New Analytical Tools to Interrogate Inositol Pyrophosphate Signaling

Dissertation zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.) im Fach Chemie

Eingereicht an der Mathematisch-Naturwissenschaftlichen Fakultät der Humboldt-Universität zu Berlin

von

M. Sc. Robert Klaus Harmel

Präsidentin der Humboldt-Universität zu Berlin

Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät

Prof. Dr. Elmar Kulke

1. Gutachterin: Prof. Dr. Dorothea Fiedler
2. Gutachter: Prof. Dr. habil. Hans Börner
3. Gutachter: Prof. Dr. Adolfo Saiardi

Tag der mündlichen Prüfung: 20.02.2020

Erklärung des Autors

Diese Arbeit wurde vom 01.09.2015 bis zum 20.9.2019 unter der Aufsicht von Prof. Dorothea Fiedler am Institute für Chemie der Humboldt-Universität zu Berlin und am Leibniz-Forschungsinstitut für Molekulare Pharmakologie angefertigt.

Deklaration

Ich erkläre, dass ich diese Dissertation selbständig und nur unter Verwendung der von mir gemäß § 7 Abs. 3 der Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät, veröffentlicht im Amtlichen Mitteilungsblatt der Humboldt-Universität zu Berlin Nr. 42/2018 am 11.7.2018 angegebenen Hilfsmittel angefertigt habe.

Robert Harmel

Berlin, 18.09.2019

Declaration

Hereby, I declare that I have authored the present dissertation independently and only by means of the stated resources in accordance to § 7 Abs. 3 der Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät, published in the Amtlichen Mitteilungsblatt der Humboldt-Universität zu Berlin Nr. 42/2018 at 11.7.2018. All citations are marked as such. This thesis was not submitted to other examination boards in the same or similar form before.

Robert Harmel

Berlin, 18.09.2019

Acknowledgments

All of the work in this thesis would not have been possible without my family, friends and colleagues that supported me all the way through my scientific development. First of all, I want to thank my parents, Andrea and Klaus, and my sister Cornelia who continuously motivated me to follow my personal dreams and who supported me throughout my whole life. You always helped me through difficult situations but we also had and have a lot of fun together. You also opened up new perspectives for me by encouraging me to go to the US, the Netherlands and Sweden. Thank you for everything!

Next, I want to thank Saskia who is not only my wife but also one of my best friends. You always have to deal with all of my projects and listen to my ideas but also problems and keep giving good feedback. Thank you for keeping the ears open for such a long time and thank you for always backing me up if necessary! I also want to acknowledge my little son Jannik who was born in the middle of my PhD and who does a tremendous job in getting my mind of work. With you around, the world became a little more easygoing!

Of course, I want to give credit to my best friends Ken and Matze who I know forever and who always help me to ground myself. I enjoy that you at least pretend to have interest in what I am doing in the lab, even though, I often talk way too long about it and with too much detail. I hope this friendship will be continued for an even longer time. With this I also want to thank Thomas Meffert who also became a good friend over the past couple years and that shares his project management skills with me to organize myself even though in some cases I did not ask.

Tjeerd and Gaston, we spent a great time together during our studies in Nijmegen and it is fantastic that we still continue to meet each other on a regular basis after such a long time. Tjeerd, you are always very honest and direct which I really appreciate and Gaston, you as my personal science rival made me more competitive and your enthusiasm for science always motivated me. However, I had also great times with you guys outside of work, especially while drinking special bears in a gezellig bruin café.

Importantly, I won't forget to mention my former colleagues from previous labs. First, Floris Rutjes and Ferdi van der Pijl who supervised my work in the synthetic organic chemistry group in Nijmegen. You were always fair, gave me good advice and supervision, and helped me to get my first two publications. That was a great start! Second, Satish Kitambi, Lars Hammaström and Annika Jenmalm-Jensen who guided my transition from organic synthetic chemistry to chemical biology and medicinal chemistry. You showed me the next level in science and made me aware of the importance of setting up good projects. This really saved me from making the wrong choices during my PhD. Third, I also want to thank Thomas Penkert and Christian Kimmich who helped me through a tough time and who created a positive atmosphere in the lab. On top of all that, I want to generally be grateful to all the members of the Rutjes lab, LCBKI and the Ernfors lab who supported me at my stays there.

Coming to my colleagues from the Fiedler group that became also friends for the most part. Dorothea, you gave me the opportunity to develop myself into a true chemical biologist in your lab. When I started, you gave me a project that did not even make it into the thesis but that did not matter because I always had the feeling you had some sort of plan and we will work it out. You give great feedback for lab work and for writing articles and it is fun to talk with you about private and scientific topics. Thank

you so much that you always care about everyone's projects and that you are such a fair and nice supervisor. You made my PhD a fun experience and you showed be many important skills that will strongly support me in situations concerning my private and scientific life. I don't want to forget to mention that you are never too cool for something that makes fun and participate in all events even though they might be a little bit weird sometimes. Special thanks goes definitely to Robi, my work wife, who collaborated on a lot of projects with me and who always gave valuable input during discussions. In addition, you became a good friend that I like to spend time with. I truly think there could not have been a better partner for the time here at the FMP. So to say: Robertis Rock!!! I also want to mention my former students Minh and Femke who worked hard under my supervision but hopefully got mainly positive memories from that. Thanks for the enthusiasm and all the valuable time that you spent to execute my project ideas. Some credit needs to be given to Nate, Alan and Ceasar who friendly welcomed me in the group and who set me up with everything what I needed and who usually gave great feedback. I also want to acknowledge Jeremy Morgan who took over the proteomics project from me but who also gave outmost important feedback for writing this thesis but also on many occasions concerning project progress. With this, I also want to thank the "mean postdoc" Sarah Hostachy to proofread chapter 3 which was basically half of this manuscript. Importantly, I also need to thank my lab predecessor that have done such a nice work preparing tools and data that were reliable and that I could use to accomplish my projects. To conclude with the Fiedler group, I want to also thank all the people that I did not mention in detail previously. I really appreciated your feedback, all the fun we had together and the nice hours during lunch, coffee breaks and so on.

Now, I want to take the opportunity to be grateful to all the collaborators that I encountered during my time. Most importantly, I want to thank Peter Schmieder and the NMR core facility who spent a great amount of time with me on the NMR instrument and gave extremely valuable input to help and develop all the great methods we have set up within the last couple of years. I want to thank the Schaaf group and the Hothorn group for involving us in the plant work. I also want to thank the Saiardi group for letting us work on slime mold project and especially Adolfo who gave us heads up information that were very useful avoiding some decent pitfalls that were we maneuvering towards. Also, I want to thank you that you take the time and became a part of my thesis committee. I want to thank the Krause/Liu group, the MS core facility and especially Martin Penkert who was extremely motivated to find endogenously pyrophosphorylated peptides. Next, I want to thank the Hackenberger group who was a really good neighbor and supported us whenever necessary. Here also special thanks to Christian for participating in my thesis committee.

Last but no least I want to thank, my thesis committee for careful reading of this thesis and all the people from the FMP and the Humboldt Universität that helped me during my PhD, and I want to apologize to all the people that I forgot in my definitely too short acknowledgment to give everyone personal credit.

Table of Contents

Abstract

Inositol pyrophosphates (PP-InsPs) are an important group of second messengers that intersect with a wide range of processes in eukaryotic cells including phosphate homeostasis, insulin signaling and apoptosis. However, despite their discovery more than two decades ago, elucidating the underlying signaling mechanisms still remains a significant challenge due to the limited access of synthetic PP-InsPs and a lack of widely available analytical tools. Therefore, a new set of chemical and analytical methods was developed to improve our understanding of these intriguing molecules on the biochemical and cellular level.

To overcome the shortage of synthetic PP-InsPs, a highly efficient and scalable chemoenzymatic approach was designed. Coupled to a simple purification method based on the precipitation with Mg^{2+} ions, the major mammalian PP-InsPs could be obtained in hundreds of milligrams quantities and high purity circumventing labor-intensive classical organic synthesis. In parallel, a new analytical tools was developed to quantify levels of PP-InsPs in complex samples. Chemoenzymatic access to ^{13}C-labeled *myo*-inositol and ^{13}C-labeled PP-InsPs enabled the detection of low concentrations of PP-InsPs using nuclear magnetic resonance (NMR) spectroscopy. *In vitro*, these compounds were of great use for the biochemical characterization of PP-InsPs kinases from mammals and plants. Endogenous pools of PP-InsPs from human cell lines could be identified and quantified by metabolic labeling with ^{13}C-labeled *myo*-inositol and followed by assignment of characteristic signals using synthetic PP-InsP standards. Finally, a new proteomics workflow towards the detection of protein pyrophosphorylation, a posttranslational modification (PTM) mediated by PP-InsPs, using mass spectrometry (MS) was optimized and endogenously modified mammalian proteins could be identified for the first time and with high confidence. In

total, 36 pyrophosphorylated proteins could be detected and gene ontology analysis suggests a function of this PTM in ribosome biogenesis, cell-cell adhesion and cell division.

Taken together, the chemical and analytical tools presented have great potential to accelerate the understanding of PP-InsP signaling and metabolism. The supply of large amounts of PP-InsPs to the research community will strongly encourage scientists to interrogate these molecules in their biological context. In addition, the quantification of PP-InsPs with NMR and the identification of endogenous pyrophosphorylation sites by MS are complementary methods that will be essential to unraveling the signaling mechanisms underlying the diverse phenotypes associated with these metabolites.

Zusammenfassung

Bei den Inositolpyrophosphaten (PP-InsPs) handelt es sich um eine wichtige Gruppe eukaryotischer sekundärer Botenstoffe, die mit einer Reihe von verschiedenen Prozessen wie Apoptose, Phosphathomeostase und Insulinsignalkaskaden verknüpft sind. Trotz ihrer Entdeckung vor mehr als 20 Jahren bleibt es eine große Herausforderung, die Mechanismen hinter den Signalwirkungen dieser Moleküle zu verstehen. Die Ursache dafür ist der limitierte Zugang zu synthetischen PP-InsPs und es einen Mangel an allgemein zugänglichen analytischen Methoden. Daher wurden in dieser Arbeit chemische und analytische Methoden entwickelt, um unser Verständnis von diesen Molekülen sowohl auf ein biochemischer als auch auf zelluläre Ebene zu verbessern.

Um der Knappheit an synthetischen PP-InsPs entgegen zu wirken, wurde ein hocheffizientes synthetisches Verfahren entwickelt. In Kombination mit einer einfachen Aufreinigungsmethode, durch Fällung mittels Zugabe von Mg^{2+} Ionen, konnten mehr als 100 mg aller wesentlichen PP-InsPs von Säugern hergestellt werden. Parallel wurde ein neues analytisches Werkzeug entwickelt, dass Konzentrationen von PP-InsPs in komplexen Proben quantifizieren konnte. Mittel Enzymkatalyse konnten ^{13}C-markiertes *myo*-inositol und ^{13}C-markierte PP-InsPs hergestellt werden und niedrige Konzentrationen mit nuklearer Magnetresonanzspektroskopie (NMR) detektiert werden. *In vitro* waren diese Verbindungen sehr nützlich, um PP-InsP Kinasen von Pflanzen und Säugern biochemisch zu charakterisieren. Endogene Konzentrationen von PP-InsPs konnten durch metabolisches Markieren mit ^{13}C-markiertem *myo*-inositol in humanen Zelllinien identifiziert und quantifiziert werden. Letztendlich wurde auch ein neuer proteomischer Ansatz zur Detektion von Proteinpyrophosphorilierung, eine von PP-InsP eingebaute posttranslationale Proteinmodifikation (PTM), mittels

Massenspektrometrie (MS) optimiert und endogen modifizierte Proteine von Säugertieren konnten zum ersten Mal nachgewiesen werden. Insgesamt wurden so 36 pyrophosphorilierte Proteine gefunden und eine Genanalyse suggerierte Funktionen dieser PTM in Ribosombiogenese, Zell-Zell Adhäsion und Zellteilung.

Zusammenfassend haben die aufgelisteten chemischen und analytischen Werkzeuge ein hohes Potenzial das Verständnis von PP-InsPs bedingten Prozessen zu beschleunigen. Die Bereitstellung von größeren Mengen an PP-InsPs für die breite Forschung wird Wissenschaftler ermutigen diese Moleküle in neuen biologischen Kontexten zu studieren. Zusätzlich werden die komplementären Methoden, Quantifizierung von PP-InsPs mittels NMR und Identifikation von Proteinpyrophosphorilierung mittels MS, essenziell sein, um die Signalmechanismen, die den diversen Phänotypen der PP-InsPs zu Grunde liegen zu verstehen.

List of Figures, Tables and Schemes

List of Abbreviations

[$^{13}C_6$]	uniformly ^{13}C-labeled on 6 carbon atoms (of *myo*-inositol)
[$^{13}C_6$]1,5(PP)$_2$-InsP$_4$	[$^{13}C_6$]1,5-bisdiphosphoinositol tetrakisphosphate
[$^{13}C_6$]1PP-InsP$_5$	[$^{13}C_6$]1-diphosphoinositol pentakisphosphate
[$^{13}C_6$]5PP-InsP$_5$	[$^{13}C_6$]5-diphosphoinositol pentakisphosphate
[$^{13}C_6$]InsP$_5$	[$^{13}C_6$]inositol pentakisphosphate
[$^{13}C_6$]InsP$_6$	[$^{13}C_6$]inositol hexakisphosphate
$^1H,^1H$-COSY	$^1H,^1H$ correlation spectroscopy
$^1H,^{13}C$-HMQC	$^1H,^{13}C$ Heteronuclear Multiple-Quantum Correlation
$^1H,^{31}P$-HMBC	$^1H,^{31}P$ Heteronuclear Multiple Bond Correlation
1,5(PP)$_2$-InsP$_4$	1,5-bisdiphosphoinositol tetrakisphosphate
1PP-InsP$_5$	1-diphosphoinositol pentakisphosphate
4,5(PP)$_2$-InsP$_4$	4,5-bisdiphosphoinositol tetrakisphosphate
5PP-InsP$_5$	5-diphosphoinositol pentakisphosphate
6,5(PP)$_2$-InsP$_4$	6,5-bisdiphosphoinositol tetrakisphosphate
AC	adenylyl cyclase
ADP	adenosine diphosphate
Amp	ampicillin
AMP	adenosine monophosphate
Akt	protein kinase B
AP-3	adaptor protein 3
Asp1	diphosphoinositol pentakisphosphate kinase from *S. pombe*
Arg82	*myo*-inositol polyphosphate multikinase from *S. cerevisiae*
ATP	adenosine triphosphate
βCN	β-cyanoethyl
BIRD	bilinear rotation decoupling
BIRD-HMQC	HMQC experiment with BIRD-pulse
Bn$_2$	dibenzyl
BSTFA	*N,O*-bis(trimethylsilyl)trifluoroacetamide
cAMP	cyclic adenosine monophosphate
CID	collision-induced dissociation
CK2	casein kinase 2
CREB	cAMP response element binding protein
CrK	creatine kinase
CSA	camphorsulfonic acid
CV	column volume
DAPI	4',6-diamidin-2-phenylindol
DBU	1,8-Diazabicyclo(5.4.0)undec-7-ene
DCI	4,5-dicyanoethyl imidazole
DCM	dichloromethane

Ddp1	diphosphoinositol-polyphosphate diphosphatase from *S. cerevisiae*
DIPP1	diphosphoinositol-polyphosphate diphosphatase from *H. sapiens*
DIBAL-H	diisobutylaluminiumhydrid
DMEM	Dulbecco's Modified Eagle Medium
DMF	*N,N*-dimethylformamide
DMSO	dimethylsulfoxide
DPA	dipicolylamine
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EThcD	electron-transfer dissociation combined with higher energy collision dissociation
Fm	fluorenylmethyl
FPLC	fast protein liquid chromatography
Gen	Gentamycin
GDP	guanosine diphosphate
GLUT4	glucose transporter type 4
GPCR	G-protein coupled receptor
GTP	guanosine triphosphate
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HPLC	high performance liquid chromatography
hSAX	hydrophilic interaction strong anion exchange chromatography
IMAC	immobilized metal affinity chromatography
Ins	*myo*-inositol
$InsP_3$	*myo*-inositol 1,4,5-trisphosphate
InsP	*myo*-inositol phosphate
$InsP_5$	*myo*-inositol pentakisphosphate
$InsP_6$	*myo*-inositol hexakisphosphate
IP6K1	inositol hexakisphosphate kinase 1 from *H. sapiens* (Genbank: AC099668)
IP6KA	inositol hexakisphosphate kinase A from *E. histolytica* (Genbank: XP_648490.2)
Ipk1	inositol-pentakisphosphate-2-kinase from *S. cerevisiae*
IPMK	inositol polyphosphate multikinase from *H. sapiens*
IPPK	inositol-pentakisphosphate 2-kinase from *H. sapiens*
Ipk1	inositol-pentakisphosphate 2-kinase *S. cerevisiae*
IPS	inositol-3-phosphate synthase (WP_010879290.1)
IPTG	Isopropyl-β-D-thiogalactopyranosid
IRS1	insulin receptor substrate 1
LC-MS	liquid chromatography coupled to mass spectrometry
LIC	ligation independent cloning
Kan	Kanamycin

k_{cat}	turnover number
Kcs1	inositol hexakisphosphate kinase from *S. cerevisiae* and *S. pombe*
Kif3A	kinesin-like protein 3A
K_M	Michaelis-Menten constant
MBP	maltose binding protein
*m*CPBA	*meta*-chloroperoxybenzoic acid
MDD	metal dye detection
MES	2-(*N*-morpholino)ethanesulfonic acid
MOPS	3-morpholinopropane-1-sulfonic acid
MS	mass spectrometry
MS/MS	tandem mass spectrometry
mTOR	mammalian target of rapamycin
NADH	nicotinamide adenine dinucleotide (reduced form)
NAD^+	nicotinamide adenine dinucleotide
NMR	nuclear magnetic resonance (spectroscopy)
NTA	nitrilotriacetic acid
OD_{600}	optical density at 600 nm
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
pCr	creatine phosphate
PDK1	phosphoinositide-dependent kinase 1
PH domain	plekstrin homology domain
Pi	phosphate
PI3K	phosphoinositide kinase 3-kinase
PHR	Pi starvation response
PLC	phospholipase C
PMB	*para*-methoxybenzyl
PPIP5K	diphosphoinositol pentakisphosphate kinase from human
$PPIP5K2^{KD}$	(human) diphosphoinositol pentakisphosphate kinase 2 kinase domain (Genbank: NP_001263206.1)
$PtdInsP_2$	phosphatidyl *myo*-inositol-4,5-diphosphate
PtdInsP	phosphatidyl *myo*-inositol phosphate
PTM	posttranslational modification
rt	room temperature
SAX	strong anion exchange
SILAC	stable isotope labeling with amino acids in cell culture
SIMAC	sequential elution immobilized metal affinity chromatography
Siw14	protein tyrosine phosphatase
SMIT1	sodium-*myo*-inositol co-transporter
SPX domain	domain found in SYG1/Pho81/XPR1 proteins
sumo	small ubiquitin like modifier

TB	Terrific broth medium
TFA	trifluoroacetic acid
TLC	thin layer chromatography
TRIS	tris(hydroxymethyl)-aminomethan
V_0	initial specific velocity
VIH2	Vip1 homolog kinase
$VIH2^{KD}$	kinase domain of Vip1 homolog kinase
Vip1	diphosphoinositol pentakisphosphate kinase from *S. cerevisiae*
V_{max}	maximum specific velocity
wt	wild type
Xy	xylyl

Publications, Posters and Talks

Publications

[1] F. van der Pijl, R. K. Harmel, G. J. J. Richelle, P. Janssen, F. L. van Delft, and F. P. J. T. Rutjes, “Organocatalytic Entry into 2,6-Disubstituted Aza-Achmatowicz Piperidinones: Application to (−)-Sedacryptine and Its Epimer,” Org. Lett., vol. 16, no. 7, pp. 2038–2041, Apr. 2014.

[2] R. K. Harmel, M. M. E. Delville, and F. P. J. T. Rutjes, “Experimental procedures for conducting organic reactions in continuous flow,” in Flow Chemistry: Organic Synthesis in Motion, Band 1, DE GRUYTER, 2014, pp. 191–250.

[3] L. G. J. Hammarström, R. K. Harmel, M. Granat, R. Ringom, Y. Gravenfors, K. Färnegårdh, P. H. Svensson, D. Wennman, G. Lundin, Y. Roddis, S. S. Kitambi, A. Bernlind, F. Lehmann, P. Ernfors, “The Oncolytic Efficacy and in Vivo Pharmacokinetics of [2-(4-Chlorophenyl)quinolin-4-yl](piperidine-2-yl)methanol (Vacquinol-1) Are Governed by Distinct Stereochemical Features,” J. Med. Chem., vol. 59, no. 18, pp. 8577–8592, Sep. 2016.

[4] R. Harmel and D. Fiedler, “Features and regulation of non-enzymatic post-translational modifications,” Nat. Chem. Biol., vol. 14, p. 244, Feb. 2018.

[5] R. K. Harmel*, R. Puschmann*, M. Nguyen Trung, A. Saiardi, P. Schmieder, and D. Fiedler, “Harnessing 13 C-labeled myo -inositol to interrogate inositol phosphate messengers by NMR,” Chem. Sci., vol. 10, pp. 5267–5274, 2019.

[6] M. Penkert, A. Hauser, R. Harmel, D. Fiedler, C. P. R. Hackenberger, and E. Krause, “Electron Transfer/Higher Energy Collisional Dissociation of Doubly Charged Peptide Ions: Identification of Labile Protein Phosphorylations,” J. Am. Soc. Mass Spectrom., vol. 30, no. 9, pp. 1578–1585, 2019.

[7] J. Zhu, K. Lau, R. Puschmann, R. K. Harmel, Y. Zhang, V. Pries, P. Gaugler, L. Broger, A. K. Dutta , H. J. Jessen, G. Schaaf, A. R. Fernie, L. A. Hothorn, D. Fiedler, M. Hothorn, “Two bifunctional inositol pyrophosphate kinases/phosphatases control plant phosphate homeostasis,” Elife, vol. 8, p. e43582, Aug. 2019.

[8] R. Puschmann*, R. K. Harmel*, and D. Fiedler, “Scalable Chemoenzymatic Synthesis of Inositol Pyrophosphates,” Biochemistry, p. Article ASAP, Aug. 2019. Highlighted in ACS Editor’s Choice.

(* contributed equally)

Posters

[1] R. Harmel*, R. Puschmann*, M. Nguyen Trung, A. Saiardi, P. Schmieder, D.Fiedler. ^{13}C-Labeled myo-inositol as a tool to interrogate inositol polyphosphate messengers. Poster presentation at the EMBO conference Chemical Biology **2018**, Heidelberg, Germany

[2] R. Harmel*, R. Puschmann*, M. Nguyen Trung, A. Saiardi, P. Schmieder, D.Fiedler. ^{13}C-Labeled myo-inositol as a tool to interrogate inositol polyphosphate messengers. Poster presentation at 1st UCSF-Berlin Integrative Structural Biology Meeting (ISB Meeting) **2019**, Berlin, Germany

(* contributed equally)

Talks

[1] R. Harmel, J. A. M. Morgan, M. Penkert, E. Krause, F. Liu, D. Fiedler, Detection of Endogenous Protein Pyrophosphorylation Sites by a Tailored Proteomics Approach. 15 min talk presented on the Jungchemikerforum (JCF) **2018**, Berlin, Germany. Awarded for best talk.

[2] R. Harmel, Session host at 1st UCSF-Berlin Integrative Structural Biology Meeting (ISB Meeting) **2019**, Berlin, Germany

Chapter 1 Introduction: Signal Transduction by Inositol Pyrophosphates

Cellular Signaling of Second Messengers

The transfer of information from the exterior to the interior of a cell is pivotal to all biological processes including cell proliferation, differentiation and even cell death. Disruption of signaling pathways and incorrect processing of cellular information are responsible for many diseases including cancer, diabetes and Alzheimer's disease. A better understanding of the mechanisms underlying signal transduction can guide the development of bespoke and effective medical therapies, as well as prevention strategies.

Converting extracellular information into a specific cellular response is a very complex process. Many signals must be recognized simultaneously, and the sum of this information must be combined into an appropriate feedback. Throughout the course of evolution, organisms have developed various signal transduction strategies that broadly follow three fundamental steps: signal perception, transduction relay and response. (**Figure 1.1**) At first, extracellular stimuli, usually in the form of small molecules or proteins (ligands), are recognized by a detector on the cell surface or within a cell (receptor).

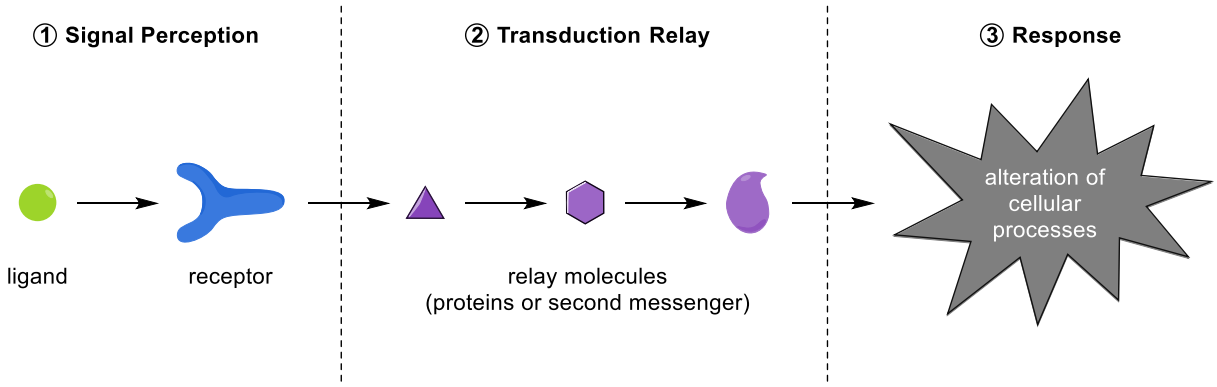


Figure 1.1: Fundamental steps of signal transduction: signal perception, transduction relay and response. Extracellular stimuli (ligands) are recognized by a receptor and the information is transferred through the cell by second messengers and/or proteins. These events lead to a cellular response, altering cellular processes.

The observed signal is then relayed by the receptor through intracellular signaling pathways *via* proteins and/or second messengers. The end of the transduction cascade results in an alteration of cellular processes which facilitates an adequate phenotypic change tailored to the initial extracellular stimulus.

An integral part of signaling is the transduction of information by second messenger molecules. In the second half of the 20th century, Earl Sutherland (Nobel Prize 1971) and colleagues have postulated the idea of a "second messenger" while investigating the function of cyclic adenosine monophosphate (cAMP).[1] Together with the pioneering work of Rodbell and Gilman (Nobel Prize 1994) on G-proteins, we now have a detailed mechanistic understanding of this process.[2] (**Figure 1.2**)

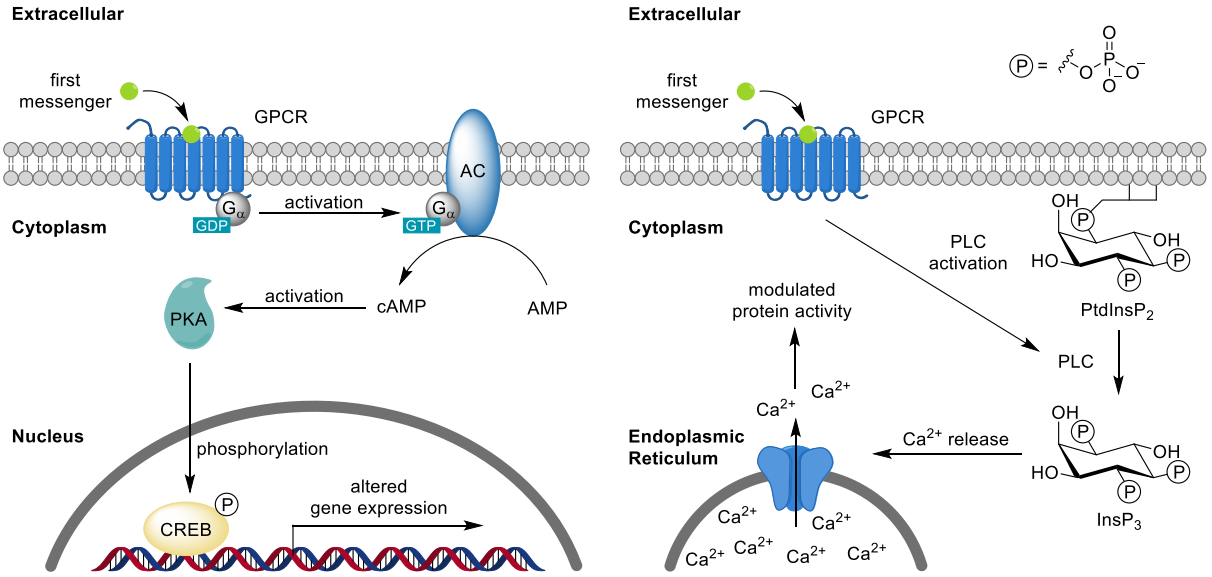


Figure 1.2: Cellular signaling of second messengers cAMP and $InsP_3$. (left) Proteome alteration *via* cAMP signaling. A first messenger is recognized by the GPCR and stimulates AC through the translocation of the G_α-protein. Activated PKA can phosphorylate CREB to alter gene expression. (right) Modulation of protein activity *via* $InsP_3$ signaling. A first messenger is recognized by the GPCR and PLC is activated in an analogous fashion as AC. $PtdInsP_2$ is cleaved to form $InsP_3$ which can interaction with the $InsP_3$-gated calcium channel. The release of Ca^{2+} from the endoplasmic reticulum leads to a modulation of protein activity inside the cell.

In their studies, they showed that neurotransmitters including dopamine and epinephrine bound to G protein-coupled receptors (GPCRs) on the cell surface and activated the production of cAMP from adenosine monophosphate (AMP) by adenyl cyclase (AC).[3–8] (**Figure 1.2 left**)The messenger cAMP then activated protein kinase A (PKA) which phosphorylated transcription factors, such as the cAMP response element (CREB) to alter gene expression. Today, it is well established that cAMP intersects with a plethora of biological functions such as the regulation of transcription factors, is associated with diseases such as cancer, and plays a wide role in metabolism.[9]

The example of cAMP shows how second messengers can directly create a response by changing the gene expression of a given cell. However, signal transduction chains also include downstream events where the function of already existing proteins is modulated. A molecule involved in such a mechanism is inositol 1,4,5-triphosphate ($InsP_3$) which participates in the regulation of cell proliferation and muscle contraction.[10] (**Figure 1.2 right**) Upon ligand binding to the GPCR, phospholipase C (PLC) activity is upregulated and cleaves phosphatidyl *myo*-inositol-4,5-bisphosphate ($PtdInsP_2$) to liberate $InsP_3$ into the cytosol. Binding of $InsP_3$ to an $InsP_3$-gated calcium channel releases Ca^{2+} from the endoplasmic reticulum into the cytosol which subsequently modulates the activity of several proteins.

In the last five decades, scientists have come to a general understanding of second messenger systems and their underlying mechanisms. Particularly, the examples of cAMP and $InsP_3$ illustrated the enormous impact that these molecules can have on cells. However, over the years, new signaling molecules were found where the underlying mechanisms are not well established, yet. Therefore, continuous work

in this field will provide future opportunities to find new targets and strategies for therapeutic intervention.

Inositol Polyphosphates and Pyrophosphates

The inositol phosphates are one of the most important and versatile classes of second messengers and are based on the *myo*-inositol scaffold. They consist of two major subclasses, characterized by the lipid-anchored phosphatidyl *myo*-inositol phosphates (PtdInsPs) and the diffusible *myo*-inositol phosphates (InsPs). Biochemically, these messengers are synthesized directly from *myo*-inositol that can either be taken up by cells through the sodium-*myo*-inositol co-transporter (SMIT1) or synthesized *de novo* from glucose.[11,12] (**Figure 1.3**) *myo*-Inositol is then incorporated into the cell membrane as PtdInsPs through a series of enzymatic transformations. PtdInsPs consist of several structurally related molecules (not shown) that fulfill functions such as synaptic vesicle trafficking and recruitment of proteins to the cell membrane, and represent an interesting group of signaling molecules in their own right.[13,14] Cleavage of $PtdInsP_2$ by PLC leads to the formation of $InsP_3$ which can be further elaborated by myo-inositol polyphosphate multikinase (*Homo sapiens*: IPMK; *Saccharomyces cerevisiae*: Arg82) and inositol-pentakisphosphate-2-kinase (*H. sapiens*: IPPK; *S. cerevisiae*: Ipk1) to the most abundant InsPs: $InsP_5$ and $InsP_6$ (10–100 μM in mammalian cells).[15] These molecules are densely negatively charged, which provides them with distinct properties that enable specific functions. For example, $InsP_6$ can strongly chelate di- and trivalent cations but can also bind positively charged amino acid side chains.[16] In the latter case, it can act as scaffolding molecule that is proposed to be involved in the assembly of the Gag hexamer and therefore the regulation of capsid stability in HIV.[17,18]

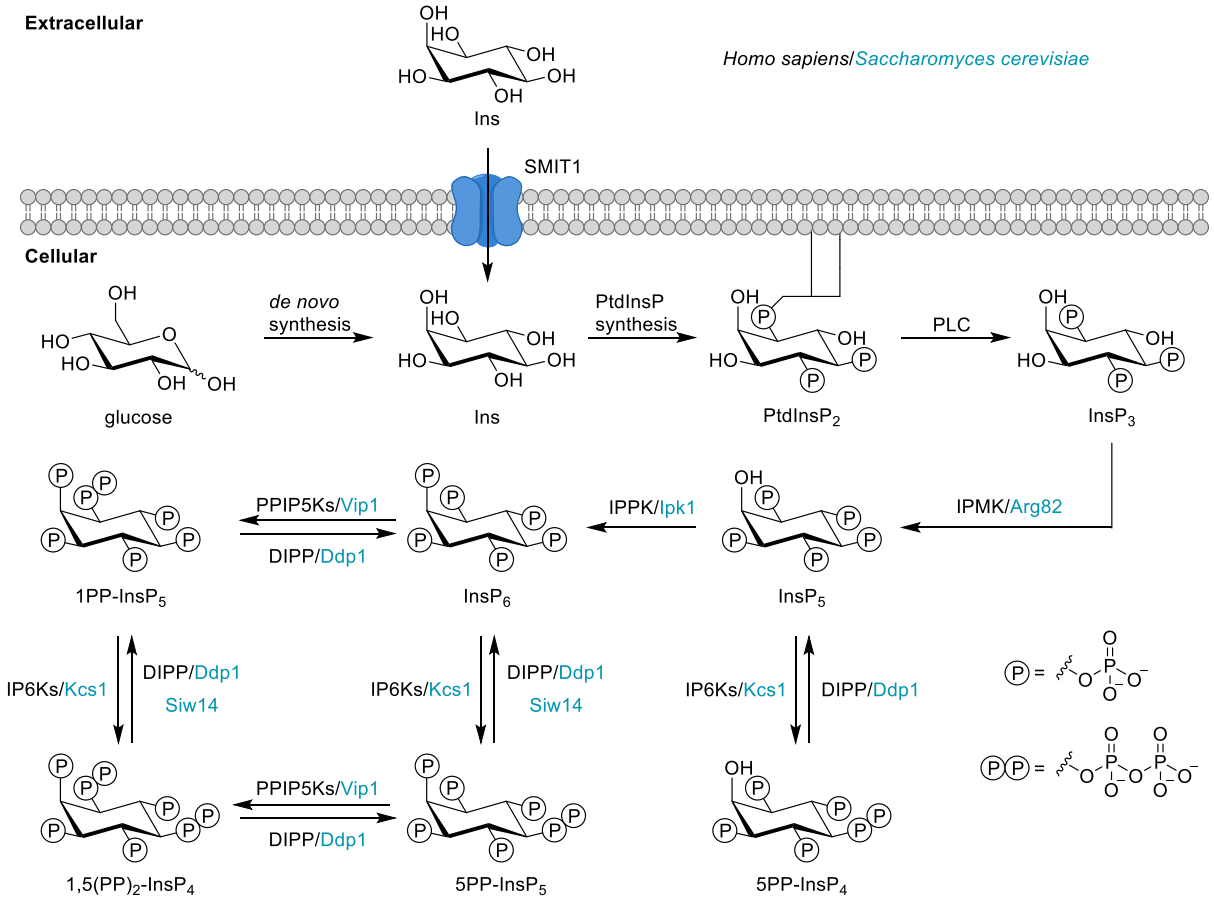


Figure 1.3: Simplified biochemical pathway of *myo*-inositol-based molecules starting from *myo*-inositol and glucose.

For some time, $InsP_6$ had been considered the highest phosphorylated InsP species but in the 1990's even higher phosphorylated substances were observed using ion exchange chromatography.[19,20] These experiments revealed a unique subclass of the InsPs which are now referred to as the inositol pyrophosphates (PP-InsPs). These molecules bear up to 8 phosphate groups that are obtained by installing one or two phosphoanhydride bonds around the *myo*-inositol ring. In mammals and yeast, two distinct classes of kinases were identified that synthesize PP-InsPs: the inositol hexakisphosphate kinases (*H. sapiens*: IP6K1/2/3; *S. cerevisiae*: Kcs1) and the diphosphoinositol-pentakisphosphate kinases (*H. sapiens*: PPIP5K1/2; *S. cerevisiae*: Vip1). IP6Ks phosphorylate $InsP_5$ and $InsP_6$ at the 5-position to form 5PP-$InsP_4$ and

5PP-InsP$_5$, respectively.[21,22] PPIP5Ks modify InsP$_6$ at the 1-position to deliver 1PP-InsP$_5$.[23,24] Finally, the combined actions of these two kinases produces 1,5(PP)$_2$-InsP$_4$ which possesses eight phosphoryl groups including two pyrophosphate moieties. Theoretically, 1,5(PP)$_2$-InsP$_4$ could be synthesized from either 1PP-InsP$_5$ or 5PP-InsP$_5$ but it is believed that the main biosynthetic pathway proceeds *via* a 5PP-InsP$_5$ intermediate as it is the most abundant PP-InsP (1–3 μM) and a much better substrate for PPIP5Ks than InsP$_6$.[15,23–26] Compared to the kinases, the PP-InsP phosphatases are less well understood. One of the best characterized representatives is diphosphoinositol-polyphosphate diphosphatase (*H. sapiens*: DIPP1–5; *S. cerevisiae*: Ddp1), a member of the nudix hydrolyses, which can dephosphorylate both the 1- as well as the 5-position.[27–32] Quite recently, protein tyrosine phosphatase (*S. cerevisiae*: Siw14) was shown to dephosphorylate PP-InsPs at the 5-position specifically.[33,34]

The initial identification and characterization of PP-InsPs immediately raised speculation about their function as signaling molecules, as the cell must invest substantial energy in the production of these metabolites. This hypothesis was based on the observation that the biosynthesis of PP-InsPs requires at least seven or eight equivalents of ATP and that a rapid metabolic turnover of PP-InsP pools had been observed.[19,20] PP-InsPs were found in yeast (*S. cerevisiae*, *Schizosaccharomyces pombe*), in the social amoeba *Dictyostelium discoideum* and in higher organisms like *Drosophila melanogaster*, *Mus musculus* and *H. sapiens* and even in plants suggesting a conserved and vital role in eukaryotic signaling.[35–38] The last two decades have amplified the interest in these messengers due to a plethora of strong phenotypic associations that were identified, specifically through genetic intervention of the PP-InsPs biosynthetic pathway. (**Table 1.1**) For example, knockouts of the murine IP6K isoforms led to defects in insulin signaling, impaired spermatogenesis, and caused

resistance to fat-induced obesity.[39–41] In mammalian cell lines, genetic perturbation of IP6Ks reduced migration and invasion of cells, and apoptosis was promoted.[42–52] In yeast, the overexpression of Kcs1 perturbed the length of telomeres and changed cell cycle progression.[53–55] The activity of PPIP5Ks has also been associated with distinct phenotypes. A PPIP5K missense mutation in humans was identified as a locus of hearing loss which could be recapitulated in PPIP5K2$^{-/-}$ deficient mice.[56] In yeast, changes of Asp1 (the PPIP5K of *S. pombe*) or Vip1 (the PPIP5K of *S. cerevisiae*) were identified to regulate the dimorphic switch.[57,58]

Table 1.1: Examples of identified phenotypes associated with IP6Ks and PPIP5Ks.

Enzyme	Organism	Effect	Reference
IP6K1	mouse	insulin signaling and sensitivity	35,40,59,60
	mouse	impaired spermatogenesis	40,41,61
	mouse, human	reduced cell migration and invasion	42,43
IP6K2	mouse, human	promoting apoptosis	44–52
	mouse	cancer metastasis and cell migration	43,62
IP6K3	mouse	life span	63
	human	susceptibility to Alzheimer's disease	64
Kcs1	yeast	regulation of rRNA transcription	65
	yeast	cell cycle progression	55
	yeast	phosphate homeostasis	66
	yeast	telomere maintenance	53,54
PPIP5K1	human	bioenergetics homeostasis	67
	human	downregulation of apoptosis	68
	human	cell migration	69
PPIP5K2	mouse, human	hearing loss	56
Vip1	yeast	regulation of dimorphic shift	57,58

Together, these discoveries have painted *myo*-Inositol-based messengers as an interesting class of signaling molecules that are involved in many cellular processes. In particular, the unique PP-InsP subclass is associated with many intriguing phenotypes, which emerge upon genetic perturbation of the biochemical pathway, and remain an exciting area of investigation today.

Towards a Better Understanding of Inositol Pyrophosphate Signaling

The development of genetic model systems and the identification of PP-InsP-specific phenotypes was greatly aided by analytical tools that could elucidate the metabolism of these messengers and monitor levels of PP-InsP. Due to the absence of a UV-absorbing group in these molecules, the main methods used to visualize PP-InsPs strongly relied on radioactive labeling with ^{3}H and ^{32}P.[19,20] *In vitro*, radioactively-labeled PP-InsP became the tool of choice for the identification of IP6Ks and PPIP5Ks in metazoan and yeast cells. Here, the purification of the enzyme could be followed by assaying the conversion of ^{3}H or ^{32}P-labeled $InsP_6$ or PP-InsPs to the corresponding PP-InsP product on polyethyleneimine-cellulose TLC plates. This method established a thorough understanding of PP-InsPs metabolism. *In vivo,* one of the major tools to analyze PP-InsPs pools was radioactive labeling of cells (yeast, *D. doscoideum,* human) with [^{3}H]*myo*-inositol or [^{32}P]orthophosphate. Upon incoporation of these radioactive labels, PP-InsPs were extracted using perchloric acid and separated by strong anion exchange (SAX) high performance liquid chromatography (HPLC) followed by their detection through scintilation counting of the collected fractions. (**Figure 1.4**). This method facilitated the discovery of these messengers and identified cellular conditions that modulate the PP-InsPs levels *in vivo*. For example, studies in

yeast cells indicated that PP-InsP levels are dependent on cell cycle progression and on phosphate availabililty.[55,70]

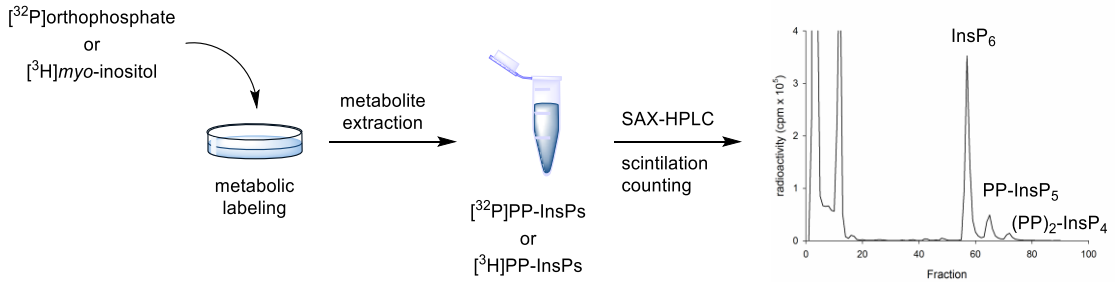


Figure 1.4: Metabolic labeling of eukaryotic cells with radioactive isotope tracers.

Genetic perturbation of PP-InsP metabolism and the analytical detection of these molecules by radioactivity had led to a good understanding of the processes PP-InsPs are involved in, but how these molecules induce their multifaceted phenotypes remained unclear in most cases. What were their underlying signaling mechanisms? A more elaborate understanding of this fundamental question required new biochemial methods that would offer a more detailed analysis of PP-InsPs signal transduction. The development of such techniques had been supported by the preparation of PP-InsPs as standards, reagents and even chemical probes. Traditionally, radioactively labeled PP-InsPs had been synthesized with the help of enzymes using short synthetic routes on a small scale.[71,72] In order to produce highly pure material on a larger scale, chemists designed multistep syntheses that are challenging but delivered low milligram quantities.[73,74] Regardless of their synthetic origin, PP-InsPs standards and reagents were provided to a wider scientific community by both methods. Unfortunately, the distribution of these small molecule messengers in sufficient amounts, especially in the case of the unsymmetrical PP-InsPs, is far from optimal and remains a major bottle neck in this field.

In the meantime, scientists that had access to these molecules started to investigate their underlying signaling mechanisms. Based on the chemical structure of these compounds it was obvious to draw a connection to the lower phosphorylated InsPs. Both classes have a high charge density, can chelate divalent metals and the pKa of the phosphoryl groups is perturbed.[16,75] InsPs like $InsP_3$, $InsP_4$ and $InsP_6$ are known to interact with proteins and modulate their activity through binding of positively charged domains.[76–78] Consequently, this mechanism of action had also been proposed for PP-InsPs. So far, only a few cases were identified where PP-InsPs specifically modulated biological processes as compared to other InsPs. A prominent example is the regulation of protein kinase B (Akt) by 5PP-$InsP_5$. (**Figure 1.5**)

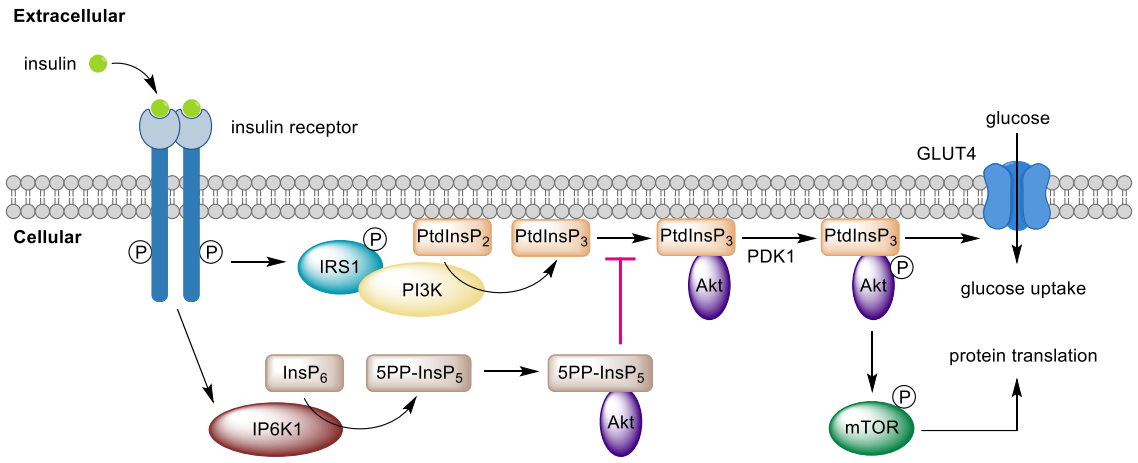


Figure 1.5: Insulin/Akt signaling pathway is regulated *via* inositol pyrophosphates. Akt is recruited to the membrane upon insulin signaling by $PtdInsP_3$ and activates GLUT4 an mTOR which leads to glucose uptake and protein translation. Simultaneously, the insulin receptor stimulates the production of 5PP-$InsP_5$ which competes with $PtdInsP_3$ for the PH domain of Akt and inhibits its translocation to the membrane. (Figure adapted from 'Insulin Signaling: Inositol Phosphates Get Into the Akt'.[79]).

After the observation of insulin hypersensitivity in $IP6K1^{-/-}$ mice, the underlying mechanism was investigated.[59] Canonically, upon binding of insulin to its receptor, insulin receptor substrate 1 (IRS1) activates phosphoinositide 3-kinase (PI3K). $PtdInsP_2$ is phosphorylated at the 3-position by PI3K to form $PtdInsP_2$ which recruits

Akt and phosphoinositide-dependent kinase-1 (PDK1) to the cell membrane by binding to their plekstrin homology (PH) domain. Here, Akt is activated by phosphorylation through PDK1 and stimulates downstream protein translation *via* mammalian target of rapamycin (mTOR) signaling and supports glucose uptake *via* glucose transporter type 4 (GLUT4). 5PP-InsP$_5$ is a physiological regulator of this process and its biosynthesis is elevated through insulin. It acts by binding to the PH domain of Akt and prevents its translocation to the membrane. In IP6K1$^{-/-}$ mice, Akt activity cannot be perturbed by 5PP-InsP$_5$ leading to insulin hypersensitivity. Conversely, increased stimulation of 5PP-InsP$_5$ synthesis might lead to reduced activity of Akt and therefore insulin resistance. A similar effect was obtained upon delivery and subsequent photorelease of exogenous 5PP-InsP$_5$ into cells.[80] Beside the PH domain, PP-InsP are known to bind to C2B domains and SPX domains suggesting a role in exocytosis and phosphate homeostasis.[77,81,82]

In an effort to identify more binding interactions between PP-InsPs and proteins, an unbiased proteomics pull-down approach using a non-hydrolyzable analog of 5PP-InsP$_5$ identified more than 100 proteins in yeast that are selectively enriched over immobilized phosphate.[83] Interestingly, most of these proteins had an affinity for InsP$_6$ and only in a few cases, selectivity for 5PP-InsP$_5$ was observed. In this study, two distinct sets of proteins were found to bind 5PP-InsP$_5$ either in the absence or presence of Mg^{2+} ions with only little overlap in between. This observation indicated the presence of two different sets of interacting partners – and potentially signaling mechanisms – where one of these was essentially dependent on divalent metal cations.

Indeed, besides binding to proteins in a manner comparable to the lower InsPs, a second mechanism of action was proposed for PP-InsPs that utilizes their unique phosphoanhydride bond. Interestingly, it was found that cell invest a lot of energy into

the regeneration of this moiety, resulting in up to 50% turnover of $InsP_6$ into PP-InsPs every hour.[19,84] Moreover, the energy of hydrolysis of the pyrophosphate bond in 5PP-$InsP_5$ and 5,6(PP)$_2$-$InsP_4$ was predicted to be in between the values of ATP (–30.5 kJ/mol) and ADP (–26.7 kJ/mol) or even higher.[85] These observations led to the hypothesis that PP-InsPs might be able to transfer their β-phosphoryl group onto a given substrate.

This proposed mechanism had been investigated by Snyder and coworkers in the mid 2000's. Using 5PP-$InsP_5$-β^{32}P, they detected *in vitro* phosphorylation of proteins in cell lysate samples from yeast, fruit flies, and mice but not in bacterial extracts.[86] (**Figure 1.6**) In this report, it was shown that the phosphoryl transfer is likely to occur *via* an enzyme-independent mechanism, requiring only Mg^{2+} or Ca^{2+} as a co-factor. In a follow-up article, the phosphorylation of proteins by PP-InsPs was found to be resistant to some phosphatases, like λ-phosphatase, and was localized at polyacidic stretches that bear a consensus motive for casein kinase 2 (CK2).[87]

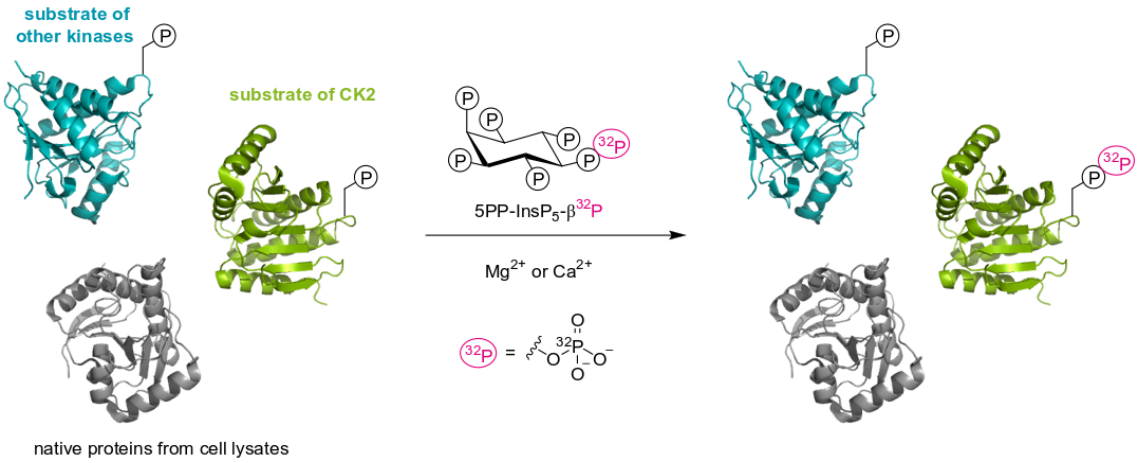


Figure 1.6: *In vitro* pyrophosphorylation of proteins from cell lysates using 5PP-$InsP_5$-β^{32}P. Proteins that are substrates of CK2 become pyrophosphorylated while non-CK2 site and unphosphorylated proteins remain unmodified.

Finally, more detailed analysis of the modified sites showed that the phosphoryl transfer was only possible on CK2-phosphorylated serine residues. Together with the previous findings, this new posttranslational modification (PTM) became known as protein pyrophosphorylation. Besides mechanistic aspects, the work from the Snyder laboratory also identified several *in vitro* pyrophosphorylated proteins from yeast (NSR1, SRP40, YGR130c, APL6) and mammals (Nopp140, AP3β3A, and TCOF1). These proteins are associated with essential cellular processes such as ribosome biogenesis, nucleocytoplasmic transport and regulation of translation.

Since the discovery of protein pyrophosphorylation, there had been various efforts to show the significance of this modification *in vivo*. This was achieved in an indirect fashion by comparing the *in vitro* pyrophosphorylation of proteins from lysates with different genetics backgrounds: wild type and IP6K$^{-/-}$. (**Figure 1.7**)

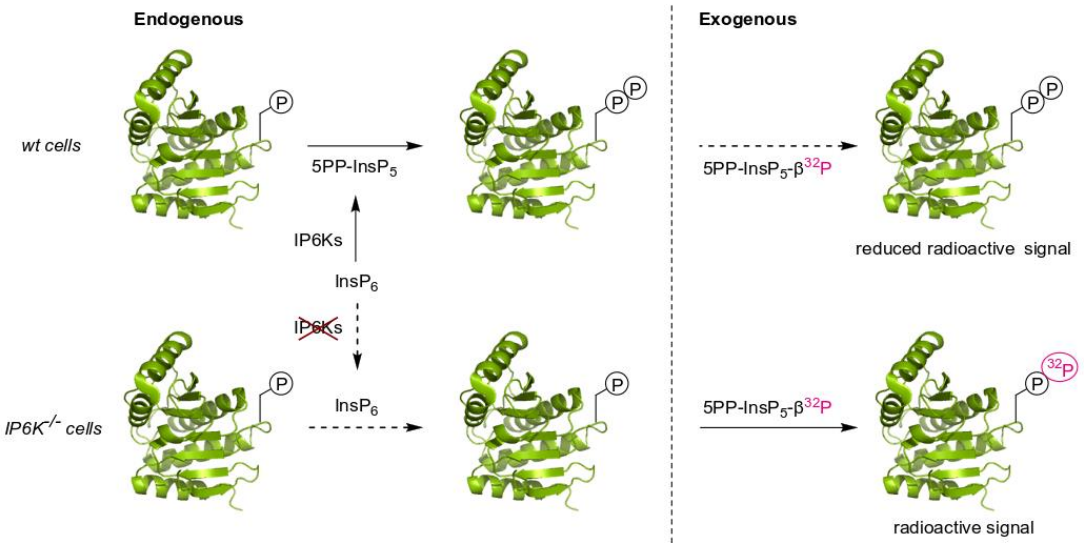


Figure 1.7: Detection of potential endogenous pyrophosphorylation sites by the 'back-pyrophosphorylation assay'. In wild type cells, 5PP-InsP$_5$ pyrophosphorylates proteins thereby reducing pyrophosphorylation by 5PP-InsP5-β^{32}P. In IP6K$^{-/-}$ cells, 5PP-InsP$_5$ is not present which maximizes the signal intensity induced by the pyrophosphorylation with 5PP-InsP5-β^{32}P. Radioactive proteins are usually analyzed by PAGE.

The so-called ‘back-pyrophosphorylation assay’ was based on the assumption that IP6K$^{-/-}$ cell lines were unable to produce 5PP-InsP$_5$ which should significantly lower the pyrophosphorylation of proteins, allowing more *in vitro* phosphoryl group transfer to subsequently occur. Consequently, *in vitro* pyrophosphorylation by 5PP-InsP$_5$-β^{32}P should create a higher radioactive signal compared to the wild type sample. Using this method, it was demonstrated that protein pyrophosphorylation can modulate protein-protein interactions. Specifically, pyrophosphorylation prevented the interaction of the adaptor protein 3 (AP-3) complex with kinesin-like protein 3A (Kif3A), thus modulating its ability to regulate the HIV-1 gag release.[88] On the other hand, pyrophosphorylation of dynein enhanced its interaction with dynactin, thereby regulating cytoplasmic dynein-driven transport.[89] Interestingly, protein pyrophosphorylation was also identified as a regulator of RNA transcription.[65] These results were not obtained by the back-pyrophosphorylation assay but by investigating a phenotype in kcs1 Δ cell lines. Transcription by RNA polymerase 1 was compromised in these cells and three subunits of this enzyme were identified as substrates for *in vitro* pyrophosphorylation. Taken together, all of these finding slowly solidified pyrophosphoserine residues as signaling entities but up until now endogenous protein pyrophosphorylation has never been detected directly. Therefore, some scientists in the field raised reasonable concerns over the current methods, and the overall impact of this modification *in vivo* remains controversial.[90,91]

The field of PP-InsP research has come from correlating PP-InsP levels with different genetic backgrounds to a more mechanistic understanding of how PP-InsP fulfil their multifaceted functions. This transition was possible through the development of more elaborate biochemical techniques that disentangle different signaling aspects of PP-InsPs such as protein binding or protein pyrophosphorylation. A deeper

understanding of these processes will enable future scientists to identify new checkpoints in signal transduction pathways and potentially lead to the discovery of new targets for pharmacological intervention.

Conclusions

Second messenger-based signaling is a very important process that plays a central role in many cellular functions as well as in severe diseases. Among these messengers, the PP-InsPs family is associated with many intriuging phenotypes suggesting a significant impact within signaling cascades. In the last 25 years, scientists have pursued a deeper understanding of the underlying mechanisms of these small molecules and have proposed binding interactions with proteins and protein pyrophosphorylation as the two major mechanisms of action. So far, many aspects of the binding-related modulation of proteins by PP-InsPs remain unclear and many proteins that were identified as potential binders have not been fully characterized yet. In the case of protein pyrophosphorylation, we are even further away from a thorough understanding. Neither the existence nor the biological significance of this PTM *in vivo* has been proven so far and the mechanisms that regulates temporal and spacial distribution or dynamics are unknown. To enable a broader scientific community to engage with PP-InsP research and to accelerate this field, widely available and sophisticated methods are necessary to study new aspects of PP-InsPs signaling *in vitro* and especially *in vivo*. This should be accompanied by the replacement of the traditional methods that are based on radioactivity. Even though radiolabeling is one of the most sensitive techniques, *in vivo* applications are limited and combinations with modern high resolution techniques like nuclear magnetic resonance (NMR) spectroscopy, mass spectrometry (MS) or cryo electron microscopy

are not possible. Finally, the development of novel methods and their usage will be highly dependent on the availability of the PP-InsPs molecules, so novel procedures to produce large amounts of these messengers are desperately needed.

References

1. Kresge, N., Simoni, R. D. & Hill, R. L. Earl W. Sutherland's Discovery of Cyclic Adenine Monophosphate and the Second Messenger System. *J. Biol. Chem.* **280**, e39–e39 (2005).
2. Sunahara, R. K. & Insel, P. A. The Molecular Pharmacology of G Protein Signaling Then and Now: A Tribute to Alfred G. Gilman. *Mol. Pharmacol.* **89**, 585–592 (2016).
3. Sutherland, E. W., Rall, T. W. & Menon, T. Adenyl Cyclase: I. DISTRIBUTION, PREPARATION, AND PROPERTIES . *J. Biol. Chem.* **237**, 1220–1227 (1962).
4. Rall, T. W. & Sutherland, E. W. FORMATION OF A CYCLIC ADENINE RIBONUCLEOTIDE BY TISSUE PARTICLES. *J. Biol. Chem.* **232**, 1065–1076 (1958).
5. Sutherland, E. W. & Rall, T. W. Fractionation and Characterization of a Cyclic Adenine Ribonucleotide Formed by Tissue Particles. *J. Biol. Chem.* **232**, 1077–1091 (1958).
6. Butcher, R. W. & Sutherland, E. W. Adenosine 3',5'-Phosphate in Biological Materials. *J. Biol. Chem.* **237**, 1244–1251 (1962).
7. Ross, E. M., Maguire, M. E., Sturgill, T. W., Biltonen, R. L. & Gilman, A. G. Relationship between the beta-adrenergic receptor and adenylate cyclase. *J. Biol. Chem.* **252**, 5761–5775 (1977).
8. May, D. C., Ross, E. M., Gilman, A. G. & Smigel, M. D. Reconstitution of catecholamine-stimulated adenylate cyclase activity using three purified proteins. *J. Biol. Chem.* **260**, 15829–15833 (1985).
9. Yan, K., Gao, L.-N., Cui, Y.-L., Zhang, Y. & Zhou, X. The cyclic AMP signaling pathway: Exploring targets for successful drug discovery (Review). *Mol. Med. Rep.* **13**, 3715–3723 (2016).
10. Berridge, M. J. The Inositol Trisphosphate/Calcium Signaling Pathway in Health and Disease. *Physiol. Rev.* **96**, 1261–1296 (2016).
11. Balla, T. Phosphoinositides: Tiny Lipids With Giant Impact on Cell Regulation. *Physiol. Rev.* **93**, 1019–1137 (2013).
12. Bevilacqua, A. & Bizzarri, and M. Review Article Inositols in Insulin Signaling and Glucose Metabolism. *Int. J. Endocrinol.* **2018**, (2018).
13. JAMES, S. R. *et al.* Specific binding of the Akt-1 protein kinase to phosphatidylinositol 3,4,5-trisphosphate without subsequent activation. *Biochem. J.* **315**, 709–713 (1996).
14. Leitner, M. G., Halaszovich, C. R., Ivanova, O. & Oliver, D. Phosphoinositide dynamics in the postsynaptic membrane compartment: Mechanisms and experimental approach. *Eur. J. Cell Biol.* **94**, 401–414 (2015).
15. Wundenberg, T. & Mayr, G. W. Synthesis and biological actions of diphosphoinositol phosphates (inositol pyrophosphates), regulators of cell homeostasis. **393**, 979–998 (2012).
16. Veiga, N. *et al.* The behaviour of myo-inositol hexakisphosphate in the presence of magnesium(II) and calcium(II): Protein-free soluble InsP6 is limited to 49 micro M under cytosolic/nuclear conditions. *J. Inorg. Biochem.* **100**, 1800–1810 (2006).
17. Dick, R. A. *et al.* Inositol phosphates are assembly co-factors for HIV-1. *Nature* **560**, 509–512 (2018).
18. Mallery, D. L. *et al.* IP6 is an HIV pocket factor that prevents capsid collapse and promotes DNA synthesis. *Elife* **7**, e35335 (2018).
19. Menniti, F. S., Miller, R. N., Putney Jr., J. W. & Shears, S. B. Turnover of inositol polyphosphate pyrophosphates in pancreatoma cells. *J Biol Chem* **268**, 3850–3856 (1993).
20. Stephens, L. *et al.* The Detection, Purification, Structural Characterization, andMetabolism of DiphosphoinositoPl entakisphosphate(s) andBisdiphosphoinositol Tetrakisphosphate(s. *J. Biol. Chem.* **268**, 4009–4015 (1993).
21. Voglmaier, S. M. *et al.* Diphosphoinositol pentakisphosphate as a high-energy phosphate donor Purified inositol hexakisphosphate kinase is an ATP synthase : **93**, 4305–4310 (1996).
22. Saiardi, A., Caffrey, J. J., Snyder, S. H. & Shears, S. B. The Inositol Hexakisphosphate Kinase

Family: CATALYTIC FLEXIBILITY AND FUNCTION IN YEAST VACUOLE BIOGENESIS . *J. Biol. Chem.* **275**, 24686–24692 (2000).

23. Fridy, P. C., Otto, J. C., Dollins, D. E. & York, J. D. Cloning and Characterization of Two Human VIP1 -like Inositol Hexakisphosphate and Diphosphoinositol Pentakisphosphate Kinases *. **282**, 30754–30762 (2007).
24. Mulugu, S. *et al.* A conserved family of enzymes that phosphorylate inositol hexakisphosphate. *Science* **316**, 106–109 (2007).
25. Albert, C. *et al.* Biological variability in the structures of diphosphoinositol polyphosphates in Dictyostelium discoideum and mammalian cells. **560**, 553–560 (1997).
26. Barker, C. J., Wright, J., Hughes, P. J., Kirk, C. J. & Michell, R. H. Complex changes in cellular inositol phosphate complement accompany transit through the cell cycle. *Biochem. J.* **380**, 465–473 (2004).
27. Kilari, R. S., Weaver, J. D., Shears, S. B. & Safrany, S. T. Understanding inositol pyrophosphate metabolism and function: Kinetic characterization of the DIPPs. *FEBS Lett.* **587**, 3464–3470 (2013).
28. Safrany, S. T. *et al.* The Diadenosine Hexaphosphate Hydrolases fromSchizosaccharomyces pombe andSaccharomyces cerevisiae Are Homologues of the Human Diphosphoinositol Polyphosphate Phosphohydrolase: OVERLAPPING SUBSTRATE SPECIFICITIES IN A MutT-TYPE PROTEIN . *J. Biol. Chem.* **274**, 21735–21740 (1999).
29. Leslie, N. R., McLennan, A. G. & Safrany, S. T. Cloning and characterisation of hAps1 and hAps2, human diadenosine polyphosphate-metabolising Nudix hydrolases. *BMC Biochem.* **3**, 20–20 (2002).
30. Safrany, S. T. *et al.* A novel context for the 'MutT' module, a guardian of cell integrity, in a diphosphoinositol polyphosphate phosphohydrolase. *EMBO J* **17**, 6599–6607 (1998).
31. Hidaka, K. *et al.* An adjacent pair of human NUDT genes on chromosome X are preferentially expressed in testis and encode two new isoforms of diphosphoinositol polyphosphate phosphohydrolase. *J. Biol. Chem.* **277**, 32730–32738 (2002).
32. Caffrey, J. J., Safrany, S. T., Yang, X. & Shears, S. B. Discovery of molecular and catalytic diversity among human diphosphoinositol-polyphosphate phosphohydrolases. An expanding NUDT family. *J. Biol. Chem.* **275**, 12730–12736 (2000).
33. Wang, H., Gu, C., Rolfes, R. J., Jessen, H. J. & Shears, S. B. Structural and biochemical characterization of Siw14: a protein-tyrosine phosphatase fold that metabolizes inositol pyrophosphates. *J. Biol. Chem.* **293**, 6905–6914 (2018).
34. Steidle, E. A. *et al.* A novel inositol pyrophosphate phosphatase in Saccharomyces cerevisiae: Siw14 selectively cleaves the β-phosphate from 5-diphosphoinositol pentakisphosphate (5PP-IP5). *J. Biol. Chem.* **291**, 6772–6783 (2016).
35. Chakraborty, A., Kim, S. & Snyder, S. H. Inositol pyrophosphates as mammalian cell signals. *Sci. Signal.* **4**, 1–11 (2011).
36. Wilson, M. S. C., Livermore, T. M. & Saiardi, A. Inositol pyrophosphates: between signalling and metabolism. *Biochem. J.* **452**, 369–79 (2013).
37. Laha, D. *et al.* VIH2 Regulates the Synthesis of Inositol Pyrophosphate InsP8 and Jasmonate-Dependent Defenses in Arabidopsis. *Plant Cell* **27**, 1082–1097 (2015).
38. Wilson, M. S. C., Bulley, S. J., Pisani, F., Irvine, R. F. & Saiardi, A. A novel method for the purification of inositol phosphates from biological samples reveals that no phytate is present in human plasma or urine. *Open Biol.* **5**, 150014 (2015).
39. Chakraborty, A. *et al.* Inositol pyrophosphates inhibit Akt signaling, regulate insulin sensitivity and weight gain. *Cell* **143**, 897–910 (2010).
40. Bhandari, R., Juluri, K. R., Resnick, A. C. & Snyder, S. H. Gene deletion of inositol hexakisphosphate kinase 1 reveals inositol pyrophosphate regulation of insulin secretion, growth, and spermiogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 2349–2353 (2008).
41. Malla, A. B. & Bhandari, R. IP6K1 is essential for chromatoid body formation and temporal regulation of Tnp2 and Prm2 expression in mouse spermatids. *J. Cell Sci.* **130**, 2854–2866 (2017).

42. Jadav, R. S. *et al.* Deletion of inositol hexakisphosphate kinase 1 (IP6K1) reduces cell migration and invasion, conferring protection from aerodigestive tract carcinoma in mice. *Cell. Signal.* **28**, 1124–1136 (2016).

43. Fu, C. *et al.* Neuronal migration is mediated by inositol hexakisphosphate kinase 1 via α-actinin and focal adhesion kinase. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 2036–2041 (2017).

44. Moriya, Y. *et al.* Inositol Hexakisphosphate Kinase 2 is a Presymptomatic Biomarker for Amyotrophic Lateral Sclerosis. *Tokai J. Exp. Clin. Med.* **42**, 13–18 (2017).

45. Nagata, E. *et al.* Inositol hexakisphosphate kinases induce cell death in Huntington disease. *J. Biol. Chem.* **286**, 26680–6 (2011).

46. Nagata, E. *et al.* Inositol hexakisphosphate kinase-2, a physiologic mediator of cell death. *J. Biol. Chem.* **280**, 1634–40 (2005).

47. Nagata, E. *et al.* Inositol Hexakisphosphate Kinase 2 Promotes Cell Death in Cells with Cytoplasmic TDP-43 Aggregation. *Mol. Neurobiol.* **53**, 5377–5383 (2016).

48. Koldobskiy, M. A. *et al.* p53-mediated apoptosis requires inositol hexakisphosphate kinase-2. *Proc Natl Acad Sci U S A* **107**, 20947–20951 (2010).

49. Morrison, B. H., Bauer, J. A., Kalvakolanu, D. V & Lindner, D. J. Inositol hexakisphosphate kinase 2 mediates growth suppressive and apoptotic effects of interferon-beta in ovarian carcinoma cells. *J. Biol. Chem.* **276**, 24965–24970 (2001).

50. Morrison, B. H., Tang, Z., Jacobs, B. S., Bauer, J. A. & Lindner, D. J. Apo2L/TRAIL induction and nuclear translocation of inositol hexakisphosphate kinase 2 during IFN-beta -induced apoptosis in ovarian carcinoma. *Biochem. J.* **385**, 595–603 (2005).

51. Morrison, B. H. *et al.* Gene deletion of inositol hexakisphosphate kinase 2 predisposes to aerodigestive tract carcinoma. *Oncogene* **28**, 2383–2392 (2009).

52. Morrison, B. H. *et al.* Inositol hexakisphosphate kinase 2 sensitizes ovarian carcinoma cells to multiple cancer therapeutics. *Oncogene* **21**, 1882–1889 (2002).

53. Banfic, H. *et al.* Inositol pyrophosphates modulate cell cycle independently of alteration in telomere length. *Adv. Biol. Regul.* **60**, 22–28 (2016).

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

55. Banfic, H., Bedalov, A., York, J. D. & Visnjic, D. Inositol pyrophosphates modulate S phase progression after pheromone-induced arrest in saccharomyces cerevisiae. *J. Biol. Chem.* **288**, 1717–1725 (2013).

56. Yousaf, R. *et al.* Mutations in Diphosphoinositol-Pentakisphosphate Kinase PPIP5K2 are associated with hearing loss in human and mouse. *PLoS Genet.* **14**, e1007297–e1007297 (2018).

57. Pohlmann, J. & Fleig, U. Asp1, a Conserved 1/3 Inositol Polyphosphate Kinase, Regulates the Dimorphic Switch in Schizosaccharomyces pombe. *Mol. Cell. Biol.* **30**, 4535–4547 (2010).

58. Pöhlmann, J. *et al.* The Vip1 Inositol Polyphosphate Kinase Family Regulates Polarized Growth and Modulates the Microtubule Cytoskeleton in Fungi. *PLoS Genet.* **10**, e1004586 (2014).

59. Chakraborty, A. *et al.* Inositol Pyrophosphates Inhibit Akt Signaling , Thereby Regulating Insulin Sensitivity and Weight Gain. *Cell* **143**, 897–910 (2010).

60. Zhu, Q. *et al.* Adipocyte-specific deletion of Ip6k1 reduces diet-induced obesity by enhancing AMPK-mediated thermogenesis. *J. Clin. Invest.* **126**, 4273–4288 (2016).

61. Fu, C. *et al.* Multiple aspects of male germ cell development and interactions with Sertoli cells require inositol hexakisphosphate kinase-1. *Sci. Rep.* **8**, 7039 (2018).

62. Rao, F. *et al.* Inositol pyrophosphates promote tumor growth and metastasis by antagonizing liver kinase B1. *Proc. Natl. Acad. Sci.* **112**, 1773 LP – 1778 (2015).

63. Moritoh, Y. *et al.* Inositol Hexakisphosphate Kinase 3 Regulates Metabolism and Lifespan in Mice. *Sci. Rep.* **6**, 32072 (2016).

64. Crocco, P. *et al.* Contribution of polymorphic variation of inositol hexakisphosphate kinase 3

(IP6K3) gene promoter to the susceptibility to late onset Alzheimer's disease. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1862**, 1766–1773 (2016).

65. Thota, S. G., Unnikannan, C. P., Thampatty, S. R., Manorama, R. & Bhandari, R. Inositol pyrophosphates regulate RNA polymerase I-mediated rRNA transcription in Saccharomyces cerevisiae. *Biochem. J.* **466**, 105–114 (2015).

66. Auesukaree, C., Tochio, H., Shirakawa, M., Kaneko, Y. & Harashima, S. Plc1p, Arg82p, and Kcs1p, Enzymes Involved in Inositol Pyrophosphate Synthesis, Are Essential for Phosphate Regulation and Polyphosphate Accumulation in Saccharomyces cerevisiae. *J. Biol. Chem.* **280**, 25127–25133 (2005).

67. Gu, C. *et al.* KO of 5-InsP(7) kinase activity transforms the HCT116 colon cancer cell line into a hypermetabolic, growth-inhibited phenotype. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 11968–11973 (2017).

68. Machkalyan, G., Hèbert, T. E. & Miller, G. J. PPIP5K1 Suppresses Etoposide-triggered Apoptosis. *J. Mol. Signal.* **11**, 4 (2016).

69. Machkalyan, G., Trieu, P., Pétrin, D., Hébert, T. E. & Miller, G. J. PPIP5K1 interacts with the exocyst complex through a C-terminal intrinsically disordered domain and regulates cell motility. *Cell. Signal.* **28**, 401–411 (2016).

70. Gu, C. *et al.* The Significance of the Bifunctional Kinase/Phosphatase Activities of Diphosphoinositol Pentakisphosphate Kinases (PPIP5Ks) for Coupling Inositol Pyrophosphate Cell Signaling to Cellular Phosphate Homeostasis. *J. Biol. Chem.* **292**, 4544–4555 (2017).

71. Loss, O., Azevedo, C., Szijgyarto, Z., Bosch, D. & Saiardi, A. Preparation of quality inositol pyrophosphates. *J. Vis. Exp.* **2**, e3027 (2011).

72. Wilson, M. S. C. & Saiardi, A. Importance of Radioactive Labelling to Elucidate Inositol Polyphosphate Signalling. *Top. Curr. Chem.* **375**, 14 (2017).

73. Capolicchio, S., Wang, H., Thakor, D. T., Shears, S. B. & Jessen, H. J. Synthesis of Densely Phosphorylated Bis-1,5-Diphospho-myo-Inositol Tetrakisphosphate and its Enantiomer by Bidirectional P-Anhydride Formation. *Angew. Chemie Int. Ed.* **53**, 9508–9511 (2014).

74. Capolicchio, S., Thakor, D. T., Linden, A. & Jessen, H. J. Synthesis of unsymmetric diphospho-inositol polyphosphates. *Angew. Chem. Int. Ed. Engl.* **52**, 6912–6 (2013).

75. Hager, A. *et al.* Cellular Cations Control Conformational Switching of Inositol Pyrophosphate Analogues. *Chem. – A Eur. J.* **22**, 12406–12414 (2016).

76. Seo, M.-D. *et al.* Structural and functional conservation of key domains in InsP3 and ryanodine receptors. *Nature* **483**, 108–112 (2012).

77. Wild, R. *et al.* Control of eukaryotic phosphate homeostasis by inositol polyphosphate sensor domains. *Science* **352**, 986–990 (2016).

78. Watson, P. J., Fairall, L., Santos, G. M. & Schwabe, J. W. R. Structure of HDAC3 bound to co-repressor and inositol tetraphosphate. *Nature* **481**, 335–340 (2012).

79. Manning, B. D. Insulin signaling: inositol phosphates get into the Akt. *Cell* **143**, 861–863 (2010).

80. Pavlovic, I. *et al.* Cellular delivery and photochemical release of a caged inositol-pyrophosphate induces PH-domain translocation in cellulo. *Nat. Commun.* **7**, 10622 (2016).

81. Lee, T.-S. *et al.* Inositol pyrophosphates inhibit synaptotagmin-dependent exocytosis. *Proc. Natl. Acad. Sci.* **113**, 8314–8319 (2016).

82. Gerasimaite, R. *et al.* Inositol Pyrophosphate Specificity of the SPX-Dependent Polyphosphate Polymerase VTC. *ACS Chem. Biol.* **12**, 648–653 (2017).

83. Wu, M., Chong, L. S., Perlman, D. H., Resnick, A. C. & Fiedler, D. Inositol polyphosphates intersect with signaling and metabolic networks via two distinct mechanisms. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E6757–E6765 (2016).

84. Stephens, L. *et al.* The detection, purification, structural characterization, and metabolism of diphosphoinositol pentakisphosphate(s) and bisdiphosphoinositol tetrakisphosphate(s). *J Biol Chem* **268**, 4009–4015 (1993).

85. Hand, C. E. & Honek, J. F. Phosphate transfer from inositol pyrophosphates InsP5PP and

InsP4(PP)2: A semi-empirical investigation. *Bioorg. Med. Chem. Lett.* **17**, 183–188 (2007).

86. Saiardi, A., Bhandari, R., Resnick, A. C., Snowman, A. M. & Snyder, S. H. Phosphorylation of Proteins by Inositol Pyrophosphates. *Science* **306**, 2101–2105 (2004).

87. Bhandari, R. *et al.* Protein pyrophosphorylation by inositol pyrophosphates is a posttranslational event. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 15305–10 (2007).

88. Azevedo, C., Burton, A., Ruiz-Mateos, E., Marsh, M. & Saiardi, A. Inositol pyrophosphate mediated pyrophosphorylation of AP3B1 regulates HIV-1 Gag release. *Proc. Natl. Acad. Sci.* **106**, 21161–21166 (2009).

89. Chanduri, M. *et al.* Inositol hexakisphosphate kinase 1 (IP6K1) activity is required for cytoplasmic dynein-driven transport. *Biochem. J.* **473**, 3031 LP – 3047 (2016).

90. Azevedo, C. & Saiardi, A. Reply to Shears: As knowledge of inositol pyrophosphates advances, wonder recedes. *Proc. Natl. Acad. Sci.* **107**, E18 LP-E18 (2010).

91. Shears, S. The long-awaited demonstration of protein pyrophosphorylation by IP7 in vivo? *Proc. Natl. Acad. Sci. U. S. A.* **107**, E17–E18 (2010).

Chapter 2 Scalable Chemoenzymatic Synthesis of Inositol Pyrophosphates

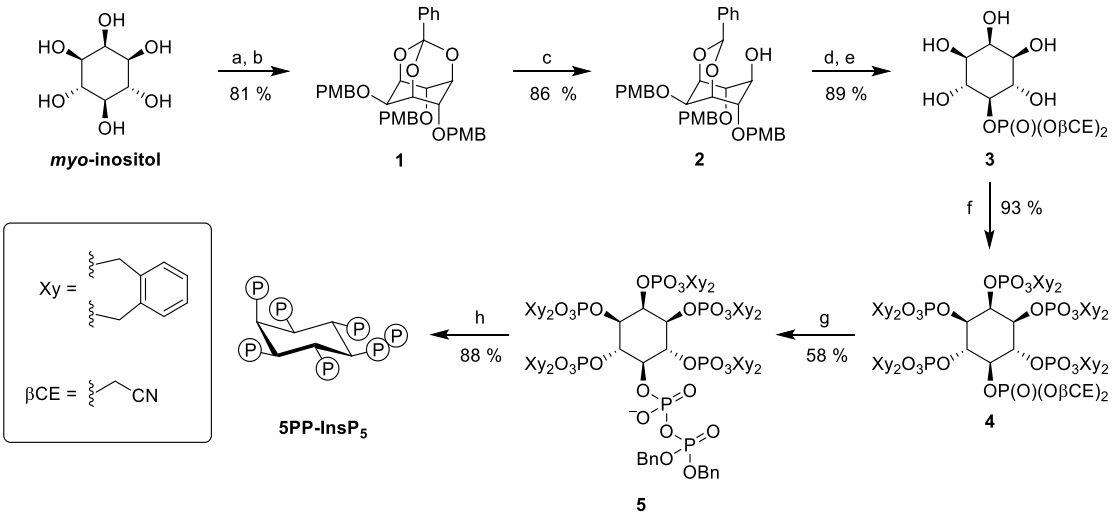
Reproduced in part with permission from (1) R. Puschmann*, R. K. Harmel*, D. Fiedler, *Biochem.* **2019,** Article ASAP. (2) R. K. Harmel*, R. Puschmann*, M. Nguyen Trung, A. Saiardi, P. Schmieder, D. Fiedler, *Chem. Sci.* **2019,** 10, 5267–5274.

(* contributed equally)

Project Collaborators: R. Puschmann performed the cloning of the IP6KA and PPI5K2 kinases into pET vectors and the optimization of the protein expression. Protein expression and synthesis of $InsP_6$ were conducted by both R. K. Harmel and R. Puschmann. PAGE gel analysis was performed by R. Puschmann. Preliminary optimization experiments on the 5PP-$InsP_5$ synthesis were performed by M. Nguyen Trung under the supervision of R. K. Harmel. R. Puschmann optimized the synthesis of 5PP-$InsP_5$-$\beta^{32}P$.

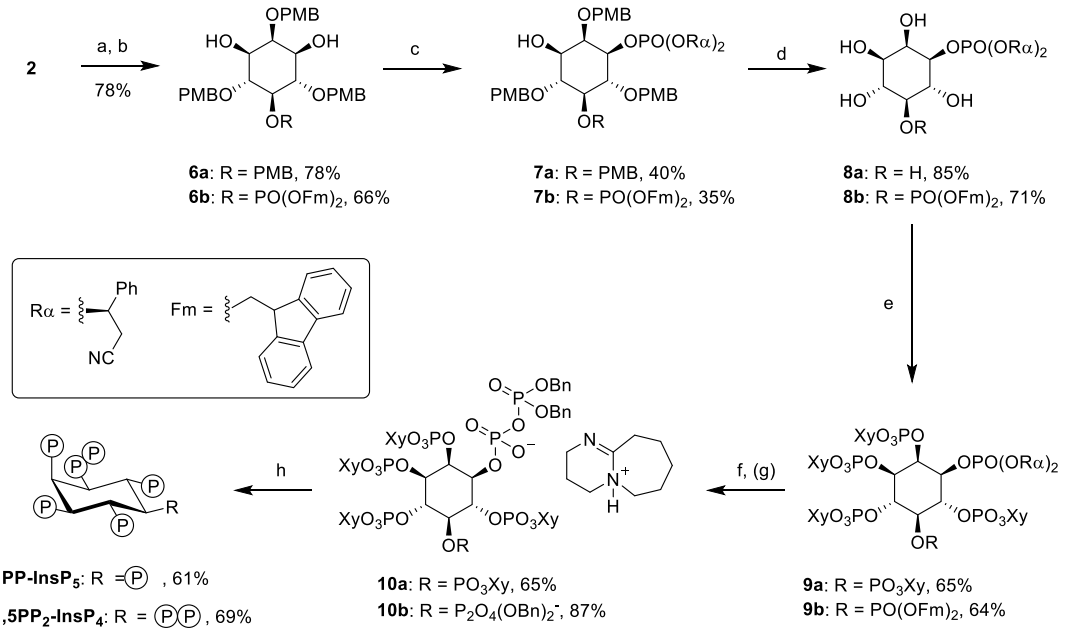
Introduction: Design of a Scalable Synthesis of PP-InsPs

The synthesis of PP-InsPs has been pursued using two main approaches: classical organic chemical synthesis and chemoenzymatic synthesis. The chemical syntheses employed orthogonally protected *myo*-inositol to introduce the pyrophosphate groups at the desired positions. In case of 5PP-InsP$_5$, *myo*-inositol was converted into an orthobenzoate ester followed by PMB protection. (**Scheme 2.1**) The resulting compound **1** was selectively deprotected by DIBAL-H at the 5-position to yield acetal **2**. The 5-position was phosphorylated with β-cyanoethyl (βCN) phosphoamidite followed by global acid deprotection. Pentaol **3** was then phosphorylated on positions 1–4 and 6 using xylyl (Xy) phosphoamidite to give **4**, followed by βCN deprotection and pyrophosphorylation at the 5-position with dibenzyl (Bn$_2$) phosphoamidite. Global hydrogenation of compound **5** afforded 7 mg of 5PP-InsP$_5$ in 30% yield over 7 steps starting from *myo*-inositol.[1–5]



Scheme 2.1: Chemical Synthesis of 5PP-InsP$_5$ with 30% overall yield. a) PhC(OMe)$_3$, CSA, DMSO b) NaH, PMB-Cl, DMF c) DIBAL-H, DCM d) βCE phosphoamidite, DCI, MeCN, then *t*BuOOH e) 5 % TFA, DCM f) Xy phosphoamidite, DCI, MeCN/DMF, then *m*CPBA g) DBU, BSTFA, MeCN, then MeOH, TFA, then Bn$_2$- phosphoamidite, tetrazole, MeCN then mCPBA h) NaHCO$_3$, Pd(black)/H$_2$, H$_2$O/*t*BuOH.

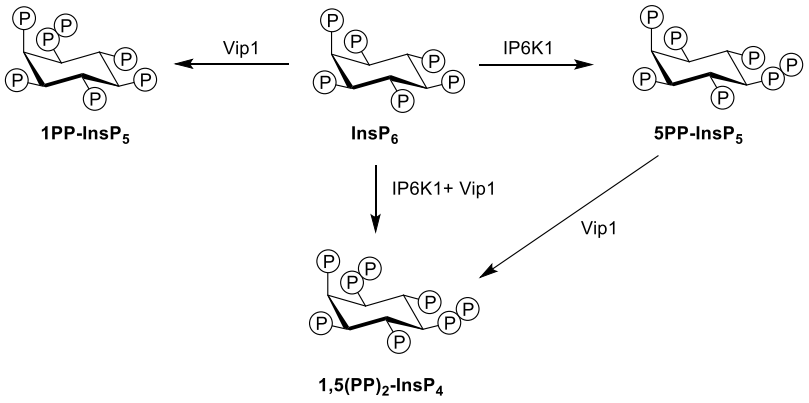
Compared to 5PP-InsP$_5$, the synthesis of the asymmetric 1PP-InsP$_5$ and 1,5(PP)$_2$-InsP$_4$ molecules was more challenging and lower yielding.[2,3] (**Scheme 2.2**) For 1PP-InsP$_5$, **2** was protected with PMB at the 5-position and the acetal was removed by acid hydrolysis to give **6a**. Phosphorylation at the 1- and the 3-positions with a chiral phosphoamidite led to a mixture of diastereoisomers which could be separated using silica chromatography. The PMB groups of the obtained **7a** were deprotected with acid and global phosphorylation of **8a** yielded the pentaphosphate ester **9a**. Selective deprotection and pyrophosphorylation of the 1-position yielded **10a** as the DBU salt which could be globally deprotected to afford 6 mg of 1PP-InsP$_5$ in 4.8% yield over 9 steps from *myo*-inositol. The chemical synthesis of 1,5(PP)$_2$-InsP$_4$ was performed in an analogous fashion.



Scheme 2.2: Chemical Synthesis of 1PP-InsP$_5$ and 1,5(PP)$_2$-InsP$_4$. a) **6a**: NaH, PMB-Cl, DMF; **6b**: Fm phosphoamidite, phenyl tetrazole, MeCN, then *t*BuOOH b) pTsOH, H$_2$O/MeOH or DCM/MeOH c) Rα phosphoamidite, tetrazole, MeCN, then *m*CPBA d) 2.5–5% TFA in CHCl$_3$ e) Xy phosphoamidite, DCI, MeCN, then *m*CPBA f) DBU, BSTFA, MeCN, then TFA, MeOH g) then Bn$_2$ phosphoamidite, tetrazole, MeCN, then *m*CPBA h) NaHCO$_3$, Pd(black)/H$_2$, H$_2$O/*t*BuOH.

Here, a third orthogonal phosphoryl protecting group was introduced at the 5-position that facilitated the introduction of the pyrophosphate in compound **10b**. Global deprotection afforded 11 mg of 1,5$(PP)_2$-InsP_4 in 4.4% yield over 9 steps from *myo*-inositol. Overall, organic syntheses of the PP-InsPs have been improved in recent years, but they still require multi-step procedures and remain viable only in the hands of skilled organic chemists in the context of an appropriately equipped laboratory. Organic chemistry, however, remains the only scalable solution to obtain milligram quantities of PP-InsPs in high purity.

An alternative route to generate PP-InsPs is their synthesis from InsP_6 using two distinct classes of small molecule kinases: IP6Ks and PPIP5Ks.[6,7] (**Scheme 2.3**) IP6K1 has been applied in the synthesis of 5PP-InsP_5 and the kinase domain (^{KD}) of PPIP5Ks either from *S. cerevisiae or H. sapiens* can be utilized for the synthesis of 1PP-InsP_5. The combined action of IP6K1 and PPIP5K^{KD} can provide access to 1,5$(PP)_2$-InsP_4. This approach has been in particular useful for the synthesis of radioactively labeled compounds where long syntheses are not suitable.[6] While the number of synthetic steps is strongly reduced compared to chemical synthesis, the challenge of this chemoenzymatic approach lies in separation of the PP-InsPs from other reaction components such as salts, buffer, proteins or other small molecules. HPLC and preparative polyacrylamide gel electrophoresis (PAGE) have been valuable tools to purify PP-InsP (discussed in detail in Chapter 3). Unfortunately, these techniques lack scalability and have thus far disqualified chemoenzymatic approaches as major sources for PP-InsPs.



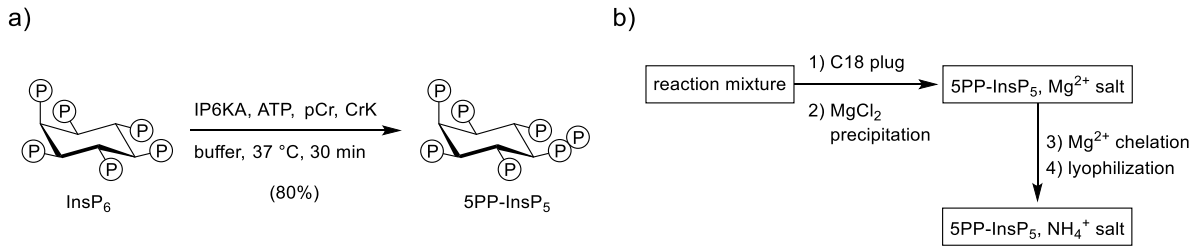
Scheme 2.3: Chemoenzymatic synthesis of PP-InsPs utilizes Vip1 or/and IP6K1.

In Chapter 2, a new strategy that merges the high efficiency of the enzymatic synthesis with the scalability of a classical synthetic approach is reported. High yielding procedures for the chemoenzymatic synthesis of 5PP-InsP$_5$, 1PP-InsP$_5$ and 1,5(PP)$_2$-InsP$_4$ on a 100–350 mg scale were developed, using the recently identified inositol hexakisphosphate kinase A (IP6KA) from *Entamoeba histolytica* and the kinase domain of human PPIP5K2 (PPIP5K2^{KD}). Coupled to a purification method that relies on precipitation of the PP-InsP-Mg complexes and strong anion exchange chromatography on a SAX fast protein liquid chromatography (FPLC) system, the PP-InsPs could be obtained in high purity. In the long term, easily accessible PP-InsP syntheses will enable researchers in fields removed from synthetic chemistry to investigate these molecules in a variety of biological contexts, enhancing our molecular understanding of PP-InsP signaling in healthy and diseased states.

Chemoenzymatic Synthesis of 5PP-InsP$_5$

To synthesize 5PP-InsP$_5$ chemoenzymatically, we recombinantly expressed IP6KA, an InsP$_6$ kinase from the parasitic amoeba *E. histolytica*.[8,9] The advantage of this enzyme, compared to the commonly used human IP6K1, is the fast enzyme reaction kinetics, good stability, and high expression yield (ca. 100 mg of protein per 1L of *E. coli* culture). Initially, we obtained an MBP-tagged IP6KA construct from Shears and co-workers that was only catalytically active upon TEV-mediated removal of the MBP-tag. The MBP-tag was replaced by a His-tag which allowed expression of a catalytically active IP6KA, circumventing the TEV cleavage. The substrate for IP6KA, InsP$_6$, was prepared from *myo*-inositol using a straightforward two step literature procedure.[10]

With substrate and enzyme in hand, reaction conditions were optimized (pH, time, enzyme- and Mg^{2+}-concentration) by monitoring PP-InsP kinase activity using NMR spectroscopy and uniform ^{13}C-labeled ([$^{13}C_6$]) InsP$_6$. This new spectroscopic approach was developed in parallel with the synthetic effort reported here, and is explained in detail in Chapter 3.[11] Full conversion of InsP$_6$ to the 5PP-InsP$_5$ product was achieved within 30 min, applying 0.3 µM of IP6KA at pH 6.4 in combination with an ATP recycling system (**Scheme 2.4 a**). The recycling of the ATP was necessary because accumulation of ADP favors the reverse reaction, resulting in an overall conversion to 5PP-InsP$_5$ of only 30%. Determining the optimal ratio of phosphate-containing species, Mg^{2+} ions, and pH was an integral consideration in preventing the precipitation of InsP$_6$ during the reaction. Therefore, the Mg^{2+} ion concentration was kept slightly below or equal to the combined concentration of phosphocreatine and ATP.



Scheme 2.4: Chemoenzymatic synthesis and purification of 5PP-$InsP_5$. a) Optimized conditions for the conversion of $InsP_6$ by IP6KA. Conditions: 250 µM $InsP_6$, 2 mM ATP, 7 mM $MgCl_2$, 5 mM creatine phosphate (pCr) and 1 U/mL creatine kinase (CrK), 0.3 µM IP6KA, 50 mM NaCl, 20 mM MES pH 6.4. b) Separation of the 5PP-$InsP_5$ from the reaction components. Enzymes were removed by a C18 plug and the product precipitation as a PP-InsP-Mg complex by addition of excess $MgCl_2$. that forms an insoluble PP-InsP-Mg complex. Mg^{2+}ions were subsequently exchanged by solid phase chelation in ammonium carbonate buffer and lyophilization afforded the product as the ammonium salt.

Employing the optimized conditions, the reaction was scaled up to 350 mg of $InsP_6$ starting material (based on the free acid) and a total reaction volume of 2 L. To accurately control the reaction time, it was crucial to preincubate the reaction mixture at 37 °C before the addition of IP6KA, and to quench the solution by cooling to 4 °C within 3–5 min using a –78 °C dry ice bath. For large scale purification, the proteins were removed by passing the solution through a fritted filter carrying a short plug of C18 reversed phase silica gel (**Scheme 2.4 b**). To separate 5PP-$InsP_5$ from other components in the mixture, we took advantage of the magnesium-chelating properties of InsPs and PP-InsPs.[12,13] At basic pH these highly negatively charged molecules form strong Mg-complexes and precipitate almost quantitatively. Therefore, an excess of $MgCl_2$ was added to the reaction mixture and the pH was raised to 8.8–9.0. The resulting precipitate could easily be isolated by centrifugation and separated from soluble impurities. To redissolve the 5PP-$InsP_5$-Mg complex, and to remove the Mg^{2+} ions, we applied an immobilized chelator (Amberlite® IRC 748) in NH_4HCO_3 buffer (pH 7.5). Initial attempts using water instead of buffer led to substantial hydrolysis of the

phosphoanhydride bond during this step, due to exposure of the product to elevated pH (up to 12). The instability of 5PP-InsPs in unbuffered solutions was a particular problem during upscaling because the incubation times with the chelator resin became longer compared to the smaller scale reactions. Finally, 5PP-InsP$_5$ was isolated by lyophilization as the ammonium salt in 80% yield (312 mg, >95% purity determined by NMR).

While 5PP-InsP$_5$ had been synthesized enzymatically before by Saiardi and co-workers, the scale of those reactions was substantially smaller. [7] The purification of the product was accomplished by preparative PAGE, followed by gel extraction (100–200 µg 5PP-InsP$_5$ per gel). We have reproduced this purification step on 5PP-InsP$_5$ mixed together with all reaction components except IP6KA. The gel band was cut, extracted and analyzed by NMR spectroscopy. (**Figure 2.1**)

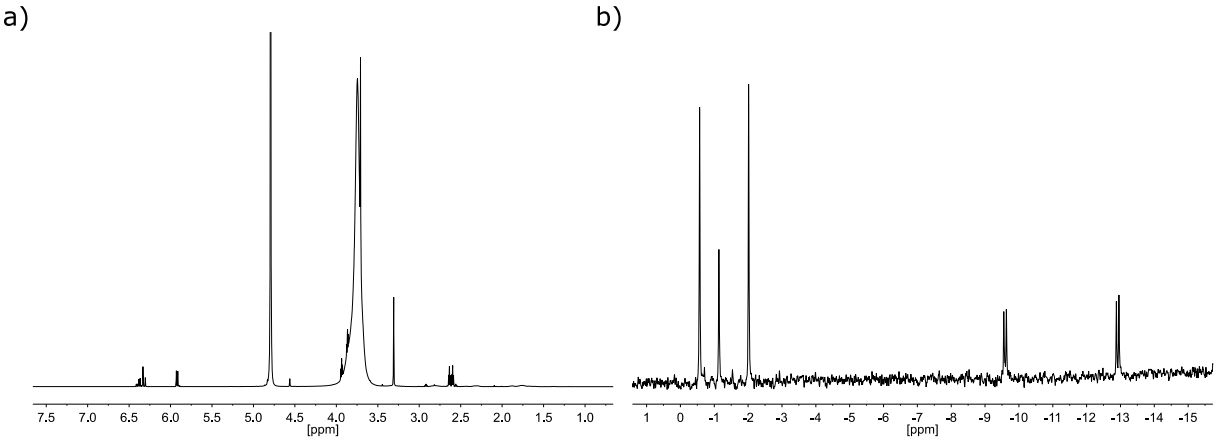


Figure 2.1: PAGE purified 5PP-InsP$_5$ is contaminated with buffer components. Pure chemoenzymatically prepared 5PP-InsP$_5$ was mixed with all kinase reaction components except IP6KA and purified by PAGE followed by gel extraction. a) ^{1}H NMR spectrum shows several highly abundant species that are a 10–100 fold higher in intensity which precludes the observation of the inositol protons. b) ^{31}P NMR spectrum shows 5PP-InsP$_5$ and no other PP-InsP species.

While the ^{31}P NMR showed only the expected 5PP-InsP$_5$ peaks, several impurities were observed in the ^{1}H NMR that are 10–100 fold higher in intensity and precluded the observation of the *myo*-inositol signals. These results indicate that our precipitation-based purification is superior to the isolation by preparative PAGE with respect to scalability and purity of the final product.

During the optimization of the 5PP-InsP$_5$ synthesis, we were surprised to observe the formation of one major side product during extended reaction times. It was speculated that this species could be another PP-InsP$_5$ regioisomer or even an (PP)$_2$-InsP$_4$ species. The unknown compound was isolated by SAX-FPLC using NH_4HCO_3 (pH 7.5) as elution buffer. After lyophilization, analysis by one- and two-dimensional NMR spectroscopy revealed the unknown product as 4,5(PP)$_2$-InsP$_4$ or 6,5(PP)$_2$-InsP$_4$ (the enantiomers are not distinguishable by NMR). (**Figure 2.2**)

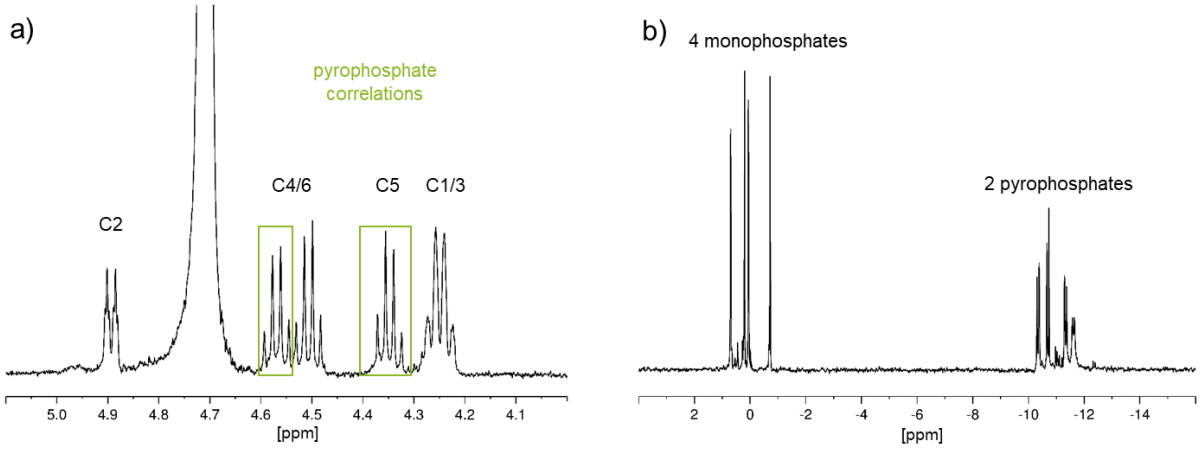


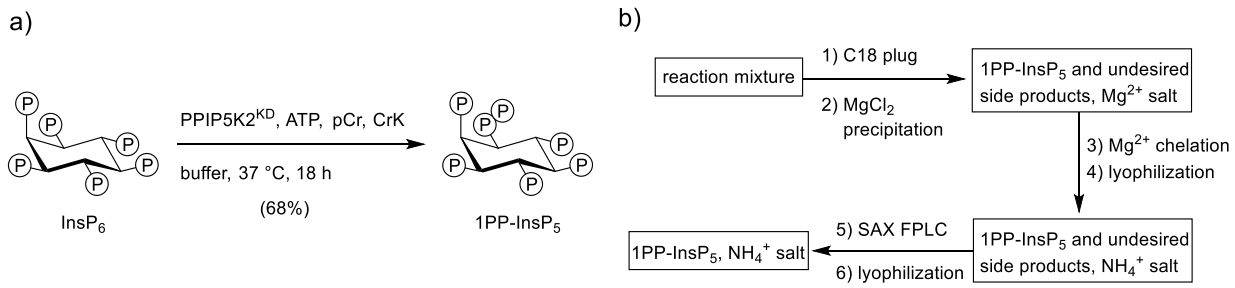
Figure 2.2: NMR spectroscopic analysis of 4/6,5(PP)$_2$-InsP$_4$. a) ^{1}H NMR spectrum with the positions on the *myo*-inositol ring depicted as C1–6 as determined by ^{1}H,^{1}H-COSY NMR analysis. Green boxes show the positions that correlate with the pyrophosphate region of the ^{31}P NMR spectrum. b) ^{31}P NMR spectrum shows 4 monophosphates around 0 ppm and 2 pyrophosphates between –10 and –12 ppm.

Interestingly, previous reports that utilized IP6KA or IP6K1 in enzymatic reactions did not mention the formation of this product, and whether this compound has a physiological relevance has not been determined yet. However, 6,5(PP)$_2$-InsP$_4$ is the dominant (PP)$_2$-InsP$_4$ regioisomer in the slime mold *D. discoideum*.[14–16] Synthetic access to this compound could therefore be of use to study PP-InsP signaling in this organism.

In summary, our chemoenzymatic synthesis of 5PP-InsP$_5$ outperforms classical organic chemistry approaches in yield and accessibility. The precipitation of the 5PP-InsP$_5$-Mg complex offers a scalable purification procedure that can deliver hundreds of milligrams large amounts of high purity 5PP-InsP$_5$.

Chemoenzymatic Synthesis of 1PP-InsP$_5$ and 1,5(PP)$_2$-InsP$_4$

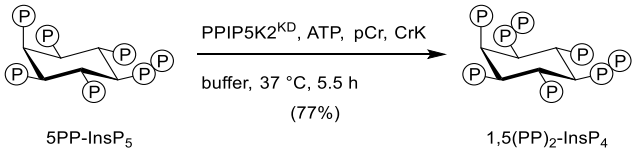
The successful, large-scale chemoenzymatic synthesis of 5PP-InsP$_5$ encouraged us to apply our strategy to the synthetically more challenging asymmetric PP-InsPs: 1PP-InsP$_5$ and 1,5(PP)$_2$-InsP$_4$. Pyrophosphorylation at the 1-position of InsP$_6$ and 5PP-InsP$_5$ is known to be catalyzed by PPIP5Ks. We thus expressed the His-tagged kinase domain of human PPIP5K2 (PPIP5K2^{KD}), but low solubility resulted in low overall protein yields. To increase the solubility, the His-tag of PPIP5K2^{KD} was replaced with a solubility-enhancing SUMO-tag, leading to the isolation of the desired protein in sufficient yield (25 mg of protein per 1L of *E. coli* culture). Next, the biochemical reaction for the formation of 1PP-InsP$_5$ was optimized with labeled [$^{13}C_6$]InsP$_6$ as substrate. Extended reaction time of 18 hours and an increased enzyme concentration of 2 µM were necessary to guarantee full conversion. (**Scheme 2.5 a**)



Scheme 2.5: Chemoenzymatic synthesis and purification of 1PP-InsP$_5$. a) Optimized conditions for the synthesis of 1PP-InsP$_5$: 250 µM InsP$_6$, 2 mM ATP, 6 mM MgCl$_2$, 5 mM creatine phosphate (pCr) and 1 U/mL creatine kinase (CrK), 2 µM PPIP5K2^{KD}, 250 mM NaCl, 20 mM MES pH 6.4. b) Separation of the 1PP-InsP$_5$ from the reaction components was performed analogous to 5PP-InsP$_5$ followed by SAX FPLC to remove remaining side products.

Following precipitation and Mg^{2+}-chelation, 1PP-InsP$_5$ still contained impurities, probably due to the formation of side products under the prolonged reaction times. To remove these compounds, 1PP-InsP$_5$ was subjected to SAX-FPLC (**Scheme 2.5 b**), allowing for the isolation of ca. 100 mg of 1PP-InsP$_5$ in two purification runs. After lyophilization, 1PP-InsP$_5$ was isolated as the ammonium salt in 68% yield (77 mg, >95% purity determined by NMR).

With robust procedures for enzymatic pyrophosphorylation at the 5- and the 1-position of the inositol ring in hand, we lastly sought to synthesize the most densely phosphorylated PP-InsP: 1,5(PP)$_2$-InsP$_4$. Full conversion from InsP$_6$ to 1,5(PP)$_2$-InsP$_4$ using both enzymes in a one-pot reaction proved to be challenging. Therefore, a one-step protocol starting from purified 5PP-InsP$_5$ was pursued. (**Scheme 2.6**) Reaction optimization with [$^{13}C_6$]5PP-InsP$_5$ confirmed that 5PP-InsP$_5$ is the preferred substrate for human PPIP5K2^{KD}, compared to InsP$_6$, and full conversion could be obtained with 1.5 µM enzyme in 5.5 hours. 1,5(PP)$_2$-InsP$_4$ was purified analogously to 5PP-InsP$_5$ by precipitation with Mg^{2+} ions. The product was isolated as the ammonium salt in 77% yield (111 mg, >85% purity determined by NMR).



Scheme 2.6: Chemoenzymatic synthesis and purification of 1,5(PP)$_2$-InsP$_4$. Optimized conditions for the synthesis of 1,5(PP)$_2$-InsP$_4$: 250 µM 5PP-InsP$_5$, 2 mM ATP, 5 mM $MgCl_2$, 5 mM creatine phosphate (pCr) and 1 U/mL creatine kinase (CrK), 1.5 µM PPIP5K2^{KD}, 250 mM NaCl, 20 mM MES pH 6.4.

Overall, the synthesis of asymmetric 1,5(PP)$_2$-InsP$_4$ was achieved in an analogous fashion to 5PP-InsP$_5$ with only minor changes in the purification protocol. The purification of 1PP-InsP$_5$, by contrast, required an additional step, limiting access to this compound to ca.100 mg batches. Preparation of similar amounts of 1PP-InsP$_5$ by chemical means, however, also remains challenging and demands specialized chemistry equipment that is typically not present in biochemistry and molecular biology laboratories.

Conclusion and Outlook

Over the last decade we have witnessed significant progress in the synthesis of PP-InsPs. Still, the general availability of these compounds in good quantities, especially of the asymmetric regioisomers, has been limited. Here, we have established a robust workflow for PP-InsP synthesis and purification. With the required enzymes and relevant substrates in hand, all characterized mammalian PP-InsPs were synthesized on at least a 100 mg scale in one or two steps, without the need for specialized chemistry equipment. While our method is efficient in synthesizing the reported PP-InsPs, some limitations remain to be addressed in the future. The current protocol is restricted by the availability and activity of the small molecule kinases that selectively install the pyrophosphate groups. For example, *D. discoideum*, a slime mold

containing the highest concentration of PP-InsPs measured to date, produces PP-InsP regioisomers that bear pyrophosphate groups at the 5- or 6-position or on both.[15] In addition, our method is not suited to synthesize analogs of PP-InsPs, such as non-hydrolyzable bisphosphonate analogs or photo-caged derivatives. Especially in the case of large structural perturbations of the PP-InsPs we envision that those syntheses remain in the realm of classical organic chemistry. Overall, the scalable biochemical syntheses of the major mammalian PP-InsPs make these compounds readily available to a broader research community and will greatly aid a thorough analysis of the varied properties of PP-InsPs. Just as access to other second messengers greatly aided the structural and biochemical elucidation of their signaling function, the method presented here will facilitate a thorough and granular analysis of the varied properties of PP-InsPs. Such detailed investigations are much needed, to determine the molecular mechanisms underlying the pleiotropic phenotypes associated with PP-InsP signaling, and ultimately guide the development of novel therapeutic strategies.

Thus far, the biochemically prepared PP-InsPs have enabled a broad use of these molecules in many research projects. The facile precipitation procedure of PP-InsPs proved to be equally useful on a small scale, as demonstrated for the synthesis of 5PP-$InsP_5$-$\beta^{32}P$.[17] In this case, HPLC purification of the radioactive compound was no longer necessary to obtain high purity 5PP-$InsP_5$-$\beta^{32}P$. The route to this material can be easily reproduced without high instrumental hurdles and will be a valuable source to study protein pyrophosphorylation. Access to large amounts of 5PP-$InsP_5$ also enabled the development of a stringent high throughput screen with 35,000 compounds to identify inhibitors of IP6K. The rationale for these experiments were developed in Chapter 3 and will be explained in detail. Using 1,5$(PP)_2$-$InsP_4$ as a substrate an analogous approach is planned for the PPIP5Ks in the near future.

Recently, providing larger amounts of PP-InsPs established several new collaborations. Together with the Moldoveanu laboratory, we started to investigate the role of PP-InsPs in necroptosis through activation of MLKL, a known allosteric substrate of InsPs.[18] Furthermore, 5PP-InsP$_5$ was provided to the Hothorn group to aid crystallization studies of plant proteins. We also working together with the Saiardi group on PP-InsPs in *D. discoideum* and therefore pursue the synthesis to the specific regioisomers from this organism. It is planned to optimize the reaction conditions for the synthesis of 4/6,5(PP)$_2$-InsP$_4$ using IP6KA and together with Siw14, a selective 5PP-InsP pyrophosphatase, we will try to synthesize also 4/6PP-InsP$_5$.[19,20]

Finally, the access to large amounts of PP-InsPs can also contribute to the study of protein pyrophosphorylation. With this material in hand, the energy of hydrolysis can be experimentally determined to confirm and rank the ability of these molecules to transfer the β-phosphoryl group onto proteins. At the moment, it is only known that all PP-InsPs have pyrophosphorylation potential but the differences between PP-InsPs – whether these are different substrate specificities or distinct reaction rates – have not been investigated. In line with this, we also want to observe and optimize the *in vitro* pyrohosphorylation of proteins in combination with MS or/and NMR spectroscopy. Currently, radiolabeling is the only method that can detect pyrophosphorylation *in vitro,* probably due to the low efficiency of this reaction.[6,21] Flexible use of PP-InsPs will allow a thorough identification and analysis of important parameters that govern the formation of this modification. Overall, the chemoenzymatic synthesis of PP-InsPs expands the use of the molecules to many new applications and will allow more groups to study their biological role and function.

Methods

General Information

All chemicals were purchased from Sigma Aldrich, VWR, Carl Roth, Thermo Fisher Scientific, Alfa Aesar, TCI and used without further purification unless stated otherwise. All dry solvents were purified using a solvent purification system MBRAUN MB-SPS-5 by passing through activated alumina columns. Deuterated solvents were purchased from Euriso-Top. The C18 reversed phase silica was purchased from Carl Roth. Telos® was ordered from Kinesis. Automated flash chromatography was performed using gradient grade solvents on a CombiFlash® Rf from Teledyne Isco using prepacked CombiFlash® columns (40–63 µm). LC-MS analysis was carried out with an Agilent 1260 Infinity Binary LC system connected to an Agilent 6130 Quadrupole LC-MS system with a ZORBAX Rapid Resolution HT Narrow Bore SB-C18 1.8 µm column (2.1 x 50mm) at 30 °C using API-ESI (atmospheric pressure ionization-electrospray) in positive ion mode. The eluent consisted of 10% ACN in water with 0.1% formic acid at 0.7 mL/min flow rate.

NMR spectra were recorded on Bruker spectrometers operating at 300 or 600 MHz for proton nuclei, 75 or 151 MHz for carbon nuclei or 122 and 244 MHz for phosphorous nuclei. NMR data are given as follows: chemical shift δ in ppm (multiplicity, coupling constant(s) J Hz, relative integral) where multiplicity is defined as: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad or combinations of the above. Measurement for the determination of enzyme activity was performed on Bruker AV-III spectrometers (Bruker Biospin, Rheinstetten, Germany) at 310 K using cryogenically cooled 5 mm TCI-triple resonance probe equipped with one-axis self-shielded gradients. The software used to control the spectrometer was topspin 3.5 pl6. Temperature had been calibrated using d_4-methanol and the formula

of Findeisen et al.[22] High-resolution mass spectrometer was performed by direct inject on an OrbitrapTM Q-Exactive mass spectrometer (Thermo Fisher Scientific). For the purification *via* FPLC an NGC Quest™ 10 Chromatography System from Bio-Rad was used with an integrated NGC™ Sample Pump Module and a BioFrac™ Fraction Collector. For the spin filtration Amicon Ultra 0.5 mL centrifugal filters with a cut off of 10 kDa or 3 kDa from Merck Millipore were used.

Recombinant Protein Expression

Inositol Hexakisphosphate Kinase A (IP6KA)

The procedure was adopted from.[11] In short: The codon optimized *Entamoeba histolytica* IP6KA-gene in a pET15b plasmid was transformed into *E. coli* BL21 Arctic Express (DE3) and an overnight culture of this strain was diluted into 1 L of TB-Amp/Gen to a final density of OD_{600} 0.1. The cells were grown for 6 h at 37 °C. The culture was then switched to 18 °C for 30 min before induction with 0.1 mM IPTG for 18 hours. The cells were harvested by centrifugation (3,000 g for 10 min at 4 °C) and washed with ice cold water. The cell pellet was resuspended in lysis buffer (25 mM Tris HCl pH 7.4, 500 mM NaCl, 50 mM imidazole). For 1 g wet weight 10 mL lysis buffer was used. The cell suspension was supplemented with lysozyme, DNase I, and 1 tablet cOmplete™ protease inhibitor (Roche), and incubated for 15 min on ice. The cells were lysed with a Microfluidizer™ LM10 at 15,000 psi with five iterations. The lysate was clarified by centrifugation (30,000 g for 30 min at 4 °C). The supernatant was filtered (VWR® vacuum filter, PES 0.45 µm), and loaded onto a Ni-NTA column (GE, 5 mL, HiTrap IMAC HP) equilibrated with lysis buffer with a flowrate of 2.5 mL/min. The column was washed with lysis buffer until the absorption was constant. IP6KA was eluted with a gradient of elution buffer (25 mM Tris HCl pH 7.4, 200 mM NaCl, 500 mM

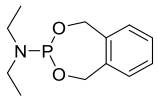
imidazole) in lysis buffer from 0–100 % over 10 CV. 1.5 mL fractions were collected. The fractions containing IP6KA were concentrated by spin filtration (Amicon® Ultra 0.5 mL 10K) and dialyzed overnight against dialysis buffer (20 mM Tris HCl pH 7.4, 200 mM NaCl, 1 mM DTT). The following day the protein was adjusted to 25 % glycerol, frozen in liquid nitrogen and stored at –80 °C.

Diphosphoinositol Pentakisphosphate Kinase 2 Kinase Domain (PPIP5K2^{KD})

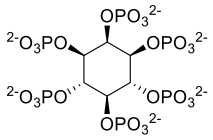
The codon optimized human PPIP5K2^{KD}-gene was subcloned from a pET21a vector into a pSumo vector using the restriction sites NdeI and XhoI. The resulting vector was transformed into *E. coli* Arctic Express (DE3) and an overnight culture of this strain was diluted into 2 L of TB-Kan/Gen to a final density of OD_{600} 0.1. The cells were grown for 4 h at 37 °C. The culture was then switched to 15 °C for 30 min before induction with 0.2 mM IPTG for 18 hours. The cells were harvested by centrifugation (3,000 g for 10 min at 4 °C) and washed with ice cold water. The cell pellet was resuspended in lysis buffer (20 mM Tris HCl pH 7.5, 500 mM NaCl). For 1 g wet weight 10 mL lysis buffer was used. The cell suspension was supplemented with lysozyme, DNase I, and 2 tablets cOmplete™ protease inhibitor (Roche), and incubated for 15 min on ice. The cells were lysed using a Microfluidizer™ LM10 at 15,000 psi with five iterations. The lysate was clarified by centrifugation (30,000 g for 30 min at 4 °C). The supernatant was filtered (VWR® vacuum filter, PES 0.45 µm), and loaded onto a Ni-NTA column (GE, 5 mL, HiTrap IMAC FF) equilibrated with lysis buffer with a flowrate of 3.5 mL/min. The column was washed with 3 CV lysis buffer and 10 CV of 10 % elution buffer (20 mM Tris HCl pH 7.5, 500 mM NaCl, 600 mM imidazole) in lysis buffer or until the absorption was constant. PPIP5K2 was eluted with 100 % elution buffer over 10 CV. 1.4 mL fractions were collected. The fractions containing PPIP5K2 were

dialyzed overnight against dialysis buffer (20 mM Tris HCl pH 7.4, 500 mM NaCl, 1 mM DTT). The following day the protein was adjusted to 20 % glycerol, frozen in liquid nitrogen and stored at –80 °C.

Chemical Synthesis of $InsP_6$



The compound was synthesized as described before in two steps starting from *myo*-inositol. Analytical data were identical with the values reported in the literature.[23,24]



$InsP_6$

The compound was synthesized as described before in two steps starting from *myo*-inositol. Analytical data were identical with the values reported in the literature.[10]

Chemoenzymatic Synthesis of Inositol Pyrophosphates

For all enzymatic syntheses the following stock solutions were prepared:

Creatine kinase stock solution: 350 U/mL creatine kinase in 200 mM MOPS pH 6.5, 20 mM $MgCl_2$, 20 mM DTT

IP6KA stock solution: 10 mg/mL IP6KA in 20 mM Tris HCl pH 7.4, 200 mM NaCl, 1 mM DTT

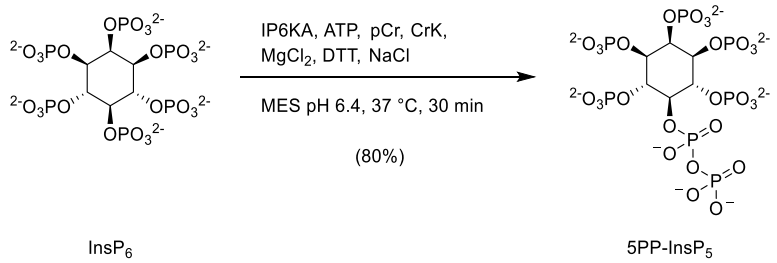
$PPIP5K2^{KD}$ stock solution: 10 mg/mL 20 mM Tris HCl pH 7.4, 500 mM NaCl, 1 mM DTT, 40% glycerol

ATP stock solution: 100 mM in MilliQ® water pH 6.4; (Note: Concentration was determined *via* UV-Vis analysis at 259 nm; ε_{259} = 15.4 E/mmol/cm)

Amberlite® IRC-748 resin was prepared as follows:

500 mL resin was placed on a big fritted filter and washed with 1 L MeOH and 1 L H_2O and stored at 4 °C in 20 % MeOH. For a 100 mg reaction roughly 30 mL of resin were placed on a filter, washed with 500 mL H_2O and acidified with 250 mL of 1 M HCl. The resin was washed with H_2O (usually 1–2 L) until it reached pH 6–7 and the resin was equilibrated with 250 mL 1 M NH_4HCO_3 pH 7.5–8. The resin was washed with 1 L H_2O and left dry on the filter until use.

Synthesis of 5PP-InsP$_5$



InsP$_6$ 5PP-InsP$_5$

MilliQ® water was prewarmed to 37 °C by incubation in a water bath. A solution of InsP$_6$ (250 µM, 350 mg based on free acid 647.9 g/mol), MES (20 mM, pH 6.4), NaCl (50 mM), ATP (disodium salt, 2 mM), creatine phosphate (5 mM), $MgCl_2$ (7 mM), DTT (1 mM) in 2.16 L prewarmed MilliQ® water was prepared. The solution was evenly split among 3 × 1 L Schott bottles and the bottles were incubated in the water bath at 37 °C for 10 min. IP6KA (1 µM) and creatine kinase (1 U/mL) were added and the Schott bottles were gently inverted several times to homogenize and left to react for 30 min without shaking. (Note: The correct temperature was essential to assure full conversion of the starting material within 30 min. Prolonged reaction times above 1 h led to side reactions. However, the speed of the reaction will depend on the batch and quality of the recombinantly expressed IP6KA.)

Purification: The reaction was stopped by cooling the reaction mixture down to 4 °C within 5 min with the help of a dry ice isopropanol bath. A fritted filter was loaded with 21 g of C18 reversed phase silica gel suspended in MeCN and sand was added on top. The C18 plug was washed with 100 mL MeCN and 100 mL H_2O and the complete reaction mixture was passed through the filter under vacuum. The C18 plug was washed with 2 × 100 mL H_2O and all the combined flow through was supplemented with 12.5 mM $MgCl_2$. The pH was adjusted to 8.8–9.0 by drop wise addition of 10 mM NaOH solution which leads to precipitation of the PP-InsPs as magnesium complex within 1 h at room temperature. (The precipitation can also be performed overnight at

room temperature.) The suspension was collected in 4 canonical 50 mL tubes by centrifugation (2 min at 3000 g) and the supernatant was removed. The precipitate of each tube was washed 3 times with 15 mL $MgCl_2$ solution (8 mM, pH 9 adjusted with NaOH).

The precipitate of each tube was resuspended in 20 mL NH_4HCO_3 buffer (10 mM, pH 8) and vortexed with Amberlite® IRC-748 (20 mL wetted bed volume, pre-equilibrated with NH_4HCO_3, pH 7.5–8) until the precipitate dissolved. The buffer/resin suspension of each tube was added to a short Amberlite® IRC-748 column (5 mL bed volume, pre-equilibrated with NH_4HCO_3, pH 8) to remove excess Mg^{2+}. The product was flushed through the column with 50 mL water and all eluents were collected, combined and lyophilized in a round-bottom flask to afford the ammonium salt of the product as white solid. The solids were dissolved in D_2O and the concentration of the solution was determined by NMR against a standard (phosphonoacetic acid) to determine yield and purity: 30 mL of 14.3 mM solution were obtained which corresponds to 312 mg (80% yield, purity > 95%) 5PP-InsP_5 based on the free acid 726.9 g/mol.

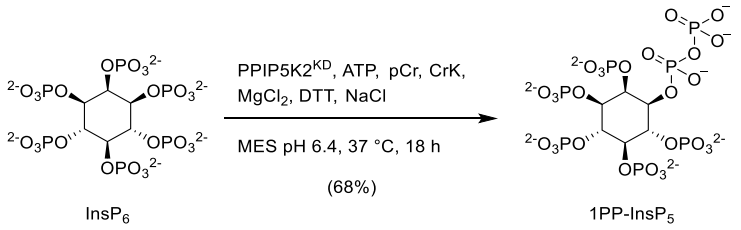
^{1}H NMR (300 MHz, Deuterium Oxide, pD 5.0) δ 4.46 (q, *J* = 9.4 Hz, 2H), 4.27 (q, *J* = 9.4 Hz, 1H), 4.10 (t, *J* = 9.6 Hz, 2H).

^{13}C NMR (151 MHz, D_2O, pD 5.0) δ 77.57, 75.81, 75.69, 73.20.

^{31}P NMR (243 MHz, Deuterium Oxide, pD 5.0) δ 0.60, 0.53, -0.42, -10.31 (d, *J* = 19.9 Hz), -10.86 (d, *J* = 19.9 Hz).

HRMS (ESI) calculated for $C_6H_{17}O_{27}P_7(M^-)$ 738.8204, found 738.8222.

Synthesis of 1PP-InsP$_5$



MilliQ® water was prewarmed to 37 °C by incubation in a water bath. A solution of InsP$_6$ (250 µM, 100 mg based on free acid 647.9 g/mol), MES (20 mM, pH 6.4), NaCl (250 mM), ATP (disodium salt, 2 mM), creatine phosphate (5 mM), $MgCl_2$ (6 mM), DTT (1 mM) in 617 mL prewarmed MilliQ® water was prepared in a 1 L Schott bottles and incubated in a water bath at 37 °C for 10 min. PPIP5K2^{KD} (2 µM) and creatine kinase (1 U/mL) were added and bottle gently inverted several times to homogenize and left to react overnight for 18 h without shaking.

Purification: The reaction was stopped by cooling the reaction mixture down to 4 °C within 5 min with the help of a dry ice isopropanol bath. A fritted filter was loaded with 6 g of C18 reversed phase silica gel suspended in MeCN and sand was added on top. The C18 plug was washed with 30 mL MeCN and 30 mL H_2O and the complete reaction mixture was passed through the filter under vacuum. The C18 plug was washed with 2 × 30 mL H_2O and all the combined flow through was supplemented with 12.5 mM $MgCl_2$. The pH was adjusted to 8.8–9.0 by drop wise addition of 10 mM NaOH solution which leads to precipitation of the PP-InsPs as magnesium complex within 1 h at room temperature. (The precipitation can also be performed overnight at room temperature.) The suspension was collected in 2 canonical 50 mL tubes by centrifugation (2 min at 3000 g) and the supernatant was removed. The precipitates of each tube were washed 3 times with 15 mL $MgCl_2$ solution (8 mM, pH 9 adjusted with NaOH).

The precipitate of each tube was resuspended in 15 mL NH_4HCO_3 buffer (10 mM, pH 8) and vortexed with Amberlite® IRC-748 (15 mL wetted bed volume, pre-equilibrated with NH_4HCO_3, pH 7.5–8) until the precipitate dissolved. The buffer/resin suspension of each tube was added to a short Amberlite® IRC-748 column (5 mL bed volume, pre-equilibrated with NH_4HCO_3, pH 8) to remove excess Mg^{2+}. The product was flushed through the column with 40 mL H_2O and all eluents were collected, combined and lyophilized in a round-bottom flask. The resulting white solid was purified in 2 runs using a SAX-FPLC (HiPrep™ Q HP 16/10, GE Healthcare) and H_2O as buffer A and 1 M NH_4HCO_3 as buffer B. The column was washed with 100% B and equilibrated at 1% B. The sample was dissolved and loaded in 1% B and, followed by a gradient from 1% to 20% in 1 CV. The product was eluted in a gradient from 20%–40% over 10 CV. Fractions were analyzed by a metal dye detection assay[25] in a 96-well plate format and product containing fractions were combined and lyophilized in a round-bottom flask. The solids were dissolved in D_2O and the concentration of the solution was determined by NMR against a standard (phosphonoacetic acid) to determine yield and purity: 20 mL of 5.3 mM solution were obtained which corresponds to 77 mg (68% yield, purity > 95%) 1PP-InsP_5 based on the free acid 726.9 g/mol.

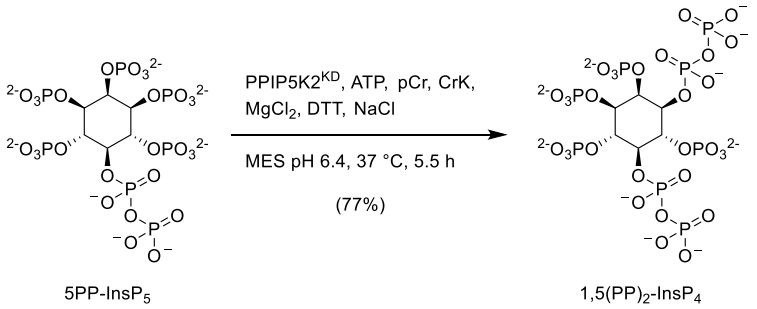
^{1}H NMR (600 MHz, Deuterium Oxide, pD 5.0) δ 5.08 (d, *J* = 9.8 Hz, 1H), 4.40 (p, *J* = 9.7 Hz, 2H), 4.21 (t, *J* = 9.8 Hz, 1H), 4.15 (q, *J* = 9.5 Hz, 1H), 4.09 (t, *J* = 9.2 Hz, 1H).

^{13}C NMR (151 MHz, D_2O, pD 5.0) δ 77.32, 76.11, 75.68, 75.58, 73.70, 73.08.

^{31}P NMR (243 MHz, Deuterium Oxide, pD 5.0) δ 1.64, 0.90, 0.40, 0.32, -0.94, -8.70, -10.89 (d, *J* = 18.7 Hz).

HRMS (ESI) calculated for $C_6H_{17}O_{27}P_7(M^-)$ 738.8204, found 738.8198.

Synthesis of 1,5(PP)$_2$-InsP$_4$



5PP-InsP$_5$ 1,5(PP)$_2$-InsP$_4$

MilliQ® water was prewarmed to 37 °C by incubation in a water bath. A solution of 5PP-InsP$_5$ (250 µM, 100 mg based on free acid 726.9 g/mol), MES (20 mM, pH 6.4), NaCl (250 mM), ATP (disodium salt, 2 mM), creatine phosphate (5 mM), MgCl$_2$ (5 mM), DTT (1 mM) in 715 mL prewarmed MilliQ® water was prepared in a 1 L Schott bottles and incubated in the water bath at 37 °C for 10 min. PPIP5K2^{KD} (1.5 µM) and creatine kinase (1 U/mL) were added and bottle gently inverted several times to homogenize and left to react overnight for 5.5 h without shaking.

Purification: The reaction was stopped by cooling the reaction mixture down to 4 °C within 5 min with the help of a dry ice isopropanol bath. A fritted filter was loaded with 6 g of C18 reversed phase silica gel suspended in MeCN and sand was added on top. The C18 plug was washed with 30 mL MeCN and 30 mL H$_2$O and the complete reaction mixture was passed through the filter under vacuum. The C18 plug was washed with 2 × 30 mL H$_2$O and all the combined flow through was supplemented with 12.5 mM MgCl$_2$. The pH was adjusted to 8.8–9.0 by drop wise addition of 10 mM NaOH solution which leads to precipitation of the PP-InsPs as magnesium complex within 1 h at room temperature. (The precipitation can also be performed overnight at room temperature.) The suspension was collected in 2 canonical 50 mL tubes by centrifugation (2 min at 3000 g) and the supernatant was removed. The precipitates of

each tube were washed 3 times with 15 mL $MgCl_2$ solution (8 mM, pH 9 adjusted with NaOH).

The precipitate of each tube was resuspended in 15 mL NH_4HCO_3 buffer (10 mM, pH 8) and vortexed with Amberlite® IRC-748 (15 mL wetted bed volume, pre-equilibrated with NH_4HCO_3, pH 7.5–8) until the precipitate dissolved. The buffer/resin suspension of each tube was added to a short Amberlite® IRC-748 column (5 mL bed volume, pre-equilibrated with NH_4HCO_3, pH 8) to remove excess Mg^{2+}. The product was flushed through the column with 40 mL H_2O and all eluents were collected, combined and lyophilized in a round-bottom flask to afford the ammonium salt of the product as white solid. The solids were dissolved in D_2O and the concentration of the solution was determined by NMR against a standard (phosphonoacetic acid) to determine yield and purity: 12 mL of 12.7 mM solution were obtained which corresponds to 111 mg (77% yield, purity > 85%) 1,5$(PP)_2$-InsP_4 based on the free acid 805.9 g/mL.

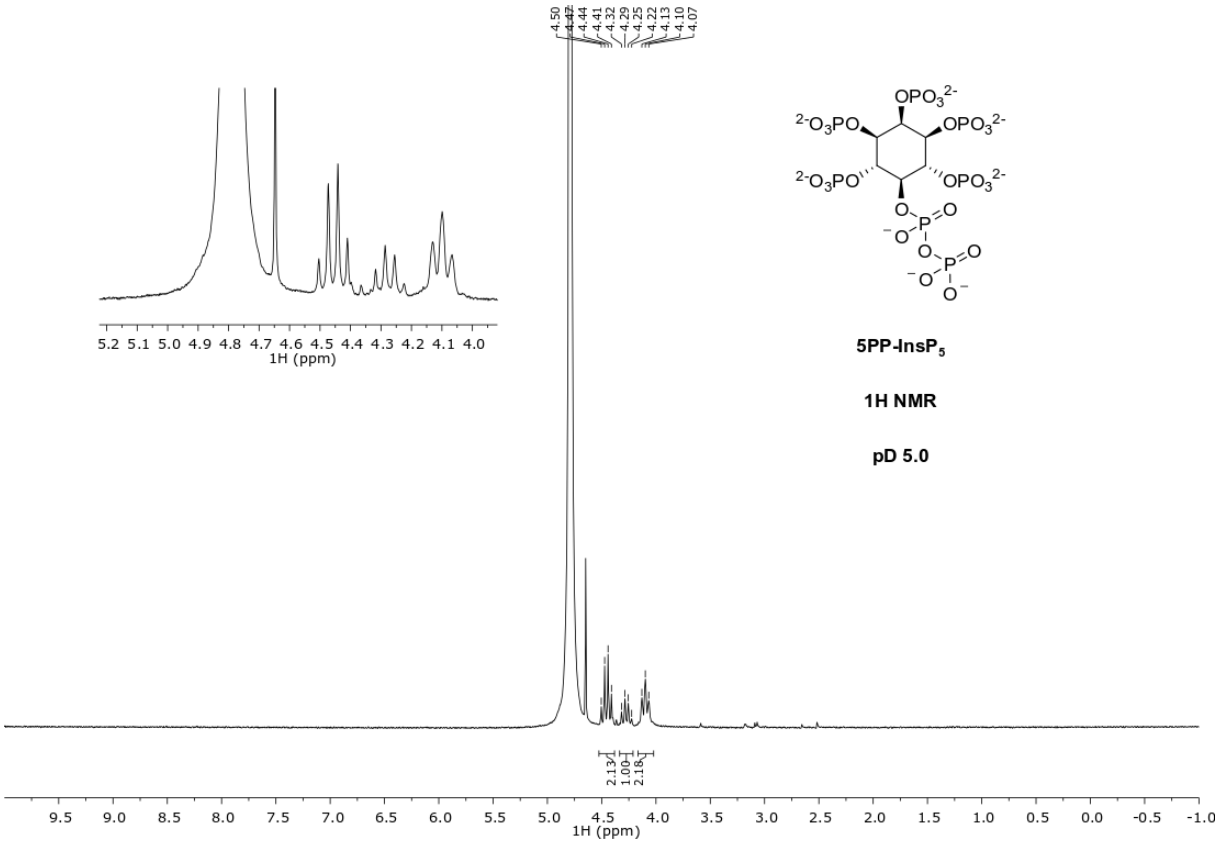
^{1}H NMR (600 MHz, Deuterium Oxide, pD 5.0) δ 5.01 (d, *J* = 9.6 Hz, 1H), 4.55 (p, *J* = 9.9 Hz, 2H), 4.36 (q, *J* = 9.8 Hz, 1H), 4.25 (t, *J* = 9.4 Hz, 1H).

^{31}P NMR (243 MHz, Deuterium Oxide, pD 5.0) δ 0.56, 0.33, -0.26, -1.04, -10.34 (dd, *J* = 29.9, 19.2 Hz), -11.03 (dd, *J* = 60.0, 19.3 Hz).

^{13}C NMR (151 MHz, D_2O, pD 5.0) δ 77.42, 76.05, 75.53, 74.78, 73.53, 73.18.

HRMS (ESI) calculated for $C_6H_{17}O_{27}P_7(M^-)$ 818.7868, found 818.7865.

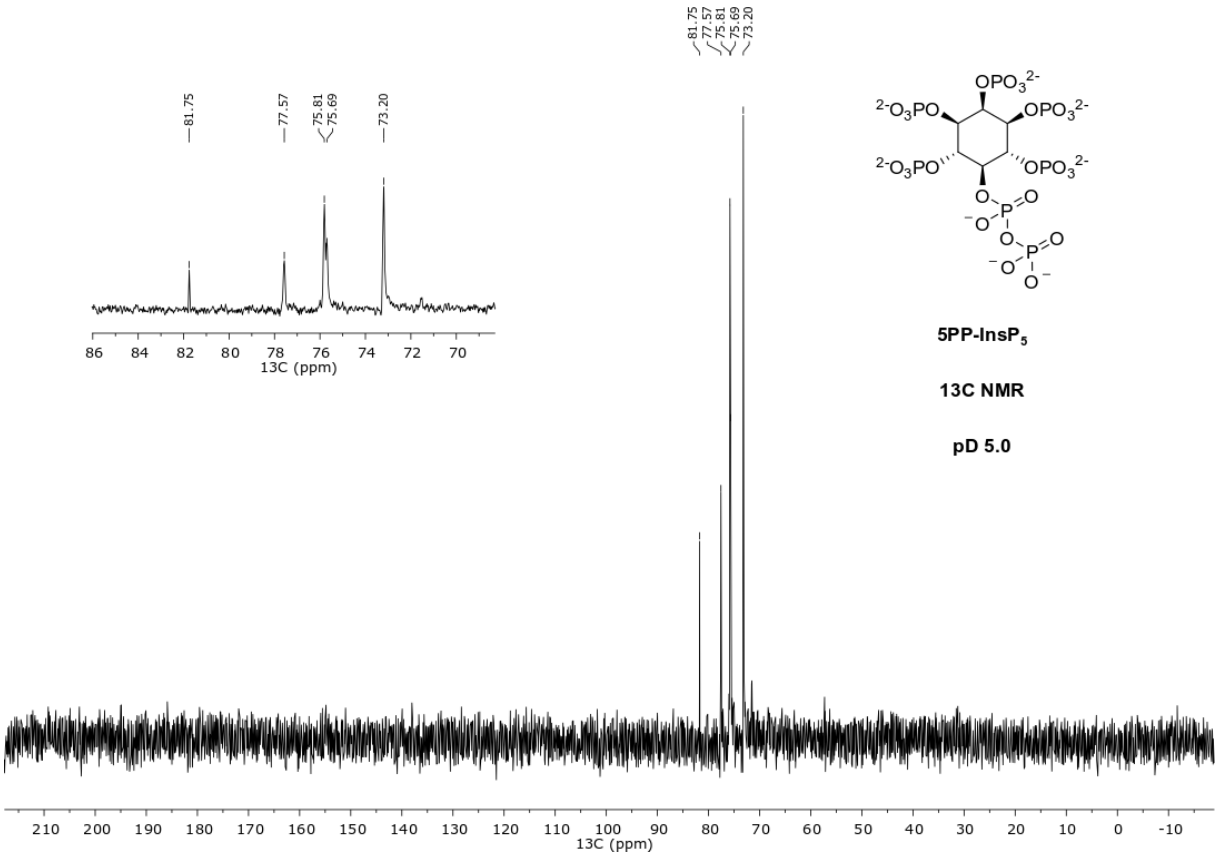
NMR Spectra



5PP-InsP$_5$

1H NMR

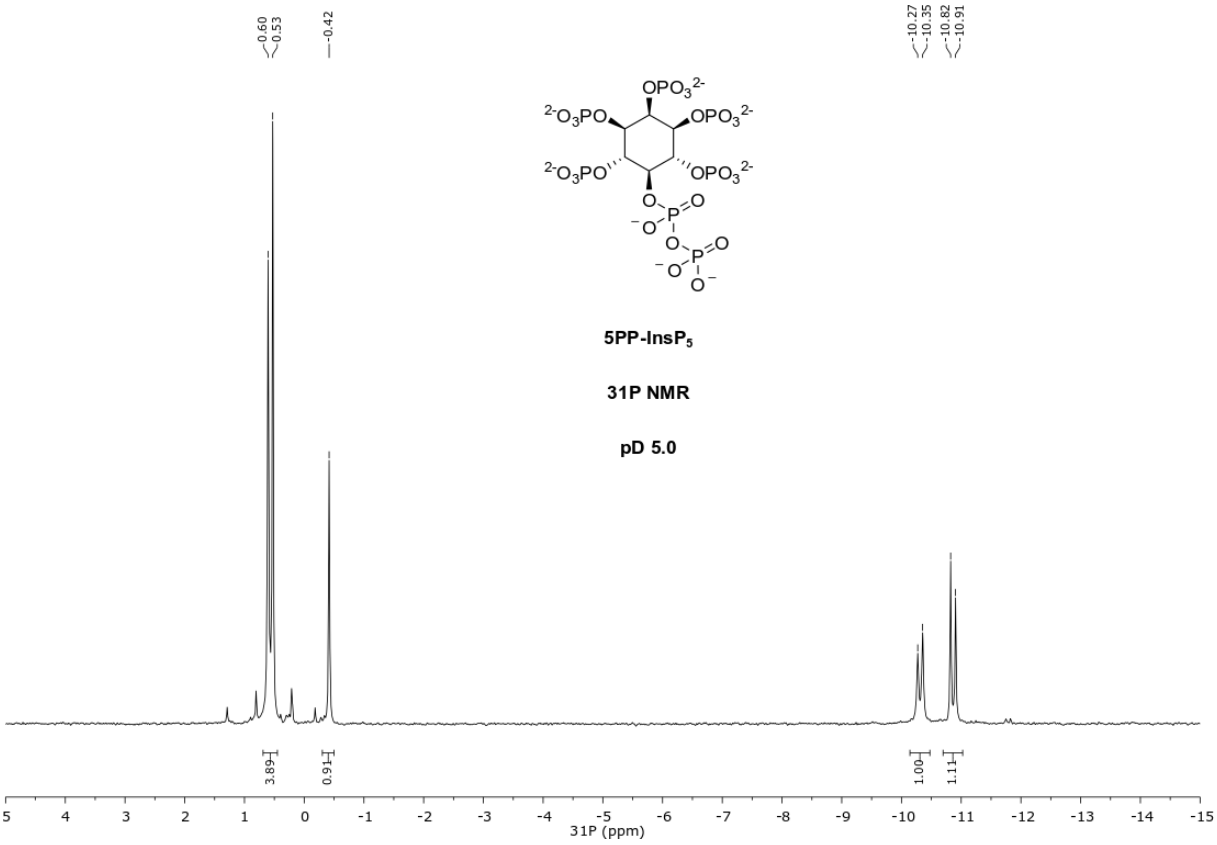
pD 5.0



5PP-InsP$_5$

13C NMR

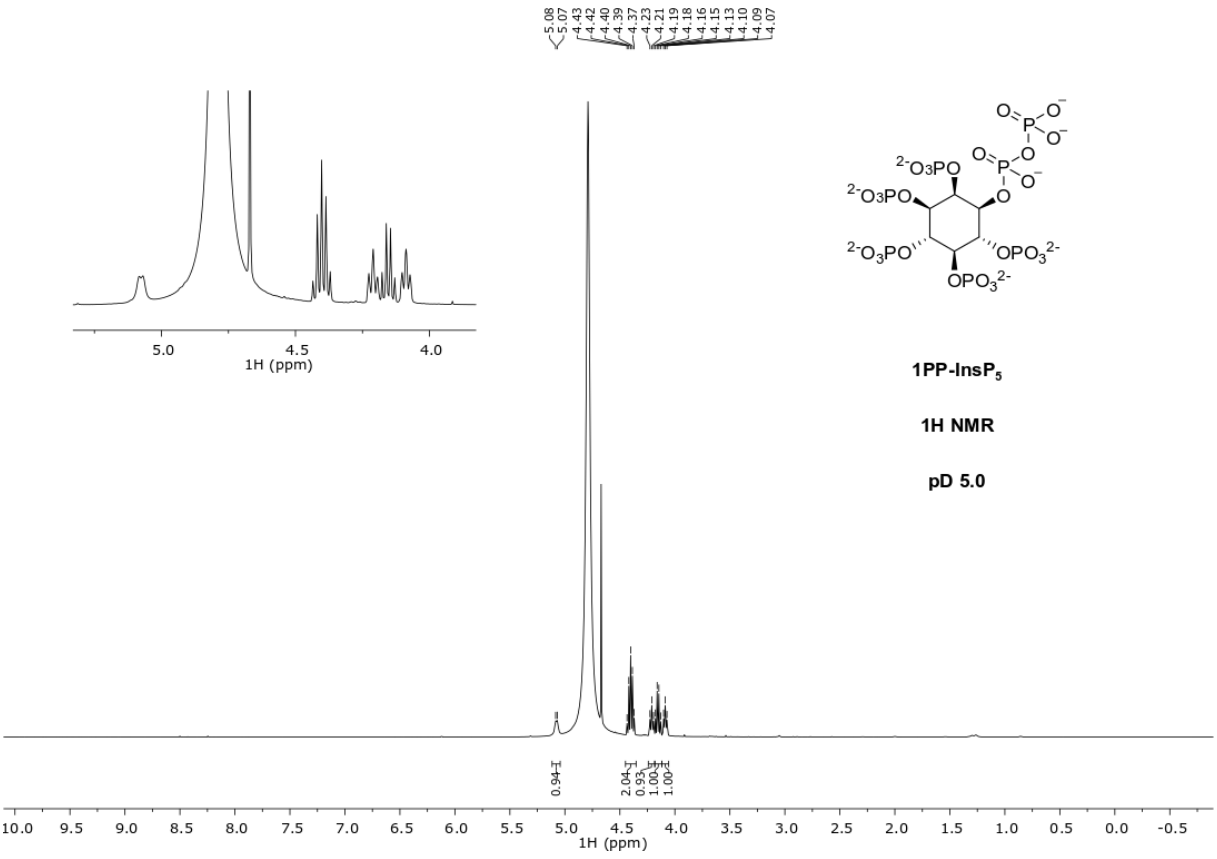
pD 5.0



5PP-InsP$_5$

31P NMR

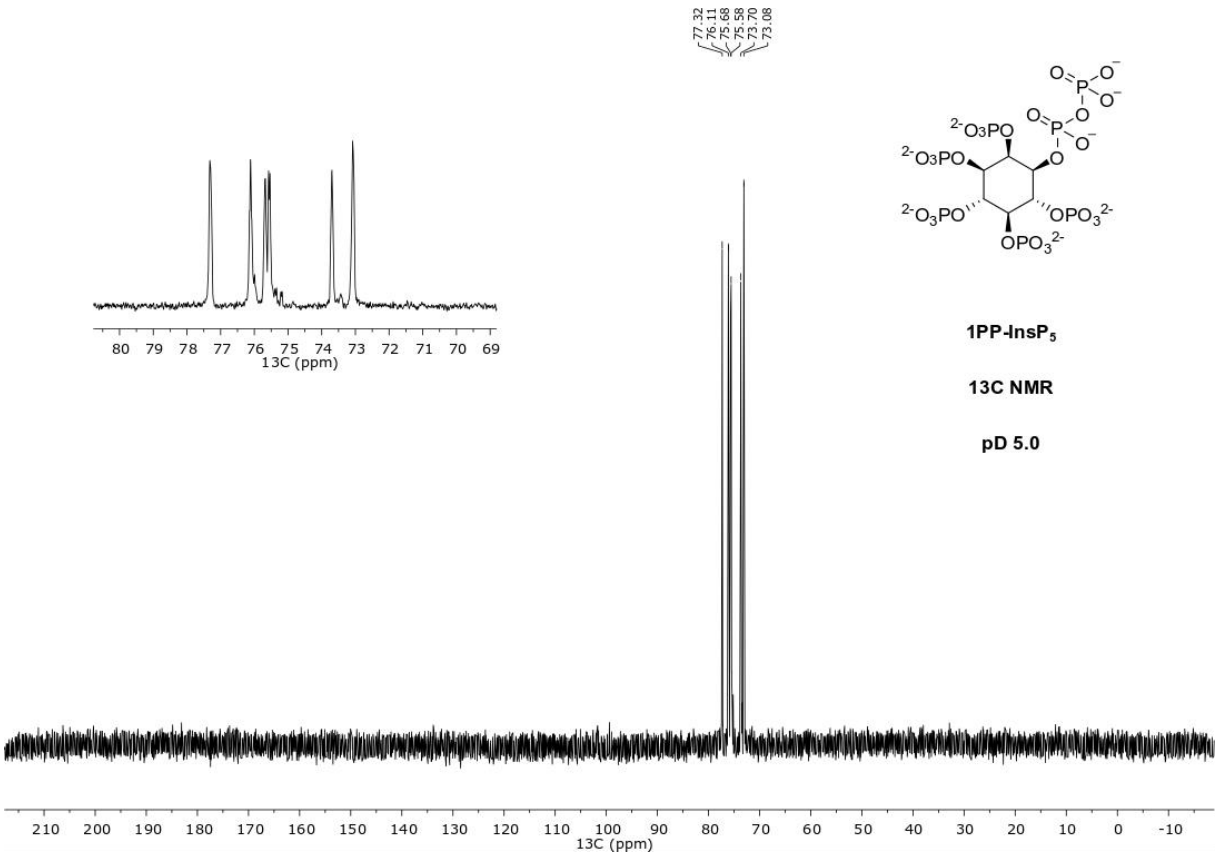
pD 5.0

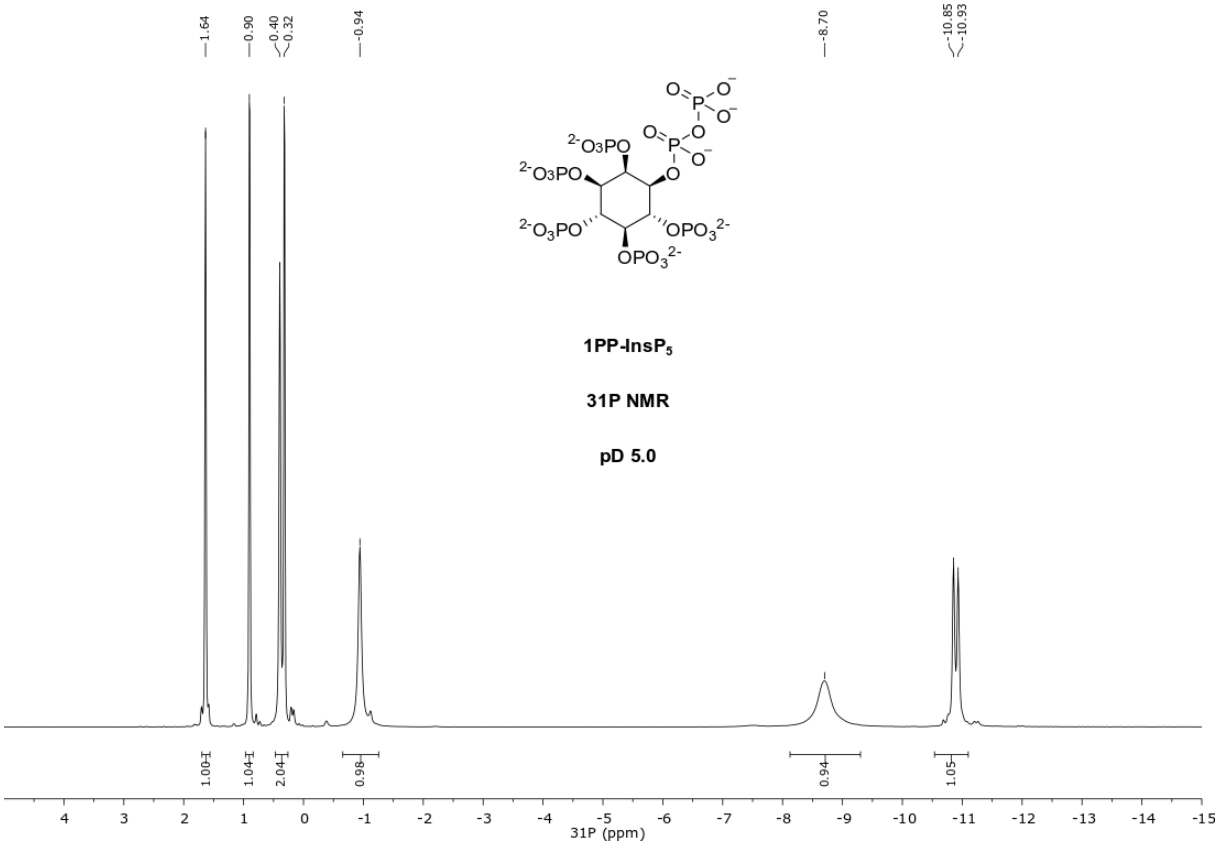


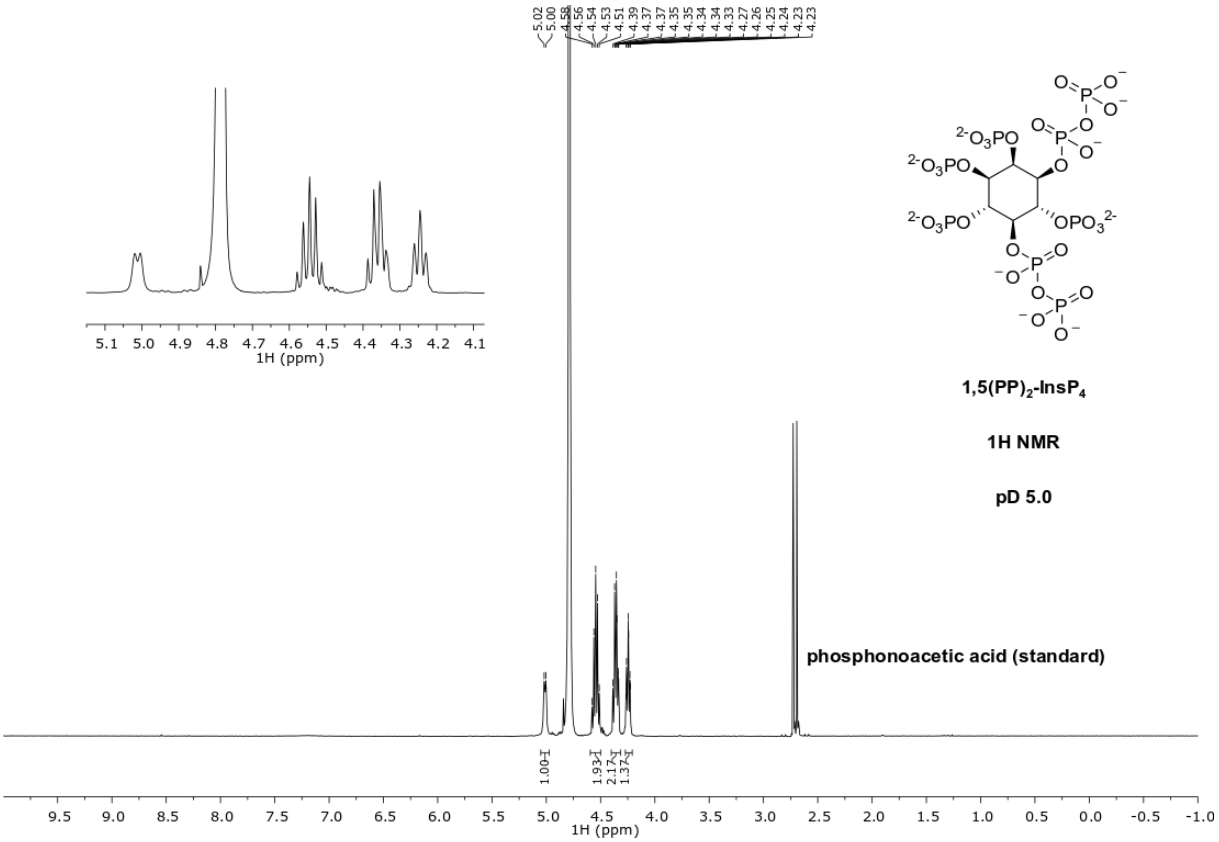
1PP-InsP$_5$

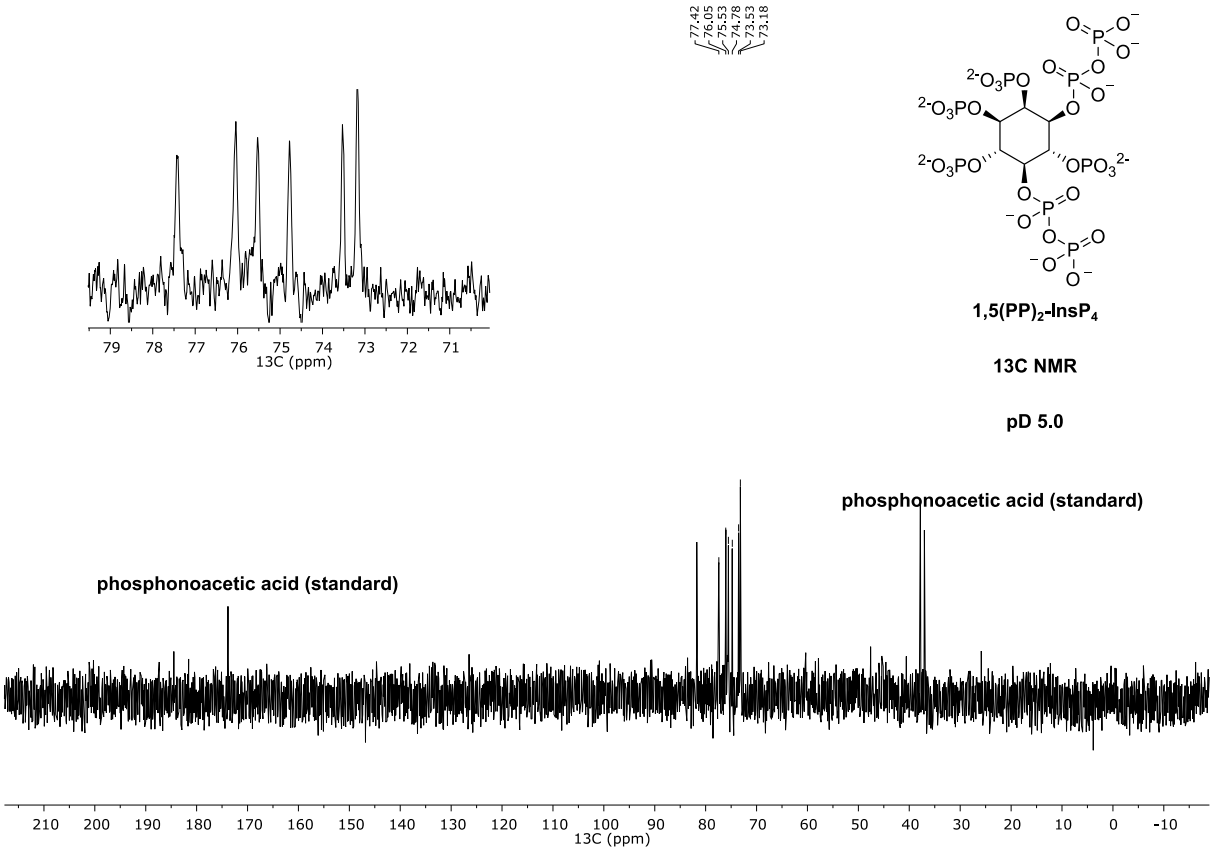
1H NMR

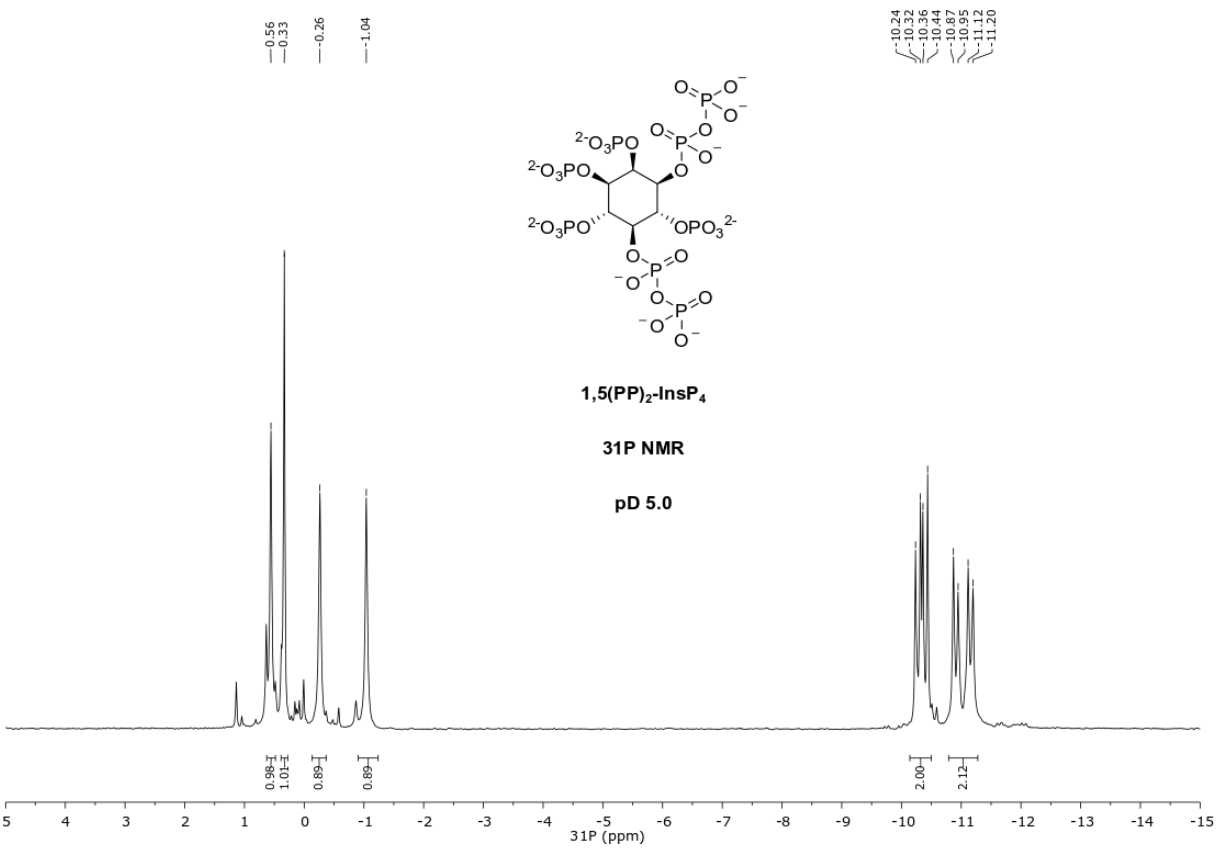
pD 5.0


77.32
76.11
75.68
75.58
73.70
73.08
80 79 78 77 76 75 74 73 72 71 70 69
13C (ppm)
1PP-InsP5
13C NMR
pD 5.0
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
13C (ppm)


1.64
0.90
0.40
0.32
0.94
8.70
10.85
10.93
1PP-InsP5
31P NMR
pD 5.0
1.00
1.04
2.04
0.98
0.94
1.05
4 3 2 1 0 -1 -2 -3 -4 -5 -6 -7 -8 -9 -10 -11 -12 -13 -14 -15
31P (ppm)


5.02
5.00
4.58
4.56
4.54
4.53
4.51
4.39
4.37
4.37
4.35
4.35
4.34
4.34
4.33
4.27
4.26
4.25
4.24
4.23
4.23
5.1 5.0 4.9 4.8 4.7 4.6 4.5 4.4 4.3 4.2 4.1
1H (ppm)
1,5(PP)2-InsP4
1H NMR
pD 5.0
phosphonoacetic acid (standard)
1.00
1.93
2.17
1.37
9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5 -1.0
1H (ppm)


77.42
76.05
75.53
74.78
73.53
73.18
79 78 77 76 75 74 73 72 71
13C (ppm)
1,5(PP)2-InsP4
13C NMR
pD 5.0
phosphonoacetic acid (standard)
phosphonoacetic acid (standard)
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
13C (ppm)


0.56
0.33
-0.26
-1.04
10.24
10.32
10.36
10.44
10.87
10.95
11.12
11.20
1,5(PP)2-InsP4
31P NMR
pD 5.0
0.98
1.01
0.89
0.89
2.00
2.12
5 4 3 2 1 0 -1 -2 -3 -4 -5 -6 -7 -8 -9 -10 -11 -12 -13 -14 -15
31P (ppm)

References

1. Riley, A. M., Godage, H. Y., Mahon, M. F. & Potter, B. V. L. Chiral desymmetrisation of myo-inositol 1,3,5-orthobenzoate gives rapid access to precursors for second messenger analogues. *Tetrahedron: Asymmetry* **17**, 171–174 (2006).
2. Capolicchio, S., Thakor, D. T., Linden, A. & Jessen, H. J. Angewandte Synthesis of Unsymmetric Diphospho-Inositol Polyphosphates **. 6912–6916 (2013). doi:10.1002/anie.201301092
3. Capolicchio, S., Wang, H., Thakor, D. T., Shears, S. B. & Jessen, H. J. Synthesis of Densely Phosphorylated Bis-1,5-Diphospho-myo-Inositol Tetrakisphosphate and its Enantiomer by Bidirectional P-Anhydride Formation. *Angew. Chemie Int. Ed.* **53**, 9508–9511 (2014).
4. Pavlovic, I. *et al.* Prometabolites of 5-Diphospho-myo-inositol Pentakisphosphate. *Angew. Chemie Int. Ed.* **54**, 9622–9626 (2015).
5. Rebekka Wild *et al.* Control of eukaryotic phosphate homeostasis by inositol polyphosphate sensor domains. *Science* **353**, 986–990 (2016).
6. Saiardi, A., Bhandari, R., Resnick, A. C., Snowman, A. M. & Snyder, S. H. Phosphorylation of proteins by inositol pyrophosphates. *Science* **306**, 2101–5 (2004).
7. Loss, O., Azevedo, C., Szijgyarto, Z., Bosch, D. & Saiardi, A. Preparation of quality inositol pyrophosphates. *J. Vis. Exp.* **2**, e3027 (2011).
8. Löser, B. *et al.* A novel Entamoeba histolytica inositol phosphate kinase catalyzes the formation of 5PP-Ins(1,2,3,4,6)P5. *Mol. Biochem. Parasitol.* **181**, 49–52 (2012).
9. Wang, H., Derose, E. F., London, R. E. & Shears, S. B. IP6K structure and the molecular determinants of catalytic specificity in an inositol phosphate kinase family. *Nat. Commun.* **5**, 1–12 (2014).
10. Podeschwa, M. *et al.* Stereoselective Synthesis of myo-, neo-, L-chiro, D-chiro, allo-, scyllo-, and epi-Inositol Systems via Conduritols Prepared from p-Benzoquinone. *European J. Org. Chem.* **2003**, 1958–1972 (2003).
11. Harmel, R. K. *et al.* Harnessing 13 C-labeled myo -inositol to interrogate inositol phosphate messengers by NMR. *Chem. Sci.* **10**, 5267–5274 (2019).
12. Veiga, N. *et al.* The behaviour of myo-inositol hexakisphosphate in the presence of magnesium(II) and calcium(II): Protein-free soluble InsP6 is limited to 49 micro M under cytosolic/nuclear conditions. *J. Inorg. Biochem.* **100**, 1800–1810 (2006).
13. Hager, A. *et al.* Cellular Cations Control Conformational Switching of Inositol Pyrophosphate Analogues. *Chem. A Eur. J.* **5**, 12406–12414 (2016).
14. Laussmann, T., Eujen, R., Weisshuhn, C. M. & Thiel, U. Structures of diphospho-myo-inositol pentakisphosphate and bisdiphospho- myo-inositol tetrakisphosphate from Dictyostelium resolved by NMR analysis. *Biochem. J.* **315**, 715–720 (1996).
15. Laussmann, T., Reddy, K. M., Reddy, K. K., Falck, J. R. & Vogel, G. Diphospho-myo-inositol phosphates from Dictyostelium identified as D-6-diphospho-myo-inositol pentakisphosphate and D-5,6-bisdiphospho-myo-inositol tetrakisphosphate. *Biochem. J.* **322**, 31–33 (1997).
16. Pisani, F. *et al.* Analysis of Dictyostelium discoideum Inositol Pyrophosphate Metabolism by Gel Electrophoresis. *PLoS One* **9**, e85533 (2014).
17. Puschmann, R., Harmel, R. K. & Fiedler, D. Scalable Chemoenzymatic Synthesis of Inositol Pyrophosphates. *Biochemistry* Article ASAP (2019). doi:10.1021/acs.biochem.9b00587
18. McNamara, D. E. *et al.* Direct Activation of Human MLKL by a Select Repertoire of Inositol Phosphate Metabolites. *Cell Chem. Biol.* **26**, 863-877.e7 (2019).

19. Steidle, E. A. *et al.* A novel inositol pyrophosphate phosphatase in Saccharomyces cerevisiae: Siw14 selectively cleaves the β-phosphate from 5-diphosphoinositol pentakisphosphate (5PP-IP5). *J. Biol. Chem.* **291**, 6772–6783 (2016).

20. Wang, H., Gu, C., Rolfes, R. J., Jessen, H. J. & Shears, S. B. Structural and biochemical characterization of Siw14: a protein-tyrosine phosphatase fold that metabolizes inositol pyrophosphates. *J. Biol. Chem.* **293**, 6905–6914 (2018).

21. Bhandari, R. *et al.* Bhandari, R., Saiardi, A., Ahmadibeni, Y., Snowman, A. M., Resnick, A. C., Kristiansen, T. Z., … Snyder, S. H. (2012). No Title, 109(51). *Proc. Natl. Acad. Sci.* **104**, 15305–15310 (2007).

22. Findeisen, M., Brand, T. & Berger, S. A 1H-NMR thermometer suitable for cryoprobes. *Magn. Reson. Chem.* **45**, 175–178 (2007).

23. Gregory, M. *et al.* Synthesis of a Tethered myo -Inositol (1,3,4,5,6)Pentakisphosphate (IP5) Derivative as a Probe for Biological Studies. *Synlett* **27**, 121–125 (2016).

24. Drent, E., Dijk, R. Van, Ginkel, R. Van, Oort, B. Van & Pugh, R. I. The first example of palladium catalysed non-perfectly alternating copolymerisation of ethene and carbon monoxide. *Chem. Commun.* 964–965 (2002).

25. Mayr, G. W. A novel metal-dye detection system permits picomolar-range h.p.l.c. analysis of inositol polyphosphates from non-radioactively labelled cell or tissue specimens. *Biochem. J.* **254**, 585–591 (1988).

Chapter 3 ^{13}C-Labeled *myo*-Inositol as a Tool to Study Inositol Pyrophosphates *In Vitro* and *Ex Vivo*

Reproduced in part with permission from (1) R. K. Harmel*, R. Puschmann*, M. Nguyen Trung, A. Saiardi, P. Schmieder, D. Fiedler, *Chem. Sci.* **2019,** 10, 5267–5274. (2) Jinsheng Zhu *et al*. *eLIFE* **2019**, 8, e43582.

(* contributed equally)

Project Collaborators: The experiments in this chapter were performed in close collaboration with R. Puschmann. Therefore, contributions of these authors are indicated more specifically as footnotes in the individual sections. Under supervision of R. K. Harmel, M. Nguyen Trung performed preliminary optimization experiments on the [$^{13}C_6$]5PP-InsP_5 synthesis and recorded time-resolved progress curves on IP6KA reactions. Experiments with enzymes from *A. thaliana* were performed in collaboration with the Hothorn and the Schaaf groups. *D. discoideum* was metabolically labeled by Saiardi and coworkers.

Introduction: New Methods to Monitor PP-InsPs in Complex Mixtures

Monitoring PP-InsPs levels and dynamics is one of the major readouts to identify and characterize biological functions of these second messengers. Since the first discovery of these small molecules, radioactively-labeled *myo*-inositol and PP-InsPs have been among the most applied tools to interrogate PP-InsP signaling. (**Chapter 1**, **Figure 1.4**) This approach established the existence of PP-InsPs as well as their participation in phosphate homeostasis and cell cycle progression.[1–4] The use of radioactive methods prevented many groups from entering the PP-InsP field and limited analysis of these small molecules to *yeast*, *D. discoideum* and a few mammalian cell lines. Therefore, new methods were developed over the last three decades. Early on, Mayr and coworkers used an alternative HPLC-based method that allowed label-free analysis of PP-InsP content from cell extracts by metal-dye detection (MDD).[5] (**Figure 3.1 a**)

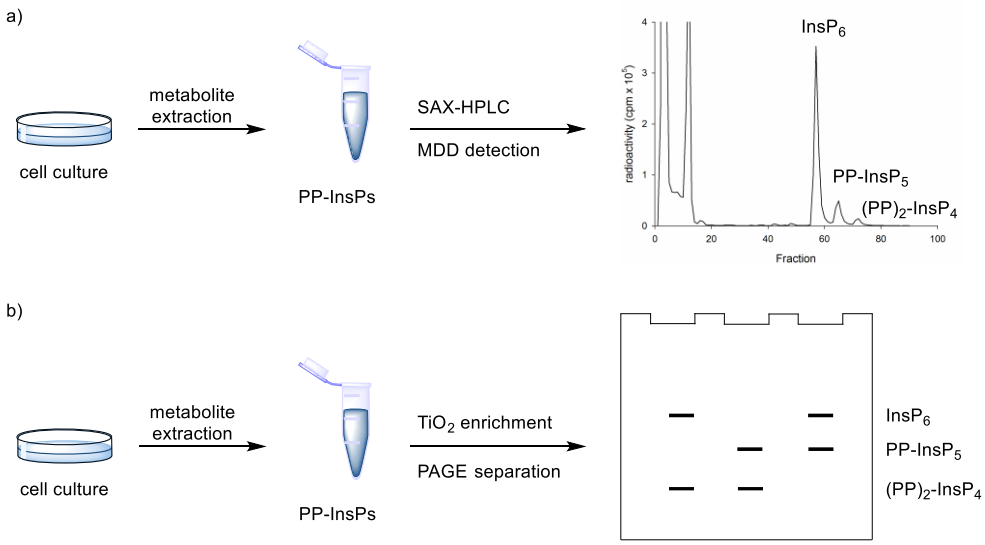


Figure 3.1: Alternative techniques to monitor PP-InsPs in cell extracts avoiding radioactivity. a) Analysis of PP-InsPs by SAX-HPLC using MDD detection. b) Analysis of PP-InsP by TiO_2 enrichment followed by PAGE separation and staining with toluidine blue.

The method relies on the detection of the specific absorption of a metal dye complex. In presence of InsPs and PP-InsPs, the metal-chelating dye is displaced from the metal centre, resulting in a decrease in absorbance of the complex. This approach allowed relative and absolute quantification and has been used since the initial characterization of PP-InsPs in *D. discoideum*.[2] MDD requires a three pump HPLC system which is not readily implemented. Eversince this method was published, only the Mayr laboratory had this setup available which made it overall sparsly used in the field. More recently, a new method avoided the use of rather harsh HPLC conditions for the separation.[6] PP-InsPs were enriched from acidified cell extracts by TiO_2 beads, followed by their separation on high percentage acrylamide gels. (**Figure 3.1 b**) Unlabeled PP-InsPs were visualized using a cationic dye like toluidine blue that binds PP-InsPs and were analyzed quantitatively using a gel imager. This method showed systematically increased PP-InsP concentrations compared to radiolabeling and MDD, which can be attributed to the milder separation conditions.[6] The relatively low technical requirements and the independence from radioactive tracers have allowed more laboratories to employ this approach, but the lack of an InsP-specific analytical handle entailed a more elaborate sample preparation and is limited to the detection of only the most highly phosphorylated InsPs.

While all methods mentioned above have been instrumental to understand PP-InsP function, they are only able to discriminate InsP and PP-InsP based on their charge and lack a clear distinction of the structural isomers of InsPs and PP-InsPs. These methods also lack the ability to monitor conversion of InsP species in real time *in vitro*, thereby losing informative kinetic insight contained within biochemical experiments. Furthermore, various separation and resolution steps were required and direct measurements in complex environments like cell extracts have not been possible so

far. Therefore, novel approaches that discriminate PP-InsPs based on their structure and that generate reliable information quickly with limited sample preparation are highly needed.

NMR spectroscopy is a non-invasive analytical method that can provide detailed information about the chemical and structural environment of a nucleus solely based on the chemical shift and magnetic couplings between nuclei. For example, two-dimensional $^{1}H,^{31}P$-NMR has been used to elucidate the structure of PP-InsPs.[7,8] Since experiments utilizing the $^{3}J(^{1}H,^{31}P)$ coupling are of limited sensitivity due to the inefficient magnetization transfer *via* these couplings, high analyte concentrations were necessary. In addition, the low chemical shift dispersion and the broad lines of ^{31}P can limit the information content.[7,8] In contrast, the chemical shift dispersion of the NMR-active nucleus ^{13}C is superior to ^{31}P and magnetization transfer *via* $^{1}J(^{1}H,^{13}C)$ one-bond couplings is much more efficient. Unfortunately, the sensitivity of NMR experiments including ^{13}C is strongly hampered by the low natural abundance of this isotope of approx. 1 %. In the past, this disadvantage could be overcome by the use of ^{13}C isotope-labeled small molecules and has been widely applied to elucidate various metabolic pathways and processes.[9–11] The 100-fold higher concentration of ^{13}C nuclei enhanced the detectability of labeled compounds with respect to sample background and experimental noise, thereby providing targeted information on these molecules in complex environments. Consequently, we envision that ^{13}C-labeled *myo*-inositol, and the corresponding labeled InsPs and PP-InsPs, could enable NMR measurements at low, biologically relevant analyte concentrations and within intricate samples.

In this chapter, a new a tool that combines NMR spectroscopy and ^{13}C-labeled *myo*-inositol to study InsP and PP-InsPs is reported. To this end, a high yielding, gram-scale synthesis of highly pure uniformly ^{13}C-labeled ([$^{13}C_6$]) *myo*-inositol was developed. In combination with synthetic procedures from the literature and the chemoenzymatic efforts, provided in chapter 2, [$^{13}C_6$]InsPs and [$^{13}C_6$]PP-InsPs were prepared. We employed the [$^{13}C_6$]InsPs and [$^{13}C_6$]PP-InsPs to monitor *in vitro* activity of PP-InsP metabolizing enzymes from *H. sapiens* and *A. thaliana* and characterized the kinetic parameters of these kinases. Moreover, we demonstrate metabolic labeling of two mammalian cell lines (HCT116 and HEK293T) and the social amoeba *D. discoideum* with [$^{13}C_6$]*myo*-inositol, and subsequent detection of the *in vivo* generated [$^{13}C_6$]InsPs and [$^{13}C_6$]PP-InsPs species within a complex cell extract. These findings highlight the utility and applicability of ^{13}C-labeled InsPs to improve our understanding of InsP and PP-InsP biochemistry and metabolism, and to interrogate and elucidate their cellular functions.

Chemoenzymatic Synthesis of ^{13}C-labeled PP-InsPs

Enzymatic Synthesis of [$^{13}C_6$]myo-Inositol on a Gram Scale[1]

To analyze inositol-based signaling molecules by NMR we first sought to prepare pure [$^{13}C_6$]*myo*-inositol on a large scale. The approaches to obtain ^{13}C-labeled *myo*-inositol reported to date have either relied on expensive [$^{13}C_6$]glucose-6-phosphate as a starting material (7000 €/g) or – when starting from [$^{13}C_6$]glucose – were low yielding (9.9 % and 6.7 %, respectively).[12–14] In addition, all procedures lacked a scalable purification strategy for the product, precluding the isolation of large

[1] Optimization of the enzymatic synthesis was conducted by R. Puschmann. Optimization of the chemical derivatization and purification was performed by R. K. Harmel. [$^{13}C_6$]*myo*-inositol was synthesized on a large scale by R. Puschmann and R. K. Harmel.

and pure amounts of product. Consequently, these methods have only been applied on a small scale. To circumvent these limitations, and to subsequently synthesize sufficient quantities of [$^{13}C_6$]InsPs, we pursued a 3-step enzymatic synthesis that uses affordable [$^{13}C_6$]glucose (80 €/g) as a starting material and employed a chemical purification strategy at the end. (**Figure 3.2**)

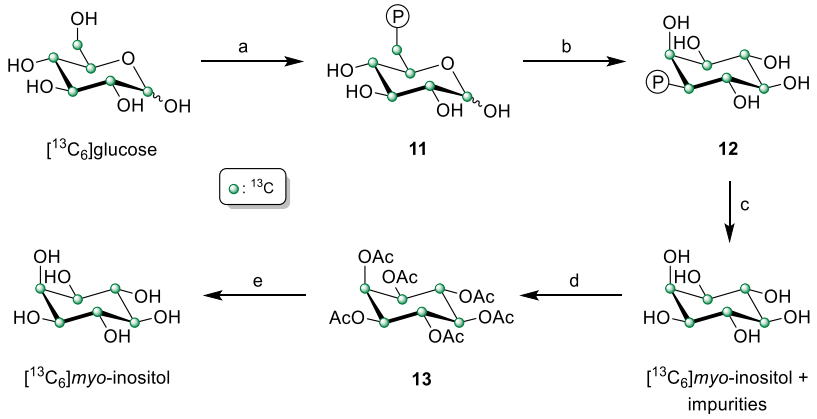


Figure 3.2: Chemoenzymatic synthesis of [$^{13}C_6$]*myo*-inositol followed by derivatization and purification. a) [$^{13}C_6$]glucose, MOPS pH 6.5, creatine phosphate, ATP, DTT, $MgCl_2$, hexokinase, creatine kinase, H_2O. b) Tris pH 8.0, NAD^+, NaCl, $MgCl_2$, inositol phosphate synthase in H_2O. c) Glycine pH 9.8, $ZnCl_2$, alkaline phosphatase in H_2O. d) Ac_2O, pyridine. e) NaOMe in MeOH.

As a first step, [$^{13}C_6$]glucose was phosphorylated at the 6-position using commercially available hexokinase in a modified literature procedure.[15] Next, we optimized and scaled up the isomerization reaction of [$^{13}C_6$]glucose-6-phosphate (**11**) to [$^{13}C_6$]*myo*-inositol-3-phosphate (**12**) published by Saiardi and coworkers.[12] During this step the conditions were altered to use Mg^{2+} rather than paramagnetic Mn^{2+}, so the progress of the reaction could be monitored by ^{13}C NMR spectroscopy. Securing full conversion to **12** was crucial for the overall synthetic strategy, as it bypasses the discrimination between **11** and **12** in the subsequent dephosphorylation step. Thereby, [$^{13}C_6$]*myo*-inositol could be obtained by dephosphorylation with robust and commercially available alkaline phosphatase instead of using recombinantly

expressed inositol monophosphatase. While the desired product was clearly visible in the ^{13}C NMR spectrum, the isolated yield exceeded 100%, indicative of contamination with salts and buffer components. Because purification of [$^{13}C_6$]*myo*-inositol by ion-exchange, normal- phase, and reverse-phase chromatography proved difficult, we applied a chemical derivatization strategy in which the product mixture containing [$^{13}C_6$]*myo*-inositol was treated with an excess of acetic anhydride in pyridine. The resulting [$^{13}C_6$]*myo*-inositol hexakisacetate (**13**) could be purified by normal-phase chromatography and subsequent saponification and precipitation provided gram amounts of [$^{13}C_6$]*myo*-inositol in excellent purity with a 50% overall yield (with respect to [$^{13}C_6$]glucose). In comparison to previous syntheses, our approach provides a simpler and shorter route with greatly improved yield.

Chemical Synthesis and Characterization of [$^{13}C_6$]InsPs and [$^{13}C_6$]PP-InsPs

With large quantities of pure [$^{13}C_6$]*myo*-inositol in hand, we prepared [$^{13}C_6$]$InsP_5$ and [$^{13}C_6$]$InsP_6$; as they are the most abundant inositol polyphosphates.[16] (**Figure 3.3**) [$^{13}C_6$]$InsP_6$ was obtained analogous to the procedure used in chapter 2 in two steps with 35% yield *via* the xylyl protected intermediate **14**.[17] Unfortunately, efforts to optimize the solvent system or use of other reported protecting groups (9-fluoenylmethyl- or benzyl) led to no significantly improved yield.[18] The synthesis of [$^{13}C_6$]$InsP_5$ consists of 5 steps and was optimized using unlabeled *myo*-inositol, then performed with [$^{13}C_6$]*myo*-inositol. The synthetic route started with the formation of the orthobenzoate ester **15**.[19] While the published procedure suggested recrystallization from EtOAc, we found that precipitation with water simplified the product purification due to the easy removal of the DMSO. Acidic hydrolysis of **15** followed by phosphitylation with xylyl phosphoamidite generated and oxidation by *m*CPBA

compounds **16**. Removal of the xylyl- and the benzoyl groups afforded [$^{13}C_6$]InsP$_5$ in 38% overall yield which is consistent with the literature yields.

Compared to the syntheses of [$^{13}C_6$]InsP$_5$ and [$^{13}C_6$]InsP$_6$, obtaining [$^{13}C_6$]PP-InsPs by chemical means would require large amounts of starting material and be quite challenging.[20–22] Fortunately, we could easily adopt the new findings described in chapter 2 to produce large amounts PP-InsPs enzymatically in high purity. Using this method, 100 mg of [$^{13}C_6$]5PP-InsP$_5$ and 50 mg of each [$^{13}C_6$]1PP-InsP$_5$ and [$^{13}C_6$]1,5(PP)$_2$-InsP$_4$ were synthesized chemoenzymatically.

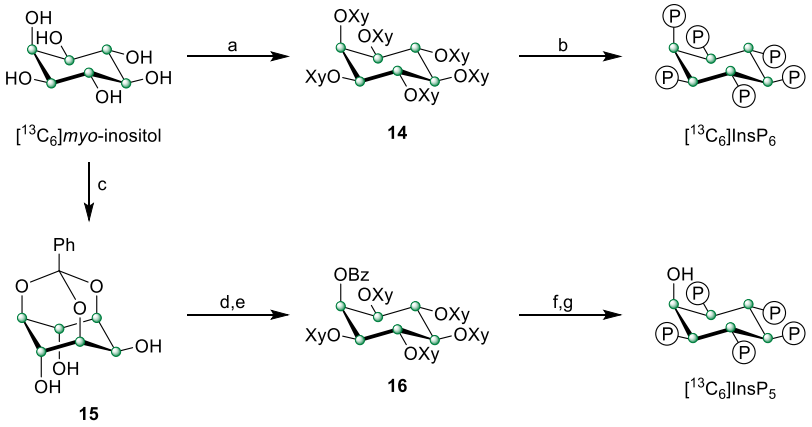


Figure 3.3: Chemical synthesis of [$^{13}C_6$]InsP$_5$ and [$^{13}C_6$]-InsP$_6$. a) Xy phosphoamidite, 5-Phenyltetrazol, DCM, then mCPBA. b) Pd(OH)$_2$/H$_2$, MeOH/H$_2$O. c) trimethyl orthoformate, camphorsulfonic acid, DMSO. d) 10% TFA. e) Xy phosphoramidite, 1H-tetrazole, mCPBA, DCM. f) Pd(OH)$_2$/H$_2$, MeOH/H$_2$O. g) conc. NH$_{3(aq)}$

We examined all prepared ^{13}C-labeled compounds by ^{1}H,^{13}C-HMQC spectroscopy. In general, ^{13}C and ^{1}H nuclei of phosphorylated positions are more deshielded compared to their unphosphorylated counterparts and displayed higher chemical shifts which was most pronounced in the ^{1}H-dimension. (**Figure 3.4 a, b**) The spectra of [$^{13}C_6$]*myo*-inositol, [$^{13}C_6$]InsP$_5$, [$^{13}C_6$]InsP$_6$ and [$^{13}C_6$]5PP-InsP$_5$ showed four resonances due to the plane of symmetry in these molecules. (**Figure 3.4 a–d**) [$^{13}C_6$]1PP-InsP$_5$ and [$^{13}C_6$]1,5(PP)$_2$-InsP$_4$ exhibited six signals as the symmetry is

broken through the pyrophosphate group at the 1-position and the signals at the 1/3 positions are clearly split. (**Figure 3.4 e, f**) Importantly, the spectra revealed unique peaks for each of the InsPs which was very promising with respect to the analysis of complex mixtures and unpurified cell extracts. Especially, the chemical shift of the 2-position was strongly affected by the given phosphorylation pattern. Remarkably, spectra could be obtained using only 32 scans (2 h) at 5 µM for [$^{13}C_6$]InsP$_6$ which illustrates the high visibility provided by labeling with ^{13}C.

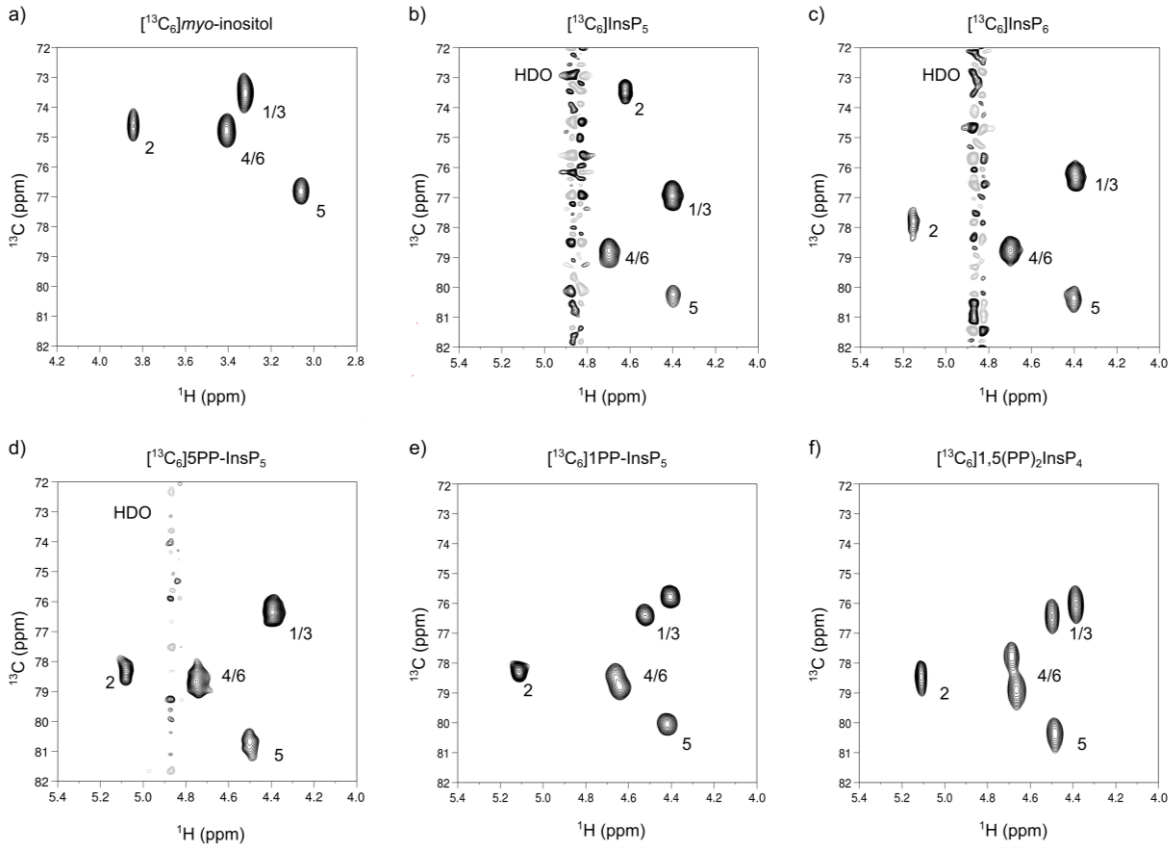


Figure 3.4: $^{1}H,^{13}C$-HSQC spectra of all synthesized ^{13}C-labeled compounds. a) [$^{13}C_6$]*myo*-inositol b) [$^{13}C_6$]InsP$_5$ c) [$^{13}C_6$]InsP$_6$ d) [$^{13}C_6$]5PP-InsP$_5$ e) [$^{13}C_6$]1PP-InsP$_5$ f) [$^{13}C_6$]1,5(PP)$_2$-InsP$_4$.

In Vitro Characterization of PP-InsP Metabolizing Enzymes

Biochemical Characterization of IP6KA and IP6K1[2]

With the synthetic ^{13}C-labeled standards in hand, we pursued the *in vitro* characterization of PP-InsP kinases by NMR spectroscopy. To date, enzymatic activities of IP6Ks and PPIP5Ks were mostly assessed by monitoring PP-InsP conversion using radioactivity or MDD, or were measured indirectly *via* quantification of ADP content using bioluminescent assays.[1–5,23] We envisioned that an NMR-based method could resemble a valuable alternative to these methods. The NMR workflow would avoid the use of radioactive compounds and cumbersome HPLC purification. In addition, atomic resolution provided by the NMR spectra would give high confidence about the identity of the substrate. Overall, this could simplify sample preparation and analysis, and could minimize the risk of false interpretation of the obtained data.

As proof of principle, the kinase activities of IP6KA and IP6K1 were investigated. While optimizing the enzymatic synthesis of [$^{13}C_6$]5PP-InsP$_5$ with IP6KA, we noted that the progression of the reaction could be followed readily by NMR, even in the presence 10% H_2O and high concentrations of non-deuterated buffer and ATP (**Figure 3.5**). Specifically, the diagnostic signals for the protons at the 2-position of the inositol ring of [$^{13}C_6$]InsP$_6$ and [$^{13}C_6$]5PP-InsP$_5$ were baseline-separated in the ^1H-dimension (5.2 ppm for [$^{13}C_6$]InsP$_6$, 5.1 ppm for [$^{13}C_6$]5PP-InsP$_5$) and showed no overlap with signals from other compounds within the mixture.

[2] Optimization and biochemical characterization of IP6K1 was performed by both R. K. Harmel and R. Puschmann. Kinetic data were analyzed and illustrated by R. Puschmann.

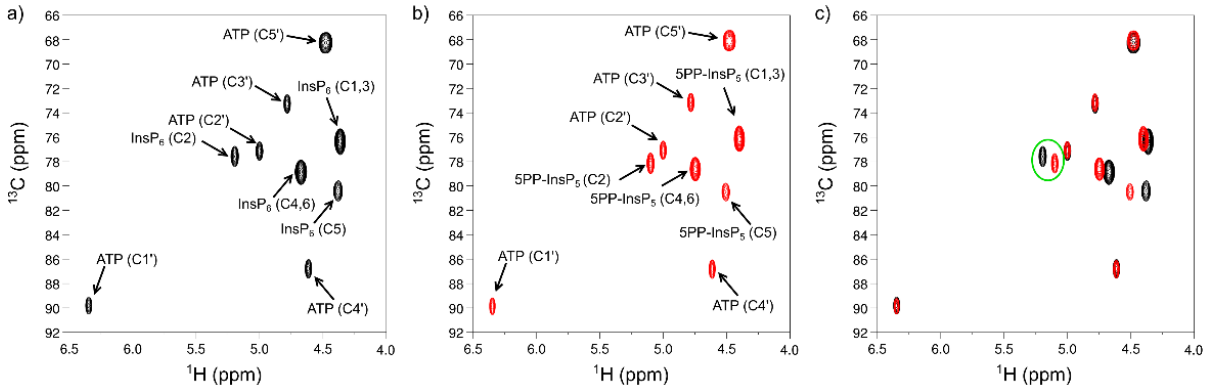


Figure 3.5: $^{1}H,^{13}C$-HMQC-Spectra to determine conversion to [$^{13}C_6$]5PP-InsP$_5$. a) Control reaction without IP6KA shows no conversion of [$^{13}C_6$]InsP$_6$. b) Full conversion of [$^{13}C_6$]InsP$_6$ to [$^{13}C_6$]5PP-InsP$_5$. c) Overlay of a and b to illustrate that the C2 peaks (highlighted with green circle) are well resolved and can be used to monitor the reaction progress.

The ability to resolve the ^{1}H-signals at the 2-position provided the opportunity to monitor the enzymatic reaction in a time-resolved fashion, using a pseudo-2D spin-echo difference experiment. This pulse sequence was similar to the $^{1}H,^{13}C$-HMQC experiments and resulted in a series of one-dimensional spectra that only displayed resonances for ^{1}H-nuclei that are bound to ^{13}C-nuclei. By plotting the peak intensity of the ^{1}H-nuclei at the 2-position against time (each experiment required only 75 seconds at a substrate concentration of 175 µM) the progress of the kinase reaction could be observed. (**Figure 3.6**) The non-invasive nature of the NMR measurements allowed for continuous reaction monitoring, providing time-resolved data from one sample.

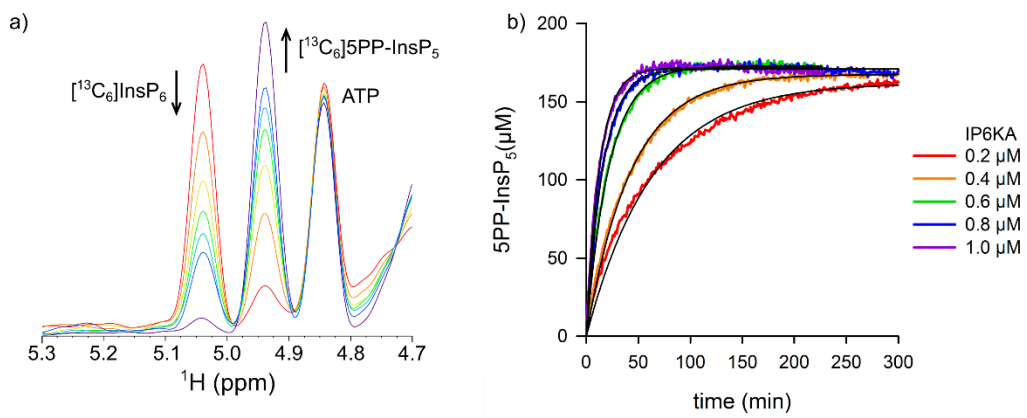


Figure 3.6: NMR-based measurements of IP6KA activity. a) Superimposed pseudo-2D spin-echo difference spectra at different time points (6 min, 28 min, 50 min, 73 min, 95 min, 118 min, 455 min). The $InsP_6$ and 5PP-$InsP_5$ peaks are labeled and the change over time is indicated by the arrows. The intensity of the ATP-peak hardly changed due to the ATP regenerating system. b) Progress curves of [$^{13}C_6$]5PP-$InsP_5$ synthesis at different IP6KA concentrations. An individual NMR-spectrum was recorded every 75 seconds. The [$^{13}C_6$]5PP-$InsP_5$ peak height was normalized against the HEPES signal and the data was fitted with a first order decay model.

With a fast and robust workflow to determine kinase activity set up, the kinetic parameters for human IP6K1 were determined to validate our NMR analysis. (**Figure 3.7**) Measurement of the initial rates at constant [$^{13}C_6$]$InsP_6$ concentration (175 µM) and varying ATP concentrations (62.5 µM to 8 mM) provided the Michalis-Menten constant. In this assay, we decided to use time points rather than continuous time curves in order to save experimental and instrument time. IP6K activity is dependent on Mg^{2+} ions and therefore we chose to quench each time point by addition of EDTA. The K_M measured for ATP (353 ± 167 µM) was comparable to previous reports and is close to cellular ATP concentrations, which supports the assertion that the IP6Ks are uniquely sensitive to ATP availability.[23,24] The V_{max} value (192 ± 41 nmol min^{-1} mg^{-1}) for IP6K1 was also within the range of published literature values, which varied greatly depending on the protein-tag and purification strategy employed (37.3 nmol min^{-1} mg^{-1} to 1410 nmol min^{-1} mg^{-1}).[24,25] Interestingly, we found that ATP concentrations above

2 mM led to substrate inhibition of IP6K1 (K_i 7.48 mM), which had not been quantified before.[25] The robust measurements of InsP conversion by NMR spectroscopy has provided kinetic data with low experimental error, which, in turn, allowed the observation of substrate inhibition by ATP.

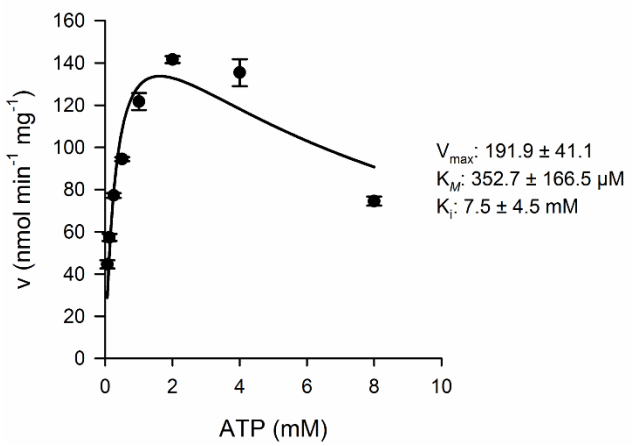


Figure 3.7: Kinetic characterization of IP6K1. Michaelis-Menten kinetics for IP6K1. The initial velocity was measured in triplicates at different ATP concentrations and fitted against a model for substrate inhibition.

All in all, biochemical characterization of IP6K1 showed great potential to monitor conversion of PP-InsPs in general. Consequently, this approach became the method of choice to optimize reaction conditions for the large-scale synthesis of PP-InsPs described in chapter 2 and to compensate for batch-dependent variations in enzyme activity. In the future, the developed NMR workflows can be adapted for the analysis of other PP-InsPs metabolizing enzymes.

Evaluation of IP6K1 Inhibitors[3]

IP6Ks regulate important aspects of metabolism and signaling in mammals and in pathogenic microorganisms like *Trypanosoma brucei* and therefore increased efforts to identify selective inhibitors have been made.[23,26,27] Thus far, two molecules, TNP (N^2-(*m*-Trifluorobenzyl)-N^6-(*p*-nitrobenzyl)purine), and Myricetin were shown to inhibit IP6K1 activity at low micromolar concentrations. TNP was identified in a study that tested $InsP_3$-3-kinase inhibitors on IP6Ks.[26] Here, the conversion of radioactive [^{3}H]$InsP_6$ was followed by TLC or HPLC assays. In an attempt to find more potent inhibitors for IP6K1, a high-throughput-screening approach using a bioluminescence-based ADP detection assay identified myricetin among a few others hits.[23] Given that we could directly measure product formation, we expanded the scope of our NMR method towards the characterization of IP6K1 inhibitors. (**Figure 3.8**)

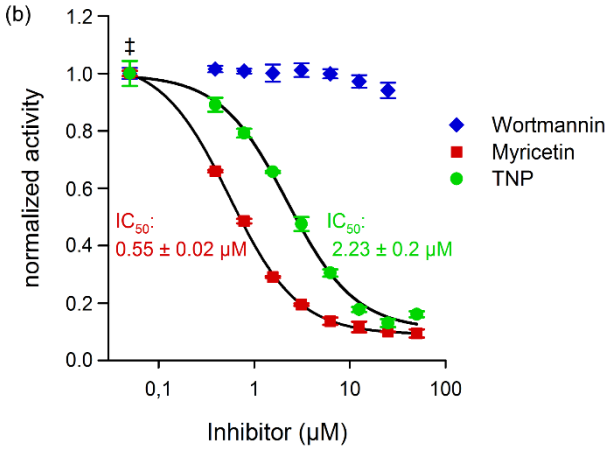


Figure 3.8: The inhibitory effect of TNP (N^2-(*m*-Trifluorobenzyl)-N^6-(*p*-nitrobenzyl)purine), Myricetin, and Wortmannin were tested at different concentrations and the IC_{50} values were determined. The ATP concentration was 2.5 mM for all inhibitor measurements. (‡ enzyme activity at 0 µM inhibitor).

[3] Evaluation of inhibitors was performed by both R. K. Harmel and R. Puschmann. Inhibition data were analyzed and illustrated by R. Puschmann.

Indeed, we found that TNP and myricetin inhibited IP6K1 with IC_{50} values of 2.25 µM and 0.6 µM, respectively, while Wortmannin (a PI3K inhibitor) did not affect enzyme activity even at high concentrations. All in all, we found that the NMR-based method represented a good alternative to the radioactive assay and the direct read-out of the substrate and the reaction product provided data with high certainty.

Evaluation of the ATP Synthase Reaction as a Tool for High-Throughput-Screening of IP6K1 Inhibitors

Our newly established NMR assay was not suitable to rapidly screen for new inhibitors of PP-InsPs kinases. However, the high reliability of our approach inspired us to utilize this technique as a tool to develop and validate new high-throughput-screening platforms for the discovery of selective IP6K inhibitors. Such inhibitors are potentially of value as lead compounds for pharmaceutical intervention in PP-InsP-mediated diseases or as chemical probes for further interrogation of PP-InsPs signaling. In 2017, Wei and coworkers developed a homogeneous high-throughput screen for IP6K1 inhibitors.[23] Kinase activity was assessed by the detection of the produced ADP using a commercially available ADP-Glo™ assay kit. (**Figure 3.9**)

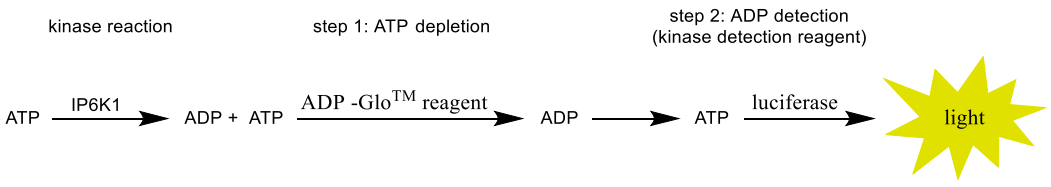


Figure 3.9: ADP detection by commercially available ADP-Glo™ assay.

Unfortunately, the most potent hits identified in this screen were flavonoids like myricetin which are known to interact with many other kinases.[28–30] This polypharmacological behaviour disqualifies these compounds as chemical probes to

study PP-InsP signaling in cells. Therefore, new approaches to identify IP6K specific or even IP6K isoform specific inhibitor are highly needed.

In previous NMR experiments to optimize the synthesis of [$^{13}C_6$]5PP-InsP$_5$, it was found that the kinase reaction would reach an equilibrium at roughly 30% conversion of [$^{13}C_6$]InsP$_6$, even though the concentration of the generated ADP was low. According to Snyder and coworkers, the reverse reaction, also referred to as ATP synthase reaction, was strongly dependent on the ADP concentration (K_M for ADP: 1.57 mM) and displayed a V_{max} value roughly twice as high, compared to the forward reaction.[31] Together these results indicated that the ATP synthase reaction could be readily favored over the kinase reaction. Since we were now able to produce large amounts of 5PP-InsP$_5$ with the methods developed in Chapter 2, the reverse reaction of IP6K1 could be explored as a platform to screen inhibitors for this enzyme.

Most kinase inhibitors compete for the nucleotide binding site and we postulated that the K_D for ATP and ADP should similar so our new approach could work. Unfortunately, these values were not published so there was no biochemical evidence that the aforementioned assumption was correct. However, an array of experiments was conducted to support the initial hypothesis. To this end, we first validated that ADP is not an inhibitor of the forward reaction but rather induces the reversed reaction. Therefore, the ATP synthase reaction had been performed with ADP and [$^{13}C_6$]5PP-InsP$_5$ and the formation of ATP and InsP$_6$ was observed by NMR. (**Figure 3.10 a**, **b**) A reaction with 20 µM TNP was included in one of these samples and showed a reduction of ATP synthase activity by 78% compared to the DMSO control.

The indication that TNP is indeed an inhibitor of the reverse reaction motivated us to investigate the compatibility of this reaction with the Kinase Glo^{TM} assay that monitors ATP synthesis. (**Figure 3.10 c**) The commercial kit was adapted to a 96-well

plate format using total reaction volumes of 200 µL. Triplicate experiments were set-up to test the activity of IP6K1 against DMSO and TNP, including two controls either lacking 5PP-InsP$_5$ or enzyme. (**Figure 3.10 d**) After 30 min, the Kinase Glo^{TM} reagent was added directly to the wells, and luminescence intensity was analyzed using a gel imager.

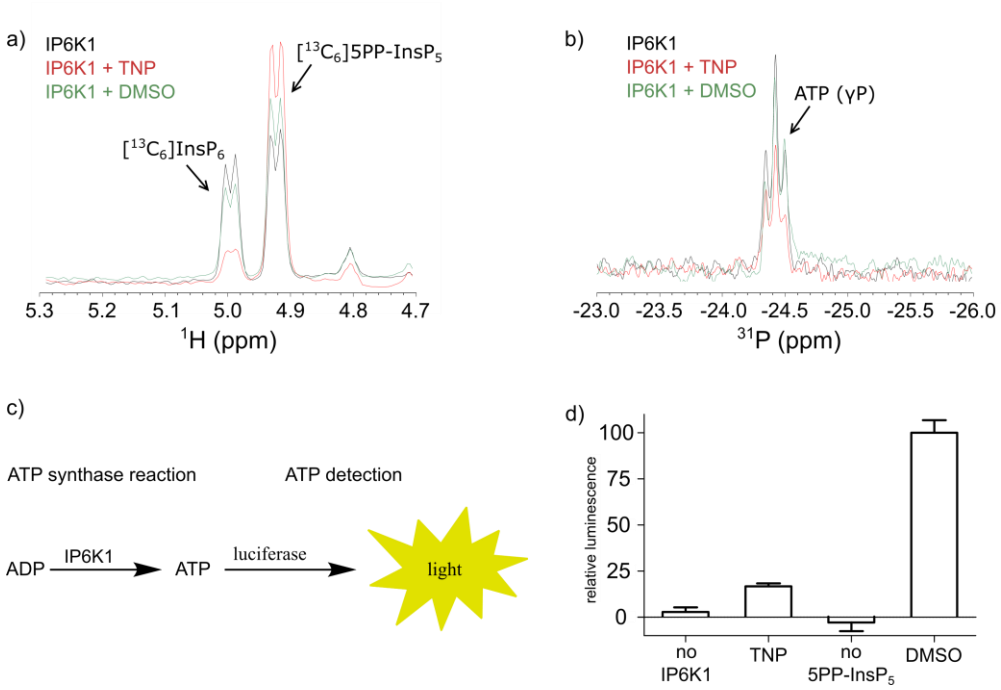


Figure 3.10: ATP synthase reactions of IP6K1 analyzed by NMR and ATP kinase Glo^{TM} assay. Formation of a) [^{13}C$_6$]InsP$_6$ and b) ATP analyzed by NMR. c) ATP detection by commercially available Kinase Glo^{TM} assay. d) Quantification of luminescence obtained from Kinase Glo^{TM} assay corrected to the average of no IP6K1 and no 5PP-InsP$_5$ controls and normalized to average of kinase reaction with DMSO added.

As for the NMR experiment, we observed robust kinase reactivity in the positive control and strongly reduced activity for the samples containing the inhibitor. The reduction in kinase activity by TNP was in line with the NMR experiments. Both negative controls showed no reaction and only background signal from the already yellowish Kinase Glo^{TM} assay solution. Notably, the error between replicates is quite small compared to the differences found in these samples.

All in all, these preliminary experiments paved the way for the optimization and setup of a high-throughput-screening assay to find specific inhibitors against IP6K1 using the ATP synthase reaction. Previously, a comparable assay would not have been a valid option because the high amounts of 5PP-InsPs required for large screen where inaccessible by organic synthesis. However, together with the newly developed enzymatic synthesis described earlier in chapter 2, this screen became feasible and is currently pursued by a postdoc and a PhD student in our lab.

Biochemical Characterization of VIH2 and ITPK1 from Arabidopsis Thaliana[4]

The *in vitro* characterization of IP6K1 showed great potential to be expanded towards other PP-InsP metabolizing enzymes. Therefore, we sought to apply our new NMR approach to investigate PP-InsP metabolism in *A. thaliana*, in collaboration with the Hothorn (University of Geneva) and Schaaf groups (University of Bonn). In plants, the biosynthesis of PP-InsPs as well as their biological function remains an understudied field compared to yeast and mammalian cells. So far, only a kinase called VIH2 was demonstrated to have PPIP5K-like behavior using a genetic model in yeast.[32,33] The IP6K family enzymes were completely unknown. Consequently, only a few phenotypes had been recognized in plants that intersect with PP-InsPs. For examples, it had been shown that jasmonate specifically increased 1,5$(PP)_2$-InsP_4 pools and that these changes were dependent on VIH2.[33] These experiments indicated that this kinase is involved in jasmonate perception and plant defenses against herbivorous insects and necrotrophic fungi. Another intriguing observation in *A. thaliana* was the role of PP-InsPs in phosphate (Pi) homeostasis.[21] PP-InsPs strongly

[4] Biochemical characterization of full length Vip1 was performed by R- Puschmann.

interact with SPX domains. These evolutionary conserved elements are either stand-alone proteins or located in proteins that support Pi uptake, storage and signaling.[21,34–42] In plants, the Pi starvation response (PHR) transcription factor binds SPX protein in a PP-InsPs dependent fashion.[43–46] (**Figure 3.11**) During low Pi availability PHR attaches to the DNA and mediates the expression of Pi starvation-induced genes. During high Pi supply, the PP-InsP levels are elevated compared to low Pi availability. Thereby, the formation of the PHR-SPX complex is promoted and PHR dissociates from its target promoter. While these studies indicated a vital role of PP-InsPs in Pi homeostasis in *A. thaliana*, it remained unclear how Pi availability was sensed and communicated within these cells. To get more insight in the underlying mechanisms a thorough biochemical characterization of PP-InsPs metabolizing enzymes is highly needed.

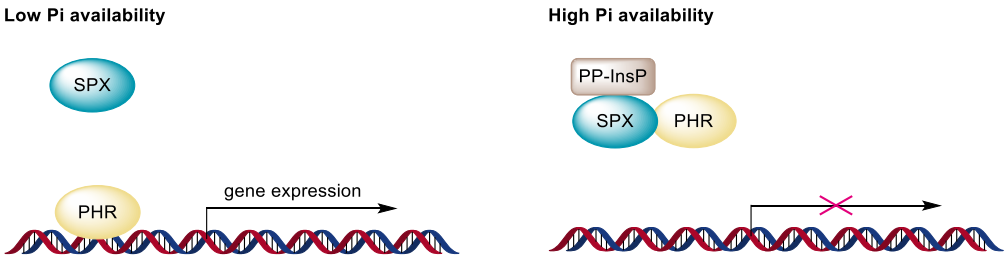


Figure 3.11: Regulation of PHR by SPX protein and PP-InsPs.

In recent attempt to understand Pi starvation response in plants, the Hothorn group found that VIH2 knockout plants responded sensitively to external Pi levels and suspected a major switch for Pi starvation response in this enzymes. To better understand how this enzyme senses levels of Pi, VIH2 was characterized biochemically. Like all PPIP5K enzymes, VIH2 contained a phosphatase domain and a kinase domain that was expected to be catalytically active on PP-InsPs. We focused

on the kinase domain (VIH2^{KD}) and identified substrates and reaction products by NMR spectroscopy. (**Figure 3.12**)

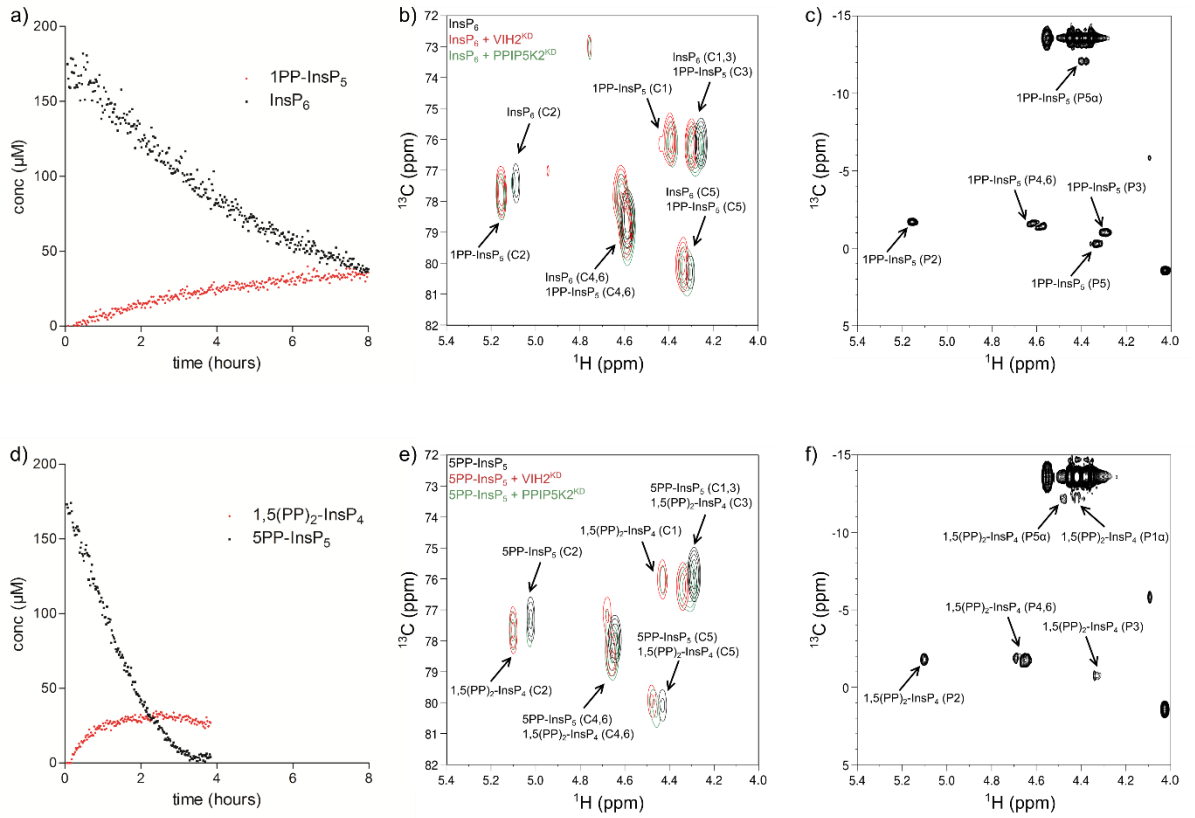


Figure 3.12: Biochemical characterization of VIH2^{KD} using [$^{13}C_6$]InsP$_6$ and [$^{13}C_6$]5PP-InsP$_5$. a) Time-resolved analysis of [$^{13}C_6$]InsP$_6$ consumption and [$^{13}C_6$]1PP-InsP$_5$ synthesis. b) ^{1}H,^{13}C-HMQC analysis of reaction product from VIH^{KD} with [$^{13}C_6$]InsP$_6$ as substrate and comparison to reaction product from PPIP5K2^{KD} with [$^{13}C_6$]InsP$_6$ as substrate. c) ^{1}H,^{31}P-HMBC analysis of reaction product from VIH^{KD} with [$^{13}C_6$]InsP$_6$ as substrate. d) Time-resolved analysis of [$^{13}C_6$]5PP-InsP$_5$ consumption and [$^{13}C_6$]1,5(PP)$_2$-InsP$_4$ synthesis. e) ^{1}H,^{13}C-HMQC analysis of reaction product from VIH^{KD} with [$^{13}C_6$]5PP-InsP$_5$ as substrate and comparison to reaction product from PPIP5K2^{KD} with [$^{13}C_6$] 5PP-InsP$_5$ as substrate. f) ^{1}H,^{31}P-HMBC analysis of reaction product from VIH^{KD} with [$^{13}C_6$]5PP-InsP$_5$ as substrate.

Continuous reaction monitoring was performed and estimated the initial rates by fitting data points within 20% conversion. A 4.9-fold higher activity for [$^{13}C_6$]5PP-InsP$_5$ (21.1 nmol min^{-1} mg^{-1}) compared to [$^{13}C_6$]InsP$_6$ (4.3 nmol min^{-1} mg^{-1}) was found. (**Figure 3.12 a**, **d**) Next, the identity of the reaction products was elucidated by

comparison to an analogous, well-characterized reaction with PPIP5K2^{KD}, as no purified standards were available at this time. Additionally, the position of the pyrophosphate was verified by ^{1}H,^{31}P-HMBC. (**Figure 3.12 b**,**c**,**e**,**f**) The ^{1}H,^{13}C-HSQC spectra of PPIP5K2 and VIH2^{KD} samples showed good overlap between the reaction products. The pyrophosphate could be localized at the 1- or 3-position. Notably, no further products were observed in theses reactions even after long incubations of more than 10 h. All in all, these experiments confirm that VIH2 belongs indeed to the PPIP5K family.

After completion of the analysis of the VIH2 kinase activity, the Hothorn group wanted to further investigated the influence of Pi levels on the activity of VIH2 full length proteins including the phosphatase domain. Unfortunately, they were unable to purify active protein and therefore all of the experiments were performed using the yeast homolog Vip1 (full length) as a mimic. In brief, we could show that high Pi availability led to an increase of [$^{13}C_6$]1PP-InsP$_5$ and [$^{13}C_6$]1,5(PP)$_2$-InsP$_4$ and a significant decrease of [$^{13}C_6$]5PP-InsP$_5$ compared to samples without Pi. Especially 1,5(PP)$_2$-InsP$_4$ is expected to bind strongly to SPX domain containing proteins leading to a downregulation of the Pi starvation response. These results corroborated the hypothesis that VIH2 activity was regulated in response to Pi availability.

At this point, all phenotypes associated with PP-InsPs in plants were explained by the regulation of VIH2. On the other hand, phenotypes associated with IP6K family enzymes were completely unknown because no IP6K-like activity had been discovered. In an effort to identify this missing kinase, the Schaaf group examined the sequence homology of different InsP kinases and found an analog of human ITPK1, an InsP$_3$ kinase, that could effectively convert InsP$_6$ and 1PP-InsP$_5$ into 5PP-InsP$_5$ and 1,5(PP)$_2$-InsP$_4$, respectively. Using our approach we enabled the determination of the

V_{max} (106,1 ± 3.7 nmol · min^{-1} · mg^{-1}) and the K_M (965.5 ± 107.4 µM) for ATP for this enzyme called now *At*ITPK1. (**Figure 3.13**) At this time, no [$^{13}C_6$]1PP-InsP_5 was available and only the reaction with [$^{13}C_6$]InsP_6 as substrate was characterized. As expected for enzymes of the IP6K family, the K_M was remarkably high indicative of a strong dependence of the enzyme activity on cellular ATP levels. Unlike IP6K1, we found no substrate inhibition at high ATP concentrations for *At*ITPK1.

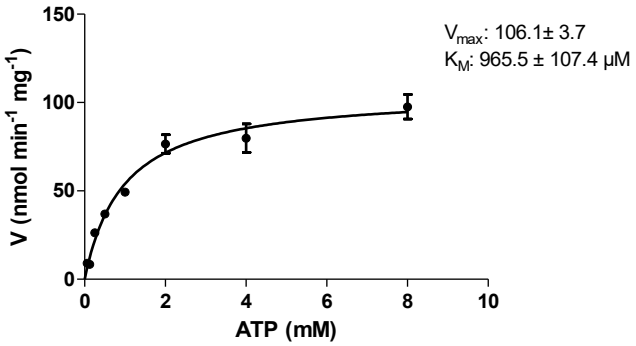


Figure 3.13: Determination of K_M for ATP and V_{max} for InsP_6. Initial rates (v) plotted against multiple ATP concentrations in triplicate.

All In all, the robust and straightforward analysis of PP-InsP metabolizing enzymes by NMR spectroscopy allowed us to produce qualitative and quantitative data with low experimental error and high confidence. Especially, the discrimination between the different PP-InsP_5 isomers in complex samples was useful to aid the identification of the Pi starvation response in plants and to confirm the identity of VIH2 as PPIP5K class enzyme.

Ex Vivo Analysis of Metabolically Labeled Eukaryotic Cells

Metabolic Labeling of Mammalian Cells[5]

In the previous sections we showed that ^{13}C-labeled PP-InsPs are useful tools to study and monitor kinase activity of PPIP5Ks and IP6Ks *in vitro* by NMR spectroscopy without the need for isolation of PP-InsPs. Encouraged by these results, we decided to apply our NMR analysis to much more complex samples. As explained in chapter 1, metabolic labeling of yeast and mammalian cells by ^{3}H-inositol is one of the most widely used methods to assay cellular InsP and PP-InsP ratios until today. Considering our access to [$^{13}C_6$]*myo*-inositol, we envisioned an analogous method that analyzes full cell extracts by NMR spectroscopy and tested the metabolic labeling of the human colon cancer cell line HCT116. (**Figure 3.14 a**) HCT116 cells were grown in the presence of 100 µM [$^{13}C_6$]*myo*-inositol and subsequently lysed using 1 M $HClO_4$. The extract was neutralized, lyophilized, dissolved in D_2O, and a ^{1}H,^{13}C-HMQC spectrum was recorded. (**Figure 3.14 b**)

[5] Metabolic labeling of HCT116 cells was performed by R. Puschmann and R. K. Harmel. Spike in experiments, TiO_2 enrichments, and preliminary experiment for the identification of unknown NMR signals were performed by R. Puschmann. Calibration curves with Me_4PBr were produced by R. Puschmann.

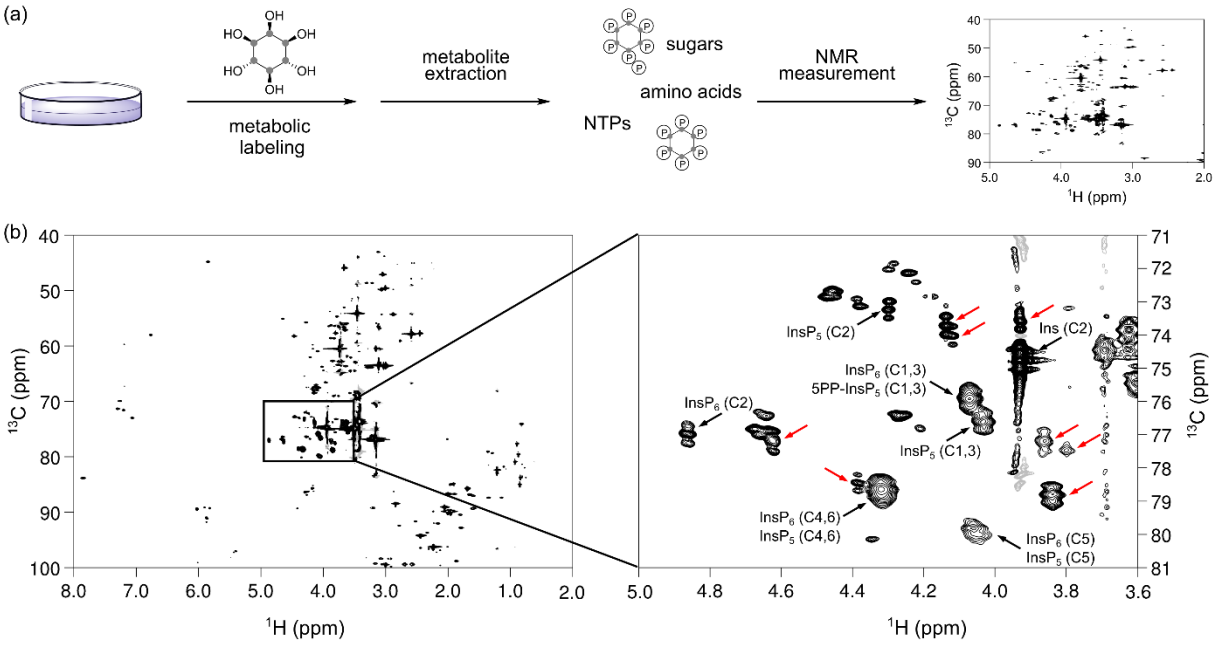


Figure 3.14: Metabolic labeling of mammalian cell line HCT116, followed by NMR analysis. (a) General workflow for the preparation of whole cell extracts for NMR spectroscopy. (b) Left: $^{1}H,^{13}C$-HMQC spectrum of an HCT116 extract. The peaks from 80 to 100 ppm in the F1 and between 0 and 3 ppm are folded into an empty region of the spectrum. Right: the inositol phosphate region of the spectrum is depicted in more detail. The identified InsPs are annotated while peaks that exhibit the expected splitting pattern in the carbon dimension but could not be attributed to either *myo*-inositol (Ins), $InsP_5$, $InsP_6$, nor 5PP-$InsP_5$ are highlighted by red arrows.

The spectrum displayed several strong peaks in the characteristic chemical shift region for InsPs and PP-InsPs (^{1}H: 5.0–3.6 ppm; ^{13}C: 81–71 ppm). At high spectral resolution a characteristic triplet pattern emerged due to the coupling of neighboring ^{13}C nuclei in the *myo*-inositol ring. Comparison to an extract prepared with [$^{12}C_6$]*myo*-inositol confirmed that these peaks were a result of labeling with [$^{13}C_6$]*myo*-inositol (**Figure 3.15 a, b**). To verify that the labeling success was not limited to HCT116 cells, we next treated human embryonic kidney cells (HEK293T) with [$^{13}C_6$]*myo*-inositol and observed robust labeling again. (**Figure 3.15 c**)

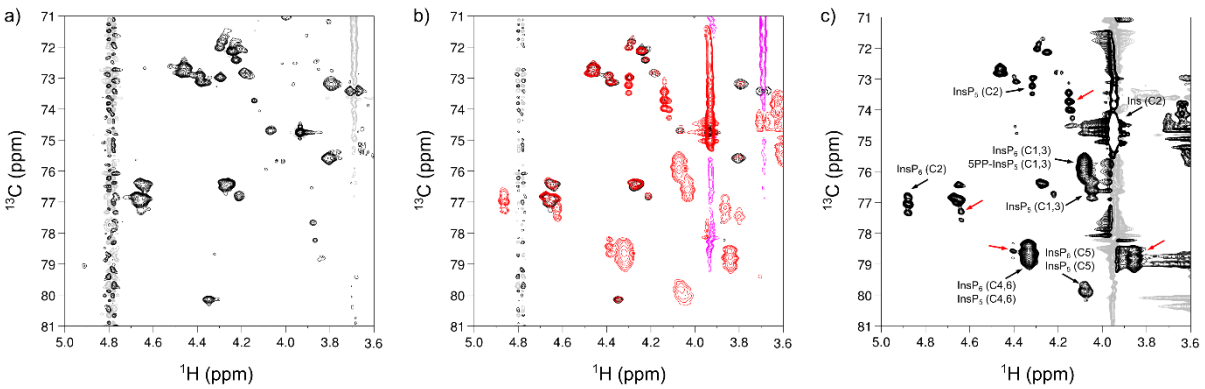


Figure 3.15: Metabolic labeling of HCT116 with [$^{12}C_6$]*myo*-inositol and HEK293T with [$^{13}C_6$]*myo*-inositol. a) Inositol phosphate region of the unlabeled HCT116 wt extract. b) Inositol phosphate regions of overlay between [$^{13}C_6$]*myo*-inositol-labeled (red) and unlabeled (black) HCT116 wt extracts. c) HEK293T cell extract. HEK293T cells were extracted with 1 M $HClO_4$ and the soluble fraction was measured by NMR-spectroscopy. The relevant InsP peaks are labeled and additional peaks displaying the triplet pattern are highlighted by red arrows.

To annotate characteristic resonances for endogenous $InsP_5$, $InsP_6$ and 5PP-$InsP_5$, we performed spike-in experiments with synthetic standards on the labeled extracts (**Figure 3.16 a–c**) and confirmed the presence of the above mentioned InsPs. We further added a mixture containing synthetic $InsP_5$, $InsP_6$ and 5PP-$InsP_5$ to an unlabeled extract from HCT116 and detected all signals at their expected chemical shifts (**Figure 3.16 d**).

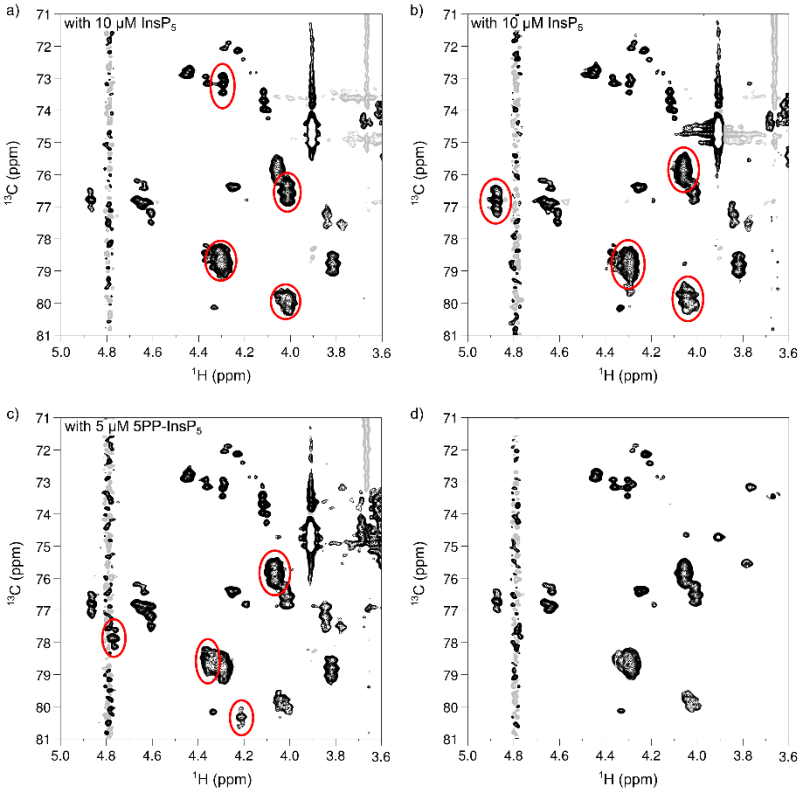


Figure 3.16: Spike in experiments to confirm the identity of $InsP_5$, $InsP_6$, and 5PP-$InsP_5$ in HCT116 extracts. a) Extract + $InsP_5$. (10 µM final concentration). b) Extract + $InsP_6$. (10 µM final concentration). c) Extract + 5PP-$InsP_5$ (5 µM final concentration). d) spike in of PP-InsPs into unlabeled cell extract.

For absolute quantification of $InsP_5$, $InsP_6$ and 5PP-$InsP_5$, we prepared [$^{13}C_6$]*myo*-inositol labeled HCT116 extracts containing a known concentration of an internal standard (tetramethyl phosphonium bromide, PMe_4Br). Resonances corresponding to the C2-position were integrated and referenced to a set of calibration curves (**Figure 3.17 a**). Assuming full replacement of the unlabeled InsP pools, we calculated respective concentrations of 26.7 ± 2.3 µM $InsP_5$, 29.4 ± 7.6 µM $InsP_6$, and 1.9 ± 0.5 µM PP-$InsP_5$, based on packed cell volume (**Figure 3.17 b**; **Table 3.1**). The measured concentrations are consistent with previous publications with respect to $InsP_5$ and $InsP_6$, however, the amount of 5PP-$InsP_5$ determined by NMR is at least 2-fold higher.[16,47] These results underline the mild sample preparation and the

quantitative nature of NMR spectroscopy and highlight its potential for future measurements of labeled cells or tissues.

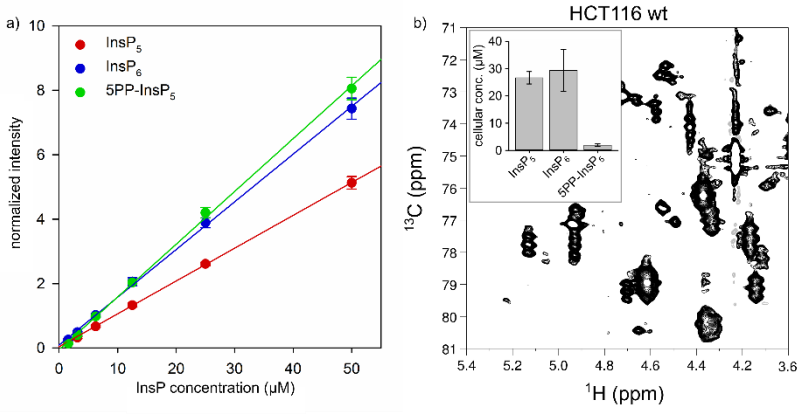


Figure 3.17: Quantification of PP-InsP pools in HCT116 wt. a) Calibration curves for InsP C2 peak intensities normalized to the reference standard PMe_4Br. Changes in 5PP-$InsP_5$ levels can be observed by NMR spectroscopy. b) $^1H,^{13}C$-HMQC spectrum of a HCT116 wild type cell extract for quantitative analysis.

Table 3.1: Absolute quantification of $InsP_5$, $InsP_6$, and 5PP-$InsP_5$ in different cell lines.

Cell line	Condition	Cellular Concentration (µM)[a]		
		$InsP_5$	**$InsP_6$**	**5PP-$InsP_5$**
HCT116 wt	no treatment	26.7 ± 2.3	29.4 ± 7.6	1.9 ± 0.5
	10 mM NaF[b]	18.6 ± 0.7	11.0 ± 0.9	14.7 ± 2.0
PPIP5K$^{-/-}$	no treatment	19.4 ± 6.1	33.7 ± 7.8	5.1 ± 2.3
	10 mM NaF[b]	11.2 ± 1.0	9.3 ± 0.7	23.8 ± 3.1

(a) The C2 peak was used for quantification. The concentration in the NMR sample was calculated based on PMe_4Br as an internal standard and the cellular concentration was calculated based on the packed cell volume.
(b) Cells were treated with 10 mM NaF 1 hour before extraction. Packed cell volume was based on untreated sample.

The ability to quantify mixtures of InsP messengers prompted us to analyze the concentrations of $InsP_5$, $InsP_6$ and 5PP-$InsP_5$ in HCT116 cells with perturbed PP-InsP metabolism. When HCT116 cells were treated for one hour with 10 mM sodium fluoride (NaF), an agent previously shown to elevate 5PP-$InsP_5$ levels,[48] the content of 5PP-$InsP_5$ increased dramatically to an absolute concentration of 14.7 ± 2.0 µM (**Figure 3.18 a**; **Table 3.1**). Elevated concentrations of 5PP-$InsP_5$ were also reported in cells

lacking PPIP5K1 and PPIP5K2.[49] We therefore grew *PPIP5K*$^{-/-}$ cells and indeed, a general increase of 5PP-InsP$_5$ was detected compared to HCT116 wt cells (**Figure 3.18 b**, **Table 3.1**). Addition of NaF to this knock-out cell line resulted in extracts which contained more 5PP-InsP$_5$ than InsP$_6$ or InsP$_5$ (**Figure 3.18 c**, **Table 3.1**). These trends of increasing 5PP-InsP$_5$ concentrations had been observed before and the ratio of 5PP-InsP$_5$ to InsP$_6$ in all our samples was consistent with previous publications.[49]

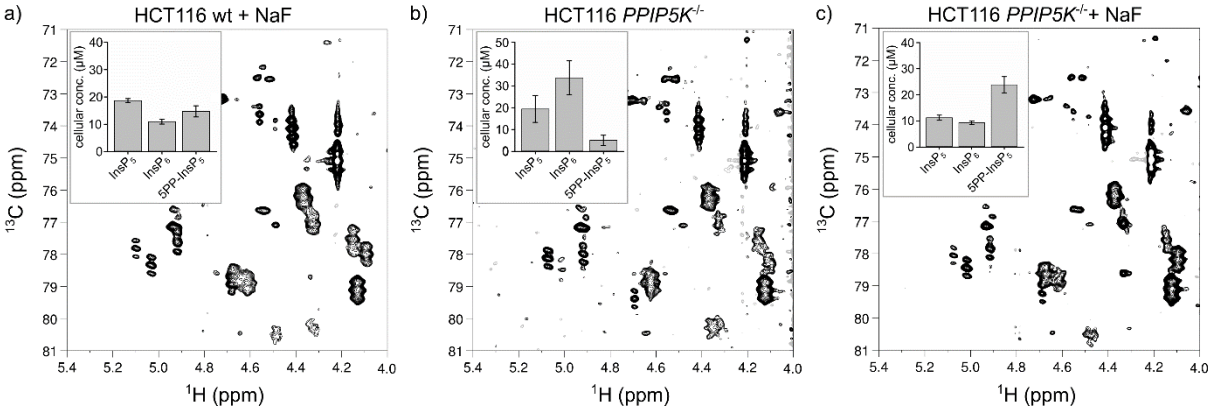


Figure 3.18: Changes in 5PP-InsP$_5$ levels can be observed by NMR spectroscopy. a) ^{1}H,^{13}C-HMQC spectrum of an extract of HCT116 wt cells treated with 10 mM NaF before extraction. b) ^{1}H,^{13}C-HMQC spectrum of HCT116 PPIP5K$^{-/-}$ cell extract. c) ^{1}H,^{13}C-HMQC spectrum of an extract of HCT116 PPIP5K$^{-/-}$ cells treated with 10 mM NaF before extraction. The inserts display triplicates of absolute cellular concentration of InsP$_5$, InsP$_6$ and 5PP-InsP$_5$ based on packed cell volume.

To our surprise, the analysis of the mammalian extracts uncovered additional NMR signals that displayed a triplet-pattern in the carbon-dimension. These signals were absent in the corresponding samples prepared with unlabeled *myo*-inositol and must therefore stem from metabolic labeling with [^{13}C$_6$]*myo*-inositol (**Figure 3.14 b**, **Figure 3.15 a**). Interestingly, several of these signals increased upon NaF treatment (**Figure 3.18 a**). To determine whether these signals correspond to other InsP or PP-InsP species, the ^{13}C-labeled, NaF treated HCT116 wt extract was incubated with TiO$_2$ beads, which enrich phosphate containing molecules.[6,50] In addition to retaining InsP$_5$,

$InsP_6$ and 5PP-$InsP_5$, a putative asymmetrical InsP species ($InsP_1$ or $InsP_2$) could be observed (**Figure 3.19**).

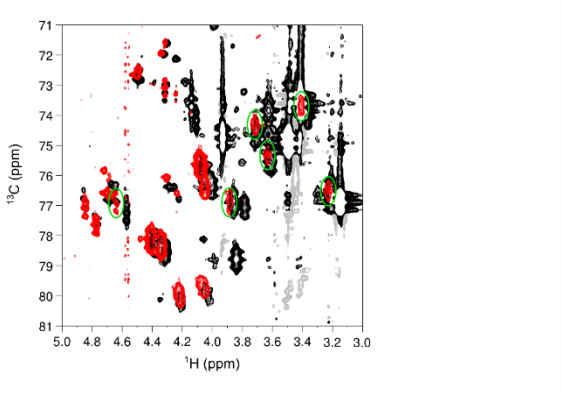


Figure 3.19: Overlay of TiO_2 enrichment of NaF treated, ^{13}C-labeled HCT116 wt cell extract (red) and the original NaF treated, ^{13}C-labeled HCT116 wt cell extract (black). One set of 6 peaks (green circles) was retained that corresponds to a putative unsymmetrical $InsP_1$ or $InsP_2$ species. All other labeled unassigned peaks were removed.

Nevertheless, most unassigned labeled resonances were not bound by the TiO_2 beads. To test for conversion of *myo*-inositol to related metabolites, spectra of D-*chiro*-inositol, s*cyllo*-inositol and glucuronic acid were recorded (**Figure 3.20 a–c**), none of which were superimposable with our unassigned labeled species.[51,52] Lastly, we considered that the additional peaks corresponded to PtdInsPs. However, after a phospholipid extraction, the aqueous layer still contained the unidentified resonances, strongly implying a non-lipid, hydrophilic character of these molecules (**Figure 3.20 a–d**). These observations suggest that the supplemented [$^{13}C_6$]*myo*-inositol has been diverted into an unknown metabolic pathway. Detection of these unanticipated species was only possible because the NMR-analysis was conducted on the crude extract and did not require a separation step. Elucidating the structure and function of these molecules connected to inositol metabolism will be of interest in the future.

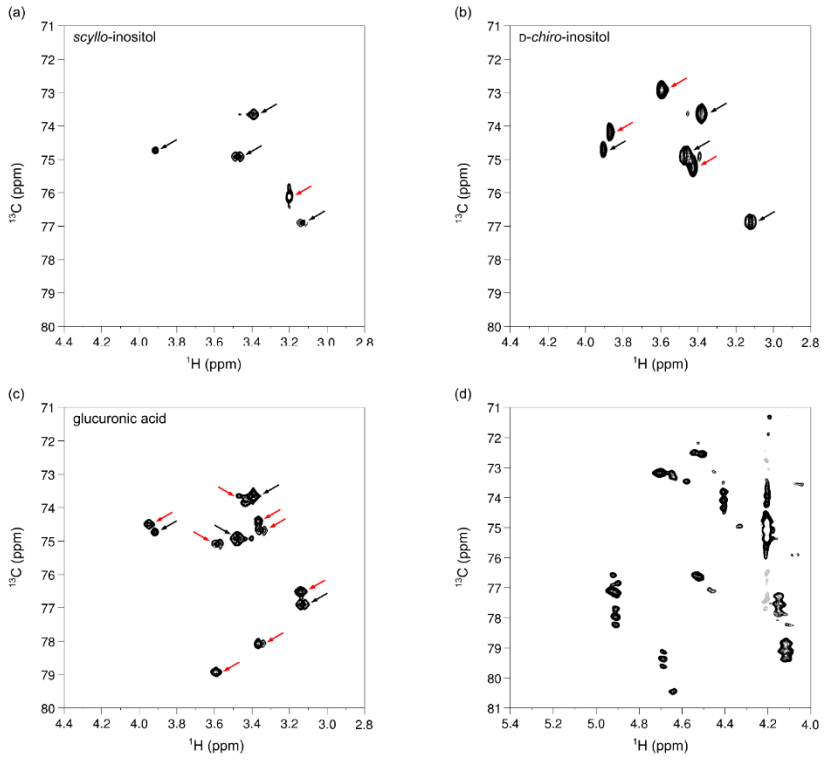


Figure 3.20: Preliminary analysis of unannotated NMR signals. HMQC spectrum of 1 mM a) scyllo-inositol, b) D-chiro-inositol, and c) glucuronic acid (highlighted with red arrows) in $KClO_4$ saturated D_2O. 1 mM *myo*-inositol was included as an internal reference (black arrows). d) Lipid extraction of HCT116 wt retained all unassigned peaks except for the putative unsymmetrical $InsP_1$ or $InsP_2$ species.

Metabolic Labeling of D. Discoideum

After we established the metabolic labeling with [$^{13}C_6$]*myo*-inositol in mammalian cells, we started to expand our method towards the model organism *D. discoideum,* together with A. Saiardi and coworkers (UCL). This social amoeba marks an evolutionary transition between unicellular and multicellular organisms and has been applied as a model to study the transition from uniform to differentiated cells.[53,54] Interestingly, *D. discoideum* was the first organism in which PP-InsPs were identified, and were subsequently shown to play a role in chemotaxis.[2,55] Opposed to mammalian cells and yeast, the main PP-InsP species were characterized as 6PP-$InsP_5$ and 5,6$(PP)_2$-$InsP_4$ and were present at a much higher concentration of 60 µM and 180 µM

respectively.[7,8,56] Retrospectively, the presence of another set of regioisomers was confounding because the kinases that form the pyrophosphate moieties show a high sequence homology between *D. discoideum* and *H. sapiens*. Therefore, the Saiardi group wanted to revisit these findings and hypothesized that either the isomers were incorrectly assigned or that the enzymology differs to other eukaryotic cells.

To this end, wild type *D. discoideum* and genetically modified stains (ppip5k null and ip6k/ppip5k null amoeba) were labeled with [$^{13}C_6$]*myo*-inositol and unpurified acidic extracts from wild type samples were measured. The $^{1}H,^{13}C$-HSQC spectrum showed four intense signals between 3.4 and 4.2 ppm and four signals between 4.2 and 5.2 ppm, corresponding to *myo*-inositol and $InsP_6$, respectively.(**Figure 3.21 a**)

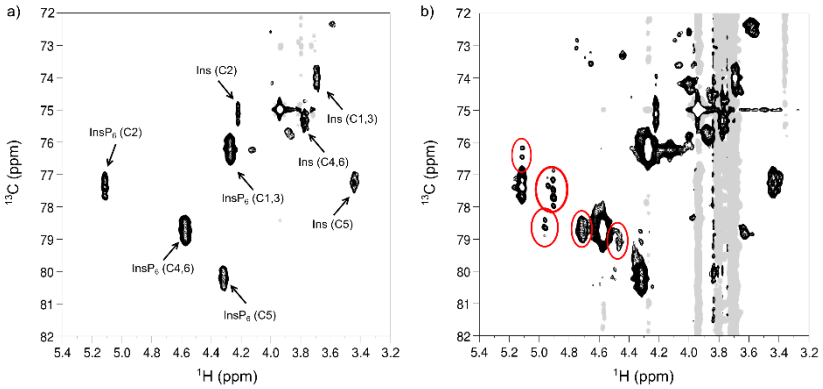


Figure 3.21: $^{1}H,^{13}C$-HSQC spectra of extracts from metabolically labeled *D. discoideum*. a) High abundant species: $InsP_6$ and *myo*-inositol. b) Lower abundant ^{13}C-labeled compounds circled in red.

In contrast to mammalian extracts, $InsP_6$ was present at a higher concentration than *myo*-inositol in these samples and there was no indication of an $InsP_5$ species. Upon increasing the counter levels, lower abundant ^{13}C-labeled species became visible, which we suspected to be PP-InsPs. (**Figure 3.21 b**) Unfortunately, the potentially different isomeric PP-InsP forms and the high sample complexity, made a clear analysis difficult. To deconvolute the sample, the Saiardi group provided extracts that

were resolved by PAGE and purified by TiO_2 beads. Individual gel bands were cut, extracted and analyzed by NMR spectroscopy. (**Figure 3.22 a**) In the wild type, four different InsP- or PP-InsP species were observed (column 1−4). The $InsP_5$ species showed 6 resonances and did not match the synthetically prepared (1,3,4,5,6)-$InsP_5$ which was mostly found in mammalian cells. (**Figure 3.22 b**) The chemical shift of the resonance corresponding to the unphosphorylated position suggested that the free OH-group is most likely located at the 1- or 3-position. HSQC-COSY experiments confirmed this proposal, as strong cross peaks with the 4/6-position were displayed. (**Figure 3.22 c**)

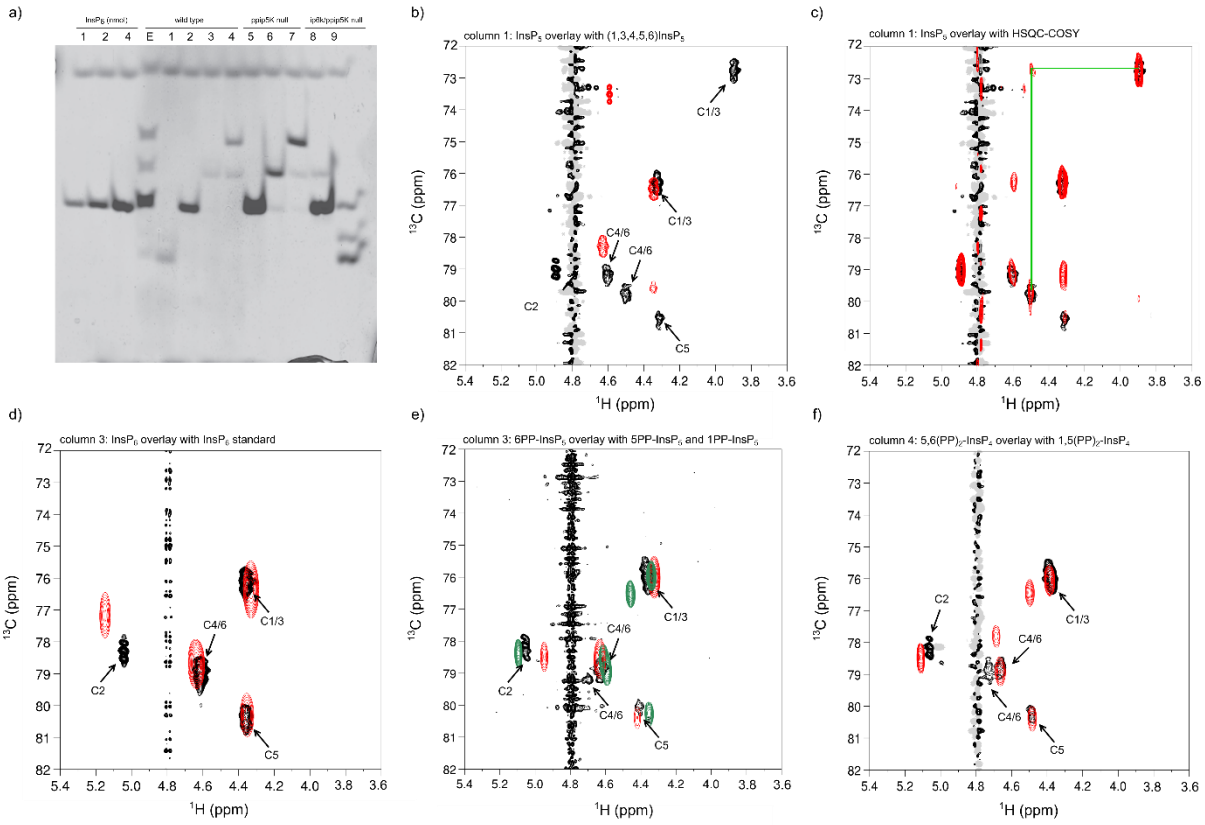


Figure 3.22: $^{1}H,^{13}C$-HSQC NMR analysis of TiO_2/gel-purified bands from wild type *D. discoideum.* a) PAGE analysis of individual bands from wild type, ppip5k null and ip6k/ppip5k null amoeba b) black: column 1, $InsP_5$; red: synthetic (1,3,4,5,6)$InsP_5$ standard. c) black: column 1, $InsP_5$; red: $^{1}H,^{13}C$-HSQC-COSY of $InsP_5$; green: cross correlation between 4/6- and putative 1/3-positions. d) black: column 2, $InsP_6$ species; red: synthetic $InsP_6$ standard e) black: column 3, 6PP-$InsP_5$; red: synthetic 5PP-$InsP_5$ standard; green: synthetic 1PP-$InsP_5$ standard. f) black: column 4, 5,6$(PP)_2$-$InsP_4$; red: synthetic 1,5$(PP)_2$-$InsP_4$ standard.

The $InsP_6$ of wild type amoeba nearly matched the human $InsP_6$. (**Figure 3.22 d**). Spectral differences especially visible at the 2-position were probably due to different ionic strength of the samples. No evidence of PP-InsPs like PP-$InsP_4$ was found. In agreement with previous reports we found 6PP-$InsP_5$ and 5,6$(PP)_2$-$InsP_4$. (**Figure 3.22 e−g**) Compared to the mammalian PP-InsPs we saw a splitting of the 4/6-positions rather than of the 1/3-positions as exemplified for 5,6$(PP)_2$-$InsP_4$. For 6PP-$InsP_5$ we observed 5 signals which did not match the spectra of 5PP-$InsP_5$ which typically have only 4 signals.

After the identification of the different InsP and PP-InsP species in the wild type, we further investigated the ppip5k null and the ip6k/ppip5k null amoeba. Surprisingly, NMR analysis of the PP-InsPs in the ppip5k knockout showed no difference as compared to the wild type but a higher concentration of PP-InsP overall, which could also be observed *via* PAGE analysis. (**Figure 3.22 a**, column 5–7) The presence of 5,6$(PP)_2$-$InsP_4$ in the ppip5k null amoeba suggested that Ppip5k is either not the main PPIP5K species in *D. discoideum,* or a complementary biosynthetic pathway exists. As explained already for VIH2 of *A. thaliana* Ppip5k is expected to have a kinase and a phosphatase domain. Apparently, the main effect observed from the absence of Ppip5k was the depletion of the phosphatase domain that could no longer dephosphorylate PP-InsPs, thereby raising PP-InsP levels in general. Compared to the knockout of Ppip5k, the change in PP-InsPs in the ip6k/ppip5k null amoeba was rather dramatic. These samples showed barely detectable amounts of PP-InsPs and mainly $InsP_6$ and several $InsP_5$ species as shown in the PAGE analysis. (**Figure 3.22 a**, column 8, 9) An NMR analysis of the $InsP_5$ species in a mixture was performed, suggested the presence of four different InsP species which could be derived from the observation of 4 peaks for the 2-position in this sample. (**Figure 3.23 a**)

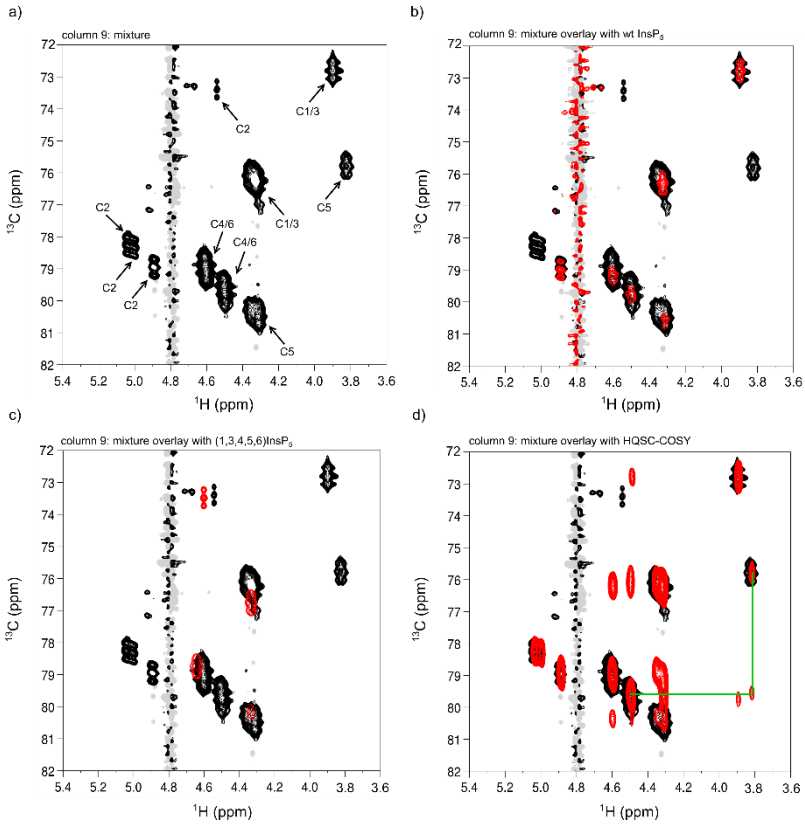


Figure 3.23: $^{1}H,^{13}C$-HSQC NMR analysis of TiO_2/gel-purified bands from ppip5k-ip6k null *D. discoideum.* a) column 9; mixture of $InsP_6$ and $InsP_5$ species. b) black: column 9, mixture of $InsP_6$; red: column 1, wild type $InsP_5$. c) black: column 9, mixture of $InsP_6$ and $InsP_5$ species; red: synthetic (1,3,4,5,6)$InsP_5$ standard. d) black: column 9, mixture of $InsP_6$ and $InsP_5$ species; red: $^{1}H,^{13}C$-HSQC-COSY of mixture of $InsP_6$ and $InsP_5$ species; green: cross correlation between 4/6- and putative 5-position.

In line with the PAGE analysis we saw $InsP_6$ in these samples and probably 3 different $InsP_5$ species. As observed previously for the wild type, we presumed one of the asymmetric species to be (2,3,4,5,6)$InsP_5$ or (1,2,4,5,6)$InsP_5$. (**Figure 3.23 b**) Furthermore, we suggest the presence of (1,3,4,5,6)$InsP_5$ by comparison to our synthetically prepared standard. (**Figure 3.23 c**) Finally, we performed HSQC-COSY analysis on this sample and identified an additional signal at 3.8 ppm that coupled to the 4/6 positions and that did not belong to any of the other species. (**Figure 3.23 d**) Since unphosphorylated 1/3-positions carbons tend to appear at a lower chemical shift in the carbon dimension, this resonance might be an unphosphorylated 5-position

hinting towards the presence of (1,2,3,4,6)$InsP_5$ in the sample. The analysis of the mixture should only be recognized as preliminary because the strong overlap between the resonances of different InsP species in the $^{1}H,^{13}C$-HSQC made a thorough and fully conclusive analysis difficult. However, the presence of these 3 different $InsP_5$ species was surprising and warrants further investigation.

The analysis of [$^{13}C_6$]*myo*-inositol labeled *D.discoideum* samples that were purified with TiO_2 and SDS PAGE provided a good method to structurally analyze PP-InsPs. It was confirmed that this slime mold can produce different $InsP_5$ species compared to mammalian cells and that the presence of the PP-InsPs is not affected by knockout of Ppip5k but only in combination with an Ip6k deletion. These observations suggest substantially altered biosynthetic pathways for PP-InsPs compared to other eukaryotic organisms.

Conclusions and Outlook

Inositol-containing molecules are ubiquitous in nature. Yet, their detection and quantification has proven to pose a formidable challenge, as these molecules typically lack a handle that would facilitate their analysis. Therefore, researchers in the field rely strongly on separation techniques and radioactive tools to answer the biological questions surrounding these metabolites.[57] We now provide a new set of methods, which emanate from access to benign ^{13}C-labeled compounds, in combination with commonly accessible NMR technology. This approach has enabled the detection and quantification of InsPs and PP-InsPs within complex mixtures including cell extracts and minimized sample purification. Specifically, the ^{13}C-labeled compounds could be employed to biochemically characterize the enzymatic activity of kinases of the IP6K and the PPIP5K family from *H. sapiens* and *A. thaliana*. In the case of IP6K1, we

investigated forward and reversed enzyme activity in response to inhibitors and established the foundation for the development of a new bioluminescent high-throughput-screening assay. In plants, we determined kinetic parameters for *At*ITPK1 and VIH2 and enabled the identification of VIH2 in combination with PP-InsP as a possible sensor of phosphate availability. In addition, the non-toxic nature of [$^{13}C_6$]*myo*-inositol allowed for metabolic labeling of mammalian cells and *D. discoideum*, followed by detection of the generated InsP and PP-InsP species. Furthermore, the structural information gained by the NMR analysis ensured reliable assignment of the identity of the detected InsP species but revealed previously unknown small molecules.

In the future, the facile applicability of this method will allow for the interrogation of new, or more complex, biological questions. On the biochemical side, for example, further characterization of kinase like human PPIP5Ks could aid the validation and identification of inhibitors as we have done for IP6Ks. The characterization of InsP- and PP-InsP phosphatases will benefit greatly from the structural information provided by the NMR measurements, and the substrate specificity of these enzymes can be unequivocally determined. Additionally, new enzymes that have been, or will be, found to metabolize InsPs can be characterized more easily, and the reaction products can be assigned with certainty.

We also envision to an important application for the labeled InsPs and PP-InsPs to gain better structural insights into their interactions with proteins. To date, the only a few canonical binding domains are known for the highly phosphorylated inositols, but many other proteins physically interact with these metabolites.[58] Instead of isotopically labeling these proteins to investigate their ligand-binding *via* NMR, applying a labeled ligand can have advantages. For example, using a ^{13}C-labeled mannose trisaccharide has recently informed the binding mode between the sugar and its protein binding

domain.[59] By analogy, the ^{13}C-labeled compounds reported here could help to obtain structural information on the conformation and the biochemical environment of the InsP ligands.

Closely related to InsP messengers are the lipid-anchored phosphatidyl inositols (PtdInsPs), which are crucial in signaling and occupy a pivotal role in health and disease.[60] Following the procedure established here for metabolic labeling, but isolating and analyzing the membrane fractions instead, may provide important information on this group of messengers. Considering that the PtdInsPs and InsPs are metabolically linked, concomitant analysis of both groups of molecules could uncover unprecedented information on the coordination of these two messenger networks. While NMR suffers from an innate low sensitivity, advances using hyperpolarization methods in recent years have greatly enhanced the signal of ^{13}C nuclei and could help to increase the sensitivity of our method in order to reliably detect inositol phosphates of low abundance, such as 1PP-InsP_5 and 1,5$(PP)_2$-InsP_4, or PtdInsPs *via* NMR.[61,62] Overall, [$^{13}C_6$]InsPs will expand the analytical tool box to interrogate inositol based messengers and their role and regulation at a biochemical and cellular level.

Methods

General Information

Please see method section of chapter 2.

Recombinant Protein Expression

Inositol Phosphate Synthase (IPS)

The IPS-gene from *A. fulgidus* cloned into a pET23a vector (Prof. Helena Santos, Universidade Nova de Lisboa)[63] was transformed into *E. coli* BL21(DE3). Two 0.8 liter overnight cultures (37 °C) in TB supplemented with Ampicillin were each diluted with 400 mL TB, prewarmed to 37 °C. After 30 min the expression was induced with 0.25 mM IPTG. After 4 hours the cells were harvested by centrifugation (3,000 g for 10 min at 4 °C) and washed with ice cold water. The pellet was stored at –80 °C until further use. The frozen cells were resuspended in lysis buffer (50 mM Tris HCl pH 8, 250 mM NaCl). For 1 g cell pellet wet weight, 5 mL lysis buffer was used. The cell suspension was supplemented with lysozyme and DNase I, and incubated for 15 min on ice. The cells were lysed with a Microfluidizer™ LM10 at 15,000 psi with five iterations. The lysate was clarified by centrifugation (30,000 g for 30 min at 4 °C). The supernatant was incubated for 30 min at 80 °C and the precipitate was removed by centrifugation (30,000 g for 30 min at 4 °C). The volume of the heat-treated supernatant was reduced to 22 mL with spin filtration (Amicon® Ultra 15 mL 10K) and the solution was dialyzed against lysis buffer overnight. The protein solution was adjusted to 33 % (v/v) glycerol, aliquoted, frozen in liquid nitrogen, and stored at –80 °C.

Inositol Hexakisphosphate Kinase 1 (IP6K1)

The procedure is adapted from.[64] In short: The codon optimized human IP6K1-gene was subcloned via LIC into a in a pET-MBP vector and transformed into *E. coli* arctic express (DE3). An 800 mL expression culture (37 °C) was inoculated at OD_{600} 0.1 in TB-Amp/Gen and grown for 4 h. The culture was switched to 15 °C for 30 min the expression was induced with 0.1 mM IPTG. The Expression occurred overnight. The cells were harvested by centrifugation (3,000 g for 10 min at 4 °C) and washed with ice cold water. The pellet was stored at –80 °C until further use. The frozen cells were resuspended in lysis buffer (20 mM Tris HCl pH 7.4, 150 mM NaCl). For 1 g wet weight 10 mL lysis buffer was used. The cell suspension was supplemented with lysozyme, DNase I and 1 tablet cOmplete™ protease inhibitor (Roche), and incubated for 15 min on ice. The cells were lysed with a Microfluidizer™ LM10 at 15,000 psi with five iterations. The lysate was clarified by centrifugation (30,000 g for 30 min at 4 °C). The supernatant was adjusted to 0.1 % (v/v) Triton X-100, filtered (VWR® vacuum filter, PES 0.45 µm), and loaded onto a Co-NTA column (GE, 1 mL, HiTrap IMAC HP) that was equilibrated with lysis buffer with a flowrate of 1 mL/min. The column was washed with wash buffer (20 mM Tris HCl pH 7.4, 500 mM NaCl, 50 mM imidazole, 0.1 % (v/v) Triton X-100) until the absorption was constant. IP6K1 was eluted with a gradient of elution buffer (20 mM Tris HCl pH 7.4, 500 mM NaCl, 500 mM imidazole, 0.1 % (v/v) Triton X-100) in wash buffer from 0–100 % over 10 CV. 1 mL fractions were collected. The fractions containing IP6K1 were concentrated and dialyzed overnight against dialysis buffer (20 mM Tris HCl pH 7.4, 500 mM NaCl, 1 mM DTT). The protein was adjusted to 20 % (v/v) glycerol and stored at –80 °C.

NMR experiments

Progress Curve Measurements with IP6KA

Progress curves were recorded for different IP6KA concentrations ranging from 0.2–1 µM. The reactions contained 20 mM HEPES NaOD pH* 7.0, 50 mM NaCl, 1 mM DTT, 10 mM ATP, 11 mM $MgCl_2$, 176 µM [$^{13}C_6$]$InsP_6Na_{12}$ (**6**), 5 mM creatine phosphate, 2 mM Me_4PBr, and 1 U/mL creatine kinase in D_2O. All stock solutions were prepared in D_2O. The reactions were run in a total volume of 650 µl and equilibrated at 37 °C. The reactions were transferred into a 5 mm NMR tube and the reaction was started by the addition of the appropriate amount of IP6KA (in D_2O buffer). The NMR tubes were inserted into the NMR instrument, locked, tuned and matched, and shimmed. A spin echo difference pulse was used to measure consecutively 75 sec spectra until the reaction was finished.

Kinetic Characterization of IP6K1

The $K_{M,ATP}$ and V_{max} were determined for IP6K1. The buffer contained 20 mM HEPES NaOD pH* 7.0, 50 mM NaCl, 1 mM DTT, 1 mg/mL BSA, 1 µM IP6K1, 5 mM creatine phosphate, and 1 U/mL creatine kinase in D_2O. The ATP concentration ranged from 8 mM to 62.5 µM (two-fold dilution series) and the $MgCl_2$ concentration was adjusted to be 5 mM plus the ATP concentration. The reaction volume was 500 µL and the reaction was started after 10 min equilibration at 37 °C by the addition of 175 µM [$^{13}C_6$]$InsP_6$. The reaction was quenched with 38 µL 0.7 M EDTA (in D_2O, pH* 8.0) after approximately 20 % conversion (5 min for IP6K1). Before the NMR measurements the pH* was adjusted to 8, if needed. The conversion of [$^{13}C_6$]$InsP_6$ at different ATP

concentration was determined by NMR spectroscopy. The data was fitted against a kinetic model for substrate inhibition: $v = \frac{V_{max}}{1+\frac{K_m}{S}+\frac{S}{K_i}}$ using SigmaPlot 12.5.

Determination of IC_{50} Values for IP6K1 Inhibitors

For the IC_{50} value determination a two-fold dilution series of the appropriate inhibitor in DMSO-d_6 was used. The reactions were run in a total volume of 500 µL and contained 20 mM HEPES NaOD pH* 7.0, 50 mM NaCl, 1 mM DTT, 2.5 mM ATP, 7.5 mM $MgCl_2$, 0.2 µM IP6K1, 1 mg/mL BSA, 5 mM creatine phosphate, and 1 U/mL creatine kinase in D_2O. The inhibitor concentration ranged from 50 µM to 195 nM (200 x stock solutions were used). The reactions were equilibrated to 37 °C for 10 min and initiated by the addition of 175 µM [$^{13}C_6$]$InsP_6$. The reactions were quenched after 3 hours by the addition of 38 µL of 0.7 M EDTA (in D_2O, pH* 8.0). Before the NMR measurements the pH* was adjusted to 8, if needed. The conversion of [$^{13}C_6$]$InsP_6$ at different inhibitor concentrations was determined by NMR spectroscopy. The data was fitted against a kinetic model for dose response inhibition: $v = \frac{bottom+(top-bottom)}{1+\frac{x}{IC_{50}}^{Hillslope}}$ using SigmaPlot 12.5.

Metabolic Labeling of Mammalian Cell Lines

HCT116 wt, HCT116 PPIP5K$^{-/-}$ (a kind gift from Stephen Shears)[49] and HEK293T cells were grown in DMEM lacking *myo*-inositol and supplemented with 10 % dialyzed fetal bovine serum, 100 µM either [$^{12}C_6$]- or [$^{13}C_6$]*myo*-inositol, and Penicillin/Streptomycin at 37 °C and 5 % CO_2. 1 liter DMEM w/o inositol was prepared from pre-mixed medium components (8.1 g, Dulbecco), $NaHCO_3$ (3.7 g), HEPES (10

mL of a 1 M stock, pH 7.4), L-glutamine (584 mg), L-serine (42 mg), D-glucose (4.5 g), and NaH_2PO_4 (125 mg). The cells were grown in 15 cm dishes (for one experiment we used five to ten plates) until they reached 80–90 % confluency. For harvesting, cells were washed with PBS and 0.9 % NaCl solution, and trypsinized. The trypsin was quenched with regular DMEM. The cells were collected and washed two more times with PBS before lysis and extraction. The packed cell volume and the cell number of the preparations were determined. If needed, the cells were incubated with 10 mM NaF for one hour before harvest.

Inositol Phosphate Extraction

All steps were performed at 4 °C. The protocol was adapted from Azevedo et al.[65] For HCT116 wt we used ten 15 cm dishes and for HCT116 wt with NaF treatment, HCT116 PPIP5K$^{-/-}$ and HCT116 PPIP5K$^{-/-}$ with NaF treatment we used three 15 cm dishes. Cell pellets were lysed with 5 mL of 1 M $HClO_4$, containing 3 mM EDTA, by vortexing. The lysate was incubated on ice for 30 min before the precipitate was removed by centrifugation. The supernatant was neutralized with 2 M KOH, containing 3 mM EDTA, and the pH was adjusted to 5.8–6. The resulting $KClO_4$ precipitate was removed by centrifugation and the supernatant was lyophilized. After lyophilization the supernatant was redissolved in 1 mL D_2O and the resulting precipitate was removed by centrifugation. The supernatant was lyophilized again and the residue was redissolved in 600 µL D_2O containing the 50–150 µM Me_4PBr standard for absolute quantification.

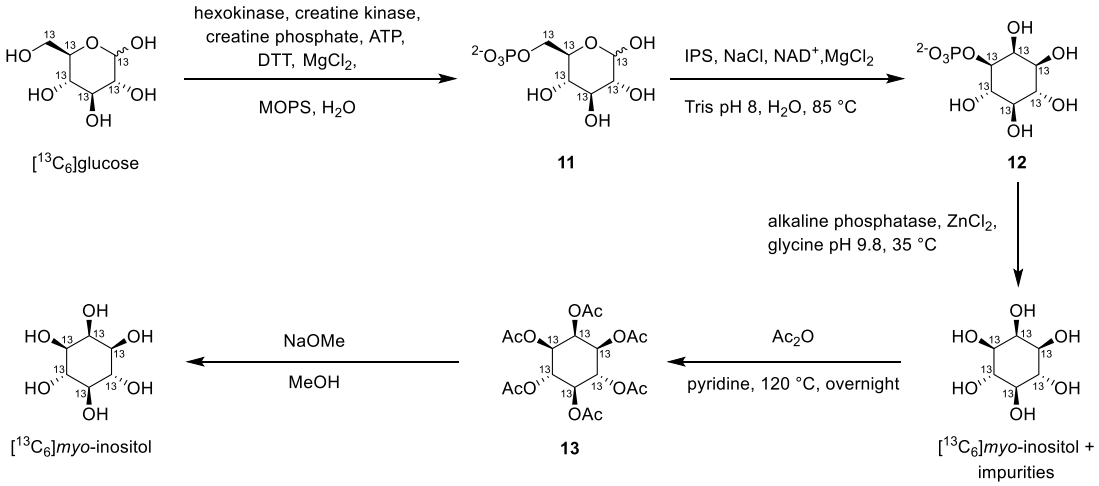
TiO_2 Enrichment

All steps were performed at 4 °C. The protocol was adapted from Wilson *et al*.[6] TiO_2 beads 4-5 mg were washed with 500 µL of water and 500 µL of 1 M $HClO_4$. Cell extract was adjusted to pH 1 with 1 M $HClO_4$ and added to the beads. The beads were rotated for 5 min and after centrifugation, the supernatant was discarded. The beads were washed twice with 500 µL of 1 M $HClO_4$ and the InsPs were eluted by incubating the beads twice with 200 µL 2.8 % NH_4OH solution. The eluate was lyophilized and the dry residue was dissolved in 600 µL of 1 mM MES (pH* 6.0) in D_2O.

Lipid Extraction

The lipid extraction was adopted from Clark et al.[66] In summary: The cells from 3 15 cm dishes were trypsinized and washed twice with PBS. The cells were resuspended in 1.7 mL Milli-Q water and 7.5 ml quench mix (48.4 mL MeOH, 24.2 mL $CHCl_3$, 2.4 mL 1 M $HClO_4$) was added. The cells were lysed by vortexing for 30 sec. 1.7 mL 1 M $HClO_4$ and 7.25 mL $CHCl_3$ were added to induce phase separation and the mixture was vortexed for 30 sec. The sample was centrifuged at 3000 g for 5 min to facilitate phase separation. The top aqueous layer was collected and washed with 7 mL $CHCl_3$. The aqueous layer was collected and neutralized with 2 M KOH containing 3 mM EDTA. The $KClO_4$ salt was removed by centrifugation and the sample was lyophilized. After lyophilization the supernatant was redissolved in 1 mL D_2O and the resulting precipitate was removed by centrifugation. The supernatant was lyophilized again and the residue was redissolved in 600 µL D_2O.

Synthesis of [$^{13}C_6$]myo-Inositol



The following stock solutions were prepared:

Hexokinase stock solution: 1000 U/mL in 50 mM citrate pH 7, 10 mM $MgCl_2$, 1 mg/mL BSA.

Creatine kinase stock solution: 350 U/mL in 200 mM MOPS pH 6.5, 20 mM $MgCl_2$, 20 mM DTT.

Inositol phosphate synthase stock solution: 4 mg/mL in 20 mM Tris HCl pH 7.4, 200 mM NaCl, 1 mM DTT.

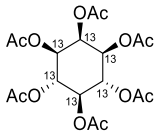
Alkaline phosphatase stock solution: 100 U/mL, 10 mM Tris pH 8.5, 5 mM $MgCl_2$, 0.2 mM $ZnCl_2$, 50 % (v/v) glycerol.

A solution containing D-[$^{13}C_6$]glucose (83 mM, 1000 mg, 5.4 mmol), MOPS (100 mM, pH 6.5), creatine phosphate (87 mM), ATP (1 mM), DTT (20 mM) and $MgCl_2$ (20 mM) in MilliQ® water (65 mL total reaction volume) was prepared and evenly split (32.5 mL each) into two 50 mL conical tubes. Hexokinase (1 U/mL) and creatine kinase (1.75 U/mL) were added and the reaction incubated at 30 °C overnight. The resulting solution was monitored by TLC (MeOH:H_2O:NH_4OH:AcOH, 50:30:15:5; stained by $KMnO_4$ or excessive heating) and upon completion diluted 5-fold with 320 mL of water. An anion-

exchange column (DOWEX® 1X8) was equilibrated with 1 M NH_4HCO_3 and washed with water. The reaction mixture was loaded onto the column followed by washing with water to remove unreacted D-[$^{13}C_6$]glucose. [$^{13}C_6$]glucose-6-phosphate (**11**) was eluted by 0.1 M $(NH_4)_2CO_3$ and lyophilized to obtain product **11** in combination with high amount of salts (4 g) as a white solid.

The product/salt mixture was redissolved in 40 mL MilliQ® water and added to a solution of Tris (50 mM, pH 8.0), NAD^+ (0.5 mM), NaCl (50 mM) and $MgCl_2$ (2 mM) in MilliQ® (196 mL total reaction volume). The pH was adjusted to 8.0, if necessary, and the mixture was evenly split (49 mL each) into four 50 mL conical tubes. Recombinantly expressed inositol-3-phosphate synthase (IPS) (500 µL of 4 mg/mL stock) was added to each tube and the reaction mixture was incubated at 85 °C monitoring conversion by NMR. After 4 h, NAD^+ (10 mg, 0.8 mM total) and IPS (500 µL, 4 mg/mL) were added to each tube and the mixture was incubated for additional 4 h at 85 °C. After full conversion of **11** into [$^{13}C_6$]*myo*-inositol-3-phosphate (**12**) was observed by ^{13}C NMR the reaction mixture was directly used for the subsequent dephosphorylation reaction.

For the dephosphorylation reaction, the 200 mL reaction mixture was split evenly across 5 conical tubes (40 mL each). For a total reaction volume of 50 mL, glycine (50 mM, pH 9.8) was added and the pH adjusted to 9.8. $ZnCl_2$ (0.25 mM) and alkaline phosphatase (0.5 U/mL) were added and the solution was filled up to 50 mL with MilliQ® water. After incubation at 35 °C overnight, all tubes were combined (250 mL total) and diluted with 300 mL MilliQ® water and the reaction mixture was applied to an ion exchange column (DOWEX® 1×8 in HCO_3^- form). The flow-through was collected and lyophilized to afford [$^{13}C_6$]*myo*-inositol with salts (7.5 g) as a brown solid which was directly used for the acetylation reaction.



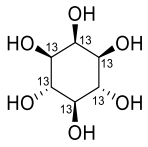
[$^{13}C_6$]Inositol Hexakisacetate (**13**)

The crude and unpurified [$^{13}C_6$]*myo*-inositol (7.5 g) was suspended in pyridine (161 mL) and treated with acetic anhydride (68.5 mL) at 120 °C overnight. The black solution was concentrated to minimize the amount of pyridine, re-dissolved in DCM (500 mL), and 1 M HCl was added. The resulting suspension was filtered and the filter washed with DCM and water. All filtrates were combined and organic and aqueous phase were separated. (Note: Phase separation is hard to see because of the black solution so use a flash light to find the separation layer.) The aqueous layer was extracted twice with DCM and the combined organic layers were washed with 1 M $NaHCO_3$. The aqueous layer was washed twice with DCM and the combined organic layers were washed with brine, dried with Na_2SO_4 and concentrated under reduced pressure. The residue was immobilized on Telos® and purified by column chromatography on silica gel changing the eluent step wise (hexane:EtOAc, 10:1 → 5:1 → 1:1 → 1:2) assisted by analysis of the fractions by LCMS to afford compound **13** (1200 mg, 2.7 mmol) as a white solid in 50% overall yield from D-[$^{13}C_6$]glucose.

^{1}H NMR (600 MHz, $CDCl_3$) δ 5.79 – 5.54 (m, 1.5H), 5.53 – 5.32 (m, 1.5H), 5.32 – 5.12 (m, 1.5H), 5.11 – 4.82 (m, 1.5H), 2.20 (s, 3H), 2.01 (s, 3H), 2.01 (s, 6H), 2.00 (s, 6H).

^{13}C NMR (151 MHz, $CDCl_3$) δ 169.94, 169.81, 169.56, 71.68 – 70.65 (m), 70.11 – 69.15 (m), 69.11 – 67.90 (m), 20.89, 20.69, 20.60.

HRMS (ESI/Orbitrap) m/z: $[M + K]^+$ calcd. for $C_{12}{}^{13}C_6H_{24}KO_{12}$ 477.1101; Found 477.1091.



[$^{13}C_6$]*myo*-Inositol

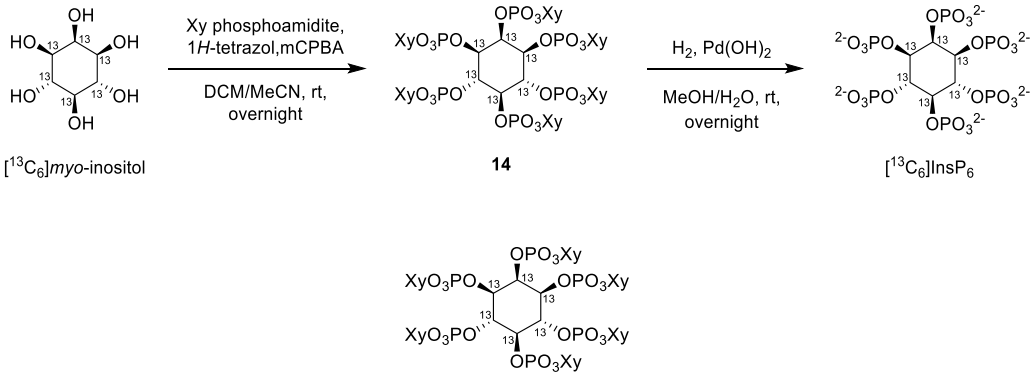
To a solution of **13** (1200 mg, 2.7 mmol) in methanol (249 mL), 5.4 M NaOMe in MeOH (3.35 mL, 18.1 mmol) was added and the reaction was left to stir for 2 hours. The resulting suspension was neutralized by the addition of DOWEX 50W x 8 (H^+ form) followed by filtration. The residue was washed with methanol (100 mL) and water (100 mL) to dissolve all precipitated [$^{13}C_6$]*myo*-inositol and the filtrate was evaporated. The resulting solids were redissolved in H_2O and precipitated by the addition of MeCN. Subsequently, the solids were redissolved in H_2O and lyophilized to afford [$^{13}C_6$]*myo*-inositol (505 mg, 2.7 mmol) as white solid in 99% yield.

^{1}H NMR (600 MHz, D_2O, pD 7.0) δ 4.29 – 4.14 (m, 0.5H), 4.03 – 3.90 (m, 0.5H), 3.82 – 3.60 (m, 2H), 3.59 – 3.34 (m, 2.5H), 3.25 – 3.12 (m, 0.5H).

^{13}C NMR (151 MHz, D_2O, pD 7.0) δ 74.96 – 73.80 (m), 72.86 – 71.71 (m), 71.60 – 70.50 (m).

HRMS (ESI/Orbitrap) m/z: $[M - H]^-$ Calcd for $^{13}C_6H_{11}O_6$ 185.0762; Found 185.0814.

Synthesis of [$^{13}C_6$]InsP$_6$



Xylyl-protected [$^{13}C_6$]InsP$_6$ (**14**)

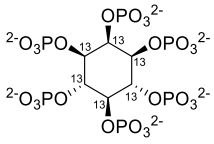
Xylyl-protected [$^{13}C_6$]InsP$_6$ (**14**) was synthesized according to a modified procedure of Podeschwa and coworkers.[17] Under nitrogen atmosphere a suspension of [$^{13}C_6$]*myo*-inositol (100 mg, 0.54 mmol) and Xy phosphoamidite (900 mg, 3.76 mmol) in dry DCM (38.65 mL) was prepared and sonicated for 1 min. 1*H*-tetrazole in anhydrous MeCN (14.33 mL, 6.45 mmol, 0.45 M) was added and the solution was stirred at rt overnight. For workup the solution was cooled to –40 °C and an anhydrous solution of *m*CPBA in DCM (30 mL dried with Na_2SO_4) was added. The solution was allowed to warm to rt, and the stirring was continued for another hour. The reaction mixture was diluted with DCM (300 mL) and washed consecutively with aqueous sodium bisulfite (20%, 2 × 50 mL), saturated $NaHCO_3$ (3 × 100 mL), and then with brine. After evaporation of the solvents, the crude was immobilized on Telos® and purified by CombiFlash® chromatography on silica gel (12 g column, gradient: 0% → 2% → 4% → 10% MeOH in DCM) and afforded the product (221 mg, 0.173 mmol) as a white solid in 32% yield.

^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.60 – 7.49 (m, 18H), 7.44 (d, *J* = 7.7 Hz, 6H), 5.93 (dd, *J* = 13.8, 9.5 Hz, 2.5H), 5.84 (dd, *J* = 13.7, 9.1 Hz, 2H), 5.79 – 5.70 (m, 6H), 5.66 (s, 0.5H), 5.55 (dd, *J* = 13.9, 12.4 Hz, 3H), 5.50 – 5.16 (m, 15H), 5.12 (s, 1H).

^{13}C NMR (151 MHz, Chloroform-*d*) δ 135.94, 135.91, 135.73, 135.69, 135.52, 134.62, 129.72, 129.60, 129.47, 129.45, 129.42, 129.40, 129.39, 129.36, 129.19, 129.13, 78.48 – 76.20 (m), 73.92 (dd, *J* = 43.9, 34.0 Hz), 69.84, 69.79, 69.68, 69.62, 69.57, 69.51, 69.45.

^{31}P NMR (122 MHz, Chloroform-*d*) δ -2.80, -3.57, -4.48, -4.71.

HRMS (ESI/Orbitrap) m/z: $[M + H]^+$ Calcd for $C_{48}{}^{13}C_6H_{54}O_{24}P_6$ 1279.1705; Found 1279.1676.



$[^{13}C_6]InsP_6$

$[^{13}C_6]InsP_6$ was synthesized according to a modified procedure of Godage and coworkers.[19] Compound **14** (148 mg 0.12 mmol) was dissolved in methanol (10.53 mL) and water (2.63 mL) and 20% $Pd(OH)_2/C$ (50% wetted with water) (68 mg, 0.49 mmol) was added. The resulting suspension was stirred at room temperature overnight under hydrogen atmosphere. The reaction mixture was passed through a PTFE syringe filter and the filtrate was evaporated under reduced pressure. The free acid was treated with NaOH (1 M) to afford the dodecasodium salt of $[^{13}C_6]InsP_6$ as a white solid (136 mg, 0.10 mmol, 66 w/w%) in 83% yield.

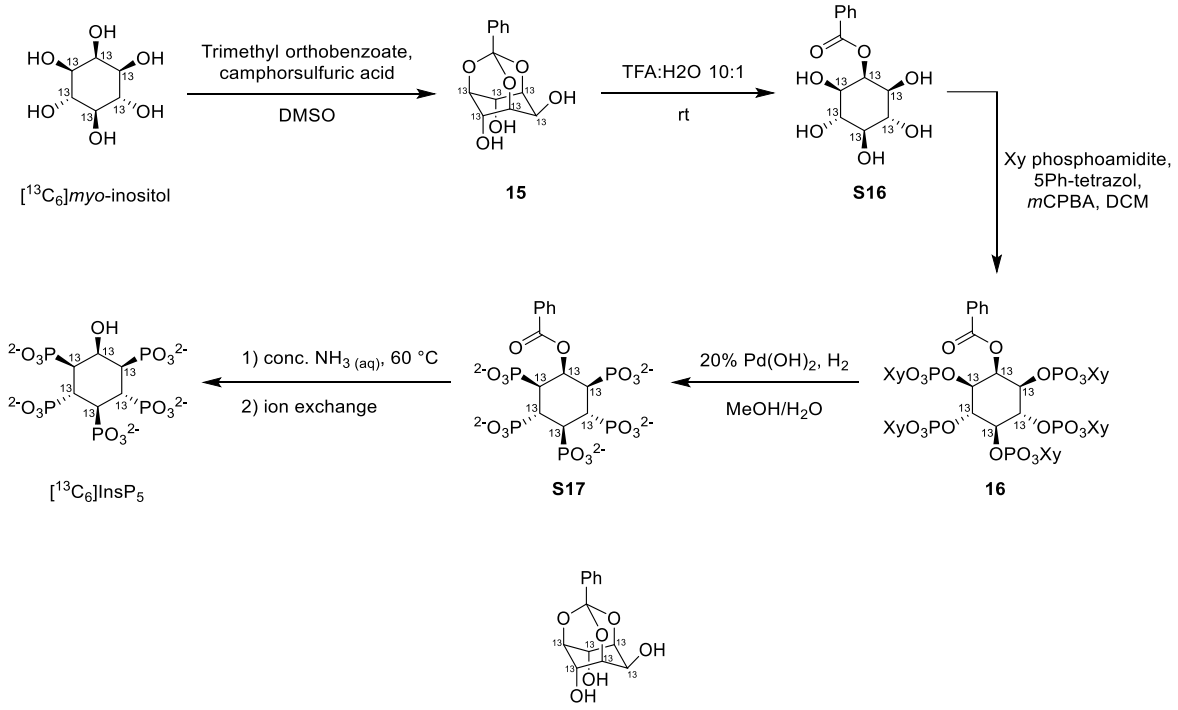
^{1}H NMR (600 MHz, Deuterium Oxide, pD 7.0) δ 5.01 – 4.90 (m, 0.5H), 4.71 – 4.63 (m, 0.5H), 4.48 – 4.34 (m, 1H), 4.22 – 4.04 (m, 2.5H), 3.95 – 3.78 (m, 1.5H).

^{13}C NMR (151 MHz, Deuterium Oxide, pD 7.0) δ 79.99 (d, *J* = 35.7 Hz), 78.52 (t, *J* = 40.0 Hz), 77.60 – 74.58 (m).

^{31}P NMR (122 MHz, Deuterium Oxide, pD 7.0) δ 1.99, 1.04, 0.77.

HRMS no ion detected.

Synthesis of [$^{13}C_6$]InsP$_5$



[$^{13}C_6$]*myo*-Inositol-(1,3,5)-orthobenzoate ester (**15**)

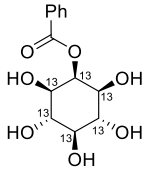
[$^{13}C_6$]inositol-(1,3,5)-orthobenzoate (**15**) ester was synthesized according to a modified procedure of Godage and coworkers.[19] A suspension of [$^{13}C_6$]*myo*-inositol (1.045 g, 5.61 mmol) and CSA (26.08 mg, 0.11 mmol) in DMSO (3.74 mL) was heated to 80 °C under vacuum (Rotovap, 30–40 mbar) to remove residual water. Upon the addition of trimethyl orthobenzoate (1.125 g, 6.17 mmol, 1.061 mL), the mixture was left at 80 °C under vacuum (Rotovap, 30–40 mbar) until the suspension became clear. The resulting solution was quenched with TEA (62.5 mg, 0.62 mmol, 86 µL). (Note: Quenching of acid is highly important to assure the stability of the formed orthoester during the workup.) The product was slowly precipitated through the addition of water

(16 mL) at 4 °C and the solids were filtered and washed with ice cold water 3 times. The mother lye was combined with the washings and applied to reversed phase chromatography changing the eluent step wise (5 % → 25 % → 50 % MeCN in H_2O) to separate residual product from DMSO. The product containing fractions were lyophilized and the precipitates were dissolved in 50 % MeCN in H_2O and lyophilized to afford the product **15** (1.300 g, 4.78 mmol) as white solid in 85% yield.

^{1}H NMR (600 MHz, Methanol-d_4) δ 7.71 – 7.64 (m, 2H), 7.42 – 7.30 (m, 3H), 4.70 (s, 1H), 4.45 (s, 2.5H), 4.37 (s, 0.5H), 4.18 (s, 1.5H), 4.13 (s, 0.5H).

^{13}C NMR (151 MHz, Methanol-d_4) δ 130.12, 128.58, 126.74, 77.50 (t, *J* = 37.2 Hz), 71.63 (t, *J* = 37.7 Hz), 68.98 (t, *J* = 37.7 Hz), 60.20 (t, *J* = 36.8 Hz).

HRMS no ions detected.



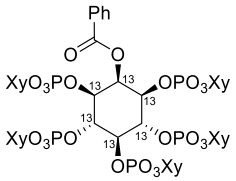
[$^{13}C_6$]2-Benzoyl inositol (**S16**)

[$^{13}C_6$]2-Benzoyl inositol (**S16**) was synthesized according to a procedure of Godage and coworkers.[19] A mixture of TFA (1.8 mL) and water (180 µL) was added to **15** (300 mg, 1.1 mmol) and the solution was stirred for 1 h and the conversion was followed by TLC (100 % EtOAc; starting material R_f 0.6; product R_f 0.0). The reaction mixture was then co-evaporated with water *in vacuo*, redissolved in water and lyophilized to obtain the product **S16** (314 mg, 1.08 mmol) as a white solid in quantitative yield.

^{1}H NMR (600 MHz, Deuterium Oxide, pD 7.0) δ 8.10 (d, *J* = 7.8 Hz, 2H), 7.74 (d, *J* = 7.5 Hz, 1H), 7.59 (t, *J* = 7.7 Hz, 2H), 5.86 (s, 0.5H), 5.60 (s, 0.5H), 4.02 – 3.89 (m, 2H), 3.78 – 3.66 (m, 2H), 3.58 – 3.51 (m, 0.5H), 3.36 – 3.28 (m, 0.5H).

^{13}C NMR (151 MHz, Deuterium Oxide, pD 7.0) δ 170.73, 136.84, 132.46, 131.52, 77.85 (t, *J* = 38.6 Hz), 77.09 (t, *J* = 38.6 Hz), 75.62 (t, *J* = 38.8 Hz), 72.63 (td, *J* = 38.8, 6.6 Hz).

HRMS (ESI/Orbitrap) m/z: $[M + H]^+$ calcd. for $C_7{}^{13}C_6H_{17}O_7$ 291.1170; Found 291.1166.



Xylyl-protected [$^{13}C_6$]benzoyl *myo*-inositol (**16**)

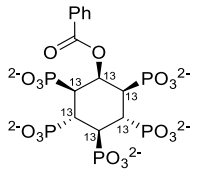
Xylyl-protected [$^{13}C_6$]benzoyl *myo*-inositol (**16**) was synthesized according to a procedure of Godage and coworkers.[19] To a solution of **S16** (200 mg, 0.7 mmol) and 5-phenyltetrazole (1.01 g, 6,89 mmol) in dry DCM (4.9 mL) under nitrogen atmosphere was added Xy phosphoamidite (1.24 g, 5.17 mmol). The suspension was sonicated (1 min) and further stirred overnight at rt. The reaction mixture was cooled to -40 °C, and *m*CPBA (1.70 g, 6.89 mmol) was added portion-wise while stirring. The cooling bath was removed, and the mixture was allowed to reach rt and diluted with DCM (50 mL), washed with 10% sodium sulfite solution (2 × 100 mL), dried and solvent evaporated *in vacuo*. The crude was immobilized on Telos® and purified by CombiFlash® chromatography on silica gel (120 g column, gradient: 1% → 4% → 10% MeOH in DCM) to afford the product **16** (651 mg, 0.54 mmol) as white solid in 79% yield.

^{1}H NMR (600 MHz, Chloroform-*d*) δ 8.06 (d, *J* = 7.1 Hz, 2H), 7.61 – 7.53 (m, 1H), 7.49 – 7.45 (m, 2H), 7.43 – 7.35 (m, 10H), 7.35 – 7.30 (m, 6H), 7.29 – 7.27 (m, 1H), 7.26 – 7.23 (m, 2H), 6.53 (s, 0.5H), 6.27 (s, 0.5H), 5.64 (ddd, *J* = 17.9, 13.9, 8.1 Hz, 4H), 5.58 – 5.47 (m, 3H), 5.41 (dd, *J* = 13.7, 9.4 Hz, 2H), 5.27 (s, 1H), 5.21 (dd, *J* = 13.5, 10.9 Hz, 3.5H), 5.16 (d, *J* = 9.8 Hz, 1H), 5.12 (d, *J* = 8.7 Hz, 1H), 5.05 (ddd, *J* = 22.1, 13.8, 5.8 Hz, 8H), 4.96 (s, 1.5H).

^{13}C NMR (151 MHz, Chloroform-*d*) δ 78.00 – 76.46 (m), 74.25 (td, *J* = 38.4, 36.8, 6.5 Hz), 70.59 (t, *J* = 38.5 Hz).

^{31}P NMR (122 MHz, $CDCl_3$) δ -3.32, -4.25, -4.90.

HRMS (ESI/Orbitrap) m/z: $[M + Na]^+$ Calcd for $C_{47}{}^{13}C_6H_{51}NaO_{22}P_5$ 1223.1613; Found 1223.1613.



[$^{13}C_6$]Benzoyl $InsP_5$ (**S17**)

[$^{13}C_6$]Benzoyl $InsP_5$ (**S17**) inositol was synthesized according to a procedure of Godage and coworkers.[19] **16** (50 mg, 0.04 mmol) was dissolved in methanol (3.8 mL), and water (0.95 mL) and 20% $Pd(OH)_2/C$ (50% wetted with water) (25 mg, 0.18 mmol) were added. The resulting suspension was stirred at rt overnight under hydrogen atmosphere. The catalyst was filtered through a PTFE syringe filter, and the filtrate neutralized by the addition of 1 M Et_3NHCO_3 until the pH 7.5 was reached. Lyophilization afforded the product **S17** (50 mg, 0.037 mmol, 8 × EtN_3 salt) as white

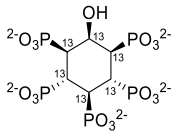
solid in 88% yield. The amount of the TEA-counter ion was determined by NMR spectroscopy.

^{1}H NMR (600 MHz, Deuterium Oxide, pD 7.0) δ 8.19 (dd, *J* = 8.2, 1.9 Hz, 2H), 7.74 (t, *J* = 7.2 Hz, 1H), 7.61 (dd, *J* = 9.0, 6.6 Hz, 2H), 6.10 (s, 0.5H), 5.84 (s, 0.5H), 4.58 – 4.45 (m, 2H), 4.39 (s, 0.5H), 4.27 (s, 1H), 4.16 (d, *J* = 11.1 Hz, 0.5H), 3.21 (qd, *J* = 7.3, 2.5 Hz, EtN_3), 1.30 (td, *J* = 7.3, 2.6 Hz, EtN_3).

^{13}C NMR (151 MHz, Deuterium Oxide, pD 7.0) δ 77.65 (t, *J* = 39.7 Hz), 76.46 (t, *J* = 39.7 Hz), 73.93 (t, *J* = 38.6 Hz), 72.53 (t, *J* = 39.9 Hz), 46.60, 8.23.

^{31}P NMR (122 MHz, Deuterium Oxide, pD 7.0) δ 0.33, -0.00, -0.80.

HRMS (ESI/Orbitrap) m/z: $[M - 2H]^{2-}$ calcd. for $C_7{}^{13}C_6H_{19}O_{22}P_5$ 343.9634; Found 343.9954.



$[^{13}C_6]InsP_5$

$[^{13}C_6]InsP_5$ was synthesized according to a procedure of Godage and coworkers.[19] Compound **S17** (50 mg, 0.037 mmol) was dissolved in concentrated aqueous ammonia solution (2.0 mL) and heated at 60 °C overnight in a Pyrex pressure tube. After evaporation of the solution under vacuum, the residue was dissolved in water and the benzamide byproduct was removed by washing with DCM. The ammonium salt of the product was obtained by evaporation of the ammonia and converted into the free acid by quick filtration (Note: Prolonged exposure causes phosphoryl group migration.) through DOWEX 50W x 8 (H^+ form) (10-fold molar

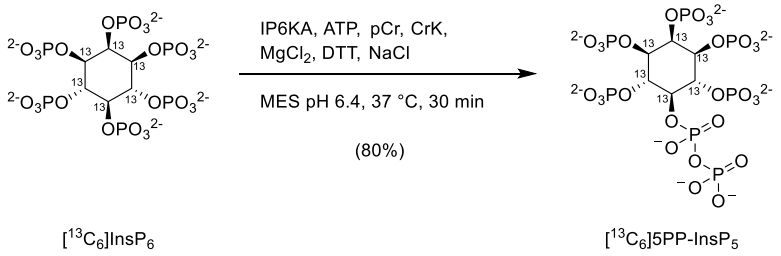
excess, previously washed with MilliQ® water) and then to its hexasodium salt by titration to pH 7.40 with 0.1 M sodium hydroxide solution. Lyophilization afforded [$^{13}C_6$]InsP$_5$ (17 mg, 0.024 mmol) as white solid in 65% yield.

^{1}H NMR (300 MHz, Deuterium Oxide, pD 7.5) δ 4.72 – 4.53 (m, 1.5H), 4.47 – 4.25 (m, 1.5H), 4.24 – 4.02 (m, 1.5H), 3.96 – 3.74 (m, 1.5H).

^{13}C NMR (75 MHz, Deuterium Oxide, pD 7.5) δ 78.28 – 74.98 (m), 73.92 (t, *J* = 37.7 Hz), 70.68 (t, *J* = 37.6 Hz).

^{31}P NMR (122 MHz, D_2O, pD 7.5) δ 1.49, 1.23, 0.86.

Chemoenzymatic Synthesis of [$^{13}C_6$]Inositol pyrophosphates



[$^{13}C_6$]InsP$_6$ [$^{13}C_6$]5PP-InsP$_5$

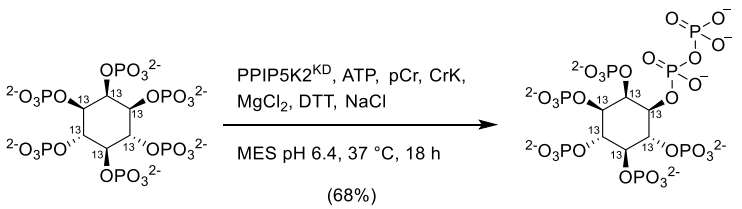
[$^{13}C_6$]5PP-InsP$_5$

The synthesis of [$^{13}C_6$]5PP-InsP$_5$ was done according the synthesis of 5PP-InsP$_5$ which is described in the method section of chapter 2.

^{1}H NMR (600 MHz, Deuterium Oxide, pD 7.5) δ 5.05 – 4.95 (m, 0.5H), 4.69 – 4.58 (m, 1H), 4.49 – 4.36 (m, 1.5H), 4.35 – 4.26 (m, 1H), 4.25 – 4.15 (m, 0.5H), 4.12 – 4.00 (m, 1H).

^{13}C NMR (151 MHz, Deuterium Oxide, pD 7.5) δ 80.74 – 79.46 (m), 79.32 – 77.17 (m), 76.03 (t, *J* = 38.7 Hz).

^{31}P NMR (243 MHz, D_2O, pD 7.5) δ -1.45, -2.00, -2.45, -10.32, -13.41.

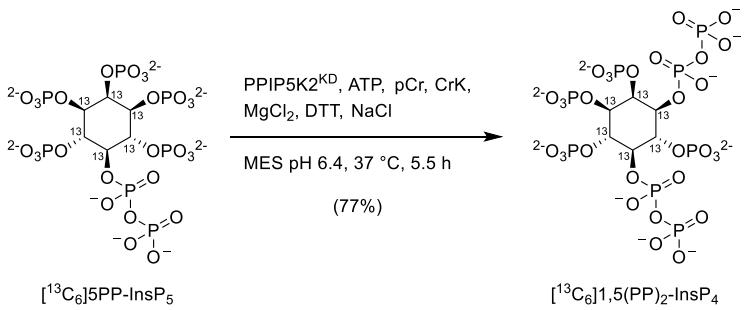


[$^{13}C_6$]InsP$_6$ [$^{13}C_6$]1PP-InsP$_5$

[$^{13}C_6$]1PP-InsP$_5$

The synthesis of [$^{13}C_6$]1PP-InsP$_5$ was done according the synthesis of 1PP-InsP$_5$ which is described in the method section of chapter 2.

^{1}H NMR (600 MHz, Deuterium Oxide) δ 5.28 – 5.21 (m, 0.5H), 5.04 – 4.96 (m, 0.5H), 4.64 – 4.57 (m, 0.5H), 4.55 – 4.45 (m, 2H), 4.41 – 4.33 (m, 0.5H), 4.31 – 4.21 (m, 1H).
^{13}C NMR (151 MHz, Deuterium Oxide) δ 77.47 (t, *J* = 40.0 Hz), 76.70 – 74.61 (m), 73.50 (dt, *J* = 82.6, 40.4 Hz).
^{31}P NMR (243 MHz, Deuterium Oxide) δ -1.82, -2.35, -2.95, -3.91, -13.25 (d, *J* = 21.9 Hz), -14.14 (d, *J* = 19.9 Hz).



[$^{13}C_6$]5PP-InsP$_5$ [$^{13}C_6$]1,5(PP)$_2$-InsP$_4$

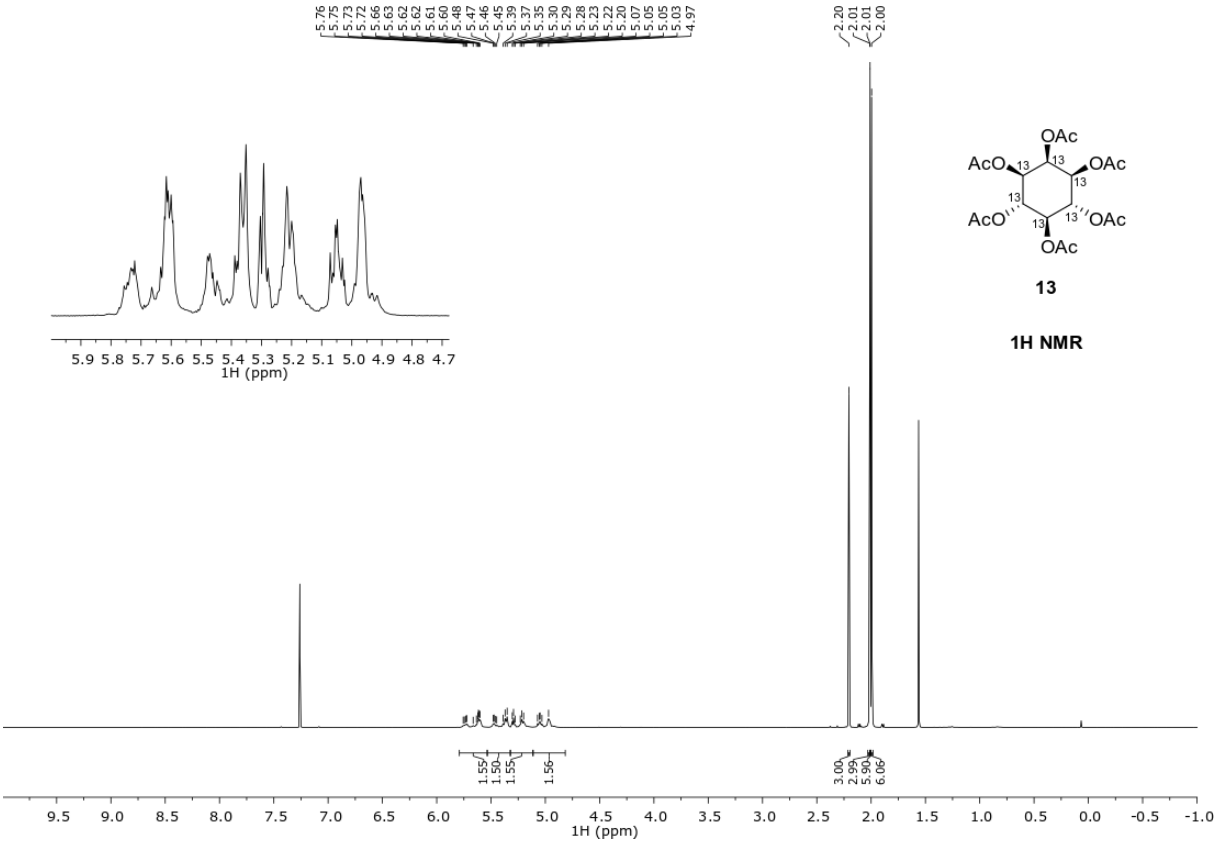
[$^{13}C_6$]1,5(PP)$_2$-InsP$_4$

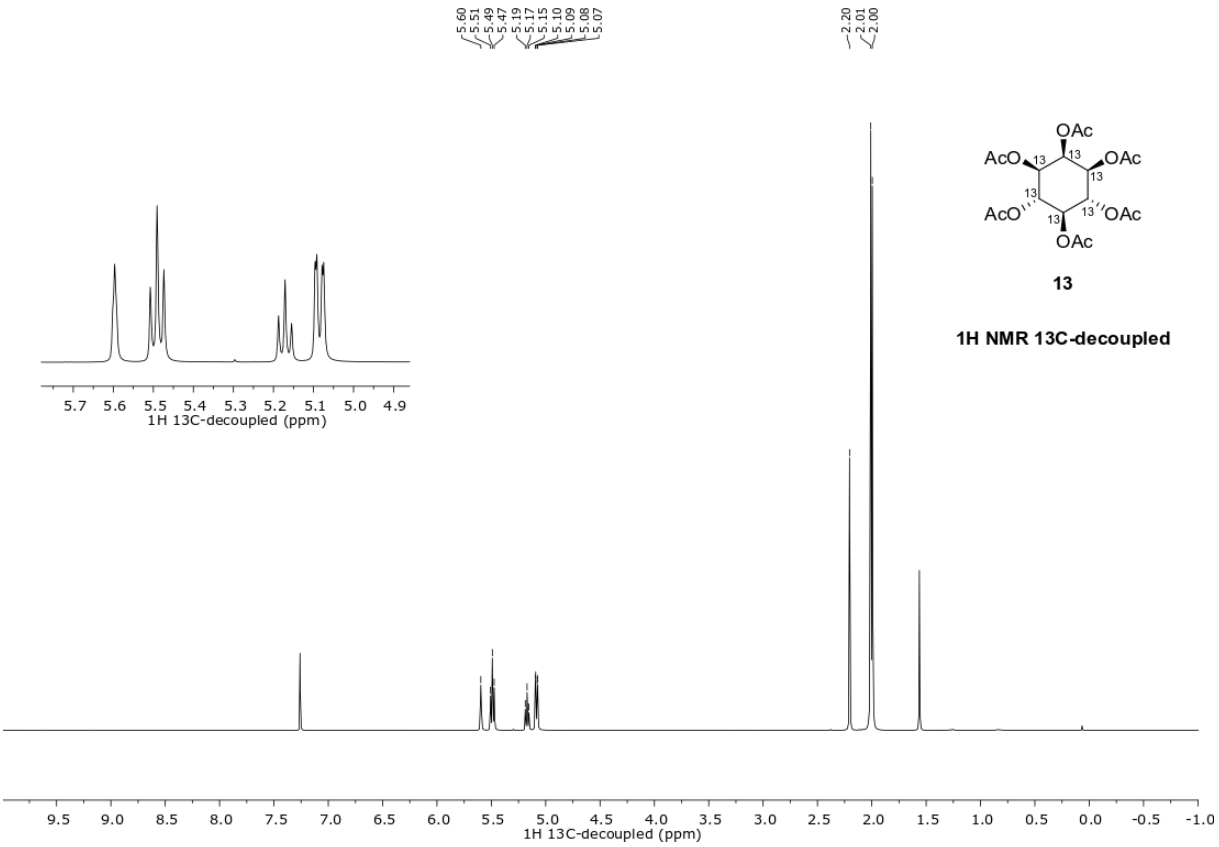
The synthesis of [$^{13}C_6$]1,5(PP)$_2$-InsP$_4$ was done according the synthesis of 1,5(PP)$_2$-InsP$_4$ which is described in the method section of chapter 2.

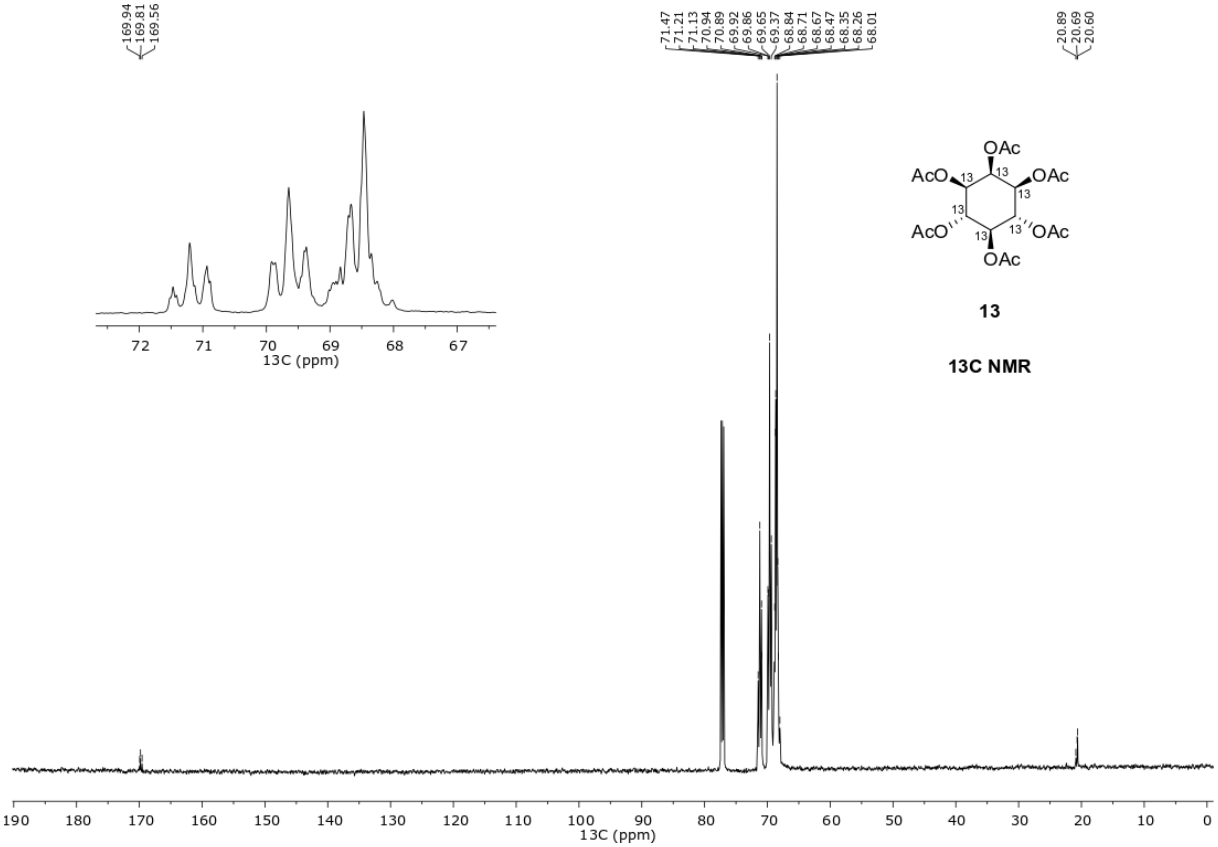
^{1}H NMR (600 MHz, Deuterium Oxide) δ 5.27 – 5.20 (m, 0.5H), 5.01 – 4.94 (m, 0.5H), 4.65 – 4.58 (m, 1H), 4.56 – 4.47 (m, 1.5H), 4.41 – 4.33 (m, 1H), 4.31 – 4.21 (m, 0.5H).
^{13}C NMR (151 MHz, Deuterium Oxide, pD 5.0) δ 77.53 (t, *J* = 39.4 Hz), 76.68 – 74.55 (m), 73.48 (q, *J* = 42.5, 39.9 Hz).

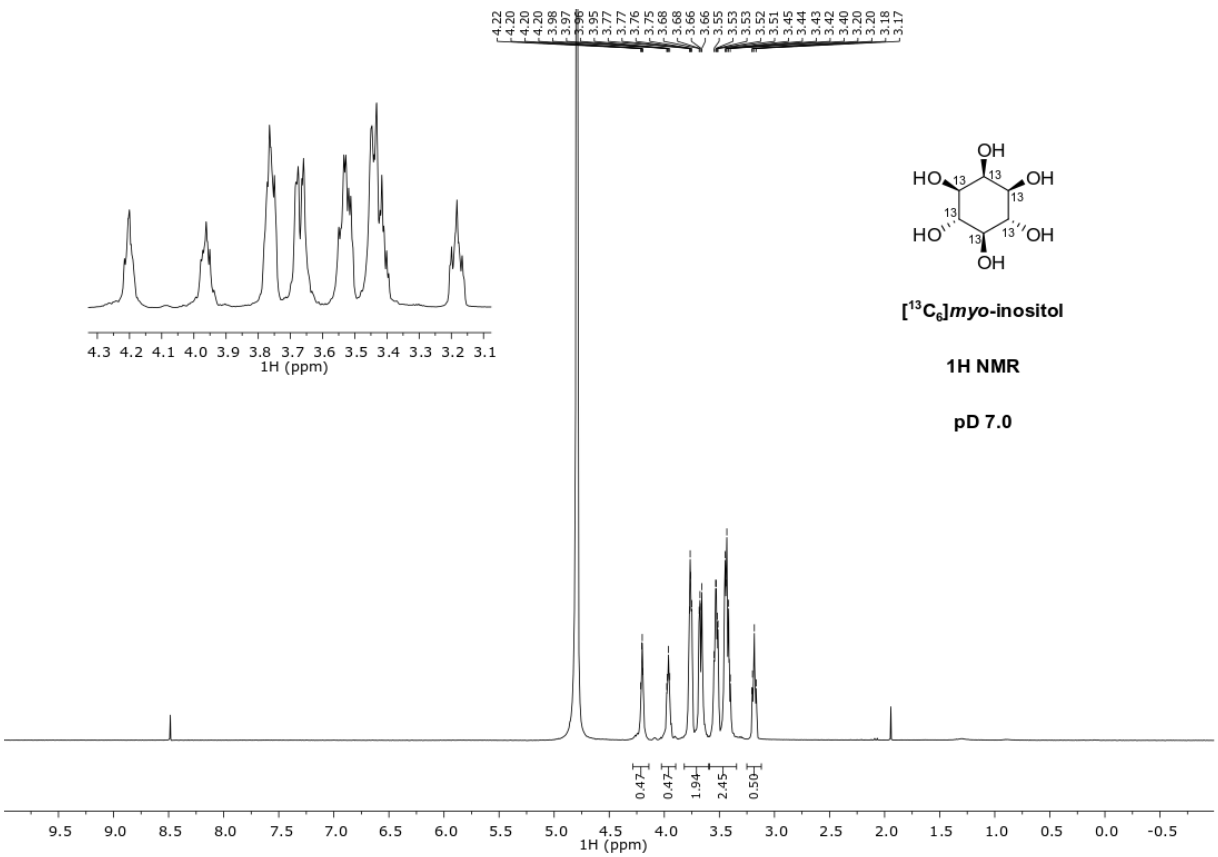
^{31}P NMR (243 MHz, Deuterium Oxide, pD 5.0) δ 0.55, 0.30, -0.19, -1.04, -10.53 (t, *J* = 21.3 Hz), -11.10 (dd, *J* = 50.4, 20.0 Hz).

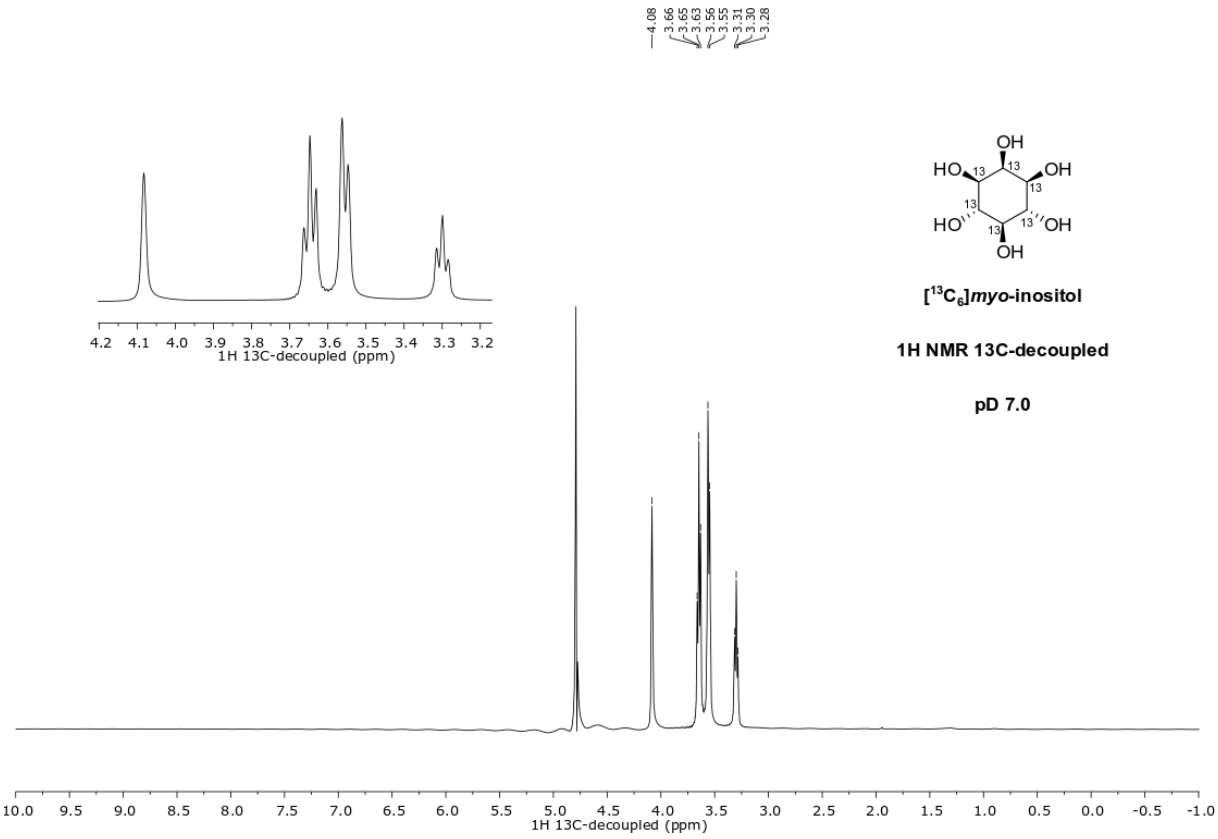
NMR Spectra

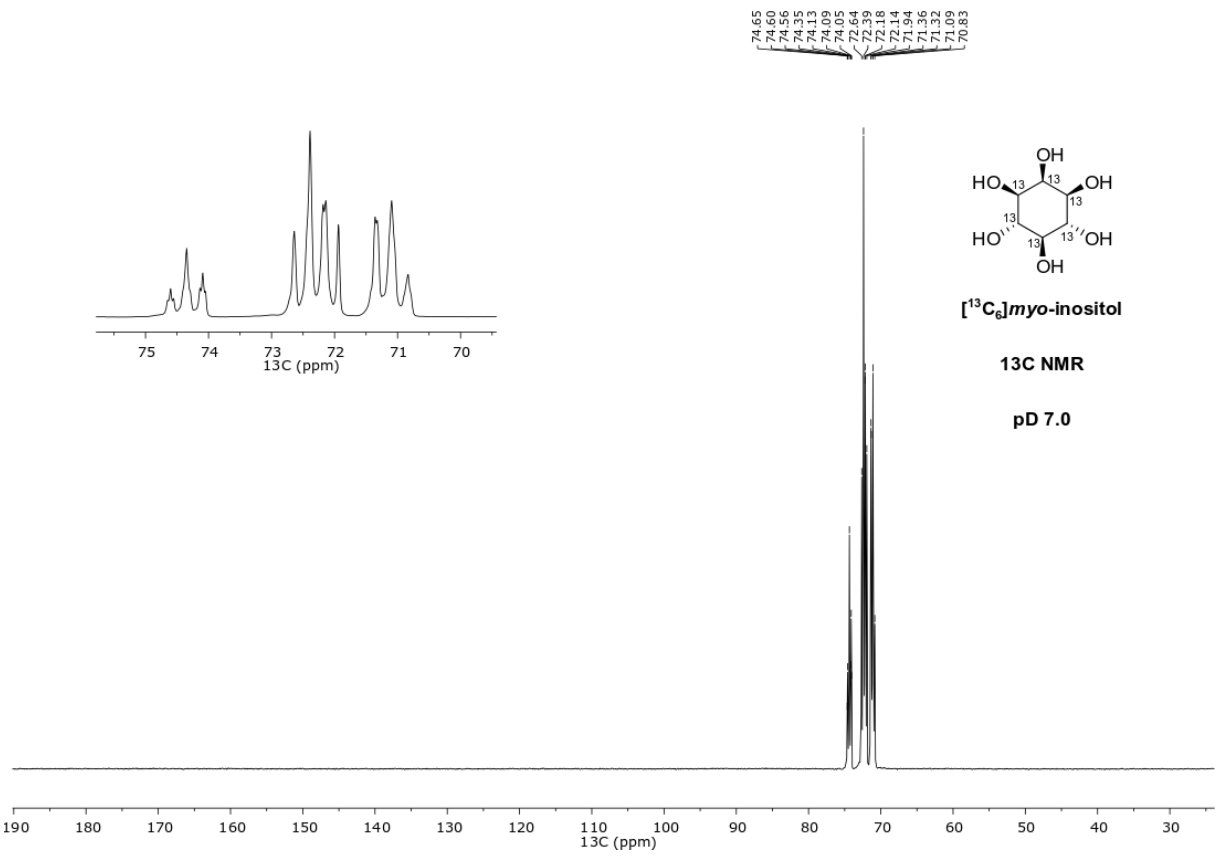

5.76 5.75 5.73 5.72 5.66 5.63 5.62 5.61 5.60 5.48 5.47 5.46 5.45 5.39 5.37 5.35 5.30 5.29 5.28 5.23 5.22 5.20 5.07 5.05 5.03 4.97
2.20 2.01 2.01 2.00
5.9 5.8 5.7 5.6 5.5 5.4 5.3 5.2 5.1 5.0 4.9 4.8 4.7
1H (ppm)
1.55 1.50 1.55 1.56
3.00 2.99 5.90 6.06
9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5 -1.0
1H (ppm)
OAc
AcO
OAc
AcO
OAc
OAc
13
1H NMR

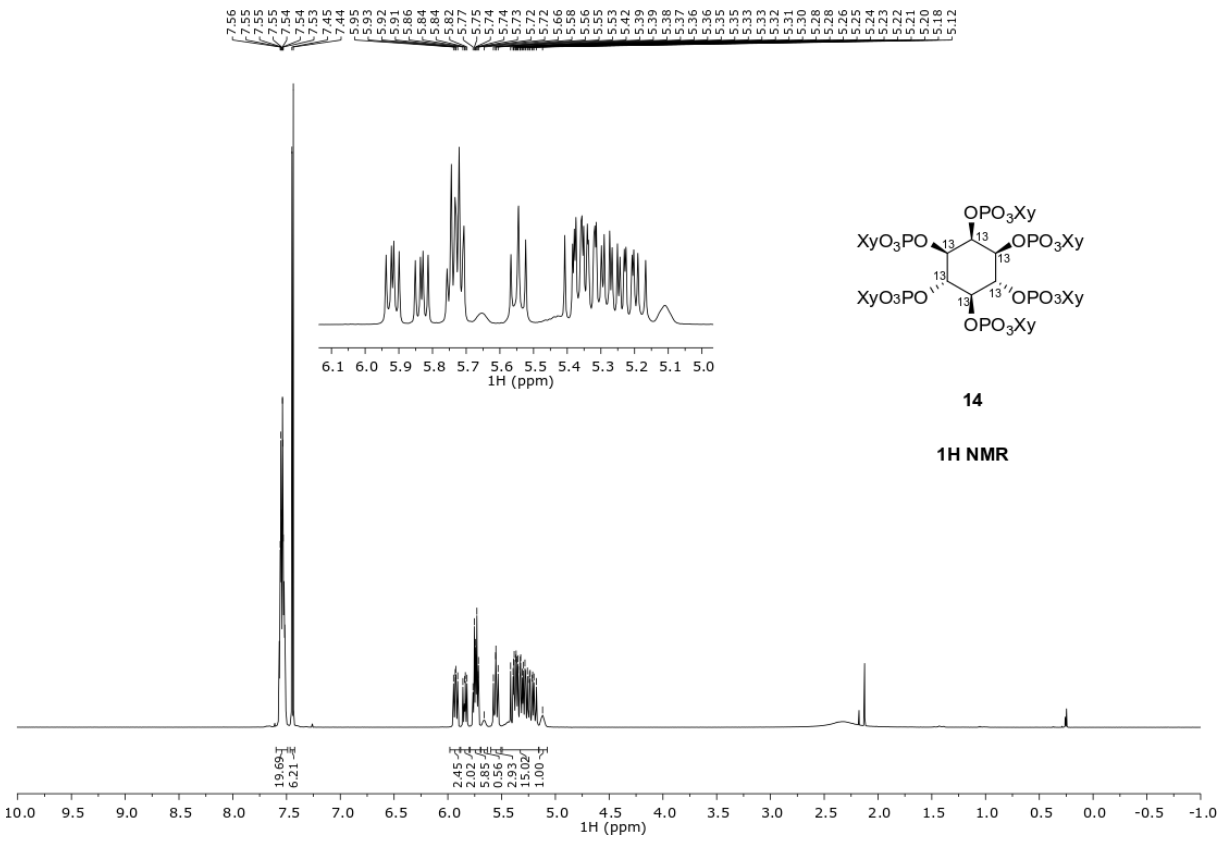

5.60 5.51 5.49 5.47 5.19 5.17 5.15 5.10 5.09 5.08 5.07
2.20 2.01 2.00
5.7 5.6 5.5 5.4 5.3 5.2 5.1 5.0 4.9
1H 13C-decoupled (ppm)
9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5 -1.0
1H 13C-decoupled (ppm)
OAc
AcO
OAc
AcO
OAc
OAc
13
1H NMR 13C-decoupled

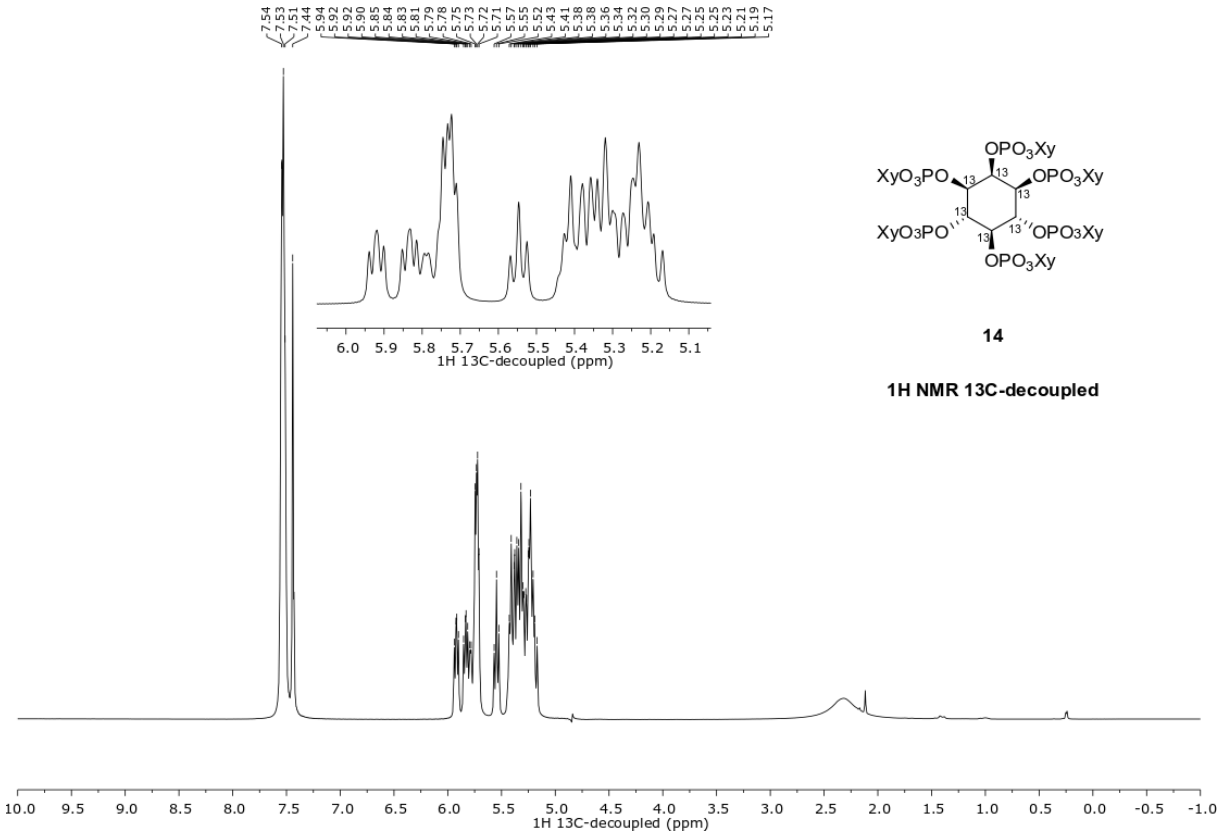

169.94
169.81
169.56
71.47
71.21
71.13
70.94
70.89
69.92
69.86
69.65
69.37
68.84
68.71
68.67
68.47
68.35
68.26
68.01
20.89
20.69
20.60
72 71 70 69 68 67
13C (ppm)
OAc
AcO
OAc
AcO
OAc
OAc
13
13C NMR
190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
13C (ppm)

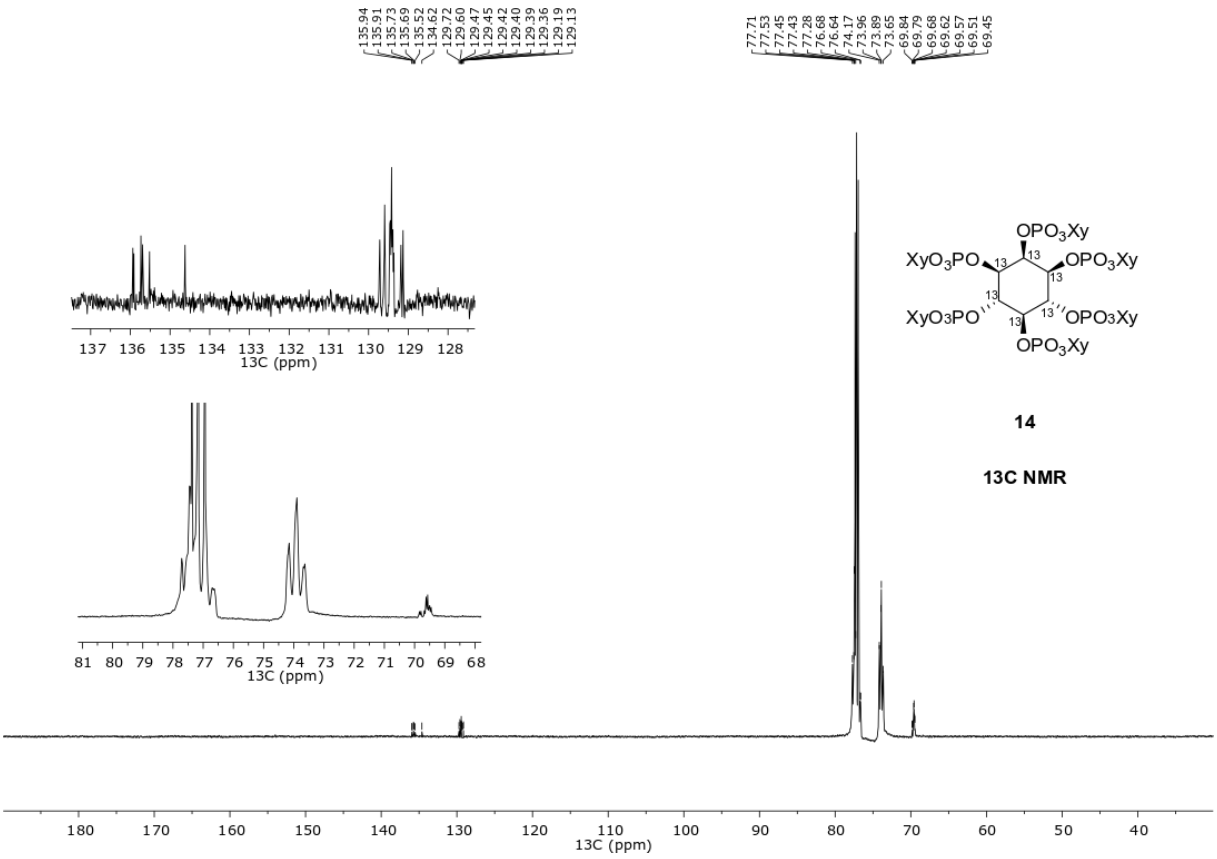

4.22
4.20
4.20
4.20
3.98
3.97
3.96
3.95
3.77
3.77
3.76
3.75
3.68
3.68
3.66
3.66
3.55
3.53
3.53
3.52
3.51
3.45
3.44
3.43
3.42
3.40
3.20
3.20
3.18
3.17
4.3 4.2 4.1 4.0 3.9 3.8 3.7 3.6 3.5 3.4 3.3 3.2 3.1
1H (ppm)
OH
HO
OH
HO
OH
OH
[13C6]myo-inositol
1H NMR
pD 7.0
0.47
0.47
1.94
2.45
0.50
9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
1H (ppm)

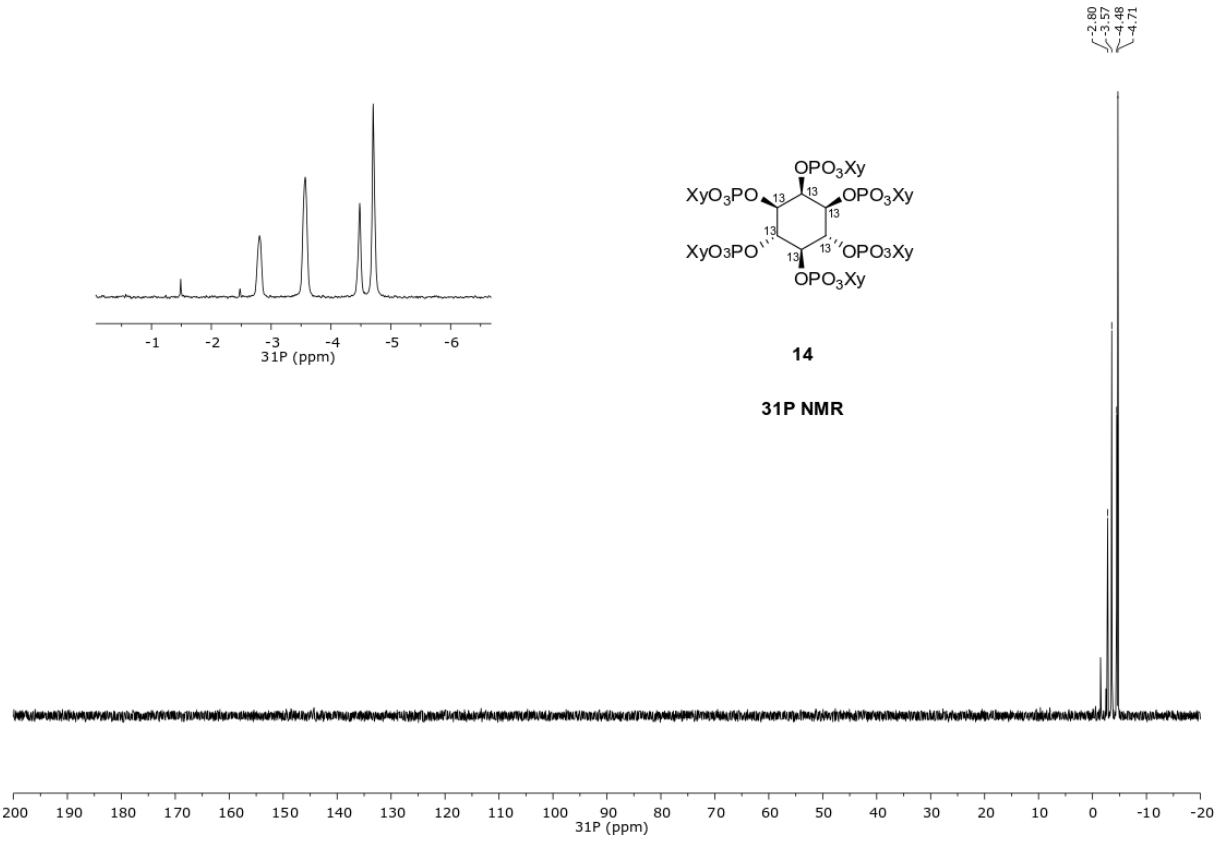

4.08
3.66
3.65
3.63
3.56
3.55
3.31
3.30
3.28
4.2 4.1 4.0 3.9 3.8 3.7 3.6 3.5 3.4 3.3 3.2
1H 13C-decoupled (ppm)
OH
HO
OH
HO
OH
13
[13C6]myo-inositol
1H NMR 13C-decoupled
pD 7.0
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5 -1.0
1H 13C-decoupled (ppm)

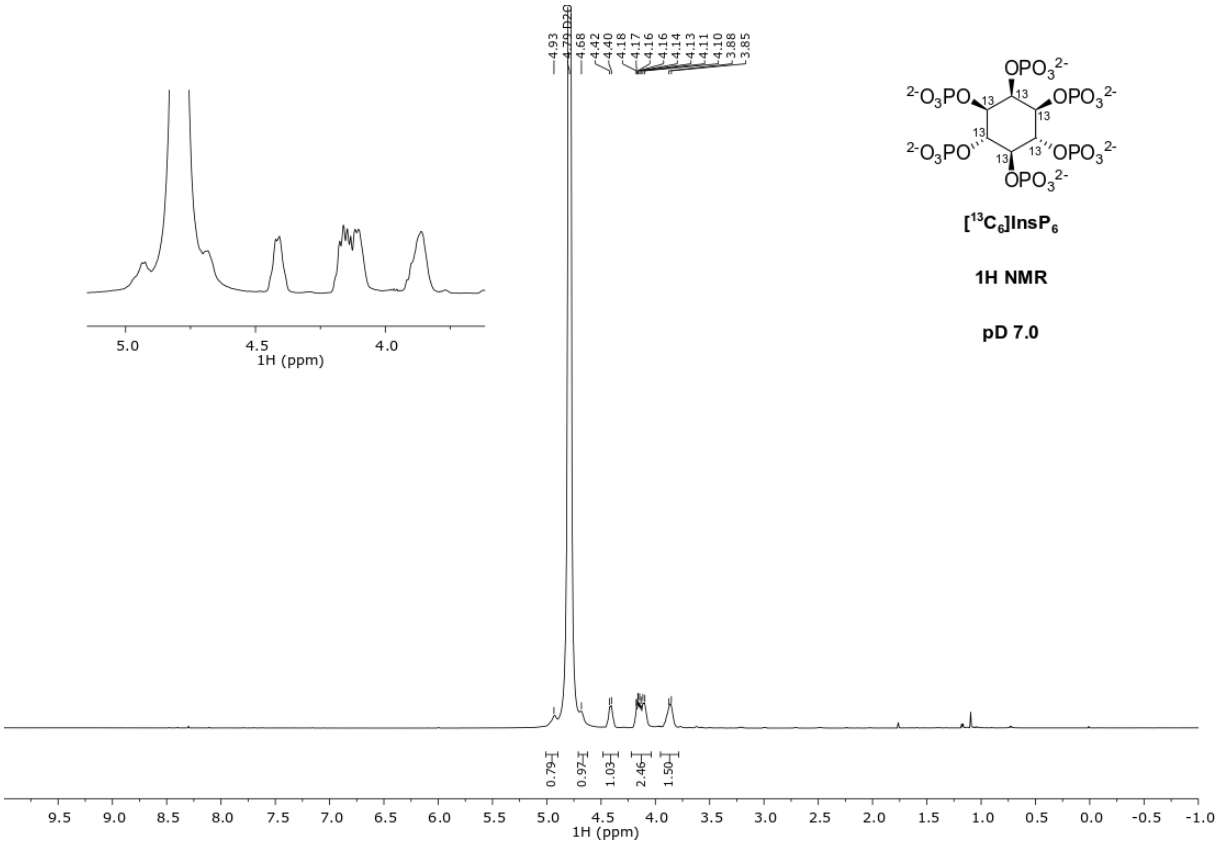

74.65
74.60
74.56
74.35
74.13
74.09
74.05
72.64
72.39
72.18
72.14
71.94
71.36
71.32
71.09
70.83
75 74 73 72 71 70
13C (ppm)
OH
HO
OH
HO
OH
13
[13C6]myo-inositol
13C NMR
pD 7.0
190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30
13C (ppm)


OPO3Xy
XyO3PO
OPO3Xy
XyO3PO
OPO3Xy
OPO3Xy
14
1H NMR
1H (ppm)


OPO3Xy
XyO3PO
OPO3Xy
XyO3PO
OPO3Xy
OPO3Xy
14
1H NMR 13C-decoupled
1H 13C-decoupled (ppm)


135.94
135.91
135.73
135.69
135.52
134.62
129.72
129.60
129.47
129.45
129.42
129.40
129.39
129.36
129.19
129.13
77.71
77.53
77.45
77.43
77.28
76.68
76.64
74.17
73.96
73.89
73.65
69.84
69.79
69.68
69.62
69.57
69.51
69.45
137 136 135 134 133 132 131 130 129 128
13C (ppm)
81 80 79 78 77 76 75 74 73 72 71 70 69 68
13C (ppm)
180 170 160 150 140 130 120 110 100 90 80 70 60 50 40
13C (ppm)
OPO3Xy
XyO3PO
OPO3Xy
XyO3PO
OPO3Xy
OPO3Xy
14
13C NMR

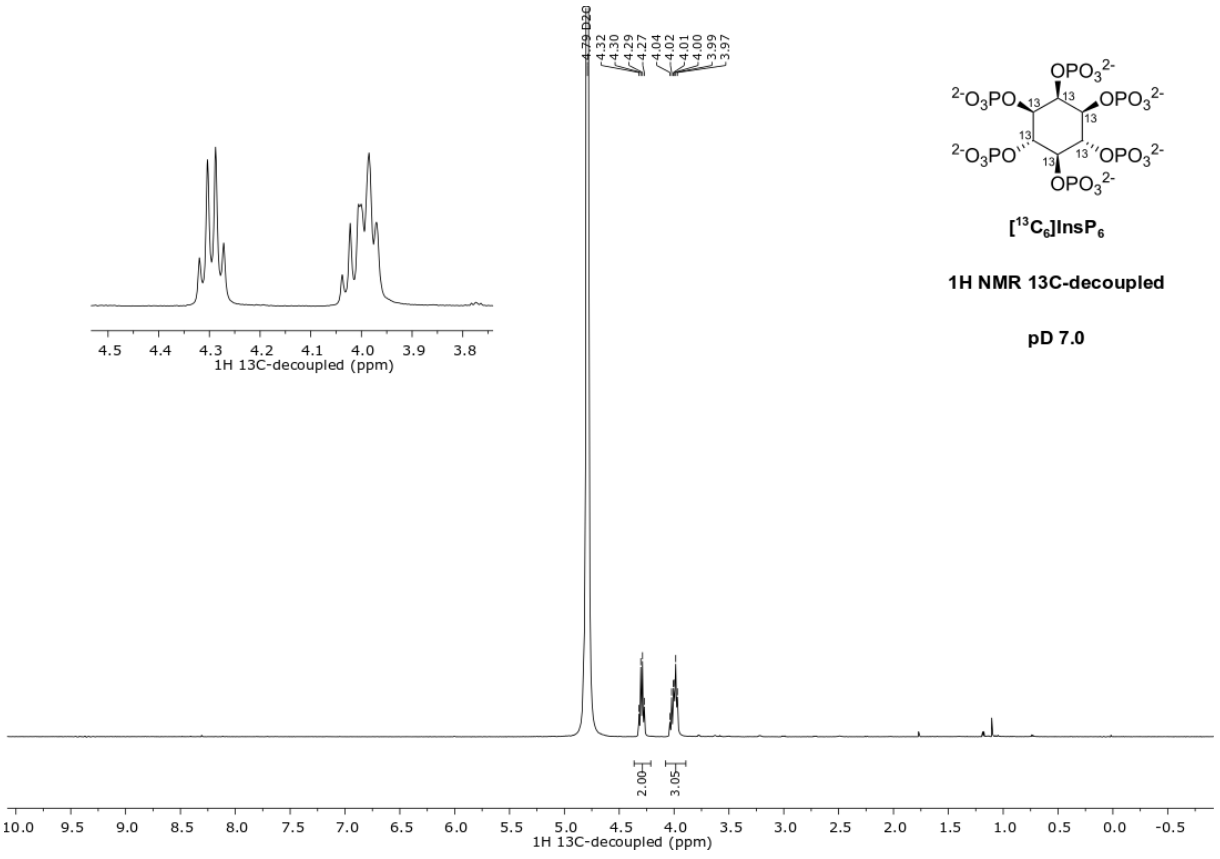

-2.80
-3.57
-4.48
-4.71
-1 -2 -3 -4 -5 -6
31P (ppm)
OPO3Xy
XyO3PO
OPO3Xy
XyO3PO
OPO3Xy
OPO3Xy
14
31P NMR
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10 -20
31P (ppm)



[$^{13}C_6$]InsP$_6$

1H NMR

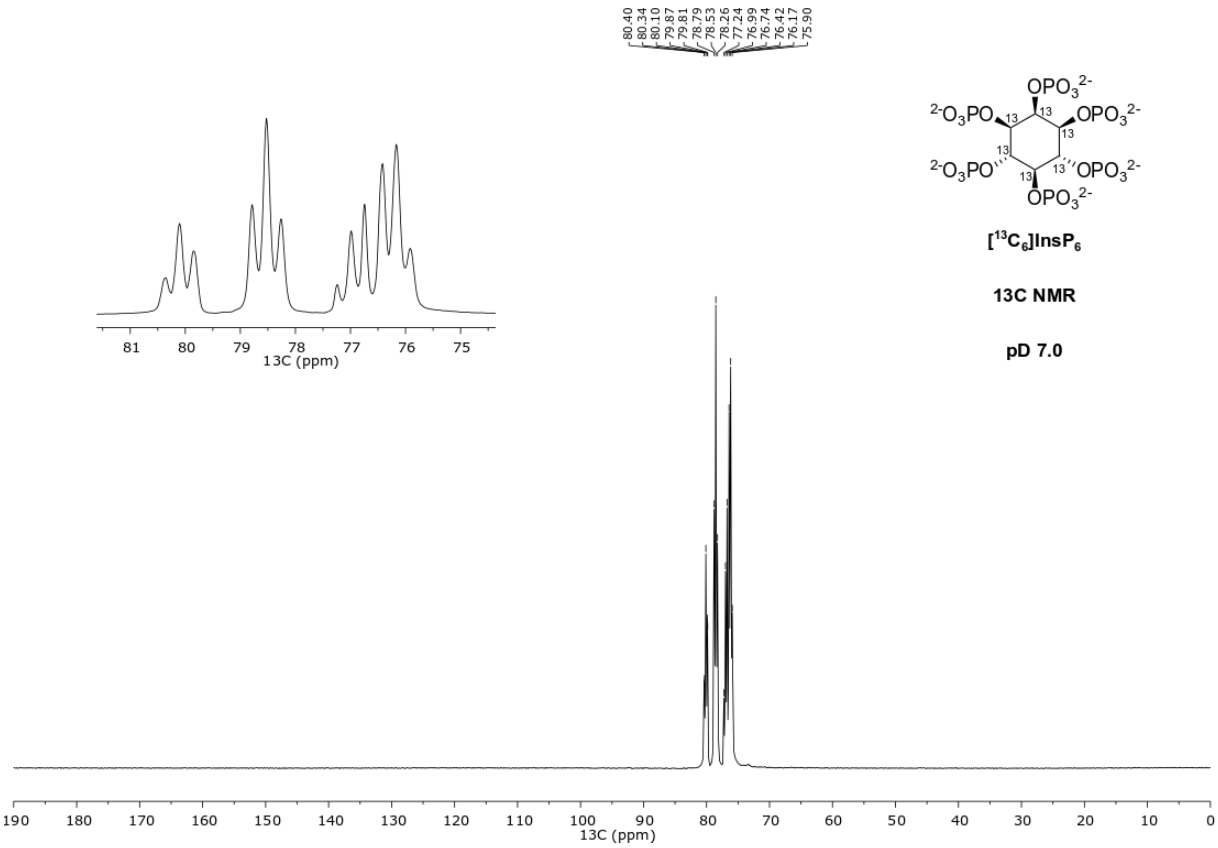
pD 7.0



[$^{13}C_6$]InsP$_6$

1H NMR 13C-decoupled

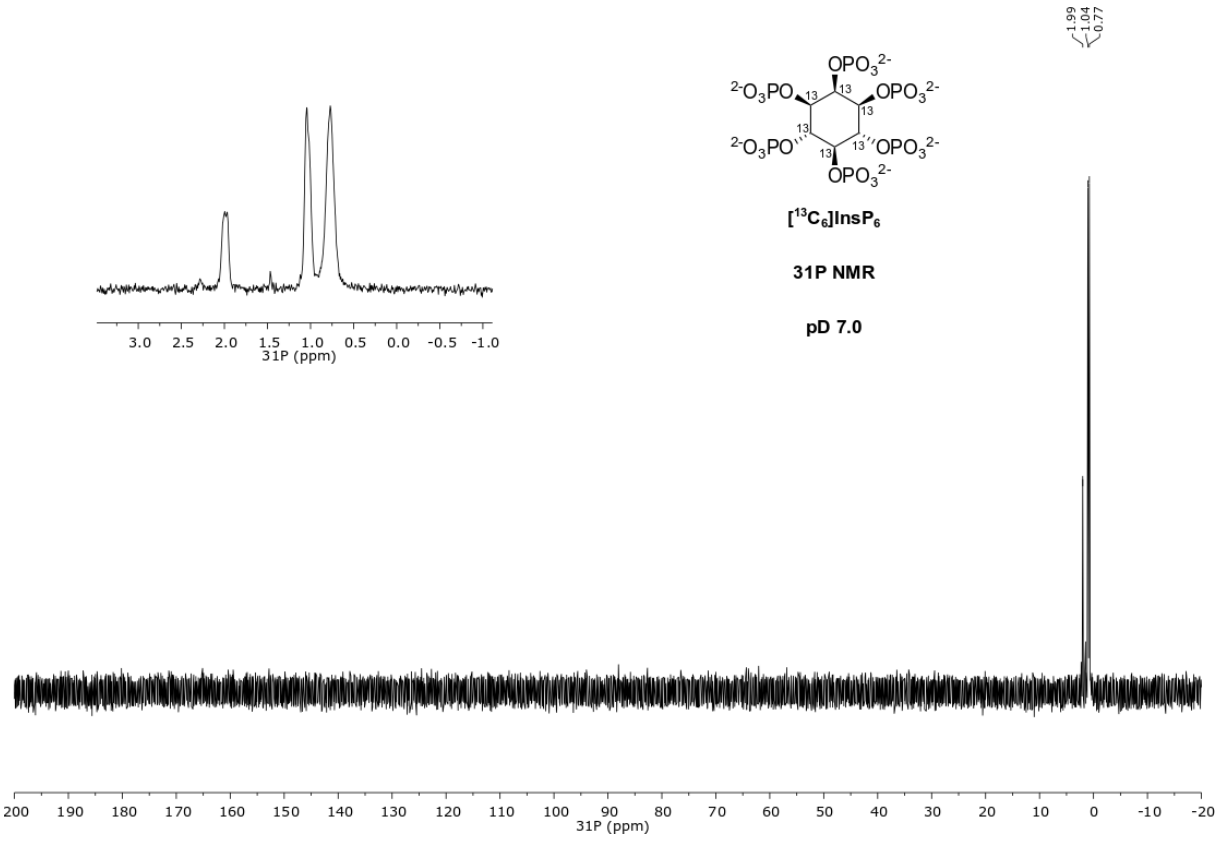
pD 7.0



[$^{13}C_6$]InsP$_6$

13C NMR

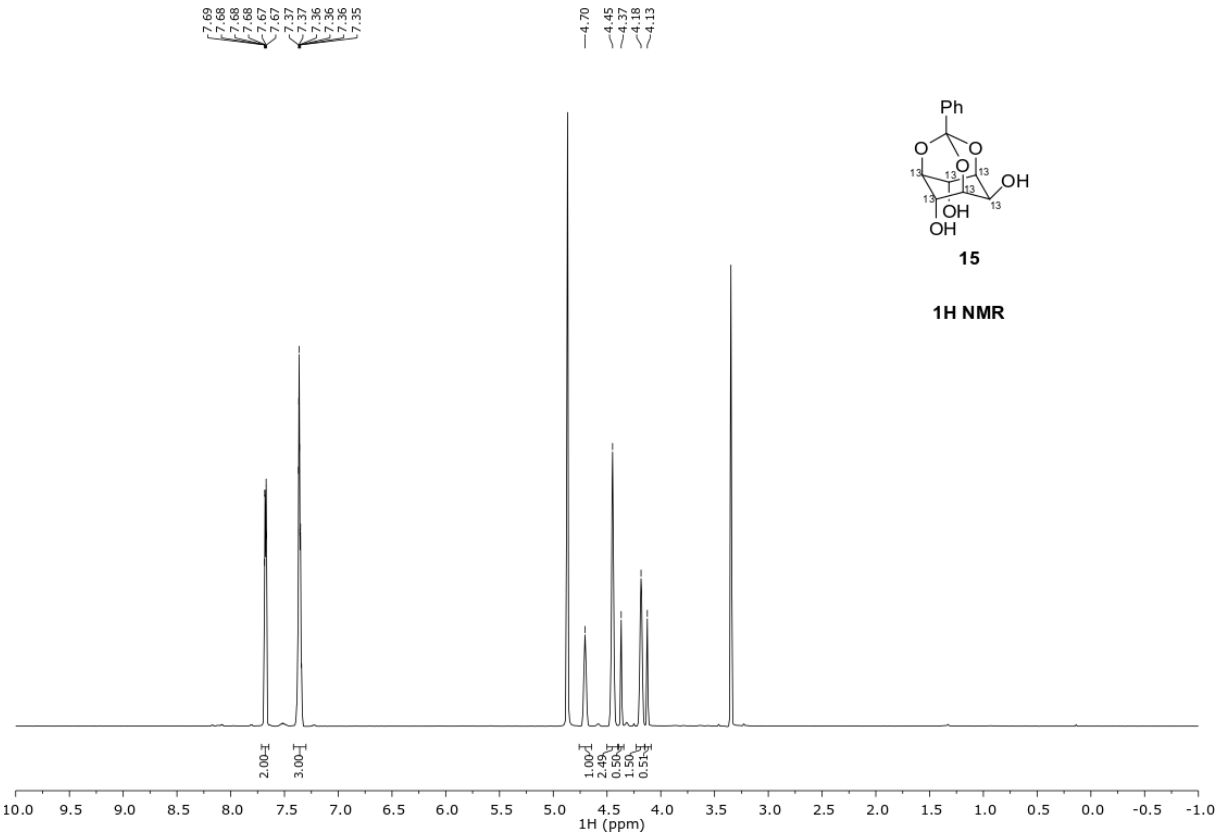
pD 7.0



[$^{13}C_6$]InsP$_6$

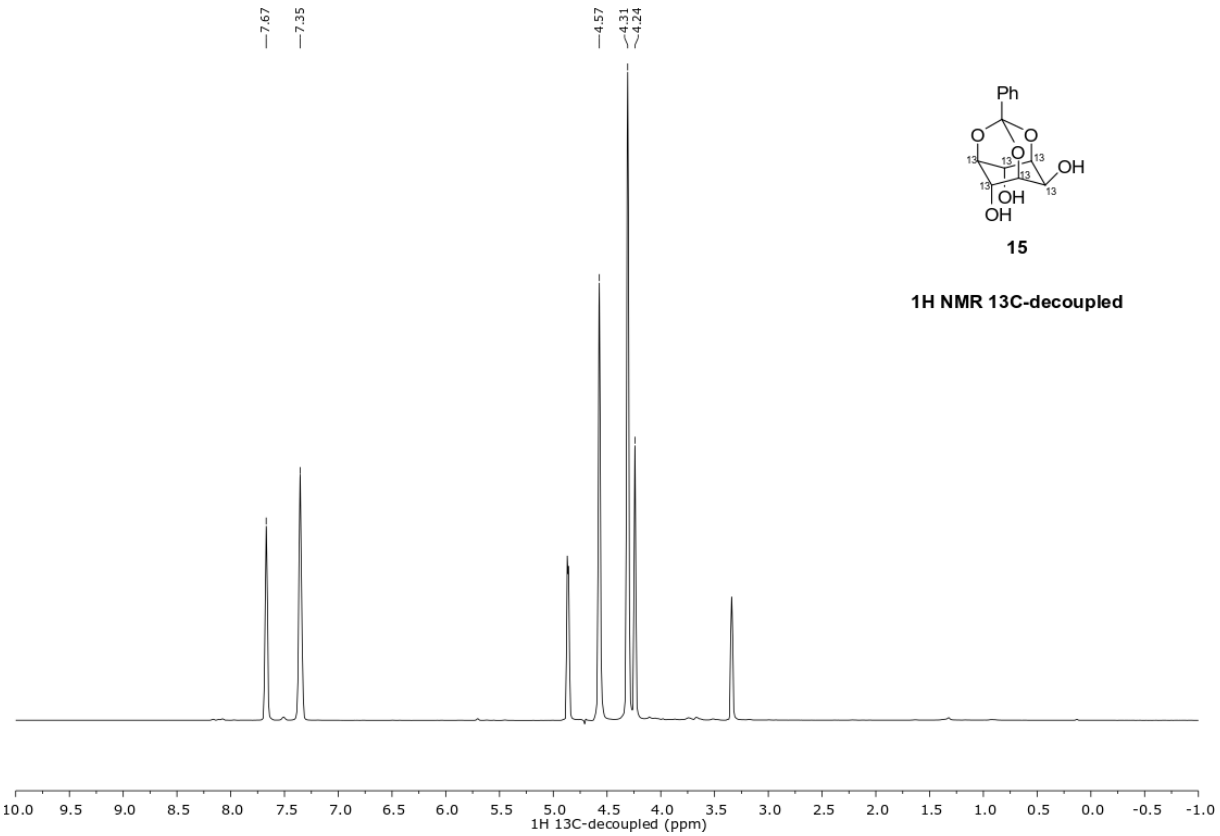
31P NMR

pD 7.0



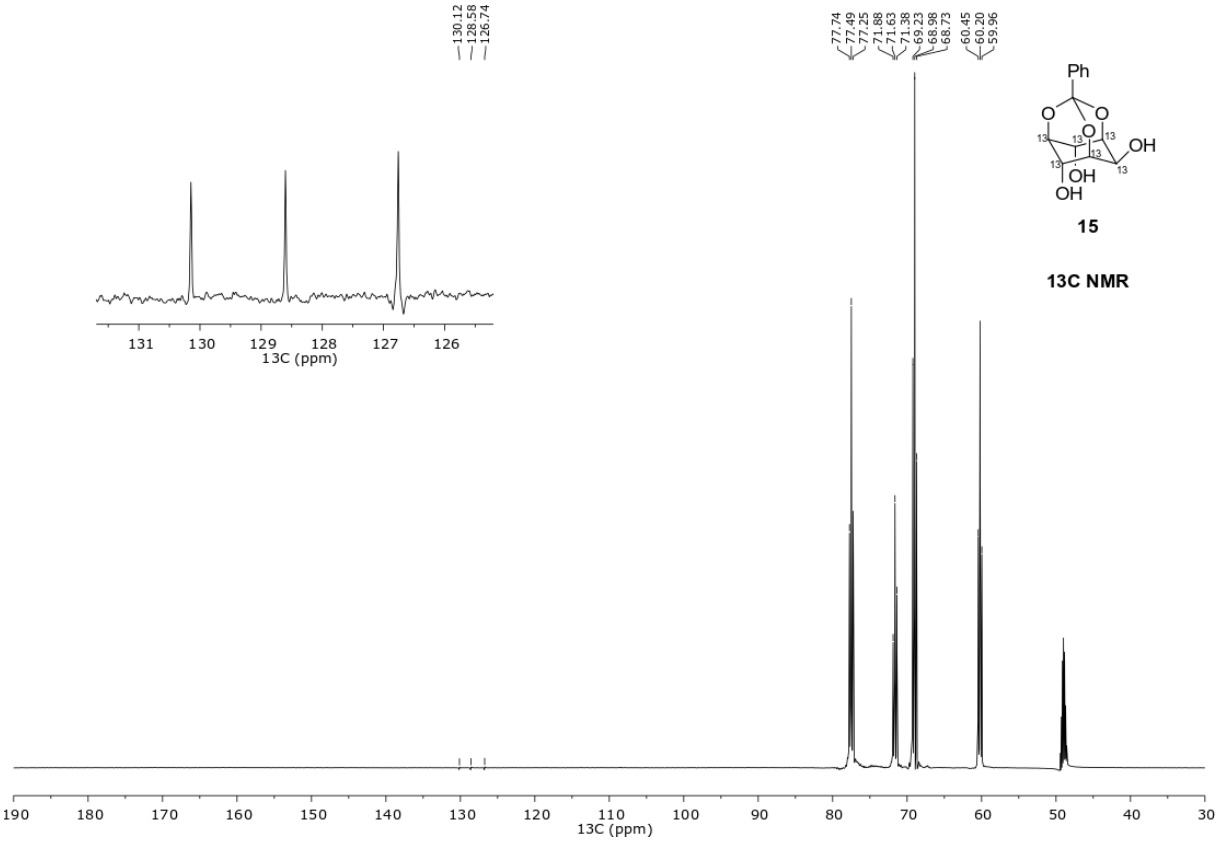
15

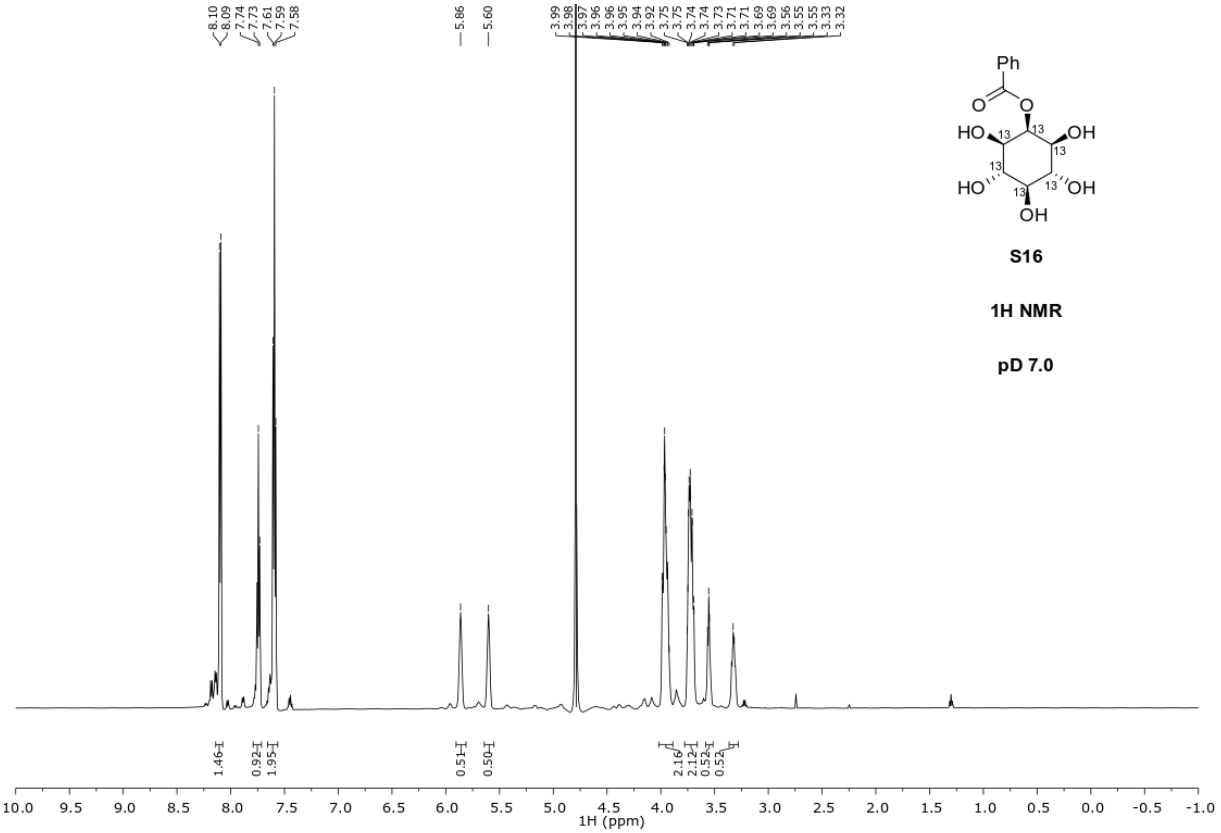
1H NMR

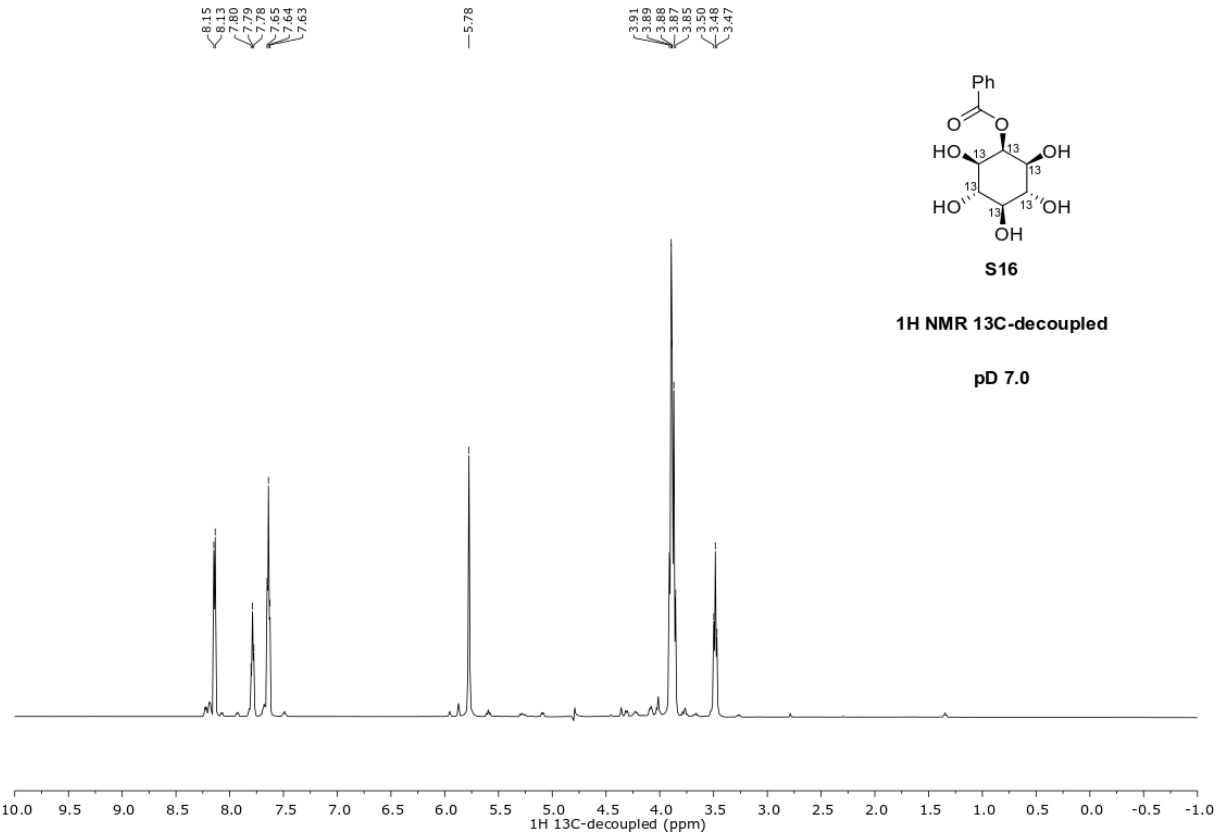


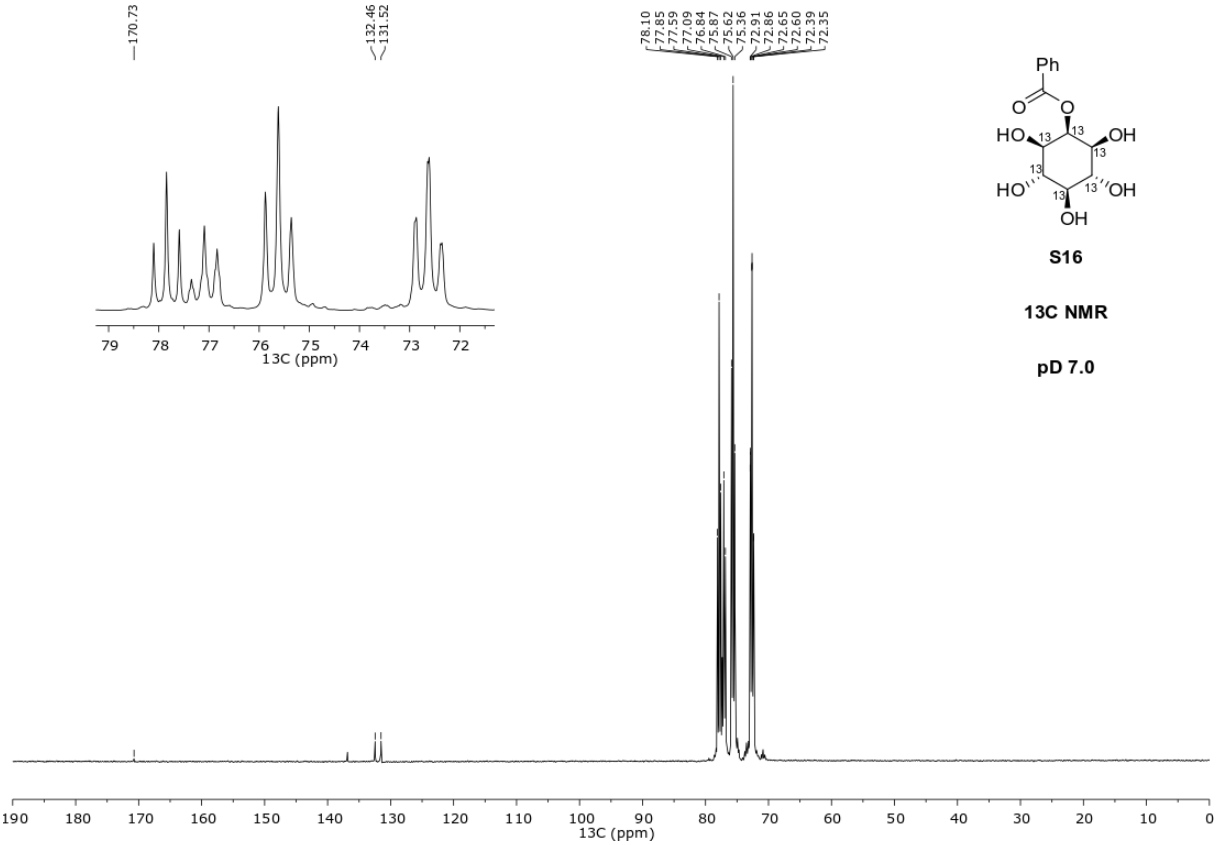
15

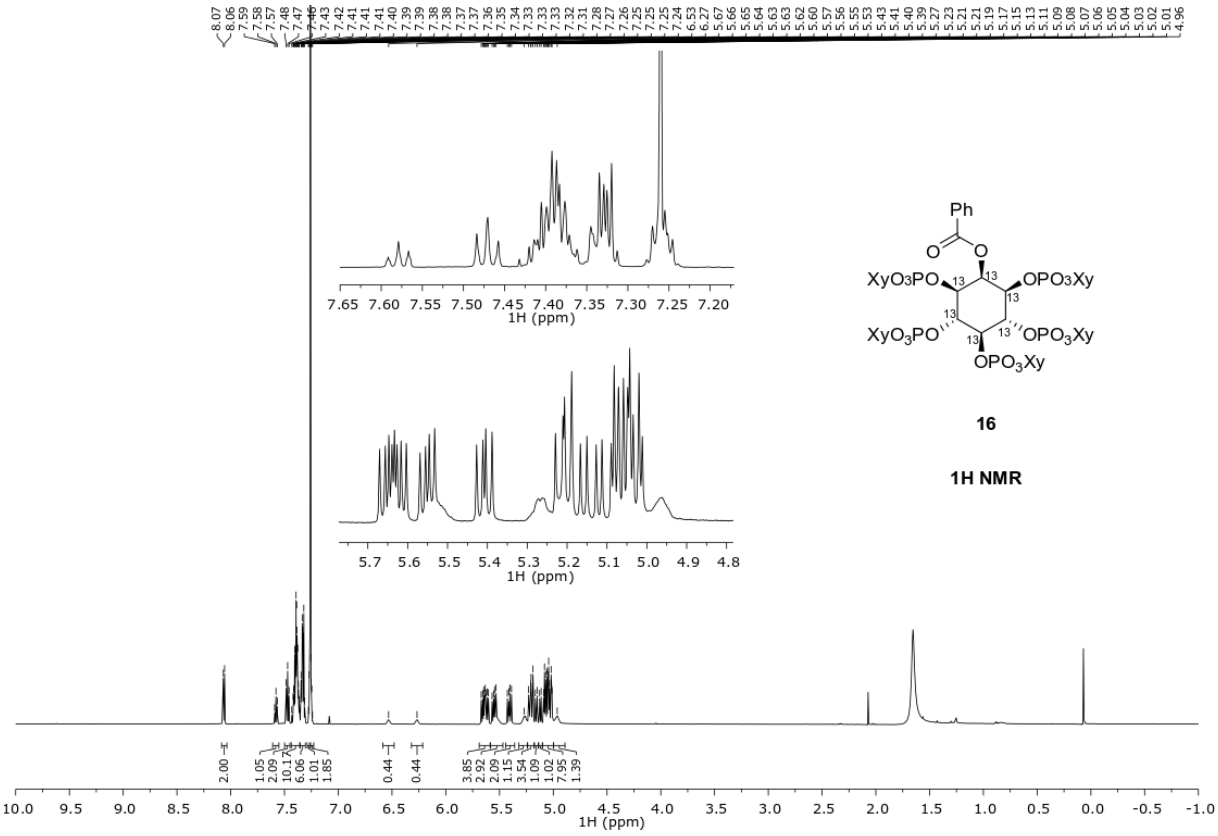
1H NMR 13C-decoupled

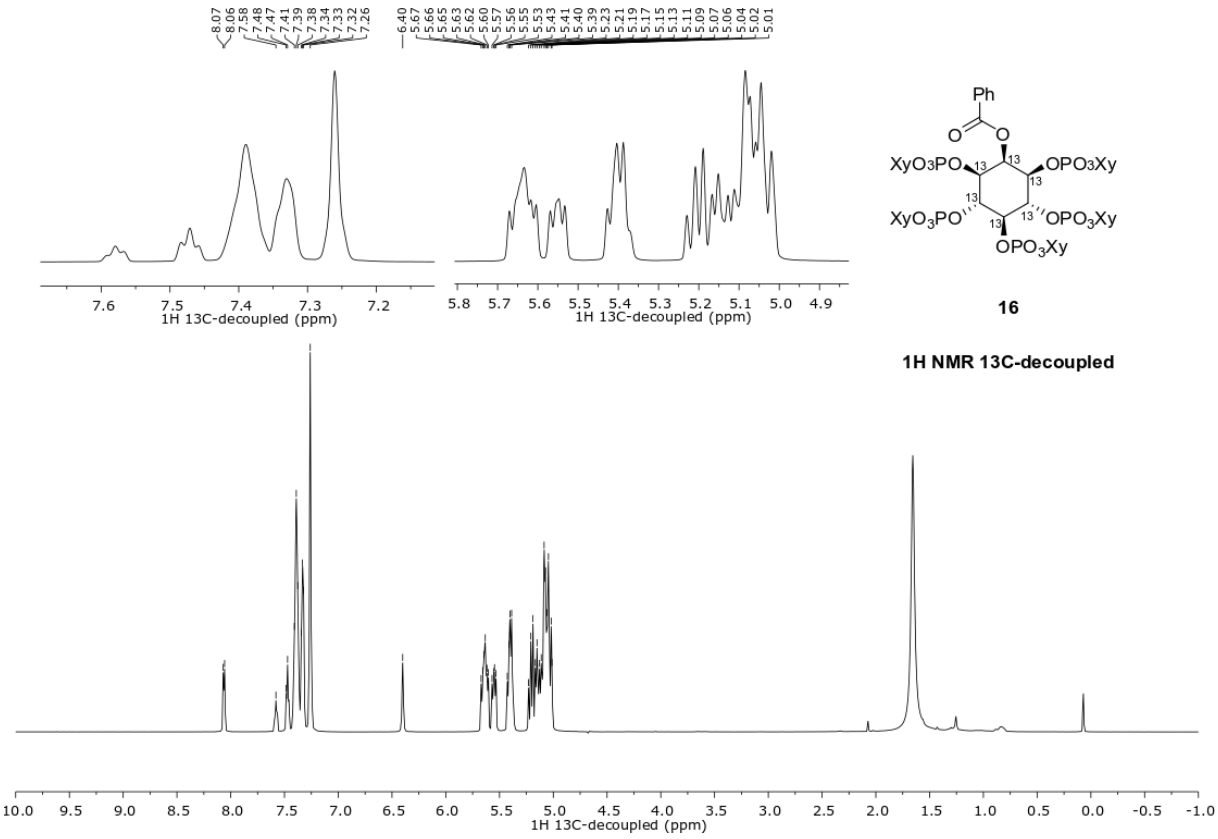

130.12
128.58
126.74
77.74
77.49
77.25
71.88
71.63
71.38
69.23
68.96
68.73
60.45
60.20
59.96
Ph
O
O
O
13
13
13
13
13
13
OH
OH
OH
15
13C NMR
131
130
129
128
127
126
13C (ppm)
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
13C (ppm)

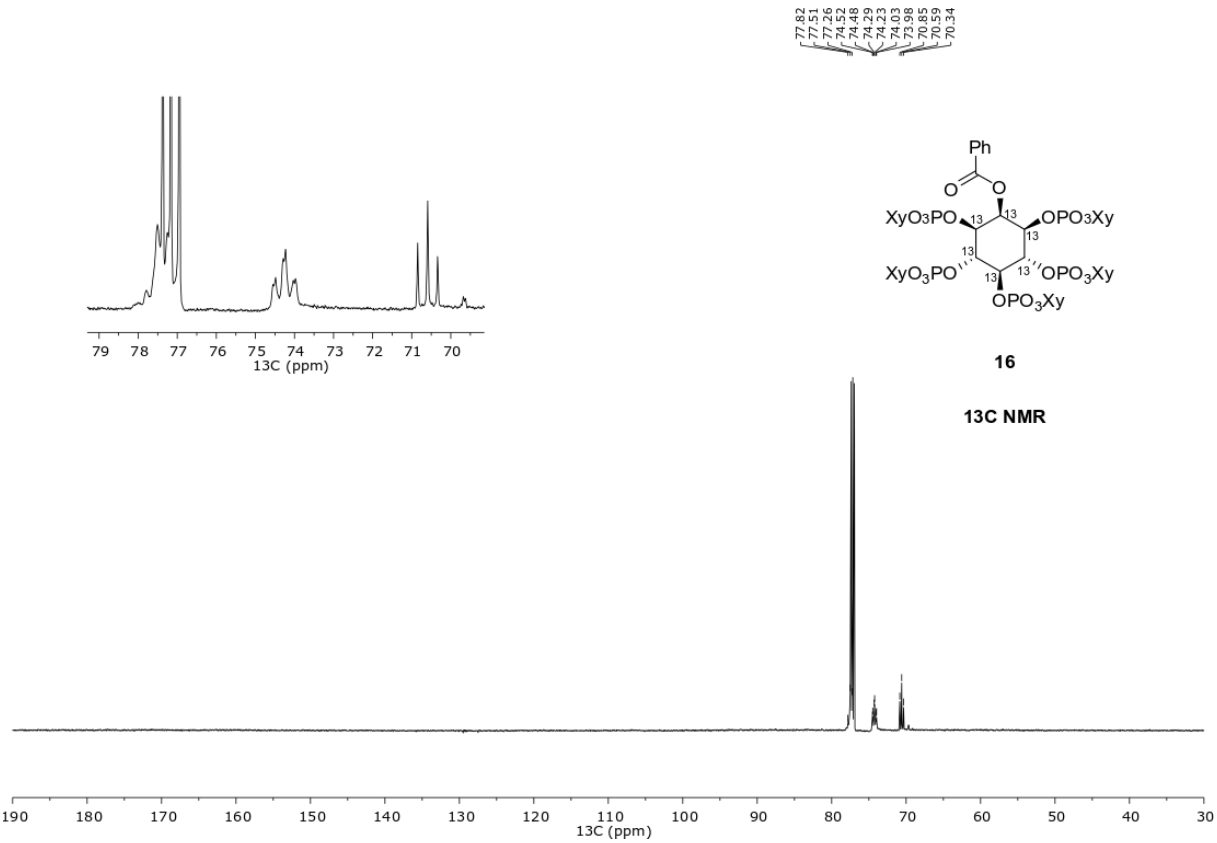

8.10
8.09
7.74
7.73
7.61
7.59
7.58
5.86
5.60
3.99
3.98
3.97
3.96
3.96
3.95
3.94
3.92
3.75
3.75
3.74
3.74
3.73
3.71
3.71
3.69
3.69
3.56
3.55
3.55
3.33
3.32
Ph
O
O
HO
13
13
OH
13
13
HO
13
13
OH
OH
S16
1H NMR
pD 7.0
1.46
0.92
1.95
0.51
0.50
2.16
2.12
0.52
0.52
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
1H (ppm)


8.15
8.13
7.80
7.79
7.78
7.65
7.64
7.63
5.78
3.91
3.89
3.88
3.87
3.85
3.50
3.48
3.47
Ph
O
O
HO
OH
HO
OH
OH
13
S16
1H NMR 13C-decoupled
pD 7.0
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5 -1.0
1H 13C-decoupled (ppm)


170.73
132.46
131.52
78.10
77.85
77.59
77.09
76.84
75.87
75.62
75.36
72.91
72.86
72.65
72.60
72.39
72.35
Ph
O
O
HO
OH
HO
OH
OH
13
S16
13C NMR
pD 7.0
79 78 77 76 75 74 73 72
13C (ppm)
190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
13C (ppm)

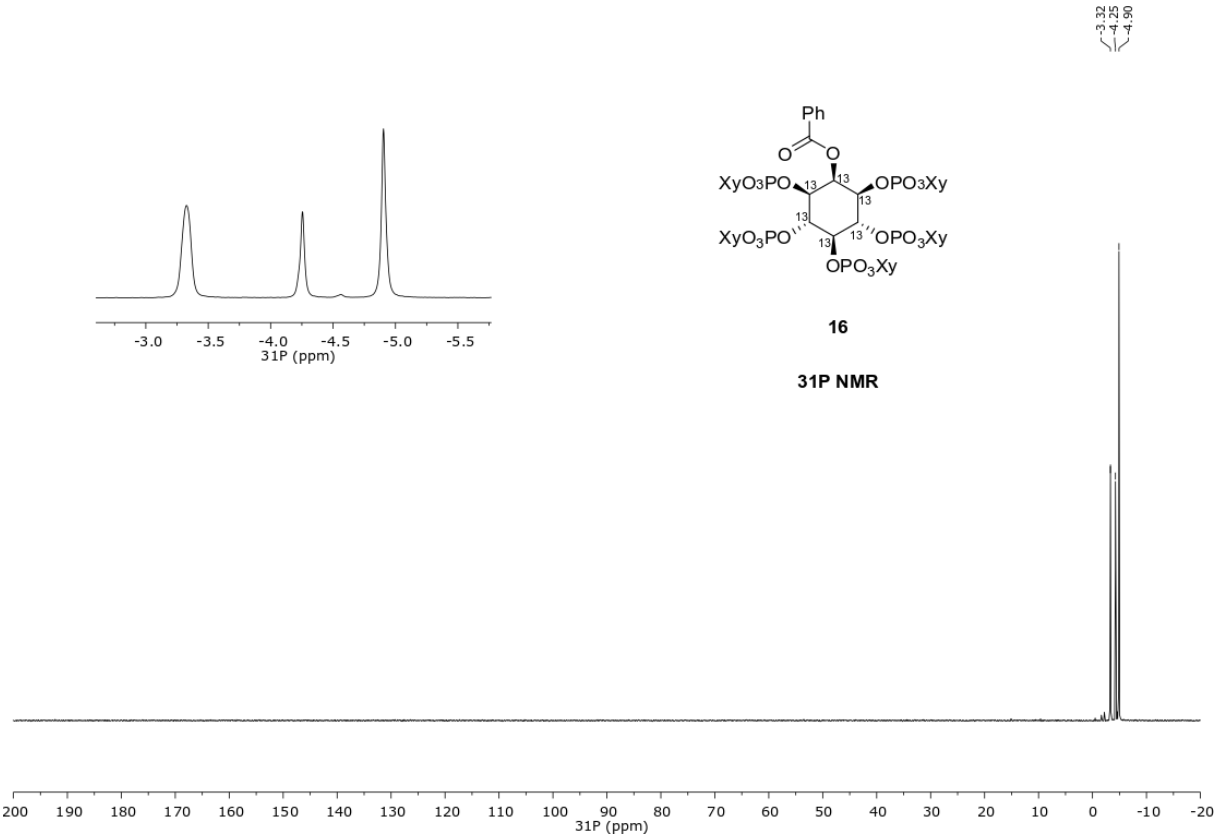

Ph
O
O
XyO3PO
OPO3Xy
XyO3PO
OPO3Xy
OPO3Xy
16
1H NMR
1H (ppm)
1H (ppm)
1H (ppm)


Ph
O
O
XyO3PO
OPO3Xy
XyO3PO
OPO3Xy
OPO3Xy
16
1H NMR 13C-decoupled
1H 13C-decoupled (ppm)
1H 13C-decoupled (ppm)
1H 13C-decoupled (ppm)



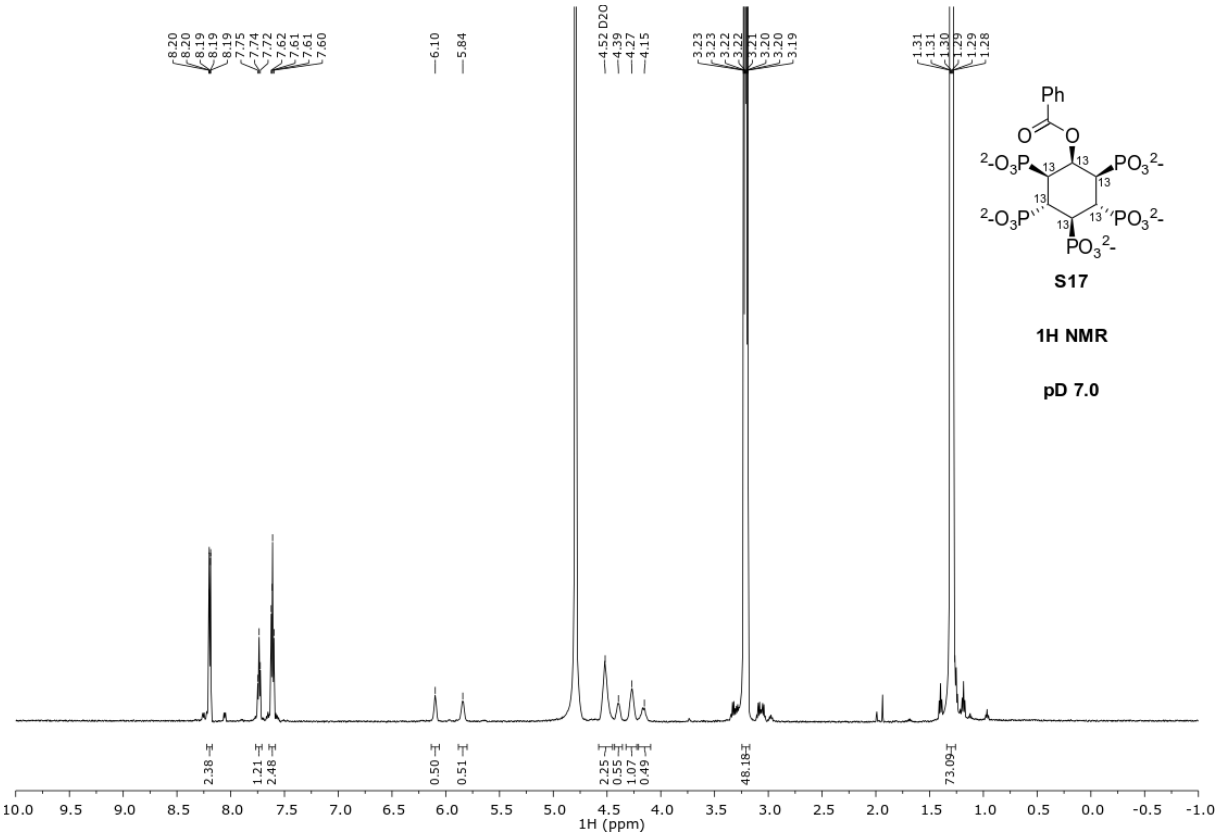
16

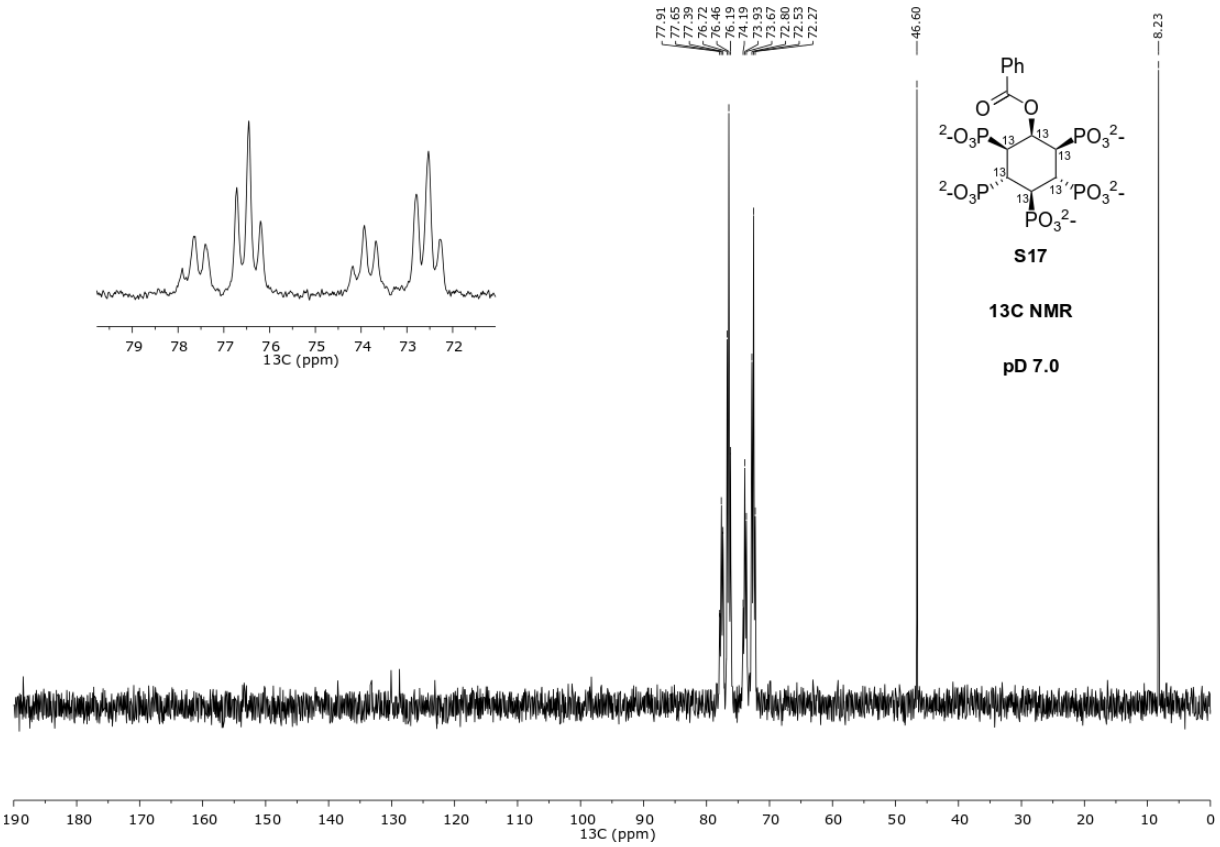
13C NMR

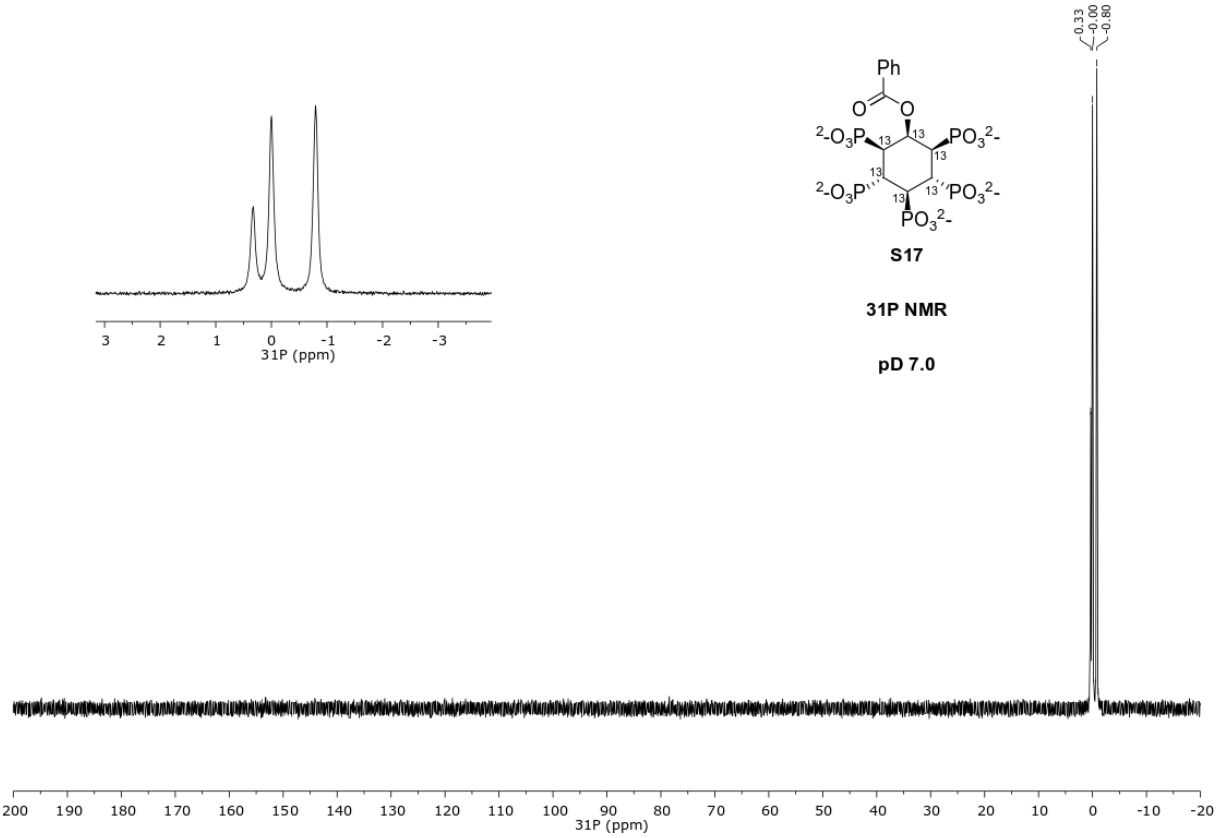


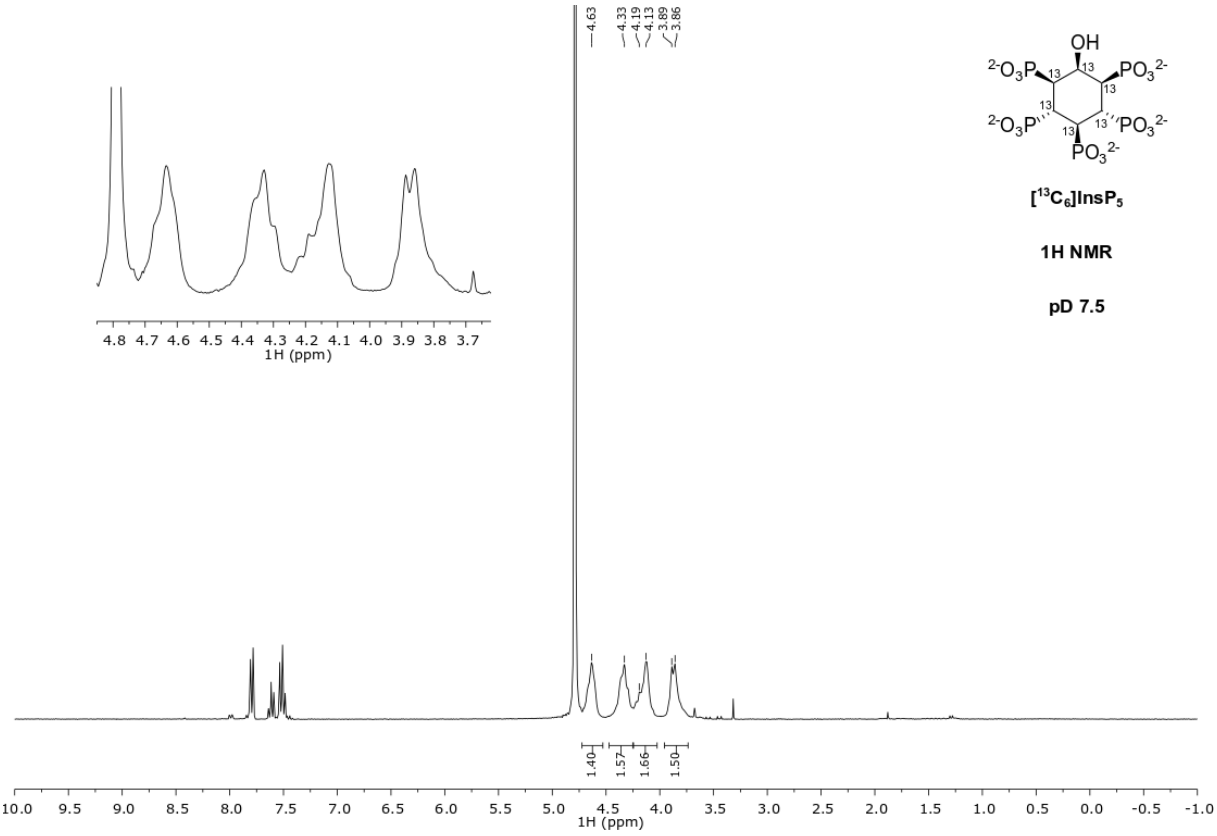
16

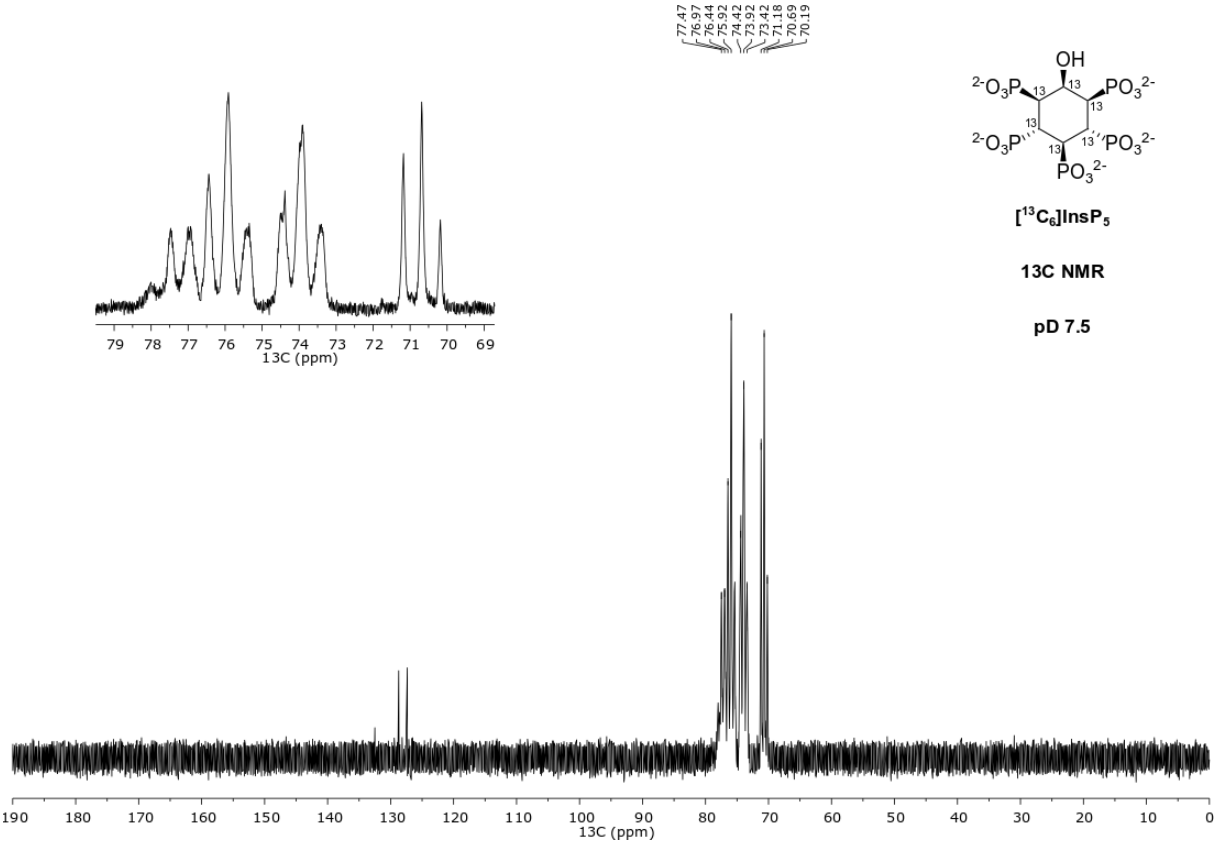
31P NMR

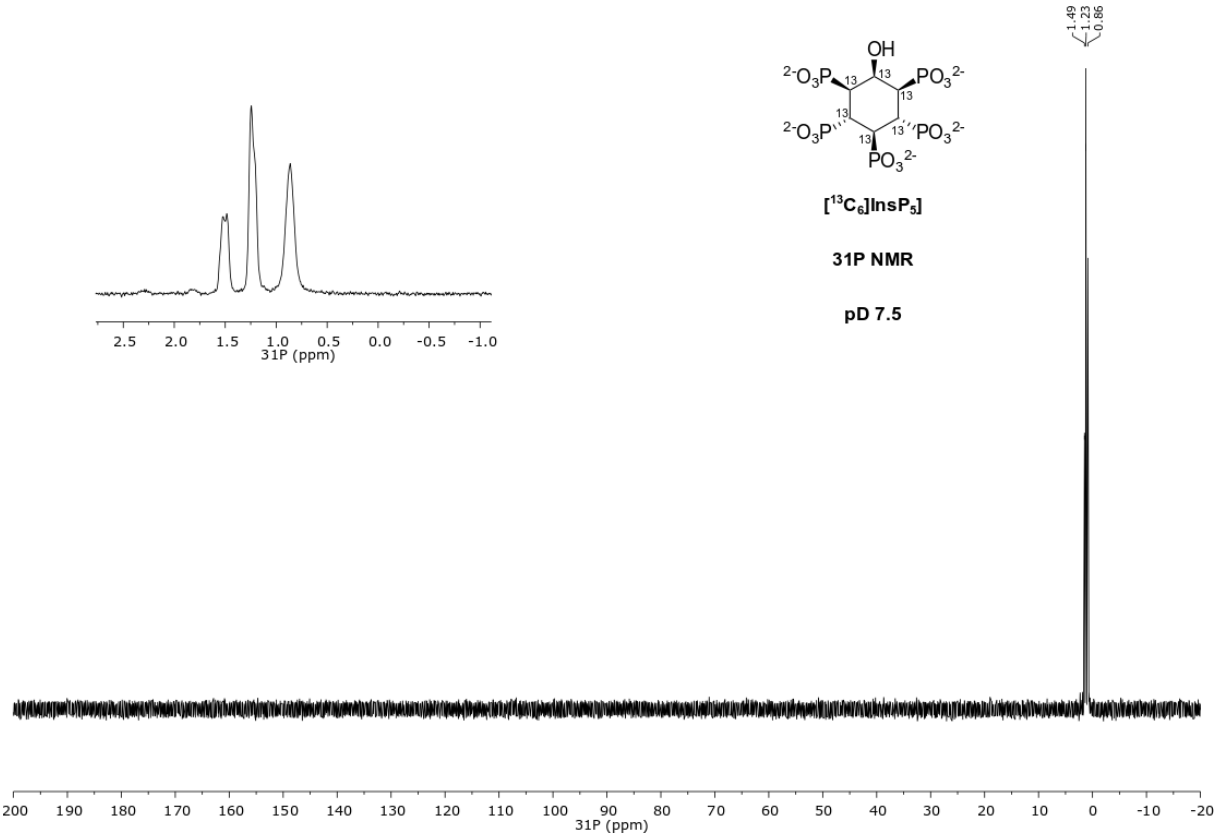

8.20
8.20
8.19
8.19
8.19
7.75
7.74
7.72
7.62
7.61
7.61
7.60
6.10
5.84
4.52 D2O
4.39
4.27
4.15
3.23
3.23
3.22
3.22
3.21
3.20
3.20
3.19
1.31
1.31
1.30
1.29
1.29
1.28
2.38
1.21
2.48
0.50
0.51
2.25
0.55
1.07
0.49
48.18
73.09
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5 -1.0
1H (ppm)
Ph
O
O
2-O3P
PO3 2-
2-O3P
PO3 2-
PO3 2-
13
S17
1H NMR
pD 7.0

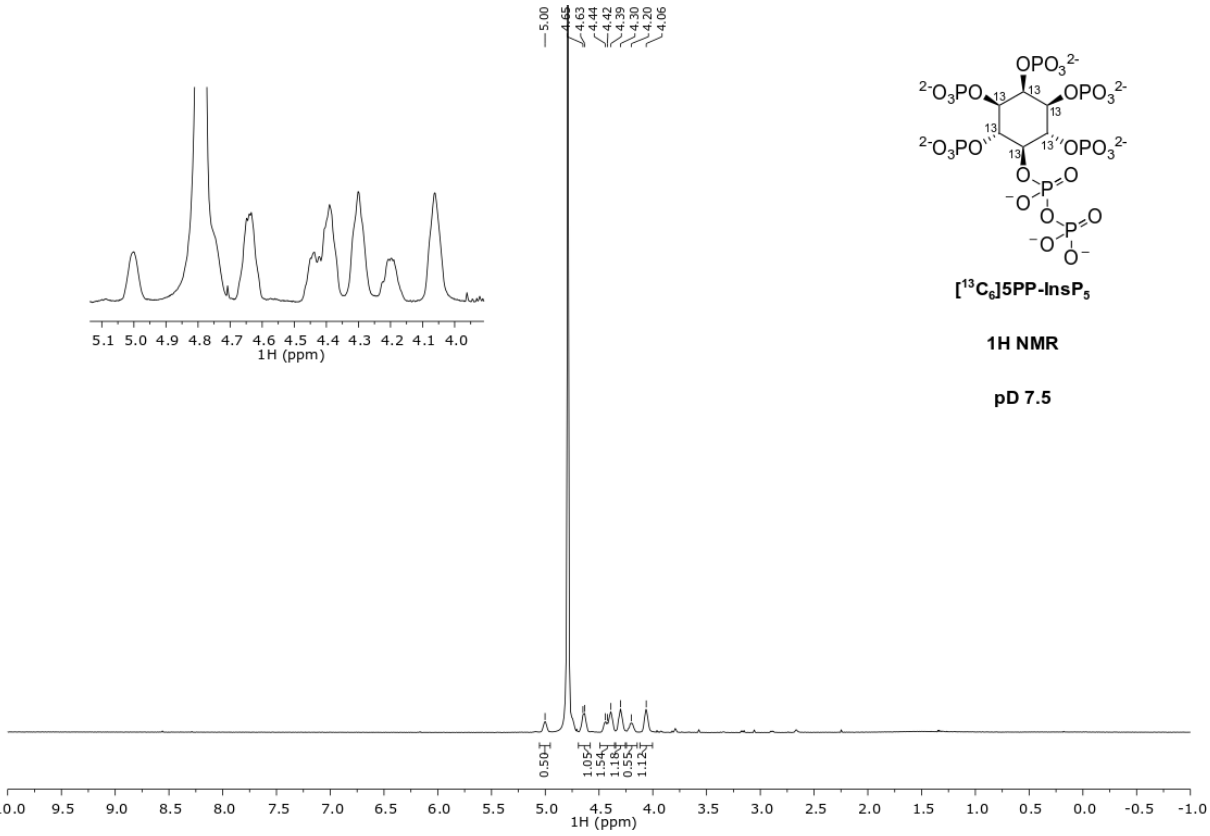

77.91
77.65
77.39
76.72
76.46
76.19
74.19
73.93
73.67
72.80
72.53
72.27
46.60
8.23
79 78 77 76 75 74 73 72
13C (ppm)
190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
13C (ppm)
Ph
O
O
2-O3P
PO3 2-
2-O3P
PO3 2-
PO3 2-
13
S17
13C NMR
pD 7.0


0.33
-0.00
-0.80
Ph
O
O
2-O3P
PO3 2-
S17
31P NMR
pD 7.0
3 2 1 0 -1 -2 -3
31P (ppm)
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10 -20
31P (ppm)


4.63
4.33
4.19
4.13
3.89
3.86
OH
2-O3P
PO3 2-
[13C6]InsP5
1H NMR
pD 7.5
4.8 4.7 4.6 4.5 4.4 4.3 4.2 4.1 4.0 3.9 3.8 3.7
1H (ppm)
1.40
1.57
1.66
1.50
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5 -1.0
1H (ppm)


77.47
76.97
76.44
75.92
74.42
73.92
73.42
71.18
70.69
70.19
OH
2-O3P
PO3 2-
PO3 2-
PO3 2-
13C (ppm)
[13C6]InsP5
13C NMR
pD 7.5

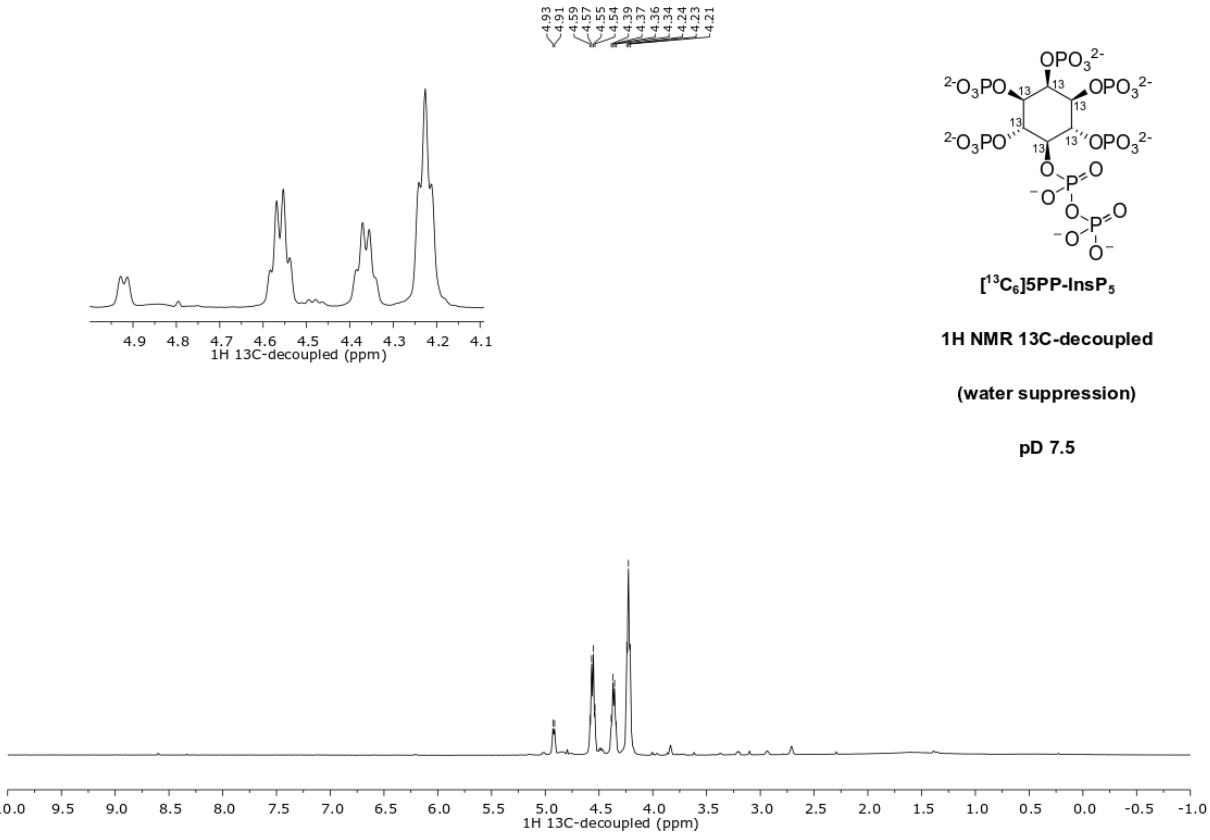

1.49
1.23
0.86
OH
2-O3P
PO3 2-
PO3 2-
PO3 2-
31P (ppm)
[13C6]InsP5]
31P NMR
pD 7.5



[$^{13}C_6$]5PP-InsP$_5$

1H NMR

pD 7.5

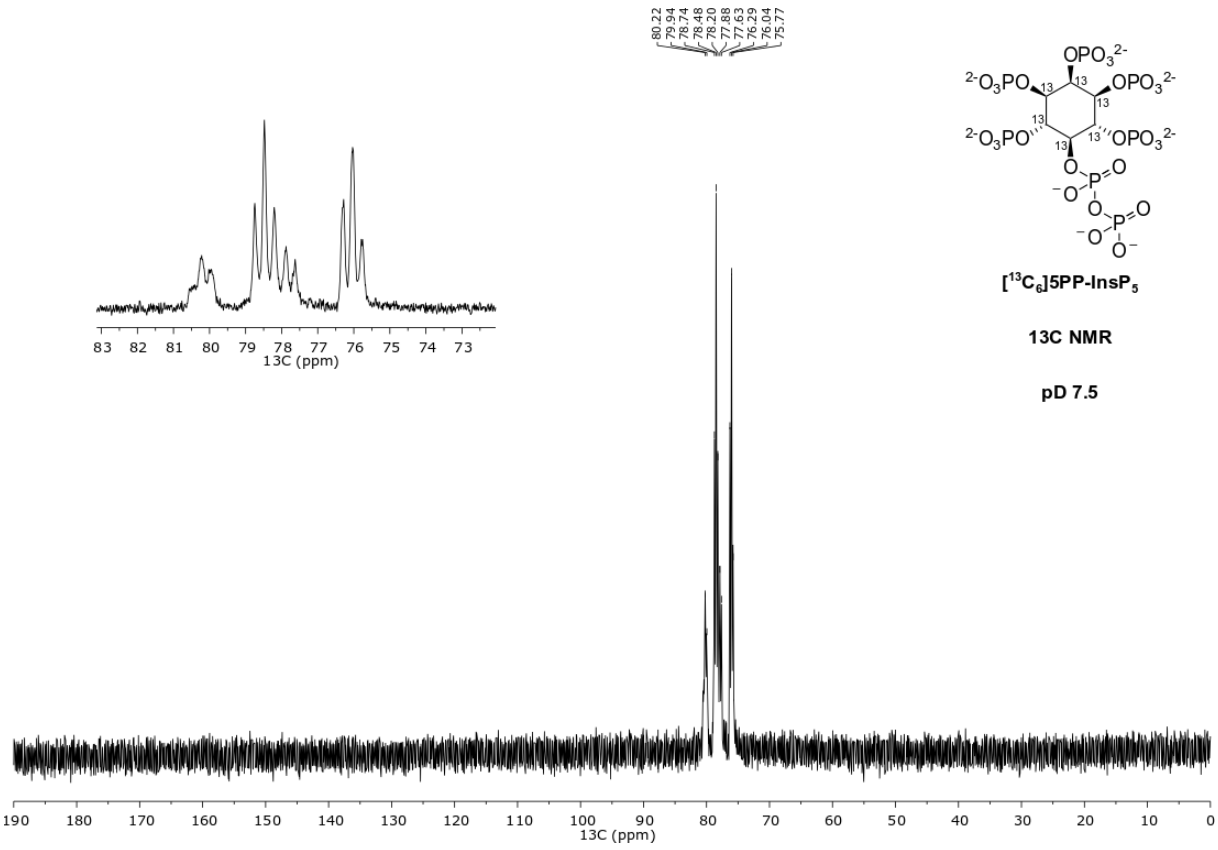


[$^{13}C_6$]5PP-InsP$_5$

1H NMR 13C-decoupled

(water suppression)

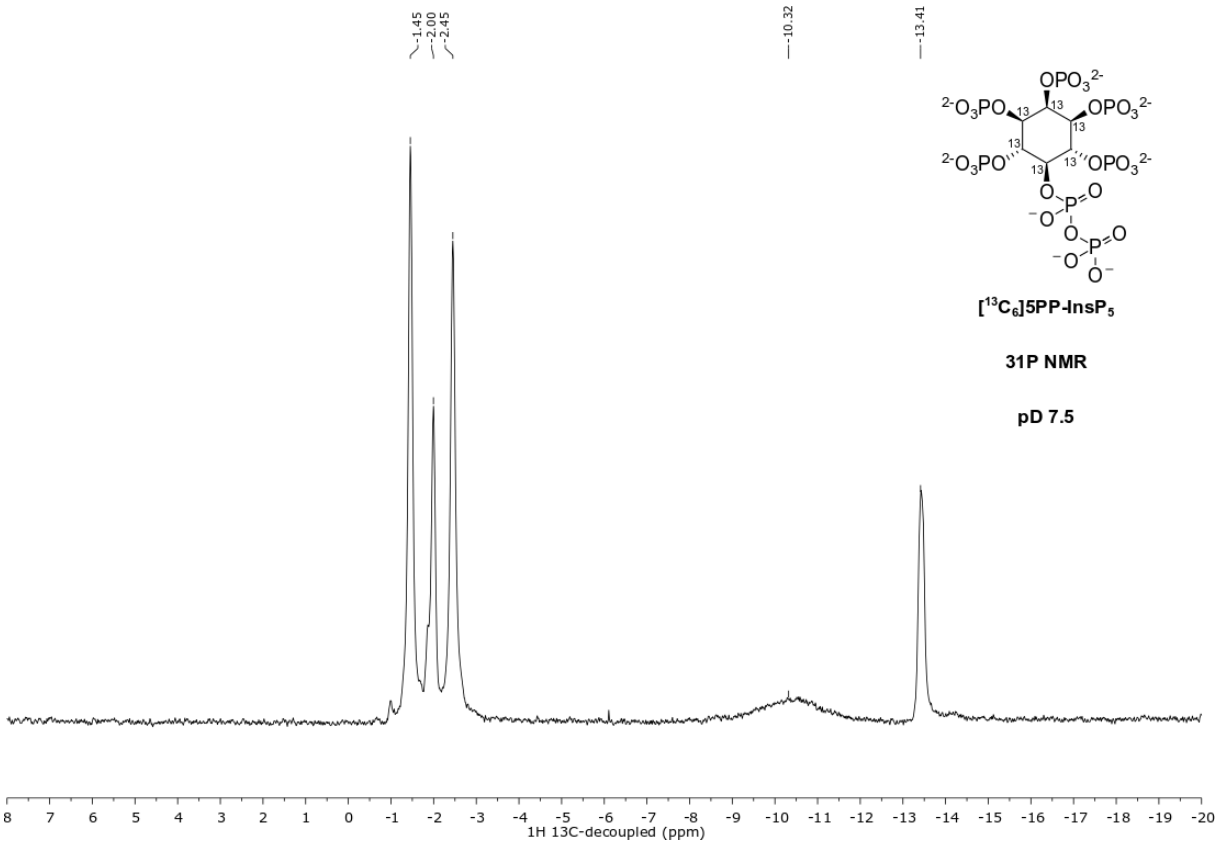
pD 7.5



[$^{13}C_6$]5PP-InsP$_5$

13C NMR

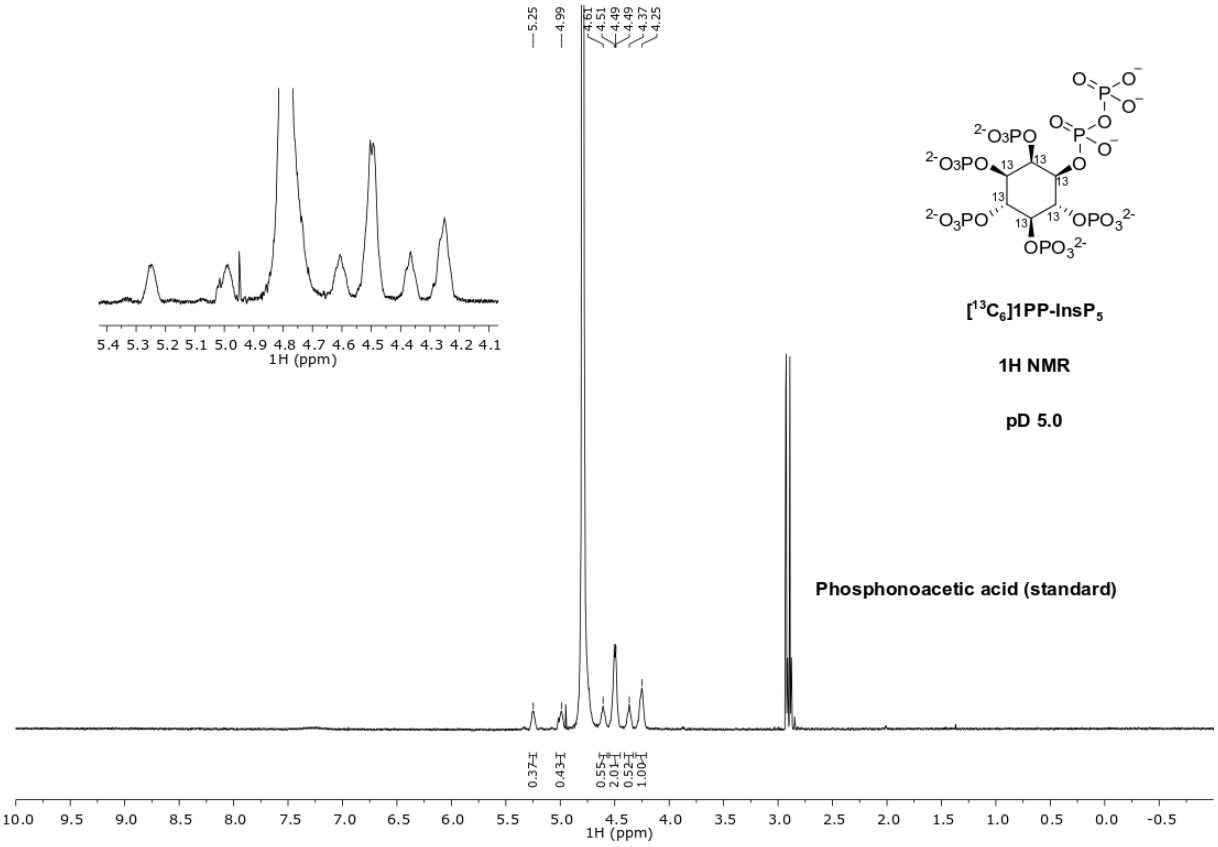
pD 7.5

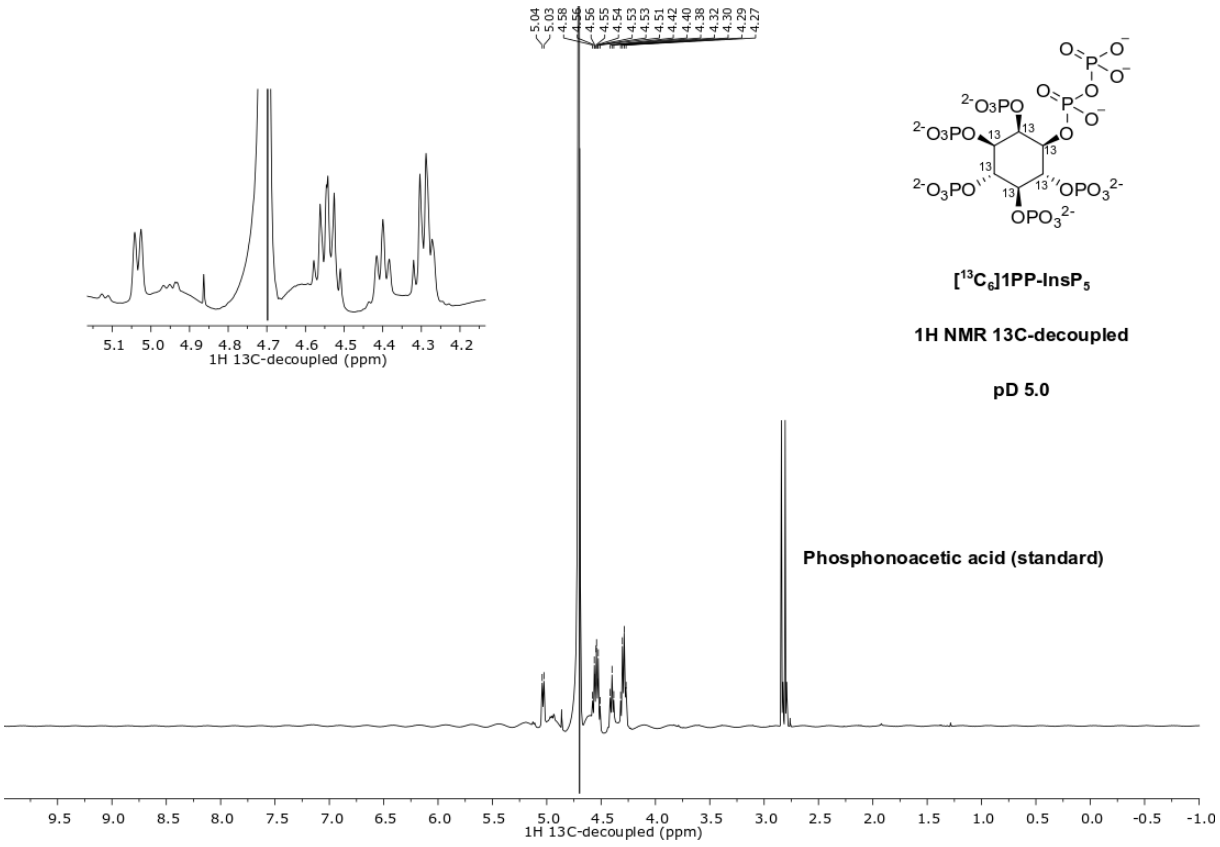


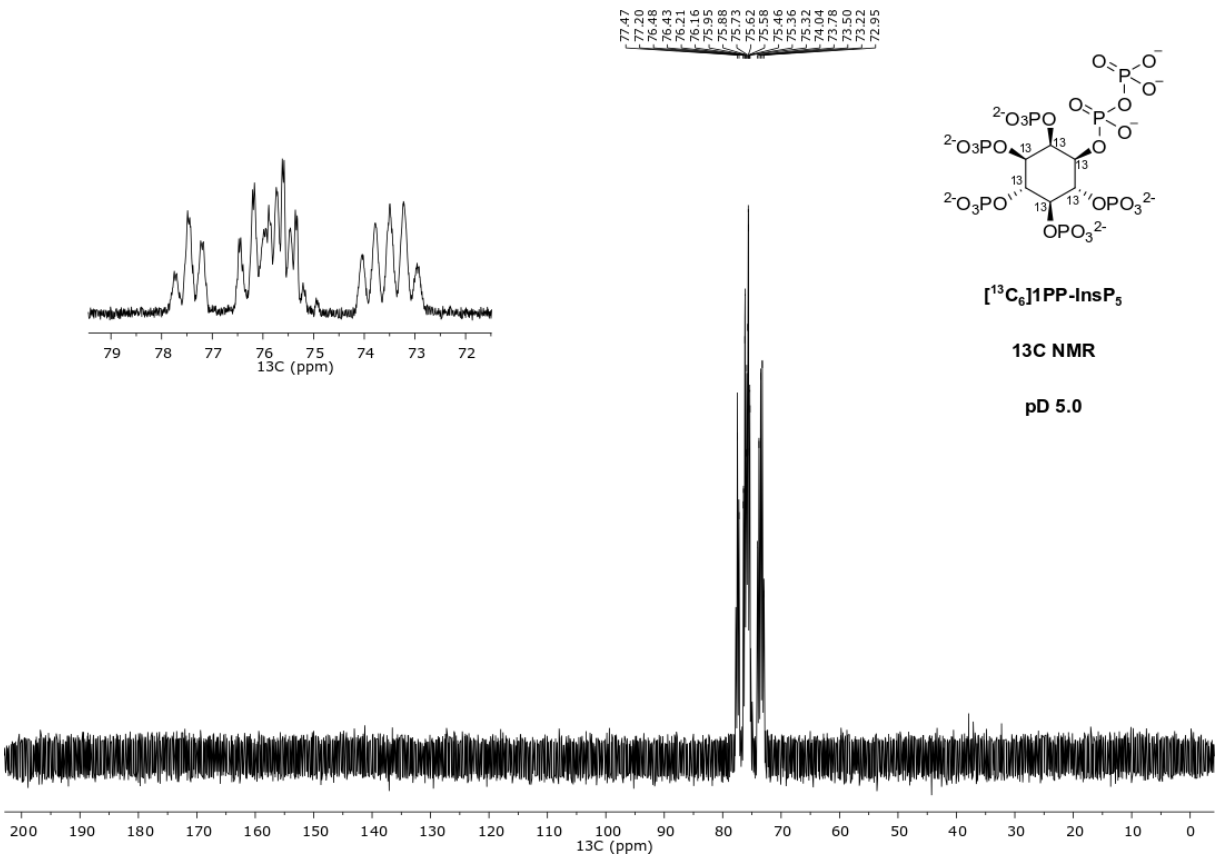
[$^{13}C_6$]5PP-InsP$_5$

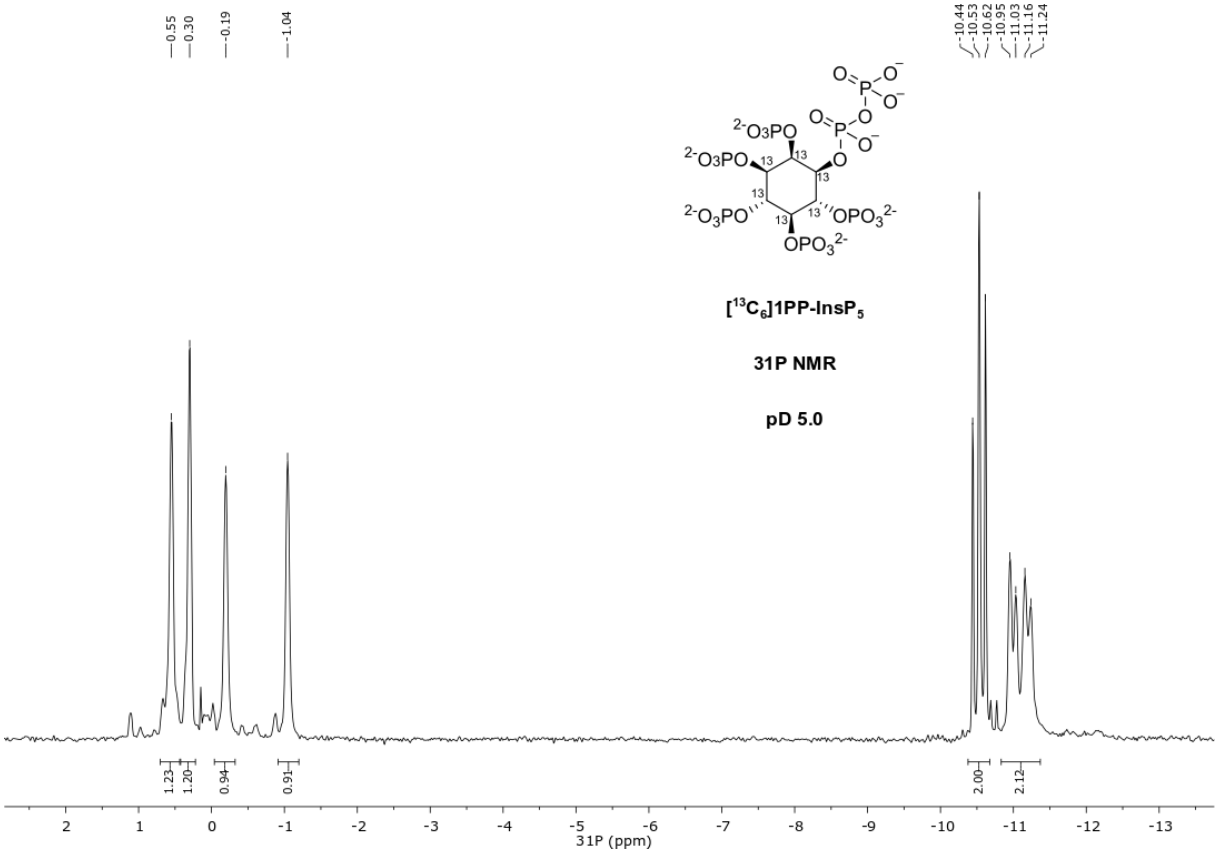
31P NMR

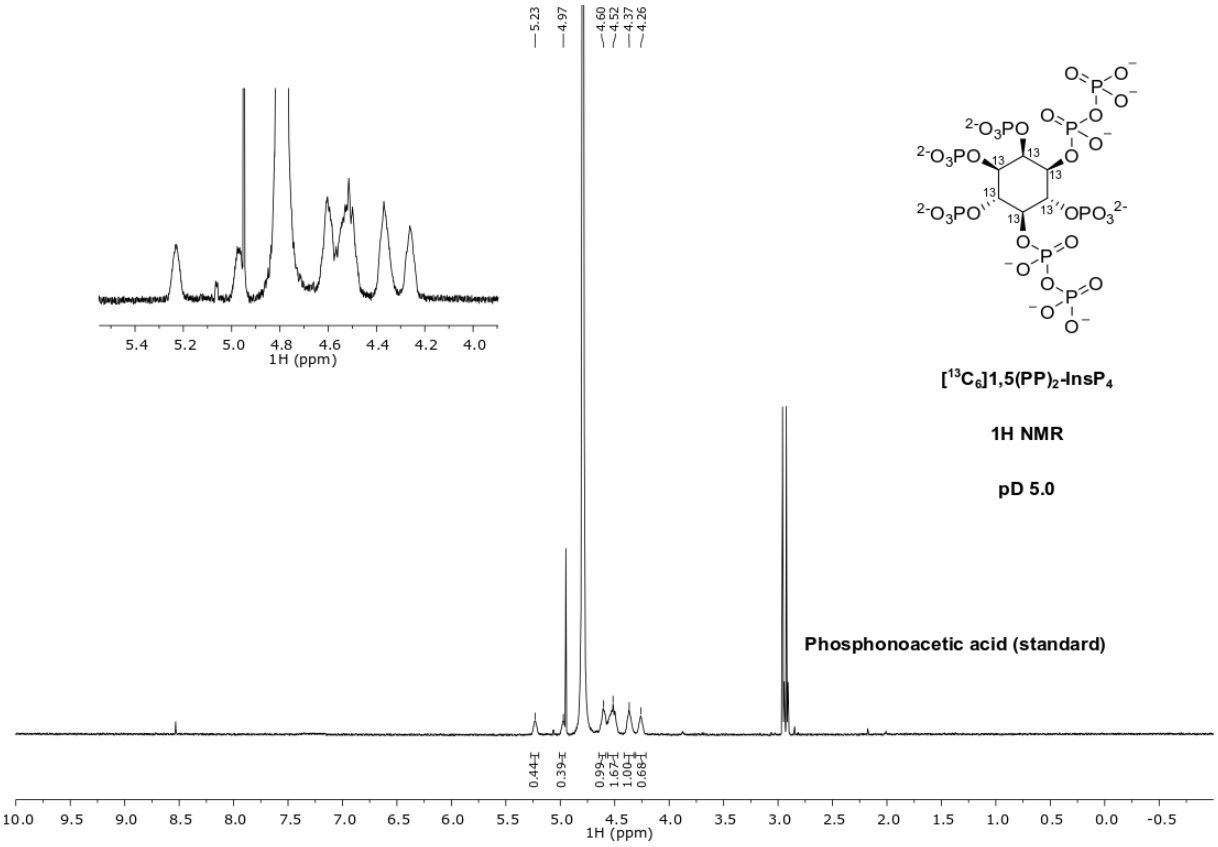
pD 7.5

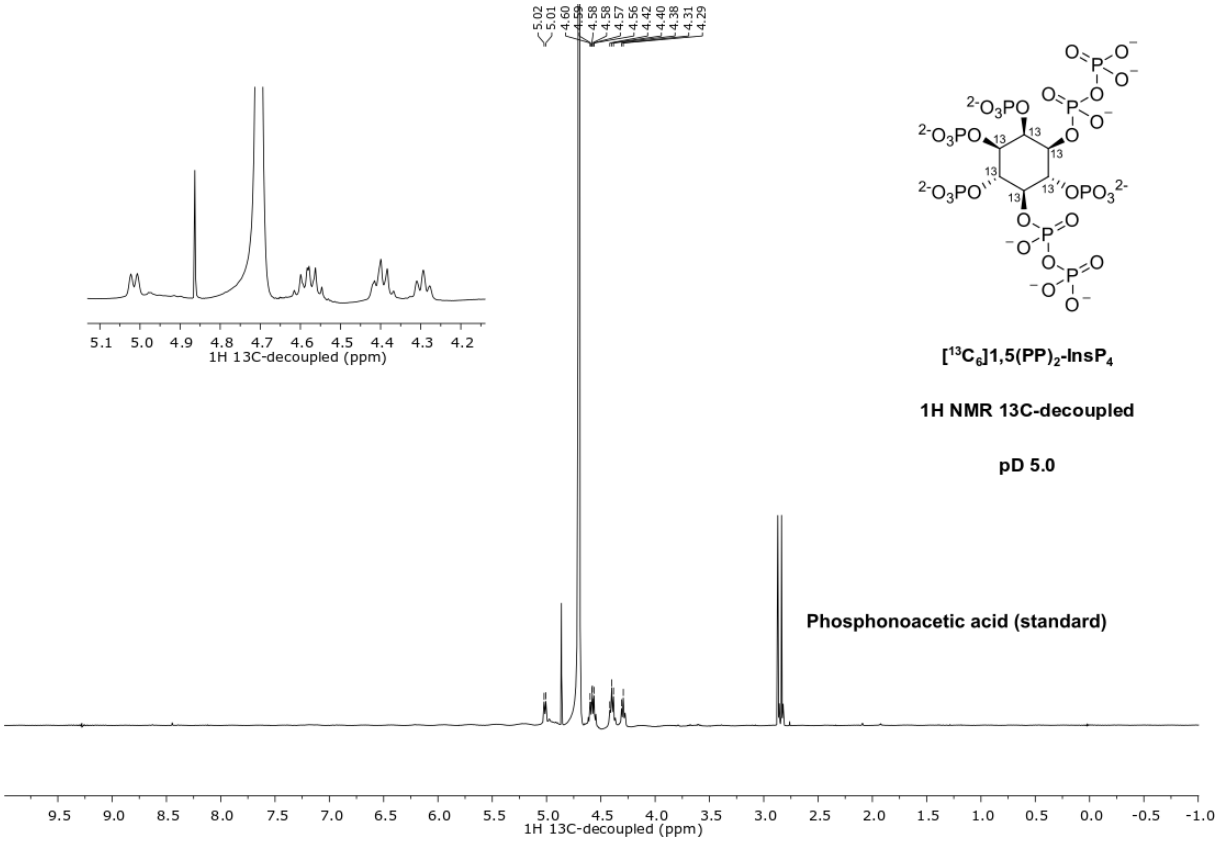

5.25
4.99
4.61
4.51
4.49
4.49
4.37
4.25
5.4 5.3 5.2 5.1 5.0 4.9 4.8 4.7 4.6 4.5 4.4 4.3 4.2 4.1
1H (ppm)
[13C6]1PP-InsP5
1H NMR
pD 5.0
Phosphonoacetic acid (standard)
0.37
0.43
0.55
2.01
0.52
1.00
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
1H (ppm)

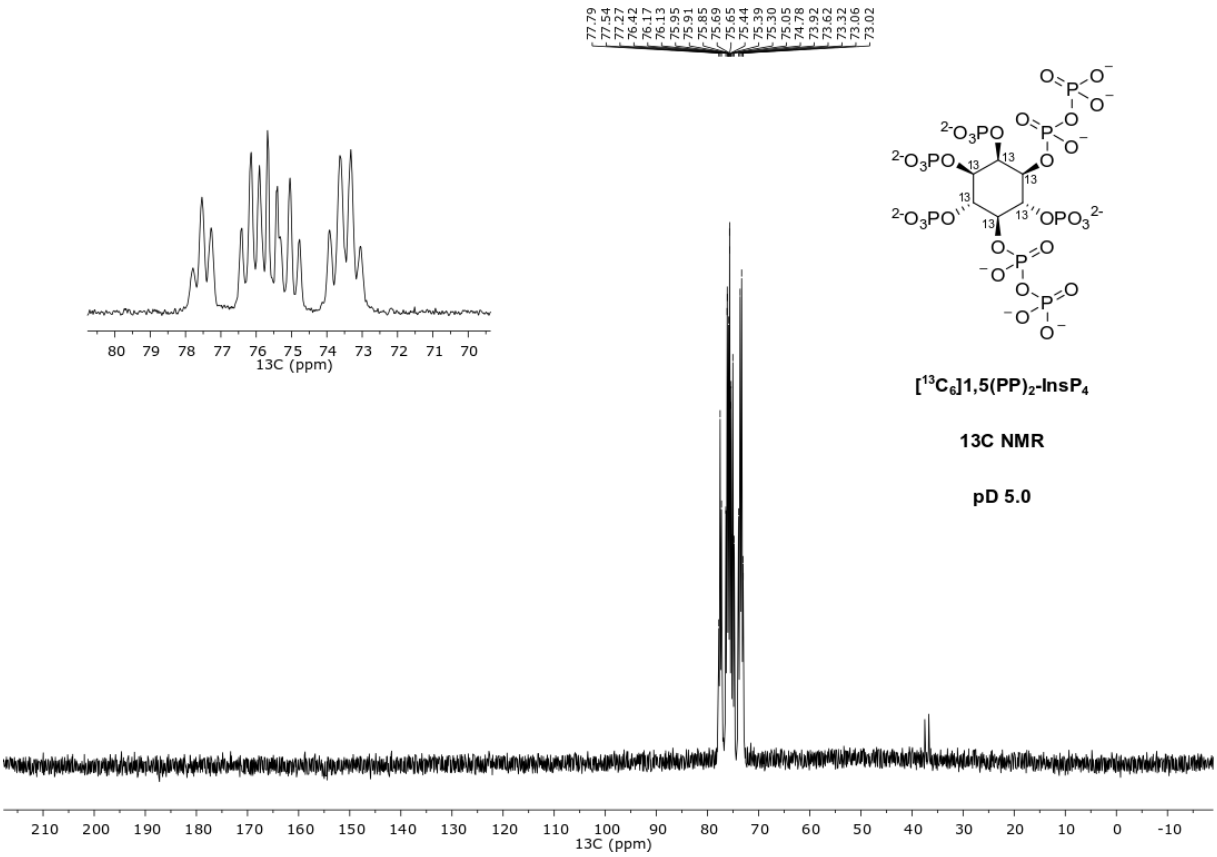

5.04
5.03
4.58
4.56
4.56
4.55
4.54
4.53
4.53
4.51
4.42
4.40
4.38
4.32
4.30
4.29
4.27
5.1 5.0 4.9 4.8 4.7 4.6 4.5 4.4 4.3 4.2
1H 13C-decoupled (ppm)
[13C6]1PP-InsP5
1H NMR 13C-decoupled
pD 5.0
Phosphonoacetic acid (standard)
9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5 -1.0
1H 13C-decoupled (ppm)

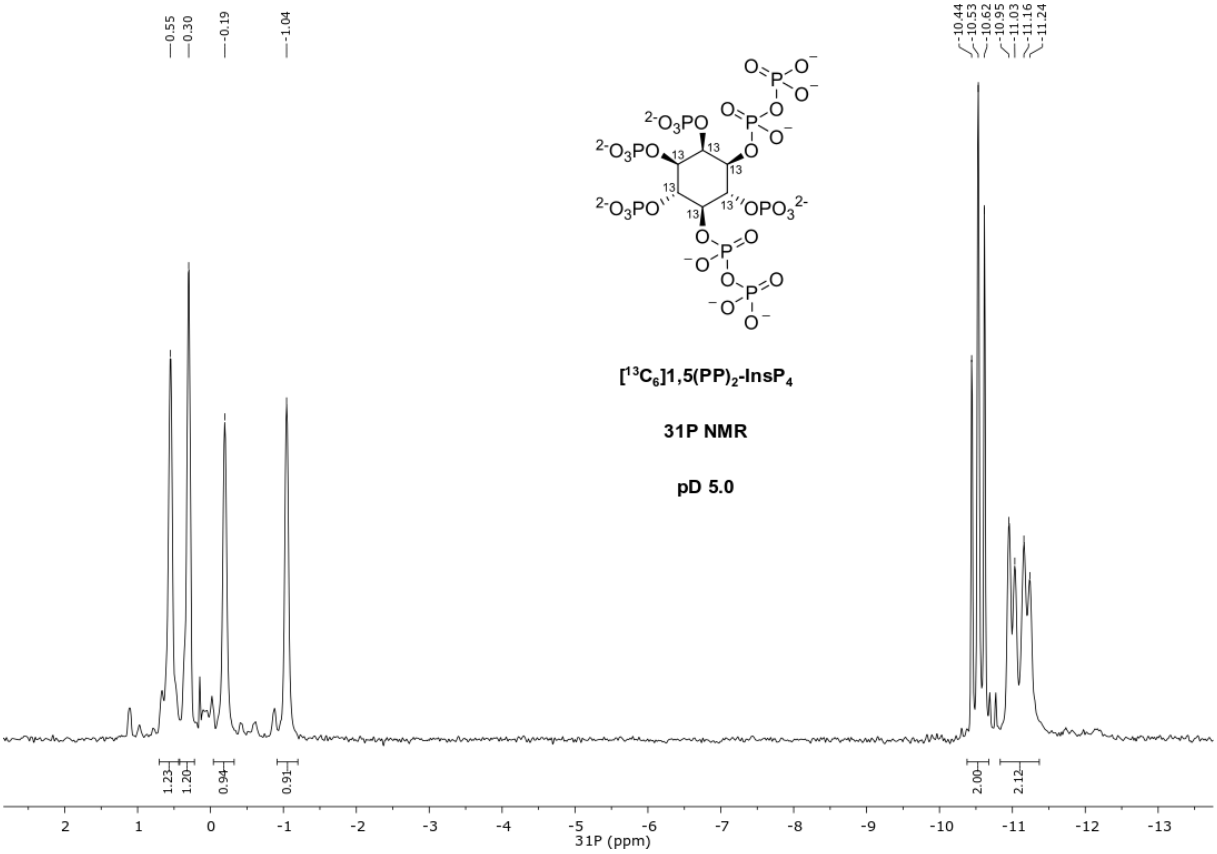

77.47
77.20
76.48
76.43
76.21
76.16
75.95
75.88
75.73
75.62
75.58
75.46
75.36
75.32
74.04
73.78
73.50
73.22
72.95
79 78 77 76 75 74 73 72
13C (ppm)
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
13C (ppm)
[13C6]1PP-InsP5
13C NMR
pD 5.0


0.55
0.30
-0.19
-1.04
10.44
10.53
10.62
10.95
11.03
11.16
11.24
1.23
1.20
0.94
0.91
2.00
2.12
2 1 0 -1 -2 -3 -4 -5 -6 -7 -8 -9 -10 -11 -12 -13
31P (ppm)
[13C6]1PP-InsP5
31P NMR
pD 5.0


5.23
4.97
4.60
4.52
4.37
4.26
1H (ppm)
[13C6]1,5(PP)2-InsP4
1H NMR
pD 5.0
Phosphonoacetic acid (standard)
0.44
0.39
0.99
1.67
1.01
0.68
1H (ppm)


5.02
5.01
4.60
4.58
4.58
4.57
4.56
4.42
4.40
4.38
4.31
4.29
1H 13C-decoupled (ppm)
[13C6]1,5(PP)2-InsP4
1H NMR 13C-decoupled
pD 5.0
Phosphonoacetic acid (standard)
1H 13C-decoupled (ppm)


77.79
77.54
77.27
76.42
76.17
76.13
75.95
75.91
75.85
75.69
75.65
75.44
75.39
75.30
75.05
74.78
73.92
73.62
73.32
73.06
73.02
80 79 78 77 76 75 74 73 72 71 70
13C (ppm)
[13C6]1,5(PP)2-InsP4
13C NMR
pD 5.0
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
13C (ppm)


0.55
0.30
-0.19
-1.04
-10.44
-10.53
-10.62
-10.95
-11.03
-11.16
-11.24
[13C6]1,5(PP)2-InsP4
31P NMR
pD 5.0
1.23
1.20
0.94
0.91
2.00
2.12
2 1 0 -1 -2 -3 -4 -5 -6 -7 -8 -9 -10 -11 -12 -13
31P (ppm)

References

1. Menniti, F. S., Miller, R. N., Putney Jr., J. W. & Shears, S. B. Turnover of inositol polyphosphate pyrophosphates in pancreatoma cells. *J Biol Chem* **268**, 3850–3856 (1993).
2. Stephens, L. *et al.* The Detection, Purification, Structural Characterization, andMetabolism of DiphosphoinositoPl entakisphosphate(s) andBisdiphosphoinositol Tetrakisphosphate(s. *J. Biol. Chem.* **268**, 4009–4015 (1993).
3. Banfic, H., Bedalov, A., York, J. D. & Visnjic, D. Inositol pyrophosphates modulate S phase progression after pheromone-induced arrest in saccharomyces cerevisiae. *J. Biol. Chem.* **288**, 1717–1725 (2013).
4. Gu, C. *et al.* The Significance of the Bifunctional Kinase/Phosphatase Activities of Diphosphoinositol Pentakisphosphate Kinases (PPIP5Ks) for Coupling Inositol Pyrophosphate Cell Signaling to Cellular Phosphate Homeostasis. *J. Biol. Chem.* **292**, 4544–4555 (2017).
5. Mayr, G. W. A novel metal-dye detection system permits picomolar-range h.p.l.c. analysis of inositol polyphosphates from non-radioactively labelled cell or tissue specimens. *Biochem. J.* **254**, 585–591 (1988).
6. Wilson, M. S. C., Bulley, S. J., Pisani, F., Irvine, R. F. & Saiardi, A. A novel method for the purification of inositol phosphates from biological samples reveals that no phytate is present in human plasma or urine. *Open Biol.* **5**, 150014 (2015).
7. Laussmann, T., Eujen, R., Weisshuhn, C. M. & Thiel, U. Structures of diphospho-myo-inositol pentakisphosphate and bisdiphospho- myo-inositol tetrakisphosphate from Dictyostelium resolved by NMR analysis. *Biochem. J.* **315**, 715–720 (1996).
8. Laussmann, T., Reddy, K. M., Reddy, K. K., Falck, J. R. & Vogel, G. Diphospho-myo-inositol phosphates from Dictyostelium identified as D-6-diphospho-myo-inositol pentakisphosphate and D-5,6-bisdiphospho-myo-inositol tetrakisphosphate. *Biochem. J.* **322**, 31–33 (1997).
9. Ijare, O. B., Baskin, D. S., Sharpe, M. A. & Pichumani, K. Metabolism of fructose in B-cells: A13C NMR spectroscopy based stable isotope tracer study. *Anal. Biochem.* **552**, 110–117 (2018).
10. Keshari, K. R. *et al.* Hyperpolarized [2-13C]-fructose: A hemiketal DNP substrate for in vivo metabolic imaging. *J. Am. Chem. Soc.* **131**, 17591–17596 (2009).
11. Park, J. M. *et al.* Hyperpolarized Sodium [1-13C]-Glycerate as a Probe for Assessing Glycolysis in Vivo. *J. Am. Chem. Soc.* **139**, 6629–6634 (2017).
12. Saiardi, A., Guillermier, C., Loss, O., Poczatek, J. C. & Lechene, C. Quantitative imaging of inositol distribution in yeast using multi-isotope imaging mass spectrometry (MIMS). *Surf. Interface Anal.* **46**, 169–172 (2014).
13. Sahai, P., Chawla, M. & Vishwakarma, R. A. 13C Labelling and electrospray mass spectrometry reveal a de novo route for inositol biosynthesis in Leishmania donovani parasite. *J. Chem. Soc. Perkin Trans. 1* **0**, 1283–1290 (2000).
14. Moure, M. J., Zhuo, Y., Boons, G. J. & Prestegard, J. H. Perdeuterated and 13 C-enriched myo-inositol for DNP assisted monitoring of enzymatic phosphorylation by inositol-3-kinase. *Chem. Commun.* **53**, 12398–12401 (2017).
15. Morin, L. G. Creatine kinase: re-examination of optimum reaction conditions. *Clin. Chem.* **23**, 1569–1575 (1977).
16. Barker, C. J., Wright, J., Hughes, P. J., Kirk, C. J. & Michell, R. H. Complex changes in cellular inositol phosphate complement accompany transit through the cell cycle. *Biochem. J.* **380**, 465–473 (2004).
17. Podeschwa, M. *et al.* Stereoselective Synthesis of myo-, neo-, L-chiro, D-chiro, allo-, scyllo-, and epi-Inositol Systems via Conduritols Prepared from p-Benzoquinone. *European J. Org. Chem.* **2003**, 1958–1972 (2003).
18. Blüher, D. *et al.* A 1-phytase type III effector interferes with plant hormone signaling. *Nat. Commun.* **8**, 1–14 (2017).
19. Godage, H. Y. *et al.* Regioselective opening of myo-inositol orthoesters: Mechanism and

synthetic utility. *J. Org. Chem.* **78**, 2275–2288 (2013).

20. Capolicchio, S., Thakor, D. T., Linden, A. & Jessen, H. J. Synthesis of unsymmetric diphospho-inositol polyphosphates. *Angew. Chem. Int. Ed. Engl.* **52**, 6912–6 (2013).
21. Rebekka Wild *et al.* Control of eukaryotic phosphate homeostasis by inositol polyphosphate sensor domains. *Science* **353**, 986–990 (2016).
22. Pavlovic, I. *et al.* Prometabolites of 5-Diphospho-myo-inositol Pentakisphosphate. *Angew. Chemie Int. Ed.* **54**, 9622–9626 (2015).
23. Wormald, M., Liao, G., Kimos, M., Barrow, J. & Wei, H. Development of a homogenous high-throughput assay for inositol hexakisphosphate kinase 1 activity. *PLoS One* **12**, e0188852 (2017).
24. Wundenberg, T., Grabinski, N., Lin, H. & Mayr, G. W. Discovery of InsP $_6$ -kinases as InsP $_6$ -dephosphorylating enzymes provides a new mechanism of cytosolic InsP $_6$ degradation driven by the cellular ATP/ADP ratio. *Biochem. J.* **462**, 173–184 (2014).
25. Voglmaier, S. M. *et al.* Purified inositol hexakisphosphate kinase is an ATP synthase: diphosphoinositol pentakisphosphate as a high-energy phosphate donor. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 4305–10 (1996).
26. Padmanabhan, U., Dollins, D. E., Fridy, P. C., York, J. D. & Downes, C. P. Characterization of a selective inhibitor of inositol hexakisphosphate kinases. Use in defining biological roles and metabolic relationships of inositol pyrophosphates. *J. Biol. Chem.* **284**, 10571–10582 (2009).
27. Saiardi, A., Azevedo, C., Desfougères, Y., Portela-Torres, P. & Wilson, M. S. C. Microbial inositol polyphosphate metabolic pathway as drug development target. *Adv. Biol. Regul.* **67**, 74–83 (2018).
28. Jung, S. K. *et al.* Myricetin Suppresses UVB-Induced Skin Cancer by Targeting Fyn. *Cancer Res.* **68**, 6021 LP – 6029 (2008).
29. Kumamoto, T., Fujii, M. & Hou, D.-X. Myricetin directly targets JAK1 to inhibit cell transformation. *Cancer Lett.* **275**, 17–26 (2009).
30. Lee, K. W. *et al.* Myricetin is a novel natural inhibitor of neoplastic cell transformation and MEK1. *Carcinogenesis* **28**, 1918–1927 (2007).
31. Voglmaier, S. M. *et al.* Purified inositol hexakisphosphate kinase is an ATP synthase: diphosphoinositol pentakisphosphate as a high-energy phosphate donor. *Proc Natl Acad Sci U S A* **93**, 4305–4310 (1996).
32. Desai, M. *et al.* Two inositol hexakisphosphate kinases drive inositol pyrophosphate synthesis in plants. *Plant J.* **80**, 642–653 (2014).
33. Laha, D. *et al.* VIH2 Regulates the Synthesis of Inositol Pyrophosphate InsP8 and Jasmonate-Dependent Defenses in Arabidopsis. *Plant Cell* **27**, 1082–1097 (2015).
34. Wilson, M. S., Jessen, H. J. & Saiardi, A. The inositol hexakisphosphate kinases IP6K1 and -2 regulate human cellular phosphate homeostasis, including XPR1-mediated phosphate export. *J. Biol. Chem.* **294**, 11597–11608 (2019).
35. Azevedo, C. & Saiardi, A. Eukaryotic Phosphate Homeostasis: The Inositol Pyrophosphate Perspective. *Trends Biochem. Sci.* **42**, 219–231 (2017).
36. Hürlimann, H. C., Pinson, B., Stadler-Waibel, M., Zeeman, S. C. & Freimoser, F. M. The SPX domain of the yeast low-affinity phosphate transporter Pho90 regulates transport activity. *EMBO Rep.* **10**, 1003–1008 (2009).
37. Hamburger, D., Rezzonico, E., MacDonald-Comber Petétot, J., Somerville, C. & Poirier, Y. Identification and characterization of the Arabidopsis PHO1 gene involved in phosphate loading to the xylem. *Plant Cell* **14**, 889–902 (2002).
38. Liu, J. *et al.* A vacuolar phosphate transporter essential for phosphate homeostasis in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E6571–E6578 (2015).
39. Giovannini, D., Touhami, J., Charnet, P., Sitbon, M. & Battini, J.-L. Inorganic Phosphate Export by the Retrovirus Receptor XPR1 in Metazoans. *Cell Rep.* **3**, 1866–1873 (2013).
40. Hothorn, M. *et al.* Catalytic Core of a Membrane-Associated Eukaryotic Polyphosphate Polymerase. *Science (80-.).* **324**, 513 LP – 516 (2009).

41. Ogawa, N. *et al.* Functional domains of Pho81p, an inhibitor of Pho85p protein kinase, in the transduction pathway of Pi signals in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **15**, 997–1004 (1995).

42. Park, B. S., Seo, J. S. & Chua, N.-H. NITROGEN LIMITATION ADAPTATION Recruits PHOSPHATE2 to Target the Phosphate Transporter PT2 for Degradation during the Regulation of Arabidopsis Phosphate Homeostasis. *Plant Cell* **26**, 454 LP – 464 (2014).

43. Duan, K. *et al.* Characterization of a sub-family of Arabidopsis genes with the SPX domain reveals their diverse functions in plant tolerance to phosphorus starvation. *Plant J.* **54**, 965–975 (2008).

44. Rubio, V. *et al.* A conserved MYB transcription factor involved in phosphate starvation signaling both in vascular plants and in unicellular algae. *Genes Dev.* **15**, 2122–2133 (2001).

45. Puga, M. I. *et al.* SPX1 is a phosphate-dependent inhibitor of Phosphate Starvation Response 1 in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 14947–14952 (2014).

46. Wang, Z. *et al.* Rice SPX1 and SPX2 inhibit phosphate starvation responses through interacting with PHR2 in a phosphate-dependent manner. *Proc. Natl. Acad. Sci.* **111**, 14953 LP – 14958 (2014).

47. Albert, C. *et al.* Biological variability in the structures of diphosphoinositol polyphosphates in Dictyostelium discoideum and mammalian cells. **560**, 553–560 (1997).

48. Lin, H. *et al.* Structural analysis and detection of biological inositol pyrophosphates reveal that the family of VIP/Diphosphoinositol pentakisphosphate kinases Are1/3-kinases. *J. Biol. Chem.* **284**, 1863–1872 (2009).

49. Gu, C. *et al.* KO of 5-InsP 7 kinase activity transforms the HCT116 colon cancer cell line into a hypermetabolic, growth-inhibited phenotype. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 11968–11973 (2017).

50. Wilson, M. & Saiardi, A. Inositol Phosphates Purification Using Titanium Dioxide Beads. *BIO-PROTOCOL* **8**, e2959 (2018).

51. Arner, R. J. *et al.* myo-Inositol oxygenase: molecular cloning and expression of a unique enzyme that oxidizes myo-inositol and D-chiro-inositol. *Biochem. J.* **360**, 313–20 (2001).

52. Thomas, M. P., Mills, S. J. & Potter, B. V. L. The 'Other' Inositols and Their Phosphates: Synthesis, Biology, and Medicine (with Recent Advances in myo-Inositol Chemistry). *Angew. Chemie - Int. Ed.* **55**, 1614–1650 (2016).

53. Bozzaro, S. The Model Organism Dictyostelium discoideum BT - Dictyostelium discoideum Protocols. in (eds. Eichinger, L. & Rivero, F.) 17–37 (Humana Press, 2013). doi:10.1007/978-1-62703-302-2_2

54. Annesley, S. J. & Fisher, P. R. Dictyostelium discoideum—a model for many reasons. *Mol. Cell. Biochem.* **329**, 73–91 (2009).

55. Luo, H. R. *et al.* Inositol pyrophosphates mediate chemotaxis in Dictyostelium via pleckstrin homology domain-PtdIns(3,4,5)P3 interactions. *Cell* **114**, 559–572 (2003).

56. Pisani, F. *et al.* Analysis of Dictyostelium discoideum inositol pyrophosphate metabolism by gel electrophoresis. *PLoS One* **9**, e85533 (2014).

57. Wilson, M. S. C. & Saiardi, A. Importance of Radioactive Labelling to Elucidate Inositol Polyphosphate Signalling. *Top. Curr. Chem.* **375**, 14 (2017).

58. Wu, M., Chong, L. S., Perlman, D. H., Resnick, A. C. & Fiedler, D. Inositol polyphosphates intersect with signaling and metabolic networks via two distinct mechanisms. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E6757–E6765 (2016).

59. Nestor, G., Anderson, T., Oscarson, S. & Gronenborn, A. M. Exploiting Uniformly13C-Labeled Carbohydrates for Probing Carbohydrate-Protein Interactions by NMR Spectroscopy. *J. Am. Chem. Soc.* **139**, 6210–6216 (2017).

60. Balla, T. Phosphoinositides: Tiny Lipids With Giant Impact on Cell Regulation. *Physiol. Rev.* **93**, 1019–1137 (2013).

61. Kjeldsen, C., Ardenkjær-Larsen, J. H. & Duus, J. O. Discovery of Intermediates of lacZ β-

Galactosidase Catalyzed Hydrolysis Using dDNP NMR. *J. Am. Chem. Soc.* **140**, 3030–3034 (2018).

62. Hirsch, M. L., Kalechofsky, N., Belzer, A., Rosay, M. & Kempf, J. G. Brute-Force Hyperpolarization for NMR and MRI. *J. Am. Chem. Soc.* **137**, 8428–8434 (2015).

63. Rodrigues, M. V. *et al.* Bifunctional CTP:inositol-1-phosphate cytidylyltransferase/CDP-inositol: Inositol-1-phosphate transferase, the key enzyme for di-myo-inositol-phosphate synthesis in several (hyper)thermophiles. *J. Bacteriol.* **189**, 5405–5412 (2007).

64. Harmel, R. K. *et al.* Harnessing 13 C-labeled myo -inositol to interrogate inositol phosphate messengers by NMR. *Chem. Sci.* **10**, 5267–5274 (2019).

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

66. Clark, J. *et al.* Quantification of PtdInsP3 molecular species in cells and tissues by mass spectrometry. *Nat. Methods* **8**, 267–272 (2011).

Chapter 4 Identification of Endogenous Protein Pyrophosphorylation by Mass Spectrometry

Reproduced in part with permission from (1) R. K. Harmel, D. Fiedler, *Nat. Chem. Biol.* **2018,** 14, 244–252.

Project Collaborators: The projects in this chapter were performed in an intensive collaboration with Dr. J. A. M Morgan. Therefore, contributions of these authors are indicated more specifically as footnotes in the text. Preliminary enrichment experiments in *S. cerevisiae* were performed by Dr. L. Yates and Dr. J. Conway. MS was performed together with the Liu group (MS core facility; former: Krause group) by Dr. M. Nadler-Holly and Dr. M. Penkert. MS data were analyzed by Dr. M. Penkert or Dr. J. A. M. Morgan. Proteomic and phosphoproteomics experiments on *S. cerevisiae* lysates were conducted by the MS core facility. High pH fractionation was performed by Dr. J. A. M. Morgan or Dr. M. Penkert. Mammalian cell lysates were prepared by Dr. J. A. M. Morgan or by L. von Oertzen.

Introduction: Design of a Pyrophosphoproteomics Workflow

PP-InsPs transmit information *via* two distinct signaling mechanisms.[1,2] **Figure 4.1**) They can either directly modulate protein activity by binding to positively charged domains or transfer their structurally unique β-phosphoryl group onto phosphoproteins, a process coined protein pyrophosphorylation.

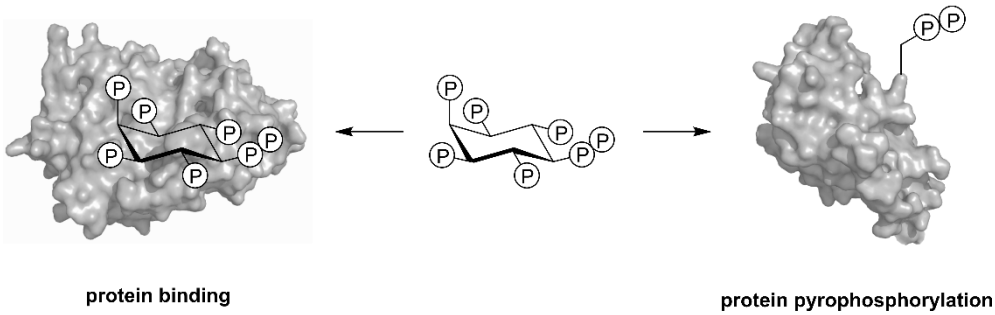


Figure 4.1: Proposed mechanisms of PP-InsP signaling. Protein binding and protein pyrophosphorylation, a covalent posttranslational modification.

Especially, the latter process is poorly understood. Radiolabeling experiments indicated that PP-InsPs could phosphorylated proteins *in vitro* and with this technique, a handful of substrates were identified.[3–7] However, the existence of this posttranslational modification *in vivo* and its dependence on PP-InsPs is still under debate because there are currently no methods available to directly detect endogenous pyrophosphorylation sites. To establish the biological significance for this putative PTM, new analytical approaches must be explored that can unambiguously detect pyrophosphorylated proteins in complex samples.

Modern MS techniques have succeeded in sensitive, rapid identification and interrogation of many PTMs in a high-throughput format.[8,9] The development of these methods required the design and optimization of bespoke proteomic workflows that overcame challenges associated with the chemical properties inherent to particular modifications. One of the best characterized modifications is protein phosphorylation.

In most cases, this PTM is investigated using a bottom-up phosphoproteomics approach.[10] (**Figure 4.2**) In this process, cells are lysed and the denatured proteins are enzymatically digested into peptides with specific proteases such as trypsin, which cleaves after the C-terminus of arginine and lysine. Like many other PTMs, phosphopeptides are usually present in substoichiometric amounts as compared to the pool of non-modified peptides.

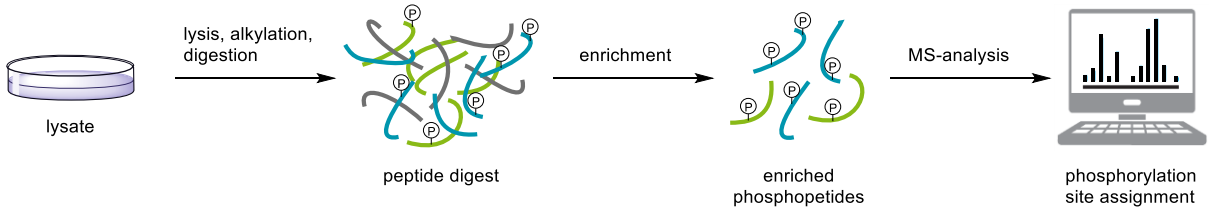


Figure 4.2: Schematic representation of the phosphoproteomics workflow. Denatured proteins are extracted from cells and enzymatically digested into peptides using specific proteases. Subsequently, the phosphopeptides are enriched and further analyzed by LC-MS/MS followed by database analysis of the obtained fragments.

In addition, the negative charge of the phosphoryl group reduces the ionization efficiency during traditional positive ion mode MS. Therefore, non-phosphorylated peptides had to be separated from tryptic digests to maximize discovery rates of phosphopeptides during MS analysis. Selective enrichment of phosphate-containing peptides is achieved by positively charged affinity reagents like IMAC or MOAC, through immunoprecipitation, chemical derivatization or chromatography.[10] Finally, enriched samples are mostly fractionated based on hydrophobicity and sequenced by tandem-MS (MS/MS). Depending on the residue attached to the phosphoryl group, specific fragmentation strategies provide fingerprint fragments that enabled the identification of phosphorylated amino acids. Today, phosphoproteomic analyses of serine, threonine and tyrosine sites are highly optimized procedures and can detect several thousand phosphorylated peptides from a single biological sample.

Establishing robust phosphoproteomic workflows with MS strongly accelerated the understanding of protein phosphorlylation and their regulation by kinase/phosphatase networks. An analogous procedure for unbiased identification of protein pyrophosphorylation by MS would probably have a large impact in this fairly new field. The chemical similarity of the phosphoryl- and pyrophosphoryl moiety inspired our research group to adopt the general phosphoproteomics approach to identify endogenous pyrophosphopeptides. It was expected that establishing this workflow would be even more challenging than analyzing phosphopeptides. Since protein pyrophosphorylation was shown to depend on a priming phosphorylation event, the levels of pyrophosphopeptides were expected to be substoichiometric to that of phosphopeptides.[3,4] Furthermore, pyrophosphopeptides bear an additional negative charge and have so far been localized to polyacidic peptide fragments that further reduce ionization. Finally, the labile phosphoanhydride bond and the anticipated difficulty in distinguishing isobaric pyrophospho- from bisphosphopeptides was expected to complicate interpretation of tandem-MS data. To overcome challenges associated with pyrophosphorylated peptides, a new set of tools was developed in our group. (**Figure 4.3**)

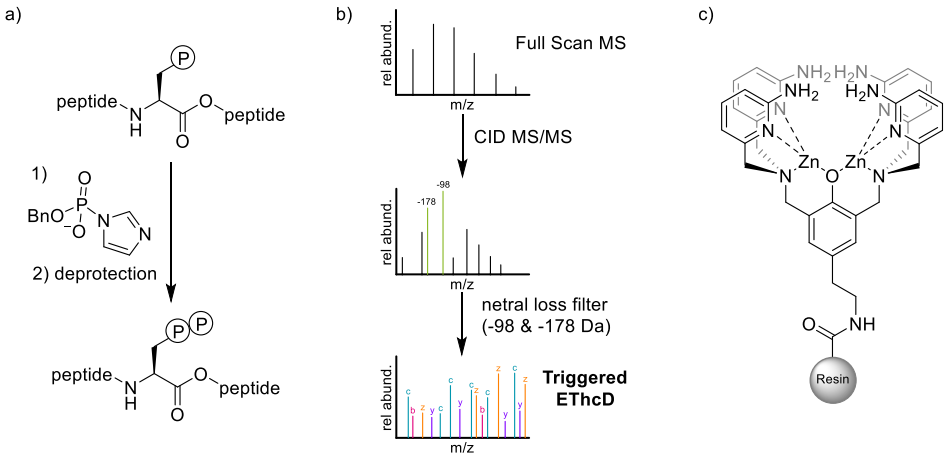


Figure 4.3: New chemical biology tools to study protein pyrophosphorylation. a) Peptide pyrophosphorylation by phosphoimidazolide reagents. b) CID neutral loss triggered EThcD fragmentation to identify and sequence pyrophosphopeptides. c) Zn-DPA affinity reagent for the enrichment of pyrophosphopeptides

More specifically, a new methodology was developed to grant access to synthetic pyrophosphopeptides using phosphoimidazolide reagents.[11] (**Figure 4.3 a**) In turn, this allowed the study of the pyrophosphate moiety during ionization and fragmentation, and the development of tailored analysis methodologies.[12,13] Gratifyingly, a unique neutral loss pattern (pyrophosphoric acid; m/z -178) during collision-induced dissociation (CID) fragmentation was identified that could be used to rapidly characterize pyrophosphorylated peptides. (**Figure 4.3 b**) Coupled to electron-transfer dissociation combined with higher energy collision dissociation (EThcD) high quality MS/MS spectra were obtained that enabled unambiguous annotation of pyrophosphosites. This data driven CID-EThcD method was later applied to identify pyrophosphorylated peptide standards spiked into HeLa protein digests where 100 fmol peptide could be reliably detected.[12] The detection sensitivity for a given pyrophosphopeptide was approximate 100-fold less than for the corresponding phosphopeptide. In parallel, pyrophosphopeptides were utilized to develop a selective enrichment method based on a Zn-DPA complex.[14] (**Figure 4.3 c**) This novel affinity reagent was able to recover spiked-in pyrophosphopeptides from complex cell lysates, and showed a 30-fold stronger binding for these peptides compared to phosphopeptides. Finally, a preliminary trial to identify endogenous protein pyrophosphorylation was conducted using lysates from *S. cerevisiae* by adopting a protocol from Gygi and coworkers.[15] Remarkably, high amounts of tryptic digest (6 mg) were necessary to find any evidence of sites above the limit of detection. Three pyrophosphopeptides were identified with high confidence, derived from proteins RRP5, RPA34 and PUF6. (unpublished results) The sequence of these peptides could be fully annotated *via* their representative ion series. Unfortunately, these results were not reproducible in further replicates and therefore more optimization appeared necessary to produce a stable result.

In this Chapter, we optimized the pyrophosphoproteomics workflow and detected the first endogenous protein pyrophosphorylation sites in mammalian cells. With the improved workflow, we could identify a total of 39 endogenous pyrophosphosites from 36 proteins with high confidence. While the corresponding proteins were located in a range of cellular compartments, the nucleus and cytoplasm were the two most common ones. Gene ontology analysis revealed an enrichment of RNA- and cadherin binding proteins and identified regulation of telomerase activity and ribonucleoprotein complex assembly among the top 3 enriched biological processes. Collectively, these initial results suggest a function for protein pyrophosphorylation in ribosome biogenesis, cell-cell adhesion and cell division.

Evaluation of the Initial Pyrophosphoproteomics Workflow

To reproducibly detect endogenous pyrophosphorylation sites from eukaryotic cells, we first isolated potential bottlenecks by further dissecting the pyrophosphoproteomics workflow used previously. The procedure consisted of 7 steps as depicted. (**Table 4.1**) One of the most crucial variables was the cell culture. Conceptually, to maximize the discovery of pyrophosphosites, an organism or cell line with a high degree of pyrophosphorylated proteins should be chosen as a starting point. Moreover, the pyrophosphoproteome should be kept stable to obtain reproducible results. Unfortunately, the abundance of pyrophosphorylation as well as the fluctuation of this PTM was unknown. Previous work had already successfully identified the first endogenous pyrophosphosites in *S. Cerevisiae*. Therefore, analyzing the proteome of this organism was a rational starting point.

Table 4.1: Potential bottlenecks that may hamper the detection of protein pyrophosphorylation in the first generation pyrophosphoproteomic workflow.

Entry	Step	Potential Bottlenecks
1	cell culture	low levels of pyrophosphorylated proteins
2	cell lysis	enzymatic pyrophosphate degradation
3	protein alkylation	pyrophosphate instability
4	tryptic digestion	inefficient tryptic digestion
5	peptide desalting	loss of peptides
6	Zn-DPA enrichment	no pyrophosphopeptide binding
7	enriched peptide desalting	loss of peptides

Apart from ensuring the presence of *in vivo* pyrophosphosites amenable to detection in the biological sample, we were also concerned about the stability and recovery of pyrophosphorylated peptides throughout the course of the workflow. During cell lysis and alkylation, residual phosphatase activity and elevated temperatures at pH 8.0 could possibly degrade the endogenous pyrophosphates. Seminal work from Yates *et al.* investigated the hydrolysis of pyrophosphopeptides in the presence of different buffer conditions and in *S. cerevisiae* cell lysates.[16] Only minor degradation was observed in the presence of Tris at pH 7.5, even at 37 °C over 24 hours. Complete phosphatase activity in cellular lysates was however only abolished through a combination of EDTA and phosphatase inhibitor treatment. The original phosphoproteomics workflow was EDTA-free and the addition of this chelator potentially complicates subsequent steps. Therefore, we refrained from using EDTA in initial experiments as we expected no phosphatase activity under the high denaturing conditions generated by the 8 M Urea in the lysis buffer.

Improving the recovery of peptides during processing is particularly relevant to steps 4–7. The trypsin/peptide ratio and the desalting conditions were not optimized for the highly acidic sequences expected. In addition, pyrophosphopeptides could be lost due to a failed enrichment step. To control for these possible risks, we envisioned to apply pyrophosphopeptide standards during the workflow or during the validation of single steps to estimate the loss of peptides. Together, these measures should help to isolate problems of the pyrophosphoprotemics workflow and allow further optimization.

Attempts to Optimize the Pyrophosphoproteomics workflow in *S. Cerevisiae*

After assessing the possible pitfalls of the pyrophosphoproteomics workflow, fresh immobilized Zn-DPA reagent was prepared in 15 chemical steps according to a procedure from Conway *et al*. and we tested its ability to bind and elute two fluorescently labeled pyrophosphopeptides.[14] About 60% of the initial peptide concentration was recovered, consistent with the literature. Secondly, we prepared three biological replicates of wild type *S. cerevisiae* and subjected these samples to proteomic and phosphoproteomic analysis. While the unmodified proteome remained very similar, there were minor differences found within phosphopeptides. These results suggested that our growth conditions produce a fairly stable background at the proteomic and phosphoproteomic levels.

Next, we applied the complete pyrophosphoproteomics workflow to one of the *S. cerevisiae* samples, but MS analysis showed low levels of peptide in this sample. We noticed that the enriched peptide desalting (step 7) included a washing step with neutral milliQ that might elute the majority of the peptides. Avoiding this step allowed the recovery of phosphorylated and unmodified peptides. Notably, MS analysis of these samples was complicated by polymer contaminations and missed tryptic

cleavage sites were commonly encountered. These problems were reduced by switching from plastic to glass tubes and raising the ratio of trypsin/protein from 1:100 to 1:20 but these changes did not lead to the identification of pyrophosphopeptides. In an attempt to identify a specific step within the proteomic workflow where highly acidic pyrophosphorylated were lost, we produced four samples where a synthetic standard (1–10 pmol) was added at four different steps: during lysis (step 2), before enrichment (before step 6); before enriched peptide desalting (before step 7); to the enriched and desalted peptide solution. We employed a modified version of the previously found PUF6, with the aim of modeling an endogenous peptide sequence but without interfering with the detection of the natural peptide. In all cases, the peptide standard could be readily detected in the final MS analysis but no endogenous pyrophosphopeptides were observed. This suggested two possibilities: either pyrophosphorylated peptides were too low in abundance compared to the synthetic standards and below the limit of detection, or the enrichment efficiency varied greatly between endogenous sequences and the PUF6 model peptide. Finally, two more samples were produced either using more lysate (6 mg instead of 1 mg), or treated with λ-phosphatase before enrichment to specifically degrade background from phosphorylated peptides. However, neither condition appeared to improve detection, as no pyrophosphopeptides were detected.

Identification of Endogenous Mammalian Pyrophosphopeptides[6]

We speculated that *S. cerevisiae* might have low levels of endogenous pyrophosphorylated proteins and turned to explore mammalian HEK293T cells instead. We enriched 6 mg of tryptic peptides with Zn-DPA affinity reagent and included the PUF6 standard during lysis. After performing the complete workflow and analyzing the output, no pyrophosphorylated peptides were detected. In a parallel approach undertaken by our collaborator at the MS core facility, Zn-DPA-enriched peptides were subjected to high pH reversed fractionation prior to LC-MS/MS.[17,18] Remarkably, this approach removed enough background signals to identify the first pyrophosphorylated peptide from a mammalian cell line. The observed peptide sequence was derived from the dual specificity kinase 1 or 4 (CLK1/4) and the majority of the amino acids, including the pyrophosphoserine, were annotated with high confidence by the MS. (**Figure 4.4 a**) Surprisingly, the identified site was not phosphorylated by CK2 which was previously reported for all *in vitro* characterized pyrophosphoproteins. Currently, there are no reports of other kinases modifying this site. Instead CLK proteins are known to widely autophosphorylate itself and this could be a potential priming phosphorylation in this case.[19] No function has been associated with phosphorylation at this site, despite it being among the most prevalent phosphosites observed in CLK1/4 by phosphoproteomics experiments. Finally, a crystal structure of CLK1 was published that located two phosphorylated residues, one at the identified position of the proteomic analysis and a second one immediately adjacent.[20] (**Figure 4.4 b**) Interestingly, the phosphoserine is in close proximity to the active site and

[6] hSAX fractionation was developed and performed by Dr. J. A. M. Morgan. Optimization of the residual workflow was performed by R. K. Harmel including the IMAC or Zn-DPA enrichment in combination with hSAX. Preliminary experiments with λ-phosphatase treatment were performed by R. K. Harmel. Further optimization including λ-phosphatase treatment and SIMAC enrichment, and identification of the residual 32 sites was performed by Dr. J. A. M. Morgan.

pyrophosphorylation potentially results in a clash with substrate suggesting a putative function in kinase inhibition.

a)

C^+	Seq.	$Z+1^+$
117,10224	V	
216,17065	V	2381,93229
331,19760	D	2282,86388
478,26601	F	2167,83694
535,28747	G	2020,76852
622,31950	S	1963,74706
693,35662	A	1876,71503
794,40429	T	1805,67792
957,46762	Y	1704,63024
1072,49457	D	1541,56691
1187,52151	D	1426,53997
1316,56410	E	1311,51302
1453,62301	H	1182,47043
1590,68193	H	1045,41152
1677,71395	S	908,35261
1778,76163	T	821,32058
1891,84570	L	720,27290
1990,91411	V	607,18884
2237,87880	S-pyropho	508,12042
2338,92648	T	261,15573
	R	160,10805

b)

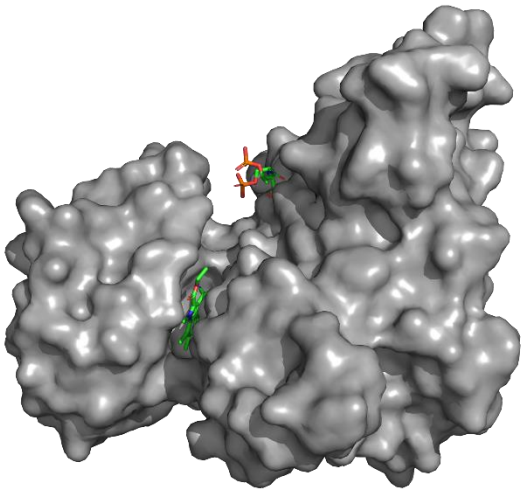


Figure 4.4: Analysis of CLK1/4 pyrophosphosite. a) Observed C^+ and Z^{+1} series for pyrophosphorylated peptide of CLK1/4. Colored m/z values were detected. b) Crystal structure of CLK1 reported by Knapp and coworkers.[20]

Encouraged by these results, the fractionation method was further optimized, and we switched to a hydrophilic hSAX column with NH_4HCO_2 buffer at pH 9 that was suitable to enrich highly acidic peptides. Combining this method with Zn-DPA enrichment, the identification of CLK1/4 was reproduced in a second biological replicate and additionally, a peptide originating from mitochondrial structural protein and transcription factor MICOS complex subunit (MIC19) was identified. The hSAX chromatography significantly reduced background signals during the MS analysis and therefore we reconsidered traditional IMAC as replacement for Zn-DPA and prepared two samples for a direct comparison.[21] In the IMAC enriched sample we found only CLK1 but the ion intensity was approximately 10 fold higher compared to the Zn-DPA enriched sample. By contrast, the Zn-DPA displayed 6 hits that were confidently

assigned as pyrophosphopeptides. This comparison suggested to us that recovery of the peptide was less important than its enrichment and separation from background phosphopeptides. To obtain more hits in a single pyrophosphoproteomics experiment, the workflow was further optimized. (**Figure 4.5**) More site were observed by raising the amount of mammalian cell lysate to 20 mg. Previously, our group showed that phosphorylated peptides were readily dephosphorylated in the presence λ-phosphatase while pyrophosphorylated peptides were not affected.[16] Therefore, an incubation step with λ-phosphatase was included which significantly improved data quality. This effect was probably due to the removal of a significant number of phosphopeptides and the depletion of phosphorylation sites close to pyrophosphorlation sites which complicate analysis. In the enrichment step, sequential IMAC (SIMAC) spin columns offered the best results.

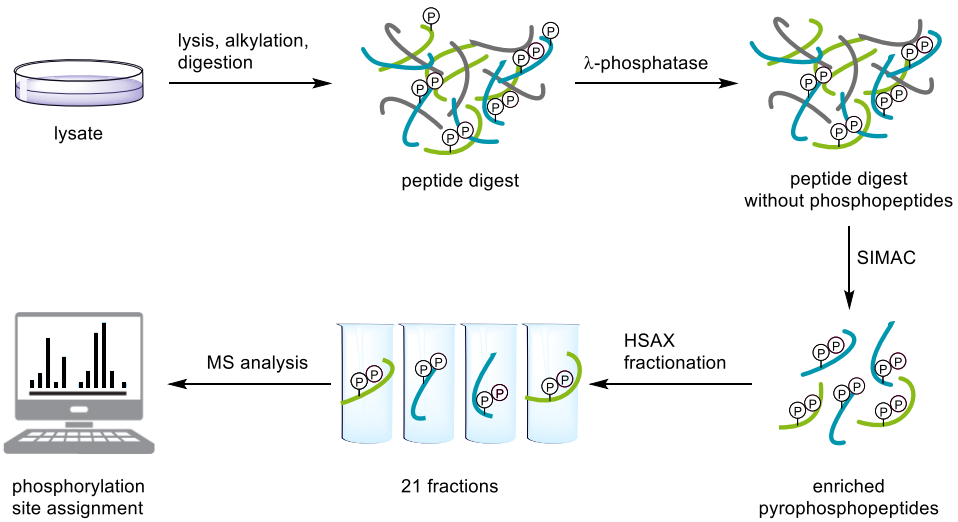


Figure 4.5: Optimized pyrophosphoproteomics workflow. A peptide digest is isolated from mammalian cells by lysis, alkylation and enzymatic digestion. Treatment with λ-phosphatase strongly reduces phosphopeptides followed by enrichment of the pyrophosphopeptides using SIMAC. hSAX fractionation produces 21 fractions that are individually analyzed by LC-MS/MS.

Compared to traditional IMAC this techniques includes an additional elution step to separate monophosphorylated from multiphosphorylated peptides.[22] Pyrophosphorylated peptides were eluted in the basic elution step and background from phosphorylated peptides was reduced. Finally, the buffer for the hSAX fractionation was changed to $(NH_4)_2CO_2$ at pH 9 to minimize the amount of salt after lyophilization. Notably, the low content in fractionated samples eliminated the need for the final desalting step, further reducing the risk of peptide loss during the workflow.

Analysis of Identified Pyrophosphorylated Proteins

With the improved pyrophosphoproteomics workflow we were able to identify 5–15 pyrophosphorylated peptides from HEK293 or HCT116 cell lysates in a single analysis, and together, 39 sites among 36 proteins were found. (**Appendix: Table 4.2**, **Table 4.3**) These include therapeutically interesting examples such as tau protein, histone deacetylase 2 (HDAC2), phosphohistidine kinase NME1 and the chaperone HSP90. The most common hit within all samples so far remained CLK1/4. It was the only protein that was consistently found in all our optimization attempts. An abundant pyrophosphorylation site in CLK1/4 may be promoted through its ability to install the priming phosphorylation by its autophosphorylation activity.[19] This modification is likely to occur at a high stoichiometry as it is independent of the interaction with other kinases. In contrast, prephosphorylation of most target sites was dependent on CK2 and CDK1 (cyclin dependent kinase 1) . Even though the initial list of hits was relatively short, identification of the consensus sequence and gene ontology analysis was performed.[23] (**Figure 4.6**)

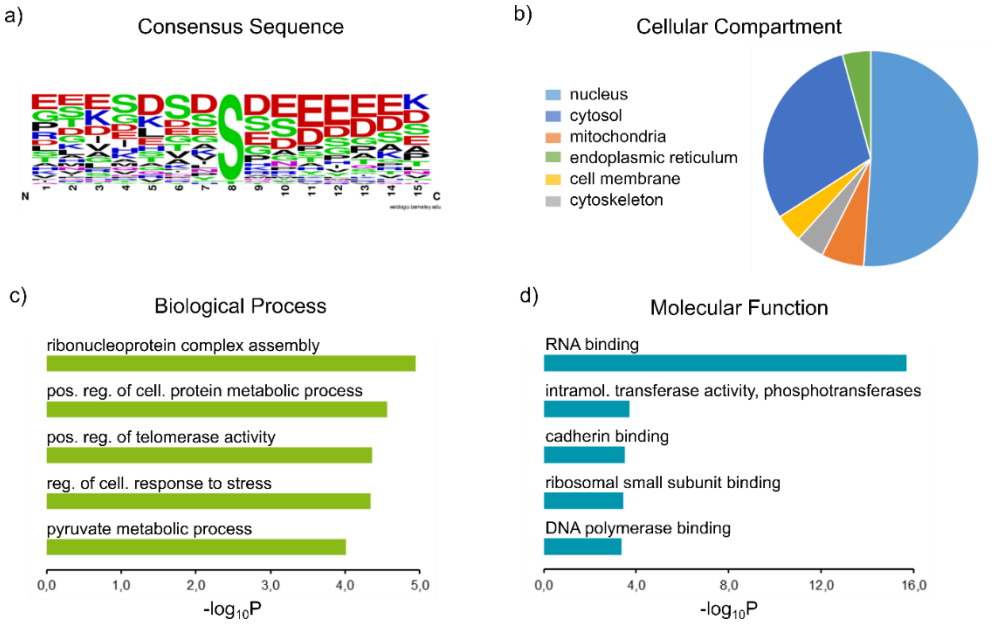


Figure 4.6: Sequence and gene ontology analysis of pyrophosphoproteins. a) Consensus sequence. b) Distribution of proteins among cellular compartments. c) GO analysis of biological processes. d) GO analysis of molecular function.

The consensus sequence revealed serine residues enclosed in a highly acidic environment, consistent with sites previously identified in biochemical assays via mutagenesis. (**Figure 4.6 a**) Furthermore, protein pyrophosphorylation was not specific to a certain group of proteins and occurred in a range of cellular compartments but mostly in the nucleus and the cytosol suggesting a fundamental role of this modification in cell signaling. Gene ontology analysis showed an enrichment of biological processes such as ribonucleoprotein complex assembly and regulation of telomerase activity, and among the top enriched molecular functions was binding to cadherin and RNA.[23] These data suggested that protein pyrophosphorylation is involved in ribosome biogenesis, cell-cell adhesion and cell division.

Conclusion and Outlook

Posttranslational phosphorylation of proteins is a highly regulated process that governs many biological processes and is an integral part of signal transduction. Consequently, powerful proteomic workflows have been developed to grant thorough analyses of this PTM. Here, the development and optimization of a pyrophosphoproteomic approach was described and the first endogenous mammalian pyrophosphopeptides were identified. Current efforts have accumulated a list of 38 endogenous targets of pyrophosphorylation and gene ontology analysis revealed a putative function in ribosome biogenesis, cell-cell adhesion and cell division.

Future work will be directed towards further optimizing the workflow. In particular, the reproducibility within technical and biological triplicates need to be addressed to create consistency. If we want to understand the biological function of this PTM, cells that are either treated differently or possess different genetic backgrounds need to be compared, ideally using SILAC to quantify differences. Unless background variability can be distinguished from experimentally-induced change, a meaningful interpretation of these data will be elusive.

Mechanistically, protein pyrophosphorylation is not well understood yet. Currently, it is unclear whether PP-InsPs are responsible for protein pyrophosphorylation *in vivo*. Genetic knockouts are useful tools to correlate PP-InsP levels with changes in pyrophosphoprotein content, and depletion or increase of intracellular PP-InsP concentrations is expected to result in a response on the PTM level with a similar trend. To further understand the impact of pyrophosphorylation on the structure and function of modified proteins, it is of utmost importance to pick pyrophosphoproteins from our current data set and characterize their properties with respect to unmodified- and phosphoproteins *in vitro*. To this end, our group recently

established a procedure to produce pyrophosphoproteins by combining the expression of phosphoproteins with genetic code expansion followed by pyrophosphorylation with phosphoimidazolide reagents.[24]

Finally, to further understand the regulation of protein pyrophosphorylation, factors that facilitate the introduction and removal of this modification need to be assessed. Previously, pyrophosphopeptides were reported to be degraded by several unspecific phosphatases *in vitro* and within *S. cerevisiae* and HeLa lysates suggesting reversibility of protein pyrophosphorylation.[25] Identification of a putative pyrophosphatase could be achieved by activity enrichment or proteomic pulldowns with identified pyrophosphopeptides and would uncover new functional aspects of protein pyrophosphorylation. Genetic knockout of the pyrophosphatase potentially leads to a massive increase in protein pyrophosphorylation, further enhancing the detectability of this PTM. Analogous to phosphatases, substrate-trapping pyrophosphatases might be prepared through mutation of the catalytic residues to alanine and used a highly selective enrichment strategy for pyrophosphopeptides.[26,27]

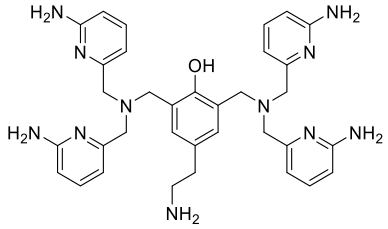
Taken together, our novel pyrophosphoproteomics workflow will significantly aid the identification and the understanding of protein pyrophosphorylation and − assuming PP-InsPs install this modification − identify new intersections of PP-InsPs within metabolism and signaling networks.

Methods

General Information

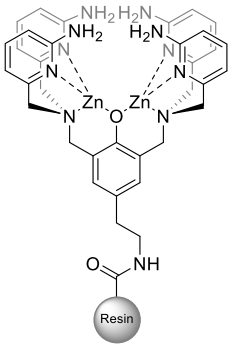
Please see method section of chapter 2.

Chemical Synthesis of Zn-DPA Resin



DPA ligand was synthesized in 13 steps starting described before by Conway and Fiedler.[14] Analytical data were identical with the values reported in the literature.

^{1}H NMR (300 MHz, Deuterium Oxide, pD 7.0) δ 7.53 (dd, *J* = 8.9, 7.3 Hz, 4H), 6.68 (s, 2H), 6.58 (dd, *J* = 14.5, 8.1 Hz, 9H), 3.76 (s, 8H), 3.55 (s, 4H), 2.88 (dd, *J* = 10.7, 6.3 Hz, 2H), 2.58 (dd, *J* = 10.8, 6.2 Hz, 2H).



Zn-DPA resin was synthesized in 2 steps starting described before by Conway and Fiedler.[14] Immobilization was followed by disappearance of starting material in LC-MS spectra. Zn^{2+} was freshly loaded right before use.

Yeast Cell Culture and Harvest

S. cerevisiae cells were prepared according to Villen and Gygi.[21] Wild type *S. cerevisiae* cells were grown on an agar plate overnight at 30 °C. One colony was picked, suspended in YPDA medium and incubated at 30 °C while shaking. The next day, cells were diluted to an OD_{600nm} of 0.1 and grown over 5–6 to mid-log phase (OD_{600nm} = 0.600–0.700) at 30 °C. Cells were harvested by centrifugation at 3,500 g for 10 min at 4 ˚C. The pellet was rinsed with 50 mM Tris-HCl (pH 7.50; 10.0 mL) and isolated by centrifugation at 3,200 g for 5 min at 4 ˚C. 1 L of culture gave approximately 100 mg of undigested protein.

Mammalian Cell Culture and Harvest

HCT116 or HEK293T cells were grown according to Villen and Gygi.[21] Cells were grown in 15 cm dishes in DMEM supplemented with FBS and Penicillin/Streptomycin at 37 °C and 5 % CO_2. Cells were split at 70 % confluency 1:10 and distributed on 5–10 plates and grown until they reached 80–90% confluency. For harvesting, cells were washed with PBS and 0.9 % NaCl solution, and trypsinized. The trypsin was quenched with DMEM. The cells were collected and washed two more times with PBS before lysis and extraction. One 15 cm dish of culture gave approximately 10 mg of undigested protein.

Yeast Lysate Preparation

S. cerevisiae lysate was prepared according to Villen and Gygi.[21] To prepare yeast lysates, two EDTA-free protease inhibitor cOmplete tablet (Roche Life Sciences) was added to the mass spec (MS) lysis buffer containing phosphatase inhibitors (8 M urea, 75 mM NaCl, 50 mM Tris (pH 8.2), 50 mM NaF, 50 mM β–glycerophosphate, 1 mM sodium orthovanadate, 10 mM sodium pyrophosphate, 1 mM PMSF; 10.0 mL) at 4 ˚C. All steps of the lysis were performed at 4 ˚C. The cell pellets of 50 mL media were treated with the MS lysis solution (300 µL lysis buffer per cell pellet), and mixed by vortexing. The cell suspension (300 µL) was transferred to screw-capped tubes, and glass beads were added. The cells were lysed by bead beating (Precellys Evolution, Bertin Technologies) at 6500 rpm using 2 cylcles of 15s with 1 min pause in between cycles. After the lysis was complete, a hole was poked in the bottom of the tubes, and the sample was removed from the glass beads by centrifugation at 16,000 g for 1 min. Cellular debris was removed by centrifugation at 16,000 g for 10 min. The supernatant was removed, and the protein concentration was determined by BCA kit (Pierce).

Mammalian Lysate Preparation

HCT116 or HEK293T lysates were prepared according to Villen and Gygi.[21] To prepare mammalian lysates, one EDTA-free protease inhibitor cOmplete tablet (Roche Life Sciences) was added to the mass spec (MS) lysis buffer containing phosphatase inhibitors (8 M urea, 75 mM NaCl, 50 mM Tris (pH 8.2), 1 mM NaF, 1 mM β–glycerophosphate, 1 mM sodium orthovanadate, 10 mM sodium pyrophosphate, 1 mM PMSF; 10.0 mL) at 4 ˚C. All steps of the lysis were performed at 4 ˚C. Cells in 15 cm dishes were washed with ice cold PBS (5 mL). After aspiration of the PBS, lysis buffer was added and the cells chilled for 10 min at 4 °C and the cells were scraped off by a

cell lifter and collected into 15 mL conical tube. Sonication of 3 × 1 min with 2 min rests at 15 % output was applied during lysis and cellular debris could be removed by 2,500 g for 10 min. The supernatant was collected, and the protein concentration was determined by BCA kit (Pierce).

Protein Reduction and Alkylation

Reduction and alkylation was performed according to Villen and Gygi.[21] To reduce and alkylate the proteins in the lysate, 5 mM DTT (50 µL of a 100 mM frozen stock) was added to each 1 mL lysate aliquot, and the samples were incubated at 56 ˚C. After incubation for 25 min, the samples were removed and cooled to room temperature. 14 mM iodoacetamide was added to each 1 mL aliquot, and the samples were incubated at room temperature in the dark for 30 min. The samples were then quenched with another portion of DTT (10 mM final concentration) and incubated for 15 min at room temperature in the dark.

Tryptic Protein Digestion

Tryptic digestion was performed according to Villen and Gygi.[21] For the trypsin digestion of the lysate, sequencing grade modified trypsin (Promega, cat. V5111) was used. The reduced and alkylated samples were diluted 1:5 in 25 mM Tris pH 8.0. To each sample was added 1 mM $CaCl_2$ and trypsin at a ratio of 1:50 protease to protein from the lysate (as determined by BCA protein concentration). The samples were incubated overnight at 37 ˚C with agitation. After 14–16 hours, the samples were removed and cooled to room temperature.

Desalting of Tryptic Peptides

Desalting of tryptic peptides was performed according to Villen and Gygi.[21] Neat TFA to a final concentration of 0.4% was added to each sample, which was spun for 10 min at 2,500 g to remove any cellular debris. For the peptide desalting step 500 mg Waters C18 SepPak cartridges was used (up to 10 mg loaded). The Waters C18 SepPak was washed with MeCN (9 mL) and then with 3 mL of 50% MeCN in H_2O with 0.5% AcOH. The column was equilibrated with 0.1% TFA in H_2O (9 mL) and then loaded with the trypsinized peptide samples in 0.4% TFA in H_2O. The column was washed with 0.1% TFA in H_2O (9 mL) to desalt the peptides. The counter-ion was exchanged by a final wash with 0.5% AcOH in H_2O (900 µL). Peptides were eluted with 5–6 mL of 50% MeCN in H_2O with 0.5% AcOH. The eluent was concentrated by freeze-drying to isolate the peptides as a white to off-white solid. (For 10 mg of undigested protein we usually obtained 6–7 mg of desalted tryptic digest.)

Desalting of Zn-DPA and IMAC Enriched Peptides

Desalting of Zn-DPA or IMAC enriched peptides was performed according to Villen and Gygi.[21] Elution fractions from IMAC or Zn-DPA were combined and acidified (pH < 2) by the addition of 10% formic acid for desalting by an Oasis HLB column. For the optimized workflow, the acidified elution fraction was then desalted using a 10 mg 1 cc Oasis HLB cartridge. The Oasis HLB column was conditioned with MeOH (1 mL) and then equilibrated with H_2O (1 mL). The acidified sample was loaded onto the cartridge, and the sample was washed with 5% formic acid in H_2O (1 mL). The peptides were eluted with 5% NH_4OH in 50% MeOH in H_2O (1 mL). The eluent was concentrated by freeze-drying to yield the enriched peptides that can be directly resuspended for mass spectrometry analysis.

Zn-DPA Enrichment

Zn-DPA enrichment was performed according to Conway et al.[14] To begin the enrichment procedure with the Zn-DPA resin, the incubation buffer (25 mM HEPES, pH 7.4, 100 mM KCl), Pi wash buffer (1 mM Pi in incubation buffer), and PPi elution buffer (5 mM PPi in incubation buffer) were prepared. Resin loading as determined by proton NMR was between 1.0 and 1.7 μmol/mL. All steps of the enrichment were carried out at 4 ˚C. Zn-DPA resin slurry (200 μL, ~200–300 nmol) was transferred to Eppendorf tubes (1 tubes for ~6.5 mg of desalted tryptic peptides). The resin was washed with incubation buffer (3 × 500 μL) and spun at 3,000 g for 3 minutes to precipitate resin. The supernatant was removed and discarded. The Zn-DPA resin was incubated with fresh $ZnBr_2$ (10 mM, 26.1 mg in 10 mL incubation buffer) for 10 minutes with agitation. The samples were spun, and the supernatant was discarded. The resin was washed with incubation buffer (2 × 950 μL). The samples were spun, and the supernatant was discarded. 1 mL of tryptic digested peptides (6.5 mg/mL) in incubation buffer were added to each 200–300 nmol resin portion and agitated for 1 hour. The samples were spun at 3,000 g for 3 minutes, and the supernatant was saved as the flow-through containing non-phosphorylated peptides. The resin was washed with incubation buffer (2 × 950 μL), spun, and the supernatant was discarded. 1 mM Pi wash buffer (500 μL) was added to the resin, and the samples were incubated for 10 minutes with mixing. The samples were spun, and the supernatant was saved. The Pi wash step was repeated 2 times. The pyrophosphorylated peptides were then eluted from the resin by treating the Zn-DPA reagent with 5 mM PPi elution buffer (500 μL) and mixing for 10 minutes. The samples were spun at 3,000 g for 3 minutes, and the supernatant was saved as the elution fraction. The PPi elution step was repeated once.

IMAC Enrichment

IMAC enrichment was performed according to Villen and Gygi.[21] The binding buffer containes 40% MeCN in H_2O with 25 mM formic acid. The elution buffer contained 50 mM K_2HPO_4, pH adjusted to 10 with NH_4OH. All steps of this enrichment will be conducted at room temperature. IMAC beads (3 × 120 µL for 3.4 mg of tryptic peptide) are washed with 1.2 mL of binding buffer and will be spun at 10,000 rpm for 30 sec to precipitate beads. The supernatant will be removed, and this step will be repeated 3 times. After the last wash supernatant will be removed, each bead portion is resuspended in 120 µL of binding buffer to make a 50% resin slurry. 120 µL of the slurry will be added to Eppendorf tubes and then treated with 1 mL of tryptic digested peptides in binding buffer (6.5 mg peptide in 1 mL of binding buffer, 6.5 mg/mL concentration). The samples will be incubated with agitation for 60 minutes. At the completion of the incubation, the samples will be spun at 10,000 rpm for 2 minutes, and the supernatant will be removed containing the non-phosphorylated peptides. The resin is washed (3 × 720 µL) with binding buffer. The peptides are eluted from the resin using the elution buffer (3 × 240 µL). The elution fractions will be combined, and further desalted.

SIMAC Enrichment

SIMAC enrichment was performed with A ‘High-Select™ Fe-NTA Phosphopeptide Enrichment Kit’ from Thermo. The digest from the previous step was dissolved in SIMAC loading buffer (1348 µl, 0.1% TFA, 50% ACN). 800 µl (20 mg) of the digest solution was loaded onto 4 supplied spin columns (4 columns used in total, 200 µl loading per column) and the manufacturer’s protocol followed with minor adjustments. (The remaining digest solution was frozen at –80 and lyophilized.)

Spin column equilibration: Remove the bottom closure and loosen screw cap. The column was placed in a 2 mL microcentrifuge tube and storage buffer was removed at 1,000 g for 30 sec. Remove the screw cap and save it for "Bind phosphopeptides" step 2. Add 200 µL of Binding/Wash Buffer. Centrifuge at 1000 × g for 30 seconds and discard the flowthrough. Repeat this step once. Cap the bottom of the column with a white Luer plug. Place the column with the plug into the empty microcentrifuge tube.

Bind phosphopeptides: Add 200 µL of the suspended peptide sample to the equilibrated spin column. Close the screw cap. Mix the resin with the sample by holding the screw cap and very gently tapping the bottom plug for 10 seconds until the resin is insuspension. Do not vortex or invert the column to avoid splashing the resin inside the column wall. (Note: Mixing samples by vortexing or inversion results in significantly higher nonspecific peptide binding.) Incubate for 30 minutes. Mix the resin gently every 10 minutes as described in step 2. Carefully remove the bottom plug and the screw cap. (Note: Do not squeeze the bottom of the plug during removal as liquid inside the plug can backflow into the spin column.) Place the column into the microcentrifuge tube. Centrifuge at 1000 × g for 30 seconds. Discard the flow through.

Wash column: Wash column by adding 200 µL of Binding/Wash Buffer. Centrifuge at 1000 × g for 30 seconds. Repeat this step two additional times for a total of 3 washes.

Discard the flowthrough. Wash column by adding 200 µL of LC-MS grade water. Centrifuge at 1000 × g for 30 seconds. Wash with SIMAC buffer A (1% TFA, 20% MeCN, 2 × 100 µL).

Elute column: Place column in a new microcentrifuge tube. Add 100 µL SIMAC buffer B (10 mL of 25% $NH_{3(aq)}$ in 500 mL MilliQ, 0.5%). Centrifuge at 1000 × g for 30 seconds. Repeat this step once. (Note: Resin color may turn brown, but this is normal.) Combine eluents from all column in a 2 mL low binding tube, flash freeze and lyophilize.

λ-Phosphatase Treatment

Using a NEB λ −phosphatase kit, the desalted protein digest from the previous step were dissolved in 5 ml of reaction buffer (homemade, 50 mM HEPES, 100 mM NaCl, 2 mM DTT, 0.01% Brij 35) and treated with 40000 units of λ −phosphatase (10000 units per 10 mg) in a 50 ml falcon tube. The sample was incubated at 37°C at 300 RPM for 5 hours in a heater shaker. The sample was then desalted. Neat TFA was added (0.4%, 40 µL), the pH checked (less than pH 2), and the sample spun for 10 min at 2,500 g to remove any cellular debris. For the peptide desalting step, one 500 mg Waters tC18 SepPak cartridge were used (38 mg/ 5 ml loading per cartridge). Air pressure was used speed up washing and elution steps but was not used during the loading steps. The SepPak columns were washed with MeCN (9 mL) and then with 3 mL of 50% MeCN in H2O with 0.5% AcOH. The columns were equilibrated with 0.1% TFA in H2O (9 mL) and then loaded with the peptide samples in 0.4% TFA in H2O. The column was washed with 0.1% TFA in H2O (9 mL) to desalt the peptides. The counter-ion was exchanged by a final wash with 0.5% AcOH in H2O (1 mL). Peptides were eluted with 6 mL of 50% MeCN in H2O with 0.5% AcOH sequentially into 15 ml

falcon tubes, with 15-20 min equilibration time before air pressure was used to elute. The eluent was combined in a single 50 ml falcon tube and freeze-dried to isolate the peptides as a white solid.

hSAX fractionation

Enriched samples were fractionated on a hSAX column (dionex ionpac As24) using an Agilent 1260 Infinity Quaternary analytical HPLC system with quaternary solvent delivery system, an auto sampler, a diode-array detector, and a fraction collector. Buffer A: 20 mM $(NH_4)_2CO_3$ pH 9.0, 20% MeCN. Buffer B: 500 mM $(NH_4)_2CO_3$ pH 9.0, 1% MeCN. Gradient: 0% B for 5 min, 0–50% B from 5–65 min, 50–100% B from 65–95 min, 100% B from 95–105 min.

The lyophilized and enriched peptides were dissolved in 100 µL buffer A and fractionated in one HPLC run using a gradient depicted and collecting 21 fraction (5 min per fraction). The fractions were flash frozen and lyophilized and handed to the MS corefacility.

LC-MS/MS

LC-MS/MS was performed based on Penkert *et al*.[12] Prior to every measured sequence of samples, the trapping and analytical column was flushed with 0.1% citric acid in 1% acetonitrile to chelate any residual metal ions. For LC-MS/MS experiments, enriched peptides were dissolved in water and analyzed by a reversed-phase capillary liquid chromatography system (Dionex Ultimate3000 NCS-3500RS Nano, Thermo Scientific, San Jose, CA) connected to a Thermo Scientific Orbitrap Fusion massspectrometer (Thermo Fisher Scientific). LC separation was performed on an in-

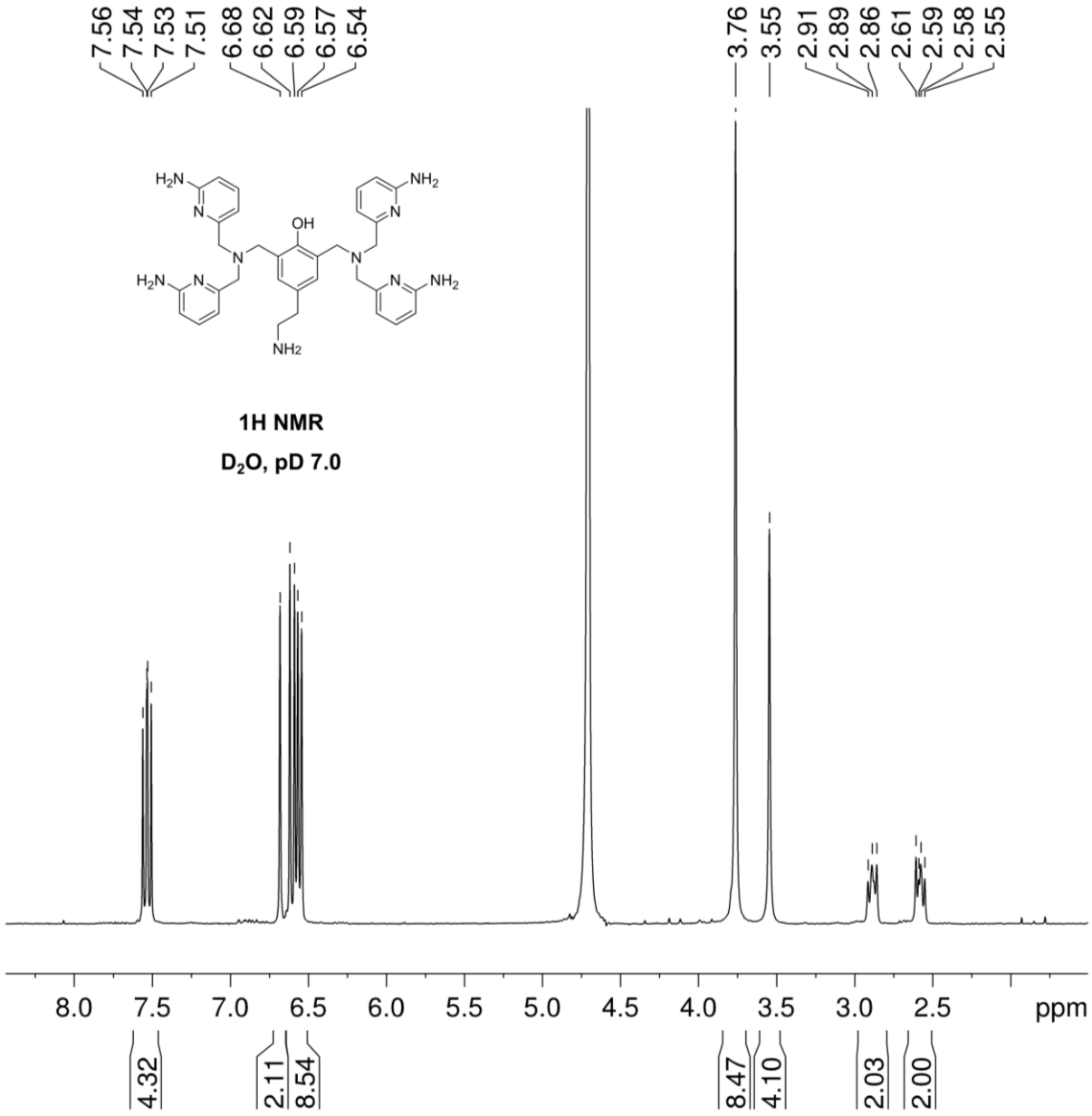
house packed 75µm inner diameter Pico Tip column (25 cm) packed with ReproSil-Pur C18AQ particles, 3µm, 120 Å (Dr. Maisch, Germany). The flow rate was 200 nL/min using gradient of 2−30% B in 60 min. Mobile phase solvent A contained 0.1% formic acid in water, and mobile phase solvent B contained 0.1% formic acid in acetonitrile.

Mass spectra were acquired in data dependent acquisition mode with a FT survey scan in the range from 350 to 1500 m/z at a resolution of 120 000 (full width at half-maximum, fwhm). The maximum injection time for MS scans was set to 50 ms and the AGC target value, to 2×10^5. Doubly and triply charged precursor ions were selected with precursor priority to the higher charge state. Precursor ions were isolated with a mass selecting quadrupole within an isolation window of 1.6m/z. Dynamic exclusion was enabled with an exclusion interval of 10 or 40 s. CID MS/MS scans were performed with a normalized collision energy of 25% and an AGC target value of 5 × 103. MS/MS spectra were acquired in the ion trap with a maximum injection time of 50 ms. If neutral losses of 49 and 89 m/z from doubly charged precursor ions were measured above a threshold of 15% and belonged to the five most abundant ions, an additional EThcD spectrum of the same precursor ions was acquired. If triply charged precursor ions showed neutral losses of 32.7 and 59.3 m/z during CID fragmentation, an additional EThcD scan was only performed if the observed signals exhibit relative intensities above 15% and were ranked among the 6 most intense signals. EThcD spectra were measured in the Orbitrap with a resolution of 15 000 and the maximum accumulation time for an AGC target value of 1×10^5 was set to 1 s. Calibrated charge dependent ETD parameters were used for the ETD process, and HCD supplemental activation was enabled. SA collision energy was set to 50% for charge state 2 and between 25% and30% for charge state 3 (calculated on m/z and charge state of the precursor ion).

MS/MS data analysis

MS/MS data analysis was performed based on Penkert *et al*.[12] MS raw data were analyzed with Proteome Discoverer 2.1 software (Thermo Fisher Scientific, Bremen, Germany). To simplify EThcD spectra, the nonfragment filter was used with the following parameters: Precursor ions and charged reduced precursors were removed within a 4 Da window and neutral losses within a 2 Da window. MS/MS spectra were searched against the Yeast proteome data base using SEQUEST or Mascot (Matrix, Science, UK). Up to two missed cleavages were allowed for protease digestion with trypsin. Precursor mass tolerance and fragment mass deviation was set to 10 ppm and 20 mmu, respectively. Oxidation of methionine, phosphorylation (STY), pyrophosphorylation (S,T), and carbamidomethyl (C) were searched as variable modifications. Results were filtered with the Target Decoy PSM Validator and a Target false discovery rate (FDR) of 0.01 at peptide identification level. EThcD MS/MS spectra of pyrophosphorylated peptides were additionally, manually verified. Manual verification was performed considering potential gas-phase rearrangements to other phosphoacceptors, the degree of neutral loss of labile modifications, the sequence coverage of the peptide, and the signal-to-noise ratio.

NMR Spectra


7.56
7.54
7.53
7.51
6.68
6.62
6.59
6.57
6.54
3.76
3.55
2.91
2.89
2.86
2.61
2.59
2.58
2.55
1H NMR
D2O, pD 7.0
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
ppm
4.32
2.11
8.54
8.47
4.10
2.03
2.00

References

1. Park, S. J., Lee, S., Park, S. E. & Kim, S. Inositol pyrophosphates as multifaceted metabolites in the regulation of mammalian signaling networks. *Animal Cells Syst. (Seoul).* **22**, 1–6 (2018).

2. Wilson, M. S. C., Livermore, T. M. & Saiardi, A. Inositol pyrophosphates: between signalling and metabolism. *Biochem. J.* **452**, 369–79 (2013).

3. Saiardi, A., Bhandari, R., Resnick, A. C., Snowman, A. M. & Snyder, S. H. Phosphorylation of Proteins by Inositol Pyrophosphates. *Science (80-.).* **306**, 2101 LP – 2105 (2004).

4. Bhandari, R. *et al.* Protein pyrophosphorylation by inositol pyrophosphates is a posttranslational event. *Proc. Natl. Acad. Sci.* **104**, 15305–15310 (2007).

5. Thota, S. G., Unnikannan, C. P., Thampatty, S. R., Manorama, R. & Bhandari, R. Inositol pyrophosphates regulate RNA polymerase I-mediated rRNA transcription in Saccharomyces cerevisiae. *Biochem. J.* **466**, 105–114 (2015).

6. Chanduri, M. *et al.* Inositol hexakisphosphate kinase 1 (IP6K1) activity is required for cytoplasmic dynein-driven transport. *Biochem. J.* **473**, 3031–3047 (2016).

7. Azevedo, C., Burton, A., Ruiz-Mateos, E., Marsh, M. & Saiardi, A. Inositol pyrophosphate mediated pyrophosphorylation of AP3B1 regulates HIV-1 Gag release. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 21161–6 (2009).

8. Silva, A. M. N., Vitorino, R., Domingues, M. R. M., Spickett, C. M. & Domingues, P. Post-translational Modifications and Mass Spectrometry Detection. *Free Radic. Biol. Med.* **65**, 925–941 (2013).

9. Witze, E. S., Old, W. M., Resing, K. A. & Ahn, N. G. Mapping protein post-translational modifications with mass spectrometry. *Nat. Methods* **4**, 798–806 (2007).

10. Jünger, M. A. & Aebersold, R. Mass spectrometry-driven phosphoproteomics: patterning the systems biology mosaic. *Wiley Interdiscip. Rev. Dev. Biol.* **3**, 83–112 (2014).

11. Marmelstein, A. M., Yates, L. M., Conway, J. H. & Fiedler, D. Chemical Pyrophosphorylation of Functionally Diverse Peptides. *J. Am. Chem. Soc.* **136**, 108–111 (2014).

12. Penkert, M. *et al.* Unambiguous Identification of Serine and Threonine Pyrophosphorylation Using Neutral-Loss-Triggered Electron-Transfer/Higher-Energy Collision Dissociation. *Anal. Chem.* **89**, 3672–3680 (2017).

13. Penkert, M. *et al.* Electron Transfer/Higher Energy Collisional Dissociation of Doubly Charged Peptide Ions: Identification of Labile Protein Phosphorylations. *J. Am. Soc. Mass Spectrom.* **30**, 1578–1585 (2019).

14. Conway, J. H. & Fiedler, D. An Affinity Reagent for the Recognition of Pyrophosphorylated Peptides. *Angew. Chemie - Int. Ed.* **13**, 3941–3945 (2015).

15. Gygi, S. P. The SCX / IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. (2008). doi:10.1038/nprot.2008.150

16. Yates, L. M. & Fiedler, D. Establishing the Stability and Reversibility of Protein Pyrophosphorylation with Synthetic Peptides. *ChemBioChem* **16**, 415–423 (2015).

17. Yang, F., Shen, Y., Camp 2nd, D. G. & Smith, R. D. High-pH reversed-phase chromatography with fraction concatenation for 2D proteomic analysis. *Expert Rev. Proteomics* **9**, 129–134 (2012).

18. Batth, T. S., Francavilla, C. & Olsen, J. V. Off-Line High-pH Reversed-Phase Fractionation for In-Depth Phosphoproteomics. *J. Proteome Res.* **13**, 6176–6186 (2014).

19. Ben-David, Y., Letwin, K., Tannock, L., Bernstein, A. & Pawson, T. A mammalian protein kinase with potential for serine/threonine and tyrosine phosphorylation is related to cell cycle regulators. *EMBO J.* **10**, 317–325 (1991).

20. Fedorov, O. *et al.* Specific CLK Inhibitors from a Novel Chemotype for Regulation of Alternative

Splicing. *Chem. Biol.* **18**, 67–76 (2011).

21. Villén, J. & Gygi, S. P. The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. *Nat. Protoc.* **3**, 1630–1638 (2008).

22. Thingholm, T. E., Jensen, O. N., Robinson, P. J. & Larsen, M. R. SIMAC (Sequential Elution from IMAC), a Phosphoproteomics Strategy for the Rapid Separation of Monophosphorylated from Multiply Phosphorylated Peptides. *Mol. & Cell. Proteomics* **7**, 661 LP – 671 (2008).

23. Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).

24. Marmelstein, A. M. *et al.* Pyrophosphorylation via selective phosphoprotein derivatization. *Chem. Sci.* **9**, 5929–5936 (2018).

25. Gmbh, C. W. V., Kgaa, C., Yates, L. M. & Fiedler, D. Establishing the Stability and Reversibility of Protein Pyrophosphorylation with Synthetic Peptides Table of Contents. (2015).

26. Blanchetot, C., Chagnon, M., Dubé, N., Hallé, M. & Tremblay, M. L. Substrate-trapping techniques in the identification of cellular PTP targets. *Methods* **35**, 44–53 (2005).

27. Trentini, D. B., Fuhrmann, J., Mechtler, K. & Clausen, T. Chasing Phosphoarginine Proteins: Development of a Selective Enrichment Method Using a Phosphatase Trap. *Mol. Cell. Proteomics* **13**, 1953–1964 (2014).

Appendix

Table 4.2*:* Protein pyrophosphorylation sites identified in HEK293T.

	Protein	GN	Sequence (red: PP-site)	PP-Site	Cell Comparment (Uniprot)	Priming Kinase
1	Dual specifity protein kinase 1/4	CLK1/4	VVDFGSATYDDEHHSTLVST RHYR	Ser 341	Nucleus	Auto?
2	MICOS complex subunit	CHCHD3	YSGAYGASVSDEELK	Ser 50	Nucleus, Mitochondria	
3	Elongation factor 1-delta	EEF1D	KPATPAEDDEDDDEDLFGSD NEEEDKEAAQLR	Ser 162	Nucleus	CK2A1
4	Nucleophosmin	NPM1	CGSGPVHISGQHLVAVEEDA ESEDEEEEDVK	Ser 125	Nucleus, Cytoskeleton	ATR, PRP4
5	Heat shock protein HSP 90-alpha	HSP90AA1	ESEDKPEIEDVGSDEEEEKK DGDK	Ser 263	Nucleus, cell membrane, cytoplasm	
6	Nucleoside disphosphate kinase A	NME1	VMLGETNPADSKPGTIR	Thr 94	Nucleus	CDK1
7	Myosin- 9	MYH9	KGAGDGSDEEVDGKADGAE AKPAE	Ser 1943	Cytoskeleton	CK2A1
8/9	Heterogeneous nuclear ribonucleoprotein D0	HNRNPD	IDASKNEEDEGHSNSSPR	Ser 82/ Ser 83	Nucleus	GSK3B
10	Nucleolar RNA helicase 2	DDX21	NEEPSEEEIDAPKPK	Ser 121	Nucleus	
11	Negative Elongation Factor E	NELFE	SISADDDLQESSR	Ser 115	Nucleus	PRP4, TBK1
12	Membrane associated progesterone receptor component 1	PGRMC1	LLKEGEEPTVYSDEEEPKDE SAR	Ser 181	Endoplasmic reticulum	
13	Coiled-coil domain containing protein 168	CCDC168	LENGTSAVTSASEMLLPHTL QNHSVK	Ser 1901	-	

	Protein	GN	Sequence (red: PP-site)	PP-Site	Cell Comparment (Uniprot)	Priming Kinase
14	Calnexin	CANX	QKSDAEEDGGTVSQEEEDR KPK	Ser 564	Endoplasmic reticulum	PLK1
15	Histone Deacetylase 2	HDAC2	MLPHAPGVQMQAIPEDAVHE DSGDEDGEDPDKR	Ser 394	Nucleus	CK2A1
16	Heat shock protein HSP 90-beta	HSP90AB1	IEDVGSDEEDDSGKDKK	Ser 255	Nucleus, cell membrane, cytoplasm,	CK2A1, AurB
17/ 18	Proline-rich protein PRCC	PRCC	IAAPELHKGDSDSEEDEPTK	Ser 157/ Ser159	Nucleus	
19	THUMP domain-containing protein 1	THUMPD1	FTDKDQQPSGSEGEDDDAE AALKK	Ser 86	Nucleus	
20	Plasminogen activator inhibitor 1 RNA-binding protein	SERBP1	SKSEEAHAEDSVMDHHFR	Ser 330	Nucleus, cytoplasm	PRKD1
21	Formin-binding protein 4	FNBP4	IDENSDKEMEVEESPEKIK	Ser 499	Lysosome, cell membrane	
22	Hepatoma-derived growth factor	HDGF	GNAEGSSDEEGKLVIDEPAK EK	Ser 133	Nucleus, cytoplasm	
23	Pyruvate dehydrogenase E1 component subunit alpha	PDHA1	YHGHSMSDPGVSYR	Ser 293	Mitocondron matrix	PDHK1-4, PDK1, PDP1
24	Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	NUCKS1	NSQEDSEDSEDKDVK	Ser 61	Nucleus	
25	Charged multivesicular body protein 2b	CHMP2B	ATISDEEIER	Ser 199	Cytosol	TBK1
26	N-acetyl-D-glucosamine kinase	NAGK	SLGLSLSGGDQEDAGR	Ser 76	Cytosol	AurB, CDK1
27	Nucleolar protein 58	NOP58	HIKEEPLSEEEPCTSTAIASPE KK	Ser 502	Nucleolus, nucleoplasm	PRP4
28	Monocarboxylate transporter 1	SLC16A1	KESKEEETSIDVAGKPNEVTK	Ser 461	Cell membrane	

	Protein	GN	Sequence (red: PP-site)	PP-Site	Cell Comparment (Uniprot)	Priming Kinase
29	Gamma-soluble NSF attachment protein	NAPG	KKSPATPQAKPDGVTATAADEEEDEYSGGLC	Ser 284	Membrane	
30	Nucleosome assembly protein 1-like 1	NAP1L1	LDGLVETPTGYIESLPR	Thr 62	Cytoplasm, nucleus	
31	Phosphoglucomutase-1	PGM1	AIGGIILTASHNPGGPNGDFGIK	Ser 117	Cytoplasm	
32	Microtubule-associated protein tau	MAPT	TDHGAEIVYKSPVVSGDTSPR	Ser 713	Cytoplasm	

Table 4.3: Protein pyrophosphorylation sites identified in HCT116.

	Protein	GN	Sequence (red: PP-site)	PP-Site	Cell Comparment (Uniprot)	Priming Kinase
1	Microtubule-associated protein tau	MAPT	AKTDHGAEIVYKppSPVVSGDTSPR	Ser 713	Cytoplasm	
2	Dual specifity protein kinase CLK1	CLK1	VVDFGSATYDDEHHSTLVppSTR	Ser 341	Nucleus	Auto?
3	Nucleolar Protein 58	NOP58	HIKEEPLppSEEEPCTSTAIASPEK	Ser 502	Nucleus/nucleoplasm	PRP4
4	Nucleoside disphosphate kinase A /B	NME1	VMLGEppTNPADSKPGTIR	Thr 94	Nucleus	CDK1
5	Serine/Arginine-rich splicing factor 5	SRSF5	SVppSRSPVPEK	Ser 231	Nucleus	
6	Ubiquitin-conjugating enzyme E2	UBE2Z	LHNENAEMDSDSSppSSGTETDLHGSLR	Ser 341	Nucleus/cytoplasm	-
7	Activated RNA polymerase II transcriptional coactivator p15	SUB1	SKELVSSSSppSGSDSDSEVDKK	Ser 13	Nucleus	-
8	Nucleolar and coiled-body phosphoprotein 1	NOLC1	KKASppSSDSEDSSEEEEEVQGPPAK	Ser 84	Nucleus/cytoplasm	
9/10	Pre-mRNA 3'-end-processing factor FIP1	FIP1L1	VTETEDDSDppSDSDDDEDDVHVTIGDIK VTETEDDppSDSDSDDDEDDVHVTIGDIK	Ser 85/87	Nucleus	
11	Nucleophosmin	NPM1	CGSGPVHISGQHLVAVEEDAEppSEDEEEEDVK	Ser 125	Nucleus, Cytoskeleton	ATR, PRP4
12	N-acetyl-D-glucosamine kinase	NAGK	SLGLSLppSGGDQEDAGR	Ser 76	Cytosol, secreted	AurB, CDK1

13	MICOS complex subunit MIC19	CHCHD3	SQRYppSGAYGASVSDEELK	Ser 50	Nucleus, Mitochondria
14	Hepatoma-derived growth factor-related protein 2	HDGFL2	APSASDSDppSKADSDGAKP EPVAMAR	Ser 236	Nucleus