2009

Structure and mechanism of protein tyrosine phosphatase-like phytases

Gruninger, Robert J.

Lethbridge, Alta. : University of Lethbridge, Dept. of Chemistry and Biochemistry, c2009

STRUCTURE AND MECHANISM OF PROTEIN TYROSINE PHOSPHATASE-LIKE PHYTASES

ROBERT J. GRUNINGER
Bachelor of Science, University of Lethbridge, 2005

A Thesis
Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfilment of the
Requirements for the Degree

DOCTOR OF PHILOSOPHY

Department of Chemistry and Biochemistry
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

Abstract

The structure and mechanism of the Protein Tyrosine Phosphatase-like Phytases (PTPLPs) from *Selenomonas ruminantium* (PhyAsr) and *Mitsuokella multacida* (PhyAmm) were investigated using a combination of enzyme kinetics, site-directed mutagenesis, and X-ray crystallography. I show that PTPLPs use a classical protein tyrosine phosphatase catalytic mechanism and adopt a core PTP fold. Several unique structural features of PTPLPs confer specificity for inositol phosphates. The effect of ionic strength and oxidation on the kinetics and structure of PTPLPs was investigated. The structural consequences of reversible and irreversible oxidation on PTPLPs and PTPs are compared and discussed. We determine the structural basis of substrate specificity in PTPLPs and propose a novel reaction mechanism for the hydrolysis of inositol polyphosphates by PTPLPs. Finally, the structure and function of a unique tandemly repeated phytase has been determined. We show that the active sites of the tandem repeat possess significantly different specificities for inositol polyphosphate.

Acknowledgements

I would like to thank my co-supervisors Dr. Brent Selinger and Dr. Steven Mosimann for all of their guidance, support and encouragement. I would also like to thank the members of my supervisory committee Dr. Marc Roussel and Dr. Hans-Joachim Wieden and my external examiner Dr. Zongchao Jia for their helpful comments and thought provoking discussions.

I would like to acknowledge financial support from the Natural Sciences and Engineering Research Council of Canada, Alberta Ingenuity and the School of Graduate Studies at the University of Lethbridge.

X-ray diffraction data was collected at beamline 8.3.1 of the Advanced Light Source (ALS) at Lawrence Berkeley Lab, under an agreement with the Alberta Synchrotron Institute (ASI). The ALS is operated by the Department of Energy and supported by the National Institutes of Health. Beamline 8.3.1 was funded by the National Science Foundation, the University of California and Henry Wheeler. The ASI synchrotron access program is supported by grants from the Alberta Science and Research Authority and AHFMR.

I would like to thank the members of the Selinger and Mosimann lab both past and present for their support and friendship. Finally I would like to thank my wife, family and friends for their love, support, and encouragement in all aspects of my life.

Table of contents

List of Tables:

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of figures:

Chapter 1

Chapter 2

Chapter 3

LIST OF ABBREVIATIONS

5-IAF	5-iodoactamidofluorescein
A. niger	*Aspergillus niger*
ALS	Advanced Light Source
ASU	Asymmetric unit
ATP	Adenosine triphosphate
βME	β -mercaptoethanol
BPPs	β -propeller phytases
CCD	Charge-coupled device
CDC25B	Cell division cycle 25 homolog B
CF	Correction factor
c_i	concentration of i
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
Ext	Extinction coefficient
G3P	Glyceroinositol(3,4,5)P3
GA	General acid
HAPs	Histidine acid phosphatases
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IAA	Iodoacetate
$InsP_6$	*Myo*-inositol 1,2,3,4,5,6 hexa*kis*phosphate

InsPs	Inositol phosphates
IPPase	Inositol polyphosphate phosphatase
IPPs	Inositol polyphosphates
IPTG	isopropyl-β-D-thiogalactopyranoside
LAR	Leukocyte Antigen Related
M. elsdenii	*Megashpaere elsdenii*
M. multacida	*Mitsuokella multacida*
MAD	Multi-wavelength anomalous dispersion
MAPK	Mitogen-activated protein kinase
MIHS	*Myo*-inositol hexakissulfate
MTMR2	Myotubularin-related 2
NCBI	National Center for Biotechnology Information
ND	Not detectable
NMR	Nuclear magnetic resonance
ORF	Open reading frame
P-loop	Phosphate binding loop
PAPs	Purple acid phosphatases
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
PhyA	Phytase A
PhyAmm	*Mitsuokella multacida* PhyA
PhyAsr	*Selenomonas ruminantium* PhyA

PISA	Protein Interfaces, Surfaces and Assemblies
pNPP	*p*-nitrophenyl phosphate
PTEN	Phosphatase and tensin homolog
PtInsPs	Phosphatidylinositol phosphates
PtIns(4,5)P_2	Phosphatidylinositol 4,5-bisphosphate
PtIns(1,4,5)P_3	Phosphatidylinositol 1,4,5-triphosphate
PTPs	Protein tyrosine phosphatases
PTPLPs	Protein tyrosine phosphatase-like phytases
pv.	Pathovar
Pyst1	MAP-kinase phosphatase 3
RMSD	Root mean square deviation
RPTPs	Receptor protein tyrosine phosphatases
RTX	Repeat in toxin
S. lacticifex	*Selenomonas lacticifex*
S. ruminanitum	*Selenomonas ruminantium*
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
Subsp.	Subspecies
Sp.	Species
TCA	Trichloroacetic acid
TR	Tandem repeat
TRI	Tandem repeat interface
TRIS	Tris(hydroxymethyl)aminomethane
U	Unit
v/v	volume/volume

w/v	weight/volume
Yop51	*Yersinia* protein tyrosine phosphatase
Z_i	charge of species i

Chapter 1. Literature Review

1.1 Inositol polyphosphates

Inositol phosphates (InsPs) were first recognized in 1919 (Irvine and Schell 2001); however, it was not until 1983 that a biological function was attributed to them (Streb et al. 1983). Inositol phosphates are universally conserved in eukaryotes and function as second messengers in signal transduction pathways. The related membrane anchored phosphoinositides, phosphatidylinositol phosphates (PtInsPs) consist of a phosphatidic acid linked to an inositol phosphate head group at the O1 position. Like inositol phosphates, phosphatidylinositol phosphates function in a variety of signalling pathways. Inositol phosphates, and phosphatidylinositol phosphates, are not found in bacteria or archaea (Michell 2008). Despite this, many prokaryotes produce inositol phosphate degrading enzymes that serve diverse functions including phosphate scavenging and pathogenesis (Norris et al. 1998; Chatterjee et al. 2003).

The best characterized phosphoinositide pathway is the Ca^{2+} mobilizing pathway involving phosphatidylinositol 4,5-bisphosphate (PtIns(4,5)P_2) and inositol 1,4,5-triphosphate (Ins(1,4,5)P_3). In response to several hormonal stimuli, phospholipase C hydrolyzes PtIns(4,5)P_2 to generate the second messenger Ins(1,4,5)P_3 which in turn leads to the release of Ca^{2+} from the endoplasmic reticulum (Irvine and Schell 2001). The released Ca^{2+} stimulates a wide range of processes through the action of calmodulin and related Ca^{2+}-binding proteins (Irvine and Schell 2001). Numerous inositol and phosphatidylinositol phosphates have been identified and implicated in many important biological process including cell survival, cytoskeleton organization, vesicular

trafficking, regulation of ion channels and pumps, and signal transduction (Boss et al. 2006).

1.1.1 *Myo*-Inositol 1,2,3,4,5,6 hexa*kis*phosphate

Myo-inositol 1,2,3,4,5,6 hexa*kis*phosphate ($InsP_6$) is the most abundant cellular inositol phosphate (Raboy 2003) and is commonly referred to as "phytic acid" or "phytate" when complexed with cations. $InsP_6$ is ubiquitous in eukaryotic cells and functions in a wide range of biological processes. Initially, $InsP_6$ was thought to function solely as a storage molecule for phosphate, cations, amino acids and *myo*-inositol. This was based on the observation that during development and germination, plants used $InsP_6$ as a source of these essential nutrients. More recently, it has become clear that the biological role of $InsP_6$ is much more diverse and it has been found to be important in several essential processes including DNA repair (Hanakahi et al. 2000), RNA processing (Macbeth et al. 2005), mRNA export (York et al. 1999), plant development (Tan et al. 2007), apoptosis (Majerus et al. 2008), and pathogenicity (Chatterjee et al. 2003; Lupardus et al. 2008). The importance of $InsP_6$ is exemplified by the observation that deletion of enzymes involved in $InsP_6$ biosynthesis has a lethal phenotype in mouse embryos (Frederick et al. 2005; Verbsky et al. 2005).

$InsP_6$ accounts for up to 90% of the phosphorus content in the grains and seeds that make up the high concentrate diets used in intensive livestock operations. $InsP_6$ is recalcitrant to digestion by monogastric animals (poultry, swine, fish) and is excreted in animal waste where it is subsequently broken down in the environment. This presents a number of problems: 1) livestock producers must supplement animal diets with expensive

inorganic phosphorus, 2) the breakdown of the excreted phytate releases phosphorus into the water system and contributes to eutrophication of aquatic environments, and 3) phytate forms complexes with essential nutrients (i.e. cations and amino acids) which negatively impacts livestock growth (Rao et al. 2009). The addition of $InsP_6$ degrading enzymes, such as phytase, to livestock feed has been successfully used since 1991 to ameliorate these problems. Supplementation of feed with phytase has been found to reduce phosphate excretion by as much as 50% (Rao et al. 2009).

In contrast to monogastric animals, ruminants can break down $InsP_6$. Ruminant animals are characterized by a unique, complex multi-chamber digestive system. The rumen is a large anaerobic chamber that is inhabited by a variety of microbes that secrete a range of hydrolytic enzymes to break down the plant material consumed by the animal. The ability of ruminants to breakdown $InsP_6$ was first reported in 1912 (Fingerling 1912) but the microbial basis of this enzyme activity has only recently been extensively explored (Yanke et al. 1998).

1.1.2 Structure and chemistry of *myo*-inositol 1,2,3,4,5,6 hexa*kis*phosphate

$InsP_6$ is a highly phosphorylated derivative of the cyclitol *myo*-inositol. The six phosphate groups possess 12 sites that can be protonated. Six protons have pKa values between 1.1 and 2.1, three protons have pKa's in the range of 6.0 to 7.6, and the last three protons have pKa's between 9.2 and 9.6. The precise pKa values have been found to depend on the counter-ion present in solution (Isbrandt and Oertel 1980). At physiological pH $InsP_6$ will have a charge of approximately -6 to -9. The large negative charge of this molecule makes it a strong chelator of cations and proteins. Additionally,

the negative charge on this molecule is important for many of its biological functions (Macbeth et al. 2005; Tan et al. 2007; Lupardus et al. 2008).

Myo-inositol nomenclature has long been a source of confusion largely due to problems encountered between chemists and biochemists assigning a consistent numbering system to these molecules (Irvine and Schell 2001). An easy method for determining the absolute configuration of *myo*-inositols is the Arganoff "turtle" analogy (Irvine and Schell 2001). Arganoff recognized that a Haworth projection of the energetically favored 5-equatorial/1-axial *myo*-inositol in a chair conformation resembles a turtle with the axial O2 position representing the head, the two pro-chiral pairs O1/O3 and O4/O6 representing the flippers and O5 representing the tail (Figure 1.1). If one orients the ring such that the axial O2 group is on the left, and pointing up then if the numbering increases in a counter-clockwise fashion it is a D-isomer. The majority of biologically relevant inositol phosphates are D-isomers and as a result, biochemists and biological journals exclusively use the D-numbering system (Murthy 2006).

The structure of $InsP_6$ has been determined both in solid state, and in solution. X-ray crystallographic studies of $InsP_6$ salts revealed that this molecule adopts a 5-axial/1-equatorial conformation that is stabilized through interactions with cations (Blank et al. 1971). Subsequent solution NMR studies revealed that the $InsP_6$ ring undergoes a pH dependent flip (Isbrandt and Oertel 1980). $InsP_6$ adopts a 5-equatorial/1-axial conformation at pH values < 9.2 and a 5-axial/1-equatorial conformation at pH values > 9.6 (Isbrandt and Oertel 1980). This conformational change likely corresponds to the deprotonation of one or all of the three least acidic protons.

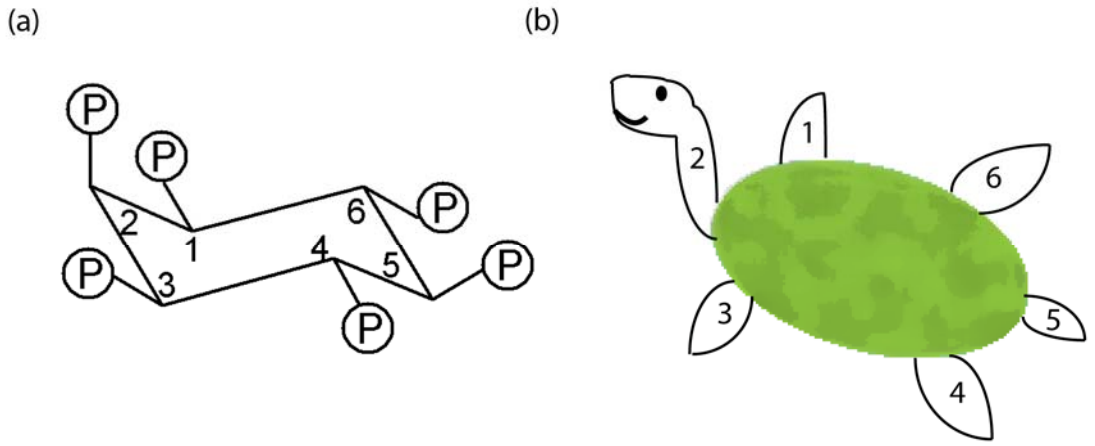

Figure 1.1. Arganof's Turtle. (a) Haworth projection of *myo*-inositol 1,2,3,4,5,6-hexa*kis*phosphate in the energetically favored chair conformation with five equatorial phosphates and one axial phosphate. For clarity, the phosphates are illustrated with the symbol Ⓟ. (b) The molecule resembles a turtle with the axial 2-position representing the head, the 1, 3, 4, and 6 positions representing the flippers and the 5-position representing the tail. The numbering shown inside the ring follows the nomenclature for the D stereoisomer.

1.1.3 Inositol pyrophosphate

$InsP_6$ and $Ins(1,3,4,5,6)P_5$ can be further phosphorylated to generate the inositol pyrophosphate containing molecules PP-$InsP_4$, $InsP_7$, and $InsP_8$ (Bennett et al. 2006). At least five unique inositol pyrophosphates have been identified to date. The enzymes that generate these molecules are highly conserved in eukaryotes and three kinases that phosphorylate $InsP_6$ (IP6K1, IP6K2, IP6K3) have been identified (Bennett et al. 2006). The pyrophosphate bonds in these molecules have a free energy of hydrolysis that is similar to that in ATP (Bennett et al. 2006). The high energy in these bonds confers $InsP_7$ with the ability to phosphorylate specific proteins in an ATP and enzyme independent manner and suggests that these molecules are involved in a novel intracellular signaling mechanism (Saiardi et al. 2004). The levels of $InsP_7$ change

rapidly at different stages of the cell cycle (Bennett et al. 2006), possibly due to its involvement in regulating this process. In addition to directly phosphorylating proteins, inositol pyrophosphates have been shown to be important in many essential processes such as endocytosis (Saiardi et al. 2002), exocytosis (Illies et al. 2007), regulation of telomere length and cell death (Saiardi et al. 2005).

1.2 Phytases

Enzymes that degrade $InsP_6$ are generically referred to as phytases and have been identified in prokaryotes, protists, fungi, animals and plants (Mullaney et al. 2000). The most extensively studied phytases are of microbial and fungal origin. This is due to the extensive use of these enzymes as agricultural feed supplements and the favorable economics of enzyme production in these systems. Until recently, little research had been done on the biological role of $InsP_6$ in animals and as such few phytases of animal origin have been identified. This will likely change in the coming years as the central role that $InsP_6$ plays in several essential cellular processes has renewed interest in $InsP_6$ and the enzymes involved in $InsP_6$ metabolism. Two dual function phytases have recently been described: 1) a highly conserved eukaryotic inositol enzyme with both inositol polyphosphatase and phytase domains (Mulugu et al. 2007), and 2) a tandemly repeated β-propeller phytase in *Shwanella oneidensis* (Cheng and Lim 2006). Another recently discovered role for $InsP_6$ degrading enzymes is the discovery of a phytase that functions as an important virulence factor in the plant pathogen *Xanthomonas oryzae* pv. *oryzae* (Chatterjee et al. 2003). $InsP_6$ has been found to play a role in pathogen resistance in plants and it is likely that the *Xanthomonas* phytase has been adapted to interfere with

this process (Murphy et al. 2008). $InsP_6$ has also been found to play a role in the virulence of the human pathogen *Vibrio cholera* by allosterically activating the protease activity of the secreted toxin RTX (Lupardus et al. 2008).

1.2.1 $InsP_6$ degradation

Phytases catalyze the stepwise removal of phosphates to generate lower inositol phosphates (i.e. inositol phosphates with less than six phosphates) in a highly ordered, specific manner (Konietzny and Greiner 2002). The intermediates generated during the breakdown of $InsP_6$ are released from the enzyme and serve as substrate for subsequent cycles of hydrolysis. One exception is the phytase expressed by *Bacillus subtilis* which sequentially removes both the 6 and 4 phosphate groups before releasing $Ins(1,2,3,5)P_4$ (Konietzny and Greiner 2002). Based on the initial position of hydrolysis, the Enzyme Nomenclature Committee of the International Union of Biochemistry recognizes three types of phytase namely, 3-phytases (EC 3.1.3.8), 4 -phytases (3.1.3.26), and 5-phytases (EC 3.1.3.72). Recently, several phytases that can hydrolyze phytate at more than one site on the phytate ring have been characterized. These include a D-3/D-4 phytase from *Klebsiella terrigena* (Greiner and Carlsson 2006) and *Megasphaera elsdeni* (Puhl et al. 2009), and D-3/D-6 phytases from the basidiomycete fungi *Ceriporia* sp. and *Trametes pubescens* (Lassen et al. 2001).

Although phytases initiate the hydrolysis of $InsP_6$ at a limited number of positions, the identity of the intermediates generated during the degradation pathway vary greatly. Many lower phosphorylated inositol phosphates have important physiological roles and hence they are of interest in the pharmaceutical industry. These products can be

used in both kinetic and physiological studies. Both chemical and enzymatic approaches have been used to generated lower inositol phosphates; however, the enzymatic approach is far superior both in its simplicity, yield, and ability to generate isomerically pure products (Haefner et al. 2005). This has been exploited as a means of enzymatically generating a variety of isomerically pure inositol phosphates. Unfortunately, to date this approach has had limited application due to difficulties purifying large quantities of pure products (Konietzny and Greiner 2002).

1.2.2 Classes of phytase

To date four classes of phytases have been described including: (1) histidine acid phosphatases (HAPs), (2) β-propeller phytases (BPPs), (3) purple acid phosphatases (PAPs) and (4) protein tyrosine phosphatase-like phytases (PTPLPs). Although these enzymes all degrade $InsP_6$ to lower inositol phosphates and inorganic phosphate they are structurally and mechanistically diverse.

1.2.2.1 Histidine acid phosphatases

The majority of phytases characterized to date are members of the histidine acid phosphatase (HAP) family (EC 3.1.3.2). These enzymes are characterized by a highly conserved RHGXRXP active-site signature sequence, and a conserved HD sequence that is located near the C-terminus. The reaction mechanism of HAPs has been extensively studied using site-directed mutagenesis, enzyme kinetics and X-ray crystallography. HAPs utilize a two step catalytic mechanism involving nucleophilic attack of the scissile phosphate by the histidine in the RHGXRXP signature sequence, and protonation of the

leaving group by the conserved aspartic acid (Ostanin et al. 1992; Ostanin and Van Etten 1993). The conserved arginine in the RHG tripeptide is involved in binding the scissile phosphate and positioning it for nucleophilic attack (Lim et al. 2000). The first step in the catalytic mechanism results in the formation of a phospho-histidine intermediate (Xiang et al. 2004). In the second step of the reaction, the conserved aspartic acid acts as a general base to abstract a proton from a water molecule that hydrolyzes the covalent intermediate (Liu et al. 2004).

The first three dimensional structure of a phytate degrading HAP was determined by Kostrewa and colleagues in 1997 (Kostrewa et al. 1997). Since then, structures of phytate degrading HAPs from *E. coli* (Lim et al. 2000; Lee et al. 2003), *Aspergillus fumigatus* (Liu et al. 2004; Xiang et al. 2004), and *Debaryomyces castellii* (Ragon et al. 2009) have been determined. All HAPs adopt a fold consisting of a conserved α/β catalytic domain and a variable α domain. Some HAPs also contain small additional domains involved in substrate specificity (Lee et al. 2003). The positioning of the catalytic residues in the catalytic domain is conserved between HAPs regardless of the enzyme's specificity. In contrast, the α domain is involved in conferring substrate specificity to these enzymes and as such, the structure and orientation of this domain is variable within this phosphatase family (Kostrewa et al. 1999).

Structures of HAPs in complex with substrate (Lim et al. 2000), the phospho-histidine intermediate (Xiang et al. 2004), and in complex with product (Liu et al. 2004) have been determined. These structures provide the unique opportunity to structurally understand all stages of the reaction mechanism of this enzyme family. $InsP_6$ binds deep in an electropositive cleft, making extensive contacts with the protein.

Substrate binding in the *E. coli* phytase is accompanied by a conformational change in an active-site loop that clamps down on the substrate (Lim et al. 2000) (Figure 1.2). The conformational change and contacts to the base of the ligand (portion of ligand closest to the nucleophile) are important for stabilizing the enzyme-substrate complex and positioning the scissile phosphate for nucleophilic attack. The phospho-histidine intermediate is stabilized by a network of interactions to several conserved polar residues in the active-site (Xiang et al. 2004). Hydrolysis of the phosphoamide bond is accompanied by conformational changes in several active-site residues to facilitate product release (Liu et al. 2004). Based on the identification of multiple phosphate binding sites in and around the active-site, it was postulated that the phosphate is passed from one binding site to another and then released from the enzyme (Liu et al. 2004).

1.2.2.2 β-propeller phytases

β-propeller phytases (BPPs, EC 3.1.3.8) were initially isolated from *Bacillus* species and much of what is known about the structure and function of this class of enzyme is based on these enzymes (Ha et al. 2000; Kerovuo et al. 2000; Shin et al. 2001; Oh et al. 2006). The initial discovery of BPPs in a limited number of *Bacillus* species gave the impression that they were not as widespread as HAPs; however, recent analysis of BPP diversity indicates that BPPs are actually the major class of phytate degrading enzymes in nature. A recent survey of the distribution of phytate degrading enzymes in marine environments revealed that the vast majority of phytases present in aquatic habitats belong to this phytase class (Lim et al. 2007). Another study examining the diversity of phytases in the intestine of grass carp identified a number of BPPs mainly

from unculturable bacteria (Huang et al. 2009). Interestingly, the phytase important for *Xanthomonas oryzae* pv *oryzae* virulence shows similarities to the *Bacillus* enzyme suggesting that it may be a BPP (Chatterjee et al. 2003).

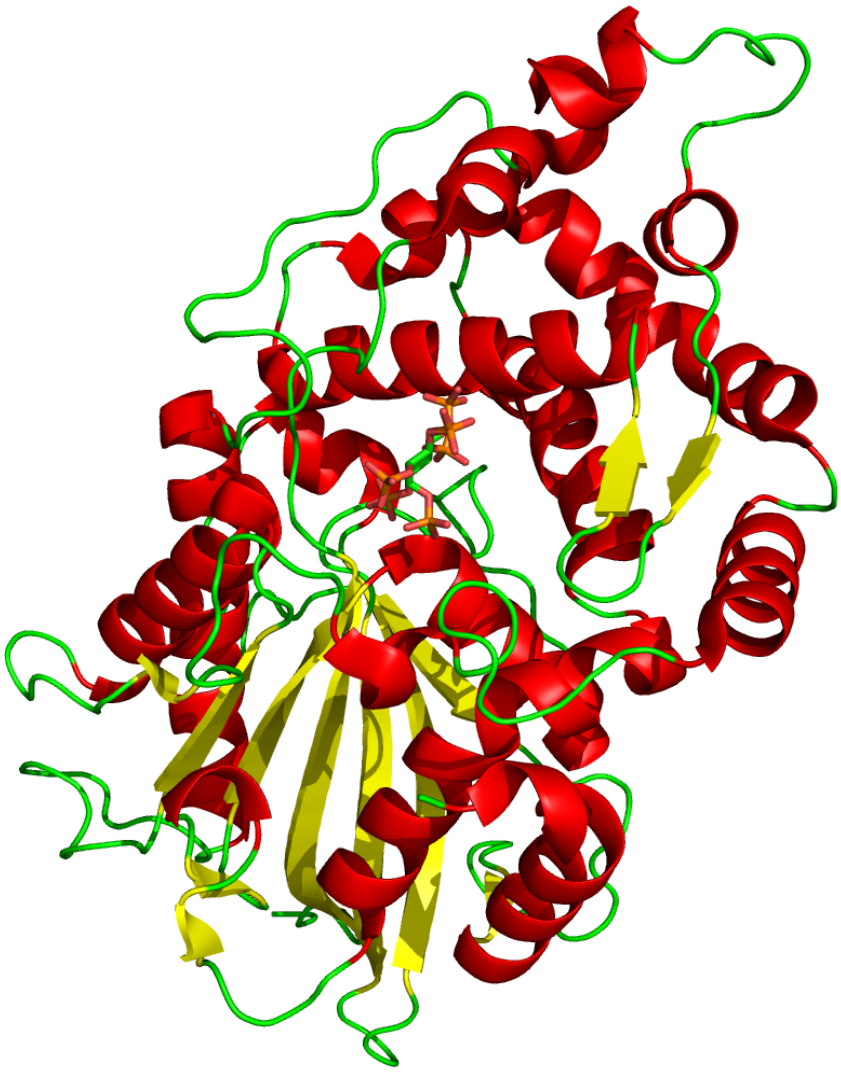

Figure 1.2: Structure of the phytate degrading histidine acid phosphatase from *E. coli* in complex with $InsP_6$ (PDB 1DKP). The ligand is bound in a cleft formed between a conserved α/β domain and a variable α domain.

BPPs adopt a β-propeller fold that is found in a wide range of functionally diverse proteins. The thermostable BPP from *Bacillus amyloliquefaciens* adopts a six bladed β-

propeller fold with five four-stranded and one five-stranded antiparallel blades (Figure 1.3). There is an electronegative, solvent accessible central channel that binds seven Ca^{2+} ions. BPPs have been found to be dependent on Ca^{2+} for both catalytic activity and substrate binding (Oh et al. 2006). Several divalent cations can substitute for Ca^{2+} however none of them are as effective (Oh et al. 2001). Mutation of residues involved in coordinating the cations results in complete loss of catalytic activity (Oh et al. 2001). The Ca^{2+} ions serve several functions, including activating a water molecule, coordinating the scissile phosphate, and stabilizing the negative charge that develops on the transition state. Several lines of evidence point to the direct attack of the phosphate by a hydroxyl, including the presence of a water molecule in close proximity to the scissile phosphate, the lack of a suitable nucleophilic side-chain near the scissile phosphate, and inhibition of catalysis by fluoride ions (Shin et al. 2001). A structure of a BPP in complex with ligand has not been solved, however the presence of ordered phosphates in the active site provided a guide with which to model $InsP_6$ (Shin et al. 2001). BPPs have two binding sites, a cleavage site and an affinity site that enhances substrate binding. Only substrates that simultaneously fill both binding sites are hydrolyzed by BPPs which explains why these enzymes can only remove three phosphates from $InsP_6$ (Kerovuo et al. 2000).

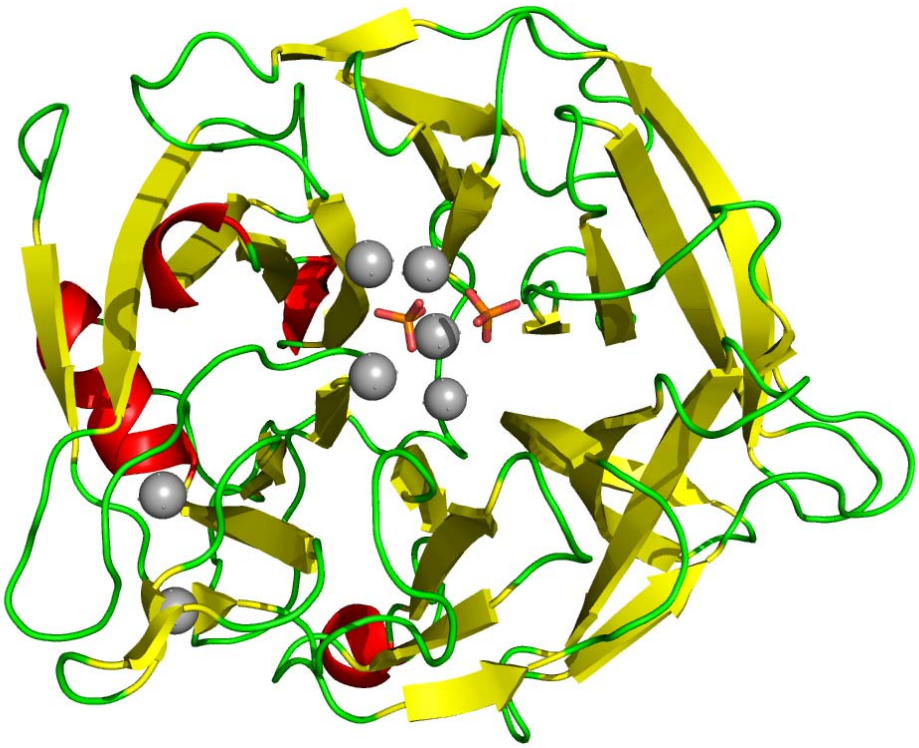

Figure 1.3: Structure of the β propeller phytase from *B. subtilus* in complex with Ca^{2+} and phosphate (PDB 1H6L). Ca^{2+} ions are shown as grey spheres and phosphate ions are shown as sticks.

1.2.2.3 Purple acid phosphatases

Phytate degrading enzymes with homology to purple acid phosphatases (PAPs) have been purified from both germinating *Glycine max* (soybean) (Hegeman and Grabau 2001) and *A. niger* (Ullah and Cummins 1988). The PAP from soybean is the only member of this enzyme class that has been shown to have significant levels of activity against $InsP_6$. The *Aspergillus niger* PAP displays only weak catalytic activity against this substrate (Mullaney and Ullah 2003). PAPs are a class of metalloenzyme and their catalytic mechanism and structural features have been extensively studied. These enzymes possess a binuclear metal centre that can be either Fe^{3+}-Fe^{2+}, Fe^{3+}-Mn^{2+}, or Fe^{3+}-Zn^{2+}. The catalytic mechanism involves the direct attack of the scissile phosphate by an

activated water molecule coordinated by the metal centre. PAPs exist as a homodimer and each monomer is composed of two domains: a small N-terminal antiparallel β sandwich and a larger C-terminal α+β domain (Figure 1.4). The C-terminal domain is characterized by two large mixed central β-sheets that form a β-sandwich. Two α-helices from each monomer interact to form the homodimer interface.

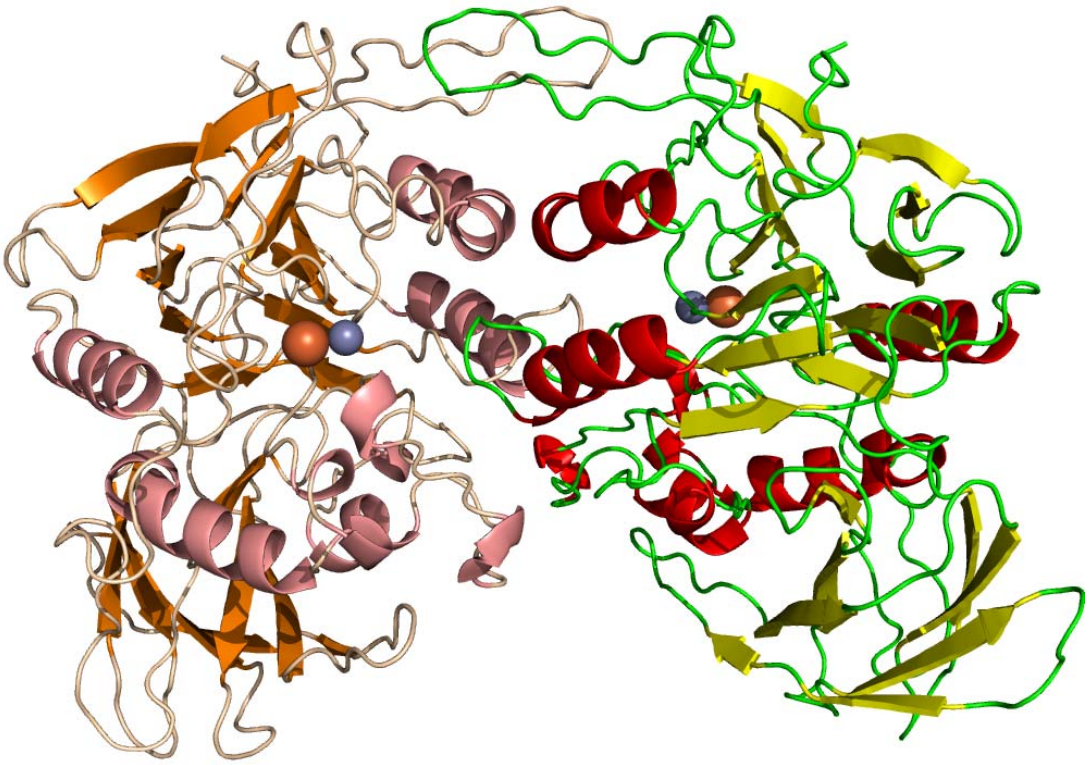

Figure 1.4: Structure of the purple acid phosphatase homodimer from *Phaseolus vulgaris* (PDB 1KBP). The Fe^{3+} and Zn^{2+} atoms in the binuclear center are shown as orange and blue spheres, respectively. The protein is oriented so that the N-terminal β domain is at the bottom of the homodimer, and the C-terminal α+β domain is at the top of the homodimer.

1.3 Protein tyrosine phosphatase-like phytases (PTPLPs)

A class of phytate degrading enzyme related to protein tyrosine phosphatases (PTPs) has recently been described (Yanke et al. 1998; Chu et al. 2004; Puhl et al. 2007; Puhl et al. 2008a; Puhl et al. 2008b; Puhl et al. 2009). These enzymes are characterized by the $CX_5R(S/T)$ PTP active-site signature sequence. This sequence forms the loop at the base of the PTP active-site and is important for coordinating the scissile phosphate for nucleophilic attack. Site-directed mutagenesis and kinetics studies have shown that PTPLPs use a PTP catalytic mechanism (Puhl et al. 2007). PhyAsr adopts an α+β PTP fold, and has a small β-domain that is unique to this class of enzyme (Chu et al. 2004) (Figure 1.5). The β-domain is implicated in the unique substrate specificity of these enzymes (Chu et al. 2004; Puhl et al. 2007).

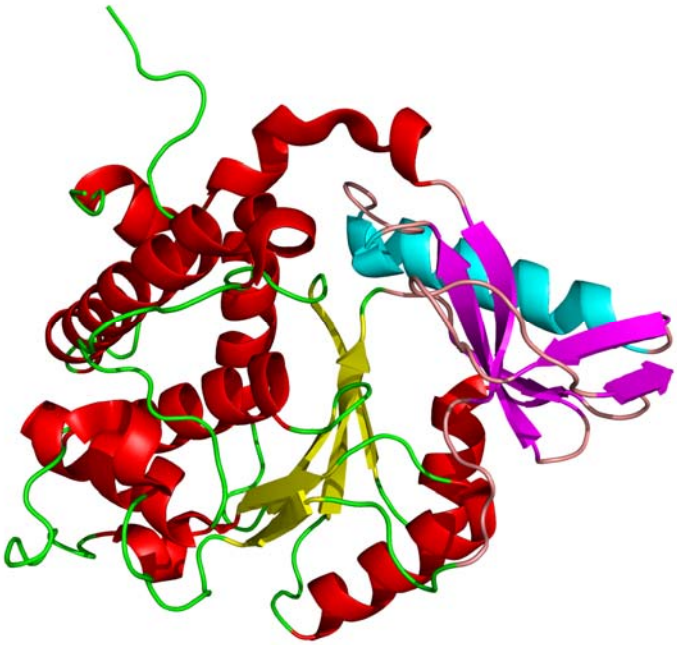

Figure 1.5: Structure of the PTPLP from *S. ruminantium* (2B4P). The helices and sheets of the core PTP domain are shown in red and yellow, respectively. The helices and sheets of the unique β-domain are shown in cyan and magenta, respectively.

1.3.1 Protein tyrosine phosphatase-like phytase diversity

PTPLPs were first discovered while screening rumen microbes for phytase activity (Yanke et al. 1998). This study identified high levels of phytase activity in *S. ruminantium*, *Mitsuokella multacida*, *Megasphera elsdenii*, and *Prevotella ruminicola*. A subsequent study identified a number of PTPLP genes from *S. ruminantium*, *S. ruminantium sp. lactillytica*, *S. lacticifex*, *M. multacida*, and *M. elsdenii* (Nakashima et al. 2007). The results of these two studies indicate that PTPLPs may be responsible for the phytase activity in the rumen. To my knowledge, the presence of HAPs, BPPs, or PAPs in the rumen environment has not been investigated. PTPLPs have also been identified in a diverse range of bacteria including the pathogens *Clostridium* spp., *Legionella pneumophila*, *Pseudomonas syringae*, *Xanthomonas campestris* and the predatory bacterium *Bdellovibrio bacteriovorus* (Lim et al. 2007).

While the function of PTPLPs is not well understood, the protein in *Pseudomonas syringae* has been shown to be involved in pathogenesis (Bretz et al. 2003; Espinosa et al. 2003). It is possible that the PTPLPs in the other pathogenic bacteria may serve a similar function. Further studies are required to understand the biological function of PTPLPs and whether these proteins function as phytases *in vivo* and/or *in vitro*.

1.3.2 Biochemical properties of PTPLPs

The biochemical properties of PTPLPs from *S. ruminantium* (Puhl et al. 2007), *S. ruminantium* subsp. *lactolytica* (Puhl et al. 2008b), *S. lacticifex* (Puhl et al. 2008a) and *M. elsdenii* (Puhl et al. 2009) have been characterized (Table 1.1). Similar to HAPs and PAPs, PTPLPs have optimal enzyme activity at acidic pH's. This preference for acidic

pH is a result of the PTPLP catalytic mechanism (discussed below). The optimal pH for BPPs is generally between 7.0 and 8.0 (Konietzny and Greiner 2002; Rao et al. 2009). The temperature optima of phytases vary from 25 to 77 °C with most displaying an optimal temperature between 50 and 65 °C (Konietzny and Greiner 2002; Rao et al. 2009). In general, phytases from plants have lower temperature optima whereas microbial phytases have higher temperature optima (Konietzny and Greiner 2002). This is likely due to differences in the natural environments of plant and microbial phytases. Most HAPs characterized to date show a rather broad substrate specificity and low specific activity (Konietzny and Greiner 2002). In contrast, BPPs are highly specific for Ca^{2+}-$InsP_6$, due to the role that Ca^{2+} has in substrate binding and catalysis. PTPLPs can be categorized as having high activity and high specificity for $InsP_6$, or low activity and broad specificity.

Table 1.1: Biochemical properties of characterized PTPLPs.

	pH optimum	Temperature optimum (°C)	Phytase activity[A]	Specificity	k_{cat} (s^{-1})	K_m[B] (μM)	$InsP_6$[C] hydrolysis
PhyAsr	5.0	50	High	High	264	425	3-1-6-5-4[D]
PhyAsrl	4.5	55	Low	Broad	65	7	5-4-6-3-1[E]
PhyAsl	4.5	40	High	High	256	309	3-4-5-6-1[F]
PhyBsl	4.5	37	Low	Broad	18	582	3-4-5-6-1[F]
PhyAme	5.0	60	Medium	High	122	64	3-4-5-6-1[G] 4-5-6-3-1[G]

A. Specific activity (Units/mg), Low < 100, Medium > 100 and < 300, High > 300
B. K_m for $InsP_6$
C. numbers indicate the identity and order that phosphates are removed
D. Puhl et al. 2007
E. Puhl et al. 2008b
F. Puhl et al. 2008a
G. Puhl et al. 2009

Interestingly, PhyAsrl shows higher catalytic activity against ATP than $InsP_6$ (Puhl et al. 2008b). Most PTPLPs preferentially initiate $InsP_6$ hydrolysis at the 3-

position however the identity of the lower order inositol phosphates generated during the degradation pathway vary greatly. In all cases, PTPLPs do not remove the 2-position (Puhl et al. 2007; Puhl et al. 2008a; Puhl et al. 2008b; Puhl et al. 2009).

1.4 The protein tyrosine phosphatase (PTP) superfamily

PTPs are characterized by the presence of a highly conserved $CX_5R(S/T)$ active-site signature sequence that is referred to as the 'phosphate binding loop' or 'P-loop'. PTPs are important regulators of essential cellular processes. In concert with tyrosine kinases, PTPs are involved in regulating the level of tyrosine phosphorylation in the cell (Alonso et al. 2004). Abnormal levels of tyrosine phosphorylation have been implicated in numerous human diseases, including cancer, immune deficiencies, diabetes and obesity (Zhang 2003; Alonso et al. 2004). There are 107 PTP genes in the human genome that can be categorized into four classes: (1) Classical cysteine based PTPs, (2) Low molecular weight PTPs, (3) CDC25's, and (4) Asp-based PTPs. The largest class is the cysteine based PTPs which includes non-receptor PTPs, the transmembrane PTPs, and dual specificity phosphatases (e.g. PTEN, Cdc14 and myotubularins) (Alonso et al. 2004). Despite the functional similarity of PTPs, the different classes evolved independently (Alonso et al. 2004). Unlike the class 1, 2, and 3 PTPs, Asp-based PTPs have been shown to use a catalytic aspartic acid and are dependent on a cation (Rayapureddi et al. 2003). This class has not been extensively studied though they have been shown to be important in ocular development in *Drosophila* (Rayapureddi et al. 2003).

1.4.1 PTP catalytic mechanism

The catalytic mechanism of PTPs has been extensively studied using site-directed mutagenesis, kinetic analysis, and X-ray crystallography. PTPs catalyze the hydrolysis of phospho-diester bonds using a two step mechanism involving the formation of a covalent phospho-enzyme intermediate (Guan and Dixon 1991; Zhang 2003). In the first step, the invariant cysteine in the P-loop nucleophilically attacks the scissile phosphate. At the same time, an invariant aspartic acid functions as a general acid (D181 in PTP1B and D223 in PhyAsr) by protonating the leaving group. The metaphosphate-like transition state is stabilized by the conserved active-site arginine, the main-chain amines of the P-loop, and the general acid (Hengge et al. 1995; Zhang 2003). The result of the first step of the reaction mechanism is the formation of a phospho-cysteine intermediate (Guan and Dixon 1991; Pannifer et al. 1998). The existence of this intermediate is well established and it has been isolated *in vitro* (Guan and Dixon 1991) and visualized crystallographically (Pannifer et al. 1998). The second step in the reaction mechanism is rate limiting and involves the hydrolysis of the phospho-enzyme intermediate (Zhang et al. 1994a; Zhang et al. 1995). The general acid from the first step serves as a general base in the second step by abstracting a proton from an activated water molecule that then hydrolyzes the phospho-enzyme intermediate (Wu and Zhang 1996; Zhang 2003). In PTP1B this water is coordinated by glutamine 262 from the so called Q-loop (Zhao et al. 1998). Mutation of Q262 prevents hydrolysis of the phospho-cysteine intermediate and enabled the structural characterization of the phospho-cysteinyl form of PTP1B (Pannifer et al. 1998). An equivalent residue is not present in PTPLPs, however there are

conserved water molecules in the active site near the general acid that could function in the hydrolysis of the covalent intermediate.

1.4.2 PTP structure

PTPs have been extensively characterized structurally and all members of the PTP superfamily, with the exception of CDC25s, adopt a conserved core structure. The CDC25 family shows no sequence similarity to other PTPs outside of the active-site signature sequence, and adopts a fold that is identical to rhodanese (Fauman et al. 1998). The core PTP fold is an α+β fold with a highly twisted central β-sheet flanked by a variable number of α-helices (Barford et al. 1994; Lee et al. 1999; Nam et al. 1999; Song et al. 2001) (Figure 1.6a). The active-site is an electropositive cleft with the P-loop located at the base of the pocket. The size of the active-site is variable and is important in substrate specificity (Yuvaniyama et al. 1996). Substrate binding in classical PTPs is accompanied by a large conformational shift in the general acid (WPD) loop (Jia et al. 1995) (Figure 1.6b). This loop goes from an 'open' catalytically incompetent conformation in the absence of ligand to a 'closed' conformation that positions the general acid for catalysis in the presence of ligand. This movement is not observed in other PTPs including PTEN (Lee et al. 1999), Cdc14 (Gray et al. 2003), and PhyAsr (Chu et al. 2004). These PTPs have pre-formed active-sites and altered substrate specificity relative to classical PTPs which suggests that this movement in the general acid loop, or lack thereof, contributes to substrate specificity. Interestingly, the CDC25 PTP family does not have a general acid loop (Fauman et al. 1998) and it has been suggested that a

residue on the protein substrate of CDC25's functions as the general acid (Rudolph 2002).

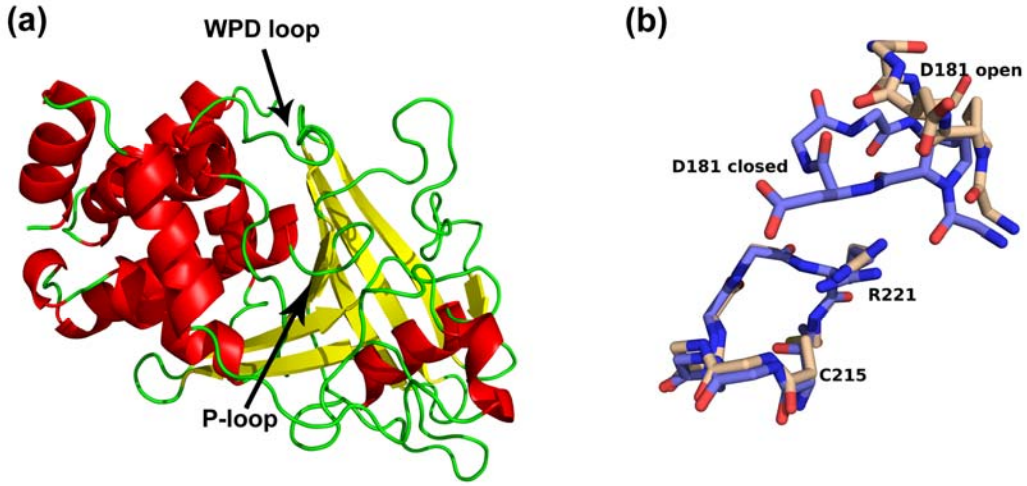

Figure 1.6 Classical PTP fold and ligand induced conformational change in the general acid loop (a) Structure of the classical PTP, PTP1B (PDB 1PTV). (b) Conformational change in the general acid (WPD) loop in PTP1B. In the absence of ligand the loop is in an 'open' conformation (Beige; PDB 2HNQ) and in the presence of ligand the loop adopts a 'closed' conformation (Light blue; PDB 1PTV). The general acid (D181), cysteine nucleophile (C215) and conserved arginine (R221) are shown as sticks.

1.4.3 Receptor PTPs

Classical PTPs can be divided into soluble non-receptor PTPs and transmembrane receptor PTPs (RPTPs). The human genome encodes 17 non-receptor PTPs and 21 receptor PTPs (Alonso et al. 2004). Receptor PTPs have a highly modular structure consisting of both cytoplasmic and extracellular domains. The extracellular domains function in a receptor-like fashion transmitting extracellular signals to the cytoplasmic PTP domain to stimulate or inhibit phosphatase activity (Alonso et al. 2004). Unlike other members of the PTP superfamily, RPTPs possess tandemly repeated phosphatase domains (Andersen et al. 2001). Interestingly, only the N-terminal (D1) repeat in many

RPTPs is catalytically active. The function of the inactive C-terminal (D2) repeat is not clear but it is thought to be involved in regulation and substrate recognition (Ng et al. 1995; Kashio et al. 1998; Blanchetot and den Hertog 2000; Yang et al. 2007). The structure of several full length RPTPs (with both D1 and D2 repeats) have been determined, including CD45 (Nam et al. 2005) (PDB 1YGU), LAR (Nam et al. 1999) (PDB 1LAR), RPTPσ (Almo et al. 2007) (PDB 2FH7), RPTPγ (Barr et al. 2009) (PDB 2NLK), and RPTPε (PDB 2JJD). The orientation of the dimer is the same in all of the RPTPs. The active-sites are orientated ~ 90˚ relative to one another with a six residue linker located at the interface between the repeats and oriented almost perpendicular to the long axis of the dimer. The dimer interface covers a surface area of ~1300-1500 $Å^2$ and is made up of an extensive network of salt bridges, hydrogen bonds, and van der Waals contacts. The conserved threonine and glutamate within the domain linker (consensus sequence G(D/E)TE(V/I/L)) are involved in conserved hydrogen bonds and salt-bridges to both repeats (Nam et al. 1999).

1.4.4 Regulation of PTPs

PTPs can undergo several posttranslation modifications that regulate phosphatase activity including phosphorylation, oxidation and nitrosylation (Salmeen et al. 2003; Alonso et al. 2004; Chen et al. 2008). The low pK_a of the active-site cysteine in PTPs makes this residue highly susceptible to oxidation and as such several mechanisms have evolved to protect PTPs from irreversible inactivation due to oxidation (Peters et al. 1998; Groen et al. 2005). Two mechanisms of reversible oxidative regulation are known: 1) formation of a cyclic sulphenyl-amide bond with the main-chain amine (Salmeen et al.

2003; van Montfort et al. 2003; Yang et al. 2007) and 2) formation of a disulfide bond with a backside (Buhrman et al. 2005) or vicinal cysteine (Caselli et al. 1998). As a result of forming these bonds the PTP active-site undergoes dramatic structural rearrangements. A comparison of the structural consequences of oxidation in PhyAsr CDC25B, RPTPα, and PTP1B suggests that oxidation of the catalytic cysteine has predictable effects on the conformation of the P-loop, general acid loop, and the conserved active-site arginine (Gruninger et al. 2008). Oxidation of the catalytic cysteine results in the formation of a semi-stable cysteine sulfenic acid which is rapidly converted to a disulfide or a sulphenyl-amide. If the cysteine sulfenic acid cannot form these reversible intermediates it is rapidly oxidized to sulfinic or sulfonic acid and the enzyme is irreversibly inactivated (Denu and Tanner 2002).

Chapter 2. Kinetic and Structural Analysis of a Bacterial Protein Tyrosine Phosphatase-Like *myo*-Inositol Polyphosphatase*.

2.1 Introduction

Protein tyrosine phosphatase (PTP) superfamily enzymes have been discovered in a range of prokaryotes, and most appear to serve roles that mimic their better-known eukaryotic counterparts as regulators of cellular function (Shi et al. 1998; Kennelly and Potts 1999). Some bacteria utilize PTPs as virulence factors by secreting them into infected host cells where they assist in the progression of infection (Bliska et al. 1991; Fu and Galan 1998; Bretz et al. 2003). The recently described PTP-like inositol polyphosphate phosphatase (IPPase) from *Selenomonas ruminantium*, PhyAsr, contains a PTP-like active-site signature sequence ($CX_5R(S/T)$), but lacks significant primary sequence identity with known IPPases and PTPs (< 20%) (Chu et al. 2004; Puhl et al. 2007). The biological function of PhyAsr is not clear. However it is the first example of a PTP-like enzyme with activity towards *myo*-inositol hexa*kis*phosphate ($InsP_6$), the most abundant cellular inositol polyphosphate (IPP) (Sasakawa et al. 1995).

* This chapter is an adapted version of the manuscript " Kinetic and Structural Analysis of a Bacterial Protein Tyrosine Phosphatase-like *myo*-inositol polyphosphatase. 2007. Puhl A.P.[1], Gruninger R.J.[1], Greiner R., Janzen T.W., Mosimann S.C. and Selinger L.B. Protein Sci. 161: 1368-1378."
[1]These authors contributed equally to this manuscript

The X-ray crystallographic structure of PhyAsr (Chu et al. 2004) reveals a PTP-like fold and several novel catalytic properties have been inferred. In particular, it was suggested that this enzyme is the first example of (1) a processive *myo*-inositol polyphosphatase (IPPase), (2) an IPPase that binds $InsP_6$ in a 5-axial/1-equatorial conformation and initiates hydrolysis at the 5-phosphate, and (3) a PTP-like enzyme that undergoes a significant substrate-induced P-loop conformational change. A mechanism has been proposed in which substrate moves between the catalytic site and a second substrate binding or standby site without being released as the successive intermediates are hydrolyzed (Chu et al. 2004). All previously characterized IPPases are distributive enzymes that release their products after each hydrolysis (Konietzny and Greiner 2002). These products can then act as substrates in further hydrolysis reactions. PTP and PTP-like enzymes frequently undergo substrate induced conformational changes in the WPD-loop that contains the conserved general acid (Jia et al. 1995; Zhang 1998; Zhang 2003). In contrast, substantial substrate induced conformational changes in the P-loop, responsible for binding the scissile phosphate, have not been previously reported. Finally, preference for the 5-phosphate position of $InsP_6$ has only been reported in one other IPPase (Barrientos et al. 1994) and to our knowledge, the 5-axial/1-equatorial conformation of $InsP_6$ has not been observed in structures of IPPase complexes.

The catalytic mechanism of PTP superfamily enzymes has been extensively studied. The active-site signature sequence, $CX_5R(S/T)$, is required for activity (Zhang et al. 1994b; Zhang 1998; Zhang 2002; Zhang 2003) and the enzymes follow a two-step, general acid-general base mechanism of dephosphorylation. The invariant cysteine residue exists as a thiolate and catalysis involves the formation of a phosphocysteine

intermediate (Cirri et al. 1993; Zhang et al. 1994c; Zhou et al. 1994). Main-chain amines and the guanidinium group of the conserved arginine coordinate the scissile phosphate in the catalytic site and stabilize the negative charge of the substrate (Barford et al. 1994; Zhang et al. 1994c) while an invariant aspartic acid serves as the general acid (Zhang et al. 1994b; Jia et al. 1995; Lohse et al. 1997).

Given the movement of the P-loop, processive mechanism, and the specificity for the 5-phosphate position of $InsP_6$, that have been inferred from previous structural studies (Chu et al. 2004), we have investigated the catalytic mechanism of PhyAsr using a combination of kinetic, site-directed mutant and structural studies. We present experimental data that indicates PhyAsr follows a classical PTP mechanism of catalysis and that in the presence of high salt concentrations and pH $\geq$ 6.5, the P-loop of PhyAsr is observed in the "closed" or active conformation in both the presence and absence of ligand.

2.2 Experimental procedures

2.2.1 Expression construct production

The region coding for the mature *S. ruminantium* IPPase (*phyAsr*; GeneBank accession number AF177214) was amplified from genomic DNA using polymerase chain reaction (PCR). PhyAsr Forward and Reverse primers included an NdeI site (CAT ATG) for cloning and a 5' GC cap. The signal peptide sequence was determined using SignalP 3.0 (Nielsen et al. 1997; Bendtsen et al. 2004). PhyAsr numbering begins with 1 at the N-terminus of the protein sequence found in GeneBank including the predicted signal peptide (first 27 residues). The PCR product was digested with *Nde*I and ligated into

similarly digested pET28b vector (Novagen). Mutant proteins C252S, C252A, R258K and D223N were prepared from *phyA*sr using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's instructions. Constructs were sequenced by automated cycle sequencing at the University of Calgary, Core DNA and Protein services facilities. Sequence data were analyzed with the aid of SEQUENCHER version 4.0 (Gene Codes Corp.) and MacDNAsis version 3.2 (Hitachi Software Engineering Co., Ltd).

2.2.2 Protein production and purification

The *S. ruminantium* phyAsr ORF was expressed as a translational gene fusion in pET28b. Protein expression was carried out in *Escherichia coli* BL21(DE3) cells (Novagen) for 18 hours at 310 K. Cells were grown to an optical density of 0.6-0.8 and protein expression was induced by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Induced cells were harvested and resuspended in lysis buffer: 20 mM KH_2PO_4 (pH 7.0), 300 mM NaCl, 1 mM *β*-mercaptoethanol (βME), 5% glycerol and one Complete Mini, EDTA-free protease inhibitor tablet (Roche Applied Science). Cells were lysed by sonication and cell debris was removed by centrifugation at 20 000 × g. Recombinant 6×His tagged PhyAsr was purified to homogeneity by metal chelating affinity chromatography (Ni^{2+}-NTA-agarose). Protein was washed on the column with lysis buffer containing 15 mM imidazole (pH 8.0) and eluted with lysis buffer containing 400 mM imidazole (pH 8.0). PhyAsr was further purified by cation exchange (Macro-Prep High S) and size exclusion chromatography. The homogeneity of the purified protein was confirmed by SDS-polyacrylamide gel electrophoresis (Laemmli

1970), and Coomassie Brilliant Blue R-250 staining. Protein concentrations were determined using the extinction coefficient calculated by PROT-PARAM (Gasteiger et al. 2005). Protein was either used immediately or dialyzed into 20 mM ammonium bicarbonate (pH 8.0) and lyophilized for long term storage.

2.2.3 Crystallization

Initial crystallization conditions were identified using the Hampton Research Crystal Screen (Hampton Research). Crystals were obtained in 100 mM sodium cacodylate (pH 6.5), 1.35 M $(NH_4)_2SO_4$, 0.80 M NaCl, and 1% 1,6-hexanediol by sitting-drop vapour diffusion. Drops were mixed using 2 μl of 30 mg/mL protein with 2 μl of mother liquor. Crystals were transferred to a solution containing 3.1 M sodium malonate (pH 6.0), 1.3 M $(NH_4)_2SO_4$, 700 mM NaCl for cryoprotection and subsequently frozen in liquid nitrogen. Data was collected at the Advanced Light Source (ALS) beamline 8.3.1 for crystals of PhyAsr C252S and D223N. Data was collected from the R258K mutant using an FR591 Bruker-Nonius rotating anode X-ray generator (45 kV, 130 mA) and a Kappa CCD detector. Monochromatic Cu K_α radiation was generated using osmic mirrors. Data was integrated and scaled using HKL 2000 (Otwinowski and Minor 1997), the structures were determined by molecular replacement with AMoRe (Navaza 1994) and the CCP4 suite of programs (CCP4 1994). Structure refinement of all three mutants was carried out using CNS 1.0 (Brunger et al. 1998). All figures were generated using PyMOL (DeLano 2008). Statistics for the data collection and the refinement of PhyAsr C252S, D223N and R258K are shown in Table 2.1.

Table 2.1. Data collection and refinement statistics for the X-ray crystallographic structures of the C252S , D223N and R258K active-site mutants of PhyAsr.

	C252S	**D223N**	**R258K**
Wavelength (Å)	1.0217	1.1178	1.5418
Spacegroup	$P2_1$	$P2_1$	$P2_1$
Unit cell a, b, c (Å)	46.1, 137.1, 80.3	46.2, 137.4, 79.6	45.8, 137.2, 80.0
B (°)	102.9	102.6	103.5
Z / Solvent content	2 / 64%	2 / 64%	2 / 64%
Resolution (Å) / Completeness	40.0-2.0 / 86.0%	50.0-1.8 / 87.9%	40.0-2.3 / 92.3%
Total / Unique reflections	158 627 / 76 771	148 269 / 77 249	167 603 / 39 348
Average I/σ	13.5	14.9	10.5
R-merge	0.096	0.074	0.078
Reflections (>1.5σ)	56 017	77 160	39 322
R_{work} / R_{free}	19.7 / 22.3	19.6 / 21.4	19.2 / 22.7
Protein / Solvent atoms	5 114 / 722	5 103 / 610	5 094 / 595
RMSD bond length (Å)	0.006	0.008	0.005
RMSD bond angle (°)	1.05	1.14	0.99
Average Protein B ($Å^2$)	19.2	18.7	16.7
Average Solvent B ($Å^2$)	26.7	24.4	29.8

2.2.4 Assay of enzymatic activity and quantification of the liberated phosphate

Activity measurements were carried out at 310 K. Enzyme reaction mixtures consisted of a 700 μL buffered substrate solution and 150 μL of enzyme solution. The buffered substrate solution contained 50 mM sodium acetate (pH 5.0) and $InsP_6$ at the appropriate concentration (0.025 – 4 mM). Ionic strength was held constant at 200 mM with the addition of NaCl. Following the appropriate, empirically determined incubation period, the reactions were stopped and the liberated phosphate was quantified using the

ammonium molybdate method previously described (Yanke et al. 1998). A 750 μL aliquot of 5% (w/v) trichloroacetic acid was added to stop the reaction, followed by the addition of 750 μL of phosphomolybdate coloring reagent. The coloring reagent was prepared by the addition of 4 volumes 1.5% (w/v) ammonium molybdate solution in 5.5% (v/v) sulfuric acid to 1 volume 2.7% (w/v) ferrous sulfate solution. The color reaction was allowed to proceed for ten minutes and then phytase activity was quantified by measuring the absorbance of the reaction at 700 nm (A_{700}). Liberated inorganic phosphate was measured as A_{700} on the spectrophotometer. Background phosphate in each sample was determined by adding TCA to the assay mixture prior to the addition of enzyme. To quantify the amount of phosphate released, a standard curve was constructed by measuring the A_{700} of samples containing a known amount of phosphate. The steady-state kinetic constants (K_m, k_{cat}) for the hydrolysis of $InsP_6$ by PhyAsr were calculated using non-linear regression (Sigma-plot 8.0; Systat Software Inc)

2.2.5 Alkylation of catalytic cysteine

Alkylation reactions were performed in the dark with 10 mM iodoacetate buffered to pH 7.0 with 20 mM HEPES. The reaction was carried out in the presence of 150 mM guanidine-HCl for 25 min. A control reaction was carried out similarly except iodoacetate was omitted. Following alkylation, samples were assayed for phytase activity against 2 mM $InsP_6$ in 0.2 M sodium acetate (pH 5.0) with the ionic strength of the substrate solution standardized to 350 mM with NaCl. The competition assays were carried out by performing the alkylation reaction in the presence of $InsP_6$.

2.2.6 Accession numbers

The nucleotide sequence for *phyAsr* has been deposited in the GenBank database under GenBank Accession Number AF177214. The amino acid sequence of this protein can be accessed through the NCBI Protein Database under NCBI Accession number AAQ13669. The coordinates for the original crystal structure of PhyAsr showing the P-loop in an 'open' and 'closed' conformation are available from the Protein Data Bank with the accession numbers 1U24 and 1U26, respectively. Structure factors and coordinates for C252S, D223N, and R258K mutants have been deposited to the Protein Data Bank with the accession numbers 2B4U, 2B4P, and 2B4O, respectively.

2.3 Results

2.3.1 Catalytic mechanism

The catalytic properties of recombinant wild-type PhyAsr towards the naturally occurring $InsP_6$ substrate have been determined at its pH optimum (Table 2.2). The rate of phosphate release can be saturated and the initial rates of reaction suggest a classical Michaelis-Menten enzyme mechanism. This is consistent with the kinetic properties of other PTP superfamily enzymes towards the commonly used phosphatase substrate *p*-nitrophenyl phosphate (pNPP) (Zhang et al. 1992; Guo et al. 2002). The apparent k_{cat} and K_m of PhyAsr for $InsP_6$ are 264 ± 19 s^{-1} and 425 ± 28 μM respectively. PhyAsr has some activity towards pNPP, phosphotyrosine (< 1% specific activity relative to $InsP_6$).

Table 2.2: Kinetic parameters for phytase activity of PhyAsr and various PhyAsr mutants. Values given are the average ± standard deviation of at least three separate experimental runs. ND – Not detectable

	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($\mu M^{-1}s^{-1}$)
PhyAsr	264 ± 19	425 ± 28	0.621 ± 0.061
C252S	ND	ND	ND
C252A	ND	ND	ND
D223N	0.84 ± 0.06	40.0 ± 9.7	0.021 ± 0.005
R258K	0.25 ± 0.01	410 ± 131	$6.1 \times 10^{-4} \pm 2.0 \times 10^{-4}$

In order to characterize the PhyAsr catalytic mechanism and compare it to existing PTP enzymes, a series of site-directed mutants have been produced and subjected to kinetic and structural studies. In particular, the C252 nucleophile, D223 general acid and the invariant R258 of the P-loop have been conservatively mutated to serine, asparagine and lysine, respectively. The kinetic constants associated with these mutant enzymes and C252A are shown in Table 2.2. The C252S and C252A mutants have no detectable IPPase activity in our assay, while the D223N and R258K mutant k_{cat}/K_m values are reduced by 30 and 1000 fold respectively. The D223N K_m decreases by more than 10 fold while the R258K K_m is essentially unchanged. The effect of the R258K and both cysteine mutants on the activity of PhyAsr are identical to those observed for equivalent mutations in PTPs (Zhang et al. 1994c; Flint et al. 1997). This suggests the P-loop movements previously observed in PhyAsr do not alter the roles of the invariant residues in the classical PTP-like catalytic mechanism. The decrease in K_m for the D223N mutant is a novel feature of PhyAsr as equivalent mutations result in an increased K_m for PTP1, *Yersinia* PTPase and yeast low molecular weight PTPase (Wu and Zhang 1996; Lohse et al. 1997; Keng et al. 1999). The lower K_m likely arises from favorable electrostatic interactions between N223 and $InsP_6$. This suggests the low pH optima of

wild-type PhyAsr is required to maintain D223 in its acid form and minimize unfavorable electrostatic interaction with the large net negative charge (-6 at pH 5) of $InsP_6$. This interpretation is supported by a shift in the pH optima of D223N from 5 to 6 (Puhl et al. 2007). A solvent molecule likely serves as the general acid in the D223N mutant and its higher pK_a is responsible for the shift in the pH optimum.

To further confirm the role of C252 as the catalytic nucleophile PhyAsr was chemically inactivated using iodoacetate (IAA). Treatment of the PhyAsr with 10 mM IAA for 25 min abolished PhyAsr activity (Table 2.3). Treatment of PhyAsr with IAA in the presence of 2 mM $InsP_6$ prevented inactivation of the enzyme indicating that $InsP_6$ protects the catalytic cysteine from the alkylating agent.

Table 2.3: Effect of alkylation by iodoacetic acid (IAA) on the catalytic activity of PhyAsr. Alkylation reactions were performed at pH 7.0 for 25 minutes. The specific activity is expressed as the amount of phosphate released (μmol) per min per mg of protein.

	Specific Activity
PhyAsr	520
PhyAsr + 10 mM IAA	ND
PhyAsr + 10 mM IAA + 2 mM $InsP_6$	516

2.3.2 X-ray crystallographic structures

Attempts to grow crystals of the PhyAsr mutants under the conditions used by Chu et al. (2004) were unsuccessful. A crystallization screen identified high salt conditions that yield crystals that diffract to a higher resolution than the previously published structure (1.8 Å). Despite the differences in the crystallization conditions, the space group, unit cell, crystal packing, and crystal contacts are nearly identical to the previously reported crystals (Chu et al. 2004).

The *S. ruminantium* phytase monomer is composed of two domains: a core PTP domain, and a small highly twisted partial β-barrel domain that is unique to protein tyrosine phosphatase-like phytases (Figure 2.1). The *S. ruminantium* phytase crystallized as a dimer in the asymmetric unit (Figure 2.2a). Size exclusion chromatography identified that the molecular weight of the phytase in solution was approximately 70 kDa. Analysis of the dimer interface (Laskowski et al. 1997) buries a surface area of 1342 $Å^2$ and involves an extensive network of van der Waals contacts, hydrogen bonds and salt bridges. The dimer interface forms between the N-terminal strand, β1, the top two strands (β7 and β8) of the partial β-barrel domain, and linker between α7 and α8 (Figure 2.2b).

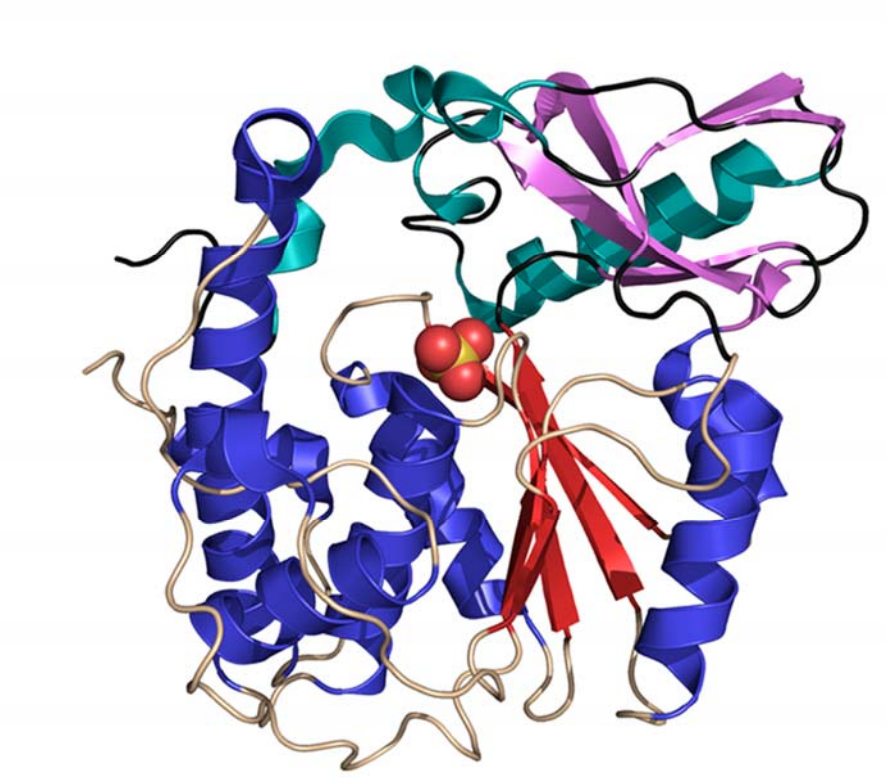

Figure 2.1: Structure of the PTP-like phytase from *Selenomonas ruminantium*, PhyAsr. Fold of PhyAsr monomer consists of a core PTP domain (helices, strands, and loops are colored blue, red, and beige, respectively) and a small partial β-barrel domain (helices, strands, and loops are colored cyan, purple, and black, respectively)

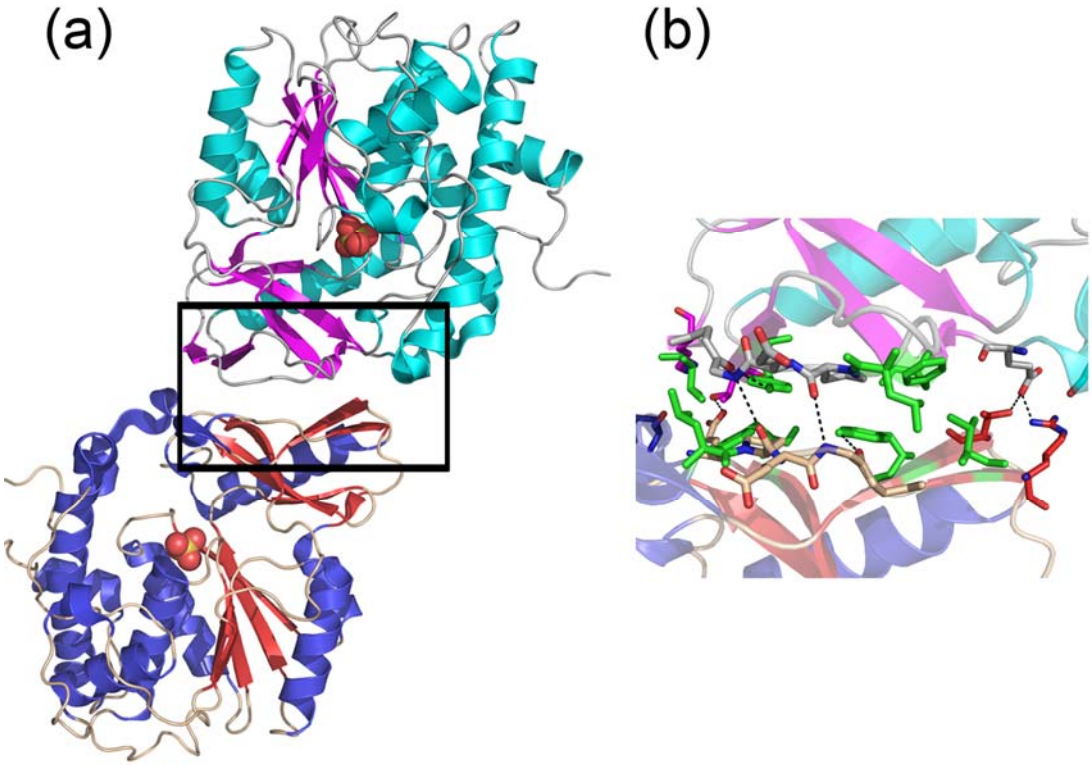

Figure 2.2 PhyAsr forms a dimer. (a) PhyAsr dimer observed in the asymmetric unit of the crystal and (b) close up of dimer interface in the boxed region in (a). Residues that make hydrogen bonds or salt bridges are shown as sticks. Residues that make van der Waals contacts are shown in green.

To determine whether any of the mutations affected the overall fold of PhyAsr we compared the structures of the mutants to the structure of wild-type PhyAsr. Least squares superposition of the main-chain atoms of the C252S, D223N, and R258K structures with wild-type PhyAsr (1U24) yield root mean square deviations (RMSDs) in the range of 0.49 – 0.55 Å, indicating the overall structure of the protein is not disrupted by these active-site mutations (Figure 2.3).

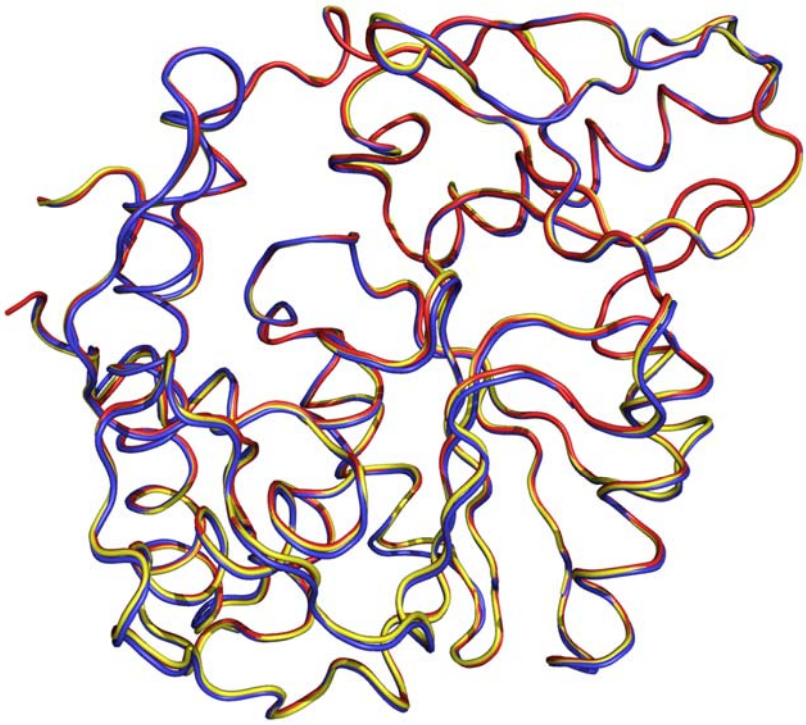

Figure 2.3: P-loop mutations do not affect the overall fold of PhyAsr. Least squares superposition of the main-chains of the C252S (Yellow), D223N (Blue), and R258K (red) mutants of PhyAsr. A trace of the superimposed Cα atoms (RMSD between 0.49-0.55 Å) is shown for each mutant.

The catalytic centre of PTPs is composed of two highly conserved loops called the phosphate binding loop (P-loop) and the WPD-loop. The WPD loop contains the conserved general acid and undergoes a well characterized conformational change during catalysis (Jia et al. 1995; Zhang 1998; Zhang 2003). The equivalent loop in PhyAsr contains an equivalent aspartic acid (D223), but cannot undergo a conformational change due to steric clashes which would result from contacts with the partial β-barrel domain which is absent in other PTPs. Instead, the P-loop of PhyAsr has been observed in two distinct conformations; open (ligand free) and closed (ligand bound) (Chu et al. 2004). We have determined the structure of PhyAsr in the presence and absence of ligand

(Figure 2.4). A large positive F_o-F_c peak was observed in the active-site of the C252S mutant that was determined to be a sulfate ion (Figure 2.4a). In contrast, the D223N (Figure 2.4b) and R258K (Figure 2.4c) structures do not have ligand bound in the active-site. Least squares superposition of the P-loop residues (251-259) of the three mutants reveals that they have virtually identical conformations with RMSDs of 0.13 Å (D223N onto C252S) and 0.19 Å (R258K onto C252S). In contrast to Chu et al (2004), we do not observe movement in the P-loop between an open, catalytically inactive and a closed, catalytically active form upon ligand binding. Least squares superposition of the P-loops in our mutant structures and in the PhyAsr-*myo*-inositol hexasulfate complex (1U26; closed or active conformation) yield RMSDs below 0.30 Å. The equivalent superpositions between our mutants and 1U24 (open or inactive conformation) have RMSDs greater than 1.00 Å (Figure 2.4d). Consequently, each of our mutant structures have P-loop conformations equivalent to the ligand bound, closed conformation reported for 1U26. This is the same P-loop conformation observed in multiple PTP X-ray structures both in the presence and absence of ligand (Burke and Zhang 1998).

While the P-loop remains in the catalytically competent conformation in each of our mutants, there are subtle differences in the positions of active-site residues in the R258K structure. R258 forms a conserved salt bridge with E136 that orients the guanidinium relative to the scissile phosphate. The replacement of Arg with Lys results in a shorter side-chain and the K258 salt bridge with E136 is lengthened by 0.88 Å. To compensate for both the size and a small movement in the K258 side-chain, the Cγ D223 shifts by 0.95 Å towards the K258 Nζ and the D223 carboxylate fills the volume vacated as a result of the mutation.

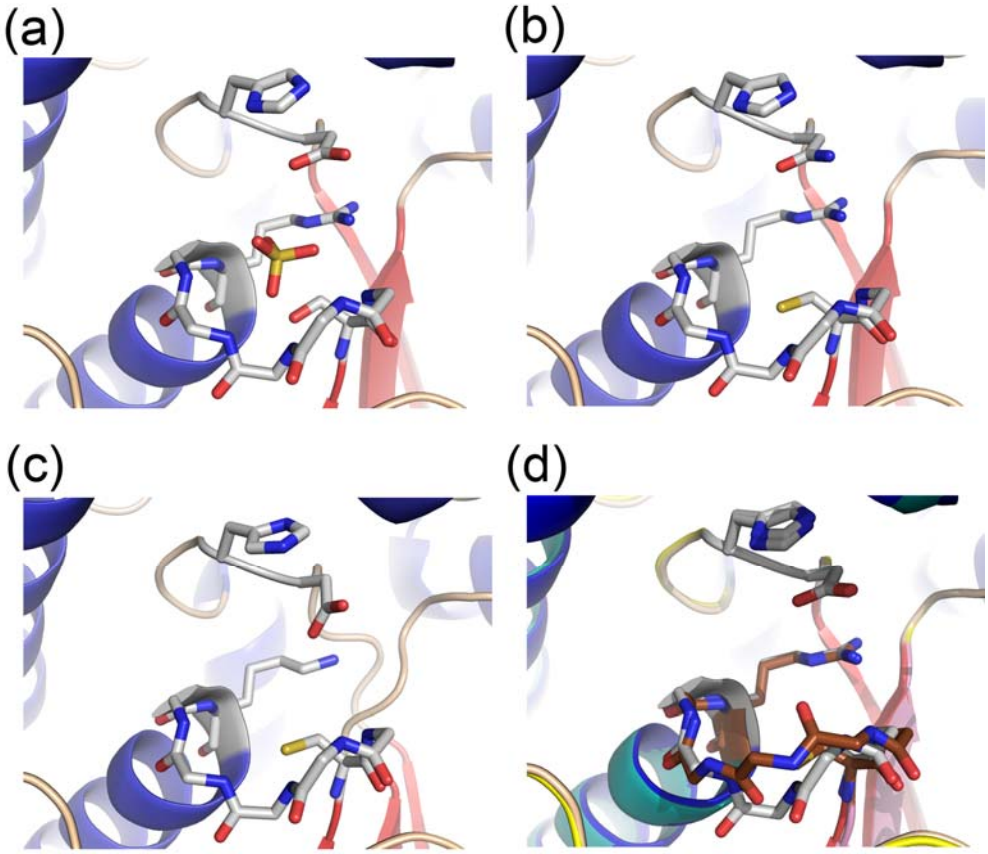

Figure 2.4. Structures of three active-site mutants of PhyAsr. (a) C252S (b) D223N (c) R258K and (d) least squares superposition of the C252S mutant structure (yellow loop) with the apo structure (1U24) (grey loop) previously solved (Chu et al. 2004). D223 and H224 of the general acid loop, and the P-loop residues, including the catalytic cysteine (C252) are shown as sticks. Least squares superposition of the P-loop main-chain atoms (residues 251-259) of the D223N and R258K mutants, 1U26, and the 1U24 with P-loop of the C252S mutant results in RMSDs of 0.13, 0.19, 0.11, and 1.15 Å (respectively). Oxygens are shown in red, nitrogens are shown in blue and sulfur is shown in yellow.

2.4 Discussion

The active-site of PTP superfamily enzymes can undergo conformational rearrangements upon substrate binding or as a regulatory mechanism. In PTPs, the main chain of the conserved WPD loop undergoes significant conformational changes upon substrate binding that bring the invariant Asp (general acid) within hydrogen bond distance of the leaving group (Jia et al. 1995; Zhang 1998; Zhang 2003). Although it has

long been known that this general acid loop undergoes a large induced fit movement, only recently has the phosphate binding loop (P-loop) been observed to undergo conformational changes (Chu et al. 2004). In several PTPs, reversible oxidation of the catalytic cysteine serves as a regulatory mechanism to reversibly inactivate the protein and is accompanied by a conformational change in the P-loop (Salmeen et al. 2003; van Montfort et al. 2003; Buhrman et al. 2005). More recently the P-loop of PhyAsr has been observed in two distinct conformations; a catalytically inactive open conformation and a catalytically active closed conformation (Chu et al. 2004). The conformational change from an open to a closed P-loop conformation in PhyAsr has been proposed to be a result of ligand binding (Chu et al. 2004).

Isostructural and conservative mutations of the three PhyAsr active-site residues that are invariant among PTP superfamily enzymes suggest the conformational differences observed in previous structural studies of PhyAsr do not affect the functional role of these residues. The PTP catalytic mechanism has been extensively studied and has been found to be a two-step process (Zhang et al. 1994b). In the first step the phosphate in the active site undergoes a nucleophilic attack by the catalytic cysteine, while a conserved aspartic acid acts as the general acid, protonating the leaving group and cleaving the scissile phosphodiester bond (Lohse et al. 1997). The first step of this reaction mechanism generates a phospho-enzyme intermediate (Guan and Dixon 1991; Zhang 1995). In the second step of the reaction the aspartic acid that acted as a general acid in the first step activates a nucleophilic water molecule that hydrolyzes the thioester linkage between the enzyme and the phosphate (Zhang 1995; Lohse et al. 1997). Each of the mutations (C252S, D223N, and R258K) produce comparable kinetic effects when

compared to equivalent mutations in the mammalian receptor-like PTP LAR, the *Yersinia* PTP and PTP1B indicating that PhyAsr utilizes a classical PTP reaction mechanism (Cirri et al. 1993; Zhang et al. 1994b; Zhang et al. 1994c; Flint et al. 1997; Lohse et al. 1997; Orchiston et al. 2004). The sole exception is a 10-fold decrease in K_m associated with the D223N mutation that is unique to PhyAsr. This difference is attributed to the large net charge (-6) associated with $InsP_6$ under physiological conditions and the polar, uncharged character of N223. The observation that chemical modification of C252 inactivates the enzyme is also consistent with its role as the nucleophile (Table 2.3). Additionally, alkylation of PhyAsr in the presence of $InsP_6$ did not inactivate the enzyme indicating the $InsP_6$ binding blocks IAA from reacting with the catalytic cysteine (Table 2.3).

In addition to the X-ray crystallographic and site directed mutagenesis results presented here, the $InsP_6$ degradation pathway and substrate specificity of PhyAsr has been investigated (Puhl et al. 2007). High performance ion exchange chromatography was used to identify the inositol phosphate intermediates generated by PhyAsr. In contrast to the classification of PhyAsr as a 5-phytase (Chu et al. 2004), Puhl and colleagues (2007) found that the 3-phosphate is preferentially removed from $InsP_6$ first. Molecular docking calculations also predicted that PhyAsr binds $InsP_6$ with the 3-phosphate in the active site (Puhl et al 2007). High performance ion exchange separation of the inositol phosphate intermediates also revealed that $InsP_6$ was degraded to Ins(2)P in an ordered, sequential fashion with a major degredation pathway (representing >80% of the total inositol phosphates) and two minor pathways (Figure 2.5). Interestingly, all of the $InsP_6$ is degraded prior to the appearance of $Ins(1,2,4,5,6)P_5$. This latter

observation is inconsistent with a processive mechanism in which $InsP_6$ is hydrolyzed to less phosphorylated inositol phosphates without releasing the intermediates.

Figure 2.5. Dephosphorylation pathways of $InsP_6$ by PhyAsr. Larger arrows indicate the major pathway and the smaller arrows indicate the minor pathways. The major pathway accounts for approximately 80% of the degradation pathway.

X-ray crystallographic structures of each PhyAsr mutant demonstrate that the overall fold of the enzyme is unchanged in comparison to the wild-type enzyme. In contrast to previous suggestions that the P-loop undergoes a conformational change upon substrate binding, the P-loop in each mutant structure adopts the closed or functional conformation in both the presence (C252S) and absence (D223N, R258K) of bound ligand. These differences cannot be due to differences in the crystalline lattice as each mutant structure has the same unit cell, space group and crystalline contacts as 1U24 (open or inactive). Furthermore, the D223N and R258K residues do not make contacts with P-loop residues and cannot account for the observed closed conformation. Finally, there are no conformational differences between our mutant structures and 1U24 that would prevent the P-loop from adopting the open conformation. This suggests that factors other than substrate binding promote the functional conformation of the PhyAsr P-loop.

Chapter 3. Effect of Ionic Strength and Oxidation on the P-loop Conformation of the Protein Tyrosine Phosphatase-Like Phytase, PhyAsr[*].

3.1 Introduction

Enzymes that degrade *myo*-inositol 1,2,3,4,5,6-hexa*kis*phosphate ($InsP_6$) are ubiquitous in nature and have been identified in prokaryotes, protists, fungi, animals, and plants (Caffrey et al. 1999; Mullaney et al. 2000). $InsP_6$ is the most abundant inositol phosphate in the cell, and has been implicated in important cellular processes including DNA repair, mRNA export, RNA editing, cellular signaling, endocytosis and vesicular trafficking (Gaidarov et al. 1999; York et al. 1999; Hanakahi et al. 2000; Raboy 2003). The generic term phytase is applied to enzymes that hydrolyze $InsP_6$ into inorganic phosphate and various lower phosphorylated *myo*-inositols. The recently described protein tyrosine phosphatase (PTP)-like phytase from *Selenomonas ruminantium*, PhyAsr, contains the PTP active-site signature sequence ($CX_5R(S/T)$), is structurally similar to PTPs, and utilizes the same catalytic mechanism as PTPs to hydrolyze phosphodiester bonds (Chu et al. 2004; Puhl et al. 2007). Although the biological

[*] Parts of this chapter has been published in FEBS Journal. Title: Effect of Ionic Strength and Oxidation on the P-loop Conformation of the Protein Tyrosine Phosphatase-like Phytase, PhyAsr. 2008. Gruninger R.J., Selinger L.B., and Mosimann S.C. FEBS J 275: 3783-3792.

function of these PTP-like phytases is unclear, they are the first example of an enzyme with a PTP fold that is capable of hydrolyzing $InsP_6$.

Initial structural studies of the *S. ruminantium* phytase report a unique conformational change in the active-site phosphate binding loop (P-loop) that takes place upon substrate binding**.** This movement is distinct from the major structural change in the general acid (WPD) loop of many PTPs that accompanies substrate binding (Stuckey et al. 1994; Jia et al. 1995). At near physiological ionic strength, the P-loop of PhyAsr adopts a catalytically inactive, open conformation in the absence of ligand, and a catalytically active, closed conformation upon substrate binding (Chu et al. 2004). P-loop movements have been observed in PTPs as a result of 1) mutation (Scapin et al. 2001), 2) oxidative regulation (Salmeen et al. 2003; van Montfort et al. 2003; Buhrman et al. 2005; Yang et al. 2007) and 3) crystal contacts (Stewart et al. 1999).

In a recent structural study, the conformational change in the P-loop was not observed at high ionic strength (Puhl et al. 2007). In this work, we examine the structure of the PhyAsr P-loop as a function of ionic strength, and upon oxidation of the catalytic cysteine. We also explore the possibility that PhyAsr is regulated by reversible oxidation. Our examination of the P-loop movement in PhyAsr and its comparison to several PTP structures provides an understanding of the factors that influence P-loop movements within the PTP superfamily. A comparison of the structural consequences of oxidation in PhyAsr CDC25B, RPTPα, and PTP1B suggests that oxidation of the catalytic cysteine has predictable effects on the conformation of the P-loop, general acid loop, and the conserved active-site arginine.

3.2 Experimental Procedures

3.2.1 Purification and crystallization

The *S. ruminantium phyA* (PhyAsr) ORF (minus putative signal peptide) was expressed as a translational gene fusion in pET28b. Amino acids were numbered according to the complete coding sequence of the *S. ruminantium* protein sequence (AAQ13669) including the putative signal peptide. This numbering scheme differs by 11 residues from that used by Chu et al (2004) but is consistent with the numbering in Puhl et al (2007). Recombinant His tagged PhyAsr was purified to homogeneity by metal chelating affinity (Ni^{2+}-NTA-agarose, Qiagen), cation exchange (Macro-Prep High S, BioRad), and size exclusion chromatography. The purified protein was dialyzed into 10 mM ammonium bicarbonate (pH 8.0), lyophilized, and stored at 253 K. Crystallization experiments were conducted using sitting-drop vapor diffusion with drop ratio of 2 μL of 30 mg/mL protein solution and 2 μL of reservoir. Crystals were grown in 8 – 10 % PEG 8000, 200 – 500 mM NaCl, 50 mM sodium acetate pH 4.8. Crystals were cryoprotected using a solution containing the crystallization reagents and 25 % glycerol. The catalytic cysteine was oxidized by treating the crystals with 100 μM H_2O_2 for 45 minutes prior to freezing.

3.2.2 Data collection and structure determination

Data was collected at 100 K on beamline 8.3.1 at the Advanced Light Source on crystals with approximate dimensions of 0.1 mm × 0.1 mm × 0.4 mm. Data was integrated and scaled with HKL 2000 (Otwinowski and Minor 1997), and structure refinement was done with CNS 1.0 (Brunger et al. 1998). The D223N structure (2B4P)

(Puhl et al. 2007) was used to solve the structures of PhyAsr at ionic strengths of 200 mM ($PhyAsr_{I200}$, 2PSZ), 300 mM ($PhyAsr_{I300,}$ 3D1O), 400 mM ($PhyAsr_{I400,}$ 3D1Q), 500 mM ($PhyAsr_{I500,}$ 3D1H), and with the catalytic cysteine (C252) oxidized ($PhyAsr_{ox}$, 2PT0). The space group and unit cell parameters of the crystals used in this study were identical to those in 2B4P. The isomorphous nature of the crystals allowed us to use the coordinates of 2B4P to calculate phases with the program SFall (CCP4 1994). Statistics for the data collection and refinement of $PhyAsr_{I200}$, $PhyAsr_{I300}$, $PhyAsr_{I400}$, $PhyAsr_{I500}$ and $PhyAsr_{ox}$, are shown in Table 3.1.

3.2.3 Kinetic assays

Kinetic assays were performed at 310 K using the standard phytase assay as previously described (Yanke et al. 1998), at ionic strengths of 0.10, 0.20, 0.35, 0.50, and 1.0 M using substrate concentrations ranging from 0.10 mM to 4 mM. This method was found to give consistent, although slightly larger, kinetic parameters than those obtained using the method of Heinonen and Lahti (Heinonen and Lahti 1981). The substrate's contribution to ionic strength was calculated assuming a net charge of -6 based on the pK_a values for $InsP_6$ (Isbrandt and Oertel 1980). The ionic strength was calculated using the equation, $I = \frac{1}{2}\sum c_i Z_i^2$, where I is the ionic strength of the solution, c_i and Z_i are the concentration and charge of species i, respectively. The sum is taken over all ionic species in the reaction or crystallization buffer. Ionic strength of the assays was standardized using NaCl. Kinetic data was fit to the Michaelis-Menten equation using non-linear regression (Sigma-plot 8.0; Systat Software Inc.).

Table 3.1: Data collection and refinement statistics for the structure of PhyAsr at ionic strengths of 200 mM ($PhyAsr_{I200}$), 300 mM ($PhyAsr_{I300}$), 400 mM ($PhyAsr_{I400}$), and 500 mM ($PhyAsr_{I500}$) and after oxidation of the catalytic cysteine ($PhyAsr_{ox}$). The PDB codes for the 200, 300, 400, and 500 mM ionic strength and oxidized structures are 2PSZ, 3D1O, 3D1Q, 3D1H and 2PT0, respectively.

Data collection statistics*

	$PhyAsr_{I200}$	**$PhyAsr_{I300}$**	**$PhyAsr_{I400}$**	**$PhyAsr_{I500}$**	**$PhyAsr_{ox}$**
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Cell (Å) (a,b,c/β)	45.9,137.1,80.0 /102.8˚	46.0,138.5,80.7 /102.4˚	45.9,137.4,79.9 /102.9˚	45.9,137.2,79.8 / 103˚	46.0,137.9,80.3 /102.8˚
Resolution (Å)	50 - 2.0	50 - 2.07	50 – 2.1	50 – 2.1	50 - 1.7
Reflections (total)	217 582	112 992	107 389	99 759	205 871
Reflections (unique)	61 360 (4135)	52 151(2115)	50 546 (2171)	50 868 (2360)	90 985 (5219)
Complete (%)	93.8 (63.8)	90.9 (50.6)	89.9 (47.8)	91.0 (56.4)	89.9 (54.0)
Average I/σ	15.6 (2.9)	13.2 (3.0)	16.6 (4.8)	13.3 (3.2)	23.7 (3.8)
R_{merge}[a] (%)	7.4 (22.8)	5.7 (25.0)	4.4 (16.0)	6.3 (20.7)	3.8 (23.7)
Refinement Statistics					
Protein atoms	5103	5098	5098	5098	5130
Non-protein	635	518	546	514	685
R_{factor}[b]	19.9	17.5	17.0	18.4	17.7
R_{free}[b]	22.5	21.2	20.3	21.4	18.9
RMSD bonds (Å)	0.006	0.013	0.014	0.010	0.009
RMSD Angle (˚)	1.22	1.44	1.46	1.25	1.24
B-factors					
Main-chain B	19.1	16.5	14.4	16.5	17.9
Side-chain B	19.9	20.0	18.2	19.6	21.1
Solvent B	25.2	24.6	22.6	22.9	28.5
***Ramachandran Plot* (%)**					
Most favoured	91.6	91.8	92.3	91.8	91.8
Additional allowed	8.0	8.0	7.3	7.8	8.2
Generously allowed	0.4	0.2	0.4	0.4	0

*Values in parentheses are for the highest resolution shell

[a] $R_{\text{merge}} = \sum \; | I_{\text{hkl}} - < I_{\text{hkl}} > | \; / \sum I_{\text{hkl}}$

[b] $R_{\text{factor}} = \sum_{\text{hkl}} \; \big| \; | F_{\text{obs}} | \; - | F_{\text{calc}} | \; \big| / \sum | F_{\text{obs}} |$

3.2.4 Oxidation sensitivity assays

The method of Denu and Tanner (Denu and Tanner 2002) was employed to examine the sensitivity of PhyAsr to oxidation and to determine if the enzyme is regulated by reversible oxidation. PhyAsr (25 μM) was incubated with 100 μM or 1 mM H_2O_2 to oxidize the catalytic cysteine. Aliquots were withdrawn 5, 10, 15, 30, and 60 minutes after addition of H_2O_2 and the reaction was quenched by the addition of catalase. The protein was then either directly assayed or added to protein storage buffer containing 10 mM DTT for 30 minutes and then assayed. The level of inactivation was determined by comparing the specific activity of oxidized PhyAsr to enzyme that had not been exposed to H_2O_2. Phytase activity was measured at 310 K using the standard phytase assay as described previously (Yanke et al. 1998). Kinetic assays were performed with the ionic strength standardized to 200 mM with NaCl.

3.3 Results

3.3.1 Ionic strength affects the catalytic efficiency of PhyAsr

To test the hypothesis that ionic strength affects the P-loop conformation of PhyAsr we determined the steady state kinetic parameters at several ionic strengths (Table 3.2). There is a 7-fold increase in catalytic efficiency (k_{cat}/K_m) and a 9-fold decrease in K_m as the ionic strength is increased from 100 to 500 mM. The increase in catalytic efficiency and decrease in K_m that is observed as a result of increasing ionic strength is consistent with the P-loop movement occurring in this range. Alternatively, the increase in ionic strength may favorably alter the electrostatic interactions between

the protein and substrate and enhance enzymatic efficiency. Further increases in ionic strength result in a decrease in k_{cat}/K_m, primarily due to an increase in K_m.

Table 3.2: Effect of ionic strength on the hydrolysis of $InsP_6$ by PhyAsr. Standard error is shown for at least six measurements.

I (mM)	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1}s^{-1}$)
100	1.29 ± 0.24	515 ± 45	398 ± 82
200	0.76 ± 0.05	678 ± 19	893 ± 66
350	0.36 ± 0.03	608 ± 17	1675 ± 144
500	0.14 ± 0.01	369 ± 7	2599 ± 202
1000	1.00 ± 0.18	164 ± 13	163 ± 32

3.3.2 Structure of PhyAsr under low and high ionic strength conditions

To examine the structural effect of ionic strength on the P-loop, X-ray crystal structures of PhyAsr were determined at several ionic strengths ranging from 200 mM to 500 mM (Table 3.1) using conditions almost identical to those reported by Chu et al (2004). Differences in the crystallization conditions are subtle and were necessary to produce optimal diffraction-quality crystals. The resulting space group, unit cell and crystal contacts are identical to those previously observed (Chu et al. 2004; Puhl et al. 2007). The structure of wild type PhyAsr ($PhyAsr_{I200}$) was determined in the absence of a ligand at low ionic strength (Figure 3.1a, 2PSZ)**.** Surprisingly, unlike the structure previously determined under similar conditions (1U24) (Chu et al. 2004) the P-loop is in a catalytically competent closed conformation (Figure 3.1b). Structures of PhyAsr at ionic strengths of 300, 400, and 500 mM (3D1O, 3D1Q, 3D1H, respectively) were also determined and in all cases the P-loop adopts the closed conformation (Figure 3.2).

These results are consistent with the P-loop conformation observed by Puhl et al (2007) at an ionic strength of > 2 M (P-loop residues 251-259 < 0.1 Å RMSD) and indicate that the closed P-loop conformation is stable over a broad ionic strength range.

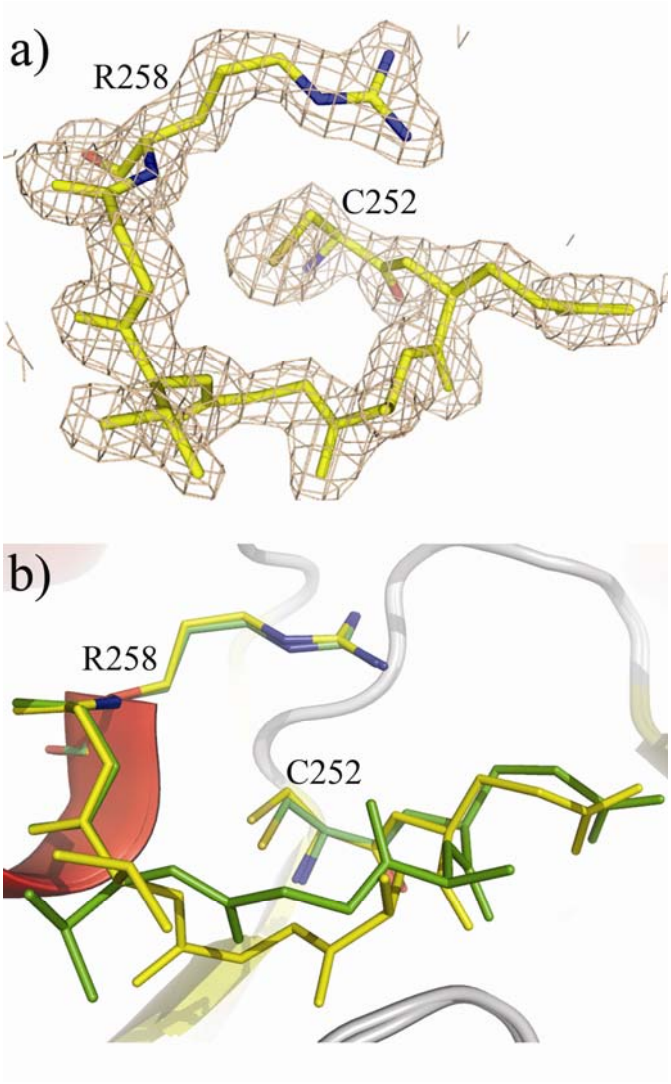

Figure 3.1: In the absence of ligand at low ionic strength PhyAsr P-loop adopts a closed conformation. (a) Conformation of the P-loop at low ionic strength in the absence of ligand. The P-loop is observed in the catalytically competent conformation. Electron density from a sigma weighted $2F_o - F_c$ map is shown at a contour level of 1σ. (b) Least squares superposition of PhyAsr with the open P-loop (green) (1U24) and PhyAsr under low ionic strength conditions (yellow) (2PSZ). RMSD of the P-loop main-chain atoms is 1.18 Å. All figures were generated with PyMOL (Delano W.L. The PyMOL Molecular Graphics System (2002) DeLano Scientific, Palo Alto, CA, USA).

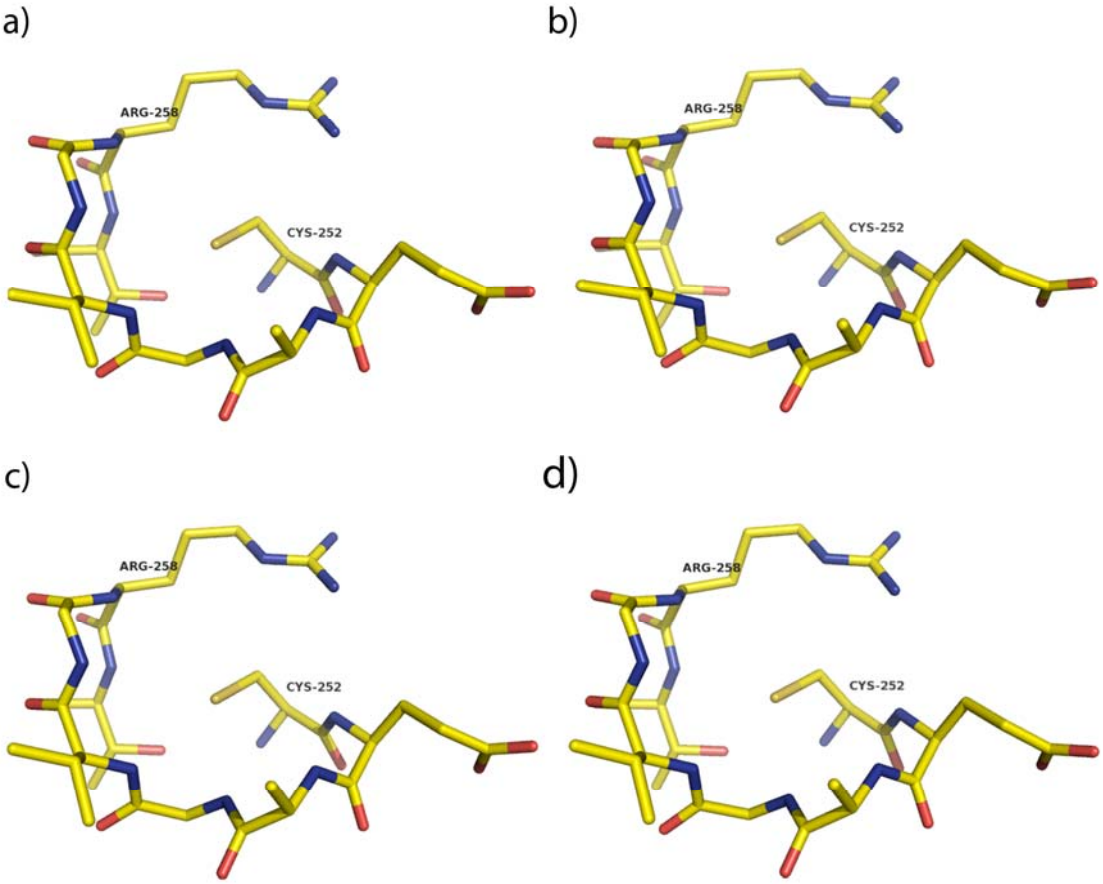

Figure 3.2: The structure of the P-loop in PhyAsr is observed in the catalytically competent conformation at ionic strengths of (a) 200 mM (2PSZ), (b) 300 mM (3D1O), (c) 400 mM (3D1Q), and (d) 500 mM (3D1H).

3.3.3 Structure of PhyAsr upon oxidation of the catalytic cysteine

A systematic comparison of the open P-loop conformation in PhyAsr to all unliganded PTP structures in the Protein Data Bank found that CDC25B adopts a roughly similar P-loop conformation upon oxidation of the catalytic cysteine (Buhrman et al. 2005). To test whether the movement of the P-loop in PhyAsr is due to oxidation of the catalytic cysteine, we oxidized crystals of PhyAsr with 100 μM H_2O_2 for

45 minutes and solved the structure of oxidized protein ($PhyAsr_{ox}$, 2PT0). The P-loop in $PhyAsr_{ox}$ adopts the open conformation as a result of oxidation of the catalytic cysteine to cysteine sulfonic acid (Figure 3.3a). Least squares superposition of the P-loop main-chain atoms of 1U24 and our oxidized structure (0.16 Å RMSD) clearly shows that the open P-loop conformation previously observed (Chu et al. 2004) is identical to the P-loop conformation after oxidation of the catalytic cysteine (Figure 3.3b).

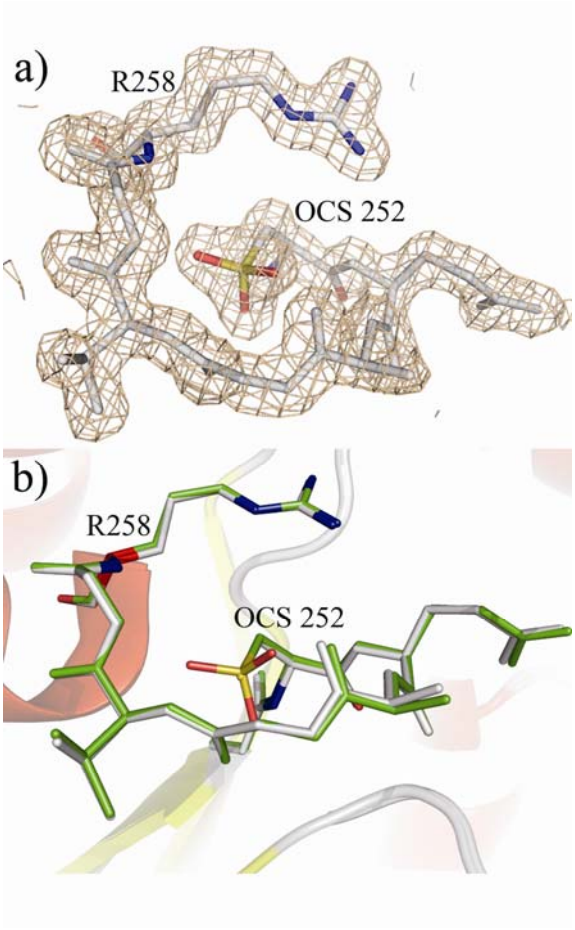

Figure 3.3: The P-loop of PhyAsr adopts an open conformation upon oxidation of the catalytic cysteine. (a) The P-loop is observed in the catalytically inactive, open conformation upon oxidation of the catalytic cysteine. Electron density from a sigma weighted $2F_o - F_c$ map is shown at a contour level of 1σ. The residue OCS 252 corresponds to the active-site cysteine oxidized to cysteine sulfonic acid. (b) Least squares superposition of PhyAsr with the open P-loop (green) (1U24) and with C252 fully oxidized (grey) (2PT0).

Modeling the cysteine as cysteine sulfenic or sulfinic acid in alternate conformations results in positive difference density around the oxygens and indicates that the observed residue is cysteine sulfonic acid**.** After obtaining the open P-loop conformation we examined the electron density of 1U24 using the deposited structure factors. This analysis reveals relatively large electron density and positive difference density surrounding the sulfur atom suggesting the cysteine is oxidized (Figure 3.4). To verify that the observed P-loop conformation is a result of oxidation, we omitted the P-loop from 1U24 and 2PT0, carried out a refinement cycle, and calculated omit maps. For both 1U24 and 2PT0, the model of PhyAsr with a cysteine sulfonic acid produces the best fit to the unbiased electron density (Figure 3.5a and 3.5b, respectively), again indicating that the open P-loop conformation is a result of oxidation of the cysteine.

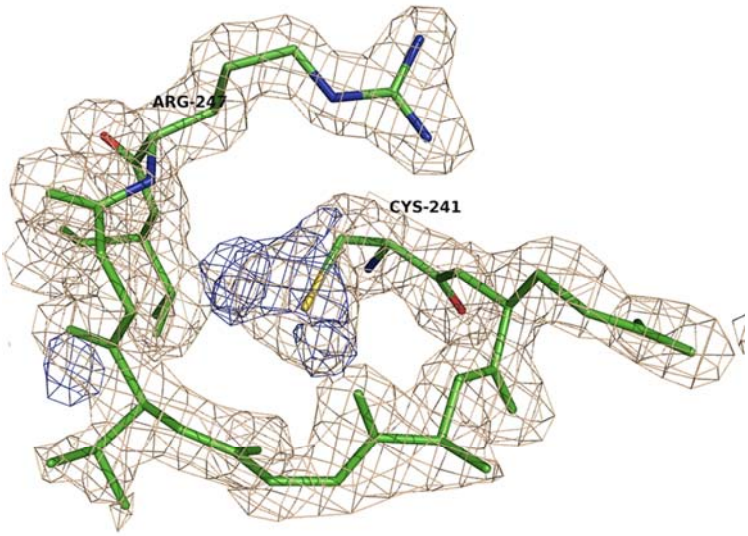

Figure 3.4: The catalytic cysteine of 1U24 is oxidized to cysteine sulfonic acid. Sigma weighted electron density calculated using the coordinates and structure factor amplitudes deposited with 1U24. Sigma weighted $2F_o$-F_c (wheat) and F_o-F_c (light blue) is shown at 1σ and 4σ, respectively. There are three clear positive difference electron density peaks equally spaced about the cysteine sulfur.

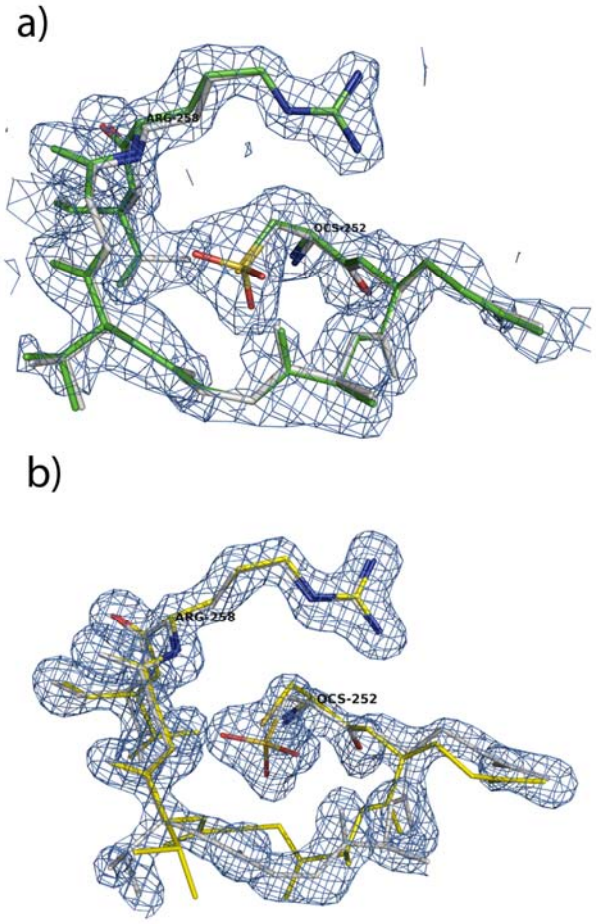

Figure 3.5: Omit electron density fits an open P-loop with the catalytic cysteine sulfonic oxidized to cysteine sulfonic acid. (a) Least squares superposition of the P-loops of 1U24 (green) and 2PT0 (grey) fit into unbiased sigma weighted F_o-F_c omit electron density (2.0 Å resolution) contoured at 1σ (cyan). (b) Least squares superposition of the P-loops of 2PSZ (yellow) and 2PT0 (grey) fit into unbiased sigma weighted F_o-F_c omit electron density (1.7 Å resolution) contoured at 4σ. The residue OCS-252 corresponds to the active-site cysteine oxidized to a cysteine sulfonic acid.

3.3.4 Comparison of contacts to the P-loop in the unoxidized and oxidized conformations

To accommodate the larger size of the cysteine sulfonic acid the P-loop must undergo a conformational change. In the absence of the large P-loop movement the main-chain amine of G255 makes a 2.32 Å contact with S_γ, a 2.21 Å contact with $O_\delta 1$, and a 1.81 Å contact with $O_\delta 2$ of the cysteine sulfonic acid. In addition to these contacts there is a 1.82 Å contact between G255 C_α and $O_\delta 2$ of the oxidized cysteine. To further understand the structural consequences of the P-loop transition that occurs upon oxidation, the program Contact (CCP4 1994) was used to compare all the interactions made to the catalytic cysteine or cysteine sulfonic acid that are less than 4 Å (Table 3.3). This analysis identified six contacts that are made directly with the cysteine S_γ in the unoxidized conformation. Oxidation of the catalytic cysteine decreases the contacts made to the cysteine S_γ but results in the formation of 17 additional interactions with the oxygens of the cysteine sulfonic acid (Table 3.3). The large number of contacts made with the cysteine sulfonic acid oxygens stabilizes the open P-loop conformation (Figure 3.6). The average B-factors of the P-loop residues in the unoxidized and oxidized structures are 14.3 $Å^2$ and 15.4 $Å^2$, respectively. This is 4-5 $Å^2$ lower than the overall B-factors of the structures (19.5 $Å^2$) and indicates the P-loop adopts a stable conformation in both the oxidized and unoxidized enzyme. To further support our conclusion that the previously observed open conformation is a result of oxidation of the catalytic cysteine we compared the contacts made to the cysteine S_γ in 1U24 and our oxidized structure and found them to be nearly identical (Table 3.3) further suggesting that the cysteine is oxidized in 1U24.

Table 3.3: Comparison of all sub-four angstrom contacts between cysteine and the P-loop in the structures of PhyAsr.

Contacts	1U24[a]	Low I[b]	Oxidized Cys[b]			
P-loop	Sγ	Sγ	Sγ	$O\delta_1$	$O\delta_2$	$O\delta_3$
C252	-	-	-	-	2.70[1]	-
E253	3.63[1]	3.49[1]	3.47[1]	3.14[1]	3.66[1]	-
A254	-	-	-	3.27[1]	-	-
G255	3.55[2]	3.22[1]	3.62[2]	-	3.31[2]	3.68[2]
V256	-	3.65[1]	-	-	-	-
G257	-	-	-	-	-	3.44[1]
R258	-	-	-	-	-	3.30[1]
T259	-	3.63[1]	-	-	-	3.04[1], 3.08[3]
Solvent	3.55, 3.78	3.15, 3.93	3.53, 3.83, 3.94	2.71, 3.14, 3.26	2.93	3.05, 3.27
Non-P-loop	-	-	-	-	3.09[4]	-

[a] PDB code for PhyAsr with the P-loop in the open conformation (Chu et al. 2004)
[b] PDB codes for the Low I and Oxidized Cys structures are 2PSZ and 2PT0, respectively
[1] contact with main-chain amine, [2] contact with a main-chain carbonyl, [3] contact with a side-chain, [4] contact with a non P-loop side-chain

3.3.5 Oxidation of cysteine affects the conformation of several residues

The P-loop conformational change is primarily due to a large shift in the ϕ/ψ torsion angles in A254 (ϕ/ψ = -88.7/-19.7 to ϕ/ψ = -146.5/136.4) and the ψ torsion angle of G255 (18.9 to -157.9) upon oxidation. This large rotation of the peptide bond between A254 and G255 results in a 4.12 Å movement of G255 C_α and a 2.27 Å movement in V256 C_β which is accompanied by a rotation of 110˚ about χ_1. The highly conserved T259 undergoes a rotation about χ_1 of 123˚ that breaks a hydrogen bond formed with C252 Sγ in the unoxidized conformation, and results in the formation of two hydrogen bonds with the main-chain carbonyl oxygen of G257 and $O_\delta 3$ of the oxidized cysteine. This conformation of T259 is also stabilized by R71 which normally makes a bidentate interaction with the carbonyl oxygen of G255. These movements away from the oxidized

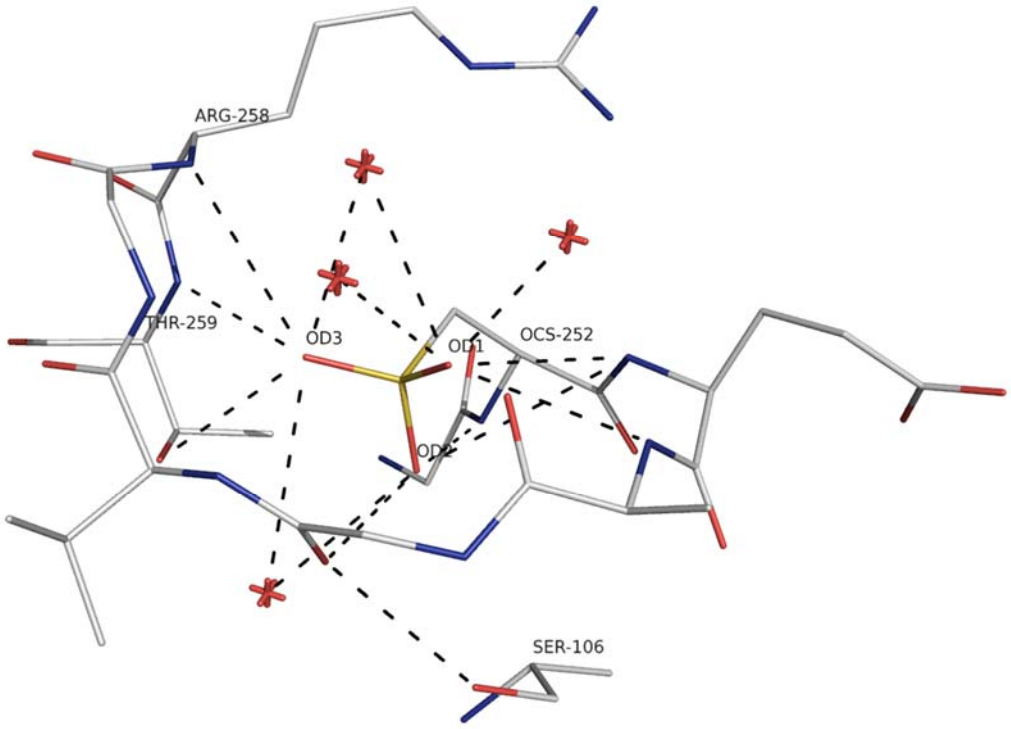

Figure 3.6: Oxidation of the catalytic cysteine to cysteine sulfonic acid (OCS-252) results in the formation of many inter-residue contacts to the OCS-252 oxygens and the P-loop. Waters are shown as red stars.

cysteine increase the space inside the P-loop to accommodate the large sulfonic acid group. The movements in the P-loop main chain are accompanied by a rotation of S106 χ_1 by 172^o to form a 3.09 Å hydrogen bond with the carbonyl oxygen of G255. This movement breaks two hydrogen bonds S106 makes with the main-chain carbonyl oxygen of A107 and the R68 main-chain amine in the unoxidized enzyme. It also appears that the movement in S106 fills the void that forms as a result of the P-loop movement. The P-loop conformation in the oxidized form results in a rearrangement of the hydrogen bonding pattern seen in the unoxidized form. Upon oxidation, the number of hydrogen bonds formed with solvent doubles from 5 to 10. Four of the additional solvent contacts are made by the ordered water 461 (numbering in PhyAsr$_{ox}$) which

makes two bidentate hydrogen bonds with the P-loop main-chain, and $O_{\delta}1$ and $O_{\delta}2$ of the oxidized cysteine.

Although there are movements in the P-loop, there are no other major conformational changes in the protein. Most notably, the loop containing the general acid (D223) does not move upon oxidation. Unlike most PTPs the general acid loop in PhyAsr cannot undergo an open to closed movement upon substrate binding due to the presence of a small β-barrel domain that is unique to this protein and involved in substrate binding (Chu et al. 2004; Puhl et al. 2007). No other residues in the protein were found to be modified by the treatment with H_2O_2.

3.3.6 Structural consequences of oxidation in the PTP superfamily

Structures of PTP1B (van Montfort et al. 2003), CDC25B (Buhrman et al. 2005), RPTPα (Yang et al. 2007) and PhyAsr (this work) have been determined with the cysteine oxidized to cysteine sulfenic (SO), sulfinic (SO_2) and sulfonic (SO_3) acid. Interestingly, oxidation of the catalytic cysteine to SO, SO_2 or SO_3 has been found to have different effects on the P-loop conformation in different enzymes. In PTP1B and RPTPα the P-loop is unchanged, whereas in CDC25B, the P-loop adopts a conformation that is similar, but not identical, to that observed in PhyAsr (Figure 3.7). Although the movements in the P-loops of PhyAsr and CDC25B are not identical they both serve to provide room for the larger oxidized cysteine. The key feature that dictates whether the P-loop moves upon oxidation of the catalytic cysteine is the ability of the conserved active-site arginine to move in a concerted fashion with the general acid loop (Figure 3.7). The general acid loop (WPD loop in PTP1B, and RPTPα) also undergoes a

large conformational change in many, but not all, PTPs (Stuckey et al. 1994; Jia et al. 1995). In the absence of ligand, the WPD loop of PTP1B and RPTPα adopts an open (inactive) conformation and upon ligand binding, a closed (active) conformation (Figure 3.7a). In oxidized PTP1B (Salmeen et al. 2003; van Montfort et al. 2003) and RPTPα (Yang et al. 2007) the position of the active-site arginine, and the WPD loop are in the open (general acid) conformation. The general acid loop and active-site arginine in PhyAsr are not free to undergo a similar conformational change (Chu et al. 2004; Puhl et al. 2007). As a result, the P-loop must move to provide room for the larger oxidized cysteine (Figure 3.7a). In CDC25B, Y428 and M531 occupy the region that corresponds to the general acid loop in PhyAsr and prevent the active-site arginine from moving. As a consequence of this steric constraint, the P-loop of CDC25B also undergoes a large conformational change upon oxidation of the catalytic cysteine (Figure 3.7b).

3.3.7 Sensitivity and reversibility of PhyAsr oxidation

Many PTPs undergo reversible oxidation *in vivo* as a regulatory mechanism. To examine whether the catalytic cysteine in PhyAsr can be reversibly oxidized we performed an oxidation time course as described by Denu and Tanner (Denu and Tanner 2002). Treatment of PhyAsr with 100 μM H_2O_2 resulted in 33 % of the enzyme being inactivated after 10 minutes of treatment and 65 % inactivated after 30 minutes. Incubation of the enzyme for up to an hour with 100 μM H_2O_2 does not further inactivate the enzyme. Incubation of PhyAsr with 1 mM H_2O_2 results in an 85 % decrease in activity after 10 minutes and a loss of approximately 95 % of its activity after 30 minutes. Treatment of PhyAsr with lower levels (50 μM) of H_2O_2 also resulted in a loss of activity

(approximately 20 % after 10 minutes). If the inactivation of the protein is due to the formation of a stable sulfenic acid, sulphenyl-amide, or disulfide then the addition of a reducing agent will restore enzymatic activity. In all cases the addition of 10 mM DTT did not restore any enzyme activity indicating that the inactivation is due to irreversible oxidation.

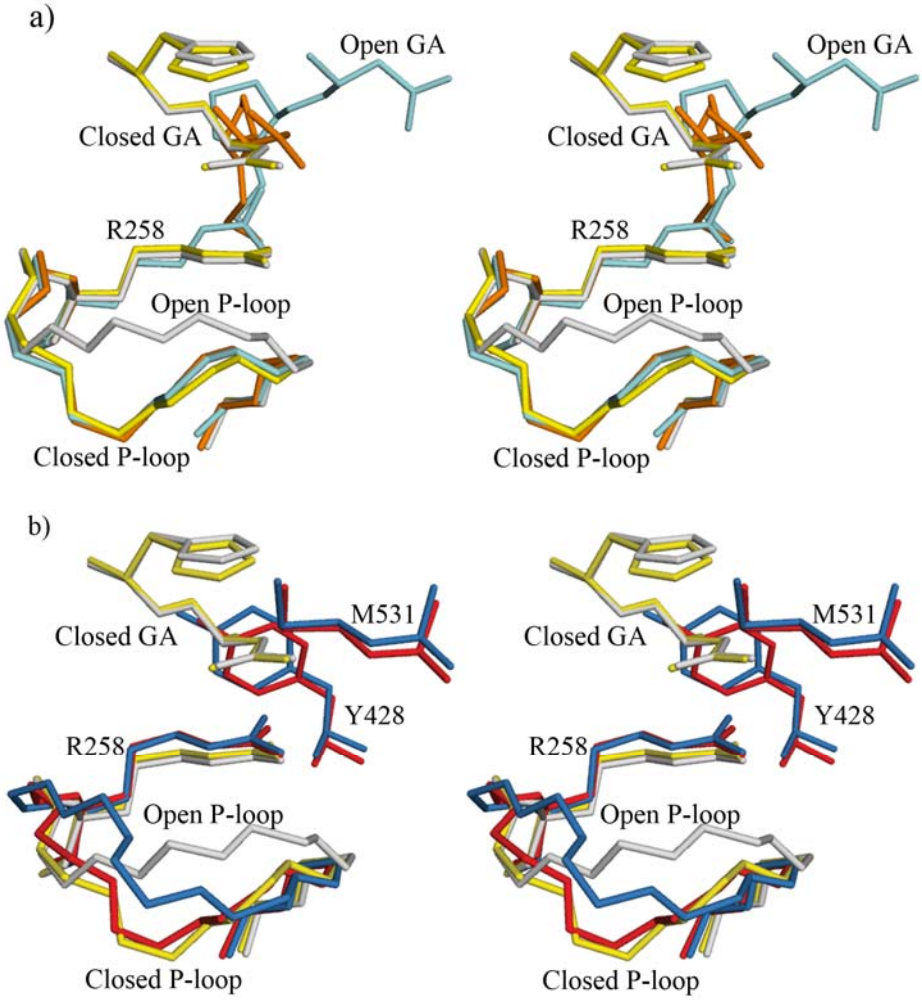

Figure 3.7: Concerted movements in the P-loop and general acid loop in PTPs upon oxidation of the catalytic cysteine. (a) Divergent stereoview of a least squares superposition of unoxidized (yellow) (2PSZ) and oxidized (grey) (2PT0) PhyAsr, with oxidized PTP1B (light blue) (1OEO), and PTP1B with the general acid loop (GA) and active-site arginine in the closed, active conformation (orange) (1PTV). (b) Divergent stereoview of a least squares superposition of unoxidized (yellow) (2PSZ) and oxidized (grey) (2PT0) PhyAsr, with unoxidized (red) (1YMK) and oxidized (blue) (1YMD) CDC25B.

3.4 Discussion

3.4.1 Effect of ionic strength on PhyAsr catalysis and P-loop structure

Changes in ionic strength have been observed to affect the catalytic efficiency of some PTPs. For example, at high ionic strength the k_{cat}/K_m of Yop51 (Zhang et al. 1992), and PTP1 (Zhang 1995), decrease 24 and 132 fold (respectively), primarily due to an increase in the K_m. The increase in K_m was attributed to a weakening of the electrostatic interactions between the substrate and the highly charged active-site (Zhang et al. 1992; Zhang 1995). In contrast to Yop51 and PTP1, increasing ionic strength from 100 mM to 500 mM enhanced the binding of $InsP_6$ to PhyAsr (Table 3.2)**.** Given the absence of structural changes as a function of ionic strength and ligand binding, the effect of ionic strength on catalytic activity is likely not of a structural nature. Instead, we suggest that the enhanced catalytic efficiency is due to the shielding of unfavorable electrostatic interactions between the active-site and the highly charged substrate. This is consistent with previous kinetic studies that found mutating the general acid, D223, to asparagine resulted in a 10-fold decrease in K_m, which was hypothesized to be due to unfavorable electrostatic interactions between the more electronegative aspartic acid and $InsP_6$ (Puhl et al. 2007).

Analysis of the PDB database identified at least 25 protein tyrosine phosphatase structures that were determined in the absence of a ligand in the active-site. Least squares superposition of the P-loop main chain (starting at the residue prior to catalytic cysteine and ending after the conserved arginine) results in RMSD values of < 0.75 Å (Table 3.4). In 23 of these PTPs the P-loop adopts the closed catalytically competent P-loop conformation observed in PhyAsr. Two exceptions were observed: 1) the apo

structure of the PTP1B C215S (Scapin et al. 2001) and 2) the MAPK phosphatase Pyst1 (Stewart et al. 1999). Interestingly, the P-loop of the C215S mutant of PTP1B has also been observed in the closed catalytically competent conformation (Jia et al. 1995), while the P-loop conformation of Pyst1 was attributed to a crystal contact. These findings indicate that in the absence of a ligand the P-loop adopts the closed conformation. The only other P-loop movements that have been observed in PTPs are a result of oxidation of the catalytic cysteine.

3.4.2 Sensitivity of PhyAsr to oxidation

The low pK_a of the active-site cysteine in PTPs makes this residue highly susceptible to oxidation (Peters et al. 1998; Groen et al. 2005). Reversible oxidation is an important regulatory mechanism in PTPs and two mechanisms of reversible oxidative regulation are known: 1) formation of a cyclic sulphenyl-amide bond with the main-chain amine (Salmeen et al. 2003; van Montfort et al. 2003; Yang et al. 2007) and 2) formation of a disulfide bond with a backside (Buhrman et al. 2005) or vicinal cysteine (Caselli et al. 1998). As a result of forming these bonds, the PTP active site undergoes dramatic structural rearrangements. Oxidation of the catalytic cysteine results in the formation of a semi-stable cysteine sulfenic acid which is rapidly converted to a disulfide or a sulphenyl-amide. If the cysteine sulfenic acid cannot form these reversible intermediates it is rapidly oxidized to sulfinic or sulfonic acid and the enzyme is irreversibly inactivated (Denu and Tanner 2002). Our examination of the sensitivity and reversibility of oxidation in PhyAsr indicates that this protein is moderately resistant to oxidation, and

that it does not undergo oxidative regulation. The reversibility and sensitivity to oxidation varies throughout the PTP superfamily and it has been suggested that some

Table 3.4: Least squares superposition of main-chain atoms of the P-loop (HCX_5RS/T) of PTP structures determined in the absence of an active-site ligand. The low ionic strength structure of PhyAsr (2PSZ) was used as the target model.

PDB ID[a]	***RMSD (Å)***
2PSZ	Target
1G4U	0.27
1I9S	0.27
1JLN	0.37
1L8K	0.32
1LAR	0.36
1OHC	0.36
1P15	0.28
1RPM	0.24
1C25	0.75
2A8B	0.38
2AHS	0.32
2B30	0.42
2BIJ	0.38
2FH7	0.37
1T48	0.37
1PA1	0.36
1YFO	0.40
2C7S	0.31
1YGR	0.66
2CFV	0.37
2BZL	0.38
2B49	0.40

[a] 1G4U = *Salmonella* SPTP, 1I9S = PTP domain mouse mRNA capping enzyme, 1JLN = PTP-SL, 1L8K = T-cell PTP, 1LAR = LAR, 1OHC = CDC14, 1P15 = PTPα-D2, 1RPM = PTPμ domain 1, 1C25 = Cdc25A, 2A8B = PCPTP1, 2AHS = RPTP-β, 2B30 = SHP-1, 2BIJ = PTPN5, 2FH7 = PTPσ, 1T48 = PTP1B, 1PA1 =PTP1B C215D, 1YFO = RPTPα-D1, 2C7S = PTPκ, 1YGR = CD45, 2CFV = RPTP-Type J, 2BZL =PTPN14, 2B49 =PTP-non-receptor type 3

PTPs have evolved an intrinsic resistance to oxidation (Ross et al. 2007). PTEN is readily oxidized and has an active site that is narrower, and half as deep (5 Å x 11 Å opening, and 8 Å depth) as the PhyAsr active site (6 Å x 14 Å opening, and 14 Å depth). In contrast, MTMR2 is highly resistant to oxidation and has an active site with a similar depth but a smaller opening (9 Å x 8 Å opening, and 13 Å depth). This suggests that the size and shape of the PTP active site not only influences substrate specificity (Yuvaniyama et al. 1996), but may also be involved in resistance to oxidation.

3.4.3 Oxidation of PTPs and the role of P-loop flexibility

Conformational changes in the P-loop of PTPs have been observed as a result of both reversible and irreversible oxidation events. The conformational changes observed in PhyAsr are due to the irreversible oxidation of the catalytic cysteine to a cysteine sulfonic acid. A comparison of the structural consequences of oxidation in PhyAsr to those in oxidatively regulated PTPs (CDC25B, RPTPα, and PTP1B) suggests that oxidation of the catalytic cysteine has predictable effects on the active-site conformation. When the catalytic cysteine is irreversibly oxidized the P-loop will only move when steric constraints prevent the movement of the general acid loop and the active-site arginine. Interestingly, this is also observed upon oxidation to the reversible sulfenic (SO) form, an intermediate in the formation of a sulphenyl-amide or disulfide (Salmeen et al. 2003; van Montfort et al. 2003; Buhrman et al. 2005; Yang et al. 2007). The formation of a reversible intra-molecular covalent bond (sulphenyl-amide or disulfide) requires the cysteine to undergo a significant conformational change. For this to occur, the P-loop must undergo a separate and distinct conformation rearrangement regardless of the

position of the general acid and active-site arginine. In summary, the conformation of the P-loop only changes 1) when the general acid loop and active-site arginine are sterically constrained and 2) upon intra-molecular bond formation.

Chapter 4. Structural Analysis of a Multifunctional, Tandemly Repeated Inositol Polyphosphatase [*].

4.1 Introduction

Inositol polyphosphates (IPPs) are ubiquitous in nature and their central role in eukaryotic cellular signaling is well known (Irvine and Schell 2001). *Myo*-inositol 1,2,3,4,5,6-hexa*kis*phosphate ($InsP_6$) is the most abundant cellular inositol phosphate and has been implicated in numerous important cellular processes including DNA repair, RNA processing and export, development, apoptosis, and pathogenicity (York et al. 1999; Hanakahi et al. 2000; Chatterjee et al. 2003; Macbeth et al. 2005; Tan et al. 2007; Majerus et al. 2008). The important biological role that $InsP_6$ plays is exemplified by the observation that deletion of enzymes involved in $InsP_6$ biosynthesis has a lethal phenotype in mouse embryos (Frederick et al. 2005; Verbsky et al. 2005). Inositol phosphatases (IPPases) that degrade $InsP_6$ are ubiquitous in nature and have been identified in prokaryotes, protists, fungi, animals, and plants and are generically referred to as phytases (Caffrey et al. 1999; Mullaney et al. 2000). Four distinct classes of phytate degrading IPPases have been identified to date and include the histidine acid phosphatases, β-propeller phytase, purple acid phosphatase, and protein tyrosine

[*] Parts of this chapter has been published in Journal of Molecular Biology. Title: Structural Analysis of a Multifunctional, Tandemly Repeated Inositol Polyphosphatase. 2009. Gruninger R.J., Selinger L.B., and Mosimann S.C. *in press*.

phosphatase-like phytase (PTPLP), also known as cysteine phytases (Mullaney and Ullah 2003; Chu et al. 2004; Puhl et al. 2007). Phytases catalyze the stepwise removal of phosphates to generate lower inositol phosphates in a highly ordered, specific manner (Konietzny and Greiner 2002). Although phytases are functionally similar, they are structurally and mechanistically diverse.

The recently described PTPLP class of IPPases possesses a protein tyrosine phosphatase (PTP) active-site signature sequence ($CX_5R(S/T)$), is structurally similar to PTPs and utilizes a classical PTP reaction mechanism (Chu et al. 2004; Puhl et al. 2007). Interestingly, these enzymes display no catalytic activity against classical PTP substrates due to several unique structural features that confer specificity for inositol polyphosphates (Puhl et al. 2007; Puhl et al. 2008a; Puhl et al. 2008b; Puhl et al. 2009). Although the biological function of these enzymes is unclear, they have been found in a wide range of bacteria including plant and human pathogens (Lim et al. 2007; Nakashima et al. 2007). A PTPLP identified in the rumen and gastro-intestinal inhabitant *Mitsuokella multacida*, is unique among known PTPLPs because it is composed of tandemly repeated PTPLP domains. This tandem arrangement is commonly observed in receptor protein tyrosine phosphatases (RPTPs) (Andersen et al. 2001) but not in IPPases. Interestingly, only the N-terminal (D1) repeat in many RPTPs is catalytically active. The function of the inactive C-terminal (D2) repeat in RPTPs is not fully understood but has been shown to be involved in regulation and substrate recognition in some RPTPs (Ng et al. 1995; Kashio et al. 1998; Blanchetot and den Hertog 2000; Yang et al. 2007).

In this work we describe the structure of the tandemly repeated $InsP_6$ degrading IPPase from *M. multacida*. Using a combination of macromolecular docking, site-

directed mutagenesis and kinetic assays we demonstrate that each repeat has a unique substrate specificity. These results shed light upon the biological function of PTPLPs and may have implications for RPTP function. The results of our analysis also suggest that PhyAmm may utilize a novel mechanism of $InsP_6$ dephosphorylation.

4.2 Experimental procedures

4.2.1 Cloning and mutagenesis

The *M. multacida* (PhyAmm) open reading frame (minus putative signal peptide) was amplified from genomic DNA using polymerase chain reaction (PCR) with Phusion DNA polymerase (Finnzymes). Amino acids were numbered according to the complete coding sequence of the *M. multacida* protein sequence (ABA18187) including the putative signal peptide. The amplified product was ligated into the *NdeI* site of pET28b and transformed into *Escherichia coli* DH5α. Mutagenesis was carried out using counter PCR amplification of the expression plasmid as previously described (Street et al. 1991). All PCR products were sequenced at the University of Calgary Core DNA and Protein services facilities. Sequence data was analyzed using the programs SEQUENCHER 4.0 (Gene Codes Corp.) and MacDNAsis 3.2 (Hitachi Software Engineering Co.). All of the primers used in this study were purchased from Integrated DNA technologies.

4.2.2 Purification of PhyAmm

Protein expression was carried out with *E. coli* BL21(DE3) cells transformed with the pET28b expression construct. Cells were grown to an optical density (600 nm) of 0.6 - 0.8 in Luria-Bertani broth supplemented with 50 μg/mL of kanamycin. Protein

expression was induced by adding isopropyl-β-D-thiogalactopyranoside to the culture at a final concentration of 1 mM. The over-expression was carried out at 310 K in an incubating shaker for 18 h. Induced cells were harvested and resuspended in lysis buffer (20 mM KH_2PO_4 (pH 7.0), 300 mM NaCl, 10 mM β-mercaptoethanol (BME), 25 mM imidazole (pH 8.0)). Cells were lysed by sonication, and cell debris was removed by centrifugation. PhyAmm was purified to homogeneity by metal chelating affinity (Ni^{2+}-NTA-agarose), cation exchange (Macro-Prep High S, BioRad), and size exclusion chromatography. The purity of the protein was assessed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (Laemmli 1970) and Coomassie Brilliant Blue R-250 staining. Protein concentration was determined by measuring absorbance at 280 nm using the extinction coefficient calculated with PROT-PARAM (Gill and von Hippel 1989). The purified protein was stored in 10 mM KH_2PO_4 (pH 7.0), 100 mM NaCl, 10 mM BME, and 0.1 mM EDTA. Purified protein was used immediately, or flash frozen in liquid nitrogen and stored at 193 K.

4.2.3 Crystallization

Crystallization experiments were conducted using sitting-drop vapor diffusion with a drop ratio of 2 μL of 5-10 mg/mL protein solution and 2 μL of reservoir. Crystals were grown at room temperature, in 6 – 10% PEG 8000, 1-5% ethylene glycol, 100 mM TRIS (pH 8.0). After several rounds of streak seeding, large, fragile crystals were obtained. Crystals were cryo-protected by gradually transferring them into a solution containing the crystallization reagents and 25% ethylene glycol, followed by flash freezing in liquid nitrogen. Heavy atom substituted crystals were obtained by soaking the

crystals in a solution consisting of the crystallization reagents and 20 mM sodium tungstate for 1 h prior to freezing.

4.2.4 Data collection and structure determination

Raw diffraction data was processed with MOSFLM (Leslie 1992) and scaled with SCALA (Evans 1997) within the CCP4 program suite (CCP4 1994). Ellipsoidal truncation and anisotropic correction of the data was performed by the diffraction anisotropy server to account for anisotropy in the diffraction data (Strong et al. 2006). Native and derivative data sets were collected at 100 K on beamline 8.3.1 at the Advanced Light Source and the structure was solved by the multi-wavelength anomalous dispersion (MAD) method with a WO_4^{2-} soaked crystal. A three wavelength dataset (1.2150 Å, 1.2146 Å and 1.1579 Å) was collected at the tungsten LIII edge. Phases (20 - 2.8 Å) were determined using SOLVE (Terwilliger and Berendzen 1999) and density modification was carried out using RESOLVE (Terwilliger 2001). SOLVE identified four tungsten atoms and obtained phases with a figure of merit of 0.51. Solvent flattening with RESOLVE resulted in phases with a figure of merit of 0.76 and a readily interpretable map. The resulting electron density map was used for automated model building with RESOLVE and ARPWARP (Perrakis et al. 1999). This model served as a starting point for iterative cycles of simulated annealing refinement, positional and B-factor refinement against a native dataset (20 - 2.3 Å) in CNS 1.2 (Brunger 2007) followed by manual model building with XFIT (McRee 1999). Continuous electron density was observed for amino acids 46-636 with the remaining residues (11-45, 637-640) located at the termini and assumed to be disordered (the expression tag is encoded

by residues 11-34). Electron density for the N and C terminal active-sites is shown in Figure 4.1. Unless indicated otherwise, figures were prepared with PyMOL (version 0.99) (DeLano 2008). PROCHECK was used throughout refinement to assess the stereochemistry of the model (Laskowski et al. 1993). Statistics for the data collection and refinement are shown in Table 4.1.

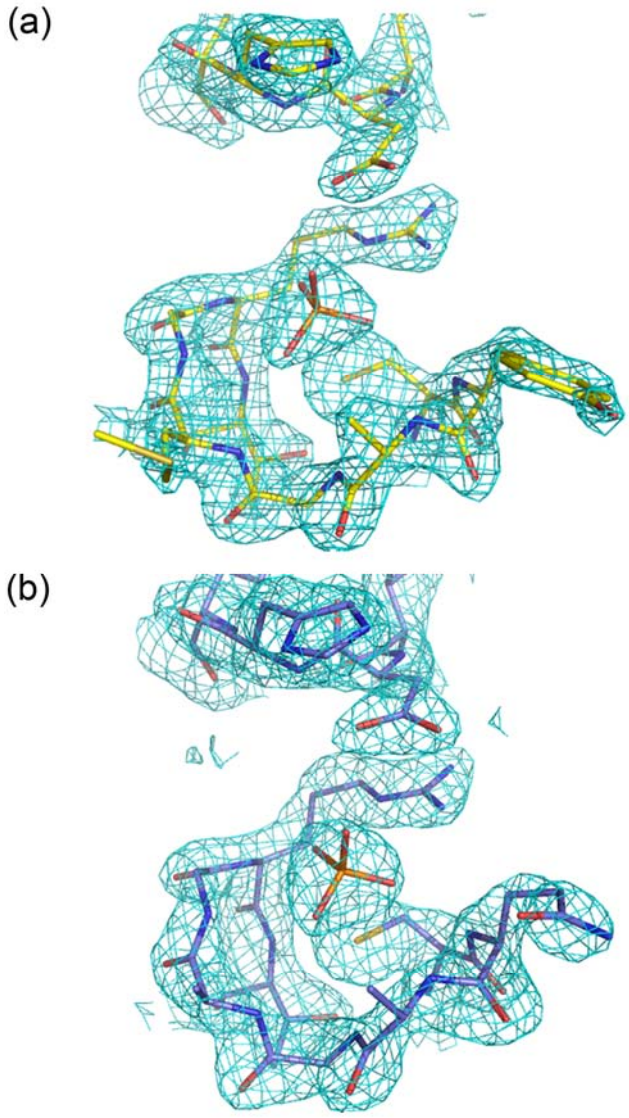

Figure 4.1: Sigma-A weighted 2Fo-Fc electron density (2.3 Å resolution) for the (a) D1 and (b) D2 active-sites contoured at 1.8 σ.

Table 4.1: Data collection and refinement statistics for PhyAmm[1]

	Native	WO_4^{2-} derivative		
Data collection				
Space group	$P2_1$	$P2_1$		
Unit cell	74.39, 73.30	74.25, 74.00, 160.83/93.9		
a, b, c (Å)/β (°)	161.31/93.7			
		Inflection	Peak	High Energy
Wavelength (Å)	1.1159	1.2150	1.2146	1.1579
Resolution (Å)	80-2.3	80-2.8	80-2.8	80-2.8
Observed reflections	241,259	210,895	170,071	169,776
Unique reflections	74,432	53,991	43,324	43,324
Completeness (%)	96.5 (91.4)	100 (100)	100 (100)	100 (100)
Redundancy	3.9 (4.0)	3.9 (3.8)	3.9 (4.0)	3.9 (4.0)
R_{merge}	0.088 (0.293)	0.050 (0.203)	0.046 (0.157)	0.049 (0.194)
I/σI	10.0 (2.1)	18.2 (5.4)	19.8 (7.3)	18.6 (6.1)
Refinement Statistics				
Resolution (Å)	20-2.3			
No. reflections work set	72,395			
No. reflections test set	2,256			
R_{work}/R_{free} (%)	20.8/25.4			
Protein atoms	9,613			
Solvent atoms	745			
PO_4^{2-}	4			
Wilson B ($Å^2$)	42.7			
Average B protein ($Å^2$)	39.2			
Average B solvent ($Å^2$)	44.2			
Average B PO_4^{2-} ($Å^2$)	58.3			
RMSD Bonds (Å)	0.0113			
RMSD Angle (°)	1.46			
Ramachandran distribution				
Most favored (%)	85.8			
Additionally allowed (%)	13.5			
Generously allowed (%)	0.4			
Disallowed (%)	0.3			

[1] values in parenthesis are for the highest resolution shell (2.42-2.30 Å for native and 2.95-2.80 Å for derivative)

4.2.5 Analysis of enzymatic activity

Analysis of enzymatic activity was carried out at 310 K as previously described (Yanke et al. 1998). All assays were carried out in 50 mM sodium acetate (pH 5.0), the pH optimum of the enzyme, with the ionic strength adjusted to 200 mM by the addition of NaCl. Activity assays were carried out in triplicate and repeated a minimum of two times. *Myo*-inositol hexa*kis*phosphate was purchased from Sigma-Aldrich. All other inositol phosphates were purchased from Echelon Biosciences. One unit of activity is defined as one μmol of phosphate released per minute. Steady state kinetic data for the hydrolysis of $InsP_6$ was fit to the Michaelis-Menten equation using non-linear regression (Sigma-plot 8.0; Systat Software Inc.).

4.2.6 Docking calculations

The $InsP_4$, $InsP_5$ and $InsP_6$ models were docked with the D1 and D2 repeats of PhyAmm using the Lamarckian Genetic algorithm provided by AutoDock version 4.0 (Morris et al. 1998). Flexible ligand models were derived from a refined $InsP_6$ model (1DKQ). Hydrogen atoms and Gasteiger charges were added and rotatable bonds (all exocyclic bonds) were defined within AutoDock. The resulting ligand models ($Ins(3,4,5,6)P_4$, $Ins(1,3,4,5,6)P_5$, $InsP_6$) have net charges of -4, -5 and -6 with 12, 15 and 18 rotatable torsion angles, respectively. Polar hydrogen atoms and Gasteiger charges were added to the protein (receptor) model prior to calculating grid maps in a 15 Å cube ($60 \times 60 \times 60$ points with a 0.25 Å spacing) centered about the active site. For each active site, 50 docking simulations were carried out with an initial population of 300 individuals and a maximum of 2.5×10^6 energy evaluations. Using these parameters, the docking

simulations for each ligand-receptor pair produced significant 'clusters' or groups of related structures suggesting adequate sampling of both ligand conformation and the active site. The best ligand-receptors structure for each of the docking simulation was chosen based upon the lowest energy.

4.3 Results

4.3.1 Structure of PhyAmm

PhyAmm crystallizes as a homodimer in the asymmetric unit (ASU), and has an approximate two fold non-crystallographic symmetry axis between the chains making up the dimer parallel to the b axis (Figure 4.2a). The Protein Interfaces, Surfaces and Assemblies (PISA) server (Krissinel and Henrick 2007) calculates that the surface area buried from formation of the PhyAmm dimer is 4040 $Å^2$ and indicates that the dimer in the ASU is stable. The dimer interface consists of an extensive network of hydrogen bonds and van der Waals contacts. Most of these contacts are made between structural features that are unique to PTPLPs including an extended Ω loop, a small partial β-barrel domain (IPP domain), and an extended C-terminal helix. These structures have been proposed to be involved in conferring substrate specificity in the related enzyme from *Selenomonas ruminantium*, PhyAsr (Chu et al. 2004; Puhl et al. 2007). Size exclusion chromatography indicates that PhyAmm also behaves as a dimer *in vitro*, suggesting that the crystallographic dimer is also formed in solution.

Each PhyAmm monomer that makes up the dimer in the ASU is composed of tandemly repeated PTP domains, D1 (residues 35-338) and D2 (residues 347-640) (according to the RPTP convention (Andersen et al. 2001)), connected by a flexible nine

residue linker. Each PTP core domain has a central five stranded β-sheet (β2/2', β3/3', β10/10', β4/4', β9/9') with two α-helices on one side (α1/1', α3/3') and five α-helices (α4/4', α5/5', α6/6', α7/7', α8/8') on the other (Figure 4.2b). The highly conserved strand-loop-helix active-site motif found in all PTP superfamily members is formed by β10/10' and α5/5'. The IPP domain that is unique to PTPLPs is composed of a highly twisted, five stranded partial β-barrel (β1/1', β5/5', β6/6', β7/7', β8/8') and a single α-helix (α2/2').

A search for structural homologues was carried out using the DALI server (Holm et al. 2008) and identified several members of the PTP superfamily including the related PTPLP, PhyAsr (Z score = 35.1) (Chu et al. 2004), Phosphatase and Tensin Homologue, PTEN (Z score = 12.1) (Lee et al. 1999), and Leukocyte Antigen Related Protein Tyrosine Phosphatase, LAR (Z score = 9.5) (Nam et al. 1999) but no members of the three other phytase classes. A similar DALI search using only the IPP domain (residues 46-63 & 132-202 in D1 and 344-356 & 431-500 in D2) identified a single homologous structure corresponding to the equivalent domain in the closely related PTPLP, PhyAsr, (Z = 5.3) suggesting that this domain is unique to this protein family.

4.3.2 The D1 and D2 tandem repeats adopt the same fold

The sequence identity of D1 and D2 is 36% internally and 34% and 48% with PhyAsr, respectively. Least squares superposition of the main-chain atoms of D1 and D2 with PhyAsr results in a root mean square deviation (RMSD) of 1.46 Å over 258 residues and 1.36 Å over 272 residues, respectively. A similar superposition between the repeats

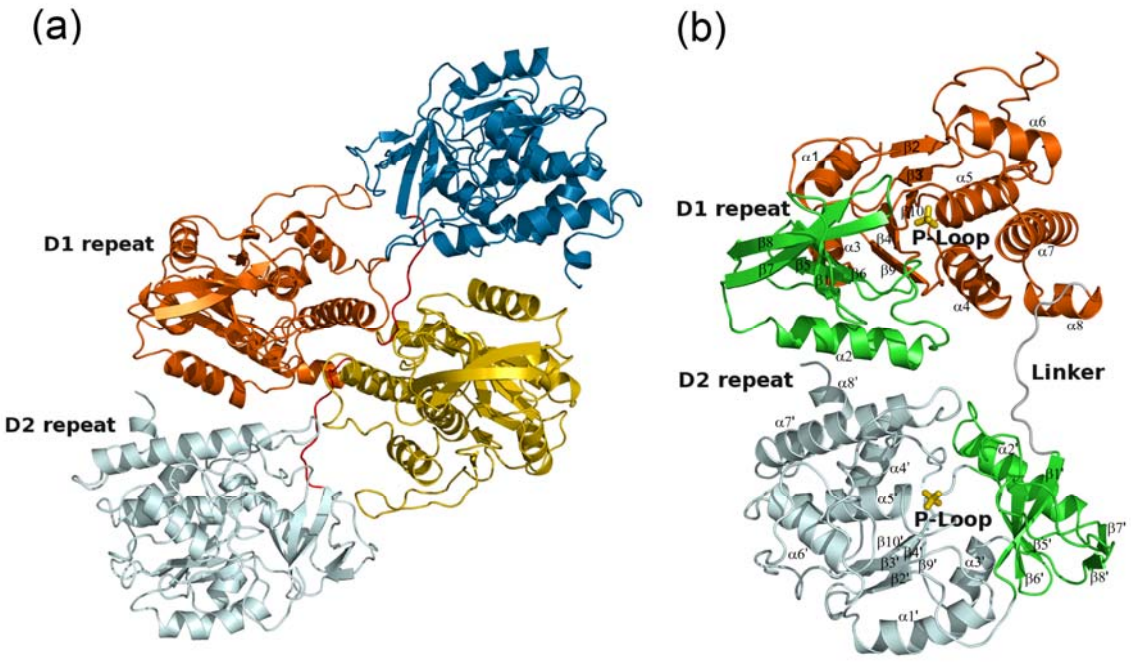

Figure 4.2: Structure of the tandemly repeated inositol polyphosphatase PhyAmm. (a). Quaternary structure of the *M. multacida* PTPLP dimer observed in the asymmetric unit. The D1 repeats are colored orange and yellow and the D2 repeats are colored light blue and blue for chains A and B, respectively. (b) Tertiary structure of PhyAmm. The tandem repeats of chain A are shown rotated by ~45° relative to their orientation in panel (a). The core PTP domains of D1 and D2 are shown in light blue and orange, respectively. The IPP domain is colored green, the linker is shown in black, and phosphate molecules found in the active site of both repeats are shown as yellow sticks.

results in an RMSD of 1.44 Å over 257 matching residues (Figure 4.3a). The regions that show the greatest deviation are unique to the PTPLP family and suggested to be involved in conferring specificity for inositol polyphosphates in these enzymes (Chu et al. 2004). These regions include β7/7' and β8/8' of the IPP domain, the Ω loop connecting β2/2' and β3/3' ($\Omega_{\beta2\text{-}\beta3}$), and the N-terminal end of α7/7'. The structural variability in the β7/7' and β8/8' β-hairpin is due to a two-residue insertion in D1 and an antiparallel G1 β-bulge involving D183, N187 and V188 that results in D184 and N187 pointing deep into the binding pocket while K185 and K186 point away from the binding pocket. The equivalent region of D2 also forms a β-hairpin but the β-strands are much shorter and

K483-K485 point towards the binding pocket to make contacts with the negatively charged phosphates on $InsP_6$. In D1, $\Omega_{\beta2\text{-}\beta3}$ is shorter than in D2 due to a seven-residue deletion and it adopts an extended conformation due to intermolecular contacts within the PhyAmm homodimer in the ASU. In D2 $\Omega_{\beta2\text{-}\beta3}$ adopts the same conformation as seen in PhyAsr, and contributes to the substrate binding pocket with K379 pointing into the active site. The variation in the N-terminal end of α7 in D1 is due to the involvement of D295, K298, N299, and Y300 in the dimer interface of the PhyAmm homodimer and a 3-residue insertion in the loop connecting α6 and α7.

4.3.3 Comparison of D1 and D2 active-sites

The size of the active-site is an important determinant of substrate specificity in the PTP superfamily (Yuvaniyama et al. 1996). Similar to PhyAsr, both the D1 and D2 repeats of PhyAmm have a very large, electropositive active site to enable the binding of the large, highly negatively charged substrate, $InsP_6$. Interestingly, comparison of the surface potentials calculated using GRASP (Nicholls et al. 1991) shows that the D1 repeat is less electropositive than the D2 repeat (Figure 4.3b). PDBsum (Laskowski et al. 1997) was used to compare the size of the active-sites of both repeats and the related PTPLP PhyAsr. The active site in the D1 repeat (volume of 4862 $Å^3$ and a depth of 13 Å) is smaller than both the active sites of D2 (volume of 7333 $Å^3$ and a depth of 17 Å) and PhyAsr (volume of 5370 $Å^3$ and depth of 14 Å). The front and sides of the binding pocket are formed by $\Omega_{\beta2\text{-}\beta3}$, α7/7', and the IPP domain, respectively. These structures are also observed in PhyAsr and make several specific contacts to the substrate (Chu et al. 2004).

The base of the active site is formed by the general acid (GA) loop (equivalent to the WPD loop in PTP1B) and the highly conserved phosphate binding loop (P-loop). In both repeats, the GA loop is in a closed conformation placing the general acid (D221 and D519) in a position equivalent to the general acid in PhyAsr, and PTEN. The P-loop contains the catalytic cysteine (C250 and C548) and the invariant arginine that coordinates the scissile phosphate for nucleophilic attack (R256 and R554). While the D1 P-loop ($HC_{250}YAGMGR_{256}T$) differs in sequence from the D2 P-loop ($HC_{548}QAGAGR_{554}T$), least squares superposition of the Cα atoms results in an RMSD of 0.22 Å indicating the P-loop conformations are essentially identical (Figure 4.3c). There are small differences in some of the P-loop side-chain conformations and the GA loop, including a 0.45 Å shift in H222 Cα and a rotation about χ_2 of 13˚ that results in a 0.81 Å movement of $N\delta_1$ away from the catalytic site in D1. In addition, Asp221 χ_2 is rotated by 26˚ resulting in a 0.64 Å movement of $O\delta_1$ into the D1 active-site. Despite these small differences, overall the structures of the active sites are essentially identical.

4.3.4 The PhyAmm tandem repeat is unique from the RPTP tandem repeats

We compared the orientation of the tandem repeat (TR) in PhyAmm to the structure of all tandemly repeated receptor PTPs (RPTP) in the protein data bank. Only RPTPs whose structure had been determined with both the D1 and D2 repeats were used in the comparison. These included CD45 (1YGU), LAR (1LAR), RPTPσ (2FH7), RPTPγ (2nlk), and RPTPε (2JJD). The orientation of the dimer is the same in all of the RPTPs with the active sites oriented ~ 90˚ relative to one another about the long axis of the TR. A six residue linker is located at the interface between the repeats and oriented

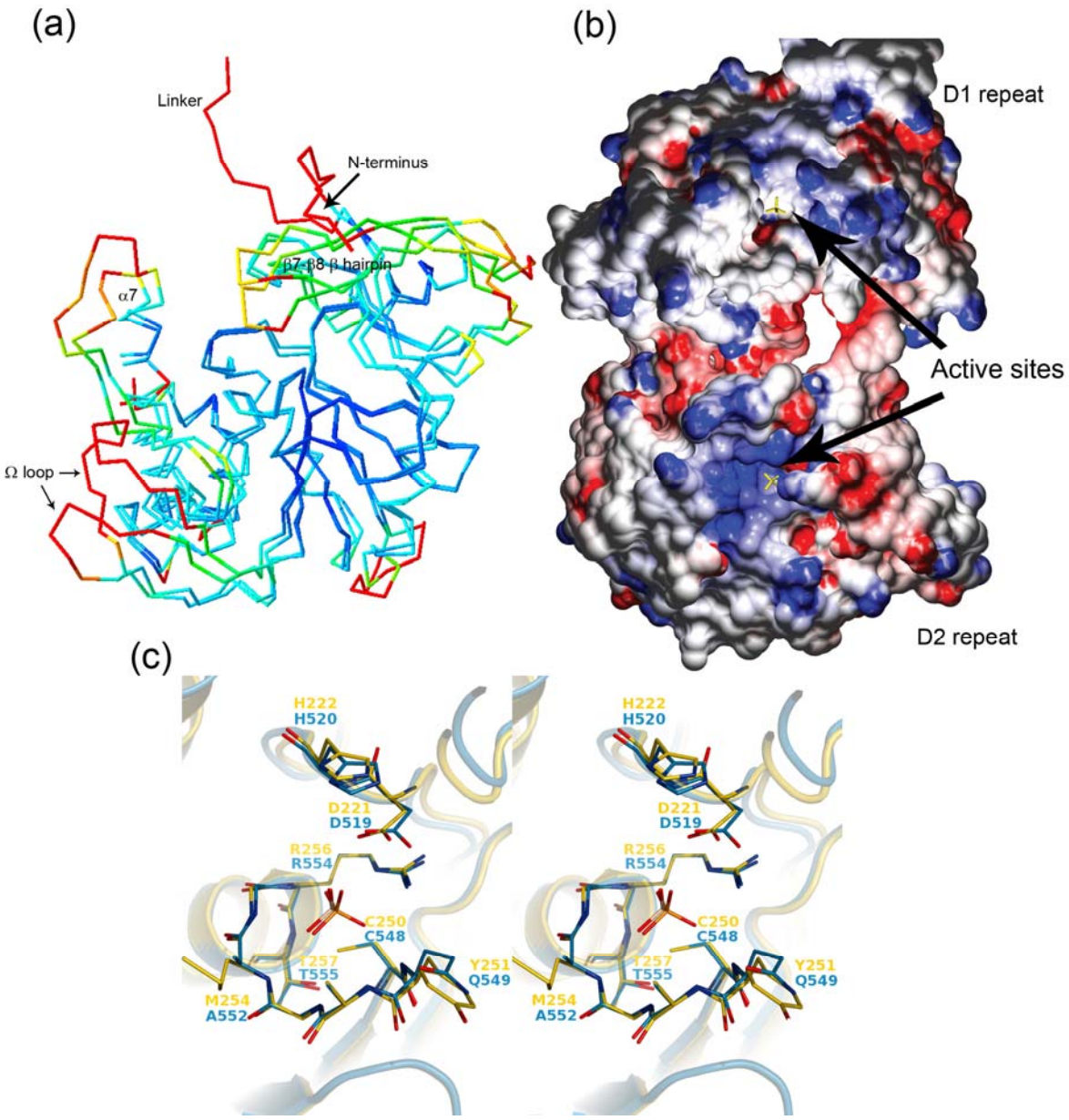

Figure 4.3: Comparison of the tandem repeats of PhyAmm. (a) Structural alignment of the D1 and D2 repeats colored according to RMSD (Dark blue; < 0.5 Å → Red; > 4.0 Å). The alignment utilized all Cα atoms and was carried out with SPDBV (Guex and Peitsch 1997). The image was created using SPDBV (Guex and Peitsch 1997). (b) Electrostatic surface potential (± 10 $kT \cdot e^{-1}$) of the PhyAmm tandem repeat calculated using GRASP2 (Petrey and Honig 2003). Electropositive regions are colored in blue and electronegative regions are colored in red. (c) Superposition of the active-site of the D1 (yellow) and D2 (blue) repeats of PhyAmm. Residues showing conformational variability or that are mutated between repeats are shown as sticks. Image is shown in divergent stereo view.

almost perpendicular to the long axis of the TR (Figure 4.4a). The tandem repeat interface (TRI) covers a surface area of ~1300-1500 $Å^2$ and is made up of an extensive network of salt bridges, hydrogen bonds, and van der Waals contacts. The conserved threonine and glutamate within the domain linker (consensus sequence G(D/E)TE(V/I/L)) are involved in specific, conserved hydrogen bonds and salt-bridges to the D2 repeat (Nam et al. 1999). In PhyAmm, the linker is orientated parallel to the long axis of the TR with the D1 and D2 active sites located on the same side (Figure 4.4b). The TRI of PhyAmm covers a surface area of ~1100 $Å^2$ and is composed of van der Waals contacts, salt bridges and hydrogen bonds between α1 and α7'. Interestingly, unlike RPTPs, no residues in the linker (N_{339}SWEPDYN) make specific contacts to either repeat at the TRI.

Structural superposition of the D1 and D2 repeats of PhyAmm onto the RPTP domains reveals that PhyAmm cannot adopt the RPTP tandem dimer due to significant steric clashes between $\Omega_{\beta2\text{-}\beta3}$ and α7. The related PTPLP PhyAsr also exists as a dimer both in solution and crystallographically (Puhl et al. 2007). However, unlike PhyAmm, the PTPLP domains in PhyAsr are not covalently linked. The PhyAsr dimer is unique from the RPTP and PhyAmm tandem repeats due to steric clashes between $\Omega_{\beta2\text{-}\beta3}$ and the extended C-terminal helix (equivalent to α7) in the RPTP TR, and steric clashes between the N-terminus and the IPP domain in the PhyAmm TR. Additionally, the N and C termini of the repeats of PhyAmm are 46 Å apart in the PhyAsr dimer, which would require a large insertion to enable a covalent linkage between the repeats.

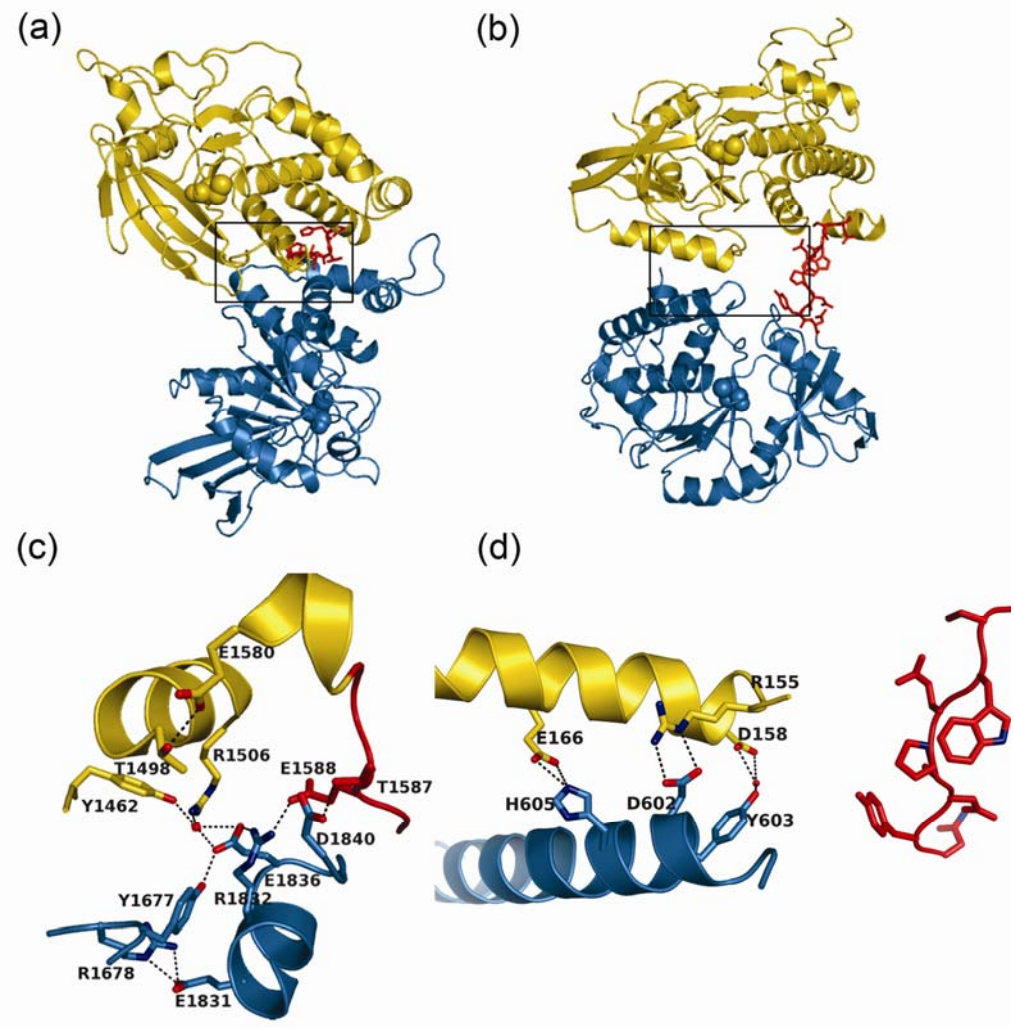

Figure 4.4: Comparison of the tandem repeats in PhyAmm and receptor PTPs. (a) Structure of the receptor PTP LAR tandem repeat (Nam et al. 1999). (b) Structure of the PhyAmm tandem repeat. Close-up of the boxed region showing the tandem repeat interface of (c) LAR and (d) PhyAmm. In panels (a) and (b), the catalytic cysteines are shown as spheres to illustrate the relative rotation of the D1 and D2 repeats. For clarity, the closeup of the LAR dimer interface is rotated ~ 45˚ about the long axis of the repeat relative to the orientation shown in panel (c). Residues making productive interactions are shown as sticks. Stabilizing interactions are shown as dashed lines. In all panels, the D1 repeat is colored yellow, the D2 repeat is colored blue and the repeat linker is shown in red.

4.3.5 Catalytic activity of PhyAmm D1 and D2 repeats

In many RPTPs, only the N-terminal, D1, repeat is catalytically active (Andersen et al. 2001). To test whether the PhyAmm repeats both possess catalytic activity, we

used site-directed mutagenesis to change the catalytic cysteine to a serine in the D1 and D2 repeats (C250S and C548S, respectively) and then tested the mutants for catalytic activity against $InsP_6$. The k_{cat} and K_m values for the hydrolysis of $InsP_6$ by wild type PhyAmm are 1109 ± 64 s^{-1} and 347 ± 48 μM, respectively. Mutation of the catalytic cysteine in D1 (C250S) had little effect on the rate of $InsP_6$ hydrolysis (k_{cat}, 1048 ± 48 s^{-1}) but resulted in a slight increase in K_m (K_m, 459 ± 35 μM). Mutation of the cysteine in D2 (C548S) completely abolished enzyme activity. Even after increased incubation times with the C548S mutant there was no detectable catalytic activity above the limit of detection for our assay system (0.003 μmol of phosphate). These results indicate that the D2 repeat is solely responsible for hydrolyzing $InsP_6$ and suggest that the D1 repeat is catalytically inactive against this substrate. It is noteworthy that in contrast to PhyAmm, the D2 repeat is catalytically inactive in RPTPs, and that the lack of activity is due to mutation of catalytic residues. All of the catalytic residues are conserved in both PhyAmm repeats indicating that another feature of PhyAmm is responsible for the lack of activity against $InsP_6$ in the D1 repeat.

4.3.6 The PhyAmm D1 repeat is active against lower inositol phosphates

Although D1 does not hydrolyze $InsP_6$, there is no obvious reason that it cannot carry out catalysis and it may be possible that the specificity of the two repeats differ. To examine whether the D1 repeat of PhyAmm has catalytic activity against lower inositol polyphosphates (i.e. less than 6 phosphates) we tested the catalytic activity of PhyAmm against several of the numerous $InsP_5$, $InsP_4$, and $InsP_3$ substrates (Table 4.2). Removal of the phosphate at the 2-position of $InsP_6$ is sufficient to observe catalytic activity in the

D1. This confirms the structure-based prediction that essential active-site residues are in a catalytically competent conformation. This repeat displayed a specific activity of 8 units/mg for Ins(1,3,4,5,6)P_5, ~ 5 times greater than the limit of detection of the assay. In contrast, the specific activity for Ins(1,3,4,5,6)P_5 compared to $InsP_6$ is approximately 1.4 fold lower in the D2 repeat. Removal of an additional phosphate resulted in a further increase in catalytic activity with the D1 repeat exhibiting a specific activity of 44 units/mg and 54 units/mg towards Ins(1,2,4,5)P_4 and Ins(3,4,5,6)P_4, respectively. Interestingly, the presence of a phosphate at the 2-position of Ins(1,2,4,5)P_4 hinders the catalytic activity of D1 supporting the hypothesis that this position is important for the differential specificity of the D1 and D2 repeats. Removal of an additional phosphate to $InsP_3$ did not result in a further increase in the catalytic activity of the D1 repeat. In all cases the D2 repeat displays catalytic activity against the lower inositol phosphates with the D1 repeat having at most 22 % relative activity against Ins(1,4,5)P_3. Interestingly, the catalytic activity in wild type PhyAmm does not equal the sum of the D1 and D2 activities possibly because the D1 and D2 work together in the wild type enzyme. Despite the higher level of catalytic activity in D2, it is clear that the D1 repeat is catalytically active and displays a marked preference for lower phosphorylated inositols. In contrast, the D2 repeat displays a preference for highly phosphorylated inositols indicating the D1 and D2 repeat have evolved different substrate specificities.

Table 4.2: Catalytic activity of wild type, C250S (D1 repeat) and C548S (D2 repeat) mutants of PhyAmm against lower order inositol phosphates.

Substrate	C250S (Units[a]/mg)	C548S (Units/mg)	Wild type (Units/mg)
Ins(1,4,5)P_3	183 ± 11	53 ± 2	237 ± 14
Ins(1,3,4)P_3	237 ± 16	53 ± 2	335 ± 23
Ins(1,2,4,5)P_4	225 ± 14	44 ± 3	316 ± 19
Ins(3,4,5,6)P_4	246 ± 19	54 ± 4	375 ± 26
Ins(1,3,4,5,6)P_5	449 ± 28	8 ± 0.6	514 ± 43
Ins(1,2,3,4,5,6)P_6	608 ± 38	1 ± 0.5[b]	836 ± 52

[a] Unit = μmol phosphate released per minute per mg of enzyme

[b] This is at the limit of detection of this assay

4.3.7 Basis for the different specificity in D1 and D2 inferred from docking calculations

We used an *in silico* approach to understand the different substrate specificities of the D1 and D2 repeats (Table 4.3). Docking $InsP_6$, Ins(1,3,4,5,6)P_5 and Ins(3,4,5,6)P_4 into the active sites of both repeats suggests that the D1 repeat cannot bind $InsP_6$ in a catalytically competent conformation due to steric interactions with the β7-β8 hairpin (Figure 4.5a). The $InsP_6$ is placed near the edge of the binding pocket in D1 with the closest phosphate being 6.32 Å from the nucleophile. Docking of Ins(1,3,4,5,6)P_5 and Ins(3,4,5,6)P_4 placed the substrate 3.11 Å and 3.20 Å from the cysteine (Figure 4.5b,c). In contrast, D2 can bind all three substrates in a conformation that supports catalysis (Figure 4.5d-f). The binding energies calculated by AutoDock agree with the substrate specificity observed *in vitro* (Table 4.2). The binding energy for $InsP_6$ is 3 kcal/mol lower in the D2 repeat than in the D1 repeat. This difference in binding energy likely reflects the different positioning of $InsP_6$ in the D1 and D2 repeats. When $InsP_5$ was used in the docking, both D1 and D2 show very similar binding energies (-6.2 kcal/mol versus

-5.9 kcal/mol, respectively). In contrast, docking of $InsP_4$ suggests that the D1 repeat shows an energetic preference for binding $InsP_4$ relative to the D2 repeat (-8.1 kcal/mol versus -7.2 kcal/mol, respectively). The differences in electrostatic surface potential and steric constraints from the smaller active-site in D1 likely both account for the binding differences between the active-sites.

Table 4.3: Docking of inositol polyphosphates into the D1 and D2 active-sites. The binding energy was calculated by Autodock 4.0. The cluster size represents the number of docked structures that are equivalent (all atom RMSD < 2.0 Å) to the lowest energy docking from 50 separate Lamarckian Genetic Algorithm trials. The best ligand-receptor structure for each of the docking simulations was chosen based upon the lowest binding energy. The S to P distance is the distance (Å) from the catalytic cysteine Sγ to the closest phosphate on the substrate.

Docking	Binding Energy (kcal/mol)	Cluster size	S to P distance (Å)
D1 $InsP_6$	-5.95	13	6.32
D2 $InsP_6$	-8.95	6	3.47
D1 $InsP_5$	-6.21	24	3.11
D2 $InsP_5$	-5.91	12	3.13
D1 $InsP_4$	-8.06	9	3.20
D2 $InsP_4$	-7.19	19	3.14

To examine if the β7-β8 β-hairpin or differences in the P-loop sequence is responsible for preventing hydrolysis of $InsP_6$ in the D1 repeat of PhyAmm, we constructed four separate site-directed mutants (in all cases the mutants had the D2 catalytic cysteine mutated to serine to abolish D2 activity). The Y251Q and M254A mutants restore residues observed in the P-loop of D2, the deletion of the β7-β8 β-hairpin loop (residues 183-187) was aimed at increasing the size of the D1 active site, and a double mutant containing both the loop deletion and the Y251Q mutation were produced. Only mutants deleting the β7-β8 β-hairpin loop produce a detectable signal in our assay, however the signal is sufficiently close to the limits of

detection that we cannot conclude the mutants are active towards $InsP_6$. These results suggest that additional structural determinants including the difference in electrostatic surface potential, mutations in the vicinity of the active site, and steric constraints are likely all involved in the altered substrate specificity of D1.

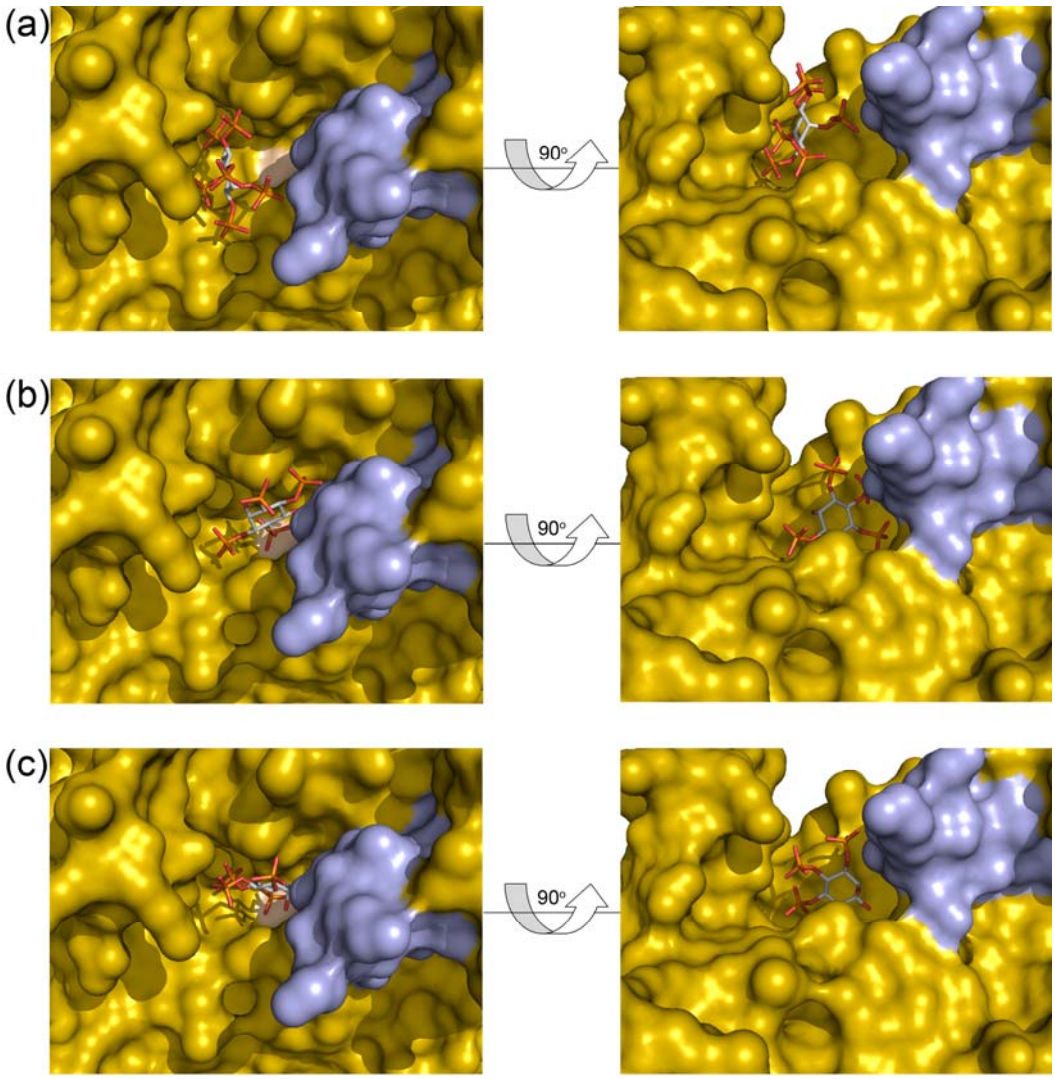

Figure 4.5: Docking of inositol phosphates in the D1 and D2 active sites of PhyAmm. Docking Ins(1,2,3,4,5,6)P_6, Ins(1,3,4,5,6)P_5, and Ins(3,4,5,6)P_4 into the D1 (a, b, c respectively) active sites reveals that Ins(1,2,3,4,5,6)P_6 cannot fit in D1. The active site is shown as a solid surface and colored yellow. The β7- β8 loop is highlighted in light blue and the catalytic cysteine is shown in beige. The docked inositol phosphate is shown in sticks.

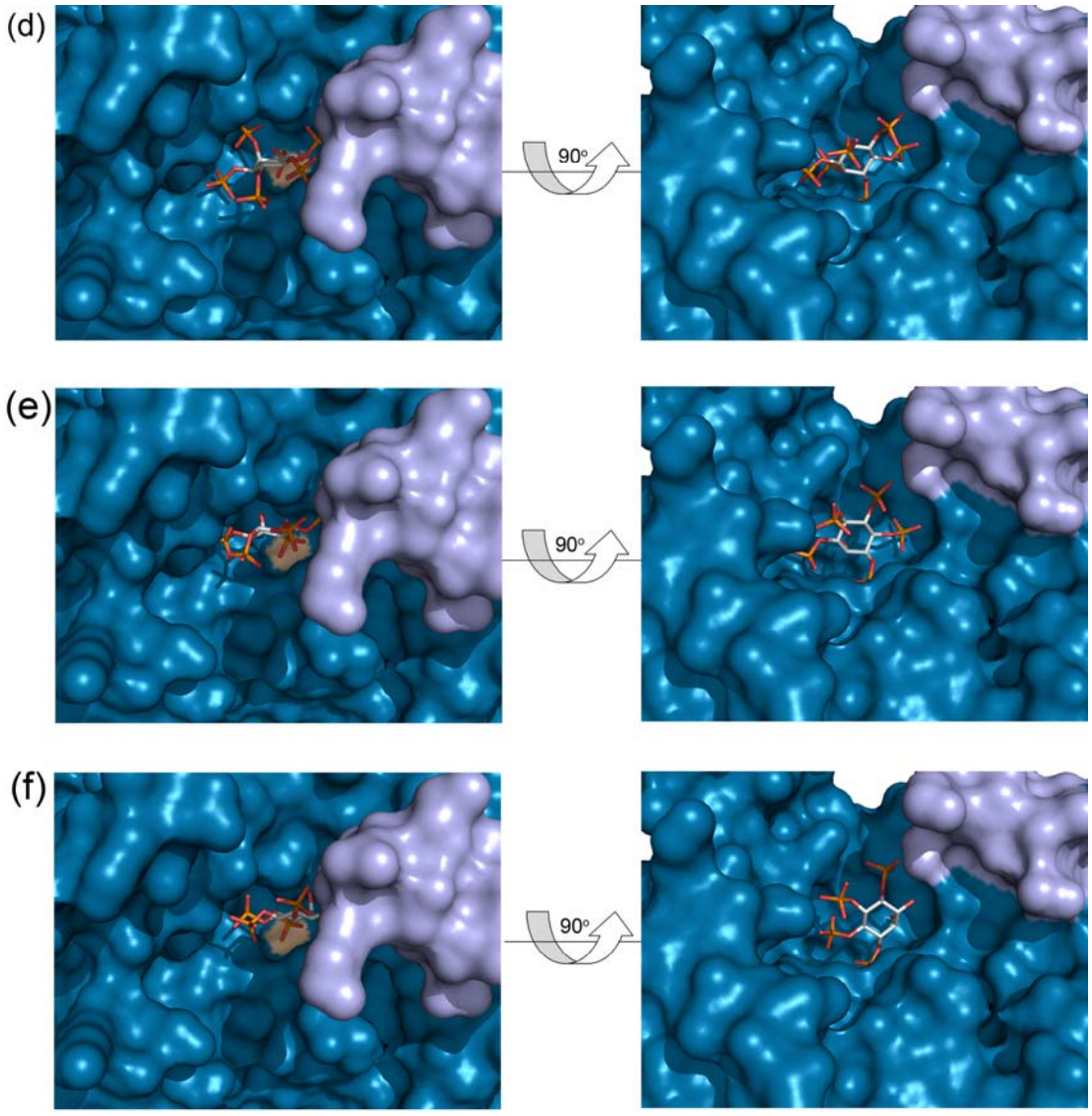

Figure 4.5 continued: Docking Ins(1,2,3,4,5,6)P_6, Ins(1,3,4,5,6)P_5, and Ins(3,4,5,6)P_4 into the D2 (d, e, f respectively) active sites. The active site is shown as a solid surface and colored blue. The β7- β8 loop is highlighted in light blue and the catalytic cysteine is shown in beige. The docked inositol phosphate is shown in sticks.

4.4 Discussion

4.4.1 The structure of PhyAmm

This work represents the first structural study carried out on a tandemly repeated inositol polyphosphatase. Similar to the other known PTPLP (PhyAsr), PhyAmm has unique structural features that contribute to the observed specificity and structural

differences amongst PTPLPs, and between PTPLPs and other members of the PTP superfamily. Like PhyAsr, PhyAmm is a homodimer both crystallographically and in solution. However, due to steric constraints the association of the PTPLP folds in these enzymes are distinct showing that the organization of PTPLP folds within this enzyme family is flexible. Interestingly, tandemly repeated PTP domains are common within the RPTP family and the occurrence of this feature in PTPLPs provides an additional level of similarity between these classes of enzyme. In all RPTPs structurally characterized to date, the association of PTP domains in the TR is structurally equivalent, but significantly different from that observed in PhyAmm and PhyAsr (Figure 4.4). Interestingly, in both PhyAmm and PhyAsr, all of the active sites are presented on one side of their respective dimers. Given that PhyAmm and PhyAsr are known to associate with the outer membrane of *M. multacida* and *S. ruminantium in vivo* (D'Silva et al. 2000), the arrangement of active sites is likely functionally significant.

4.4.2 PhyAmm D1 active site

Many tandemly repeated RPTPs have a catalytically inactive C-terminal (D2) repeat due to mutations to catalytic residues. In several cases, the mutations are conservative and do not involve the catalytic cysteine. RPTPα, RPTPε, RPTPσ, RPTPδ and LAR have conservative mutations in two essential motifs: 1) the aspartate in the WPD loop is mutated to glutamate and 2) a conserved tyrosine that is important for substrate binding is mutated to valine or leucine. Mutation of these residues back to the consensus sequence restores full catalytic activity to this repeat in RPTPα, RPTPε and LAR (Buist et al. 1999; Lim et al. 1999; Nam et al. 1999). In contrast to RPTPs, the

different catalytic properties of the PhyAmm repeats are likely a result of steric constraints on substrate binding caused by a decrease in the size of the binding pocket (Figure 4.5). Interestingly, RPTPα, RPTPε, RPTPσ, and LAR also have smaller binding pockets in the inactive repeat. Additionally, differences in the electrostatic potential of the D1 active-site may contribute to the lack of activity toward $InsP_6$ (Figure 4.3). Electrostatic interactions with substrate are important in PhyAsr (Puhl et al. 2007) and likely play an equally important role in the catalytic activity of PhyAmm. Interestingly, the electrostatic surface potential of the inactive repeats of RPTPα, RPTPε, and RPTPσ are also less electropositive than the catalytically active repeat.

The active site of PhyAmm is largely composed of structural elements that are unique to PTPLPs. These same regions are sites of an insertion (β7-β8 β-hairpin loop) and deletions (Ω loop and α7 helix) in a sequence alignment of the D1 and D2 repeats (Figure 4.6). These regions also display the largest deviations in a least-squares structural superposition of the repeats (Figure 4.3a). The inositol phosphatase activity of several PTPLPs has recently been reported (Puhl et al. 2007; Puhl et al. 2008a; Puhl et al. 2008b; Puhl et al. 2009) and can be separated into low activity and high activity forms. Interestingly, each of the low activity PTPLPs (*S. ruminantium* subsp. l*actolytica* and *S. lacticifex,*) have a similar pattern of insertions and deletion to those observed in D1 (Figure 4.6). This suggests the substrate specificity of D1 and low activity PTPLPs depend upon each of these structural elements.

4.4.3 PhyAmm D1 activity

While it is impractical to assess D1 activity against all possible *myo*-inositol polyphosphates, it is clear that D1 exhibits catalytic activity against several less phosphorylated forms, in particular $InsP_4$ and $InsP_3$. At the same time, the D2 repeat preferentially hydrolyzes $InsP_5$ and $InsP_6$, the highly phosphorylated forms (Table 4.2). Docking studies suggest the axial C2 phosphate is at least partly responsible for the inability of D1 to hydrolyze $InsP_6$. This is supported by the observation that this repeat displays catalytic activity against Ins(1,3,4,5,6)P_5 (Table 4.2). The eukaryotic IPPase, PTEN shows high levels of catalytic activity against Ins(1,3,4,5,6)P_5, but no activity against $InsP_6$ due to steric clashes between the protein and the axial phosphate at the two position (Caffrey et al. 2001). This is an important finding and suggests that the axial C2 phosphate position can serve to distinguish inositol phosphate pools in the cell, ensuring that $InsP_6$, the most abundant cellular inositol phosphate, does not interfere with signaling pathways involving other inositol phosphates. Furthermore, the activity of PhyAmm towards *myo*-inositols lacking a C2 phosphate suggests that related PTPLPs expressed by pathogenic bacteria may directly affect *myo*-inositol signaling pathways.

What function might the low IPP repeat (D1) play in PhyAmm? Phytate degrading inositol polyphosphatases hydrolyze $InsP_6$ by the sequential, stepwise removal of phosphates to generate lower order inositol polyphosphates that can serve as substrate for the removal of additional phosphates in an ordered pathway. Many acid phosphatases (Konietzny and Greiner 2002) and all characterized PTPLPs (Puhl et al. 2007; Puhl et al. 2008a; Puhl et al. 2008b; Puhl et al. 2009) can remove all of the equatorial phosphates to yield inorganic phosphate and Ins(2)P_1. The presence of repeats with specificity

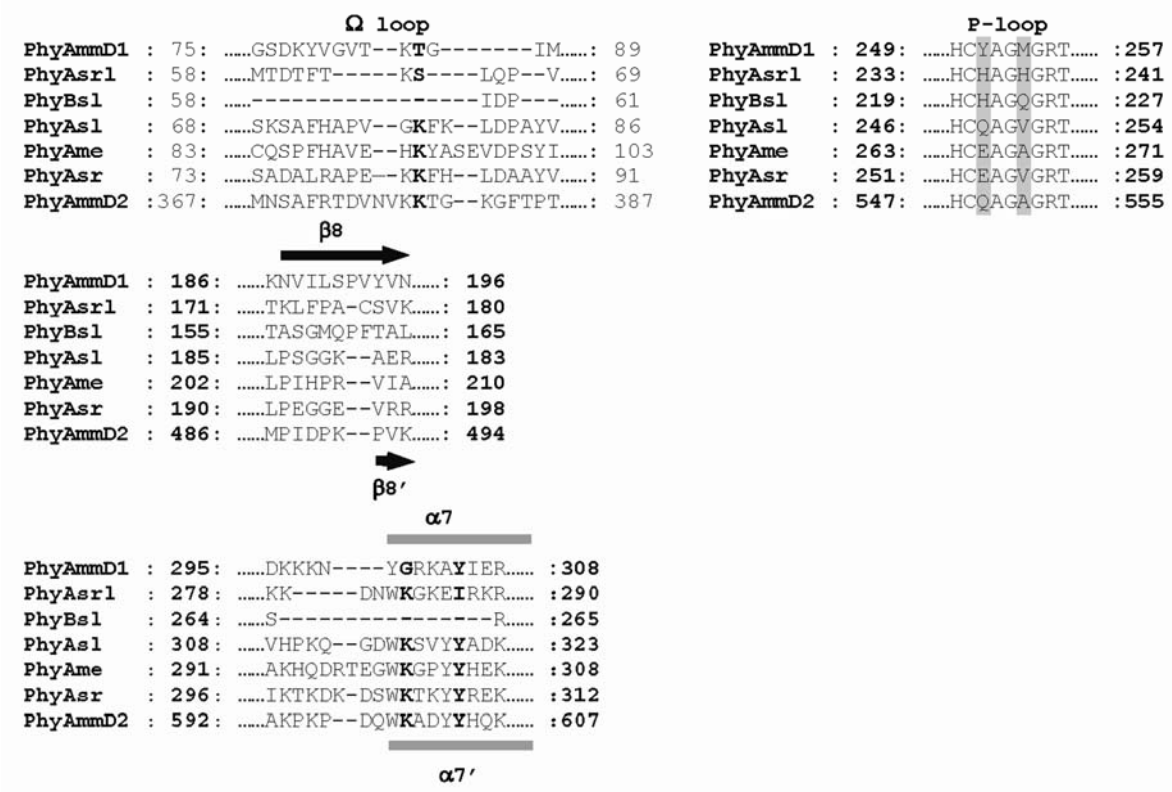

Figure 4.6: ClustalW alignment of the variable regions of bacterial protein tyrosine phosphatase-like phytases. Numbers at the beginning and end of each sequence represent the residue number for the first and last amino acid in that sequence, respectively. Residues in PhyAsr that have been identified as being involved in substrate binding are highlighted in bold and the variable positions in the phosphate binding loop are highlighted in light grey. Secondary structures of the D1 and D2 repeats of PhyAmm are shown on the top and bottom of the alignment, respectively. The protein abbreviations, source, and GeneBank accession number are as follows: PhyAmmD1, *M. multacida* D1 repeat, ABA18187; PhyAsrl, *S. ruminantium* subsp. *lactylitica*, ABC69359; PhyBsl, *S. lacticifex*, ABC69361; PhyAsl, *S. lacticifex*, ABC69367; *Measphaera elsdenii*, ABC69358; PhyAsr, *S. ruminantium*, AAQ13669; PhyAmmD2, *M. multacida* D1 repeat, ABA18187.

for both highly (D2) and less phosphorylated *myo*-inositols (D1) suggests this enzyme may have evolved a 'divide and conquer' approach to hydrolyzing the numerous *myo*-inositol derivatives. In this mechanism, D2 targets and hydrolyzes highly phosphorylated IPPs including $InsP_6$ in an ordered pathway. At the same time, D1 targets less phosphorylated IPPs including IPP intermediates released by D2 and unrelated $InsP_4$,

$InsP_3$ and $InsP_2$ molecules present in the environment. This would enable the enzyme to hydrolyze multiple IPP species simultaneously which presumably is more efficient than hydrolyzing a single species at a time. Interestingly, a highly conserved eukaryotic enzyme possessing both inositol polyphosphate kinase and phytase domains has recently been described (Mulugu et al. 2007). The presence of a dual domain enzyme with both inositol kinase and phosphatase activity, and an enzyme with domains that exhibit unique IPP phosphatase activity, suggest that dual domain proteins may play important roles in inositol phosphate metabolism.

4.4.4 PhyAmm and tandemly repeated RPTPs

What is the biological advantage of having tandemly repeated PTPLP or PTP folds that exhibit different specificities? In the case of PhyAmm, it may function in a 'divide and conquer' mechanism of *myo*-inositol hydrolysis. Due to the higher level of activity in the D2 repeat, the D1 repeat likely only functions catalytically when significant levels of lower phosphorylated IPPs are present. Alternatively, the D1 repeat may function to target an unidentified substrate *in vivo* or, like RPTPs, play a role in substrate recognition and/or recruitment. Given the association of PhyAmm with the outer membrane of *M. multacida in vivo* (D'Silva et al. 2000), the D1 repeat may act as a membrane anchor or target a membrane phospholipid.

In addition to substrate recognition, targeting, or regulation (Ng et al. 1995; Kashio et al. 1998; Blanchetot and den Hertog 2000; Yang et al. 2007), our results suggest another role for the D2 repeat in RPTPs retaining all essential catalytic residues. These repeats may not be catalytically inactive but rather possess a unique specificity for

substrates that have yet to be identified. The differences between the active and inactive repeats of RPTPs are remarkably similar to those observed between the less active and more active repeats of PhyAmm. In particular, mutations, insertions and deletions within loops that form the active site reduce the volume of the active site and significantly alter the electrostatic surface potential making it less electropositive. While these differences are likely to alter substrate specificity, the underlying catalytic machinery remains intact and may well retain some form of activity against a substrate that has yet to be identified.

Chapter 5. Examination of Substrate Binding Provides Mechanistic Insight into Inositol Phosphate Degradation by Protein Tyrosine Phosphatase-like Phytases.

5.1 Introduction

Inositol polyphosphates (IPPs) are ubiquitous in nature and are involved in a number of important cellular signaling events (Irvine and Schell 2001). *Myo*-inositol 1,2,3,4,5,6-hexa*kis*phosphate (Phytic acid, $InsP_6$) is the most abundant cellular inositol phosphate and plays a central role in numerous essential cellular processes including DNA repair, RNA processing, mRNA export, plant development, apoptosis, and bacterial pathogenicity (York et al. 1999; Hanakahi et al. 2000; Chatterjee et al. 2003; Macbeth et al. 2005; Tan et al. 2007; Majerus et al. 2008). Inositol phosphatases (IPPases) that degrade $InsP_6$ are ubiquitous in nature and have been identified in prokaryotes, protists, fungi, animals, and plants and are generically referred to as phytases (Caffrey et al. 1999; Mullaney et al. 2000; Rao et al. 2009). To date four classes of phytases have been described, and these include, histidine acid phosphatases (HAPs), β-propeller phytases (BPPs), purple acid phosphatases (PAPs) and protein tyrosine phosphatase-like phytases (PTPLPs) (Mullaney and Ullah 2003; Chu et al. 2004; Puhl et al. 2007). Phytases catalyze the stepwise removal of phosphates to generate lower inositol phosphates in a highly ordered, specific manner (Konietzny and Greiner 2002). The intermediates

generated during the breakdown of $InsP_6$ are released from the enzyme and serve as substrates for subsequent cycles of hydrolysis. The mechanism by which phytases bind and discriminate between the various lower inositol phosphates has not been structurally characterized and is largely unknown. Although all of the phytase classes are functionally similar, they are structurally and mechanistically diverse.

The PTPLP class of phytases possesses a protein tyrosine phosphatase (PTP) active-site signature sequence ($CX_5R(S/T)$), is structurally similar to PTPs and utilizes a classical PTP reaction mechanism (Chu et al. 2004; Puhl et al. 2007). Interestingly, these enzymes display no catalytic activity against classical PTP substrates due to several unique structural features that confer specificity for IPPs (Puhl et al. 2007; Puhl et al. 2008a; Puhl et al. 2008b; Puhl et al. 2009). Although the biological function of these enzymes is unclear, they have been found in a wide range of bacteria including plant and human pathogens (Lim et al. 2007; Nakashima et al. 2007).

The structural basis of IPP binding has been investigated in HAPs and BPPs. The substrate binding pocket in HAPs is a deep electropositive cleft that makes extensive contacts to the base of the substrate (Lim et al. 2000). Substrate binding in the *Escherichia coli* phytase is accompanied by a conformational change that stabilizes the enzyme substrate complex (Lim et al. 2000). The substrate binding site in BPPs is an electronegative, solvent filled channel that binds seven Ca^{2+} ions. The Ca^{2+} ions are essential for both catalytic activity and substrate binding (Oh et al. 2006). The structure of a BPP in complex with substrate has not been solved but the identification of phosphate ions in the binding pocket enabled $InsP_6$ to be modeled into the active site (Shin et al. 2001). BPPs have two binding sites, a cleavage site that binds the scissile

phosphate and an affinity site that enhances substrate binding (Shin et al. 2001). Only substrates that simultaneously fill both binding sites are hydrolyzed by BPPs, explaining why these enzymes can only remove every second phosphate from $InsP_6$ (Kerovuo et al. 2000). The structure of the PTPLP from *Selenomonas ruminantium* (PhyAsr) bound to the competitive inhibitor *myo*-inositol hexasulfate (MIHS) has been determined (Chu et al. 2004). Like HAPs, PhyAsr has a deep, electropositive substrate binding site and makes extensive contacts to MIHS (Chu et al. 2004). Interestingly, the PhyAsr-MIHS model is inconsistent with the $InsP_6$ degradation pathway of PhyAsr (Puhl et al. 2007) suggesting that the binding modes of substrate and inhibitor are unique.

In this chapter, we describe the high resolution structures of PhyAsr in complex with several IPPs. Our structural studies reveal that: 1) substrate binds in multiple conformations in the binding pocket, and 2) substrate binding induces strain on the ligand. Binding studies reveal that the K_d for $InsP_6$ is much lower than the K_m suggesting PhyAsr does not utilize a Michaelis-Menten mechanism. Preliminary results from a rapid kinetics study of substrate binding is biphasic binding involving a rapid first phase followed by a slower phase. Based on our binding and structural studies and the results of previous studies (Puhl et al. 2007), we propose a reaction mechanism in which substrate binding occurs in two steps: 1) initial binding, and 2) substrate reorientation to position it in a catalytic conformation. An examination of the catalytic properties of other classes of $InsP_6$-degrading enzymes suggests that this reaction mechanism may be shared by other $InsP_6$-degrading enzymes.

5.2 Experimental procedures

5.2.1 Cloning and mutagenesis

The pET28b expression construct from Puhl et al (2007) containing the full-length *S. ruminantium* ORF (minus the putative signal peptide) was used as the template for all cloning and mutagenesis procedures. The signal peptide sequence was determined using SignalP 3.0 (Nielsen et al. 1997; Bendtsen et al. 2004). PhyAsr numbering begins with 1 at the N-terminus of the protein sequence found in GeneBank (AAQ13669) including the predicted 27 residue signal peptide. Mutant proteins C252S, and C252A, H188C, H188C/C252S were prepared by site direct mutagenesis using counter PCR amplification of the expression plasmid as previously described (Street et al. 1991). To verify the identity of the construct and the presence of the desired mutations, all PCR products were sequenced at the University of Calgary Core DNA and Protein services facilities. Sequence data was analyzed using the programs SEQUENCHER 4.0 (Gene Codes Corp.) and MacDNAsis 3.2 (Hitachi Software Engineering Co.). All of the primers used in this study were purchased from Integrated DNA technologies.

5.2.2 Protein production and purification

Protein expression was carried out in *E. coli* BL21(DE3) cells (Novagen) for 18 hours at 310 K. Cells were grown to an optical density of 0.6-0.8 and protein expression was induced by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Induced cells were harvested and resuspended in lysis buffer: 20 mM KH_2PO_4 (pH 7.0), 300 mM NaCl, 1 mM β-mercaptoethanol (BME), 15 mM imidazole (pH 8.0). Cells were lysed by sonication and cell debris was removed by

centrifugation at 20 000 × g. PhyAsr was purified to homogeneity by metal chelating affinity chromatography (Ni^{2+}-NTA-agarose, Qiagen). Bound protein was washed with lysis buffer containing 15 mM imidazole (pH 8.0) and eluted with lysis buffer containing 400 mM imidazole (pH 8.0). PhyAsr was further purified by cation exchange (Macro-Prep High S) and size exclusion chromatography. The homogeneity of the purified protein was confirmed by SDS-polyacrylamide gel electrophoresis (Laemmli 1970), and Coomassie Brilliant Blue R-250 staining. Protein was either used immediately or dialyzed into 20 mM ammonium bicarbonate (pH 8.0) and lyophilized for long term storage.

5.2.3 Crystallization and ligand soaking

Crystallization experiments were carried out using sitting-drop vapor diffusion with drop ratio of 2 μL of 20 mg/mL protein solution and 2 μL of reservoir. Crystals were grown in 8 – 10 % PEG 8000, 200 – 300 mM NaCl, 50 mM sodium acetate pH 4.8. Ligand solutions were prepared by dissolving Ins(1,2,3,4,5,6)P_6 ($InsP_6$) and Glyceroinositol(3,4,5)P_3 (G3P) in mother liquor to a final concentration of 2 mM and 1 mM, respectively. $InsP_6$ and G3P were purchased from Sigma-Aldrich and Echelon Biosciences, respectively. Ligand solutions were added to drops containing crystals and incubated for approximately one hour prior to freezing. Crystals were cryo-protected using a solution containing the crystallization reagents, ligand, and 25 % glycerol. Cryo-protected crystals were flash frozen in liquid nitrogen.

5.2.4 Data collection, image processing and structure refinement

Diffraction data (λ = 1.1159 Å) for all of the complexes was collected at 100 K on beamline 8.3.1 at the Advanced Light Source. Diffraction images were processed with MOSFLM (Leslie 1992) and scaled with SCALA (Evans 1997) within the CCP4 program suite (CCP4 1994). The D223N structure (PDB 2B4P) (Puhl et al. 2007) was used to solve the structures of the PhyAsr ligand complexes. The isomorphous nature of the crystals allowed us to use the coordinates of 2B4P to calculate phases with the program SFall (CCP4 1994). CNS 1.2 (Brunger 2007) was used to iteratively refine the models using simulated annealing, positional and B-factor refinement followed by manual model building in XFIT (McRee 1999). Unless indicated otherwise, figures were prepared with PyMOL (version 0.99) (DeLano 2008). Topology and parameter files for the ligands were generated using the structure of $InsP_6$ bound to the *E. coli* phytase (1DKQ) as a model. For the refinement of the $InsP_5$ and $InsP_4$ complexes, these files were edited to reflect the structures of the ligands. PROCHECK was used throughout refinement to assess the stereochemistry of the model (Laskowski et al. 1993). Statistics for the data collection and refinement are shown in Table 1.

5.2.5 Fluorescent labeling

150 nmol of PhyAsr H188C or PhyAsr H188C/C252S was batch bound to Ni-NTA resin (Qiagen) in binding buffer (20 mM Tris-HCl (pH 8.0), 100 mM NaCl, 10 mM βME). Bound protein was buffer exchanged with 4 column volumes of labeling buffer (20 mM Tris-HCl (pH 8.0), 100 mM NaCl, 10 % Glycerol). A twenty fold molar excess

of 5-iodoactamidofluorescein (5-IAF) was added to 2 ml of labeling buffer and added to the column. The resin was incubated for 48 h at 310 K with gentle agitation and then

Table 5.1: Data collection and refinement statistics for PhyAsr in complex with Ins(1,2,3,4,5,6)P_6, Ins(1,3,4,5)P_4, and Ins(1,2,3,5,6)P_5[1].

	PhyAsr C252S $InsP_6$	PhyAsr C252S $InsP_4$	PhyAsr C252A $InsP_5$
Data collection			
Space group	$P2_1$	$P2_1$	P21
Unit cell a, b, c (Å) /β (°)	46.1, 136.9, 79.9/102.9	46.0, 137.2, 80.0/103.0	46.0, 137.3, 80.0/103.0
Wavelength (Å)	1.1159	1.1159	1.1159
Resolution (Å)	51 - 1.60	51 - 1.85	37.5 - 1.55
Observed reflections	560,803	382,424	456,682
Unique reflections	120,565	77,809	126,093
Completeness (%)	92.9 (77.0)	94.1 (89.4)	89.9 (56.8)
Redundancy	4.6 (4.2)	4.9 (5.0)	3.6 (2.8)
R_{merge}	0.083 (0.76)	0.12 (0.75)	0.074 (0.38)
I/σI	10.7 (1.5)	8.4 (2.0)	15.5 (2.8)
Refinement Statistics			
Resolution (Å)	50 - 1.60	51 - 1.85	50 - 1.55
No. reflections work set	113,134	72,898	119,687
No. reflections test set	6,015	3,894	6,346
R_{work}/R_{free} (%)	17.0/18.5	16.2/17.0	16.4/16.9
Protein atoms	5103	5134	5304
Solvent atoms	932	771	886
Ligand atoms	82	112	74
Wilson B ($Å^2$)	17.9	19.1	15.1
Average B protein ($Å^2$)	16.0	19.0	14.6
Average B solvent ($Å^2$)	30.6	28.8	29.3
Average B ligand ($Å^2$)	39.1	25.2	23.2
RMSD Bonds (Å)	0.0134	0.0144	0.0110
RMSD Angle (°)	1.66	1.56	1.65
Ramachandran distribution			
Most favored (%)	93.1	93.1	92.7
Additionally allowed (%)	6.6	6.3	6.9
Generously allowed (%)	0.4	0.6	0.4

[1] values in parenthesis are for the highest resolution shell

washed with labeling buffer until the flow through was colorless. Following washing, the labeled protein was eluted with elution buffer (20 mM TrisCl (pH 8.0), 100 mM NaCl, 2.5 M imidazole (pH 8.0) and subsequently buffer exchanged into storage buffer (20 mM sodium acetate (pH 5.0), 300 mM NaCl, 20 % glycerol) and stored at 193 K.

5.2.6 Binding studies

Dissociation constants (K_d) were determined by titrating a fixed amount of 5-IAF labeled PhyAsr with $InsP_6$ or MIHS. All fluorescence measurements were made using a Cary Eclipse fluorimeter (Varian). The 5-IAF fluorophore was excited at 480 nm, and the resulting emission between 490 nm and 550 nm was measured. Measurements were taken in 0.5 nm increments at a rate of 300 nm/min in quadruplicate at 293 K after a 2 min equilibration period following injection of titrant. Blank titrations were performed by titrating 5-IAF labelled PhyAsr with buffer. Fluorescent data points were obtained by averaging data from fluorescent emissions across 20 nm, centered on the fluorescence emission maxima (517 nm). The change in fluorescence was obtained by subtracting the blank titrations from the titration curves for $InsP_6$ or MIHS. The resulting difference curves were evaluated using the following equation:

$$\Delta Fl = 0.5 \times (Bmax / [Prot]t) \times [Prot]t + [Lig]t - \sqrt{(Kd + [Prot]t + [Lig]t) - 4 \times [Prot]t \times [Lig]}$$

Where B_{max} is amplitude, Prot is the total concentration of PhyAsr, Lig is the total concentration of $InsP_6$/MIHS, and K_d is the dissociation constant (Wilden et al. 2006).

5.3 Results

5.3.1 Conformation of $InsP_6$ bound to PhyAsr

To examine the interactions between PhyAsr and its substrate, $InsP_6$, we have determined the high resolution structure (1.6 Å) of the inactive C252S mutant of PhyAsr in complex with $InsP_6$. Analyses of $2F_o$-F_c omit and difference density maps clearly identified density for $InsP_6$ bound in the active site. The $InsP_6$ molecule displays excellent electron density. However, refinement of the ligand with full occupancy resulted in negative difference density for all of the atoms in the ligand with the exception of the 3-phosphate. To account for this density, we modeled PhyAsr as binding either $InsP_6$ or inorganic phosphate (an impurity in the $InsP_6$) with occupancies of 75 % and 25 %, respectively.

The structure of PhyAsr in complex with $InsP_6$ is significantly different than the structure of PhyAsr in complex with the competitive inhibitor MIHS (Chu et al. 2004). The first major difference is that the substrate $InsP_6$ does not bind to PhyAsr in a 5-axial/1-equatorial conformation as was observed with the inhibitor (Figure 5.1). One consequence of this conformational difference is that binding of $InsP_6$ in a 5-equatorial/ 1-axial conformation better positions the leaving group oxygen on the scissile phosphate for protonation by the general acid (D223). Binding an equatorial phosphate in the active site positions the bridging oxygen 0.60 Å closer to the general acid than in the inhibitor complex (Chu et al. 2004). This was previously predicted from substrate docking studies (Puhl et al. 2007).

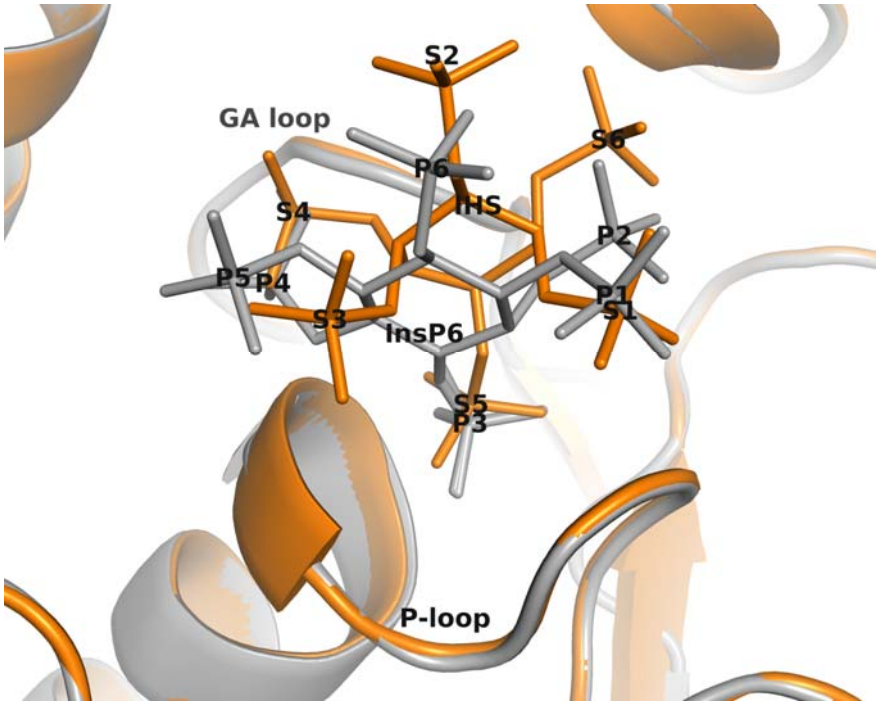

Figure 5.1: Comparison of the structures of PhyAsr bound to $InsP_6$ (grey) and MIHS (orange; 1U26). $InsP_6$ binds in a 5-equatorial/1-axial conformation with the 3-phosphate bound in the active site. MIHS is bound in a 5-axial/1-equatorial conformation with the 5-sulfate bound in the active site.

The second major difference that we observe is that $InsP_6$ binds to PhyAsr with the 3-phosphate in the active site (Figure 5.2a). This agrees with the PhyAsr degradation pathway which showed that $InsP_6$ dephosphorylation is initiated at and is highly specific for the 3-phosphate (Puhl et al. 2007). Finally, we observe that $InsP_6$ binding induces strain on the ligand by forcing the ring into a partially twisted chair conformation and increasing the C-O-P angle for the active-site phosphate by 27° relative to the other phosphates (Figure 5.2b). The twisted conformation of the ligand maximizes the agreement with the electron density for the ligand (Figure 5.2c,d). The $InsP_6$ molecule bound to the *E. coli* phytase (1DKQ) adopts a chair conformation. Manual superposition of this ligand onto the $InsP_6$ bound to PhyAsr reveals that steric clashes between the GA loop and the 2-phosphate appear to be involved in forcing the ring into the twisted

conformation. Additionally, the increase in the C-O-P angle of the phosphate in the active site relieves steric clashes between the *myo*-inositol ring and P-loop, specifically A254. This conformation is stabilized by the extensive contacts made to both the 2-phosphate and the scissile phosphate.

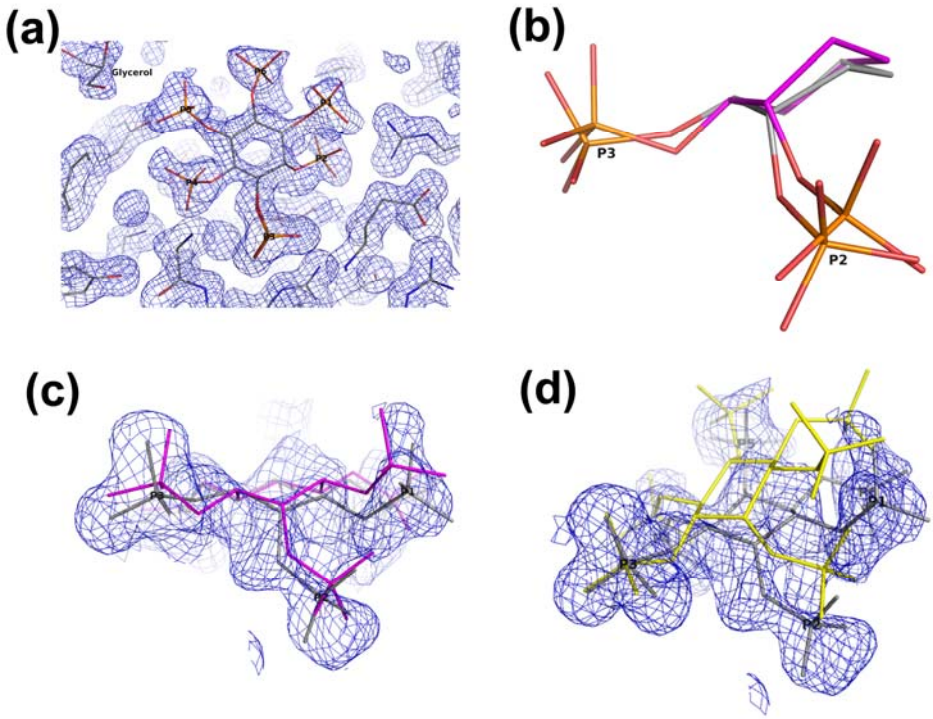

Figure 5.2: $InsP_6$ ring conformation and electron density. (a) σ_A $2F_o$-F_c density for $InsP_6$ bound in the active site with the 3-phosphate. A glycerol is bound in close proximity to the 5-phosphate. (b) Superposition of the $InsP_6$ from PhyAsr complex (grey) onto the ideal chair conformation of the phytate bound to the *E. coli* phytase (magenta) to show the partially twisted ring conformation and increased C-O-P angle for the active site phosphate. For clarity only the 2 and 3-phosphates are shown. (c) Superposition of the $InsP_6$ ring comparing the fit of a strained $InsP_6$ (grey) and an ideal chair $InsP_6$ (magenta) into σ_A $2F_o$-F_c electron density. (d) Superposition of the active site phosphate comparing the fit of a strained $InsP_6$ (grey) and an ideal chair $InsP_6$ (yellow) into σ_A $2F_o$-F_c electron density. All electron density is contoured at 1σ.

5.3.2 PhyAsr makes extensive contacts to $InsP_6$

The binding pocket of PhyAsr is an electropositive cleft that is large enough to accommodate the highly negative $InsP_6$ molecule (Figure 5.3a). Extensive contacts with the protein are made to the 2-, 3- and 4-phosphates at the base of the ligand (i.e. the portion of ligand closest to the nucleophile) in the P2, P3 and P4 sites, respectively (Figure 5.3b). In contrast, few contacts between the protein and substrate are made to the solvent-exposed 1-, 5- and 6-phosphates. These phosphates point out of the active site and occupy the P1, P5 and P6 sites, respectively. Residues that form hydrogen bonds and ion pairs with the substrate are R57, R68, D153, D223, H224, the P-loop main chain (residues 253-258), R258, K301, Y309, and K312 (Figure 5.3c). In addition, several ordered solvent molecules are observed coordinating the phosphates in the P2, P4 and P5 sites. K83, H224, K305 and K312 are involved in coordinating these ordered waters. A detailed listing of the specific contacts made to each phosphate is shown in Table 5.2. Many of the residues that are involved in the binding of $InsP_6$ are also involved in binding MIHS (Chu et al. 2004). R68 was not involved in inhibitor binding but undergoes a large conformational change upon substrate binding. In the absence of substrate, R68 adopts a conformation pointing into the solvent and away from the active site. This induced-fit movement results in R68 adopting a conformation pointing towards the substrate and coordinating the phosphate in the P1 site. This movement is a result of a large rotation in $\chi 1$ and $\chi 2$ that leads to a movement in the $N\eta 1$ and $N\eta 2$ atoms of the guanidium group by to 8.3 and 9.8 Å, respectively. This movement results in the formation of a 3.1 Å hydrogen bond to the 1-phosphate and may be important for the initial binding of substrate.

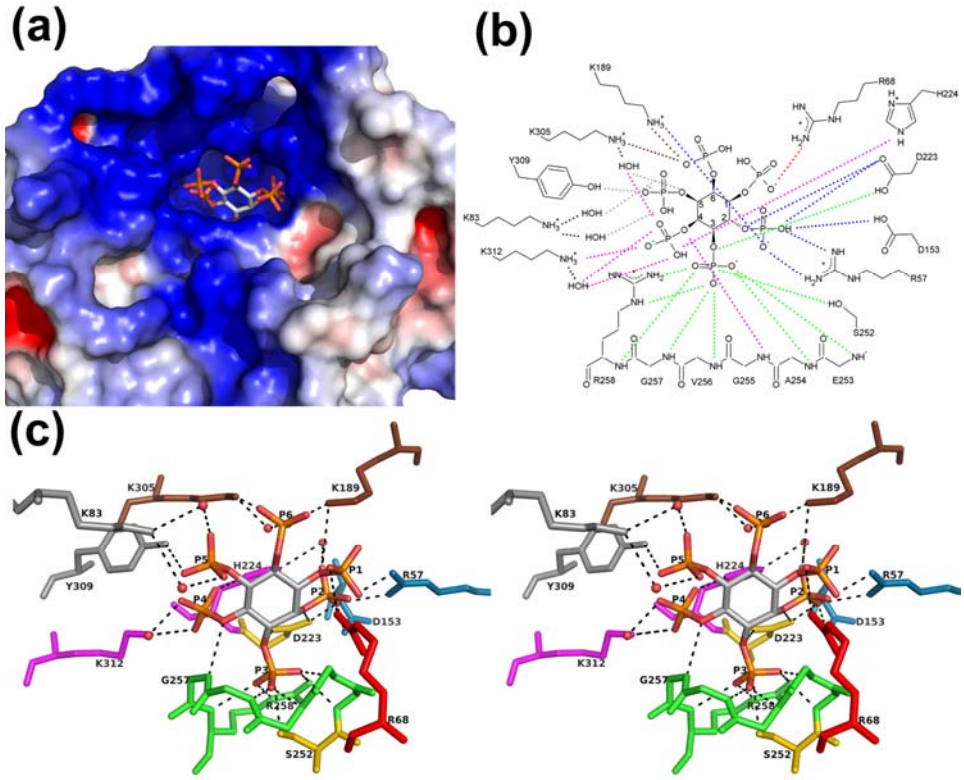

Figure 5.3: Binding of $InsP_6$ to PhyAsr **(a)** $InsP_6$ binds deep in the highly basic PhyAsr binding pocket. The electrostatic surface potential was calculated using the APBS plugin in PyMOL (DeLano 2008). Positively and negatively charged regions are colored blue and red, respectively. **(b)** Two-dimensional representation of the contacts made by PhyAsr to each of the phosphates. Contacts are shown as dashed lines and are colored according to the phosphate binding site. P1 site – Red, P2 site – Blue, P3 site – Green, P4 site – purple, P5 site – light grey, P6 site – Brown. **(c)** Divergent stereo-view of the interactions between $InsP_6$ and PhyAsr. Residues involved in $InsP_6$ binding are shown as sticks and color coded by site as in panel b. Contacts are shown as dashed lines.

Table 5.2: Contacts between PhyAsr and Ins(1,2,3,4,5,6)P_6

Residue	Phosphate	Contact (Å)
R57	P2	2.73, 2.88, 2.48
R68	P1	3.12
D153	P2	3.09
K189	P6/	3.47/
	P2	3.37
D223	P2	3.10, 2.57, 2.74
H224	P4	2.97
S252	P3	2.50, 3.50
E253	P3	3.00
A254	P3	3.14
G255	P3	2.83
V256	P3	2.67
G257	P4	3.24
R258	P3	2.89, 2.87, 2.79
K305	P6	3.60
Y309	P5	3.26
K312	P4	2.83

A glycerol molecule is observed in the binding pocket of PhyAsr packing against a hydrophobic patch composed of F84, F289, F294 and Y309 (Figure 5.4). The oxygens are coordinated by Nζ of K83, the main-chain amine of K297, and the 5-phosphate of $InsP_6$ (3.0 Å hydrogen bond). The close proximity of a glycerol to the inositol phosphate, and the presence of distinct glycerol binding sites may confer on PhyAsr the ability to bind phosphatidylinositol phosphates (PtInsPs).

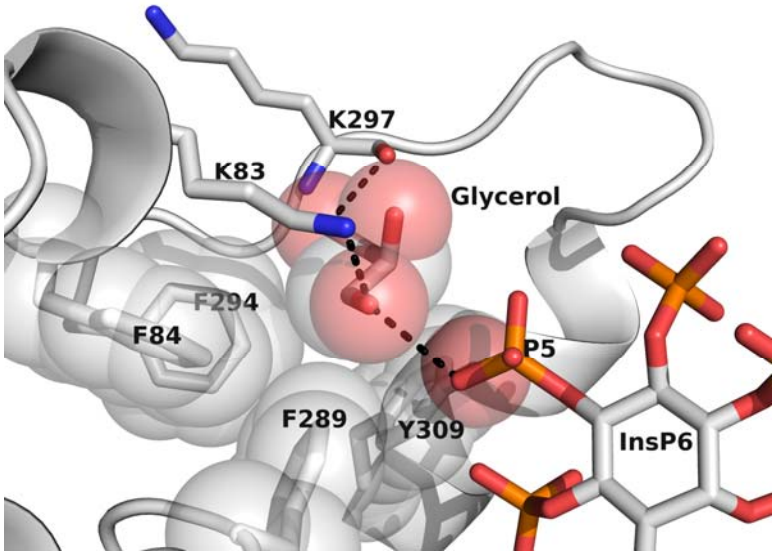

Figure 5.4: Hydrophobic glycerol binding site in the binding pocket of PhyAsr. The glycerol is positioned 3 Å from the 5-phosphate of $InsP_6$. Hydrophobic residues that make van der Waals interactions with the glycerol backbone are shown as sticks surrounded by transparent spheres. Interactions to glycerol oxygens are shown as dashed lines.

5.3.3 Interactions to the general acid loop and P-loop

The scissile 3-phosphate is bound in the active site and makes 5 hydrogen bonds to the P-loop main-chain amines and two hydrogen bonds to the guanidium group of the conserved active-site arginine (R258) (Figure 5.3b, c). This positions the active-site phosphate for nucleophilic attack by C252. An additional hydrogen bond is formed between the main-chain amine of G257 and the bridging oxygen of the 4-phosphate (Figure 5.3b, c). This latter hydrogen bond is not observed in the inhibitor complex and appears to help stabilize the base of the ligand.

The HD motif in the general acid loop of PTPLPs was shown to be involved in binding MIHS (Chu et al. 2004). H224 makes a bridging interaction between two of the axial sulfate groups on the inhibitor (Chu et al. 2004). In the $InsP_6$ complex, H224 makes

a single hydrogen bond to the 4-phosphate and a solvent bridged interaction to the axial 2-phosphate.

5.3.4 $InsP_6$ and MIHS binding studies

To date, the reported binding affinities of phytases for $InsP_6$ or any other inositol phosphates have been based on Michaelis-Menten kinetic parameter K_m. To examine the binding affinity (K_d) of PhyAsr for $InsP_6$ and MIHS we used a fluorescent based assay to measure substrate binding. The K_d for MIHS was determined to be between 2 μM and 5 μM by steady state titrations, and stopped-flow binding studies (Figure 5.5a). The binding curve obtained for $InsP_6$ is characteristic of a very tight binding interaction and indicates that the concentration of protein (5 μM) in the assay was much greater than the K_d for $InsP_6$ (Figure 5.5b) (Goodrich and Kugel 2007). The sensitivity of the current fluorescence based assay precludes us from significantly decreasing the protein concentration to enable us to obtain a K_d value, however the curve is consistent with PhyAsr having a sub-μM binding affinity for $InsP_6$. The K_m for $InsP_6$ is dependent on ionic strength and has been found to vary from 1290 ± 240 μM at 100 mM ionic strength to 150 ± 10 μM at 500 mM ionic strength (Gruninger et al. 2008). At all ionic strengths, the K_d for $InsP_6$ and MIHS is two orders of magnitude less than the K_m for $InsP_6$. These finding suggest that PhyAsr does not utilize a classical Michaelis-Menten reaction mechanism.

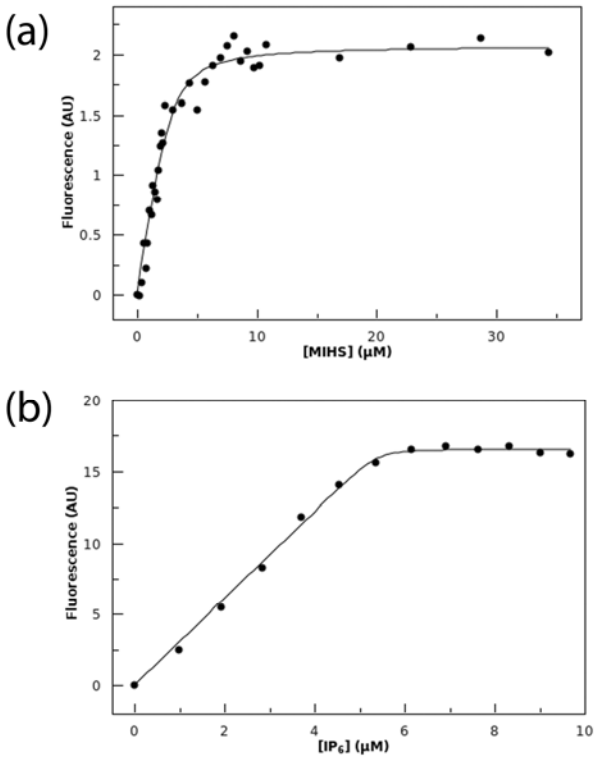

Figure 5.5 Equilibrium binding studies of PhyAsr. (a) Titration of PhyAsr H188C with MIHS. (b) Titration of PhyAsr H188C/C252S with $InsP_6$. Both binding curves were fit using the equation shown in experimental procedures.

5.3.5 Structure of PhyAsr bound to Ins(1,3,4,5)P_4 in two conformations

PhyAsr is known to associate with the outer membrane of *S. ruminantium* (D'Silva et al. 2000) and the identification of a glycerol binding site in close proximity to a ligand phosphate led us to speculate that PhyAsr can bind phosphatylinositol phosphates. The related inositol phosphatase PTEN targets the membrane phospholipid PtIns(3,4,5)P_3 so it was hypothesized that PhyAsr may also bind this substrate. To obtain a complex of PtIns(3,4,5)P_3 the soluble analogue of this substrate, glycerol-Ins(3,4,5)P_3 (G3P), was added to a crystal of the PhyAsr C252S mutant. The structure of this complex was determined to a resolution of 1.85 Å. Clear continuous electron density is observed

for the inositol ring and the phosphates (Figure 5.6d). The glycerol was not visible so it was omitted from the model. The Ins(1,3,4,5)P_4 head group binds in a conformation identical to that observed in the $InsP_6$ complex (Figure 5.6a). The phosphates occupy the binding sites observed in the $InsP_6$ complex with the 3-phosphate positioned in the active site and the remaining phosphates occupying the P1, P4 and P5 sites (Figure 5.6b, c). The protein makes 19 hydrogen bonds to the ligand involving residues R57, R68, K198, D223, H224, the P-loop, and K312 (Table 5.3).

Interestingly, we observe two distinct Ins(1,3,4,5)P_4 conformations bound in the active site. The first conformation is the same as $InsP_6$ (described above), while the second is bound with the 1-phosphate in the active-site and a ring conformation that is distinct from that observed in $InsP_6$ (Figure 5.7a). The ring of the latter conformation (referred to as the $InsP_4$ conformation), is flipped relative to $InsP_6$ and tilted by approximately 75° so that it packs against the GA loop (Figure 5.7a). $2F_o$-F_c and F_o-F_c electron density calculated after omitting the ligands is fully accounted for by modeling Ins(1,3,4,5)P_4 in both the $InsP_4$ and $InsP_6$ conformations. However, the density for the $InsP_4$ conformation is much stronger than the $InsP_6$ conformation (Figure 5.7d). These two ligand conformations are bound with occupancies of 25% and 75%, respectively.

The binding of the 1-phosphate in the active site is only possible in the absence of glycerol at this position. According to the manufacturer (Echelon Biosciences), hydrolysis of the glycerol moiety can occur at low pH values < 4.0. The crystallization pH is 4.8 so we did not expect this to be a problem. However, the local pH of the active site may be lower and/or the electropositive binding site may have promoted hydrolysis

of the glycerol. Alternatively, a small amount of substrate without the glycerol moiety may have been present which preferentially bound to the protein. It is unclear whether

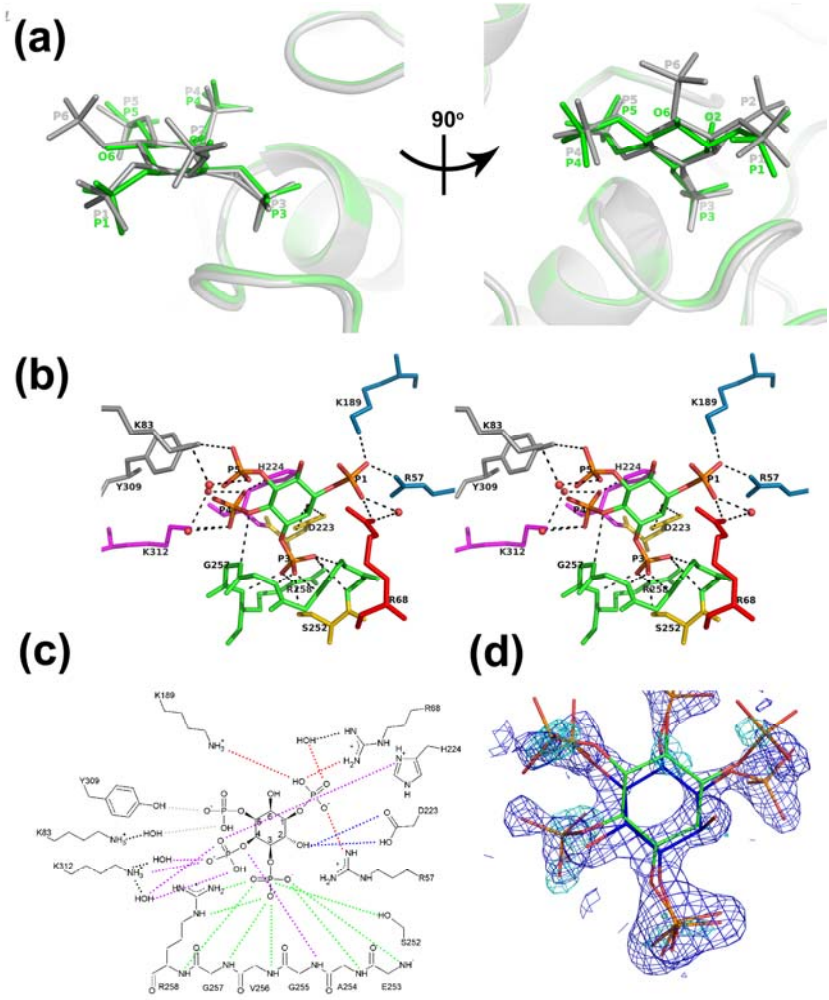

Figure 5.6: Binding of Ins(1,3,4,5)P_4 to PhyAsr in the InsP_6 conformation. **(a)** Least squares superposition of the PhyAsr InsP_6 complex (Grey) and PhyAsr bound to Ins(1,3,4,5)P_4 in the InsP_6 conformation (Green). **(b)** Divergent stereo-view of the interactions between Ins(1,3,4,5)P_4 and PhyAsr. Residues involved in binding are shown as sticks and color coded by phosphate binding site as in Figure 5.3. Contacts are shown as dashed lines. **(c)** Two dimensional representation of the contacts made by PhyAsr to each phosphate. Contacts are shown as dashed lines and are colored by phosphate binding site as shown in panel b. **(d)** σ_A weighted omit density calculated after omitting the InsP_6 conformation (green). Positive F_o-F_c density is contoured at 4σ and shown in cyan. $2F_o$-F_c density is contoured at 1 σ and shown in blue.

the ligand with the 3-phosphate in the active-site is G3P, as the inability to resolve the glycerol may be due to disorder.

The protein makes extensive contacts to all four phosphates of Ins(1,3,4,5)P_4 in the $InsP_4$ conformation (Figure 5.7b, c). With the exception of R68, the same residues involved in binding the $InsP_6$ conformation are involved in binding the $InsP_4$ conformation (Table 5.3). In this conformation the 3-phosphate occupies the P2 binding site and the 4- and 5- phosphates occupy binding sites not observed in the $InsP_6$ conformation. The 4-phosphate is coordinated by K189 and K305 while the 5-phosphate makes specific contacts to H224, K305 and Y309. The binding sites occupied by the 1- and 4- phosphates in the $InsP_6$ conformation are occupied by solvent molecules. The main-chain amine of G257 interacts with the hydroxyl at the 6-position in a fashion that

Table 5.3: Contacts between PhyAsr and Ins(1,3,4,5)P_4 bound in two conformations. The residue, ligand phosphate/hydroxyl and contact distance (Å) are shown.

PhyAsr C252S $InsP_4$ - $InsP_6$ conformation			**PhyAsr C252S $InsP_4$ - $InsP_4$ conformation**		
R57	P1	2.87	R57	P3	2.94, 2.83, 2.63
R68	P1	2.99	R68	-	-
K189	P1	2.95	K189	P3/ P4	2.61/ 2.71
D223	O2	3.13, 2.87	D223	P3	2.89, 2.66
H224	P4	2.91	H224	P3/ P5	2.84/ 2.92
S252	P3	2.80, 3.30	S252	P1	2.44
E253	P3	3.19	E253	P1	3.89
A254	P3	3.30	A254	P1	3.17
G255	P3	2.99	G255	P1	2.84
V256	P3	2.86	V256	P1	2.64
G257	P3/ P4	3.45/ 3.38	G257	P1/ O6	3.49/ 2.92
R258	P3	2.93, 2.68, 2.83	R258	P1	2.81, 3.04, 2.96
K305	-	-	K305	P4/ P5	2.54/ 2.67, 2.93
Y309	P5	3.59	Y309	P5	2.05, 3.22
K312	P4	3.00	K312	-	-

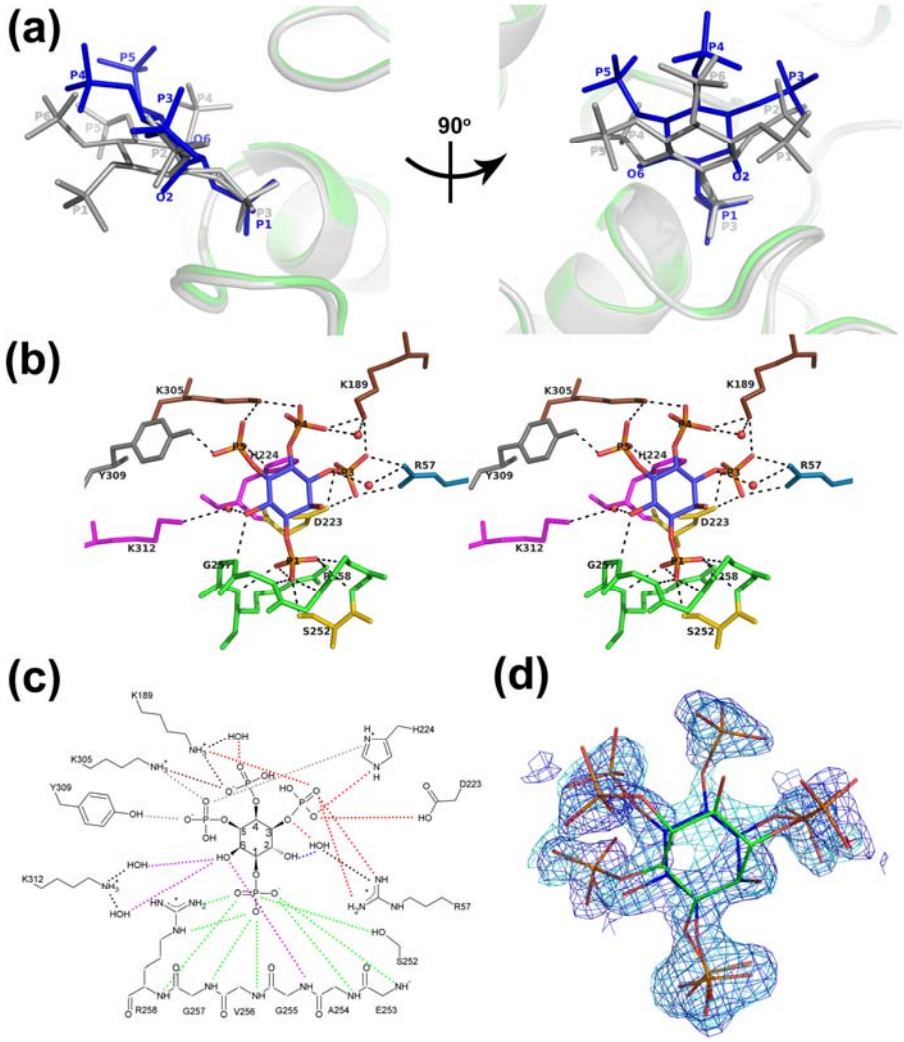

Figure 5.7: Binding of Ins(1,3,4,5)P_4 to PhyAsr in the InsP_4 conformation. **(a)** Le/ast squares superposition of the PhyAsr InsP_6 complex (Grey) and PhyAsr bound to Ins(1,3,4,5)P_4 in the InsP_4 conformation (Blue). The InsP_4 ring is flipped relative to InsP_6 and shifted towards the GA loop. **(b)** Divergent stereo-view of the interactions between Ins(1,3,4,5)P_4 and PhyAsr. Residues involved in binding are shown as sticks and color coded by phosphate binding site as in Figure 5.3. Contacts are shown as dashed lines. **(c)** Two dimensional representation of the contacts to each of the phosphates. Contacts are shown as dashed lines and are colored by phosphate binding site as in panel b. **(d)** σ_A weighted omit density calculated after omitting the InsP_6 conformation (blue). Positive F_o-F_c density is contoured at 4σ and shown in cyan. $2F_o$-F_c density is contoured at 1 σ and shown in blue.

is analogous to that observed in the InsP_6 complex. The C-O-P angle of the scissile phosphate in the InsP_4 conformation kinks away from D223, positioning the leaving

group oxygen 0.7 Å farther from the general acid. This conformation is similar to that observed for the active-site sulfate in the PhyAsr MIHS complex (Chu et al. 2004).

5.3.6 Structure of PhyAsr bound to Ins(1,2,3,5,6)P_5

In an attempt to obtain a structure of the PhyAsr C252A mutant in complex with $InsP_6$ we unexpectedly observed the binding of the inositol pentakisphosphate, Ins(1,2,3,5,6)P_5, and inorganic phosphate to the protein (Figure 5.8a). This inositol polyphosphate is a contaminant of $InsP_6$ purified from plant sources and is generated from intrinsic phytase activity in plants. This compound has previously been observed in the structure of a pleckstrin homology domain soaked with $InsP_6$ (Jackson et al. 2007). The smaller alanine side-chain enables the phosphate in the active-site to bind 0.8 Å deeper than the phosphate in the $InsP_6$ complex. The binding of this phosphate sterically prevents $InsP_6$ binding but still enabled binding of Ins(1,2,3,5,6)P_5. The phosphate in the active site makes a 2.6 Å hydrogen bond with the 4-hydroxyl of Ins(1,2,3,5,6)P_5. The $InsP_5$ and $InsP_6$ rings are positioned similarly in both complexes, however the ring of $InsP_5$ is shifted away from the base of the active-site by 1 Å and flipped relative to $InsP_6$ (Figure 5.8a).

The protein makes 14 contacts to $InsP_5$ with R57, K83, K189, D223, H224, K305, Y309, and K312 with phosphates bound in the P2, P4, P5 and P6 sites (Figure 5.8b, c). An additional 10 contacts are made with the phosphate ion in the active site by the main-chain amines of the P-loop and the side chains of R258 and T259 (Table 5.4). The 1-phosphate is solvent-exposed and occupies a binding site not observed in the $InsP_6$, $InsP_4$,

or MIHS complexes. Similar to both the $InsP_4$ and $InsP_6$ complexes, G257 makes a contact to the base of the ligand.

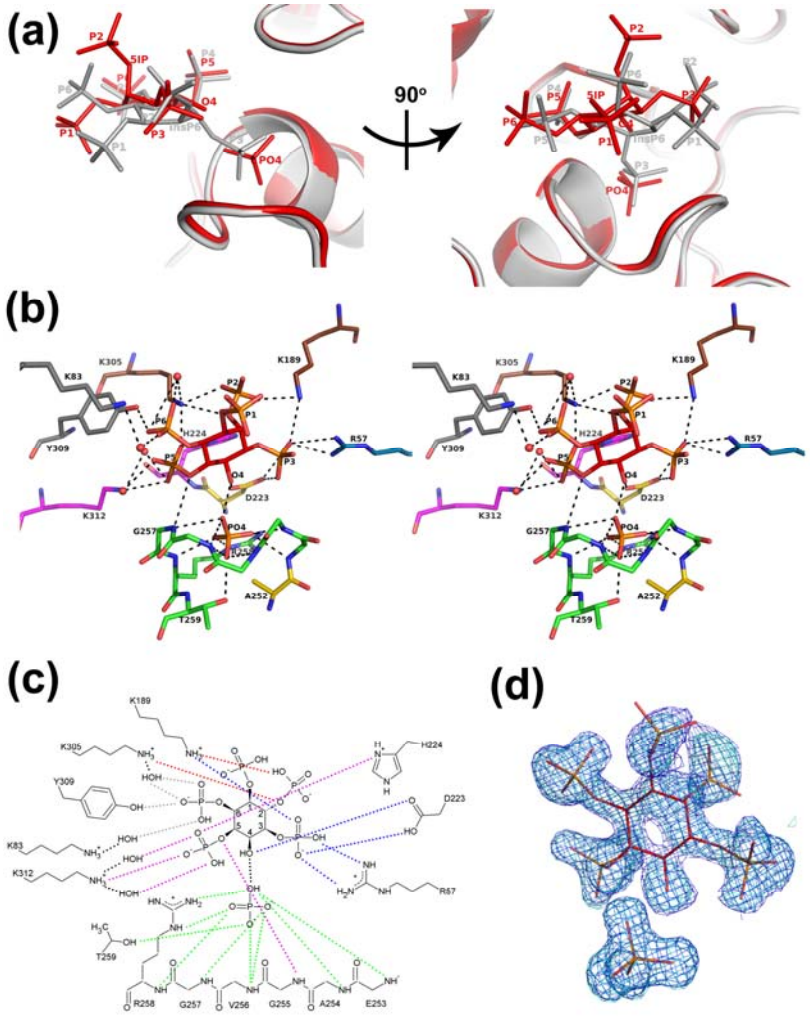

Figure 5.8: Binding of Ins(1,2,3,4,5)P_5 and inorganic phosphate to PhyAsr. **(a)** Least squares superposition of the PhyAsr $InsP_6$ complex (Grey) and PhyAsr bound to Ins(1,2,3,5,6)P_5 and inorganic phosphate (Red). **(b)** Divergent stereo-view of the interactions between Ins(1,2,3,5,6)P_5, inorganic phosphate and PhyAsr. Residues involved in binding are shown as sticks and color coded by phosphate binding site as in Figure 5.3. Contacts are shown as dashed lines. **(c)** Two-dimensional representation of the contacts made to each phosphate. Contacts are shown as dashed lines and are colored by phosphate binding site as shown in panel b. **(d)** σ_A weighted omit density calculated after omitting the $InsP_5$ and inorganic phosphate. Positive F_o-F_c density is contoured at 4σ and shown in cyan. $2F_o$-F_c density is contoured at 1 σ and shown in blue.

Table 5.4: Contacts between PhyAsr and Ins(1,2,3,5,6)P_5 and phosphate.

PhyAsr C252A $InsP_5$		
R57	P3	2.91, 3.06, 3.35
K83	P6	3.53
K189	P2/ P3	2.66, 3.16
D223	P3/ P4	2.63, 2.87
H224	P5	2.73
E253	Pi	2.82
A254	Pi	2.96
G255	Pi	3.02
V256	Pi	3.17, 2.97
G257	Pi/ P5	3.11/ 3.42
R258	Pi	2.78, 2.83, 2.94
T259	Pi	2.78
K305	P2/ P5	2.86/ 2.71
Y309	P6	2.60
K312	P5	3.07

5.3.7 Similarity between the PhyAsr C252A complex and PTP1B covalent intermediate

Based on the shift in the phosphate position, and the smaller size of the alanine side chain, we thought that the phosphate in the active site of the C252A mutant could be positioned similarly to the cysteinyl-phosphate intermediate in PTP1B (Pannifer et al. 1998). Superposition of the PhyAsr C252A complex with the PTP1B phospho-enzyme intermediate (PDB 1A5Y) confirmed this hypothesis (Figure 5.9). There is a small difference in the position of the phosphates that is likely due to the non-covalent nature of the phosphate in our structure. The similarity between the PhyAsr C252A structure and

the PTP1B cysteinyl-phosphate indicates that this model provides insight into the nature of the PhyAsr covalent intermediate.

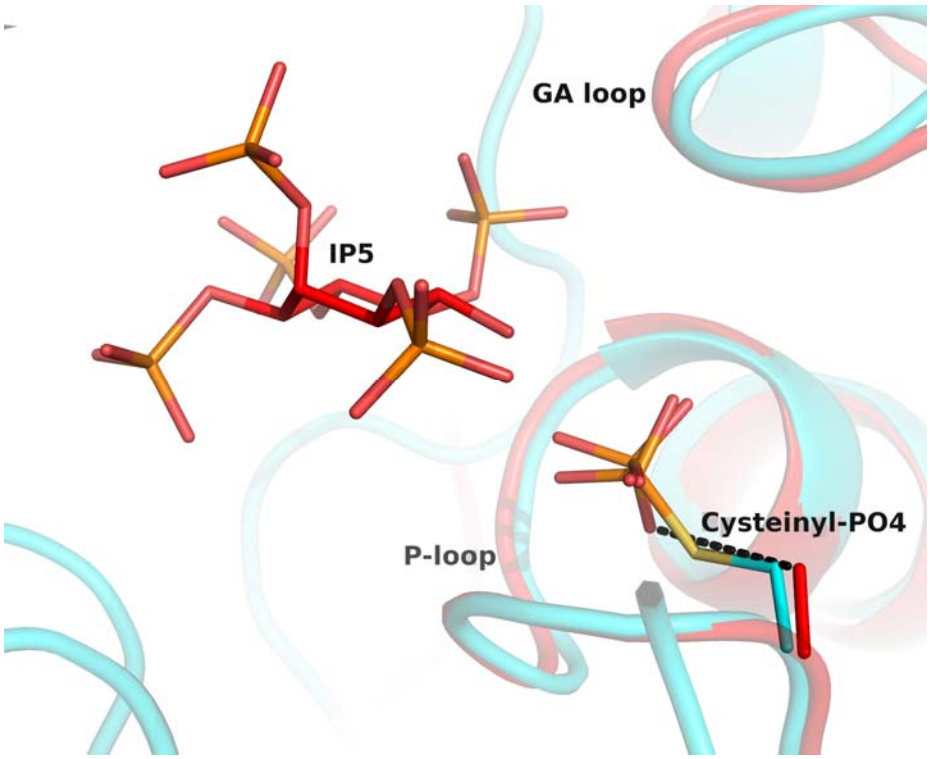

Figure 5.9: The PhyAsr C252A $InsP_5$ complex resembles the PTP1B phospho-enzyme intermediate. Least squares superposition of PhyAsr C252A mutant (Red) bound to Ins(1,2,3,5,6)P_5 (IP5) and phosphate and the phospho-cysteine intermediate of PTP1B (Cyan, PDB 1A5Y). The general acid and phosphate binding loops are labeled GA loop and P-loop, respectively.

5.4 Discussion

5.4.1 $InsP_6$ binding by PhyAsr

We have determined the first structure of a protein tyrosine phosphatase-like phytase in complex with $InsP_6$ and have identified novel features not observed in the inhibitor complex previously described (Chu et al. 2004). Our results show that PhyAsr binds $InsP_6$ in a 5-equatorial/1-axial conformation with the 3-phosphate in the active site.

This is consistent both with the known $InsP_6$ degradation pathway and the preferred ring conformation under physiological conditions (Puhl et al. 2007). Furthermore, we observe that the binding of $InsP_6$ induces strain on the ligand. There are no reported examples of substrate distortion in other inositol phosphatases. Substrate distortion has been observed in several glycosidase and glycosyltransferases (Zou et al. 1999; Davies et al. 2003; Shah et al. 2003). In these enzymes, the ligand is distorted into a high-energy conformation that resembles the transition state and helps position the substrate optimally for catalysis (Davies et al. 2003). The glycosidases and PhyAsr both distort the substrate to relieve steric clashes with the protein. The distortion in the C-O-P angle of the scissile phosphate forces the ligand into a conformation closer to the transition state. The ring strain may function to minimize unfavorable steric and electrostatic clashes with the protein, while maximizing the stabilizing contacts with the ligand.

The only other phytate degrading enzyme that has been structurally characterized in complex with $InsP_6$ is the *E. coli* phytase (Lim et al. 2000). This enzyme hydrolyses the 6-phosphate preferentially but was observed to bind the 3-phosphate. There was no structural explanation as to why the 3-phosphate bound in the active site (Lim et al. 2000). The results of Lim and colleagues (2000) indicate that the *E. coli* phytase can bind $InsP_6$ in multiple conformations. As in the case of PhyAsr, the binding pocket of this enzyme is a highly basic cleft and the protein makes extensive contacts to the phosphates near the catalytic site at the base of pocket. Two of the phosphates point out of the active site and make minimal contacts to the protein. In contrast to PhyAsr and the *E. coli* phytase, β-propeller phytases utilize both divalent cations and basic residues to bind the substrate (Shin et al. 2001).

$InsP_6$ has also been observed to bind to a wide range of other proteins important in an array of biological processes including signaling, vesicular trafficking, RNA processing, glycolysis, and bacterial virulence (Rigden et al. 1999; Ford et al. 2001; Collins et al. 2002; Macbeth et al. 2005; Milano et al. 2006; Tan et al. 2007; Lupardus et al. 2008). The biological role of $InsP_6$ varies widely in these proteins. It has been found to serve as a co-factor important for substrate binding, a structural co-factor, an allosteric regulator, and an inhibitor. An examination of the interactions made to $InsP_6$ in these proteins reveals that there appear to be two mechanisms of $InsP_6$ binding. In proteins that utilize $InsP_6$ as a co-factor, $InsP_6$ binding involves extensive contacts between basic residues and all of the phosphate groups on the ligand. In contrast, proteins that utilize $InsP_6$ as a substrate (i.e. PhyAsr and the *E. coli* phytase), and those involved in signaling pathways make extensive contacts to only some of the phosphates, presumably to facilitate dissociation of the product/ligand. In all cases, $InsP_6$ binds in large, highly basic clefts on the surface of the protein.

5.4.2 Functional significance of binding sites

We have determined the structure of PhyAsr in complex with multiple inositol phosphates that differ in both the number, and position of the phosphate groups. This is the first such study carried out with a phytate degrading enzyme and provides insight into the substrate specificity of phytases. We have identified several conserved phosphate binding sites (the P2, P3 and P4 sites) near the base of the active site, and several variable, solvent accessible binding sites (P1, P5 and P6 sites). The P3 site is the catalytic center and as such is directly involved in catalysis. The P2 and P4 sites are

located adjacent to the catalytic site and appear to function as anchor points that ensure the scissile phosphate is correctly oriented for nucleophilic attack. The function of the variable sites is less clear but they may be important for binding and discriminating between the lower phosphorylated inositol phosphates that are generated during the degradation of $InsP_6$. An examination of the structure of the *E. coli* phytase bound to $InsP_6$ reveals that it also has binding sites that anchor the substrate near the active site and binding sites that are solvent exposed and make few contacts to the substrate (Lim et al. 2000). The complete lack of homology or structural similarity between PhyAsr and the *E. coli* phytase suggests that PTPLPs and HAPs have convergently evolved a similar approach to substrate binding.

5.4.3 Conservation of substrate binding residues

With the exception of the remote standby site identified by Chu et al (2004) the same residues are involved in binding the $InsP_4$, $InsP_5$, $InsP_6$ and MIHS inhibitor conformations. In particular, R57, K189, H224, Y309, and the P-loop are important for binding all of these conformations, while R68, K305, and K312 make contacts in most of the complexes. The catalytic properties of several PTPLPs have recently been described (Puhl et al. 2007; Puhl et al. 2008a; Puhl et al. 2008b; Puhl et al. 2009) and these enzymes can be separated into high activity and low activity groups. Interestingly, all of the binding residues identified in PhyAsr are conserved in the high activity PTPLPs while the low activity PTPLPs have conservative and non-conservative mutations, as well as deletions, at several of these sites (Figure 5.10). Conservation of these binding residues also appears to play a role in defining substrate specificity as two of the enzymes.

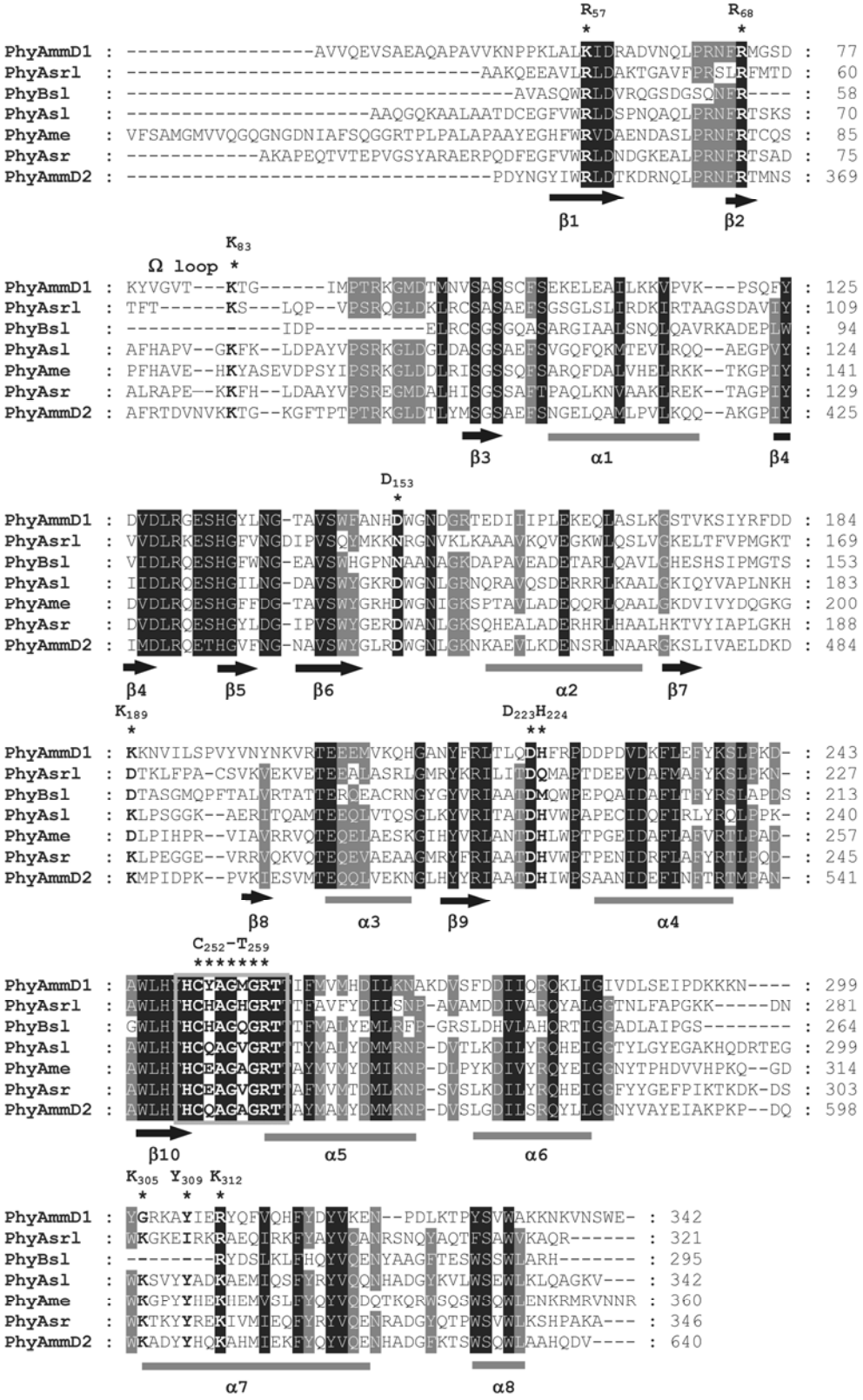

Figure 5.10: Alignment of bacterial protein tyrosine phosphatase-like phytases. Amino acids are shaded according to the level of conservation (black, 100%, grey, 80%). The

Figure 5.10 continued: protein abbreviations, source, GeneBank accession number, and residues included are as follows: PhyAmmD1, *M. multacida* D1 repeat, ABA18187, 35-342; PhyAsrl, *S. ruminantium* subsp. *lactylitica*, ABC69359, 32-321; PhyBsl, *S. lacticifex*, ABC69361, 38-295; PhyAsl, *S. lacticifex*, ABC69367, 33-342; *Measphaera elsdenii*, ABC69358, 26-360; PhyAsr, *S. ruminantium*, AAQ13669, 28-346; PhyAmmD2, *M. multacida* D1 repeat, ABA18187, 343-640. Numbers at the end of each sequence represent the residue number for the last amino acid in that sequence. Substrate binding residues in PhyAsr are shown above the alignment (*), the general acid loop (GA loop) and phosphate binding loop (grey box) are highlighted. Secondary structure of PhyAsr is shown on the bottom of the alignment.

Mutations to residues involved in substrate binding have significantly altered substrate specificity in several PTPLPs (Puhl et al. 2008a; Puhl et al. 2008b; Gruninger et al. 2009). One of the strictly conserved residues is R68. This residue undergoes a large induced fit movement upon substrate binding and interacts with the 1-phosphate of $InsP_6$. Although this residue does not appear to be involved in binding the $InsP_4$ conformation or Ins(1,2,3,5,6)P_5, it is also observed in a closed conformation. It is possible that this residue is involved in the initial substrate binding stages and guides the ligand into the active site. The interaction between the main-chain amine of G257 and the bridging oxygen in the P4 site is observed in all of our complexes and appears to help anchor the base of the substrate. This contact is particularly important for binding less phosphorylated inositol phosphates as it stabilizes the base of the ring in ligands that lack a phosphate at this position. An examination of the P-loop sequence in PTPLPs and PTPs reveals that G257 is invariant (Andersen et al. 2001). Interestingly, this residue adopts a favorable conformation in the Ramachandran plot that is allowed for non-glycine residues (φ/ψ are both ~ -60°) so why is this residue universally glycine? Mutation of this residue to an amino acid with a larger side chain would result in the formation of a close contact with the ligand that would likely sterically inhibit or preclude substrate binding.

5.4.4 Mechanistic insight into PhyAsr catalysis

A comparison of the K_d values for $InsP_6$ and MIHS (this study) to the previously reported K_m values reveals a large discrepancy (Puhl et al. 2007; Gruninger et al. 2008). The K_d values for $InsP_6$ and MIHS are at least two orders of magnitude lower than the K_m for $InsP_6$ indicating that PhyAsr utilizes a non-Michaelis-Menten mechanism. The K_d is representative of all binding events whether they are catalytic or not whereas the K_m is only representative of binding events that lead to catalysis. If PhyAsr binds substrate in multiple conformations but only some are catalytic then the K_m would underestimate the binding affinity. It was initially suggested that PhyAsr utilizes a processive mechanism for $InsP_6$ degradation (Chu et al. 2004). In this mechanism, the intermediate inositol phosphates are not released from the enzyme until $InsP_6$ is completely hydrolyzed; however, this was shown to be incompatible with the $InsP_6$ degradation pathway (Puhl et al. 2007). Based on our structural analysis, and binding studies, we suggest an alternative reaction mechanism for PhyAsr in which substrate binding involves two phases: 1) rapid initial binding of the substrate, and 2) slow reorientation of the substrate to position it in a catalytic conformation. The observation of five distinct binding modes in PhyAsr clearly indicates that substrate does bind to PhyAsr in multiple conformations. Furthermore, preliminary pre-steady state binding studies suggest that substrate binding involves a rapid initial phase followed by a slower phase (Figure 5.11). This slower phase may be the slow reorientation of the substrate into a catalytic conformation. However, we cannot rule out the possibility that it is due to another process such as a conformational change that occurs during substrate binding. Detailed binding and kinetic

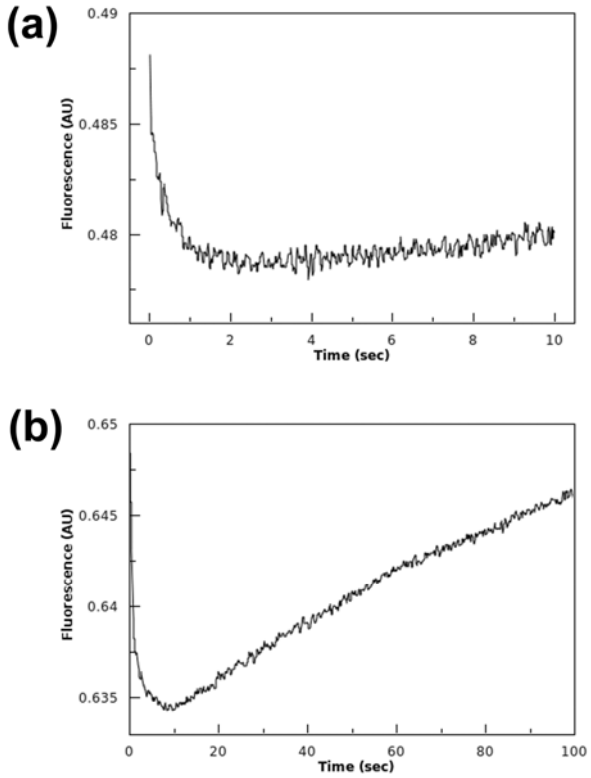

Figure 5.11 Substrate binding to PhyAsr involves more than one phase. The kinetics of inhibitor binding by PhyAsr were measured over (a) a 10 second time scale and (b) a 100 second time scale.

studies are underway to further characterize this multi-phasic binding behavior.

To our knowledge there have been no studies examining the binding of $InsP_6$ to phytases; however, the dissociation constant for the binding of the competitive inhibitor MIHS (K_i) has been determined for both PhyA and PhyB from *Aspergillus ficuum* (Ullah and Sethumadhavan 1998) and for PhyA from *A. fumigates* (Ullah et al. 2000). Interestingly, in both of these enzymes the K_i is much smaller than the K_m for $InsP_6$ by at least an order of magnitude (Ullah and Sethumadhavan 1998; Ullah et al. 2000). These phytases belong to the HAP family of phytase and are structurally similar to the *E. coli* phytase. It is unclear whether the K_d for the substrate, $InsP_6$, would also be much lower than the K_m. However, it should be noted that PhyAsr was observed to have a much

higher affinity for $InsP_6$ than for MIHS. The discrepancy between the binding affinity of MIHS and the K_m of $InsP_6$ in HAPs, the similarities in substrate binding by PhyAsr and the *E. coli* phytase and the finding that the *E. coli* phytase can bind $InsP_6$ in multiple conformations, all suggest that HAPs and PTPLPs may both utilize a similar mechanism involving substrate binding and re-orientation. It will be necessary to examine substrate binding and revisit the enzyme kinetics of HAPs to determine if they do have a reaction mechanism similar to the mechanism proposed for PTPLPs.

Chapter 6 Conclusions and Future Directions

6.1 Overview

This thesis describes the structural and kinetic analysis of the PTPLPs from *S. ruminantium* (PhyAsr) and *M. multacida* (PhyAmm). Using a multi-disciplinary approach, I have shown that these enzymes use a classical PTP catalytic mechanism, adopt a core PTP fold, and have several unique features that contribute to the altered substrate specificity of PTPLPs. I have also determined the first structure of a tandemly repeated inositol polyphosphatase and shown that it has unique substrate specificity. Additionally, I have carried out a detailed structural analysis of inositol polyphosphate binding to PhyAsr. This work provides a basis for understanding the structure-function relationship of this expanding class of enzymes (Table 6.1).

At the outset of this work, a structure of PhyAsr had been reported and several unique features were identified (Chu et al. 2004). These included (1) a substrate induced P-loop movement (2) a 5-axial/1-equatorial substrate binding mode and (3) a processive mechanism of $InsP_6$ hydrolysis. In chapters 2 and 3, I characterized this P-loop movement and conclusively showed that it is a result of oxidation of the catalytic cysteine. Furthermore, examination of the electron density of the previous structure (1U24) revealed that this protein was also oxidized; however, this modification was not reported or modeled.

Table 6.1: Proteins related to PhyAsr PTPLP as identified by BLAST-P.

NCBI Accession	Organism	Identity (%)	E-value
AAQ13669.1	*Selenomonas ruminantium*	100	0.0
ABC69367.2	*Selenomonas lacticifex*	56	1×10^{-99}
ABC69358.4	*Megasphaera elsdenii*	49	2×10^{-86}
ABA18187.1	*Mitsuokella multacida D2*	50	4×10^{-84}
ABA18187.1	*Mitsuokella multacida D1*	38	1×10^{-49}
ZP_04658998	*Selenomonas flueggei*	48	3×10^{-80}
ZP_03929107.1	*Acidaminococcus sp.*	44	2×10^{-67}
ABC69359.4	*Selenomonas ruminantium subsp. Lactilytica*	39	3×10^{-67}
NP_149178.1	*Clostridium acetobutylicum*	37	2×10^{-46}
YP_696211.1	*Clostridium perfringens**	37	1×10^{-44}
YP_001394001.1	*Clostridium kluyveri*	38	6×10^{-44}
ZP_03474300.1	*Mitsuokella multacida DSM 20544*	38	2×10^{-43}
ZP_02995608.1	*Clostridium sporogenes*	34	3×10^{-43}
NP_968118.1	*Bdellovibrio bacteriovorus*	38	7×10^{-41}
YP_001787593.1	*Clostridium botulinum**	32	1×10^{-39}
YP_008827.1	*Candidatus Protochlamydia amoebophila*	35	7×10^{-39}
NP_782216.1	*Clostridium tetani*	33	2×10^{-38}
YP_002953065.1	*Desulfovibrio magneticus*	36	3×10^{-37}
ZP_02951610.1	*Clostridium butyricum*	31	1×10^{-36}
YP_001310746.1	*Clostridium beijerinckii*	30	2×10^{-35}
YP_002850639.1	*Clostridium sp.*	37	7×10^{-35}
ZP_03472163.1	*Mitsuokella multacida DSM 20544*	34	4×10^{-33}
YP_361664.1	*Xanthomonas campestris pv. vesicatoria str. 85-10*	34	2×10^{-32}
YP_125176.1	*Legionella pneumophila**	31	2×10^{-27}
ZP_01462826.1	*Stigmatella aurantiaca*	33	3×10^{-27}
YP_971831.1	*Acidovorax avenae subsp. Citrulli*	34	2×10^{-26}
ZP_04586937.1	*Pseudomonas syringae pv. Oryzae*	29	1×10^{-29}
NP_794465.1	*Pseudomonas syringae pv. tomato str. DC3000*	29	1×10^{-23}

*multiple strains were identified with a related PTPLP

Chapter 5 examines the interactions between PhyAsr and several inositol polyphosphates. This work identified several unique features in the PhyAsr $InsP_6$ complex that are not observed in the inhibitor complex. Specifically, (1) substrate binds in the energetically favoured 5-equatorial/1-axial conformation with the 3-phosphate in the active site, (2) binding induces strain on the ligand, and (3) PhyAsr binds substrate in several distinct conformations. The first observation is consistent with the known PhyAsr catalyzed $InsP_6$ degradation pathway and the conformation of $InsP_6$ at low pH. The second and third observations are novel and have led to the proposal of a non-Michaelis-Menten reaction mechanism that involves an initial binding event, substrate rearrangement, phospho-diester bond hydrolysis and product release. Further studies are needed to dissect the individual steps in the reaction mechanism and to evaluate the validity of this mechanism.

I have also carried out the first structural analysis of a tandemly repeated inositol phosphatase. Mutagenesis and kinetic analysis were used to show that each repeat possesses unique substrate specificity. Interestingly, one of the repeats is specific for highly phosphorylated IPPs (D2) while the other repeat is specific for less phosphorylated IPPs (D1). This has led to the proposal of a 'divide and conquer' approach to inositol polyphosphate metabolism which involves the D2 repeat targeting highly phosphorylated IPPs and the D1 repeat targeting less phosphorylated IPPs. It is noteworthy that the catalytic activity of the D1 repeat is much lower than the D2 repeat which may indicate that D1 is not catalytically active *in vivo*. Alternatively, this repeat may function by targeting a completely unrelated substrate or it may be involved in substrate recruitment and/or protein localization. Interestingly, in addition to the tandem PTPLP, two

additional PTPLPs have been identified in the genome of *M. multacida* (Table 6.1). The other two PTPLPs from *M. multacida* show similar levels of identity to the D2 repeat of PhyAmm as the PhyAmm D1 repeat. This suggests that these enzymes may also have altered substrate specificity. These proteins may not be bona-fide phytases but rather have specificity for lower order inositol phosphates or a completely different substrate. It is unclear why the *M. multacida* genome encodes three related PTPLPs. However, it is possible that they have arisen through multiple gene duplication events. It is also not clear whether these proteins are even expressed or if they are pseudo-genes.

6.2 Biological function of PTPLPs

One of the major unanswered questions about PTPLPs is the biological function of these enzymes. Most of the related proteins have been identified in bacteria belonging to the class Clostridia and many of these bacteria inhabit the gastrointestinal tract of animals (Table 6.1). The prevalence of these proteins in gastrointestinal microbes suggests that they play an important role in the metabolism of $InsP_6$ in animals. These proteins have also been identified in several plant pathogens. Given the role that $InsP_6$ plays in pathogen defense in plants (Murphy et al. 2008) and the identification of virulence factors with phytase activity (Chatterjee et al. 2003) it is possible that these proteins are involved in pathogenesis. The protein expressed by *Pseudomonas syringae pv. tomato str.* DC3000 has been shown to be involved in pathogenesis (Bretz et al. 2003; Espinosa et al. 2003) however the *in vivo* target of this protein has yet to be identified. This protein may in fact act as an inositol phosphatase *in vivo* not as a PTP as has been previously suggested (Bretz et al. 2003; Espinosa et al. 2003). These proteins have also

been identified in the human pathogen *Legionella pneumophila*. This bacterium is interesting because it is an intracellular parasite. SignalP (Bendtsen et al. 2004) identifies the first 20 amino acids of this protein as a signal peptide indicating that it is secreted by the bacterium. Secretion of this PTPLP into the cytoplasm of the host cell would enable it to directly interfere with inositol phosphate signalling pathways. The intracellular pathogen *Salmonella* has been found to secrete an inositol phosphatase into host cells to aid in the cytoskeletal rearrangements that are involved in the internalization of the bacterium (Zhou et al. 2001). It is possible that the *Legionella* protein has a similar function.

Due to the obligate anaerobic nature of *S. ruminantium* and *M. multacida*, it has been difficult to carry out *in vivo* studies to elucidate the biological function of PhyAsr and PhyAmm. Additionally, the genetics of these microbes are not well understood and there are currently no tools for genetically manipulating these bacteria, further hampering attempts to examine the *in vivo* function of these enzymes. One approach to circumvent this problem is to choose an alternative model system that is amenable to these types of studies. Once a suitable model system is obtained, gene knock-outs and complementation assays could be used to identify the biological function of these proteins. Attractive model systems that should be investigated are the plant pathogen *P. syringae* and the animal pathogen *L. pneumophila*. Both of these organisms are aerobic, have sequenced genomes and an observable phenotype (i.e. virulence). The protein from *P. syringae* has been identified as a virulence factor (Bretz et al. 2003; Espinosa et al. 2003) making it a particularly interesting target. The question of whether the *in vivo* targets of these proteins are inositol phosphates could be addressed by labeling cells with

radiolabelled inositol phosphates and comparing the inositol phosphate profile of the cell before and after infection by *P. syringae* or *L. pneumophila*. By identifying changes in the inositol phosphate pool, the targeted molecule and/or signaling pathway could be identified. This approach has been used to study the role of inositol phosphate kinases in yeast and an established protocol is available (Azevedo and Saiardi 2006).

Although the above system could be used to identify putative *in vivo* targets of the PTPLPs expressed by pathogenic microbes, it is unlikely that the proteins expressed by the non-pathogenic bacteria function similarly. The prevalence of these proteins in gastrointestinal tracts suggests that these proteins may function as phosphate and/or nutrient scavenging proteins. It has been shown that these proteins are constitutively expressed in low- and high-phosphate environments making it unlikely that they function as phosphate scavengers (Yanke et al. 1998; Puhl et al. 2008a). An alternative possibility that has not been explored is that these proteins are involved in divalent cation uptake. $InsP_6$ has been shown to have siderophore activity, and bacteria can import iron from $InsP_6$-Fe complexes (Smith et al. 1994; Hirst et al. 1999). It has not been determined if the $InsP_6$-Fe complex or just the iron is transported into the cell. Note that no inositol polyphosphate transporters have been identified to date. It seems more likely that the iron is being released from the $InsP_6$ complex and subsequently transported into the cell by an iron specific transport system. $InsP_6$ has been shown to associate with the outer cell membrane (Poyner et al. 1993) in a metal dependent fashion. Interestingly, Smith and colleagues (1994) indicate that they detected degradation of the $InsP_6$-Fe complex although they did not identify why or how it was being broken down. The latter result suggests that membrane-associated phytases could be involved in the $InsP_6$ mediated

uptake of iron in bacteria. The phytase would break down the $InsP_6$-Fe complex thus releasing the chelated iron so that it can be transported. A similar mechanism could be used for the uptake of any essential divalent cation that is chelated by $InsP_6$. Interestingly, both PhyAsr and PhyAmm are associated with the bacterial outer membrane. To identify the biological role of these proteins it will be necessary to develop the tools to work with these anaerobic bacteria.

Another interesting system that may prove amenable to *in vivo* studies is *Bdellovibrio bacteriovorus*. This bacterium is interesting in that it is a parasitic bacterium that preys on gram-negative bacteria. *Bdellovibrio* enters the periplasm of gram-negative bacteria where it replicates until it breaks open the cell releasing its progeny into the environment. The released cells can then attack other prey cells. An interesting question that has arisen is why does a bacterial predator have an inositol phosphatase? Bacteria do not have inositol phosphate signalling pathways like eukaryotic cells so it is unlikely that this protein functions analogously to the PTPLPs expressed by plant and human pathogens. One possibility is that it functions in divalent cation uptake similar to the function proposed for the gastro-intestinal tract microbes. Alternatively, it may not function as an inositol phosphatase *in vivo*. This system could potentially serve as a suitable *in vivo* system as it would be easy to follow the effect of a gene knockout on predation.

6.3 Mechanism of $InsP_6$ degradation by PTPLPs and HAPs

The results of our binding studies suggest that PTPLPs do not utilize a classical Michaelis-Menten reaction mechanism. This is based on the observed discrepancy

between the K_d and K_m for $InsP_6$ and MIHS in PhyAsr and the multi-phasic binding kinetics observed when examining the pre-steady state kinetics of substrate binding. In addition, five distinct ligand binding conformations have been observed in PhyAsr indicating that there is a degree of flexibility in substrate binding. This flexibility is likely important for binding the multiple related inositol phosphates that are generated during the degradation of $InsP_6$.

A comparison of the binding affinity for MIHS (K_i) and the K_m for $InsP_6$ in the HAP phytase class reveals a similar discrepancy and suggests that these phytases may use a similar reaction mechanism. Both HAPs and PTPLPs appear to bind substrate similarly and go through a phospho-enzyme intermediate during their catalytic mechanisms. The similarities between PTPLPs and HAPs is striking and suggests that the kinetics of substrate binding should be examined in HAPs to determine whether they also utilize a non-Michaelis-Menten reaction mechanism. This could be done using a fluorescence based assay similar to that used to study substrate binding to PhyAsr, or by doing binding assays with radiolabeled substrate. One of the problems encountered with the fluorescence based assay is that it is not sensitive enough to reach the protein concentrations that are needed to measure the tight binding interaction between PhyAsr and $InsP_6$. Radiolabeled assays are much more sensitive and should allow us to measure this interaction. There are several protocols for the production and purification of radiolabeled inositol phosphates (Stephens et al. 1991; Azevedo and Saiardi 2006). If HAPs are also found to use a non-Michaelis-Menten mechanism, it will be necessary to re-evaluate the reaction kinetics of this class of phytase.

References:

Almo, S.C., Bonanno, J.B., Sauder, J.M., Emtage, S., Dilorenzo, T.P., Malashkevich, V., Wasserman, S.R., Swaminathan, S., Eswaramoorthy, S., Agarwal, R. et al. 2007. Structural genomics of protein phosphatases. *J Struct Funct Genomics* **8**(2-3): 121-140.

Alonso, A., Sasin, J., Bottini, N., Friedberg, I., Osterman, A., Godzik, A., Hunter, T., Dixon, J., and Mustelin, T. 2004. Protein tyrosine phosphatases in the human genome. *Cell* **117**(6): 699-711.

Andersen, J.N., Mortensen, O.H., Peters, G.H., Drake, P.G., Iversen, L.F., Olsen, O.H., Jansen, P.G., Andersen, H.S., Tonks, N.K., and Moller, N.P. 2001. Structural and evolutionary relationships among protein tyrosine phosphatase domains. *Mol Cell Biol* **21**(21): 7117-7136.

Azevedo, C. and Saiardi, A. 2006. Extraction and analysis of soluble inositol polyphosphates from yeast. *Nat Protoc* **1**(5): 2416-2422.

Barford, D., Flint, A.J., and Tonks, N.K. 1994. Crystal structure of human protein tyrosine phosphatase 1B. *Science* **263**(5152): 1397-1404.

Barr, A.J., Ugochukwu, E., Lee, W.H., King, O.N., Filippakopoulos, P., Alfano, I., Savitsky, P., Burgess-Brown, N.A., Muller, S., and Knapp, S. 2009. Large-scale structural analysis of the classical human protein tyrosine phosphatome. *Cell* **136**(2): 352-363.

Barrientos, L., Scott, J.J., and Murthy, P.P. 1994. Specificity of hydrolysis of phytic acid by alkaline phytase from lily pollen. *Plant Physiol* **106**(4): 1489-1495.

Bendtsen, J.D., Nielsen, H., von Heijne, G., and Brunak, S. 2004. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* **340**(4): 783-795.

Bennett, M., Onnebo, S.M., Azevedo, C., and Saiardi, A. 2006. Inositol pyrophosphates: metabolism and signaling. *Cell Mol Life Sci* **63**(5): 552-564.

Blanchetot, C. and den Hertog, J. 2000. Multiple interactions between receptor protein-tyrosine phosphatase (RPTP) alpha and membrane-distal protein-tyrosine phosphatase domains of various RPTPs. *J Biol Chem* **275**(17): 12446-12452.

Blank, G.E., Pletcher, J., and Sax, M. 1971. The structure of *myo*-inositol hexaphosphate dodecasodium salt octatricontahydrate: A single crystal X-ray analysis. *Biochem Biophys Res Commun* **44**: 319-325.

Bliska, J.B., Guan, K.L., Dixon, J.E., and Falkow, S. 1991. Tyrosine phosphate hydrolysis of host proteins by an essential *Yersinia* virulence determinant. *Proc Natl Acad Sci U S A* **88**(4): 1187-1191.

Boss, W.F., Davis, A.J., Im, Y.J., Galvo, M., and Perera, I.Y. 2006. Phosphoinositide Metabolism: Towards an Understanding of Subcellular Signalling. In *Biology of Inositol and Phosphoinositides*, Vol 39, pp. 181-205. Springer, New York, NY.

Bretz, J.R., Mock, N.M., Charity, J.C., Zeyad, S., Baker, C.J., and Hutcheson, S.W. 2003. A translocated protein tyrosine phosphatase of *Pseudomonas syringae* pv. *tomato* DC3000 modulates plant defence response to infection. *Mol Microbiol* **49**(2): 389-400.

Brunger, A.T. 2007. Version 1.2 of the Crystallography and NMR system. *Nat Protoc* **2**(11): 2728-2733.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S. et al. 1998. Crystallography & NMR system (CNS): A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**(Pt 5): 905-921.

Buhrman, G., Parker, B., Sohn, J., Rudolph, J., and Mattos, C. 2005. Structural mechanism of oxidative regulation of the phosphatase CDC25B via an intramolecular disulfide bond. *Biochemistry* **44**(14): 5307-5316.

Buist, A., Zhang, Y.L., Keng, Y.F., Wu, L., Zhang, Z.Y., and den Hertog, J. 1999. Restoration of potent protein-tyrosine phosphatase activity into the membrane-distal domain of receptor protein-tyrosine phosphatase alpha. *Biochemistry* **38**(3): 914-922.

Burke, T.R., Jr. and Zhang, Z.Y. 1998. Protein-tyrosine phosphatases: structure, mechanism, and inhibitor discovery. *Biopolymers* **47**(3): 225-241.

Caffrey, J.J., Darden, T., Wenk, M.R., and Shears, S.B. 2001. Expanding coincident signaling by PTEN through its inositol 1,3,4,5,6-penta*kis*phosphate 3-phosphatase activity. *FEBS Lett* **499**(1-2): 6-10.

Caffrey, J.J., Hidaka, K., Matsuda, M., Hirata, M., and Shears, S.B. 1999. The human and rat forms of multiple inositol polyphosphate phosphatase: functional homology with a histidine acid phosphatase up-regulated during endochondral ossification. *FEBS Lett* **442**(1): 99-104.

Caselli, A., Marzocchini, R., Camici, G., Manao, G., Moneti, G., Pieraccini, G., and Ramponi, G. 1998. The inactivation mechanism of low molecular weight phosphotyrosine-protein phosphatase by H_2O_2. *J Biol Chem* **273**(49): 32554-32560.

CCP4. 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**(Pt 5): 760-763.

Chatterjee, S., Sankaranarayanan, R., and Sonti, R.V. 2003. PhyA, a secreted protein of *Xanthomonas oryzae* pv. *oryzae*, is required for optimum virulence and growth on phytic acid as a sole phosphate source. *Mol Plant Microbe Interact* **16**(11): 973-982.

Chen, Y.Y., Chu, H.M., Pan, K.T., Teng, C.H., Wang, D.L., Wang, A.H., Khoo, K.H., and Meng, T.C. 2008. Cysteine S-nitrosylation protects protein-tyrosine phosphatase 1B against oxidation-induced permanent inactivation. *J Biol Chem* **283**(50): 35265-35272.

Cheng, C. and Lim, B.L. 2006. Beta-propeller phytases in the aquatic environment. *Arch Microbiol* **185**(1): 1-13.

Chu, H.M., Guo, R.T., Lin, T.W., Chou, C.C., Shr, H.L., Lai, H.L., Tang, T.Y., Cheng, K.J., Selinger, B.L., and Wang, A.H. 2004. Structures of *Selenomonas ruminantium* phytase in complex with persulfated phytate: DSP phytase fold and mechanism for sequential substrate hydrolysis. *Structure* **12**(11): 2015-2024.

Cirri, P., Chiarugi, P., Camici, G., Manao, G., Raugei, G., Cappugi, G., and Ramponi, G. 1993. The role of Cys12, Cys17 and Arg18 in the catalytic mechanism of low-M(r) cytosolic phosphotyrosine protein phosphatase. *Eur J Biochem* **214**(3): 647-657.

Collins, B.M., McCoy, A.J., Kent, H.M., Evans, P.R., and Owen, D.J. 2002. Molecular architecture and functional model of the endocytic AP2 complex. *Cell* **109**(4): 523-535.

D'Silva, C.G., Bae, H.D., Yanke, L.J., Cheng, K.J., and Selinger, L.B. 2000. Localization of phytase in *Selenomonas ruminantium* and *Mitsuokella multiacidus* by transmission electron microscopy. *Can J Microbiol* **46**(4): 391-395.

Davies, G.J., Ducros, V.M., Varrot, A., and Zechel, D.L. 2003. Mapping the conformational itinerary of beta-glycosidases by X-ray crystallography. *Biochem Soc Trans* **31**(Pt 3): 523-527.

DeLano, W.L. 2008. The PyMOL Molecular Graphics System. . DeLano Scientific LLC Palo Alto, CA, USA.

Denu, J.M. and Tanner, K.G. 2002. Redox regulation of protein tyrosine phosphatases by hydrogen peroxide: detecting sulfenic acid intermediates and examining reversible inactivation. *Methods Enzymol* **348**: 297-305.

Espinosa, A., Guo, M., Tam, V.C., Fu, Z.Q., and Alfano, J.R. 2003. The *Pseudomonas syringae* type III-secreted protein HopPtoD2 possesses protein tyrosine phosphatase activity and suppresses programmed cell death in plants. *Mol Microbiol* **49**(2): 377-387.

Evans, P.R. 1997. Scala. *Joint CCP4 and ESF-EAMCB Newsletter* (33): 22-24.

Fauman, E.B., Cogswell, J.P., Lovejoy, B., Rocque, W.J., Holmes, W., Montana, V.G., Piwnica-Worms, H., Rink, M.J., and Saper, M.A. 1998. Crystal structure of the catalytic domain of the human cell cycle control phosphatase, CDC25A. *Cell* **93**(4): 617-625.

Fingerling, G. 1912. Utilization by ruminants of the phosphorous compounds of feedstuffs. *Biochem Z* **37**: 265.

Flint, A.J., Tiganis, T., Barford, D., and Tonks, N.K. 1997. Development of "substrate-trapping" mutants to identify physiological substrates of protein tyrosine phosphatases. *Proc Natl Acad Sci U S A* **94**(5): 1680-1685.

Ford, M.G., Pearse, B.M., Higgins, M.K., Vallis, Y., Owen, D.J., Gibson, A., Hopkins, C.R., Evans, P.R., and McMahon, H.T. 2001. Simultaneous binding of PtdIns(4,5)P_2 and clathrin by AP180 in the nucleation of clathrin lattices on membranes. *Science* **291**(5506): 1051-1055.

Frederick, J.P., Mattiske, D., Wofford, J.A., Megosh, L.C., Drake, L.Y., Chiou, S.T., Hogan, B.L., and York, J.D. 2005. An essential role for an inositol polyphosphate multikinase, Ipk2, in mouse embryogenesis and second messenger production. *Proc Natl Acad Sci U S A* **102**(24): 8454-8459.

Fu, Y. and Galan, J.E. 1998. The *Salmonella typhimurium* tyrosine phosphatase SptP is translocated into host cells and disrupts the actin cytoskeleton. *Mol Microbiol* **27**(2): 359-368.

Gaidarov, I., Krupnick, J.G., Falck, J.R., Benovic, J.L., and Keen, J.H. 1999. Arrestin function in G protein-coupled receptor endocytosis requires phosphoinositide binding. *EMBO J* **18**(4): 871-881.

Gasteiger, E., Hoogland, C., Gattiker, C., Duvaud, S., Wilkins, M.R., Appel, R.D., and Bairoch, A. 2005. Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*, (ed. J.M. Walker), pp. 571-607. Humana Press.

Gill, S.C. and von Hippel, P.H. 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Analytical Biochemistry* **182**(2): 319-326.

Goodrich, J.A. and Kugel, J.F. 2007. *Binding and Kinetic for Molecular Biologists*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Gray, C.H., Good, V.M., Tonks, N.K., and Barford, D. 2003. The structure of the cell cycle protein Cdc14 reveals a proline-directed protein phosphatase. *EMBO J* **22**(14): 3524-3535.

Greiner, R. and Carlsson, N.G. 2006. *Myo*-inositol phosphate isomers generated by the action of a phytate-degrading enzyme from *Klebsiella terrigena* on phytate. *Can J Microbiol* **52**(8): 759-768.

Groen, A., Lemeer, S., van der Wijk, T., Overvoorde, J., Heck, A.J., Ostman, A., Barford, D., Slijper, M., and den Hertog, J. 2005. Differential oxidation of protein-tyrosine phosphatases. *J Biol Chem* **280**(11): 10298-10304.

Gruninger, R.J., Selinger, L.B., and Mosimann, S.C. 2008. Effect of ionic strength and oxidation on the P-loop conformation of the protein tyrosine phosphatase-like phytase, PhyAsr. *FEBS J* **275**(15): 3783-3792.

Gruninger, R.J., Selinger, L.B., and Mosimann, S.C. 2009. Structural Analysis of a Multifunctional, Tandemly Repeated Inositol Polyphosphatase. *J Mol Biol* **In press**.

Guan, K.L. and Dixon, J.E. 1991. Evidence for protein-tyrosine-phosphatase catalysis proceeding via a cysteine-phosphate intermediate. *J Biol Chem* **266**(26): 17026-17030.

Guex, N. and Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714-2723.

Guo, X.L., Shen, K., Wang, F., Lawrence, D.S., and Zhang, Z.Y. 2002. Probing the molecular basis for potent and selective protein-tyrosine phosphatase 1B inhibition. *J Biol Chem* **277**(43): 41014-41022.

Ha, N.C., Oh, B.C., Shin, S., Kim, H.J., Oh, T.K., Kim, Y.O., Choi, K.Y., and Oh, B.H. 2000. Crystal structures of a novel, thermostable phytase in partially and fully calcium-loaded states. *Nat Struct Biol* **7**(2): 147-153.

Haefner, S., Knietsch, A., Scholten, E., Braun, J., Lohscheidt, M., and Zelder, O. 2005. Biotechnological production and applications of phytases. *Appl Microbiol Biotechnol* **68**(5): 588-597.

Hanakahi, L.A., Bartlet-Jones, M., Chappell, C., Pappin, D., and West, S.C. 2000. Binding of inositol phosphate to DNA-PK and stimulation of double-strand break repair. *Cell* **102**(6): 721-729.

Hegeman, C.E. and Grabau, E.A. 2001. A novel phytase with sequence similarity to purple acid phosphatases is expressed in cotyledons of germinating soybean seedlings. *Plant Physiol* **126**(4): 1598-1608.

Heinonen, J.K. and Lahti, R.J. 1981. A new and convenient colorimetric determination of inorganic orthophosphate and its application to the assay of inorganic pyrophosphatase. *Anal Biochem* **113**(2): 313-317.

Hengge, A.C., Sowa, G.A., Wu, L., and Zhang, Z.Y. 1995. Nature of the transition state of the protein-tyrosine phosphatase-catalyzed reaction. *Biochemistry* **34**(43): 13982-13987.

Hirst, P.H., Riley, A.M., Mills, S.J., Spiers, I.D., Poyner, D.R., Freeman, S., Potter, B.V., and Smith, A.W. 1999. Inositol polyphosphate-mediated iron transport in *Pseudomonas aeruginosa*. *J Appl Microbiol* **86**(3): 537-543.

Holm, L., Kaariainen, S., Rosenstrom, P., and Schenkel, A. 2008. Searching protein structure databases with DaliLite v.3. *Bioinformatics* **24**(23): 2780-2781.

Huang, H., Shi, P., Wang, Y., Luo, H., Shao, N., Wang, G., Yang, P., and Yao, B. 2009. Diversity of beta-propeller phytase genes in the intestinal contents of grass carp provides insight into the release of major phosphorus from phytate in nature. *Appl Environ Microbiol* **75**(6): 1508-1516.

Illies, C., Gromada, J., Fiume, R., Leibiger, B., Yu, J., Juhl, K., Yang, S.N., Barma, D.K., Falck, J.R., Saiardi, A. et al. 2007. Requirement of inositol pyrophosphates for full exocytotic capacity in pancreatic beta cells. *Science* **318**(5854): 1299-1302.

Irvine, R.F. and Schell, M.J. 2001. Back in the water: the return of the inositol phosphates. *Nat Rev Mol Cell Biol* **2**(5): 327-338.

Isbrandt, L.R. and Oertel, R.P. 1980. Conformational states of *myo*-inositol hexa*kis*phosphate in aqueous solution. A ^{13}C NMR, ^{31}P NMR and Raman spectroscopy investigation. *J Am Chem Soc* **102**: 3144-3148.

Jackson, S.G., Zhang, Y., Haslam, R.J., and Junop, M.S. 2007. Structural analysis of the carboxy terminal PH domain of pleckstrin bound to D-*myo*-inositol 1,2,3,5,6-penta*kis*phosphate. *BMC Struct Biol* **7**: 80.

Jia, Z., Barford, D., Flint, A.J., and Tonks, N.K. 1995. Structural basis for phosphotyrosine peptide recognition by protein tyrosine phosphatase 1B. *Science* **268**(5218): 1754-1758.

Kashio, N., Matsumoto, W., Parker, S., and Rothstein, D.M. 1998. The second domain of the CD45 protein tyrosine phosphatase is critical for interleukin-2 secretion and substrate recruitment of TCR-zeta *in vivo*. *J Biol Chem* **273**(50): 33856-33863.

Keng, Y.F., Wu, L., and Zhang, Z.Y. 1999. Probing the function of the conserved tryptophan in the flexible loop of the *Yersinia* protein-tyrosine phosphatase. *Eur J Biochem* **259**(3): 809-814.

Kennelly, P.J. and Potts, M. 1999. Life among the primitives: protein O-phosphatases in prokaryotes. *Front Biosci* **4**: D372-385.

Kerovuo, J., Rouvinen, J., and Hatzack, F. 2000. Analysis of *myo*-inositol hexa*kis*phosphate hydrolysis by *Bacillus* phytase: indication of a novel reaction mechanism. *Biochem J* **352 Pt 3**: 623-628.

Konietzny, U. and Greiner, R. 2002. Molecular and catalytic properties of phytate-degrading enzymes (phytases). *Int J Food Sci Technol* **37**(7): 791-812.

Kostrewa, D., Gruninger-Leitch, F., D'Arcy, A., Broger, C., Mitchell, D., and van Loon, A.P. 1997. Crystal structure of phytase from *Aspergillus ficuum* at 2.5 Å resolution. *Nat Struct Biol* **4**(3): 185-190.

Kostrewa, D., Wyss, M., D'Arcy, A., and van Loon, A.P. 1999. Crystal structure of *Aspergillus niger* pH 2.5 acid phosphatase at 2. 4 Å resolution. *J Mol Biol* **288**(5): 965-974.

Krissinel, E. and Henrick, K. 2007. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**(3): 774-797.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**(5259): 680-685.

Laskowski, R.A., Hutchinson, E.G., Michie, A.D., Wallace, A.C., Jones, M.L., and Thornton, J.M. 1997. PDBsum: a Web-based database of summaries and analyses of all PDB structures. *Trends Biochem Sci* **22**(12): 488-490.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* **26**: 283-291.

Lassen, S.F., Breinholt, J., Ostergaard, P.R., Brugger, R., Bischoff, A., Wyss, M., and Fuglsang, C.C. 2001. Expression, gene cloning, and characterization of five novel phytases from four basidiomycete fungi: *Peniophora lycii*, *Agrocybe pediades*, a *Ceriporia sp*., and *Trametes pubescens*. *Appl Environ Microbiol* **67**(10): 4701-4707.

Lee, D.C., Cottrill, M.A., Forsberg, C.W., and Jia, Z. 2003. Functional insights revealed by the crystal structures of *Escherichia coli* glucose-1-phosphatase. *J Biol Chem* **278**(33): 31412-31418.

Lee, J.O., Yang, H., Georgescu, M.M., Di Cristofano, A., Maehama, T., Shi, Y., Dixon, J.E., Pandolfi, P., and Pavletich, N.P. 1999. Crystal structure of the PTEN tumor suppressor: implications for its phosphoinositide phosphatase activity and membrane association. *Cell* **99**(3): 323-334.

Leslie, A.G.W. 1992. Recent Changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* (No 26).

Lim, B.L., Yeung, P., Cheng, C., and Hill, J.E. 2007. Distribution and diversity of phytate-mineralizing bacteria. *ISME J* **1**(4): 321-330.

Lim, D., Golovan, S., Forsberg, C.W., and Jia, Z. 2000. Crystal structures of *Escherichia coli* phytase and its complex with phytate. *Nat Struct Biol* **7**(2): 108-113.

Lim, K.L., Ng, C.H., and Pallen, C.J. 1999. Catalytic activation of the membrane distal domain of protein tyrosine phosphatase epsilon, but not CD45, by two point mutations. *Biochim Biophys Acta* **1434**(2): 275-283.

Liu, Q., Huang, Q., Lei, X.G., and Hao, Q. 2004. Crystallographic snapshots of *Aspergillus fumigatus* phytase, revealing its enzymatic dynamics. *Structure* **12**(9): 1575-1583.

Lohse, D.L., Denu, J.M., Santoro, N., and Dixon, J.E. 1997. Roles of aspartic acid-181 and serine-222 in intermediate formation and hydrolysis of the mammalian protein-tyrosine-phosphatase PTP1. *Biochemistry* **36**(15): 4568-4575.

Lupardus, P.J., Shen, A., Bogyo, M., and Garcia, K.C. 2008. Small molecule-induced allosteric activation of the *Vibrio cholerae* RTX cysteine protease domain. *Science* **322**(5899): 265-268.

Macbeth, M.R., Schubert, H.L., Vandemark, A.P., Lingam, A.T., Hill, C.P., and Bass, B.L. 2005. Inositol hexa*kis*phosphate is bound in the ADAR2 core and required for RNA editing. *Science* **309**(5740): 1534-1539.

Majerus, P.W., Zou, J., Marjanovic, J., Kisseleva, M.V., and Wilson, M.P. 2008. The role of inositol signaling in the control of apoptosis. *Adv Enzyme Regul* **48**: 10-17.

McRee, D.E. 1999. XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. *J Struct Biol* **125**(2-3): 156-165.

Michell, R.H. 2008. Inositol derivatives: evolution and functions. *Nat Rev Mol Cell Biol* **9**: 151-161.

Milano, S.K., Kim, Y.M., Stefano, F.P., Benovic, J.L., and Brenner, C. 2006. Nonvisual arrestin oligomerization and cellular localization are regulated by inositol hexa*kis*phosphate binding. *J Biol Chem* **281**(14): 9812-9823.

Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., and Olson, A.J. 1998. Automated docking using a Lamarckain genetic algorithm and an empirical binding free energy function. *J Comput Chem* **19**: 1639-1662.

Mullaney, E.J., Daly, C.B., and Ullah, A.H. 2000. Advances in phytase research. *Adv Appl Microbiol* **47**: 157-199.

Mullaney, E.J. and Ullah, A.H. 2003. The term phytase comprises several different classes of enzymes. *Biochem Biophys Res Commun* **312**(1): 179-184.

Mulugu, S., Bai, W., Fridy, P.C., Bastidas, R.J., Otto, J.C., Dollins, D.E., Haystead, T.A., Ribeiro, A.A., and York, J.D. 2007. A conserved family of enzymes that phosphorylate inositol hexa*kis*phosphate. *Science* **316**(5821): 106-109.

Murphy, A.M., Otto, B., Brearley, C.A., Carr, J.P., and Hanke, D.E. 2008. A role for inositol hexa*kis*phosphate in the maintenance of basal resistance to plant pathogens. *Plant J* **56**(4): 638-652.

Murthy, P.P.N. 2006. Structure and Nomenclature of Inositol Phosphates, Phosphoinositides, and Glycosylphosphatidylinositols. In *Biology of Inositols and Phosphoinositides*, Vol 39, pp. 1-20. Springer, New York, NY.

Nakashima, B.A., McAllister, T.A., Sharma, R., and Selinger, L.B. 2007. Diversity of phytases in the rumen. *Microb Ecol* **53**(1): 82-88.

Nam, H.J., Poy, F., Krueger, N.X., Saito, H., and Frederick, C.A. 1999. Crystal structure of the tandem phosphatase domains of RPTP LAR. *Cell* **97**(4): 449-457.

Nam, H.J., Poy, F., Saito, H., and Frederick, C.A. 2005. Structural basis for the function and regulation of the receptor protein tyrosine phosphatase CD45. *J Exp Med* **201**(3): 441-452.

Navaza, J. 1994. AMoRe: an automated package for molecular replacement. *Acta Crystallogr A* **50**: 157-163.

Ng, D.H., Maiti, A., and Johnson, P. 1995. Point mutation in the second phosphatase domain of CD45 abrogates tyrosine phosphatase activity. *Biochem Biophys Res Commun* **206**(1): 302-309.

Nicholls, A., Sharp, K.A., and Honig, B. 1991. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**(4): 281-296.

Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* **10**(1): 1-6.

Norris, F.A., Wilson, M.P., Wallis, T.S., Galyov, E.E., and Majerus, P.W. 1998. SopB, a protein required for virulence of *Salmonella dublin*, is an inositol phosphate phosphatase. *Proc Natl Acad Sci U S A* **95**(24): 14057-14059.

Oh, B.C., Chang, B.S., Park, K.H., Ha, N.C., Kim, H.K., Oh, B.H., and Oh, T.K. 2001. Calcium-dependent catalytic activity of a novel phytase from *Bacillus amyloliquefaciens* DS11. *Biochemistry* **40**(32): 9669-9676.

Oh, B.C., Kim, M.H., Yun, B.S., Choi, W.C., Park, S.C., Bae, S.C., and Oh, T.K. 2006. Ca(2+)-inositol phosphate chelation mediates the substrate specificity of beta-propeller phytase. *Biochemistry* **45**(31): 9531-9539.

Orchiston, E.A., Bennett, D., Leslie, N.R., Clarke, R.G., Winward, L., Downes, C.P., and Safrany, S.T. 2004. PTEN M-CBR3, a versatile and selective regulator of inositol 1,3,4,5,6-penta*kis*phosphate (Ins(1,3,4,5,6)P_5). Evidence for Ins(1,3,4,5,6)P_5 as a proliferative signal. *J Biol Chem* **279**(2): 1116-1122.

Ostanin, K., Harms, E.H., Stevis, P.E., Kuciel, R., Zhou, M.M., and Van Etten, R.L. 1992. Overexpression, site-directed mutagenesis, and mechanism of *Escherichia coli* acid phosphatase. *J Biol Chem* **267**(32): 22830-22836.

Ostanin, K. and Van Etten, R.L. 1993. Asp304 of *Escherichia coli* acid phosphatase is involved in leaving group protonation. *J Biol Chem* **268**(28): 20778-20784.

Otwinowski, Z. and Minor, W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**: 307-326.

Pannifer, A.D., Flint, A.J., Tonks, N.K., and Barford, D. 1998. Visualization of the cysteinyl-phosphate intermediate of a protein-tyrosine phosphatase byX-ray crystallography. *J Biol Chem* **273**(17): 10454-10462.

Perrakis, A., Morris, R., and Lamzin, V.S. 1999. Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* **6**(5): 458-463.

Peters, G.H., Frimurer, T.M., and Olsen, O.H. 1998. Electrostatic evaluation of the signature motif (H/V)CX_5R(S/T) in protein-tyrosine phosphatases. *Biochemistry* **37**(16): 5383-5393.

Petrey, D. and Honig, B. 2003. GRASP2: visualization, surface properties, and electrostatics of macromolecular structures and sequences. *Methods Enzymol* **374**: 492-509.

Poyner, D.R., Cooke, F., Hanley, M.R., Reynolds, D.J., and Hawkins, P.T. 1993. Characterization of metal ion-induced [^{3}H]inositol hexa*kis*phosphate binding to rat cerebellar membranes. *J Biol Chem* **268**(2): 1032-1038.

Puhl, A.A., Greiner, R., and Selinger, L.B. 2008a. Kinetics, substrate specificity, and stereospecificity of two new protein tyrosine phosphatase-like inositol polyphosphatases from *Selenomonas lacticifex*. *Biochem Cell Biol* **86**(4): 322-330.

Puhl, A.A., Greiner, R., and Selinger, L.B. 2008b. A protein tyrosine phosphatase-like inositol polyphosphatase from *Selenomonas ruminantium* subsp. *lactilytica* has specificity for the 5-phosphate of *myo*-inositol hexa*kis*phosphate. *Int J Biochem Cell Biol* **40**(10): 2053-2064.

Puhl, A.A., Greiner, R., and Selinger, L.B. 2009. Stereospecificity of *myo*-inositol hexa*kis*phosphate hydrolysis by a protein tyrosine phosphatase-like inositol polyphosphatase from *Megasphaera elsdenii*. *Appl Microbiol Biotechnol* **82**(1): 95-103.

Puhl, A.A., Gruninger, R.J., Greiner, R., Janzen, T.W., Mosimann, S.C., and Selinger, L.B. 2007. Kinetic and structural analysis of a bacterial protein tyrosine phosphatase-like *myo*-inositol polyphosphatase. *Protein Sci* **16**(7): 1368-1378.

Raboy, V. 2003. *Myo*-inositol-1,2,3,4,5,6-hexa*kis*phosphate. *Phytochemistry* **64**(6): 1033-1043.

Ragon, M., Hoh, F., Aumelas, A., Chiche, L., Moulin, G., and Boze, H. 2009. Structure of *Debaryomyces castellii* CBS 2923 phytase. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **65**(Pt 4): 321-326.

Rao, D.E., Rao, K.V., Reddy, T.P., and Reddy, V.D. 2009. Molecular characterization, physiochemical properties, known and potential applications of phytases: An overview. *Crit Rev Biotechnol* **29**(2): 182-198.

Rayapureddi, J.P., Kattamuri, C., Steinmetz, B.D., Frankfort, B.J., Ostrin, E.J., Mardon, G., and Hegde, R.S. 2003. Eyes absent represents a class of protein tyrosine phosphatases. *Nature* **426**(6964): 295-298.

Rigden, D.J., Walter, R.A., Phillips, S.E., and Fothergill-Gilmore, L.A. 1999. Polyanionic inhibitors of phosphoglycerate mutase: combined structural and biochemical analysis. *J Mol Biol* **289**(4): 691-699.

Ross, S.H., Lindsay, Y., Safrany, S.T., Lorenzo, O., Villa, F., Toth, R., Clague, M.J., Downes, C.P., and Leslie, N.R. 2007. Differential redox regulation within the PTP superfamily. *Cell Signal* **19**(7): 1521-1530.

Rudolph, J. 2002. Catalytic mechanism of CDC25. *Biochemistry* **41**(49): 14613-14623.

Saiardi, A., Bhandari, R., Resnick, A.C., Snowman, A.M., and Snyder, S.H. 2004. Phosphorylation of proteins by inositol pyrophosphates. *Science* **306**(5704): 2101-2105.

Saiardi, A., Resnick, A.C., Snowman, A.M., Wendland, B., and Snyder, S.H. 2005. Inositol pyrophosphates regulate cell death and telomere length through phosphoinositide 3-kinase-related protein kinases. *Proc Natl Acad Sci U S A* **102**(6): 1911-1914.

Saiardi, A., Sciambi, C., McCaffery, J.M., Wendland, B., and Snyder, S.H. 2002. Inositol pyrophosphates regulate endocytic trafficking. *Proc Natl Acad Sci U S A* **99**(22): 14206-14211.

Salmeen, A., Andersen, J.N., Myers, M.P., Meng, T.C., Hinks, J.A., Tonks, N.K., and Barford, D. 2003. Redox regulation of protein tyrosine phosphatase 1B involves a sulphenyl-amide intermediate. *Nature* **423**(6941): 769-773.

Sasakawa, N., Sharif, M., and Hanley, M.R. 1995. Metabolism and biological activities of inositol penta*kis*phosphate and inositol hexa*kis*phosphate. *Biochem Pharmacol* **50**(2): 137-146.

Scapin, G., Patel, S., Patel, V., Kennedy, B., and Asante-Appiah, E. 2001. The structure of apo protein-tyrosine phosphatase 1B C215S mutant: more than just an S → O change. *Protein Sci* **10**(8): 1596-1605.

Shah, N., Kuntz, D.A., and Rose, D.R. 2003. Comparison of kifunensine and 1-deoxymannojirimycin binding to class I and II alpha-mannosidases demonstrates different saccharide distortions in inverting and retaining catalytic mechanisms. *Biochemistry* **42**(47): 13812-13816.

Shi, L., Potts, M., and Kennelly, P.J. 1998. The serine, threonine, and/or tyrosine-specific protein kinases and protein phosphatases of prokaryotic organisms: a family portrait. *FEMS Microbiol Rev* **22**(4): 229-253.

Shin, S., Ha, N.C., Oh, B.C., Oh, T.K., and Oh, B.H. 2001. Enzyme mechanism and catalytic property of beta propeller phytase. *Structure* **9**(9): 851-858.

Smith, A.W., Poyner, D.R., Hughes, H.K., and Lambert, P.A. 1994. Siderophore activity of *myo*-inositol hexa*kis*phosphate in *Pseudomonas aeruginosa*. *J Bacteriol* **176**(12): 3455-3459.

Song, H., Hanlon, N., Brown, N.R., Noble, M.E., Johnson, L.N., and Barford, D. 2001. Phosphoprotein-protein interactions revealed by the crystal structure of kinase-associated phosphatase in complex with phosphoCDK2. *Mol Cell* **7**(3): 615-626.

Stephens, L.R., Hawkins, P.T., Stanley, A.F., Moore, T., Poyner, D.R., Morris, P.J., Hanley, M.R., Kay, R.R., and Irvine, R.F. 1991. *myo*-inositol penta*kis*phosphates. Structure, biological occurrence and phosphorylation to *myo*-inositol hexa*kis*phosphate. *Biochem J* **275 (Pt 2)**: 485-499.

Stewart, A.E., Dowd, S., Keyse, S.M., and McDonald, N.Q. 1999. Crystal structure of the MAPK phosphatase Pyst1 catalytic domain and implications for regulated activation. *Nat Struct Biol* **6**(2): 174-181.

Streb, H., Irvine, R.F., Berridge, M.J., and Schulz, I. 1983. Release of Ca^{2+} from a nonmitochondrial intracellular store in pancreatic acinar cells by inositol-1,4,5-trisphosphate. *Nature* **306**(5938): 67-69.

Street, I.P., Coffman, H.R., and Poulter, C.D. 1991. Isopentenyl diphosphate isomerase. Site-directed mutagenesis of Cys139 using "counter" PCR amplification of an expression plasmid. *Tetrahedron* **47**(31): 5919-5924.

Strong, M., Sawaya, M.R., Wang, S., Phillips, M., Cascio, D., and Eisenberg, D. 2006. Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* **103**(21): 8060-8065.

Stuckey, J.A., Schubert, H.L., Fauman, E.B., Zhang, Z.Y., Dixon, J.E., and Saper, M.A. 1994. Crystal structure of *Yersinia* protein tyrosine phosphatase at 2.5 Å and the complex with tungstate. *Nature* **370**(6490): 571-575.

Tan, X., Calderon-Villalobos, L.I., Sharon, M., Zheng, C., Robinson, C.V., Estelle, M., and Zheng, N. 2007. Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature* **446**(7136): 640-645.

Terwilliger, T.C. 2001. Map-likelihood phasing. *Acta Crystallogr D Biol Crystallogr* **57**(Pt 12): 1763-1775.

Terwilliger, T.C. and Berendzen, J. 1999. Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr* **55**(Pt 4): 849-861.

Ullah, A.H. and Cummins, B.J. 1988. *Aspergillus ficuum* extracellular pH 6.0 optimum acid phosphatase: purification, N-terminal amino acid sequence, and biochemical characterization. *Prep Biochem* **18**(1): 37-65.

Ullah, A.H. and Sethumadhavan, K. 1998. *Myo*-inositol hexasulfate is a potent inhibitor of *Aspergillus ficuum* phytase. *Biochem Biophys Res Commun* **251**(1): 260-263.

Ullah, A.H., Sethumadhavan, K., Lei, X.G., and Mullaney, E.J. 2000. Biochemical characterization of cloned *Aspergillus fumigatus* phytase (phyA). *Biochem Biophys Res Commun* **275**(2): 279-285.

van Montfort, R.L., Congreve, M., Tisi, D., Carr, R., and Jhoti, H. 2003. Oxidation state of the active-site cysteine in protein tyrosine phosphatase 1B. *Nature* **423**(6941): 773-777.

Verbsky, J., Lavine, K., and Majerus, P.W. 2005. Disruption of the mouse inositol 1,3,4,5,6-penta*kis*phosphate 2-kinase gene, associated lethality, and tissue distribution of 2-kinase expression. *Proc Natl Acad Sci U S A* **102**(24): 8448-8453.

Wilden, B., Savelsbergh, A., Rodnina, M.V., and Wintermeyer, W. 2006. Role and timing of GTP binding and hydrolysis during EF-G-dependent tRNA translocation on the ribosome. *Proc Natl Acad Sci U S A* **103**(37): 13670-13675.

Wu, L. and Zhang, Z.Y. 1996. Probing the function of Asp128 in the lower molecular weight protein-tyrosine phosphatase-catalyzed reaction. A pre-steady-state and steady-state kinetic investigation. *Biochemistry* **35**(17): 5426-5434.

Xiang, T., Liu, Q., Deacon, A.M., Koshy, M., Kriksunov, I.A., Lei, X.G., Hao, Q., and Thiel, D.J. 2004. Crystal structure of a heat-resilient phytase from *Aspergillus fumigatus*, carrying a phosphorylated histidine. *J Mol Biol* **339**(2): 437-445.

Yang, J., Groen, A., Lemeer, S., Jans, A., Slijper, M., Roe, S.M., den Hertog, J., and Barford, D. 2007. Reversible oxidation of the membrane distal domain of receptor PTPalpha is mediated by a cyclic sulfenamide. *Biochemistry* **46**(3): 709-719.

Yanke, L.J., Bae, H.D., Selinger, L.B., and Cheng, K.J. 1998. Phytase activity of anaerobic ruminal bacteria. *Microbiology* **144**: 1565-1573.

York, J.D., Odom, A.R., Murphy, R., Ives, E.B., and Wente, S.R. 1999. A phospholipase C-dependent inositol polyphosphate kinase pathway required for efficient messenger RNA export. *Science* **285**(5424): 96-100.

Yuvaniyama, J., Denu, J.M., Dixon, J.E., and Saper, M.A. 1996. Crystal structure of the dual specificity protein phosphatase VHR. *Science* **272**(5266): 1328-1331.

Zhang, Z.Y. 1995. Kinetic and mechanistic characterization of a mammalian protein-tyrosine phosphatase, PTP1. *J Biol Chem* **270**(19): 11199-11204.

Zhang, Z.Y. 1998. Protein-tyrosine phosphatases: biological function, structural characteristics, and mechanism of catalysis. *Crit Rev Biochem Mol Biol* **33**(1): 1-52.

Zhang, Z.Y. 2002. Protein tyrosine phosphatases: structure and function, substrate specificity, and inhibitor development. *Annu Rev Pharmacol Toxicol* **42**: 209-234.

Zhang, Z.Y. 2003. Mechanistic studies on protein tyrosine phosphatases. In *Progress in Nucleic Acid Research and Molecular Biology, Vol 73*, pp. 171-220. ACADEMIC PRESS INC, SAN DIEGO.

Zhang, Z.Y., Clemens, J.C., Schubert, H.L., Stuckey, J.A., Fischer, M.W., Hume, D.M., Saper, M.A., and Dixon, J.E. 1992. Expression, purification, and physicochemical characterization of a recombinant *Yersinia* protein tyrosine phosphatase. *J Biol Chem* **267**(33): 23759-23766.

Zhang, Z.Y., Malachowski, W.P., Van Etten, R.L., and Dixon, J.E. 1994a. Nature of the rate-determining steps of the reaction catalyzed by the *Yersinia* protein-tyrosine phosphatase. *J Biol Chem* **269**(11): 8140-8145.

Zhang, Z.Y., Wang, Y., and Dixon, J.E. 1994b. Dissecting the catalytic mechanism of protein-tyrosine phosphatases. *Proc Natl Acad Sci U S A* **91**(5): 1624-1627.

Zhang, Z.Y., Wang, Y., Wu, L., Fauman, E.B., Stuckey, J.A., Schubert, H.L., Saper, M.A., and Dixon, J.E. 1994c. The $Cys(X)_5Arg$ catalytic motif in phosphoester hydrolysis. *Biochemistry* **33**(51): 15266-15270.

Zhang, Z.Y., Wu, L., and Chen, L. 1995. Transition state and rate-limiting step of the reaction catalyzed by the human dual-specificity phosphatase, VHR. *Biochemistry* **34**(49): 16088-16096.

Zhao, Y., Wu, L., Noh, S.J., Guan, K.L., and Zhang, Z.Y. 1998. Altering the nucleophile specificity of a protein-tyrosine phosphatase-catalyzed reaction. Probing the function of the invariant glutamine residues. *J Biol Chem* **273**(10): 5484-5492.

Zhou, D., Chen, L.M., Hernandez, L., Shears, S.B., and Galan, J.E. 2001. A *Salmonella* inositol polyphosphatase acts in conjunction with other bacterial effectors to promote host cell actin cytoskeleton rearrangements and bacterial internalization. *Mol Microbiol* **39**(2): 248-259.

Zhou, G., Denu, J.M., Wu, L., and Dixon, J.E. 1994. The catalytic role of Cys124 in the dual specificity phosphatase VHR. *J Biol Chem* **269**(45): 28084-28090.

Zou, J., Kleywegt, G.J., Stahlberg, J., Driguez, H., Nerinckx, W., Claeyssens, M., Koivula, A., Teeri, T.T., and Jones, T.A. 1999. Crystallographic evidence for substrate ring distortion and protein conformational changes during catalysis in cellobiohydrolase Ce16A from *Trichoderma reesei*. *Structure* **7**(9): 1035-1045.