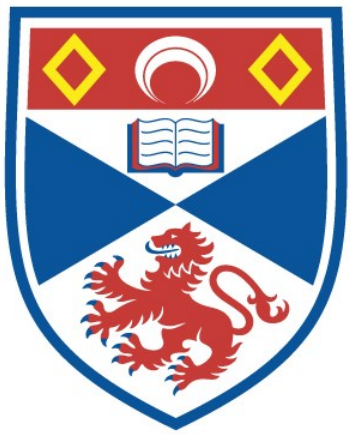
A NOVEL APPROACH TOWARDS THE STEREOSELECTIVE SYNTHESIS OF INOSITOLS AND ITS APPLICATION IN THE SYNTHESIS OF BIOLOGICALLY IMPORTANT MOLECULES

Lloyd Sayer

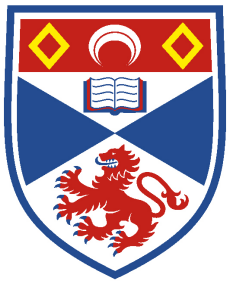
A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



2016

A novel approach towards the stereoselective synthesis of inositols and its application in the synthesis of biologically important molecules

Lloyd Sayer



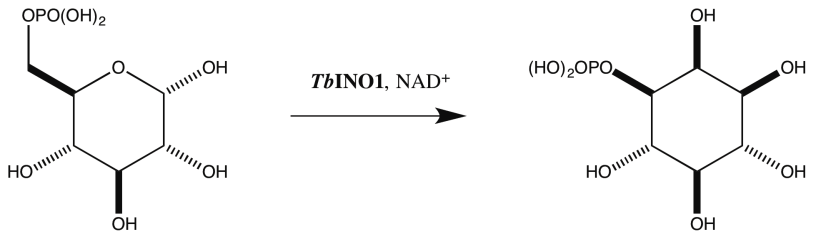
University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of Doctor of Philosophy
at the
University of St Andrews

January 2016

Abstract

myo-Inositol is ubiquitous in nature and is found at the structural core of a diverse range of biologically important derivatives, including phosphatidylinositols, inositol phosphates and mycothiol. The synthesis of *myo*-inositol derivatives is notoriously difficult due to the need to control both regio- and enantioselectivity. As a result, synthetic routes to derivatives of this type are often lengthy and low yielding. The first biosynthetic step in the production of all *myo*-inositol metabolites is the isomerisation of D-glucose 6-phosphate to L-*myo*-inositol 1-phosphate as mediated by L-*myo*-inositol 1-phosphate synthase (INO1). For the protozoan parasite *Trypanosoma brucei*, INO1 is essential for survival and its version of the enzyme (*Tb*INO1) has a high turnover. This makes *Tb*INO1 an attractive candidate for the biocatalytic production of L-*myo*-inositol 1-phosphate, and a potential starting point for drastically shortened syntheses of important *myo*-inositol derivatives.



D-glucose 6-phosphate L-*myo*-inositol 1-phosphate

The production of L-*myo*-inositol 1-phosphate by *Tb*INO1 has been optimised to achieve complete conversion in reaction conditions that facilitate product isolation. Due to problems with an in-batch process, the *Tb*INO1 enzyme was immobilised and the process was transferred to a flow system. This has allowed for production of significant quantities of L-*myo*-inositol 1-phosphate with a high level of purity. L-*myo*-inositol 1-phosphate obtained from the flow system has been used to prepare mycothiol glycosylation acceptor, 1,2,4,5,6-penta-*O*-acetyl-D-*myo*-inositol, in a concise synthesis with a greatly improved yield over the literature.

Declarations

I, Lloyd Henry Sayer, hereby certify that this thesis, which is approximately 29,328 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in October 2011 and as a candidate for the degree of Doctor of Philosophy in September 2012; the higher study for which this is a record was carried out in the University of St Andrews between 2011 and 2015.

Date …… Signature of candidate ……

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Doctor of Philosophy in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date …… Signature of supervisor ……

Date …… Signature of supervisor ……

Copyright Declaration

Restricted access.

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. We have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisors regarding the electronic publication of this thesis:

Access to all or part of printed copy but embargo of all or part of electronic publication of the thesis for a period of 2 years on the grounds that publication may preclude future publication in peer-reviewed journals.

Date …… Signature of candidate ……

Date …… Signature of supervisor ……

Date …… Signature of supervisor ……

Acknowledgements

I would firstly like to thank my supervisors, Dr Gordon Florence and Prof. Terry Smith, for all of their advice, guidance and encouragement throughout my PhD. I am deeply appreciative for everything they have done for me, and the time that they spent proof reading my thesis.

I acknowledge the BBSRC for funding this work, and I would like to thank all of the technical staff at St Andrews who have been fantastic during my time here. Especially, Melanja and Thomas for running an amazing NMR service, and Catherine and Caroline for helping with mass spectrometry.

A massive thank you to all of the members of the Florence and Smith labs, past and present. You have really made me feel at home in St Andrews and I know I've made some life-long friends. Coron, from unorthodox dance moves in the kitchen to the quest for perfection in Mario Kart, living with you has been a pleasure. I don't think I'll look at a White Russian the same way again. Greg, road tripping in America was a personal highlight, even if we did miss a whole 2 floors of the Natural History Museum. Thanks for welcoming me into the lab, and being a friend throughout my time in St Andrews. I owe a big thank you to Barry, Andrew, Stef and Ross for providing me with somewhere to crash for the final few months. Jo, Lindsay and Eoin, your ideas and knowledge were invaluable during this project. A major thank you to Louise for her crash course in protein expression and purification to get me started in the lab. Remember guys, you can't spell 'discovery' without 'disco'!

I'd like to thank Prof. Andy Smith and the members of his group for their thoughts and ideas on my project during group meetings, and Madras FP Hockey Club for providing me with a distraction from my studies.

To my family I am indebted for all your support and encouragement, even if you never really understood what I did. Adam, thank you for the honour of being your best man and for being my big brother. My biggest thanks go to Katie. Her love and support has been endless, and has helped keep me sane whilst writing this thing.

Table of Contents

Abbreviations

<	less than
>	greater than
°C	degrees Celsius
Ac	acetyl
Ac_2O	acetic anhydride
AIBN	azobisisobutyronitrile
All	allyl
app.	apparent
aq.	aqueous
Ar	aromatic
ATP	adenosine triphosphate
ATR	attenuated total reflectance
BEMP	2-*tert*-butylimino-2-diethylamino-1,3-dimethylperhydro-1,3,2-diazaphosphine
Bn	benzyl
Boc	*tert*-butyloxycarbonyl
br. s	broad singlet
Bz	benzoyl
BOM	benzyloxymethyl acetal
c	concentration
Ca^{2+}	calcium ion
cat.	catalytic
CSA	camphor sulfonic acid
cm^{-1}	wavenumbers
δ	delta
d	doublet
DABCO	1,4-diazabicyclo[2.2.2]octane
DAG	diacylglycerol
DAST	diethylaminosulfur trifluoride
DCC	*N*,*N*'-dicyclohexylcarbodiimide
DCM	dichloromethane
dd	doublet of doublets

ddd	doublet of doublets of doublets
DDQ	2,3dichloro-5,6-dicyano-1,4-benzoquinone
DIBAL-H	diisobutylaluminium hydride
DMAP	4-dimethylaminopyridine
DMF	*N*,*N*'-dimethylformamide
DMP	Dess-Martin periodane
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dt	doublet of triplets
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	ethylenediaminetetraacetic acid
ee	enantiomeric excess
equiv.	equivalent
ER	endoplasmic reticulum
ES	electrospray
Et	ethyl
g	grams
GPCR	G protein-coupled receptor
GPI	glycosylphosphatidylinositol
HAT	Human African Trypanosomiasis
HATU	*O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate
His	histidine
HIV	human immunodeficiency virus
hr	hours
HRMS	high resolution mass spectrometry
HOAt	1-hydroxy-7-azabenzo-triazole
HPLC	high performance liquid chromatography
Hz	hertz
^{i}Bu	isobutyl
IMPase	*myo*-inositol monophosphatase
INO1	L-*myo*-inositol 1-phosphate synthase

$InsP_3$	D-*myo*-inositol 1,4,5-trisphosphate
$InsP_3R$	D-*myo*-inositol 1,4,5-trisphosphate receptor
$InsP_5$	*myo*-inositol 1,3,4,5,6-pentakisphosphate
$InsP_6$	*myo*-inositol 1,2,3,4,5,6-hexakisphosphate
^{i}Pr	isopropyl
IR	infrared
J	coupling constant
kDa	kilodaltons
Km	Michaelis-Menten constant
L	litres
m	multiplet
M	mol dm^{-3}
M. tuberculosis	*Mycobacterium tuberculosis*
MALDI	matrix-assisted laser desorption/ionisation
mbar	millibar
MCA	mycothiol *S*-conjugate amidase
mCPBA	*m*-chloroperbenzoic acid
MDR-TB	multi-drug-resistant tuberculosis
Me	methyl
mg	milligrams
Mg^{2+}	magnesium ions
MHz	megahertz
min	minutes
mL	millilitres
mM	millimol dm^{-3}
mmol	millimoles
mol	moles
MshA	mycothiol glycosyltransferase
MshA2	mycothiol phosphatase
MshB	mycothiol deacetylase
MshC	mycothiol ligase
MshD	mycothio synthase
NAD^{+}/H	nicotinamide adenine dinucleotide^{+} / hydrogen
Naph	napthyl

^{n}Bu	*n*-butyl
Ni^{2+}	nickel ion
nm	nanometre
nmol	nanomole
NMR	nuclear magnetic resonance
NTA	nitrilotriacetic acid
PCR	polymerase chain reaction
Pd-C	palladium on carbon
Ph	phenyl
PI	phosphatidylinositol
PI-PLC	phosphatidylinositol phospholipase C
PIPs	phosphatidylinositol phosphates
PLC	phospholipase C
PMB	*para*-methoxybenzyl
ppm	parts per million
pTSA	*para*-toluenesulfonic acid
rcf	relative centrifugal force
RT	room temperature
s	singlet
S. cerevisiae	*Saccharomyces cerevisiae*
sat.	saturated
t	triplet
T. b. brucei	*Trypanosoma brucei brucei*
T. b. gambiense	*Trypanosoma brucei gambiense*
T. b. rhodesiense	*Trypanosoma brucei rhodesiense*
T. brucei	*Trypanosoma brucei*
TB	tuberculosis
TBABr	tetrabutylammonium bromide
TBAF	tetrabutylammonium fluoride
TBAI	tetrabutylammonium iodide
*Tb*INO1	*Trypanosoma brucei* L-*myo*-inositol 1-phosphate synthase
TES	triethylsilyl
Tf	trifluoromethansulfonyl
TFA	trifluoroacetic acid

THF	tetrahydrofuran
TIPS	triisopropylsilyl
TLC	thin-layer chromatography
TMEDA	tetramethylethylenediamine
TMS	tetramethylsilane
Tris	tris(hydroxymethyl)methylammonium
μg	micrograms
μL	microlitres
μM	micromole dm^{-3}
U	μmol min^{-1}
VSG	variant surface glycoprotein
WHO	World Health Organization
wt %	weight percent
XDR-TB	extremely drug-resistant tuberculosis

Chapter One

1 Introduction

1.1 Structure and Nomenclature of Inositols

The inositols are the nine isomeric forms of cyclohexane-1,2,3,4,5,6-hexol (Figure 1.1.1), which are small, polar and chemically very stable molecules with versatile properties.[1] They were discovered in 1850, when Scherer[2] isolated an optically inactive isomer of cyclohexanehexol from muscle. He named the molecule "inosit" and as more isomers were discovered or synthesized, the –ol suffix was added to give the name as it stands today.[3]

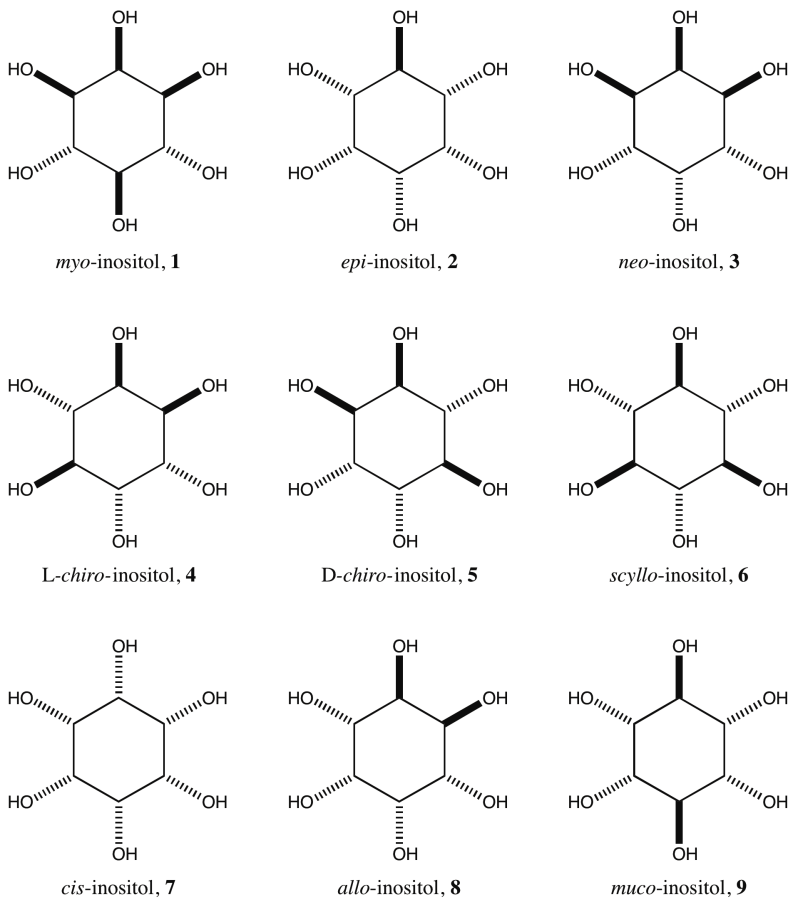
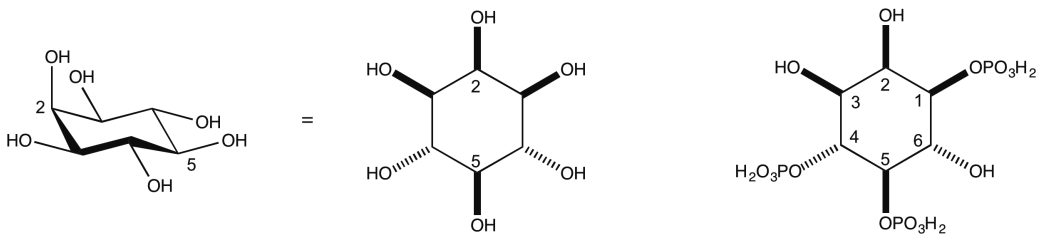


Figure 1.1.1 – The nine isomeric forms of cyclohexane-1,2,3,4,5,6-hexol, or inositol.

Six of the nine isomers have been found to occur within nature, these are *myo*-inositol (**1**), *epi*-inositol (**2**), *neo*-inositol (**3**), L-(+)-*chiro*-inositol (**4**), D-(-)-*chiro*-inositol (**5**), and

scyllo-inositol (**6**). The other three isomers (*cis*-inositol (**7**), *allo*-inositol (**8**) and *muco*-inositol (**9**)) are unnatural synthetic compounds.[1] *myo*-Inositol (**1**) is the most abundant isomer of inositol within nature, and it is accepted that when literature refers to inositol without a prefix it is to be assumed that this is the isomer being described. This isomer is a *meso*-isomer of inositol, as although it contains more than two chiral centres, it can be superimposed upon a mirror image of itself. It possesses a plane of symmetry that runs through the centre of the molecule as it has one axial alcohol and five equatorial alcohols in its lowest energy conformer.[3]

The nomenclature of *myo*-inositol derivatives was outlined by the International Union of Biochemistry Nomenclature Committee in 1989.[4] These rules were put forward in an attempt to address the difficulties in nomenclature that the subtle differences in *myo*-inositol derivatives can produce. Placing the *myo*-inositol ring in its most thermodynamically stable conformation (with one axial and five equatorial substituents, Figure 1.1.2), the ring is numbered by designating C-2 to the carbon bearing the axial substituent. The other carbons within the ring are then numbered from C-1 to C-6 by choosing to label one of the carbons adjacent to C-2 as C-1 and continuing to number in a clockwise or anti-clockwise fashion (directions taken as if looking down onto the molecule from the face containing the axial substituent).[3] Independent of which direction the ring is numbered, the carbon opposite the C-2 position will be labelled as C-5. This means that the plane of symmetry in *myo*-inositol (**1**) passes through C-2 and C-5, and substitution at either of these two positions will not disturb the symmetry of the ring.[5]



1 **10**

Figure 1.1.2 – *myo*-Inositol (**1**) contains a plane of symmetry that passes through C-2 and C-5, whilst D-*myo*-inositol 1,4,5-trisphosphate (**10**) is chiral due to the substitution around the ring.

Substitution at any of the other carbon atoms (C-1, C-3, C-4, or C-6) will disrupt the symmetry of the ring, producing a chiral compound. Therefore, a D- or L- prefix must be given to the name of the molecule in order to account for this. An anti-clockwise numbering is given a D- prefix, whereas numbering in a clockwise fashion gives a L- prefix. The prefix, and thus the direction to number the ring, is usually determined by choosing the numbering that gives the ring substituents the lowest numbering possible.[3] For example, the *myo*-inositol trisphosphate **10** (Figure 1.1.2) is labelled with a D- prefix to give the phosphate groups in the positions with the lowest numbering. It is thus named D-*myo*-inositol 1,4,5-trisphosphate and not L-*myo*-inositol 3,5,6-trisphosphate.

1.2 *myo*-Inositol Derivatives in Nature

Derivatives of *myo*-inositol (**1**) have acquired diverse functions over the course of evolution. Their innate chemical versatility coupled with the stability of the *myo*-inositol core allow these derivatives to play a central role in a plethora of cellular signalling processes and recognition events[6] that take place inter- and intracellularly, also between different cell types.[7] The vast majority of natural *myo*-inositol derivatives are part of the extremely diverse family of signalling molecules comprised of inositol phosphates, diphosphoinositol polyphosphates and phosphatidylinositol phosphates. Variation of biological properties within these biosynthetically interrelated species leads to control over numerous cellular processes.[8] *myo*-Inositol (**1**) can also form part of a larger structure such as a glycosylphosphatidylinositol (GPI) anchor or bioactive natural products such as discoside.

1.2.1 Inositol phosphates

The essence of what makes inositol phosphates so useful within biology is the versatility provided by compounds that are seemingly very simple. It is very easy for a cell to interchange between different inositol phosphates using various kinases and phosphatases, controlling which inositol phosphates are present at any one time. This allows the cell to control the wide-ranging processes that inositol phosphates govern as needed.

Probably the most well known example of an inositol phosphate is D-*myo*-inositol 1,4,5-trisphosphate (**10**, $InsP_3$) as a secondary messenger used within signal transduction and lipid signalling in cells. The role of $InsP_3$ (**10**) within a cell is very much intertwined with calcium signalling.

Modulation of the cytoplasmic concentration of calcium ions (Ca^{2+}) is incredibly important for cellular signalling. Cells receive signals via electrical, hormonal or mechanical stimulation and use a "toolkit" of channels, pumps and cytosolic buffers to control Ca^{2+} levels within the cytoplasm.[9] An increase in Ca^{2+} ion concentration will exert allosteric regulatory effects on many enzymes, which can be reverted back to normal once Ca^{2+} levels have been lowered.[10] The processes that are controlled by this mechanism are as diverse as cell motility, gene transcription, muscle contraction, and exocytosis.[9,11]

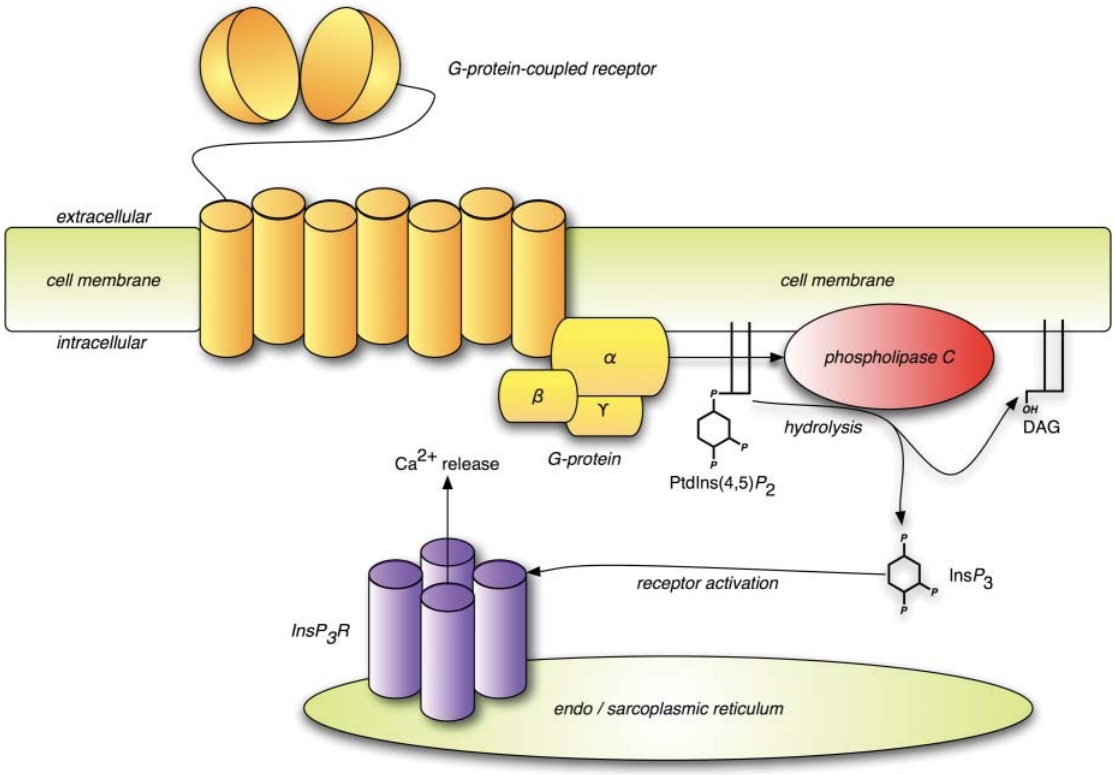
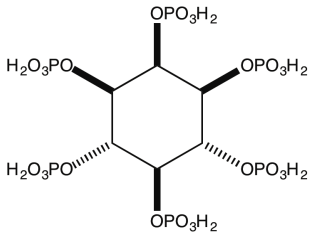


Figure 1.2.1 – Modulation of calcium ion concentration via action of D-*myo*-inositol 1,4,5-trisphosphate (**10**, InP_3).[12]

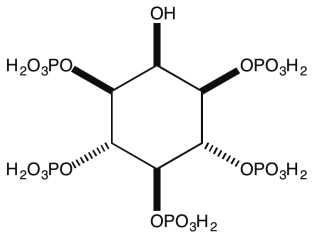
$InsP_3$ (**10**) has been well established as a ubiquitous and essential secondary messenger in Ca^{2+} release. The pathways that link the molecule to Ca^{2+} signalling are generally well

understood. The $InsP_3$ (**10**)-mediated Ca^{2+} signalling cascade (Figure 1.2.1) is initiated by the binding of hormones to specific G-protein coupled receptors (GPCRs) on the cell's outer surface. This leads to the activation of phospholipase C (PLC)-β and PLC-γ, which both act to hydrolyse membrane-bound phosphatidylinositol 4,5-bisphosphate, producing the secondary messengers diacylglycerol (DAG) and $InsP_3$ (**10**).[10] DAG stays tethered within the cell membrane where it can activate protein kinase C enzymes, with the downstream effect of potentially opening voltage-gated Ca^{2+} channels. $InsP_3$ (**10**) diffuses into the cytoplasm and can bind to its receptor. The $InsP_3$ receptor ($InsP_3R$) is located in the endoplasmic reticulum (ER) membrane, and acts as an intracellular ligand-gated Ca^{2+} release channel.[9]

$InsP_3$ (**10**)-mediated Ca^{2+} signals are responsible for a wide variety of roles within a cell. This versatility stems from the complexity that can be involved in how these signals are organised in both time and space, utilising repetitive spikes or oscillations and propagating waves that initiate in specific areas within the cell.[13]



11



12

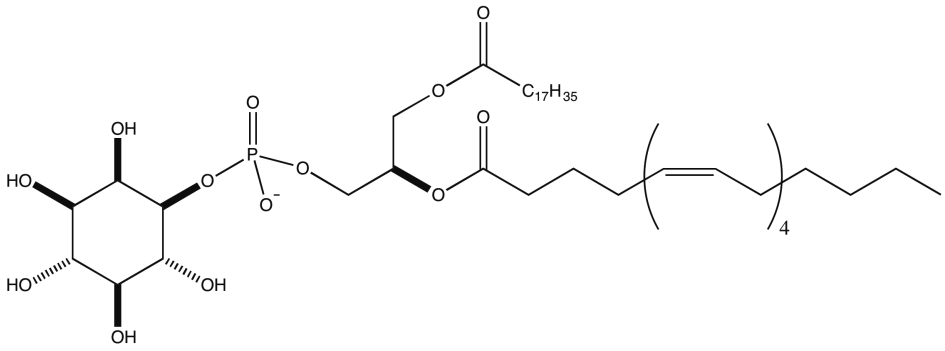
Figure 1.2.2 – The most abundant *myo*-inositol derivatives within cells: *myo*-Inositol 1,2,3,4,5,6-hexakisphosphate (**11**) and *myo*-inositol 1,3,4,5,6-pentakisphosphate (**12**).

Without consideration of pyrophosphate examples, there are 63 possible structural permutations of inositol phosphates, and at least half of these compounds have been identified as metabolites in various biological systems.[8] *myo*-Inositol 1,2,3,4,5,6-hexakisphosphate (**11**, $InsP_6$, phytate, Figure 1.2.2) and *myo*-inositol 1,3,4,5,6-pentakisphosphate (**12**, $InsP_5$) are the most abundant inositol phosphate isomers with cellular concentrations generally maintained between 10 and 100 μM.[8,14] Both of these compounds are very versatile, with $InsP_5$ associated with cell proliferation, viral assembly, chromatin remodelling and regulation of calcium channels. $InsP_6$ has roles within neurotransmission, immune responses, regulation of protein kinases and

phosphatases, and calcium channel activation. It also has antioxidant properties and acts as an anti-neoplastic agent for several different carcinomas.[8]

1.2.2 Phosphatidylinositol and Glycosylphosphatidylinositol

Phosphatidylinositol (**13**, PI, Figure 1.2.3) and phosphatidylinositol phosphates (PIPs) are another biologically important class of *myo*-inositol derivatives. The inositol ring forms the head group for the structure, with attachment to the glycerophospholipid through a phosphodiester linkage at the D-1 position. There are seven known examples of PIPs within nature, with phosphate decoration at the 3-, 4- or 5- positions of the *myo*-inositol ring.[8]



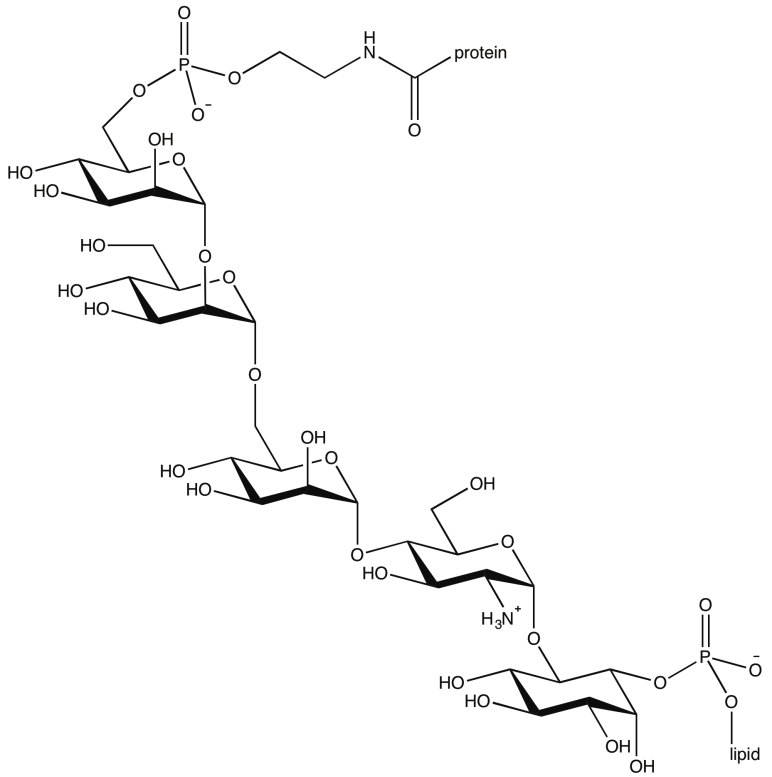
PI, **13**

Figure 1.2.3 – Structure of phosphatidylinositol (**13**, PI).

The lipid portion of PIs and PIPs anchor these compounds within the lipid bilayer of a cell membrane, presenting the *myo*-inositol headgroup at the membrane surface. Therefore, when a protein binds to a headgroup of one of these compounds, it will be reversibly tethered to the cell surface.[8] Interactions between proteins and PIPs regulate many crucial cellular processes through the regulation of protein function and subcellular localisation. Control is exerted by either direct conformational changes to the protein upon binding or the protein being led to interaction with binding partners situated on the membrane surface.[8,15]

A glycosylphosphatidylinositol (GPI) anchor (**14**, Figure 1.2.4) is a glycolipid that is added posttranslationally to the C-terminus of a protein to tether it to the cell membrane. The proteins involved are very diverse in function, playing important roles within signal

transduction, prion disease pathogenesis, immune response and the pathobiology of trypanosomal parasites.[16] The structure of a GPI anchor (**14**) is highly complex with three domains: a phosphoethanolamine linker, glycan core and phospholipid tail.[17]



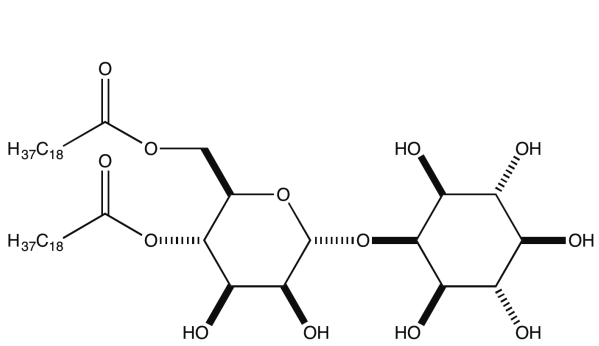
14

Figure 1.2.4 – General structure of a glycosylphosphatidylinositol (GPI) anchor (**14**).

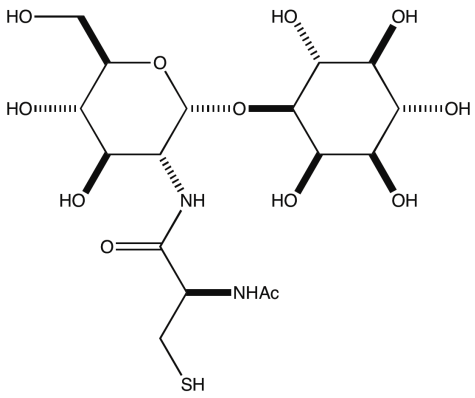
The phosphoethanolamine linker is used to covalently bind the protein to the third mannose of the glycan core, which is a highly conserved structure comprising three mannose, one glucosamine and one inositol ring. Whilst the core of the glycan is highly conserved, a plethora of modifications is possible, including additional phosphoethanolamine groups or sugars, such as mannose, galactose or sialic acid.[16] The transition between the glycan core and phospholipid is achieved via a PI moiety. The lipid moiety can be a diacylglycerol, alkylacylglycerol or ceramide depending on protein and species of origin.[18] An additional fatty acid is present on the OH-2 of the *myo*-inositol ring in many GPI anchors, which makes them resistant to the enzyme phosphoinositide phospholipase C (PI-PLC) that normally cleaves the head group from PI and PIPs.[16,18]

1.2.3 Bioactive Natural Products

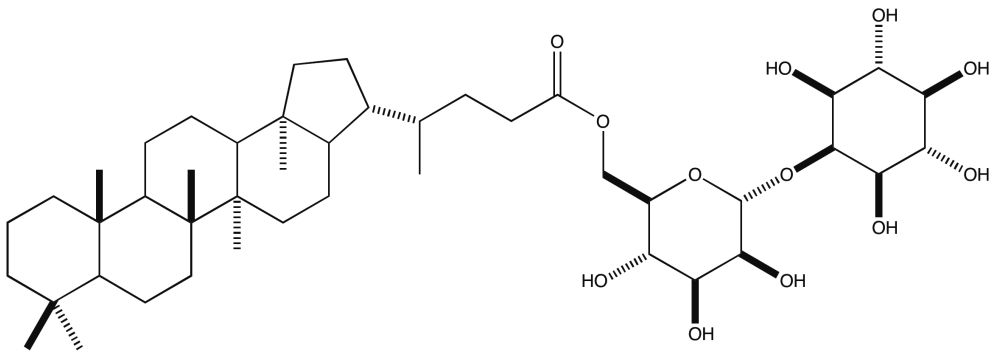
The roles discussed so far have shown the importance of *myo*-inositol derivatives within the cell, and how they can be used to control a large number of cellular processes. There are other examples in the literature of *myo*-inositol derivatives that have been isolated in small quantities from nature. Discoside[19] (**15**, Figure 1.2.5) and Plakohopanoid[20] (**17**) are natural products that have been isolated from the marine sponges *Discodermia dissoluta* and *Plakortis simplex*, respectively. Both contain a mannose attached to the *myo*-inositol via the OH-2. Mycothiol[21] (**16**) is a low molecular weight thiol that has a high prevalence within mycobacteria, where it is used to relieve oxidative stress within cells, and will be discussed in more detail in Chapter 4.



Discoside, **15**



Mycothiol, **16**

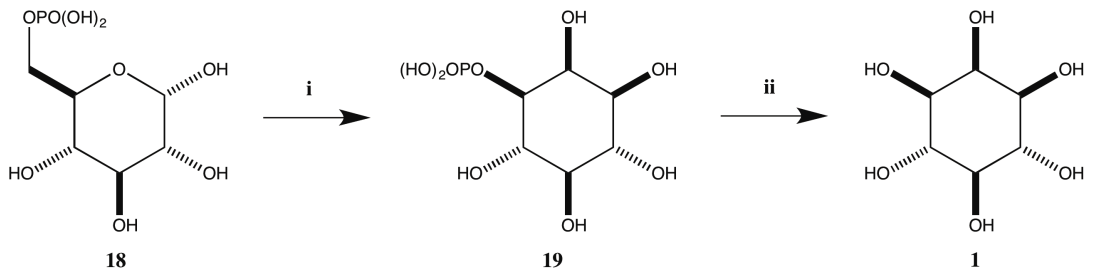


Plakohopanoid, **17**

Figure 1.2.5 – *myo*-Inositol containing natural products Discoside (**15**), Mycothiol (**16**) and Plakohopanoid (**17**).

1.3 Biosynthesis of *myo*-Inositol

Cells can obtain *myo*-inositol (**1**) from the extracellular environment, the turnover of inositol phospholipids and dephosphorylation of inositol phosphates, or it can be *de novo* synthesised.[22] The *de novo* synthesis (Scheme 1.3.1) utilises two enzymes: L-*myo*-inositol 1-phosphate synthase (INO1) and *myo*-inositol monophosphatase (IMPase). INO1 acts to convert D-glucose 6-phosphate (**18**) into L-*myo*-inositol 1-phosphate (**19**) in a three step process that utilises nicotinamide adenine dinucleotide (NAD^+) as a cofactor.[23] IMPase is a specific phosphatase that acts to cleave the phosphate from L-*myo*-inositol 1-phosphate (**19**) to produce *myo*-inositol (**1**).[3] Either of these two compounds, L-*myo*-inositol 1-phosphate (**19**) and *myo*-inositol, can then be modified by the organism to produce compounds of need or used without further elaboration.



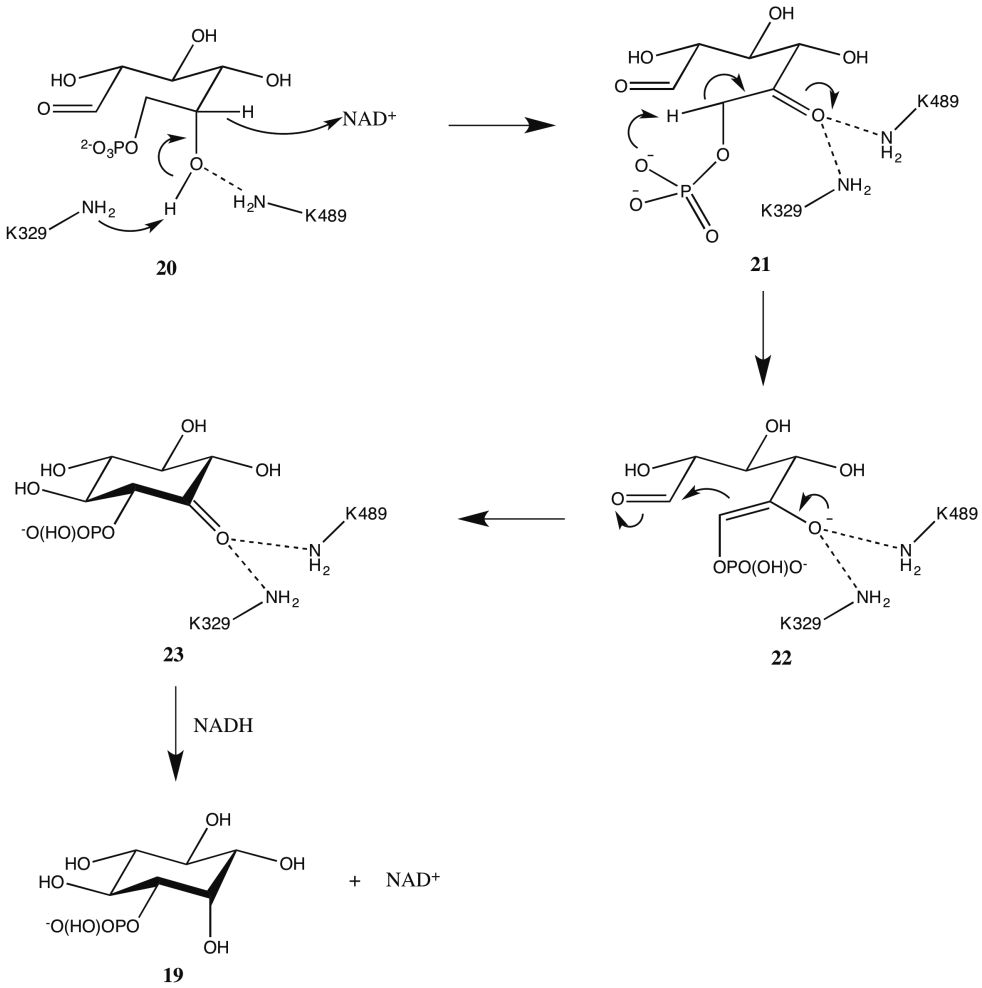
Scheme 1.3.1 – Biosynthesis of *myo*-inositol (**1**). *Reagents and conditions:* **i**. INO1, NAD^+; **ii**. IMPase, Mg^{2+}.

INO1 has been identified in a wide variety of evolutionarily diverse organisms including Cyanobacteria, eubacteria, Archaebacteria, plants, fungi and animals.[24,25] The 'core catalytic structure' of the enzyme is highly conserved throughout the biological kingdom, highlighting the essential role within cellular metabolism.[24]

1.3.1 L-*myo*-Inositol 1-phosphate Synthase Mechanism

It had been hypothesised that the *myo*-inositol core is formed via a cyclisation of D-glucose. However, it was not until 1962[26] that isotopic labelling of the 1 and 6 carbons of D-glucose proved that they condensed to form *myo*-inositol (**1**). The most studied INO1 gene is that of *Saccharomyces cerevisiae* (*S. cerevisiae*).[25] Scheme 1.3.2 shows the

mechanism of this enzyme. The highly conserved structure of the catalytic pocket across the biological kingdoms suggests that this is the general mechanism for all INO1 genes.



Scheme 1.3.2 – Proposed mechanism of *S. cerevisiae* L-*myo*-inositol 1-phosphate synthase.[25]

Studies have shown that the enzyme only binds the acyclic form of D-glucose 6-phosphate (**20**), even though this form represents less than 1 % of D-glucose 6-phosphate molecules.[25] All INO1 proteins within eukaryotes contain a Rossmann fold, which is characterised by a GXGGXXG motif.[24] This is typical of an oxidoreductase, and functions to bind the cofactor, NAD^+, more tightly.

Once the acyclic D-glucose 6-phosphate (**20**) has been bound, a lysine residue will deprotonate the alcohol at C-5. This leads to the formation of keto intermediate **21** through hydride addition to NAD^+. The enol **22** is then formed via deprotonation at C-6.

It is possible that the phosphate oxygen will act as a base for this step (as shown), however it is also possible that it could be performed by a second lysine residue. pKa values of these two groups would not normally be considered within the range to perform this extraction, however phosphate ester-mediated deprotonation has been observed in dehydroquinate synthase,[27,28] and lysines that are buried within enzymes can have dramatically different pKa's than those expected.[25,29,30] An intramolecular aldol cyclisation between C-1 and C-6 forms the key carbon-carbon bond to give inosose **23**, before selective hydride transfer from NADH produces the product, L-*myo*-inositol 1-phosphate (**19**).

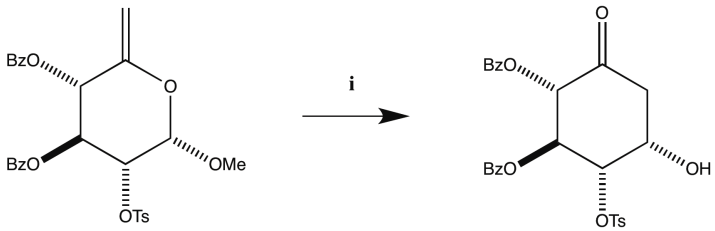
1.4 Synthetic Routes to *myo*-Inositol Derivatives

Synthetic routes to *myo*-inositol derivatives are notoriously difficult. Not only are there complex regioselectivity issues associated with the six hydroxyl groups, but most derivatives are also chiral. There are four main methods used to produce *myo*-inositol derivatives. Syntheses can start from a carbohydrate such as D-glucose, D-xylose or D-mannitol, and utilise some type of rearrangement to produce the *myo*-inositol core, generally via a Ferrier carbocyclisation. Alternatively, a D-pinitol or L-quebrachitol derivative can be converted into a *myo*-inositol derivative by inversion of the necessary stereocentre. Starting from the microbial oxidation of benzene, it is possible to produce *myo*-inositol derivatives through careful control of epoxidations and subsequent ring openings. It is also possible to synthesise derivatives from *myo*-inositol (**1**) itself, with various protection strategies and optical resolutions available for use.

1.4.1 Ferrier Transformations

A popular approach to obtain *myo*-inositol derivatives is to build up functionality on a non-inositol skeleton and then transform the compound into the *myo*-inositol core via a carbocyclisation or dihydroxylation reaction. Methods to do this involve examples of Ferrier type-II reactions,[31] ring closing metathesis[32] and a pinacol type rearrangement.[33] When utilising these methods, it is possible to use protecting group strategies and orders of protection that are different from those associated with syntheses starting from *myo*-inositol (**1**) itself.

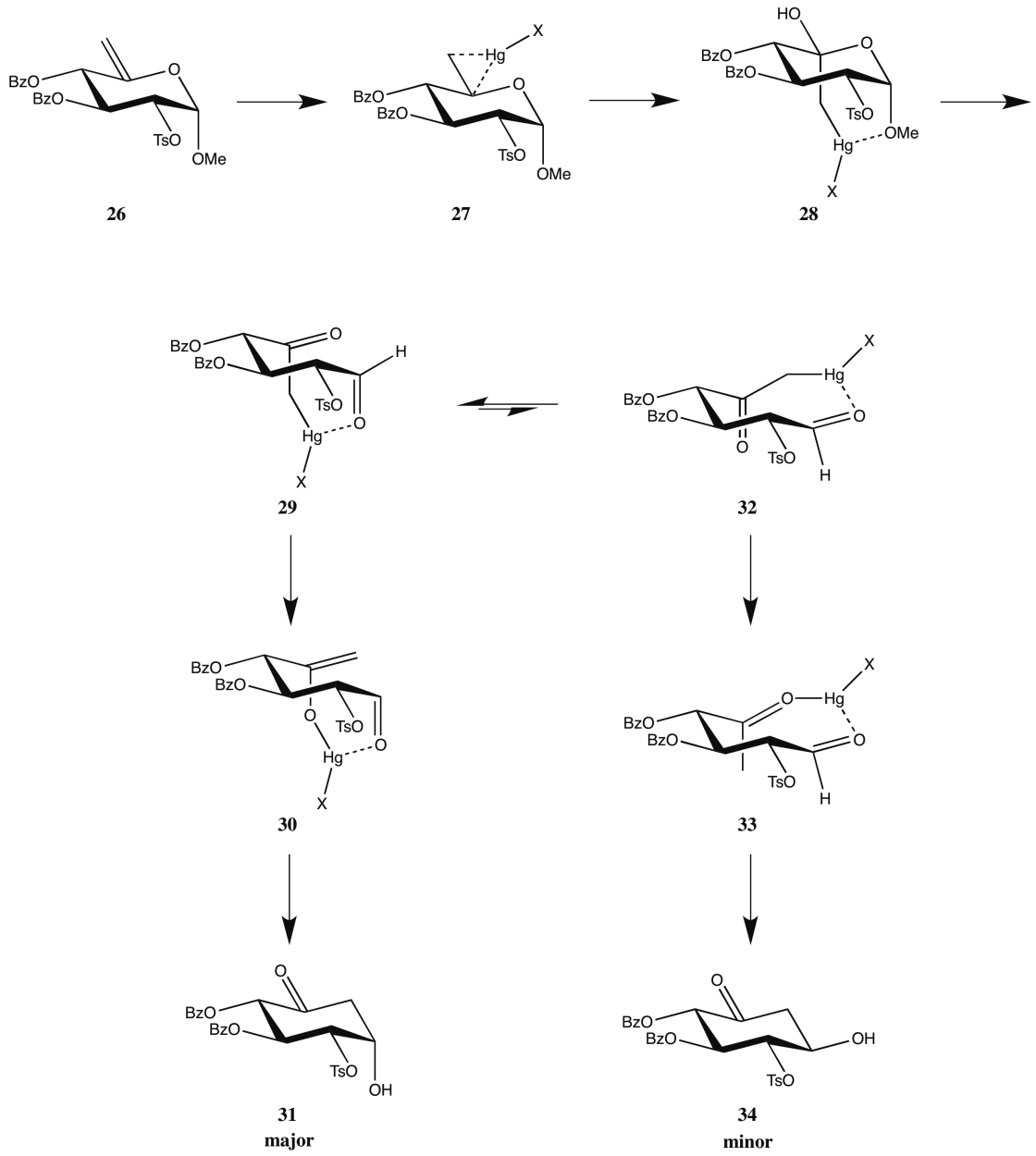
The most popular methods within this approach are based upon the Ferrier-II reaction (also known as the Ferrier carbocyclisation).[31] This reaction provides an efficient, one-step conversion of 5,6-unsaturated hexopyranose derivatives into functionalised cyclohexanones. The cyclohexanones produced are convenient for the synthesis of enantiopure inositols and their amino-, deoxy-, unsaturated and selectively *O*-substituted derivatives.[34] Ferrier[31] developed the reaction when he was trying to synthesise cyclohexane derivatives via a biomimetic approach. As detailed previously, the INO1 enzyme converts D-glucose 6-phosphate (**18**) into L-*myo*-inositol 1-phosphate (**19**) via a mechanism that involves an intramolecular aldol cyclisation to form an inosose.[25] This was the key step that Ferrier emulated in his original work, producing cyclohexanone **25** (Scheme 1.4.1) from pyranoside **24** in a mercury-mediated ring closure.[31] Reaction did not proceed in the absence of mercury ions, confirming that the mercury ions were responsible for transformation and not simply the acidic reaction conditions.



24 **25**

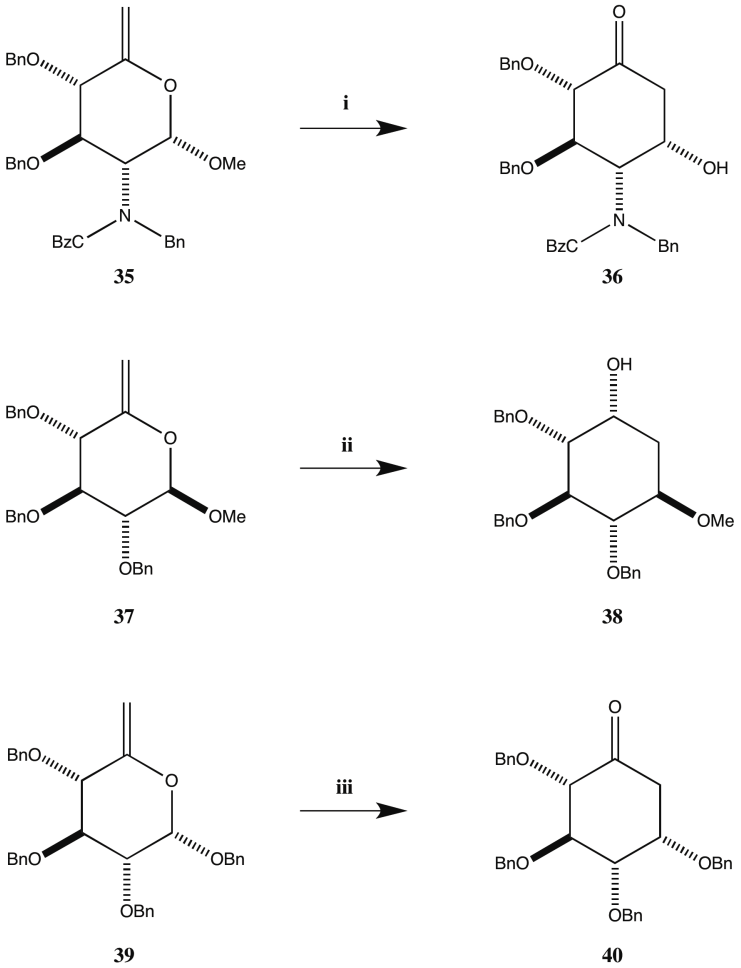
Scheme 1.4.1 – Ferrier-II reaction as originally conducted by Ferrier in 1979.[31] *Reagents and conditions:* **i**. $HgCl_2$, acetone:water (2:1), 4.5 hr, reflux, 83 %.

There have been several modifications and improvements to the procedure since it was originally published. Mercuric (II) acetate improved the yield of the system to 89 %, and also allowed for a larger scope than mercuric (II) chloride.[35] Scope was increased even further by the use of mercuric (II) trifluoroacetate to convert azide-containing pyranosides into the corresponding inososes.[36] It was shown[37] that mercuric (II) sulfate could be used sub-stoichiometrically to produce the same result, and further to this, work by Ogawa[38] demonstrated the catalytic use of mercuric (II) trifluoroacetate and mercuric (II) oxide in a dramatic improvement to the methodology. With his synthesis of Hygromycin A[39] and (+)-Lycoricidine[40], Ogawa showed that as little as 1 mol % mercuric (II) trifluoroacetate could be used to effect Ferrier-II reactions.



Scheme 1.4.2 – Proposed mechanism for the mercury-mediated Ferrier-II reaction.[37]

The most well accepted mechanism for the Ferrier-II reaction was proposed by Machado *et al.* (Scheme 1.4.2).[37] Taking deoxyhex-5-enopyranoside **26** as an example, the proposed mechanism begins with oxymercuration of the alkene. This leads to reaction intermediate **28**, where the hydroxyl group has been introduced in a Markovnikov fashion. Ring-opening is accomplished through loss of methanol, to furnish ketoaldehydes **29** and **32**. The equilibrium between these two compounds is thought to lie far in the favour of **29**, giving cyclohexanone **31** as the major product for this reaction. It is proposed that **31** is formed via the pseudo chair intermediate **30** in an intramolecular aldol-type reaction.



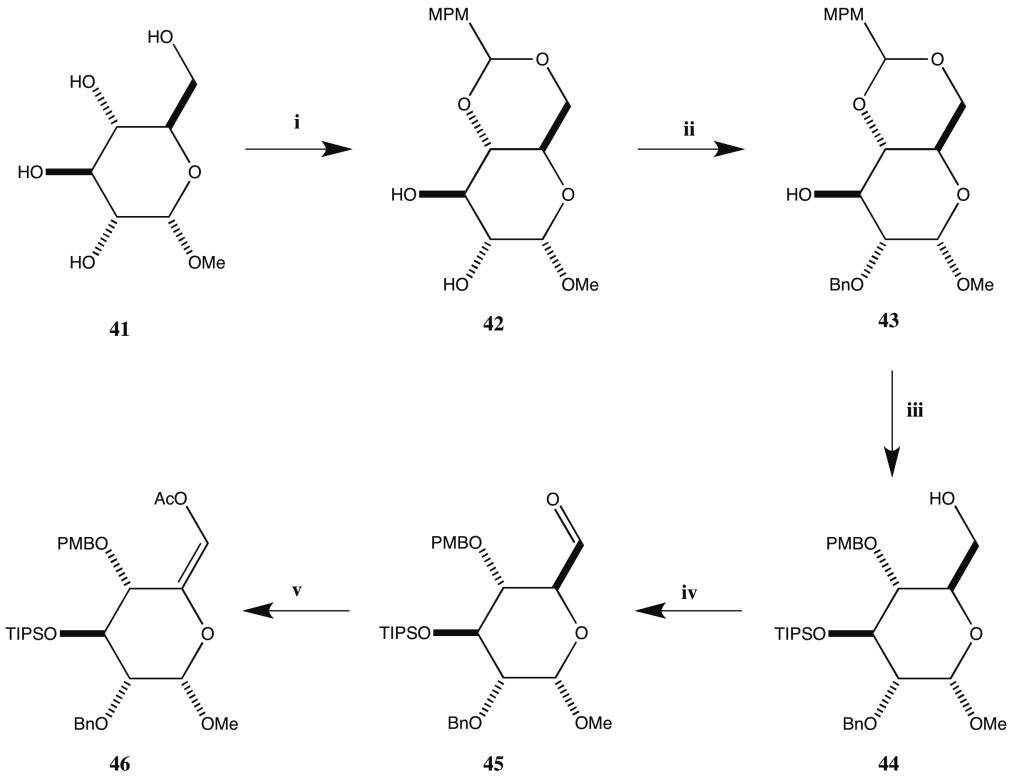
Scheme 1.4.3 – Examples of analogous Ferrier-II reactions utilising palladium-, aluminium- and titanium-mediated reactions. *Reagents and conditions:* **i**. $Pd(OAc)_2$ or $PdCl_2$ (10-20 mol %), 1,4-dioxane or acetone, H_2SO_4, 60 °C, 60-80 %;[41] **ii**. $^{i}Bu_3Al$, toluene, 40 °C, 70 %;[42] **iii**. $Ti(O^{i}Pr)Cl_3$, DCM, -78 °C, 98 %.[43]

The Ferrier-II reaction is not limited to the use of mercury to facilitate the transformation (Scheme 1.4.3). Work by Adam *et al.*[41] discovered that the transformation occurs rapidly when using palladium (II) salts in the presence of acid. The reaction is thought to proceed via a different mechanism, however in an analogous manner to the mercury (II)-mediated process. Further work by Ikegami *et al.*[44] utilised the palladium-mediated reaction in studies towards all nine isomers of inositol. Their work showed that palladium (II) chloride could be used in as little as 5 mol % to effect the transformations of 6-deoxypyranosides in dioxane-water mixtures.[45] Palladium-mediated reactions were generally higher yielding than those mediated by mercury. On the other hand, there was also a higher number of isomeric products observed in the palladium-mediated reactions, the majority of which could be separated by chromatography.[46]

It has also been shown that other metals can facilitate Ferrier-type transformations. Use of triisobutylaluminium ($^{i}Bu_3Al$),[42] and isopropoxytitanium trichloride ($Ti(O^{i}Pr)Cl_3$)[43] both complete the transformation with retention of the glycosidic bond and the major product maintains the original configuration at the anomeric position (Scheme 1.4.3). Use of trimethylaluminium[47] in place of triisobutylaluminium introduces an additional methyl group at the C-2 equatorial position of alcohol **38**.

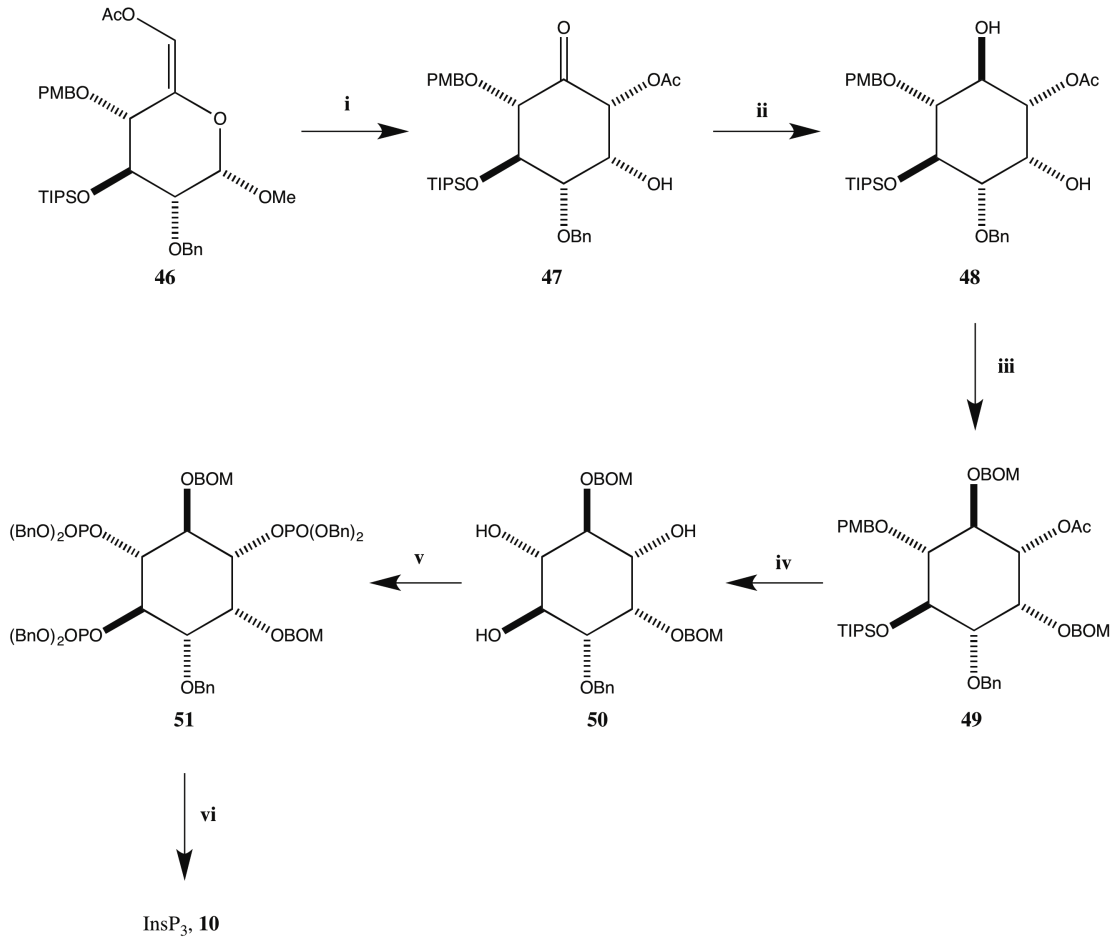
The Ferrier-II reaction has proved to be a useful tool in the synthesis of *myo*-inositol derivatives. Conway *et al.*[48] utilised this method in their synthesis of $InsP_3$ (**10**). They synthesised the Ferrier carbocyclisation precursor **46** (Scheme 1.4.4) starting from α–D-glucopyranose **41**, which they initially selectively protected at the 4- and 6-positions to give anisylidene acetal **42**. Selective benzylation arising through the tin acetal gave the desired alcohol **43** in a 2:1 ratio in over the unwanted regioisomer. Silyl protection and subsequent diisobutylaluminium hydride (DIBAL-H) mediated cleavage of the anisylidene acetal gave primary alcohol **44**. Oxidation using Dess-Martin periodane (DMP) gave aldehyde **45**, which was subsequently converted into the enol acetate **46** using potassium carbonate and acetic anhydride (Ac_2O).[48]

They initially attempted the Ferrier carbocyclisation using palladium (II) chloride, but discovered that they produced a large number of inseparable isomeric products. Treatment of enol acetate **46** (Scheme 1.4.5) with mercuric (II) acetate gave a cleaner reaction, producing the desired inososе **47** in a fairly low yield, however with minimal isomeric side products.



Scheme 1.4.4 – Synthesis of Ferrier carbocyclisation precursor **46**. *Reagents and conditions:* **i**. Anisaldehyde dimethyl acetal, amberlyst-15, DMF, 200 mbar, 80 °C, 58 %; **ii**. $^{n}Bu_2SnO$, TBABr, BnBr, MeCN, reflux, 54 %; **iii**. **a**. Et_3N, TIPSOTf, DCM, 92 %; **b**. DIBAL-H, DCM, 88 %; **iv**. DMP, DCM, 82 %; **v**. K_2CO_3, Ac_2O, DMAP, MeCN, reflux, 65 %.[48]

The reduction of inososе **47** was conducted using tetramethylammonium triacetoxyborohydride and gave a good yield of desired diol **48**. They attribute the selectivity in this reaction to the coordination of the reducing agent to the axial OH-2 of the inososе **47**. This leads to delivery of the hydride from the top face of the molecule. Now that they had created the *myo*-inositol core of the molecule, a series of protecting group manipulations afforded them the triol **50**, containing the necessary symmetry for production of $InsP_3$ (**10**). Phosphitylation of the triol **50**, followed by oxidation using m-chloroperbenzoic acid (mCPBA), furnished the perbenzylated trisphosphate **51**, which was taken through to the desired $InsP_3$ (**10**) via hydrogenolysis.[48]

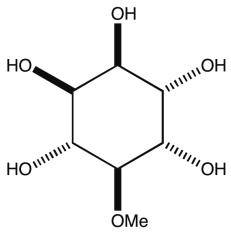


Scheme 1.4.5 – Completion of the synthesis of $InsP_3$ (**10**). *Reagents and conditions:* **i**. $Hg(OAc)_2$, acetone, water, 35 %; **ii**. Tetramethylammonium triacetoxyborohydride, acetic acid, MeCN, 89 %; **iii**. BOMCl, Hünig's base, 85 °C, 82 %; **iv**. **a**. LiOH, MeOH, THF, 97 %; **b**. DDQ, DCM, water, 95 %; **c**. TBAF, THF, DCM, 90 %; **v**. **a**. $(BnO)_2PN(^iPr)_2$, 1*H*-tetrazole, MeCN, DCM; **b**. mCPBA, -78 °C to RT, 93 % over 2 steps; **vi**. H_2, *tert*-butanol, water, Palladium black, $NaHCO_3$, 95 %.[48]

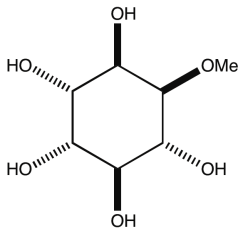
Whilst the Ferrier-II reaction is a useful tool for synthesis, it faces criticism for several reasons. Firstly, it is still necessary to synthesise a highly functionalised and enantiopure 5,6-unsaturated hexopyranose derivative, which is not a trivial task. Converting the produced inosose into the desired *myo*-inositol derivative will also require careful manipulation of protection groups, meaning that the completed syntheses are generally still very lengthy and tedious. Mercury is a very toxic element, and its use will generally be avoided if possible. The palladium-mediated reaction, whilst safer and higher yielding, gives no guarantee of separation from isomeric side products from the reaction.

1.4.2 Routes from D-Pinitol and L-Quebrachitol

D-Pinitol (**52**, Figure 1.4.1) and L-quebrachitol (**53**) are both optically active cyclitol natural products.[49] They are methyl ethers of D-*chiro*-inositol and L-*chiro*-inositol respectively, and have been used as building blocks for the synthesis of polyhydroxylated natural products and their derivatives.[50,51] Within the synthesis of *myo*-inositol derivatives,[49,52] the optical purity of these compounds gives them an advantage as no chemical or enzymatic resolution is needed to produce optically pure products. Only a single inversion of stereochemistry is necessary to furnish the *myo*-inositol core from either of the *chiro*-inositols.[50]



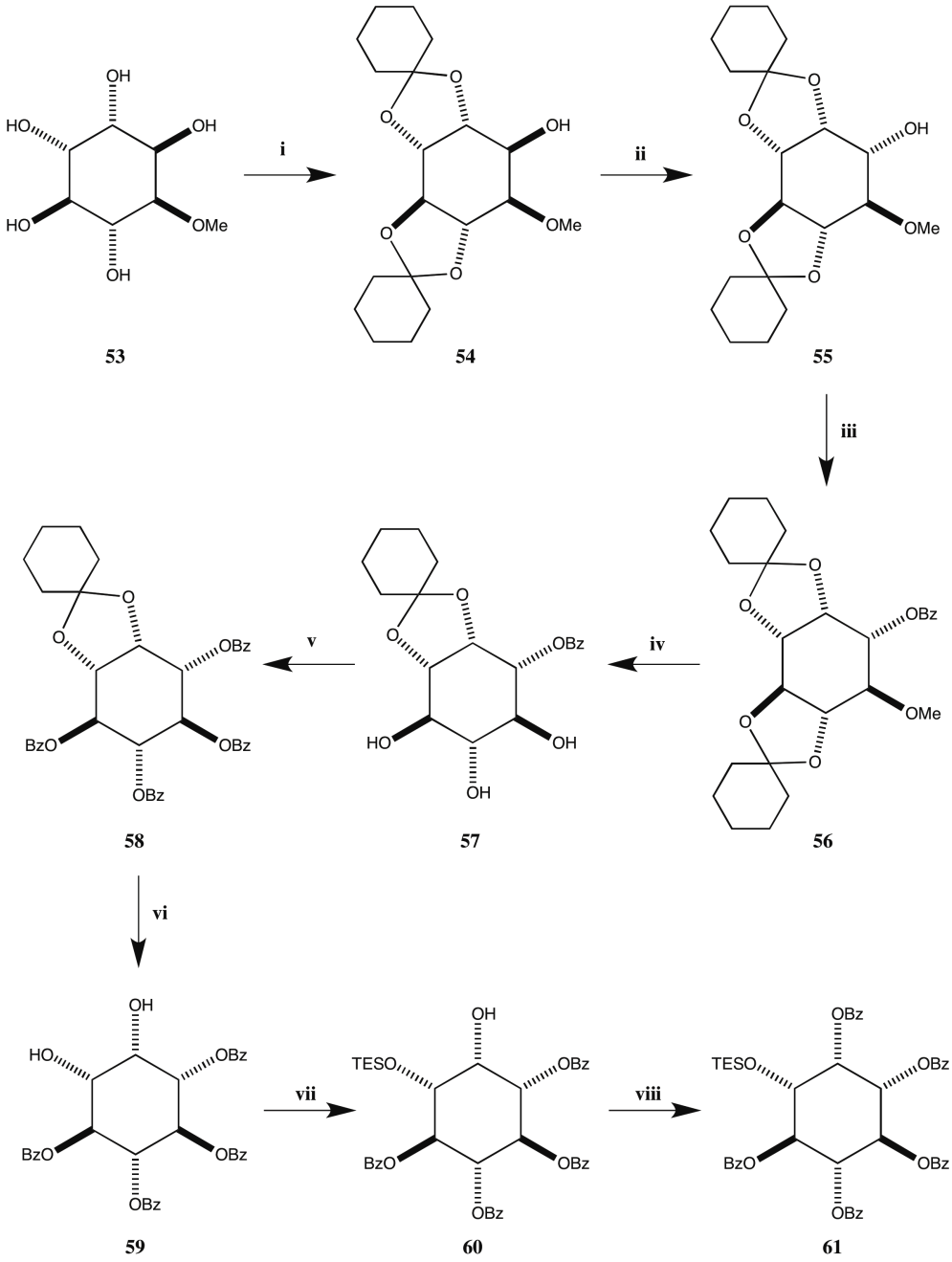
D-Pinitol, **52**



L-Quebrachitol, **53**

Figure 1.4.1 – D-Pinitol (**52**) and L-quebrachitol (**53**).

Ozaki *et al.*[53] used L-quebrachitol (**53**) as a starting material in their synthesis of D-*myo*-inositol 1-phosphate (Scheme 1.4.6). They furnished the final product in 14 steps, in an overall yield of 31 %. L-Quebrachitol (**53**) was treated with cyclohexanone and *para*-toluenesulfonic acid (pTSA) in benzene to afford the L-*chiro*-inositol **54**. Oxidation using Ac_2O and dimethyl sulfoxide (DMSO) furnished a ketone, which could be subsequently stereoselectively reduced using lithium borohydride ($LiBH_4$) to provide *myo*-inositol **55**. Protection with benzoyl chloride (BzCl) and triethylamine (Et_3N) gave benzoyl ester **56**, which was treated with aluminium chloride and sodium iodide to selectively cleave the *trans*-cyclohexylidene and methyl ether to furnish triol **57**. Benzoylation was accomplished before acid hydrolysis of the *cis*-cyclohexylidene gave diol **59**. Subsequent treatment with triethylsilyl chloride (TESCl) in the presence of pyridine exclusively gave the silyl ether **60**. Benzoylation furnished the pentakisbenzoylate **61**, which could be transformed in 5 steps to the desired D-*myo*-inositol 1-phosphate.[53]



Scheme 1.4.6 – Synthetic route to pentakisbenzoylate (**61**), an intermediate in the synthesis of D-*myo*-inositol 1-phosphate from L-quebrachitol (**53**) by Ozaki *et al.*.[53] Reagents and conditions: **i**. cyclohexanone, pTSA, benzene, reflux, 70 %; **ii**. Ac_2O, DMSO, DCM, reflux, 100 %; **iii**. $LiBH_4$, THF, -78 °C, 92 %; **iv**. BzCl, Et_3N, DMAP, DCM, 0 °C, 88 %; **v**. $AlCl_3$-NaI, MeCN, 83 %; **vi**. **a**. BzCl, Et_3N, DMAP, DCM, 99 %; **b**. TFA, MeOH, 96 %; **vii**. TESCl, pyridine, 0 °C, 100 %; **viii**. BzCl, Et_3N, DMAP, DCM, 93 %.

The main strength of synthetic routes to *myo*-inositol derivatives starting from D-pinitol (**52**) or L-quebrachitol (**53**) is that they bypass the need for a chemical or enzymatic resolution to achieve an optically pure product, allowing for a higher overall yield. The long and tedious synthetic routes associated with the synthesis of *myo*-inositol derivatives are still necessary, along with the detailed planning needed to consider the reactivity of each protecting group being utilised. The synthetic routes used do not readily lend themselves to the synthesis of analogues or similar *myo*-inositol derivatives.

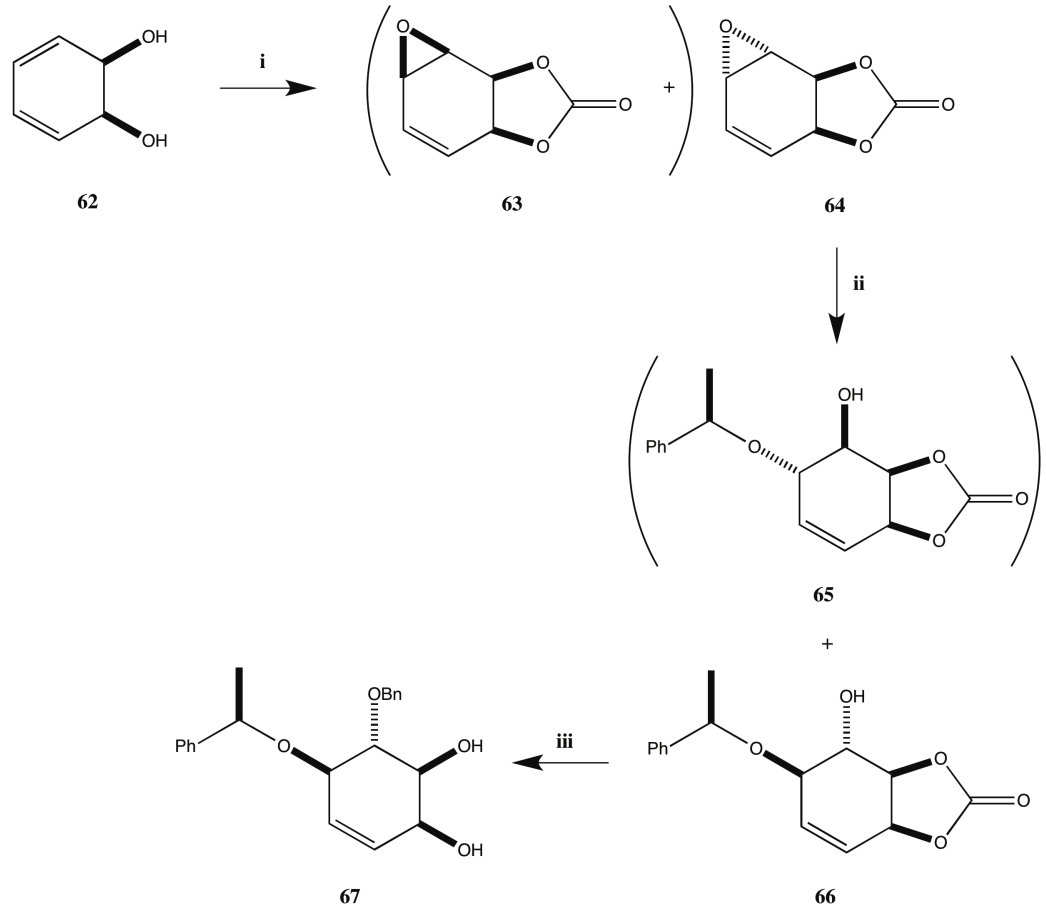
1.4.3 *myo*-Inositol Derivatives from Benzene

Ley *et al.*[54] produced syntheses of $InsP_3$ (**10**), racemic *myo*-inositol 1-phosphate (±**19**) and other *myo*-inositol derivatives, based upon the oxidation of benzene by Gram-negative soil bacterium *Pseudomonas putida*.[55] *cis*-Cyclohexadiene-1,2-diol (**62**, Scheme 1.4.7), the product of the microbial oxidation, was utilised as the starting point for these syntheses that introduced the six stereogenic centres sequentially. The two hydroxyl groups present in **62** would eventually become hydroxyl groups OH-2 and OH-3 within the derivatives synthesised.

Their synthesis of $InsP_3$ (**10**) is completed with an overall yield of 3.9 % over 12 steps starting from *cis*-cyclohexadiene-1,2-diol (**62**, Scheme 1.4.7). It was necessary to use a chiral auxiliary in order to achieve optical purity in the final product. They began their synthesis with the formation of the cyclic carbonate and subsequent epoxidation using mCPBA to give α- and β-epoxides (**64** and **63** respectively) in a 4.6:1 ratio. Epoxide **64** was regioselectively ring-opened using (*R*)-1-phenylethanol, derived from (*S*)-(+)-mandelic acid, giving a separable pair of diastereomeric alcohols (**65** and **66**) in a 1:1 ratio. Alcohol **66** was taken on to form diol **67** via benzylation and subsequent carbonate group removal.

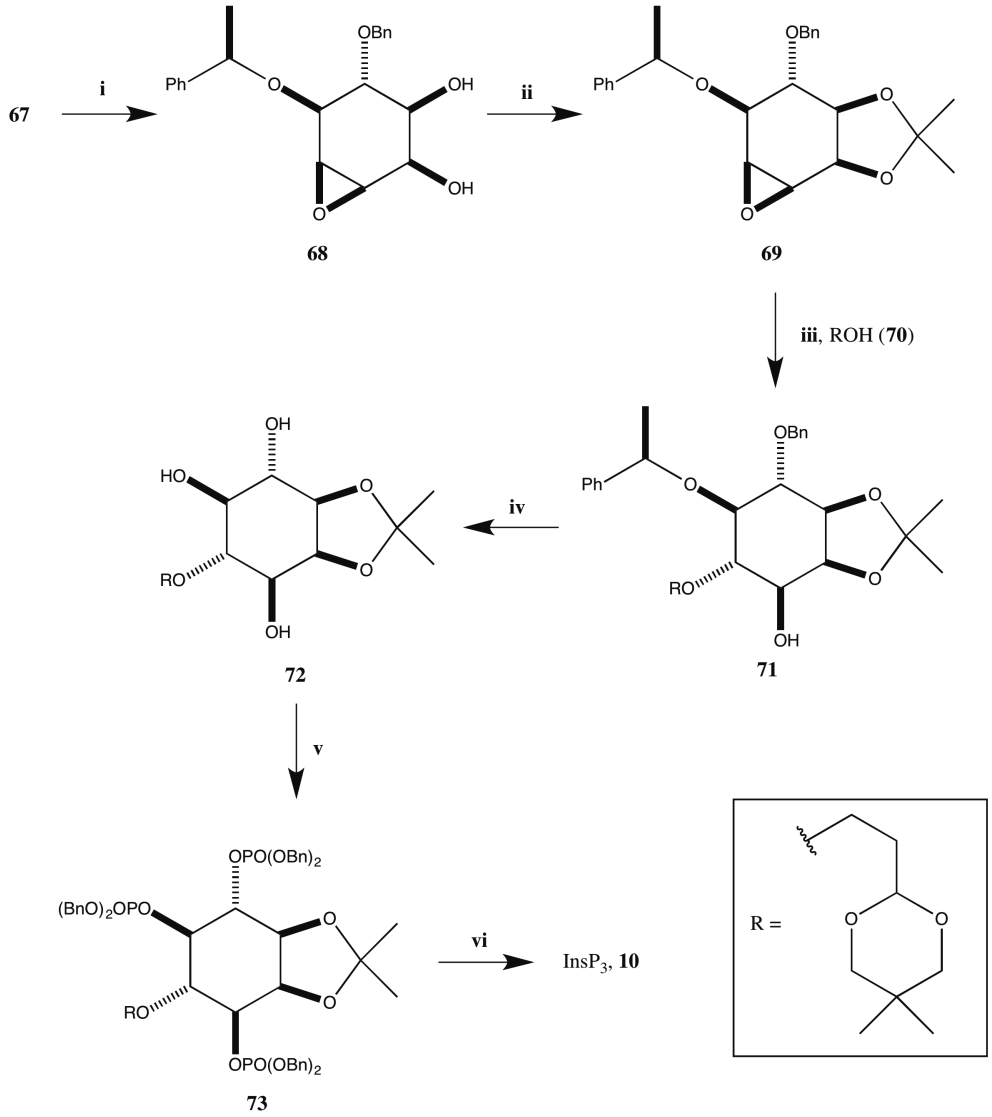
The diol **67** was treated with mCPBA to form almost exclusively β-epoxide **68** (Scheme 1.4.8), which was subsequently converted to the acetinide **69** using 2,2-dimethoxypropane. The epoxide of **69** was opened in a facial-selective manner using alcohol **70**. This gave rise to a mixture of *myo*- and *muco*-inositol rings, from which the *myo*-inositol **71** was taken on. Removal of the chiral auxiliary and the benzyl group via hydrogenolysis afforded the 1,4,5-triol **72**. Phosphorylation gave the protected

trisphosphate **73**, which, after global deprotection, furnished the desired optically pure $InsP_3$ (**10**).[54]



Scheme 1.4.7 – Synthesis of D-*myo*-inositol 1,4,5-trisphosphate (**10**, $InsP_3$) by Ley *et al.* (part 1).[54] *Reagents and conditions:* **i**. **a**. Dimethyl carbonate, NaOMe, MeOH; **b**. mCPBA, DCM, 47 % over 2 steps; **ii**. (*R*)-1-phenylethanol, $HBF_4.OEt_2$ (cat.), DCM, 67 % (racemic); **iii**. **a**. BnBr, Ag_2O, DMF, 100 %; **b**. Et_3N, MeOH, water, 3 days 99 %.

Ley *et al.*[54] have shown that it is possible to set and control each position of the *myo*-inositol ring through careful stereocontrol of epoxidations and their subsequent ring openings. The introduction of a chiral auxiliary allows the separation of diasteromeric alcohols **65** and **66**. However, this drastically lowers the yield for the synthesis, as the diastereomers are present in a 1:1 ratio. Further to this, due to the restricted arrangement of protection groups within this synthesis, the route does not easily facilitate the synthesis of analogues. Work by Trost[56] and Altenbach[57] has utilised the dihydroxylation concepts proposed within this synthesis to produce $InsP_3$ from Conduritol B derivatives.

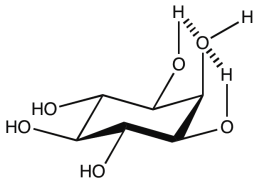


Scheme 1.4.8 – Synthesis of D-*myo*-inositol 1,4,5-trisphosphate (**10**, $InsP_3$) by Ley *et al.* (part 2).[54] *Reagents and conditions:* **i**. mCPBA, DCM, 87 %; **ii**. 2,2-Dimethoxypropane, CSA, DCM, 89 %; **iii**. Alcohol **70**, NaH, TMEDA, 110 °C, 72 hr, 58 %; **iv**. H_2, 10 % Pd-C, EtOH, 16 hr, 100 %; **v**. ^{n}BuLi, $^{i}Pr_2NH$, tetrabenzylpyrophosphate, THF, -30 °C to RT, 67 %; **vi**. **a**. H_2, 10 % Pd-C, EtOH, 48 hr; **b**. TFA (80 %, aq.), 4 hr, 88 % over 2 steps.

1.4.4 Routes from *myo*-Inositol

The synthesis of *myo*-inositol derivatives starting from *myo*-inositol (**1**) itself is probably the most favoured route. *myo*-Inositol (**1**) is inexpensive and readily available, making it an ideal starting point in synthesis. There are now several methods available to effect the desymmetrisation of *myo*-inositol derivatives and the resolution of racemic derivatives, enabling the synthesis of enantiomerically pure products.[50]

Within the literature, there has been a remarkable push towards the synthesis of partially protected *myo*-inositol derivatives containing free hydroxyl groups at specific positions. This work has allowed a deep understanding of the chemistry and reactivity of *myo*-inositol (**1**) and its derivatives. Researchers have been able to produce various selectivity patterns by carefully controlling different factors within their syntheses. The synthetic routes involved rely upon the inherent acidity of the free hydroxyl groups and their interactions with neighbouring functional groups (generally through hydrogen bonding), the protecting groups that are present, the conformation of the inositol ring, and the reaction conditions and reagents used. The reason for the observed selectivity is not always obvious, and there is some debate over how some methods produce their selectivity.[50]



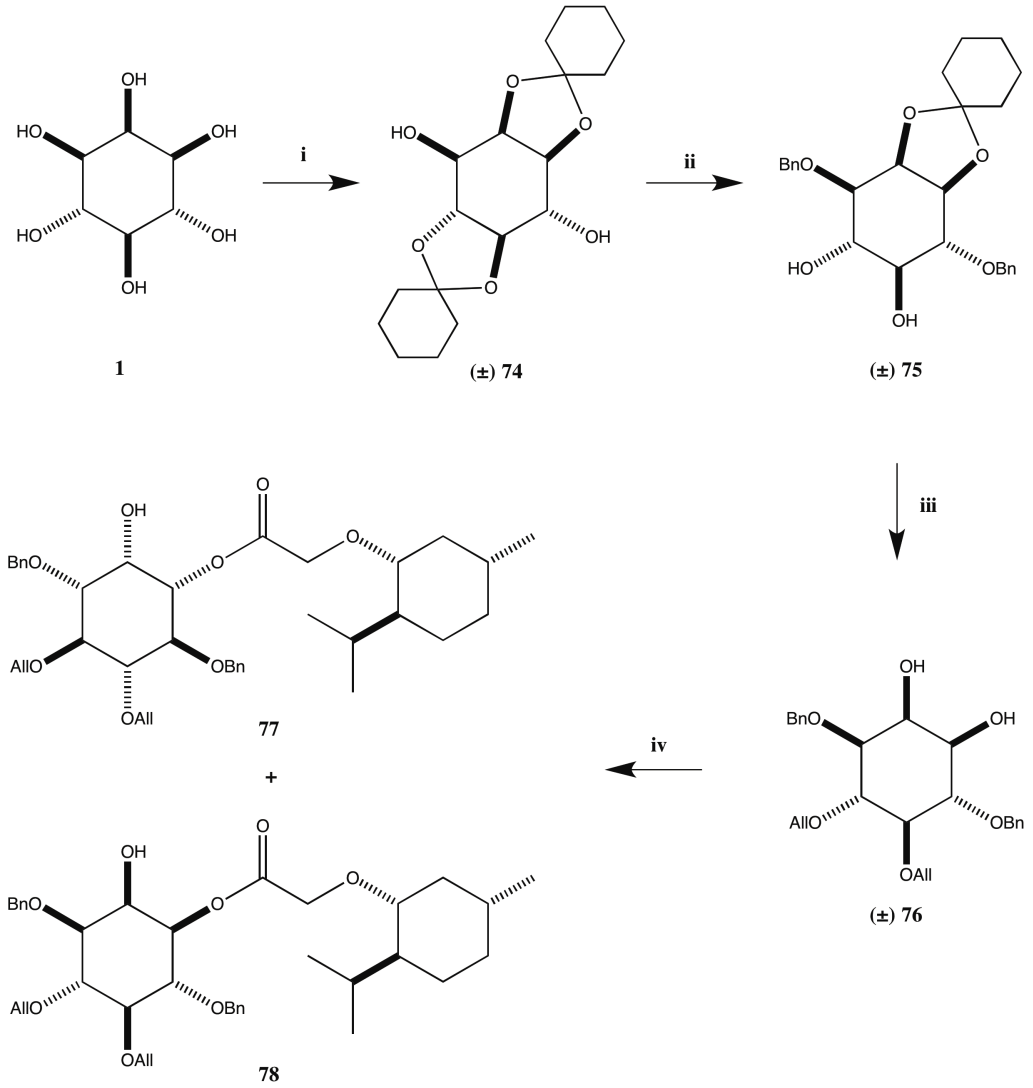
1

Figure 1.4.2 – Hydrogen bonding between *cis*-vicinal hydroxyls within *myo*-inositol (**1**).

From work conducted by Gigg *et al.*,[58] the order of reactivity for the hydroxyl groups around the *myo*-inositol (**1**) ring is OH-1 ≈ OH-3 > OH-4 > OH-5. The higher reactivity of hydroxyl groups OH-1 and OH-3 can be attributed to the hydrogen bonding interaction between the proton of the reacting hydroxyl group and the *cis*-vicinal oxygen of OH-2 (Figure 1.4.2).[59,60]

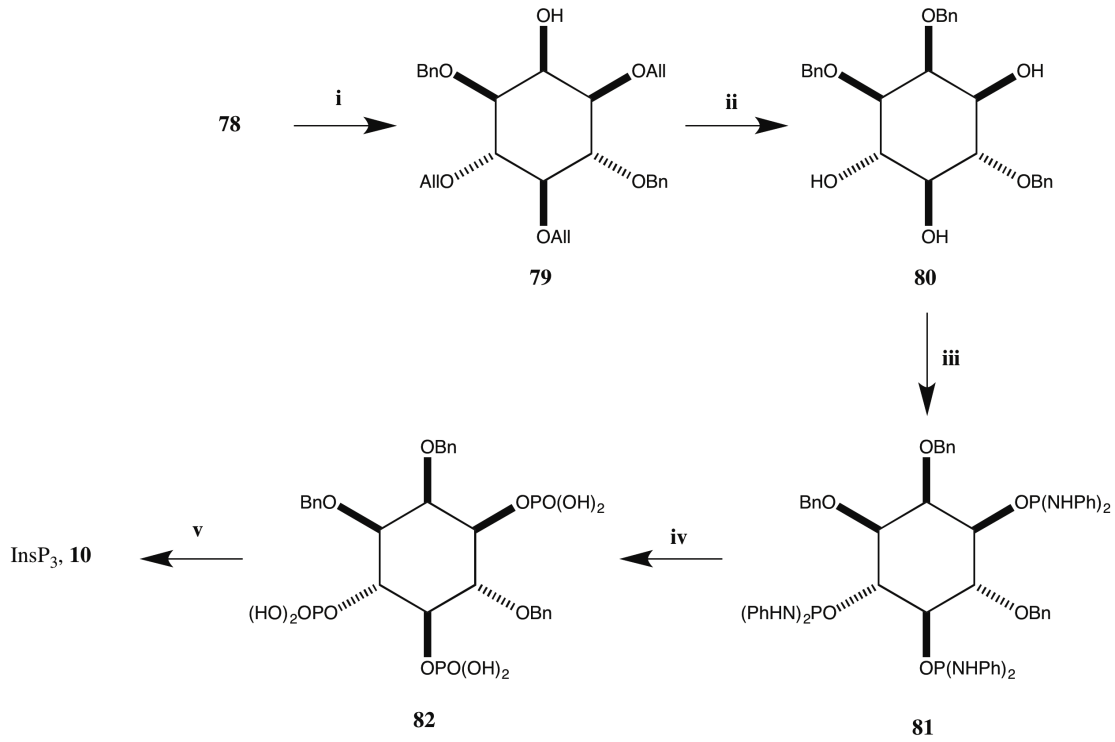
1.4.4.1 Routes Using Ketals

One of the first *myo*-inositol (**1**) protection strategies was the formation of ketals from acetone,[61] cyclohexanone or cyclopentanone.[62] Since then, this method has become widely used in the synthesis of *myo*-inositol derivatives. An example of ketals being utilised in this way comes from Ozaki *et al.*,[63] in what was the first synthesis of $InsP_3$ (**10**) in 1986 (Schemes 1.4.9 and 1.4.10).



Scheme 1.4.9 – Ozaki *et al.*'s synthesis of D-*myo*-inositol 1,4,5-trisphosphate (**10**, $InsP_3$) starting from *myo*-inositol (**1**, part 1).[63] *Reagents and conditions:* **i**. 1-Ethoxycyclohexane, pTSA, DMF, reflux, 26 %; **ii**. **a**. BnCl, NaH, DMF, 90 %; **b**. ethylene glycol, pTsOH, chloroform, 80 %; **iii**. **a**. AllBr, NaH, DMF, 100 %; **b**. AcOH, water, 90 °C, 88 %; **iv**. 1-menthoxyacetylchloride, pyridine, 39 % of **78**.

Starting from *myo*-inositol (**1**), they formed the racemic bisketal (±**) 74** through reflux in DMF with 1-ethoxcyclohexene and catalytic pTSA. Subsequent perbenzylation and selective 4,5-trans ketal cleavage[62] produced the diol (±**) 75**. Allylation of both hydroxyl groups was followed by removal of the *cis*-ketal to furnish the racemic diol (±**) 76**. Addition of a chiral auxiliary and formation of diastereoisomers **77** and **78**, allowed for separation by fractional crystallisation and column chromatography to produce single diastereoisomer **78** in 39 % yield.



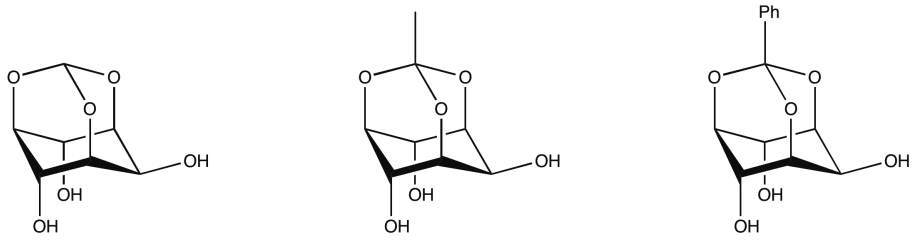
Scheme 1.4.10 – Ozaki *et al.*'s synthesis of D-*myo*-inositol 1,4,5-trisphosphate (**10**, $InsP_3$) starting from *myo*-inositol (**1**, part 2).[63] *Reagents and conditions:* **i**. **a**. NaOH, water, MeOH, 98 %; **b**. AllBr, NaOH, benzene, reflux, 76 %; **ii**. **a**. BnCl, NaH, DMF, 98 %; **b**. $RhCl(PPh_3)_3$, DABCO, then HCl, MeOH, 58 %; **iii**. $(PhNH)_2P(O)Cl$, DMAP, pyridine, 41 %; **iv**. isoamylnitrite, AcOH, pyridine, Ac_2O; **v**. 5 % Pd-C, H_2, MeOH, water.

The chiral auxiliary was removed by alkaline methanolysis, before selective allylation produced alcohol **79** (Scheme 1.4.10). Benzylation and subsequent removal of allyl groups using Wilkinson's catalyst and acidic methanolysis, furnished triol **80** which could be converted to $InsP_3$ in three steps.[63]

Unfortunately, their report on this synthesis did not cite yields for the final two steps as they had not been optimised. The overall yield of the protected trisphosphate **81** was only 1.1 % and, in total, their synthesis was 13 steps. The yield of the synthesis was dramatically reduced not only by the resolution step, but also by the poor yield in the first step of the synthesis.

1.4.4.2 Routes Using Orthoesters

A popular method of modifying *myo*-inositol (**1**) to produce biologically important derivatives is via their orthoesters. Lee *et al.*[64] were the first to use this method in their synthesis of enterobactin analogues and since then it has been used increasingly within *myo*-inositol derivative synthesis. Orthoformates, orthoacetates and orthobenzoates (**83**, **84** and **85**, respectively, Figure 1.4.3) are some of the most commonly used orthoesters for this purpose.

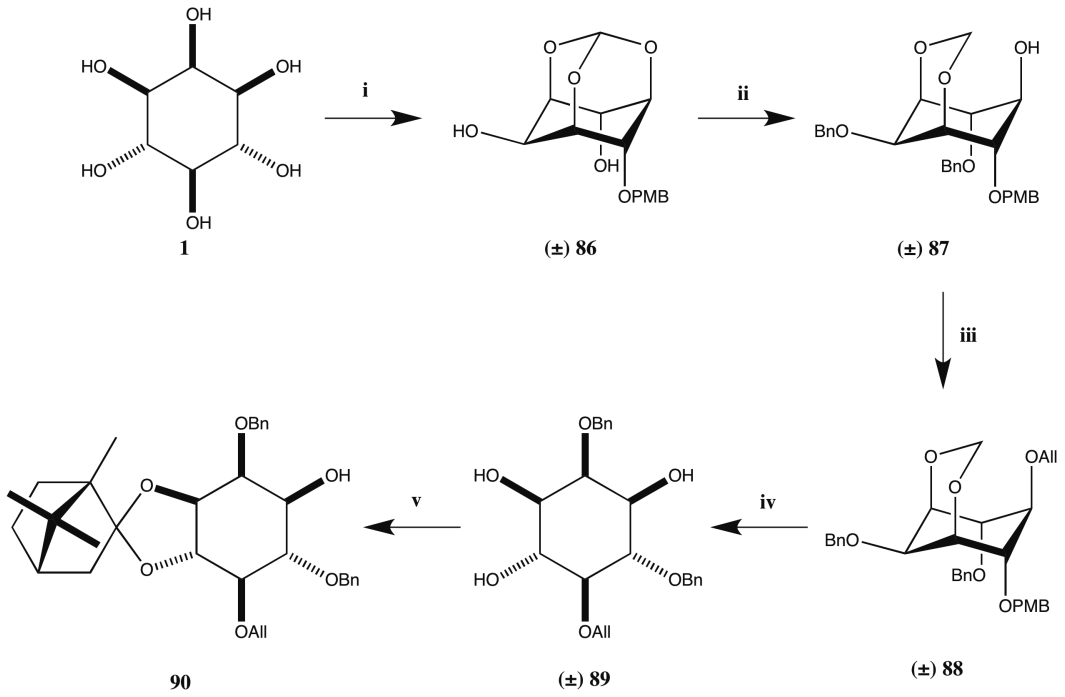


83 **84** **85**

Figure 1.4.3 – Simultaneous protection of OH-1, OH-3 and OH-5 using orthoesters.

Billington *et al.*[65] identified this method as a way to avoid the low-yielding 1,2;4,5-bisketal formation undertaken by Ozaki *et al.*.[63] Orthoesters of this type also allow for the simultaneous protection of OH-1, OH-3 and OH-5, whilst inverting the axial/equatorial relationship of the remaining free hydroxyl groups. They produced syntheses using this method for *myo*-inositol 2-phosphate, *myo*-inositol 1,3-bisphosphate, *myo*-inositol 1,3,4,5-tetrakisphoshate and *myo*-inositol 4-phosphate. The syntheses were considerably higher yielding than those published previously, however, for the latter two syntheses, the racemic product was produced. *myo*-Inositol 2-phosphate and *myo*-inositol 1,3-bisphosphate are both meso compounds so this was not an issue.

As previously seen in the enantioselective synthesis of *myo*-inositol derivatives starting from *myo*-inositol (**1**), it is typically necessary to conduct a resolution. Holmes *et al.*[66,67] provided a synthetic route to enantiopure alcohol (**90**, Scheme 1.4.11), which they used as a common intermediate in the production of a range of inositol polyphosphates.

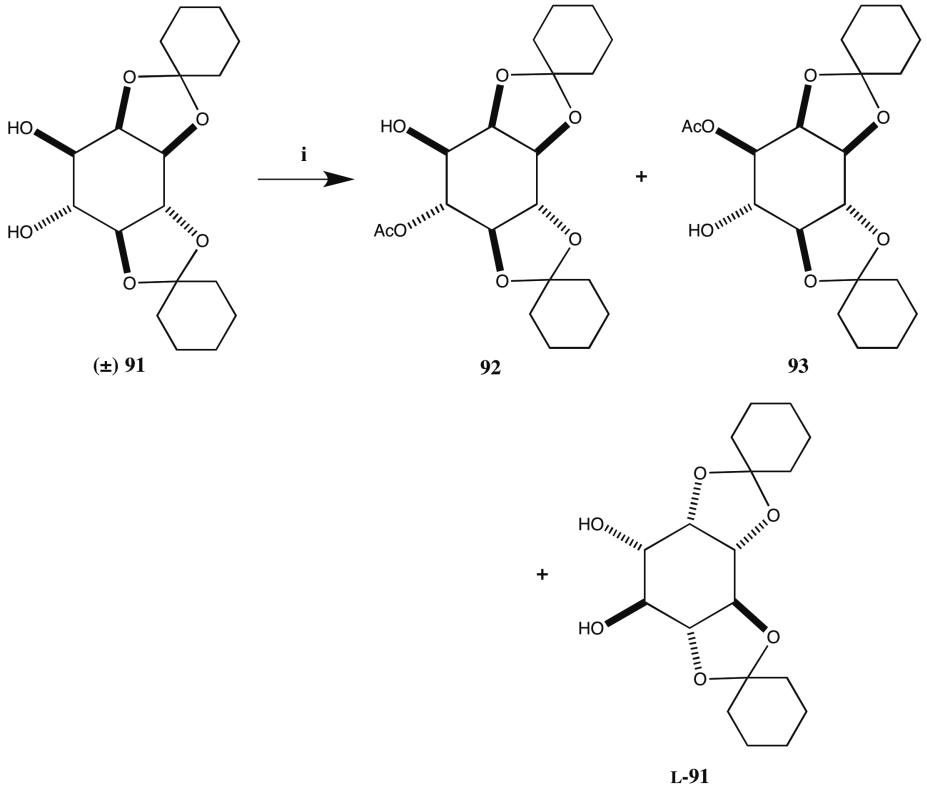


Scheme 1.4.11 – Holmes *et al*.'s route to enantiopure synthesis intermediate alcohol **90**.[67] *Reagents and conditions:* **i**. **a**. triethylorthoformate, pTSA, DMF, 100 °C, 63 %; **b**. PMBCl, NaH, DMF, 67 %; **ii**. **a**. BnBr, NaH, DMF, 99 %; **b**. DIBAL-H, DCM, 96 %; **iii**. AllBr, NaH, imidazole, DMF, 85 %; **iv**. HCl, MeOH, reflux, 78 %; **v**. 1-(*S*)-(-)-camphor dimethylacetal, pTSA, DCM, reflux, 31 %.

The synthesis of alcohol **90** is relatively long considering that this is an intermediate to be used in synthesis and not the product itself. Starting from *myo*-inositol (**1**), they form the orthoformate under the same conditions as Billington *et al.*.[65] Protection of a single axial hydroxyl group, using one equivalent of sodium hydride and paramethoxybenzyl chloride (PMBCl), furnished racemic diol **(±) 86**. Perbenzylation followed by treatment with DIBAL-H induced regioselective orthoformate opening and produced alcohol **(±) 87**. Allylation produced the fully O-protected *myo*-inositol **(±) 88** before removal of the 1,3-methylidene bridge and paramethoxybenzyl (PMB) group by acidic methanolysis afforded triol **(±) 89**. The optical resolution was then conducted by protection of the vicinal 3,4-diol using the camphor acetal strategy. This produced four diastereoisomers, from which the desired alcohol **90** could be isolated in a 31 % yield. This procedure gave the desired intermediate alcohol **90** in 8.2 % yield over 7 steps.[67]

1.4.4.3 Enzymatic Resolutions

As shown so far, routes to *myo*-inositol derivatives starting from *myo*-inositol (**1**) are often long and tedious, with various protecting group manipulations necessary to produce the desired product. The optical purity of the products has been achieved in chemical resolution steps, by adding a chiral auxiliary to a racemic mixture and then separating the diastereoisomers. Enzymes can also be used to effect optical resolutions on *myo*-inositol derivatives. These procedures generally work by the enzyme involved only turning over one of the enantiomers of the racemic mixture. Separation of the enzymatic product and unreacted starting material therefore leads to production of a single enantiomer.



Scheme 1.4.12 – Ozaki's enzymatic resolution of *myo*-inositol derivative **(±) 91**.[68] *Reagents and conditions:* **i**. Ac_2O, Amano Lipase AY, Et_2O, 48 % of **92** and 50 % of **L-91**.

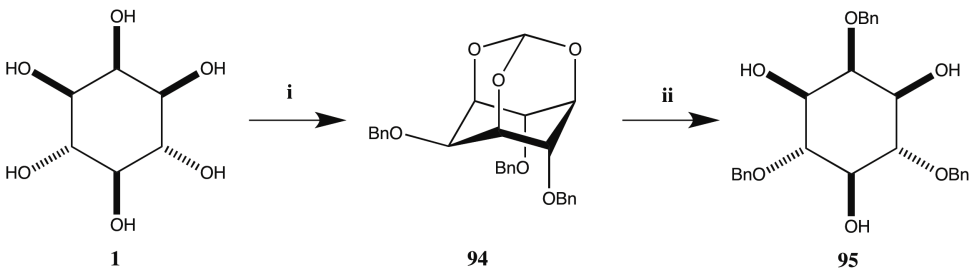
In further work by Ozaki,[68] he was able to provide the first enzymatic resolution of a racemic *myo*-inositol derivative. A small screen of commercially available hydrolytic enzymes provided a lipase from *Candida cylindracea* (Amano Lipase AY) that could regio- and enantioselectively acetylate biscyclohexylidene *myo*-inositol derivatives

(Scheme 1.4.12). Racemate **(±) 91** was treated with acetic anhydride in the presence of Amano Lipase AY and selectively produced acetate **92** in 48 % yield. A trace amount of acetate **93**, the product from acetylation at OH-3 of the D-enantiomer, was also produced. However, the L-enantiomer of **(±) 91** was left unreacted.

1.4.4.4 Enantioselective Catalysis

The synthetic routes to *myo*-inositol derivatives shown so far are generally long and suffer from poor yields due to the critical steps to achieve control over stereochemistry and enantioselectivity. Miller and coworkers[69,70] set out to develop low-molecular weight peptide-based catalysts that could overcome these problems and perform enantioselective phosphorylation of a *meso myo*-inositol derivative. They hoped to produce catalysts that could control regio- and stereoselectivity through their secondary structures much in the way that proteins do.

Miller *et al.*[69] chose to work with 2,4,6-tribenzyl *myo*-inositol (**95**, Scheme 1.4.13) as this compound maintained the *meso* nature of *myo*-inositol (**1**) whilst reducing the number of potential reaction sites from six to three. They had previously seen that they could obtain high levels of selectivity with this type of reaction in non-polar solvents, and using **95** instead of *myo*-inositol (**1**) facilitated use of less polar solvents. The starting material **95** can be accessed relatively straightforwardly from *myo*-inositol (**1**) in a 52 % yield (Scheme 1.4.13).[64,65,71] Formation of the orthoformate of **1** was followed by perbenzylation to furnish the tribenzyl ether **94**. Removal of the orthoformate produced the desired triol **95**.



1 **94** **95**

Scheme 1.4.13 – Synthesis of 2,4,6-tribenzyl *myo*-inositol (**95**).[64,65,71] *Reagents and conditions:* **i**. **a**. $HC(OEt_3)$, pTSA, DMF, 100 °C, 76 %; **b**. BnBr, NaH, DMF, 79 %; **ii**. HCl, MeOH, reflux, 87 %.

With this compound in hand, Miller *et al.*[69] wished to discover peptide catalysts that could selectively phosphorylate either the 1- or 3-position of the molecule. This would desymmetrise the molecule and give an enantiopure product. They screened a small library of random peptides based on modified histidine residues, designed to be small kinase mimics. Screening for the regio- and enantioselectivity of the monophosphorylation of triol **95**, they showed that peptide **96** (Figure 1.4.4) and peptide **97** had excellent selectivity of the 1- and 3-positions respectively, in >98 % ee.[69] It is important to note that these peptides are not enantiomers of each other, but produce opposing selectivities due to their secondary structure.

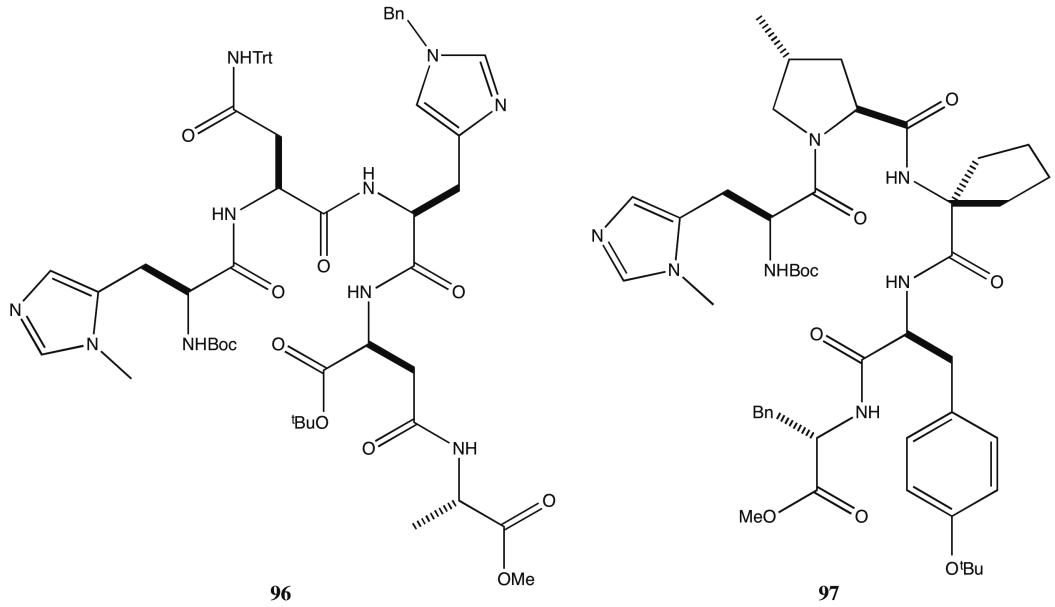
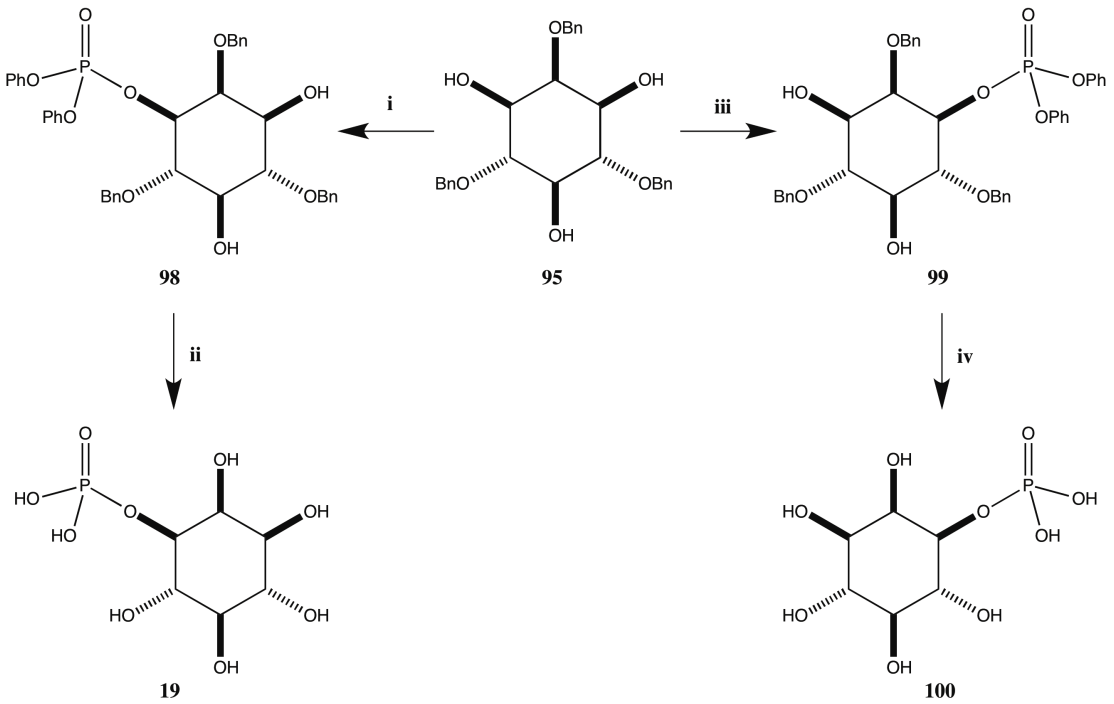


Figure 1.4.4 – Peptide catalysts selected from the screen to have selectivity for phosphorylation at OH-1 (**96**) and OH-3 (**97**).

Upon the discovery of these catalysts, Miller *et al.*[69] completed the synthesis of L-*myo*-inositol 1-phosphate (**19**) and D-*myo*-inositol 1-phosphate (**100**) in concise syntheses from triol **95** (Scheme 1.4.14). Treatment of **95** with diphenylchlorophosphate, Et_3N and either peptide **96** or **97**, gave protected L-*myo*-inositol 1-phosphate **98** and protected D-*myo*-inositol 1-phosphate **99**, respectively. Removal of protection groups via Birch reduction, produced the desired enantiopure *myo*-inositol phosphates **19** and **100**.



Scheme 1.4.14 – Use of peptide catalysts to produce L-*myo*-inositol 1-phosphate (**19**) and D-*myo*-inositol 1-phosphate (**100**).[69] *Reagents and conditions:* **i**. Diphenylchlorophosphate, peptide **96** (2.5 mol %), Et_3N, toluene, 0 °C, 56 % (>98 % ee); **ii**. Li, $NH_{3(l)}$, THF, -78 °C, 95 %; **iii**. Diphenylchlorophosphate, peptide **97** (2.0 mol %), Et_3N, toluene, 0 °C, 65 % (>98 % ee); **iv**. Li, $NH_{3(l)}$, THF, -78 °C, 96 %.

Miller has extended this work to the synthesis of *myo*-inositol polyphosphates,[72] phosphatidylinositols and their phosphates[73–75] and even dimeric *myo*-inositol phosphates.[76,77] Within each piece of work, the catalytic desymmetrising phosphorylation was the key step to provide the desired *myo*-inositol derivatives. This step could only furnish a moderate 65 % yield, however, the remaining starting material could be recovered from the reaction. The requirement to start with 2,4,6-tribenzyl *myo*-inositol (**95**) is a small inconvenience, and the desired *myo*-inositol derivatives are generally obtained with good yields in relatively short syntheses. Availability of the peptide catalysts is a potential issue for this methodology.

1.5 Biocatalysis

Biocatalysis comprises the use of bacteria, fungi and, more specifically, enzymes to catalyse chemical reactions.[78] The benefits of using biocatalysis over traditional chemical methods are the high levels of selectivity achieved within reactions, the mild reaction conditions needed, and the fact that biocatalysts are obtainable from renewable resources and are biodegradable.[78] This field has developed enormously in the past decade and it can now offer solutions to many of synthetic chemistry's challenging problems in the pharmaceutical, fine chemical and lower value industrial sectors.[79]

In its crudest forms, biocatalysis has been used for thousands of years as a way of producing and preserving food such as beer, wine and cheese. It wasn't until the late 1800s that a more refined use of biocatalysis was demonstrated when Louis Pasteur used *Penicillin glaucum* as a way to enantiomerically enrich an aqueous solution of racemic tartaric acid. Since then, the use of biocatalysis within organic synthesis has steadily grown as scientists have developed methods to overcome the inherent difficulties associated with using enzymes or whole cell organisms as a synthetic tool.[79]

One of the major problems hurdles to overcome in the use of enzymes in synthesis was the synthesis of the enzymes themselves. With no fixed method for the production of enzymes on a large scale, the utility of enzymes was very much restricted. However, in the 1970s, the development of a recombinant DNA methodology allowed scientists to obtain large quantities of a given enzyme via overexpression within a host organism, such as *Escherichia coli* (*E. coli*) or *Bacillus subtilis*.[79] The exact conditions needed for the methodology vary from protein to protein, but, since the introduction of the method, modifications and new techniques have vastly improved the diversity of proteins that can be produced in this way.

By the 1990s, improvements in the processes involved in biocatalysis meant that it was increasingly viable within an industrial setting, and many enzymes had already been commercialised.[79] It had found particular use in the preparation of chiral compounds, and was being used in preference of chemical methods to produce complex intermediates within the pharmaceutical, agrochemical and other industries.[80] The biotechnological

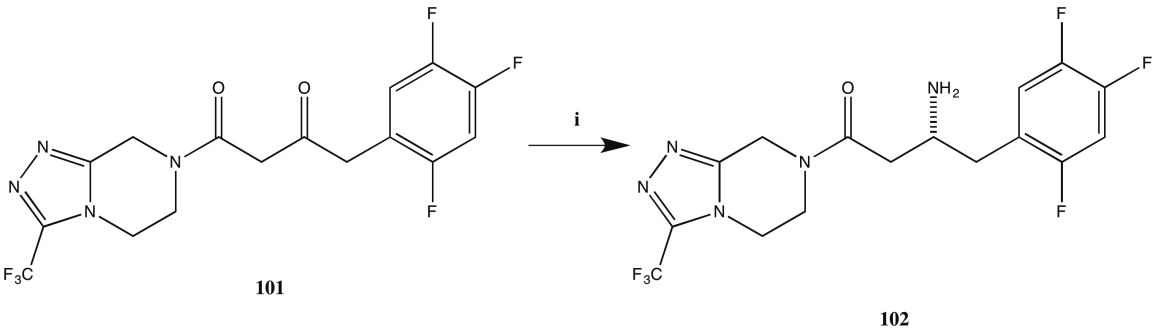
design of reactors, immobilisation of enzymes and optimization of downstream work ups can all be attributed to aiding the rapid development of biocatalysis within industry.[79]

There are a vast number of enzymes within nature that conduct a large variety of transformations, however it is very rare that they lend themselves for use in industrial applications. The problems associated with using enzymes in synthesis are the limited solvent tolerance, poor stability at temperatures outside their normal range, low levels of expression, poor activity, and high selectivity for substrates, which limits their ability to turnover non-natural substrates.[81] As new methods have been developed it has become possible to overcome the majority of these issues, more often than not by altering the properties of the enzyme itself.

Engineering of enzymes is generally conducted at the DNA sequencing level, involving modification of amino acids by chemical or genetic methods. The enzyme produced can then be tested for novel, optimal or improved properties.[82] There are two main methods used to engineer proteins, and it is common to conduct a mixture of both methods to obtain desired enzyme modifications.

Firstly, enzymes can be modified by rational design. Mutations are introduced in the protein-encoding gene at specific postions based upon knowledge of sequence, structure, function and catalytic mechanism. In order to conduct this method, a lot of information about the enzyme's structure and energy functions is needed to correctly predict the outcome of modification.[83]

The second method is via directed evolution.[84] Repeated cycles of random mutagenesis are conducted with variants of the gene to create a library of genes that have slightly different sequences. The resulting enzymes are submitted to high throughput screening or genetic selection to determine which of the enzymes display an improvement in the desired property. This process is repeated on the improved enzymes / gene sequences until an enzyme is produced that has the desired properties. No structural or mechanistic information is needed for the enzyme, and it is likely that changes that occur far from the active site of the enzyme would not have been selected by rational design.[83]



Scheme 1.5.1 – Biocatalytic production of sitagliptin (**102**).[85] *Reagents and conditions:* **i.** *R*-Selective transaminase (R-ATA, ATA-117), isopropylamine, 50 % DMSO in water, 40 °C, 92 %.

Sitagliptin (**102**, Scheme 1.5.1) is a drug used to treat type II diabetes. It is currently produced from the amination of prositagliptin ketone (**101**) using a transaminase to achieve the desired enantioselectivity.[86] *R*-Selective transaminase (R-ATA, ATA-117) from *Arthrobacter* sp. was engineered by Codexis and Merck using directed enzyme evolution to overcome the large substrate size.[85] Coupled with process engineering, the system produced can convert 200 g L^{-1} of the ketone **101** into sitagliptin (**102**) with an enantiopurity of >99.5 %.[85] The process can tolerate large quantities of isopropylamine and runs with 50 % DMSO as the solvent at temperatures greater than 40 °C.[87]

This biocatalytic step provides a number of advantages over the previously used rhodium-catalysed method. Total waste production was reduced and the need for a rare heavy metal was eliminated, making the process more "green". The overall yield of the process was increased by 10 %, and the overall productivity within an industrial setting could be increased by 53 % (measured by Kg L^{-1} per day).[85,86]

1.5.1 Isolated Enzymes versus Whole Cell Systems

There are numerous approaches within biocatalysis that use either isolated enzymes or whole cell systems to achieve a desired transformation. The selection and utility of one of these methods ultimately depends on the nature of the enzyme, substrate, product and reaction involved.[78]

Use of isolated enzymes over whole cell systems is advantageous as there is greater access for the substrate and enzyme with respect to each other, and unwanted side reactions from interference of additional cellular enzymes is eliminated.[88] However, the enzyme may have dramatically reduced stability in an isolated state, and isolated enzymes do not tolerate extremes in temperature, pH or ion concentration, shear stress or organic solvents very well.[88] Within a whole cell system, the cell can act to protect enzymes from harsh external conditions and allow the enzyme to keep its structure and activity. On the other hand, the conditions used may also be toxic to the cell.

The preparation of whole cell reactions is simpler than with isolated enzymes. Potentially lengthy enzyme purification procedures are not needed, and the whole cell is able to conduct multi-step syntheses involving multiple enzymes.[88] Enzymes that require a cofactor, such as NAD^+ or adenosine triphosphate (ATP), also benefit from a whole cell system as these cofactors are present as part of the cell's natural metabolism.[89] If cofactors are needed for an isolated enzyme, it could be necessary to use stoichiometric amounts of potentially very expensive materials or set up a method for cofactor recycling.[89]

It is possible to improve isolated enzymes by increasing their structural rigidity, either by directed evolution[90] or multipoint immobilisation.[91] Immobilisation of enzymes has other advantages, including ease of removal from a reaction vessel, enzyme recycling and use of continuous flow reactors.[88] Unfortunately, increasing the rigidity of the enzyme may have a detrimental effect on enzyme performance. Whereas this is not the case in all enzymes, too much conformational stability can restrict the flexibility of the enzyme that may be needed for activity.[88]

Whole cell processes are well established commercially, but they are mostly limited to batch processes and cannot be reused.[78] Immobilisation of whole cells is possible, but this reduces the mass to catalytic activity ratio and requires more laborious preparation, which will add to the cost.[88]

In general, whole cell systems are used for reactions that involve unstable, multi-component or membrane-bound enzymes, and / or require cofactors. On average, within an industrial setting whole cell systems are less expensive than isolated enzymes.[78]

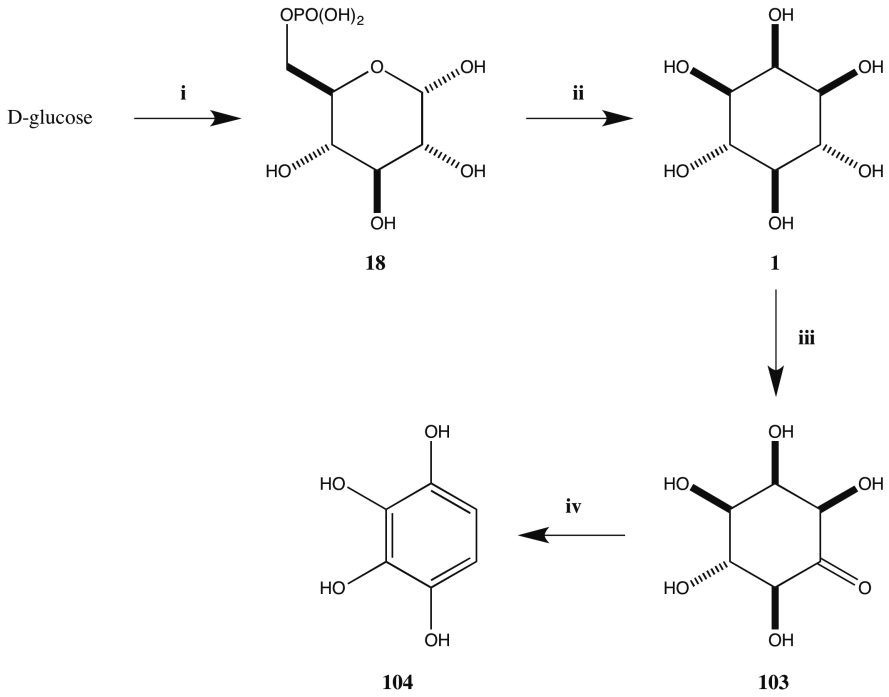
However, there are various factors that need to be taken into account, and the immobilisation of isolated enzymes has provided new methodologies with attractive advantages. The modification of enzymes using directed evolution or rational design and site-directed mutagenesis has provided scientists to overcome the many obstacles that biocatalysis faces.

1.5.2 Synthesis of *myo*-Inositol Derivatives using Biocatalysis

Within the synthesis of *myo*-inositol derivatives biocatalysis has seen very limited use. As mentioned previously (section 1.4.4.3), enzymes have been used to perform resolutions on *myo*-inositol (**1**)-based synthetic intermediates to produce an enantiopure product. The isomerisation of D-glucose 6-phosphate (**18**) to L-*myo*-inositol 1-phosphate (**19**) as accomplished by INO1 has scarcely been used within synthesis[92,93] even though the product is highly desirable.

Both whole cell systems and isolated enzymes have been used to produce L-*myo*-inositol 1-phosphate (**19**). In an effort to provide a convenient synthetic route to polyhydroxybenzenes that contain the oxygenation pattern of 1,2,3,4-tetrahydroxybenzene (**104**), Frost *et al.*[92] utilised a whole cell synthesis of *myo*-inositol (**1**) within *E. coli* (Scheme 1.5.2).

D-Glucose is taken up by *E. coli* JWF1/pAD1.88A and converted to D-Glucose 6-phosphate (**18**) by the phosphotransferase system of the organism, which uses phosphoenolpyruvate as a source of the phosphoryl group. Expression of the *Saccharomyces cerevisiae* INO1 gene within the *E. coli* JWF1/pAD1.88a on plasmid pAD1.88A allowed isomerisation of D-Glucose 6-phosphate (**18**) to L-*myo*-inositol 1-phosphate (**19**). Unidentified phosphatase activity then hydrolysed the phosphoester to produce *myo*-inositol (**1**). Under fed-batch fermentor conditions, this whole-cell system could produce 21 g L^{-1} *myo*-inositol (**1**) and 4 g L^{-1} L-*myo*-inositol 1-phosphate (**19**) in a combined yield of 11 % from D-glucose.[92] The *myo*-inositol (**1**) obtained in this process could be oxidised to *myo*-2-inosose (**103**) by incubation with *Gluconobacter oxidans* ATCC 621, before reflux in sulphuric acid (0.5 M, aq.) produced 1,2,3,4-tetrahydroxybenzene (**104**).[92]



Scheme 1.5.2 – Synthesis of 1,2,3,4-tetrahydroxybenzene (**104**) using the whole cell biocatalytic synthesis of *myo*-inositol (**1**).[92] *Reagents and conditions:* **i.** *E. coli* phosphotransferase system; **ii. a.** *Saccharomyces cerevisiae* INO1; **b.** phosphatase activity, 9 % over 3 steps; **iii.** dehydrogenase activity, 95 %; **iv.** sulphuric acid, water, reflux, 66 %.

This whole cell system is able to produce large quantities of L-*myo*-inositol 1-phosphate (**19**, 4 g L^{-1}), however this work clearly demonstrates a major issue with using a whole cell organism for the production of this compound. Phosphatase enzymes already present within the *E. coli* quickly hydrolyse L-*myo*-inositol 1-phosphate (**19**) to *myo*-inositol (**1**), thus a large quanitity (84 %) of the product is lost. With a combined yield of only 11 % from the D-glucose fed into the system, this means that 4 g L^{-1} is a less than 2 % yield of L-*myo*-inositol 1-phosphate (**19**). Whereas a large quanitity could still be produced, no comment is made on the isolation of L-*myo*-inositol 1-phosphate (**19**) from the supernatant and this may not be a trivial process.

As an isolated enzyme, INO1 has been used to prepare small quantities of L-*myo*-inositol 1-phosphate (**19**) within studies of mycothiol glycosyltransferase (MshA).[93] MshA catalyses the transfer of *N*-acetylglucosamine onto L-*myo*-inositol 1-phosphate (**19**) in the first step of mycothiol (**16**) biosynthesis (discussed in more detail in Chapter 4). To study the structure and mechanism of MshA, the investigators needed a source of L-*myo*-

inositol 1-phosphate (**19**) and chose the INO1 gene from hyperthermophile *Archaeoglobus fulgidus* as it had previously been used to clone and express *Archaeoglobus fulgidus* INO1 (*Af*INO1) in a 9 mg L^{-1} preparation.[93,94]

As the INO1 gene was taken from a hyperthermophile, the optimum temperature for enzyme activity was 85 °C. Their reaction mixture contained 125 mM D-glucose 6-phosphate (**18**), 0.625 mM zinc (II) chloride, 1.25 mM NAD^+ and 3.6 mg of *Af*INO1 in 50 mM tris(hydroxymethyl)methylammonium (Tris) buffer (pH 7.5, NH_4OH) and was held at 85 °C in a heat block for the course of the reaction. Every 45 minutes the concentration of NAD^+ was increased by 0.5 mM and a further 2 mg of *Af*INO1 was added. The reaction was monitored via a BgMshA/pyruvate kinase/lactate assay and was completed after 3 hours.[93]

There is no comment on the quantity of L-*myo*-inositol 1-phosphate (**19**) produced in this procedure, but they were able to produce enough to determine the crystal structure of MshA with L-*myo*-inositol 1-phosphate (**19**) bound within the active site. The reaction is conducted at a very high temperature, which would complicate procedures within an industrial setting and also potentially have a detrimental effect on the product or cofactors. The *Af*INO1 was expressed obtaining only 9 mg L^{-1}.[94] This is quite a low amount of protein to be produced when such a large amount is used in the procedure (9.6 mg). Continuous dosing of NAD^+ and *Af*INO1 was also needed to furnish full conversion to the product.

The use of INO1 within the literature has shown that there is precedence for this enzyme to be used in a biocatalytic setting to produce large quantities of the highly desirable L-*myo*-inositol 1-phosphate (**19**). In an industrial setting these methods may fail to provide an efficient and economical route to the product.

1.6 Conclusions and Project Aims

Routes to *myo*-inositol derivatives have long been unsatisfactory. The complexity of these compounds has led to the development of a vast number of techniques to try and provide the shortest possible synthetic routes to a high yield of enantiopure product. The

biosynthesis of L-*myo*-inositol 1-phosphate (**19**) from D-glucose 6-phosphate (**18**) is a one step process catalyzed by INO1 and furnishes an optically pure *myo*-inositol derivative. Whilst INO1 has been used to produce L-*myo*-inositol 1-phosphate (**19**) synthetically, the processes involved have not been optimised to produce large quantities of this product. With careful consideration of the processes involved, biocatalysis could provide an efficient route to the highly desirable product. The recycling of NAD^+ is an added bonus as it allows for a lower amount of the relatively expensive cofactor to be used.

Further to this, with the production of an optically pure *myo*-inositol derivative, we believe that protection of the phosphate and subsequent selective protection of hydroxyl groups will produce dramatically reduced syntheses of *myo*-inositol derivatives. Access to large quantities of these derivatives and analogues to be used as novel chemical probes, will mean that important biological questions can be addressed.

The work undertaken here details the optimisation and scale up of the biotransformation to a scale that can produce synthetically viable quantities (~500 mg) of pure L-*myo*-inositol 1-phosphate (**19**). Subsequent demonstration of the utility of this process within synthesis is an important target to achieve. Thus providing evidence for the power of this biotransformation within organic synthesis.

Chapter Two

2 Identification and Utility of a Suitable INO1

2.1 *Trypanosoma brucei*

Roughly 15 % of the world's population is affected by infectious diseases caused by parasitic protozoa, with millions of fatalities.[95] Human African Trypanosomiasis (HAT), also known as African sleeping sickness, is an infectious disease caused by the extracellular protozoan parasite *Trypanosoma brucei* (*T. brucei*).[96] The parasite is transmitted to the mammalian host via the bite of an infected tsetse fly and the disease is endemic in sub-Saharan Africa, where the tsetse fly thrives.[97]

The life cycle of *T. brucei* is relatively complex, with the parasite adopting different morphologies depending upon environment (Figure 2.1.1).[98] The life cycle can be seen to begin with the procyclic form inside the tsetse fly's midgut. This is where the procyclic form of the parasite replicates and the number of parasites proliferates. Migration of the parasite into the salivary gland of the insect induces transformation into the metacyclic form. When the tsetse fly takes a blood-meal, it is the metacyclic form of the parasite that is injected into the mammalian host. Upon entry into the bloodstream of the mammalian host, the parasites morph into the bloodstream form and can replicate as the slender form. Some parasites will morph into a non-replicative stumpy form, which is pre-adapted to the environment presented to the parasite after a tsetse fly blood-meal. This allows the cycle to continue. Eventually the bloodstream form of the parasite will cross the blood-brain barrier and invade the host's central nervous system.[96] Once the parasite has crossed this barrier, the more obvious symptoms of the disease appear. These include changes in behaviour, slurred speech and disturbance to the sleep cycle.[99] Without treatment, HAT will lead to progressive mental deterioration before death of the host.

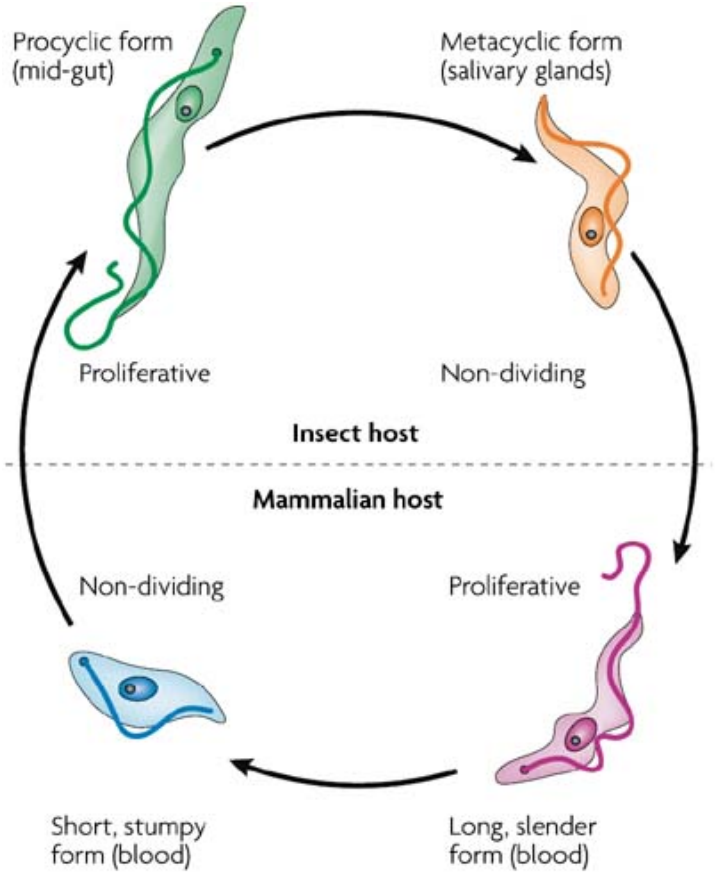


Figure 2.1.1 – Life cycle of protozoan parasite *Trypanosoma brucei*.[98]

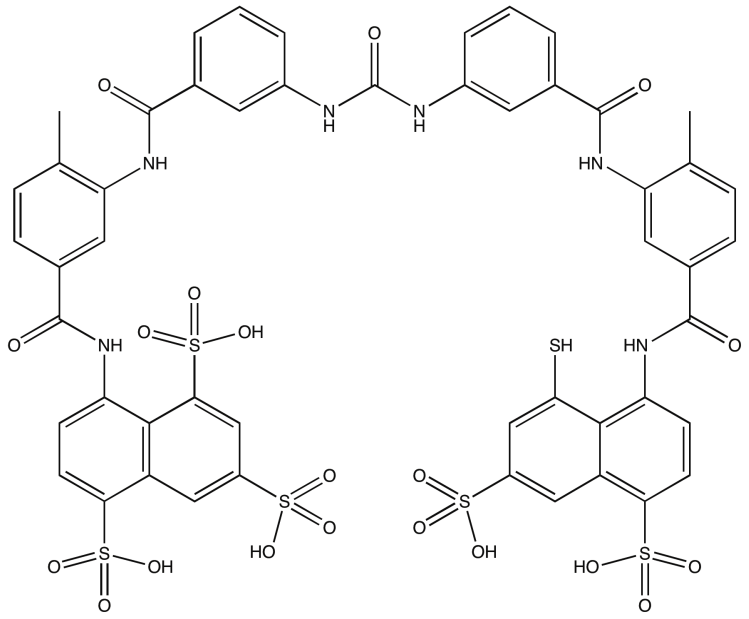
T. brucei has three subspecies: *Trypanosoma brucei gambiense* (*T. b. gambiense*), *Trypanosoma brucei rhodesiense* (*T. b. rhodesiense*) and *Trypanosoma brucei brucei* (*T. b. brucei*), which only infects humans very rarely.[100] Over 95 % of current cases of HAT are caused by *T. b. gambiense*,[97] which is endemic in Western and Central Africa. *T. b. rhodesiense* is responsible for the remaining cases and endemic in Eastern and Southern Africa.[97]

The disease came close to eradication in the late 1960s, however in the last 50 years the disease has re-emerged and the number of cases has risen to levels as high as they were in the 1920s.[96] This can be attributed to the rarity of cases leading to a loss in interest in sustained surveillance and the possible re-emergence of the disease being overlooked.[101] Further to this, control of HAT was complicated during the 1990s due to civil wars and social upheavals that prevented access to a large number of HAT-endemic areas.[101] Since 2001, drugs to treat HAT have been supplied for free by the pharmaceutical companies

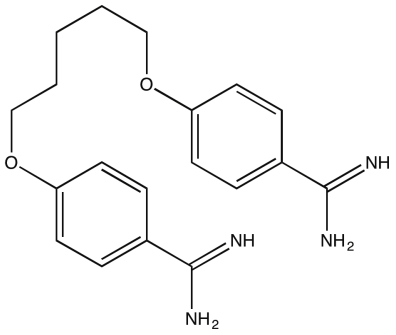
that produce them. This, coupled with increased funding for WHO led sleeping sickness control programmes, has helped decrease the prevalence of HAT. In 2012, the number of reported cases dropped to 8,000, even though it is thought that due to incomplete surveillance this number may be as high as 20,000.[97]

The World Health Organization (WHO) reports that HAT is a serious health risk to roughly 60 million people in sub-Saharan Africa.[95] However, more recently, increased efforts in vector control have decreased the annual rate of infections to 50,000 per year.[95] The estimated number of fatalities has also dropped to 7,000 per year.[95] In addition to the human risk, the animal form of the disease, called Nagana, continues to affect the cattle population of Africa. This has a major impact on African economy with the WHO estimating a loss of roughly 4 billion US dollars per year.[95]

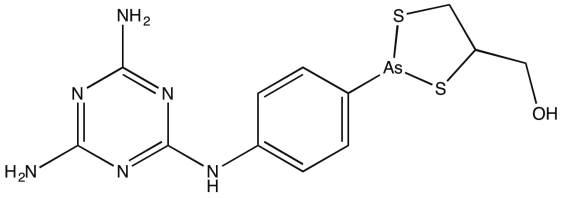
The current human and animal drug therapies have long been considered unsatisfactory, and with *T. brucei* having a growing resistance to established drugs, the situation is only growing worse.[95] Different stages of the disease require different treatment regimens, and it is currently necessary to perform a lumbar puncture (or spinal tap) to determine the appropriate drug.[97]



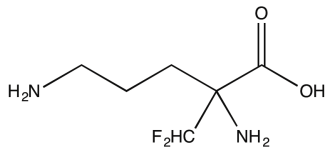
Suramin, **105**



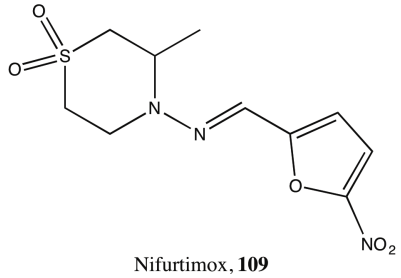
Pentamidine, **106**



Melarsoprol, **107**



Eflornithine, **108**



Nifurtimox, **109**

Figure 2.1.2 – Current drugs used to treat HAT.

The first-line drugs used for the treatment of *T. b.* gambiense and *T. b. rhodesiense* are suramin (**105**, Figure 2.1.2) and pentamidine (**106**), respectively. The first-line drugs are less toxic and also easier to administer than the second-line drugs melarsoprol (**107**) and eflornithine (**108**).[97] Melarsoprol (**107**) was introduced in 1949, however it fails to cure between 10 % and 30 % of patients. On top of this, serious side effects from the arsenic content of the drug kill between 4 % and 10 % of those who take it.[96] Eflornithine (**108**) is the most recent anti-HAT drug, and faces criticism because it is expensive, difficult to

administer and only effective against *T. b. gambiense*. A combination therapy of eflornithine (**108**) and nifurtimox (**109**) was introduced in 2009 which simplifies the administration of eflornithine (**108**), however the treatment is still not effective for *T. b. rhodesiense*.[97] The poor quality of the current therapies for this disease means that there is an urgent need for new therapeutic approaches for HAT and Nagana. The treatment of other Third World diseases caused by closely related protozoan parasites, such as *Leishmania spp.* and *Trypanosoma cruzi*, may also benefit from new therapies for HAT.[95]

When designing new drugs for the treatment of parasitic disease, the obvious place to target the parasite is somewhere that it differs with the host. Melarsoprol (**107**) does not do this as it works by giving the host a dose of arsenic that is large enough to kill the parasite, but not large enough to kill the host. This means that the drug is incredibly toxic to the host, and has many side effects associated with it. By targeting an area which is completely different to the host, not only will the parasite be killed, but the host should encounter a fewer number of side effects.

As part of its life cycle, *T. brucei* needs to be able to survive vastly different environments of the mammalian host's bloodstream and various tissues within the tsetse vector.[95] To account for this, the outer surface of the parasite is covered in a dense cell-surface coat of roughly 5 million variant surface glycoprotein (VSG) dimers, which act as a diffusion barrier (Figure 2.1.3).[96] The parasite has a repertoire of roughly 1000 immunologically distinct VSG genes that it can utilize in a process called antigenic variation to adapt and constantly evade the host's immune system,[96,102] thus making vaccines non-viable. Independent of which variant is present, the VSG is linked to the *T. brucei* cell membrane via GPI anchors.[96] Genetic and chemical validations have shown that *T. brucei* GPI anchor biosynthesis is a valid drug target against HAT.[96]

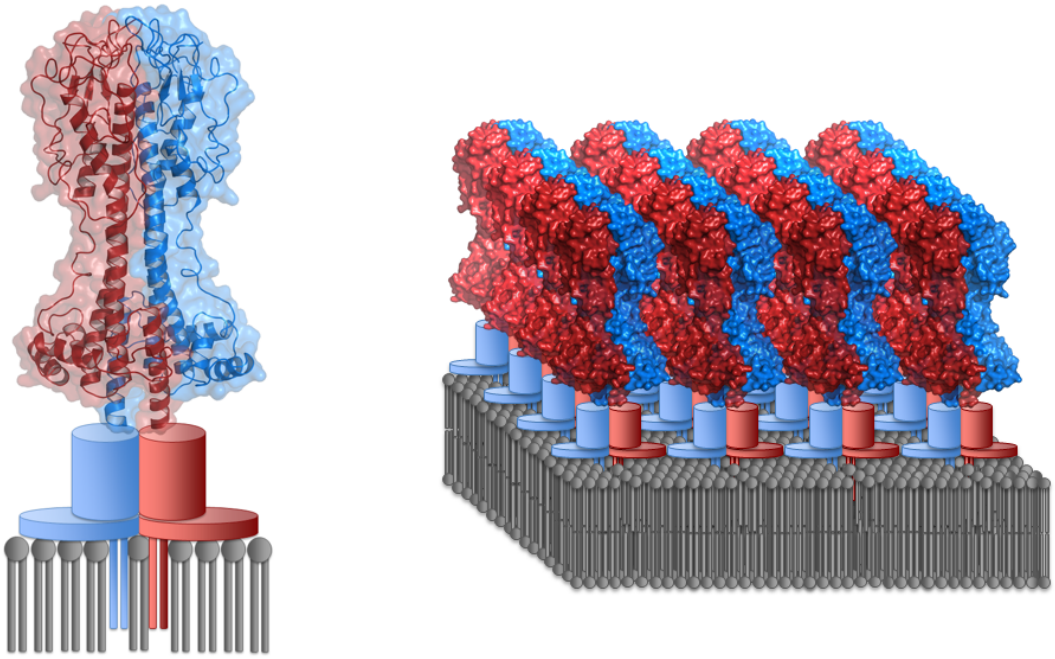
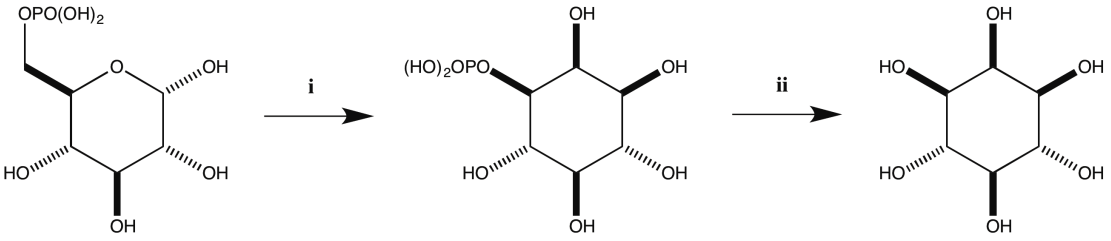


Figure 2.1.3 – Dimeric N-terminal domain of VSG from *T. brucei* variant MITat shown embedded within the lipid bilayer via a GPI anchor and an example of the dense cell surface coat.[103]

Metabolism of *myo*-inositol (**1**) within *T. brucei* is compartmentalised.[104] *myo*-Inositol (**1**) scavenged from the environment is used for the synthesis of bulk phosphatidylinositol within the golgi, whereas *de novo* synthesised *myo*-inositol is used solitarily for GPI anchor biosynthesis within the ER.[104] Coupled with the fact that GPI anchor biosynthesis is essential for the survival of *T. brucei*, this suggests that the *de novo* synthesis of *myo*-inositol (**1**) is also essential and therefore a valid drug target. Previous work conducted within the Smith group has demonstrated this through creation of a conditional double knockout.[22] Increasing the concentration of *myo*-inositol within the growth media from 40 μM to 100 mM did not allow conditional knockout cells to overcome the deletion of the gene, further strengthening the hypothesis that GPI biosynthesis only uses *de novo* synthesised *myo*-inositol (**1**).[22]

As detailed previously, the *de novo* synthesis of *myo*-inositol (**1**) is conducted by the action of two enzymes: INO1 and IMPase. The rate-limiting step is the conversion of D-glucose 6-phosphate (**18**) to L-*myo*-inositol 1-phosphate (**19**) by INO1 in the presence of NAD^+ (Scheme 2.1.1). Within the previous work of the Smith group,[22] an INO1 was

identified within the *T. brucei* genome database (Sanger centre) and the putative open reading frame was PCR-amplified, cloned and sequenced (GenBank, Accession #AJ86670). Alignment of the predicted translated sequence for the *T. brucei* INO1 (*Tb*INO1) and INO1s from other organisms showed several similarities. Firstly, the *Tb*INO1 has a perfect copy of the GWGGNNG motif within a Rossmann fold, associated with NAD^+ binding.[24] A further three motifs were also identified from other eukaryotic INO1s: LWTANTERY, NGSPQNTFVPGL and SYNHLGNNDG.[24] At 58 kDa, the size of the predicted TbINO1 is also in good agreement with other INO1s, which typically fall within the range of 58-67 kDa.[22,23]



Scheme 2.1.1 – Biosynthesis of *myo*-inositol (**1**). *Reagents and conditions:* **i**. INO1, NAD^+; **ii**. IMPase, Mg^{2+}.

*Tb*INO1 was overexpressed in *E. coli* using a pBAD TA vector, causing the encoding of a C-terminal hexa-His-tag. This allowed for purification of the recombinant protein through affinity chromatography using a Ni^{2+}-charged Sepharose column and elution using an increasing concentration of imidazole. MALDI analysis showed the molecular weight of the purified recombinant protein to be 64 kDa, as predicted.[22]

They took advantage of the IMPase enzyme to develop a coupled assay and show catalytic activity of the recombinant *Tb*INO1. Saturation kinetics were conducted, giving an apparent Km of 0.58 mM. This is very close to the human INO1, which has a Km of 0.57 mM.[105] However, at 756 nmol mL^{-1} mg^{-1}, the specific activity of the recombinant *Tb*INO1 is over 10 times higher than the human recombinant INO1. This high activity may be due to the importance of the *de novo* synthesis of *myo*-inositol (**1**) to the survival of *T. brucei*.[22]

2.1.1 Suitability of *Tb*INO1 for Biocatalysis

As discussed previously (Section 1.5.2), previous methods in the biocatalytic production of L-*myo*-inositol 1-phosphate (**19**) have been conducted using a whole cell system[92] and isolated *Af*INO1.[93] The whole cell system faces criticism due to low yields of L-*myo*-inositol 1-phosphate (**19**) compared to the amount of D-glucose that is fed into the system. A 2 % yield of the desired product is obtained, with the majority (84 %) of the L-*myo*-inositol 1-phosphate (**19**) being produced being hydrolysed by unidentified phosphatases within the cell.[92] There is no comment on isolation of the product from the supernatant either. With such a low yield of L-*myo*-inositol 1-phosphate (**19**) this system is not viewed as efficient enough for industrial scale up.

The use of isolated enzyme *Af*INO1[93] does have advantages. The enzyme comes from a hyperthermophile organism and is therefore extremely stable at high temperatures. The enzyme also has an incredibly high specific activity of 13.1 μmol mL^{-1} mg^{-1}.[94] Coupled with a Km of 0.12 mM, this makes *Af*INO1 a more attractive candidate than *Tb*INO1 for biocatalysis when considering the enzyme kinetics. Kinetics is not the only consideration when choosing which enzyme to use for biocatalysis, and other, more practical, considerations have to be taken into account.

The *Af*INO1 enzyme is expressed and purified in a preparation that produces 9 mg mL^{-1} of culture.[94] Compared with *Tb*INO1, which is expressed and purified from preparations that produce 75 mg mL^{-1},[106] this is a low level of expression. The high temperatures used within the reaction with *Af*INO1 (85 °C) are harder to achieve on an industrial scale than the 37 °C needed for *Tb*INO1. The high temperature also has the potential to cause decomposition of starting materials or products. When *Af*INO1 was used previously to prepare L-*myo*-inositol 1-phosphate (**19**) for studies on the MshA enzyme, the reaction required additional dosing of both NAD^+ and *Af*INO1 to furnish full conversion.

The reaction conditions required for *Af*INO1 differ from all other previously studied INO1 enzymes. Instead of ammonium ions the archaeal INO1 requires divalent metal ions to be present, with manganese or zinc ions being preferred.[94] Using *Tb*INO1 instead of *Af*INO1 therefore avoids the need to use, and subsequently remove, metal ions from the reaction.

*Tb*INO1 has a relatively high specific activity of 756 nmol mL^{-1} mg^{-1} when compared to those from humans (7 nmol mL^{-1} mg^{-1}) and other organisms.[105,107,108] Compared to the INO1 from *Archaeoglobus fulgidus* the enzyme activity is optimal at 37 °C and does not require metal ions for the reaction to proceed. For these reasons we believe that *Tb*INO1 is an attractive candidate for biocatalysis. L-*myo*-Inositol 1-phosphate (**19**), the product of the biotransformation, has a market value of £163 mg^{-1} (Cayman Chemical, at time of writing). On top of this, the optical purity of the product means that the biocatalysis could be used as a starting point for the synthesis of inositol derivatives. Offering dramatically shortened synthetic routes to access the highly valuable *myo*-inositol derivatives important for research.

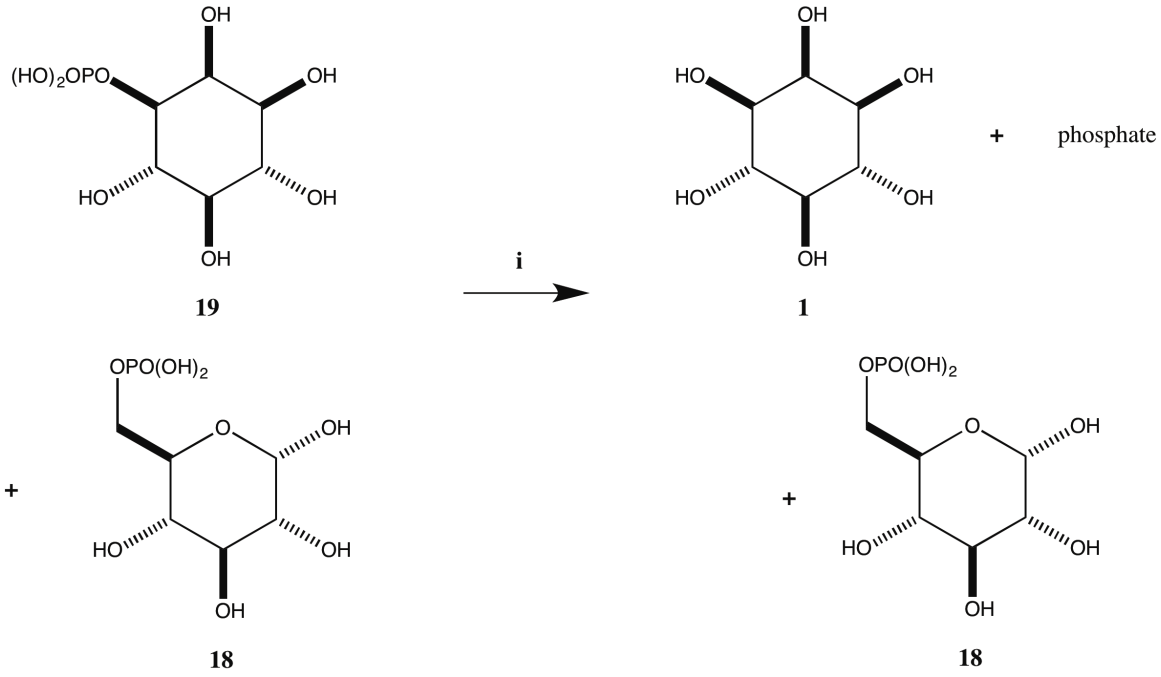
2.2 Considerations for Scale-Up

The initial work conducted on the biotransformation was undertaken under an investigative mindset. The protein was being studied and screened for inhibitors so only a very small scale was necessary to see the desired results. Therefore, the conditions used within the experiments were optimised without taking into account a potential scale up of the process. This presented several obstacles when faced with the task of scaling the process up to produce a synthetically viable amount of L-*myo*-inositol 1-phosphate (**19**).

The biggest problem with this biotransformation is the high polarity of the starting material and product. D-Glucose 6-phosphate (**18**) and L-*myo*-inositol 1-phosphate (**19**) are so similar in polarity that separation of the two compounds by conventional methods is going to be incredibly difficult, if not impossible. This means that the only way to produce high purity L-*myo*-inositol 1-phosphate (**19**), without a convoluted purification method, will require the biotransformation to proceed with 100 % conversion. Maintaining this conversion throughout the scale up of the process proved to be a difficult task.

Calculation of conversion itself also needed to be taken into consideration. Direct measurement of conversion on the initial system was not possible due to the experiments being conducted on such a small scale (0.4 mg of D-glucose 6-phosphate (**18**) was being

used per experiment). Fortunately, the work conducted on this biotransformation previously had developed an assay to measure the conversion. They took advantage of the IMPase enzyme that selectively cleaves phosphate from L-*myo*-inositol 1-phosphate (**19**), whilst leaving D-glucose 6-phosphate (**18**) untouched (Scheme 2.2.1). By pushing this second biotransformation to completion, it can be assumed that any free phosphate present in the sample is an indication of the amount of L-*myo*-inositol 1-phosphate (**19**) that was present. The amount of phosphate that is present can be measured by addition of Malachite Green. This reagent turns green in the presence of free phosphate, emitting light at 650 nm. Absorbance readings can then be used to determine the conversion as these readings are directly proportional to the free phosphate concentration, and hence the concentration of L-*myo*-inositol 1-phosphate (**19**) within the sample.



Scheme 2.2.1 – IMPase assay used to calculate INO1 reaction conversion. IMPase does not hydrolyse phosphate of D-glucose 6-phosphate (**18**). *Reagents and conditions:* **i.** IMPase (1 mU), $MgCl_2$ (4 mM), Tris-acetate (50 mM, pH 8.0), 37 °C.

In the initial work, the biotransformation was conducted in a tris(hydroxymethyl)methylammonium (Tris) acetate buffer. This buffer was ideal for their work as its pH range is consistent with that of the protein's activity, whilst it doesn't contain any free phosphate to interfere with the IMPase assay. The removal of the associated Tris acetate salt at the end of the biotransformation was highlighted as a

potential problem. Therefore, a different buffer was sought that maintained high enzyme activity and a lack of free phosphate, whilst being easy to remove. Ammonium bicarbonate seemed like an ideal candidate, as the buffer salt should easily decompose into ammonia, carbon dioxide and water.

The final factor to take into consideration was protein removal. With an "in-batch" process this would be an issue, however if immobilisation of the protein was necessary then it would not be a problem. Two methods could be used to remove the protein from the solution: acidification and subsequent filtration to remove the denatured protein, or passing the reaction mixture through a centrifugal size-exclusion filter. Both of these methods were found to be satisfactory during this project, even though the size-exclusion filtration took longer to conduct. Whilst choosing between these two methods, it was necessary to think about the next step in the process: selective phosphate benzylation using phenyldiazomethane. This step requires the phosphate to be in its free acid form, which can be achieved by acidification. Therefore, when an "in-batch" process was conducted the preferred method of protein removal was acidification followed by filtration.

2.3 Preliminary Results

2.3.1 *Tb*INO1 Protein Overexpression and Purification

*Tb*INO1 was recombinantly expressed and purified on a large scale using the construct pET15b-*Tb*INO1 in BL21 Rosetta (DE3) cells.[102] The protein was purified by Ni affinity chromatography, eluted with an increasing imidazole gradient (10 mM to 400 mM) in a 20 mM Tris acetate (pH 7.5) buffer containing 300 mM NaCl. Removal of imidazole was achieved by dialysis against 20 mM Tris acetate (pH 7.5), 50 mM NaCl and 5 mM dithiothreitol (DTT). The protein was then stored containing 20 % glycerol at -80 °C.[102] When conducted, 8 L of *E. coli* cultures were grown and, typically, 600 mg of *Tb*INO1 was collected after purification. This gives 75 mg L^{-1} of *Tb*INO1, a very high yielding expression. Electrophoresis gels provided evidence of a good level or purity from the procedure (Figure 2.3.1).

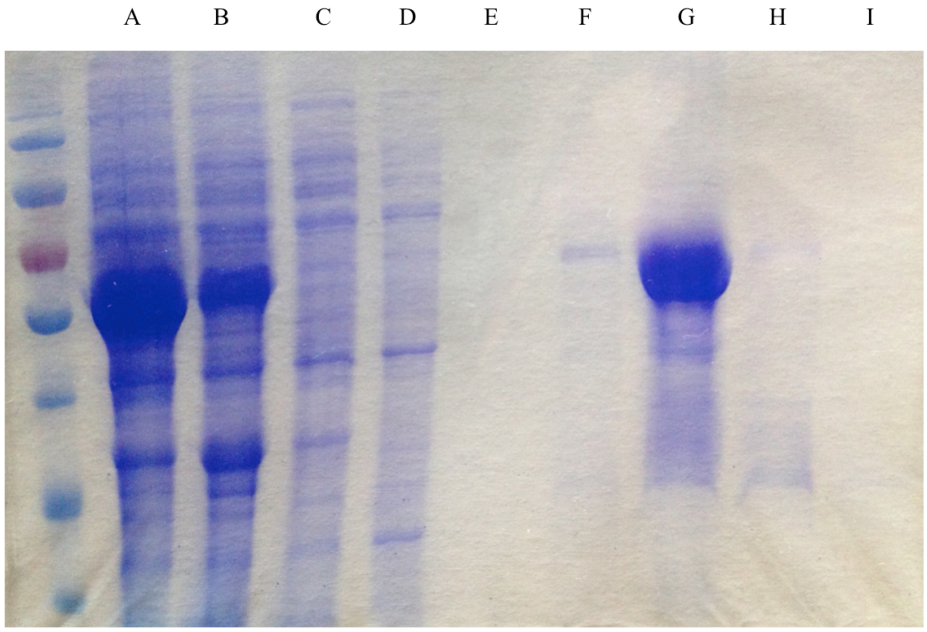


Figure 2.3.1 – Protein electrophoresis gel from a standard protein purification. A – Whole cell, B – supernatant, C – Flow through, D – 10 mM imidazole wash, E – 20 mM imidazole wash, F – 50 mM imidazole wash, G – 100 mM imidazole wash, H – 250 mM imidazole wash, I – 400 mM imidazole wash.

2.3.2 Initial Optimisation

Before any work could be done on the scale up of the biotransformation, it was necessary to optimise the process on a small scale. As high conversion had been highlighted as an important target, work on the scale up of the process would not be started until a conversion of >95 % had been achieved. The original system[22] had achieved such conversions, but switching to an ammonium bicarbonate buffer could have a big effect on the proficiency of the process. The first experiments used the same amount of protein (87 μg), D-glucose 6-phosphate (**18**, 10 mM) and DTT (1 mM) throughout, and the total volume of the reaction mixture was 150 μL. The concentration of ammonium bicarbonate in the buffer was kept at 50 mM and the temperature of the reaction was always 37 °C. The pH and concentrations of NAD^+ and ammonium acetate were investigated.

Initial results were poor, with the IMPase assay demonstrating conversions as low as 8 %. The first variable that was investigated was the pH of the ammonium bicarbonate buffer. Experiments were conducted at three pH values: 8.0, 8.5 and 9.0. Overnight conversions

for these reactions were again very low, with only 16 %, 23 % and 20 % being observed respectively (Table 2.3.1). With conversions this low, it was hypothesised that the *Tb*INO1 used had lost some activity after a long period of storage. Therefore, a new batch of *Tb*INO1 was expressed and purified. Utilising the new *Tb*INO1, the reaction gave a 53 % conversion with the buffer at pH 8.5.

*Tb*INO1	pH	[NAD^+]	Conversion
Old	8.0	1 mM	16 %
Old	8.5	1 mM	23 %
Old	9.0	1 mM	20 %
New	8.5	1 mM	53 %
New	8.5	2 mM	56 %
New	8.5	4 mM	50 %

Table 2.3.1 – Calculated conversions showing differences in pH, old and new *Tb*INO1 batches and NAD^+ concentration. Other variables were set to 87 μg of TbINO1, 10 mM D-glucose 6-phosphate (**18**), 2 mM ammonium acetate and 50 mM ammonium bicarbonate as buffer. Reactions incubated at 37 °C overnight.

As the cofactor, NAD^+ plays an essential role in the reaction. Potential optimisation of the NAD^+ concentration was conducted by running three overnight experiments in parallel, with varying concentrations of NAD^+ (Table 2.3.1). With only a small variation in conversion observed between these reactions, it is difficult to draw any conclusions as to the optimal concentration of NAD^+. A concentration of 2 mM was chosen for remaining experiments. The concentration of ammonium acetate within the reaction mixture, seemed to have little or no effect on conversion. The previous work[22] had shown that ammonium ions were needed for the biotransformation to progress. The selected ammonium bicarbonate buffer now contained this ion in high abundance. Doubling the ammonium acetate concentration to 4 mM had little to no effect, and indeed, in later experiments, completely omitting the addition of ammonium acetate had no effect on conversion. The DTT is included in the reaction mixture to prolong the life of the protein by relieving oxidative stress. It reduces and prevents enzyme S-S bonds

forming intra- or intermolecularly with cysteine residues. Reaction conversion was therefore not affected by an increase in DTT concentration.

Having investigated each of the components of the reaction, it was hypothesised that there could be an issue with the IMPase assay. If the IMPase was not efficiently cleaving the phosphate from L-*myo*-inositol 1-phosphate (**19**) then a true representation of the conversion would not be observed. The reaction conditions used for the IMPase reactions were initially 50 mM ammonium bicarbonate (pH 8.5), 5 mM magnesium chloride and 1.25 mU IMPase (bovine brain, Sigma) for 3 hours at 37 °C. Increasing the reaction time had little to no effect on the conversions observed. Leaving the reaction overnight led to the reaction standard (containing no *Tb*INO1) being observably very green in colour after addition of Malachite Green, suggesting that IMPase is able to cleave the phosphate from D-glucose 6-phosphate (**18**). Although this suggests that IMPase is not selective for the reaction product, D-glucose 6-phosphate (**18**) is hydrolysed at a much slower rate than L-*myo*-inositol 1-phosphate (**19**), and within the preferred 3 hour reaction time for the IMPase reaction, the amount of free phosphate present from the starting material would be negligible. Using an increased amount of IMPase gave no discernible increase in conversion, and changes to magnesium chloride concentration also had little to no effect.

Finally, the Tris acetate buffer was used for this reaction. This was the original buffer used by the previous work,[22] but had been changed to ammonium bicarbonate in error causing the IMPase to act less efficiently. Using the original buffer for this reaction, pushed the conversion of 56 % that had been observed previously, up to 80 %.

On this initial 150 μL scale, the desired conversion of >95 % had still not been observed up to this point. It was hypothesised that switching the buffer of the stored enzyme prior to addition to the reaction mixture, may have a desirable effect. The enzyme had been stored in the previously used Tris acetate buffer, and it was discovered that by switching this to the ammonium bicarbonate buffer the reaction conversion was pushed to 92 % on the first occasion and since then, conversions >98 % have been observed. The buffer was switched using centrifugal filters with 30 kDa cut off, washing the protein with ammonium bicarbonate buffer three times before adding to the reaction.

Variable	Optimum
D-Glucose 6-phosphate (**18**)	10 mM
*Tb*INO1	87 μg
NAD^+	2 mM
DTT	1 mM
Ammonium bicarbonate	50 mM (pH 8.5)

Table 2.3.2 – Table showing the required conditions to achieve full conversion of D-glucose 6-phosphate (**18**) to L-*myo*-inositol 1-phosphate (**19**) at 37 °C overnight. Total reaction volume of 150 μL.

The final system used to obtain >95 % conversion on this 0.4 mg scale (10 mM D-glucose 6-phosphate (**18**) in 150 μL reaction) was therefore 87 μg *Tb*INO1, 50 mM ammonium bicarbonate (pH 8.5), 2 mM NAD^+ and 1 mM DTT conducted at 37 °C overnight (Table 2.3.2).

2.4 Scale-Up "In-Batch"

2.4.1 10 mg

With the biotransformation optimised to obtain a conversion >95 % on 0.4 mg of D-glucose 6-phosphate (**18**), it was now possible to investigate the scale-up of the process. The first target was to scale-up to using 10 mg of starting material, which would be a scale-up of 25 times magnitude.

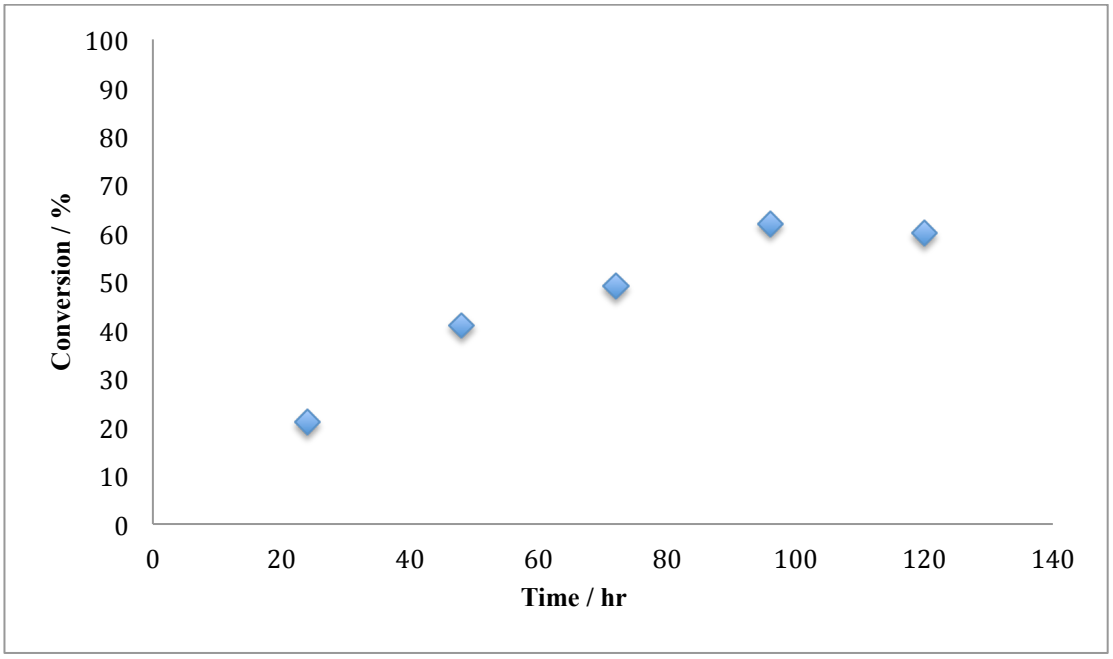
There were no observed problems with this scale up. Using 300 μg of *Tb*INO1, increasing the reaction volume to 600 μL and extending the reaction time to 3 days gave quantitative conversions.

2.4.2 100 mg

The initial attempt using 100 mg of D-glucose 6-phosphate (**18**) was not as successful as the 10 mg scale. In this experiment, the total reaction volume was increased to 4 mL, whereas the amount of *Tb*INO1 was left at 300 μg. This amount of protein should be more than enough turnover 100 mg of substrate in a reasonable time period. The reaction was assayed every 24 hours, and the initial conversion calculated, 26 %, looked promising. However, the next day the reaction had only progressed to 32 % conversion and, after a further 2 days, the reaction did not progress any further. At this point, an extra 300 μg of *Tb*INO1 was added to the reaction. Unfortunately, the additional protein only pushed conversion to 40 %. The D-glucose 6-phosphate (**18**) being used was in the form of a disodium salt, meaning that the high concentration of the starting material would not dramatically alter the buffer pH. Potential differences in pH were, therefore, not hypothesised as a reason for low conversion.

The second experiment on this scale was conducted with a higher concentration. Using only 1 mL for the total volume would make the D-glucose 6-phosphate (**18**) and protein 4 times as concentrated. To better accommodate this increase in enzyme concentration, the amount of NAD^+ present was increased from 2 mM to 4 mM. The reaction was also placed within an incubator with shaking.

The results for this experiment are shown in Graph 2.4.1. There was a steady increase in conversion over the first 4 days, however after this point the conversion plateaued at 62 %. Extra NAD^+ was added at this point, however this produced no further reaction. This addition of NAD^+ was conducted because within the literature[109] it is stated that NAD^+ is subject to decomposition for prolonged periods in aqueous solution at basic pH. At pH 8.5, this reaction is slightly more basic than the recommended pH of 7 to 7.5.[109] the reaction had gradually turned yellow over the course of the experiment. The major mechanism for NAD^+ decomposition is from nucleophilic attack on C-4 of the nicotinamide moiety, leading to a 1,4-dihydropyridine.[109]



Graph 2.4.1 – Calculated conversion as time progressed. Reaction conditions: 100 mg D-glucose 6-phosphate (**18**), 4 mM NAD^+, 1 mM DTT, 50 mM ammonium bicarbonate (pH 8.5), 300 µg *Tb*INO1, 37 °C with shaking, and 1 mL total reaction volume.

Conducting the reaction in the dark and under an inert atmosphere had a positive effect and experiment showed a steady increase of conversion for 6 days, allowing the reaction to reach 76 % conversion. At this point, the reaction gave no further increase in conversion and a white precipitate had formed within the reaction mixture. It is hypothesised that the white precipitate is denatured *Tb*INO1 and that six days is the longest that this protein is stable to the conditions of the reaction. Indeed, with further experiments, the protein was never observed to survive longer than this. This problem meant that there was only a window of 6 days to achieve a quantitative conversion with this process.

The obvious way to achieve this would be to increase the protein to substrate ratio. Using 600 µg of *Tb*INO1 in this system caused the protein to denature quicker, in only 2 days, and halted the reaction at 26 %. To accommodate the increase in protein, a separate experiment was conducted in which the total volume was increased to 5 mL. This allowed the reaction to progress to 50 % over the course of 5 days before the protein precipitated.

One factor that needed to be looked into was the mixing of the reaction mixture. The reaction definitely had a higher conversion when the reaction mixture was placed in an incubator with shaking, as evidenced in experiments after the initial one on this scale. However, with a total volume of 5 mL the reaction vessel being used was a 15 mL Falcon tube, which are very narrow. The Falcon tube was being stood vertically within the shaker, which did not allow for very thorough mixing. An experiment was conducted by placing the Falcon tube on its side within the shaking incubator, with the thought that this would lead to a more effective agitation of the reaction mixture. This was indeed the case, however the mixing was now too vigorous and all that this accomplished was the protein becoming denatured more quickly. Within 2 hours the protein had precipitated.

A potential solution to this problem was conducting the reaction mixture within a 50 mL Falcon tube. This vessel is much wider than its 15 mL counterpart and would allow the reaction mixture to mix more effectively, without being mixed too aggressively. Unfortunately, this had no positive effect and the reaction only gave a 32 % conversion.

A rotating wheel was also assessed as a potential solution to this problem. Placing the Falcon tube in this piece of apparatus allows it to be gently inverted, allowing for a non-vigorous mixing of the reaction mixture. Unfortunately, an incubated rotating wheel was not available, however the reaction was conducted at room temperature in parallel with a reaction being run in an incubator with shaking. After 2 days, the reaction in the incubator showed a conversion of 30 %, whereas the reaction in the rotating wheel gave a conversion of 28 %. Although this experiment contained two variables (method of mixing and temperature), it is known[22] that *Tb*INO1 has an optimum temperature of 37 °C, and a 12 °C decrease in temperature will slow down the reaction rate. The fact that the two conversion values are comparable suggests that the rotating wheel is a better method to mix the reaction.

As stated, an incubated rotating wheel was not available for use. As time progressed it became apparent that an "in-batch" process may not be viable for the scale-up of this process. Even if this piece of apparatus was available, there was no guarantee that this would allow for the achievement of quantitative conversion in this system. Protein lifetime under the reaction conditions would always remain the same, and scaling up further would exacerbate the problem. It is possible that the enzyme is experiencing

product inhibition, and as the concentration of product increases in the solution the protein becomes less active. The high concentration of D-glucose 6-phosphate (**18**) being used in these reactions could possibly explain why this was not observed on a smaller scale. Alternatives to an "in-batch" process were therefore sought, and the immobilisation of *Tb*INO1 was identified as an attractive and accessible alternative.

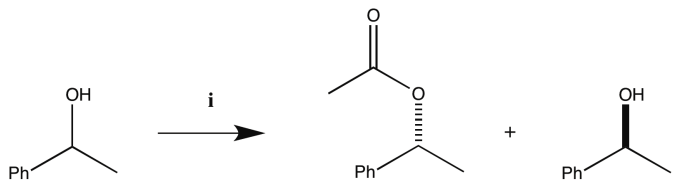
Chapter Three

3 Biocatalysis in Flow

3.1 Advantages of a Flow System

The use of continuous flow processes in biocatalysis has become more common in recent years. As the ability of scientists to modify and control enzymes has increased with techniques such as site-directed mutagenesis and directed evolution, the technology implemented in utilising these enzymes has needed to improve. Isolated enzymes are generally criticised as a synthetic method due to their instability and difficult recovery and re-use.[110] Immobilisation of an enzyme can overcome these issues. With the enzyme embedded on a solid it can be easily removed from reaction mixtures, therefore aiding enzyme re-use and purity of the product. Immobilised enzymes are normally more stable in general, with increased tolerance to temperature changes or organic solvents. This technique has greatly improved enzyme performance with higher catalyst productivities.[110]

There have been a number of enzymatic resolutions of racemic substrates using flow processes reported in recent years.[111] One such example is the kinetic resolution of racemic 1-phenylethanol (**(±) 110**, Scheme 3.1.1) by an immobilised lipase (*Candida antartica*, Novozym 435).[112]



(±) 110 **111** **(*S*)-110**

Scheme 3.1.1 – Kinetic resolution of 1-phenylethanol (**(±) 110**) using immobilised *Candida antartica*, Novozym 435.[112] *Reaents and conditions:* **i.** Immobilised lipase, vinyl acetate, $scCO_2$.

In continuous flow, this system provided the (*R*)-acetate **111** in 99.7 % e.e. and 47 % yield, and the (*S*)-alcohol **(*S*)-110** in 99.8 % e.e. and 47 % yield. At 25 mmol hr^{-1}, in a

5 mL flow reactor, 221 g of the (*S*)-alcohol **(*S*)-110** was produced over 3 days. This is a great improvement over the batch conditions, which could only produce 0.83 mmol in 7 hours.[111,112]

Immobilisation of *Tb*INO1 and conducting the reaction within a flow system has several advantages over an in-batch process. Firstly, it removes the issue of being unable to obtain full conversion within a certain time frame. The in-batch process was not able to maintain stability of the protein for long enough to facilitate desired substrate conversion. Issues with concentration and mixing meant that this was not a problem that could be easily overcome and would only be exacerbated on further scale up. With the introduction of a flow system, the reaction mixture should be passed through the system at a concentration that ensures full conversion in one pass. The exact concentration being dependent on the amount of *Tb*INO1 immobilised, the reactivity of the bound *Tb*INO1 and the flow rate of the system (how long the substrate has in contact with the protein). Once the immobilised protein loses it's activity, the reaction can be stopped. By collecting fractions from the column, it is possible to combine and continue with only those containing a fully converted product.

If product inhibition of the enzyme is an issue then conducting the reaction in a flow process will overcome this. As the substrate is converted into product, the product is then washed away from the immobilised enzyme, meaning that it cannot inhibit the reaction.

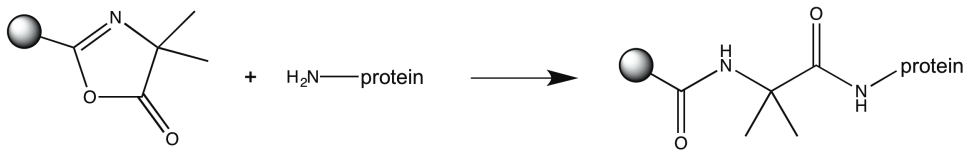
It is possible to automate flow systems with much less effort than in-batch processes, and they also lend themselves to scale-up procedures. As long as the ratio of substrate to protein remains the same and they have the same contact time, a flow system will easily accommodate scale up. Additional considerations, such as cofactor concentration, would also come into play, but these should be trivial to resolve.

Within the flow system, the *Tb*INO1 will be used to it's full potential. Using an in-batch process it may be necessary to use a non-catalytic amount or even excess of protein to effect full conversion. The flow system effectively matches the amount of substrate to the longevity and activity of the protein, allowing for a more efficient reaction stoichiometry and system. On top of these points, the immobilisation of the protein aids

the purification of the final product. The tedious process of removing protein from the reaction by acidification and filtration through centrifugal filter columns will be removed.

3.1.1 Resin Choice and Binding of *Tb*INO1

Originally, an azlactone-based polyacrylamide resin was trialled for the immobilisation of *Tb*INO1. This resin should bind to the protein (**113**, Scheme 3.1.1) through the azlactone residues situated on the beads (**112**). Following the manufacturer's protocol for protein immobilisation, the resulting beads had no activity in a test reaction. Increasing the contact time between the beads and protein within the binding step from 1 to 16 hours produced no improvement in the results. This suggests that the binding *Tb*INO1 via this method produced an inactive protein, or the protein was simply not binding to the beads.



112 **113** **114**

Scheme 3.1.2 – Attachment of a protein to an azlactone-based resin.

At the same time, another avenue was being explored. The His-Tag that was used for purification of the *Tb*INO1 had not been cleaved from the protein. It was hypothesised that this moiety could be used to immobilise the protein to a Ni^{2+}-Sepharose column, in very much the same way as when the protein is purified. This would not be as strong an attachment as the azlactone resin would produce as it is non-covalent, but it may be sufficient for proof of principle studies in this biocatalytic flow system.

Ni^{2+}-charged Sepharose beads that make up the protein purification columns were mixed with a *Tb*INO1 containing solution for 30 minutes. They were then washed with ammonium bicarbonate buffer (50 mM, pH 8.5) to remove any unattached protein. The prepared *Tb*INO1 beads were added to a reaction solution containing 10 mM D-glucose 6-phosphate (**18**), 2 mM NAD^+, 1 mM 2-mercaptoethanol, and 50 mM ammonium bicarbonate (pH 8.5). The DTT for protein stabilisation was replaced with 2-mercaptoethanol as the Ni^{2+} ions are known to react with DTT and this would hamper the reaction process. After incubation with shaking at 37 °C for 3 days, the IMPase assay

showed 80 % conversion. This was seen as sufficient evidence to begin working on the development of a flow system.

3.2 Experiments with a 1 mL Ni^{2+}-Sepharose Column

Within the flow system, the columns being used were pre-packed with Ni^{2+}- Sepharose and reusable. Initial work was conducted using a column size of 1 mL, which should hold up to 40 mg of protein as per the manufacturers specifications. All of the experiments being conducted were completed with an excess of *Tb*INO1 being loaded onto the column, in order to ensure the column was holding the full capacity of protein. Washing the protein from the 1 mL column after reaction and measuring the amount collected gave a reading of 39 mg of protein. Meaning that the columns were loaded and saturated with *Tb*INO1 protein during these experiments.

One of the main practical issues for the flow system was how to maintain the reaction at an optimal temperature of 37 °C. The first experiments conducted were done at room temperature as a method of determining the feasibility of this system, before more expensive equipment (a column oven) was invested in for the project. A solution containing D-glucose 6-phosphate (**18**), 2 mM NAD^+, 1 mM 2-mercaptoethanol and 50 mM ammonium bicarbonate (pH 8.5) was passed through the 1 mL Ni^{2+}-Sepharose column loaded with *Tb*INO1. A peristaltic pump powered the flow through the system, and the slowest setting on the pump was used in order to allow the substrate an increased contact time with the protein.

As experiments progressed, a gradual decrease in conversion was noted and this was attributed to the reuse of the Ni^{2+}-Sepharose columns. Whilst the drop in conversion was only slight, for the optimisation of the process a new column was used each time to achieve reproducible results. Each new column was treated the same way in the set up of the flow system. The Ni^{2+}-Sepharose column comes stored in a 20 % ethanol solution (aq.), which must be washed off prior to use. Passing five column volumes of water through the column is seen as a thorough wash of the column. This volume was used for each wash of the column being conducted. The columns were prepared by washing with water, followed by 50 mM ammonium bicarbonate (pH 8.5), water again, 200 mM nickel (II) sulfate solution, water, and finally 50 mM ammonium bicarbonate (pH 8.5). As

previously stated, the initial water wash is needed to remove the 20 % ethanol solution from the column. The first ammonium bicarbonate wash adjusts the pH of the column to 8.5 as desired for the reaction. The next water wash will remove any remaining ammonium bicarbonate solution. The wash with nickel (II) sulfate recharges the column with Ni^{2+} ions, and is generally conducted on a cycle for 1 hour to ensure the column is holding as many of these ions as possible. The following water wash removes any of the remaining nickel (II) sulfate solution, before the column can be washed with ammonium bicarbonate prior to protein loading. It was important to establish a procedure that would be used in each experiment, to avoid any inconsistencies that this may produce on results.

The first results at this temperature are shown in Table 3.2.1. This experiment was conducted by passing a 10 mM D-glucose 6-phosphate (**18**) solution through the *Tb*INO1 column at a flow rate of 2.5 mL hr^{-1}. A sample was taken and a second pass of the eluted solution was conducted through the column, before a third pass was completed. The eluting tube and feeding tube were then connected to make a closed system, which would slowly pump the reaction mixture through the column overnight, repeatedly.

Passes	Conversion
1	50 %
2	55 %
3	60 %
Cycle	60 %

Table 3.2.1 – Calculated conversion at room temperature with a 1 mL *Tb*INO1 loaded Ni^{2+}-Sepharose column and flow rate of 2.5 mL hr^{-1}. Reaction solution contained 10 mM D-glucose 6-phosphate (**18**), 2 mM NAD^{+}, 1 mM 2-mercaptoethanol, 50 mM ammonium bicarbonate (pH 8.5).

As can be seen from the results in Table 3.2.1, the first pass through the system was able to convert 50 % of the substrate. Passing the same solution through the column again increased the conversion to 55 %, and repeating the process again gained an extra 5 % conversion. The overnight cycle plateaued at 60 % conversion. This, coupled with the incremental increases in conversion noted in subsequent passes, suggests that possibly

product inhibition could be playing a role in the reaction. The immobilisation of the protein could lower the activity, due to potential conformational restriction of the enzyme, or where the active site is positioned in respect to where the protein is tethered. The lower temperature will undoubtedly cause a lower activity because the optimum temperature for the protein is 37 °C. It was hypothesised that by increasing the temperature of the system to 37 °C, the protein would be more active and potentially be able to overcome the product inhibition.

With no specific equipment to run the flow system at 37 °C, a system was trialled that contained a reservoir for the reaction solution. The idea was that by sitting the reservoir in a water bath at 37 °C, the heat would carry through the reaction and be sufficient to increase enzyme activity and achieve full conversion. After 24 hours of reaction, the calculated conversion was only 33 %. After 4 days, the conversion increased to 40 %, but no progress was observed after this time. In subsequent attempts, this procedure remained poor, with only a 45 % conversion being achieved at best.

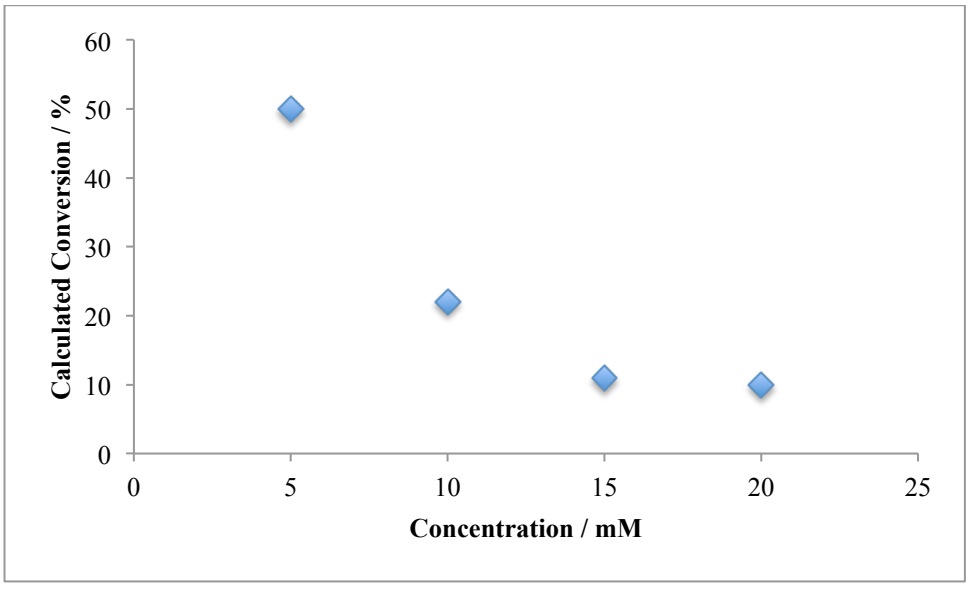
The introduction of a column oven for the *Tb*INO1 loaded Ni^{2+}-Sepharose column meant that progress could be made with the flow system. This provided a consistent method of maintaining the reaction at 37 °C and allowed for demonstration of an effective flow system. The first experiments conducted at this temperature demonstrated the systems ability to produce full conversion to L-*myo*-inositol 1-phosphate (**19**), but also highlighted the need to scale-up this process further.

Passes	Conversion
1	81 %
2	100 %
Cycle	100 %

Table 3.2.2 - Calculated conversion at 37 °C with a 1 mL *Tb*INO1 loaded Ni^{2+}-Sepharose column and flow rate of 2.5 mL hr^{-1}. Reaction solution contained 5 mM D-glucose 6-phosphate (**18**), 2 mM NAD^{+}, 1 mM 2-mercaptoethanol, 50 mM ammonium bicarbonate (pH 8.5).

Taking the same system used in the first 1 mL *Tb*INO1 loaded Ni^{2+}-Sepharose column experiments, a solution containing 5 mM D-glucose 6-phosphate (**18**) was passed through the system. The same solution was taken and passed through the system again, before being cycled through it overnight. The results are shown in Table 3.2.2.

A subsequent experiment completed single passes of varying D-glucose 6-phosphate (**18**) concentrations through the system. The same Ni^{2+}-Sepharose column was used throughout the experiment, with freshly made reaction solutions being passed through. Each of the different concentration solutions were thoroughly washed from the system before the new solution was passed through. It was ensured that samples were only collected from the system once a sufficient amount of each solution had passed through, maintaining the substrate / product concentration. The results are shown in Graph 3.2.1. It should be noted that this experiment was conducted on a non-fresh Ni^{2+}-Sepharose column, and this is the reason for the drop in conversion for the 5 mM result, and potential lower conversions for the other concentration results.



Graph 3.2.1 – The ability of the 1 mL *Tb*INO1 loaded Ni^{2+}-Sepharose flow system to convert different concentrations of D-glucose 6-phosphate (**18**) in one pass.

With the amount of protein remaining constant, as the substrate concentration is increased the percentage of the substrate being converted will be decreased. The protein is only able to turnover a certain number of moles of substrate in a single pass, and as the amount of substrate increases this limit will be reached. Once this is reached as the concentration

of substrate doubles, the % conversion will half, resulting in a graph that looks like Graph 3.2.1. It is possible that high amounts of substrate will cause substrate inhibition of the protein, but it does not look like this has been observed in this experiment as it is a flow system.

From these results it is possible to postulate that a 4 mM concentration of D-glucose 6-phosphate (**18**) should be able to pass through this system and achieve full conversion to L-*myo*-inositol 1-phosphate (**19**). At this concentration and rate of flow through the system, it is calculated that 42 mg of product would be produced within a 16 hour time period. This was extrapolated from the system being able to convert 81 % of a 5 mM D-glucose 6-phosphate (**18**) solution, and a flow rate of 2.5 mL hr^{-1}. With a scale up to a 5 mL column this amount should be increased to 210 mg, a much more synthetically viable quantity.

It should be noted that at this stage in the project, Sigma Aldrich discontinued their source of IMPase enzyme. This presented an issue because this was the source of the IMPase that was being used for the assay to determine conversion of the reaction. Fortunately, other members of the group were investigating the IMPase from *T. brucei*. This parasite has two forms of the IMPase protein, IMPase 1 and IMPase 2, which have been shown to work together in a 3:1 ratio to cleave the phosphate from L-*myo*-inositol 1-phosphate (**19**).[113] These two proteins could be expressed and purified for use within the IMPase assay used within this project. However, the low levels of these proteins being expressed and purified meant that this had to be conducted quite often and was quite time consuming. A new method to determine reaction conversion was sought, and with a scale-up in reaction this was possible.

3.3 Ni^{2+}-Sepharose Column Scale-Up

By scaling the process up to using a 5 mL Ni^{2+}-Sepharose column, the amount of material being passed through the system means that it is now suitable to use ^{1}H Nuclear Magnetic Resonance (NMR) to monitor the reaction. This means that the calculation of conversion will now be a direct measurement of the reaction, and not reliant upon the action of IMPase.

According to the manufacturer, the Ni^{2+}-Sepharose columns being used (IMAC Sepharose™ 6 Fast Flow) can hold 40 mg of protein per 1 mL of Sepharose. This means that a 5 mL column should be able to hold 200 mg of protein. The overexpression of *Tb*INO1 produces roughly 600 mg of protein per 8 L of culture. Therefore, by taking the *Tb*INO1 produced from 4 L of culture and loading this amount onto the 5 mL column, there will be an excess of protein and the column will be saturated. In the purification of *Tb*INO1, the protein is washed off the column with an increasing gradient of imidazole. As can be seen in Figure 3.3.1, the first wash containing 10 mM imidazole removes all of the undesired protein and cell extracts that remain on the column after loading. After this wash, all that remains on the column is the desired *Tb*INO1 protein. By leaving the protein on the column at this stage, the purification is effectively complete and column is ready for use in the flow system.

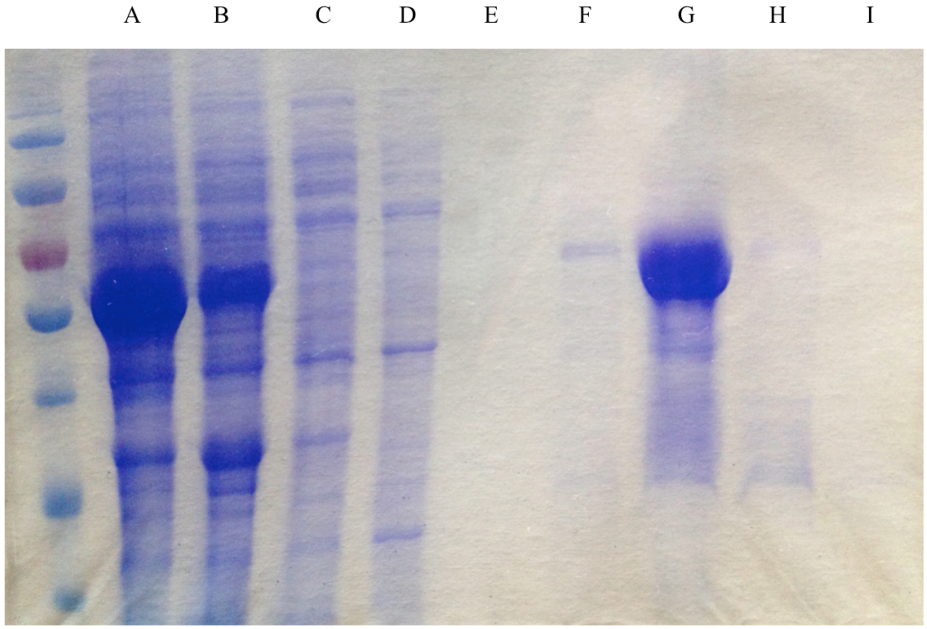


Figure 3.3.1 – Protein gel from a standard protein purification. A – Whole cell, B – supernatant, C – Flow through, D – 10 mM imidazole wash, E – 20 mM imidazole wash, F – 50 mM imidazole wash, G – 100 mM imidazole wash, H – 250 mM imidazole wash, I – 400 mM imidazole wash.

Aside from the larger *Tb*INO1 loaded Ni^{2+}-Sepharose column, the system used remained the same as the 1 mL experiments. A solution containing D-glucose 6-phosphate (**18**), 1 mM NAD^+, 1 mM 2-mercaptoethanol and 50 mM ammonium bicarbonate (pH 8.5) was

slowly passed (2.5 mL hr^{-1}, single pass) through the new Ni^{2+}-Sepharose column at 37 °C. With a 5 times increase in the amount of *Tb*INO1 present immobilised in the system, it was hypothesised that 5 times the amount of substrate could be converted by the system as the flow rate was remaining the same. This meant that if 4 mM D-glucose 6-phosphate (**18**) could be tolerated by the 1 mL system, then 20 mM D-glucose 6-phosphate (**18**) should be tolerated by the 5 mL system. During the first experiment, the concentration of D-glucose 6-phosphate (**18**) was initially chosen as 10 mM, and after 12 hours the concentration of this substrate was increased to 20 mM. This was to check that 10 mM could be tolerated, and then test 20 mM.

The reaction was then allowed to continue for 48 hours in total (Figure 3.3.2), with 4 fractions being collected: A – initial 12 hours with 10 mM D-glucose 6-phosphate (**18**), B – 12 to 24 hours with a switch to 20 mM D-glucose 6-phosphate (**18**), C – 24 to 36 hours with 20 mM D-glucose 6-phosphate (**18**), and D – 36 to 48 hours with 20 mM D-glucose 6-phosphate (**18**). Each of the fractions collected were acidified to pH 2 using concentrated hydrochloric acid before being lyophilised. NMR samples were prepared in D_2O of the resulting solid products.

As can be seen from Figure 3.3.2, the ^1H NMR of sample A shows 100 % conversion to L-*myo*-inositol 1-phosphate (**19**), with no D-glucose 6-phosphate (**18**) peaks present within the spectrum. The NMR of sample B looks like it may contain a small amount of D-glucose 6-phosphate (**18**). This is the sample that would have contained the remainder of the 10 mM D-glucose 6-phosphate (**18**) solution pass. This suggests that the flow system was not able to fully convert the 20 mM substrate solution in one pass, unless the protein had begun to lose its activity. Samples C and D show that as time progressed the *Tb*INO1 steadily lost activity, with sample D containing solely D-glucose 6-phosphate (**18**).

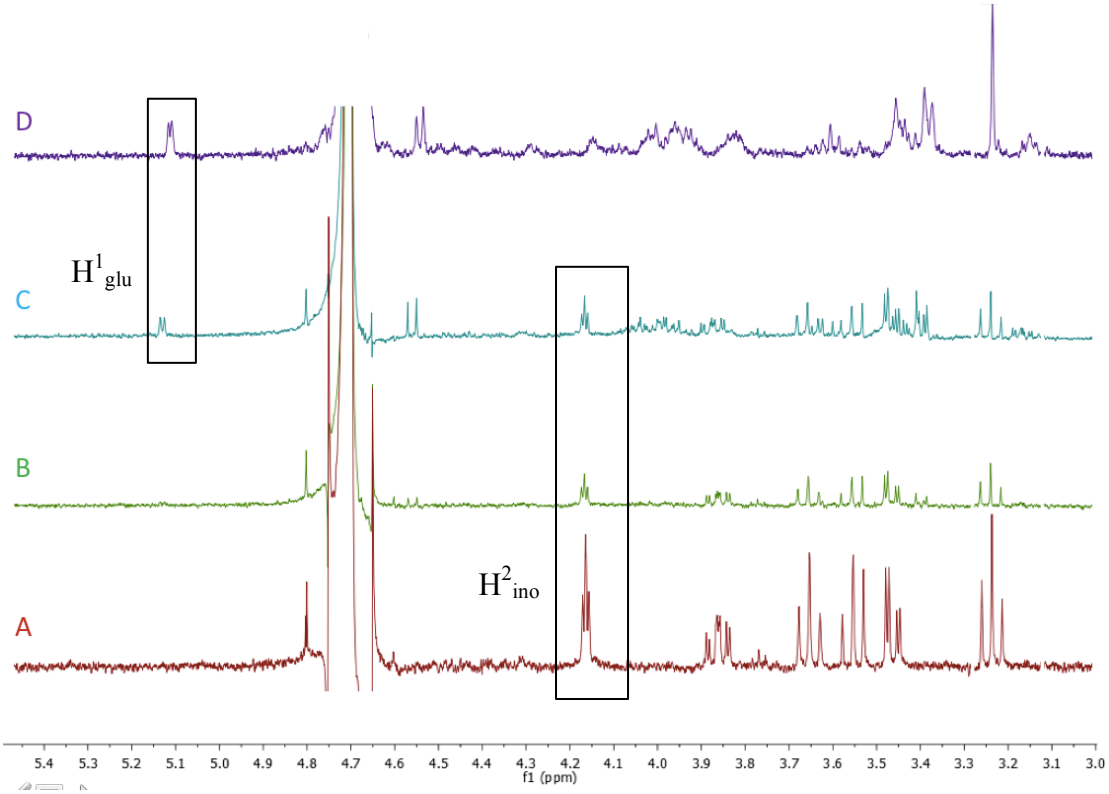


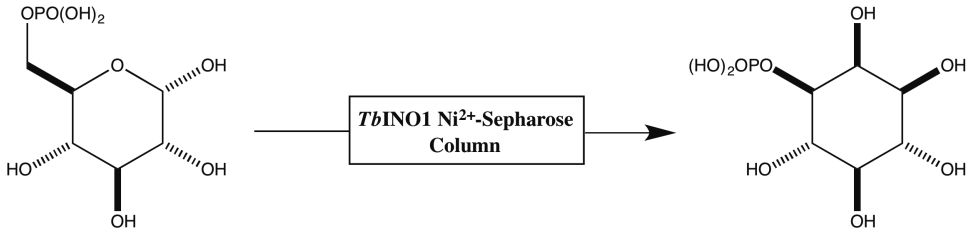
Figure 3.3.2 – ^{1}H NMR data showing the activity of immobilised *Tb*INO1 over time. A – initial 12 hours with 10 mM D-glucose 6-phosphate (**18**), B – 12 to 24 hours with a switch to 20 mM D-glucose 6-phosphate (**18**), C – 24 to 36 hours with 20 mM D-glucose 6-phosphate (**18**), and D – 36 to 48 hours with 20 mM D-glucose 6-phosphate (**18**).

It may have been the case that the drop in activity of the system was due to the protein being slowly stripped from the column. Tests using the Bradford reagent suggested that this was not the case, as the reagent did not turn blue when added to the solution eluting from the system. Coupled with the results from the experiment, this suggests that the protein begins to significantly lose activity after roughly 24 hours under these conditions. This is shorter than the in-batch reaction conditions, and this is thought to be due to the immobilisation of the protein. Although the *Tb*INO1 is less stable when bound to the column, the advantages that the flow system offers completely outweigh this issue.

Fraction	Conversion
0 – 10 hr	100 %
10 – 20 hr	100 %
20 – 30 hr	100 %
30 – 40 hr	58 %

Table 3.3.1 – Table showing protein activity as time passed. Solution containing 18 mM D-glucose 6-phosphate (**18**), 1 mM NAD^+, 1 mM 2-mercaptoethanol, 50 mM ammonium bicarbonate solution (aq., pH 8.5) was passed through the column at a flow rate of 2.5 mL hr^{-1} and 37 °C. Conversions were calculated from ^1H NMR data.

In subsequent experiments, 18 mM D-glucose 6-phosphate (**18**) was chosen as a concentration that the immobilised *Tb*INO1 could fully convert to L-*myo*-inositol 1-phosphate (**19**) in one pass. At this concentration the system is able to achieve full conversion for a time period of roughly 30 hours, after which the protein loses activity (Table 3.3.1). Fractions are collected every 10 hours and, before they are combined, the conversion of each is checked by ^1H NMR. Each flow reaction produces roughly 350 mg of L-*myo*-inositol 1-phosphate (**19**, Scheme 3.3.1).



18 **19**

Scheme 3.3.1 – Synthesis of L-*myo*-inositol 1-phosphate (**19**) using immobilised *Tb*INO1 (~200 μg) on a 5 mL Ni^{2+}-Sepharose column. Solution containing 18 mM D-glucose 6-phosphate (**18**), 1 mM NAD^+, 1 mM 2-mercaptoethanol, 50 mM ammonium bicarbonate solution (aq., pH 8.5) was passed through the column at a flow rate of 2.5 mL hr^{-1} and 37 °C.

3.3.1 Further Considerations

Throughout this process, one factor that needed to be considered was product isolation. The need to obtain complete conversion stems from a product isolation standpoint rather than a process efficiency one. Conducting the biotransformation "in-flow" has major advantages when it comes to isolation of the product. Obviously, the protein does not need to be removed from the reaction mixture, which bypasses one step in purification. By collecting fractions from the biotransformation it is possible to deduce the exact point when the protein is unable to turnover all of the substrate, so only fractions containing no D-glucose 6-phosphate (**18**) can be collected.

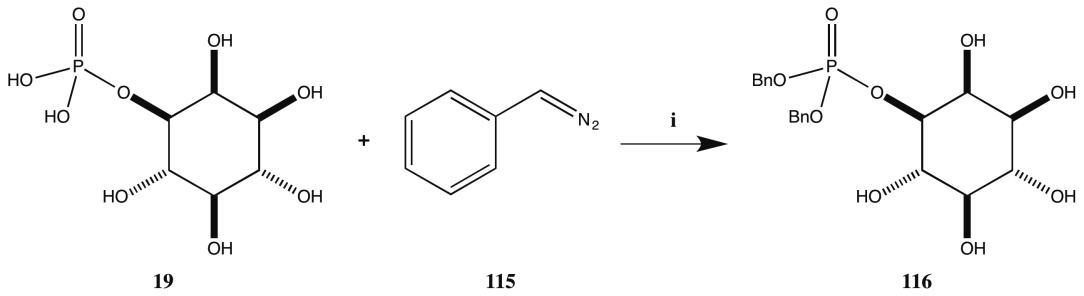
The solution that elutes from the column is acidified to pH 2 using concentrated hydrochloric acid. This ensures that the L-*myo*-inositol 1-phosphate (**19**) is in the free acid form, and not the sodium salt, which is necessary for the next step of the process. Use of Amberlite H^+ was also trialled, which was efficient at forming the free acid. However, using this method had no added benefits and ultimately just led to more water having to be removed from the final product.

After acidification, the solution is lyophilised to remove the water, buffer salt and 2-mercaptoethanol. Having previously considered the isolation of L-*myo*-inositol 1-phosphate (**19**) in the scale up of the process, the product obtained from this process is relatively pure with the only real contaminant as NAD^+/NADH. As soon as any synthetic manipulation is conducted and subsequent purification takes place, this contaminant will be removed.

3.4 Selective Phosphate Benzylation

One potential problem with using L-*myo*-inositol 1-phosphate (**19**) as a starting material is the reactivity of the phosphate. We obviously need this phosphate present, at least until substitution of the inositol ring in order to retain optical purity. To avoid complications with the phosphate, it was decided to protect it using benzyl groups that would also allow for better organic solvent solubility. Phenyldiazomethane (**115**, $PhCHN_2$) was highlighted as a potential reagent to use in order to achieve selective phosphate

benzylation and produce 1-bis(benzyloxy)phosphoryl L-*myo*-inositol (**116**, Scheme 3.4.1). There is precedence in the literature of diazomethane as a selective methylating reagent of phosphates in the presence of alcohols.[114,115]

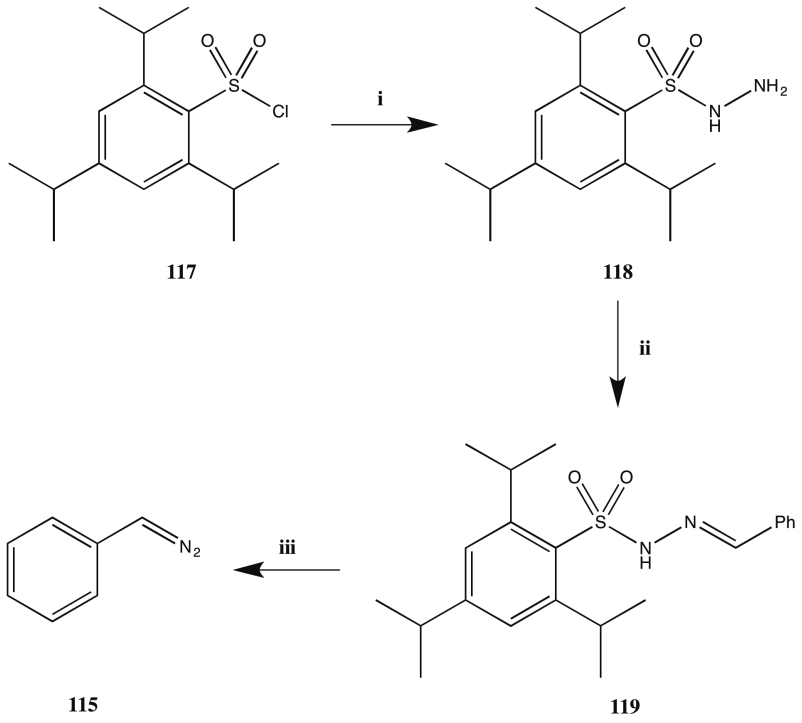


Scheme 3.4.1 – Proposed selective benzylation using phenyldiazomethane (**115**). *Reagents and conditions:* **i**. Et_2O, 0 °C.

Phenyldiazomethane (**115**) was synthesised via the route outlined in Scheme 3.4.2. Starting from the sulfonyl chloride **117**, addition of hydrazine afforded the hydrazide **118**. This could then be converted into hydrazide **119** through reaction with benzaldehyde. Phenyldiazomethane (**115**) was synthesised as needed, which was feasible as the formation of this reagent only requires reflux in methanol with potassium hydroxide for 20 minutes.[116] In initial experiments, the reaction was cooled to room temperature and ice-cold water was added to quench the reaction. The reagent was then extracted into diethyl ether and used as a crude ethereal solution. In subsequent work, washing the combined diethyl ether layers with sodium carbonate solution (sat., aq.), water and brine gave a much cleaner reaction for the benzylation.

Before conducting the benzylation on the highly precious L-*myo*-inositol 1-phosphate (**19**), preliminary experiments were conducted on D-glucose 6-phosphate (**18**). This allowed for development of the benzylation process on a material that is very similar in terms of solubility and reactivity. As previously stated, for reaction with a diazo compound, the phosphate must be present in its free acid form. The disodium salt of D-glucose 6-phosphate (**18**) was dissolved in water, acidified to pH 2 using concentrated hydrochloric acid and finally lyophilised to furnish the free acid. On a small scale (<10 mg) the benzylation could be accomplished via addition of an ethereal solution of phenyldiazomethane in small aliquots at 0 °C, until the orange colour of the reagent persisted (showing an excess of reagent). After 10 minutes at this temperature, the

reaction was allowed to warm to room temperature and the highly volatile reagent was able to evaporate from the reaction. Mass spectrometry showed the success of the benzylation on this small scale, with presence of a peak for 1-bis(benzyloxy)phosphoryl L-*myo*-inositol (**116**) and no peaks representing starting material or monobenzylated products within the data.



Scheme 3.4.2 – Synthesis of phenyldiazomethane (**115**). *Reagents and conditions:* **i**. Hydrazine monohydrate, THF, 0 °C, 4 hr, 95 %; **ii**. benzaldehyde, MeOH, 5 °C, 16 hr, 91 %; **iii**. KOH, MeOH, reflux, 30 min.

On a larger scale (>20 mg) incomplete benzylation of the phosphate was observed (Table 3.4.1), with mass spectrometry data showing only a trace of benzylated products compared D-glucose 6-phosphate (**18**). This is thought to be due to poor mixing of the ethereal reagent solution and the insoluble D-glucose 6-phosphate (**18**). A biphasic reaction with an aqueous solution of D-glucose 6-phosphate (**18**) was attempted with vigorous mixing. Unfortunately, this did not solve the problem. Dissolving the D-glucose 6-phosphate (**18**) in methanol had a better effect on the reaction, allowing for some benzylation to occur, (16 % monobenzylation and 16 % dibenzylation). Addition of the D-glucose 6-phosphate (**18**) to the reaction mixture would not be productive as the reaction mixture is highly basic and would convert the phosphate away from it's reactive

free phosphate form. It could also lead to cleavage of the phosphate completely, with hydroxide ions being present in excess.

$PhCHN_2$ Solvent	$PhCHN_2$ washed?	D-Glucose 6-phosphate solvent	Amount of monobenzylation	Amount of dibenzylation
Et_2O	No	Neat	Trace	Trace
Et_2O	No	Water	None	None
Et_2O	Yes	Methanol	16 %	16 %
DCM	Yes	Methanol	15 %	15 %
Et_2O	Yes	Methanol (high dilution)	None	>99 %

Table 3.4.1 – Reagents and conditions used in the selective phosphate benzylation of D-glucose 6-phosphate (**18**) using phenyldiazomethane ($PhCHN_2$, **115**). Conversions to monobenzylated and dibenzylated products were determined using ^1H NMR spectroscopy.

Introduction of partial purification of the phenyldiazomethane (**115**) by washing with sodium carbonate solution (sat., aq.), water and brine, had a slight improvement on the reaction outcome. Extraction of the phenyldiazomethane into dichloromethane instead of diethyl ether could have been advantageous for dissolution within methanol, however this had little to no effect on reactivity. With both diethyl ether and dichloromethane as solvents for the reagent it was noted that upon addition of the phenyldiazomethane solution, the D-glucose 6-phosphate (**18**) crashed out of the methanol and diethyl ether or dichloromethane mixture. With the hypothesis that this was responsible for the incomplete benzylation, a very highly diluted solution of D-glucose 6-phosphate (**18**) was used in an attempt to keep it in solution throughout the reaction. Upon addition of ethereal phenyldiazomethane, the D-glucose 6-phosphate (**18**) remained in solution, and complete phosphate benzylation was observed by mass spectrometry.

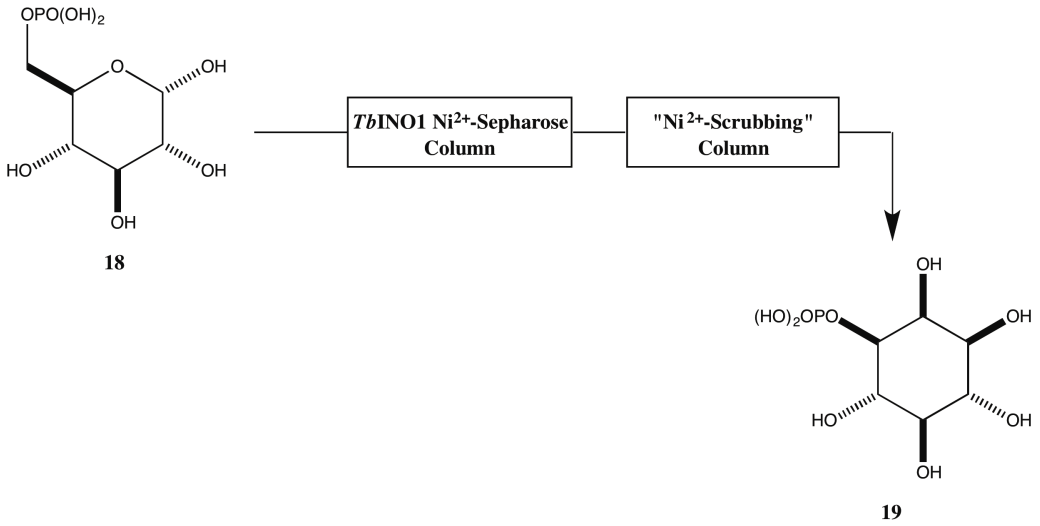
With a method developed for selective phosphate benzylation, it was possible to move onto the actual substrate for the reaction, L-*myo*-inositol 1-phosphate (**19**) from the biotransformation. The crude material obtained from the biocatalytic flow system was treated with phenyldiazomethane under the same conditions developed for the

benzylation of D-glucose 6-phosphate (**18**). Unfortunately, this process did not achieve any benzylation, and only starting material was recovered from the reaction.

With this starting material, achieving any benzylation at all proved to be an issue. Methods using benzyl bromide and a base, such as Et_3N or diisopropylethylamine (Hünig's base), failed to achieve benzylation at any position of the L-*myo*-inositol phosphate (**19**). Even moving to the highly reactive bromomethyl acetate failed to furnish a benzylated product.

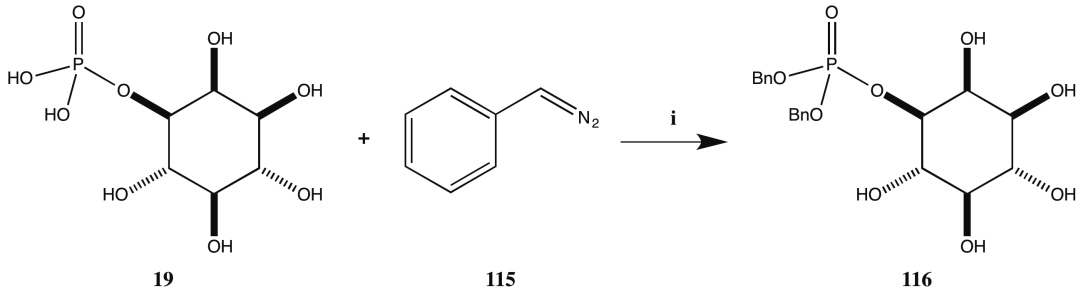
With all of these benzylation methods failing, it was hypothesised that there must be some artifact from the biocatalytic step that was interrupting subsequent reactions. This was supported by the fact that unreacted D-glucose 6-phosphate (**18**) from the biocatalytic step also did not react under the same conditions that benzylated the substrate previously. The reaction solution was passed through the system after *Tb*INO1 had lost its activity. This sample was then treated in the same way that the L-*myo*-inositol 1-phosphate (**19**) had been (acidification to pH 2 and lyophilisation). The resulting product was subjected to the benzylation reaction and furnished none of the desired product. Within the flow system, the His-tagged *Tb*INO1 protein is tethered to the Ni^{2+}-Sepharose column via coordination to the Ni^{2+} ions that are embedded within the Sepharose's nitrilotriacetic acid (NTA) residues. It is possible that within the reaction conditions, the Ni^{2+} ions could leach from the Sepharose column, and subsequently contaminate the L-*myo*-inositol 1-phosphate (**19**) produced.

A simple test and remedy for this problem was found by placing a Ni^{2+} stripped Sepharose column at the end of the flow system (Scheme 3.4.3). This "Ni^{2+}-scrubbing" column was prepared by washing a Ni^{2+}-Sepharose column with 50 mM ethylenediaminetetraacetic acid (EDTA) solution, which binds and removes any Ni^{2+} ions from the column. The L-*myo*-inositol 1-phosphate (**19**) produced from the modified flow system, was successfully benzylated using the previously developed phenyldiazomethane procedure.



Scheme 3.4.3 – Synthesis of L-*myo*-inositol 1-phosphate (**19**) using immobilised *Tb*INO1 (~200 μg) on a 5 mL Ni^{2+}-Sepharose column. Solution of 18 mM D-glucose 6-phosphate (**18**), 1 mM NAD^+, 1 mM 2-mercaptoethanol, 50 mM ammonium bicarbonate solution (aq., pH 8.5) was passed through the column at 2.5 mL hr^{-1} and 37 °C. Including a "Ni^{2+}-scrubbing" column in the flow set up allows for removal of unwanted Ni^{2+} ions.

The benzylation of L-*myo*-inositol 1-phosphate (**19**) was optimised to produce a 97 % yield of 1-bis(benzyloxy)phosphoryl L-*myo*-inositol (**116**, Scheme 3.4.4). The modified flow system produced a white fluffy solid after lyophilisation rather than the observed dark brown or black product achieved from earlier attempts. When dissolving the L-*myo*-inositol 1-phosphate (**19**) in methanol for the benzylation, it was noted that a significant proportion of the white solid did not dissolve. NMR data showed that this was NAD^+ or NADH, which are only sparingly soluble in methanol. A quick filtration therefore led to the removal of the vast majority of this from the L-*myo*-inositol 1-phosphate (**19**).



Scheme 3.4.4 – Selective benzylation of L-*myo*-inositol 1-phosphate (**19**) using phenyldiazomethane (**115**). *Reagents and conditions:* **i**. Et_2O, MeOH (large excess), 0 °C, 97 %.

With the L-*myo*-inositol 1-phosphate (**19**) dissolved in a large volume of methanol (300 mL per 100 mg), a freshly prepared and washed ethereal solution of phenyldiazomethane could be added in excess at 0 °C without precipitation of the starting material. Stirring the reaction mixture for 4 hours allowed complete benzylation of the phosphate, with selectivity over the inositol hydroxyl groups. The reaction was quenched by addition of acetic acid before being concentrated. At this stage, side products from the formation of phenyldiazomethane remained within the crude product. These could be removed by dissolution of the product in water, and washing with diethyl ether. Lyophilisation then produced substantially pure 1-bis(benzyloxy)phosphoryl L-*myo*-inositol (**116**).

The L-*myo*-inositol 1-phosphate (**19**) produced from this methodology is achieved with complete conversion. As a result of this, and the way in which it is worked up, the L-*myo*-inositol 1-phosphate (**19**) produced is highly pure, as can be seen in the ^{1}H NMR in Figure 3.4.1.

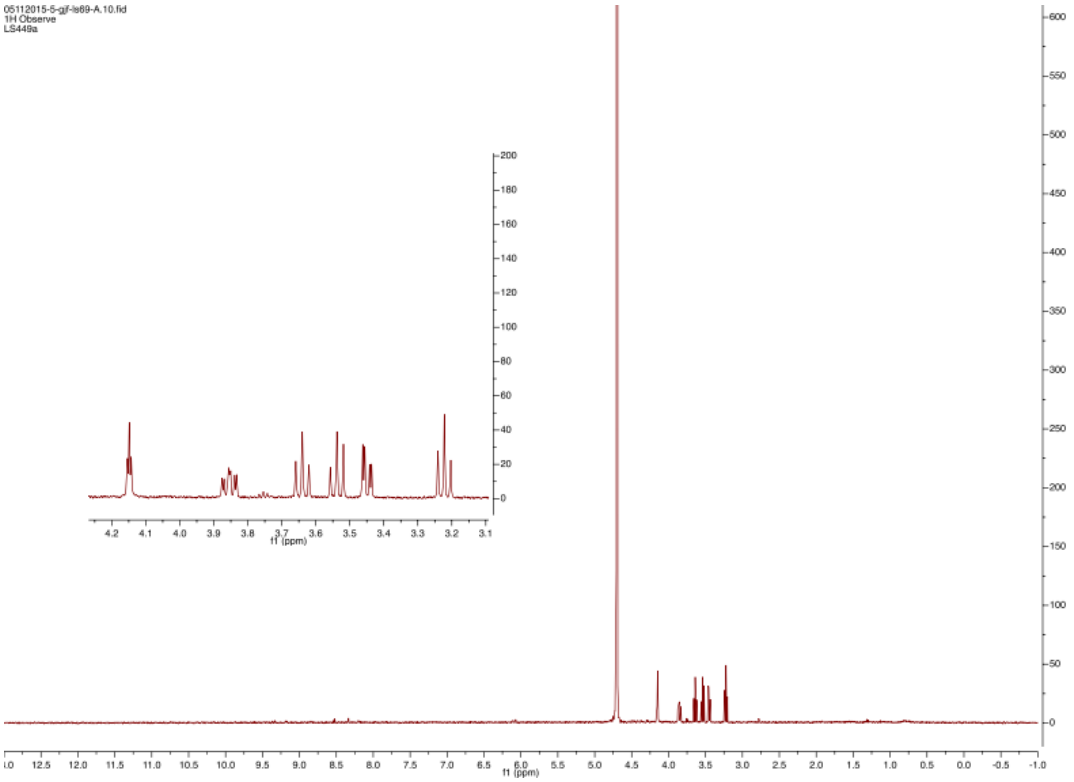


Figure 3.4.1 – ^{1}H NMR of L-*myo*-inositol 1-phosphate (**19**) obtained from the biocatalytic flow system.

Possibly the most important factor to consider with the product is the enantiopurity. The literature suggests[57,117] that the measurement of optical rotation of inositol phosphates is dependent upon the pH of the solution being measured. The L-*myo*-inositol 1-phosphate (**19**) produced in this system has an $[\alpha]_D$ of -4.6 (pH 9), which is in reasonable agreement with the measurement of Miller *et al.*,[69] $[\alpha]_D$ -3.4 (pH 9).

3.5 Secondary Function of INO1

The versatility of *myo*-inositol (**1**) allows the compound and its derivatives to play central roles within a plethora of cellular processes. This is a phenomenon that is seen across nature and, as such, INO1 genes have been characterised from a small number of prokaryotic and various eukaryotic organisms. These include *Porteresia coarctata*,[118] *Arabidopsis thaliana*,[108] *Mycobacterium tuberculosis* (*M. tuberculosis)*,[119,120] *Mycobacterium smegmatis*,[121] *Drosphila melanogaster*,[122] *S. cerevisiae*,[107] *Entamoeba histolytica*,[123] *Leishmania mexicana*[124] and *Homo sapiens*.[105,125] It has been revealed by genetic studies that within both *Leishmania mexicana*[124] and *M. tuberculosis*[120] INO1 plays a major role in regulation of *myo*-inositol (**1**) levels. In addition to *T. brucei*,[22] this highlights INO1 as a potential drug target for *Leishmania mexicana* and *M. tuberculosis*.

In work being conducted parallel to this project, the Smith group was screening compounds as potential inhibitors of *Tb*INO1. They were initially using the IMPase assay to determine reaction conversion and inhibition data for the *Tb*INO1 reaction. However, this assay had two major disadvantages within this work: potential inhibition of IMPase and the high levels of free phosphate in some of the potential inhibitors being tested. They decided to employ high performance ion-exchange chromatography to quantify the enzyme activity and overcome these issues.

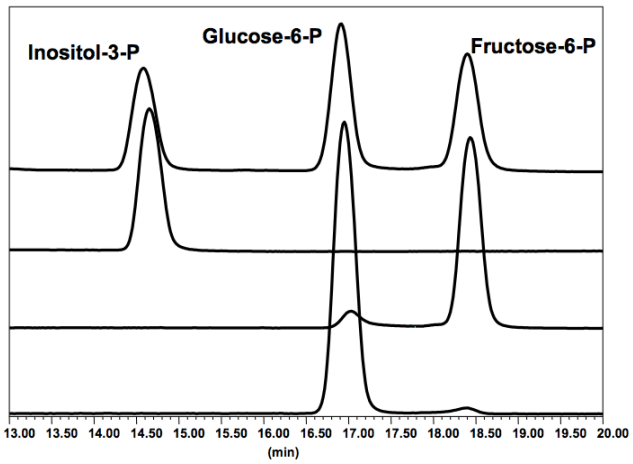
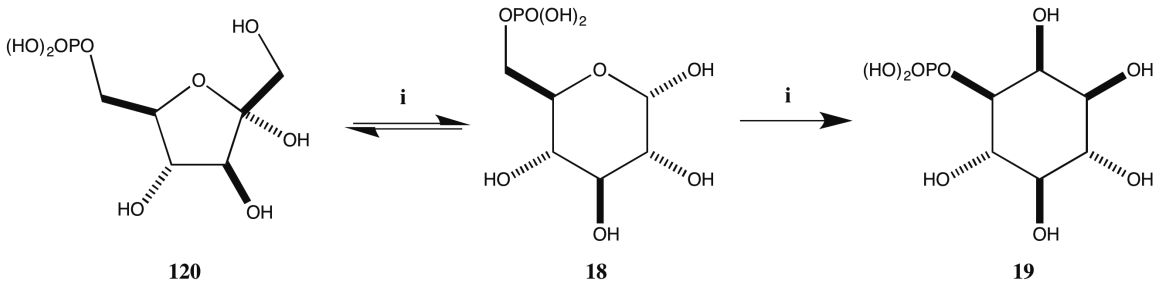


Figure 3.5.1 – Reaction standards D-glucose 6-phosphate (**18**), D-fructose 6-phosphate (**120**) and L-*myo*-inositol 1-phosphate (**19**) by high performance anion-exchange chromatography using a Dionex HPLC system with a CarboPacPA-1 column and PA-1 guard column (Dionex).

This method was able to separate D-glucose 6-phosphate (**18**) and L-*myo*-inositol 1-phosphate (**19**), as well as the other components of the reaction mixture, allowing them to obtain a quantitative reading of conversion. Within these experiments, they often observed an additional peak in the spectra from the assays. This peak was not present within control reactions for enzyme and substrate blanks (Figure 3.5.1), which suggests that this peak was present as a result of enzyme activity. It correlates with the expected peak for D-fructose 6-phosphate (**120**, Scheme 3.5.1), and suggests that *Tb*INO1 was able to form D-fructose 6-phosphate (**120**) in addition to L-*myo*-inositol 1-phosphate (**19**).



120 **18** **19**

Scheme 3.5.1 – Conversion of D-fructose 6-phosphate (**120**) into D-glucose 6-phosphate (**18**) and L-*myo*-inositol 1-phosphate (**19**) via action of *Tb*INO1. *Reagents and conditions:* **i**. *Tb*INO1, NAD^+.

By switching the substrate for the reaction to D-fructose 6-phosphate (**120**) they were able to observe peaks with identical elution times to D-fructose 6-phosphate (**120**), D-glucose 6-phosphate (**18**) and L-*myo*-inositol 1-phosphate (**19**, Figure 3.5.2). Extension of the reaction for longer periods only produced L-*myo*-inositol 1-phosphate (**19**). So *Tb*INO1 is able to convert D-fructose 6-phosphate (**114**) into L-*myo*-inositol 1-phosphate (**19**), via conversion to D-glucose 6-phosphate (**18**). It is thought that the production of D-glucose 6-phosphate (**18**) is reversible due to the initial discovery of D-fructose 6-phosphate (**114**) being produced in the reaction.

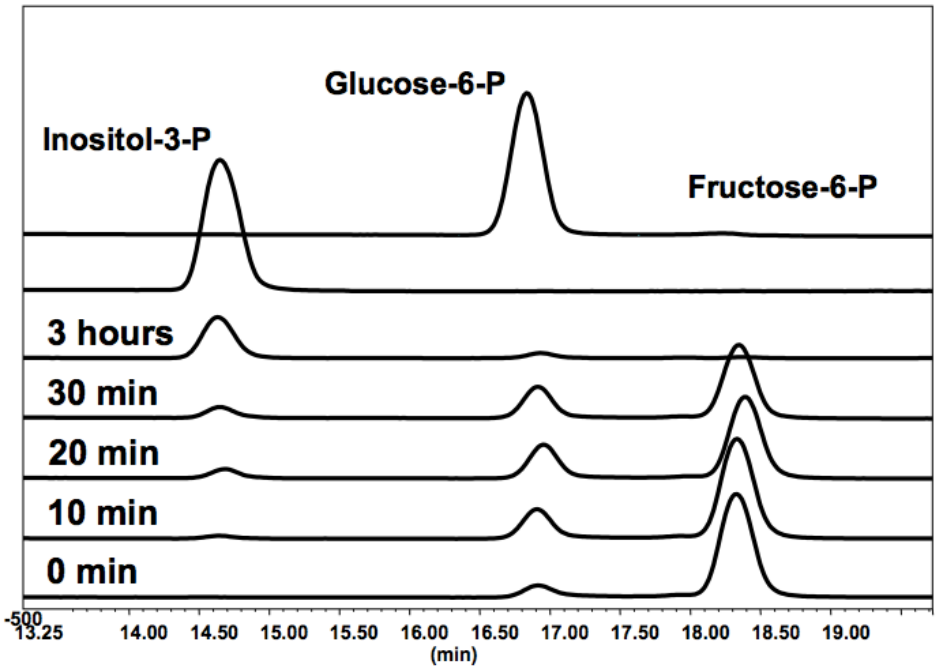


Figure 3.5.2 – Conversion of D-fructose 6-phosphate (**120**) to both D-glucose 6-phosphate (**18**) and L-*myo*-inositol 1-phosphate (**19**), shown by high performance anion-exchange chromatography using a Dionex HPLC system with a CarboPacPA-1 column and PA-1 guard column (Dionex). Reaction mixture contained 2 mM Tris acetate (pH 8), 2 mM D-fructose 6-phosphate (**120**), 1 mM NAD^+, 1 mM DTT, 2 mM ammonium acetate and 5 μg *Tb*INO1 in a final volume of 20 μL. The reaction was incubated at 37 ^oC. Samples were taken at selected time points and terminated by heating at 100 ^oC for 10 minutes and the addition of 80 μL of 10 mM NaOH (aq.).

The flow system developed within this project was being conducted on a scale large enough for ^1H NMR to be used as a method of conversion determination. Feeding D-fructose 6-phosphate (**120**) into the system in place of D-glucose 6-phosphate (**18**),

provided further demonstration of the protein's ability to convert D-fructose 6-phosphate (**120**) into L-*myo*-inositol 1-phosphate (**19**) via D-glucose 6-phosphate (**18**). The biocatalytic flow system was set up as usual and the substrate was replaced in the reaction mixture by D-fructose 6-phosphate (**120**) at a 15 mM concentration. By conducting a slow pass through the system, at the normal rate of 2.5 mL hr^{-1}, conversion to L-*myo*-inositol 1-phosphate (**19**) was observed by ^{1}H NMR (Spectrum A, Figure 3.5.3). Increasing the rate to 7.5 mL hr^{-1} allowed partial conversion of the substrate into D-glucose 6-phosphate (**18**) and L-*myo*-inositol 1-phosphate (**19**, Spectrum B).

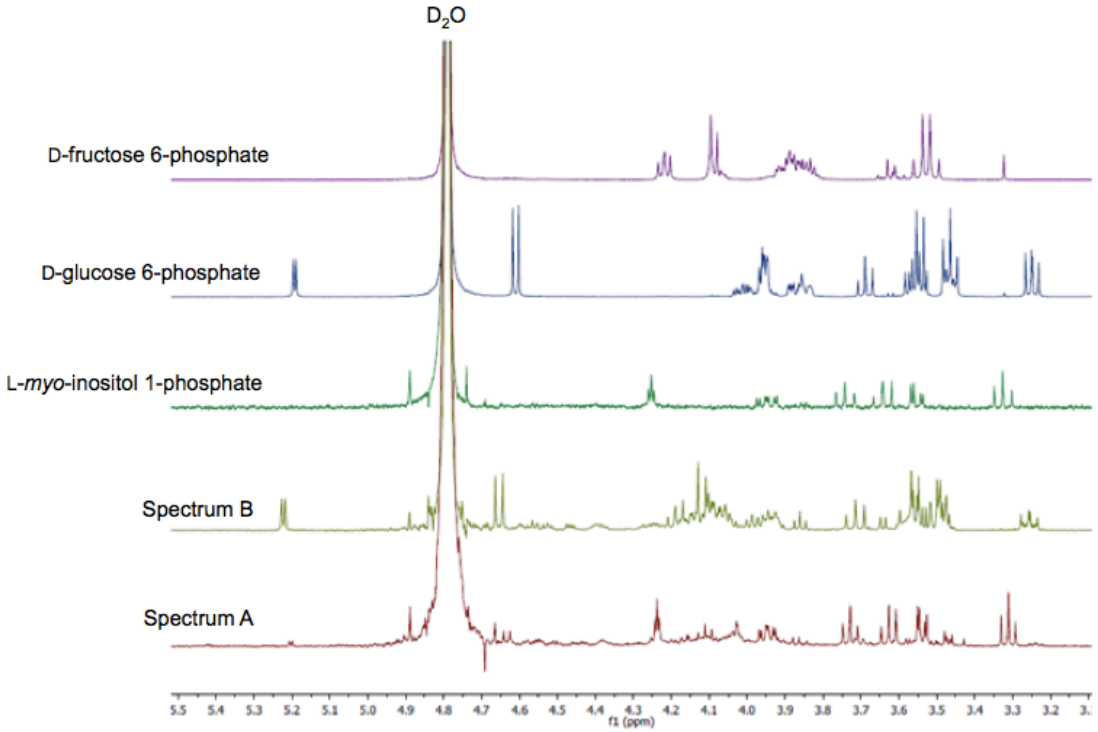


Figure 3.5.3 – ^{1}H NMR spectra for standards D-fructose 6-phosphate (**120**), D-glucose 6-phosphate (**18**) and L-*myo*-inositol 1-phosphate (**19**), and products of the reaction conducted at different flow rates through the system, 2.5 mL hr^{-1} (Spectrum A) and 7.5 mL hr^{-1} (Spectrum B).

These results show that the recombinant *Tb*INO1 is able to catalyse the reversible aldose-ketose isomerisation of D-glucose 6-phosphate (**18**) to D-fructose 6-phosphate (**120**). This activity is normally only observed by D-glucose 6-phosphate isomerases,[126] and has not been reported for INO1s from any other organisms. The scale of the biocatalysis in the flow system has allowed NMR spectroscopy to further show this transformation. Also, D-fructose 6-phosphate (**120**) could be used instead of D-glucose 6-phosphate (**18**)

as a starting material for the production of L-*myo*-inositol 1-phosphate (**19**) if that was advantageous.

3.6 Conclusions

The biocatalytic production of L-*myo*-inositol 1-phosphate (**19**) on a synthetically viable scale has been a success. Using a flow system for this process has dramatically improved the scale up of the process, and has meant that further scale up of the process can be conducted more easily. Use of a larger Ni^{2+}-Sepharose column or setting up two columns in series, would allow for a higher concentration and larger amounts of D-glucose 6-phosphate (**18**) to be passed through and be fully converted into L-*myo*-inositol 1-phosphate (**19**).

The process as a whole, whilst able to produce significant quantities of L-*myo*-inositol 1-phosphate (**19**), is not fully optimised. Further work would be necessary to make this a commercially viable process. The protein is very crudely immobilised, with the His-tag being used to attach it to a Ni^{2+}-Sepharose column non-covalently. There are many methods available to immobilise proteins and these should be fully investigated to achieve the most efficient system possible.

Enzymes can be immobilised by binding to a support, entrapment in a polymer network, or by cross-linking of enzyme aggregates or crystals.[109] Enzymes can be bound to a support using physical, ionic or covalent interactions. The type of supports involved are generally resins, biopolymers, or inorganic solids, such as silicas or zeolites. Physical binding is generally too weak to be used in industrial applications.[109] In this work, *Tb*INO1 is bound to a Ni^{2+}-Sepharose column in an ionic binding to a resin. This type of interaction is stronger than physical, but in this system Ni^{2+} leaching was observed so a new ionic method or covalent method of enzyme attachment should be investigated.

The current system uses a large amount of NAD^+ within the reaction solution, and whilst this can be easily removed from the product, it may be possible to remove it and recycle it. Aside from the incorporation of a His-tag into the protein, the *Tb*INO1 gene being used has not been modified in any way. It is now possible to modify proteins to allow for

a wider substrate scope, have activity within a larger range of solvents and temperatures, improve lifetime, or switch the stereoselectivity of the reaction.[127] Directed evolution could be used to produce an INO1 gene that produces an enzyme that is much more stable or active. With no crystal structure of *Tb*INO1 available at this time, design and mutagenesis of specific sites would not be viable.

The starting material for the biocatalysis, D-glucose 6-phosphate (**18**), is inexpensive when compared to the highly valuable product that it is used to produce. However, within a commercial setting it is always better to start from a material that is as low in cost as possible. D-Glucose 6-phosphate (**18**) could be obtained from a glucose phosphorylase and fed directly into the *Tb*INO1 reaction. The D-glucose itself could even be produced from starch. This really would be a cheap starting material that could be converted into a highly valuable product in as little as three steps. It is probable that the most expensive component for this process would be adenosine triphosphate (ATP) in the phosphorylation of D-glucose. However, methods can be used in parallel to recycle cofactors, such as ATP, and produce an efficient system.[128]

Chapter Four

4 Studies Towards The Synthesis Of *epi*-Mycothiol

4.1 Introduction

The actinomycetes form a diverse group of Gram-positive bacteria containing species that are of high pharmacological interest.[129] A member of this family of organisms is mycobacteria, which contain a number of pathogenic organisms that cause serious diseases in mammals, including *M. tuberculosis*, *Mycobacterium leprae*, *Mycobacterium bovis* and *Mycobacterium avium*.[130] The disease tuberculosis (TB) is caused by *M. tuberculosis*, and remains a major global health problem. An estimated 9 million people developed TB in 2013 and it is second to only human immunodeficiency virus (HIV) as the leading cause of death from an infectious disease worldwide (1.5 million deaths estimated in 2013).[131]

Treatment of diseases that are caused by mycobacteria is very difficult.[130] These organisms are incredibly adept at surviving in macrophages and can tolerate a variety of environmental conditions. Changes in oxidative levels caused by oxidative stress or hypoxia are well tolerated by these organisms, as well as varying pH conditions and changes to metal ion concentrations. Variation of available nutrients can be easily dealt with, and they are also well protected against reactive oxygen and nitrogen species.[130,132,133] Treatment for TB therefore requires multidrug therapy over extended periods of time (typically 6-9 months) and it is often unsuccessful.[134,135]

The first drug used to treat TB was streptomycin, which was introduced in 1944. Drug resistance emerged not long after its introduction, thought mainly to be due to its use as a monotherapy. Several other drugs have now been identified as having anti-TB activity, and it is now common practise to use multidrug therapy to treat the disease.[136] The current recommended treatment for drug-susceptible TB requires a minimum of 6 months of treatment with two phases: the intensive phase of taking four different drugs (isoniazid (**121**, Figure 4.1.1), rifampicin (**122**), pyrazinamide (**123**) and ethambutol (**124**)) for a 2 month period, and the continuation phase, typically consisting of taking isoniazid and rifampicin for 4 months.[137] This therapy achieves cure rates of >95 % when administered

under directly observed therapy. There are, however, significant issues associated with the current therapeutic regime: drug intolerance and associated toxicities, pharmacokinetic drug-drug interactions (particularly amongst patients co-infected with TB and HIV), and, due to the lengthy treatment duration needed for a non-relapsing cure, patient adherence to the regimen.[137]

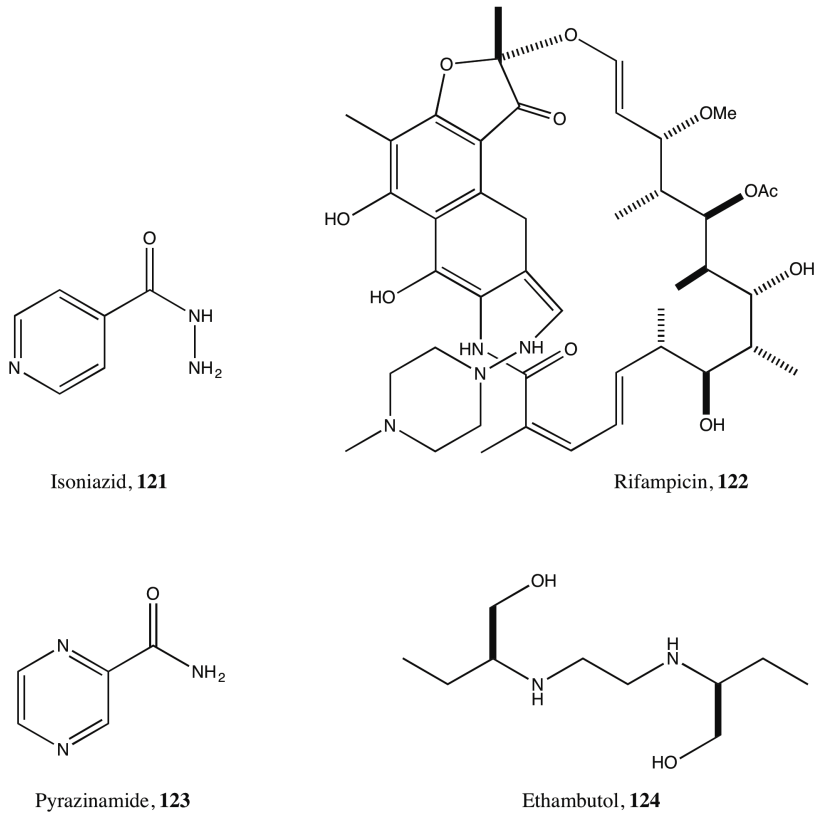


Figure 4.1.1 – Drugs currently used in the first-line treatment of TB.

One of the biggest obstacles in the treatment of TB is drug resistance. Multi-drug-resistant TB (MDR-TB) is caused by strains of *M. tuberculosis* that are resistant to at least rifampicin and isoniazid.[136] It can be treated by the second-line drugs: kanamycin, capreomycin, amikacin, *para*-aminosalicylic acid, cycloserine, terizidone, thionamine, protionamide and fluoroquinolones: levofloxacin, moxifloxacin and ofloxacin (Table 4.1.1).[138] 3.5 % of new and 20.5 % of previously treated TB cases were determined to have MDR-TB in 2013, giving a total of 480,000 people with MDR-TB in 2013.[131] Out of these patients, 9.0 % were determined to have extensively drug-resistant TB (XDR-TB).[131] This is categorised by TB strains that are resistant to both rifampicin and isoniazid, and are further resistant to any of the fluoroquinolones and to at least one of the second-line drugs kanamycin, capreomycin and amikacin.[136] With aggressive treatment

regimens it is possible to achieve cure rates of up to 60 % of patients with XDR-TB. Instances of drug resistance beyond that of XDR-TB have been described and highlight the need for new drugs for drug-resistant TB.[137]

First-line	Second-line
Isoniazid	Kanamycin
Rifampicin	Capreomycin
Pyrazinamide	Amikacin
Ethambutol	*para*-Aminosalicylic acid
Streptomycin	Cycloserine
	Terizidone
	Thionamide
	Protionamine
	Fluoroquinolones: levofloxacin, moxifloxacin and ofloxacin.

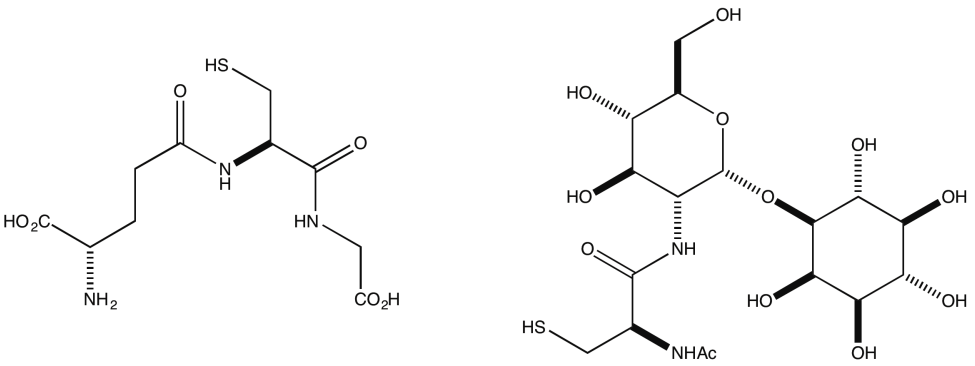
Table 4.1.1 - Drugs used to treat TB, MDR-TB and XDR-TB.

4.1.1 Mycothiol

Many organisms make use of low molecular weight thiols to help cells deal with oxidative stress and maintain redox homeostasis.[139] These molecules primarily act as redox buffers and reducing cofactors, aiding important biological functions such as DNA synthesis and formaldehyde reduction.[139] Most higher organisms utilise glutathione (**125**, Figure 4.1.2), a cysteine containing tripeptide, as the dominant thiol for this purpose. However, in some eukaryotic parasites glutathione is not the dominant thiol and a range of different thiols, such as trypanothione in kinetoplastids, are utilised instead of glutathione for this role.[140]

Studies of *Streptomyces clavuligerus*[141] led to the discovery of a novel cysteine derivative. It was noted that this thiol was produced throughout the growth cycle of the organism in millimolar levels, suggesting that this compound acts as an antioxidant thiol.[142] Sakuda *et al.*[143] were the first to report the structure of this compound, first as the symmetrical disulfide isolated from *Streptomyces sp. AJ 9463* and later as the sulfide **16**

derived from *Mycobacterium bovis* and *Streptomyces clavuligerus*.[142,144] The sulfide **16**, 1-O-(2-[*N*-acetyl-L-cysteinyl]amido-2-deoxy-α-D-glucopyranosyl)-D-*myo*-inositol was given the common name mycothiol due to its prevalence in mycobacteria.[129]

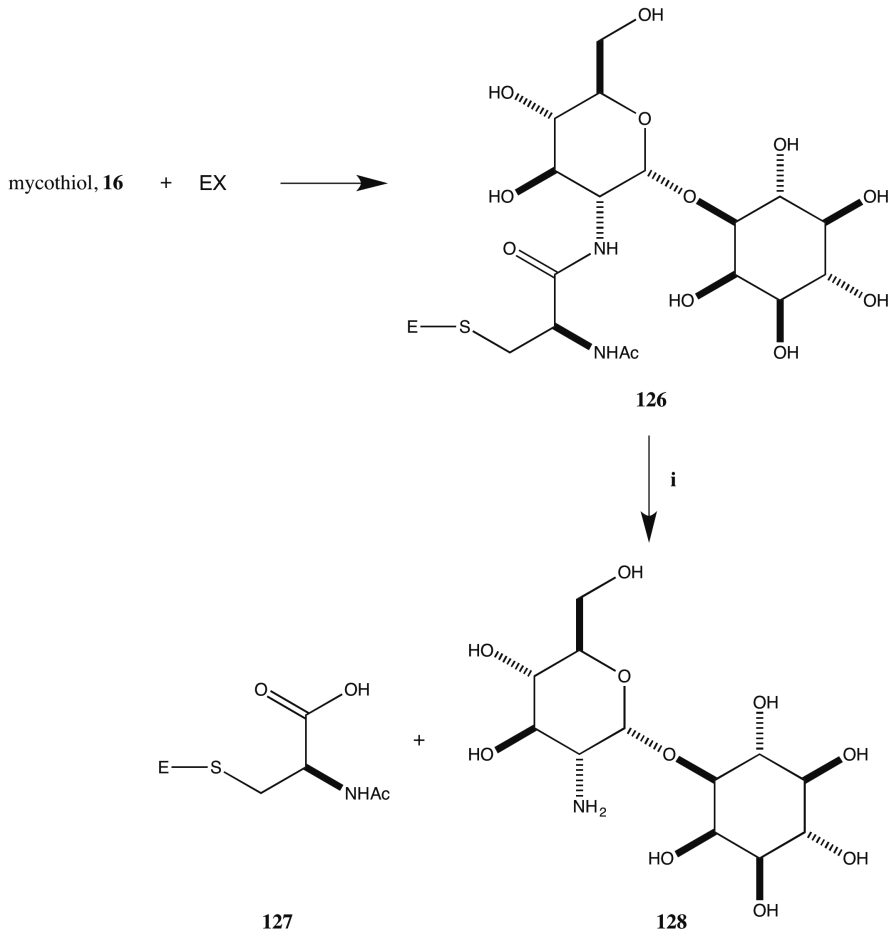


glutathione, **125** mycothiol, **16**

Figure 4.1.2 – Structures of glutathione (**125**) and mycothiol (**16**).

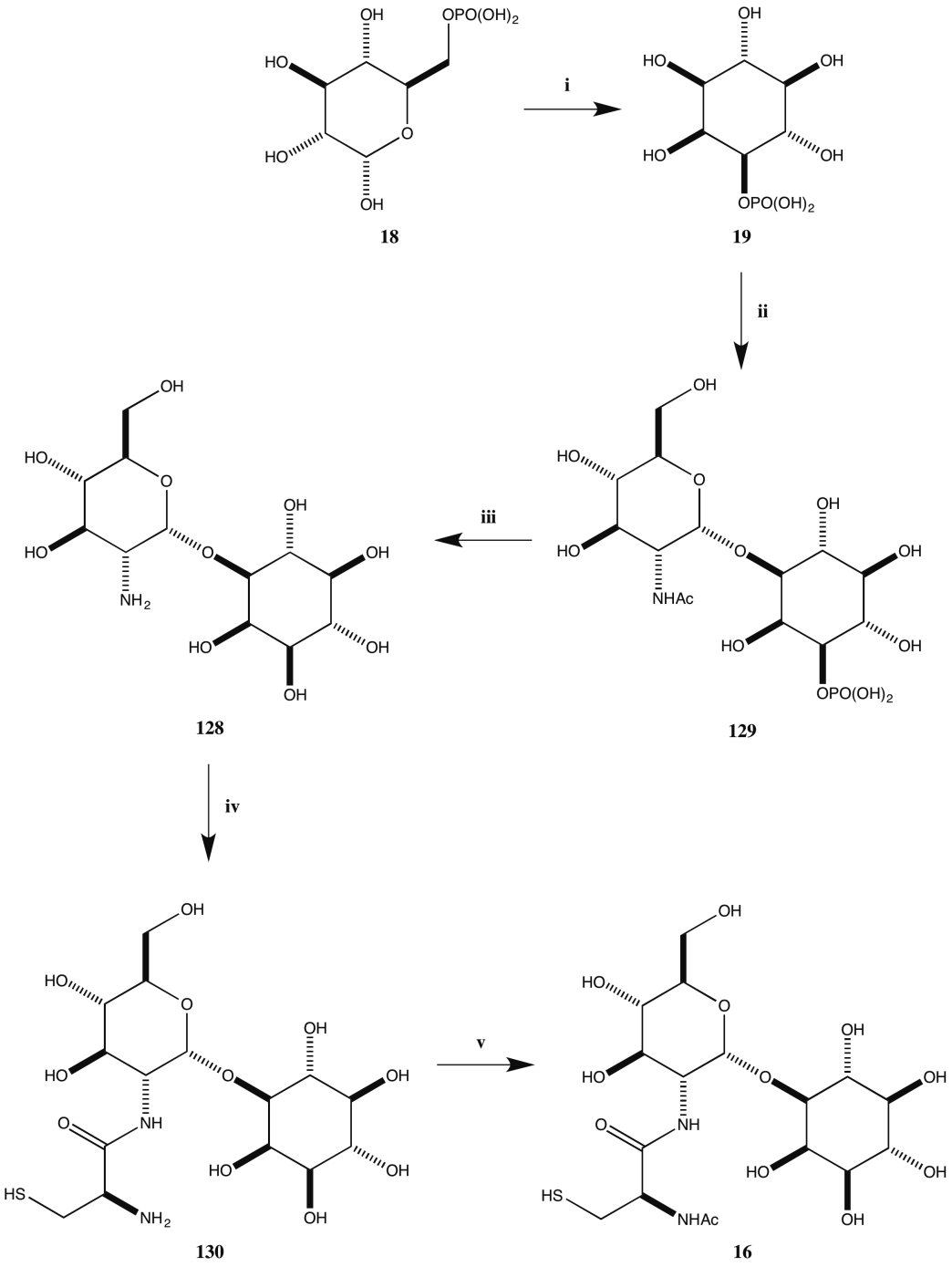
Studies[141,145] have shown that mycothiol **16** is the major low molecular weight thiol in most actinomycetes, but is not even present in other representative bacteria or eukaryotes.[129] The actinomycetes form a diverse group of Gram-positive bacteria containing species that are of high pharmacological interest. One of the species with the highest mycothiol **16** content has been shown to be *M. tuberculosis*.[145]

As a low molecular weight thiol, mycothiol (**16**) acts to help reduce toxic oxidants and maintain an intracellular reducing environment within mycobacteria thereby playing a key role in oxidative stress management and drug resistance.[129] Mycothiol (**16**) can detoxify electrophilic species, including drugs, within the cell by formation of a conjugate **126** (Scheme 4.1.1). The enzyme mycothiol *S*-conjugate amidase (MCA) or mycothiol deacetylase (MshB) will then cleave the conjugated cysteinyl residue from the rest of the molecule, producing a mercapturic acid **127**.[146] This is just one method by which mycothiol (**16**) is used by mycobacteria to maintain a healthy cellular environment.



Scheme 4.1.1 – Detoxification of electrophilic species (EX) within a cell by mycothiol (**16**). *Reagents and conditions:* **i**. Mycothiol *S*-conjugate amidase (MCA) or mycothiol deacetylase (MshB).

It is hypothesised that *M. tuberculosis* could become vulnerable to therapeutics and other stress factors if the enzymatic pathways in the biosynthesis of mycothiol (**16**) or it's roles within the cell are interrupted.[147] For this reason, mycothiol (**16**) and analogues thereof are important synthetic targets that could lead to therapeutic treatments for tuberculosis.[146] With access to mycothiol (**16**) and intermediates within its biosynthesis, scientists can investigate the properties and mechanisms of the enzymes involved in this process. With a deeper understanding of the enzymes, inhibitors can be designed to interrupt this biosynthesis.



Scheme 4.1.2 – Biosynthesis of mycothiol (**16**). *Reagents and conditions:* **i**. INO1; **ii**. MshA; **iii**. a. MshA2 b. MshB; **iv**. L-cysteine, MshC, adenosine triphosphate (ATP); **v**. MshD, acetyl-coenzyme A.

The biosynthesis of mycothiol (**16**) is accomplished in six steps, starting from D-glucose 6-phosphate (**18**, Scheme 4.1.2). In the first step of the biosynthesis, *myo*-inositol 1-phosphate synthase (INO1) acts to convert this compound into L-*myo*-inositol 1-phosphate (**19**). Mycothiol glycosyltransferase (MshA) produces a glycosidic linkage

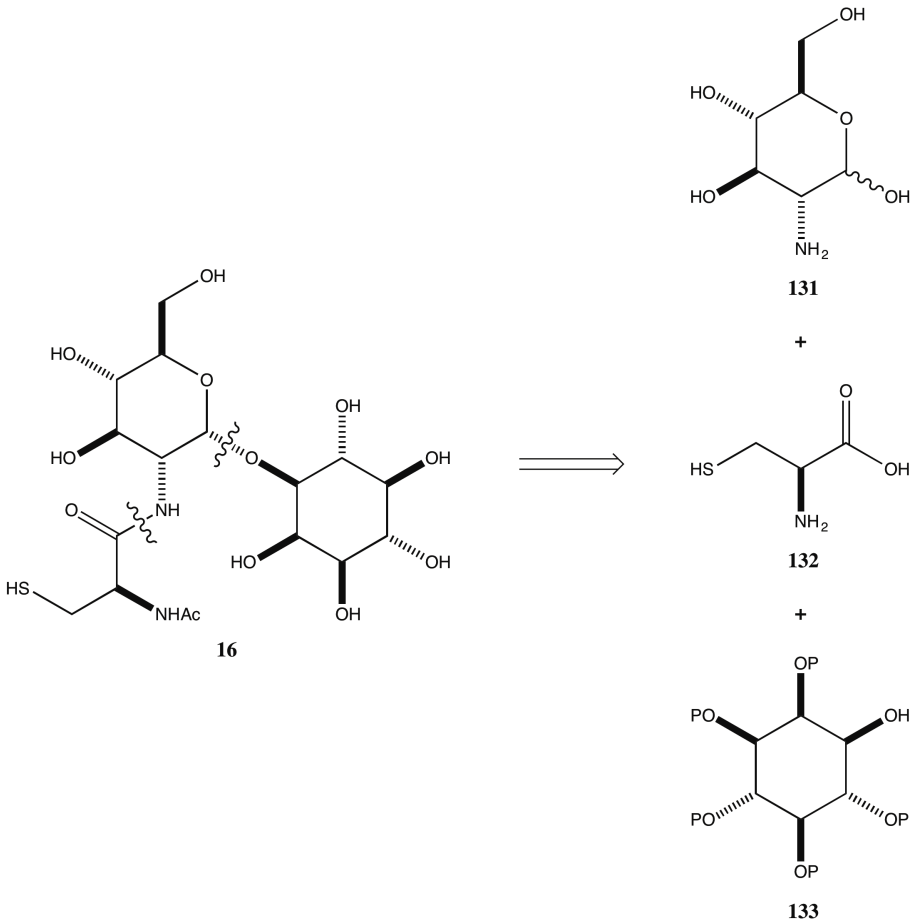
between *N*-acetylglucosamine and the 3-position of L-*myo*-inositol 1-phosphate (**129**) before the phosphate and acetyl groups are cleaved by mycothiol phosphatase (MshA2) and MshB respectively. Coupling of amine **128** and L-cysteine by action of mycothiol ligase (MshC) in the presence of adenosine triphosphate (ATP) produces **130**, which undergoes acetylation by mycothiol synthase (MshD) and acetyl coenzyme A to produce mycothiol (**16**).[148]

From previous studies[149,150] on the biosynthesis of mycothiol (**16**) within *M. tuberculosis*, MshA and MshC have been shown to be crucial to this process. With convenient access to the substrates and products of these enzymes mechanistic and inhibitory studies can be conducted to potentially furnish new therapeutics for tuberculosis. The synthesis of mycothiol (**16**) and its biosynthetic intermediates is not trivial, with the presence of a *myo*-inositol ring attached through the 1-position.

4.1.2 Synthetic Routes to Mycothiol and Analogues

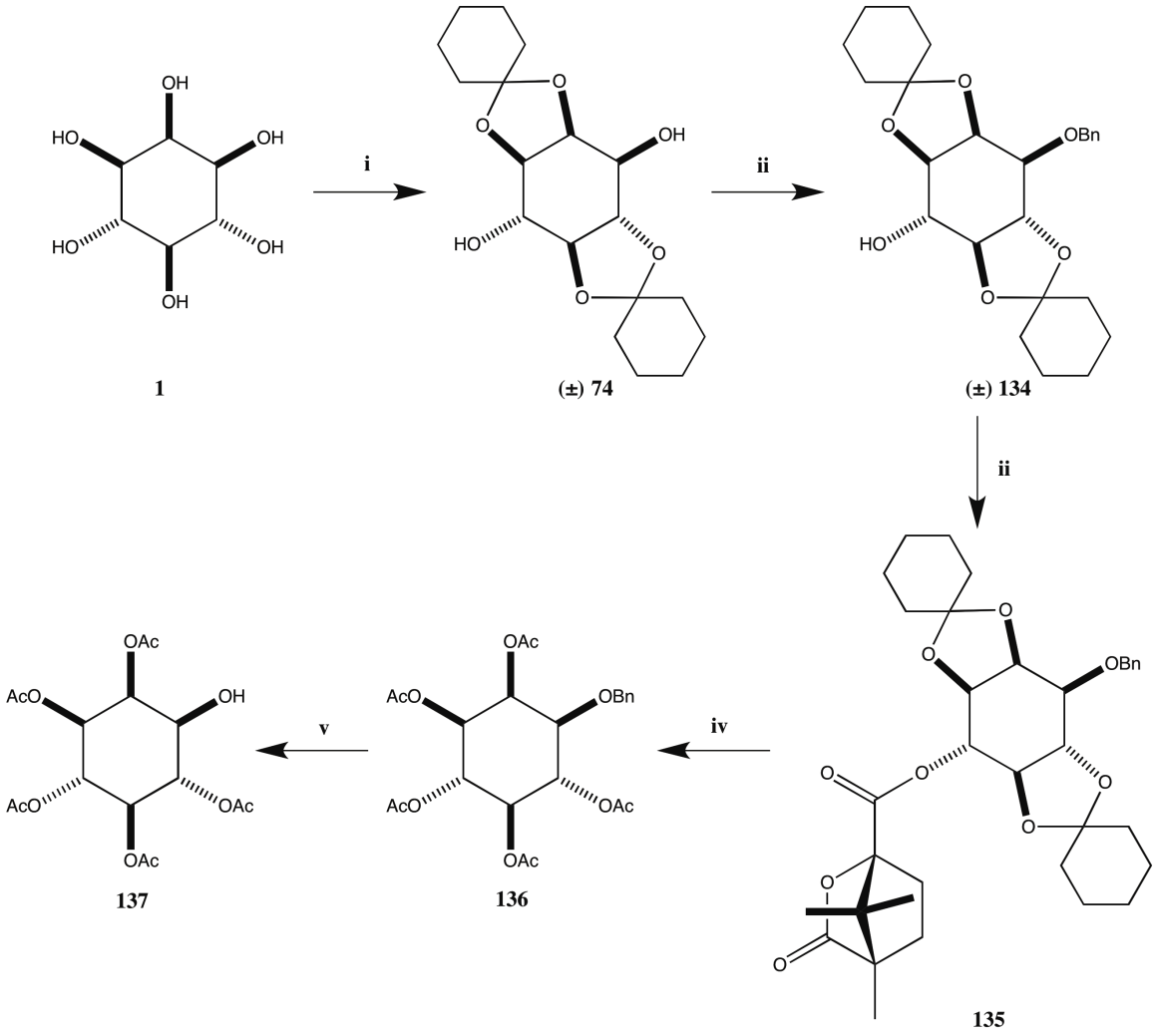
With mycothiol (**16**) highlighted as an important synthetic target in the search for therapeutics for TB, there are published syntheses of the natural product and analogues. The routes are quite lengthy and are also criticized because they have only been conducted on small scales. Both of these issues can be attributed to the requirement of introducing an enantiopure *myo*-inositol moiety within the structure.

Synthetic routes to mycothiol (**16**) all follow the same general retrosynthetic approach (Scheme 4.1.3). Disconnections at the glycosidic and amide bonds lead to glucosamine (**131**), L-cysteine (**132**) and a protected *myo*-inositol derivative (**133**).[129] There are three main issues in the synthesis of mycothiol (**16**) that synthetic routes need to overcome: anomeric selectivity within the glycosylation reaction, cysteine epimerisation during side chain addition, and synthesis of the enantiopure *myo*-inositol derivative (**133**).[129]



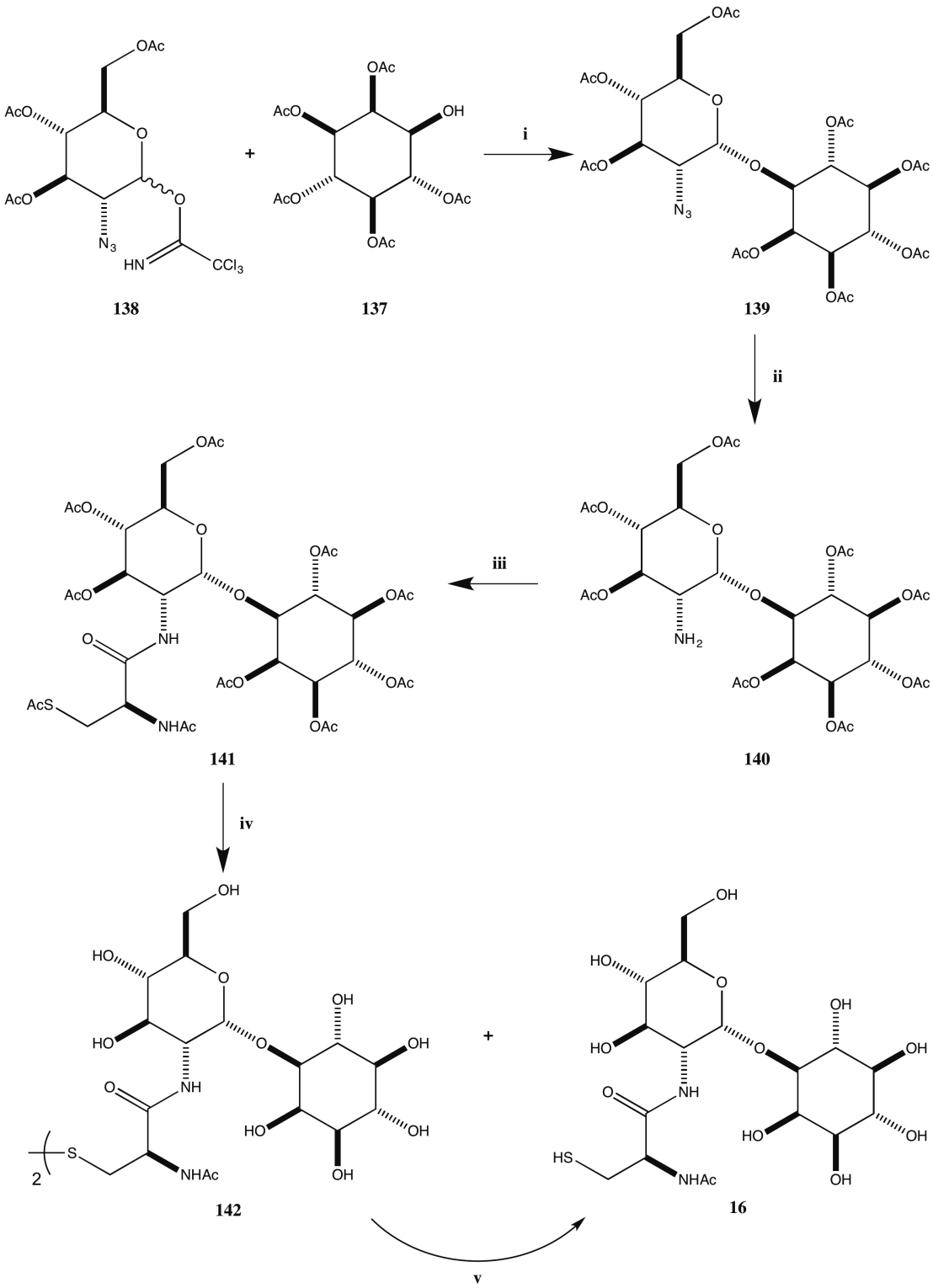
Scheme 4.1.3 – General retrosynthetic route for mycothiol (**16**).

The first total synthesis of mycothiol (**16**) was conducted in 2004[21] with a total of 11 steps in the longest linear sequence, and a total yield of only 0.4 % from *myo*-inositol (**1**). The first key building block in the synthesis was the enantiopure *myo*-inositol derivative **137** (Scheme 4.1.4). They formed the racemic diol **(±) 74** with 1-ethoxycyclohexene before selective benzylation gave alcohol **(±)134**. The D- and L-*myo*-inositol isomers were separated by conversion to their diastereoisomeric camphanate esters using a chiral auxiliary. Taking single diastereoisomer **135**, they produced pentaacetate **136** via cleavage of the camphanate ester and ketals, and peracetylation. Removal of the benzyl group afforded enantiopure alcohol **137**.[21]



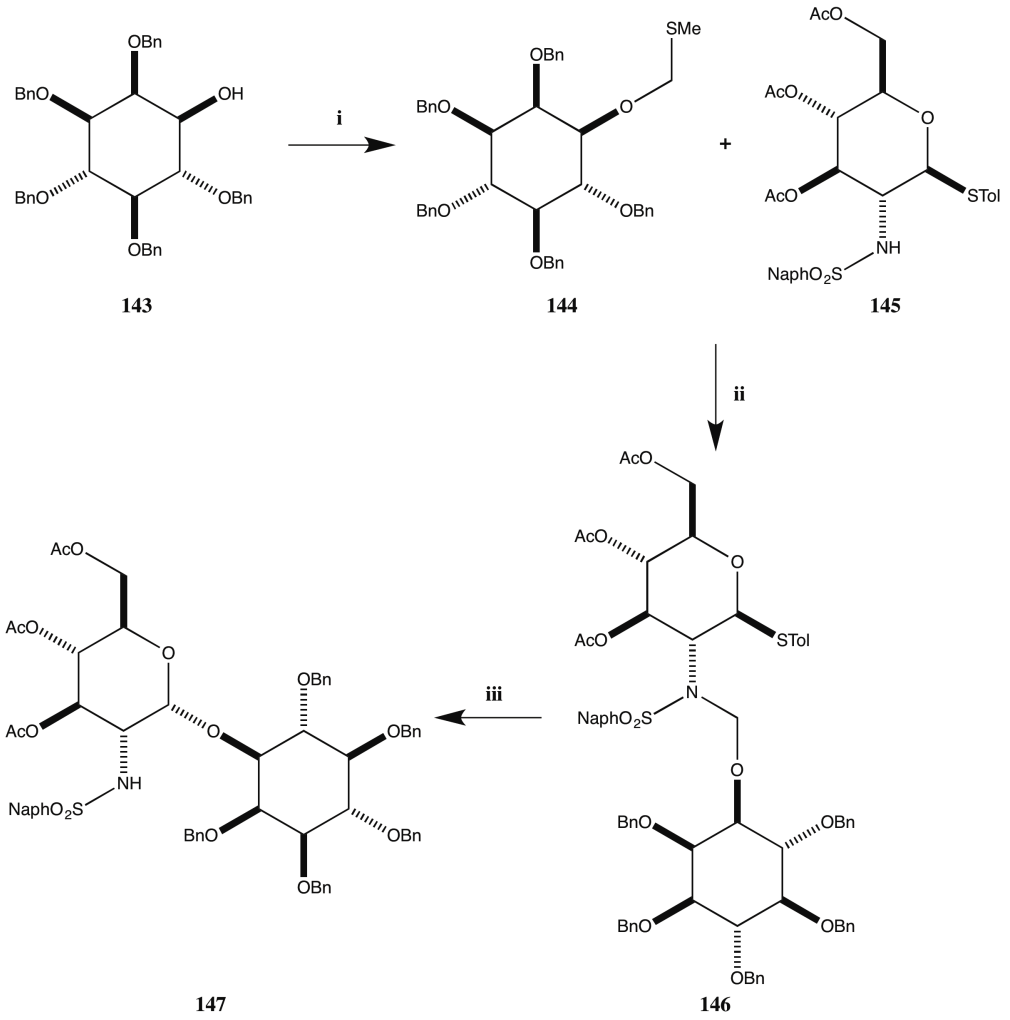
Scheme 4.1.4 – Synthesis of glycosylation acceptor **137**.[21] *Reagents and conditions:* **i**. 1-Ethoxycyclohexene, *p*TsOH, DCM, 100 °C, 23 %; **ii**. NaH, BnBr, toluene, 130 °C then RT, 54 %; **iii**. (1*R*)-(+)-camphanic chloride, DMAP, Et_3N, DCM, 49 %; **iv**. a. KOH, EtOH, 97 %; b. AcOH, water, 100 °C; c. Ac_2O, pyridine, 89 % over 2 steps; **v**. H_2, Pd-C, EtOAc, 97 %.

With the desired alcohol **137** in hand, the completion of mycothiol (**16**) is outlined in Scheme 4.1.5. Trimethylsilyl trifluoromethanesulfonate (TMSOTf) mediated glycosylation of alcohol **137** and trichloroacetamidate **138** (synthesised in 3 steps, 13 % yield) gave a 9:1 ratio of anomers (α:β) in a 65 % yield. Reduction of the azide **139** gave the amine **140**, which could be coupled to *N*,*S*-diacetyl-L-cysteine using *O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HATU) with 1-hydroxy-7-azabenzo-triazole (HOAt) and collidine. It should be noted that 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) and *N*,*N'*-dicyclohexylcarbodiimide (DCC)



Scheme 4.1.5 – The first total synthesis of mycothiol (**16**).[21] *Reagents and conditions:* **i**. TMSOTf, molecular sieves, DCM, 0 °C, 56 %; **ii**. H_2, Pd-C, EtOAc, HCl (2 M, aq.), 81 %; **iii**. *N*,*S*-diacetyl-L-cysteine, HATU, HOAt, collidine, DMF, 0 °C to RT, 25 %; **iv**. $Mg(OMe)_2$, MeOH, 69 %; **v**. bis(2-mercaptoethyl)sulfone, water, 100 %.

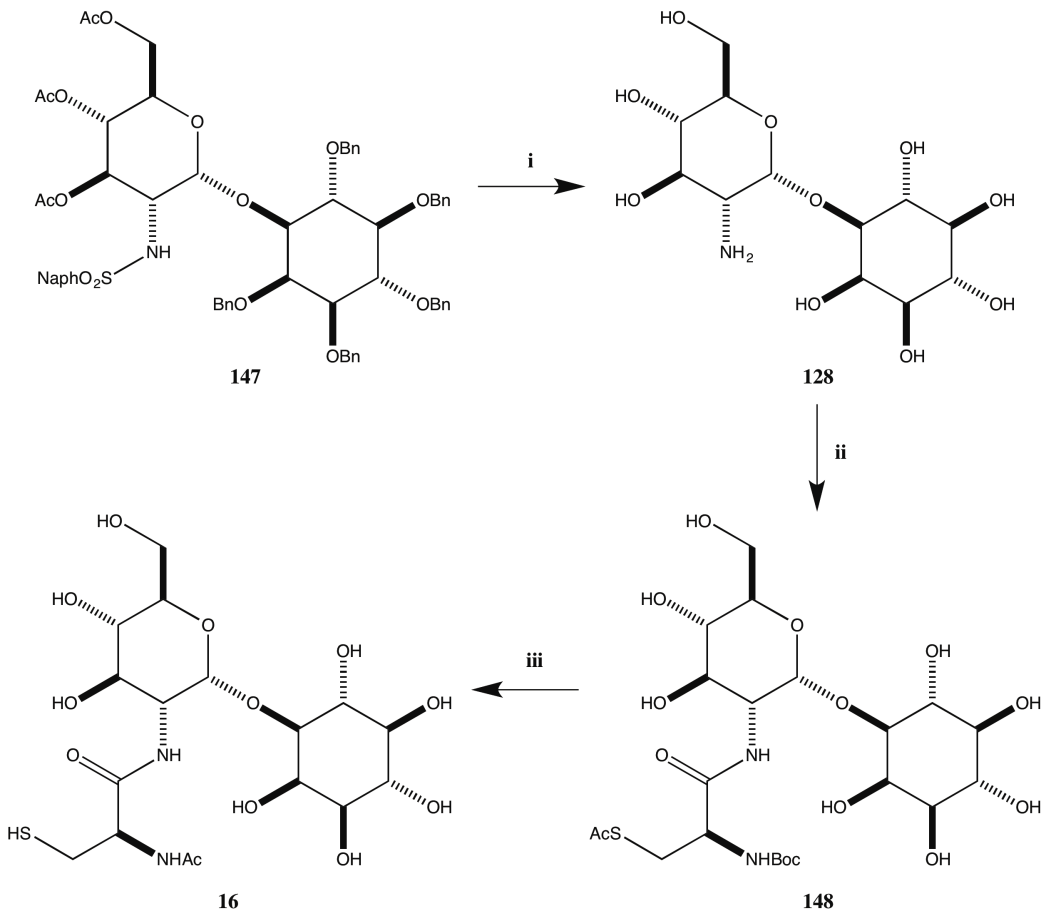
were also used for this coupling, however they both gave epimerisation of the cysteine and lower yields. Treatment of protected mycothiol **141** with dimethoxymagnesium removed the acetate groups from the hydroxyls and thiol, giving mycothiol (**16**) and mycothiol disulfide (**142**), which could be converted into mycothiol (**16**) by treatment with bis(2-mercaptoethyl)sulfone in water for 5 days.[21]



Scheme 4.1.6 – α–Selective glycosylation.[151] *Reagents and conditions:* **i**. MeSCHCl, NaI, NaH, THF, 74 %; **ii**. SO_2Cl_2, DCM, 0 °C, then **145**, BEMP, THF, 98 %; **iii**. PhSCl, AgOTf, DCM, MeCN, -78 to -20 °C, 93 %.

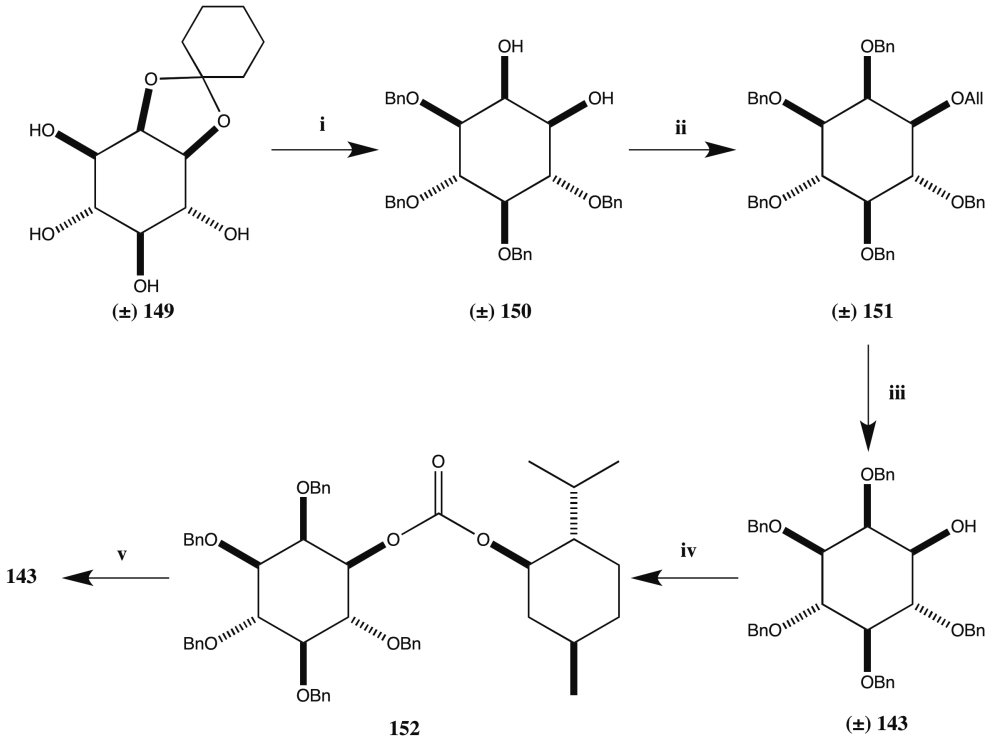
Since then a considerable effort has been made to improve the glycosylation reaction. Knapp *et al.*[151] synthesised mycothiol (**16**) utilising an intramolecular glycosylation reaction. When complex glycosylation donors and acceptors are required, this method has several advantages over intermolecular glycosylation. Firstly, the donor and acceptor are absolutely present in a 1:1 ratio, meaning that there will be less potential for loss of

highly valuable starting materials. The tether between the donor and acceptor means that the aglycon is delivered in a *syn* fashion. It holds the two species within close proximity, allowing for rapid reaction and less likelihood of nonproductive donor elimination or hydrolysis.[151] By careful design and synthesis of glycosylation substrates **144** and **145** (Scheme 4.1.6), Knapp *et al.* demonstrated that selective intramolecular α-glycosylation was indeed possible.[151] For the synthesis of the glycosylation substrate **146**, the protected *myo*-inositol **143** was initially converted into the methylthiomethyl ether **144**. 2-*tert*-Butylimino-2-diethylamino-1,3-dimethylperhydro-1,3,2-diazaphosphine (BEMP) was used to couple sulfonamide **145** (prepared in 2 steps from 1,3,4,6-tetra-*O*-acetyl glucosamine) with *in situ* formed chloromethylether of **144**, producing desired substrate **146**. They could furnish the α-glycosylation product **147** in 93 % yield by treatment with phenylselenyl chloride and catalytic silver (I) trifluoromethanesulfonate (AgOTf).[151]



Scheme 4.1.7 – Completion of mycothiol (**16**).[151] *Reagents and conditions:* **i**. **a**. NaOMe, MeOH; **b**. Na(Hg), Na_2HPO_4, MeOH, 74 % over 2 steps; **c**. H_2, $Pd(OH)_2$, HCl (2 M, aq.), *tert*-butanol (aq.); **ii**. *N*-Boc-*S*-acetyl-L-cysteine, HATU, Hünig's base, DMF, 0 to 20 °C, 77 %; **iii**. **a**. trifluoroacetic acid, DCM; **b**. pyridine; c. concentrate.

They went on to complete the synthesis of mycothiol (**16**, Scheme 4.1.7) via global deprotection to produce amine **128**, before coupling with *N*-*tert*-butyloxycarbonyl-*S*-acetyl-L-cysteine gave protected thiol **148**. *tert*-Butyloxycarbonyl (Boc) deprotection and dissolution in pyridine allowed acetyl group migration from the thiol to the amine, and subsequent removal of pyridine furnished mycothiol (**16**).[151]



Scheme 4.1.8 – Synthesis of optically pure protected *myo*-inositol **143**.[151] *Reagents and conditions:* **i**. **a**. BnCl, KOH, reflux, 16 hr; **b**. AcOH, water, 75 % over 2 steps; **ii**. **a**. $^{n}Bu_2SnO$, benzene, reflux, 24 hr, then AllBr, TBABr, 60 °C, 36 hr, 96 %; **b**. NaH, BnBr, DMF, 24 hr, 97 %; **iii**. Wilkinson's catalyst, DABCO, ethanol (90 %, aq.), reflux, 3 hr, then AcOH, water, THF, reflux, 4 hr, 93 %; **iv**. (+)-menthoxycarbonyl chloride, pyridine, 72 %; **v**. HCl, MeOH, 93 %.

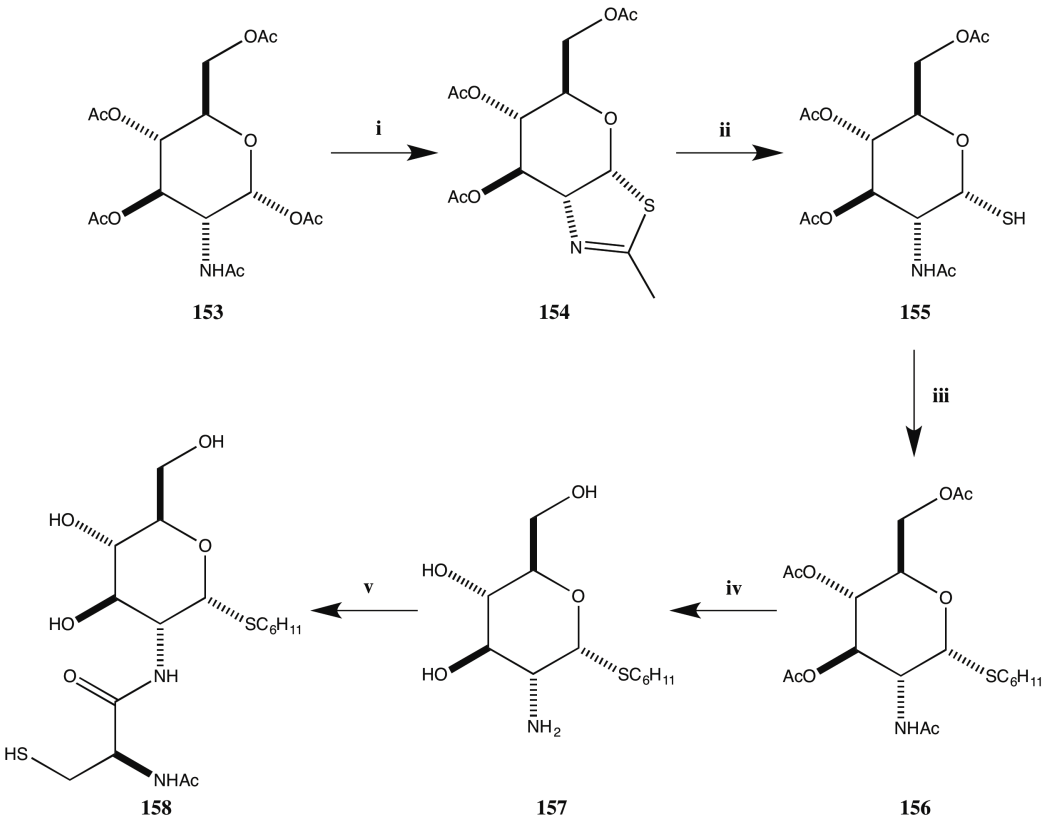
The protected *myo*-inositol **143** was prepared via a resolution methodology based upon purification of a mixed (+)-menthol carbonate (Scheme 4.1.8).[151,152] Racemic **(±) 143** was synthesised in six steps from *myo*-inositol (**1**).[153] Initially the OH-1 and OH-2 were protected as the ketal **(±) 149** in 74 % yield. Benzylation, followed by ketal cleavage produced diol **(±) 150**, which was selectively allylated at the more reactive equatorial hydroxyl. Subsequent benzylation produced the fully protected *myo*-inositol **(±) 151**,

before removal of the allyl group gave the racemic protected *myo*-inositol **(±) 143**. Introduction of the (+)-menthol carbonate allowed separation of diastereoisomers and isolation of carbonate **152**.[152] Removal of the carbonate group then furnished the optically pure protected *myo*-inositol **143**.

In this work, Knapp *et al.*[151] completed the synthesis of the natural product mycothiol (**16**) more efficiently than the first total synthesis. They were able to conduct a more efficient glycosylation using an intramolecular methodology, and their incorporation of the *N*-acetyl-L-cysteine moiety was completed in an elegant fashion with the observed acetyl migration. However, their route to synthesise the optically pure protected *myo*-inositol **143** that they required for the synthesis of their glycosylation substrate **144**, was long, tedious and low yielding. This moiety of the mycothiol (**16**) structure continues to be a stumbling point for synthesis.

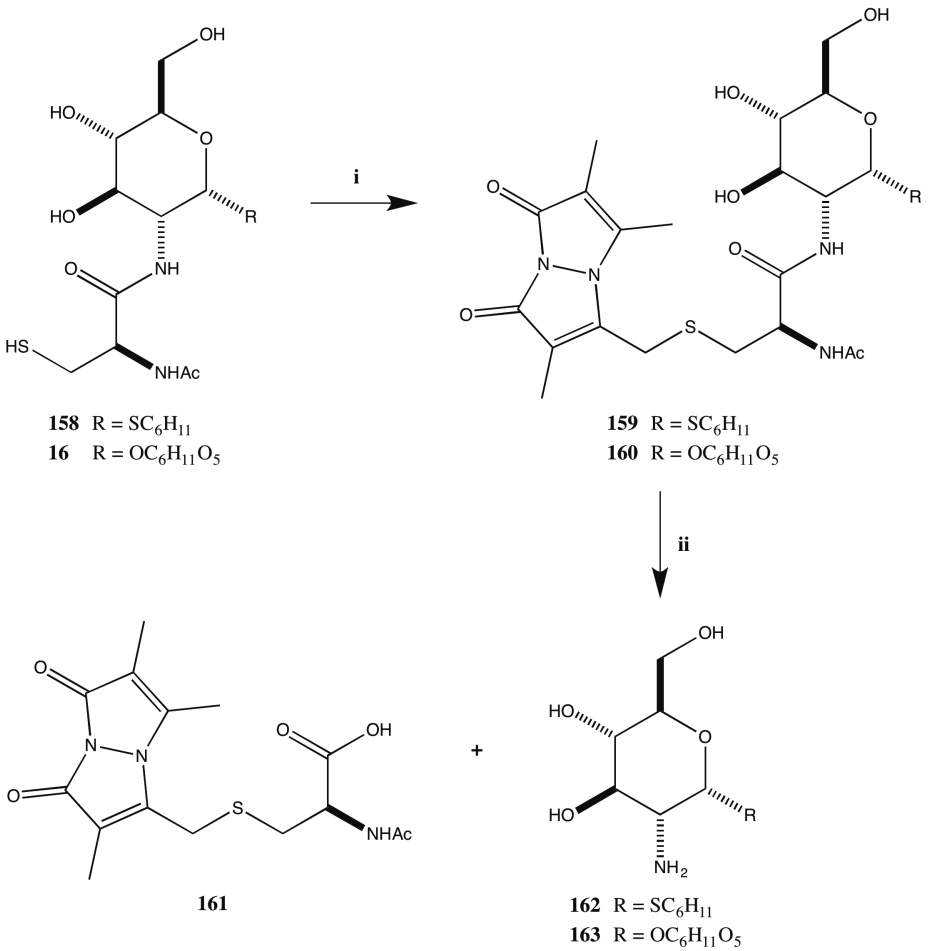
When initially designing potential inhibitors of mycothiol (**16**) biosynthesis, Knapp *et al.*[147] replaced this moiety with a thioglycoside. This not only removed the issue of *myo*-inositol derivative synthesis, but thioglycosides have been shown to be more resistant to degradation by glycosidases.[154] They were able to synthesise thioglycoside **158** (Scheme 4.1.9) in an overall yield of 54 %.

Treatment of peracetylated glucosamine **153** with Lawesson's reagent gave the thiazoline **154**. Hydrolysis furnished the thiol **155**, which could be converted into a free radical using azobisisobutyronitrile (AIBN) and added into cyclohexene to produce the cyclohexyl thioglycoside **156**. Removal of the acetyl groups using hydrazine furnished the triol **157**, before introduction of *N*-Boc-*S*-acetyl-L-cysteine and subsequent treatment, as conducted previously (Scheme 4.1.7), gave the thioglycoside **158** in excellent yield.[147]



Scheme 4.1.9 – Synthesis of thioglycoside **158**.[147] *Reagents and conditions:* **i**. Lawesson's reagent, toluene, 80 °C, 100 %; **ii**. TFA, wet MeOH, 100 %; **iii**. cyclohexene, AIBN, chloroform, reflux, 79 %; **iv**. hydrazine monohydrate, 120 °C, 84 %; **v**. **a**. *N*-Boc-*S*-acetyl-L-cysteine, EDCI, DMF, 82 %; **b**. TFA; c. pyridine; **d**. Sephadex, 100 % over 3 steps.

Using thioglycoside **158** they were able to show that thioglycosides of this type were turned over by the MCA enzyme, and therefore have the potential to serve as a good foundation to build inhibitors.[147] Testing with this enzyme was conducted by initially converting thioglycoside **158** and mycothiol **16** into their bimane derivatives **159** and **160** (Scheme 4.1.10), respectively. They could then subject these derivatives to cleavage by the *M. tuberculosis* MCA enzyme, and use a fluorescence-detected HPLC assay to monitor production of the cysteine-*S*-bimane product **161**. The specific activities for the thioglycoside-*S*-bimane **159** and mycothiol-S-bimane **160** were determined to be 7500 and 14200 nmol min^{-1} mg^{-1} $protein^{-1}$ respectively.[147]



Scheme 4.1.10 – Assay used to determine thioglycoside **158** activity with the *M. tuberculosis* mycothiol S-conjugate amidase (MCA) enzyme.[147] *Reagents and conditions:* **i**. Bromobimane, Tris hydrochloride buffer (20 mM, pH 8.0); **ii**. MCA (recombinant *M. tuberculosis*), Tris buffer (50 mM, pH 7.4), 32 °C.

Whilst the thioglycoside-*S*-bimane **159** is a good substrate for the *M. tuberculosis* MCA enzyme, the mycothiol-*S*-bimane **160** has a greater specific activity. This can be attributed to better binding within the enzyme active site, which can only be made possible by the natural product's *myo*-inositol moiety. If an antibiotic was produced based upon the thioglycoside structure then it would have to compete with the natural mycothiol derived substrates. Inclusion of the *myo*-inositol moiety will be important to increase the efficacy of any future antibiotics produced.

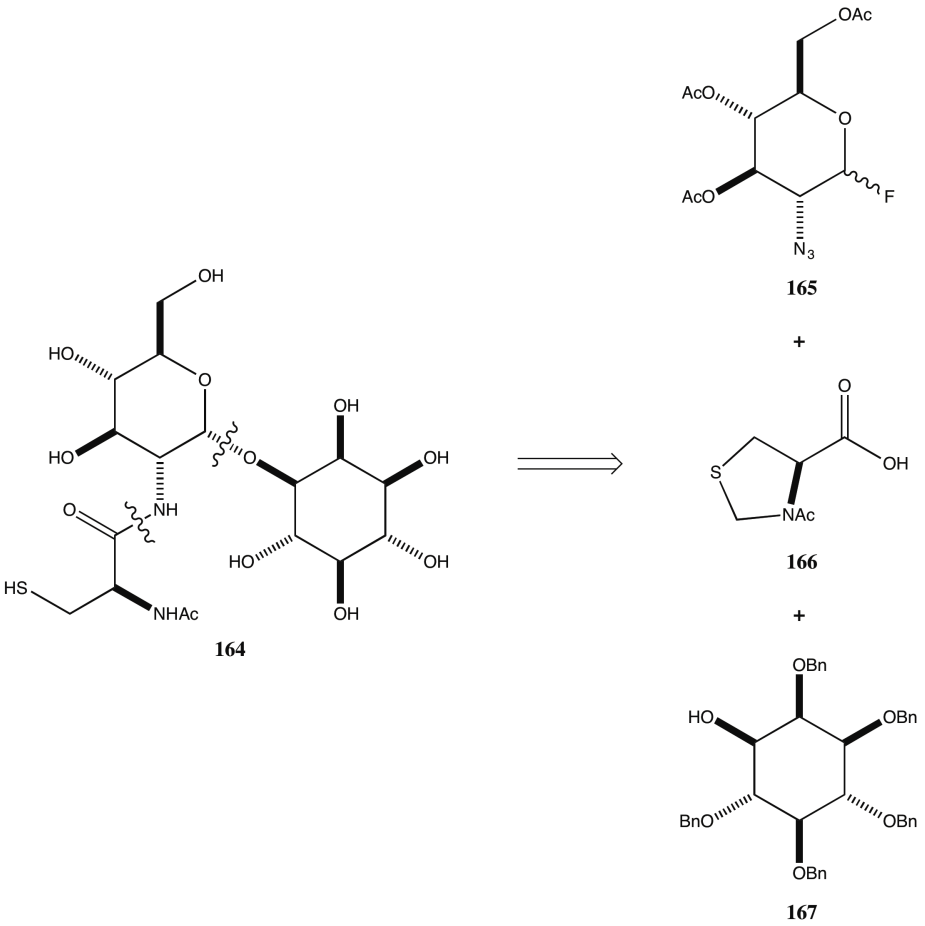
4.2 *epi*-Mycothiol Synthesis

The large scale biocatalytic production of L-*myo*-inositol 1-phosphate (**19**) as conducted in this project, provides vast quantities of an important biosynthetic intermediate in the biosynthesis of mycothiol (**16**). MshA has been highlighted as a crucial enzyme in the synthesis of mycothiol (**16**) within *M. tuberculosis*.[150] The previous work[93] on this enzyme had used a very inefficient method of producing L-*myo*-inositol 1-phosphate (**19**) and, now, with access to large quantities of both of the substrates for this enzyme, further studies on enzyme structure and mechanism can be conducted with more ease.

With 1-bis(phenyloxy)phosphoryl-L-*myo*-inositol (**116**) in hand, the synthesis of mycothiol (**16**) and, in particular, *epi*-mycothiol (**164**, Scheme 4.2.1) should be rapidly achieved. Intermediates within a mycothiol (**16**) synthesis, can be used to form amine **128**, an advanced intermediate in the biosynthesis of mycothiol (**16**). This compound is the substrate for MshC, an enzyme that has also been shown as crucial for the biosynthesis of mycothiol (**16**) within *M. tuberculosis*.[149] This is a better point to target the biosynthesis of mycothiol (**16**) than MshA as it is later in the process and should have less resemblance to molecules present within the human cell.

Although it is not the original natural product, *epi*-mycothiol (**164**) is also desirable as it could help answer questions about the biosynthesis or chemistry of mycothiol (**16**) within a cell. Within the synthesis of *epi*-mycothiol (**164**), we would therefore have access to a diastereomer of **128** and a diastereomer of the natural product. This may be useful in studying the binding site of the enzymes involved, as the hydroxyls of the *myo*-inositol moiety will be presented slightly differently.

Our retrosynthesis of these compounds uses the same initial strategy as those routes within the literature. Retrosynthetic disconnections of the glycosidic bond and the amide bond of *epi*-mycothiol (**164**) initially gave the fluoride **165**, *N*-acetyl thiazolidine **166** and the inositol glycosylation acceptor **167** (Scheme 4.2.1).



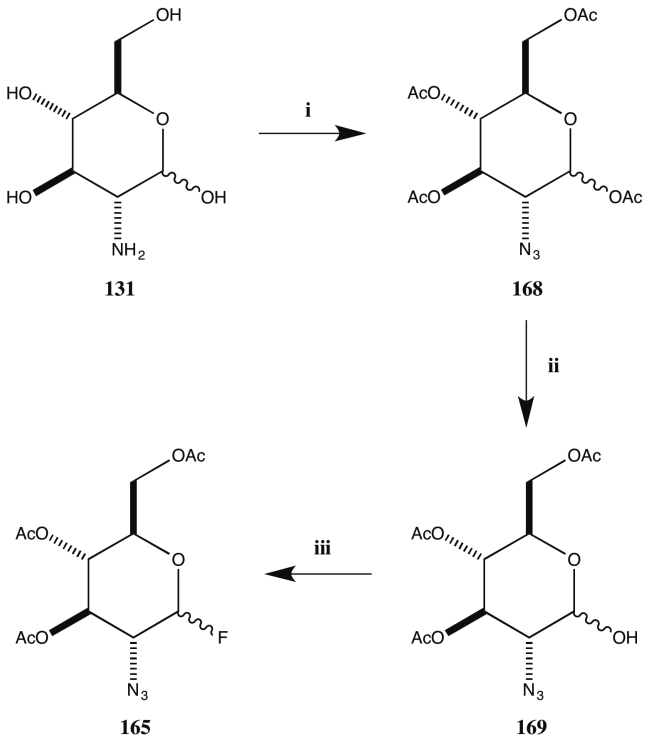
Scheme 4.2.1 – Retrosynthetic analysis of *epi*-mycothiol (**164**).

As previously discussed in section 4.1.2, the synthesis of a protected inositol glycosylation acceptor, such as **167**, is extremely difficult to achieve. We envisage that by utilising the L-*myo*-inositol 1-phosphate (**19**) produced from the biocatalysis, we can furnish the glycosylation acceptor **167** in a concise and efficient synthesis. The synthesis of a glycosylation acceptor for mycothiol (**16**) would require extra steps, this is something that we hope to achieve once a proof of principle is conducted with **167**.

4.2.1 Glucosamine Section

Flouride **165** is a common intermediate in the synthesis of mycothiol (**16**), and as a result there are already a number of synthetic routes to this compound. Following the work of Seeberger *et al.*,[155] introduction of the azide functionality was initially accomplished by treatment of glucosamine (**131**, Scheme 4.2.2) with trifluoromethanesulfonyl azide (triflic azide, TfN_3) with catalytic copper (II) sulfate ($CuSO_4$). Subsequent peracetylation gave

the tetraacetate **168**, which could be selectively deprotected at the anomeric position using benzylamine. The resulting alcohol **169**, could then be converted into the fluoride **165** via treatment with diethylaminosulfur trifluoride (DAST).



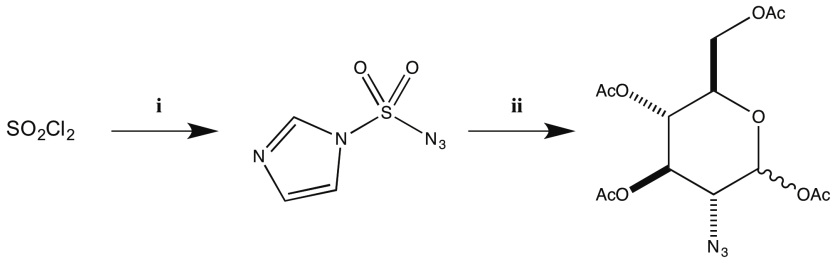
Scheme 4.2.2 – Synthesis of fluoride **165**. *Reagents and conditions:* **i**. **a**. TfN_3, $CuSO_4.5H_2O$, K_2CO_3, DCM, MeOH, water; **b**. Ac_2O, pyridine, DMAP, 73 % over 2 steps **ii**. Benzylamine, Et_2O, 77 %; **iii**. DAST, THF, -30 °C then RT, 87 %.

The initial step in the synthesis, the azidation of glucosamine (**131**), furnished the desired tetraacetate **168** in good yield after peracetylation. This procedure can achieve very high yields in mild reaction conditions, with retention of configuration and good compatibility with most functional groups.[156] However, this procedure has many issues associated with it.[156,157]

Firstly, triflic azide is explosive when not in solution and has a fairly poor shelf-life, meaning that it must be prepared immediately prior to use. The original procedure to prepare triflic azide was to add trifluoromethanesulfonyl anhydride (triflic anhydride) to a vigorously stirred biphasic mixture of dichloromethane and saturated aqueous sodium azide solution. With these conditions, hydrolysis of triflic anhydride will occur and, therefore, it is necessary to use a large excess of the reagent and sodium azide.[156] These

are both highly toxic reagents and triflic anhydride is relatively expensive.[157] This procedure also leaves triflic azide as a solution in dichloromethane after extraction, a solvent that is not very versatile for use in general diazotransfer reactions.[156] The diazotransfer reaction being conducted here is carried out in a mixture of methanol, water and dichloromethane. The ratio of these solvents must be exactly 3:10:3 (methanol:water:dichloromethane) to minimise precipitation of salts from the reaction.[158] This ratio can have a dramatic influence the reaction, and the reaction can still be unpredictable and troublesome.[156] In replicate experiments, the yield of this reaction varied greatly, with a 72 % yield after acetylation only being produced on one occasion.

Ultimately, a different method was needed to accomplish the azidation with a consistent yield. Work by Stick *et al.*,[157] had already addressed this problem. They were able to develop a method of diazotransfer that bypasses the majority of these problems. By using imidazole-1-sulfonyl azide (**170**, Scheme 4.2.3) as a replacement for triflic azide, they discovered a method that is inexpensive, robust and safe. Conversion of imidazole-1-sulfonyl azide (**170**) to the hydrochloride salt makes the new reagent a crystalline solid and very easy to handle. Impact tests, vigorous grinding and prolonged heating did not produce an explosive reaction. They used imidazole-1-sulfonyl azide (**170**) in a very similar system to that used for triflic azide and were able to affect the azidation of a range of amines under mild conditions with retention of stereochemistry and in high yields.[157]



170 **168**

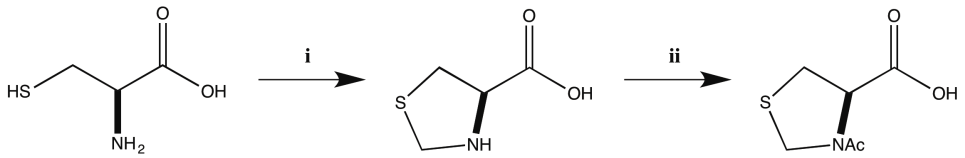
Scheme 4.2.3 – Synthesis and utilisation of imidazole-1-sulfonyl azide (**168**) as a diazotransfer reagent. *Reagents and conditions:* **i**. Sodium azide, MeCN, then imidazole, 0 °C, 90 %; **ii**. a. glucosamine hydrochloride, $CuSO_4.5H_2O$, K_2CO_3, MeOH; b. Ac_2O, pyridine, DMAP, 68 % over 2 steps.

Using imidazole-1-sulfonyl azide (**170**) in the synthesis of the desired azide **168** gave good results (Scheme 4.2.3), and was preferred over the use of triflic anhydride mainly

due to the consistency that this method provided. Synthesis of the reagent (**170**) was achieved readily from imidazole, sodium azide and sulfonyl chloride.

4.2.2 Cysteine Section

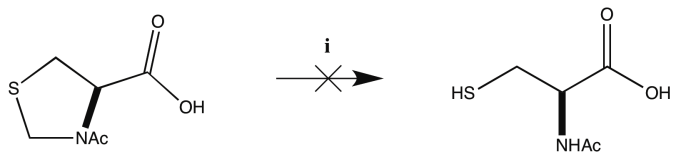
In the retrosynthesis, it was envisaged that the *N*-acetyl L-cysteine moiety could be introduced into *epi*-mycothiol (**164**) via an amide coupling. Routes from the literature utilise a protected thiol in this step to avoid side reactions. There are many options available for this, but a thiazolidine was initially chosen. In work by Payne *et al.*,[159] they utilised the thiazolidine functionality to protect cysteine residues as they built long peptide chains. Treatment with methoxyamine after completion of the chain converted the thiazolidine back into the required cysteine residue. This route was favoured in our synthesis of *epi*-mycothiol (**164**) as it would allow protection of the amine with an acetyl group that could be carried through the synthesis. Final unmasking of the thiazolidine should then leave the required *N*-acetyl group. With this route in mind, *N*-acetyl-L-thiazolidine-4-carboxylic acid (**172**, Scheme 4.2.4) was synthesised.



132 **171** **172**

Scheme 4.2.4 – Synthesis of *N*-acetyl thiazolidine **172**. *Reagents and conditions:* **i**. Formaldehyde, water, 78 %; **ii**. Ac_2O, pyridine, 96 %.

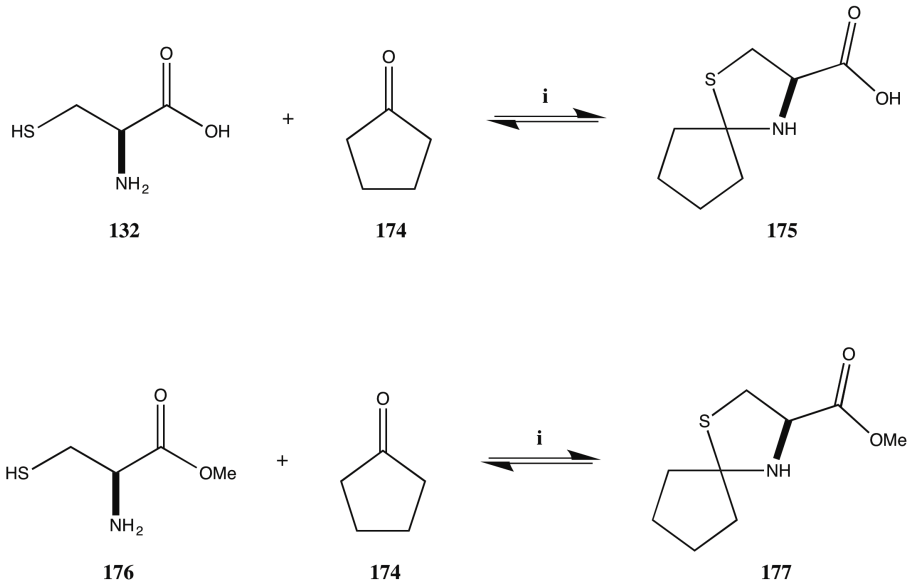
Starting from L-cysteine (**132**), treatment with formaldehyde formed the thiazolidine **171**. Acetylation then produced desired acid **172** in good yield. The final steps in the synthesis of *epi*-mycothiol (**164**) via this route would be the removal of the thiazolidine functionality. In the literature,[159] this is accomplished using methoxyamine. However, in practise reactions on the *N*-acetyl thiazolidine **172**, this procedure did not furnish any reaction and the desired *N*-acetyl cysteine **173** was not produced (Scheme 4.2.5).



172 **173**

Scheme 4.2.5 – Unproductive thiazolidine deprotection reaction. *Reagents and conditions:* **i**. Methoxyamine, water, pH 4.

It was hypothesised that this problem could be overcome by using cyclopentanone (**174**, Scheme 4.2.6) instead of formaldehyde to form a spirothiazolidine (**175**). Within this type of moiety the thiazolidine ring will be more strained, as it has to accommodate the bond geometries needed for the attached cyclopentane ring, and as a result much easier to remove. Indeed, this proved to be the case. However, the spirothiazolidine being formed (**175**) was too unstable, and would ring open within the reaction conditions for its formation. This was observed by ^{1}H NMR spectroscopy, with the ratio of starting material to product fluctuating throughout the course of the reaction. A starting material to product ratio of 2.5:1 was generally observed, with no improvement upon heating the reaction to 40 °C or reflux. Starting with the methyl ester of L-cysteine (**176**) gave no advantage, and the fluctuation between starting material and product (**177**) was still observed. Therefore, using a thiazolidine functionality to protect the thiol of the L-cysteine was abandoned.

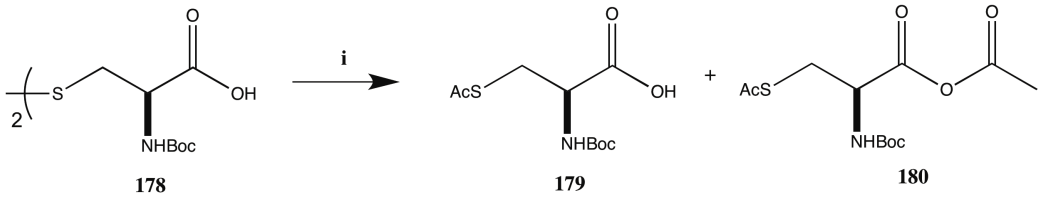


132 **174** **175**

176 **174** **177**

Scheme 4.2.6 – Spirothiazolidine formation using cyclopentanone (**174**). *Reagents and conditions:* **i**. Water, pH 2, 16 hr.

In work by Knapp *et al.*,[151] they accomplished introduction of the L-cysteine moiety into mycothiol (**16**) via addition of *N*-Boc-*S*-acetyl-L-cysteine (**179**, Scheme 4.2.7). Subsequent Boc deprotection and dissolution in pyridine allowed for the transfer of the acetyl group from the thiol to the amine, as required for the natural product. Following this route (Scheme 4.2.7), *N*-Boc-*S*-acetyl-L-cysteine (**179**) was prepared from *N,N'*-bis(Boc)-L-cystine (**178**) via zinc reduction and *in situ S*-acetylation. Unfortunately, the product was only obtained in a 27 % yield as a mixture of the acid **179** and mixed anhydride **180**.

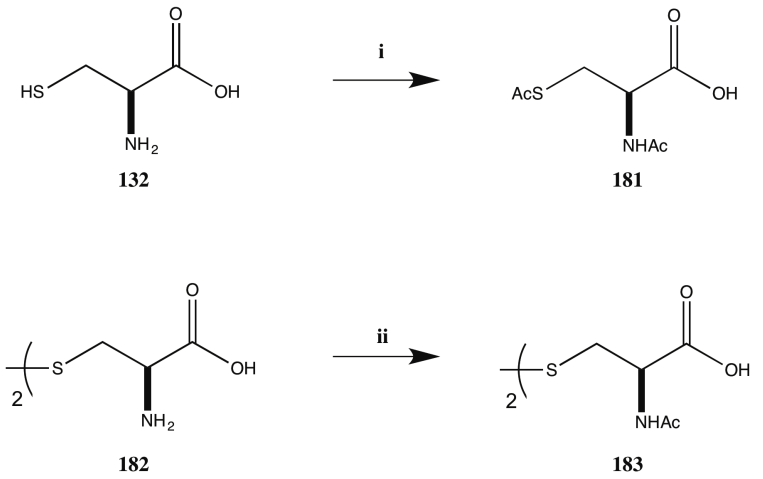


Scheme 4.2.7 – Formation of *N*-Boc-*S*-acetyl-L-cysteine (**179**) and its mixed anhydride **180**. *Reagents and condition:* **i**. a. Zinc (dust), AcOH; b. Ac_2O, pyridine.

The idea of using a protected derivative of cystine (**182**, Scheme 4.2.8) as a means to introduce the cysteine moiety is attractive as it would allow amide couplings on both ends of the molecule, before reduction can produce the free thiol. Procedures from the literature[160,161] were able to furnish *N,N'*-diacetyl-L-cystine (**183**, Scheme 4.2.8) from L-cysteine (**132**) by treatment with acetic anhydride under aqueous basic conditions. Attempts to follow these procedures resulted in the formation of *N,S*-diacetyl-L-cysteine (**181**) and the desired cystine **183** was not observed.

Initial attempts at the straight acetylation of L-cystine (**182**) were not productive, with no acetylation occurring from treatment with acetic anhydride in saturated sodium bicarbonate solution. In the end, acetylation of L-cystine (**182**) was achieved by careful manipulation of the pH of the reaction. The starting material was dissolved in sodium hydroxide solution (15 %, aq.) and cooled to 0 °C. Drops of acetic anhydride and sodium hydroxide solution (30 %, aq.) were added alternately until complete addition of the acetic anhydride was completed. The reaction was then stirred at the same temperature for 2 hours, before being concentrated. Fortuitously, the product was solely soluble in

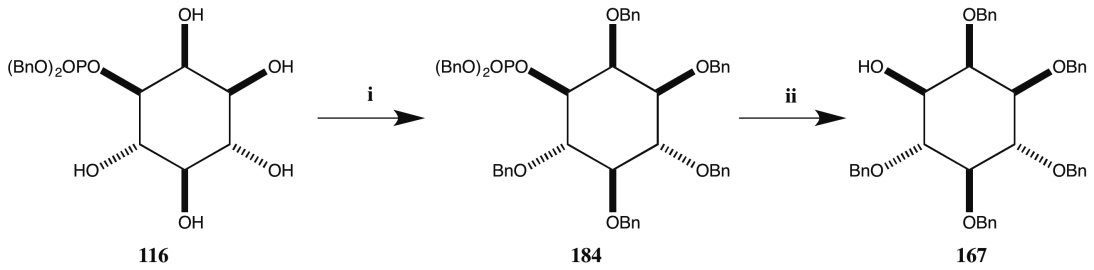
acetic acid and dissolution within this solvent followed by filtration and concentration furnished pure *N*,*N'*-diacetyl-L-cystine (**183**).



Scheme 4.2.8 – Synthesis of *N*,*S*-diacetyl-L-cysteine (**181**) and *N*,*N'*-diacetyl-L-cystine (**183**). *Reagents and conditions:* **i**. Ac_2O, KOH, water, pH 10-10.5, 0 °C, 24 %; **ii**. Ac_2O, NaOH (30 %, aq.), 0 °C, 98 %.

4.2.3 Inositol Section

The synthesis of *epi*-mycothiol (**164**) requires a L-2,3,4,5,6-O-protected glycosylation acceptor, such as **167** (Scheme 4.2.9). Synthesis of *myo*-inositol derivatives that have this enantiopure pattern of protection has been extremely difficult to achieve within the literature (see Section 4.1.2).[21,151] With 1-bis(phenyloxy)phosphoryl-L-*myo*-inositol (**116**) readily available, it was possible to envisage a concise route to an optically pure derivative of this type (Scheme 4.2.9).



Scheme 4.2.9 – Proposed route to a protected glycosylation acceptor (**167**). *Reagents and conditions:* **i**. Benzylation; **ii**. hydrolysis.

Protection of the five hydroxyls with a suitable group, would allow cleavage of the dibenzyl phosphate to reveal the free alcohol (**167**). Ideally, the cleavage of the dibenzyl phosphate would be conducted in one hydrolysis step. The group used to protect the hydroxyl groups should therefore be stable to acid or base hydrolysis. Initially, a benzyl group was chosen as an ideal candidate.

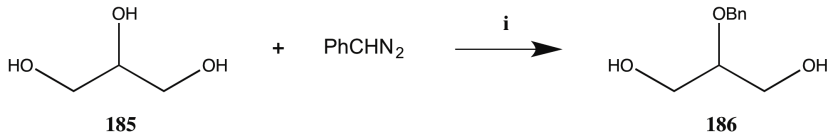
The benzylation of the five hydroxyl groups proved to be extremely difficult to achieve. One of the main issues associated with the attempted benzylation was the poor solubility of the starting material. The 1-bis(phenyloxy)phosphoryl-L-*myo*-inositol (**116**) could only be dissolved in water and methanol. Whilst methods of benzylation do exist for reaction in protic polar solvents, they are not ideal due to the potential quenching of reactive intermediates.

Initial attempts at the perbenzylation were made using pyridine as a potential solvent. The starting material was not soluble in pyridine, but it was hypothesised that if some reaction was to occur then the partially benzylated product would be taken up by this solvent and the reaction could then proceed more rapidly. Using an excess of benzyl bromide (50 equivalents) as a benzylation reagent did not furnish any reaction at all. Addition of tetrabutylammonium iodide (TBAI) did nothing to help the situation, and again no reaction was observed.

Previously, phenyldiazomethane (**115**) had been used as a reagent for the selective benzylation of the phosphate group. It was hypothesised that with a tweak to the reaction conditions, this reagent could be used to achieve perbenzylation on either the L-*myo*-inositol 1-phosphate (**19**) or 1-bis(phenyloxy)phosphoryl-L-*myo*-inositol (**116**). Attempts at the perbenzylation of both of these substrates were conducted by treatment with an excess of the reagent in methanol at room temperature overnight. 1-Bis(phenyloxy)phosphoryl-L-*myo*-inositol (**116**) failed to react with phenyldiazomethane and L-*myo*-inositol 1-phosphate (**19**) was converted efficiently into 1-bis(phenyloxy)phosphoryl-L-*myo*-inositol (**116**), but the reaction did not progress past this point.

Within the literature, Lewis acids have been shown to facilitate the benzylation of secondary hydroxyl groups by phenyldiazomethane with methanol as solvent.

Chittenden[162] was able to effect the selective benzylation of the secondary hydroxyl group within glycerol (**185**) using tin (II) chloride as a Lewis acid (Scheme 4.2.10).



185 **186**

Scheme 4.2.10 – Selective secondary alcohol benzylation using phenyldiazomethane and tin (II) chloride.[162] *Reagents and conditions:* **i**. Tin (II) chloride, MeOH, DCM, 48 hr, 28 %.

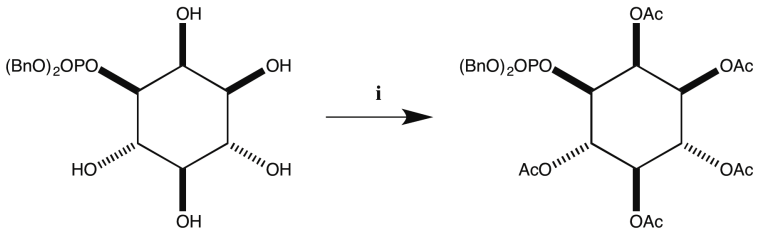
Unfortunately, efforts to conduct the Lewis base promoted benzylation of 1-bis(phenyloxy)phosphoryl-L-*myo*-inositol (**116**) using phenyldiazomethane proved unsuccessful. Tin (II) chloride and cerium (III) chloride were trialled for this purpose as they were readily accessible, but catalytic addition of neither of these Lewis bases afforded any benzylation of the substrate.

An example was found in the literature of fluoroboric acid being used in conjunction with phenyldiazomethane to efficiently benzylate alcohols.[163] Addition of this reagent to the reaction, however, proved unsuccessful with 1-bis(phenyloxy)phosphoryl-L-*myo*-inositol (**116**).

Formation of a tin acetal with dibutyl tin oxide ($^{n}Bu_2SnO$) and subsequent treatment with benzyl bromide is known as a method of selectively benzylating diols.[48] This method can be used to obtain selectivity in which of the two hydroxyl groups are being protected, but also requires a diol for reaction due to the formation of the tin acetal. This method should therefore lend itself to reaction within methanol. 1-bis(phenyloxy)phosphoryl-L-*myo*-inositol (**116**) was refluxed with $^{n}Bu_2SnO$ in a mixure of methanol and toluene for 6 hours to form the tin acetal. The reaction was concentrated before dissolution in toluene and addition of benzyl bromide and tetrabutylammonium bromide. The mixture was refluxed overnight, however, no benzylated product was observed, and it appears as though the tin acetal was not formed in the first instance. The reaction was also attempted in acetonitrile with the same result.

With the benzylation of 1-bis(phenyloxy)phosphoryl-L-*myo*-inositol (**116**) proving elusive, a different protecting group was sought to effect the protection of the five inositol hydroxyls. Ideally, a protecting group would be chosen that is stable at pH extremes whilst being able to be introduced with water or methanol as the solvent. There was no ideal candidate with this chemistry that could be removed at the end of the *epi*-mycothiol (**164**) synthesis without major effects on the rest of the molecule.

With reconsideration of the phosphate cleavage, acetylation of the hydroxyl groups looked like a viable option. Treatment of 1-bis(phenyloxy)phosphoryl-L-*myo*-inositol (**116**) with acetic anhydride, pyridine and DMAP furnished the pentaacetate **187** (Scheme 4.2.11) in a 22 % yield. Repetition of the reaction with freshly distilled pyridine and a new bottle of acetic anhydride pushed the yield up to only 27 %. Unfortunately, it did not seem possible to achieve a higher yield for this step. As the reaction progressed, a side product was being formed in the reaction, which meant that formation of the product seemed to be capped at 27 %. The side product was not observed by thin-layer chromatography (TLC) until roughly 5 hours after beginning, but after this point it appeared that no more of the desired product was being produced. Exclusion of DMAP from the reaction mixture had little to no effect on the reaction outcome. Conducting the reaction overnight, with the DMAP absence still produced the side product and a similarly low yield of the desired pentaacetate **187**. It is hypothesised that this side product is the result of some type of rearrangement occurring with the phosphate moiety, although the identity of this compound is not clear from crude ^{1}H NMR data. If time had allowed, further investigation into this side product would have been investigated.



116 **187**

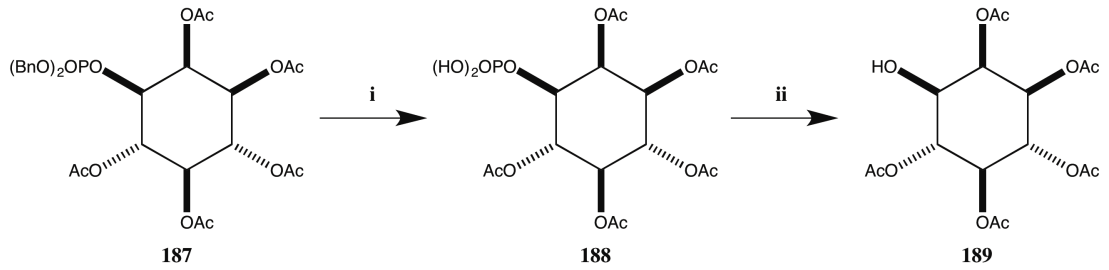
Scheme 4.2.11 – Peracetylation of 1-bis(phenyloxy)phosphoryl-L-*myo*-inositol (**116**). *Reagents and conditions:* **i**. Ac_2O, pyridine, DMAP, 27 %.

Alternative methods to acetylate 1-bis(phenyloxy)phosphoryl-L-*myo*-inositol (**116**) proved unsuccessful. These methods involved trying to acetylate in aqueous conditions, as conducted for the cystine **183**, and using different solvents, DMF and DMSO. Whilst not exhaustive, time constraints meant that the low yielding acetylation product was taken forward.

With the five acetyl groups in place, it was necessary to remove the phosphate group. With hydrolysis originally being highlighted as an efficient method to do this, this was attempted first. Obviously, the acetyl groups will cleave under basic hydrolysis conditions, so it was envisaged that by using aqueous hydrofluoric acid the phosphate could potentially be selectively cleaved. The fluoride ion in aqueous hydrofluoric acid is much less nucleophilic, it therefore presents a much milder method for hydrolysis. Unfortunately, treatment of the pentaacetate **187** with hydrofluoric acid (48 %, aq.) at -15 °C did not produce phosphate hydrolysis. An increase in temperature to 0 °C and an extended reaction time of 3 days, was still not successful. Even treatment of the substrate with hydrofluoric acid (48 %, aq.) at room temperature overnight, failed to produce any reaction at all, as monitored by TLC.

An increase in nucleophilicity of the acid's anion was sought, and it was hypothesised that by switching to hydrochloric acid in methanol and lowering the temperature to -15 °C it may be possible to obtain selective phosphate hydrolysis. However, this was not the case and, even at this lower temperature, the acetyl groups were also removed.

Turning once again to biology, phosphatases are a class of enzymes that have been widely used within synthesis to cleave phosphate groups from substrates. Acid phosphatase from potato is commercially available as a lyophilised powder, and it was envisaged that this could be used as a mild and selective method of phosphate cleavage. Initial attempts using the acid phosphatase were conducted with the phosphate still protected as the dibenzyl species **187**. No reaction was observed with this substrate and it was necessary to remove the benzyl groups prior to reaction.



Scheme 4.2.12 – Synthesis of glycosylation acceptor **189**. *Reagents and conditions:* **i**. H_2, Pd-C, EtOAc; **ii**. acid phosphatase from potato, sodium acetate buffer (50 mM, pH 5.2), DMSO, 37 °C, 79 % over 2 steps.

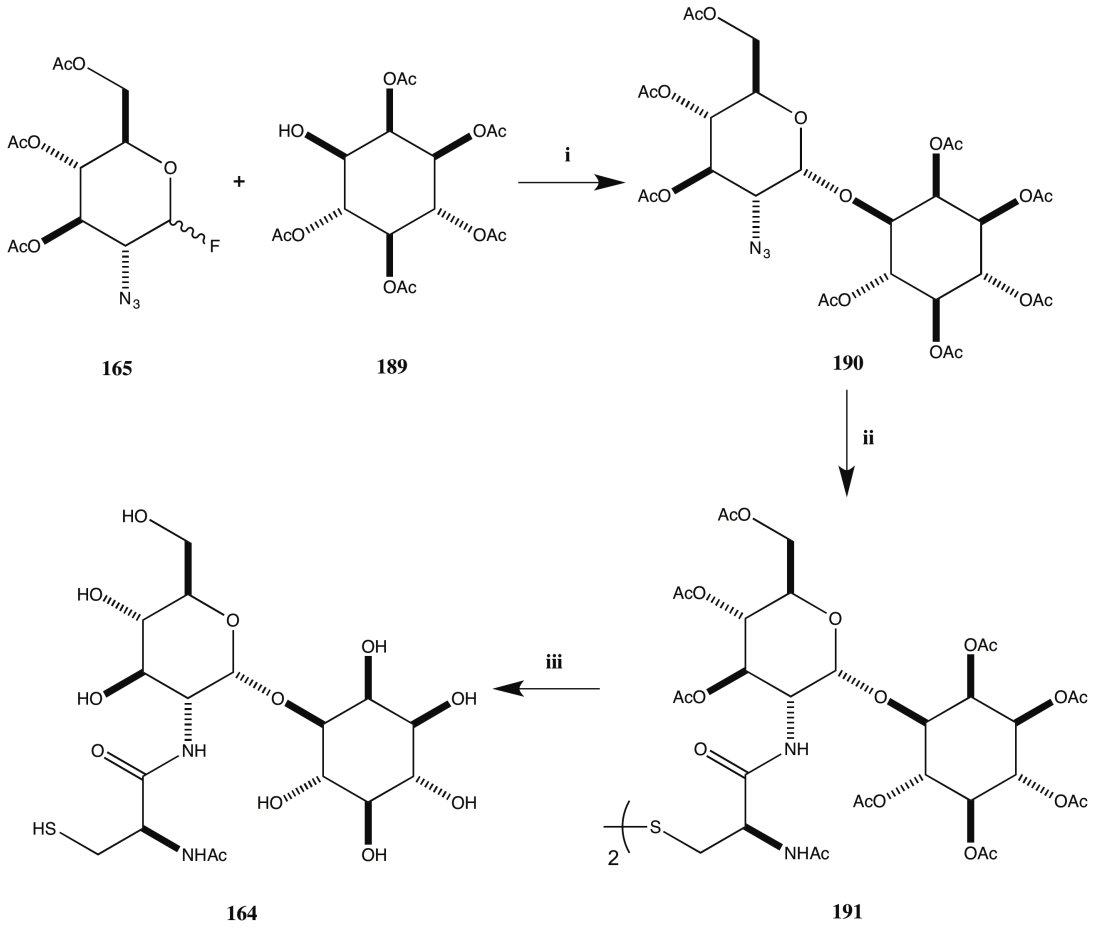
The benzyl groups were removed by hydrogenolysis, furnishing the free phosphate **188** (Scheme 4.2.12). The acid phosphatase could then be used to cleave the phosphate and produce alcohol **189**. The reaction was conducted with shaking at 37 °C and pH 5.2 in a sodium acetate buffer (50 mM). Addition of DMSO was necessary to aid the solubility of the phosphate **188** in the reaction, and the final product could be extracted into ethyl acetate. Overall, this gives the desired glycosylation acceptor **189** in 4 steps and 21 % yield from the biocatalytically produced L-*myo*-inositol 1-phosphate (**19**).

4.3 Conclusions

Each of the three components outlined within the retrosynthesis of *epi*-mycothiol (**164**) have been produced. The route to the glucosamine derived building block very closely follows work from within the literature, with the more reliable imidazole-1-sulfonyl azide (**170**) being used to introduce the azide functionality. Use of *N*,*N'*-diacetylcystine (**183**) as a method of introducing the L-cysteine moiety has not been attempted within the literature. The real strength of our synthesis has been demonstrated with the concise and enantiopure synthesis of the glycosylation acceptor **189**.

The 4-step route conducted to furnish **189** is a vast improvement on the long, tedious and low yielding routes within the literature. Unfortunately, the yield of the acetylation is low and causes an overall yield of 21 %. However, even with this taken into account, our route is still more concise and efficient. With further work on this project, it is believed that a more efficient method of protection of the five hydroxyl groups could be

discovered. One possibility is the use of benzyl trichloroacetamidate as a benzylating reagent.[164]



Scheme 4.3.1 – Envisaged route to *epi*-mycothiol (**164**). *Reagents and conditions:* **i**. Glycosylation; **ii**. **a**. reduction; **b**. amide coupling; **iii**. **a**. selective O-acetate deprotection; **b**. disulfide reduction.

With the three building blocks in hand, it is envisaged that the synthesis of *epi*-mycothiol (**164**) can be completed rapidly (Scheme 4.3.1). Unfortunately, time constraints meant that this work could not be completed.

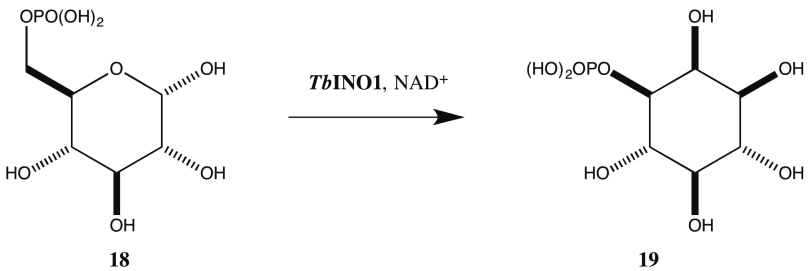
Chapter Five

5 Conclusions

Derivatives of *myo*-inositol (**1**) can be observed within a plethora of important and fundamental biological roles throughout nature. They are incredibly important within cellular signalling as inositol phosphates and phosphatidylinositols and their phosphates.[8] Glycosylphosphatidylinositol (GPI) tethers proteins to the cell membrane,[17] and the *myo*-inositol core can be found within natural products such as mycothiol (**16**).[21]

The synthesis of *myo*-inositol derivatives is challenging. Inositol is a carbohydrate, although it contains no anomeric carbon, and as such faces the regioselective difficulties of classical carbohydrate chemistry. In addition to this, *myo*-inositol chemistry is further complicated by the complex symmetry associated with this molecule, and the associated need for enantioselectivity within synthesis.[8]

Trypanosoma brucei L-*myo*-inositol 1-phosphate synthase (*Tb*INO1) is particularly efficient at converting D-glucose 6-phosphate (**18**) into L-*myo*-inositol 1-phosphate (**19**), due to the parasite's dependence on this protein for survival (Scheme 5.1).[22] This protein was highlighted as a potential synthetic tool to produce highly valuable L-*myo*-inositol 1-phosphate (**19**). This enantiopure material should then be amenable to classical sugar chemistry protecting group strategies to produce highly complex protected *myo*-inositols for use within the synthesis of biologically important molecules.



Scheme 5.1 – Action of *Tb*INO1.

Previous work within the Smith group[22] had recombinantly expressed *Tb*INO1 and shown the protein's activity within a screen for potential inhibitors.[102] The amount of D-glucose

6-phosphate (**18**) being used within each of the reactions was as small as 0.4 mg. We wanted to be working on a scale that was 1000 times greater than this, and several considerations had to be taken into account.

The main issue with the process is achieving complete conversion to L-*myo*-inositol 1-phosphate (**19**). Without full conversion to the product, the remaining D-glucose 6-phosphate (**18**) would be very difficult to remove. It was, therefore, paramount that full conversion within the biocatalysis was achieved in order to obtain a pure product. With the introduction of ammonium bicarbonate as the buffer salt, the process could be streamlined and optimised, giving complete conversion to L-*myo*-inositol 1-phosphate (**19**) on a 10 mg scale.

Further scale-up of the process was not viable within the in-batch process being used. Although modifications were made to the procedure and 76 % conversion could be achieved on a 100 mg scale, ultimately, *Tb*INO1 was not stable at 37 °C for enough time to achieve complete conversion or in an in-batch process product inhibition was an issue.

It was hypothesised that a flow system could facilitate the scale-up of this process. The immobilisation of *Tb*INO1 onto a Ni^{2+}-Sepharose column via its His-Tag provided a proof of this principle, even with the immobilisation method being unoptimised. Using a 5 mL Ni^{2+}-Sepharose column saturated with *Tb*INO1 (~200 mg), it was possible to produce 400 mg of enantiopure L-*myo*-inositol 1-phosphate (**19**) in flow at 37 °C.

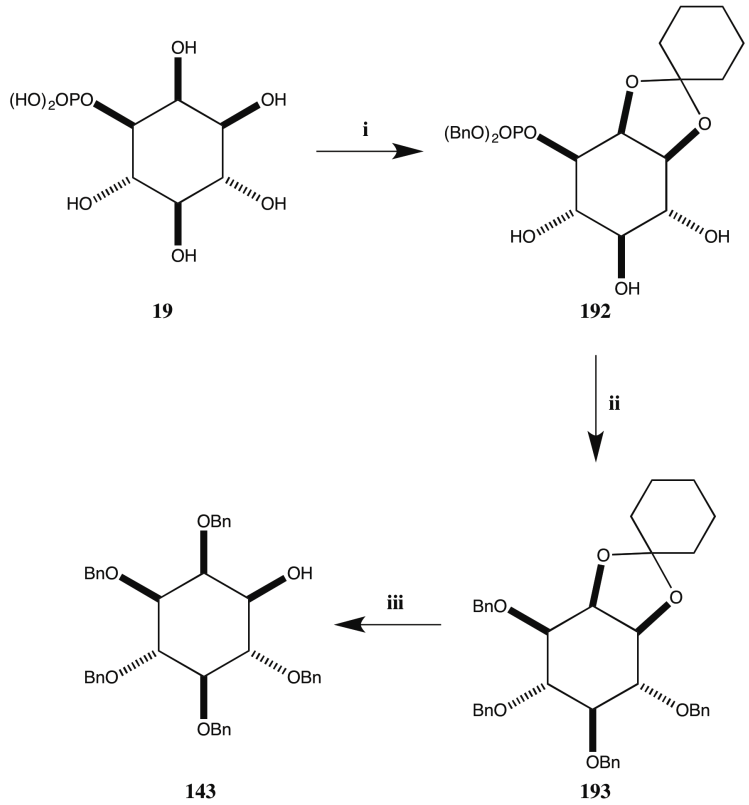
The current system is far from optimal, and suffers significant loss of catalytic activity after 36 hours. With investigation into immobilisation methodologies and precise reaction and stability tuning, the biocatalytic flow system would become more efficient and viable on an industrial scale.

This biotransformation will be a great asset to the synthesis of *myo*-inositol derivatives. The production of 400 mg of L-*myo*-inositol 1-phosphate (**19**) over a 30 hour period (13.3 mg hr^{-1}) is on a scale that has never been seen before. The enzyme is able to produce the enantiopure product in one step, whereas synthetic methods have only been able to reduce the number of steps down to five. Even then, this route requires the synthesis of a polypeptide catalyst.[70]

An industrial production of L-*myo*-inositol 1-phosphate (**19**) would greatly benefit research into a vast number of biological areas. This biotransformation has the potential to produce L-*myo*-inositol 1-phosphate (**19**), which costs £163 mg^{-1} (Cayman Chemical), from starting materials that effectively cost nothing. The substrate for this enzyme, D-glucose 6-phosphate (**18**), could be produced by a glucose phosphorylase. In turn, the D-glucose used in this reaction could be produced from starch. Ultimately, the most expensive part of this process would be the cofactors adenosine triphosphate (ATP) and nicotinamide adenine dinucleotide (NAD^+). NAD^+ is reduced and reoxidised in the course of the *Tb*INO1 reaction, and is tightly held by the enzyme in a Rossman fold.[22] This means that it may be possible to dramatically reduce the amount of NAD^+ required for this process, and recycle it without the need for any modification. ATP can be recycled using enzymes in parallel to the process,[128] so this is not a major concern.

The concept of using a chiral inositol phosphate derivative as a synthetic starting point has not been realised within the literature. This can be attributed to lack of availability and the prohibitive cost of such phosphates. This process provides a cheap and efficient route to an optically pure *myo*-inositol phosphate, which has been shown to undergo selective phosphate benzylation to yield 1-bis(phenyloxy)phosphoryl-L-*myo*-inositol (**116**).

The synthetic advantages of this process have been demonstrated with the synthesis of 1,2,4,5,6-penta-*O*-acetyl-D-*myo*-inositol (**189**) from the biocatalytically produced L-*myo*-inositol 1-phosphate (**19**). This compound can act as a glycosylation acceptor within the synthesis of *epi*-mycothiol (**164**). This diastereomer of mycothiol (**16**) could help address important questions in the function of mycothiol (**16**) within *Mycobacterium tuberculosis*. It differs from the natural product in only the point of connection to the *myo*-inositol moiety, and its synthesis was completed in only 4 steps from L-*myo*-inositol 1-phosphate (**19**) in 21 % yield. We envisage that the synthetic route to the mycothiol glycosylation acceptor (**143**), could be completed in 6 steps (Scheme 5.2). Without the need for a chemical resolution, this synthetic route would be a drastic improvement over the literature.



Scheme 5.2 – Envisaged route to mycothiol glycosylation acceptor **143**. *Reagents and conditions:* **i**. **a**. $PhCHN_2$, Et_2O, MeOH, 0 °C; **b**. acetal formation; **ii**. **a**. phosphate cleavage; **b**. benzylation; **iii**. **a**. acetal cleavage; **b**. selective benzylation.

The *cis*-protection of alcohols on C-2 and C-3 of the *myo*-inositol ring using an acetal (Scheme 5.2) is desirable because it would allow access to different protection patterns of the alcohols. Within the literature,[50] there are a large number of protection strategies that can be used to selectively protect and produce different protection patterns on *myo*-inositol derivatives. With careful consideration of these protection methodologies, we believe that access can be granted to any desired *myo*-inositol derivative from L-*myo*-inositol 1-phosphate (**19**). In some cases, the resulting routes may be lengthy, but with an optically pure starting point the need for a low yielding resolution is removed.

Chapter Six

6 Experimental

6.1 General Experimental Protocols

^{1}H NMR

Spectra were recorded on Bruker Avance 300 (300.1 MHz), Bruker Avance II 400 (400.1 MHz) or Bruker Avance III 500 (500.1 MHz) instruments, and referenced using deuterated chloroform to 7.26 ppm or deuterated water at 4.79 ppm. The chemical shifts were recorded in parts per million (ppm) on the delta (δ) scale ($\delta_{TMS} = 0$). The multiplicity of each signal is indicated by: s (singlet), br. s (broad singlet), d (doublet), dd (doublet of doublets), ddd (doublet of doublets of doublets), dt (doublet of triplets), t (triplet), app. t (apparent triplet) or m (multiplet). The coupling constant (*J*) is measured in Hz and reported to the nearest 0.1 Hz. The number of protons (n) for a given resonance is indicated by nH.

^{13}C NMR

Spectra were recorded on Bruker Avance 300 (75 MHz) or Bruker Avance III 500 (126 MHz) instruments, at ambient room temperature and internal deuterium lock. The chemical shifts were recorded in ppm on the δ scale ($\delta_{TMS} = 0$) and referenced using deuterated chloroform to 77.00 ppm. The coupling constant (*J*) is measured in Hz and reported to the nearest 0.1 Hz.

^{31}P NMR

Spectra were recorded on a Bruker Avance III 500 (202 MHz) instrument using broadband proton decoupling pulse sequences and deuterium internal lock. The chemical shifts were recorded in ppm on the δ scale ($\delta_{TMS} = 0$).

Infrared (IR)

Spectra were recorded on a Shimadzu IRAffinity-1 FTIR spectrometer, using attenuated total reflectance (ATR) as the sampling technique. Absorption maxima are reported in wave numbers (cm^{-1}).

Optical Specific Rotations

Measured using a Perkin Elmer Model 341 automatic polarimeter instrument and 589 nm (sodium D line), in a cell with a path length of 1 dm and are reported as: $[\alpha]_D^{20}$, concentration (*c* in g / 100 mL) and solvent.

High Resolution Mass Spectrometry (HRMS)

Recorded by the EPSRC National Mass Spectrometry Service at Swansea on a Thermo Fisher LTQ Orbitrap XL mass spectrometer using the Electrospray Ionisation (ES) technique.

Analytical Thin Layer Chromatography (TLC)

Conducted on pre-coated (25 μm) Merck Kieselgel 60 F_{254} plates with visualization by ultraviolet (UV) light at 254 nm and/or heating the plate after staining with a solution of 20 % ceric ammonium molybdate w/v in water.

Flash Column Chromatography

Conducted on Merck Silica gel 60 (40-63 μm) under a positive pressure of compressed air.

Reagents and Solvents

Purified by standard methods. Diethyl ether (Et_2O), dichloromethane (DCM) and tetrahydron furan (THF) were dried by passage through two columns of alumina using a MBRAUN SPS-800 solvent purification system. Methanol and acetonitrile were distilled from calcium hydride in a recycling still under argon. Pyridine was purchased from Fisher Scientific and dried by distillation from 4 Å molecular sieves. Dimethylsulfoxide (DMSO) was dried over calcium hydride, distilled under high vacuum and stored over 4 Å molecular sieves. All other reagents were purchased from Sigma Aldrich UK, Alfa Aesar UK, Acros Organics UK or TCI Europe and were used as received, unless otherwise stated. Reactions were performed under anhydrous conditions when appropriate and in an inert atmosphere of argon, using a vacuum manifold with argon passed through a calcium chloride and self-indicating silica gel column. Room temperature (RT) refers to a temperature of approximately 20 °C.

Miscellaneous

The term "brine" refers to a saturated aqueous solution of sodium chloride in deionized water.

6.2 Chapter 2 Experimental

6.2.1 *Trypanosoma brucei* L-*myo*-Inositol 1-phosphate Synthase (*Tb*INO1) Expression and Purification[22]

Recombinant expression of *Tb*INO1 was conducted using the pET15b-*Tb*INO1 construct within BL21 Rosetta (DE3) cells (Invitrogen). *Tb*INO1 was purified by Ni^{2+} affinity chromatography, eluted with an increasing step gradient of imidazole (10, 20, 50, 100, 250 and 400 mM) in 20 mM ammonium bicarbonate (pH 8.5) and 250 mM NaCl. The His-tagged protein was dialysed against 20 mM ammonium bicarbonate (pH 8.5), 50 mM NaCl and 5 mM DTT. The resulting solution could be stored with 10 % glycerol at -80 °C for up to 6 months. In a typical expression, 600 mg of *Tb*INO1 was collected from 8 L of cell culture (75 mg L^{-1}).

6.2.2 In-Batch Reactions

IMPase assay[22]

A sample from the *Tb*INO1 reaction mixture was added to a solution containing Tris-acetate (50 mM, pH 8.0), 1.25 mU of IMPase from bovine brain (Sigma) and $MgCl_2$ (4 mM). The reaction was incubated at 37 °C for 3 hours. The amount of free phosphate was determined by the malachite green method[165] and compared with a standard curve of KH_2PO_4 solution.

Initial Experiments

To water (70 μL) were added solutions of ammonium bicarbonate (500 mM, pH 8.5, 15 μL), D-glucose 6-phosphate (**18**, 100 mM, 15 μL), NAD^+ (20 mM, 15 μL), DTT (10 mM, 15 μL) and *Tb*INO1 (4.35 μg / μL, 20 μL) in order. The mixture was incubated at 37 °C overnight. The IMPase assay determined the conversion of the reaction as 92 %.

10 mg Scale

To D-glucose 6-phosphate (**18**, 10 mg, 38 μmol) in water (458.0 μL) were added solutions of ammonium bicarbonate (500 mM, pH 8.5, 60 μL), NAD^+ (200 mM, 6 μL), DTT (100 mM, 6 μL) and *Tb*INO1 (4.35 μg / μL, 70 μL) in order. The mixture was incubated at 37 °C overnight. The IMPase assay determined the conversion of the reaction as 98 %.

100 mg Scale

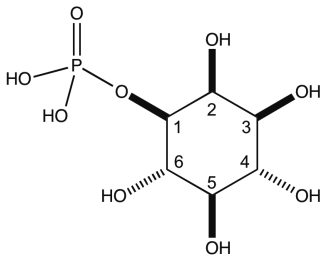
To D-glucose 6-phosphate (**18**, 100 mg, 384 μmol) in water (1664 μL) were added solutions of ammonium bicarbonate (500 mM, pH 8.5, 240 μL), NAD^+ (200 mM, 48 μL), DTT (100 mM, 24 μL) and *Tb*INO1 (4.35 μg / μL, 400 μL) in order. The mixture was incubated for 6 days with shaking at 37 °C in the dark and under an argon atmosphere. The IMPase assay determined the conversion of the reaction as 76 % and no further reaction progress was observed.

6.3 Chapter 3 Experimental

6.3.1 Set up of flow system

Frozen pellets from 4 L of growth media were suspended in a lysis buffer of 20 mM ammonium bicarbonate (pH 8.5), 200 mM NaCl and 3 mM 2-mercaptoethanol. Cells were lysed by sonication (7 x 20 second bursts, 40 second rests) and the supernatant cleared by centrifugation (40,000 rcf, 20 minutes, 4 °C). The supernatant was filtered (0.45 μm) prior to loading onto a freshly charged, new 5 mL Ni^{2+}-Sepharose column. The column was washed with 25 mL of a solution containing 10 mM imidazole, 50 mM ammonium bicarbonate (pH 8.5) and 1 mM 2-mercaptoethanol to elute the unwanted cell extracts. The column was finally washed with a solution containing 50 mM ammonium bicarbonate (pH 8.5) and 1 mM 2-mercaptoethanol, before fitting a Sepharose column that had been previously stripped of any Ni^{2+} ions by washing with 50 mM EDTA solution.

6.3.2 L-*myo*-Inositol 1-phosphate (19)



19

Procedure A

A solution containing 18 mM D-glucose 6-phosphate (**18**), 1 mM NAD^+, 50 mM ammonium bicarbonate (pH 8.5) and 1 mM 2-mercaptoethanol was passed through the previously prepared flow system at a rate of 2.5 mL hr^{-1}. 25 mL fractions were collected and each fraction was separately acidified to pH <2 using concentrated HCl, before being lyophilised to produce crude L-*myo*-inositol 1-phosphate (**19**). ^{1}H NMR was used to determine which fractions contained no D-glucose 6-phosphate (**18**), which were combined and dissolved in methanol, filtered and concentrated under reduced pressure to yield L-*myo*-inositol 1-phosphate (**19**, 341 mg) with a small impurity.

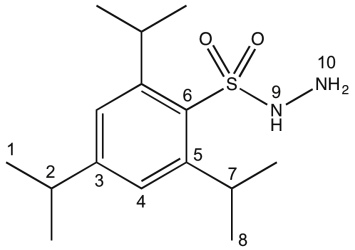
Procedure B

A solution containing 15 mM D-fructose 6-phosphate (**120**), 1 mM NAD^+, 50 mM ammonium bicarbonate (pH 8.5) and 1 mM 2-mercaptoethanol was passed through the previously prepared flow system at a rate of 2.5 mL hr^{-1}. The eluted solution was acidified to pH <2 using concentrated HCl, before being lyophilised to produce crude L-*myo*-inositol 1-phosphate (**19**).

$[\alpha]_D^{20}$ -4.60 (*c* 1.00, H_2O at pH 9); **IR** (ATR) 3262, 3041, 2965, 2936, 1634, 1033, 999; **^{1}H NMR** (500 MHz, D_2O) δ 4.24 (1H, t, *J* = 2.8 Hz, H-2), 3.94 (1H, ddd, *J* = 9.8, 8.8, 2.8 Hz, H-1), 3.73 (1H, t, *J* = 9.7 Hz, H-6), 3.63 (1H, t, *J* = 9.7 Hz, H-4), 3.54 (1H, dd, *J* = 10.0, 2.9 Hz, H-3) and 3.31 (1H, t, *J* = 9.4 Hz, H-5); **^{13}C NMR** (126 MHz, D_2O) δ 75.7, 73.8, 72.2, 71.4 (d, *J* = 6.1 Hz), 71.1 and 70.7; **^{31}P NMR** (202 MHz, D_2O) δ -0.19;

HRMS (ES^-) Calc. for $C_6H_{12}O_9P$ $[M-H]^-$ 259.0224, found 259.0222. The data are in good agreement with the literature values.[69,166]

6.3.3 2,4,6-Triisopropylbenzenesulfonyl hydrazide (118)[167]

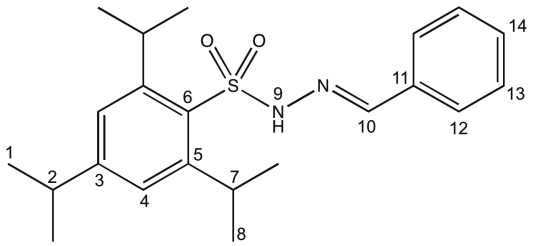


118

To a stirred solution of 2,4,6-triisopropylbenzenesulfonyl chloride (**117**, 2.00 g, 6.60 mmol) in THF (10 mL) was added hydrazine monohydrate (0.64 mL, 13.21 mmol, 2 equiv.) slowly dropwise at 0 °C. The reaction was stirred for 4 hours at this temperature. The precipitate was dissolved by addition of water (20 mL) and the product was extracted in diethyl ether (2 x 30 mL). The organic phases were combined and washed with water (30 mL) and brine (30 mL) before being dried over magnesium sulfate, filtered and concentrated under reduced pressure to give hydrazide **118** as a white solid (1.88 g, 6.28 mmol, 95 %).

^{1}H NMR (300 MHz, $CDCl_3$) δ 7.20 (2H, s, H-4), 5.49 (1H, s, H-9), 4.17 (2H, septet, J = 6.8 Hz, H-7), 3.66 (2H, br. s, H-10), 2.92 (1H, septet, J = 7.0 Hz, H-2), 1.27 (12H, d, J = 6.8 Hz, H-8) and 1.26 (6H, d, J = 7.0 Hz, H-1); **^{13}C NMR** (75 MHz, $CDCl_3$) δ 154.0, 152.0, 128.7, 124.2, 122.9, 34.4, 30.0, 25.1 and 23.7. The data are in good agreement with the literature values.[167]

6.3.4 *N'*-Benzylidine-2,4,6-triisopropylbenzenesulfonyl hydrazide (119)[168]

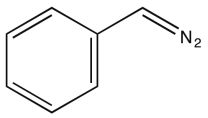


119

To a solution of 2,4,6-triisopropylbenzenesulfonyl hydrazide (**118**, 1.50 g, 5.03 mmol) in methanol (20 mL) was added benzaldehyde (0.51 mL, 5.03 mmol, 1 equiv.). The reaction was stirred at room temperature for 30 minutes before being stirred at 4 °C for 16 hours. The precipitate was collected and washed with ice-cold methanol before dried *in vacuo*. Recrystallisation in ethanol furnished pure hydrazide **119** (1.77 g, 4.59 mmol, 91 %).

^{1}H NMR (300 MHz, $CDCl_3$) δ 7.85 (1H, br. s, H-10), 7.60-7.50 (2H, m, H-13), 7.41-7.29 (3H, m, H-12 and H-14), 7.18 (2H, s, H-4), 4.27 (2H, septet, *J* = 6.8 Hz, H-7), 2.89 (1H, septet, *J* = 6.9 Hz, H-2), 1.31 (12H, d, J = 6.8 Hz, H-8) and 1.24 (6H, d, *J* = 6.9 Hz, H-1); **^{13}C NMR** (75 MHz, $CDCl_3$) δ 153.5, 151.4, 146.4, 130.3, 128.6, 127.3, 123.9, 34.2, 30.1, 24.9 and 23.5. The data are in good agreement with the literature values.[116]

6.3.5 Phenyldiazomethane (115)[116]

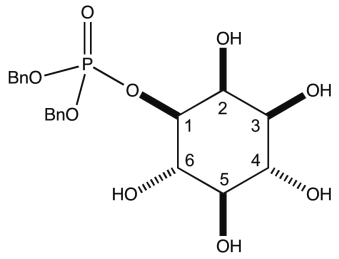


115

To a stirred suspension of *N*'-benzylidene-2,4,6-triisopropylbenzenesulfonyl hydrazide (**119**, 1.78 g, 4.61 mmol) in methanol (30 mL) was added potassium hydroxide (0.52 g, 9.22 mmol, 2 equiv.) and the mixture heated to reflux for 30 minutes in apparatus with flame-polished joints. The reaction was allowed to cool to room temperature before ice-cold water (30 mL) was added. The product was extracted into diethyl ether (3 x 30 mL)

and subsequently washed with sodium bicarbonate solution (sat., aq., 30 mL), water (30 mL) and brine (30 mL) in glassware with flame-polished joints. The phenyldiazomethane (**115**) was then used as an ethereal solution in further reactions.

6.3.6 1-Bis(benzyloxy)phosphoryl-L-*myo*-inositol (116)



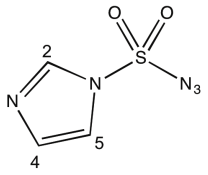
116

To a stirred solution of L-*myo*-inositol 1-phosphate (**19**, 200 mg, 0.77 mmol) in methanol (600 mL) at 0 °C was added an excess of ethereal phenyldiazomethane solution (freshly prepared, 6 equivalents of hydrazide **119** used in preparation). The reaction was stirred for 4 hours at 0 °C before being quenched with acetic acid (0.50 mL) and stirred for a further 10 minutes. The reaction was concentrated before being dissolved in water (100 mL) and washed with diethyl ether (100 mL x 3). The water was removed by lyophilisation to yield 1-bis(benzyloxy)phosphoryl-L-*myo*-inositol (**116**, 329 mg, 0.75 mmol, 97 %) as a pale yellow solid.

$[\alpha]_D^{20}$ +6.50 (*c* 0.10, H_2O); **IR** (ATR) 3323, 3134, 3035, 2808, 1402, 1010, 999; **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.48-7.34 (m, 10H, H-Ar), 5.12-5.06 (m, 4H, H-7), 4.16-4.10 (m, 1H, H-1), 4.07 (t, *J* = 2.9 Hz, 1H, H-2), 3.76 (t, *J* = 9.7 Hz, 1H, H-6), 3.60 (t, *J* = 9.6 Hz, 1H, H-4), 3.40 (dd, *J* = 10.0, 2.8 Hz, 1H, H-3) and 3.25 (t, *J* = 9.5 Hz, 1H, H-5); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 135.1, 129.1, 128.8, 128.4, 78.7, 73.6, 71.9, 70.8, 70.8, 70.7, 70.7, 70.6, 70.5, 70.4, 70.3; **^{31}P NMR** (202 MHz, $CDCl_3$) δ -1.95; **HRMS** (ES^+) Calc. for $C_{20}H_{26}O_9P$ $[M+H]^+$ 441.1309, found 441.1308.

6.4 Chapter 4 Experimental

6.4.1 Imidazole-1-sulfonyl azide (170)[157]

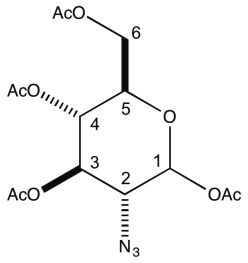


170

To a suspension of sodium azide (3.20 g, 50.0 mmol, 1 equiv.) in dry acetonitrile (50 mL) was added sulfuryl chloride (4.00 mL, 50.0 mmol, 1 equiv.) slowly dropwise at 0 ^{o}C. The reaction was stirred at room temperature overnight. Imidazole (6.80 g, 100.0 mmol, 2 equiv.) was added in portions over a 1 hour period at 0 ^{o}C, before the reaction was allowed to warm to room temperature. The resulting slurry was stirred for 3 hours before being diluted with ethyl acetate (50 mL) and water (50 mL). The phases were separated and the organic washed with water (50 mL), sodium bicarbonate solution (sat., aq., 50 mL) and brine (50 mL) before being dried over magnesium sulfate, filtered and concentrated under reduced pressure to yield the pure product **170** (7.81 g, 45.1 mmol, 90 %).

Rf 0.88 (ethyl acetate); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.02 (s, 1H, H-2), 7.38 (app. t, J = 1.5 Hz, 1H, H-5) and 7.22 (d, J = 0.9 Hz, 1H, H-4); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 136.8, 132.0 and 117.8. The data are in good agreement with the literature values.[169]

6.4.2 1,3,4,6-Tetra-O-acetyl-2-azido-D-glucopyranose (168)



168

Procedure A[155,158]

To a stirred solution of sodium azide (6.03 g, 92.8 mmol, 10 equiv.) in water (32 mL) was added DCM (16 mL). The biphasic mixture was cooled to 0 ^{0}C before triflic anhydride (3.12 mL, 18.6 mmol, 2 equiv.) was added slowly dropwise. The reaction was stirred at this temperature for 2 hours before the phases were separated and the aqueous layer washed with DCM (2 x 8 mL). The organic phases were combined and washed with Na_2CO_3 solution (sat., aq., 10 mL).

To a stirred solution of glucosamine hydrochloride (**131**, 2.00 g, 9.23 mmol, 1 equiv.) in water (32 mL) was added copper sulfate pentahydrate (23.0 mg, 0.09 mmol, 0.01 equiv.), potassium carbonate (1.92 g, 13.9 mmol, 1.5 equiv.) and methanol (32 mL). The triflic azide solution (32 mL) was added slowly at 0 ^{0}C before more methanol (75 mL) was added. The reaction was allowed to warm to room temperature and stirred overnight.

The reaction was concentrated before being taken up in pyridine (30 mL) and acetic anhydride (6.98 mL, 73.8 mmol, 8 equiv.) was added at 0 °C. The mixture was stirred for 16 hours before being quenched with sodium bicarbonate solution (sat., aq., 10 mL), extracted into DCM (3 x 50 mL), washed with sodium bicarbonate (sat., aq., 50 mL), HCl (2 M, aq., 50 mL), water (50 mL) and brine (50 mL). The organic layer was dried over magnesium sulfate, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (silica gel, 25 % ethyl aetate in hexane) to yield **162** as a yellow oil (2.52 g, 6.75 mmol, 73 %).

Procedure B[157]

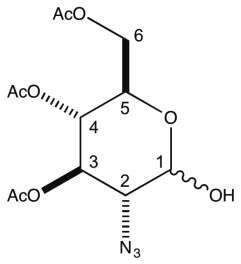
To a stirred suspension of D-glucosamine hydrochloride (**131**, 2.08 g, 9.64 mmol, 1 equiv.) in methanol (40 mL) was added potassium carbonate (2.00 g, 14.5 mmol, 1.5 equiv.), copper sulfate pentahydrate (20.8 mg, 0.10 mmol, 0.01 equiv.) and imidazole-1-sulfonyl azide (**170**, 2.00 g, 11.6 mmol, 1.2 equiv.). The reaction was stirred at room temperature for 2 hours before being concentrated under reduced pressure and co-evaporated with toluene (2 x 20 mL).

Acetic anhydride (9.11 mL, 96.4 mmol, 10 equiv.) was added to the residue in pyridine (30 mL) and the mixture was stirred for 16 hours. The reaction was concentrated before being diluted with water (20 mL) and extracted with ethyl acetate (3 x 20 mL). The

combined organic phases were washed with water (50 mL) and brine (50 mL) before being dried over magnesium sulfate, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (silica gel, 25 % ethyl acetate in hexane) to give **168** as a colourless oil (2.46 g, 6.58 mmol, 68 %).

Rf 0.33 (30 % ethyl acetate in hexane); **^{1}H NMR** (500 MHz, $CDCl_3$, α/β = 33:67) δ 6.29 (d, *J* = 3.7 Hz, 0.33H, H-1α), 5.55 (d, *J* = 8.6 Hz, 0.67H, H-1β), 5.45 (dd, *J* = 10.5, 9.4 Hz, 0.33H, H-3α), 5.14-5.02 (m, 1.67H, H-3β and H-4), 4.30 (dt, *J* = 12.7, 4.3 Hz, 1H, H-6a), 4.11-4.02 (m, 1.33H, H-5α and H-6b), 3.80 (ddd, *J* = 9.7, 4.4, 2.1 Hz, 0.67H, H-5β), 3.70-3.63 (m, 1H, H-2), 2.19 (s, 1H, AcO), 2.18 (s, 2H, AcO), 2.10 (s, 1H, AcO), 2.09 (s, 2H, AcO), 2.07 (s, 3H, AcO), 2.04 (2, 1H, AcO) and 2.02 (s, 2H, AcO); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 170.7, 170.2, 169.9, 169.8, 169.7, 168.7, 168.6, 92.7, 90.1, 72.9, 72.8, 70.9, 69.9, 68.0, 67.9, 62.7, 61.5, 60.4, 21.1, 21.0, 20.9, 20.8, 20.7 and 20.6. The data are in good agreement with the literature values.[170]

6.4.3 3,4,6-Tri-O-acetyl-2-azido-2-deoxy-D-glucopyranose (169)[171]

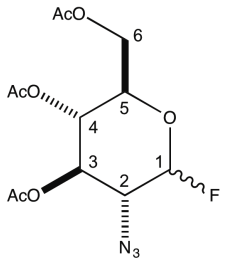


169

To a stirred solution of 1,3,4,6-tetra-O-acetyl-2-azido-2-deoxy-D-glucopyranose (**168**, 0.20 g, 0.54 mmol, 1 equiv.) in diethyl ether (5 mL) was added benzylamine (0.12 mL, 1.08 mmol, 2 equiv.). The reaction was stirred for 16 hours before being concentrated under reduced pressure and the residue dissolved in chloroform (10 mL). The solution was washed with sodium bicarbonate solution (sat., aq., 10 mL), water (10 mL) and brine (10 mL) before being dried over magnesium sulfate, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (silica gel, 50 % ethyl acetate in hexane) to yield **169** (0.14 g, 0.41 mmol, 77 %).

Rf 0.23 (30 % hexane in ethyl acetate); **^{1}H NMR** (500 MHz, $CDCl_3$, α/β = 36:64) δ 5.51 (dd, *J* = 10.5, 9.3 Hz, 0.65H, H-3α), 5.39 (t, *J* = 3.5 Hz, 0.65H, H-1α), 5.08-5.00 (m, 1.35H, H-3β and H-4), 4.73 (dd, *J* = 8.1, 3.8 Hz, 0.35H, H-1β), 4.29-4.08 (m, 3H, H-5, and H-6), 3.74-3.69 (m, 0.35H, H-5β), 3.42 (dd, *J* = 10.5, 3.4 Hz, 1H, H-2), 2.10-2.07 (m, 6H, AcO), 2.04 (s, 2H, AcO) and 2.02 (s, 1H, AcO). The data are in good agreement with the literature values.[172–174]

6.4.4 3,4,6-Tri-O-acetyl-2-azido-2-deoxy-D-glucopyranosyl fluoride (165)[171]



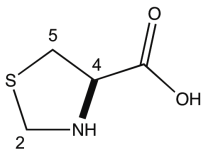
165

To a stirred solution of 3,4,6-tri-*O*-acetyl-2-azido-2-deoxy-D-glucopyranose (**169**, 0.85 g, 2.55 mmol) in THF (50 mL) at -30 °C was added DAST (0.40 mL, 3.06 mmol, 1.2 equiv.) dropwise. The reaction was stirred at this temperature for 10 minutes, before being allowed to warm to room temperature and stirred for 1 hour. The reaction was quenched by addition of methanol (20 mL) at 0 °C and stirred for 20 minutes before being concentrated. The crude material was dissolved in chloroform (50 mL) and washed with sodium bicarbonate solution (sat., aq., 2 x 50 mL) and brine (50 mL), before being dried over magnesium sulfate, filtered and concentrated under reduced pressure to yield the fluoride **165** (0.74 g, 2.22 mmol, 87 %).

Rf 0.27 (30 % ethyl acetate in hexane); **^{1}H NMR** (500 MHz, $CDCl_3$, α/β = 18:82) δ 5.72 (dd, *J* = 51.8, 2.7 Hz, 0.18H, H-1α), 5.48 (dd, *J* = 10.5, 9.4 Hz, 0.82H, H-3α), 5.16 (dd, *J* = 51.4, 7.3 Hz, 0.82H, H-1β), 5.16-5.03 (m, 1.82H, H-3β and H-4), 4.33-4.28 (m, 0.18H, H-6aα), 4.28 (ddd, *J* = 12.5, 4.9, 0.9 Hz, 0.82H, H-6aβ), 4.23-4.20 (m, 0.18H, H-5α), 4.17 (dd, *J* = 12.6, 2.4 Hz, 0.82H, H-6bβ), 3.84-3.76 (m, 1H, H-6bα and H-5β), 3.71-3.63 (m, 0.82H, H-2β), 3.51 (ddd, *J* = 25.6, 10.5, 2.6 Hz, 0.18H, H-2α) and 2.11-2.03 (m, 9H, AcO); **^{19}F NMR** (282 MHz, $CDCl_3$, α/β = 18:82) δ -139.42 (dd, *J* = 51.5,

12.5 Hz, Fβ) and -147.27 (dd, *J* = 52.0, 25.8 Hz, Fα). The data are in good agreement with the literature values.[175]

6.4.5 L-Thiazolidine-4-carboxylic acid (171)[176]

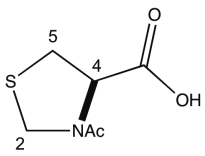


171

Formaldehyde (37 wt %, aq., 1 mL) was added to a stirred solution of L-cysteine (**132**, 1.16 g, 9.64 mmol) in water (4 mL), and the reaction was stirred for 2 hours. Pyridine was added until a pH of 7 was obtained and the product crashed out of solution. L-Thiazolidine-4-carboxylic acid (**171**, 1.00 g, 7.51 mmol, 78 %) was collected by filtration and washed with ice-cold ethanol.

^{1}H NMR (300 MHz, D_2O) δ 4.47-4.29 (m, 3H, H-2 and H-4), 3.40 (dd, *J* = 12.1, 7.4 Hz, 1H, H-5) and 3.29 (dd, *J* = 12.1, 5.8 Hz, 1H, H-5); **^{13}C NMR** (126 MHz, D_2O) δ 172.0, 64.3, 49.1 and 32.1. The data are in good agreement with the literature values.[176]

6.4.6 *N*-Acetyl-L-thiazolidine-4-carboxylic acid (172)[177]

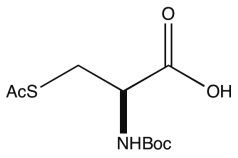


172

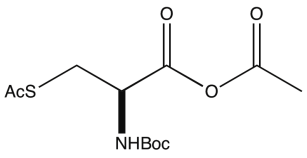
To a stirred solution of L-thiazolidine-4-carboxylic acid (**171**, 50 mg, 0.38 mmol) in pyridine (4 mL) was added acetic anhydride (0.5 mL, 5.30 mmol, 14 equiv). The reaction was stirred for 1 hour before being concentrated under reduced pressure and the residual pyridine removed by co-evaporation with toluene three times to yield *N*-acetyl-L-thiazolidine-4-carboxylic acid (**172**, 66 mg, 0.37 mmol, 99 %) as a white solid.

M.P. 143-144 °C; **^{1}H NMR** (500 MHz, $CDCl_3$) δ 5.04 (dd, *J* = 7.0, 3.5 Hz, 1H, H-4), 4.66-4.49 (m, 2H, H-2), 3.47 (dd, *J* = 12.1, 3.5 Hz, 1H, H-5), 3.24 (dd, *J* = 12.1, 7.0 Hz, 1H, H-5) and 2.24 (s, 3H). The data are in good agreement with the literature values.[178,179]

6.4.7 *N*-Boc-*S*-acetyl cysteine (179) and *N*-Boc-*S*,*O*-diacetyl cysteine (180)[151]



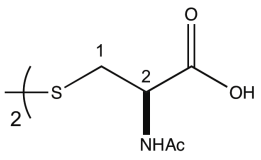
179



180

Zinc (dust, 417 mg, 6.42 mmol, 10 equiv.) was added in portions over a 1 hour period to *N,N'*-bis(Boc)-L-cystine (**178**, 283 mg, 0.64 mmol) in acetic acid (4.5 mL). The reaction was stirred for 2 hours before being filtered and concentrated under reduced pressure. The residue was taken up in pyridine (4.5 mL) before acetic anhydride (0.61 mL, 6.45 mmol, 1.05 equiv.) was added. The reaction was stirred overnight before being quenched with HCl solution (2 M, aq., 2 mL). The product was extracted into ethyl acetate (3 x 10 mL), washed with HCl solution (1 M, aq., 20 mL) and brine (20 mL), and dried over magnesium sulfate, filtered and concentrated under reduced pressure. *N*-Boc-*S*-acetyl-L-cysteine (**179**) was obtained in a mixture with its anhydride **180** in 27 % yield.

6.4.8 *N,N'*-Diacetyl-L-cystine (183)



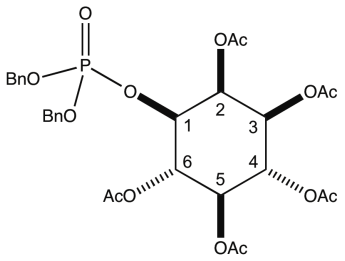
183

To a stirred solution of L-cystine (**182**, 1.00 g, 4.16 mmol) in water (5 mL) was added sodium hydroxide solution (30 % aq., 5 mL) at 0 °C. At this temperature, acetic

anhydride (1 mL) was added slowly dropwise before sodium hydroxide solution (30 % aq., 1 mL) was added slowly dropwise. This process was repeated six times before stirring the reaction at 0 °C for 2 hours. The reaction was concentrated under reduced pressure and the residue taken up in acetic anhydride. Filtration and concentration under reduced pressure yielded *N,N'*-diacetyl-L-cystine (**183**, 1.32 g, 4.08 mmol, 98 %) as a white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 4.68 (dd, *J* = 8.8, 4.4 Hz, 1H, H-2), 3.21 (dd, *J* = 14.3, 4.4 Hz, 1H, H-1a), 3.00 (dd, *J* = 14.3, 8.8 Hz, 1H, H-1b) and 2.06 (s, 3H, AcN); **HRMS** (ES^+) Calc. for $C_{10}H_{17}N_2O_6S_2$ $[M+H]^+$ 325.0523, found 325.0527. The data are in good agreement with the literature values.[180]

6.4.9 2,3,4,5,6-Penta-*O*-acetyl 1-bis(benzyloxy)phosphoryl L-*myo*-inositol (187)

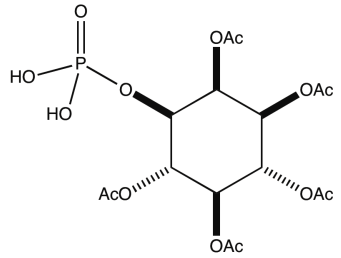


187

To a stirred suspension of 1-bis(benzyloxy)phosphoryl L-*myo*-inositol (**116**, 100 mg, 0.23 mmol, 1 equiv.) in pyridine (5 mL) was slowly added acetic anhydride (1.06 mL, 11.21 mmol, 50 equiv.) and 4-dimethylaminopyridine (3 mg, 0.02 mmol, 0.1 equiv.) at 0 °C. The reaction was stirred at this temperature for 10 minutes before being stirred at room temperature for 5 hours. The reaction was concentrated under reduced pressure and the residue dissolved in dichloromethane (10 mL). The solution was washed with water (10 mL), HCl solution (1 M, 2 x 10 mL), water (10 mL) and brine (10 mL) before being dried over magnesium sulfate, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (silica gel, 50 % ethyl acetate in hexane) to yield **187** as a colourless oil (40 mg, 0.06 mmol, 27 %).

Rf 0.35 (50 % ethyl acetate in hexane); $[\alpha]_D^{20}$ +13.00 (*c* 1.00, $CHCl_3$); **IR** (ATR) 3036, 2965, 1749, 1210, 1009, 999; **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.34-7.18 (10H, m, Ar), 5.66 (1H, t, *J* = 3.0 Hz, H-2), 5.46-5.36 (2H, m, H-6 and H-4), 5.03 (1H, t, *J* = 9.9 Hz, H-5), 5.98-4.83 (5H, m, H-3 and 2 x CH_2), 4.50 (1H, dt, *J* = 3.0, 9.8 Hz, H-1), 2.07 (3H, s, Ac-2), 1.94 (3H, s, Ac-4), 1.93 (6H, s, Ac-3 and Ac-5) and 1.79 (3H, s, Ac-6); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 169.9, 169.7, 169.7, 169.7, 169.5, 128.7, 128.7, 128.7, 128.6, 128.0, 128.0, 73.0 (d, *J* = 5.3 Hz), 70.7, 70.0 (d, *J* = 5.0 Hz), 69.9 (d, *J* = 5.9 Hz), 69.6 (d, *J* = 5.7 Hz), 69.3, 68.9 (d, *J* = 2.81 Hz), 68.5, 20.7, 20.6 and 20.5; **^{31}P NMR** (202 MHz, $CDCl_3$) δ -1.76; **HRMS** (ES^+) Calc. for $C_{30}H_{36}O_{14}P$ $[M+H]^+$ 651.1837, found 651.1823.

6.4.10 2,3,4,5,6-penta-*O*-acetyl 1-phosphoryl L-myo-inositol (188)

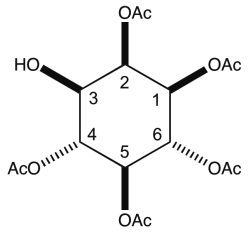


188

To a stirred solution of pentaacetate **187** (6.3 mg, 9.73 μmol) in ethyl acetate (3 mL) was added palladium on carbon (10 wt % loading, 10 mg, 0.09 mmol). The reaction was purged with hydrogen and left under a positive pressure of hydrogen with stirring overnight. The reaction was filtered through Celite and washed through with ethyl acetate, before being concentrated under reduced pressure. The product **188** (4 mg) was used in the next step without further purification.

Rf 0.05 (10 % methanol in DCM).

6.4.11 1,2,4,5,6-penta-*O*-acetyl D-myo-inositol (189)



189

Crude phosphate **188** (4 mg, 8.9 μmol) was taken up in a mixture of sodium acetate buffer (50 mM, pH 5.2, 860 μL) and DMSO (40 μL) before lyophilised acid phosphatase (from potato, 3.3 Umg^{-1}, 1 mg) was added and the reaction incubated with shaking at 37 °C overnight. The product was extracted into ethyl acetate (1 mL x 5) before being concentrated under reduced pressure to yield **189** (3 mg, 7.7 μmol, 79 % over 2 steps).

Rf 0.56 (10 % methanol in DCM); $[\alpha]_D^{20}$ +3.00 (*c* 0.20, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 5.62 (1H, t, *J* = 3.0 Hz, H-2), 5.48 (1H, t, *J* = 10.2 Hz, H-4), 5.33 (1H, t, *J* = 10.0 Hz, H-6), 5.17 (1H, t, *J* = 9.8 Hz, H-5), 4.99 (1H, dd, *J* = 10.6, 2.8 Hz, H-3), 3.91 (1H, dd, *J* = 10.1, 3.0 Hz, H-1), 2.24 (3H, s, Ac), 2.12 (3H, s, Ac), 2.05 (3H, s Ac), 2.04 (3H, s, Ac) and 2.02 (3H, s, Ac); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 171.1, 170.3, 169.9, 169.8, 169.8, 72.5, 70.7, 70.5, 69.4, 69.1, 69.0, 20.9, 20.8, 20.6, 20.6 and 20.6. The data are in good agreement with the literature values.[181,182]

7 References

1. Michell, R. H., *Nat. Rev. Mol. Cell Biol.* **9,** 151–161 (2008).
2. Scherer, J., *Justus Liebigs Ann. Chem.* **73,** 322–328 (1850).
3. Potter, B. V. L., Lampe, D., *Angew. Chem. Int. Ed. Engl.* **34,** 1933–1972 (1995).
4. Nomenclature Committee of the International Union of Biochemistry. *Biochem. J.* **258,** 1–2 (1989).
5. Parthasarathy, R., Eisenberg, F., *Biochem. J.* **235,** 313–322 (1986).
6. Pasari, S., Ismail, S. M., Wenk, M. R., Lear, M. J., *Tetrahedron Lett.* **56,** 2597–2601 (2015).
7. Agranoff, B. W., Roy, M., *J. Biol. Chem.* **233,** 1077–1083 (1958).
8. Best, M. D., Zhang, H., Prestwich, G. D., *Nat. Prod. Rep.* **27,** 1403–1430 (2010).
9. Bootman, M. D., *Cold Spring Harb. Perspect. Biol.* **4,** a011171–a011171 (2012).
10. Foskett, J. K., White, C., Cheung, K., Mak, D. D., *Physiol Rev.* **87,** 593–658 (2007).
11. Orrenius, S., Zhivotovsky, B., Nicotera, P., *Nat. Rev. Mol. Cell Biol.* **4,** 552–565 (2003).
12. Conway, S. J., Miller, G. J., *Nat. Prod. Rep.* **24,** 687–707 (2007).
13. Mak, D. O., McBride, S., Foskett, J. K., *J. Gen. Physiol.* **117,** 435–446 (2001).
14. Sasakawa, N., Sharif, M., Hanley, M. R., *Biochem. Pharmacol.* **50,** 137–146 (1995).
15. Cho, W., Stahelin, R. V., *Annu. Rev. Biophys. Biomol. Struct.* **34,** 119–151 (2005).
16. Paulick, M. G., Bertozzi, C. R., *Biochemistry* **47,** 6991–7000 (2008).
17. Ferguson, M. A., *J. Cell Sci.* **112,** 2799–2809 (1999).
18. McConville, M. J., Ferguson, M. A., *Biochem. J.* **294 (Pt 2),** 305–24 (1993).
19. Barbieri, L., Costantino, V., Fattorusso, E., Mangoni, A., *J. Nat. Prod.* **68,** 1527–1530 (2005).
20. Costantino, V., Della Sala, G., Mangoni, A., Perinu, C., Teta, R., *Eur. J. Org. Chem.* **2012,** 5171–5176 (2012).
21. Lee, S., Rosazza, J. P. N., *Org. Lett.* **6,** 365–368 (2004).
22. Martin, K. L., Smith, T. K., *Mol. Microbiol.* **61,** 89–105 (2006).
23. Majumder, A. L., Johnson, M. D., Henry, S., *Biochim. Biophys. Acta - Lipids Lipid Metab.* **1348,** 245–256 (1997).

24. Majumder, A. L., Chatterjee, A., Ghosh Dastidar, K., Majee, M., *FEBS Lett.* **553,** 3–10 (2003).
25. Geiger, J. H., Jin, X., *Biol. Inositols Phosphoinositides* 157–180 (2006).
26. Loewus, F. A., Kelly, S., *Biochem. Biophys. Res. Commun.* **7,** 204–208 (1962).
27. Bender, S. L., Widlanski, T., Knowles, J. R., *Biochemistry* **28,** 7560–72 (1989).
28. Widlanski, T., Bender, S. L., Knowles, J. R., *J. Am. Chem. Soc.* **111,** 2299–2300 (1989).
29. Lee, J. K., Houk, K. N., *Science* **276,** 942–945 (1997).
30. Schmidt, D. E., Westheimert, F. H., *Biochemistry* **10,** 1249–1253 (1971).
31. Ferrier, R. J., *J. Chem. Soc. Perkin Trans. 1* 1455 (1979).
32. Conrad, R. M., Grogan, M. J., Bertozzi, C. R., *Org. Lett.* **4,** 1359–61 (2002).
33. Luchetti, G., Ding, K., Kornienko, A., d'Alarcao, M., *Synthesis* **2008,** 3148–3154 (2008).
34. Li, J. J., *Name Reactions* 245–246 (2014).
35. Blattner, R., Ferrier, R. J., Haines, S. R., *J. Chem. Soc. Perkin Trans. 1* 2413 (1985).
36. Chretien, F., Chapleur, Y., *J. Chem. Soc. Chem. Commun.* 1268 (1984).
37. Machado, A. S., Dubreuil, D., Cleophax, J., Gero, S. D., Thomas, N. F., *Carbohydr. Res.* **233,** C5–C8 (1992).
38. Chida, N., Ohtsuka, M., Ogawa, S., Ogura, K., *Bull. Chem. Soc. Jpn.* **64,** 2118 (1991).
39. Chida, N., Ohtsuka, M., Nakazawa, K., Ogawa, S., *J. Org. Chem.* **56,** 2976–2983 (1991).
40. Chida, N., Ohtsuka, M., Ogawa, S., *Tetrahedron Lett.* **32,** 4525–4528 (1991).
41. Adam, S., *Tetrahedron Lett.* **29,** 6589–6592 (1988).
42. Das, S. K., Mallet, J.-M., Sinaÿ, P., *Angew. Chem. Int. Ed. Engl.* **36,** 493–496 (1997).
43. Sollogoub, M., Mallet, J., Sinaÿ, P., *Tetrahedron Lett.* **39,** 3471–3472 (1998).
44. Takahashi, H., Kittaka, H., Ikegami, S., *J. Org. Chem.* **66,** 2705–2716 (2001).
45. Iimori, T., Takahashi, H., Ikegami, S., *Tetrahedron Lett.* **37,** 649–652 (1996).
46. Takahashi, H., Kittaka, H., Ikegami, S., *Tetrahedron Lett.* **39,** 9703–9706 (1998).
47. Jia, C., Pearce, A. J., Blériot, Y., Zhang, Y., Zhang, L.-H., Sollogoub, M., Sinaÿ, P., *Tetrahedron Asymmetry* **15,** 699–703 (2004).
48. Keddie, N. S., Bultynck, G., Luyten, T., Slawin, A. M. Z., Conway, S. J.,

Tetrahedron Asymmetry **20,** 857–866 (2009).

49. Tegge, W., Ballou, C. E., *PNAS U. S. A.* **86,** 94–8 (1989).
50. Sureshan, K. M., Shashidhar, M. S., Praveen, T., Das, T., *Chem. Rev.* **103,** 4477–4503 (2003).
51. Kiddle, J. J., *Chem. Rev.* **95,** 2189–2202 (1995).
52. Qiao, L., Hu, Y., Nan, F., Powis, G., Kozikowski, A. P., *Org. Lett.* **2,** 115–117 (2000).
53. Akiyama, T., Takechi, N., Ozaki, S., *Tetrahedron Lett.* **31,** 1433–1434 (1990).
54. Ley, S. V, Pzura, M., Redgrave, A. J., *Tetrahedron* **46,** 4995–5026 (1990).
55. Ley, S. V., Sternfeld, F., *Tetrahedron Lett.* **29,** 5305–5308 (1988).
56. Trost, B. M., Patterson, D. E., Hembre, E. J., August, R. V., *J. Am. Chem. Soc.* **121,** 10834–10835 (1999).
57. Podeschwa, M. A. L., Plettenburg, O., Altenbach, H.-J., *Eur. J. Org. Chem.* **2005,** 3101–3115 (2005).
58. Desai, T., Gigg, J., Gigg, R., Payne, S., Penades, S., Rogers, H. J., *Carbohydr. Res.* **216,** 197–209 (1992).
59. Haines, A. H., *Adv. Carbohydr. Chem. Biochem* **33,** 11–109 (1976).
60. Knapp, S., Kukkola, P. J., Sharma, S., Dhar, T. G. M., Naughton, A. B. J., *J. Org. Chem.* **55,** 5700–5710 (1990).
61. Angyal, S. J., Macdonald, C. G., *J. Chem. Soc.* 686 (1952).
62. Angyal, S. J., Tate, M. E., *J. Chem. Soc.* 4116–4122 (1961).
63. Ozaki, S., Watanabe, Y., Ogasawara, T., Kondo, Y., Shiotani, N., Nishii, H., Matsuki, T., *Tetrahedron Lett.* **27,** 3157–3160 (1986).
64. Lee, H. W., Kishi, Y., *J. Org. Chem.* **50,** 4402–4404 (1985).
65. Billington, D. C., Baker, R., Kulagowski, J. J., Mawer, I. M., Vacca, J. P., deSolms, S. J., Huff, J. R., *J. Chem. Soc. Perkin Trans. 1* 1423 (1989).
66. Grove, S. J. A., Gilbert, I. H., Holmes, A. B., Painter, G. F., Hill, M. L., *Chem. Commun.* **10,** 1633–1634 (1997).
67. Painter, G. F., Grove, S. J. A., Gilbert, I. H., Holmes, A. B., Raithby, P R., Hill, M. L., Hawkins, P. T., Stephens, L. R., *J. Chem. Soc. Perkin Trans. 1* 923–936 (1999). doi:10.1039/a900278b
68. Ling, L., Watanabe, Y., Akiyama, T., Ozaki, S., *Tetrahedron Lett.* **33,** 1911–1914 (1992).
69. Sculimbrene, B. R., Morgan, A. J., Miller, S. J., *J. Am. Chem. Soc.* **124,** 11653–

11656 (2002).

70. Sculimbrene, B. R., Morgan, A. J., Miller, S. J., *Chem. Commun.* 1781–1785 (2003).
71. Chandler, B. D., Burkhardt, A. L., Foley, K., Cullis, C., Driscoll, D., D'Amore, N. R., Miller, S. J., *J. Am. Chem. Soc.* **136,** 412–8 (2014).
72. Morgan, A. J., Komiya, S., Xu, Y., Miller, S. J., *J. Org. Chem.* **71,** 6923–31 (2006).
73. Xu, Y., Sculimbrene, B. R., Miller, S. J., *J. Org. Chem.* **71,** 4919–28 (2006).
74. Kayser-Bricker, K. J., Jordan, P. A., Miller, S. J., *Tetrahedron* **64,** 7015–7020 (2008).
75. Sculimbrene, B. R., Xu, Y., Miller, S. J., *J. Am. Chem. Soc.* **126,** 13182–3 (2004).
76. Ciulla, R. a., Burggraf, S., Stetter, K. O., Roberts, M. F., *Appl. Environ. Microbiol.* **60,** 3660–3664 (1994).
77. Longo, C. M., Wei, Y., Roberts, M. F., Miller, S. J., *Angew. Chem. Int. Ed. Engl.* **48,** 4158–61 (2009).
78. Rosche, B., Li, X. Z., Hauer, B., Schmid, A., Buehler, K., *Trends Biotechnol.* **12,** 636–643 (2009).
79. Reetz, M. T., *J. Am. Chem. Soc.* **135,** 12480–12496 (2013).
80. Schoemaker, H. E., *Science* **299,** 1694–1697 (2003).
81. Bornscheuer, U. T., Kazlauskas, R. J., *Angew. Chem. Int. Ed. Engl.* **43,** 6032–6040 (2004).
82. Ulmer, K. M., *Science* **219,** 666–671 (1983).
83. Damian-Almazo, J. Y., Saab-Rincon, G.. *Genetic Manipulation of DNA and Protein - Examples from Current Research, First Edition,* 303–304 (2013).
84. Chen, K., Arnold, F. H., *PNAS U.S.A.* **90,** 5618–5622 (1993).
85. Desai, A. A., *Angew. Chem. Int. Ed. Engl.* **50,** 1974–1976 (2011).
86. Choi, J., Han, S., Kim, H., *Biotechnol. Adv.* **33,** 1443–1454 (2015).
87. Jarvis, W. R., Colbeck, J. C., Krebber, A., Fleitz, F. J., Brands, J., *Science* **329,** 305–309 (2010).
88. Winn, M., Foulkes, J. M., Perni, S., Simmons, M. J. H., Overton, T. W., Goss, R. J. M, *Catal. Sci. Technol.* **2,** 1544–1547 (2012).
89. Hudlicky, T., Reed, J. W., *Chem. Soc. Rev.* **38,** 3117–3132 (2009).
90. Mateo, C., Palomo, J. M., Fernandez-lorente, G., Guisan, J. M., Fernandez-lafuente, R., *Enzym. Microb. Technol.* **40,** 1451–1463 (2007).

91. Hernandez, K., Fernandez-lafuente, R., *Enzym. Microb. Technol.* **48,** 107–122 (2011).
92. Hansen, C. A., Dean, A. B., Draths, K. M., Frost, J. W., *J. Am. Chem. Soc.* 3799–3800 (1999).
93. Catalysis, S., Vetting, M. W., Frantom, P. A., Blanchard, J. S., *J. Biol. Chem.* **283,** 15834–15844 (2008).
94. Chen, L., Zhou, C., Yang, H., Roberts, M. F., *Biochemistry* 12415–12423 (2000).
95. Smith, T. K., Bütikofer, P., *Mol. Biochem. Parasitol.* **172,** 66–79 (2010).
96. Smith, T. K., Young, B. L., Denton, H., Hughes, D. L., Wagner, G. K., *Bioorganic Med. Chem. Lett.* **19,** 1749–1752 (2009).
97. World Health Organization. Control and surveillance of human African trypanosomiasis. *World Health Organ. Tech. Rep. Ser.* 1–237 (2013).
98. Pays, E., van Hollebeke, B., van Hamme, L., Patrieux-Hanocq, F., Nolan, D. P., Perez-Morga, D., *Nat. Rev. Microbiol.* **4,** 477–486 (2006).
99. Wolburg, H., Mogk, S., Acker, S., Frey, C, Meinert, M., Schönfeld, C., Lazarus, M., Urade, Y., Kubata, B. K., Duszenko, M., *PLoS One* **7,** e34304 (2012).
100. Deborggraeve, S., Koffi, M., Jamonneau, V., Bonsu, F. A., Queyson, R., Simarro, P. P., Herdewijn, P., Büscher, P., *Diagn. Microbiol. Infect. Dis.* **61,** 428–433 (2008).
101. Simarro, P. P., Diarra, A., Ruiz Postigo, J. A., Franco, J. R., Jannin, J. G., *PLoS Negl. Trop. Dis.* **5,** e1007 (2011).
102. Major, L. L., Smith, T. K., *Mol. Biol. Int.* **2011,** 1–14 (2011).
103. Freymann, D., Down, J., Carrington, M., Roditi, I., Turner, M., Wiley, D., *J. Mol. Biol.* **216,** 141–160 (1990).
104. González-Salgado, A., Steinmann, M., Major, L. L., Sigel, E., Reymond, J.-L., Smith, T. K., Bütikofer, P., *Eukaryot. Cell* **14,** 616–624 (2015).
105. Ju, S., Shaltiel, G., Shamir, A., Agam, G., Greenberg, M. L., *J. Biol. Chem.* **279,** 21759–65 (2004).
106. Major, L. L., *personal communication* (2011).
107. Donahue, F. T., Henry, A. S., *J. Biol. Chem.* **256,** 7077–7085 (1981).
108. Johnson, M. D., Sussex, I. M., *Plant Physiol.* **107,** 613–619 (1995).
109. Wong, C. H., Whitesides, G. M. *Enzymes in Synthetic Organic Chemistry*, *First Edition.* Elsevier Science Ltd (1994).
110. Sheldon, R. A., van Pelt, S., *Chem. Soc. Rev.* **42,** 6223–6235 (2013).

111. Mak, X. Y., Laurino, P., Seeberger, P. H., *Bellstein J. Org. Chem.* **11,** 1–11 (2009).
112. Matsuda, T., Watanabe, K., Harada, T., Nakamura, K., Arita, Y., *Chem. Commun.* 2286–2287 (2004).
113. Denton, H., *personal communication* (2015).
114. Koy, C., Michalik, M., Oehme, G., Fischer, C., Tillack, A., Baidisch, H., Kempe, R., *Phosphorus. Sulfur. Silicon Relat. Elem.* **152,** 203–228 (2016).
115. Janiak, A. M., Hoffmann, M., Milewska, J., Milewski, S., *Bioorganic Med. Chem.* **11,** 1653–1662 (2003).
116. Dudman, C. C., Reese, C. B., *Synthesis* **1982,** 419–421 (1982).
117. Pizer, F. L., Ballou, C. E., *J. Am. Chem. Soc.* **81,** 915–921 (1959).
118. Majee, M., Maitra, S., Dastidar, K. G., Pattnick, S., Chatterjee, A., Hait, N C., Das, K. P., Majumder, A. L., *J. Biol. Chem.* **279,** 28539–28552 (2004).
119. Norman, R. A., McAllister, M. S. B., Murray-Rust, J., Movahedzadeh, F., Stoker, N. G., McDonald, N. Q., *Structure* **10,** 393–402 (2002).
120. Movahedzadeh, F., Norman, R. A., Dinadayala, P., Murray-Rust, J., Russell, D. G., Kendall, S. L., Rison, S. C., McAllister, M. S., Bancroft, G. J., McDonald, N. Q., Daffe, M., Av-Gay, Y., Stoker, N. G., *Mol. Microbiol.* **51,** 1003–1014 (2004).
121. Haites, R. E., Morita, Y. S., McConville, M. J., Billman-Jacobe, H., *J. Biol. Chem.* **280,** 10981–10987 (2005).
122. Park, S., Kim, J., *J. Microbiol.* **42,** 20–24 (2004).
123. Lohia, A., Hait, N. C., Majumder, A. L., *Mol. Biochem. Parasitol.* **98,** 67–79 (1999).
124. Ilg, T., *Mol. Biochem. Parasitol.* **120,** 151–156 (2002).
125. Guan, G., Dai, P., Shechter, I., *Arch. Biochem. Biophys.* **417,** 251–259 (2003).
126. Swan, M. K., Hansen, T., Schönheit, P., Davies, C., *J. Biol. Chem.* **279,** 39838–45 (2004).
127. Bornscheuer, U. T., *OCL - Ol. Corps Gras Lipides* **20,** 45–49 (2013).
128. Zhao, H., van der Donk, W. A., *Curr. Opin. Biotechnol.* **14,** 583–589 (2003).
129. Jothivasan, V. K., Hamilton, C. J., *Nat. Prod. Rep.* **25,** 1091–1117 (2008).
130. Hernick, M., *Expert Rev. Anti. Infect. Ther.* **11,** 49–67 (2013).
131. World Health Organization. Global tuberculosis report 2014. (2014).
132. Dube, D., Agrawal, G. P., Vyas, S. P., *Drug Discov. Today* **17,** 760–773 (2012).
133. Mukhopadhyay, S., Nair, S., Ghosh, S., *FEMS Microbiol. Rev.* **36,** 463–485 (2012).

134. American Thoracic Society/Centers for Disease Control/Infectious Diseases Society of America. *Morb. Mortal. Wkly. Rep. Recomm. reports* **52,** 1–77 (2003).
135. Centers for Disease Control and Prevention (CDC). *Morb. Mortal. Wkly. Rep.* **55,** 301–5 (2006).
136. da Silva, P. E. A., Palomino, J. C., *J. Antimicrob. Chemother.* **66,** 1417–1430 (2011).
137. Zumla, A., Nahid, P., Cole, S. T., *Nat. Rev. Drug Discov.* **12,** 388–404 (2013).
138. World Health Organization. *Treatment of tuberculosis: guidelines*. *4th Edition* (2010).
139. Patel, M. P., Blanchard, J. S., *J. Am. Chem. Soc.* **120,** 11538–11539 (1998).
140. Rawat, M., Av-Gay, Y., *FEMS Microbiol. Rev.* **31,** 278–292 (2007).
141. Newton, G. L., Fahey, R. C., Cohen, G., Aharonowitz, Y., *J. Bacteriol.* **175,** 2734–2742 (1993).
142. Fahey, R. C., *Annu. Rev. Microbiol.* **55,** 333–56 (2001).
143. Sakuda, S., Zhou, Z. Y., Yamada, Y., *Biosci. Biotechnol. Biochem.* **58,** 1347–1348 (1994).
144. Newton, G. L., Bewley, C. A., Dwyer, T. J., Horn, R., Aharonowitz, Y., Cohen, G., Davies, J., Faulkner, D. J., Fahey, R. C., *Eur. J. Biochem.* **230,** 821–825 (1995).
145. Newton, G. L., Arnold, K., Price, M. S., Sherill, C., Delcardayre, S. B., Aharonowitz, Y., Cohen, G., Davies, J., Fahey, R. C., Davis, C., *J. Bacteriol.* **178,** 1990–1995 (1996).
146. Riordan, S. W., Field, J. J., Corkran, H. M., Dasyam, N., Stocker, B. L., Timmer, M. S., Harvey, J. E., Teesdale-Spittle, P. H., *Bioorg. Med. Chem. Lett.* **25,** 2152–2155 (2015).
147. Knapp, S., Gonzalez, S., Myers, D. S., Eckman, L. L., Bewley, C. A., *Org. Lett.* **4,** 4337–4339 (2002).
148. Newton, G. L., Buchmeier, N., Fahey, R. C., *Microbiol. Mol. Biol. Rev.* **72,** 471–494 (2008).
149. Buchmeier, N., Fahey, R. C., *FEMS Microbiol. Lett.* **264,** 74–79 (2006).
150. Sareen, D., Newton, G. L., Fahey, R. C., Buchmeier, N. A., *J. Bacteriol.* **185,** 6736–6740 (2003).
151. Ajayi, K., Thakur, V. V, Lapo, R. C., Knapp, S., *Org. Lett.* **12,** 2630–3 (2010).
152. Koto, S., Hirooka, M., Yoshida, T., Takenaka, K., Asai, C., Nagamitsu, T., Sakuma, H., Sakurai, M., Masuzawa, M., Komiya, M., Sato, T., Zen, S., Yago, K.,

Tomonaga, F., *Bull. Chem. Soc. Jpn.* **73,** 2521 (2000).

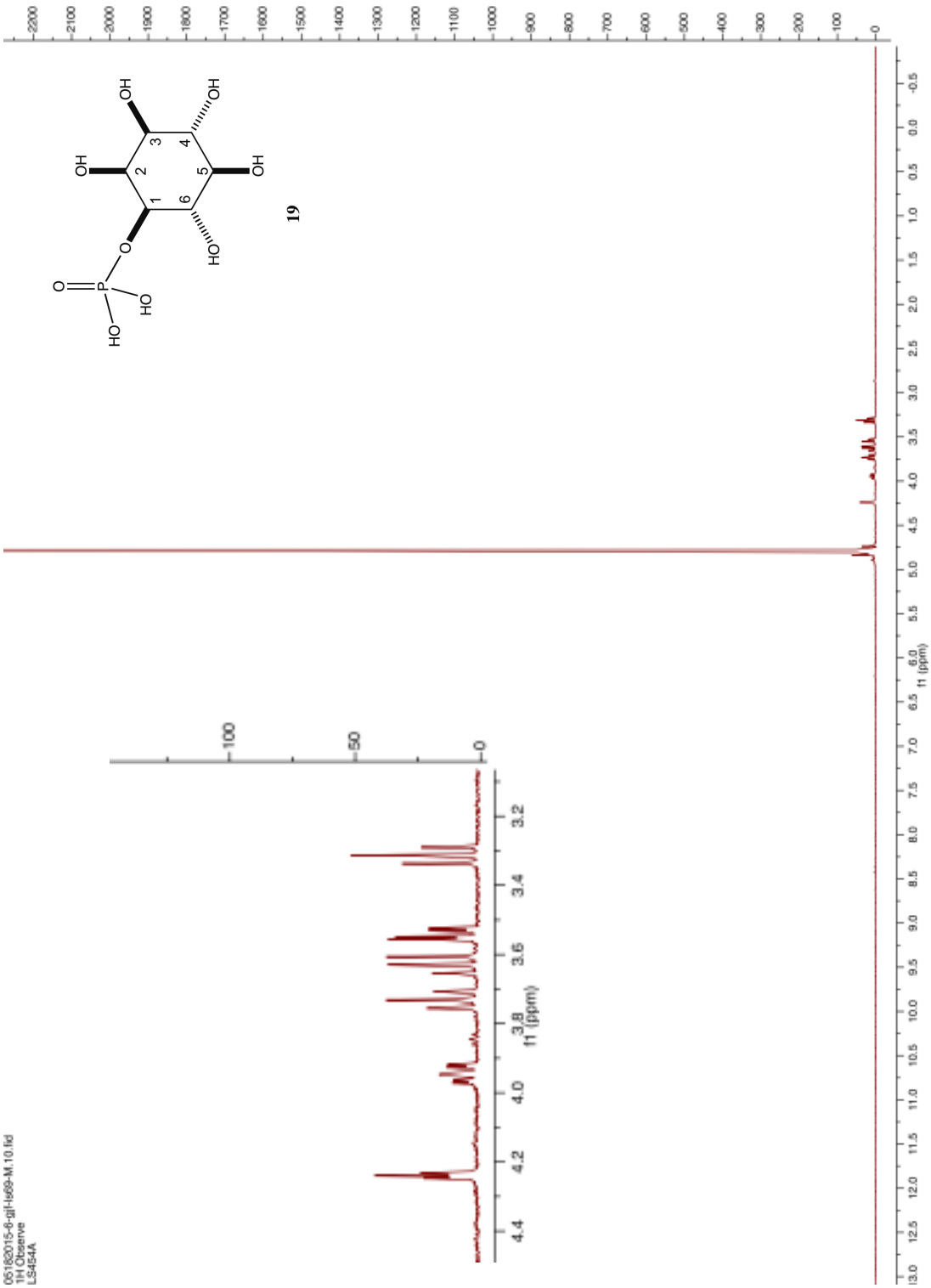
153. Billington, D. C., Baker, R., Kulagowski, J. J., Mawer, I. M., *J. Chem. Soc. Chem. Commun.* **1,** 314 (1987).
154. Bousquet, E., Spadaro, A., Pappalardo, M. S., Bernardini, R., Romeo, R., Panza, L., Ronsisvalle, G. J., *J. Carbohydr. Chem.* **19,** 527–541 (2000).
155. Orgueira, H. A., Bartolozzi, A., Schell, P., Litjens, R. E. J. N., Palmacci, E. R., Seeberger, P. H., *Chem. Eur. J.* **9,** 140–169 (2003).
156. Yan, R.-B., Yang, F., Wu, Y., Zhang, L.-H., Ye, X.-S., *Tetrahedron Lett.* **46,** 8993–8995 (2005).
157. Goddard-Borger, E. D., Stick, R. V., *Org. Lett.* **9,** 3797–800 (2007).
158. Nyffeler, P. T., Liang, C. H., Koeller, K. M., Wong, C. H., *J. Am. Chem. Soc.* **124,** 10773–10778 (2002).
159. Garner, J., Jolliffe, K. a, Harding, M. M., Payne, R. J., *Chem. Commun.* 6925 (2009). doi:10.1039/b918021d
160. Watzig, D., Aigner, W., *Pharmazie* **49,** 249–252 (1994).
161. Bloom, D. M., Park, M., Hill, M., *United States Patent No. 4827016* (1982).
162. Chittenden, G. J. F., *Carbohydr. Res.* **91,** 85–88 (1981).
163. Liotta, L., Ganem, B., *Tetrahedron Lett.* 4759–4762 (1989).
164. Eckenberg, P., Groth, U., Huhn, T., Richter, N., Schmeck, C., *Tetrahedron* **49,** 1619–1624 (1993).
165. Itaya, K., Ui, M., *Clin. Chim. Acta.* **14,** 361–6 (1966).
166. Capolicchio, S., Thakor, D. T., Linden, A., Jessen, H. J., *Angew. Chem. Int. Ed. Engl.* **52,** 6912–6916 (2013).
167. Pedersen, P. J., Henriksen, J., Gotfredsen, C. H., Clausen, M. H., *Tetrahedron Lett.* **49,** 6220–6223 (2008).
168. Kabalka, G. W., Maddox, J. T., Bogas, E., Kelley, S. W., *J. Org. Chem.* **62,** 3688–3695 (1997).
169. Ye, H., Liu, R., Li, D., Liu, Y., Yuan, H., Guo, W., Zhou, L., Cao, X., Tian, H., Shen, J., Wang, P. G., *Org. Lett.* **15,** 18–21 (2013).
170. Vasella, A., Witzig, C., Chiara, J. L., Martin-Lomas, M., *Helv. Chim. Acta* **74,** 2073–2077 (1991).
171. Dasgupta, F., Masada, R. I., *Carbohydr. Res.* **337,** 1055–1058 (2002).
172. Chen, M.-Y., Patkar, L. N., Chen, H.-T., Lin, C.-C., *Carbohydr. Res.* **338,** 1327–32 (2003).

173. Rele, S. M., Iyer, S. S., Baskaran, S., Chaikof, E. L., *J. Org. Chem.* **69,** 9159–9170 (2004).
174. Iyer, S. S., Rele, S. M., Baskaran, S., Chaikof, E. L., *Tetrahedron* **59,** 631–638 (2003).
175. Palme, M., Vasella, A., *Helv. Chim. Acta* **78,** 959–969 (1995).
176. Kaupp, G., Schmeyers, J., Boy, J., *Tetrahedron* **56,** 6899–6911 (2000).
177. Ratner, S., Clarke, H. T., *J. Am. Chem. Soc.* **59,** 200–206 (1937).
178. West, N. P., Cergol, K. M., Xue, M., Randall, E. J., Britton, W. J., Payne, R J., *Chem. Commun.* **47,** 5166 (2011).
179. Croce, P. D., Ferracciol, R., *Tetrahedron* **4020,** 9385–9392 (1995).
180. Oba, M., Tanaka, K., Nishiyama, K., Ando, W., *J. Org. Chem.* **76,** 4173–4177 (2011).
181. Sibrikov, Y., Stepanov, A., Shvets, V., *J. Org. Chem. USSR (English Transl.)* **26,** 899–904 (1990).
182. Anraku, K., Inoue, T., Sugimoto, K., Morii, T., Mori, Y., Okamoto, Y., Otsuka, M., *Org. Biomol. Chem.* **6,** 1822–1830 (2008).

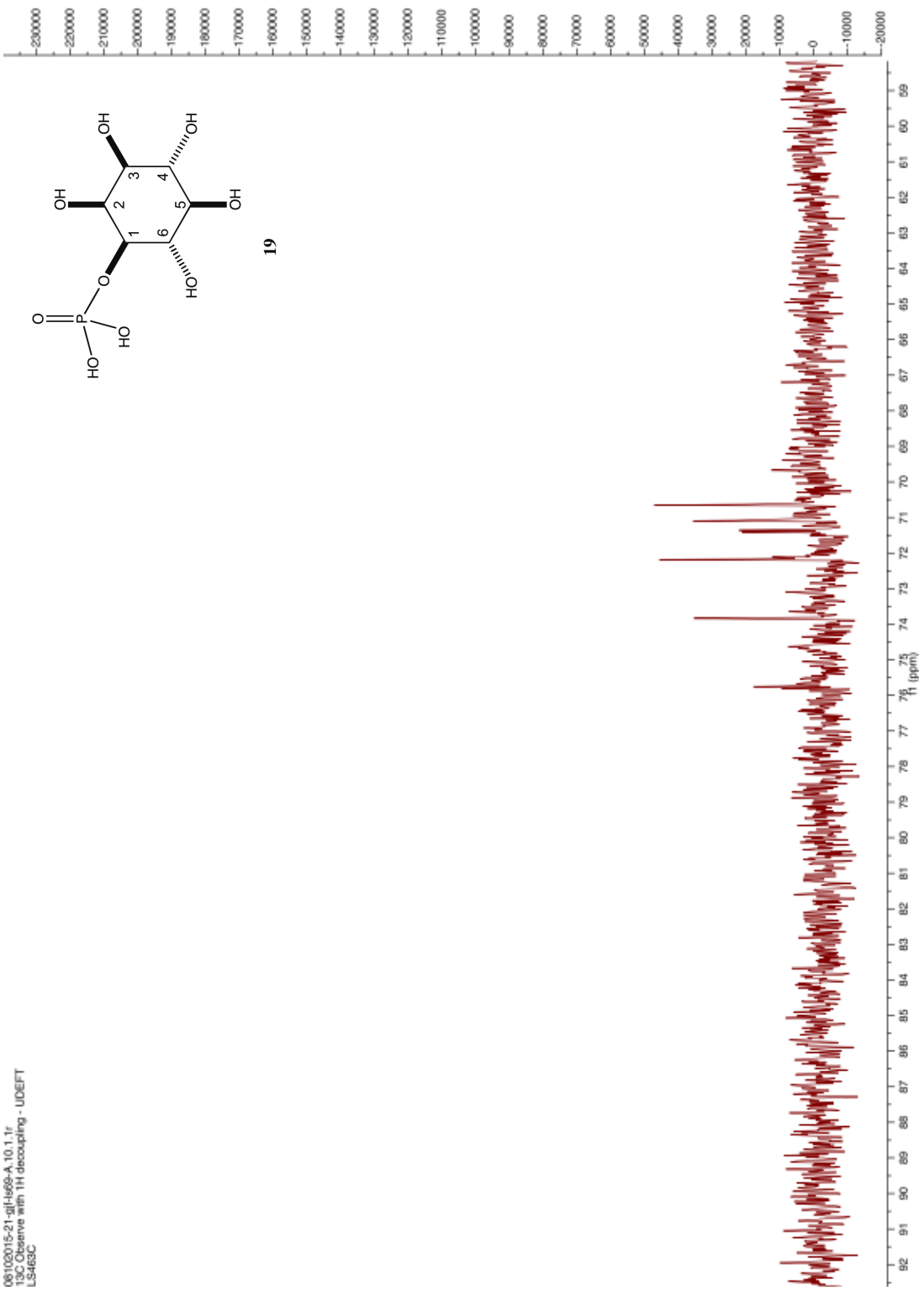
Appendix

Selected NMR Spectra

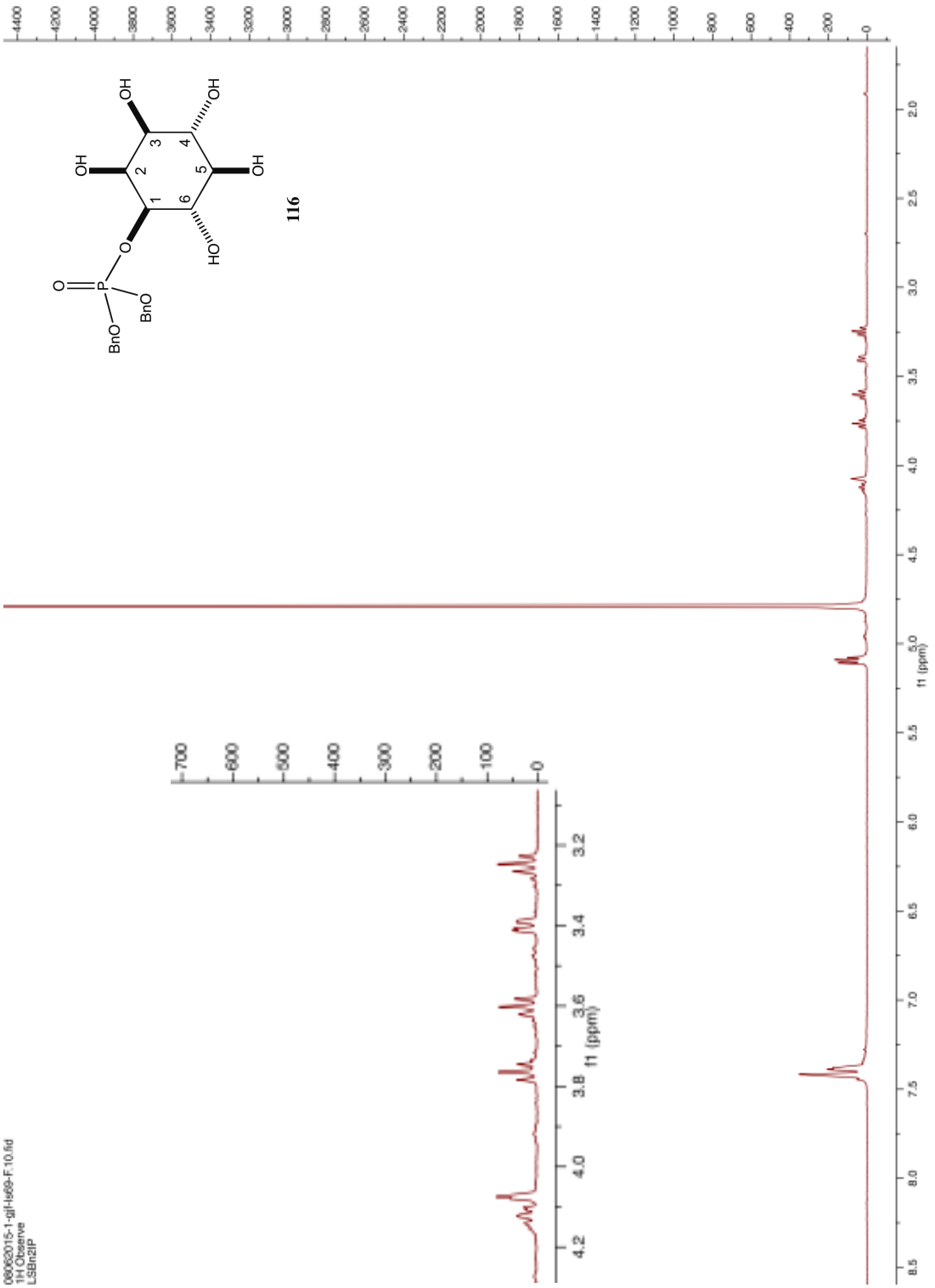
L-*myo*-Inositol 1-phosphate (19) ^{1}H NMR



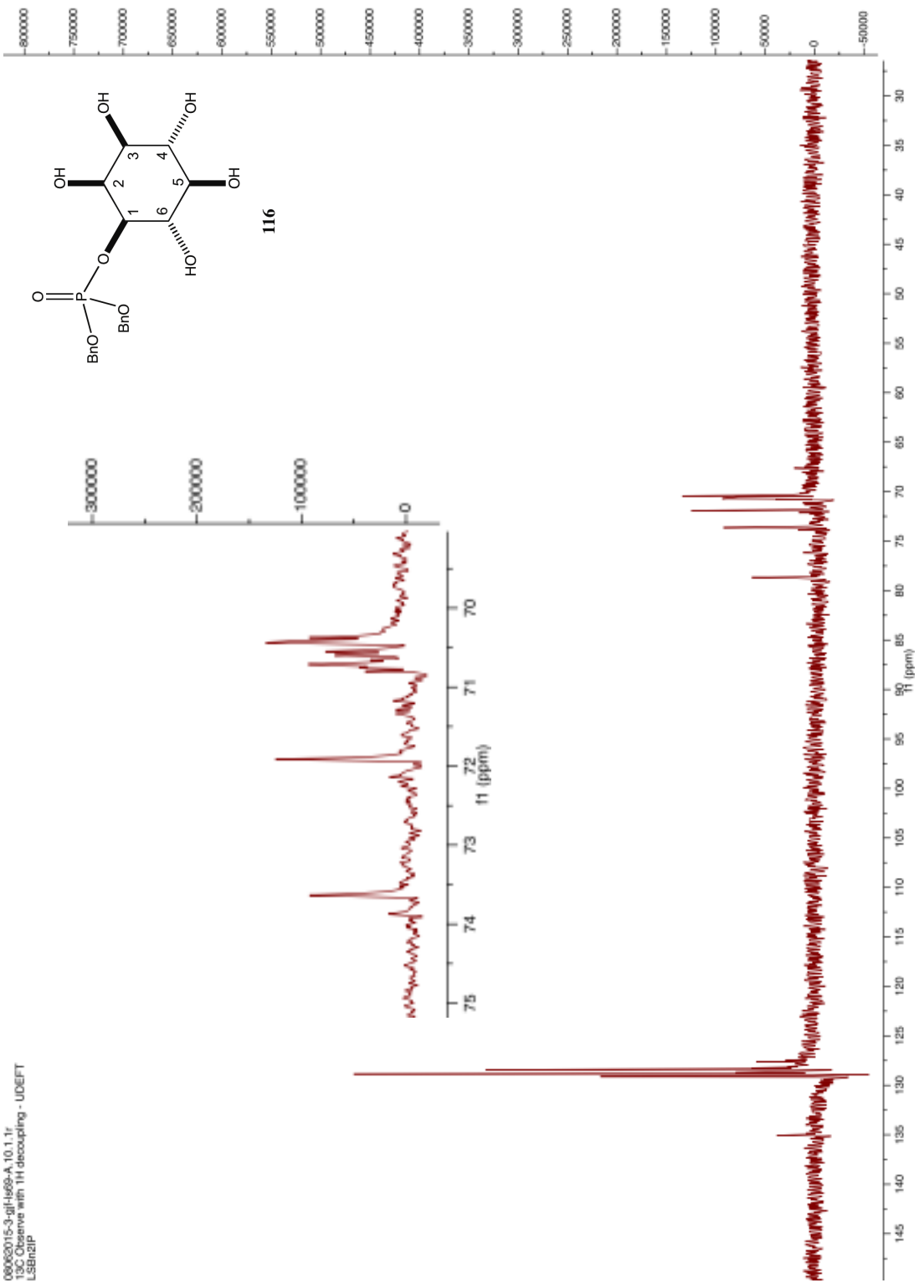
L-*myo*-Inositol 1-phosphate (19) ^{13}C NMR



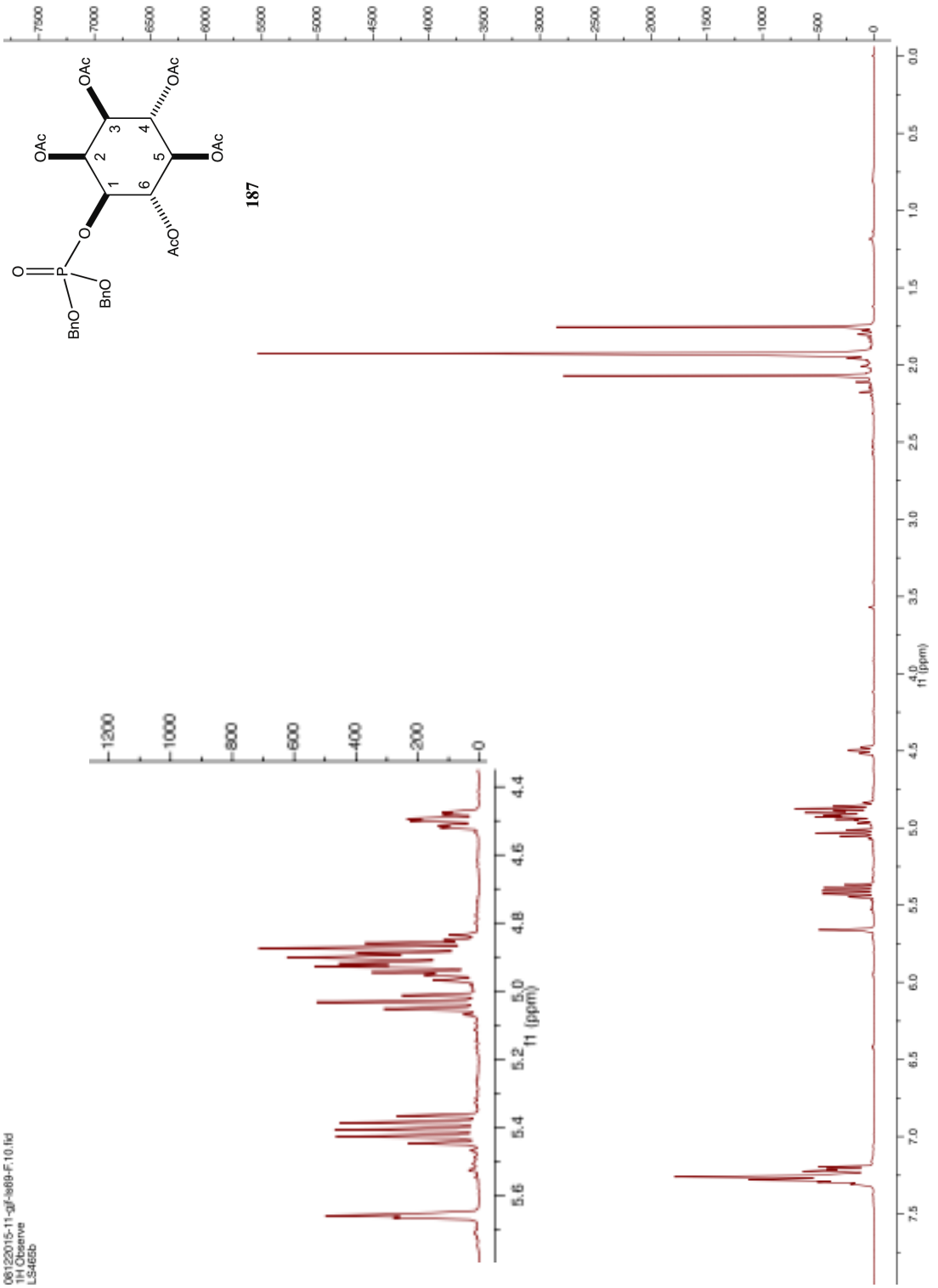
1-Bis(benzyloxy)phosphoryl L-*myo*-inositol (116) ^{1}H NMR



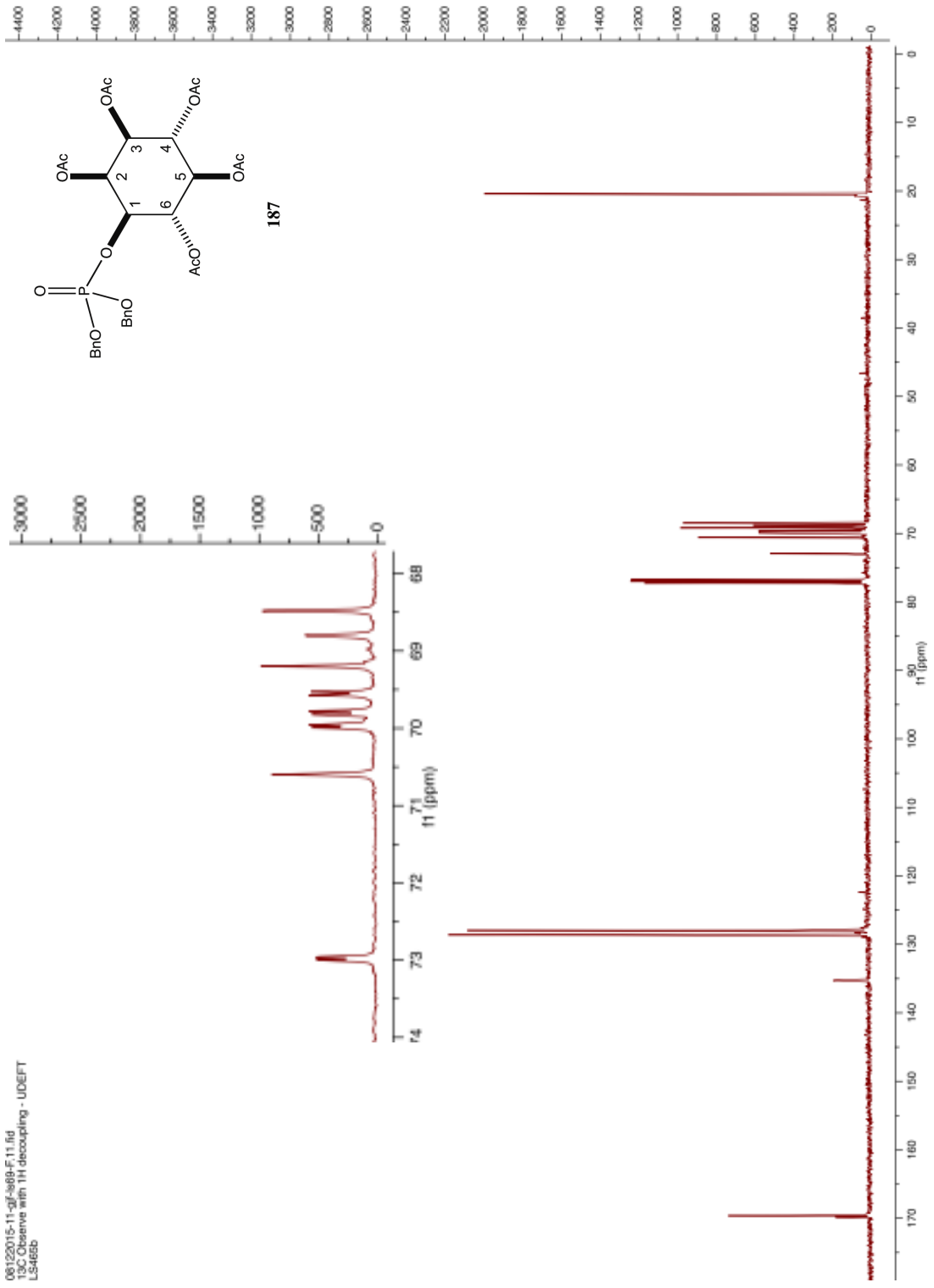
1-Bis(benzyloxy)phosphoryl L-*myo*-inositol (116) ^{13}C NMR



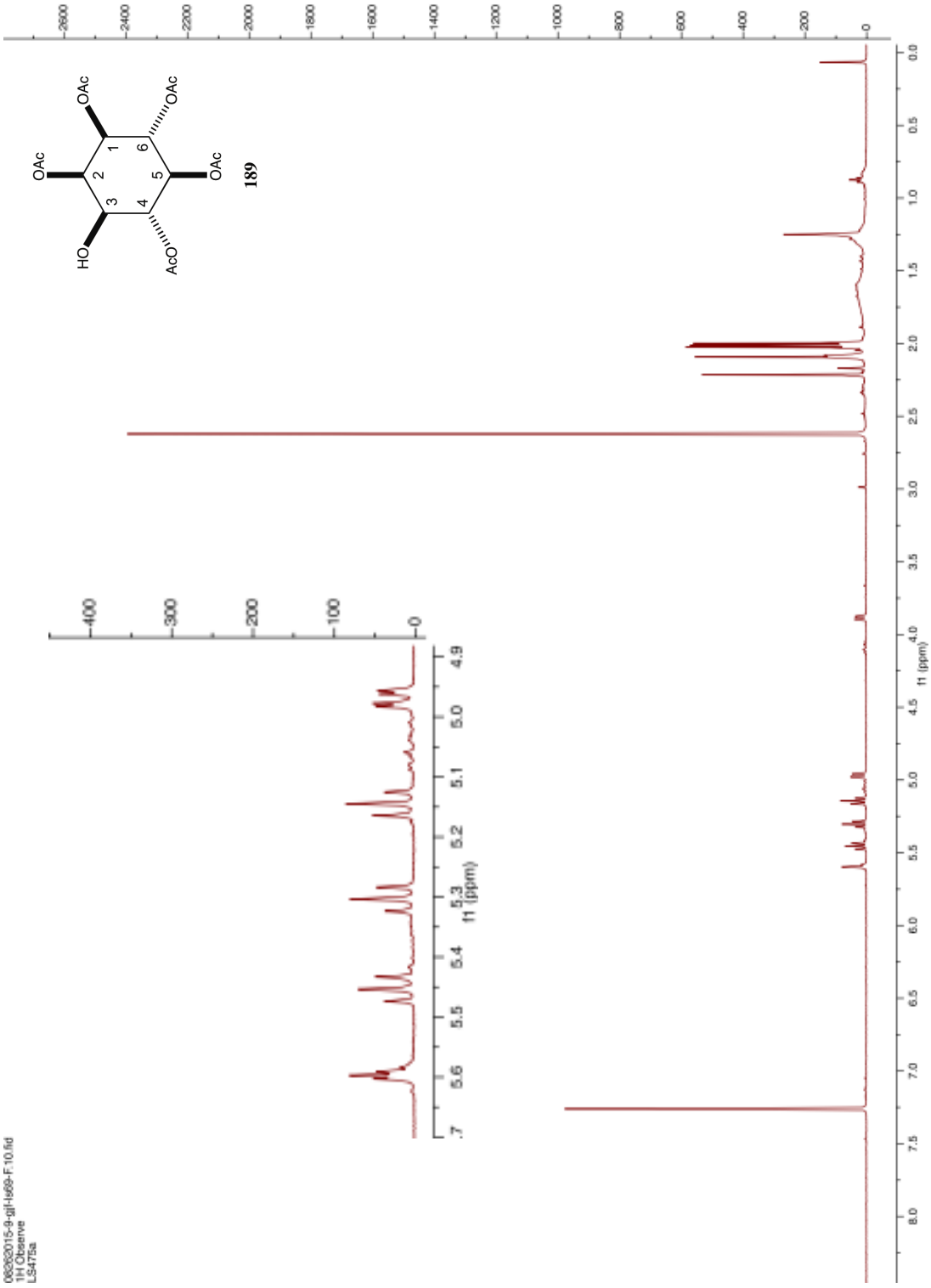
2,3,4,5,6-Penta-*O*-acetyl 1-bis(benzyloxy)phosphoryl L-*myo*-inositol (187) ^{1}H NMR



2,3,4,5,6-Penta-*O*-acetyl 1-bis(benzyloxy)phosphoryl L-*myo*-inositol (187) ^{13}C NMR



1,2,4,5,6-penta-*O*-acetyl D-myo-inositol (189) ^{1}H NMR



1,2,4,5,6-penta-*O*-acetyl D-myo-inositol (189) ^{13}C NMR