.
S. propinquum	*SpPPCK*	BE918643	Contig of 8 ests matching the full-length SbPPCK. Putative PPCK molecular weight is 32.5 kDa.
		BE918996	
		BG558033	
		BG558320	
		BG558617	
		BG558697	
		BG560592	
		BG560660	
T. aestivum	*TaPPCK2*	BE517315	The 3' end of the contig seems to represent an unprocessed transcript as it contains an intron as found in other PPCK genes. This intron has been ignored in the prediction of aminoacid sequence.
		BF484446	
		BM134904	
Z. mays	*ZmPPCK3*	See table 2	Contig of 5 overlapping genomic fragments gives 345 bp of 'convincing' sequence.

'Convincing' sequence means that the predicted aminoacid sequence is very similar to other PPCKs without frameshifts, additions, deletions, or inframe stop codons.

Figure 5.1 Alignment of the incomplete aminoacid sequences of *B. olerace* PPCK, *E. esula* PPCK, *G. arboreum* PPCK, *H. vulgare* PPCK1, *H. vulgare* PPCK2, *L. japonicus* PPCK, *Musa acuminata* PPCK, *M. truncatula* PPCK, *O. sativa* PPCK3, *P. vulgaris* PPCK, *S. propinqum* PPCK, *T. aestivum* PPCK2 and *Z. mays* PPCK3


10 20 30 40 50
T. aestivum PPCK2
S. propinqum PPCK MS-----AELKRDYEIGAEIGRGRFGVVHRCTSRATGEAFAVKSVDRSQLADDLDREL
P. vulgaris PPCK M-CETLKT----QGHVNEEIGRGRFGTIFRCFHPLSNEPYACKVIDKSLLADSTDREC
O. sativa PPCK3 MS-----EELKRDFEIGEEIGRGRFGVVHRCASRSTGEAYAVKSVDRSRLSDDLDRSL
M. truncatula PPCK M-CETLKT----KYQLSEEIGRGRFGTIFRCYHPDSAVPTAVKVIDKSLLADSTDREC
Z. mays PPCK3 MT-----EQLNREYDIGDEIGRGRFGVVRRCVSRATGDAFAVKSVDRSLLGDDLDRAL
L. japonicus PPCK M-CEALKS----QYQLCEEIGRGRFGTIFRCFHPISTDTFAVKLIDKSLLADSTDRHC
G. arboreum PPCK MLSETLKN----KYQLCEEIGRGRFGVISRVFCPLTESSFACKSIDKRLLSDPTDREC
B. olerace PPCK MTCSQNNNSNDNKYQICEEIGRGRFGTVTRVYAPSTGDFXACKTIDKTSLTDALDRAC
M. acuminata PPCK CEEIGRGRFGVVRRCRSVATGEEFAVKSVDKRLLADSVDREC
E. esula PPCK M-CSTLKN----NYQFCDEIGRGRFGTISRCFSPINNEFFACKIIDKNLLTDPTDREC
H. vulgare PPCK1 MVYGG-EEALMRQYSIGDEIGRGRFGTVRRCHXNATGEALALKTTPKAPLRDPLDLAL
H. vulgare PPCK2 MS-----GELNRDYEIGEEIGRGRFGVVHRCTSRSTGEAFAVKSVDRSNLSDDLDREL
60 70 80 90 100 110
T. aestivum PPCK2
S. propinqum PPCK AELEPKLAQLAGA---GNPGVVQTHAVYEDETWTHTVMDLCAGP-DLLEWVRLRRGAP
P. vulgaris PPCK LQNEPKFMTLLSP----HPNILQIFDVFEDDDYLSIVMDLCQP-HTLFDRIVNG---P
O. sativa PPCK3 AALEPKLARLAAA---GNPGVVQVHAVYEDDDWTHTVMDLCTGP-DLLDWVRLRCGQP
M. truncatula PPCK LENEPKFLSLLSP----HPNILQIFDVFENDDFLSIVLELCQP-HTLLDRIVAN---P
Z. mays PPCK3 AELEPKLARLAGGEGDGNPGVLQVRAVYEDDAWTHTVMDLCTGP-DLLDWI
L. japonicus PPCK LENEPKYMSLLSP----HPNILQIFDVFEDDDVLSMVIELCQP-LTLLDRIVAANGTS
G. arboreum PPCK LENEPKVMTLLSP----HSNIVQIYDMFESDDTLQLIIDLCQP-FTLYDKILEPD---
B. olerace PPCK IDTEPKLMALXXY----HPNIXQIHDLVXTDSTLSIYMELVDPSVSIYDRLVSS--GT
M. acuminata PPCK AAREAKVHGLAAAR---NPYAAQIYDAYEDDHWVHLVVELLDGP-DLCDRIAARGGTP
E. esula PPCK LENEAKIMTILSP----HPNIVXIHDVFDSDDSVTVVMELCEQ-YTLYDRIMKSNGGG
H. vulgare PPCK1 AEQEPKLHLLASSPP-CSPHLVALHAAFDDANAVHLVVDLCAGG-DLLSLLSAR-GGR
H. vulgare PPCK2 AEIEPKLAQLAAA---GNPGVVQVHAVYEDDSWTHMVMDLCSGP-DLLDWVRLRRGAP
120 130 140 150 160 170
T. aestivum PPCK2 QLAQALALCHRRGVAHRDVKPDNVVLDVEDDGEN----------
S. propinqum PPCK VPEPLAAAIVAQVAQALALCHRRGVAHRDVKPDNILIDTAPAGSDDGEDEEEDSGEAE
P. vulgaris PPCK ISEPQAAHLIKSLLEAVAHCHRLGVAHRDIKPD
O. sativa PPCK3 VPEPDAAAVVAQIAEALALCHRRGVAHRDVKPDNVLLDATGDGP----------
M. truncatula PPCK LTEQQAAAIIKKLLEAVVHCHRLGVAHRDIKPDNILFDSE-D----------
Z. mays PPCK3
L. japonicus PPCK IPEVEAAGLMKQLLEAVAHCHRLGVAHRDVKPDNVLFGGG-G----------
G. arboreum PPCK MSEPKAAAYMQQLMLGLAHCHRFGIVHRDIKPDNIFFDFR-G----------
B. olerace PPCK FSESQTASFAKQILXGLSHCHRYGVVXRDIKPENILXDLRND----------
M. acuminata PPCK FPEPEAAAVVEALAEALAGCHRRGVAHRDVKPDNVLFDALG----------
E. esula PPCK LCESKSASIIKQLLSAISHCHRFNIVHRDIKPDNILFDSR-G----------
H. vulgare PPCK1 LPEREAAGLAAQLASVLAACHRRGVAHRDVKPDNLLFDAATGA----------
H. vulgare PPCK2 VPEPVAADIVAQLAQALALCHRRGVIHRDVKPDNVVLDV---------DDEGENG---

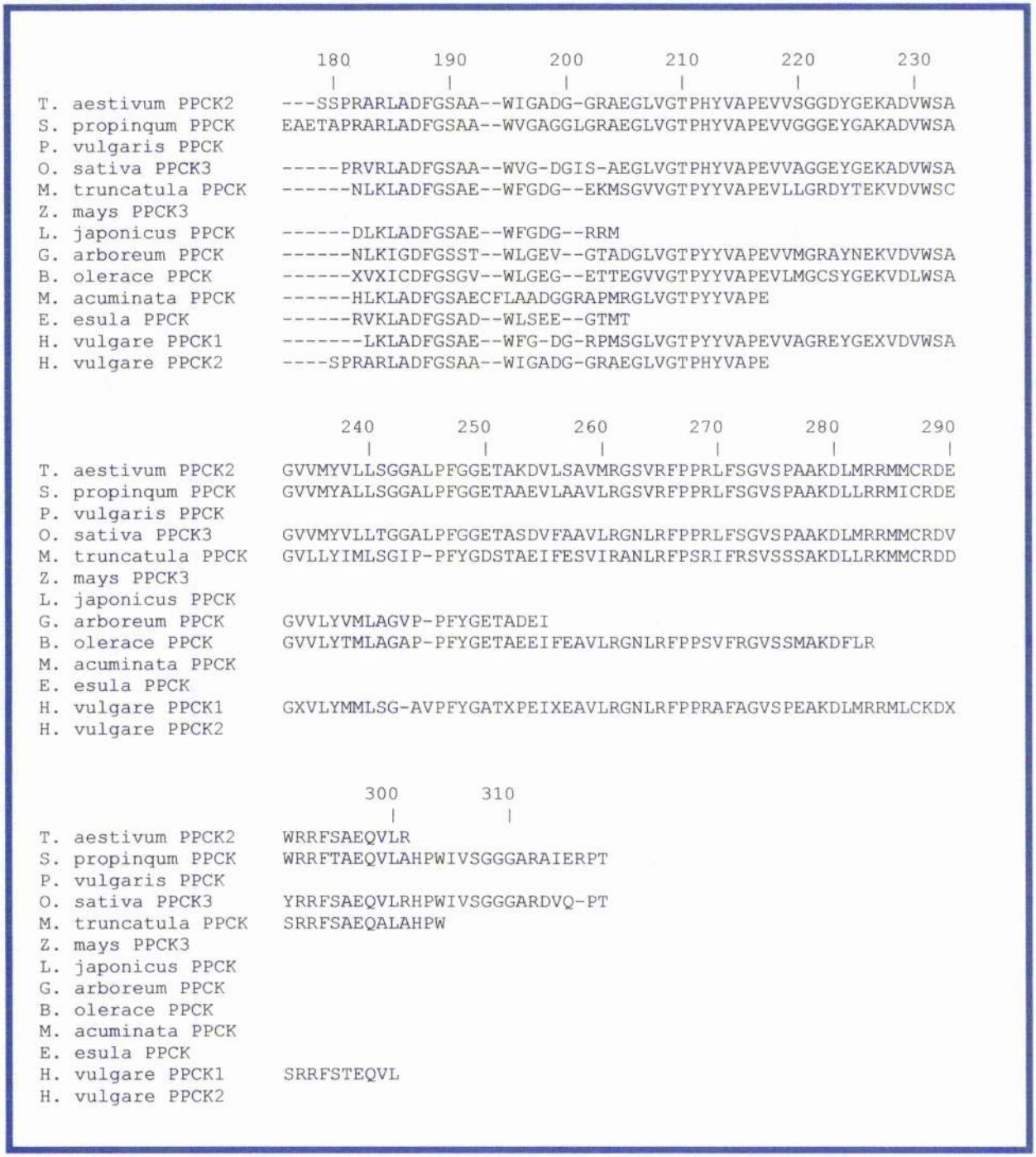

180 190 200 210 220 230
T. aestivum PPCK2 ---SSPRARLADFGSAA--WIGADG-GRAEGLVGTPHYVAPEVVSGGDYGEKADVWSA
S. propinqum PPCK EAETAPRARLADFGSAA--WVGAGGLGRAEGLVGTPHYVAPEVVGGGEYGAKADVWSA
P. vulgaris PPCK
O. sativa PPCK3 -----PRVRLADFGSAA--WVG-DGIS-AEGLVGTPHYVAPEVVAGGEYGEKADVWSA
M. truncatula PPCK ------NLKLADFGSAE--WFGDG--EKMSGVVGTPYYVAPEVLLGRDYTEKVDVWSC
Z. mays PPCK3
L. japonicus PPCK ------DLKLADFGSAE--WFGDG--RRM
G. arboreum PPCK ------NLKIGDFGSST--WLGEV--GTADGLVGTPYYVAPEVVMGRAYNEKVDVWSA
B. olerace PPCK ------XVXICDFGSGV--WLGEG--ETTEGVVGTPYYVAPEVLMGCSYGEKVDLWSA
M. acuminata PPCK ------HLKLADFGSAECFLAADGGRAPMRGLVGTPYYVAPE
E. esula PPCK ------RVKLADFGSAD--WLSEE--GTMT
H. vulgare PPCK1 -------LKLADFGSAE--WFG-DG-RPMSGLVGTPYYVAPEVVAGREYGEXVDVWSA
H. vulgare PPCK2 ----SPRARLADFGSAA--WIGADG-GRAEGLVGTPHYVAPE
240 250 260 270 280 290
T. aestivum PPCK2 GVVMYVLLSGGALPFGGETAKDVLSAVMRGSVRFPPRLFSGVSPAAKDLMRRMMCRDE
S. propinqum PPCK GVVMYALLSGGALPFGGETAAEVLAAVLRGSVRFPPRLFSGVSPAAKDLLRRMICRDE
P. vulgaris PPCK
O. sativa PPCK3 GVVMYVLLTGGALPFGGETASDVFAAVLRGNLRFPPRLFSGVSPAAKDLMRRMMCRDV
M. truncatula PPCK GVLLYIMLSGIP-PFYGDSTAEIFESVIRANLRFPSRIFRSVSSSAKDLLRKMMCRDD
Z. mays PPCK3
L. japonicus PPCK
G. arboreum PPCK GVVLYVMLAGVP-PFYGETADEI
B. olerace PPCK GVVLYTMLAGAP-PFYGETAEEIFEAVLRGNLRFPPSVFRGVSSMAKDFLR
M. acuminata PPCK
E. esula PPCK
H. vulgare PPCK1 GXVLYMMLSG-AVPFYGATXPEIXEAVLRGNLRFPPRAFAGVSPEAKDLMRRMLCKDX
H. vulgare PPCK2
300 310
T. aestivum PPCK2 WRRFSAEQVLR
S. propinqum PPCK WRRFTAEQVLAHPWIVSGGGARAIERPT
P. vulgaris PPCK
O. sativa PPCK3 YRRFSAEQVLRHPWIVSGGGARDVQ-PT
M. truncatula PPCK SRRFSAEQALAHPW
Z. mays PPCK3
L. japonicus PPCK
G. arboreum PPCK
B. olerace PPCK
M. acuminata PPCK
E. esula PPCK
H. vulgare PPCK1 SRRFSTEQVL
H. vulgare PPCK2

Figure 5.2 Phylogenetic analysis of PPCK and other protein kinases

Sequences of catalytic domains were aligned and the tree was constructed using Clustal X, where N=1000. The tree was displayed using TreeView. Human cyclic AMP-dependent kinase (PKA) and protein kinase C (PKC), members of the AGC family of protein kinases were defined as the outgroup. The SNF1 kinases are those from *A. thaliana* and *S. cervisiae* . The animal CaMK sequences are those from rat phosphorylase b kinase and mouse Ca2+/Cam-dependent protein kinase II. The thirty PPCK sequences are from *A. thaliana, B. olerace, B. vulgaris, E. esula, F. trinervia, G. arboreum, G. max, H. vulgare, K. fedtschenkoi, L. esculentum, L. japonicus, M. crystallinum, M. truncatula, O. sativa, P. vulgaris, S. bicolor, S. tuberosum, T. aestivum* and *Z. mays*. The plant calcium-dependent kinase sequences are from *A. thaliana, M. crystallinum, N. tabacum, T. aestivum* and *Z. mays*.

The asterix marked [*] PPCK sequences represent partial aminoacid sequences.

The bar and number represents the difference in sequence identity.

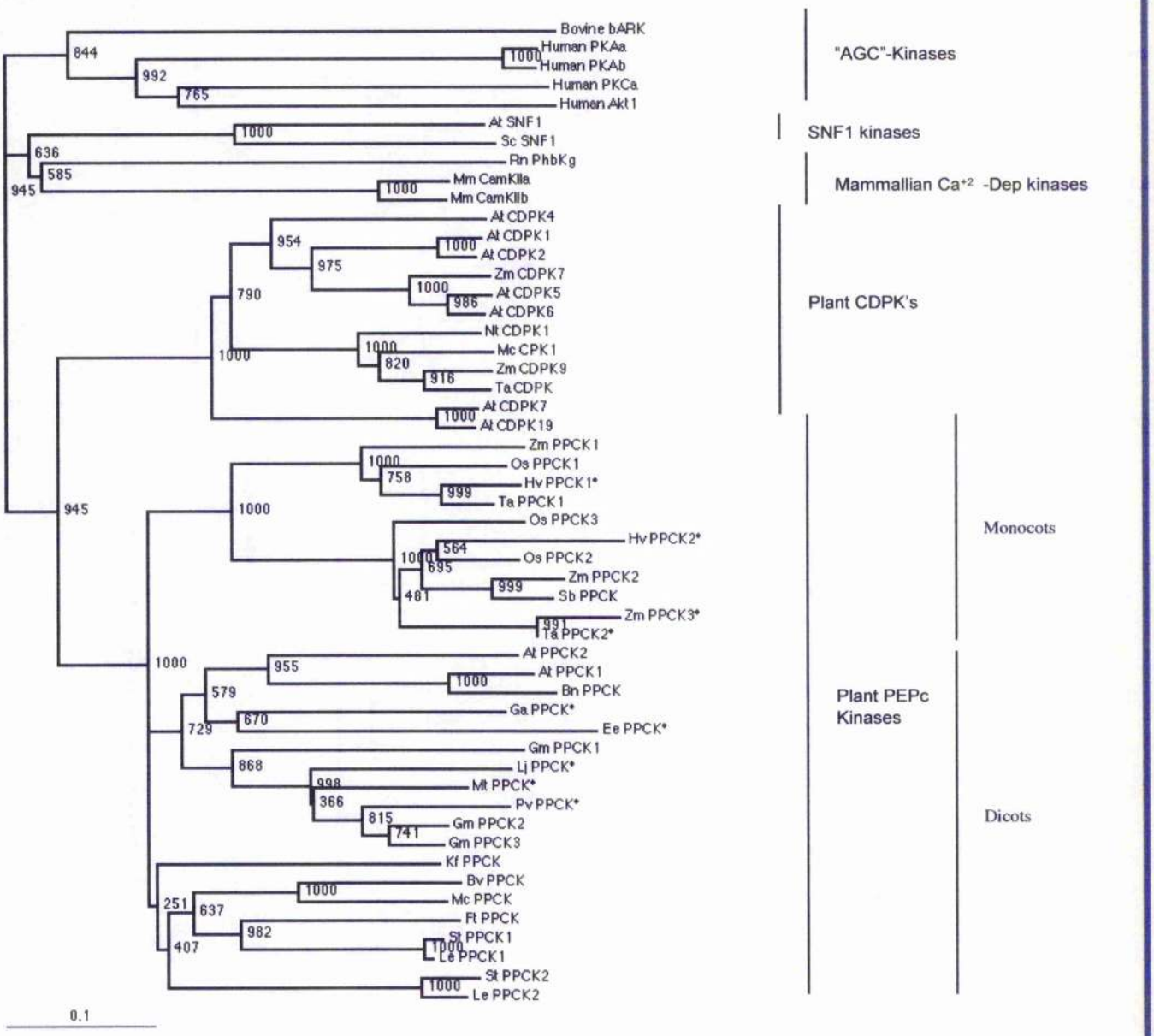
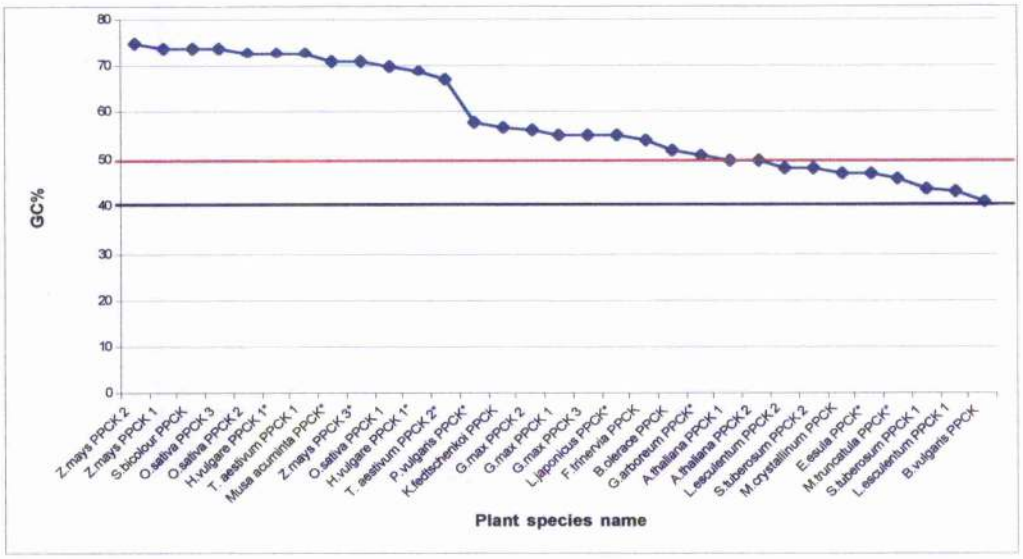
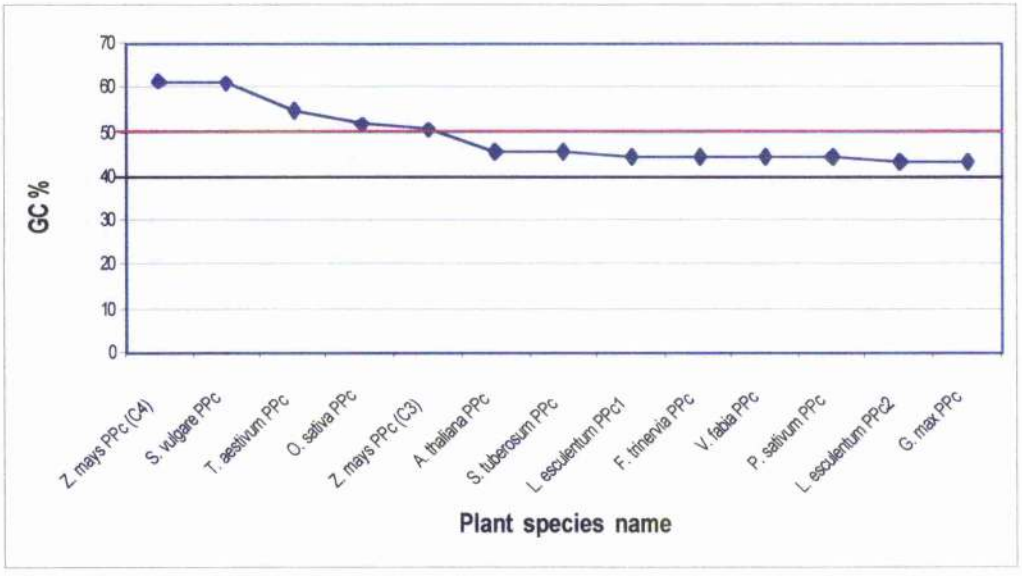

Bovine bARK
Human PKAa
Human PKAb
Human PKCa
Human Akt1
At SNF1
Sc SNF1
Rn PhbKg
Mm CamKIIa
Mm CamKIIb
At CDPK4
At CDPK1
At CDPK2
Zm CDPK7
At CDPK5
At CDPK6
Nt CDPK1
Mc CPK1
Zm CDPK9
Ta CDPK
At CDPK7
At CDPK19
Zm PPCK1
Os PPCK1
Hv PPCK1*
Ta PPCK1
Os PPCK3
Hv PPCK2*
Os PPCK2
Zm PPCK2
Sb PPCK
Zm PPCK3*
Ta PPCK2*
At PPCK2
At PPCK1
Bn PPCK
Ga PPCK*
Ee PPCK*
Gm PPCK1
Lj PPCK*
Mt PPCK*
Pv PPCK*
Gm PPCK2
Gm PPCK3
Kf PPCK
Bv PPCK
Mc PPCK
Ft PPCK
St PPCK1
Le PPCK1
St PPCK2
Le PPCK2
0.1
"AGC"-Kinases
SNF1 kinases
Mammallian Ca+2 -Dep kinases
Plant CDPK's
Monocots
Plant PEPc Kinases
Dicots

Figure 5.3 Graphical representation of GC content of PPCK and PEPc open reading frames.


GC%
80
70
60
50
40
30
20
10
0
Z.mays PPCK 2
Z.mays PPCK 1
S.bicolour PPCK
O.sativa PPCK 3
O.sativa PPCK 2
H.vulgare PPCK 1*
T. aestivum PPCK 1
Musa acuminita PPCK*
Z.mays PPCK 3*
O.sativa PPCK 1
H.vulgare PPCK 1*
T. aestivum PPCK 2*
P.vulgaris PPCK*
K.fedtschenkoi PPCK
G.max PPCK 2
G.max PPCK 1
G.max PPCK 3
L.japonicus PPCK*
F.trinervia PPCK
B.oleracea PPCK
G.arboreum PPCK*
A.thaliana PPCK 1
A.thaliana PPCK 2
L.esculentum PPCK 2
S.tuberosum PPCK 2
M.crystallinum PPCK
E.esula PPCK*
M.truncatula PPCK*
S.tuberosum PPCK 1
L.esculentum PPCK 1
B.vulgaris PPCK
Plant species name
Monocot PPCK
Dicot PPCK


GC %
70
60
50
40
30
20
10
0
Z. mays PPc (C4)
S. vulgare PPc
T. aestivum PPc
O. sativa PPc
Z. mays PPc (C3)
A. thaliana PPc
S. tuberosum PPc
L. esculentum PPc1
F. trinervia PPc
V. fabia PPc
P. sativum PPc
L. esculentum PPc2
G. max PPc
Plant species name
Monocot PEPc
Dicot PEPc

Chapter 6

General discussion and future research

In recent years research into the control of plant PEPc has focussed upon the identification and regulation of plant PPCKs. A number of important advances have been made, with the past few years seeing the identification of a number of PPCK genes. This has allowed the study of PPCK transcript levels, which in turn has lead to the hypothesis that PPCK activity is generally controlled at the level of gene expression rather than by second messengers such as Ca^{2+}, cAMP or diacylglycerol. The next few years will see more effort directed at heterologous expression of the newly identified PPCKs, allowing anti-PPCK antibodies to be raised and crystallisation studies to be carried out. These core experiments have not been possible until now due to the low level of PPCK *in vivo*, stunting classical biochemical approaches.

The aims of the work described in this thesis have been twofold. My first objective was to identify the PPCK responsible for light-activation of PEPc in maize. During the course of my work, a chance observation about tomato PPCK ests grew into an analysis of alternative splicing in an unusual PPCK gene.

My direct attempts to identify PPCK genes from maize did not succeed, though subsequent work described in Chapter 3 has provided an explanation for this. During my work a light-up regulated PPCK gene was identified in the C_4 dicot plant, *Flaveria trinervia* (Tsuchida *et al.*, 2001). Our lab had previously identified a PPCK gene from the C_4 monocot, *Sorghum*. This kinase, though, showed no up regulation in illuminated mature leaves and its role is presumably not to phosphorylate leaf PEPc in the light. The work described in Chapter 3 highlights that a very similar PPCK gene is also present in maize; the expression of the maize *PPCK2* gene, like the *Sorghum* gene, is not up regulated by light. This part of the results also shows the existence of two other putative maize PPCK genes. One of these, *ZmPPCK1*, has been fully sequenced by our lab. The expression of this gene is up regulated by light in leaves. Analysis of the recently sequenced rice genome also revealed three putative PPCKs similar to those found in maize. This demonstrates that a small PPCK gene family exists within monocot plants. Interestingly, the highly acidic extension present in the centre of SbPPCK and ZmPPCK2 is absent from the rice PPCKs. This unique PPCK region would be worth investigating, particularly in terms of the solubility of PPCK

and its recognition of PEP. Protein modelling of PPCK by our lab suggests that the acidic region in SbPPCK and ZmPPCK2 is located at the surface of the protein on the opposite side to the catalytic site (H.G. Nimmo, unpublished). This raises the idea that this region may interact with a protein without affecting the catalytic ability of PPCK. PPCK is thought to be highly specific to PEPc, but unusually, for a protein kinase, it does not phosphorylate short peptides well. It is possible, though unlikely that the addition of an acidic domain to PPCK may cause the enzyme to recognise substrates other than PEPc. Alternatively, this extra region may be involved in regulation of PPCK. PPCK is known to have a short half-life *in vivo*. This would suggest some form of targeted degradation, perhaps by the ubiquination pathway. Such a process may involve direct recognition of PPCK. It would therefore be interesting to analyse whether SbPPCK or ZmPPCK2 has a different half-life than other PPCKs lacking the acidic insertion. It is also worth noting that SbPPCK is dramatically induced by cycloheximide (Fontaine *et al.*, unpublished). One plausible explanation for this is the existence of a protein factor required for turnover of this PPCK that itself has a short half-life.

This work also highlights the curious GC-richness of the monocot PPCKs (Chapter 5). The importance of this trait requires further research, possibly directed at the importance of mRNA secondary structure formation *in planta* and *in vitro*. One such experiment is outlined in Chapter 5.

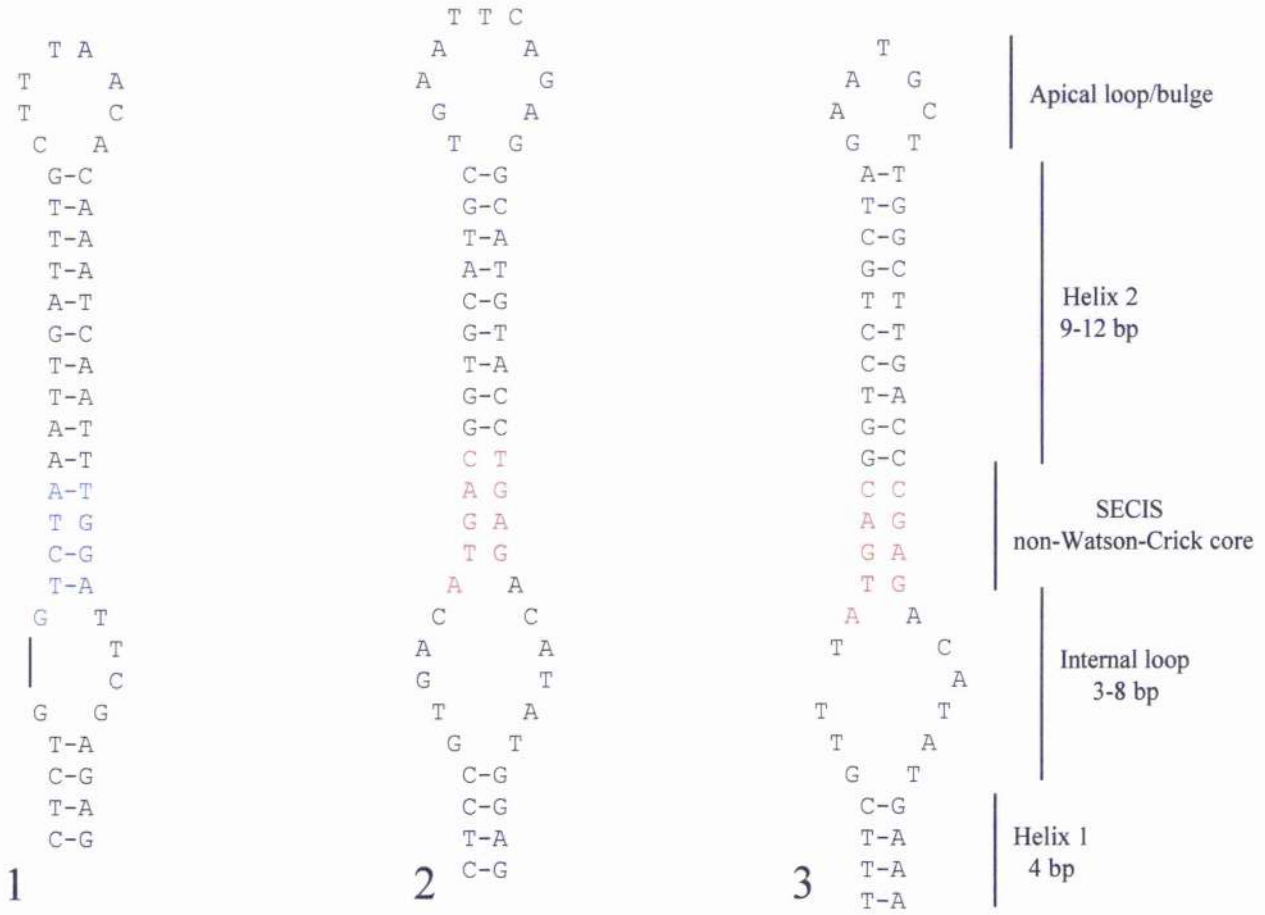
The second aspect of this work has been to investigate C_3 dicot plant PPCK, using tomato. Tomato is the model system used in fruit ripening, and thus allowed the investigation of PPCK in fruit, an area much ignored in plant PPCK research. Chapter 4 describes how I identified two PPCKs from tomato. Building upon this, expression studies were carried out on a range of tissues and fruit compartments revealing that these PPCKs are controlled differentially and are not constitutive. It remains to be seen where both isoforms of PPCK are expressed in a single cell type and, if so, whether both can phosphorylate the various isoforms of PEPc equally well. The control of PPCK gene expression and expression is likely to be multi-faceted, influenced by a range of factors.

I have already discussed the possible functions of the two PPCK genes and of the alternative splicing of *LePPCK2* in Chapter 4. One further, very speculative, idea about the alternative splicing of LePPCK2 can be explored. Analysis of the LePPCK2

transcript-β reveals that the only premature stop codon is UGA. LePPCK2 transcript-γ has further stop codons downstream. In animals the read through of the stop codon UGA is possible if the translational machinery interprets the codon as a selenocysteine (Sec). This is only possible if the 3' untranslated region of the mRNA holds a Sec insertion element (SECIS). The SECIS is a cis-acting mRNA structure that recruits the trans-acting Sec tRNA, selenophosphate synthetase, Sec synthetase, Sec-specific elongation factor and a SECIS-binding protein (Novoselov *et al.*, 2002). So far, higher plants and yeast have been assumed not to produce selenoproteins, as Sec insertion machinery has not been identified in the genome of *A. thaliana*. Recently in *Chlamydomonas reinhardtii*, a member of the plant kingdom, selenoproteins have been identified and their SECIS sequences characterised (Novoselov *et al.* 2002). SECIS sequences are essentially a helix separated by an internal loop that holds a SECIS conserved core that is a quartet of non Watson-Crick paired nucleotides and an apical loop. A comparison of the 3'-untranslated region of LePPCK2 and the SECIS of *C. reinhardtii* selenoprotein K 1 (CrSelK1) and *C. reinhardtii* thioredoxin reductase 1 (CrTR1) showed that a putative SECIS is present in the 3' untranslated region of LePPCK2. Figure 6.1 shows a structural comparison of these three sequences. Interestingly, the secondary structure is also present in potato PPCK 2. This may mean that the LePPCK2 transcript-β forms a 37 kDa seleno-PPCK. However, the absence of a conserved non Watson-Crick core suggests that the putative SECIS sequence is either non-functional or a new form of a SECIS. Even if non-functional, the presence of a near complete SECIS suggests that this gene may have had a selenoprotein role in evolutionary history. Alternatively, this region of secondary structure has another role. The identification of potato and aubergine homologues of tomato PPCK1 and 2, as described in Chapter 4, strongly suggests that other closely related plants such as tobacco may have a similar alternatively spliced *PPCK2*-like gene. I conducted a number of experiments to investigate the possibility that LePPCK-transcript-β encoding a selenoprotein. These experiments (not shown) were inconclusive. The experimental strategy involved the transcription and translation of both linearized LePPCK-transcript-β cDNA and linearized mammalian seleno-binding protein 2 (SBP2) cDNA. However, no 37 kDa translation product, the

Figure 6.1 Secondary structure of 3' untranslated regions of *LePPCK2* and the selenocysteine proteins CrTR1 and CrSelK1

The red indicates the SECIS conserved non-Watson-Crick core, while the blue indicates a similar LePPCK2 core.


Apical loop/bulge
Helix 2
9-12 bp
SECIS
non-Watson-Crick core
Internal loop
3-8 bp
Helix 1
4 bp
1
2
3
Helix 1
Quartet
Apical loop/bulge
Quartet
Helix 1
1) Le PPCK 2 CTCT G G TCTA AATTGATTTG CTTTAACA CAAATCAATT TGGA TTCG AGAG
2) Cr TR 1 CTCC GTGAC A TGAC GGTGCATGC TGAATTCAGAG GCATGTACC TGAG ACATAT GCAG
3) Cr SelK 1 TTTC GTTT A TGAC GGTCCTGGCTA GAATGCT TGGCTTGGACC CGAG CTACA GAAA
Helix 2
Helix 2

expected size of a putative LePPCK-transcript-β selenoprotein, was detected. Similarly, the translation products were not able to phosphorylate PEPc.

The research described in this thesis raises a number of interesting questions that require further research. Firstly, it is important to confirm the function of the putative maize and rice PPCK genes by *in vitro* translation of full-length cDNAs, followed by PPCK activity assays. Building on this, the expression pattern of all three maize and rice PPCKs must be analysed. This would involve first semi-quantitative RT-PCR and subsequently generation and analysis of genetically modified plants with PPCK gene knockouts, antisense PPCK or expressing GFP fusion constructs or PPCK promoter:reporter constructs. This would allow us to determine the roles of each PPCK. Rice is the obvious organism for such studies. However one key question must be resolved in maize (or other C_4 plants) namely the nature and mode of regulation of the PPCKs involved in control of the C_4 pathway. This needs to be established for both C_4 monocots such as maize and C_4 dicots such as *Flaveria trinervia*.

The existence of a PPCK multi-gene family raises the question of whether a single isoform can perform multiple functions in specialised tissues or if several isoforms are expressed together. This question would be relevant for example, in the study of PPCK expression in illuminated maize mesophyll cells. This cell type would be expected to carry out both nitrogen assimilation **and the** fixation of CO_2 in the C_4 pathway. Therefore there are at least two roles for PPCK in this tissue. Expression analysis may indicate whether these roles are carried out by one, two or three PPCKs. Accumulation of one or more forms may suggest that the PPCKs are up regulated by similar signalling pathways.

Another important question that has yet to be answered is the identity of the 37 kDa form of PPCK that has been reported in some species by in-gel assays (as discussed in Chapter 1). Anti-PPCK antibodies, raised from an over-expressed PPCK would help to answer this question. A band corresponding to a 37 kDa form on a western blot using this antibody would imply that this is post-translationally modified PPCK, possibly ubiquinated. This experiment could best be conducted in *Arabidopsis* as the complete sequencing of its genome suggests that no gene encoding a 37 kDa form of PPCK exists. Anti-PPCK antibodies and heterologous expression of the three rice and maize PPCKs would also aid research into any specificity and differences in the catalytic ability of these kinases to phosphorylate the multiple PEPc isoforms.

The studies upon the tomato PPCKs also raise several possibilities for further research. The expression of both PPCK1 and PPCK2 should be studied at much higher resolution. The possible expression of PPCK and PEPc in vascular bundles has already been mentioned. This could be assessed for PPCK1 by using *in situ* RT-PCR or hybridisation on tomato fruit sections. Investigating LePPCK2 expression at a cell-specific level would be more difficult, as *in situ* RT-PCR and hybridisation would most likely be unable to differentiate between LePPCK2 transcripts-α, -β and -γ. A more productive, but difficult, strategy would be to conduct RT-PCR upon RNA purified from single cells of the tomato fruit, by use of single cell sampling (for an example see Laval *et al.*, 2002). An obvious problem for this approach is the reproducibility between experiments.

The question of the possible role of the translated product of LePPCK2 transcripts-β and -γ may be answered by the generation of an anti-LePPCK2 antibody against a peptide corresponding to the region not present in the translated product of LePPCK2 transcript-α. Such an anti-peptide antibody would allow a western blot analysis of tomato tissue to reveal whether a truncated protein accumulates (16 kDa) and/or a 37 kDa form is present. The accumulation of the truncated protein would be indirect evidence that this protein has some unidentified role in tomato, while the presence of a 37 kDa form might suggest that the product of LePPCK2 transcripts-β is a selenoprotein. It thus becomes crucial that an anti-PPCK antibody is acquired for the next phase of studies on PPCK. This will require over-expression of PPCK, which should be achieved by 2003, opening up a number of new areas of study.

References

Andreo, C.S., Gonzalez, D.H. and Iglesias, A.A. (1987) Higher plant phosphoenolpyruvate carboxylase: structure and regulation. *FEBS Letts.* **213**, 1-8.

Araus, J.L, Bort, J., Brown, R.H., Bassett, C.L. and Cortadellas, N. (1993) Immunocytochemical localization of PEP carboxylase and photosynthesis gas-exchange charecteristics in ears of *Triticum durum*. *Planta* **191**, 507-514.

Ashton, A.R. and Hatch, M.D. (1983) Regulation of C_4 photosynthesis: regulation of pyruvate, Pi dikinase by ADP-dependent phosphorylation and dephosphorylation. *Biochem Biophys Res Commun.* **115**, 53-60.

Ausenhus, S.L. and O'Leary, M.H. (1992) Hydrolysis of phosphoenolpyruvate by phosphoenolpyruvate carboxylase from *Zea mays*. *Biochem.* **31**, 6427-6431.

Bakrim, N., Brulfert, J., Vidal, J. and Chollet, R. (2001) Phosphoenolpyruvate carboxylase kinase is controlled by a similar signalling cascade in CAM and C_4 plants. *Biochem Biophys Res Commun.* **286**, 1158-1162.

Bakrim, N., Prioul J.L., Deleens E., Rocher, J.P, Arrio-Dupont, M., Pierre, J.N., Vidal, J., Chollet, R. and Gadal, P. (1992) Regulatory phosphorylation of *Sorghum* leaf phosphoenolpyruvate carboxylase: identification of the protein-serine kinase and some elements of the signal-transduction cascade. *Eur. J. Biochem.* **204**, 821-830.

Bakrim, N., Prioul, J.L., Deleens, E., Rocher, J.P., Arrio-Dupont, M., Vidal, J., Gadal, P. and Chollet R. (1993) Regulatory phosphorylation of C_4 phosphoenolpyruvate carboxylase. A cardinal event influencing the photosynthesis rate in *Sorghum* and maize. *Plant Physiol.* **101**, 891-897.

Baur B., Dietz, K.J. and Winter, K. (1992) Regulatory protein phosphorylation of phosphoenolpyruvate carboxylase in the facultative crassulacean-acid-metabolism plant *Mesembryanthemum crystallinum* L. *Eur. J. Biochem.* **209**, 95-101.

Blanke, M.M. and Lenz, F. (1989) Fruit photosynthesis. *Plant Cell Environ.* **12**, 31-46.

Borland A.M. and Griffiths H. (1997) A comparative study on the regulation of C_3 and C_4 carboxylation processes in the constitutive crassulacean acid metabolism (CAM) plant *Kalanchoë daigremontiana* and the C_3-CAM intermediate *Clusia minor*. *Planta* **201**, 368-378.

Borland, A.M., Hartwell, J., Jenkins, G.I., Wilkins, M.B. and Nimmo, H.G. (1999) Metabolite control overrides circadian regulation of phosphoenolpyruvate carboxylase

kinase and CO_2 fixation in Crassulacean acid metabolism. *Plant Physiol.* **121**, 889-896.

Bowes G, Rao SK, Estavillo GM and Reiskind JB. 2002. C-4 mechanisms in aquatic angiosperms: comparisons with terrestrial C_4 systems. *Funct. Plant Biol.* **29**, 379-392.

Brown, J.W.S. and Simpson C.G. (1998) Splice site selection in plant pre mRNA splicing. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* **49**, 77-95.

Brulfert, J., Vidal, J. Keryer, E. Thomas, M., Gadal, P. and Queiroz, O. (1986) Phosphorylation-dephosphorylation process as a probable mechanism for the diurnal regulatory changes of phosphoenolpyruvate carboxylase in CAM plants. *Biochem. Biophys.Res. Commun.* **136**, 151-159

Buchanan-Bollig, I.C. and Smith J.A.C. (1984) Kinetic changes with temperature of phosphoenolpyruvate carboxylase from a CAM plant. *Plant Cell Environ.* **7**, 63-70.

Budde, R.J.A. and Chollet, R. (1986) *In vitro* phosphorylation of maize leaf phosphoenolpyruvate carboxylase. *Plant Physiol.* **82**, 1107-1114.

Burnell, J.N. (1984) Regulation of C_4 photosynthesis catalytic dephosphorylation and Pi-mediated activation of pyruvate Pi dikinase. *Biochem Biophys Res Commun.* **120**, 559-565.

Burnell, J.N. and Ludwig, M. (1997) Characterisation of two cDNAs encoding carbonic anhydrase in maize leaves. *Aust. J. Plant Physiol.* **24**, 451-458.

Canvin, D.T. (1990) In: *Plant Physiology, Biochemistry and Molecular Biology.* (D. T. Dennis and D. H. Turpin, eds.), pp253-273, Longman Scientific and Technical, Singapore.

Carroll, L.J., Dunaway-Mariano, D., Smith, C.M. and Chollet, R. (1990) Determination of the catalytic pathway of C_4-leaf pyruvate, orthophosphate dikinase from maize. *FEBS Lett.* **274**, 178-180.

Carter, P.J., Nimmo, H.G., Fewson, C.A. and Wilkins, M.B. (1990) *Bryophyllum fedtschenkoi* protein phosphatase type 2A can dephosphorylate phosphoenolpyruvate carboxylase. *FEBS lett.* **263**, 233-236.

Carter P.J., Nimmo, H.G., Fewson, C.A., and Wilkins, M.B. (1991) Circadian rhythms in the activity of a plant protein kinase. *EMBO J.* **10**, 2063-2068.

Carter P.J., Wilkins, M.B., Nimmo, H.G. and Fewson, C.A. (1995a) Effects of temperature on the activity of phosphoenolpyruvate carboxylase and on the control of CO_2 fixation in *Bryophyllum fedtschenkoi*. *Planta* **196**, 375-380.

Carter P.J., Wilkins, M.B., Nimmo, H.G. and Fewson, C.A. (1995b) The role of temperature in the regulation of the circadian rhythm of CO_2 fixation in *Bryophyllum fedtschenkoi*. *Planta* **196**, 381-386

Carter P.J., Wilkins, M.B., Nimmo, H.G. and Fewson, C.A. (1996) Role of circadian rhythms, light and temperature in the regulation of phosphoenolpyruvate carboxylase in crassulacean acid metabolism. In: *Crassulacean Acid Metabolism*. (K. Winter and J.A.C. Smith eds.), pp46-52. Springer-Verlag

Chollet, R., Vidal, J. and O'Leary, M.H. (1996) Phosphoenolpyruvate carboxylase: a ubiquitous, highly regulated enzyme in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **47**, 273-298.

Christeller, J.T., Laing, W.A. and Sutton, D. (1977) Carbon dioxide fixation by lupin nodules. **1**. Characterisation, association with phosphoenolpyruvate carboxylase, and correlation with nitrogen fixation with nodule development. *Plant Physiol.* **60**, 47-50.

Coker, G. T. and Schubert K. R. (1981) Carbon dioxide fixation in soybean roots and nodules.1. Characterization and comparison with N_2 fixation and composition of xylem exudate during early nodule development. *Plant Physiol.* **67**, 691-696.

Colombo, S.L., Pairoba, C.F. and Andreo, C.S. (1996) Inhibitory effect of shikimic acid on phosophoenolpyruvate carboxylase activity. *Plant Cell Physiol.* **37**, 870-872.

Cottelle, V., Pierre, J-N. and Vavasseur, A. (1999) Potential strong regulation of guard cell phosophoenolpyruvate carboxylase through phosphorylation. *J. Exp. Bot.* **50**, 777-783.

Coursol, S., Giglioli-Guivarc'h, N., Vidal, J. and Pierre, J.N. (2000) An increase in phosphoinositide-specific phospholipase C activity precedes induction of C_4 phosphoenolpyruvate carboxylase phosphorylation in illuminated and NH_4Cl-treated protoplasts from *Digitaria sanguinalis*. *Plant J.* **23**, 497-506.

Coursol, S., Pierre, J.N. and Vidal, J. (2000) Role of the phosphoinositide pathway in the light-dependent C_4 phosphoenolpyruvate carboxylase phosphorylation cascade in *Digitaria sanguinalis* protoplasts. *Biochem Soc. Trans.* **28**, 821-823.

Cretin, C., Santi, S, Keryer, E. Lepiniec, L., Tagu, D., Vidal, J. and Gadal, P. (1991) The phosphoenolpyruvate carboxylase gene family of *Sorghum*: promoter structures, amino acid sequences and expression of genes. *Gene* **99**, 87-94.

Cushman, J.C. (2001) Crassulacean acid metabolism. A plastic photosynthetic adaptation to arid environments. *Plant Physiol.* **127**, 1439-1448.

Cushman, J.C. and Bohnhert H.J. (1997) Molecular genetics of Crassulacean acid metabolisim. *Plant Physiol.* **113**, 667-676.

Cushman, J.C. and Bohnert, H.J. (1999) Crassulacean acid metabolism: Molecular genetics. *Ann. Rev. Plant Phys. Plant Mol. Biol.* **50**, 305-332.

Deroche, M.E. and Carrayol, E. (1988) Nodule phosphoenolpyruvate carboxylase: a review. *Physiol Plant.* **74**, 775-782.

Dittrich, P., Campbell, W.H. and Black Jr., C.C. (1973) Phosphoenolpyruvate carboxykinase in plants exhibiting crassulacean acid metabolism. *Plant Physiol.* **52**, 357-361.

Dodd AN, Borland AM, Haslam RP, Griffiths H, Maxwell K. (2002) Crassulacean acid metabolism: plastic, fantastic. *J. Exp. Bot.* **53**. 569-580.

Doncaster, H.D. and Leegood, R.C. (1987) Regulation of phosphoenolpyruvate carboxylase activity in maize leaves. *Plant Physiol.* **84**, 82-87.

Dong, L-Y., Ermolova, N.V. and Chollet, R. (2001) Partial purification and biochemical characterization of a heteromeric protein phosphatase 2A holoenzyme from maize (*Zea mays* L.) leaves that dephosphorylates C_4 phosphoenolpyruvate carboxylase. *Planta* **213**, 379-389.

Dong, L.-Y., Masuda, T., Kawamura, T., Hata, S. and Izui, K. (1998) Cloning, expression and characterzation of a root-form phosphoenolpyruvate carboxylase from *Zea mays*: comparison with C_4-form enzyme. *Plant Cell Physiol.* **39**, 865-873.

Drozdowicz, Y.M. and Jones, R.L. (1995) Hormonal regulation of organic and phosphoric acid release by barley aleurone layers and scutella. *Plant Physiol.* **108**, 769-776.

Du, Z., Aghoram, K. and Outlaw, J., W.H. (1997) *In vivo* phosphorylation of phosphoenolpyruvate carboxylase in guard cells of *Vicia faba* L. is enhanced by fusicoccin and suppressed by abscisic acid. *Arch. Biochem. Biophys.*, **337**, 345-350.

Duff, M.G.S. and Chollet, R. (1995) *In vivo* regulation of wheat leaf phosphoenolpyruvate carboxylase by reversible phosphorylation. *Plant Physiol.* **107**, 775-782.

Duff, M.G.S., Giglioli-Guivarc'h, N., Pierre, J-N, Vidal, J., Conden, S.A. and Chollet, R. (1996) *In situ* evidence for the involvement of calcium and bundle-sheath-derived photosynthetic metabolites in the C_4 phosphoenolpyruvate carboxylase kinase signal transduction chain. *Planta* **199**, 467-474.

Echevarria, C., Garcia-Maurino, S., Alvarez, R., Soler, A. and Vidal, J. (2001) Salt stress increases the Ca^{2+}-independent phosphoenolpyruvate carboxylase kinase activity in *Sorghum* leaves. *Planta* **214**, 283-287.

Echevarria, C., Vidal, J., Jiao, J. and Chollet, R. (1988) The phosphorylation of *Sorghum* leaf phosphoenolpyruvate carboxylase is a Ca^{2-}calmodulin dependent process. *Biochem. Biophys. Res. Commun.* **155**, 835-840.

Edwards, G.E., Nakamoto, H., Burnell, J.N. and Hatch, M.D. (1985) Pyruvate, Pi dikinase and NADP-malate dehydrogenase in C_4 photosynthesis - properties and mechanism of light dark regulation. *Ann Rev Plant Physiol Plant Mol Biol.* **36**, 255-286.

Ernst, K. and Westhoff, P. (1997) The phosphoenolpyruvate carboxylase (ppc) gene family of *Flaveria trinervia* (C_4) and *F. pringlei* (C_3): molecular charecterisation and expression analysis of the ppcB and ppcC genes. *Plant Mol. Biol.* **34**, 427-443.

Etienne, C., Moing, A., Dirlewanger, E., Raymond, P.,Monet, R. and Rothan, C. (2002) Isolation and characterization of six peach cDNAs encoding key proteins in organic acid metabolism and solute accumulation: involvement in regulating peach fruit acidity. *Physiol. Plantarum* **114**, 259-270.

Famiani, F., Walker, R.P, Tesci, L., Chen, Z.H., Proietti, P. and Leegood, R.C. (2000) An immunohistochemical study of the compartmentation of metabolism during the development of grape (*Vitis vinifera* L.) berries. *J. Exp. Bot.* **51**, 675-683.

Fontaine, V., Hartwell, J., Jenkins, G.I. and Nimmo, H.G. (2002) *Arabidopsis thaliana* contains two phosphoenolpyruvate carboxylase kinase genes with different expression patterns. *Plant Cell Environ.* **25**, 115-122.

Gao, Y. and Woo, K.C. (1996) Regulation of phosphoenolpyruvate carboxylase in *Zea mays* by protein phosphorylation and metabolites and their roles in photosynthesis. *Ann. J. Plant Physiol.* **23**, 25-32.

Gao, Y. and Woo, K.C. (1996) Site-directed mutagenesis of *Flaveria trinervia* phosphoenolpyruvate carboxylase: Arg(450) and Arg(767) are essential for catalytic activity and Lys(829) affects substrate binding. *FEBS lett.* **392**, 285-288.

Giglioli-Guivarc'h, N., Pierre, J-N., Brown, S., Chollet, R., Vidal, J. and Gadal, P. (1996) The light-dependent transduction pathway controlling the regulatory phosphorylation of C_4 phosphoenolpyruvate carboxylase in protoplasts in *Digitaria sanguinalis*. *Plant Cell* **8**, 573-586.

Glackin, C. and Grula, J. (1990) Organ-specific transcripts of different size and abundance derive from the same pyruvate orthophosphate dikinase gene in maize. *Proc. Natl. Acad. Sci. USA* **87**, 3004-3008.

Gniadkowski, M., Hemmings-Mieszczak, M., Klahre, U., Liu, H.X. and Filipowicz, W. (1996) Charecterization of intronic uridine-rich sequence elements acting as possible targets for nuclear proteins during pre-mRNA splicing in *Nicotiana plumbaginifolia*. *Nucleic Acids Res.* **24**, 619-627.

Golombek, S., Heim, U., Horstmann, C., Wobus, U. and Weber, H. (1999) Phosphoenolpyruvate carboxylase in developing seeds of *Vicia faba* L: gene expression and metabolic regulation. *Planta* **208**, 66-72.

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

Guillet, C., Just, D., Benard, N., Destrac-Irvine, A., Baldet, P., Hernould, M., Causse, M., Raymond, P. and Rothan, C. (2001) A fruit-specific phosphoenolpyruvate carboxylase is related to rapid growth of tomato fruit. *Planta* **214**, 717-26.

Hanks S.K. and Hunter T. (1995) The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* **9**, 576-596.

Hartwell, J., Jenkins, G.I., Wilkins, M.B. and Nimmo, H.G. (1999a) The light induction of maize phosphoenolpyruvate carboxylase kinase translatable mRNA requires transcription but not translation. *Plant Cell Environ.* **22**, 883-889.

Hartwell, J., Gill, A., Nimmo, G.A., Wilkins, M.B., Jenkins, G.I. and Nimmo, H.G. (1999b) Phosphoenolpyruvate carboxylase kinase is a novel protein kinase regulated at the level of expression. *Plant J.* **20**, 333-342.

Hartwell, J., Smith, L.H., Wilkins, M.B., Jenkins, G.I. and Nimmo, H.G. (1996) Higher plant phosphoenolpyruvate carboxylase kinase is regulated at the level of translatable mRNA in response to light or circadian rhythm. *Plant J.* **10**, 1071-1078.

Hartwell, J., Nimmo, G.A., Wilkins, M.B., Jenkins, G.I. and Nimmo, H.G. (2002) Probing the circadian control of phosphoenolpyruvate carboxylase kinase expression in *Kalanchoe fedtschenkoi*. *Funct. Plant Biol.* **29**, 663-668.

Hata, S., Izui, K. and Kouchi, H. (1998) Expression of a soybean nodule-enhanced phosphoenolpyruvate carboxylase gene that shows striking similarity to another gene for a house-keeping isoform. *Plant J.* **13**, 267-273.

Hatch M.D. and Slack. C.R. (1970) Photosynthetic CO_2 - fixation pathways. *Ann.Rev. Plant. Phys.* **21**, 141-147.

Herppich, W.B. and Herppich, M. (1997) The interrelationship between changes in PEPc activity and organic acid accumulation during the C_3-CAM shift in *Mesembryanthemum crystallinum*. *Plant Physiol.* **151**, 373-378.

Huber, S.C., Huber, J.L. and McMichael, R.W. (1994) Control of plant enzyme activity by reversible protein phosphorylation. *Int. Rev. Cytology.* **149**, 47-98.

Hudsperth, R.L. and Grula, J.W. (1989) Structure and expression of the maize gene encoding phosphoenolpyruvate carboxylase isozyme involved in C_4 photosynthesis. *Plant Mol. Biol.* **12**, 579-589.

Imaizumi, N., Ku, M.S., Ishihara, K., Samejima, M., Kaneko, S. and Matsuoka, M. (1997) Characterisation of the gene for pyruvate, orthophosphate dikinase from rice, a C_3 plant and a comparison of structure and expression between C_3 and C_4 genes for this protein. *Plant Mol. Biol.* **34**, 701-716.

Israel, D. W. and Jackson, W.A. (1982) Ion balance, uptake, and transport processes in N_2-fixing and nitrate- and urea-dependent soybean plants. *Plant Physiol.* **69**, 171-178.

Jeanneau, M., Vidal, J., Gousset-Dupont, A., Lebouteiller, M., Gerentes, D. and Perez, P. (2002) Manipulating PEPc levels in plants. *J. Exp. Bot.* **53**, 1837-1845.

Jiao, J. and Chollet, R. (1988) Light/dark regulation of maize leaf phosphoenolpyruvate carboxylase by in vivo phosphorylation. *Arch. Biochem. Biophys.* **261**, 409-417.

Jiao, J. and Chollet, R. (1989) Regulatory seryl-phosphorylation of C_4 phosphoenolpyruvate carboxylase by a soluble protein kinase from maize leaves. *Arch. Biochem. Biophys.* **269**, 529-535.

Jiao, J. and Chollet, R. (1990) Regulatory phosphorylation of serine-15 in maize phosphoenolpyruvate carboxylase by a C_4-leaf protein-serine kinase. *Arch. Biochem. Biophys.* **283**, 300-305.

Jiao, J. and Chollet, R. (1991) Posttranslational modification of phosphoenolpyruvate carboxylase in C_4 and crassulacean acid metabolism plants. *Plant Physiol.* **95**, 981-985.

Jiao J., Echevarria, C., Vidal, J. and Chollet, R. (1991a) Protein turnover as a component in the light/dark regulation of phosphoenolpyruvate carboxylase protein-serine kinase activity in C_4 plants. *Proc. Natl. Acad. Sci USA* **88**, 2712-2715.

Jiao J., Vidal, J., Echevarria, C. and Chollet, R. (1991b) *In vivo* regulatory phosphorylation site in C_4-leaf phosphoenolpyruvate carboxylase from maize and *Sorghum*. *Plant Physiol.* **96**, 297-301.

Jenkins, C.L.D., Anderson, L.E. and Hatch, M.D. (1986) NADP-malate dehydrogenase from *Zea Mays* leaves - amino-acid-composition and thiol content of active and inactive forms. *Plant Sci.* **45**, 1-7.

Kagawa, T. and Bruno, P.L. (1988) NADP-malate dehydrogenase from leaves of *Zea mays*: purification and physical, chemical, and kinetic properties. *Arch Biochem Biophys*. **260**, 674-695.

Kai, Y., Matsumura, H., Inoue, T., Terada, K., Nagara, Y., Yoshinaga, T., Kihara, A., Tsumura, K. and Izui, K. (1999) Three-dimensional structure of phosphoenolpyruvate carboxylase: A proposed mechanism for allosteric inhibition. *Proc. Natl. Acad. Sci. USA* **96,** 823-828.

Kawamura, T., Shigesada, K., Toh, H., Okumura, S., Yanagisawa, S. and Izui, K. (1992) Molecular evolution of phosphoenolpyruvate carboxylase for C_4 photosynthesis in maize: comparison of its cDNA sequence with a newly isolated cDNA encoding an isozyme involved in the anaplerotic function. *J Biochem*. **112,** 147-54.

Kloos, D. U., Oltmanns, H., Dock, S., Stahl, D. and Hehl, R. (2002) Isolation and molecular analysis of six taproot expressed genes from sugar beet. *J. Exp. Bot.* **53**, 1533-1534.

Kluge, M., Maier, P., Brulfert, J., Faist, K. and Wollny, E. (1988) Regulation of phosphoenolpyruvate carboxylase in crassulacean acid metabolism: *in vitro* phosphorylation of the enzyme. *J. Plant Physiol.* **133**, 252-256.

Ko, C.H., Brendel, V., Taylor, R.D. and Walbot, V. (1998) U-richness is a defining feature of plant introns and may function as an intron recognition signal in maize. *Plant Mol. Biol.* **36**, 573-583.

Kopriva, S., Chu, C.C. and Bauwe, H. (1995) H-protein of the cleavage system in *Flaveria*: alternative splicing of the pre-mRNA occurs exclusively in advanced C_4 species. *Plant J.* **10**, 396-374.

Kramer, A. (1996) The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu. Rev. Biochem.* **65**, 367-409.

Laing, W.A. Christeller, J.T., and Sutton, D. (1977) Carbon dioxide fixation by lupin nodules. 2. Studies with ^{14}C-labelled glucose, the pathway of glucaose catabolism and

the effects of some treatments that inhibit nitrogen fixation. *Plant physiol.* **63**, 450-454.

Latzko, E. and Kelly G.J. (1983) The many-faceted function of phosphoenolpyruvate carboxylase in C_3 plants. *Physiol. Veg.* **21**, 805-815.

Laval-Martin, D., Farineau, J. and Diamond, J. (1977) Light versus dark carbon metabolism in cherry tomato fruits. *Plant Physiol.* **60**, 872-876.

Law R.D. and Plaxton W.C. (1997) Regulatory phosphorylation of banana fruit phosphoenolpyruvate carboxylase by a copurifying phosphoenolpyruvate carboxylase-kinase. *Eur J. Biochem.* **247**, 642-651.

Law R.D. and Plaxton W.C. (1995) Phosphorylation *in vitro* and *in vivo* of banana fruit phosphoenolpyruvate carboxylase. *Plant Pysiol.* **108**, 32-32.

Laval, V., Koroleva, O.A., Murphy, E., Lu, C., Milner, J.J., Hooks, M.A. and Tomos, A.D. (2002) Distribution of actin gene isoforms in *Arabidopsis* leaf measured in microsamples from intact cells. *Planta* **15**, 287-292.

Leegood, R.C. (1997) The regulation of C_4 photosynthesis. *Advan. Bot. Res.* **26**, 251-316.

Leegood, R.C. (2002) C_4 photosynthesis: principles of CO_2 concentration and prospects for its introduction into C_3 plants. *J. Exp. Bot.* **53**, 581-590.

Leegood, R.C. and Osmond, C.B. (1990) In: *Plant physiolgy, biochemistry and molecular biology.* (D.T. Dennis and D.H. Turpin, eds.), pp. 284-308, Longman Scientific and Technical, Singapore.

Lepiniec, L., Vidal, J., Chollet, R., Gadal, P. and Cretin, C. (1994) Phosphoenolpyruvate carboxylase-structure, regulation and evolution. *Plant Sci.* **99**, 111-124.

Lepiniec, L., Keryer, E., Philippe, H., Gadal, P. and Cretin, C. (1993) *Sorghum* phosphoenolpyruvate carboxylase gene family: structure, function and molecular evolution. *Plant. Mol. Biol.* **21**, 487-502.

Li, B., Pacquit, V., Jaio, J.A., Duff, S.M.G, Maralihalli, G.B., Sarath, G., Condon, S.A., Vidal, J. and Chollet, R. (1997) Structural requirements for phosphorylation of C_4 leaf phosphoenolpyruvate carboxylase by its highly regulated protein-serine kinase. A comparitive study with synthetic-peptide substrates and mutant target protein. *Aust. J. Plant Physiol.* **24**, 443-449.

Li, B., Zhang, X.Q. and Chollet, R. (1996) Phosphoenolpyruvate carboxylase kinase in tobacco leaves is activated by light in a similar but not identical way as in maize. *Plant Pysiol.* **111**, 497-505.

Li, B. and Chollet, R. (1993) Resolution and identification of C_4 phosphoenolpyruvate carboxylase protein-kinase polypeptides and their reversible light activation in maize leaves. *Arch. Biochem. Biophys.* **307**, 416-419.

Lillo, C., Smith, L.H., Nimmo, H.G. and Wilkins, M.B. (1996) Regulation of nitrate reductase and phosphoenolpyruvate carboxylase activities in barley leaf protoplasts. *Planta* **200**, 181-185.

Liu, H.X., Goodall, G.J., Kole, R. and Filipowicz, W. (1995) Effects of secondary structure on pre-mRNA splicing: hairpins sequestering the 5' but not the 3' splice site inhibit intron processing in *Nicotiana plumbaginifolia*. *EMBO J.* **14**, 377-88.

Luchetta, P., Cretin, C. and Gadal, P. (1990) Structure and characterisation of the *Sorghum vulgare* gene encoding NADP-malate dehydrogenase. *Gene* **89**, 171-117.

Ludwig, M. and Burnell, J.N. (1995) Molecular comparison of carbonic anhydrase from *Flaveria* species demonstrating different photosynthetic pathways. *Plant. Mol. Biol.* **29**, 353-365.

Lorkovic, Z.J., Kirk, D.A.W., Lambermon, M.H.L. and Filipowicz, W. (2000) Pre-mRNA splicing in higher plants. *Trends Plant Sci.* **5**,160-167.

Macdonald, F.D. and Buchanan, B.B. (1990) In: *Plant Physiology, Biochemistry and Molecular Biology*. (Dennis, D. T. and Turpin, D. H. eds.), pp. 239-252, Longman Scientific and Technical, Singapore.

MacKnight, R., Bancroft, I., Page, T., Lister, C., Schmidt, R., Love, K. Westphal, L., Murphy, G., Sherson, S., Corbett, C., and Dean, C. (1997) FCA, a gene controlling flowering time in Arabidopsis, encodes a protein containing RNA-binding domains. *Cell* **89**, 737-745.

MacKnight R., Duroux, M., Laurie, R., Dijkwel, P., Simpson, G. and Dean, C. (2002) Functional significance of the alternative transcript processing of the *arabidopsis* floral promoter FCA. *Plant Cell* **14**, 877-888.

Manh, C.T., Bismuth, E., Boutin, J.P., Provot, M. and Champigny, M.L. (1993) Metabolite effectors for short-term nitrogen-dependent enhancement of phosphoenolpyruvate carboxylase activity and decrease of net sucrose synthesis in wheat leaves. *Physiol Plant* **89**, 460-466.

Mano, S., Hayashi, M. and Nishimura, M. (1999) Light regulates alternative splicing of hydroxypyruvate reductase in pumpkin. *Plant J.* **17**, 309-320.

Marshall, J., Stubbs, J. and Taylor, W. (1996) Two genes encode highly similar chloroplastic NADP-malic enzymes in *Flaveria*: implications for the evolution of C_4 photosynthesis. *Plant Physiol.* **111**, 125-161.

Martin, T., Hellmann, H., Schmidt, R., Willmitzer, L. and Frommer, W.B. (1997) Identification of mutants in metabolically regulated gene expression. *Plant J.* **11**, 53-62.

Matsumura, H., Nagata, T., Terada, M., Shirakata, S., Inoue, T., Yoshinaga, T., Ueno, Y., Saze, H., Izui, K. and Kai, Y. (1999) Crystallization and preliminary X-ray diffraction studies of C_4-form phosphoenolpyruvate carboxylase from maize. *Acta Crystallographica* **55**, 1937-1938.

Matsumura, H., Terada, M., Shirakata, S., Inoue, T., Yoshinaga, T., Izui, K. and Kai, Y. (1999) Plausible phosphoenolpyruvate binding site revealed by 2.6 angstrom structure of Mn^{2+}-bound phospoenolpyruvate carboxylase from *Escherichia coli*. *FEBS lett.* **458**, 93-96.

McGonigle, B. and Nelson, T. (1995) C_4 isoform of NADP-malate dehydrogenase. cDNA cloning and expression in leaves of C_4, C_3 and C_3-C_4 intermediate species of *Flaveria*. *Plant Physiol.* **108**, 1119-1126.

McNaughton, G.A.L., MacKintosh, C., Fewson, C.A., Wilkins, M.B. and Nimmo, H.G. (1991) Illumination increases the phosphorylation stae of maize leaf phosphoenolpyruvate carboxylase by causing an increase in the activity of a protein kinase. *Biochim. Biophys. Acta* **1093**, 189-195.

Meinhard, M. and Schnabl, H. (2001) Fusicoccin- and light-induced activation and *in vivo* phosphorylation of phosphoenolpyruvate carboxylase in *Vicia* guard cell protoplasts. *Plant Sci.* **160**, 635-646.

Metlzer, E. and O'Leary, M.H. (1987) Anaplerotic CO_2 fixation by phosphoenolpyruvate carboxylase in C_3 plants. *Plant Physiol.* **84**, 58-60.

Metzler, M.C., Rothermel, B.A. and Nelson, T. (1989) Maize NADP-malate dehydrogenase: cDNA cloning sequence, and mRNA characterisation. *Plant. Mol. Biol.* **12**, 713-722.

Mohr, W.P. and Stein, M. (1969) Fine structure of fruit development in tomato. *Can. J. Plant Sci.* **49**, 549-553.

Monselie, S.P. (1986) Closing remarks. In: *CRC handbook of fruit set and development.* (Monselie S.P., ed) pp 521-537, CRC Press, Boca Raton, FL,

Moore, R., Clark, D.W. and Vodopich, D.S. (1998) In *Botany*, 2nd edition, WCB McGrawHill.

Muller-Rober, B., Ehrhardt, T. and Plesch, G. (1998) Molecular features of stomatal guard cells. *J. Exp. Bot.* **49**, 293-304.

Nhiri, M., Bakrim, N., Bakrim, N., El Hachimi-Messouak, Z., Echevarria, C. and Vidal, J. (2000) Posttranslational regulation of phosphoenolpyruvate carboxylase during germination of *Sorghum* seeds: influence of NaCl and L-malate. *Plant Sci.* **151**, 29-37.

Nhiri, M., Bakrim, N., Pacquit, V., Osuna, L. and Vidal, J. (1998) Calcium-dependent and -independent phosphoenolpyruvate carboxylase kinases in *Sorghum* leaves: further evidence for the involvement of calcium-independent protein kinase in the *in situ* regulatory phosphorylation of C_4 phosphoenolpyruvate carboxylase. *Plant Cell Physiol.* **39**, 241-246.

Nimmo, G.A., Wilkins, M.B. and Nimmo, H.G. (2001) Partial purification and characterization of a protein inhibitor of phosphoenolpyruvate carboxylase kinase. *Planta* **213**, 250-257.

Nimmo, G.A, Nimmo, H.G., Fewson, C.A. and Wilkins, M.B. (1984) Diurnal changes in the properties of phosphoenolpyruvate carboxylase in *Bryophyllum* leaves: a possible covalent modification. *FEBS Lett.* **178**, 199-203.

Nimmo, G.A., Nimmo, H.G., Hamilton, I.D., Fewson, C.A. and Wilkins, M.B. (1986) Purification of the phosphorylated night form and dephosphorylated day form of phosphoenolpyruvate carboxylase from *Bryophyllum fedtschenkoi*. *Biochem. J.* **239**, 213-220.

Nimmo, G.A., Wilkins, M.B., Fewson, C.A. and Nimmo, H.G. (1987) Persistent circadian rhythms in the phosphorylation state of phosphoenolpyruvate carboxylase from *Bryophyllum fedtschenkoi* leaves and in its sensitivity to inhibition by malate. *Planta* **170**, 408-415.

Nimmo, G.A., McNaughton, G.A.L., Fewson, C.A., Wilkins, M.B. and Nimmo, H.G. (1987) Changes in the kinetic properties and phosphorylation state of phosphoenolpyruvate carboxylase in *Zea mays* leaves in response to light and dark. *FEBS Lett.* **213**, 18-22.

Nimmo, H.G., Fontaine, V., Hartwell, J., Jenkins, G.I., Nimmo, G.A., Wilkins, M.B. (2001) PEP carboxylase kinase is a novel protein kinase controlled at the level of expression. *New Phytologist* **151**, 91-97.

Nimmo, H.G. (2000) The regulation of phosphoenolpyruvate carboxylase in CAM plants. *Trends. Plant. Sci.* **5**, 75-80.

O'Leary, M.H. (1982) Phosphoenolpyruvate carboxylase: an enzymologist's view. *Ann. Rev. Plant Physiol.* **33**, 297-315.

Ogawa, N., Yabuta, N.; Ueno, Y. and Izui, K. (1998) Characterization of a maize Ca^{2+} dependent protein kinase phosphorylating phosphoenolpyruvate carboxylase, *Plant cell Phys.* **39**, 1010-1019.

Ogawa, N., Okumura, S. and Izui, K. (1992) A Ca^{2+}-dependent protein kinase phosphorylates phosphoenolpyruvate carboxylase in maize. *FEBS Lett.* **302**, 86-88.

Or, E., Baybik, J., Sadka, A. and Saks, Y. (2000) Isolation of mitochondrial malate dehydrogenase and phosphoenolpyruvate carboxylase cDNA clones from grape berries and analysis of their expression pattern throughout berry development. *J. Plant. Physiol.* **157**, 527-534.

Osuna, L., Gonzalez, M.C., Cejudo, F.J., Vidal, J. and Echevarria, C. (1996) *In vivo* and *in vitro* phosphorylation of the phosphoenolpyruvate carboxylase from wheat seeds during germination. *Plant Physiol.* **111**, 551-558.

Osmond, C.B. (1978) Crassulacean acid metabolism: a curiosity in context. *Ann. Rev. Plant Physiol.* **29**, 379-414.

Osmond, C.B. (1984) In: *Advances in Photosynthesis Research III.* (C. Sybesma, ed.), pp. 557-564, Martinus Nijhoff/Dr. W. Junk Publishers, The Hague/Boston/Lancaster.

Osuna, L., Pierre, J-N., Gonzalez, M-C., Alvarez, R., Cejud, F.J., Echevarria, C. and Vidal, J. (1999) Evidence for a slow-turnover of the Ca^{2+}-independent phosphoenolpyruvate carboxylase kinase in the aleurone-endosperm tissue of germinating barley seeds. *Plant Physiol.* **119**, 511-520.

Pacquit, V., Santi, S., Cretin, C., Bui, V.T., Vidal, J. and Gadal, P. (1993) Production and properties of recombinant C_3-type phosphoenolpyruvate carboxylase from *Sorghum vulgare*: *in vitro* phosphorylation by leaf and root PyrPCprotein serine kinases. *Biochem. Biophys. Res. Commun.* **197**, 1415-1423.

Page R.D.M. (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* **12**, 357-358.

Pairoba, C.F., Colombo, S.L. and Andreo, C.S. (1996) Flavonoids as inhibitors of NADP-malic enzyme and PEP carboxylase from C_4 plants. *Biosci., Biotech. Biochemi.* **60**, 779-783.

Pathirana, M.S., Samac, D.A., Roeven, R., Yoshioka, H., Vance, C.P., Gantt, J.S. (1997) Analyses of phosphoenolpyruvate carboxylase gene structure and expression in alfalfa nodules. *Plant J.* **12**, 293-304.

Robinson, D.L., Pathirana, S.M., Gantt, J.S. and Vance, C.P. (1996) Immunogold localization of nodule-enhanced phosphoenolpyruvate carboxylase in alfalfa. *Plant Cell Environ.* **19**, 602-608.

Roeske, C.A. and Chollet, R. (1987) Chemical modification of the bifunctional regulatory protein of maize leaf pyruvate, orthophosphate dikinase. Evidence for two distinct active sites. *J Biol Chem.* **262**, 12575-12582.

Roeske, C.A., Kutny, R.M., Budde, R.J. and Chollet, R. (1988) Sequence of the phosphothreonyl regulatory site peptide from inactive maize leaf pyruvate, orthophosphate dikinase. *J Biol Chem.* **263**, 6683-6687.

Rothermel, B.A. and Nelson, T. (1989) Primary structure of the maize NADP-dependent malic enzyme. *J. Biol. Chem.* **264**, 19587-19592.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). In Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, NY, Vol. 1, 2, 3.

Sangwan, R.S., Singh, N. and Plaxton, W.C. (1992) Phosphoenolpyruvate carboxylase activity and concentration in the endosperm of developing and germinating castor oil seeds. *Plant Physiol.* **99**, 445-449.

Saze, H., Ueno, Y., Hisabori, T., Hayashi, H. and Izui, K. (2001) Thioredoxin-mediated reductive activation of a protein kinase for the regulatory phosphorylation of C_4-form phosphoenolpyruvate carboxylase from maize. *Plant Cell Physiol.* **42**, 1295-302.

Schaffner, A.R and Sheen, J. (1990) Metabolic repression of transcription in higher plants. *Plant Cell* **2**, 221-232.

Schuller, K.A., Turpin, D.H. and Plaxton, W.C. (1990) Metabolite regulation of partially purified soybean nodule phosphoenolpyruvate carboxylase. *Plant Physiol.* **94**, 1429-1435.

Schuller, K.A. and Werner, D. (1993) Phosphorylation of soybean (*Glycine max* L.) nodule phosphoenolpyruvate carboxylase *in vitro* decreases sensitivity to inhibition by L-malate. *Plant Physiol.* **101**, 1267-1273.

Schulz, M., Hunte, C. and Schnabl, H. (1992) Multiple forms of phosphoenolpyruvate carboxylase in mesophyll, epidermal and guard cells of *Vicia faba* L. *Physiol. Plant.* **86**, 315-321.

Sheen, J. (1999) C_4 gene expression. *Ann. Rev. Plant. Physiol. Plant. Mol. Biol*, **50**, 187-217.

Simpson, G.G. and Filipowicz, W. (1996) Splicing of precursors to mRNA in higher plants: mechanisms, regulation and sub-nuclear organization of the spliceosomal machinery. *Plant Mol. Biol.* **32**, 1-41.

Smith, C.M., Duff, S.M. and Chollet, R. (1994) Partial purification and characterization of maize-leaf pyruvate, orthophosphate dikinase regulatory protein: a low-abundance, mesophyll-chloroplast stromal protein. *Arch Biochem Biophys*. **308**, 200-206.

Smith, L.H.; Langdale, J.A. and Chollet, R. (1998) A functional Calvin cycle is not indespensable for the light activation of C_4 phosphoenolpyruvate carboxylase kinase and its target enzyme in the maize mutant bundle-sheath defective 2-mutable 1. *Plant Physiol.* **118**, 191-197.

Smith LH, Lillo C, Nimmo HG, Wilkins MB. (1996) Light regulation of phosphoenolpyruvate carboxylase in barley mesophyll protoplasts is modulated by protein synthesis and calcium, and not necessarily correlated with phosphoenolpyruvate carboxylase kinase activity. *Planta* **200**, 174-180.

Streeter, J.G. (1987) Carbohydrate, organic acid, and amino acid composition of bacteroids and cytosol from soybean nodules. *Plant Physiol.* **85**, 768-773.

Sugimoto, T, Kawasaki, T., Kato, T., Whittier, R.F., Shibata, D. and Kawamura, Y. (1992) cDNA sequence and expression of a phosphoenolpyruvate carboxylase gene from soybean. *Plant Mol. Biol.* **20**, 743-747.

Sugimoto, T., Tanaka, K., Monma, M., Kawamura, Y. and Saio, K. (1989) Phosphoenolpyruvate carboxylase level in soybean seed highly correlates to its contents of protein and lipid. *Agric. Biol. Chem.* **55**, 885-887.

Taybi. T., Patil, S., Chollet, R. and Cushman, J.C. (2000) A minimal serine/threonine protein kinase circadianly regulates phosphoenolpyruvate carboxylase activity in crassulacean acid metabolism-induced leaves of the common ice plant. *Plant Physiol.* **123**, 1471-1481.

Terada, K., Kai, T., Okuno, S., Fujisawa, H. and Izui, K. (1990) Maize leaf phosphoenolpyruvate carboxylase: phosphorylation of Ser^{15} with a mammalian cyclic

AMP-dependent protein kinase dimishes sensitivity to inhibition by malate. *FEBS Lett.* 259, 241-244.

Thompson J.D. *et al.* (1997) The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25, 4876-4882.

Toh, H., Kawamura, T. and Izui, K. (1994) Molecular evolution of phosphoenolpyruvate carboxylase. *Plant Cell Environ.* **17**, 31-43.

Tsuchida, Y., Furumoto, T., Izumida, A., Hata, S. and Izui, K. (2001) Phosphoenolpyruvate carboxylase kinase involved in C_4 photosynthesis in *Flaveria trinervia*: cDNA cloning and characterization. *FEBS Lett.* **507**, 318-22.

Ueno, Y., Imanari, E., Emura, J., Yoshizawa-Kumagaye, K., Nakajima, K., Inami, K., Shiba, T., Sakakibara, H., Sugiyama, T. and Izui, K. (2000) Immunological analysis of the phosphorylayion state of maize of C_4-form phosphoenolpyruvate carboxylase with specific antibodies raised against a synthetic phosphorylated peptide. *Plant J.* **21**, 17-26.

Usuda, H. (1988) Adenine-nucleotide levels, the redox state of the NADP system, and the assimilatory force in nonaqueously purified mesophyll chloroplasts from maize leaves under different light intensities. *Plant Physiol.* **88**, 1461-1468

Valcarcel, J. and Green, M.R. (1996) The SR protein family: pleiotropic functions in pre-mRNA splicing. *Trends Biochem. Sci.* **21**, 296-301.

Van Quy, L., Foyer, C. and Champigny, M.L. (1991a) Effect of light and NO_3^- on wheat leaf phosphoenolpyruvate carboxylase activity: evidence for the covalent modification of the C_3 enzyme. *Plant Physiol.* **97**, 1476-1482

Van Quy, L., Lamaze, T. and Champigny, M.L. (1991b) Short-term effects of nitrate on sucrose synthesis in wheat leaves. *Planta* **185**, 53-57.

Vance, C. P. (1983) *Rhizobium* infection and nodulation: A beneficial plant disease? *Ann. Rev. Microbiol.* **37**, 399-424.

Vance, C.P., Gregerson, R.G., Robinson, D.L., Miller, S.S. and Grant J.S. (1994) Primary assimilation of nitrogen in alfalfa nodules: molecular features of the enzymes involved. *Plant Sci.* **101**, 51-64.

Vance, C. P. and. Heichel G. H. (1991) Carbon in N_2 fixation: Limitation or exquisite adaptation. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* **42**, 373-392.

Varga, J. and Bruinsma, J. (1986) Tomato. In: CRC handbook of fruit set and development (Monselie S.P., ed.), pp. 461-491.CRC Press, Boca Raton, FL,

Vazquez-Tello, A., Whittier, R.F., Kawasaki, T., Sugimoto, T., Kawamura, Y., Shibata, D. (1993) Sequence of a soybean (*Glycine-Max* L) Phosphoenolpyruvate Carboxylase cDNA. *Plant Physiol.* **103**, 1025-1026.

Vidal, J. and Chollet, R. (1997) Regulatory phosphorylation of C_4 PEP carboxylase. *Trends Plant Sci.* **2**, 230-237.

Vidal, J., Nguyen, J., Perrot-Rechenmann, C. and Gadal, P. (1986) Phosphoenolpyruvate carboxylase in soybean root nodules. An immunochemical study. *Planta* **169**, 198 201.

Voznesenskaya, E.V., Franceschi, V.R., Kiirats, O., Freitag, H. and Edwards, G.E. (2001) Kranz anatomy is not essential for terrestrial C_4 plant photosynthesis. *Nature* **414**, 543-546.

Voznesenskaya, E.V., Franceschi, V.R., Kiirats, O., Artyusheva, E.G., Freitag, H. and Edwards, G.E. 2002. Proof of C_4 photosyhnthesis without kranz anatomy in *Bienertia cyclptera* (Chenopodiaceae). *Plant J.* **31**, 649-662.

Wadham, C., Winter, H. and Schuller, K.A. (1996) Regulation of soybean nodule phosphoenolpyruvate carboxylase *in vivo*. *Physiol Plant.* **97**, 531-535.

Wang, Y.H. and Chollet, R.(1993a) *In vitro* phosphorylation of purified tobacco-leaf phosphoenolpyruvate carboxylase. *FEBS Lett.* **328**, 215-218.

Wang, Y.H. and Chollet, R. (1993b) Partial purification and characterization of phosphoenolpyruvate carboxylase protein-serine kinase from illuminated maize leaves. *Arch. Biochem. Biophys.* **304**, 496-502.

Wang, Y.H., Duff, S.M.G., Lepiniec, L., Cretin, C., Sarath, G., Condon, S.A., Vidal, J., Gadal, P. and Chollet, R. (1992) Site-directed mutagenesis of a phosphorylatable serine (Ser^8) in C_4 phosphoenolpyruvate carboxylase from *Sorghum*: the effect of negative charge at position 8. *J. Biol. Chem.* **267**, 16759-16762.

Weber, H., Borisjuk, L., and Wobus, U. (1997) Sugar import and metabolism during seed development. *Trends Plant Sci.* **22**, 169-174.

Weigend, M. (1994) *In vivo* phosphorylation of phosphoenolpyruvate carboxylase from the facultative CAM plant *Mesembryanthemum crystallinum*. *J.Plant Physiol.* **144**, 654-660.

Werneke, J.M., Chatfield, J.M., and Ogren, W.L. (1989) Alternative mRNA splicing generates the two ribulosebisphosphate carboxylase/oxygenase activase polypeptides in spinach and *Arabidopsis*. *Plant Cell* 1, 815-825.

Wilkins, M.B. (1992) Circadian rhythms: their origin and control. *New Phytol.* **121**, 347-375.

Winter, K. (1982) Properties of phosphoenolpyruvate carboxylase in rapidly prepared, desalted leaf extracts of the Crassulacean acid metabolism plant *Mesembryanthemum crystallinum* L. *Planta* **154**, 298-308.

Winter, K. and Smith, J.A.C. (1996) In: *Crassulacean Acid Metabolism - biochemistry, ecophysiology and evolution.* (K. Winter and J. A. C. Smith, eds.), Ecological Studies 114, pp. 1-13, Springer-Verlag, Berlin Heidelberg.

Woo, K. C. and Xu S. (1996) Metabolite regulation of phosphoenolpyruvate carboxylase in legume nodules. *Austral. J. Plant Physiol.* **23**, 413-419.

Yannagisawa, S., Izui, K., Yamaguchi, Y., Shigesuda, K. and Katsuki, H. (1988) Further analysis of cDNA clones for maize phosphoenolpyruvate carboxylase involved in C_4 photosynthesis: nucleotide sequence of entire open reading frame and evidence for polyadenylation of mRNA at multiple sites *in vivo*. *FEBS Lett.* **229**, 107-110.

Yano, Y., Terada, K., Umiji, K. and Izui, K. (1995) Catalytic role of an arginine residue in the highly conserved and unique sequence of phosphoenolpyruvate carboxylase. *J. Biochem.* **117**, 1196-1200.

Yu, J., Hu, S.N., Wang, J., Wong, G.K.S., Li, S.G., Liu, B., Deng, Y.J., Dai, L., Zhou, Y., Zhang, X.Q., Cao, M.L., Liu, J., Sun, J.D., Tang, J.B., Chen, Y.J., Huang, X.B., Lin, W., Ye, C., Tong, W., Cong, L.J., Geng, J.N., Han, Y.J., Li, L., Li, W., Hu, G.Q., Huang, X.G., Li, W.J., Li, J., Liu, Z.W., Li, L., Liu, J.P., Qi, Q.H., Liu, J.S., Li, L., Li, T., Wang, X.G., Lu, H., Wu, T.T., Zhu, M., Ni, P.X., Han, H., Dong, W., Ren, X.Y., Feng, X.L., Cui, P., Li, X.R., Wang, H., Xu. X,, Zhai, W.X., Xu, Z., Zhang, J.S., He, S.J., Zhang, J.G., Xu, J.C., Zhang, K.L., Zheng, X.W., Dong, J.H., Zeng, W.Y., Tao, L., Ye, J., Tan, J., Ren, X.D., Chen, X.W., He, J., Liu, D.F., Tian, W., Tian, C.G., Xia, H.G., Bao, Q.Y., Li, G., Gao, H.,Cao, T., Wang, J., Zhao, W.M., Li, P., Chen, W., Wang, X.D., Zhang, Y., Hu, J.F., Wang, J., Liu, S., Yang, J., Zhang, G.Y., Xiong, Y.Q., Li, Z.J., Mao, L., Zhou, C.S., Zhu, Z.,Chen, R.S., Hao, B.L., Zheng, W.M., Chen, S.Y., Guo, W., Li, G.J., Liu, S.Q., Tao, M., Wang, J., Zhu, L.H., Yuan, L.P. and Yang, H.M. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp indica). *Science* **296**, 79-92.

Zhang, X.Q. and Chollet, R. (1997) Phosphoenolpyruvate carboxylase protein kinase from soybean root nodules: Partial purification, characterization, and up/down-

regulation by photosynthate supply from the shoots. *Arch. Biochem. Biophys.* **343**, 260-268.

Zhang, X.Q., Li, B. and Chollet, R. (1995) *In vivo* phosphorylation of soybean nodule phosphoenolpyruvate carboxylase. *Plant Physiol.* **108**, 1561-1568.

Zhang, S.Q., Outlaw, W.H. and Chollet, R. (1994) Lessened inhibition of guard-cell phosphoenolpyruvate carboxylase velocity during stomatal opening. *FEBS Letts.* **352**, 45-48.

Appendix A: The panel shows a partial nucleotide sequence of a cloned PCR product obtained from maize gDNA using the universal PPCK primers designed from the conserved domains of *KfPPCK* (as outlined in section 3.2.1). The sequence shows no significant homology to any known gene, and is therefore thought to be intergenic maize DNA.

The red font indicates one of the primer hybridisation sites.

TGCGAGGAGA TCGGCCGKG

TGCGAGGAGA TCGGCCGGGC GTACTGTCTG TCACTGAGGC TCACCTGTCG
GCGACGTGGT CCAGCTGCGG CAGCGACGAC TGCCACCGGC GCGGGAGCAG
AAGGGAGCAT CCGCCGCCTC CTCCTCCTCC TCCGGCCTGC GAACAGCGAG
CAGCGAAAGG GAAGAGGTTA CACCGACCTC TGAGCATGGA TGATCTCATC
GCGGCGGGGA GAGCCATCGA TCGTCCTGGA CGCGCGGCGG CGGCCACGGC
CGAATTAAAG AAGAAACGAT CGGTTACGGA AACGAGTAAA AGCGGGCCGC
ACAGTCCAGC TTCCACACGG CGAGAGGTTA ACGCATGGGA TTGGGATGTG
ATGTGAACGG TGCGCACAGG CACACACTAC ACACGGCGAG AATCGCGCAC
GCACCGTAGC CAACGCCTTC CCAGGACCTC CTCAAGTC

Appendix B: Panel A shows the full nucleotide sequence of a cloned PCR product obtained from illuminated leaf maize cDNA using the universal PPCK primers designed from the conserved domains of *KfPPCK* (as outlined in section 3.2.1). The sequence is identical to the *Zea mays CDPK1* gene

The red font indicates both of the primer hybridisation sites.

Panel B shows the aminoacid sequence of ZmCDPK1.

The blue font indicates the region cloned and sequenced.

A

```
TCCGCYTGCA ARWCSATCG
TCCGCCTGCA AAACGATCGC CAAGCGCAAG CTGCTCACCC CCGAGGACGT
CGACGACGTG CGGCGCGAGA TCCAGATCAT GCACCACCTC GCCGGCCACA
AGAGCGTCGT CACCATCAAG GGCGCCTACG AGGACCCGCT CTACGTCCAC
ATCGTGATGG AGCTCTGCGA GGGCGGCGAG CTCTTCGACC GCATCGTCGA
CAGGGGGTAC TTCTCCGAGC GCAAGGCCGC CGAGATCGCG CGCGTCATCG
TCGGCGTCGT GGAGGCGTGC CACTCGCTGG GCGTCATGCA CCGGGACCTC
AAGCCCGAGA ACTTCTTGCT CAAGGACAGG GGCCACGACG CGTCGCTCAA
GGCCATCGAC TTCGGCCTCT CCGTCTTCTT CAAGCCCGGC CAGGTGTTCA
CCGACGTGGT GGGCTCCCCC TACTACGTGG CGCCGGAATC
                    G TAYTACGTGG CGCCGGAGGT
```

B

```
Zm CDPK1
MRRGGAGAPP DLGSVLGHTT PNLRDLYALG RKLGQGQFGT TYLCTELATG
IDYACKSISK RKLITKEDVD DVRREIQIMH HLSGHKNVVA IKGAYEDQVY
VHIVMELCAG GELFDRIIQR GHYSERKAAA LTRIIVGVVE ACHSLGVMHR
DLKPENFLLA NRDDDLSLKA IDFGLSVFFK PGQVFTDVVG SPYYVAPEVL
LKSYGPAADV WTAGVILYIL LSGVPPFWAE TQQGIFDAVL KGAIDFDSDP
WPVISDSAKD LIRRMLNPRP AERLTAHEVL CHPWIRDHGV APDRPLDPAV
LSRIKQFSAM NKLKKMALRV IAESLSEEEI AGLKEMFQTM DTDNSGAITY
DELKEGLRKY GSTLKDTEIR DLMDAADIDN SGTIDYIEFI AATLHLNKLE
REEHLVAAFS YFDKDGSGYI TVDELQLACK EHNMPDAFLD DVINEADQDN
DGRIDYGEFV AMMTKGNMGV GRRTMRNSLN ISMRDDLVCS ET
```

Appendix C: Panel A shows the partial nucleotide sequence of a cloned PCR product obtained from illuminated leaf maize cDNA using a universal PPCK primer, designed from the conserved domains of *KfPPCK*, and a reverse dT oligo primer. The PCR product was shown to be preferentially expressed in the 'light' pool of maize leaf RNA by DDrtPCR (as outlined in section 3.2.1). The sequence is very similar to the *Glycine max SPK-3* gene.

The red font indicates one of the primer hybridisation sites.

Panel B shows the aminoacid sequence of the predicted translated product of the PCR product that bears similarity to GmSPK-3.

A

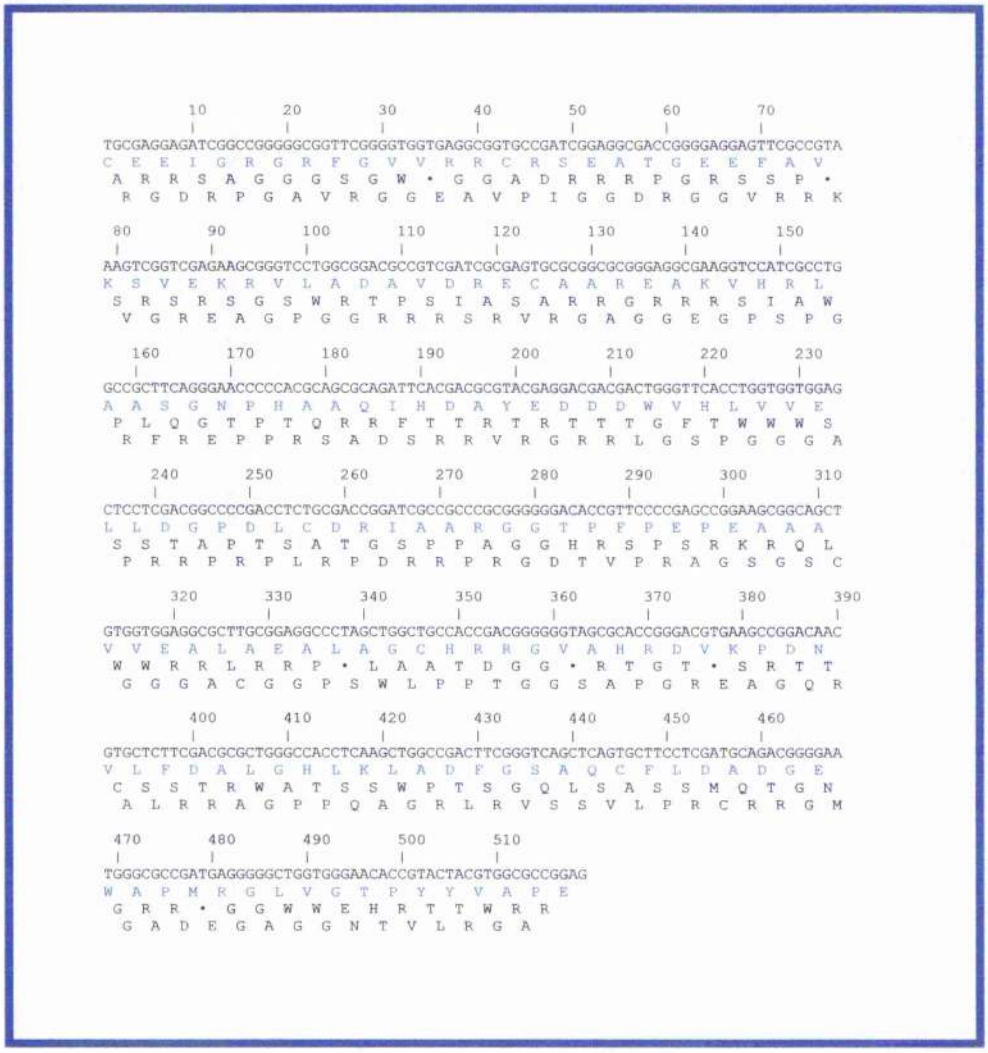
```
GGCCACCGGG ACATCAARCC NGANAA
GGCCACCGGG ACATCAAGCC GGATAACACG CTGCTGGACG GCAGCCCGGC
GCCGTGCCTC AAGATCTGCG ACTTCGGCTA CTCCAAGTCC TCGGTGCTGC
ACTCGCGACC GAAGTCAACG GTGGGCACGC CGGCGTACAT CGCGCCCGAG
GTGCTGTCGC GGCGAGAGTA CGACGGCAAG CACGCCGACG TCTGGTCGTG
CGGAGTCACG CTCTACGTCA TGCTGGTGGG GGCGTACCCG TTCGAGGATC
CAAAGGACCC AAAGAACTTC AGGAAGACAA TACAGCGCAT CATGTCGGTT
CAGTACAAGA TCCCCGAGTA CGTGCACGTC TCCCACAACT GCCGCCACCT
CCCCTCCCGC ATCTTCGTCC AGAACCCGTA CAAGAGGATC ACCATGAGCG
AGATCAAGAC CCACCCCTGG TACCTCAAGA ACCTGCCGAG GGAGCTCAAG
GAGGAGGCGC AGGCAGCCTA CTACTACAGC CGCCAGGGAG ACAGCAGCAG
CAGCAGCAGC AATAATGCGA CAGCAGCACC CGCCTACTCG TCGCAGAGCG
TGGAGGAGAT CTTGAGGATC GTGCAGGAGG CGCAGACGGT GCCCAAGCCA
CCGTCCAGGC CCGACGAAGC AGAGCAACAG GAAGAAGACG ACGACTATGA
CGATGACGAC GACTACGACA GAACGGTGCG GCAGGTGCAC GCCAGCGGAG
AGTTCGACAT GATTATGAGC AGACTCCAAA TCTGACTCTT AGTATCTTGC
CTGCCATTAT CCGGCTT
```

B

```
NSDGHRDIKP DNTLLDGSPA PCLKICDFGY SKSSVLHSRP KSTVGTPAYI
APEVLSRREY DGKHADVWSC GVTLYVMLVG AYPFEDPKDP KNFRKTIQRI
MSVQYKIPEY VHVSHNCRHL PSRIFVQNPY KRITMSEIKT HPWYLKNLPR
ELKEEAQAAY YYSRQGDSSS SSSNNATAAP AYSSQSVEEI LRIVQEAQTV
PKPPSRPDEA EQQEEDDDYD DDDDYDRTVR QVHASGEFDM IMSRLQI•
```

Appendix D: Translation map of the partial sequence of a putative OsPPCK

Blue font indicates the correct reading frame.


10 20 30 40 50 60 70
TGCGAGGAGATCGGCCGGGGGCGGTTCGGGGTGGTGAGGCGGTGCCGATCGGAGGCGACCGGGGAGGAGTTCGCCGTA
C E E I G R G R F G V V R R C R S E A T G E E F A V
A R R S A G G G S G W • G G A D R R R P G R S S P •
R G D R P G A V R G G E A V P I G G D R G G V R R K
80 90 100 110 120 130 140 150
AAGTCGGTCGAGAAGCGGGTCCTGGCGGACGCCGTCGATCGCGAGTGCGCGGCGCGGGAGGCGAAGGTCCATCGCCTG
K S V E K R V L A D A V D R E C A A A R E A K V H R L
S R S R S G S W R T P S I A S A R R G R R R S I A W
V G R E A G P G G R R R S R V R G A G G E G P S P G
160 170 180 190 200 210 220 230
GCCGCTTCAGGGAACCCCCACGCAGCGCAGATTCACGACGCGTACGAGGACGACGACTGGGTTCACCTGGTGGTGGAG
A A S G N P H A A Q I H D A Y E D D D W V H L V V E
P L Q G T P T Q R R F T T R T R T T T G F T W W W S
R F R E P P R S A D S R R V R G R R L G S P G G G A
240 250 260 270 280 290 300 310
CTCCTCGACGGCCCCGACCTCTGCGACCGGATCGCCGCCCGCGGGGGGACACCGTTCCCCGAGCCGGAAGCGGCAGCT
L L D G P D L C D R I A A R G G T P F P E P E A A A
S S T A P T S A T G S P P A G G H R S P S R K R Q L
P R R P R P L R P D R R P R G D T V P R A G S G S C
320 330 340 350 360 370 380 390
GTGGTGGAGGCGCTTGCGGAGGCCCTAGCTGGCTGCCACCGACGGGGGGTAGCGCACCGGGACGTGAAGCCGGACAAC
V V E A L A E A L A G C H R R G V A H R D V K P D N
W W R R L R R P • L A A T D G G • R T G T • S R T T
G G G A C G G P S W L P P T G G S A P G R E A G Q R
400 410 420 430 440 450 460
GTGCTCTTCGACGCGCTGGGCCACCTCAAGCTGGCCGACTTCGGGTCAGCTCAGTGCTTCCTCGATGCAGACGGGGAA
V L F D A L G H L K L A D F G S A Q C F L D A D G E
C S S T R W A T S S W P T S G Q L S A S S M Q T G N
A L R R A G P P Q A G R L R V S S V L P R C R R G M
470 480 490 500 510
TGGGCGCCGATGAGGGGGCTGGTGGGAACACCGTACTACGTGGCGCCGGAG
W A P M R G L V G T P Y Y V A P E
G R R • G G W W E H R T T W R R
G A D E G A G G N T V L R G A

Appendix E: Panel A shows the partial nucleotide sequence of a cloned product isolated from a maize seedling library using a putative partial rice PPCK sequence as a probe. (as outlined in section 3.2.3). The sequence is very similar to a gene encoding the 6.1 kDa polypeptide of photosystem II of maize.
Panel B shows the aminoacid sequence of the predicted translated product of the isolated product that bears similarity to 6.1 kDa polypeptide of photosystem II.

A

```
TGAACGCCCA CCGGAGAACA ACGCACTTGT GTTGCGTACG TGCGGTCACA
TGCATTGCAT CTGCATCCGG AGTAGCGAGT TCTAGACACA GCAGGCAACA
GACTGAGGTA GCCCAGCGAG CGAGCGAGCA ATGGCAACCC TCAGCGCCGC
CCCCCTCGTC GGCGCGGCCG CCGTCGCCAG GCCGTGTCAG GCTCAAGGTT
TGCCGCAGCT GAGGGTGAGA GCCGAGAAGG CGAGGTGCGG CGCCGCCCAC
TCGAGGCGGC CGAGCCAGCG GCGGGACGGC AACAACGGCG CATCGTCGTC
GCTTCTGGCC GTGGCCACCA GCGCGGTGAC CACGTCGCCC GCGCTGGCGC
TGGTGGACGA GCGGATGTCG ACGGAGGGCA CCGGGCTCAG CCTGGGGCTC
AGCAACAACC TGCTGGGGTG GATCCTGCTG GGCGTCTTCG GCCTCATCTG
GTCCCTCTAC ACCGTCTACA CCTCCACGCT CGACGAGGAC GACGACTCCG
GCCTCTCGCT CTGAACTCAT CCGTCGATTA TTCCTAGCTA CGACTGTATT
TTTTGTTCAG CCTGACCCGT GAGTTCGATC GATCTATGCA TGCGTTGGTG
CAGACGAAGC TTAAGACCGT GGTTGTATGC GACCGTAATT CAGTTGTTGC
ATCTCAGATG GTAATGATAA TGACGAGTAC GTAGAACCAT TCCAGCAAAA
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA A
```

B

```
MATLSAAPLV GAAAVARPCQ AQGLPQLRVR AEKARCGAAH SRRPSQRRDG
NNGASSSLLA VATSAVTTSP ALALVDERMS TEGTGLSLGL SNNLLGWILL
GVFGLIWSLY TVYTSTLDED DDSGLSL
```