

**SUBSTRATE SPECIFICITY OF PROTEIN TYROSINE PHOSPHATASE-LIKE
MYO-INOSITOL PHOSPHATASES: PHYA IN COMPLEX WITH INS(1,2,4,5,6)P₅,
INS(1,3,4,5)P₄, AND INS(1,4,5)P₃**

LISZA M. BRUDER

Bachelor of Science, University of Lethbridge, 2011

Master of Science, University of Lethbridge, 2013

A Thesis

Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfillment of the
Requirements for the Degree

DOCTOR OF PHILOSOPHY

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

© Lisza M. Bruder, 2018

SUBSTRATE SPECIFICITY OF PROTEIN TYROSINE PHOSPHATASE-LIKE
MYO-INOSITOL PHOSPHATASES: PHYA IN COMPLEX WITH INS(1,2,4,5,6)P₅,
INS(1,3,4,5)P₄, AND INS(1,4,5)P₃

LISZA M. BRUDER

Date of Defence: December 10, 2018

Dr. Steven C. Mosimann Supervisor	Associate Professor	Ph.D.
Dr. L. Brent Selinger Committee Member	Professor	Ph.D.
Dr. Marc Roussel Committee Member	Professor	Ph.D.
Dr. Dmytro Yevtushenko Internal External Examiner	Associate Professor	Ph.D.
Dr. Kenneth Ng External Examiner University of Calgary	Professor	Ph.D.
Dr. Michael Gerken Chair, Thesis Examination Committee	Professor	Ph.D.

Abstract

In order to understand the substrate specificity of protein tyrosine phosphatase-like *myo*-inositol phosphatases (PTPLPs), I have determined the structure of PhyA from *Mitsuokella multacida* (PhyAmm) and *Selenomonas ruminantium* (PhyAsr) in complex with multiple *myo*-inositol phosphates (IPs). These are the first atomic resolution structures of PTPLPs in complex with InsP₅, InsP₄, or InsP₃ substrates, and represent four of the five known PTPLP:IP complex structures. Based on these structures, I have identified three structural features that determine the substrate specificity of PTPLPs. As part of this work, I demonstrate that the PhyAmm C-terminal repeat and PhyAsr bind substrates using conserved phosphoryl-binding sites. Further, I have determined the InsP₆ hydrolysis pathways for several PTPLPs and present a novel IP specificity assay. With this assay, I identify that the two repeats of PhyAmm have divergent activities that clearly indicate a “divide and conquer” approach to maximize phosphoryl group removal from InsP₆ and mixed IPs.

Acknowledgments

I would like to thank my supervisor Dr. Steven Mosimann for giving me the opportunity to learn and grow as a student in his laboratory, and for his guidance, support and encouragement.

I would also like to thank the members of my supervisory committee, Dr. Marc Roussel, Dr. Elizabeth Schultz, and Dr. L. Brent Selinger, for their helpful comments and discussions.

I am grateful for the financial support from the Natural Sciences and Engineering Research Council of Canada, Alberta Innovate Technologies, and the School of Graduate Studies at the University of Lethbridge.

I would like to thank the members of the Mosimann and Selinger labs, both past and present, for their support and friendship. I would specifically like to thank Dr. L. Brent Selinger for providing the material for the IP specificity assays, and Colyn L. Cleland for providing the InsP₆ hydrolysis pathway of PhyAbb.

Finally, I would like to thank my family and friends for their love, support and encouragement. Without you I would not have made it this far.

Contents

Contents	v
List of Tables	viii
List of Figures	ix
List of Abbreviations	xi
1 Literature Review	1
1.1 <i>myo</i> -inositol	1
1.2 <i>myo</i> -inositol phosphates	1
1.2.1 <i>myo</i> -inositol-1,2,3,4,5,6-hexakisphosphate	3
1.2.2 <i>myo</i> -inositol pentakisphosphate	4
1.2.3 <i>myo</i> -inositol tetrakisphosphate and <i>myo</i> -inositol triakisphosphate	5
1.2.4 Phosphoinositides	5
1.2.5 <i>myo</i> -inositol pyrophosphate	6
1.3 <i>myo</i> -inositol phosphatases	6
1.3.1 Purple acid phosphatases	8
1.3.2 β -propeller phytases	9
1.3.3 Histidine acid phosphatases	10
1.3.4 Protein tyrosine phosphatase-like <i>myo</i> -inositol phosphatases	11
1.4 Objective	15
2 Bacterial PhyA protein tyrosine phosphatase-like <i>myo</i>-inositol phosphatases in complex with the Ins(1,3,4,5)P₄ and Ins(1,4,5)P₃ second messengers	16
2.1 Introduction	16
2.2 Experimental procedures	18
2.2.1 Expression and purification	18
2.2.2 Crystallization	18
2.2.3 Data collection and image processing	19
2.2.4 Structure refinement and model validation	19
2.2.5 Structure analysis	21
2.2.6 Identification of hydrolysis products	21
2.3 Results	22
2.3.1 Active sites of PTPLPs are pre-formed	22
2.3.2 Alternative binding modes of IPs	24
2.3.3 PTPLP specificity differences	26
2.4 Discussion	30

2.4.1	<i>myo</i> -inositol ring movements compensate for PTPLP active-site rigidity	30
2.4.2	Structural determinants of the substrate specificity of PTPLPs	32
2.4.3	Ins(1,3,4,5)P ₄ is not bound at the catalytic site of the PhyAmm N-terminal repeat	34
2.4.4	Biological implications	34
3	The tandemly repeated protein tyrosine phosphatase-like <i>myo</i>-inositol phosphatase from <i>Mitsuokella multacida</i> in complex with Ins(1,2,4,5,6)P₅ displays divergent <i>myo</i>-inositol phosphate activity	36
3.1	Introduction	36
3.2	Experimental procedures	38
3.2.1	Cloning and mutagenesis	38
3.2.2	Expression and purification	38
3.2.3	IP production	39
3.2.4	Crystallization	39
3.2.5	Data collection and image processing	40
3.2.6	Structure refinement and model validation	40
3.2.7	Structure analysis	40
3.2.8	Identification of hydrolysis products	42
3.2.9	IP Specificity Assay	42
3.3	Results	43
3.3.1	Ins(1,2,4,5,6)P ₅ binding is consistent with the hydrolysis pathway	43
3.3.2	Structural features determining specificity of the C-terminal repeat	44
3.3.3	N-terminal active-site structure predicts different substrate specificity than the C-terminal active site	47
3.3.4	The PhyAmm N- and C-terminal repeats have divergent IP activity	50
3.4	Discussion	56
3.4.1	<i>myo</i> -inositol ring position allows for optimal use of the PhyAmm C-terminal repeat phosphoryl-binding sites	56
3.4.2	Residues stabilizing Ins(1,2,4,5,6)P ₅ binding to the C-terminal repeat diverge from PhyAsr	58
3.4.3	Factors influencing the conformations of the variable loops	59
3.4.4	Substrate specificity of PhyAmm	61
3.4.5	InsP ₆ specificity of the N-terminal repeat	63
4	Conclusions and future directions	65
4.1	Overview	65
4.2	Summary	66
4.3	PTPLPs for producing IPs	68
4.4	Functions of PTPLPs	69
	Bibliography	71

A	Appendix A	88
A.1	Identification of PhyAsr, PhyAmm and PhyAbb hydrolysis products	88
A.2	Figures and tables	88
B	Appendix B	93
B.1	Figures and tables	93

List of Tables

2.1	Data collection and refinement statistics for the PhyAsrC252S·Ins(1,3,4,5)P ₄ , PhyAsrC252S·Ins(1,4,5)P ₃ , and PhyAmmC250S/C548S·Ins(1,3,4,5)P ₄ structures	20
2.2	Electrostatic and hydrogen bond distances in the PhyAsrC252S·Ins(1,3,4,5)-P ₄ , PhyAsrC252S·Ins(1,4,5)P ₃ and PhyAmmC250S/C548S·Ins(1,3,4,5)P ₄ structures	27
3.1	Data collection and refinement statistics for the PhyAmmC250S/C548S·Ins-(1,2,4,5,6)P ₅ structure	41
3.2	Electrostatic and hydrogen bond distances in the PhyAmmC250S/C548S·-InsP ₅ structure	46
3.3	Structurally equivalent residues of the PhyAmm N- and C-terminal active sites that interact with IP substrates	49
3.4	Summary of the activity of wild-type PhyAmm, PhyAmmC250S and PhyAmmC548S against the 11 hour acid hydrolysis mixture	53
3.5	IP ring shifts relative to InsP ₅ in LSQ superposed structures	57
A.1	Pairwise RMSD comparison of least squares superposition of the PhyAsrC252S monomer alone and in complex with IPs	88
A.2	Pairwise RMSD comparison of LSQ superposition of the PhyAsrC252S active site alone and in complex with IPs	89
A.3	Pairwise RMSD comparison of LSQ superposition of the PhyAmm monomers from the P2 ₁ , and P1 space groups, in complex with Ins(1,3,4,5)P ₄ and without	90
A.4	Pairwise RMSD comparison of LSQ superposition of the C-terminal repeat active site of PhyAmmC250S/C548S in complex with Ins(1,3,4,5)P ₄ and without	91
A.5	Less-phosphorylated IP ring shifts relative to InsP ₆ in LSQ superposed structures	91
B.1	Pairwise RMSD (Å) comparison of LSQ superposition of the PhyAmm monomers from the InsP ₅ , Ins(1,3,4,5)P ₄ and apo structures	93
B.2	Pairwise RMSD comparison of LSQ superposition of the C-terminal repeat active site of PhyAmmC250S/C548S alone and in complex with InsP ₅ and Ins(1,3,4,5)P ₄	93

List of Figures

1.1	Diagram of <i>myo</i> -inositol in the energetically favoured chair conformation with five equatorial and one axial hydroxyl group	2
1.2	Diagram of <i>myo</i> -inositol-1,2,3,4,5,6-hexakisphosphate in the energetically favoured chair conformation with five equatorial and one axial phosphoryl group	3
1.3	Structure of the purple acid phosphatase homodimer from <i>Phaseolus vulgaris</i>	8
1.4	Structure of the β -propeller phytase from <i>Bacillus subtilis</i> in complex with <i>myo</i> -inositol-1,2,3,4,5,6-hexakisulfate	9
1.5	Structure of the histidine acid phosphatase from <i>Aspergillus niger</i> in complex with <i>myo</i> -inositol-1,2,3,4,5,6-hexakisulfate	11
1.6	Structure of the protein tyrosine phosphatase-like <i>myo</i> -inositol phosphatase from <i>Selenomonas ruminantium</i> in complex with InsP ₆	12
1.7	Schematic representation of InsP ₆ bound to the active site of PhyAsr	14
2.1	Ribbon diagrams of the overall fold of PhyAsrC252S and PhyAmmC250S/-C548S in complex with Ins(1,3,4,5)P ₄	23
2.2	Ligand density for Ins(1,3,4,5)P ₄ and Ins(1,4,5)P ₃ bound to PhyAsrC252S and PhyAmmC250S/C548S	25
2.3	Stereo view of the IP substrates bound to PhyAsrC252S and the C-terminal repeat of PhyAmmC250S/C548S	26
2.4	The superposition of PhyAsrC252S in complex with InsP ₆ and Ins(1,3,4,5)P ₄ demonstrating the <i>myo</i> -inositol ring tilt and rotation	28
2.5	The variable loops implicated in the substrate specificity of PTPLPs	29
2.6	HPIC chromatograms of the PhyAsr, PhyAmm and PhyAbb InsP ₆ hydrolysis products	30
3.1	³¹ P (¹ H decoupled) spectra of 5 mM purified Ins(1,2,4,5,6)P ₅	39
3.2	Ribbon diagram of the overall fold of PhyAmmC250S/C548S	44
3.3	Electron density for inorganic phosphate and Ins(1,2,4,5,6)P ₅ bound to the N- and C-terminal repeats, respectively, of PhyAmmC250S/C548S	45
3.4	Ins(1,2,4,5,6)P ₅ bound to the C-terminal repeat of PhyAmmC250S/C548S	45
3.5	Superposition of the N- and C-terminal repeats of the PhyAmmC250S/-C548S·InsP ₅ structure	47
3.6	Space-filling stereo view of the N- and C-terminal repeats of the PhyAmmC250S/C548S·InsP ₅ structure	51
3.7	HPIC chromatograms of the PhyAmmC250S and PhyAmmC548S InsP ₆ hydrolysis products	52

3.8	HPIC chromatograms of the activities of wild-type PhyAmm, PhyAmm-C250S and PhyAmmC548S against the 11 hour acid hydrolysis mixture . . .	54
3.9	Modeling of PhyAmmC250S/C548S·Ins(1,2,4,5,6)P ₅ with the substrates from the PhyAmmC250S/C548S·Ins(1,3,4,5)P ₄ and PhyAsrC252S·InsP ₆ structures demonstrating the intermediate position of Ins(1,2,4,5,6)P ₅	56
3.10	Ribbon diagram of the superposition of the C-terminal repeat of the PhyAmm-C250S/C548S·InsP ₅ structure and PhyAsr	59
3.11	Dimer interface of PhyAmm and PhyAsr influencing the conformation of the variable loops	61
3.12	Variable loops implicated in the substrate specificity of PTPLPs	62
A.1	Identification of products from the hydrolysis of InsP ₆ by PhyAsr and PhyAmm	92
B.1	HPIC chromatograms of the activity of PhyAmmC548S against the 11 hour acid hydrolysis mixture.	94

List of Abbreviations

Apo	Structure without IP ligand
BME	β -mercaptoethanol
BPP	β -propeller phytase
CCD	Charge-coupled device
CLS	Canadian Light Source
GA-loop	General acid loop
HAP	Histidine acid phosphatase
HPIC	High performance ion chromatography
Ins(1,2,4,5,6)P ₅	<i>myo</i> -inositol-1,2,4,5,6-pentakisphosphate
Ins(1,3,4,5)P ₄	<i>myo</i> -inositol-1,3,4,5-tetrakisphosphate
Ins(1,4,5)P ₃	<i>myo</i> -inositol-1,4,5-trikisphosphate
InsP ₂	<i>myo</i> -inositol dikisphosphate
InsP ₃	<i>myo</i> -inositol trikisphosphate
InsP ₄	<i>myo</i> -inositol tetrakisphosphate
InsP ₅	<i>myo</i> -inositol pentakisphosphate
InsP ₆	<i>myo</i> -inositol-1,2,3,4,5,6-hexakisphosphate or phytate
IPase	<i>myo</i> -inositol phosphatase
IP	<i>myo</i> -inositol phosphate
IPTG	Isopropyl β -D-1-thiogalactopyranoside
LSQ	Least squares
MIHS	<i>myo</i> -inositol-1,2,3,4,5,6-hexakisulfate
LppA	Phytase A from <i>Legionella pneumophila</i>
PAP	Purple acid phosphatase
PDB	Protein data bank
PEG	Polyethylene glycol
PhyA	Phytase A
PhyAbb	Phytase A from <i>Bdellovibrio bacteriovorus</i>
PhyAmm	Phytase A from <i>Mitsuokella multacida</i>
PhyAsr	Phytase A from <i>Selenomonas ruminantium</i>
PI	Phosphoinositides
P-loop	Phosphate-binding loop

PP	Pyrophosphate or diphosphoryl group
PP-InsP ₅	Diphosphoinositol pentakisphosphate
PtInsPs	Phosphatidylinositol phosphates
PtIns(4,5)P ₂	Phosphatidylinositol-4,5-bisphosphate
PTP	Protein tyrosine phosphatase
PTPLP	Protein tyrosine phosphatase-like <i>myo</i> -inositol phosphatase
RMSD	Root-mean-square deviations
TCA	Trichloroacetic acid
Tris	Tris(hydroxymethyl)aminomethane
XopH	Phytase A from <i>Xanthamonas campestris</i> pv. <i>vesicatoria</i>

Chapter 1

Literature Review

1.1 *myo*-inositol

Inositol is a cyclic carbohydrate with hydroxyl groups attached to the six carbons of the ring (Figure 1.1). *myo*-inositol is the stereoisomer most often used in biology where one hydroxyl is in the axial position and the other five are equatorial in the lowest energy chair conformation [1, 2]. *myo*-inositol is also known as *meso*-inositol as it has a plane of symmetry through the C2 and C5 atoms [3]. The remaining atoms are prochiral pairs (C1/C3 and C4/C6) and numbers can be assigned relative to C2 in either the D- or L- configuration (Figure 1.1). Generally, unless the L- configuration is specified, the abbreviation Ins is regarded as *myo*-inositol in the D- configuration as per IUPAC recommendations. Since the original isolation of *myo*-inositol in 1850, it has been shown to have a central role in cellular metabolism [4, 5]. *myo*-inositol is now known to be ubiquitous in eukaryotes, present in many archaea, and within a few bacteria [2, 3, 6–8]. *myo*-inositol is present in the cells as a variety of derivatives including *myo*-inositol phosphates (IPs) [5].

1.2 *myo*-inositol phosphates

In 1872, Pfeffer demonstrated that an organic phosphate salt was present in wheat grain which was later identified to contain *myo*-inositol and phosphate [9–12]. The different phosphorylation levels of *myo*-inositol phosphates (IPs) were indistinguishable leading to all IPs being referred to as phytic acid or phytin, where phytin refers to IPs in the salt form [13, 14]. In 1951, the heterogeneity of phytic acid was revealed by the application of

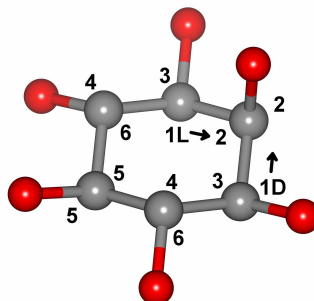


Figure 1.1: Ball and stick diagram of *myo*-inositol in the energetically favoured chair conformation with five equatorial and one axial hydroxyl group, carbon is in grey and oxygen in red. *myo*-inositol has a plane of symmetry that rotates the structure about C2 and C5 as fixed positions. The remaining positions are two prochiral pairs: C1/C3 and C4/C6. When numbered clockwise (inside the ring) the assigned number on carbon 1 is 1L. When numbered counterclockwise (outside the ring) the assigned number is 1D. Generally, the abbreviation Ins should be regarded as *myo*-inositol in the D- configuration, unless the L-configuration is specified.

ion exchange procedures and later shown to included a mixture of IPs and esters of inositols other than *myo*-inositol [10, 15, 16]. In 1962, experiments following phytate formation during rice grain maturation demonstrated that levels of phosphoinositides (PIs), IPs, and *myo*-inositol fluctuate [17].

Since their initial discovery, phosphorylated *myo*-inositols containing from one to eight phosphoryl groups have been identified and many have demonstrated biological activity [2, 18, 19]. In general, highly-phosphorylated IPs serve as cofactors, whereas less-phosphorylated IPs are utilized as second messengers in signal transduction pathways. Less-abundant IPs are difficult to detect and identify at sub-micromolar concentrations and novel functional roles are still being discovered [20, 21]. IPs are particularly challenging to purify or synthesize and only a small subset of known IPs are commercially available [22]. Consequently, improved methods for producing, detecting, and identifying IPs are likely to lead to additional discoveries regarding the biological roles of this family of metabolites.

1.2.1 *myo*-inositol-1,2,3,4,5,6-hexakisphosphate

myo-inositol-1,2,3,4,5,6-hexakisphosphate (InsP₆ or phytic acid) is the most abundant phosphorylated *myo*-inositol derivative (Figure 1.2) [23]. First identified as a storage source for *myo*-inositol and inorganic phosphate in seeds, InsP₆ has since been shown to have multiple roles in the cellular processes such as regulating plant hormone receptors, iron transport, translation, DNA repair, RNA processing, mRNA export, plant development, apoptosis, and pathogenicity [3, 17, 24–32]. The importance of InsP₆ in cells is emphasized by the observation that deletion of enzymes involved in InsP₆ biosynthesis has a lethal phenotype in mouse embryos [33, 34].

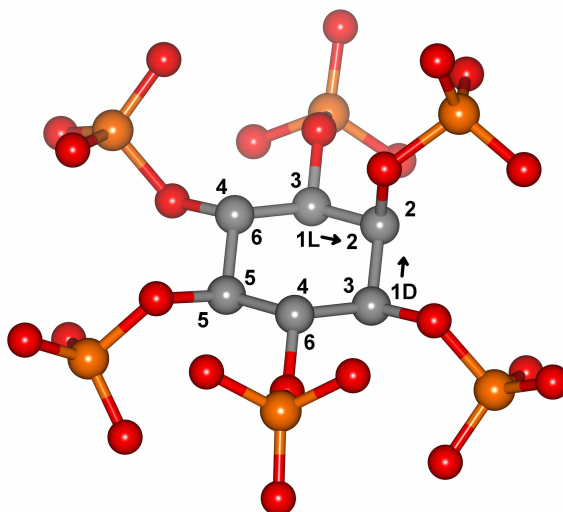


Figure 1.2: Ball and stick diagram of *myo*-inositol-1,2,3,4,5,6-hexakisphosphate (InsP₆ or phytic acid) in the energetically favoured chair conformation with five equatorial and one axial phosphoryl group, carbon is in grey, oxygen in red and phosphorus in orange. InsP₆ has a plane of symmetry that rotates the structure about the C2 and C5 as fixed positions. The remaining positions are two prochiral pairs: C1/C3 and C4/C6. When numbered clockwise (inside the ring) the assigned number on carbon 1 is 1L. When numbered counterclockwise (outside the ring) the assigned number is 1D. Generally, the abbreviation Ins should be regarded as *myo*-inositol in the D- configuration, unless otherwise specified in the L-configuration.

The structure of InsP₆ was given in 1907 and later confirmed with ³¹P-NMR [11, 14, 35, 36]. Subsequently, it was determined that InsP₆ in solution preferentially adopts the five equatorial and one axial phosphoryl group conformation [37, 38]. InsP₆ undergoes a

conformational inversion between pH 9.2 and 9.6 from the five equatorial and one axial conformation (pH <9.2) to one equatorial and five axial phosphoryl groups (pH >9.6) [38]. Between pH 9.2 and 9.6 InsP₆ is in dynamic equilibrium, likely caused by the association/dissociation of one or multiple of the least acidic protons [38]. InsP₆ has twelve sites that can be protonated across the six phosphoryl groups [37, 38]. Of the twelve sites, six are strongly acidic with pK_a values between 1.1 and 2.1, three sites are slightly acidic with pK_a values between 6.0 to 7.6 and the last three are weakly acidic with pK_a values between 9.2 and 9.6 [37, 38]. The pK_a values are dependent on the counter-ion present and their concentration in solution as InsP₆ is a strong chelator [38–41]. More recent studies demonstrate the preferred chair conformation of less-phosphorylated IPs and the pH at which they undergo conformational inversion is unique to the particular isomer [42].

1.2.2 *myo*-inositol pentakisphosphate

When the method to resolve different IPs was discovered, the presence of *myo*-inositol pentakisphosphate (InsP₅) was identified in soil samples [16]. Subsequently, *myo*-inositol-1,3,4,5,6-pentakisphosphate (Ins(1,3,4,5,6)P₅) was shown to be the most abundant IP in avian erythrocytes where it functions as a hemoglobin effector [14, 43–45]. Following the identification the role of *myo*-inositol-1,4,5-trikisphosphate (Ins(1,4,5)P₃) as a releaser of intracellular calcium (discussed below), the presence of Ins(1,3,4,5,6)P₅ was identified in pituitary cells and found to respond to hormone stimulation [46, 47]. As methods were developed to distinguish between the different InsP₅ isomers, they were shown to respond to hormone signaling, affect rates of cell proliferation, modulate apoptotic responses, regulate viral assembly, chromatin remodeling, calcium channel activity, cytoskeleton reorganization, function as a hemoglobin effector, and functions as a precursor to several metabolites [18, 45, 47–57].

1.2.3 *myo*-inositol tetrakisphosphate and *myo*-inositol triakisphosphate

Generally, the less-phosphorylated IPs, *myo*-inositol tetrakisphosphate (InsP₄) and *myo*-inositol triakisphosphate (InsP₃), function as second messengers. Discovery of their role in signal transduction pathways followed work demonstrating that the cleavage of phosphoinositol stimulated calcium release, leading to the identification of Ins(1,4,5)P₃ as a second messenger in 1983 [46, 58, 59]. In this calcium mobilizing pathway, phosphatidylinositol-4,5-bisphosphate (PtdIns(4,5)P₂) is hydrolyzed to generate the secondary messenger Ins(1,4,5)P₃ in response to several hormonal stimuli resulting in the release of calcium from the endoplasmic reticulum [60]. Further work demonstrated a related second messenger, Ins(1,3,4,5)P₄, increases the sensitivity of cells towards Ins(1,4,5)P₃ in a cooperative manner and results in a longer-lasting signal [18, 61, 62]. Additional less-phosphorylated IPs have characterized roles in cellular processes such as chloride secretion, iron accumulation, chromatin remodeling, gene expression, and hormone signaling [49, 53, 63–66]

1.2.4 Phosphoinositides

In 1930, the presence of inositol within the lipid fraction of bacteria was discovered and the presence of inositol containing phospholipids (phosphatidylinositol: PtdIns) has been identified in many archaea, all eukaryotes and a few bacteria [2, 67, 68]. The phosphorylated forms of PtdIns are collectively known as phosphoinositides (PIs) with phosphatidylinositol-4,5-bisphosphate (PtdIns(4,5)P₂) being the first identified PI [69–71]. In 1953, the “phospholipid effect” was first observed when external stimulation effected changes in the PI turnover rates [72–75]. PIs typically make up less than 15% of the total cell phospholipids with asymmetric distributions within membranes [72, 76]. Roles of PIs include acting as precursors for second messengers, membrane anchoring sites for proteins, organelle identification and targeting, exocytosis, endocytosis, signal transduction pathways, regulation of integral plasma membrane proteins, and plasma membrane-cytoskeleton interactions [46, 72, 74, 77–84].

1.2.5 *myo*-inositol pyrophosphate

myo-inositol pyrophosphates (PP-InsPs) contain a diphosphoryl group (pyrophosphate; PP) attached to the *myo*-inositol ring. After their initial identification in 1993, PP-InsPs have been found in all eukaryotes tested to date [56, 57, 85]. PP-InsPs are produced by the phosphorylation of InsP₆ and Ins(1,3,4,5,6)P₅, creating high-energy bonds similar to adenosine diphosphate (ADP) and adenosine triphosphate (ATP) [57, 86]. PP-InsPs are able to phosphorylate specific proteins in an ATP and enzyme independent manner, and participate in processes such as endocytosis, exocytosis, ribosome biogenesis, chemotaxis, apoptosis, regulation of telomere length, response to environmental stresses, and chromatin remodeling [53, 87–95].

1.3 *myo*-inositol phosphatases

The discovery of InsP₆ in 1907 led to the identification of *myo*-inositol phosphatases (IPases; also known as phytases) which hydrolyzed InsP₆ into inorganic phosphate and a less-phosphorylated IP or *myo*-inositol [10, 11, 96]. Since their discovery, IPases have been identified in all eukaryotic organisms as components of biosynthetic and regulatory pathways [97, 98]. In a variety of microorganisms, IPases serve as phosphate scavenging proteins and some have been implicated in pathogenesis [2, 32, 99, 100].

Interest in IPases initially arose within the field of nutrition with the observation that cereals are rachitogenic for puppies due to the chelating properties of InsP₆ impairing calcium absorption [101–103]. To increase the nutritional value of cereals, research was aimed to identify conditions that led to InsP₆ hydrolysis [103]. By 1936, researchers understood some ingested InsP₆ was broken down in mammal intestines with three possible IPase sources considered: digestive secretions, the food, or intestinal microorganisms [103–106]. It was ruled out that IPases were produced by digestive secretions and demonstrated that feed containing high levels of endogenous IPase activity promoted weight gain in livestock [104, 107–112]. Subsequent studies indicated the nutrient availability of InsP₆ was

better for ruminants than non-ruminants leading to the discovery in 1956 that ruminant microorganisms were responsible for the majority of InsP₆ breakdown [113]. These observations made it evident that the IPase activity was primarily the actions of the intestinal flora within ruminants [103, 104, 111]. It is now common practice to supplement livestock feed with microbial IPases as it increases the nutritional value of cereal crops, especially in the case of non-ruminants [103, 114]. It also reduces the demand for inorganic phosphorus supplements and decreases phosphate pollution by reducing the amount of InsP₆ excreted into the environment [114–117]. IPase supplements are also found in human foods where they enhance mineral availability and lower the risk of mineral deficiencies caused by the chelating effects of InsP₆ [118–120].

In addition to their use in the feed industry, IPases have been utilized in the biosynthesis of various less-phosphorylated IPs [114]. The importance of IPs within biological systems is becoming more apparent and to better understand their various roles a reliable source is needed. Chemical synthesis is difficult due to the large number of stereoisomers, whereas an enzymatic synthesis has the advantage of high stereospecificity and mild reaction conditions [22, 121]. IPases have been successfully used to produce the less-phosphorylated IPs by incubation with InsP₆ [120–122]. To date the enzymatic approach has had limited application due to difficulties purifying large quantities of pure products [96]. A greater understanding of IPases will provide a means for improving the process.

To date, the majority of IPase research has been directed at improved livestock and human feed supplements resulting in extensive studies of microbial IPases [114]. Four classes of microbial IPases have been described based on primary sequence and structural similarities. These four enzyme classes are purple acid phosphatases (PAPs), β -propeller phytases (BPPs), histidine acid phosphatases (HAPs), and protein tyrosine phosphatase-like *myo*-inositol phosphatases (PTPLPs). Each IPase class is capable of removing multiple phosphoryl groups from InsP₆ and display varying degrees of specificity. 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).

1.3.1 Purple acid phosphatases

In 1973, the first purple acid phosphatase (PAP) was identified in pig allantoic fluid (Uteroferrin) and GmPhy from germinating soybeans was the first PAP that exhibited IPase activity [123–126]. Since then, few PAPs demonstrated IPase activity and they appear to represent a small subset of the PAPs class [126–129]. PAPs consist of homodimers with each monomer composed of an N-terminal antiparallel β -sandwich and a larger C-terminal α + β domain (Figure 1.3) [130]. Each monomer contributes two α -helices to the dimer interface. PAPs have five signature sequences (**DXG**, **GDXXY**, **GNH(D/E)**, **VXXH**, and **GHXH**; bold letters indicate metal ligating residues) that are involved in coordinating the Fe(III)-M(II) centre, where M can be Fe, Mn or Zn [126, 129–132]. The catalytic mechanism involves the direct attack of the scissile phosphate by an activated water coordinated by the metal centre [130]. The optimal pH range for PAPs that function as IPases is between 4.5 and 6 [126, 129].

1.3.2 β -propeller phytases

Originally isolated from *Bacillus* species in 1998, the β -propeller phytases (BPPs) have since been shown to be widely distributed among bacteria [133–136]. BPPs adopt a β -propeller fold with a solvent-accessible central channel that binds calcium ions (Figure 1.4) [137, 138]. The bound calcium increases the stability of the protein and is required for catalytic activity as it provides a favourable electrostatic environment that permits substrate binding [139]. The substrate-binding site is located at the top of the β -propeller and consists of several charged residues to facilitate calcium and InsP₆ binding [138]. The crystallographic structure of *Bacillus subtilis* in complex with the inhibitor myo-inositol-1,2,3,4,5,6-hexakisulfate (MIHS) and the hydrolysis data indicate the enzyme is a 4-phytase and that there are no large-scale conformational changes upon substrate bind-

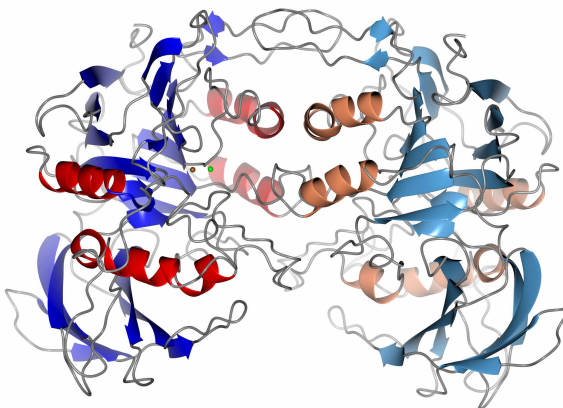


Figure 1.3: Structure of the purple acid phosphatase (PAP) homodimer from *Phaseolus vulgaris* (PDB 1KBP). The iron and zinc ions of the metal centre are shown as brown and green spheres, respectively. The protein is oriented so the N-terminal antiparallel β -sandwich is at the bottom of the homodimer and the C-terminal α + β domain is at the top. The proteins are shown as ribbons with the β -sheets in dark and light blue, α -helices are red and orange, and the loops are in grey.

ing [137–139]. BPP contain two phosphoryl-binding sites where one is the cleavage site and the other binds an adjacent phosphoryl group [140, 141]. Both phosphoryl-binding sites appear to need to be filled for successful hydrolysis explaining why BPPs only remove three phosphoryl groups from InsP_6 [140, 141]. The optimal pH range for the BPPs is between 7 and 8 due to the protonation state of the calcium-binding residues [133, 134, 137].

1.3.3 Histidine acid phosphatases

In 1911, the first and most extensively studied histidine acid phosphatase (HAP) was identified in *Aspergillus niger* and the crystallographic structure determined in 1997 [142, 143]. The HAP class of IPases contain two domains: an α / β -domain from the acid phosphatase family and an α -domain unique to IPase HAPs with the substrate-binding pocket located at the domain interface (Figure 1.5) [143, 144]. The greatest conservation among HAPs is within the α / β -domain that contains the active-site signature sequence (RH(G/N)-XRXP) for binding and catalysis and the downstream HD sequence involved in product binding and release [143, 145]. The α -domain is involved in conferring substrate specificity

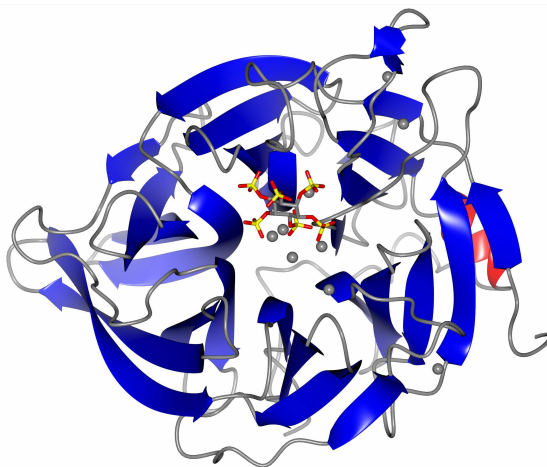


Figure 1.4: Structure of the β -propeller phytase (BPP) from *Bacillus subtilis* in complex with *myo*-inositol-1,2,3,4,5,6-hexakisulfate (MIHS) (PDB 3AMR). Calcium ions are shown as grey spheres and MIHS as cylinders with carbon in grey, sulfur in yellow and oxygen in red. The protein is shown as a ribbon with β -sheets in blue, α -helices in red and the loops in grey.

by contributing multiple interactions with phosphoryl groups not adjacent to the scissile phosphate that stabilizes the orientation of InsP_6 within the active site [144, 146, 147]. The conformation of the single axial phosphoryl group dictates the InsP_6 specificity of both the 3- and 4-phytases [144, 145]. In the case of the 3-phytases, binding of the C3-phosphoryl group for hydrolysis is due to the steric interactions restricting the axial C2-phosphoryl group to a specific location adjacent to the scissile phosphoryl [144, 145]. In the case of the 4-phytases, steric restraints limit the phosphoryl groups adjacent to the scissile phosphoryl group to be in the equatorial conformation placing the axial group away from the binding site [148]. HAPs continue to dephosphorylate the products of the hydrolyzed InsP_6 to produce InsP [145, 148]. HAPs utilize a two-step catalytic mechanism where the first step is the formation of the phospho-histidine intermediate between the scissile phosphate and the histidine in the $\text{RH}(\text{G/N})\text{XRX}\text{P}$ signature sequence [114, 149–152]. In the second step of the reaction the aspartic acid of the HD motif protonates the bridging oxygen of the leaving group as the phospho-histidine intermediate is hydrolyzed by an incoming water molecule, releasing the histidine for subsequent reactions. The optimal pH of the IPase HAPs are 2.5 or the within range of 4.5 to 6 [131, 153].

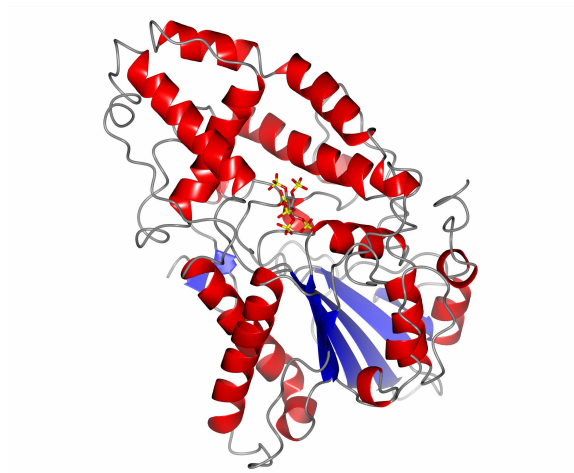


Figure 1.5: Structure of the histidine acid phosphatase (HAP) from *Aspergillus niger* (PhyA) in complex with *myo*-inositol-1,2,3,4,5,6-hexakisulfate (MIHS) (PDB 3K4Q). The ligand is bound in the substrate-binding pocket located at the domain interface with the α -domain at the top and the α/β -domain at the bottom. MIHS is shown as cylinders with carbon in grey, sulfur in yellow and oxygen in red. The protein is shown as a ribbon with β -sheets in blue, α -helices in red and the loops in grey.

1.3.4 Protein tyrosine phosphatase-like *myo*-inositol phosphatases

In 1998, investigation of the rumen organisms with the highest InsP₆ activity resulted in the identification of the first protein tyrosine phosphatase-like *myo*-inositol phosphatases (PTPLPs) [154–156]. Since their initial identification in *Selenomonas ruminantium* and *Mitsuokella multacida*, PTPLPs have been found in a range of bacterial organisms and have been shown to participate in phosphate scavenging, reversing iron-restricted growth inhibition, and pathogenesis [156–161]. The PTPLPs of *S. ruminantium* (PhyAsr) and *M. multacida* (PhyAmm) are associated with the outer membrane which is also the case for the PTPLP of the predatory bacterium *Bdellovibrio bacteriovorus* (PhyAbb), while the PTPLPs of pathogenic bacteria are injected into host cells [155, 157, 159–161]. PTPLPs consist of two domains: the large catalytic PTP domain and a smaller phytase-specific (Phy) domain (Figure 1.6) [162]. The PTP domain is an α - β - α sandwich with two signature sequences that are responsible for catalysis [162–165]. The phosphate-binding loop (P-loop; CXXGXGR(S/T)) provides the nucleophilic cysteine and binds the scissile phosphoryl group, while the general-acid loop (GA-loop; HD) provides an acidic aspartate.

The PTP domain almost exclusively contributes contacts with the IP substrates through the P-loop and GA-loop [166]. The Phy domain is an antiparallel α - β sandwich which has been identified as contributing to the specificity of the PTPLPs [162, 166]. The PTPLP active site is located at the domain interface between the domains with the P-loop forming the base of the active site coordinating the scissile phosphoryl group and the GA-loop at the back [162, 166].

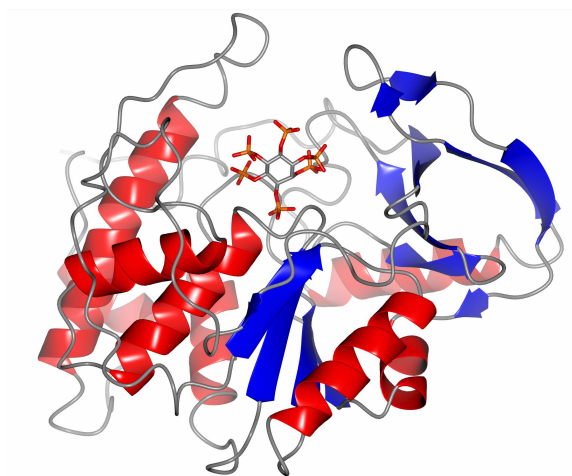


Figure 1.6: Structure of the protein tyrosine phosphatase-like *myo*-inositol phosphatase (PTPLP) from *S. ruminantium* (PhyAsr) in complex with InsP₆ (PDB 3MMJ). The protein is orientated so that the α - β - α sandwich PTP domain is at the bottom and the antiparallel α - β sandwich Phy domain is located at the top right with InsP₆ bound at the domain interface. InsP₆ is shown as cylinders with carbon in grey, phosphorus in orange and oxygen in red. The protein is shown as a ribbon with β -sheets in blue, α -helices in red and the loops in grey.

PTPLPs catalyze the hydrolysis of phosphodiester bonds following a two-step PTP-like mechanism [162–165]. In the first step, the invariant P-loop cysteine thiolate nucleophilically attacks the scissile phosphate. This generates a phospho-cysteine intermediate that is subsequently hydrolyzed. The aspartate of the GA-loop serves as a general base in step two, generating a hydroxyl by abstracting a proton from a water molecule that attacks the phospho-cysteine intermediate, releasing an inorganic phosphate and regenerating the enzyme [164, 167]. Isosteric site-directed mutations of the nucleophile cysteine to serine and the aspartate general acid to asparagine result in complete and partial inac-

tivation of PTPLPs, respectively [165]. This is consistent with the above mechanism and similar studies examining the catalytic mechanism of PTPs. Further, the crystallographic structure of PhyAsrC252A in complex with inorganic phosphate and Ins(1,2,3,5,6)P₅ (PDB 3MOZ) closely resembles the structure of the PTP1B phospho-enzyme intermediate (PDB 1A5Y) [166]. In this work, the smaller alanine side chain allows the inorganic phosphate to bind deeper within the P-loop of the PhyAsrC252A complex structure, where it occupies a position that is structurally equivalent to the phosphoryl group of the PTP1B phospho-enzyme intermediate.

The active sites of PTPLPs have an overall positive electrostatic surface potential to offset the negative charge of the IP substrates [162]. With the predominant electrostatic interactions between the protein and substrates, the activity of PTPLPs vary strongly with ionic strength and have pH optima of 4.5 to 6 [162, 165, 168, 169]. The noted exception is the PTPLP from *Xanthomonas campestris* pv. *vesicatoria* (XopH) with a pH optimum of 7 that is optimized for activity within host cells [161]. PTPLPs classified as 1-, 3-, and 5-phytases have been identified [157, 158, 161, 165, 170]. With the noted exception of XopH which stops at producing Ins(2,3,4,5,6)P₅ from InsP₆, each PTPLP hydrolyzed InsP₆ to Ins(2)P through different dephosphorylation pathways [157, 158, 161, 165, 170]. In each case, PTPLPs have been shown to have activity for IPs other than InsP₆ with varying specificities [157–161, 165, 169, 170].

To facilitate the discussion of substrate conformation within the active site, the six phosphoryl-binding sites identified in the structure of PhyAsrC252S in complex with InsP₆ complex are referred to as the P_s, P_a, P_b, P_{a'}, and P_{b'} sites (Figure 1.7) [166]. The scissile phosphoryl group binds the P_s site formed by P-loop and GA-loops which orients the phosphoryl group for hydrolysis and is conserved in all PTPLPs. The remaining sites are more variable in terms of number of contacts and steric restrictions but have been identified in PhyAsr, PhyAmm and PhyAbb [160, 166, 169]. In PhyAsr, the P_{a'} site forms multiple direct contacts with the only axial phosphoryl group of InsP₆ suggesting strong interactions, and

is likely responsible for the strict C3-phosphoryl group specificity of the enzyme as it can only accommodate the axial phosphoryl or a hydroxyl group [165]. The P_a site forms multiple direct contacts with the substrate also suggesting a strong interaction, while the P_b site makes multiple solvent-mediated contacts suggesting a weaker interaction. In the case of the $P_{b'}$ and P_c sites, they have relatively few contacts which are typically mediated by solvent suggesting a smaller contribution to InsP_6 binding. Taken together, this suggests that the $P_{a'}$, P_a and to a lesser extent the P_b sites are responsible for the bulk of the favourable interactions between the protein and non-scissile phosphoryl groups of PhyAsr .

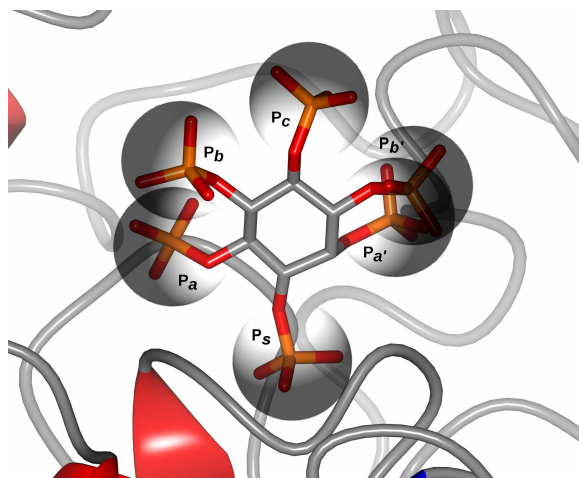


Figure 1.7: Schematic representation of InsP_6 bound to the active site of PhyAsr (PDB 3MMJ). The grey spheres represent the phosphoryl group within the binding sites labelled P_s (scissile phosphate), P_a , $P_{a'}$, P_b , $P_{b'}$, and P_c for reference. The C3-phosphoryl group is bound to the P_s site with the axial C2-phosphoryl group bound to the $P_{a'}$ site. InsP_6 is shown as cylinders with carbon in grey, phosphorus in orange and oxygen in red. The protein is shown as a ribbon with β -sheets in blue, α -helices in red and the loops in grey.

PhyAmm is a tandem repeat that contains two PTPLP folds with distinct activities towards IPs [169]. The C-terminal repeat closely resembles PhyAsr with nearly identical residues available to interact with the substrates. The N-terminal repeat has residue substitutions, insertions and deletions that restrict the space available for IP binding compared to both the C-terminal repeat and PhyAsr . Through activity studies it has been demonstrated that the size of the active site has an influence on the substrate preference, with the C-terminal repeat having a preference for InsP_6 and InsP_5 and the N-terminal repeat hav-

ing almost no activity towards these larger substrates while retaining activity towards the InsP₄ and InsP₃ substrates [169]. Mutagenesis studies suggest that the difference in electrostatic surface potential, mutations in the vicinity of the active sites and steric constraints are involved in the altered substrate specificity of the N-terminal repeat of PhyAmm [169].

1.4 Objective

In order to understand the substrate specificity of protein tyrosine phosphatase-like *myo*-inositol phosphatases (PTPLPs), I have determined the structure the PTPLPs from *Mitsukella multacida* (PhyAmm) and *Selenomonas ruminantium* (PhyAsr) in complex with multiple *myo*-inositol phosphates (IPs). In Chapter 2, I present the structures of PhyAsr in complex with Ins(1,3,4,5)P₄ and Ins(1,4,5)P₃, and PhyAmm in complex with Ins(1,3,4,5)-P₄. Based on these structures, I identify three structural features I predict contribute to the substrate specificity of PTPLPs. I test this hypothesis in Chapter 3 by presenting the structure of PhyAmm in complex with Ins(1,2,4,5,6)P₅. Supporting the structural studies, I present InsP₆ hydrolysis data of the two repeats of PhyAmm as well as the PTPLP from *Bdellovibrio bacteriovorus* (PhyAbb). The structure and biochemical data support the hypothesis that the three structural features I have identified contribute to the specificity of these enzymes. Further, I present a novel IP specificity assay that demonstrates the two repeats of PhyAmm have divergent activities that clearly indicate a “divide and conquer” approach to maximize phosphoryl group removal from mixed IPs.

Chapter 2

Bacterial PhyA protein tyrosine phosphatase-like *myo*-inositol phosphatases in complex with the Ins(1,3,4,5)P₄ and Ins(1,4,5)P₃ second messengers¹

2.1 Introduction

Myo-inositol phosphates (IPs) containing between one and eight phosphoryl groups are ubiquitous in eukaryotic species and have diverse biological activities [18]. The most abundant, *myo*-inositol-1,2,3,4,5,6-hexakisphosphate (InsP₆ or phytate), has multiple important roles in eukaryotic cellular processes including regulating plant hormone receptors, DNA repair, RNA processing, mRNA export, plant development, apoptosis, and pathogenicity [24, 27–32, 98]. Less-abundant, less-phosphorylated IPs have also been implicated in many important biological processes including second messenger activities [18, 19]. For example, the role of Ins(1,4,5)P₃ and Ins(1,3,4,5)P₄ in calcium mobilization has been well characterized [18, 62]. Ins(1,4,5)P₃ stimulates the release of calcium from the endoplasmic reticulum resulting in further cellular responses, and Ins(1,3,4,5)P₄ acts to increase sensitivity and generate a longer-lasting signal. In general, highly-phosphorylated IPs (InsP₆ and

¹This chapter is an adapted version of the manuscript “Bacterial PhyA protein tyrosine phosphatase-like *myo*-inositol phosphatases in complex with the Ins(1,3,4,5)P₄ and Ins(1,4,5)P₃ second messengers. 2017. Bruder, L.M., Gruninger, R.J., Cleland, C.P., and Mosimann, S.C. J Biol Chem. 292: 17302-17311.” L.M.B. wrote the paper and performed the work presented here, with the exception of the PhyAbb pathway which was determined by C.P.C.

InsP₅) serve as cofactors, whereas less-phosphorylated IPs are utilized as second messengers in signal transduction pathways [19]. In contrast to eukaryotic cells, IPs are typically absent from prokaryotes [8, 171], despite the presence of enzymes specific for IPs (IPases) in many organisms. Known microbial IPases have few characterized roles in prokaryotes [99]. Most serve as phosphate scavenging proteins, and more recently others have been implicated in pathogenesis [2, 32, 100].

Microbial IPases that specifically hydrolyze the C3- and C4-phosphoryl groups of InsP₆ (D-*myo* numbering) are common, and experimental structures of both classes of enzyme are available [145, 148, 166]. Structural features that determine C3-phosphoryl group specificity of IPases (3-phytases) have been identified for two distinct enzyme families: protein tyrosine phosphatase-like *myo*-inositol phosphatase (PTPLPs) and histidine-acid phosphatases (HAPs) [145, 166]. In each of these cases, the InsP₆ substrate adopts its lowest energy conformation, and steric interactions restrict the only axial phosphoryl group (C2) to a specific location adjacent to the scissile phosphoryl group [144, 145, 166].

PTPLPs and HAPs hydrolyze InsP₆ to produce *myo*-inositol-2-monokisphosphate (Ins(2)P) and inorganic phosphate after prolonged incubation [148, 157, 158, 165, 170]. Characterized PTPLPs have been shown to hydrolyze InsP₆ to Ins(2)P via different dephosphorylation pathways [157, 158, 165, 170]. An example is PhyA from *Selenomonas ruminantium* (PhyAsr), which first removes the C3-phosphoryl group (P3) followed (in order) by P1, P6, P5, and P4 [165]. PTPLPs such as PhyA from *Mitsuokella multacida* (PhyAmm) and *Legionella pneumophila* (LppA) have demonstrated activity towards the Ins(1,4,5)P₃ second messenger [159, 169]. It is likely these enzymes can also hydrolyze additional IP second messengers and that this activity is shared by other PTPLPs.

Here we have presented the first structures of bacterial PTPLPs in complex with eukaryotic IP second messengers, representing crystallographic evidence that PTPLPs bind and hydrolyze these compounds. These IPs bind in a different conformation than InsP₆ while utilizing a subset of the previously determined phosphoryl-binding sites [166]. Fur-

ther, we have identified several variable loops adjacent to the active site that influence their dephosphorylation pathways. Consequently, this work both refines and extends previous studies aimed at understanding the function and specificity of these enzymes and likely applies to other IPase families.

2.2 Experimental procedures

2.2.1 Expression and purification

The *phyA* genes of the *S. ruminantium* (phyAsrC252S), *M. multacida* (phyAmmC250S/C548S), and *Bdellovibrio bacteriovorus* HD100 (PhyAbb; previously Bd1204), minus the putative signal peptide, were previously cloned into the *NdeI* site of the pET28b^{Kan} expression vector (EMD Biosciences) [160, 165, 169]. The PhyAsrC252S and PhyAmmC250S/C548S proteins are catalytically inactive as a result of the cysteine to serine mutation (an isosteric substitution). All contained an N-terminal His₆ tag and were produced and purified as described previously [160, 165, 166, 169]. PhyAsrC252S was dialyzed into 20 mM ammonium bicarbonate (pH 8.0) and lyophilized, while PhyAmmC252S/C548S was dialyzed into 100 mM Tris-HCl (pH 8.0), 100 mM NaCl, 1 mM β -mercaptoethanol (BME), and 0.1 mM ethylenediaminetetraacetic acid (EDTA; pH 8.0) followed by the addition of glycerol to 20% v/v, and PhyAbb was dialyzed into 50 mM Na acetate (pH 5.0) 300 mM NaCl, 5 mM BME, and 0.1 mM EDTA (pH 8.0). The protein was used immediately or flash-frozen and stored at 193 K.

2.2.2 Crystallization

Crystallization experiments were conducted at room temperature using sitting-drop vapour diffusion with drop ratios of 2 μ L protein solution to 2 μ L reservoir solution. PhyAsrC252S protein solutions were prepared at 20 mg/mL, and crystals were grown as described previously [168]. PhyAmmC250S/C548S protein solutions were concentrated to 4.5 mg/mL using a Millipore Ultracel 10-kDa Centrifugal Filter, and crystals were grown

in 10% w/v polyethylene glycol (PEG) 8000, 100 mM Tris-HCl (pH 8.0), 1 mM BME, 4% v/v ethylene glycol, and 20% v/v glycerol. Following a 24-h equilibration, the reservoir solution was supplemented with 100 μ L of glycerol. After 30 days PhyAsrC252S grew rod-like crystals with approximate dimensions of $30 \times 30 \times 100 \mu\text{m}$, and after 10 days PhyAmmC250S/C548S grew rod-like crystals with approximate dimensions of $100 \times 100 \times 500 \mu\text{m}$. In each case, crystals were soaked in mother liquor supplemented with 10 mM Ins(1,3,4,5)P₄ (15 minutes; Echelon Bioscience) or 10 mM Ins(1,4,5)P₃ (15 minutes; Sigma-Aldrich). Following the introduction of ligand, PhyAsrC252S crystals were transferred to mother liquor containing 22 to 25% v/v glycerol (cryoprotectant) and flash-frozen in liquid nitrogen. PhyAmmC250S/C548S crystals were flash-frozen directly following the introduction of ligand.

2.2.3 Data collection and image processing

Diffraction data ($\lambda = 0.97934 \text{ \AA}$) was collected from frozen crystals (100 K) using a Rayonix MX300 CCD detector at beamline 08ID-1 located at the Canadian Light Source (CLS; Saskatoon, SK, Canada). The space group and unit cell parameters of the PhyAsrC252S crystals in complex with ligand are equivalent to those of the PhyAsrC252S·InsP₆ structure (PDB 3MMJ), whereas the PhyAmmC250S/C548S·Ins(1,3,4,5)P₄ crystal has a novel P1 unit cell. All diffraction image data were processed interactively with MOSFLM prior to scaling and merging within AIMLESS of the CCP4 program suite, version 6.3.0 [172–175]. Data collection statistics are shown in Table 2.1.

2.2.4 Structure refinement and model validation

Phases derived from the PhyAsrC252S·InsP₆ structure (PDB 3MMJ) and wild-type PhyAmm (PDB 3F41) were used to solve the structures by molecular replacement (MOLREP). The refined structures have continuous electron density for main-chain atoms of amino acids 33 to 346 of PhyAsrC252S and 46 to 636 of PhyAmmC250S/C548S, with the remaining residues located at the termini assumed to be disordered. This includes the

Table 2.1: Data collection and refinement statistics for the PhyAsrC252S·Ins(1,3,4,5)P₄, PhyAsrC252S·Ins(1,4,5)P₃, and PhyAmmC250S/C548S·Ins(1,3,4,5)P₄ structures.¹

	PhyAsrC252S Ins(1,3,4,5)P ₄	PhyAsrC252S Ins(1,4,5)P ₃	PhyAmmC250S/C548S Ins(1,3,4,5)P ₄
PDB code	4WTY	4WU2	4WU3
Data collection			
Space group	P2 ₁	P2 ₁	P1
a, b, c (Å)	45.8, 138.2, 80.6	46.0, 137.7, 80.0	73.9, 86.7, 124.2
α, β, γ (°)	90.0, 102.3, 90.0	90.0, 102.4, 90.0	107.3, 91.7, 90.0
Wavelength (Å)	0.97934	0.97934	0.97934
Resolution (Å)	43.0 - 2.10 (2.16 - 2.10)	44.9 - 2.15 (2.21 - 2.15)	44.5 - 2.20 (2.24 - 2.20)
Observed reflections	214 850	175 251	323 812
Unique reflections	57 016	52 695	146 418
Completeness (%)	100 (100)	99.9 (100)	98.1 (96.6)
Redundancy	3.8 (3.7)	3.3 (3.3)	2.2 (2.2)
Rmerge (%)	13.5 (48.1)	12.8 (52.3)	8.0 (16.0)
I / σI	6.2 (2.3)	6.1 (1.9)	7.2 (4.3)
Refinement Statistics			
No. reflections work set	55 187	50 541	144 299
No. reflections test set	1 797	2 154	2 118
Rwork / Rfree (%)	15.7 / 17.5	19.7 / 21.9	19.0 / 21.3
Asymmetric unit	Dimer	Dimer	Dimer of Dimers
Protein atoms	5 057	5 096	19 188
Solvent atoms	663	484	2 262
Ligand atoms	93	83	270
Wilson B (Å ²)	27.51	31.04	19.52
Average B protein (Å ²)	26.2	30.5	18.4
Average B solvent (Å ²)	35.3	35.3	27.4
Average B ligand (Å ²)	42.4	35.3	39.4
RMSD Bonds (Å)	0.005	0.008	0.007
RMSD Angle (°)	1.067	1.153	1.106
Ramachandran distribution			
Preferred (%)	97.04	95.9	96.28
Allowed (%)	2.46	3.44	2.69
Outliers (%)	0.49	0.66	1.03

¹ values in parenthesis are for the highest resolution shell

N-terminal histidine tag of both proteins, residues 28 to 32 of PhyAsrC252S and residues 31 to 45 of PhyAmmC250S/C548S. Refinement was performed using REFMAC, version 5.7, within the CCP4 program suite, while interactive fitting of the models to the electron density were performed in COOT, version 0.6.2, [173, 176]. PROCHECK and structure validation tools with COOT were used throughout refinement to assess the stereochemistry of the model [177]. Unless indicated otherwise, figures were prepared with CCP4mg, version 2.10.8 [178]. Key data processing and refinement statistics for each PhyAsrC252S and PhyAmmC250S/C548S complex structures are presented in Table 2.1.

2.2.5 Structure analysis

These and previously determined structures were compared by least squares (LSQ) superposition using LSQKAB from the CCP4 program suite [173]. The main-chain atoms of residues 35 to 346 of PhyAsrC252S, and 47 to 342 and 343 to 636 for the N- and C-terminal repeats of PhyAmmC250S/C548S, respectively, were used in overall fold comparisons. Active-site comparisons were made using residues 56 to 58, 152 to 154, 189 to 190, 221 to 226, 249 to 262, 304 to 309, and 311 to 313 of PhyAsrC252S, and 448 to 450, 484 to 486, 517 to 522, 545 to 558, 584 to 586, and 599 to 605 of the PhyAmmC250S/C548S C-terminal repeat. Differences in the relative position of the *myo*-inositol ring of bound ligands of PhyAsrC252S were calculated using the superposed structure coordinates and GEOMCALC in the CCP4 program suite [173, 176].

2.2.6 Identification of hydrolysis products

Hydrolysis of 5 mM InsP₆ (Sigma-Aldrich) was carried out at room temperature in the presence of 10 nM wild type PhyAsr or PhyAmm (50 mM sodium acetate pH 5.0, 200 mM NaCl, 1 mM BME, 0.1 mM EDTA). In the case of PhyAbb, 100 nM protein and 10 mM InsP₆ were used. Aliquots of 200 μ L were taken, heat-denatured at 95°C for 2 min, and subjected to high performance ion chromatography (HPIC; Waters 1525 Binary HPIC Pump; Milford, MA) utilizing a CarboPac PA-100 (4 x 240 mm) analytical

column (Dionex; Sunnyvale, CA) [179] at room temperature using a post-column reactor flow rate of 0.2 mL/min. Identification of hydrolysis products utilized a standard hydrolysis chromatogram (Appendix A experimental procedures).

2.3 Results

2.3.1 Active sites of PTPLPs are pre-formed

The overall fold of PTPLPs are composed of a catalytic α - β - α sandwich PTP domain and an antiparallel α - β sandwich Phy domain. The active site includes the general acid loop (GA-loop) and phosphate-binding loop (P-loop) from the catalytic PTP domain as well as residues from both domains that participate in substrate binding [166]. The fold and active site of PhyAsr is shown in Figure 2.1 A. Least squares (LSQ) superpositions of PhyAsr crystal structures in the presence and absence of ligand yield root-mean-square deviations (RMSDs) of approximately 0.2 Å (over 1248 main-chain atoms), clearly demonstrating there are no large-scale main-chain conformational changes associated with ligand binding (Appendix Table A.1) [165, 166]. The same is true within the active site, as the main- and side-chain conformation of the residues alone or in complex with InsP₆, Ins(1,3,4,5)P₄ or Ins(1,4,5)P₃ are essentially identical (Appendix Table A.2).

The PhyAmm monomer is an example of a tandemly repeated IPase that contains two copies of the catalytic PTP and Phy domains (Figure 2.1 B). Although the N- and C-terminal repeats of PhyAmm have different IP substrate specificities, they are both active towards the Ins(1,4,5)P₃ second messenger [169]. Presented here is the complex structure of PhyAmmC252S/C548S with Ins(1,3,4,5)P₄, a related second messenger. The PhyAmmC250S/C548S·Ins(1,3,4,5)P₄ structure was solved in a different space group (P1) than PhyAmm without ligand (P2₁). Between the two structures, the individual repeats of PhyAmm are nearly identical as judged by LSQ superpositions (Appendix Table A.3) [169]. Small movements in the linker region between repeats subtly alters the relative orientation of the tandem repeats and account for the bulk of the observed differences between the

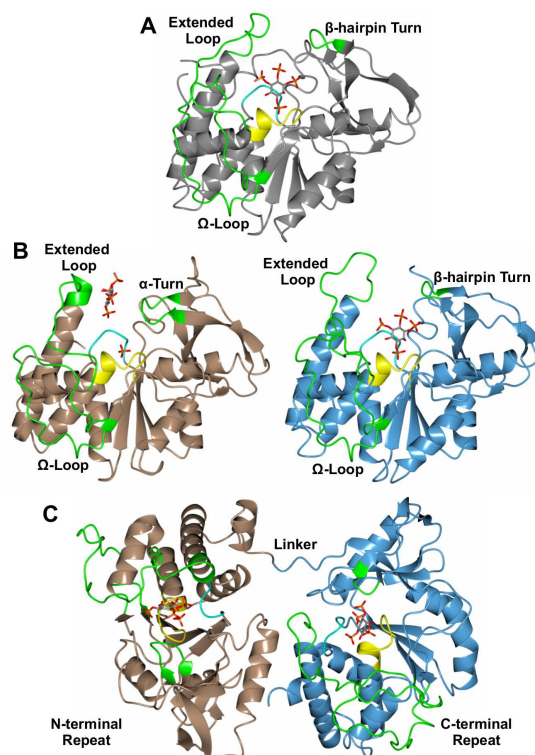


Figure 2.1: Ribbon diagrams of the overall fold of PhyAsrC252S (grey) and PhyAmmC250S/C548S (brown/blue) in complex with Ins(1,3,4,5)P₄. *A*, PhyAsrC252 is oriented so the Ins(1,3,4,5)P₄ ring and phosphoryl groups are clearly visible, with the C1-phosphoryl group bound to the phosphate-binding loop (P-loop; yellow). *B*, the individual N-terminal (brown) and C-terminal (blue) repeats of PhyAmm are shown in the same orientation as *A*. *C*, PhyAmm oriented to view the tandem repeats and linker resulting in the Ins(1,3,4,5)P₄ ligand viewed from the top with the P-loop below the ligand. The ligand atoms are shown as sticks with carbon shown in grey, oxygen in red and phosphorus in orange. The variable loops are shown in green. These include the extended loops prior to the penultimate helix, the α -turn (N-terminal repeat) and β -hairpin turns (PhyAsr and C-terminal repeat), and the extended Ω -loops. The P-loop (yellow) and the general-acid loop (GA-loop; cyan) contribute catalytic residues. Mutation of the cysteine nucleophile to the isosteric serine prevents thiolate formation and renders the enzyme inactive.

individual monomers and dimers. As reported previously, the dimer interface of PhyAmm is almost exclusively formed between the N-terminal repeats, and these interactions are an extensive network of hydrogen bonds, salt bridges, and van der Waals contacts [169]. As seen with PhyAsr, the C-terminal active site of PhyAmmC250S/C548S is virtually identical in the presence or absence of Ins(1,3,4,5)P₄ (Appendix Table A.4). Taken together, the structures of both PhyAsr and PhyAmm support previous suggestions that the PTPLP

family of enzymes have pre-formed active sites [166].

2.3.2 Alternative binding modes of IPs

The electron density for bound ligand is clearly visible in initial $2mF_o - DF_c$ (1.5σ) and $mF_o - DF_c$ (3.5σ) maps of each PhyAsrC252S structure. Refined electron density ($2mF_o - DF_c$) for the Ins(1,3,4,5)P₄ and Ins(1,4,5)P₃ ligands at 1σ are shown in Figure 2.2 A and B. Obvious electron density for the axial C2-hydroxyls and each phosphoryl group clearly identify the conformation of each ligand bound within the active site. Interestingly, PhyAsrC252S binds Ins(1,4,5)P₃ in a position identical to that of the Ins(1,3,4,5)P₄ less one phosphate (Figure 2.3). The modeled IPs adopt the lowest-energy chair conformation with five equatorial hydroxyl/phosphoryl groups and an axial C2-hydroxyl. The C1-phosphoryl groups of the ligands are positioned for hydrolysis in the P_s phosphoryl-binding site by forming extensive interactions with the P-loop (Figure 2.3; Table 2.2) [162, 166]. The P-loop (residues 252 to 259), GA-loop (residues 222 to 225), and Lys-312 are the only interactions that originate from the catalytic PTP domain. The remaining contacts are mediated by residue side chains derived from the Phy domain (residue ranges 51 to 59 and 136 to 203) and the Phy-specific extension of the penultimate helix (residues 291 to 307). The less-phosphorylated IPs utilize a subset of the phosphoryl-binding sites identified previously in the PhyAsrC252S-InsP₆ structure [166].

The *myo*-inositol rings of the less-phosphorylated IP complexes have different relative orientations within the active site when compared with the PhyAsrC252S-InsP₆ structure. In particular, the *myo*-inositol rings are rotated by 180° , resulting in the opposite face contacting the enzyme. For the C1-phosphoryl group to maintain contact with the P-loop, the rotated *myo*-inositol rings tilt toward the GA-loop (Figure 2.4). Overall, the IPs shift by more than 1.2 Å (centre of mass to centre of mass; Appendix Table A.5) and fill the space occupied by ordered solvent in the InsP₆ complex structure.

The N- and C-terminal active sites of PhyAmmC250S/C548S are non-equivalent and

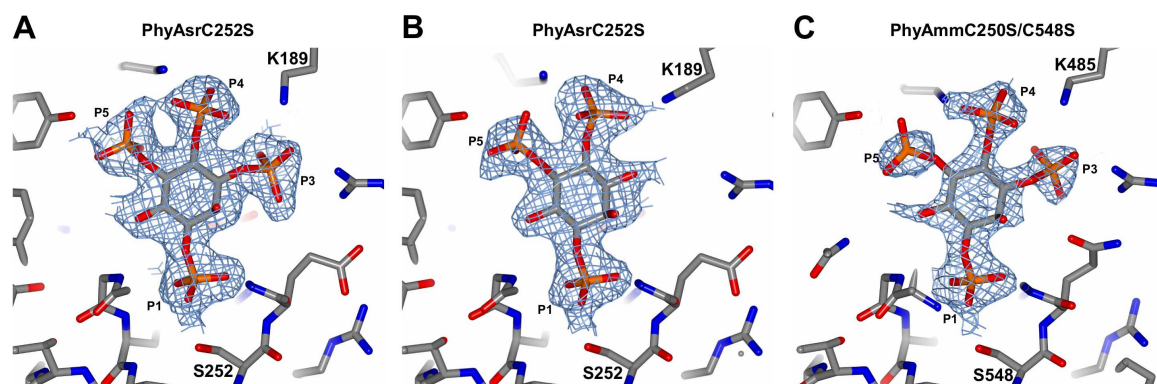


Figure 2.2: Clear electron density for the phosphoryl groups and C2-hydroxyl allow for an unambiguous fit of the ligands and places the C1-phosphoryl group (P1) above the cysteine to serine mutations at position 252 (PhyAsr) and 548 (PhyAmm). The refined $2mF_o - DF_c$ electron density, contoured at 1σ (blue mesh) for PhyAsrC252S in complex with Ins(1,3,4,5)P₄ (A) and Ins(1,4,5)P₃ (B) and for the PhyAmmC250S/C548S C-terminal repeat in complex with Ins(1,3,4,5)P₄ (C). Ligand and protein are shown as sticks with oxygen in red, nitrogen in blue, phosphorus in orange, and carbon in grey.

have different substrate specificities [169]. The C-terminal active site of PhyAmm is highly active toward InsP₆ and 12 of the 14 residues contacting the ligand are conserved when compared with PhyAsr (Table 2.2). Not unexpectedly, Ins(1,3,4,5)P₄ binding within the PhyAmmC250S/C548S C-terminal active site is nearly identical to that observed in the PhyAsrC252S·Ins(1,3,4,5)P₄ structure. In contrast, the N-terminal active site of PhyAmm does not bind ligand in a catalytically competent manner. Instead, an inorganic phosphate is bound by the catalytic P-loop, and a partially occupied (0.6) Ins(1,3,4,5)P₄ is bound at a novel site more than 8 Å (phosphorus to phosphorus) from the inorganic phosphate (Figure 2.3 C). The N-terminal active site has residue substitutions as well as a two-residue insertion. The insertion is in the Phy-specific domain and generates a larger α -turn (residues 182 to 188) that extends into the active site. The equivalent region of PhyAsr and the C-terminal repeat of PhyAmm have smaller β -hairpin turns. The mutations in the N-terminal repeat of PhyAmm reduce the positive electrostatic surface potential and introduce additional steric limitations.

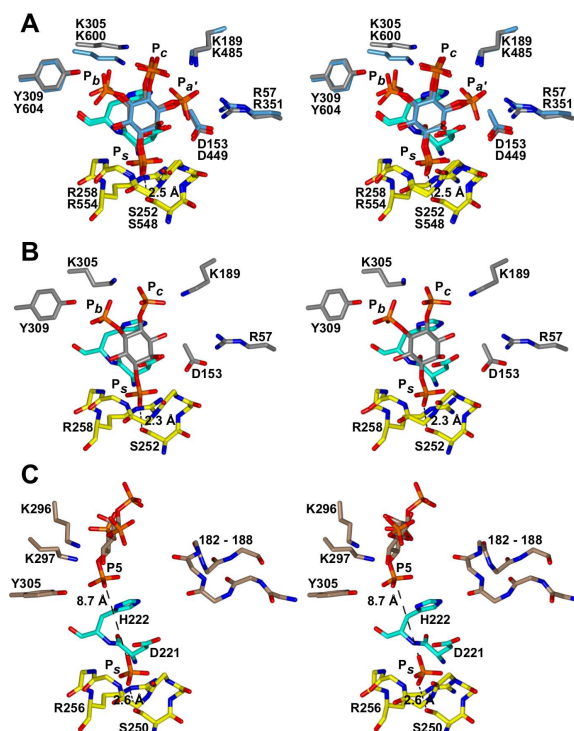


Figure 2.3: PhyAsrC252S and the C-terminal repeat of PhyAmmC250S/C548S bind the IP substrates nearly identically using a subset of the contacts identified in the PhyAsrC252-InsP₆ structure (PDB 3MMJ) [166]. The phosphoryl-binding sites are labeled according to the InsP₆ structure (P_s, P_b, P_a, and P_c). A, stereo view of the superposition of PhyAsrC252S (grey) and the C-terminal repeat of PhyAmmC250S/C548S (blue) in complex with Ins(1,3,4,5)P₄. B, PhyAsrC252S in complex with Ins(1,4,5)P₃ in the same conformation as A less one phosphoryl group. C, the N-terminal active site does not bind Ins(1,3,4,5)P₄ in a catalytically competent manner. Instead, the ligand is over 8 Å (long dashed line) from the inorganic phosphate bound by the P-loop. The short dashed lines represent distances between the phosphoryl group bound by the P-loop and the serine hydroxyl. Residues that interact with the ligands are derived from the P-loop (yellow), GA-loop (cyan), Phy domain, and penultimate helix. Oxygen is shown in red, nitrogen in blue, phosphorus in orange, and carbon in grey.

2.3.3 PTPLP specificity differences

The active sites of PhyAsr and the C-terminal repeat of PhyAmm have conserved residues that contact the ligands (Figure 2.3; Table 2.2) [169]. However, they contain significant differences in the main-chain conformation of three loops that contribute to the active sites of these enzymes: the extended loop prior to the penultimate helix (PhyAsr 287 to 305 and PhyAmm 583 to 600), the extended Ω-loop (PhyAsr 73 to 102 and PhyAmm 367 to

Table 2.2: Electrostatic and hydrogen bond distances in the PhyAsrC252S·Ins(1,3,4,5)-P₄, PhyAsrC252S·Ins(1,4,5)P₃ and PhyAmmC250S/C548S·Ins(1,3,4,5)P₄ structures are highly similar. Contact distances (<3.4 Å) between PhyA and the ligand phosphoryl (Phos) and hydroxyl (-OH) groups are shown. Bolded distances are main-chain interactions.

PhyAsrC252S						PhyAmmC250S/C548S			
Residue	Site	Ins(1,3,4,5)P ₄		Ins(1,4,5)P ₃		Residue	Site	Ins(1,3,4,5)P ₄	
		Phos/ -OH	Distance (Å)	Phos/ -OH	Distance (Å)			Phos/ -OH	Distance (Å)
R57	P _{at}	P3	2.80			R351	P _{at}	P3	3.16
									2.85
			3.17						2.64
D153	P _{at}					D449	P _{at}	P3	3.24
K189	P _{at} P _c	P3 P4	2.60	P4	2.33	K485	P _{at} P _c	P3 P4	2.73
			2.96						
									3.03
D223	P _{at}	P3	3.16			D519	P _{at}	P3	2.67
			2.81						
H224	P _a	P3	2.95			H520	P _a	P3	3.02
		P4	3.40						
		P5	2.51						
S252	P _s	P1	2.51	P1	2.27	S548	P _s	P1	2.67
E253	P _s	P1	3.06	P1	3.00	Q549	P _s	P1	3.22
A254	P _s	P1	3.17	P1	3.24	A550	P _s		
G255	P _s	P1	2.91	P1	2.79	G551	P _s	P1	3.15
V256	P _s	P1	2.69	P1	2.92	A552	P _s	P1	2.90
G257	P _a	O6	3.16	O6	3.08	G553	P _a	O6	3.34
								P1	3.31
R258	P _s	P1	2.92		2.72	R554	P _s	P1	2.73
			3.00	P1	2.90				3.11
			2.99		2.90				2.76
K305	P _c	P4	2.52	P4	2.41	K600	P _c	P4	2.90
			2.51						
	P _b	P5	2.82	P5	3.02		P _b	P5	3.15
									2.99
Y309	P _b	P5	2.82			Y604	P _b	P5	2.31
			2.47	P5	2.68				3.20
			3.18		3.13				3.38

398), and a β -hairpin turn of the Phy-specific domain (PhyAsr 186 to 189 and PhyAmm 482 to 485) (Figure 2.5). As the observed differences in these loop conformations may

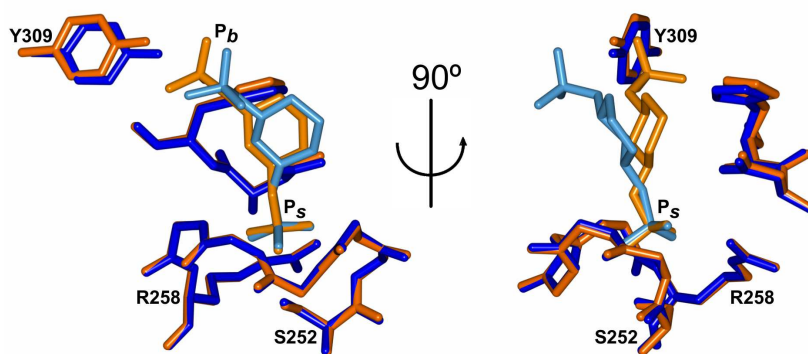


Figure 2.4: The superposition of PhyAsrC252S in complex with InsP₆ (blue; PDB 3MMJ) and Ins(1,3,4,5)P₄ (orange) demonstrates the 180° rotation and tilt of the *myo*-inositol ring towards the GA-loop of the less-phosphorylated IPs relative to InsP₆. The P-loop, GA-loop and Tyr-309 are shown. The *myo*-inositol ring and only two of the phosphoryl groups are displayed to simplify the diagram. Despite the rotation and tilt of the *myo*-inositol ring, the C1-phosphoryl group of Ins(1,3,4,5)P₄ remains bound by the P-loop, and the remaining phosphoryl groups are bound by equivalent residues. For example, the C5-phosphoryl groups in these structures form similar hydrogen bonds with Tyr-309 that originates from the opposite side of the residue.

affect substrate access to the active site, we determined the InsP₆ hydrolysis pathway for these enzymes. The time courses of the InsP₆ hydrolysis products separated by high performance ion chromatography (HPIC) clearly demonstrates that even though the enzymes are highly specific for the C3-phosphoryl group of InsP₆, they have different specificities for the Ins(1,2,4,5,6)P₅ substrate (Figure 2.6 A and B). Comparisons with standard chromatograms demonstrate PhyAmm is specific for the C4-phosphoryl group of Ins(1,2,4,5,6)P₅, whereas PhyAsr hydrolyzes the C1-phosphoryl (Appendix Figure A.1).

To further assess the contribution of these variable loops to the substrate specificity of PTPLPs, we examined the structure and hydrolysis pathway of PhyAbb. PhyAbb is one of the smallest known PTPLPs and has large deletions in the the extended loop prior to the penultimate helix, the extended Ω-loop and the Phy-specific domain (Figure 2.5) [160]. As a result, the active site of PhyAbb is more open and accessible than in PhyAsr or PhyAmm. Taken together, it is expected that PhyAbb has an altered pathway and would have a broader specificity for IPs than either PhyAsr or PhyAmm, which is confirmed by the hydrolysis

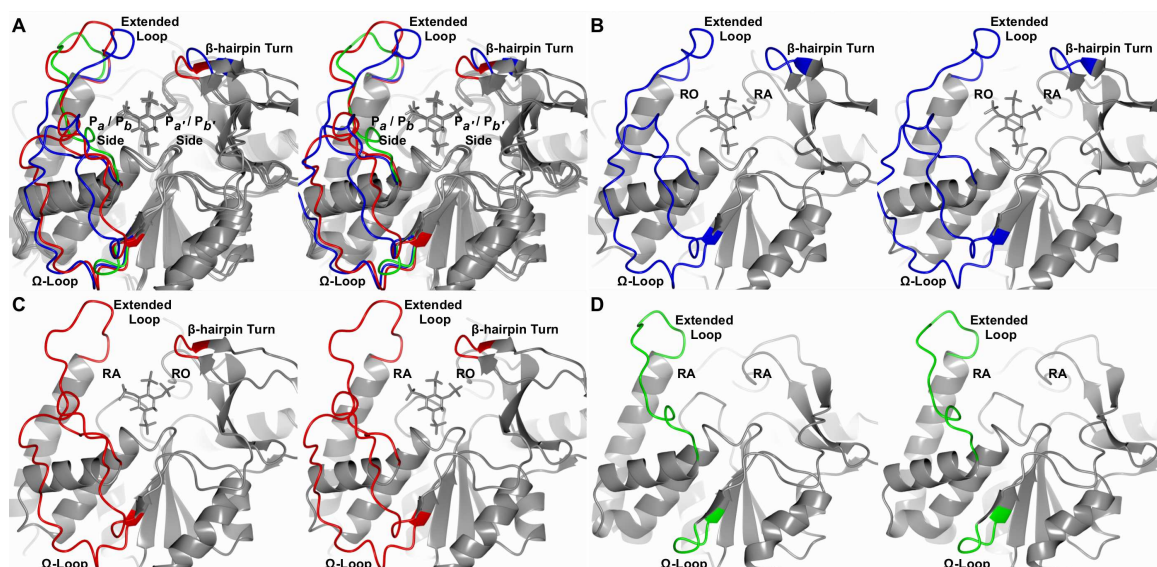


Figure 2.5: Variable loops implicated in the substrate specificity of PTPLPs. **A**, stereo view of the superposition of PhyAsrC252S·Ins(1,3,4,5)P₄ (blue), PhyAmmC250S/C548S·Ins(1,3,4,5)P₄ (red) and PhyAbb (green; PDB 4NX8) as a ribbon diagram with the ligand as sticks. The variable loops (colored segments) that influence substrate specificity include the extended loop prior to the penultimate helix, the extended Ω-loop, and the β-hairpin turn within the Phy-specific domain. **B**, the PhyAsr (blue) active site is relatively occluded (RO) on the P_a/P_b side and relatively accessible (RA) on the P_{at}/P_{bt} side. **C**, in the case of PhyAmm (red), the relatively occluded and relatively accessible sides are reversed. **D**, PhyAbb (green) has a 13 residue deletion, which removes the loop that contains the β-hairpin turn, resulting in an accessible (RA) active site on the P_{at}/P_{bt} side. Additionally, the position of the extended loop prior to the penultimate helix of PhyAbb is similar in position to the equivalent loop of PhyAmm, and the extended Ω-loop is deleted, leaving the P_a/P_b side more accessible (RA). As a result, PhyAbb produces four different InsP₄ products in contrast to PhyAsr and PhyAmm.

pathway (Figure 2.6 C). Though PhyAbb is also specific for the C3-phosphoryl group of InsP₆, making it a 3-phytase, PhyAbb produces four different InsP₄, which includes the PhyAsr (Ins(2,4,5,6)P₄) and PhyAmm (Ins(1,2,5,6)P₄) products. The much broader specificity of PhyAbb for IPs indicates that the loops that contribute to the active site influence substrate specificity.

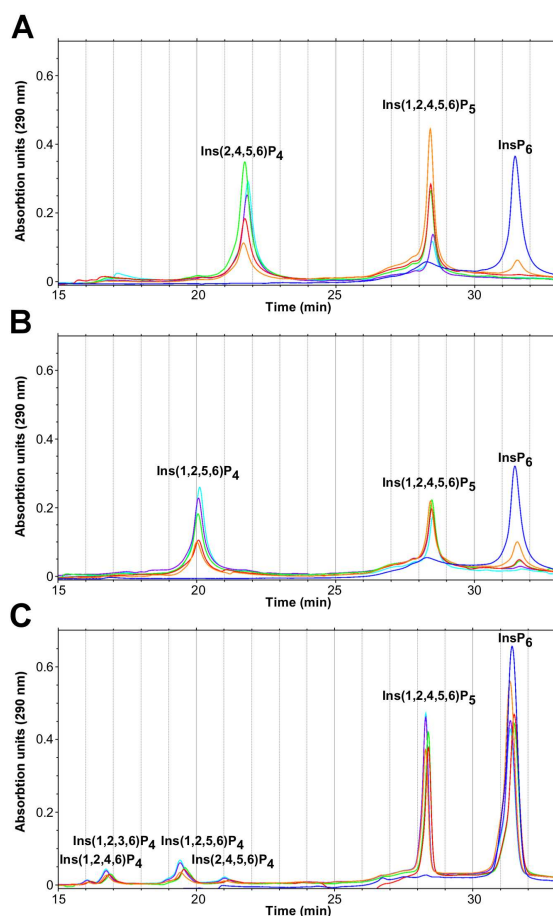


Figure 2.6: HPIC chromatograms of the PhyAsr *A*, PhyAmm *B* and PhyAbb *C* InsP₆ hydrolysis products demonstrating that the hydrolysis pathways diverge to produce alternative InsP₄ products. InsP₆ (5 mM) was incubated with 10 nM PhyAsr and PhyAmm at room temperature and 100 nM PhyAbb with 10 mM InsP₆. Samples were taken at 0 min (blue), 20 min (orange), 30 min (red), 40 min (green), 50 min (purple) and 60 min (cyan) and separated using a CarboPac PA-100 analytical column with a methanesulfonic acid gradient [179]. IPs were visualized using a post-column reactor with 0.1% (m/v) Fe(NO₃)₃ in a 2% (m/v) HClO₄ solution (0.2 mL/min).

2.4 Discussion

2.4.1 *myo*-inositol ring movements compensate for PTPLP active-site rigidity

There are no changes in the conformation of the active site upon binding of IP ligands to either PhyAsr or PhyAmm (Appendix Tables A.2 and A.4). The apparent rigidity of these enzymes suggest this may be a feature of the PTPLPs, and their specificity can be understood in simple structural terms. Further, the PhyAmmC250S/C548S C-terminal active site binds Ins(1,3,4,5)P₄ using the same phosphoryl-binding sites identified in PhyAsr (Fig-

ure 2.3 A; Table 2.2) [166]. In the PhyAsrC252S·InsP₆ structure the orientation of InsP₆ was such that in the P_{ax} site, only a hydroxyl or an axial phosphoryl group could be accommodated. Consequently, the P_{ax} site was identified as the primary structural determinant giving rise to the specificity of PhyAsr for the C3-phosphoryl group of InsP₆ [165, 166]. Assuming the ring orientation remains constant, the inability of the P_{ax} site to accommodate an equatorial phosphoryl group is sufficient to explain the known hydrolysis pathway of PhyAsr, but does not account for all minor hydrolysis products [165]. For example, PhyAsr produces small amounts of Ins(1,2,4,6)P₄ by removing the C5-phosphoryl group from Ins(1,2,4,5,6)P₅, which requires an equatorial phosphoryl group adjacent to the scissile phosphoryl group to bind within the P_{ax} site.

This work identifies a significant shift and a 180° rotation of the *myo*-inositol ring in the less-phosphorylated Ins(1,3,4,5)P₄ and Ins(1,4,5)P₃ complex structures compared with the PhyAsrC252S·InsP₆ structure (Figure 2.4). The ring-shift and rotation allow smaller substrates to utilize a different subset of phosphoryl binding sites. For example, in the PhyAsrC252S·InsP₆ structure and the less-phosphorylated IP complexes, hydrogen bonds to the C5-phosphoryl group originate from opposite sides of Tyr-309 (Figure 2.4). The *myo*-inositol ring shift also allows phosphoryl groups not adjacent to the scissile phosphoryl group to occupy the P_{ax} site. Importantly, simple modeling studies based on the observed ring-shift suggest that, provided the P_{bx} site contains a hydroxyl, the P_{ax} site can accommodate equatorial phosphoryl groups adjacent to the scissile phosphate. Alternatively, both the P_{ax} and P_{bx} sites may be able to accommodate equatorial phosphoryl groups if the ring is allowed to adopt ring orientations that are intermediate to those observed in the InsP₆ and less-phosphorylated IP complex structures. These observations are sufficient to rationalize the formation of minor products of the PhyAsr hydrolysis pathway and suggest PTPLPs may be able to hydrolyze a wide-range of less-phosphorylated IPs.

2.4.2 Structural determinants of the substrate specificity of PTPLPs

PhyAsr and PhyAmm hydrolyze Ins(1,2,4,5,6)P₅ to different InsP₄ products despite having the same specificity for the C3-phosphoryl group of the InsP₆ and binding Ins(1,3,4,5)-P₄ nearly identically (Figures 2.3 and 2.6) [165]. The residues that contact the ligands are conserved between PhyAsr and PhyAmm except for residue substitutions in the P-loop (Table 2.2). In the PhyAmm C-terminal repeat, Gln-549 immediately follows the catalytic Cys-548 (Ser-548 in our structure), whereas the equivalent residue in the PhyAsr is Glu-253. The Gln-549 side chain of PhyAmm is directed towards the active site and may directly or indirectly contact highly phosphorylated IPs providing an additional favourable electrostatic interaction in the P_{α'} site. In contrast, both in the presence and absence of ligand, the side chain of Glu-253 (PhyAsr) is directed away from the active site. The charge difference and the spatial orientation in the P_{α'} site are capable of influencing the difference in the activity and specificity of these enzymes but does not fully account for the pathway divergence.

The extended loop prior to the penultimate helix, extended Ω-loop, and β-hairpin in the Phy domain are loops around the active site that are different in PhyAsr and PhyAmm and contribute to the divergent pathways (Figure 2.5). The main-chain conformational differences in these connecting segments result from residue substitutions, insertions/deletions, and the distinct homodimers formed by each enzyme. The extended loop prior to the penultimate helix of PhyAsr located on the P_α/P_β side of the active site contain a single residue insertion, and the C-terminal end of the loop folds into the active site of the enzyme. In contrast, the equivalent loop in PhyAmm is pulled away from the active site and participates in a two-fold symmetric, homodimer interface. Also located on the P_α/P_β side of the active sites are the extended Ω-loops. These loops have similar conformations, and there is a structurally equivalent lysine (PhyAsr Lys-83 and PhyAmm Lys-379) pointing toward the active site. However, the extended Ω-loop of PhyAmm has a two-residue insertion that leaves the active site more accessible on the P_α/P_β side than PhyAsr. On the P_{α'}/P_{β'} side of

the active site is the β -hairpin turn, which differs both in type (PhyAsr Type II; PhyAmm Type I) and orientation between the two active sites. The conformational difference is primarily a result of homodimer formation even though the turn and associated β -strands have several residue substitutions. In particular, PhyAsr residues 191 to 193 are part of the two-fold symmetric homodimer interface that features an intermolecular antiparallel β -sheet, whereas the equivalent residues in PhyAmm interact with solvent. As a consequence of these differences, the β -hairpin and following residues are pulled away from the active site in PhyAsr to facilitate homodimer formation, whereas the equivalent residues of PhyAmm extend into the active site on the P_a/P_b side.

Taken together, the residue substitution and main-chain conformational differences indicate that the PhyAsr active site is relatively occluded on the P_a/P_b side of the active site and relatively accessible on the P_a/P_b' side (Figure 2.5). In the case of the PhyAmm active site, the converse is true. This is consistent with the specificity of PhyAsr for the C1-phosphoryl group of Ins(1,2,4,5,6)P₅, as it would place the C3-hydroxyl on the relatively occluded side of its active site. Likewise, the specificity of PhyAmm for the C4-phosphoryl group of Ins(1,2,4,5,6)P₅ would place the C3-hydroxyl on the opposite side of its active site, which corresponds to its occluded side.

The structure and pathway of PhyAbb support the occlusion theory as an explanation of the divergent pathways of PhyAsr and PhyAmm. PhyAbb lacks the β -hairpin in the Phy-specific domain on the P_a/P_b' side of the active site, leaving it more accessible than in either PhyAsr or PhyAmm (Figure 2.5). Further, the deletions of the extended Ω -loop, in the extended loop prior to the penultimate helix, and of two turns of the penultimate helix result in a more accessible P_a/P_b side of the active site. The result is a more open and accessible active site that allows Ins(1,2,4,5,6)P₅ to bind in multiple orientations and gives rise to four distinct InsP₄ ligands.

2.4.3 Ins(1,3,4,5)P₄ is not bound at the catalytic site of the PhyAmm N-terminal repeat

The PhyAmm N-terminal repeat shares 36% and 34% sequence identity with the C-terminal repeat and PhyAsr, respectively [169]. Not surprisingly, the main-chain conformation of the N-terminal active site shares a high degree of similarity with both PhyAsr and the C-terminal repeat of PhyAmm (0.71 and 0.28 Å, respectively, 128 atoms). Despite the closely similar main-chain conformations, in our structure only the C-terminal repeat binds Ins(1,3,4,5)P₄ in a catalytically competent manner. Differences between the N- and C-terminal repeat active sites have been discussed previously [169]. At present, there is no clear and unambiguous rationale for the lack of binding of Ins(1,3,4,5)P₄ to the N-terminal repeat P-loop. We note that Ins(1,3,4,5)P₄ is not a natural substrate for this enzyme, and the observed binding sterically prevents another Ins(1,3,4,5)P₄ from binding to the P-loop. Further, the natural InsP₄ substrates, which contain a C2-phosphoryl, cannot bind in the same manner because of steric clashes involving this axial phosphoryl group. This suggests that what we observed was a binding site that is preferred by non-native substrates. Alternatively, the structure was produced by soaking the substrate into a pre-formed crystal, which may prevent binding if there is an induced fit associated with IP binding to the N-terminal active site.

2.4.4 Biological implications

PhyAsr, PhyAmm and PhyAbb all function as 3-phytases yet have different specificities for Ins(1,2,4,5,6)P₅, thus producing alternative InsP₄ products. In the case of PhyAsr and PhyAmm, we were unable to predict their different specificities by identification of residues directly interacting with the substrates, as they are essentially identical. This suggests substrate specificity is influenced by structural features that do not directly interact with the bound ligand. We identified three variable loops unique to PTPLPs that alter substrate access to the PhyAsr and the PhyAmm C-terminal repeat active sites, which may explain their

observed specificities. This role of the variable loops is supported by our ability to predict and demonstrate that PhyAbb, an enzyme with large deletions in these loops, has a broad specificity towards IP substrates.

The activity of PTPLPs towards second messengers has been demonstrated, and this work confirms that two second messengers, Ins(1,3,4,5)P₄ and Ins(1,4,5)P₃, are capable of binding to the PTPLP active sites in a catalytically competent manner [159, 169]. Further, various IPases have been demonstrated as important for the function and survival of pathogenic bacteria in host systems [2, 32, 100]. They function by either providing the phosphate for growth or by derangement of the host phosphatidylinositol signalling pathway [32, 100]. This suggest PTPLP virulence factors disrupt phosphatidylinositide or inositol phosphate signaling pathways as opposed to phosphotyrosine-mediated pathways [2, 180, 181].

Chapter 3

The tandemly repeated protein tyrosine phosphatase-like *myo*-inositol phosphatase from *Mitsuokella multacida* in complex with Ins(1,2,4,5,6)P₅ displays divergent *myo*-inositol phosphate activity

3.1 Introduction

myo-inositol phosphates (IPs) are a derivative of *myo*-inositol that contain between one and eight phosphoryl groups with the most abundant being *myo*-inositol-1,2,3,4,5,6-hexakisphosphate (InsP₆ or phytate) [5, 23]. InsP₆ has diverse roles within eukaryotic cells and the importance is emphasized by the observation that deletions of enzymes involved in InsP₆ biosynthesis have lethal phenotypes in mouse embryos [18, 33, 34, 182]. The roles of *myo*-inositol pentakisphosphates (InsP₅) have more recently been investigated and shown to be involved in hormone signaling, affecting cell proliferation rates, modulating apoptotic responses, regulating viral assembly, chromatin remodeling, calcium channel activity, cytoskeleton reorganization, functioning as a hemoglobin effector, and functioning as a precursor to several metabolites [18, 45, 47, 49–57]. Enzymes specific for IPs are found within prokaryotic organisms despite the absence of IPs from most prokaryotic cells [8, 171]. These enzymes have characterized roles in phosphate scavenging, pathogenesis, and hydrolyzing dietary InsP₆ in the mammalian gut [2, 32, 99, 100, 183].

Protein tyrosine phosphatase-like *myo*-inositol phosphatases (PTPLPs) can remove one

or more phosphoryl groups from IPs, producing inorganic phosphate and Ins(2)P or another lesser phosphorylated IP [157, 158, 165, 170, 184]. The phosphoryl group in the 2-position is unique as it is axial while the other five are equatorial in the lowest energy chair conformation [37, 38]. Of the PTPLPs characterized to date, five remove the C3-phosphoryl group from InsP₆ (3-phytases) [158, 165, 170, 184, 185], one removes the C5-phosphoryl group (5-phytases) [157], and one the C1-phosphoryl group (1-phytase) [161].

PTPLPs are composed of a catalytic α - β - α sandwich PTP domain and an antiparallel α - β sandwich Phy-specific domain [162]. The catalytic PTP domain contributes the general acid loop (GA-loop) and phosphate-binding loop (P-loop) from the catalytic PTP domain, while residues from both the PTP and Phy-specific domains participate in substrate binding [166, 184]. Three loops adjacent to the active site have been shown to contribute to the specificity of PTPLPs: the Ω -loop, the β -hairpin (C-terminal repeat) and α -turn (N-terminal repeat) of the Phy-specific domain, and the extended loop prior to the penultimate helix [184]. These loops share limited sequence identity, vary in length, and adopt distinct conformations in structurally characterized PTPLPs [184]. The different loop conformations affect substrate access and available space within the active site. In general, the more open and accessible PTPLP active sites have a more relaxed substrate specificity towards less-phosphorylated substrates [184]. The PTPLP from *Mitsuokella multacida* (PhyAmm) is a homodimer and each monomer is a tandemly repeated PTPLP that contains two copies of the catalytic PTP and Phy-specific domains [169, 184]. The PhyAmm dimer interface is almost exclusively formed between the N-terminal repeats as an extensive network of hydrogen bonds, salt bridges, and van der Waals contacts [169, 184].

Here we present the first structure of a PTPLP (PhyAmm) in complex with Ins(1,2,4,5,6)-P₅. In this structure, an inorganic phosphate is bound to the N-terminal repeat and Ins(1,2,4,5,6)P₅ is bound to the C-terminal repeat of the inactive PhyAmmC250S/C548S mutant. This work presents the structural features that determine the conformation of the variable loops adjacent to the active site that influence the substrate specificity of PTPLPs. Further,

a novel IP specificity assay is presented which was used to identify the preferred substrates of the PhyAmm repeats, revealing that their specificities diverge and display a “divide and conquer” approach to InsP₆ hydrolysis.

3.2 Experimental procedures

3.2.1 Cloning and mutagenesis

The *phyA* gene of *M. multacida*, minus the putative signal peptide, was previously cloned into the *Nde*I site of the pET28b^{Kan} expression vector (EMD Biosciences) to add an N-terminal His₆ tag [169]. All mutant proteins (PhyAmmC250S/C548S, PhyAmmC250S, and PhyAmmC548S) were created by site-directed mutagenesis using counter-PCR amplification of the expression plasmid as described previously [186]. To confirm the presence of the desired mutations the nucleotide sequences were analyzed at the McGill University and Genome Quebec Innovation Centre.

3.2.2 Expression and purification

Protein expression was carried out in *Escherichia coli* T7 Express Competent cells (New England Biolabs; NEB) in ZYM-505 media supplemented with 50 µg/mL kanamycin. Cells were grown to an optical density (600 nm) of 3.0 at 310 K when protein expression was induced by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and grown overnight at 293 K. The induced cells were harvested and resuspended in lysis buffer (20 mM KH₂PO₄ pH 7.0, 300 mM NaCl, 5 mM β-mercaptoethanol (BME), 25 mM imidazole (pH 8.0), lysed by sonication, and cell debris was removed by centrifugation at 24 700 × g (45 min).

All proteins were purified to homogeneity by metal chelating affinity chromatography (Ni²⁺-NTA-agarose, Bio-Rad) as previously described [165, 169]. PhyAmmC250S/C548S was further purified for crystallization by cation exchange (Bio-Scale S Column, Bio-Rad) and size exclusion chromatography (S200, GE Healthcare) as described previously [165,

166, 169, 184]. The PhyAmm (C250S and C548S) cysteine to serine mutations (isosteric substitution) result in inactivation of both active sites.

3.2.3 IP production

The Ins(1,2,4,5,6)P₅ was produced enzymatically and purified to greater than 99% purity as shown in Figure 3.1. A 700 MHz Bruker Avance III NMR Spectrometer solution state nuclear magnetic resonance (NMR) spectrometer was used. After referencing the spectrometer with 85% phosphoric acid, the ³¹P (¹H decoupled) spectra for the 10 mM Ins(1,2,4,5,6)P₅ samples in the presence of 0.1 M NaOH were collected at 293 K.

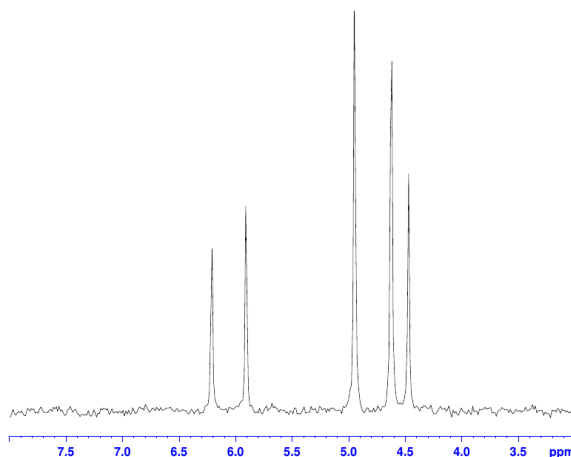


Figure 3.1: ³¹P (¹H decoupled) spectra of 5 mM purified Ins(1,2,4,5,6)P₅.

3.2.4 Crystallization

Crystallization experiments were conducted at room temperature using sitting-drop vapour diffusion with drop ratios of 2 μ L protein solution to 2 μ L reservoir solution. PhyAmm-C250S/C548S protein solutions were prepared at 4.5 mg/mL, and crystals were grown as described previously [184]. After 30 days rod-like crystals grew with approximate dimensions of 100 \times 100 \times 500 μ m. Crystals were soaked in mother liquor supplemented with 10 mM Ins(1,2,4,5,6)P₅ (produced enzymatically) for 15 minutes, and the crystals were flash-frozen in liquid nitrogen.

3.2.5 Data collection and image processing

Diffraction data ($\lambda = 0.97925 \text{ \AA}$) was collected from frozen crystals (100 K) using a Rayonix MX300 CCD detector at beamline 08ID-1 located at the Canadian Light Source (CLS; Saskatoon, SK, Canada). The space group and unit cell parameters are equivalent to those of the PhyAmmC250S/C548S·Ins(1,3,4,5)P₄ structure (PDB 4WU3). All diffraction image data were processed interactively with MOSFLM prior to scaling and merging within AIMLESS of the CCP4 program suite, version 6.3.0 [172–175]. Data collection statistics are shown in Table 3.1.

3.2.6 Structure refinement and model validation

Phases derived from the PhyAmmC250S/C548S·Ins(1,3,4,5)P₄ structure (PDB 4WU3) were used to solve the structures by molecular replacement. The refined structures have continuous electron density for main-chain atoms of amino acids 47 to 637, with the remaining residues located at the termini assumed to be disordered. This includes the N-terminal histidine tag and residues 31 to 46. Refinement was performed using PHENIX, version 1.11.1-2527, while interactive fitting of the models to the electron density were performed in COOT, version 0.8.8 [176, 187]. PROCHECK and structure validation tools with COOT were used throughout refinement to assess the stereochemistry of the model [177]. Unless indicated otherwise, figures were prepared with CCP4mg, version 2.10.8 [178]. Key data processing and refinement statistics are presented in Table 3.1.

3.2.7 Structure analysis

PhyAmmC250S/C548S and previously determined structures were compared by least squares (LSQ) superposition using LSQKAB from the CCP4 program suite [173]. The main-chain atoms of residues 47 to 342 and 343 to 636 for the N- and C-terminal repeats of PhyAmmC250S/C548S, respectively, were used in overall fold comparisons. Active-site comparisons were made using residues 350 to 351, 448 to 450, 484 to 486, 517 to 522, 545 to 558, 584 to 586, and 599 to 608 of the PhyAmmC250S/C548S C-terminal repeat. Dif-

Table 3.1: Data collection and refinement statistics for the PhyAmmC250S/C548S·Ins-(1,2,4,5,6)P₅ structure.¹

	PhyAmmC250S/C548S Ins(1,2,4,5,6)P ₅
Data collection	
Space group	P1
a, b, c (Å)	74.4, 86.5, 123.5
α , β , γ (°)	107.3, 91.6, 90.1
Wavelength (Å)	0.97925
Resolution (Å)	45.51 - 2.4 (2.486 - 2.4)
Observed reflections	205 848
Unique reflections	103 981
Completeness (%)	90.4 (83.6)
Redundancy	2.0
Rmerge (%)	11.2 (37.4)
I / σ I	2.1
Refinement Statistics	
No. reflections work set	103 953
No. reflections test set	4950
Rwork / Rfree (%)	20.7 / 25.2 (28.7 / 36.6)
Asymmetric unit	Dimer of Dimers
Protein atoms	19 172
Solvent atoms	681
Ligand atoms	168
Wilson B (Å ²)	26.1
Average B protein (Å ²)	35.1
Average B solvent (Å ²)	29.7
Average B ligand (Å ²)	46.5
RMSD Bonds (Å)	0.004
RMSD Angle (°)	0.900
Ramachandran distribution	
Preferred (%)	96.77
Allowed (%)	3.23
Outliers (%)	0

¹ values in parenthesis are for the highest resolution shell

ferences in the relative position of the *myo*-inositol ring of bound ligands of PhyAsrC252S were calculated using the superposed structures produced by secondary structure matching (SSM) within COOT, and GEOMCALC in the CCP4 program suite [173, 176, 177].

3.2.8 Identification of hydrolysis products

Hydrolysis of 5 mM InsP₆ (Sigma-Aldrich) was carried out at room temperature in the presence of PhyAmmC250S (10 nM) or PhyAmmC548S (500 nM) (50 mM sodium acetate pH 5.0, 200 mM NaCl, 1 mM BME, 0.1 mM EDTA). Reactions were quenched with an equal volume of 5% trichloroacetic acid (TCA) and subjected to high performance ion chromatography (HPIC; Waters 1525 Binary HPIC Pump; Milford, MA) utilizing a CarboPac PA-100 (4 × 240 mm) analytical column (Dionex; Sunnyvale, CA) [179] at room temperature using a post-column reactor flow rate of 0.2 mL/min. Identification of hydrolysis products utilized a standard hydrolysis chromatogram (appendix A experimental procedures).

3.2.9 IP Specificity Assay

The 11 hour acid hydrolysis sample was produced by following the methodology of Blaabjerg et al. (2010) [179]. Briefly, partial hydrolysis of InsP₆ was performed by boiling 1.5 g InsP₆ with 100 mL of 0.5 M HCl for 11 hours. The solution was evaporated to dryness and the residue was dissolved to 30 mg/mL in a 50 mM sodium acetate (pH 5.0), 200 mM NaCl, 1 mM BME, and 0.1 mM EDTA buffer. Peak assignment was performed by referencing the chromatograms to the peak identities determined by Blaabjer et al. (2010) [179, 184].

The activities of PhyAmm (100 nM), PhyAmmC250S (100 nM), and PhyAmmC548S (200 nM) were determined against the 11 hour acid hydrolysis mixture. The proteins were incubated with 30 mg/mL of the 11 hour acid hydrolysis, the reaction quenched with an equal volume of 5% TCA and analyzed using a CarboPac PA-100 analytical column with a methanesulfonic acid gradient [179, 184]. IPs were visualized using a post-column reactor with 0.1% (m/v) Fe(NO₃)₃ in a 2% (m/v) HClO₄ solution (0.2 mL/min).

3.3 Results

3.3.1 Ins(1,2,4,5,6)P₅ binding is consistent with the hydrolysis pathway

The PhyAmmC250S/C548S-Ins(1,2,4,5,6)P₅ structure presented here is isomorphous with the PhyAmmC250S/C548S-Ins(1,3,4,5)P₄ structure previously published (Figure 3.2; Table B.1) [184]. The C-terminal active site is virtually identical in the presence and absence of ligand, and between space groups, supporting the observation that the PTPLP family of enzymes have pre-formed active sites (Table B.2) [166, 184]. The electron density for Ins(1,2,4,5,6)P₅ and inorganic phosphate bound within the C-terminal and N-terminal active sites, respectively, is visible in initial $2mF_o - DF_c$ (1.5σ) and $mF_o - DF_c$ (3.5σ) maps, consistent with previous studies [169, 184]. Modeling Ins(1,2,4,5,6)P₅ as partial occupancy (0.70 occupancy) with inorganic phosphate (0.30 occupancy) in the C-terminal active-site reduces the difference density found within the active site and improves the overall quality of the density resulting in the refined $2mF_o - DF_c$ (1σ) electron density seen in Figure 3.3 [166]. Consistent with previous PTPLP structures in complex with ligand, the modeled Ins(1,2,4,5,6)P₅ adopts the lowest-energy chair conformation with five equatorial hydroxyl/phosphoryl groups and an axial C2-phosphoryl group.

Clear electron density for the axial C3-hydroxyl and each phosphoryl group identifies the Ins(1,2,4,5,6)P₅ ligand and its orientation within the active site of the C-terminal repeat. To facilitate the discussion of substrate conformation within the active site, the six phosphoryl-binding sites identified in the structure of PhyAsrC252S in complex with InsP₆ complex will be used and are referred to as the P_s, P_a, P_b, P_{a'}, and P_{b'} sites [166]. Consistent with the known InsP₆ hydrolysis pathway, the C4-phosphoryl group is located within the catalytic P_s site where it forms extensive contacts with P-loop (Figure 3.4; Table 3.2) [184]. The C3-hydroxyl of Ins(1,2,4,5,6)P₅ is located on the P_{a'}/P_{b'} side of the active site and makes a novel interaction with Gln-549 of the P-loop. This allows the axial C2-phosphoryl group to fill the P_{a'} site where it makes multiple contacts with Arg-351 and Lys-485 of the Phy-specific domain. The remaining C5, C6 and C1 phosphoryl groups occupy the P_a, P_b

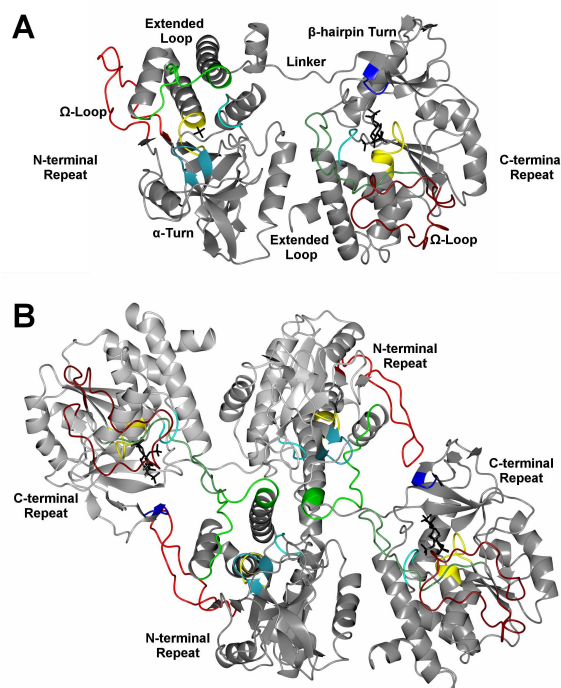


Figure 3.2: Ribbon diagram of the overall fold of PhyAmmC250S/C548S in complex with InsP₅. *A* PhyAmm oriented to view the tandem repeats and linker resulting in the InsP₅ ligand viewed from the top with the P-loop below the ligand. Within each catalytic PTP domain are the P-loop (yellow; residues 250 to 257 and 548 to 555) and GA-loop (cyan; residues 220 to 223 and 518 to 521). The Phy-specific domain (residue ranges 53 to 61 and 132 to 201, and 345 to 353 and 432 to 499) and the Phy-specific extension of the penultimate helix (green; residues 286 to 301 and 583 to 600) are unique to PTPLPs. *B* The PhyAmm dimer oriented to view the dimer interface formed primarily between the N-terminal repeats. The ligand atoms are shown as sticks in black. The polypeptide chains are shown in grey. The variable loops include the extended loops prior to the penultimate helix (green), the α -turn and β -hairpin turns (blue), and the extended Ω -loops (red). The P-loop (yellow) and the general-acid loop (GA-loop; cyan) contribute catalytic residues. Mutation of the cysteine nucleophile to the isosteric serine prevents thiolate formation and renders the enzyme inactive.

and P_c sites, respectively, while the P_{b'} site remains empty.

3.3.2 Structural features determining specificity of the C-terminal repeat

The structure of Ins(1,2,4,5,6)P₅ in complex with PhyAmmC250S/C548S is the first example of an InsP₆ hydrolysis product in complex with a PTPLP and directly tests the hypothesis that the variable loops adjacent to the active site influence the substrate specificity of PTPLPs (Figure 3.2). In the C-terminal repeat, substrate access to the P_{a'}/P_{b'}

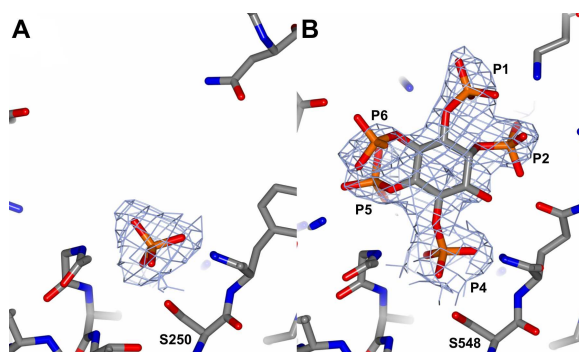


Figure 3.3: Clear electron density for *A* the inorganic phosphate, and *B* the phosphoryl groups and C3-hydroxyl of Ins(1,2,4,5,6)P₅ allow for an unambiguous fit of the ligands above the cysteine to serine mutations at position 250 (N-terminus) and 548 (C-terminus). The fit of Ins(1,2,4,5,6)P₅ places the C4-phosphoryl group (P4) in direct contact with the P-loop in a catalytically competent manner. The refined $2mF_o - DF_c$ electron density, contoured at 1σ (blue mesh) for PhyAmmC250S/C548S in complex with inorganic phosphate and InsP₅. Ligand and protein are shown as sticks with oxygen shown in red, nitrogen in blue, phosphorus in orange, and carbon in grey.

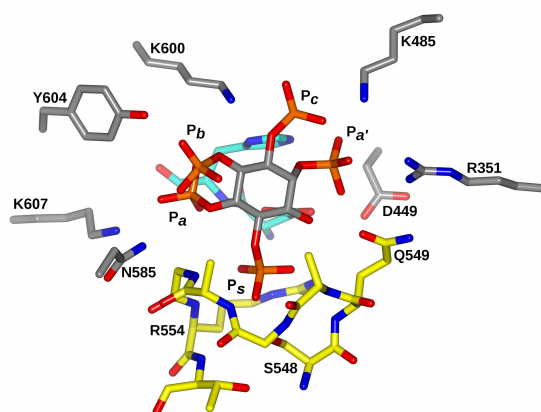


Figure 3.4: The C-terminal repeat of PhyAmmC250S/C548S binds Ins(1,2,4,5,6)P₅ using a subset of the contacts identified in the PhyAsrC252-InsP₆ (PDB 3MMJ) structure [166]. The phosphoryl-binding sites are labeled according to the InsP₆ structure (P_s , P_a , P_b , $P_{a'}$, and P_c). Residues that interact with the ligands are derived from the P-loop (yellow), GA-loop (cyan), Phy-specific domain, and penultimate helix. Oxygen is shown in red, nitrogen in blue, phosphorus in orange, and carbon in grey.

phosphoryl-binding sites is reduced as a result of the conformation of the β -hairpin loop of the Phy-specific domain that extends over the $P_{b'}$ site (Figure 3.5). In contrast, the P_a and P_b sites are relatively accessible as the extended loop prior to the penultimate helix is directed away from the active site. Consistent with the model, the Ins(1,2,4,5,6)P₅ substrate

Table 3.2: Electrostatic and hydrogen bond distances in the PhyAmmC250S/C548S·InsP₅ structure. Contact distances (<3.4 Å) between PhyA and the ligand phosphoryl (Phos) and hydroxyl (-OH) groups are shown. Bolded distances are main-chain interactions.

Residue	Site	Phos / -OH	Distance (Å)
Arg-351	P _{at}	P2	3.40 / 2.99
Lys-485	P _{at}	P2	3.10 / 3.13
	P _c	P1	2.64
Asp-519	P _{at}	P2	3.04
Ser-548	P _s	P4	2.19
Gln-549	P _s	P4	3.25
	P _{at}	P2	3.17
		O3	2.90
Gly-551	P _s	P4	3.27
Ala-552	P _s	P4	3.09
Gly-553	P _s	P4	3.29 / 3.19
	P _a	P5	3.37
Arg-554	P _s	P4	2.80 / 3.09 / 2.81
Asn-585	P _b	P6	3.31
Lys-600	P _c	P1	3.18
Lys-607	P _a	P5	3.22

is bound with the C2-phosphoryl group and C3-hydroxyl on the relatively occluded P_{at}/P_{bt} side of the active site while the P_a and P_b sites are filled by the C5- and C6-phosphoryl groups, respectively (Figure 3.4). The orientation of the Ins(1,2,4,5,6)P₅ *myo*-inositol ring and the C1-C2-O2-P2 torsion angle allows the C2-phosphoryl to fill the P_{at} site. This maximizes the number of favourable electrostatic contacts with the substrate as the P_s, P_a, P_b and P_{at} sites account for virtually all direct contacts in this structure (Table 3.2).

In addition to the variation of the loops adjacent to the active site, there are two residues that are unique to the C-terminal repeat that would likely affect substrate specificity: Gln-549 and Asn-585. In the case of Gln-549, the side chain is directed towards the active site where it contributes to the occlusion of the P_{at}/P_{bt} side of the active site and makes multiple interactions with the axial C2-phosphoryl group and the C3-hydroxyl (Table 3.2). Asn-585 contacts the C6-phosphoryl group bound in the P_b site, stabilizing the orientation of the bound Ins(1,2,4,5,6)P₅ substrate with a pair of equatorial phosphoryl groups in the P_a and

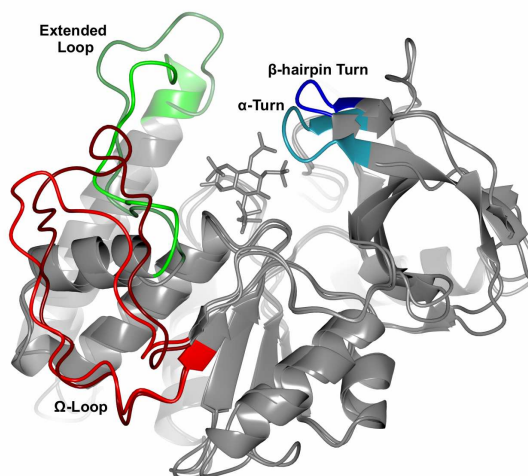


Figure 3.5: Superposition of the N- (light coloured loops) and C-terminal (dark coloured loops) repeats of the PhyAmmC250S/C548S-InsP₅ structure as a ribbon diagram with the ligand as sticks. The variable loops (colored segments) that influence substrate specificity include the extended loop prior to the penultimate helix (green), the extended Ω -loop (red), and the β -hairpin and α turns (blue) within the Phy-specific domain.

P_b sites and the C4-phosphoryl group positioned for hydrolysis in the P_s site. Consequently, the relatively small Asn-585 results in an increased accessibility on the P_a/P_b side of the active site while contributing a favourable interaction that selects for the observed orientation of the substrate. Together, these residues contribute to the more occluded $P_{a'}/P_{b'}$ side and more accessible P_a/P_b side of the active site, while providing favourable interactions that stabilize the observed orientation of the substrate with the C4-phosphoryl in the P_s site.

3.3.3 N-terminal active-site structure predicts different substrate specificity than the C-terminal active site

As demonstrated both in previous and current work, the N- and C-terminal repeats have distinct activities towards IPs [169, 184]. In accordance with these differences, the size and conformation of the variable loops adjacent to the active site of the N-terminal repeat diverge from those of the C-terminal repeat, predicting divergent substrate specificities (Figure 3.5). The Ω -loop of the N-terminal repeat is seven residues shorter than Ω -loop of the C-terminal repeat and does not directly affect the active-site structure. The size dif-

ference is smaller in the remaining two loops, with the N-terminal repeat α -turn of the Phy-specific domain containing two additional residues than the equivalent β -hairpin turn of the C-terminal repeat, while the extended loop prior to the penultimate helix contains two fewer residues in the N-terminal repeat than the C-terminal repeat. The additional residues of the α -turn extend over the $P_{b'}$ site and are shifted towards the P-loop (Figure 3.5). In this conformation the α -turn severely restricts access to the $P_{b'}$ site and would sterically prevent InsP₆ binding. Consequently, the α -turn conformation of the Phy-specific domain is likely responsible for the low activity of the N-terminal repeats towards InsP₆ and InsP₅ substrates [169]. Finally, conformation of the extended loop prior to the penultimate helix and the longer penultimate helix of the N-terminal repeat is stabilized by the dimer interface which prevents the loop from restricting access to the P_a/P_b side of the active site as is seen in the C-terminal repeat (Figure 3.5). The conformation of the variable loops result in a more occluded $P_{a'}/P_{b'}$ side and a more accessible P_a/P_b side of the active site.

In addition to variations in the loops adjacent to the active site and the penultimate helix, there are multiple point mutations within the active site that likely influence substrate binding (Table 3.3). Three of the point mutations cluster about the $P_{a'}/P_{b'}$ sites where they reduce the number of potential interactions from the $P_{a'}$ site and occlude the $P_{b'}$ site. In particular, Tyr-251 in the P-loop replaces the structurally equivalent Gln-549 of the C-terminal repeat, while Lys-59 and Asn-187 of the N-terminal repeat Phy-specific domain replace Arg-351 and Lys-485, respectively. The larger Tyr-251 side chain is directed away from the active site and forms a hydrogen bond with Asp-184 of the α -turn. As a result, Tyr-251 cannot replicate the multiple favourable interactions between Gln-549 and the substrate observed in the C-terminal repeat (Table 3.2). Likewise, the shorter Lys-59 side chain forms a hydrogen bond with Asp-61 and is unlikely to form the interactions seen in the C-terminal between Arg-351 and the substrate. As a result of these substitutions, the $P_{a'}$ site of the N-terminal repeat is expected to make half as many interactions with bound substrates than seen in the more active C-terminal repeat. The remaining mutation on the

$P_{a'}/P_{b'}$ side of the active site occurs in the α -turn of the Phy-specific domain. The different conformations of the α -turn and β -hairpin turn of the Phy-specific domain allow the amide side chain of the shorter Asn-187 in the N-terminal repeat to superpose with the Lys-485 amine of the C-terminal repeat. While this substitution preserves a likely $P_{a'}$ interaction, it facilitates a shift of the main chain towards the P-loop, which severely restricts access to the $P_{b'}$ site.

Table 3.3: Structurally equivalent residues of the PhyAmm N- and C-terminal active sites that interact with the IP substrates (or are expected to do so in the N-terminal active site). Bolded residues show differences in the active sites.

PhyAmm N-terminus	Phosphoryl Binding site	PhyAmm C-terminus
Lys-59	$P_{a'}$	Arg-351
Arg-70	$P_{b'}$	Arg-362
Asp-149	$P_{a'}$	Asp-449
Asn-187	$P_{a'}/P_c$	Lys-485
Asp-221	$P_{a'}$	Asp-519
His-222	P_a	His-520
Cys-250	P_s	Cys-548
Tyr-251	P_s	Gln-549
Ala-252	P_s	Ala-550
Gly-253	P_s	Gly-551
Met-254	P_s	Ala-552
Gly-255	P_a	Gly-553
Arg-256	P_s	Arg-554
Val-288	P_b	Asn-585
Gly-301	P_c/P_b	Lys-600
Tyr-305	P_b	Tyr-604
Arg-308	P_a	Lys-607

The remaining mutations are located on the P_a/P_b side of the active and involve Met-254, Val-288, Gly-301 and Arg-308 (Table 3.3). As a result of these mutations, it is likely that phosphoryl groups bound in the P_a and P_b sites of the N-terminal repeat are shifted towards to GA-loop resulting in the substrate binding deeper in the active site. Gly-301 of the extended loop replaces Lys-600 of the C-terminal repeat and removes any potential interaction from this position to the P_b site. Together with the differences in the extended loop

conformation, this substitution creates a large space adjacent to the GA-loop (Figure 3.6). At the same time, the large Met-254 side chain replaces Ala-552 of the P-loop and in the observed conformation it is likely to make close contacts with phosphoryl groups bound to the P_b site. While the conformation of the Met-254 side chain may change in the presence of substrate, the Gly-301 substitution suggests that phosphoryl groups bound to the P_b site may shift towards the open space adjacent to the GA-loop. The N-terminal repeat Arg-308 replacement of Lys-607 is also likely to create too close contacts with phosphoryl groups bound to the P_a site. Since the conformation of the Arg-308 side chain is fixed by multiple packing interactions, phosphoryl groups bound to the P_a site are likely to shift towards the GA-loop. Finally, Val-288 cannot replace the Asn-585 contribution to the P_a site. Taken together, the mutations on the P_a/P_b side of the active site significantly alter the shape of the active site. Close contacts involving the P_a and P_b sites with the additional open space adjacent to the GA-loop suggest the P_a and P_b sites are shifted towards the GA-loop of the N-terminal repeat.

3.3.4 The PhyAmm N- and C-terminal repeats have divergent IP activity

The N- and C-terminal repeats of PhyAmm have different activities towards InsP₆ and InsP₅ substrates. Specifically, the N-terminal repeat has less than 2% of the activity of the C-terminal repeat towards InsP₆ and InsP₅ [169]. Given the observed difference in the active site structure of the N- and C-terminal repeats, the InsP₆ hydrolysis pathway of both domains were determined (Figure 3.7). The active C-terminal repeat (PhyAmmC250S) produces a major (>90%) and minor InsP₅ product that have been identified as DL-Ins(1,2,4,5,6)P₅ and DL-Ins(1,2,3,5,6)P₅, respectively (Figure 3.7 A). Subsequent reactions convert the major DL-Ins(1,2,4,5,6)P₅ product to DL-Ins(1,2,5,6)P₄ and DL-Ins(1,2,6)P₃ along the major hydrolysis pathway and DL-Ins(1,2,3,5,6)P₅ to DL-Ins(1,2,3,6)P₄ along the minor pathway. These results are nearly identical to those previously generated for the fully active enzyme, less the minor pathway products [184]. The active N-terminal re-

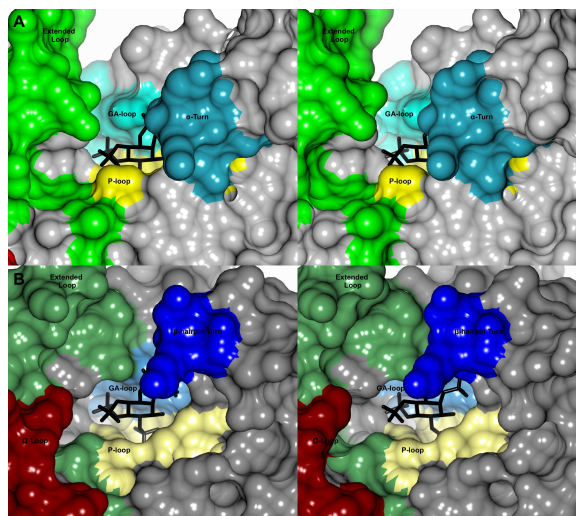


Figure 3.6: Stereo view of the A N- and B C-terminal repeats of the PhyAmmC250S/-C548S·InsP₅ structure as a space-filling diagram with the ligand as sticks. The view emphasizes the space at the back of the active site adjacent to the GA-loop which is either filled (C-terminal repeat) or open (N-terminal repeat). The variable loops (colored segments) that influence substrate specificity include the extended loop prior to the penultimate helix (green), the extended Ω -loop (red), and the β -hairpin and α turns (blue) within the Phy-specific domain. The ligand has been superposed onto the N-terminal repeat for comparison.

peat (PhyAmmC548S) produces the same InsP₅, InsP₄ and InsP₃ peaks, though the precise enantiomers have yet to be determined (Figure 3.7 B).

The accumulation and subsequent hydrolysis of InsP₅ through InsP₂ differs for the individual repeats indicating divergent activities (Figure 3.7). In the case of the C-terminal repeat, each of the major pathway hydrolysis products accumulate to relatively high levels prior to the formation of the subsequent IP product. The sequential accumulation of C-terminal repeat products as a function of time suggest the rates of catalysis of the individual steps are comparable or decrease for each lesser-phosphorylated product. The InsP₆ hydrolysis pathway of the N-terminal repeat has a similar pattern of product accumulation and subsequent hydrolysis for InsP₅ and InsP₄. This pattern diverges for InsP₃ and InsP₂, as the InsP₂ peak accumulates more rapidly. This suggests the efficiency of InsP₃ hydrolysis is significantly larger than that of the other IPs.

To further understand the substrate specificity of the individual repeats of PhyAmm,

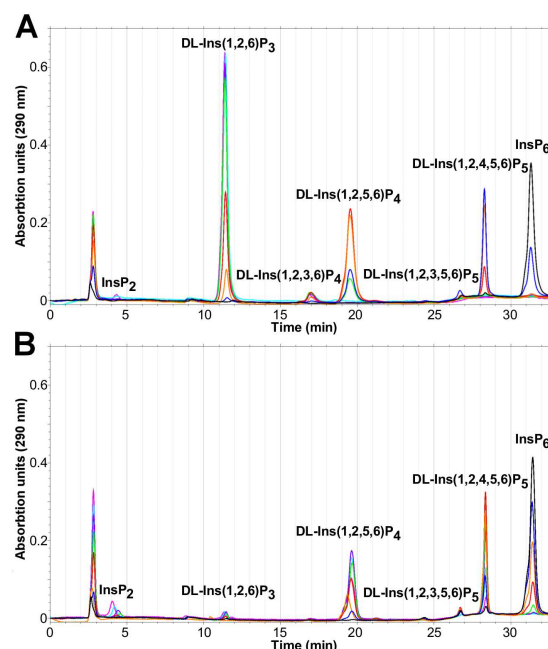


Figure 3.7: HPIC chromatograms of the PhyAmmC250S (A) and PhyAmmC548S (B) InsP_6 hydrolysis products. InsP_6 (5 mM) was incubated with 10 nM PhyAmmC250S and 500 nM PhyAmmC548S at room temperature. Samples were taken at 0 min (black), 10 min (blue), 20 min (orange), 30 min (red), 40 min (green), 50 min (purple), and 60 min (cyan), and separated using a CarboPac PA-100 analytical column with a methanesulfonic acid gradient [20, 179]. IPs were visualized using a post-column reactor with 0.1% (m/v) $\text{Fe}(\text{NO}_3)_3$ in a 2% (m/v) HClO_4 solution (0.2 mL/min).

each repeat was incubated with an IP mixture generated by incomplete acid hydrolysis and followed as a function of time (Figure 3.8; Table 3.4). While the analysis of hydrolysis data for complex mixtures is more challenging, chromatographic peaks that increase in area are products, those that decrease are substrates, and those that remain constant are either unaffected by the enzyme or are produced and consumed at equal rates. The time course of hydrolysis for the fully active PhyAmm (Figure 3.8 A) suggests the vast majority of the possible IPs can be hydrolyzed to a greater or lesser extent and clearly demonstrates the enzyme has activity towards substrates that are not present in the InsP_6 hydrolysis pathway. As summarized in Table 3.4, InsP_6 and the InsP_5 substrates are rapidly hydrolyzed and their concentrations decrease at each time point until they have been completely hydrolyzed. The InsP_4 and InsP_3 substrates typically accumulate during initial time points and are subsequently hydrolyzed, while the InsP_2 (peaks 1-5) and several InsP_3 substrates

accumulate at all measured time points. The overall trend is consistent with an enzyme that has a higher activity towards highly phosphorylated substrates and a lower activity towards less phosphorylated substrates. While there is no evidence the InsP_2 and several InsP_3 substrates are hydrolyzed under these conditions, prolonged incubation will likely result in additional hydrolysis of these substrates as has been demonstrated in previous work [169].

Table 3.4: Summary of the activity of wild-type PhyAmm, PhyAmmC250S and PhyAmmC548S against the 11 hour acid hydrolysis mixture. The wild-type PhyAmm and PhyAmmC250S at 100 nM, and PhyAmmC548S at 200 nM were incubated with 20 mg/mL of the 11 hour acid hydrolysis for up to one hour (Figure 3.8). \uparrow and \downarrow indicates an increase and decrease in peak size, respectively, \emptyset indicates no detectable activity, and C indicates complete peak disappearance in 60 minutes. The peaks were identified according to Blaabjerg et al. (2010) [179, 184].

Chromatogram Peak Number ¹	<i>myo</i> -inositol phosphate ²	PhyAmm wild-type		PhyAmmC250S		PhyAmmC548S	
0	PO_4	\uparrow		\uparrow		\uparrow	
1 - 4	InsP_2	\uparrow		\uparrow		$\uparrow\downarrow$	
5	InsP_2	\uparrow		\uparrow		$\uparrow\downarrow$	
6 - 11	InsP_3	\uparrow		\uparrow		\downarrow	
	InsP_3	$\downarrow\uparrow$		\uparrow		$\downarrow\uparrow$	
	InsP_3	$\uparrow\downarrow$		$\uparrow\downarrow$		\downarrow	
12	DL- $\text{Ins}(1,5,6)\text{P}_3$	\downarrow		\downarrow		\downarrow	
13	DL- $\text{Ins}(4,5,6)\text{P}_3$	$\uparrow\downarrow$		\uparrow		\downarrow	C
14 - 15	$\text{Ins}(1,2,3,5)\text{P}_4$	\downarrow		\downarrow	C	\downarrow	
	DL- $\text{Ins}(1,2,4,6)\text{P}_4$						
16 - 17	DL- $\text{Ins}(1,2,3,6)\text{P}_4$	$\uparrow\downarrow$		$\uparrow\downarrow$		\downarrow	
	$\text{Ins}(1,3,4,6)\text{P}_4$						
18	DL- $\text{Ins}(1,2,4,5)\text{P}_4$	\downarrow	C	\downarrow	C	\emptyset	
19	DL- $\text{Ins}(1,3,4,5)\text{P}_4$	\downarrow	C	\downarrow	C	\emptyset	
20	DL- $\text{Ins}(1,2,5,6)\text{P}_4$	$\uparrow\downarrow$		$\uparrow\downarrow$	C	$\downarrow\uparrow$	
21	$\text{Ins}(2,4,5,6)\text{P}_4$	$\uparrow\downarrow$	C	$\uparrow\downarrow$	C	\uparrow	
22	DL- $\text{Ins}(1,4,5,6)\text{P}_4$	$\uparrow\downarrow$		$\uparrow\downarrow$		\downarrow	
23	$\text{Ins}(1,2,3,4,6)\text{P}_5$	\downarrow	C	\downarrow	C	\downarrow	
24	DL- $\text{Ins}(1,2,3,4,5)\text{P}_5$	\downarrow	C	\downarrow	C	$\uparrow\downarrow$	
25	DL- $\text{Ins}(1,2,4,5,6)\text{P}_5$	\downarrow	C	\downarrow	C	\uparrow	
26	$\text{Ins}(1,3,4,5,6)\text{P}_5$	\downarrow	C	\downarrow	C	\emptyset	
27	InsP_6	\downarrow	C	\downarrow	C	\downarrow	

¹identified in Figure 3.8

²as identified by Blaabjerg et al. (2010) [179]

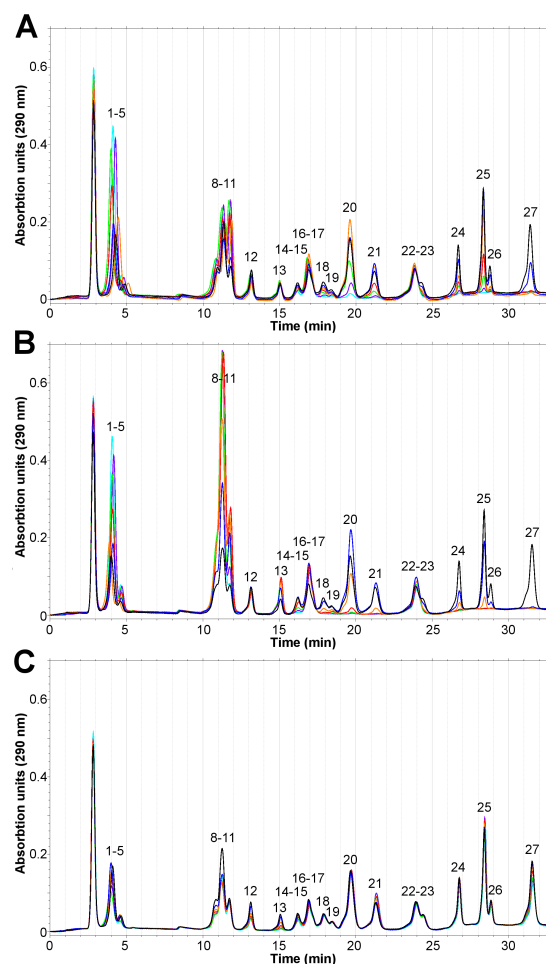


Figure 3.8: HPIC chromatograms of the activities of wild-type PhyAmm (A), PhyAmmC250S (B) and PhyAmmC548S (C) against the 11 hour acid hydrolysis mixture. The wild-type PhyAmm and PhyAmmC250S at 100 nM, and PhyAmmC548S at 200 nM were incubated with 20 mg/mL of the 11 hour acid hydrolysis for 10 min (blue), 20 min (orange), 30 min (red), 40 min (green), 50 min (purple), and 60 min (cyan). Production of the 11 hour acid hydrolysis sample (black) as performed previously [179, 184]. The peaks are identified by Blaabjerg et al. (2010) [179] as: (1-5) InsP_2 , (6-11) InsP_3 , (12) DL- $\text{Ins}(1,5,6)\text{P}_3$, (13) DL- $\text{Ins}(4,5,6)\text{P}_3$, (14) $\text{Ins}(1,2,3,5)\text{P}_4$, (15) DL- $\text{Ins}(1,2,4,6)\text{P}_4$, (16) DL- $\text{Ins}(1,2,3,6)\text{P}_4$, (17) $\text{Ins}(1,3,4,6)\text{P}_4$, (18) DL- $\text{Ins}(1,2,4,5)\text{P}_4$, (19) DL- $\text{Ins}(1,3,4,5)\text{P}_4$, (20) DL- $\text{Ins}(1,2,5,6)\text{P}_4$, (21) $\text{Ins}(2,4,5,6)\text{P}_4$, (22) DL- $\text{Ins}(1,4,5,6)\text{P}_4$, (23) $\text{Ins}(1,2,3,4,6)\text{P}_5$, (24) DL- $\text{Ins}(1,2,3,4,5)\text{P}_5$, (25) DL- $\text{Ins}(1,2,4,5,6)\text{P}_5$, (26) $\text{Ins}(1,3,4,5,6)\text{P}_5$, (27) InsP_6 . The IPs were separated using a CarboPac PA-100 analytical column with a methanesulfonic acid gradient [179, 184]. IPs were visualized using a post-column reactor with 0.1% (m/v) $\text{Fe}(\text{NO}_3)_3$ in a 2% (m/v) HClO_4 solution (0.2 mL/min).

In the case of the active C-terminal repeat (PhyAmmC250S; Figure 3.8 B), the time course of hydrolysis is similar to that observed for the fully active PhyAmm, in agree-

ment with the C-terminal repeat having higher activity when compared with the N-terminal repeat. The most notable differences include a greater accumulation of InsP_3 and InsP_2 products that are not subsequently hydrolyzed under these conditions. Of particular interest, peaks that differ between the fully active PhyAmm and the active C-terminal repeat include those peaks that are efficiently hydrolyzed by the N-terminal repeat. This strongly suggests the activity of the fully active PhyAmm is approximately the sum of the activity of the individual repeats and likely represents a mechanism to maximize the removal of phosphoryl groups in a mixed IP environment.

The active N-terminal repeat (PhyAmmC548S) has limited or no activity towards InsP_6 , InsP_5 and InsP_4 substrates, and exhibits its greatest activity towards the InsP_3 and InsP_2 substrates (Figure 3.8 C). The InsP_3 peaks decrease throughout the time course with the exception of a single InsP_3 peak representing $\text{Ins}(1,2,3)\text{P}_3$, $\text{DL-Ins}(1,2,6)\text{P}_3$ and $\text{DL-Ins}(1,4,6)\text{P}_3$ (middle peak of peak 8-11), which undergoes a rapid initial decrease and followed by a much smaller accumulation late in the time course. The same InsP_3 peak is produced in the active N-terminal repeat InsP_6 hydrolysis pathway where it is rapidly converted to an InsP_2 product. The small accumulation of this peak late in the time course of the IP mixture likely arises from the limited activity of the N-terminal repeat towards InsP_4 peaks that are not generated as part of the InsP_6 hydrolysis pathway. This suggests at least one of the InsP_3 components of this peak are not efficiently hydrolyzed.

Finally, while there is evidence the C-terminal repeat has a higher activity towards InsP_6 and InsP_5 substrates than the fully active enzyme, this may simply reflect batch-to-batch variation in the effective concentration of the enzyme preparations. At this time, we cannot rule out alternative explanations, including differential sensitivity to oxidants, differences in potential inhibitory mechanisms and/or cooperativity between active sites.

3.4 Discussion

3.4.1 *myo*-inositol ring position allows for optimal use of the PhyAmm C-terminal repeat phosphoryl-binding sites

The PhyAmmC250S/C548S·Ins(1,3,4,5)P₄ structure presented here is isomorphous with the PhyAmmC250S/C548S·Ins(1,3,4,5)P₄ structure published previously, and a comparison of the substrates bound in the two structures reveals changes in the relative orientation of the *myo*-inositol ring (Figure 3.9; Table 3.5) [184]. The Ins(1,3,4,5)P₄ substrate binds with the opposite face of the *myo*-inositol ring towards the GA-loop when compared to the Ins(1,2,4,5,6)P₅ binding observed in this work [184]. This directs the axial C2-hydroxyl away from the GA-loop and allows the equatorial C3-phosphoryl group to fill the P_{a'} site. In the Ins(1,2,4,5,6)P₅ structure, the *myo*-inositol ring is rotated away from the GA-loop, allowing the axial C2-phosphoryl to fill the P_{a'} site while the C4-phosphoryl group occupies the P_s site (Figure 3.9; Table 3.5). As a result, Ins(1,2,4,5,6)P₅ binds to the P_s, P_a, P_b, P_{a'} and P_c sites of PhyAmmC250S/C548S with an orientation that is more similar to that observed in the PhyAsrC252S·InsP₆ complex structure (Figure 3.9; Table 3.5).

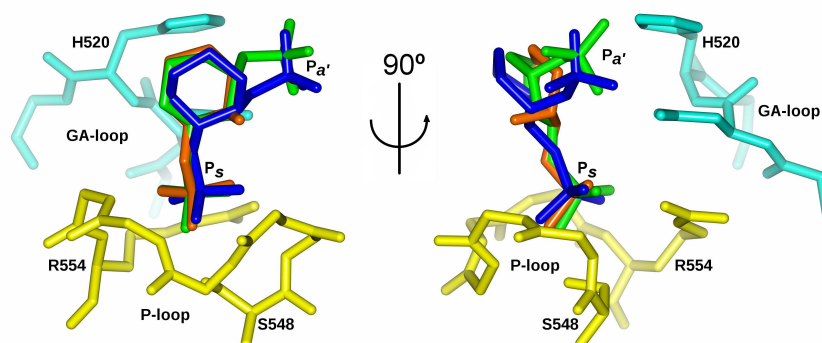


Figure 3.9: Modeling of PhyAmmC250S/C548S·Ins(1,2,4,5,6)P₅ (green) with the substrates from the PhyAmmC250S/C548S·Ins(1,3,4,5)P₄ (orange; PDB 4VU3) and PhyAsrC252S·InsP₆ (blue; PDB 3MMJ) structures demonstrates the intermediate position of Ins(1,2,4,5,6)P₅ relative to InsP₆ and Ins(1,3,4,5)P₄. The P-loop (yellow) and GA-loop (cyan) are shown. The *myo*-inositol ring and only two of the phosphoryl/hydroxyl groups are displayed to simplify the diagram.

In the PhyAsrC252S·InsP₆ structure, the C2-phosphoryl group occupies the same P_{a'}

Table 3.5: IP ring shifts relative to InsP₅ in LSQ superposed structures (active site). Distances (Å) between the C1 of InsP₅ and the structurally equivalent carbons, the angles between the plane of the 6-carbon rings, and the distances between the centres of mass of the 6-carbon rings. The 6-carbon ring angle is the rotation angle between the planes of the rings.

	PhyAmmC250S/C548S Ins(1,3,4,5)P ₄ ¹	PhyAsrC252S InsP ₆ ²	PhyAsrC252S Ins(1,3,4,5)P ₄ ³	PhyAsrC252S Ins(1,4,5)P ₃ ⁴
C6 distance	1.00 Å	0.41 Å	1.31 Å	1.33 Å
6-carbon ring angle	18.5°	-10.4°	21.0°	28.4°
6-carbon ring distance	0.60 Å	0.60 Å	0.80 Å	0.85 Å

¹PhyAmmC250S/C548S·Ins(1,3,4,5)P₄ structure: PDB 4WU3

²PhyAsrC252S·InsP₆ structure: PDB 3MMJ

³PhyAsrC252S·Ins(1,3,4,5)P₄ structure: PDB 4WTY

⁴PhyAsrC252S·Ins(1,4,5)P₃ structure: PDB 4WU2

site, despite the C3-phosphoryl group location in the P_s site [166]. In order for the C2-phosphoryl group to fill the P_{a'} site when either the C3- or C4-phosphoryl group of InsP₆ or Ins(1,2,4,5,6)P₅, respectively, is in the P_s site, there are compensatory changes in the C1-C2-O2-P2 torsion angles ($\sim 60^\circ$ InsP₆ and $\sim 140^\circ$ Ins(1,2,4,5,6)P₅) that complement the change in relative orientation of the *myo*-inositol ring within the active site (Figure 3.9). In particular, the Ins(1,2,4,5,6)P₅ *myo*-inositol ring is tilted and twisted towards the GA-loop of PhyAmmC250S/C548S when compared to the PhyAsrC252S·InsP₆ structure and the Ins(1,2,4,5,6)P₅ orientation relative to the GA-loop is intermediate to that in the InsP₆ and Ins(1,3,4,5)P₄ structures (Table 3.5). This maximizes the number of favourable electrostatic contacts with substrate as the P_s, P_a, P_b and P_{a'} site account for virtually all direct contact in the PhyAmmC250S/C548S·Ins(1,2,4,5,6)P₅ and PhyAsrC252S·InsP₆ complex structures. Previous work has demonstrated that the PhyAsr active site is preformed and changes in the *myo*-inositol ring orientation allow different IP substrates to optimize their use of the existing phosphoryl-binding sites [166, 184]. This work demonstrates the C-terminal repeat of PhyAmm uses a similar mechanism to bind various IP substrates and suggests the mechanism is common within the PTPLP family of enzymes.

3.4.2 Residues stabilizing Ins(1,2,4,5,6)P₅ binding to the C-terminal repeat diverge from PhyAsr

The InsP₆ hydrolysis pathway of the C-terminal repeat of PhyAmm and PhyAsr diverge following the removal of the C3-phosphoryl group of InsP₆, with PhyAmm removing the C4-phosphoryl group and PhyAsr removing the C1-phosphoryl group of Ins(1,2,4,5,6)P₅ [165, 184]. The position of the C4-phosphoryl group in the P_s site of the C-terminal repeat of PhyAmm is consistent with the hydrolysis pathway that produces Ins(1,2,5,6)P₄ from the hydrolysis of Ins(1,2,4,5,6)P₅ (Figures 3.4 and 3.7) [184]. Residue substitutions in the two active sites that alter substrate contacts and likely influence substrate specificity include Gln-549 and Asn-585 of PhyAmm, with Glu-253 and Phe-289 being the equivalent PhyAsr residues (Figure 3.10). In PhyAmm, Gln-549 and Asn-585 provide favourable interactions that stabilize the orientation and conformation of Ins(1,2,4,5,6)P₅ within the active site and facilitate the binding of the C4-phosphoryl in the P_s site. Glu-253 of PhyAsr is directed away from the active site in both the presence and absence of bound ligands [162, 166, 184]. While this prevents a steric clash with bound InsP₆, the same conformation is observed the Ins(1,3,4,5)P₄ complex structure where steric interaction should not be problematic. Given the PhyAsr structures have been determined at the pH optima of the enzymes, it is likely the Glu-253 carboxylate is anionic and the observed conformation removes an unfavourable interaction with bound substrate. In the case of Asn-585, the structural equivalent Phe-289 of PhyAsr is incapable of providing a favourable electrostatic contact while decreasing the accessibility of the P_a and P_b sites. As a result, Gln-549 and Asn-585 of PhyAmm provide favourable interactions that stabilize the orientation and conformation of the substrate with the C4-phosphoryl group in the P_s site while the equivalent residues in PhyAsr provide none of the favourable interactions and potentially introduce unfavourable interactions.

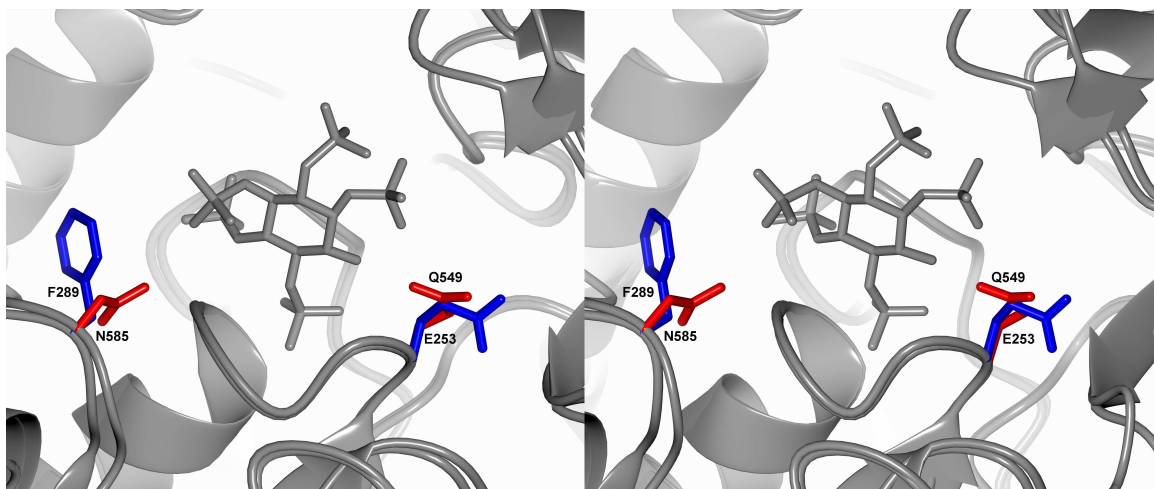


Figure 3.10: Superposition of the C-terminal repeat of the PhyAmmC250S/C548S-InsP₅ structure and PhyAsr (PDB 4WTY) as a ribbon diagram with the ligand as sticks. Gln-548 and Asn-585 (red sticks) of the C-terminal repeat are visible and the structurally equivalent Glu-253 and Phe-289 (blue sticks) of PhyAsr. Gln-549 makes additional interactions not possible by Glu-253 due to side-chain conformation and also provides interactions with the C3-hydroxyl. Asn-585 makes an additional contact with the C6-phosphoryl group bound in the P_b site that the Phe-289 of PhyAsr is incapable of providing. These interactions stabilize the conformation that results in positioning the C4-phosphoryl group of Ins(1,2,4,5,6)P₅ for hydrolysis.

3.4.3 Factors influencing the conformations of the variable loops

The Ω -loop, the β -hairpin (C-terminal repeat) and α -turn (N-terminal repeat) of the Phy-specific domain, and the extended loop prior to the penultimate helix are all adjacent to the active sites of PTPLPs and have been demonstrated to contribute to their specificity as they vary in sequence identity, length and conformation [184]. These differences coupled with the observation that the active-site residues that directly contact the IP ligands are highly conserved, led to the hypothesis that the above variable loops are responsible for specificity differences among PTPLPs. These loops are important in determining the size, shape, and accessibility of the P_a/P_b and P_{a'}/P_{b'} sides of the active sites, where the more open and accessible the active site is the more relaxed substrate specificity is displayed by the protein [184].

The quaternary structure of PhyAmm and PhyAsr directly influence the observed conformation of the extended loop (P_a/P_b side) and the β -hairpin turn of the Phy-specific do-

main ($P_a/P_{b'}$). In PhyAmm, the dimer interface includes an intermolecular contact between the extended loop of the C-terminal repeat and the same loop of the N-terminal repeat (Figure 3.11 A). The two loops pack against one another and are indirectly stabilized by a hydrogen bond, between Asp-295 and Arg-448, and a stacking interaction between Arg-448 and Trp-599. These intermolecular interactions, pull the extended loop away from the active site of the C-terminal repeat and increase accessibility to the P_a/P_b side of the active site. The dimer interface in the PhyAsr homodimer is distinct from that observed in PhyAmm and involves the β -hairpin turn of the Phy-specific domain (Figure 3.11 B). In particular, Asp-192 and Gly-193 of chain A and the identical residues of chain B are related by an approximate 2-fold rotation that facilitates the formation of several intermolecular main-chain hydrogen bonds. This interaction fixes the conformation of the β -hairpin and allows the Asp-192 side chain to form a hydrogen bond with the $i+1$ main chain (Lys-187) of the β -hairpin turn. These interactions pull the β -hairpin turn away from the active site and increase the accessibility on the $P_a/P_{b'}$ side of the PhyAsr active site.

In the case of the PhyAmm C-terminal repeat, the P_a/P_b side of the active site is relatively accessible due to the conformation of the extended loop leaving more space as described, while the $P_a'/P_{b'}$ side of the active site is relatively occluded due to the conformation of the β -hairpin turn extending over the $P_{b'}$ site (Figure 3.5). In the PhyAmmC250S/-C548S·InsP₅ complex structure, the C3-hydroxyl is bound on the relatively occluded $P_a'/P_{b'}$ side of the active site which is consistent with the proposed model. In the case of PhyAsr, the extended loop prior to the penultimate helix extends over P_b while the β -hairpin turn of the Phy loop is further from the active site. The model predicts the C3-hydroxyl will be located on the relatively occluded P_a/P_b side of the active site which is compatible with the C1-phosphoryl group specificity of PhyAsr for Ins(1,2,4,5,6)P₅.

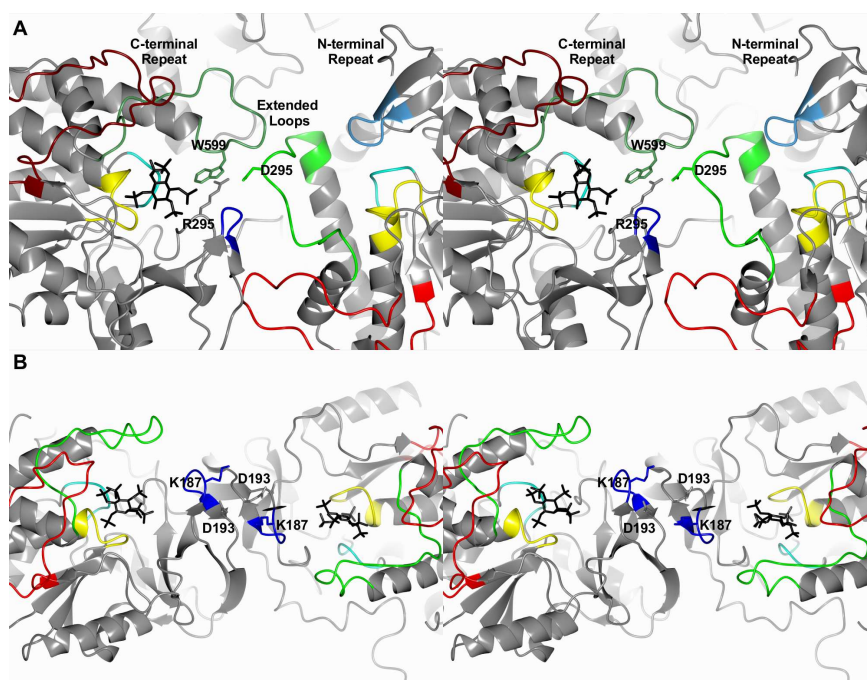


Figure 3.11: Conformation of the variable loops of both PhyAmm and PhyAsr are influenced by their quaternary structure. *A*, The dimer interface of PhyAmm includes contacts between the extended loop prior to the penultimate helix of both the N-terminal (light coloured loops) and C-terminal (dark coloured loops) repeats. The contacts are stabilized by hydrogen bonds between Asp-295 and Arg-448, and stacking interactions between Arg-448 and Trp-599. *B*, The dimer interface of PhyAsr involves the β -hairpin turn of the Phy-specific domain with Asp-192 and Gly-193 participating in multiple main-chain hydrogen bonds. Lys-187 forms a hydrogen bond with Asp-192 which locks the conformation of the β -hairpin turn. The variable loops (colored segments) that influence substrate specificity include the extended loop prior to the penultimate helix (green), the extended Ω -loop (red), and the β -hairpin and α turns (blue) within the Phy-specific domain. The active-site GA-loops (cyan) and P-loops (yellow) are displayed with Ins(1,2,4,5,6)P₅ (PhyAmm C-terminal repeat) and InsP₆ (PhyAsr) bound as black sticks to show protein orientation.

3.4.4 Substrate specificity of PhyAmm

The IP specificity assay accurately identifies the relative activities of the individual repeats of PhyAmm and their preferred IP substrates (Figure 3.7 and 3.8) [166, 184]. More importantly, it strongly suggests they follow a “divide and conquer” approach to completely hydrolyze InsP₆. The highly active C-terminal repeat rapidly hydrolyzes InsP₆ and the various InsP₅ substrates while multiple InsP₃ peaks accumulate throughout the time course (Figure 3.8 *B*; Table 3.4). These InsP₃ peaks accumulate to a greater extent than observed

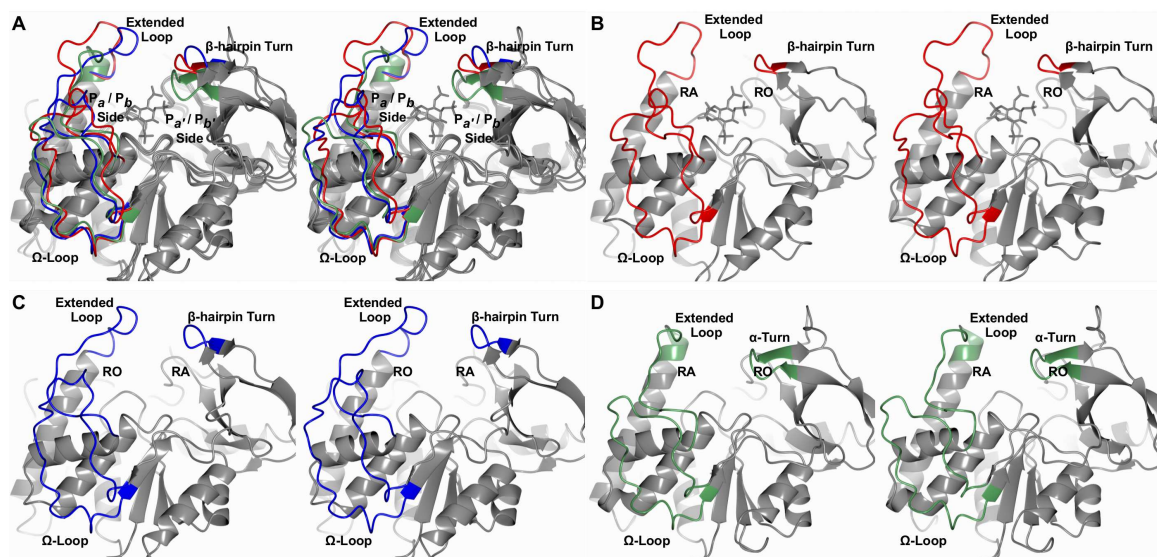


Figure 3.12: Variable loops implicated in the substrate specificity of PTPLPs. *A*, stereo view of the superposition of N-terminal (green) and C-terminal (red) repeats of PhyAmmC250S/C548S·Ins(1,2,4,5,6)P₅ and PhyAsrC252S·InsP₆ (red; PDB 3MMJ), as a ribbon diagram with the ligand as sticks. The variable loops (colored segments) that influence substrate specificity include the extended loop prior to the penultimate helix, the extended Ω -loop, and the β -hairpin (C-terminal repeat and PhyAsr) and α -turn (N-terminal repeat) within the Phy-specific domain. *B*, the active site of the C-terminal repeat (red) is relatively accessible (RA) on the P_a/P_b side and relatively occluded (RO) on the $P_{a'}/P_{b'}$ side. *C*, the PhyAsr active site (blue) is relatively occluded (RO) on the P_a/P_b side and relatively accessible (RA) on the $P_{a'}/P_{b'}$ side. *D*, the active site of the N-terminal repeat (green) is relatively accessible (RA) on the P_a/P_b side and relatively occluded (RO) on the $P_{a'}/P_{b'}$ side.

for the fully active PhyAmm suggesting they are targeted by the N-terminal repeat. The IP assay in the presence of the N-terminal repeat not only confirms its greatest activity is towards these same peaks, it also demonstrates the N-terminal repeat has limited or no activity towards the remaining peaks (Figure 3.8 C; Table 3.4). This novel assay also suggests the N-terminal repeat has an additional substrate preference. In particular, it appears to prefer InsP₄ substrates that contain three adjacent phosphoryl groups and the axial C2-phosphoryl group. This is demonstrated by the lack of activity towards the InsP₄ substrates in peaks 18, 19 and 22 while peaks 14 to 17, 20 and 21 exhibit some activity (Figure 3.8 C and B.1; Table 3.4). Peak 18 (DL-Ins(1,2,4,5)P₄) lacks three adjacent phosphoryl groups while peak 19 (DL-Ins(1,3,4,5)P₄) and peak 22 (DL-Ins(1,4,5,6)P₄) lack an axial C2-phosphoryl. The

only IP lacking an axial C2-phosphoryl not included in this list is Ins(1,3,4,6)P₄ which overlaps peak 16 (DL-Ins(1,2,3,6)P₄). The remaining InsP₄ peaks all contain both a C2-phosphoryl group and three adjacent phosphoryl groups. While the C-terminal repeat may show a slight preference for IPs containing an axial C2-phosphoryl group, the high activity of this repeat limits the detection of additional substrate preferences under the present reaction conditions.

3.4.5 InsP₆ specificity of the N-terminal repeat

The InsP₆ hydrolysis pathway of the individual repeats of PhyAmm generate the same peaks as judged by HPIC (Figure 3.7). However, we must point out that this technique cannot separate all *myo*-inositol enantiomers as indicated by their designation as DL-Ins(1,2,4,5,6)P₅ and DL-Ins(1,2,5,6)P₄ [48, 161]. In the case of the C-terminal repeat and PhyAsr, the precise enantiomer is known from crystallographic complex structures and has been confirmed for PhyAsr by alternative methods [165, 166, 184]. Consequently, PhyAsr and the C-terminal repeat are known to generate Ins(1,2,4,5,6)P₅ and are classified as 3-phytases. In the case of the N-terminal repeat, the DL-Ins(1,2,4,5,6)P₅ enantiomer generated from InsP₆ hydrolysis has yet to be identified. Based on the observed accessibility within the N-terminal repeat active site and differences in the P_a and P_{a'} sites, we suggest the N-terminal repeat is a 1-phytase that produces Ins(2,3,4,5,6)P₅.

The active sites of the N- and C-terminal repeats differ in size, shape and residue composition. As a result of these differences, the P_{a'}/P_{b'} side of the N-terminal repeat active site has less available space and the P_a/P_b side has more available space than observed in the C-terminal repeat. Consequently, the available space on each side of the active site of the N-terminal repeat is opposite that observed in PhyAsr and represents a more extreme version of the accessibility observed in the C-terminal repeat (Figure 3.12). The relatively occluded P_{a'}/P_{b'} side of the active site is likely responsible for the lack of activity of the N-terminal repeat towards InsP₆ and InsP₅ substrates as superpositions of the N-

and C-terminal repeats, and similar superpositions with PhyAsr suggest that highly phosphorylated substrates such as InsP₆ and InsP₅ will form too close contacts on this side of the active site (Figures 3.5 and 3.6). At the same time, conformational differences in the extended loop prior to the penultimate helix and the substitution of Gly-301, for Lys-600 of the C-terminal repeat, opens a large space adjacent to the GA-loop on the P_a/P_b side of the active site of the N-terminal repeat that can accommodate the axial C2-phosphoryl group. In addition to changes in active-site accessibility, residue substitutions in the N-terminal repeat active site significantly reduce the number of contacts within the P_{a'} site (Table 3.3). Specifically, three potential hydrogen bonding interaction are lost as Arg-351 and Gln-549 of the C-terminal repeat are replaced by Lys-59 and Tyr-251, respectively. Finally, the presence of the bulky Met-254 of the N-terminal repeat P-loop forms close contacts with phosphoryl groups bound to the P_b site in superpositions with the PhyAsrC252S·InsP₆ and PhyAmmC250S/C548S·Ins(1,2,4,5,6)P₅ structures. This suggests that the P_a and P_b sites of the N-terminal repeat are shifted towards the GA-loop. This would allow an axial C2-phosphoryl group to fill the P_a site, interact with the GA-loop after a small reorientation of the *myo*-inositol ring, and minimize steric interaction within the P_{a'} and P_{b'} sites. Each of the above differences in the size and shape of the N-terminal repeat active site suggest the axial C2-phosphoryl of InsP₆ can be accommodated within the P_a site, consistent with 1-phytase activity.

Chapter 4

Conclusions and future directions

4.1 Overview

Protein tyrosine phosphatase-like *myo*-inositol phosphatases (PTPLPs) are bacterial enzyme members of the protein tyrosine phosphatase (PTP) superfamily [162, 165]. In eukaryotes, PTPs remove phosphoryl groups from phosphorylated proteins, contributing to the regulation of a wide variety of signal transduction pathways [188, 189]. Originally named for their activity towards proteins containing phosphotyrosine residues, variants of classic PTPs are known to target phosphoserine and phosphothreonine containing proteins and serve as transmembrane receptors involved in cell-cell communication [188, 190]. Unlike other known PTPs, PTPLPs typically have little or no activity towards phosphoproteins and preferentially hydrolyze *myo*-inositol phosphates (IPs), an important family of eukaryotic metabolites [2, 157, 158, 160, 161, 165, 170]. These compounds have an enormous array of activities in biological systems from cofactors and second messengers to phosphorus storage compounds and lipid components [2, 18, 19].

PTPLPs were originally detected and identified in rumen bacteria as enzymes with high activity towards InsP_6 (or phytate), and represent the fourth structural class of microbial *myo*-inositol phosphatases (IPases) [114, 154, 162]. Microbial IPases have important agricultural applications and are commonly used as grain based diet feed supplements where they increase the availability of inorganic phosphate and *myo*-inositol, and decrease phosphate pollution [103, 114–117]. Consequently, initial research involving PTPLPs examined their nutritional applications and targeted highly-active enzymes believed to be phosphate

scavengers [154, 156]. Subsequent research has shown multiple PTPLPs are present in microbial pathogens where they are believed to selectively inhibit signal transduction pathways [159–161]. More recently, gut microbes expressing IPases have been shown to alter calcium mobilization in host epithelial cells, suggesting interkingdom cell-cell communication [183]. Clearly, the discovery of each additional biological function suggests the PTPLPs, and IPases in general, are more than simple phosphate scavengers.

Finally, PTPLPs sequentially remove phosphoryl groups from InsP_6 , the only abundant IP in the environment, via enzyme-specific pathways [157, 158, 165, 170]. For example, PhyAsr removes the C3-phosphoryl from InsP_6 , the C1-phosphoryl from the resulting $\text{Ins}(1,2,4,5,6)\text{P}_5$, followed by the C6-phosphoryl from $\text{Ins}(2,4,5,6)\text{P}_4$, the C5-phosphoryl from $\text{Ins}(2,4,5)\text{P}_3$ and finally the C4-phosphoryl from $\text{Ins}(2,4)\text{P}_2$ [165]. While PhyAsr does not have a strict specificity for these IPs, in the presence of saturating concentrations of InsP_6 more than 85% of the starting material is hydrolyzed by this pathway. Other characterized PTPLPs have distinct pathways and include the only known IPase to remove the C1-phosphoryl group and one of only two shown to remove the C5-phosphoryl from InsP_6 [157, 158, 161, 165, 170, 184].

4.2 Summary

In Chapter 2, I present structures of the PTPLP from *Selenomonas ruminantium* (PhyAsr) in complex with $\text{Ins}(1,3,4,5)\text{P}_4$ and $\text{Ins}(1,4,5)\text{P}_3$, and the PTPLP from *Mitsuokella multacida* (PhyAmm) in complex with $\text{Ins}(1,3,4,5)\text{P}_4$. These represent the first PTPLP structures in complex with InsP_4 and InsP_3 substrates and the first complex structure of PhyAmm. The IPs bind to each of the proteins using phosphoryl-binding sites that were initially identified in the structure of PhyAsr in complex with InsP_6 . Further, PhyAsr and the PhyAmm C-terminal repeat bind $\text{Ins}(1,3,4,5)\text{P}_4$ in identical conformations utilizing conserved residues. These complex structures demonstrate that there are no conformational changes upon substrate binding regardless of the phosphorylation level of the substrate.

However, there is a marked difference between the *myo*-inositol ring orientations of the Ins(1,3,4,5)P₄ and Ins(1,4,5)P₃ substrates when compared to InsP₆. The changes in *myo*-inositol ring orientation allow these substrates to utilize the identified phosphoryl-binding sites without changes to the preformed active site. The hydrolysis pathway of PhyAmm was determined and demonstrates that PhyAsr and PhyAmm have divergent specificities for the Ins(1,2,4,5,6)P₅, with PhyAsr producing Ins(2,4,5,6)P₄ and PhyAmm producing Ins(1,2,5,6)P₄. The distinct substrate specificity of these enzymes and the conserved interactions they make with Ins(1,3,4,5)P₄ suggests that structural features that do not directly contact the bound substrate influence substrate specificity. Three structural features that vary in sequence composition, length, and conformation were identified: the extended loop prior to the penultimate helix, the Ω -loop, and a β -hairpin turn of the Phy-specific domain. A simple hypothesis suggests the available space within the active site affects substrate specificity, which was confirmed by determining the InsP₆ hydrolysis pathway of the PT-PLP from *Bdellovibrio bacteriovorus* (PhyAbb). PhyAbb has large deletions in several of these loops and demonstrates a broad specificity towards Ins(1,2,4,5,6)P₅ by producing four different InsP₄ products.

In Chapter 3, I present the structure of PhyAmm in complex with Ins(1,2,4,5,6)P₅ that represents the first complex of a PTPLP with an InsP₅ and the first of an IP along the InsP₆ hydrolysis pathway. I have also presented the first IP specificity assay. In this assay, IPs are produced by a partial acid hydrolysis of InsP₆ and are then incubated with the protein. By producing a time-course of the enzyme hydrolysis reaction, observations can be made about the activity of the protein against the various IPs. In the case of PhyAmm, it is evident that the two repeats have divergent activities towards the IPs, with the N-terminal repeat having higher activity towards InsP₄ and InsP₃ substrates, while the C-terminal repeat has the highest activity towards InsP₆ and InsP₅ substrates. In addition, this novel assay clearly demonstrates the N-terminal repeat has its greatest activity towards IPs that accumulate in the presence of the C-terminal repeat. This clearly indicates the two repeats have a

“divide and conquer” approach to maximizing phosphoryl group removal from InsP_6 and mixed IPs. This work further confirms the hypothesis that the variable loops adjacent to the active site are influencing substrate specificity and elaborates on the structural features of PhyAmm and PhyAsr that confer the conformation of these loops.

4.3 PTPLPs for producing IPs

IPs are particularly challenging to purify or synthesize due to their low abundance and large number of stereoisomers, respectively, and only a small subset of known IPs are commercially available [22]. PTPLPs are capable of producing InsP_5 , InsP_4 , InsP_3 , InsP_2 and $\text{Ins}(2)\text{P}$ products when supplied with InsP_6 [157, 158, 161, 165, 170]. Of special interest is that the PTPLPs characterized to date have unique InsP_6 hydrolysis pathways and activities toward the various IPs. As the number of and our knowledge of PTPLPs increases, there is the potential to produce any IP by engineering the proteins to produce the specific product. Producing IPs that are currently unavailable will likely lead to additional discoveries regarding the biological roles of this family of metabolites. As it is, the characterized PTPLPs can be used to produce large quantities of IPs that are currently unavailable by halting the enzyme when the target IP is produced.

To better our knowledge of PTPLPs, increasing our understanding of the structural features that contribute to the substrate specificity will aid in our ability to identify a PTPLP that produces a desired IP, or to engineer one. Ultimately, the ability to accurately predict substrate specificity from primary sequence information would be an invaluable tool. As this work represents the first examples of PTPLPs in complex with an InsP_4 , InsP_3 , or an IP along a hydrolysis pathway, it would be beneficial to produce additional complex structures. Of special interest would be to produce PhyAsr and PhyAmm in complex with an on-pathway InsP_4 or InsP_3 so that 1) we would see if the structural features that distinguish the InsP_5 specificity translate to the other substrates, and 2) we could identify the contacts that the N-terminal repeat makes when in complex with an IP. Additional complex struc-

tures of other members of the PTPLP class of IPases would provide additional information about the structural features that have been identified in this work. Of particular interest would be PhyAbb and the PTPLP from *Legionella pneumophila* (PhyAlp) as their crystal structures have been determined making them strong candidates for producing a successful complex structure [159, 160]. Further, these proteins have larger variations in their PTPLP specific loops and residues that are predicted to interact with the substrates than PhyAmm and PhyAsr and would provide great insight into the function of these features.

4.4 Functions of PTPLPs

With their broad specificity and presence in rumen organisms, PhyAmm and PhyAsr have been predicted to function as phosphate scavengers [154, 155]. A *myo*-inositol phosphatase (IPase) from the prominent human gut bacteria *Bacteroides thetaiotaomicron* (BtMinpp) is packaged into outer membrane vesicles (OMVs) to prevent degradation from proteases within the gut [183]. Interestingly, BtMinpp-OMVs have been shown to interact with epithelial cells to promote intracellular calcium signalling. As both PhyAsr and PhyAmm are associated with the outer membrane, it is possible that with more studies, a similar cross-kingdom cell-to-cell signalling as seen with BtMinpp will be identified [155, 157, 159, 160].

The N-terminal repeat of PhyAmm displays structural features that indicate it possibly removes the C1-phosphoryl group from InsP_6 . Due to the inability to distinguish between enantiomers via HPIC, further studies must be conducted to determine the InsP_6 specificity of the N-terminal repeat. Interestingly, the first PTPLP specific for the C1-phosphoryl group of InsP_6 (1-phytase) has been identified in the pathogenic organism *Xanthomonas campestris* pv. *vesicatoria* (XopH) [161]. In addition to the unique C1-phosphoryl group specificity, XopH does not hydrolyze the $\text{Ins}(2,3,4,5,6)\text{P}_5$ product, unlike the other PTPLPs characterized to date. Further biochemical characterizations of XopH revealed that, though it does not hydrolyze the InsP_5 it produces, it successfully hydrolyzed the other five InsP_5 isomers. This indicates XopH is not functioning as a scavenging enzyme, sug-

gesting the Ins(2,3,4,5,6)P₅ product directly or indirectly affects the host organism during *Xanthomonas* infections and clearly demonstrates the importance of PTPLP substrate specificity. The IP specificity assay would help determine if XopH has narrow specificity towards IPs or if it is capable of hydrolyzing some or all of lesser phosphorylated forms revealing other possible cellular targets. The novel IP hydrolysis assay in the presence of a mixture of IPs is an obvious tool to rapidly identify the substrate specificity of PTPLPs. In turn, this information would allow research programs to target specific PTPLPs with unique specificities and more rapidly advance our understanding of this class of enzyme. Structural studies targeting pathogen PTPLPs are ongoing and aimed at generating complex structures that can be utilized to understand the structural determinants that govern their substrate specificity. Recently, a novel class of PTPLP containing an additional glycosyl hydrolase domain has been identified in *Clostridia*. Very little is known about these enzymes making them interesting candidates for both structural studies and our IP hydrolysis assay.

As interest in IPases initially arose in the field of nutrition, it is natural to ask if PTPLPs potentially have a role in human or animal nutrition. PTPLPs have favourable pH profiles and high catalytic efficiency; however, their application is restricted due to the nature of their catalytic mechanism [162–165, 168, 169]. The thiolate of the catalytic cysteine can be irreversibly oxidized leading to enzyme inactivation [168]. A possible remedy would be to modify the enzyme active site according to other PTP members that exhibit a lower tendency to oxidation or higher ability to reverse oxidation [114, 191]. As the understanding of PTPLPs increases, genetically engineering them to exhibit a lower tendency to oxidation or higher ability to reverse oxidation will be more effective. As a result, PTPLPs would become better candidates as feed additives. A large variety of PTPLPs are still being discovered and a better candidate than those we currently know may still arise.

Bibliography

- [1] T. Posternak, "The phosphoric esters of the cyclitols," *The Cyclitols*. Holden-Day, San Francisco, CA, pp. 221–243, 1965.
- [2] R. H. Michell, "Inositol derivatives: evolution and functions," *Nature Reviews Molecular Cell Biology*, vol. 9, no. 2, pp. 151–161, 2008.
- [3] L. Anderson and K. E. Wolter, "Cyclitols in plants: biochemistry and physiology," *Annual Review of Plant Physiology*, vol. 17, no. 1, pp. 209–222, 1966.
- [4] J. Scherer, "Ueber eine neue, aus dem muskelfleische gewonnene zuckerart," *Liebigs Annalen der Chemie*, vol. 73, pp. 322–328, 1850.
- [5] P. P. N. Murthy, "Structure and nomenclature of inositol phosphates, phosphoinositides, and glycosylphosphatidylinositols," in *Biology of Inositols and Phosphoinositides*, pp. 1–19, Springer, 2006.
- [6] H. Vohl, "Schwefelsaures magnesiamainganoxydulzinkoxyd, ein neues gepaartes salz (tripelgepaartes salz)," *European Journal of Organic Chemistry*, vol. 99, no. 1, pp. 124–125, 1856.
- [7] A. Darbre and F. W. Norris, "Vitamins in germination. determination of free and combined inositol in germinating oats," *Biochemical Journal*, vol. 64, no. 3, p. 441, 1956.
- [8] R. H. Michell, "Evolution of the diverse biological roles of inositols," *Cell Biology of Inositol Lipids and Phosphates*, vol. 74, pp. 223–246, 2007.
- [9] W. Pfeffer, "Zur blutenentwicklung der primulaceen und ampelideen," *Jahrbcher fr Wissenschaftliche Botanik*, vol. 8, pp. 194–215, 1872.
- [10] D. J. Cosgrove, "Chemistry and biochemistry of inositol polyphosphates," *Reviews of Pure and Applied Chemistry*, vol. 16, no. DEC, pp. 209–+, 1966.
- [11] U. Suzuki, K. Yoshimura, and M. Takaishi, "About the enzyme phytase, which splits anhydro-oxy-methylene diphosphoric acid," *Bulletin of the College of Agriculture, Tokyo Imperial University*, vol. 7, pp. 503–512, 1907.
- [12] E. Winterstein, "Ueber einen phosphorhaltigen pflanzenbestandtheil, welcher bei der spaltung inosit liefert," *European Journal of Inorganic Chemistry*, vol. 30, no. 2, pp. 2299–2302, 1897.

- [13] S. Posternak, "On a new phospho-organic principle of vegetation," *Comptes Rendus Des Seances De La Societe De Biologie Et De Ses Filiales*, vol. 55, pp. 1190–1192, 1903.
- [14] L. F. Johnson and M. E. Tate, "Structure of phytic acids," *Canadian Journal of Chemistry*, vol. 47, no. 1, pp. 63–&, 1969.
- [15] D. H. Smith and F. E. Clark, "Chromatographic separations of inositol phosphorus compounds," *Soil Science Society of America Proceedings*, vol. 16, no. 2, pp. 170–172, 1952.
- [16] D. H. Smith and F. E. Clark, "Anion-exchange chromatography of inositol phosphates from soil," *Soil Science*, vol. 72, no. 5, pp. 353–360, 1951.
- [17] K. Asada and Z. Kasai, "Formation of myo-inositol and phytin in ripening rice grains," *Plant and Cell Physiology*, vol. 3, no. 4, pp. 397–406, 1962.
- [18] R. F. Irvine and M. J. Schell, "Back in the water: the return of the inositol phosphates," *Nature Reviews Molecular Cell Biology*, vol. 2, no. 5, pp. 327–338, 2001.
- [19] M. S. Wilson, T. M. Livermore, and A. Saiardi, "Inositol pyrophosphates: between signalling and metabolism," *Biochemical Journal*, vol. 452, no. 3, pp. 369–379, 2013.
- [20] M. J. Sun, J. Alikhani, A. Massoudieh, R. Greiner, and D. P. Jaisi, "Phytate degradation by different phosphohydrolase enzymes: Contrasting kinetics, decay rates, pathways, and isotope effects," *Soil Science Society of America Journal*, vol. 81, pp. 61–75, Jan. 2017.
- [21] A. Chakraborty, "The inositol pyrophosphate pathway in health and diseases," *Biological Reviews*, vol. 93, pp. 1203–1227, May 2018.
- [22] D. C. Billington, *The inositol phosphates: chemical synthesis and biological significance*. VCH Verlagsgesellschaft mbH, 1993.
- [23] D. J. Cosgrove and G. Irving, *Inositol phosphates: their chemistry, biochemistry, and physiology*, vol. 4. Elsevier Science & Technology, 1980.
- [24] X. Tan, L. I. A. Calderon-Villalobos, M. Sharon, C. Zheng, C. V. Robinson, M. Estelle, and N. Zheng, "Mechanism of auxin perception by the tir1 ubiquitin ligase," *Nature*, vol. 446, no. 7136, pp. 640–645, 2007.
- [25] A. W. Smith, D. R. Poyner, H. K. Hughes, and P. A. Lambert, "Siderophore activity of myo-inositol hexakisphosphate in pseudomonas-aeruginosa," *Journal of Bacteriology*, vol. 176, pp. 3455–3459, June 1994.
- [26] T. A. Bolger, A. W. Folkmann, E. J. Tran, and S. R. Went, "The mrna export factor gle1 and inositol hexakisphosphate regulate distinct stages of translation," *Cell*, vol. 134, no. 4, pp. 624 – 633, 2008.

- [27] L. A. Hanakahi, M. Bartlett-Jones, C. Chappell, D. Pappin, and S. C. West, "Binding of inositol phosphate to dna-pk and stimulation of double-strand break repair," *Cell*, vol. 102, no. 6, pp. 721–729, 2000.
- [28] M. R. Macbeth, H. L. Schubert, A. P. VanDemark, A. T. Lingam, C. P. Hill, and B. L. Bass, "Inositol hexakisphosphate is bound in the adar2 core and required for rna editing," *Science*, vol. 309, no. 5740, pp. 1534–1539, 2005.
- [29] J. D. York, A. R. Odom, R. Murphy, E. B. Ives, and S. R. Went, "A phospholipase c-dependent inositol polyphosphate kinase pathway required for efficient messenger rna export," *Science*, vol. 285, no. 5424, pp. 96–100, 1999.
- [30] P. W. Majerus, J. Zou, J. Marjanovic, M. V. Kisseleva, and M. P. Wilson, "The role of inositol signaling in the control of apoptosis," *Advances in Enzyme Regulation*, vol. 48, p. 10, 2008.
- [31] P. J. Lupardus, A. Shen, M. Bogoy, and K. C. Garcia, "Small molecule-induced allosteric activation of the vibrio cholerae rtx cysteine protease domain," *Science*, vol. 322, no. 5899, pp. 265–268, 2008.
- [32] S. Chatterjee, R. Sankaranarayanan, and R. V. Sonti, "Phya, a secreted protein of xanthomonas oryzae pv. oryzae, is required for optimum virulence and growth on phytic acid as a sole phosphate source," *Molecular Plant-Microbe Interactions*, vol. 16, no. 11, pp. 973–982, 2003.
- [33] J. Verbsky, K. Lavine, and P. W. Majerus, "Disruption of the mouse inositol 1,3,4,5,6-pentakisphosphate 2-kinase gene, associated lethality, and tissue distribution of 2-kinase expression," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 24, pp. 8448–8453, 2005.
- [34] J. P. Frederick, D. Mattiske, J. A. Wofford, L. C. Megosh, L. Y. Drake, S. T. Chiou, B. L. M. Hogan, and J. D. York, "An essential role for an inositol polyphosphate multikinase, ipk2, in mouse embryogenesis and second messenger production," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 24, pp. 8454–8459, 2005.
- [35] S. J. Angyal and A. F. Russell, "Cyclitols .28. methyl esters of inositol phosphates . structure of phytic acid," *Australian Journal of Chemistry*, vol. 22, no. 2, pp. 383–&, 1969.
- [36] R. J. Anderson, "Concerning phytin in oats.," *Journal of Biological Chemistry*, vol. 17, pp. 151–163, Mar. 1914.
- [37] A. J. R. Costello, T. Glonek, and T. C. Myers, "P-31 nuclear magnetic resonance-ph titrations of myoinositol hexaphosphate," *Carbohydrate Research*, vol. 46, no. 2, pp. 159–171, 1976.

- [38] L. R. Isbrandt and R. P. Oertel, "Conformational states of myoinositol hexakis(phosphate) in aqueous-solution - a c-13 nmr, p-31 nmr, and raman-spectroscopic investigation," *Journal of the American Chemical Society*, vol. 102, no. 9, pp. 3144–3148, 1980.
- [39] E. Graf, "Applications of phytic acid," *Journal of the American Oil Chemists Society*, vol. 60, no. 11, pp. 1861–1867, 1983.
- [40] P. T. Hawkins, D. R. Poyner, T. R. Jackson, A. J. Letcher, D. A. Lander, and R. F. Irvine, "Inhibition of iron-catalyzed hydroxyl radical formation by inositol polyphosphates - a possible physiological-function for myoinositol hexakisphosphate," *Biochemical Journal*, vol. 294, pp. 929–934, Sept. 1993.
- [41] A. R. D. Boland, G. B. Garner, and B. L. Odell, "Identification and properties of phytate in cereal-grains and oilseed products," *Journal of Agricultural and Food Chemistry*, vol. 23, no. 6, pp. 1186–1189, 1975.
- [42] L. G. Barrientos and P. P. N. Murthy, "Conformational studies of myo-inositol phosphates," *Carbohydrate Research*, vol. 296, pp. 39–54, Dec. 1996.
- [43] J. H. Steward and M. E. Tate, "Gel chromatography of inositol polyphosphates and avian haemoglobin-inositol pentaphosphate complex," *Journal of Chromatography*, vol. 45, no. 3-4, pp. 400–&, 1969.
- [44] S. Rapoport and G. M. Guest, "Distribution of acid-soluble phosphorus in the blood cells of various vertebrates," *Journal of Biological Chemistry*, vol. 138, pp. 269–282, Mar. 1941.
- [45] T. A. Borgese and R. L. Nagel, "Differential effects of 2,3-dpg, atp and inositol pentaphosphate (ip5) on the oxygen equilibria of duck embryonic, fetal and adult hemoglobins," *Comparative Biochemistry and Physiology Part A: Physiology*, vol. 56, no. 4, pp. 539 – 543, 1977.
- [46] H. Streb, R. Irvine, M. Berridge, and I. Schulz, "Release of ca^{2+} from a nonmitochondrial intracellular store in pancreatic acinar cells by inositol-1,4,5-trisphosphate," *Nature*, vol. 306, pp. 67–69, 1983.
- [47] J. P. Heslop, R. F. Irvine, A. H. Tashjian, and M. J. Berridge, "Inositol tetrakisphosphate and pentakisphosphate in gh4 cells," *Journal of Experimental Biology*, vol. 119, pp. 395–401, Nov. 1985.
- [48] B. Q. Phillippy and J. M. Bland, "Gradient ion chromatography of inositol phosphates," *Analytical Biochemistry*, vol. 175, pp. 162–166, Nov. 1988.
- [49] L. B. Sheard, X. Tan, H. B. Mao, J. Withers, G. Ben-Nissan, T. R. Hinds, Y. Kobayashi, F. F. Hsu, M. Sharon, J. Browse, S. Y. He, J. Rizo, G. A. Howe, and N. Zheng, "Jasmonate perception by inositol-phosphate-potentiated coil-jaz co-receptor," *Nature*, vol. 468, pp. 400–U301, Nov. 2010.

- [50] E. A. Orchiston, D. Bennett, N. R. Leslie, R. G. Clarke, L. Winward, C. P. Downes, and S. T. Safrany, "Pten m-cbr3, a versatile and selective regulator of inositol 1,3,4,5,6-pentakisphosphate (ins(1,3,4,5,6)p-5) - evidence for ins(1,3,4,5,6)p-5 as a proliferative signal," *Journal of Biological Chemistry*, vol. 279, pp. 1116–1122, Jan. 2004.
- [51] E. Piccolo, S. Vignati, T. Maffucci, P. F. Innominato, A. M. Riley, B. V. L. Potter, P. P. Pandolfi, M. Broggini, S. Iacobelli, P. Innocenti, and M. Falasca, "Inositol pentakisphosphate promotes apoptosis through the pi3-k/akt pathway," *Oncogene*, vol. 23, pp. 1754–1765, Mar. 2004.
- [52] S. Campbell, R. J. Fisher, E. M. Towler, S. Fox, H. J. Issaq, T. Wolfe, L. R. Phillips, and A. Rein, "Modulation of hiv-like particle assembly in vitro by inositol phosphates," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, pp. 10875–10879, Sept. 2001.
- [53] D. J. Steger, E. S. Haswell, A. L. Miller, S. R. Wentz, and E. K. O'Shea, "Regulation of chromatin remodeling by inositol polyphosphates," *Science*, vol. 299, pp. 114–116, Jan. 2003.
- [54] S. G. Jackson, Y. Zhang, R. J. Haslam, and M. S. Junop, "Structural analysis of the carboxy terminal ph domain of pleckstrin bound to d-myo-inositol 1,2,3,5,6-pentakisphosphate," *Bmc Structural Biology*, vol. 7, p. 80, Nov. 2007.
- [55] J. F. Quignard, L. Rakotoarisoa, J. Mironneau, and C. Mironneau, "Stimulation of l-type ca^{2+} channels by inositol pentakis- and hexakisphosphates in rat vascular smooth muscle cells," *Journal of Physiology-london*, vol. 549, pp. 729–737, June 2003.
- [56] L. Stephens, T. Radenberg, U. Thiel, G. Vogel, K. H. Khoo, A. Dell, T. R. Jackson, P. T. Hawkins, and G. W. Mayr, "The detection, purification, structural characterization, and metabolism of diphosphoinositol pentakisphosphate(s) and bisdiphosphoinositol tetrakisphosphate(s)," *Journal of Biological Chemistry*, vol. 268, pp. 4009–4015, Feb. 1993.
- [57] F. S. Menniti, R. N. Miller, J. W. Putney, and S. B. Shears, "Turnover of inositol polyphosphate pyrophosphates in pancreatoma cells," *Journal of Biological Chemistry*, vol. 268, pp. 3850–3856, Feb. 1993.
- [58] D. Allan and R. H. Michell, "Phosphatidylinositol cleavage in lymphocytes. requirement for calcium ions at a low concentration and effects of other cations," *Biochemical Journal*, vol. 142, no. 3, pp. 599–604, 1974.
- [59] J. N. Fain and M. J. Berridge, "Relationship between phosphatidylinositol synthesis and recovery of 5-hydroxytryptamine-responsive ca^{2+} flux in blowfly salivary glands," *Biochemical Journal*, vol. 180, no. 3, pp. 655–661, 1979.

- [60] M. D. Bootman, T. J. Collins, C. M. Peppiatt, L. S. Prothero, L. MacKenzie, P. De Smet, M. Travers, S. C. Tovey, J. T. Seo, M. J. Berridge, F. Ciccolini, and P. Lipp, "Calcium signalling - an overview," *Seminars in Cell & Developmental Biology*, vol. 12, pp. 3–10, Feb. 2001.
- [61] I. R. Batty, S. R. Nahorski, and R. F. Irvine, "Rapid formation of inositol 1,3,4,5-tetrakisphosphate following muscarinic receptor stimulation of rat cerebral cortical slices," *Biochemical Journal*, vol. 232, no. 1, pp. 211–215, 1985.
- [62] D. J. Gawler, B. Potter, and S. R. Nahorski, "Inositol 1, 3, 4, 5-tetrakisphosphate-induced release of intracellular Ca^{2+} in sh-sy5y neuroblastoma cells," *Biochemical journal*, vol. 272, no. 2, pp. 519–524, 1990.
- [63] M. Vajanaphanich, C. Schultz, M. T. Rudolf, M. Wasserman, P. Enyedi, A. Craxton, S. B. Shears, R. Y. Tsien, K. E. Barrett, and A. Traynorkaplan, "Long-term uncoupling of chloride secretion from intracellular calcium levels by $ins(3,4,5,6)p-4$," *Nature*, vol. 371, pp. 711–714, Oct. 1994.
- [64] P. H. Hirst, A. M. Riley, S. J. Mills, I. D. Spiers, D. R. Poyner, S. Freeman, B. V. L. Potter, and A. W. Smith, "Inositol polyphosphate-mediated iron transport in *Pseudomonas aeruginosa*," *Journal of Applied Microbiology*, vol. 86, pp. 537–543, Mar. 1999.
- [65] X. T. Shen, H. Xiao, R. Ranallo, W. H. Wu, and C. Wu, "Modulation of atp-dependent chromatin-remodeling complexes by inositol polyphosphates," *Science*, vol. 299, pp. 112–114, Jan. 2003.
- [66] A. R. Odom, A. Stahlberg, S. R. Went, and J. D. York, "A role for nuclear inositol 1,4,5-trisphosphate kinase in transcriptional control," *Science*, vol. 287, pp. 2026–2029, Mar. 2000.
- [67] R. J. Anderson, "The chemistry of the lipoids of tubercle bacilli. xiv. the occurrence of inositol in the phosphatide from human tubercle bacilli," *Journal of the American Chemical Society*, vol. 52, pp. 1607–1608, Jan. 1930.
- [68] H. E. Carter, W. D. Celmer, D. S. Galanos, R. H. Gigg, W. E. M. Lands, J. H. Law, K. L. Mueller, T. Nakayama, H. H. Tomizawa, and E. Weber, "Biochemistry of the sphingolipides .10. phytoglycolipide, a complex phytosphingosine-containing lipid from plant seeds," *Journal of the American Oil Chemists Society*, vol. 35, no. 7, pp. 335–343, 1958.
- [69] J. Folch, "The nature of the glycerophosphoric acid present in phosphatides," *Journal of Biological Chemistry*, vol. 146, pp. 31–33, Nov. 1942.
- [70] J. Folch, "Brain cephalin, a mixture of phosphatides. separation from it of phosphatidyl serine, phosphatidyl ethanolamine, and a fraction containing an inositol phosphatide," *Journal of Biological Chemistry*, vol. 146, pp. 35–44, Nov. 1942.

- [71] J. Folch, "Brain diphosphoinositide, a new phosphatide having inositol metadiphosphate as a constituent," *Journal of Biological Chemistry*, vol. 177, no. 2, pp. 505–519, 1949.
- [72] G. Di Paolo and P. De Camilli, "Phosphoinositides in cell regulation and membrane dynamics," *Nature*, vol. 443, pp. 651–657, Oct. 2006.
- [73] L. E. Hokin, "The role of ribonucleic acids in amylase secretion by pancreas slices," *Biochimica Et Biophysica Acta*, vol. 8, no. 2, pp. 225–226, 1952.
- [74] M. R. Hokin and L. E. Hokin, "Enzyme secretion and the incorporation of p-82 into phospholipides of pancreas slices," *Journal of Biological Chemistry*, vol. 203, no. 2, pp. 967–977, 1953.
- [75] L. E. Hokin and M. R. Hokin, "Metabolism of phospholipids invitro," *Canadian Journal of Biochemistry and Physiology*, vol. 34, no. 2, pp. 349–360, 1956.
- [76] R. H. Michell, "Inositol phospholipids and cell-surface receptor function," *Biochimica Et Biophysica Acta*, vol. 415, no. 1, pp. 81–147, 1975.
- [77] M. J. Berridge, R. M. C. Dawson, C. P. Downes, J. P. Heslop, and R. F. Irvine, "Changes in the levels of inositol phosphates after agonist-dependent hydrolysis of membrane phosphoinositides," *Biochemical Journal*, vol. 212, no. 2, pp. 473–482, 1983.
- [78] J. A. Creba, C. P. Downes, P. T. Hawkins, G. Brewster, R. H. Michell, and C. J. Kirk, "Rapid breakdown of phosphatidylinositol 4-phosphate and phosphatidylinositol 4,5-bisphosphate in rat hepatocytes stimulated by vasopressin and other Ca^{2+} -mobilizing hormones," *Biochemical Journal*, vol. 212, no. 3, pp. 733–747, 1983.
- [79] P. Garcia, R. Gupta, S. Shah, A. J. Morris, S. A. Rudge, S. Scarlata, V. Petrova, S. McLaughlin, and M. J. Rebecchi, "The pleckstrin homology domain of phospholipase c-delta(1) binds with high affinity to phosphatidylinositol 4,5-bisphosphate in bilayer membranes," *Biochemistry*, vol. 34, pp. 16228–16234, Dec. 1995.
- [80] I. Gaidarov and J. H. Keen, "Phosphoinositide-ap-2 interactions required for targeting to plasma membrane clathrin-coated pits," *Journal of Cell Biology*, vol. 146, pp. 755–764, Aug. 1999.
- [81] H. C. Lin, B. Barylko, M. Achiriloaie, and J. P. Albanesi, "Phosphatidylinositol (4,5)-bisphosphate-dependent activation of dynamins i and ii lacking the proline/arginine-rich domains," *Journal of Biological Chemistry*, vol. 272, pp. 25999–26004, Oct. 1997.
- [82] K. Takei, V. Haucke, V. Slepnev, K. Farsad, M. Salazar, H. Chen, and P. De Camilli, "Generation of coated intermediates of clathrin-mediated endocytosis on protein-free liposomes," *Cell*, vol. 94, pp. 131–141, July 1998.

- [83] D. W. Hilgemann and R. Ball, "Regulation of cardiac Na^+ , Ca^{2+} exchange and K-ATP potassium channels by PIP_2 ," *Science*, vol. 273, pp. 956–959, Aug. 1996.
- [84] R. Rohatgi, H. Y. H. Ho, and M. W. Kirschner, "Mechanism of N-WASP activation by CDC42 and phosphatidylinositol 4,5-bisphosphate," *Journal of Cell Biology*, vol. 150, pp. 1299–1309, Sept. 2000.
- [85] M. C. Glennon and S. B. Shears, "Turnover of inositol pentakisphosphates, inositol hexakisphosphate and diphosphoinositol polyphosphates in primary cultured-hepatocytes," *Biochemical Journal*, vol. 293, pp. 583–590, July 1993.
- [86] S. B. Shears, N. A. Gokhale, H. C. Wang, and A. Zaremba, "Diphosphoinositol polyphosphates: What are the mechanisms?," 2011.
- [87] A. Saiardi, R. Bhandari, A. C. Resnick, A. M. Snowman, and S. H. Snyder, "Phosphorylation of proteins by inositol pyrophosphates," *Science*, vol. 306, pp. 2101–2105, Dec. 2004.
- [88] A. Saiardi, C. Sciambi, J. M. McCaffery, B. Wendland, and S. H. Snyder, "Inositol pyrophosphates regulate endocytic trafficking," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, pp. 14206–14211, Oct. 2002.
- [89] C. Illies, J. Gromada, R. Fiume, B. Leibiger, J. Yu, K. Juhl, S. N. Yang, D. K. Barma, J. R. Falck, A. Saiardi, C. J. Barker, and P. O. Berggren, "Requirement of inositol pyrophosphates for full exocytotic capacity in pancreatic beta cells," *Science*, vol. 318, pp. 1299–1302, Nov. 2007.
- [90] W. L. Ye, N. Ali, M. E. Bembenek, S. B. Shears, and E. M. Lafer, "Inhibition of clathrin assembly by high-affinity binding of specific inositol polyphosphates to the synapse-specific clathrin assembly protein AP-3," *Journal of Biological Chemistry*, vol. 270, pp. 1564–1568, Jan. 1995.
- [91] C. Horigome, R. Ikeda, T. Okada, K. Takenami, and K. Mizuta, "Genetic interaction between ribosome biogenesis and inositol polyphosphate metabolism in *Saccharomyces cerevisiae*," *Bioscience Biotechnology and Biochemistry*, vol. 73, pp. 443–446, Feb. 2009.
- [92] H. R. Luo, Y. E. Huang, J. C. Chen, A. Saiardi, M. Iijima, K. Ye, Y. Huang, E. Nagata, P. Devreotes, and S. H. Snyder, "Inositol pyrophosphates mediate chemotaxis in *Dictyostelium* via pleckstrin homology domain-PTDins (3, 4, 5) P3 interactions," *Cell*, vol. 114, no. 5, pp. 559–572, 2003.
- [93] B. H. Morrison, J. A. Bauer, D. V. Kalvakolanu, and D. J. Lindner, "Inositol hexakisphosphate kinase 2 mediates growth suppressive and apoptotic effects of interferon-beta in ovarian carcinoma cells," *Journal of Biological Chemistry*, vol. 276, pp. 24965–24970, July 2001.

- [94] A. Saiardi, A. C. Resnick, A. M. Snowman, B. Wendland, and S. H. Snyder, "Inositol pyrophosphates regulate cell death and telomere length through phosphoinositide 3-kinase-related protein kinases," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, pp. 1911–1914, Feb. 2005.
- [95] E. Dubois, B. Scherens, F. Vierendeels, M. M. W. Ho, F. Messenguy, and S. B. Shears, "In *saccharomyces cerevisiae*, the inositol polyphosphate kinase activity of *kcs1p* is required for resistance to salt stress, cell wall integrity, and vacuolar morphogenesis," *Journal of Biological Chemistry*, vol. 277, pp. 23755–23763, June 2002.
- [96] U. Konietzny and R. Greiner, "Molecular and catalytic properties of phytate-degrading enzymes (phytases)," *International Journal of Food Science and Technology*, vol. 37, no. 7, pp. 791–812, 2002.
- [97] J. Stevenson-Paulik, S.-T. Chiou, J. P. Frederick, A. M. Seeds, J. C. Otto, and J. D. York, "Inositol phosphate metabolomics: merging genetic perturbation with modernized radiolabeling methods," *Methods*, vol. 39, no. 2, pp. 112–121, 2006.
- [98] G. E. Gillaspay, "The cellular language of myo-inositol signaling," *New Phytologist*, vol. 192, no. 4, pp. 823–839, 2011.
- [99] E. J. Mullaney, C. B. Daly, and A. H. Ullah, "Advances in phytase research," *Advances in Applied Microbiology*, vol. 47, pp. 157–199, 2000.
- [100] F. A. Norris, M. P. Wilson, T. S. Wallis, E. E. Galyov, and P. W. Majerus, "Sopb, a protein required for virulence of *salmonella dublin*, is an inositol phosphate phosphatase," *Proceedings of the National Academy of Sciences*, vol. 95, no. 24, pp. 14057–14059, 1998.
- [101] E. Mellanby, "Accessory food factors (vitamines) in the feeding of infants," *Lancet*, vol. 1, pp. 856–862, 1920.
- [102] H. M. Bruce and R. K. Callow, "Cereals and rickets. the role of inositolhexaphosphoric acid.," *Biochemical Journal*, vol. 28, pp. 517–528, 1934.
- [103] T. G. Taylor, "Availability of calcium and phosphorus of plant materials for animals," *Proceedings of the Nutrition Society*, vol. 24, no. 1, pp. 105–&, 1965.
- [104] J. T. Lowe and H. Steenbock, "Cereals and rickets. vii. the role of inorganic phosphorus in calcification on cereal diets.," *Biochemical Journal*, vol. 30, pp. 1126–1134, 1936.
- [105] V. N. Patwardhan, "The occurrence of a phytinsplitting enzyme in the intestines of albino rats," *Biochemical Journal*, vol. 31, pp. 560–564, 1937.
- [106] R. R. Spitzer and P. H. Phillips, "Enzymatic relationships in the utilization of soybean oil meal phosphorus by the rat: Two figures," *The Journal of Nutrition*, vol. 30, no. 3, pp. 183–192, 1945.

- [107] E. Singsen and H. Mitchell, "Soybean meal chick rations need no inorganic phosphorus supplements," *Poultry Science*, vol. 23, no. 2, pp. 152–153, 1944.
- [108] J. H. Moore and C. Tyler, "Studies on the intestinal absorption and excretion of calcium and phosphorus in the pig .1. a critical study of the bergeim technique for investigating the intestinal absorption and excretion of calcium and phosphorus," *British Journal of Nutrition*, vol. 9, no. 1, pp. 63–80, 1955.
- [109] J. H. Moore and C. Tyler, "Studies on the intestinal absorption and excretion of calcium and phosphorus in the pig .2. the intestinal absorption and excretion of radioactive calcium and phosphorus," *British Journal of Nutrition*, vol. 9, no. 1, pp. 81–93, 1955.
- [110] R. Hill and C. Tyler, "The influence of time, temperature, ph and calcium carbonate on the activity of the phytase of certain cereals," *Journal of Agricultural Science*, vol. 44, no. 3, pp. 306–310, 1954.
- [111] R. A. McCance and E. M. Widdowson, "Activity of the phytase in different cereals and its resistance to dry heat," *Nature*, vol. 153, pp. 650–650, Jan. 1944.
- [112] H. Mollgaard, "On phytic acid, its importance in metabolism and its enzymic cleavage in bread supplemented with calcium," *Biochemical Journal*, vol. 40, no. 4, pp. 589–603, 1946.
- [113] A. Raun, E. Cheng, and W. Burroughs, "Phytate phosphorus hydrolysis and availability to rumen microorganisms," *Journal of Agricultural and Food Chemistry*, vol. 4, no. 10, pp. 869–871, 1956.
- [114] C. C. Chen, K. J. Cheng, T. P. Ko, and R. T. Guo, "Current progresses in phytase research: Three-dimensional structure and protein engineering," *Chembioeng Reviews*, vol. 2, pp. 76–86, Apr. 2015.
- [115] Z. Yi, E. T. Kornegay, V. Ravindran, and D. M. Denbow, "Improving phytate phosphorus availability in corn and soybean meal for broilers using microbial phytase and calculation of phosphorus equivalency values for phytase," *Poultry Science*, vol. 75, pp. 240–249, Feb. 1996.
- [116] W. H. Jordan, E. F. Hart, and A. J. Patten, "A study of the metabolism and physiological effects of certain phosphorus compounds with milch cows.," *American Journal of Physiology*, vol. 16, pp. 268–313, June 1906.
- [117] R. J. Wodzinski and A. H. J. Ullah, "Phytase," *Advances in Applied Microbiology*, Vol 42, vol. 42, pp. 263–302, 1996.
- [118] I. Egli, L. Davidsson, C. Zeder, T. Walczyk, and R. Hurrell, "Dephytinization of a complementary food based on wheat and soy increases zinc, but not copper, apparent absorption in adults," *Journal of Nutrition*, vol. 134, pp. 1077–1080, May 2004.

- [119] R. F. Hurrell, M. B. Reddy, M. A. Juillerat, and J. D. Cook, "Degradation of phytic acid in cereal porridges improves iron absorption by human subjects," *American Journal of Clinical Nutrition*, vol. 77, pp. 1213–1219, May 2003.
- [120] R. Greiner, U. Konietzny, D. M. Blackburn, and M. A. Jorquera, "Production of partially phosphorylated myo-inositol phosphates using phytases immobilised on magnetic nanoparticles," *Bioresource Technology*, vol. 142, pp. 375–383, Aug. 2013.
- [121] S. Haefner, A. Knietsch, E. Scholten, J. Braun, M. Lohscheidt, and O. Zelder, "Biotechnological production and applications of phytases," *Applied Microbiology and Biotechnology*, vol. 68, pp. 588–597, Sept. 2005.
- [122] J. Dvorakova, J. Kopecky, V. Havlicek, and V. Kren, "Formation of myo-inositol phosphates by *aspergillus niger* 3-phytase," *Folia Microbiologica*, vol. 45, no. 2, pp. 128–132, 2000.
- [123] T. T. Chen, F. W. Bazer, J. J. Cetorelli, W. E. Pollard, and R. M. Roberts, "Purification and properties of a progesterone-induced basic glycoprotein from uterine fluids of pigs," *Journal of Biological Chemistry*, vol. 248, no. 24, pp. 8560–8566, 1973.
- [124] D. C. Schlosnagle, F. W. Bazer, J. C. M. Tsibris, and R. M. Roberts, "Iron-containing phosphatase induced by progesterone in uterine fluids of pigs," *Journal of Biological Chemistry*, vol. 249, no. 23, pp. 7574–7579, 1974.
- [125] J. B. Vincent, G. L. Oliver-Lilley, and B. A. Averill, "Proteins containing oxo-bridged dinuclear iron centers - a bioinorganic perspective," *Chemical Reviews*, vol. 90, pp. 1447–1467, Dec. 1990.
- [126] C. E. Hegeman and E. A. Grabau, "A novel phytase with sequence similarity to purple acid phosphatases is expressed in cotyledons of germinating soybean seedlings," *Plant Physiology*, vol. 126, pp. 1598–1608, Aug. 2001.
- [127] A. H. J. Ullah and B. J. Cummins, "Aspergillus-ficuum extracellular ph 6.0 optimum acid-phosphatase - purification, n-terminal amino-acid sequence, and biochemical-characterization," *Preparative Biochemistry*, vol. 18, no. 1, pp. 37–65, 1988.
- [128] R. Kuang, K.-H. Chan, E. Yeung, and B. L. Lim, "Molecular and biochemical characterization of atpap15, a purple acid phosphatase with phytase activity, in arabidopsis," *Plant Physiology*, vol. 151, no. 1, pp. 199–209, 2009.
- [129] G. Schenk, N. Mitic, G. R. Hanson, and P. Comba, "Purple acid phosphatase: A journey into the function and mechanism of a colorful enzyme," *Coordination Chemistry Reviews*, vol. 257, pp. 473–482, Jan. 2013.
- [130] T. Klabunde, N. Strater, R. Frohlich, H. Witzel, and B. Krebs, "Mechanism of fe(iii)-zn(ii) purple acid phosphatase based on crystal structures," *Journal of Molecular Biology*, vol. 259, pp. 737–748, June 1996.

- [131] E. J. Mullaney and A. H. J. Ullah, "The term phytase comprises several different classes of enzymes," *Biochemical and Biophysical Research Communications*, vol. 312, no. 1, pp. 179–184, 2003.
- [132] D. P. Li, H. F. Zhu, K. F. Liu, X. Liu, G. Leggewie, M. Udvardi, and D. W. Wang, "Purple acid phosphatases of *Arabidopsis thaliana* - comparative analysis and differential regulation by phosphate deprivation," *Journal of Biological Chemistry*, vol. 277, pp. 27772–27781, Aug. 2002.
- [133] J. Kerovuo, M. Lauraeus, P. Nurminen, N. Kalkkinen, and J. Apajalahti, "Isolation, characterization, molecular gene cloning, and sequencing of a novel phytase from *Bacillus subtilis*," *Applied and Environmental Microbiology*, vol. 64, no. 6, pp. 2079–2085, 1998.
- [134] Y. O. Kim, H. K. Kim, K. S. Bae, J. H. Yu, and T. K. Oh, "Purification and properties of a thermostable phytase from *Bacillus* sp. ds11," *Enzyme and Microbial Technology*, vol. 22, pp. 2–7, Jan. 1998.
- [135] B. L. Lim, P. Yeung, C. Cheng, and J. E. Hill, "Distribution and diversity of phytate-mineralizing bacteria," *Isme Journal*, vol. 1, pp. 321–330, Aug. 2007.
- [136] H. Q. Huang, P. J. Shi, Y. R. Wang, H. Y. Luo, N. Shao, G. Z. Wang, P. L. Yang, and B. Yao, "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," *Applied and Environmental Microbiology*, vol. 75, pp. 1508–1516, Mar. 2009.
- [137] S. Shin, N. C. Ha, B. C. Oh, T. K. Oh, and B. H. Oh, "Enzyme mechanism and catalytic property of beta propeller phytase," *Structure*, vol. 9, pp. 851–858, Sept. 2001.
- [138] Y. F. Zeng, T. P. Ko, H. L. Lai, Y. S. Cheng, T. H. Wu, Y. H. Ma, C. C. Chen, C. S. Yang, K. J. Cheng, C. H. Huang, R. T. Guo, and J. R. Liu, "Crystal structures of *Bacillus* alkaline phytase in complex with divalent metal ions and inositol hexasulfate," *Journal of Molecular Biology*, vol. 409, pp. 214–224, June 2011.
- [139] N. C. Ha, B. C. Oh, S. Shin, H. J. Kim, T. K. Oh, Y. O. Kim, K. Y. Choi, and B. H. Oh, "Crystal structures of a novel, thermostable phytase in partially and fully calcium-loaded states," *Nature Structural Biology*, vol. 7, pp. 147–153, Feb. 2000.
- [140] R. Greiner, B. L. Lim, C. W. Cheng, and N. G. Carlsson, "Pathway of phytate dephosphorylation by beta-propeller phytases of different origins," *Canadian Journal of Microbiology*, vol. 53, pp. 488–495, Apr. 2007.
- [141] J. Kerovuo, J. Rouvinen, and F. Hatzack, "Analysis of myo-inositol hexakisphosphate hydrolysis by *Bacillus* phytase: indication of a novel reaction mechanism," *Biochemical Journal*, vol. 352, pp. 623–628, Dec. 2000.
- [142] A. W. Dox and R. Golden, "Phytase in lower fungi," *Journal of Biological Chemistry*, vol. 10, pp. 183–186, Oct. 1911.

- [143] D. Kostrewa, F. G. Leitch, A. D'Arcy, C. Broger, D. Mitchell, and A. P. G. M. van Loon, "Crystal structure of phytase from *aspergillus ficuum* at 2.5 angstrom resolution," *Nature Structural Biology*, vol. 4, pp. 185–190, Mar. 1997.
- [144] A. J. Oakley, "The structure of *aspergillus niger* phytase phya in complex with a phytate mimetic," *Biochemical and Biophysical Research Communications*, vol. 397, pp. 745–749, July 2010.
- [145] D. C. Lee, M. A. Cottrill, C. W. Forsberg, and Z. Jia, "Functional insights revealed by the crystal structures of *escherichia coli* glucose-1-phosphatase," *Journal of Biological Chemistry*, vol. 278, no. 33, pp. 31412–31418, 2003.
- [146] D. Kostrewa, M. Wyss, A. D'Arcy, and A. P. G. M. van Loon, "Crystal structure of *aspergillus niger* ph 2.5 acid phosphatase at 2.4 angstrom resolution," *Journal of Molecular Biology*, vol. 288, pp. 965–974, May 1999.
- [147] D. Lim, S. Golovan, C. W. Forsberg, and Z. C. Jia, "Crystal structures of *escherichia coli* phytase and its complex with phytate," *Nature Structural Biology*, vol. 7, pp. 108–113, Feb. 2000.
- [148] A. Ariza, O. V. Moroz, E. V. Blagova, J. P. Turkenburg, J. Waterman, S. M. Roberts, J. Vind, C. Sjøholm, S. F. Lassen, L. De Maria, *et al.*, "Degradation of phytate by the 6-phytase from *hafnia alvei*: A combined structural and solution study," *PLoS One*, vol. 8, no. 5, p. e65062, 2013.
- [149] T. Xiang, Q. Liu, A. M. Deacon, M. Koshy, I. A. Kriksunov, X. G. Lei, Q. Hao, and D. J. Thiel, "Crystal structure of a heat-resilient phytase from *aspergillus fumigatus*, carrying a phosphorylated histidine," *Journal of Molecular Biology*, vol. 339, pp. 437–445, May 2004.
- [150] K. Ostanin, E. H. Harms, P. E. Stevis, R. Kuceil, M. M. Zhou, and R. L. Vanetten, "Overexpression, site-directed mutagenesis, and mechanism of *escherichia coli* acid-phosphatase," *Journal of Biological Chemistry*, vol. 267, pp. 22830–22836, Nov. 1992.
- [151] K. Ostanin and R. L. Van Etten, "Asp(304) of *escherichia coli* acid-phosphatase is involved in leaving group protonation," *Journal of Biological Chemistry*, vol. 268, pp. 20778–20784, Oct. 1993.
- [152] O. Liu, Q. Q. Huang, X. G. Lei, and Q. Hao, "Crystallographic snapshots of *aspergillus fumigatus* phytase, revealing its enzymatic dynamics," *Structure*, vol. 12, pp. 1575–1583, Sept. 2004.
- [153] B. C. Oh, W. C. Choi, S. Park, Y. O. Kim, and T. K. Oh, "Biochemical properties and substrate specificities of alkaline and histidine acid phytases," *Applied Microbiology and Biotechnology*, vol. 63, pp. 362–372, Jan. 2004.
- [154] L. J. Yanke, H. D. Bae, L. B. Selinger, and K. J. Cheng, "Phytase activity of anaerobic ruminal bacteria," *Microbiology-uk*, vol. 144, pp. 1565–1573, June 1998.

- [155] C. G. D'Silva, H. D. Bae, L. J. Yanke, K. J. Cheng, and L. B. Selinger, "Localization of phytase in *selenomonas ruminantium* and *mitsuokella multiacidus* by transmission electron microscopy," *Canadian Journal of Microbiology*, vol. 46, pp. 391–395, Apr. 2000.
- [156] B. A. Nakashima, T. A. McAllister, R. Sharma, and L. B. Selinger, "Diversity of phytases in the rumen," *Microbial Ecology*, vol. 53, pp. 82–88, Jan. 2007.
- [157] A. A. Puhl, R. Greiner, and L. B. Selinger, "A protein tyrosine phosphatase-like inositol polyphosphatase from *selenomonas ruminantium* subsp. *lactilytica* has specificity for the 5-phosphate of myo-inositol hexakisphosphate," *The International Journal of Biochemistry & Cell Biology*, vol. 40, no. 10, pp. 2053–2064, 2008.
- [158] A. A. Puhl, R. Greiner, and L. B. Selinger, "Stereospecificity of myo-inositol hexakisphosphate hydrolysis by a protein tyrosine phosphatase-like inositol polyphosphatase from *megaspheera elsdenii*," *Applied Microbiology and Biotechnology*, vol. 82, no. 1, pp. 95–103, 2009.
- [159] S. Weber, C. U. Stirnimann, M. Wieser, D. Frey, R. Meier, S. Engelhardt, X. Li, G. Capitani, R. A. Kammerer, and H. Hilbi, "A type iv translocated legionella cysteine phytase counteracts intracellular growth restriction by phytate," *Journal of Biological Chemistry*, vol. 289, no. 49, pp. 34175–34188, 2014.
- [160] R. J. Gruninger, J. Thibault, M. J. Capeness, R. Till, S. C. Mosimann, R. E. Sockett, B. L. Selinger, and A. L. Lovering, "Structural and biochemical analysis of a unique phosphatase from *bdellovibrio bacteriovorus* reveals its structural and functional relationship with the protein tyrosine phosphatase class of phytase," *PLoS One*, vol. 9, no. 4, 2014.
- [161] D. Bluher, D. Laha, S. Thieme, A. Hofer, L. Eschen-Lippold, A. Masch, G. Balcke, I. Pavlovic, O. Nagel, A. Schonsky, R. Hinkelmann, J. Worner, N. Parvin, R. Greiner, S. Weber, A. Tissier, M. Schutkowski, J. Lee, H. Jessen, G. Schaaf, and U. Bonas, "A 1-phytase type iii effector interferes with plant hormone signaling," *Nature Communications*, vol. 8, p. 2159, 2017.
- [162] H.-M. Chu, R.-T. Guo, T.-W. Lin, C.-C. Chou, H.-L. Shr, H.-L. Lai, T.-Y. Tang, K.-J. Cheng, B. L. Selinger, and A. H.-J. Wang, "Structures of *selenomonas ruminantium* phytase in complex with persulfated phytate: Dsp phytase fold and mechanism for sequential substrate hydrolysis," *Structure*, vol. 12, no. 11, pp. 2015–2024, 2004.
- [163] K. L. Guan and J. E. Dixon, "Evidence for protein-tyrosine-phosphatase catalysis proceeding via a cysteine-phosphate intermediate," *Journal of Biological Chemistry*, vol. 266, no. 26, pp. 17026–17030, 1991.
- [164] Z. Y. Zhang, "Chemical and mechanistic approaches to the study of protein tyrosine phosphatases," *Accounts of Chemical Research*, vol. 36, no. 6, pp. 385–392, 2003.

- [165] A. A. Puhl, R. J. Gruninger, R. Greiner, T. W. Janzen, S. C. Mosimann, and L. B. Selinger, "Kinetic and structural analysis of a bacterial protein tyrosine phosphatase-like myo-inositol polyphosphatase," *Protein Science*, vol. 16, no. 7, pp. 1368–1378, 2007.
- [166] R. J. Gruninger, S. Dobing, A. D. Smith, L. M. Bruder, L. B. Selinger, H.-J. Wieden, and S. C. Mosimann, "Substrate binding in protein-tyrosine phosphatase-like inositol polyphosphatases," *Journal of Biological Chemistry*, vol. 287, no. 13, pp. 9722–9730, 2012.
- [167] L. Wu and Z. Y. Zhang, "Probing the function of asp128 in the low molecular weight protein-tyrosine phosphatase-catalyzed reaction. a pre-steady-state and steady-state kinetic investigation," *Biochemistry*, vol. 35, pp. 5426–5434, Apr. 1996.
- [168] R. J. Gruninger, L. B. Selinger, and S. C. Mosimann, "Effect of ionic strength and oxidation on the p-loop conformation of the protein tyrosine phosphatase-like phytase, phyasr," *The FEBS Journal*, vol. 275, no. 15, pp. 3783–3792, 2008.
- [169] R. J. Gruninger, L. B. Selinger, and S. C. Mosimann, "Structural analysis of a multifunctional, tandemly repeated inositol polyphosphatase," *Journal of Molecular Biology*, vol. 392, no. 1, pp. 75–86, 2009.
- [170] A. A. Puhl, R. Greiner, and L. B. Selinger, "Kinetics, substrate specificity, and stereospecificity of two new protein tyrosine phosphatase-like inositol polyphosphatases from *Selenomonas laticifex*," *Biochemistry and Cell Biology*, vol. 86, no. 4, pp. 322–330, 2008.
- [171] L. Chen and M. F. Roberts, "Characterization of a tetrameric inositol monophosphatase from the hyperthermophilic bacterium *Thermotoga maritima*," *Applied and Environmental Microbiology*, vol. 65, no. 10, pp. 4559–4567, 1999.
- [172] T. G. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell, and A. G. W. Leslie, "imosflm: a new graphical interface for diffraction-image processing with mosflm," *Acta Crystallographica Section D-biological Crystallography*, vol. 67, pp. 271–281, 2011.
- [173] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, and K. S. Wilson, "Overview of the ccp4 suite and current developments," *Acta Crystallographica Section D-biological Crystallography*, vol. 67, pp. 235–242, 2011.
- [174] P. Evans, "Scaling and assessment of data quality," *Acta Crystallographica Section D: Biological Crystallography*, vol. 62, no. 1, pp. 72–82, 2006.
- [175] P. R. Evans, "An introduction to data reduction: space-group determination, scaling and intensity statistics," *Acta Crystallographica Section D: Biological Crystallography*, vol. 67, no. 4, pp. 282–292, 2011.

- [176] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan, "Features and development of coot," *Acta Crystallographica Section D: Biological Crystallography*, vol. 66, no. 4, pp. 486–501, 2010.
- [177] A. A. Vaguine, J. Richelle, and S. Wodak, "Sfcheck: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model," *Acta Crystallographica Section D: Biological Crystallography*, vol. 55, no. 1, pp. 191–205, 1999.
- [178] S. McNicholas, E. Potterton, K. Wilson, and M. Noble, "Presenting your structures: the ccp4mg molecular-graphics software," *Acta Crystallographica Section D: Biological Crystallography*, vol. 67, no. 4, pp. 386–394, 2011.
- [179] K. Blaabjerg, J. Hansen-Møller, and H. D. Poulsen, "High-performance ion chromatography method for separation and quantification of inositol phosphates in diets and digesta," *Journal of Chromatography. B, Analytical Technologies in the Biomedical and Life Sciences*, vol. 878, no. 3, pp. 347–354, 2010.
- [180] A. Espinosa, M. Guo, V. C. Tam, Z. Q. Fu, and J. R. Alfano, "The pseudomonas syringae type iii-secreted protein hopptod2 possesses protein tyrosine phosphatase activity and suppresses programmed cell death in plants," *Molecular Microbiology*, vol. 49, no. 2, pp. 377–387, 2003.
- [181] J. R. Bretz, N. M. Mock, J. C. Charity, S. Zeyad, C. J. Baker, and S. W. Hutcheson, "A translocated protein tyrosine phosphatase of pseudomonas syringae pv. tomato dc3000 modulates plant defence response to infection," *Molecular Microbiology*, vol. 49, pp. 389–400, 2003.
- [182] V. Raboy, "myo-inositol-1,2,3,4,5,6-hexakisphosphate," *Phytochemistry*, vol. 64, no. 6, pp. 1033–1043, 2003.
- [183] R. Stentz, S. Osborne, N. Horn, A. W. H. Li, I. Hautefort, R. Bongaerts, M. Rouyer, P. Bailey, S. B. Shears, A. M. Hemmings, C. A. Brearley, and S. R. Carding, "A bacterial homolog of a eukaryotic inositol phosphate signaling enzyme mediates cross-kingdom dialog in the mammalian gut," *Cell Reports*, vol. 6, pp. 646–656, Feb. 2014.
- [184] L. M. Bruder, R. J. Gruninger, C. P. Cleland, and S. C. Mosimann, "Bacterial phya protein tyrosine phosphatase-like myo-inositol phosphatases in complex with the ins(1,3,4,5)p4 and ins(1,4,5)p3 second messengers," *Journal of Biological Chemistry*, vol. 292, no. 42, pp. 17302–17311, 2017.
- [185] A. A. Puhl, R. Greiner, and L. B. Selinger, "Stereospecificity of myo-inositol hexakisphosphate hydrolysis by a protein tyrosine phosphatase-like inositol polyphosphatase from megasphaera elsdenii," *Applied Microbiology and Biotechnology*, vol. 82, pp. 95–103, Feb. 2009.
- [186] I. P. Street, H. R. Coffman, and C. D. Poulter, "Isopentenyl diphosphate isomerase - site-directed mutagenesis of cys139 using counter pcr amplification of an expression plasmid," *Tetrahedron*, vol. 47, no. 31, pp. 5919–5924, 1991.

- [187] P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, and P. H. Zwart, “Phenix: a comprehensive python-based system for macromolecular structure solution,” *Acta Crystallographica Section D-biological Crystallography*, vol. 66, pp. 213–221, 2010.
- [188] A. Alonso, J. Sasin, N. Bottini, I. Friedberg, I. Friedberg, A. Osterman, A. Godzik, T. Hunter, J. Dixon, and T. Mustelin, “Protein tyrosine phosphatases in the human genome,” *Cell*, vol. 117, pp. 699–711, June 2004.
- [189] T. Tiganis and A. M. Bennett, “Protein tyrosine phosphatase function: the substrate perspective,” *Biochemical Journal*, vol. 402, pp. 1–15, Feb. 2007.
- [190] Y. G. Shi, “Serine/threonine phosphatases: Mechanism through structure,” *Cell*, vol. 139, pp. 468–484, Oct. 2009.
- [191] A. Claiborne, J. I. Yeh, T. C. Mallett, J. Luba, E. J. Crane, V. Charrier, and D. Parsonage, “Protein-sulfenic acids: Diverse roles for an unlikely player in enzyme catalysis and redox regulation,” *Biochemistry*, vol. 38, pp. 15407–15416, Nov. 1999.

Appendix A

A.1 Identification of PhyAsr, PhyAmm and PhyAbb hydrolysis products

Hydrolysis of 10 mM InsP_6 (Sigma-Aldrich) was carried out at room temperature in the presence of 50 nM wild type PhyAsr and PhyAmm, and 100 nM wild type PhyAbb in 50 mM Na-acetate pH 5.0, 300 mM NaCl, 5 mM BME, 0.1 mM EDTA. Aliquots of 100 μL were taken, quenched with 5% trichloroacetic acid and subjected to HPIC (Waters 1525 Binary HPIC Pump; Milford, MA) utilizing a CarboPac PA-100 (4 x 240 mm) analytical column (Dionex; Sunnyvale, CA) [179], at room temperature using a post-column reactor flow rate of 0.2 mL/min. The reference sample was produced by following the methodology of Blaabjerg et al. (2010) [179]. Briefly, partial hydrolysis of InsP_6 was performed by boiling 1.5 g InsP_6 with 100 mL of 0.5 M HCl for 11 hours. The solution was evaporated to dryness and the residue was dissolved to 100 mg/mL. Peak assignment was performed by referencing the chromatograms to the peak identities determined by Blaabjerg et al. (2010) [179]. Spiking experiments were performed by mixing the reference sample (30%) and the hydrolysis product (70%) of PhyAsr (30 minute time point), PhyAmm (60 minute time point) and PhyAbb (60 minute time point).

A.2 Figures and tables

Table A.1: Pairwise RMSD comparison of least squares (LSQ) superposition of the PhyAsrC252S monomer alone and in complex with IPs. RMSDs (\AA) for the monomer main- (mc; 1248 atoms) and side-chain (sc; 1293 atoms) atoms are shown. In all cases, chain A of the indicated structures were used in the comparison. The main-chain conformation of these structures are essentially identical as the RMSDs are comparable to the co-ordinate errors in the structure determinations.

	InsP_6^2		$\text{Ins}(1,3,4,5)\text{P}_4^3$		$\text{Ins}(1,4,5)\text{P}_3^4$	
	mc	sc	mc	sc	mc	sc
Apo ¹	0.14	0.57	0.18	0.87	0.22	0.77
InsP_6^2			0.19	0.86	0.21	0.64
$\text{Ins}(1,3,4,5)\text{P}_4^3$					0.19	0.77

¹PhyAsr wild-type structure: PDB 2PSZ

²PhyAsrC252S- InsP_6 structure: PDB 3MMJ

³PhyAsrC252S- $\text{Ins}(1,3,4,5)\text{P}_4$ structure: PDB 4WTY

⁴PhyAsrC252S- $\text{Ins}(1,4,5)\text{P}_3$ structure: PDB 4WU2

Table A.2: Pairwise RMSD comparison of LSQ superposition of the PhyAsrC252S active site alone and in complex with IPs. Main- (mc; 148 atoms) and side-chain (sc; 182 atoms) RMSDs (Å) of the active-site residues are shown. In all cases, chain A of the indicated structures were used in the comparison.

	InsP ₆ ²		Ins(1,3,4,5)P ₄ ³		Ins(1,4,5)P ₃ ⁴	
	mc	sc	mc	sc	mc	sc
Apo ¹	0.13	0.36	0.19	0.57	0.18	0.44
InsP ₆ ²			0.17	0.50	0.14	0.28
Ins(1,3,4,5)P ₄ ³					0.12	0.49

¹PhyAsr wild-type structure: PDB 2PSZ

²PhyAsrC252S·InsP₆ structure: PDB 3MMJ

³PhyAsrC252S·Ins(1,3,4,5)P₄ structure: PDB 4WTY

⁴PhyAsrC252S·Ins(1,4,5)P₃ structure: PDB 4WU2

Table A.3: Pairwise RMSD (Å) comparison of LSQ superposition of the PhyAmm monomers from the P2₁, and P1 space groups, in complex with Ins(1,3,4,5)P₄ (PDB 4WU3) and without (apo; PDB 3F41). Superpositions were performed using amino acids 47 to 636 (m; 2360 atoms), 47 to 342 (n; 1184 atoms) and 343 to 636 (c; 1176 atoms).

	Apo Chain B			Ins(1,3,4,5)P ₄ Chain A			Ins(1,3,4,5)P ₄ Chain B			Ins(1,3,4,5)P ₄ Chain C			Ins(1,3,4,5)P ₄ Chain D		
	m	n	c	m	n	c	m	n	c	m	n	c	m	n	c
Apo Chain A	0.64	0.28	0.38	0.51	0.38	0.38	0.46	0.36	0.40	0.50	0.41	0.38	0.57	0.35	0.39
Ins(1,3,4,5)P ₄ Chain A				0.68	0.38	0.43	0.59	0.36	0.42	0.70	0.40	0.43	0.42	0.35	0.38
Ins(1,3,4,5)P ₄ Chain B							0.26	0.16	0.27	0.12	0.13	0.09	0.43	0.16	0.30
Ins(1,3,4,5)P ₄ Chain C										0.25	0.13	0.27	0.30	0.16	0.18
Ins(1,3,4,5)P ₄ Chain D													0.45	0.19	0.31

Table A.4: Pairwise RMSD comparison of LSQ superposition of the C-terminal repeat active site of PhyAmmC250S/C548S in complex with Ins(1,3,4,5)P₄ (PDB 4WU3) and without (PDB 3F41). Main- (mc; 184 atoms) and side-chain (sc; 216 atoms) RMSDs (Å) of the active-site residues are shown. The main-chain conformation of these structures are essentially identical as the RMSDs are comparable to the co-ordinate errors in the structure determinations.

	Apo		Ins(1,3,4,5)P ₄		Ins(1,3,4,5)P ₄		Ins(1,3,4,5)P ₄		Ins(1,3,4,5)P ₄	
	Chain B		Chain A		Chain B		Chain C		Chain D	
	mc	sc	mc	sc	mc	sc	mc	sc	mc	sc
Apo										
Chain A	0.21	0.49	0.19	0.34	0.18	0.34	0.19	0.35	0.17	0.34
Ins(1,3,4,5)P ₄										
Chain A			0.30	0.57	0.27	0.56	0.30	0.58	0.22	0.53
Ins(1,3,4,5)P ₄										
Chain B					0.12	0.16	0.07	0.10	0.14	0.20
Ins(1,3,4,5)P ₄										
Chain C							0.11	0.14	0.11	0.14
Ins(1,3,4,5)P ₄										
Chain D									0.15	0.19

Table A.5: Less-phosphorylated IP ring shifts relative to InsP₆ (PDB 3MMJ) in LSQ superposed structures (active site). Distances (Å) between the C6 of InsP₆ and the structurally equivalent carbons, the angles between the plane of the 6-carbon rings, and the distances between the centres of mass of the 6-carbon rings. The 6-carbon ring angle is the rotation angle between the planes of the rings.

	Ins(1,3,4,5)P ₄ ¹	Ins(1,4,5)P ₃ ²
C6 Distance	1.53 Å	1.51 Å
6-carbon ring angle	30.79°	36.66°
6-carbon ring distance	1.24 Å	1.25 Å

¹PhyAsrC252S·Ins(1,3,4,5)P₄ structure: PDB 4WTY

²PhyAsrC252S·Ins(1,4,5)P₃ structure: PDB 4WU2

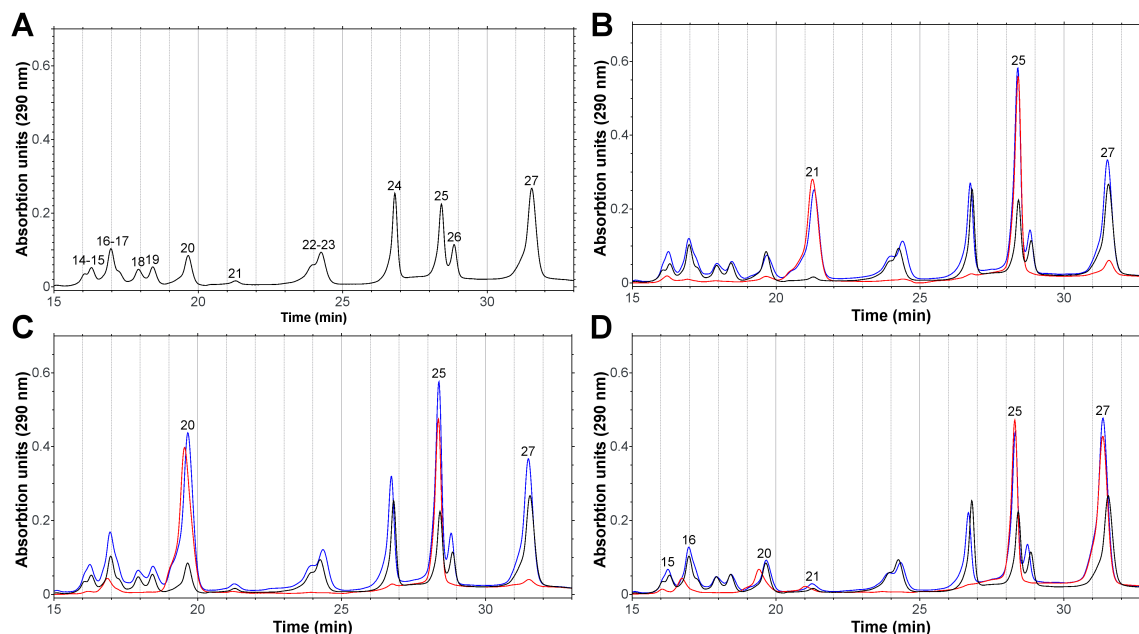


Figure A.1: Identification of products from the hydrolysis of InsP_6 by PhyAsr and PhyAmm. (A) Chromatogram of the reference samples produced by an 11 hour acid hydrolysis as performed by and identified by Blaabjerg et al. (2010) [179]. The peaks are identified as: (14) $\text{Ins}(1,2,3,5)\text{P}_4$, (15) $\text{DL-Ins}(1,2,4,6)\text{P}_4$, (16) $\text{DL-Ins}(1,2,3,6)\text{P}_4$, (17) $\text{Ins}(1,3,4,6)\text{P}_4$, (18) $\text{DL-Ins}(1,2,4,5)\text{P}_4$, (19) $\text{DL-Ins}(1,3,4,5)\text{P}_4$, (20) $\text{DL-Ins}(1,2,5,6)\text{P}_4$, (21) $\text{Ins}(2,4,5,6)\text{P}_4$, (22) $\text{DL-Ins}(1,4,5,6)\text{P}_4$, (23) $\text{Ins}(1,2,3,4,6)\text{P}_5$, (24) $\text{DL-Ins}(1,2,3,4,5)\text{P}_5$, (25) $\text{DL-Ins}(1,2,4,5,6)\text{P}_5$, (26) $\text{Ins}(1,3,4,5,6)\text{P}_5$, (27) InsP_6 . (B) Identification of the PhyAsr hydrolysis product by incubating 50 nM wild type PhyAsr with 10 mM InsP_6 for 30 minutes (red), spiked with 30 g/L of the reference sample (blue) and the reference sample (30 g/L; black). (C) Identification of the PhyAmm hydrolysis product by incubating 50 nM wild type PhyAmm with 10 mM InsP_6 for 60 minutes (red), spiked with 30 g/L of the reference sample (blue) and the reference sample (30 g/L; black). (D) Identification of the PhyAbb hydrolysis product by incubating 100 nM wild type PhyAbb with 10 mM InsP_6 for 60 minutes (red), spiked with 30 g/L of the reference sample (blue) and the reference sample (30 g/L; black). The IPs were separated using a CarboPac PA-100 analytical column with a methanesulfonic acid gradient [179]. IPs were visualized using a post-column reactor with 0.1% (m/v) $\text{Fe}(\text{NO}_3)_3$ in a 2% (m/v) HClO_4 solution (0.2 mL/min).

Appendix B

B.1 Figures and tables

Table B.1: Pairwise RMSD (Å) comparison of LSQ superposition of the PhyAmm monomers from the InsP₅ (P1), Ins(1,3,4,5)P₄ (P1) and apo (without substrate; P2₁) structures. Main-chain atom superpositions were performed using amino acids 47 to 636 (m; 2360 atoms), 47 to 342 (n; 1184 atoms) and 343 to 636 (c; 1176 atoms).

	Ins(1,3,4,5)P ₄ ²			Apo ³		
	m	n	c	m	n	c
Ins(1,2,4,5,6)P ₅ ¹	0.33	0.27	0.31	0.53	0.44	0.47
Ins(1,3,4,5)P ₄ ²				0.51	0.38	0.38

¹PhyAmmC250S/C548S·InsP₅ structure

²PhyAmmC250S/C548S·Ins(1,3,4,5)P₄ structure: PDB 4WU3

³PhyAmm Apo structure: PDB 3F41

Table B.2: Pairwise RMSD comparison of LSQ superposition of the C-terminal repeat active site of PhyAmmC250S/C548S alone and in complex with InsP₅ and Ins(1,3,4,5)P₄. Main- (mc; 164 atoms) and side-chain (sc; 194 atoms) RMSDs (Å) of the active-site residues are shown. The main-chain conformation of these structures are essentially identical as the RMSDs are comparable to the co-ordinate errors in the structure determinations.

	Ins(1,3,4,5)P ₄ ²		Apo ³	
	mc	sc	mc	sc
Ins(1,2,4,5,6)P ₅ ¹	0.22	0.50	0.25	0.59
Ins(1,3,4,5)P ₄ ²			0.18	0.46

¹PhyAmmC250S/C548S·InsP₅ structure

²PhyAmmC250S/C548S·Ins(1,3,4,5)P₄ structure: PDB 4WU3

³PhyAmm Apo structure: PDB 3F41

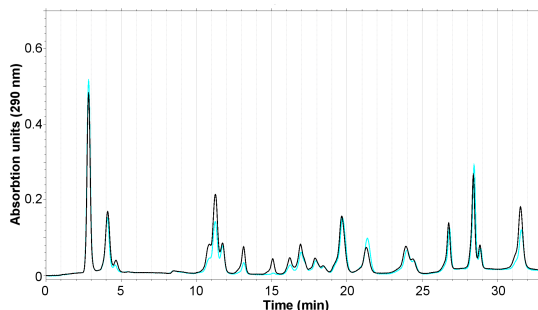


Figure B.1: HPIC chromatograms of the activities PhyAmmC548S against the 11 hour acid hydrolysis mixture. The PhyAmmC548S at 200 nM was incubated with 20 mg/mL of the 11 hour acid hydrolysis for 60 min (cyan). Production of the 11 hour acid hydrolysis sample (black) as performed previously [179, 184]. The peaks are identified by Blaabjerg et al. (2010) [179] as: (1-5) InsP_2 , (6-11) InsP_3 , (12) DL- $\text{Ins}(1,5,6)\text{P}_3$, (13) DL- $\text{Ins}(4,5,6)\text{P}_3$, (14) $\text{Ins}(1,2,3,5)\text{P}_4$, (15) DL- $\text{Ins}(1,2,4,6)\text{P}_4$, (16) DL- $\text{Ins}(1,2,3,6)\text{P}_4$, (17) $\text{Ins}(1,3,4,6)\text{P}_4$, (18) DL- $\text{Ins}(1,2,4,5)\text{P}_4$, (19) DL- $\text{Ins}(1,3,4,5)\text{P}_4$, (20) DL- $\text{Ins}(1,2,5,6)\text{P}_4$, (21) $\text{Ins}(2,4,5,6)\text{P}_4$, (22) DL- $\text{Ins}(1,4,5,6)\text{P}_4$, (23) $\text{Ins}(1,2,3,4,6)\text{P}_5$, (24) DL- $\text{Ins}(1,2,3,4,5)\text{P}_5$, (25) DL- $\text{Ins}(1,2,4,5,6)\text{P}_5$, (26) $\text{Ins}(1,3,4,5,6)\text{P}_5$, (27) InsP_6 . The IPs were separated using a CarboPac PA-100 analytical column with a methanesulfonic acid gradient [179, 184]. IPs were visualized using a post-column reactor with 0.1% (m/v) $\text{Fe}(\text{NO}_3)_3$ in a 2% (m/v) HClO_4 solution (0.2 mL/min).