Functional characterisation of a predicted chloroplastic plant protein phosphatase

Michael David John Seymour

Functional Characterisation of a Predicted Chloroplastic Plant Protein Phosphatase

A thesis submitted by Michael David John Seymour B.Sc in accordance with the requirements of Durham University for the degree of Doctor of Philosophy.

Department of Biological and Biomedical Sciences

Durham University

2009

Abstract

The phosphatase AtPTPKIS1 is involved in the control of starch metabolism in *Arabidopsis thaliana* leaves at night. The SEX4 (Starch Excess 4) mutants, lacking this predicted phosphatase, have strongly reduced rates of starch metabolism. It is shown that this chloroplastic protein can bind to glucans through a carbohydrate binding domain (CBM) located within its previously predicted kinase interaction sequence (KIS), while another novel KIS containing protein (AKINβγ) shows no such interaction. Further analysis of the CBM identifies conserved residues vital for carbohydrate binding and common to CBMs, as well as sugar tongs, not present in similar CBMs or the GBD/KIS domain of the previously studies AMPKβ, but found within the binding domain of the PTPKIS family proteins.

While PTPKIS1 shows activity to generic phosphatase substrates, it is unable to dephosphorylate either phosphotyrosine or phosphothreonine containing peptides. It does however show phosphatase activity towards phosphorylated starch and amylopectin, comparable to that of the mammalian protein laforin. Remarkably, the most closely related protein to PTPKIS1 outside the plant kingdom is laforin, required for the metabolism of the mammalian storage carbohydrate glycogen and implicated in a severe form of epilepsy (Lafora disease) in humans, through the formation of insoluble starch like polyglucans (lafora bodies).

In addition to PTPKIS1, PTPKIS2 (At3g015180) is identified, a predicted phosphatase, with a domain structure homologous to that of PTPKIS1, termed. The PTPKIS2-SALK (PTPKIS2 knockout) mutant, lacking this predicted phosphatase, has a reduced rate of starch metabolism. It is shown that this mutant causes a phenotype similar to SEX4, but less extreme. It is further shown that this protein can bind to glucans through a carbohydrate binding domain (CBM), but unlike PTPKIS1 shows no activity towards any phosphate substrates. PTPKIS2 does however modulate the activity of PTPKIS1, causing a 4-fold increase in the activity of PTPKIS1 against phosphorylated starch, when both enzymes are present in equimolar concentrations.

Finally, a hypothesis is proposed as to the roles of PTPKIS1 and PTPKIS2 in starch metabolism, and the similarity of function seen in the mammalian protein Laforin.

Declaration

No material contained herein has been previously submitted for any other degree. Except where acknowledged, all material is the work of the author.

Statement of Copyright

Acknowledgements

For his guidance and assistance throughout this piece of work, I would like to thank Dr John Gatehouse. I would also like to thank Dr Tony Fordham-Skelton, whose work laid the foundation for this studentship. Thanks also goes to all the members of Lab 1 past and present who have ensured a pleasant working environment and who have, at some stage, all offered advice and support.

For their support, advice and discussion during my time at the Starch group, KVL, I thank Drs. Andreas Blennow, Lone Baunsgaard and Rene Mikkelsen. Thanks must also go to all those members of KVL who created such a wonderful working environment.

I would like to thank the staff and students of St Cuthbert's Society for their support and friendship.

I especially thank my family, who have given support and encouragement throughout my time at university.

Finally thank you to my parents, whose constant support and encouragement have been invaluable to me during this time, and all those preceding it.

Table of Contents

List of Figures

List of Tables

Glossary

AGPase	ADP-glucose pyrophosphorylase
AMPK	AMP-activated protein kinases
APBD	Adult polyglucosan body disease
CaMK	Calmodulin-dependent protein kinases
CAZy	Carbohydrate-Binding Module Family Server
CBM	Carbohydrate Binding Module
CCaMK	calcium and calmodulin-dependent protein kinases
CDPs	cysteine-dependent phosphatases
CDPK	calcium-dependent protein kinase
CNS	Central Nervous System
CRK	CDPK-related kinases
DBE	Debranching Enzyme
DEPC	Diethyl Pyrocarbonate
DNaseI	Deoxyribonuclease1
DsPTP	Dual-Specificity PTPs
DSP	Dual-Specificity Phosphatases
DPE1	Disproportionating Enzyme 1
DPE2	Cytosolic transglucosidase Disproportionating Enzyme 2
ECL	Enhanced chemoluminescence
EST	Expressed Sequence Tag
G-1-P	Glucose-1-Phosphate
G-6-P	Glucose-6-Phosphate
GBE	Glycogen Branching Enzyme
GBSS	Granule Bound Starch Synthase
GDBE	Glycogen Debranching Enzyme
GP	Glycogen Phosphorylase
GS	Glycogen Synthase
GST	Glutathione S-Transferase
GWD	Glucan Water Dikinase
GWD1	Glucan Water Dikinase 1
GWD2	Glucan Water Dikinase 2

GSD	Glycogen Storage Disease
ISA2	Arabidopsis knockout mutants in dbe1
KIS	Kinase Interaction Sequence
LBM	Luria-Bertani Medium
LD	Lafora Disease
MAPK	Mitogen-Activated Protein Kinase
MEX1	Chloroplastic maltose transporter
PAGE	Polyacrylamide gel electrophoresis
PEPRK	PEP carboxylase kinase-related kinases
PG	Phospho-glucoisomerase
PGM	Phosphoglucomutase
PME	Progressive Myoclonus Epilepsy
PP1	Serine/Threonine Phosphatases Type 1
PP1c	PP1 catalytic subunit
PP2	Serine/Threonine Phosphatases Type 2
PPCK	phosphoenolpyruvate carboxylase kinases
PTP	Protein Tyrosine Phosphatase
PWD1	Phosphoglucan Water Dikinase 1
RE	Restriction endonuclease
SBD	Starch Binding Domain
SBE	Starch Branching Enzyme
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEX4	Starch Excess 4
SNF1	Sucrose Non-Fermenting kinase
SnRK	SNF1 related kinases
SSS	Soluble Starch Synthase
TEMED	*N, N, N', N'*-tetramethylethylenediamine
UTP	Uridine Triphosphate

Chapter 1

INTRODUCTION

1.1 PHOSPHATASES

Phosphatases are a class of enzyme, which catalyse the removal of a phosphate group from a substrate. The catalysis results in the formation of a phosphate ion and a molecule with a free hydroxy group. This hydrolysis is the reverse reaction of that mediated by phosphorylases and kinases, which attach phosphate groups to a substrate, often via the use of energetic molecules such as ATP.

Phosphatases can be categorised into two main groups, cysteine-dependent phosphatases (CDPs) and metallo-phosphatases (Barford, 1996). As their names suggest, CDPs require an active site cystine to catalyse the hydrolysis reaction, while metallo-phosphatases require coordination of a specific metal ion(s) in their active site for activity. There are additional enzymes, which fall outside of their groupings, such as those using an active site histidine in place of a cysteine, for example the hPAP protein, but these are much less common.

1.1.1 Protein Phosphatases

The addition or removal of a phosphate group on a protein can have a significant effect, on the structural and thereby functional properties of a protein. Phosphorylation can regulate the interactions between protein partners in a protein complex (Pawson, 1995; Aparicio-Fabre *et al.*, 2006), act directly to regulate enzymatic activity through the initiation of conformation changes in the protein (Johnson *et al.*, 1993; Johnson and O'Reilly, 1996), and is also required by some proteins in order to target regions of the cell in which they function (Guan *et al.*, 2006).

Due to this capability, a large majority of cell functions involve reversible phosphorylation. This includes stress response (Maya *et al.*, 2001; Park *et al.*, 2006), ion transport (Kahle *et al.*, 2005; Delpire and Gagnon, 2006), metabolism, cell cycle, and developmental control.

The majority of protein phosphorylation occurs at serine and threonine residues, with phosphorylation at tyrosine residues being less common. In a small number of cases histidine residues may be phosphorylated. This range of phosphorylation types also results in a range of different phosphatases, each one able to deal with a different phosphorylates residue.

1.1.2 Serine/Threonine Phosphatases

Phosphorylation and dephosphorylation of proteins most commonly occurs at serine and threonine residues. Phosphorylation and dephosphorylation of serine and threonine residues in plant signal transduction accounts for around 97% of the protein phosphorylation events in this pathway (Smith and Walker, 1996).

Serine/Threonine phosphatases were initially divided into 2 groups through their pharmacological properties and substrate specificity (Cohen, 1989). Serine/Threonine phosphatase type 1 (PP1), prefers the Beta-subunit of phosphorylase kinase substrate and is inhibited by nanomolar concentrations of two small peptide inhibitors, inhibitor 1 and 2. In PP1 there is a diverse array of different holoenzymes in which the same catalytic subunit (PP1c) is complexed to distinct regulatory and targeting subunits, which control the activity and specificity of the catalytic subunit (Faux and Scott, 1996; Hubbard and Cohen, 1993). Serine/Threonine phosphatases type 2 (PP2) is preferentially dephosphorylated by the alpha-subunit of phosphorylase kinase and is insensitive to inhibitor 1 and 2. PP2 is further split into 3 subgroups: PP2A, which does not require divalent cat ions for activity (like PP1's); PP2B, which is regulated by Ca2+; PP2C, which is regulated by Mg2+.

Following sequence and structural analysis of the products of identified serine/threonine phosphatases an alternative nomenclature has been suggested. PP1, PP2A, and PP2B are closely related and termed the PPP family, while PP2C, Pyruvate dehydrogenase phosphatase, and several other Mg2+ dependent Ser/Thr phosphatases are similar, and so termed the PPM family (Barford, 1996; Cohen, 1997).

1.1.3 Tyrosine Phosphatases

The tyrosine phosphatase superfamily includes all the protein phosphatases that are capable of hydrolysing phosphotyrosine residues. They have a conserved catalytic motif, $VHCX_5R$, containing a cysteine residue that requires reduced conditions for catalytic activity. Tyrosine phosphatases are divided into tyrosine specific and dual specific phosphatases. The role of tyrosine phosphorylation has not been studied in the same detail as serine/threonine phosphorylation. In animals and lower eukaryotes such as yeast, transduction of signals such as hormones, growth factors and cytokines (Fauman and Saper, 1996; Darnell, 1997; Hunter, 1998), involves the reversible phosphorylation of tyrosine residues. Tyrosine phosphatases, which form a large superfamily that can be divided into families, based on sequence homology and functional properties (Barford, 1995; Barford *et al.*, 1995; Neel and Tonks, 1997), can be further catagorised by their phosphoamino acid specificity.

Tyrosine-specific PTP can dephosphorylate phosphotyrosine but not phosphoserine/threonine, and Dual-specificity PTPs (DsPTP) can dephosphorylate phosphotyrosine and phosphoserine/threonine (also termed DSP - dual-specificity phosphatases). Examples of DsPTPs include MAP kinase phosphatases, cell cycle regulator cdc25 phosphatase (Dunphy, 1994), and tumour suppressor PTEN (Maehama *et al.*, 2001).

Tyrosine-specific PTP's can further be divided into 2 catagories: Receptor like, which generally have extracellular putative ligand-binding domain, a single Trans-membrane region and one or two cytoplasmic PTP domains; Intracellular, which contain single catalytic domains and various amino or carboxyl terminal extensions

including SH2 domains (e.g. SHP1 from Neel and Tonks (1997)) that have targeting or regulatory functions. All PTP's are characterized by a lack of metal ion requirement for catalytic activity, the ability to hydrolyze p-nitrophenyl phosphate, insensitivity to okadaic acid and sensitivity to vanadate.

Until recently no PTP had been characterised in plant systems and the understanding of their roles in the plant was limited. Xu *et al* (1998) reported the characterisation of the first plant PTP, AtPTP1, from *Arabidopsis* cDNA. Soon afterwards this characterisation was improved upon (the first published sequence was found to be chimeric) and homologues of AtPTP1 from non-Arabidopsis species (pea and soybean) were reported (Fordham-Skelton *et al.*, 1999). The first DsPTP, AtDsPTP1 was characterised by (Gupta *et al.*, 1998) and found to negatively regulate an *Arabidopsis* mitogen-activated protein kinase (MAPK), AtMPK4. Similarly, another DsPTP, AtMKP1, has been shown to dephosphorylate and deactivate a MAPK in *Arabidopsis* having a role in genotoxic stress relief (Ulm *et al*, 2001).

1.2 CARBOHYDRATE BINDING MODULES

While there are a number of different carbohydrate binding proteins, such as lectins, antibodies, and a number of glycolytic enzymes, not all are classed as containing a carbohydrate binding module (CBM). CBMs were originally classified as cellulose binding domains, due to the initial discovery of several cellulose binding modules (Gilkes *et al.*, 1988; Tomme *et al.*, 1988). Following the identification of a number of binding modules specific for carbohydrates other than cellulose, these binding modules were reclassified under the name Carbohydrate Binding Modules.

A CBM can be defined as a continuous amino acid sequence within a carbohydrate-associating enzyme with a discrete fold having carbohydrate-binding activity (Boraston *et al.*, 1999; Coutinho *et al.*, 1999). The classifications for CBMs can be found in the Carbohydrate-Binding Module Family Server (CAZy) (http://afmb.cnrs-mrs.fr/~pedro/CAZY/cbm.html), with the binding domains having been classified into over 40 families, from over 50 different species, based on binding specificity, amino acid sequence, and structure, with CBMs containing anywhere from 30 to 200

amino acids. The CBMs can be located anywhere within the protein, but occur with higher frequency at the C- or N- terminals, while also existing as a single, double, or in some cases triple domain within a single protein.

In order to greater understand the mechanism of CBM-carbohydrate recognition and interaction, the three-dimensional structure of members of over 20 CBM families have been studied. These data indicate structural similarities between the families (review see (Boraston *et al.*, 2004), with their carbohydrate binding ability due in part to several aromatic amino acids forming a hydrophobic region within the binding fold.

Using the three dimensional data, it has been possible to further classify CBMs, based on their structure, into one of seven "fold families": "b-Sandwich" (family 1), "b-Trefoil" (family 2), "Cysteine knot" (family 3), "Unique" (family 4), "OB fold" (family 5), "Hevein fold" (family 6), and "Unique; contains hevein-like fold" (family 7) (Boraston *et al.*, 2004). This categorisation however fails to take into account functional similarities, and so they are also classified into three types: "surface binding" (type A), "glycan chain binding" (type B), and "small sugar binding" (type C) (Boraston *et al.*, 2004). While the members of each fold family usually occur in the same functional type, members of the b-sandwich family (family 1) can be found in all of the functional types.

1.2.1 Starch and Glycogen Binding Domains

There are currently seven starch binding domain (SBD) families in the CAZy database (http://afmb.cnrs-mrs.fr/~pedro/CAZY/cbm.html): CMB20, CBM21, CBM25, CBM26, CBM34 CBM41 and CBM45. Those families in which one or more members have had the three dimensional structure studied (CBM20 and CBM34) both fall into the b-Sandwich fold family (family1) as well as the same functional type; "glycan chain binding" (type B) (Boraston *et al.*, 2004). It has recently been proposed that CBM20 and CBM21 be combined to form a larger 'clan' due to their sequence similarity and specificities (Machovic *et al.*, 2005). Initially CBM20 was believed to be located exclusively at the C-terminal end of their parent

proteins, wish CBM21 being located exclusively at the N-terminal end (and often referred to as glycogen binding domain), following work almost 20 years ago, showing the sequence similarities of the two families (Svensson *et al.*, 1989).

The CBM20 family is the most heavily studied of the SBD families, with in excess of 100 proteins identified and the three dimensional structures of 13 studied. The module contains ~100 amino acid residues and has been found linked to a diverse range of catalytic domains. CBM21 is not as heavily characterised as CBM20, with considerably fewer proteins identified, and three-dimensional structure of only two proteins, having been studied.

As with CBM20 and CBM21, CBM25 and CBM26 show some degree of sequence homology, with both families containing less than 20 members. Members of these families often occur in tandem repeats, such as those seen in *Paenibacillus polymyxa* multidomain amylase (Kawazu *et al.*, 1987) and in two Streptomyces a-amylases (Yin *et al.*, 1998; Bentley *et al.*, 2002) in the CBM25 family. In CBM26, there are tandem repeats present in *Lactobacillus amylovorus* and *Lactobacillus plantarum* A6 α-amylases (Giraud and Cuny, 1997).

CBM45 is the most recently catagorised family, and containing only 20 members, all of which are from the plant kingdom. It is composed of CBMs from a number of amylases, but primarily contains the CBMs from the Glucan Water Dikinase (GWD) proteins (GWD1, GWD2, and GWD3/PWD1) (Mikkelsen *et al.*, 2006).

1.3 KINASE INTERACTION SEQUENCE DOMAINS

Kinase Interaction Sequence (KIS) domains are found throughout animal, plant and yeast systems, where they act in the assembly of the Plant SNF1 related kinases (SnRK) multiprotein complexs responsible for regulation of multiple metabolic processes. In plants SnRKs have been shown to regulate the activity of rate limiting metabolic enzymes, including 3-hydroxy-3-methylglutaryl-CoA reductase, nitrate reductase, and sucrose phosphate synthase (Sugden *et al.*, 1999), as well as the

transcription of glucose and stress-regulated genes (Bhalerao *et al.*, 1999). SnRK's are, as their name suggests related to the sucrose non-fermenting kinase (SNF1) of yeast (*Saccharomyces cerevisiae)* and animal AMP-activated protein kinases (AMPK)(Alderson *et al.*, 1991; Mitchelhill *et al.*, 1994; reviewed in Halford *et al.*, 2000). In animals and yeast SNF1 and AMPK's regulate many of the cellular responses to environmental and nutritional stress (Hardie *et al.*, 1998; Halford *et al.*, 2000). The SNF1 and AMPK complexes consist of the kinase and two additional subunits, termed the β and γ subunits, these additional units fulfil both scaffolding and regulatory roles (Jiang and Carlson, 1997; Hardie *et al.*, 1998). The kinase subunits are composed of an N-terminal serine/threonine kinase domain and a C-terminal regulatory domain (Celenza and Carlson, 1986; Gancedo, 1998).

During evolution, the family of plant SnRKs has diverged into three subfamilies: SnRK1a/1b, SnRK2, and SnRK3. Only the SnRK1 subfamily represents direct structural and functional homologues of the SNF1/AMPK family (Halford and Hardie, 1998), while members of the SnRK2 and SnRK3 subfamilies show unique function among plants. The SnRK1 homologs from various plant species have been shown to complement yeast *snf1* defective mutant phenotypes, suggesting its function is evolutionary conserved (Rolland *et al.*, 2006).

The β subunit (SIP1/SIP2/GAL83 in yeast) is made up of two protein interaction domains, allowing it to bind to the kinase and the γ subunit. The first of these domains, termed the KIS (kinase interaction sequence) domain (Jiang and Carlson, 1997), binds to the C-terminal region of the kinase. The second domain, termed the ASC domain (association with SNF1complex) interacts with the γ regulatory subunit (SNF4 in yeast) (Fig. 1.1).

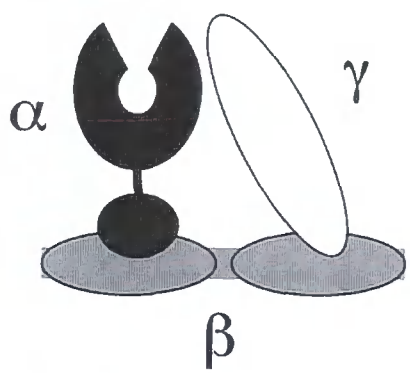


Figure 1.1 Standard Configuration of SnRK/AMPK/SNF1 Complexes. Diagramatic representation of the predicted interactions between the members of the SnRK, AMPK, and SNF1 complexes. Where α represents the catalytic subunit, β represents the scaffold subunit, and γ represents the regulatory subunit

Plant SnRK's form part of the CDPK-SnRK superfamily of protein kinases (Hrabak *et al.*, 2003), this consists of seven types of serine-threonine protein kinases: calcium-dependent protein kinase (CDPKs), CDPK-related kinases (CRKs), phosphoenolpyruvate carboxylase kinases (PPCKs), PEP carboxylase kinase-related kinases (PEPRKs), calmodulin-dependent protein kinases (CaMKs), calcium and calmodulin-dependent protein kinases (CCaMKs), and SnRKs.

Homologues of β and γ SnRK subunits have been identified and characterised in plants (Bouly *et al.*, 1999; Lakatos *et al.*, 1999; Kleinow *et al.*, 2000). However, higher plants have been shown to possess novel proteins, which contain regions similar in sequence and function as the KIS domain, but are fused to other domains not found in the β subunit. The first of these was designated AKINβγ (Lumbreras *et al.*, 2001) and contained an N-terminal KIS domain fused to an SNF4-like polypeptide. The second of these was designated PTPKIS1 (Fordham-Skelton *et al.*, 2002) and contained a C-terminal KIS domain fused to a PTP catalytic domain (Fig. 1.2). This suggests that higher plants have evolved SnRK complexes, which have a different subunit composition from the yeast and animal models (Lumbreras *et al.*, 2001; Fordham-Skelton *et al.*, 2002). This implies that higher plants may have evolved unique regulatory mechanisms, using their SnRKs, to respond to metabolic and environmental stress unique to plants.

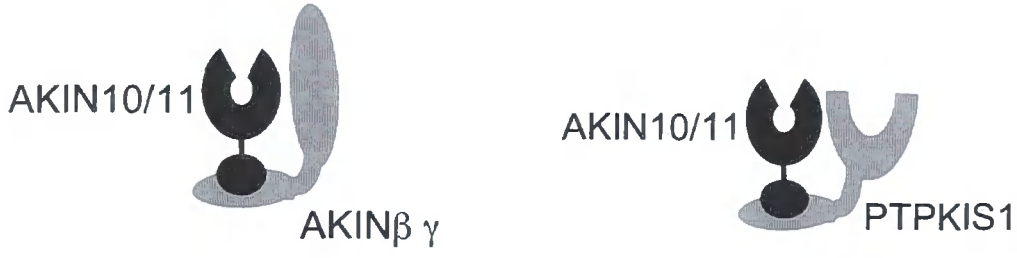


Figure 1.2 Alternative SnRK Conformations Unique to Plants. Diagramatic representation of the predicted interaction between the catalytic subunit of the plant SnRK complex, and AKINβγ or PTPKIS1

A recent paper has reported the presence of a functional glycogen-binding domain in the mammalian AMPKβ subunit of the SNF1 complex (Polekhina *et al.*, 2003). It was found that the new glycogen-binding domain overlapped with the KIS domain and that it was the glycogen-binding portion of this region that was most conserved across species. It was found that the KIS domain was not necessary for the formation of the heterotrimeric complex and it was thus suggested that calling this region of the protein the glycogen binding domain rather than the KIS domain would be more appropriate (Hudson *et al.*, 2003).

There is evidence for variable specificity across the KIS domains of various proteins in plants (Gissot *et al.*, 2006). One such example is AtHSPRO1 and AtHSPRO2 which interact with the AKINβγ KIS domain but not with KIS domains of AKINB1, B2, and B3. AKINB3 contains a KIS domain lacking the previously characterised glycogen binding domain in Polekhina *et al* (2003) and Hudson *et al* (2003) but is still capable of rescuing yeast sip1Δsip2Δgal83Δ mutant. Interaction between AKINβ3 and other AKIN complex subunits from *A. thaliana* were detected by two-hybrid experiments and in vitro binding assays (Gissot *et al.*, 2004) supporting the suggestion that KIS domains do not always posses carbohydrate binding abilities.

1.4 TRANSIENT LEAF STARCH METABOLISM

The energy storing macromolecule α-glucans, glycogen and starch, are widespread throughout a variety of kingdoms: in vertebrates, plants and fungi (Ganesh *et al.*, 2003; Kerk *et al.*, 2006). Their structure is optimised to minimize storage space and maximize energy concentration, and provide a further layer of regulation of biosynthesis and degradation (Manners, 1988; Manners, 1990). In animals and fungi the α-glucan is glycogen, and in plants it is starch. The two forms of energy storage share a lot of common features by being macromolecules, built solely from glucose residues (Manners, 1988; Manners, 1990).

In plants, carbohydrate is stored as large starch particles, consisting only of glucose residues joined by α-1,4 linkages, with some degree of α-1,6 branch points. Starch

contains far fewer α-1,6 branch points than glycogen, with starch usually contains two distinct components; amylose, containing no branch points at all, linear polymer; and amylopectin, containing infrequent branch points. The ratio of amylose and amylopectin varies between plant species. The crystalline starch granules are found in the amyloplasts in heterotrophic tissue and chloroplasts in autotrophic tissue. Starch granules form with internal growth rings, where crystalline lamellae alternate with amorphous lamellae with a frequency around 9nm (Baker *et al.*, 2001; Pilling and Smith, 2003). Buléon *et al.* (1998) state that the crystalline structure is based on interactions between the double helices formed by the branched side chains, while the amorphous regions are formed by the branch points in the amylopectin structure. The crystalline lamellae are approximately 5-7nm, while the amorphous regions are approximately 2-4nm (Blennow *et al.*, 2002). Two different kinds of starch can be identified, namely storage starch such as that present in rice grains, and transient starch present in photosynthetic and autotrophic tissue such as leaves (Buléon *et al.*, 1998; Smith *et al.*, 2005).

1.4.1 Starch Granule Formation

Starch is synthesized in the stroma of the chloroplasts through photosynthetic reactions during the light period. Glucose-1-phosphate is synthesized from triose phosphate via the intermediate fructose-1,6-bisphosphate from the Calvin cycle. The first committed step in starch synthesis is the conversion of glucose-1-phosphate to the precursor of starch synthesis ADP-glucose, catalysed by the enzyme ADP-glucose pyrophosphorylase (AGPase) (E.C.2.7.7.27). This reaction requires ATP and generates pyrophosphate and ADP. This is the critical step in starch synthesis, as it is an irreversible reaction. AGPase is allosterically regulated by direct binding of either the activator 3-phospho-glycerate or the inhibitor orthophosphate.

The next step in the synthesis of starch is the addition of the glucose moiety of ADP-glucose to the nonreducing C4 of the terminal glucose residue of the growing chain (Fig. 1.3). This reaction is catalysed by two different enzymes; for amylopectin it is catalysed by soluble-starch synthase (SSS) (E.C.2.4.1.21), for amylose it is catalysed by granule-bound starch synthase (GBSS) inside the granule (reviewed by

Buléon *et al.*, 1998). There are at least five different isoforms of the SSS, termed SSSI – SSSV.

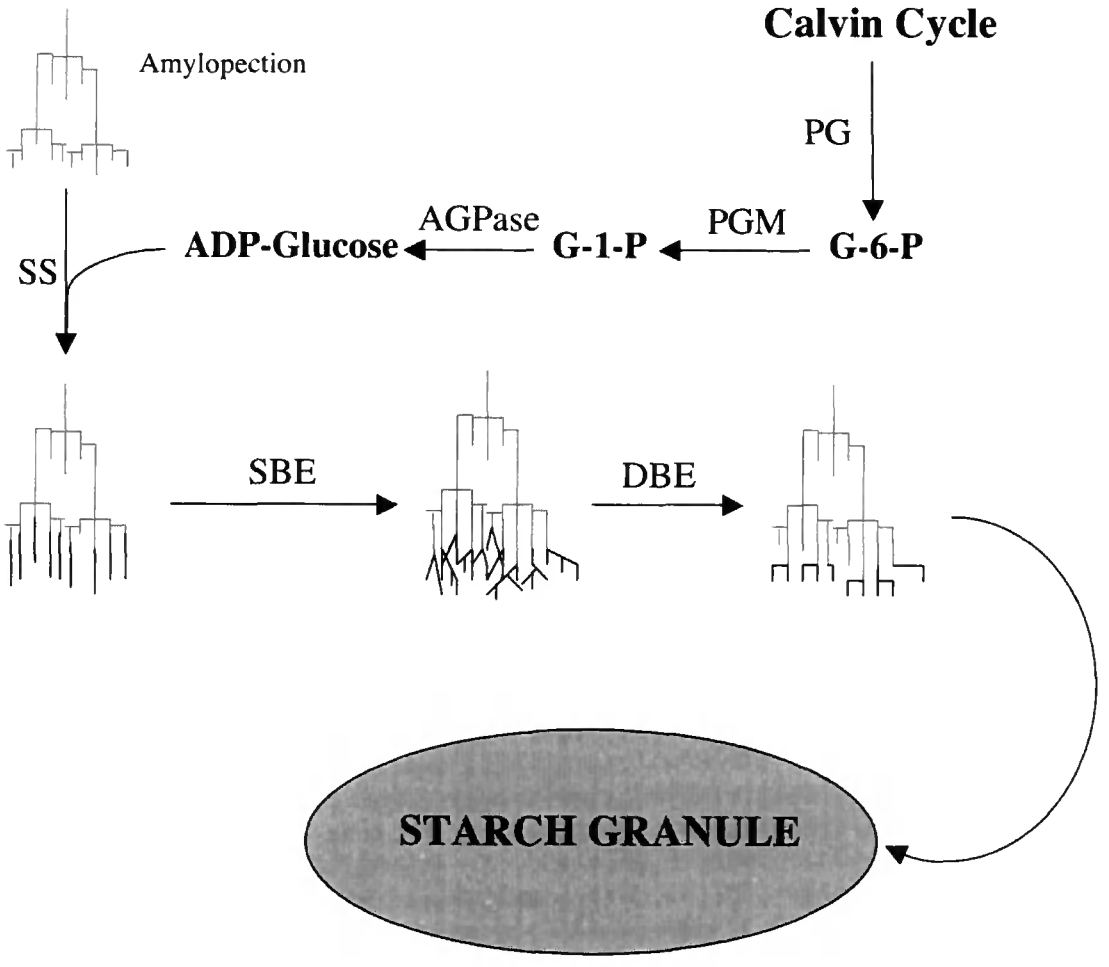


Figure 1.3 Starch Granule Synthesis. Diagramatic representation of starch granule synthesis from the products of the calvin cycle. Glucose-6-Phosphate (G-6-P), Glucose-1-Phosphate (G-1-P), Phospho-glucoisomerase (PG), Phosphoglucomutase (PGM), ADP-glucose pyrophosphorylase (AGPase), Starch Synthase (SS), Starch Branching Enzyme (SBE), Debranching Enzyme (DBE).

An additional enzyme is needed to synthesize the α-1,6 linkage in amylopectin; starch branching enzyme (SBE) (E.C.2.4.1.18). SBE cleaves an α-1,4 glycosidic bond in the reducing end of the glucan chain, and transfers the glucan fragment to a new position, establishing the α-1,6 linkage between the glucan and a glucose residue, resulting in the formation of a new branch. There are several different isoforms of the SBE, which can be divided into two types or families: The A family (or SBEII) and the B family (or SBEI) (Burton *et al.*, 1995). The two isoforms cooperate in forming amylopectin, where the A-family transfers shorter glucan-chains than the B-family. Recently, a third group has been identified in *Arabidopsis*, rice and *Populus* (Han *et al.*, 2007).

According to the so-called glucan trimming model, there are specific isoamylases (ISA) involved in maturation of the starch molecule (Ball *et al.*, 1996; Myers *et al.*, 2000). In *Arabidopsis thaliana*, the debranching enzymes (DBE) ISA1 and ISA2 are suggested to hydrolyse the excess α-1,6 branches from the outer side chains in the outer layers of the granule synthesised by SBE. If these surplus branches are not hydrolysed, the resulting glucan will not be able to crystallise. This can cause disturbance in storing the starch in a granular structure (Reviewed by Ball *et al.*, 1998, and Smith *et al.*, 2005).

1.4.2 Starch Granule Degradation

At night, transient starch is degraded to release carbohydrate compounds that are accumulated through photosynthesis in the daytime. This provides nutrients to the non-photosynthetic parts of the plant, as well as carbon skeletons, energy and reductants in the leaves (Smith *et al.*, 2005). It is a composite process, differing slightly among species, involving several enzymes, some of which are not known, poorly characterised, or the roles of which remain to be identified. This section is concentrating on the mobilisation of transient starch in *Arabidopsis thaliana*, from plastidic granular starch to cytosolic hexose phosphates.

In recent research, the following pathway has been suggested in *Arabidopsis thaliana*, however, the exact order of reactions is not clarified. Based on starch

excess (SEX) phenotypes of knock-out plants (as reviewed by Smith *et al.*, 2005), it has been established, that GWD, Isoamylase 3 (ISA3) and β-amylase are required for starch degradation, while α-amylase activity has been shown not to be required for degradation of leaf starch.

In *Arabidopsis thaliana*, starch degradation starts with starch phosphorylation, through GWDs. An autophosphorylation of an internal histidine takes place in the enzyme glucan water dikinase 1 (GWD1) (E.C.2.7.9.4), using the β-phosphate from ATP. The autophosphorylated enzyme then transfers the phosphate group to the C3- or C6-position of the glucose residue in the amylopectin chain (Reviewed by Smith *et al.*, 2005). After this phosphorylation, glucan water dikinase 3 (GWD3) (E.C.2.7.9.4), which requires a pre-phosphorylated substrate, can also act upon the glucan chain. GWD3 phosphorylates glucose residues at the C-3 position in the glucan molecule (Baunsgaard *et al.*, 2005; Kotting *et al.*, 2005). The phosphorylation of the starch molecule is required for the break down of starch. This requirement may be due to phosphorylation increasing starch solubility (Blennow *et al.*, 2002; Baunsgaard *et al.*, 2005). The phosphorylation of the starch granule by GWD1 and GWD3, allows other enzymes to degrade the molecule (Blennow *et al.*, 2002; Lloyd *et al.*, 2005; Edner *et al.*, 2007).

β-amylase (E.C.3.2.1.2) is an exo-acting enzyme responsible for hydrolysing α-1,4 glycosidic linkages in the non-reducing end of polyglucans, yielding β-maltose However, the β-amylase lacks the ability to cleave α-1,6 bonds. This is believed to be carried out by the isoamylase 3 (ISA3, E.C. 3.2.1.68), which hydrolyses the small α-1,6 glycosyl branching points on the chain, allowing the β-amylase to further degrade the amylopectin to maltose units (Delatte *et al.*, 2006). The maltose molecules are then exported from the chloroplast to the cytosol via the maltose exporter MEX1 (Niittyla *et al.*, 2004) for further metabolism (Smith *et al.*, 2005).

As β-amylase can only work on chains longer than four glycosyl residues, this enzymatic digestion is also expected to yield a second product, maltotriose. One of the few reported genes in the plant genome producing enzymes capable of degrading maltotriose is disproportionating enzyme 1 (DPE1), an α-1,4 glucanotransferase

(E.C.2.4.1.25). DPE1 transfers one glucose unit from the non-reducing end of the donor molecule to the acceptor, a linear glucan, thereby releasing one glucose molecule (Reviewed by Smith *et al.*, 2005). DPE1 is located in the stroma of the chloroplast, whereas another disproportionating enzyme also taking part in the starch degradation process has been localised to the cytosol. This is the cytosolic transglucosidase (DPE2) (E.C.2.4.1.25) (reviewed by Smith *et al.*, 2005). DPE2 acts on the maltose transfered to the cytosol through the MEX1 transporter in the chloroplast membrane, transferring one glucose unit from maltose to a glucan acceptor, suggested to be an α-1,4, α-1,6 heteroglucan. The reaction yields the release of one free glucose unit for further cellular metabolism, and the lengthening of the heteroglucan by one glucose residue (Smith *et al.*, 2005), which may be degraded by cytosolic amylases.

1.4.3 Starch Metabolism Related Mutations

One method used in order to understand the proteins involved in starch granule degradation and formation is to identify mutant plants in which starch metabolism is altered, and identify the mutated gene responsible for this. This approach has identified a number of genes that would be expected to be involved in starch metabolism as well as some where links were less obvious.

Mutations in dbe1 (ISA2) lead to the formation of starch that is highly branched, has shorter branches and is water-soluble (Zeeman *et al.*, 1998). Since these altered properties of the glucan resemble those of glycogen, the polymer is termed phytoglycogen. The content of glucans is less then half of the starch content in the corresponding wildtype, however about 20% of the glucans is normal starch. Of interest, the phytoglycogen appears to be as accessible to the plant in the same way as starch, while it is the reduced amount of glucan, which affects the phenotype negatively (Zeeman *et al.*, 1998).

The starch excess phenotype of *sex1* is caused by a mutation in the GWD1 locus. This hinders the initial phosphorylation of the starch granule and results in the accumulation of starch. In addition, the *sex1* mutant has reduced growth in

comparison to the corresponding wildtype, further emphasising that the plant has diminished access to its energy reserves (Yu *et al.*, 2001).

Analysis of mutants in the single chloroplastic α-amylase in *A.thaliana* (AtAMY3) showed no alteration in transitory starch breakdown in leaves, something also seen in triple knockout mutants of all three *A.thaliana* α-amylase genes (Yu *et al.*, 2005). This suggested that starch degradation in leaves was likely to require a combined action of β-amylase and debranching enzyme, instead of the previously predicted activity of α-amylase.

The identification of the chloroplastic maltose transporter (MEX1) (Niittyla *et al.*, 2004) supported the surprising findings of the α-amylase mutants. Mutations of MEX1 cause an increase in chloroplastic maltose levels, due to an inability to transport maltose into the cytosol. This supports the hypothesis that β-amylase acts upon the starch granule (producing maltose) instead of α-amylase, which would result in glucose production. Allowing a means for the maltose produced by β-amylase activity upon starch granules, to be transported to the cytosol for further degradation.

1.5 COMPARISON OF STARCH AND GLYCOGEN METABOLISM

Glycogen is a heterogeneous, irregular structure containing up to 120,000 glucose units per molecule, interlinked by α-1,4 bonds and branched on average on every 10^{th} to 14^{th} residue by α-1,6 glycosidic linkages (Manners, 1990). Structurally, glycogen resembles amylopectin, one of the two polyglucans in starch (Buléon *et al.*, 1998). About 20-30% of a starch granule is made from the linear α-1,4 amylose chains consisting of 600-6000 glucose residues internally linked by α-1,4 glycosidic linkages. The remaining 70-80% of starch is made from α-1,6 branched polyglucan amylopectin, where the glucan chains is built from 15-30 α-1,4 linked glucose units, connected by α-1,6 linked branch points (Manners, 1988).

Both glycogen and starch are built in a comparable tree-like structure, with one reducing end, and many nonreducing ends. The nonreducing ends have a free hydroxyl group on the C4 of the last glucose unit, from which the molecule can be elongated. The relatively long side chains of amylopectin cause the starch granule to crystallise, while the relatively shorter side chains of glycogen do not produce this structure. The glycogen molecule is therefore non-crystalline, and stored in granules measuring 10 – 40 nm in diameter. These are localised in the cytosol of somatic and some neuronal cells.

Glycogen reserves are found predominantly in the skeletal muscle and the liver. With a glycogen content of 95mg per g tissue, the liver is the organ with the highest concentration of glycogen in the body, followed by skeletal muscles, which have a glycogen content of 12.8 mg per g tissue, and the heart, which has a glycogen content of 5.9mg per g tissue (Alonso *et al.*, 1995). But as the skeletal muscles are much more abundant in the body, in assembly, the largest amount of glycogen is localised here.

To synthesize glycogen from glucose, glucose must be in its activated form, UDP-glucose, unlike in starch synthesis where ADP-glucose is used. This metabolite is synthesized from glucose-1-phosphate and uridine triphosphate (UTP), catalyzed by the enzyme UDP-glucose pyrophosphorylase (E.C. 2.7.7.9). The newly formed UDP-glucose is transferred to a nonreducing end of the glycogen molecule, and this reaction is catalyzed by the glycosyltransferase, glycogen synthase (GS, E.C. 2.4.1.11). In mammals, two isoforms of GS exist. One isoform is expressed in many tissues, muscles being the predominant tissue, and the other isoform is expressed in the liver (reviewed by Roach, 2002). The GS proteins show homology with the SS proteins found in plants, which carry out a similar role.

GS can only add glucose residues to an existing chain. Hence, the GS needs an initiating protein. The priming protein of glycogen synthesis is a glycosyltransferase termed glycogenin (E.C.2.4.1.186), on which an oligosaccharide chain of up to 10 glucose residues can be synthesised. This oligosaccharide chain attached to the glycogenin is the primer of glycogen synthesis. Humans have two isoforms of the glycogenin. Glycogenin 1 is the predominant form, and is expressed in the muscles.

Glycogenin 2 is expressed in the liver, in the heart, and to a less degree in the pancreas (reviewed by Roach, 2002). So, this tissue specificity supports that glycogen metabolism is regulated differently in the respective tissues, with the different isoforms of the enzymes. This differs from starch formation in plants, where no protein primer has been identified.

The branch points of the glycogen molecule are synthesised by the glycogen branching enzyme (GBE, E.C. 2.4.1.18). GBE cleaves an α1,4 glycosidic linkage, thus excising an oligosaccharide which GBE subsequently transfers back on the glycogen branch and ends up forming a new α-1,6 glycosidic linkage between the liberated chain and another outer chain of the glycogen molecule in the same way as the starch branching enzyme of plants. This reaction provides new non-reducing ends for the glycogen synthase to act upon (reviewed by Roach, 2002).

The enzyme primarily responsible for the degradation of glycogen is the glycosyltransferase termed glycogen phosphorylase (GP, E.C.2.4.1.1). This enzyme phosphorolyses the end glucose residue from the non-reducing end by adding orthophosphate to the C-1 of the glucose unit, thus cleaving the glycosidic α-1,4 bond, yielding glucose 1-phosphate and a glycogen molecule shortened by one glucose residue. While similar proteins are present in plants (termed starch phosphorylase) they have been shown to play only a minor role in starch degredation.

The GP cannot cleave α1,4 glycosidic linkages close to the branch points, and additional enzyme activity is needed to continue the degradation. This 'trimmed' glycogen molecule, with its shortened branches acts as the substrate for glycogen debranching enzyme (GDBE) (E.C.3.2.1.68). The GDBE has two catalytic activities, a transferase activity and an α-1,6 glucosidase activity. The transferase part of the enzyme catalyses the hydrolysis of an α-1,4 glycosidic linkage in a branch of the glycogen molecule, excising a maltotriose and leaving a single glucose residue on the branch. Then, the enzyme transfers the excised maltotriose to the longer chain, forming an α1,4 glycosidic linkage. The second activity of the GDBE is an α-1,6 glucosidase activity, removing the remaining glucose unit by cleaving the branch point. When the glycogen molecule is linearized, the GP can carry on the

degradation (Roach, 2002), as is seen in starch degredation with the inter play of beta-amylases and ISA3.

The enzyme acid α-glucosidase (E.C.3.2.1.20) catalyses the hydrolysis of both α-1,4 and α-1,6 glycosidic linkages, (Roach, 2002) and is thus capable of complete degradation of glycogen in place of the above enzymes.

Apart from the obvious similarity between the two energy storage forms, there are differences in their synthesis and degradation, however, as comparison of the two systems continues, it grows more obvious that what was often seen as very different systems, share a great deal of commonality, not just at the level of the primary enzymatic proteins, but also through their regulation, and the other proteins and mechanisms involved in the macromolecular carbohydrates degradation and formation

1.5.1 Glycogen Metabolism Related Mutations

Mutations in enzymes involved in mammalian glycogen metabolism generally tend to cause a glycogen excess phenotype, in preference to a glycogen lacking phenotype. This may indicate that a storage disease leading to insufficient amounts of glycogen may be fatal to mammals. There are a number of genetically determined glycogen metabolism related disorders, which make up the family of glycogen storage diseases (GSD), along with Lafora disease.

GSD type IV, also termed Andersens disease, is a deficiency of GBE caused by mutations of the coding gene. This leads to the formation of glycogen containing long unbranched glucose chains, which have a low solubility, and can precipitate out. As a result, polyglucosan bodies are formed in the central nervous system (CNS) in astrocytes, but not in the neurons. Disease onset occurs in infancy, and manifests as liver cirrhosis, cardiomegaly and muscle involvement (Raben *et al.*, 2001). The patients generally die in early childhood (Minassian *et al.*, 2002).

Adult polyglucosan body disease (APBD) is a progressive neurologic disorder. In Ashkenazi Jewish patients, the disease caused by mutations in GBE leads to less severe symptoms; in non-Ashkenazi Jewish patients, the disease is not observed in association with reduced activity of the GBE, indicating that other genes may be involved. Polyglucosan bodies accumulate in the axons, and not in the perikarya and the dendrites as in Lafora Disease (Minassian *et al.*, 2001). The disease has a late onset, and manifests as peripheral neuropathy and CNS dysfunction (Raben *et al.*, 2001), motor and sensor control is affected along with the onset of dementia, but seizures or myoclonias do not occur (Minassian *et al.*, 2001).

Raben *et al.*, (2001) argue that since all these glycogen storage diseases, characterized by the presence of polyglucosan bodies, are caused by mutations affecting either the GBE or the GS, polyglucosan bodies arise from an imbalance in the level of activity between these enzymes. This theory may help to not only understand glycogen storage disease, but also other carbohydrate storage diseases such as the starch excess mutations in plants.

1.5.2 Lafora Disease and Laforin

Lafora Disease (LD) is an autosomal, recessive Progressive Myoclonus Epilepsy (PME), which was first described by Gonzales Lafora in 1911. The disease is associated with myoclonic seizures, intellectual decline and a fatal outcome (Reviewed by Minassian, 2002). The lafora bodies consist of 80-93% crystallized α1,4- α1,6 glucans and 6% protein (Ganesh *et al.*, 2006). They are mainly found in the cytoplasm in the perikaryal region of neurons and in the dendrites, but not in the axons (Chan *et al.*, 2004). The Lafora bodies are believed too dense to be degraded by enzymes during mobilisation of glycogen, while its precipitation in the cytoplasm suggests at features in common with plant starch (Wang and Roach, 2004). Unlike other glycogen storage disorders, it is not caused by mutation in genes encoding GBE or GS, but by one in a gene encoding at least one of the proteins; Laforin (product of EPM2A gene), Malin (product of EPM2B gene) (Gentry *et al.*, 2005; Worby *et al.*, 2008) or a yet to be identified protein (product of EPM2C gene).

LD sets in during the first or second decade of life, with a clinical course of epileptic seizures, rapid intellectual decline, myoclonic and tonic-clonic seizures caused by progressive central nervous system degeneration. Also severe motor and coordination deterioration, focal occipital seizures and constant myoclonus is observed, and death will set in within ten years of onset (Gentry *et al.*, 2005). One of the characteristics of the disease is the presence of lafora bodies in the neurons of several tissues including brain, liver, skeletal muscle and heart (Reviewed by Minassian, 2002).

The laforin protein is a phosphatase with a carbohydrate binding domain, which preferentially binds to polyglucans (Chan *et al.*, 2004). It is involved in glycogen metabolism in mammals, and the gene coding for Laforin has been localised to chromosome 6q24 by homozygosity mapping and linkage analysis (Minassian *et al.*, 1998; Fernandez-Sanchez *et al.*, 2003). The gene product is a 331 amino acid protein, in the C-terminal encoding a DSP, while the N-terminal contains a CBM20 motif.

Using alignments of the Homologous laforin sequences from the mammalian, the avian and the fish genome, 13 of the 14 known LD missense mutations in EPM2A were located in residues conserved among the laforin homologues. Furthermore, it was found that these mutations for the most part are in the phosphatase domain or in the CBD (Ganesh *et al.* 2004)

Previous work (Fordham-Skelton *et al.*, 2002) has shown sequence similarity between the N-terminal PTP domain of PTPKIS1, and the C-terminal phosphatase domain of laforin. At the time, this was the only similarity drawn, since the PTP domain in laforin was associated with a carbohydrate binding domain and that of PTPKIS1 associated to the predicted KIS domain. With the evidence that the KIS domain of PTPKIS1 can itself act as a carbohydrate binding module (Kerk *et al.*, 2006; Niittyla *et al.*, 2006), the homology seen in the PTP domain sequence becomes even more relevant, since the two proteins both contain both PTP and CBM domains. Phosphatases with carbohydrate binding modules, similar in structure to laforin or PTPKIS1, are present throughout a range of different organisms including protists, evolved from a progenitor red alga (Gentry *et al.*, 2007). The assumed

functional similarity between laforin and PTPKIS1 is also reflected in a similar phenotype resulting from inactivation of the proteins. While a deleterious mutation in the gene encoding AtPTPKIS1 leads to the SEX4 phenotype, in which large starch granules form with high levels of un-branched amylose are present, similar mutations in the genes producing mammalian laforin result in the formation of lafora bodies, themselves high in un-branched amylose (Chan *et al.*, 2004).

In addition to its similarity to laforin, the phosphatase domain of PTPKIS1 was also shown to have sequence similarity to two other proteins in *Arabidopsis thaliana.* These proteins, encoded by genes At3g01510 and At3g10940 were also identified in the publication by Fordham-Skelton *et al,* (2002) as having homologous PTP regions, with At3g01510, containing a C-terminal domain similar to that of the KIS domain in PTPKIS1.

1.6 AIMS AND OBJECTIVES

The aim of this project was to study the previously identified PTPKIS1 proteins as well as the paralogues encoded by At3g01510 and At3g10940, with regards to their role as phosphatases and any carbohydrate binding ability they may exhibit. Through expression of these as recombinant proteins it was possible to study their ability to dephosphorylate a number of substrates and characterise their binding to carbohydrates *in vitro.* In addition, analysis of their role *in vivo* was carried out through analysis of transcription levels and metabolite profiling of mutant lines in which the PTPKIS1 enzyme or the product of the At3g01510 gene were not present.

Chapter 2

MATERIALS AND METHODS

2.0.1 Chemicals and Reagents

All chemicals and reagents were supplied by BDH Chemicals Ltd (Poole, Dorset, UK) and Sigma Chemical Company (St Louis, USA). Chemicals and reagents were of analytical grade, or best commercially available.

2.0.2 Plant Material

Arabidopsis thaliana ecotype Columbia and mutant lines were grown from seed in growth chambers at 20°C, 70% humidity and a photon flux density of 120 μmol photons m^{-2} sec^{-1} with an 8-h or 12-h photoperiod. SEX4-1, SEX4-3 and PTPKIS2-SALK lines were kindly donated by Alison Smith (JIC).

2.0.3 Standard Molecular and Biochemical Techniques

All standard techniques are as used in the Department of Biological and Biomedical Sciences, Durham University, and were based on protocols in Molecular Cloning: A Laboratory Manual (Sambrook *et al.*, 2001); unless stated otherwise. Equipment for routine DNA work was sterilised by autoclaving. While working with RNA all equipment was treated with 0.1% (v/v) diethyl pyrocarbonate (DEPC) overnight at 37°C and then autoclaved.

2.1 STANDARD DNA WORK

2.1.1 Plasmid DNA Mini-Preparations

Plasmid DNA was isolated from up to 10 ml of overnight *E.coli* culture with the Wizard® Plus SV Miniprep kit (Promega), according to manufacturers instructions. In brief bacteria were alkaline lysed and genomic DNA was precipitated. Plasmid DNAs were recovered in a spin column and washed with ethanol. Plasmid DNAs were eluted in water. This method consistently gave plasmid DNA concentrations of 100-250 ng/μl. Isolated plasmid DNA was stored at –20°C.

2.1.2 Restriction Endonuclease Digestion

Restriction endonuclease (RE) digests were carried out using buffers and temperatures recommended by the manufacturers. Typical digests were carried out at 37°C on 2 μg of DNA using 2-10 units of RE (under optimal conditions 1 U of RE will completely digest 1 μg of DNA in a 50 μl reaction volume). Restriction products were separated by agarose gel electrophoresis and visualised under UV light following ethidium bromide staining.

2.1.3 Agarose Gel Electrophoresis of DNA

Mixtures of DNA were separated according to size by submarine agarose gel electrophoresis (NBL medium size gel apparatus and Pharmacia GNA 100 apparatus), as described by Sambrook (2000). Gels contained agarose up to a concentration of 2% (w/v) in 1X TAE buffer (20 mM glacial acetic acid, 0.2 mM EDTA, 40 mM Tris, pH 7.2) and ethidium bromide (0.5 μg/ml). Prior to loading, DNA samples were mixed with 1/6 sample volume of 6X loading dye (10 mM Tris-HCl at pH 8.0, 10 mM EDTA, 30% (w/v) glycerol, 0.1% (w/v) orange G). Samples were loaded onto horizontal gels submerged in 1X TAE buffer containing ethidium bromide (0.5 μg/ml) and run at room temperature between 50-100V (constant voltage). Size markers covering the appropriate molecular range were run alongside DNA samples. Size markers included HindIII or Eco471 digested lambda DNA

(MBI-Fermentas). The ethidium bromide stained gels were viewed using either a Bio-Rad Gel Doc 2000 imaging system with images captured on a PC running Quantity One software (Bio-Rad), or a Gene Flash gel imaging system (Syngene).

2.1.4 Recovery of DNA Fragments From Agarose Gels

DNA gels were visualised on a trans-illuminator (UVB, λ 302 nm) and appropriate bands were excised by using a single edged razor. Excess agarose was trimmed from the gel piece. DNA was purified from the agarose by using QIAquick gel extraction kit (Qiagen), or a Perfectprep Gel Cleanup kit (Eppendorf AG) according to manufacturers instructions. In brief, gel pieces up to 400 mg were dissolved in a chaotropic agent at 50°C. The dissolved agarose was loaded onto a silica spin column, which binds up to 10μg DNA. Following ethanol washes the DNA was eluted in 50μl nuclease-free water 30μl elution buffer. Eluted DNAs were stored at -20°C.

2.1.5 DNA Precipitation

DNA was precipitated from solution by adding 1/10 volume of 3 M sodium acetate (pH 5.2) and 2 volumes of ice-cold ethanol. The tube was vigorously mixed and then placed at –20°C for at least 1 hour. For small amounts of DNA (<0.1 μg/ml) glycogen was added as a carrier molecule (final concentration 50μg/ml) and MgCl2 (final concentration 10mM). Precipitated DNA was pelleted by centrifugation (13,000 x g for 15 minutes, 0°C). The supernatant was removed and the pellet was washed in 70% (v/v) ethanol. Following centrifugation (13,000 x g for 5 minutes, 0°C) the supernatant was removed and the pellet was dried in a 37°C block for 10 minutes OR in a vacuum. The DNA pellet was then dissolved in an appropriate volume of nuclease free water.

2.1.5 DNA Ligation

Following restriction enzyme digestion, DNA for ligation was separated by agarose gel electrophoresis and purified from the gel. DNAs with compatible ends for

ligation were set up in standard 10 μl reactions using commercially available T4 ligase and buffer (Promega), according to manufacturers instructions. A typical ligation reaction used 100 ng of plasmid DNA, in a 1:3 molar ratio of vector:insert. Sticky-end ligations were incubated at 22°C for a minimum of 3 hours or at 4°C overnight, while blunt end ligations were left overnight at 4°C. Following ligation the DNA was used to transform *E.coli* cells of an appropriate strain.

2.1.6 Nucleotide Sequence Analysis

Nucleotide sequence was determined by DBS Genomics (School of Biological and Biomedical Sciences, University of Durham), using BigDye Terminator with AmpliTaq DNA polymerase (ABI Biosciences). Reaction products were analysed on automated sequencers (ABI Prism 373 STRETCH and ABI Prism 377 XL). Expression constructs were completely sequenced on both strands of the DNA. Contiguous sequences were produced in SequencherTM Version 4.1.2 (Gene Codes Corporation) on a Macintosh computer. Edited nucleotide data was used in BLASTX similarity searches against the NCBI database (http://www.ncbi.nlm.nih.gov/BLAST/) and identification of sequence features in encoded polypeptides was performed using the CBS prediction servers (http://www.cbs.dtu.dk/services/).

2.1.7 Oligonucleotides

Oligonucleotides were chemically synthesised on a solid support by TAGN Ltd (International Centre for Life, Newcastle; http://www.vhbio.com/tagn/) or Sigma-Genosys Ltd. Upon receipt primers were resuspended in sterile water to a final concentration of 100pmol/μl and stored at –20°C. In PCR reactions non-degenerate primers were used at a final concentration of 0.2 μM. Melting temperature (Tm) was calculated using the following formula: Tm = 69.3+ 0.41x(%G+C)-650/(nA+nG+nC+nT). The annealing temperature (Ta) of a particular oligonucleotide was commonly 3°C below the calculated Tm.

2.1.8 Standard Polymerase Chain Reaction

PCR reactions were set up on ice according standard procedures. PCR reactions (25-100 μl) were set up in thin walled PCR tubes and consisted of 0.2mM each of dATP, dGTP, dCTP and dTTP, 2.5mM $MgCl_2$, 1X PCR buffer (50mM KCl, 0.1% (v/v) Triton X-100, 10mM Tris-HCl, pH 9.0), DNA template (5-100 ng) and 1 Unit of Taq polymerase (Promega) per 50µl reaction (where 1U calatalyses the incorporation of 10 nmol of dNTP into an acid insoluble form at 74°C). Gene specific oligonucleotide primers were used at 0.2μM. For multiple PCR reactions a master mix containing all common components was mixed in a 1.5ml eppendorf tube, and then aliquoted to individual PCR tubes. Setting up multiple PCR reactions in this way reduced pipette error and ensured consistency across all reactions. PCR thermocycling was performed on a Perkin Elmer 2400 thermal cycler.

2.1.9 Subcloning PCR Products for pCR2.1 TOPO Cloning System

PCR products were resolved by agarose gel electrophoresis and visualised by ethidium bromide staining. Bands corresponding to the predicted size were cut from the gel and purified. Purified DNA was resuspended in a minimal volume of water and subcloned into the pCR2.1 TOPO vector using the TOPO cloning kit (Invitrogen) according to the manufacturers instructions. Transformed clones were selected on the basis of blue/white colour selection, picked using sterile tooth picks and grown overnight at 37°C in 10 ml LBM broth containing 50μg/ml kanamycin. Liquid cultures were placed on a rotary shaker set at 200 rpm. Plasmid DNA from positive transformants was isolated by mini-preparation (Promega) and checked by restriction enzyme digest and sequencing.

2.1.10 Subcloning PCR Products for Gateway pENTR TOPO System

Primers were designed according to the guidelines provided in the pENTR directional TOPO cloning kit (Invitrogen). PCR was carried out using Phusion polymerase (Finnzyme), According to manufacturers instructions, PCR products were resolved by agarose gel electrophoresis and visualised under UV light

following ethidium bromide staining. Bands corresponding to the predicted size were cut from the gel and purified. Purified DNA was resuspended in a minimal volume of water and subcloned into the pENTR/D-TOPO vector using the pENTR directional TOPO cloning kit (Invitrogen) according to the manufacturers instructions. Transformed clones were selected on plates containing 50μg/ml kanamycin, picked using sterile tooth picks and grown overnight at 37°C in 10 ml LBM broth containing 50μg/ml kanamycin. Liquid cultures were placed on a rotary shaker set at 200 rpm. Plasmid DNA from positive transformants was isolated by mini-preparation (Promega) and checked by restriction enzyme digest and determining of the nucleotide sequence.

2.1.11 Gateway BP Reaction

Primers were designed according to the guidelines provided in the Gateway Technology Manual (Invitrogen) containing attB sequences in each primer. PCR was carried out using Phusion polymerase, PCR products were resolved by agarose gel electrophoresis and visualised by ethidium bromide staining. Bands corresponding to the predicted size were cut from the gel and purified. Purified DNA was then cloned into the pDONR201 vector (Invitrogn) Using the BP Clonase Enzyme Mix (Invitrogen) according to the manufacturers instructions. Transformed clones were selected on plates containing 50μg/ml kanamycin, picked using sterile tooth picks and grown O/N at 37°C in 10 ml LBM broth containing 50μg/ml kanamycin. Liquid cultures were placed on a rotary shaker set at 200 rpm. Plasmid DNA from positive transformants was isolated by mini-preparation (Promega) and checked by restriction enzyme digest and sequencing.

2.1.12 Gateway LR Reaction

Plasmids produced through either the pENTR TOPO method or BP reaction method were used in order to subclone the PCR product they cloned into the expression vectors pDEST15 and pDEST17. The pENTR or BP clones were mixed with the pDEST15 or pDEST17 plasmid, in the presence of the LR clonase mixture, according to the manufacturers instructions. Following transformation into *E.coli*,

Transformed clones were selected on plates containing 75μg/ml carbenicillin, picked using sterile tooth picks and grown O/N at 37°C in 10 ml LBM broth containing 75μg/ml carbenicillin. Liquid cultures were placed on a rotary shaker set at 200 rpm. Plasmid DNA from positive transformants was isolated by mini-preparation (Promega) and checked by restriction enzyme digest and sequencing.

2.1.13 Site Directed Mutagenesis

Specific mutations were generated using the Stratagene "Quick Change" mutagenesis kit. Primers were designed according to the Quick Change Mutagenesis Kit instructions (Stratagene). Following mutagenesis plasmid DNA from positive transformants was isolated by mini-preparation (Promega) and checked by restriction enzyme digest and sequencing.

2.1.14 Colony PCR

Colony PCR was used as a method for rapid screening of recombinant plasmids. Ten to twenty colony transformants from recombinant *E. coli* were picked and resuspended in 10 ml distilled water; this was then used in a standard PCR reaction as previously described.

2.2 STANDARD RNA WORK

2.2.1 RNA Isolation

RNA was isolated from leaf material using the RNeasy Plant Mini Kit (Qiagen) according to the manufacturers instructions. RNA was quantified by spectral absorbance at 260 and 280 nm (A260/280).

2.2.2 Reverse Transcription-PCR (RT-PCR)

RT-PCR was carried out using the Access To RT-PCR kit (Promega) according to the manufacturers instructions. PCR thermocycling was performed on a Perkin Elmer 2400 thermal cycler. Reaction products were seperated using agarose gels and ethidium bromide stained gels were viewed using either a Bio-Rad Gel Doc 2000 imaging system with images captured on a PC running Quantity One software (Bio-Rad), or a Gene Flash gel imaging system (Syngene).

2.2.3 Real-Time RT-PCR

Following quantification by spectral absorbance, 2μg of total RNA was treated with deoxyribonuclease1 (DNaseI) according to the manufacturers instructions. 2μg RNA, 2μl 10x reaction buffer and 2μl amplification grade DNaseI were combined in a total volume of 20μl with RNase free water, and incubated for 15 minutes at room temperature. Following incubation, 2μl of stop solution were added and the reaction heated at 70°C for 10 minutes before being chilled on ice. 2μl was retained for use as controls, while the remainder was used in cDNA synthesis.

cDNA synthesis was carried out using iScript (Bio Rad), using 10μl of the DNaseI treated RNA in each reaction according to the manufacturers instructions. Reactions were first incubated at 25°C for 5 minutes, followed by a 30 minute incubation at 42°C and finally 5 minutes at 85°C prior to being frozen at -20°C until used in real-time RT-PCR reactions.

The real-time RT-PCR reaction was performed on an iCycler Instrument (Bio-Rad) using the iQ SYBR Green Supermix kit for PCR (Bio-Rad) according to the manufacturer's instructions. Each reaction was performed with 5μl of between 1:10 and 1:100 dilutions of the first-strand cDNA (concentrations being standardised following quantification by spectral absorbance) in a total volume of 20μl.

The reactions were incubated at 96°C for 2 min to activate the hot start recombinant Taq DNA polymerase, followed by 50 cycles of 30 sec at 96°C, 30 sec at 60°C and 1 min at 72°C. Gene-specific primers (listed in the relevant sections) were used. Specificity of the PCR amplification was confirmed with a heat dissociation protocol (melt curve) (from 65 to 95°C) following the final cycle of the PCR. The efficiency of the primer sets was calculated by performing real-time PCR on several dilutions of first strand cDNA. The efficiency of the different primer sets was found to be comparible. The results obtained for the different conditions analyzed were standardized to the ACTIN1 Real Time RT-PCR product level.

2.3 STANDARD BACTERIAL WORK

2.3.1 *Escherichia coli* Strains

Escherichia coli strains were used for plasmid propagation. Plasmid DNAs were maintained at high copy numbers in the following *E.coli* strains: DH5α (genotype: F$^-$ ϕ80d*lac*ZΔM15, Δ(*lac*ZYA-*arg*F)U169, *deo*R *rec*A1, *end*A1, *hsd*R17(r_K^-, m_K^+), *pho*A, *sup*E44, , λ$^-$, *thi*-1, *gyr*A96, *rel*A1) and TOP10 (genotype: F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *deo*R *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G). DH5α and TOP10 cells were purchased from Gibco BRL and Invitrogen, respectively.

Escherichia coli strain DB3.1 was used for propagation of Gateway Destination Vectors containing the CcdB gene (genotype: F- gyrA462 endA1 Δ(sr1-recA) mcrB mrr hsdS20(rB-, mB-) supE44 ara14 galK2 lacY1 proA2 rpsL20(Smr) xyl5 Δleu mtl1). DB3.1 cells were purchased from Invitrogen.

For over expression of recombinant proteins expression plasmids were transferred to an *E.coli* host containing a chromosomal copy of the T7 RNA polymerase gene under *lacUV5* control or the arabinose-inducible *araBAD* promoter. These strains were BL21 STAR (DE3) (genotype: F$^-$ *ompT hsdS*$_B$ ($r_B^-m_B^-$) *gal dcm rne131* (DE3)) and BL21-AI (genotype: F$^-$ *omp*T *hsd*S$_B$ (r_B^-*m*B$^-$) *gal dcm ara*B::T7RNAP-tetA)

Respectively. BL21 STAR (DE3) and BL21-AI cells were purchased from Invitrogen.

2.3.2 Bacterial Culture

For routine work, bacteria were propagated in sterile Luria-Bertani Medium (LBM) (10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl). Liquid cultures were inoculated with single bacterial clone and grown overnight (approx. 16 hours) at 37°C, with shaking at 220 rpm. For growth of bacteria on solid LBM agar, liquid media was prepared and Bacto Agar (15 g/l) was added prior to autoclaving, upon cooling agar was poured into sterile petri dishes. Inoculated dishes were inverted and incubated at 37°C overnight. Where necessary antibiotics were added to the culture medium.

2.3.3 Transformation of Competent *E.coli* Cells

Competent cells were removed from –80°C storage and thawed on ice. Plasmid DNA for transformation (10 ng plasmid DNA in a maximum of 3 μl) was added to the cells and gently mixed. Cells were incubated on ice for 30 minutes and then heat shocked for 30-60 seconds at 42°C followed by a return to ice for 2 minutes. Following transformation 0.25 ml of LBM or SOC broth was added to each transformation and tubes were placed at 37°C for 1 hour, with shaking (220 rpm). Transformants were selected by plating cells (100-200 μl) on LBM-agar containing an appropriate antibiotic. For plasmid DNAs allowing blue/white screening LBM-agar was supplemented with 40 μg/ml 5-bromo-4-chloro-3-indoyl-β-D-galatoside (X-Gal)

2.4 STANDARD PROTEIN WORK

2.4.1 SDS-PAGE

Protein samples were analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), using 12.5%, 15% or 17.5% resolving gel (12.5%, 15% or 17.5% (w/v) Protogel (37.5:1 acrylamide:bisacrylamide) (National Diagnostics, www.nationaldiagnostics.com), 0.375 M Tris/HCl (pH 8.8), 0.1% (w/v) SDS, 0.075% (w/v) ammonium persulphate, 0.05% (v/v) TEMED (*N, N, N', N'*-tetramethylethylenediamine) and 2.5% stacking gel (2.5% Protogel, 0.125M Tris/HCl (pH 6.8), 0.1% (w/v) SDS, 0.1% (w/v) ammonium pershulphate and 0.075% (v/v) TEMED according to Låemmli (Låemmli, 1970)). Minigels (8 x 10cm) were prepared and run in 1 x reservoir buffer (0.025M Tris/HCl pH 8.3, 0.192 M glycine, 0.1% (w/v) SDS) at 70 V through the stacking gel and 120 V through the resolving gel using an ATTO AE-6450 gel tank apparatus (Genetic Research Instrumentation Ltd. http://www.gri.co.uk/). Protein samples were prepared by adding SDS sample buffer to a 1 x concentration (from a 2 x or 5 x stock solution depending on the volume of protein solution in the sample) (0.1 M Tris/HCl (pH 6.8), 10% (v/v) glycerol, 1% (w/v) SDS, 0.001% (w/v) bromophenol blue, 1% β-mercaptoethanol) and heated in boiling water for 5 minutes before loading onto gel. A molecular weight marker was used to calibrate gels. Proteins were visualised on gel using either Coomassie blue or Kenacid blue stain solution.

2.4.2 Coomassie Blue / Kenacid Blue Staining of PAGE Gels

After polyacrylamide electrophoresis proteins in the μg range were visualised by staining with 0.25% (w/v) Coomassie Brilliant Blue G250 (or Kenacid blue), in 10% (v/v) glacial acetic acid, 40% (v/v) methanol, followed by destaining with 10% (v/v) glacial acetic acid, 40% (v/v) methanol. Gels were stained for a minimum of 3 hours, both staining and destaining were carried out at room temperature with gentle agitation.

2.4.3 Western blotting

Proteins were transferred from polyacrylamide gel to nitrocellulose membrane (Protran BA85, Whatman Ltd. www.whatman.com) by semi-dry blotting. Six sheets of 3MM paper (Whatman Ltd. www.whatman.com) and one sheet of nitrocellulose membrane were cut to the same size as the gel and soaked for 10 minutes in Towbins buffer (25mM Tris, 20% (v/v) MeOH, 192mM Glycine). The blotting sandwich was then prepared in an ATTO blotting apparatus (Genetic Research Instrumentation Ltd. http://www.gri.co.uk/) as follows: ANODE; 3x sheets of 3MM paper; nitrocellulose membrane; gel; 3x sheets of 3MM paper; CATHODE. Electroblotting was carried out at 150-200 mA for 1 hour. For visualization of protein standard marker transfer prior to immunodetection, the nitrocellulose membrane was soaked in Ponceau S stain (0.1% Ponceau S, 5% Acetic Acid) for 5 minutes and rinsed thoroughly in water.

For immunodetection, the nitrocellulose membrane was blocked for either 1 hour at room temperature or at 4°C overnight in blocking buffer (1.5 mM KH_2PO_4, 8mM Na_2HPO_4, 0.137 M NaCl, 1% Tween-20, 5% (w/v) non-fat milk powder). Following 3 x 5 minute washes in antisera buffer (1.5 mM KH_2PO_4, 8mM Na_2HPO_4, 0.137 M NaCl, 0.1% Tween-20, 5% (w/v) Non-fat milk powder), primary antibody was diluted 1:3000 (anti-GST) in 10 ml anti-sera buffer and the membrane incubated overnight at 4°C. The membrane was then washed in anti-sera buffer for 3 x 5 minutes at room temperature. IgG horseradish peroxidase conjugate (Biorad, www.biorad.com) was used as a secondary antibody in 10 ml blocking buffer at a 1:5000 dilution and incubated with the membrane at room temperature for 3 hours. This was followed by 1 x 5 wash in anti-sera buffer then 1 x 15 minute and 2 x 5 minute washes in PBS-T (1.5 mM KH_2PO_4, 8mM Na_2HPO_4, 0.137 M NaCl, 0.1% Tween-20) and several rinses in distilled water.

Enhanced chemoluminescence (ECL) reagents (GE healthcare, www.gehealthcare.com) were used for detection according to the manufacturer's instructions and proteins were visualized by exposure to X-ray film (Fuji SuperRX, Fuji Photo. Film Ltd, www.fujifilm.co.uk). Autoradiographs were either developed

using an automatic developer (Xograph Imaging Systems Compact X4, www.xograph.com) or manually.

2.4.4 Estimation of Protein Concentration

Protein concentrations were determined by using the Bio-Rad protein assay. Assays were set up in micro-titre plates and bovine serum albumin (BSA) was used as a protein standard, with dilutions prepared in the same buffer as the sample. BSA dilutions were used to plot a standard curve. Protein solutions of an unknown concentration were predicted using the standard curve. In microtitre plates 10μl of each standard or unknown sample was added to 150μl of water in separate wells and then mixed with 40μl of Bradford reagent. Absorbance was read at 570 nm using a Dynatech MT 5000 microtitre plate reader

2.4.5 Expression of GST-Tagged Proteins in BL21 AI Cells

BL21 AI (Invitrogen) were transformed with the expression constructs for AtPTPKIS1, AtPTPKIS2, AtPTPKIL1 or StPTPKL1. Clonal transformants were grown in liquid LBM broth containing carbenicillin (50μg/ml) at 37°C with shaking (200 rpm) until an OD λ_{600nm} 0.6-0.8. Mid-log cells were induced for expression with the addition of L-arabinose at a concentration of 0.02% (w/v) and continued growth for 16 hours post induction under similar growth conditions, with the temperature reduced to 28°C. Cells were collected by centrifugation (6,000 x *g* for 30 min, 4°C) and resuspended in 150mM HEPES to 1/20th the initial culture volume. Cells were lysed through sonication followed by centrifugation (10,000x g for 20 minutes, 4°C) to remove cell debris. Soluble protein was filtered and purified using 1ml Glutothione Sepharose FF from Amersham Biosciences following the manufacturers guidelines.

2.4.6 Expression of GST-Tagged Proteins in BL21 (DE3) STAR Cells

BL21(DE3) STAR (Invitrogen) were transformed with the expression constructs encoding the predicted KIS/CBM domains from AtPTPKIS1 and AtPTPKIS2 or the

control plasmid pGEX5x1(Amersham Biosciences). Clonal transformants were grown in liquid LBM broth containing carbenicillin (50 μg/ml) at 37°C with shaking (200 rpm) until an OD λ_{600nm} 0.6-0.8. Mid-log cells were induced for expression with isopropyl β-D-thiogalactoside (0.1 mM) and continued growth for 3 hours post induction under similar growth conditions. Cells were collected by centrifugation (6,000 x *g* for 30 min, 4°C) and resuspended in PBS to 1/20th the initial culture volume. Cells were lysed through sonication followed by centrifugation (10,000x g for 20 minutes, 4°C) to remove cell debris. Soluble protein was filtered and purified using 1ml GSTrap FF from Amersham Biosciences following the manufacturers guidelines.

2.5 METABOLITE ANALYSIS

2.5.1 Glucose, Fructose and Sucrose Analysis

For measurement of glucose, fructose and sucrose, leaf tissue was harvested and frozen in liquid N_2. Frozen leaf material was then powdered using a shaking steel ball mill at 30Hz for 30 seconds, and extracted in 200μl of 0.7 M perchloric acid for 30 minutes on ice. After centrifugation (5900g for 5 minutes at 4°C), 75μl of the supernatant was neutralised with 35μl 2M KOH. Following centrifugation (5900g for 5 minutes at 4°C), 50μl of the supernatant was analysed for glucose, fructose and sucrose content. Glucose, fructose and sucrose levels were identified through enzymatic assay, using the conversion of NAD to NADH as a method for quantifying glucose, fructose, and sucrose content sequentially as described in Muller-Rober *et al.* (1992).

2.5.2 Glucose-6-Phosphate Analysis

Leaf tissue was harvested throughout the day/night period from plants grown in short day conditions (8h light, 16h dark), flash frozen in N_2 and stored at -80°C prior to analysis. 300mg frozen material and 10μl 100mM glucose 1,6 biphosphate (for use as an internal standard) was incubated with 5ml 80% ethanol for 15 minutes prior to

centrifugation at 3000g for 10 minutes. Supernatant is collected and stored, before plant material was resuspended in 5ml 25% ethanol at 0°C for 30 minutes. Following centrifugation at 3000g for 10 minutes, the supernatant was removed and stored, and plant material resuspended in 2ml water at 0°C for 15minutes. Following centrifugation at 3000g for 10 minutes, supernatant was removed and stored, while the plant material was resuspended in 2ml water at 0°C for 5 minutes. Following centrifugation at 3000g for 10 minutes, the supernatant was removed and combined with all the other supernatants at room temperature. To this mixture, 2ml of dichloromethane was added and mixed gently before the mixture was centrifuged for 10 minutes at 3000g. The upper water phase was then removed, and re-extracted with 2ml dichloromethane. The final water phase was evaporated using a speedvac at room temperature overnight and the remaining material resuspended in 200μl water.

Phosphorylated sugars were then subjected to HPAEC-PAD on a PA-1 column (Dionex, Rødovre, Denmark) according to Blennow *et al.* (1998) using glucose-1, 6-bisphosphate as an internal standard.

2.5.3 Starch Content Analysis

Samples of leaf tissue were taken throughout the day/night period from plants grown in short day conditions (8h light, 16h dark) and their starch content quantified. Soluble carbohydrates were extracted using water and ethanol, before the remaining starch was isolated and determined through enzymatic assay as described in Muller-Rober *et al.* (1992).

2.5.4 Starch Bound Phosphate Analysis

Leaf tissue (400mg) was extracted 3 times for 5 minutes with 90% ethanol at 80°C, and then washed 3 times for 5 minutes with 50mM Na Acetate, pH 4.0 at room temperature. The leaf material was then washed twice in water, before 100μl Termamyl (Novozymes) was added to degrade the starch at 100°C for 30 minutes with agitation. Following centrifugation for 5 minutes at 20,000g, the supernatant was removed and stored, while the plant material was extracted twice more with

water. All supernatant fractions were combined and the water evaporated off in a speedvac, before the remaining material was dissolved in 100μl water. This material was then subject to acid hydrolysis with HCl at a final concentration of 0.7M for 4 hours at 100°C. Following acid hydrolysis the solution was neutralised with the addition of NaOH and the neutralised hydrolysate subjected to HPAEC-PAD on a PA-1 column (Dionex, Rødovre, Denmark) according to Blennow *et al.* (1998).

2.5.5 Iodine Staining of Leaf Starch

Mature leaves were harvested at the end of the light and dark periods from plants grown in short day conditions (8 hours light, 16 hours dark). Leaves were discoloured in 3 incubations with 90% ethanol at 80°C for 10 minutes. They where then incubated in 20% ethanol at room temperature for 5 minutes, before being rehydrated in 3 incubations with water at room temperature for 5 minutes. Following removal of the supernatants, the leaf material was incubated with Lugol solution (5g iodine, 10g potassium iodide in 100ml distilled water) for 10 minutes. The leaf material was then washed 3 times in water to remove excess iodine, before imaging on a scanner.

2.6 PHOSPHATASE ASSAYS

In all assays results were shown as the activity with the mass of enzyme used discounting and additional sequences (such as GST tag) introduced through recombinant expression.

2.6.1 *P*-nitrophenyl Phosphate Assays

p-nitrophenyl phosphate assays were carried out on clear flat bottomed micro-titre plates at room temperature. Each reaction was to a final volume of 100µl and contained 10mM *p*-nitrophenyl phosphate and 0.1mM DTT unless stated otherwise. Reactions were terminated with the addition of NaOH to a final concentration of 0.2M. The absorbance at 405nm was measured using a Dynatech MR5000.

Standards were generated using p-nitrophenyl to calculate the total product generated. Phosphatase inhibitors phenylarsine oxide and vanadate were used as described previously (Fordham-Skelton *et al.*, 1999).

2.6.2 Malachite Green Assays

Malachite Green solution was made according to Baykov *et al.*, (1988). Concentrated sulphuric acid (60ml) was slowly added to water (300ml) and allowed to cool to room temperature, before malachite green (0.44g) was added. Prior to use 7.5% ammonium molybdate (2.5ml) and 11% Tween 20 (0.2ml) were added to the malachite green solution (10ml). Following the phosphatase reaction, four volume of this coloured reagent was mixed with one volumes of the reaction solution. Reactions were kept at room temperature for 5 minutes and the absorbance measured. The absorbance at 620nm was measured using a Dynatech MR5000.

Phosphorylated peptides (Biomol) were resuspended in distilled water and used at a final concentration of 100μM with free phosphate release measured using Malachite Green. Carbohydrate substrates were used at ~100μg in 80μl and reactions lightly agitated throughout the assay. Buffers used were as stated in the results section. Following reaction, the assays were centrifuged (10,000g for 4 minutes at 4°C) and the supernatant removed for assay with malachite green. The reaction was stopped through heating or the addition of a terminator such as N-ethylmaleimide (to a final concentration of 50mM).

2.6.3 Radiolabeled Phosphoglucan Assay

Phosphoglucans were created using standard techniques (Mikkelsen *et al.*, 2004), with the scaffold of potato amylopectin being replaced by liver glycogen. Elongated glycogen was phosphorylated using GWD1, incorporating a radioactive label, as described by Mikkelsen *et al.* (2004). Following washing, phosphoglucans were incubated for 30 minutes at room temperature with the assayed protein. The reaction was terminated, and polyglucans precipitated with the addition of 18x reaction volume of 75% Methanol. The supernatant was removed and dried to remove

methanol before re-suspension in 400μl of water. The released radioactivity was measured using a MicroBeta Trilux liquid scintillation counter (Wallac).

2.7 CARBOHYDRATE BINDING ASSAYS

2.7.1 Glycogen Binding Pull-Down Assay

Protein-free glycogen (5 mg/ml) in 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% (v/v) 2-mercaptoethanol, 0.02% (w/v) Brij-35, 0.1 mg/ml BSA (Armstrong *et al.*, 1998) was mixed with GST fusion protein. After incubation on ice for 30 min, the samples were centrifuged for 90 min at 100000 g. Pellets were washed in 200ul 50 mM Tris-HCl, pH 7.5,150 mM NaCl, before resuspension in 4x SDS sample buffer (8% β-mercaptoethanol), and run on SDS-PAGE gel.

2.7.2 Starch Binding Pull-Down Assay

Potato Starch (sigma) (5 mg/ml) in 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% (v/v) 2-mercaptoethanol/0.02% (w/v) Brij-35/0.1 mg/ml BSA (Armstrong *et al.*, 1998) was mixed with GST fusion protein. After incubation on ice for 30 min, the samples were centrifuged for 10 min at 10,000g. Pellets were washed in 200ul 50 mM Tris-HCl, pH 7.5,150 mM NaCl, before resuspension in 4x SDS sample buffer (8% β-mercaptoethanol), and run on SDS-PAGE gel.

Quantification of protein content in protein bands on SDS gel was carried out using Quantity One (BioRad) to calculate the protein content of a band relative to that of known bands on the same gel.

2.7.3 Native PAGE

Protein samples were analysed by polyacrylamide gel electrophoresis (PAGE), using 10% or 7.5% resolving gel (10% or 7.5% (w/v) Protogel (37.5:1 acrylamide:bisacrylamide) (National Diagnostics, www.nationaldiagnostics.com), 0.375 M Tris/HCl (pH 8.8), 0.075% (w/v) ammonium persulphate, 0.05% (v/v) TEMED (*N, N, N', N'*-tetramethylethylenediamine) and 2.5% stacking gel (2.5% Protogel, 0.125 M Tris/HCl (pH 6.8), 0.1% (w/v) ammonium persulphate and 0.075% (v/v) TEMED, modified from Låemmli (Låemmli, 1970). Minigels (8 x 10cm) were prepared and run in 1 x native reservoir buffer (0.025 M Tris/HCl pH 8.3, 0.192 M glycine) at 70 V through the stacking gel and 120 V through the resolving gel using an ATTO AE-6450 gel tank apparatus (Genetic Research Instrumentation Ltd. http://www.gri.co.uk/). Protein samples were prepared by adding native sample buffer to a 1x concentration (from a 2 x or 5 x stock solution depending on the volume of protein solution in the sample) (0.1 M Tris/HCl (pH 6.8), 10% (w/v) sucrose, 0.001% (w/v) bromophenol blue).

Where starch or glycogen was used in native gels, it was included in the resolving gel at the concentration given in the results section. Soluble starch and bovine glycogen were purchased from Sigma-Aldritch.

2.8 BIOINFORMATICS

2.8.1 Protein Sequence Analysis

Where possible sequence information corresponding to these proteins was retrieved from the NCBI database (http://www.ncbi.nlm.nih.gov/). Using DNASTAR software (version 3.16) run on a Macintosh computer, protein sequences were aligned within the MegAlign programme using ClustalV algorithm. Prediction of secondary structure was carried out using Protean, part of the DNASTAR software suite.

2.8.2 Nucleotide Sequence Analysis

Identification of homologous or similar sequences was carried out using similarity searches against the NCBI database (http://www.ncbi.nlm.nih.gov/BLAST/), including its Nucleotide collection database, and its Expressed Sequence Tag (EST) database, using nucleotide and translated nucleotide searches. Contiguous sequences were produced in SequencherTM Version 4.1.2 (Gene Codes Corporation) on a Macintosh computer.

Chapter 3

PHOSPHATASE ACTIVITY OF PTPKIS FAMILY PROTEINS

3.1 INTRODUCTION

PTPKIS1 was originally identified by Fordham-Skelton *et al.*, (2002) and was described as "a novel higher plant protein tyrosine phosphatase", due to the presence of a C-terminal domain classed as a kinase interaction sequence (KIS) domain in combination with the N-terminal PTP domain. The study showed evidence of *in vitro* interaction between this sequence and the plant SnRK AKIN11, a protein related to AMPKs and SNF1 in the budding yeast *Saccharomyces cerevisiae* (Alderson *et al..*, 1991; Mitchelhill *et al..*, 1994; reviewed in Halford *et al..*, 2000). This led the authors to surmise that PTPKIS1 may act to regulate AKIN11 activity in the plant through dephosphorylation of the AKIN11 kinase. This may itself, be responsible for regulating a number of metabolic pathways, as has been identified with its homologues in budding yeast and animals (Hardie *et al.*, 1998; Halford *et al..*, 2000). Unfortunately the authors were unable to show significant biological activity upon phosphorylated tyrosine peptides, and were unable to express recombinant full-length protein from the *Arabidopsis thaliana* sequence.

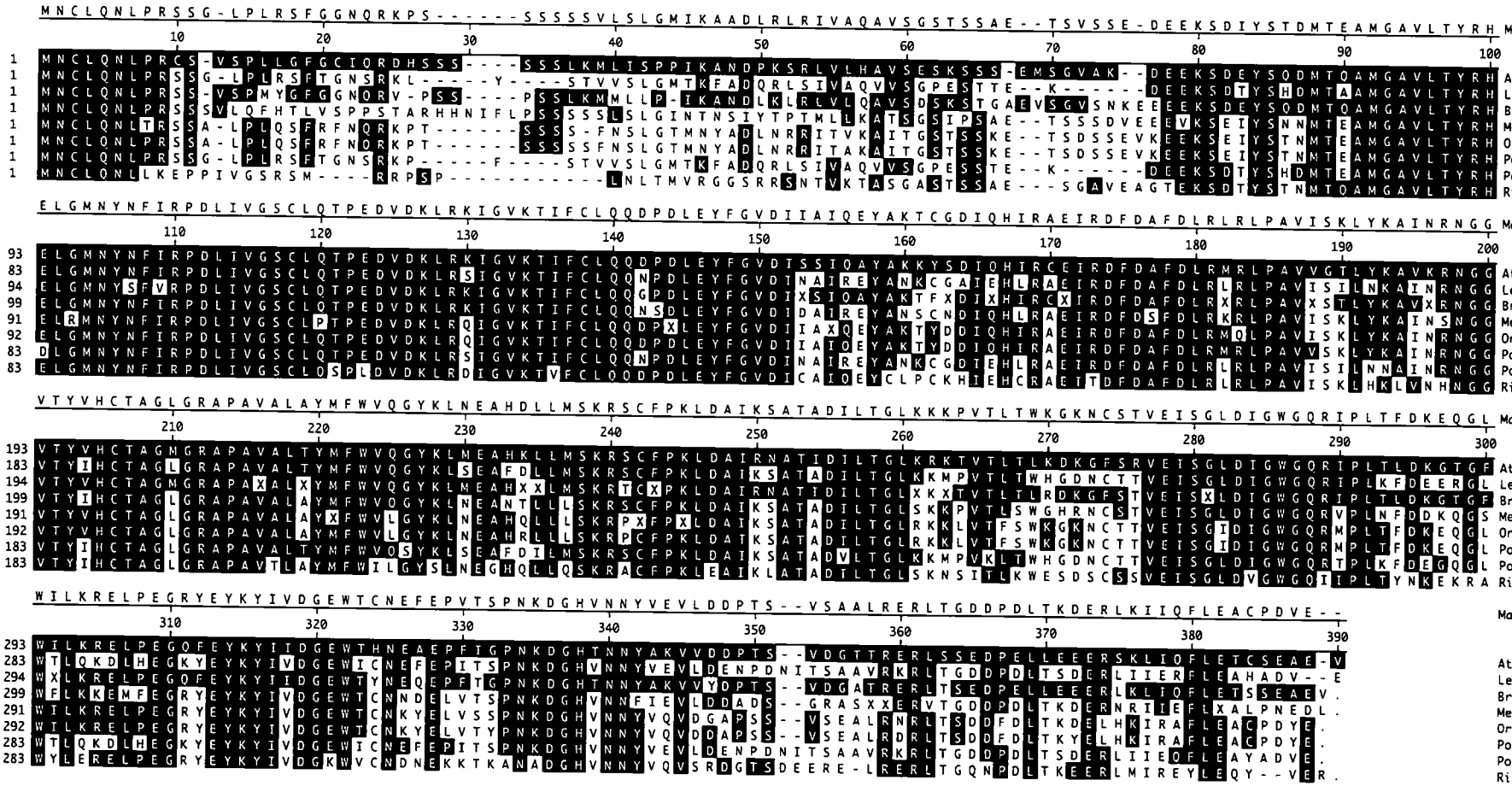
More recently it has been shown that the previously identified **S**tarch **EX**cess **4** phenotype (SEX4) was not in fact caused through mutation to a chloroplastic starch-hydrolysing enzyme (Zeeman et al., 1998), but due to mutation of the At3g52180 gene encoding AtPTPKIS1 (Niittyla et al., 2006). This publication as well as two others (Kerk et al., 2006; Sokolov et al., 2006) also showed the ability of the previously identified KIS domain to act as a carbohydrate binding module as well as association of PTPKIS1 to starch granules. It was also shown that PTPKIS1 localised to the chloroplast, making it unlikely it would be able to interact with AKIN11 (Kerk et al., 2006; Niittyla et al., 2006; Sokolov et al., 2006).

In Fordham-Skelton *et al.*, (2002), comparison was drawn between the sequences of the N-terminal protein tyrosine phosphatase (PTP) domain of PTPKIS1, and the phosphatase domain of the animal protein laforin. At the time, this was the only similarity drawn, since the PTP domain in laforin was C-terminal to an N-terminal CBM that showed little similarity to the PTPKIS1 sequence. With the evidence that the KIS domain of PTPKIS1 can itself act as a carbohydrate binding module, the similarity in PTP domain sequence becomes even more relevant, since the two proteins both contain two domains with similar functions. While this piece of work discusses the PTPKIS1 proteins present in higher plants, recent publications have also shown the presence of phosphatases with carbohydrate binding modules, similar in structure to laforin, present throughout a range of different organisms including protists, evolved from a progenitor red alga (Gentry et al., 2007). The assumed functional similarity between laforin and PTPKIS1 is also reflected in a similar phenotype resulting from inactivation of the proteins. While a deleterious mutation in the AtPTPKIS1 leads to the SEX4 phenotype, in which large starch granules form with high levels of un-branched amylose are present, similar mutations in mammalian laforin result in the formation of lafora bodies (Chan et al., 2004). These lafora bodies are formed from insoluble un-branched glycogen, in a similar fashion to starch granules (Chan et al., 2004).

In addition to its similarity to laforin, the phosphatase domain of PTPKIS1 was also shown to have sequence similarity to two other proteins in *Arabidopsis thaliana.* These proteins, encoded by genes At3g01510 and At3g10940, were also identified in the publication by Fordham-Skelton et al, (2002) as having homologous PTP regions, with At3g01510 (from now on referred to as AtPTPKIS2), containing a C-terminal domain similar to that of the KIS domain in PTPKIS1. Using the sequence from AtPTPKIS1 it was possible to search the GeneBank dbEST database of EST's using the 'TBLASTN' program, which generates a large set of sequences with potential homology to AtPTPKIS1 (Appendix 1). Where possible, multiple sequences from one species were compiled to generate a predicted full length sequence, which showed homology to PTPKIS1 (Fig. 3.1).

This chapter will aim to characterise the phosphatase activity of PTPKIS1 from both *Arabidopsis thaliana* and *Solanum tuberosum,* as well as the activities of the protein products of the At3g10940 and PTPKIS2 genes, in order to better understand their roles within the plant, and the relevance of these phosphatases to the degradation of leaf starch, as shown through the SEX4 mutation.

A


Majority
AtPTPKIS1 Seq
LePTPKIS seq
Brassica PTPKIS1
Medicago PTPKIS1
Orange PTPKIS1
Poncirus PTPKIS1
Potato PTPKIS1
Rice PTPKIS1a

B

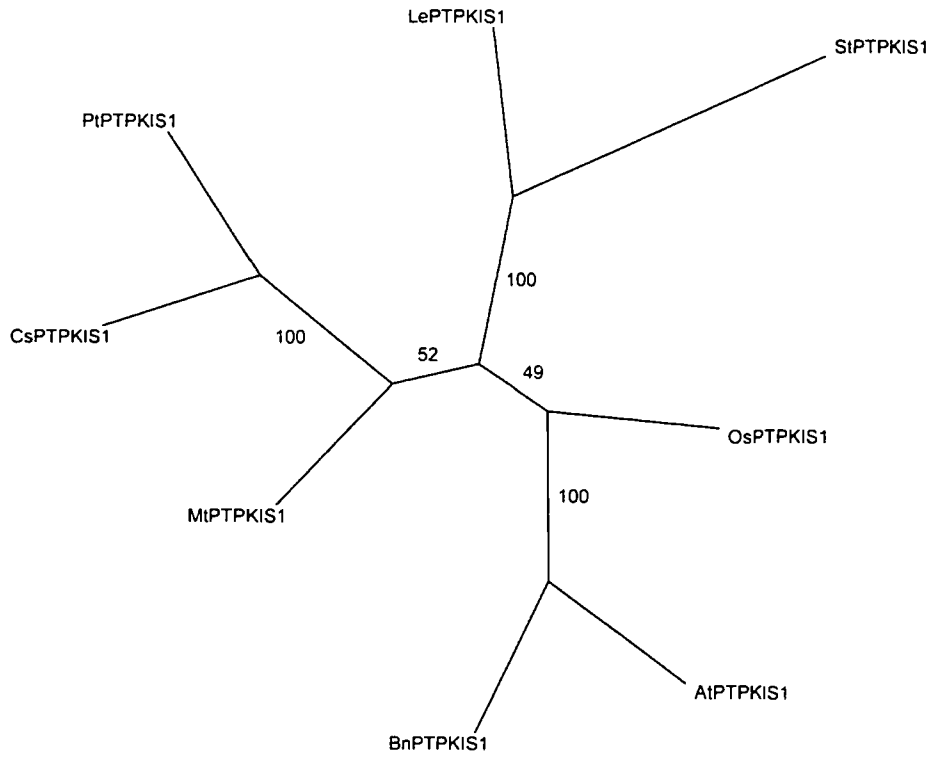


Figure 3.1 PTPKIS1 Homologues.

A. Alignment of predicted full length PTPKIS1 homologue sequences generated through similarity searches of the EST database. ESTs used in the generation of these sequences are found in Appendix I. Those residues common to AtPTPKIS1 are highlighted in black

B. Phylogenetic tree of predicted full length PTPKIS1 homologue sequences generated through similarity searches of the EST database. ESTs used in the generation of these sequences are found in Appendix I. Bootstrap values are show.

3.2 GENERAL PHOSPHATASE ACTIVITY OF PTPKIS1

3.2.1 Expression and Purification of AtPTPKIS1

cDNA corresponding to the previously identified AtPTPKIS1 (Fordham-Skelton et al., 2002) (accession number: NP_566960) was used to generate a BP clone (see materials and methods) using the Gateway System (Invitrogen), Plasmid pDONR201 and primers:

Forward: 5' – GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAATTGT CTTCAGAATCTTCCC – 3'

Reverse: 5' – GGGGACCACTTTGTACAAGAAAGCTGGGTTATGAACTTCT GCCTCAGAA CAAGTCTC – 3'

The resulting clone, termed RBP1, which contained the complete coding sequence for PTPKIS1 was then used in the gateway reaction with pDEST15 to generate a plasmid encoding AtPTPKIS1 with an additional N-terminal GST Tag for purification (Fig. 3.2). This plasmid was transformed into *E.coli* strain BL21AI, which was used for protein expression. Cultures were grown at 28°C and induced with 0.01% L-Arabinose for 16 hours before being pelleted and stored at -20°C. The low growth temperature and low level of induction were selected to maximise the amount of soluble recombinant protein. GST-PTPKIS1 was purified from cell supernatant by affinity batch absorption on glutathione-Sepharose, followed by washing and elution with Tris-HCl buffer at pH 8.0 (Fig. 3.2).

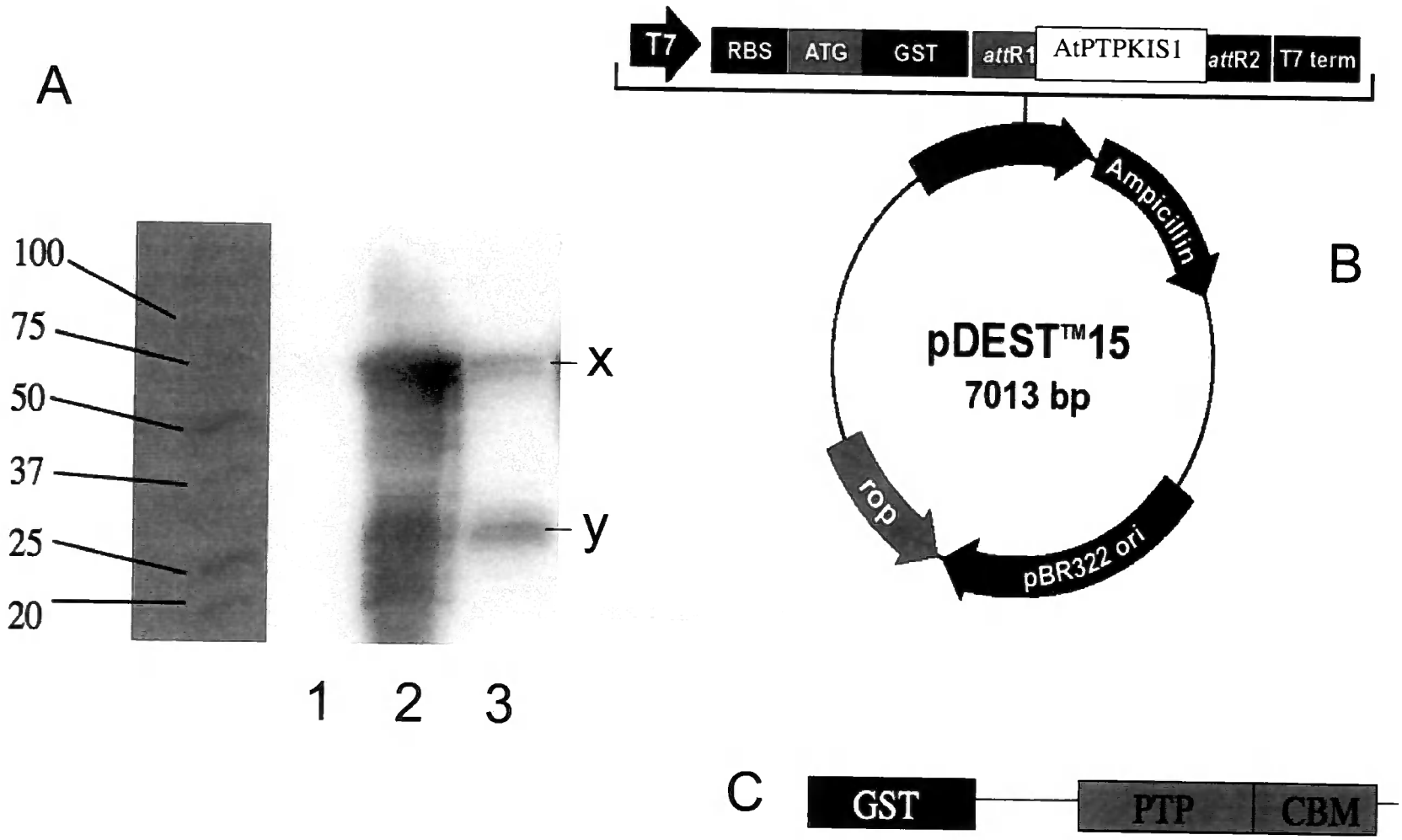


FIGURE 3.2 Expression of Recombinant AtPTPKIS1.

(A) Western blot using anti-GST antibody (invitrogen) showing expression of AtPTPKIS1. Lane 1,uninduced whole cell protein. Lane 2, induced whole cell protein. Lane 3, Purified protein. Band x, represents full length recombinant protein. Band y represents cleaved GST protein. Left hand lane is standard size markers, value shown are in kDa.

(B) Diagrammatic representation of the plasmid construct used for expression of recombinant AtPTPKIS1.

(C) Diagrammatic representation of recombinant protein, containing GST tag, and with highlighted predicted phosphatase domain (PTP) and carbohydrate binding / KIS domain (CBM)

3.2.2 Identification of PTPKIS1 in *Solanum tuberosum*

The sequence corresponding to the previously identified AtPTPKIS1 (Fordham-Skelton et al., 2002) (accession number: NP_566960) was used to search the GeneBank dbEST database of EST's using the 'TBLASTN' program. This produced 11 overlapping EST sequences for *Solanum tuberosum* (Appendix I), which were assembled into an open reading frame of 1113 base pairs, encoding an AtPTPKIS1 orthologue, termed StPTPKIS1 (Fig. 3.1). The predicted sequence of StPTPKIS1 was very similar to the previously identified tomato orthologue LePTPKIS1 (Fordham-Skelton et al., 2002) (accession number: CAC44460).

To isolate a sequence encoding StPTPKIS1, RT-PCR was carried out using primers:
Forward: 5' – ATGAATTGCCTTCAGAATCTTCCC – 3'
Reverse: 5' – CTTCAAGAAATCGTTCAATTATGA – 3'
DNA produced from the RNA extracted from young potato tuber gave a PCR product of the expected size, showing that StPTPKIS1 is expressed in this tissue, while DNA produced from the RNA extracted from mature, freshly wounded tubers, and mature wounded tubers after 16 hours showed no expression of StPTPKIS1 (Fig. 3.3). Young tubers were those tubers still developing on young potato plants prior to RNA extraction, while mature tubers were those harvested and stored before RNA isolation was carried out. EST databases suggested expression of StPTPKIS1 was affected by the developmental stage of the tuber and also by stress conditions.
The PCR product was cloned, and was confirmed to correspond to the expected product by DNA sequence determination. Following sequence confirmation the cDNA was cloned into pDONR201 using the Gateway BP reaction and primers:
Forward: 5' – GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAATTGC CTTCAGAATCTTCCC – 3'
Reverse: 5' – GGGGACCACTTTGTACAAGAAAGCTGGGTTCTTCAAGAAAT CGTTCAATTATGA – 3'
An expression clone for a recombinant GST-StPTPKIS1 protein was prepared in the same way as described for recombinant GST-AtPTPKIS1 clone. The resultant clone was used to produce recombinant protein, which was purified as described above for AtPTPKIS1 (Fig.3.3).

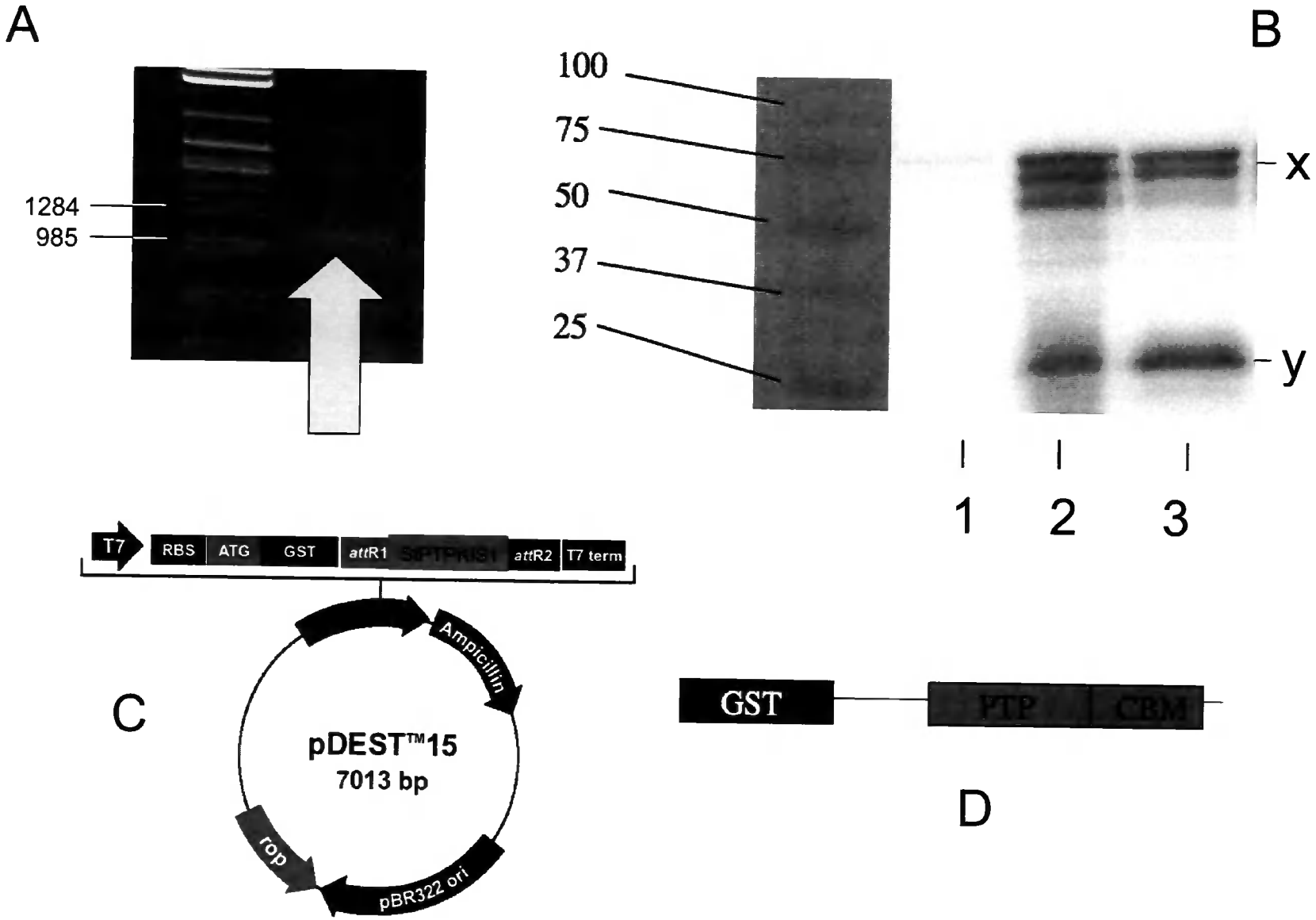


FIGURE 3.3 Expression of Recombinant StPTPKIS1.

(A) Agarose gel image of RT-PCR amplification of StPTPKIS1 coding sequence. Arrow indicates band corresponding to StPTPKIS1 coding sequence. Left hand lane contains Lambda DNA/Eco47I (AvaII) Marker, numbers shown are in nucleic acid base pairs (bp)

(B) Western blot using anti-GST antibody (Invitrogen) showing expression of StPTPKIS1. Lane 1, uninduced whole cell protein. Lane 2, induced whole cell protein. Lane 3, Purified protein. Band x, represents full length recombinant protein. Band y represents cleaved GST protein. Left had lane is standard size markers, value shown are in kDa.

(C) Diagrammatic representation of the plasmid construct used for expression of recombinant StPTPKIS1.

(D) Diagrammatic representation of recombinant protein, containing GST tag, and with highlighted predicted phosphatase domain (PTP) and carbohydrate binding / KIS domain (CBM)

3.2.3 Phosphatase Activity of PTPKIS1

Purified recombinant AtPTPKIS1 and StPTPKIS1 protein hydrolysed the generic phosphatase substrate *p*-nitrophenyl phosphate at a maximum rate similar to that previously estimated for the recombinant tomato enzyme LePTPKIS1 (Fordham-Skelton et al., 2002). Estimated values for Vmax were 8.2 ± 1.3nmoles/min/ug protein and 8.7 ± 1.1nmoles/min/ug protein (at pH7), respectively, for the Arabidopsis and potato recombinant proteins. The presence of a strong reducing agent such as DTT was necessary for maximal activity, with little activity in its absence, indicating the necessity of maintaining the active site cysteine residue in reduced form, and its sensitivity to oxidation. Against p-nitrophenyl phosphate AtPTPKIS and StPTPKIS1 have Km values of 0.445mM and 0.451mM respectively (Table 3.1) at pH7.

Enzyme	Substrate	Vmax	Km
AtPTPKIS1	*p*-nitrophenyl phosphate	8.2 ± 1.3 (nmoles/min/ug)	0.445mM
	Peptide Substrate1	0.41 ± 0.43 (pmoles/min/ug)	
	Peptide Substrate 2	0.37 ± 0.40 (pmoles/min/ug)	
StPTPKIS1	*p*-nitrophenyl phosphate	8.7 ± 1.1 (nmoles/min/ug)	0.451mM
	Peptide Substrate1	0.52 ± 0.50 (pmoles/min/ug)	
	Peptide Substrate 2	0.38 ± 0.39 (pmoles/min/ug)	

Table 3.1 Phosphatase Activity against *p*-nitrophenyl phosphate and Peptide Substrates. Table showing the activity of recombinant AtPTPKIS1 and StPTPKIS1 against the general phosphatase substrate *p*-nitrophenyl phosphate and 2 synthetic phosphopeptide substrates.

The similarity in the kinetic parameters for AtPTPKIS1 and StPTPKIS1 is also reflected in the pH optimum; both enzymes had a pH optimum around pH 7, as can be seen in figure 3.4. This is slightly above the pH optimum of 6.8 in Fordham-Skelton *et al.*, (2002), and below the pH 8.0 suggested as an optimum in the recently published paper (Gentry et al., 2007). This does however place the optimum activity in the biological range of pH 7-8 found in the plant chloroplast throughout the day and night cycle, with the enzymes having >50% activity in the range of pH6.25-pH8.

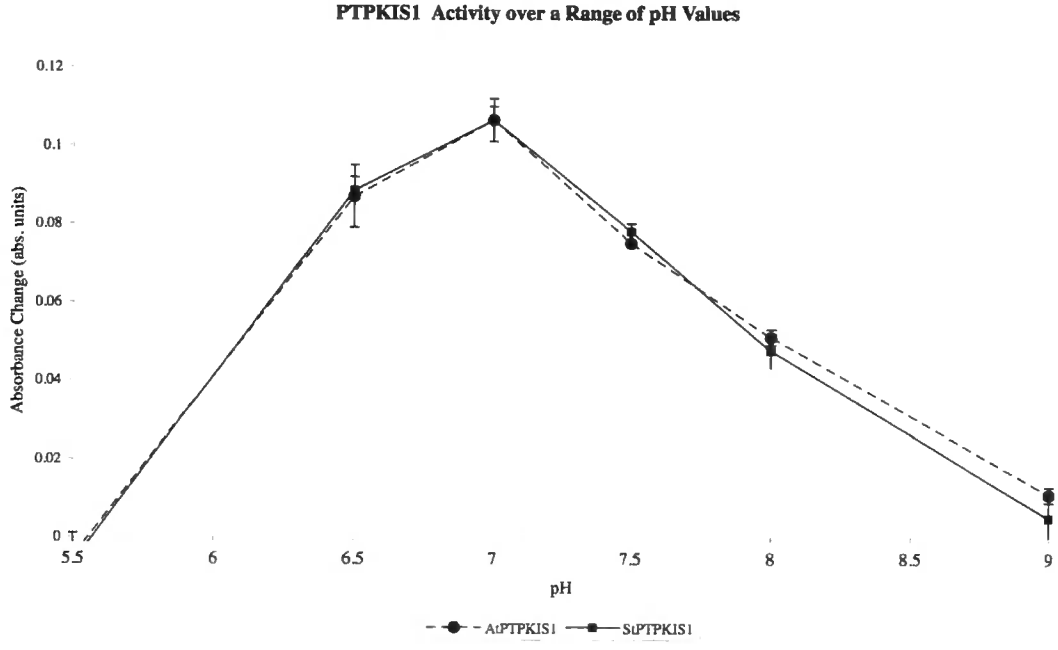


Figure 3.4 pH Activity Curve of AtPTPKIS1 and StPTPKIS1. Graph showing the activity of AtPTPKIS1 (●) and StPTPKIS1(■) over a range of pH, against *p*-nitrophenyl phosphate, using standard assay conditions.

In addition to the use of the standard phosphatase substrate *p*-nitrophenyl phosphate, the phosphatase activity of AtPTPKIS1 and StPTPKIS1 towards two phosphopeptide substrates was also studied (see Table 3.2). Substrate 1, was a protein phosphatase peptide substrate containing a phosphorylated threonine residue, commonly used as a substrate for several seine/threonine phosphatases. Substrate 2 was a protein tyrosine phosphatase peptide substrate, containing a phosphorylated tyrosine residue, which acts specifically as a peptide substrate for tyrosine phosphatases. Assays were carried out at room temperature, at pH 7.0, and free phosphate content was measured using the malachite green method.

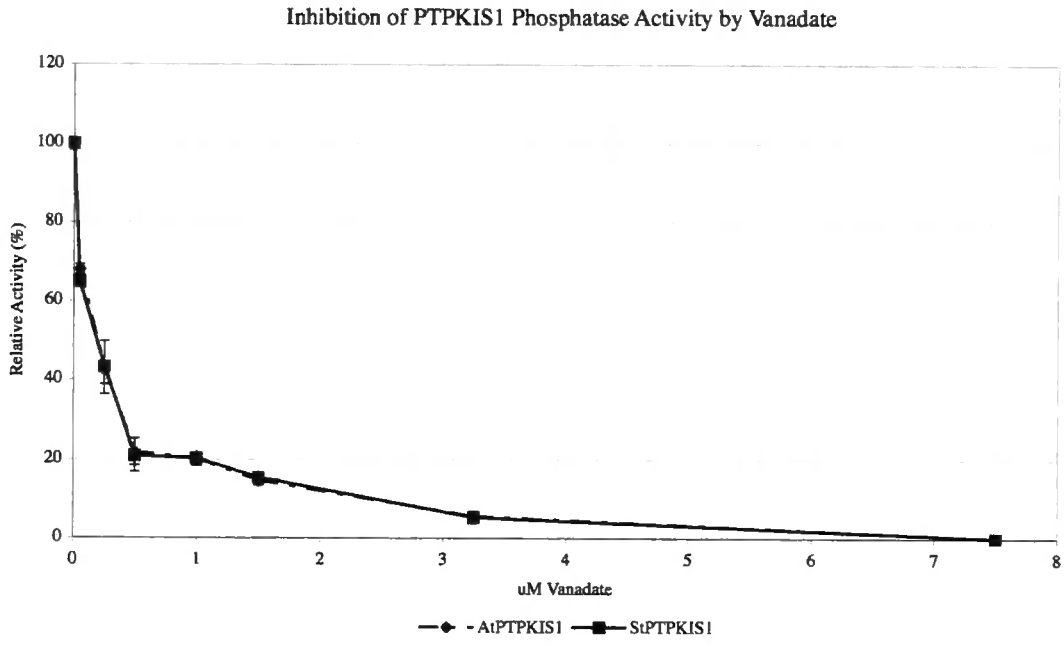
	Sequence
Substrate1 Protein Phosphatase Substrate	H-Arg-Arg-Ala-**pThr**-Val-Ala-OH (RRA**pT**VA)
Substrate2 Protein Tyrosine Phosphatase Substrate	H-Glu-Asn-Asp-**pTyr**-Ile-Asn-Ala-Ser-Leu-OH (END**pY**INASL)

Table 3.2 Peptide substrates used in phosphopeptide phosphatase assay.

In both cases the phosphatase activity shown by both AtPTPKIS1 and StPTPKIS1 was not significant, with any phosphate release being barely detectable above background. The activity observed was orders of magnitude lower than that of a true protein phosphatase used as a positive control, (recombinant soybean tyrosine-specific PTP; (Fordham-Skelton et al., 1999)), and was similar to that of the previously characterised LePTPKIS1 (Fordham-Skelton et al., 2002) (Table 3.1) showing little or no significant activity.

3.2.4 Inhibition Of Phosphatase Activity

The phosphatase activity of AtPTPKIS1 and StPTPKIS1 towards *p*-nitrophenyl phosphate was assayed in the presence of either vanadate or phenylarsinine oxide (PAO), both of which are used commonly to distinguish PTP's from other protein phosphatases. Both enzymes were inhibited in a concentration dependent manner by both agents (Fig 3.5). Vanadate was the more potent inhibitor of both enzymes, with an IC50 value of approx. 0.2μM, while the IC50 value for PAO was approx. 2.5μM. The inhibition observed was similar to that previously characterised for LePTPKIS1 (Fordham-Skelton *et al.*, 2002).



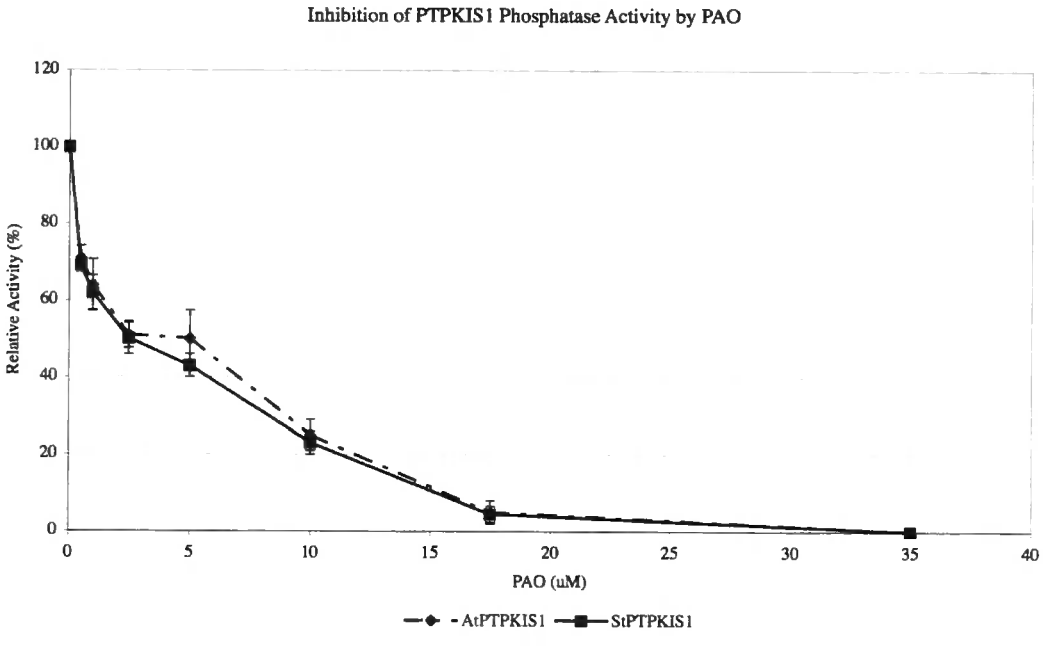


Figure 3.5 Inhibition of PTPKIS1 by Vanadate and PAO. Graphs showing the effect of increasing concentrations of Vanadate or PAO upon the activity of AtPTPKIS1 (●) and StPTPKIS1(■), with activity measured relative to the enzymes activity without either Vanadate or PAO present.

In addition to the standard inhibition assays using phosphatase inhibitors carried out for PTP's, additional assays were carried out in recognition of the presence of the carbohydrate binding domain in PTPKIS1 and its possible effect upon phosphatase activity. Using *p*-nitrophenyl phosphate to measure activity, the effects of reagents

which could interact with the carbohydrate-binding domain were assayed. Maltooligosaccharides are predicted to be capable of binding to the protein and could hinder access to the active site of the phosphatase domain, and thus could have an inhibitory effect. Assays showed that while no maltooligsacharide caused complete inhibition of the enzyme's phosphatase activity, unlike vanadate or PAO, the presence of the carbohydrates had a significant effect on phosphatase activity, dependent on the size of maltooligosacharide used (Fig. 3.6).

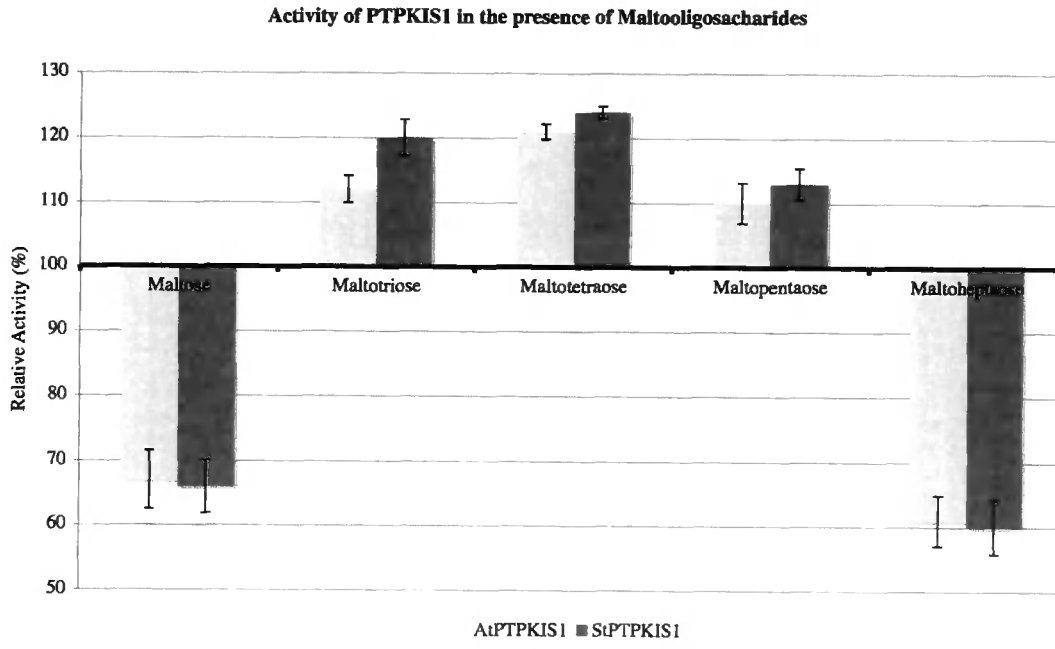


Figure 3.6 Activity of PTPKIS1 in the Presence of Maltooligosacharides. Effect on AtPTPKIS1 (light grey) and StPTPKIS1 (dark grey) activity in standard phosphatase assays using *p*-nitrophenyl phosphate as a substrate, of the inclusion of 50ug/ml maltooligosacharide agent at 50ug/ml. Activity is given relative to PTPKIS1 activity in the absence of any maltooligosacharide agent.

In the case of both AtPTPKIS1 and StPTPKIS1 activity increases in activity of up to 20% were observed in the presence of polysaccharides containing 3, 4 and 5 glucose units (maltotriose, maltotetraose and maltopentaose), but activity was reduced by up to 40% in the presence of polysaccharides containing 2 and 7 glucose units (maltose and maltoheptaose). This may have biological significance if the activity of PTPKIS1 is in part regulated by the ratios of different maltooligosacharides.

3.3 PHOSPHOGLUCAN PHOSPHATASE ACTIVITY OF PTPKIS1

Following the observation that maltooligosacharides of variable length have different effects on the activity of the PTPKIS1 phosphatase domain, it was important to study the possibility that PTPKIS1 was not a protein phosphatase, but a phosphoglucan phosphatase, capable of dephosphorylating the glucans already phosphorylated by glucose-water dikinase (GWD1). Inactivation of GWD1 gives a starch excess phenotype designated SEX1 which is distinct from, and non-allelic to, the SEX4 phenotype produced by PTPKIS1 inactivation.

3.3.1 Activity On Elongated Phosphorylated Glucan Substrates

In collaboration with the Starch group at KVL (Denmark), phosphoglucans were created using standard techniques (Mikkelsen et al., 2004), with the scaffold of potato amylopectin being replaced by liver glycogen. Elongated glycogen was phosphorylated using GWD1, incorporating a radioactive label, as described by Mikkelsen *et al.* (2004). Following washing, phosphoglucans were incubated for 30 minutes at room temperature with StPTPKIS1. To estimate the released phosphate, the polysaccharides were precipitated, and the radioactivity in the supernatant was estimated by scintillation counting after drying down and resuspension.

The results of this assay do not support a role for PTPKIS1 as a phosphoglucan phosphatase. While there is a slight increase in released radioactivity in the presence of PTPKIS1, at no point does this increase become significant (Fig. 3.7). While this result shows it is unlikely that there is any PTPKIS1 activity upon this substrate, it does not rule out activity towards naturally formed substrates, such as starch, which have a complex branching structure which may be important for binding and phosphatase activity. In addition, phosphorylation of glucans by GWD1 alone *in vitro* may not replicate the phosphorylation of starch *in vivo*.

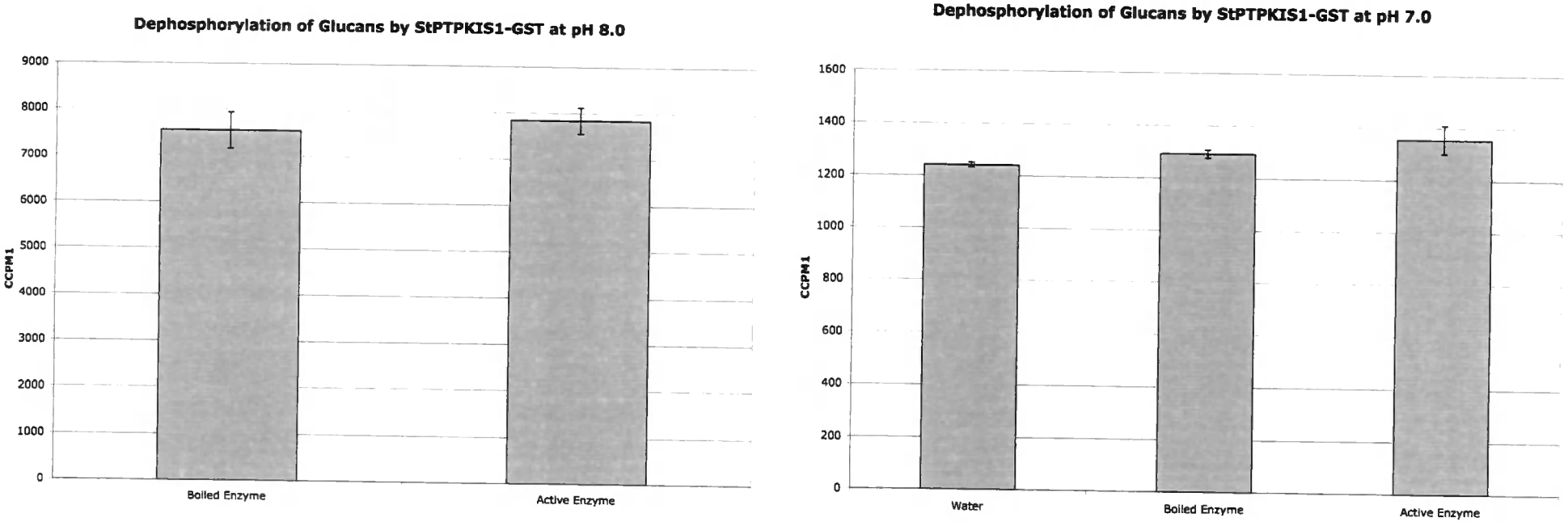


Fig 3.7 – Phosphate Release from Phosphorylated Glucans. Phosphoglucans created using standard techniques (Mikkelsen et al., 2004), using elongated glycogen phosphorylated using GWD1, were used as a substrate for StPTPKIS1. Assays were carried out in standard conditions. Activity is shown as CCPM1, counts per minute corrected for background and efficiency.

3.3.2 Activity on 'Natural' Glucan Substrates

In order to ascertain the potential for PTPKIS1 to act upon phosphorylated glucan substrates *in vivo*, its activity against a number of naturally generated glucan substrates was studied. Activity of AtPTPKIS1 and StPTPKIS1 was compared with that of a mutated AtPTPKIS1, in which the aspartic acid residue at position 166 of AtPTPKIS1 was mutated to form an alanine residue in order to produce a substrate trap (Sylviane Comparot-Moss, JIC. Personal communication). This mutant was termed D166A-FL and generated using the Stratagene "Quick Change" mutagenesis kit (see Methods section), using primers:

Forward: GC TGT GAA ATT AGA GCC TTT GAT GCA TTT GAT TTG
Reverse: CAA ATC AAA TGC ATC AAA GGC TCT AAT TTC ACA GC

The mutated form of AtPTPKIS1 was produced as a recombinant protein as described for AtPTPKIS1. It showed no phosphatase activity towards P-nitrophenyl phosphate.

Release of phosphate from samples of amylopectin (Sigma), soluble potato starch (Sigma) and starch prepared from leaves of SEX4 Arabidopsis plants (see methods section) was assayed using the malachite green method (see materials and methods).

It was shown that AtPTPKIS1 and StPTPKIS1 efficiently liberate phosphate from all of the naturally occurring polysaccharides, with similar activities for both enzymes against all three substrates (Fig. 3.8). The inactive D166A-FL mutant of AtPTPKIS1 gave levels of released phosphate that were not significantly above background, showing that enzyme activity was necessary for phosphate release.

This activity is ~40x less than towards *p*-nitrophenyl phosphate, which could be predicted due to *p*-nitrophenyl phosphates solubility, and the ease at which it is hydrolysed by other phosphatases.

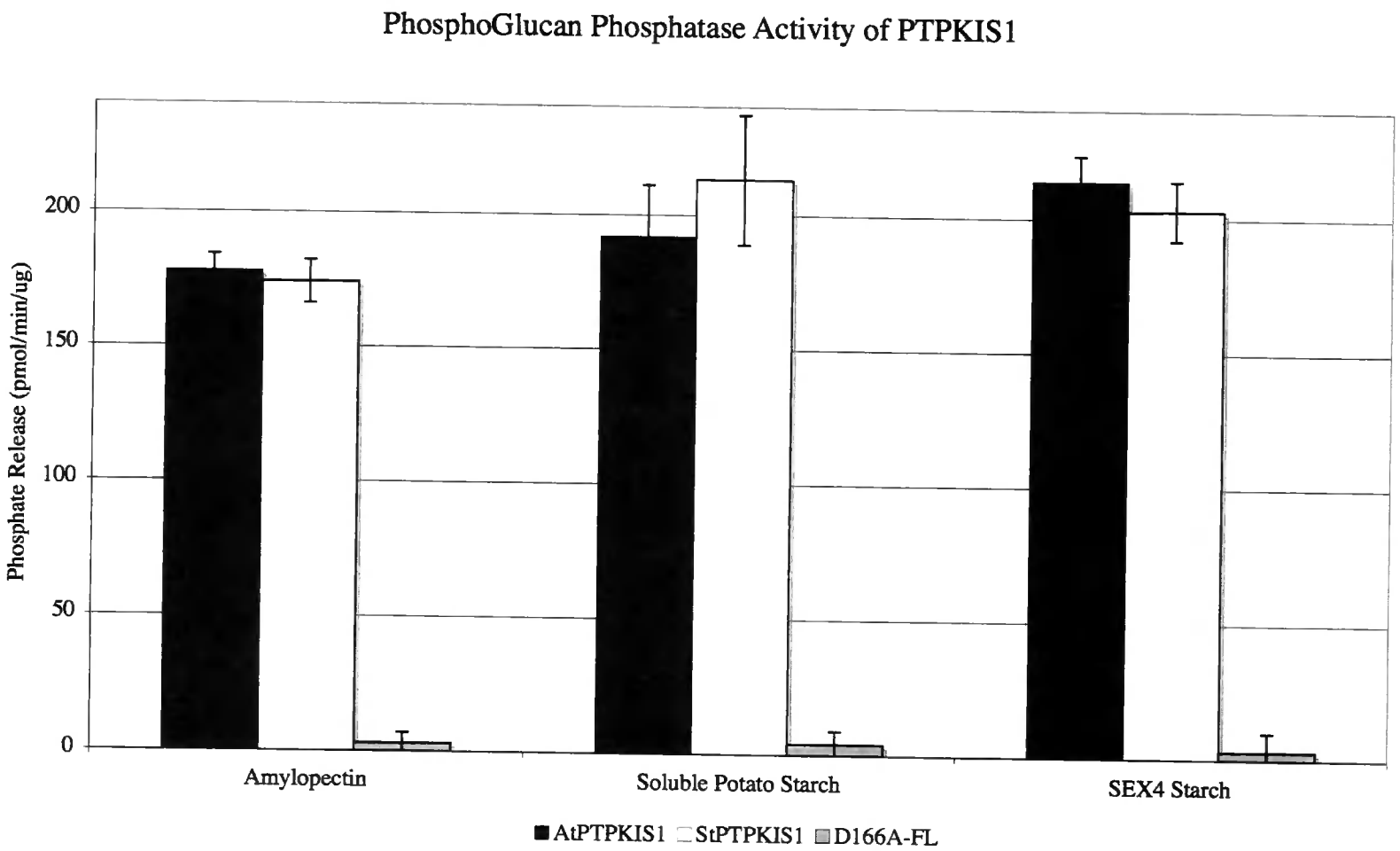
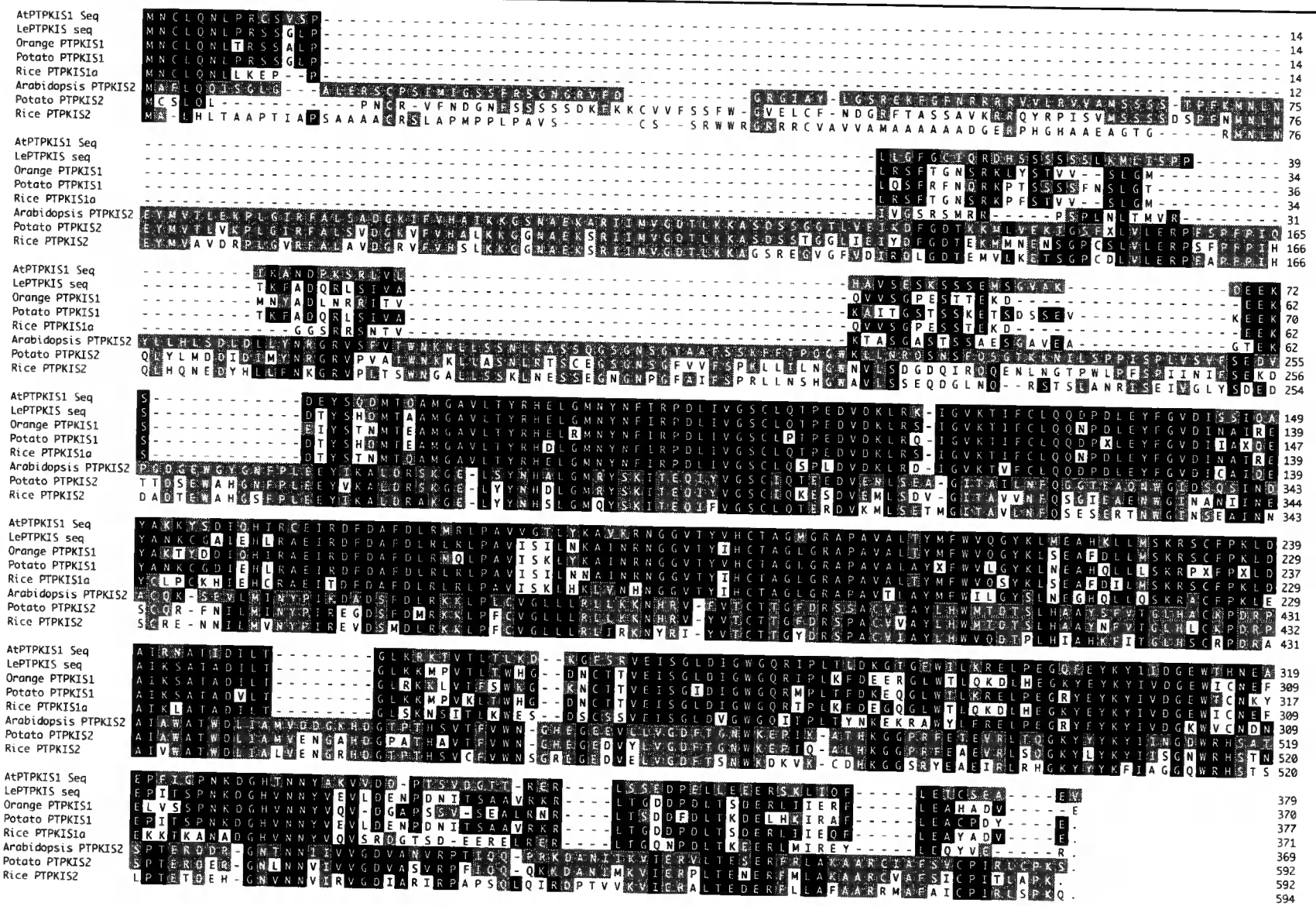


Figure 3.8 PhosphoGlucan Phosphatase Activity of PTPKIS1. Graph showing the activity of AtPTPKIS1 (black), StPTPKIS1 (white) and the AtPTPKIS1 mutant D166A-FL (grey), Against 3 substrates, potato amylopectin, soluble potato starch and starch isolated from SEX4-3 plants at the beginning of the dark period.

3.4 PHOSPHATASE ACTIVITY OF PTPKIS2

AtPTPKIS2 was identified in previous publications (Fordham-Skelton *et al.* 2002; Kerk *et al.* 2002) as having a high degree of similarity in its predicted PTP domain to that of AtPTPKIS1, while also having a similar predicted KIS/carbohydrate binding module at its C-terminal region (Fig3.9). As with AtPTPKIS1, the predicted sequence of AtPTPKIS2 (Accession number: NP_566139) was used to search the GeneBank dbEST database of EST's using the 'TBLASTN' program. From the EST sequences identified, it was possible to generate full-length sequences encoding a single protein in *Solanium tuberomun* (StPTPKIS2) and two in *Oryza sativa* (OsPTPKIS2a and OsPTPKIS2b). Analysis of the *Oryza sativa* sequenced genome showed the OsPTPKIS2a and OsPTPKIS2b were splice variants of the same gene (Os08g29160) (see Appendix 2), with OsPTPKIS2b containing additional sequences in its N-terminal region. When predicted protein sequences for AtPTPKIS2, StPTPKIS2 and OsPTPKIS2a are compared they show great similarity throughout the PTP and KIS/carbohydrate binding module regions as well as in a region at the N-terminus between residues 70 and 135 (Fig 3.10). This N-terminal region was identified by Fordham-Skelton *et al.* (2002) as having similarity to a PDZ domain, often used for protein- protein interaction. In addition to these full-length sequences, additional EST's were found for 10 other higher plants, including barley, spruce, soya, and tomato (see appendix 1).


AtPTPKIS1 Seq
LePTPKIS seq
Orange PTPKIS1
Potato PTPKIS1
Rice PTPKIS1a
Arabidopsis PTPKIS2
Potato PTPKIS2
Rice PTPKIS2

A

B

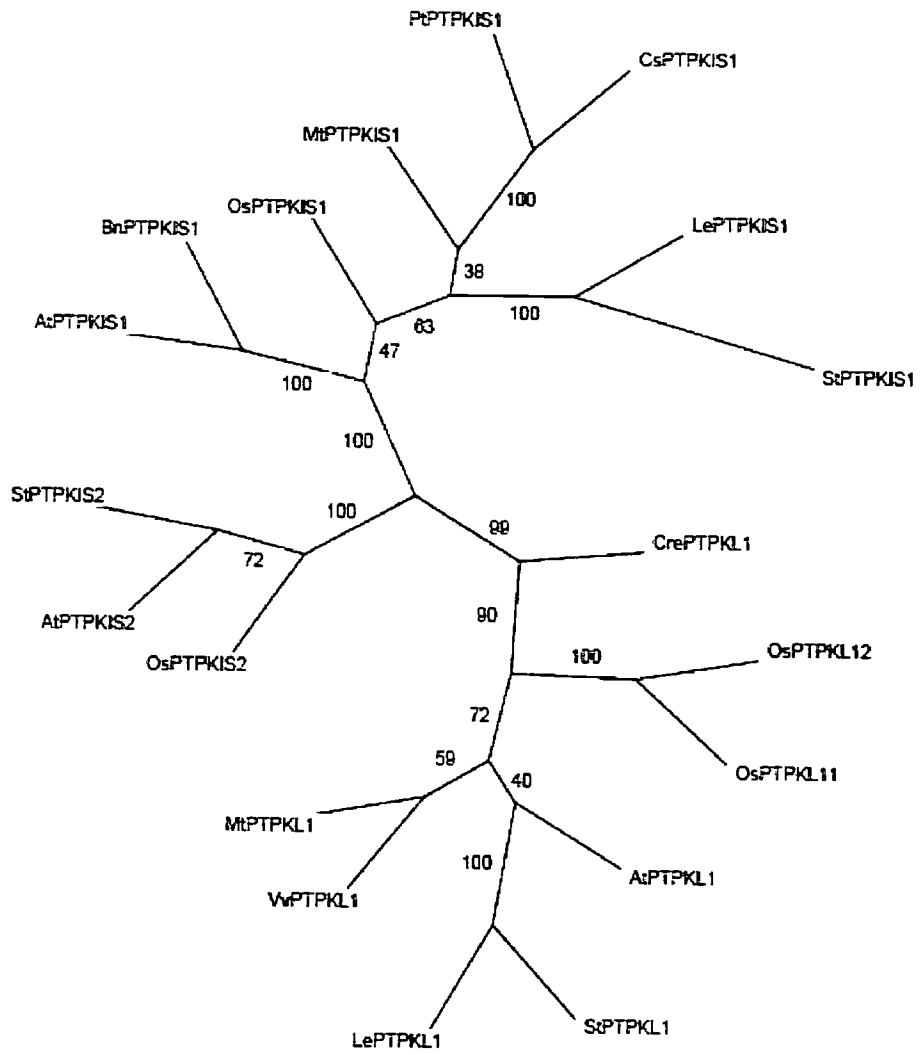


Figure 3.9 Sequence similarity of PTPKIS1 and PTPKIS2.

A.

Alignment of some of the predicted PTPKIS2 and PTPKIS1 proteins. Those residues identical to the consensus sequence are highlighted in black. Those that differ from the consensus, but are identical to either AtPTPKIS1 or AtPTPKIS2 are shown highlighted in grey.

B.

Phylogenetic tree of predicted full length PTPKIS1, PTPKIS2 and PTPKL homologue sequences generated through similarity searches of the EST database. ESTs used in the generation of these sequences are found in Appendix I. Numbers shown are bootstrap values.

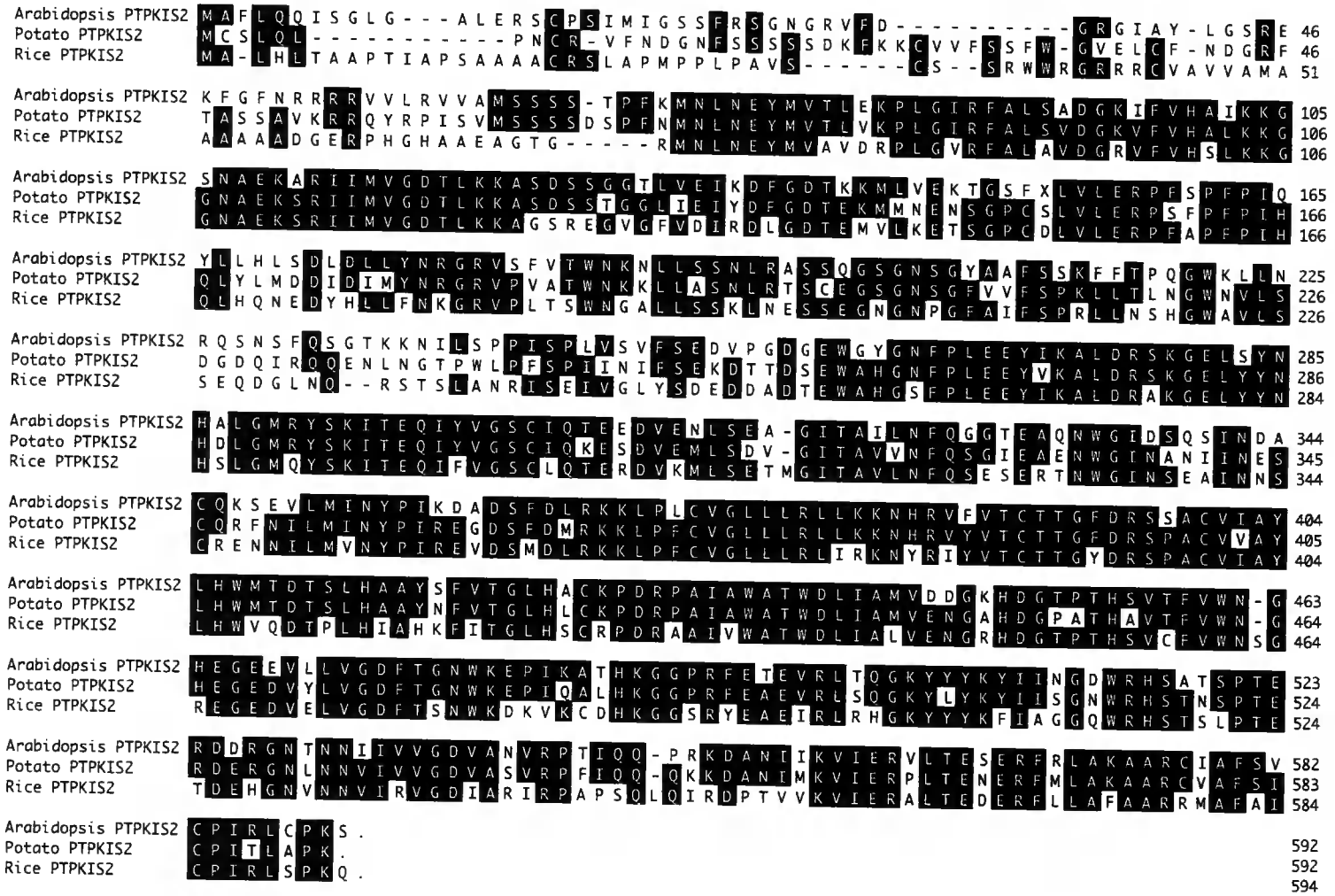


Figure 3.10 PTPKIS2 Homologues. Alignment of the predicted PTPKIS2 protein homologues. Those residues identical to the consensus sequence are highlighted in black.

3.4.1 Expression of Recombinant AtPTPKIS2

The coding sequence of AtPTPKIS2 was isolated from cDNA generated from RNA isolated from leaf tissue under light conditions using the primers:
Forward: ATGGCGTTTCTTCAACAAATCTCCGG
Reverse: CTACGACTTTGGGCATAGTCTTATTGG

These primers amplified the complete coding sequence including the start codon. The predicted amplification product was a fragment of 1776bp. When the *A.thaliana* cDNA was amplified with these primers, a single fragment was observed following agarose gel electrophoresis, of approximately 1700bp – 1800bp (Fig 3.11).

The fragment was purified by excision of band from gel, isolation of DNA, and cloning in a specialised vector for PCR products (pCR2.1), using the TOPO TA cloning system (Invitrogen). Clones resulting from this operation were characterised by DNA sequencing, which confirmed the clone as being identical to the predicted sequence, apart from a small number of single nucleotide changes, which did not alter the encoded amino acids.

Using the characterised clone encoding AtPTPKIS2 as a template, the coding sequence was amplified using primers:
Forward: 5'- GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCGTTTC TTCAACAAATCTCCGG – 3'
Reverse: 5' - GGGGACCACTTTGTACAAGAAAGCTGGGTTATGCGACTTT GGCGATAGTCTTAT – 3'
These contained the addition of an attB1 and attB2 site respectively in order to allow the insertion into the gateway donor vector pDONR201 (Invirogen). This clone, termed ABP1, was then used in the gateway reaction with pDEST15 (Invitrogen) to generate a plasmid encoding AtPTPKIS2 with an additional N-terminal GST Tag for purification (Fig. 3.11). This plasmid was transformed into *E.coli* strain BL21AI, which was used for protein expression, employing a similar method to that used for PTPKIS1 production. Cultures were grown at 28°C and induced with 0.01% L-Arabinose for 16 hours. GST-PTPKIS2 was purified from cell supernatant by

affinity batch absorption on glutathione-Sepharose, followed by washing and elution with Tris-HCl buffer at pH 8.0 (Fig. 3.11).

3.4.2 Phosphatase Activity of AtPTPKIS2

Purified recombinant AtPTPKIS2 was assayed for phosphatase activity using the same methodology as AtPTPKIS1 and StPTPKIS1. When assayed against the generic substrate *p*-nitrophenyl phosphate and the phosphorylated peptides used to assay PTPKIS1, there was no activity above background. When phosphatase activity was assayed using amylopectin, soluble starch and starch from SEX4 plants there was also no measurable activity above background.

This inability to show any activity towards any of the phosphatase substrates provided suggests that AtPTPKIS2 is catalytically inactive. This may be due to a significant difference in the amino acid sequence in its active site region compared to that of PTPKIS1. A conserved histidine residue in PTPKIS1 is replaced with threonine in PTPKIS2. This histidine residue forms part of the conserved diagnostic motif (HCX_5R) shared by PTPKIS1, tyrosine-specific PTPs and dsPTPs (Denu and Dixon, 1998)

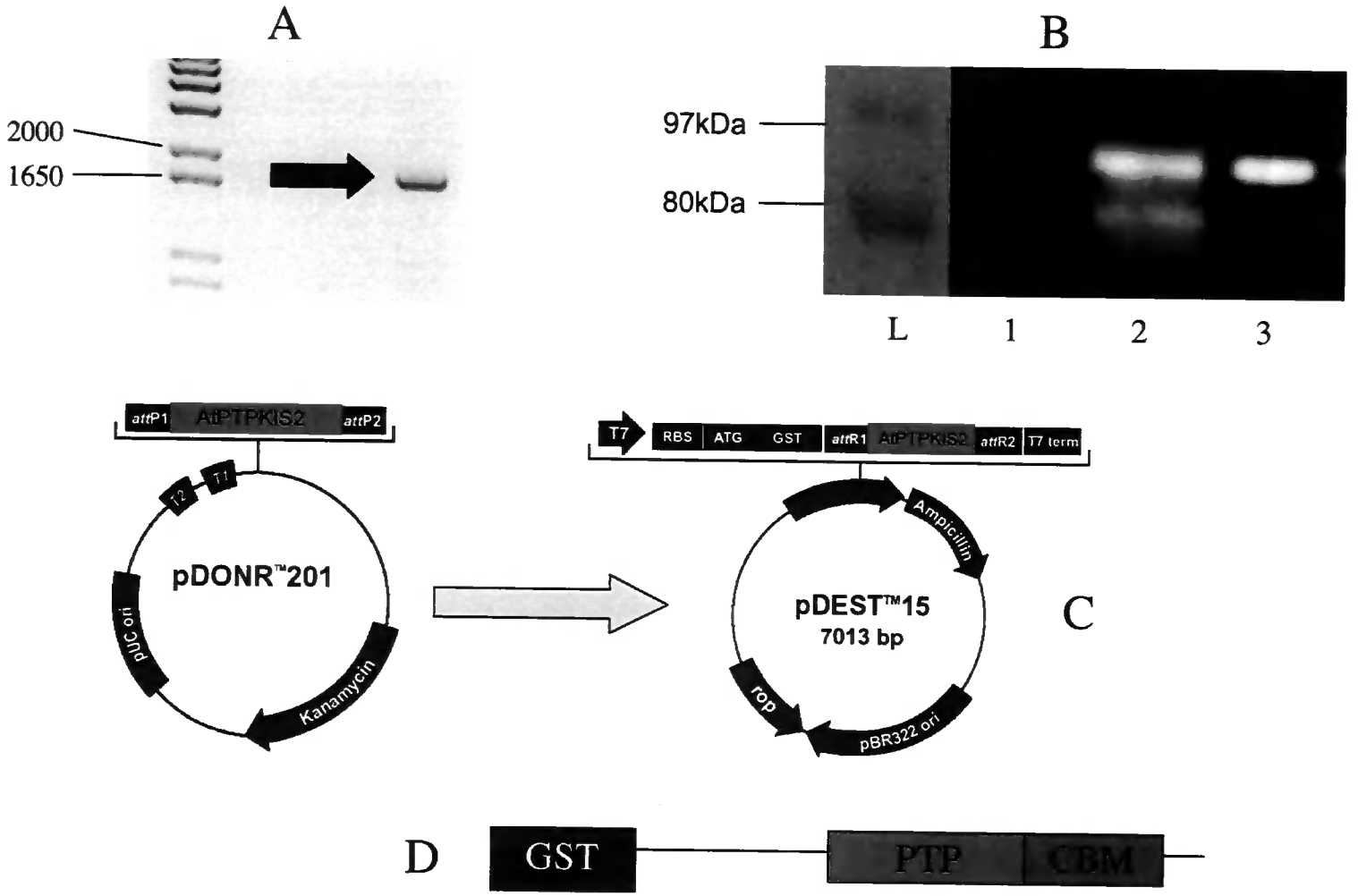


Figure 3.11 Expression and Purification of Recombinant AtPTPKIS2.

(A) Agarose gel image of RT-PCR amplification of AtPTPKIS2 coding sequence. Arrow indicates band corresponding to AtPTPKIS2 coding sequence. Left hand lane contains 1kb Plus DNA Ladder (Invitrogen), numbers shown are in nucleic acid base pairs (bp)

(B) Western blot using anti-GST antibody (Invitrogen) showing expression of AtPTPKIS1. Lane 1, uninduced whole cell protein. Lane 2, induced whole cell protein. Lane 3, Purified protein. Lane L contains Protein Molecular Weight Standards (Molecular Probes).

(C) Diagrammatic representation of the plasmid constructs used for cloning and expression of recombinant AtPTPKIS2.

(D) Diagrammatic representation of recombinant protein, containing GST tag, and with highlighted predicted phosphatase domain (PTP) and carbohydrate binding / KIS domain (CBM)

3.5 MODULATION OF THE PHOSPHATASE ACTIVITY OF PTPKIS1 BY PTPKIS2

Due to the inability of AtPTPKIS2 to dephosphorylate generic substrates, or phosphoglucans, it was important to consider other possible roles for the protein. Plants in which PTPKIS2 had been knocked out show a starch excess phenotype, similar, but not as extreme as that of SEX4 (see Chapter 5), and so it is important to try and identify the potential role of PTPKIS2.

3.5.1 Activity on Generic Substrates

When combined with AtPTPKIS2 at an equimolar concentration, Purified recombinant AtPTPKIS1 and StPTPKIS1 protein hydrolysed the generic phosphatase substrate *p*-nitrophenyl phosphate at a comparative rate to that seen without AtPTPKIS2. AtPTPKIS1 + AtPTPKIS2 released phosphate at a maximal rate of 8.4±1.6 (nmoles/min/ug protein AtPTPKIS1), and StPTPKIS1 + AtPTPKIS2 released phosphate at 8.8±1.5 (nmoles/min/ug protein StPTPKIS1), both at pH7. As before the assays required the presence of a strong reducing agent, such as DTT, in order to show this activity.

When phosphatase activity was assayed using the previously used phosphopeptides, the results remained unchanged, independent of the amount of AtPTPKIS2 added to the reaction.

These results show that the addition of the AtPTPKIS2 protein to these proteins has no effect on phosphatase activity against general phosphate substrates, or general phosphopeptides. While this does not preclude an interaction between PTPKIS1 and PTPKIS2 it suggests that such an interaction is unlikely to alter the activity of PTPKIS1 towards phosphorylated proteins. This suggests that the phenotype seen through mutation of the AtPTPKIS2 gene cannot be the result of the AtPTPKIS2 protein directly affecting the ability of PTPKIS1 to dephosphorylate proteins.

3.5.3 Activity on Glucan Substrates

In order to ascertain the potential for PTPKIS2 to alter the activity of PTPKIS1 towards phosphorylated glucan substrates *in vivo*, the activity against the naturally generated glucan substrates already examined was studied in the presence of AtPTPKIS2. Activity of AtPTPKIS1, StPTPKIS1 and D166A-FL, were compared with and without the presence of AtPTPKIS2 at equimolar concentrations.

As previously shown, AtPTPKIS1 and StPTPKIS1 efficiently liberate phosphate from amylopectin, starch, and starch isolated from SEX4 plants (Fig. 3.8) on their own. In presence of AtPTPKIS2 the activities of AtPTPKIS1 and StPTPKIS1 towards these phosphoglucan substrates both increased by approximately 4-fold (Fig 3.12). The rate of phosphate release in the presence of AtPTPKIS2 is comparable to the recorded values for the phosphatase activity of laforin upon amylopectin (Worby et al., 2006; Gentry et al., 2007) suggesting the possibility of similar roles *in vivo* for laforin and the PTPKIS1/PTPKIS2 combination. As expected, the inactive D166A-FL mutant of AtPTPKIS1 showed no significant rate of phosphate release whether or not AtPTPKIS2 was present, confirming that AtPTPKIS2 itself has no phosphatase activity towards these substrates.

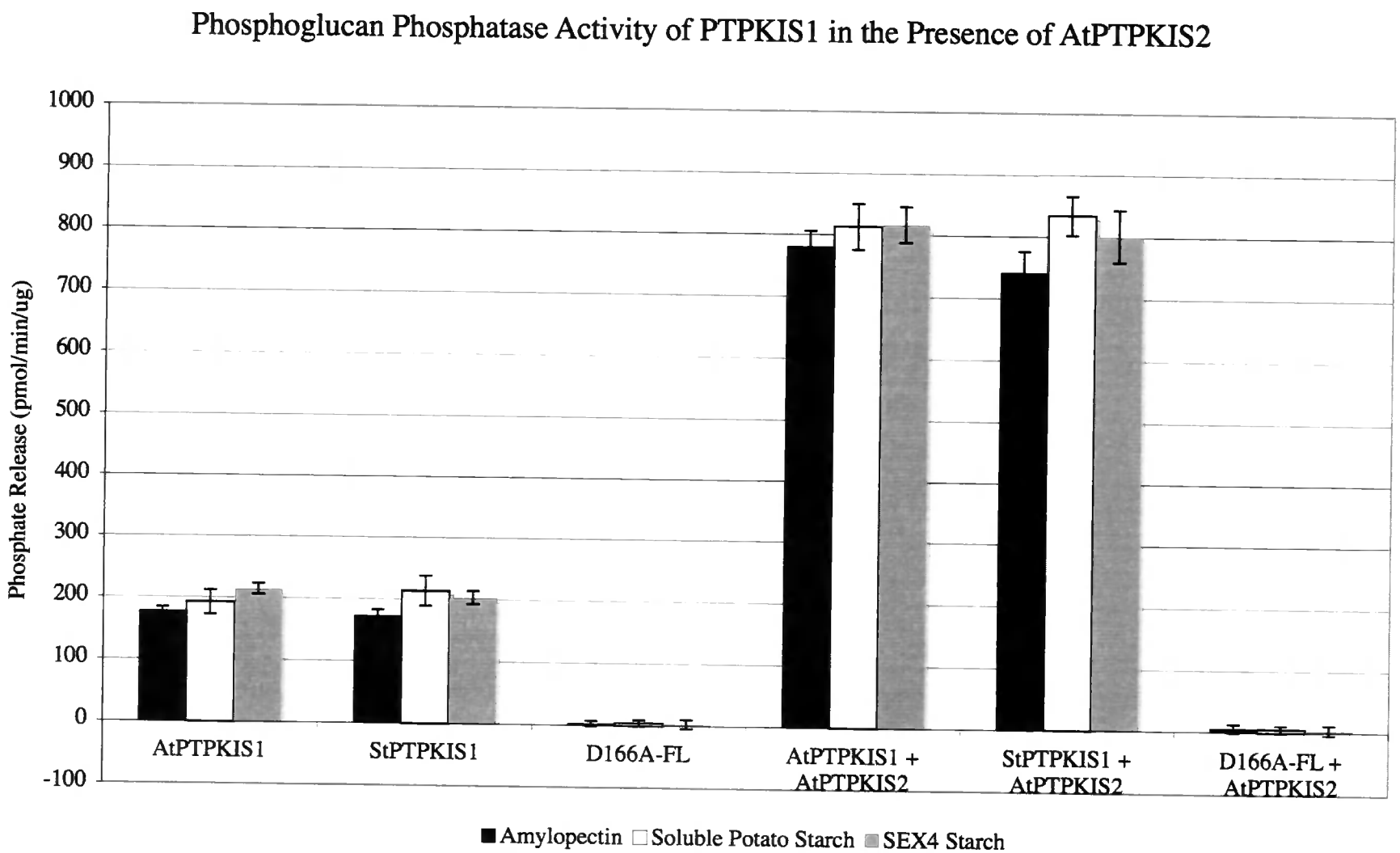


Figure 3.12 PhosphoGlucanPhosphatase activity of PTPKIS1 and PTPKIS2. Graph showing the activity of AtPTPKIS1, StPTPKIS1 and the AtPTPKIS1 mutant D166A-FL, with or without an equimolar concentration of recombinant AtPTPKIS2, against potato amylopectin (black), soluble potato starch (white) and starch isolated from SEX4-3 plants at the beginning of the dark period (grey).

3.6 PHOSPHATASE ACTIVITY OF PTPKL1

3.6.1 Identification and Isolation of PTPKL1

Analysis of data published in Fordham-Skelton *et al.* (2002) and Kerk *et al.* (2002) showed that an additional predicted phosphatase with sequence similarity to AtPTPKIS1 existed in the Arabidopsis genome (accession number: NP_566383), and was identified as the gene product of At3g10940. In Kerk *et al.*, (2002) it was shown that AtPTPKIS1 (At3g52180), AtPTPKIS2 (At3g01510) and At3g10940 also had sequence similarity to the animal laforin dual specificity protein phosphatases, as had been noted in Fordham-Skelton *et al.*, (2002). The At3g10940 gene encodes a protein predicted to be 282 amino acids long, containing a phosphatase domain homologous to that of PTPKIS1 and to a lesser extent PTPKIS2 (Fig. 3.9) , but lacking the C-terminal KIS/CBM region. This protein was termed PTPKL1 (**PTPK**is1 **L**ike **1**), and like other members of the PTPKIS family seemed to be present, and expressed in other plant species, through identification of homologues (Fig. 3.13) in the EST database (see Appendix 1.)

The coding sequence of AtPTPKL1 was isolated from cDNA generated from RNA isolated from leaf tissue under light conditions using the primers:
Forward: ATGAGTGTGATTGGAAGCAAGAGC
Reverse: TCAGGTTCCACGGAGGGCCCGAAC
These primers amplified the complete coding sequence including the start and stop codons. The predicted amplification product was a fragment of 852bp. When the *A.thaliana* cDNA was amplified with these primers, a single fragment was observed following agarose gel electrophoresis, of approximately 850bp – 860bp (Fig 3.14).

The fragment was purified by excision of band from gel, isolation of DNA, and cloning in a specialised vector for PCR products (pCR2.1), using the TOPO TA cloning system (Invitrogen). Clones resulting from this operation were characterised by DNA sequencing, which confirmed the clone as being identical to the predicted sequence, apart from a single nucleotide change, which did not alter the encoded amino acid.

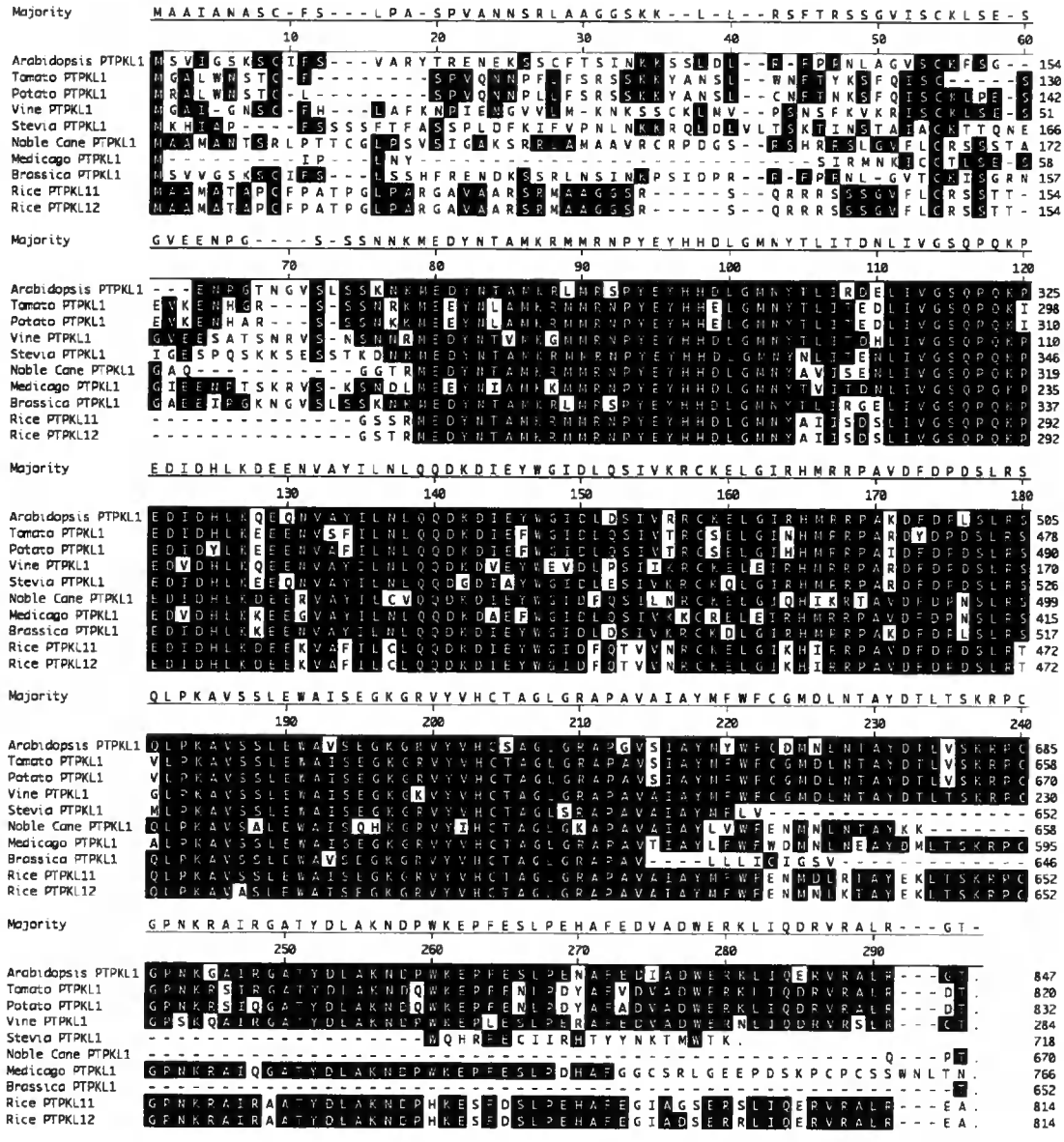


FIGURE 3.13 PTPKL1 Homologues. Alignment of the predicted PTPKL1 protein homologues. Those residues identical to the consensus sequence are highlighted in black

3.6.2 Expression Of Recombinant AtPTPKL1

Using the characterised clone encoding AtPTPKL1 as a template, the coding sequence was amplified using primers:

Forward -

GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAGTGTGATTGGAAGC AAGAGC

Reverse -

GGGGACCACTTTGTACAAGAAAGCTGGGTTCATTCCACGGAGGGCCCGA AC

These contained the addition of an attB1 and attB2 site respectively in order to allow the insertion into the gateway donor vector pDONR201 (Invirogen). This clone, termed ABP1, was then used in the gateway reaction with pDEST15 (Invitrogen) to generate a plasmid encoding AtPTPKL1 with an additional N-terminal GST Tag for purification (Fig. 3.14). This plasmid was transformed into *E.coli* strain BL21AI, which was used for protein expression. Cultures were grown at 28°C and induced with 0.01% L-Arabinose for 16 hours. GST-PTPKL1 was purified from cell supernatant by affinity batch absorption on glutathione-Sepharose, followed by washing and elution with Tris-HCl buffer at pH 8.0 (Fig. 3.14).

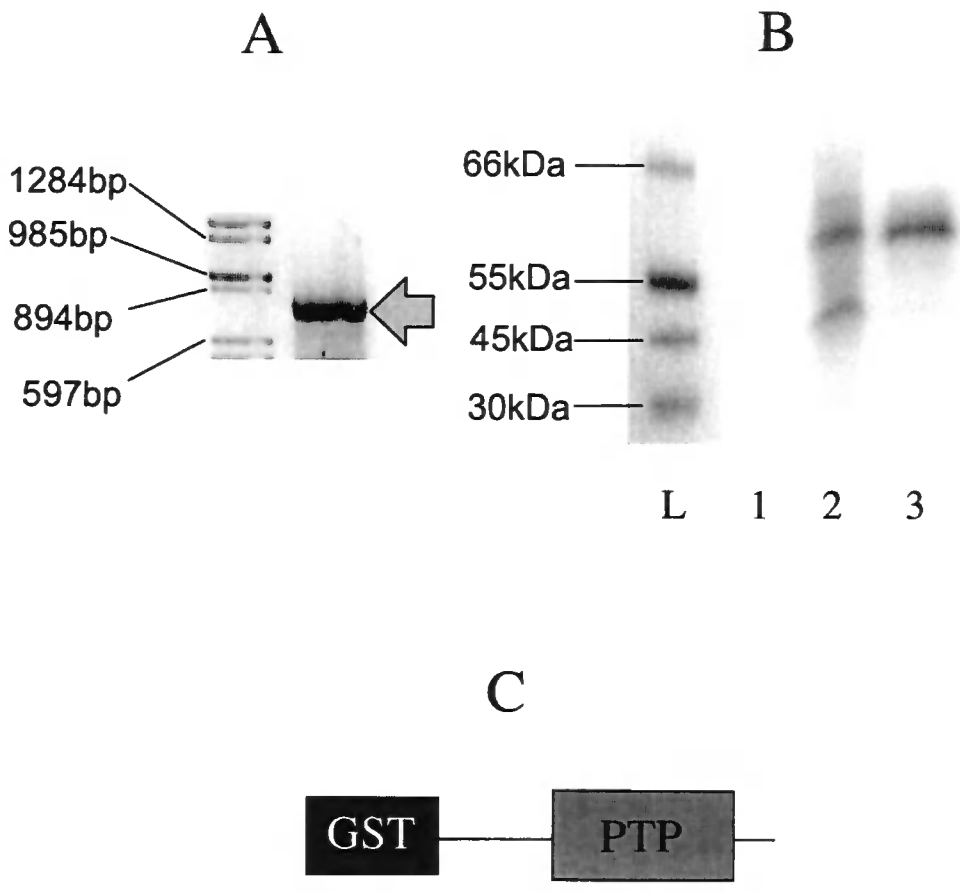


Figure 3.14 Expression of Recombinant AtPTPKL1. (A) Agarose gel image of RT-PCR amplification of AtPTPKL1 coding sequence. Arrow indicates band corresponding to AtPTPKL1 coding sequence. Left hand lane contains Lambda DNA/Eco47I (AvaII) Marker, numbers shown are in nucleic acid base pairs (bp) (B) Western blot using anti-GST antibody (Invitrogen) showing expression of AtPTPKL1. Lane 1, uninduced whole cell protein. Lane 2, induced whole cell protein. Lane 3, Purified protein. Left had lane is standard size markers, value shown are in kDa. (D) Diagrammatic representation of recombinant protein, containing GST tag, and with highlighted predicted phosphatase domain (PTP).

3.6.3 Dephosphorylation Of Standard Substrates

Purified recombinant AtPTPL1 protein hydrolysed the generic phosphatase substrate *p*-nitrophenyl phosphate at a maximum rate of 9.1±0.8 nmoles/min/ug protein This rate is comparable to the maximal rate for the already characterised LePTPKIS1 (Fordham-Skelton et al., 2002), as well as those of AtPTPKIS1 and StPTPKIS1 which are characterised above. Like these enzymes, AtPTPKL1 also required the presence of a strong reducing agent such as DTT to show activity, most likely to enable the reduction of an active site cysteine as found in the PTPKIS1 proteins.

Like AtPTPKIS1 and AtPTPKIS2, AtPTPKL1 showed no measurable phosphatase activity above background towards either of the phosphopeptide substrates used, suggesting it may play a role that does not require it to dephosporylate protein, or that the sequence of peptide it acts upon is highly specific, so that it does not show activity on the peptides used as general phosphoprotein phosphatase substrates.

3.6.4 Inhibition Of Phosphatase Activity

The phosphatase activity of AtPTPKL1 towards *p*-nitrophenyl phosphate was assayed in the presence of either vanadate or phenylarsinine oxide (PAO), both of which are used commonly to distinguish PTP's from other protein phosphatases. Both enzymes were inhibited in a concentration dependent manner by both agents (Fig 3.15). Vanadate was the more potent inhibitor of both enzymes, with an IC50 value of approx. 0.2μM, while the IC50 value for PAO was approx. 5μM. The inhibition observed was similar to that previously characterised for LePTPKIS1 (Fordham-Skelton *et al.*, 2002) and AtPTPKIS1 and StPTPKIS1 assayed earlier in this chapter, although the IC50 with PAO was slightly higher.

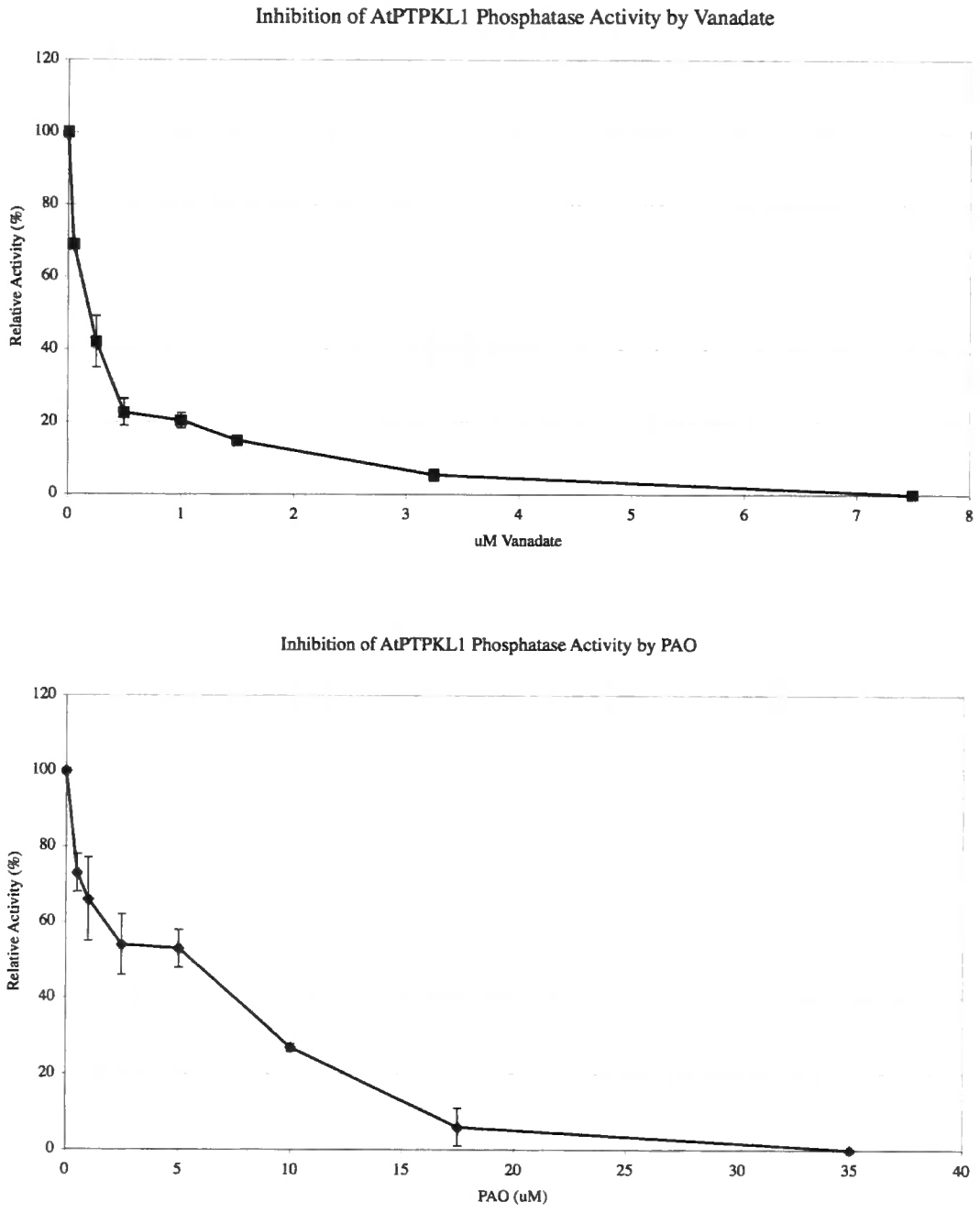


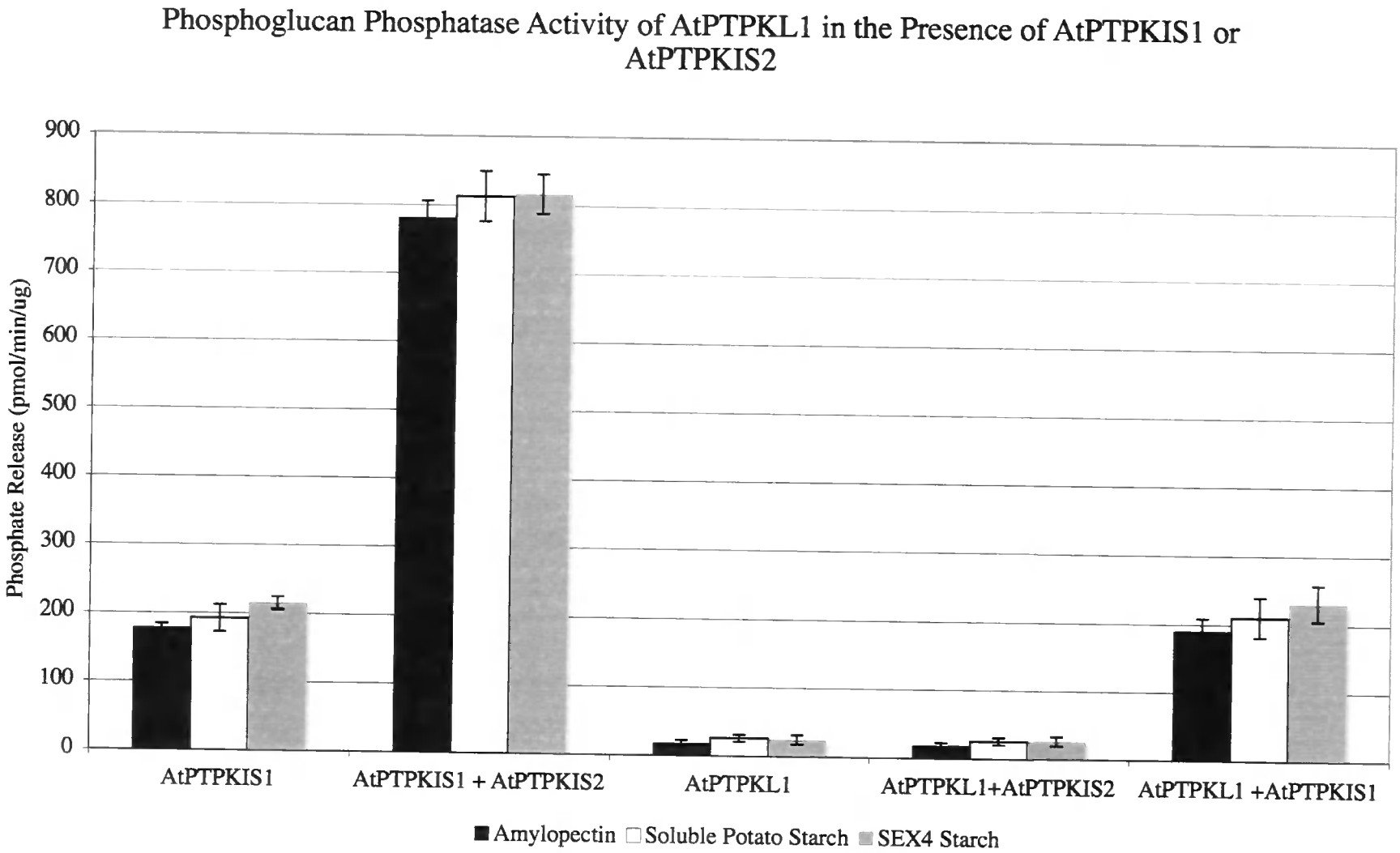
Figure 3.15 Inhibition of AtPTPKL1 by Vanadate and PAO. Graphs showing the effect of increasing concentrations of Vanadate or PAO upon the activity of AtPTPKL1, with activity measured relative to the enzymes activity without either Vanadate or PAO present

3.6.5 Phosphatase Activity On Phosphoglucans

Following the identification of phosphoglucan phosphatase activity in PTPKIS1, it was important to ascertain if such activity was present in PTPKL, given its lack of carbohydrate binding module.

As with PTPKIS1 and PTPKIS2, AtPTPKL1 was assayed for phosphatase activity using soluble starch. Analysis showed activity an order of magnitude lower than that of both AtPTPKIS1 and StPTPKIS1 (Fig. 3.16). Unlike PTPKIS1, addition of AtPTPKIS2 did not increase the phosphoglucans phosphatase activity seen. Combination of recombinant AtPTPKIS1 and AtPTPKL1, in equimolar concentrations, cause no increase in activity over what would be predicted as a sum of the individual activities.

Additional phosphatase assays, using Glucose 6-Phosphate as a substrate, were carried out with AtPTPKIS1, StPTPKIS1, AtPTPKIS2 and AtPTPKL1. It was not possible to show phosphatase activity against that particular phosphomonosacharide for any of the proteins tested (data not shown).



Figue 3.16 PhosphoGlucanPhosphatase activity of AtPTPKL1. Graph showing the activity of AtPTPKL1 with or without an equimolar concentration of recombinant AtPTPKIS1 or AtPTPKIS2, against potato amylopectin (black), soluble potato starch (white) and starch isolated from SEX4-3 plants at the beginning of the dark period (grey).

3.6.6 Identification Of PTPKL1 Homologues

Through analysis of the global sequence databases (genomic DNA sequences, cDNA databases, and Expressed Sequence Tag (EST) databases) it has been possible to identify sequences from other plant species, which are orthologous to those of AtPTPKL1 (Fig. 3.13). For most plant species only one PTPKL1 was identified, however in Rice two PTPKL proteins were identified. These differed in only 4 amino acid substitutions. Genomic study localised these sequences to chromosome 11 and 12 and so they have been termed PTPKL11 and PTPKL12 respectively. EST data suggests that both proteins are expressed in roots under the same conditions (see appendix 1). The presence of EST's encoding PTPKL1 homologues throughout the animal kingdom suggests the PTPKL1 protein is used in a common process throughout the plant kingdom, while the lack of homologues within other kingdoms, or at least homologues which do not contain carbohydrate binding modules, suggests a evolutionary role unique to plants.

Study of EST databases for lower plants provided a group of EST's from *Chlamydomonas reinhardtii*, a green algae, which encode a protein, of 298 amino acids in length, showing great similarity to that of AtPTPKL1 (Fig. 3.17). This protein shows greater similarity to AtPTPKL1 than to previously identified dsPTP's in green alga (e.g. *Chlamydomonas eugameto* VH-PTP13, accession: Q39491). However, *C. reinhardtii* may contain several dsPTPs, since EST searches have also suggested at the presence of a VH-PTP13 like protein in this species (see appendix1).

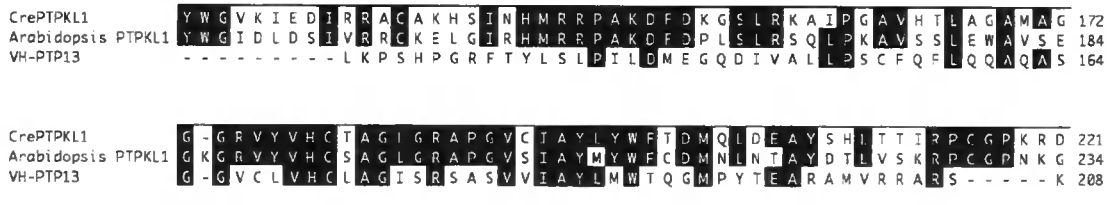


Figure 3.17 *Chlamydomonas reinhardtii* PTPKL1. Alignment of the PTP domains of AtPTPKL1 and VH-PTP13 with the predicted *Chlamydomonas reinhardtii* PTPKL1 (CrePTPKL1). Residues matching the consensus are highlighted in black.

3.7 DISCUSSION

Work described here on the phosphatase activity of AtPTPKIS1 and StPTPKIS1 shows great levels of consistency with previously published results (Fordham-Skelton et al., 2002; Kerk et al., 2006; Sokolov et al., 2006; Gentry et al., 2007). For example both enzymes show similar Km values with *p*-nitrophenyl phosphate to that recently published (Gentry et al., 2007). In particular, the identification of PTPKIS1 phosphoglucan phosphatase activity (Gentry et al., 2007) and the absence of significant protein phosphatase activity towards phosphorylated peptides is consistent with the published data. While the ability to dephosphorylate substrates, such as starch and amylopectin, seems to be contradicted by its inability to dephosphorylate elongated glucans previously phosphorylated by GWD1, there are a number of reasons to account for this.

It may be possible that the location of phosphate along the carbohydrate backbone is important to the activity of PTPKIS1. This possibility is supported by the effect variation in maltooligosacharide length has on the phosphatase activity of PTPKIS1 towards the general phosphatase substrate *p*-nitrophenyl phosphate. Due to GWD1's ability to phosphorylate glucose residues at both the C6 and C3 position, it is possible to discard the notion of PTPKIS1 only acting upon one of these phosphatase sites as the reason for its inability to dephosphorylate the synthetic substrate. *In vivo*, however GWD3 is also present to increase the percentage of phosphorylation at the C3 position (Baunsgaard et al., 2005; Kotting et al., 2005).

While the location of the phosphate residue may be of importance, so may be the structure of the surrounding glucan. It is possible that the CBM of PTPKIS1 is specific to positions within a starch granule, with branch points, and amylose content playing an important role. The elongation of the glucan chains on a glycogen backbone may not provide the required structure for PTPKIS1 binding. Recent work by Hejazi *et al.* (2008) has shown that the physical arrangements of glucans to effect activity of GWD *in vitro*, this may be the case with PTPKIS1.

Expression and analysis of the AtPTPKIS2 showed a protein with no identifiable phosphatase activity towards any of the protein peptides or carbohydrates used previously. In addition it also failed to dephosphorylate the general phosphatase substrate *p*-nitrophenyl phosphate. It is likely this inactivity is due to the replacement of its active site histidine with threonine. This histidine usually forms part of the conserved diagnostic motif (HCx5R) found in other PTPs. The absence of this histidine residue, while being a possible explanation of the inactivity of this predicted phosphatase, does not help to explain its role *in vivo*.

The increase in phosphoglucan phosphatase activity when PTPKIS1 and PTPKIS2 were combined may suggest a role for the PTPKIS2 protein as a partner for PTPKIS1. While combining both shows an increase in activity against phosphoglucans, it does not increase activity against peptide substrates or *p*-nitrophenyl phosphate, suggesting that the increase in activity is likely to be due to a more favourable association with the phosphoglucans carbohydrate structure, caused by the presence of the CBM on both PTPKIS1 and PTPKIS2. When full length PTPKIS2 is replaced in this assay with just its CBM as used in Chapter 4, there is no noticeable increase in PTPKIS1 phosphoglucan phosphatase activity (data not shown) suggesting that the interaction between the 2 proteins is at some other region. Possibly this involves the predicted PDZ domain proposed in Fordham-Skelton *et al.* (2002).

There are other dimeric phosphatases, given the functional similarity to PTPKIS1, it is of particular interest that laforin, which also contains a carbohydrate-binding domain, albeit at its N-terminus, has also been shown to form dimers which exhibit increased activity over the monomeric form (Liu et al., 2006).

Expression and analysis of the AtPTPKL1 protein, a predicted PTP domain containing protein, lacking the CBM/KIS domain of PTPKIS1 or PTPKIS2 showed a protein with comparable activity towards *p*-nitrophenyl phosphate to PTPKIS1. AtPTPKL1 however showed a greatly reduced activity towards starch, most likely due to absence of a CBM like that found in PTPKIS1. AtPTPKL1 activity towards starch was comparable to that shown for activity of AtPTPKIS1 with mutations

inactivating its CBM (Gentry et al, 2007), suggesting that this reduced activity is solely due to its lack of a CBM.

While PTPKIS1 is likely to act upon the starch granule, it may be possible that PTPKL1 acts upon shorter chain phosphoglucans produced during the degradation of starch granules in the dark period (Smith et al., 2005). This however would require the generation of shorter length maltooligosacharides with specific phosporylation patterns in order to study further.

As previously identified by Fordham-Skelton *et al.*, (2002) and supported by the work of Sokolov *et al.*, (2006) it is shown that reducing conditions are required in order for PTPKIS1 to show phosphatase activity. It has also shown that AtPTPKL1 requires reducing conditions in order to show phosphatase activity. The reduction of the active site may be one of the mechanisms (*in vivo*) for regulation of activity. It has been shown that other proteins involved in starch degradation, such as GWD1, require reducing conditions to become active (Mikkelsen et al., 2005). This redox control may be a common element in regulation of starch metabolism.

For all the above phosphatase, it may be possible that they do in fact act upon specific protein sequences, however recent assays using the D166A protein as a substrate trap have yielded no likely protein candidate for dephosphorylation by AtPTPKIS1 (Nana Chougule, School of Biological and Biomedical Sciences, Durham University, Personal Communication), supporting its classification as a phosphoglucan phosphatase.

Chapter 4

CARBOHYDRATE BINDING ACTIVITY OF PTPKIS FAMILY PROTEINS

4.1 INTRODUCTION

In the original publication identifying PTPKIS1, Fordham-Skelton *et al.* (2002) proposed that this protein, which was assumed to be a protein phosphatase on the basis of sequence similarity, was novel due to its C-terminal domain, which was similar in sequence to a domain designated KIS. KIS domains are present in a variety of proteins and are known to mediate protein-protein interactions; specifically, they are responsible for binding the scaffold protein(s) to the regulatory SNF1 kinase in yeast as part of the SNF1 complex. Fordham-Skelton *et al* (2002) showed interaction between the recombinant KIS domain of AtPTPKIS1 and the *A.thaliana* SnRK, AKIN11 (a protein related to AMPKs in animals (Halford *et al.*, 2000)) in both pull down and yeast-2-hybrid analysis. These results supported a role for PTPKIS1 in regulating the activity of the plant regulatory kinase complex. However the pull down assay, included another protein containing a KIS domain, designated AKINβγ (Lumbreras *et al.*, 2001). The KIS domain from AKINβγ appeared to showed a greater affinity for AKIN11 than the KIS domain from AtPTPKIS1. In itself, this result would not be sufficient to call the hypothesised role for the KIS domain in PTPKIS1 into doubt, but further data have since appeared which suggest that the binding to AKIN11 observed by Fordham-Skelton *et al.* (2002) *in vitro* may not be relevant *in vivo*.

Hudson *et al.* (2003) and Polekhina *et al.* (2003) showed that the KIS domain region of the mammalian AMPKβ interacted with glycogen and thus acted as a carbohydrate-binding domain. The designation "KIS" domain may therefore be incorrect in this case. The sequence similarity between the "KIS" domains in

AMPKβ and PTPKIS1 suggests that the "KIS" domain in the latter may also have a carbohydrate-binding function. Subsequent work by this author (Niittyla *et al.*, 2006) as well as additional work (Kerk *et al.*, 2006; Sokolov *et al.*, 2006; Gentry *et al.*, 2007), has confirmed this prediction, and shown that PTPKIS1 has the ability to bind to carbohydrates through the C-terminal region, previously termed the KIS domain. This region has therefore been redesignated as a Carbohydrate Binding Module (CBM), and will be referred to as this subsequently.

This chapter will deal with the functional characterisation of the Carbohydrate Binding Module (CBM) found in AtPTPKIS1 and AtPTPKIS2, the identification of key residues, and the study of a conserved 'sugar tong' region within the CBMs of PTPKIS family proteins.

4.2 IDENTIFICATION OF A CBM IN PTPKIS FAMILY MEMBERS

Using the work by Hudson *et al.*, (2003) and Polekhina *et al.*, (2003) as a starting point, an amino acid sequence alignment was created using the AMPKβ subunits from mammals, AMPKβ subunits homologues from *A.thaliana* and GAL83 from *S.cerevisiae*. In addition the predicted KIS domains from 7 PTPKIS1 proteins, identified in chapter 3, and the glycogen binding domain from the GBE protein found in *E.coli* were added to the alignment (Fig. 4.1). All the sequences used for the alignment contained a clearly defined CBM; this domain is not defined in the Interpro database, but is present as an unclassified domain in the Panther database (PTHR10353) (www.pantherdb.org). The domain has sequence similarity to carbohydrate-binding domains CBM20 and CBM21, as defined in the CAZY database (http://afmb.cnrs-mrs.fr/~pedro/CAZY/cbm.html).

This alignment (Fig. 4.1) showed a number of residues conserved among the aligned sequences, in particular those identified by Hudson *et al.*, (2003) and Polekhina *et al.*, (2003) as being required by AMPKβ subunits to bind to glycogen. Previous alignments of PTPKIS1 and PTPKIS2 (Chapter 3) showed a high degree of homology in those regions, supporting the potential of the predicted PTPKIS2 KIS

domain also acting as a CBM. Alignment of predicted sequences for PTPKIS1 and PTPKIS2 show conservation of residues with potential roles in carbohydrate binding (Fig. 4.2).

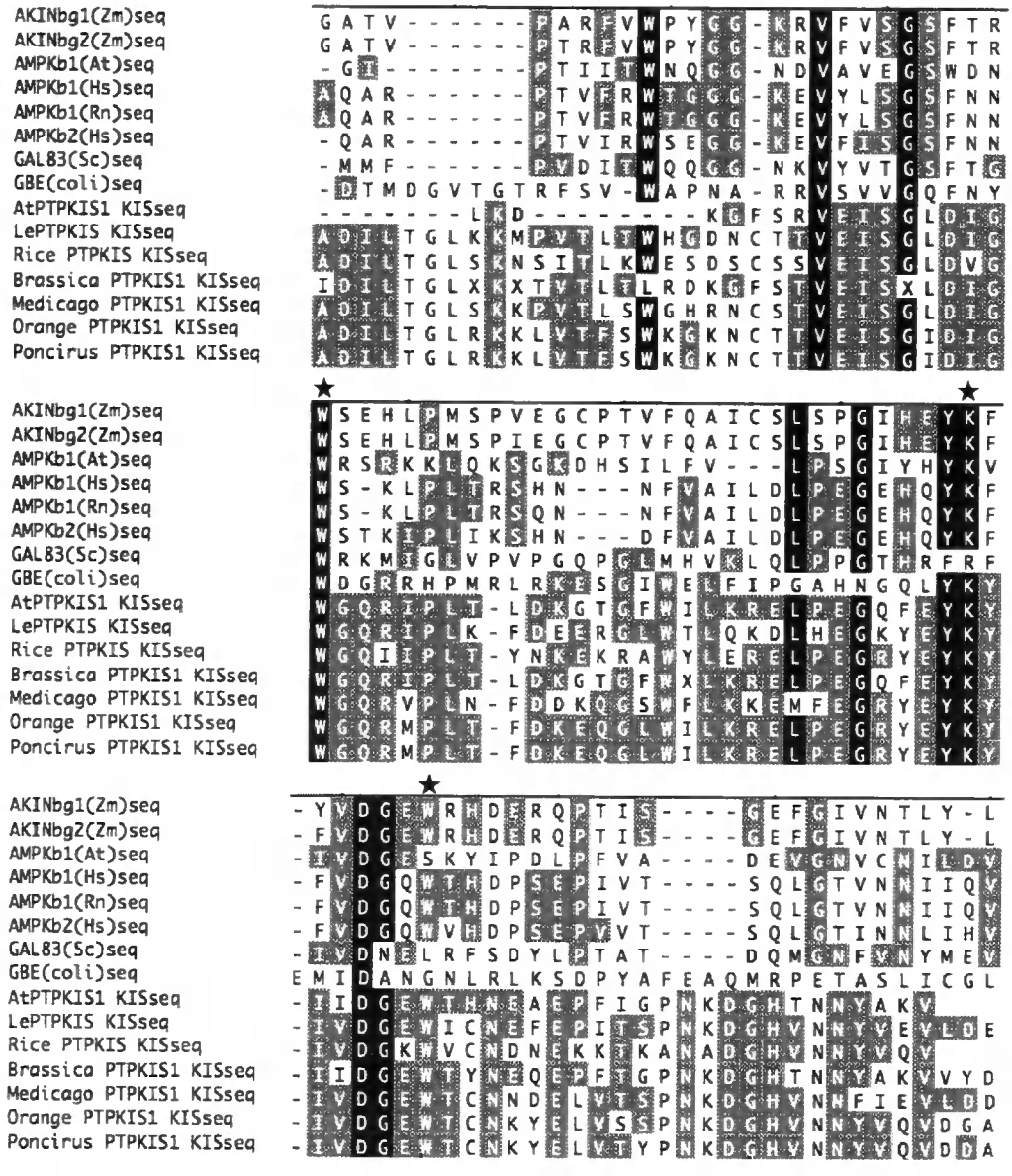


Figure 4.1Alignment of CBMs from PTPKIS1, *E. coli* GBD, AKINβγ, and Snf1 family of β subunits. Zea mays AKINβγ1 (AKINbg1(Zm), accession number NP_001105005), Zea mays AKINβγ2 (AKINbg2(Zm), accession number NP_001105555), *A. thaliana* AMPKβ1, (AMPKb1(At), accession number NP_197615), Human AMPKβ1 (AMPKb1(Hs), accession number Q9Y478), Rat AMPKβ1 (AMPKb1(Rn), accession number P80386), Human AMPKβ2 (AMPKb2(Hs), accession number NP_005390), S. cerevisiae GAL83 (Gal83(Sc), accession number Q04739), *E. coli* glycogen branching enzyme (GBE(*E.coli*), accession number P07762). PTPKIS1 sequences were generated from EST alighnments (see appendix 1). Amino acid sequences were aligned with clustalw. Amino acids identical to the concensous sequence were highlighted in black, amino acids similar to 50% of other sequences are shaded in grey. Residues highlighted with a star (★) were shown by Polekhina *et al.*, (2003) to be required for carbohydrate binding.

Figure 4.2 Alignment of predicted CBMs for PTPKIS1 and PTPKIS2. Alignment of predicted CBMs using sequences generated from alignment of EST data (appendix1). Amino acid sequences were aligned with clustalw. Amino acids conserved throughout all sequences were highlighted in black. Residues highlighted with a star (★) were shown by Polekhina *et al.*, (2003) to be required for carbohydrate binding. Numbers shown refer to the position of the first amino acid on each line in the wild type protein.

4.2.1 Expression Of Predicted CBMs

Constructs to produce CBMs from AtPTPKIS1 and AKINβγ have been described by Fordham-Skelton *et al.*, (2002). These constructs produce CBM as a C-terminal fusion to glutathione-S-transferase (GST); GST itself, produced by the vector, (pGEX5) was used as a negative control for binding assays. Assembly of the constructs, and expression of the recombinant proteins, is described in Materials and Methods. GST fusions were purified from cell supernatant of recombinant bacterial clones expressing the proteins by affinity absorption on glutathione-Sepharose columns (Amersham), followed by washing and elution with Tris-HCl buffer at pH8.0

Using the full length sequence for AtPTPKIS2 as a template, primers were designed in order to amplify the predicted KIS/CBM and allow insertion into the cloning vector pGEX5X1 using restriction enzymes (EcoRI and XhoI) as had been used

previously for the KIS/CBM of AtPTPKIS1. The primers used to amplify the CBM and containing the required restriction sites used were:

Forward: 5' – GAATTCTCTGTAACCTTTGTG – 3'

Reverse: 5' – CTCGAGATCACCAACTACAAT – 3'

This GST-KIS (AtPTPKIS2) construct was used to generate recombinant protein for use in binding assays (see materials and methods). The protein was purified by affinity chromatography on glutathione-Sepaharose as described above

4.2.2 Carbohydrate Binding

In order to determine if predicted CBMs from AtPTPKIS1 and AKINβγ showed binding to carbohydrate as seen in Polekhina *et al.*, (2003), pull down assays using bovine glycogen (Sigma) were carried out (as described in materials and methods) using GST-KIS (AtPTPKIS1), GST-KIS (AKINβγ), and GST as soluble proteins. Bovine serum albumin (BSA) was used as a non-specific binding blocker in these assays. Glycogen was pelleted by centrifugation after incubation of the recombinant proteins with soluble glycogen. The presence, in identifiable quantities, of GST-KIS (AtPTPKIS1) in the 100,000g pellets only when incubated in the presence of glycogen suggests binding of the GST-KIS (AtPTPKIS1) protein to glycogen (Fig. 4.3). The inability to identify any bands corresponding to GST alone gave no band in the 100,000g pellet both with and without glycogen showing that this co-sedimentation of the GST-KIS(AtPTPKIS1) protein with glycogen is due to the KIS region of the GST-KIS (AtPTPKIS1) protein. GST-KIS (AKINβγ) showed no presence in the 100,000g pellets with and without glycogen incubation, suggesting that it has no direct binding to glycogen.

While this assay showed the presence of glycogen binding, additional pull-down assays were carried out using granular potato starch as a more physiologically relevant material. Pull down assays using granular starch (Sigma), were carried out (as described in materials and methods) using GST-KIS (AtPTPKIS1), GST-KIS (AKINβγ), GST-KIS (AtPTPKIS2), and GST. The presence, in identifiable quantities, of GST-KIS (AtPTPKIS1) and GST-KIS (AtPTPKIS2) bands in the 10,000g pellets only when incubated in the presence of granular potato starch

suggests binding of the GST-KIS (AtPTPKIS1) and GST-KIS (AtPTPKIS2) proteins to granular potato starch (Fig. 4.4). The inability to identify any bands corresponding to GST alone with and without starch suggest this co-sedimentation with starch is due to the KIS region of the GST-KIS (AtPTPKIS1) or GST-KIS (AtPTPKIS1) proteins. GST-KIS (AKINβγ) showed no presence in the 10,000g pellets with and without starch incubation, suggesting that it has no direct binding to starch.

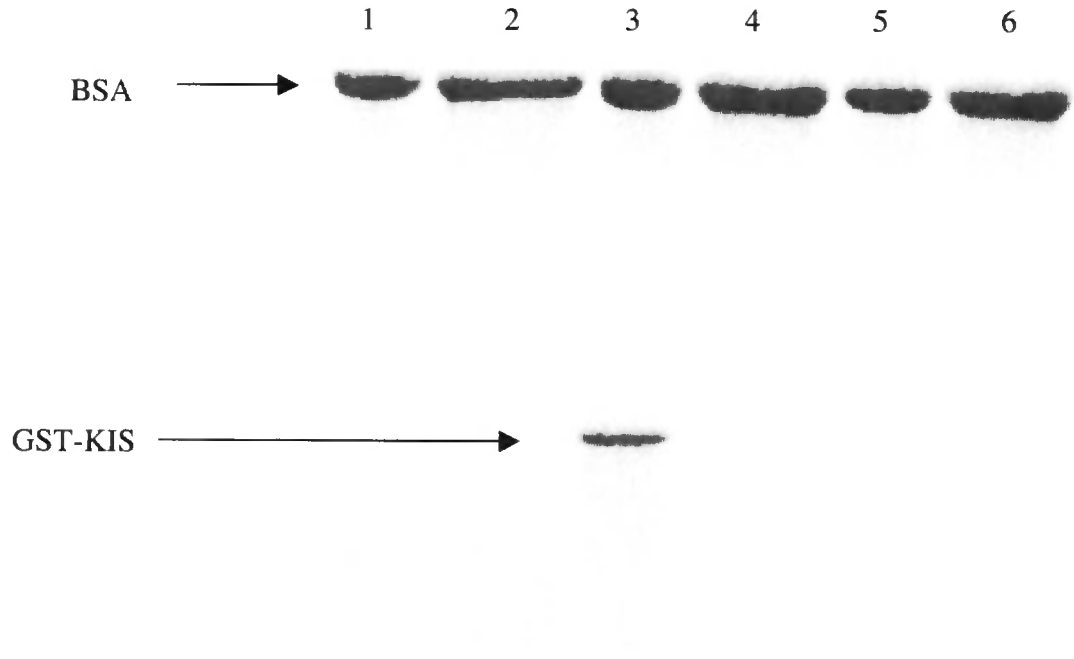


Figure 4.3. Glycogen pull down assay with GST-KIS (AtPTPKIS1) and GST-KIS (AKINβγ) and GST. As published by the author previously in Niittyla *et al.*, (2006) SDS-PAGE gel, stained with Coomassie Brilliant Blue G250, showing protein content of pellets from glycogen binding assay. 1&2 are from tubes containing GST and BSA only, 3&4 are from tubes containing GST-KIS and BSA only, and 5&6 are from tubes containing GST-AKIN βγ(KIS) and BSA only. Tubes 1, 3, and 5 contained glycogen, while tubes 2,4, and 6 did not contain glycogen. (This figure was previously published in Niittyla *et al.*, (2006))

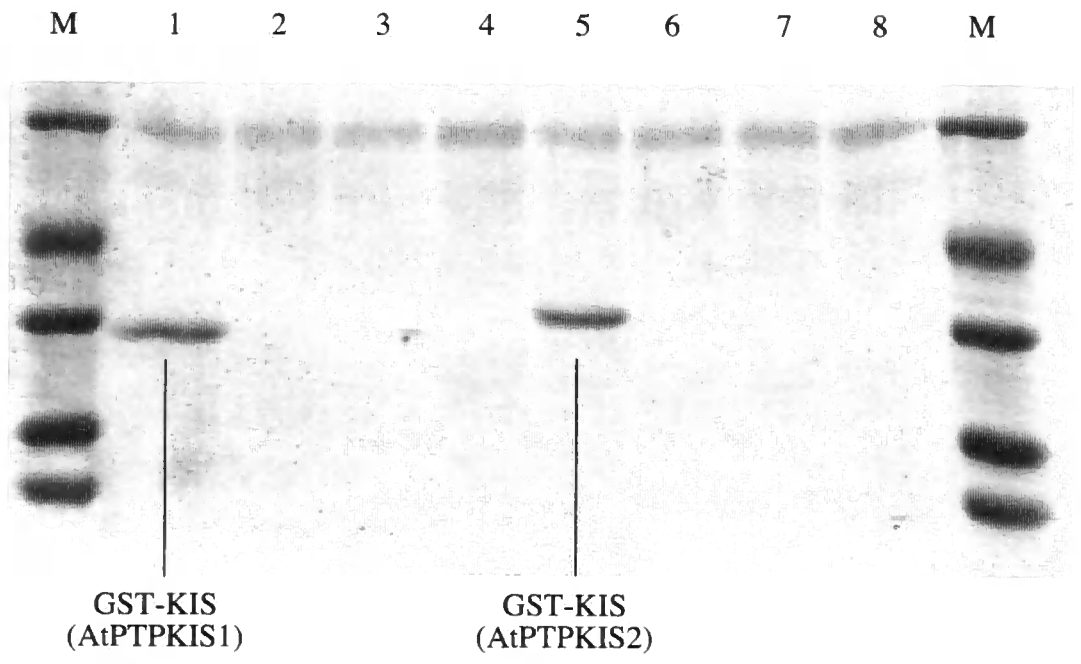


Figure 4.4. Starch pull down assay with GST-KIS (AtPTPKIS1), GST-KIS (AKINβγ), GST and GST-KIS (AtPTPKIS2). SDS-PAGE gel, stained with Coomassie Brilliant Blue G250, showing protein content of pellets from granular starch binding assay. 1&2 are from tubes containing GST-KIS (AtPTPKIS1) and BSA only, 3&4 are from tubes containing GST-KIS (AKINβγ) and BSA only, 5&6 are from tubes containing GST-KIS (AtPTPKIS2) and BSA only, and 7&8 are tubes containing GST and BSA only. Tubes 1, 3, 5, and 7 contained granular starch, while tubes 2,4, 6, and 8 did not contain granular starch. Lane M is standard protein size markers.

The inability of the CBM in GST-KIS (AKINβγ) to bind to glycogen or starch, suggests that this carbohydrate-binding ability is not common to all KIS domains. Using the Gamier-Robson method for predicting beta sheets in protein tertiary structure, it was possible to compare KIS/CBM regions of AtPTPKIS1, AKINβγ and a fungal starch-binding domain (accession no. AAP04499) (Fig. 4.5)

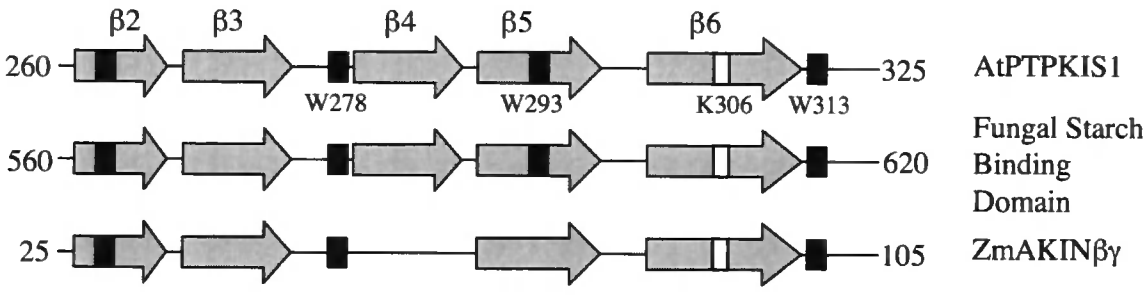


Figure 4.5. Diagrammatic representation of beta sheet arrangement in AtPTPKIS1, ZmAKINβγ and a fungal starch binding domain. Grey arrows represent beta sheets, black bands represent conserved tryptophan residues thought to be involved in carbohydrate binding, and white bands represent conserved lysine residues thought to be involved in carbohydrate binding. *Arabidopsis thaliana* PTPKIS1 (AtPTPKIS1, accession number CAC17593, residues 260-325), *Aspergillus niger* Glucoamylase (accession number AAP04499, residues 560-620), *Zea mays* AKINβγ1 (ZmAKINβγ1, accession number NP_001105005, residues 25-105). β2-β6 refer to the β-sheets described previously (Polekhina *et al.*, 2003).

The Gamier Robson method shows that AKINβγ does not contain a predicted β4 region common to both the AtPTPKIS1 and Fungal Starch Binding Domain. It also shows that AKINβγ lacks the conserved tryptophan residue found in or around the β5 region when compared with AtPTPKIS1 and the fungal starch binding domain. If this prediction is correct, the shape of the binding site may be altered from that of a CBM. Combined with the loss of the conserved tryptophan residue, this could account for its inability to bind carbohydrates, while still being classified as a KIS domain on the basis of sequence similarity.

4.3 BINDING PARAMETERS OF PTPKIS CBM

Following the identification of carbohydrate binding activity in the "KIS" domains from AtPTPKIS1 and AtPTPKIS2, the binding parameters of these domains with physiological substrates were studied to obtain potential insights into the possible role of these proteins in vivo, and to compare these CBMs with previously characterised CBMs found in proteins involved in starch metabolism, such as GWD1 (Smith *et al.*, 2005; Mikkelsen *et al.*, 2006).

Interactions between CBMs and carbohydrate were studied using the co-sedimentation (pull-down) method used previously, using granular potato starch as a substrate. In addition, interactions with solublised starch were studied using PAGE gels containing variable soluble starch concentrations, measuring retardation in the protein's mobility relative to gels containing no starch. This method has previously been used in the characterisation of interactions between proteins and carbohydrates in starch-binding enzymes (Blennow *et al.*, 1998; Mikkelsen *et al.*, 2006).

4.3.1 Co-Sedimentation Assay Binding Parameters

Co-sedimentation experiments were carried out in the presence of varying concentrations of potato starch. The fraction of protein present in the pellet (i.e. bound to starch) was estimated by quantitating the protein present in bands formed on SDS-PAGE gels (see materials and methods). This allowed a curve of fraction of protein bound against starch concentration to be plotted; results are shown in figs. 4.6 and 4.7. Using curve-fitting, it was possible to calculate the K_d values (estimated as the concentration of starch at which half the maximum amount of protein was bound) of 4.88mg/ml and 5.41mg/ml for GST-KIS(AtPTPKIS1) and GST-KIS(AtPTPKIS2) respectively. B_{max}, the maximum fraction of protein bound was estimated as 0.53 and 0.43 for GST-KIS(AtPTPKIS1) and GST-KIS(AtPTPKIS2) respectively (Table 4.1). Using an equation equivalent to the Hill equation (Mikkelsen *et al.*, 2006) it was possible to evaluate the degree of co-operativity for granular starch binding. A Hill factor (h) of 1.03 and 1.08 was calculated for GST-KIS(AtPTPKIS1) and GST-KIS(AtPTPKIS2) respectively (Fig. 4.8). The results

quoted above are close to 1, indicating that each binding event is independent. The Hill factors support the hypothesis that the predicted KIS/CBM domain contains only one binding site for polysaccharide molecules, since multiple binding events would involve significant steric hindrance giving Hill factors significantly <1.

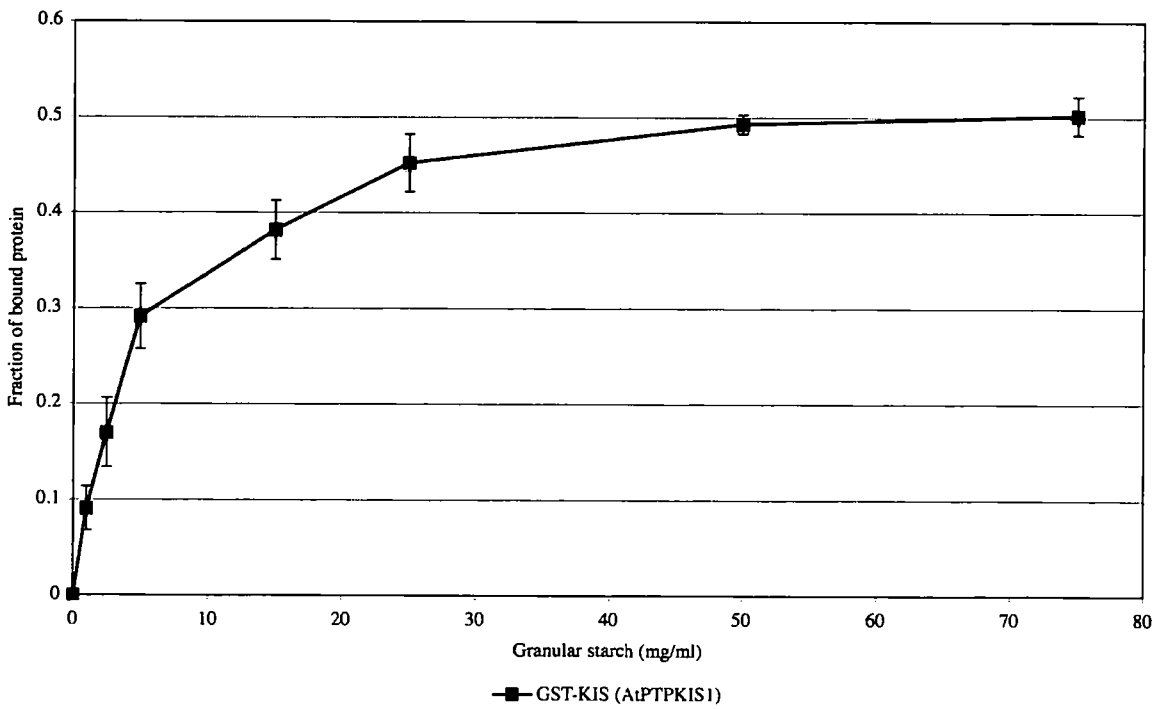


Figure 4.6 Sedimentation of GST-KIS(AtPTPKIS1) with Starch. Graph showing the fraction of bound GST-KIS(AtPTPKIS1) protein as the amount of granular starch was increased.

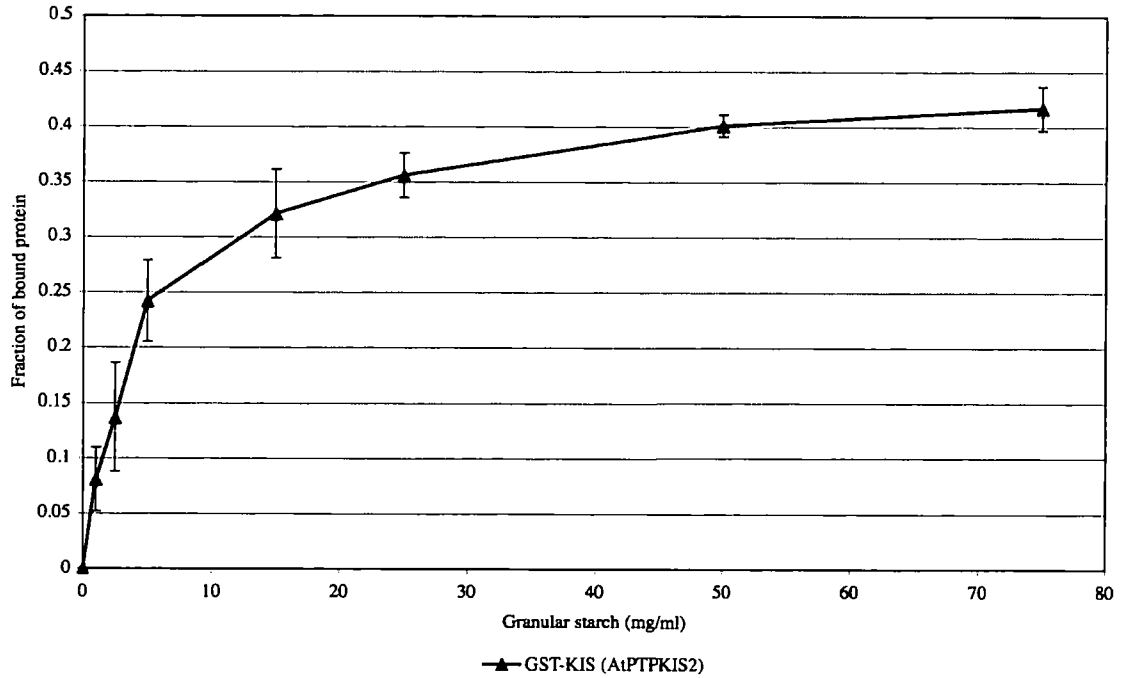


Figure 4.7 Sedimentation of GST-KIS(AtPTPKIS2) with Starch. Graph showing the fraction of bound GST-KIS(AtPTPKIS2) protein as the amount of granular starch was increased.

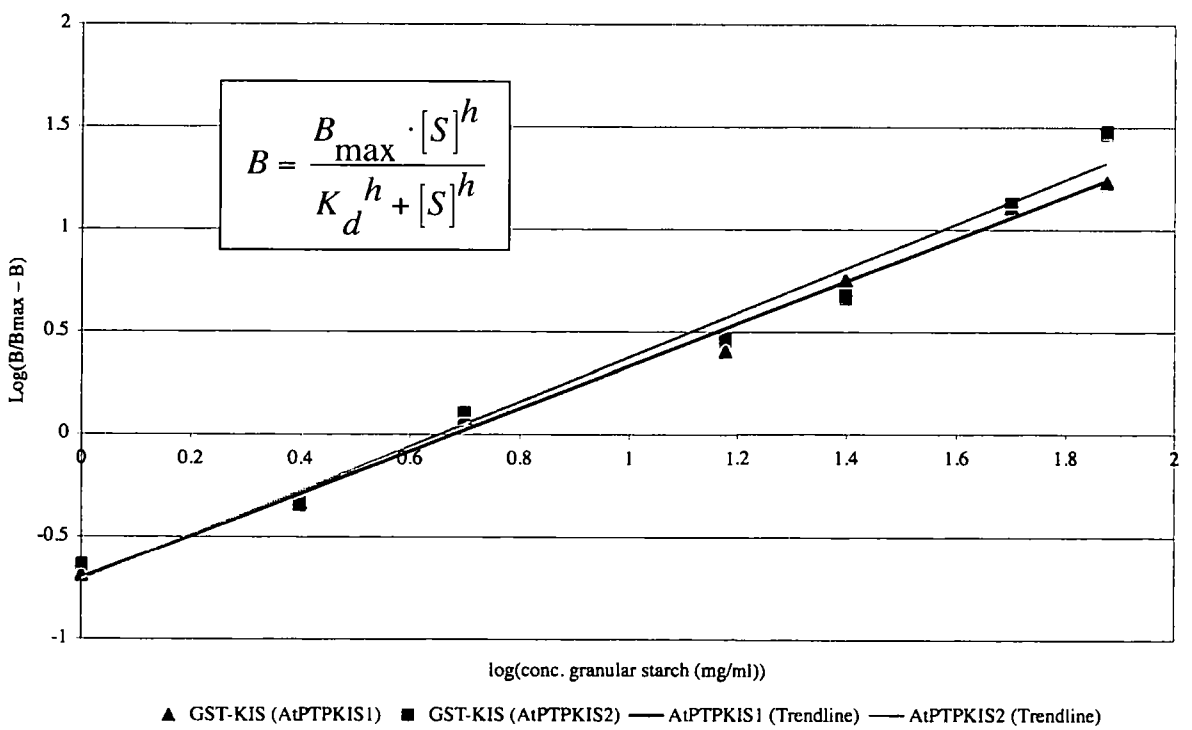


Figure 4.8 Graphical Representation of Hill Factor as Calculated for GST-KIS(AtPTPKIS1) and GST-KIS(AtPTPKIS2). The gradient of the trendline gives the Hill factor (h) for each of the assays. The insert is the Hill equation, as shown in Mikkelsen *et al* (2006).

The values for K_d and B^{max} determined for GST-KIS(AtPTPKIS1) and GST-KIS(AtPTPKIS2) are comparable to those observed for other CBMs (Penninga *et al.*, 1996; Paldi *et al.*, 2003; Mikkelsen *et al.*, 2006). This comparability of binding parameters supports a role in carbohydrate binding for the predicted KIS/CBM domain.

In order to further characterise the binding parameters of these predicted KIS/CBM domains, inhibition of binding by β-cyclodextrin, a commonly used inhibitor of carbohydrate binding in proteins containing CBM20 and CBM21 domains (Paldi *et al.*, 2003) was determined. Using the co-sedimentation assay as a basis for this analysis, protein was pre-incubated with variable concentrations of β-cyclodextrin before granular starch (Sigma) was added. The fraction of protein in the sedimented pellet was determined as above, by SDS-PAGE gel analysis and quantitation of protein bands on the gel (see materials and methods) (Fig. 4.9 & 4.10). Inhibition of binding was found to be similar to that of previously characterised KIS/CBM domains (Polekhina *et al.*, 2003). Half maximal inhibition occurring with ~ 1.1mM β-cyclodextrin for GST-KIS(AtPTPKIS1) and with ~ 0.9mM β-cyclodextrin for GST-KIS(AtPTPKIS2).

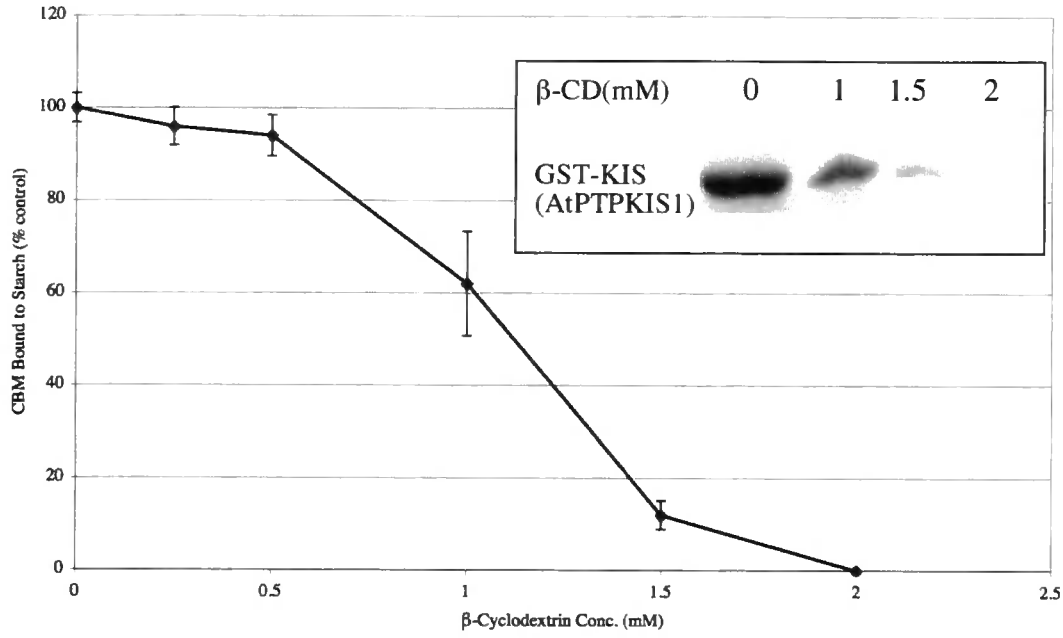


Figure 4.9 Inhibition of GST-KIS(AtPTPKIS1) Starch Binding. Graph showing the effect increasing concentration of the starch binding domain inhibitor, β-Cyclodextrin (β-CD), has upon GST-KIS(AtPTPKIS1) binding and subsequent sedimentation with granular starch. Insert shows Coomassie staining of SDS gel containing fractions from the experiment.

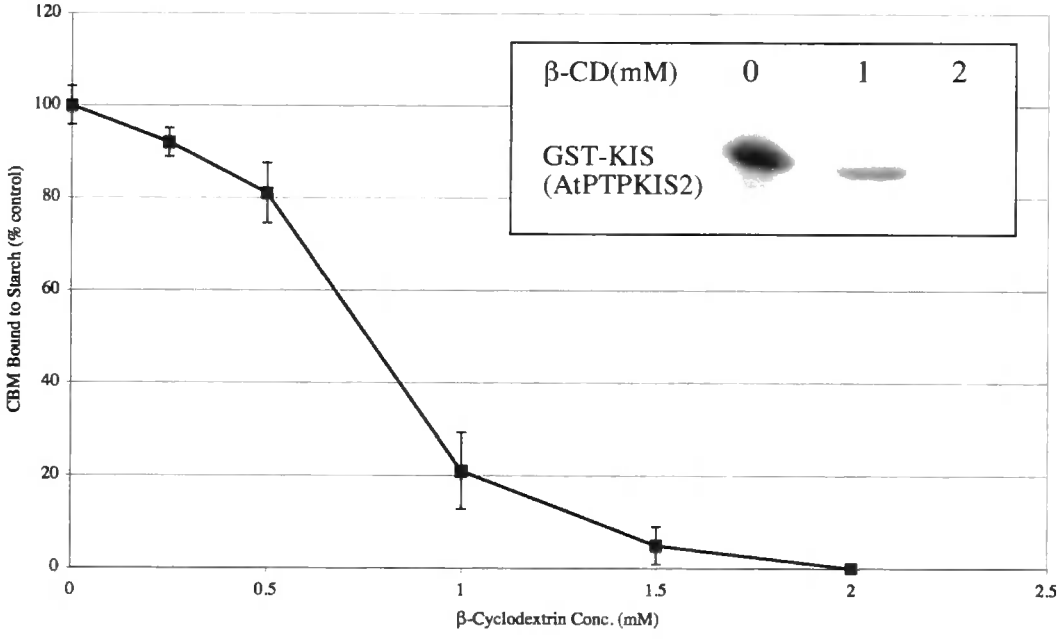


Figure 4.10 Inhibition of GST-KIS(AtPTPKIS2) Starch Binding. Graph showing the effect increasing concentration of the starch binding domain inhibitor, β-Cyclodextrin (β-CD), has upon GST-KIS(AtPTPKIS2) binding and subsequent sedimentation with granular starch. Insert shows Coomassie staining of SDS gel containing fractions from the experiment.

4.3.2 Estimation of Carbohydrate Binding Parameters by PAGE Mobility Shift Assays

In addition to the co-sedimentation study, binding of CBMs to carbohydrates was studied using PAGE gels containing soluble starch. Using a set of native PAGE gels, containing a range of soluble starch concentrations, it is possible to measure the degree of retardation relative to a native PAGE gels without added starch. The relative mobility of a protein in this assay gives an estimate of the strength of interaction with the carbohydrate, and can be used to calculate its binding parameters. This method has been used previously in the characterisation of starch binding enzymes (Blennow *et al.*, 1998; Mikkelsen *et al.*, 2006). Using a set of native PAGE gels, containing a range of soluble starch concentrations, it is possible to measure the degree of retardation relative to native PAGE gels without added starch. Previous analysis of the recombinant CBMs showed that addition of glycerol affected the sedimentation assays, and so this was omitted from the native protein loading dye used. In addition markers were run within the gels, as guides to ensure accuracy in comparison across the native PAGE gels (see materials and methods).

Analysis of gel image data showed retardation of both GST-KIS(AtPTPKIS1) and GST-KIS(AtPTPKIS2) by inclusion of soluble starch or glycogen in the gels, with the degree of retardation increasing in a concentration dependent manner (Fig. 4.11 & 4.12). Starch (higher proportion of unbranched chains) was more effective than glycogen (higher proportion of branched chains) in retarding the CBMs.

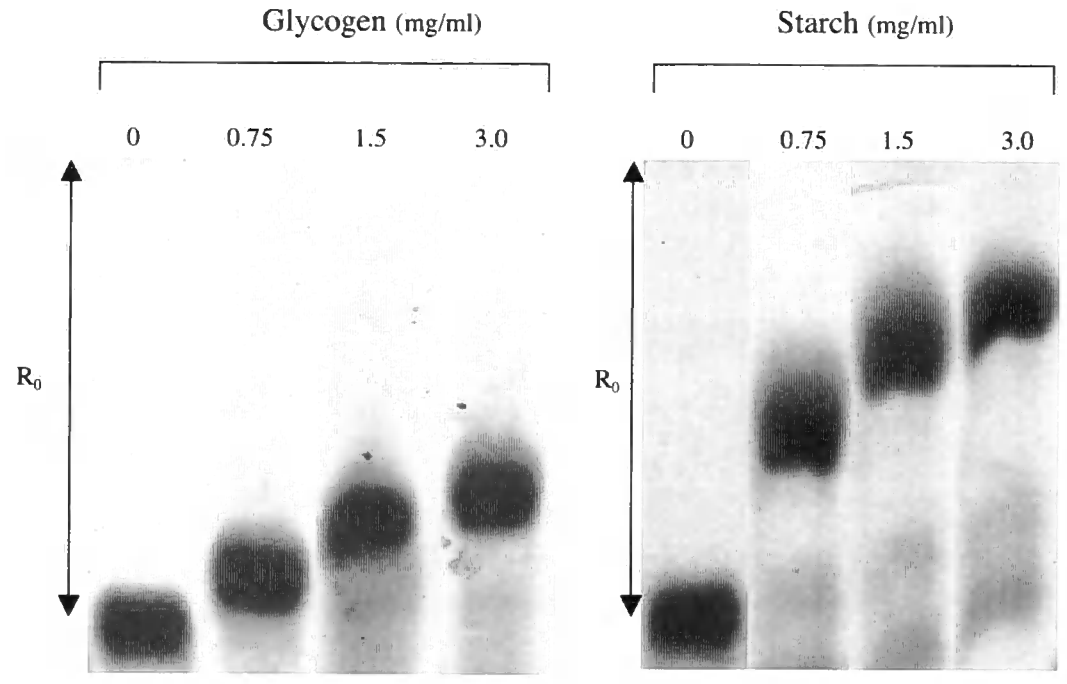


Figure 4.11 Mobility Gels of GST-KIS(AtPTPKIS1). Native gels containing variable concentrations of either soluble starch or glycogen.

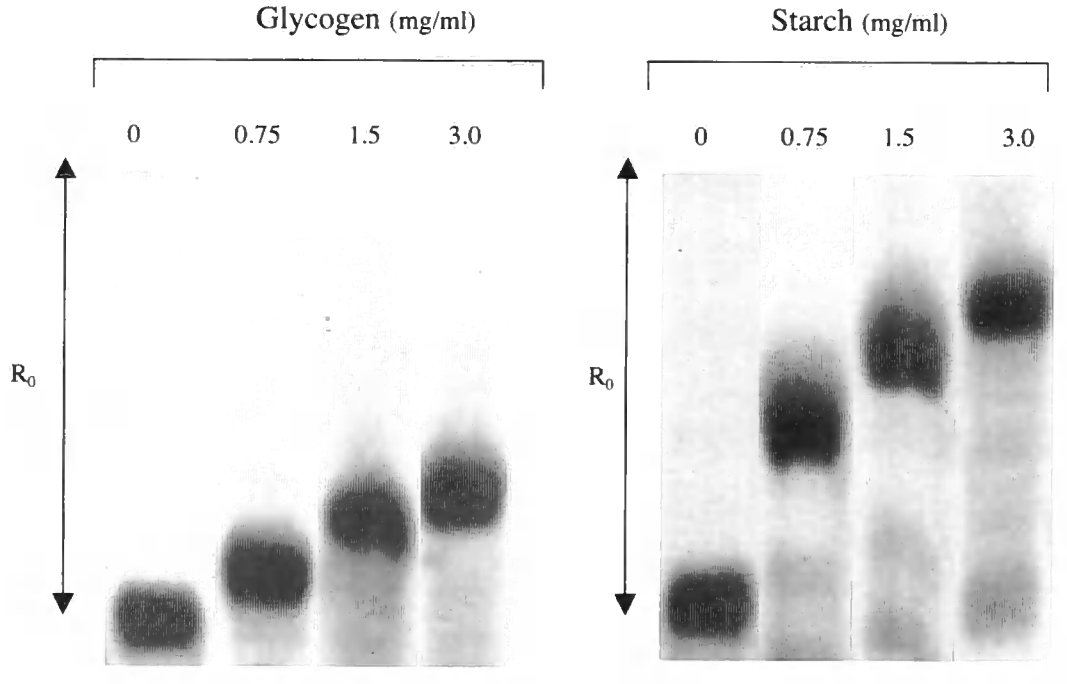


Figure 4.12 Mobility Gels of GST-KIS(AtPTPKIS2). Native gels containing variable concentrations of either soluble starch or glycogen.

For each lane the relative mobility (R_m) of the protein was calculated for the given substrate concentration (S) and the relative mobility of protein (R_0) in gels without substrate (S) determined. Using these data, the K_d was calculated from $1/R_m$ versus [S] plots as described in previous publications (Blennow *et al.*, 1998; Mikkelsen *et al.*, 2006) using the expression:

$$\frac{1}{R_m} = \frac{1}{R_0}\left(1 + \frac{[S]}{K_d}\right)$$

This gave a K_d of 1.37mg/ml and 1.43mg/ml for GST-KIS (AtPTPKIS1) and GST-KIS (AtPTPKIS2) respectively for soluble starch (Table 4.1). When glycogen was used as a substrate this gave a K_d of 9.0mg/ml and 9.9mg/ml for GST-KIS (AtPTPKIS1) and GST-KIS (AtPTPKIS2) respectively. While this interaction with soluble starch is a weaker interaction than other CBMs (Penninga *et al.*, 1996; Paldi *et al.*, 2003) it does show similarity to that identified for GWD1 (Mikkelsen *et al.*, 2006).

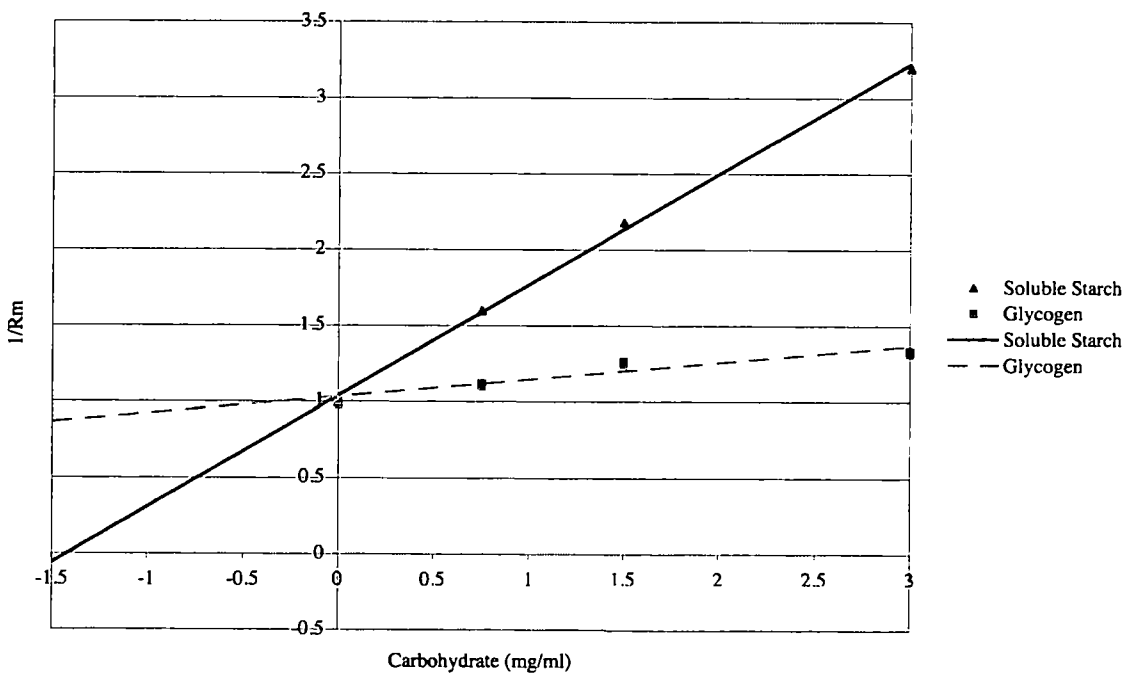


Figure 4.13 Relative Mobility of GST-KIS(AtPTPKIS1). Graphical representation of relative mobility of GST-KIS(AtPTPKIS1) against variable carbohydrate concentrations.

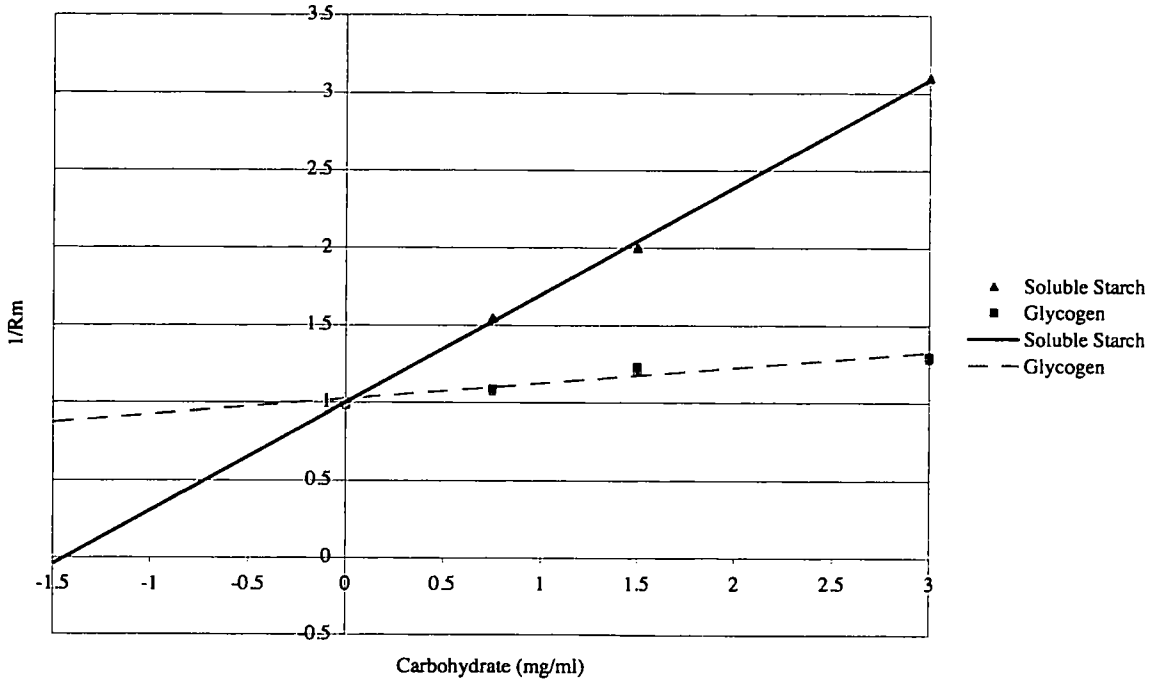


Figure 4.14 Relative Mobility of GST-KIS(AtPTPKIS1). Graphical representation of relative mobility of GST-KIS(AtPTPKIS1) against variable carbohydrate concentrations.

Substrate		GST-KIS (AtPTPKIS1)	GST-KIS (AtPTPKIS2)
Granular Starch	K_d	4.88 (mg/ml)	5.41 (mg/ml)
	B^{max}	0.53	0.43
Soluble Starch	K_d	1.37 (mg/ml)	1.43 (mg/ml)
Glycogen	K_d	9.0 (mg/ml)	9.9 (mg/ml)

Table 4.1 Binding Parameters. GST-KIS(AtPTPKIS1) and GST-KIS(AtPTPKIS2) binding parameters to granular starch, soluble starch and glycogen.

4.4 IDENTIFICATION OF KEY RESIDUES CONSISTENT WITH THE CLASSIC MODEL OF CBM's

In order to investigate the roles of the conserved tryptophan residues within the predicted KIS/CBM, consistent with other similar sequences shown in figure 4.1, a mutational analysis was carried out. Using the GST-KIS (AtPTPKIS1) plasmid as a basis, the following mutants were constructed: W278L, W293L, W313L and W293A (Fig. 4.15). The mutation of W293 was carried out with two different amino acids due to concerns over the low levels of soluble protein from the W293L construct. In both cases the constructs produced soluble protein, which performed in the same manner during analysis.

In addition to the mutation of the conserved tryptophan residues, a conserved lysine residue, homologous to one identified in AMPKβ as being involved in carbohydrate binding (Polekhina *et al.*, 2003), and conserved through the identified PTPKIS1 and PTPKIS2 sequences studied, was carried out. Using the GST-KIS(AtPTPKIS1) plasmid as a basis the mutant K307Q was constructed (Fig. 4.15).

Specific mutations in the predicted KIS/CBM region were generated using the Stratagene "Quick Change" mutagenesis kit (see Materials and Methods chapter), using primers shown in Table 4.2. As with the previous GST-KIS fusions, the resulting recombinant proteins were purified from cell supernatants of bacterial clones by affinity absorption on glutathione-Sepharose columns (Amersham), followed by washing and elution with Tris-HCl buffer at pH 8.0 (see Materials and Methods chapter).

Mutation	Primer Name	Primer Sequence
K307Q	K307Q Fd	5′ - GGA CAG TTT GAA TAT **c**AA TAC ATC ATA GAT GGT G - 3′
	K307Q Rv	5′ - C ACC ATC TAT GAT GTA TT**g** ATA TTC AAA CTG TCC - 3′
W278L	W278L Fd	5′ - CT GGC CTT GAC ATT GGA T**t**G GGA CAG AGG ATA CC - 3′
	W278L Rv	5′ - GG TAT CCT CTG TCC C**a**A TCC AAT GTC AAG GCC AG - 3′
W293L	W293L Fd	5′ - G GAC AAG GGA ACA GGA TTC T**t**G ATC CTA AAG AGA G - 3′
	W293L Rv	5′ - C TCT CTT TAG GAT C**a**A GAA TCC TGT TCC CTT GTC C - 3′
W314L	W314L Fv	5′ - C ATC ATA GAT GGT GAA T**t**G ACA CAC AAT GAG GCC G - 3′
	W314L Rv	5′ - C GGC CTC ATT GTG TGT C**a**A TTC ACC ATC TAT GAT G - 3′
W293A	W293A Fd	5′ - G GAC AAG GGA ACA GGA TTC **gc**G ATC CTA AAG AGA G - 3′
	W293A Rv	5′ - C TCT CTT TAG GAT C**gc** GAA TCC TGT TCC CTT GTC C - 3′

Table 4.2 Site Directed Mutagenesis Primers 1. Table of primers to be used for site directed mutagenesis of key residues in the GST-KIS(AtPTPKIS1) recombinant protein. Nucleic acids, which are to be mutated, are shown in lower care and highlighted in bold.

GAA GGT CGT GGG ATC CCC GAA TTC AGG AAG ACT GTT ACT CTG ACA
E G R G I P E F R K T V T L T

CTG AAA GAT AAG GGG TTC TCC AGA GTA GAA ATT TCT GGC CTT GAC
L K D K G F S R V E I S G L D

ATT GGA TGG GGA CAG AGG ATA CCT CTA ACA CTG GAC AAG GGA ACA
I G W G Q R I P L T L D K G T
W278

GGA TTC TGG ATC CTA AAG AGA GAA CTG CCT GAA GGA CAG TTT GAA
G F W I L K R E L P E G Q F E
W293

TAT AAA TAC ATC ATA GAT GGT GAA TGG ACA CAC AAT GAG GCC GAA
Y K Y I I D G E W T H N E A E
K307 **W314**

CCG TTT ATA GGA CCT AAC AAA GAC GGC CAT ACC AAC AAT TAC GCT
P F I G P N K D G H T N N Y A

AAA GTA GTG GAC TGA *CTC GAG CGG CCG CAT CGT GAC TGA*
K V V D . *L E R P H R D .*

Figure 4.15 GST-KIS(AtPTPKIS1) Sequence. The coding sequence from the KIS/CBM of region of AtPTPKIS1 in the pGEX5x1. Regions from the pGEX5x1 vector are show in italics. Residues that are mutated are highlighted in bold, numbers shown are the residues position if the native AtPTPKIS1 protein.

In order to determine if GST-KIS (AtPTPKIS1) mutant proteins showed binding to carbohydrate, pull down assays using granular potato starch (Sigma) were carried out, using the same methodology used to identify the carbohydrate binding of the predicted KIS/CBMs (as described in materials and methods).

The co-sedimentation assay showed that besides the positive control GST-KIS(AtPTPKIS1) only the K307Q mutant gave a band indicating that protein was present in the 10,000g pellets when incubated in the presence of granular potato starch. The inability to identify any bands corresponding to W278L, W293L, W313L and W293A in the pellet with and without starch suggests co-sedimentation, and thus binding to starch, requires the involvement of these tryptophan residues (Fig. 4.18). Binding of the K307 protein to granular potato starch (Fig. 4.16) was significantly weaker than for the positive control, showing that this mutation had decreased the binding ability of the CBM.

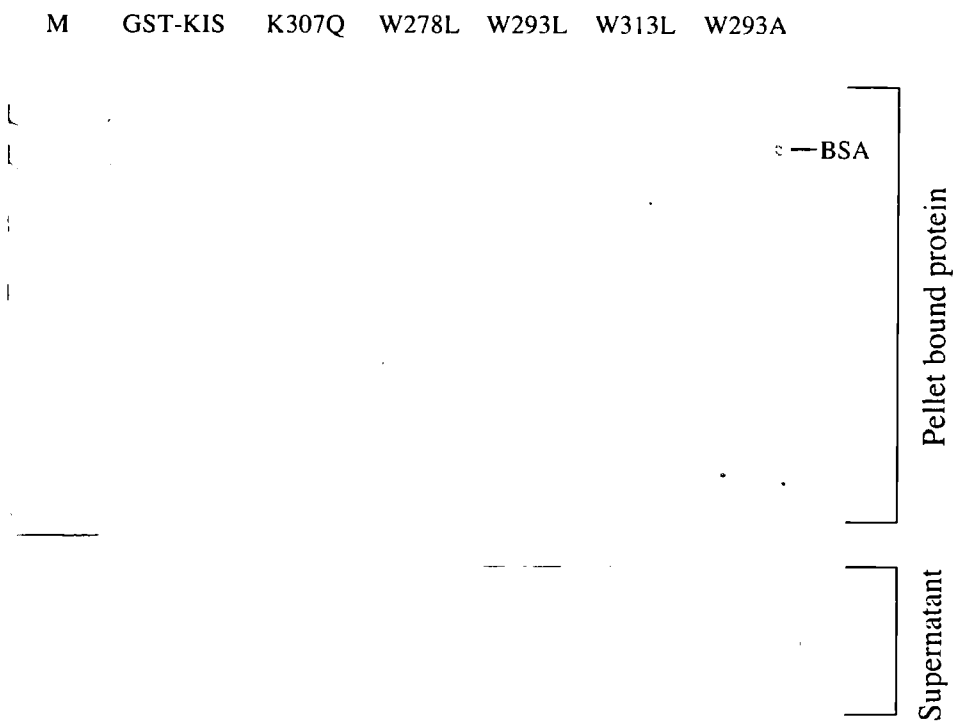


Figure 4.16. Starch pull down assay with GST-KIS(AtPTPKIS1)Mutants. SDS-PAGE gel, stained with Coomassie Brilliant Blue G250, showing protein content of pellets and the supernatant, from granular starch binding assay. Lane M is standard protein size markers.

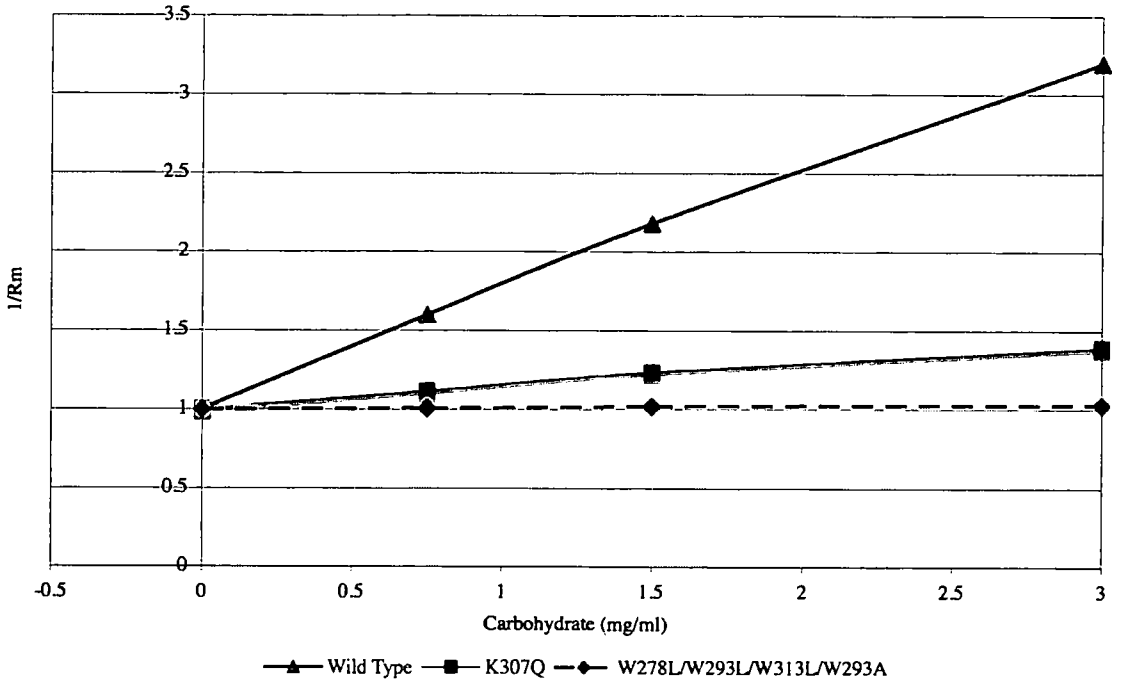


Figure 4.17 Relative Mobility of GST-KIS(AtPTPKIS1) Mutants. Graphical representation of relative mobility of GST-KIS(AtPTPKIS1) and the K307Q, W278L, W293L, W313L, and W293A mutants against variable carbohydrate concentrations.

A retardation assay using native PAGE in gels containing soluble starch was used to confirm the results of the cosedimentation assay. No retardation was observed for the W278L, W293L, W313L and W293A mutants (Fig. 4.17), showing an absence of starch binding. Both techniques show that residues W278, W293 and W313 are required for starch binding, and mutation in any of these leads to an inability to bind starch. The result of PAGE gel retardation assay with the K307Q mutant showed a slight alteration in mobility with starch concentration (Fig. 4.17), but not as pronounced as that seen with the "wild-type" protein, indicating weaker binding of the K307Q mutant to soluble starch, in agreement with the co-sedimentation assay. Both techniques show that the K307 residue plays a role in starch binding, but its role is less vital to the binding than the tryptophan residues.

4.5 NOVEL "SUGAR TONGS" IN THE BINDING SITE OF PTPKIS CBM

Study of the alignment of the PTPKIS CBMs, shows a conserved pair of tyrosine residues flanking the highly conserved lysine residue (K307 in AtPTPKIS1) in predicted PTPKIS1 and PTPKIS2 proteins in plants with the exception of the predicted *O.sativa* homologue, which only contains one of the tyrosine residues flanking the conserved lysine residue (Fig. 4.18). Analysis of proteins involved in carbohydrate degradation have shown regions away from the active site, containing such residues, termed sugar tongs (Robert *et al.*, 2003). While these 'sugar tongs' are usually found away from the catalytic site of the enzymes they have been shown to increase the enzyme's activity towards its substrate (Bozonnet *et al.*, 2007), with mutations in this region causing a reduction in enzymatic activity. While many residues have been shown to be conserved between PTPKIS family proteins and other carbohydrate binding proteins (Fig. 4.1), the pair of 'sugar tongs' seems only to be common within the PTPKIS family (Fig. 4.19)

```
StPTPKIS1        287 KDLHEGKYEYKYIVDGEWICNEFEPITSPNKDGHVNNY 324
LePTPKIS1        287 KDLHEGKYEYKYIVDGEWICNEFEPITSPNKDGHVNNY 324
MtPTPKIS1        303 KEMFEGRYEYKYIVDGEWTCNNDELVTSPNKDGHVNNF 340
AtPTPKIS1        295 RELPEGQFEYKYIIDGEWTHNEAEPFIGPNKDGHTNNY 332
OsPTPKIS1        287 RELPEGRYEYKYIVDGKWVCNDNEKKTKANADGHVNNY 324
PtPTPKIS1        296 RELPEGRYEYKYIVDGEWTCNKYELVTYPNKDGHVNNY 333
CsPTPKIS1        296 RELPEGRYEYKYIVDGEWTCNKYELVSSPNKDGHVNNY 333
AtPTPKIS2        505 VRLTQGKYYYKYIINGDWRHSATSP-TERDDRGNTNNI 542
StPTPKIS2        497 VRLSQGKYLYKYIISGNWRHSTNSP-TERDERGNLNNV 534
OsPTPKIS2        497 IRLRHGKYYYKFIAGGQWRHSTSLP-TETDEHGNVNNV 534
```

Figure 4.18 Alignment of Predicted Sugar Tong Region in PTPKIS1 and PTPKIS2. Alignment of predicted CBMs using sequences generated from alignment of EST data (appendix1). Amino acid sequences were aligned with clustalw. Amino acids making up the predicted sugar tongs (Y306 & Y308 in AtPTPKIS1) and sandwiched lysine residue (K307 in AtPTPKIS1) conserved throughout all sequences were highlighted in black, Additional tyrosine residues were highlighted in grey. *Solanum tuberosum* PTPKIS1 (StPTPKIS1), *Lycopersicon esculentum* PTPKIS1 (LePTPKIS1), *Medicago truncatula* PTPKIS1 (MtPTPKIS1), *Arabidopsis thaliana* PTPKIS1 (AtPTPKIS1), *Oryza sativa* PTPKIS1 (OsPTPKIS1), *Poncirus trifoliata* PTPKIS1 (PtPTPKIS1), *Citrus sinensis* PTPKIS1 (CsPTPKIS1), *Arabidopsis thaliana* PTPKIS2 (AtPTPKIS2), *Solanum tuberosum* PTPKIS2 (StPTPKIS2), *Lycopersicon esculentum* PTPKIS2 (LePTPKIS2). Numbers refer to the position of the first and last amino acid in the native protein.

```
StPTPKIS1      287  KDLHEGKYEYKYIVDGEWICNE 309
LePTPKIS1      287  KDLHEGKYEYKYIVDGEWICNE 309
AtPTPKIS1      295  RELPEGQFEYKYIIDGEWTHNE 317
AtPTPKIS2      505  VRLTQGKYYYKYIINGDWRHSA 527
AKINβγ1(Zm)     67  CSLSPGIHEYKFYVDGEWRHDE  89
AKINβγ2(Zm)     66  CSLSPGIHEYKFFVDGEWRHDE  89
AMPKβ1(Hs)     116  LDLPEGEHQYKFFVDGQWTHDP 138
AMPKβ2(Hs)     116  LDLPEGEHQYKFFVDGQWVHDP 138
AMPKβ1(Rn)     116  LDLPEGEHQYKFFVDGQWTHDP 138
GAL83(Sc)      203  LQLPPGTHRFRFIVDNELRFSD 225
```

Figure 4.19 Alignment of Predicted Sugar Tong Region in KIS domains. Alignment of predicted CBMs using sequences generated from alignment of EST data (appendix1). Amino acid sequences were aligned with clustalw. Amino acids making up the predicted sugar tongs (Y306 & Y308 in AtPTPKIS1) and sandwiched lysine residue (K307 in AtPTPKIS1) conserved throughout sequences were highlighted in black, Additional tyrosine residues were highlighted in grey. *Solanum tuberosum* PTPKIS1 (StPTPKIS1), *Lycopersicon esculentum* PTPKIS1 (LePTPKIS1), *Arabidopsis thaliana* PTPKIS1 (AtPTPKIS1), *Arabidopsis thaliana* PTPKIS2 (AtPTPKIS2), Zea mays AKINβγ1 (AKINbg1(Zm), accession number NP_001105005), Zea mays AKINβγ2 (AKINbg2(Zm), accession number NP_001105555), Human AMPKβ1 (AMPKb1(Hs), accession number Q9Y478), Human AMPKβ2 (AMPKb2(Hs), accession number NP_005390), Rat AMPKβ1 (AMPKb1(Rn), accession number P80386), S. cerevisiae GAL83 (Gal83(Sc), accession number Q04739). Numbers refer to the position of the first and last amino acid in the native protein.

In order to investigate the possibility that these conserved residues may play an important role in carbohydrate association, mutant protein was expressed, in which the predicted sugar tongs of the AtPTPKIS1 CBM were mutated. Using the GST-KIS (AtPTPKIS1) construct generated in Fordham-Skelton *et al.*, (2002), specific mutations in the predicted sugar tong region were generated using the Stratagene "Quick Change" mutagenesis kit (see Materials and Methods chapter), using primers shown in Table 4.3. As with the previous GST-KIS fusions, this was purified from cell supernatant by affinity absorption on glutathione-Sepharose columns (Amersham), followed by washing and elution with Tris-HCl buffer at pH 8.0 (see Materials and Methods chapter).

Mutation	Primer Name	Primer Sequence
Y306F-Y308F	Y306F-Y308F Fd	5′ - CCT GAA GGA CAG TTT GAA T**t**T AAA T**t**C ATC ATA GAT GGT G - 3′
	Y306F-Y308F Rv	5′ - C ACC ATC TAT GAT G**a**A TTT A**a**A TTC AAA CTG TCC TTC AGG - 3′
Y306F	Y306F Fd	5′ - CCT GAA GGA CAG TTT GAA T**t**T AAA TAC ATC ATA GAT GGT G - 3′
	Y306F Rv	5′ - C ACC ATC TAT GAT GTA TTT A**a**A TTC AAA CTG TCC TTC AGG - 3′
Y308F	Y308F Fd	5′ - CCT GAA GGA CAG TTT GAA TAT AAA T**t**C ATC ATA GAT GGT G - 3′
	Y308F Rv	5′ - C ACC ATC TAT GAT G**a**A TTT ATA TTC AAA CTG TCC TTC AGG - 3′
Y306Q-Y308Q	Y306Q-Y308Q Fd	5′ - CCT GAA GGA CAG TTT GAA **cAa** AAA **cAa** ATC ATA GAT GGT G - 3′
	Y306Q-Y308Q Rv	5′ - C ACC ATC TAT GAT **tTg** TTT **tTg** TTC AAA CTG TCC TTC AGG - 3′

Table 4.3 Site Directed Mutagenesis Primers 2. Table of nucleotide primers used for site directed mutagenesis of key residues in the predicted "sugar tong" region of GST-KIS(AtPTPKIS1). Nucleic acids, which are different from the wild type form, are shown in lower care and highlighted in bold.

As with the other mutants studied previously, pull down assays using granular potato starch (Sigma) were carried out to identify the carbohydrate binding of the predicted KIS/CBMs (as described in materials and methods) using GST-KIS(AtPTPKIS1) and mutant protein as generated above.

The co-sedimentation assay showed that the positive control GST-KIS(AtPTPKIS1) was present in the 10,000g pellets when incubated with potato starch, as were, to a much smaller level, the Y306F and Y308F mutants (Fig. 4.20). Bands corresponding to Y306F-Y308F and Y306Q-Y308Q mutants were not observed in the pellets both with and without starch. These data suggest that the Y306F and Y308F proteins are

able to bind to granular potato starch, but binding is weaker than for the "wild type" sequence or for the K307Q mutant. The double mutants do not co-sediment with starch (Fig. 4.20). These data suggest that these 'sugar tong' residues are important in binding to starch. Disruption of one residue causes a marked decrease in binding affinity, and disruption of both causes a loss of binding ability.

As with the previous analysis of predicted KIS/CBMs, PAGE retardation assays were carried out using a range of starch concentrations. As was expected from the sedimentation analysis, the Y306F-Y308F and Y306Q-Y308Q mutants showed no retardation by starch. The Y306F and Y308F showed very minimal retardation, indicative of weak interaction with starch, in agreement with the results of the co-sedimentation assay (Fig 4.21).

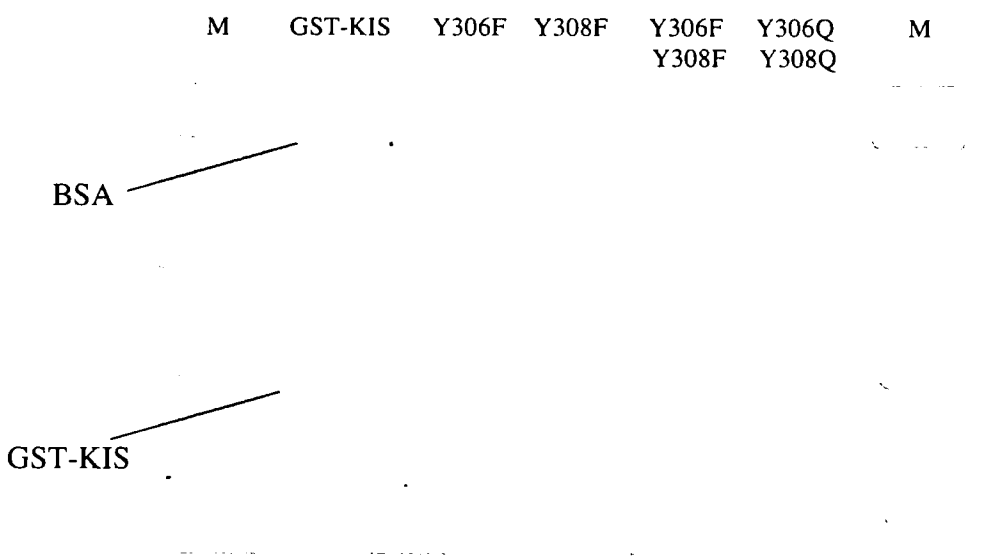


Figure 4.20 Starch pull down assay with GST-KIS(AtPTPKIS1) Sugar Tong Mutants. SDS-PAGE gel, stained with Coomassie Brilliant Blue G250, showing protein content of pellets, from granular starch binding assay. Lane M is standard protein size markers.

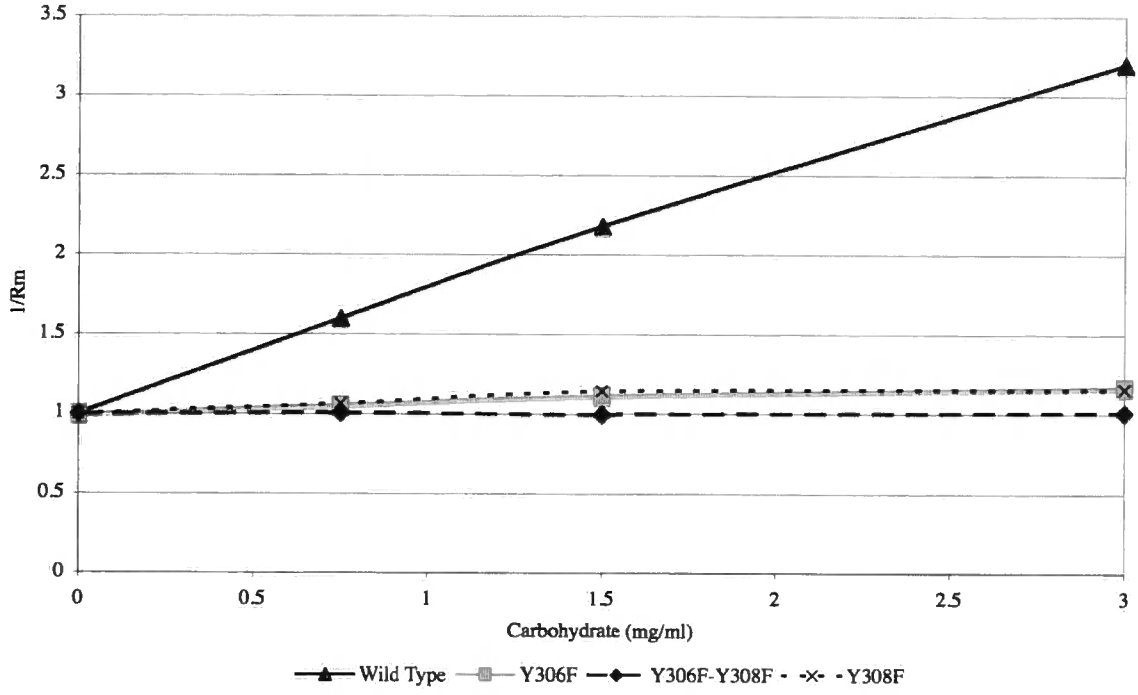


Figure 4.21 Relative Mobility of GST-KIS(AtPTPKIS1) Sugar Tong Mutants. Graphical representation of relative mobility of GST-KIS(AtPTPKIS1) and the Y306F, Y308F, and Y306F-Y308F mutants, against variable carbohydrate concentrations.

4.6 DISCUSSION

Work described here identifies the previously predicted KIS domains of AtPTPKIS1 and AtPTPKIS2 as being CBMs, while showing that not all KIS domains can act as CBMs. Analysis of the identified PTPKIS1 and PTPKIS2 sequences shows the existence of C-terminal sequences, overlapping previously predicted KIS domains, with significant similarity to CBMs of family 20 and 21, and that identified in AMPKβ.

Starch binding properties of the identified CBMs were demonstrated using recombinant protein expressed and purified from *E. coli*. Experiments using granular potato starch showed a clear binding of the predicted KIS/CBM, in addition to the clear interaction with soluble starch as observed using native PAGE experiments. Inhibition of this binding was shown to occur using β-cyclodextrin as is common in members of the CBM 20 and 21 family (Paldi *et al.*, 2003; Polekhina *et al.*, 2003). While the binding kinetics are comparable to previously studied CBMs involved in starch metabolism, the kinetics of the CBMs in both AtPTPKIS1 and AtPTPKIS2 show a great deal of similarity to that of the previously characterised GWD1 (Mikkelsen *et al.*, 2006). Mutation in either of the PTPKIS genes, or the GWD1 gene results in a starch excess phenotype.

Aromatic residues of well characterised CBMs have been shown to play key roles in substrate binding (Svensson *et al.*, 1989; Williamson *et al.*, 1997; Mikkelsen *et al.*, 2006). This was seen in AMPKβ (Polekhina *et al.*, 2003) and substitution of W278, W293, and W313 in AtPTPKIS1, resulted in loss of starch binding function. This identifies these tryptophan residues as being actively involved in carbohydrate recognition. The requirement for all these residues to be present, in combination with the calculated Hill factor of almost 1, supports the assumption that this CBM contains only one binding site, employing a binding mechanism similar to that of other CBMs, with tryptophan residues binding to the glucose residues via stacking interactions (Simpson *et al.*, 2000; Polekhina *et al.*, 2003; Boraston *et al.*, 2004).

Previous analysis of the phosphatase domain of PTPKIS1 and PTPKIS2 showed great similarity in sequence to Laforin, a protein containing both phosphatase and CBM domains, mutation of which results in formation of starch like lafora bodies in animals. Identification of a CBM in both AtPTPKIS1 and AtPTPKIS2 suggests a possible similarity in function between Laforin and the PTPKIS proteins.

Study of the alignment of the PTPKIS CBMs, shows a conserved pair of tyrosine residues flanking the highly conserved lysine residue (K307 in AtPTPKIS1) in PTPKIS1 and PTPKIS2. Previous analysis of some enzyme's involved in carbohydrate metabolism has identified regions, away from catalytic active sites, containing tyrosine residues, which act to bind carbohydrates (Robert *et al.*, 2003; Bozonnet *et al.*, 2007). Mutation of these sugar tongs reduces the enzymes carbohydrate binding ability, while not affecting the active site. Substitution of Y306 and Y308 in AtPTPKIS1, resulted in loss of starch binding function, while substitution of a single residue (either Y306 or Y308) resulted in a reduction in binding ability, greater than substitution on the conserved lysine residue (K307) between them. This identifies these tyrosine residues as being actively involved in carbohydrate recognition. This suggests a possible role in effecting the specificity of the CBMs of the PTPKIS proteins, giving them a preference to binding regions of some carbohydrates more readily than those CBMs without the identified 'sugar tongs'.

The predicted OsPTPKIS2 sequence, confirmed through analysis of both genomic and EST sequences lacks one of the conserved lysine flanking tyrosine residues (region 306-308 in AtPTPKIS1). While this may alter the enzymes affinity to carbohydrate, this may be counteracted by the increase other tyrosine residues local to the predicted 'sugar tong' region.

In this chapter it has been shown that the KIS domains of AtPTPKIS1 and AtPTPKIS2 act as CBMs in a physiologically relevant manner, while showing that carbohydrate binding is not common to all KIS domains. The presence of this CBM and the previously characterised phosphatase activity of the enzymes, suggests a comparable role with that of Laforin.

Chapter 5

ANALYSIS OF SEX MUTANTS: CARBOHYDRATES AND EXPRESSION OF SELECTED GENES

5.1 INTRODUCTION

The Arabidopsis SEX4 phenotype, characterised by accumulation of higher than normal amounts of starch in leaves, was originally identified by Zeeman *et al.* (1998). These authors suggested that the phenotype was the result of mutation in a gene encoding a "chloroplastic starch-hydrolysing enzyme", resulting in an inability to break down the carbohydrate end-product of photosynthesis in chloroplasts, and thus accumulation of excess starch. Analysis of starch levels showed that the increase in leaf starch changed with aging, the variation between starch levels of wild type and SEX4 plants being smaller in younger leaves than older tissues (Zeeman and Rees, 1999). More recent publications have shown that accumulation of excess starch in Arabidopsis lines showing the SEX4-1 phenotype is due to deletion of the N-terminal region of gene At3g52180 (Niittyla *et al.*, 2006). The polypeptide encoded by At3g52180, designated AtPTPKIS1, was originally identified by Fordham-Skelton *et al.*, (2002) as a novel plant protein tyrosine phosphatase, whose sequence shows no evidence for glycohydrolase activity, in terms of sequence similarity to known glycohydrolases, or presence of sequence motifs typical of glycohydrolases. The SEX4-1 phenotype cannot therefore be a direct consequence of failure to hydrolyse starch. However, it could result from an effect of the activity of AtPTPKIS1, possibly as part of a metabolic regulatory system. In addition to the SEX4-1 line, further Arabidopsis lines have been identified in which mutations in the At3g52180 gene cause a starch excess phenotype (Fig 5.1) (Niittyla *et al.*, 2006). Analysis of metabolite levels in SEX4 lines, and other knockouts, can throw light on the effects of the PTPKIS1, and similar enzymes, on metabolic regulation.

As determined in the previous chapters and in a number of recent publications (Kerk *et al.*, 2006; Niittyla *et al.*, 2006; Sokolov *et al.*, 2006; Gentry *et al.*, 2007), AtPTPKIS1 (and homologues in other species) has the ability to bind to carbohydrates through a region of its C-terminus showing similarity to characterised carbohydrate binding modules, which can thus be identified as a carbohydrate-binding domain. These results provide a basis for models which link the activity of PTPKIS1 to the regulation of starch metabolism, through an association of the enzyme with starch granules, but leave open the problem of whether the regulatory effect of the enzyme is through its action on phosphorylated proteins associated with starch, or on phosphorylated sugar residues in starch itself. The N-terminal region of the protein, identified as the phosphatase domain, shows sequence similarity to protein phosphatases, suggesting that the substrate is phosphorylated protein(s), but recent results have provided evidence that phosphatase activity towards phosphoglucans can be detected, suggesting that the enzyme acts directly on starch. Analysis of metabolites in lines containing PTPKIS knockouts may provide useful evidence to distinguish between these possibilities.

As described in previous chapters, a second gene with sequence similarity to AtPTPKIS1, designated AtPTPKIS2 (At3g01510) is present in the Arabidopsis genome, and has been shown to increase AtPTPKIS1 phosphatase activity towards phosphoglucans. While containing a similar domain structure to AtPTPKIS1, with a predicted PTP domain and a C-terminal carbohydrate binding domain, AtPTPKIS2 showed no independent phosphatase activity. Analysis of a SALK insertion mutant in the At3g01510 gene (termed PTPKIS2-SALK) showed that a starch excess phenotype was present, similar to that of the At3g52180 insertion line SEX4-3 (Fig 5.1). The basis for this phenotype is as yet unclear, but analysis should show whether the similar overall phenotype results from similar changes to metabolism.

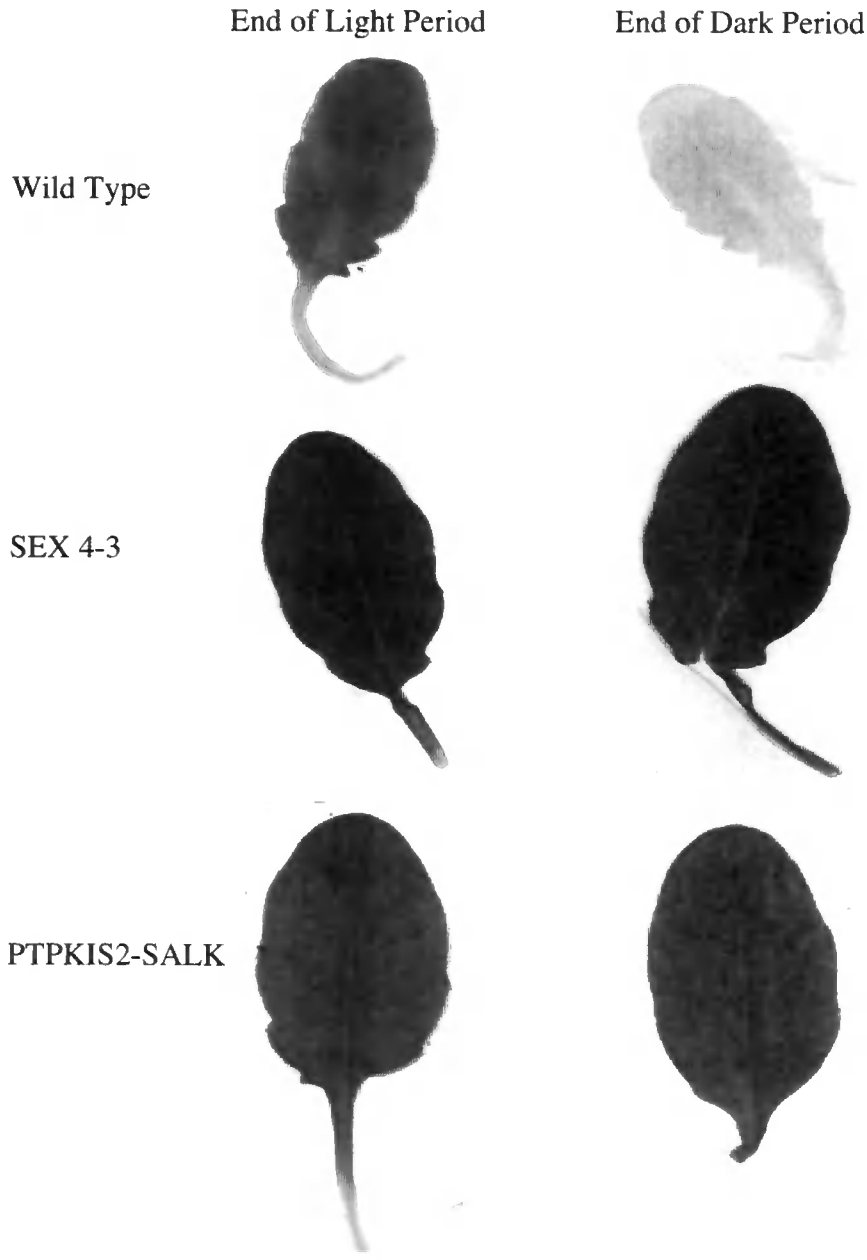


Figure 5.1 Iodine Stained Starch in Leaves of Mutant and Wild Type Plants. Images of leaves harvested at the end of the dark or light period, stained with lugols solution, which turns black in the presence of starch. Leaves were taken from mature plants, at the end of the light and dark periods, before being stained as described in materials and methods.

This chapter aims to characterise the change in glucose, fructose and sucrose levels caused by mutations in the AtPTPKIS1 and AtPTPKIS2 genes, in addition to the analysis of phosphate levels in the starch of these mutant lines. The levels of relevant transcripts were also studied, along with a study of the expression profiles of the previously studied members of the PTPKIS family.

5.2 METABOLITE ANALYSIS OF STARCH EXCESS LINES

5.2.1 Sugar Levels of AtPTPKIS1 Mutants SEX4-1 and SEX4-3 During the Diurnal Cycle

Initial published metabolomic data for the SEX4 mutation (AtPTPKIS1 knockout) was generated using the SEX4-1 mutant line. This line was generated by X-ray-mutagenesis (Zeeman *et al.*, 1998), which resulted in a deletion on the first 6 exons of At3g52180 as well as the 5' UTR and the region up stream of this. Further investigation has shown that this deletion extends into the upstream gene At3g52170 (Sylviana Comparot-Moss, Personal Communication), which shows homology to a DNA-binding protein PD2 from *Pisum sativum*. It is therefore possible that the phenotype observed could also result from disruption of gene At3g52170. In order to confirm the original results, a comparison of metabolites from SEX4-1 plants and SEX4-3 (SALK line SALK_102567) plants was carried out. The SEX4-3 line contains a T-DNA insertion within the 7th exon, resulting in disruption within the phosphatase domain of the protein (Fig 5.2).

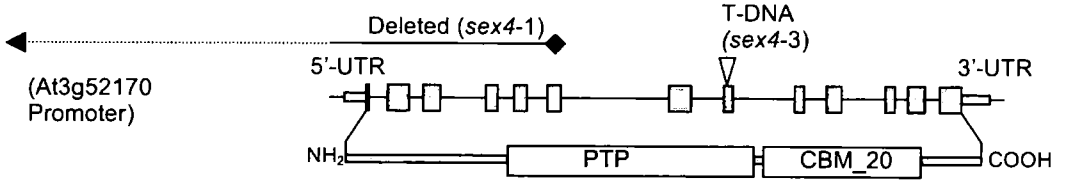


Figure 5.2 SEX4-1 and SEX4-3 mutations in At3g52180. Diagrammatic representation of the AtPTPKIS1 gene (At3g52180) and the location of the SEX4-1 and SEX4-3 disruptions. Large grey boxes represent the exons, with the 5' and 3' UTR labelled.

In order to determine if there were any significant changes in the glucose, fructose and sucrose levels, samples of leaf tissue were taken throughout the day/night period from plants grown in short day conditions (8h light, 16h dark). Sugars were extracted (as stated in the materials and methods) and their concentrations were determined through enzymatic assay (Fig 5.3 - 5.5).

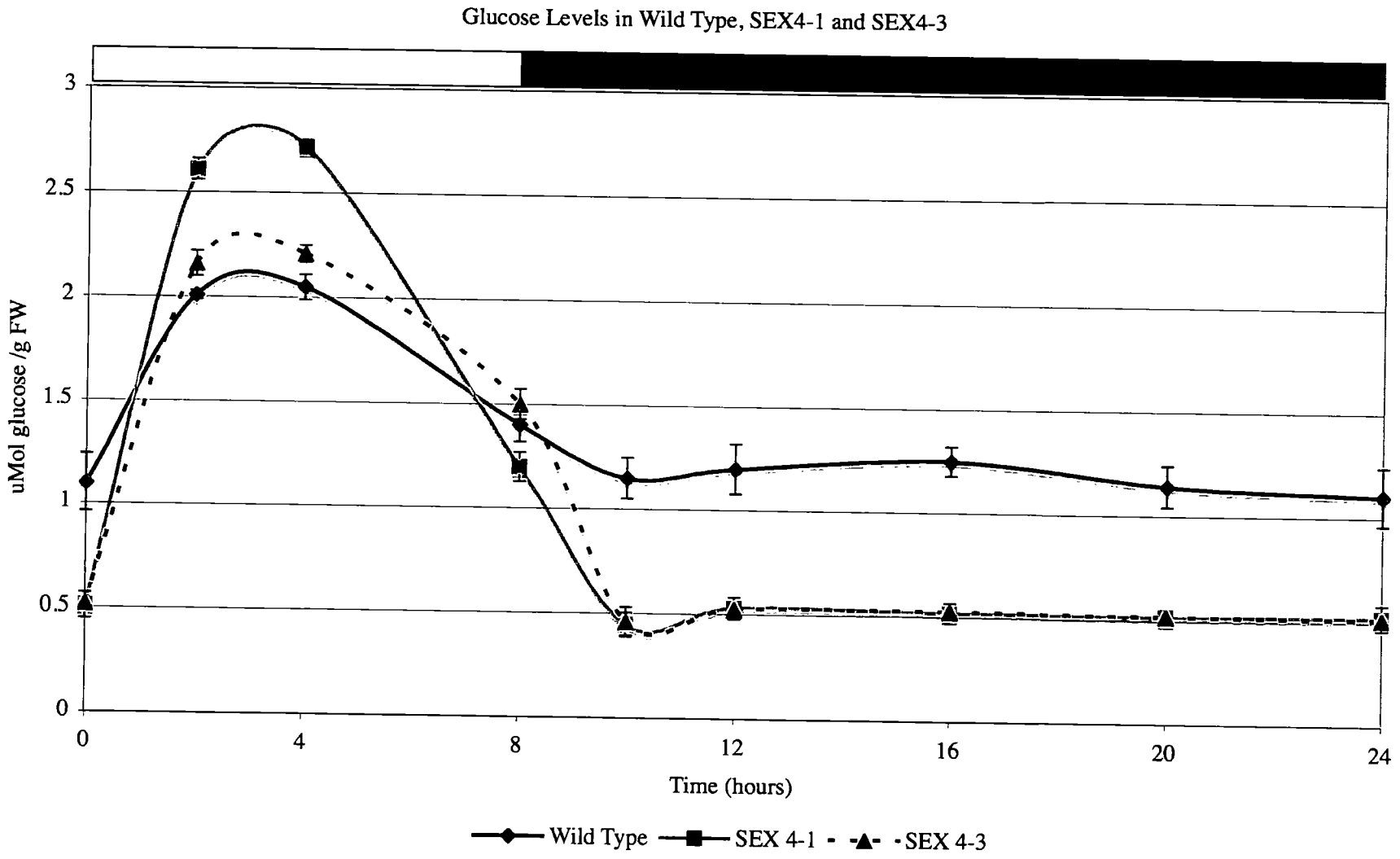


Figure 5.3 Glucose Levels in Wild Type, SEX4-1 and SEX4-3. Graph showing the levels of leaf glucose, isolated from mature leaves in Wild Type (◆), SEX4-1 (■) and SEX4-3 (▲) plants throughout a 24-hour period. Samples were taken though the 8 hours of light and 16 hours darkness. Error bars represent standard error at each time point.

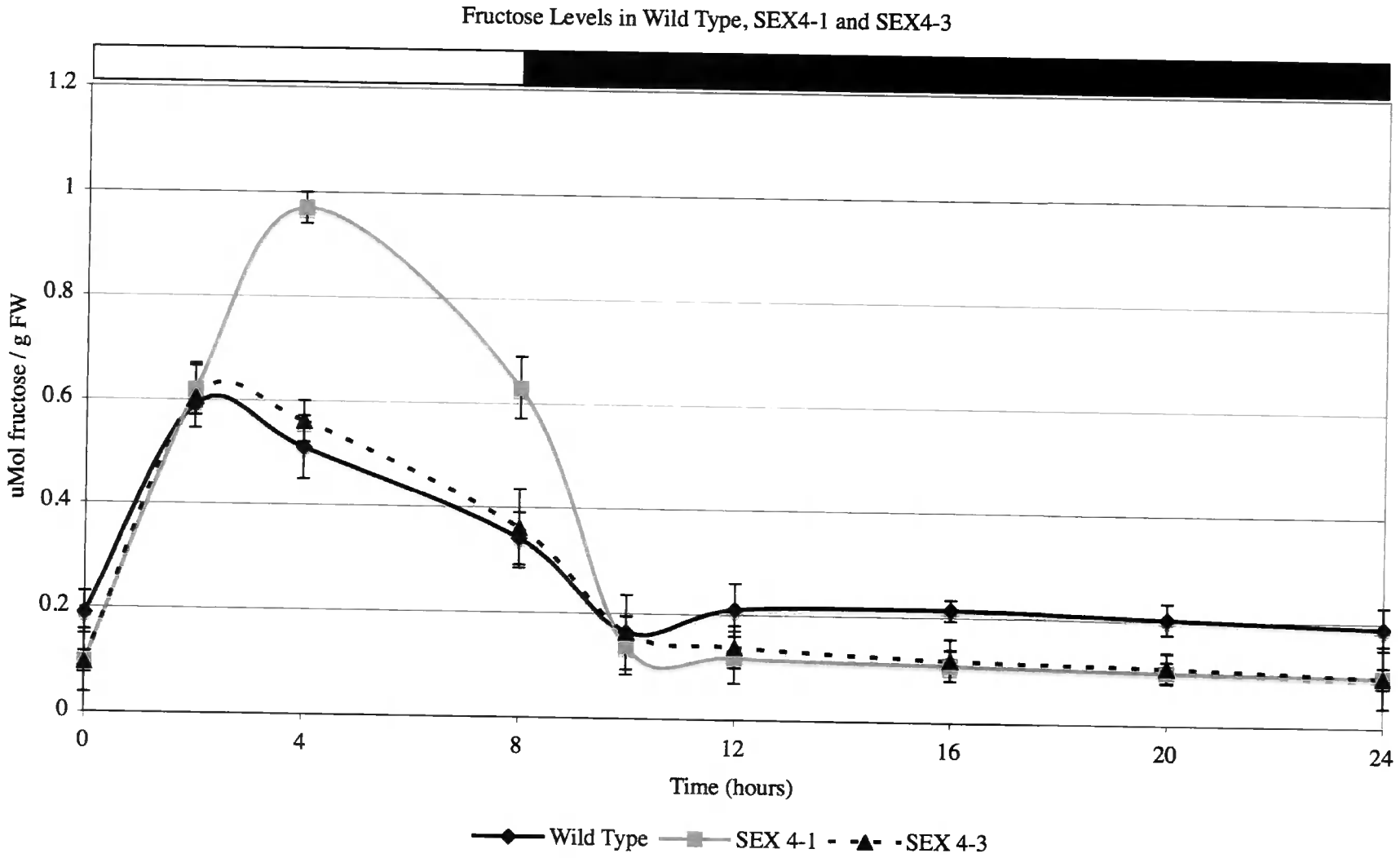


Figure 5.4 Fructose Levels in Wild Type, SEX4-1 and SEX4-3. Graph showing the levels of leaf fructose, isolated from mature leaves in Wild Type (◆), SEX4-1 (■) and SEX4-3 (▲) plants throughout a 24-hour period. Samples were taken though the 8 hours of light and 16 hours darkness. Error bars represent standard error at each time point.

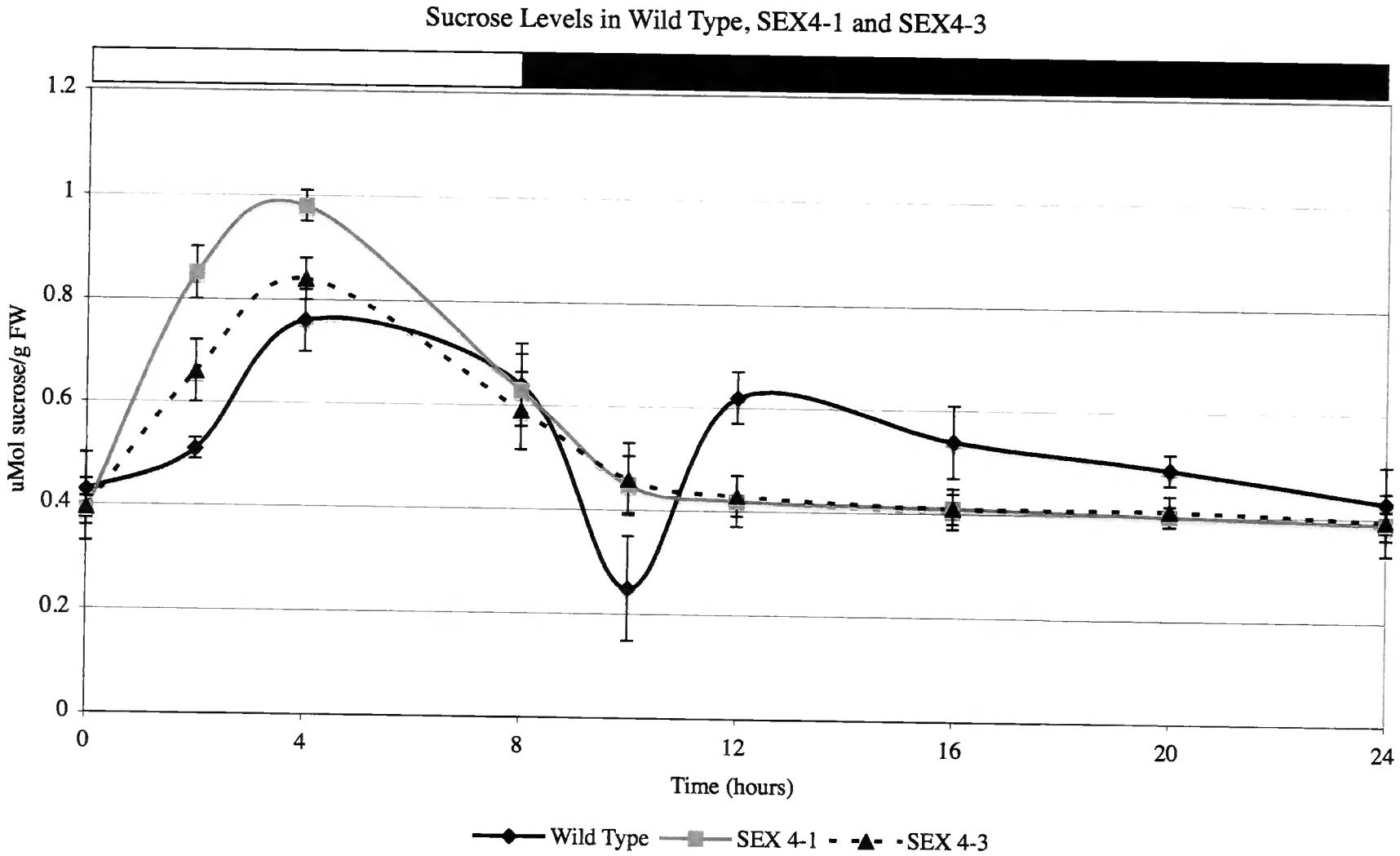


Figure 5.5 Sucrose Levels in Wild Type, SEX4-1 and SEX4-3. Graph showing the levels of leaf sucrose, isolated from mature leaves in Wild Type (◆), SEX4-1 (■) and SEX4-3 (▲) plants throughout a 24-hour period. Samples were taken though the 8 hours of light and 16 hours darkness. Error bars represent standard error at each time point.

During the light period, levels of sucrose in wild type, and both SEX4 mutants rise, with SEX4-1 showing a significantly higher level of sucrose in the middle of the light period than wild type and SEX4-3 plants. Upon the onset of darkness, sucrose levels drop in wild type plants, before rising again to levels similar to those before the onset of darkness,, and then declining during the rest of the dark period. This however is not mirrored in either of the SEX4 lines, which show a consistent decrease in sucrose levels during the latter part of the light period and the early part of the dark period, followed by a steady level during the rest of the dark period, with consistently lower sucrose levels than those in wild type plants (Fig. 5.5). If sucrose levels during the dark period are regarded as products of starch breakdown, these results are consistent with decreased hydrolysis in the SEX4 lines. Higher levels of sucrose during the light period in the mutant lines establish that sucrose synthesis is not impaired, supporting this conclusion.

Glucose levels of SEX4-1 during the light period are significantly higher than those of both wild type and SEX4-3 plants. This significantly higher glucose level through the day is consistent with the findings of Zeeman *et al*, (1999). Following the onset of darkness, the glucose levels in all plants decreases, with both SEX4-1 and SEX4-3 showing similarly low levels, significantly lower than those seen in wild type plants (Fig5.3). This is again consistent with a failure to break down starch in the SEX4 mutants, and the concomitant inability to generate glucose at the same rate as the wild type plants.

Fructose levels in all plants follow a similar pattern to that seen in glucose levels, with fructose being, in general, half as common as glucose. Of note is the fact that, as with glucose and sucrose, during the light period SEX4-3 is similar to wild type plants in its sugar levels, but with the onset of night, shows greater similarity to the sugar levels of SEX4-1. This may suggest that the mutation in SEX4-1 affecting the upstream gene At3g52170 may cause additional sucrose, glucose and fructose to be present in plant leaves during the light period. It is not clear whether this is due to mutation in At3g52170 and At3g52180 genes, or just the At3g52170 gene.

5.2.2 Comparison Of The Sugar Levels in AtPTPKIS1 Mutant SEX4-3 And The AtPTPKIS2 SALK Mutant

The SALK insertional mutant SALK_053285 causes interruption of the 5' UTR of gene At3g01510 (Fig. 5. 6), referred to as AtPTPKIS2 previously in this work. This mutant line shows a starch excess phenotype similar to that of SEX4 when leaves, harvested at the end of the dark period, are stained for starch content with iodine (Fig 5.1).

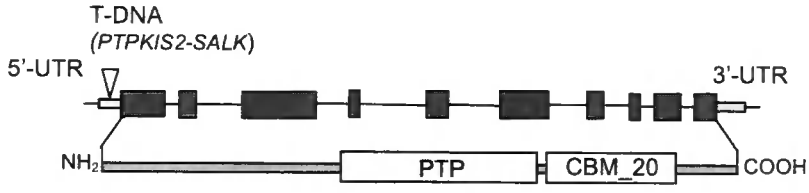


Figure 5.6 PTPKIS2-SALK mutation in At3g01510. Diagrammatic representation of the AtPTPKIS2 gene (At3g01510) and the location of the PTPKIS2-SALK insertion. Large grey boxes represent the exons, with the 5' and 3' UTR labelled.

In Chapter 3 it was shown that AtPTPKIS2 alone was incapable of showing any phosphatase activity, but was instead able to increase phosphatase activity of PTPKIS1 towards phosphoglucans. In order to further understand the role of AtPTPKIS2 it was important to determine if there were any significant changes in the glucose, fructose and sucrose levels between the SEX4-3 mutant line and the PTPKIS2-SALK mutant line. To do this samples of leaf tissue were taken throughout the day/night period from plants grown in short day conditions (8h light, 16h dark) and their levels of these sugars studied. Sugars were extracted (as stated in the materials and methods) and their concentrations were determined through enzymatic assay (Fig 5.7 – 5.9).

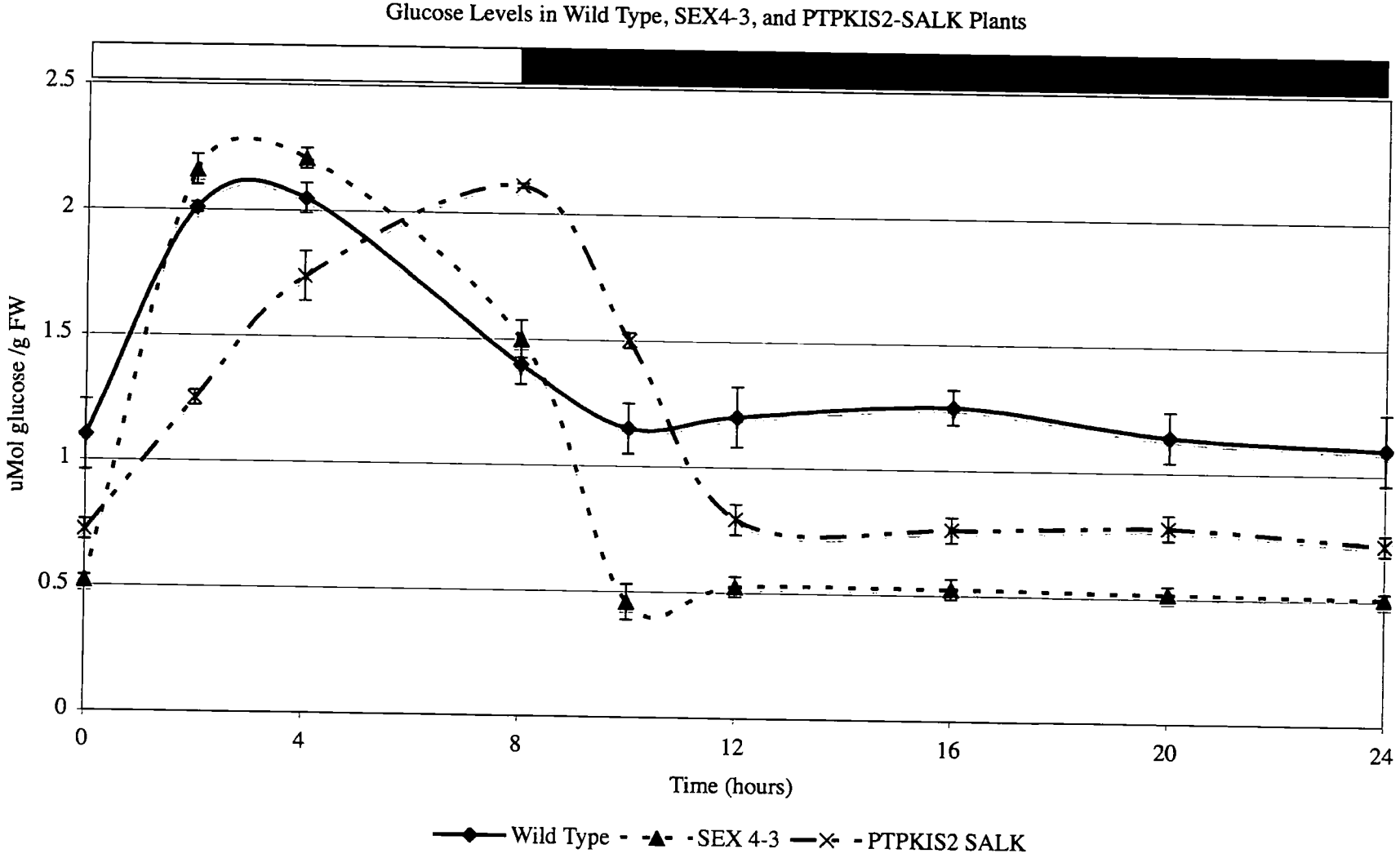


Figure 5.7 Glucose Levels in Wild Type, SEX4-3, and PTPKIS2-SALK Plants. Graph showing the levels of leaf fructose, isolated from mature leaves in Wild Type (◆), SEX4-3 (▲) and PTPKIS2-SALK (×) plants throughout a 24-hour period. Samples were taken though the 8 hours of light and 16 hours darkness. Error bars represent standard error at each time point.

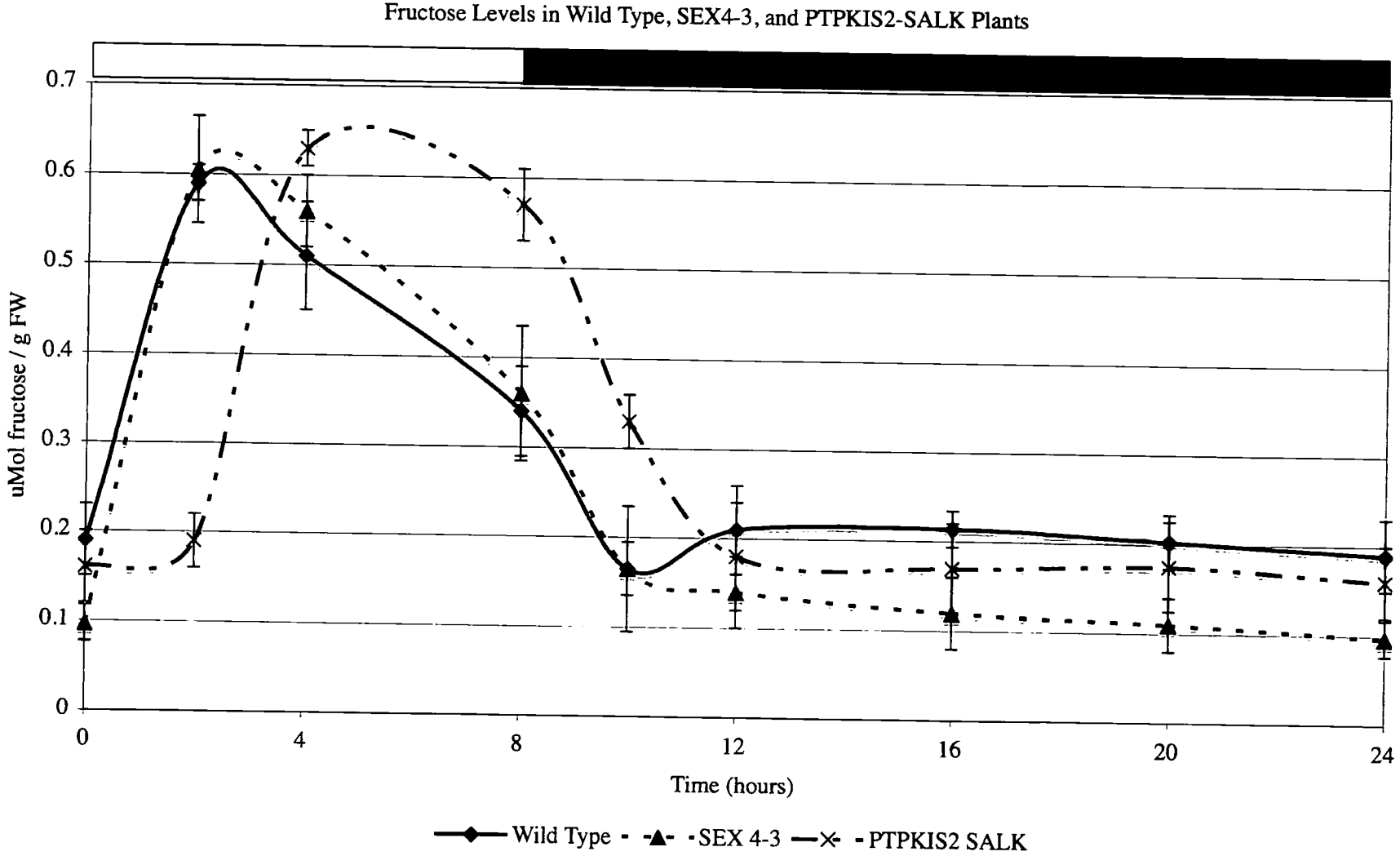


Figure 5.7 Fructose Levels in Wild Type, SEX4-3, and PTPKIS2-SALK Plants. Graph showing the levels of leaf fructose, isolated from mature leaves in Wild Type (◆), SEX4-3 (▲) and PTPKIS2-SALK (×) plants throughout a 24-hour period. Samples were taken though the 8 hours of light and 16 hours darkness. Error bars represent standard error at each time point.

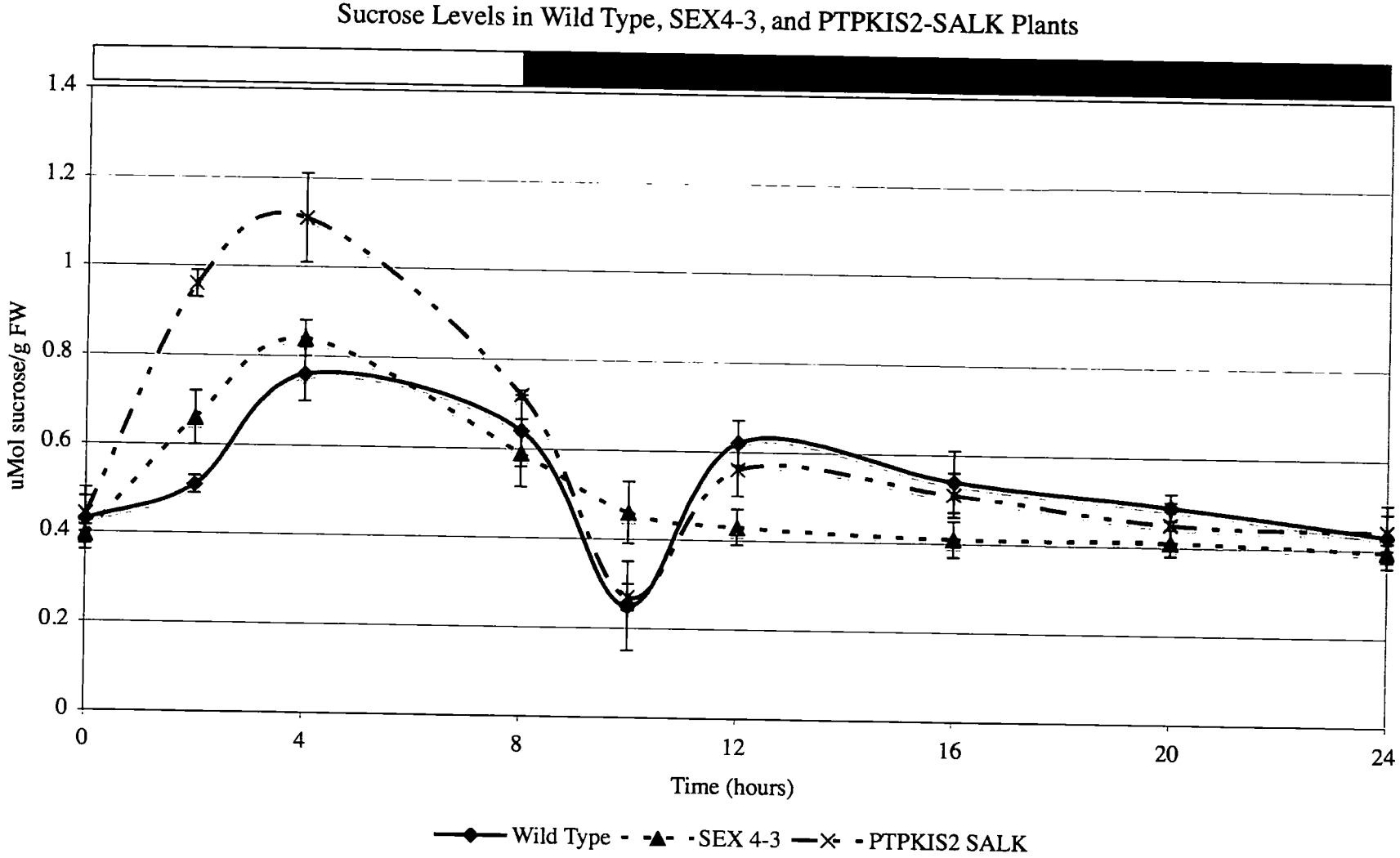


Figure 5.7 Sucrose Levels in Wild Type, SEX4-3, and PTPKIS2-SALK Plants. Graph showing the levels of leaf sucrose, isolated from mature leaves in Wild Type (◆), SEX4-3 (▲) and PTPKIS2-SALK (×) plants throughout a 24-hour period. Samples were taken though the 8 hours of light and 16 hours darkness. Error bars represent standard error at each time point.

These data show, that throughout the dark period, the PTPKIS2-SALK line shows significantly lower sucrose, glucose and fructose levels than wild type plants, but also significantly higher levels than found in SEX4-3 lines. While having similar peak levels of glucose and fructose during the daytime, hexose levels in PTPKIS2-SALK appear to peak 4 hours later compared to wild type and SEX4-3 (Fig 5.7 & 5.8). While the sucrose levels of PTPKIS2-SALK follow a similar pattern to wild type, they however show an increase in sucrose levels during the light period, being up to 30% higher in the middle of the period (Fig 5.9).

Comparison of sugar levels over the diurnal period in PTPKIS2-SALK and SEX4-3 show significant differences in both the amounts and the patterns with time, suggesting that the effect of knocking out either gene has a different effect on the plant. Furthermore it suggests that the enzymes PTPKIS1 and PTPKIS2 are not interchangeable, and fulfil unique roles within starch metabolism.

5.2.3 Glucose-6-Phosphate Levels In Wild Type And SEX4-3 Plant Lines

It has previously been shown that the levels of sucrose, fructose, and glucose in the SEX4-1 mutant line differ significantly from those of wild type plants. In order to determine if there were any significant changes in the hexose phosphate levels, samples of leaf tissue were taken throughout the day/night period and the levels of glucose-1-phosphate, glucose-6-phosphate, and fructose-6-phosphate were determined. Phosphorylated sugars were extracted in ethanol and water, and then subjected to HPAEC-PAD on a PA-1 column (Dionex, Rødovre, Denmark). During the extraction glucose-1, 6-bisphosphate was added as an internal standard.

Results showed that on average, through out the day/night cycle, SEX4-3 plants contained around 10 times the concentration of glucose-6-phosphate present in wild type plants. Whereas glucose-6-phosphate levels in wild type plants showed little evidence of variation over the day/night cycle, in SEX4-3 plants there was a marked diurnal cycle in levels, with a peak of glucose-6-phosphate during the light period, and a decrease during the dark periods (Figure 5.10). In comparison levels of

glucose-1-phosphate and fructose-6-phosphate showed no notable variation between wild type and SEX4-3 plants (data not shown).

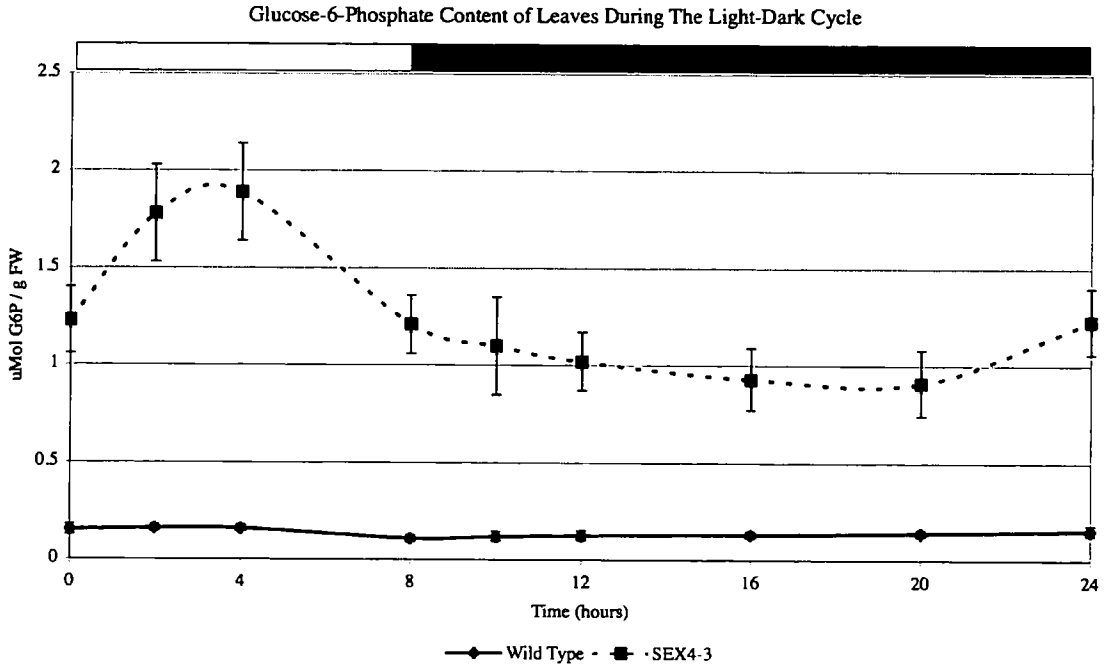


Figure 5.10 Glucose-6-Phosphate Content of Leaves. Graph showing the glucose-6-phosphate content of leaves from mature Wild type (◆) and SEX4-3 (■) plants throughout a 24-hour period. Samples were taken though the 8 hours of light and 16 hours darkness. Error bars represent standard error at each time point. Note: some of the error bars are too small to see clearly on the graph

5.2.4 Starch Levels in Wild Type, SEX4-3 and PTPKIS2-SALK plants

While the mutant lines for AtPTPKIS1 and AtPTPKIS2 (SEX4-3 and PTPKIS2-SALK respectively) show a starch excess phenotype (Fig. 5.1), the degree to which they retain starch in comparison to wild type plants, and to each other, may help in understanding the *in vivo* role for both proteins. To do this samples of leaf tissue were taken throughout the day/night period from plants grown in short day conditions (8h light, 16h dark) and their starch content determined. Soluble carbohydrates were extracted using water and ethanol (as stated in the materials and methods) before the remaining starch was isolated and determined through enzymatic assay.

The data shows increased levels of starch in both SEX4-3 and PTPKIS2-SALK lines when compared to wild type plants, with PTPKIS2-SALK showing lower levels than seen in SEX4-3 (Fig5.11). While both SEX4-3 and PTPKIS2-SALK have significant levels of starch following the dark period, there is evidence for some starch degradation activity in both mutant lines, but not at the same rate as that of wild type. The PTPKIS2-SALK line shows less starch accumulation, and more evidence of breakdown, that the SEX4-3 line.

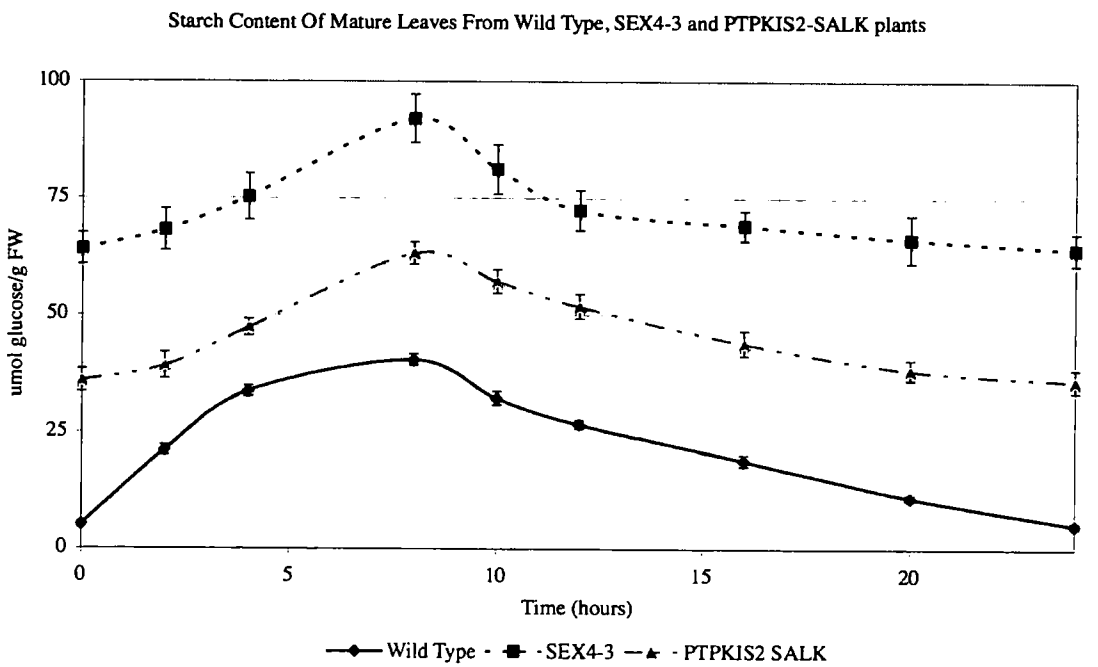


Figure 5.11 **Starch Content of Leaves**. Graph showing the starch content of leaves from mature Wild type (◆), SEX4-3 (■) and PTPKIS2-SALK (▲) plants throughout a 24-hour period. Samples were taken though the 8 hours of light and 16 hours darkness. Error bars represent standard error at each time point. Note: some of the error bars are too small to see clearly on the graph

5.2.5 Starch-Linked Glc-3-P And Glc-6-P In Wild Type and AtPTPKIS1 Mutant (SEX4-3)

The SEX4-1 mutant accumulates starch gradually during consecutive photoperiods with a low starch degradation rate during the dark (Zeeman and Rees, 1999). It has been previously shown that differences in phosphorylation of the C-6 (Yu *et al.*, 2001; Ritte *et al.*, 2002) and C-3 (Baunsgaard *et al.*, 2005; Kotting *et al.*, 2005) hydroxyl group can effect starch mobilisation in transient leaf starch.

To determine any difference in C-6 and C-3 hydroxyl group phosphorylation between AtPTPKIS1 Mutant SEX4-3 and wild type starch, starch isolated from plants grown in short photoperiod at the end of the light period was used. This starch was subject to acid hydrolysis and the neutralised hydrolysate subjected to HPAEC-PAD on a PA-1 column (Dionex, Rødovre, Denmark) according to (Blennow *et al.*, 1998). The concentration of Glc-6-P and Glc-3-P linked to starch was 5.4 ± 0.6 nmol mg^{-1} starch, 0.21 ± 0.03 nmol mg^{-1} starch for wild type and 5.1 ± 0.8 nmol mg^{-1} starch and 0.18 ± 0.05 nmol mg^{-1} starch for SEX4-3 respectively (Fig. 5.12); although there is a small decrease in phosphorylated glucose residues in SEX4-3 starch, the differences are not significant

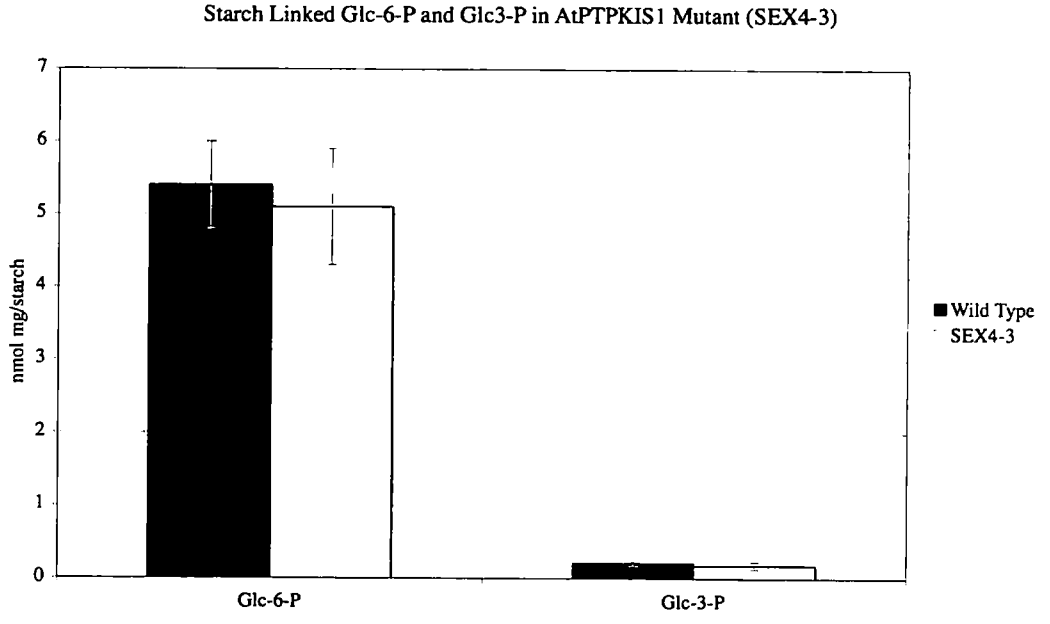


Figure 5.12 Starch linked Glc-6-P and Glc -3-P Levels. Graph showing the starch linked Glc-6-P and Glc -3-P Levels content of leaves from mature Wild type (black bars) and SEX4-3 (white bars) plants at the end of the light period. Error bars represent standard error at each time point. Note: some of the error bars for Glu-3-P are too small to see clearly on the graph.

5.3 SELECTED GENE EXPRESSION ANALYSIS OF STARCH EXCESS LINES

5.3.1 Transcript Levels Of AtPTPKIS1 In Wild Type And AtPTPKIS2 Mutant Backgrounds Through the Light-Dark Cycle.

Total RNA was isolated from the leaves of wild type and AtPTPKIS2 mutant plants (SALK line SALK_053285) during the diurnal cycle. Plants were grown in short day conditions (16h dark, 8h light) with RNA pooled from 6 independent plants. Specific primers were designed (Table 5.1) for the ORF of AtPTPKIS1 and Actin1 (used to standardise samples) and Real Time PCR was conducted.

Gene	Forward Primer	Reverse Primer
Actin1	TGGAACTGGAATGGTTAAGGCTGG	TCTCCAGAGTCGAGCACAATACCG
AtPTPKIS1 (At3g52180)	TGGTGAATGGACACACAATGAGGC	CCACACTTGTTGGATCGTCCACT

Table 5.1 Primers used in Real Time PCR

The transcript level for both groups of plants increased initially during the light period, reaching a peak at the middle of the light period before decreasing. Wild type plants transcript levels remain consistently low throughout the dark period. In comparison the transcript levels of the mutant plants increases rapidly after 4 hours of the dark period, reaching a peak after 8 hours of darkness, with transcripts around 6 times higher than during the peak in the light period (Fig. 5.13).

These data indicate that the transcription of the AtPTPKIS1 gene is affected during the dark period in plants lacking the AtPTPKIS2 protein. Even with this increase in expression of AtPTPKIS1, the mutant plants still show a starch excess phenotype suggesting either an independent role for both proteins or a requirement for both to be present in order to fulfil all of their required activities.

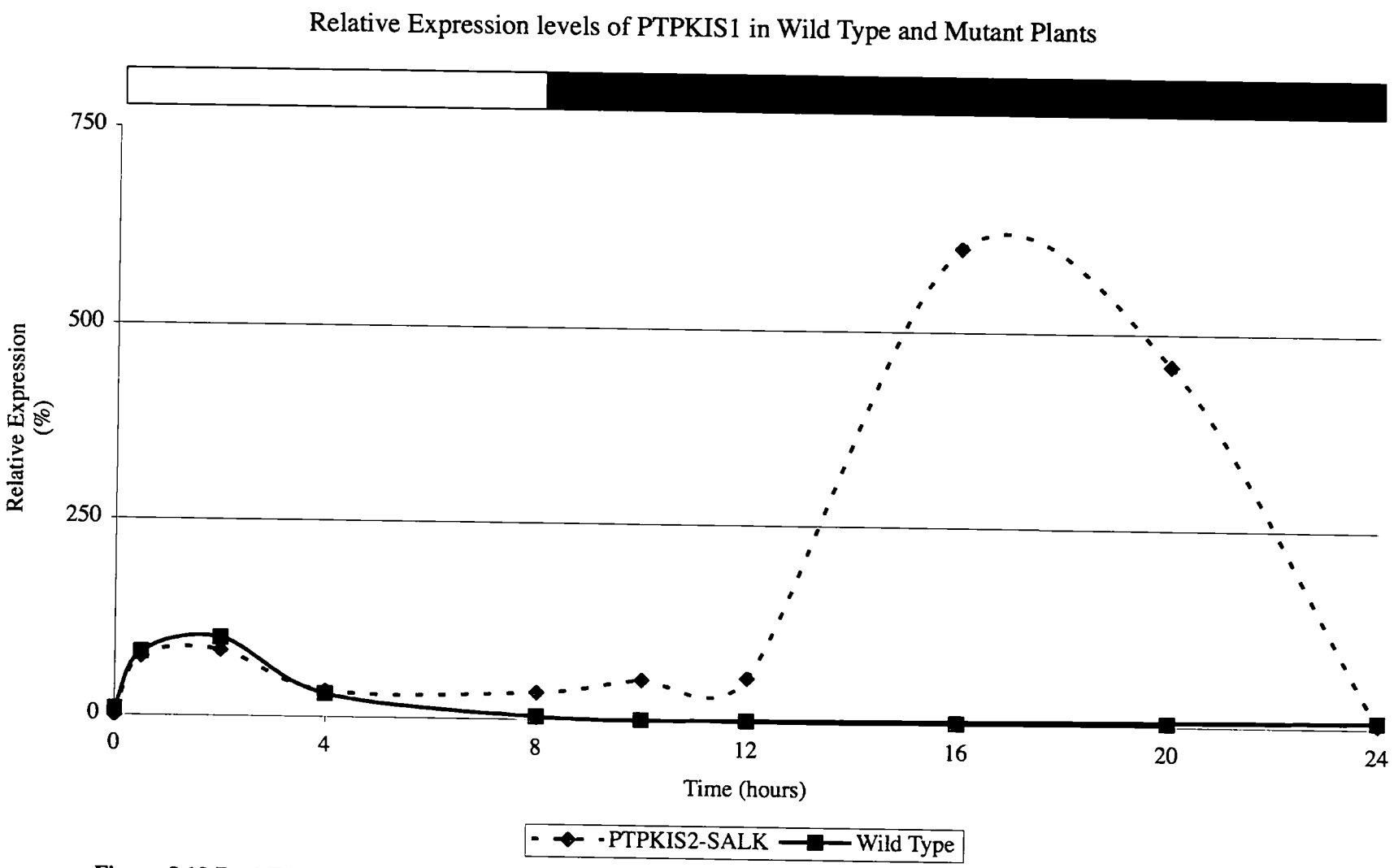


Figure 5.13 Real Time RT-PCR Analysis of AtPTPKIS1 Expression. Graph showing expression levels of the AtPTPKIS1 gene in wild type (■) and PTPKIS2-SALK (◆) plants, as studied through Real Time RT-PCR analysis, throughout a 24-hour period. Samples were taken though the 8 hours of light and 16 hours darkness. All expression levels are shown relative to the product of the ACTIN1 gene.

5.3.2 Correlation In Gene Expression Of AtPTPKIS1, AtPTPKIS2 And AtPTPKL1

Gene expression data were mined from public databases for AtPTPKIS1 (At3g52180), AtPTPKIS2 (At3g01510) and AtPTPKL1 (At3g10940). Using data from a developmental survey of Arabidopsis (Schmid *et al.*, 2005) obtained using the Affimetrix ATH1 Arabidopsis Genome array, a transcript co-response analysis was carried out through the Compatative Systems Biology Project (CSB) (Steinhauser *et al.*, 2004) with AtPTPKIS1 (At3g52180), AtPTPKIS2 (At3g01510) and AtPTPKL1 (At3g10940). This showed a strong correlation between the expression patterns of transcripts encoded by the loci of the three putative phosphatases. The correlation coefficients for single gene comparisons, shown as spearman coefficients following non-parametric Spearman's Rho rank correlation, are: AtPTPKIS1/ AtPTPKIS2 (0.8023); AtPTPKIS1/ AtPLKIS1 (0.7432); AtPTPKIS2/ AtPLKIS1 (0.8717) (Table 5.2). These data suggest similar mechanisms for regulation of expression for the above proteins, supporting their involvement in a common process.

Gene 1	*Gene 2*	*Spearman's Rho rank*	*p value*
AtPTPKIS2	AtPTPKL1	0.8717	0.00e+00
AtPTPKIS1	AtPTPKIS2	0.8023	2.66e-15
AtPTPKIS1	AtPTPKL1	0.7432	3.05e-12

Table 5.2 Transcript Co-Response Analysis. Table showing the correlation coefficients for single gene comparisons, as spearman coefficients following non-parametric Spearman's Rho rank correlation on AtPTPKIS1, AtPTPKIS2 and AtPTPKL1. Data used was from a developmental survey of Arabidopsis obtained using the Affimetrix ATH1 Arabidopsis Genome array (Schmid *et al.*, 2005).

5.4 DISCUSSION

The similarity in starch phosphorylation levels in wild type and SEX4-3 mutants greatly reduces the likelihood that the starch excess phenotype seen in SEX4 mutants is due to an effect on the activity of GWD1 or GWD3/PWD1. It is however important to consider the fact that this method of measuring starch phosphorylation fails to discriminate between the phosphorylation of starch on the surface of the granule and that found deep within the granule. Due to the increased size of the starch granules in SEX4 plants (Zeeman *et al.*, 2002), any variation in phosphorylation local to the starch granule surface could quite easily be hidden by the phosphoylation levels of the starch in the interior of the granule. While it is possible to rule out variation in GWD1 or GWD3/PWD1 activity as causing the starch excess in SEX4, as this is identifiable in lower levels of total phosphate (Glc-6-P and Glc-3-P) and Glc-3-P respectively throughout the starch (Yu *et al.*, 2001; Baunsgaard *et al.*, 2005; Kotting *et al.*, 2005), it is not possible to rule out some effect on phosphorylation/dephosphorylation of the starch granule in localised regions.

The starch accumulation phenotype is a result of decreased degradation, which was shown by the assays of starch content carried out. Nevertheless, all lines do show some starch degradation, with the relative activity being wild type line > PTPKIS2-SALK line > SEX4-3 line (PTPKIS1 mutant). The severity of the mutation, with regards starch degradation, is consistently less when PTPKIS2 is knocked out than if the knockout is in PTPKIS1. Since PTPKIS2 does not appear to be an active phosphatase, it is less likely to have a direct effect in metabolic regulation and may play the role of a modulator. The decreased degradation of starch in the mutant lines is also shown by lower levels of hexoses during the dark period, when photosynthesis is not taking place, and starch degradation is the primary source of free hexoses. Once again, the PTPKIS1 knockout shows a more marked decrease in hexose levels than the PTPKIS2 knockout. The shift of diurnal regulation in the latter line is again indicative of a modulator function for PTPKIS2. The altered transcription of the PTPKIS1 gene in the PTPKIS2 knockout line, compared to wild type, once again suggests a modulatory function.

In the analysis of the variation of sugar and starch levels throughout the day/night period between wild type, SEX4-1 and PTPKIS2-SALK lines show a pattern similar to those previously identified in the starch excess phenotype by Zeeman *et al.*, (1999), with the exception that the SEX4-3 line does not show the high daytime levels of leaf glucose and fructose previously identified in the SEX4-1 line. There is a small increase in the levels of hexoses, but <10% over wild-type, whereas levels in the SEX4-1 mutant are increased by >50%. Further research will be necessary to show whether the increased levels of free hexoses in the SEX4-1 line result solely from PTPKIS1 knockout, in view of the mutation in this line possibly affecting a neighbouring gene. Nevertheless, effects of PTPKIS1 knockout on hexose levels are seen in the SEX4-3 line, in the form of accumulation of glucose-6-phosphate.

The presence of high levels of glucose-6-phosphate in SEX4-3 plants may suggest that PTPKIS1 acts as a phosphatase towards this metabolite; however no activity was seen when glucose-6-phosphate was used as a phosphate substrate for any of the studied proteins in chapter 3. It is more likely this high glucose-6-phosphate level is due to the general effects of the starch excess phenotype upon the plant, or an inability to dephosphorylate longer chain phosphoglucans, such as phosphomaltooligosaccharides, or starch. Hydrolysis of starch containing glucose residues phosphorylated at the C-6 –OH would lead to production of glucose-6-phosphate, but it is not clear why this metabolite is not then immediately converted to glucose-1-phosphate by phosphoglucomutase for starch synthesis, or fructose-6-phosphate by phosphoglucose isomerase for glycolysis.

The similar expression patterns seen between PTPKIS1, PTPKIS2 and PTPKL1 suggest their expression is controlled in a similar way, supporting involvement in a common process. The PTPKIS1, PTPKIS2 and PTPKL1 genes in Arabidopsis also show similarity in expression pattern to other genes involved in starch degradation, (Smith *et al.*, 2004; Niittyla *et al.*, 2006) which supports them having a role in the process of starch degradation.

Chapter 6

GENERAL DISCUSSION

Energy storing macromolecules of α-glucans are widespread in a variety of kingdoms; in animals and fungi the α-glucan is glycogen, and in plants it is starch. The structure of these macromolecules is optimised to minimize storage space and maximize energy concentration while providing a further layer of regulation of biosynthesis and degradation (Manners, 1988; Manners, 1990). While glycogens highly branched and soluble structure seems greatly different from the insoluble starch with its limited branching, the two forms of energy storage share a lot of common features. Mutations in enzymes involved in the metabolism of glycogen, such as glycogen branching enzyme (Raben *et al.*, 2001) or Laforin (Minassian, 2002), can result in a phenotype similar to that of starch where large insoluble α-glucans form in the cells. Alternatively, mutations in starch metabolism enzymes can result in highly branched phytoglycogen (Zeeman *et al.*, 1998b). Within different starch there is a range of branching, from some showing high levels of branching to those containing high levels of amylose. These similarities between glycogen and starches should not be unexpected when considering the similarity in the enzymes responsible for their formation (see introduction).

The aim of this work was to study AtPTPKIS1 and its homologues in order to better understand its function within plants. Originally it was suggested that PTPKIS1 formed a group of novel KIS domain containing proteins which acted to regulate SnRKs in a novel manner when compared to the standard SnRK complexes previously identified in plants, yeast, and animals (Lumbreras *et al.*, 2001; Fordham-Skelton *et al.*, 2002). Recently published data has shown that AtPTPKIS1 localises to the chloroplast in plants (Niittyla *et al.*, 2006), reducing the likelihood of it interacting with SnRKs found in the cytoplasm. Results presented here and in recent publications (Kerk *et al.*, 2006; Niittyla *et al.*, 2006; Sokolov *et al.*, 2006; Gentry *et al.*, 2007) have shown AtPTPKIS1 to be a phosphoglucan phosphatase with a C-

terminal CBM. Mutations in this gene are shown to be responsible for the previously identified SEX4 phenotype. Previous comparisons drawn between the predicted PTP domain of AtPTPKIS1 and Laforin have become more significant with the identification of phosphoglucan phosphatase activity in both enzymes, and the identification of a C-terminal CBM in PTPKIS1 comparable to that found at the N-terminus of Laforin. While their phosphatase domains are homologous, the CBMs lack a similarly high degree of homology while being located at opposite ends of the protein.

Mutation in the laforin gene not only results in decreased branching of the glycogen molecule, but also an increase in the glycogen bound phosphate (Tagliabracci *et al.*, 2007) while mutation of the AtPTPKIS1 gene (SEX4) results in an increase in starch and like mutations in laforin, a decrease in the branching of that starch (Zeeman *et al.*, 2002). The data presented here shows no significant increase in starch bound phosphate in plants unable to produce active PTPKIS1 protein. This may suggest a variation in role between Laforin and PTPKIS1, or it may be due to the insoluble nature of starch and any change in starch bound phosphate may only be present in the surface of the granule.

While no GWD1 or GWD3/PWD1 homologues have been identified in animals, when laforin is mutated there is evidence for an increase in the phosphate levels of glycogen, suggesting the presence of an enzyme with a similar role to GWD1/GWD3. Analysis of genes whose mutation is responsible for the formation of lafora bodies has identified Laforin (EPM2A) and Malin (NHLRC1) but these alone do not account for all the cases of lafora disease (Chan *et al.*, 2003; Ganesh *et al.*, 2006). It may be possible that a protein with a similar role to that of GWD1 and GWD3/PWD1 is involved in glycogen metabolism, and it may be mutation in this, which results in the additional cases of Lafora disease. While it may not have homologous structure to GWD1 or GWD3/PWD1, it is possible that it has a similar function.

Redox regulation is seen in a number of enzymes involved in starch metabolism, such as ADP-glucose pyrophosphorylase in which post-translational redox modification regulates activity (Tiessen *et al.*, 2002; Hendriks *et al.*, 2003). Results

presented here and previously (Fordham-Skelton *et al.*, 2002; Sokolov *et al.*, 2006) show that PTPKIS1 activity is redox regulated, requiring reduction in order to be active. GWD1 has been shown previously to show a similar expression pattern to that of AtPTPKIS1 (Smith *et al.*, 2004), and has also been shown to require reduction in order to phosphorylate glucans (Mikkelsen *et al.*, 2005) with reduction by thioredoxin activating oxidised starch bound GWD1. Together this suggests that redox regulation of activity plays an important role in the regulation of a number of the enzymes involved in starch metabolism.

Evidence presented in this work shows that PTPKIS2 may act to modulate the activity of PTPKIS1 and that when present, it can increase the phosphatase activity of PTPKIS1 against phosphoglucans to a comparable level of Laforin (Gentry *et al.*, 2007), which in itself has been shown to be more active in its homodimeric form (Liu *et al.*, 2006). Work in which the SEX4 phenotype was partially rescued through transformation with a vector encoding a chloroplastic localised laforin (Gentry *et al.*, 2007) suggests at a similarity of function. However the lack of complete rescue makes it likely that either the regulation of their phosphoglucans phosphatase activity, or the role of the protein in some other capacity, is not being duplicated in laforin. The variation in the SEX4 and PTPKIS2-SALK phenotypes supports this theory, as it suggests a similarity in function, but not a co-dependence for activity.

Identification of the phosphatase activity of the AtPTPKL1 protein, given its similarity in expression when compared to AtPTPKIS1 and AtPTPKIS2 suggests a role in starch metabolism. Its inability, like PTPKIS1, to show activity against synthetic phosphopeptides, while showing some activity against phosphoglucans, supports this. While the activity seen against macromolecular phosphoglucans was significantly lower than that of AtPTPKIS1, this could be attributed to the latter containing a carbohydrate binding domain. It may be possible that AtPTPKL1 acts within the chloroplastic stroma upon shorter phospho-maltooligosacharides, released through normal starch degradation.

6.1 Phosphate Wave Hypothesis

Using the information provided within this section of work, in combination with that previously published, it is possible to postulate a number of theories as to the mechanism of starch degradation. While a lot has been published recently on the processing of linear glucans into maltose and its subsequent export into the cytosol (Smith *et al.*, 2005), there is still a lot of conjecture around the initial attack upon the transient leaf starch granule.

The SEX1 mutants, deficient in GWD1, have shown a requirement for starch phosphorylation in order to facilitate starch mobilisation (Yu *et al.*, 2001; Ritte *et al.*, 2002). With it being postulated that phosphate increases the surface solubility of starch (Baunsgaard *et al.*, 2005), allowing other enzymes to act upon it more freely. It may be possible that unphosphorylated starch is difficult for β-amylase to act upon, and in turn the insolubility would limit the accessibility of ISA3 to the branched regions below the granules surface. Recent publications have shown that activity of β-amylase is greatly increased in the presence of GWD1 (Edner *et al.*, 2007) and that phosphorylation of glucans resulting in greater beta-amylase activity upon them (Hajezi *et al* 2008).

In contrast, the SEX4 mutation would remove, or at least reduce, the potential to dephosphorylate starch. This would suggest an increase in the surface solubility of the starch granule, a situation likely to be favourable for β-amylase to act upon the granule. While this work shows that the level of phosphate within the granule as a whole in comparable to wild type plants, this does not rule out a redistribution, in which phosphorylated starch is more common at the granule surface. While it is possible to predict that a more soluble starch granule would be more susceptible to attack by β-amylase, it is possible that the increased phosphate levels would act to inhibit the β-amylase activity, in a similar way to that seen by branch points. This is supported by data from SEX4 mutants which shows an increase in β-amylase levels, while there is only limited evidence of starch degradation (Zeeman *et al.*, 1998a).

Starch levels do decrease in both SEX4 and PTPKIS2-SALK, this may be due to the ability to degrade the surface slowly with this increase in β-amylase levels. Alternatively this increased phosphorylation may allow ISA3 greater access to the bands of highly branched starch. In SEX4 mutants, levels of unbranched starch was seen to be higher than both wild type and SEX1 mutants lines (Zeeman *et al.*, 1998a) This increased percentage of unbranched starch, may be due to reduced efficiency of β-amylase acting upon the crystalline lamellae, while the amorphous lamellae is still susceptible to degradation by ISA3 and β-amylase so removing the branched regions, leaving behind unbranched regions.

This may suggest that for optimal degradation of starch granules, a balance must be met between the need to be phosphorylated to increase accessibility, and the need to have unphosphorylated regions so as to not inhibit degradation. Similarly, is could be seen as an equilibrium between GWD1 and GWD3/PWD1 on one side, and PTPKIS1 and PTPKIS2 on the other. An example of such a mechanism can be seen in figure 6.1.

Figure 6.1 Phosphate Wave Hypothesis. (Following page) Starch chains on the surface of the granule are phosphorylated by GWD1 and GWD3 increasing the solubility of these regions. With this increase in solubility, the PTPKIS1/PTPKIS2 complex is able to more readily access the surface of the starch granule, dephosphorylating those phosphate residues near the surface. This dephosphorylated starch, with phosphate residues at its base, is attacked more easily by β amylase. When the β amylase reaches the phosphorylated glucose residues, its activity is greatly reduced. At this point, GWD1 and GWD3 again phosphorylated the deeper surface residues, allowing the cycle to continue

When amorphous, branched regions are reached by degrading enzymes, ISA3 and β-amylase act together to break down the highly branched region into maltose and other maltooligosacherides. Released phosphomaltooligosacharides may be dephosphorylated in the stroma by general phosphatases, PTPKL1, or PTPKIS1/PTPKIS2.

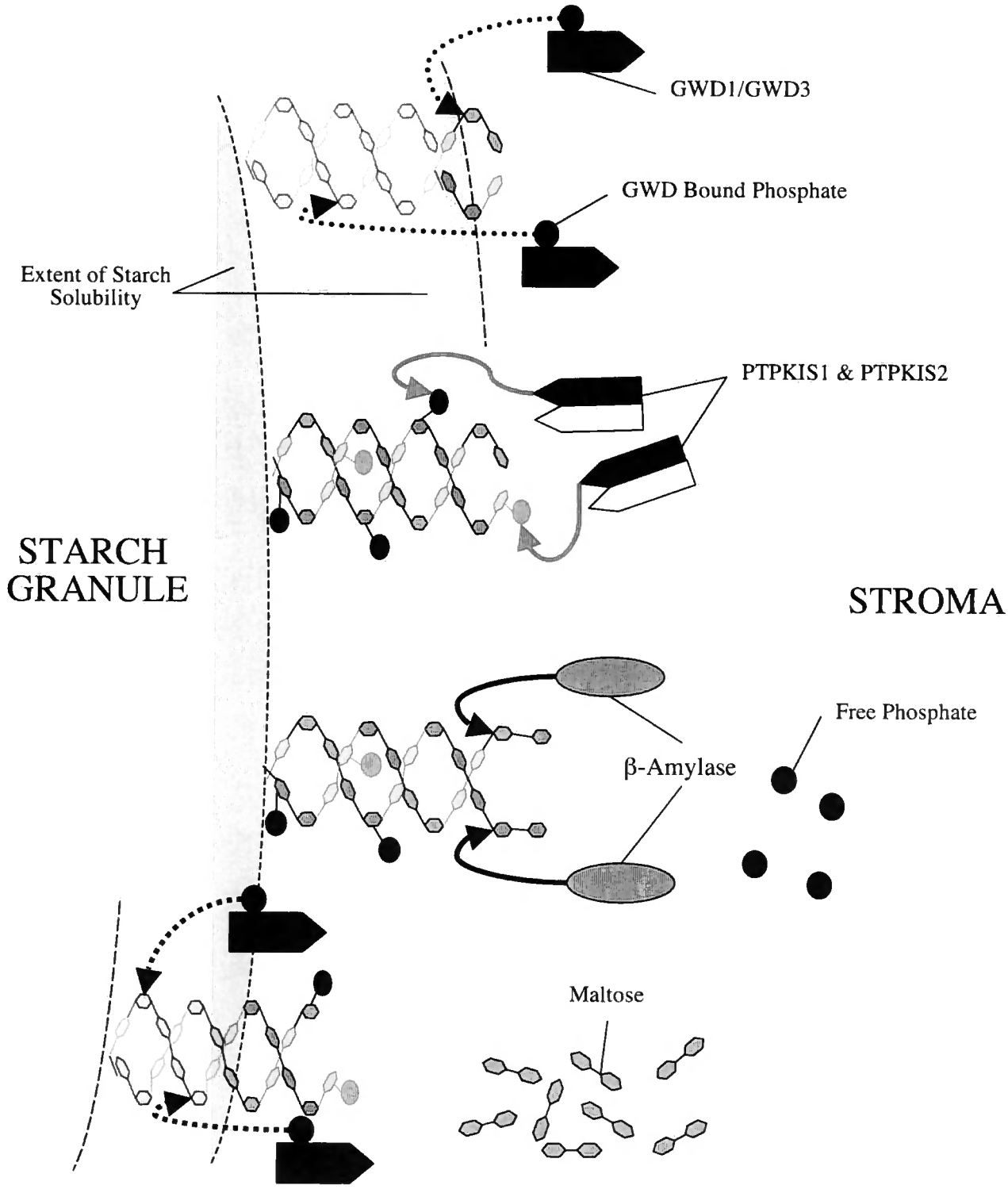


Figure 6.1 Phosphate Wave Hypothesis.

6.2 Branch Regulation Hypothesis

Mutation in PTPKIS1 and PTPKIS2 causes a phenotype in which starch degradation is reduced. This suggests a role in starch degradation, but it is also possible that PTPKIS1 and PTPKIS2 play a role in starch synthesis, and it is alteration in the structure of the starch during synthesis, which results in an altered rate of degradation. The similarities between the structure and activity of Laforin and PTPKIS1, combined with the similarity of affect mutations in either have on glycogen and starch structure respectively, provide us with an insight into alternative roles in starch synthesis.

Recent publications have shown that the glycogen bound phosphate levels in mice in which the laforin gene (Epm2a) is knocked out, was around 40% higher (Tagliabracci *et al.*, 2007). It has been proposed that this phosphorylation may result in aberrant branching resulting in the formation of lafora bodies. This may suggest a similar action in starch, where without PTPKIS1 to dephosphorylate the starch, branching occurs less frequently, resulting in a granule with a higher amylose content, which is in turn more insoluble, requiring greater β-amylase activity, as is seen in the SEX4 mutants (Zeeman *et al.*, 1998a), in order to degrade the starch. Unlike mutations in Laforin, SEX4 shows no elevated levels in starch bound phosphate. This may be accounted for by the variation in the ratio of amylose to amylopectin found between starch from wild type and SEX4 mutants, as it may be possible that it occurs more in one than the other.

Branching may be inhibited by the presence of phosphate residues. GWD1 has a higher activity near branch points, as it gets further from branch points during starch synthesis, its activity drops. This combined with the activity of the phosphatase (PTPKIS1 and PTPKIS2) means that there is a decrease in phosphorylated residues as the distance from the initial branch point increases. This allows branching enzymes to act, resulting in a branched region, forming the layer of highly branched starch of the amorphous lamellae. In this way, the width of the crystalline lamellae and the amorphous lamellae could be described in terms of the equilibrium between GWD1 and GWD3/PWD1 activity as it decreases over distance from branch point

and the activity of PTPKIS1 and PTPKIS2. An example of such a mechanism can be seen in figure 6.2.

Alternatively to regulating directly the position of branch points, phosphorylation may regulate the removal of branch points. It has been shown branches are formed and then removed throughout starch synthesis by very short pulse methods. In many ways this could act in a similar fashion to the previous phosphate wave theory, with the activity occurring during branch formation with phosphate being removed to stop branching enzyme activity

Figure 6.2 Branch Point Hypothesis. (Following page) Starch chains on the surface of the granule are phosphorylated by GWD1 and GWD3, resulting in an inability for Starch Branching Enzyme (SBE) to form branch points, while Starch Synthase (SS) continues to elongate the starch chains. The phosphoglucans phosphatase activity of the PTPKIS1/PTPKIS2 complex is much lower than the activity of GWD1/GWD3, resulting in a higher occurrence of phosphorylated glucose residues.

As the length of the un-branched starch chain increases, GWD acts further from the original branched region, resulting in a decrease in GWD activity. The ratio of GWD:PTPKIS activity moves in favor of PTPKIS activity. Dephosphorylation by PTPKIS1/PTPKIS2 results in lower numbers of phosphorylated residues further from the branched region. Which allows SBE to act, forming branch points, with the distance between this branched region and the previous one a result of the interplay of activities of GWD1/GWD3 and PTPKIS1/PTPKIS2

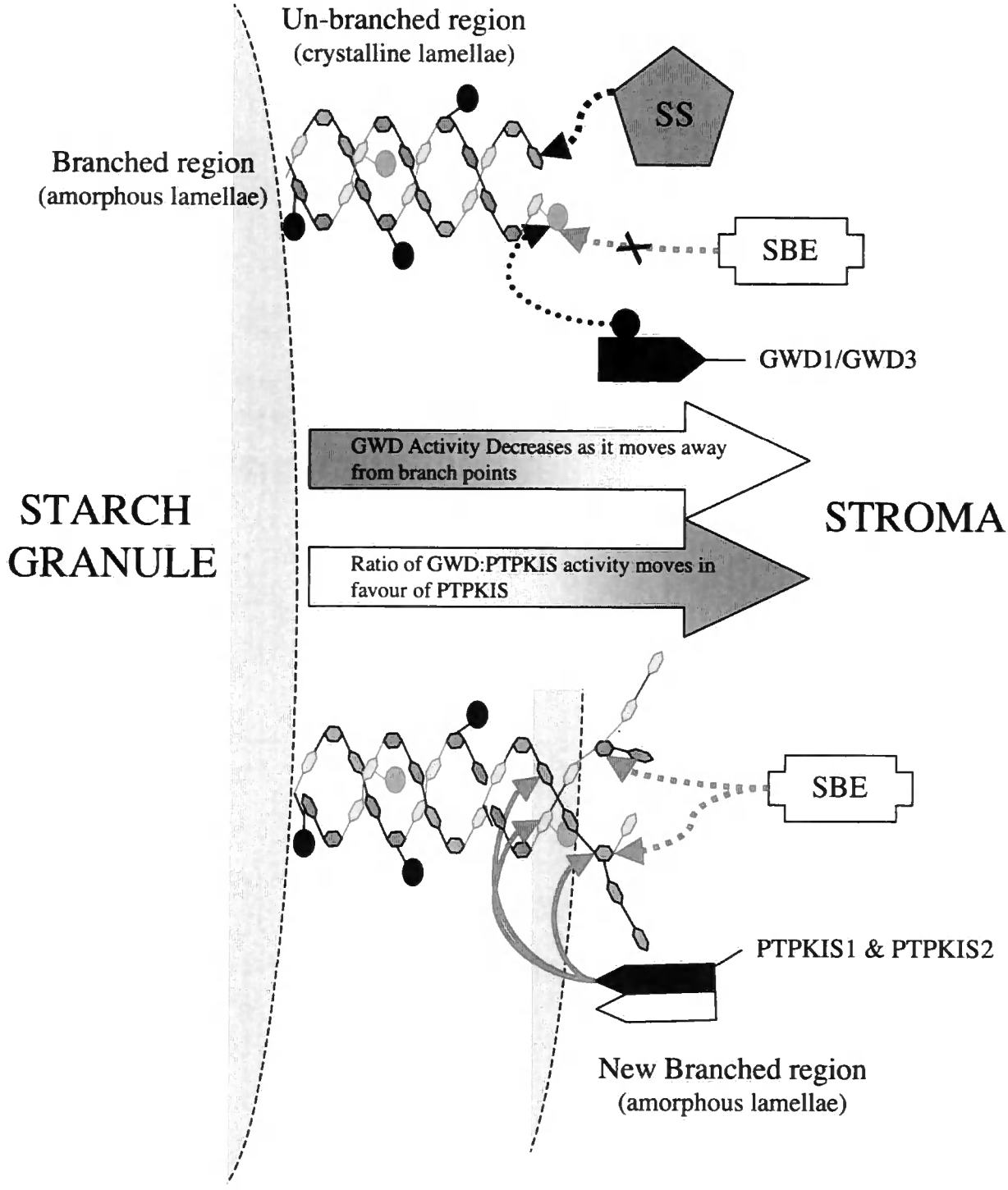


Figure 6.2 Branch Point Hypothesis.

6.3 Starch and Glycogen Metabolism

The mutations in Laforin and PTPKIS1 show the importance of phosphoglucan phosphatases in glycogen and starch metabolism respectively. As with mutations in other glycogen/starch metabolism enzymes, mutations in either the PTPKIS1 or Laforin gene cause of similar physiological chance in the carbohydrate storage molecule. While starch and glycogen have a number of differences in their structure, there are a great deal of similarities in the enzymes involved, and the regulatory mechanisms used.

A combination of the above branching theory, during macromolecule formation and phosphate wave during its degradation could account for the role of both laforin and PTPKIS1/PTPKIS2 in starch and glycogen metabolism. In glycogen the major determinant of structure would be the branch theory during formation, as seen by the limited phosphate content in native glycogen. In plants, the higher activity of GWD1 and GWD3/PWD1 would reduce the branching on formation, and give the phosphate wave theory during degradation a bigger role in the determination of the granule structure.

6.4 Conclusions

The results of this work show:

- AtPTPKIS1 acts as a phosphoglucans phosphatase.
- AtPTPKIS2 modulates AtPTPKIS1 activity against phosphoglucans.
- AtPTPKIS2 shows no phosphatase activity to any generic substrates.
- AtPTPKL1 shows similar levels of phosphatase activity to AtPTPKIS1 against generic substrates
- AtPTPKL1 shows a significantly lower activity against phosphoglucans when compared to AtPTPKIS1
- AtPTPKIS1, AtPTPKIS2 and AtPTPKL1 show no activity towards phosphorylated peptides.
- The previously predicted KIS domains on AtPTPKIS1 acts as carbohydrate binding domain
- Mutational analysis confirms the CBM in AtPTPKIS1 contains residues consistent with standard CBMs from the CBM20 family
- The CBM of AtPTPKIS1 contains a novel pair of sugar tongs required for carbohydrate binding activity.
- Plants lacking AtPTPKIS1 or AtPTPKIS2 expression both show a starch excess phenotype.
- Analysis of sugars shows variation between plants lacking AtPTPKIS1 or AtPTPKIS2 expression.
- Starch bound phosphate shows no variation in the ration of glucose-3-phosphate and glucose-6-phosphate in starch of AtPTPKIS1 mutant plants.

Using this and recently published material it is possible to hypothesis a number of roles for PTPKIS1 and PTPKIS2 in vivo, such as a role in regulation of branch points, or in dephosphorylation of starch during starch degradation.

There are a number of ways in which this work may be carried on. Generation of SEX4/PTPKIS2-SALK double mutant would enable conformation of AtPTPKIS2 as only acting to modulate AtPTPKIS1 activity Double mutants showing a SEX4 would suggest PTPKIS2 only acted to modulate PTPKIS1. If mutations were cumulative it

would suggest alternative roles for AtPTPKIS2 not identified by this work. Site directed mutagenesis of the AtPTPKIS2 CBM to further characterise its affinity to carbohydrate would act to support that data generated through site directed mutagenesis of the CBM of AtPTPKIS1. Through generation of maltooligosacharides with specific phosphorylation patterns, it would be possible to better characterise the phosphoglucan phosphatase activity of PTPKIS1, PTPKIS2 and PTPKL1. This may also allow us to identify a role of AtPTPKL1.

Short pulse label experiments with radiolabeled phosphate in *in-vitro* assays on starch degradation in the presence of PTPKIS and GWD would act to provide data as to the role of PTPKIS and GWD in the proposed phosphate wave theory. In a similar manner, assays on starch branching enzyme activity against substrates with phosphorylation at specific glucose residues would act to support the branch point theory.

REFERENCES

Alderson, A., Sabelli, P.A., Dickinson, J.R., Cole, D., Richardson, M., Kreis, M., Shewry, P.R. and Halford, N.G. (1991) Complementation of snf1, a mutation affecting global regulation of carbon metabolism in yeast by a plant protein kinase cDNA. Proceedings of the National Academy of Science of the United States of America. **88**, 8602-8605.

Alonso, M.D., Lomako, J., Lomako, W.M., and Whelan, W.J. (1995). A new look at the biogenesis of glycogen. FASEB Journal **9**, 1126-1137.

Aparicio-Fabre, R., Guillen, G., Estrada, G., Olivares-Grajales, J., Gurrola, G., and Sanchez, F. (2006). Profilin tyrosine phosphorylation in poly-l-proline-binding regions inhibits binding to phosphoinositide 3-kinase in Phaseolus vulgaris. Plant Journal **47**, 491-500.

Alonso, M.D., Lomako, J., Lomako, W.M., and Whelan, W.J. (1995). A new look at the biogenesis of glycogen. FASEB Journal **9**, 1126-1137.

Armstrong, C.G., Doherty, M.J., and Cohen, P.T. (1998). Identification of the separate domains in the hepatic glycogen-targeting subunit of protein phosphatase 1 that interact with phosphorylase a, glycogen and protein phosphatase 1. Biochemical Journal **336 (Pt 3)**, 699-704.

Baker, A.A., Miles, M.J., and Helbert, W. (2001). Internal structure of the starch granule revealed by AFM. Carbohydrate Research **330**, 249-256.

Ball, S., Guan, H.P., James, M., Myers, A., Keeling, P., Mouille, G., Buleon, A., Colonna, P., and Preiss, J. (1996). From glycogen to amylopectin: a model for the biogenesis of the plant starch granule. Cell **86**, 349-352.

Barford, D. (1995). Protein phosphatases. Current Opinion in Structural Biology **5**, 728-734.

Barford, D. (1996). Molecular mechanisms of the protein serine/threonine phosphatases. Trends in Biochemical Sciences **21**, 407-412.

Barford, D., Jia, Z., and Tonks, N.K. (1995). Protein tyrosine phosphatases take off. Nature Structural Biology **2**, 1043-1053.

Baunsgaard, L., Lutken, H., Mikkelsen, R., Glaring, M.A., Pham, T.T., and Blennow, A. (2005). A novel isoform of glucan, water dikinase

phosphorylates pre-phosphorylated alpha-glucans and is involved in starch degradation in Arabidopsis. Plant Journal **41,** 595-605.

Baykov, A.A., Evtushenko, O.A., and Avaena, S.M. (1988). A Malachite Green Procedure for Orthophosphate Determination and Its Use in Alkaline Phosphatase-Based Enzyme Immunoassay. Analytical Biochemistry **171,** 266-270.

Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C.W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C.H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M.A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B.G., Parkhill, J., and Hopwood, D.A. (2002). Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). Nature **417,** 141-147.

Bhalerao, R.P., Salchert, K., Bako, L., Okresz, L., Szabados, L., Muranaka, T., Machida, Y., Schell, J., and Koncz, C. (1999). Regulatory interaction of PRL1 WD protein with Arabidopsis SNF1-like protein kinases. Proceedings of the National Academy of Sciences of the United States of America **96,** 5322-5327.

Blennow, A., Bay-Smidt, A.M., Olsen, C.E., and Møller, B.L. (1998). Analysis of starch-bound glucose 3-phosphate and glucose 6-phosphate using controlled acid treatment combined with high-performance anion-exchange chromatography. Journal of Chromatography A. **829,** 385-391.

Blennow, A., Vikso-Nielsen, A., and Morell, M.K. (1998). Alpha-glucan binding of potato-tuber starch-branching enzyme I as determined by tryptophan fluorescence quenching, affinity electrophoresis and steady-state kinetics. European Journal of Biochemistry / FEBS **252,** 331-338.

Blennow, A., Nielsen, T.H., Baunsgaard, L., Mikkelsen, R., and Engelsen, S.B. (2002). Starch phosphorylation: a new front line in starch research. Trends in Plant Science **7,** 445-450.

Boraston, A. B., McLean, B. W., Kormos, J. M., Alam, M., Gilkes, N. R., Haynes, C. A., Tomme, P., Kilburn, D. G., Warrer, R. A. J. (1999)

Carbohydrate-binding modules: diversity of structure and function, p. 202-211. *In* **H. J. Gilbert, G. J. Davies, B. Henrissat, and B, Svensson** (ed.), Recent advances in carbohydrate bioengineering. The Royal Society of Chemistry, Cambridge, United Kingdom.

Boraston, A.B., Bolam, D.N., Gilbert, H.J., and Davies, G.J. (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. The Biochemical Journal **382,** 769-781.

Bozonnet, S., Jensen, M.T., Nielsen, M.M., Aghajari, N., Jensen, M.H., Kramhoft, B., Willemoes, M., Tranier, S., Haser, R., and Svensson, B. (2007). The 'pair of sugar tongs' site on the non-catalytic domain C of barley alpha-amylase participates in substrate binding and activity. The FEBS Journal **274,** 5055-5067.

Bouly, J.P., Gissot, L., Lessard, P., Kreis, M., and Thomas, M. (1999). Arabidopsis thaliana proteins related to the yeast SIP and SNF4 interact with AKINalpha1, an SNF1-like protein kinase. Plant Journal **18,** 541-550.

Buléon, A., Colonna, P., Planchot, V., Ball, S. (1998). Starch granules: structure and biosynthesis. International Journal of Biological Macromolecules. **23,** 85-112.

Burton, R.A., Bewley, J.D., Smith, A.M., Bhattacharyya, M.K., Tatge, H., Ring, S., Bull, V., Hamilton, W.D., and Martin, C. (1995). Starch branching enzymes belonging to distinct enzyme families are differentially expressed during pea embryo development. Plant Journal **7,** 3-15.

Bradford, M. M. (1976) Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principles of protein-dye binding. Analytical Biochemistry. **72,** 248-254

Celenza, J.L., and Carlson, M. (1986). A yeast gene that is essential for release from glucose repression encodes a protein kinase. Science. **233,** 1175-1180.

Chan, E.M., Ackerley, C.A., Lohi, H., Ianzano, L., Cortez, M.A., Shannon, P., Scherer, S.W., and Minassian, B.A. (2004). Laforin preferentially binds the neurotoxic starch-like polyglucosans, which form in its absence in progressive myoclonus epilepsy. Human Molecular Genetics. **13,** 1117-1129.

Chan, E.M., Young, E.J., Ianzano, L., Munteanu, I., Zhao, X., Christopoulos, C.C., Avanzini, G., Elia, M., Ackerley, C.A., Jovic, N.J., Bohlega, S., Andermann, E., Rouleau, G.A., Delgado-Escueta, A.V., Minassian, B.A.,

and Scherer, S.W. (2003). Mutations in NHLRC1 cause progressive myoclonus epilepsy. Nature Genetics. **35,** 125-127.

Cohen P. (1989). Structure and regulation of protein phosphatases. Annual Review of Biochemistry. **58**: 453-508

Cohen PTW. (1997) Novel protein serinethreoninephosphatases: variety is the spice of life. Trends in Biochemical Science. **22**: 245-251

Coutinho, P. M., Henrissat, B. (1999) Carbohydrate-activated enzyme: an integrated database approach, p. 3-12. *In* **H. J. Gilbert, G. J. Davies, B. Henrissat, and B, Svensson** (ed.), Recent Advances in Carbohydrate Bioengineering. The Royal Society of Chemistry, Cambridge, United Kingdom.

Darnell, J.E., Jr. (1997). STATs and gene regulation. Science. **277,** 1630-1635.

Delatte, T., Umhang, M., Trevisan, M., Eicke, S., Thorneycroft, D., Smith, S.M., and Zeeman, S.C. (2006). Evidence for distinct mechanisms of starch granule breakdown in plants. The Journal of Biological Chemistry **281,** 12050-12059.

Delpire, E., and Gagnon, K.B. (2006). SPAK and OSR1, key kinases involved in the regulation of chloride transport. Acta physiologica. **187,** 103-113.

Denu, J.M., and Dixon, J.E. (1998). Protein tyrosine phosphatases: mechanisms of catalysis and regulation. Current opinion in chemical biology **2,** 633-641.

Dunphy, W.G. (1994). The decision to enter mitosis. Trends in Cell Biology **4,** 202-207.

Edner, C., Li, J., Albrecht, T., Mahlow, S., Hejazi, M., Hussain, H., Kaplan, F., Guy, C., Smith, S.M., Steup, M., and Ritte, G. (2007). Glucan, water dikinase activity stimulates breakdown of starch granules by plastidial beta-amylases. Plant Physiology. **145,** 17-28.

Fauman, E.B., and Saper, M.A. (1996). Structure and function of the protein tyrosine phosphatases. Trends in Biochemical Sciences. **21,** 413-417.

Faux MC, Scott JD. (1996) More on target with protein phosphorylation: conferring specificity by location. Trends in Biochemical. Sciences. **21**, 312–315

Fernandez-Sanchez, M.E., Criado-Garcia, O., Heath, K.E., Garcia-Fojeda, B., Medrano-Fernandez, I., Gomez-Garre, P., Sanz, P., Serratosa, J.M., and Rodriguez de Cordoba, S. (2003). Laforin, the dual-phosphatase responsible for Lafora disease, interacts with R5 (PTG), a regulatory subunit

of protein phosphatase-1 that enhances glycogen accumulation. Human Molecular Genetics. **12,** 3161-3171.

Fordham-Skelton, A.P., Skipsey, M., Eveans, I.M., Edwards, R., and Gatehouse, J.A. (1999). Higher plant tyrosine-specific protein phosphatases (PTPs) contain novel amino-terminal domains: expression during embryogenesis. Plant Molecular Biology. **39,** 593-605.

Fordham-Skelton, A.P., Chilley, P., Lumbreras, V., Reignoux, S., Fenton, T.R., Dahm, C.C., Pages, M., and Gatehouse, J.A. (2002). A novel higher plant protein tyrosine phosphatase interacts with SNF1-related protein kinases via a KIS (kinase interaction sequence) domain. Plant Journal. **29,** 705-715.

Gancedo, J.M. (1998). Yeast carbon catabolite repression. Microbiology and Molecular Biology Review. **62,** 334-361.

Ganesh, S., Puri, R., Singh, S., Mittal, S., and Dubey, D. (2006). Recent advances in the molecular basis of Lafora's progressive myoclonus epilepsy. Journal of Human Genetics. **51,** 1-8.

Ganesh, S., Tsurutani, N., Amano, K., Mittal, S., Uchikawa, C., Delgado-Escueta, A. V., Yamakawa, K., (2005). Transcriptional profiling of a mouse model for Lafora disease reveals dysregulation of genes involved in the expression and modification of proteins. Neuroscience Letters. **387**, 62-67.

Gentry, M.S., Worby, C.A., and Dixon, J.E. (2005). Insights into Lafora disease: malin is an E3 ubiquitin ligase that ubiquitinates and promotes the degradation of laforin. Proceedings of the National Academy of Sciences of the United States of America. **102,** 8501-8506.

Gentry, M.S., Dowen, R.H., 3rd, Worby, C.A., Mattoo, S., Ecker, J.R., and Dixon, J.E. (2007). The phosphatase laforin crosses evolutionary boundaries and links carbohydrate metabolism to neuronal disease. The Journal of Cell Biology. **178,** 477-488.

Gilkes, N.R., Warren, R.A., Miller, R.C., Jr., and Kilburn, D.G. (1988). Precise excision of the cellulose binding domains from two Cellulomonas fimi cellulases by a homologous protease and the effect on catalysis. The Journal of Biological Chemistry. **263,** 10401-10407.

Giraud, E., and Cuny, G. (1997). Molecular characterization of the alpha-amylase genes of Lactobacillus plantarum A6 and Lactobacillus amylovorus reveals

an unusual 3' end structure with direct tandem repeats and suggests a common evolutionary origin. Gene. **198,** 149-157.

Gissot, L., Polge, C., Bouly, J.P., Lemaitre, T., Kreis, M., and Thomas, M. (2004). AKINbeta3, a plant specific SnRK1 protein, is lacking domains present in yeast and mammals non-catalytic beta-subunits. Plant Molecular Biology. **56,** 747-759.

Gissot, L., Polge, C., Jossier, M., Girin, T., Bouly, J.P., Kreis, M., and Thomas, M. (2006). AKINbetagamma contributes to SnRK1 heterotrimeric complexes and interacts with two proteins implicated in plant pathogen resistance through its KIS/GBD sequence. Plant Physiology. **142,** 931-944.

Guan, X., Chen, L., Wang, J., Geng, H., Chu, X., Zhang, Q., Du, L., and De, W. (2006). Mutations of phosphorylation sites Ser10 and Thr187 of p27Kip1 abolish cytoplasmic redistribution but do not abrogate G0/1 phase arrest in the HepG2 cell line. Biochemical and Biophysical Research Communications. **347,** 601-607.

Gupta, R., Huang, Y., Kieber, J., and Luan, S. (1998). Identification of a dual-specificity protein phosphatase that inactivates a MAP kinase from Arabidopsis. Plant Journal. **16,** 581-589.

Halford, N.G., Bouly, J.-P. and Thomas, M. (2000) SNF-1 related protein kinases (SnRKs) regulators at the heart of the control of carbon metabolism and partitioning. In: Plant Protein Kinases (Kreis, M. and Walker, J.C., eds). London: Academic Press, pp. 405-434.

Han, Y., Sun, F.J., Rosales-Mendoza, S., and Korban, S.S. (2007). Three orthologs in rice, Arabidopsis, and Populus encoding starch branching enzymes (SBEs) are different from other SBE gene families in plants. Gene. **401,** 123-130.

Hardie, D.G., Carling, D. and Carlson, M. (1998) The AMPactivated / SNF1 protein kinase subfamily: metabolic sensors of the eukaryotic cell. Annual Review of Biochemistry. **67**, 821-855.

Hejazi, M., Fettke, J., Haebel, S., Edner, C., Paris, O., Frohberg, C., Steup, M., and Ritte, G. (2008) Glucan Water Dikinase Phosphorylates Crystalline Maltodextrins and thereby initiates Solubilization. The Plant Journal. **55**, 323–334)

Hendriks, J.H., Kolbe, A., Gibon, Y., Stitt, M., and Geigenberger, P. (2003). ADP-glucose pyrophosphorylase is activated by posttranslational redox-modification in response to light and to sugars in leaves of Arabidopsis and other plant species. Plant Physiology. **133,** 838-849.

Hrabak, E.M., Chan, C.W., Gribskov, M., Harper, J.F., Choi, J.H., Halford, N., Kudla, J., Luan, S., Nimmo, H.G., Sussman, M.R., Thomas, M., Walker-Simmons, K., Zhu, J.K., and Harmon, A.C. (2003). The Arabidopsis CDPK-SnRK superfamily of protein kinases. Plant Physiology. **132,** 666-680.

Hubbard MJ, Cohen P. (1993) On target with a new mechanism for the regulation of protein phosphorylation. Trends in Biochemical Sciences. **18**, 172–177

Hudson, E.R., Pan, D.A., James, J., Lucocq, J.M., Hawley, S.A., Green, K.A., Baba, O., Terashima, T., and Hardie, D.G. (2003). A novel domain in AMP-activated protein kinase causes glycogen storage bodies similar to those seen in hereditary cardiac arrhythmias. Current Biology. **13,** 861-866.

Hunter, T. (1998). The role of tyrosine phosphorylation in cell growth and disease. Harvey lectures. **94,** 81-119.

Jiang, R. and Carlson, M. (1997) The Snf1 protein kinase and its activating subunit, Snf4, interact with distinct domains of the Sip1 / Sip2 / Gal83 component in the kinase complex. Molecular and Cellular Biology. 17, 2099-2106.

Johnson, L.N., and O'Reilly, M. (1996). Control by phosphorylation. Current Opinion in Structural Biology. **6,** 762-769.

Johnson, L.N., Snape, P., Martin, J.L., Acharya, K.R., Barford, D., and Oikonomakos, N.G. (1993). Crystallographic binding studies on the allosteric inhibitor glucose-6-phosphate to T state glycogen phosphorylase b. Journal of Molecular Biology. **232,** 253-267.

Kahle, K.T., Rinehart, J., de Los Heros, P., Louvi, A., Meade, P., Vazquez, N., Hebert, S.C., Gamba, G., Gimenez, I., and Lifton, R.P. (2005). WNK3 modulates transport of Cl- in and out of cells: implications for control of cell volume and neuronal excitability. Proceedings of the National Academy of Sciences of the United States of America. **102,** 16783-16788.

Kawazu, T., Nakanishi, Y., Uozumi, N., Sasaki, T., Yamagata, H., Tsukagoshi, N., and Udaka, S. (1987). Cloning and nucleotide sequence of the gene

coding for enzymatically active fragments of the Bacillus polymyxa beta-amylase. Journal of bacteriology **169,** 1564-1570.

Kerk, D., Bulgrien, J., Smith, D.W., Barsam, B., Veretnik, S., and Gribskov, M. (2002). The complement of protein phosphatase catalytic subunits encoded in the genome of Arabidopsis. Plant Physiology. **129,** 908-925.

Kerk, D., Conley, T.R., Rodriguez, F.A., Tran, H.T., Nimick, M., Muench, D.G., and Moorhead, G.B. (2006). A chloroplast-localized dual-specificity protein phosphatase in Arabidopsis contains a phylogenetically dispersed and ancient carbohydrate-binding domain, which binds the polysaccharide starch. Plant Journal. **46,** 400-413.

Kleinow, T., Bhalerao, R., Breuer, F., Umeda, M., Salchert, K., and Koncz, C. (2000). Functional identification of an Arabidopsis snf4 ortholog by screening for heterologous multicopy suppressors of snf4 deficiency in yeast. Plant Journal. **23,** 115-122.

Kotting, O., Pusch, K., Tiessen, A., Geigenberger, P., Steup, M., and Ritte, G. (2005). Identification of a novel enzyme required for starch metabolism in Arabidopsis leaves. The phosphoglucan, water dikinase. Plant Physiology. **137,** 242-252.

Låemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of Bacteriophage T4. Nature. **227,** 680-685.

Lakatos, L., Klein, M., Hofgen, R., and Banfalvi, Z. (1999). Potato StubSNF1 interacts with StubGAL83: a plant protein kinase complex with yeast and mammalian counterparts. Plant Journal. **17,** 569-574.

Liu, Y., Wang, Y., Wu, C., Liu, Y., and Zheng, P. (2006). Dimerization of Laforin is required for its optimal phosphatase activity, regulation of GSK3beta Phosphorylation, and Wnt Signalling. Journal of Biological Chemistry. **281** (46). 34768 - 34774

Lloyd, J.R., Kossmann, J., and Ritte, G. (2005). Leaf starch degradation comes out of the shadows. Trends in Plant Science. **10,** 130-137.

Lumbreras, V., Alba, M.M., Kleinow, T., Koncz, C., and Pages, M. (2001). Domain fusion between SNF1-related kinase subunits during plant evolution. EMBO Reports. **2,** 55-60.

Machovic, M., Svensson, B., MacGregor, E.A., and Janecek, S. (2005). A new clan of CBM families based on bioinformatics of starch-binding domains from families CBM20 and CBM21. The FEBS Journal. **272,** 5497-5513.

Maehama, T., Taylor, G.S., and Dixon, J.E. (2001). PTEN and myotubularin: novel phosphoinositide phosphatases. Annual Review of Biochemistry **70,** 247-279.

Manners, D.J. (1989) Recent Developments in Our understanding of Amylopectin Structure. Carbohydrate Polymers. **11,** 87-112.

Manners, D.J. (1990). Recent Developments in Our Understanding of Glycogen Structure. Carbohydrate Polymers. **16**, 37-82

Maya, R., Balass, M., Kim, S.T., Shkedy, D., Leal, J.F., Shifman, O., Moas, M., Buschmann, T., Ronai, Z., Shiloh, Y., Kastan, M.B., Katzir, E., and Oren, M. (2001). ATM-dependent phosphorylation of Mdm2 on serine 395: role in p53 activation by DNA damage. Genes and Development. **15,** 1067-1077.

Mikkelsen, R., Baunsgaard, L., and Blennow, A. (2004). Functional characterization of alpha-glucan,water dikinase, the starch phosphorylating enzyme. The Biochemical Journal. **377,** 525-532.

Mikkelsen, R., Mutenda, K.E., Mant, A., Schurmann, P., and Blennow, A. (2005). Alpha-glucan, water dikinase (GWD): a plastidic enzyme with redox-regulated and coordinated catalytic activity and binding affinity. Proceedings of the National Academy of Sciences of the United States of America. **102,** 1785-1790.

Mikkelsen, R., Suszkiewicz, K., and Blennow, A. (2006). A novel type carbohydrate-binding module identified in alpha-glucan, water dikinases is specific for regulated plastidial starch metabolism. Biochemistry. **45,** 4674-4682.

Minassian, B.A. (2002). Progressive myoclonus epilepsy with polyglucosan bodies: Lafora disease. Advances in Neurology. **89,** 199-210.

Minassian, B.A., Andrade, D.M., Ianzano, L., Young, E.J., Chan, E., Ackerley, C.A., and Scherer, S.W. (2001). Laforin is a cell membrane and endoplasmic reticulum-associated protein tyrosine phosphatase. Annals of Neurology. **49,** 271-275.

Minassian, B.A., Aiyar, R., Alic, S., Banwell, B., Villanova, M., Fardeau, M., Mandell, J.W., Juel, V.C., Rafii, M., Auranen, M., and Kalimo, H. (2002). Narrowing in on the causative defect of an intriguing X-linked myopathy with excessive autophagy. Neurology. **59,** 596-601.

Minassian, B.A., Lee, J.R., Herbrick, J.A., Huizenga, J., Soder, S., Mungall, A.J., Dunham, I., Gardner, R., Fong, C.Y., Carpenter, S., Jardim, L., Satishchandra, P., Andermann, E., Snead, O.C., 3rd, Lopes-Cendes, I., Tsui, L.C., Delgado-Escueta, A.V., Rouleau, G.A., and Scherer, S.W. (1998). Mutations in a gene encoding a novel protein tyrosine phosphatase cause progressive myoclonus epilepsy. Nature Genetics. **20,** 171-174.

Mitchelhill, K. I., Stapleton, D., Gao, G., House, C., Michell, B., Katsis, F., Witters, L. A., Kemp, B. E., (1994) Mammalian AMP-activated protein kinase shares structural and functional homology with the catalytic domain of yeast Snf1 protein kinase. Journal of Biological Chemistry. **269**, 4, 2361-2364

Muller-Rober, B., Sonnewald, U., and Willmitzer, L. (1992). Inhibition of the ADP-glucose pyrophosphorylase in transgenic potatoes leads to sugar-storing tubers and influences tuber formation and expression of tuber storage protein genes. The EMBO Journal. **11,** 1229-1238.

Myers, A.M., Morell, M.K., James, M.G., and Ball, S.G. (2000). Recent progress toward understanding biosynthesis of the amylopectin crystal. Plant Physiology. **122,** 989-997.

Neel, B.G., and Tonks, N.K. (1997). Protein tyrosine phosphatases in signal transduction. Current Opinion in Cell Biology. **9,** 193-204.

Niittyla, T., Messerli, G., Trevisan, M., Chen, J., Smith, A.M., and Zeeman, S.C. (2004). A previously unknown maltose transporter essential for starch degradation in leaves. Science. **303,** 87-89.

Niittyla, T., Comparot-Moss, S., Lue, W.L., Messerli, G., Trevisan, M., Seymour, M.D.J., Gatehouse, J.A., Villadsen, D., Smith, S.M., Chen, J., Zeeman, S.C., and Smith, A.M. (2006). Similar protein phosphatases control starch metabolism in plants and glycogen metabolism in mammals. The Journal of Biological Chemistry. **281,** 11815-11818.

Paldi, T., Levy, I., and Shoseyov, O. (2003). Glucoamylase starch-binding domain of Aspergillus niger B1: molecular cloning and functional characterization. The Biochemical Journal. **372,** 905-910.

Park, K.M., Kim, D.J., Paik, S.G., Kim, S.J., and Yeom, Y.I. (2006). Role of E2F1 in endoplasmic reticulum stress signaling. Molecules and Cells. **21,** 356-359.

Pawson, T. (1995). Protein modules and signalling networks. Nature **373,** 573-580.

Penninga, D., van der Veen, B.A., Knegtel, R.M., van Hijum, S.A., Rozeboom, H.J., Kalk, K.H., Dijkstra, B.W., and Dijkhuizen, L. (1996). The raw starch binding domain of cyclodextrin glycosyltransferase from Bacillus circulans strain 251. The Journal of Biological Chemistry. **271,** 32777-32784.

Pilling, E., and Smith, A.M. (2003). Growth ring formation in the starch granules of potato tubers. Plant Physiology. **132,** 365-371.

Polekhina, G., Gupta, A., Michell, B.J., van Denderen, B., Murthy, S., Feil, S.C., Jennings, I.G., Campbell, D.J., Witters, L.A., Parker, M.W., Kemp, B.E., and Stapleton, D. (2003). AMPK beta subunit targets metabolic stress sensing to glycogen. Current Biology. **13,** 867-871.

Raben, N., Danon, M., Lu, N., Lee, E., Shliselfeld, L., Skurat, A.V., Roach, P.J., Lawrence, J.C., Jr., Musumeci, O., Shanske, S., DiMauro, S., and Plotz, P. (2001). Surprises of genetic engineering: a possible model of polyglucosan body disease. Neurology. **56,** 1739-1745.

Ritte, G., Lloyd, J.R., Eckermann, N., Rottmann, A., Kossmann, J., and Steup, M. (2002). The starch-related R1 protein is an alpha -glucan, water dikinase. Proceedings of the National Academy of Sciences of the United States of America. **99,** 7166-7171.

Roach, P.J. (2002). Glycogen and its metabolism. Current Molecular Medicine. **2,** 101-120.

Robert, X., Haser, R., Gottschalk, T.E., Ratajczak, F., Driguez, H., Svensson, B., and Aghajari, N. (2003). The structure of barley alpha-amylase isozyme 1 reveals a novel role of domain C in substrate recognition and binding: a pair of sugar tongs. Structure. **11,** 973-984.

Rolland, F., Baena-Gonzalez, E., and Sheen, J. (2006). Sugar sensing and signaling in plants: conserved and novel mechanisms. Annual Review of Plant Biology. **57,** 675-709.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (2001). Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press **Cold Spring Harbor**.

Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Scholkopf, B., Weigel, D., and Lohmann, J.U. (2005). A gene expression map of Arabidopsis thaliana development. Nature Genetics. **37,** 501-506.

Simpson, P.J., Xie, H., Bolam, D.N., Gilbert, H.J., and Williamson, M.P. (2000). The structural basis for the ligand specificity of family 2 carbohydrate-binding modules. The Journal of Biological Chemistry. **275,** 41137-41142.

Smith, S.M., Fulton, D.C., Chia, T., Thorneycroft, D., Chapple, A., Dunstan, H., Hylton, C., Zeeman, S.C., and Smith, A.M. (2004). Diurnal changes in the transcriptome encoding enzymes of starch metabolism provide evidence for both transcriptional and posttranscriptional regulation of starch metabolism in Arabidopsis leaves. Plant Physiology **136,** 2687-2699.

Smith, A.M., Zeeman, S.C., and Smith, S.M. (2005). Starch degradation. Annual Review of Plant Biology. **56,** 73-98.

Smith, R.D., and Walker, J.C. (1996). Plant Protein Phosphatases. Annual Review of Plant Physiology and Plant Molecular Biology. **47,** 101-125.

Sokolov, L.N., Dominguez-Solis, J.R., Allary, A.L., Buchanan, B.B., and Luan, S. (2006). A redox-regulated chloroplast protein phosphatase binds to starch diurnally and functions in its accumulation. Proceedings of the National Academy of Sciences of the United States of America. **103,** 9732-9737.

Steinhauser, D., Usadel, B., Luedemann, A., Thimm, O., and Kopka, J. (2004). CSB.DB: a comprehensive systems-biology database. Bioinformatics (Oxford, England). **20,** 3647-3651.

Sugden, C., Donaghy, P.G., Halford, N.G., and Hardie, D.G. (1999). Two SNF1-related protein kinases from spinach leaf phosphorylate and inactivate 3-hydroxy-3-methylglutaryl-coenzyme A reductase, nitrate reductase, and sucrose phosphate synthase in vitro. Plant Physiology. **120,** 257-274.

Svensson, B., Jespersen, H., Sierks, M.R., and MacGregor, E.A. (1989). Sequence homology between putative raw-starch binding domains from different starch-degrading enzymes. The Biochemical Journal. **264,** 309-311.

Tagliabracci, V.S., Turnbull, J., Wang, W., Girard, J.M., Zhao, X., Skurat, A.V., Delgado-Escueta, A.V., Minassian, B.A., Depaoli-Roach, A.A., and

Roach, P.J. (2007). Laforin is a glycogen phosphatase, deficiency of which leads to elevated phosphorylation of glycogen in vivo. Proceedings of the National Academy of Sciences of the United States of America. **104,** 19262-19266.

Tiessen, A., Hendriks, J.H., Stitt, M., Branscheid, A., Gibon, Y., Farre, E.M., and Geigenberger, P. (2002). Starch synthesis in potato tubers is regulated by post-translational redox modification of ADP-glucose pyrophosphorylase: a novel regulatory mechanism linking starch synthesis to the sucrose supply. The Plant Cell. **14,** 2191-2213.

Tomme, P., Van Tilbeurgh, H., Pettersson, G., Van Damme, J., Vandekerckhove, J., Knowles, J., Teeri, T., and Claeyssens, M. (1988). Studies of the cellulolytic system of Trichoderma reesei QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis. European Journal of Biochemistry / FEBS. **170,** 575-581.

Ulm, R., Revenkova, E., di Sansebastiano, G-P., Bechtold, N. and Paszkowski, J. (2001) *Mitogen-activated protein kinase phosphatase is required for genotoxic stress relief in Arabidopsis.* Genes and Development. **15,** 699-709.

Van Tilbeurgh, H., Tomme, P., Claeyssens, M., Bhikhahai, R. (1986) Limited proteolysis of the cellubiohydrolase I from *Trichoderma reesi.* Seperation of functional domains. FEBS Letters. **204**:223-227.

Wang, W., and Roach, P.J. (2004). Glycogen and related polysaccharides inhibit the laforin dual-specificity protein phosphatase. Biochemical and Biophysical Research Communications. **325,** 726-730.

Williamson, M.P., Le Gal-Coeffet, M.F., Sorimachi, K., Furniss, C.S., Archer, D.B., and Williamson, G. (1997). Function of conserved tryptophans in the Aspergillus niger glucoamylase 1 starch binding domain. Biochemistry. **36,** 7535-7539.

Worby, C.A., Gentry, M.S., and Dixon, J.E. (2006). Laforin, a dual specificity phosphatase that dephosphorylates complex carbohydrates. The Journal of Biological Chemistry. **281,** 30412-30418.

Worby, C.A., Gentry, M.S., and Dixon, J.E. (2008). Malin decreases glycogen accumulation by promoting the degradation of protein targeting to glycogen (PTG). The Journal of Biological Chemistry. **283,** 4069-4076.

Xu, Q., Fu, H-F., Gupta, R. and Luan, S. (1998) Molecular cloning and characterisation of a tyrosine-specific protein phosphatase encoded by a stress-responsive gene in Arabidopsis. Plant Cell. **10**, 849-857.

Yin, X.H., Gerbaud, C., Francou, F.X., Guerineau, M., and Virolle, M.J. (1998). amlC, another amylolytic gene maps close to the amlB locus in Streptomyces lividans TK24. Gene. **215,** 171-180.

Yu, T.S., Kofler, H., Hausler, R.E., Hille, D., Flugge, U.I., Zeeman, S.C., Smith, A.M., Kossmann, J., Lloyd, J., Ritte, G., Steup, M., Lue, W.L., Chen, J., and Weber, A. (2001). The Arabidopsis sex1 mutant is defective in the R1 protein, a general regulator of starch degradation in plants, and not in the chloroplast hexose transporter. Plant Cell. **13,** 1907-1918.

Yu, T.S., Zeeman, S.C., Thorneycroft, D., Fulton, D.C., Dunstan, H., Lue, W.L., Hegemann, B., Tung, S.Y., Umemoto, T., Chapple, A., Tsai, D.L., Wang, S.M., Smith, A.M., Chen, J., and Smith, S.M. (2005). alpha-Amylase is not required for breakdown of transitory starch in Arabidopsis leaves. The Journal of Biological Chemistry **280,** 9773-9779.

Zeeman, S.C., and Rees, T.a. (1999). Changes in Carbohydrate metabolism and assimilate export in starch-excess mutants of Arabidopsis. Plant, Cell and Environment. **22,** 1445-1453.

Zeeman, S.C., Northrop, F., Smith, A.M., and Rees, T. (1998). A starch-accumulating mutant of Arabidopsis thaliana deficient in a chloroplastic starch-hydrolysing enzyme. Plant Journal. **15,** 357-365.

Zeeman, S.C., Tiessen, A., Pilling, E., Kato, K.L., Donald, A.M., and Smith, A.M. (2002). Starch synthesis in Arabidopsis. Granule synthesis, composition, and structure. Plant Physiology. **129,** 516-529.

Zeeman, S.C., Umemoto, T., Lue, W.L., Au-Yeung, P., Martin, C., Smith, A.M., and Chen, J. (1998b). A mutant of Arabidopsis lacking a chloroplastic isoamylase accumulates both starch and phytoglycogen. Plant Cell. **10,** 1699-1712.

Appendix 1

ESTs and sequences predicted through database searches to encode AtPTPKIS1, AtPTPKIS1 and AtPTPKL1 homologues

PTPKIS1 HOMOLOGUES

Oryza sative PTPKIS1

Synonym: OsPTPKIS1, RicePTPKIS1

Accession No.	Tissue	Conditions
CB652367	Leaf	Rice Blast
CB652422	Leaf	Rice Blast
AU225812	Pistil	Tri-nucleate pollen stage
CB631851	Leaf	Rice Blast
C73079	Panicle	Flowering Stage
CA762883	Panicle	Flowering - Water Stress
CB645161	Leaf	Rice Blast
CB646492	Leaf	Rice Blast
CB209746	Immature Leaf	4 Weeks after germination

Predicted amino acid sequence:
MNCLQNLLKEPPIVGSRSMRRPSPLNLTMVRGGSRRSNTVKTASGASTSSAS
GAVEAGTEKSDTYSTNMTQAMGAVLTYRHELGMNYNFIRPDLIVGSCLQSP
LDVDKLRDIGVKTVFCLQQDPDLEYFGVDICAIQEYCLCKIEHCRAEIDFDAF
DLRLRLPAVISKLHKLVNHNGGVTYIHCTAGLGRAPAVTLAYMFWILGYSL
NEGHQLLQSKRACFPKLEAIKLATADILTGLSKNSITLKWESDSCSSVEISGL
DVGWGQIIPLTYNKEKRAWYLERELPEGRYEYKYIVDGKWVCNDNEKKTK
ANADGHVNNYVQVSRDGTSDEERELRERLTGQNPDLTKEERLMIREYLEQY
VER

Solanum tuberosum PTPKIS1

Synonym: StPTPKIS1, PotatoPTPKIS1

Accession No.	Tissue	Conditions
BG889548	Dormant Tuber	1 Month post harvest - 4oC
BM406110	Roots	In Vitro grown Stem cuttings
BI406739	Tuber	
CK276028	Leaf & Root Mixed	Abiotic Stress (Heat, Cold, Salt, Drought)
BQ511893	Tissue	Phytophthora Infestans
BE343106	Tuber Formation	Introduction of tuber formation as described in Bachem et al. (Plant Journal 1996)
CK276029	Leaf & Root Dormant	Abiotic Stress (Heat, Cold, Salt, Drought)
BG890712	Tuber	1 Month post harvest - 4oC

BF459745	Tuber	
BI405969	Tuber	
BQ511892	Mixed Tissue	Phytophthora Infestans

Predicted amino acid sequence:
MNCLQNLPRSSGLPLRSFTGNSRKPFSTVVSLGMTKFADQRLSIVAQVVSGP
ESSTEKDEEKSDTYSHDMTEAMGAVLTYRHDLGMNYNFIRPDLIVGSCLQT
PEDVDKLRSIGVKTIFCLQQNPDLEYFGVDINAIREYANKCGDIEHLRAEIRD
FDAFDLRLRLPAVISILNNAINRNGGVTYIHCTAGLGRAPAVALTYMFWVQS
YKLSEAFDILMSKRSCFPKLDAIKSATADVLTGLKKMPVKLTWHGDNCTTV
EISGLDIGWGQRTPLKFDEGQGLWTLQKDLHEGKYEYKYIVDGEWICNEFE
PITSPNKDGHVNNYVEVLDENPDNITSAAVRKRLTGDDPDLTSDERLIIEQFL
EAYADVE.

Citrus sinensis PTPKIS1

Synonym: CsPTPKIS1, OrangePTPKIS1

Accession No.	Tissue	Conditions
CB292627	Rind - Mature fruit	Cold Temperature
CB292628	Rind - Mature fruit	

Predicted amino acid sequence:
MNCLQNLTRSSALPLQSFRFNQRKPTSSSSFNSLGTMNYADLNRRITVKAITG
STSSKETSDSSEVKEEKSEIYSTNMTEAMGAVLTYRHELRMNYNFIRPDLIVG
SCLPTPEDVDKLRQIGVKTIFCLQQDPXLEYFGVDIIAXQEYAKTYDDIQHIR
AEIRDFDAFDLRMQLPAVISKLYKAINRNGGVTYVHCTAGLGRAPAVALAY
XFWVLGYKLNEAHQLLLSKRPXFPXLDAIKSATADILTGLRKKLVTFSWKG
KNCTTVEISGIDIGWGQRMPLTFDKEQGLWILKRELPEGRYEYKYIVDGEWT
CNKYELVSSPNKDGHVNNYVQVDGAPSSVSEALRNRLTSDDFDLTKDELHK
IRAFLEACPDYE

Poncirus trifoliata PTPKIS1

Synonym: PtPTPKIS1, PoncirusPTPKIS1

Accession No.	Tissue	Conditions
CD575981	Phloem - 10 to 30cm shoots	Rootstock infected with citrus tristeza
CD575173	Phloem - 10 to 30cm shoots	Rootstock infected with citrus tristeza

Predicted amino acid sequence:
MNCLQNLPRSSALPLQSFRFNQRKPTSSSSSFNSLGTMNYADLNRRITAKAIT
GSTSSKETSDSSEVKEEKSEIYSTNMTEAMGAVLTYRHELGMNYNFIRPDLI
VGSCLQTPEDVDKLRQIGVKTIFCLQQDPDLEYFGVDIIAIQEYAKTYDDIQH
IRAEIRDFDAFDLRMRLPAVVSKLYKAINRNGGVTYVHCTAGLGRAPAVAL
AYMFWVLGYKLNEAHRLLLSKRPCFPKLDAIKSATADILTGLRKKLVTFSW
KGKNCTTVEISGIDIGWGQRMPLTFDKEQGLWILKRELPEGRYEYKYIVDGE

WTCNKYELVTYPNKDGHVNNYVQVDDAPSSVSEALRDRLTSDDFDLTKYE
LHKIRAFLEACPDYE

Medicago truncatula PTPKIS1

Synonym: MtPTPKIS1, MedicagoPTPKIS1

Accession No.	Tissue	Conditions
BI267614	Local and Systemic Leaves	Wounded leaves - Spodoptera Exigua (Beet armyworm)
BG581666	Root nodule	1 Month post inoculation with Sinorhizobium Meliloti
BG581814	Root nodule	1 Month post inoculation with Sinorhizobium Meliloti

Predicted amino acid sequence:
MNCLQNLPRSSSVLQFHTLVSPPSTARHHNIFLPSSSSSLSLGINTNSIYTPTML
LKATSGSIPSAETSSSDVEEEVKSEIYSNNMTEAMGAVLTYRHELGMNYNFI
RPDLIVGSCLQTPEDVDKLRKIGVKTIFCLQQNSDLEYFGVDIDAIREYANSC
NDIQHLRAEIRDFDSFDLRKRLPAVISKLYKAINSNGGVTYIHCTAGLGRAPA
VALAYMFWVQGYKLNEANTLLLSKRSCFPKLDAIKSATADILTGLSKKPVT
LSWGHRNCSTVEISGLDIGWGQRVPLNFDDKQGSWFLKKEMFEGRYEYKYI
VDGEWTCNNDELVTSPNKDGHVNNFIEVLDDADSGRASXXERVTGDDPDL
TKDERNRIIEFLXALPNEDL

Brassica napus PTPKIS1

Synonym: BnPTPKIS1, BrassicaPTPKIS1

Accession No.	Tissue	Conditions
BQ704623	Fourth Leaf	3 week seedlings
CD814324	Seed	

Predicted amino acid sequence:
MNCLQNLPRSSVSPMYGFGGNQRVPSSPSSLKMMLLPIKANDLKLRLVLQA
VSDSKSTGAEVSGVSNKEEEEKSDEYSQDMTQAMGAVLTYRHELGMNYSF
VRPDLIVGSCLQTPEDVDKLRKIGVKTIFCLQQGPDLEYFGVDIXSIQAYAKT
FXDIXHIRCXIRDFDAFDLRXRLPAVXSTLYKAVXRNGGVTYVHCTAGMGR
APAXALXYMFWVQGYKLMEAHXXLMSKRTCXPKLDAIRNATIDILTGLXK
XTVTLTLRDKGFSTVEISXLDIGWGQRIPLTLDKGTGFWXLKRELPEGQFEY
KYIIDGEWTYNEQEPFTGPNKDGHTNNYAKVVYDPTSVDGATRERLTSEDP
ELLEEERLKLIQFLETSSEAEV

ESTs which generate partial sequences

Organism	**EST Accession No.**
Descurainia sophia	BU238151
Hordeum vulgare	AV91738 BJ465420 BJ468667 BQ763771 CB870250 CD053966
Zea mays	AY106017 AI920608 BG267495 BI542743 BQ619480 CF014674 CO452300
Picea glauca	CK437744

PTPKIS2 HOMOLOGUES

Oryza sative PTPKIS2

Synonym: OsPTPKIS2, RicePTPKIS2
(containing splice variants a and b)

OsPTPKIS2a

Accession No.	Tissue	Conditions
CB629190	LEAF	Rice Blast
CB634603	LEAF	Rice Blast
CB665979	LEAF	Rice Blast
CB665807	LEAF	Rice Blast
CB664108	LEAF	Rice Blast
CB664109	LEAF	Rice Blast
CB665980	LEAF	Rice Blast
CB665982	LEAF	Rice Blast

OsPTPKIS2b

Accession No.	Tissue	Conditions
BF430624	LEAF	150mM NaCl
CB634604	LEAF	Rice Blast
CB654633	LEAF	Rice Blast
CB641215	LEAF	Rice Blast
CB654633	LEAF	Rice Blast
CB664109	LEAF	Rice Blast
AK119345		Science. 2003 Jul 18;301(5631):376-9
NM_188913		

Predicted amino acid sequence:

OsPTPKIS2a-
MALHLTAAPTIAPSAAAACRSLAPMPPLPAVSCSSRWWRGRRRCVAVVAM
AAAAAADGERPHGHAAEAGTGRMNLNEYMVAVDRPLGVRFALAVDGRVF
VHSLKKGGNAEKSRIIMVGDTLKKAGSREGVGFVDIRDLGDTEMVLKETSG
PCDLVLERPFAPFPIHQLHQNEDYHLLFNKGRVPLTSWNGALLSSKLNESSE
GNGNPGFAIFSPRLLNSHGWAVLSSEQDGLNQRSTSLANRISEIVGLYSDEDD
ADTEWAHGSFPLEEYIKALDRAKGELYYNHSLGMQYSKITEQIFVGSCLQTE
RDVKMLSETMGITAVLNFQSESERTNWGINSEAINNSCRENNILMVNYPIRE
VDSMDLRKKLSFCVGLLLRLIRKNYRIYVTCTTGYDRSPACVIAYLHWVQD
TPLHIAHKFITGLHSCRPDRAAIVWATWDLIALVENGRHDGTPTHSVCFVWN
SGREGEDVELVGDFTSNWKDKVKCDHKGGSRYEAEIRLRHGKYYYKFIAG
GQWRHSTSLPTETDEHGNVNNVIRVGDIARIRPAPSQLQIRDPTVVKVIERAL
TEDERFLLAFAARRMAFAICPIRLSPKQ

OsPTPKIS2b -
MALHLTAAPTIAPSAAAACRSLAPMPPLPAVSCSSRWWRGRRRCVAVVAM
AAAAAADGERPHGHAAEAGTGRMNLNEYMVAVDRPLGVRFALAVDGRVF
VHSLKKGVKHLYHLRLARCGYDATSFVYCLAQSIQSAAAQCSPHPTSLGAV
QLPQLLLRSKIHRQHHAAHGNAEKSRIIMVGDTLKKAGSREGVGFVDIRDLG
DTEYGTSKVSLAETKSEALHRIELANVTSATLAMSKYYIMVLKETSGPCDLV
LERPFAPFPIHQLHQNEDYHLLFNKGRVPLTSWNGALLSSKLNESSEGNGNP
GFAIFSPRLLNSHGWAVLSSEQDGLNQRSTSLANRISEIVGLYSDEDDADTE
WAHGSFPLEEYIKALDRAKGELYYNHSLGMQYSKITEQIFVGSCLQTERDVK
MLSETMGITAVLNFQSESERTNWGINSEAINNSCRENNILMVNYPIREVDSM
DLRKKLSFCVGLLLRLIRKNYRIYVTCTTGYDRSPACVIAYLHWVQDTPLHI
AHKFITGLHSCRPDRAAIVWATWDLIALVENGRHDGTPTHSVCFVWNSGRE
GEDVELVGDFTSNWKDKVKCDHKGGSRYEAEIRLRHGKYYYKFIAGGQWR
HSTSLPTETDEHGNVNNVIRVGDIARIRPAPSQLQIRDPTVVKVIERALTEDER
FLLAFAARRMAFAICPIRLSPKQ

Solanum tuberosum PTPKIS2

Synonym: StPTPKIS2, PotatoPTPKIS2

Accession No.	Tissue	Conditions
BE921193	Leaflets & Petioles	8 Week old
BQ112146	Mixed Tissue	
BF054387	Leaflets & Petioles	8 Week old
BG600243	Sprouting Eyes from Tubers	12-14 weeks post harvest
BI436255	axillary buds of stem explants; growing sink-tubers	
BI436290	axillary buds of stem explants; growing sink-tubers	
BI178258	axillary buds of stem explants; growing sink-tubers	
BI178265	axillary buds of stem explants; growing sink-tubers	
BQ112147	Mixed Tissue	
BE921818	Leaflets & Petioles	8 Week old
BI78737		

Predicted amino acid sequence:
MCSLQLPNCRVFNDGNFSSSSSDKFKKCVVFSSFWGVELCFNDGRFTASSAV
KRRQYRPISVMSSSSDSPFNMNLNEYMVTLVKPLGIRFALSVDGKVFVHALK
KGGNAEKSRIIMVGDTLKKASDSSTGGLIEIYDFGDTEKMMNENSGPCSLVL
ERPSFPFPIHQLYLMDDIDIMYNRGRVPVATWNKKLLASNLRTSCEGSGNSG
FVVFSPKLLTLNGWNVLSDGDQIRQQENLNGTPWLPFSPIINIFSEKDTTDSE
WAHGNFPLEEYVKALDRSKGELYYNHDLGMRYSKITEQIYVGSCIQKESDV
EMLSDVGITAVVNFQSGIEAENWGINANIINESCQRFNILMINYPIREGDSFD
MRKKLPFCVGLLLRLLKKNHRVYVTCTTGFDRSPACVVAYLHWMTDTSLH
AAYNFVTGLHLCKPDRPAIAWATWDLIAMVENGAHDGPATHAVTFVWNG
HEGEDVYLVGDFTGNWKEPIQALHKGGPRFEAEVRLSQGKYLYKYIISGNW
RHSTNSPTERDERGNLNNVIVVGDVASVRPFIQQQKKDANIMKVIERPLTEN
ERFMLAKAARCVAFSICPITLAPK

PTPKL1 HOMOLOGUES

Oryza sative PTPKL11 (Located on Chromosome 11)

Synonym: OsPTPKIS11, RicePTPKIS11

Accession No.	Tissue	Conditions
C72777	Panicle	Flowering Stage
AK072567		
AU173453	Root	Seedling Root
CA762603	Panicle	Drought Stress

Predicted amino acid sequence:
MAAMATAPCFPATPGLPARGAVAARSRMAAGGSRSQRRRSSSGVFLCRSST
TGSSRMEDYNTAMKRMMRNPYEYHHDLGMNYAIISDSLIVGSQPQKPEDID
HLKDEEKVAFILCLQQDKDIEYWGIDFQTVVNRCKELGIKHIRRPAVDFDPD
SLRTQLPKAVSSLEWAISEGKGRVYVHCTAGLGRAPAVAIAYMFWFENMD
LRTAYEKLTSKRPCGPNKRAIRAATYDLAKNDPHKESFDSLPEHAFEGIAGS
ERSLIQERVRALREA

Oryza sative PTPKL12 (Located on Chromosome 12)

Synonym: OsPTPKIS12, RicePTPKIS12

Accession No.	Tissue	Conditions
AU173452	Root	Seedling Root
CB626087	Leaf	3 Week old - lesion mimic SPL11
CB626088	Leaf	3 Week old - lesion mimic SPL11

Predicted amino acid sequence:
MAAMATAPCFPATPGLPARGAVAARSRMAAGGSRSQRRRSSSGVFLCRSST
TGSTRMEDYNTAMKRMMRNPYEYHHDLGMNYAIISDSLIVGSQPQKPEDID
HLKDEEKVAFILCLQQDKDIEYWGIDFQTVVNRCKELGIKHIRRPAVDFDPD
SLRTQLPKAVASLEWAISEGKGRVYVHCTAGLGRAPAVAIAYMFWFENMN
LKTAYEKLTSKRPCGPNKRAIRAATYDLAKNDPHKESFDSLPEHAFEGIADS
ERRLIQERVRALREA

Solanum tuberosum PTPKL1

Synonym: StPTPKL1, PotatoPTPKL1

Accession No.	Tissue	Conditions
	Tuber	Introduction of tuber formation as described in Bachem et al. (Plant
AW906716	Formation	Journal 1996)
BG350600	Tuber	Field Grown

Predicted amino acid sequence:
MRALWNSTCLSPVQNNPLLFSRSSKKYANSLCNFTNKSFQISCKLPESEVKE
NHARSSSNKKMEEYNLAMKRMMRNPYEYHHELGMNYTLITEDLIVGSQPQ
KIEDIDYLKEEENVAFILNLQQDKDIEFWGIDLQSIVTRCSELGIHHMRRPAID
FDPDSLRSVLPKAVSSLEWAISEGKGRVYVHCTAGLGRAPAVSIAYMFWFC
GMDLNTAYDTLVSKRPCGPNKRSIQGATYDLAKNDQWKEPFENLPDYAFA
DVADWERKLIQDRVRALRDT

Vitis vinifera PTPKL1

Synonym: VitisPTPKIS1, VinePTPKIS1, VvPTPKL1

Accession No.	Tissue	Conditions
CA808154	Leaf	Late Season - Infected with bacterial pathogen (Xylella fastidiosa)
CA809091	Leaf	Late Season - Infected with bacterial pathogen (Xylella fastidiosa)
CA813296	Leaf	Late Season - NON-Infected with bacterial pathogen (Xylella fastidiosa)
CA813425	Leaf	Late Season - NON-Infected with bacterial pathogen (Xylella fastidiosa)
CA814373	Leaf	Late Season - NON-Infected with bacterial pathogen (Xylella fastidiosa)
CD799215	Pedicle	Green stage
CF210907	Flower	Bloom
CF210993	Flower	Bloom

Predicted amino acid sequence:
MGAIGNSCFHLAFKNPIENGVVLMKNKSSCKLMVPSNSFKVKRISCKLSESG
VEESATSNRVSNSNNRMEDYNTVMKGMMRNPYEYHHDLGMNYTLITDHLI
VGSQPQKPEDVDHLKQEENVAYILNLQQDKDVEYWEVDLPSIIKRCKELEIR
HMRRPARDFDPDSLRSGLPKAVSSLEWAISEGKGKVYVHCTAGLGRAPAVA
IAYMFWFCGMDLNTAYDTLTSKRPCGPSKQAIRGATYDLAKNDPWKEPLES
LPERAFEDVADWERNLIQDRVRSLRGT

Chlamydomonas reinhardtii PTPKL1

Synonym: CrePTPKL1

Accession No.	Tissue	Conditions
BQ823280	Total	Re Synthesizing Flagella
BE337148	Total	Grown in TAP Conditions
BI873597	Total	Stress Conditions
BP093950	Mixed	Standard Conditions
BE337148	Mixed	Standard Conditions

Predicted amino acid sequence:

MQAHLQSRPVALQQRARPAAQFPSHAGALGRSRRRQAVVVSARHKVDQN
AKDDAYNRNMQREMGWSHLNPYQYHWDRGLYYHEIIPNLICGTQPRNAGE
VDTLADNEGITHILNLQEDKDMHYWGVKIEDIRRACAKHSINHMRRPAKDF
DKGSLRKAIPGAVHTLAGAMAGGGRVYYVHCTAGLGRAPGVCIAYLYWFTD
MQLDEAYSHLTTIRPCGPKRDAIRGATYDVLVGSGVPHNXHNSNGHHGHH
XGHGGHGGPQVAHAPPPLPFESLPEQAYATLSEDDRFALQYRVLKGLC

Appendix 2

Rice PTPKIS2 (A & B Splice) (Os08g29160) Intron/Exon Splice

(B in italics & underlined)

```
                                        (-20) GACGACGAGACCCGCCGGCC

ATGGCGCTCCATCTCACCGCCGCCCCGACCATTGCCCCCTCCGCGGCCGCCGCCTGCAGG    60
 M  A  L  H  L  T  A  A  P  T  I  A  P  S  A  A  A  A  C  R

TCGCTGGCGCCGATGCCGCCTCTCCCCGCCGTTTCGTGCTCGAGTAGGTGGTGGAGGGGG   120
 S  L  A  P  M  P  P  L  P  A  V  S  C  S  S  R  W  W  R  G

AGGCGGCGGTGTGTCGCGGTGGTGGCCATGGCGGCGGCGGCGGCGGCGGACGGGGAGAGG   180
 R  R  R  C  V  A  V  V  A  M  A  A  A  A  A  A  D  G  E  R

CCGCATGGGCACGCGGCGGAGGCTGGCACTGGGAGGATGAACCTCAACGAGTACATGGTC   240
 P  H  G  H  A  A  E  A  G  T  G  R  M  N  L  N  E  Y  M  V

GCCGTCGACCGCCCGCTCGGCGTCCGCTTCGCGCTCGCCGTCGACGGACGCGTCTTCGTC   300
 A  V  D  R  P  L  G  V  R  F  A  L  A  V  D  G  R  V  F  V

CACTCCCTCAAGAAAGGGGtaaagcatttgtatcatctccggctcgcccgctgcggttac   360
--------------------TAAAGCATTTGTATCATCTCCGGCTCGCCCGCTGCGGTTAC
 H  S  L  K  K  G  V  K  H  L  Y  H  L  R  L  A  R  C  G  Y

gacgctacttcatttgtttactgtttagcgcaatccatccagtctgctgctgctcaatgc   420
GACGCTACTTCATTTGTTTACTGTTTAGCGCAATCCATCCAGTCTGCTGCTGCTCAATGC
 D  A  T  S  F  V  Y  C  L  A  Q  S  I  Q  S  A  A  A  Q  C

tcacctcatcccacttcacttggtgctgtgcaattgccacagctcctgctgcgcagcaaa   480
TCACCTCATCCCACTTCACTTGGTGCTGTGCAATTGCCACAGCTCCTGCTGCGCAGCAAA
 S  P  H  P  T  S  L  G  A  V  Q  L  P  Q  L  L  L  R  S  K


attcacaggcagcatcatgctgcccacgttggtgtttacggcatttcatttcatctgatg   540
ATTCACAGGCAGCATCATGCTGCCCACG------------------------------
 I  H  R  Q  H  H  A  A  H
                        I
tgaattttactgtaaccactcgatgttagtgtcatgagcaaggggatactgggaaattgt   600
cgtctggtcaatagttgacatactagtttacttaacttttcttcttgttcgttgtggaac   660
taagatctttatcatttttaactatctctagttgggtatggttgtggaggccttgcaatt   720
ttgctgactgttggtgctgcatccgtgtgacaggGGAATGCGGAGAAATCACGGATTATC   780
                                   G  N  A  E  K  S  R  I  I

ATGGTAGGGGACACTCTGAAGAAGGCTGGCAGTCGTGAGGGCGTGGGTTTTGTTGACATC   840
 M  V  G  D  T  L  K  K  A  G  S  R  E  G  V  G  F  V  D  I

AGAGACCTAGGTGACACGGAgtatggaactagtaaagtttctctagctgaaacagtattc   900
--------------------GTATGGAACTAGTAAAGTTTCTCTAGCTGAAACA------
 R  D  L  G  D  T  E  Y  G  T  S  K  V  S  L  A  E  T
```

IIA

```
ctttgctcttctgtctacaccattttctgcaaaatttaataaaaaaaatgtgatatgact   960
atttggaactagtaagatgtactgtatctattggaacttagagctcatatttgattacat  1020
ttgataatttaaggtcacgtcaaaattttatgactgtaagatgatgtaaaatggtttagt  1080
tcaaatgttccagaaatcagaagctctgcacagaatagagttagcaaatgtgacaagtgc  1140
-------------AAATCAGAAGCTCTGCACAGAATAGAGTTAGCAAATGTGACAAGTGC
              K  S  E  A  L  H  R  I  E  L  A  N  V  T  S  A

aacattggcaatgtctaaatattacatgtcaactaaaaaataataccagcagtggagtaa  1200
AACATTGGCAATGTCTAAATATTACAT--------------------------------
  T  L  A  M  S  K  Y  Y  I
```

IIB

```
catttctagttgattgtcaaaatgaatacaacttttcttatgttgcctttcaactttcat  1260
gtagAATGGTGTTGAAGGAAACGTCAGGGCCATGCGATCTTGTCCTTGAGAGGCCATTTG  1320
      M  V  L  K  E  T  S  G  P  C  D  L  V  L  E  R  P  F

CTCCTTTCCCGATACATCAGTTGCATCAAAATGAAGATTATCATCTCCTATTTAACAAGG  1380
A  P  F  P  I  H  Q  L  H  Q  N  E  D  Y  H  L  L  F  N  K

GTAGGGTTCCTCTTACTAGCTGGAACGGTGCTCTATTATCATCAAAGCTGAATGAATCAT  1440
G  R  V  P  L  T  S  W  N  G  A  L  L  S  S  K  L  N  E  S

CTGAGGGGAACGGAAATCCTGGATTTGCCATATTCTCGCCAAGGCTGCTAAATTCACATG  1500
S  E  G  N  G  N  P  G  F  A  I  F  S  P  R  L  L  N  S  H

GATGGGCAGTTTTGTCTAGTGAGCAAGATGGACTTAATCAGCGCAGTACTAGCCTTGCAA  1560
G  W  A  V  L  S  S  E  Q  D  G  L  N  Q  R  S  T  S  L  A

ATCGTATAAGTGAGATTGTTGGTTTGTACTCTGATGAGGATGATGCAGATACTGAATGGG  1620
N  R  I  S  E  I  V  G  L  Y  S  D  E  D  D  A  D  T  E  W

CACATGGTAGCTTTCCTTTGGAGGAGTACATTAAAGCACTAGACCGTGCTAAAGGTGAAC  1680
A  H  G  S  F  P  L  E  E  Y  I  K  A  L  D  R  A  K  G  E

TGTACTACAATCATTCACTTGGTATGCAATACAGCAAGgtaattgtcctcctataactaa  1740
L  Y  Y  N  H  S  L  G  M  Q  Y  S  K
```

III

```
ttagttcttggatgtgcatcaaaagattttaatattaacagtgtttttgtcttgttgata  1800
tcctctcaggactttaagcttactcattttccgtatatttttttaacagATTACAGAACAA  1860
                                                  I  T  E  Q

ATATTTGTTGGATCATGCCTACAAACAGAAAGAGATGTGAAAATGCTATCAGAGACTATG  1920
 I  F  V  G  S  C  L  Q  T  E  R  D  V  K  M  L  S  E  T  M

Gtaggttcctcacatatgcaatgacaaatctcataactggtacttcgaaaggctttttgc  1980
```

IV

```
gtgtcctttaattatcaaattaagtttctgatatttcaaaaggcttcaggGTATCACTGC  2040
                                                    G  I  T  A

TGTTCTGAATTTTCAAAGTGAAAGTGAGCGCACCAATTGGGGAATCAATTCAGAGGCAAT  2100
  V  L  N  F  Q  S  E  S  E  R  T  N  W  G  I  N  S  E  A  I

CAACAATTCTTGTCGCGAGAACAACATTTTGATGGTTAACTACCCTATACGgtaggatat  2160
  N  N  S  C  R  E  N  N  I  L  M  V  N  Y  P  I  R
```

V

```
ggttttctcttttgtaaaagatcatagtgaaaaacctatttgactcttgtagcataata  2220
tatttgtagttgttattggcatgctaagcttatcagtctaatattcatcaaactcaagtg  2280
gcatttctgatttaatttttttcagAGAGGTTGATTCAATGGACCTGAGGAAGAAGCTTTC  2340
                          E  V  D  S  M  D  L  R  K  K  L  S

TTTCTGTGTTGGTCTTCTACTGCGGCTTATAAGGAAGAACTACCGCATATATGTGACTTG  2400
  F  C  V  G  L  L  L  R  L  I  R  K  N  Y  R  I  Y  V  T  C

TACCACTGGATATGATAGATCACCAGCATGTGTGATTGCATATCTACATTGGGTGCAGGA  2460
  T  T  G  Y  D  R  S  P  A  C  V  I  A  Y  L  H  W  V  Q  D

TACGCCTCTCCATATTGCTCACAAGTTCATCACTGGTTTGCACTCCTGTAGACCTGACAG  2520
  T  P  L  H  I  A  H  K  F  I  T  G  L  H  S  C  R  P  D  R
```

VI

```
gtgttgttagttgacataattttcttaattgactgggcttgatccatctgcataagcagt  2580
gttgatccttaatgtccacatttgtctgaatatgcagAGCTGCAATTGTGTGGGCAACAT  2640
                                    A  A  I  V  W  A  T

GGGATCTCATTGCACTAGTTGAAAACGGAAGACATGATGGTACTCCCACACATTCAGTAT  2700
W  D  L  I  A  L  V  E  N  G  R  H  D  G  T  P  T  H  S  V

GCTTTGTTTGGAACAGTGGTCGGGAGGtacattttgcctaatgcggaatgcatcaattag  2760
C  F  V  W  N  S  G  R  E
```

VII

```
cagtgtacttttgtctaaaaaatcactattttagcttgatattcagtaaactgatgaag  2820
ttgaattcacaatttttttctaggGTGAGGATGTGGAATTGGTGGGGGATTTTACAAGTA  2880
                       G  E  D  V  E  L  V  G  D  F  T  S

ACTGGAAAGACAAAGTAAAGTGTGACCACAAAGGTGGGTCAAGATATGAAGCTGAAATTC  2940
N  W  K  D  K  V  K  C  D  H  K  G  G  S  R  Y  E  A  E  I

GACTTCGACATGGGAAGTAagtagtcagtccttcgatcttatattttctaacccagaagc  3000
R  L  R  H  G  K  Y
```

VIII

```
atgtaacattcttatatgccttgggctcaacttcatcattggctgaaaaggataaggaac  3060
atcatcttccttccatttctccatttctgatcccttcactgctgatcaagagttatctgc  3120
aactaacaacattttattcaggtaCTATTACAAATTCATAGCAGGGGGCCAGTGGAGGCA  3180
                       Y  Y  K  F  I  A  G  G  Q  W  R  H

CTCGACTTCATTGCCAACAGAGACTGATGAACATGGGAATGTCAACAATGTTATCAGGGT  3240
  S  T  S  L  P  T  E  T  D  E  H  G  N  V  N  N  V  I  R  V

TGGTGACATCGCTCGTATTCGGCCTGCTCCCAGCCAACTGCAGATAAGGGtatgcatccc  3300
  G  D  I  A  R  I  R  P  A  P  S  Q  L  Q  I  R
```

IX

```
caattactcacttgctctagtgaccccaactgagaatatttgtgtcttagcccacctgcc  3360
tatatgtacgccagacattcattctcctcgctccattgtatcattcatgcatatcaattt  3420
ttataggtatatcttattcgatttttttaataggtataggtttaacatataaaaaaaatt  3480
gtgttttctcccatgtcaggACCCAACTGTTGTCAAGGTCATAGAGAGGGCACTAACTGA  3540
                    D  P  T  V  V  K  V  I  E  R  A  L  T  E

GGACGAGCGATTCTTACTGGCCTTCGCTGCACGCCGCATGGCATTTGCAATCTGCCCAAT  3600
  D  E  R  F  L  L  A  F  A  A  R  R  M  A  F  A  I  C  P  I

CAGATTGTCTCCCAAGCAATGAACACATCATAATACCTACTGAGATGAACTGCAGCGAAA  3660
  R  L  S  P  K  Q  *
```