Structural and functional studies of phosphoenolpyruvate carboxylase kinase

by

Gregor Robert Lennie

Thesis submitted for the degree of Master of Science

School of Life Sciences

College of Medical, Veterinary and Life Sciences

University of Glasgow

September 2010

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Abstract

Phosphoenolpyruvate carboxylase (PEPc) catalyses the irreversible β-carboxylation of phosphoenolpyruvate (PEP) to form oxaloacetate and inorganic phosphate. PEPc plays an anaplerotic role in all plants and bacteria, providing citric acid cycle intermediates. The enzyme also catalyses the primary carboxylation step in C_4 photosynthesis and Crassulacean Acid Metabolism (CAM), and has more specialised roles in guard cell opening and fruit ripening. Due to the importance of the role of PEPc, the enzyme must be tightly controlled. This is achieved in two ways: allosteric regulation and covalent modification. The enzyme can be activated by glucose 6-phosphate and inhibited by malate. Superimposed on this, phosphorylation of PEPc alters the allosteric properties, making it more susceptible to activation and less susceptible to inhibition. The enzyme that catalyses phosphorylation of PEPc is PEPc kinase.

The initial objective of this project was to solve the structure of PEPc kinase by X-ray crystallography. Crystallisation conditions in the presence and absence of an ATP analogue were investigated using PEPc kinase from soybean expressed heterologously in yeast, but only one "hit" of microcrystals was obtained. Using a fresh batch of protein, optimisation around the conditions that gave microcrystals was unsuccessful. Owing to time constraints, further attempts at cystallisation were abandoned. The use of NMR to solve the structure of the kinase in solution was not practicable owing to the relatively low amounts of kinase that can be obtained.

The quaternary structure of the protein was analysed using size exclusion chromatography on Superose 12. This showed that the protein, despite migrating

as a single band on an SDS-PAGE gel, was composed of two peaks with different specific activities. The relative intensities of the two peaks were not influenced by temperature or the presence of ATP/Mg. Since the two peaks could be due to oligomerisation, analytical ultracentrifugation analysis of recombinant PEPc kinase was carried out. Sedimentation velocity analysis gave a lower sedimentation coefficient than a homology model of soybean PEPc kinase would suggest. Sedimentation equilibrium analysis showed no evidence for oligomerisation of the protein. Circular dichroism spectroscopy showed the presence of significant amounts of α-helix and ß-sheet. Thus it was suggested that the two peaks in size exclusion chromatography arise due to the existence of two different monomeric conformations, one more open than the other.

All plants contain a small family of PEPc genes and a small family of PEPc kinase genes, raising the possibility that specific PEPc kinases could target specific PEPcs. The kinetic properties of PEPc kinases from soybean, Sorghum and Arabidopsis were measured using PEPcs from Kalanchoë, maize and rice. The data show that all of the PEPc kinases can use all of the PEPc isoforms, arguing against kinetically determined specific PEPc/PEPc kinase pairs.

It has recently been discovered that an isoform of PEPc is targeted to the chloroplast of rice. This isoform contains the consensus phosphorylation sequence for PEPc kinase, which raises the question of whether this protein is a substrate for the known rice PEPc kinases. The rice PEPc kinases were expressed *in vitro* and used in a kinase assay with rice chloroplast PEPc as substrate. This confirmed that rice PEPc kinases can phosphorylate rice chloroplast PEPc *in vitro*.

Table of Contents

List of Figures

Acknowledgements

I would first of all like to acknowledge my supervisor, Professor Hugh Nimmo, for giving me the opportunity to carry out my research in his laboratory. His feedback and understanding have kept me going throughout my PhD and I will always be grateful for his support.

I would also like to acknowledge the fellow members of the Bond Lab, past and present, for their technical assistance and helpful discussions (and for sometimes tidying my desk).

On the technical side of things, I would like to thank Dr Niall Fraser and Dr Mads Gabrielsson for help with crystallography, Dr Sharon Kelly for help with circular dichroism, and Dr Olwyn Byron for help with analytical ultracentrifugation.

Finally, I would like to thank my family for their endless support during my education.

Author's Declaration

I declare that the work presented in this thesis is my own and that it has not been submitted for a degree at another institution.

Abbreviations

ATP	Adenosine 5'-triphosphate
AUC	Analytical ultracentrifugation
CAM	Crassulacean acid metabolism
cAMP	Cyclic 3,5'-monophosphate
CD	Circular dichroism
DNAse	Deoxyribonuclease
DTT	Dithiothreitol
LB	Luria-Bertani
NAD	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide (reduced form)
NMR	Nuclear magnetic resonance
PAGE	Polyacrylamide gel electrophoresis
PEG	Polyethylene glycol
PEP	Phosphoenolpyruvate
PEPc	Phosphoenolpyruvate carboxylase
PKA	Protein kinase A
Rubisco	Ribulose-1,5-bisphosphate carboxylase/oxygenase

SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
YPD	Yeast-peptone-dextrose

1 Introduction

Phosphoenolpyruvate carboxylase (PEPc, E.C. 4.1.1.31) catalyses the irreversible β-carboxylation of phosphoenolpyruvate (PEP), in the presence of Mg^{2+} or Mn^{2+}, to yield oxaloacetate and inorganic phosphate (Chollet *et al.*, 1996). PEPc is ubiquitous in photosynthetic organisms, such as plants, algae and cyanobacteria, and also present in nonphotosynthetic organisms, such as bacteria and protozoa. PEPc is not found in animals, yeast and fungi however. PEPc plays an anaplerotic role in all plants and bacteria, supplying citric acid cycle intermediates. In addition to this, the enzyme has a more specialised role in C_4 photosynthesis and Crassulacean Acid Metabolism (CAM), catalysing the primary carboxylation step in these metabolic pathways. Also, PEPc is involved in fruit ripening and guard cell opening in higher plants, and provides carbon skeletons as acceptors for nitrogen fixation. As it has been shown to have a varied number of roles, the activity of PEPc must be tightly controlled. This control is achieved in two ways: allosteric regulation by small molecules (such as malate and aspartate) and covalent modification by reversible phosphorylation.

The aims of the first chapter in this thesis are: (1) to describe the structures and roles of PEPc in various organisms; (2) to summarise the molecular details of catalysis by PEPc; (3) to explain the regulation of PEPc at the molecular level; and (4) to describe the features of the enzyme responsible for regulation of PEPc.

1.1 PEPc in higher plants

Higher plant PEPc isoforms are typically homotetramers of ~110 kDa polypeptides, encoded by a small gene family (Chollet *et al.*, 1996). For example, the model plant *Arabidopsis thaliana* has four PEPc genes (Sanchez & Cejudo, 2003). These genes encode isoforms with tissue-specific distribution and functions. PEPc is thought to be cytosolic but there one exception to this is noted in Chapter 5.

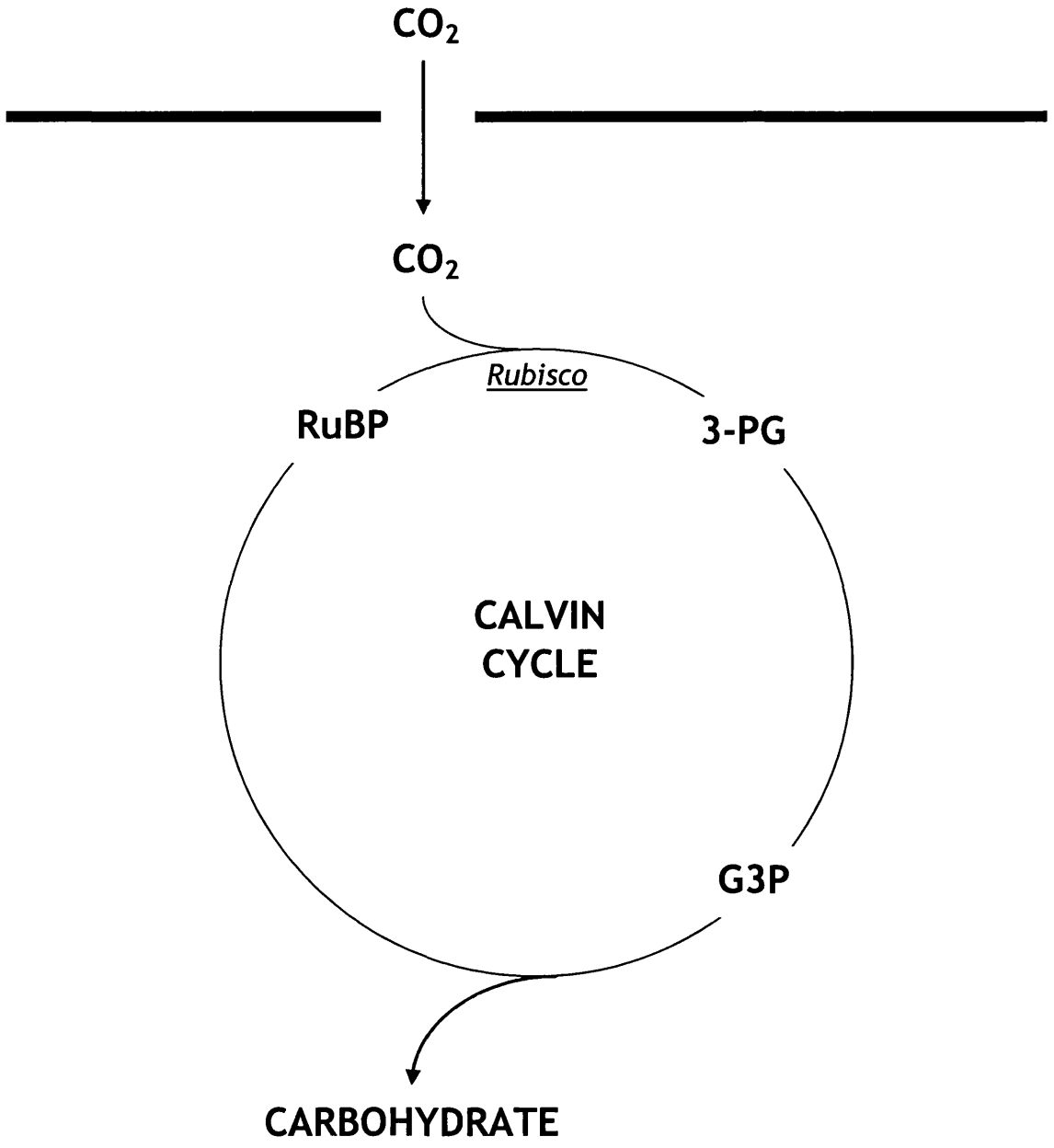
The primary role of PEPc in the leaves of C_4 and CAM plants is to fix CO_2. C_4 photosynthesis is an adaptation to environments that are hot, dry and have high light intensities while CAM is an adaptation to arid environments. The primary carboxylation step in C_3 photosynthesis (Fig. 1.1A) is catalysed by ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco). This enzyme is extremely inefficient, with a turnover of 3 molecules of CO_2 per second, and in higher temperatures the ratio of oxygenase activity to carboxylase activity increases (Andersson, 2008). This leads to an accumulation of glyoxylic acid, which must be removed by a process called photorespiration (Foyer *et al.*, 2009). Energy, in the form of adenosine 5'-triphosphate (ATP), is consumed by this process so if photorespiration can be avoided the organism would be at an evolutionary advantage. More importantly, the reactions required for recycling the products of photorespiration leads to a loss of fixed nitrogen and fixed carbon.

In C_4 photosynthesis, CO_2 is concentrated in a spatial manner (Leegood, 2002). C_4 plants have a number of specific features, including a highly expressed PEPc isoform in mesophyll cells. C_4 plants also have a specialised leaf anatomy, known as Kranz anatomy (Fig.1.1B), which allows CO_2 to be fixed by mesophyll

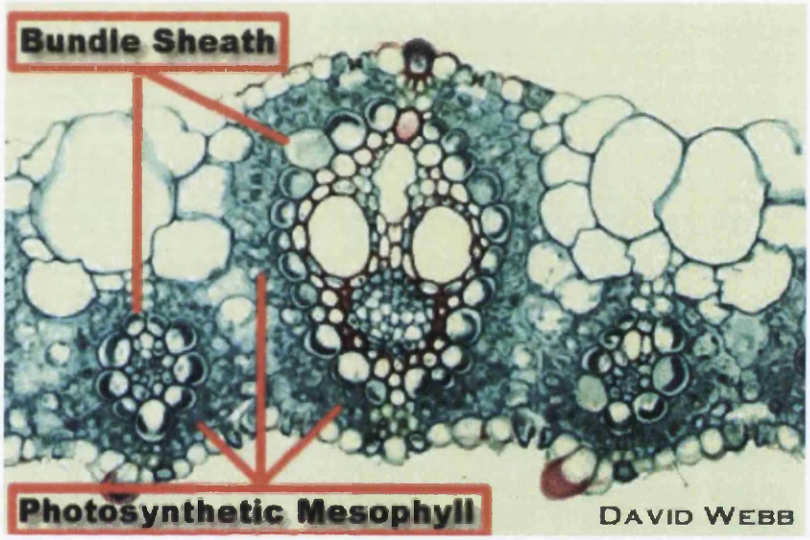
Figure 1.1 Photosynthesis

In C_3 photosynthesis, CO_2 is fixed by Rubisco to produce carbohydrate (A). C_4 plants have a specialised leaf anatomy compared with C_3 and CAM plants. The word "kranz" is German for "wreath", and is used to describe the organisation of mesophyll cells around the bundle-sheath cells (B). CO_2 is taken up by the mesophyll cells and fixed by PEPc, and then transported to the bundle-sheath cells where CO_2 fixation by Rubisco takes place (C). In CAM, CO_2 assimilation and carbon fixation are separated temporally (D). RuBP = ribulose 1,5-bisphosphate; 3-PG = 3-phosphoglycerate; G3P = glyceraldehyde 3-phosphate; OAA = oxaloacetate. Image in B taken from http://www.botany.hawaii.edu/faculty/Webb/BOT311/BOT311-00/PSyn/PsynDark2.htm

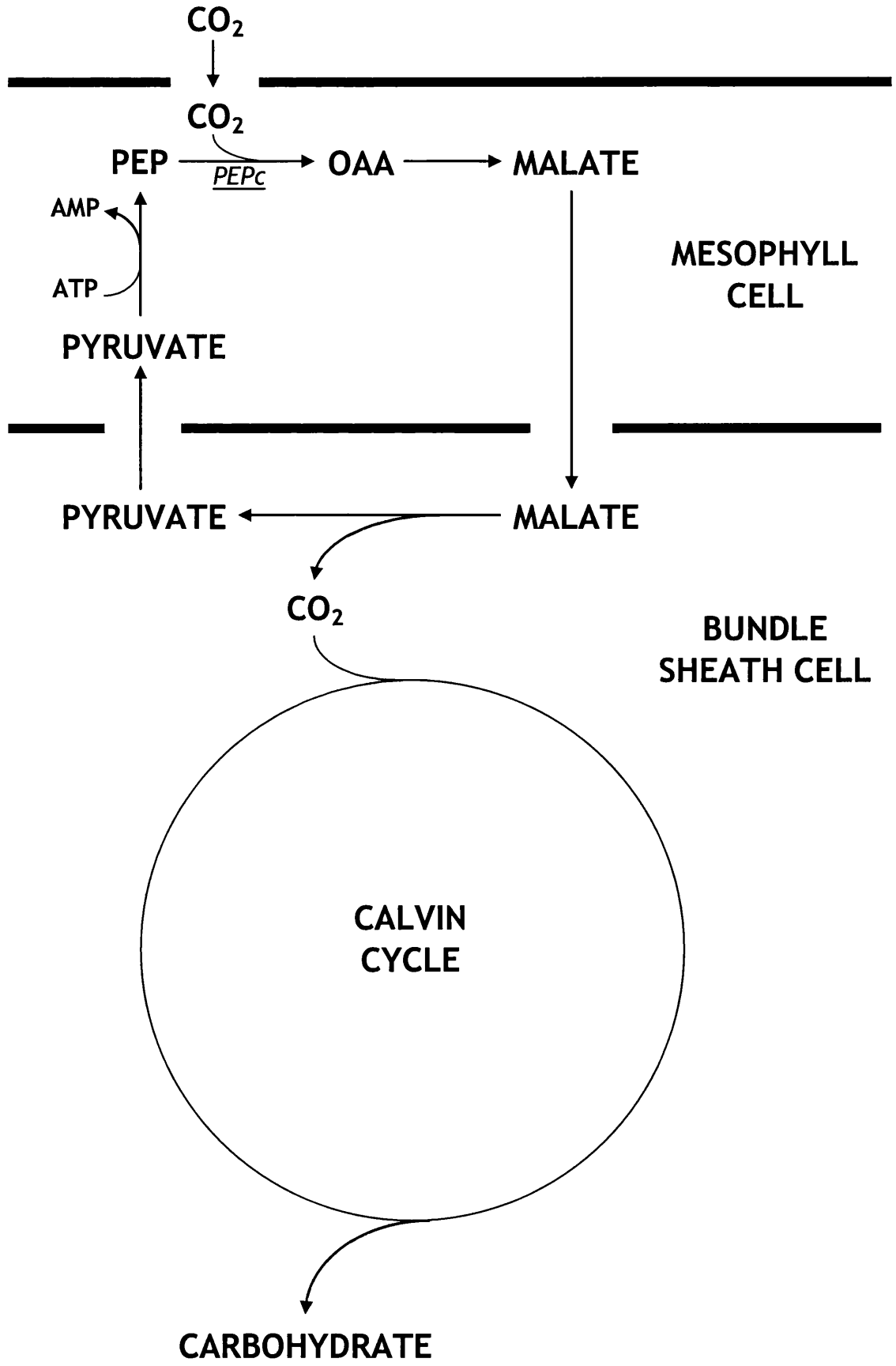
A



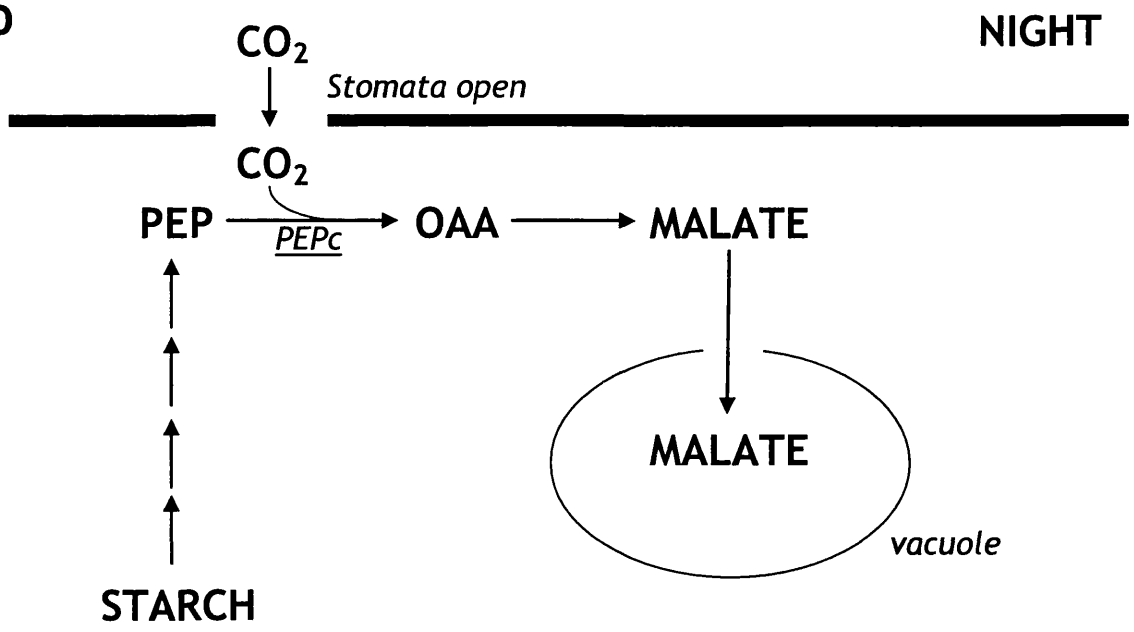
B

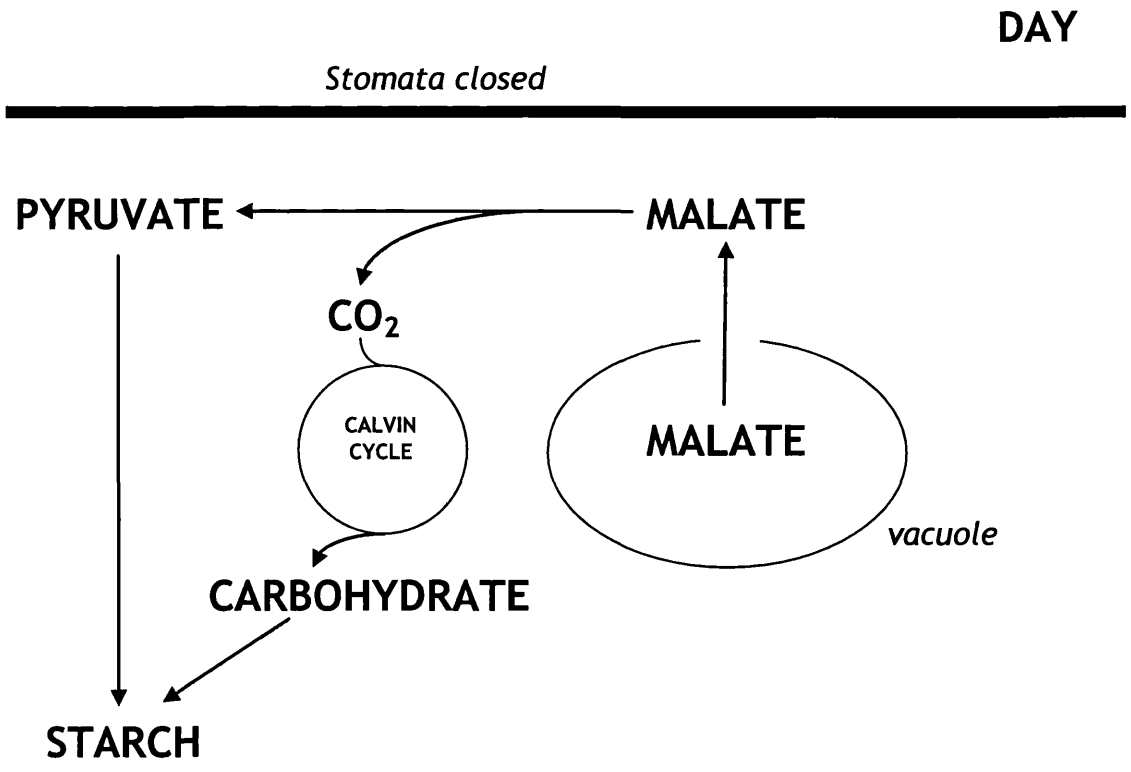


C


CO_2
CO_2
PEP
PEPc
OAA
MALATE
AMP
ATP
MESOPHYLL CELL
PYRUVATE
PYRUVATE
MALATE
CO_2
BUNDLE SHEATH CELL
CALVIN CYCLE
CARBOHYDRATE

D


NIGHT
CO_2
Stomata open
CO_2
PEP
PEPc
OAA
MALATE
MALATE
vacuole
STARCH


DAY
Stomata closed
PYRUVATE
MALATE
CO_2
CALVIN CYCLE
MALATE
vacuole
CARBOHYDRATE
STARCH

cells and then transported to the bundle-sheath cells where it is released and fixed by Rubisco. There is considerable interest in genetically engineering C_3 plants such as rice to use the C_4 photosynthesis pathway, to minimise photorespiration and raise water-use efficiency (Miyao & Fukayama, 2003).

In CAM, CO_2 is concentrated in a temporal manner (Cushman, 2001). Stomata allow gas and water exchange with the environment. Unlike C_3 and C_4 photosynthesis, in CAM stomata open only at night when temperatures are generally lower (Fig. 1.1D). PEPc catalyses the fixation of CO_2, producing oxaloacetate. This is then reduced to form malate, and stored in the vacuole. Then, during the day when light is available for C_3 photosynthesis, malate is mobilised from the vacuole and CO_2 is released by the action of malate dehydrogenase. This mechanism of CO_2 concentration minimises water loss by transpiration, and CAM plants have water-use efficiency several fold higher than C_3 and C_4 plants, while C_4 plants in turn have water-use efficiency higher than C_3 plants.

In addition to the important role that it plays in carbon fixation in alternative photosynthetic pathways, PEPc is also involved in a number of other key processes in all plants. One such process is guard cell opening, where increases and decreases in the turgour of guard cells determine stomatal aperture. Guard cells exhibit a very different organelle content compared with mesophyll cells, in that they contain high numbers of mitochondria and low numbers of chloroplasts. Guard cell chloroplasts are smaller than mesophyll cell chloroplasts, with a lower chlorophyll content. In terms of enzyme content, guard cells have low Rubisco and high PEPc compared with mesophyll cells. In guard cells, stomatal aperture is regulated by ion transport into the cytoplasm.

Malate, the reduced form of oxaloacetate, acts as a counterion to balance potassium uptake.

1.2 Molecular properties of plant PEPc isoforms

The three dimensional structure of a C_4 isoform of PEPc from maize leaves was solved at 3.0 Å resolution by the group of Katsura Izui (Matsumura *et al.*, 2002). The overall quaternary structure of maize PEPc was found to be a "dimer of dimers" (Fig. 1.2). A monomer of PEPc contains an 8-stranded β-barrel structure and a high (57%) α-helical content. The tertiary structure of plant PEPc gives an insight into the function of conserved residues in the primary structure of this enzyme. For example, Gln493 and Arg498 are found at the interface between subunits, and form an ion pair holding the tetramer together. Indeed, mutating Arg498 to cysteine leads to partial dissociation of the tetramer into a dimer, presumably by disrupting this ion pair.

Common to all plant PEPc isoforms is the C-terminal tetrapeptide sequence Gln-Asn-Thr-Gly (QNTG). Site-specific mutagenesis of the glycine residue at the end of this peptide in sorghum C4-specific PEPc revealed its importance in catalysis (Xu *et al.*, 2006). While the specific activity of wild-type sorghum PEPc was found to be between 20-25 units mg^{-1}, mutation or deletion of the C-terminal glycine residue led to a substantial decrease in specific activity (~5 units mg^{-1} when glycine was deleted or mutated to alanine). Deletion of the entire C-terminal tetrapeptide led to elimination of PEPc activity. None of the mutations or deletions were found to affect the oligomerisation state of the enzyme.

Figure 1.2 The three-dimensional structure of maize PEPc

The crystal structure of maize PEPc, solved by Matsumura et al., (2002) shows that the overall quaternary structure of the protein is a "dimer of dimers."

Another key feature unique to the primary structure of plant PEPc is the presence of a consensus sequence that directs covalent modification of the N-terminus. Analysis of the three-dimensional structure of the maize isoform of PEPc has led to a proposed structural mechanism of how this modification regulates activity and leads to a reduced desensitisation to malate or aspirate inhibition (Kai *et al.*, 2003). The N-terminal region of PEPc contains a channel for aspartate binding, and when the serine residue in the consensus sequence is modified by phosphate, the negative charges and bulk of the phosphate group are proposed to block this channel. However, this does not explain why loss of the N-terminal region also leads to marked desensitisation to malate/aspartate inhibition.

This covalent modification will be discussed in considerably more detail later in this chapter.

1.3 The structure and function of bacterial PEPc

In contrast to plants, non-photosynthetic bacteria typically contain only one PEPc gene encoding a PEPc isoform that differs from PEPc from higher plants. Bacterial PEPc lacks the consensus sequence for covalent modification in the N-terminus of the PEPc polypeptide (Vidal & Chollet, 1997), thus bacterial PEPc appears to be regulated solely by allostery. The function of bacterial PEPc appears to be solely anaplerotic, providing dicarboxylic acids (oxaloacetate and malate) for the citric acid cycle. Similar to plant PEPc, bacterial PEPc has a signature tetrapeptide sequence at the C-terminus of the polypeptide. In contrast to plant PEPc, this sequence is Arg-Gln-Thr-Gly (RNTG).

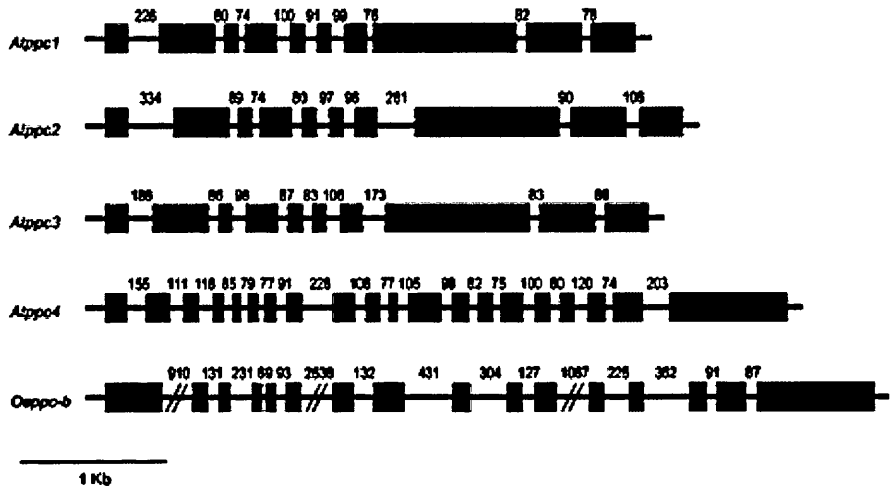
The three-dimensional structure of the *E. coli* PEPc enzyme has also been determined by X-ray crystallography (Kai *et al.*, 1999; Matsumura *et al.*, 1999; Matsumura *et al.*, 2002). Three enzyme-ligand structures were solved: a PEPc-Asp complex; a PEPc-Mn^{2+}-Asp complex; and a PEPc-Mn^{2+}-DCDP-Asp complex. DCDP, 3,3-dichloro-2-dihydroxyphosphinoylmethyl-2-propenoate, is a substrate analogue of PEP while Asp is a regulator of PEPc activity (see later in this chapter). Similar to maize C_4-PEPc, *E. coli* PEPc is a "dimer-of-dimers".

Analysis of the genome sequence of Arabidopsis led to the identification of a bacterial-type PEPc in plants (Sanchez & Cejudo, 2003). As stated earlier, Arabidopsis has four PEPc genes, termed *AtPpc1-4*. The predicted gene products of three of these isoforms (*AtPpc1-3*) were found to contain a serine phosphorylation motif in the N-terminus of their sequences, typical of plant PEPc isoforms. The predicted *AtPpc4* gene product, however, lacked this motif. The predicted product is much closer in sequence to the *E. coli* PEPc than are the predicted products of the other three genes, and the AtPpc4 polypeptide has the bacterial C-terminal amino acid sequence (RNTG) rather than the plant C-terminal amino acid sequence (QNTG). Also, the *AtPpc4* gene has an unusual intron structure compared with the other Arabidopsis PEPc genes (Fig. 1.3). This suggests *AtPpc4* has a completely different evolutionary origin than the other three PEPc genes in Arabidopsis. Rice was also found to contain a bacterial-type PEPc isoform (Sanchez and Cejudo, 2003).

A later study by the same group confirmed using antibodies raised against plant PEPc that the gene products of *AtPpc1-3* are immunologically distinct from the gene product of *AtPpc4* (Sanchez *et al.*, 2006). They also showed that, while AtPpc4 is expressed at low levels in Arabidopsis, high salinity and drought led to

Figure 1.3 Exon organisation of Arabidopsis PEPc genes

Arabidopsis thaliana has four PEPc genes, annotated *Atppc1-4* in this diagram. Three (*Atppc1-3*) contain ten exons, while the fourth (*Atppc4*) contains twenty exons, suggesting similarities to the bacterial PEPc genes and a different evolutionary origin. Black boxes represent exons; black lines represent introns; numbers above the black lines represent the number of base pairs between each exon. Data from Sanchez and Cejudo (2003).


Atppc1
226 80 74 100 91 99 76 82 78
Atppc2
334 89 74 89 97 96 261 90 106
Atppc3
186 86 98 87 83 106 173 83 88
Atppc4
155 111 118 85 79 77 91 228 108 77 105 98 82 75 100 80 120 74 203
Osppc-b
910 131 231 89 93 2538 132 431 304 127 1087 225 362 91 87
1 Kb

the induction of expression of this gene. One role of bacterial-type PEPc is therefore thought to be in plant adaptations to environmental stress.

The Plaxton group identified a role for bacterial-type PEPc in the endosperm of developing castor oil seeds (COS). When isolating PEPc from developing COS in the presence of malate and an inhibitor of thiol endopeptidase, they noted that two peaks of PEPc activity were resolved on size exclusion chromatography (Blonde & Plaxton, 2003). One peak corresponded to a molecular mass of 410 kDa, the expected mass of a PEPc homotetramer. However, the other peak was considerably larger, corresponding to a molecular mass of 681 kDa. Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of this larger complex showed that it consisted of two polypeptides, one with a mass of 107 kDa, the other with a mass of 64 kDa. The 64 kDa band did not react with PEPc antibody raised against *Brassica napus* PEPc and mass spectrometric sequencing showed that it most closely matched bacterial-type PEPc from Arabidopsis and rice. Further work later established that the 64 kDa polypeptide was actually a proteolytic fragment of full length bacterial-type PEPc, and that plant-type and bacterial-type form a hetero-octamer during COS development giving a 910 kDa complex (O'Leary *et al.*, 2009). They classified homotetrameric PEPc as "Class-1" PEPc and hetero-octameric PEPc as "Class-2" PEPc, due to the analogy with Class-1 and Class-2 PEPc isoforms in algae (see later in this chapter).

Kinetic characterisation of Class-2 PEPc revealed it to have different regulatory properties to Class-1 PEPc. Also, thermal stability studies of showed that Class-2 PEPc is far less heat-labile: when incubated at 50°C for 3 min, Class-2 PEPc lost only 20% of its activity, while Class-1 PEPc lost 82% of its activity

under the same conditions (Blonde & Plaxton, 2003). Non-denaturing polyacrylamide gel electrophoresis was used in the same study to show that ratio of Class-1 PEPc:Class-2 PEPc increases during COS development, until only Class-1 PEPc can be detected. Thus, Class-2 PEPc may be playing a role in development, and becomes redundant as development progresses.

1.4 Algal PEPc

The original classification of PEPc into Class-1 and Class-2 isoforms was made in studies of algal PEPc enzymes. Four active isoforms of PEPc were purified from *Selenastrum minutum*, a green alga (Rivoal *et al.*, 1996). One isoform, PEPc1, was found to be a homotetramer of a 102 kDa polypeptide, similar to higher plant PEPc. The other three, PEPc2, PEPc3 and PEPc4, were found to be hetero-oligomers, and had molecular masses of 984, 1186, and 1590 kDa respectively. SDS-PAGE analysis of all four isoforms showed that they all contained the 102 kDa polypeptide with varying amounts of lower molecular mass polypeptides. The increase in molecular mass of the PEPc complex correlated with a decrease in the relative amounts of 102 kDa polypeptide. A decrease in reactivity with anti-PEPc1 antibody also correlated with the increase in molecular mass. As was found for plant Class-2 PEPc, hetero-oligomeric PEPc from *S. minutum* is less heat labile than homotetrameric PEPc. PEPc1 was classified as "Class-1" PEPc, while PEPc2, PEPc3 and PEPc4 were classified as "Class-2" PEPc. In the same paper, the kinetic properties of Class-1 and Class-2 PEPc isoforms from *S. minutum* were reported (Rivoal *et al.*, 1996).

Later studies by the Plaxton group determined the polypeptide composition of *S. minutum* Class-2 PEPc isoforms and showed that algae contain

a unique PEPc polypeptide distinct from bacteria and higher plants (Rivoal *et al.*, 2001; Rivoal *et al.*, 1998). Purification of *S. minutum* Class-2 PEPc in the presence of an array of protease inhibitors and subsequent SDS-PAGE/immunoblot analysis showed the molecular masses of the polypeptides to be 130 kDa, 102 kDa, 73 kDa and 65 kDa. Stoichiometric analysis confirmed the earlier finding that the representation of the 102 kDa polypeptide decreases as the molecular mass of the PEPc complex increases.

1.5 PEPc in marine diatoms

An interesting finding in a major group of algae led to the conclusion that C_4 photosynthesis can also take place in marine diatoms (Reinfelder *et al.*, 2000). Marine diatoms fix more than ten billion tons of inorganic carbon per annum despite existing in an environment where CO_2 is limited. Thus, marine diatoms must make use of carbon concentrating mechanisms to maximise the CO_2:O_2 ratio around Rubisco. There is considerable evidence that the C4 photosynthetic pathway is the major mechanism of concentrating carbon. The evolution of C_4 photosynthesis in unicellular organisms has been proposed to pre-date evolution of multi-cellular C_4 photosynthesis (Raven *et al.*, 2008).

1.6 Catalysis by PEPc

The true substrate for PEPc is bicarbonate anion, rather than the more chemically reactive CO_2. The reaction catalysed by PEPc is highly exergonic, and the ΔG° of the reaction is -30 kJ mol^{-1} making it essentially irreversible. The catalytic mechanism has been proposed by studies of the hydrolysis of PEP

analogues and high resolution structural studies of the bacterial and C4-isoforms of the enzyme (Fig. 1.4). First, the reactants bind in the preferred order: divalent cation, PEP, and then bicarbonate. Then a carboxyphosphate/enolate intermediate is formed, stabilised by a histidine residue in the hydrophobic pocket of the active site. This histidine residue extracts a proton from carboxyphosphate. The third step of the reaction is the transfer of CO_2 to carbon-3 of the enolate intermediate by nucleophilic attack. This liberates oxaloacetate and orthophosphate, which is protonated by the histidine residue in the hydrophobic pocket.

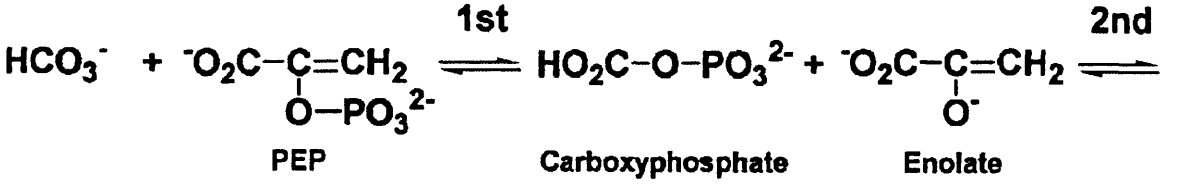
1.7 Regulation of PEPc by allostery

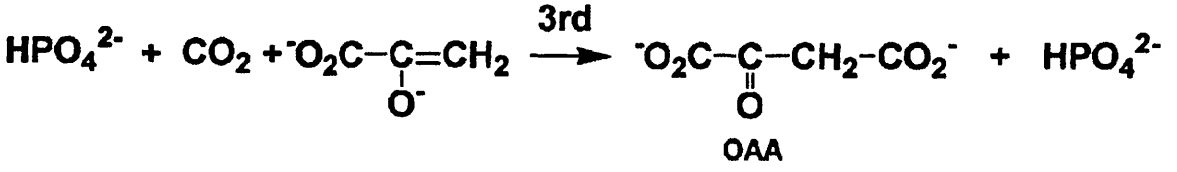
Given that PEPc has a range of vital roles and functions, it follows that the activity of the enzyme is tightly controlled. A common feature of all PEPc isoforms is that they are regulated by an allosteric mechanism. Allosteric regulation involves binding of a small molecule to an enzyme at a site distinct from the site of catalysis. This binding causes a conformational change in the enzyme which results in either inhibition or activation of activity.

The classic allosteric inhibitor of higher plant PEPc is L-malate. This compound is formed from the reduction of oxaloacetate by NAD- or NADP-malate

Figure 1.4 The catalytic mechanism of PEPc

Phosphoenolpyruvate (PEP) and bicarbonate ion bind to the active site of PEPc in the presence of divalent cation. The reaction proceeds via carboxyphosphate and enolate intermediates. The first two steps of the reaction are reversible, while the third is irreversible. Oxaloacetate (OAA) and inorganic phosphate are the products of the final reaction step (Kai *et al.*, 2003).


1st
2nd
PEP
Carboxyphosphate
Enolate


3rd
OAA

dehydrogenase. In this way, the activity of PEPc can be regulated by feedback inhibition. Another key inhibitor of higher plant PEPc is the amino acid aspartate. Glucose 6-phosphate is an activator of higher plant PEPc. Glycine has also been shown to activate the isoform of PEPc involved in photosynthetic CO_2 fixation in C_4 plants.

Allosteric regulation of bacterial and algal PEPc enzymes is considerably more complex than in higher plants. A range of allosteric activators of the *E. coli* enzyme have been identified, such as acetyl-CoA, long chain fatty acids and fructose 1,6-bisphosphate.

Comparisons of the *E. coli* and maize PEPc structures have revealed the molecular basis of the allosteric control of PEPc activity (Kai *et al.*, 1999). The structure of the *E. coli* enzyme was solved in complex with aspartate, while the maize enzyme was crystallised with a sulphate ion in the binding site for glucose 6-phosphate. When the two structures are compared, both global and local structural changes that occur upon activation and inhibition by allosteric effectors can be observed. For example, the essential arginine residue that is part of the GRGGXXG**R**GG motif (R647 in maize, R587 in *E. coli*) moves 15 Å towards the catalytic site in the active state. An essential histidine residue (H177 in maize, H138 in *E. coli*), that is thought to stabilise the carboxyphosphate intermediate in the catalytic reaction, was found to move 10 Å towards the active site in the active state. Thus it can be seen that particular regions of the polypeptide undergo large conformational changes upon binding of allosteric effectors. In fact, comparison of the basic dimeric units of the maize and *E. coli* enzymes show a 10° rotation when superimposed upon each other.

Despite intense study on the control of PEPc activity, the allosteric model could not explain all of the features of PEPc regulation seen in higher plants. It had been observed that the enzyme from the CAM plant *Kalanchoë fedtschenkoi* displayed a different affinity for malate during the day compared with at night (Nimmo *et al.*, 1986). This was shown to be as a result of covalent modification of the enzyme, leading to an increase in the affinity to glucose 6-phosphate and a decrease in the affinity to malate. *In vivo* labelling studies using $^{32}P_i$ showed this covalent modification to be phosphorylation. This work led to intensive investigation of the phosphorylation of PEPc in several plant systems.

1.8 The discovery and identification of PEPc kinase

Further experiments by several groups showed the existence of a specific PEPc kinase (Nimmo, 2000). However, PEPc kinase is present in very low abundance *in vivo*, precluding the direct derivation of amino acid sequence.

The Nimmo group successfully cloned PEPc kinase from *Kalanchoë fedtschenkoi* using an *in vitro* transcription/translation approach (Hartwell *et al.*, 1999). A cDNA library was constructed from *K. fedtschenkoi* leaves and the library was screened for the ability of the translation products to phosphorylate PEPc. By successive sub-divisions of this library, the cDNA for PEPc kinase was isolated. When the cDNA was expressed in *E. coli* the gene product was found to be almost exclusively in inclusion bodies. Nevertheless, this was the first reported cloning of a PEPc kinase gene and allowed identification of PEPc kinase genes in many other plants. Detailed analysis of PEPc kinase genes showed that all plants studied to date contain a small gene family (Nimmo, 2003; Echevarria and Vidal, 2003) and that the major means of controlling PEPc kinase activity is

expression of PEPc kinase genes in response to a range of signals in different tissue types. For example, in CAM plants the expression of PEPc kinase genes is controlled by the central circadian oscillator (Nimmo, 2003).

PEPc kinase is the smallest known eukaryotic protein kinase identified thus far (Hartwell *et al.*, 1999). The protein consists of solely a catalytic serine/threonine protein kinase domain, with minimal N- and C-terminal extensions. Primary structure analysis of the protein has shown that it is most closely related to the calcium/calmodulin-dependent protein kinases. However, unlike these proteins, PEPc kinase does not have EF-hands, is not calcium dependent and does not have the autoinhibitory mechanism.

1.9 Control of protein function by phosphorylation

Protein phosphorylation is the transfer of the γ-phosphate group of ATP to the hydroxyl group of a serine, threonine or tyrosine residue on the target protein. This reaction is catalysed by a group of enzymes called protein kinases. These enzymes generally exist in two classes: serine/threonine protein kinases, and tyrosine protein kinases. It should be noted that there are some dual specificity protein kinases, capable of phosphorylating all three amino acids (Mercer & Friedman, 2006). Histidine kinases are also known to exist, though these are rare in eukaryotes (Kowluru, 2008).

Common to the serine/threonine and tyrosine protein kinases is the protein kinase domain, a stretch of amino acid residues that is approximately 250 amino acids long. Genome sequencing has shown that 2% of all eukaryotic genes contain a protein kinase domain (Gomase & Tagore, 2008). Also, it is

estimated that 30% of proteins in a eukaryotic cell are phosphorylated at any given time.

Protein phosphorylation can have a number of marked effects on the function of proteins. For example, addition of a phosphate group to a serine, threonine or tyrosine residue can create new protein-protein interaction sites, modulate sub-cellular localisation, and modulate activity.

Phosphorylation of proteins is not a permanent modification, and can be reversed by the activity of protein phosphatases. Just like protein kinases, there are two classes of protein phosphatase: serine/threonine protein phosphatases and tyrosine protein phosphatases. An in-depth review of protein phosphatases is beyond the scope of this thesis (but see Luan, 2003 for further details).

Protein kinases exist in two functional states: a catalytically-active "on" state, and an inactive "off" state. Since the reaction that they catalyse is the same, the structures of the "on" states of protein kinases are remarkably similar. However, as there are less catalytic constraints on the "off" states, a number of different mechanisms for controlling protein kinase activity have evolved.

Some protein kinases contain a threonine residue in their catalytic loop, which can be phosphorylated to stimulate kinase activity. Other protein kinases contain a negatively charged amino acid residue in place of the threonine residue, mimicking the effects of threonine phosphorylation.

Protein-protein interactions can also regulate protein kinase activity. The paradigm for this method of regulation is protein kinase A (PKA). In the "off" state, PKA exists as a heterotetramer of two catalytic subunits and two regulatory subunits. The regulatory subunits act to block the active site of the

catalytic subunits, rendering the protein kinase inactive. Upon elevation of cellular cyclic adenosine 3',5'-monophosphate (cAMP) levels, the regulatory subunits dissociate from the catalytic subunits.

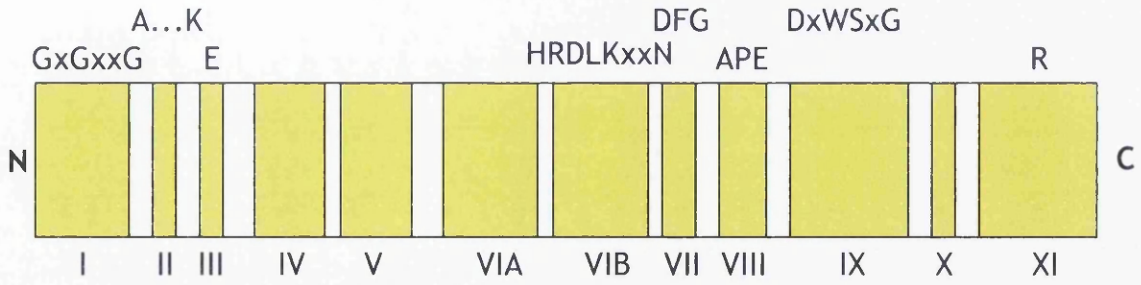
1.10 The structure of protein kinase domains

The primary structure of protein kinase domains is divided into 12 subdomains (Fig. 1.5). These 12 subdomains contain a number of absolutely conserved amino acid residues essential for protein kinase function (Hanks & Hunter, 1995). For example, subdomain I contains a triad of glycine residues; subdomain VIB, meanwhile, contains the asparagine and aspartate residues involved in catalysis.

The first high-resolution three-dimensional structure of a protein kinase was obtained in 1991, when the crystal structure of the catalytic subunit of PKA was reported (Knighton *et al.* 1991a; Knighton *et al.* 1991b). This structure revealed the importance of the strictly conserved residues in catalysis and substrate binding. The three-dimensional structure of the protein kinase domain of PKA can be viewed as bi-lobal. The N-terminal lobe comprises of subdomains I-IV, comprises of mostly β-sheet and is responsible for nucleotide binding and orientation. The C-terminal lobe is made up of subdomains VIA-XI, is predominantly α-helical, and is involved in substrate binding and initiation of catalysis. The conserved glycine triad in subdomain I helps the polypeptide to form a β-sheet-turn-β-sheet structure, that acts as a flexible clamp over the β- and γ-phosphates of ATP. The conserved asparagine and aspartate residues are in the active site of the enzyme and participate in catalysis.

Figure 1.5 Schematic representation of a serine/threonine protein kinase domain

A schematic diagram representing the polypeptide chain of a serine/threonine protein kinase domain, not shown to scale. Conserved amino acids are shown at the top of the diagram, and domain numbering is shown at the bottom. Yellow boxes = domains; N = N-terminus; C = C-terminus. Adapted from Hanks and Hunter (1995).


A...K
GxGxxG
E
HRDLKxxN
DFG
APE
DxWSxG
R
N
C
I
II
III
IV
V
VIA
VIB
VII
VIII
IX
X
XI

1.11 Control of PEPc kinase activity

PEPc kinase is unique in that it is not regulated by any of the mechanisms mentioned above. In fact, for a considerable amount of time, it was thought that the enzyme was regulated solely by its synthesis and degradation (Nimmo, 2003). However, this has now been shown not to be the case, and several medium-term regulatory mechanisms for PEPc kinase have since been discovered.

PEPc kinase is thought to be sensitive to changes in intracellular pH as the enzyme is active over a very narrow pH range. For example, the activity of partially purified maize PEPc kinase was found to be 80% lower at pH 7 than at pH 8 (Wang & Chollet, 1993).

There is also *in vitro* evidence that the activity of PEPc kinase can be regulated by allosteric effectors of PEPc (Wang & Chollet, 1993). When assayed with increasing concentrations of malate, phosphorylation of PEPc by PEPc kinase decreased. This effect was shown to be pH-dependent as when the pH of the assay was lowered, inhibition by malate was more pronounced. However, PKA phosphorylation of PEPc was also inhibited by malate, suggesting that it is the conformation of protein substrate rather than PEPc kinase that is being affected.

The group of Katsura Izui purified maize PEPc kinase to homogeneity and observed that activity could be lost and recovered by changing buffer conditions (Saze *et al.*, 2001). Starting with 2.6 kg of maize leaves, they purified PEPc kinase using a range of chromatography steps, including hydrophobic interaction, size exclusion and ion exchange chromatography, which resulted in less than 1 μg of pure protein. They discovered that omission of reducing agents from the

purification buffers led to a loss of PEPc kinase activity; activity could be restored by re-addition of thioredoxin or dithiothreitol (DTT). This led to the proposal that the redox state of PEPc kinase is important in regulating activity of the protein.

A fourth medium-term regulatory mechanism has been identified in the leaves of CAM and C_4 plants. It was noted that the specific activity of PEPc kinase in desalted extracts from *Kalanchoë fedtschenkoi* and maize leaves increased upon dilution, while the specific activity of purified PEPc kinase remained constant (Nimmo *et al.*, 2001). This led them to conclude that plant leaves contain a specific inhibitor of PEPc kinase. Treatment with a protease and heat denaturation demonstrated that the inhibitor is a protein. It was also shown that the protein is not a phosphatase, an ATPase or a protease itself. The analogy was made with the regulatory subunits of protein kinase A, in that the inhibitor masks the basal activity of PEPc kinase and upon an activation signal this suppression is lifted, perhaps involving proteolytic degradation of the inhibitor as well as increased expression of the kinase. The inhibitor and PEPc kinase could be separated by hydrophobic interaction chromatography on Phenyl Sepharose and the inhibitor was partially purified; however, the inhibitor-PEPc kinase complex could not be isolated. The inhibitor appeared to be highly specific to PEPc kinase as incubation with casein kinase I or histone kinase had no effect on the activities of these proteins. Further characterisation of this inhibitor is awaited.

1.12 Analysis of PEPc kinase genes

Just as they contain a small family of PEPc genes, nearly all higher plants contain a small family of PEPc kinase (*PPCK*) genes. There is less tissue specificity in the expression patterns of PEPc kinase. Induction of the expression of specific PEPc kinase isoforms by environmental stressors can also occur. One of the best characterised families of *PPCK* genes is in soybean (Sullivan *et al.*, 2004; Xu *et al.*, 2007).

PPCK genes from a number of other organisms have been analysed and characterised (Nimmo, 2003). Generally, *PPCK* genes contain a single intron in the 3' end of the coding region, but there are exceptions to this. For example, tomato *PPCK2* (*LePPCK2*) contains a second intron that exhibits alternative splicing (Marsh *et al.*, 2003). *LePPCK2* can give rise to three transcripts: correctly spliced, incorrectly spliced and unspliced. The first of these transcripts is functional, while the other two are not.

The structure of rice *PPCK2* (*OsPPCK2*) is also remarkable. The *OsPPCK2* gene has three in-frame "start" codons (ATG codons), two of which act as translation initiation sites (Fukayama *et al.*, 2006). *OsPPCK2* also lacks a "stop" codon in the intron, so if the intron is not spliced out a longer transcript is obtained. Thus, *OsPPCK2* can give rise to four different translation products and *in vitro* transcription/translation showed that all of these are active except the product that contains the expressed intron.

One of the Sorghum PEPc kinase isoforms contains an insert of acid amino acids relative to other PEPc kinase polypeptides (Nimmo, 2003). This insert modifies the overall hydrophobicity of the polypeptide considerably. However,

molecular modelling of this isoform shows that the acidic insert is not part of the active site of the enzyme.

1.13 Questions arising from the literature

This chapter has summarised the current literature on PEPc, its regulation by PEPc kinase, and the control of the activity of PEPc kinase. Evaluation of this chapter raises a number of specific questions regarding PEPc kinase:

(1) How does a minimal protein kinase domain fold?

While the structures of the catalytic domains of serine/threonine protein kinases are known, they have been determined as part of a larger polypeptide chain (Johnson and Lewis, 2001). It is possible that a minimal protein kinase domain has a different "off" state than that of a larger protein kinase polypeptide. By solving the three-dimensional structure of PEPc kinase alone, the basic structural unit of a protein kinase will be known.

(2) Why does PEPc not phosphorylate peptide substrates?

Peptide substrates based on PEPc are very poor substrates for PEPc kinase *in vitro* (Li *et al.*, 1997). This is in contrast to many other protein kinases, which can often phosphorylate very efficiently peptides corresponding to the sequence round the phosphorylation site in their targets. This suggests the entire PEPc polypeptide must be present to allow efficient phosphorylation by PEPc kinase. Determination of the three dimensional structure of PEPc kinase alone and in complex with PEPc may show why this is the case.

(3) Do specific PEPc-PEPc kinase pairs exist?

PEPc and PEPc kinase are both encoded by small gene families in higher plants. Specific isoforms are expressed in specific tissues or induced by particular environmental factors - for example, the C_4-isoform of PEPc is highly expressed in leaves, where it can carry out carbon fixation. By determining the Michaelis constant (K_M) of particular PEPc kinases for particular PEPc enzymes, a measure of the enzyme-substrate interaction can be made.

(4) What is the function of the acidic insert in Sorghum PPCK1?

Analysis of PPCK genes have shown that one of the Sorghum *PPCK* genes encode a PEPc kinase protein with an insert of acidic amino acid residues relative to other PEPc kinases (Nimmo, 2003). While molecular modelling studies have shown that this is unlikely to be involved in catalysis, the role of this insert is still unclear. Determination of kinetic constants of the interaction of Sorghum PEPc kinase with PEPc may give a clue to the function of the insert.

The aim of this project is to investigate the questions posed above, and related issues concerning the structure and function of PEPc kinase.

2 Materials and Methods

2.1 Materials

All chemicals used were from VWR (Poole, Dorset, UK), Sigma-Aldrich (Poole, Dorset, UK) or Fisher Scientific (Southampton, UK) unless otherwise stated.

Restriction enzymes, DNA and RNA size markers, and dNTPs were from Promega UK (Southampton, UK).

[γ-^{32}P]-ATP (3000 Ci mmol^{-1}) and [^{35}S]-L-methionine (1100 Ci mmol^{-1}) were obtained from Perkin-Elmer.

Rabbit reticulocyte lysate *in vitro* translation kit and most chromatography and protein purification materials were obtained from GE Healthcare (Amersham, Bucks, UK).

Zeocin and DNase I was obtained from Invitrogen (Paisley, UK).

Hydroxyapatite (Bio-gel HTP) and 37.5:1 acrylamide:bis-acrylamide solution were from Bio-Rad (Bramley, Kent, UK).

mMessage T3 *in vitro* transcription kit was from Ambion (Huntingdon, UK).

Okadaic acid was from Moana Bioproducts (Honolulu, USA).

Ni-NTA agarose was obtained from Qiagen (Crawley, West Sussex, UK).

2.2 Plant, yeast and bacteria

2.2.1 Kalanchoë fedtschenkoi

Kalanchoë fedtschenkoi Hamet et Perrier plants were maintained and harvested as outlined by Nimmo *et al.* (1984). Leaves were harvested approximately 2 h after dawn.

2.2.2 Zea mays

Maize (*Zea mays* L. var. Bastille) plants were grown in a field. Mature plants were harvested under full sunlight in August/September. Leaves were stored at -80°C prior to purification of PEPc.

2.2.3 Pichia pastoris

Pichia pastoris cells expressing PEPc kinase constructs were previously constructed by Dr. Justin Marsh or Dr. Gillian Nimmo (e.g. Chen *et al.*, 2008). Colonies were grown from frozen glycerol stocks on solid YPD agar + 0.2 mg ml^{-1} Zeocin. Biomass increase was achieved by growing in BMGY medium (2% (w/v) peptone, 1% (w/v) yeast extract, 100 mM phosphate buffer pH 6.0, 1.34% yeast nitrogen base, 1% (v/v) glycerol). Expression of heterologous protein was achieved by growing in BMMY medium (2% (w/v) peptone, 1% (w/v) yeast extract, 100 mM phosphate buffer pH 6.0, 1.34% yeast nitrogen base, 1% (v/v) methanol) for the number of days specified in the text.

2.2.4 *Escherichia coli*

Escherichia coli cells were maintained in LB medium (1% (w/v) tryptone, 1% (w/v) NaCl, 0.5% (w/v) yeast extract) plus the appropriate antibiotic. Recombinant protein expression was carried out using BL21 (DE3) cells (Stratagene).

2.3 Protein expression and purification methods

2.3.1 *Expression of recombinant PEPc kinase*

Frozen glycerol stock of *Pichia pastoris* transfected with the GmPPCK3 cDNA was plated out on YPD agar + 0.2 mg ml^{-1} zeocin, and grown for 4 days at 28°C. A single colony of cells was used to inoculate a starter culture of 10 ml of BMGY + 0.2 mg ml^{-1} zeocin, which was grown overnight at 28°C with shaking at 200 rpm. Five 1 litre flasks, each containing 200 ml BMGY, were each inoculated with 1 ml starter culture, and grown overnight at 28°C, shaking at 200 rpm with two layers of muslin over the necks of the flasks. Each 200 ml aliquot of BMGY was then split into two, centrifuged at 2,500 x g, and each cell pellet resuspended in 400 ml BMMY in 2 litre flasks. The resulting ten flasks were returned to the incubator, and cells were grown for four days at 28°C with shaking at 200 rpm. Muslin was placed over the necks of the flasks to allow adequate aeration of cultures, and 1/200 volume of methanol was added every 24 h.

2.3.2 Purification of recombinant PEPc kinase

After expression of recombinant PEPc kinase, *Pichia pastoris* cells were pelleted by centrifugation at 4000 x g for 30 min at 4°C. The volume of supernatant was measured, and stock solutions were added to give final concentrations of 1 mM DTT, 2 mM EDTA, 1 mM benzamidine/HCl, 10 $\mu g\ ml^{-1}$ chymostatin, 10 $\mu g\ ml^{-1}$ leupeptin, 10 $\mu g\ ml^{-1}$ antipain and 0.1 mM PMSF. The pH of the supernatant was adjusted to 7.4 and it was loaded onto a 200 ml column of phenyl sepharose (GE Healthcare) pre-equilibrated with buffer A (50 mM Tris/HCl, 1 mM DTT, 1 mM benzamidine/HCl, 1 mM EDTA, pH 7.4).

The column was washed with 500 ml buffer A, then 5 litres of buffer B (50 mM Tris/HCl, 20 mM β-mercaptoethanol, 1 mM benzamidine/HCl, 30% (v/v) ethylene glycol, pH 7.4). The A_{280} of the effluent was monitored using buffer B as the reference, and once the A_{280} had fallen to 0.01, 300 ml of buffer C (50 mM Tris/HCl, 20 mM β-mercaptoethanol, 1 mM benzamidine/HCl, 60% (v/v) ethylene glycol, pH 7.4) was used to elute bound protein.

The peak of A_{280} absorbance was pooled and loaded onto a Ni^{2+}-NTA agarose column (Qiagen) pre-equilibrated with buffer D (50 mM Tris/HCl, 0.5 M NaCl, 10 mM imidazole, 1 mM benzamidine/HCl, 0.1% (v/v) Tween-20, pH 7.4). The column was washed with buffer E (50 mM Tris/HCl, 0.5 M NaCl, 70 mM imidazole, 1 mM benzamidine/HCl, 0.1% (v/v) Tween-20, pH 7.4) and the A_{280} of the effluent was monitored using buffer E as the reference. Once the A_{280} of the effluent had fallen close to that of buffer E, the bound protein was eluted using buffer F (50 mM Tris/HCl, 0.5 M NaCl, 500 mM imidazole, 1 mM benzamidine/HCl, 0.1% (v/v) Tween-20, pH 7.4). The peak of A_{280} absorbance was pooled and concentrated in a centrifugal concentrator (Sartorius). Glycerol

was added to a final concentration of 50% (v/v) and purified PEPc kinase was stored at -20°C until required.

2.3.3 Expression of recombinant OsPpc4

E. coli BL21(DE3) cells were transformed with plasmid containing the OsPpc4 cDNA (Matsumoto *et al.*, 2007) according to standard techniques and plated out on LB agar plus 50 $\mu g\ ml^{-1}$ kanamycin. A single colony was used to inoculate a starter culture of LB medium plus 50 $\mu g\ ml^{-1}$ kanamycin, which was grown overnight at 37°C with shaking at 200 rpm. The starter culture was diluted 1:50 with fresh LB medium plus 50 $\mu g\ ml^{-1}$ kanamycin and grown to an OD_{600} = 0.5-0.7 at 37°C with shaking at 200 rpm. Expression of OsPpc4 was induced by addition of IPTG to a final concentration of 1 mM, and the culture was incubated at 37°C for a further three hours. Cells were harvested by centrifugation at 4,000 x g for 20 min at 4°C, and stored at -80°C until required. Typically, cells were grown in batches of 500 ml in two litre shake flasks.

2.3.4 Purification of recombinant OsPpc4

One batch of cells, corresponding to five litres of cell culture, was thawed on ice for 15 min and resuspended in 15 ml binding buffer (50 mM Tris/HCl pH 7.5, 1 mM DTT, 1 mM benzamidine, 10 mM imidazole, 300 mM NaCl) plus one Complete Mini protease inhibitor tablet (Roche) and DNase I (Invitrogen). Cells were lysed by four passages through a French pressure cell, and the insoluble material was removed by centrifugation at 17,000 x g for 30 min at 4°C.

The supernatant was incubated with 10 ml Ni-NTA resin (20 ml of a 50% slurry) for 1 hour at 4°C on a rotary shaker. The resin was then poured into a column and washed with binding buffer until the A280 of the effluent was <0.01. Protein was then eluted with elution buffer (50 mM Tris/HCl pH 7.5, 1 mM DTT, 1 mM benzamidine, 500 mM imidazole, 300 mM NaCl).

The elution fractions were desalted on a 10 ml Sephadex G25-M column and then loaded onto a Mono Q 5/5 column (GE Healthcare) at a flowrate of 1 ml min^{-1}. PEPc protein was eluted using a 0 - 400 mM KCl gradient over 20 column volumes.

2.3.5 Isolation of PEPc from Kalanchoë fedtschenkoi

Dephospho-PEPc was purified from *K. fedtschenkoi* leaves according to the method of Nimmo *et al.* (1986). All steps were carried out at 4°C, with the exception of the Superose 6 step, which was carried out at room temperature using a BioCAD 700E system.

2.3.6 Isolation of PEPc from Zea mays

Phospho-PEPc was purified from *Z. mays* leaves according to the method of McNaughton *et al.* (1989). All steps were carried out at 4°C, with the exception of the Mono Q step, which was carried out at room temperature using a BioCAD 700E system. 200 µg Phospho-PEPc was treated with 0.02 units of alkaline phosphatase (Stratagene) for 1 h and then re-chromatographed on Mono Q prior to use in enzyme assays.

2.4 *In vitro* transcription/translation of OsPPCKs

2.4.1 In vitro transcription of template DNA

Plasmids containing OsPPCK genes were provided by Dr Stuart Sullivan (Fukayama *et al.*, 2006) and 5 μg of plasmid DNA was linearised by an overnight digestion with EcoRI. The reaction was initiated by addition of EcoRI and reaction mixtures were incubated in a waterbath at 37°C overnight. After 16 h, 12 extra units of EcoRI was added to each reaction, and incubated at 37°C for a further 3 h.

Reactions were terminated by addition of 1/20 volume 0.5 M EDTA, then DNA was precipitated by adding 1/10 volume of 7.5 M ammonium acetate and 2 volumes of 100% ethanol. These mixtures were incubated at -20°C overnight, then centrifuged at 16,200 x g in a benchtop centrifuge for 15 min. The supernatant was removed, and the pellet resuspended in 10 μl dH_2O and stored at -20°C.

Linearised plasmid DNA containing OsPPCK cDNA was used for synthesis of capped messenger RNA (mRNA) using the mMessage mMachine T3 kit according to the manufacturer's protocol. Reaction mixtures were incubated at 37°C in a waterbath for 2 h, after which template DNA was removed by treatment with 1 μl TURBO DNase for 15 min at 37°C.

RNA was recovered by phenol:chloroform extraction and isopropanol precipitation. 115 μl nuclease-free water and 15 μl ammonium acetate stop solution were added to reaction mixtures and mixed thoroughly. An equal volume of phenol/chloroform and then an equal volume of chloroform was added, mixed and then centrifuged at full speed in a benchtop centrifuge for 5

min. The aqueous phase was recovered and transferred to a clean tube. One volume of 100% isopropanol was added and mixed well. The mixture was incubated at -20°C for 1 h, and then centrifuged at full speed in a benchtop centrifuge at 4°C for 15 min. Isopropanol was removed, pellet was air-dried for 15 min at room temperature and resuspended in 30 μl dH_2O. RNA was quantified on a spectrophotometer and checked for purity and integrity by denaturing agarose gel electrophoresis.

2.4.2 In vitro translation of mRNA

The mRNA obtained from the *in vitro* transcription reactions was translated using the Rabbit Reticulocyte Lysate System (GE Healthcare) according to the manufacturer's instructions. Protein synthesis was monitored using [^{35}S]-methionine and analysis by SDS-PAGE.

2.4.3 Assay of translated OsPPCK mRNA activity

The products of the *in vitro* translation reaction were assayed for the ability to phosphorylate OsPpc4 protein according to the method of Hartwell *et al.* (1996).

2.5 Biophysical analysis of PEPc kinase

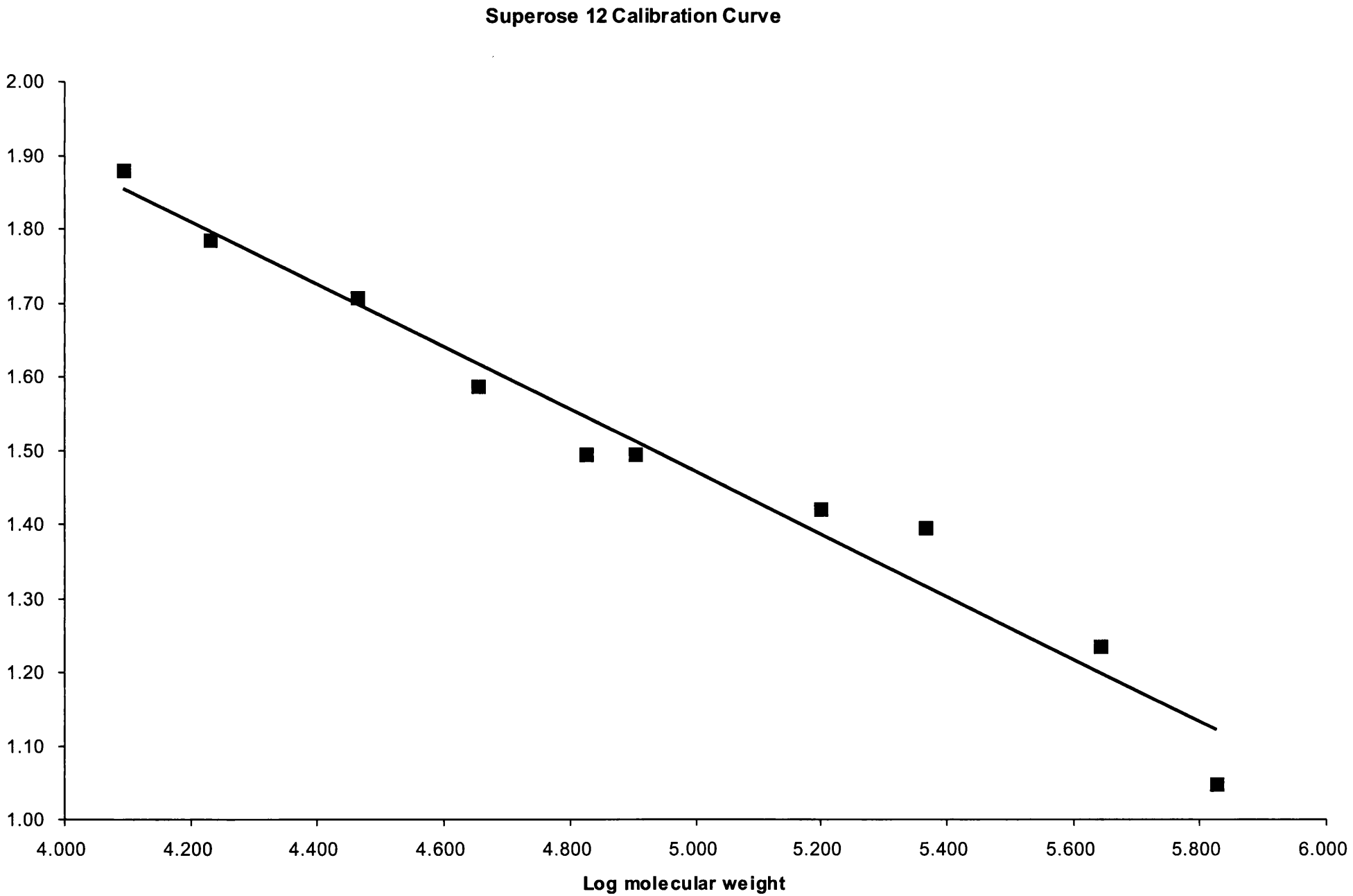
2.5.1 Analytical size exclusion chromatography

Analytical size exclusion chromatography of recombinant PEPc kinase was carried out using a Superose 12 column attached to a BioCAD 700E system. 100

μl of protein solution was applied to an equilibrated Superose 12 column at a flowrate of 0.3 ml min^{-1} and 1 min fractions collected. The absorbance of the effluent was measured at 280 nm. The column was calibrated by measuring the elution volume of molecular weight standards (V_e) relative to the elution volume of blue dextran (V_o) and plotting this against the logarithm of molecular weight (Fig. 2.1).

Figure 2.1 Superose 12 calibration curve

Proteins of known molecular weight were applied to a Superose 12 column as outlined in the text. The elution volumes of standard proteins (V_e) were related to the void volume of the column (V_o, the elution volume of blue dextran) and plotted against $\log_{10}$ molecular weight. Molecular weight markers used were: thyroglobulin (669 kDa), apoferritin (440 kDa), catalase (232 kDa), aldolase (158 kDa), transferring (80 kDa), bovine serum albumin (67 kDa), ovalbumin (45 kDa), carbonic anhydrase (29 kDa), myoglobin (17 kDa) and cytochrome c (12.4 kDa).


Superose 12 Calibration Curve
2.00
1.90
1.80
1.70
1.60
1.50
1.40
1.30
1.20
1.10
1.00
4.000
4.200
4.400
4.600
4.800
5.000
5.200
5.400
5.600
5.800
6.000
Log molecular weight

2.5.2 Sedimentation velocity analysis of PEPc kinase

Sedimentation velocity ultracentrifugation was carried out using a Beckman XL-1 ultracentrifuge at 49,000 rpm. Three concentrations of PEPc kinase were used (0.1 mg ml^{-1}, 0.2 mg ml^{-1} and 0.3 mg ml^{-1}). The progress of sedimentation was analysed using interference optics and results were analysed using SEDFIT.

2.5.3 Sedimentation equilibrium analysis of PEPc kinase

Sedimentation equilibrium ultracentrifugation was carried out using a Beckman XL-1 ultracentrifuge at 15,000 rpm. Three concentrations of PEPc kinase were used (0.1 mg ml^{-1}, 0.2 mg ml^{-1} and 0.3 mg ml^{-1}). The progress of sedimentation was analysed using interference optics. Results were analysed using SEDPHAT.

2.6 Structural techniques

2.6.1 Crystallisation of PEPc kinase

Recombinant soybean PEPc kinase was expressed and purified as outlined in section 2.3. Protein was buffer exchanged using an ultrafiltration centrifugal concentrator and concentrated to 4 mg ml^{-1} prior to addition to crystal screens. See Appendix 1 for details of crystallisation buffers.

2.6.2 CD spectroscopy

Recombinant soybean PEPc kinase was purified as outlined in section 2.3. Protein was buffer exchanged using an ultrafiltration centrifugal concentrator. The final concentration soybean PEPc kinase was measured using a spectrophotometer. CD spectra were recorded using a Jasco V-800 spectropolarimeter and analysed using CONTIN.

2.7 Protein assays

2.7.1 Assays of recombinant PEPc kinase activity

Recombinant PEPc kinase was assayed using γ-labelled ATP in a discontinuous radiometric kinase assay. A typical assay contained (in a volume of 25 μl) 100 mM Tris/HCl pH 8.0, 0.5% (w/v) PEG 20,000, 0.1 mM EDTA, 1 mM DTT, 1 mM benzamidine, 10 μg μl^{-1} chymostatin, 10 μg μl^{-1} leupeptin, 10 μg μl^{-1} antipain, 10 μM ATP, 2.5 mM $MgCl_2$, 1 μCi [γ-^{32}P] ATP, 2 μg dephospho-PEPc and diluted fraction of PEPc kinase. The assay was initiated either by addition of PEPc kinase or ATP, depending on the variable. Assays were incubated at 30°C for 5-15 min and terminated by addition of 20 μl 2X SDS-PAGE sample loading buffer and heating for 3 min at 95°C. Unincorporated ATP was removed by electrophoresis on an 8% SDS-PAGE gel, and the incorporation of radioactivity was measured by phosphorimaging. Unless stated otherwise, the PEPc substrate was isolated from *Kalanchoë fedtschenkoi*.

2.7.2 Assay of PEPc activity

PEPc activity was assayed in a continuous NADH-coupled assay. Briefly, 1 ml assay mixture (50 mM Tris/HCl pH 7.8, 5 mM $MgCl_2$, 10 mM HCO_3, 2 mM PEP, 0.2 mM NADH, 5 units malate dehydrogenase) was incubated at 25°C for 10 min, and then assay was started by addition of a small volume (0.5-3 μl) of PEPc-containing fraction. The change in absorbance at 340 nm of the mixture over time was measured using a Jasco V-550 spectrophotometer, and converted to μmol min^{-1} ml^{-1} using the extinction coefficient of NADH.

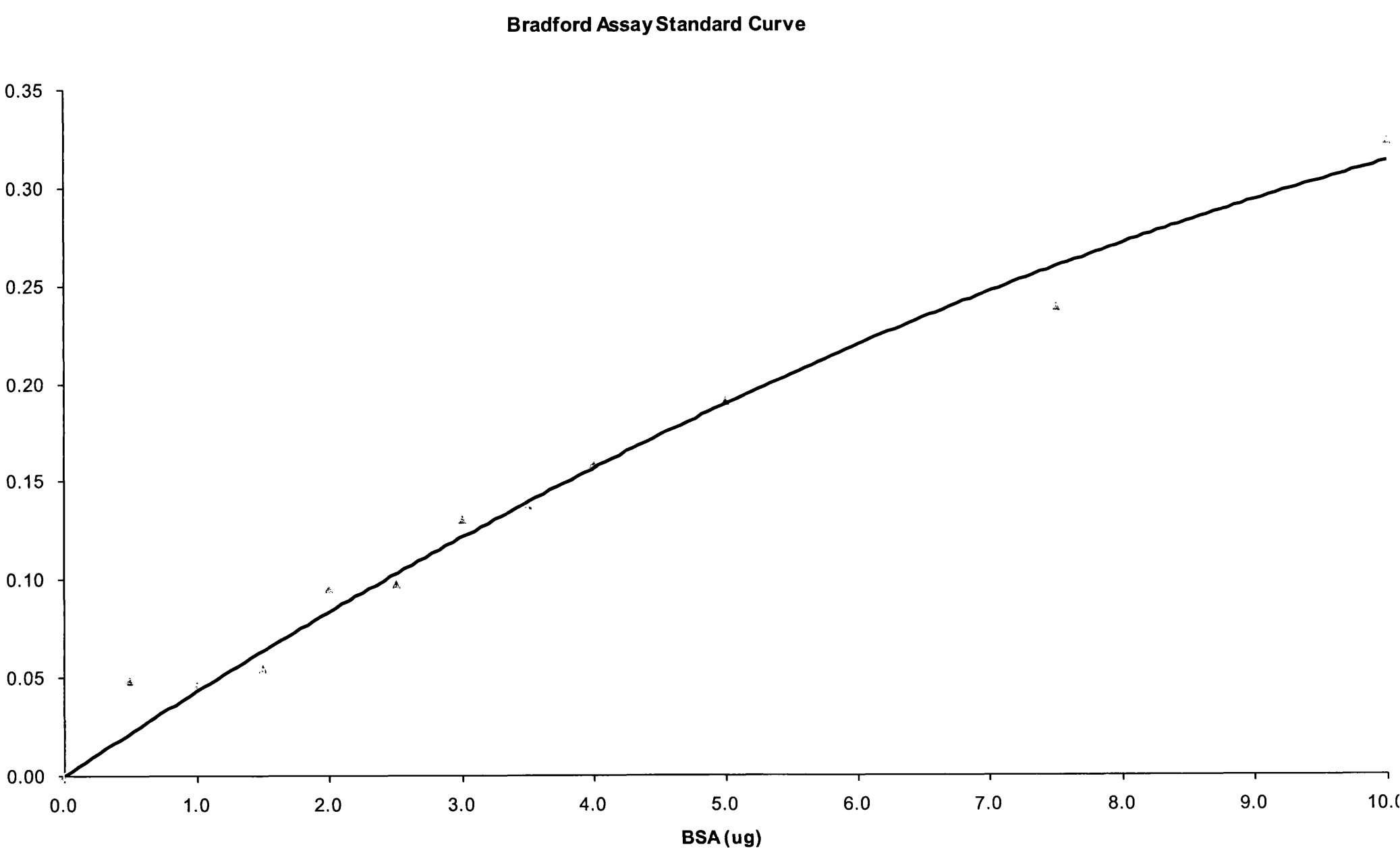
2.7.3 Protein quantification

Protein concentrations were normally measured by the method of Bradford (1976). Briefly, 100 μl of protein solution was mixed with 900 μl of Bradford's reagent and incubated at room temperature for 10 min. The absorbance of the solution at 595 nm was measured using a Jasco V-550 spectrophotometer, using 900 μl Bradford's reagent plus 100 μl dilution buffer as the reference. Using a stock solution of 0.1 mg ml^{-1} bovine serum albumin, a standard curve of A_{595} vs. μg of protein was plotted, and used to calculate the concentration of protein in the unknown solution. A typical standard curve is shown in Fig. 2.2.

When a non-destructive method of quantification was required, the absorbance of protein solution at 280 nm was measured. The PROTPARAM algorithm was used to determine the molar extinction coefficient, using the amino acid sequence of the corresponding protein, obtained from the NCBI

Figure 2.2 Bradford assay standard curve

A solution of bovine serum albumin was diluted in the reference buffer and added to Bradford's reagent as outlined in the text. After a 10 min incubation at room temperature, the absorbance of the mixture was measured at 595 nm. Samples were measured in duplicate and each point represents an average of two readings.


Bradford Assay Standard Curve
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
BSA (ug)

website. When an exact sequence could not be obtained a closely related homologue was used.

2.8 Gel electrophoresis techniques

2.8.1 SDS-polyacrylamide gel electrophoresis

Sodium dodecyl sulphate (SDS)-polyacrylamide gel electrophoresis (PAGE) was carried out according to the method of Laemmli (1970). All gels contained a 3% stacking gel and an appropriate separating gel. In general, PEPc was electrophoresed on an 8% separating gel while PEPc kinase was electrophoresed on a 12% separating gel.

2.8.2 SDS-PAGE gel staining

After electrophoresis, stacking gels were discarded and separating gels incubated in Coomassie staining solution (50% (v/v) methanol, 10% (v/v) acetic acid, 0.1% (w/v) Coomassie Brilliant Blue) for 30 min at 40°C. Gels were destained using destaining solution (10% (v/v) methanol, 10% (v/v) acetic acid) overnight at 40°C.

To detect nanogram amounts of protein, silver staining was used, either straight after electrophoresis or after Coomassie staining. Gels were incubated in 50% (v/v) methanol for at least 4 h to remove acetic acid and/or SDS. 100 ml of silver staining solution was made by adding 4 ml of 0.4 mg/ml silver nitrate dropwise to 21 ml of 0.36% sodium hydroxide with 1.4 ml concentrated ammonium hydroxide. The resulting solution was made up to 100 ml with deionised water and gels were incubated in this mixture for 15 min. Gels were

washed for at least 30 min with deionised water and then developed with developing solution (0.005% (w/v) citric acid, 0.019% (v/v) formaldehyde) until bands were of the desired intensity. Gels were then washed thoroughly with deionised water and kept in 10% (v/v) methanol.

2.8.3 Drying and phosphorimaging of protein gels

Destained gels were dried onto Whatmann 3MM chromatography paper under vacuum using a SCIE-PLAS gel dryer at 80°C for 1 h. Dried radioactive gels were placed in a phosphorimaging cassette and incubated at room temperature typically for 24 h.

2.8.4 Agarose gel electrophoresis

RNA was analysed by denaturing gel electrophoresis using formaldehyde-MOPS as the buffer system. Samples were mixed 1:1 with RNA loading dye and heated to 50°C for 30 min prior to electrophoresis.

3 Purification and crystallisation of recombinant PEPc kinase

3.1 Introduction

One of the main initial aims of this thesis was to characterise the three-dimensional structure of PEPc kinase. The "gold standard" in terms of determining protein structure at atomic resolution is x-ray crystallography, and this technique was chosen for the analysis of the three-dimensional structure of PEPc kinase. X-ray crystallography requires protein that is >95% pure, as judged by SDS-PAGE (Thomas *et al.*, 1998). This is then precipitated in a controlled manner to generate protein crystals, which can be placed in an X-ray beam. A brief outline of the theory of crystallisation is given here.

Proteins are labile molecules, and can be easily denatured by extremes of temperature and buffer conditions. Therefore, crystallisation methods used for inorganic substances are inappropriate, and more gentle methods of protein precipitation must be employed. The protein of interest is dissolved in a buffer containing a precipitant, and allowed to evaporate over time. Evaporation increases the concentration of both protein and precipitant and the protein should therefore form either an ordered crystalline structure or an amorphous precipitate depending on the solution conditions. The speed of this evaporation process may affect the crystal quality: crystals grown more slowly tend to be of a higher quality.

There are two stages of crystal formation: nucleation and growth. Nucleation is the initial formation of crystal "seeds" from which the crystal grows. The initial crystal seed must reach a critical size for it to go on and grow

into a macroscopic crystal, suitable for X-ray analysis. If nucleation is allowed to continue, multiple microscopic crystals will be formed.

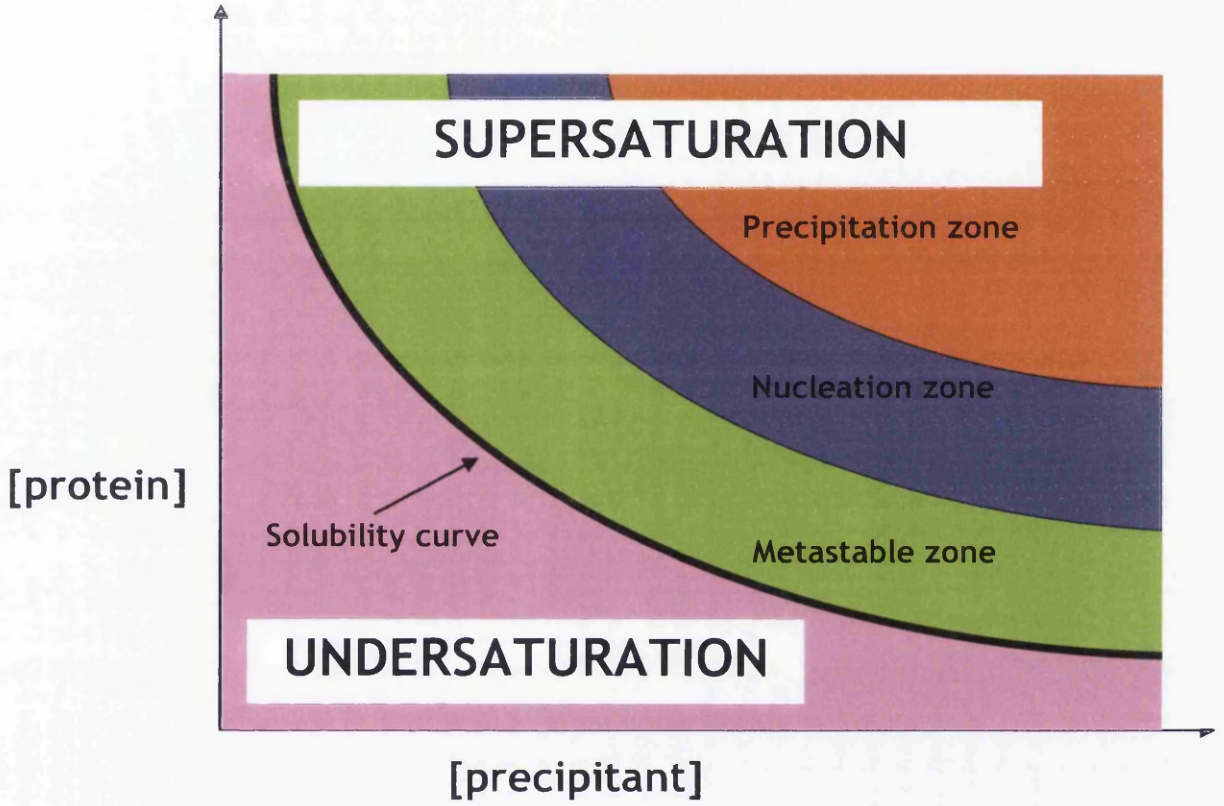
The object of crystallisation screening is to find the optimum protein and precipitant concentrations that promote nucleation for a given time, and then allow growth to become the dominant process. Formation of protein crystals by precipitants can be visualised using a phase diagram (Fig. 3.1). Below the solubility curve, the protein-precipitant solution is undersaturated, and no crystallisation will occur. Above the solubility curve, the solution becomes supersaturated and crystallisation can take place, depending on the relative concentrations of precipitant and protein. These conditions must be empirically determined for each protein, and there is no set of rules that determines which protein crystallises under which conditions.

There are a number of three-dimensional structures of protein kinase catalytic domains already present in the literature (e.g. Knighton *et al.*, 1991). Most of these structures are part of larger polypeptide chains. The three-dimensional structure of PEPc kinase would be of particular interest, as at the time of beginning this work, the smallest protein kinase to have its structure solved was the catalytic subunit of protein kinase A. However, PEPc kinase is significantly smaller and its structure should reveal the bare minimum that is required for conventional protein kinase activity.

This chapter describes expression and purification to homogeneity of recombinant PEPc kinase and conditions for producing microcrystals of recombinant soybean PEPc kinase.

Figure 3.1 Two dimensional representation of a crystallisation phase diagram

The diagram is divided into undersaturated and supersaturated parts by a solubility curve. The supersaturated part is further divided into three zones: the metastable zone, where growth occurs; the nucleation zone, where the formation of crystal "seeds" is promoted; and the precipitation zone, where amorphous precipitation happens. The object of protein crystallisation is to find the correct concentration of protein and precipitant to reach the nucleation zone, and then the metastable zone. Adapted from Wiencek (1999).


SUPERSATURATION
Precipitation zone
Nucleation zone
[protein]
Solubility curve
Metastable zone
UNDERSATURATION
[precipitant]

3.2 Results

3.2.1 Expression and purification of recombinant PEPc kinase

As outlined in the first chapter, PEPc kinase has an extremely low natural abundance. Therefore, to obtain levels of protein required for high-resolution structural analysis, a set of recombinant PEPc kinases were expressed using the methylotrophic yeast *Pichia pastoris*. Five clones of PEPc kinase had been generated in our laboratory: GmPPCK3 (soybean), SbPPCK1 (Sorghum), LePPCK1 (tomato), AtPPCK1 (Arabidopsis) and KfPPCK (Kalanchoë). All of these proteins have a FLAG-tag and a decahistidine-tag at the N-terminus (Fraser, 2006), facilitating their identification and purification. A general purification scheme for recombinant PEPc kinases is described here.

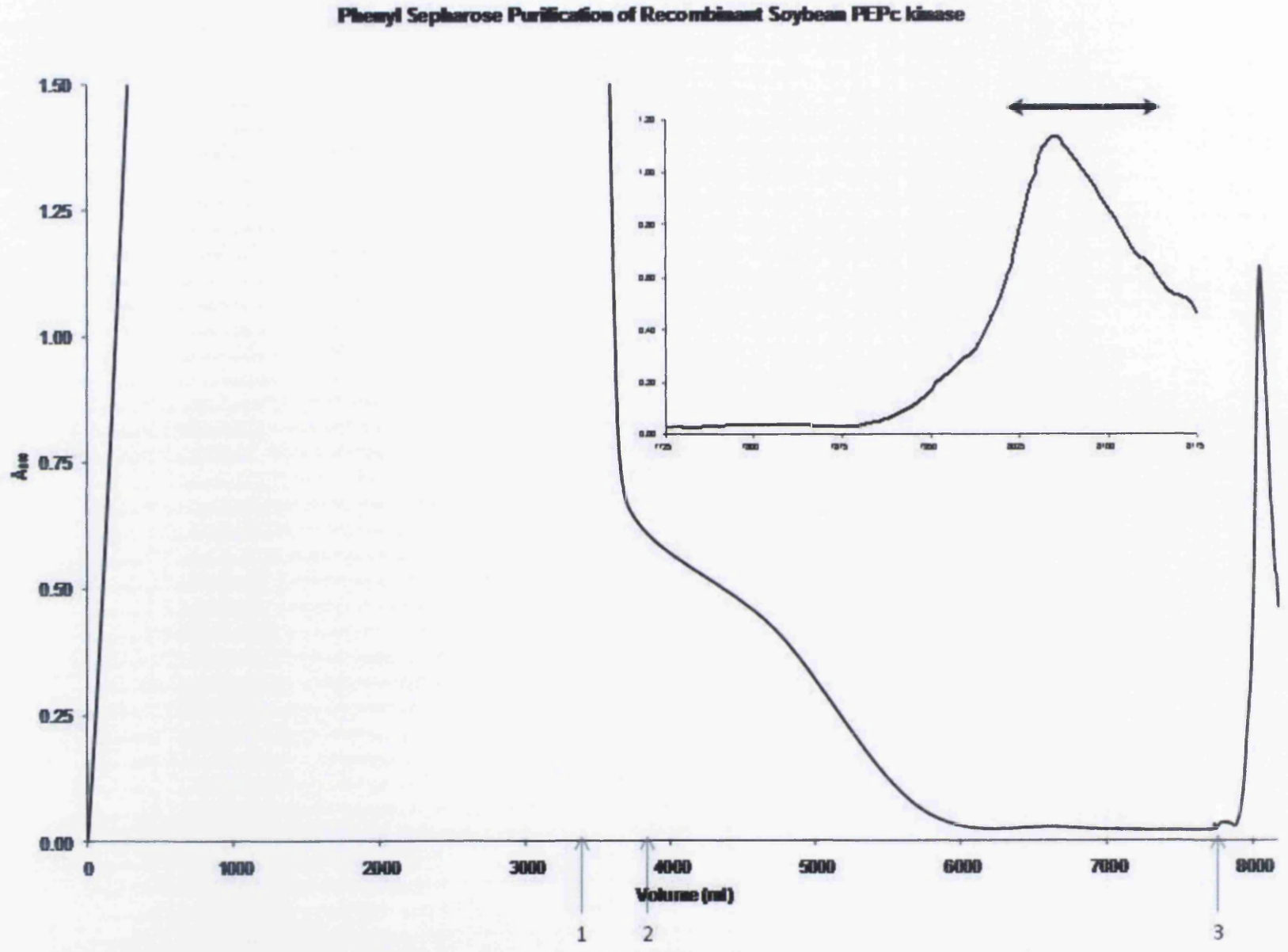
The first stage of purification was hydrophobic interaction chromatography using Phenyl Sepharose. Supernatant from four litres of an induced culture grown for four days in 1% (v/v) methanol was loaded onto a 200 ml column of Phenyl Sepharose. The column was washed with Buffer B containing 30% (v/v) ethylene glycol and then the protein was eluted with Buffer C containing 60% (v/v) ethylene glycol (for more detailed information, including expression conditions, buffer compositions and flow rates, see Chapter 2). Figure 3.2A shows the absorbance trace of the purification steps. Peak fractions were pooled and stored overnight at 4°C in the presence of protease inhibitors (1 μg ml^{-1} each of leupeptin, chymostatin, and antipain).

Next, the Phenyl Sepharose pool was loaded onto a 10 ml Ni-NTA agarose column. Column was washed with 70 mM imidazole buffer before protein was eluted with 500 mM imidazole buffer. Figures 3.2B and 3.2C show the

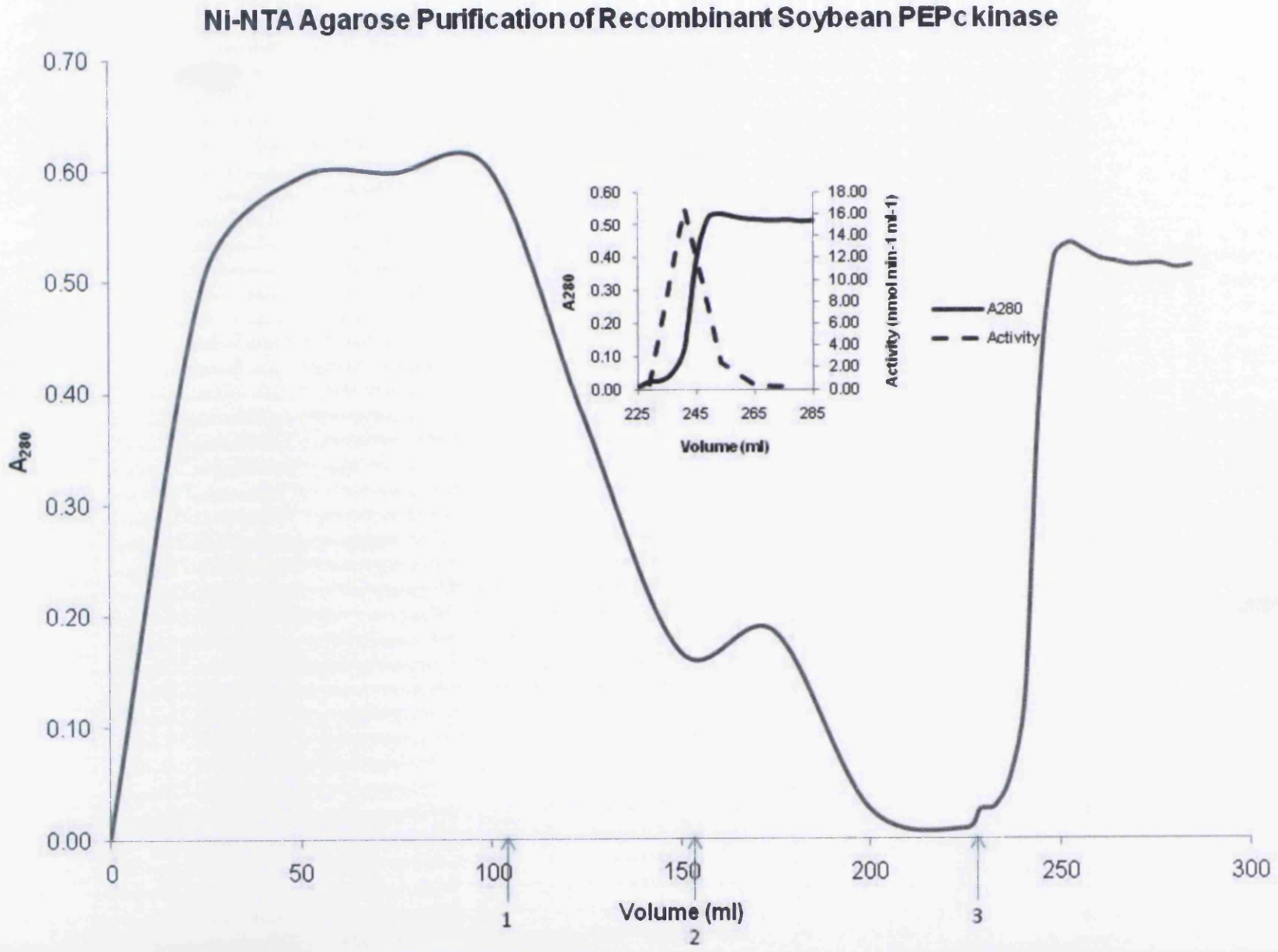
Figure 3.2 Purification of recombinant soybean PEPc kinase

Recombinant soybean PEPc kinase was expressed in four litres of *P. pastoris* cell culture as outlined in the text. The supernatant was applied to a 200 ml column of Phenyl Sepharose pre-equilibrated with Buffer A, at a flowrate of 2 l h^{-1} (A). The column was then washed with 500 ml Buffer A (arrow 1), and 5 litres Buffer B containing 30% (v/v) ethylene glycol (arrow 2). The protein was eluted with Buffer B containing 60% (v/v) ethylene glycol at a flowrate of 5 ml min^{-1} (arrow 3). The pooled fractions are indicated by the black horizontal line (inset). The pooled fractions from Phenyl Sepharose were then loaded onto a 10 ml Ni-NTA agarose column at a flowrate of 5 ml min^{-1} (B). The column was washed with 50 ml of 10 mM imidazole buffer (arrow 1) and then with 70 mM imidazole buffer until the A_{280} of the effluent fell to <0.01 (arrow 2). The protein was eluted with 500 mM imidazole buffer (arrow 3). Fractions were assayed for PEPc kinase activity as outlined in Chapter 2. PEPc kinase activity is given by the dashed line (inset). The purification scheme was analysed and is given in C. SDS-PAGE analysis of the purified and concentrated protein is also shown (D).

A


Phenyl Sepharose Purification of Recombinant Soybean PEPc kinase
1.50
1.25
1.00
0.75
0.50
0.25
0.00
A₂₈₀
0
1000
2000
3000
4000
5000
6000
7000
8000
Volume (ml)
1
2
3

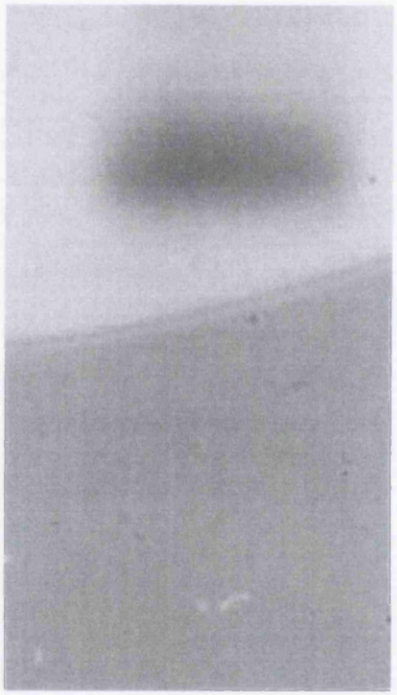
B


Ni-NTA Agarose Purification of Recombinant Soybean PEPc kinase
A280
Activity (nmol min-1 ml-1)
Volume (ml)
A280
Activity
1
2
3

C

Fraction	Volume (ml)	Activity (nmol min^{-1} ml^{-1})	Protein concentration (mg ml^{-1})	Total protein (mg)	Total activity (nmol min^{-1})	Specific activity (nmol min^{-1} mg^{-1})	% yield	Purification fold
Supernatant	3350	0.13	0.315	1055.25	435.50	0.41	100	1
Phenyl Sepharose	106	3.23	0.28	29.68	342.49	11.54	79	28
Ni-NTA agarose	15	9.21	0.05	0.75	138.09	184.12	32	449

D



absorbance trace of the purification steps. Peak fractions were pooled and concentrated using a centrifugal concentrator. The flowthrough from the concentration step was used as a reference to measure the protein concentration. An equal volume of glycerol was added to the concentrated protein and protein was stored at -20°C until required for crystal growth.

PEPc kinase activity was assayed at each step in the purification scheme, using ATP and *Kalanchoë fedtschenkoi* PEPc as substrates, and the results are displayed in Fig. 3.2C. A yield of 32% was achieved using Phenyl Sepharose and Ni-NTA agarose as purification steps. The protein was judged to be sufficiently pure for crystallisation studies by analysis using SDS-PAGE (Fig. 3.2D).

3.2.2 Crystallisation of recombinant soybean PEPc kinase

Concentrated purified recombinant PEPc kinase from soybean was used in crystal trials, since this clone of PEPc kinase gave a consistently higher level of purified protein than the other available clones. A number of commercially available crystal screens were used to determine optimum conditions for PEPc kinase crystallisation (see Appendix 1 for compositions of crystallisation buffers).

Soybean PEPc kinase was expressed in *Pichia pastoris* and purified as outlined in Section 3.2.1. The pool of fractions from Ni-NTA agarose was concentrated to approximately 4 mg ml^{-1} using an ultrafiltration device. The method of crystallisation chosen was the sitting-drop method. 1 μl of crystallisation solution was mixed with 1 μl protein solution in 96-well plates, using a robotic pipetting system. The final concentration of PEPc kinase in the crystal screens was 2 mg ml^{-1}. The ATP analogue AMP-PNP was also used as an

additive in each condition. Therefore, a total of 576 buffer conditions ± AMP-PNP were tried.

Most of the conditions gave amorphous precipitate, but at least one condition gave microcrystals (Fig. 3.3). These were found to be too small to mount in an X-ray beam. This hit was generated using Hampton Crystal Screen 2, which is based on the original sparse matrix screen by Jancarik and Kim (1991) but contains novel precipitants. The solution in this well is 100 mM Tris/HCl pH 8.5, 200 mM $MgCl_2.6H_2O$, 3.4 M 1,6-hexanediol. Optimisation of this condition was tried using a combination of different concentrations of magnesium chloride and 1,6-hexanediol. However, this did not yield any successful results (data not shown). Also, the original hit was not reproduced in the optimisation trials.

3.2.3 Purification and crystallisation of recombinant Sorghum PEPc kinase

The purification scheme in Section 3.2.1 describes a general purification for recombinant PEPc kinases and should be applicable to all clones except the Sorghum isoform. This is because Sorghum PEPc kinase has an acidic insert of residues relative to other PEPc kinases (Nimmo, 2003), which changes its affinity to Phenyl Sepharose considerably.

Recombinant Sorghum PEPc kinase was expressed using *Pichia pastoris*. Supernatant from four litres of cell culture was applied to a 200 ml column of Phenyl Sepharose as before. This time, column was washed with Buffer B with 15% (v/v) ethylene glycol and protein was eluted with Buffer C with 45% (v/v) ethylene glycol (Fig. 3.4A). Ni-NTA agarose purification was then carried out as

Figure 3.3 Crystals of recombinant soybean PEPc kinase

Recombinant soybean PEPc kinase was expressed in Pichia pastoris and purified as described in the text. 1 μl of protein solution (4 mg ml^{-1}) was mixed with 1 μl of crystallisation buffer and incubated at 22°C for a number of weeks. The composition of the crystallisation buffer was 100 mM Tris/HCl pH 8.5, 200 mM $MgCl_2.6H_20$, 3.4M 1,6-hexanediol. The edge of the well is shown in the top left hand corner.

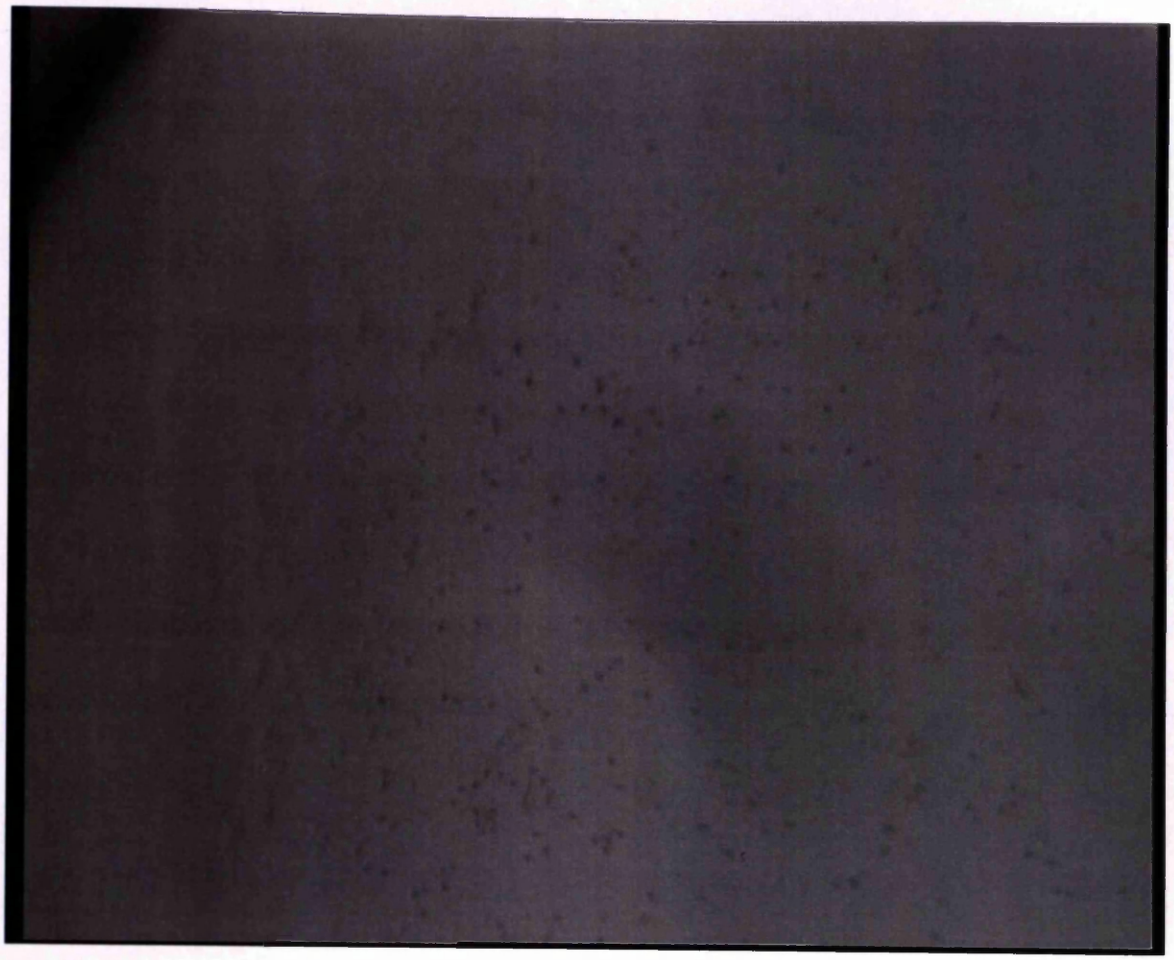
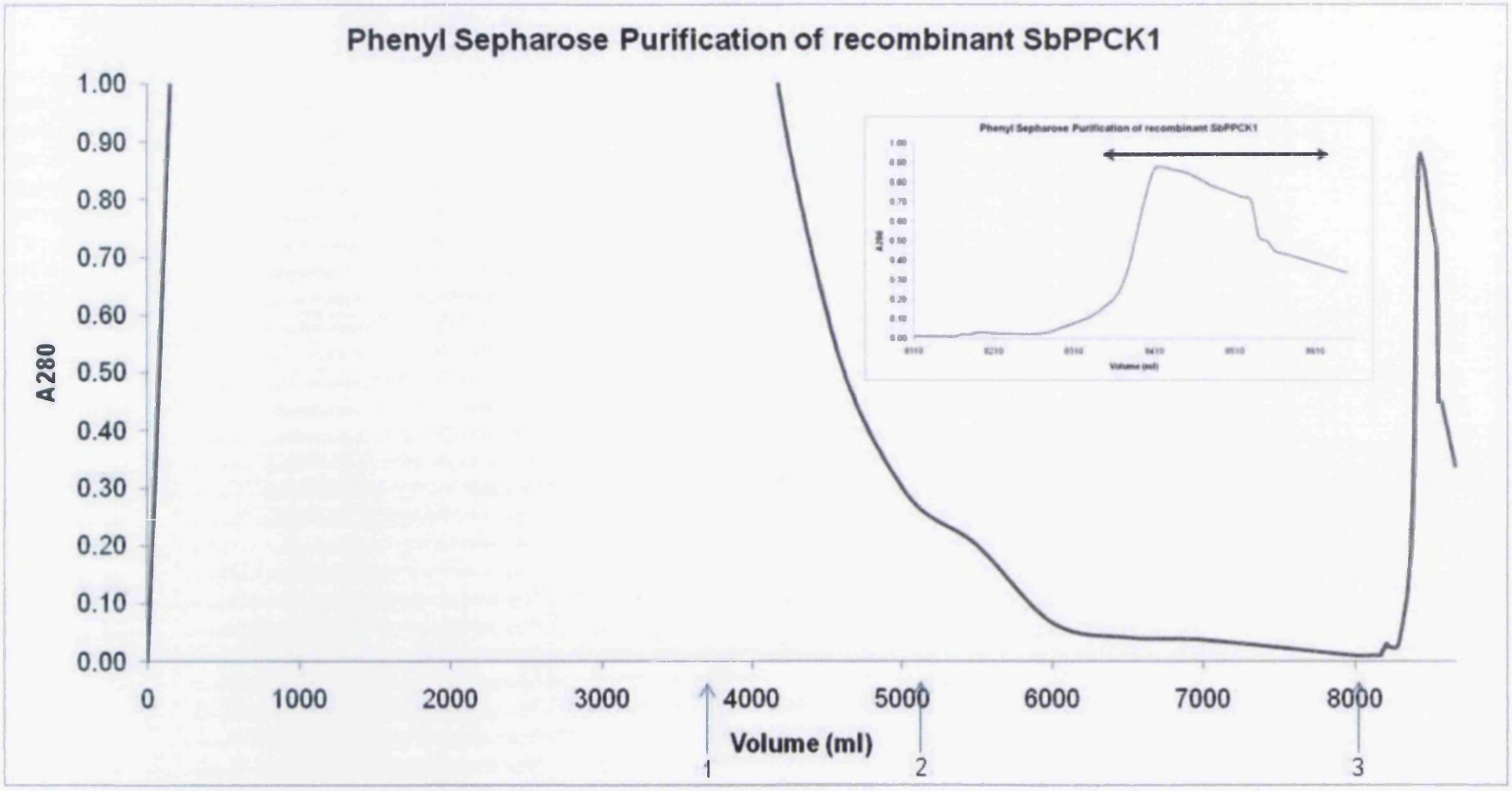


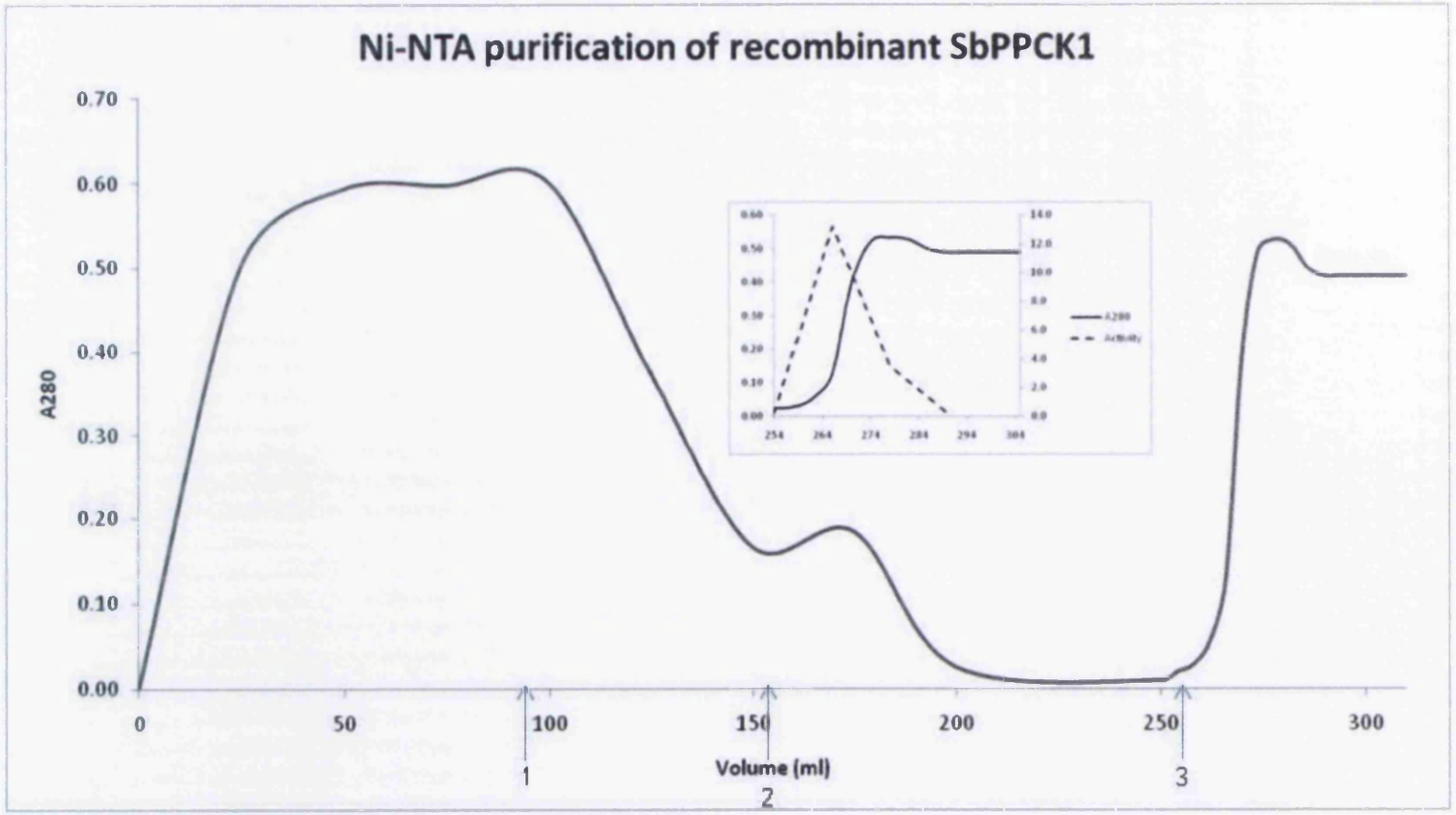
Figure 3.4 Purification of recombinant sorghum PEPc kinase

Recombinant sorghum PEPc kinase was expressed in four litres of *P. pastoris* cell culture as outlined in the text. The supernatant was applied to a 200 ml column of Phenyl Sepharose pre-equilibrated with Buffer A, at a flowrate of 2 l h^{-1} (A). The column was then washed with 500 ml Buffer A (arrow 1), and 5 litres Buffer B containing 15% (v/v) ethylene glycol. The protein was eluted with Buffer B containing 45% (v/v) ethylene glycol at a flowrate of 5 ml min^{-1}. The pooled fractions are indicated by the black horizontal line (inset). The pooled fractions from Phenyl Sepharose were then loaded onto a 10 ml Ni-NTA agarose column at a flowrate of 5 ml min^{-1} (B). The column was washed with 50 ml of 10 mM imidazole buffer and then with 70 mM imidazole buffer until the A_{280} of the effluent fell to <0.01. The protein was eluted with 500 mM imidazole buffer. The purification scheme was analysed and is given in C. SDS-PAGE analysis of the purified concentrated protein is also shown (D).

A


Phenyl Sepharose Purification of recombinant SbPPCK1
A280
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0
1000
2000
3000
4000
5000
6000
7000
8000
Volume (ml)
1
2
3

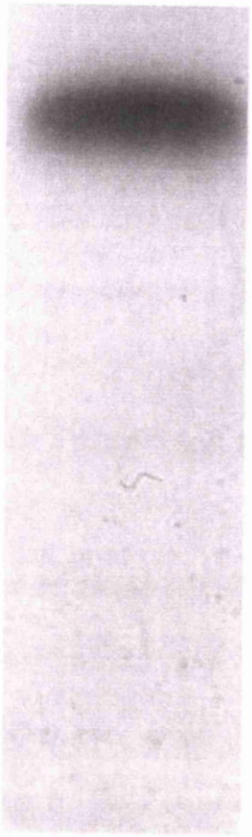
B


Ni-NTA purification of recombinant SbPPCK1
A280
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0
50
100
150
200
250
300
Volume (ml)
1
2
3

C

Fraction	Volume (ml)	Activity (nmol min^{-1} ml^{-1})	Protein concentration (mg ml^{-1})	Total protein (mg)	Total activity (nmol min^{-1})	Specific activity (nmol min^{-1} mg^{-1})	% yield	Purification fold
Supernatant	3520	0.17	0.49	1724.8	598.4	0.347	100	1
Phenyl Sepharose	210	0.63	0.14	30.0	131.6	4.38	22	13
Ni-NTA Agarose	15	9.21	0.03	0.45	72.7	161.56	12	466

D



before (Fig. 3.4B). The purification scheme was analysed as before and is presented in Figure 3.4C.

Recombinant Sorghum PEPc kinase was used in a smaller number of crystal trials due to the failure of soybean PEPc kinase to give suitable crystals. However, no suitable crystals were obtained (data not shown) and due to time constraints, X-ray crystallography analysis of recombinant PEPc kinase was not pursued further.

3.3 Discussion

This chapter has described the expression and purification of recombinant PEPc kinase, and also conditions that promote formation of microcrystals of this protein. These results are now discussed in the context of the current literature.

X-ray crystallography was chosen as a structural determination tool over Nuclear Magnetic Resonance (NMR) spectroscopy. The size limitation of NMR spectroscopy is well documented in the literature. Estimates of this limitation vary; currently it is thought to be approximately 50 kDa. PEPc kinase, with a molecular mass of ~32 kDa, is well within this limit; however, NMR spectroscopy requires multimilligram amounts of protein, uniformly labelled with both carbon-13 and nitrogen-15. While this could be achieved using the *Pichia pastoris* expression system, the levels of PEPc kinase protein expressed (<0.5 mg litre^{-1}) would not make this approach economically viable.

The low natural abundance of PEPc kinase meant that a recombinant protein was required for high-resolution structural analysis. Generally, the first choice of recombinant host is *Escherichia coli*, due to the ease of manipulation of this organism. However, when expressed in this host, PEPc kinase is found

almost exclusively in inclusion bodies (Hartwell *et al.*, 1999). In the same paper, Hartwell *et al.* (1999) reported solubilisation of recombinant PEPc kinase with detergent to give active enzyme. However, all attempts to remove the detergent, or perform detergent exchange, failed (H. Nimmo, personal communication). *Pichia pastoris* was chosen as the host as it is a eukaryotic organism and is more likely to make the correct posttranslational modifications of PEPc kinase. However, this could also be detrimental, as hyperglycosylation of secreted proteins is sometimes reported.

The purification scheme for recombinant PEPc kinase described here was effective in generating sufficiently pure protein for crystal trials (Fig 3.2D). The first stage of purification was a hydrophobic interaction step, using Phenyl Sepharose. This method of purification also features in most natural PEPc kinase purification schemes, and is generally thought to be the most effective step in such schemes (Zhang and Chollet, 1997). Primary structure analyses of PEPc kinases show them to be largely hydrophobic proteins, which would explain their high affinity for hydrophobic chromatography resins. It also explains why Sorghum PEPc kinase, with an insert of acidic residues relative to other PEPc kinases, elutes from Phenyl Sepharose in a lower concentration of ethylene glycol.

The second stage of the recombinant PEPc kinase purification scheme was an immobilised metal affinity chromatography step on Ni-NTA agarose. Normally, four to six histidine residues are used in affinity tags to allow binding to metal resins. The constructs of PEPc kinase used here all have ten histidine residues, allowing the column to be washed with a higher stringency before elution, and therefore removing any potential contaminants.

There is considerable debate in the literature whether recombinant proteins are true representations of their native counterparts. For example, studies using circular dichroism spectroscopy have shown structural differences between native and recombinant proteins. Introducing extra amino acid residues, such as epitope tags and affinity tags, facilitates identification and purification of the target protein. However, extra amino acids may disrupt protein folding and the recombinant protein may have significantly different secondary, tertiary and even quaternary structure. An assay of recombinant soybean PEPc kinase activity, the best method of determining the correct folded state, gave a specific activity of approximately 750 nmol min^{-1} mg^{-1}. This is comparable with protein kinases purified from native sources, suggesting that protein folding has not been disrupted by addition of tags.

Only small crystals of recombinant PEPc kinase were obtained from the crystal screens described here (Fig. 3.3). The presence of the FLAG-tag and his_{10}-tags may be acting as a hindrance to crystallisation, as the increased mobility of the N-terminus might prevent tight packing into a crystalline array. Nevertheless, there are examples in the literature of recombinant proteins that have been crystallised with epitope and purification tags as part of the polypeptide chain. In this case, the recombinant PEPc kinase constructs used do not have a cleavage site between the C-terminus of the his_{10}-tag and the N-terminus of the PEPc kinase polypeptide.

Many discrete microcrystals can be seen in Fig. 3.3, suggesting that multiple nucleation events have occurred with little growth. As outlined in Section 3.1, the object of crystallisation screening is to first find conditions that promote nucleation; but crystals do not grow under these conditions. Growth

conditions were sought by screening around the nucleation conditions; however most of the new permutations of conditions generated amorphous precipitate. Also, the original hit could not be reproduced in the optimisation screen, suggesting variability between protein preparations.

A second approach to the problem of failure of a protein to crystallise is to try the same protein from another species. In this case the Sorghum enzyme was chosen, for two reasons. Firstly, the expression levels of our clone are acceptable for structural studies. Secondly, the insert of acidic amino acid residues may be biologically significant. This protein also failed to crystallise

One of the key criteria of obtaining protein crystals is sample homogeneity. Even microheterogeneity can have a detrimental effect on crystal size and quality (Thomas *et al.*, 1998). Not only must the protein be >95% pure as judged by SDS-PAGE, there must be homogeneity in the structure of the protein i.e. all the protein molecules must have the same conformation. Subsequent work, described in the next chapter, showed that this is not the case with recombinant PEPc kinase. With hindsight, it would have been better to check whether the protein was monodisperse after purification conditions were established and before crystal screening was begun. However, the success of the initial screen led us to believe that recombinant PEPc kinase was indeed monodisperse, and it was not until later that the converse was established.

4 Structural characterisation of recombinant PEPc kinase

4.1 Introduction

The previous chapter covered purification and initial crystallisation conditions for recombinant PEPc kinase. Discrete microcrystals of PEPc kinase were obtained but these were unsuitable for X-ray analysis. It was proposed in the Discussion section of Chapter 3 that this is due to a lack of monodispersity in the protein solution. The first aim of the work described in this chapter was to characterise the properties of PEPc kinase in solution, to investigate further why macroscopic crystals were not obtained from the crystallisation screens. The second aim was to characterise the structure of recombinant PEPc kinase at a lower resolution than afforded by X-ray crystallography.

A simple way to analyse the characteristics of a protein in solution, and to characterise quaternary structure, is to use size exclusion chromatography (SEC). SEC involves applying a protein solution to a column of porous beads, such as Superose, Sephacryl or Superdex. Larger macromolecules have a higher excluded volume than smaller macromolecules, partition more to the mobile phase, and therefore travel through the column quicker and elute in a lower volume. Thus, proteins can be separated by their mass and, by calibrating the column with proteins of known molecular weight, their mass can be estimated. Since this is a non-denaturing technique, it provides an estimate of native molecular weight and comparison with subunit molecular weight can allow quaternary structure to be described.

This chapter made use of analytical ultracentrifugation (AUC) to examine the properties of recombinant PEPc kinase in solution. The analytical ultracentrifuge is used in two modes: sedimentation velocity centrifugation and sedimentation equilibrium centrifugation. In sedimentation velocity (SV), the rotor is spun at a very high speed, and the sedimentation coefficient reflects the size and shape of the molecule. In sedimentation equilibrium (SE), the rotor is spun at a lower speed over a longer period of time until diffusion and sedimentation reach equilibrium. From this, the absolute molecular mass of a protein can be determined, and hence the quaternary structure characterised. Both AUC techniques are used here to obtain maximum data.

Structural characterisation of recombinant PEPc kinase in this chapter was carried out by circular dichroism spectroscopy. Circular dichroism is a low-resolution structural technique that can be used to estimate the amount of secondary structure in a protein molecule. Kelly *et al.* (2005) cover structural analysis of proteins by CD in greater detail and the reader is directed to their review for further information.

The results in this chapter reveal an interesting feature of recombinant soybean PEPc kinase in solution. When the protein was analysed by SEC, two peaks were resolved, with different specific activities. AUC analysis of the protein did not show any evidence for oligomerisation. Attempts to resolve this apparent anomaly are made in the Discussion.

4.2 Results

4.2.1 Size exclusion chromatography of recombinant PEPc kinase

The quaternary structure of recombinant soybean PEPc kinase was analysed by SEC on a Superose 12 column. Recombinant soybean PEPc kinase was expressed and purified as outlined in Chapter 3. In the standard conditions 100 μl of protein solution was then applied to a pre-equilibrated Superose 12 column at a rate of 0.3 ml min^{-1} using a BioCAD 700E system. The A_{280} of the effluent was measured, and 1 min fractions were collected for kinetic analysis.

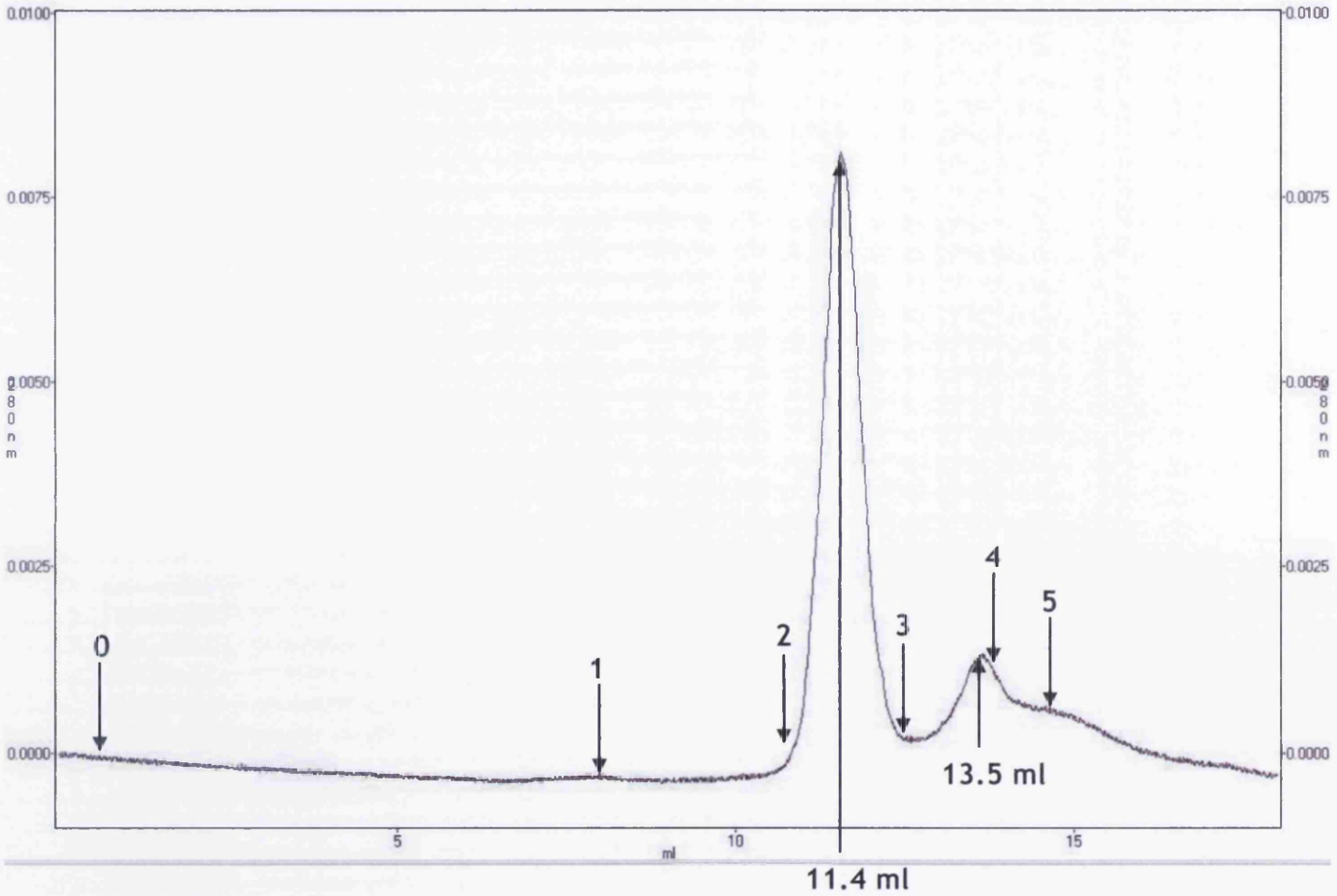
The A_{280} trace of Superose 12 chromatography of recombinant PEPc kinase shows two peaks of absorbance with elution volumes of 11.4 ml and 13.5 ml (Fig. 4.1A). Calibration of this column with proteins of known molecular weight (Fig. 2.1) shows that the lower elution volume peak corresponds to a mass of 120 kDa, while the higher elution volume peak corresponds to a mass of 30 kDa. The molecular weight of recombinant PEPc kinase is approximately 32 kDa so 120 kDa and 30 kDa peaks could correspond to tetramer and monomer respectively. This behaviour is reproducible and has been seen with several different preparations of soybean PEPc kinase. It has also been seen regardless of the starting concentration of the solution of PEPc kinase. PEPc kinases from other species (Sorghum and Arabidopsis) have also been analysed and show a similar behaviour. SEC was carried out with and without glycerol and the same result was obtained in both cases.

To examine whether these two forms are in equilibrium, fractions corresponding to each peak were pooled, concentrated, and then re-chromatographed. When the 120 kDa peak was pooled, concentrated and then

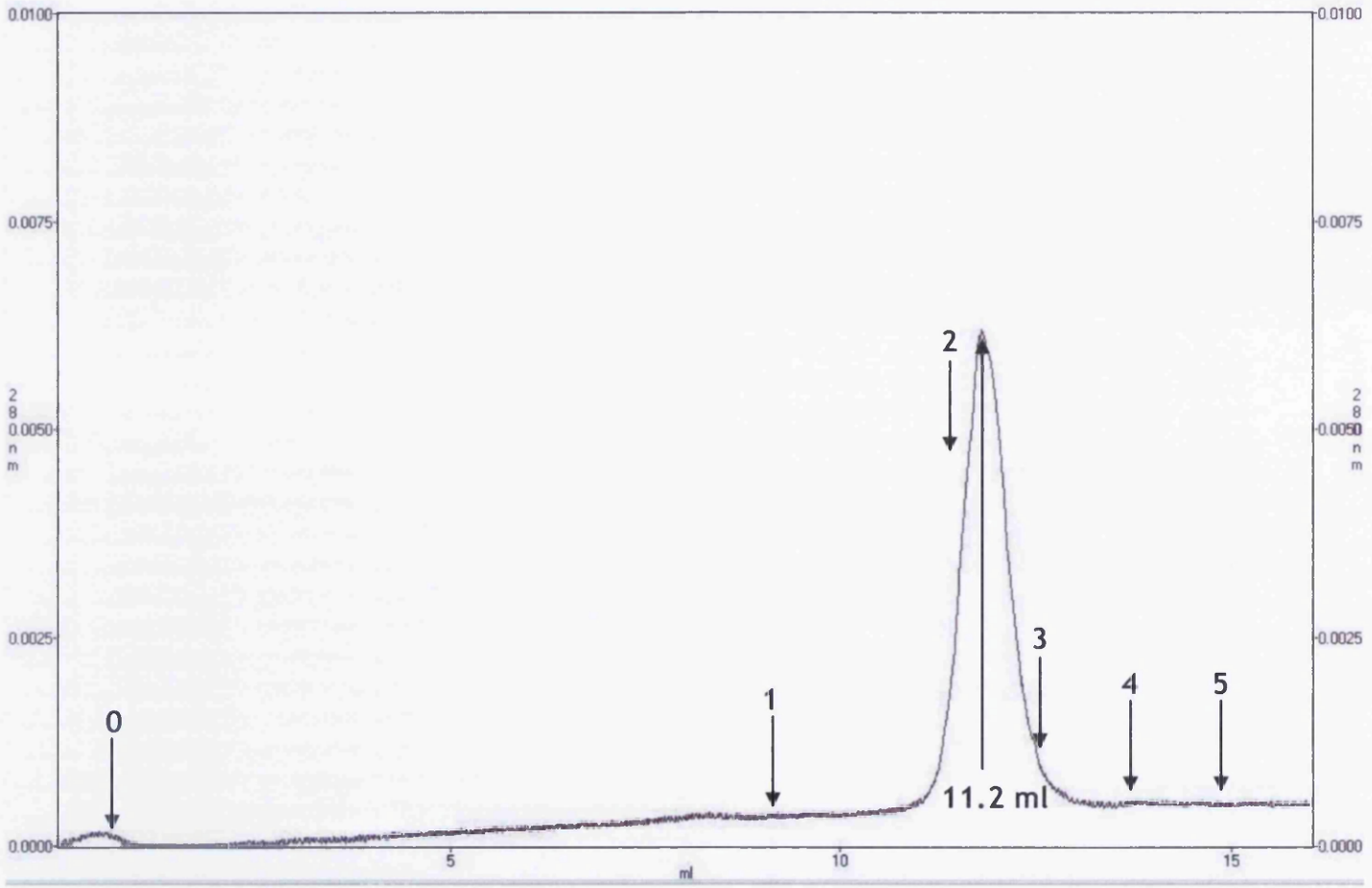
Figure 4.1 Superose 12 chromatography of recombinant soybean PEPc kinase

Recombinant soybean PEPc kinase was expressed and purified as outlined in the text. The buffer was exchanged to 50 mM HEPES pH 7.4, 1 mM DTT, 150 mM NaCl using a centrifugal concentrator and the final protein concentration was determined as 0.2 mg ml^{-1} by the method of Bradford (1978). 100 μl of protein solution was applied to a Superose 12 column pre-equilibrated with 50 mM HEPES pH 7.4, 1 mM DTT, 150 mM NaCl at a flowrate of 0.3 ml min^{-1} using a BioCAD 700E system. The A_{280} of the effluent was monitored to give the resulting trace (A). Fractions corresponding to the 120 kDa peak were pooled, concentrated and re-chromatographed (B). The same treatment was carried out on the fractions corresponding to the 30 kDa peak (C). The arrows numbered "0" indicate the volume at which the sample was applied to the column and the elution volumes of the peaks relative to this point are given. The elution volumes of standard proteins are indicated by the numbered arrows: thyroglobulin, 669 kDa (1); aldolase, 158 kDa (2); bovine serum albumin, 67 kDa (3); carbonic anhydrase, 29 kDa (4); myoglobin, 17 kDa (5).

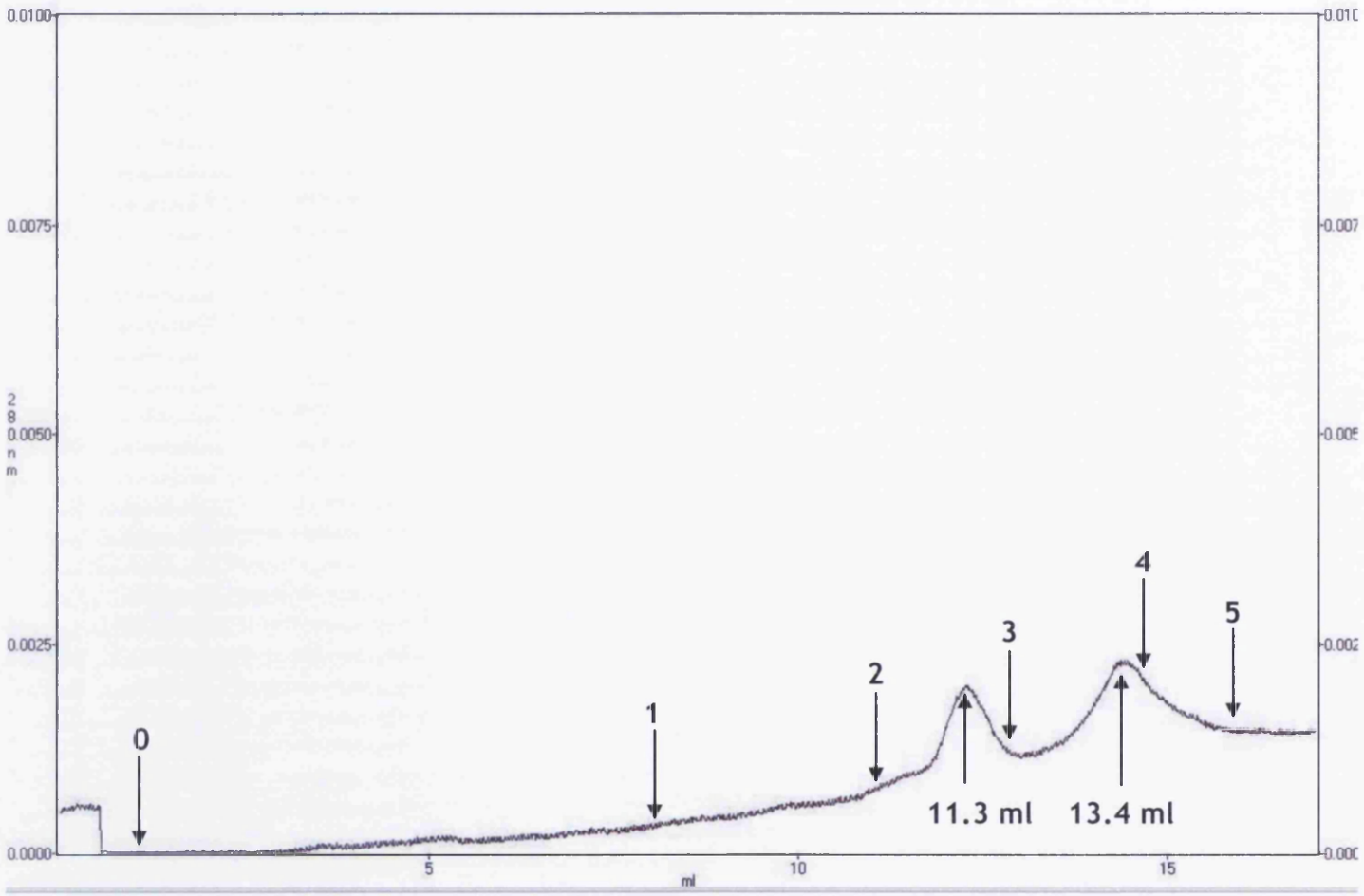
A


0.0100
0.0075
0.0050
0.0025
0.0000
280nm
0
1
2
3
4
5
13.5 ml
5
10
15
ml
11.4 ml

B



C


0.0100
0.0075
0.0050
0.0025
0.0000
280 nm
5
10
15
ml
0
1
2
3
4
5
11.3 ml
13.4 ml

re-chromatographed, in some runs a single 120 kDa peak was obtained (Fig. 4.1B). However, on other occasions some splitting into two peaks was observed (data not shown). When the 30 kDa peak was pooled, concentrated and then re-chromatographed, two peaks were consistently obtained - one at 30 kDa and one at 120 kDa (Fig. 4.1C).

The PEPc kinase specific activity of the protein in each peak was determined using a radiometric kinase assay, using *Kalanchoë fedtschenkoi* PEPc as the substrate, as outlined in Chapter 2. The specific activity of the putative tetramer was 30 nmol min^{-1} mg^{-1} irrespective of protein concentration. The specific activity of the putative monomer varied markedly with dilution factor, and as the concentration of protein increases, the specific activity was found to decrease (Fig. 4.2). Taking the SEC data and the activity data into account, the following hypothesis was proposed: *"PEPc kinase forms concentration-dependent oligomers in solution, which leads to a decrease in specific activity."* This hypothesis was tested using analytical ultracentrifugation, which can give information on the quaternary structure of molecules in solution.

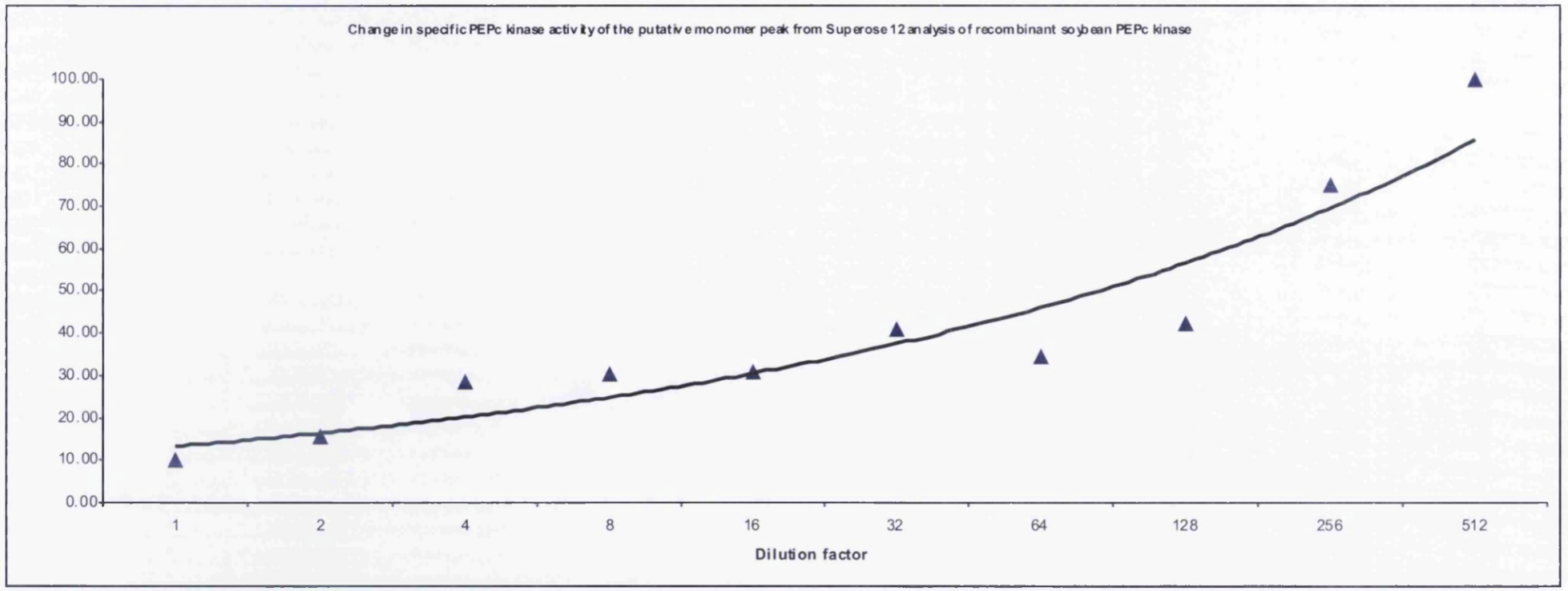
4.2.2 Sedimentation velocity analysis of recombinant PEPc kinase

The oligomerisation state of PEPc kinase in solution was further investigated by AUC, to test the hypothesis proposed in Section 4.2.1. Both sedimentation velocity and sedimentation equilibrium analysis was carried out to maximise the amount of data acquired.

Recombinant soybean PEPc kinase was expressed and purified as outlined in Chapter 3. Protein was buffer exchanged into 50 mM HEPES/NaOH pH 7.5,

Figure 4.2 Specific PEPc kinase activity of the putative monomer peak from Superose 12 chromatography of recombinant soybean PEPc kinase

Recombinant soybean PEPc kinase was analysed by Superose 12 chromatography as outlined in the text. The peak fraction corresponding to the putative monomer was assayed for PEPc kinase activity at increasing dilution as outlined in Chapter 2 and specific activity was calculated. Specific activity at the highest dilution measured was set as 100% and specific activity at lower dilutions calculated as a percentage of the highest dilution.


Change in specific PEPc kinase activity of the putative monomer peak from Superose 12 analysis of recombinant soybean PEPc kinase
100.00
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
1
2
4
8
16
32
64
128
256
512
Dilution factor

1 mM DTT, 150 mM NaCl using a HiPrep 26/10 Desalting column, concentrated using a centrifugal concentration device, and then loaded into the AUC cells. All SV experiments were carried out at a rotor speed of 49,000 rpm in a Ti-60 rotor at 4°C. Three loading concentrations were used: 0.3 mg ml^{-1}, 0.2 mg ml^{-1} and 0.1 mg ml^{-1} as determined by UV spectroscopy. 203 scans were collected using interference optics (Fig. 4.3A) and a c(S) distribution was generated using the analysis program SEDFIT with all scans (Fig. 4.3B). A c(S) distribution is the concentration of molecules in the sample with a particular sedimentation coefficient.

From the c(S) distribution for the sample of PEPc kinase, it can be seen that the theoretical sedimentation coefficient peaks at 1.1 S at 4°C, with a smaller peak at 2.0 S. When the 1.1 S value is corrected to 20°C it becomes 1.5 S. This is largely due to the higher viscosity at low temperatures. Therefore, the experimental sedimentation coefficient of recombinant soybean PEPc kinase is 1.5 S.

4.2.3 Homology modelling of soybean PEPc kinase

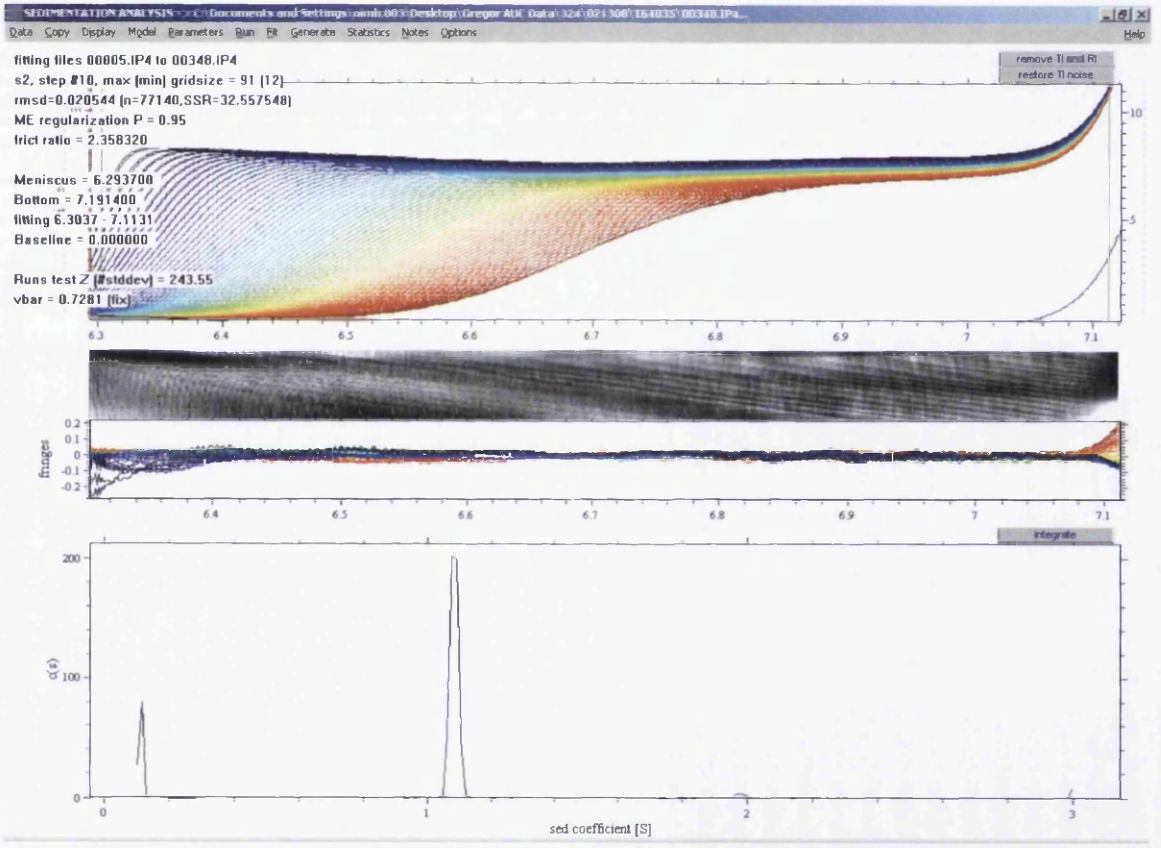
In an attempt to throw further light on the discrepancy between the SEC and the sedimentation velocity experiments, a homology model of soybean PEPc kinase was created using the SWISS-MODEL algorithm. This allowed an estimate to be made of what the sedimentation coefficient would be if PEPc kinase were a monomer.

SWISS-MODEL gives the user two options for modelling the three dimensional structure of proteins from their primary sequence. The algorithm can find homologous proteins in the Protein Data Bank itself, or the user can

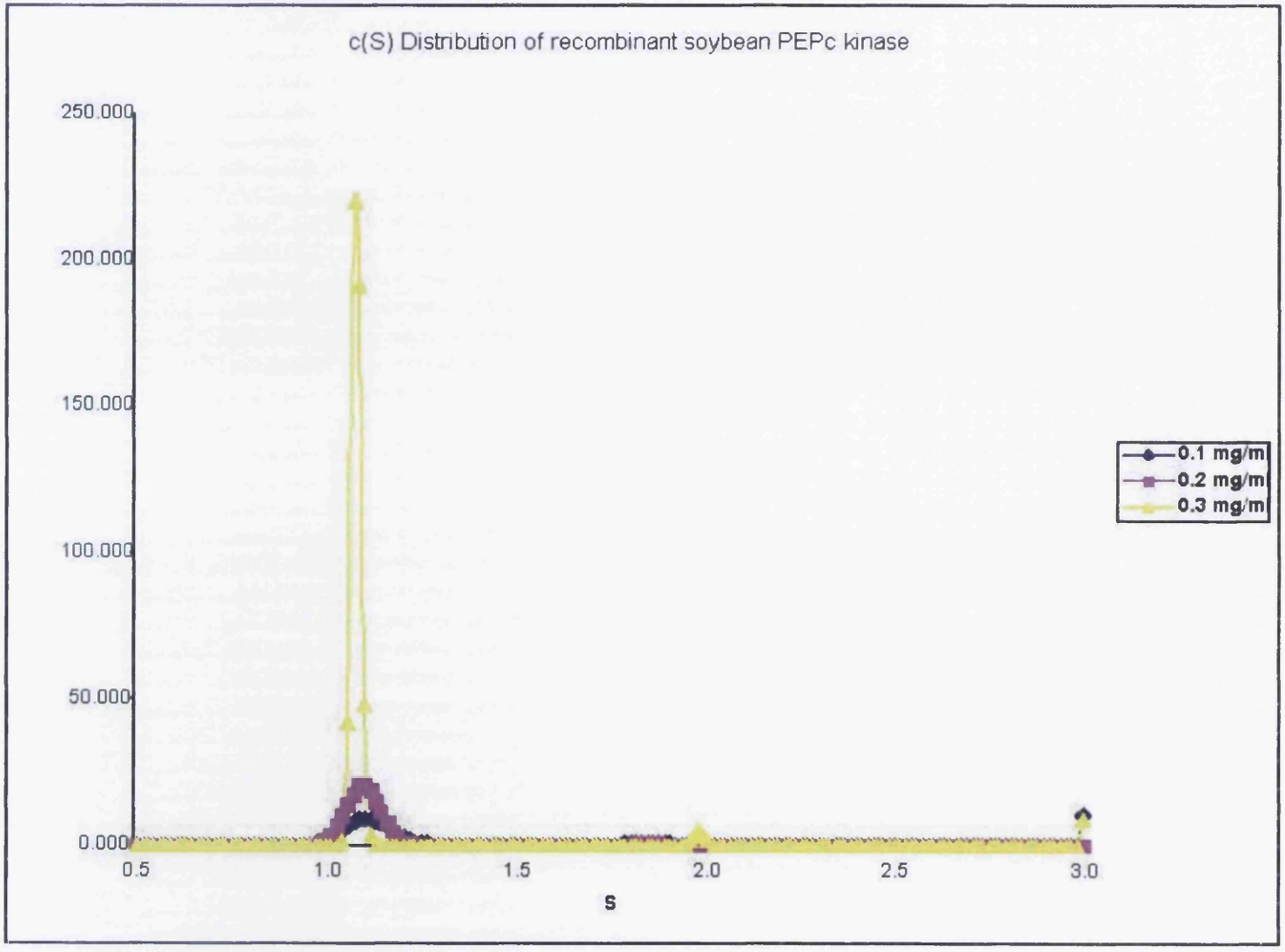
Figure 4.3 Sedimentation velocity ultracentrifugation of recombinant soybean PEPc kinase

Recombinant soybean PEPc kinase was expressed and purified as outlined in Chapter 3. The buffer was exchanged to 50 mM HEPES pH 7.4, 1 mM DTT, 150 mM NaCl using a HiPrep 26/10 column attached to a BioCAD 700E system. Protein was concentrated using a centrifugal concentration device and loaded into the analytical ultracentrifugation cells at 3 different protein concentrations: 0.1 mg ml^{-1}, 0.2 mg ml^{-1} and 0.3 mg ml^{-1}. The cells were spun at 49,000 rpm at 4°C and the rate of sedimentation measured using interference optics. Representative data of the 0.3 mg ml^{-1} loading is shown (A). 203 scans were collected and analysed using SEDFIT to give a c(S) distribution (B).

A


Data Copy Display Model Parameters Run Fit Generate Statistics Notes Options Help
fitting files 00005.IP4 to 00348.IP4
s2, step #10, max (min) gridsize = 91 (12)
rmsd=0.020544 (n=77140,SSR=32.557548)
ME regularization P = 0.95
frict ratio = 2.358320
Meniscus = 6.293700
Bottom = 7.191400
fitting 6.3037 - 7.1131
Baseline = 0.000000
Runs test Z (#stddev) = 243.55
vbar = 0.7281 (fix)
remove TI and RI
restore TI noise
fringes
Integrate
c(s)
sed coefficient [S]

B


c(S) Distribution of recombinant soybean PEPc kinase
250.000
200.000
150.000
100.000
50.000
0.000
0.5
1.0
1.5
2.0
2.5
3.0
s
0.1 mg/m
0.2 mg/m
0.3 mg/m

specify files that the primary sequence can be modelled on. In this case, the soybean PEPc kinase primary sequence was modelled on homologous proteins found in the Protein Data Bank by the SWISS-MODEL program. The homology model is shown in Fig. 4.4A. It can be seen from this figure that the resulting three-dimensional structure is characteristic of a serine/threonine protein kinase domain, with a bi-lobal structure.

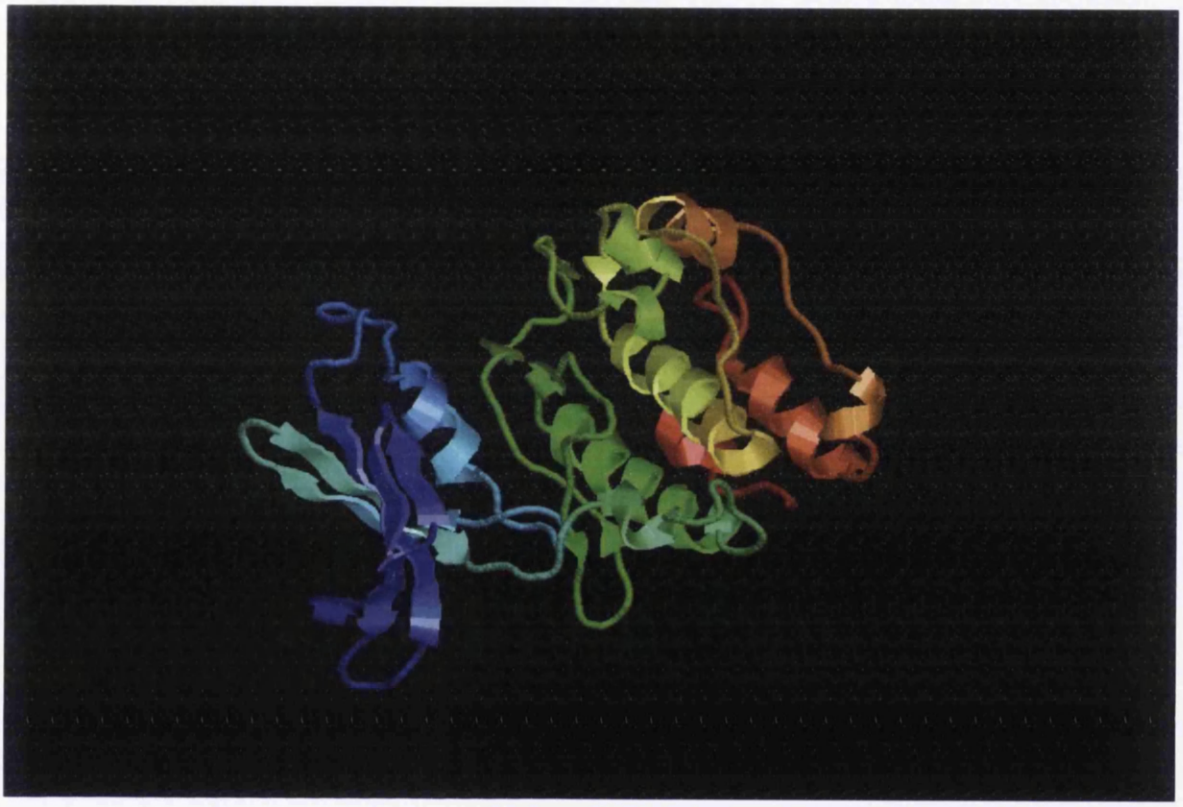
The model was more closely inspected using a number of different methods (Fig. 4.4B). Verify3D looks at the secondary structures in which each amino acid is found and compares these with the likelihood of that amino acid appearing in that structure in all known proteins. This analysis merely showed that all the residues in the homology model were in "good" conformations (>0). ANOLEA and Gromos calculate the energy associated with each residue given its local environment (negative = favourable, positive = unfavourable). There is fairly good agreement between the two methods. In particular, regions 56-72 and 132-147 are calculated to be relatively unfavourable by both programs. Both of these regions correspond to surface exposed sites on the homology model (Fig. 4.4C). It is interesting to note that there is one region of some disagreement between the methods. This is region 217-229, another region that is surface exposed.

One way of testing the model is to compare experimental and calculated sedimentation coefficients. The calculated sedimentation coefficient for this model is 2.67 S. This is considerably higher than the experimental sedimentation coefficient obtained for recombinant soybean PEPc kinase (1.5 S). A possible explanation for the large discrepancy between the two sedimentation coefficients is that the PEPc kinase molecule observed in the experiment is

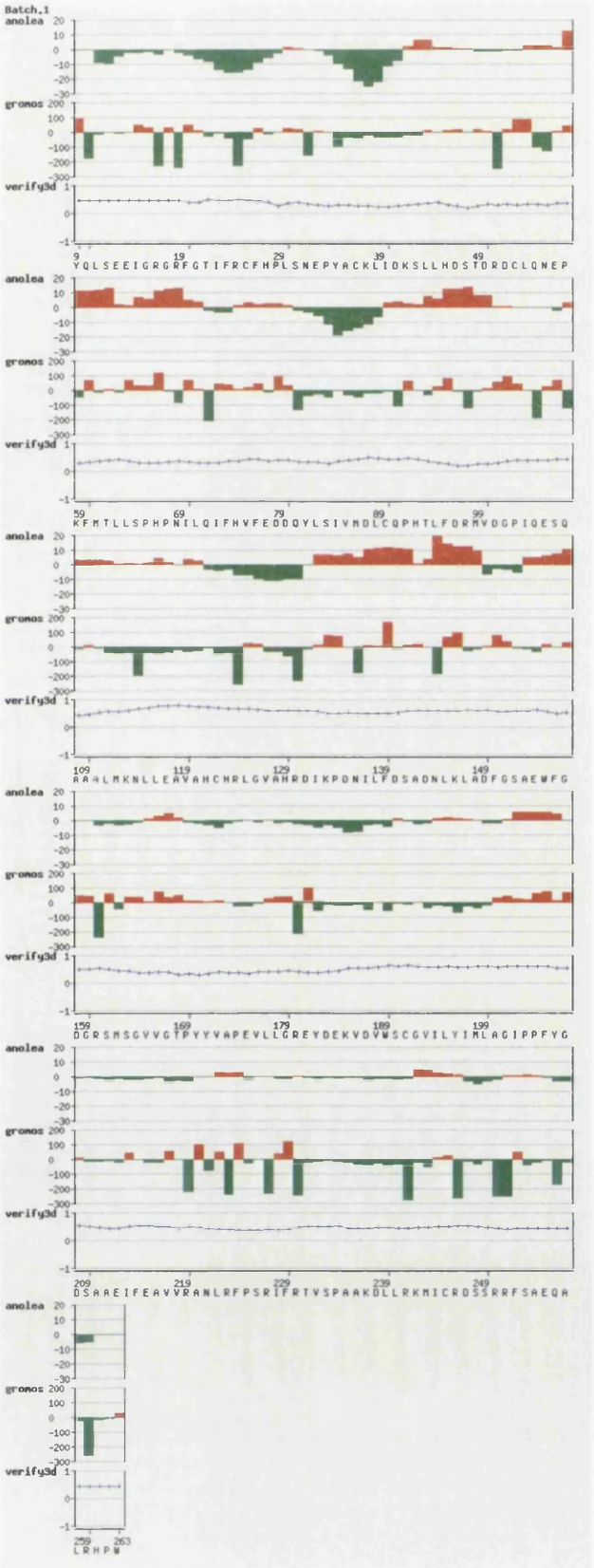
Figure 4.4 Homology model of GmPPCK3

Using the primary sequence of GmPPCK3 and the SWISS-MODEL algorithm, a theoretical model of GmPPCK3 was created (A). The model was analysed using a number of different programs (B): Verify3D showed that all of the residues are in good conformations; ANOLEA and Gromos showed that there are a number of regions of unfavourable energy. These are coloured (C) green (regions 56-72), red (regions 132-147) and cyan (regions 217-229).

A



B


Batch.1
anolea
gromos
verify3d
9 19 29 39 49
YQLSEEIGRGRFGTIFRCFHPLSNEPYACKLIQKSLLHDSTDRDCLQNEP
59 69 79 89 99
KFMTLLSPHPNILQIFHVFEDDQYLSIVMDLCQPHTLFDRMVDGPIQESQ
109 119 129 139 149
AAALMKNLLEAVAHCHRLGVAHRDIKPDNILFDSADNLKLADFGSAEWFG
159 169 179 189 199
DGRSMSGVVGTPYYVAPEVLLGREYDEKVDVWSCGVILYIMLAGIPPFYG
209 219 229 239 249
DSAAEIFEAVVRANLRFPSRIFRTVSPAAKDLLRKMICRDSSRRFSAEQA
259 263
LRHPW

C

asymmetric. Also, there is a relatively large tag at the N-terminus of the molecule, which may also influence the sedimentation coefficient.

4.2.4 Sedimentation equilibrium analysis of recombinant PEPc kinase

Sedimentation equilibrium (SE) experiments were carried out at a rotor speed of 20,000 rpm in a Ti-60 rotor at 4°C for 24 h. Interference optics were used to scan the cells and the SE data was analysed using the SEDPHAT algorithm. The SE data obtained from SEDPHAT are shown in Fig. 4.5A.

The apparent molecular mass at each loading concentration was plotted on a graph (Fig. 4.5B). When the line of best fit is extrapolated back to infinite dilution, the apparent molecular mass is approximately 31,000 Da. Therefore it would appear that recombinant soybean PEPc kinase is monomeric.

4.2.5 Further size exclusion chromatography of recombinant PEPc kinase

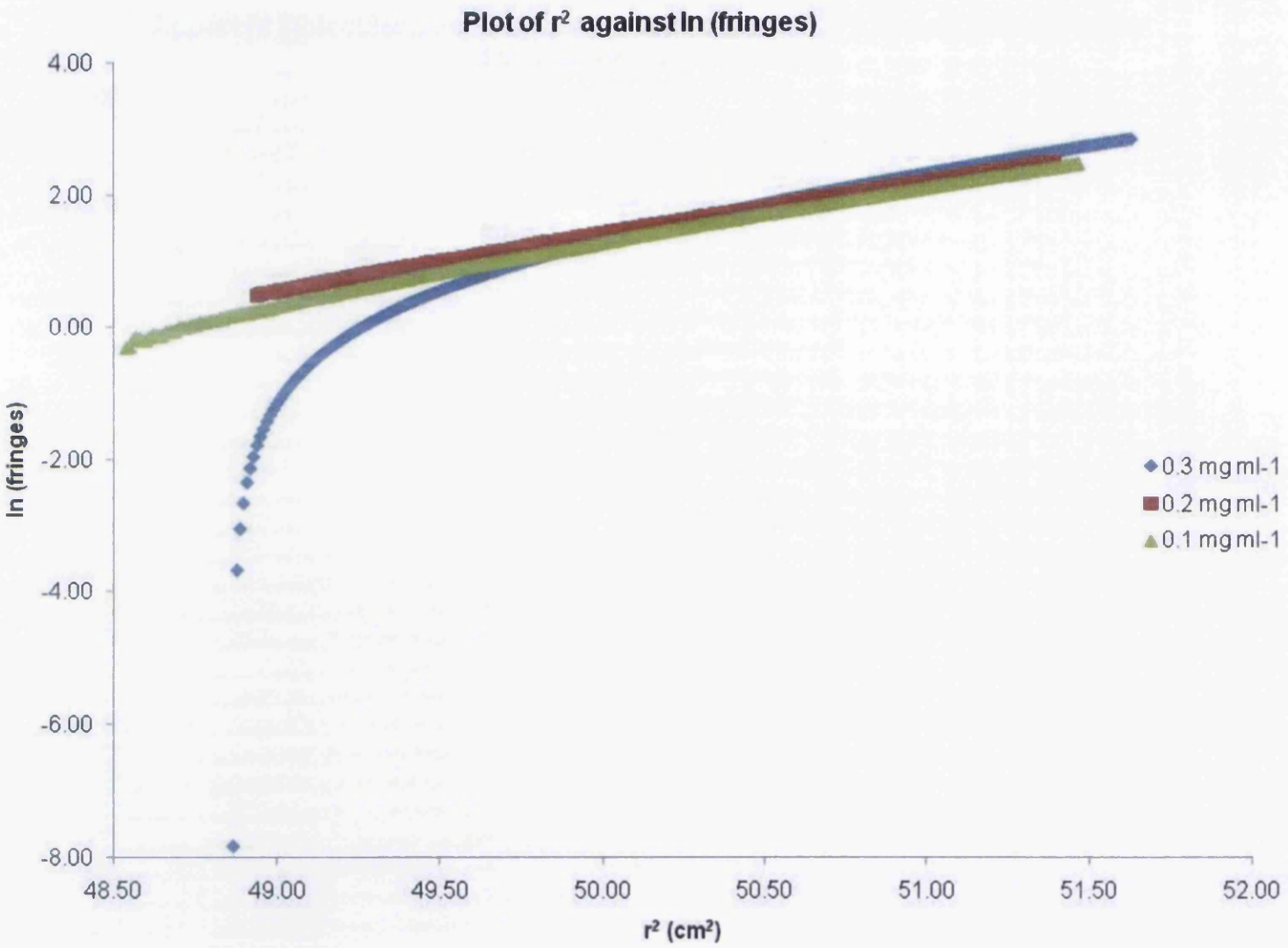
In the standard experiment, SEC analysis of recombinant PEPc kinase was carried out at room temperature. However, AUC analysis of recombinant PEPc kinase was carried out at 4°C to minimise sample degradation over the extended period of time that an AUC experiment is carried out. To make a more valid comparison between the AUC data and the SEC data, SEC analysis of recombinant PEPc kinase was also carried out at 4°C.

Recombinant soybean PEPc kinase was expressed, purified and concentrated as outlined previously. 100 µl of protein solution was applied to a pre-equilibrated Superose 12 column at 4°C at a flowrate of 0.3 ml min^{-1} using

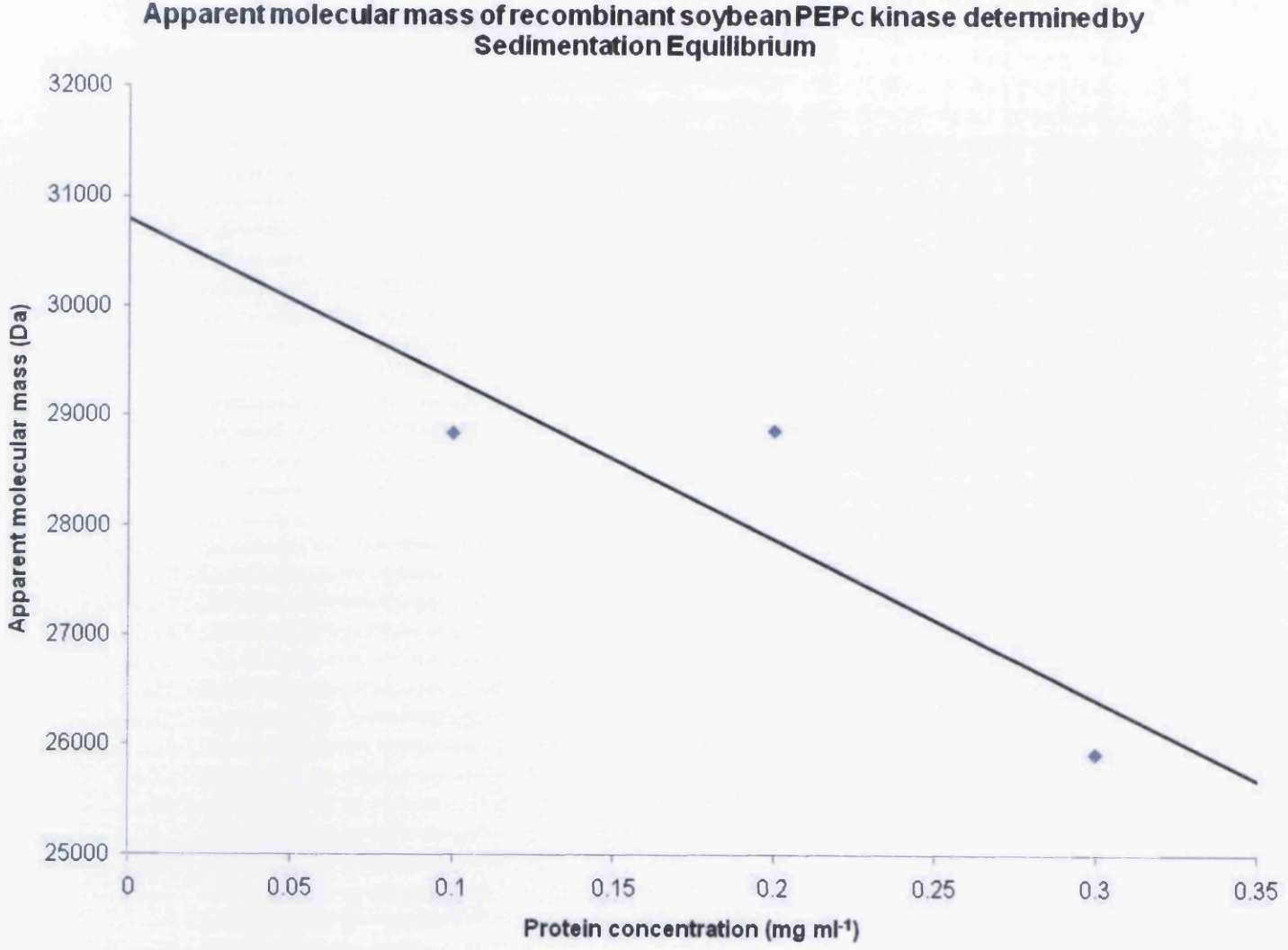
Figure 4.5 Sedimentation equilibrium ultracentrifugation of recombinant soybean PEPc kinase

Recombinant soybean PEPc kinase was expressed and purified as outlined in Chapter 3. The buffer was exchanged to 50 mM HEPES pH 7.4, 1 mM DTT, 150 mM NaCl using a HiPrep 26/10 column attached to a BioCAD 700E system. Protein was concentrated using a centrifugal concentration device and loaded into the analytical ultracentrifugation cells at 3 different protein concentrations: 0.1 mg ml^{-1}, 0.2 mg ml^{-1} and 0.3 mg ml^{-1}. The cells were spun at 20,000 rpm and scanned using interference optics every 3 h until equilibrium had been reached. The data was analysed using the SEDPHAT algorithm (A). Apparent molecular mass at each loading concentration was calculated and plotted on a graph (B). Extrapolation of the best fit straight line shows that the mass of recombinant soybean PEPc kinase at infinite dilution is ~31 kDa.

A


Plot of r² against ln (fringes)
4.00
2.00
0.00
-2.00
-4.00
-6.00
-8.00
ln (fringes)
48.50
49.00
49.50
50.00
50.50
51.00
51.50
52.00
r² (cm²)
0.3 mg ml-1
0.2 mg ml-1
0.1 mg ml-1

B


Apparent molecular mass of recombinant soybean PEPc kinase determined by Sedimentation Equilibrium
32000
31000
30000
29000
28000
27000
26000
25000
Apparent molecular mass (Da)
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
Protein concentration (mg ml-1)

an ÄKTA Explorer system (GE Healthcare). As a control, 100 µl of protein solution was applied to the same Superose 12 column at room temperature at a flowrate of 0.3 ml min^{-1} also using an ÄKTA Explorer system. The A_{280} of the effluents at both temperatures were monitored and overlayed (Fig 4.6A). It can be seen from the figure that there is no difference between the relative intensities of the two peaks.

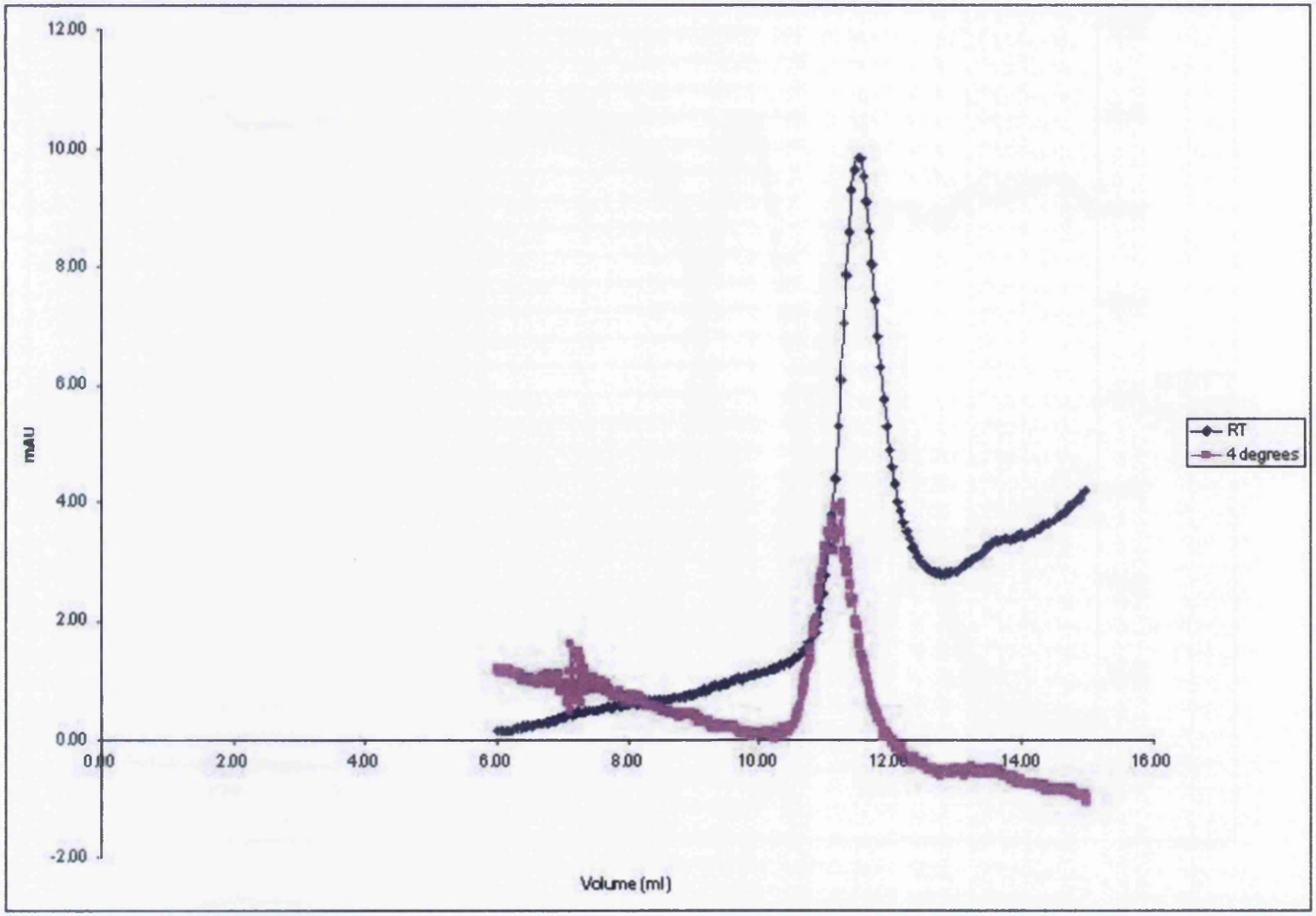
When using the ÄKTA Explorer system as opposed to the BioCAD 700E system for SEC analysis, it was noted that a small change in sample conductivity could be seen when the higher molecular weight peak eluted from the column (Fig 4.6B). However, there was no change in sample conductivity when the lower molecular weight peak eluted from the column. This is further evidence of a difference between the two molecular weight species. A possible explanation for the observed change in sample conductivity is offered in the Discussion section of this chapter.

To determine whether two peaks of A_{280} absorbance could be seen with other recombinant PEPc kinases, the SEC analysis experiments were repeated using purified recombinant Sorghum PEPc kinase and purified recombinant Arabidopsis PEPc kinase. Recombinant Sorghum PEPc kinase was expressed, purified and concentrated as outlined in Chapter 3. Recombinant Arabidopsis PEPc kinase was expressed, purified and concentrated as outlined in Chapter 5. 100 µl of each protein solution was applied to the Superose 12 column at room temperature and 4°C. As before, the A_{280} of the effluents were monitored and overlayed (Fig 4.6D). It can be seen that both recombinant Sorghum PEPc kinase and recombinant Arabidopsis PEPc kinase also show two peaks of absorbance and that this effect is temperature independent. As a final

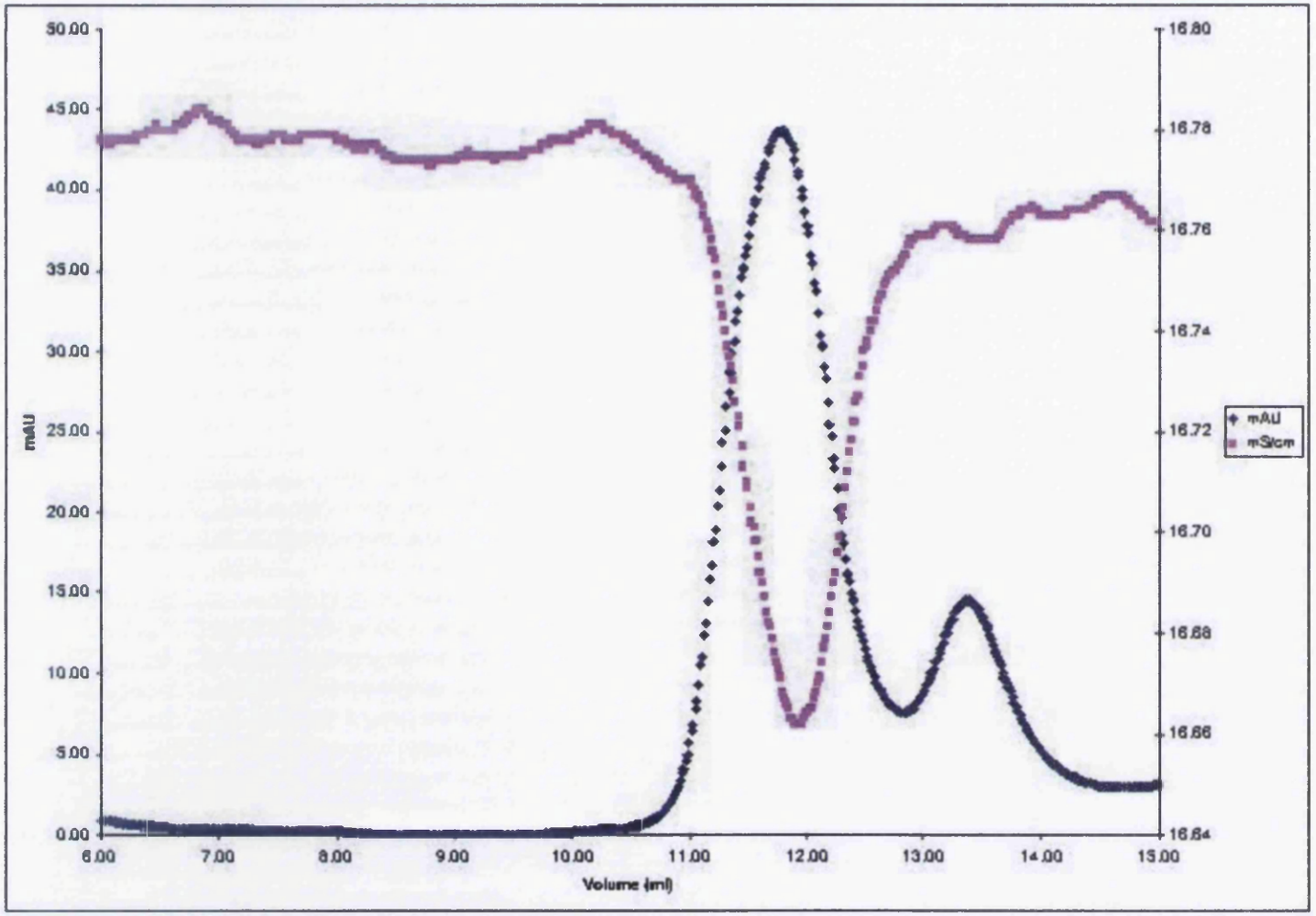
Figure 4.6 Further size exclusion chromatography of recombinant PEPc kinase

Recombinant PEPc kinase was expressed, purified and concentrated as outlined in the text. 100 µl of recombinant soybean PEPc kinase solution was loaded onto a pre-equilibrated Superose 12 column at a flowrate of 0.3 ml min^{-1} at 4°C and room temperature. The A_{280} of the effluents were monitored and overlayed (A). Conductivity was also monitored, and a small decrease in conductivity was seen when the larger molecular weight peak eluted from the column (B). The experiments were repeated with recombinant sorghum PEPc kinase and recombinant Arabidopsis PEPc kinase (C). Also, recombinant soybean PEPc kinase was chromatographed with and without ATP/Mg in the column buffer (D).

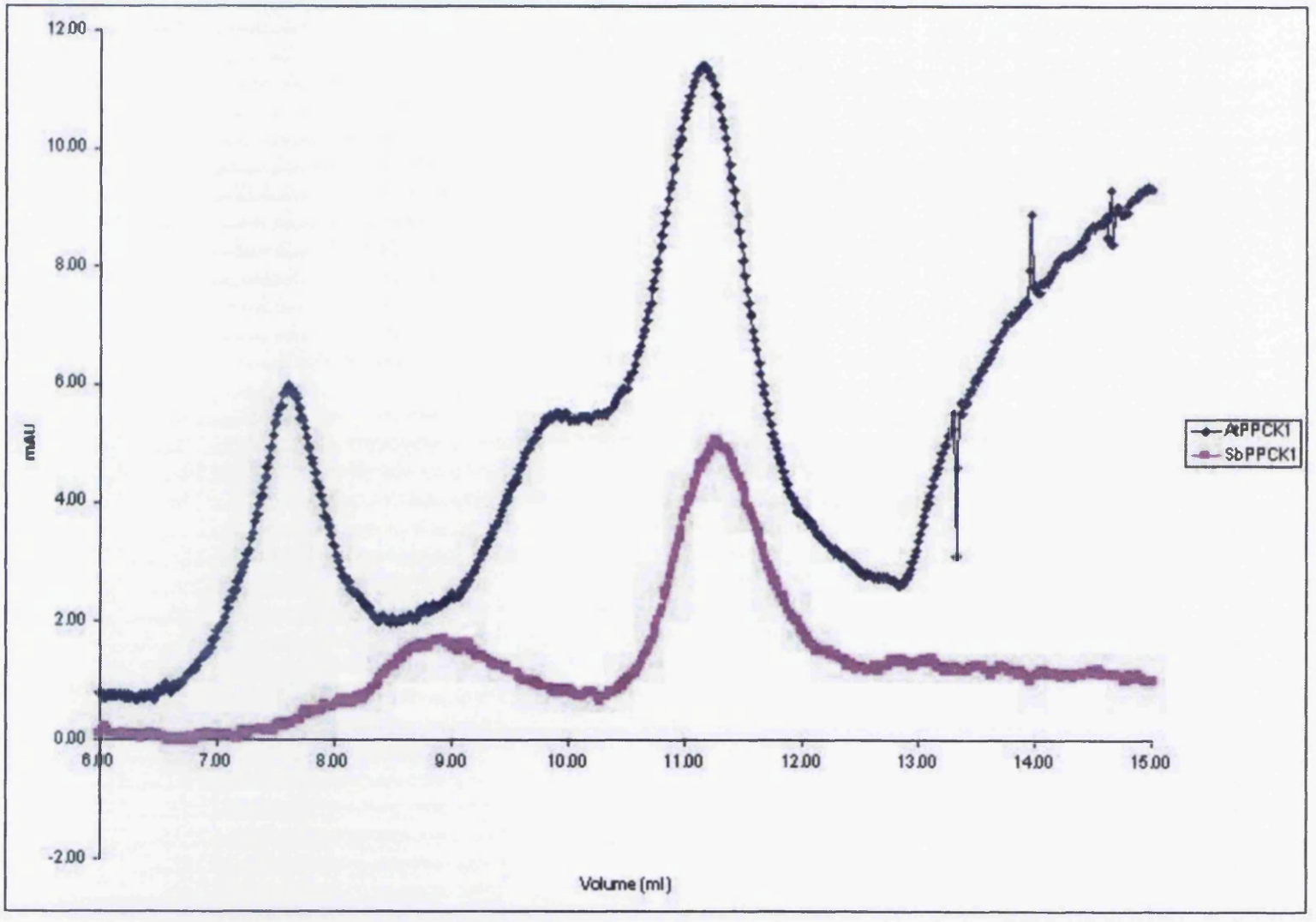
A


12.00
10.00
8.00
6.00
4.00
2.00
0.00
-2.00
mAU
0.00
2.00
4.00
6.00
8.00
10.00
12.00
14.00
16.00
Volume (ml)
RT
4 degrees

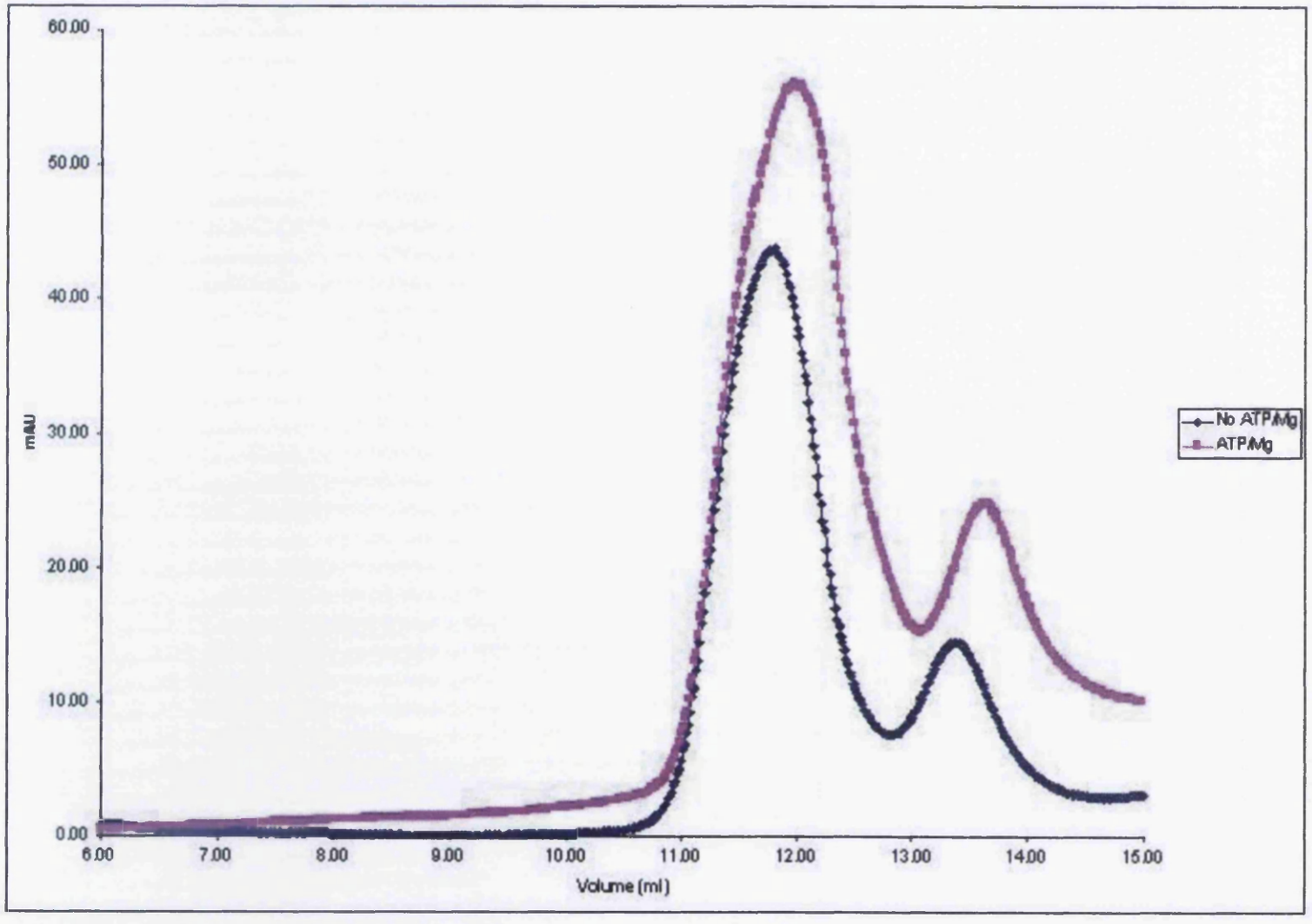
B


50.00
45.00
40.00
35.00
30.00
25.00
20.00
15.00
10.00
5.00
0.00
mAU
16.80
16.78
16.76
16.74
16.72
16.70
16.68
16.66
16.64
mS/cm
6.00
7.00
8.00
9.00
10.00
11.00
12.00
13.00
14.00
15.00
Volume (ml)
mAU
mS/cm

C


12.00
10.00
8.00
6.00
4.00
2.00
0.00
-2.00
mAU
6.00
7.00
8.00
9.00
10.00
11.00
12.00
13.00
14.00
15.00
Volume (ml)
AtPPCK1
SbPPCK1

D


60.00
50.00
40.00
30.00
20.00
10.00
0.00
mAU
6.00
7.00
8.00
9.00
10.00
11.00
12.00
13.00
14.00
15.00
Volume (ml)
No ATP/Mg
ATP/Mg

experiment, SEC analysis was carried out with ATP and magnesium in the column running buffer to determine whether their presence could influence the relative populations of recombinant soybean PEPc kinase molecules in each peak. Recombinant soybean PEPc kinase was expressed, purified and concentrated as before. The Superose 12 column was equilibrated with 50 mM HEPES pH 7.5, 1 mM DTT, 150 mM NaCl, 25 mM $MgCl_2$, 100 µM ATP at room temperature. 100 µl of protein solution was applied to the Superose 12 column at a flowrate of 0.3 ml min^{-1} and the A_{280} of the effluent was monitored. As a control, recombinant soybean PEPc kinase was applied to the same Superose 12 column at room temperature in the absence of ATP and magnesium, and the A280 traces were overlayed (Fig 4.6D). From the resulting chromatographs, it would appear that the presence of ATP and magnesium has no effect on the relative intensities of the two peaks.

4.2.6 Circular dichroism spectroscopy of PEPc kinase

X-ray crystallography was unsuccessful in determining the three dimensional structure of PEPc kinase, and recombinant PEPc kinase protein is not expressed at a high enough level in *P. pastoris* to justify heavy isotope labelling for heteronuclear NMR spectroscopy. Therefore, a lower resolution approach to determining the structure of recombinant PEPc kinase was taken using CD spectroscopy.

Recombinant soybean PEPc kinase was expressed and purified as outlined in Chapter 3. To facilitate far-UV CD studies, buffer was exchanged to 50 mM sodium phosphate pH 7.4, 1 mM DTT, 100 mM NaF using an centrifugal concentrator. Sodium fluoride was used instead of sodium chloride as chloride

ions give a signal in the far-UV CD spectrum. The final concentration of protein was determined by UV absorbance to be 0.12 mg/ml. The structure of the protein was then analysed in the far-UV using a spectropolarimeter in a 0.02 cm pathlength cell (Fig 4.7A). The acquired data was analysed by the CONTIN algorithm and a percentage of secondary structure for recombinant PEPc kinase was obtained (Fig. 4.7B). CONTIN estimates the percentage composition of secondary structure by comparing the experimental CD spectrum with the CD spectrum of proteins with known three-dimensional structure.

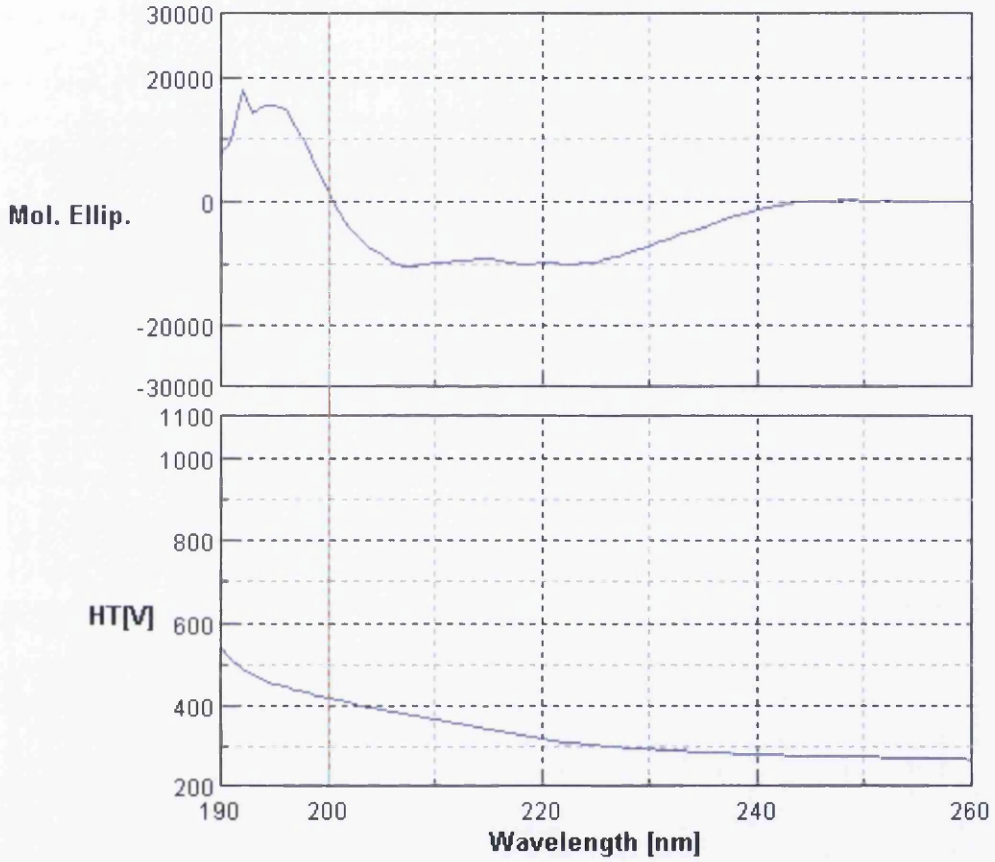
When the CD spectrum of recombinant PEPc kinase was analysed using reference set 4, the percentage composition of α-helix, β-sheet, turns and unordered polypeptide was 36.7%, 14.3%, 23.5% and 25.5% respectively. The root mean squared deviation for this analysis was 0.133. When the spectrum was analysed using reference set SP175, the percentage composition of α-helix, β-sheet, turns and unordered polypeptide was 32.4%, 17.0%, 13.5% and 37% respectively. However, the root mean squared deviation for this analysis was higher than the previous one, at 0.165.

The experimental percentage composition values were compared with the percentage composition values for homologous proteins that have had their three dimensional structures solved by X-ray crystallography. The primary structure of soybean PEPc kinase was entered into the Protein Data Bank search facility, to find closely related homologues with known structures. The closest match was the kinase domain of human calcium calmodulin dependent protein kinase type IIα (CAMK2A) in complex with indirubin E804 (PDB ID: 2VZ6), which has 39% identity with GmPPCK3. Secondary structure analysis of CAMK2A shows that the percentage composition of α-helix and β-sheet is 41% and 16%

Figure 4.7 Circular dichroism spectroscopy of recombinant soybean PEPc kinase

Recombinant soybean PEPc kinase was expressed, purified and concentrated as outlined in the text. The concentration of protein was determined by spectrophotometer and the far-UV CD spectrum of the protein was measured (A). Using the CONTIN algorithm, an estimation of secondary structure was obtained (B).

A


30000
20000
0
-20000
-30000
Mol. Ellip.
1100
1000
800
600
400
200
HT[V]
190
200
220
240
260
Wavelength [nm]

B

	α-helix	B-sheet	Turns	Unordered
GmPPCK3 (reference set 4)	36.7	14.3	23.5	25.5
GmPPCK3 (reference set SP)	32.4	17.0	13.5	37.0

respectively. This is slightly higher than the values obtained from the CD spectrum of GmPPCK3 when analysed using reference set 4.

4.3 Discussion

In this chapter, the solution characteristics of recombinant PEPc kinase have been investigated, using a combination of SEC and AUC. While SEC appeared to show the presence of oligomers in solution, no evidence for this could be found using AUC. *In silico* analysis was carried out to determine whether PEPc kinase could be classed as unstructured, and no evidence was found for this either. Finally, the structure of PEPc and PEPC kinase were analysed using CD spectroscopy. The results presented here will now be interpreted in light of the current literature.

SEC was chosen for initial analysis of the solution characteristics of recombinant PEPc kinase as it is a simple, reproducible method. One potential drawback of this method is that the protein may interact with the stationary phase, both via hydrophobic interactions and ionic interactions. The ionic strength of the mobile phase is therefore important, as a high ionic strength promotes hydrophobic interactions, while low ionic strength promotes ionic interactions. Generally 0.1-0.5 M NaCl is considered to be a suitable range of salt concentration for minimising interactions with the stationary phase in SEC, and a salt concentration at the lower end of this range was chosen since PEPc kinase is a particularly hydrophobic protein. Discrete peaks were resolved suggesting there is little interaction between the mobile and stationary phases in this case. Also, if non-specific aggregation of PEPc kinase occurred, a much lower elution volume would be expected.

When analysing proteins by SEC, it is assumed that the proteins that are being analysed are globular. However, if a protein is in an elongated conformation, it would have a larger excluded volume and elute earlier than a globular protein of the same molecular mass. Two peaks were consistently resolved when purified recombinant PEPc kinase was analysed by SEC - one peak corresponding to 30 kDa, and the other corresponding to 120 kDa. Pooling and re-chromatographing each peak appears to show that the "default" state of PEPc kinase is the 120 kDa form.

When each peak was assayed for the ability to phosphorylate PEPc, both peaks were found to contain PEPc kinase activity. However, the specific activity of the 120 kDa peak was much lower than that of the 30 kDa peak. Interestingly, depending on the dilution factor, the specific activity of the 30 kDa peak varied from 100 nmol min^{-1} mg^{-1} to 835 nmol min^{-1} mg^{-1}. This effect is similar to that seen in assays of PEPc kinase in plant leaf extracts (Nimmo *et al.*, 2001) but it is unlikely that an inhibitor of PEPc kinase is present in *Pichia pastoris* culture.

To summarise the SEC data, there are a number of key differences between the proteins found in the two peaks. First, and most obvious, the two peaks have different apparent molecular weights. Second, the two peaks have different specific activities, and while the 120 kDa peak has a specific activity that is independent of protein concentration, the 30 kDa peak specific activity varied with dilution factor. Third, elution of the 120 kDa from the Superose 12 column is associated with a small decrease in conductivity, while elution of the 30 kDa peak is not. Thus, an attractive hypothesis is that the oligomerisation state of PEPc kinase changes to regulate activity of the enzyme.

However, the analytical ultracentrifugation data provides no evidence to support this hypothesis. There is only one predominant species seen in the sedimentation velocity experiments, and this has a much lower sedimentation coefficient than the homology model would suggest. The sedimentation equilibrium data, the best indicator of molecular mass, shows no evidence for oligomerisation of recombinant PEPc kinase.

The homology model of soybean PEPc kinase, while similar in structure to other protein kinase domains, appears to have a cluster of hydrophobic residues on the surface of the molecule. This contradicts the principles of protein folding, where the hydrophobic residues are supposed to be buried in the inside of the molecule to maximise entropy.

CD spectroscopy was chosen as the method to investigate this phenomenon for a number of reasons. Firstly, it is a relatively facile procedure and the results can be obtained and analysed quickly. The protein does not have to be at as high a concentration as in X-ray crystallography or NMR, which is useful when solubility of protein is an issue. Also, the protein structure is examined in solution, unlike X-ray crystallography where the protein is in the solid state.

The use of CD spectroscopy has a number of potential drawbacks. It is a low resolution approach, and cannot be used as a substitute for the atomic resolution obtained using X-ray crystallography or NMR. It can only be used to estimate secondary structure and to make comparisons of tertiary structure under different conditions. Also, there are a number of interfering substances that give a signal in the far-UV CD, such as chloride ions and Tris buffer.

The data obtained from CD spectroscopy of recombinant soybean PEPc kinase in the far-UV showed a slightly lower composition of secondary structure compared with a homologous protein of known three-dimensional structure. The differences in α-helix and β-sheet were too small to be considered significant. Nevertheless, recombinant soybean PEPc kinase may still lack tertiary structure while maintaining secondary structural features. The protein concentration was too low to determine the near-UV CD spectrum of recombinant soybean PEPc kinase. However, absolute tertiary structure of a protein cannot be obtained from near-UV CD measurements anyway and near-UV CD is only useful in making comparisons between different conditions.

By examining all of the data obtained in this chapter, a picture of recombinant PEPc kinase in solution can be built. First, from the SEC data, it can be said that two forms of recombinant PEPc kinase exist in solution, with different elution volumes and different specific activities. From the AUC data, it can be said that the apparently larger form is not an oligomer of the apparently smaller form. From the CD spectroscopy data, recombinant PEPc kinase may adopt a loose tertiary structure while maintaining secondary structure. Thus, it is proposed that recombinant PEPc kinase exists in two folded conformations - a tight, compact, fully folded and fully functional state, and a loose, compact, partially folded state with little or no activity.

One important question is whether this "two state model" can be applied only to the recombinant PEPc kinase protein or if it can also be applied to the native protein in plant tissue. The specific activity of the 30 kDa form of recombinant PEPc kinase increases upon dilution and this is also true of PEPc kinase in desalted crude leaf extracts. It is possible that the loose form of PEPc

kinase has, until now, escaped detection due to its lack of activity. The main methods that are employed for purification of PEPc kinase (Phenyl Sepharose chromatography and Blue Dextran Agarose chromatography) select for fully folded protein. Also, if the progress of the purification was being followed by monitoring activity, any partially or non-active kinase would be discarded.

The "loose" conformation of PEPc kinase could be a key feature of its regulation. Being only partially folded may expose potential sites for proteases that would otherwise be hidden in the fully folded protein, and make it more susceptible to degradation. It may also expose sites for ligation of ubiquitin molecules, targeting the protein to the proteosome. It has been proposed that PEPc kinase is rapidly turned over. Therefore, if the protein predominantly existed in a protease-sensitive conformation, this would accelerate its degradation. This is further explored in Chapter 6.

The work in this chapter has characterised the properties of recombinant PEPc kinase in solution and has proposed that this protein exists in two folded states; this helps to explain why suitable crystals were not obtained from the crystal screens used in Chapter 3. The final results chapter of this thesis will examine the functional properties and kinetics of PEPc phosphorylation.

5 Kinetic anaiysis of the phosphorylation of PEPc

5.1 Introduction

In the previous results chapters, structural analysis of recombinant PEPc kinase was carried out, and it was found that the protein exists in two independent conformations. The aim of this chapter was to characterise the functional properties of PEPc kinase, and determine the kinetic parameters of the interaction between PEPc and PEPc kinase. Also, phosphorylation of a novel isoform of PEPc in rice was examined.

Studies of PEPc and PEPc kinase genes have shown that there is a small family of both genes in most plant species. For example, there are four PEPc genes and two PEPc kinase genes in Arabidopsis. Also, there are five PEPc genes and three PEPc kinase genes in rice. In both of these organisms the different PEPc and PEPc kinase genes have tissue-specific distributions, which leads to the question of whether there are specific PEPc/PEPc kinase pairs *in vivo*. One way to address this is to look for kinetic evidence of specificity. Thus, in this chapter the K_m values for different PEPc proteins of individual PEPc kinases were determined.

A novel PEPc that is localised to the chloroplast of rice has recently been discovered by the group of Mitsue Miyao. Further localisation studies showed that this isoform of PEPc is in the stroma of the chloroplast, while kinetic analysis showed that it has very similar properties to other plant-type isoforms of PEPc in terms of allosteric regulation. Analysis of the N-terminus of the

polypeptide shows that it contains the consensus PEPc kinase phosphorylation sequence. This raises the question of whether this isoform of PEPc can be phosphorylated by any or all of the rice PEPc kinases. While the *in vivo* site of phosphorylation would also be useful to know, a first step towards determining whether this sequence is phosphorylated is to study the system in vitro. Thus, an *in vitro* transcription/translation approach is taken to answer this question.

This chapter reports the determination of the kinetic parameters of the PEPc-PEPc kinase interaction. The phosphorylation of rice chloroplast PEPc was also demonstrated.

5.2 Results

5.2.1 Expression and purification of recombinant Arabidopsis PEPc kinase

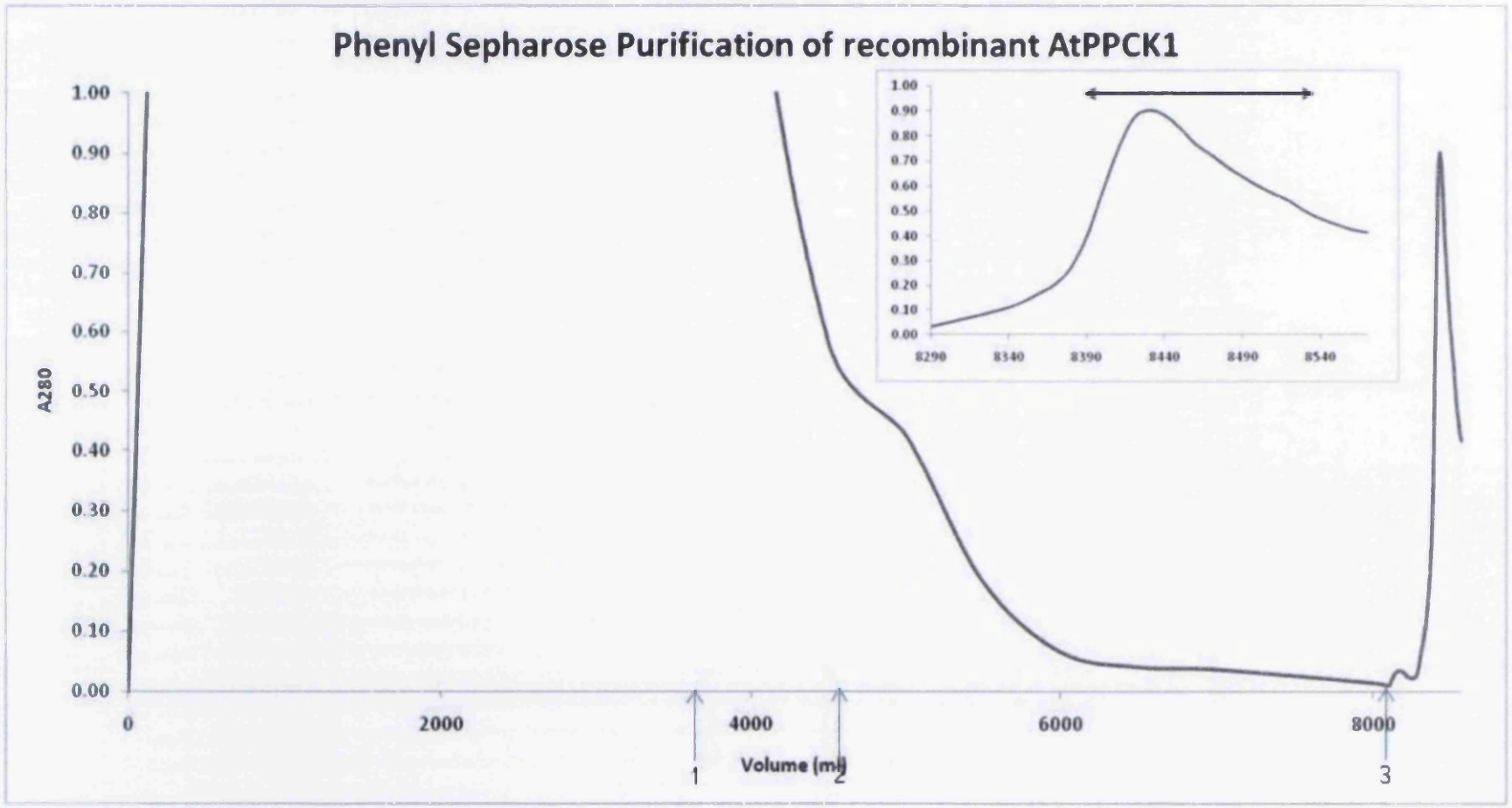
Due to the low abundance of PEPc kinase *in planta*, recombinant PEPc kinases were used to determine the kinetic parameters of the interaction between PEPc and PEPc kinase. Expression and purification of recombinant soybean and Sorghum PEPc kinases was previously outlined in Chapter 3. Here, expression and purification of recombinant Arabidopsis PEPc kinase is described.

Pichia pastoris cells containing the Arabidopsis PPCK1 cDNA were cultured from a frozen glycerol stock as outlined in Chapter 2. Expression was induced by growth in BMMY medium for 3 days, after which the supernatant was harvested. This was applied to a 200 ml Phenyl Sepharose column, the column was washed with 30% (v/v) ethylene glycol buffer and protein eluted with 60% (v/v) ethylene glycol buffer (Fig. 5.1A). The eluate from the Phenyl Sepharose column was then applied to a 10 ml Ni-NTA agarose column, the column washed with 70 mM

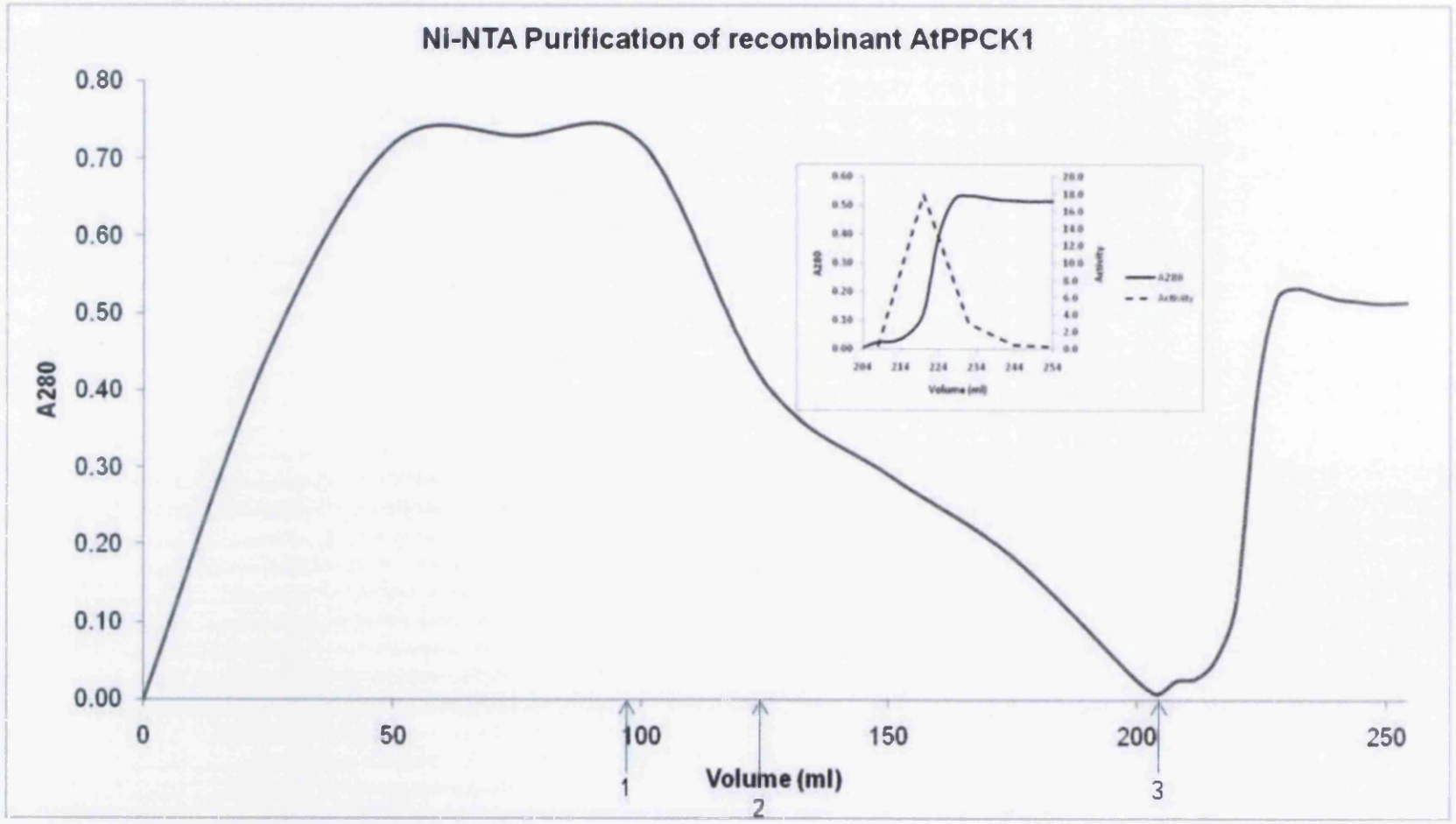
Figure 5.1 Purification of recombinant Arabidopsis PPCK1

Recombinant Arabidopsis PEPc kinase was expressed in four litres of *P. pastoris* cell culture as outlined in the text. The supernatant was applied to a 200 ml column of Phenyl Sepharose pre-equilibrated with Buffer A, at a flowrate of 2 l h^{-1} (A). The column was then washed with 500 ml Buffer A, and 5 litres Buffer B containing 30% (v/v) ethylene glycol. The protein was eluted with Buffer B containing 60% (v/v) ethylene glycol at a flowrate of 5 ml min^{-1}. The pooled fractions are indicated by the black horizontal line (inset). The pooled fractions from Phenyl Sepharose were then loaded onto a 10 ml Ni-NTA agarose column at a flowrate of 5 ml min^{-1} (B). The column was washed with 50 ml of 10 mM imidazole buffer and then with 70 mM imidazole buffer until the A_{280} of the effluent fell to <0.01. The protein was eluted with 500 mM imidazole buffer. Fractions were assayed for PEPc kinase activity as outlined in Chapter 2. PEPc kinase activity is given by the dashed line (inset). The purification scheme was analysed and is given in (C).

A


Phenyl Sepharose Purification of recombinant AtPPCK1
A280
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0
2000
4000
6000
8000
Volume (ml)
1
2
3
8290
8340
8390
8440
8490
8540

B


Ni-NTA Purification of recombinant AtPPCK1
A280
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0
50
100
150
200
250
Volume (ml)
1
2
3

C

Step	Volume (ml)	Total Protein (mg)	Total Enzyme Activity (units)	Specific Activity (units/mg)	Yield (%)	Purification (n-fold)
Supernatant	3480	365.40	1388.55	3.80	100	1.00
Phenyl Sepharose	96	12.96	67.36	5.20	5	1.37
Ni-NTA	17.5	0.61	21.82	35.77	2	9.41

imidazole buffer and the protein eluted with 500 mM imidazole buffer (Fig 5.1B). The supernatant and pooled fractions from each column were assayed for PEPc kinase activity, and the purification scheme assessed (Fig. 5.1C). When compared with the purification scheme for soybean PEPc kinase, the yield of recombinant Arabidopsis PPCK1 was low (2%).

5.2.2 Determination of the kinetic parameters of the PEPc-PEPc kinase interaction

The kinetic parameters of the interaction between PEPc and PEPc kinase were determined by a radiometric kinase assay. This requires purified enzymes and radiolabelled ATP, which are incubated together for a set amount of time. The unincorporated radioactivity is separated from the incorporated radioactivity by gel electrophoresis and the amount of radioactivity incorporated is quantified by phosphorimaging.

Maize and *Kalanchoë fedtschenkoi* PEPc enzymes were isolated from plant leaves as outlined in Chapter 2. Recombinant rice chloroplast PEPc was expressed and purified as outlined later in this chapter. Expression and purification of recombinant soybean and Sorghum PEPc kinases was described in Chapter 3, while expression and purification of recombinant Arabidopsis PEPc kinase was described earlier in this chapter.

First, the K_m for ATP of each recombinant PEPc kinase was measured using *K. fedtschenkoi* PEPc as substrate. Kinase assays were carried out as outlined in Chapter 2. The SDS-PAGE gel was stained using Coomassie blue to check for equal loading of PEPc, and then the gel was dried and phosphorimaged (Fig. 5.2A). The Eadie-Hofstee plots for ATP of each recombinant PEPc kinase are

shown in Fig. 5.2B. The K_m values for ATP of each recombinant PEPc kinase are reported in the table in Fig. 5.2C. When using ATP concentrations in excess of the K_m value, the initial velocity decreased as the ATP concentration increased (data not shown). This is indicative of substrate inhibition.

Next, the K_m for PEPc of each recombinant PEPc kinase was measured using each PEPc as substrate. Kinase assays were carried out as before using 50 μM ATP. The reactions were started by addition of PEPc, terminated after 10 min by addition of SDS-PAGE sample loading buffer and boiled for 3 min. The samples were electrophoresed on an 8% SDS-PAGE gel with a 3% stacking gel until the dye front had left the bottom of the gel. The gel was stained and destained as before, dried and then exposed to a phosphorimage plate for 24 h. The resulting Eadie-Hofstee plots are shown in Fig. 5.3A while the K_m values are shown in Fig. 5.3B.

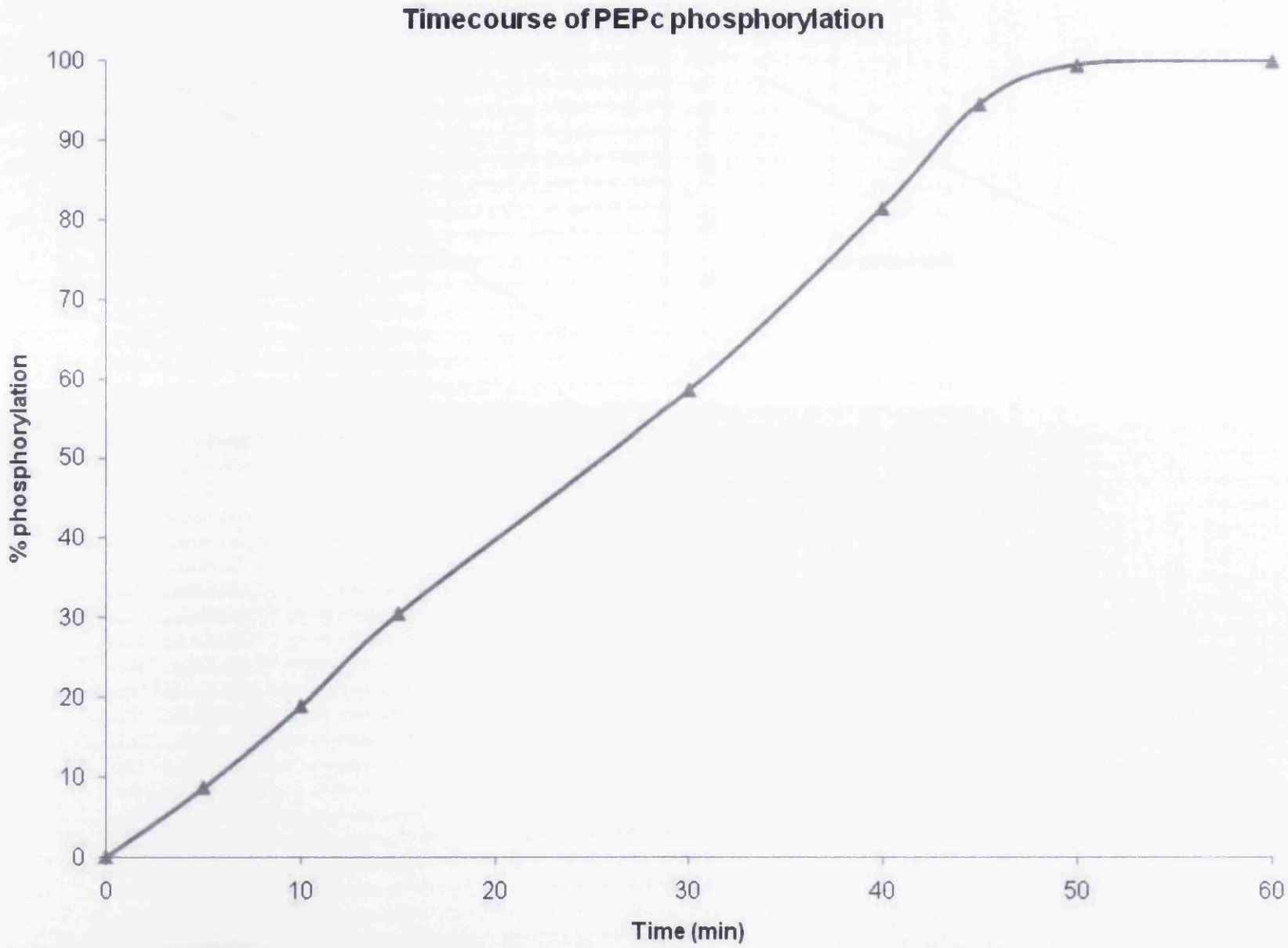
5.2.3 Phosphorylation of recombinant rice chloroplast PEPc

pET-28a vector containing the rice chloroplast PEPc (OsPpc4) cDNA was obtained from Mitsue Miyao (Matsumoto *et al*, 2007) and transformed into E. coli BL21 (DE3) cells using standard techniques. A single colony was used to grow an overnight culture of cells in LB medium, which was then used to inoculate 5 litres of fresh LB. This culture was grown to an OD_{600} = 0.6, after which expression of OsPpc4 was induced by addition of IPTG to a final concentration of 1 mM. After 3 h at 37°C, cells were harvested, snap frozen in liquid nitrogen and stored at -80°C until required for purification.

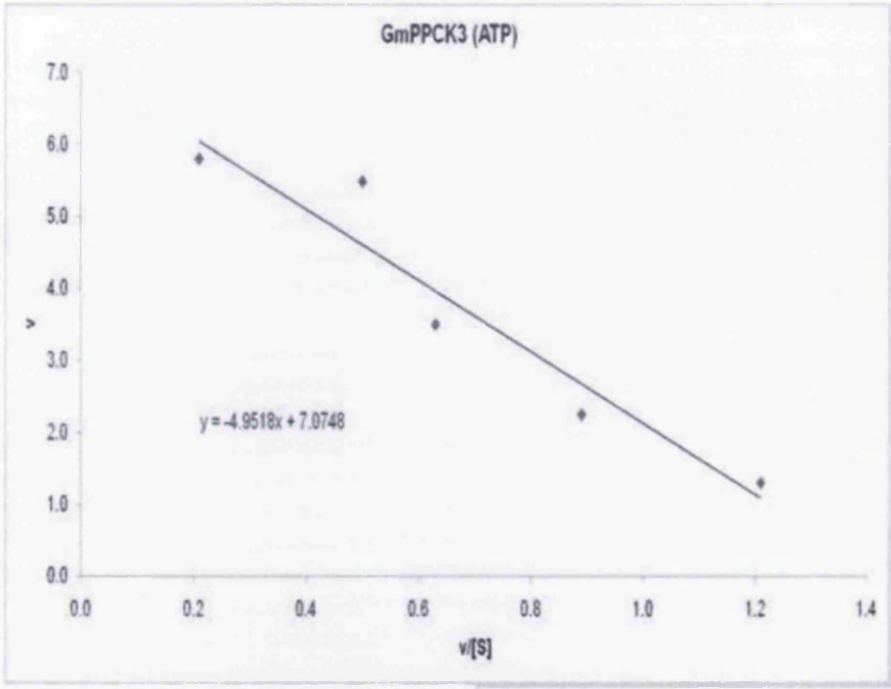
Figure 5.2 Determination of the K_m for ATP of recombinant PEPc kinases

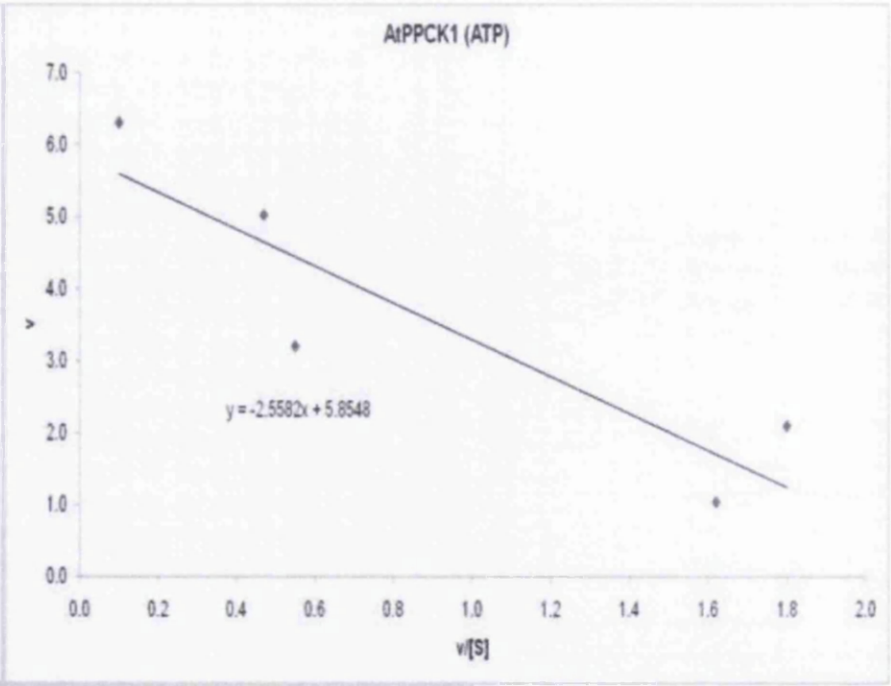
Recombinant PEPc kinases were expressed and purified as outlined in the text. The rates of phosphorylation of *Kalanchoe fedtschenkoi* PEPc at varying ATP concentrations were measured using a radiometric kinase assay. Incorporated ^{32}P was separated from unincorporated by SDS-PAGE. The assay was checked for linearity using recombinant soybean PEPc kinase (A). The data were analysed using Eadie-Hofstee plots (B) and a K_m value for ATP was determined by obtaining the slope of the best fit straight line (C).

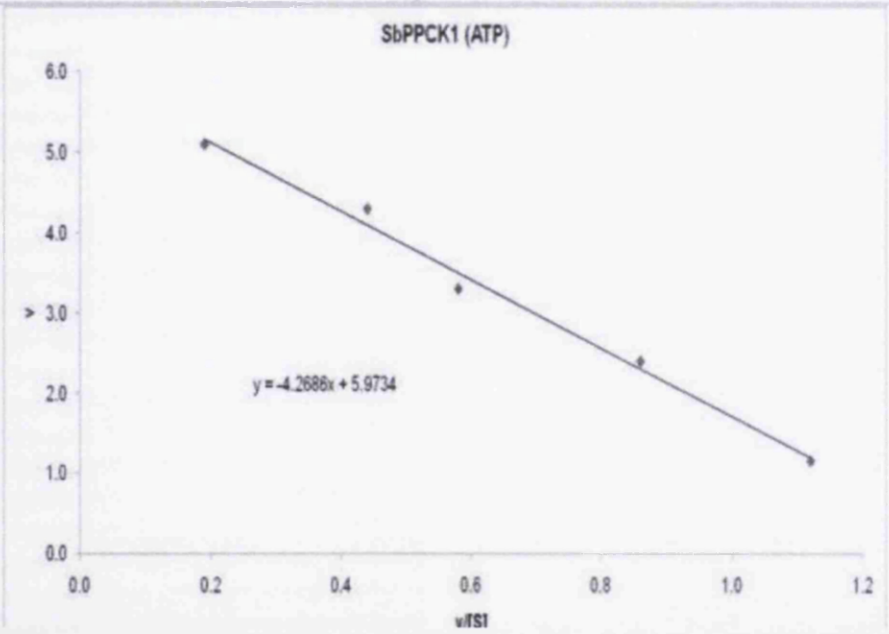
A


Timecourse of PEPc phosphorylation
% phosphorylation
100
90
80
70
60
50
40
30
20
10
0
0
10
20
30
40
50
60
Time (min)

B


GmPPCK3 (ATP)
v
v/[S]
y = -4.9518x + 7.0748
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4


AtPPCK1 (ATP)
v
v/[S]
y = -2.5582x + 5.8548
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0


SbPPCK1 (ATP)
v
v/[S]
y = -4.2686x + 5.9734
6.0
5.0
4.0
3.0
2.0
1.0
0.0
0.0
0.2
0.4
0.6
0.8
1.0
1.2

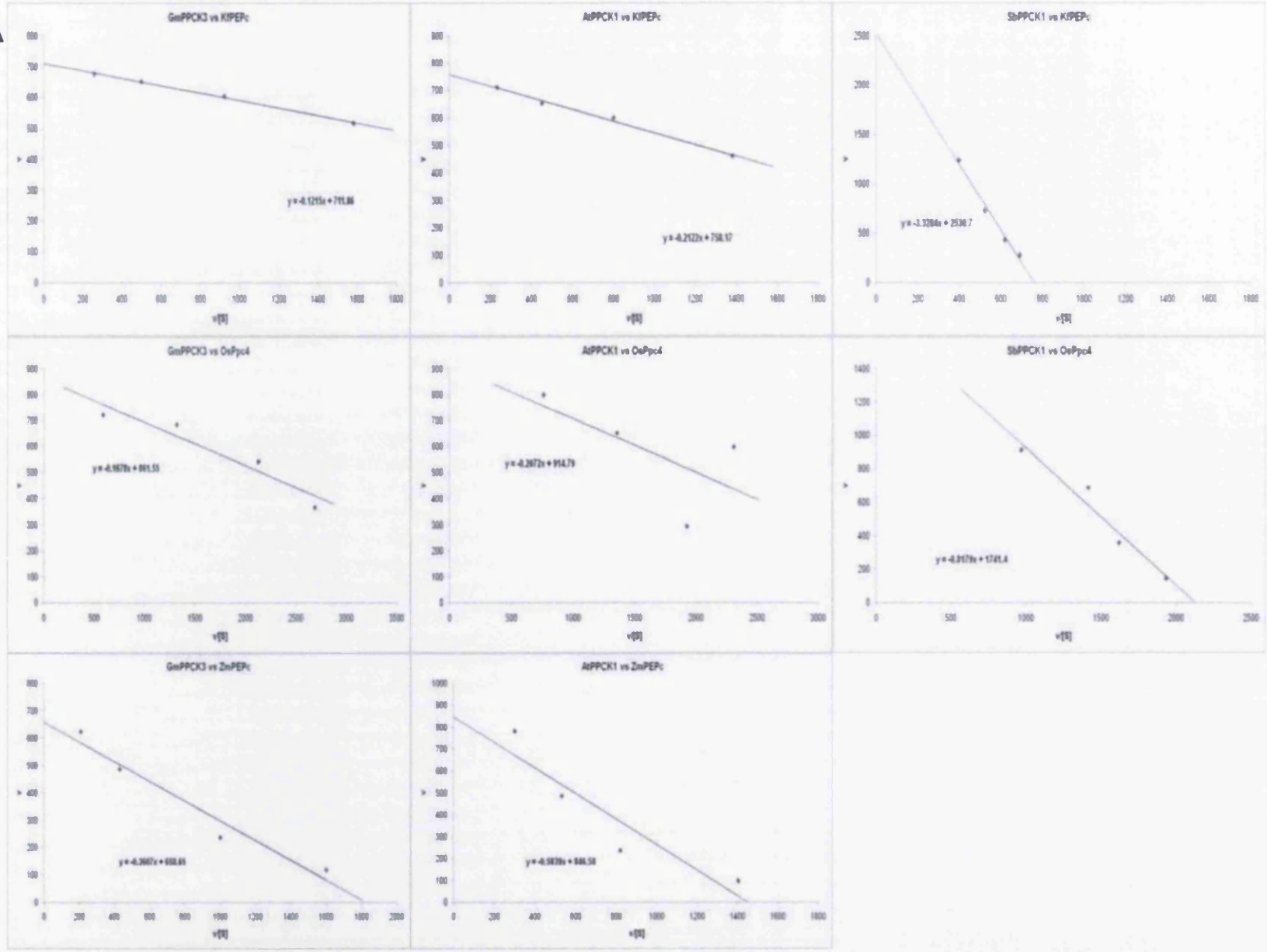
C

PEPc kinase	K_m for ATP (μM)
GmPPCK3	4.95
SbPPCK1	4.27
AtPPCK1	2.56

Figure 5.3 Determination of the K_m for PEPc of recombinant PEPc kinases

Recombinant PEPc kinases and isoforms of PEPc were purified as outlined in the text. The concentration of PEPc was determined by Bradford assay, and the rates of phosphorylation in the presence of increasing concentrations of PEPc were measured by radiometric kinase assays. Eadie-Hofstee plots of the data were obtained (A) and the slope of the best-fit straight lines used to calculate the K_m for PEPc (B).

A


GmPPCK3 vs KfPEPc
y = -0.1215x + 711.86
AtPPCK1 vs KfPEPc
y = -0.2122x + 758.17
SbPPCK1 vs KfPEPc
y = -3.3284x + 2536.7
GmPPCK3 vs OsPpc4
y = -0.1837x + 891.55
AtPPCK1 vs OsPpc4
y = -0.2672x + 914.79
SbPPCK1 vs OsPpc4
y = -0.9173x + 1741.4
GmPPCK3 vs ZmPEPc
y = -0.3667x + 658.65
AtPPCK1 vs ZmPEPc
y = -0.5839x + 846.58
v[S]

B

PEPc kinase isoform	K_m for PEPc (μM)		
	Kalanchoe	Maize	Rice chloroplast
GmPPCK3	0.12	0.36	0.17
SbPPCK1	3.32	-	0.82
AtPPCK1	0.21	0.58	0.28

Purification of OsPpc4 was carried out using immobilised metal affinity chromatography on Ni-NTA agarose and ion exchange chromatography using a Mono Q column as outlined in Chapter 2. While extra bands of protein are seen in the Coomassie blue analysis of the kinase assay gel, these bands co-elute at the same intensity on a salt gradient from the Mono Q column and are also recognised by anti-PEPc antiserum (data not shown).

To answer the question of whether rice chloroplast PEPc (OsPpc4) can be phosphorylated by rice PEPc kinases (OsPPCKs), an *in vitro* transcription/translation approach was taken. This approach has been successfully employed by Fukayama *et al.* (2006) to determine whether OsPPCK genes encode functional proteins.

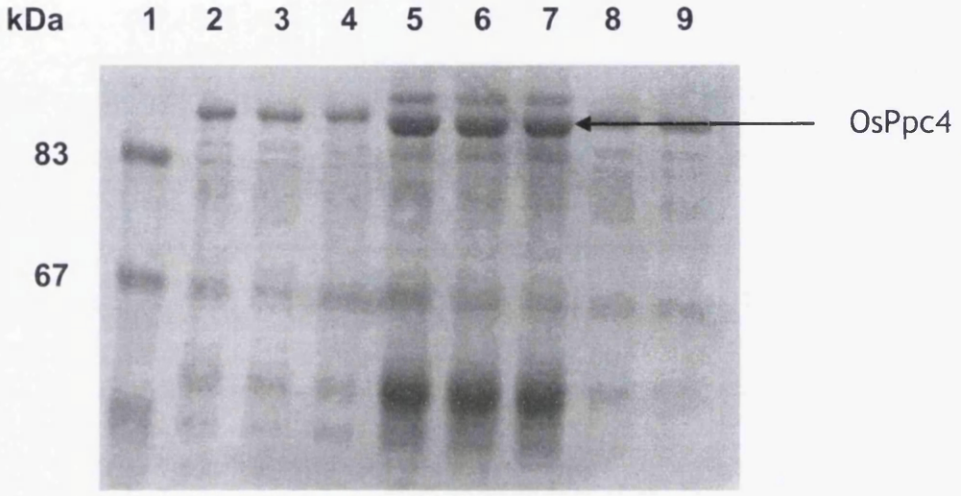
DNA plasmids containing the OsPPCK genes (OsPPCK1, OsPPCK2, OsPPCK3) were linearised completely using EcoRI. After the plasmids had been gel purified, they were transcribed *in vitro* using T3 RNA polymerase. The resulting mRNA was analysed by denaturing agarose gel electrophoresis for integrity and, using rabbit reticulocyte lysate supplemented with [^{35}S]-methionine, the OsPPCK mRNAs were used to direct synthesis of OsPpck proteins. Incorporation of phosphate into recombinant OsPpc4 by these proteins was observed (Fig. 5.4A); however, the strong band on the phosphorimage did not correspond to the major band on the Coomassie blue stained gel. This led us to believe that a degradation product of OsPpc4 was being phosphorylated. To investigate this further, a sample of the recombinant OsPpc4 protein was treated with thrombin for 3 h (since there is a thrombin cleavage site between the his$_6$-tag and the protein), and then re-purified on Ni-NTA. The protein in the flowthrough was collected and used as substrate in an *in vitro* kinase assay, using the translation

Figure 5.4 *In vitro* phosphorylation of recombinant OsPpc4

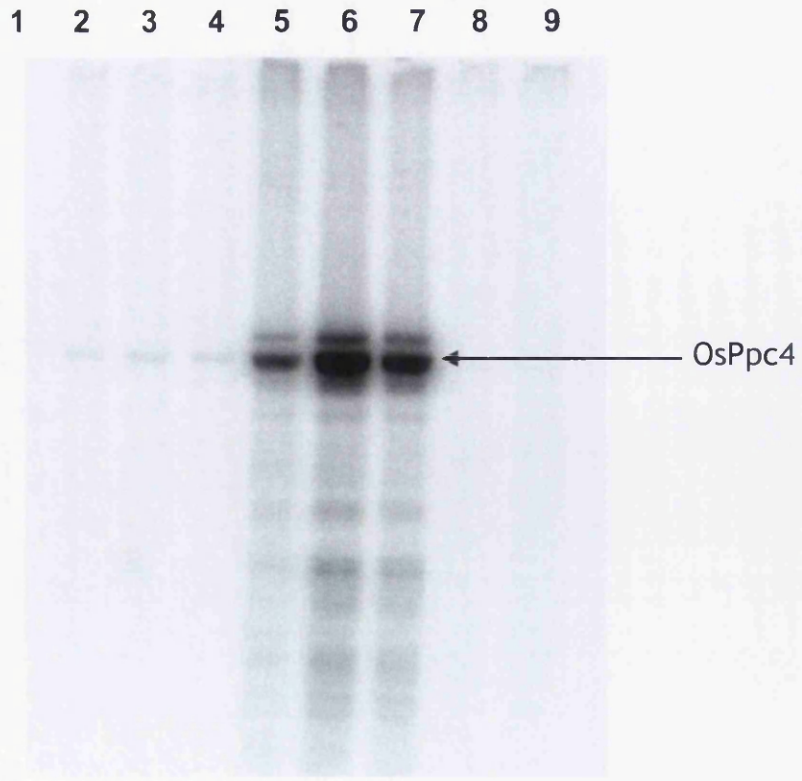
Recombinant OsPpc4 was expressed and purified as outlined in the text. Plasmids containing the OsPPCK cDNAs were linearised and mRNA synthesis was carried out using the T3 promoter. The produced mRNA was analysed for integrity by denaturing agarose gel electrophoresis. Protein synthesis was carried out using rabbit reticulocyte lysate and the translation products were assayed for PEPc kinase activity as outlined in the text (A). Lane 1 = markers; lane 2 = translated OsPpck1 mRNA minus OsPpc4; lane 3 = translated OsPpck2 mRNA minus OsPpc4; lane 4 = translated OsPpck3 mRNA minus OsPpc4; lane 5 = translated OsPpck1 mRNA plus OsPpc4; lane 6 = translated OsPpck2 mRNA plus OsPpc4; lane 7 = translated OsPpck3 mRNA plus OsPpc4; lane 8 = translated control RNA minus OsPpc4; lane 9 = translated control DNA minus OsPpc4. Phosphorylation of OsPpc4 in the presence and absence of the his_6-tag was also investigated (B).

A

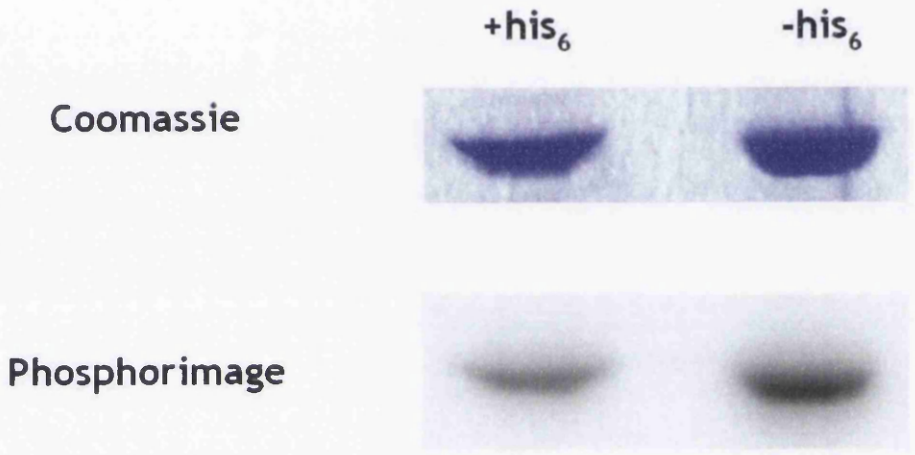
Coomassie



Phosphorimage



B



products of the *in vitro* transcription reaction of OsPPCK2 cDNA. It can be seen that incorporation of ^{32}P into the non-tagged OsPpc4 is higher than in the tagged OsPpc4 protein (Fig 5.4B). This suggests that the his_6-tag is blocking phosphorylation of PEPc in some way.

5.3 Discussion

In this chapter functional aspects of phosphorylation of PEPc by PEPc kinase were explored. The Michaelis constant for ATP of each recombinant PEPc kinase was measured. Also, the Michaelis constant for different PEPc isoforms of each recombinant PEPc kinase was determined.

The first part of this chapter described expression and purification of recombinant Arabidopsis PPCK1. Examination of the primary structure of Arabidopsis PPCK1 led us to believe that it has similar biophysical properties as soybean PEPc kinase and would also require 60% (v/v) ethylene glycol to elute the protein from the Phenyl Sepharose column. Yet the yield of recombinant Arabidopsis PPCK1 was very low (2%) and most of the activity was lost at the Phenyl Sepharose stage. Thus, it is feasible that there are differences in surface charge between Arabidopsis and soybean PEPc kinase, and this leads to different binding behaviour on hydrophobic interaction columns.

The K_m values for ATP of each recombinant PEPc kinase were determined using *Kalanchoë fedtschenkoi* PEPc as substrate. Generally, kinases have a K_m value of 5-25 μM for ATP and the values for recombinant PEPc kinases were at the low end of this range. Initial experiments to determine the range of ATP concentrations to use for K_m determination showed that recombinant PEPc kinase is susceptible to substrate inhibition. One explanation of high substrate

inhibition is that the active site of an enzyme has got two or more subsites for the substrate. For example, ATP could bind to the enzyme via the adenine ring or by the phosphate groups. It would be possible for two ATP molecules to bind, one via each subsite, and this would clearly cause high substrate inhibition.

Radiometric kinase assays were used to determine the value of the Michaelis constants for both substrates of recombinant PEPc kinase. Radiometric kinase assays are considered to be the "gold standard" method for assaying kinase activity as they are highly sensitive and a large number of samples can be processed at one time (Hastie *et al.*, 2001). However, radiometric kinase assays are an example of a discontinuous assay *i.e.* the assay is terminated before being analysed. The preferred method of enzyme analysis is a continuous assay, which allows enzyme activity to be monitored directly. This can be achieved spectrophotometrically by coupling the reaction to another reaction, which causes a change in light absorbance. In this case, a continuous assay cannot be done because it is not sensitive enough.

In this chapter, three different PEPc isoforms were used: a C_3 isoform (rice), a C_4 isoform (maize) and a CAM isoform (*Kalanchoë fedtschenkoi*). Comparison of the K_m values for each PEPc for each individual recombinant PEPc kinase suggests that there is no kinetic evidence for specific PEPc-PEPc kinase pairs. Also, there is no appreciable difference between the K_m values for PEPc of recombinant sorghum PEPc kinase compared with the other recombinant PEPc kinases. This would suggest that the acidic insert of amino acids does not play a direct role in substrate recognition.

This chapter has also described phosphorylation of a novel PEPc isoform discovered in the chloroplast of rice. Currently, rice is the only plant species

where PEPc has been found in the chloroplast. This could be due to the fact that the genome of rice has been sequenced and that there are other plant species with a chloroplast PEPc that have not had their genomes sequenced yet. For example, it is possible that all grasses might have a chloroplastic PEPc. The experiments in this chapter have shown that rice chloroplast PEPc can be phosphorylated by rice PEPc kinases. This then leads to the question of whether phosphorylation of this isoform of rice PEPc can occur inside the chloroplast and whether rice PEPc kinases can also be targeted to the chloroplast. Currently there is no evidence for localisation of PEPc kinases to the chloroplast of rice, but this requires further study. A possible further experiment would be to isolate chloroplasts from rice cells, supply them with $^{32}P_i$ and then use phosphoproteomics to identify phosphorylation of OsPpc4. However, it is possible that only some OsPpc4 is actually localised to the chloroplast and that the remainder is phosphorylated in the cytoplasm.

The final chapter of this thesis will summarise the findings of the results chapters and put these in the context of the current literature.

6 General Discussion

6.1 Introduction

The three results chapters in this thesis have reported stuctural and functional characterisation of recombinant PEPc kinase. Chapter 3 reported a general purification scheme and the failure to obtain diffraction quality crystals of recombinant PEPc kinase. Chapter 4 explored why recombinant PEPc kinase failed to crystallise showing that the protein exists in two different conformations *in vitro*. Investigation of the structural properties of the recombinant protein by circular dichroism spectroscopy estimated that secondary structure was still intact. Chapter 5 looked at the functional characteristics of phosphorylation by PEPc kinase, and demonstrated phosphorylation of a rice chloroplast isoform of PEPc. The conclusions of this thesis are now interpreted more broadly than in the Discussion sections of the results chapters.

6.2 Crystallisation of PEPc kinase

In chapter 3, a purification scheme for recombinant PEPc kinase was established and conditions which promote crystallisation of PEPc kinase were investigated. The purification scheme made use of the unusually high hydrophobicity of the protein to separate it from the bulk of proteins in the *P. pastoris* culture supernatant. Then, the decahistidine tag was used in an affinity purification step, and the resulting protein was homogeneous on an SDS-PAGE gel. The pure protein was used in crystallisation trials in the presence and

absence of the ATP analogue AMP-PNP. The crystal trials were largely unsuccessful, and only one "hit" was generated. Optimisation of this condition was also unsuccessful, and during optimisation it was found that the original condition could not be replicated.

There are a number of reasons why some proteins fail to crystallise. If there are flexible regions in the protein, they may hinder crystal packing. Recombinant PEPc kinase has two recombinant "tags" at the N-terminal end of the polypeptide chain. While these tags facilitated purification and identification of the protein, they also added to the disorder in this region. If this region is highly mobile in solution, then crystallisation may not occur.

Another reason why some proteins do not crystallise is that they are not monodisperse and that there are regions of microheterogeneity. While a protein may be determined to be pure in terms of migrating as a single band on an SDS-PAGE gel, it may not be pure in terms of quaternary structure. At first, it was assumed that recombinant PEPc kinase was homogeneous due to a sample of the protein migrating as a single band on SDS-PAGE and the initial success of crystallisation. SEC analysis later proved that this was not the case, and that there was heterogeneity in the protein solution.

Further work is therefore needed to develop better methods for the expression and purification of PEPc kinases. These should include the use of vectors which introduce proteolytic cleavage sites between the tag(s) and the kinase, and investigation of other expression systems such as insect cells. If as the data in Chapter 4 suggest, all PEPc kinases can exist in two (or more) conformations, crystallisation trials may prove unsuccessful until conditions are found that stabilise the enzyme in one conformation. Hence, NMR may prove a

more useful technique than crystallography for solving the structure of PEPc kinase. However, this will only be realised if a significantly better expression system is found.

6.3 The structure of PEPc kinase in solution

In chapter 4, the solution characteristics of recombinant PEPc kinase were analysed using SEC and AUC. The initial SEC data appeared to show that PEPc kinase was not monodisperse, and two discrete peaks were resolved. This was thought to be due to oligomerisation which was further investigated by AUC. Yet, no evidence was found for oligomerisation and an alternative explanation was proposed: that recombinant PEPc kinase exists in two differently folded states.

Viewed at a basic level, proteins exist in two states of folding: a fully folded, fully functional state; or a fully unfolded state that lacks function. However, experiments on protein folding have shown the picture to be more complex. Certain proteins exist in partially folded states under particular conditions, and these intermediates can be partially active or completely inactive. It appears from the anomalies between the SEC data and the AUC data that PEPc kinase exists in a partially folded state under the conditions examined.

Protein kinases must have a structurally conserved "on" state, as they all carry out the same reaction (the transfer of the γ-phosphate of ATP to a hydroxyl group on the target protein). However, their "off" state allows for much more flexibility and here it is proposed that the "off" state of PEPc kinase is a loosely folded conformer of the properly folded protein. Using NMR, the high resolution structure of the protein in solution could be examined. This would

provide unequivocal evidence for conformational changes proposed in this thesis.

Is there any reason why PEPc kinase might exist in a less structured form in the cell? This may be a mechanism of minimising aberrant kinase activity. PEPc kinase is simply a catalytic protein kinase domain with no intrinsic regulatory sequences upstream or downstream. Correct phosphorylation is vital to the maintenance of an organism and if the default state of the protein is an inactive conformation, then this would protect the organism from phosphorylating the wrong targets. Also, if PEPc kinase is only partially folded *in vivo*, this may confer substrate specificity. Rather like an induced fit mechanism, it is plausible that PEPc is the only protein in the cell with the correct protein-protein interaction surfaces to induce the folding of PEPc kinase from the inactive form to the active form. This could explain the observation in Chapter 1 that short peptides are very poor substrates for PEPc kinase (Li *et al.*, 1996). Furthermore, the two conformation model offers a plausible explanation for the mode of action of the PEPc kinase inhibitor protein reported by Nimmo *et al.* (2001). A protein that bound selectively to the loosely folded, low activity conformer of PEP kinase would act as an inhibitor.

Future work could involve solving the structures of the two different conformations by X-ray crystallography or NMR. Comparisons of the two structures could reveal the residues involved in modulating the transitions between the loose and compact conformations. This could then be further explored using mutagenesis to lock the enzyme in its fully active state. If this were successful, it would open an approach to one of the general problems that arise in studies of regulatory enzymology: to what extent are observations and

mechanisms observed *in vitro* relevant *in vivo*? In this case, it could be envisaged that introducing a mutant PEPc kinase that was constitutively correctly folded and active into a plant might result in the aberrant phosphorylation of proteins other than PEPc; this would be detectable via phosphoproteomics.

6.4 Kinetics of phosphorylation by PEPc

The kinetic experiments in Chapter 5 showed that all of the PEPc kinases tested were able to phosphorylate all of the PEPc enzymes that were available. This argues against the view that particular PEPc kinases might only be able to phosphorylate some PEPcs. It is of course possible that particular kinases are responsible for particular PEPc phosphorylation events *in vivo* by virtue of tissue specificity of expression. Ideally all the kinases from one plant should be tested against all the PEPcs from the same plant, but the required expression clones are not yet available.

Clearly, there is further work to be done on investigating aspects of the structure and function of PEPc kinase.

7 References

Blonde JD, Plaxton WC (2003) Structural and kinetic properties of high and low molecular mass phosphoenolpyruvate carboxylase isoforms from the endosperm of developing castor oilseeds. *Journal of Biological Chemistry* **278** pp11867-11873.

Bradford MM (1976) A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72** pp248-254.

Chen ZH, Nimmo GA, Jenkins GI, Nimmo HG (2007) BHLH32 modulates several biochemical and morphological processes that respond to Pi starvation in Arabidopsis. *Biochemical Journal* **405** pp191-198.

Chollet R, Vidal J, O'Leary MH (1996) Phosphoenolpyruvate carboxylase: A ubiquitous, highly regulated enzyme in plants. *Annual Review of Plant Biology* **47** pp 273-298.

Cushman JC (2001) Crassulacean acid metabolism: a plastic photosynthetic adaptation to arid environments. *Plant Physiology* **127** pp1439-1448.

Echevarria C, Vidal J (2003) The unique phosphoenolpyruvate carboxylase kinase. *Plant Physiology and Biochemistry* **41** pp541-547.

Fraser NJ (2006) Expression and functional purification of a glycosylation deficient version of the human adenosine 2a receptor for structural studies. *Protein Expression and Purification* **49** pp129-137.

Foyer CH, Bloom AJ, Queval G, Noctor G (2009) Photorespiratory metabolism: genes, mutants, energetics and redox signalling. *Annual Review of Plant Biology* **60** pp455-484.

Fukayama H, Tamai T, Taniguchi Y, Sullivan S, Miyao M, Nimmo HG (2006) Characterization and functional analysis of phosphoenolpyruvate carboxylase kinase genes in rice. *Plant Journal* **47** pp258-268.

Gomase VS, Tagore S (2008) Kinomics. *Current Drug Metabolism* **9** pp255-258.

Hanks SK, Hunter T (1995) Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB Journal* **9** pp576-596

Hartwell J, Smith LH, Wilkins MB, Jenkins GI, Nimmo HG (1996) Higher plant phosphoenolpyruvate carboxylase kinase is regulated at the level of translatable mRNA in response to light or a circadian rhythm. *Plant Journal* **10** pp1071-1078.

Hartwell J, Gill A, Nimmo GA, Wilkins MB, Jenkins GI, Nimmo HG (1999) Phosphoenolpyruvate carboxylase kinase is a novel protein kinase regulated at the level of expression. *Plant Journal* **20** pp333-342.

Hastie CJ, McLauchlan HJ, Cohen P (2001) Assay of protein kinases using radiolabelled ATP: a protocol. *Nature Protocols* **1** pp968-971.

Jancarik J, Kim SH (1991) Sparse matrix sampling: a screening method for crystallization of proteins. *Journal of Applied Crystallography* **24** pp409-411.

Johnson LN, Lewis RJ (2001) Structural basis for control by phosphorylation. *Chemcial Reviews* **101** pp.2209-2242.

Kai Y, Matsumura H, Inoue T, Terada K, Nagara Y, Kihara A, Tsumura K, Izui K (1999) Three-dimensional structure of phosphoenolpyruvate carboxylase: a proposed mechanism for allosteric inhibition. *Proceedings of the National Academy of Sciences USA* **96** pp823-828.

Kai Y, Matsumura H, Izui K (2003) Phosphoenolpyruvate carboxylase: three-dimensional structure and molecular mechanisms. *Archives of Biochemistry and Biophysics* **2** pp170-179.

Kelly SM, Tess TJ, Price NC (2005) How to study proteins by circular dichroism. *Biochimica et Biophysica Acta - Proteins and Proteomics* **2** pp119-139.

Knighton DR, Zheng JH, Ten Eyck LF, Ashford VA, Xuong NH, Taylor SS, Sowadski JM (1991a) Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253** pp407-414.

Knighton DR, Zheng JH, Ten Eyck LF, Xuong NH, Taylor SS, Sowadski JM (1991b) Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253** pp414-420.

Kowluru A (2008) Emerging roles for protein histidine phosphorylation in cellular signal transduction: lessons from the islet beta-cell. *Journal of Cellular and Molecular Medicine* **12** pp1885-1908.

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227** pp680-685.

Leegood RC (2002) C_4 photosynthesis: principles of CO_2 concentration and prospects for its introduction into C_3 plants. *Journal of Experimental Botany* **53** pp581-590.

Li B, Pacquit V, Jiao JA, Duff SMG, Maralihalli GB, Sarath G, Condon SA, Vidal J, Chollet R (1997) Structural requirements for phosphorylation of C_4-leaf phosphoenolpyruvate carboxylase by its highly regulated protein-serine kinase. A comparative study with synthetic peptide substrates and mutant target proteins. *Australian Journal of Plant Physiology* **24** pp443-449.

Luan S (2003) Protein phosphatases in plants. *Annual Review of Plant Biology* **54** pp63-92.

Marsh JT, Sullivan S, Hartwell J, Nimmo HG (2003) Structure and expression of phosphoenolpyruvate carboxylase kinase genes in Solanaceae. A novel gene exhibits alternative splicing. *Plant Physiology* **133** pp2021-2028.

Masumoto (2007) *Photosynthesis Research* **91** pp

Matsumura H, Terada M, Shirakata S, Inoue T, Yoshinaga T, Izui K, Kai Y (1999) Plausible phosphoenolpyruvate binding site revealed by 2.6 Å structure of Mn^{2+}-bound phosphoenolpyruvate carboxylase from *Escherichia coli*. *FEBS Letters* **458** pp93-96.

Matsumura H, Xie Y, Shirakata S, Inoue T, Yoshinaga T, Ueno Y, Izui K. Kai Y (2002) Crystal structures of C_4 form maize and quaternary complex of *E. coli* phosphoenolpyruvate carboxylases. *Structure* **10** pp1721-1730.

Mercer SE, Friedman E (2006) Mirk/Dyrk1B: a multifunctional dual-specificity kinase involved in growth arrest, differentiation and cell survival. *Cell Biochemistry and Biophysics* **45** pp303-315.

Miyao M, Fukayama M (2003) Metabolic consequences of overproduction of phosphoenolpyruvate carboxylase in C_3 plants. *Archives of Biochemistry and Biophysics* **414** pp197-203

McNaughton GA, Fewson CA, Wilkins MB, Nimmo HG (1989) Purification, oligomerization state and malate sensitivity of maize leaf phosphoenolpyruvate carboxylase. *Biochemical Journal* **261** pp349-355.

Nimmo GA, Nimmo HG, Fewson CA, Wilkins MB (1984) Diurnal changes in the properties of phosphoenolpyruvate carboxylase in *Bryophyllum* leaves: a possible covalent modification. *FEBS Letters* **178** pp199-203.

Nimmo GA, Nimmo HG, Hamilton ID, Fewson CA, Wilkins MD (1986) Purification of the phosphorylated night form and dephosphorylated day form of phosphoenolpyruvate carboxylase from *Bryophyllum fedtschenkoi*. *Biochemical Journal* **239** pp213-220.

Nimmo GA, Wilkins MB, Nimmo HG (2001) Partial purification and characterization of a protein inhibitor of phosphoenolpyruvate carboxylase kinase. *Planta* **213** pp250-257.

Nimmo HG (2000) The regulation of phosphoenolpyruvate carboxylase in CAM plants. *Trends in Plant Science* **5** pp75-80.

Nimmo HG (2003) Control of the phosphorylation of phosphoenolpyruvate carboxylase in higher plants. *Archives of Biochemistry and Biophysics* **414** pp189-196.

O'Leary B, Rao SK, Kim J, Plaxton WC (2009) Bacterial-type phosphoenolpyruvate carboxylase (PEPc) functions as a catalytic and regulatory subunit of the novel Class-2 PEPc complex of vascular plants. *Journal of Biological Chemistry* **284** pp24797-24805.

Raven JA, Cockell CS, De La Rocha CL (2008) The evolution of inorganic carbon concentrating mechanisms in photosynthesis. *Philosophical Transactions of the Royal Society (London) B - Biological Sciences* **363** pp2641-2650.

Reinfelder JR, Kraepiel AML, Morel FMM (2000) Unicellular C_4 photosynthesis in a marine diatom. *Nature* **407** pp996-999.

Rivoal J, Dunford R, Plaxton WC, Turpin DH (1996) Purification and properties of four phosphoenolpyruvate carboxylase isoforms from the green alga *Selanastrum minutum*: evidence that association of the 102-kDa catalytic subunit with unrelated polypeptides may modify the physical and kinetic properties of the enzyme. *Archives of Biochemistry and Biophysics* **332** pp47-57.

Rivoal J, Plaxton WC, Turpin DH (1998) Purification and characterization of high- and low-molecular-mass isoforms of phosphoenolpyruvate carboxylase from *Chlamydomonas reinhardtii*. Kinetic, structural and immunological evidence that the green algal enzyme is distinct from the prokaryotic and higher plant enzymes. *Biochemical Journal* **331** pp201-209.

Rivoal J, Trzos S, Gage DA, Plaxton WC, Turpin DH (2001) Two unrelated phosphoenolpyruvate carboxylase polypeptides physically interact in the high molecular mass isoforms of this enzyme in the unicellular green alga *Selenastrum minutum*. *Journal of Biological Chemistry* **276** pp12588-12597.

Sanchez R, Cejudo FJ (2003) Identification and expression analysis of a gene encoding bacterial-type phosphoenolpyruvate carboxylase from Arabidopsis and rice. *Plant Physiology* **132** pp949-957.

Saze H, Ueno Y, Hisabori T, Hayashi H, Izui K (2001) Thioredoxin-mediated reductive activation of a protein kinase for the regulatory phosphorylation of C_4-form phosphoenolpyruvate carboxylase from maize. *Plant and Cell Physiology* **42** pp1295-1302.

Sullivan S, Jenkins GI, Nimmo HG (2004) Roots, cycles and leaves. Expression of the phosphoenolpyruvate carboxylase kinase gene family in soybean. *Plant Physiology* **135** pp2078-2087.

Thomas BR, Vekilov PG, Rosenberger F (1998) Effects of microheterogeneity in hen egg-white lysozyme crystallization. *Acta Crystallographica* D**54** pp226-236.

Vidal J, Chollet R (1997) Regulatory phosphorylation of C_4 PEP carboxylase. *Trends in Plant Science* **2** pp230-237.

Wang YH, Chollet R (1993) Partial purification and characterization of phosphoenolpyruvate protein-serine kinase from illuminated maize leaves. *Archives of Biochemistry and Biophysics* **304** pp496-502.

Wiencek JM (1999) New strategies for protein crystal growth. *Annual Review of Biomedical Engineering* **1** pp505-534.

Xu W, Sato SJ, Clemente TE, Chollet R (2007) The PEP-carboxylase kinase gene family in *Glycine max* (GmPpcK1-4): an in-depth molecular analysis with nodulated, non-transgenic and transgenic plants. *Plant Journal* **49** pp910-923.

Zhang XQ, Chollet R (1997) Phosphoenolpyruvate carboxylase protein kinase from soybean root nodules: partial purification, characterization and up/down-regulation by photosynthate supply from the shoots. *Archives of Biochemistry and Biophysics* **343** pp260-268.

Appendix 1

Compositions of crystallisation buffers

Structure Screen 1 (Molecular Dynamics)

Tube #	Salt	Buffer	pH	Precipitant
1	0.02 M calcium chloride	0.1 M sodium acetate	4.6	30 % v/v MPD
2	0.2 M ammonium acetate	0.1 M sodium acetate	4.6	30 % w/v PEG 4K
3	0.2 M ammonium sulfate	0.1 M sodium acetate	4.6	25 % w/v PEG 4K
4	None	0.1 M sodium acetate	4.6	2.0 M sodium formate
5	None	0.1 M sodium acetate	4.6	2.0 M ammonium sulfate
6	None	0.1 M sodium acetate	4.6	8 % w/v PEG 4K
7	0.2 M ammonium acetate	0.1 M tri-sodium citrate	5.6	30 % w/v PEG 4K
8	0.2 M ammonium acetate	0.1 M tri-sodium citrate	5.6	30 % v/v MPD
9	None	0.1 M tri-sodium citrate	5.6	20 % v/v 2-propanol, 20%w/v PEG 4K
10	None	0.1 M tri-sodium citrate	5.6	1.0 M ammonium dihydrogen phosphate
11	0.2 M calcium chloride	0.1 M sodium acetate	4.6	20 % v/v 2-propanol
12	None	0.1 M sodium cacodylate	6.5	1.4 M sodium acetate
13	0.2 M tri-sodium citrate	0.1 M sodium cacodylate	6.5	30 % v/v 2-propanol
14	0.2 M ammonium sulfate	0.1 M sodium cacodylate	6.5	30 % w/v PEG 8K
15	0.2 M magnesium acetate	0.1 M sodium cacodylate	6.5	20 % w/v PEG 8K
16	0.2 M magnesium acetate	0.1 M sodium cacodylate	6.5	30 % v/v MPD
17	None	0.1 M imidazole	6.5	1.0 M sodium acetate
18	0.2 M sodium acetate	0.1 M sodium cacodylate	6.5	30 % w/v PEG 8K
19	0.2 M zinc acetate	0.1 M sodium cacodylate	6.5	18 % w/v PEG 8K
20	0.2 M calcium acetate	0.1 M sodium cacodylate	6.5	18 % w/v PEG 8K
21	0.2 M tri-sodium citrate	0.1 M Na HEPES	7.5	30 % v/v MPD
22	0.2 M magnesium chloride	0.1 M Na HEPES	7.5	30 % v/ v 2-propanol
23	0.2 M calcium chloride	0.1 M Na HEPES	7.5	28 % v/v PEG 400
24	0.2 M magnesium chloride	0.1 M Na HEPES	7.5	30 % v/v PEG 400
25	0.2 M tri-sodium citrate	0.1 M Na HEPES	7.5	20 % v/v 2-propanol
26	None	0.1 M Na HEPES	7.5	0.8 M K/Na tartrate
27	None	0.1 M Na HEPES	7.5	1.5 M lithium sulfate
28	None	0.1 M Na HEPES	7.5	0.8 M sodium dihydrogen phosphate/ 0.8 M K dihydrogen phosphate
29	None	0.1 M Na HEPES	7.5	1.4 M tri-sodium citrate
30	None	0.1 M Na HEPES	7.5	2 % v/v PEG 400, 2.0 M ammonium sulfate
31	None	0.1 M Na HEPES	7.5	10 % v/v 2-propanol, 20% w/v PEG 4K
32	None	0.1 M Tris	8.5	2.0 M ammonium sulfate
33	0.2 M magnesium chloride	0.1 M Tris	8.5	30 % w/v PEG 4K
34	0.2 M tri-sodium citrate	0.1 M Tris	8.5	30 % v/v PEG 400
35	0.2 M lithium sulfate	0.1 M Tris	8.5	30 % w/v PEG 4K
36	0.2 M ammonium acetate	0.1 M Tris	8.5	30 % v/v 2-Propanol
37	0.2 M sodium acetate	0.1 M Tris	8.5	30 % w/v PEG 4K
38	None	0.1 M Tris	8.5	8 % w/v PEG 8K
39	None	0.1 M Tris	8.5	2.0 M ammonium dihydrogen phosphate
40	None	None	-	0.4 M K/Na Tartrate
41	None	None	-	0.4 M ammonium dihydrogen phosphate
42	0.2 M ammonium sulfate	None	-	30 % w/v PEG 8K
43	0.2 M ammonium sulfate	None	-	30 % w/v PEG 4K
44	None	None	-	2.0 M ammonium sulfate
45	None	None	-	4.0 M sodium formate
46	0.05 M potassium dihydrogen phosphate	None	-	20 % w/v PEG 8K
47	None	None	-	30 % w/v PEG 1.5K
48	None	None	-	0.2 M magnesium formate
49	1.0 M lithium sulfate	None	-	2 % w/v PEG 8K
50	0.5 M lithium sulfate	None	-	15 % w/v PEG 8K

Structure Screen 2 (Molecular Dynamics)

Tube No.	Salt	Buffer	pH	Precipitant
1	0.1 M sodium chloride	0.1 M Bicine	9.0	30 % v/v PEG 550 MME
2	None	0.1 M Bicine	9.0	2.0 M magnesium chloride
3	None	0.1 M Bicine	9.0	2 % v/v 1,4-Dioxane/10 % w/v PEG 20,000
4	0.2 M magnesium chloride	0.1 M Tris	8.5	3.4 M 1,6-hexanediol
5	None	0.1 M Tris	8.5	25 % v/v tert-butanol
6	0.01 M nickel chloride	0.1 M Tris	8.5	1.0 M lithium sulfate
7	1.5 M ammonium sulfate	0.1 M Tris	8.5	12 % v/v glycerol
8	0.2 M ammonium dihydrogen phosphate	0.1 M Tris	8.5	50 % v/v MPD
9	None	0.1 M Tris	8.5	20 % v/v ethanol
10	0.01 M nickel chloride	0.1 M Tris	8.5	20 % w/v PEG 2000 MME
11	0.5 M ammonium sulfate	0.1 M Na HEPES	7.5	30 % v/v MPD
12	None	0.1 M Na HEPES	7.5	10 % w/v PEG 6000, 5% v/v MPD
13	None	0.1 M Na HEPES	7.5	20 % v/v Jeffamine M-600
14	0.1 M sodium chloride	0.1 M Na HEPES	7.5	1.6 M ammonium sulfate
15	None	0.1 M Na HEPES	7.5	2.0 M ammonium formate
16	0.05 M cadmium sulfate	0.1 M Na HEPES	7.5	1.0 M sodium acetate
17	None	0.1 M Na HEPES	7.5	70 % v/v MPD
18	None	0.1 M Na HEPES	7.5	4.3 M sodium chloride
19	None	0.1 M Na HEPES	7.5	10 % w/v PEG 8000, 8 % v/v ethylene glycol
20	None	0.1 M MES	6.5	1.6 M magnesium sulfate
21	0.1 M potassium phosphate + 0.1 M sodium phosphate	0.1 M MES	6.5	2.0 M sodium chloride
22	None	0.1 M MES	6.5	12 % w/v PEG 20,000
23	1.6 M ammonium sulfate	0.1 M MES	6.5	10 % v/v Dioxane
24	0.05 M caesium chloride	0.1 M MES	6.5	30 % v/v Jeffamine M-600
25	0.01 M cobalt chloride	0.1 M MES	6.5	1.8 M ammonium sulfate
26	0.2 M ammonium sulfate	0.1 M MES	6.5	30 % w/v PEG 5000 MME
27	0.01 M zinc sulfate	0.1 M MES	6.5	25 % v/v PEG 550 MME
28	None	0.1 M Na HEPES	7.5	20 % w/v PEG 10,000
29	0.2 M potassium sodium tartrate	0.1 M Na citrate	5.6	2.0 M ammonium sulfate
30	0.5 M ammonium sulfate	0.1 M Na citrate	5.6	1.0 M lithium sulfate
31	0.5 M sodium chloride	0.1 M Na citrate	5.6	4 % v/v polyethyleneimine
32	None	0.1 M Na citrate	5.6	35 % v/v tert-butanol
33	0.01 M ferric chloride	0.1 M Na citrate	5.6	10 % v/v Jeffamine M-600
34	0.01 M manganese chloride	0.1 M Na citrate	5.6	2.5 M 1,6-hexanediol
35	None	0.1 M Na acetate	4.6	2.0 M sodium chloride
36	0.2 M sodium chloride	0.1 M Na acetate	4.6	30 % v/v MPD
37	0.01 M cobalt chloride	0.1 M Na acetate	4.6	1.0 M 1,6-hexanediol
38	0.1 M cadmium chloride	0.1 M Na acetate	4.6	30 % v/v PEG 400
39	0.2 M ammonium sulfate	0.1 M Na acetate	4.6	30 % w/v PEG 2000 MME
40	2.0 M sodium chloride	None	None	10 % w/v PEG 6000
41	0.01 M CTAB	None	None	0.5 M sodium chloride, 0.1 M magnesium chloride
42	None	None	None	25 % v/v ethylene glycol
43	None	None	None	35 % v/v dioxane
44	2.0 M ammonium sulfate	None	None	5 % v/v 2-propanol
45	None	None	None	1.8 M imidazole pH 7.0
46	None	None	None	10 % w/v PEG 1000, 10 % w/v PEG 8000
47	1.5 M sodium chloride	None	None	10 % v/v ethanol
48	None	None	None	1.6 M sodium citrate pH 6.5
49	15 % w/v polyvinylpyrrolidone	None	None	None
50	2.0 M urea	None	None	None

Crystal Screen 1 (Hampton Research)

Tube #	Salt	Tube #	Buffer	Tube #	Precipitant
1.	0.02 M Calcium chloride dihydrate	1.	0.1 M Sodium acetate trihydrate pH 4.6	1.	30% v/v (+/-)-2-Methyl-2,4-pentanediol
2.	None	2.	None	2.	0.4 M Potassium sodium tartrate tetrahydrate
3.	None	3.	None	3.	0.4 M Ammonium phosphate monobasic
4.	None	4.	0.1 M TRIS hydrochloride pH 8.5	4.	2.0 M Ammonium sulfate
5.	0.2 M Sodium citrate tribasic dihydrate	5.	0.1 M HEPES sodium pH 7.5	5.	30% v/v (+/-)-2-Methyl-2,4-pentanediol
6.	0.2 M Magnesium chloride hexahydrate	6.	0.1 M TRIS hydrochloride pH 8.5	6.	30% w/v Polyethylene glycol 4,000
7.	None	7.	0.1 M Sodium cacodylate trihydrate pH 6.5	7.	1.4 M Sodium acetate trihydrate
8.	0.2 M Sodium citrate tribasic dihydrate	8.	0.1 M Sodium cacodylate trihydrate pH 6.5	8.	30% v/v 2-Propanol
9.	0.2 M Ammonium acetate	9.	0.1 M Sodium citrate tribasic dihydrate pH 5.6	9.	30% w/v Polyethylene glycol 4,000
10.	0.2 M Ammonium acetate	10.	0.1 M Sodium acetate trihydrate pH 4.6	10.	30% w/v Polyethylene glycol 4,000
11.	None	11.	0.1 M Sodium citrate tribasic dihydrate pH 5.6	11.	1.0 M Ammonium phosphate monobasic
12.	0.2 M Magnesium chloride hexahydrate	12.	0.1 M HEPES sodium pH 7.5	12.	30% v/v 2-Propanol
13.	0.2 M Sodium citrate tribasic dihydrate	13.	0.1 M TRIS hydrochloride pH 8.5	13.	30% v/v Polyethylene glycol 400
14.	0.2 M Calcium chloride dihydrate	14.	0.1 M HEPES sodium pH 7.5	14.	28% v/v Polyethylene glycol 400
15.	0.2 M Ammonium sulfate	15.	0.1 M Sodium cacodylate trihydrate pH 6.5	15.	30% w/v Polyethylene glycol 8,000
16.	None	16.	0.1 M HEPES sodium pH 7.5	16.	1.5 M Lithium sulfate monohydrate
17.	0.2 M Lithium sulfate monohydrate	17.	0.1 M TRIS hydrochloride pH 8.5	17.	30% w/v Polyethylene glycol 4,000
18.	0.2 M Magnesium acetate tetrahydrate	18.	0.1 M Sodium cacodylate trihydrate pH 6.5	18.	20% w/v Polyethylene glycol 8,000
19.	0.2 M Ammonium acetate	19.	0.1 M TRIS hydrochloride pH 8.5	19.	30% v/v 2-Propanol
20.	0.2 M Ammonium sulfate	20.	0.1 M Sodium acetate trihydrate pH 4.6	20.	25% w/v Polyethylene glycol 4,000
21.	0.2 M Magnesium acetate tetrahydrate	21.	0.1 M Sodium cacodylate trihydrate pH 6.5	21.	30% v/v (+/-)-2-Methyl-2,4-pentanediol
22.	0.2 M Sodium acetate trihydrate	22.	0.1 M TRIS hydrochloride pH 8.5	22.	30% w/v Polyethylene glycol 4,000
23.	0.2 M Magnesium chloride hexahydrate	23.	0.1 M HEPES sodium pH 7.5	23.	30% v/v Polyethylene glycol 400
24.	0.2 M Calcium chloride dihydrate	24.	0.1 M Sodium acetate trihydrate pH 4.6	24.	20% v/v 2-Propanol
25.	None	25.	0.1 M Imidazole pH 6.5	25.	1.0 M Sodium acetate trihydrate
26.	0.2 M Ammonium acetate	26.	0.1 M Sodium citrate tribasic dihydrate pH 5.6	26.	30% v/v (+/-)-2-Methyl-2,4-pentanediol
27.	0.2 M Sodium citrate tribasic dihydrate	27.	0.1 M HEPES sodium pH 7.5	27.	20% v/v 2-Propanol
28.	0.2 M Sodium acetate trihydrate	28.	0.1 M Sodium cacodylate trihydrate pH 6.5	28.	30% w/v Polyethylene glycol 8,000
29.	None	29.	0.1 M HEPES sodium pH 7.5	29.	0.8 M Potassium sodium tartrate tetrahydrate
30.	0.2 M Ammonium sulfate	30.	None	30.	30% w/v Polyethylene glycol 8,000
31.	0.2 M Ammonium sulfate	31.	None	31.	30% w/v Polyethylene glycol 4,000
32.	None	32.	None	32.	2.0 M Ammonium sulfate
33.	None	33.	None	33.	4.0 M Sodium formate
34.	None	34.	0.1 M Sodium acetate trihydrate pH 4.6	34.	2.0 M Sodium formate
35.	None	35.	0.1 M HEPES sodium pH 7.5	35.	0.8 M Sodium phosphate monobasic monohydrate 0.8 M Potassium phosphate monobasic
36.	None	36.	0.1 M TRIS hydrochloride pH 8.5	36.	8% w/v Polyethylene glycol 8,000
37.	None	37.	0.1 M Sodium acetate trihydrate pH 4.6	37.	8% w/v Polyethylene glycol 4,000
38.	None	38.	0.1 M HEPES sodium pH 7.5	38.	1.4 M Sodium citrate tribasic dihydrate
39.	None	39.	0.1 M HEPES sodium pH 7.5	39.	2% v/v Polyethylene glycol 400 2.0 M Ammonium sulfate
40.	None	40.	0.1 M Sodium citrate tribasic dihydrate pH 5.6	40.	20% v/v 2-Propanol 20% w/v Polyethylene glycol 4,000
41.	None	41.	0.1 M HEPES sodium pH 7.5	41.	10% v/v 2-Propanol 20% w/v Polyethylene glycol 4,000
42.	0.05 M Potassium phosphate monobasic	42.	None	42.	20% w/v Polyethylene glycol 8,000
43.	None	43.	None	43.	30% w/v Polyethylene glycol 1,500
44.	None	44.	None	44.	0.2 M Magnesium formate dihydrate
45.	0.2 M Zinc acetate dihydrate	45.	0.1 M Sodium cacodylate trihydrate pH 6.5	45.	18% w/v Polyethylene glycol 8,000
46.	0.2 M Calcium acetate hydrate	46.	0.1 M Sodium cacodylate trihydrate pH 6.5	46.	18% w/v Polyethylene glycol 8,000
47.	None	47.	0.1 M Sodium acetate trihydrate pH 4.6	47.	2.0 M Ammonium sulfate
48.	None	48.	0.1 M TRIS hydrochloride pH 8.5	48.	2.0 M Ammonium phosphate monobasic
49.	1.0 M Lithium sulfate monohydrate	49.	None	49.	2% w/v Polyethylene glycol 8,000
50.	0.5 M Lithium sulfate monohydrate	50.	None	50.	15% w/v Polyethylene glycol 8,000

Crystal Screen 2 (Hampton Research)

Tube #	Salt	Tube #	Buffer c	Tube #	Precipitant
1.	2.0 M Sodium chloride	1.	None	1.	10% w/v Polyethylene glycol 6,000
2.	0.5 M Sodium chloride 0.01 M Magnesium chloride hexahydrate	2.	None	2.	0.01 M Hexadecyltrimethylammonium bromide
3.	None	3.	None	3.	25% v/v Ethylene glycol
4.	None	4.	None	4.	35% v/v 1,4-Dioxane
5.	2.0 M Ammonium sulfate	5.	None	5.	5% v/v 2-Propanol
6.	None	6.	None	6.	1.0 M Imidazole pH 7.0
7.	None	7.	None	7.	10% w/v Polyethylene glycol 1,000 10% w/v Polyethylene glycol 8,000
8.	1.5 M Sodium chloride	8.	None	8.	10% v/v Ethanol
9.	None	9.	0.1 M Sodium acetate trihydrate pH 4.6	9.	2.0 M Sodium chloride
10.	0.2 M Sodium chloride	10.	0.1 M Sodium acetate trihydrate pH 4.6	10.	30% v/v (+/-)-2-Methyl-2,4-pentanediol
11.	0.01 M Cobalt (II) chloride hexahydrate	11.	0.1 M Sodium acetate trihydrate pH 4.6	11.	1.0 M 1,6-Hexanediol
12.	0.1 M Cadmium chloride hydrate	12.	0.1 M Sodium acetate trihydrate pH 4.6	12.	30% v/v Polyethylene glycol 400
13.	0.2 M Ammonium sulfate	13.	0.1 M Sodium acetate trihydrate pH 4.6	13.	30% w/v Polyethylene glycol monomethyl ether 2,000
14.	0.2 M Potassium sodium tartrate tetrahydrate	14.	0.1 M Sodium citrate tribasic dihydrate pH 5.6	14.	2.0 M Ammonium sulfate
15.	0.5 M Ammonium sulfate	15.	0.1 M Sodium citrate tribasic dihydrate pH 5.6	15.	1.0 M Lithium sulfate monohydrate
16.	0.5 M Sodium chloride	16.	0.1 M Sodium citrate tribasic dihydrate pH 5.6	16.	2% v/v Ethylene imine polymer
17.	None	17.	0.1 M Sodium citrate tribasic dihydrate pH 5.6	17.	35% v/v tert-Butanol
18.	0.01 M Iron (III) chloride hexahydrate	18.	0.1 M Sodium citrate tribasic dihydrate pH 5.6	18.	10% v/v Jeffamine M-600®
19.	None	19.	0.1 M Sodium citrate tribasic dihydrate pH 5.6	19.	2.5 M 1,6-Hexanediol
20.	None	20.	0.1 M MES monohydrate pH 6.5	20.	1.6 M Magnesium sulfate heptahydrate
21.	0.1 M Sodium phosphate monobasic monohydrate 0.1 M Potassium phosphate monobasic	21.	0.1 M MES monohydrate pH 6.5	21.	2.0 M Sodium chloride
22.	None	22.	0.1 M MES monohydrate pH 6.5	22.	12% w/v Polyethylene glycol 20,000
23.	1.6 M Ammonium sulfate	23.	0.1 M MES monohydrate pH 6.5	23.	10% v/v 1,4-Dioxane
24.	0.05 M Cesium chloride	24.	0.1 M MES monohydrate pH 6.5	24.	30% v/v Jeffamine M-600®
25.	0.01 M Cobalt (II) chloride hexahydrate	25.	0.1 M MES monohydrate pH 6.5	25.	1.8 M Ammonium sulfate
26.	0.2 M Ammonium sulfate	26.	0.1 M MES monohydrate pH 6.5	26.	30% w/v Polyethylene glycol monomethyl ether 5,000
27.	0.01 M Zinc sulfate heptahydrate	27.	0.1 M MES monohydrate pH 6.5	27.	25% v/v Polyethylene glycol monomethyl ether 550
28.	None	28.	None	28.	1.6 M Sodium citrate tribasic dihydrate pH 6.5
29.	0.5 M Ammonium sulfate	29.	0.1 M HEPES pH 7.5	29.	30% v/v (+/-)-2-Methyl-2,4-pentanediol
30.	None	30.	0.1 M HEPES pH 7.5	30.	10% w/v Polyethylene glycol 8,000 5% v/v (+/-)-2-Methyl-2,4-pentanediol
31.	None	31.	0.1 M HEPES pH 7.5	31.	20% v/v Jeffamine M-600®
32.	0.1 M Sodium chloride	32.	0.1 M HEPES pH 7.5	32.	1.6 M Ammonium sulfate
33.	None	33.	0.1 M HEPES pH 7.5	33.	2.0 M Ammonium formate
34.	0.05 M Cadmium sulfate hydrate	34.	0.1 M HEPES pH 7.5	34.	1.0 M Sodium acetate trihydrate
35.	None	35.	0.1 M HEPES pH 7.5	35.	70% v/v (+/-)-2-Methyl-2,4-pentanediol
36.	None	36.	0.1 M HEPES pH 7.5	36.	4.3 M Sodium chloride
37.	None	37.	0.1 M HEPES pH 7.5	37.	10% w/v Polyethylene glycol 8,000 8% v/v Ethylene glycol
38.	None	38.	0.1 M HEPES pH 7.5	38.	20% w/v Polyethylene glycol 10,000
39.	0.2 M Magnesium chloride hexahydrate	39.	0.1 M Tris pH 8.5	39.	3.4 M 1,6-Hexanediol
40.	None	40.	0.1 M Tris pH 8.5	40.	25% v/v tert-Butanol
41.	0.01 M Nickel (II) chloride hexahydrate	41.	0.1 M Tris pH 8.5	41.	1.0 M Lithium sulfate monohydrate
42.	1.5 M Ammonium sulfate	42.	0.1 M Tris pH 8.5	42.	12% v/v Glycerol
43.	0.2 M Ammonium phosphate monobasic	43.	0.1 M Tris pH 8.5	43.	50% v/v (+/-)-2-Methyl-2,4-pentanediol
44.	None	44.	0.1 M Tris pH 8.5	44.	20% v/v Ethanol
45.	0.01 M Nickel (II) chloride hexahydrate	45.	0.1 M Tris pH 8.5	45.	20% w/v Polyethylene glycol monomethyl ether 2,000
46.	0.1 M Sodium chloride	46.	0.1 M BICINE pH 9.0	46.	20% v/v Polyethylene glycol monomethyl ether 550
47.	None	47.	0.1 M BICINE pH 9.0	47.	2.0 M Magnesium chloride hexahydrate
48.	None	48.	0.1 M BICINE pH 9.0	48.	2% v/v 1,4-Dioxane 10% w/v Polyethylene glycol 20,000

Wizard Screen I (Emerald Biosystems)

	crystallant	buffer (0.1 M)	salt (0.2 M)	
1	20% (w/v) PEG-8000	CHES pH 9.5	none	1
2	10% (v/v) 2-propanol	HEPES pH 7.5	NaCl	2
3	15% (v/v) ethanol	CHES pH 9.5	none	3
4	35% (v/v) 2-methyl-2,4-pentanediol	imidazole pH 8.0	$MgCl_2$	4
5	30% (v/v) PEG-400	CAPS pH 10.5	none	5
6	20% (w/v) PEG-3000	citrate pH 5.5	none	6
7	10% (w/v) PEG-8000	MES pH 6.0	$Zn(OAc)_2$	7
8	2.0 M $(NH_4)_2SO_4$	citrate pH 5.5	none	8
9	1.0 M $(NH_4)_2HPO_4$	acetate pH 4.5	none	9
10	20% (w/v) PEG-2000 MME	Tris pH 7.0	none	10
11	20% (v/v) 1,4-butanediol	MES pH 6.0	Li_2SO_4	11
12	20% (w/v) PEG-1000	imidazole pH 8.0	$Ca(OAc)_2$	12
13	1.26 M $(NH_4)_2SO_4$	cacodylate pH 6.5	none	13
14	1.0 M sodium citrate	cacodylate pH 6.5	none	14
15	10% (w/v) PEG-3000	imidazole pH 8.0	Li_2SO_4	15
16	2.5 M NaCl	Na/K phosphate pH 6.2	none	16
17	30% (w/v) PEG-8000	acetate pH 4.5	Li_2SO_4	17
18	1.0 M K/Na tartrate	imidazole pH 8.0	NaCl	18
19	20% (w/v) PEG-1000	Tris pH 7.0	none	19
20	0.4 M NaH_2PO_4/1.6 M K_2HPO_4	imidazole pH 8.0	NaCl	20
21	20% (w/v) PEG-8000	HEPES pH 7.5	none	21
22	10% (v/v) 2-propanol	Tris pH 8.5	none	22
23	15% (v/v) ethanol	imidazole pH 8.0	$MgCl_2$	23
24	35% (v/v) 2-methyl-2,4-pentanediol	Tris pH 7.0	NaCl	24
25	30% (v/v) PEG-400	Tris pH 8.5	$MgCl_2$	25
26	10% (w/v) PEG-3000	CHES pH 9.5	none	26
27	1.2 M NaH_2PO_4/0.8 M K_2HPO_4	CAPS pH 10.5	Li_2SO_4	27
28	20% (w/v) PEG-3000	HEPES pH 7.5	NaCl	28
29	10% (w/v) PEG-8000	CHES pH 9.5	NaCl	29
30	1.26 M $(NH_4)_2SO_4$	acetate pH 4.5	NaCl	30
31	20% (w/v) PEG-8000	phosphate-citrate pH 4.2	NaCl	31
32	10% (w/v) PEG-3000	Na/K phosphate pH 6.2	none	32
33	2.0 M $(NH_4)_2SO_4$	CAPS pH 10.5	Li_2SO_4	33
34	1.0 M $(NH_4)_2HPO_4$	imidazole pH 8.0	none	34
35	20% (v/v) 1,4-butanediol	acetate pH 4.5	none	35
36	1.0 M sodium citrate	imidazole pH 8.0	none	36
37	2.5 M NaCl	imidazole pH 8.0	none	37
38	1.0 M K/Na tartrate	CHES pH 9.5	Li_2SO_4	38
39	20% (w/v) PEG-1000	phosphate-citrate pH 4.2	Li_2SO_4	39
40	10% (v/v) 2-propanol	MES pH 6.0	$Ca(OAc)_2$	40
41	30% (w/v) PEG-3000	CHES pH 9.5	none	41
42	15% (v/v) ethanol	Tris pH 7.0	none	42
43	35% (v/v) 2-methyl-2,4-pentanediol	Na/K phosphate pH 6.2	none	43
44	30% (v/v) PEG-400	acetate pH 4.5	$Ca(OAc)_2$	44
45	20% (w/v) PEG-3000	acetate pH 4.5	none	45
46	10% (w/v) PEG-8000	imidazole pH 8.0	$Ca(OAc)_2$	46
47	1.26 M $(NH_4)_2SO_4$	Tris pH 8.5	Li_2SO_4	47
48	20% (w/v) PEG-1000	acetate pH 4.5	$Zn(OAc)_2$	48

Wizard Screen II (Emerald Biosystems)

	crystallant	buffer (0.1 M)	salt (0.2 M)	
1	10% (w/v) PEG-3000	acetate pH 4.5	$Zn(OAc)_2$	1
2	35% (v/v) 2-methyl-2,4-pentanediol	MES pH 6.0	Li_2SO_4	2
3	20% (w/v) PEG-8000	Tris pH 8.5	$MgCl_2$	3
4	2.0 M $(NH_4)_2SO_4$	cacodylate pH 6.5	NaCl	4
5	20% (v/v) 1,4-butanediol	HEPES pH 7.5	NaCl	5
6	10% (v/v) 2-propanol	phosphate-citrate pH 4.2	Li_2SO_4	6
7	30% (w/v) PEG-3000	Tris pH 7.0	NaCl	7
8	10% (w/v) PEG-8000	Na/K phosphate pH 6.2	NaCl	8
9	2.0 M $(NH_4)_2SO_4$	phosphate-citrate pH 4.2	none	9
10	1.0 M $(NH_4)_2HPO_4$	Tris pH 8.5	none	10
11	10% (v/v) 2-propanol	cacodylate pH 6.5	$Zn(OAc)_2$	11
12	30% (v/v) PEG-400	cacodylate pH 6.5	Li_2SO_4	12
13	15% (v/v) ethanol	citrate pH 5.5	Li_2SO_4	13
14	20% (w/v) PEG-1000	Na/K phosphate pH 6.2	NaCl	14
15	1.26 M $(NH_4)_2SO_4$	HEPES pH 7.5	none	15
16	1.0 M sodium citrate	CHES pH 9.5	none	16
17	2.5 M NaCl	Tris pH 7.0	$MgCl_2$	17
18	20% (w/v) PEG-3000	Tris pH 7.0	$Ca(OAc)_2$	18
19	1.6 M NaH_2PO_4/0.4 M K_2HPO_4	phosphate-citrate pH 4.2	none	19
20	15% (v/v) ethanol	MES pH 6.0	$Zn(OAc)_2$	20
21	35% (v/v) 2-methyl-2,4-pentanediol	acetate pH 4.5	none	21
22	10% (v/v) 2-propanol	imidazole pH 8.0	none	22
23	15% (v/v) ethanol	HEPES pH 7.5	$MgCl_2$	23
24	30% (w/v) PEG-8000	imidazole pH 8.0	NaCl	24
25	35% (v/v) 2-methyl-2,4-pentanediol	HEPES pH 7.5	NaCl	25
26	30% (v/v) PEG-400	CHES pH 9.5	none	26
27	10% (w/v) PEG-3000	cacodylate pH 6.5	$MgCl_2$	27
28	20% (w/v) PEG-8000	MES pH 6.0	$Ca(OAc)_2$	28
29	1.26 M $(NH_4)_2SO_4$	CHES pH 9.5	NaCl	29
30	20% (v/v) 1,4-butanediol	imidazole pH 8.0	$Zn(OAc)_2$	30
31	1.0 M sodium citrate	Tris pH 7.0	NaCl	31
32	20% (w/v) PEG-1000	Tris pH 8.5	none	32
33	1.0 M $(NH_4)_2HPO_4$	citrate pH 5.5	NaCl	33
34	10% (w/v) PEG-8000	imidazole pH 8.0	none	34
35	0.8 M NaH_2PO_4/1.2 M K_2HPO_4	acetate pH 4.5	none	35
36	10% (w/v) PEG-3000	phosphate-citrate pH 4.2	NaCl	36
37	1.0 M K/Na tartrate	Tris pH 7.0	Li_2SO_4	37
38	2.5 M NaCl	acetate pH 4.5	Li_2SO_4	38
39	20% (w/v) PEG-8000	CAPS pH 10.5	NaCl	39
40	20% (w/v) PEG-3000	imidazole pH 8.0	$Zn(OAc)_2$	40
41	2.0 M $(NH_4)_2SO_4$	Tris pH 7.0	Li_2SO_4	41
42	30% (v/v) PEG-400	HEPES pH 7.5	NaCl	42
43	10% (w/v) PEG-8000	Tris pH 7.0	$MgCl_2$	43
44	20% (w/v) PEG-1000	cacodylate pH 6.5	$MgCl_2$	44
45	1.26 M $(NH_4)_2SO_4$	MES pH 6.0	none	45
46	1.0 M $(NH_4)_2HPO_4$	imidazole pH 8.0	NaCl	46
47	2.5 M NaCl	imidazole pH 8.0	$Zn(OAc)_2$	47
48	1.0 M K/Na tartrate	MES pH 6.0	none	48