Unravelling the Biological Roles of Charged Carbohydrates

A thesis submitted in partial fulfilment of the degree of
Doctor of Philosophy at the University of Oxford

Mitul K. Patel

Trinity 2011

Unravelling the Biological Roles of Charged Carbohydrates

Mitul K. Patel

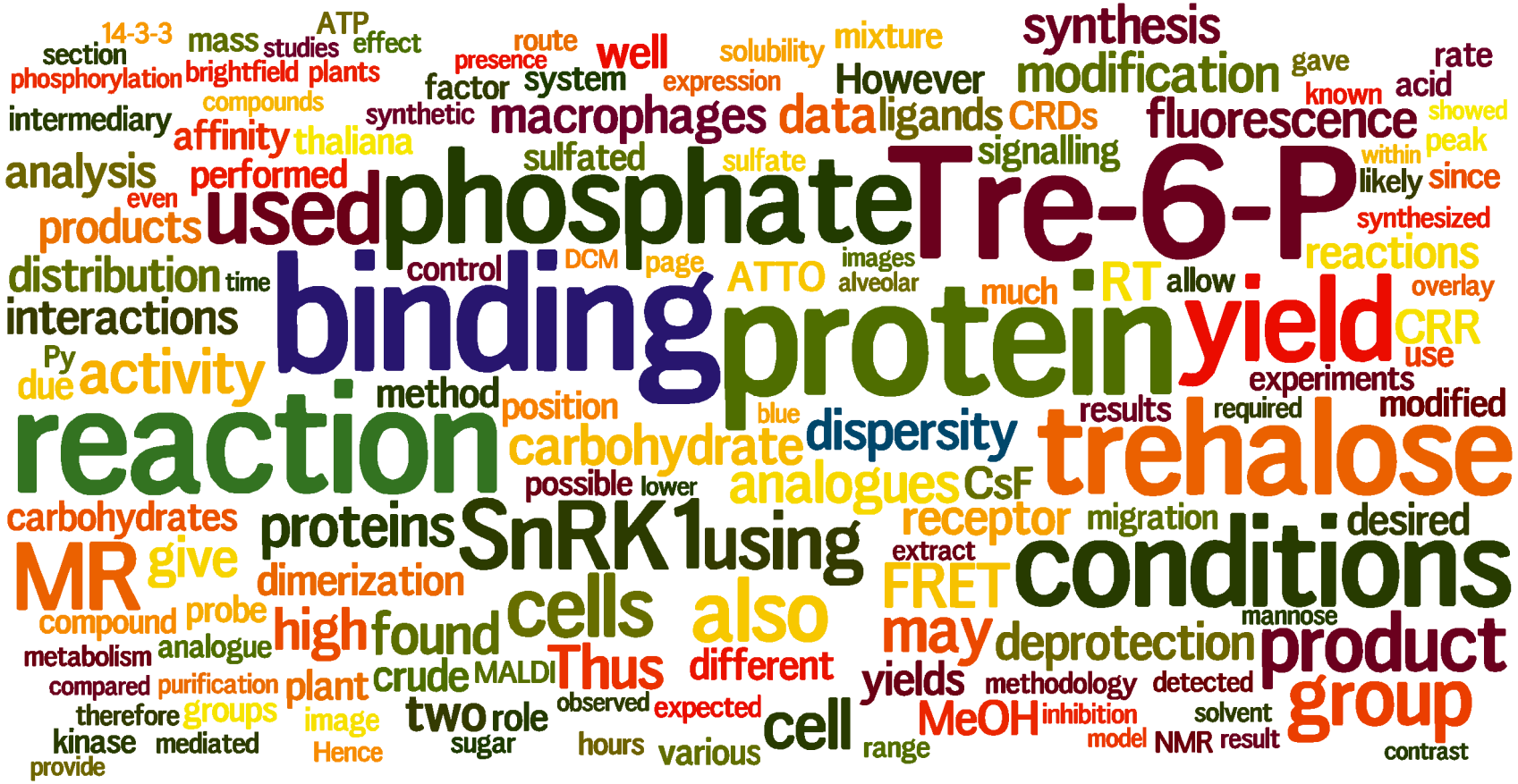
Magdalen College
University of Oxford
Submitted for the degree of Doctor of Philosophy in Organic Chemistry
Trinity 2011

The phosphate and sulfate esters of carbohydrates are widespread in nature, where they partake in a variety of signalling roles. This thesis focuses on the involvement of trehalose-6-phosphate (Tre-6-P) in regulating plant growth, and the function of sulfated carbohydrates in the immune response.

In plants, Tre-6-P is a critical messenger that conveys environmental information for the control of metabolic state. Beneficial natural conditions lead to an increase in the *in vivo* concentration of Tre-6-P, which in turn triggers biosynthetic processes facilitating plant growth. However, the mechanism by which this effect is exerted is largely unknown. A thorough understanding of this fundamental signalling pathway could allow considerable improvements in agriculture. Recent evidence has found an interaction between Tre-6-P and a protein kinase system (SnRK1), mediated via an intermediary protein whose identity is unknown. Our efforts to identify this elusive protein are presented. Methodology is developed to allow the synthesis of Tre-6-P analogues for defining structure-activity relationships, with particular attention focused towards the desymmetrization of trehalose and the control of phosphate migrations. Tre-6-P is tethered to a solid phase resin allowing affinity extraction of putative intermediary factors. The 14-3-3 class of regulatory proteins is identified as a potential binding partner of Tre-6-P.

The mannose receptor (MR) is an essential element of the mammalian immune system. Expressed on macrophages, the protein can bind and internalize pathogens through their surface exposed carbohydrates. However, the receptor can also dimerize, upon which it is able to divalently bind to sulfated carbohydrates. The biological significance of this sulfate binding is poorly understood. Current hypotheses invoke a role for the sulfate mediated trafficking of MR dimers to areas of B-cell maturation for antigen presentation, thus linking the innate and adaptive defence mechanisms. In any case, the dimerization of the MR is an essential prerequisite for interaction with carbohydrate sulfates. Methodology is required to probe this dimerization on the surface of cells under their native conditions. In this thesis, we use fluorescence microscopy to study the binding of various probes to macrophage and endothelial liver derived cell lines. Results of our initial investigations are presented that show internalization of the probes. The use of FRET imaging to directly probe the presence of MR dimers is studied.

For My Parents and Grandparents....


14-3-3
ATP
mass
studies
effect
route
well
solubility
mixture
synthesis
section
presence
phosphorylation
brightfield
plants
factor
system
expression
However
modification
gave
rate
known
acid
compounds
intermediary
synthetic
macrophages
data
ligands
CRDs
fluorescence
showed
affinity
thaliana
peak
analysis
sulfated
sulfate
signalling
within
performed
Tre-6-P
likely
since
even
used
phosphate
products
synthesized
reactions
distribution
time
control
DCM
page
ATTO
images
alveolar
much
RT
allow
overlay
binding
protein
yield
interactions
CRR
Py
use
due
activity
experiments
method
position
blue
results
required
modified
reaction
carbohydrate
dispersity
trehalose
possible
lower
analogues
CsF
carbohydrates
receptor
migration
desired
proteins
SnRK1
using
extract
MR
give
dimerization
FRET
conditions
mannose
compound
probe
high
found
cells
also
may
deprotection
product
metabolism
analogue
Thus
different
yields
methodology
detected
plant
crude
MALDI
compared
purification
MeOH
inhibition
solvent
group
therefore
groups
image
two
role
observed
expected
cell
kinase
mediated
sugar
hours
various
range
model
NMR
result
contrast
provide
Hence

Acknowledgements

I've had a great time over the last five years in the BGD group and there are a lot of people I'd like to thank for making it so. Thank you to Ben for letting me into the group and for the stream of ideas (sometimes crazy ideas – sorry I never quite got around to growing a tree in an MRI machine). Most of all, I appreciate the freedom I have been given to do whatever I have wanted despite some of my more stupid ideas ending in spectacular failure (the reader is directed to the next 300 pages).

A very big thank you to my collaborators Matthew and Lucia at Rothamsted Research. It has been great working with you and this thesis would not have been the same without your input. I'd also like to thank the IAVI for funding over the last 4 years and for making my DPhil possible.

A big thanks to my proofreaders – Justin, Conor, Sgt. Matt, Smita and Chris for correcting all of my typos and making this thesis at least semi-readable.

Lastly, and most importantly, I want to thank all the members of the BGD group past and present. The best part of the DPhil has been working with you guys. I have learnt a whole range of different things from people in the group and I have appreciated your help over the last five years. There are way too many of you to thank individually so I'm not going to make a massive list. Thank you everyone! Oh and one last thing…all the strange mysterious things that happened in our group…it wasn't me.

Abbreviations

The following are a list of important or unusual abbreviations used in this thesis. All other abbreviations follow the standard conventions.

A(M/D/T)P	Adenosine mono/di/tri phosphate
BSA	Bovine serum albumin
CD	Circular dichroism
CRD	Carbohydrate recognition domain
CRM	Cross reacting material fragment of diphtheria toxin
CRR	Cysteine rich region
DIAD	Diisopropyl azodicarboxylate
DPCP	Diphenyl chlorophosphate
DTT	Dithiothreitol
E_A	Apparent efficiency
EDAC	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide)
ELS	Evaporative light scattering
FWHM	Full width at half maximum
Glc-1-P	Glucose-1-phosphate
Glc-6-P	Glucose-6-phosphate
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HPLC	High performance liquid chromatography
IAVI	International AIDS Vaccine Initiative
ITC	Isothermal titration calorimetry
LC-MS	Liquid chromatography mass spectrometry
LUT	Look-up table
MALDI	Matrix-assisted laser desorption/ionization
MOWSE	Molecular weight search
MR	Mannose receptor
MS	Mass spectrometry
nOe	nuclear Overhauser effect
PAPS	3'-Phosphoadenosine-5'-phosphosulfate
PBS	Phosphate buffered saline
PRK	Phophoribulosekinase
RGB	Red-green-blue
Rib-5-P	Ribose-5-phosphate
SAX	Strong anion exchange
SIR	Single ion recording
STD	Saturation transfer difference
TIR	Total internal reflectance
TOF	Time of flight
TPS	Trehalose phosphate synthase
TPP	Trehalose phosphate phosphatase
Tre-6-P	Trehalose-6-phosphate

Contents

Acknowledgements iii
Abbreviations iv
Chapter I - Introduction 1
1.1 Carbohydrate phosphate and sulfate esters 1
1.2 Trehalose-6-phosphate signalling in plants 2
1.2.1 The trehalose pathway 3
1.2.2 Communication with kinase cascades 8
1.2.3 Research objectives 11
1.3 The mannose receptor 11
1.3.1 Carbohydrate recognition and innate defence 12
1.3.2 Sulfated carbohydrate recognition 14
1.3.3 Carbohydrate sulfates in adaptive immunity 17
1.3.4 Research objectives 19
1.4 Thesis layout 20
1.5 References 21
Chapter II - The desymmetrization of trehalose 33
2.1 Introduction 33
2.2 Initial explorations into mono-functionalization reactions 35
2.3 Kinetics of consecutive reactions 36
2.4 Desymmetrizations under continuous flow 37
2.5 A reverse modification strategy 40
2.6 Synthesis of trehalose-6-phosphate 43
2.7 Conclusions 45
2.8 Experimental 46
2.8.1 General considerations 46
2.8.2 Batch trehalose mono-functionalizations 47
2.8.3 Flow chemistry and kinetic analysis 51
2.8.4 Optimized synthesis of trehalose-6-phosphate 59
2.9 References 62
Chapter III - Phosphate group manipulations 65
3.1 Introduction 65
3.2 Nucleophilic substitutions of the diphenylphosphate group 67
3.3 Diphenylphosphate deprotections 72
3.4 Phosphate transesterfication 74
3.5 Control of phosphate migrations with CsF 76

3.6 Extending the scope of transesterfications 81
3.7 Synthetic route for 6' modification of trehalose-6-phosphate 83
3.8 Conclusions 85
3.9 Experimental 85
3.9.1 Initial nucleophilic substitutions 86
3.9.2 Nucleophilic substitutions on the glucose model 91
3.9.3 Optimized phosphate S_N2 route 97
3.9.4 Diphenylphosphate deprotections by hydrogenation 101
3.9.5 Diphenylphosphate deprotections by basic hydrolysis 103
3.9.6 Probing the mechanism of CsF transesterfications 108
3.9.7 Installation of diphenylphosphate functionality 113
3.9.8 CsF transesterfication reactions 116
3.9.9 Dimethylphosphate deprotections 121
3.9.10 Synthetic rout for 6' modifications 123
3.10 References 128
Chapter IV – Investigations on the inhibition of SnRK1 133
4.1 Introduction 133
4.2 Metabolism of exogenous sugars *in vitro* 136
4.3 Metabolism of ribose-5-phosphate 140
4.4 Library of analogues modified at the phosphate centre 146
4.5 Library of analogues modified at the C6' centre 148
4.6 Building a Tre-6-P affinity column 153
4.7 Identification of Tre-6-P binding proteins 157
4.8 Probing the binding of 14-3-3s with Tre-6-P 160
4.9 Tre-6-P microarrays 167
4.10 Conclusions 168
4.11 Experimental 169
4.11.1 Metabolism experiments on exogenous sugars 169
4.11.2 Tre-6-P library synthesis 172
4.11.3 Synthesis of affinity columns 185
4.11.4 Proteomics 197
4.11.5 Studies on recombinant 14-3-3 197
4.11.6 Microarray compounds 201
4.12 References 208
Chapter V - Sulfate binding to the mannose receptor 213
5.1 Introduction 213

5.2 Synthesis of the mannose probe 216
5.3 Mannosylated quantum dots 217
5.4 Fluorescent mannosylated proteins 221
5.5 Attempted synthesis of an *N*-acetyl-galactose-4-sulfate probe 227
5.6 Synthesis of a generic sulfate probe 229
5.7 Fluorescent sulfonated proteins 230
5.8 Probing receptor dimerization with FRET microscopy 236
5.9 Endothelial liver derived cell line 243
5.10 Conclusions 250
5.11 Experimental 251
5.11.1 Synthesis of the mannose ligand 251
5.11.2 Attempted synthesis of GalNAc-4-SO_4 ligand 256
5.11.3 Synthesis of the sulfate ligand 270
5.11.4 Modification of quantum dots 271
5.11.5 Modification of BSA 272
5.11.6 Modification of ATTO dyes 274
5.11.7 Growth of cells 274
5.11.8 Confocal microscopy 276
5.11.9 FRET microscopy 277
5.12 References 277
Chapter VI – Characterization of synthetic glycoproteins 281
6.1 Introduction 281
6.2 The Gaussian approximation 282
6.3 Analysis of MALDI peak broadening 284
6.4 Validation of the methodology 286
6.5 Dispersity analysis on synthetic glycoproteins 289
6.6 Conclusions 294
6.7 Experimental 295
6.7.1 Synthetic protocols 295
6.7.2 Protein sequences 298
6.7.3 Protein modification 299
6.7.4 MALDI mass spectrometric analysis 300
6.7.5 MATLAB M-file scripts 300
6.8 References 301
Chapter VII – Conclusions and Future Directions 305

Chapter I - Introduction

1.1 Carbohydrate phosphate and sulfate esters

The importance of phosphate esters in nature is indisputable.[1] For instance, the unique properties of the P–O bond has allowed adenosine triphosphate (ATP) to adopt a central role in the transport and storage of cellular energy across the natural world. Due to the very high kinetic stability of phosphodiester linkages,[1,2] this functionality is commonly used as a structural component, for example in the backbone of DNA, or as a part of phospholipids within bilayer membranes. Moreover, under physiological conditions, the phosphate moiety exists predominantly as an anion which allows strong ionic bonding interactions that are ideally suited for signalling purposes, such as with inositol phosphates.[3] Protein phosphorylation is a particularly widespread example of phosphate mediated regulation, where the strong ionic interactions are exploited to induce conformational changes that modulate the underlying activity of the protein.[4]

The phosphorylation of carbohydrates is particularly prevalent.[5] The phosphate moiety is critical for the biochemistry of sugars, where it acts as a biological "activating agent" facilitating various metabolic transformations.[6] Apart from their role as reactive intermediates, phosphorylated carbohydrates are often potent signalling molecules in their own right. Thus, carbohydrate phosphate esters are involved in enzyme trafficking,[7] lymphocyte homing,[8] protein translocation[9] and in the immune[10] and inflammatory responses.[11]

Compared to the prevalence of the phosphates, the sulfate functionality is much rarer. For the biosynthesis of such sulfated molecules, 3'-phosphoadenosine-5'-phosphosulfate (PAPS) is needed as a metabolic precursor. PAPS is biochemically synthesized from ATP at relatively low concentrations within the cell, which allows for a controlled deployment of the sulfate group resulting in more specialized biological applications compared to phosphorylated adducts.[12] Thus the use of PAPS represents a deliberate barrier by which the phosphorylation and sulfation systems can be separated.[1,12] Whereas the phosphate group is used for a wide range of roles that are ubiquitous across nature, the sulfate group is utilized in a much more specialized way. For instance, compared to the purely structural role of phospholipids, sulfolipids

such as sulfated cholesterols frequently have additional signalling roles.[13] As an additional example, protein sulfation, in contrast to ubiquitous protein phosphorylation, is a far rarer modification that is used exclusively in an extracellular environment for secreted or membrane proteins, and is often employed very specifically to modulate protein-protein interactions.[14]

Reflecting these specialist roles for sulfate esters, carbohydrate sulfates are almost exclusively synthesized for tailored signalling roles.[15] Examples of such applications include nodulation factors that initiate nodule formation on legumes, heparin sulfate chains that are involved in a variety of cell regulatory processes, and keratin sulfates which are implicated in the regulation of corneal collagen fibrils.[15] However, there exist a plethora of other sulfated carbohydrates whose biological functionality is yet to be elucidated and it remains likely that this class of sugars will have a much wider role than is currently appreciated.

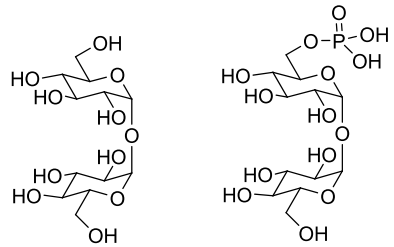
In this thesis, two different carbohydrate signalling systems are explored. Initially, the role of trehalose-6-phosphate (Tre-6-P) in plant development was investigated (section 1.2, page 2). Next the involvement of sulfated carbohydrates in the mammalian immune response was explored (section 1.3, page 11).

1.2 Trehalose-6-phosphate signalling in plants

The control of metabolism and growth is particularly challenging for the plant kingdom. As sessile life forms, they are at the whims of their immediate surroundings and must adapt to utilize the available resources. Not surprisingly, plants have developed elaborate signalling networks that allow versatile responses to changes in the natural environment.[16] Phosphorylation cascades mediated by phosphotransferases (more colloquially known as kinases) form the bulk of these signalling pathways.[17] Compared to the 518 protein kinase genes present in humans,[18] *Arabidopsis thaliana* has over 1000 reflecting a much more developed signalling network.[19]

The biochemical transformations of carbohydrates within a plant are particularly interesting. As photosynthetic organisms, plants have two opposing directions of carbohydrate flux which must be adequately regulated. These carbohydrates are not only sources of energy, but also act as sensors that convey information on the

metabolic state of the plant.[20] Over the past decade, attention has focused on the role of the disaccharide trehalose **1.1** and its biosynthetic precursor trehalose-6-phosphate **1.2** (Tre-6-P) in plant signalling.[21] There is now a significant body of evidence that implicates Tre-6-P as a central signalling molecule which regulates overall plant metabolism and controls critical parameters such as growth rates and crop biomass.[21,22] A thorough understanding of this signalling pathway could lead to considerable benefits in agriculture and further research into this field is therefore essential.



1.1 **1.2**

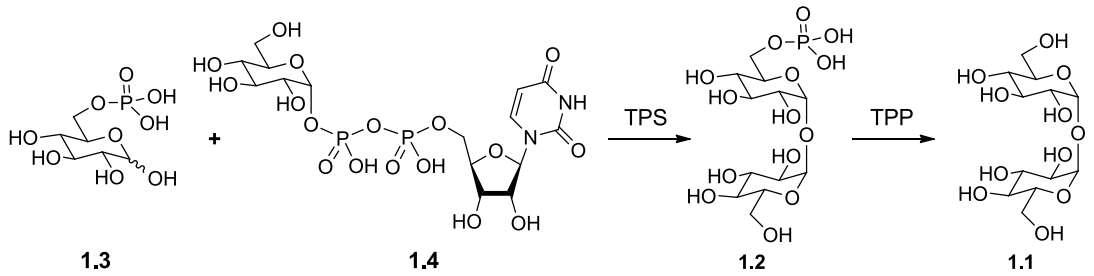
Figure 1.1. *Chemical structures of trehalose and Tre-6-P.*

1.2.1 The trehalose pathway

Trehalose is naturally synthesized in all invertebrates fulfilling a variety of functions depending on the species.[21] The unique α,α glycosidic linkage allows trehalose to bind to the head groups of lipid bilayers with trapped water thus acting as a stabilizing agent to protect against extreme stresses such as desiccation or high osmolarity, which can help preserve cell integrity.[23] Thus, trehalose is synthesized in bacteria as a storage sugar to protect against such stresses, but is also used as the sole carbon source during normal conditions.[21,24] Trehalose is the major haemolymph (blood) sugar in insects, where it is present at high concentrations as an efficient source of energy to power flight.[25] At such high concentrations, glucose is very toxic due to the reducing power of the free anomeric centre and therefore cannot be used.[25]

The role of trehalose in plants is far less clear. A minority of desert dwelling species known as resurrection plants that are under the constant risk of desiccation produce significant quantities of trehalose as a protectant. For instance, *Selaginella lepidophylla* can contain trehalose up to 10% of the plant's total weight.[26] However, for the majority of plants, trehalose is not accumulated to any significant extent. As a carbon source, the sugar is completely redundant since this role is performed by

sucrose. Organisms for which trehalose is essential posses multiple biosynthetic routes for its synthesis as a contingency against the loss of an external supply of metabolic precursors.[27] In contrast, plants contain only one biosynthetic path which would seem to imply that trehalose is non-essential. This view has largely been upheld, and as a result, the sugar was considered an obscurity and was mostly ignored among the plant sciences community. More recently, with the sequencing of the *A. thaliana* genome and modern transgenic experiments, trehalose has once again been thrust into prominence. Now however, attention has shifted to its metabolic precursor Tre-6-P as a potent signalling molecule.



Scheme 1.1. *The trehalose biosynthetic pathway in plants.*

The Tre-6-P pathway, which is the most prevalent of the five known biosynthetic routes for trehalose,[27] involves the glycosylation of glucose-6-phosphate **1.3** with UDP-glucose **1.4** using a trehalose phosphate synthase (TPS) to produce Tre-6-P, which can then be dephosphorylated with a trehalose phosphate phosphatase (TPP) to produce trehalose as the end-product. Interestingly, despite the low accumulation of trehalose, there are a plethora of putative genes for this pathway in *A. thaliana*, with 11 TPSs and 10 TPPs having been identified.[28] This is in stark contrast to the 4 analogous sucrose phosphate synthases and 4 sucrose phosphate phosphatases described in the same genome.[29] Furthermore, for the breakdown of trehalose, there exists only one trehalase enzyme.[21] Given that trehalose serves no apparent function, this wide range of genes encoding its synthesis is somewhat surprising and suggests a signalling role for this pathway. Consistent with this hypothesis, genetic studies using microarrays found that the expression of the TPS and TPP genes is highly dependent on a range of environmental factors, such as the availability of light, nutrients and essential ions.[22,30] Furthermore, compared to yeast protein sequences, the active *A. thaliana* TPS1 protein has an *N*-terminal extension that limits its synthase activity which reduces the build up of trehalose and Tre-6-P, suggestive of an evolutionary change to a signalling responsibility for these carbohydrates.[22,31]

Interestingly, a detailed analysis of the *A. thaliana* genome indicates that there is a large amount of proliferation amongst the genes of the trehalose pathway.[32] In fact, of the 11 TPS genes, only one has retained synthase activity (AtTPS1).[27] The exact function of the inactive TPS proteins is unclear, but current evidence points to a regulatory function.[30] Gene expression studies have shown that these inactive TPS proteins are differentially expressed in growing tissues, although their exact mechanism of action is yet to be elucidated.[33]

The earliest studies on live plants focused on the addition of exogenous trehalose to the growth medium, where the sugar was found to have a toxic effect that stunted growth.[34] More detailed studies demonstrated that trehalose reduced root growth and led to starch accumulation in cotyledons (a part of the embryo that would ultimately become the first leaves after germination).[35] While these results indicated that the trehalose pathway had some bearing on overall plant growth and carbohydrate metabolism, their relevance was not fully appreciated at the time. A more complete picture on the role of the trehalose pathway has emerged from the creation of genetically modified plants which have allowed for controlled changes in the *in vivo* concentration of Tre-6-P.[36] This has been achieved by either the generation of AtTPS1 knockout mutants, or the incorporation of the *Escherichia coli* genes for TPS and TPP (OtsA and OtsB respectively) to create transgenic plants. Goddijn *et al.* engineered OtsA and OtsB into *Nicotiana tobaccum*, but this failed to induce the accumulation of trehalose which provided the first evidence to associate trehalose as part of a well-regulated signalling network.[37]

Most importantly, such transgenic experiments have been able to conclusively demonstrate that rather than trehalose, it is in fact Tre-6-P that is the active signalling molecule in plants.[38] Thus, the phenotypes observed from the feeding of trehalose are in fact due to the perturbation of the equilibrium with TPP, whose phosphorylase activity is restricted resulting in an accumulation of Tre-6-P.[34,35,39] *A. thaliana* mutants containing the OtsA gene also have a greater Tre-6-P concentration than the wild-type and provide a much more controlled method for modifying the *in vivo* concentration of Tre-6-P. These OtsA plants are found to have small, dark green leaves while those of the OtsB mutants, which have a considerably reduced concentration of Tre-6-P, were larger and paler.[38] Crucially, a control mutant expressing an *E. coli* trehalase which depletes the concentration of trehalose but

leaves Tre-6-P unaffected had an identical phenotype to the wild-type plant, thus proving that Tre-6-P rather than trehalose is the vital regulatory element.[38] Furthermore, the feeding of exogenous carbohydrates (specifically glucose, fructose and sucrose) to OtsB mutants was found to have a toxic effect. These plants exhibited a severe reduction in growth rate and a build up of unmetabolized carbohydrate intermediates clearly illustrating that Tre-6-P is an indispensable signal for carbohydrate utilization and is vital for normal plant development. In sharp contrast, the supply of sugars to OtsA mutants led to a significant enhancement in the rate of growth compare to the wild-type and illustrates an improved utilization of carbohydrates.[38] Such a feat has never previously been accomplished through either the genetic engineering or selective breeding of crops and thus, emphatically demonstrates the potential importance of this pathway to agriculture.

Moreover, the biosynthesis of Tre-6-P correlates exactly with the supply of exogenous sucrose and increases almost 30 fold within 3 hours of feeding, suggesting that Tre-6-P is a messenger of overall carbon availability which triggers downstream processes for carbohydrate metabolism.[39,40] In agreement with this hypothesis, growth for *A. thaliana* AtTPS1 knockout mutants is arrested at the torpedo stage of developments due to sucrose accumulation.[41] This stage of plant embryonic growth, just prior to root and cotyledon development, is associated with an increase in cell biosynthetic activities which are compromised by the lack of an adequate Tre-6-P signal. Histochemical staining experiments have shown that the AtTPS1 knockout mutant affects the pot-translational modification of a vast range of important developmental enzymes[42] leading to lower cell division, altered cellular architecture and ultimately death.[43] Not surprisingly, Tre-6-P is also essential in other high growth situations such as during flowering, where AtTPS1 in constitutively expressed in buds.[44] Furthermore, these signalling effects are not exclusive to *A. thaliana* and similar studies on transgenic potatoes have shown a Tre-6-P mediated effect on tuber growth.[45] Given the ubiquity of the Tre-6-P genes, this signalling pathway is expected to persist across the entire plant kingdom.[20-22]

While plants engineered for reduced Tre-6-P assimilation suffer from growth defects, increased concentrations of the sugar can potentially bestow a range of desirable characteristics on such plants. For example, an OtsA mutant of *N. tobaccum* was found to have a greatly improved photosynthetic capacity due to an increase in the

activity of Rubisco (a vital enzyme in the Calvin cycle for the fixation of CO_2 during photosynthesis), which allows the plant's energy requirements to be serviced by a much smaller leaf area thus explaining previously observed phenotypes where OtsA mutants had smaller and greener leaves than the wild-type.[46] Small leaves are potentially very useful for reducing water loss through transpiration, which has implications for agriculture in drought-affected areas. Alternatively, when coupled with the improvement of growth rate observed for OtsA mutants on feeding exogenous sugars,[38] such changes have the potential to vastly improve crop yields. As well as the effect on plant growth and metabolism, Tre-6-P has a much wider range of effects including the modulation of abscisic acid (a plant hormone controlling many developmental processes)[47] and enhancing tolerance to stresses such as drought, for which OtsA mutants have been particularly useful.[48] Interestingly, there is also an example of a naturally occurring TPP knockout mutant in maize which, as a result of higher Tre-6-P levels, has a higher degree of branching at the meristem resulting in higher yields illustrating the possible evolutionary advantage of such mutations.[49]

There is clearly a vast amount of evidence demonstrating the role played by Tre-6-P as a central signalling molecule, but the mechanism of its action is at present not known. For yeast, Tre-6-P has a similar function that is effected through the inhibition of hexokinase.[50] However, plant hexokinase is not inhibited by Tre-6-P,[51] and therefore constitutes a separate, independent signalling system.[52]

One mode of plant control occurs through Tre-6-P mediated regulation of the enzyme ADP-glucose pyrophosphorylase (AGPase).[53] AGPase is a vital enzyme for starch synthesis[54] and its *in vivo* activity is controlled by post-translational redox activation to give a more active, disulfide-bond tethered, dimeric form of the protein,[55] which is found to increase in *A. thaliana* OtsA mutants.[53] Thus, increased levels of *in planta* sucrose induce starch production via a Tre-6-P mediated AGPase activation signal.[40,53] While this hypothesis provides a reasonable explanation for the Tre-6-P control of carbohydrate utilization, it fails to account for the vast range of other factors that the sugar seems to influence.[20-22] Additionally, the precise mechanism by which AGPase activity is controlled is also lacking.

Now however, recent research has shown that Tre-6-P signalling is intimately involved with a protein kinase cascade that controls plant development at the most fundamental level.[56]

1.2.2 Communication with kinase cascades

The regulation of energy supply for all eukaryotes is performed by a class of ubiquitous calcium independent serine/threonine protein kinases.[17,57] These kinases can modulate the activity of a vast array enzymes and transcription factors through protein phosphorylation and can control various downstream developmental processes. In response to an energy deficit, these kinases are activated and are able to repress biosynthetic activities while simultaneously triggering catabolic pathways that provide alternative sources of nutrients, thus diverting resources to essential processes for survival. In contrast, during an energy surplus, the kinase activity is repressed and biosynthesis is allowed to proceed, facilitating growth and development.

The plant analogue, [sucrose non-fermenting 1] related protein kinase 1, more succinctly known as SnRK1, is responsible for the control of carbohydrate and nitrate metabolism,[17,57,58] as well as being involved in the abscisic acid hormone[59,60] and stress pathways.[60] Several genetics experiments have been used to demonstrate the broad role played by this central kinase. Incredibly, microarray based experiments on SnRK1 over-expressing *A. thaliana* mesophyll cells have demonstrated that the kinase alters the expression of over a thousand genes.[61] Under energy starvation conditions, 300 genes mostly representing protein and ribosome synthesis as well as other biosynthetic pathways were repressed, while 300 catabolic genes for amino acid, cell wall and lipid degradation were activated. In contrast, the feeding of exogenous sugars produced the exact opposite expression profile. Additional experiments have identified 165 genes that encode various transcriptional factors and redox regulators which are responsible for the transition from catabolic to anabolic metabolism in response to nutrient availability.[62] This change is incredibly sensitive and is initiated by even small changes in sucrose concentration.[63]

Antisense RNA experiments to silence the expression of SnRK1 have proved especially useful for studying the effect of the kinase and have conclusively demonstrated an important role in the control of carbohydrate flux. Thus, antisense

SnRK1 potatoes have a reduced activity of sucrose synthase,[64] and a lower amylase activity preventing starch breakdown.[65] Ordinarily, active SnRK1 would activate catabolic processes to provide alternative sources of sucrose, but when muted by an antisense gene, such processes are inhibited. Not surprisingly, the loss of adequate energy management results in several developmental defects in various antisense species including peas which showed maturation problems and poor cotyledon growth,[66] barley which produced deficient pollen,[67] and potatoes which exhibited poor root growth and tuber development.[68] A complete SnRK1 knockout mutant of the moss *Physcomitrella patens* showed considerable developmental defects.[69] Most interestingly, the SnRK1 deficient moss was only able to grow under constant illumination and rapidly died if exposed to a normal day-light cycle, indicative of the role played by the kinase in the starvation response. As well as responding to changes in carbohydrate concentrations, it appears that SnRK1 can also actively relocate available resources throughout the plant. For instance, under herbivore attack, SnRK1 is responsible for the reallocation of carbon to the roots as part of the plants natural defence mechanism.[70] Moreover, there is evidence to suggest that SnRK1 has a wider role in plant defence against pathogenic attack.[71]

There is significant divergence in the characteristics of plant SnRK1 compared to the other analogous kinases found in eukaryotes.[57,58] The mammalian form of the protein, AMP-activated protein kinase (AMPK), acts as a sensor for the ATP:AMP ratio and is directly activated by AMP under conditions of stress such as hypoxia or exercise.[72] Similarly, the yeast analogue, sucrose non-fermenting 1 (SNF1), responds to glucose availability and under certain conditions, also seems to respond to AMP concentrations.[73,74] In contrast, SnRK1 does not directly interact with AMP and its mechanism of sensing the energy deficit is very different to both AMPK and SNF1.[75]

This difference is likely attributed to the structural differences between these proteins. As with AMPK and SNF1, SnRK1 is likely to exist as a trimeric complex.[76] Within this complex, the α subunit is responsible for kinase activity and as a result, this domain is well conserved across the family.[57,77] For SnRK1, there are two possible forms for this α domain known as KIN10 and KIN11, with both exhibiting kinase activity.[57] It is worth noting that there also exist far more divergent forms of this subunit giving rise to protein complexes known as SnRK2 and SnRK3, but these lack any kinase activity and are unable to complement a yeast SNF1 knockout and are

therefore likely to have alternative roles in plant signalling.[17,78] In contrast to the α domain, the β and γ subunits are far more diverse for SnRK1 than for AMPK and SNF1.[57,79] This difference is particularly important since the AMP binding site is on the γ domain of AMPK, and seems to have been evolved out of plant SnRK1.[57,80] While AMP prevents dephosphorylation of a critical "activation loop" on SnRK1 (which is itself phosphorylated by the upstream kinases GRIK1 and GRIK2 as part of a kinase signalling cascade),[81] it is not directly involved in communicating the plant energy status.[82] Therefore, an alternative sensing mechanism must exist.

There are clear parallels between the effect of SnRK1 on plant development to that of Tre-6-P, which would seem to imply some level of cross-talk between the two signalling networks. Such rationale is supported by the fact that of the numerous genes controlled by SnRK1, several are involved in the trehalose biosynthetic pathway.[58,62] These associations are maintained when studying OtsA and OtsB mutants of *A. thaliana*, where the SnRK1 (KIN11) transcripts levels correlate with the *in vivo* concentration of Tre-6-P,[39] and the activation of AGPase by Tre-6-P occurs in an SnRK1 dependent manner.[83]

Pioneering work by Zhang *et al.* conclusively demonstrated this interplay between the two systems by showing that Tre-6-P strongly inhibits SnRK1 at micromolar concentrations.[56] A microarray analysis on *A. thaliana* OtsA mutants showed the opposite gene expression profile to SnRK1 over-expression, as would be expected with such inhibition, with a general up-regulation of biosynthetic genes and repression of those encoding degradatory pathways. Most effected by this Tre-6-P inhibition were genes responsible for photosynthesis, mitochondrial electron transport, amino acid metabolism, nucleotide metabolism and the citric acid cycle. In addition to Tre-6-P, glucose-6-phosphate (Glc-6-P)[84] and glucose-1-phosphate (Glc-1-P)[85] also inhibit SnRK1, but much less potently requiring millimolar concentrations of the sugar to obtain the same effect as Tre-6-P, reflecting their naturally higher concentration *in vivo*. In this model, environmental factors such as light and the availability of nutrients trigger the production of Tre-6-P (and to lesser extent Glc-6-P and Glc-1-P), which inhibits the normal activity of SnRK1, allowing the utilization of resources for biosynthesis and growth. Thus, the inhibition of SnRK1 by Tre-6-P represents a pivotal signalling system that controls the overall plant development which clearly warrants further investigations.

1.2.3 Research objectives

The mechanism of binding between Tre-6-P and SnRK1 is poorly understood. Preliminary data has shown that the kinase binds Tre-6-P at a site distinct from ATP and is non-competitive.[56] In fact, based on the computational fitting of various binding models to dose response curves, the interaction with SnRK1 is believed to have been mediated by an intermediary protein which simultaneously binds Tre-6-P and the protein kinase to reversibly form a ternary complex.[56] The involvement of an intermediary factor is not entirely surprising since SnRK1 is known to interact with other proteins in a similar manner.[86] However, the identity of this protein has remained elusive. To fully understand this fundamental signalling system, and for the eventual exploitation of this network for agricultural gains, it is essential to determine the exact mechanism of action for kinase inhibition. The ultimate goal would be the identification of this intermediary factor and the determination of its binding sites with Tre-6-P and SnRK1 using methods such as x-ray crystallography or photoaffinity probes.[87] With this ambition, we synthesized analogues of Tre-6-P to elucidate structure-activity relationships with the intermediary factor, and used this data to design an affinity column for the attempted isolation and identification of this elusive protein.

1.3 The mannose receptor

Vertebrates have evolved an intricate immune system that comprises both innate and adaptive defence mechanisms.[88] At the heart of this system are pattern recognition receptors that can identify foreign antigens and invoke the immune response.[89] The mannose receptor (MR) that is ubiquitous across the mammalian family is an example of such a receptor.[90-92] As well as its direct pattern recognition involvement in innate defence, recent research has uncovered a potential role for the MR in the adaptive immune response which has ignited considerable interest in the protein.

The MR is an endocytotic C-type lectin expressed at the surface of a variety of cells, including macrophages, hepatic endothelial cells and possibly even dendritic cells.[91-94] This 180 kDa transmembrane protein shares its general structure with the three other members of this protein family.[95] These proteins are made up of multiple domains linked together in tandem that, for the MR, confer a multifunctional role (Figure 1.2).

The protein contains a transmembrane domain, a short cytosolic tail that is involved in endosomal sorting and receptor recycling after endocytosis,[96] and a segment containing a fibronectin type II repeat (FTII) which binds collagen.[97] The most important domains in the context of ligand binding are the eight C-type lectin-like carbohydrate recognition domains (CRDs) and the terminal cysteine rich region (CRR).

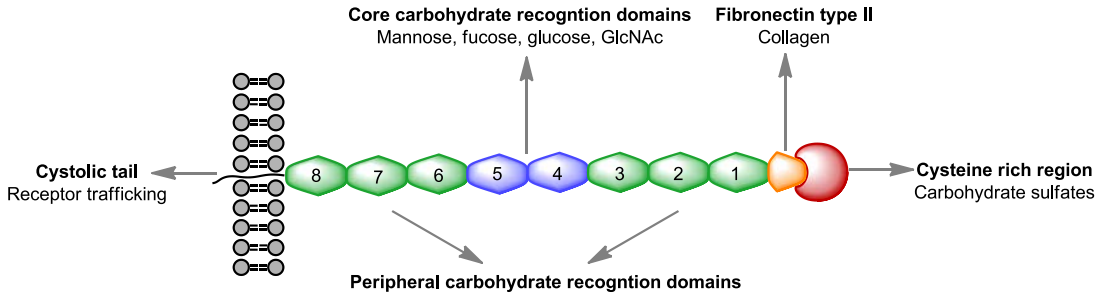


Figure 1.2. *The major domains of the MR and their ligands or function.*

1.3.1 Carbohydrate recognition and innate defence

The CRDs are primarily involved in the binding of oligosaccharides that are frequently found on the surface of pathogens and constitutes one of the first steps of the immune response.[92] Thus, MR interacts with mannose, fucose, *N*-acetylglucosamine (GlcNAc) and glucose which are prevalent on bacteria, yeasts and parasites, but far less common on mammalian cells.[98] In this way, the receptor can discriminate between "self" and "non-self" antigens.

These CRDs have been studied using a variety of techniques and as a result, a thorough understanding of their binding properties has been developed. There is considerable sequence variation between the individual CRDs and consequently these domains have differing interactions with carbohydrate ligands.[99] Fragment expression techniques have shown that CRD4 and CRD5 are the most important for binding to sugars, and together they form the core carbohydrate recognition element of the receptor.[100,101] To replicate the binding of the complete MR, CRDs 6-8 are also required although their binding affinity to carbohydrates is relatively small and they are only implicated in peripheral interactions with oligosaccharides. CRDs 1-3 do not show any avidity for ligands and are completely dispensable.

Interestingly, of all the active CRDs, only CRD4 retains carbohydrate binding activity when expressed alone.[100] While this domain interacts well with monosaccharides,

natural glycoproteins are not recognized, which implies that multivalent interactions involving other CRDs are essential for binding to oligosaccharides.[100] The receptor exists in a rigid and extended conformation where all domains are freely accessible, facilitating such clustered binding.[102] Since each CRD has only one binding site, the geometry of carbohydrate presentation to the receptor is a critical parameter controlling binding affinity. Thus branched oligosaccharides have far stronger binding compared to linear analogues.[103] However, the spacing between adjacent CRDs is variable so that different inter-domain distance permutations are possible allowing the receptor to accommodate different levels of branching.[103] In this way, the MR is able to strongly interact with a wide range of oligosaccharides.

Sequence comparisons with other well-characterized C-type lectins, such as mannose binding protein or transmembrane selectins,[104,105] to which the CRDs show considerable similarity, have proved useful for rationalizing the observed binding properties on a molecular level. For instance, the comparisons have shown that the critical amino acid motif required for carbohydrate interactions is only present on CRDs 4 and 5, explaining their significance to the overall binding properties of the receptor.[92,99] Furthermore, the predicted conformation only allows interaction with hexoses bearing equatorial hydroxyls at the 3 and 4 positions, thus explaining the observed ligand selectivity.[92,105] The other CRDs of the MR show much more divergence and a resultant reduction in affinity, although CRDs 6-8 have retained at least some of the amino acids that directly interact with carbohydrates enabling a marginal role for these domains.[99]

Given the importance of CRD4 to overall binding, research efforts have been directed toward this fragment as a simplified system for more detailed studies.[106-108] Along these lines, x-ray crystallography[107] and site-directed mutagenesis[108] have revealed a unique calcium binding dependence that is notably different to other C-type lectins. As well as a primary Ca^{2+} binding site that is common among all C-type lectins, CRD4 contains an additional binding position at an extended loop. For binding to carbohydrates, Ca^{2+} is essential at both sites, but its mechanism of action is very different at the two positions. While the primary Ca^{2+} site directly interacts with carbohydrates through ionic interactions, the secondary site operates through a calcium induced conformational change.[106] Crucially, the binding at this second site is pH dependent and Ca^{2+} is unbound in an acidic environment providing a mechanism

for receptor recycling after endocytosis, where glycoproteins are unloaded in the acidic lysosomes leaving the receptor free to return to the surface.

This pH dependence is suggestive of a primary endocytotic role for the CRDs, which is responsible for the innate immune response.[90,92,109] The MR expressed in macrophages (as well as other lymphatic cells) can interact with non-self sugars on a wide variety of pathogenic microorganisms and allow phagocytosis and subsequent degradation of the pathogen in the lysosome.[90,92,110] Furthermore, infected cells release glycoproteins such as lysosomal hydrolases[111] and tissue plasminogen activator[92,112] that are removed by the MR and in doing so, assist in the reduction of tissue damage during inflammation. As well as this direct attack on pathogens, there is some evidence to suggest that the receptor may be involved in signal transduction.[90,92] After the detection of pathogenic antigens, it is proposed that the receptor may trigger secretion of pro-inflammatory cytokines[90,92,113] and may modulate other cell surface receptors.[92,114] Thus, interaction with the MR may trigger a cascade of other innate defence mechanisms.

More recent research, however, contradicts this traditional viewpoint.[115] Experiments conducted with $MR^{-/-}$ knockout mice have shown a normal host defence when challenged with systematic *Candidia albicans* infections, which would seem to imply that the MR is a dispensable component during the immune response.[116] In fact, there is now some evidence to suggest that binding to the MR suppresses immune activity. For instance, binding of alveolar macrophage MRs with *Pneumocystis* resulted in a suppression of inflammatory cytokine release demonstrating a potential regulatory role for the receptor.[117] Further evidence supporting this view is the illustration that the binding of mucins to the MR has a similar inhibitory phenotype in tumour associated macrophages.[118] Thus, as well triggering the innate defences, the MR may have a further role in the regulation of the immune response and this area will require further research in the future.[119]

1.3.2 Sulfated carbohydrate recognition

The binding of carbohydrates with the CRDs has been extensively studied and the biological role of this interaction is now fairly well understood. However, the discovery that the CRR can also interact with ligands has opened up the possibility of

a much wider role for the MR. As a result, research efforts have been directed towards discovering potential ligands for the CRR and elucidating their biological relevance.

The potential for ligand binding at the CRR was originally discovered in endothelial rat liver cells, which were found to rapidly internalize pituitary hormones bearing a terminal GalNAc-4-SO_4 moiety.[120] Through the expression of MR protein fragments, the binding was found to be specific for the CRR and completely independent from carbohydrate interactions at the CRDs, where binding to galactose sugars is not tolerated.[121] As with the CRDs, the binding affinity was pH dependent and lost in an acidic environment, but in contrast, was independent of calcium ions.[120,121]

The GalNAc-4-SO_4 moiety was found to be essential for the correct physiological effect of the pituitary hormones lutropin, which is responsible for ovulation, and follitropin, which directs follicular development.[122] The sulfated sugar serves as a specific tag for rapid clearance from the bloodstream and thus controls the circulatory half-life of the glycoproteins.[122] In this way, a rapid increase in hormonal concentration can be followed by an equally rapid clearance through hepatic endothelial cells. Such pulsing in concentration is critical for hormonal activity and is reliant on the highly specific interaction of GalNAc-4-SO_4 with the CRR.[122,123] In support of this viewpoint, $MR^{-/-}$ deficient mice were found to have either a greatly reduced clearance of lutropin,[124] or the knockout was found to be lethal *in utero*.[122] Thus, sulfated oligosaccharides are a vital component in the reproductive system.

An important facet of sulfate binding to the MR is the modulation in ligand selectivity according to cell type.[120,121,125] While hepatic endothelial cells show strong binding to GalNAc-4-SO_4, alveolar macrophages have no such interaction despite both cells having a high expression of the receptor. This discrepancy is attributed to MR dimerization in the former, which allows divalent binding that greatly improves the avidity for GalNAc-4-SO_4.[125] Moreover, the selectivity of the receptor is also dramatically improved and the dimer is able to better discriminate between the 3 and 4 sulfates of *N*-acetyl galactosamine; both of which show reasonable affinity for the monomer, but for multivalent interactions, are magnitudes apart.[126] In pituitary hormones, the glycosylation patterns may be used to fine tune the binding affinity to the MR. For example, lutropin uses biantennary oligosaccharides with three possible sites of attachment on the protein chain.[127] The oligosaccharides can be sulfated at one

or both arms, but binding is significantly better when GalNAc-4-SO_4 moieties from two different oligosaccharides are involved.[125] Thus, the geometry of sulfate presentation is an important parameter in dictating binding affinity with the MR.

The mechanism of MR dimerization is unclear, though current evidence indicates that the dimers are held together through weak or transient interactions that allow reversible dimerization.[128] There are currently two schools of thought and both invoke the glycosylation of the MR itself as the critical factor that controls the receptor's fate. One theory relies upon receptor modification with high mannose oligosaccharides, followed by Ca^{2+} mediated interactions with the CRDs of a neighbouring receptor.[102] An alternative mechanism is based upon modification with sialic acid terminating oligosaccharides. In this theory, which has been demonstrated in an *in vitro* model, charged sialic acid residues prevent receptor dimerization through mutual charge repulsion.[129] Samples of the MR from different tissue types have differing masses which may indicate variances in post-translational glycosylation, consistent with both of the above theories.[128,129]

Such oligomerization is a common feature of lectin biology, where multivalency is used to enhance binding and selectivity to carbohydrate ligands.[130] For the binding of neutral carbohydrates, the array of CRDs facilitates multivalent interactions without the need for dimerization. The effect that this dimerization has on the binding properties of the CRDs is somewhat unclear. While dimeric MR from hepatic endothial cells of a Sprague-Dawley rat did not bind to mannose,[120,121,125] monocyte derived macrophages and dendritic cells expressing the dimeric form of the receptor did.[128] In fact, for these macrophages and dendritic cells, binding to the HIV envelope spike protein gp120 was only possible for the dimer through interaction with the CRDs.[128] This inconsistency may originate from the specific mechanism of dimerization. Oligomerization mediated though high mannose oligosaccharides would directly compete for binding at the CRDs and would prevent interactions with other carbohydrates, whereas other mechanisms for dimerization may sidestep this problem. There may be more than one dimerization mechanism, which raises the interesting possibility of a dimer mediated switch in ligand selectivity. The biochemical mechanism of this dimerization clearly needs further investigations, but in any case, the presence of dimers is essential for the binding of sulfates.[125,126]

The expression of the receptor is known to be heavily regulated according to the specific cellular environment (such as cell type, age and functional state)[92,131] through standard genetic mechanisms such as operons[132] and expression of transcription factors,[133] but is also modulated by other factors such as the presence of cytokines,[134] pathogenic products[135] and immunoglobin receptors.[136] Along these lines, it is reasonable to expect that receptor dimerization would be similarly regulated. Thus, analysis of the MR glycans has revealed a tissue specific sialylation which may indicate a role for the sialic acid regulation of dimerization.[137] In some cases, the functional activity of the MR is found to vary according to the environment despite consistent levels of expression, and may be as a result of receptor dimerization. For example, in cultured immature dendritic cells the MR is able to internalize mannosylated ligands, but upon maturation, the receptor has a greatly reduced endocytotic capacity, despite a continued high presence at the cell surface.[92,94,138] As a second example, the MR on monocyte-derived macrophages exhibits an increased propensity for phagocytosis after treatment with the cytokine interferon-gamma.[139] Both of these observations may be related to the oligomerization state of the MR.

1.3.3 Carbohydrate sulfates in adaptive immunity

As well as acting as a glycoprotein scavenger, there is now a significant body of evidence to show that the CRR may have a much broader biological role. Firstly, the GalNAc-4-SO_4 moiety has now been discovered on proteins that are completely unrelated to the pituitary hormones.[140] In addition to binding GalNAc-4-SO_4, recent studies have demonstrated that the CRR also interacts with chondroitin-4-sulfate chains and sulfated blood group oligosaccharides such as sialyl Lewis^a and Lewis^x (Figure 1.3).[141] Crystal structures of the CRR binding to various ligands have been solved which demonstrate that the majority of strong interactions are mediated by the sulfate group which binds in a neutral pocket of the protein.[142] Importantly, this allows the CRR to tolerate a variety of carbohydrate structures, since they engage in only minor interactions with the protein. There exist many glycoproteins bearing sulfated oligosaccharide chains that could be potential ligands for the CRR.[143] The biological relevance of the MR interaction with these other sulfated carbohydrates and glycoproteins is not fully understood. The high degree of conservation in the amino acid sequences of the CRR and the FTII region between the human and murine forms

of the MR seem to indicate an important function for these domains.[144] Moreover, the presence of the MR on a such wide variety of unrelated cell types,[91-93] such as smooth muscle cells[145] or retinal epithelial cells,[146] also remains to be explained. These observations raise the possibility of a much wider spectrum of activity for carbohydrate sulfates.

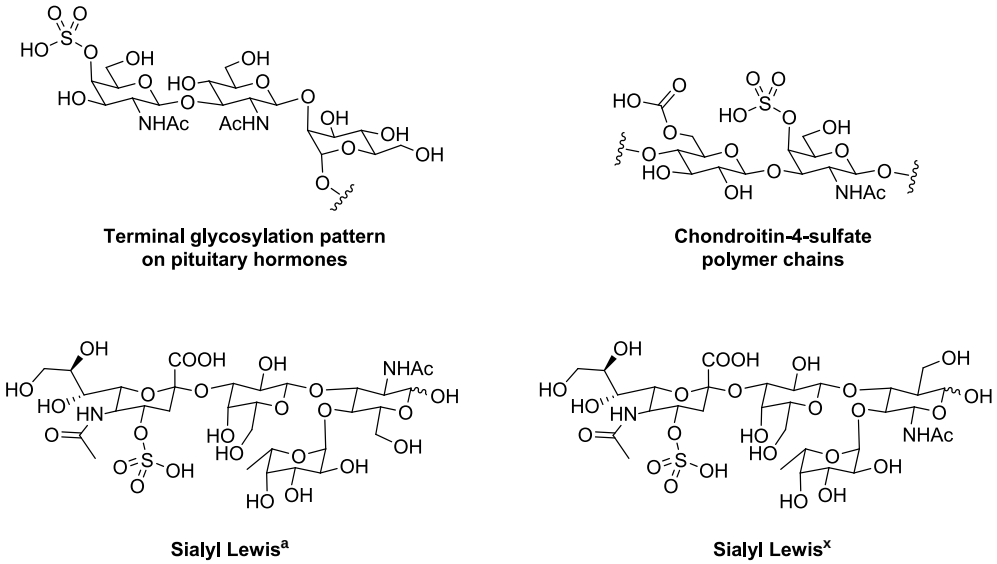


Figure 1.3. *Sulfated carbohydrates that are known to bind with the CRR of the MR.*

Current hypotheses favour a role for carbohydrate sulfates in linking the innate and adaptive immune responses, via their interaction with the MR.[91,92,141] In such a model, the CRR would facilitate the trafficking of MR-bearing cells for antigen presentation to areas of the lymphatic system that are important for the adaptive response.

Pioneering studies by Martinez-Pomarez *et al.* have demonstrated that a chimeric form of the receptor bearing the CRR and FTII domains was specifically targeted to metallophilic macrophages in the marginal zones of the spleen and the lymph node subcapsular sinus.[147] Importantly, these areas are directly adjacent to germinal centres where B-cell maturation occurs and are therefore vitally important components of the adaptive immune system. Thus, MR-bearing cells directed to these regions would be well placed to deliver foreign antigens and trigger adaptive immune defences.

Chondroitin-4-sulfate chains are found on secreted and cell-membrane associated proteoglycans of cells of the hematopoietic system, such as lymphocytes, monocytes and macrophages and could account for the observed targeting.[148,149] Indeed, two of the ligands for the CRR were found to be the membrane proteins sialoadhesin and

CD45, both of which contained sulfated structures.[150] Another likely candidate ligand is serglycin, which is a major component of the secreted proteoglycans of immune cells,[151] and is also know to contain chondroitin-4-sulfate chains.[149] Various other chondroitin sulfate proteoglycans are known and could be possible ligands.[152] Alternatively, the enzymatic machinery required for the installation of the GalNAc-4-SO_4 moiety has been found in various tissue types including the spleen which may provide a second source of ligands for the CRR.[150,153]

Moreover, the secretion of these sulfated ligands is highly regulated during the immune attack.[154] In the period of the immune response, the normal expression of sulfated ligands in metallophilic macrophages[155] is accompanied by sulfate production in dendritic cell that reside directly inside B cell follicles.[156] In addition, ligand producing cells can themselves migrate into B cell follicles,[157] and the secretion of proteoglycans such as serglycin, which is regulated according to various external stimuli, may also be altered.[149,158] Together, these changes allow direct targeting of the antigen into germinal centres for an improved adaptive immune response.[147]

Thus, it is hypothesized that sulfated carbohydrates form a network of markers for the navigation of components of the immune systems to areas of B cell maturation for antigen presentation. Some of these sulfated ligands, such as the sulfates of sialyl Lewis^a and Lewis^x, may have alternate binding targets with selectins that are expressed on leukocytes which may compete with the targeting of these cells.[159] These sulfated blood sugars are expressed on a variety of cells, but their distribution in the immune system remains to be determined.[141] However, these selectins bind only poorly to chondroitin sulfate chains and so this mechanism of targeting is specific for the MR.

1.3.4 Research objectives

For this sulfate mediated trafficking system to proceed, dimerization of the MR is a critical prerequisite. Thus, *in vivo* experiments with a cross-linked chimeric CR-containing protein allowed a 10 fold improvement in targeting efficiency.[160] A soluble form of the MR has been discovered that could potentially transport antigens, but a method for its dimerization in the solution phase is not currently known.[161] Alternatively, cells may themselves be transported to B cell follicles. To investigate

this possibility, and to better understand the apparent functional changes in MR activity in response to the cellular environment, a method to probe the dimerization state of the receptor is needed. Given the highly regulated and delicate nature of the system, the method would need to cause minimal disruption to the native environment, so as to be representative of the true *in vivo* state of the cells. Ultimately, we would hope to use such methodology to probe the dimerization state of the MR on various hematopoietic cells during an immune response to investigate the importance of sulfate mediated trafficking in adaptive defences, and to explore the effect of MR dimerization on overall cell function. Along these lines, we explore the use of fluorescence microscopy and Förster resonance energy transfer experiments to achieve these goals.

1.4 Thesis layout

Progress in these two different research directions is now presented. Initially, attention is focused on methodology development for the synthesis of Tre-6-P analogues. The syntheses of such compounds are hindered by two major obstacles which are addressed at the outset. Firstly, the modification of trehalose to give non-symmetrical analogues can prove difficult due to the C_2 symmetrical nature of the sugar resulting in poor yields for the desired compounds. In Chapter II, we explore options to improve on such desymmetrization reactions and ultimately present a high yielding route for the synthesis of Tre-6-P. A second problem involves the manipulation of the phosphoryl moiety which can prove troublesome in many synthetic routes due to orthoganality of protecting groups as well as the more detrimental problem of phosphate migrations and cyclizations. In Chapter III, methodology to control phosphate migrations is developed and a general route for the synthesis of Tre-6-P analogues is presented. These methodologies form the basis of our efforts to identify the intermediary factor for the binding of SnRK1 to Tre-6-P, details of which are documented in Chapter IV. The following chapter is dedicated to the exploration of MR, where we develop and apply a FRET imaging technique to probe the dimerization state of the receptor. Finally in Chapter VI, we move away from the main body of this thesis and investigate the use of matix-assisted laser desorption/ionization mass spectroscopy for better characterization of synthetic protein neoglycoconjugates including some of the fluorescent protein probes that are

used in Chapter V. In many ways, this work is a generalized theoretical extension of the desymmetrization kinetics explored in Chapter II. Each chapter concludes with a discussion on future perspectives for the various topics discussed herein.

1.5 References

1. (a) M. W. Bowler, M. J. Cliff, J. P. Waltho, G. M. Blackburn, *New. J. Chem.*, **2010**, 34, 784. (b) F. H. Westheimer, *Science*, **1987**, 235, 1173.
2. G. K. Schroeder, C. Lad, P. Wyman, N. H. Williams, R. Wolfenden, *Proc. Natl. Acad. Sci. U. S. A.*, **2006**, 103, 4052.
3. (a) C. J. Barker, C. Illies, G. C. Gaboardi, P. Berggren, *Cell Mol. Life Sci.*, **2009**, 66, 3851. (b) P. W. Majerus, *Annu. Rev. Biochem.*, **1992**, 61, 225.
4. P. Cohen, *Nature Cell Biol.*, **2002**, 4, E127.
5. (a) D. E. C. Cornbridge, *Phosphorus: An Outline of its Chemistry, Biochemistry, and Technology*, **1990**, Elsievier. (b) J. Emsley, D. Hall, *The Chemistry of Phosphorus: Environmental, Organic, Inorganic, Biochemical and Spectroscopic Aspects*, **1976**, Harper & Row.
6. (a) J. Lengeler, G. Drews, H. Schlegel, *The Biology of the Prokaryotes*, **1998**, Wiley-Blackwell. (b) B. Axelrod, *Metabolic Pathways*, **1967**, Academic.
7. (a) A. Jeanjean, M. Garcia, A. Leydet, J-L. Montero, A. Morere, *Biorg. Med. Chem.*, **2006**, 14, 3575. (b) S. Kornfield, *Annu. Rev. Biochem.*, **1992**, 61, 307. (b) S. Kornfield, I. Mellman, *Annu. Rev. Cell. Biol.*, **1989**, 5, 483.
8. (a) L. M. Stoolman, T. S. Tenforde, S. D. Rosen, *J. Cell Biol.* **1984**, 99, 1535. (b) T. A. Yednock, L. M. Stoolman, S. D. Rosen, *J. Cell Biol.* **1987**, 104, 713.
9. (a) P. Ghosh, N. M. Dahms, S. Kornfeld, *Nat. Mol. Cell Bio.*, **2003**, 4, 202. (b) W. S. Sly, H. D. Fischer, *J. Cell. Biochem.*, **1982**, 18, 67. (b) S. R. Pfeffer, *J. Membrane Biol.* **1988**, 103, 7.
10. (a) P. Chong, N. Chan, A. Kandil, B. Tripet, O. James, Y. P. Yang, S. P. Shi, M. Klein, *Infect. Immun.*, **1997**, 65, 4918. (b) H. J. Jennings, *The Molecular Immunology of Complex Carbohydrates*, **1988**, Plenum. (c) D. Evenberg, P. Hoogerhout, C. A. A. van Boeckel, G. T. Rijkers, E. C. Beuvery, J. H. van Boom, J. T. Poolman, *J. Infect. Dis.*, **1992**, 165, S152. (d) C. C. A. M. Peeters, D. Evenberg, P. Hoogerhout, H. Kayhty, L. Saarinen, C. A. A. van Boeckel, G.

A. van der Marel, J. H. van Boom, J. T. Poolman, *Infect. Immunol*. **1992**, 60, 1826.

11. (a) N. Gao, J. Shang, D. Huynh, V. L. Manthati, C. Arias, H. P. Harding, R. J. Kaufman, I. Mohr, D. Ron, J. R. Falk, M. A. Lehrman, *Mol. Cell. Biol.*, **2011**, 22, 2994. (b) D. O. Willenborg, C. R. Parish, W. B. Cowden, *FASEB J*., **1989**, 3, 1968.
12. C. A. Strott, *Endocr. Rev.*, **2002**, 23, 703.
13. (a) C. A. Strott, Y. Higashi, *J. Lipid Res.*, **2003**, 44, 1268. (b) A. A. Farooqui, L. A. Horrocks, *Mol. Cell Biochem.*, **1985**, 66, 87.
14. (a) F. Monigatti, B. Hekking, H. Steen, *Biochim. Biophys. Acta*, **2006**, 1764, 1904. (b) J. W. Kehoe, C. R. Bertozzi, *Chem. Biol.*, **2000**, 7, R57.
15. (a) K. Honke, N. Taniguchi, *Med. Res. Rev.*, **2002**, 22, 637. (b) L. V. Hopper, S. M. Manzella, J. U. Baenziger, *FASEB J.*, **1996**, 10, 1137.
16. N. G. Halford, M. J. Paul, *Plant Biotechnol. J.*, **2003**, 1, 381.
17. N. G. Halford, S. J. Hey, *Biochem. J.*, **2009**, 419, 247.
18. G. Manning, D. B. Whyte, R. Martinez, T. Hunter, S. Sudarsanam, *Science*, **1997**, 298, 1912.
19. The Arabidopsis Genome Initiative, *Nature*, **2000**, 408, 796.
20. (a) F. Rolland, E. Baena-Gonzalez, J. Sheen, *Annu. Rev. Plant. Biol.*, **2006**, 57, 675. (b) M. Ramon, F. Rolland, J. Sheen, *Sugar Sensing and Signalling*, The Arabidopsis Book, **2008**, American Society of Plant Biologists.
21. (a) M. J. Paul, L. F. Primavesi, D. Jhurreea, Y. Zhang, *Annu. Rev. Plant. Biol.*, **2008**, 59, 417. (b) O. J. M. Goddijn, K. Van Dun, *Trends Plant Sci.*, **1999**, 4, 314.
22. M. J. Paul, *Curr. Opin. Plant. Biol.*, **2007**, 10, 303.
23. (a) J. H. Crowe, J. F. Carpenter , L. M. Crowe, *Annu. Rev. Physiol.*, **1998**, 60, 73. (b) F. Albertorio, V. A. Chapa, X. Chen, A. J. Diaz, P. S. Cremer, *J. Am. Chem. Soc.*, **2007**, 129, 10567.
24. J. C. Arguelles, *Arch. Microbiol.*, **2000**, 174, 217.
25. A. Becker, P. Schloder, J. E. Steele, G. Wegener, *Cell Mol. Life Sci.*, **1996**, 52, 433.
26. R. P. Adams, E. Kendall, K. K. Kartha, *Biochem. Syst. Ecol.*, **1990**, 18, 107.

27. N. Avonce, A. Mendoza-Vargas, E. Morett, G. Iturriaga, *BMC Evol. Biol.*, **2006**, 6, 109.
28. B. Leyman, D. Van Dijk, J. M. Thevelein, *Trends Plant. Sci.*, **2001**, 6, 510.
29. G. Langekamper, R. W. M. Fung, R. D. Newcomb, R. G. Atkinson, R. C. Gardner, E. A. MacRae, *J. Mol. Evol.*, **2002**, 54, 322.
30. (a) M. Ramon, I. De Smet, L. Vandesteene, M. Naudts, B. P. Leyman, P. Van Dijk, F. Rolland, J. Beeckman, J. M. Thevelein, *Plant Cell Environ.*, **2009**, 32, 1015. (b) L. Vandesteene, M. Ramon, K. L. Roy, P. Van Dijck, F. Rolland, *Mol. Plant*, **2010**, 3 406.
31. P. Van Dijk, J. O. Mascorro-Gallardo, M. De Bus, K. Royakers, G. Itturiaga, J. M. Thevelein, *Biochem. J.*, **2002**, 366, 63.
32. J. E. Lunn, *Funct. Plant Biol.*, **2007**, 34, 550.
33. M. Schmid, T. S. Davidson, S. R. Henz, U. J. Pape, M. Demar, M. Vingron, B. Scholkopf, D. Weigel, J. U. Lohmann, *Nat. Genet.*, **2005**, 37, 501.
34. V. Veluthambi, S. Mahadevan, R. Maheshwari, *Plant Physiol.*, **1982**, 69. 1247.
35. A. Wingler, T. Fritzius, A. Wiemken, T. Boller, R. A. Aeshbacher, *Plant Physiol.*, **2000**, 124, 105.
36. (a) S. Penna, *Trends Plant Sci.*, **2003**, 8, 355. (b) M. Ramon, F. Rolland, *Trends Plant Sci.*, **2007**, 12, 185.
37. O. M. J. Goddijn, T. C. Verwoerd, E. Voogd, R. W. Kurtwagen, P. T. de Graaf, K. Van Dun, J. Poels, A. S. Ponstein, B. Damm, J. Pen, *Plant. Physiol.*, **1997**, 113, 181.
38. H. Schluepmann T. Pellney, A. van Dijken, S, Smeekens, M. J. Paul, *Proc. Natl. Acad. Sci. USA*, **2003**, 100, 6849.
39. H. Schluepmann, A. van Dijken, M. Aghdasi, B. Wobbes, M. Paul, S. Smeekens, *Plant Physiol.*, **2004**, 135, 879.
40. J. E. Lunn, R. Feil, J. H. M. Hendricks, Y. Gibon, R. Morcuende, D. Osuna, W. Scheible, P. Carillo, M. Hajirezaei, M. Stitt, *Biochem. J.*, **2006**, 397, 139.
41. P. J. Eastmond, A. J. van Dijken, M. Spielman, A. Kerr, A. F. Tissier, H. G. Dickinson, J. D. Jones, S. C. Smeekens, I. A. Graham, *Plant J.*, **2002**, 29, 223.
42. S. Baud, I. A. Graham, *Plant J.*, **2006**, 46, 155.

43. (a) L. D. Gomez, S. Baud, I. A. Graham, *Biochem. Soc. Trans.*, **2005**, 33, 280. (b) S. N. Chary, G. R. Hicks, Y. G. Choi, D. Carter, N. V. Raikhel, *Plant Physiol.*, **2008**, 146, 97.
44. A. J. H. Van Dijken, H. Schluepmann, S. C. M. Smeekens, *Plant Physiol.*, **2004**, 135, 969.
45. S. Debast, A. Nune-Nesi, M. R. Hajirezai, J. Hofmann, U. Sonnewald, A. R. Fernie, F. Bornke, *Plant Physiol.*, **2011**, 156, 1754.
46. T. K. Pellny, O. Ghannoum, J. P. Conroy, H. Schluepmann, S. Smeekens, J. Amdralojc, K. Krause, O. Goddijn, M. J. Paul, *Plant. Biotechnol. J.*, **2004**, 2, 71.
47. N. Avonce, B. Leyman, . O. Mascorro-Gallardo, P. Van Dijk, J. M. Thevelein, G. Itturiaga, *Plant Physiol.*, **2004**, 136, 3649.
48. (a) M. Kondrak, F. Marincs, F. Antal, Z. Juhasz, Z. Banfalvi, *BMC Plant Biol.*, **2012**, 12. (b) C. Romero, J. L. Belles, J. L. Vaya, R. Serrano, F. A. Cullianes-Macia, *Planta*, **1997**, 201, 293. (c) E. A. H. Pilon-Smits, N. Terry, T. Seors, H. Kim, A. Zayed, S. Hwang, K. Van Don, E. Voogd, T. C. Verwoerd, R. W. Krytwagen, O. M. J. Goddijn, *J. Plant. Physiol.*, **1998**, 152, 525.
49. N. Satoh-Nagasawa, N. Nagasawa, S. Malcomber, H. Sakai, D. Jackson, *Nature*, **2006**, 441, 227.
50. M. A. Balzquez, R. Lagunas, C. Gancedo, J. M Gancedo, *FEBS Lett.*, **1993**, 329, 51.
51. M Kandel-Kfir, H. Damari-Weissler, M. A. German, D. Gidoni, A. Mett, E. Belausov, M. Petreikov, N. Adir, D. Granot, *Planta*, **2006**, 224, 1341.
52. (a) J. Sheen, E. Baena-Gonzalez, *Trends Plant Sci.*, **2008**, 13, 474. (b) N. G. Halford, P. C. Purcell, D. G. Hardie, *Trends Plant Sci.*, **1999**, 4, 117. (c) B. Moore, L. Zhou, F. Rolland, Q. Hall, W. Cheng, Y. Liu, I. Hwang, T. Jones, J. Sheen, *Science*, **2003**, 300, 332.
53. A. Kolbe, A. Tiessen, H. Schluepmann, M. Paul, S. Ulrich, P. Geigenberger, *Proc. Natl. Acad. Sci.*, **2005**, 102, 31, 11118.
54. (a) H. Neuhaus, M. Stitt, *Planta*, **1990**, 182, 445. (b) P. Geigenberger, M. Stitt, A. R. Fernie, *Plant Cell Environ.*, **2004**, 27, 655.
55. M. A. Ballicora, J. B. Frueauf, Y. Fu, P. Schurmann, J. Preiss, *J. Biol. Chem.*, **1998**, 275, 1315.

56. Y. Zhang, L. F. Primavesi, D. Jhurreea, P. J. Andralojc, R. A. C. Mitchell, S. J. Powers, H. Schluepmann, T. Delatte, A. Wingler, M. J. Paul, *Plant Physiol.* **2009**, 149, 1860.
57. C. Polge, M. Thomas, *Trends Plant Sci.*, **2007**, 12, 1360.
58. E. Baena-Gonzalez, J. Sheen, *Trends Plant Sci.*, **2008**, 13, 474.
59. M. Jossier, J. Bouly, P. Meimoun, A. Arjmand, P. Lessard, S. Hawley, D. G. Hardie, M. Thomas, *Plant J.*, **2009**, 59, 316.
60. S. J. Hey, E. Byrne, N. G. Halford, *Annals. Of Botany*, **2010**, 105, 197.
61. E. Baena-Gonzalez, F. Rolland, J. M. Thevelein, J. Sheen, *Nature*, **2007**, 488, 938.
62. D. Osuna, B. Usadel, R. Morcuende, Y. Gibon, O. E. Bläsing, M. Höhne, M. Günter, B. Kamlage, R. Trethewey, W-R Scheible, M. Stitt, *Plant J.*, **2007**, 49, 463.
63. B. Usadel, *Plant Physiol.*, **2008**, 46, 1834.
64. P. C. Purcell, A. M. Smith, N. G. Halford, *Plant J.*, **1998**, 14, 195.
65. S. Laurie, R. S. McKibbin, N. G. Halford, *J. Exp. Bot.*, **2003**, 54, 739.
66. R. Radchuk, V. Radchuk, W. Weschke, L. Borisjuk, H. Weber, *Plant Physiol.*, **2006**, 140, 263.
67. Y. Zhang, P. R. Shewry, H. Jones, P. Barcelo, P. A. Lazzeri, N. G. Halford, *Plant J.*, **2001**, 28, 431.
68. A. Lovas, A. Bimbó, L. Szabó, Z. Bánfalvi, *Plant J.*, **2003**, 33, 139.
69. M. Thelander, T. Olsson, H. Ronne, *EMBO J.*, **2004**, 23, 1900.
70. J. Schwachtje, P. E. H. Minchin, S. Jahnke, J T. van Dongen, U. Schittko, I. T. Baldwin, *Proc. Natl. Acad. Sci. USA*, **2006**, 103, 12935.
71. (a) L. Hao, H. Wang, G. Sunter, D. M. Bisaro, *Plant Cell*, **2003**, 15, 1034. (b) L. Gissot, C. Polge, M. Jossier, T. Girin, J-P Bouly, M. Kreis, M. Thomas, *Plant Physiol.*, **2006**, 142, 931.
72. (a) D. G. Hardie, *Nature*, **2007**, 8, 774. (b) S. L. McGee, M. Hargreaves, *Front. Biosci.*, **2008**, 13, 3022.
73. K. Hedbacker, M. Carlson, *Front. Biosci.*, **2008**, 13, 2408.
74. W. A. Wilson, S. A. Hawley, D. G. Hardie, *Curr. Biol.*, **1996**, 6, 1426.
75. N. G. Halford, S. Hey, D. Jhurreea, S. Laurie, R. S. McKibbin, M. Paul, Y. Zhang, *J. Exp. Bot.*, **2003**, 54, 467.

76. J. P. Bouly, L. Gissot, P. Lessard, M. Kreis, M. Thomas, *Plant J.*, **1999**, 18, 541.
77. N. G. Halford, J. Boulyz, M. Thomas, *Advances in botanical research*, **2000**, Academic Press.
78. E. M. Hraback, C. W.M. Chan, M. Gribskov, J. F. Harper, J. H. Choi, N. Halford, J. Kudla, S. Luan, H. G. Nimmo, M. R. Sussman, M. Thomas, K Walker-Simmons, J-K Zhu, A. C. Harmon, *Plant. Physiol.*, **2003**, 132, 666.
79. (a) L. Gissot, C. Polge, J-P Bouly, T. Lemaitre, M. Kreis, M. Thomas, *Plant Mol. Biol.*, **2004**, 56, 747. (b) V. Lumbreras, M. M. Albà, T. Kleinow, C. Koncz M. Pagès, *EMBO Rep.*, **2001**, 2, 55.
80. B. E. Kemp, *J. Clin. Invest.*, **2004**, 113, 182.
81. (a) S Hey, H. Mayerhofer, N. G. Halford, J. R. Dickinson, *J. Biol. Chem.*, **2007**, 282, 10479. (b) W. Shen, L. Hanley-Bowdoin, *Plant Physiol.*, **2006**, 142, 1642.
82. C. Sugden, R. M. Crawford, N. G. Halford, D. G. Hardie, *Plant J.*, **1999**, 19, 433.
83. A. Tiessen, K. Preschka, A. Branscheid, N. Palacios, R. McKibbin, N. G. Halford, P. Geigenberger, *Plant J.*, **2003**, 35, 490.
84. D. Toroser, Z. Plaut, S. C. Huber, *Plant Physiol.*, **2000**, 123, 403.
85. M. J. Paul, L. F. Primavesi, *Personal Communication*.
86. (a) E. A. Ananieva, G. E. Gillapsy, A. Ely, R. N. Burnette, F. L. Erickson, *Plant Physiol.*, **2008**, 148, 1868. (b) L. Gissot, C. Polge, M. Jossier, T. Girin, J-P. Bouly, M. Kreis, M. Thomas, *Plant Physiol.*, **2006**, 142, 931.
87. (a) L. Dubinsky, B. P. Krom, M. M. Meijler, *Bioorg. Med. Chem.*, **2012**, 20, 554. (b) S. A. Fleming, *Tetrahedron*, **1995**, 51, 46, 12479. (c) J. Brunner, *Annu. Rev. Biochem*, **1993**, 62, 483. (d) V. Chowdhry, *Annu. Rev. Biochem.*, **1979**, 48, 293.
88. C. A. Janeway, P. A. Travers, M. Walport, M. J. Schlomchik, *Immunobiology*, **2005**, Garland Science.
89. A. Pluddemann, S. Mukhopadhyay, S. Gordon, *Immunol. Rev.*, **2011**, 240, 11.
90. P. R. Taylor, L. Martinez-Pomarez, M. Stacey, H. Lin, G. D. Brown, S. Gordon, *Annu. Rev. Immunol.*, **2005**, 23, 901.
91. (a) P. R. Taylor, S. Gordon, L. Matinez-Pomares, *Trends Immunol.*, **2005**, 26, 104. (b) L. Matinez-Pomares, S. A. Linehan, P. R. Taylor, S. Gordon,

Immunobiol., **2001**, 204, 527. (c) D. S. Roseman, J. U. Baenziger, *Methods Enzymol.*, **2003**, 363, 121

92. P. D. Stahl, R. A. B. Ezekowitz, *Curr. Opin. Immunol.*, **1998**, 10, 50.
93. (a) J. H. Martens, J. Kzhyshkowska, M. Falkowski-Hansen, K. Schledzewki, A. Gratchev, U. Mansamann, C. Schmutermaier, E. Dippel, *et al., J. Pathol.*, **2006**, 208,574. (b) S. A. Linehan, L. Martinez-Pomares, P. D. Stahl, S. Gordon, *J. Exp. Med.*, **1999**, 189, 1961. (c) L. East, C. M. Isacke, *Biochim. Biophys. Acta*, **2002**, 1572, 364.
94. (a) A. Avrameas, D. Mcillroy, A. Hosmalin, B. Autran, P. Debre, M. Monsigny, A. C. Roche, P. Mildoux, *Eur. J. Immunol*., **1996**, 26, 394. (b) F. Sallusto, M. Cella, C. Danieli, A. Lanzavecchia, *J. Exp. Med.*, **1995**, 182, 389. (c) A. J. Engering, M. Cella, D. M. Fluitsma, E. C. Hoefsmit, A. Lanzavecchia, *Adv. Exp. Med. Biol.*, **1997**, 417, 183.
95. (a) L. East, C. M. Isacke, *Biochim. Biophys. Acta*, **2002**, 1572, 364. (b) M. E. Taylor, *Glycobiology*, **1997**, 7, v.
96. A. Schweizer, P. D. Stahl, J. Rohrer, *J. Biol.Chem*., **2000**, 275, 29694.
97. C. E. Napper, K. Drickamer, M. E. Taylor, *Biochem. J.*, **2006**, 395, 579.
98. (a) Y. V. Kook, A. Engering, A. N. Lekkerkerker, I. S. Ludwig, T. B. H. Geijtenbeek, *Curr. Opin. Immunol*., **2004**, 16, 488. (b) R. T. Lee, T-L. Hsu, S. K. Huang, S-L. Hsieh, C-H. Wong, Y. C Lee, *Glycobiology*, **2011**, 21, 512. (c) B. L. Largent, K. M. Walton, C. A. Hoppe, Y. C. Lee, R. L. Schnaar, *J. Biol. Chem.*, **1984**, 259, 1764. (d) E. R. Parise, M. E. Taylor, J. A. Summerfield, *J. Lab. Clin. Med.*, **1984**, 104, 908.
99. M. E. Taylor, J. T. Conray, M. R. Lennartz, P. D. Stahl, K. Drickamer, *J. Biol. Chem.*, **1990**, 265, 12156.
100. M. E. Taylor, K. Bezouska, K. Drickamer, *J. Biol. Chem.*, **1992**, 267, 1719.
101. (a) M. E. Taylor, K. Drickamer, *J. Biol. Chem.*, **1993**, 268, 399. (b) N. P. Mullin, K. T. Hall, M. E. Taylor, *J. Biol. Chem.*, **1994**, 269, 28405.
102. C. E. Napper, M. H. Dyson, M. E. Taylor, *J. Biol. Chem.*, **2001**, 276, 14759.
103. (a) V. Kery, J. J. Krepinsky, C. D. Warren, P. Capek, P. D. Stahl, *Biochim. Biophys.*, **1992**, 298, 49. (b) V. Kery, J. J. f. Krepinsky, C. D. Warren, P. Capek, S. D. Stahl, *Arch. Biochem. Biophys.*, **1992**, 298, 49.
104. K. Drickamer, *Curr. Opin. Struct. Biol.*, **1993**, 3, 393.

105. S. T. Iobst, K. Drickamer, *J. Biol. Chem.*, **1994**, 269, 15512.
106. H. Feinberg, S. Park-Snyder, A. R. Kolatkar, C. T. Heisse, M. E. Taylor, W. I. Weis, *J. Biol. Chem.*, **2000**, 275, 21539.
107. Z. Otwinowski, W. Minor, *Methods Enzymol.*, **1997**, 276, 307.
108. N. P. Mullin, P. G. Hitchen, M. E. Taylor, *J. Biol. Chem.*, **1997**, 272, 5668.
109. (a) G. J. Doherty, H. T. McMahon, *Annu. Rev. Biochem.*, **2009**, 78, 31. (b) F. R. Maxfield, T. E. McGraw, *Nat. Rev. Mol. Cell Biol.*, **2004**, 5, 121. (c) P. D. Stahl, P. H. Schlessinger, E. Sigardson, J. S. Rodman, Y. C. Lee, *Cell*, **1980**, 19, 207. (d) C. Tietze, P. Schlessinger, P. Stahl, *J. Cell Biol.*, **1982**, 92, 417.
110. (a) J. P. Upham, D. Pickett, T. Irimura, E. M. Anders, P. C. Reading, *J. Virol.*, **2010**, 84, 3730. (b) M. D. Tate, A. G. Brooks, P. C. Reading, *Immunol. Cell Biol.*, **2011**, 89, 482. (c) R. A. B. Ezekovitz, K. Sastry, P. Bailly, A. Warner, *J. Exp. Med.*, **1990**, 172, 1785. (d) R. A. B. Ezekovitz, D. J. Williams, H. Koziel, M. Y. K. Armstrong, A. Warner, F. F. Richards, R. M. Rose, *Nature*, **1991**, 351, 155. (e) C. Peters, T. Aebischer, Y.D. Stierhof, M. Fuchs, P. Overath, *J. Cell Sci.*, **1995**, 108, 3715.
111. (a) P. D. Stahl, P. H. Schlesinger, *Trends Biochem. Sci.*, **1980**, 5, 194. (b) S. E. Ponntow, V. Kery, P. D. Stahl, *Int. Rev. Cytol.*, **1992**, 137, 221.
112. M. Otter, M. M. Barrett-Bergshoeff, D. C. Rijken, *J. Biol. Chem.*, **1990**, 265, 12156.
113. (a) Y. Yamamoto, T. W. Klein, H. Friedman, *Infect. Immun.*, **1997**, 65, 1077. (b) Y. Shibata, W. J. Metzger, Q. N. Myrvik, *J. Immunol.*, **1997**, 159, 2462.
114. (a) M. Murai, Y. Aranaki, S. Tsuchiya, *Immunology*, **1996**, 89, 436. (b) M. Murai, Y. Aranaki, S. Tsuchiya, *J. Leuk. Biol.*, **1995**, 57, 687. (c) J. Bernardo, A. M. Billingslea, R. M. Blumenthal, K. F. Seetoo, E. R. Simons, M. J. Fenton, *Infect. Immun.*, **1998**, 66, 28.
115. U. Gazi, M. Martinez-Pomares, *Immunobiology*, **2009**, 214, 554.
116. S. J. Lee, N-Y. Zheng, M. Clavijo, M. C. Nussenzweig, *Infect. Immun.*, **2003**, 71, 437.
117. J. Zhang, S. D. Tachado, N. Patel, J. Zhu, A. Imrich, P. Manfruelli, M. Cushion, T. B. Kinane, H. Koziel, *J. Leukoc. Biol.*, **2005**, 78, 665.
118. P. Allavena, M. Chieppa, G. Bianchi, G. Solinas, M. Fabbri, G. Laskarin, A.Mantovani, *Clin. Dev. Immunol.*, **2010**, 1.

119. (a) P. Redelinghuys, G. D. Brown, *Immunol. Lett.*, **2011**, 136, 1. (b) J. V. Ravetch, L. L. Lanier, *Science*, **2000**, 290, 84.
120. (a) D. Fiete, V. Srivastava, O. Hindsgaul, J. U. Baenziger, *Cell*, **1991**, 67, 1103. (b) D. Fiete, J. U. Baenziger, *J. Biol. Chem.*, **1997**, 272, 14629.
121. (a) D. Fiete, M. C. Beranek, J. U. Baenziger, *Proc. Natl. Acad. Sci. USA*, **1997**, 94, 11256, (b) D. Fiete, M. C. Beranek, J. U. Baenziger, *Proc. Natl. Acad. Sci. USA*, **1998**, 95, 2089.
122. Y. Mi, S. D. Shapiro, J. U. Baenziger, *J. Clin. Invest.*, **2002**, 109, 269.
123. S. Mukherjee, K. Palczewski, V. Gurevich, J. L. Benovic, J. P. Banga, M. Hunzicker-Dunn, *Proc. Natl. Acad. Sci. USA*, **1999**, 96, 493.
124. S. J. Lee, S. Evers, D. Roeder, A. F. Parlow, J. Risteli, L. Risteli, Y. C. Lee, T. Feizi, H. Langen, M. C. Nussenzweig, *Science*, **2002**, 295, 1898.
125. D. S. Roseman, J. U. Baenziger, *Proc. Natl. Acad. Sci. USA*, **2000**, 97, 9949.
126. D. S. Roseman, J. U. Baenziger, *J. Biol. Chem.*, **2001**, 276, 17052.
127. (a) E. D. Green, J. U. Baenziger, *J. Biol. Cehm.*, **1988**, 263, 25. (b) E. D. Green, J. U. Baenziger, *J. Biol. Cehm.*, **1988**, 263, 36. (c) J. U. Baenziger, E. D. Green, *Biochim. Biophys. Acta*, **1988**, 947, 287.
128. J. Lai, O. K. Bernhard, S. G. Turville, A. N. Harman, J. Wilkinson, A. L. Cunningham, *J. Biol. Chem.*, **2009**, 284, 11027.
129. Y. Su, T. Bakker, J. Harris, C. Tsang, G. D. Brown, M. R. Wormald, S. Gordon, R. A. Dwek, P. M. Rudd, L. Martinez-Pomarez, *J. Biol. Chem.*, **2005**, 280, 32811.
130. (a) N. Jayaraman, *Chem. Soc. Rev.*, **2009**, 38, 3463. (b) J. J. Lundquist, E. J. Toone, *Chem. Rev.*, **2002**, 102, 555. (c) M. Mammen, S. Choi, G. M. Whitesides, *Angew. Chem. Int. Ed.*, **1998**, 37, 2754.
131. Z. Chroneos, V. L. Shepherd, *Am. J. Physiol. Lung Cell Mol. Physol.*, **1995**, 269, L721.
132. F. Rouleux, M. Monsigny, A. Legrand, *Exp. Cell Res.*, **1994**, 214, 113.
133. O. Eichbaum, D. Heney, D. Raveh, M. Chung, M. Davidson, J. Epstein, R. A. B. Ezekovitz, *Blood*, **1997**, 90, 4135.
134. (a) M. Stein, S. Keshav, N. Harris, S. Gordon, *J. Exp. Med.*, **1992**, 176, 287. (b) K. M. DeFife, C. R. Jenney, A. K. McNally, E. Colton, J. M. Anderson, *J.*

Immunol., **1997**, 158, 3385. (c) S. Scheiber, S. L. Perkins, S. L. Teitelbaum, J. Chappel, P. D. Stahl, J. S. Blum, *J. Immunol.*, **1993**, 151, 4973.

135. (a) D. L. Lefkovitz, J. A. Lincoln, S. S. Lefkovitz, A. Bollen, N. Moguilevsky, *Immunol. Cell Biol.*, **1997**, 75, 136. (b) V. L. Shepherd, K. B. Lane, R. Abdolrasulnia, *Arch. Biochem. Biophys.*, **1997**, 344, 350.
136. (a) S. Schreiber, W. F. Stenson, R. P. McDermitt, P. D. Stahl, S. L. Teitelbaum, S. L. Perkins, *Proc. Natl. Acad. Sci. USA*, **1991**, 88, 1616. (b) S. Schreiber, W. F. Stenson, R. P. MacDermott, J. C. Chappel, S. L. Teitelbaum, S. L. Perkins, *J. Immulnol.*, **1991**, 147, 1377.
137. Y. Su, L. Royle, C. M. Radcliffe, D. J. Harvey, R. A. Dwek, L. Martinez-Pomares, P. M. Rudd, *Biochem. Biophys Res. Commun.*, **2009**, 384, 436.
138. C. R. E. Sousa, P. D. Stahl, J. M. Austin, *J. Exp. Med.*, **1993**, 178, 509. (b)
139. L. Marodi, S. Schreiber, D. C. Anderson, R. P. MacDermott, H. M. Korchak, R. B. Johnston Jr, *J. Clin. Invest.*, **1993**, 91, 2596.
140. D. Fiete, Y. Mi, E. L. Oats, M. C. Beranek, J. U. Baenziger, *J. Biol. Chem.*, **2007**, 282, 1873.
141. C. Leteux, W. Chai, R. W. Loveless, C. Yuen, L. Uhlin-Hansen, Y. Combarnous, M. Jankovic, S. C. Maric, Z. Misulovin, M. C. Nussenzweig, T. Feizi, *J. Exp. Med.*, **2000**, 191, 1117.
142. (a) Y. Liu, A. J. Chirino, Z. Misulovin, C. Leteux, T. Feizi, M. C. Nussenzweig, P. J. Bjorkman, *J. Exp. Med.*, **2000**, 191,1105. (b) Y. Liu, Z. Misulovin, P. J. Bjorkman, *J. Mol. Biol.*, **2001**, 305, 481.
143. (a) V. K. Karaivanova, R. G. Spiro, *Biochem. J.*, **1998**, 329, 511. (b) N. Kawasaki, Y. Haishima, M. Ohta, S. Itoh, M. Hyuga, S. Hyuga, T. Hayakawa, *Glycobiology*, **2001**, 11, 1043. (c) J. Kovensky, *Curr. Med. Chem.*, **2009**, 16, 2338.
144. N. Harris, M. Super, M. Rits, G. Chang, R. A. B. Ezekovitz, *Blood*, **1992**, 80, 2363.
145. D. B. Lew, E. Songu-Mize, S. E. Ponto, P. D. Stahl, M. C. Rattezzi, *J. Clin. Invest.*, **1994**, 94, 1855.
146 V. L. Shepherd, B. I. Tarnowski, B. J. McLaughlin, *Invest. Opthamol. Vis. Sci.*, **1991**, 32. 1779.

147. L. Martinez-Pomares, M. Kosco-Vilbois, E. Darley, P. Tree, S. Herren, J. Bonnefoy, S. Gordon, *J. Exp. Med.*, **1996**, 184, 1927.

148. (a) S. Haylock-Jackobs, M. B. Keough, L. Lau, V. W. Yong, *Autoimm. Rev.*, **2011**, 10, 766. (b) P. Du Souich, A. G. Garcia, J. Verges, E. Montell, *J. Cell Mol. Med.*, **2009**, 13, 1451. (a) S. O. Kolset, J. T. Gallagher, *Biochim, Biophys. Acta.*, **1990**, 1032, 191. (b) H. Kresse, H. Hauser, E. Schonherr, *Experimentia*, **1993**, 49, 403. (c) E. Ruoslahti, *J. Biol. Chem.*, **1989**, 264, 13369.

149. (a) S. O. Kolsetm, G. Pejler, *J. Immunol.*, **2011**, 187, 4927. (b) N. Toyama-Sorimachi, N. F. Kitamura, H. Habuchi, Y. Tobita, K. Kimata, M. Miyasaka, *J. Biol. Chem.*, **1997**, 272, 26714.

150. L. Martinez-Pomares, P. R. Crocker, R. D. Silva, N. Holmes, C. Colominas, P. Rudd, R. Dwek, S. Gordon, *J. Biol. Chem.*, **1999**, 274, 35211.

151. M. A. Kulseth, S. L. Mustorp, L. Uhlin-Hansen, F. Oberg, S. O. Kolset, *Glycobiology*, **1998**, 8, 747.

152. (a) J. D Esko, K. Kimata, U. Lindahl, *Essentials of Glycobiology*, Chapter 16: Proteoglycans and Sulfated Glycosaminoglycans, **2009**, *Cold Spring Harbor*. (b) M. Piepkorn, P. Hovingh, K. L. Bennett, A. Aruffo, A. Linker, *Biochem. J.*, **1997**, 327, 499.

153. (a) R. G. Spiro, V. D. Bhoyroo, *Biochem. J.*, **1998**, 331, 265. (b) S. M. Dharmesh, T. P. Skelton, J. U. Baenziger, *J. Biol. Chem.*, **1993**, 268, 17096.

154. P. R. Taylor, S. Zamze, R. J. Stillion, S. Y. C. Wong, S. Gordon, L. Martinez-Pomares, *Proc. Natl. Acad. Sci. USA*, **2004**, 101, 1963.

155. P. Yu, Y. Wang, R. K. Chin, L. Martinez-Pomares, S. Gordon, M. H. Kosco-Vibois, J. Cyster, Y. X. Fu, *J. Immunol.*, **2002**, 168, 5117.

156. C. Berney, S. Herren, C. A. Power, S. Gordon, L. Martinez-Pomares, M. H. Kosco-Vibois, *J. Exp. Med.*, **1999**, 190, 851.

157. (a) F. Martin, J. F. Kearney, *Nat. Rev. Immunol.*, **2002**, 2, 323. (b) S. Pillai, A. Cariappa, S. T. Moran, *Annu. Rev. Immunol.*, **2005**, 23, 161. (c) P. H. Groeneveld, T. Erich, G. Kraal, *Immunology*, **1986**, 58, 285.

158. P. D. Kittlick, D. Engelmann, *Exp. Toxicol. Pathol.*, **1992**, 44, 407.

159. (a) C. Ehrhardt, C. Kneuer, U. Barkowsky, *Adv. Drug Deliver. Rev.*, **2004**, 56, 527. (b) S. Wedepohl, F. Beceren-Braun, S. Riese, K. Buscher, S. Enders, G. Bernhard, K. Kilian, V. Blanchard, J. Dernedde, R. Tauber, *Eur. J. Cell Biol.*,

2012, 91, 257. (c) P. R. Crocker, T. Feizi, *Curr. Opin. Sruct. Biol.*, **1996**, 6, 679. (d) A. Varki, *Proc. Natl. Acad. Sci. USA*, **1994**, 91, 7390. (e) R. P. McEver, K. L. Moore, R. D. Cummings, *J. Biol. Chem.*, **1995**, 270, 11025.

160. P. R. Taylor, S. Zamze, R. J. Stillion, S. Y. C. Wong, S. Gordon, L. Martinez-Pomares, *Proc. Natl. Acad. Sci. USA*, **2004**, 101. 1963.
161. L. Martinez-Pomares, J. A. Mahoney, R. Kaposzta, S. A. Linehan, P. D. Stahl, S. Gordon, *J. Biol. Chem.*, **1998**, 273, 23376.

Chapter II - The desymmetrization of trehalose

2.1 Introduction

Analogues of D-trehalose have found uses in a variety of applications. Trehalose is naturally synthesized by bacteria and invertebrates as a storage sugar and a source of carbon.[1] During metabolism, the sugar is broken down to glucose by the enzyme trehalase. Trehalose analogues that inhibit trehalase can prevent this vital step and may be potent fungicides or antibiotics.[2] In the pathogen *Mycobacterium tuberculosis*, the Antigen 85 enzyme family converts trehalose into its mycolic acid ester derivative which is subsequently incorporated into the cell wall.[3] This lipid wall is vital for bacterium-host interactions during infection, and furthermore, is a barrier for potential drugs.[4] Analogues of trehalose that inhibit Antigen 85 can disrupt cell wall synthesis and may be novel anti-tubercular drugs.[5] Aside from these biological applications, the α,α linkage gives trehalose a unique structure that facilitates its use in nanotechnology[6] and as a novel scaffold for asymmetric reactions.[7]

For all of these applications, access to non-symmetrical trehalose analogues would be particularly valuable. The glycosylation of two glucose units is a conceivable strategy, but would require the simultaneous formation of two α anomeric linkages. While methods to achieve this transformation have been reported, there are often limitations or drawbacks. For example, the dehydrative glycosylation of 2-iodoglycals can give disaccharides with the α,α linkage, but only for sugars in the manno or talo configuration.[8] The glycosylation of methyl ketosides can also give the desired α,α linkage in high yield, but only furnishes 1-methyl trehalose analogues as the end-product, which may not be suitable for all applications.[9] One interesting approach utilizes acylfurans as the starting material to give the desired α,α linkage via a diastereoselective palladium mediated glycosylation, but so far this route has only been demonstrated for manno and 2,3-deoxy analogues of trehalose.[10] A promising method uses intramolecular aglycone delivery where the two monosaccharides are covalently tethered together in an orientation that, upon activation, favours formation of the desired linkage.[11] While this methodology may be the most generally applicable, several time-consuming manipulations are required to obtain the precursors for the critical glycosylation step.

The direct modification of trehalose would be a much more pragmatic approach that circumvents the need for difficult glycosylations. Several methods for the large scale enzymatic conversion of starch or maltose to trehalose have been developed,[12] and consequently, trehalose is now available at low cost. Due to this cheap availability and the unique preservative properties of trehalose,[13] the sugar has found use as a stabilizer in the food industry, as a component in cosmetics, and in pharmaceuticals as a cryoprotectant for vaccines and organs during transplantation.[12] Similarly, trehalose would be a viable starting material for organic synthesis, not just for laboratory scale experiments, but potentially also for the industrial production of trehalose-based products.

A drawback to this approach is the C_2 rotational symmetry of the trehalose molecule, which leads to a lack of differentiation between the two glucose rings. Hence, attempts to make non-symmetrical analogues, for instance by mono-functionalization, would lead to a statistical mixture of products and low yields. Despite this downside, the simple nature of this transformation has led to many non-symmetrical trehalose analogues being synthesized with this method.[5-7,14] Single modifications at the 2, 3, 4 and 6 positions have been reported proving the viability of this approach.[15] Methodology for the high yielding mono-functionalization of trehalose would be highly desirable. The mono-functionalized analogues that would result could be useful intermediates for the synthesis of more complicated non-symmetrical trehalose compounds.

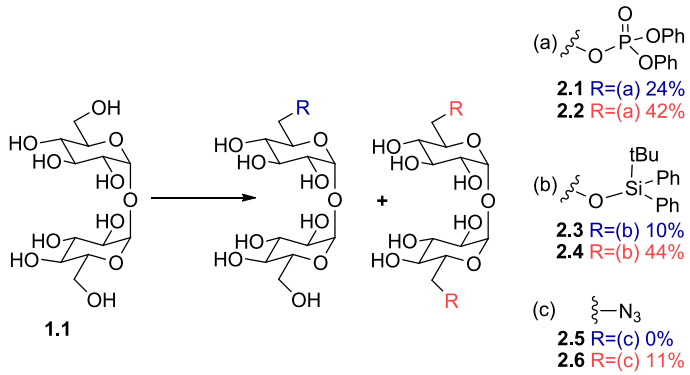
The mono-functionalization of a polyol such as trehalose can be envisioned as two separate strategic levels of selectivity. Regioselective modification of the primary hydroxyls over secondary alcohols is, of course, well established in carbohydrate chemistry.[16] The second mode of selectivity requires differentiation of the symmetry units within the C_2 symmetric trehalose. For these *desymmetrization* reactions, the source of the selectivity must arise from physical differences between the two hydroxyls rather than inherent chemical differences, such as electronic properties, that would be identical for a symmetrical molecule. For example, a well-established method for the desymmetrization of symmetric diols utilizes the difference in solubility of the sodium salts of the diol.[17] Treatment with sodium methoxide gives the monosodium alkoxide of the diol which, due to its low solubility, precipitates from the solution and prevents further modification. An alternative method for

modification with halides,[18] or tosyl groups[19] uses silver salts that differentiate the two alcohols through intramolecular chelation of one hydroxyl, leaving the other free to react. Regioselective access to the primary alcohols of trehalose simplifies the overall problem to that of a symmetric diol and methods based on physical selectivity could allow successful desymmetrization.

In this chapter, the mono-functionalization of trehalose is explored. Data is presented to show that these seemingly simple reactions in fact lead to very poor yields for the desired mono-functionalized product. This anomaly is investigated with microreactor technology that furnishes detailed kinetic information. From this data, the poor yields for trehalose mono-functionalizations witnessed in the literature can be explained with a physical model based on solubility. Armed with this knowledge, a route is demonstrated that exploits these physical effects to maximise the yield of desymmetrization.

2.2 Initial explorations into mono-functionalization reactions

At first sight, the synthesis of non-symmetrical trehalose analogues differentiated at the 6 position would seem like a facile undertaking. Regioselective access to the primary alcohols would circumvent the need for protecting groups and mono-functionalization would give an expected 50% statistical yield for the non-symmetrical analogue, which while modest, would give access to enough material for further steps. With this in mind, simple regioselective mono-functionalizations were attempted on trehalose (Scheme 2.1). Three different reactions were performed: (a) a phosphorylation with diphenylchlorophosphate (DPCP), (b) a TBDPS protection, and (c) a Mitsunobu reaction with hydrazoic acid for instalment of an azide group. In all cases, the yield of the desired mono-functionalized compound was unexpectedly low, with the doubly modified adducts being preferentially formed. For the phosphorylation, the desired product **2.1** was only formed in 24% yield whereas **2.2** was produced in 42%. This theme was exaggerated for the silylation, where the yield for **2.3** was only 10% while **2.4** was significantly higher at 44%. The desired product **2.5** was not even detected for the Mitsunobu reaction, which solely produced the diazide **2.6** in 10% yield.



Scheme 2.1 *Attempted monofunctionalization of trehalose. All reactions were performed with 1eq of reagent and were given sufficient time to allow the system to reach the end-point. Conditions: (a) Diphenylchlorophosphate, Py, RT; (b) tert-butyldiphenylchlorosilane, Im, DMF, RT; (c) HN_3, DIAD, PPh_3, 1,4-dioxane, RT.*

Such poor mono-functionalization yields are consistent with previous literature reports where trehalose is used as a starting material. For example, the halogenation of the primary hydroxyls with *N*-halosuccinimides gave 13%, 37% and 21% for the desired mono-functionalized chloro, bromo and iodo compounds respectively.[20] Similarly, an attempted reaction with tosyl chloride gave only the doubly modified product.[21] These consistent poor yields, under a variety of different reaction conditions, highlighted a fundamental inefficiency in the modification of trehalose.

2.3 Kinetics of consecutive reactions

The kinetics of mono-functionalizations can be modelled as two consecutive second order reactions. The corresponding rate equations can be solved computationally (or analytically under certain conditions[22]) to relate the product distribution to the reaction coordinate and rate constants, allowing better understanding of the reaction at a quantitative level.

$$\mathrm{Tre} + \mathrm{R} \xrightarrow{k_1} \mathrm{Mono} + \mathrm{R} \xrightarrow{k_2} \mathrm{Di}$$

$$\frac{d[\mathrm{Tre}]}{dt} = -k_1[\mathrm{Tre}][\mathrm{R}] \qquad (1)$$

$$\frac{d[\mathrm{Mono}]}{dt} = k_1[\mathrm{Tre}][\mathrm{R}] - k_2[\mathrm{Mono}][\mathrm{R}] \qquad (2)$$

$$\frac{d[\mathrm{Di}]}{dt} = k_2[\mathrm{Mono}][\mathrm{R}] \qquad (3)$$

$$\frac{d[\mathrm{R}]}{dt} = -k_1[\mathrm{Tre}][\mathrm{R}] - k_2[\mathrm{Mono}][\mathrm{R}] \qquad (4)$$

For a symmetric diol where the two hydroxyls are spatially and electronically separated from each other, one would expect the modification of the first alcohol to not perturb the reactivity of the other. Under these conditions, $k_1 = 2k_2$ since the relative rates are dependent solely on the number of available reactive groups at each reaction step. For 1 equivalent of reagent R, this gives a 50% theoretical yield of the mono-functionalized product at the reaction end point. Under regioselective reaction conditions, the situation for trehalose would approximate a symmetric diol and mono-functionalization yields of 50% would be expected. A simplified pseudo first order analysis shows that even if the second reaction is faster than the statistical expectation so that $k_1 = k_2$, the expected yield would still be a reasonable 37%.[23] For trehalose, the observed yields have been significantly lower implying that $k_{2a} > k_{1a}$.[*] Hence, the trehalose mono-functionalizations were particularly anomalous and warranted further investigation.

We hypothesized that these non statistical yields were due to the poor solubility of trehalose in organic solvents rather than any intrinsic chemical reactivity problems. In our hands, 1g of trehalose has required at least 100ml of pyridine for full solubilisation and was largely insoluble in other organic solvents such as DMF, methanol or DCM. This low solubility leads to an inherently slow first reaction step. The mono-functionalized trehalose adduct that results is considerably more soluble due to the presence of the non-polar functional group. Hence, the second step occurs at a faster rate than the first, and in this scenario $k_{2a} > k_{1a}$ resulting in low mono-functionalization yields. With this in mind, we sought to test the validity of this hypothesis and to devise synthetic strategies that bypass the problem.

2.4 Desymmetrizations under continuous flow

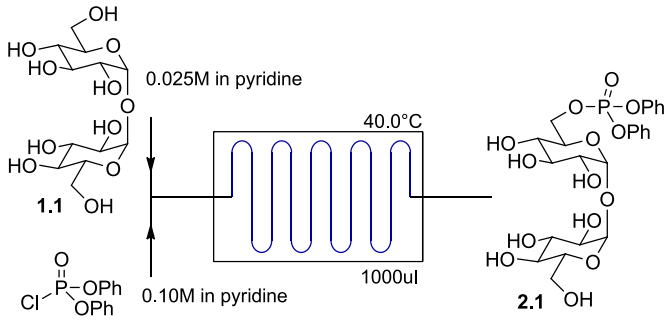
Flow chemistry is an emerging technology that offers many advantages over traditional batch reactions.[24] Microreactors with fast heat transfer properties operating in the laminar flow regime allow strict temperature and temporal control. Furthermore the time-scale of mixing can also be precisely managed through methods such as chaotic advection (where rapid reactant mixing is achieved by encouraging turbulent

[*] The rate constants k_{1a} and k_{2a} refer to the *apparent* rate constants. Due to the poor solubility of the trehalose analogues, the actual concentration of these compounds in solution cannot be easily determined. Thus, while the actual rate constants cannot be measured, it is convenient in the context of these discussions to express the reductions in rate as apparent rate constants.

flow) or radial diffusion (where there is slower diffusion limited mixing of reagents in a laminar flow regime).[24] This precise control over reaction conditions often results in improved yields and selectivity compared to batch reactions. Automation can allow facile reaction optimization and on-line monitoring can furnish detailed quantitative data over a large area of reaction space.[25] A distinct advantage of flow chemistry is the possibility of reliable scale-up, because previously optimized conditions can simply be run over a longer period of time to increase quantities.

Flow chemistry is particularly advantageous for desymmetrization reactions. Slow reactions are often sped up by using excess reagent, but the optimal end point can be difficult to gauge especially when requiring visual estimation of product distributions by TLC. In view of this, trehalose mono-functionalizations were studied under continuous flow in the hope that reliable and scalable conditions could be found.

A primary aim in this thesis is to study the biological relevance of the sugar trehalose-6-phosphate (Tre-6-P), and so naturally, this was chosen as our primary synthetic target. Therefore the phosphorylation of trehalose with DPCP to give **2.1**, a precursor to Tre-6-P, was studied in a microreactor. Trehalose was reacted with DPCP in pyridine on a 1.0ml microreactor chip over a range of different conditions (Scheme 2.2). To find optimal conditions, reactions were initially conducted on a 250µl scale and the crude reaction mixtures were per-*O*-acetylated immediately as a quench and analysed by C_{18} HPLC to elucidate the product distributions (see section 2.8.3, page 51). Various temperatures were screened to ensure a reasonable rate of reactivity that was suitable for flow conditions. It was found that temperatures in excess of 100˚C resulted in decomposition of trehalose, so were avoided. Reactions at 80˚C had lower regioselectivity, so were also avoided. The most favourable conditions were at 40˚C with 4 equivalents of DPCP which gave reaction times of between 1-10 minutes; ideally suited to flow chemistry.



Scheme 2.2. *Reaction conditions for trehalose phosphorylation in a microreactor. Reactions were run over a residence time of between 1-10mins and analysed by C_{18} HPLC.*

These conditions were used to screen the reaction over a range of residence times to evaluate the product distribution as a function of the reaction coordinate (Figure 2.1). Non-linear regression analysis was used to fit equations (1)-(4) to these data and revealed that, under these conditions, k_{1a} = 0.157 $mol^{-1}dm^3s^{-1}$ and k_{2a} = 0.133 $mol^{-1}dm^3s^{-1}$.

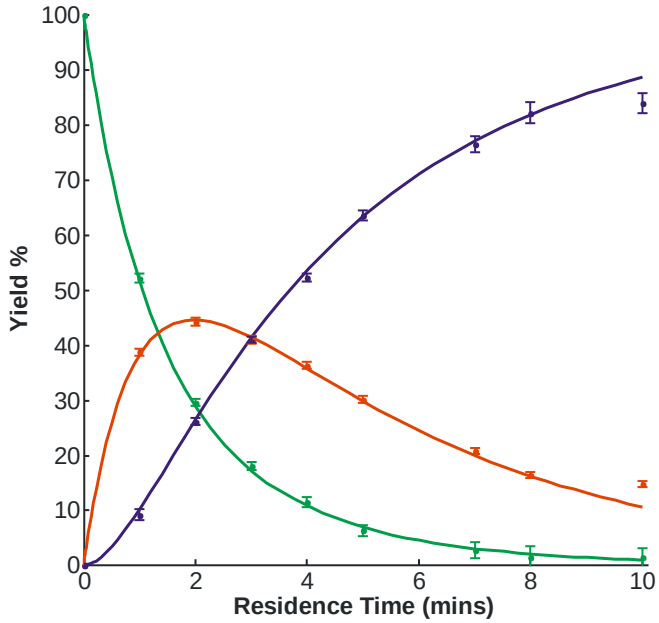


Figure 2.1. *Graph showing product distribution versus residence time for trehalose phosphorylation with DPCP in py at 40°C. Data fitted to eqn (1)-(4), r^2 = 0.986. Legend: Green (trehalose), red (2.1), blue (2.2). Error bars show standard error of the mean.*

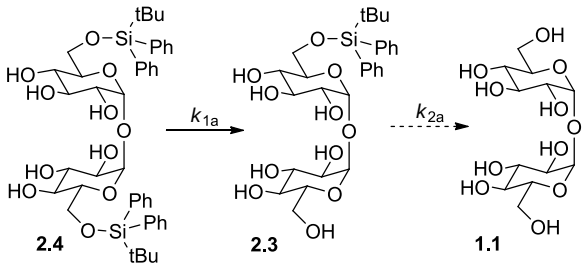
The data have revealed that a maximum yield of 44% for compound **2.1** is attainable with a 2 minute residence time. The marked improvement in yield over the corresponding batch reaction may be at least in part due to the higher volume of solvent used. Flow reactors call for fully solubilised reagents which, for trehalose, necessitates a dilute 25mM solution in pyridine. Compared to the batch reaction where the trehalose is not fully dissolved in pyridine, the first modification would be

faster in a fully solubilised system leading to a higher apparent k_1 and a yield that better resembles the statistical expectation. As such, this observation provides support to our solubility hypothesis concerning trehalose mono-functionalizations.

By running the system for a longer period while maintaining the residence time at 2 minutes, the test reactions could be reliably scaled up. A reaction was performed on a 2 gram scale which resulted in a 39% yield after purification. While this improvement in yield is clearly beneficial, the requirement for high volumes of solvent is an obvious drawback. Methods that minimize solvent usage would be preferable.[26] Moreover, rather than simply using excess solvent to sidestep the poor solubility of trehalose, a method that could *exploit* this property would be far more satisfying and could potentially deliver even greater yields for mono-functionalizations.

2.5 A reverse modification strategy

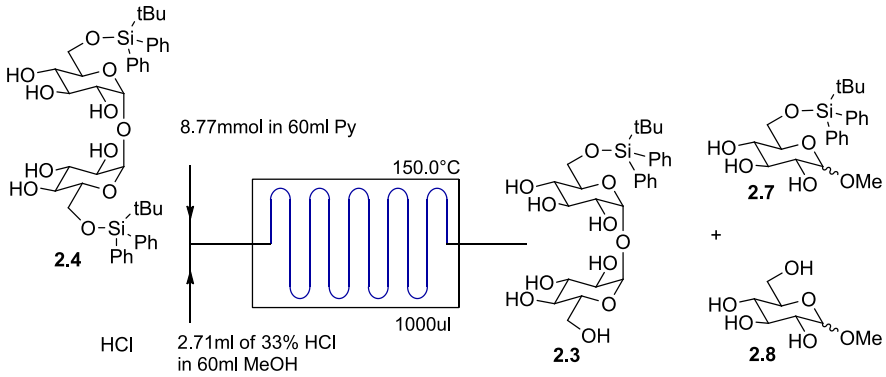
To achieve better yields for desymmetrizations, we envisaged a system where instead of attempting to mono-functionalize trehalose, the reverse process of removing functionality from a doubly modified trehalose adduct, or mono-*de*functionalization, could be conducted (Scheme 2.3). Studies were focused on **2.4** because the silyl group can be easily removed under a number of mild conditions.[27] In this scenario, removal of a silyl group from **2.4** gives **2.3**, which crucially, is less soluble than the starting carbohydrate. Hence, the second deprotection to give trehalose occurs at a slower rate. Now $k_{1a} > k_{2a}$ and higher yields for the mono-functionalized adduct can be attained.



Scheme 2.3. *General concept for the reverse modification strategy.*

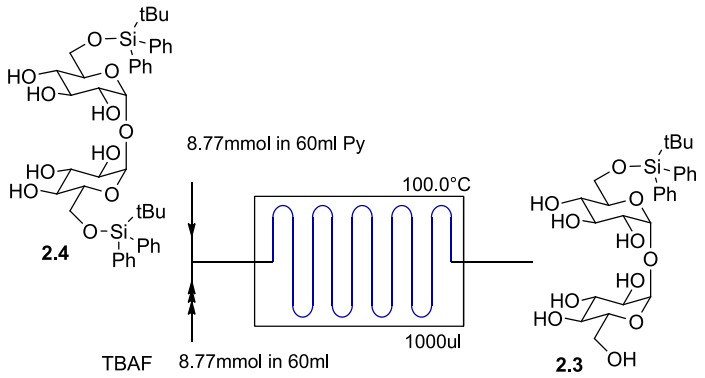
As before, we studied these reactions under flow conditions in the hope that we could get reliable conditions for desymmetrizations. Reactions were performed on a 250µl scale and analysed by C_{18} HPLC after per-*O*-acetylation (section 2.8.3 page 51). Initially, the deprotection of **2.4** was attempted under acidic conditions. However, the

sluggish reaction required elevated temperatures of 150°C which resulted in hydrolysis of the glycosidic linkage giving compounds **2.7** and **2.8**.



Scheme 2.4. *Attempted deprotection of a silyl group in a microreactor. Reactions were run over a residence time of between 2-20mins and analysed by C_{18} HPLC.*

Instead, the deprotection was conducted using excess TBAF. The reaction was still slow, so was run with residence times varying from 10 to 70 minutes. The reaction was investigated in two different solvent systems: (i) 1 methanol : 1 pyridine representing high solubility conditions where both **2.3** and **2.4** show reasonable solubility, and (ii) 1 THF : 1 petrol representing low solubility conditions where the solubility of **2.3** is significantly reduced compared to **2.4**.



Scheme 2.5. *Reaction conditions for desilylation in a microreactor. Reactions were performed in either 1 MeOH : 1 Py, or 1 THF : 1 petrol, and were run over a residence time of between 10-70mins and analysed by C_{18} HPLC. The flow rate of TBAF was 3× that of 2.4.*

Under the high solubility conditions, a maximum yield of 41% was achieved with a 20 minute residence time (Figure 2.2). As before, non-linear regression analysis was used to obtain the rate constants and it was found that $k_{1a} = 21.9 \times 10^{-3}$ mol^{-1}dm^{3}s^{-1} while $k_{2a} = 17.9 \times 10^{-3}$ mol^{-1}dm^{3}s^{-1}.

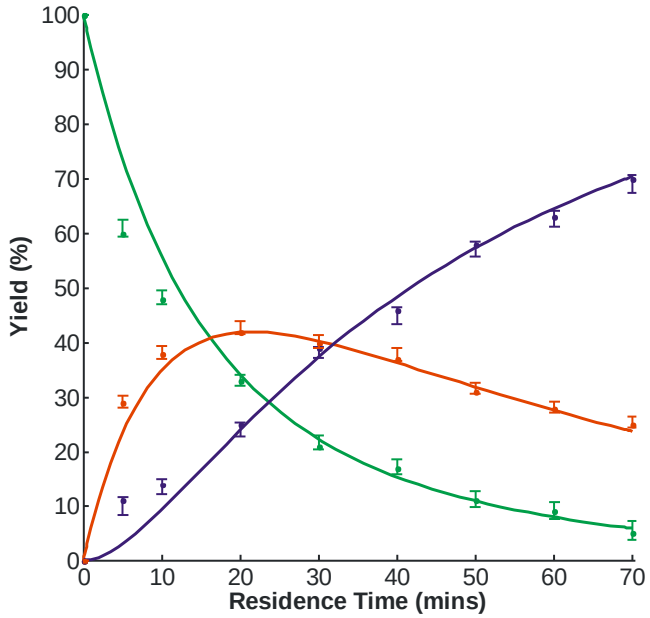


Figure 2.2. *Graph showing product distribution versus residence time for the desilylation of 2.4 with TBAF in 1 MeOH : 1 Py at 40°C. Data fitted to eqn (1)-(4), r^2 = 0.965. Legend: Green (2.4), red (2.3), blue (trehalose). Error bars show standard error of the mean.*

Under low solubility conditions the maximum yield was improved to 57% (Figure 2.3). Analysis of the data showed that k_{1a} = 21.6 × 10^{-3} $mol^{-1}dm^3s^{-1}$ and was essentially unaltered by the solvent. In contrast, k_{2a} = 7.67 × 10^{-3} $mol^{-1}dm^3s^{-1}$ and was substantially lower than under the high solubility conditions.

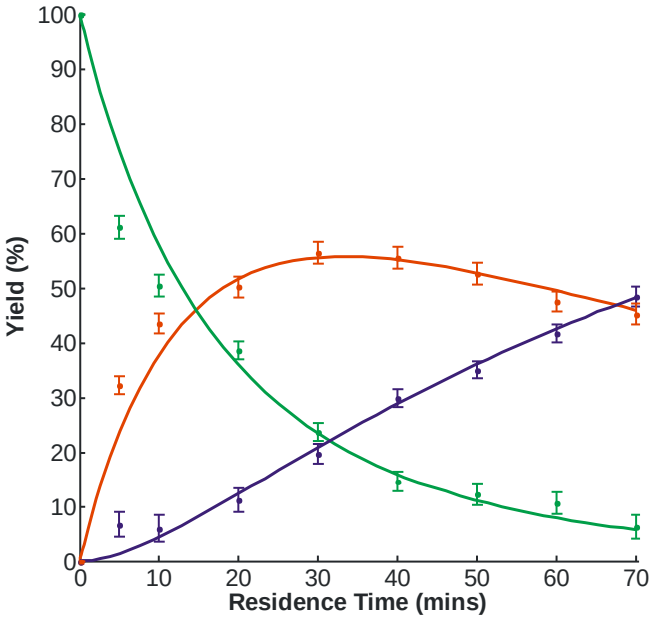


Figure 2.3. *Graph showing product distribution versus residence time for the desilylation of 2.4 with TBAF in 1 THF : 1 petrol at 40°C. Data fitted to eqn (1)-(4), r^2 = 0.969. Legend: Green (2.4), red (2.3), blue (trehalose). Error bars show standard error of the mean.*

Thus, the solvation properties of trehalose analogues has provided an effective means to manipulate k_{2a}, which in turn has led to improved yields for desymmetrization

(Figure 2.4). This is analogous to the desymmetrization of symmetric diols with sodium methoxide, where solubility was a critical parameter in improving yields (section 2.1, page 33).[17]

Unfortunately, the flow conditions developed here could not be used on a significantly large scale. Ironically, the poor solubility of **2.3** which facilitates the improved yields also results in precipitation and subsequent blocking of the chip. The clogging of microreactors with precipitated material is still an outstanding problem in flow chemistry. While several potential solutions have been developed, such as ultrasonication,[28,29] liquid/gas slug flow,[29] liquid/liquid slug flow,[30] droplet flow[31] and sheath flow,[32] a generally applicable method is still needed and further effort needs to be directed in this area.

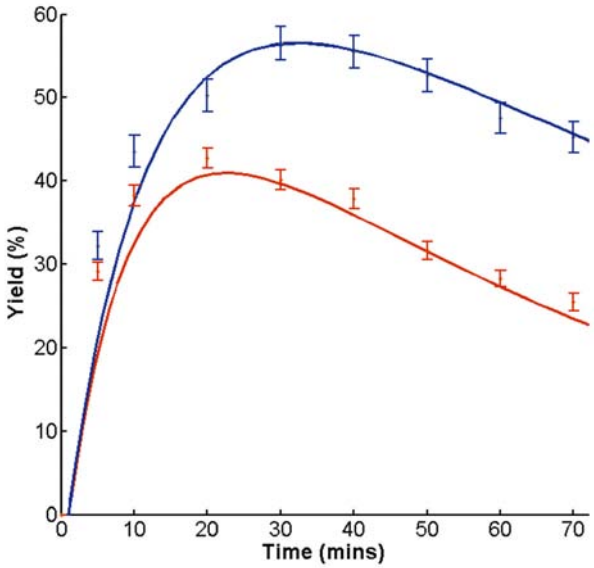
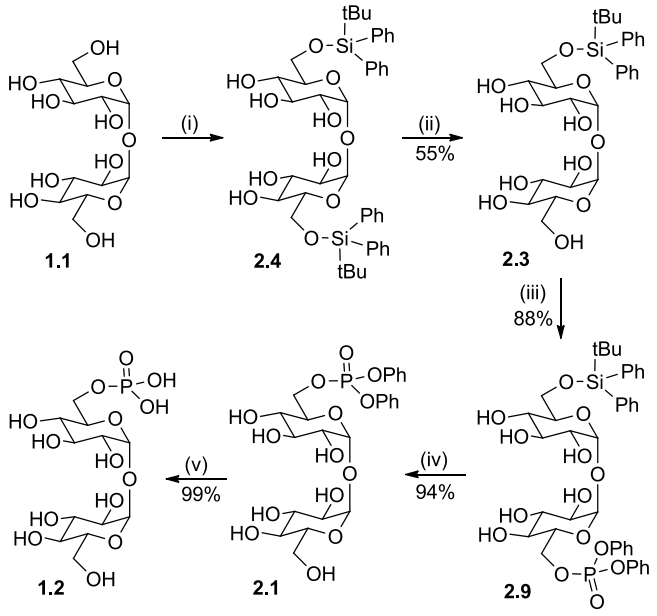


Figure 2.4. *Graph comparing the formation of 2.3 under different solvent conditions. Legend: Red (1 MeOH : 1 py), blue (1 THF : 1 petrol). Error bars show standard error of the mean.*

2.6 Synthesis of trehalose-6-phosphate

Despite this shortcoming, the conditions and kinetic data derived from the flow conditions could still be used to create a more generally applicable route for a large scale synthesis in batch. The conditions for conversion of **2.4** to **2.3** were replicated in batch mode where the vital low solubility conditions were employed. We envisaged that **2.3** could be a versatile starting material for the synthesis of other non-symmetrical trehalose analogues, since the compound can be easily made on a large scale giving selective access to the remaining primary alcohol and because the silyl

group can be removed under mild conditions towards the end of the synthesis. To illustrate this, we designed a route for the synthesis of Tre-6-P using the efficient desymmetrization of trehalose to **2.3** to achieve high yields (Scheme 2.6).



Scheme 2.6. *Reaction scheme to exploit for solubility of trehalose for monofunctionalization. Conditions: (i)* tert*-butyldiphenylchlorosilane, Py, 14h; (ii) AcCl, MeOH:Et_2O 1:3 18 h; (iii) DPCP, Py, 15h; (iv) AcCl, MeOH, 18h; (v) H_2, PtO_2, 1% AcOH, EtOH, 14h.*

Regioselective silylation of trehalose smoothly gave **2.4** which was used without further purification. Acidic conditions were chosen for the mono-desilylation of **2.4** since the rate of reaction was slower than with TBAF, therefore the reaction was easier to monitor and control. Low solubility conditions were achieved using 1 methanol : 3 diethyl ether in which both **2.3** and **2.4** are poorly dissolved, and acid was generated *in situ* by addition of acetyl chloride. Pleasingly, this method replicated the flow conditions and through regular monitoring of the reaction by TLC, a yield of 55% over 2 steps was achieved for **2.3**. With selective access to the remaining primary alcohol, the subsequent phosphorylation with DPCP gave **2.9** in a high yield of 88%. Silyl deprotection under mild acid conditions went smoothly producing **2.1** in 94% yield. Finally, phosphate deprotection was effected in near quantitative yield using hydrogen over a PtO_2 catalyst to furnish Tre-6-P in 45% over 5 steps. In contrast, the direct synthesis of **2.1** from trehalose only occurs in 24% yield. Hence, we obtained phosphorylation yields similar to those achieved in flow using a more convenient batch synthesis.

The facile and scalable synthesis of **2.3** can readily be incorporated into the synthetic routes of other non-symmetrical trehalose analogues. For Tre-6-P, an arguably more convenient two-step route via **2.1** may be favoured despite the lower yields. However, shorter routes may not be available for other analogues such as the azide **2.5**. In these cases, and for other more structurally complicated analogues, our reverse modification strategy could lead to substantial improvements in yields.

2.7 Conclusions

In organic chemistry, reactions are often performed in biphasic mixtures involving reagents in a solid phase and a solution phase. Manipulating the equilibrium between the two phases is a common method for controlling such reactions. For example, this "phase control" is used in enzymatic reactions where judicious choice of solvent can selectively precipitate the product, thus pushing the reaction in the forward direction.[33] Attrition grinding deracemizations represent another example of phase controlled reactions.[34] Here, a racemic mixture of a compound in equilibrium with a racemic mixture of its distinct enantiopure crystals is stirred with mechanical grinding. This forces a continual cycle of dissolution and re-crystallization which, through Ostwald ripening, eventually results in a single enantiopure solid phase. Understanding the physics of this attrition grinding process[35] has allowed for the enantiopure synthesis of the non-steroidal anti-inflammatory drug Naproxen.[36] Analogously, by understanding the kinetics of trehalose desymmetrizations, we have optimized the synthesis of biologically relevant analogues such as Tre-6-P.

Moreover, our data can explain trends for trehalose desymmetrizations published in the literature over the last 40 years. Anomalously low yields for mono-functionalizations on fully deprotected trehalose can now be rationalized in terms of solubility of the starting sugar.[20,21] On the other hand, desymmetrizations using partially protected trehalose analogues proceed with much better yield. For instance, the synthesis of a tri-*O*-cyclohexylidene analogue occurs in 42% yield.[37] Selective mono-debenzylidenation of 4,6,4',6'-di-*O*-benzylidene-2,2',3-3'-tetrabenzyl analogue proceeds in 47% yield.[38] Such high yields are also consistent with our model, because for protected trehalose analogues, the differences in solubility before and after modification are much less marked. Hence, more statistical yields result.

Using secondary hydroxyl protection, or dilute reaction solutions are possible workarounds to give statistical desymmetrization yields. Alternatively, the poor solubility can be exploited to improve yields beyond this barrier. As we have shown, solubility can have a large influence on the chemistry of trehalose and this principle can likely be extended to other pseudo-symmetric disaccharides. For these sugars, due consideration should be given to solubility when designing synthetic routes.

2.8 Experimental

2.8.1 General considerations

Proton nuclear magnetic resonance spectra (^{1}H NMR) were recorded on a Bruker AVC500 (500MHz), a Bruker AV400 (400 MHz), or a Bruker DPX 400 (400MHz), spectrometer. Carbon nuclear magnetic resonance spectra (^{13}C NMR) were recorded on a Bruker AVC500 (125.6 MHz), or a Bruker AV400 (100.7 MHz) and are proton decoupled. Spectra were assigned using COSY, DEPT-135, HMQC, HSQC, and HMBC if required. Phosphorus nuclear magnetic resonance spectra (^{31}P NMR) were recorded on a Bruker AV400 (162 MHz). All chemical shifts are quoted on the δ scale in ppm using residual solvent as an internal standard.

Low resolution mass spectra were recorded on a Micromass Platform 1 spectrometer using electrospray ionization (ES), or on a Bruker Daltronic MicroTOF spectrometrer. High resolution mass spectra were recorded on a Bruker Daltronic MicroTOF spectrometer. m/z values are reported in Daltons. Infrared spectra (FT-IR) were recorded on a Bruker Tensor 27 Fourier Transform spectrophotometer using KBr discs, thin films on NaCl plates, or using a total internal reflectance (TIR) module as stated. Absorption maxima are reported in wavenumbers (cm^{-1}). Only signals representing functional groups are reported; C-H absorptions and the fingerprint region are not listed.

Melting points were recorded on a Kofler hot block. Optical rotations were measured on a Perkin-Elmer 241 polarimeter at 589nm (Na D-line) with a path length of 1.0 dm and are reported in units of deg $dm^{-1}cm^{3}g^{-1}$. Concentrations are given in g/100 mL.

Thin Layer Chromatography (TLC) was carried out using Merck aluminum backed sheets coated with Kieselgel $60F_{254}$ silica gel. Visualization of the sheets was

achieved using a UV lamp (λ_{max} = 254 or 365 nm) and either ammonium molybdate (5% in 2M H_2SO_4), potassium permanganate (5% in 1M NaOH), sulfuric acid (0.2M in 1 MeOH : 1 H_2O), or ninhydrin (1% in ethanol). Silica gel chromatography was carried out using Fluka Kieselgel 60 220-240 mesh silica. Automated silica gel chromatography was performed on a Biotage SP4 purification system using KP-Sil SNAP cartridges.

High Performance Liquid Chromatography was conducted on a Dionex HPLC system at ambient temperature, with an in-line UV absorbance detector, and/or an in-line or parallel to the main flow path Varian PLS400 Evaporative Light Scattering (ELS) detector as stated.

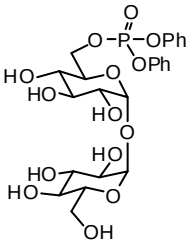
Anhydrous THF, DCM, toluene, and diethyl ether were dried through a column of alumina. Other anhydrous solvents were purchased from Fluka, and stored under argon over molecular sieves. All other solvents were used as supplied (analytical or HPLC grade). “Petrol” refers to the fraction of light petroleum ether boiling in the range 40-60°C. “Brine” refers to a saturated aqueous solution of sodium chloride.

Dowex® cation exchange resins (H^+ form) were conditioned by washing with methanol, water, 1M HCl, and water until the filtrate was neutral. Amberlite™ anion exchange resins (OH^- form) were conditioned by washing with methanol, water, concentrated aqueous ammonia, and water until the filtrate was neutral.

2.8.2 Batch trehalose mono-functionalizations

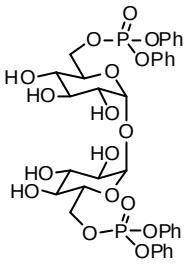
Phosphorylation

To a suspension of D-trehalose (7.50 g, 21.9 mmol, 1 eq) in anhydrous pyridine (100 mL) was added dropwise DPCP (4.54 mL, 21.9 mmol, 1 eq). TLC (1 water : 4 isopropanol : 4 ethyl acetate) after 18 hours showed the presence of two products. The reaction was quenched with methanol (10 mL) and the reaction mixture concentrated *in vacuo*. The residue was co-evaporated with toluene to remove pyridine (x3). Silica gel chromatography (1 water : 3 isopropanol : 13 ethyl acetate) allowed separation of the two products. Lyophilization yielded **2.1** (3.02g, 24%) and **2.2** (7.41g, 42%) as white amorphous solids.



6-*O*-(diphenoxyphosphoryl)-D -trehalose (2.1)

R_f 0.68 (1 water : 4 isopropanol : 4 ethyl acetate), $[\alpha]_D^{22}$ +63.9 (c = 1.0, MeOH); ^{1}H NMR (500MHz, MeOD) δ ppm 3.34 (1H, t, $J_{H3'\text{-}H4'}$ 9.1Hz, $J_{H4'\text{-}H5'}$ 9.1Hz, H4'), 3.38 (1H, t, $J_{H3\text{-}H4}$ 9.1Hz, $J_{H4\text{-}H5}$ 9.1Hz, H4), 3.43 (1H, dd, $J_{H2'\text{-}H3'}$ 9.8Hz, $J_{H1'\text{-}H2'}$ 3.5Hz, H2'), 3.48 (1H, dd, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.70 (1H, dd, $J_{H6'a\text{-}H6'b}$ 12.0Hz, $J_{H5'\text{-}H6'b}$ 5.4Hz, H6'b), 3.80 (1H, t, $J_{H2'\text{-}H3'}$ 9.5Hz, $J_{H3'\text{-}H4'}$ 9.5Hz, H3'), 3.81 (1H, t, $J_{H2\text{-}H3}$ 9.1Hz, $J_{H3\text{-}H4}$ 9.1Hz, H3), 3.80-3.83 (1H, m, H6'a), 3.84 (1H, m, $J_{H5'\text{-}H6'b}$ 4.1Hz, $J_{H5'\text{-}H6'a}$ 2.2Hz, H5'), 4.09 (1H, dt, $J_{H4\text{-}H5}$ 10.1Hz, $J_{H5\text{-}H6a}$ 2.1Hz, $J_{H5\text{-}H6b}$ 2.1Hz, H5), 4.48 (1H, ddd, $J_{H6a\text{-}H6b}$ 11.5Hz, $J_{H6b\text{-}31P}$ 7.1Hz, $J_{H5\text{-}H6b}$ 3.5Hz, H6b), 4.55 (1H, ddd, $J_{H6a\text{-}H6b}$ 11.5Hz, $J_{H6a\text{-}31P}$ 6.8Hz, $J_{H5\text{-}H6a}$ 1.9Hz, H6a), 5.09 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.10 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 7.21-7.31 (3H, m, ArH_{ortho}, ArH_{para}), 7.40-7.43 (2H, m, ArH_{meta}); ^{13}C NMR (126MHz, MeOD) δ ppm 62.6 (1C, C6'), 69.8 (1C, d, $J_{C6\text{-}31P}$ 6.7Hz, C6), 71.2 (1C, C4'), 71.9 (1C, C4), 72.0 (1C, d, $J_{C5\text{-}31P}$ 6.7Hz, C5), 73.0 (1C, C5'), 73.2 (1C, C2'), 73.9 (1C, C2), 74.4 (1C, C3'), 74.6 (1C, C3), 95.2 (1C, C1'), 95.3 (1C, C1), 121.4 (2C, d, $J_{C\text{-}31P}$ 4.8Hz, ArC_{ortho}), 126.8 (1C, ArC_{para}), 131.1 (2C, ArC_{meta}), 151.9 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, ArC_{ipso}), 151.9 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, ArC_{ipso}); ^{31}P{^{1}H} NMR (162MHz, MeOD) δ ppm -12.0 (1P, $P(O)(OPh)_2$); FT-IR (KBr disc) υ 1287 (P=O), 3271 br (OH); HRMS m/z (ES^+) 596.1357 $[M + Na]^+$ required (596.1344); elemental C (49.87%) H (5.19%), required C (50.18%) H (5.44%).



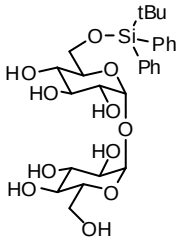
6,6'-*O*-di(diphenoxyphosphoryl)-D-trehalose (2.2)

R_f 0.71 (1 water : 4 isopropanol : 4 ethyl acetate), $[\alpha]_D^{22}$ +63.8 (c = 1.0, MeOH); ^{1}H NMR (500MHz, MeOD) δ ppm 3.37 (2H, t, $J_{H3\text{-}H4}$ 9.1Hz, $J_{H4\text{-}H5}$ 9.1Hz, H4), 3.40 (2H, dd, $J_{H2\text{-}H3}$ 9.9Hz, $J_{H1\text{-}H2}$ 3.6Hz, H2), 3.79 (2H, t, $J_{H2\text{-}H3}$ 9.3Hz, $J_{H3\text{-}H4}$ 9.3Hz, H3), 4.07 (2H, ddd, $J_{H4\text{-}H5}$ 9.9Hz, $J_{H5\text{-}H6b}$ 4.7Hz, $J_{H5\text{-}H6a}$ 2.0Hz, H5), 4.48 (2H, ddd, $J_{H6a\text{-}H6b}$ 11.1Hz, $J_{H6b\text{-}31P}$ 7.6Hz, $J_{H5\text{-}H6b}$ 4.7Hz, H6b), 4.55 (2H, ddd, $J_{H6\text{-}H6'}$ 11.1Hz, $J_{H6a\text{-}31P}$ 6.6Hz, $J_{H5\text{-}H6a}$ 1.9Hz, H6a), 5.00 (2H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 7.21-7.29 (6H, m, ArH_{ortho}, ArH_{para}), 7.37-7.43 (4H, m, ArH_{meta}); ^{13}C NMR (126MHz, MeOD) δ ppm 69.8 (2C, d, $J_{C6\text{-}31P}$ 6.7Hz, C6), 51.3 (2C, C4), 72.1 (2C, d, $J_{C5\text{-}31P}$ 6.7Hz, C5), 73.0 (2C, C2), 74.5 (2C, C3,), 95.3 (2C, C1), 121.2 (8C, d, $J_{C\text{-}31P}$ 4.8Hz, ArC_{ortho}), 126.9 (4C, ArC_{para}), 131.1 (8C, d, $J_{C\text{-}31P}$ 2.9Hz, ArC_{meta}), 151.9 (2C, d, $J_{C\text{-}31P}$ 1.9Hz, ArC_{ipso}), 151.9 (2C, d, $J_{C\text{-}31P}$ 2.9Hz,

ArC_{ipso}); $^{31}P\{^{1}H\}$ NMR (162MHz, MeOD) δ ppm -11.9 (1P, $P(O)(OPh)_2$); FT-IR (KBr disc) υ 1271 (P=O), 3364 br (OH); HRMS m/z (ES^+) 829.1629 $[M + Na]^+$ required (829.1633).

Silylation

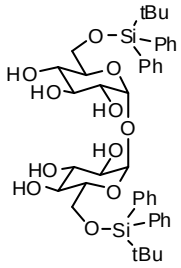
To a stirred suspension of D-trehalose (0.5g, 1.46mmol, 1eq) and imidazole (0.099g, 1.46mmol, 1eq) in dry DMF (10ml) was added *tert*-butyl diphenylchlorosilane (0.38ml, 1.46mmol, 1eq) at room temperature. After stirring for 26 hours, TLC (1 water : 2 isopropanol : 2 ethyl acetate) indicated the formation of two products. The reaction mixture was concentrated *in vacuo* and purified by silica gel chromatography (2 ethyl acetate : 1 isopropanol) to give **2.3** (85mg, 10%) and **2.4** (0.53g, 44%) as white amorphous solids after lyophilization.



6-*O*-terbutyldiphenylsilyl-D-trehalose (2.3)

R_f 0.29 (1 water : 4 isopropanol : 4 ethyl acetate), $[\alpha]_D^{18}$ +20.3 (c = 1.0, dioxane); ^{1}H NMR (500 MHz, DMSO-d_6) δ ppm 1.00 (9H, s, ^{t}Bu), 3.15 (1H, td, $J_{H3'\text{-}H4'}$ 9.3Hz, $J_{H4'\text{-}H5'}$ 9.3Hz, $J_{H4'\text{-}OH}$ 5.7Hz, H4'), 3.25 (1H, td, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, $J_{H4\text{-}OH}$ 5.0Hz, H4), 3.26 (1H, m, H2'), 3.29 (1H, dt, $J_{H2\text{-}H3}$ 10.7Hz, $J_{H1\text{-}H2}$ 4.4Hz, $J_{H2\text{-}OH}$ 4.4Hz, H2), 3.49 (1H, dt, $J_{H6'a\text{-}H6'b}$ 11.4Hz, $J_{H5'\text{-}H6'b}$ 5.5Hz, $J_{H6'b\text{-}OH}$ 5.5Hz, H6'b), 3.55-3.64 (3H, m, H3', H5', H6'a), 3.68 (1H, ddd, $J_{H3\text{-}H4}$ 9.9Hz, $J_{H3\text{-}OH}$ 4.7Hz, $J_{H2\text{-}H3}$ 1.9Hz, H3), 3.79 (1H, dd, $J_{H6a\text{-}H6b}$ 10.7Hz, $J_{H5\text{-}H6b}$ 5.4Hz, H6b), 3.83 (1H, dd, $J_{H6a\text{-}H6b}$ 10.7Hz, $J_{H5\text{-}H6a}$ 1.3Hz, H6a), 3.87 (1H, ddd, $J_{H4\text{-}H5}$ 9.9Hz, $J_{H5\text{-}H6b}$ 5.0Hz, $J_{H5\text{-}H6a}$ 1.3Hz, H5), 4.39 (1H, t, $J_{H6a\text{-}OH}$ 5.8Hz, $J_{H6b\text{-}OH}$ 5.8Hz, 6'-OH), 4.58 (1H, d, $J_{H2\text{-}OH}$ 6.3Hz, 2-OH), 4.66 (1H, d, $J_{H2'\text{-}OH}$ 6.3Hz, 2'-OH), 4.75 (1H, d, $J_{H3'\text{-}OH}$ 4.7Hz, 3'-OH), 4.78 (1H, d, $J_{H4'\text{-}OH}$ 5.4Hz, 4'-OH), 4.84 (1H, d, $J_{H3\text{-}OH}$ 5.0Hz, 3-OH), 4.86 (1H, d, $J_{H4\text{-}OH}$ 5.4Hz, 4-OH), 4.95 (1H, d, $J_{H1'\text{-}H2'}$ 3.5Hz, H1'), 4.97 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 7.40-7.46 (6H, m, ArH_{ortho}, ArH_{para}), 7.67 (4H, ad, *J* 6.9Hz, ArH_{meta}); ^{13}C NMR (126 MHz, DMSO-d_6) δ ppm 19.0 (1C, $\underline{C}(CH_3)_3$), 26.6 (1C, $C(\underline{C}H_3)_3$), 60.7 (1C, C6'), 63.4 (1C, C6), 69.9 (1C, C4), 70.1 (1C, C4'), 71.6 (1C, C2), 71.7 (1C, C2'), 72.3 (1C, C5), 72.5 (1C, C3), 73.0 (1C, C5'), 73.0 (1C, C3'), 93.0 (1C, C1), 93.1 (1C, C1'), 127.8 (4C, ArC_{ortho}), 129.6 (1C, ArC_{para}), 129.7 (1C, ArC_{para}), 133.3 (2C, ArC_{ipso}), 133.5 (2C, ArC_{ipso}), 135.1 (1C, ArC_{meta}), 135.2 (1C, ArC_{meta}); FT-IR (KBr disc) υ 3400 br (OH); m/z (ES^+) 603.2221

$[M + Na]^+$ (required 603.2232); elemental C (58.00%) H (6.97%), required C (57.91%) H (6.94%).



6,6'-*O*-diterbutyldiphenylsilyl-D-trehalose (2.4)

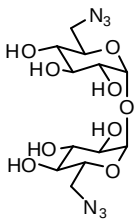
R_f 0.35 (1 water : 2 isopropanol : 2 ethyl acetate), $[\alpha]_D^{17}$ +71.5 (c = 1.0, dioxane), ^{1}H NMR (500MHz, DMSO) δ ppm 1.00 (18H, s, 2 x ^{t}Bu), 3.26 (2H, td, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, $J_{H4\text{-}OH}$ 5.4Hz, H4), 3.30 (2H, ddd, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H2\text{-}OH}$ 6.3Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.63 (2H, td, $J_{H2\text{-}H3}$ 9.1Hz, $J_{H3\text{-}H4}$ 9.1Hz, $J_{H3\text{-}OH}$ 5.0Hz, H3), 3.80 (2H, dd, $J_{H6a\text{-}H6b}$ 11.0Hz, $J_{H5\text{-}H6b}$ 5.7Hz, H6b), 3.86 (4H, m, H5, H6a), 4.62 (2H, d, $J_{H2\text{-}OH}$ 6.3Hz, 2-OH), 4.84 (2H, d, $J_{H3\text{-}OH}$ 4.7Hz, 3-OH), 4.91 (2H, d, $J_{H2\text{-}OH}$ 5.4Hz, 4-OH), 5.06 (1H, d, $J_{H1\text{-}H2}$ 3.5Hz, H1), 7.36-7.46 (12H, ArH_{ortho}, ArH_{para}), 7.70 (8H, m, ArH_{meta}); ^{13}C NMR (126MHz, DMSO) δ ppm 19.0 (2C, 2 x $\underline{C}(CH_3)_3$), 26.6 (6C, 2 x $C(\underline{C}H_3)_3$), 63.4 (2C, C6), 70.0 (2C, C4), 71.8 (2C, C2), 72.3 (2C, C5), 73.2 (2C, C3), 92.8 (2C, C1), 127.5 (4C, ArC_{ortho}), 127.7 (4C, ArC_{ortho}), 129.2 (2C, ArC_{para}), 129.6 (2C, ArC_{para}), 134.5 (4C, ArC_{ipso}), 135.2 (4C, ArC_{ipso}), 133.3 (2C, ArC_{meta}), 133.5 (2C, ArC_{meta}); FT-IR (KBr disc) υ 3350 br (OH); m/z (ES^+) 841.3410 [$M + Na^+$] (required 841.3410).

Mitsunobu reaction

Synthesis of hydrazoic acid: Hydrazoic acid was prepared by the careful addition of concentrated sulphuric acid (0.41ml, 7.7mmol, 0.5eq) to a solution of NaN_3 (1.0g, 15.4mmol, 1eq) in toluene (12ml) and water (1.5ml) at 0 ºC. The toluene layer containing the hydrazoic acid was decanted off via cannula and stored over Na_2SO_4. Titration against a standardized solution of NaOH with phenolphthalein as indicator was used to determine the concentration of HN_3 as 0.75M.

Mitsunobu reaction: Diisopropyl azodicarboxylate (DIAD, 0.057ml, 0.292mmol, 1eq) was added to a stirred solution of D-trehalose (0.10g, 0.292mmol, 1eq) and PPh_3 (0.077g, 0.292mmol, 1eq) in dry 1,4-dioxane (10ml) to form a pale orange solution. After stirring for 30 minutes at room temperature, the solution decolourized and hydrazoic acid (0.750M solution in toluene, 0.39ml, 0.292mmol, 1eq) was added dropwise. TLC (1 water : 2 isopropanol : 2 ethyl acetate) after 66 hours indicated the formation of a product. The reaction was quenched with 2M NaOH (1ml) and the solvent removed *in vacuo*. Silica gel chromatography (1 water : 4 isopropanol : 8

ethyl acetate) followed by lyophilization allowed isolation of **2.6** as a white amorphous solid (13.3mg, 11%).



6,6'-diazido-D-trehalose (2.6)

R_f 0.69 (1 water : 2 isopropanol : 2 ethyl acetate), $[\alpha]_D^{17}$ +154.8 (c = 0.5, H_2O), lit.[39] $[\alpha]_D^{20}$ +156 (c = 1.0, H_2O); ^{1}H NMR (500MHz, D_2O) δ ppm 3.37 (2H, t, $J_{H3\text{-}H4}$ 9.1Hz, $J_{H4\text{-}H5}$ 9.1Hz, H4), 3.47 (2H, dd, $J_{H6a\text{-}H6b}$ 13.7Hz, $J_{H5\text{-}H6b}$ 5.8Hz, H6b), 3.58 (2H, as, H2), 3.58 (2H, dd, $J_{H6a\text{-}H6b}$ 10.4Hz, $J_{H5\text{-}H6a}$ 3.8Hz, H6a), 3.74 (2H, t, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H3\text{-}H4}$ 9.5Hz, H3), 3.89 (2H, ddd, $J_{H4\text{-}H5}$ 10.0Hz, $J_{H5\text{-}H6b}$ 5.9Hz, $J_{H5\text{-}H6a}$ 2.4Hz, H5), 5.11 (2H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 50.8 (2C, C6), 70.4 (2C, C4), 70.9 (2C, C2), 71.1 (2C, C5), 72.3 (2C, C3), 93.8 (2C, C1); FT-IR (KBr disc) υ 2106 (N_3), 3428 br (OH); HRMS m/z (ES^-) 391.1218 $[M - H]^-$ required (391.1219).

2.8.3 Flow chemistry and kinetic analysis

Reactions were performed using a Syrris Africa flow chemistry system using a 1000µl or 250µl microreactor chip. The flow system was dried by rinsing reagent flasks and microreactor chip with anhydrous DCM (x3). The reagent flasks were subsequently dried with a stream of nitrogen. The flow path was primed with nitrogen, prior to addition of any solvents or reagents.

HPLC was conducted with an in-line Varian PLS400 ELS detector for eluant detection. Injections were performed manually using a 20µl Rheodyne sample loop as per the manufacturer's instructions.

Quantification of eluants required determination of ELS detector response curves as reported previously.[40] 5µl of standardised concentrations of eluants (between 0 – 2.0mg/ml) were injected and the response of the detector measured and fitted to Area = am^b . ELS detector settings, chromatographic conditions and values of a and b for the different compounds are reported below.

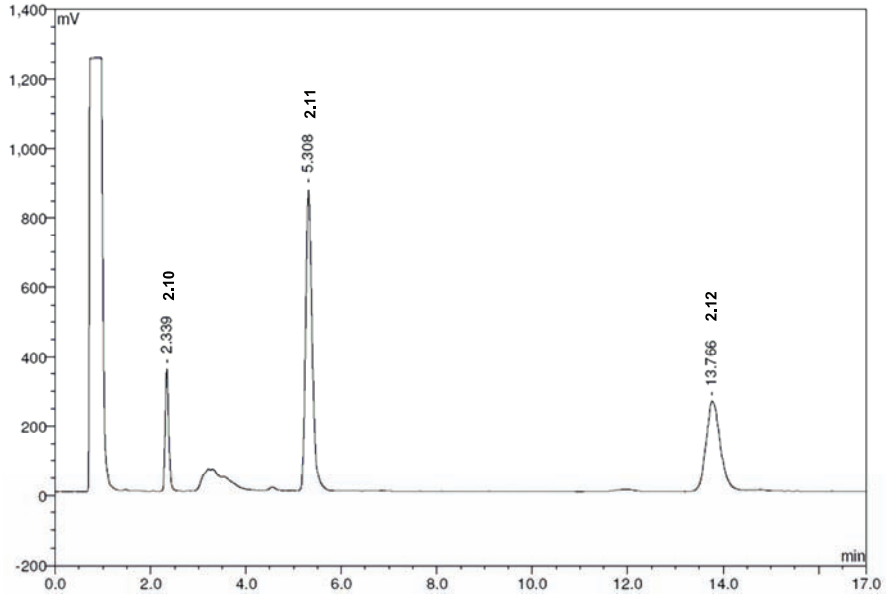
Kinetic parameters were estimated by minimisation of the objective function:

$$Q = \sum_i \sum_t \left(x_{i,t,exp} - x_{i,t,theo} \right)^2$$

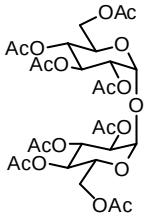
where $x_{i,t,exp}$ and $x_{i,t,theo}$ are experimental and theoretical concentrations respectively of the different reaction components, *i*, at different time points *t*. Theoretical values were determined by computationally solving equations (1)-(4) with Matlab 2009 using a Runge-Kutta algorithm.

Trehalose phosphorylation

Reactions were performed using a 1000µl microreactor chip at 40.0 °C. Solutions of D-trehalose (0.025M in pyridine) and DPCP (0.1M in pyridine) were pumped to the chip at identical flow rates. Reactions were performed with a residence time between 1-10 minutes (1 minute reactions were performed on a 250µl chip). 250µl of the reaction mixture was collected in a vial containing methanol (1ml). The solvents were removed *in vacuo* and the crude residue dissolved in pyridine (0.5ml). Acetic anhydride (0.5ml) was added and the reaction mixture stirred at room temperature for 8 hours. The reactions were quenched with methanol (1ml) and the solvents removed *in vacuo*. The crude products were dissolved in methanol (5ml) and the product distribution analysed by HPLC using 5µl injections into a Phenomenex Synergi Hydro C_{18} column (250 x 4.8mm, 4µm) with 40 water : 60 acetonitrile as the mobile phase at a flow rate of 1.0 ml/min. Eluants were detected using an in-line ELS detector (gain 5.0, nebulizer 30 °C, drift tube 80 °C, 1.65 Bar N_2).

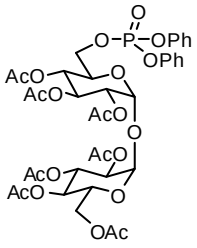


	a	b
2.10	261.5	1.435
2.11	255.8	1.421
2.12	120.7	1.309



2,2',3,3',4,4',5,5',6,6'-O-acetyl-D-trehalose (2.10)

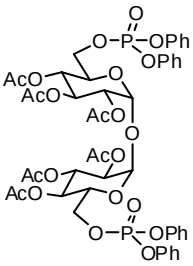
R_f 0.02 (2 petrol : 1 ethyl acetate); $[\alpha]_D^{25}$ +160 (c = 1.0, $CHCl_3$), lit.[41] $[\alpha]_D^{22}$ +160 (c = 1.0, $CHCl_3$), m.p. 97-98°C (methanol), lit.[41] m.p. 98-100°C; ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.02, 2.04, 2.07, 2.07 (4 x 6H, 4 x s, 8 x OAc), 4.00 (2H, dd, $J_{H6a-H6b}$ 12.1, J_{H5-H6b} 2.0Hz, H6b), 4.04 (2H, ddd, J_{H4-H5} 10.1Hz, J_{H5-H6a} 5.4Hz, J_{H5-H6b} 2.7Hz, H5), 4.23 (1H, dd, $J_{H6a-H6b}$ 12.0Hz, J_{H5-H6a} 5.7Hz, H6a), 5.03 (1H, dd, J_{H2-H3} 10.4Hz, J_{H1-H2} 3.8Hz, H2), 5.03 (1H, t, J_{H3-H4} 9.9Hz, J_{H4-H5} 9.9Hz, H4), 5.28 (1H, d, J_{H1-H2} 3.8Hz, H1), 5.48 (1H, t, J_{H2-H3} 9.8Hz, J_{H3-H4} 9.8Hz, H3); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.5, 20.6, 20.6, 20.6 (4 x 2C, 8 x CO<u>C</u>H_3), 61.7 (2C, C6), 68.1 (2C, C5), 68.5 (2C, C4), 69.8 (2C, C2), 69.9 (2C, C3), 92.2 (2C, C1), 169.5, 169.6, 169.9, 170.5 (4 x 2C, 8 x C=O); m/z 696.6 $[M + NH_4]^+$, 701.6 $[M + Na]^+$, 737.6 $[M + CH_3CN + NH_4]^+$.



6-*O*-(diphenoxyphosphoryl)-2,2',3,3',4,4',6'-*O*-acetyl-D-trehalose (2.11)

R_f 0.8 (ethyl acetate), $[\alpha]_D^{25}$ +50.7 (c = 0.71, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.03, 2.05, 2.06, 2.06, 2.09, 2.10, 2.11 (7 x 3H, 7 x s, 7 x OAc), 4.01 (1H, dd, $J_{H6'a-H6'b}$ 12.1Hz, $J_{H5-H6'b}$ 2.0Hz, H6'b), 4.06 (1H, ddd, $J_{H4'-H5'}$ 10.2Hz, $J_{H5'-H6'a}$ 5.5Hz, $J_{H5'-H6'b}$ 1.9Hz, H5'), 4.14 (1H, ddd, J_{H4-H5} 10.2Hz, J_{H5-H6a} 5.1Hz, J_{H5-H6b} 2.1Hz, H5), 4.24 (1H, dd, $J_{H6'a-H6'b}$ 12.3Hz, $J_{H5'-H6'a}$ 5.7Hz, H6'a), 4.25 (1H, add, $J_{H6a-H6b}$ 11.0Hz, $J_{H6b-31P}$ 6.3Hz, H6b), 4.31 (1H, ddd, $J_{H6a-H6b}$ 11.4Hz, $J_{H6a-31P}$ 6.6Hz, J_{H5-H6a} 5.7Hz, H6a), 4.96 (1H, dd, J_{H2-H3} 10.2Hz, J_{H1-H2} 3.9Hz, H2), 5.05 (1H, t, J_{H3-H4} 9.3Hz, J_{H4-H5} 9.3Hz, H4), 5.05 (1H, dd, $J_{H2'-H3'}$ 10.1Hz, $J_{H1'-H2'}$ 4.7Hz, H2'), 5.06 (1H, t, $J_{H3'-H4'}$ 9.6Hz, $J_{H4'-H5'}$ 9.6Hz, H4'), 5.13 (1H, d, J_{H1-H2} 3.8Hz, H1), 5.21 (1H, d, $J_{H1'-H2'}$ 4.1Hz, H1'), 5.46 (1H, t, $J_{H2'-H3'}$ 9.8Hz, $J_{H3'-H4'}$ 9.8Hz, H3'), 5.49 (1H, t, J_{H2-H3} 9.8Hz, J_{H3-H4} 9.8Hz, H3), 7.20-7.26 (6H, m, ArH_{ortho}, ArH_{para}), 7.35 (2H, t, $J_{ortho-meta}$ 7.9Hz, $J_{meta-para}$ 7.9Hz, ArH_{meta}), 7.36 (2H, t, $J_{ortho-meta}$ 7.9Hz, $J_{meta-para}$ 7.9Hz, ArH_{meta}); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.5, 20.5, 20.6 (x3), 20.7, 20.7 (7 x 1C, 7 x CO<u>C</u>H_3), 61.7 (1C, C6'), 66.6 (1C, d, J_{C-31P} 4.8Hz, C6), 68.2 (1C, C5'), 68.4 (1C, C4'), 68.4 (1C, C4), 68.7 (1C, d, J_{C-31P} 7.6Hz, C5), 69.5 (1C, C2'), 69.7 (1C, C2), 69.9 (1C, C3'), 70.1 (1C, C3), 92.5 (1C, C1'), 92.8 (1C, C1), 120.0 (2C, d, J_{C-31P} 2.9Hz, ArC_{ortho}), 120.0 (2C, d, J_{C-31P} 2.9Hz, ArC_{ortho}), 125.5 (2C,

ArC_{para}), 129.8 (4C, ArC_{meta}), 150.3 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, ArC_{ipso}), 150.4 (1C, d, $J_{C\text{-}31P}$ 4.8Hz, ArC_{ipso}), 169.3, 169.5, 169.5, 169.6, 170.0, 170.0, 170.6 (7 x 1C, 7 x C=O); $^{31}P\{^{1}H\}$ NMR (162MHz, $CDCl_3$) δ ppm -11.8 (1P, $P(O)(OPh)_2$); FT-IR (thin film) υ 1241 (P=O), 1723 (C=O); HRMS (ES^+) m/z 891.2083 $[M + Na]^+$ (required 891.2083).



6,6'-*O*-di(diphenoxyphosphoryl), 2,2',3,3',4,4'-*O*-acetyl-D-trehalose (2.12)

R_f 0.33 (2 ethyl acetate : 1 petrol), $[\alpha]_D^{19}$ +80.4 (c = 1.0, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.02, 2.03, 2.04 (3 x 6H, 3 x s, 6 x CH_3), 4.12 (2H, ddd, $J_{H4\text{-}H5}$ 10.1Hz, $J_{H5\text{-}H6a}$ 6.6Hz, $J_{H5\text{-}H6b}$ 2.0Hz, H5), 4.22 (2H, ddd, $J_{H6a\text{-}H6b}$ 11.1Hz, $J_{H6b\text{-}31P}$ 6.6Hz, $J_{H5\text{-}H6b}$ 2.2Hz, H6b), 4.29 (2H, ddd, $J_{H6a\text{-}H6b}$ 11.4Hz, $J_{H6a\text{-}31P}$ 6.6Hz, $J_{H5\text{-}H6a}$ 5.4Hz, H6a), 4.96 (2H, dd, $J_{H2\text{-}H3}$ 10.2Hz, $J_{H1\text{-}H2}$ 3.9Hz, H2), 5.00 (2H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 5.03 (2H, dd, $J_{H4\text{-}H5}$ 10.2, $J_{H3\text{-}H4}$ 9.3Hz, H4), 5.43 (2H, dd, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H3\text{-}H4}$ 9.5Hz, H3), 7.16-7.22 (12H, m, ArH_{ortho}, ArH_{para}), 7.32-7.36 (8H, m, ArH_{meta}); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.5, 20.6, 20.7 (3 x 2C, CH_3), 66.6 (2C, d, $J_{C6\text{-}31P}$ 5.7Hz, C6), 68.3 (2C, C4), 68.7 (2C, d, $J_{C5\text{-}31P}$ 8.6Hz, C5), 69.3 (2C, C2), 70.1 (2C, C3), 93.0 (2C, C1), 120.0 (8C, d, $J_{C\text{-}31P}$ 4.8Hz, ArC_{ortho}), 125.5 (4C, ArC_{para}), 129.8 (8C, ArC_{meta}), 150.3 (2C, d, $J_{C\text{-}31P}$ 6.7Hz, ArC_{ipso}), 150.4 (2C, d, $J_{C\text{-}31P}$ 7.6Hz, ArC_{ipso}), 169.3, 169.4, 170.0 (3 x 2C, C=O); $^{31}P\{^{1}H\}$ NMR (162MHz, $CDCl_3$) δ ppm -11.9 (1P, $P(O)(OPh)_2$); FT-IR (thin film) υ 1248 (P=O), 1755 (C=O); HRMS m/z (ES^+) 1081.23 (100%), 1082.23 (52%), 1083.24 (17%), 1084.24 (4%) required 1081.23 (100%), 1082.23 (53%), 1083.23 (19%), 1084.24 (5%).

Large scale trehalose phosphorylation

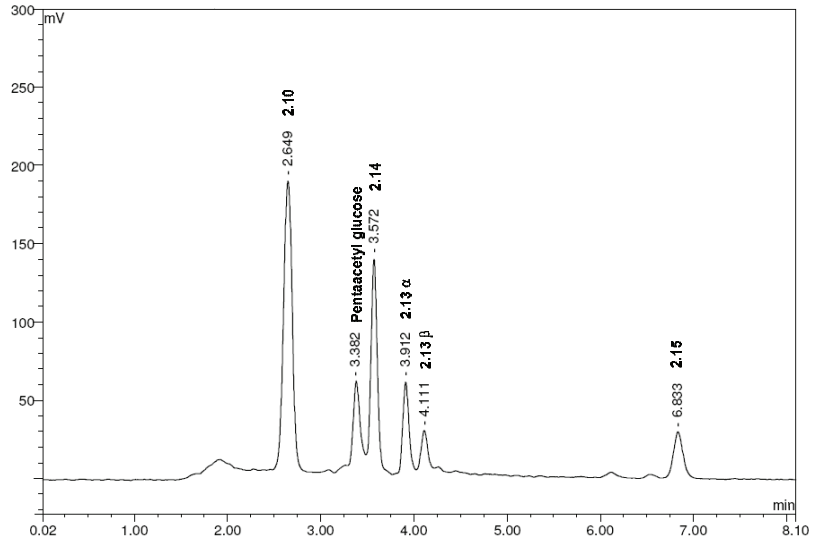
A 0.025M solution of anhydrous D-trehalose in pyridine was reacted with a 0.10M solution of DPCP in pyridine at 40.0 °C through a 1000μl microreactor with a residence time of 120 seconds. The effluent from the chip was quenched by addition into water (500ml). 468ml of reaction mixture (equating to 2.0g of trehalose, 5.85mmol) was collected. The solvents were removed *in vacuo* and co-evaporated with toluene to remove pyridine (x3). The desired product was purified by silica gel chromatography (1 water : 4 isopropanol : 6 ethyl acetate + 1% NH_4OH for 4 column volumes, then a gradient to 1 water : 4 isopropanol : 6 ethyl acetate over 6 column volumes) to give **2.1** as a white amorphous solid (1.31g, 39%).

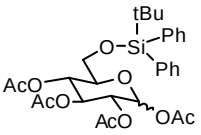
*Synthesis of **2.4***

To a stirred suspension of D-trehalose (3.00g, 8.77mmol, 1eq) in dry pyridine (60ml) was added *tert*-butyl diphenylchlorosilane (4.79ml, 18.4mmol, 2.1eq) at room temperature. After stirring for 12 hours, TLC (1 water : 2 isopropanol : 2 ethyl acetate) indicated the complete consumption of starting. The reaction mixture was concentrated *in vacuo* and co-evaporated with toluene to remove pyridine(x3). The product was purified by silica gel chromatography (9 THF : 1 methanol) to give the desired compound as a white amorphous solid after lyophilization (6.82g, 95%).

*Desilylation of **2.4** under acidic conditions*

Reactions were performed using an Africa Syris flow chemistry system using a 1000µl microreactor chip at 150.0°C. Solutions of **2.4** (8.77mmol in 60ml) and HCl (2.71ml of a 33% solution in 60ml) were pumped through the chip at equal flow rates. Reactions were performed with a residence time between 2-20 minutes. Approximately 0.2 Bar back pressure was applied to prevent boiling of the solvent. 250µl of the reaction mixture was collected and diluted with pyridine (1ml). To this mixture was added acetic anhydride (1ml) and the reaction mixtures maintained at room temperature for 10 hours. The reactions were quenched with methanol (1ml) and the solvents removed *in vacuo*. The crude products were dissolved in methanol (5ml) and the product distribution analysed by HPLC using 5µl injections into a Pheneomenex Jupiter Proteo C_{12} column (300 x 4.8mm, 5µm) with acetonitrile as the mobile phase at a flow rate of 1.0 ml/min. Eluants were detected using an in-line ELS detector (gain 2.0, nebulizer 30°C, drift tube 80°C, 1.50 Bar N_2).





2.13. α/β-methyl-6-O-tertbutyldiphenylsilyl-2,3,4-acetyl-d-glucopyranoside

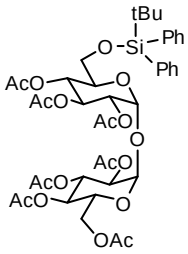
α:β 1:0.64, R_f 0.21 (2 petrol : 1 ethyl acetate), $[\alpha]_D^{25}$ +44.7 (c = 1.0, $CHCl_3$), FT-IR (thin film) 1754 (C=O); HRMS m/z (ES^+) 581.2175 $[M + Na]^+$ (required 581.2177).

α anomer

^{1}H NMR (500MHz, $CDCl_3$) δ ppm 1.05 (9H, s, ^{t}Bu), 1.88, 2.00, 2.08 (3 x 3H, 3 x s, 3 x OAc), 3.41 (3H, s, OMe), 3.62-3.77 (2H, m, H6, H6'), 3.87 (1H, ddd, J_{H4-H5} 10.2Hz, J_{H5-H6} 5.0Hz, $J_{H5-H6'}$ 2.4Hz, H5), 4.89 (1H, dd, J_{H2-H3} 10.2Hz, J_{H1-H2} 3.6Hz, H2), 4.97 (1H, d, J_{H1-H2} 3.5Hz, H1), 5.08 (1H, t, J_{H3-H4} 9.8Hz, J_{H4-H5} 9.8Hz, H4), 5.46 (1H, t, J_{H2-H3} 9.8Hz, J_{H3-H4} 9.8Hz, H3), 7.39 (4H, dd, $J_{ortho-meta}$ 7.6Hz, $J_{ortho-ortho}$ 3.5Hz, ArH_{ortho}), 7.43 (1H, t, $J_{meta-para}$ 7.3Hz, ArH_{para}), 7.66 (1H, ddd, $J_{ortho-meta}$ 8.2Hz, $J_{meta-para}$ 7.5Hz, ^4J 1.6Hz, ArH_{meta}); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 19.2 (1C, $\underline{C}(CH_3)_3$), 20.5, 20.6, 20.8, (3 x 1C, 3 x CO$\underline{C}H_3$) 26.7 (3C, C($\underline{C}H_3)_3$), 55.1 (1C, OMe), 62.6 (1C, C6), 68.9 (1C, C4), 69.9 (1C, C5), 70.5 (1C, C3), 71.0 (1C, C2), 96.5 (1C, C1), 127.7 (4C, ArC_{ortho}), 129.7 (2C, ArC_{para}), 133.1 (2C, ArC_{ipso}), 135.7 (4C, ArC_{meta}), 169.5, 170.2, 170.2 (3 x 1C, 3 x C=O);

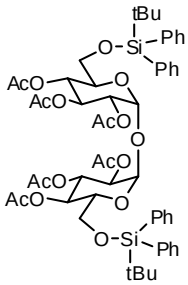
β anomer

^{1}H NMR (500MHz, $CDCl_3$) δ ppm 1.05 (9H, s, ^{t}Bu), 1.92, 2.00, 2.06 (3 x 3H, 3 x s, 3 x OAc), 3.52 (3H, s, OMe), 3.56 (1H, ddd, J_{H4-H5} 9.6Hz, J_{H5-H6} 3.6Hz, $J_{H5-H6'}$ 3.5Hz, H5), 3.63-3.76 (2H, m, H6, H6'), 4.42 (1H, d, J_{H1-H2} 7.9Hz, H1), 5.00 (1H, t, J_{H1-H2} 9.1Hz, J_{H2-H3} 9.1Hz, H2), 5.15 (1H, t, J_{H3-H4} 9.5Hz, J_{H4-H5} 9.5Hz, H4), 5.20 (1H, t, J_{H2-H3} 9.1Hz, J_{H3-H4} 9.1Hz, H3), 7.31-7.48 (6H, m, ArH_{ortho}, ArH_{para}), 7.69 (2H, ddd, ^3J 12.3Hz, ^3J 8.0Hz, ^4J 1.3Hz, ArH_{meta}); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 19.2 (1C, $\underline{C}(CH_3)_3$), 20.5, 20.6, 20.7 (3 x 1C, 3 x CO$\underline{C}H_3$), 26.6 (3C, C($\underline{C}H_3)_3$), 56.6 (1C, OMe), 68.6 (1C, C4), 70.6 (1C, C6), 71.4 (1C, C2), 73.3 (1C, C3), 74.5 (1C, C5), 101.4 (1C, C1), 127.6 (1C, ArC_{ortho}), 129.7 (1C, ArC_{para}), 133.2 (1C, ArC_{ipso}), 135.6 (1C, ArC_{meta}), 169.2, 169.7, 169.7 (3 x 1C, 3 x C=O).



6-*O*-tertbutyldiphenylsilyl, 2,2',3,3',4,4',6'-acetyl-D-trehalose (2.14)

R_f 0.17 (2 petrol : 1 ethyl acetate), $[\alpha]_D^{25}$ +119.4 (c = 1.0, $CHCl_3$), m.p. 154-156°C (ethyl acetate/petrol); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 1.03 (9H, s, ^{t}Bu), 1.92, 1.96, 2.01, 2.04, 2.05, 2.10, 2.10 (7 x 3H, 7 x s, 7 x OAc), 3.65 (2H, ad, *J* 3.2Hz, H6'a, H6'b), 3.94 (1H, dt, $J_{H4'\text{-}H5'}$ 10.2Hz, $J_{H5'\text{-}H6'a}$ 3.0Hz, $J_{H5'\text{-}H6'b}$ 3.0Hz, H5'), 4.04 (1H, dd, $J_{H6a\text{-}H6b}$ 12.0Hz, $J_{H5\text{-}H6b}$ 1.9Hz, H6b), 4.07 (1H, ddd, $J_{H4\text{-}H5}$ 10.4Hz, $J_{H5\text{-}H6a}$ 5.5Hz, $J_{H5\text{-}H6b}$ 2.0Hz, H5), 4.24 (1H, dd, $J_{H6a\text{-}H6b}$ 11.8Hz, $J_{H5\text{-}H6a}$ 5.2Hz, H6a), 5.03 (1H, dd, $J_{H2\text{-}H3}$ 10.2Hz, $J_{H1\text{-}H2}$ 3.9Hz, H2), 5.06 (1H, t, $J_{H3\text{-}H4}$ 8.7Hz, $J_{H4\text{-}H5}$ 8.7Hz, H4), 5.06 (1H, dd, $J_{H2'\text{-}H3'}$ 10.7Hz, $J_{H1'\text{-}H2'}$ 3.5Hz, H2'), 5.26 (1H, t, $J_{H3'\text{-}H4'}$ 9.8Hz, $J_{H4'\text{-}H5'}$ 9.8Hz, H4'), 5.25 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.32 (1H, d, $J_{H1\text{-}H2}$ 4.1Hz, H1), 5.47 (1H, t, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H3\text{-}H4}$ 9.8Hz, H3), 5.49 (1H, t, $J_{H2'\text{-}H3'}$ 10.1Hz, $J_{H3'\text{-}H4'}$ 10.1Hz, H3'), 7.31-7.47 (6H, m, ArH_{ortho}, ArH_{para}), 7.62 (4H, ddd, ^3J 10.9Hz, ^3J 8.0Hz, ^4J 1.3Hz, ArH_{meta}); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 19.1 (1C, $\underline{C}(CH_3)_3$), 20.4, 20.6, 20.7 (3 x 1C, 3 x CO$\underline{C}H_3$), 20.6, 20.6 (2 x 2C, 4 x CO$\underline{C}H_3$), 26.7 (3C, C($\underline{C}H_3)_3$), 61.8 (1C, C6), 61.8 (1C, C6'), 68.1 (1C, C5), 68.4 (1C, C4'), 68.4 (1C, C4), 69.6 (1C, C2'), 70.1 (1C, C2), 70.2 (1C, C3), 70.4 (1C, C3'), 70.6 (1C, C5'), 92.4 (1C, C1'), 92.5 (1C, C1), 127.7 (4C, ArC_{ortho}), 129.8 (2C, ArC_{para}), 129.8 (2C, ArC_{para}), 132.8 (2C, ArC_{ipso}), 132.8 (2C, ArC_{ipso}), 135.6 (2C, ArC_{meta}), 135.6 (2C, ArC_{meta}), 169.3, 169.4, 169.5, 169.7, 169.9, 170.2, 170.6 (7 x 1C, 7 x C=O); FT-IR (KBr disc) υ 1754 (C=O); HRMS m/z (ES^+) 897.2980 $[M + Na]^+$ (required 897.2972).



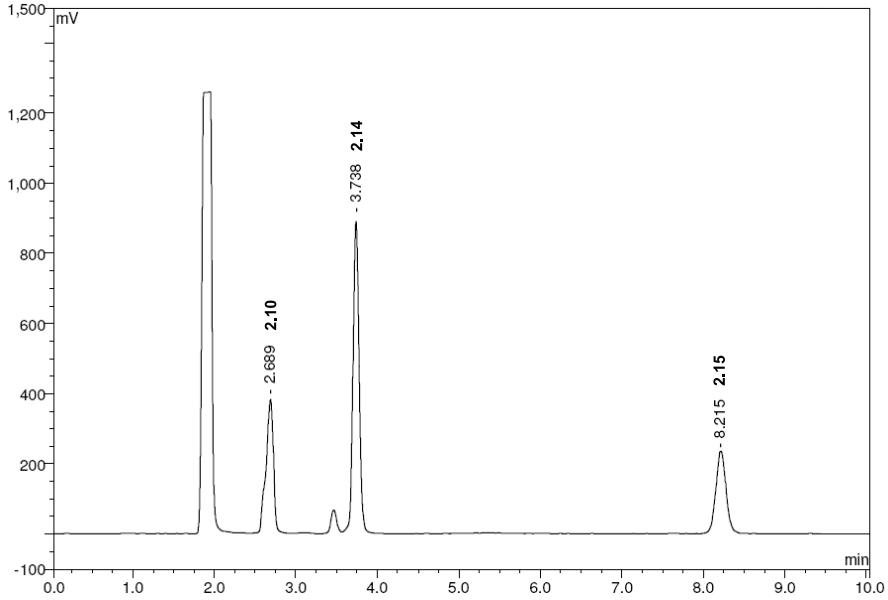
6,6'-*O*-ditertbutyldiphenylsilyl, 2,2',3,3',4,4'-acetyl-D-trehalose (2.15)

R_f 0.35 (2 petrol : 1 ethyl acetate), $[\alpha]_D^{25}$ +94.6 (c = 1.4, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 1.06 (18H, s, 2 x ^{t}Bu), 1.92, 1.96, 2.01 (3 x 6H, 3 x s, 6 x OAc), 3.67 (4H, ad, *J* 3.2Hz, H6a, H6b), 3.93 (2H, dt, $J_{H4\text{-}H5}$ 10.3Hz, $J_{H5\text{-}H6a}$ 2.3Hz, $J_{H5\text{-}H6b}$ 2.3Hz, H5), 5.04 (2H, dd, $J_{H2\text{-}H3}$ 10.4Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 5.28 (2H, t, $J_{H3\text{-}H4}$ 9.8Hz, $J_{H4\text{-}H5}$ 9.8Hz, H4), 5.27 (2H, d, $J_{H1\text{-}H2}$ 4.1Hz, H1), 5.46 (2H, t, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H3\text{-}H4}$ 10.1Hz, H3), 7.34-7.47 (12H, m, ArH_{ortho}, ArH_{para}), 7.64 (8H, ddd, ^3J 13.7Hz, ^3J 7.9Hz, ^4J 1.4Hz, ArH_{meta}); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 19.2 (2C, $\underline{C}(CH_3)_3$), 20.5, 20.6, 20.7 (3 x 2C, 6 x CO$\underline{C}H_3$), 26.7 (6C, C($\underline{C}H_3)_3$), 61.8 (2C, C6), 68.4 (2C, C4), 69.9 (2C, C2), 70.5 (2C,

C5), 70.8 (2C, C3), 92.6 (2C, C1), 127.7 (8C, ArC_{ortho}), 129.8 (2C, ArC_{para}), 129.8 (2C, ArC_{para}), 132.8 (2C, ArC_{ipso}), 132.9 (2C, ArC_{ipso}), 135.6 (4C, ArC_{meta}), 135.6 (4C, ArC_{meta}), 169.3, 169.6, 170.2 (3 x 2C, 6 x C=O); FT-IR (KBr disc) υ 1756 (C=O); HRMS m/z (ES^+) $[M + Na]^+$ peaks 1790.55 (100%), 1791.56 (88%), 1792.55 (41%), 1793.54 (13%), 1794.57 (5.0%), required, 1790.55 (100%), 1791.56 (84%), 1792.56 (44%), 1793.56 (18%), 1794.57 (6%); elemental C (57.72%) H (6.25%), required C (57.65%) H (6.22%).

*Desilylation of **2.4** with TBAF*

Reactions were performed using a Syrris Africa flow chemistry system using a 1000µl microreactor chip at 25.0°C. TBAF (8.77mmol in 60ml of solvent) was pumped through the chip at a flow rate three times that of **2.4** (8.77mmol in 60ml of solvent). For the high solubility system, **2.4** was dissolved in pyridine and TBAF was dissolved in methanol. For the low solubility system, **2.4** was dissolved in THF and TBAF was dissolved in petrol with 1% methanol. Reactions were performed with a residence time between 0-70 minutes. 250µl of the reaction mixture was collected and diluted with pyridine (1ml). To this mixture was added acetic anhydride (1ml) and the reaction mixtures maintained at room temperature for 10 hours. The reactions were quenched with methanol (1ml) and the solvents removed *in vacuo*. The crude products were dissolved in methanol (5ml) and the product distribution analysed by HPLC using 5µl injections into a Pheneomenex Jupiter Proteo C12 column (300 x 4.8mm, 5µm) with acetonitrile as the mobile phase at a flow rate of 1.0 ml/min. Eluants were detected using an in-line ELS detector (gain 2.0, nebulizer 30°C, drift tube 80°C, 1.65 Bar N_2)



	a	b
2.10	175.3	1.361
2.14	175.5	1.360
2.15	112.7	1.288

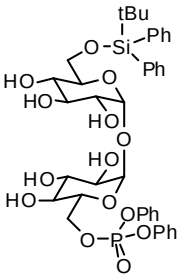
2.8.4 Optimized synthesis of trehalose-6-phosphate

Step (i) and (ii)

To a suspension of D-trehalose (1.0g, 2.92mmol, 1eq) in anhydrous pyridine (50ml) was added *tert*-butyl diphenyl chlorosilane (1.52ml, 5.84mmol, 2eq). TLC (1 water : 4 isopropanol : 10 ethyl acetate) after 14 hours indicated complete consumption of the starting materials. The solvent was removed *in vacuo* and the residue co-evaporated with toluene to remove pyridine (x3). To the crude solid, stirred at 0 ºC in methanol (5ml) and diethyl ether (15ml), was added acetyl chloride (0.40ml, 5.84mmol, 2eq). The reaction mixture was warmed to room temperature, and regularly monitored by TLC and mass spectrometry. After 18 hours, a significant amount of the D-trehalose by-product was seen (~10%) and the reaction quenched with saturated $NaHCO_3$ solution until the mixture was at pH 7 (pH paper), and the solvent removed *in vacuo*. Two separate silica gel columns were required to purify the desired compound (1 water : 4 isopropanol : 10 ethyl acetate followed by 2 isopropanol : 1 ethyl acetate). Lyophilization gave **2.3** as a white amorphous solid (0.93g, 55%).

Step (iii)

To a solution of **2.3** (0.90g, 1.55mmol, 1eq) in anhydrous pyridine (30ml) was added DPCP (0.36ml, 1.70mmol, 1.1eq). The reaction mixture was stirred at room temperature for 15 hours, after which TLC (1 water : 4 isopropanol : 10 ethyl acetate) indicated complete consumption of starting material. The reaction was quenched with methanol (2ml), the solvent removed *in vacuo*, and the crude solid co-evaporated with toluene to remove pyridine (x3). Silica gel chromatography (9 ethyl acetate : 1 methanol) followed by lyophilization gave the desired compound as a white solid (1.11g, 88%).



6-*O*-ditertbutyldiphenylsilyl-6'-*O*-diphenoxyphosphoryl-D-trehalose (2.9)

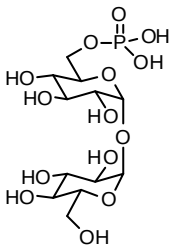
R_f 0.9 (1 water : 4 isopropanol : 10 ethyl acetate), $[\alpha]_D^{22}$ +31.4 (c = 1.0, MeOH), ^{1}H NMR (500MHz, MeOD) δ ppm 3.34 (2H, t, $J_{H3\text{-}H4}$ 9.7Hz, $J_{H4\text{-}H5}$ 9.7Hz, H4, H4'), 3.38 (1H, dd, $J_{H2\text{-}H3}$ 9.7Hz, $J_{H1\text{-}H2}$ 3.5Hz, H2), 3.46 (1H, dd, $J_{H2'\text{-}H3'}$ 9.9Hz, $J_{H1'\text{-}H2'}$ 3.9Hz, H2'), 3.78 (1H, t, $J_{H2'\text{-}H3'}$ 9.4Hz, $J_{H3'\text{-}H4'}$ 9.4Hz, H3'), 3.80 (1H, t, $J_{H2\text{-}H3}$ 9.4Hz, $J_{H3\text{-}H4}$ 9.4Hz, H3), 4.05 (2H, m, H5, H5'), 4.45 (1H, dd, $J_{H5\text{-}H6b}$ 4.1Hz, $J_{H6a\text{-}H6b}$ 7.3Hz, H6b), 4.47 (1H, dd, $J_{H6a\text{-}H6b}$ 7.1Hz, $J_{H5\text{-}H6a}$ 3.9Hz, H6a), 4.52 (1H, m, H6'b), 4.55 (1H, ddd, $J_{H6'a\text{-}H6'b}$ 11.2Hz, $J_{H6'b\text{-}31P}$ 6.6Hz, $J_{H5'\text{-}H6'b}$ 2.0Hz, H6'b), 7.19-7.26 (6H, m, P-OPh ArH_{ortho}, P-OPh ArH_{para}), 7.35-7.41 (10H, m, P-OPh ArH_{ortho}, P-OPh ArH_{meta}, Si-Ph Ar_{ortho}, Si-Ph Ar_{para}), 7.71-7.74 (4H, m, Si-Ph ArH_{meta}); ^{13}C NMR (126MHz, MeOD) δ ppm 20.3 (1C, $\underline{C}(CH_3)_3$), 27.4 (3C, $C(\underline{C}H_3)_3$), 69.9 (1C, d, $J_{C6'\text{-}31P}$ 6.7Hz, C6'), 69.9 (1C, C6), 71.3 (1C, C5), 71.7 (1C, C4), 72.0 (1C, d, $J_{C5'\text{-}31P}$ 6.7Hz, C5'), 73.1 (1C, C4'), 73.3 (1C, C2), 74.3 (1C, C2'), 74.6 (1C, C3), 74.9 (1C, C3'), 94.9 (1C, C1), 95.0 (1C, C1'), 121.2 (4C, d, $J_{C\text{-}31P}$ 4.7Hz, P-OPh ArC_{ortho}), 126.8 (2C, Si-Ph ArC_{para}) 128.7 (2C, d, $J_{C\text{-}31P}$ 4.8Hz, P-OPh ArC_{para}), 130.8 (4C, Si-Ph ArC_{ortho}), 131.1 (4C, d, $J_{C\text{-}31P}$ 2.9Hz, P-OPh ArC_{meta}), 134.8, 135.0 (2 x 1C, 2 x Si-Ph ArC_{ipso}), 136.8, 137.9 (2 x 2C, 4 x Si-Ph ArC_{meta}), 151.9 (2C, d, $J_{C\text{-}31P}$ 6.7Hz, P-OPh ArC_{ipso}); $^{31}P\{^1H\}$ NMR (162MHz, $CDCl_3$) δ ppm -12.1 (1P, $P(O)(OPh)_2$); FT-IR (KBr disc) υ 3369 br (OH); HRMS m/z (ES^-) 811.2556 $[M - H]^-$ required (811.2556).

Step (iv)

To a solution of **2.9** (1.10g, 1.35mmol, 1eq) in methanol (20ml) at room temperature, was added acetyl chloride (0.39ml, 5.41mmol, 4eq). The reaction was stirred at this temperature for 18 hours, after which TLC (1 water : 4 isopropanol : 4 ethyl acetate) indicated completion of the reaction. The reaction mixture was quenched with saturated $NaHCO_3$ solution to pH 7, and the solvent removed *in vacuo*. Silica gel chromatography (1 water : 4 isopropanol : 6 ethyl acetate) followed by lyophilization afforded **2.1** as a white amorphous solid (0.73g, 94%).

Step (v)

A suspension of **3** (0.75g, 1.31mmol, 1eq) and PtO_2 (15 mg, 0.066mmol, 0.05eq) in 75% aqueous ethanol (10 mL) with 0.5% glacial acetic acid (50 μL) was repeatedly degassed under high vacuum and the reaction vessel charged with hydrogen. The reaction was maintained at room temperature with aggressive stirring under an atmospheric pressure of hydrogen for 14 hours after which TLC (5 ethanol : 3 NH_4OH : 1 water) showed the complete consumption of the starting material and the formation of a single product. The reaction mixture was filtered through Celite® and the solvent was removed *in vacuo*. Lyophilization afforded Tre-6-P as a white amorphous solid (0.55g, 99%).



Trehalose-6-phosphate (1.1)

R_f 0.2 (5 ethanol : 3 NH_4OH : 1 water), $[\alpha]_D^{18}$ + 150.3 (c = 1.0, H_2O) lit.[42] $[\alpha]_D^{21}$ +151.2 (c = 0.8, H_2O); ^{1}H NMR (500MHz, D_2O) δ ppm 3.36 (1H, t, $J_{H3'\text{-}H4'}$ 9.5Hz, $J_{H4'\text{-}H5'}$ 9.5Hz, H4'), 3.50 (1H, t, $J_{H3\text{-}H4}$ 9.6Hz, $J_{H4\text{-}H5}$ 9.6Hz, H4), 3.56 (1H, dd, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.59 (1H, dd, $J_{H2'\text{-}H3'}$ 9.8Hz, $J_{H1'\text{-}H2'}$ 3.8Hz, H2'), 3.67 (1H, dd, $J_{H6'a\text{-}H6'b}$ 11.8Hz, $J_{H5'\text{-}H6'b}$ 5.4Hz, H6'b), 3.71-3.79 (4H, m, H3, H3', H5', H6'a), 3.82 (1H, d, $J_{H4\text{-}H5}$ 10.1Hz, H5), 3.86 (ddd, $J_{H6a\text{-}H6b}$ 12.1Hz, $J_{H6b\text{-}31P}$ 5.4Hz, $J_{H5\text{-}H6b}$ 1.7Hz, H6b), 3.94 (1H, ddd, $J_{H6a\text{-}H6b}$ 11.9Hz, $J_{H6a\text{-}31P}$ 7.7Hz, $J_{H5\text{-}H6a}$ 4.1Hz, H6a), 5.09 (1H, d, $J_{H1'\text{-}H2'}$ 4.1Hz, H1'), 5.12 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 60.5 (1C, C6'), 63.0 (1C, d, $J_{C6\text{-}31P}$ 4.8Hz, C6), 69.1 (C1, C4'), 69.7 (1C, C4), 70.9 (1C, C2), 71.1 (1C, C2'), 71.6 (1C, d, $J_{C5\text{-}31P}$ 6.7Hz, C5), 72.1 (1C, C5'), 72.2 (1C, C3), 72.4 (1C,

C3'), 93.3 (1C, C1), 93.4 (1C, C1'); $^{31}P\{^{1}H\}$ NMR (162MHz, D_2O) δ ppm 3.6 (1P, $ROP(O)(OH)_2$); m/z (ES^-) 421.5 $[M - H]^-$.

2.9 References

Figures in this chapter were reproduced by permission from The Royal Society of Chemistry from: M. K. Patel, B. G. Davis, *Org. Biomol. Chem.*, **2010**, 8, 4232.

1. M. J. Paul, L. F. Primavesi, D. Jhurreea, Y. Zhang, *Annu. Rev. Plant. Biol.*, **2008**, 59, 417.
2. (a) J. L. Chiara, I. S. de Gracia, A. Garcia, A. Bastida, S. Bobo, M. D. Martin-Ortega, *ChemBioChem*, **2005**, 6, 186. (b) C. Hui, C. T. Chang, *Org. Lett.* **2002**, 4, 2245.
3. (a) D. R. Ronning T. Klabunde, G. S. Besra, V. D. Vissa, J. T. Belisle, J. C. Sacchettini, *Nature Struct. Biol.*, **2000**, *7*, 141. (b) J. T. Belisle, V. D. Vissa, T. Sievert, K. Takayama, P. J. Brennan, G. S. Besra, *Science* **1997**, *276*, 1420.
4. P. J. Brennan, H. Nikaido, *Annu. Rev. Biochem.* , **1995,** *64*, 29.
5. (a) J. D. Rose, J. A. Maddry, R. N. Comber, W. J. Suling, L. N. Wilson, R. C. Reynolds, *Carbohdr. Res.* **2002**, *337*, 105. (b) J. Wang, B. Elchert, Y. Hui, J. Y. Takemoto, M. Bensaci, J. Wennergren, H. Chang, R. Rai, C. T. Chang, *Bioorg. Med. Chem.*, **2004**, *12*, 6397.
6. D. Rodrıguez-Lucena, C. O. Mellet, C. Jaime, K. K. Burusco, J. M. G. Fernandez, J. M. Benito, *J. Org. Chem.*, **2009**, *74*, 2997.
7. S. Shin, T. V. RajanBabu, *Org. Lett.* **1999**, *1*, 1229.
8. M. A. Rodríguez, O. Boutureira, M. I. Matheu, Y. Díaz, S. Castillón, P. H. Seeberger, P. H, *J. Org. Chem.* **2007**, *72*, 8998.
9. R. Namme, T. Mitsugi, H. Takahashi, S. Ikegami, *Eur. J. Org. Chem* **2007**, *2007*, 3758.
10. R. S. Babu, G. A. O'Doherty, *J. Carbohydr. Chem.* **2005**, *24*, 169.
11. M. R. Pratt, C. D. Leigh, C. R. Bertozzi, *Org. Lett.* **2003**, *5*, 3185.
12. (a) C. Schiraldi, I. D. Lernia, M. D. Rosa, *Trends Biothechnol.*, **2002**, 20, 420. (b) H. Chou, S. Chang, G. Lee, Y. Chen, T. Yeh, C. C. Akoh, J. Shaw, *Food Chem.*, **2010**, 119, 1017.

13. (a) J. H. Crowe, J. F. Carpenter , L. M. Crowe, *Annu. Rev. Physiol.*, **1998**, 60, 73. (b) F. Albertorio, V. A. Chapa, X. Chen, A. J. Diaz, P. S. Cremer, *J. Am. Chem. Soc.*, **2007**, 129, 10567.
14. D. Rodríguez-Lucena, J. M. Benito, E. Álvarez, C. Jaime, J. Perez-Miron, C. O. Mellet, J. M. G. Fernandez, *J. Org. Chem.*, 2008, **73**, 2967.
15. F. L. Lin, H. van Halbeek, C. R. Bertozzi, *Carbohydr. Res.*, 2007, **342**, 2014.
16. T. Tsuchiya, *Glycoscience: Chemistry and Biology* (ed., B. Fraser-Reid, K. Tatsuta, J. Thiem), **1**, (2001) Springer-Verlag, Berlin, pp117-194.
17. P. G. McDougal, J. G Rico, Y. I. Oh, B. D. Condon, *J. Org. Chem.*, 1986, **51**, 3388.
18. A. Bouzide, G. Sauvé, *Tetrahedron Lett.*, 1997, **38**, 5945.
19. A. Bouzide, G. Sauvé, *Org. Lett.*, 2002, **4**, 2329.
20. S. Hanessian, P. Lavallée, *Carbohydr. Res.*, 1973, **28**, 303.
21. G. Birch and A. C. Richardson, *Carbohydr. Res.*, 1968, **8**, 411.
22. (a) A. A. Frost, W. C. Schwemer, *J. Am. Chem. Soc.*, **1952**, 74, 1268. (b) H. J. R. Maget, *J. Polym. Sci., Part A: Polym. Chem.*, 1964, **2**, 1281.
23. D. A. Deranleau, *Cell Mol. Life Sci.*, **1982**, 38, 661.
24. (a) B. P. Mason, K. E. Price, J. L. Steinbacher, A. R. Bogdan, D. T. McQuade, *Chem. Rev.*, 2007, 107, 2300. (b) K. Jähnisch, V. Hessel, H. Löwe, M. Baerns, *Angew. Chem., Int. Ed.,* 2004, 43, 406. (c) P. Watts, C. Wiles, *Org. Biomol. Chem.*, 2007, 5, 727. (d) G. Jas, A. Kirschning, *Angew. Chem. Eur. J.*, 2003, 9, 5708. (e) F. Benito-López, R. J. M. Egberink, D. N. Reinhoudt, W. Verboom, *Tetrahedron*, 2008, 64, 10024. (f) S. V. Ley, I. R. Baxendale, R. N. Bream, P. S. Jackson, A. G. Leach, D. A. Longbottom, M. Nesi, J. S. Scott, R. I. Storer, S. J. Taylor, *J. Chem. Soc., Perkin Trans. 1*, **2000**, 3815.
25. D. G. Blackmond, *Angew. Chem. Int. Ed.*, **2005**, 44, 4302.
26. C. Capello, U. Fischer and K. Hungerbühler, *Green Chem.*, 2007, **9**, 927.
27. E. M. Nashed, C. P. J. Glaudemans, *J. Org. Chem*, **1987**, 52, 5255.
28. *Ricoh Co. Ltd.*, H. Takahiro, W. Yoichiro, S. Kazumi, O. Shinji, N. Yoshihiro, *Patent of Japan*, **2009**, patent number: 2009-122173.
29. T. Horie, M. Sumino, T. Tanaka, Y Matsushita, T Ichimura, J. Yushida, *Org. Process Res. Dev.*, **2010**, 14, 405.

30. (a) B. Zheng, L. S. Roach, R. F. Ismagilov, *J. Am. Chem. Soc*, **2003**, 125, 11170. (b) M. Takagi, T. Maki, M. Miyahara, K. Mae, *Chem. Eng. J.* **2004**, 101, 269. (c) A. Ufer, D. Sudhoff, A. Mescher, D.W. Agar, *Chem. Eng. J.*, **2011**, 167, 468.
31. S. L. Poe, M. A. Cummings, M. P. Haaf, D. T. McQuade, *Angew. Chem., Int. Ed.*, **2006**, 45, 1544.
32. M. Takagi, T. Maki, M. Miyahara, K. Mae, *Chem. Eng. J.*, **2004**, 101, 269.
33. (a) R.V. Ulijn and P. J. Halling, *Green Chem.*, 2004, **6**, 488. (b) B. Cho, J. Seo, T. Kang and B. Kim, *Biotechnol. Bioeng.*, 2003, **83**, 226. (c) V. N. Barai, S. V. Kvach, A. I. Zinchenko and I. A. Mikhailopulo, *Biotechnol. Lett.*, 2004, **24**, 1847.
34. D. G. Blackmond, *Chem. Eur. J.*, 2007, **13**, 3290.
35. (a) Y. Saito and H. Hyuga, *J. Phys. Soc. Jpn.*, 2009, **78**, 104001. (b) C. Viedma, *Phys. Rev. Lett.*, 2005, **94**, 065504.
36. Y. Saito and H. Hyuga, *J. Phys. Soc. Jpn.*, 2009, **78**, 104001; C. Viedma, *Phys. Rev. Lett.*, 2005, **94**, 065504.
37. P. A. Wallace, D. E. Minnikin, *J. Chem. Soc., Chem.Commun.*, **1993**, 1292.
38. A. C. Richardson, E. Tarelli, *J. Chem. Soc.*, **1971**, 3733.
39. Fernández, J.M.G. Mellet, C.O. Blanco, J.L.J. Mota, J.F. Gadelle, A. Coste-Sarguet, A.Defaye J. *Carbohydr. Res.*, **2005**, 268, 57.
40. Vervoort, N. Daemen, D. Török, G. *J. Chromatogr. A.*, **2008**, *1189*, 92.
41. Chittenden, G.J.F. *Carbohydr. Res.*, **1969**, 9, 323.
42. Viebel, *Biochem. Z.*, **1931**, *239*, 350

Chapter III - Phosphate group manipulations

3.1 Introduction

Phosphorylated carbohydrates are widespread in the natural world (Section 1 page 1). Their central role in metabolism and signalling make them a natural target for the synthetic chemist. However, carbohydrates and other poly-hydroxylated systems (such as inositols) have an inherent structural complexity which makes their phosphorylation particularly difficult.[1-5]

For the installation of a phosphoryl group, multistep sequences using protecting groups are generally required resulting in poor overall yields. Furthermore, the high polarity of a deprotected phosphate group can make purification difficult.[6] To combat this, the phosphate groups are also protected to simplify handling, but this can cause further problems with orthogonality to other functional groups in the molecule.[1,2,7] To shorten procedures, phosphorylations can also be performed on fully deprotected carbohydrates, but this strategy often results in multiple modifications.[4,8] Here, a balance has to be found between regioselectivity and the efficiency of phosphorylation. Modification at the anomeric centre is particularly difficult, as the phophoryl group is highly labile and therefore very sensitive to even moderately harsh reaction conditions.[5,9] For all of the above cases, migration and cyclization of the phosphate group is a common problem requiring particular attention.[1] Neighbouring hydroxyls, especially those in a vicinal position, are well placed to attack the phosphorous centre resulting in a vast range of possible side reactions. This is particularly problematic, since these reactions often occur during late stage deprotections and the resultant polar compounds can be very difficult to purify.

In response to these problems, a plethora of phosphorylation methodologies have been developed. Earlier methods utilized reagents with the P(V) oxidation states, such as $POCl_3$, phosphoesters and chlorophosphates.[8,10,11] These reagents tend to react with hard nucleophiles by virtue of the P=O bond and allowing for simple installation of the phosphoryl group. The use bulky side groups can afford high levels of regioselectivity, however a slow rate of reaction that is often observed. Older methods used alkoxides to activate the alcohols, or a proton scavenger to facilitate the reaction while more modern methods utilize metal catalysis, such as with $Ti(t\text{BuO})_4$ or

$Cu(OTf)_2$.[12,13] This enhances the reaction rate and allows the use of bulky phosphoryl transfer groups that improve regioselectivity.[13]

The development of phosphorylation techniques has been accelerated by oligonucleotide chemistry and the desire to have efficient syntheses that are amenable to solid-phase automation.[14] To this end, reagents in the P(III) oxidation state have found increased usage.[11,15] Such reagents are generally activated with amine side-groups giving phosphoramidite derivatives.[11] This activation is vital to improve the electrophilicity of these phosphites and such compounds yield very efficient phosphorylations.[11] A novel variant to this methodology is the use of polymer supported P(III) reagents which may have advantages in product purification.[4,16] A drawback to the phosphoramidite approach is the necessity for subsequent oxidation to the +5 states which may not always be compatible with the remainder of the molecule. Furthermore, the +3 state has a natural propensity to oxidation, therefore these compounds require careful handling. An alternative P(III) methodology that is receiving recent interest involves H-phosphonates which provide the reactivity of phosphoramidates, but with much improved stability.[11,17] These versatile reagents will likely find considerable use in the future.

Although the improvements in efficiency and regioselectivity of phosphorylations are an important advantage, the problem of phosphate group migrations is still largely unsolved for poly-hydroxylated systems.[1] While the above techniques give effective phosphorylations, subsequent manipulations (such as deprotections) may still expose neighbouring hydroxyls to conditions favouring phosphate cyclization and migration. For some compounds, enzymatic syntheses are effective workarounds which circumvent the need for any protecting groups.[18] An interesting method utilizing short peptide sequences to simulate the stereochemical environment of a protein allowing precise regio- and steroselective control over phosphorylations may be able to fulfil the role of enzymes.[19] However, for the majority of carbohydrates and inositol derivatives, phosphate migrations remain a significant problem that can only be alleviated by careful planning of the synthetic route.

Tre-6-P has been synthesized using the above chemistry. For example, the previous chapter described the use of DPCP and previous literature reports have utilized $POCl_3$ (which required a multi-step protecting group sequence to access the free 6 position)[20]

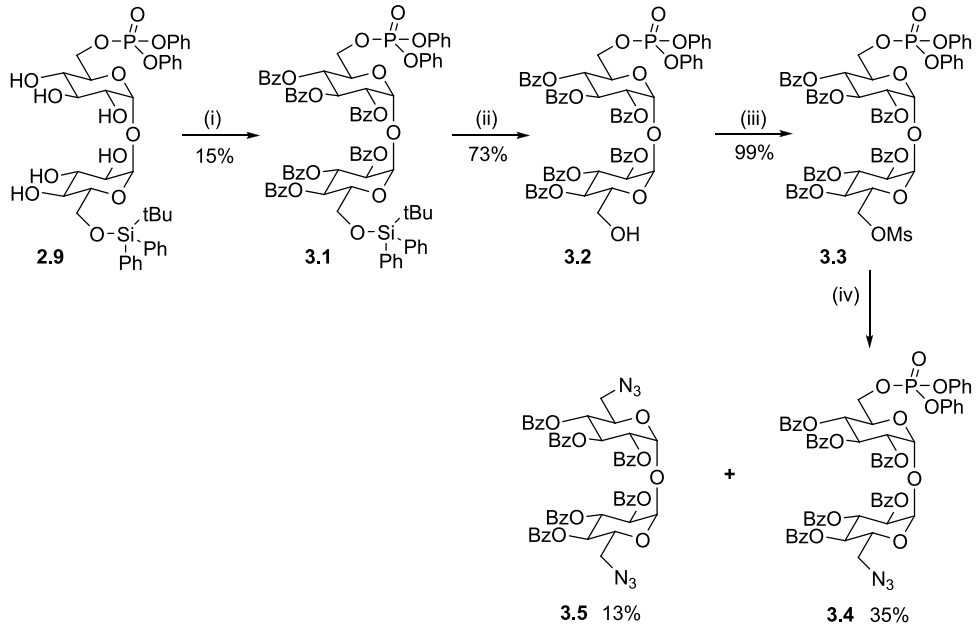
and even a reaction with sodium phosphate (which generated a mixture of phosphorylated analogues).[21] Enzymatic methods also exist for the synthesis using either *Saccharomyces cerevisiea*,[22] or trehalose phosphate synthase (TPS) from *Escherichia coli*.[23] Synthetic analogues of Tre-6-P, however, have not been reported.

In this chapter, some preliminary investigations on the synthesis of Tre-6-P analogues is described. The problems encountered with various phosphate group manipulations are discussed and their solutions are demonstrated. Initially, we show that C6' modifications can be easily performed by using the phosphate group itself as a leaving group. We then move onto phosphate deprotections where novel conditions are developed to prevent phosphate group migrations during transesterfications. Mechanistic studies are conducted on this reaction and a model is proposed to explain the observations. Finally, the scope of this reaction is investigated with other biologically important carbohydrates.

3.2 Nucleophilic substitutions of the diphenylphosphate group

We chose to modify the C6' position, since this primary centre is easily accessible using standard protecting group chemistry. The initial modification we chose to investigate was the installation of an azide group. Such a modification would provide a good model system to study the chemistry of Tre-6-P and would replicate the properties of photoaffinity probes[24] (such as compatibility with hydrogenation and other reaction conditions), which we would ultimately hope to use to elucidate the site of action between Tre-6-P and the SnRK1 intermediary factor (section 1.2.3, page 11).

Standard protecting group chemistry was used to gain selective access to the C6' centre (Scheme 3.1). Starting with **2.9**, protection of the secondary alcohols, to simplify handling of these compounds, was achieved with benzoyl chloride in pyridine in a poor 15% yield due to difficulty in purifying **3.1** freom residual benzoic acid. Benzoyl protection was chosen over acetyl protection to minimize silyl group migration in the subsequent deprotection step,[25] which was achieved using HCl generated *in situ* to give **3.2** in 73%.

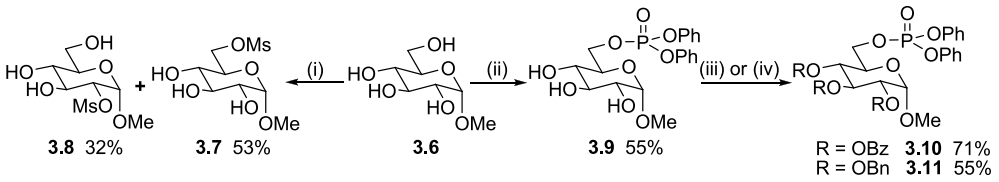


Scheme 3.1. *Initial reaction scheme for installation of azide group at the C6' position. Conditions: (i) BzCl, Py; (ii) AcCl, MeOH, Et_2O; (iii) MsCl, Py; (iv) NaN_3, DMF, 80°C.*

We envisaged a simple S_N2 reaction at the C6' position to install the azide functionality. The free hydroxyl was mesylated with mesyl chloride in pyridine to furnish **3.3** in 99% yield. The nucleophilic substitution was performed with sodium azide in DMF. The desired compound **3.4** was produced in 35% yield, but surprisingly, it was accompanied by the dizaide **3.5** in 13% yield. This implied that S_N2 substitution of the phosphate was occurring simultaneously, and from the yields it appears that phosphate displacement occurs at a similar (albeit slightly slower) rate to mesyl displacement.

While sulfonates are often used as leaving groups in organic syntheses, phosphates are used more rarely. Examples of phosphates acting as leaving groups include LiDBB mediated reductive cyclizations,[26] β eliminations,[27] and displacements involving Grignard reagents and other organometallics.[28] Other displacements have generally required soft nucleophiles such as thiols.[29] The few reported cases of phosphate displacement by azide have required the use of crown ethers and have benefited from catalysis by $^nBu_4NI$.[30] The facile nature of the phosphate displacement was potentially very useful, so we briefly investigated the reaction further.

Glucose analogues were synthesized as a model system to develop reaction conditions compatible with the phosphoryl group, and also to probe the propensity for phosphate displacement in a variety of different environments (Scheme 3.2).



Scheme 3.2. *Synthesis of precursors for testing the leaving group ability of diphenyl phosphate. Conditions: (i) MsCl, Py; (ii) DPCP, Py; (iii) BzCl, Py; (iv) Benzyl trichloroacetimidate, TMSOTf, DCM, 3Å molecular sieves, -40°C to RT.*

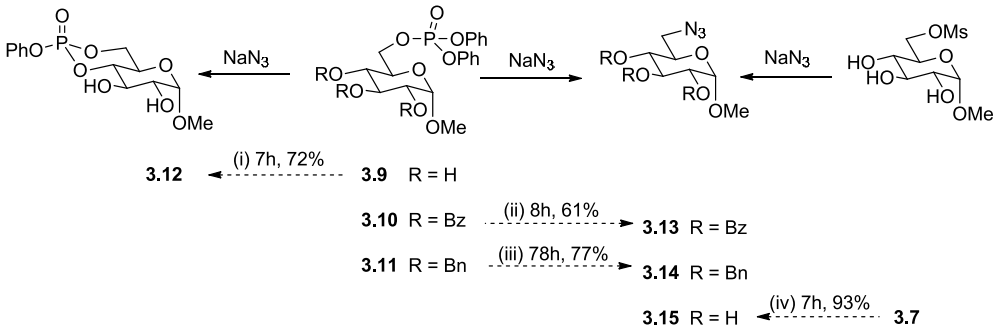
Mesylation of the 6-OH of α-methyl glucopyranoside **3.6** with mesyl chloride gave in 53% yield the desired compound **3.7**, which was to be used as a control against which other results were compared. The 2-mesylate **3.8** was also produced as a side product. Phosphorylation of **3.6** with DPCP gave **3.9** in 55% yield. Benzoylation of secondary alcohols proceeded smoothly to furnish **3.10** in 71% yield. The benzylation of **3.9** was more troublesome and a range of conditions had to be screened to obtain the target compound (Table 3.1).

Entry	Reagents	Solvent	Temperature	Yield
A	BnBr + NaH	DMF	RT	0%
B	BnBr + NaH + $^{n}Bu_4NI$	DMF	RT	0%
C	BnBr + $^{n}Bu_4NI$	DMF	70°C	0%
D	BnBr + Ag_2O	DMF	60°C	0%
E	BnBr + Ag_2O	DMF	RT	0%
F	BnTCA	Dioxane	RT	8%
G	BnTCA	Dioxane	50°C	0%
H	BnTCA	DCM	RT	2%
I	BnTCA	DCM	-40°C to RT	55%

Table 3.1. *Table summarising various attempted reaction conditions used for the synthesis of **3.11**. Note BnTCA = benzyl trichloro acetimidate.*

Use of sodium hydride (entries A and B) was found to be incompatible with **3.9** due to rapid phosphate cyclization under basic conditions to give **3.12**. Use of $^{n}Bu_4NI$ alone was also unsuccessful (entry C) and use of Ag_2O to aid bromide cleavage consistently resulted in decomposition of the starting material (entries D and E). Benzylation was attempted using benzyl trichloroacetimidate (BnTCA) with dioxane as solvent which gave a modest 8% yield (entry F). Heating the reaction mixture led to decomposition of the BnTCA, presumably to the amide (entry G).[31] While **3.9** exhibits lower solubility in DCM, the aforementioned amide rearrangement occurs to

a lesser extent in non-polar solvents since the reaction proceeds via an ionic mechanism.[31] Therefore, the reaction was attempted in DCM which at room temperature gave only a 2% yield (entry H). When the reaction was initiated at -40°C and allowed to warm to room temperature, much greater yields of 55% were achieved (entry I).*



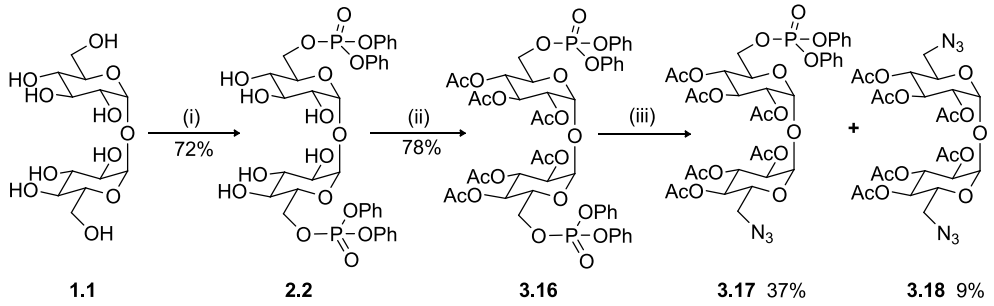
Scheme 3.3. *Comparison of the rate of displacement of mesyl group compared to diphenyl phosphate. Reactions were performed with 1eq of NaN_3 in DMF at 80°C until TLC showed complete consumption of the starting material.*

The glucose analogues were subjected to azide S_N2 reaction conditions (Scheme 3.3). For the wholly deprotected analogue **3.9**, azide substitution did not occur. The reaction was instead dominated by phosphate cyclization yielding **3.35** as the sole product. Thus, protection of the secondary alcohols is essential for a successful phosphate S_N2. As expected, azide nucleophilic substitution proceeded smoothly for the benzoyl analogue **3.10** furnishing **3.13** within 8 hours. The corresponding benzyl analogue **3.11**, however, reacted at a much slower rate requiring 78 hours for completion. In contrast, the mesylate was completely converted to **3.14** within 7 hours. From these results, it appears that phosphate displacements can be comparable to mesyl substitution with a suitably protected carbohydrate. The significantly faster displacements with benzoyl as compared with benzyl group may be as a result of anchimeric assistance from the 4-position carboxyl group, which is well placed to assist in leaving group expulsion via a 6-membered ring.

With an efficient phosphate displacement, other leaving groups are no longer needed and a shortened reaction sequence can be envisaged (Scheme 3.4). Now, the synthesis

* Since performing these reactions, methodology for benzylation in the presence of phosphoryl groups has been demonstarted using a stable benzyl pyridinium salt. See: (a) C. M. Longo, Y. Wei, M. F. Roberts, S. J. Miller, *Angew. Chem. Int. Ed.*, **2009**, 48, 1; and (b) K. W. C. Poon, G. B. Dudley, *J. Org. Chem.*, **2006**, 71, 3923.

can go via **2.2** where the diphenylphosphate functionality can serve a dual purpose of phosphate precursor and leaving group.



Scheme 3.4. *Optimized reaction scheme for installation of phosphate and azide functionality. Conditions: (i) 1.75eq DPCP, Py; (ii) Ac_2O, Py; (iii) NaN_3, DMF, 80˚C, unused* ***3.16*** *was re-subjected to reaction conditions. Total yield of 3.17 after 2 cycles was 51%.*

The reaction of trehalose, **1.1**, with DPCP was not perfectly regioselective giving multiple phosphorylation products as well as the desired diphosphate **2.2**. Hence, the reaction was screened with varying equivalents of DPCP to optimize yields (Table 3.2).

Entry	Eq of DPCP	% 2.1	% 2.2
A	4	0	0
B	2	0	57
C	1.75	5	72
D	1	24	36

Table 3.2. *Table showing product distribution for varying equivalents of DPCP. Entries A and B led to multiple phosphorylations. Entry C was used for the synthesis of* **2.2**.

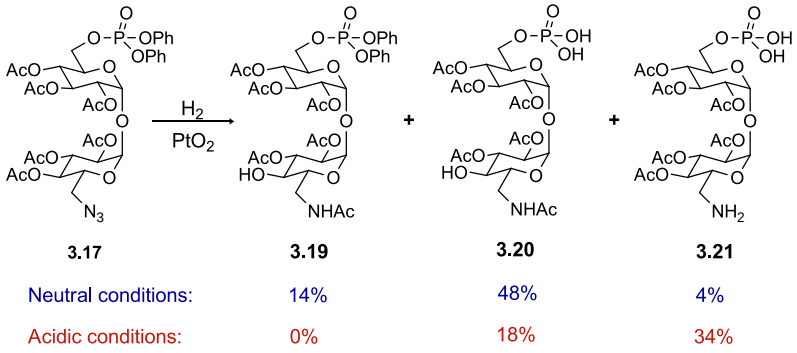
With excess DPCP the reaction produced only multiply phosphorylated products (entry A). Similarly, with 2 equivalents, significant over phosphorylation still occurred lowering the yield for **2.2** (entry B). The maximal yield for the desired product required 1.75 equivalents of DPCP (entry C) and these conditions were used for the synthesis of **2.2**. Analysis of the over-phosphorylated products showed that the third phosphorylation occurs preferentially at the 2 position (see section 3.9, page 85).

For secondary alcohol protection, the benzoyl group was abandoned due to persistently low yields and since silyl group migration was no longer a problem. Instead, acetate protection afforded **3.16** in 78% yield. This compound was subjected to azide S_N2 conditions and the reaction was continuously monitored by TLC. Upon formation of significant quantities of **3.18** (9%), the reaction was stopped affording 37% of the desired product **3.17**. Unused starting material was also recovered and re-

subjected to the reaction conditions giving a second crop of the desired compound. Thus, after two cycles, **3.17** was produced in a healthy 51% total yield.

3.3 Diphenylphosphate deprotections

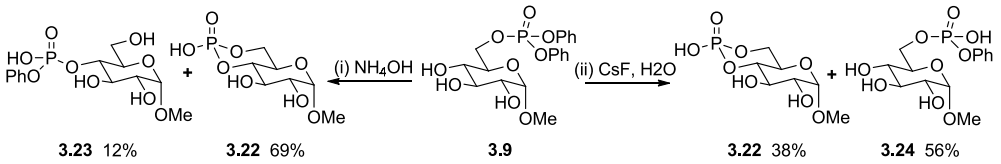
The most common method for deprotecting phenylphosphates is by hydrogenation over PtO_2 (section 2.6 page 41). This, however, is incompatible with the azide functionality in **3.17** which is rapidly reduced to an amine (Scheme 3.5). Under neutral conditions, this results in acetyl migration from the 4' hydroxyl giving compound **3.19**. Eventual phosphate deprotection gives **3.20** as the major product in 48% yield. Only a small amount of product is isolated (**3.21**) where no acetyl migration occurs to the amine.



Scheme 3.5. *Attempted deprotection of phosphate group by hydrogenation results in concominant azide reduction.Reaction conditions: H_2, PtO_2, 75% aqueous EtOH, 90-120mins.*

In our hands, the rate of phenylphosphate deprotection increased in an acidic environment. Hence, we repeated the above reaction at pH 4 to improve phosphate deprotection relative to azide reduction. In this case, phosphate deprotection was much faster and the reaction could be halted in 90 minutes, minimizing acetyl migration (compound **3.20**) and giving the amine **3.21** in 34% yield.

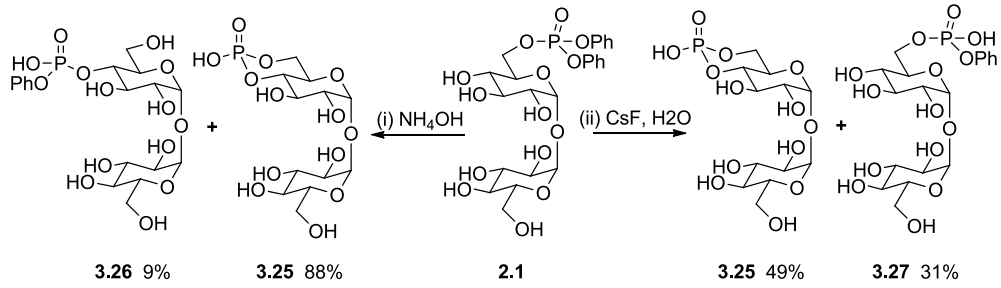
Under no conditions was the azide left intact. Thus azide reduction was always preferential to phosphophate deprotection and an alternative to hydrogenation was clearly needed. Literature precedent suggested that deprotection could be accomplished with basic hydrolysis.[32] To develop deprotection conditions, **3.9** was used as a model system for further studies (Scheme 3.6).



Scheme 3.6. *Testing non-hydrogenic deprotection of diphenyl phosphate on a glucose model system. Conditions: (i) NH_4OH, H_2O, ^tBuOH, 10mins; (ii) 10eq CsF, H_2O, 80˚C, 12h.*

On exposure of **3.9** to ammonium hydroxide, the reaction produces the cyclic phosphate **3.22** in 69% yield within 10 minutes of initiation. With such short reaction times, 12% of the partially deprotected, but migrated, 4-phosphate **3.23** was also isolated. It is worth noting that over longer reaction times, **3.23** was not isolated and **3.22** was the sole product.

Inspired by the work of Ogilvie *et al.*, who preformed deprotections and transesterfications on phenylphosphates in various nucleotides using CsF,[33] we decided to employ similar conditions for the deprotection of **3.9**. Just as before, phosphate cyclization was the major pathway for the reaction yielding **3.22** in 38%. However, these condition gave substantially more of the partially protected adduct **3.24**, which crucially, had the phosphoryl group retained at the 6-position.

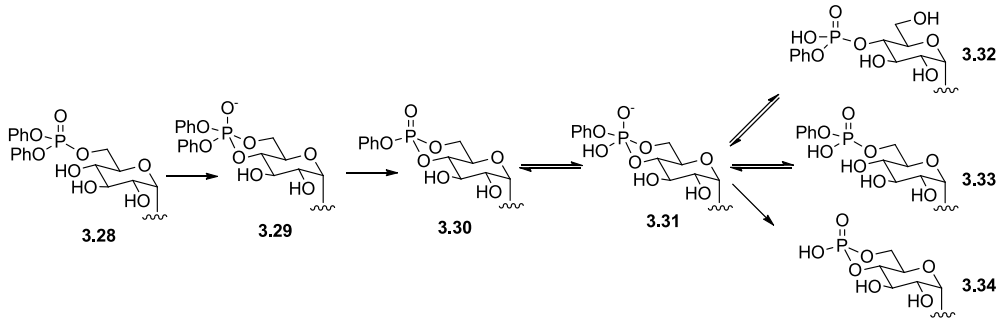


Scheme 3.7. *Testing non-hydrogenic deprotection on trehalose. Conditions: (i) NH_4OH, H_2O, ^tBuOH, 10mins; (ii) 10eq CsF, H_2O, 80˚C, 12h.*

These reactions were repeated for the equivalent trehalose analogue **2.1** (Scheme 3.7). Once again, treatment with ammonium hydroxide led predominantly to phosphate cyclization yielding **3.25** in 88% yield. Over the short time course, the 4-phosphate **3.26** was isolated in 9%, but over longer periods, only **3.25** was produced. Reaction with CsF did not result in phosphate migration, producing the partially deprotected analogue **3.27** in a good 31% yield, accompanied by 38% of **3.25**.

The dominance of the cyclic products can be rationalized by considering the consecutive steps in the phosphate transesterfication sequence (Scheme 3.8). Initially, the diphenylphosphate **3.28** is prone to intramolecular attack by the neighbouring 4-

hydroxyl producing the intermediate **3.29**. Intramolecular cyclization is entropically more favourable than attack by external nucleophiles, and so this step is likely to be rapid. From **3.29**, the phenoxy group extrusion would be the most favourable, giving cyclic compound **3.30**. At this point, external attack of a hydroxide is possible yielding the intermediate **3.31**. Phosphate migrations are known to occur via such an intermediate,[34] and the final product distribution is dependent on the loss of one out of three possible leaving groups. Expulsion of the 6-OH would give the 4-phosphate **3.32**, while loss of the 4-OH would return to the 6-phosphate **3.33**. However, loss of the phenoxy group is more favourable and therefore the reaction predominantly gives **3.34**. These diesters exist in the deprotonated form and the repulsive effects of this anionic charge are the source of extreme hydrolytic stability.[35] As a result, external nucleophilic attack on **3.34** is not possible and its formation is therefore irreversible. However, both **3.32** and **3.33** are still inclined to internal cyclization and therefore their conversion from **3.31** is reversible. Hence, given enough time, the cyclic phosphate **3.34** would be the sole reaction product.



Scheme 3.8. *The phosphate transesterfication sequence resulting in three possible products.*

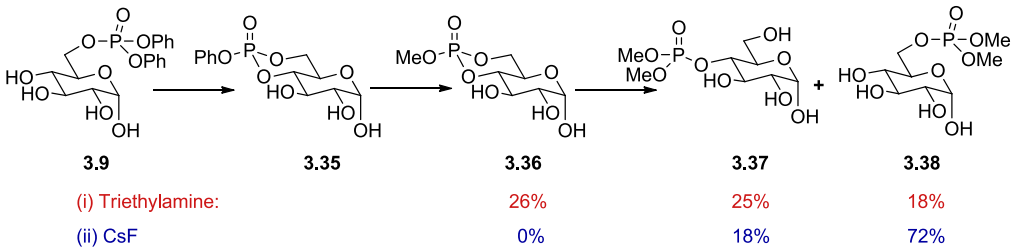
A striking feature of the above deprotections is the lower propensity for phosphoryl group migrations in the presence of CsF – a feature that would be continually replicated in other experiments.

3.4 Phosphate transesterfication

Because conditions for complete phosphate deprotection could not be found, an alternative strategy was sought. Rather than attempting an alkaline deprotection of the phenylphosphates, a transesterfication could be performed to produce a methoxyphosphate, which could then be deprotected by treatment with soft nucleophiles.[7] Unlike the above deprotections which irreversibly produce cyclic

phosphates (**3.34**), a transesterfication produces an analogous methoxy species, which crucially, is not charged and is therefore formed reversibly. Hence, by performing the reaction with methanol as the solvent, the reaction can be pushed towards the desired dimethylphosphates. To perform this reaction, we initially used triethylamine in methanol heated to reflux, but this resulted in a plethora of unwanted side products. Under the strongly basic conditions, it was found that after methoxy transesterfication, nucleophilic attack of the solvent at the methyl centre resulted in net deprotection giving compounds of the type **3.32** and **3.33**.

To prevent such deprotections, triethylamine was slowly added to the reaction mixture over a period of 2 hours. Careful monitoring of the products by TLC allowed the reaction to be stopped before any deprotection occurred. These conditions were tested with the glucose model system **3.9** (Scheme 3.9).

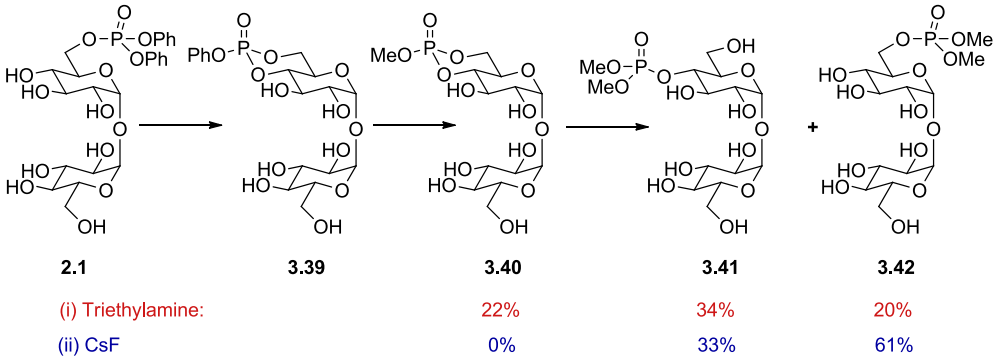


Scheme 3.9. *Product distribution for the deprotection of the diphenyl phosphate group in the glucoe model system. Conditions: (i) NEt_3, in MeOH added over 2h and reflux for 1h; reaction quenched on appearance of deprotected side products; (ii) 10eq CsF, MeOH, reflux, 16h.*

The reaction products could be monitored by TLC and mass spectrometry. The diphenylphosphate quickly cyclized to give **3.35**, which was then converted into the methyl ester **3.36**. This compound persisted for longer than its precursors, but over time, was converted to the 4-phosphate **3.37** and the 6-phosphate **3.38**. The reaction was stopped once deprotection was detected. As with the hydroxyl deprotections (section 3.3, page 72), the major product was the migrated 4-phosphate **3.37** in 25% yield with only 18% formation of the desired product **3.38**.

As before, we attempted the transesterficatiosn with CsF using slightly modified conditions from Ogilvie *et al.*[33] Compound **3.36** was quickly formed and was a persistent entity within the reaction mixture which, over time, reacted further to give **3.37** and **3.38**. Under these milder conditions, methyl deprotection did not occur and the reaction could be allowed to reach completion without side product formation. Moreover, CsF was once again found to reduce the amount of phosphate migration.

We then extended the principle to **2.1** (Scheme 3.10).



Scheme 3.10. *Product distribution for the deprotection of the diphenyl phosphate group. Conditions: (i) NEt_3, in MeOH added over 2h and reflux for 1h; reaction quenched on appearance of deprotected side products; (ii) 10eq CsF, MeOH, reflux, 16h.*

The general patterns of reactivity were replicated for trehalose. For both reaction conditions, **2.1** quickly cyclized to give **3.39**, which subsequently underwent transesterfication to the methoxy analogue **3.40**. Over the course of the reaction, **3.40** was ring-opened to either **3.41** or **3.42**. Once again, under basic conditions the 4-phophate was preferred over the 6-phosphate, whereas the opposite was true when CsF was used. This unique feature clearly warranted further investigations.

3.5 Control of phosphate migrations with CsF

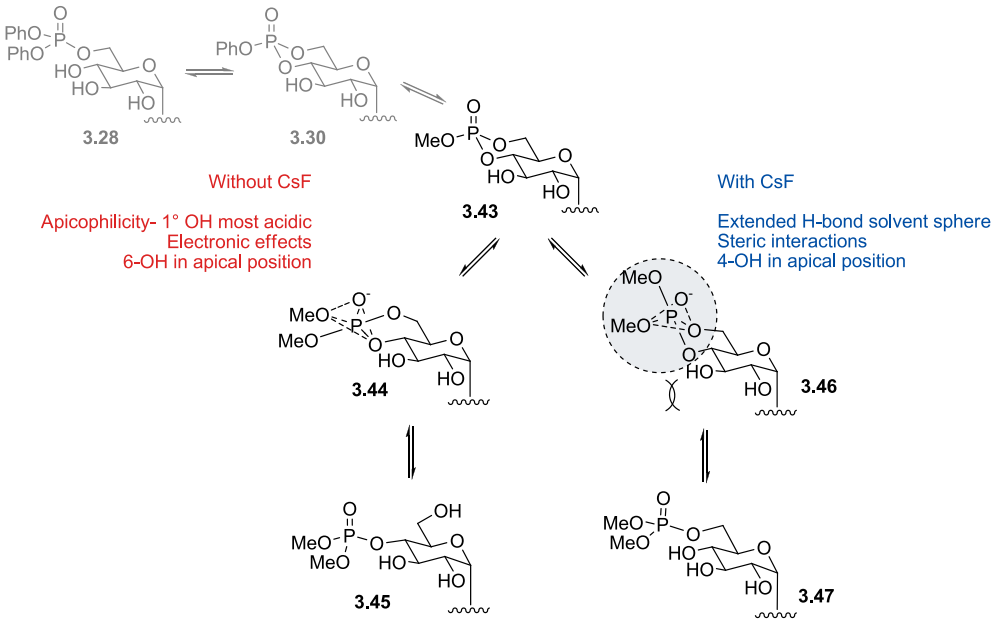
A series of transesterfication reactions were carried out on **2.1** under various conditions to probe possible reasons for the altered migration properties (Table 3.3). The proportions of **3.41** and **3.42** were determined by ^{1}H NMR of the crude mixture and C_{18} HPLC with ELS detection (section 3.9 page 85). From this data, various patterns could be elucidated. Firstly, we found that a decreasing concentration of CsF (going from entry A to entry C) led to a greater amount of phosphate migration to the 4 position. Other fluoride salts were also capable of transesterfication, but the nature of the salt was found to have a large effect on migration. In particular, it was found that fluoride salts with a higher solubility in organic solvents (going from entry D to entry F) reduced the amount of phosphoryl migration. Thirdly, the solvent was also found to have a sizeable effect on the reaction outcome. Our analysis of the data shows that the observed migration decreases broadly with the increasing H-bond acceptor and donor strengths of the solvent (entry G to J).[36] The reaction was found to be essentially insensitive to the pH (entries K and L).

Entry	Conditions	Yield %	3.41%	3.42%	Side Products
A	10eq CsF	94	35	65	-
B	2.5eq CsF	91	42	58	-
C	NEt_3	54	63	37	**3.40** (22%)
D	KF	93	39	61	-
E	CsF	94	35	65	-
F	TBAF	51	15	85	**3.25** (24%)
G	DMF	59	45	55	**3.25** (22%)
H	^{t}BuOH	62	41	59	**3.25** (11%)
I	MeOH	94	35	65	-
J*	H_2O	31	0	100 (**3.27**)	**3.25** (49%)
K	CsF pH 13	92	36	64	-
L	CsF pH 2	93	34	68	-
M	CsCl	0	0	0	**3.25** (82%)
N	$MgCl_2$	0	0	0	**3.25** (85%)
O	10eq CsF, 50˚C	90	37	63	-

Table 3.3. *Product distributions for transesterfications under different reaction conditions, grouped together according to various patterns of reactivity. The yield column represents the combined yield of **3.41** and **3.42**. The corresponding columns for **3.41** and **3.42** represent the proportion of that yield that is the 4-phosphate or the 6-phosphate. Other side products, where applicable, are also reported together with their yields. All reactions were performed under reflux, with the exception of entry O. *Entry J is a transesterfication with OH^- rather than MeOH (Scheme 3.7) and is added for comparison.*

Finally, the reaction was investigated in the absence of fluoride, but with differing cations to probe the system's behaviour to Lewis acidity (entry M and N). Under these conditions, **3.25** was formed exclusively, presumably by attack at the methyl by the more nucleophilic chloride. Finally, a reaction was also performed at a lower temperature of 50˚C (as opposed to reflux conditions as for the other reactions) and while the reaction required longer for completion, the product distribution was essentially identical to the higher temperature case (entry O). The fluoride ion clearly has an important role to play in the above reactions. However, in all of these experiments, a fluorophosphonate has never been detected or isolated. Such compounds are generally quite stable and readily isolable, so if they were being formed in the reaction, we would expect to detect their presence. Furthermore, when a control reaction was performed with **2.1** and CsF in DMF in the absence of methanol, no fluorophosphonates were detected and only phosphate cyclization occurs to produce **3.39** as the sole product. Hence, it is believed that fluoride attack onto phosphorous is not occurring to any significant extent and is therefore unlikely to be responsible for controlling the product distribution.

Rather than direct fluoride participation, we suggest a model based on increased hydrogen bonding in the solvent mediated by the fluoride ion. The mechanism of phosphate migrations is known to occur through a cyclic intermediate,[34] following the stereochemical rules proposed by Westheimer (Scheme 3.11).[37]



Scheme 3.11. *Proposed model to explain the outcome of phosphate migrations during transesterfication reactions under normal conditions (left) and with CsF (right).*

The kinetic preference for phosphate migration to the 4 position is well documented[38,39] and can be explained through the concept of apicophilicity.[38,40,41] Under normal conditions, nucleophilic attack onto the cyclic phosphate **3.43** gives an intermediate with trigonal bypyramidal geometry **3.44**, where the most electronegative atom takes the apical position. In this case, the primary 6-OH is less electron rich than the secondary 4-OH and therefore assumes the apical position from which ligand extrusion is more facile, since the longer apical bond is easier to break. Thus the 4-phosphate **3.45** is formed.

The preference of electronegative atoms for the apical position can be rationalized by considering the molecular orbitals of the hypothetical molecule PH_5 in the trigonal bypramidal geometry.[42] A simple group theoretical analysis of the relative changes in molecular orbital energy levels on switching a hydrogen with a more electronegative substituent can reveal the fundamental causes underlying apicophilicity (Figure 3.1).

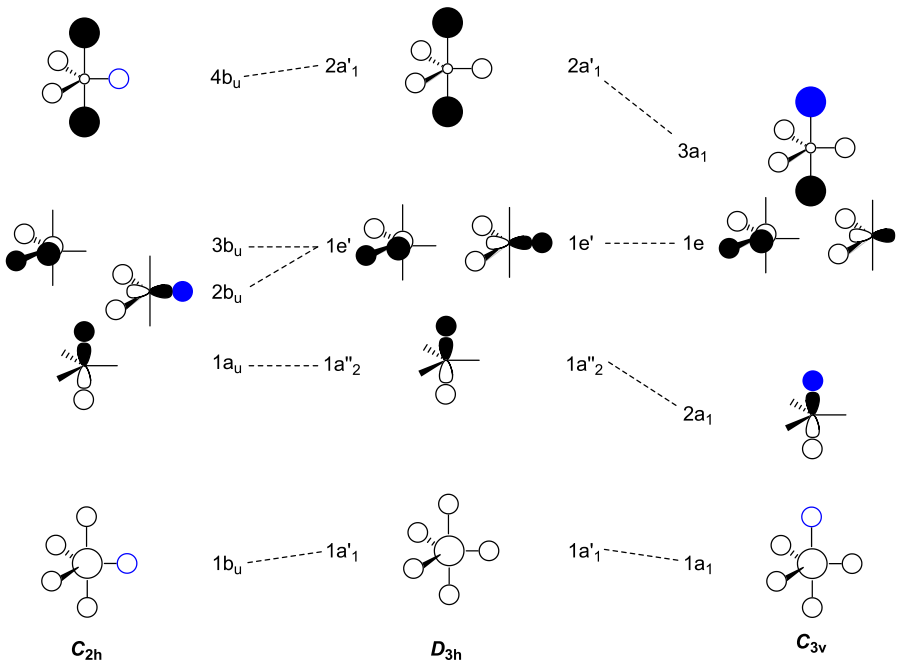


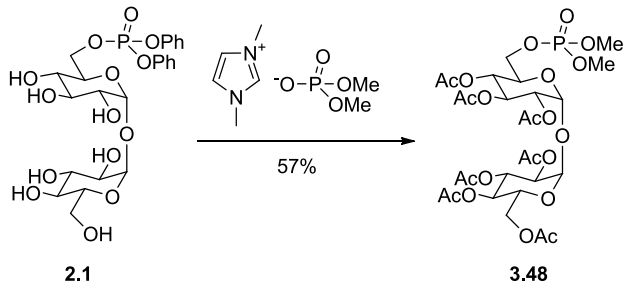
Figure 3.1. *Correlation diagram showing the relative changes in the energy levels of the occupied molecular orbitals on switching a hydrogen atom in the hypothetical molecule PH_5 with a more electronegative element at either the equatorial position (left) or apical position (right).*

For a standard PH_5 molecule, there exist five occupied molecular orbitals in the D_{3h} point group. Substitution of an equatorial hydrogen with a more electronegative substituent (here a hypothetical H* atom with the electronic configuration $1s_2$) results in the C_{2h} geometry. The degeneracy of the 1e' orbitals is lost and $2b_u$ is stabilized by a greater bonding interaction from a p_x orbital on the phosphorous pointing directly at the H* 1s orbitals. In contrast, a larger stabilization occurs by substitution at the apical position for the $1a''_2$ orbital (now $2a_1$). This stabilization is larger since the delocalised orbital is over only 2 hydrogen atoms as opposed to 3 in the C_{2h} case and thus, the orbital coefficients at these positions will be bigger. However, the dominating factor favouring apical substitution is the behaviour of the non-bonding orbital $2a'_1$. This orbital (which is mostly derived from the two in-phase apical H* $1s_2$, with slight mixing with the PH_3 bonding and anti-bonding orbitals) has much larger coefficients at the apical positions rather than the equatorial ones. Hence, $3a_1$ in the C_{3v} point group is far more stabilized than the corresponding $4b_u$ orbital in the C_{2h} group.

However, sterical interactions can also have an important effect on apicophilicity that may override the above electronic arguments.[40] The dissolution of fluoride ions is dominated by hydrogen bonding interactions.[43] In this way, fluoride ions act as

"structure markers" that increase solvent ordering.[44] In particular, for methanolic solutions of CsF, conductivity studies[45] have confirmed this effect where enhancements in the secondary solvent sheath have been detected.[46] Recent work using dielectric spectroscopy[45] has identified the formation of vertex linked solvent shared ion pairs.[47] Moreover, NMR studies have shown that fluoride ions interact strongly with the hydroxyls in glucose[48] and it is therefore conceivable that these alcohols could be involved in hydrogen bonding networks with the solvent. Hence it is proposed that in the presence of fluoride, hydrogen bonding is likely to be much greater resulting in increased solvent ordering and a large solvation sphere around the charged phosphoryl group (**3.46**, Figure 3.1). This increased steric bulk would be able to overcome the subtle electronic effects of apicophilicity, because the secondary 4-OH would assume the more spatial apical position to reduce steric interactions, allowing the primary 6-OH to make a stronger equatorial bond. Thus, extrusion of the 4-OH would give the 6-phosphate **3.47**.

To further validate this solvation sphere mechanism, a reaction was performed in an ionic liquid (Scheme 3.12). Ionic liquids have a highly ordered long range structure[49] and would therefore give large solvation spheres. Under these conditions, the 6-phosphate species **3.48** was formed exclusively (after per-*O*-acetylation to aid in purification).

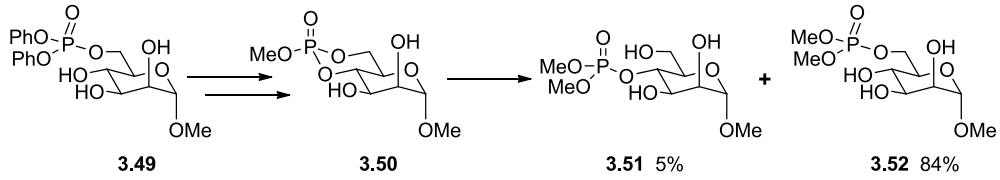


Scheme 3.12. *CsF mediated transesterfication in an ionic liquid produces the 6-phosphate exclusively. Conditions: NEt_3, MeOH, $MMIM^+PO(OMe)_2O^-$, 80˚C, 18h then Ac_2O, Py.*

This unusual ability of CsF to control reaction selectivity has been reported once before, where the diasteroselctivity of an intramolecular Michael addition was greatly enhanced by the presence of the salt.[50] In this case, a Cs-F ion pair was hypothesized as the cause, which has some parallels with our above solvation model.

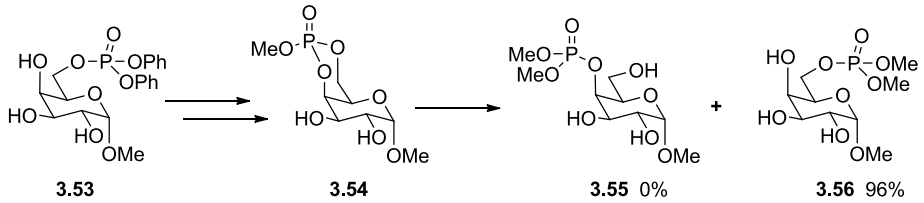
3.6 Extending the scope of transesterfications

To demonstrate the scope of our methodology, we used the CsF transesterfication in various other biologically important carbohydrates. As well as trehalose and glucose, for which CsF transesterfications have been illustrated above (section 3.4, page 74), the conditions were also used for mannose, galactose and ribose to prevent migration of the phosphoryl group from the 6-position. For mannose (Scheme 3.13), the reaction proceeded as before with the diphenylphosphate **3.49** rapidly converting to the cyclic methyl ester **3.50**, which ring-opens to give the 4-phophate **3.51** in 5% yield as the minor product and the 6-phosphate **3.52** as the major product in an excellent 84%.



Scheme 3.13. *Product distribution for the CsF mediated transesterfication in mannose. Conditions: 10eq CsF, MeOH, reflux, 16h.*

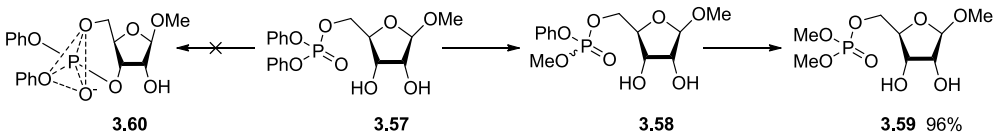
For the galactose case **3.53** (Scheme 3.14), the formation of the cyclic ester **3.54** was much slower and consequently, the reaction required 40 hours for completion. This lower propensity for cyclization is presumably due to the axial position of the 4-hydroxyl, which results in unfavourable 1,3 diaxial interactions with the phosphoryl group. This steric interaction leads to the 4-hydroxyl taking the apical position in the ring-opening step yielding the 6-phosphate **3.56** exclusively in 96%.



Scheme 3.14. *Product distribution for the CsF mediated transesterfication in galactose. Conditions: 10eq CsF, MeOH, reflux, 40h.*

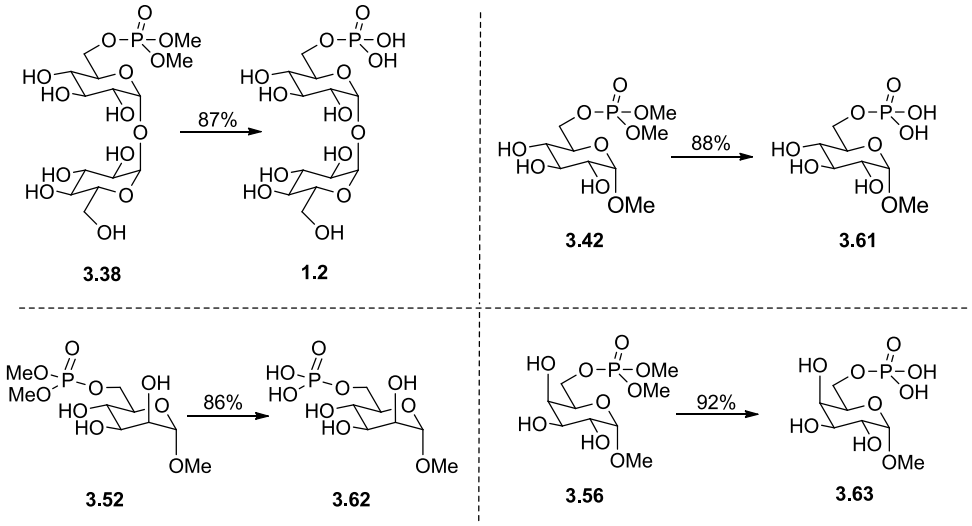
The reaction sequence was very different for the ribose diphenylphosphate **3.57**. In this case, the reaction did not proceed via a cyclic methyl ester, but was instead initiated by direct methanol attack at the phosphorous resulting in a mixed ester **3.58**. Since cylization did not occur, there was no opportunity for migration and the 5-phosphate **3.59** was produced in near quantitative yield. The lack of observed cyclization may be due to the greater rigidity of the 5-membered ribose ring. Ligand

exchanges at phosphorous require the entering group to enter from an apical position.[37] For cyclizations, this results in a 90˚ bond angle in the cyclic phosphate intermediate (**3.60**), which would be tolerable for a 6-membered ring such as glucose, but may be unfavourable for a 5-membered ring like ribose due to ring strain.

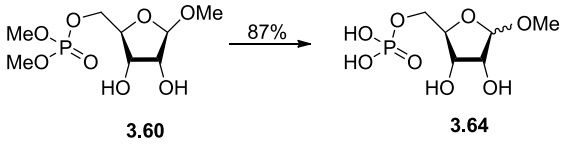


Scheme 3.15. *Product distribution for the CsF mediated transesterfication in ribose. Conditions: 10eq CsF, MeOH, reflux, 16h. The reaction does not go via a cyclic intermediate* ***3.60*** *due to ring strain, instead proceeding through direct methoxy attack.*

Deprotection of the dimethylphosphates was achieved with TMSI.[51] The reactions were rapid and were quenched within 15 minutes to prevent deprotection of the anomeric centre. For trehalose, glucose, mannose and galactose, excellent yields of greater than 85% were achieved in all cases (Scheme 3.16). For ribose, while deprotection of the phosphate was effective, the anomeric centre was racemized under the reaction conditions (Scheme 3.17).



Scheme 3.16. *Deprotection of demthyl phosphates. Conditions: TMSI, dioxane, RT, 15mins.*



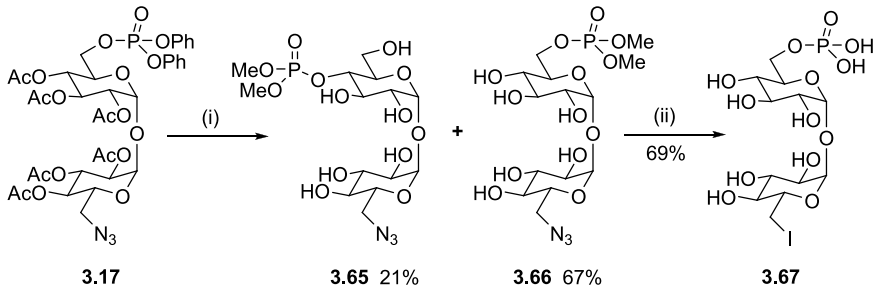
Scheme 3.17. *Deprotection of dimethyl phosphate in ribose leads to anomeric racemization. Conditions: TMSI, dioxane, RT, 5mins.*

Having developed a reaction to prevent migration from the 6-position, we briefly investigated conditions to capitalize on migration as a potentially facile route to 4-phosphate without the need for tedious protecting group manipulations. Various conditions were screened with **2.1** to maximize the yield of the 4-phosphate **3.41**. As mentioned above, triethylamine led to phosphate deprotection and so an alternative milder base was sought. Sodium carbonate was found to be the best reagent because its poor solubility in methanol produced only a mildly basic solution which prevented methoxy deprotection. Various solvents (such as DMF, dioxane and acetonitrile) were also screened and we found that a 1:1 combination of methanol and DCM produced the best results. Under these conditions, the transesterfication steps were relatively slow allowing better monitoring and greater control of the reaction mixture. This was especially important to limit migration to the 2 and 3 positions, which lowered yields and made purification more difficult.

As with other transesterfications, cyclization of the diphenyl phosphate was rapid giving species of the type **3.30**, which over time, produced the cyclic methoxy ester **3.43**. With Na_2CO_3 in DCM and methanol, **3.43** was far more stable than with CsF and required up to 80 hours for ring opening to give the desired dimethylphosphates. For galactose, the cyclic methoxy ester **3.54** only gave the 6-phosphate **3.56** due to 1,3 diaxial interactions. As with CsF, ribose did not produce a cyclic intermediate and so migration to the 3-position was not possible. For trehalose **2.1**, glucose **3.9** and mannose **3.49**, these conditions gave modest yields of 30-40% for the 4-phosphates, but in any case, the method may still offer a viable route compared to longer syntheses that are reliant on protecting groups.

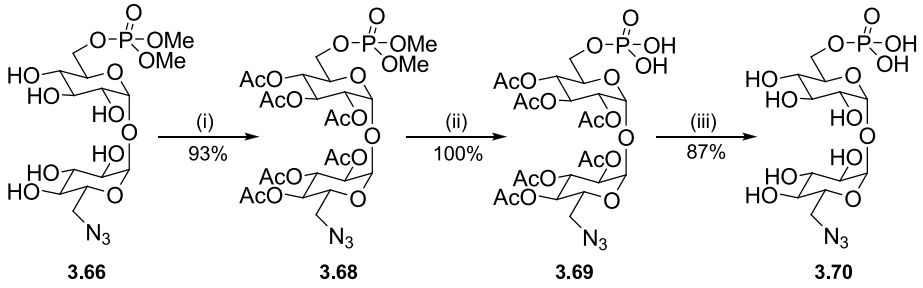
3.7 Synthetic route for 6’ modification of trehalose-6-phosphate

Having established methodology for the deprotection of diphenyl phosphate esters under non-hydrogenic conditions, we could return our attention to **3.17**. The phosphate deprotection could now be envisaged as two steps: a controlled transesterfication with methanol, followed by methoxy deprotection (Scheme 3.18).



Scheme 3.18. *Attempted deprotection of dimethyl phosphate leads to concominant azide displacement. Conditions: (i) CsF, MeOF, reflux, 15h; (ii) TMSI, dioxane, 0°C, 10mins.*

Transesterfication with CsF gave the 4-phosphate **3.65** as minor side product in 21% yield and the desired product **3.66** in a very good 67%. The deprotection of **3.66** was attempted with TMSI as above (section 3.6, page 81), but surprisingly, this led to azide displacement with an iodide giving **3.67** as the sole product. The deprotection was also attempted with the milder TMSBr, but the slower reaction time allowed phosphate migration from 6-position giving a mixture of inseparable products.



Scheme 3.19. *Alternative deprotection scheme where phosphate is deprotected prior to release of secondary hydroxyls. Conditions: (i) Ac_2O, Py; (ii) TMSBr, dioxane, RT, 14h; (ii) NaOMe, MeOH, RT.*

Acetyl protection of the secondary alcohols was required to prevent migration of the phosphate (Scheme 3.19). The protected analogue **3.68** could then be treated with TMSBr to give **3.69** in quantitative yield. Since the phosphate was fully deprotected and hence far less prone to nucleophilic attack, subsequent deacetylation with sodium methoxide could be performed without concern of phosphate migrations to give the target compound **3.70** in 87% yield.

By varying the nucleophilic substituent for the phosphate S_N2 substitution step, this general route of transesterfication and deprotection can be used for the synthesis of a wide variety of other Tre-6-P analogues which will be the subject of the next chapter.

3.8 Conclusions

The regiomeric control of reactions is important in carbohydrate systems, where the presence of multiple hydroxyls can produce many possible products. This becomes particularly imperative when working with functionality that is prone to migration such as, acetyl, silyl and phosphoryl groups. Here, we have demonstrated methodology that can offer some level of control for phosphate migrations.

These phosphate transesterfications show many parallels with regioselective reductive 4,6-*O*-benzylidene ring openings.[52] Under acid catalysis, electronic control dictates protonation at the more electron rich 4-*O* leading to the 6-ether; in phosphate migrations electronic effects dictate apical substituents. With catalysis from a bulky Lewis acid (such as TMSCl), steric effects overcome the electronic factors giving the 4-ether, via sterically more favourable 6-*O* activation. For the phosphate case, steric requirements of the surrounding solvent dictate the geometry of the central phosphorous. Use of CsF can modulate this sterical environment to maximize the yields of the 6-phosphate. Alternatively, Na_2CO_3 in DCM and methanol can be used to increase migrations, which can provide a facile route to 4-phosphates.

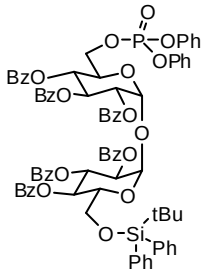
Here, we use phosphate transesterfications to exploit the differing chemical properties of the phenyl and methyl protecting groups for the phosphate moiety. Early use of phenyl protection allows for an efficient S_N2 reaction that significantly shortens the reaction sequence. Subsequently, our CsF transesterfication can allow exchange of the phenyl group for a methyl, which has orthogonal reactivity to other groups in the molecule. Crucially, our method minimizes phosphate migrations which can plague other synthetic routes. Thus, protecting group exchange is now a trivial step that can potentially be used in the synthesis of other carbohydrate phosphates.

3.9 Experimental

For general experimental conditions see section 2.8.1 page 44.

The stereochemistry at the phosphorous centre for compounds **3.40**, **3.36**, **3.50** and **3.54** were determined using NOESY and/or NOE NMR experiments.

3.9.1 Initial nucleophilic substitutions

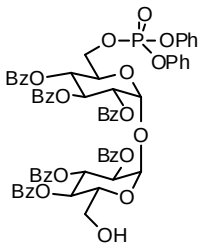


6-*O*-diphenoxyphosphoryl-6'-*O*-*tert*-butyl-diphenylsilyl-2,2',3,3',4,4'-benzoyl-D-trehalose (3.1)

Benzoyl chloride (1.25ml, 9.74mmol, 9eq) was added dropwise to a solution of **2.9** (0.88g, 1.08mmol, 1eq) in anhydrous pyridine (10ml) at room temperature under an Ar atmosphere. The reaction mixture was stirred for 16 hours, after which TLC (1 water : 4 isopropanol : 10 ethyl acetate) showed complete consumption of the starting compound. The reaction mixture was quenched with methanol (10ml) and the mixture concentrated *in vacuo*. The residue was partitioned between in ethyl acetate (25ml) and water (50ml). The aqueous layer was extracted with ethyl acetate (2 x 25ml) and the combined organics washed with 1M HCl (2 x 50ml), saturated $NaHCO_3$ solution (25ml), brine (2 x 50ml), dried over $MgSO_4$ and the solvent removed *in vacuo*. Silica gel chromatography (2 petrol : 1 ethyl acetate) gave the desired compound as a colourless oil (0.227g, 15%).

R_f 0.3 (2 petrol : 1 ethyl acetate), $[\alpha]_D^{20}$ +138.3 (c = 1.0, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 0.98 (9H, s, ^{t}Bu) 3.05 (1H, dd, $J_{H6'a\text{-}H6'b}$ 12.0Hz, $J_{H5\text{-}H6'b}$ 3.2Hz, H6'b), 3.13 (1H, d, $J_{H6'a\text{-}H6'b}$ 11.0Hz, H6'a), 3.74 (1H, ddd, $J_{H6a\text{-}H6b}$ 11.8Hz, $J_{H6b\text{-}31P}$ 6.4Hz, $J_{H5\text{-}H6b}$ 3.8Hz, H6b), 3.83 (1 H, dd, $J_{H6a\text{-}H6b}$ 12.0Hz, $J_{H6a\text{-}31P}$ 6.6Hz, H6a), 3.99 (1H, d, $J_{H4'\text{-}H5'}$ 10.1Hz, H5'), 4.13 (1H, d, $J_{H4\text{-}H5}$ 10.1Hz, H5), 5.22 (1H, dd, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H2\text{-}H3}$ 3.8Hz, H2), 5.44 (1H, dd, $J_{H2'\text{-}H3'}$ 10.2Hz, $J_{H1'\text{-}H2'}$ 3.9Hz, H2'), 5.53 (1H, d, $J_{H1'\text{-}H2'}$ 3.5Hz, H1'), 5.55 (1H, t, $J_{H3\text{-}H4}$ 10.1Hz, $J_{H4\text{-}H5}$ 10.1Hz, H4), 5.62 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 5.81 (1H, t, $J_{H3'\text{-}H4'}$ 9.9Hz, $J_{H4'\text{-}H5'}$ 9.9Hz), 6.17 (1H, t, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H3\text{-}H4}$ 9.8Hz, H3), 6.20 (1H, t, $J_{H2'\text{-}H3'}$ 10.1Hz, $J_{H3'\text{-}H4'}$ 10.1Hz, H3'), 7.11-7.63 (38H, m, ArH), 7.78, 7.81 (2 x 2H, at, $J_{ortho\text{-}meta}$ 6.9Hz, Bz ArH_{ortho}), 7.88, 7.90 (2 x 2H, at, $J_{ortho\text{-}meta}$ 8.5Hz, Bz ArH_{ortho}), 8.11, 8.13 (2 x 2H, at, $J_{ortho\text{-}meta}$ 6.3Hz, Bz ArH_{ortho}); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 14.2 (1C, $\underline{C}(CH_3)_3$), 26.6 (3C, $C(\underline{C}H_3)_3$), 60.9 (1C, C6'), 65.6 (1 C, d, $J_{C6\text{-}31P}$ 5.72Hz, C6), 67.7 (1C, C4), 67.9 (1C, C4'), 68.7 (1C, d, $J_{C5\text{-}31P}$ 7.63Hz, C5), 70.4 (1C, C3), 70.8 (1C, C5'), 70.9 (1C, C3'), 71.1 (1C, C2), 71.5 (1C, C2'), 92.6 (1C, C1'), 93.0 (1C, C1), 120.1 (2C, d, $J_{C\text{-}31P}$ 4.8Hz, P-OPh ArC_{para}), 125.3 (4C, d, $J_{C\text{-}31P}$ 8.6Hz, ArC_{ortho}), 127.4-130.1 (42C, m, ArC), 133.1, 133.2, 133.3, 133.5, 133.6, 133.6 (6 x 1C, 6 x Bz ArC_{para}), 135.4, 135.6 (2 x 2C, 4 x Si-Ph ArC_{meta}), 150.4 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, P-OPh ArC_{ipso}), 150.5 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, P-OPh ArC_{ipso}), 164.6, 165.0, 165.2, 165.5, 165.8, 170.9 (6 x 1C, C=O); FT-IR (thin film) υ 1277

(P=O), 1733 (C=O); HRMS m/z (ES^+) $[M + Na]^+$ peaks 1459.36 (100%), 1460.36 (91%), 1461.36 (43%), 1462.36 (13%), 1463.37 (4%), required 1459.41 (100%), 1460.41 (95%), 1461.42 (52%), 1462.42 (20%), 1463.42 (6%).

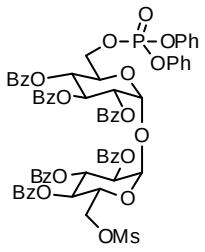


6-*O*-diphenoxyphosphoryl, 2,2',3,3',4,4'-benzoyl-D-trehalose (3.2)

Acetyl chloride (0.010ml, 1.39mmol, 10eq) was added dropwise to methanol (1ml), pre-cooled to 0ºC. After 5 minutes, the mixture was allowed to reach room temperature and transferred to a solution of **3.1** (0.20g, 0.139mmol, 1eq) in diethyl ether (5ml). The reaction mixture was heated to 50ºC for 68 hours after which TLC showed complete consumption of the starting material. The solvent was removed *in vacuo* and the residue partitioned between ethyl acetate (25ml) and water (25ml). The product was extracted from the aqueous layer with ethyl acetate (2 x 25ml) and the combined organics washed with saturated $NaHCO_3$ solution (25ml), brine (2 x 25ml), and dried over $MgSO_4$. Silica gel chromatography (2 petrol : 1 ethyl acetate) afforded the desired compound as a colourless oil (0.12g, 73%).

R_f 0.37 (1 petrol : 1 ethyl acetate), $[\alpha]_D^{19}$ +140.2 (c = 1.0, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.39 (1H, br s, 6'-OH), 2.95 (1H, dd, $J_{H6'a\text{-}H6'b}$ 13.2Hz, $J_{H5'\text{-}H6'b}$ 1.9Hz, H6'b), 3.16 (1H, d, $J_{H6'a\text{-}H6'b}$ 12.9Hz, H6'a), 3.71 (1H, ddd, $J_{H6a\text{-}H6b}$ 11.6Hz, $J_{H6b\text{-}31P}$ 6.7Hz, $J_{H5\text{-}H6b}$ 3.8 Hz, H6b), 3.82 (1H, dd, $J_{H6a\text{-}H6b}$ 10.6Hz, $J_{H6a\text{-}31P}$ 6.8Hz, H6a), 3.86 (1H, d, $J_{H4'\text{-}H5'}$ 10.1Hz, H5'), 4.16 (1H, m, H5), 5.21 (1H, dd, $J_{H2\text{-}H3}$ 10.2Hz, $J_{H1\text{-}H2}$ 3.6Hz, H2), 5.43 (1H, dd, $J_{H2'\text{-}H3'}$ 10.1Hz, $J_{H1'\text{-}H2'}$ 3.8Hz, H2'), 5.47 (1H, t, $J_{H3'\text{-}H4'}$ 9.9 Hz, $J_{H4'\text{-}H5'}$ 9.9Hz, H4'), 5.57 (1H, t, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, H4), 5.58 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 5.63 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 6.20 (1H, t, $J_{H2\text{-}H3}$ 9.8 Hz, $J_{H3\text{-}H4}$ 9.8Hz, H3), 6.29 (1H, t, $J_{H2'\text{-}H3'}$ 9.9Hz, $J_{H3'\text{-}H4'}$ 9.9Hz, H3'), 7.15-7.60 (28H, m, ArH), 7.82, 7.89, (2 x 2H, 2 x d, $J_{ortho\text{-}meta}$ 7.9Hz, 2 x Bz ArH_{ortho}), 7.93 (4H, d, $J_{ortho\text{-}meta}$ 7.9Hz, 2 x Bz ArH_{ortho}), 8.06, 8.09 (2 x 2H, 2 x d, $J_{ortho\text{-}meta}$ 7.6Hz, 2 x Bz ArH_{ortho}); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 59.8 (1C, C6'), 65.5 (1C, d, $J_{C6\text{-}31P}$ 5.7Hz, C6), 67.7 (1C, C4), 68.7 (1C, C4'), 68.8 (1C, d, $J_{C5\text{-}31P}$ 7.6Hz, C5), 70.0 (1C, C3'), 70.2 (1C, C3), 70.5 (1C, C5), 71.2 (1C, C2), 71.3 (1C, C2'), 92.7 (1C, C1), 93.1 (1C, C1'), 120.1 (2C, d, $J_{C\text{-}31P}$ 4.8Hz, P-OPh ArC_{para}), 125.3 (4C, d, $J_{C\text{-}31P}$ 8.6Hz, P-OPh ArC_{ortho}), 128.4, 128.5, 128.6, 128.8 (3 x 2C, 1 x 6C, 12 x Bz ArC_{meta}), 128.5, 128.6, 128.7, 128.7 129.0, 129.1 (6 x 1C, 6 x Bz ArC_{ipso}), 129.7, 129.7, 129.7, 129.8, 129.8

(5 x 2C, 1 x 4C, 12 x Bz ArC_{ortho}), 130.0 (4C, P-OPh ArC_{meta}), 133.2, 133.3, 133.5, 133.6, 133.7, 133.8 (6 x 1C, 6 x Bz ArC_{para}), 150.4 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, P-OPh ArC_{ipso}), 150.4 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, P-OPh ArC_{ipso}), 164.5, 165.1, 165.2, 165.6, 165.8, 166.2 (6 x 1C, C=O); FT-IR (thin film) υ 1277 (P=O), 1728 (C=O), 3681 (OH); HRMS m/z (ES^+) $[M + Na]^+$ peaks 1221.26 (100%), 1222.27 (70%), 1223.27 (24%), 1224.27 (7%), 1225.27 (2%), required 1221.29 (100%), 1222.30 (73%), 1223.30 (30%), 1224.30 (9%), 1225.30 (2%).



6-*O*-diphenoxyphosphoryl, 6'-*O*-mesyl, 2,2',3,3',4,4'-benzoyl-D-trehalose (3.3)

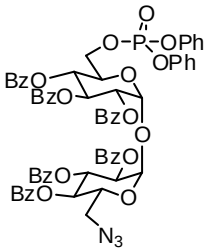
To a solution of **3.2** (0.05g, 0.042mmol, 1eq) in anhydrous pyridine (5ml) at 0ºC under argon was added dropwise mesyl chloride (0.005ml, 0.063mmol, 1.5eq). The reaction was maintained at this temperature for 60 minutes after which TLC indicated complete consumption of starting materials and formation of a single product. The reaction was quenched with methanol (5ml). The mixture was concentrated *in vacuo* and partitioned between water (25ml) and ethyl acetate (25ml). The aqueous layer was extracted with ethyl acetate (2 x 25ml) and the combined organics washed with 1M HCl (3 x 20ml), saturated $NaHCO_3$ solution (2 x 20ml), brine (2 x 20ml) and dried over $MgSO_4$. Removal of solvent under reduced pressure gave the title compound as a colourless oil (53mg, 99%).

R_f 0.27 (2 petrol : 1 ethyl acetate), $[\alpha]_D^{19}$ +119.7 (c = 0.75, $CHCl_3$); ^1H NMR (400MHz, $CDCl_3$) δ ppm 2.98 (3H, s, CH_3), 3.72 (1H, dd, $J_{H6a'\text{-}H6b'}$ 11.6Hz, $J_{H5'\text{-}H6b'}$ 2.0Hz, H6b'), 3.79 (2H, m, H6b, H6a'), 3.85 (1H, ddd, $J_{H6a\text{-}H6b}$ 11.6Hz, $J_{H6a\text{-}31P}$ 6.7Hz, $J_{H5\text{-}H6a}$ 2.0Hz, H6a), 4.15 (2H, m, H5, H5'), 5.24 (1H, dd, $J_{H2'\text{-}H3'}$ 10.2Hz, $J_{H1'\text{-}H2'}$ 3.8Hz, H2'), 5.43 (1H, dd, $J_{H2\text{-}H3}$ 10.2Hz, $J_{H1\text{-}H2}$ 4.1Hz, H2), 5.56 (1H, t, $J_{H3'\text{-}H4'}$10.1Hz, $J_{H4'\text{-}H5'}$ 10.1Hz, H4'), 5.58 (1H, t, $J_{H3\text{-}H4}$ 10.1Hz, $J_{H4\text{-}H5}$ 10.1Hz, H4), 5.60 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.64 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 6.21 (1H, t, $J_{H2'\text{-}H3'}$ 9.9Hz, $J_{H3'\text{-}H4'}$ 9.9Hz, H3'), 6.22 (1H, t, $J_{H2\text{-}H3}$ 9.9Hz, $J_{H3\text{-}H4}$ 9.9Hz, H3), 7.16-7.60 (28H, m, ArH), 7.81, 7.83 (2 x 2H, atd, $J_{ortho\text{-}meta}$ 6.7Hz, $J_{ortho\text{-}ortho}$ 1.0Hz, Bz ArH_{ortho}), 7.92, 7.95 (2 x 2H, atd, $J_{ortho\text{-}meta}$ 8.2Hz, $J_{ortho\text{-}ortho}$ 0.9Hz, Bz ArH_{ortho}), 8.10, 8.13 (2 x 2H, atd, $J_{ortho\text{-}meta}$ 9.7Hz, $J_{ortho\text{-}ortho}$ 1.4Hz, Bz ArH_{ortho}); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 37.6 (1C, CH_3), 65.5 (1C, C6'), 65.5 (1C, C6), 67.6 (1C, C4'), 67.9 (1C, C4), 68.2 (1C, C5'), 68.9 (1C, d, $J_{C5\text{-}31P}$ 7.6Hz, C5), 70.1 (1C, C3), 70.1 (1C, C3'), 71.0 (1C,

C2), 71.2 (1C, C2'), 92.5 (1C, C1) 92.6 (1C, C1'), 120.1 (2C, d, $J_{C\text{-}31P}$ 4.8Hz, P-OPh ArC_{para}), 125.4 (4C, d, $J_{C\text{-}31P}$ 8.6Hz, P-OPh ArC_{ortho}), 128.3-128.4 (12C, m, Bz ArC_{meta}), 128.6, 128.7, 128.7, 128.8, 128.9, 129.0 (6 x 1C, Bz ArC_{ipso}), 129.7-129.9 (16C, m, 12 x Bz ArC_{ortho}, 4 x P-OPh ArC_{meta}), 133.3, 133.3, 133.5, 133.7, 133.7, 133.8 (6 x 1C, 6 x Bz ArC_{para}), 150.3 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, P-OPh ArC_{ipso}), 150.4 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, P-OPh ArC_{ipso}), 164.5, 164.8, 165.1, 165.2, 165.4, 165.5 (6 x 1C, 6 x C=O); FT-IR (thin film) υ 1261 (P=O), 1731 (C=O); HRMS m/z (ES^+) $[M + Na]^+$ peaks 1299.24 (100%), 1300.24 (69%), 1301.24 (33%), 1302.24 (11%), 1303.25 (3%), required 1299.27 (100%), 1300.27 (75%), 1301.27 (37%), 1302.28 (13%), 1303.28 (4%).

6-*O*-diphenoxyphosphoryl-6'azido-2,2',3,3',4,4'-*O*-benzoyl-D-trehalose (3.4)
6,6'-diazido-2,2',3,3',4,4'-*O*-benzoyl-D-trehalose (3.5)

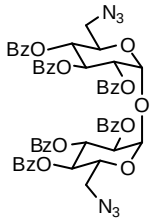
A mixture of **3.3** (50mg, 0.039mmol, 1eq) and sodium azide (10mg, 0.16mmol, 4eq) in DMF (3ml) was heated at 80°C. After 16 hours TLC (1 petrol : 1 ethyl acetate) showed the formation of two products. The reaction was allowed to cool to room temperature and was partitioned between water (20ml) and ethyl acetate (20ml). The aqueous layer was extracted with ethyl acetate (3 x 20ml) and the combined organics washed with brine (5x 20ml), dried over $MgSO_4$ and the solvent removed *in vacuo*. Separation of the products with silica gel chromatography (2 petrol : 1 ethyl acetate) gave **3.4** (16.7mg, 35%) and **3.5** (5.2mg, 13%) as colourless oils.



Compound **3.4**

R_f 0.77 (1 petrol : 1 ethyl acetate), $[\alpha]_D^{19}$ +128.6 (c = 1.0, $CHCl_3$); ^1H NMR (500MHz, $CDCl_3$) δ ppm 2.84 (1H, dd, $J_{H6'a\text{-}H6'b}$ 13.9Hz, $J_{H5'\text{-}H6'b}$ 2.8Hz, H6'b), 2.88 (1H, dd, $J_{H\text{-}6'a\text{-}H6'b}$ 13.7Hz, $J_{H5'\text{-}H6'a}$ 5.0Hz, H6'a), 3.75 (1H, ddd, $J_{H6a\text{-}H6b}$ 11.3Hz, $J_{H6b\text{-}31P}$ 6.6Hz, $J_{H5\text{-}H6b}$ 4.1Hz, H6b), 3.84 (1H, dd, $J_{H6a\text{-}H6b}$ 11.7Hz, $J_{H6a\text{-}31P}$ 7.3Hz, H6a), 4.08 (1H, dt, $J_{H4'\text{-}H5'}$ 9.8Hz, $J_{H5'\text{-}H6'a}$ 9.8Hz, $J_{H5'\text{-}H6'b}$ 3.8Hz, H5'), 4.13 (1H, d, $J_{H4\text{-}H5}$ 8.5Hz, H5), 5.26 (1H, dd, $J_{H2'\text{-}H3'}$ 10.2Hz, $J_{H1'\text{-}H2'}$ 3.6Hz, H2'), 5.42 (1H, dd, $J_{H2\text{-}H3}$ 10.2Hz, $J_{H1\text{-}H2}$ 3.6Hz, H2), 5.50-5.60 (2H, m, H4, H4'), 5.58 (2H, d, $J_{H1\text{-}H2}$ 3.8 Hz, H1, H1'), 6.20 (2H, t, $J_{H2\text{-}H3}$ 9.9Hz, $J_{H3\text{-}H4}$ 9.9Hz, H3, H3'), 7.17-7.58 (28H, m, ArH), 7.81 (2H, d, $J_{ortho\text{-}meta}$ 7.6Hz, Bz ArH_{ortho}), 7.83 (2H, d, $J_{ortho\text{-}meta}$ 7.9Hz, Bz ArH_{ortho}), 7.91, 7.93 (2 x 2H, at, $J_{ortho\text{-}meta}$ 7.6Hz, 2 x Bz ArH_{ortho}), 8.09 (2H, d, $J_{ortho\text{-}meta}$ 7.9Hz, Bz ArH_{ortho}), 8.11 (2H,

d, $J_{ortho\text{-}meta}$ 7.6Hz, Bz ArH_{ortho}); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 50.0 (1C, C6') 65.6 (1C, d, $J_{C6\text{-}31P}$ 5.7Hz, C6), 67.7 (1C, C4), 68.9 (1C, d, $J_{C5\text{-}31P}$ 7.6Hz, C5), 69.0 (1C, C4'), 69.7 (1C, C5'), 70.1 (1C, C3'), 70.2 (1C, C3), 71.1 (1C, C2), 71.1 (1C, C2'), 92.5 (1C, C1'), 92.6 (1C, C1), 120.1 (2C, d, $J_{C\text{-}31P}$ 4.8Hz, P-OPh ArC_{para}), 125.4 (4C, d, $J_{C\text{-}31P}$ 8.6Hz, P-OPh ArC_{ortho}), 128.3-128.4 (12C, m, Bz ArC_{meta}), 128.6-128.7 (6C, m, Bz ArC_{ipso}), 128.7-129.9 (16C, m, Bz ArC_{ortho}, P-OPh ArC_{meta}), 133.2, 133.5, 133.6, 133.7 (2 x 1C, 2 x 2C, 4 x Bz ArC_{para}), 150.4 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, P-OPh ArC_{ipso}), 150.4 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, P-OPh ArC_{ipso}), 164.6, 164.9, 165.1, 165.2, 165.5, 165.5 (6 x 1C, C=O); FT-IR (thin film) υ 1275 (P=O), 1733 (C=O), 2160 (N_3); HRMS m/z (ES^+) $[M + Na]^+$ peaks 1246.26 (100%), 1247.27 (72%), 1248.27 (27%), 1249.27 (8%), required 1246.30 (100%), 1247.30 (74%), 1248.30 (31%), 1249.31 (9%).



Compound **3.5**

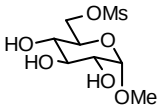
R_f 0.8 (1 petrol : 1 ethyl acetate), $[\alpha]_D^{19}$ +151.7 (c = 1.0, $CHCl_3$); ^1H NMR (400MHz, $CDCl_3$) δ ppm 2.87 (4H, d, $J_{H5\text{-}H6}$ 3.8Hz, H6, H6'), 4.11 (2H, dt, $J_{H4\text{-}H5}$ 10.1Hz, $J_{H5\text{-}H6}$ 3.8Hz, $J_{H5\text{-}H6'}$ 3.8Hz, H5), 5.44 (2H, dd, $J_{H2\text{-}H3}$ 10.2Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 5.56 (2H, t, $J_{H3\text{-}H4}$ 9.9Hz, $J_{H4\text{-}H5}$ 9.9Hz, H4), 5.75 (2H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 6.23 (2H, t, $J_{H2\text{-}H3}$ 9.9Hz, $J_{H3\text{-}H4}$ 9.9Hz, H3), 7.27-7.60 (18H, m, ArH), 7.85, 7.93 (2 x 4H, dd, $J_{ortho\text{-}meta}$ 7.3Hz, $J_{ortho\text{-}ortho}$ 1.3Hz, 4 x Bz ArH_{ortho}), 8.11 (4H, dd, $J_{ortho\text{-}meta}$ 7.3Hz, $J_{ortho\text{-}ortho}$ 1.5Hz, 2 x Bz ArH_{ortho}); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 50.0 (2C, C6), 69.0 (2C, C4), 69.6 (2C, C5), 70.1 (2C, C3), 71.2 (2C, C2), 92.8 (2C, C1), 128.4, 128.4, 128.6 (3 x 4C, 12 x ArC_{meta}), 128.5, 128.5, 129.0 (3 x 2C, 6 x ArC_{ipso}), 129.7, 129.9 (1 x 4C, 1 x 8C, ArC_{ortho}), 133.3, 133.6, 133.7 (3 x 2C, 6 x ArC_{ortho}), 164.9, 165.2, 165.5 (3 x 2C, C=O); FT-IR (thin film) υ 1732 (C=O), 2105 (N_3); HRMS m/z (ES^+) $[M + Na]^+$ peaks 1039.26 (100%), 1040.26 (60%), 1041.26 (20%), 1042.27 (5%), required 1039.28 (100%), 1040.28 (62%), 1041.28 (22%), 1042.28 (5%).

3.9.2 Nucleophilic substitutions on the glucose model

α-methyl-6-mesyl-D-glucose pyranoside[53] (3.7)

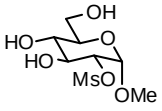
α-methyl-2-mesyl-D-glucose pyranoside[54] (3.8)

To a solution of α-methyl glucose pyranoside (0.10g, 0.52mmol, 1eq) in anhydrous pyridine (5ml) under Ar was added mesyl chloride (40µl, 0.52mmol, 1eq). The reaction mixture was stired at room temperature for 1 hour, after which TLC (1 water : 4 isopropanol : 4 ethyl acetate) showed the complete consumption of starting materials. The reaction was quenched with methanol (1ml) and the solvent removed *in vacuo*. The crude residue was co-evaporated with toluene (3 x 25ml) to remove pyridine and silica gel chromatography (95 ethyl acetate : 5 methanol) allowed the pure isolation of **3.7** (149mg, 53%) and **3.8** (90mg, 32%) as colourless oils.



*Compound **3.7***

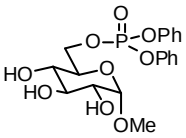
R_f 0.57 (1 water : 4 isopropanol : 4 ethyl acetate), $[\alpha]_D^{22}$ +118.2 (c = 1.0, H_2O), lit.[53] $[\alpha]_D^{20}$ +115 (c = 1.0, H_2O); ^{1}H NMR (400MHz, MeOD) δ ppm 3.11 (3H, s, S-Me), 3.31 (1H, t, $J_{H3\text{-}H4}$ 9.9Hz, $J_{H4\text{-}H5}$ 9.9Hz, H4), 3.41 (2H, dd, $J_{H2\text{-}H3}$ 9.9Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.43 (3H, s, OMe), 3.63 (1H, t, $J_{H2\text{-}H3}$ 9.3Hz, $J_{H3\text{-}H4}$ 9.3Hz, H3), 3.76 (1H, ddd, $J_{H4\text{-}H5}$ 10.0Hz, $J_{H5\text{-}H6'}$ 5.4Hz, $J_{H5\text{-}H6}$ 1.8Hz, H5), 4.39 (1H, dd, $J_{H6\text{-}H6'}$ 11.1Hz, $J_{H5\text{-}H6'}$ 5.6Hz, H6'), 4.50 (1H, dd, $J_{H6\text{-}H6'}$ 11.1Hz, $J_{H5\text{-}H6}$ 1.8Hz, H6), 4.70 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (101MHz, MeOD) δ ppm 36.3 (1C, S-Me), 54.8 (1C, OMe), 69.8 (1C, C6), 70.2 (1C, C5), 70.3 (1C, C4), 72.4 (1C, C2), 74.0 (1C, C3), 100.4 (1C, C1); m/z (ES^+) 290.1 $[M + NH_4]^+$, 336.1 $[M + CH_3CN + Na]$, 567.1 $[2M + Na]^+$ (ES^-) 271.0 $[M - H]^-$, 307.0 $[M + {}^{35}Cl]^-$, 331.1 $[M + AcO]^-$, 543.1 $[2M - H]^-$, 579.1 $[2M + {}^{35}Cl]^-$, 581.1 $[2M + {}^{37}Cl]^-$.



*Compound **3.8***

R_f 0.71 (1 water : 4 isopropanol : 4 ethyl acetate), $[\alpha]_D^{22}$ +60.2 (c = 1.0, MeOH), lit.[54] $[\alpha]_D^{25}$ +64.5 (c = 1.3, MeOH); ^{1}H NMR (400MHz, MeOD) δ ppm 3.16 (3H, s, S-Me), 3.40 (1H, dd, $J_{H4\text{-}H5}$ 9.5Hz, $J_{H3\text{-}H4}$ 9.1Hz, H4), 3.44 (3H, s, OMe), 3.57 (1H, ddd, $J_{H4\text{-}H5}$ 9.8Hz, $J_{H5\text{-}H6'}$ 5.6Hz, $J_{H5\text{-}H6}$ 2.3Hz, H5), 3.70 (1H, dd, $J_{H6\text{-}H6'}$ 11.9Hz, $J_{H5\text{-}H6'}$ 5.6Hz, H6'), 3.81 (1H, dd, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H3\text{-}H4}$ 9.0Hz, H3), 3.83 (1H, dd, $J_{H6\text{-}H6'}$ 11.8Hz, $J_{H5\text{-}H6}$ 2.3Hz, H6), 4.30 (1H, dd, $J_{H2\text{-}H3}$ 9.9Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 4.90 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (101MHz, MeOD) δ ppm 37.2 (1C, S-Me), 54.7 (1C, OMe), 61.3 (1C, C6), 70.8 (1C, C3), 71.3 (1C, C4), 72.5 (1C, C5), 80.5 (1C, C2), 98.3 (1C, C1); m/z (ES^+) 290.1 $[M + NH_4]^+$, 295.0 $[M + Na]^+$ 336.0 $[M + CH_3CN +$

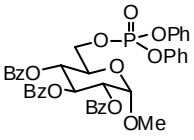
Na], (ES$^-$) 307.0 [M + ^{35}Cl]$^-$, 309.1 [M + ^{37}Cl]$^-$, 543.1 [2M – H]$^-$, 579.0 [2M + ^{35}Cl]$^-$, 581.0 [2M + ^{37}Cl]$^-$.



α-methyl-6-*O*-diphenoxyphosphoryl-D-glucose pyranoside[55] (3.9)

To a solution of α-methyl glucose pyranoside (1.00g, 5.15mmol, 1eq) and DMAP (0.066g, 0.51mmol, 0.1eq) in anhydrous pyridine (5ml) at room temperatur was addedd dropwise diphenyl chlorophosphate (1.07ml, 5.15mmol, 1eq). TLC (1 water : 4 isopropanol : 4 ethyl acetate) after 7 hours showed complete consumption of the starting materials. The reaction was quenched with methanol (5ml) and the solvent removed *in vacuo*. The residue was co-evaporated three times with toluene to remove pyridine. Purification by silica gel chromatography (97 ethyl acetate : 3 methanol) gave the title compound as a pale yellow oil (1.22g, 55%).

R_f 0.84 (1 water : 4 isopropanol : 4 ethyl acetate), $[\alpha]_D^{19}$ +60.4 (c = 1.0, $CHCl_3$) lit.[55] $[\alpha]_D^{20}$ +62.1 (c = 0.9, $CHCl_3$); ^{1}H NMR (500MHz, MeOD) δ ppm 3.31 (1H, t, J_{H3-H4} 9.1Hz, J_{H4-H5} 9.1Hz, H4), 3.34 (1H, m, H2), 3.35 (3H, s, OMe), 3.63 (1H, t, J_{H2-H3} 9.3Hz, J_{H3-H4} 9.3Hz, H3), 3.75 (1H, add, J_{H4-H5} 9.8Hz, $J_{H5-H6'}$ 5.4Hz, H5), 4.45 (1H, ddd, $J_{H6-H6'}$ 11.0Hz, $J_{H6'-31P}$ 8.0Hz, $J_{H5-H6'}$ 5.4Hz, H6'), 4.59 (1H, ddd, $J_{H6-H6'}$ 10.9Hz, J_{H6-31P} 6.9Hz, J_{H5-H6} 1.6Hz, H6), 4.64 (1H, d, J_{H1-H2} 3.5Hz, H1), 7.26-7.28 (6H, m, ArH_{ortho}, ArH_{para}), 7.47 (4H, t, $J_{ortho-meta}$ 7.9Hz, $J_{meta-para}$ 7.9Hz, ArH_{meta}); ^{13}C NMR (126MHz, MeOD) δ ppm 55.7 (1C, OMe), 70.0 (1C, d, J_{C6-31P} 5.7Hz, C6), 71.2 (1C, C4), 71.8 (1C, d, J_{C5-31P} 6.7Hz, C5), 73.3 (1C, C2), 75.0 (1C, C3), 101.4 (1C, C1), 121.2 (4C, d, J_{C-31P} 4.8Hz, ArC_{ortho}), 126.8 (2C, ArC_{para}), 131.06 (4C, ArC_{meta}), 151.88 (1C, d, J_{C-31P} 7.6Hz, ArC_{ipso}), 151.90 (1C, d, J_{C-31P} 6.7Hz, ArC_{ipso}); ^{31}P{^{1}H} NMR (162MHz, MeOD) δ ppm -11.8 (1P, $ROP(O)(OPh)_2$); m/z (ES$^+$) 427.2 [M + H]$^+$, 875.2 [2M + Na]$^+$.

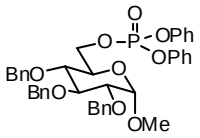


α-methyl-6-*O*-diphenoxyphosphoryl, 2,3,4-*O*-benzoyl-D-glucose pyranoside (3.10)

To a solution of **3.9** (50mg, 0.12mmol, 1eq) in dry pyridine (3ml) was added dropwise benzoyl chloride (0.08ml, 0.70mmol, 6eq). The reaction was stirred at room temperature for 2 hours after which TLC (95 ethyl acetate : 5 methanol) showed complete consumption of starting material. The reaction was quenched with methanol (2ml) and the solvent removed under reduced pressure. The

residue was partitioned between ethyl acetate (20ml) and water (20ml). The aqueous layer was extracted with ethyl acetate (2 x 10ml) and the combined organics washed with 1M HCl (3 x 20ml), saturated $NaHCO_3$ solution (20ml), brine (25ml), dried over $MgSO_4$ and the solvent removed *in vacuo*. The resulting residue was purified by silica gel chromatography (1 petrol : 1 ethyl acetate) to afford the title compound as a colourless oil (63mg, 71%).

R_f 0.2 (1 petrol : 1 ethyl acetate), $[\alpha]_D^{25}$ +40.9 (c = 1.0, $CHCl_3$); ^1H NMR (500MHz, $CDCl_3$) δ ppm 3.40 (1H, s, OMe), 4.30 (1H, ddd, J_{H4-H5} 10.6Hz, $J_{H5-H6'}$ 5.2Hz, J_{H5-H6} 1.3Hz , H5), 4.40 (1H, ddd, $J_{H6-H6'}$ 10.7Hz, $J_{H5-H6'}$ 5.1Hz, $J_{6'-31P}$ 2.5Hz, H6'), 4.47 (1H ddd, $J_{H6-H6'}$ 11.5Hz, J_{H6-31P} 7.1Hz, J_{H5-H6} 2.5Hz, H6), 5.16 (1H, dd, J_{H2-H3} 7.3Hz, J_{H1-H2} 3.8Hz, H2), 5.17 (1H, s, H1), 5.50 (1H, dd, J_{H4-H5} 10.1Hz, J_{H3-H4} 9.5 Hz, H4), 6.14 (1H, t, J_{H2-H3} 9.6Hz, J_{H3-H4} 9.6Hz, H3), 7.15-7.54 (19H, m, ArH), 7.87 (2H, dd, $J_{ortho-meta}$ 8.4Hz, $J_{ortho-ortho}$ 1.1Hz, Bz ArH_{ortho}), 7.94 (2H, dd, $J_{ortho-meta}$ 8.2Hz, $J_{ortho-ortho}$ 1.3Hz, Bz ArH_{ortho}), 7.98 (2H, dd, $J_{ortho-meta}$ 8.4Hz, $J_{ortho-ortho}$ 1.4Hz, Bz ArH_{ortho}) ; ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 55.7 (1C, OMe), 67.3 (1C, d, J_{C6-31P} 5.7Hz, C6), 68.3 (1C, d, J_{C5-31P} 6.7Hz, C5), 69.0 (1C, C4), 70.2 (1C, C3), 71.9 (1C, C2), 96.9 (1C, C1), 120.1 (2C, d, J_{C-31P} 3.8Hz P-OPh ArC_{para}), 125.4 (4C, d, J_{C-31P} 5.7Hz, P-OPh ArC_{ortho}), 128.3, 128.4, 128.4 (3 x 2C, 3 x Bz ArC_{meta}), 128.7, 129.0, 129.1 (3 x 1C, 3 x Bz ArC_{ipso}), 129.6, 129.7, 129.8 (3 x 2C, 3 x Bz ArC_{ortho}), 129.9 (2C, d, J_{C-31P} 3.8Hz, P-OPh ArC_{meta}), 133.1, 133.4, 133.5 (3 x 1C, 3 x Bz ArC_{ortho}), 150.4 (1C, d, J_{C-31P} 7.6Hz, P-OPh ArC_{ipso}), 150.5 (1C, d, J_{C-31P} 6.7Hz, P-OPh ArC_{ipso}), 165.2, 165.7, 165.7 (3 x 1C, 3 x C=O); FT-IR (thin film) υ 1279 (P=O), 1730 (C=O); HRMS m/z (ES^+) 761.1765 $[M + Na]^+$ (required 761.1758).

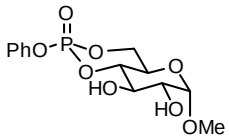


α-methyl-6-*O*-diphenoxyphosphoryl-2,3,4-benzyl-D-glucose pyranoside (3.11)

The reaction was performed under standard Schlenk conditions using argon gas. Benzyl trichloroacetimidate (0.12ml, 0.63mmol, 9eq) was added to a stirred solution of **3.9** (30mg, 0.07mmol, 1eq) in anhydrous DCM (3ml) cooled to -40°C in the presence of 3Å molecular sieves. To this solution was added trimethylsilyl triflate (0.01ml, cat.). The reaction mixture was allowed to warm to room temperatur over a period of 21 hours, after which TLC (2 petrol : 1 ethyl acetate) showed complete consumption of the starting material. The reaction was quenched with methanol (1ml)

and the solvent removed *in vacuo*. The residue was taken up in ethyl acetate (50ml) and washed with 1M NaOH (25ml), 1M HCl (25ml), saturated $NaHCO_3$ (25ml), brine (25ml), decolourised with activated charcoal, dried over $MgSO_4$ and the solvent removed *in vacuo*. Silica gel chromatography (2 petrol : 1 ethyl acetate) gave the desired compound as a pale yellow oil (27mg, 55%).

R_f 0.28 (2 petrol : 1 ethyl acetate), $[\alpha]_D^{19}$ +7.6 (c = 1.35, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 3.32 (3H, s, OMe), 3.45 (1H, dd, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.46 (1H, t, $J_{H3\text{-}H4}$ 10.1Hz, $J_{H4\text{-}H5}$ 10.1Hz, H4) 3.77 (1H, m, H5), 3.99 (1H, t, $J_{H2\text{-}H3}$ 9.1Hz, $J_{H3\text{-}H4}$ 9.1Hz, H3), 4.38 (1H, ddd, $J_{H6\text{-}H6'}$ 11.0Hz, $J_{H6'\text{-}31P}$ 6.9Hz, $J_{H5\text{-}H6'}$ 4.4Hz, H6'), 4.46 (1H, ddd, $J_{H6\text{-}H6'}$ 11.1Hz, $J_{H6\text{-}31P}$ 6.3Hz, $J_{H5\text{-}H6}$ 1.9Hz, H6), 4.49, 4.99 (2 x 1H, 2 x d, ^2J 11.0Hz, $PhCH_2$), 4.54 (1H, d, $J_{H1\text{-}H2}$ 3.5Hz, H1), 4.65, 4.78 (2 x 1H, 2 x d, ^2J 12.3Hz, $PhCH_2$), 4.81, 4.83 (2 x 1H, 2 x d, ^2J 7.3Hz, $PhCH_2$), 6.83-7.52 (25H, m, ArH); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 55.3 (1C, OMe), 67.7 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, C6), 69.3 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, C5), 73.4, 75.1, 75.7 (3 x 1C, 3 x $PhCH_2$), 76.8 (1C, C2), 79.7 (1C, 4C), 81.8 (1C, C3), 98.0 (1C, C1), 120.1 (4C, d, $J_{C\text{-}31P}$ 4.8Hz, P-OPh ArC_{ortho}), 125.3 (2C, d, $J_{C\text{-}31P}$ 6.7Hz, P-OPh ArC_{para}), 127.7-128.5 (15C, m, Bn ArC_{ortho}, ArC_{meta}, ArC_{para}), 129.7 (4C, d, $J_{C\text{-}31P}$ 6.7Hz, P-OPh ArC_{meta}), 137.9, 138.0, 138.6 (3 x 1C, Bn ArC_{ipso}), 150.5 (2C, at, $J_{C\text{-}31P}$ 7.2Hz, P-OPh ArC_{ipso}); FT-IR (thin film) υ 1260 (P=O); HRMS (ES^+) m/z 719.2380 $[M + Na]^+$ (required 719.2380).

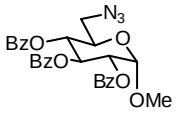


α-methyl-6-*O*-4-*O*-cyclophenoxyphosphoryl-D-glucose pyranoside[56] (3.35)

A mixture of **3.9** (0.03g, 0.07mmol, 1eq) and sodium azide (4.6mg, 0.07mmol, 1eq) in DMF (2ml) were heated to 80°C. After 7 hours TLC (95 ethyl acetate : 5 methanol) showed complete consumption of starting materials. The solvent was removed *in vacuo* and the product isolated by silica gel chromatography (98 ethyl acetate : 2 methanol). Lyophilization afforded the title compound as an amorphous white solid (17mg, 74%).

R_f 0.5 (95 ethyl acetate : 5 methanol), $[\alpha]_D^{19}$ +94.8 (c = 0.85, EtOH) lit.[56] $[\alpha]_D^{20}$ +100.8 (c = 1.0, EtOH), m.p. 170.2-171.9°C lit[56] 196-197°C; ^{1}H NMR (500MHz, MeOD) δ ppm 3.48 (3H, s, OMe), 3.55 (1H, dd, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.89 (1H, t, $J_{H2\text{-}H3}$ 9.1Hz, $J_{H3\text{-}H4}$ 9.1Hz, H3), 4.11 (1H, td, $J_{H4\text{-}H5}$ 10.2Hz, $J_{H5\text{-}H6'}$ 10.2Hz, $J_{H5\text{-}H6}$ 4.9 Hz, H5), 4.23 (1H, t, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, H4), 4.41 (1H, t, $J_{H5\text{-}H6'}$ 10.4Hz, $J_{H6\text{-}H6'}$ 10.4Hz, H6'), 4.52 (1H, ddd, $J_{H6\text{-}31P}$ 24.0Hz, $J_{H6\text{-}H6'}$ 10.5Hz, $J_{H5\text{-}H6}$

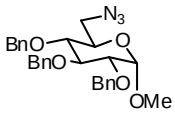
4.7Hz, H6), 4.79 (1H, d, J_{H1-H2} 3.8Hz, H1), 7.27 (1H, t, 2 x $J_{meta-para}$ 7.4Hz, ArH_{meta}), 7.33 (2H, d, $J_{meta-para}$ 7.9Hz, ArH_{para}), 7.43 (2H, t, $J_{ortho-meta}$ 7.9Hz, $J_{ortho-para}$ 7.9Hz, ArH_{ortho}); ^{13}C NMR (126MHz, MeOD) δ ppm 56.3 (1C, OMe), 63.6 (1C, d, J_{C5-31P} 6.7Hz, C5), 71.2 (1C, d, J_{C6-31P} 8.6Hz, C6), 72.1 (1C, d, J_{C3-31P} 8.6Hz, C3), 73.2 (1C, d, J_{C-31P} 2.9Hz, C2), 83.1 (1C, d, J_{C4-31P} 6.7Hz, C4), 101.9 (1C, C1), 120.9 (2C, d, J_{C-31P} 4.8Hz, ArC_{ortho}), 126.8 (1C, ArC_{para}), 131.2 (2C, ArC_{meta}), 151.6 (1C, d, J_{C-31P} 6.7Hz, ArC_{ipso}); $^{31}P\{^1H\}$ NMR (16MHz, MeOD) δ ppm -11.87 (1P, (RO)(R'O)P(O)(OPh)); FT-IR (KBr disc) υ 1207 (P=O), 3404 br (OH); m/z (ES^-) 331.4 $[M - H]^-$.



α-methyl-6-azido-2,3,4-*O*-benzoyl-D-glucose pyranoside (3.13)

A mixture of **3.10** (50mg, 0.068mmol, 1eq) and sodium azide (8.8mg, 0.14mmol, 2eq) in DMF (3ml) was aggressively stirred at 80°C. TLC (1 petrol : 1 ethyl acetate) after 8 hours showed the complete consumption of starting material. The reaction mixture was concentrated under reduced pressure, and the crude residue partitioned between ethyl acetate (20ml) and water (20ml). The aqueous layer was extracted with ethyl acetate (2 x 10ml) and the combined organics washed with brine (3 x 30ml), dried over $MgSO_4$ and the solvent removed *in vacuo*. Purification with silica gel chromatography (3 petrol : 1 ethyl acetate) afforded the desired compound as a colourless oil (22mg, 61%).

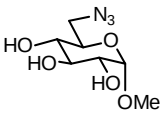
R_f 0.3 (2 petrol : 1 ethyl acetate), $[\alpha]_D^{19}$ +60.5 (c = 1.0, $CHCl_3$); ^1H NMR (500MHz, $CDCl_3$) δ ppm 3.42 (1H, dd, $J_{H6-H6'}$ 13.5Hz, $J_{H5-H6'}$ 2.5Hz, H6'), 3.51 (1H, dd, $J_{H6-H6'}$ 13.7Hz, J_{H5-H6} 6.9Hz, H6), 3.53 (1H, s, OMe), 4.25 (1H, ddd, J_{H4-H5} 9.9Hz, J_{H5-H6} 6.9Hz, $J_{H5-H6'}$ 2.5Hz, H5), 5.27 (1H, s, H1), 5.30 (1H, d, J_{H1-H2} 4.0Hz, H2), 5.52 (1H, t, J_{H3-H4} 9.8Hz, J_{H4-H5} 9.8Hz, H4), 6.16 (1H, t, J_{H3-H4} 9.6Hz, J_{H2-H3} 9.6Hz, H3), 7.07-7.55 (9H, m, ArH), 7.87, 7.96 (2 x 2H, 2 x dd, $J_{ortho-meta}$ 8.4Hz, $J_{ortho-ortho}$ 1.1Hz, 2 x ArH_{ortho}), 7.99 (2H, dd, $J_{ortho-meta}$ 8.2Hz, $J_{ortho-ortho}$ 1.3Hz, ArH_{ortho}), ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 51.2 (1C, C6), 55.8 (1C, OMe), 69.1 (1C, C5), 70.1 (1C, C3), 70.1 (1C, C4), 71.9 (1C, C2), 97.0 (1C, C1), 128.3, 128.4, 128.5 (3 x 2C, ArC_{meta}), 128.6, 129.0, 129.1 (3 x 1C, ArC_{ipso}), 129.6, 129.9, 129.9 (3 x 2C, ArC_{ortho}), 133.1, 133.4, 133.6 (3 x 1C, ArC_{para}), 165.4, 165.7, 165.8 (3 x 1C, C=O); FT-IR (thin film) υ 1726 (C=O), 2103 (N_3); HRMS m/z (ES^+) 554.1535 $[M + Na]^+$ (required 554.1534).



α-methyl-6-azido-2,3,4-*O*-benzyl-D-glucose pyranoside[57] (3.14)

A mixture of **3.11** (50mg, 0.072mmol, 1eq) and sodium azide (5mg, 0.072mmol, 1eq) in DMF (3ml) was heated to 80°C. The reaction was monitored by TLC (2 petrol : 1 ethyl acetate) and required 78 hours for completion. The reaction mixture was diluted with ethyl acetate (25ml) and partitioned with water (25ml). The aqueous layer was extracted with ethyl acetate (3 x 20ml) and the combined organics washed with brine (5 x 20ml), dried over $MgSO_4$ and the solvent removed *in vacuo*. The compound was purified by silica gel chromatography (9 petrol : 1 ethyl acetate) as a colourless oil (27mg, 77%).

R_f 0.24 (9 petrol : 1 ethyl acetate); $[\alpha]_D^{19}$ +28.8 (c = 1.35, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 3.33 (1H, dd, $J_{H6\text{-}H6'}$ 13.2Hz, $J_{H5\text{-}H6'}$ 5.7Hz, H6'), 3.41 (3H, s, OMe), 3.44 (1H, t, $J_{H3\text{-}H4}$ 9.3Hz, $J_{H4\text{-}H5}$ 9.3Hz, H4), 3.45 (1H, dd, $J_{H6\text{-}H6'}$ 13.5Hz, $J_{H5\text{-}H6}$ 2.5Hz, H6), 3.55 (1H, dd, $J_{H2\text{-}H3}$ 9.6Hz, $J_{H1\text{-}H2}$ 3.6Hz, H2), 3.79 (1H, ddd, $J_{H4\text{-}H5}$ 9.9Hz, $J_{H5\text{-}H6'}$ 5.8Hz, $J_{H5\text{-}H6}$ 2.4Hz, H5), 3.99 (1H, t, $J_{H2\text{-}H3}$ 9.3Hz, $J_{H3\text{-}H4}$ 9.3Hz, H3), 4.58, 5.00 (2 x 1H, d, ^2J 11.0Hz, $PhCH_2$), 4.62 (1H, d, $J_{H1\text{-}H2}$ 3.5Hz, H1), 4.68, 4.91 (2 x 1H, d, ^2J 12.0Hz, $PhCH_2$), 4.80 (1H, d, ^2J 10.4Hz, $PhCH_2$), 4.82 (1H, d, ^2J 9.1Hz, $PhCH_2$), 7.13-7.42 (15H, m, ArH); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 51.4 (1C, C6), 55.3 (1C, OMe), 69.9 (1C, C5), 73.4, 75.1, 75.7 (3 x 1C, 3 x $PhCH_2$), 78.3 (1C, C4), 79.9 (1C, C2), 81.8 (1C, C3), 98.0 (1C, C1) 127.4-129.4 (15C, m, ArC), 137.9, 138.0, 138.6 (3 x 1C, ArC_{ipso}); FT-IR (thin film) υ 2099 (N_3); m/z (ES^+) 507.2 $[M + NH_4]^+$ 512.2 $[M + Na]^+$.



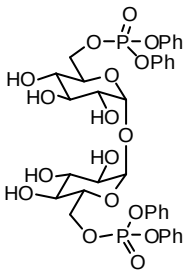
α-methyl-6-azido-D-glucose pyranoside[58] (3.15)

A mixture of **3.7** (100mg, 0.37mmol, 1eq) and sodium azide (24mg, 0.37mmol, 1eq) in DMF (5ml) were heated to 80°C. TLC (95 ethyl acetate : 5 methanol) showed complete consumption of starting material after 7 hours. The reaction mixture was allowed to cool to room temperature and the solvent removed by lyophilization. The crude product was purified by silica gel chromatography (95 ethyl acetate : 5 methanol) to give the title compound as colourless crystals (75mg, 93%).

R_f 0.33 (95 ethyl acetate : 5 methanol), $[\alpha]_D^{22}$ + 52.6 (c = 1.0, H_2O), lit.[58] $[\alpha]_D^{20}$ +54 (c = 0.9, H_2O); ^{1}H NMR (400MHz, D_2O) δ ppm 3.34 (1H, t, $J_{H3\text{-}H4}$ 9.6Hz, $J_{H4\text{-}H5}$ 9.6Hz, H4), 3.36 (3H, s, OMe), 3.49 (1H, dd, $J_{H6\text{-}H6'}$ 13.1Hz, $J_{H5\text{-}H6'}$ 7.1Hz, H6'), 3.50 (1H, dd, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.57 (1H, t, $J_{H2\text{-}H3}$ 9.1Hz, $J_{H3\text{-}H4}$ 9.1Hz, H3),

3.62 (1H, dd, $J_{H6\text{-}H6'}$ 13.5Hz, $J_{H5\text{-}H6}$ 2.5Hz, H6), 3.70 (1H, ddd, $J_{H4\text{-}H5}$ 9.9Hz, $J_{H5\text{-}H6'}$ 6.1Hz, $J_{H5\text{-}H6}$ 2.3Hz, H5), 4.74 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (101MHz, D_2O) δ ppm 51.2 (1C, C6), 55.6 (1C, OMe), 70.7 (1C, C5), 70.7 (1C, C4), 71.5 (1C, C2), 73.2 (1C, C3), 99.7 (1C, C1); m/z (ES^+) 283.1 [M + CH_3CN + Na], 461.2 [2M + Na] (ES^-) 254.1 [M + $^{35}Cl]^-$, 437.1 [2M – H], 474.3 [2M + ^{35}Cl].

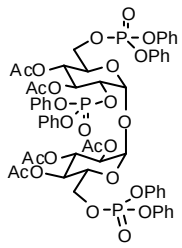
3.9.3 Optimized phosphate S_N2 route



6,6'-*O*-Di(diphenoxyphosphoryl)-D-trehalose (2.2)

To a suspension of D-trehalose (1.00g, 2.92mmol, 1eq) in dry pyridine (20ml) at 0ºC was added dropwise diphenylchlorophosphate (0.53ml, 2.55mmol, 1.75eq). TLC (1 water : 4 isopropanol : 4 ethyl acetate) after 3 hours showed the presence of a major product. The reaction was quenched with methanol (2ml), the reaction mixture concentrated *in vacuo*, and the residue co-evaporated with toluene to remove pyridine. Silica gel chromatography (1 water : 4 isopropanol : 8 ethyl acetate) was used for purification. Lyophilization yielded the title compound as a white amorphous solid (1.69g, 72%).
For characterization data see section 2.8.2 page 46.

Analysis of Multiple Phosphorylation Products

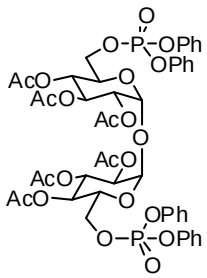


6,6',2-O-Tri(diphenoxyphosphoryl)-O-acetyl-D-trehalose

To a solution of D-trehalose (50mg, 0.15mmol, 1eq) in dry pyridine (3ml) was added diphenylchlorophosphate (60µl, 0.58mmol, 4eq) at room temperature. TLC (1 water : 4 isopropanol : 4 ethyl acetate) after 3 hours showed complete consumption of starting materials and mass spectrometry (ES^+) indicated multiple phosphorylations. To aid purification and analysis, the crude products were peracetylated by dropwise addition of acetic anhydride (0.22ml, 2.32mmol, 16eq). After 5 hours stirring at room temperature, the reaction was quenched by addition of methanol (5ml) and the solvents were removed *in vacuo*. The crude residue was partitioned between water (25ml) and ethyl acetate (25ml) and the aqueous phase extracted with ethyl acetate (2 x 25ml). The combined organics were washed with 1M HCl (3 x 20ml), saturated $NaHCO_3$ solution (20ml), brine (20ml), dried over $MgSO_4$ and concentrated *in vacuo*. Silica gel

chromatography (2 petrol : 1 ethyl acetate) allowed the pure isolation of the title compound (60mg, 32%).

R_f 0.45 (2 ethyl acetate : 1 petrol), $[\alpha]_D^{18}$ + 87.6 (c = 1.0, $CHCl_3$); ^{1}H NMR (500MHz, MeOD) δ ppm 2.00-2.13 (15H, m, 5 x OAc), 4.06 (1H, ddd, $J_{H6'a-H6'b}$ 11.3Hz, $J_{H6'b-31P}$ 7.2Hz, $J_{H5'-H6'b}$ 2.7Hz, H6'b), 4.12-4.22 (2H, m, H5, H5'), 4.26 (1H, ddd, $J_{H6'a-H6'b}$ 11.7Hz, $J_{H6'a-31P}$ 6.3Hz, $J_{H5'-H6'a}$ 1.9Hz), 4.34 (2H, m, H6a, H6b), 4.78 (1H, td, J_{H2-H3} 9.1Hz, J_{H2-31P} 9.1Hz, J_{H1-H2} 3.8Hz, H2), 4.95 (1H, dd, $J_{H2'-H3'}$ 10.1Hz, $J_{H1'-H2'}$ 3.5Hz, H2'), 5.05-5.14 (2H, m, H4, H4'), 5.16 (1H, d, $J_{H1'-H2'}$ 3.5Hz, H1'), 5.34 (1H, d, J_{H1-H2} 3.8Hz, H1), 5.46 (1H, t, $J_{H2'-H3'}$ 10.1Hz, $J_{H3'-H4'}$ 10.1Hz, H3'), 5.57 (1H, t, J_{H2-H3} 9.6Hz, J_{H3-H4} 9.6Hz, H3), 7.17-7.29 (18H, m, ArH_{orhto}, ArH_{para}), 7.37-7.45 (12H, m, ArH_{meta}) ; ^{13}C NMR (126MHz, MeOD) δ ppm 20.6, 20.6, 20.7, 20.7, 20.8 (5 x 1C, 5 x $CO\underline{C}CH_3$), 67.5 (1C, d, J_{C-31P} 5.7Hz, C6'), 68.4 (1C, d, J_{C-31P} 5.7Hz, C6), 68.7 (1C, C4), 69.3 (1C, C4'), 69.9 (1C, C5), 69.9 (1C, C5'), 70.9 (1C, C2'), 71.5 (1C, C3'), 71.6 (1C, C3), 76.1 (1C, d, J_{C-31P} 6.7Hz, C2), 93.8 (1C, d, J_{C-31P} 3.8Hz, C1), 94.0 (1C, C1'), 120.9-121.3 (12C, m, ArC_{ortho}), 126.7-127.2 (6C, m, ArC_{para}), 131.1, 131.2, 131.2, 131.3, 131.4, 131.4 (6 x 2C, 6 x ArC_{meta}), 151.3-151.9 (6C, m, ArC_{ipso}), 170.8, 171.0, 171.2, 171.6, 171.7 (5 x 1C, 5 x C=O); FT-IR (thin film) υ 1218 (P=O), 3442 br (OH); HRMS m/z $(ES)^+$ $[M + Na]^+$ peaks 1271.24 (100%), 1272.24 (60%), 1273.24 (23%), 1274.24 (6%) required 1271.25 (100%), 1272.25 (64%), 1273.25 (24%), 1274.25 (6%).



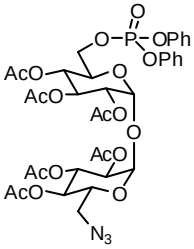
6,6'-*O*-Di(diphenoxyphosphoryl), 2,2',3,3',4,4'-*O*-acetyl-D-trehalose (3.16)

Acetic anhydride (2.18ml, 23.0mmol, 12eq) was added dropwise over a period of 5 minutes to a solution of **2.2** (1.55g, 1.92mmol, 1eq) in anhydrous pyridine (10ml). The reaction was allowed to stir at room temperature for 5 hours, after which TLC (2 ethyl acetate : 1 petrol) showed the full consumption of starting material. The reaction was quenched with methanol (5ml) and the solvent removed *in vacuo*. The crude residue was partitioned between ethyl acetate (50ml) and water (20ml). The aqueous layer was extracted with ethyl acetate (2 x 25ml) and the combined organics washed with 1M HCl (3 x 25ml), saturated $NaHCO_3$ solution (25ml), brine (25ml), dried over $MgSO_4$ and the solvent removed *in vacuo*. Purification was achieved with silica gel chromatography (1 petrol : 1 ethyl acetate) to afford the desired compound as a colourless oil (1.60g, 78%).

R_f 0.33 (2 ethyl acetate : 1 petrol), $[\alpha]_D^{19}$ +80.4 (c = 1.0, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.02, 2.03, 2.04 (3 x 6H, 3 x s, 6 x CH_3), 4.12 (2H, ddd, J_{H4-H5} 10.1Hz, J_{H5-H6}6.6Hz, $J_{H5-H6'}$ 2.0Hz, H5), 4.22 (2H, ddd, $J_{H6-H6'}$ 11.1Hz, $J_{H6'-31P}$ 6.6Hz, $J_{H5-H6'}$ 2.2Hz, H6'), 4.29 (2H, ddd, $J_{H6-H6'}$ 11.4Hz, J_{H6-31P} 6.6Hz, J_{H5-H6} 5.4Hz, H6), 4.96 (2H, dd, J_{H2-H3} 10.2Hz, J_{H1-H2} 3.9Hz, H2), 5.00 (2H, d, J_{H1-H2} 3.8Hz, H1), 5.03 (2H, dd, J_{H4-H5} 10.2, J_{H3-H4} 9.3Hz, H4), 5.43 (2H, dd, J_{H2-H3} 10.1Hz, J_{H3-H4} 9.5Hz, H3), 7.16-7.22 (12H, m, ArH_{ortho}, ArH_{para}), 7.32-7.36 (8H, m, ArH_{meta}); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.5, 20.6, 20.7 (3 x 2C, CH_3), 66.6 (2C, d, J_{C6-31P} 5.7Hz, C6), 68.3 (2C, C4), 68.7 (2C, d, J_{C5-31P} 8.6Hz, C5), 69.3 (2C, C2), 70.1 (2C, C3), 93.0 (2C, C1), 120.0 (8C, d, J_{C-31P} 4.8Hz, ArC_{ortho}), 125.5 (4C, ArC_{para}), 129.8 (8C, ArC_{meta}), 150.3 (2C, d, J_{C-31P} 6.7Hz, ArC_{ipso}), 150.4 (2C, d, J_{C-31P} 7.6Hz, ArC_{ipso}), 169.3, 169.4, 170.0 (3 x 2C, C=O); FT-IR (thin film) υ 1248 (P=O), 1755 (C=O); HRMS m/z (ES^+) 1081.23 (100%), 1082.23 (52%), 1083.24 (17%), 1084.24 (4%) required 1081.23 (100%), 1082.23 (53%), 1083.23 (19%), 1084.24 (5%).

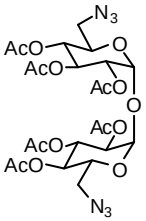
6-O-diphenoxyphosphoryl-6'azido-2,2',3,3',4,4'-O-acetyl-D-trehalose (3.17)
6,6'diazido-2,2',3,3',4,4'-O-acetyl-D-trehalose[59] (3.18)

A mixture of **3.16** (1.60g, 1.5mmol, 1eq) and sodium azide (9.8mg, 1.5mmol, 1eq) in DMF (10ml) was heated to 80ºC with aggressive stirring. The reaction was regularly monitored by TLC (2 ethyl acetate : 1 petrol) and mass spectrometry (ES^+). After 9 hours of heating, significant amounts of side-product **3.18** were detected and the reaction was allowed to cool to room temperature. The reaction mixture was concentrated under reduced pressure and the resulting residue was partitioned between ethyl acetate (50ml) and water (20ml). The aqueous layer was extracted with ethyl acetate (2 x 25ml) and combined organics washed with brine (3 x 25ml), dried over $MgSO_4$ and concentrated *in vacuo*. The three components were separated by silica gel chromatography (2 petrol : 1 ethyl acetate) to give the desired product **3.17** (0.48g, 37%) and side-product **3.18** (0.087g, 9%) as crystalline white solids. Starting material, **3.16**, was also recovered (0.72g, 45%). The starting material was re-subjected to the reaction conditions to give a second crop of **3.17** (0.66g, 51%).



*Compound **3.17***

R_f 0.39 (2 ethyl acetate : 1 petrol), $[\alpha]_D^{19}$ +97.4 (c = 1.0, $CHCl_3$), m.p. 42.2-44.6ºC, ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.02, 2.04, 2.05, 2.05, 2.06, 2.12 (6 x 3H, 6 x s, 6 x CH_3), 3.14 (1H, dd, $J_{H6'a-H6'b}$ 13.4Hz, $J_{H5'-H6'b}$ 2.4Hz, H6'b), 3.35 (1H, dd, $J_{H6'a-H6'b}$ 13.2Hz, $J_{H5'-H6'a}$ 7.3Hz, H6'a), 4.06 (1H, ddd, $J_{H4'-H5'}$ 10.2Hz, $J_{H5'-H6'a}$ 7.5Hz, $J_{H5'-H6'b}$ 2.2Hz, H5'), 4.16 (1H, add, J_{H4-H5} 10.4Hz, J_{H5-H6a} 5.4Hz, H5), 4.22 (1H, ddd, $J_{H6a-H6b}$ 11.3Hz, $J_{H6b-31P}$ 6.5Hz, J_{H5-H6b} 2.4Hz, H6b), 4.31 (1H, ddd, $J_{H6a-H6b}$ 11.4Hz, $J_{H6a-31P}$ 6.2Hz, J_{H5-H6a} 5.7Hz, H6a), 4.98 (1H, t, $J_{H3'-H4'}$ 9.8Hz, $J_{H4'-H5'}$ 9.8Hz, H4'), 4.99 (1H, dd, $J_{H2'-H3'}$ 10.0Hz, $J_{H1'-H2'}$ 3.8Hz, H2'), 5.03 (1H, dd, J_{H2-H3} 8.2Hz, J_{H1-H2} 4.0Hz, H2), 5.03 (1H, t, J_{H3-H4} 10.1Hz, J_{H4-H5} 10.1Hz, H4), 5.03 (1H, d, $J_{H1'-H2'}$ 2.8Hz, H1'), 5.18 (1H, d, J_{H1-H2} 3.8Hz, H1), 5.43 (1H, dd, J_{H2-H3} 9.6Hz, J_{H3-H4} 9.6Hz, H3), 5.47 (1H, dd, $J_{H2'-H3'}$ 10.2Hz, $J_{H3'-H4'}$ 9.3Hz, H3'), 7.13-7.23 (6H, m, ArH_{ortho}, ArH_{para}), 7.33-7.37 (4H, m, ArH_{meta}); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 50.9 (1C, C6), 66.7 (1C, d, $J_{C6'-31P}$ 5.7Hz, C6'), 68.4 (1C, C4), 68.7 (1C, d, J_{C4-31P} 8.6Hz, C5), 69.3 (1C, C2), 69.5 (1C, C4'), 69.6 (1C, C2'), 69.9 (3C, C3, C3', C5'), 92.8 (1C, C1'), 92.9 (1C, C1), 120.0 (4C, d, J_{C-31P} 5.7Hz, ArC_{ortho}), 125.5 (2C, ArC_{para}), 129.8 (4C, d, J_{C-31P} 2.9Hz, ArC_{meta}), 150.3 (1C, d, J_{C-31P} 6.7Hz , ArC_{ipso}), 150.4 (1C, d, J_{C-31P} 7.6Hz, ArC_{ipso}), 169.3, 169.4, 169.6, 169.6, 170.0, 170.03 (6 x 1C, C=O), FT-IR (KBr disc) υ 1248 (P=O), 1755 (C=O), 2105 (N_3); HRMS m/z (ES^+) 874.2043 $[M + Na]^+$ (required 874.2042).



*Compound **3.18***

R_f 0.54 (2 ethyl acetate : 1 petrol), $[\alpha]_D^{19}$ +121.1 (c = 1.0, $CHCl_3$), lit.[59] $[\alpha]_D$ +134 (c = 0.5, $CHCl_3$), m.p. 103-106ºC, lit.[59] 114-116°C; ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.03, 2.06, 2.12 (3 x 6H, 3 x s, 6 x CH_3), 3.17 (2H, dd, $J_{H6-H6'}$ 13.4Hz, $J_{H5-H6'}$ 2.4Hz, H6'), 3.37 (2H, dd, $J_{H6-H6'}$ 13.4Hz, J_{H5-H6} 7.4Hz, H6), 4.08 (2H, dd, J_{H5-H6} 7.3Hz, $J_{H5-H6'}$ 2.5Hz, H5), 4.99 (2H, t, J_{H3-H4} 9.5Hz, J_{H4-H5} 9.5Hz, H4), 5.08 (2H, dd, J_{H2-H3} 10.2Hz, J_{H1-H2} 3.9Hz, H2), 5.33 (2H, d, J_{H1-H2} 4.1Hz, H1), 5.47 (2H, t, J_{H2-H3} 10.1Hz, J_{H3-H4} 10.1Hz, H3); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.6, 20.6, 20.6 (3 x 2C, OAc), 50.9 (1C, C6), 69.6 (1C, C5) 69.7 (1C, C2), 69.8 (1C, C4), 69.9 (1C, C3), 93.0 (1C, C1), 169.6, 169.6, 169.9 (3 x 2C, C=O); FT-IR (KBr disc) υ 1748 (C=O), 2105 (N_3); m/z (ES^+) 667.1823 $[M + Na]^+$ (required 667.1818).

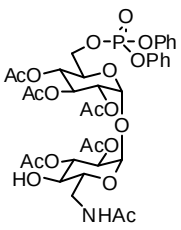
3.9.4 Diphenylphosphate deprotections by hydrogenation

Neutral hydrogenation conditions

A suspension of **3.17** (30mg, 0.035mmol, 1eq) and PtO_2 (8mg, 0.035, 1eq) in 75% aqueous ethanol (3ml) was repeatedly degassed under vacuum and flushed with hydrogen. The reaction mixture was aggressively stirred under an atmospheric pressure of hydrogen. TLC (1 water : 2 isopropanol : 2 ethyl acetate) after 90 minutes showed complete consumption of starting material and the formation of several products. The reaction mixture was filtered through Celite® and the solvent removed *in vacuo*. The compounds were separated by silica gel chromatography using gradient elution with stepwise increases in polarity of the mobile phase from pure ethyl acetate to 1 water : 1 isopropanol : 1 ethyl acetate to give **3.19** (4.0mg, 14%), **3.20** (8mg, 48%), **3.21** (1mg, 4%), and smaller quantities of other products.

Acidic hydrogenation conditions

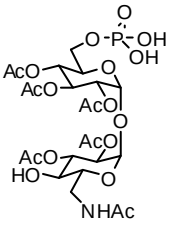
A suspension of **3.17** (30mg, 0.035mmol, 1eq) and PtO_2 (8mg, 0.035, 1eq) in 75% aqueous ethanol (3ml), adjusted to pH4 with 1M HCl (pH paper), was repeatedly degassed under vacuum and flushed with hydrogen. The reaction mixture was aggressively stirred under an atmospheric pressure of hydrogen. TLC (1 petrol : 2 ethyl acetate) after 2 hours showed complete consumption of starting material and the formation of two highly polar products. The reaction mixture was neutralized with 1M NaOH (pH paper), filtered through Celite® and the solvent removed *in vacuo*. The compounds were separated by silica gel chromatography (1 water : 1 isopropanol : 1 ethyl acetate). Fractions were concentrated and centrifuged (16000*g*, 5 minutes) and the supernatant filtered through a 45µm membrane to remove residual silica. Lyophilization gave **3.20** (4.2mg, 18%) and **3.21** (8.2mg, 34%) as amorphous white solids.



6-*O*-diphenoxyphosphoryl-6'acetimido-2,2',3,3',4,-*O*-acetyl-D-trehalose (3.19)

R_f 0.24 (ethyl acetate), $[\alpha]_D^{18}$ +66.5 (c = 0.2, $CHCl_3$); ^1H NMR (500MHz, $CDCl_3$) δ ppm 2.03, 2.03, 2.05, 2.08, 2.08, 2.09 (6 x 3H, 6 x s, 6 x OAc), 3.06 (1H, ddd, $J_{H6'a\text{-}H6'b}$ 14.7Hz, $J_{H6'b\text{-}NH}$ 5.0Hz, $J_{H5'\text{-}H6'b}$ 2.7Hz, H6'b),

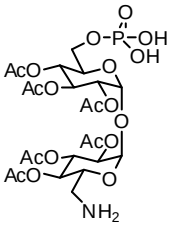
3.34 (1H, td, $J_{H3'-H4'}$ 9.6Hz, $J_{H4'-H5'}$ 9.6Hz, $J_{H4'-OH}$ 2.5Hz, H4'), 3.72 (1H, dt, $J_{H4'-H5'}$ 9.8Hz, $J_{H5'-H6'a}$ 2.8Hz, $J_{H5'-H6'b}$ 2.8Hz, H5'), 3.95 (1H, ddd, $J_{H6'a-H6'b}$ 14.9Hz, $J_{H6'a-NH}$ 8.0Hz, $J_{H5'-H6'a}$ 3.0Hz, H6'a), 4.13 (1H, m, H5), 4.24 (1H, ddd, $J_{H6a-H6b}$ 10.9Hz, $J_{H6b-31P}$ 6.6Hz, J_{H5-H6b} 2.2Hz, H6b), 4.30 (1H, ddd, $J_{H6a-H6b}$ 11.5Hz, $J_{H6a-31P}$ 6.8Hz, J_{H5-H6a} 5.2Hz, H6a), 4.84 (1H, dd, J_{H2-H3} 10.2Hz, J_{H1-H2} 3.9Hz, H2), 4.94 (1H, dd, $J_{H2'-H3'}$ 10.2Hz, $J_{H1'-H2'}$ 3.9Hz, H2'), 5.04 (1H, t, J_{H3-H4} 10.4Hz, J_{H4-H5} 10.4Hz, H4), 5.10 (1H, d, $J_{H1'-H2'}$ 4.1Hz, H1'), 5.26 (1H, d, J_{H1-H2} 3.8Hz, H1), 5.39 (1H, t, $J_{H2'-H3'}$ 9.9Hz, $J_{H3'-H4'}$ 9.9Hz, H3'), 5.45 (1H, t, J_{H2-H3} 9.8Hz, J_{H3-H4} 9.8Hz, H3), 5.85 (1H, dd, J_{H6a-NH} 7.7Hz, J_{H6b-NH} 5.2Hz, NH), 7.14-7.22 (6H, m, ArH_{ortho}, ArH_{para}), 7.33-7.37 (4H, m, ArH_{meta}); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.6, 20.6, 20.6, 20.6, 20.9 (5 x 1C, 5 x OAc), 23.0 (1C, NHAc), 39.5 (1C, C6'), 66.6 (1C, d, J_{C-31P} 5.7Hz, C6), 68.3 (1C, C4), 68.5 (1C, d, J_{C-31P} 8.6Hz, C5), 68.9 (1C, C4'), 69.8 (1C, C3), 69.9 (1C, C2), 70.2 (1C, C2'), 71.0 (1C, C5'), 71.6 (1C, C3'), 92.1 (1C, C1), 93.2 (1C, C1'), 120.0, 120.0 (2 x 2C, d, J_{C-31P} 4.8Hz, ArC_{ortho}), 125.56 (2C, ArC_{para}), 129.8(4C, ArC_{meta}), 150.3, 150.4 (2 x 1C, d, J_{C-31P} 6.7Hz, ArC_{ipso}), 169.5, 169.7, 169.8, 170.0, 170.4 (5 x 1C, 5 x OAc C=O) 172.5 (1C, $NHCOCH_3$); FT-IR (thin film) 1222 (P=O), 1651 (RNHC=O), 1750 (C=O); HRMS (ES^+) m/z 848.2139 $[M + Na]^+$ (required 848.2137).



6-O-dihydroxyphosphoryl-6'acetimido-2,2',3,3',4,-O-acetyl-D-trehalose (3.20)

R_f 0.52 (1 water : 1 isopropanol : 1 ethyl acetate), $[\alpha]_D^{18}$ +24.9 (c = 0.21, H_2O); ^1H NMR (500MHz, D_2O) δ ppm 2.01, 2.01, 2.02, 2.03, 2.04, 2.06 (6 x 3H, 6 x s, 5 x OAc, NHAc), 3.13 (1H, dd, $J_{H6'a-H6'b}$ 14.2Hz, $J_{H5'-H6'b}$ 7.4Hz, H6'b), 3.54 (1H, t, $J_{H3'-H4'}$ 9.5Hz, $J_{H4'-H5'}$ 9.5Hz, H4'), 3.62 (1H, ddd, $J_{H4'-H5'}$ 10.2Hz, $J_{H5'-H6'a}$ 8.7Hz, $J_{H5'-H6'b}$ 2.4Hz, H5'), 3.71 (1H, dd, $J_{H6'a-H6'b}$ 14.2Hz, $J_{H5'-H6'a}$ 2.2Hz, H6'a), 3.82 (2H, m, H6a, H6b), 4.04 (1H, ddd, J_{H4-H5} 10.2Hz, J_{H5-H6a} 4.8Hz, J_{H5-H6b} 2.9Hz, H5), 4.97 (1H, dd, $J_{H2'-H3'}$ 10.1Hz, $J_{H1'-H2'}$ 3.8Hz, H2'), 5.06 (1H, dd, J_{H2-H3} 10.1Hz, J_{H1-H2} 3.8Hz, H2), 5.09 (1H, t, J_{H3-H4} 9.9Hz, J_{H4-H5} 9.9Hz, H4), 5.26 (1H, t, $J_{H2'-H3'}$ 9.6Hz, $J_{H3'-H4'}$ 9.6Hz, H3'), 5.32 (2H, d, J_{H1-H2} 3.4Hz, H1, H1'), 5.50 (1H, t, J_{H2-H3} 9.7Hz, J_{H3-H4} 9.7Hz, H3); ^{13}C NMR (126MHz, D_2O) δ ppm 19.9, 20.0, 20.1, 20.2, 20.3 (5 x 1C, 5 x OAc), 22.1 (1C, NHAc), 39.5 (1C, C6'), 62.1 (1C, d, J_{C-31P} 2.9Hz, C6), 68.5 (1C, C4), 68.7 (1C, d, J_{C-31P} 2.9Hz, C5), 69.3 (1C, C2), 70.4 (1C, C2'), 70.7 (1C, C3), 71.5 (1C, C3'), 72.3 (1C, C5'), 91.4 (1C, C1'), 91.5 (1C, C1), 173.4, 173.5, 173.6, 173.9, 174.0, 174.4 (6 x 1C, 6 x C=O); FT-IR (KBr disc) υ 1631

(RNHC=O), 1744 (C=O), 3448 br (OH); HRMS (ES^+) m/z 696.1502 $[M + Na]^+$ (required 696.1511).



6-O-hydroxyoxyphosphoryl-6'-amino-2,2',3,3',4,4'-O-acetyl-D-trehalose (3.21)

R_f 0.43 (1 water : 1 isopropanol : 1 ethyl acetate), $[\alpha]_D^{18}$ +42.9 (c = 1.0, H_2O); ^1H NMR (500MHz, D_2O) δ ppm 2.01, 2.01, 2.03, 2.04, 2.06, 2.08 (6 x 3H, 6 x s, 6 x OAc), 3.84-3.86 (2H, m, H6a, H6b), 4.10 (1H, m, H5), 4.12 (1H, dd, $J_{H6'a\text{-}H6'b}$ 12.2Hz, $J_{H5'\text{-}H6'b}$ 2.2Hz, H6'b), 4.16 (1H, ddd, $J_{H4'\text{-}H5'}$ 10.2Hz, $J_{H5'\text{-}H6'a}$ 5.2Hz, $J_{H5'\text{-}H6'b}$ 2.2Hz, H5'), 4.30 (1H, dd, $J_{H6'a\text{-}H6'b}$ 12.3Hz, $J_{H5'\text{-}H6'a}$ 5.0Hz, H6'a), 5.09 (1H, t, $J_{H3\text{-}H4}$ 9.8Hz, $J_{H4\text{-}H5}$ 9.8Hz, H4), 5.14 (1H, t, $J_{H3'\text{-}H4'}$ 9.8Hz, $J_{H4'\text{-}H5'}$ 9.8Hz, H4'), 5.14 (2H, dd, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H1\text{-}H2}$ 3.5Hz, H2, H2'), 5.36 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 5.41 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.51 (1H, t, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H3\text{-}H4}$ 9.8Hz, H3), 5.52 (1H, t, $J_{H2'\text{-}H3'}$ 9.8Hz, $J_{H3'\text{-}H4'}$ 9.8Hz, H3'); ^{13}C NMR (126MHz, D_2O) δ ppm 20.0, 20.1, 20.1, 20.1, 20.1, 20.2 (6 x 1C, 6 x OAc), 62.1 (1C, C6'), 62.3 (1C, d, $J_{C\text{-}31P}$ 2.9Hz, C6), 68.0 (1C, C4'), 68.3 (1C, C4), 68.4 (1C, C5'), 69.4 (31C, d, $J_{C\text{-}31P}$ 9.5Hz, C5), 69.8 (1C, C2'), 70.0 (1C, C2), 70.7 (1C, C3'), 71.0 (1C, C3), 92.1 (1C, C1'), 92.1 (1C, C1), 172.4, 172.6, 172.8, 173.3, 173.4, 173.6 (6 x 1C, 6 x OAc C=O); FT-IR (KBr disc) 1751 (C=O), 3442 br (OH); HRMS (ES^+) m/z 696.1507 $[M + Na]^+$ (required 696.1511).

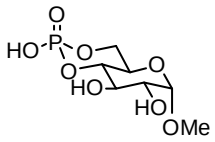
3.9.5 Diphenylphosphate deprotections by basic hydrolysis

*Deprotection of **3.9** with NH_4OH*

To a solution of **3.9** (30mg, 0.07mmol, 1eq) in water (4ml) was added NH_4OH (28% solution in water, 18µl, 0.28mmol, 4eq). The reaction was allowed to stir at room temperature and monitored by TLC (95 ethyl acetate : 5 methanol) for consumption of starting material. After 10 minutes, the reaction was neutralized (pH 7 by pH paper) with 1M HCl and the solvent and ammonium hydroxide were removed *in vacuo*. The crude solid was taken up in water (20ml) and washed with DCM (2 x 10ml). The aqueous layer was concentrated *in vacuo*. Silica gel chromatography (1 water : 2 isopropanol : 2 ethyl acetate) allowed separation of the products. Lyophilization yielded **3.22** (12.4mg, 69%) and **3.23** (3mg, 12%).

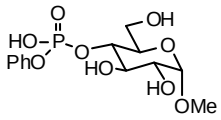
*Deprotection of **3.9** with CsF*

A solution of **3.9** (41mg, 0.096mmol, 1eq) and CsF (0.14g, 0.96mmol, 10eq) in *tert*-butanol (3ml) and water (1ml) was heated to 80°C. TLC (1 petrol : 1 ethyl acetate) after 12 hours indicated complete consumption of the starting material. The solvent was removed under reduced pressure and the products separated by silica gel chromatography (1 water : 4 isopropanol : 4 ethyl acetate). Lyophilisation of the pooled fractions afforded **3.22** (12.7mg, 38%) and **3.24** (13.4mg, 56%) as white amorphous solids.



4,6-(monohydrogen)phosphoryl-D-glucose (3.22)

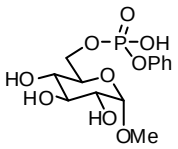
R_f 0.5 (5 ethanol : 3 NH_4OH : 1 water), $[\alpha]_D^{19}$ +52.6 (c = 1.0, H_2O); ^1H NMR (500MHz, D_2O) δ ppm 3.37 (1H, t, $J_{H3'\text{-}H4'}$ 9.5Hz, $J_{H4'\text{-}H5'}$ 9.5Hz, H4'), 3.56 (1H, dd, $J_{H2'\text{-}H3'}$ 10.1Hz, $J_{H1'\text{-}H2'}$ 3.8Hz, H2'), 3.67 (1H, dd, $J_{H2\text{-}H3}$ 9.2Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.68 (1H, t, $J_{H6'a\text{-}H6'b}$ 6.0Hz, $J_{H5'\text{-}H6'b}$ 6.0Hz, H6'b), 3.73-3.80 (3H, m, H3', H5', H6'a), 3.89 (1H, ddd, $J_{H4\text{-}H5}$ 9.5Hz, $J_{H5\text{-}H6ax}$ 9.0Hz, $J_{H5\text{-}H6eq}$ 0.9Hz, H5), 3.92 (1H, t, $J_{H2\text{-}H3}$ 9.1Hz, $J_{H3\text{-}H4}$ 9.1Hz, H3), 4.01 (1H, td, $J_{H5\text{-}H6ax}$ 8.4Hz, $J_{H6ax\text{-}H6eq}$ 8.4Hz, $J_{H6ax\text{-}31P}$ 1.7Hz, H6ax), 4.02 (1H, t, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, H4), 4.13 (1H, ddd, $J_{H6eq\text{-}31P}$ 22.3Hz, $J_{H6ax\text{-}H6eq}$ 8.4Hz, $J_{H5\text{-}H6eq}$ 3.8Hz, H6eq), 5.10 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.13 (1H, d, $J_{H1\text{-}H2}$ 4.1Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 60.4 (1C, C6'), 64.0 (1C, d, $J_{C4\text{-}31P}$ 4.8Hz), 66.5 (1C, d, $J_{C6\text{-}31P}$ 5.7Hz, C6), 69.6 (1C, C4'), 70.4 (1C, d, $J_{C3\text{-}31P}$ 8.6Hz, C3), 70.8 (1C, d, $J_{C2\text{-}31P}$ 1.9Hz, C2), 71.0 (1C, C2'), 72.2 (1C, C3'), 72.5 (1C, C5'), 78.2 (1C, d, $J_{C5\text{-}31P}$ 4.8Hz, C5), 93.7 (1C, C1'), 94.0 (1C, C1); $^{31}P\{^1H\}$ NMR (162MHz, D_2O) δ ppm -2.46 (1P, P(O)(OR)(OR')); FT-IR (KBr disc) υ 1139 (P=O), 3407 br (OH); HRMS m/z (ES^-) 403.0649 $[M - H]^-$ required 403.0647.



4-O-phenoxyphosphoryl-D-glucose (3.23)

R_f 0.66 (5 ethanol : 3 NH_4OH : 1 water), $[\alpha]_D^{18}$ +33 (c = 0.15, H_2O); ^1H NMR (500MHz, D_2O) δ ppm 3.36 (1H, t, $J_{H3'\text{-}H4'}$ 9.6Hz, $J_{H4'\text{-}H5'}$ 9.6Hz, H4'), 3.55 (1H, dd, $J_{H2'\text{-}H3'}$ 10.1Hz, $J_{H1'\text{-}H2'}$ 3.8Hz, H2'), 3.63 (1H, dd, $J_{H2\text{-}H3}$ 9.3Hz, $J_{H1\text{-}H2}$ 4.4Hz, H2), 3.67 (2H, add, $J_{H6'a\text{-}H6'b}$ 11.4Hz, $J_{H5\text{-}H6'}$ 5.0Hz, H6'a, H6'b), 3.73-3.80 (5H, m, H3, H3', H6a, H6b), 3.87 (1H, ddd, $J_{H4'\text{-}H5'}$ 9.3Hz, $J_{H5'\text{-}H6'a}$ 5.2Hz, $J_{H5\text{-}H6'b}$ 2.2Hz, H5), 3.98 (1H, m, H5), 3.99 (1H, q, $J_{H3\text{-}H4}$ 9.3Hz, $J_{H4\text{-}H5}$ 9.3Hz, $J_{H4\text{-}31P}$ 9.3Hz, H4), 5.11 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.12 (1H, d, $J_{H1\text{-}H2}$ 4.1Hz, H1),

7.01-7.20 (3H, m, ArH_{ortho}, ArH_{para}), 7.32 (2H, t, $J_{ortho\text{-}meta}$ 8.0Hz, $J_{meta\text{-}para}$ 8.0Hz, ArH_{meta}); ^{13}C NMR (126MHz, D_2O) δ ppm 60.3 (1C, C6), 60.5 (1C, C6'), 69.6 (1C, C4'), 70.8 (1C, C2), 71.1 (1C, C2'), 71.2 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, C5), 71.7 (1C, d, $J_{C\text{-}31P}$ 1.9Hz, C4), 72.1 (1C, C3), 72.1 (1C, C3'), 72.5 (1C, C5'), 93.1 (1C, C1), 93.4 (1C, C1'), 120.2 (2C, d, $J_{C\text{-}31P}$ 4.8Hz, ArC_{ortho}), 124.5 (1C, ArC_{para}), 129.8 (2C, ArC_{meta}), 151.6 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, ArC_{ipso}); $^{31}P\{^1H\}$ NMR (162MHz, D_2O) δ ppm -8.70 (1P, P(O)(OR)(OPh)(OH)); FT-IR (KBr disc) υ 1106 (P=O), 3426 br (OH); HRMS (ES^-) m/z 497.1068 $[M - H]^-$ (required 497.1066).



α-methyl-6-O-monophenoxyphosphoryl-D-glucose pyranoside[60]

(3.24)

R_f 0.44 (1 water : 2 isopropanol : 2 ethyl acetate), $[\alpha]_D^{18}$ +55.9 (c = 0.63, H_2O), lit.[60] $[\alpha]_D^{20}$ +61.9 (c = 0.5, H_2O); ^1H NMR (500MHz, D_2O) δ ppm 3.31 (3H, s, OMe), 3.38 (1H, t, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, H4), 3.44 (1H, dd, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.57 (1H, t, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H3\text{-}H4}$ 9.5Hz, H3), 3.67 (1H, add, $J_{H4\text{-}H5}$ 9.0Hz, $J_{H5\text{-}H6'}$ 3.6Hz, H5), 4.08 (1H, ddd, $J_{H6\text{-}H6'}$ 11.6Hz, $J_{H6'\text{-}31P}$ 6.7Hz, $J_{H5\text{-}H6'}$ 5.0 Hz, H6'), 4.16 (1H, ddd, $J_{H6\text{-}H6'}$ 11.5Hz, $J_{H6\text{-}31P}$ 5.9Hz, $J_{H5\text{-}H6}$ 1.9Hz, H6), 4.68 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 7.14 (3H, d, $J_{ortho\text{-}meta}$ 7.9Hz, $J_{meta\text{-}para}$ 7.9Hz, ArH_{ortho}, ArH_{para}), 7.32 (2H, t, $J_{ortho\text{-}meta}$ 7.9Hz, $J_{meta\text{-}para}$ 7.9Hz, ArH_{meta}); ^{13}C NMR (126MHz, D_2O) δ ppm 55.1 (1C, OMe), 64.8 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, C6), 69.1 (1C, C4), 70.5 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C5), 71.1 (1C, C2), 72.9 (1C, C3), 99.3 (1C, C1), 120.2 (2C, d, $J_{C\text{-}31P}$ 4.8Hz, ArC_{ortho}), 124.3 (1C, ArC_{para}), 129.7 (2C, ArC_{meta}), 151.7 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, ArC_{ipso}); m/z (ES^-) 349.1 $[M - H]^-$.

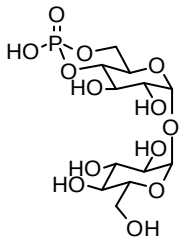
*Deprotection of **2.1** with NH_4OH*

To a solution of **2.1** (30mg, 0.052mmol, 1eq) in water (4ml) was added NH_4OH (28% solution in water, 14μl, 0.21mmol, 4eq). The reaction mixture was stirred at room temperature and monitored by TLC (1 water : 1 isopropanol : 1 ethyl acetate) for consumption of the starting material. After 10 minutes, the reaction was neutralized (pH 7 by pH paper) with 1M HCl, the solvent and ammonium hydroxide were removed *in vacuo*. The crude solid was taken up in water (25ml) and washed with DCM (3 x 10ml). The aqueous layer was retained and the water removed under reduced pressure. The compounds was separated using silica gel chromatography (1

water : 2 isopropanol : 2 ethyl acetate). Lyophilization gave **3.25** (18.4mg, 88%) and **3.23** (3mg, 9%) as amorphous white solids.

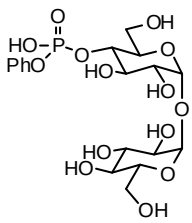
*Deprotection of **2.1** with CsF*

A solution of **2.1** (30mg, 0.052mmol, 1eq) and CsF (80mg, 0.52mmol, 10eq) in water (1ml) and *tert*-butanol (3ml) were heated to 80°C. TLC (1 water : 2 isopropanol : 2 ethyl acetate) showed complete consumption of starting material after 12 hours. The reaction mixture was allowed to cool and the solvent removed *in vacuo*. Separation of the components was achieved using silica gel chromatography (1 water : 2 isopropanol : 2 ethyl acetate). Lyophilization of the product containing fractions yielded **3.25** (10.2mg, 49%) and **3.27** (8.1mg, 31%).



4,6-(monohydrogen)phosphoryl-D –trehalose (3.25)

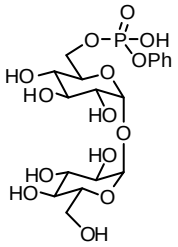
R_f 0.5 (5 ethanol : 3 NH_4OH : 1 water), $[\alpha]_D^{19}$ +52.6 (c = 1.0, H_2O); ^1H NMR (500MHz, D_2O) δ ppm 3.37 (1H, t, $J_{H3'\text{-}H4'}$ 9.5Hz, $J_{H4'\text{-}H5'}$ 9.5Hz, H4'), 3.56 (1H, dd, $J_{H2'\text{-}H3'}$ 10.1Hz, $J_{H1'\text{-}H2'}$ 3.8Hz, H2'), 3.67 (1H, dd, $J_{H2\text{-}H3}$ 9.2Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.68 (1H, t, $J_{H6'a\text{-}H6'b}$ 6.0Hz, $J_{H5'\text{-}H6'b}$ 6.0Hz, H6'b), 3.73-3.80 (3H, m, H3', H5', H6'a), 3.89 (1H, ddd, $J_{H4\text{-}H5}$ 9.5Hz, $J_{H5\text{-}H6ax}$ 9.0Hz, $J_{H5\text{-}H6eq}$ 0.9Hz, H5), 3.92 (1H, t, $J_{H2\text{-}H3}$ 9.1Hz, $J_{H3\text{-}H4}$ 9.1Hz, H3), 4.01 (1H, td, $J_{H5\text{-}H6ax}$ 8.4Hz, $J_{H6ax\text{-}H6eq}$ 8.4Hz, $J_{H6ax\text{-}31P}$ 1.7Hz, H6ax), 4.02 (1H, t, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, H4), 4.13 (1H, ddd, $J_{H6eq\text{-}31P}$ 22.3Hz, $J_{H6ax\text{-}H6eq}$ 8.4Hz, $J_{H5\text{-}H6eq}$ 3.8Hz, H6eq), 5.10 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.13 (1H, d, $J_{H1\text{-}H2}$ 4.1Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 60.4 (1C, C6'), 64.0 (1C, d, $J_{C4\text{-}31P}$ 4.8Hz), 66.5 (1C, d, $J_{C6\text{-}31P}$ 5.7Hz, C6), 69.6 (1C, C4'), 70.4 (1C, d, $J_{C3\text{-}31P}$ 8.6Hz, C3), 70.8 (1C, d, $J_{C2\text{-}31P}$ 1.9Hz, C2), 71.0 (1C, C2'), 72.2 (1C, C3'), 72.5 (1C, C5'), 78.2 (1C, d, $J_{C5\text{-}31P}$ 4.8Hz, C5), 93.7 (1C, C1'), 94.0 (1C, C1); $^{31}P\{^1H\}$ NMR (162MHz, D_2O) δ ppm -2.46 (1P, P(O)(OR)(OR')); FT-IR (KBr disc) υ 1139 (P=O), 3407 br (OH); HRMS m/z (ES^-) 403.0649 $[M - H]^-$ required 403.0647.



4-O-phenoxyphosphoryl-D-trehalose (3.23)

R_f 0.66 (5 ethanol : 3 NH_4OH : 1 water), $[\alpha]_D^{18}$ +33 (c = 0.15, H_2O); ^1H NMR (500MHz, D_2O) δ ppm 3.36 (1H, t, $J_{H3'\text{-}H4'}$ 9.6Hz, $J_{H4'\text{-}H5'}$ 9.6Hz, H4'), 3.55 (1H, dd, $J_{H2'\text{-}H3'}$ 10.1Hz, $J_{H1'\text{-}H2'}$ 3.8Hz, H2'), 3.63 (1H, dd, $J_{H2\text{-}H3}$ 9.3Hz, $J_{H1\text{-}H2}$ 4.4Hz, H2), 3.67 (2H, add, $J_{H6'a\text{-}H6'b}$ 11.4Hz, $J_{H5\text{-}H6'}$

5.0Hz, H6'a, H6'b), 3.73-3.80 (5H, m, H3, H3', H6a, H6b), 3.87 (1H, ddd, $J_{H4'-H5'}$ 9.3Hz, $J_{H5'-H6'a}$ 5.2Hz, $J_{H5-H6'b}$ 2.2Hz, H5), 3.98 (1H, m, H5), 3.99 (1H, q, J_{H3-H4} 9.3Hz, J_{H4-H5} 9.3Hz, J_{H4-31P} 9.3Hz, H4), 5.11 (1H, d, $J_{H1'-H2'}$ 3.8Hz, H1'), 5.12 (1H, d, J_{H1-H2} 4.1Hz, H1), 7.01-7.20 (3H, m, ArH_{ortho}, ArH_{para}), 7.32 (2H, t, $J_{ortho-meta}$ 8.0Hz, $J_{meta-para}$ 8.0Hz, ArH_{meta}); ^{13}C NMR (126MHz, D_2O) δ ppm 60.3 (1C, C6), 60.5 (1C, C6'), 69.6 (1C, C4'), 70.8 (1C, C2), 71.1 (1C, C2'), 71.2 (1C, d, J_{C-31P} 5.7Hz, C5), 71.7 (1C, d, J_{C-31P} 1.9Hz, C4), 72.1 (1C, C3), 72.1 (1C, C3'), 72.5 (1C, C5'), 93.1 (1C, C1), 93.4 (1C, C1'), 120.2 (2C, d, J_{C-31P} 4.8Hz, ArC_{ortho}), 124.5 (1C, ArC_{para}), 129.8 (2C, ArC_{meta}), 151.6 (1C, d, J_{C-31P} 7.6Hz, ArC_{ipso}); $^{31}P\{^1H\}$ NMR (162MHz, D_2O) δ ppm -8.70 (1P, P(O)(OR)(OPh)(OH)); FT-IR (KBr disc) υ 1106 (P=O), 3426 br (OH); HRMS (ES^-) m/z 497.1068 $[M - H]^-$ (required 497.1066).



6-O-phenoxyphosphoryl-D-trehalose (3.27)

R_f 0.63 (5 ethanol : 3 NH_4OH : 1 water), $[\alpha]_D^{19}$ +82.7 (c = 0.41, H_2O); ^1H NMR (500MHz, D_2O) δ ppm 3.35 (1H, t, $J_{H3'-H4'}$ 9.5Hz, $J_{H4'-H5'}$ 9.5Hz, H4'), 3.43 (1H, t, J_{H3-H4} 9.6Hz, J_{H4-H5} 9.6Hz, H4), 3.51 (1H, dd, $J_{H2'-H3'}$ 9.9Hz, $J_{H1'-H2'}$ 3.9Hz, H2'), 3.53 (1H, dd, J_{H2-H3} 9.8Hz, J_{H1-H2} 3.8Hz, H2), 3.67 (1H, dd, $J_{H6'a-H6'b}$ 11.9Hz, $J_{H5'-H6'b}$ 5.0Hz, H6'b), 3.70-3.79 (4H, m, H3, H3', H5', H6'a), 3.86 (2 H, ddd, J_{H4-H5} 9.9Hz, J_{H5-H6b} 2.0Hz, J_{H5-H6a} 1.9Hz, H5), 4.09 (1H, ddd, $J_{H6a-H6b}$ 11.6Hz, $J_{H6b-31P}$ 6.7Hz, J_{H5-H6b} 4.7Hz, H6b), 4.14 (1H, ddd, $J_{H6a-H6b}$ 11.5Hz, $J_{H6a-31P}$ 5.8Hz, J_{H5-H6a} 1.9Hz, H6a), 5.06 (1H, d, $J_{H1'-H2'}$ 3.8Hz, H1'), 5.08 (1H, d, J_{H1-H2} 3.8Hz, H1), 7.10-7.15 (3H, m, ArH_{ortho}, ArH_{para}), 7.32 (2H, t, $J_{ortho-meta}$ 7.8Hz, $J_{meta-para}$ 7.8Hz, ArH_{meta}); ^{13}C NMR (126MHz, D_2O) δ ppm 60.5 (1C, C6'), 64.8 (1C, d, J_{C6-31P} 4.8Hz, C6), 69.2 (1C, C4), 69.7 (1C, C4'), 70.9 (1C, C2'), 70.9 (1C, C2), 71.1 (1C, d, J_{C5-31P} 7.6Hz, C5), 72.1 (1C, C5'), 72.4 (1C, C3'), 72.4 (1C, C3), 93.3 (1C, C1'), 93.3 (1C, C1), 120.2 (2C, d, J_{C-31P} 4.8Hz, ArC_{ortho}), 124.3 (1C, ArC_{para}), 129.7 (2C, ArC_{meta}), 151.7 (1C, d, J_{C-31P} 6.7Hz, ArC_{ipso}); $^{31}P\{^1H\}$ NMR (162MHz, D_2O) δ ppm -3.79 (1P, P(O)(OPh)(OH)); FT-IR (KBr disc) υ 1156 (P=O), 3417 br (OH); HRMS (ES^-) m/z 497.1068 $[M - H]^-$ (required 497.1066).

3.9.6 Probing the mechanism of CsF transesterfications

Transesterfication reactions were performed under a range of conditions (Table 3.3). Reactions were generally performed on **2.1** (20mg, 0.035mmol) with methanol under reflux conditions and with 10 equivalents of the transesterfication reagent, unless otherwise stated below. Reactions were monitored by TLC (1 water : 4 isopropanol : 4 ethyl acetate). After 16 hours the reaction was stopped and products separated by silica gel chromatography (1 water : 4 isopropanol : 4 ethyl acetate).

For entry B, only 2.5 equivalents of CsF were used. For entry C, triethylamine (4µl, 0.052mmol, 1eq) pre-dissolved in methanol (10ml) was added over a period of 2 hours, via a syringe pump, to a solution of **2.1** (20mg, 0.035mmol, 1eq) in methanol (15ml) under reflux. After the addition was complete, the reaction mixture was allowed to heat for a further hour after which TLC showed complete consumption of the starting carbohydrate. For entry F, TBAF was used as a 1M solution in THF. Or entries G and H, a 1 : 1 mixture of methanol and the stated solvent was used in the reaction mixture. For entry K, the pH was adjusted with triethylamine. For entry L, the pH was adjusted with 1M HCl in methanol.

Fractions containing **3.41** and **3.42** were pooled together and solvents removed *in vacuo* to obtain a total yield. Product distribution of the regioisomers was calculated from ^{1}H NMR (D_2O) resonance integrals for the 4 and 6 position protons of compounds **3.41** and **3.42** respectively (Figure 3.2). The mixture was further analysed by C_{18} HPLC after per-*O*-acetylation. A sample of the acetylated mixture was injected into a Phenomenex Synergi Fusion C_{18} column (4.6 x 150mm, 4µm) eluted with a mobile phase of 1 water : 1 acetonitrile at a flow rate of 1.0ml/min. Eluents were detected with an in-line ELS detector and the peak integrals used to confirm the product distributions (Figure 3.3).

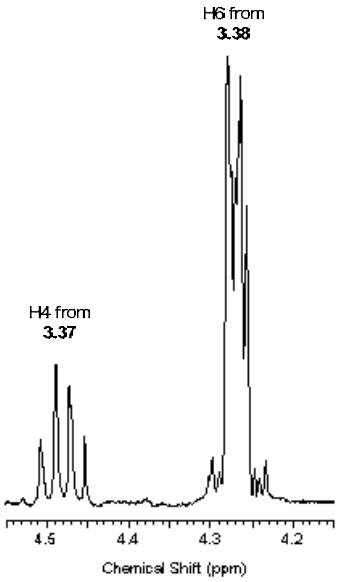


Figure 3.2. *^{1}H NMR signals for the 4 and 6 positions of compounds **3.41** and **3.42** respectively.*

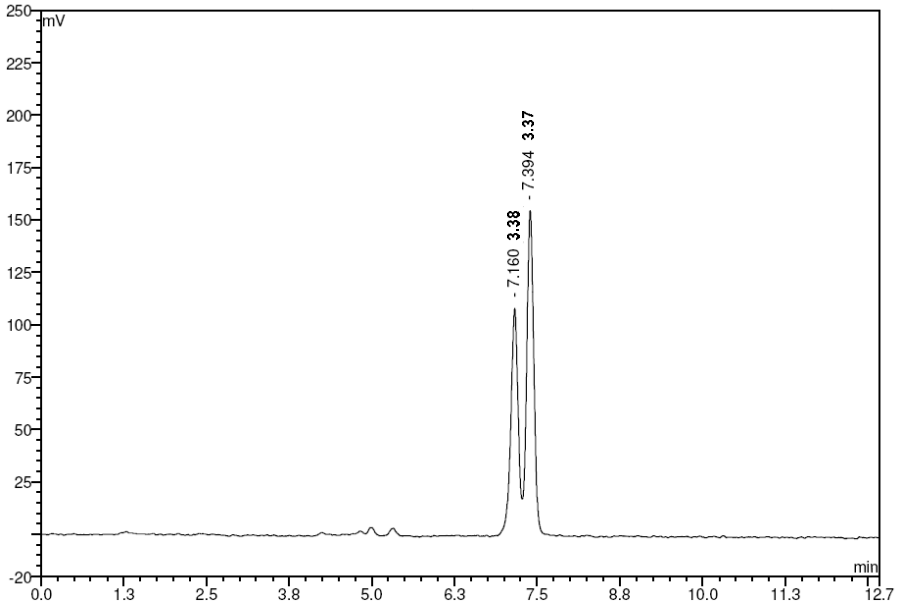
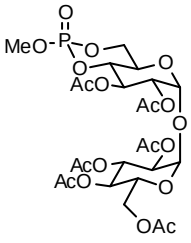


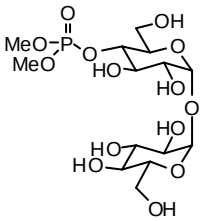
Figure 3.3. *HPLC trace of per-*O*-acetylated mixture of **3.41** and **3.42**.*

Compound **3.40** was per-*O*-acetylated to aid in characterization.



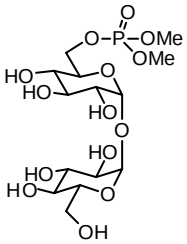
2,2',3,3',4',6'-O-acetyl-4,6-(monomethyl)phosphoryl-D-trehalose R_f 0.34 (ethyl acetate), $[\alpha]_D^{25}$ + 127 (c = 1.25, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.05, 2.06, 2.08, 2.10, 2.11, 2.15 (6 x 1C, 6 x $CO\underline{C}H_3$), 3.86 (3H, d, $J_{H\text{-}31P}$ 11.0Hz, P-OMe), 4.01 (1H, dd, $J_{H6'a\text{-}H6'b}$ 12.3Hz, $J_{H5'\text{-}H6'b}$ 1.9Hz, H6'b), 4.06 (1H, ddd, $J_{H4'\text{-}H5'}$ 10.4Hz, $J_{H5'\text{-}H6'a}$ 5.7Hz, $J_{H5\text{-}H6'b}$ 1.9Hz, H5), 4.11-4.21 (4H, m, H4, H5, H6a, H6b), 4.24 (1H, dd, $J_{H6'a\text{-}H6'b}$ 12.3Hz, $J_{H5'\text{-}H6'a}$ 5.7Hz, H6'a), 4.97 (1H, dd, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H1\text{-}H2}$ 4.1Hz, H2), 4.99 (1H, dd, $J_{H2'\text{-}H3'}$ 10.1Hz, $J_{H1'\text{-}H2'}$ 3.8Hz, H2'), 5.05 (1H, t, $J_{H3'\text{-}H4'}$ 9.9Hz,

$J_{H4'\text{-}H5'}$ 9.9Hz, H4'), 5.27 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 5.34 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.48 (1H, t, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H3\text{-}H4}$ 9.8Hz, H3), 5.58 (1H, t, $J_{H2'\text{-}H3'}$ 9.5Hz, $J_{H3'\text{-}H4'}$ 9.5Hz, H3'); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.5, 20.6, 20.6 (6C, m, CO<u>C</u>H$_3$), 54.3 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, P-OMe), 61.7 (1C, C6'), 63.0 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, C5), 68.1 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C6), 68.3 (1C, C5'), 68.3 (1C, C4') 68.7 (1C, d, $J_{C\text{-}31P}$ 9.5Hz, C3), 69.8 (1C, C2'), 69.9 (1C, C2), 70.1 (1C, C4), 92.4 (1C, C1), 92.9 (1C, C1'), 169.0, 169.6, 169.6, 169.7, 170.0, 170.5 (6 x 1C, 6 x C=O); FT-IR (thin film) υ 1224 (P=O), 1751 (C=O); HRMS (ES^+) m/z 693.1409 $[M + Na]^+$ (required 693.1408.



4-O-(dimethoxyphosphoryl)-d-trehalose (3.41)

R_f 0.36 (1 water : 4 isopropanol : 4 ethyl acetate), $[\alpha]_D^{18}$ +41.1 (c = 0.67, H_2O); ^{1}H NMR (500MHz, MeOD) δ ppm 3.37 (1H, as, H4'), 3.50 (1H, dd, $J_{H2'\text{-}H3'}$ 9.8Hz, $J_{H1'\text{-}H2'}$ 3.8Hz, H2'), 3.58 (1H, t, $J_{H2\text{-}H3}$ 9.6Hz, $J_{H3\text{-}H4}$ 9.6Hz, H3), 3.69 (1H, dd, $J_{H6'a\text{-}H6'b}$ 9.6Hz, $J_{H5'\text{-}H6b}$ 3.2Hz, H6'b), 3.72 (1H, dd, $J_{H2\text{-}H3}$ 9.6Hz, $J_{H1\text{-}H2}$ 3.7Hz, H2), 3.75 (1H, dd, $J_{H6'a\text{-}H6'b}$ 10.8Hz, $J_{H5'\text{-}H6'a}$ 5.0Hz, H6'a), 3.82 (2H, m, H6a, H6b), 3.84 (6H, d, $J_{H\text{-}31P}$ 11.0Hz, 2 x OMe), 3.85 (2H, m, H3', H5'), 3.86 (1H, m , H5), 4.59 (1H, q, $J_{H3\text{-}H4}$ 8.9Hz, $J_{H4\text{-}H5}$ 8.9Hz, $J_{H4\text{-}31P}$ 8.9Hz, H4), 5.13 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.19 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (126MHz, MeOD) δ ppm 55.3 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, OMe), 55.3 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, OMe), 62.2 (1C, C6'), 62.6 (1C, C6), 70.3 (1C, d, $J_{C\text{-}31P}$ 2.9Hz, C3), 72.0 (1C, C2), 72.0 (1C, C4), 73.2 (1C, C2'), 73.8 (1C, C3'), 73.9(1C, C5), 74.5 (1C, C5'), 83.3 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, H4), 95.1 (2C, C1, C1'); FT-IR (KBr disc) υ 1258 (P=O), 3501 br (OH); HRMS (ES^+) m/z 473.1029 $[M + Na]^+$ (required 473.1031).



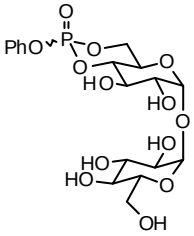
6-O-(dimethoxyphosphoryl)-D-trehalose (3.42)

R_f 0.12 (1 water : 2 isopropanol : 4 ethyl acetate), $[\alpha]_D^{18}$ +83.3 (c = 1.0, H_2O); ^{1}H NMR (500MHz, D_2O) δ ppm 3.36 (1H, t, $J_{H3'\text{-}H4'}$ 9.5Hz, $J_{H4'\text{-}H5'}$ 9.5Hz, H4'), 3.43 (1H, dd, $J_{H3\text{-}H4}$ 10.1Hz, $J_{H4\text{-}H5}$ 9.1Hz, H4), 3.56 (1H, t, $J_{H2\text{-}H3}$ 10.4Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2'), 3.57 (1H, t, $J_{H2\text{-}H3}$ 10.4Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2'), 3.67 (1H, dd, $J_{H6'a\text{-}H6'b}$ 11.5Hz, $J_{H5'\text{-}H6'b}$ 5.0Hz, H6'b), 3.72 (1H, m, H3'), 3.75 (1H, d, $J_{H\text{-}31P}$ 11.0Hz, OMe), 3.75-3.81 (3H, m, H3, H5', H6'a), 3.76 (1H, d, $J_{H\text{-}31P}$ 11.0Hz, OMe), 3.91 (1H, dt, $J_{H4\text{-}H5}$ 10.1Hz, $J_{H5\text{-}H6a}$ 2.6Hz, $J_{H5\text{-}H6b}$ 2.6Hz, H5), 4.25-4.29 (2H, m, H6a, H6b), 5.10 (1H, d, $J_{H1'\text{-}H2}$ 3.8Hz, H1'), 5.13 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 55.1 (2C, d, $J_{C\text{-}31P}$

8.8Hz, 2 x OMe), 60.5 (1C, C6'), 66.7 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, C6), 69.0 (1C, C4), 69.6 (1C, C4'), 70.5 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, C5), 70.9 (1C, C2), 71.0 (1C, C2'), 72.2 (1C, C5'), 72.4 (1C, C3), 72.5 (1C, C3'), 93.4 (1C, C1), 93.5 (1C, C1'); FT-IR (KBr disc) υ 1260 (P=O), 3486 br (OH); HRMS (ES^+) m/z 473.1027 $[M + Na]^+$ (required 473.1031).

Test for fluoride attack at the phosphorous centre

A suspension of **2.1** (20mg, 0.035mmol, 1eq) and CsF (66mg, 0.35mmol, 10eq) in DMF (3ml) was sonicated for 15 minutes to aid dissolution of CsF. The resultant solution was then heated at 80°C for 16 hours after which TLC (1 water : 4 isopropanol : 4 ethylacetate) showed complete consumption of the starting material. The solvent was removed by lyophilization. Silica gel chromatography (1 water : 4 isopropanol : 4 ethyl acetate) followed by lyophilization allowed isolation of **3.39** as a mixture of diastereomers at the phosphorus centre (15.5mg, 92%; 1 equatorial : 0.43 axial).



4,6-(monophenyl)phosphoryl-D –trehalose (3.39)

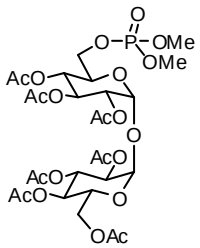
R_f 0.6 (1 water : 4 isopropanol : 4ethyl acetate), $[\alpha]_D^{25}$ +75.5 (c = 0.78, H_2O); FT-IR (KBr disc) 1221 (P=O), 3245 br (OH); HRMS m/z (ES^-) 479.0960 $[M - H]^-$ required 479.0960;

Equatorial Diastereomer ^{1}H NMR (500MHz, D_2O) δ ppm 3.38 (1H, t, $J_{H3'\text{-}H4'}$ 9.5Hz, $J_{H4'\text{-}H5'}$ 9.5Hz, H4'), 3.58 (1H, dd, $J_{H2'\text{-}H3'}$ 10.1Hz, $J_{H1'\text{-}H2'}$ 3.8Hz, H2'), 3.70 (1H, m, H6'b), 3.70 (1H, dd, $J_{H2\text{-}H3}$ 10.0Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.65-3.80 (2H, m, H3', H5', H6'a), 4.09 (1H, t, $J_{H2\text{-}H3}$ 9.1Hz, $J_{H3\text{-}H4}$ 9.1Hz, H3), 4.32 (1H, m, H5), 4.38 (1H, t, $J_{H3\text{-}H4}$ 9.8Hz, $J_{H4\text{-}H5}$ 9.8Hz, H4), 4.43 (1H, t, $J_{H5\text{-}H6ax}$ 9.8Hz, $J_{H6ax\text{-}H6eq}$ 9.8Hz, H6ax), 4.52 (1H, ddd, $J_{H6eq\text{-}31P}$ 23.9Hz, $J_{Hax\text{-}Heq}$ 9.8Hz, $J_{H5\text{-}Heq}$ 4.7Hz, H6eq), 5.12 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.18 (1H, d, $J_{H1\text{-}H2}$ 4.1Hz, H1), 7.20-7.29 (3H, m, ArH_{ortho}, ArH_{para}), 7.41 (2H, t, $J_{ortho\text{-}meta}$ 7.9Hz, $J_{meta\text{-}para}$ 7.9Hz, ArH_{meta}); ^{13}C NMR (126MHz, D_2O) δ ppm 60.4 (1C, C6), 62.4 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, C4), 69.5 (1C, C4'), 69.7 (1C, C3), 69.8 (1C, d, $J_{C\text{-}31P}$ 8.6Hz, C6), 70.6 (1C, d, $J_{C\text{-}31P}$ 1.9Hz, C2), 70.9 (1C, C2'), 72.3 (1C, C5'), 72.5 (1C, C3'), 80.9 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, C5), 94.0 (1C, C1'), 94.1 (1C, C1), 119.6 (2C, d, $J_{C\text{-}31P}$ 4.8Hz, ArC_{ortho}), 126.2 (1C, ArC_{para}), 130.4 (2C, ArC_{meta}), 149.4 (1C, $J_{C\text{-}31P}$ 6.7Hz, ArC_{ipso}).

Axial Diastereomer ^{1}H NMR (500MHz, D_2O) δ ppm 3.35 (1H, t, $J_{H3'\text{-}H4'}$ 9.5Hz, $J_{H4'\text{-}H5'}$ 9.5Hz, H4'), 3.52 (1H, dd, $J_{H2'\text{-}H3'}$ 10.1Hz, $J_{H1'\text{-}H2'}$ 3.8Hz, H2'), 3.64-3.81 (5H, m, H2, H3', H5', H6'a, H6'b), 3.99 (1H, t, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H3\text{-}H4}$ 9.5Hz, H3), 4.19 (1H, td, $J_{H3\text{-}H4}$ 10.3Hz, $J_{H4\text{-}H5}$ 10.3Hz, $J_{H4\text{-}31P}$ 6.0Hz, H4), 4.30-4.41 (3H, m, H5, H6ax, H6eq), 5.04 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.14 (1H, d, $J_{H1\text{-}H2}$ 4.1Hz, H1), 7.20-7.29 (3H, m, ArH_{ortho}, ArH_{para}), 7.37-7.43 (2H, m, ArH_{meta}); ^{13}C NMR (126MHz, D_2O) δ ppm 60.4 (1C, C6'), 62.1 (1C, d, $J_{C\text{-}31P}$ 8.6Hz, C4), 69.5 (1C, C4'), 69.6 (1C, C3), 70.0 (1C, d, $J_{C\text{-}31P}$ 8.6Hz, C6), 70.5 (1C, d, $J_{C\text{-}31P}$ 2.9Hz, C2), 70.8 (1C, C2'), 72.3 (1C, C3'), 72.4 (1C, C5'), 80.1 (1C, d, $J_{C\text{-}31P}$ 4.8Hz, C5), 93.7 (1C, C1'), 93.8 (1C, C1), 120.1 (2C, d, $J_{C\text{-}31P}$ 3.8Hz, ArC_{ortho}), 126.5 (1C, ArC_{para}), 130.3 (1C, ArC_{meta}), 149.6 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, ArC_{ipso}).

Transesterfication in ionic liquid

2.1 (15mg, 0.026mmol, 1eq) and NEt_3 (4µl, 0.12mmol, 4.4eq) was dissolved in $MMIM^{+}PO(OMe)_2O^{-}$ (0.5ml) and methanol (0.5ml) and heated at 80°C for 18 hours. The reaction mixture was cooled to room temperature and the methanol removed under reduced pressure. The crude mixture was dissolved in pyridine (1ml) and acetic anhydride was added (1ml, 9.1mmol, 350eq). The reaction was allowed to proceed at room temperature, after which TLC (ethyl acetate) showed the presence of a sinlge compound. The reaction was quenched with methanol (2ml) and the organic solvents were removed in vacuo. The ionic liquid was extracted with ethyl acetate (6 x 10ml) and the combined organics were washed with water (10ml), brine (10ml), dried over $MgSO_4$ and concentrated in vacuo. The compound was purified by silica gel chromatography (1 petrol : 4 ethyl acetate) to give the title compound as a colourless oil (11.1mg, 57%).



6-O-(dimethoxyphosphoryl)-2,2',3,3',4,4',6'-O-acetyl-D-trehalose (3.48)

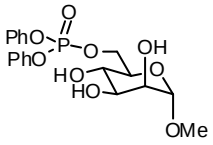
R_f 0.25 (1 petrol : 4 ethyl acetate), $[\alpha]_D^{25}$ +80.0 (c = 0.56, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.03, 2.04, 2.05, 2.08, 2.08, 2.09. 2.09 (7 x 3H, 7 x s, 7 x OAc), 3.76 (3H, d, $J_{H\text{-}31P}$ 11.0Hz, OMe), 3.77 (3H, d, $J_{H\text{-}31P}$ 11.3Hz, OMe), 4.01 (1H, dd, $J_{H6'a\text{-}H6'b}$ 12.3Hz, $J_{H5'\text{-}H6'b}$ 1.9Hz, H6'b), 4.04 (1H, m, H6b), 4.06-4.10 (2H, m, H5, H5'), 4.12 (1H, m, H6a), 4.25 (1H, dd, $J_{H6'a\text{-}H6'b}$ 11.8Hz, $J_{H5'\text{-}H6'a}$ 5.7Hz, H6'a), 5.03 (1H, dd, $J_{H2'\text{-}H3'}$ 10.4Hz, $J_{H1'\text{-}H2'}$

3.8Hz, H2'), 5.07 (1H, m, H4'), 5.07 (1H, dd, $J_{H2\text{-}H3}$ 10.7Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 5.08 (1H, t, $J_{H3\text{-}H4}$ 9.8Hz, $J_{H4\text{-}H5}$ 9.8Hz, H4), 5.28 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.31 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 5.49 (1H, t, $J_{H2'\text{-}H3'}$ 9.8Hz, $J_{H3'\text{-}H4'}$ 9.8Hz, H3'), 5.50 (1H, t, $J_{H3\text{-}H4}$ 9.8Hz, $J_{H3\text{-}H4}$ 9.8Hz, H3); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.5, 20.5, 20.6, 20.6, 20.6, 20.7, 20.7 (7 x 1C, 7 x CO<u>C</u>H$_3$), 54.6 (2C, d, $J_{C\text{-}31P}$ 5.7Hz, 2 x OMe), 61.7 (1C, C6'), 65.2 (1C, d, $J_{C\text{-}31P}$ 4.8Hz, C6), 68.2 (1C, C4'), 68.2 (1C, C4), 68.4 (1C, C5'), 68.8 (1C, d, $J_{C\text{-}31P}$ 8.6Hz, C5), 69.6 (1C, C2'), 69.8 (1C, C2), 69.9 (1C, C3'), 70.0 (1C, C3), 92.5 (1C, C1'), 92.6 (1C, C1), 169.5, 169.5, 169.6, 169.6, 170.0, 170.0, 170.6 (7 x 1C, 7 x C=O); FT-IR (thin film) υ 1221 (P=O), 1750 (C=O); HRMS m/z (ES^+) 767.1768 $[M + Na]^+$ required 767.1770.

3.9.7 Installation of diphenylphosphate functionality

For synthesis of **2.1** see section 2.8.2 page 45.

For synthesis of **3.9** see section **3.9.2** page 91.

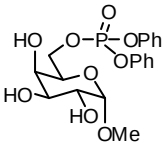


α-methyl 6-O-(diphenoxyphosphoryl)- D-mannopyranoside[61] (3.49)

DPCP (0.53ml, 2.58mmol, 1eq) was added to a stirred suspension of α-methylmannopyranoside (0.50g, 2.58mmol, 1eq) and DMAP (31mg, 0.26mmol, 0.1eq) in dry pyridine (10ml). After stirring at room temperature for 3 hours, TLC (1 water : 4 isopropanol : 4 ethyl acetate) indicated the complete consumption of starting material. The reaction was quenched with methanol (5ml), concentrated *in vacuo*, and purified by silica gel chromatography (ethyl acetate) to give the desired compound as a colourless oil (0.82g, 75%).

R_f 0.21 (ethyl acetate); $[\alpha]_D^{21.2}$ +36.6 (c = 1.0, $CHCl_3$), lit.[61] $[\alpha]_D^{25}$ +44.1 (c = 1, $CHCl_3$); ^{1}H NMR (500MHz, MeOD) δ ppm 3.31 (3H, s, OMe), 3.66 (1H, t, $J_{H3\text{-}H4}$ 9.1Hz, $J_{H4\text{-}H5}$ 9.1Hz, H4), 3.70 (1H, dd, $J_{H3\text{-}H4}$ 9.0Hz, $J_{H2\text{-}H3}$ 3.5Hz, H3), 3.74 (1H, add, $J_{H4\text{-}H5}$ 8.5Hz, $J_{H5\text{-}H6'}$ 6.9Hz, H5), 3.83 (1H, dd, $J_{H2\text{-}H3}$ 3.0Hz, $J_{H1\text{-}H2}$ 1.7Hz, H2), 4.47 (1H, ddd, $J_{H6\text{-}H6'}$ 10.9Hz, $J_{H6'\text{-}31P}$ 7.6Hz, $J_{H5\text{-}H6'}$ 6.8Hz, H6'), 4.63 (1H, ddd, $J_{H6\text{-}H6'}$ 10.9Hz, $J_{H6\text{-}31P}$ 6.6Hz, $J_{H5\text{-}H6}$ 1.7Hz, H6), 4.66 (1H, d, $J_{H1\text{-}H2}$ 1.3Hz, H1), 7.25 (2H, d, $J_{meta\text{-}para}$ 7.3Hz, ArH_{para}), 7.28 (4H, dd, $J_{ortho\text{-}meta}$ 8.7Hz, $J_{ortho\text{-}ortho}$ 1.0Hz, ArH_{ortho}), 7.41 (4H, t, $J_{ortho\text{-}meta}$ 7.3Hz, $J_{meta\text{-}para}$ 7.3Hz, ArH_{meta}); ^{13}C NMR (126MHz, MeOD) δ ppm 55.4 (1C, OMe), 68.1 (1C, C4), 70.4 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, C6), 71.9 (1C, C2),

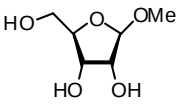
72.5 (1C, C3), 72.9 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C5), 102.9 (1C, C1), 121.3 (2C, d, $J_{C\text{-}31P}$ 4.8Hz, ArC_{ortho}), 121.3 (2C, d, $J_{C\text{-}31P}$ 4.8Hz, ArC_{ortho}), 126.8 (2C, ArC_{para}), 131.1 (4C, ArC_{meta}), 151.9 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, ArC_{ipso}), 151.9 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, ArC_{ipso}); $^{31}P\{^{1}H\}$ (162MHz, $CDCl_3$) δ ppm -11.42 (1P, $P(O)(OR)(OPh)_2$); m/z (ES^+) 448.90 $[M + Na]^+$.



α-methyl-6-O-(diphenoxyphosphoryl)-D-galactopyranoside[62]

DPCP (2.46ml, 12.9mmol, 1eq) was added dropwise to a stirred solution of α-methyl-galactopyranoside (2.30g, 12.9mmol, 1eq) in dry pyridine (50ml) over 30 minutes. After stirring at room temperature for 19 hours, TLC (1 methanol : 9 ethyl acetate) indicated the complete consumption of starting material. The reaction was quenched with methanol (5ml), concentrated *in vacuo* and co-evaporated with toluene (x3) to remove pyridine. Automated silica gel chromatography (2% to 20% methanol in ethyl acetate over 9 column volumes) gave the desired compound as a white amorphous solid (3.96g, 72%).

R_f 0.24 (1 methanol : 9 ethyl acetate), $[\alpha]_D^{25}$ +74.0 (c = 1.0, MeOH); ^1H NMR (500MHz, MeOD) δ ppm 3.31 (1H, s, OMe), 3.73 (1H, dd, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H3\text{-}H4}$ 3.2Hz, H3), 3.80 (1H, dd, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.88 (1H, d, $J_{H3\text{-}H4}$ 3.2Hz, H4), 4.00 (1H, dd, $J_{H5\text{-}H6}$ 7.9Hz, $J_{H5\text{-}H6'}$ 4.1Hz, H5), 4.44 (1H, ddd, $J_{H6\text{-}H6'}$ 10.7Hz, $J_{H6'\text{-}31P}$ 7.3Hz, $J_{H5\text{-}H6'}$ 4.1Hz, H6'), 4.49 (1H, dd, $J_{H6\text{-}H6'}$ 10.4Hz, $J_{H5\text{-}H6}$ 7.9Hz, H6), 4.72 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 7.27 (6H, t, $J_{ortho\text{-}meta}$ 7.3Hz, $J_{meta\text{-}para}$ 7.3Hz, ArH_{ortho}, ArH_{para}), 7.42 (1H, t, $J_{ortho\text{-}meta}$ 7.6Hz, $J_{meta\text{-}para}$ 7.6 Hz, Ar_{meta}), 7.42 (1H, t, $J_{ortho\text{-}meta}$ 7.6Hz, $J_{meta\text{-}para}$ 7.6 Hz, Ar_{meta}); ^{13}C NMR (126MHz, MeOD) δ ppm 55.8 (1C, OMe), 70.0 (1C, C2), 70.4 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, C6), 70.6 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C5), 70.7 (1C, C4), 71.1 (1C, C3), 101.5 (1C, C1), 121.2 (2C, d, $J_{C\text{-}31P}$ 4.8Hz, ArC_{ortho}), 121.2 (2C, d, $J_{C\text{-}31P}$ 4.8Hz, ArC_{ortho}), 126.9 (2C, ArC_{para}), 131.1 (4C, ArC_{meta}), 151.8 (2C, d, $J_{C\text{-}31P}$ 7.6Hz, ArC_{ipso}); $^{31}P\{^{1}H\}$ NMR (162MHz, MeOD) δ ppm -11.8 (1P, $P(O)(OPh)_2$); m/z $(ES)^+$ 449.1 $[M + Na]^+$, 875.2 $[2M + Na]^+$ (ES^-) 425.1 $[M - H]^-$, 461.1 $[M + {}^{35}Cl]^-$, 485.2 $[M + OAc]^-$.

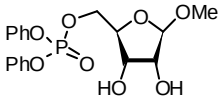


β-methyl-ribofuranoside[63] (3.71)

A mixture of D-ribose (5.0g, 33.3mmol, 1eq) and Dowex 50WX8 100-200 mesh (H^+ form) was stirred at room temperature. TLC (85 DCM : 15 methanol) after 13 hours showed complete consumption of starting material. The

reaction mixture was filtered and the solvent removed *in vacuo*. The desired product was isolated by automated silica gel chromatography (2% to 20% methanol in ethyl acetate over 9 column volumes) as a colourless oil (2.51g, 46%).

R_f 0.49 (8 ethyl acetate : 2 methanol), $[\alpha]_D^{25}$ -48.9 (c = 1.0, H_20), lit.[63] $[\alpha]_D^{20}$ -50 (c = 2.0, H_20); ^{1}H NMR (400MHz, D_2O) δ ppm 3.28 (3H, s, OMe), 3.49 (1H, dd, $J_{H5\text{-}H5'}$ 12.5Hz, $J_{H4\text{-}H5'}$ 6.5Hz, H5'), 3.68 (1H, dd, $J_{H5\text{-}H5'}$ 12.4Hz, $J_{H4\text{-}H5}$ 3.3Hz, H5), 3.88 (1H, td, $J_{H3\text{-}H4}$ 6.8Hz, $J_{H4\text{-}H5'}$ 6.8Hz, $J_{H4\text{-}H5}$ 3.4Hz, H4), 3.92 (1H, dd, $J_{H2\text{-}H3}$ 4.8Hz, $J_{H1\text{-}H2}$ 0.9Hz, H2), 4.04 (1H, dd, $J_{H3\text{-}H4}$ 6.8Hz, $J_{H2\text{-}H3}$ 4.8Hz, H3), 4.78 (1H, s, H1); ^{13}C NMR (101MHz, D_2O) δ ppm 55.4 (1C, OMe), 63.0 (1C, C5), 71.0 (1C, C3), 74.5 (1C, C2), 83.1 (1C, C4), 108.2 (1C, C1); m/z (ES^+) 187.1 $[M + Na]^+$, 351.1 $[2M + Na]^+$ (ES^-) 163.1 $[M - H]^-$.



β-methyl-5-(diphenoxyphosphoryl)-ribofuranoside (3.57)

Diphenylchlorophosphate (1.78ml, 8.59mmol, 1eq) was added dropwise to a stirred solution of **3.71** (1.40g, 8.59mmol, 1eq) in dry pyridine (40ml) over 5 minutes. After stirring at room temperature for 15 hours, TLC (1 methanol : 9 ethyl acetate) indicated the complete consumption of starting material. The reaction was quenched with methanol (10ml), concentrated *in vacuo* and co-evaporated with toluene (x3) to remove pyridine. Automated silica gel chromatography (7 ethyl acetate : 3 petrol)gave the desired compound as a colourless oil (3.26g, 96%).

R_f 0.26 (7 ethyl acetate : 3 petrol), $[\alpha]_D^{25}$ -20.3 (c = 1.0, MeOH); ^{1}H NMR (500MHz, MeOD) δ ppm 3.31 (3H, s, OMe), 3.90 (1H, d, $J_{H2\text{-}H3}$ 3.2Hz, H2), 4.14-4.17 (2H, as, H3, H4), 4.31 (1H, dddd, $J_{H5\text{-}H5'}$ 11.1Hz, $J_{H5'\text{-}31P}$ 6.3Hz, $J_{H4\text{-}H5'}$ 4.0Hz, $J_{H3\text{-}H5'}$ 1.9Hz, H5'), 4.49 (1H, ddd, $J_{H5\text{-}H5'}$ 11.1Hz, $J_{H5\text{-}31P}$ 6.3Hz, $J_{H4\text{-}H5}$ 1.9Hz, H5), 4.79 (1H, s, H1), 7.24-7.31 (1H, m, ArH_{para}), 7.28 (2H, d, $J_{ortho\text{-}meta}$ 6.9Hz, ArH_{ortho}), 7.42 (2H, t, $J_{ortho\text{-}meta}$ 7.3Hz, $J_{meta\text{-}para}$ 7.3Hz, ArH_{meta}); ^{13}C NMR (126MHz, MeOD) δ ppm 55.4 (1C, OMe), 71.4 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, C5), 72.1 (1C, C3), 76.0 (1C, C2), 82.1 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C4), 110.0 (1C, C1), 121.2 (2C, d, $J_{C\text{-}31P}$ 4.8Hz, ArC_{ortho}), 121.2 (2C, d, $J_{C\text{-}31P}$ 4.8Hz, ArC_{ortho}), 126.9 (2C, ArC_{para}), 131.1 (4C, d, $J_{C\text{-}31P}$ 2.9Hz, ArC_{meta}), 151.8 (2C, d, $J_{C\text{-}31P}$ 7.6Hz, ArC_{ipso}); $^{31}P\{^1H\}$ NMR (162MHz, MeOD) δ ppm -12.0 (1P, $P(O)(OPh)_2$); HRMS m/z (ES^+) 419.0871 $[M + Na]^+$ (required 419.0866).

3.9.8 CsF transesterfication reactions

General conditions

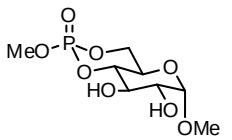
A solution of carbohydrate (50mg, 1eq) and CsF (10eq) in methanol (5ml) was heated to 80°C with vigorous stirring for 16 hours and monitored by TLC. The solvent was removed *in vacuo* and the products were purified by silica gel chromatography according to the conditions described below. The solvent was removed under reduced pressure to yield the purified compounds.

Trehalose

For transesterfications on trehalose, see section 3.9.6 page 108. The products were purified by two rounds of automated silica gel chromatography using 1 water : 4 isopropanol : 4 ethyl acetate.

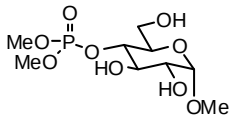
Glucose

The reaction was performed as above and the products purified by silica gel chromatography with 95 ethyl acetate : 5 methanol.



α-methyl-4,6-O-(*S*-methanephosphoryl)-D-glucopyranoside (3.36)

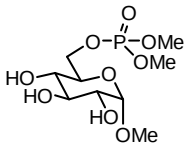
R_f 0.4 (95 ethyl acetate : 5 methanol), $[\alpha]_D^{25}$ +89.6 (c = 0.25, MeOH); ^{1}H NMR (500MHz, MeOD) δ ppm 3.46 (3H, s, C-OMe), 3.50 (1H, dd, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.83 (1H, t, $J_{H2\text{-}H3}$ 9.0Hz, $J_{H3\text{-}H4}$ 9.0Hz, H3), 3.87 (3H, d, $J_{H\text{-}31P}$ 11.0Hz, P-OMe), 3.97 (1H, t, $J_{H3\text{-}H4}$ 9.8Hz, $J_{H4\text{-}H5}$ 9.8Hz, H4), 4.00 (1H, td, $J_{H4\text{-}H5}$ 10.2Hz, $J_{H5\text{-}H6ax}$ 10.2Hz, $J_{H5\text{-}Heq}$ 4.4Hz, H5), 4.22 (1H, t, $J_{H5\text{-}H6ax}$ 9.9Hz, $J_{H6ax\text{-}H6eq}$ 9.9Hz, H6ax), 4.41 (1H, ddd, $J_{H6eq\text{-}31P}$ 23.8Hz, $J_{H6ax\text{-}H6eq}$ 10.1Hz, $J_{H5\text{-}H6eq}$ 4.4Hz, H6eq), 4.75 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (126MHz, MeOD) δ ppm 54.8 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, P-OMe), 56.2 (1C, C-OMe), 63.6 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, C5), 70.6 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C6), 72.1 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C3), 73.3 (1C, d, $J_{C\text{-}31P}$ 2.9Hz, C2), 82.5 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, C4), 101.8 (1C, C1); FT-IR (thin film) υ 1203 (P=O), 3329 br (OH); HRMS (ES^+) m/z 293.0403 $[M + Na]^+$ (required 293.0397).



α-methyl-4-*O*-(diphenylphosphoryl)-D-glucopyanoside (3.37)

R_f 0.25 (95 ethyl acetate : 5 methanol), $[\alpha]_D^{25}$ +89.7 (c = 0.22, MeOH); ^{1}H NMR (500MHz, MeOD) δ ppm 3.45 (3H, s, C-OMe), 3.52 (1H, dd, $J_{H2\text{-}H3}$ 9.9Hz, $J_{H3\text{-}H4}$ 8.8Hz, H3), 3.57 (1H, ddd, $J_{H4\text{-}H5}$ 7.7Hz, $J_{H5\text{-}H6'}$

5.1Hz, $J_{H5\text{-}H6}$ 2.5Hz, H5), 3.59 (1H, dd, $J_{H2\text{-}H3}$ 10.0Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.73 (1H, dd, $J_{H6\text{-}H6'}$ 11.0Hz, $J_{H5\text{-}H6'}$ 5.0Hz, H6'), 3.83 (1H, m, H6), 3.82 (3H, d, $J_{H\text{-}31P}$ 11.0Hz, P-OMe), 3.83 (3H, d, $J_{H\text{-}31P}$ 11.0Hz, P-OMe), 4.41 (1H, q, $J_{H3\text{-}H4}$ 8.5Hz, $J_{H4\text{-}H5}$ 8.5Hz, $J_{H4\text{-}31P}$ 8.5Hz, H4), 4.74 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (126MHz, MeOD) δ ppm 55.3 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, P-OMe), 55.3 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, P-OMe), 55.7 (1C, C-OMe), 62.3 (1C, C6), 70.3 (1C, d, $J_{C\text{-}31P}$ 2.9Hz, C3), 72.3 (1C, d, $J_{C\text{-}31P}$ 3.8Hz, C2), 73.5 (1C, C5), 83.7 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C4), 101.3 (1C, C1); FT-IR (thin film) υ 1278 (P=O), 3436 br (OH); HRMS m/z (FT-MS$^+$) 325.0659 $[M + Na]^+$ (required 325.0659).

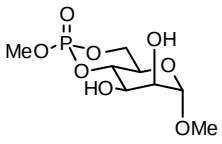


α-methyl-6-*O*-(diphenylphosphoryl)-D-glucopyanoside (3.38)

R_f 0.22 (95 ethyl acetate : 5 methanol), $[\alpha]_D^{25}$ +55.0 (c = 0.73, MeOH); ^{1}H NMR (500MHz, MeOD) δ ppm 3.32 (1H, t, $J_{H3\text{-}H4}$ 10.1Hz, $J_{H4\text{-}H5}$ 10.1Hz, H4), 3.41 (1H, dd, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.44 (3H, s, C-OMe), 3.64 (1H, t, $J_{H2\text{-}H3}$ 9.3Hz, $J_{H3\text{-}H4}$ 9.3Hz, H3), 3.71 (1H, ddd, $J_{H4\text{-}H5}$ 10.1Hz, $J_{H5\text{-}H6'}$ 5.4Hz, $J_{H5\text{-}H6}$ 1.6Hz, H5), 3.81 (3H, d, $J_{H\text{-}31P}$ 11.3Hz, P-OMe), 3.82 (3H, d, $J_{P\text{-}OMe}$ 11.0Hz, P-OMe), 4.23 (1H, ddd, $J_{H6\text{-}H6'}$ 11.0Hz, $J_{H6'\text{-}31P}$ 7.3Hz, $J_{H5\text{-}H6'}$ 5.4Hz, H6'), 4.32 (1H, ddd, $J_{H6\text{-}H6'}$ 11.0Hz, $J_{H6\text{-}31P}$ 6.1Hz, $J_{H5\text{-}H6}$ 2.2Hz, H6), 4.71 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (126MHz, MeOD) δ ppm 55.2 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, P-OMe), 55.7 (1C, C-OMe), 68.5 (1C, d, $J_{C\text{-}31P}$ 4.8Hz, C6), 71.2 (1C, C4), 71.8 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C5), 73.4 (1C, C2), 75.0 (1C, C3), 101.5 (1C, C1); FT-IR (thin film) υ 1261 (P=O), 3419 br (OH); HRMS m/z (ES$^+$) 325.0660 $[M + Na]^+$ (required 325.0659).

Mannose

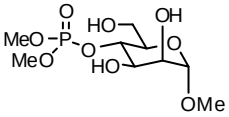
The reaction was performed as above and the products purified by two cycles of automated silica gel chromatography (2% to 20% methanol in ethyl acetate over 9 column volumes, repeated for mixed fractions).



α-methyl-4,6-(monomethyl)phosphoryl-D-mannopyranoside (3.50)

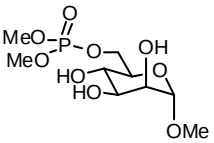
R_f 0.39 (9 ethyl acetate : 1 methanol), $[\alpha]_D^{25}$ +38.5 (c = 1.0, MeOH), m.p. 179-184°C (DCM/petrol); ^{1}H NMR (500 MHz, MeOD) δ ppm 3.42 (3H, s, C-OMe), 3.87 (3H, d, $J_{H\text{-}31P}$ 11.0Hz, P-OMe), 3.89 (1H, as, H2), 3.91 (1H, m, H3), 3.94 (1H, td, $J_{H4\text{-}H5}$ 10.3Hz, $J_{H5\text{-}H6ax}$ 10.3Hz, $J_{H5\text{-}H6eq}$ 5.0Hz, H5), 4.24 (1H, t, $J_{H5\text{-}H6ax}$ 10.4Hz, $J_{H6ax\text{-}H6eq}$ 10.4Hz, H6ax), 4.36 (1H, t, $J_{H3\text{-}H4}$ 9.1Hz, $J_{H4\text{-}H5}$ 9.1Hz, H4),

4.40 (1H, ddd, $J_{H6eq-31P}$ 23.6Hz, $J_{H6ax-H6eq}$ 10.1Hz, $J_{H5-H6eq}$ 5.0Hz, H6eq), 4.69 (1H, s, H1); ^{13}C NMR (126MHz, MeOD) δ ppm 54.7 (1C, d, J_{C-31P} 5.7Hz, P-OMe), 55.7 (1C, C-OMe), 64.6 (1C, d, J_{C-31P} 5.7Hz, C5), 69.7 (1C, d, J_{C-31P} 8.6Hz, C3), 70.6 (1C, d, J_{C-31P} 7.6Hz, C6), 72.2 (1C, d, J_{C-31P} 1.9Hz, C2), 80.6 (1C, d, J_{C-31P} 5.7Hz, C4), 103.6 (1C, C1); ^{31}P{^{1}H} NMR (162MHz, MeOD) δ ppm -4.2 (1P, P(O)OMe); FT-IR (thin film) υ 1285 (P=O), 3383 br (OH); HRMS m/z (ES^{+}) 293.0395 (required 293.0397).



α-methyl-4-*O*-(dimethoxyphosphoryl)-D-mannopyranoside (3.51)

R_f 0.25 (1 methanol : 9 chloroform), $[\alpha]_D^{25}$ +45.3 (c = 1.0, MeOH); ^{1}H NMR (500MHz, MeOD) δ ppm 3.41 (3H, s, C-OMe), 3.66 (1H, ddd, J_{H4-H5} 9.8Hz, $J_{H5-H6'}$ 5.5Hz, J_{H5-H6} 1.7Hz, H5), 3.75 (1H, dd, $J_{H6-H6'}$ 12.0Hz, $J_{H5-H6'}$ 5.7Hz, H6'), 3.82 (3H, d, J_{H-31P} 11.0Hz, P-OMe), 3.84 (3H, d, J_{H-31P} 11.3Hz, P-OMe), 3.83-3.87 (3H, m, H2, H3, H6), 4.43 (1H, q, J_{H3-H4} 9.0Hz, J_{H4-H5} 9.0Hz, J_{H4-31P} 9.0Hz, H4), 4.68 (1H, d, J_{H1-H2} 0.8Hz, H1); ^{13}C NMR (126MHz, MeOD) δ ppm 55.3 (1C, d, J_{C-31P} 5.7Hz, P-OMe), 55.4 (1C, C-OMe), 55.6 (1C, d, J_{C-31P} 6.7Hz, P-OMe), 62.3 (1C, C6), 71.1 (1C, d, J_{C-31P} 1.9Hz, C3), 72.5 (1C, C2), 72.8 (1C, d, J_{C-31P} 5.7Hz, C5), 76.6 (1C, d, J_{C-31P} 6.7Hz, C4), 102.5 (1C, C1); ^{31}P{^{1}H} NMR (162MHz, MeOD) δ ppm 0.8 (1P, P(O)(OMe)$_2$); FT-IR (thin film) υ 1259 (P=O), 3300 br (OH); HRMS m/z (ES^{+}) 325.0663 [M + Na]$^{+}$ (required 325.0659).

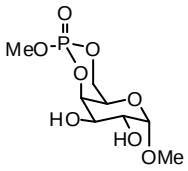


α-methyl-6-*O*-(dimethoxyphosphoryl)-D-mannopyranoside (3.52)

R_f 0.17 (9 ethyl acetate : 1 methanol), $[\alpha]_D^{25}$ +51.3 (c = 1.0, MeOH); ^{1}H NMR (500MHz, MeOD) δ ppm 3.40 (3H, s, C-OMe), 3.66 (1H, t, J_{H3-H4} 8.8Hz, J_{H4-H5} 8.8Hz, H4), 3.67 (1H, d, J_{H3-H4} 7.1Hz, H3), 3.80 (1H, m, H2), 3.82 (3H, d, J_{H-31P} 11.0Hz, P-OMe), 3.82 (3H, d, J_{H-31P} 11.3Hz, P-OMe), 4.25 (1H, ddd, $J_{H6-H6'}$ 11.0Hz, $J_{H6'-31P}$ 7.6Hz, $J_{H5-H6'}$ 5.0Hz, H6'), 4.36 (1H, ddd, $J_{H6-H6'}$ 11.0Hz, J_{H6-31P} 6.3Hz, J_{H5-H6} 1.3Hz, H6), 4.66 (1H, d, J_{H1-H2} 1.3Hz, H1); ^{13}C NMR (126MHz, MeOD) δ ppm 55.2 (1C, d, J_{C-31P} 4.8Hz, P-OMe), 55.3 (1C, d, J_{C-31P} 3.8Hz, P-OMe), 55.3 (1C, C-OMe), 68.0 (1C, C4), 68.7 (1C, d, J_{C-31P} 5.7Hz, C6), 71.9 (1C, C2), 72.5 (1C, C3), 72.9 (1C, d, J_{C-31P} 6.7Hz, C5), 102.9 (1C, C1); ^{31}P{^{1}H} NMR (162MHz, MeOD) δ ppm 1.2 (1P, P(O)(OMe)$_2$); FT-IR (thin film) υ 1255 (P=O), 3384 br (OH); HRMS m/z (ES^{+}) 325.0664 [M + Na]$^{+}$ (required 325.0659).

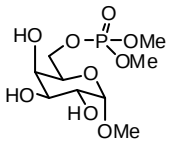
Galactose

The reaction was performed as above and the products purified by automated silica gel chromatography (2% to 20% methanol in ethyl acetate over 9 column volumes).



α-methyl-4,6-(monomethyl)phosphoryl-D-galactose (3.54)

R_f 0.17 (1 methanol : 9 DCM), $[\alpha]_D^{25}$ + 101.4 (c = 1.0, MeOH), m.p. 205-206°C; ^{1}H NMR (500MHz, MeOD) δ ppm 3.45 (3H, s, C-OMe), 3.84 (1H, dd, J_{H2-H3} 10.4Hz, J_{H1-H2} 3.8Hz, H2), 3.84 (3H, d, J_{H-31P} 11.0Hz, P-OMe), 3.91 (1H, dt, J_{H2-H3} 10.1Hz, J_{H3-H4} 3.8Hz, $J_{H3-H31P}$ 3.8Hz, H3), 3.94 (1H, m, H5), 4.42 (1H, ddd, $J_{H6eq-31P}$ 22.8Hz, $J_{H6ax-H6eq}$ 12.3Hz, $J_{H5-H6eq}$ 1.6Hz, H6eq), 4.54 (1H, d, $J_{H6ax-H6eq}$ 12.3Hz, H6ax), 4.77 (1H, d, J_{H3-H4} 2.8Hz, H4), 4.82 (1H, d, J_{H1-H2} 3.8Hz, H2); ^{13}C NMR (126MHz, MeOD) δ ppm 54.5 (1C, d, J_{C-31P} 5.7Hz, P-OMe), 56.2 (1C, C-OMe), 64.0 (1C, d, J_{C-31P} 5.7Hz, C5), 69.2 (1C, d, J_{C-31P} 6.7Hz, C3), 69.3 (1C, C2), 72.4 (1C, d, J_{C-31P} 7.6Hz, C6), 81.7 (1C, d, J_{C-31P} 6.7Hz, C4), 102.0 (1C, C1); ^{31}P{^{1}H} NMR (162MHz, MeOD) δ ppm -6.3 (1P, P(O)$(OR)_2$OH); FT-IR (KBr disc) υ 1278 (P=O), 3385 br (OH); HRMS (ES^+) m/z 293.0396 [M + Na] (required 293.0397).

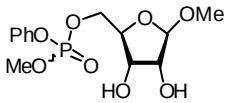


α-methyl-6-*O*-(dimethoxyphosphoryl)-D-galactopyranoside ()

R_f 0.12 (1 methanol : 9 DCM), $[\alpha]_D^{25}$ +91.4; ^{1}H NMR (500MHz, MeOD) δ ppm 3.43 (3H, s, C-OMe), 3.75 (1H, dd, J_{H2-H3} 10.1Hz, J_{H3-H4} 3.2Hz, H3), 3.76 (1H, dd, J_{H2-H3} 10.1Hz, J_{H1-H2} 3.2Hz, H2), 3.82 (6H, d, J_{H-31P} 11.0Hz, P-OMe), 3.89 (1H, dd, J_{H3-H4} 2.9Hz, J_{H4-H5} 1.2Hz, H4), 4.00 (1H, t, J_{H5-H6} 6.0Hz, $J_{H5-H6'}$ 6.0Hz, H5), 4.22 (2H, m, *J* 6.9Hz, H6, H6'), 4.74 (1H, d, J_{H1-H2} 3.5Hz, H1); ^{13}C NMR (126MHz, MeOD) δ ppm 55.2 (1C, d, J_{C-31P} 7.6Hz, P-OMe), 55.3 (1C, d, J_{C-31P} 5.7Hz, P-OMe), 55.8 (1C, OMe), 68.7 (1C, d, J_{C-31P} 4.8Hz, C6), 70.0 (1C, C2), 70.6 (1C, d, J_{C-31P} 7.6Hz, C5), 70.7 (1C, C4), 71.1 (1C, C3), 101.6 (1C, C1); ^{31}P{^{1}H} NMR (162MHz, MeOD) δ ppm 1.2 (1P, P(O)$(OPh)_2$); FT-IR (thin film) υ 1259 (P=O), 3385 br (OH); HRMS (ES^+) m/z 325.0656 $[M + Na]^+$ (required 325.0659).

Ribose

The reaction was performed as above and the products purified by automated silica gel chromatography (95 ethyl acetate : 5 methanol).



β-methyl-5-(methoxy-phenoxy-phosphoryl)-ribofuranoside (3.58)

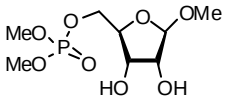
R_f 0.47 (5 methanol : 95 ethyl acetate), ^{31}P NMR (162MHz, MeOD) δ ppm -5.4 (P(O)OROPhOMe); FT-IR (thin film) υ 1203 (P=O), 3202 br (OH); HRMS (ES^+) m/z 357.0706 $[M + Na]^+$ (required 357.0710).

Major diastereomer

^{1}H NMR (500MHz, MeOD) δ ppm 3.35 (3H, s, C-OMe), 3.90 (1H, dd, J_{H2-H3} 3.9Hz, J_{H1-H2} 1.7Hz, H2), 3.93 (3H, d, J_{H-31P} 11.3Hz, P-OMe), 4.10-4.16 (2H, m, H3, H4), 4.20 (1H, td, $J_{H5a-H5b}$ 11.6Hz, J_{H4-H5b} 5.8Hz, $J_{H5b-31P}$ 5.8Hz, H5b), 4.34 (1H, ddd, $J_{H5a-H5b}$ 10.7Hz, $J_{H5a-31P}$ 6.6Hz, J_{H4-H5a} 2.2Hz, H5a), 4.79 (1H, s, H1), 7.25-7.29 (3H, m, ArH_{ortho}, ArH_{para}), 7.42 (2H, t, $J_{ortho-meta}$ 8.2Hz, $J_{meta-para}$ 8.2Hz, ArH_{meta}); ^{13}C NMR (126MHz, MeOD) δ ppm 55.4 (1C, C-O<u>C</u>H$_3$), 55.9 (1C, d, J_{C-31P} 4.8Hz, P-OMe), 70.8 (1C, d, J_{C-31P} 5.7Hz, C5), 72.2 (1C, C3), 76.0 (1C, C2), 82.2 (1C, d, J_{C-31P} 7.6Hz, C4), 110.0 (1C, C1), 121.1 (1C, d, J_{C-31P} 4.8Hz, ArC_{ortho}), 126.6 (1C, s, ArC_{para}), 131.0 (1C, s, ArC_{meta}), 151.9 (1C, d, J_{C-31P} 7.6Hz, ArC_{ipso}).

Minor diastereomer

^{1}H NMR (500MHz, MeOD) δ ppm 3.35 (3H, s, C-OMe), 3.90 (1H, dd, J_{H2-H3} 3.9Hz, J_{H1-H2} 1.7Hz, H2), 3.93 (3H, d, J_{H-31P} 11.3Hz, P-OMe), 4.10-4.16 (2H, m, H3, H4), 4.20 (1H, td, $J_{H5a-H5b}$ 11.6Hz, J_{H4-H5b} 5.8Hz, $J_{H5b-31P}$ 5.8Hz, H5b), 4.38 (1H, ddd, $J_{H5a-H5b}$ 10.7Hz, $J_{H5a-31P}$ 6.8Hz, J_{H4-H5a} 2.5Hz, H5a), 4.79 (1H, s, H1), 7.25-7.29 (3H, m, ArH_{ortho}, ArH_{para}), 7.42 (2H, t, $J_{ortho-meta}$ 8.2Hz, $J_{meta-para}$ 8.2Hz, ArH_{meta}); ^{13}C NMR (126MHz, MeOD) δ ppm 55.4 (1C, C-O<u>C</u>H$_3$), 55.9 (1C, d, J_{C-31P} 4.8Hz, P-OMe), 70.6 (1C, d, J_{C-31P} 6.7Hz, C5), 72.2 (1C, C3), 76.0 (1C, C2), 82.2 (1C, d, J_{C-31P} 7.6Hz, C4), 110.0 (1C, C1), 121.1 (1C, d, J_{C-31P} 4.8Hz, ArC_{ortho}), 126.6 (1C, s, ArC_{para}), 131.0 (1C, s, ArC_{meta}), 151.9 (1C, d, J_{C-31P} 7.6Hz, ArC_{ipso}).



β-methyl-5-(dimethoxyphosphoryl)-ribofuranoside (3.59)

R_f 0.26 (95 ethyl acetate : 5 methanol), $[\alpha]_D^{25}$ -22.5 (c = 2.0, MeOH); ^{1}H NMR (500MHz, MeOD) δ ppm 3.37 (3H, s, C-OMe), 3.82 (3H, d, J_{H-31P} 11.0Hz, P-OMe), 3.83 (3H, d, J_{H-31P} 11.0Hz, P-OMe), 3.90 (1H, d, J_{H2-H3} 4.4Hz, H2), 4.05-4.11 (2H, m, H4, H5'), 4.14 (1H, dd, J_{H3-H4} 6.9Hz, J_{H2-H3} 4.4Hz, H3), 4.25 (1H, ddd, $J_{H5-H5'}$ 14.1Hz, J_{H5-31P} 6.5Hz, J_{H4-H5} 5.7Hz, H5), 4.78 (1H, s, H1); ^{13}C NMR (126MHz, MeOD) δ ppm 55.3 (1C, d, J_{C-31P} 5.7Hz, P-OMe), 55.3 (1C, d, J_{C-31P} 5.7Hz, P-OMe), 55.4 (1C, C-OMe), 70.0 (1C, d, J_{C-31P} 5.7Hz, C5), 72.3 (1C, C3),

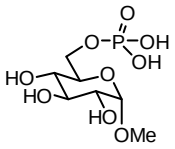
76.0 (1C, C2), 82.3 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, C4), 110.0 (1C, C1); $^{31}P\{^{1}H\}$ NMR (162MHz, MeOD) δ ppm -12.0 (1P, $P(O)(OPh)_2$); HRMS m/z (ES^+) 295.0552 $[M + Na]^+$ (required 295.0553).

3.9.9 Dimethylphosphate deprotections

General reaction conditions

To a solution of carbohydrate (70-200mg, 1eq) in anhydrous dioxane (3ml) was added TMSI (10eq) at room temperature. The reaction was monitored by TLC (5 ethanol : 3 NH_4OH : 1 water) and electrospray mass spectrometry in the negative mode. After 15 minutes (5 for the reaction with **3.59**) complete consumption of starting material and the formation of the desired product was detected. The reaction was quenched with water (1ml) and the solvent removed *in vacuo*. The crude product was partitioned between water (3ml) and ethyl acetate (3ml). The aqueous layer was retained and washed with ethyl acetate (6 x 2ml). The aqueous layer was concentrated under reduced pressure and purified using HPLC on an Applied Biosystems Poros® HQ strongly basic anion exchange column (10mm x 100mm, 50µm). A gradient from 0mM to 500mM aqueous NH_4HCO_3 was used as the mobile phase and eluants were detected with an ELS detector parallel to the main flow path. Fractions containing the product were pooled and lyophilized repeatedly to remove residual NH_4HCO_3. The purified product was dissolved in water (0.5ml) and passed through a pre-equilibrated Waters C18 SepPak® Light cartridge and the eluant collected. The cartridge washed with water (3.5ml) and the eluant combined with the previous collection. Lyophilization yielded the desired compound as a white amorphous solid.

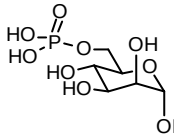
For Tre-6-P see section 2.8.4, page 59.



α-methyl-6-*O*-(dihydroxyphosphoryl)-D-glucopyranoside[64] (3.61)

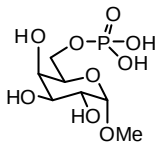
R_f 0.31 (5 ethanol : 3 NH_4OH : 1 water), $[\alpha]_D^{25}$ +89.9 (c = 1.0, H_2O), lit.[65] (Ba^{2+} salt) $[\alpha]_D^{16}$ + 95 (c = 1.0, H_20); ^1H NMR (500MHz, D_2O) δ ppm 3.30 (3H, s, OMe), 3.36 (1H, t, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, H4), 3.45 (1H, dd, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.55 (1H, t, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H3\text{-}H4}$ 9.5Hz, H3), 3.65 (1H, ad, $J_{H4\text{-}H5}$ 9.8Hz, H5), 4.02 (1H, ddd, $J_{H6\text{-}H6'}$ 11.6Hz, $J_{H6'\text{-}31P}$ 6.6Hz, $J_{H5\text{-}H6'}$ 4.7Hz, H6'), 4.05 (1H, ddd, $J_{H6\text{-}H6'}$ 11.6Hz, $J_{H6\text{-}31P}$ 5.7Hz, $J_{H5\text{-}H6}$ 2.3Hz, H6), 4.68 (1H, d, $J_{H1\text{-}H2}$ 3.5Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 55.1 (1C, OMe), 64.4 (1C, d, $J_{C\text{-}31P}$

4.8Hz, C6), 68.9 (1C, C4), 70.3 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C5), 71.1 (1C, C2), 72.9 (1C, C2), 99.3 (1C, C3); $^{31}P\{^{1}H\}$ NMR (162MHz, D_2O) δ ppm 0.4 (1P, $P(O)(OH)_2$); m/s (ES^-) 273.1 $[M - H]^-$.



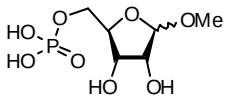
α-methyl-6-*O*-(dihydroxyphosphoryl)-D-mannopyranoside[66] (3.62)

R_f 0.36 (5 ethanol : 3 NH_4OH : 1 water), $[\alpha]_D^{25}$ +47.0 (c = 1.0, H_2O); lit.[66] (Na^+ form) $[\alpha]_D$ +35 (c = 1.0, H_20); ^{1}H NMR (500MHz, D_2O) δ ppm 3.26 (3H, s, OMe), 3.57 (1H, t, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, H4), 3.58 (1H, m, H5), 3.62 (1H, dd, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H2\text{-}H3}$ 2.8Hz, H3), 3.79 (1H, dd, $J_{H2\text{-}H3}$ 2.8Hz, $J_{H1\text{-}H2}$ 1.6Hz, H2), 4.01 (1H, ddd, $J_{H6\text{-}H6'}$ 11.6Hz, $J_{H6'\text{-}31P}$ 6.9Hz, $J_{H5\text{-}H6'}$ 4.7Hz, H6'), 4.07 (1H, ddd, $J_{H6\text{-}H6'}$ 11.5Hz, $J_{H6\text{-}31P}$ 5.7Hz, $J_{H65\text{-}H6}$ 1.1Hz, H6), 4.61 (1H, d, $J_{H1\text{-}H2}$ 0.8Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 54.7 (1C, OMe), 64.8 (1C, d, $J_{C\text{-}31P}$ 4.8Hz, C6), 66.1 (1C, C4), 69.7 (1C, C2), 70.3 (1C, C3), 71.1 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C5), 100.9 (1C, C1); $^{31}P\{^{1}H\}$ NMR (162MHz, D_2O) δ ppm 0.2 (1P, $P(O)(OH)_2$); m/z (ES^-) 273.0 $[M - H]^-$.



α-methyl-6-*O*-(dihydroxyphosphoryl)-D-galactopyranoside (3.63)

R_f 0.29 (5 ethanol : 3 NH_4OH : 1 water), $[\alpha]_D^{25}$ +92.7 (c = 1.39, H_2O); ^{1}H NMR (500MHz, D_2O) δ ppm 3.32 (3H, s, OMe), 3.72-3.73 (2H, m, *J* 2.2Hz, H2, H3), 3.85 (1H, add, $J_{H6\text{-}H6'}$ 10.4Hz, $J_{H6'\text{-}31P}$ 7.6Hz, H6'), 3.88 (1H, dd, *J* $_{H6\text{-}H6'}$12.4Hz, $J_{H5\text{-}H6}$ 6.9Hz, H6), 3.91-3.97 (2H, m, H4, H5), 4.74 (1H, d, $J_{H1\text{-}H2}$ 1.9Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 55.1 (1C, OMe), 63.9 (1C, d, $J_{C\text{-}31P}$ 4.8Hz, C6), 68.1 (1C, C2), 68.8 (1C, C4), 69.2 (1C, C3), 69.4 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C5), 81.7 (1C, C1); $^{31}P\{^{1}H\}$ NMR (162MHz, D_2O) δ ppm 0.9 (1P, $P(O)(OH)_2$); HRMS (ES^-) m/z 273.0387 $[M - H]^-$ (required 273.0381).



α/β-methyl-5-(dihydroxyphosphoryl)-ribofuranoside (3.64)

R_f 0.31 (5 ethanol : 3 NH_4OH : 1 water), ^{31}P NMR (162MHz, D_2O) δ ppm 0.1 ($P(O)OR(OH)_2$); FT-IR (KBr disc) υ 1198 (P=O), 3102 br (OH); HRMS (ES^-) m/z 243.0280 $[M - H]^-$ (required 243.0275).

Alpha anomer

^{1}H NMR (500MHz, D_2O) δ ppm 3.23 (3H, s, OMe), 3.89 (1H, m, H5b), 3.90 (1H, dd, $J_{H2\text{-}H3}$ 4.6Hz, $J_{H1\text{-}H2}$ 2.0Hz, H2), 3.96-4.00 (2H, m, H4, H5a), 4.16 (1H, t, $J_{H2\text{-}H3}$ 5.4Hz, $J_{H3\text{-}H4}$ 5.4Hz, H3), 5.13 (1H, d, $J_{H1\text{-}H2}$ 1.6Hz, H1); ^{13}C NMR (126MHz, D_2O) δ

ppm 48.8 (1C, OMe), 66.1 (1C, d, $J_{C\text{-}31P}$ 4.8Hz, C5), 70.3 (1C, C3), 75.0 (1C, C2), 80.9 (1C, d, $J_{C\text{-}31P}$ 8.6Hz, C4), 101.0 (1C, C1).

Beta anomer

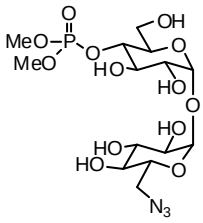
^{1}H NMR (500MHz, D_2O) δ ppm 3.23 (3H, s, OMe), 3.86-3.90 (2H, m, H5a, H5b), 4.04 (1H, t, $J_{H1\text{-}H2}$ 5.4Hz, $J_{H2\text{-}H3}$ 5.4Hz, H2), 4.06 (1H, t, $J_{H2\text{-}H3}$ 5.7Hz, $J_{H3\text{-}H4}$ 5.7Hz, H3), 4.13 (1H, m, H4), 5.28 (1H, d, $J_{H1\text{-}H2}$ 4.1Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 48.8 (1C, OMe), 65.2 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, C5), 69.9 (1C, C3), 70.7 (1C, C2), 81.7 (1C, d, $J_{C\text{-}31P}$ 8.6Hz, C4), 96.3 (1C, C1).

3.9.10 Synthetic rout for 6' modifications

4-*O*-(dimethoxyphosphoryl)-6'azido-D-trehalose (3.65)

6-*O*-(dimethoxyphosphoryl)-6'-azido-D-trehalose (3.66)

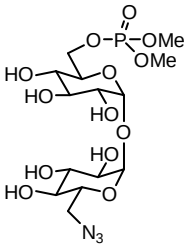
A solution of **3.17** (100mg, 0.12mmol, 1eq) and CsF (180mg, 1.2mmol, 10eq) in methanol (5ml) was heated to 80°C with vigorous stirring for 5 hours at which point TLC (1 water : 4 isopropanol : 4 ethyl acetate) showed the complete consumption of starting material. The solvent was removed *in vacuo* and the products were purified by silica gel chromatography (1 water : 3 isopropanol : 10 ethyl acetate). Lyophilisation yielded the desired compound **3.65** (11.8mg, 21%) and **3.66** (37mg, 67%) as white amorphous solids.



*Compound **3.65***

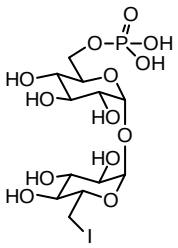
R_f 0.6 (1 water : 4 isopropanol : 4 ethyl acetate), $[\alpha]_D^{18}$ +68.0 (c = 0.49, MeOH); ^{1}H NMR (500MHz, MeOD) δ ppm 3.31 (1H, t, $J_{H3'\text{-}H4'}$ 9.8Hz, $J_{H4'\text{-}H5'}$ 9.8Hz, H4'), 3.43 (1H, dd, $J_{H6'a\text{-}H6'b}$ 14.1Hz, $J_{H5'\text{-}H6'b}$ 5.7Hz, H6'b), 3.51 (1H, dd, $J_{H2'\text{-}H3'}$ 9.5Hz, $J_{H1'\text{-}H2'}$ 5.4Hz, H2'), 3.55 (1H, dd, $J_{H6'a\text{-}H6'b}$ 13.0Hz, $J_{H5'\text{-}H6'a}$ 4.5Hz, H6'a), 3.60 (1H, td, $J_{H4\text{-}H5}$ 10.1Hz, $J_{H5\text{-}H6a}$ 10.1Hz, $J_{H5\text{-}H6b}$ 4.7Hz, H5), 3.71 (1H, dd, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.74 (1H, dd, $J_{H6a\text{-}H6b}$ 12.1Hz, $J_{H5\text{-}H6b}$ 4.3Hz, H6b), 3.82 (1H, t, $J_{H2'\text{-}H3'}$ 9.1Hz, $J_{H1'\text{-}H2'}$ 9.1Hz, H3'), 3.84 (6H, d, $J_{H\text{-}31P}$ 11.3Hz, 2 x OMe), 3.85-3.89 (2H, m, H3, H6a), 4.07 (1H, ddd, $J_{H4\text{-}H5}$ 10.0Hz, $J_{H5\text{-}H6b}$ 5.8Hz, $J_{H5\text{-}H6a}$ 2.0Hz, H5), 4.59 (1H, q, $J_{H3\text{-}H4}$ 8.5Hz, $J_{H4\text{-}H5}$ 8.5Hz, $J_{H4\text{-}31P}$ 8.5Hz, H4), 5.15 (1H, d, $J_{H1'\text{-}H2'}$ 3.5Hz, H1'), 5.18 (1H, d, $J_{H1\text{-}H2}$ 3.5Hz, H1); ^{13}C NMR (126MHz, MeOD) δ ppm 52.8 (1C, C6'), 55.3 (2C, d, $J_{C\text{-}31P}$ 5.7Hz, 2 x OMe), 62.2 (1C, C6), 70.3 (1C, d, $J_{C\text{-}31P}$ 2.9Hz, C5), 71.9 (1C, d, $J_{C\text{-}31P}$ 2.9Hz, C2), 72.8 (1C, C4'), 73.1 (1C, C5'), 73.2 (1C, C2'), 73.9 (1C, d , $J_{C\text{-}31P}$ 2.9Hz, C3), 74.3 (1C, C3'),

83.4 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, C4), 95.2 (1C, C1'), 95.5 (1C, C1); FT-IR (KBr disc) υ 1261 (P=O), 2105 (N_3), 3424 br (OH); HRMS (ES^+) 498.1095 $[M + Na]^+$ (required 498.1095).



*Compound **3.66***

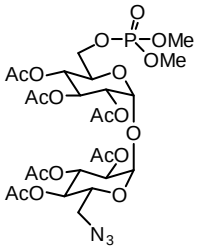
R_f 0.4 (1 water : 4 isopropanol : 4 ethyl acetate), $[\alpha]_D^{18}$ + 68.4 (c = 1.0, MeOH); ^1H NMR (500MHz, D_2O) δ ppm 3.38 (1H, t, $J_{H3'\text{-}H4'}$ 9.5Hz, $J_{H4'\text{-}H5'}$ 9.5Hz, H4'), 3.44 (1H, t, $J_{H3\text{-}H4}$ 9.8Hz, $J_{H4\text{-}H5}$ 9.8Hz, H4), 3.47 (2H, dd, $J_{H6'a\text{-}H6'b}$13.4Hz, $J_{H5'\text{-}H6'b}$ 5.5Hz, H6'b), 3.56 (1H, m, H6'a), 3.58 (2H, dd, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2, H2'), 3.75 (1H, m, H3'), 3.76 (3H, d, $J_{H\text{-}31P}$ 11.0Hz, OMe), 3.76 (3H, d, $J_{H\text{-}31P}$ 11.0Hz, OMe), 3.78 (1H, t, $J_{H2\text{-}H3}$ 9.1Hz, $J_{H3\text{-}H4}$ 9.1Hz, H3), 3.90 (1H, ddd, $J_{H4'\text{-}H5'}$ 12.6Hz, $J_{H5'\text{-}H6'b}$ 6.1Hz, $J_{H5'\text{-}H6'a}$ 2.4Hz, H5'), 3.92 (1H, dd, $J_{H4\text{-}H5}$ 10.7Hz, $J_{H5\text{-}H6a}$ 1.9Hz, H5), 4.27 (2H, m, H6a, H6b), 5.11 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.13 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 50.8 (1C, C6'), 55.2 (2C, d, $J_{C\text{-}31P}$ 7.55Hz, 2 x OMe), 66.8 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, C6), 69.1 (1C, C4'), 70.4 (1C, C4), 70.5 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, C5), 70.8 (1C, C5), 70.9 (1C, C2), 71.1 (1C, C2'), 72.3 (1C, C3), 72.4 (1C, C3'), 93.7 (1C, C1), 93.8 (1C, C1'); ^1H NMR (500MHz, MeOD) δ ppm 3.33 (1H, t, $J_{H3\text{-}H4}$ 9.3Hz, $J_{H4\text{-}H5}$ 9.3Hz, H4), 3.38 (1H, t, $J_{H3'\text{-}H4'}$ 10.1Hz, $J_{H4'\text{-}H5'}$ 10.1Hz, H4'), 3.42 (1H, dd, $J_{H6'a\text{-}H6'b}$ 14.0Hz, $J_{H6'b\text{-}31P}$ 6.1Hz, H6'b), 3.52 (2H, dd, $J_{H2\text{-}H3}$ 9.1Hz, $J_{H1\text{-}H2}$ 3.1Hz, H2, H2'), 3.50 (1H, m, H6'a), 3.79 (1H, t, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H3\text{-}H4}$ 9.5Hz, H3), 3.81 (3H, d, $J_{H\text{-}31P}$ 11.0Hz, OMe), 3.81 (3H, d, $J_{H\text{-}31P}$ 11.0Hz, OMe), 3.83 (1H, t, $J_{H2'\text{-}H3'}$ 10.4Hz, $J_{H3'\text{-}H4'}$ 10.4Hz, H3'), 4.04 (2H, dt, $J_{H4'\text{-}H5'}$ 9.7Hz, $J_{H5'\text{-}H6'a}$ 9.7Hz, $J_{H5'\text{-}H6'b}$ 2.6Hz, H5, H5'), 4.25 (1H, ddd, $J_{H6a\text{-}H6b}$ 11.1Hz, $J_{H6b\text{-}31P}$ 6.4Hz, $J_{H5\text{-}H6b}$ 5.0Hz, H6b), 4.31 (1H, ddd, $J_{H6a\text{-}H6b}$ 11.2Hz, $J_{H6a\text{-}31P}$ 6.4Hz, $J_{H5\text{-}H6a}$ 1.6Hz, H6a), 5.12 (2H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.14 (2H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (126MHz, MeOD) δ ppm 52.7 (1C, C6'), 55.3 (2C, d, $J_{C\text{-}31P}$ 6.7Hz, 2 x OMe), 68.4 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, C6), 71.3 (1C, C4'), 72.2 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, C5), 72.6 (1C, C4), 73.1 (1C, C5'), 73.1 (1C, C2'), 73.1 (1C, C2'), 74.4 (1C, C3), 74.5 (1C, C3'), 95.4 (1C, C1), 95.6 (1C, C1'); $^{31}P\{^1H\}$ (162MHz, MeOD) 1.00 (1P, $ROP(O)(OMe)_2$); FT-IR (KBr disc) υ 1261 (P=O), 2105 (N_3), 3424 br (OH); HRMS (ES^+) 498.1095 $[M + Na]^+$ (required 498.1095).



D-trehalose-6-phospate-6'-iodide (3.67)

Compound **3.66** (15mg, 0.032mmol, 1eq) was suspended in dry dioxane (2ml) and briefly sonicated for 5 minutes. The mixture was cooled to 0°C and TMSI (10µl, 0.038mmol, 2.2eq) was added dropwise. The reaction was monitored by mass spectrometry (ES^-) and after 10 minutes showed the complete conversion of starting material to the title compound. The reaction mixture was quenched with water (1ml) and the solvents removed under reduced pressure. The crude product was taken up in water (2ml) and washed with ethyl acetate (3 x 1ml). The aqueous layer was concentrated *in vacuo* and purified by HPLC on an Applied Biosystems Poros® HQ 50µm strongly basic anion exchange column (10mm x 100mm, 7.9ml). An applied gradient from 0mM to 500mM aqueous NH_4HCO_3 was used as the mobile phase and eluants were detected with an ELS detector parallel to the main flow path. Fractions containing the product were pooled and repeated lyophilization to remove residual NH_4HCO_3 afforded the title compound as a white amorphous solid (12.4mg, 69%).

R_f 0.45 (5 ethanol : 3 NH_4OH : 1 water), $[\alpha]_D^{25}$ + 10.3 (c = 0.59, H_2O); ^1H NMR (500MHz, D_2O) δ ppm 3.29 (1H, dd, $J_{H6'a-H6'b}$ 12.6Hz, $J_{H5'-H6'b}$ 7.3Hz, H6'b), 3.27 (1H, t, $J_{H3'-H4'}$ 9.1Hz, $J_{H4'-H5'}$ 9.1Hz, H4'), 3.48-3.56 (2H, m, H4, H6'a), 3.53 (2H, dd, J_{H2-H3} 9.3Hz, J_{H1-H2} 2.7Hz, H2, H2'), 3.60 (1H, ddd, J_{H4-H5} 9.8Hz, J_{H5-H6a} 5.7Hz, J_{H5-H6b} 4.1Hz, H5), 3.73-3.81 (1H, m, H5') 3.76 (1H, t, J_{H2-H3} 9.5Hz, J_{H3-H4} 9.5Hz, H3), 3.78 (1H, t, $J_{H2'-H3'}$ 9.5Hz, $J_{H3'-H4'}$ 9.5Hz, H3'), 3.86 (1H, ddd, $J_{H6a-H6b}$ 12.4Hz, $J_{H6b-31P}$ 5.4Hz , J_{H5-H6b} 1.9Hz, H6b), 3.94 (1H, ddd, $J_{H6a-H6b}$ 12.2Hz, $J_{H6a-31P}$ 8.1Hz, J_{H5-H6a} 3.8Hz, H6a), 5.11 (1H, d, $J_{H1'-H2'}$ 4.1Hz, H1'), 5.16 (1H, d, J_{H1-H2} 3.8Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 6.5 (1C, C6'), 62.4 (1C, d, J_{C-31P} 3.8Hz, C6), 69.0 (1C, C2'), 70.6 (1C, C2), 71.0 (1C, C4), 71.1 (1C, C5'), 71.8 (1C, C3'), 71.9 (1C, d, J_{C-31P} 5.7Hz, C5), 72.1 (1C, C3), 73.6 (1C, C4'), 93.3 (1C, C1'), 93.6 (1C, C1); FT-IR (KBr disc) υ 1103 (P=O), 3431 br (OH); HRMS (ES^-) m/z 530.9769 $[M - H]^-$ (required 530.9770).

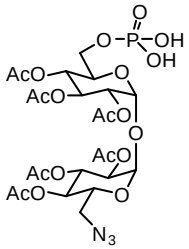


6-*O*-(dimethoxyphosphoryl)-6'azido-2,2',3,3',4,4'-acyl-D-trehalose (3.68)

To a solution of **3.67** (32mg, 0.063mmol, 1eq) in pyridine (3ml) at room temperature was added acetic anhydride (72µl, 0.76mmol, 12eq). TLC (ethyl acetate) after 16 hours showed complete

consumption of starting material. The reaction was quenched with water (2ml) and partitioned between ethyl acetate (30ml) and water (30ml). The aqueous layer was extracted with ethyl acetate (2 x 15ml) and the combined organics washed with 1M HCl (5 x 20ml), saturated $NaHCO_3$ solution (25ml), brine (25ml), dried over $MgSO_4$ and concentrated *in vacuo*. The product was purified by silica gel chromatography (9 ethyl acetate : 1 petrol) to give **3.68** (43mg, 93%).

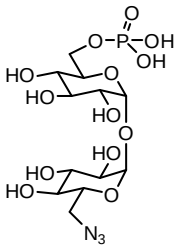
R_f 0.53 (ethyl acetate), $[\alpha]_D^{23}$ +119.9 (c = 1.0, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.02, 2.03, 2.06, 2.07, 2.08, 2.12 (6 x 3H, 6 x s, 6 x OAc), 3.16 (1H, dd, $J_{H6'a\text{-}H6'b}$ 13.2Hz, $J_{H5'\text{-}H6'b}$ 2.5Hz, H6'b), 3.36 (1H, dd, $J_{H6'a\text{-}H6'b}$ 13.4Hz, $J_{H5'\text{-}H6a'}$ 7.4Hz, H6'a), 3.76 (3H, d, $J_{H\text{-}31P}$ 11.4Hz, OMe), 3.77 (3H, d, $J_{H\text{-}31P}$ 11.4Hz, OMe), 4.03 (1H, ddd, $J_{H6a\text{-}H6b}$ 14.2Hz, $J_{H5\text{-}H6b}$ 5.0Hz, $J_{H\text{-}31P}$ 5.0Hz, H6b), 4.06-4.16 (2H, m, H5, H5'), 4.12 (1H, ddd, $J_{H6a\text{-}H6b}$ 14.3Hz, $J_{H\text{-}31P}$ 7.3Hz, $J_{H5\text{-}H6a}$ 2.2Hz, H6a), 4.99 (1H, t, $J_{H3'\text{-}H4'}$ 9.8Hz, $J_{H4'\text{-}H5'}$ 9.8Hz, H4'), 5.06 (1H, dd, $J_{H2'\text{-}H3'}$ 10.7Hz, $J_{H1'\text{-}H2'}$ 3.8Hz, H2'), 5.09 (1H, t, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, H4), 5.08 (1H, dd, $J_{H2\text{-}H3}$ 10.5Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 5.30 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.32 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H2), 5.46 (1H, t, $J_{H2'\text{-}H3'}$ 9.9Hz, $J_{H3'\text{-}H4'}$ 9.9Hz, H3'), 5.49 (1H, t, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H3\text{-}H4}$ 10.1Hz, H3); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.5, 20.6, 20.6, 20.7 (6C, 4 x as, $CO\underline{C}H_3$), 50.9 (1C, C6'), 54.6 (2C, d, $J_{C\text{-}31P}$ 5.7Hz, 2 x OMe), 65.1 (1C, d, $J_{C\text{-}31P}$ 4.8Hz, C6), 68.2 (1C, C4), 68.8 (1C, d, $J_{C\text{-}31P}$ 8.6Hz, C5), 69.6 (2C, C2, C2'), 69.6 (1C, C4'), 69.8 (1C, C3'), 69.9 (1C, C5'), 70.0 (1C, C5), 92.8 (1C, C1'), 93.0 (1C, C1), 169.5, 169.5, 169.6, 169.7, 169.9, 170.0 (6 x 1C, 6 x C=O); FT-IR (thin film) υ 1219 (P=O), 1753 (C=O), 2106 (N_3); HRMS (ES^+) m/z 750.1721 $[M + Na]^+$ (required 750.1729).



6-*O*-(dihydroxyphosphoryl)-6'azido-2,2',3,3',4,4'-acyl-D-trehalose (3.69)

To a solution of **3.70** (15mg, 0.021mmol, 1eq) in dry dioxane (2ml) was added TMSBr (27µl, 0.21mmol, 10eq). TLC (1 water : 4 isopropanol : 4 ethyl acetate) after 14 hours showed complete conversion of the starting material to the desired product. The reaction was quenched with water (1ml) and the solvents were removed under reduced pressure. The crude product was taken up in water (5ml) and washed with ethyl acetate (3 x 20ml). Lyophilization of the aqueous layer gave the desired product as a white amorphous solid (14.2mg, 100%).

R_f 0.13 (1 water : 4 isopropanol : 4 ethyl acetate), $[\alpha]_D^{25}$ +44.9 (c = 1.0, MeOH); ^{1}H NMR (500MHz, D_2O) δ ppm 2.05, 2.05, 2.07, 2.08, 2.11, 2.12 (6 x 3H, 6 x s, 6 x OAc), 3.51 (2H, m, H6'), 3.89 (2H, m, H6), 4.13 (1H, ddd, $J_{H4'\text{-}H5'}$ 10.1Hz, $J_{H5'\text{-}H6'a}$ 4.4Hz, $J_{H5'\text{-}H6'b}$ 1.3Hz, H5'), 4.16 (1H, dt, $J_{H4\text{-}H5}$ 9.9Hz, $J_{H5\text{-}H6a}$ 3.2Hz, $J_{H5\text{-}H6b}$ 3.2Hz, H5), 5.12 (1H, t, $J_{H3'\text{-}H4'}$ 9.8Hz, $J_{H4'\text{-}H5'}$ 9.8Hz, H4'), 5.17 (1H, dd, $J_{H2'\text{-}H3'}$ 9.3Hz, $J_{H1'\text{-}H2'}$ 2.5Hz, H2'), 5.17 (1H, t, $J_{H3\text{-}H4}$ 9.8Hz, $J_{H4\text{-}H5}$ 9.8Hz, H4), 5.18 (1H, dd, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 5.42 (1H, d, $J_{H1'\text{-}H2'}$ 3.5Hz, H1'), 5.45 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 5.52 (1H, t, $J_{H2'\text{-}H3'}$ 9.8Hz, $J_{H3'\text{-}H4'}$ 9.8Hz, H3'), 5.55 (1H, t, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H3\text{-}H4}$ 9.8Hz, H3); ^{13}C NMR (126MHz, D_2O) δ ppm 20.2, 20.2, 20.2, 20.2, 20.2, 20.3 (6C, 6 x CO<u>C</u>H$_3$), 50.1 (1C, C6'), 62.5 (1C, d, $J_{C\text{-}31P}$ 3.8Hz, C6), 68.4 (1C, C4), 69.2 (1C, C4'), 69.3 (1C, C5'), 69.3 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C5), 69.9 (1C, C2'), 70.1 (1C, C2), 70.7 (1C, C3'), 71.0 (1C, C3), 92.3 (1C, C1'), 92.4 (1C, C1), 172.6, 172.6, 172.7, 172.8, 173.3, 173.4 (6 x 1C, 6 x C=O); ^{31}P{^{1}H} (162MHz, D_2O) 4.71 (1P, ROP(O)(OH)$_2$); FT-IR (KBr disc) υ 1203 (P=O), 1767 (C=O), 2106 (N_3), 3106 br (OH); HRMS (ES^-) m/z 698.1467 $[M - H]^-$ (required 698.1451).



6'-azido-trehalose-6-phosphate (3.70)

Compound **3.69** (14mg, 0.02mmol, 1eq) and sodium methoxide (0.5mg, 0.002mmol, 0.5eq) were dissolved in methanol (5ml) and stirred at room temperature. After 14 hours, TLC (1 water : 2 isopropanol : 2 ethyl acetate) showed the complete consumption of starting material and the formation of a single product. The reaction mixture was opened to the atmosphere and quenched by the addition of dry ice until neutral (pH paper). The solvent was removed *in vacuo* and purified by HPLC on an Applied Biosystems, Poros® HQ 50µm strongly basic anion exchange column (10mm x 100mm, 7.9ml). A gradient from 0mM to 500mM aqueous NH_4HCO_3 was used as the mobile phase and eluants were detected with an in-line Evaporative Light Scattering detector. Fractions containing the product were pooled and repeated lyophilization to remove NH_4HCO_3 gave the target compound as a white amorphous solid (7.8mg, 87%).

R_f 0.35 (5 ethanol : 3 NH_4OH : 1 water), $[\alpha]_D^{25}$ +36.2 (c = 0.39, H_2O); ^{1}H NMR (500MHz, D_2O) δ ppm 3.37 (1H, t, $J_{H3'\text{-}H4'}$ 9.5Hz, $J_{H4'\text{-}H5'}$ 9.5Hz, H4'), 3.47 (1H, dd, $J_{H6'a\text{-}H6'b}$ 13.6Hz, $J_{H5'\text{-}H6'b}$ 5.7Hz, H6'b), 3.54 (1H, t, $J_{H3\text{-}H4}$ 9.6Hz, $J_{H4\text{-}H5}$ 9.6Hz, H4), 3.57-3.64 (2H, m, H2, H2'), 3.59 (1H, dd, $J_{H6'a\text{-}H6'b}$ 13.9Hz, $J_{H5'\text{-}H6'a}$ 3.5Hz, H6'a), 3.74

(1H, t, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H3\text{-}H4}$ 9.8Hz, H3), 3.76 (1H, t, $J_{H2'\text{-}H3'}$ 9.8Hz, $J_{H3'\text{-}H4'}$ 9.8Hz, H3'), 3.81 (1H, add, $J_{H4\text{-}H5}$ 9.5Hz, $J_{H5\text{-}H6b}$ 1.6Hz, H5), 3.85 (1H, ddd, $J_{H6a\text{-}H6b}$ 10.4Hz, $J_{H6b\text{-}31P}$ 5.2Hz, $J_{H5\text{-}H6b}$ 1.7Hz, H6b), 3.88 (1H, ddd, $J_{H4'\text{-}H5'}$ 10.2Hz, $J_{H5'\text{-}H6'b}$ 5.8Hz, $J_{H5'\text{-}H6'a}$ 2.5Hz, H5'), 3.94 (1H, ddd, $J_{H6a\text{-}H6b}$ 12.0Hz, $J_{H6a\text{-}31P}$ 7.3Hz, $J_{H5\text{-}H6a}$ 4.1Hz, H6a), 5.09 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.13 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 50.8 (1C, C6'), 62.4 (1C, d, $J_{C\text{-}31P}$ 4.8Hz, C6), 69.0 (1C, C4), 70.4 (1C, C4'), 70.8 (1C, C2), 70.9 (1C, C2'), 71.1 (1C, C5'), 71.9 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C5), 72.1 (1C, C3'), 72.2 (1C, C3), 93.5 (1C, C1'), 93.8 (1C, C1); $^{31}P\{^1H\}$ NMR (162MHz, D_2O) δ ppm 9.92 (1P, $ROP(O)(OH)_2$); FT-IR (KBr disc) υ 1262 (P=O), 2111 (N_3), 3431 br (OH); HRMS m/z (ES^-) 446.0821 $[M - H]^-$ (required 446.0817).

3.10 References

Figures in this chapter were reproduced by permission from The American Chemical Society from: M. K. Patel, B. G. Davis, *Org. Lett.*, **2013**, 15, 346.

1. (a) D. C. Billington, *Chem. Soc. Rev.*, **1989**, 18, 83. (b) D. C. Billington, *The Inositol Phosphates: Chemical Synthesis and Biological Significance*, **1993**, VCH.
2. K. M. Sureshan, M. S. Shashidhar, T. Praveen, T. Das, *Chem. Rev.*, **2003**, 103, 4477.
3. H. K. Chenault, R. F. Mandes, K. R. Hornberger, *J. Org. Chem.*, **1997**, 62, 331.
4. Y. Ahmadibeni, K. Parang, *Org. Lett.*, **2005**, 7, 1955.
5. D. Ruhela, R. A. Vishwakarma, *J. Org. Chem.*, **2003**, 68, 4446.
6. J. Feurle, H. Jomaa, M. Wilhelm, B. Gutsche, M. Herderich, *J. Chromatgr. A.*, **1998**, 803, 111.
7. T. W. Greene, P. G.M. Wuts, *Protective Groups in Organic Synthesis, Third Edition,* **1999**, John Wiley & Sons.
8. (a) V. N. Shibaev, L. L. Danilov, n *Glycopeptides and Related* Compounds, **1997**, Marcel Dekker. (b) G. Boons, B. Heskamp, *Carbohydrate Chemistry*, **1998**, Thomson Science.
9. S. Sabesan, S. Neira, *Carbohydr. Res.*, **1992**, 223, 169.

10. (a) L. A. Slotin, *Synthesis*. **1977**, 737. (b) C. B. Reese, *Tetrahedron*, **1978**, 34, 3143. (c) R. S. Edmundson, *Comprehensive Organic Chemistry*, **1979**, Pergamon Press Ltd.
11. J. Stawinsky, A. Kraszewski, *Acc. Chem. Res.*, **2002**, 35, 952.
12. (a) S. Jones, D. Selitsianos, K. J. Thomson, S. M. Toms, *J. Org. Chem.*, **2003**, 68, 5211. (b) S. Jones, D. Selitsianos, *Org. Lett.*, **2002**, 4, 3671.
13. S. Jones, C. Smanmoo, *Org. Lett.,* **2005**, 7, 15, 3271.
14. B. S. Sproat. M. J. Gait, *Oligonucleotide Synthesis*, **1984**, IRL Press.
15. (a) S. L. Beaucage, R. P. Iyer, *Tetrahedron*, **1992**, 48, 2223. (b) S. L. Beaucage, R. P. Iyer, *Tetrahedron*, **1993**, 49, 1925. (c) E. E. Nifantiev, M. K. Grachev, S. Y. Burmistrov, *Chem. Rev.*, **2000**, 100, 3755.
16. K. Parang, *Bio. Med. Chem. Lett.*, **2002**, 1863.
17. (a) A. Kraszewski, J. Stawinsk, *Pure Appl. Chem.*, **2007**, 79, 2217. (b) D. V. Yashunsky, A. V. Nikolaev, *J. Chem. Soc., Perkin Trans. 1*, **2000**, 1195.
18. (a) H. Nakajima, H. Kondo, R. Tsurutani, M. Dombou, I. Tomioka, K. Tomita, *Enzymes in Carbohydrate Synthesis*, **1991**, American Chemical Society. (b) H. K. Chenault, R. F. Mandes, K. R. Hornberger, *J. Org. Chem.*, **1997**, 62, 331.
19. B. R. Sculimbrene, S. J. Miller, *J. Am. Chem. Soc.,* **2001**, 123, 10125.
20. E. C. L. Rønnow, M. Meldal, K. Bock, *Carbohydr. Res.*, **1994**, 260, 323.
21. E. Tarelli, S. F. Wheeler, *Carbohydr. Res.*, **1994**, 261, 25.
22. J. Doi, K. Yokoigawa, Y. Isobe, H. Kawai, *Biosci., Biotech., Biochem.*, **1998**, 62, 735.
23. J. C. Errey, S. S. Lee, R. P. Gibson, C. M. Fleites, C. S. Barry, P. M. J. Jung, A. C. O. Sullivan, B. G. Davis, G. J. Davies, *Angew. Chem. Int. Ed.*, **2010**, 49, 1234.
24. (a) S. A. Fleming, *Tetrahedron*, **1995**, 51, 46, 12479. (b) J. Brunner, *Annu. Rev. Biochem*, **1993**, 62, 483. (c) V. Chowdhry, *Annu. Rev. Biochem.,* **1979**, 48, 293.
25. E. M. Nashed, C. P. J. Glaudemans, *J. Org. Chem*, **1987**, 52, 5255.
26. S. A. Wolckenhauer, S. D. Rychnovsky, *Org. Lett.*, **2004**, 6, 2745.
27. J. P. Larkin, D. C. Nonhebel, H. C. S. Wood, *J. Chem. Soc., Perkin Trans. 1*, **1976**, 2524.
28. A. Yanagisawa, H. Hibino, N. Nomura, H. Yamamoto., *J. Am. Chem. Soc.*, **1993**,115, 5879.

29. S. Araki, K. Minami, Y. Butsugan, *Bull. Chem. Soc. Jpn.*, **1981**, 54, 629.
30. (a) F. Liu, D. J. Austin, *J. Org. Chem.*, **2001**, 66, 8643. (b) F. Liu, D. J. Austin, *Tetrahedron Lett.*, **2001**, 42, 3153. (c) K. Chang, L. Lee, J. Chen, *et al.*, *Chem. Commun.*, **2006**, 629.
31. H. Wessel, T. Iversen, D. R. Bundle, *J. Chem. Soc. Perkin Trans. 1*, **1985**, 2247.
32. (a) R. Plourde, M. d'Alarco, *Tetrahedron Lett.*, **1990**, 31, 2693. (b) G. D. Nanteuil, A. Benoist, G. Remond, *et al.*, *Tetrahedron Lett.*, **1995**, 36, 1435. (c) J. H. Boom, P. M. J. Bergers, P. H. Deursen, *et al.*, *Tetrahedron Lett.*, **1974**, 43, 3785. (d) H. Rokos, A. Myles, W. Hutzenlaub, *et al.*, *Chem. Ber.*, **1975**, 108, 2872. (e) R. W. Adamiak, R. Arentzen, C. B. Reese, *Tetrahedron Lett.*, **1977**, 17, 1431.
33. K. K. Ogilvie, S. L. Beaucage, *Nucl. Acids Res.,* **1979**, 7, 3, 805.
34. (a) W. D. Fordham and J. H. Wang, *J. Am. Chem. Soc.,* **1967**, 89, 4197. (b) S. L. Buchwald, D. H. Pliura, J. R. Knowles, *J. Am. Chem. Soc.,* **1984**, 108, 4916.
35. (a) M. W. Bowler, M. J. Cliff, J. P. Waltho, G. M. Blackburn, *New J. Chem.*, **2010**, 34, 784. (b) F. H. Westheimer, *Science*, **1987**, 235, 1173. (c) G. K. Schroeder, C. Lad, P. Wyman, N. H. Williams, R. Wolfenden, *Proc. Natl. Acad. Sci. U. S. A.*, **2006**, 103, 4052.
36. (a) M. J. Kamlet, R. W. Taft, *J. Am. Hem. Soc.*, **1976**, 98, 377. (b) R. W. Taft, M. J. Kamlet, *J. Am. Hem. Soc.*, **1976**, 98, 2886.
37. (a) F. H. Westheimer, *Acc. Che. Res.,* **1968**, 1, 70. (b) E. A. Dennis, F. H. Westheimer, *J. Am. Chem. Soc.*, **1966**, 88, 3432. (c) P. Gillespie, F. Ramirez, I. Ugi, D. Marquarding, *Angew. Chem. Int. Ed.*, **1973**, 12, 91.
38. C. R. Hall, T. D. Inch, *Tetrahedron*, **1980**, 36, 2059.
39. (a) D. B. Cooper, J. M. Harrison, T. D. Inch, G. J. Lewis, *J. Chem. Soc., Perkin Trans. 1*, **1974**, 1058. (b) P. Szabo, L. Szabo, *J. Chem. Soc.,* **1960**, 3758. (c) T. D. Inch, G. J. Lewis, *J. Chem. Soc., Chem. Commun.*, **1973**, 310.
40. S. Trippett, *Phosphorous Sulfur*, **1976**, 1, 89.
41. (a) S. N. Mikhailov, M. Oivanen, P. Oksman, H. Lonnberg, *J. Org. Chem.,* **1992**, 57, 4122. (b) M. Oivanen, R. Schnell, W. Pfeiderer, H. Lönnberg, *J. Org. Chem.*, **1991**, 56, 3623. (c) P. Järvinen, M. Oivanen, H. Lönnberg, *J. Org. Chem.*, **1991**, 56, 5396.

42. R. Hoffmann, J. M. Howell, E. L. Muetterties, *J. Am. Chem. Soc.*, **1972**, 94, 3047.
43. (a) G. T. Hefter, *Pure & Appl. Chem.*, **1991**, 63, 12, 1749. (b) G. T. Hefter, P. J. McLay, *J. Sol. Chem.*, **1988**, 17, 535.
44. H. S. Frank, W. Y. Wen, *Disc. Faraday Soc.*, **1957**, 24,133.
45. Y. Marcus, G. Hefter, *Chem. Rev.*, **2006**, 106, 4585.
46. G. T. Hefter, M. Salomon, *J. Sol. Chem.*, **1996**, 25, 6, 541.
47. R. Buchner, G. Hefter, *J. Sol. Chem.,* **2002**, 31, 521.
48. R. González, P. Ortiz, E. Reguera, J. F. Bertrán, *J. Fluorine Chem.*, **2001**, 110, 5.
49. C. Hardacre, J. Holbrey, M. Nieuwenhuyzen, *et al.*, *Acc. Chem. Res.,* **2007**, 40, 1146.
50. T. Ishikawa, Y. Oku, K. Kotake, H. Ishii, *J. Org. Chem.*, **1996**, 61, 6484.
51. (a) G. M. Blackburn, D Ingleson, *J. Chem. Soc., Chem. Commun.*, **1978**, 870. (b) G. A. Olah, S. C. Narang, *Tetrahedron*, **1982**, 38, 2225.
52. P. Collins, R. Ferrier, Monosaccharides Their Chemistry and Their Roles in Natural Products, **1995**, *Wiley*.
53. R. C. Chalk, D. H. Ball, L. Long, *J. Org. Chem.,* **1966**, 31, 1509.
54. S. Hanessian, T. Ogawa, Y. Guindon, *Carbhydr. Res.,* **1974**, 48, C15.
55. E.Chelmecka, K. Pasterny, M. Gawlick-Jedrysiak, *et a.l.*, *J. Mol. Stuct.,* **2007**, 834, 498.
56. Baddiley, J, Buchanan, J. G., Szabo, L., *J. Chem. Soc.*, **1954**, 3826.
57. W. R. Kobertz, C. R. Bertozzi, M. D. Bednarski, *J. Org. Chem*, **1996**, 61, 1894.
58. K. Matsuda, T. Tsuchiya, T. Torri, *Bull. Chem. Soc. Jpn.,* **1986**, 59, 1397.
59. G. Birch, A.C. Richardson, *Carbohdr. Res.*, **1968**, 8, 411.
60. E. Chelmecka, K. Pasterny, M. Gawlik-Jedrysiak, *et al.*, *J. Mol. Struct.*, **2007**, 834, 498.
61. K. L. Matta, M. S. Chowdhary, R. K. Jain, *Carbohydr. Res.*, **1986**, 150, C1.
62. E. Chelmecka, K. Pasterny, M. Gawlik-Jedrysiak, W. Szeja, R. Wrzalik, *J. Mol. Struc.*, **2007**, 834, 498.
63. R. Barker, H. G. Fletcher, *J. Org. Chem.*, **1961**, 26, 4605.
64. C. A. Stortz, R. Cherniak, R. G. Jones, *et al.*, *Carbohydr. Res.*, **1990**, 270, 101.

65. E. E. Percival, E. G. V. Percival, *J. Chem. Soc.*, **1945**, 87.
66. C. Bernlind, S. Oscarson, G. Widmalm, *Carbohydr. Res.*, **1994**, 263, 173.

Chapter IV – Investigations on the inhibition of SnRK1

4.1 Introduction

The advent of modern proteomics has revolutionized the biological sciences.[1] The general principle involves the isolation of a protein of interest, usually via gel electrophoresis, and subsequent digestion to peptide fragments. These peptides can then be analysed by high resolution mass spectrometry and MS/MS techniques that produce a fingerprint of the peptide masses as well as basic sequence information on the protein. This data is then compared against genomic databases which contain lists of all putative protein sequences within an organism, potentially revealing the identity of the protein.[2] Crucially, the method is incredibly sensitive requiring only nanomolar concentrations of a protein, since only a small number of unique peptide hits are required for a successful identification.

The inhibition of the plant kinase SnRK1 by Tre-6-P forms a central regulatory network by which energy usage and metabolism is controlled at the whole-organism level (section 1.2, page 2). This inhibition is believed to be mediated by an intermediary protein that can bind and inhibit SnRK1 in the presence of micromolar concentrations of Tre-6-P. This signalling pathway provides an excellent opportunity to alter plant growth for agricultural benefit, so determining the identity of this elusive intermediary factor is essential. Mass spectrometry based proteomics techniques are ideal for this purpose, which simplifies the overall problem to the isolation of the intermediary factor.

The purification of transient protein complexes from a crude mixture of cell lysate can be very difficult.[3] The general strategy involves the selective targeting of a major component of the protein complex for affinity purification, in the hope that other binding partners will also be isolated in tandem. Often, organisms are engineered to express the major component with a high affinity tag that is used for specific purification.[3,4] However, the genetic manipulation that this technique necessitates is difficult to implement and is not a viable option for plants. An alternative *in vitro* technique utilizes the highly specific binding of antibodies for the immunoprecipitation of the desired protein complex, but this methodology is often plagued by various technical difficulties, such as the choice of antibody which can

give highly variable results.[3] Another generic approach involves the expression of a recombinant protein bearing a glutathione *S*-transferase tag, which allows attachment of the protein to a solid phase resin to act as a scavenger.[3] While this method allows isolation of even weak binding partners, it cannot easily be applied to SnRK1 since the surface immobilization of a trimeric protein complex would rely on the high stability of the trimer, which in the case of SnRK1, rules out this technique.

Instead of using SnRK1 as bait, it would be experimentally simpler to use Tre-6-P to which the intermediary factor is known to have an affinity. Thus, Tre-6-P could be covalently adhered onto a solid phase resin which would be used for the affinity purification of partially purified plant extract known to contain the intermediary factor. Such affinity chromatography approaches are commonly used in proteomics,[5] and could allow isolation and identification of the unknown protein factor.

For the successful application of this technique, a Tre-6-P analogue is required that can be covalently linked to a solid phase resin. This analogue must be carefully designed to ensure adequate binding with the desired protein. A critical parameter in determining this capability is the site on Tre-6-P from which the molecule could be tethered to a solid phase. Structure-activity relationships derived from binding data of a library of Tre-6-P analogues would allow the determination of the binding epitope, which would reveal the locations of the sugar at which modifications would be tolerated. Binding of Tre-6-P analogues to the intermediary factor can be indirectly probed by monitoring the effect of these sugars on the activity of SnRK1, which has a very specific recognition motif that can be utilized for an activity assay.[6] In such an assay, partially purified cell lysate containing SnRK1 and the intermediary factor is treated with a peptide containing the kinase recognition sequence and γ ^{32}P-ATP. The activity of SnRK1 can be followed by monitoring the phosphorylation of the peptide sequence, which can easily be detected by virtue of the radiolabel.[7] Thus, the inhibition of SnRK1 with various Tre-6-P analogues can easily be determined to reveal the binding epitope with the intermediary factor.

However, care must be taken in the interpretation of these results, since artefacts can easily result from the metabolism of any exogenous compounds. The crude cell lysate would likely contain a plethora of active enzymes that could facilitate the inter-conversion of carbohydrate phosphate esters, which are particularly prone to

biochemical transformations since they are often reactive intermediates in various metabolic pathways.[8] For example, Tre-6-P,[9] Glc-6-P[10] and Glc-1-P[11] have all been shown to inhibit SnRK1 in a crude plant extract assay, but are also readily inter-converted *in vivo*.[12] Whether this conversion occurs in an *in vitro* assay, and the consequential effect of this on the apparent SnRK1 activity, has yet to be addressed. Preliminary research from M. J. Paul *et al.* has hinted at such metabolic transformations for Glc-6-P, whose inhibition of SnRK1 is found to be variable depending on the exact conditions used for the assay.[11] More recently, ribose-5-phosphate (Rib-5-P) was also found to inhibit SnRK1 in an analogous crude wheat (*Triticum aestivum*) endosperm assay,[13] but this is inconsistent with unpublished data from the M. J. Paul group and so is believed to be an erroneous result caused by *in vitro* metabolism of the sugar under the assay conditions.[11,14] Investigating the metabolism of exogenous compounds is essential for eliminating rogue data in order to determine the true inhibitors for SnRK1.

In this chapter, the inhibition of SnRK1 with various compounds is explored.[*] Initially, metabolism studies are conducted that verify the stability of Tre-6-P, Glc-1-P and Glc-6-P under the assay conditions. In contrast, Rib-5-P is completely metabolized consuming ATP in the process. Evidence is presented to prove that this loss of ATP results in an apparent inhibition of SnRK1 and that, contradictory to the original report,[13] Rib-5-P itself is not involved in any direct interactions with the intermediary factor. Having established that Tre-6-P as a true inhibitor, the methodology developed in the previous two chapters is utilized to build a small library of analogues which provide a map of the binding epitope, which reveals that modifications can be tolerated at the C6' centre. Using a linker attached to this position, a Tre-6-P affinity column is synthesized and is used to extract proteins from a partially purified SnRK1 preparation. Proteomics analysis of the strongly retained components revealed the 14-3-3 class of protein as potential candidates for the intermediary factor. The omega isoform of the *A. thaliana* 14-3-3 was expressed in *E. coli* and its interaction with Tre-6-P was studied with isothermal titration calorimetry (ITC), circular dichroism (CD) spectroscopy and NMR spectroscopy, but in all cases, no binding was detected. Finally, the solid phase tethering approach was expanded to

[*] The research presented in this chapter was designed in collaboration with Dr M. J. Paul and Dr. L. F. Primavesi at Rothamsted Research, Harpenden, UK. Specific experimental contributions are highlighted as appropriate.

a carbohydrate microarray for the screening of a larger number of potential Tre-6-P binding proteins.

4.2 Metabolism of exogenous sugars *in vitro*

Initially, we sought to test the stability of Tre-6-P, Glc-1-P and Glc-6-P under the *in vitro* assay conditions, since these compounds are widely regarded as inhibitors of SnRK1. The phosphorylated sugars (at a final concentration of 1mM) were incubated with 0.2mM ATP, 5mM $MgCl_2$ and crude *A. thaliana* extract[†] in 40mM HEPES buffer at 30˚C for 10 minutes to simulate the assay conditions (section 4.11.1, page 169). To monitor the metabolism of these sugars, three different experimental approaches were used, initially focusing on Tre-6-P (Figure 4.1). ^{31}P NMR of the crude reaction mixture allowed direct probing of the exogenous sugars and any transformations to other phosphorylated products, which in the case of Tre-6-P, were found not to occur (A). These results were verified using strong anion exchange (SAX) LC-MS, where singly phosphorylated sugars are eluted as a single peak at ~15 minutes and can thus be easily probed by mass spectrometric analysis. As with the ^{31}P NMR data, Tre-6-P was the sole phosphorylated carbohydrate that was detected in the mixture (B). The integrals for the Tre-6-P peaks in both the ^{31}P NMR and SAX LC-MS experiments indicated the complete preservation of Tre-6-P. This was confirmed by probing dephosphorylation products that would not be directly detected with the previous techniques. The crude mixture was per-*O*-acetylated and analysed by C_{18} LC-MS and compared against trehalose octaacetate as a reference, which would be the expected end-product from the dephosphorylation of Tre-6-P (C). Such non-phosphorylated products were not detected and together, these data indicate the general stability of Tre-6-P under these assay conditions.

[†] All plant extracts presented herein were prepared by Dr L. F. Primavesi at Rothamsted Research, Harpenden, UK.

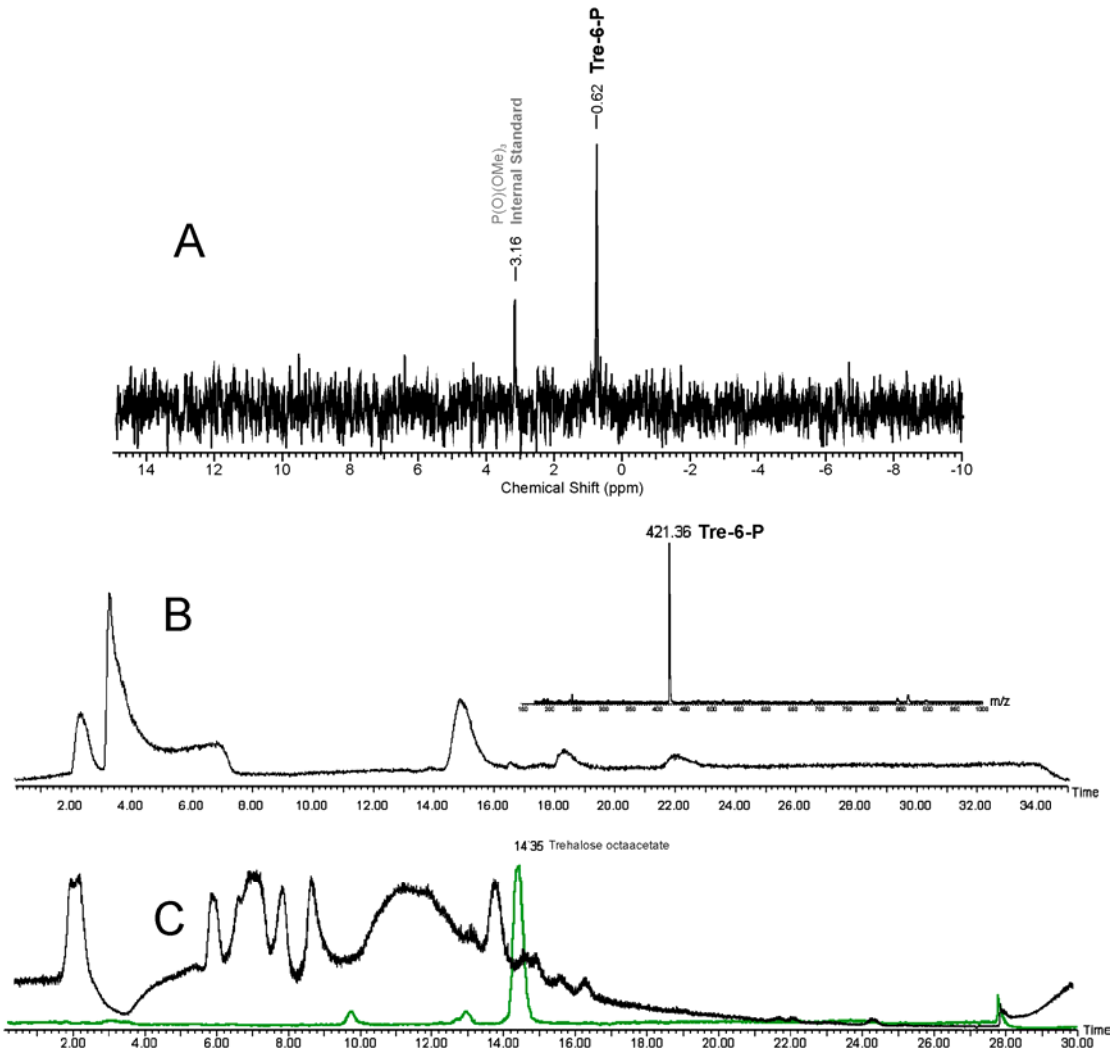


Figure 4.1. *Analysis of the* in vitro *metabolism of Tre-6-P with crude* A. thaliana *extract after 10min incubation at 30˚C. Data shows that Tre-6-P is stable in the reaction mixture and not converted to other products. (A)* ^{31}*P NMR referenced to* $P(O)(OMe)_3$*; (B) SAX LC-MS ion count trace and mass spectrum of the major peak at ~15mins; (C)* C_{18} *LC-MS ion count trace of the crude mixture after per-*O*-acetylation (black) and of trehalose octaacetate as a reference (green).*

A similar analysis was performed for Glc-1-P (Figure 4.2) where ^{31}P NMR and SAX LC-MS showed no conversion to other phosphorylated species. As with Tre-6-P, the peak integrals were indicative of complete preservation of Glc-1-P. This was confirmed by C_{18} LC-MS of the per-*O*-acetylated reaction mixture, which failed to reveal any evidence of dephosphorylation. Thus, Glc-1-P was also found to be fully stable in an *in vitro* SnRK1 activity assay.

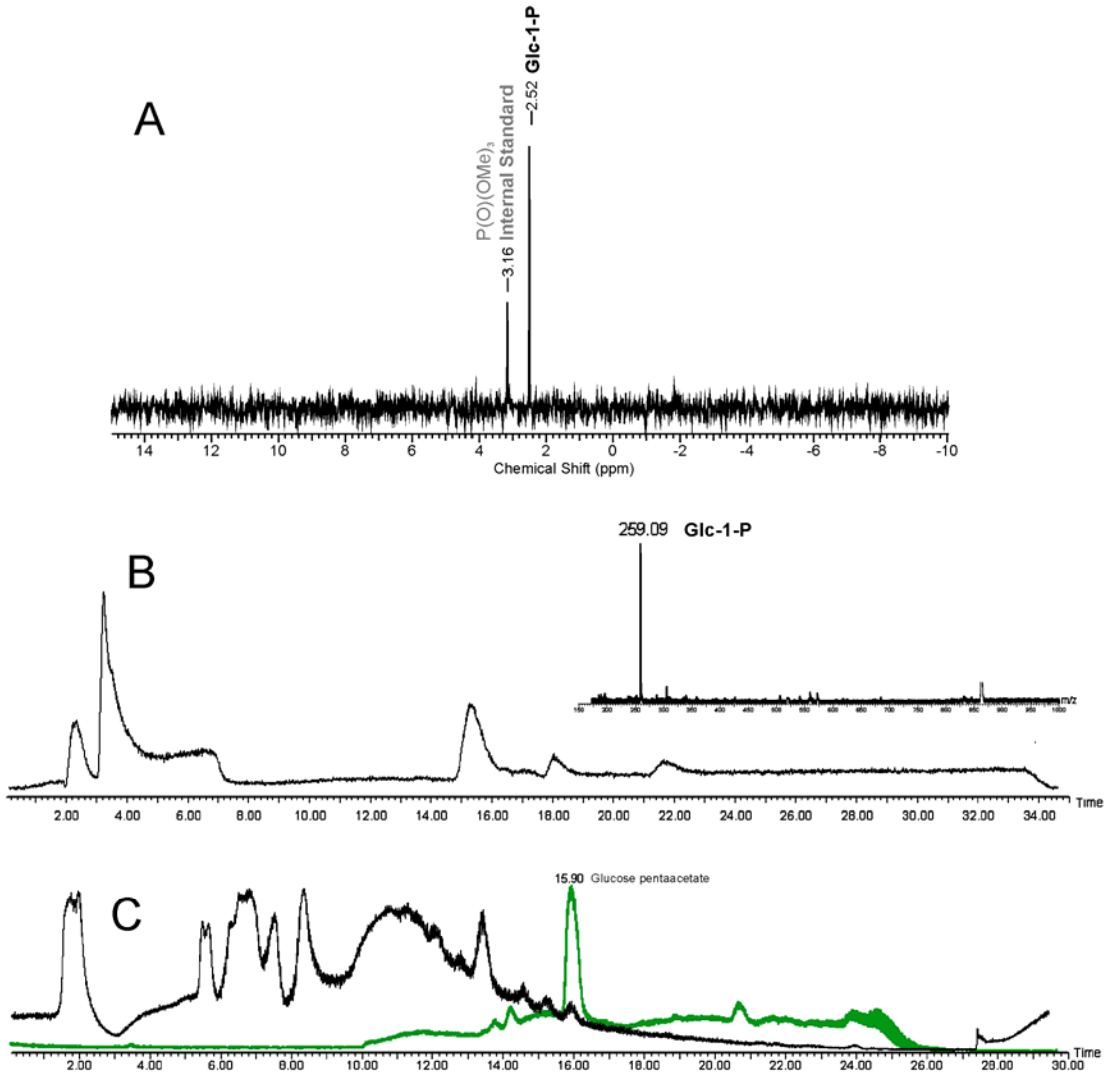


Figure 4.2. *Analysis of the* in vitro *metabolism of Glc-1-P with crude* A. thaliana *extract after 10min incubation at 30˚C. Data shows that Glc-1-P is stable in the reaction mixture and not converted to other products. (A)* ^{31}*P NMR referenced to* $P(O)(OMe)_3$*; (B) SAX LC-MS ion count trace and mass spectrum of the major peak at ~15mins; (C)* C_{18} *LC-MS ion count trace of the crude mixture after per-*O*-acetylation (black) and of glucose pentaacetate as a reference (green).*

In contrast, Glc-6-P was found to be less stable under these conditions (Figure 4.3). After 10 minutes of incubation, ^{31}P NMR revealed approximately 10% conversion to Glc-1-P. SAX LC-MS is less useful in this case since the molecular weights of Glc-1-P and Glc-6-P are identical, but nevertheless confirms the lack of conversion to any other carbohydrate phosphate esters. As with the previous cases, dephosphorylation was not detected by C_{18} LC-MS.

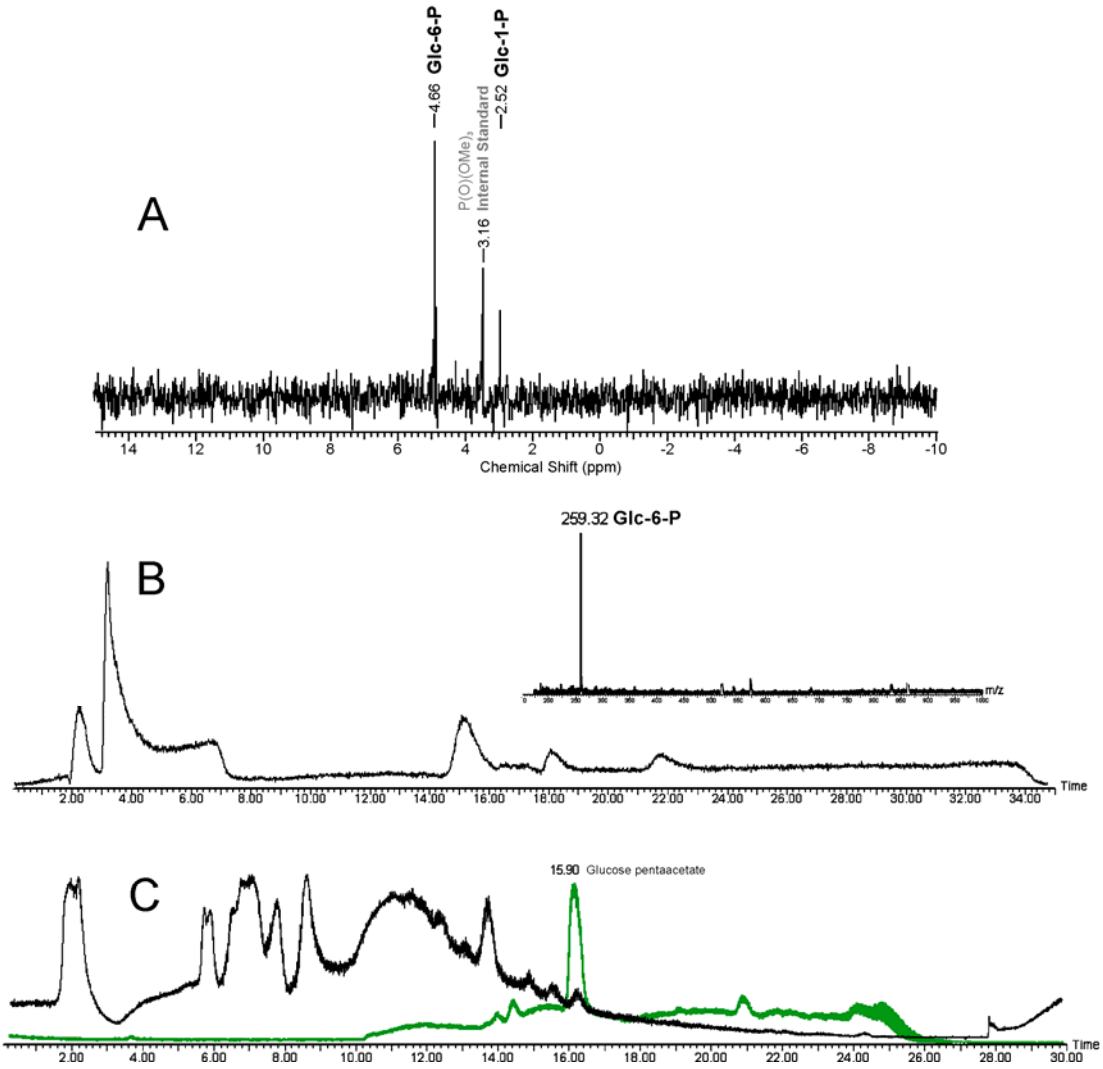


Figure 4.3. *Analysis of the* in vitro *metabolism of Glc-6-P with crude* A. thaliana *extract after 10min incubation at 30˚C. Data shows that Glc-6-P is partially converted to Glc-1-P (~10% conversion) after incubation in the assay mixture. (A)* ^{31}*P NMR referenced to P(O)(OMe)*$_3$*; (B) SAX LC-MS ion count trace and mass spectrum of the major peak at ~15mins; (C)* C_{18} *LC-MS ion count trace of the crude mixture after per-*O*-acetylation (black) and of glucose pentaacetate as a reference (green).*

Incubating Glc-6-P for longer periods increased the conversion to Glc-1-P. Thus, after 30 minutes at 30˚C, ^{31}P NMR showed approximately 50% conversion (Figure 4.4). The inhibition of SnRK1 by Glc-6-P was found to be less consistent than that of Tre-6-P and Glc-1-P,[11] and this observation can now be attributed to the metabolism of Glc-6-P under the assay conditions, which would likely produce some variation between experiments. Nevertheless, for the 10 minute incubation times that are generally used for kinase activity assays, only a minimal conversion of Glc-6-P is observed, while Tre-6-P and Glc-1-P were completely stable. This validates the previously observed inhibition of SnRK1 and confirms that Tre-6-P, Glc-1-P and Glc-6-P are in fact real inhibitors.

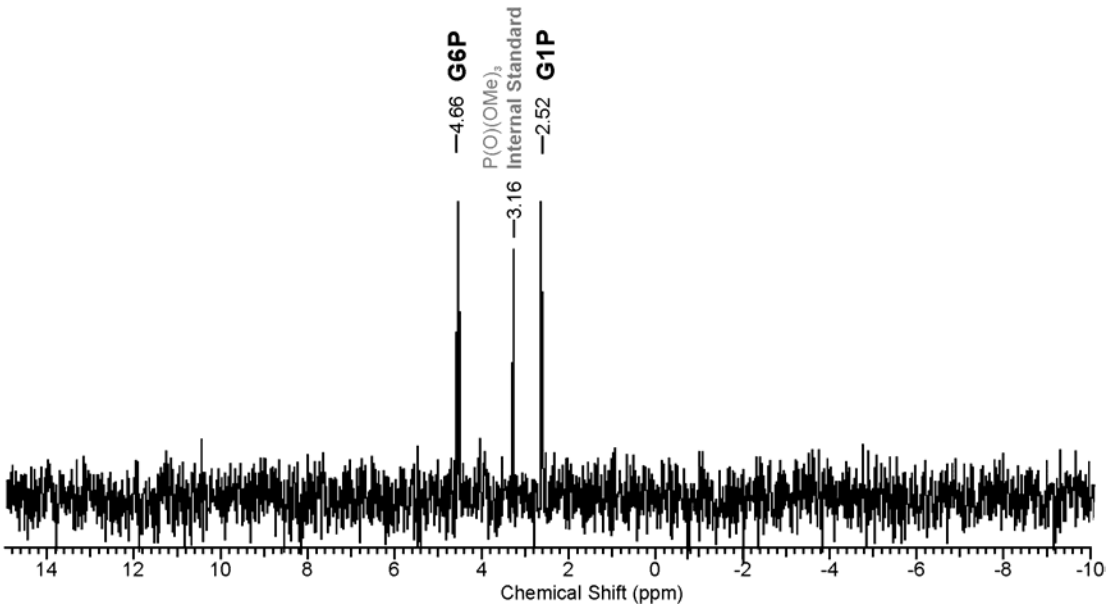
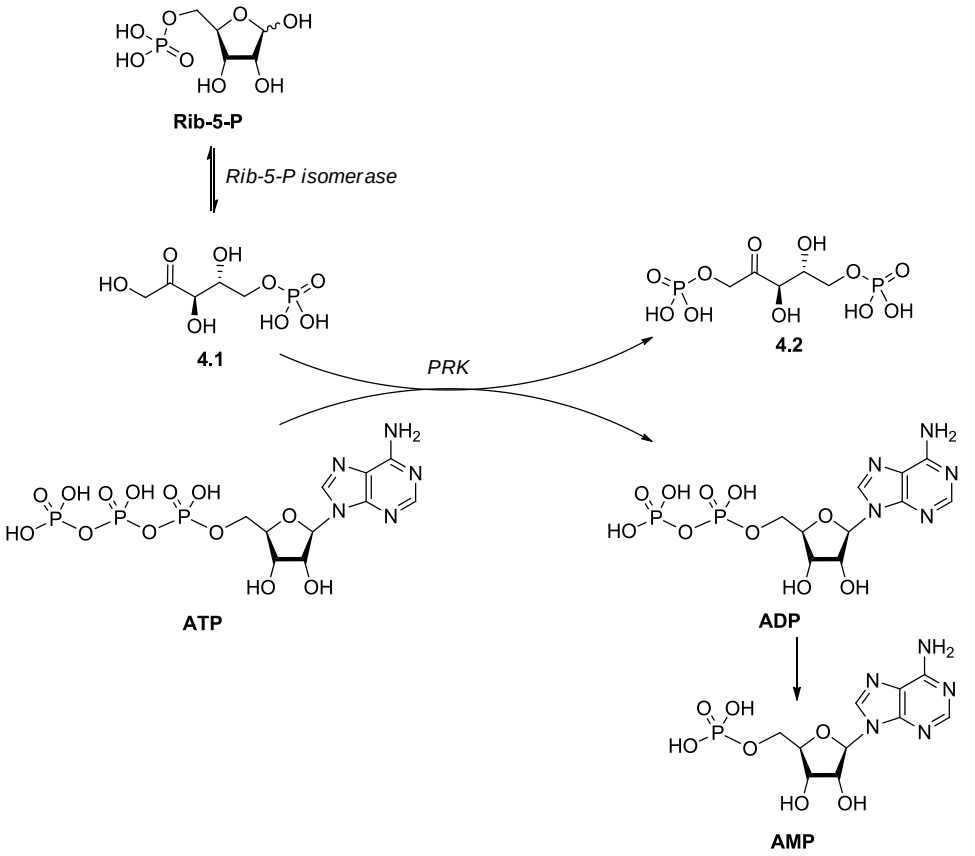


Figure 4.4. *^{31}P NMR of the metabolism of Glc-6-P in crude* A. thaliana *extract after 30mins of incubation at 30˚C. Data shows increased conversion of Glc-6-P to Glc-1-P after extended incubation times (~50% conversion)*

4.3 Metabolism of ribose-5-phosphate

Recently Piattoni *et al.* reported that SnRK1 from growing wheat endosperm is inhibited by Rib-5-P based on evidence obtained from crude extract assays as above.[13] However, such experiments are prone to interference from other active enzymes in the crude extract and are therefore not always reliable, as has been illustrated for Glc-6-P. This is especially true for Rib-5-P, which is a central synthetic intermediate involved in both the Calvin cycle[15] and the pentose phosphate pathway.[16] These fundamental metabolic pathways are essential for plant survival and enzymes for these routes are likely to be prevalent in crude plant extracts. Thus, Rib-5-P may be subjected to very rapid metabolism which must be correctly accounted for.

Rather than direct inhibition of SnRK1, we proposed an alternative route that could account for the observed reduction in kinase activity (Scheme 4.1). In this scheme, Rib-5-P is under equilibrium with its isomer ribulose-5-phosphate (**4.1**) through the action of Rib-5-P isomerase. Phosphoribulosekinase (PRK), a naturally abundant and highly active enzyme due to its essential function in the Calvin cycle, could then convert **4.1** to ribulose-1,5-biphosphate (**4.2**), consuming ATP in the process. Since ATP is the substrate for SnRK1, any decrease in its concentration would result in a lower kinase activity, which could be misinterpreted as inhibition by Rib-5-P.



Scheme 4.1. *Proposed route for ATP depletion through the metabolism of Rib-5-P with Rib-5-P isomerise and PRK.*

Consistent with this theory, recent unpublished work from the M. J. Paul group[14] has demonstrated that the activity of SnRK1 from *A. thaliana* extracts decreases over time when incubated with Rib-5-P, rather than a fixed decrease in activity that would be expected from the inhibition model. Importantly, this research has also demonstrated that the kinase activity can be recovered by using higher concentrations of ATP, providing further support for our hypothesis. We sought to rigorously test for the consumption of ATP using the tools of analytical chemistry.

The concentration of ATP and the reaction products ADP and AMP can easily be followed by HPLC, since the compounds are readily separable by SAX chromatography and can be detected using their UV absorbance at 260nm. We initially used crude *A. thaliana* extract to test the overall hypothesis (Figure 4.5). Reactions were performed at 30˚C for 6 minutes, with 0.2mM ATP and a final concentration of 1mM for the Rib-5-P under the exact conditions used in the original paper (section 4.11.1, page 169).[13] SAX HPLC indicated the complete consumption of ATP in the presence of Rib-5-P and the formation of AMP rather than ADP. A

control reaction with ADP in the place of ATP also showed full conversion to AMP indicating a naturally low metabolic stability for ADP under these conditions. Additionally, we used LC-MS in negative mode with single ion recording (SIR) to selectively detect ions of mass 229 and 309 Daltons, corresponding to Rib-5-P and **4.2** respectively. Thus, in the presence of Rib-5-P, a build-up of **4.2** was detected as would be expected from the action of PRK.

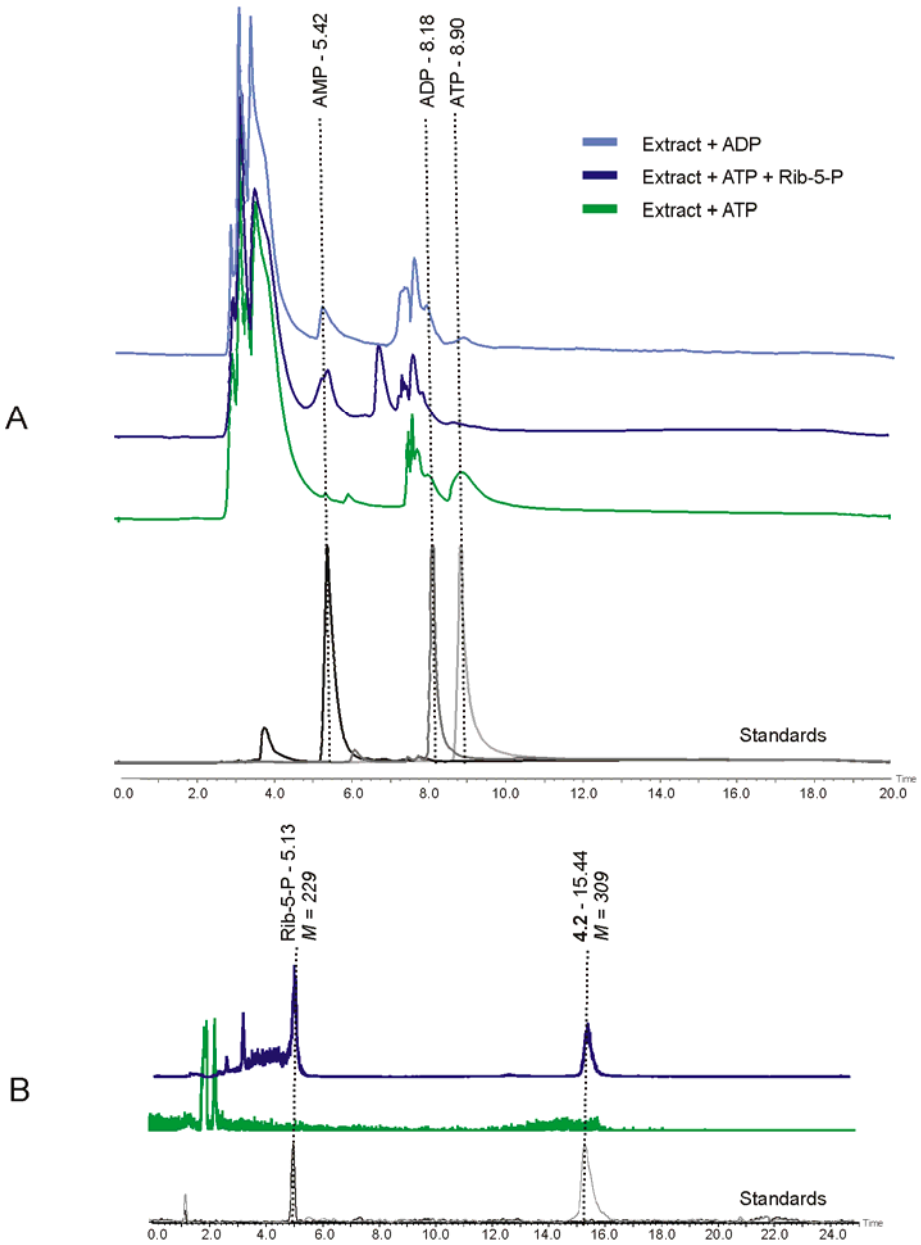


Figure 4.5. *Analysis of crude* A. thaliana *extract after 6mins incubation at 30˚C with ATP, ADP and/or Rib-5-P as indicated. Data shows that ATP is depleted and converted to AMP in the presence of Rib-5-P and crude plant extract, with build up of* ***4.2*** *detected as a reaction product. (A) SAX HPLC, UV = 260nm; (B) SAX LC-MS with SIR of masses at 229 and 309 Da.*

Having established the principle on *A. thaliana* extract, we performed analogous analyses on crude whole wheat (*Triticum aestivum*) extract (Figure 4.6) and post-anthesis wheat endosperm extract (Figure 4.7) to replicate the exact conditions used by Piattoni *et al.*[13] In both cases, as with *A. thaliana* extract, ATP was completely converted to AMP. Mass spectrometric analysis showed an accumulation of **4.2** as a metabolic end-product. Together, these data provide strong evidence to support our ATP depletion hypothesis as a viable explanation for the observed decrease in SnRK1 activity.

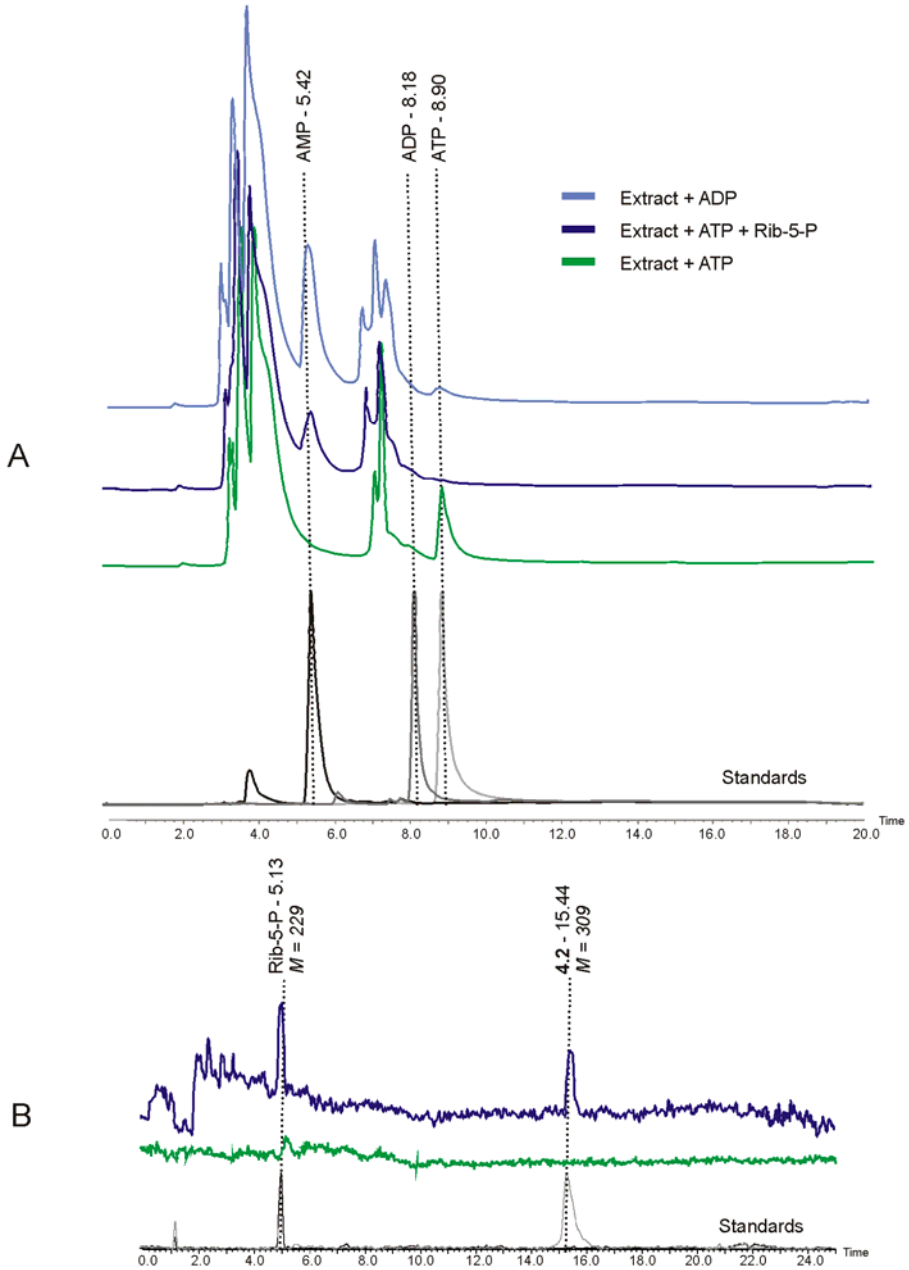


Figure 4.6. *Analysis of crude whole wheat extract after 6mins incubation at 30˚C with ATP, ADP and/or Rib-5-P as indicated. Data shows that ATP is depleted and converted to AMP in the presence of Rib-5-P and crude wheat extract, with build up of* ***4.2*** *detected as a reaction product. (A) SAX HPLC, UV = 260nm; (B) SAX LC-MS with SIR of masses at 229 and 309 Da.*

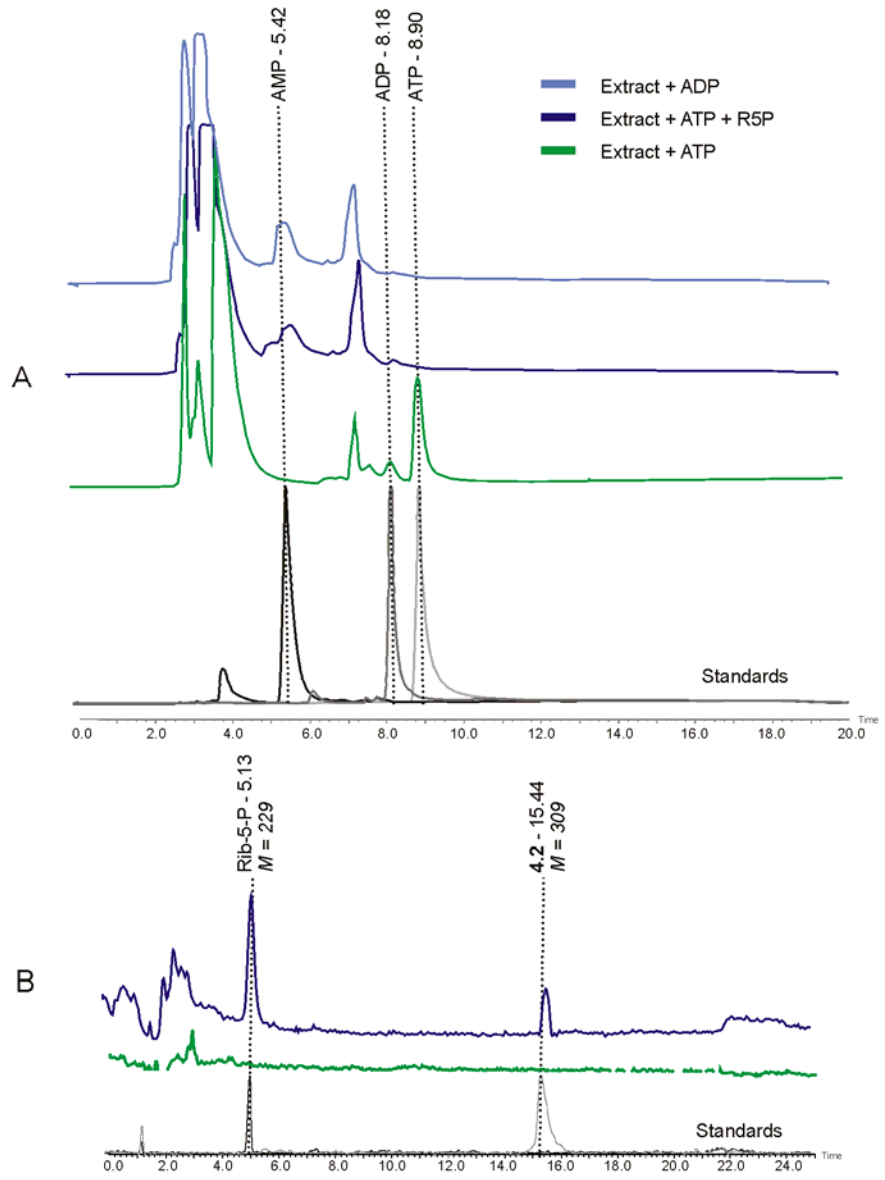


Figure 4.7. *Analysis of crude wheat endosperm extract after 6mins incubation at 30˚C with ATP, ADP and/or Rib-5-P as indicated. Data shows that ATP is depleted and converted to AMP in the presence of Rib-5-P and crude wheat endosperm extract, with build up of* ***4.2*** *detected as a reaction product. (A) SAX HPLC, UV = 260nm; (B) SAX LC-MS with SIR of masses at 229 and 309 Da.*

To demonstrate that the depletion of ATP is due to PRK activity, *N. tobaccum* plants engineered with an antisense PRK gene to have a reduced expression of this enzyme were grown.[17] Crude extract from these transgenic tobacco plants were subjected to the assay conditions and monitored for ATP depletion and accumulation of **4.2** as above (Figure 4.8). HPLC analysis showed that ATP was retained for transgenic plants, while **4.2** could not be detected using LC-MS. In contrast, wild-type tobacco extract consumed all exogenous ATP, which was converted to ADP, and was found to contain detectable levels of **4.2**.

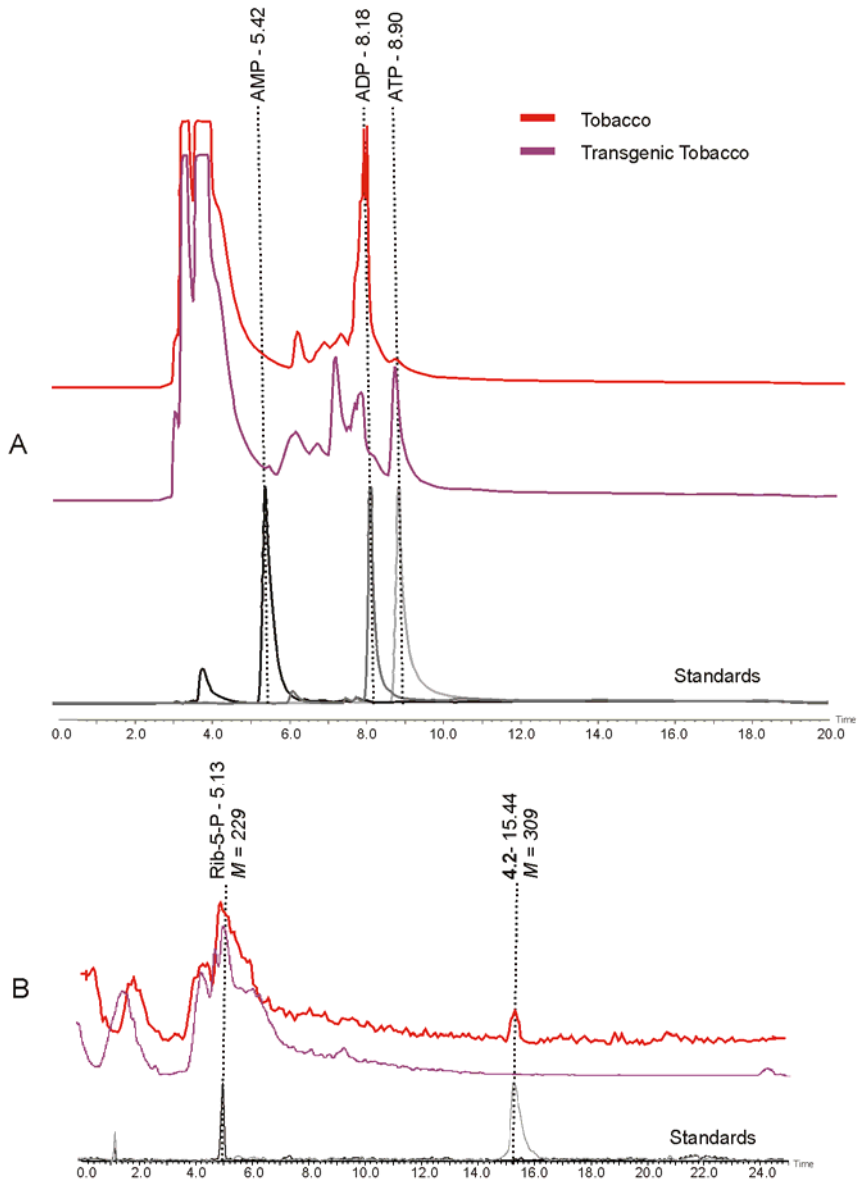


Figure 4.8. *Analysis of wild-type and transgenic tobacco crude extracts after 6mins incubation at 30˚C with 0.2mM ATP and 1mM Rib-5-P. Data shows depletion of ATP and build-up of* ***4.2*** *in wild type tobacco plant extract in the presence of Rib-5-P, but not for transgenic plants with a lower expression of PRK. (A) SAX HPLC, UV = 260nm; (B) SAX LC-MS with SIR of masses at 229 and 309 Da.*

The decrease in SnRK1 activity observed in growing wheat endosperm extract was attributed by Piattoni *et al.* to inhibition of the kinase by Rib-5-P.[13] Here, we have presented conclusive evidence to show that this viewpoint is in fact false. Our data demonstrates that this fall in activity is actually caused by a drastic depletion of ATP which starves SnRK1 of essential substrate. This emphatically illustrates the importance of conducting detailed metabolic studies when working with crude cell lysates.

4.4 Library of analogues modified at the phosphate centre

Having verified the reliability of the SnRK1 activity assay for Tre-6-P, we turned our attention to probing the binding epitope of the sugar with the intermediary factor. Our initial goal was to probe the involvement of the phosphate group in binding with the intermediary factor. Armed with our knowledge from the previous chapter, Tre-6-P analogues with modifications at the phosphate centre could be readily synthesized. In the course of our research, various useful Tre-6-P analogues were already synthesized (as documented in Chapter 3) and could be used as the basis of a "phosphorous centre" library (Figure 4.9).

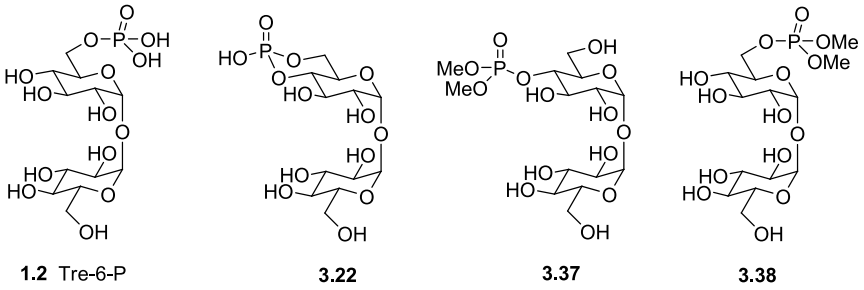
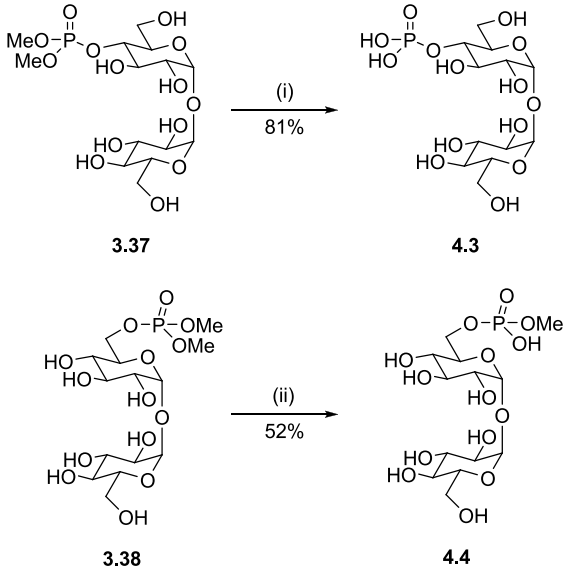


Figure 4.9. *Tre-6-P analogues previously synthesized in Chapter 3.*

Compounds **3.37** and **3.38** were deprotected by methoxy removal to give a wider range of phosphate analogues (Scheme 4.2). Compound **3.37** was fully deprotected with TMSI to give the 4-phosphate **4.3** in 81% yield, while **3.38** was partially deprotected with TMSBr to give the methoxy ester **4.4** in 52% yield.



Scheme 4.2. *Synthesis of other variants of Tre-6-P with modifications at the phosphorous centre. Conditions: (i) TMSI, dioxane, RT, 30mins; (ii) TMSBr, dioxane, RT, 3h.*

This array of analogues was tested for SnRK1 inhibition at a final concentration of 1mM (Figure 4.10).[‡] The data revealed interesting trends in the binding epitope.

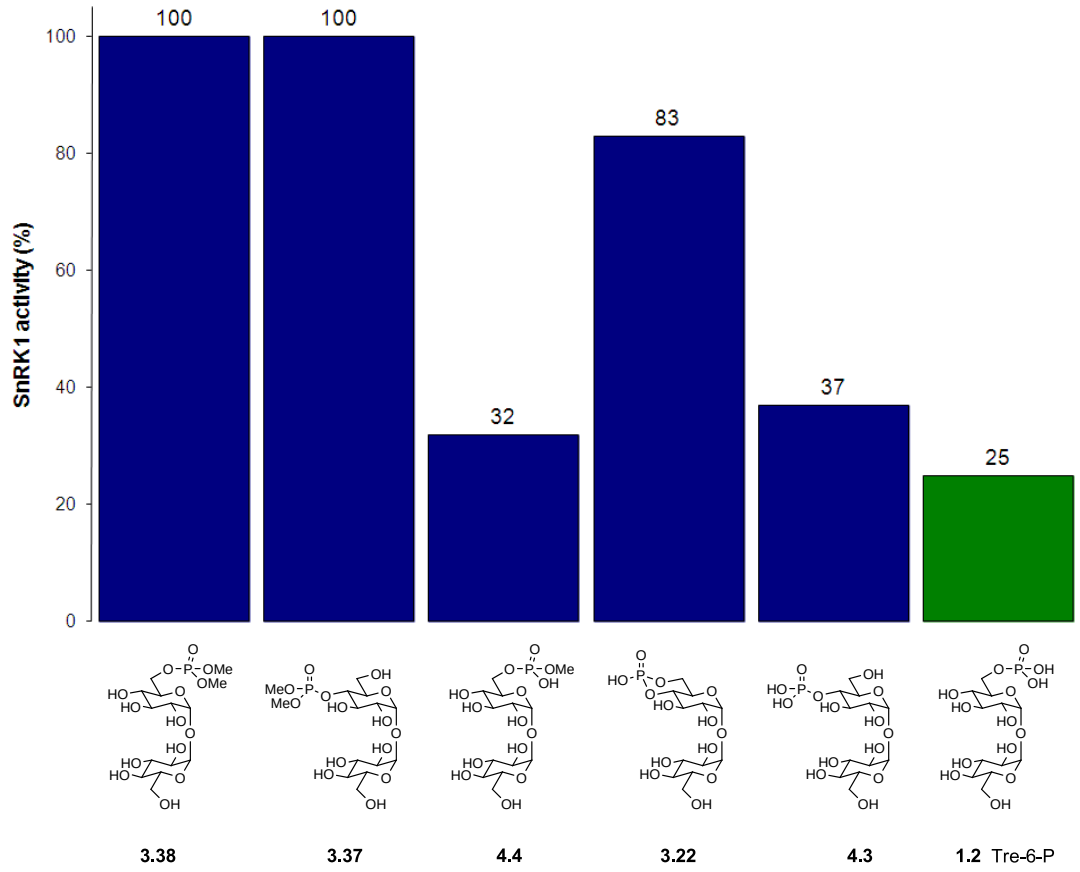


Figure 4.10. *Activity of SnRK1 in the presence of 1mM analogue concentration. 100% indicates SnRK1 activity in the absence of an inhibitor. The natural inhibitor, Tre-6-P, is shown in green. Data shows that the phosphate group is essential for inhibition of SnRK1, but loss of a P-OH or different locations of the phosphate group can be tolerated.*

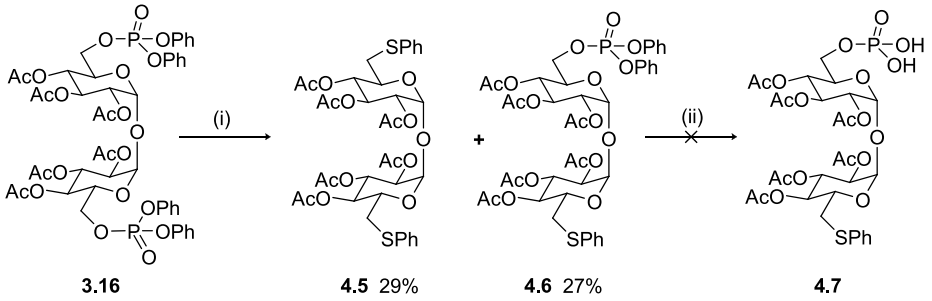
Given that trehalose does not inhibit SnRK1, one would expect the phosphate moiety of Tre-6-P to play a critical role in binding to the intermediary factor. Therefore, as expected, completely capped phosphate groups are unable to inhibit the kinase (**3.37** and **3.38**). However, capping just one hydroxyl of the phosphate (**4.4**) restores the binding to the intermediary factor. Thus, SnRK1 is inhibited to 32% of its normal activity, which compares favourably with the natural inhibitor Tre-6-P that reduces kinase activity to 25%. This implies that only one phosphate hydroxyl is needed and that the second is, to an extent, dispensable. Supporting this viewpoint, the cyclic phosphate **3.22** also inhibits SnRK1 to 83% of its activity. The reduced inhibition of this analogue may reflect the restricted flexibility of the phosphate hydroxyl that

[‡] SnRK1 activity assays were performed by Dr L. F. Primavesi at Rothamsted Research, Harpenden, UK.

would result from a cyclization. Finally, displacing the phosphate to the 4-position (**4.3**) reduces inhibition at 37% to normal activity, indicating a level of tolerance to the precise location of the phosphoryl functionality. This combined data is suggestive of a broad phosphate binding pocket around the 4 and 6 positions of trehalose, necessitating at least one strong interaction with a phospho-hydroxyl group.

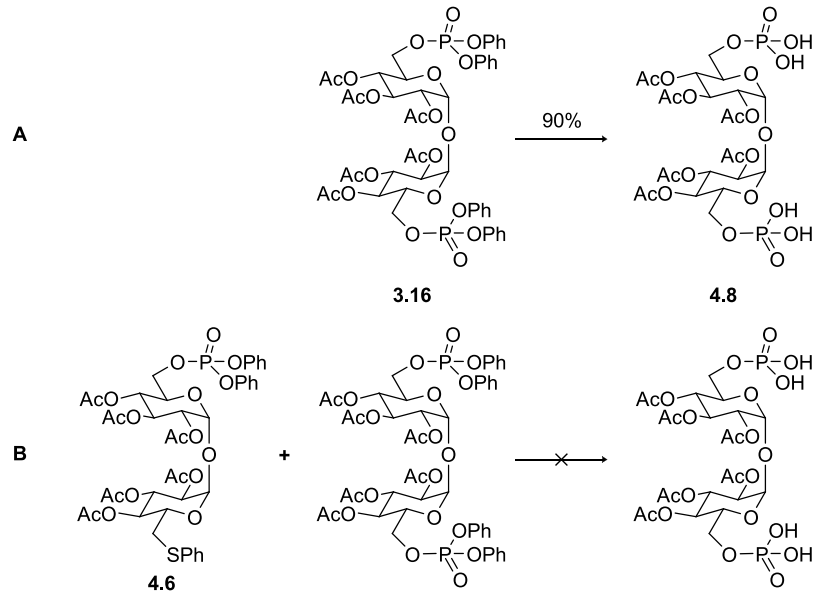
4.5 Library of analogues modified at the C6' centre

Attention was then focused towards screening the C6' centre for tolerance to different functionalities. We hypothesized that the C6' centre would be most suitable for attachment to a solid phase resin, since it is the furthest away from the critical phosphate group which would likely be intimately bound to the intermediary factor. Thus, our primary motive was to investigate the effect of bulk at this position. The general synthetic strategy mirrored the S_N2 route that was developed in the previous chapter for the azide **3.70**. Nucleophilic thiols can participate in such S_N2 reactions and provide a means to install different substituents at the C6' centre. Initially, thiophenol was used as a model to optimize the synthetic route (Scheme 4.3).



Scheme 4.3.*Attempted deprotection of diphenyl phosphate moiety in the presence of an SPh group. Conditions: (i) PhSH, NaH, THF, microwave, 150°C, 2h. (ii) H_2, PtO_2, MeOH, 1% AcOH.*

For a successful displacement reaction, deprotonation of the thiol was found to be necessary. Thiophenol was treated with sodium hydride, prior to microwave irradiation with **3.16** at 150°C in THF for 2 hours which gave the doubly modified adduct **4.5** as a side product in 29% yield as well as the desired product **4.6** in a reasonable 27%. Frustratingly, the following deprotection of the diphenyl phosphate ester to give **4.7** using hydrogen over a PtO_2 catalyst was unsuccessful. This was particularly surprising because previous analogous hydrogenations have been very rapid, which warranted further investigations into the cause of the lack of reaction.



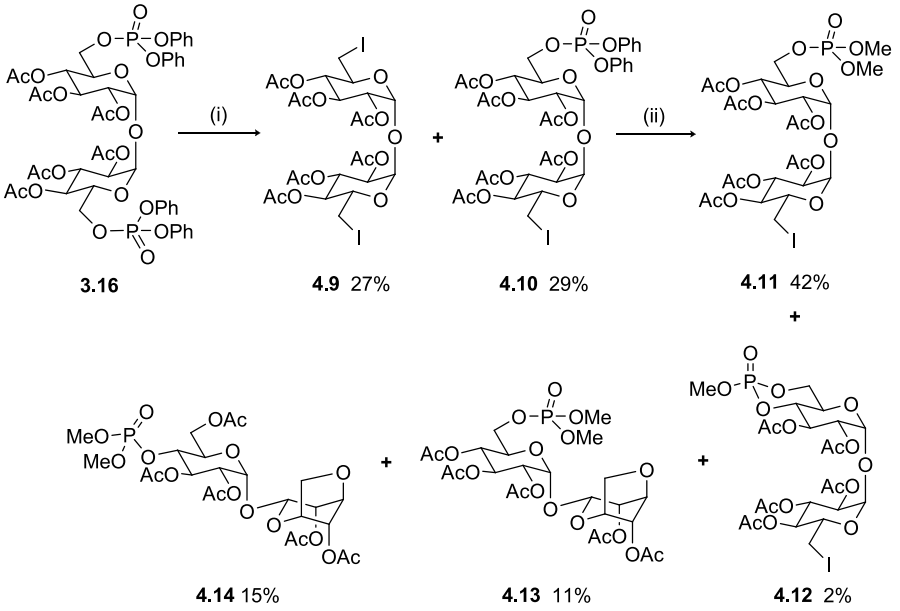
Scheme 4.4. *Competion reaction shows that under normal conditions, hydrogenic phosphate deprotection is facile, but in the presence of a thioether, the reaction does not occur. Conditions: H_2, PtO_2, MeOH, 1% AcOH, 32h.*

We argued that the poor deprotection observed for **4.6** was due to the thioether functionality that would have a strong affinity towards the PtO_2 leading to poisoning of the catalyst surface.[18] To test this, we attempted the hydrogenation of **3.16**, which ordinarily proceeds smoothly to give **4.8** in 90% yield, in the presence of **4.6** (Scheme 4.4). Under these conditions, the reaction was essentially halted which would seem to support our original hypothesis.

Fortunately, a CsF transesterfication was developed in the previous chapter to allow exchange of the phenoxy groups with methoxy functionality, which would circumvent the need for a hydrogenation step. While this would certainly be a viable route, the transesterfication step would produce a 4-phosphate as a side product that would necessitate a difficult separation to isolate the desired product. For the synthesis of a library of compounds, such tedious purification steps would ideally be limited and a more succinct route would be preferred. With this in mind, an alternative scheme was designed based on the synthesis of a common intermediate that would allow late-stage diversification with minimal purification steps.

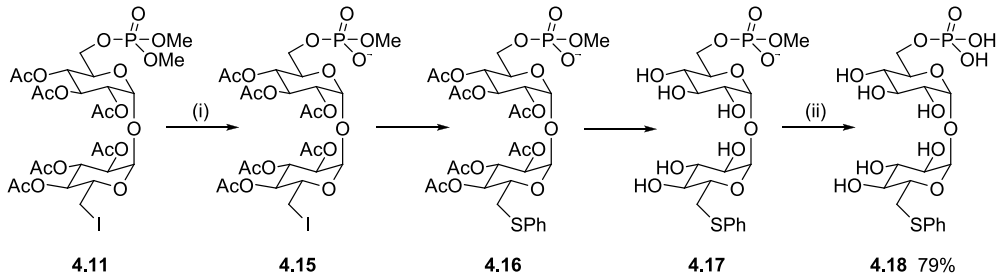
In this strategy, earlier manipulations were performed with an iodide to form the central intermediate (Scheme 4.5). Compound **3.16** was treated with tetrabutylammonium iodide under microwave irradiation to displace the phosphate groups yielding the doubly modified adduct **4.9** as a side product in 27% as well as

the desired product **4.10** in 29%. Here, the iodide would serve as a labile leaving group for later diversification to other analogues. Hence, phosphate transesterfication was conducted at this earlier stage to limit the route to only one difficult purification step. The diphenylphosphate **4.10** was treated with CsF and methanol to yield the desired dimethylphosphate **4.11** in a reasonable 42% yield as well as various other side products of which the cyclic methoxy ester **4.12** was present as a minor component. Interestingly, the iodide was prone to intramolecular cyclization after ring flipping to give the 6-phospho (**4.13**) and 4-phospho (**4.14**) bicyclic analogues in 11% and 15% respectively.



Scheme 4.5.*Alternative route for sythesis of library via intermediate **4.11**, the synthesis of which requires careful control to avoid side product formation. Conditions: (i) TBAI, THF, microwave, 150°C, 2h; (ii) CsF, MeOH, reflux, 8h followed by Ac_2O, Py, 32h.*

The elegance of this route arises from the subsequent diversification step (Scheme 4.6). Microwave irradiation of **4.11** with thiophenol under basic conditions initially leads to partial methoxy deprotection (**4.15**) which drastically lowers the reactivity of the phosphoryl group due to the resultant anionic charge.[19] Thus, the phosphate functionality is no longer prone to either migration or nucleophilic substitution. S_N2 displacement exclusively occurs at the iodide (**4.16**) followed by acetyl deprotection to yield **4.17** as a stable intermediate. The final methoxy group can be removed by treatment with TMSI to yield the target compound **4.18** in 79% yield over the two steps. Crucially, because the delicate phosphate moiety is locked in place, minimal purification is needed making the route particularly attractive for small-scale libraries.



Scheme 4.6. *One-pot reaction sequence to allow rapid diversification/deprotection of a wide variety of Tre-6-P analogues modified at the C6' position. Conditions: (i) PhSH, NaH, THF, microwave, 80°C, 30mins; (ii) TMSI, dioxane, RT, 60mins.*

This deprotection/ modification cascade can be used with other thiols to create a variety of C6' modified Tre-6-P analogues (Figure 4.11). For example, use of ethanethiol gave **4.19** in 84% yield while octanethiol furnished **4.20** in 85%. Branched thiols can also be used, but in these cases, the reaction was also accompanied by phosphate migration products which were essentially inseparable from the desired products. The bulkier thiols may have a lower reactivity altering the order of events highlighted above (Scheme 4.6). For instance, where acetyl deprotection may precede methoxy deprotection opening up an avenue for phosphate migration. Thus 2-propanethiol produced approximately 80% of the desired 6-phosphate **4.21** according to crude NMR, while **4.22** was approximately 70% pure with phosphate migration products as the major contaminants. Despite their impurity, and given that phosphate migration products are tolerated by the intermediary factor (section 4.4, page 146), these compounds may still provide crude qualitative information on the binding epitope and were included as part of our library for SnRK1 activity analysis.

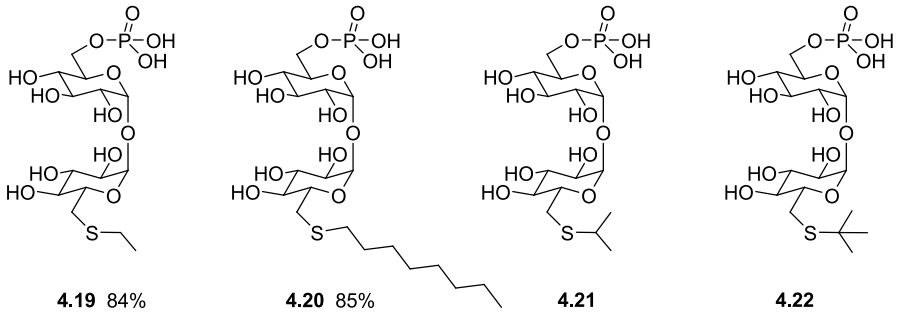
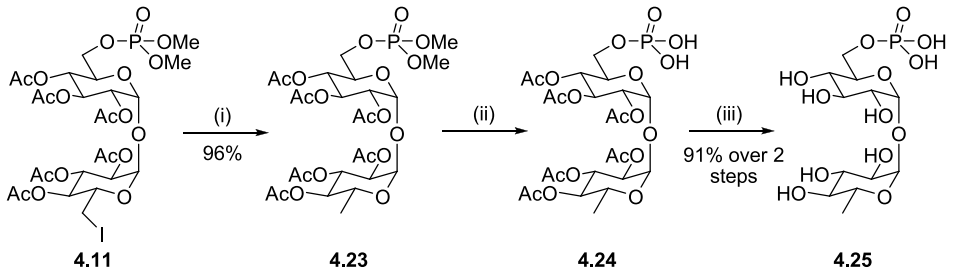


Figure 4.11. *Other Tre-6-P analogues synthesized using the S_N2/deprotection conditions.*

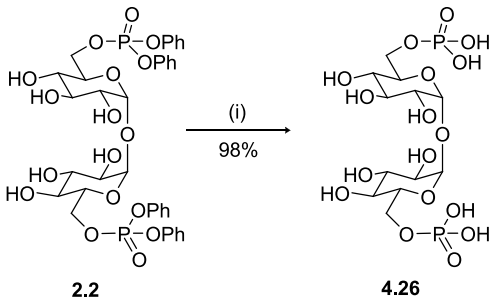
Two further analogues were included as part of our library. To probe the necessity for hydrogen bonding at the C6' position, a C6'-deoxy analogue was synthesized (Scheme 4.7). Hydrogenation of **4.11** over palladium on charcoal allowed reduction of the iodide to give the deoxy analogue **4.23** in 96% yield. Phosphate deprotection was

effected with TMSBr to give **4.24**, which was deacetylated using sodium methoxide in methanol to furnish the desired compound **4.25** in 91% over two steps.



Scheme 4.7. *Synthesis of a C6' deoxy analogue of Tre-6-P. Conditions: (i) H_2, Pd/C, NEt_3, MeOH (ii) TMSBr, dioxane, 50˚C; (iii) NaOMe, MeOH.*

For the final compound, the diphosphate **2.2** was deprotected by hydrogenation over PtO_2 in near quantitative yield to give **4.26** (Scheme 4.8), which could provide information on the tolerance of charged functionality at the C6' position.



Scheme 4.8. *Synthesis of a diphosphate Tre-6-P analogue. Conditions: (i) H_2, PtO_2, 1 MeOH : 3 H_2O, 20mM HCl.*

With the compounds in hand, the inhibition of SnRK1 was tested as before at a 1mM concentration of the exogenous compounds (Figure 4.12).[‡] The deoxy analogue **4.25** inhibited SnRK1 to 50% of its normal activity indicating that loss of a hydrogen bond donor at the C6' position could be tolerated. Similarly, the linear alkyl analogues **4.19** **4.20** also reduced activity to 58% and 57% respectively, while the more inflexible azide group was less accommodated, inhibiting SnRK1 to only 66%. Importantly, these data show that the C6' position is well suited for modification with long chain linkers and could make an ideal site from which to anchor Tre-6-P to a solid phase.

Interestingly, the much bulkier phenyl analogue **4.18** was a potent inhibitor of SnRK1, restricting its activity to 33% which is comparable to Tre-6-P itself. Furthermore, the diphosphate **4.26** also strongly inhibited the kinase to 43% of its normal activity. These surprising results may be indicative of peripheral amino acid

interactions around the C6' position that may result in additional hydrophobic stacking, or ionic bonding for such non-natural analogues.

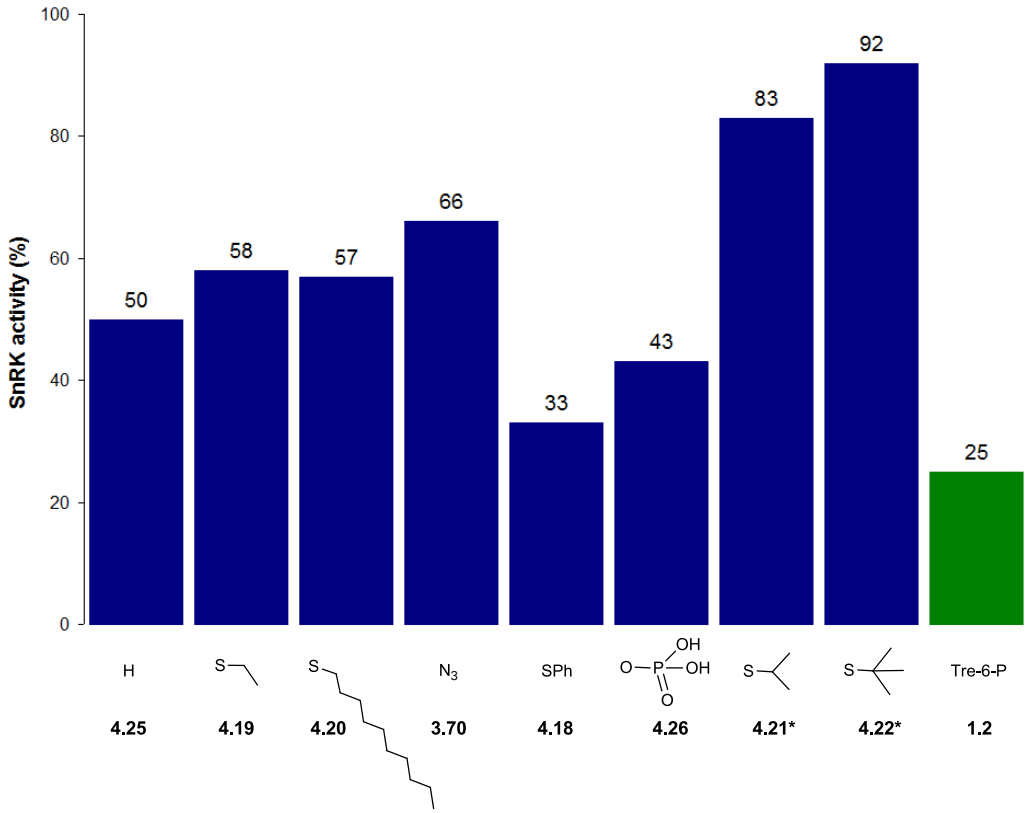


Figure 4.12. *Activity of SnRK1 in the presence of 1mM of the analogues. 100% indicates SnRK1 activity in the absence of an inhibitor. The natural inhibitor, Tre-6-P, is shown in green. *Compounds were contaminated with migration by-products as explained in the main text. Data shows that a variety of functionality, including bulky and charged groups, can be tolerated at the C6' position.*

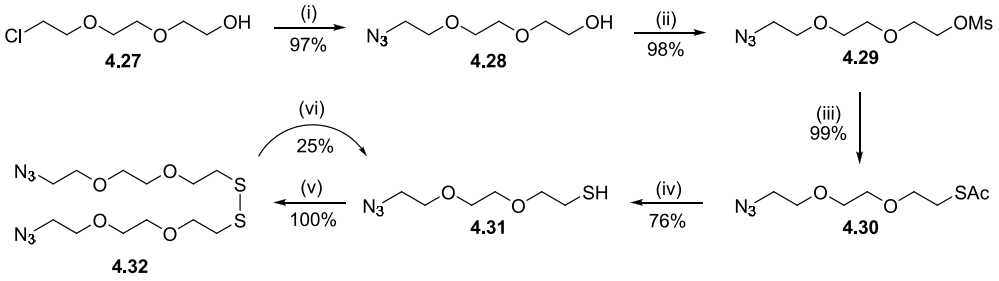
The branched alkyl analogues **4.21** and **4.22**, despite the presence of other phosphate contaminants, were also screened against SnRK1 to give a crude insight into the three dimensional spatial requirements at the C6' position. Kinase activity was slightly inhibited by these compounds (83% and 92% respectively), signifying a moderate spatial tolerance in binding to the intermediary factor at this position.

4.6 Building a Tre-6-P affinity column

Based on this structure-activity data, a Tre-6-P analogue was designed to allow tethering to a solid phase resin. We envisioned the attachment of Tre-6-P using an amine terminated, polyethylene glycol (PEG) spacer arm to provide adequate separation between the resin and the active binding motif. PEG was chosen to prevent chromatographic phase collapse in aqueous buffers,[20] thus ensuring maximal capacity

for protein binding, while the amine would serve as a chemical handle for attachment to the resin.

The synthetic route for this compound mirrored the strategy used for the previous *S*-linked analogues. Attention was initially focused on the synthesis of the PEG linker for the installation of terminal amine and thiol groups (Scheme 4.9).

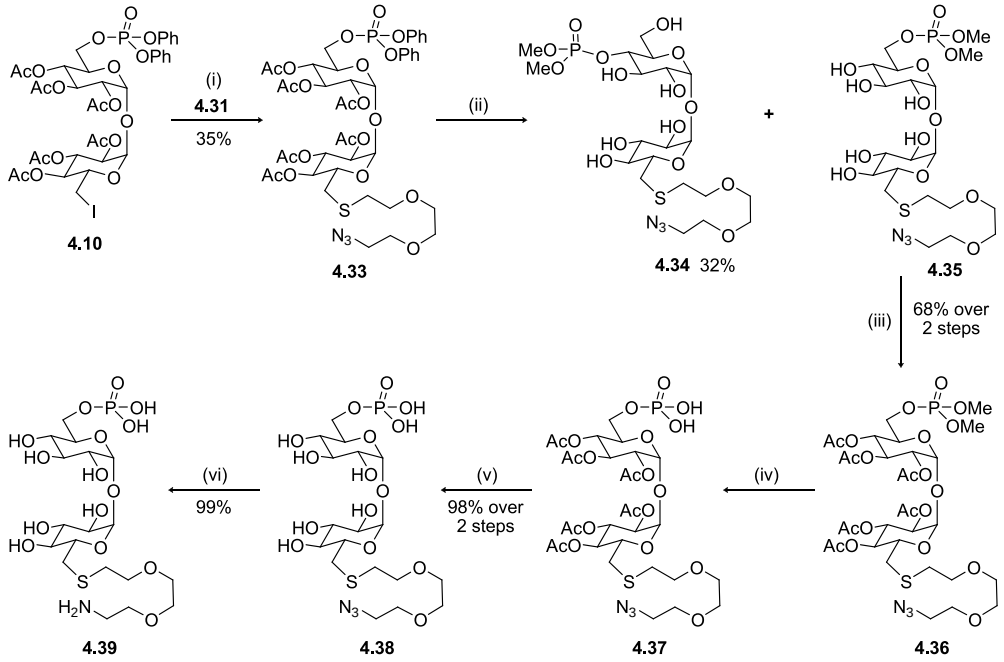


Scheme 4.9. *Conditions: (i) NaN_3, DMF, 80˚C; (ii) MsCl, NEt_3, DCM, 0˚C; (iii) KSAc, DMF, RT; (iv) NaOMe, MeOH, RT; (v) oxidation over 3h; (vi) DTT, 100mM NH_4HCO_3, MeOH, EtOAc, RT.*

Starting with the commercially available triethylene glycol analogue **4.27**, nucleophilic substitution of the chloride with sodium azide afforded **4.28** in 97% yield. The remaining free hydroxyl was mesylated in 98% to give **4.29**, which was treated with potassium thioacetate to furnish the protected thiol **4.30** in almost quantitative yield. Deacetylation with sodium methoxide in methanol gave the desired compound **4.31** in 76%, but over a period of 3 hours, the compound was oxidized to the disulfide **4.32**. The disulfide was reduced back to **4.31** with DTT in 25% yield and due to the propensity for re-oxidation of the thiol, was used immediately.

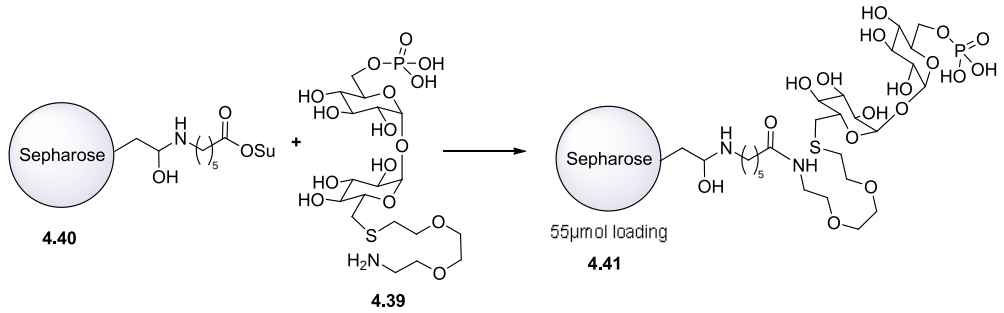
The previous route for the ligation of thiols (Scheme 4.6) was adapted for larger scale syntheses. Instead of coupling with **4.11**, nucleophilic substitution was performed on **4.10** which leads to a more straightforward purification since deprotected phosphates are avoided. Thus, **4.10** was coupled with **4.31** under microwave irradiance at 80˚C for 40 minutes to furnish the desired product **4.33** in a moderate 35%, with the remainder as unreacted starting sugar due to *in situ* oxidation of **4.31** to **4.32** which restricted the maximum attainable yield. With the linker attached to the core trehalose skeleton, subsequent protecting group manipulations could be conducted to give the desired product. Phosphate transesterfication with CsF in methanol minimized formation of the 4-phosphate side-product **4.34** to 32%, giving the desired 6-phosphate **4.35** as the major product, which was acetylated to **4.36** in a healthy 68% yield over the two steps. Subsequent phosphate methoxy deprotection with TMSBr

afforded the free phosphate **4.37**, which was deacetylated without any further purification using standard Zemplen conditions to give **4.38** in an excellent 98% yield over the two steps. The final azide reduction with hydrogen over palladium on carbon was straightforward, giving the target compound **4.39** in near quantitative yield.



Scheme 4.10. *Synthesis of a Tre-6-P analogue capable of immobilization onto a solid phase. Conditions: (i) NaH, THF, microwave, 80˚C, 40mins; (ii) CsF, MeOH, reflux 17h; (iii) Ac_2O, Py, RT; (iv) TMSBr, dioxane, RT; (v) NaOMe, MeOH, RT; (vi) H_2, Pd/C, H_2O.*

For surface immobilization of **4.39**, Sepharose decorated with terminal *N*-hydro succinimide activated esters was used in a pre-packed Hi-Trap® column (GE Healthcare), to allow superior chromatographic extractions (Scheme 4.11).

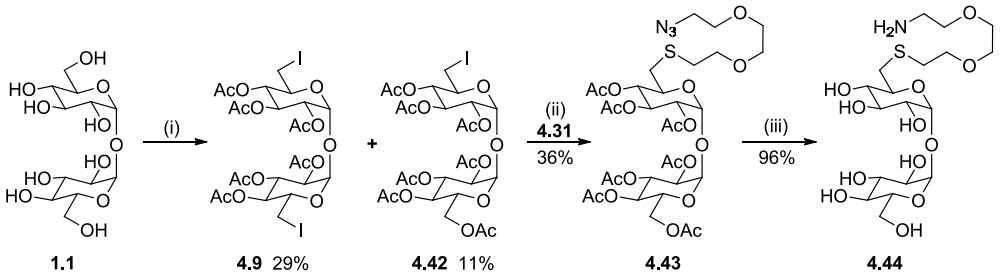


Scheme 4.11. *Reaction to decorate a sepharose resin with Tre-6-P. Conditions: 0.2M $NaHCO_3$, pH 8.3.*

The Sepharose resin **4.40** was coupled with **4.39** under basic conditions to produce a Tre-6-P affinity column (**4.41**), tethered using strong and biostable amide linkages. The extent of ligand loading, determined using the UV absorbance at 220nm of the

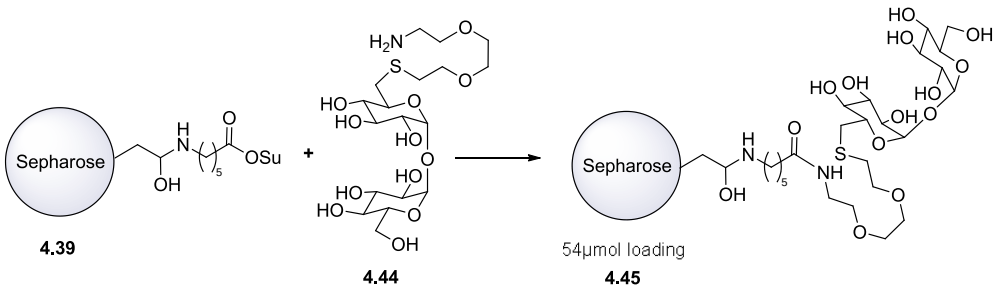
released succinimde group, was found to be 55 μmoles for a 5ml Hi-Trap® column which represents complete modification of all available reactive groups.

An analogous trehalose affinity column was also synthesized for the extraction of trehalose binding proteins from crude lysates and is briefly discussed here, although such an extraction is yet to be performed. Installation of the iodo groups were performed using a three-step one-pot synthesis starting from trehalose (Scheme 4.12). Trehalose (**1.1**) was phosphorylated using DPCP, per-*O*-acetylated and the crude mixture subjected to microwave irradiance with TBAI in THF at 80˚C which resulted in phosphate S_N2 to give the diiodide **4.9** in 29% yield and the desired product **4.42** in 11%. The monoiodide was reacted with **4.31** under basic conditions with microwave assistance to furnish the coupled adduct **4.43** in a 36%, which was deacetylated and reduced with hydrogen to furnish the free amine **4.44** in 96% yield.



Scheme 4.12. *Syntheis of a trehalose analogue for surface thethering onto a solid phase resin. Conditions: (i) DPCP, Py, RT followed by Ac_2O, Py, RT followed by TBAI, THF, microwave, 120˚C, 60mins; (ii) NaH, THF, 80˚C, 40mins; (iii) NaOMe, MeOH followed by H_2, Pd/C, MeOH.*

The trehalose analogue was coupled onto Sepharose as before, to create a 5ml affinity column with a 54 μmol loading of **4.44** (Scheme 4.13).



Scheme 4.13. *Reaction to decorate a sepharose resin with trejhalose. Conditions: 0.2M $NaHCO_3$, pH 8.3.*

4.7 Identification of Tre-6-P binding proteins

With the Tre-6-P affinity column in hand, an extraction of proteins from crude cell lysate could be attempted.[§] To maximize the quantity of isolated proteins, the purification would ideally be scaled to a large volume of plant extract especially since the *in vivo* concentration of the intermediary factor is anticipated to be low. Thus, *Brassica oleracea* (cauliflower), which is much easier to grow on a large scale, was preferred over *A. thaliana*. Partially purified preparations of SnRK1 were obtained from *B. oleracea* extract using standard gel filtration and ion exchange chromatography techniques commonly used in molecular biology. Fractions of SnRK1 that were inhibited by Tre-6-P, and therefore contained the intermediary factor, were pooled for affinity purification with **4.41**. Thus, the partially purified SnRK1 was applied to the column and eluted with an increasing concentration of sodium chloride to gradually release bound proteins from the resin. Gel electrophoresis with silver staining of the late fractions revealed several proteins that were strongly adhered to the Tre-6-P affinity column (Figure 4.13).

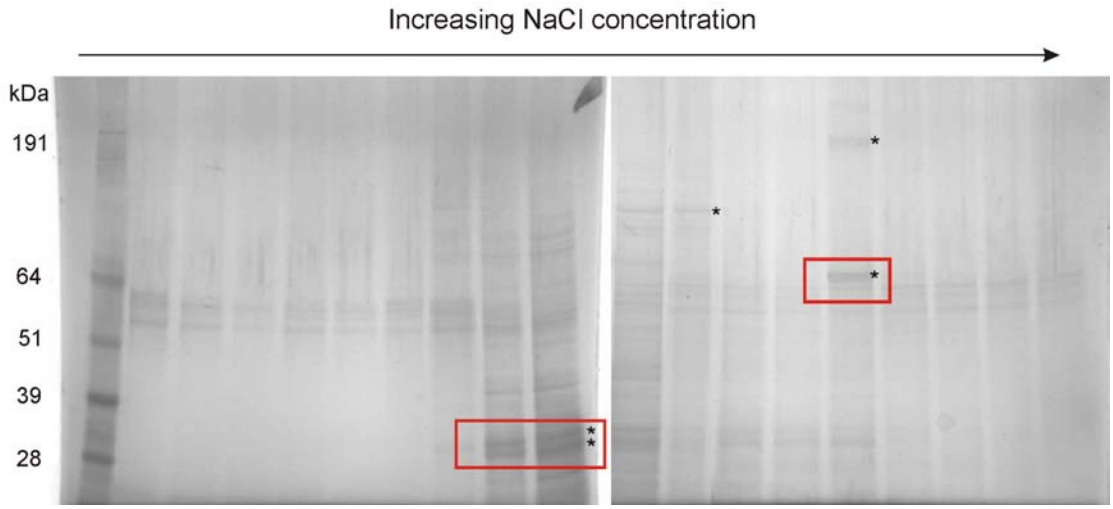


Figure 4.13. *Silver stained gel of fractions eluted from the Tre-6-P affinity purification of partially purified SnRK1. Early fractions of unbound proteins are not shown. * Bands that were excised for proteomics analysis, red boxes indicates a successful identification.*

The strongest of these bands were excised, the gel fragments were destained and proteins were digested with trypsin. The released peptides were concentrated and injected through a C_{18} nano-HPLC column for separation, and subsequently analysed by high resolution mass spectrometry and MS/MS. The peptide mass fingerprint data

[§] All plant extract preparations, protein purifications and Tre-6-P affinity chromatography steps were conducted by Dr L. F. Primavesi at Rothamsted Research, Harpenden, UK.

was compared against the genomes of *A. thaliana*, *B. oleracea* and other green plants using the Mascot search engine.[21] The identification of proteins utilized a probabilistic molecular weight search (MOWSE) score which assesses the likelihood of a particular hit arising purely by chance.[22] Using this method, the constituent proteins from three of the excised bands could be identified. Most interestingly, all three bands were composed of multiple proteins belonging to the 14-3-3 family. Each identified protein was composed of multiple peptide hits with high MOWSE scores indicating greater than 99.9% significance, providing confident certainty in the assignments. For instance, a major component in the band at ~30kDa corresponding to a 14-3-3 isoform was identified with 13 individual peptide hits covering 25% of the protein sequence (Table 4.1).

Observed	Mr(expt)	Mr(calc)	ppm	Miss	Expect	Peptide
335.2026	668.3907	668.3897	1.48	0	0.041	K.VFYLK.M
402.6849	803.3553	803.3524	3.63	0	0.46	R.YMAEFK.S + Oxidation (M)
452.2618	902.5090	902.5073	1.85	0	0.00018	R.IVSSIEQK.E
509.2406	1016.4667	1016.4637	2.93	0	69	R.DQYVYMAK.L
517.2381	1032.4617	1032.4586	3.00	0	0.047	R.DQYVYMAK.L + Oxidation (M)
595.3359	1188.6572	1188.6536	2.99	0	0.0015	K.DSTLIMQLLR.D
603.3332	1204.6519	1204.6485	2.76	0	0.14	K.DSTLIMQLLR.D + Oxidation (M)
671.3406	1340.6666	1340.6646	1.53	1	51	R.KTAAEDTMLAYK.A
679.3400	1356.6655	1356.6595	4.45	1	0.023	R.KTAAEDTMLAYK.A + Oxidation (M)
702.8719	1403.7292	1403.7256	2.60	1	0.0042	R.IVSSIEQKEESR.K
468.9176	1403.7310	1403.7256	3.82	1	2.2	R.IVSSIEQKEESR.K
882.9396	1763.8646	1763.8624	1.25	0	3.5e-05	K.AAQDIAAADMAPTHPIR.L + Oxidation (M)
588.9630	1763.8672	1763.8624	2.69	0	0.00035	K.AAQDIAAADMAPTHPIR.L + Oxidation (M)

```
  1 MAATLGRDQY VYMAKLAEQA ERYEEMVQFM EQLVTGATPA EELTVEERNL
 51 LSVAYKNVIG SLRAAWRIVS SIEQKEESRK NDEHVSLVKD YRSKVESELS
101 SVCSGILKLL DSHLIPSAGA SESKVFYLKM KGDYHRYMAE FKSGDERKTA
151 AEDTMLAYKA AQDIAAADMA PTHPIRLGLA LNFSVFYYEI LNSSDKACNM
201 AKQAFEEAIA ELDTLGEESY KDSTLIMQLL RDNLTLWTSD MQTNQMHHIR
251 DIKEHVKTEI TAKPCVLSYY YSM
```

Table 4.1. *Peptide hits corresponding to the lambda isoform of* A. thaliana *14-3-3. Ppm indicates the difference between the observed and calculated peptide masses, "miss" indicates number of miss trypsin cleavage sites and "expect" indicates the expectation value for the peptide (the number of times a better match would be expected purely by chance). The corresponding sequence is shown below, with matched peptides highlighted in red.*

The 14-3-3 proteins are essential regulatory elements that are ubiquitous across the eukaryotic kingdom.[23] The 14-3-3s complement protein phosphorylation by binding to phosphorylated serine and threonine residues, which can alter the activity and cell localization of interacting proteins. As a result, 14-3-3 proteins are involved a wide

range of biological signalling roles and are conserved across nature. In plants, the 14-3-3s are implicated in various developmental processes such as carbon and nitrogen metabolism, sucrose synthesis and the abscisic acid pathway,[24] and can even interact with TPS5 from *A. thaliana*.[25] Given that these processes are also under the control of the SnRK1 and Tre-6-P signalling mechanisms (section 1.2, page 2), it would be tempting to invoke the 14-3-3s as the intermediary factor.

Each eukaryotic species contains a plethora of protein isoforms belonging to the 14-3-3 class.[23,24] For instance, in *A. thaliana*, 13 variants are known to exist.[26] Importantly, these individual proteins can dimerize resulting in a large number of possible heterodimer permutations which may reflect finely tuned functional specificities.[27] Our affinity purification allowed isolation of both monomers (~30kDa) and dimers (~60kDa, Figure 4.13), and subsequent proteomics analysis revealed the presence of multiple isoforms indicating that any interaction with Tre-6-P is likely to be ubiquitous across all plant 14-3-3s. Moreover, the dimers were more strongly retained by the column which hints at a stoichiometric interaction between Tre-6-P and individual 14-3-3 monomers.

X-ray crystallography experiments have provided detailed structural insights into the binding of 14-3-3 proteins.[28] These have revealed that the 14-3-3s specifically recognise binding motifs of the type RSXpSXP and RX(Y/F)XpSXP which are bound in an amphipathic groove on the protein surface. However, such binding motifs are not present on the *A. thaliana* SnRK1 sequence,[29] which led us to propose a non-conventional, Tre-6-P dependent binding mode between the kinase and the 14-3-3s. We proposed that the interaction of Tre-6-P with the 14-3-3s could induce a conformational change that would allow binding with SnRK1, restricting its activity. Various small molecules have been demonstrated as binding partners for 14-3-3s (such as AMP,[30] the phytotoxin fusicoccin[31] and the lipid sphingosine[32]) and an atypical binding motif has also been proposed lending support to our hypothesis.[33]

4.8 Probing the binding of 14-3-3s with Tre-6-P

To test our hypothesis, the omega isoform of *A. thaliana* 14-3-3 was recombinantly expressed in *E. coli*.[**] The protein was initially expressed with a 20 amino acid extension bearing a His_6 tag to aid purification, but this was removed at a thrombin cleavage site to furnish purified native protein suitable for analytical studies.

Isothermal titration calorimetry (ITC) is a widely used technique for determining the thermodynamic binding parameters of biomacromolecules in their native solution phase.[34,35] In this method, a ligand is titrated into a chamber containing a solution of the protein, and the power required to maintain a constant temperature is measured, providing a direct readout of the energy changes occurring on complex formation. Various ITC experiments were attempted with the recombinant 14-3-3 titrated against Tre-6-P, but in all cases the expected binding interactions were not detected and data was instead dominated by mixing artefacts (Figure 4.14).

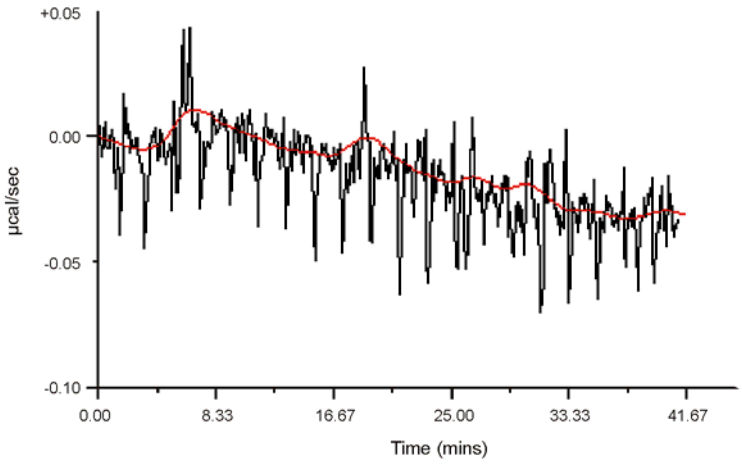


Figure 4.14. *Representative trace of raw data from an ITC experiment with Tre-6-P (2eq) titrated into 14-3-3 omega (20µM, 1eq).Data shows only mixing aretfacts and no evidence of binding.*

Since these artefacts could obscure any underlying data, alternative methods for investigation the binding interaction were sought. The binding of ligands can affect the structural stability of a protein, altering its characteristic "melting temperature" (T_m) representative of the point at which denaturation and subsequent aggregation occurs.[36] Measurement of this parameter can provide information on the extent of ligand binding. Circular dichroism (CD) spectroscopy is an effective method for

[**]14-3-3 protein expression and purification was conducted by Dr L. F. Primavesi at Rothamsted Research, Harpenden, UK.

determining the secondary structure of a protein,[37] and when performed over an increasing temperature range, can be used to establish the T_m by monitoring the gradual loss of distinguishing structural features.[35,38] A CD spectrum of the 14-3-3 omega isoform at room temperature shows two minima at 210 and 225nm which are characteristic of α helices (Figure 4.15).

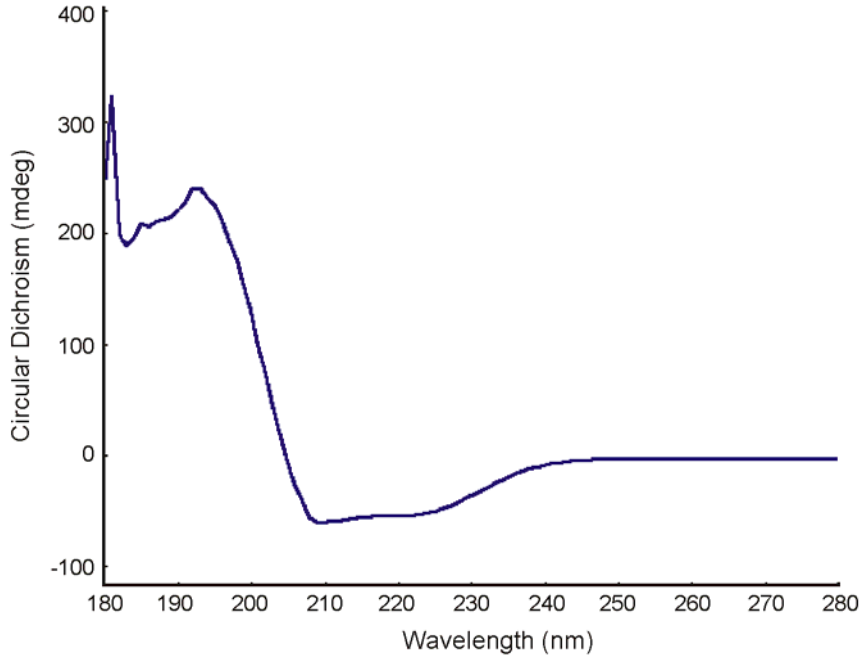


Figure 4.15. *CD spectra of 14-3-3 omega isoform at RT demonstrated folded structure dominated by α helices.*

These peaks can be used to follow the structure of the protein as a function of increasing temperature. Thus, at room temperature, where 95% of the protein is composed of α helices (as determined by deconvolution of the above data), the peaks would be at their strongest. As the temperature is increased beyond T_m, the protein would unfold to a random coil and eventually aggregate with the circular dichroism tending towards zero. Such experiments were performed on the recombinant 14-3-3 with different concentrations of Tre-6-P to explore the effect of the sugar on the thermodynamic stability of the protein (Figure 4.16).

Interestingly, it was found that the T_m decreased with increasing concentrations of Tre-6-P. If the protein bound to the ligand, an increase in T_m would be expected as the folded state is stabilized with respect to the transition state for unfolding (Figure 4.17).[36] Generally, a decrease in T_m results from the non-specific binding of ligands to the *unfolded* state. Such a result does not exclude binding to the bound state, but simply demonstrates that interactions with the unfolded state are dominant.

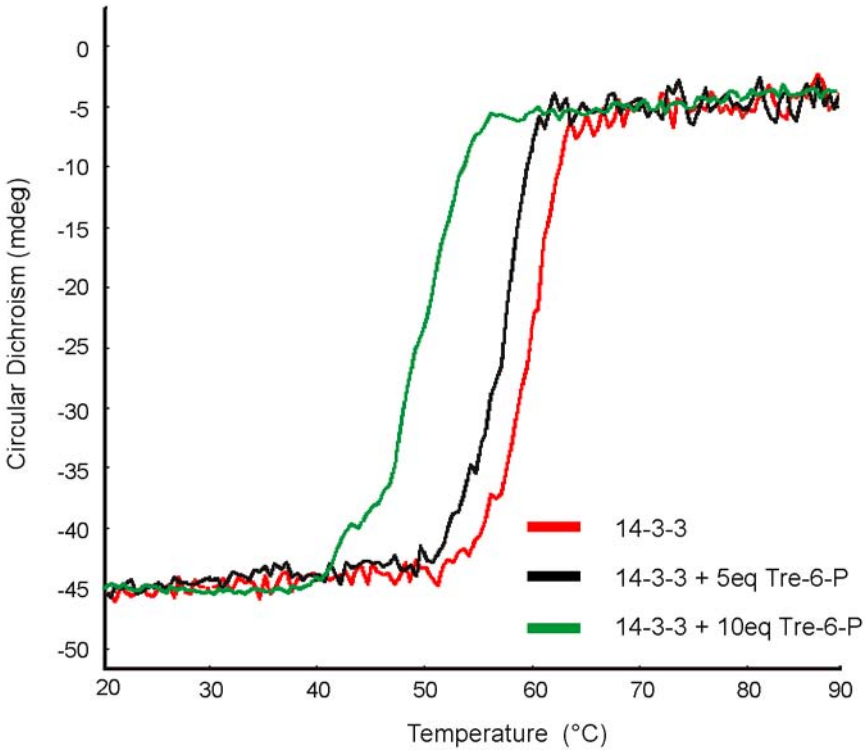


Figure 4.16. *CD of 14-3-3 omega isoform at 210nm against increasing temperature. Data shows that an increasing concentration of Tre-6-P leads to a decrease in* T_m.

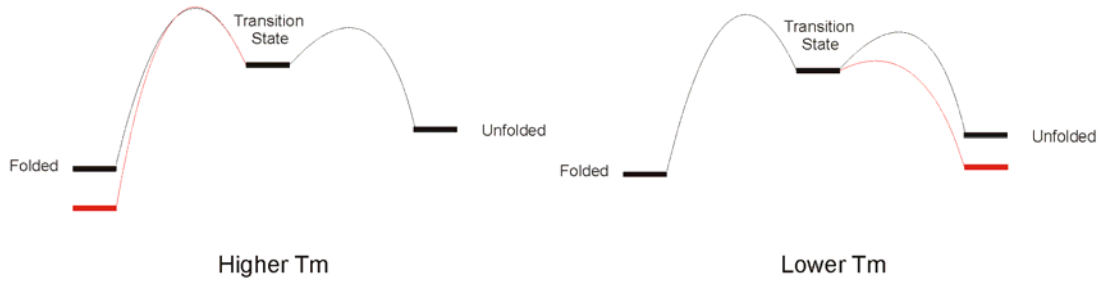


Figure 4.17. *Energy level diagrams for the process of protein denaturation illustrating outcomes from the binding of ligands with the folded state (A), the unfolded state (B), or the transition state (C).*

Due to the inconclusive nature of the CD experiments, an alternative technique was required. The saturation transfer difference (STD) NMR experiment is ideally suited to the study of weak protein-ligand interactions and for this reason, is widely used in the pharmaceutical industry for drug discovery.[39] In this technique, a pulse is applied to specifically excite protons on the protein, from which energy is transferred onto bound ligands via the nuclear Overhauser effect resulting in an enhanced NMR signal. Moreover, the enhancement of ligand proton signals is correlated with their proximity to the protein, and in this way, STD NMR facilitates direct measurement of the binding epitope.[40] STD NMR experiments were attempted with the recombinant

14-3-3 and Tre-6-P under a variety of conditions, but sadly, enhancements were not observed indicating no binding (Figure 4.18).

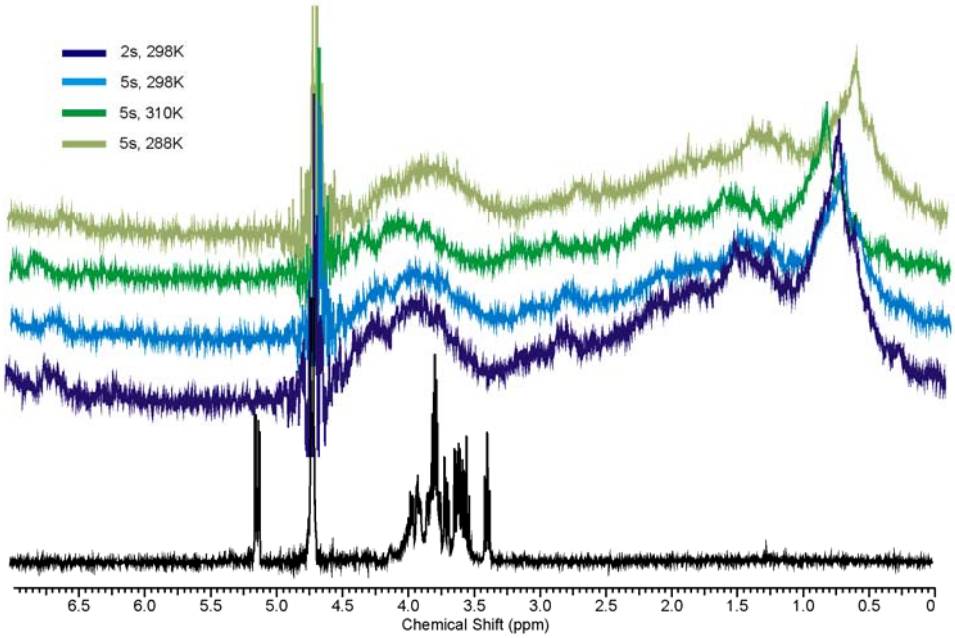


Figure 4.18. *STD NMR traces of Tre-6-P incubated with recombinant 14-3-3 omega isoform, with different saturation times and temperatures. Data shows only the background signal from the protein and no enhancements for the ligand indicating lack of binding.*

While no binding of Tre-6-P was detected, we unexpectedly detected the formation of other compounds over the course of the experiments (Figure 4.19).

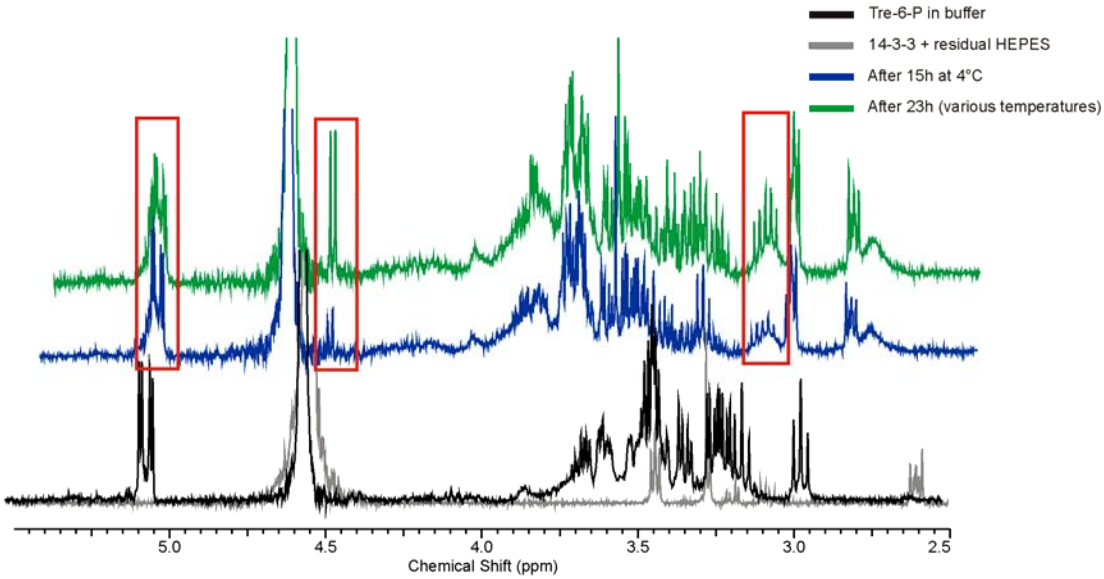


Figure 4.19. *^{1}H NMRs before and after STD NMR experiments. Samples were stored at 4˚C for 15h followed by 8h at RT. Red boxes highlight the appearance of new peaks. Data shows the formation of a new product over time.*

A detailed analysis of the product peaks indicated the presence of glucose and Glc-6-P, which likely derive from the degradation of Tre-6-P (Figure 4.20).

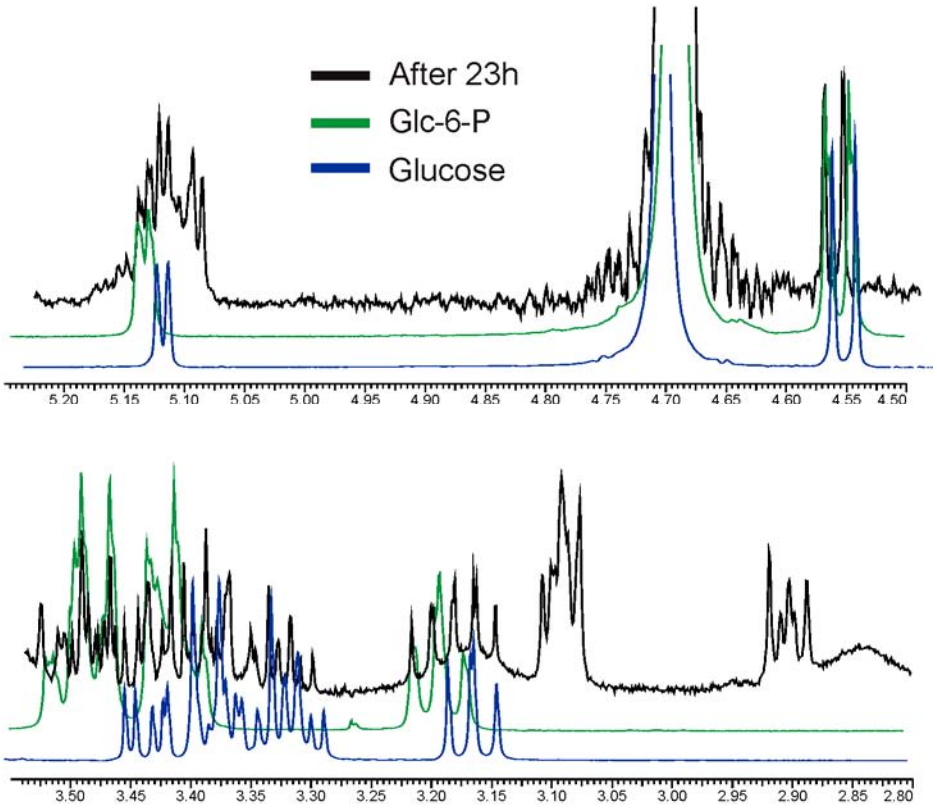


Figure 4.20. *Comparison of the ^{1}H NMR of crude reaction mixture with Glc-6-P and Glc in 20mM phosphate and 130mM NaCl at pH* 7.5. Data shows that new peaks can be attributed to glucose and Glc-6-P.*

The conversion of Tre-6-P to Glc-6-P and glucose was further demonstrated by LC-MS of the crude reaction mixture, which conclusively showed the consumption of Tre-6-P and the accumulation of Glc-6-P (Figure 4.21).

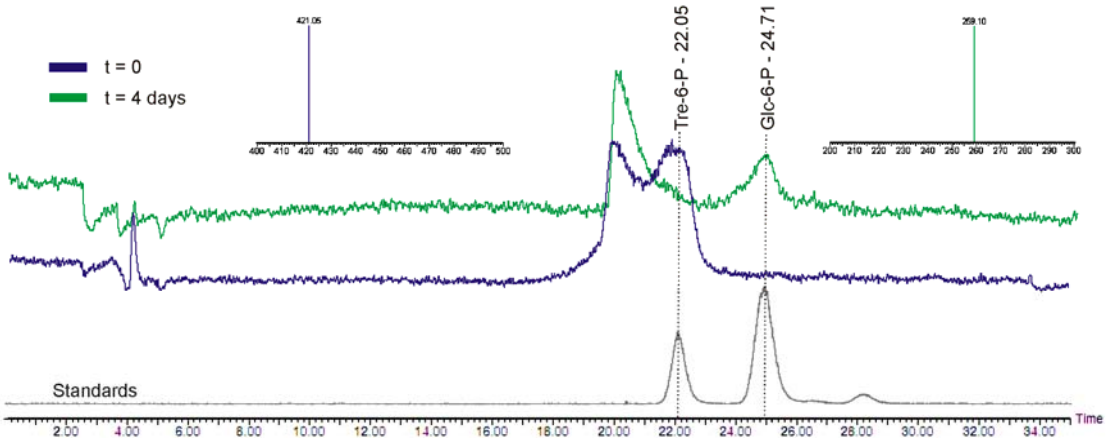


Figure 4.21. *SAX LC-MS ion trace of crude reaction mixture. Inset: combined mass spectra of the Tre-6-P and Glc-6-P peaks of the chromatogram. Data shows the consumption of Tre-6-P and formation of Glc-6-P over time in presence of recombinant 14-3-3omega isoform protein.*

The enzymatic activity of a 14-3-3 protein would be an unprecedented result and it was therefore vital to determine whether the degradation of Tre-6-P was attributable to the recombinant omega isoform, or another impurity within the mixture. An SDS-PAGE gel illustrated that the sample was pure at the Coomassie blue level of sensitivity (Figure 4.22). Furthermore, this data also excluded thrombin, a possible impurity from the His_6 cleavage step, as a contaminant.

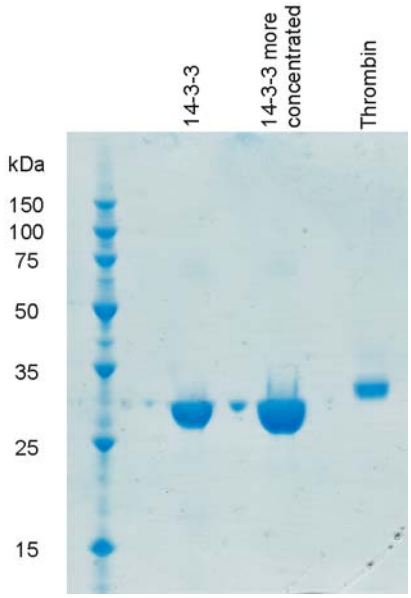


Figure 4.22. *SDS-PAGE gel (10% bis-tris) run at 220V using MOPS running buffer. Data shows that expressed 14-3-3 omega isoform appear pure and free of thrombin as an impurity.*

However, the more sensitive silver stain showed the presence of higher mass impurities that were present in the early His_6 tagged preparation and was carried through after thrombin cleavage (Figure 4.23). Interestingly, the band for the 14-3-3 showed a sizeable gel shift after room temperature exposure to Tre-6-P. Mass spectrometric analysis revealed mass reduction of 515 Da corresponding to the cleavage of the EEQQ sequence at the C terminus (section 4.11.5, page 197).

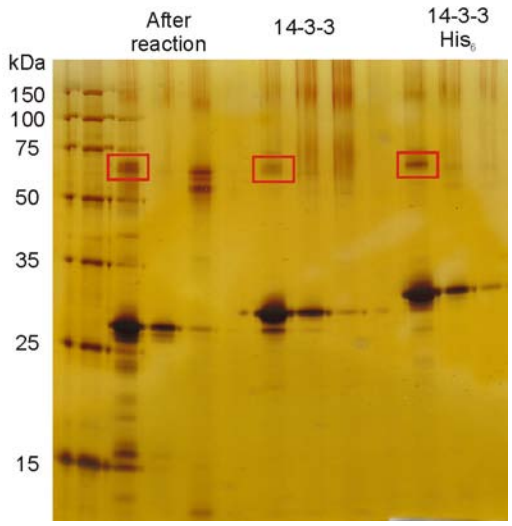


Figure 4.23. *SDS-PAGE gel (10% bis-tris) run at 220V using MOPS running buffer. Each sample was run thrice with different loadings: 10μg,1μg and 0.1μg. Red boxes highlight impurities. Data shows that 14-3-3 isoform contains an impurity using a high sensitivity silver stain. After reaction with Tre-6-P, a gel shift is seen in the protein indicating degradation of the 14-3-3.*

We conjectured that such an anomalous result was likely caused by contamination with a protease, which may also account for the unprecedented Tre-6-P hydrolysis. To test this, high mass impurities were removed by spin filtration (Figure 4.24).

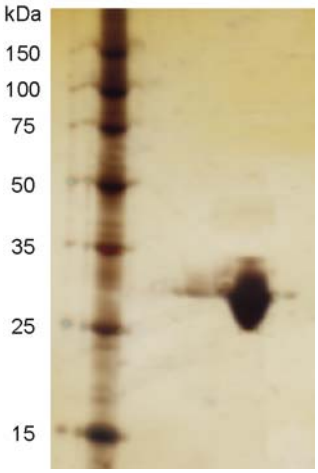


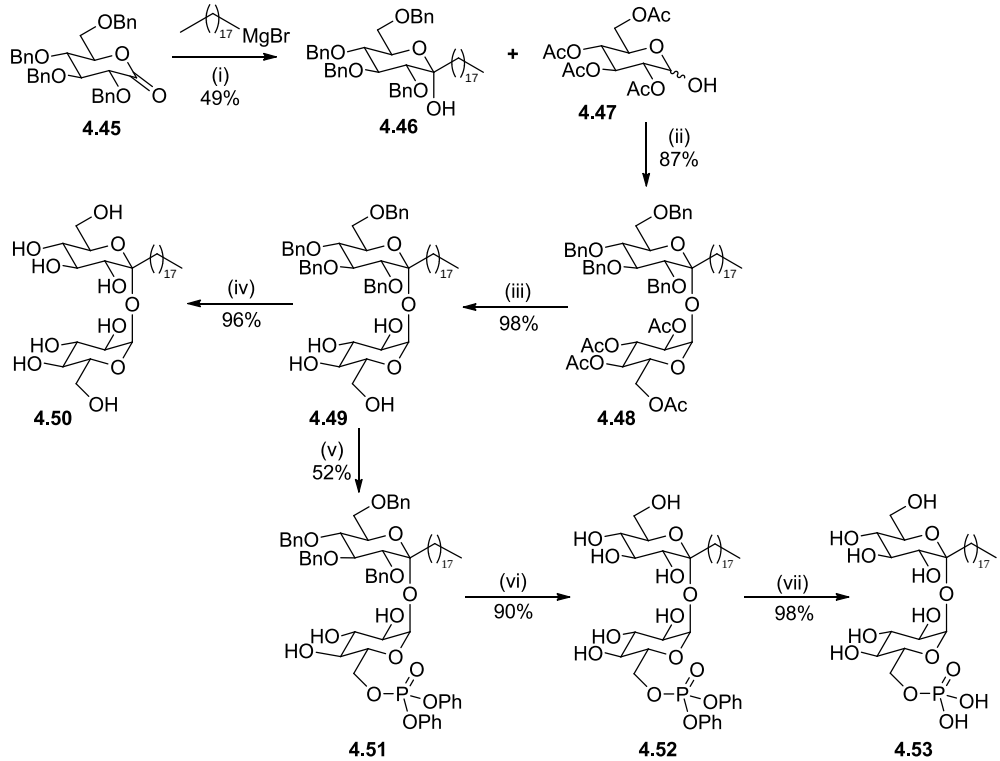
Figure 4.24. *SDS-PAGE gel (10% bis-tris) run at 220V using MOPS running buffer. Data shows 14-3-3 protein after multiple spin filtrations is purer and free of the original impurity.*

The purified sample showed essentially no activity with Tre-6-P. Thus, the recombinant omeaga isoform of the *A. thaliana* 14-3-3 protein appears to have no interaction with Tre-6-P under the tested conditions.

4.9 Tre-6-P microarrays

To complement the Tre-6-P affinity extraction we explored the use of microarray technology as an alternative method for identifying the intermediary factor. Such methodology, which is derived from analogous RNA microarray techniques, utilizes surface-immobilized carbohydrates to screen against multiple fluorescently tagged proteins.[41] In this way, a large number of potential targets can be easily probed.

There are various approaches to the immobilization of carbohydrates, but a common method involves the use of alkyl "lipid tags" for irreversible adherence onto nitrocellulose membranes. In such an approach, the orientation of carbohydrate display would be an important parameter dictating the binding to proteins. The octane tagged Tre-6-P analogue **4.20** is well suited for microarray screening, but an alternative site of lipid attachment would allow complete exploration of different presentation geometries. Thus, we envisioned the synthesis of a C1'-lipid tagged Tre-6-P analogue based on ketoside glycosylations (Scheme 4.14).[42]



Scheme 4.14. *Sythesis of Tre-6-P and trehalose analogues for surface tethering onto nitrocellulose microarray plates.Conditions: (i) Et_2O, -78˚C to RT, 2h; (ii) TMSOTf, DCM, 3Å molecular sieves,-78˚C, 30mins; (iii) NaOMe, MeOH; (iv) H_2, $Pd(OH)_2$, alumina, EtOH, 1 week; (v) DPCP, Py, DMAP; (vi) H_2, Pd/C, MeOH, 70 Bar, 25˚C; (vii) H_2, PtO_2, MeOH + 0.5% AcOH.*

A Grignard reaction with octadecylmagnesium bromide on the ketoside **4.45**[††] allowed lipid attachment at the C1 position to furnish **4.46** in a respectable 49% yield, which was subsequently glycosylated with **4.47**[††] using TMSOTf to give the desired α,α linked trehalose adduct **4.48** in an excellent 87%. Deacetylation under standard Zemplen conditions proceeded without any problems to afford **4.49**, which could be deprotected under hydrogen over palladium hydroxide to furnish a lipid-tagged trehalose analogue **4.50** that could itself be used in microarray screens. Alternatively, phosphorylation of the free primary hydroxyl with DPCP produced **4.51** in a reasonable 52% yield. The subsequent debenzylation was performed using flow hydrogenation at high pressure, which allowed facile conversion to **4.52** in 90% yield. Finally, phosphate deprotection with hydrogen over PtO_2 furnished the target compound **4.53** in an almost quantitative 98% yield.

With the successful synthesis of **4.20** and **4.53**, a platform for microarray studies on Tre-6-P binding proteins can be established which would likely provide novel insights into the biology of this interesting sugar.

4.10 Conclusions

The complex signalling and metabolic pathways operating in plants provide an opportunity for discovery and innovation within the field. However, the same complexity can hamper research efforts necessitating cautious experimental design. For instance, we have demonstrated how the plethora of active enzymes within crude cell lysate can produce results that can easily be misinterpreted. However, modern analytical chemistry provides tools that allow robust investigations of such complicated systems. With this approach, Tre-6-P has been conclusively confirmed as a vital signalling molecule within plants.

This rigorous chemical approach can be extended further to study the mechanism of action for Tre-6-P. Thus, through the use of Tre-6-P analogues, a binding epitope map was determined, from which an affinity column was tailored for isolation of the intermediary factor. This has revealed the 14-3-3 class of proteins as a potential lead which, if correct, would elegantly tie together various aspects of plant development.

[††] Kindly donated by Keriann M. Backus, University of Oxford, Oxford, UK.

Our attempts at proving this interaction were unsuccessful, but we believe that this may be attributable to the sheer complexity of the 14-3-3 regulatory system which cannot be replicated *in vitro*. For instance, the binding of certain 14-3-3s is critically dependent on a post-translational phosphorylation,[43] which would be difficult to imitate for the recombinant protein. The presence of divalent ions which is known to modulate the behaviour of 14-3-3 would be another parameter for investigation.[44] Moreover, the proteomics approach is inadequate for resolving closely homologous protein sequences, and therefore we have no clear understanding of the specificity of Tre-6-P binding to different 14-3-3 isoforms. Of the 91 possible 14-3-3 permutations for *A. thaliana*, we have only investigated one and the expression of other isoforms would therefore be worthwhile.

Immunoprecipitation experiments, where the 14-3-3s could be selectively removed from cell lysates may prove valuable for determining their role in SnRK1 inhibition. Alternatively, a larger scale Tre-6-P affinity extraction would allow determination of other proteins that bind to the sugar. In any event, there is a strong case for continued efforts in this direction.

4.11 Experimental

For general experimental conditions see section 2.8.1 page 44.

The stereochemistry at the phosphorous centre for compounds **4.46**, and **4.48 - 4.53** were determined using NOESY and/or NOE NMR experiments.

4.11.1 Metabolism experiments on exogenous sugars

General reaction protocol for Tre-6-P, Glc-1-P and Glc-6-P

Plant extract was prepared from 3 days old *A. thaliana* seedling by Dr Lucia Primavesi (Rothamsted Research). Plant extract was stored at -80°C and allowed to thaw on ice and used immediately. Reactions were performed with 50µl of a 50mM stock solution of carbohydrate and 450µl of plant extract. Reactions were further diluted with 2.0ml of 0.04M HEPES, 5.0mM $MgCl_2$, 4.0mM DTT and 0.2mM ATP at pH 7.5. The reaction mixtures were incubated at 30°C for either 10 or 30 minutes. The reaction mixtures were quenched by heating to 95°C for 5 minutes and

centrifuged at 3000 rpm for 10 seconds to remove denatured protein debris. The supernatant was retained for further analysis. Mass spectrometric analysis was conducted directly on this solution. For analysis by ^{31}P NMR, the reaction mixtures were concentrated by lyophilization and subsequently dissolved in 400µl of water and analysed as below. For positive mode analysis, 50µl of the supernatant was dissolved in pyridine (0.5ml) and acetic anhydride (0.5ml). The reaction mixtures were allowed to stir at room temperature for 5 hours after which was quenched with methanol (5ml). The solvents were removed *in vacuo* and the crude samples dissolved in methanol (50µl) and analysed by LC-MS as mentioned below.

^{31}P NMR analysis procedure

All NMRs were conducted in water. Prior to analysis, 50µl of D_2O was added as a deuterium lock. ^{31}P were referenced to $PO(OMe)_3$ ($\delta = 3.16$ ppm), which was added as an internal standard.

LC-MS analysis procedure (negative mode)

Samples were analysed through a Waters Spherisorb strong anion exchange column (250 x 4.6mm, 5µm). A gradient was applied from water (pH 7) to water + 10% formic acid (pH 2) over 30 minutes at a flow rate of 1.0ml/min. Eluants were detected using a Waters Micromass ESI mass spectrometer in negative mode. The mass spectrometer was calibrated against the NaF cluster ion series. With a desolvation flow rate = 800 L/hr, nebuliser flow rate = 100 L/hr, cone voltage = 50V, capillary = 2800V, extraction = +3V, RF lens = 200V, desolvation temperature = 200°C and source = 150V, a 10mM solution of NaF in 9 MeCN : 1 H_2O was injected at a flow rate of 10µl/min. The corresponding ion series was used for mass spectrometer calibration.

LC-MS analysis procedure (positive mode)

Samples were analysed through a Phenomenex Synergi Hydro C_{18} column (150 x 4.6mm, 4µm). A gradient was applied from water + 0.1% formic acid to acetonitrile + 0.1% formic acid over 30 minutes at a flow rate of 1.0ml/min. Eluants were detected using a Waters Micromass ESI mass spectrometer in positive mode. The mass spectrometer was calibrated against the myoglobin ion series. With a desolvation flow rate = 800 L/hr, nebuliser flow rate = 100 L/hr, cone voltage = 50V, capillary =

2800V, extraction = +3V, RF lens = 200V, desolvation temperature = 200°C and source = 150V, a 0.5mg/ml solution of myoglobin in 1 MeCN : 1 H_2O was injected at a flow rate of 10µl/min. The corresponding ion series was used for mass spectrometer calibration.

General reaction protocol for Rib-5-P

Plant extract was prepared by Dr Lucia Primavesi (Rothamsted Research) from (i) 3 days old *Arabidopsis thaliana* seedling, (ii) growing whole *Triticum aestivum* (iii) growing, post-anthesis *T. aestivum* endosperm, (iv) whole *N. tobaccum,* or (v) *N. tobaccum* engineered with PRK antisense gene.[17] Reactions were performed according to the procedures described by *Piattoni et al.*[13] Plant extract was stored at -80°C and allowed to thaw on ice and used immediately. Reactions contained a final concentration of 0.04M HEPES, 5.0mM $MgCl_2$, 4.0mM DTT, 0.2mM ATP, 0.5µM okadaic acid and protease inhibitor cocktail (1x concentration) at pH 7.5. Reactions were performed in a total volume of 500µl. To a vial containing 400µl of the above components (with concentrations adjusted for subsequent dilution to 500µl) was added a 50mM solution of Rib-5-P (10µl) and crude plant extract (90µl) in quick succession. The reaction mixtures were incubated at 30°C for exactly 6 minutes. The reaction mixtures were snap frozen in liquid nitrogen and lyophilized. The resultant solid was dissolved in water (10µl) and the entire solution injected for UV-HPLC analysis as below. The reaction was repeated for LC-MS analysis.

Control reactions were run as above without Rib-5-P, and without Rib-5-P with ADP instead of ATP.

Detection of ATP, ADP and AMP

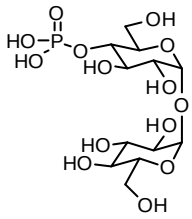
HPLC was conducted on a Dionex UltiMate 4000. Samples were analysed through a Waters Spherisorb strong anion exchange column (250 x 4.6mm, 5µm). A gradient was applied from 40mM sodium phosphate at pH 4.5 to 500mM sodium phosphate buffer at pH 2.5 over 10 minutes and then eluted for a further 5 minutes, all at a flow rate of 1.0ml/min. Eluants were detected using an in-line UV absorbance detector at λ = 260nm.

Detection of Rib-5-P and ribulose-1,5-biphosphate

HPLC was conducted on a Waters binary HPLC system. Samples were analysed through a Hichrom SiELC Primesep SM mixed mode anion exchange/C_{18} reverse phase column (150 x 4.6mm, 5μm) with an applied gradient from 10mM ammonium formate at pH 3.0 to 80mM ammonium formate at pH 3.0 over 15 minutes at a flow rate of 1.0ml/min. Eluants were fed directly[‡‡] into a Waters Quattro micro in negative mode either operating in Selected Ion Recording mode centred at the monoisotopic masses of Rib-5-P and ribulosw-1,5-biphosphate with a detection width of 0.5Da and a dwell time of 10ms (for quantitative applications). The mass spectrometer was operated with a cone voltage of 35V, a source temperature of 100°C and desolvation temperature of 400°C. All chromatograms are shown after smoothing.

4.11.2 Tre-6-P library synthesis

SnRK1 activity assays were performed by Dr Lucia Primavesi (Rothamsted Research) following the procedure described by Zhang *et al.*[9] Briefly, assays were conducted in 0.04M HEPES, 5.0mM $MgCl_2$ and 4.0mM DTT at pH 7.5, with 1.0mM of the Tre-6-P analogue, 0.2mM of γ ^{32}P-ATP and 0.2mM of the AMARA peptide sequence. Phosphorylated peptide was isolated by absorption onto Whatman P81 phosphocellulose paper and conversion measured by scintillation counting.



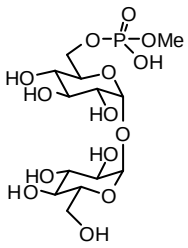
4.3. D-trehalose-4-phospate[45]

Compound **3.37** (6.5mg, 0.014mmol, 1eq) was suspended in dry dioxane (1ml) and briefly sonicated for 5 minutes. TMSI (20μl, 0.14mmol, 10eq) was added to this mixture and stirred at RT for 30 minutes, after which mass spectrometry (ES^-) showed only the desired compound. The reaction mixture was quenched with water (1ml) and the solvent removed under reduced pressure. The crude product was taken up in water (3ml) and washed with ethyl acetate (3 x 1ml). The aqueous layer was concentrated *in vacuo* and purified by HPLC with an Applied Biosystems Poros HQ50 strongly basic

[‡‡]To minimize the amount of salts being fed into the mass spectrometer, either: (a) the spectrometer can be kept parallel to the main flow path with a splitter valve ensuring only a minimal volume is diverted to the detector thus sacrificing signal to noise for long term machine health, or (b) once the expected retention times are determined with reference samples, the in-built divert valve can be utilised only at the correct time points to deliver *all* eluants to the mass spectrometer thus maximizing signal to noise.

anion exchange cartridge. A gradient from 0mM to 500mM ammonium bicarbonate was applied to the column as the mobile phase and the desired compound was eluted at approximately 75mM NH_4CO_3. Repeated lyophilization afforded the title compound as a white amorphous solid (5.2mg, 81%).

R_f 0.2 (5 ethanol: 3 NH_4OH : 1 water), $[\alpha]_D^{25}$ +16.8 (c = 0.41, H_2O); ^1H NMR (500MHz, D_2O) δ ppm 3.35 (1H, t, $J_{H3'\text{-}H4'}$ 9.6Hz, $J_{H4'\text{-}H5'}$ 9.6Hz, H4'), 3.56 (1H, dd, $J_{H2'\text{-}H3'}$ 9.6Hz, $J_{H1'\text{-}H2'}$ 3.8Hz, H2'), 3.64 (1H, dd, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.75-3.91 (5H, m, H5, H6a, H6b, H6'a,H6'b), 3.80 (1H, t, $J_{H2'\text{-}H3'}$ 10.6Hz, $J_{H3'\text{-}H4'}$ 10.6Hz, H3'), 3.81 (1H, q, $J_{H3\text{-}H4}$ 9.1Hz, $J_{H4\text{-}H5}$ 9.1Hz, $J_{H4\text{-}31P}$ 9.1Hz, H4), 3.97 (1H, t, $J_{H2\text{-}H3}$ 9.1Hz, $J_{H3\text{-}H4}$ 9.1Hz, H3), 3.90 (1H, ddd, $J_{H5'\text{-}H6'a}$ 10.3Hz, $J_{H4'\text{-}H5'}$ 9.7Hz, $J_{H5'\text{-}H6'b}$ 2.8Hz, H5), 5.10 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 5.12 (1H, d, $J_{H1'\text{-}H2'}$ 3.5Hz, H1'); ^{13}C NMR (126MHz, D_2O) δ ppm 60.5 (1C, C6'), 61.1 (1C, C6), 70.1 (1C, C4'), 71.3 (1C, C2'), 71.3 (1C, C2), 71.6 (1C, C3), 71.9 (1C, C5'), 72.0 (1C, C5), 72.8 (1C, C3'), 73.3 (1C, d, $J_{C\text{-}31P}$ 9.5Hz, C4), 93.5 (1C, C1'), 93.5 (1C, C1); FT-IR (KBr disc) υ 1178 (P=O), 3443 br (OH); HRMS (ES^-) m/z 421.0751 $[M - H]^-$ (required 421.0753).



4.4. 6-*O*-monomethoxyphosphoryl-D-trehalose

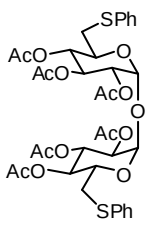
Compound **3.38** (11mg, 0.025mmol, 1eq) was suspended in dioxane and briefly sonicated for 5 minutes. TMSBr (33µl, 0.25mmol, 10eq) was added to this mixture at RT. The reaction was monitored by mass spectrometry (ES^-) and after 3 hours, the two deprotected analogues were detected. After a further 2 hours, the composition of the reaction mixture had not substantially changed according to mass spectrometry intensities. The reaction was quenched by the addition of water (1ml) and the solvents removed *in vacuo*. The crude mixture was taken up in water (3ml) and washed with ethyl acetate (3 x 1ml). The aqueous layer was concentrated *in vacuo* and the products were separated by HPLC with an Applied Biosystems Poros HQ50 strongly basic anion exchange cartridge. A gradient from 0mM to 500mM ammonium bicarbonate was applied to the column as the mobile phase. **4.4** was not bound to the column and was immediately eluted whereas Tre-6-P was eluted at approximately 75mM NH_4CO_3. Repeated lyophilization gave **4.4** (5.7mg, 52%) and Tre-6-P (1.2mg, 11%) as white amorphous solids.

R_f 0.1 (1 water : 2 isopropanol : 2 ethyl acetate), $[\alpha]_D^{25}$ + 38.8 (c = 0.24, H_2O); ^1H NMR (500MHz, D_2O) δ ppm 3.36 (1H, t, $J_{H3'\text{-}H4'}$ 9.3Hz, $J_{H4'\text{-}H5'}$ 9.3Hz, H4'), 3.45 (1H, t, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, H4), 3.52 (3H, d, $J_{H\text{-}31P}$ 11.0Hz, OMe), 3.57 (1H, dd, $J_{H2'\text{-}H3'}$ 8.3Hz, $J_{H1'\text{-}H2'}$ 3.3Hz, H2'), 3.59 (1H, dd, $J_{H2\text{-}H3}$ 9.3Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.68 (1H, dd, $J_{H6'a\text{-}H6'b}$ 11.2Hz, $J_{H5\text{-}H6'b}$ 3.9Hz, H6'b), 3.72 (1H, m. H5'), 3.76 (2H, t, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H3\text{-}H4}$ 9.5Hz, H3, H3'), 3.76 (1H, m, H6'a), 3.96 (1H, m, H5), 4.01 (2H, m, H6), 5.11 (1H, d, $J_{H1'\text{-}H2'}$ 4.1Hz, H1'), 5.12 (1H, d, $J_{H1\text{-}H2}$ 4.4Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 52.9 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, OMe), 60.5 (1C, C6'), 64.1 (1C, d, $J_{C\text{-}31P}$ 4.8Hz, C6), 69.2 (1C, C4), 69.7 (1C, C4'), 70.9 (1C, C2), 71.0 (1C, C2'), 71.7 (1C, C5), 72.1 (1C, C5'), 72.4 (1C, C3), 72.4 (1C, C3'), 93.3 (1C, C1), 93.4 (1C, C1'); FT-IR (KBr disc) υ 1137 (P=O), 3440 br (OH); HRMS (ES^-) m/z 435.0923 $[M - H]^-$ (required 435.0909).

4.5. 6,6'-dithiophenyl-2,2',3,3',4,4'-acetyl-D-trehalose

4.6. 6-O-(diphenoxyphosphoryl)-6'-thiophenyl-2,2',3,3',4,4'-acetyl-D-trehalose

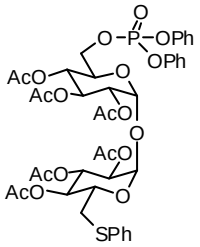
To a suspension of NaH (60% dispersion in mineral oil, 19mg, 0.47mmol, 1eq) in anhydrous THF (1ml) was added dropwise at room temperature thiophenol (48µl, 0.47mmol, 1eq). Once hydrogen production had ceased, the reaction was stirred for a further 30 minutes and a solution of **3.16** (0.50g, 0.47mmol, 1eq) in THF (4ml) was added. The reaction mixture was irradiated in a microwave cavity to 150°C at a maximum power of 300W and maximum pressure of 300Psi with a stream of cooling nitrogen for 2 hours, after which TLC (2 ethyl acetate : 1 petrol) showed the formation of two products. The solvent was removed *in vacuo* and the crude residue taken up in ethyl acetate (25ml) and washed with water (25ml) and brine (25ml), dried over $MgSO_4$ and concentrated *in vacuo*. The compounds were separated by automated silica gel chromatography (12 – 100% ethyl acetate / petrol over 10 column volumes) to give **4.6** (98mg, 27%) as a colourless oil and **4.5** (127mg, 29%) as a white crystalline solid.



*Compound **4.5***

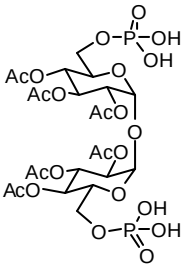
R_f 0.63 (2 ethyl acetate : 1 petrol), $[\alpha]_D^{25}$ +94.0 (c = 1.0, $CHCl_3$), m.p. 135-137°C; ^1H NMR (500MHz, $CDCl_3$) δ ppm 2.02, 2.03, 2.06 (3 x 6H, 3 x s, 6 x OAc), 3.02 (2H, dd, $J_{H6\text{-}H6'}$ 13.4Hz, $J_{H5\text{-}H6'}$ 2.8Hz, H6'), 3.08 (2H, dd, $J_{H6\text{-}H6'}$ 13.4Hz, $J_{H5\text{-}H6}$ 8.5Hz, H6), 4.14 (2H, td, $J_{H4\text{-}H5}$ 9.4Hz,

$J_{H5\text{-}H6}$ 9.4Hz, $J_{H5\text{-}H6'}$ 2.7Hz, H5), 5.03 (2H, t, $J_{H3\text{-}H4}$ 9.6Hz, $J_{H4\text{-}H5}$ 9.6Hz, H4), 5.16 (2H, dd, $J_{H2\text{-}H3}$ 9.9Hz, $J_{H1\text{-}H2}$ 4.1Hz, H2,) 5.19 (2H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 5.47 (2H, t, $J_{H2\text{-}H3}$ 9.6Hz, $J_{H3\text{-}H4}$ 9.6Hz, H3), 7.18 (2H, tt, $J_{meta\text{-}para}$ 6.6Hz, $J_{ortho\text{-}para}$ 2.2Hz, ArH_{para}), 7.23-7.31 (8H, m, ArH); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.7, 20.7, 20.8, (3 x 2C, 6 x $\underline{C}H_3CO$), 35.6 (2C, C6), 69.2 (2C, C2), 70.2 (4C, C3, C5), 71.9 (2C, C4), 91.8 (2C, C1), 126.3 (2C, ArC_{para}), 128.9 (4C, ArC_{ortho}), 129.0 (4C, ArC_{meta}), 136.2 (2C, ArC_{ipso}), 169.6, 169.7, 170.0 (3 x 2C, 6 x $CH_3\underline{C}O$); FT-IR (KBr disc) υ 1759 (C=O); HRMS (MS^+) m/z 801.1864 $[M + Na]^+$ (required 801.1857).



*Compound **4.6***

R_f 0.52 (2 ethyl acetate : 1 petrol), $[\alpha]_D^{25}$ +101.2 (c = 1.0, $CHCl_3$); ^1H NMR (500MHz, $CDCl_3$) δ ppm 2.01, 2.02, 2.03, 2.03, 2.04, 2.06 (6 x 3H, 6 x s, 6 x OAc), 3.02 (1H, dd, $J_{H6'a\text{-}H6'b}$ 13.6Hz, $J_{H5'\text{-}H6'b}$ 3.2Hz, H6'b), 3.07 (1H, dd, $J_{H6'a\text{-}H6'b}$ 13.6Hz, $J_{H5'\text{-}H6'a}$ 8.2Hz, H6'a), 4.07-4.15 (2H, m, H5, H5'), 4.22 (1H, ddd, $J_{H6a\text{-}H6b}$ 11.3Hz, $J_{H\text{-}31P}$ 6.6Hz, $J_{H5\text{-}H6b}$ 2.2Hz, H6b), 4.29 (1H, ddd, $J_{H6a\text{-}H6b}$ 11.2Hz, $J_{H\text{-}31P}$ 6.2Hz, $J_{H5\text{-}H6a}$ 5.7Hz, H6a), 5.01 (1H, t, $J_{H3'\text{-}H4'}$ 9.1Hz, $J_{H4'\text{-}H5'}$ 9.1Hz, H4'), 5.02-5.05 (3H, m, H1', H2', H4), 5.07 (1H, dd, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 5.12 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 5.44 (1H, t, $J_{H2'\text{-}H3'}$ 9.5Hz, $J_{H3'\text{-}H4'}$ 9.5Hz, H3'), 5.46 (1H, t, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H3\text{-}H4}$ 9.5Hz, H3), 7.14-7.24 (7H, m, P-OPh ArH_{ortho}, P-OPh ArH_{para}, S-Ph ArH_{para}), 7.24-7.40 (8H, m, P-OPh ArH_{meta}, SPh ArH_{ortho}, SPh ArH_{meta}); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.5, 20.6, 20.6, 20.7, 20.7, 20.8 (6 x 1C, 6 x $\underline{C}H_3CO$), 35.6 (1C, C6'), 66.6 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, C6), 68.4 (1C, C2'), 68.7 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C5), 69.0 (1C, C2), 69.5 (1C, C4'), 70.1 (1C, C3'), 70.1 (1C, C3), 70.2 (1C, C5'), 71.7 (1C, C4), 92.1 (1C, C1'), 92.3 (1C, C1), 120.0 (4C, d, $J_{C\text{-}31P}$ 4.8Hz, P-OPh ArC_{ortho}), 125.5 (2C, P-OPh ArC_{para}), 126.3 (1C, S-Ph ArC_{para}), 129.0 (2C, S-Ph ArC_{ortho}), 129.0 (2C, S-Ph ArC_{meta}), 129.8 (4C, d, $J_{C\text{-}31P}$ 1.9Hz, P-OPh ArC_{meta}), 136.3 (1C, S-Ph ArC_{ipso}), 150.3 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, P-OPh ArC_{ipso}), 150.4 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, P-OPh ArC_{ipso}), 169.4 (2C, 2 x $CH_3\underline{C}O$), 169.6, 169.7(2 x 1C, 2 x $CH_3\underline{C}O$), 170.0(2C, 2 x $CH_3\underline{C}O$); FT-IR (thin film) υ 1216 (P=O), 1757 (C=O); HRMS (ES^+) m/z 941.2067 $[M + Na]^+$ (required 941.2062).



4.8. 2,2',3,3',4,4'-O-acetyl-D-trehalose-6,6'-diphosphate

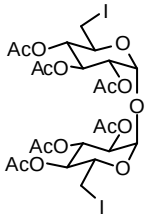
A suspension of **3.16** (50mg, 0.047mmol, 1eq) and PtO_2 (1mg, 5μmol, 0.1eq) in methanol (2ml) and water (2ml) with glacial acetic acid (20μl) was repeatedly degassed under vacuum and charged with hydrogen. The reaction mixture was vigorously stirred at room temperature for 32 hours, after which TLC (1 water : 2 isopropanol : 2 ethyl acetate) showed the complete consumption of starting material and the formation of a single product. The reaction was filtered through Celite® and the solvent removed *in vacuo* to give the desired compound as a colourless solid (32mg, 90%).

R_f 0.1 (1 water : 2 isopropanol : 2 ethyl acetate), $[\alpha]_D^{25}$ +115 (c = 1.0, MeOH); ^{1}H NMR (500MHz, MeOD) δ ppm 2.03, 2.07, 2.13 (6 x 3H, 3 x s, 6 x OAc), 4.01 (2H, m, H6'), 4.05 (2H, m, H6), 4.19 (2H, ddd, J_{H4-H5} 10.4Hz, J_{H5-H6} 5.1Hz, $J_{H5-H6'}$ 2.8Hz, H5), 5.11 (2H, dd, J_{H2-H3} 10.1Hz, J_{H1-H2} 3.1Hz, H2), 5.12 (2H, t, J_{H3-H4} 9.5Hz, J_{H4-H5} 9.5Hz, H4), 5.37 (2H, d, J_{H1-H2} 3.5Hz, H1), 5.50 (2H, t, J_{H2-H3} 9.8Hz, J_{H3-H4} 9.8Hz, H3); ^{13}C NMR (126MHz, MeOD) δ ppm 20.6, 20.7, 20.7 (3 x 2C, 6 x CO<u>C</u>H$_3$) 65.9 (2C, d, J_{C-31P} 3.8Hz, C6), 70.0 (2C, C4), 70.5 (2C, d, J_{C-31P} 3.8Hz, C5), 71.0 (2C, C2), 71.7 (2C, C3), 94.0 (2C, C1), 171.3, 171.3, 171.5 (3 x 2C, 6 x C=O); FT-IR (KBr disc) υ 1701 (C=O); HRMS m/z (ES^-) 753.1050 $[M - H]^-$ (required 753.1050).

4.9. 6,6'-dideoxyiodo-2,2',3,3',4,4'-acetyl-D-trehalose

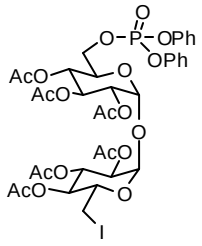
4.104.10. 6-*O*-(diphenoxyphosphoryl)-6'-deoxyiodo-2,2',3,3',4,4'-acetyl-D-trehalose

A solution of **3.16** (1.00g, 0.95mmol, 1eq) and TBAI (0.35g, 0.95mmol, 1eq) in anhydrous THF (7ml) was irradiated in a microwave cavity to 150°C at a maximum power of 300W and maximum pressure of 300Psi with a stream of cooling nitrogen for 2 hours, after which TLC (2 ethyl acetate : 1 petrol) showed the formation of two products. The solvents were removed *in vacuo* and the residue taken up in ethyl acetate (25ml) and washed with water (25ml) and brine (3 x 10ml), dried over $MgSO_4$ and concentrated *in vacuo*. The compounds were separated by automated silica gel chromatography (12 – 100% ethyl acetate / petrol over 9 column volumes) to give **4.10** (250mg, 29%) as a colourless oil and **4.9** (208mg, 27%) as a white crystalline solid.



*Compound **4.9***

R_f 0.62 (2 ethyl acetate : 1petrol), $[\alpha]_D^{25}$ +72.3 (c = 1.0, $CHCl_3$), m.p. 152-157°C; ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.01, 2.07, 2.14 (3 x 6H, 3 x s, 6 x OAc), 3.06 (2H, dd, $J_{H6\text{-}H6'}$ 11.0Hz, $J_{H5\text{-}H6'}$ 9.1Hz, H6'), 3.22 (2H, dd, $J_{H6\text{-}H6'}$ 10.9Hz, $J_{H5\text{-}H6}$ 2.4Hz, H6), 3.94 (2H, td, $J_{H4\text{-}H5}$ 9.5Hz, $J_{H5\text{-}H6'}$ 9.5Hz, $J_{H5\text{-}H6}$ 2.2Hz, H5), 4.88 (2H, t, $J_{H3\text{-}H4}$ 9.6Hz, $J_{H4\text{-}H5}$ 9.6Hz, H4), 5.19 (2H, dd, $J_{H2\text{-}H3}$ 10.4Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 5.41 (2H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 5.47 (2H, t, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H3\text{-}H4}$ 10.1Hz, H3); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 2.5 (2C, C6), 20.6, 20.7, 21.2 (3 x 2C, 6 x $\underline{C}H_3CO$), 69.2 (2C, C2), 69.7 (2C, C3), 69.9 (2C, C5), 72.3 (2C, C4), 91.7 (2C, C1), 169.4, 169.5, 169.9 (3 x 2C, 6 x $CH_3\underline{C}O$); FT-IR (KBr disc) υ 1760 (C=O); HRMS (ES^+) m/z 836.9728 $[M + Na]^+$ (required 836.9723).



*Compound **4.10***

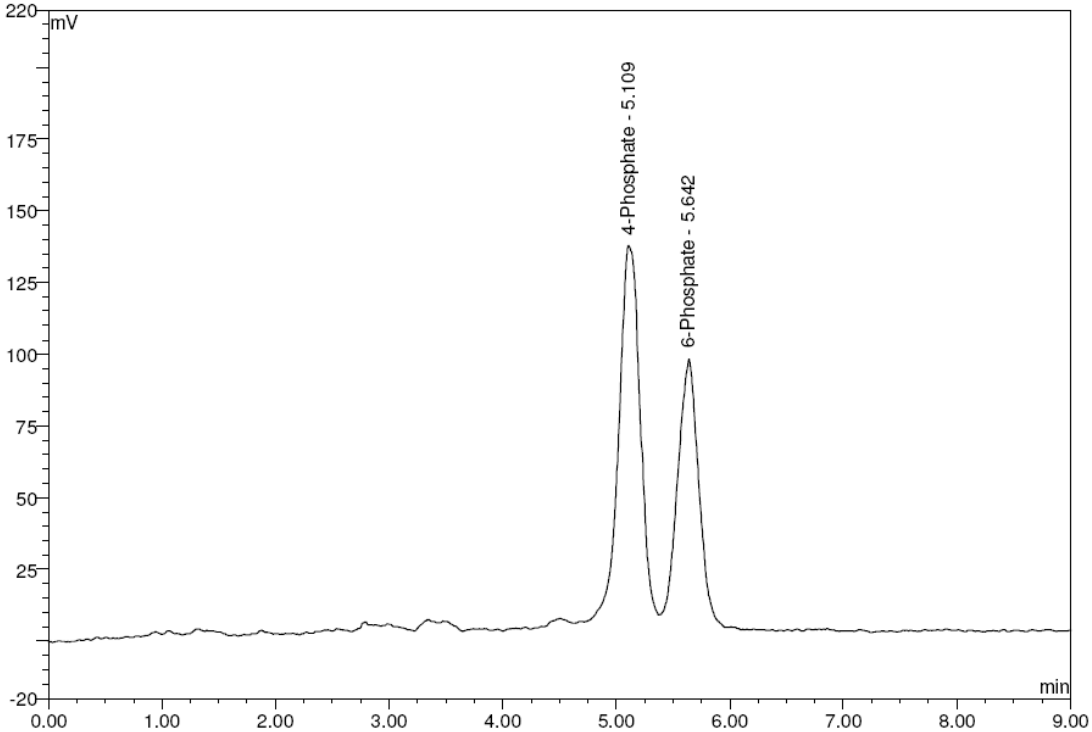
R_f 0.53 (2 ethyl acetate : 1 petrol), $[\alpha]_D^{25}$ +88.8 (c = 1.0, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.02, 2.02, 2.03, 2.05, 2.08, 2.14 (6 x 3H, 6 x s, 6 OAc), 3.06 (1H, dd, $J_{H6'a\text{-}H6'b}$ 10.9Hz, $J_{H5\text{-}H6'b}$ 8.7Hz, H6'b), 3.22 (1H, dd, $J_{H6'a\text{-}H6'b}$ 11.0Hz, $J_{H5\text{-}H6'a}$ 2.5Hz, H6'a), 3.90 (1H, td, $J_{H4'\text{-}H5'}$ 9.4Hz, $J_{H5'\text{-}H6'b}$ 9.4Hz, $J_{H5'\text{-}H6'a}$ 2.5Hz, H5'), 4.11 (1H, m, H5), 4.23 (1H, ddd, $J_{H6a\text{-}H6b}$ 11.3Hz, $J_{H\text{-}31P}$ 6.6Hz, $J_{H5\text{-}H6b}$ 2.2Hz, H6b), 4.29 (1H, ddd, $J_{H6a\text{-}H6b}$ 11.4Hz, $J_{H\text{-}31P}$ 6.8Hz, $J_{H5\text{-}H6a}$ 6.0Hz, H6a), 4.88 (1H, t, $J_{H3'\text{-}H4'}$ 9.6Hz, $J_{H4'\text{-}H5'}$ 9.6Hz, H4'), 5.04 (1H, dd, $J_{H3\text{-}H4}$ 10.2Hz, $J_{H4\text{-}H5}$ 9.3Hz, H4), 5.03 (1H, dd, $J_{H2'\text{-}H3'}$ 10.2Hz, $J_{H1'\text{-}H2'}$ 3.9Hz, H2'), 5.09 (1H, dd, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 5.11 (1H, d, $J_{H1'\text{-}H2'}$ 3.9Hz, H1'), 5.25 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 5.45 (1H, t, $J_{H2'\text{-}H3'}$ 9.8Hz, $J_{H3'\text{-}H4'}$ 9.8Hz, H3'), 5.46 (1H, t, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H3\text{-}H4}$ 9.8Hz, H3), 7.18-7.24 (6H, m, ArH_{ortho}, ArH_{para}), 7.35 (1H, t, $J_{ortho\text{-}meta}$ 7.9Hz, $J_{meta\text{-}para}$ 7.9Hz, ArH_{meta}), 7.36 (1H, t, $J_{ortho\text{-}meta}$ 7.9Hz, $J_{meta\text{-}para}$ 7.9Hz, ArH_{meta}); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 2.6 (1C, C6'), 20.5, 20.6, 20.6, 20.6, 20.7, 21.2 (6 x 1C, 6 x $\underline{C}H_3CO$), 66.6 (1C, d, $J_{C\text{-}31P}$ 4.8Hz, C6), 68.4 (1C, C2'), 68.7 (1C, d, $J_{C\text{-}31P}$ 8.6Hz, C5), 69.0 (1C, C2), 69.6 (1C, C4), 69.6 (1C, C5'), 69.7 (1C, C3), 70.1 (1C, C3'), 72.3 (1C, C4'), 92.0 (1C, C1'), 92.2 (1C, C1), 120.0 (4C, d, $J_{C\text{-}31P}$ 5.7Hz, ArH_{ortho}), 125.5 (2C, ArC_{para}), 129.8 (4C, ArC_{meta}), 150.3 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, ArC_{ipso}), 150.4 (1C, d, $J_{C\text{-}31P}$ 4.8Hz, ArC_{ipso}), 169.4, 169.4, 169.4, 169.6, 169.9, 170.0 (6 x 1C, $CH_3\underline{C}O$); FT-IR (thin film) υ 1220 (P=O), 1756 (C=O); HRMS (ES^+) m/z 959.1043 $[M + Na]^+$ (required 959.0995).

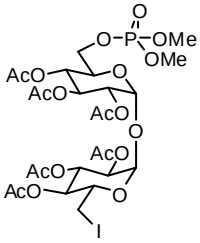
4.11. 6-O-(dimethoxyphosphoryl)-6'-iodo-2,2',3,3',4,4'-acetyl-D-trehalose

4.12. 4,6-(monomethyl)phosphoryl-6'-iodo-2,2',3,3',4',6'-acetyl-D-trehalose

4.13. 6-O-(dimethoxyphosphoryl)-2'6'-cyclo-2,3,3',4,4'-acetyl-D-trehalose

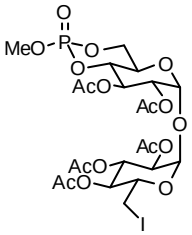
A solution of **4.10** (2.40g, 2.56mmol, 1eq) and CsF (3.88g, 25.6mmol, 10eq) was heated under reflux for 8 hours after which TLC showed the complete consumption of starting sugar. The reaction mixture was cooled to room temperature and the solvent removed *in vacuo*. The residue was dissolved in pyridine and acetic anhydride (4.35ml, 46.1mmol, 18eq) was added. The reaction mixture was stirred at room temperature for 32 hours. The reaction was quenched with methanol (5ml) and the solvents removed *in vacuo*. The crude products were taken up in ethyl acetate (50ml) and washed with 1M HCl (3 x 50ml), saturated $NaHCO_3$ (50ml), 0.25M $CuSO_4$ (50ml), brine (2 x 20ml) and dried over $MgSO_4$. The products were separated by silica gel chromatography (4 ethyl acetate : 1 petrol to pure ethyl acetate over 4 CV then ethyl acetate for 6 CV) to give the desired compound **4.11** as a white solid (0.85g, 42%), as well as side product **4.12** (38mg, 2%) and as colourless oils. Two other cyclization adducts were also isolated by silica gel chromatography (0.41g, 25%) but required HPLC using a Phenomenex Synergi Hydro C_{18} column (21.2 x 250mm, 4μm) using 3 water : 1 acetonitrile as the mobile phase at a flow rate of 23ml/min, with ELS detection of eluants.





*Compound **4.11***

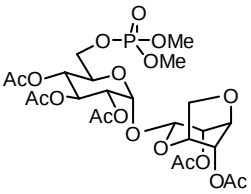
R_f 0.27 (ethyl acetate), $[\alpha]_D^{25}$ +96.2 (c = 1.0, $CHCl_3$), m.p. 59-60°C (ethyl acetate/petrol); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.02, 2.03, 2.07, 2.08, 2.09, 2.15 (6 x 3H, 6 x s, 6 x OAc), 3.07 (1H, dd, $J_{H6'a-H6'b}$ 11.0Hz, $J_{H5'-H6'b}$ 8.8Hz, H6'b), 3.23 (1H, dd, $J_{H6'a-H6'b}$ 11.0Hz, $J_{H5'-H6'a}$ 2.5Hz, H6'a), 3.76 (3H, d, J_{H-31P} 11.3Hz, OMe), 3.77 (3H, d, J_{H-31P} 11.3Hz, OMe), 3.94 (1H, td, $J_{H4'-H5'}$ 9.4Hz, $J_{H5'-H6'b}$ 9.4Hz, $J_{H5'-H6'a}$ 2.4Hz, H5'), 4.03 (1H, ddd, $J_{H6a-H6b}$ 10.7Hz, $J_{H6b-31P}$ 5.3Hz, J_{H5-H6b} 1.9Hz, H6b), 4.05 (1H, m, H5), 4.09 (1H, dt, $J_{H6a-H6b}$ 10.7Hz, $J_{H6a-31P}$ 5.7Hz, J_{H5-H6a} 5.7Hz, H6a), 4.89 (1H, t, $J_{H3'-H4'}$ 9.6Hz, $J_{H4'-H5'}$ 9.6Hz, H4'), 5.07 (1H, dd, $J_{H2'-H3'}$ 10.1Hz, $J_{H1'-H2'}$ 3.8Hz, H2'), 5.08 (1H, t, J_{H3-H4} 9.5Hz, J_{H4-H5} 9.5Hz, H4), 5.18 (1H, dd, J_{H2-H3} 10.2Hz, J_{H1-H2} 3.9Hz, H2), 5.33 (1H, d, $J_{H1'-H2'}$ 4.1Hz, H1'), 5.38 (1H, d, J_{H1-H2} 3.8Hz, H1), 5.48 (1H, t, J_{H2-H3} 9.5Hz, J_{H3-H4} 9.5Hz, H3), 5.49 (1H, dd, $J_{H2'-H3'}$ 10.1Hz, $J_{H3'-H4'}$ 9.5Hz, H3'); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 2.6 (1C, C6'), 20.5, 20.6, 20.7, 21.2 (4 x 1C, 4 x CO<u>C</u>H_3), 20.6 (2C, 2 x CO<u>C</u>H_3), 54.6 (2C, d, J_{C-31P} 6.7Hz, 2 x OMe), 65.2 (1C, d, J_{C-31P} 4.8Hz, C6), 68.2 (1C, C4), 68.8 (1C, d, J_{C-31P} 8.6Hz, C5), 69.1 (1C, C2), 69.5 (1C, C3'), 69.7 (1C, C5'), 69.8 (1C, C2'), 70.2 (1C, C3), 72.3 (1C, C4'), 91.9 (1C, C1'), 92.3 (1C, C1), 169.5, 169.5, 169.5, 169.6, 169.9, 170.0 (6 x 1C, 6 x C=O); FT-IR (KBr disc) υ 1219 (P=O), 1754 (C=O); HRMS m/z (ES^+) 835.0675 $[M + Na]^+$ (required 865.0682).



*Compound **4.12***

R_f 0.38 (ethyl acetate), $[\alpha]_D^{25}$ 93.4 (c = 1.0, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.03, 2.08, 2.09, 2.15, 2.15 (5 x 3H, 5 x s, 5 x OAc), 3.05 (1H, dd, $J_{H6'a-H6'b}$ 11.0Hz, $J_{H5'-H6'b}$ 9.1Hz, H6'b), 3.23 (1H, dd, $J_{H6'a-H6'b}$ 11.0Hz, $J_{H5'-H6'a}$ 2.5Hz, H6'a), 3.86 (3H, d, J_{H-31P} 11.0Hz, OMe), 3.95 (1H, td, $J_{H4'-H5'}$ 9.5Hz, $J_{H5'-H6'b}$ 9.5Hz, $J_{H5'-H6'a}$ 2.4Hz, H5'), 4.09 (1H, m, H5), 4.13 (1H, t, $J_{H6ax-H6eq}$ 9.5Hz, $J_{H5-H6ax}$ 9.5Hz, H6ax), 4.19 (1H, t, J_{H3-H4} 9.5Hz, J_{H4-H5} 9.5Hz, H4), 4.21 (1H, ddd, $J_{H6eq-31P}$ 22.7Hz, $J_{H6ax-H6eq}$ 9.1Hz, $J_{H5-H6eq}$ 3.5Hz, H6eq), 4.88 (1H, t, $J_{H3'-H4'}$ 9.5Hz, $J_{H4'-H5'}$ 9.5Hz, H4'), 4.98 (1H, dd, $J_{H2'-H3'}$ 10.4Hz, $J_{H1'-H2'}$ 4.1Hz, H2'), 5.11 (1H, dd, J_{H2-H3} 9.9Hz, J_{H1-H2} 3.9Hz, H2), 5.33 (1H, d, J_{H1-H2} 3.8Hz, H1), 5.39 (1H, d, $J_{H1'-H2'}$ 4.1Hz, H1'), 5.48 (1H, t, $J_{H2'-H3'}$ 9.8Hz, $J_{H3'-H4'}$ 9.8Hz, H3'), 5.54 (1H, t, J_{H2-H3} 9.6Hz, J_{H3-H4} 9.6Hz, H3); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 2.5 (1C, C6'), 20.6, 20.7, 21.2 (5C, 3 x as, 5 x CO<u>C</u>H_3), 54.3 (1C, d, J_{C-31P} 5.7Hz, OMe), 63.0 (1C, d, J_{C-31P} 5.7Hz, C5), 68.1 (1C, d, J_{C-31P} 8.6Hz, C6), 69.1

(1C, d, $J_{C\text{-}31P}$ 9.5Hz, C3), 69.2 (1C, d, $J_{C\text{-}31P}$ 2.9Hz, C2), 69.3 (1C, C3'), 70.0 (1C, C5'), 70.3 (1C, C2'), 72.1 (1C, C4'), 77.5 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, C4), 91.8 (1C, C1), 92.8 (1C, C1'), 169.1, 169.6, 169.6, 169.7, 170.1 (5 x 1C, 5 x C=O); FT-IR (thin film) υ 1220 (P=O), 1755 (C=O).



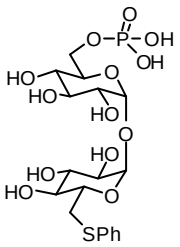
*Compound **4.13***

R_f 0.14 (ethyl acetate), $[\alpha]_D^{25}$ +87.7 (c = 0.11, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.03, 2.07, 2.11, 2.17, 2.20 (5 x 3H, 5 x s, 5 x OAc), 3.79 (6H, d, $J_{H\text{-}31P}$ 11.3Hz, 2 x OMe), 3.99 (1H, dd, $J_{H6'a\text{-}H6'b}$ 10.9Hz, $J_{H5'\text{-}H6'b}$ 3.0Hz, H6'b), 4.02 (1H, ddd, $J_{H4\text{-}H5}$ 10.2Hz, $J_{H5\text{-}H6b}$ 3.0Hz, $J_{H5\text{-}H6a}$ 1.9Hz, H5), 4.10 (1H, ddd, $J_{H6a\text{-}H6b}$ 11.0Hz, $J_{H6b\text{-}31P}$ 6.8Hz, $J_{H5\text{-}H6b}$ 4.3Hz, H6b), 4.16 (1H, ddd, $J_{H6a\text{-}H6b}$ 11.4Hz, $J_{H6a\text{-}31P}$ 5.7Hz, $J_{H5\text{-}H6a}$ 2.2Hz, H6a), 4.19 (1H, d, $J_{H6a'\text{-}H6b'}$ 10.7Hz, H6a'), 4.50 (1H, t, $J_{H4'\text{-}H5'}$ 2.7Hz, $J_{H5'\text{-}H6'b}$ 2.7Hz, H5'), 4.64 (1H, t, $J_{H2'\text{-}H3'}$ 4.7Hz, $J_{H3'\text{-}H4'}$ 4.7Hz, H3'), 4.73 (1H, dd, $J_{H3'\text{-}H4'}$ 5.0Hz, $J_{H4'\text{-}H5'}$ 2.8Hz, H4'), 4.82 (1H, dd, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 5.11 (1H, t, $J_{H3\text{-}H4}$ 9.9Hz, $J_{H4\text{-}H5}$ 9.9Hz, H4), 5.17 (1H, t, $J_{H1'\text{-}H2'}$ 3.9Hz, $J_{H2'\text{-}H3'}$ 3.9Hz, H2'), 5.35 (1H, d, $J_{H1'\text{-}H2'}$ 3.5Hz, H1'), 5.37 (1H, t, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H3\text{-}H4}$ 9.8Hz, H3), 5.53 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.5, 20.5, 20.6, 20.7, 20.7 (5 x 1C, 5 x CO$\underline{C}H_3$), 54.5 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, OMe), 54.6 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, OMe), 65.2 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, C6), 67.6 (1C, C4), 67.6 (1C, C2'), 68.2 (1C, C6'), 68.9 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C5), 69.4 (1C, C3), 70.0 (1C, C3'), 70.4 (1C, C2), 70.9 (1C, C4'), 73.5 (1C, C5'), 90.0 (1C, C1'), 92.0 (1C, C1), 169.6, 170.0, 170.1, 170.1, 170.7 (5 x 1C, 5 x C=O); FT-IR (thin film) υ 1225 (P=O), 1748 (C=O); HRMS m/z (ES^+) 665.1463 $[M + Na]^+$ (required 665.1453).

General protocol for one-pot substitution/deprotection

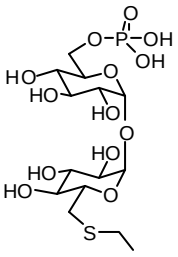
To a suspension of NaH (60% dispersion in mineral oil, 10eq) in anhydrous THF (1ml) was added dropwise at room temperature the thiol (10eq). The reaction was stirred for 30 minutes and once hydrogen production had ceased, a solution of **4.11** (1eq) in THF (4ml) was added. The reaction mixture was irradiated in a microwave cavity to 80°C at a maximum power of 300W and maximum pressure of 300Psi with a stream of cooling nitrogen for 30 minutes. The reaction was monitored by electrospray mass spectrometry (negative mode) for unwanted phosphate cyclization. The reaction was quenched by addition of glacial acetic acid. The solvents were

removed under reduced pressure, and the crude mixture was dried *in vacuo* for 1 hour. To the crude solid was added dioxane (3ml) and the suspension was sonicated for 10 minutes. TMSI (10eq) was added and the reaction mixture stirred at room temperature for 60 minutes, after which mass spectrometry and TLC showed the formation of a single product. The reaction was quenched by addition of water (1ml) and the solvents were removed *in vacuo*. The crude product was taken up in water (1ml) and washed with ethyl acetate (3 x 2ml). The aqueous layer was concentrated under reduced pressure and the product purified using HPLC on an Applied Biosystems Poros® HQ strongly basic anion exchange column (10mm x 100mm, 50µm). A gradient from 0mM to 500mM aqueous NH_4HCO_3 was used as the mobile phase and eluants were detected with an ELS detector parallel to the main flow path. Fractions containing the desired product were pooled and repeated lyophilization to removed residual NH_4HCO_3 afforded the title compounds as a white amorphous solids.



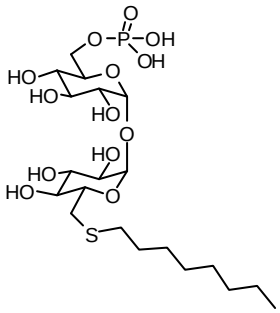
*Compound **4.18***

R_f 0.8 (5 ethanol : 3 NH_4OH : 1 water), $[\alpha]_D^{25}$ +18.9 (c = 1.12, H_2O); ^1H NMR (500MHz, D_2O) δ ppm 3.01 (1H, dd, $J_{H6'a\text{-}H6'b}$ 14.3Hz, $J_{H5'\text{-}H6'b}$ 8.7Hz, H6'b), 3.33 (1H, t, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, H4), 3.48 (1H, m, H6'a), 3.49 (1H, dd, $J_{H2'\text{-}H3'}$ 10.1Hz, $J_{H1'\text{-}H2'}$ 4.4Hz, H2'), 3.51 (1H, t, $J_{H3'\text{-}H4'}$ 10.1Hz, $J_{H4'\text{-}H5'}$ 10.1Hz, H4'), 3.59 (1H, dd, $J_{H2\text{-}H3}$ 9.9Hz, $J_{H1\text{-}H2}$ 3.9Hz, H2), 3.71 (1H, t, $J_{H2'\text{-}H3'}$ 9.6Hz, $J_{H3'\text{-}H4'}$ 9.6Hz, H3'), 3.71 (1H, t, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H3\text{-}H4}$ 9.5Hz, H3), 3.76 (1 H, add, $J_{H4\text{-}H5}$ 9.8Hz, $J_{H5\text{-}H6a}$ 3.5Hz, H5), 3.85 (1H, ddd, $J_{H6a\text{-}H6b}$ 12.0Hz, $J_{H6b\text{-}31P}$ 5.2Hz, $J_{H5\text{-}H6b}$ 1.4Hz, H6b), 3.87 (1H, m, H5'), 3.93 (1H, ddd, $J_{H6a\text{-}H6b}$ 12.0Hz, $J_{H6a\text{-}31P}$ 7.4Hz, $J_{H5\text{-}H6a}$ 4.1Hz,H6a), 5.00 (1H, d, $J_{H1'\text{-}H2'}$ 4.1Hz, H1'), 5.08 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 7.21 (1H, t, $J_{meta\text{-}para}$ 7.6Hz, ArH_{para}), 7.31 (2H, t, $J_{ortho\text{-}meta}$ 7.7Hz, $J_{meta\text{-}para}$ 7.7Hz, ArH_{meta}), 7.39 (1H, d, $J_{orth\text{-}meta}$ 7.9Hz, ArH_{ortho}); ^{13}C NMR (126MHz, D_2O) δ ppm 34.8 (1C, C6'), 62.5 (1C, d, $J_{C\text{-}31P}$ 3.8Hz, C6), 69.0 (1C, C4), 70.7 (1C, C5'), 70.9 (1C, C2), 71.1 (1C, C2'), 71.8 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, C5), 72.3 (2C, C3, C3'), 72.9 (1C, C4'), 92.9 (1C, C1), 93.3 (1C, C1'), 126.6 (1C, ArC_{para}), 129.2 (1C, ArC_{ortho}), 129.3 (1C, ArC_{meta}), 135.1 (1C, ArC_{ipso}); FT-IR (KBr disc) υ 3453 br (OH); MRMS m/z (ES^-) 513.0839 $[M - H]^-$ (required 513.0837).



*Compound **4.19***

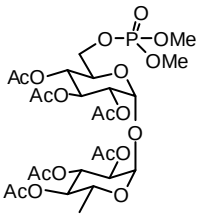
R_f 0.41 (5 ethanol : 3 NH_4OH : 1 water), $[\alpha]_D^{25}$ +15.8 (c = 0.80, H_2O); ^1H NMR (500MHz, D_2O) δ ppm 1.16 (3H, t, $J_{CH2\text{-}CH3}$ 7.4Hz, $SCH_2C\underline{H}_3$), 2.57 (2H, q, $J_{CH2\text{-}CH3}$ 7.6Hz, $SC\underline{H}_2CH_3$), 2.64 (1H, dd, $J_{H6'a\text{-}H6'b}$ 13.9Hz, $J_{H5'\text{-}H6'b}$ 8.5Hz, H6'b), 2.99 (1H, dd, $J_{H6'a\text{-}H6'b}$ 14.2Hz, $J_{H5'\text{-}H6'a}$ 2.2Hz, H6'a), 3.30 (1H, t, $J_{H3'\text{-}H4'}$ 9.3Hz, $J_{H4'\text{-}H5'}$ 9.3Hz, H4'), 3.55 (1H, t, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, H4), 3.60 (1H, dd, $J_{H2'\text{-}H3'}$ 10.1Hz, $J_{H1'\text{-}H2'}$ 4.1Hz, H2'), 3.61 (1H, dd, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H1\text{-}H2}$ 4.1Hz, H2), 3.74 (1H, t, $J_{H2'\text{-}H3'}$ 9.8Hz, $J_{H3'\text{-}H4'}$ 9.8Hz, H3'), 3.78 (1H, t, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H3\text{-}H4}$ 9.5Hz, H3), 3.82 (1H, m, H5), 3.87 (1H, m, H5'), 3.87 (1H, ddd, $J_{H6a\text{-}H6b}$ 12.6Hz, $J_{H6b\text{-}31P}$ 5.4Hz, $J_{H5\text{-}H6b}$ 1.9Hz, H6b), 3.95 (1H, ddd, $J_{H6a\text{-}H6b}$ 12.0Hz, $J_{H6a\text{-}31P}$ 7.6Hz, *J*H5-H6a 4.1Hz, H6a), 5.11 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.13 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 13.8 (1C, $SCH_2\underline{C}H_3$), 25.9 (1C, $S\underline{C}H_2CH_3$), 32.2 (1C, C6'), 62.5 (1C, d, $J_{C\text{-}31P}$ 3.8Hz, C6), 69.1 (1C, C4), 70.9 (1C, C5'), 71.0 (1C, C2'), 71.2 (1C, C2), 71.8 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C5), 72.3 (1C, C3), 72.3 (1C, C3'), 72.8 (1C, C4'), 93.1 (1C, C1'), 93.6 (1C, C1); HRMS m/z (ES^-) 465.0844 $[M - H]^-$ (required 465.0837).



*Compound **4.20***

R_f 0.8 (5 ethanol : 3 NH_4OH : 1 water), $[\alpha]_D^{25}$ 22.64 (c = 0.75, H_2O); ^1H NMR (500MHz, D_2O) δ ppm 0.79 (3H, t, $J_{H13\text{-}H14}$ 6.0Hz, H14), 1.21 (6H, br s, H13, H12, H11), 1.30 (2H, s, H10), 1.53 (2H, quin, $J_{H8\text{-}H9}$ 5.7Hz, $J_{H9\text{-}H10}$ 5.7Hz, H9), 2.56 (2H, at, $J_{H8\text{-}H9}$ 6.6Hz, H8), 2.63 (1H, dd, $J_{H7\text{-}H7'}$ 13.9Hz, $J_{H7'\text{-}H8}$ 8.8Hz, H7'), 2.98 (1H, d, $J_{H7\text{-}H7'}$ 13.2Hz, H7), 3.30 (1H, t, $J_{H3'\text{-}H4'}$ 9.1Hz, $J_{H4'\text{-}H5'}$ 9.1Hz, H4'), 3.54 (1H, t, $J_{H3\text{-}H4}$ 9.1Hz, $J_{H4\text{-}H5}$ 9.1Hz, H4), 3.60 (2H, ad, $J_{H2\text{-}H3}$ 8.8Hz, H2, H2'), 3.70 (1H, m, H6'b), 3.74 (1H, t, $J_{H2'\text{-}H3'}$ 9.5Hz, $J_{H3'\text{-}H4'}$ 9.5Hz, H3'), 3.78 (1H, t, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H3\text{-}H4}$ 9.8Hz, H3), 3.77-3.88 (2H, m, H5, H6'a), 3.85 (2H, m, H5', H6b), 3.94 (1H, ad, $J_{H6a\text{-}H6b}$ 10.7Hz, H6a), 5.11 (1H, d, $J_{H1'\text{-}H2'}$ 3.1Hz, H1'), 5.13 (1H, d, $J_{H1\text{-}H2}$ 3.1Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 13.5 (1C, C14), 22.0 (1C, C13), 27.9 (1C, C10), 28.3 (1C, C12), 28.3 (1C, C9), 28.6 (1C, C11), 31.1 (1C, C8), 31.9 (1C, C7), 61.9 (1C, C6'), 62.5 (1C, d, $J_{C\text{-}31P}$ 3.9Hz, C6), 68.3 (1C, C4), 69.1 (1C, C5'), 70.9 (1C, C2'), 71.0 (1C, C2), 71.2 (1C, C3'), 71.8 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, C5),

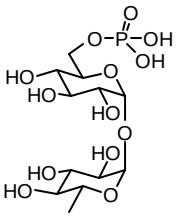
72.3 (1C, C3), 72.8 (1C, C4'), 93.1 (1C, C1'), 93.5 (1C, C1); FT-IR (KBr disc) υ 3455 br (OH); HRMS m/z (ES^-) 549.1776 $[M - H]^-$ (required 549.1776).



4.23. 2,2',3,3',4,4'-O-acetyl-6-(dimethoxy)phosphoryl-6'-deoxy-D-trehalose

A suspension of **4.11** (20mg, 0.025mmol, 1eq) and Pd/C (5mg, 0.0025mmol, 0.1eq) was vigorously stirred in methanol (5ml) and triethylamine (0.5ml). The reaction mixture was degassed under vacuum and charged with hydrogen (x3) and then stirred at room temperature for 16 hours after which TLC (2 ethyl acetate : 1 petrol) showed the complete consumption of starting material. The reaction was filtered through Celite® and the solvent removed *in vacuo*. The crude product was taken up in ethyl acetate (20ml) and washed with water (20ml), 1M HCl (2 x 20ml), saturated $NaHCO_3$ (20ml), brine (20ml) and dried over $MgSO_4$. The product was purified by automated silica gel chromatography (90% ethyl acetate in petrol for 3 column volumes, ramped to 100% ethyl acetate over 4 column volumes, then held at this composition for 8 column volumes) as a colourless oil (16.3mg, 96%).

R_f 0.34 (ethyl acetate), $[\alpha]_D^{25}$ +91.7 (c = 0.82, $CHCl_3$); ^1H NMR (500MHz, $CDCl_3$) δ ppm 1.16 (3H, d, $J_{H5\text{-}H6'}$ 6.3Hz, H6'), 2.03, 2.04, 2.07, 2.09 (4 x 3H, 4 x s, 4 x OAc), 2.08 (2 x 3H, s, 2 x OAc), 3.77 (3H, d, $J_{H\text{-}31P}$ 11.3Hz, OMe), 3.77 (3H, d, $J_{\text{-}31P}$ 11.3Hz, OMe), 3.98 (1H, m, $J_{H5'\text{-}H6'}$ 6.3Hz, H5'), 4.02 (1H, ddd, $J_{H6a\text{-}H6b}$ 14.0Hz, $J_{H5\text{-}H6b}$ 9.1Hz, $J_{H5\text{-}31P}$ 5.0Hz, H6b), 4.13 (2H, m, H5, H6a), 4.83 (1H, t, $J_{H3'\text{-}H4'}$ 9.6Hz, $J_{H4'\text{-}H5'}$ 9.6Hz, H4'), 5.01 (1H, dd, $J_{H2\text{-}H3}$ 10.4Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 5.03 (1H, dd, $J_{H2'\text{-}H3'}$ 9.5Hz, $J_{H1'\text{-}H2'}$ 3.8Hz, H2'), 5.09 (1H, t, $J_{H3\text{-}H4}$ 9.6Hz, $J_{H4\text{-}H5}$ 9.6Hz, H4), 5.20 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.32 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 5.43 (1H, t, $J_{H2'\text{-}H3'}$ 9.8Hz, $J_{H3'\text{-}H4'}$ 9.8Hz, H3'), 5.50 (1H, t, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H3\text{-}H4}$ 9.8Hz, H3); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 17.1 (1C, C6'), 20.6, 20.6, 20.7, 20.8 (4 x 1C, 4 x CO$\underline{C}H_3$), 20.7 (2C, 2 x CO$\underline{C}H_3$), 54.6 (2C, d, $J_{C\text{-}31P}$ 6.4Hz, 2 x OMe), 65.3 (1C, d, $J_{C\text{-}31P}$ 4.6Hz, C6), 66.0 (1C, C5'), 68.3 (1C, C4), 68.7 (1C, d, $J_{C\text{-}31P}$ 8.2Hz, C5), 70.1 (1C, C3'), 70.1 (1C, C3), 70.2 (2C, C2, C2'), 73.5 (1C, C4'), 92.9 (1C, C1), 93.3 (1C, C1'), 128.6, 128.7, 129.0, 129.1 (6C, C=O); ^{31}P NMR (162MHz, $CDCl_3$) δ ppm 1.3 ($P(O)(OMe)_2$); FT-IR (thin film) υ 1223 (P=O), 1720 (C=O); HRMS (ES^+) m/z 709.1715 $[M + Na]^+$ (required 709.1718).

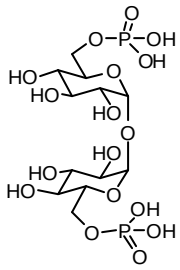


4.25. 6'-deoxy-D-trehalose-6-phosphate

To a solution of **4.23** (13mg, 0.019mmol, 1eq) in anhydrous dioxane (2ml) was added TMSBr (5.5µl, 0.042mmol, 2.2eq) and the reaction mixture heated to 50°C. After 17 hours, TLC (1 water : 4 isopropanol : 4 ethyl acetate) showed complete consumption of starting material and the formation of a single product. The reaction was quenched with methanol (1ml). The solvent was removed under reduced pressure and the crude product dried under high vacuum for 2 hours. The crude product was dissolved in dry methanol (3ml) and NaOMe (~5mg) was added until the reaction mixture was pH 10 (pH paper). The reaction mixture was stirred at room temperature for 4 hours, after which TLC (1 water : 3 NH_4OH : 5 ethanol) showed the complete consumption of starting material and the formation of a single product. The reaction was quenched by addition of solid CO_2 until the reaction mixture was at pH 7 (pH paper). The crude product was partitioned between water (3ml) and ethyl acetate (3ml). The aqueous layer was retained and washed with ethyl acetate (6 x 2ml). The aqueous layer was concentrated under reduced pressure and the crude product dissolved in water (1ml). The solution was passed through a C_{18} SepPak® Light cartridge, pre-equilibrated with water, and the cartridge was eluted with water (4ml). All effluent was collected and The aqueous layer was concentrated under reduced pressure and purified using HPLC on an Applied Biosystems Poros® HQ strongly basic anion exchange column (10mm x 100mm, 50µm). A gradient from 0mM to 500mM aqueous NH_4HCO_3 was used as the mobile phase and eluants were detected with an ELS detector parallel to the main flow path. Fractions containing the product were pooled and lyophilized repeatedly to remove residual NH_4HCO_3 to yield the title compound as an amorphous white solid (7.0mg, 91%).

R_f 0.39 (1 water : 3 NH_4OH : 5 ethanol), ^{1}H NMR (500MHz, D_2O) δ ppm 1.19 (3H, d, $J_{H5'\text{-}H6'}$ 6.3Hz, H6'), 3.11 (1H, t, $J_{H3'\text{-}H4'}$ 9.5Hz, $J_{H4'\text{-}H5'}$ 9.5Hz, H4'), 3.54 (1H, t, $J_{H3\text{-}H4}$ 9.8Hz, $J_{H4\text{-}H5}$ 9.8Hz, H4), 3.57 (1H, dd, $J_{H2'\text{-}H3'}$ 10.7Hz, $J_{H1'\text{-}H2'}$ 4.4Hz, H2'), 3.60 (1H, dd, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.71 (1H, t, $J_{H2'\text{-}H3'}$ 9.5Hz, $J_{H3'\text{-}H4'}$ 9.5Hz, H3'), 3.77 (1H, t, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H3\text{-}H4}$ 9.8Hz, H3), 3.79-3.82 (2H, m, H5, H5'), 3.86 (1H, ddd, $J_{H6a\text{-}H6b}$ 11.7Hz, $J_{H6b\text{-}31P}$ 5.4Hz, $J_{H5\text{-}H6b}$ 1.6Hz, H6b), 3.95 (1H, ddd, $J_{H6a\text{-}H6b}$ 11.9Hz, $J_{H6a\text{-}31P}$ 7.6Hz, $J_{H5\text{-}H6a}$ 4.1Hz, H6a), 5.07 (2H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1, H1'); ^{13}C NMR (126MHz, D_2O) δ ppm 16.6 (1C, C6'), 62.4 (1C, d, $J_{C\text{-}31P}$ 4.6Hz, C6), 68.2 (1C, C5'), 69.1 (1C, C4), 71.2 (1C, C2'), 71.2 (1C, C2), 71.8 (1C, d, $J_{C\text{-}31P}$

7.3Hz, C5), 72.1 (1C, C3'), 72.2 (1C, C3), 75.2 (1C, C4'), 93.2 (1C, C1), 93.5 (1C, C1'); ^{31}P NMR (162MHz, D_2O) δ ppm 4.6 ($P(O)(OH)_2$); FT-IR (KBr disc) υ 1112 (P=O), 3402 br (OH); HRMS (ES^-) m/z 405.0807 $[M - H]^-$ (required 405.0804).

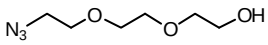


4.26. D-trehalose-6,6'-diphosphate

A suspension of **2.2** (180mg, 0.22mmol, 1eq) and PtO_2 (5mg, 0.022mmol, 0.1eq) in methanol (2.5ml) and water (7.5ml) with 1M HCl (0.2ml) was repeatedly degassed under vacuum and charged with hydrogen. The reaction mixture was vigorously stirred at room temperature for 28 hours, after which TLC (5 ethanol : 3 NH_4OH : 1 water) showed the complete consumption of starting material and the formation of a single product. The reaction mixture was filtered through Celite® and the solvent removed *in vacuo*. Lyophilization yielded the desired compound as an amorphous white solid (110mg, 98%).

R_f 0.13 (5 ethanol : 3 NH_4OH : 1 water), $[\alpha]_D^{25}$ +45.4 (c = 1.0, H_2O); ^{1}H NMR (500MHz, D_2O) δ ppm 3.46 (2H, t, $J_{H3\text{-}H4}$ 9.6Hz, $J_{H4\text{-}H5}$ 9.6Hz, H4), 3.59 (2H, dd, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 3.77 (2H, t, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H3\text{-}H4}$ 9.5Hz, H3), 3.85 (2H, dt, $J_{H4\text{-}H5}$ 10.1Hz, $J_{H5\text{-}H6}$ 3.0Hz, $J_{H5\text{-}H6'}$ 3.0Hz, H5), 3.98-4.03 (4H, m, H6, H6'), 5.12 (2H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 63.7 (2C, d, $J_{C\text{-}31P}$ 4.8Hz, C6), 69.2 (2C, C4), 70.9 (2C, C2), 71.2 (2C, d, $J_{C\text{-}31P}$ 7.6Hz, C5), 72.3 (2C, C3), 93.5 (2C, C1); $^{31}P\{^1H\}$ NMR (162MHz, D_2O) δ ppm 0.6 (1P, $P(O)(OH)_2$); FT-IR (KBr disc) υ 3308 br (OH); HRMS m/z (ES^-) 501.0413 $[M - H]^-$ (required 501.0416).

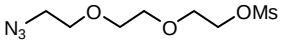
4.11.3 Synthesis of affinity columns



4.28. 1-azido- triethylene glycol[46]

A mixture of **4.27** (20.0g, 119mmol, 1eq) and sodium azide (11.1g, 131mmol, 1.1eq) in DMF (50ml) was heated to 80°C for 18 hours after which TLC (ethyl acetate) showed complete consumption of starting material. The reaction mixture was allowed to cool to room temperature and the solvent removed *in vacuo*. The crude residue was partitioned between DCM (100ml) and water (50ml). The aqueous layer was extracted with DCM (2 x 50ml) and the combined organics were dried over $MgSO_4$. The solvent was removed under reduced pressure to furnish the desired compound as a colourless liquid (20.1g, 97%).

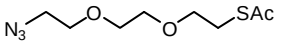
R_f 0.46 (ethyl acetate); ^{1}H NMR (400MHz, $CDCl_3$) δ ppm 2.38 (1H, t, $J_{H\text{-}OH}$ 5.6, $CH_2O\underline{H}$), 3.41 (2H, t, *J* 4.9Hz, CH_2N_3), 3.62 (2H, t, *J* 4.0Hz, $C\underline{H_2}CH_2OH$), 3.69 (6H, s, OCH_2CH_2O, $OC\underline{H_2}CH_2N_3$), 3.74 (2H, q, *J* 4.8Hz, $C\underline{H_2}OH$); ^{13}C NMR (101MHz, $CDCl_3$) δ ppm 50.6 (1C, CH_2N_3), 61.8 (1C, CH_2OH), 70.0, 70.4, 70.6 (3 x 1C, 3 x CH_2), 72.5 (1C, $\underline{C}H_2CH_2OH$); FT-IR (thin film) υ 2105 (N_3), 3453 br (OH); m/z (ES^+) 198.1 $[M + Na]^+$.



4.29. 1-azido-5-mesyl-triethylene glycol

To a solution of **4.28** (20.5g, 117mmol, 1eq) in triethylamine (33ml, 234mmol, 2eq) and anhydrous DCM (200ml) at 0°C was added mesyl chloride (13.6ml, 176mmol, 1.5eq) dropwise over 15 minute. The reaction mixture was stired at this temperature for 3 hours, after which TLC (ethyl acetate) showed complete comsumption of starting material. The reaction mixture was filtered to remove triethylamine salts and the solvent removed *in vacuo*. The reaction mixture was partitioned with DCM (200ml) and water (50ml) and the organic layer washed with 1M HCl (100ml). The aqueous layers were combined and re-extracted with DCM (3 x 50ml) and the combined organics dried over $MgSO_4$. The solvent was removed *in vacuo* to yield the desired compound as a yellow liquid (29.0g, 98%).

R_f 0.68 (ethyl acetate), ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 3.06 (3H, s, OMs), 3.38 (2H, t, *J* 4.9Hz, $OCH_2C\underline{H_2}N_3$), 3.66 (6H, m, $OC\underline{H_2}C\underline{H_2}OC\underline{H_2}CH_2N_3$), 3.78 (2H, t, *J* 4.4Hz, $MsOCH_2C\underline{H_2}O$), 4.37 (2H, t, *J* 4.4Hz, $MsOC\underline{H_2}CH_2O$); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 37.6 (1C, OMs), 50.6 (1C, $OCH_2\underline{C}H_2N_3$), 69.1 (1C, $MsOCH_2\underline{C}H_2O$), 69.2 (1C, $MsO\underline{C}H_2CH_2O$), 70.0, 70.6, 70.6 (3 x 1C, $O\underline{C}H_2\underline{C}H_2O\underline{C}H_2CH_2N_3$); FT-IR (thin film) υ 2110 (N_3); HRMS (ES^+) m/z 276.0623 $[M + Na]^+$ (required 276.0625).

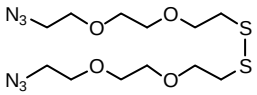


4.30. 1-azido-5-thioacetyl-triethylene glycol

To a solution of **4.29** (30.4g, 120mmol, 1eq) in dry DMF (150ml) was added portionwise over 15 minutes KSAc (20.5g, 180mmol, 1.5eq). After 15 minutes, a gel had formed within the reaction mixture. DMF (50ml) was added and the mixture stirred vigorously for 4 hours. TLC (ethyl acetate) showed complete consumption of starting material. The solvent was removed under high vacuum and the reaction mixture diluted with brine (300ml) and DCM (300ml). The aqueous layer was extracted with DCM (2 x 150ml) and the combined organics dried over $MgSO_4$. The

solvent was removed under reduced pressure to yield the desired compound as a yellow liquid (27.9g, 99%).

R_f 0.91 (ethyl acetate), ^{1}H NMR (400MHz, $CDCl_3$) δ ppm 2.31 (3H, s, SAc), 3.06 (2H, t, *J* 6.4Hz, $OCH_2C\underline{H}_2N_3$), 3.36 (2H, t, *J* 4.9Hz, $OCH_2C\underline{H}_2SAc$), 3.58 (2H, t, *J* 6.4Hz, $OC\underline{H}_2CH_2N_3$) 3.61 (4H, s, OCH_2CH_2O), 3.65 (2H, t, *J* 5.1Hz, $OC\underline{H}_2CH_2SAc$); ^{13}C NMR (101MHz, $CDCl_3$) δ ppm 28.8 (1C, $OCH_2\underline{C}H_2N_3$), 30.5 (1C, $CO\underline{C}H_3$), 50.6 (1C, $OCH_2\underline{C}H_2SAc$), 69.8 (1C, $O\underline{C}H_2CH_2SAc$), 70.0, 70.3, 70.5 (3 x 1C, $O\underline{C}H_2\underline{C}H_2O\underline{C}H_2CH_2SAc$), 195.5 (1C, C=O); FT-IR (thin film) υ 1692 (C=O), 2103 (N_3); HRMS (ES^+) 256.0727 $[M + Na]^+$ (required 256.0726).

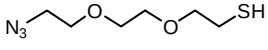


4.32. 1,18-diazido-3,6,13,16-tetraoxa-9,10-dithiaoctadecane

A solution of **4.30** (27.8g, 119mmol, 1eq) and NaOMe (7.09g, 131mmol, 1.1eq) in anhydrous methanol (50ml) was stirred at room temperature for 2 hours. TLC (ethyl acetate) was inadequate for resolving the starting material and product, hence mass spectrometry was used to follow the reaction. Electrospray mass spectrometry (positive ionisation mode) showed the complete consumption of starting material and the formation of the desired compound. The reaction was quenched with Dowex 50WX8 (H^+ form) until neutral (pH paper). The mixture was filtered and the product purified by silica gel chromatograph (10% ethyl acetate in petrol for 4CV, then ramped to 20% ethyl acetate in petrol over 2CV and maintained at this level for 3CV) as a colourless liquid (17.4g, 76%). However, over time, the product oxidized to the title compound.

R_f 0.22 (1 petrol : 9 ethyl acetate), ^{1}H NMR (400MHz, $CDCl_3$) δ ppm 2.86 (4H, t, *J* 6.7Hz, $OCH_2C\underline{H}_2N_3$), 3.35 (4H, t, *J* 5.1Hz, $OCH_2C\underline{H}_2S$), 3.62 (8H, s, OCH_2CH_2O), 3.64 (4H, t, *J* 5.1Hz, $OC\underline{H}_2CH_2S$), 3.71 (4H, t, *J* 6.6Hz, $OC\underline{H}_2CH_2N_3$); ^{13}C NMR (101MHz, $CDCl_3$) δ ppm 38.4 (2C, $OCH_2\underline{C}H_2N_3$), 50.6 (2C, $OCH_2\underline{C}H_2S$), 69.6 (2C, $O\underline{C}H_2CH_2N_3$), 70.0 (2C, $O\underline{C}H_2CH_2S$), 70.4, 70.6 (2 x 1C, OCH_2CH_2O); FT-IR (thin film) υ 2107 (N_3); HRMS (ES^+) m/z 403.1196 $[M + Na]^+$ (required 403.1193).

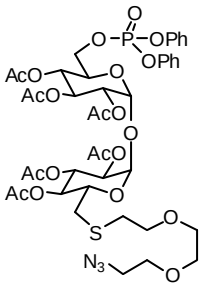
4.31. 1-azido-5-thiol-triethylene glycol



To a solution of **4.32** (1.00g, 2.63mmol, 1eq) in methanol (5ml), ethyl acetate (5ml) and 100mM NH_4HCO_3 buffer (5ml) was added dithiothreitol (405mg, 2.63mmol, 1eq). The reaction was stirred at room temperature for 10 minutes after which TLC (2 ethyl acetate : 1 petrol) showed complete consumption of the starting materials. The solvent was removed *in vacuo* and the

product purified by automated silica gel chromatography (4 petrol : 1 ethyl acetate) to give the desired compound as a colourless liquid (247mg, 25%).

R_f 0.28 (4 petrol : 2 ethyl acetate), ^{1}H NMR (400MHz, $CDCl_3$) δ ppm 1.60 (1H, br. s., SH), 2.69 (2H, t, *J* 6.3Hz, $OCH_2C\underline{H}_2N_3$), 3.38 (2H, t, *J* 5.1Hz, $OCH_2C\underline{H}_2SH$), 3.62 (2H, t, *J* 6.3Hz, $OC\underline{H}_2CH_2N_3$), 3.64 (4H, s, OCH_2CH_2O), 3.67 (2H, t, *J* 5.3Hz, $OC\underline{H}_2CH_2SH$); ^{13}C NMR (101MHz, $CDCl_3$) δ ppm 24.3 (1C, $OCH_2\underline{C}H_2N_3$), 50.7 (1C, $OCH_2\underline{C}H_2SH$), 70.1 (1C, $O\underline{C}H_2CH_2SH$), 70.2, 70.6 (2 x 1C, OCH_2CH_2O), 72.9 (1C, $O\underline{C}H_2CH_2N_3$); FT-IR (thin film) υ 2106 (N_3), 3003 (SH); HRMS (ES^+) m/z 214.0618 $[M + Na]^+$ (required 214.0621).



4.33. 6-O-(diphenoxyphosphoryl)-6'-deoxythio(triethylglycol-6dexoxyazido)-2,2',3,3',4,4'-acetyl-D-trehalose

To a solution of **4.31** (278mg, 1.46mmol, 1.05eq) in anhydrous THF (10ml) was added NaH (60% dispersion in mineral oil, 58mg, 1.46mmol, 1.05eq). The reaction mixture was stirred at room temperature for 15 minutes, until the production of hydrogen had ceased. The solution was then added to **4.10** (1.30g, 1.39mmol, 1eq) and the reaction mixture was irradiated in a microwave cavity at 80°C with a stream of cooling nitrogen at a maximum power of 300W for 40 minutes, after which TLC (4 ethyl : 3 petrol) showed the complete consumption of **4.31** and the formation of a product. The reaction mixture was quenched with glacial acetic acid (100μl) and the solvent removed *in vacuo*. The product was isolated by automated silica gel chromatography (12% ethyl acetate in petrol to 100% ethyl acetate over 10 column volumes) to give the desired compound as a colourless oil (488mg, 35%) as well as recovered starting material (360mg, 28%).

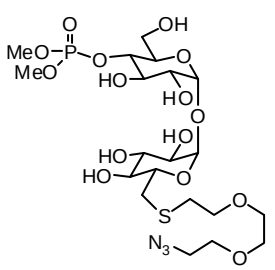
R_f 0.14 (1 petrol : 1 ethyl acetate), $[\alpha]_D^{25}$ +83.9 (c = 1.0, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.01, 2.03, 2.03, 2.04, 2.06, 2.14 (6 x 3H, 6 x s, 6 x OAc), 2.64-2.68 (2H, m, H6'a, H6'b), 2.72 (1H, dt, $J_{H7a-H7b}$ 13.2Hz, 2 x J_{H7b-H8} 6.6Hz, H7b), 2.79 (1H, dt, $J_{H7a-H7b}$ 13.6Hz, 2 x J_{H7a-H8} 6.6Hz, H7a), 3.39 (2H, t, $J_{H11-H12}$ 5.0Hz, H12), 3.60-3.66 (6H, m, H8, H9, H10), 3.68 (2H, t, $J_{H11-H12}$ 4.7Hz, H11), 4.03 (1H, ddd, $J_{H4'-H5'}$ 10.1Hz, $J_{H5'-H6'a}$ 6.0Hz, $J_{H5'-H6'b}$ 4.1Hz, H5'), 4.13 (1H, ddd, J_{H4-H5} 9.5Hz, J_{H5-H6a} 4.7Hz, J_{H5-H6b} 1.6Hz, H5), 4.23 (1H, ddd, $J_{H6a-H6b}$ 11.3Hz, $J_{H6b-31P}$ 5.6Hz, J_{H5-H6b} 2.0Hz, H6b), 4.30 (1H, ddd, $J_{H6a-H6b}$ 12.3Hz, $J_{H6a-31P}$ 6.6Hz, J_{H5-H6a} 5.7Hz, H6a), 4.99 (1H, t, $J_{H3'-H4'}$ 10.1Hz, $J_{H4'-H5'}$ 10.1Hz, H4'), 5.03 (1H, dd, $J_{H2'-H3'}$ 10.1Hz, $J_{H1'-H2'}$

2.4Hz, H2'), 5.04 (1H, t, $J_{H3\text{-}H4}$ 10.1Hz, $J_{H4\text{-}H5}$ 10.1Hz, H4), 5.05 (1H, dd, $J_{H2\text{-}H3}$ 10.4Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 5.08 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.19 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 5.43 (1H, t, $J_{H2'\text{-}H3'}$ 9.6Hz, $J_{H3'\text{-}H4'}$ 9.6Hz, H3'), 5.47 (1H, t, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H3\text{-}H4}$ 9.8Hz, H3), 7.21 (6H, ad, J 7.3Hz, ArH_{ortho}, ArH_{para}), 7.35 (4H, m, ArH_{meta}); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.5, 20.6, 20.7, 21.0 (4 x 1C, 4 x $CO\underline{C}H_3$), 20.7 (2C, 2 x $CO\underline{C}H_3$), 32.8 (1C, C7), 33.4 (1C, C6'), 50.6 (1C, C12), 66.6 (1C, d, $J_{C\text{-}31P}$ 4.8Hz, C6), 68.5 (1C, C2), 68.7 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C5), 69.2 (1C, C4), 69.6 (1C, C2'), 70.1 (2C, C3, C3'), 70.1 (1C, C11), 70.4, 70.6, 71.0 (3 x 1C, C8, C9, C10), 71.3 (1C, C5'), 71.5 (1C, C4'), 92.2 (1C, C1'), 92.3 (1C, C1), 120.0 (4C, d, $J_{C\text{-}31P}$ 4.8Hz, ArC_{ortho}), 125.5 (2C, ArC_{para}), 129.8 (4C, d, $J_{C\text{-}31P}$ 1.9Hz, ArC_{meta}), 150.3 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, ArC_{ipso}), 150.4 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, ArC_{ipso}), 169.4, 170.0 (2 x 2C, 4 x C=O),169.6, 169.7 (2 x 1C, 2 x C=O); FT-IR (thin film) 1221 (P=O), 1751 (C=O), 2107 (N_3); HRMS (ES^+) m/z 1022.2613 $[M + Na]^+$ (required 1022.2600).

4.34. 4-*O*-(dimethoxyphosphoryl)-6'-deoxythio(triethylglycol-6dexoxyazido)-D-trehalose

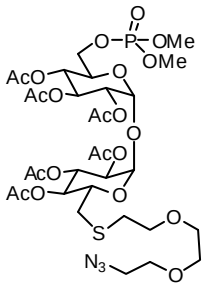
4.36. 6-*O*-(dimethoxyphosphoryl)-6'-deoxythio(triethylglycol-6dexoxyazido)-2,2',3,3',4,4'-acetyl-D-trehalose

A solution of **4.33** (160mg, 0.16mmol, 1wq) and CsF (0.24g, 1.6mmol, 10eq) in methanol (10ml) was heated under reflux for 17 hours, after which TLC (1 water : 4 isopropanol : 6 ethyl acetate) showed the formation of 2 products. The solvent was removed *in vacuo* and the products separated by automated silica gel chromatography (1 water : 4 isopropanol : 10 ethyl acetate for 4 column volumes, then ramped to 1 water : 4 isopropanol : 4 ethyl acetate over 5 column volumes) to give the side product **4.34** (32.3mg, 32%) and the desired compound **4.35** (67.0mg, 67%) as white amorphous solids. The desired compound, **4.35**, was dissolved in anhydrous pyridine (10ml) and acetic anhydride (0.27ml, 2.9mmol, 18eq) was added. The reaction mixture was stirred at room temperature for 10 hours after which TLC (ethyl acetate) showed complete consumption of starting material. The solvent was removed under reduced pressure and the crude product was co-evaporated with toluene to remove pyridine (x3). The product was dried under high vacuum to give the title compound **4.36** as a colourless oil (94.5mg, 68% from **4.33**).



*Compound **4.34***

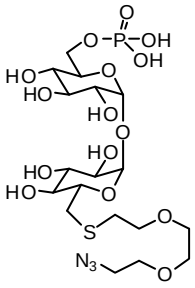
R_f 0.54 (1 water : 4 isopropanol : 6 ethyl acetate); ^{1}H NMR (500MHz, D_2O) δ ppm 2.66 (1H, dd, $J_{H6'a\text{-}H6'b}$ 14.2Hz, $J_{H5'\text{-}H6'b}$ 8.8Hz, H6'b), 2.76 (1H, t, $J_{H7\text{-}H8}$ 6.1Hz, H7), 3.02 (1H, dd, $J_{H6'a\text{-}H6'b}$ 14.2Hz, $J_{H5'\text{-}H6'a}$ 2.2Hz, H6'a), 3.29 (1H, t, $J_{H3'\text{-}H4'}$ 9.5Hz, $J_{H4'\text{-}H5'}$ 9.5Hz, H4'), 3.42 (1H, t, $J_{H11\text{-}H12}$ 5.0Hz, H12), 3.58 (1H, dd, $J_{H2'\text{-}H3'}$ 9.8Hz, $J_{H1'\text{-}H2'}$ 3.8Hz, H2'), 3.59 (1H, t, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H3\text{-}H4}$ 9.5Hz, H3), 3.63-3.65 (6H, m, H9, H10, H11), 3.66 (2H, t, $J_{H7\text{-}H8}$ 6.3Hz, H8), 3.69 (1H, dd, $J_{H6a\text{-}H6b}$ 12.9Hz, $J_{H5\text{-}H6b}$ 5.7Hz, H6b), 3.75 (3H, d, $J_{H\text{-}31P}$ 11.3Hz, OMe), 3.75-3.80 (2H, m, H5, H6a), 3.76 (3H, d, $J_{H\text{-}31P}$ 11.3Hz, OMe), 3.78 (1H, t, $J_{H2'\text{-}H3'}$ 9.6Hz, $J_{H3'\text{-}H4'}$ 9.6Hz, H3'), 3.79 (1H, dd, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H1\text{-}H2}$ 2.1Hz, H2), 3.85 (1H, td, $J_{H4'\text{-}H5'}$ 9.4Hz, $J_{H5'\text{-}H6'b}$ 9.4Hz, $J_{H5'\text{-}H6'a}$ 2.2Hz, H5'), 4.48 (1H, q, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, $J_{H4\text{-}31P}$ 9.5Hz, H4), 5.09 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.18 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 31.1 (1C, C7), 32.9 (1C, C6'), 50.1 (1C, C12), 55.1 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, OMe), 55.1 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, OMe), 60.2 (1C, C6), 68.3 (1C, d, $J_{C\text{-}31P}$ 1.9Hz, C3) 69.2, 69.3, 69.5 (3 x 1C, C9, C10, C11), 69.5 (1C, C8), 69.7 (1C, d, $J_{C\text{-}31P}$ 3.8Hz, C5), 71.1 (1C, C2'), 71.1 (1C, C5'), 72.0 (1C, C3'), 72.4 (1C, C2), 72.7 (1C, C4'), 81.5 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C4), 92.9 (1C, C1'), 93.3 (1C, C1); ^{31}P NMR (162MHz, D_2O) δ ppm 0.9 ($P(O)(OMe)_2$); FT-IR (KBr disc) 1260 (P=O), 2114 (N_3), 3424 br (OH); HRMS (ES^+) m/z 646.1656 $[M + Na]^+$ (required 646.1653).



*Compound **4.36***

R_f 0.27 (ethyl acetate), $[\alpha]_D^{25}$ +60.5 (c = 1.0, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.03 (2 x 3H, s, 2 x OAc), 2.07, 2.07, 2.09, 2.15 (4 x 3H, 4 x s, 4 x OAc), 2.67-2.68 (2H, m, H6'a, H6'b) 2.72 (1H, dt, $J_{H7a\text{-}H7b}$ 13.6Hz, 2 x $J_{H7b\text{-}H8}$ 6.6Hz, H7b), 2.79 (1H, dt, $J_{H7a\text{-}H7b}$ 13.6Hz, $J_{H7a\text{-}H8}$ 6.6Hz, H7a), 3.40 (2H, t, $J_{H11\text{-}H12}$ 5.0Hz, H12), 3.60-3.67 (6H, m, H8, H9, H10), 3.68 (2H, t, $J_{H11\text{-}H12}$ 5.2Hz, H11), 3.77 (3H, d, $J_{H\text{-}31P}$ 11.0Hz, OMe), 3.77 (3H, d, $J_{H\text{-}31P}$ 11.3Hz, OMe), 4.04 (1H, ddd, $J_{H6a\text{-}H6b}$ 10.1Hz, $J_{H6b\text{-}31P}$ 5.4Hz, $J_{H5\text{-}H6b}$ 3.8Hz, H6b), 4.05-4.08 (2H, m, H5, H5'), 4.11 (1H, dt, $J_{H6a\text{-}H6b}$ 10.7Hz, $J_{H6a\text{-}31P}$ 5.0Hz, $J_{H5\text{-}H6a}$ 5.0Hz, H6a), 5.00 (1H, t, $J_{H3'\text{-}H4'}$ 9.8Hz, $J_{H4'\text{-}H5'}$ 9.8Hz, H4'), 5.05 (1H, dd, $J_{H2'\text{-}H3'}$ 10.2Hz, $J_{H1'\text{-}H2'}$ 3.9Hz, H2'), 5.09 (1H, t, $J_{H3\text{-}H4}$ 9.6Hz, $J_{H4\text{-}H5}$ 9.6Hz, H4), 5.13 (1H, dd, $J_{H2\text{-}H3}$ 10.4Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 5.28 (1H, d, $J_{H1'\text{-}H2'}$ 4.1Hz, H1'), 5.30 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 5.46 (1H, t, $J_{H2'\text{-}H3'}$ 9.5Hz, $J_{H3'\text{-}H4'}$

9.5Hz, H3'), 5.49 (1H, t, J_{H2-H3} 9.5Hz, J_{H3-H4} 9.5Hz, H3); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.5, 20.6, 20.7, 20.7, 20.7, 21.0 (6 x 1C, 6 x CO<u>C</u>H$_3$), 32.8 (1C, C7), 33.4 (1C, C6'), 50.6 (1C, C12), 54.6 (2C, d, J_{C-31P} 6.7Hz, 2 x OMe), 65.2 (1C, d, J_{C-31P} 4.8Hz, C6), 68.3 (1C, C4), 68.8 (1C, d, J_{C-31P} 8.6Hz, C5), 69.2 (1C, C2), 69.8 (1C, C2'), 69.9, 70.1 (2 x 1C, C3, C3), 70.1 (1C, C11), 70. 4, 70.6, 71.0 (3 x 1C, C8, C9, C10), 71.3 (1C, C5'), 71.6 (1C, C4'), 92.1 (1C, C1'), 92.4 (1C, C1), 169.5, 169.6, 170.0, 170.0 (4 x 1C, 4 x C=O), 169.7 (2C, 2 x C=O); ^{31}P NMR (162MHz, $CDCl_3$) δ ppm 1.3 ($P(O)(OMe)_2$); FT-IR (thin film) υ 1221 (P=O), 1753 (C=O), 2110 (N_3); HRMS (ES^+) m/z 898.2312 $[M + Na]^+$ (required 898.2287).



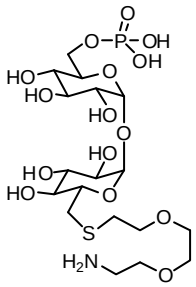
4.38. 6-O-(dihydroxyphosphoryl)-6'-deoxythio(triethylglycol-6-dexoxyazido)-D-trehalose

To a solution of **4.36** (90mg, 0.10mmol, 1eq) in anhydrous dioxane (5ml) was added TMSBr (29µl, 0.22mmol, 2.1eq) and the reaction mixture heated to 50°C. After 16 hours, TLC (1 water : 4 isopropanol : 4 ethyl acetate) showed complete consumption of starting material and the formation of a single product. The reaction was quenched with methanol (1ml). The solvent was removed under reduced pressure and the crude product dried under high vacuum for 4 hours. The crude product was dissolved in dry methanol (10ml) and NaOMe (~10mg) was added until the reaction mixture was pH 10 (pH paper). The reaction mixture was stirred at room temperature for 6 hours, after which TLC (1 water : 3 NH_4OH : 5 ethanol) showed the complete consumption of starting material and the formation of a single product. The reaction was quenched by addition of solid CO_2 until the reaction mixture was at pH 7 (pH paper). The crude product was partitioned between water (3ml) and ethyl acetate (3ml). The aqueous layer was retained and washed with ethyl acetate (6 x 2ml). The aqueous layer was concentrated under reduced pressure and the crude product dissolved in water (1ml). The solution was passed through a C_{18} SepPak® Light cartridge, pre-equilibrated with water, and the cartridge was eluted with water (4ml). All effluent was collected and lyophilized to yield the desired compound as an amorphous white solid (60.1mg, 98%).

R_f 0.31 (1 water : 3 NH_4OH : 5 ethanol), $[\alpha]_D^{25}$ +56.9 (c = 0.5, H_2O); ^{1}H NMR (500MHz, D_2O) δ ppm 2.67 (1H, dd, $J_{H6'a-H6'b}$ 14.2Hz, $J_{H5'-H6'b}$ 8.2Hz, H6'b), 2.76 (2H, t, J_{H7-H8} 6.3Hz, H7), 3.00 (1H, dd, $J_{H6'a-H6'b}$ 14.2Hz, $J_{H5'-H6'a}$ 2.2Hz, H6'a), 3.30

(1H, t, $J_{H3'\text{-}H4'}$ 9.5Hz, $J_{H4'\text{-}H5'}$ 9.5Hz, H4'), 3.42 (2H, t, $J_{H11\text{-}H12}$ 4.7Hz, H12), 3.54 (1H, t, $J_{H3\text{-}H4}$ 9.6Hz, $J_{H4\text{-}H5}$ 9.6Hz, H4), 3.57 (1H, dd, $J_{H2'\text{-}H3'}$ 10.1Hz, $J_{H1'\text{-}H2'}$ 4.1Hz, H2'), 3.59 (1H, dd, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H1\text{-}H2}$ 4.1Hz, H2), 3.62-3.68 (8H, m, H8, H9, H10, H11), 3.73 (1H, t, $J_{H2'\text{-}H3'}$ 9.5Hz, $J_{H3'\text{-}H4'}$ 9.5Hz, H3'), 3.76 (1H, t, $J_{H2\text{-}H3}$ 9.6Hz, $J_{H3\text{-}H4}$ 9.6Hz, H3), 3.79 (1H, m, H5), 3.82-3.87 (2H, m, H5', H6b), 3.94 (1H, ddd, $J_{H6a\text{-}H6b}$ 12.0Hz, $J_{H6a\text{-}31P}$ 7.9Hz, $J_{H5\text{-}H6a}$ 4.4Hz, H6a), 5.10 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.11 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1)**;** ^{13}C NMR (126MHz, D_2O) δ ppm 31.2 (1C, C7), 32.9 (1C, C6'), 50.1 (1C, C6), 62.4 (1C, d, $J_{C\text{-}31P}$ 3.8Hz, C6), 69.1 (1C, C4), 69.2, 69.3, 69.5, 69.5 (4 x 1C, C8, C9, C10, C11), 71.0 (1C, C2'), 71.1 (1C, C5'), 71.2 (1C, C2), 71.8 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, C5), 72.2 (1C, C3), 72.3 (1C, C3'), 72.6 (1C, C4'), 93.1 (1C, C1), 93.6 (1C, C1'); ^{31}P NMR (162MHz, D_2O) δ ppm 4.6 ($P(O)(OH)_2$); FT-IR (KBr disc) υ 1103 (P=O), 2113 (N_3), 3443 (OH); HRMS (ES^-) 594.1368 $[M - H]^-$ (required 594.1368).

4.39. 6-*O*-(dihydroxyphosphoryl)-6'-deoxythio(triethylglycol-6-dexoxyazido)-D-trehalose



A suspension of **4.38** (60mg, 0.10mmol, 1eq) and Pd/C (22mg, 0.01mmol, 0.1eq) was vigorously stirred in water (10ml). The reaction mixture was degassed under vacuum and charged with hydrogen (x3) and then stirred at room temperature for 3 hours after which TLC (1 water : 3 NH_4OH : 5 ethanol) showed the complete consumption of starting material. The reaction was filtered through Celite® and the solvent removed *in vacuo* and lyophilized to yield the target compound as a white amorphous solid (57mg, 99%).

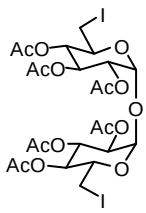
R_f 0.19 (1 water : 3 NH_4OH : 5 ethanol), $[\alpha]_D^{25}$ +65.5 (c = 0.5, H_2O); ^{1}H NMR (500MHz, D_2O) δ ppm 2.66 (1H, dd, $J_{H6'a\text{-}H6'b}$ 14.2Hz, $J_{H5'\text{-}H6'b}$ 8.5Hz, H6'b), 2.78 (1H, t, $J_{H7\text{-}H8}$ 6.1Hz, H7), 2.91 (1H, t, $J_{H11\text{-}H12}$ 5.4Hz, H12), 3.00 (1H, dd, $J_{H6'a\text{-}H6'b}$ 14.2Hz, $J_{H5'\text{-}H6'a}$ 2.2Hz, H6'a), 3.30 (1H, t, $J_{H3'\text{-}H4'}$ 9.5Hz, $J_{H4'\text{-}H5'}$ 9.5Hz, H4'), 3.54 (1H, t, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, H4), 3.59 (1H, dd, $J_{H2'\text{-}H3'}$ 9.9Hz, $J_{H1'\text{-}H2'}$ 3.1Hz, H2'), 3.58 (1H, dd, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H1\text{-}H2}$ 2.5Hz, H2), 3.59-3.64 (6H, m, H9, H10, H11), 3.66 (2H, t, $J_{H7\text{-}H8}$ 6.1Hz, H8), 3.73 (1H, t, $J_{H2'\text{-}H3'}$ 9.5Hz, $J_{H3'\text{-}H4'}$ 9.5Hz, H3'), 3.77 (1H, t, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H3\text{-}H4}$ 9.5Hz, H3), 3.78 (1H, m, H5), 3.84 (1H, td, $J_{H4'\text{-}H5'}$ 9.5Hz, $J_{H5'\text{-}H6'b}$ 9.5Hz, $J_{H5'\text{-}H6'a}$ 1.9Hz, H5'), 3.85 (1H, m, H6b), 3.94 (1H, ddd, $J_{H6a\text{-}H6b}$ 12.0Hz, $J_{H6a\text{-}31P}$ 7.9Hz, $J_{H5\text{-}H6a}$ 4.4Hz, H6a), 5.11 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.13 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 31.3 (1C, C7), 32.9 (1C,

C6'), 39.5 (1C, C12), 62.4 (1C, d, $J_{C\text{-}31P}$ 4.8Hz, C6), 69.0 (1C, C4), 69.3, 69.4 (x2), 69.4 (4C, C8, C9, C10, C11), 71.0 (1C, C2'), 71.2 (1C, C5'), 71.2 (1C, C2), 71.8 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, C5), 72.2 (1C, C3), 72.3 (1C, C3'), 72.6 (1C, C4'), 92.9 (1C, C1), 93.4 (1C, C1'); ^{31}P NMR (162MHz, D_2O) δ ppm 4.6 ($P(O)(OH)_2$); FT-IR (KBr disc) υ 1102 br (P=O, C-N), 1632 (1^o NH_2), 3430 br (OH, NH_2); HRMS (ES^+) 592.1436 $[M + Na]^+$ (required 592.1435).

4.9. 2,2',3,3',4,4'-O-acetyl-6,6'diiodo-D-trehalose

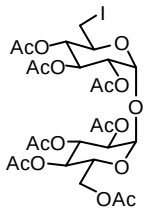
4.42. 2,2',3,3',4,4',6'-O-acetyl-6-iodo-D-trehalose

To a suspension of trehalose (5.0g, 14.6mmol, 1eq) in anhydrous pyridine (20ml) was added dropwise diphenyl chlorophosphate (3.03ml, 14.6mmol, 1eq). The reaction was allowed to stir at room temperature for 15 hours, after which TLC (1 water : 4 isopropanol : 4 ethyl acetate) showed the presence of two products. Acetic anhydride (30ml) was added to the reaction mixture and stirred at room temperature for 12 hours. The reaction was cooled to 0°C and quenched by the addition of methanol (10ml). The solvents were removed *in vacuo* and the crude residue was dissolved in ethyl acetate (250ml) and washed with 1M HCl (5 x 100ml), saturated $NaHCO_3$ solution (2 x 100ml), brine (100ml) and dried over $MgSO_4$. The solvents were removed under reduced pressure and the crude solid dried under high vacuum for 12 hours. The crude products were dissolved in dry THF (20ml) and tetrabutylammoinum iodide (16.2g, 43.9mmol, 3eq) was added. All solids were dissolved by gentle heating and the resultant solution was divided between two microwave tubes. The reaction mixtures were irradiated with microwaves at a maximum of 300W for 60 minutes with cooling nitrogen to a temperature of 120°C. TLC (2 ethyl acetate : 1 petrol) showed the presence of three products. The solvents were removed *in vacuo* and the crude residue re-dissolved in ethyl acetate (250ml). The organic solvent was washed with brine (5 x 100ml) and dried over $MgSO_4$. The compounds were separated by automated silica gel chromatography (10% ethyl acetate in petrol to 100% ethyl acetate over 12 column volumes) to give the desired compound **4.42** as a colourless oil (1.23g, 11%). Compound **4.9** was recrystallized from ethyl acetate as colourless needles (3.41g, 29%).



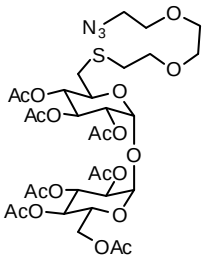
*Compound **4.9***

R_f 0.62 (2 ethyl acetate : 1petrol), $[\alpha]_D^{25}$ +72.3 (c = 1.0, $CHCl_3$), m.p. 152-157°C (ethyl acetate); ^1H NMR (500MHz, $CDCl_3$) δ ppm 2.01, 2.07, 2.14 (3 x 6H, 3 x s, 6 x OAc), 3.06 (2H, dd, $J_{H6\text{-}H6'}$ 11.0Hz, $J_{H5\text{-}H6'}$ 9.1Hz, H6'), 3.22 (2H, dd, $J_{H6\text{-}H6'}$ 10.9Hz, $J_{H5\text{-}H6}$ 2.4Hz, H6), 3.94 (2H, td, $J_{H4\text{-}H5}$ 9.5Hz, $J_{H5\text{-}H6'}$ 9.5Hz, $J_{H5\text{-}H6}$ 2.2Hz, H5), 4.88 (2H, t, $J_{H3\text{-}H4}$ 9.6Hz, $J_{H4\text{-}H5}$ 9.6Hz, H4), 5.19 (2H, dd, $J_{H2\text{-}H3}$ 10.4Hz, $J_{H1\text{-}H2}$ 3.8Hz, H2), 5.41 (2H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 5.47 (2H, t, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H3\text{-}H4}$ 10.1Hz, H3); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 2.5 (2C, C6), 20.6, 20.7, 21.2 (3 x 2C, 6 x $\underline{C}H_3CO$), 69.2 (2C, C2), 69.7 (2C, C3), 69.9 (2C, C5), 72.3 (2C, C4), 91.7 (2C, C1), 169.4, 169.5, 169.9 (3 x 2C, 6 x $CH_3\underline{C}O$); FT-IR (ATR) υ 1760 (C=O); HRMS (ES^+) m/z 836.9728 $[M + Na]^+$ (required 836.9723).



*Compound **4.42***

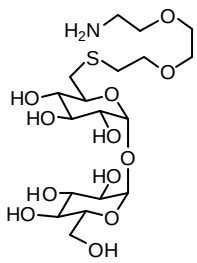
R_f 0.47 (2 ethyl acetate : 1 petrol), $[\alpha]_D^{25}$ +111.3 (c = 1.0, $CHCl_3$); ^1H NMR (500MHz, $CDCl_3$) δ ppm 2.03, 2.04, 2.05, 2.09 (4 x 3H, 4 x s, 4 x OAc), 2.09 (6H, s, 2 x OAc), 2.15 (3H, s, C6'-OAc), 3.07 (1H, dd, $J_{H6a\text{-}H6b}$ 11.0Hz, $J_{H5\text{-}H6b}$ 8.8Hz, H6b), 3.23 (1H, dd, $J_{H6a\text{-}H6b}$ 11.0Hz, $J_{H5\text{-}H6a}$ 2.5Hz, H6a), 3.94 (1H, td, $J_{H4\text{-}H5}$ 9.4Hz, $J_{H5\text{-}H6b}$ 9.4Hz, $J_{H5\text{-}H6a}$ 2.4Hz, H5), 4.02 (1H, td, $J_{H4'\text{-}H5'}$ 9.8Hz, $J_{H5'\text{-}H6'a}$ 9.8Hz, $J_{H5'\text{-}H6'b}$ 2.2Hz, H5'), 4.04 (1H, dd, $J_{H6'a\text{-}H6'b}$ 12.3Hz, $J_{H5'\text{-}H6'b}$ 1.9Hz, H6'b), 4.22 (1H, dd, $J_{H6'a\text{-}H6'b}$ 12.6Hz, $J_{H5'\text{-}H6'a}$ 6.6Hz, H6'a), 4.89 (1H, t, $J_{H3\text{-}H4}$ 9.6Hz, $J_{H4\text{-}H5}$ 9.6Hz, H4), 5.05 (1H, t, $J_{H3'\text{-}H4'}$ 10.4Hz, $J_{H4'\text{-}H5'}$ 10.4Hz, H4'), 5.05 (1H, dd, $J_{H2\text{-}H3}$ 9.1Hz, $J_{H1\text{-}H2}$ 2.5Hz, H2), 5.19 (1H, dd, $J_{H2'\text{-}H3'}$ 10.2Hz, $J_{H1'\text{-}H2'}$ 3.9Hz, H2'), 5.35 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 5.38 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 5.48 (1H, dd, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H3\text{-}H4}$ 9.1Hz, H3), 5.51 (1H, dd, $J_{H2'\text{-}H3'}$ 10.1Hz, $J_{H3'\text{-}H4'}$ 9.5Hz, H3'); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 2.5 (1C, C6), 20.5 (1C, $CO\underline{C}H_3$), 20.6, 20.6 (2 x 2C, 2 x $CO\underline{C}H_3$), 20.7 (1C, $CO\underline{C}H_3$), 21.2 (1C, C6'-$CO\underline{C}H_3$), 61.8 (1C, C6'), 68.2 (1C, C5'), 68.6 (1C, C2), 69.1 (1C, C2'), 69.4 (1C, C3'), 69.7 (1C, C5), 70.0 (1C, C4'), 70.2 (1C, C3), 72.3 (1C, C4), 91.6 (1C, C1), 92.1 (1C, C1'), 169.5, 169.6, 169.6, 169.9, 169.9, 170.6 (7C, 7 x C=O); FT-IR (ATR) υ 1753 (C=O); HRMS (ES^+) m/z 769.0816 $[M + Na]^+$ (required 769.0816).



4.43. 6-deoxythio(triethylglycol-6-dexoxyazido)-2,2',3,3',4,4',6'-acetyl-D-trehalose

To a solution of **4.42** (250mg, 1.31mmol, 1eq) in anhydrous THF (7ml) was added NaH (60% dispersion in mineral oil, 52mg, 1.31mmol, 1eq). The reaction mixture was stirred at room temperature for 15 minutes, until the production of hydrogen had ceased. The solution was then added to **4.31** (980mg, 1.31mmol, 1eq) and the reaction mixture was irradiated in a microwave cavity at 80°C with a stream of cooling nitrogen at a maximum power of 300W for 40 minutes, after which TLC (2 ethyl : 1 petrol) showed the complete consumption of starting materials. The reaction mixture was quenched with glacial acetic acid (100µl) and the solvent removed *in vacuo*. The product was isolated by automated silica gel chromatography (50% ethyl acetate in petrol to 100% ethyl acetate over 10 column volumes) to give the desired compound as an amorphous white solid (386mg, 36%).

R_f 0.35 (2 ethyl acetate : 1 petrol), $[\alpha]_D^{25}$ +113.0 (c = 0.71, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.03, 2.03, 2.05, 2.06, 2.08, 2.09 (6 x 3H, 6 x s, 6 x OAc), 2.14 (3H, s, C6'-OAc), 2.65-2.68 (2H, m, H6a, H6b), 2.72 (1H, dt, $J_{H7a-H7b}$ 13.6Hz, 2 x J_{H7b-H8} 6.6Hz, H7b), 2.79 (1H, dt, $J_{H7a-H7b}$ 13.6Hz, 2 x J_{H7a-H8} 6.6Hz, H7a), 3.40 (2H, t, $J_{H11-H12}$ 5.0Hz, H12), 3.60-3.66 (6H, m, H8, H9, H10), 3.68 (2H, t, $J_{H11-H12}$ 5.0Hz, H11), 4.02 (1H, dd, $J_{H6'a-H6'b}$ 12.9Hz, $J_{H5'-H6'b}$ 1.9Hz, H6'b), 4.02-4.07 (2H, m, H5, H5'), 4.22 (1H, dd, $J_{H6'a-H6'b}$ 12.6Hz, $J_{H5-H6'b}$ 6.0Hz, H6'b), 4.99 (1H, t, J_{H3-H4} 9.5Hz, J_{H4-H5} 9.5Hz, H4), 5.02 (1H, dd, J_{H2-H3} 10.4Hz, J_{H1-H2} 3.8Hz, H2), 5.04 (1H, t, $J_{H3'-H4'}$ 10.1Hz, $J_{H4'-H5'}$ 10.1Hz, H4'), 5.13 (1H, dd, $J_{H2'-H3'}$ 10.4Hz, $J_{H1'-H2'}$ 3.8Hz, H2'), 5.29 (1H, d, J_{H1-H2} 3.8Hz, H1), 5.29 (1H, d, $J_{H1'-H2'}$ 3.8Hz, H1'), 5.47 (1H, dd, J_{H2-H3} 10.1Hz, J_{H3-H4} 9.1Hz, H3), 5.49 (1H, dd, $J_{H2'-H3'}$ 10.2Hz, $J_{H3'-H4'}$ 9.3Hz, H3'); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.6 (1C, CO$\underline{C}H_3$), 20.6 (2C, 2 x CO$\underline{C}H_3$), 20.6, 20.7, 20.7 (3 x 1C, 3 x CO$\underline{C}H_3$), 21.0 (1C, C6'-CO$\underline{C}H_3$), 32.8 (1C, C7), 33.3 (1C, C6), 50.6 (1C, C12), 61.8 (1C, C6'), 68.2 (1C, C5'), 68.6 (1C, C4'), 69.3 (1C, C2'), 69.9 (1C, C2), 70.0 (1C, C4), 70.1 (1C, C3 or C3'), 70.1 (1C, C3 or C3'), 70.4 (1C, C11), 70.6, 71.0, 71.3 (3 x 1C, C8, C9, C10), 71.6 (1C, C5), 91.7 (1C, C1 or C1'), 92.1 (1C, C1 or C1'), 169.5, 169.6, 169.7, 169.7, 169.9, 169.9, 170.6 (7 x 1C, 7 x C=O); FT-IR (ATR) υ 1751 (C=O), 2110 (N_3); HRMS (ES^+) m/z 832.2421 $[M + Na]^+$ (required 832.2421).



4.44. 6-deoxythio(triethylglycol-6-dexoxyazido)-D-trehalose

To a solution of **4.43** (380mg, 0.47mmol, 1eq) in anhydrous methanol (20ml) was added NaOMe (25mg, 0.47mmol, 1eq). The reaction mixture was stirred at room temperature for 15 hours, after which TLC (1 water : 3 isopropanol : 8 ethyl acetate) showed the presence of only one product. The reaction was quenched with Dowex 50WX8 (H^+ form) until neutral and the mixture filtered and washed with methanol. The reaction mixture was concentrated under reduced pressure to ~20ml. 5% Pd/C (100mg, 0.05mmol, 0.1eq) was added and the reaction vessel was repeatedly degassed under vacuum and charged with hydrogen. The reaction was vigorously stirred under an atmosphere of hydrogen for 3 hours, after which TLC (1 water : 3 isopropanol : 8 ethyl acetate) showed complete consumption of the intermediate. The reaction mixture was filtered through Celite® and the solvent removed *in vacuo*. The product was dissolved in water and lyophilized to give the target compound as an amorphous white solid (220mg, 96%).

R_f 0.5 (1 NH_4OH : 3 H_2O : 5 EtOH), $[\alpha]_D^{25}$ +104.0 (c = 0.5, H_2O); ^1H NMR (500MHz, D_2O) δ ppm 2.67 (1H, dd, $J_{H6a\text{-}H6b}$ 14.2Hz, $J_{H5\text{-}H6b}$ 8.2Hz, H6b), 2.78 (2H, t, $J_{H7\text{-}H8}$ 6.3Hz, H7), 3.01 (1H, dd, $J_{H6a\text{-}H6b}$ 14.2Hz, $J_{H5\text{-}H6a}$ 2.2 Hz, H6a), 3.14 (1H, t, $J_{H10\text{-}H11}$ 5.0Hz, H11), 3.31 (1H, t, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, H4), 3.36 (1H, t, $J_{H3'\text{-}H4'}$ 9.5Hz, $J_{H4'\text{-}H5'}$ 9.5Hz, H4'), 3.55 (1H, dd, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H1\text{-}H2}$ 4.1Hz, H2), 3.58 (1H, dd, $J_{H2'\text{-}H3'}$ 10.1Hz, $J_{H1'\text{-}H2'}$ 3.8Hz, H2'), 3.63-3.70 (9H, m, H6'b, H8, H9, H10, H11), 3.74 (1H, t, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H3\text{-}H4}$ 9.5Hz, H3), 3.75 (1H, t, $J_{H2'\text{-}H3'}$ 9.5Hz, $J_{H3'\text{-}H4'}$ 9.5Hz, H3'), 3.73-3.85 (1H, m, H5'), 3.78 (1H, dd, $J_{H6'a\text{-}H6'b}$ 12.0Hz, $J_{H5'\text{-}H6'a}$ 2.2Hz, H6a'), 3.85 (1H, ddd, $J_{H4\text{-}H5}$ 9.5Hz, $J_{H5\text{-}H6b}$ 8.2Hz, $J_{H5\text{-}H6a}$ 2.2Hz, H5), 5.09 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1), 5.14 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'); ^{13}C NMR (126MHz, D_2O) δ ppm 31.2 (1C, C7), 32.8 (1C, C6), 39.1 (1C, C12), 60.5 (1C, C6'), 66.4 (1C, C11), 69.3, 69.4, 69.5 (3 x 1C, C8, C9, C10), 69.7 (1C, C4'), 71.0 (2C, C2, C2'), 71.1 (1C, C5), 72.2 (1C, C5'), 72.4 (1C, C4), 72.6 (2C, C3, C3'), 93.0 (1C, C1), 93.3 (1C, C1'); FT-IR (ATR) υ 2881 (NH_2 bend overtone), 3299 (OH); HRMS (ES^+) m/z 490.1941 $[M + H]^+$ (required 490.1953).

General resin coupling reaction

Compounds **4.39** and **4.44** were coupled to *N*-hydro succinimide activated Sepharose®, pre-packed in a 5ml HiTrap® column (GE Healthcare). All steps were performed at room temperature. The column was eluted with 1mM HCl (10ml) at 1 ml/min. Compounds **4.39** and **4.44** (0.055mmol) were dissolved in 0.2M $NaHCO_3$ at pH 8.3 and eluted through the column at 0.5 ml/min. The effluent was cycled back and forth through the column at 0.5ml for 2 hours after which the column was pumped with 0.2M $NaHCO_3$ (15ml) and the total effluent was retained. The column was then washed with 0.5M ethanolamine at pH 8.3 at 5 ml/min (30ml); 0.1M acetate at pH 4 at 5 ml/min (30ml); and 0.5M ethanolamine at pH 8.3 at 5 ml/min (30ml). The column was capped and maintained at room temperature for 30 minutes, after which the column was eluted with 0.1M acetate at pH 4 at 5 ml/min (30ml); 0.5M ethanolamine at pH 8.3 at 5 ml/min (30ml); and 0.1M acetate at pH 4 at 5 ml/min (30ml). The column was then washed with 0.05M Na_2HPO_4 with 0.01% NaN_3 at pH 7, and stored at 4°C until use. Coupling efficiency was determined by measuring A_{220} of the effluent from the column containing NHS and comparing to a standardised curve.

4.11.4 Proteomics

Proteomics experiments were conducted by the Proteomics Facility at the Dunn School of Pathology, University of Oxford, UK. Excised bands were destained, treated with DTT and digested by trypsin. Samples were injected through a Dionex nano-HPLC system running a C_{18} column, which was coupled to a LTQ-Orbitrap mass spectrometer. Peptide mass data was compared against the genomes of *A. thaliana, B. oleracea* and other green plants using the Mascot search engine.[21] Successful protein hits were determined with the MOWSE scoring algorithm.[22]

4.11.5 Studies on recombinant 14-3-3

Expression of Omega isoform

The *A. thaliana* 14-3-3 Omega isoform was prepared in *E. Coli* using the pET-15b expression vector by Dr Lucia Primavesi (Rothamsted Research). The protein was purified by Ni^{2+} affinity chromatography, dialysed into 50mM HEPES and 150mM

NaCl at pH 7.5. and concentrated to 12.5mg/ml. The samples were stored at -80°C until use.

```
MGSSHHHHHH SSGLVPRGSH MASGREEFVY MAKLAEQAER YEEMVEFMEK VSAAVDGDEL 60
TVEERNLLSV AYKNVIGARR ASWRIISSIE QKEESRGNDD HVTAIREYRS KIETELSGIC 120
DGILKLLDSR LIPAAASGDS KVFYLKMKGD YHRYLAEFKT GQERKDAAEH TLAAYKSAQD 180
IANAELAPTH PIRLGLALNF SVFYYEILNS PDRACNLAKQ AFDEAIAELD TLGEESYKDS 240
TLIMQLLRDN LTLWTSDMQD DAADEIKEAA APKPTEEQQ                       279
```

The His-tag extension was cleaved using a Thrombin cleavage kit (Novagen) by Lucia Primavesi (Rothamsted Research). The His-tagged protein solution (300µl, 12.5mg/ml, 3.77mg) and thrombin (4.2µl, 0.9 units/µl, 3.8 units) were incubated in 10x cleavage buffer (150µl) made up to a total volume of 1.50ml with deionised water (MilliQ®) for 24 hours at room temperature, after which SDS-PAGE gel showed the formation of a single product. A strptavidin-agarose slurry (122µl) was added to the mixture and rocked gently at room temperature for 45 minutes. The reaction mixture was passed through a spin filter by gentle centrifugation, as per manufacturer's instructions. The flow though was collected and Ni^{2+}-NTA agarose (160µl, Quigen, pre-equilibrated in 50mM HEPES and 150mM NaCl at pH 7.5) was added. The mixture was shaken at room temperature for 10 minutes and the resin subsequently removed by spin filtration. The protein solution was concentrated by centrifugal filtration through a 100kDa cut-off membrane (Vivaspin®) and the concentration determined a standardised BCA assay to be 10.6mg/ml. The protein was frozen in liquid nitrogen and stored at -80°C until use.

```
GSHMASGREE FVYMAKLAEQ AERYEEMVEF MEKVSAAVDG DELTVEERNL LSVAYKNVIG 60
ARRASWRIIS SIEQKEESRG NDDHVTAIRE YRSKIETELS GICDGILKLL DSRLIPAAAS 120
GDSKVFYLKM KGDYHRYLAE FKTGQERKDA AEHTLAAYKS AQDIANAELA PTHPIRLGLA 180
LNFSVFYYEI LNSPDRACNL AKQAFDEAIA ELDTLGEESY KDSTLIMQLL RDNLTLWTSD 240
MQDDAADEIK EAAAPKPTEE QQ                                         262
```

Isothermal titration calorimetry

ITC experiments were performed on a MicroCal ITC200 using 250µl of protein solution. Ligand was injected every 120s. The initial injection was 1ul over 2s. All subsequent injections were 2ul over 4 seconds. Protein was used at between 0.01mM to 0.02mM in 50mM HEPES and 150mM NaCl at pH 7.5. The protein concentration was measured by BCA assay or using a Nanodrop spectrophotometer. Ligand was prepared in the same buffer at between 0.2mM to 1.0mM concentration. Protein was buffer exchanged into the same using a PD-10 or PD-G25 size exclusion column.

Experiments were performed at 35°C with stirring (1000rpm) using buffer as the reference cell. Data were analysed using Origin Pro.

Circular dichroism

CD data were recorded on a Chirascan CD spectrophotometer using a 400ul cuvette and a path length of 1.0mm. Protein was used in HEPES and 150mM NaCl at pH 7.5 at a concentration of 0.2mg/ml. Experiments were performed with 390ul of protein solution and 10ul of ligand (or water for the control). Ligand was added so it was 5 or 10 equivalent to the protein concentration. Denaturation experiments were performed by raising the temperature at 2°C per minute and monitoring at 210nm wavelength.

STD NMR

For NMR studies, the protein was buffer exchanged into deuterated 20mM sodium phosphate and 130mM Nacl buffer at pH* 7.5 using a PD MiniTrap G-25 size exclusion column (GE Healthcare). The column was pre-equilibrated with the said buffer for 30 minutes prior to use, and then used as per manufacturer's instructions. The protein eluant was collected and its concentration determined using a NanoDrop ND-1000 spectrophotometer and theoretical extinction coefficients based on the protein sequence. The protein concentration was adjusted to 40μM and 0.5ml was used for the NMR experiments. To this solution was added Tre-6-P (0.844mg, 4.0mM final concentration) and the mixture stored at 4°C for 15 hours to allow complete deuterium exchange of all exchangeable protons.

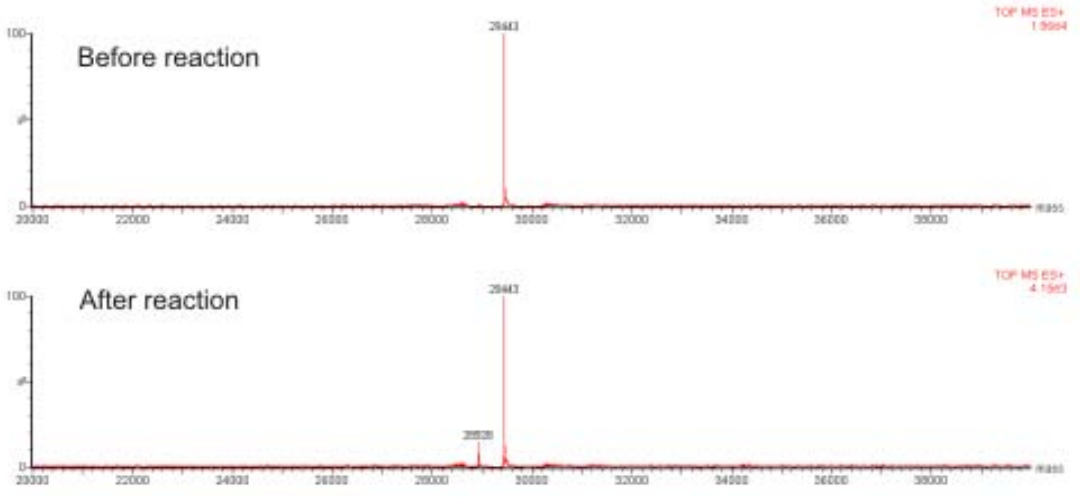
STD NMR experiments were performed by Dr Timothy Claridge (Oxford Univeristy) on a Bruker AVII 500 spectrometer equipped with a 5mm TXI probe. Acquisitions were performed using the standard STD pulse sequence with a shaped Q5 pulse train (50ms, 90°, saturation bandwidth of 120Hz, 4μs delay between pulses). The pulses were either off-resonance at 50ppm (where no NMR resonances of ligand or protein are present) or on-resonance at 0.82ppm (to saturate mostly Ile, Leu and Val methyl groups). Multiple scans were measured for each spectrum with a relaxation recovery time of 10 seconds to allow complete relaxation to thermal equilibrium. All experiments were performed on the same sample at either 288, 298 or 310K with a saturation time of either 2 or 5 seconds.

Mass spectrometric analysis of Tre-6-P degradation

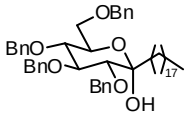
Analysis was performed on a Hichrom SiELC Primesep SM mixed mode anion exchange/C_{18} reverse phase column (150 x 4.6mm, 5µm) with an applied gradient from 0.5% formic acid in water to 1% formic acid in 3 water : 1 acetonitrile at a flow rate of 0.4ml/min. Eluants were fed directly into a Waters Quattro micro in negative mode either operating in scan mode over the mass range 150-450 (for determining reaction products) or in SIR mode centred at the monoisotopic masses of Tre-6-P, Glc-6-P or 2-deoxy-Glc-6-P with a detection width of 0.5Da and a dwell time of 1ms (for quantitative applications). The mass spectrometer was operated with a cone voltage of 35V, a source temperature of 100°C and desolvation temperature of 150°C. All chromatograms are shown after smoothing.

Mass spectrometric analysis of protein cleavage

HPLC separation was performed using a Phenomenex Jupiter C_4 column (4.6 x 250mm, 5µm) and a linear gradient of 5% acetonitrile in water to 100% acetonitrile (both solvents containing 0.1% formic acid) over 30 minutes at a flow rate of 0.6ml/min, with an injection volume of 20µl of the protein mixture at ~1mg/ml. Eluant detection was performed using a Waters Micromass LCT (reflecton TOF) mass spectrometer operating in the positive ionisation mode with a cone voltage of 30V. The mass spectrometer was calibrated against the myoglobin ion series. Raw mass spectra were deconvoluted using the MaxEnt algorithm in the MassLynx (Waters) software suite.



4.11.6 Microarray compounds

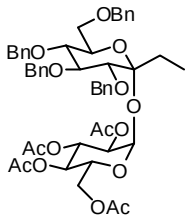


4.46. 2,3,4,6-tetra-*O*-benzyl-1-*C*-octadecyl-α-D-glucopyranose

The reaction was performed under standard Schlenk conditions using argon gas. To a suspension of magnesium turnings (2.0g, 84mmol, 1.4eq) and iodine (~5mg, cat.) in anhydrous THF (10ml) was added 5ml of 1-bromo-octadecane (20.1g, 60mmol, 1eq), pre-dissolved in THF (30ml). The reaction was initiated by gentle heating and was subsequently maintained under reflux by the dropwise addition of 1-bromo-octadecane. After addition was complete, the reaction mixture was allowed to cool to room temperature. The concentration of active Gringard was determined by titration against menthol, with 1,10-phenanthroline as indicator, to be 0.77M. The Grignard solution (3.0ml, 2.3mmol, 3.5eq) was added via syringe to a solution of **4.45** (0.36g, 0.66mmol, 1eq)[††] in anhydrous THF (10ml) at -78°C. The reaction was maintained at this temperature and after 3 hours TLC (2 petrol : 1ethyl acetate) showed the formation of a product, but with incomplete consumption of **4.45**. Addition of excess Gringard failed to drive the reaction to completion. The reaction was quenched by addition of 1M NH_4Cl (5ml). The solvent was removed *in vacuo* and the crude residue taken up in ethyl acetate (20ml) and washed with saturated $NaHCO_3$ solution (20ml), brine (20ml), dried over $MgSO_4$ and concentrated *in vacuo*. The product was purified by silica gel chromatography (5 petrol : 1 ethyl acetate) as a colourless oil (0.21g, 39%).

R_f 0.8 (5 petrol : 1 ethyl acetate), $[\alpha]_D^{24}$ + 10.2 (c = 1.0, $CHCl_3$); ^{1}H NMR (500MHz $CDCl_3$) δ ppm 0.89 (3H, t, $J_{CH3\text{-}CH2}$ 6.9Hz, CH_3), 1.27 (32H, as, $C_{16}H_{32}$), 1.62-1.67 (2H, m, $C\underline{H}_2C_{16}H_{33}$), 3.44 (1H, d, $J_{H2\text{-}H3}$ 9.1Hz, H2), 3.65 (1, t, $J_{H3\text{-}H4}$ 9.6Hz, $J_{H4\text{-}H5}$ 9.6Hz, H4), 3.67 (1H, dd, $J_{H6\text{-}H6'}$ 11.0Hz, $J_{H5\text{-}H6'}$ 2.2Hz, H6'), 3.77 (1H, dd, $J_{H6\text{-}H6'}$ 11.0Hz, $J_{H5\text{-}H6}$ 3.8Hz, H6), 3.99 (1H, ddd, $J_{H4\text{-}H5}$ 10.3Hz, $J_{H5\text{-}H6}$ 3.8Hz, $J_{H5\text{-}H6'}$ 2.0Hz, H5), 4.01 (1H, t, $J_{H2\text{-}H3}$ 9.1Hz, $J_{H3\text{-}H4}$ 9.1Hz, H3), 4.56 (1H, d, ^2J 11.9Hz, $OC\underline{H}_2Ph^1$), 4.62 (1H, d, ^2J 11.0Hz, $OC\underline{H}_2Ph^2$), 4.64 (1H, d, ^2J 11.9 Hz, $OC\underline{H}_2Ph^1$), 4.69 (1H, d, ^2J 11.3 Hz, $OC\underline{H}_2Ph^3$), 4.84 (1H, d, ^2J 11.0Hz, $OC\underline{H}_2Ph^2$), 4.88 (1H, d, ^2J 11.3Hz, $OC\underline{H}_2Ph^3$), 4.93 (1H, d, ^2J 11.1Hz, $OC\underline{H}_2Ph^4$), 4.93 (1H, d, ^2J 11.0Hz, $OC\underline{H}_2Ph^4$), 7.21-7.37 (20H, m, Ar-H); ^{13}C NMR (126MHz $CDCl_3$) δ ppm 14.1 (1C, CH_3), 22.5, 22.6, 29.3, 29.5, 29.6-29.7, 31.9, (16C, $CH_2\underline{C}_{16}H_{32}CH_3$), 38.6 (1C, $\underline{C}H_2C_{17}H_{35}$), 68.8 (1C, C6), 71.6 (1C, C5), 73.3, 74.8, 75.4, 75.6, (4 x $O\underline{C}H_2Ph$), 78.4 (1C, C4), 81.4 (1C, C2), 83.8 (1C, C3), 98.9 (1C, C1), 127.5, 127.6, 127.7, 127.9 (4 x 1C, ArC_{para}),

127.6-128.3 (8C, m, ArC_{ortho}) 126.3-128.4 (8C, m, ArC_{meta}), 137.9, 138.2, 138.4, 138.6 (4 x 1C, 4 x ArC_{ipso}); FT-IR (thin film) υ 2850 (C-H), 3032 br (OH); HRMS m/z (ES^+) 815.5221 $[M + Na]^+$ required 815.5227.

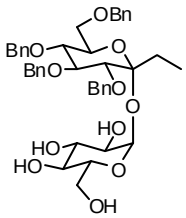


4.48. 2,3,4,6-tetra-O-benzyl-1-C-octadecyl-α-D-glucopyranosyl-2,3,4,6-O-acetyl-α-D-glucopyranoside

To a solution of **4.46** (100mg, 0.35mmol, 1.5eq) and **4.47** (150mg, 0.19mmol, 1eq) in anhydrous DCM (12ml) in the presence of 3Å molecular sieves at -40°C was added TMSOTf (5µl, 0.027mmol, 0.14eq). After stirring for 30 minutes, TLC (4 petrol : 1 ethyl acetate) revealed complete consumption of starting material. The reaction was quenched by addition of NEt_3 (1ml), filtered through Celite® and concentrated *in vacuo*. The crude residue was purified by silica gel chromatography (3 petrol : 1ethyl acetate) to afford the title compound as a colourless oil (132mg, 61%).

R_f 0.2 (4 petrol : 1 ethyl acetate), $[\alpha]_D^{25}$ +34.7 (c = 1.0, $CHCl_3$), ^1H NMR (500MHz $CDCl_3$) δ ppm 0.89 (3H, t, $J_{CH3\text{-}CH2}$ 6.9Hz, CH_3), 1.26 (30H, as, $(CH_2)_2C_{15}\underline{H}_{30}CH_3$), 1.51-1.62 (3H, m, $C\underline{H}_2C\underline{H}_2C_{16}H_{33}$), 1.83 (1H, at, $J_{CH2\text{-}CH2}$ 8.2 Hz, $C\underline{H}_2C_{17}H_{35}$), 1.98, 2.01, 2.03, 2.05 (4 x 3H, 4 x s, 4 x $COC\underline{H}_3$), 3.53-3.65 (4H, m, H2, H4, H6a, H6b), 3.83 (1H, dd, $J_{H6'a\text{-}H6'b}$ 12.5Hz, $J_{H5'\text{-}H6'b}$ 2.4Hz, H6'b), 3.90 (1H, ddd, $J_{H4\text{-}H5}$ 10.0Hz, $J_{H5\text{-}H6a}$ 4.3Hz, $J_{H5\text{-}H6b}$ 2.0Hz, H5), 4.10 (1H, dd, $J_{H6'a\text{-}H6'b}$ 12.8Hz, $J_{H5'\text{-}H6'b}$ 4.9Hz, H6'b), 4.14 (1H, t, $J_{H2\text{-}H3}$ 9.5Hz, $J_{H3\text{-}H4}$ 9.5Hz, H3), 4.36 (1H, ddd, $J_{H4'\text{-}H5'}$ 10.3Hz, $J_{H5'\text{-}H6'b}$ 4.7Hz, $J_{H5'\text{-}H6'a}$ 2.4Hz, H5'), 4.49 (1H, d, ^2J 12.0Hz, $OC\underline{H}_2Ph^1$), 4.56 (1H, d, ^2J 12.3Hz, $OC\underline{H}_2Ph^1$), 4.61 (2H, d, ^2J 10.7Hz, 2 x $OC\underline{H}_2Ph^2$), 4.86 (1H, d, ^2J 11.3Hz, $O\underline{C}H_2Ph^3$), 4.94 (1H, d, ^2J 11.3Hz, $O\underline{C}H_2Ph^3$), 4.96 (1H, d, ^2J 11.0Hz, $O\underline{C}H_2Ph^4$), 4.99 (1H, d, ^2J 11.3Hz, $O\underline{C}H_2Ph^4$), 5.05 (1H, t, $J_{H3'\text{-}H4'}$ 10.1Hz, $J_{H4'\text{-}H5'}$ 10.1Hz, H4), 5.05 (1H, dd, $J_{H2'\text{-}H3'}$ 10.1Hz, $J_{H1'\text{-}H2'}$ 3.5Hz, H2'), 5.38 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1), 5.58 (1H, t, $J_{H2'\text{-}H3'}$ 9.6Hz, $J_{H3'\text{-}H4'}$ 9.6Hz, H3'), 7.19-7.36 (20H, m, ArH); ^{13}C NMR (126MHz $CDCl_3$) δ ppm 14.1 (1C, CH_3), 20.6, 22.6, 25.7, 29.3-29.7, 31.9, 32.8, (16C, $CH_2\underline{C}_{16}H_{32}CH_3$), 34.2 (1C, $\underline{C}H_2C_{17}H_{35}$), 61.9 (1C, C6'), 67.1 (1C, C5'), 68.6 (1C, C4'), 68.6 (1C, C6), 70.4 (1C, C2'), 70.7 (1C, C3'), 72.2 (1C, C5), 73.3, 74.9, 74.9, 75.5 (4 x 1C, 4 x $O\underline{C}H_2Ph$), 78.4 (1C, C4), 79.6 (1C, C2), 82.9 (1C, C3), 88.7 (1C, C1'), 103.6 (1C, C1), 127.4, 127.5, 127.7, 127.9 (4 x 2C, ArC_{ortho}), 127.5, 127.5, 127.6, 127.8 (4 x 1C, ArC_{para}), 128.3, 128.3, 128.3, 128.4 (4 x 2C, ArC_{meta}), 138.3, 138.3, 138.6,138.6 (4 x 1C, ArC_{ipso}), 169.7, 169.8, 170.2, 170.6 (4 x 1C, 4 x C=O); IR

(thin film): FT-IR (thin film) 1751 (C=O), 2955 (C-H), 3028 br (OH); HRMS m/z (ES^+) $[M + Na]^+$ peaks 1145.59 (100%), 2246.60 (74%), 1147.60 (29%), 1148.60 (8%), 1149.60 (2%) required 1145.62 (100%), 1146.62 (73%), 1147.62 (29%), 1148.63 (8%), 1149.63 (2%).

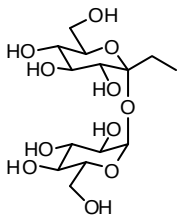


4.49. 1-octadecyl-2,3,4,6-O-benzyl-D-trehalose

Compound **4.48** (50mg, 0.045mmol, 1eq) and sodium methoxide (1.2mg, 22µmol, 0.5eq) were dissolved in methanol (5ml) and allowed to stir at room temperature. TLC (ethyl acetate) after 2 hours showed the complete consumption of starting material. The reaction mixture was opened to the atmosphere and quenched by the addition of dry ice until neutral (pH paper). The reaction mixture was concentrated *in vacuo* and the compound purified by silica gel chromatography (ethyl acetate) to give the desired compound as a colourless oil (35mg, 82%).

R_f 0.25 (ethyl acetate), $[\alpha]_D^{21}$ +32.3 (c = 0.84, $CHCl_3$); ^1H NMR (500MHz, $CDCl_3$) δ ppm 0.89 (3H, t, $J_{CH2\text{-}CH3}$ 6.9Hz, CH_3), 1.25 (32H, br s, $CH_2C_{16}\underline{H}_{32}CH_3$), 1.77 (1H, ddd, ^2J 14.0Hz, $J_{C\underline{H}H\text{-}C\underline{H}H}$ 13.6Hz, $J_{C\underline{H}H\text{-}CH\underline{H}}$ 4.1Hz, $C\underline{H}_2C_{17}H_{35}$), 1.82 (1H, ddd, ^2J 14.4Hz, $J_{CH\underline{H}\text{-}C\underline{H}H}$ 13.4Hz, $J_{CH\underline{H}\text{-}CH\underline{H}}$ 4.4Hz, $C\underline{H}_2C_{17}H_{35}$), 3.46 (1H, dd, $J_{H2\text{-}H3}$ 9.6Hz, $J_{H1\text{-}H2}$ 3.0Hz, H2), 3.49 (1H, t, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, H4), 3.56 (1H, d, $J_{H2'\text{-}H3'}$ 9.1Hz, H2'), 3.57 (1H, m, H6b), 3.60 (1H, dd, $J_{H6'a\text{-}H6'b}$ 11.2Hz, $J_{H5'\text{-}H6'b}$ 4.3Hz, H6'b), 3.62 (1H, t, $J_{H2'\text{-}H3'}$ 9.6Hz, $J_{H3'\text{-}H4'}$ 9.6Hz, H3'), 3.68 (1H, dd, $J_{H6a\text{-}H6b}$ 11.4Hz, $J_{H5\text{-}H6a}$ 3.2Hz, H6a), 3.69 (1H, dd, $J_{H6'a\text{-}H6'b}$ 11.4Hz, $J_{H5'\text{-}H6'a}$ 4.1Hz, H6'a), 3.90 (1H, dt, $J_{H4'\text{-}H5'}$ 9.8Hz, $J_{H5'\text{-}H6'a}$ 3.7Hz, $J_{H5'\text{-}H6'b}$ 3.7Hz, H5'), 3.94 (1H, t, $J_{H2\text{-}H3}$ 9.3Hz, $J_{H3\text{-}H4}$ 9.3Hz, H3), 4.10 (1H, t, $J_{H3'\text{-}H4'}$ 9.3Hz, $J_{H4'\text{-}H5'}$ 9.3Hz, H4'), 4.25 (1H, add, $J_{H4\text{-}H5}$ 9.5Hz, $J_{H5\text{-}H6b}$ 1.6Hz, H5), 4.45 (1H, d, ^2J 12.0Hz, $OC\underline{H}_2Ph^1$), 4.57 (1H, d, ^2J 12.3Hz, $OC\underline{H}_2Ph^1$), 4.60 (1H, d, ^2J 5.4Hz, $OC\underline{H}_2Ph^2$), 4.62 (1H, d, ^2J 5.7Hz, $OC\underline{H}_2Ph^2$), 4.80 (1H, d, ^2J 11.0Hz, $OC\underline{H}_2Ph^3$), 4.87 (2H, s, $OC\underline{H}_2Ph^4$), 4.90 (1H, d, ^2J 11.3Hz, $OC\underline{H}_2Ph^3$), 5.19 (1H, d, $J_{H1\text{-}H2}$ 3.5Hz, H1), 7.14-7.39 (20H, m, ArH); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 14.1 (1C, CH_3) 22.7, 24.5, 29.3, 296-29.7, 30.0, 31.9 (16C, m, $C_{16}H_{32}$), 34.6 (1C, $\underline{C}H_2C_{17}H_{35}$), 61.8 (1C, C6), 68.8 (1C, C6'), 70.4 (1C, C4), 71.2 (1C, C5), 71.2 (1C, C5'), 72.4 (1C, C2), 73.1 (1C, $O\underline{C}H_2Ph^1$), 73.8 (1C, C3), 75.1 (1C, $O\underline{C}H_2Ph^3$), 75.2 (1C, $O\underline{C}H_2Ph^2$), 75.1 (1C, $O\underline{C}H_2Ph^4$), 78.6 (1C, C3'), 80.2 (1C, C2'), 82.9 (1C, C4'), 91.3 (1C, C1), 102.9 (1C, C1'), 127.4, 127.6, 127.9 (8C, ArC_{ortho}), 127.4, 127.5, 127.7, 127.8 (4 x 1C, ArC_{para}), 128.3, 128.4 (2 x 4C, ArC_{meta}), 137.9,

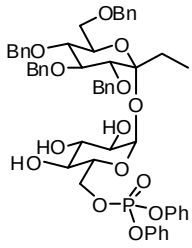
138.3, 138.5, 138.8 (4 x 1C, 4 x ArC_{ipso}); FT-IR (thin film) υ 2852, 2924 (C-H), 3385 br, 3583 (OH); HRMS (ES^+) m/z 977.5747 $[M + Na]^+$ (required 977.5749).



4.50. 1-octadecyl-2,3,4,6-O-benzyl-D-trehalose

A suspension of **4.49** (17mg, 0.018mmol, 1eq), 10% Pd/C (1mg, 0.009mmol, 0.1eq) and Brockman Grade I Alumina (5mg, as a buffer to ensure basic conditions) in methanol (1ml) was repeatedly degassed under high vacuum and the reaction vessel charged with hydrogen. The mixture was stirred under an atmospheric pressure of hydrogen for 6 days, after which TLC (1 water : 4 isopropanol : 4 ethyl acetate) showed complete consumption of starting material and the formation of a single product. The reaction mixture was filtered through Celite® and the solvent removed *in vacuo* to yield the desired compound as a white amorphous solid (10.1mg, 96%).

R_f 0.57 (1 water : 4 isopropanol : 4 ethyl acetate); $[\alpha]_D^{25}$ + 38.4 (c = 0.51, MeOH); ^1H NMR (500MHz, MeOD) δ ppm 0.92 (3H, t, $J_{CH2\text{-}CH3}$ 6.9Hz, CH_3), 1.31 (32H, as, $CH_2C_{16}\underline{H}_{32}CH_3$) 1.91-2.02 (2H, m, $C\underline{H}_2C_{17}H_{35}$), 3.23 (1H, t, $J_{H3\text{-}H4}$ 9.6Hz, $J_{H4\text{-}H5}$ 9.6Hz, H4), 3.38 (1H, t, $J_{H3'\text{-}H4'}$ 9.5Hz, $J_{H4'\text{-}H5'}$ 9.5Hz, H4'), 3.41 (1H, d, $J_{H2\text{-}H3}$ 9.8Hz, H2), 3.47 (1H, dd, $J_{H2'\text{-}H3'}$ 9.2Hz, $J_{H1'\text{-}H2'}$ 3.6Hz, H2'), 3.66 (1H, dd, $J_{H6a\text{-}H6b}$ 11.8Hz, $J_{H5\text{-}H6b}$ 5.8Hz, H6b), 3.74 (2H, m, H6'a, H6'b), 3.81 (1H, m, $J_{H5\text{-}H6a}$ 2.2Hz, H6a), 3.83 (1H, m, H5'), 3.83 (1H, $J_{H2'\text{-}H3'}$ 9.1Hz, $J_{H3'\text{-}H4'}$ 9.1Hz, H3'), 3.84 (1H, $J_{H2\text{-}H3}$ 9.8Hz, $J_{H3\text{-}H4}$ 9.8Hz, H3), 4.15 (1H, ddd, $J_{H4\text{-}H5}$ 9.9Hz, $J_{H5\text{-}H6b}$ 5.8Hz, $J_{H5\text{-}H6a}$ 2.2Hz, H5), 5.27 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1); ^{13}C NMR (126MHz, MeOD) δ ppm 14.5 (1C, CH_3), 23.8, 25.3, 30.5, 30.8, 31.3, 33.1 (16C, $CH_2\underline{C}_{16}H_{32}CH_3$), 36.3 (1C, $\underline{C}H_2C_{17}H_{35}$), 62.5 (1C, C6'), 63.1 (1C, C6), 72.1 (1C, C4'), 72.3 (1C, C4), 73.5 (1C, C5'), 73.8 (1C, C2'), 73.9 (1C, C2), 73.9 (1C, C5), 74.6 (1C, C3), 75.0 (1C, C3'), 92.8 (1C, C1); FT-IR (KBr disc) υ 2945 (C-H), 3398 br (OH); HRMS m/z (ES^-) 593.3913 $[M - H]^-$ required 593.3906.



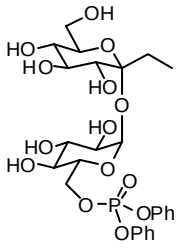
4.51. 1-octadecyl-2,3,4,6-O-benzyl-6'-diphenylphosphoryl-D-trehalose

The reaction was performed under Schlenk conditions using Argon gas. To a solution of **4.49** (50mg, 0.052mmol, 1eq) in anhydrous DCM (1ml) and anhydrous pyridine (1ml) was added diphenyl chlorophosphate (10.9µl, 0.052mmol, 1eq). The reaction was monitiored by TLC

(ethyl acetate) and the addition of excess diphenyl chlorophosphate (10eq in total) and DMAP (0.6mg, 0.005mmol, 0.1eq) was required to force the reaction to completion. TLC after stirring for 64 hours showed complete consumption of starting material and the reaction was quenched by addition of water (20ml). The reaction mixture was extracted with ethyl acetate (3 x 20ml) and the combined organics washed with 1M HCl (2 x 25ml), 0.25M $CuSO_4$ (25ml), water (25ml), brine (2 x 25ml), dried over $MgSO_4$ and the solvents removed *in vacuo*. The compound was purified by automated silica gel chromatography (1 petrol : 2 ethyl acetate) to give the desired compound as a colourless oil (32mg, 52%).

R_f 0.13 (1 petrol : 2 ethyl acetate), $[\alpha]_D^{18}$ +31.1 (c = 0.48, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 0.89 (3H, t, $J_{CH2\text{-}CH3}$ 6.9Hz, CH_3), 1.14 - 1.35 (32H, m, $C_{16}H_{32}$), 1.82 (1H, m, $J_{CH2\text{-}CH2}$ 4.4Hz, C$\underline{H}_2C_{17}H_{35}$), 1.87 (1H, m, $J_{CH2\text{-}CH2}$ 4.7Hz, C$\underline{H}_2C_{17}H_{35}$), 3.29 (1H, t, $J_{H3'\text{-}H4'}$ 9.5Hz, $J_{H4'\text{-}H5'}$ 9.5Hz, H4'), 3.39 (1H, dd, $J_{H2'\text{-}H3'}$ 9.6Hz, $J_{H1'\text{-}H2'}$ 3.6Hz, H2'), 3.57 (1H, t, $J_{H2\text{-}H3}$ 9.1Hz, $J_{H3\text{-}H4}$ 9.1Hz, H3), 3.62 (1H, t, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, H4), 3.64 (1H, m, H6b), 3.72 (1H, dd, $J_{H6a\text{-}H6b}$ 11.2Hz, $J_{H5\text{-}H6a}$ 4.1Hz, H6a), 3.89 (1H, t, $J_{H2'\text{-}H3'}$ 9.3Hz, $J_{H3'\text{-}H4'}$ 9.3Hz, H3'), 3.98 - 4.07 (2H, m, H5', H6'b), 4.03 (1H, d, $J_{H2\text{-}H3}$ 9.5Hz, H2), 4.28 (1H, ddd, $J_{H4\text{-}H5}$ 10.1Hz, $J_{H5\text{-}H6a}$ 4.3Hz, $J_{H5\text{-}H6b}$ 1.9Hz, H5), 4.40 (1H, ddd, $J_{H6'a\text{-}H6'b}$ 12.1Hz, $J_{H6'a\text{-}31P}$ 9.0Hz, $J_{H5'\text{-}H6'a}$ 2.8Hz, H6'a), 4.50 (1H, d, ^2J 12.3Hz, OC$\underline{H}_2$Ph^1), 4.57 (1H, d, ^2J 11.3Hz, OC$\underline{H}_2$Ph^2), 4.61 (1H, d, ^2J 12.3Hz, OC$\underline{H}_2$Ph^1), 4.62 (1H, d, ^2J 10.7Hz, OC$\underline{H}_2$Ph^3), 4.83 (1H, d, ^2J 11.0Hz, OC$\underline{H}_2$Ph^3), 4.86 (1H, d, ^2J 11.0Hz, OC$\underline{H}_2$Ph^4), 4.90 (1H, d, ^2J 11.0Hz, OC$\underline{H}_2$Ph^4), 4.92 (1H, d, ^2J 11.0Hz, OC$\underline{H}_2$Ph^2), 5.24 (1H, d, $J_{H1'\text{-}H2'}$ 3.8Hz, H1'), 7.19 - 7.40 (30H, m, ArH); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 14.1 (1C, $C_{17}H_{34}\underline{C}H_3$), 22.7, 29.3, 29.6, 30.0, 34.1 (16C, m, $CH_2\underline{C}_{16}H_{32}CH_3$), 31.9 (1C, $\underline{C}H_2C_{17}H_{34}$), 67.5 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, C6'), 68.9 (1C, C6), 69.2 (1C, C4'), 70.4 (1C, d, $J_{C\text{-}31P}$ 4.8Hz, C5'), 71.4 (1C, C5), 72.4 (1C, C2'), 73.0 (1C, O$\underline{C}H_2$Ph^1), 73.2 (1C, C3'), 75.0 (1C, O$\underline{C}H_2$Ph^3), 75.0 (1C, O$\underline{C}H_2$Ph^2), 75.4 (1C, O$\underline{C}H_2$Ph^4), 78.5 (1C, C4), 80.1 (1C, C3), 83.1 (1C, C2), 91.0 (1C, C1'), 103.0 (1C, C1), 120.0 (1C, d, $J_{C\text{-}31P}$ 4.8Hz, P-OPh ArC_{ortho}), 120.3 (1C, d, $J_{C\text{-}31P}$ 3.8Hz, P-OPh ArC_{ortho}), 125.6 (1C, P-OPh ArC_{para}), 125.8 (1C, P-OPh ArC_{para}), 127.4, 127.6, 127.8, 128.0, (4 x 2C, OBn ArC_{ortho}), 127.5, 127.7 (4C, OBn ArC_{para}), 128.3, 128.3, 128.4, 128.4 (4 x 2C, OBn ArC_{meta}), 129.8, 129.9 (2 x 2C, P-OPh ArC_{meta}), 138.1, 138.5, 138.6, 138.6 (4 x 1C, OBn ArC_{ipso}), 150.2 (1C, d, $J_{C\text{-}31P}$ 6.7Hz, P-OPh ArC_{ipso}), 150.4 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, P-OPh ArC_{ipso}); ^{31}P{^{1}H} NMR (162 MHz,

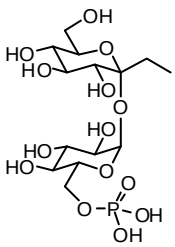
$CDCl_3$) δ ppm -9.53 (1P, $P(O)(OPh)_2$); FT-IR (thin film) υ 1190 (P=O), 2922 (C-H), 3297 br (OH); HRMS m/z (ES^+) $[M + Na]^+$ peaks 1209.60 (100%), 1210.61 (75%), 1211.61 (30%), 1212.61 (8%), required 1209.60 (100%), 1210.61 (77%), 1211.61 (32%), 1212.61 (9%).



4.52. 1-octadecyl-6'-diphenylphosphoryl-D-trehalose

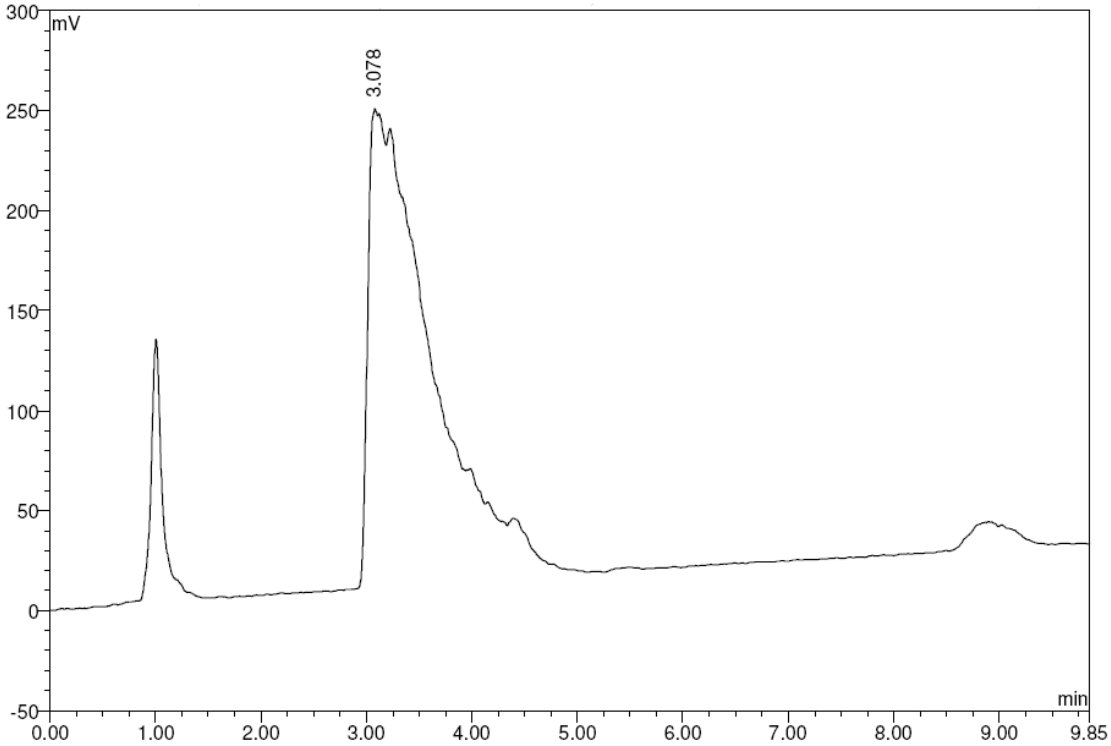
Compound **4.51** was reduced using flow hydrogenation (Thales Nano H-Cube®) with 70 Bar hydrogen pressure and methanol as the solvent at a flow rate of 1.0ml/min over Pd/C (30mm CatCart®) at 25°C. The reagent was continually cycled over the catalyst for 60 minutes after which TLC (1 water : 3 isopropanol : 14 ethyl acetate) showed complete consumption of starting materials and the formation of a single product. The solvent was removed *in vacuo* give the title compound as a colourless oil (19mg, 90%).

R_f 0.17 (1 water : 3 isopropanol : 14 ethyl acetate), $[\alpha]_D^{21}$ +51.0 (c = 0.2, MeOH); ^1H NMR (500MHz, MeOD) δ ppm 0.92 (3H, t, $J_{CH3\text{-}CH2}$ 6.9Hz, $C_{17}H_{34}C\underline{H}_3$), 1.31 (32H, br. s, $CH_2C_{16}\underline{H}_{32}CH_3$), 1.92 (1H, td, ^2J 12.4Hz, $J_{C\underline{H}H\text{-}C\underline{H}H}$ 12.4Hz, $J_{C\underline{H}H\text{-}CH\underline{H}}$ 3.8Hz, $C\underline{H}_2C_{16}H_{33}$), 2.04 (1H, td, ^2J 13.0Hz, $J_{CH\underline{H}\text{-}CH\underline{H}}$ 13.0Hz, $J_{CH\underline{H}\text{-}C\underline{H}H}$ 4.9Hz, $C\underline{H}_2C_{16}H_{33}$), 3.23 (1H, t, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, H4), 3.33 (1H, t, $J_{H3'\text{-}H4'}$ 9.5Hz, $J_{H4'\text{-}H5'}$ 9.5Hz, H4'), 3.41 (1H, d, $J_{H2\text{-}H3}$ 9.5Hz, H2), 3.45 (1H, dd, $J_{H2'\text{-}H3'}$ 9.8Hz, $J_{H1'\text{-}H2'}$ 3.5Hz, H2'), 3.66 (1H, dd, $J_{H6a\text{-}H6b}$ 11.8Hz, $J_{H5\text{-}H6b}$ 5.8Hz, H6b), 3.82 (1H, dd, $J_{H6a\text{-}H6b}$ 11.4Hz, $J_{H5\text{-}H6a}$ 2.2Hz, H6a), 3.85 (2H, t, $J_{H2\text{-}H3}$ 9.1Hz, $J_{H3\text{-}H4}$ 9.1Hz, H3, H3'), 4.13 (1H, ddd, $J_{H4'\text{-}H5'}$ 9.8Hz, $J_{H5'\text{-}H6'b}$ 5.4Hz, $J_{H5'\text{-}H6'a}$ 1.6Hz, H5'), 4.16 (1H, ddd, $J_{H4\text{-}H5}$ 9.8Hz, $J_{H5\text{-}H6b}$ 5.8Hz, $J_{H5\text{-}H6a}$ 2.0Hz, H5), 4.42 (1H, ddd, $J_{H6'a\text{-}H6'b}$ 10.9Hz, $J_{H\text{-}31P}$ 6.1Hz, $J_{H5'\text{-}H6'b}$ 6.0Hz, H6'b), 4.54 (1H, ddd, $J_{H6'a\text{-}H6'b}$ 10.9Hz, $J_{H\text{-}31P}$ 5.8Hz, $J_{H5'\text{-}H6'a}$ 1.7Hz, H6'a), 5.27 (1H, d, $J_{H1\text{-}H2}$ 3.8Hz, H1'), 7.26 (6H, m, ArH_{ortho}, ArH_{para}), 7.42 (4H, t, $J_{ortho\text{-}meta}$ 7.9Hz, $J_{meta\text{-}para}$ 7.9Hz, ArH_{meta}); ^{13}C NMR (126MHz, MeOD) δ ppm 14.5 (1C, CH_3), 23.8, 25.4, 30.5, 30.7, 30.8, 30.9, 31.3, 33.1 (16C, m, $CH_2\underline{C}_{16}H_{32}CH_3$), 36.4 (1C, $\underline{C}H_2C_{17}H_{35}$), 63.1 (1C, C6), 70.1 (1C, d, $J_{C\text{-}31P}$ 5.7Hz, C6'), 71.8 (1C, C4'), 72.0 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, C5'), 72.3 (1C, C4), 73.6 (1C, C2'), 73.8 (1C, C2), 74.0 (1C, C5), 74.4 (1C, C3'), 75.0 (1C, C3), 92.8 (1C, C1'), 104.2 (1C, C1), 121.3 (2C, d, $J_{C\text{-}31P}$ 4.8Hz, ArC_{ortho}), 126.8 (1C, ArC_{para}), 131.1 (1C, ArC_{meta}), 151.8 (1C, d, $J_{C\text{-}31P}$ 7.6Hz, ArC_{ipso}); FT-IR (KBr disc) υ 1201 (P=O), 2900 (C-H), 3337 br (OH); HRMS (ES^+) m/z 849.4140 $[M + Na]^+$ (required 849.4161).



4.53. 1-octadecyl-6'-diphenylphosphoryl-D-trehalose

A suspension of **4.52** (12mg, 15µmol, 1eq) and PtO_2 (0.4mg, 1.5µmol, 0.1eq) in methanol (5ml) with glacial acetic acid (25µl) was repeatedly deassed under vacuum and charged with hydrogen. The reaction mixture was stirred vigorously at room temperature. After 5 hours, TLC (1 water : 3 isopropanol : 14 ethyl acetate) showed complete consumption of the starting material. The reaction mixture was filtered through Celite®, which was subsequently washed with water. The solvents were removed *in vacuo*. The compound was purified by HPLC using a Pehnomenex Synergi Hydro C18 column (21.2 x 100mm, 4µm) and 3 water : 7 acetonitrile as the mobile phase at flow rate of 23ml/min , with ELS detection of eluants. Lyophilization yielded the desired compound as an amorphous white solid (9.5mg, 98%).



R_f 0.5 (5 ethanol : 3 NH_4OH : 1 water), $[\alpha]_D^{20}$ +35.7 (c = 0.23, H_2O); ^1H NMR (500MHz, D_2O + 2% MeOD) δ ppm 0.74 (3H, br. s, CH_3), 1.15 (34H, br. s, $C_{17}\underline{H}_{34}$), 3.12 (1H, t, $J_{H3\text{-}H4}$ 8.5Hz, $J_{H4\text{-}H5}$ 8.5Hz, H4), 3.36 (1H, t, $J_{H2\text{-}H3}$ 10.1Hz, $J_{H3\text{-}H4}$ 10.1Hz, H3), 3.36 (1H, m, H2), 3.41 (1H, dd, $J_{H2'\text{-}H3'}$ 9.7Hz, $J_{H1'\text{-}H2'}$ 3.8Hz, H2'), 3.43 (1H, m, H4'), 3.44 (1H, add, $J_{H5\text{-}H6b}$ 7.2Hz, $J_{H5\text{-}H6a}$ 2.8Hz, H5), 3.58 (1H, t, $J_{H2'\text{-}H3'}$ 9.6Hz, $J_{H3'\text{-}H4'}$ 9.6Hz, H3'), 3.80 (1H, ad, $J_{H4'\text{-}H5'}$ 10.3Hz, H5'), 3.90 (1H, dd, $J_{H6a\text{-}H6b}$ 11.9Hz, $J_{H5\text{-}H6b}$ 6.0Hz, H6b), 3.92-3.97 (2H, m, H6'b, H6a), 4.00 (1H, ddd, $J_{H6'a\text{-}H6'b}$ 11.7Hz, $J_{H\text{-}}$

$_{31P}$ 5.4Hz, $J_{H5'-H6'a}$ 1.6Hz, H6'a), 5.08 (1H, d, $J_{H1'-H2'}$ 3.8Hz, H1'); ^{13}C NMR (126MHz, D_2O + 2% MeOD) δ ppm 13.9 (1C, CH_3), 22.6, 29.4-30.8, 32.0 (17C, $\underline{C}_{17}H_{34}CH_3$), 64.0 (1C, d, J_{C-31P} 9.5Hz, C6'), 64.0 (1C, C6), 69.0 (1C, C5), 69.1 (1C, C3), 70.4 (1C, d, J_{C-31P} 7.6Hz, C5'), 71.4 (1C, C2'), 72.5 (1C, C3'), 74.0 (1C, C4), 74.7 (1C, d, J_{C-31P} 8.6Hz, C4'), 75.4 (1C, C2), 92.1 (1C, C1'), 95.9 (1C, C1); FT-IR (KBr disc) 1196 (P=O), 3298 br (OH); HRMS (ES^-) m/z 673.3569 $[M - H]^-$ (required 673.3570).

4.12 References

Figures in this chapter were reproduced by permission from Elsevier from: Cátia Nunes, Lucia F. Primavesi, Mitul K. Patel, Eleazar Martinez-Barajas, Stephen J. Powers, Ram Sagar, Pedro S. Fevereiro, Benjamin G. Davis, Matthew J. Paul, *Plant Physiol. Biochem.*, **2013**, 63, 89.

1. (a) R. Aebersold, M. Mann, *Nature*, **2003**, 422, 198. (b) P. R. Graves, T. A. J. Haystead, *Microbiol. Mol. Biol. Rev.*, **2002**, 66, 39. (c) A. Lovegrove, L. Salt, P. R. Shewry, *Methods Mol. Biol.*, **2009**, 478, 273.
2. E. M. Marcotte, *Nature Biotechnol.*, **2007**, 25, 755.
3. A. Bauer, B. Kuster, *Eur. J. Biochem.*, **2003**, 270, 570.
4. (a) M. O. Collins, J. S. Choudhary, *Curr. Opin. Biotechnol.*, **2008**, 19, 324. (b) J. V. Leene, E. Witters, D. Inze, G. D. Jaeger, *Trends Plant Sci.*, **2008**, 13, 517.
5. H. Zou, Q. Luo, D. Zhou, *Biochem. Biophys. Methods*, **2001**, 49, 199.
6. (a) S. Dale, W. A. Wilson, A. M. Edelman, D. G. Hardie, *FEBS Lett.*, **1995**, 361, 191. (b) J. Weekes, K. L. Ball, F. B. Caudwell, D. G. Hardie, *FEBS Lett.*, **1993**, 334, 335.
7. S. P. Davies, D. Carling, D. G. Hardie, *Eur. J. Biochem.*, **1989**, 186, 123.
8. (a) J. Lengeler, G. Drews, H. Schlegel, *The Biology of the Prokaryotes*, **1998**, Wiley-Blackwell. (b) B. Axelrod, *Metabolic Pathways*, **1967**, Academic.
9. Y. Zhang, L. F. Primavesi, D. Jhurreea, P. J. Andralojc, R. A. C. Mitchell, S. J. Powers, H. Schluepmann, T. Delatte, A. Wingler, M. J. Paul, *Plant Physiol.* **2009**, 149, 1860.
10. D. Toroser, Z. Plaut, S. C. Huber, *Plant Physiol.*, **2000**, 123, 403.

11. M. J. Paul, L. F. Primavesi, *Personal Communication*.
12. N. G. Halford, S. J. Hey, *Biochem. J.*, **2009**, 419, 247.
13. C. V. Piattoni, D. M. Bustos, S. A. Guerrero, A. Á. Iglesias, *Plant Physiol.*, **2011**, 156, 1337.
14. C. Nunes, L. F. Primavesi, A. B. Silva, P. S. Fevereiro, M. J. Paul, *Unpublished Results*.
15. C. A. Raines, *Photosynth. Res.*, **2003**, 75, 1.
16. N. J. Kruger, A. V. Schaewen, *Curr. Opin. Plant Biol.*, **2003**, 6, 236.
17. M. J. Paul, S. P. Driscoll, P. J. Andralojc, J. S. Knight, J. C. Gray, D. W. Lawlor, *Planta*, **2000**, 211, 122.
18. (a) P. J. Kocienski, *Protecting Groups*, **2005**, Thieme. (b) J. Oudar, *Catal. Rev. Sci. Eng.*, **1980**, 22, 171. (c) L. L. Hegedus, R. W. McCabe, *Catal. Rev. Sci. Eng.*, **1981**, 23, 377.
19. (a) M. W. Bowler, M. J. Cliff, J. P. Waltho, G. M. Blackburn, *New J. Chem.*, **2010**, 34, 784. (b) F. H. Westheimer, *Science*, **1987**, 235, 1173. (c) G. K. Schroeder, C. Lad, P. Wyman, N. H. Williams, R. Wolfenden, *Proc. Natl. Acad. Sci. U. S. A.*, **2006**, 103, 4052.
20. T. Enami, N. Nagae, S. Doshi, *LC-GC Europe*, **2003**, July, 2.
21. D. N. Perkins, D. J. C. Pappins, D. M. Creasy, J. S. Cottrell, *Electrophoresis*, **1999**, 20, 3551.
22. D. J. C. Pappin, P. Hojrup, A. J. Bleasby, *Curr. Biol.*, **1993**, 3, 327.
23. (a) R. J. Ferl, *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **1996**, 47, 49. (b) A. Aitken, *Semin. Cancer Biol.*, **2006**, 16, 162. (c) M. B. Yaffe, *FEBS Lett.*, **2002**, 513, 53. (d) V. Obsilova, J. Silhan, E. Boura, J. Teisinger, T. Obsil, *Physiol. Res.*, **2008**, 57, S11.
24. (a) M. R. Roberts, *Curr. Opin. Plant Biol.*, **2000**, 3, 400. (b) C. Oecking, N. Jaspert, *Curr. Opin. Plant Biol.*, **2009**, 12, 760. (c) M. R. Roberts, *Trends Plant Sci.*, **2003**, 8, 218. (d) H-J. Chung, P. C. Sehnke, R. J. Ferl, *Trends Plant Sci.*, **1999**, 4, 367. (e) P. Aducci, L. Camoni, M. Marra, S.Visconti, *IUBMB Life*, **2002**, 53, 49. (f) C-S. Oh, *Plant Pathol. J.*, **2010**, 26, 1.
25. J. E. Harthill, S. E. Meek, N. Morrice, M. W. Peggie, J. Borch, B. H. Wong, C. Mackintosh, *Plant J.*, **2006**, 47, 211.

26. (a) J. M. DeLille, P. C. Sehnke, R. J.Ferl, *Plant Physiol.*, **2001**, 126, 35. (b) K. Wu, M.F. Rooney, R. J. Ferl, *Plant Physiol.*, **1997**, 114, 1421.
27. M. Rosenquist, P. Sehnke, R. J. Ferl, M. Sommarin, C. Larsson, *J. Mol. Evol.*, **2000**, 51, 446.
28. (a) M. B. Yaffe, K. Rittinger, S. Volinia, P. R. Caron, A. Aitken, H. Leffers, S. J. Gamblin, S. J. Smerdon, L. C. Cantley, *Cell*, **1997**, 91, 961. (b) K. Rittinger, J. Budman, J. Xu, S. Volinia, L. C. Cantley, S. J. Smerdon, S. J. Gamblin, M. B. Yaffe, *Mol. Cell*, **1999**, 4, 153.
29. The Arabidopsis Information Resource (TAIR), www.arabidopsis.org.
30. G. S. Athwal, J. L. Huber, S. C. Huber, *Plant Physiol.*, **1998**, 118, 1041.
31. M. Wurtele, C. Jelich-Ottmann, A. Wittinghofer, C. Oecking, *EMBO J.*, **2003**, 22, 987.
32. J. M. Woodcock, Y. Ma, C. Coolen, D. Pham, C. Jones, A. F. Lopez, S. M. Pitson, *Cell Signal.*, **2010**, 22, 1291.
33. N. Jaspert, C. Oecking, *Planta*, **2002**, 216, 136.
34. (a) M. C. Jecklin, S. Schauer, C. E. Dumelin, R. Zenobi, *J. Mol. Recognit.*, **2009**, 22, 319. (b) L. Indyk, H. F. Fisher, *Methods Enzymol.*, **1998**, 295, 350.
35. G. A. Holdgate, W. H. J. Ward, *Drug Discov. Today*, **2005**, 10, 1543.
36. (a) T. T. Waldron, K. P. Murphy, *Biochemistry*, **2003**, 42, 5058. (b) P. Cimmperman, L. Baranauskiene, S. Jachimoviciute, J. Jachno, J. Torresan, V. Michailoviene, J. Matuliene, J. Sereikaite, V. Bumelis, D. Matulis, *Biophys. J.*, **2008**, 95, 3222. (c) J. M. Sanchez-Ruiz, *Biophys. Chem.*, **2007**, 126, 43. (d) T. O Street, D. W. Bolen, J. D. Rose, *Proc. Natl. Acad. Sci.*, **2006**, 103, 13997. (e) L. R. Singh, N. K. Poddar, T. A. Dar, R. Kumar, F. Ahmad, *Life Sci.*, **2011**, 88, 117.
37. N. G. Greenfield, *Methods Enzymol.*, **2004**, 383, 282.
38. Y. Griko, N. Sreerama, P. Osumi-Davis, R. W. Woody, A. M. Woody, *Protein Sci.*, **2001**, 10, 845.
39. (a) B. Meyer, T. Peters, *Angew. Chem. Int. Ed.*, **2003**, 42, 864. (b) J. W. Peng, J. Moore, N. Abdul-Manan, *Prog. Nucl. Mag. Res. Sp.*, **2004**, 44, 225. (c) T. Carlomagno, *Annu. Rev. Biophys. Biomol. Struct.*, **2005**, 34, 245.
40. (a) M. Mayer, B. Meyer, *J. Am. Chem. Soc.*, **2001**, 123, 6108. (b) N. R. Krishna, V. Jayalakshmi, *Prog. Nucl. Mag. Res. Sp.*, **2006**, 49, 1.

41. T. Feizi, F. Fazio, W. Chai, C-H Wong, *Curr. Opin. Struct. Biol.*, **2003**, 13, 637.
42. R. Namme, T. Mitsugi, H. Takahashi, S. Ikegami, *Eur. J. Org. Chem* **2007**, *2007,* 3758.
43. A. K. Gardino, S. J. Smerdon, M. B. Yaffe, *Semin. Cancer Biol.*, **2006**, 16, 173.
44. M. S. Manak, R. J. Ferl, *Biochemistry*, **2007**, 46, 1055.
45. E. Tarelli, S. F. Wheeler, *Carbohydr. Res.*, **1994**, 261, 25.
46. A. Natarajan, W. Du, C. Xiong, *et al.*, *Chem. Commun.*, **2007**, 7, 695.

Chapter V - Sulfate binding to the mannose receptor

5.1 Introduction

Multivalency is a common strategy that nature employs for the binding of carbohydrates to lectins.[1] The mannose receptor (MR) is a particularly interesting example of this, where polyvalent interactions are used not only to improve binding affinities, but to also control the selectivity between the many possible ligands for the receptor (section 1.3, page 11).[2-5]

The binding of oligosaccharides terminating with mannose, fucose, *N*-acetyl glucosamine and glucose to the carbohydrate recognition domains (CRDs) of the MR is fairly well understood.[3,6,7] The co-operative action of CRDs 4 and 5, which are almost exclusively responsible for the binding of these carbohydrate ligands,[8,9] permits divalent interactions resulting in strong binding. Sulfated carbohydrates can interact with a binding site at the terminal cysteine-rich region (CRR) of the MR, but such singular interactions are too weak to be biologically significant.[2,4] However, receptor dimerization brings together two CRRs, allowing for divalent interactions that enhance binding affinity and selectivity for proteins bearing terminal sulfated carbohydrates.[2,4] Several sulfated carbohydrate binding partners for the MR have been documented, but the biological implications of these interactions are poorly understood.[10-12] However, there is a growing body of evidence to support the involvement of the CRR in the adaptive immune response through sulfate mediated trafficking of various components of the lymphatic system.[4,10]

To better understand these biological roles, it is vital to have a clear awareness of the expression and post-translational modification patterns of the MR which are known to vary hugely between different cell types. Specifically, the factors responsible for receptor dimerization, upon which sulfate binding is critically dependent, need to be fully understood. Current methods to probe the extent of dimerization are rather archaic relying on covalent cross-linking,[13] receptor isolation and purification,[2] or density gradient centrifugation.[9] There is evidence to suggest that MR dimers are held together with weak, non-covalent interactions and that dimerization is likely to be reversible.[13] Hence, it is not surprising that these harsh methods have produced widely

differing results for the extent of dimerization. A method with minimal perturbation of the native cellular environment is needed for more accurate and reliable studies.

Fluorescence microscopy is an excellent method for studying live cells.[14] The high sensitivity and spatial and temporal resolution allows accurate monitoring of sub-cellular events in real time,[15] and with modern equipment, single molecule sensitivity can be attained giving unprecedented levels of insight into biological systems.[16] Importantly, cells extracted from live sources can be probed *ex vivo* with minimal disruption of the natural environment.

Förster resonance energy transfer (FRET) is a specialized application of fluorescence microscopy, where energy is transferred between appropriately chosen fluorophores resulting in an increase in fluorescence intensity for the acceptor and a concomitant decrease for the donor.[17] As a non-radiative energy transfer mechanism that results from dipolar coupling between the two fluorophores, the phenomenon is intrinsically short-ranged and rapidly falls off as the sixth power of inter-fluorophore distance. As such, the detection of short-ranged interactions that would normally be below the resolution of traditional fluorescence microscopy is possible. The method is now commonly used in biochemistry to probe the interactions of proteins on a sub 10nm scale.[17-19] FRET microscopy could be a useful technique for the study of MR dimerization where two monomers are brought within close proximity of each other.

The labelling of membrane proteins is often accomplished with immunofluorescence techniques, where the high antigen specificity of fluorescently labelled antibodies is exploited to strongly tether to proteins of interest.[19] However, such methods also have drawbacks such as non-stoichiometric concentrations of fluorophores making quantification difficult, lower FRET resolution due to the large size of the antibodies, and high costs.[20] For membrane *receptors*, an alternative is to use the binding ligand itself as a scaffold for fluorophore attachment.[18,21,22] In this way, a much smaller probe can be created that has a very specific binding site on the receptor and ensures a 1 : 1 stoichiometry between the fluorophore and the receptor. For the MR, different FRET systems can be envisaged based on the divalent binding of ligands at the CRR and the CRDs (Figure 5.1). Simultaneous use of both acceptor-labelled and donor-labelled ligands results in FRET pair creation through 50% statistically coincidental binding of acceptor and donor dyes at adjacent sites.

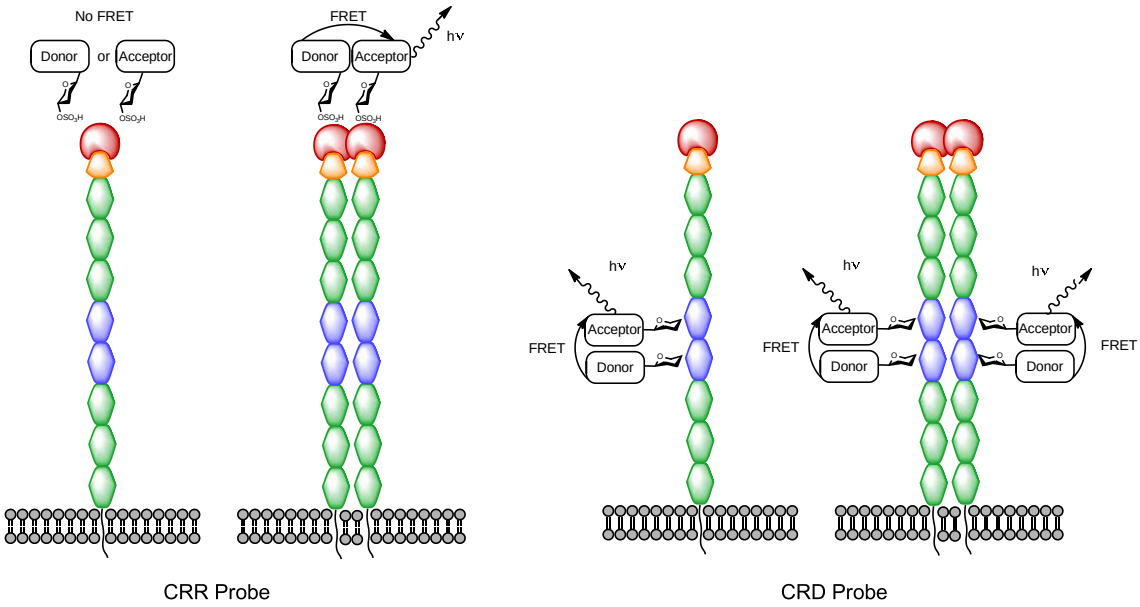


Figure 5.1. *Key: CRR (red), fibronectin domain (orange), core CRDs (blue), inactive or weakly active CRDs (green). Left: A system to probe dimerization based on the binding of sulfated carbohydrates at the CRR. A FRET signal is only possible for the dimer. Right: A system to probe the total amount of MR present based on carbohydrate binding to the CRDs. A FRET signal is produced for both monomer and dimer. The combined results would reveal the amount of monomer and dimer present at the cell surface.*

In such a scheme, dimerization could be surveyed with CRR specific, sulfated FRET probes. Since the MR contains only one CRR binding domain,[2-5] sulfated ligands would only be able to form a FRET pair on the dimer. In this way, the sulfate FRET signal would provide a direct measurement for MR dimerization at the cell surface.

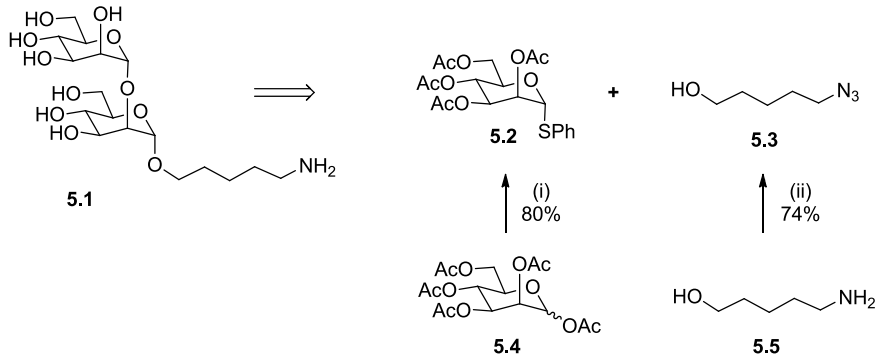
In contrast, each MR contains two active CRDs that allow divalent binding even for the monomer. Recent evidence suggests that carbohydrate binding to the CRDs remains possible even after dimerization,[13,23] and so a FRET signal would be expected for each individual CRD pair. Thus, the carbohydrate FRET signal would provide data on the total levels of MR presentation. Experimentally measuring these FRET signals would be very difficult, but if conditions could be found, ,the amounts of MR monomer and dimer could potentially be evaluated from the CRR and CRD specific probes.

In this chapter, we present some preliminary investigations on probing the MR with fluorescent microscopy. The synthesis of a mannose disaccharide is presented as a probe specific for the CRDs and our investigations into the synthesis of a GalNAc-4-SO_4 probe for the CRR are discussed. A cultured rat macrophage cell line is used to determine whether FRET experiments on live cells would be a feasible method for

determining MR dimerization. Finally, we consider further experiments that would be required to draw firmer conclusions from these early inverstigations.

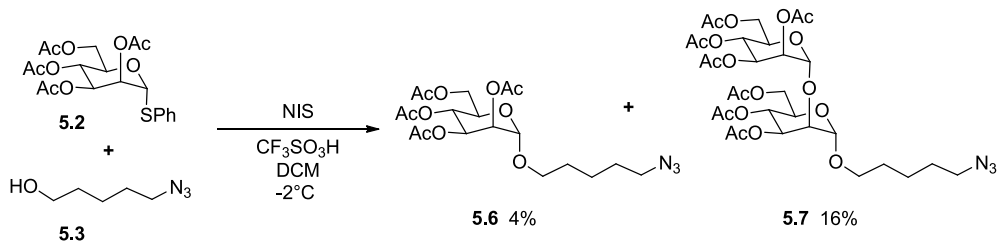
5.2 Synthesis of the mannose probe

Mannose-based oligosaccharides are recognised by the CRDs of the MR. Hence, the 1→2 α mannose disaccharide motif that is widespread in nature was chosen as the ligand for the CRD probes. A C_5 spacer with a terminal amine group was chosen as a versatile chemical handle for attachment to fluorophores and fluorescent proteins. We envisaged that this CRD probe **5.1** could be efficiently synthesized from the precursors **5.2** and **5.3** (Scheme 5.1). The monosaccharide **5.2** was synthesised by glycosylation of **5.4** with thiophenol in an excellent 80% yield, while azido transfer from triflic azide to **5.5** smoothly furnished the protected linker **5.3** in 74% yield.



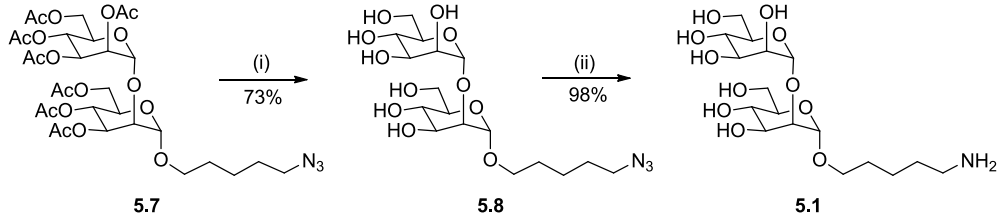
Scheme 5.1. *Retrosynthetic analysis for the synthesis of the CRD probe. Synthesis of the precursors, conditions: (i) PhSH, DCM, $BF_3{\cdot}OEt_2$, RT; (ii) NaN_3, Tf_2O, $CuSO_4$, NEt_3, H_2O, RT.*

Within our research group, a polyglycosylation has been developed that permits the synthesis of 1→2 α polymannosides directly from **5.2** and alcohols such as **5.3**.[24] It is hypothesized that this reaction, which uses *N*-iodosuccinimide and triflic acid as the activator, proceeds via the orthoester between the 1 and 2 positions and the alcohol. Further attack at the anomeric centre can result in deprotection of the 2 position, from which further glycosylation can occur. These optimized conditions were applied to **5.2** and **5.3**, which gave the monosaccharide **5.6** in 4% yield and the desired compound **5.7** in a pleasing 16% (Scheme 5.2). The method also produced the tri- and tetrasaccharides as minor side products (<5%).



Scheme 5.2. *Polyglycosylation conditions to produce the desired disaccharide.*

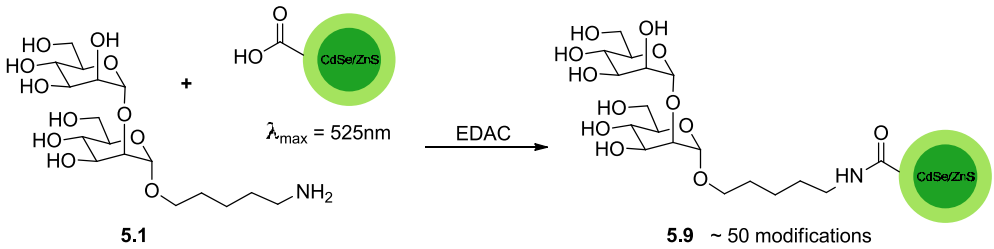
The disaccharide **5.7** was subsequently deacetylated with sodium methoxide in methanol to produce **5.8** in 73% yield. The product was then subjected to hydrogenolysis over palladium on charcoal to furnish the target compound **5.1** in near quantitative yield (Scheme 5.3).



Scheme 5.3. *Conditions: (i) NaOMe, MeOH; (ii) H_2, Pd/C, MeOH, H_2O.*

5.3 Mannosylated quantum dots

The amine functionality in **5.1** allows attachment onto a variety of fluorophores through the formation of a strong and stable amide bond. Quantum dots offer several advantages that favour their use in microscopy, including very high fluorescence quantum yields, low vulnerability to photobleaching, narrow emission bandwidths, high biostability and the multivalent display of ligands.[25] For these reasons, quantum dots were chosen for preliminary internalization studies with cultured cells. To ensure normal functioning of endocytotic pathways, small nanoparticles (~5nm) were used. CsSe/ZnS core/shell quantum dots with terminal carboxyl groups were modified with **5.1** using an EDAC mediated peptide coupling reaction (Scheme 5.4). This gave quantum dots decorated with ~50 carbohydrates (**5.9**), as determined using the phenol sulphuric acid method (section 5.11.4, page 271).[26]



Scheme 5.4. *Conditions: EDAC, 20mM sodium phosphate buffer pH 7.5, 1h, RT. Number of modifications assayed with the phenol-sulfuric acid method.*

Alveolar macrophages from rat lungs are known to have a high expression of the MR in a form that interacts strongly with mannose.[2,3,5] To emulate this environment, we used the immortal cell line NR8383, which is derived from alveolar macrophages of a Sprague-Dawley rat, as a more convenient platform for studying the MR. A 40nM solution of **5.9** was incubated with lives cells derived from the rat macrophage cell line, which were labelled within minutes after addition of the quantum dots. After 1 hour, significant internalization had occurred (Figure 5.2 and Figure 5.3).

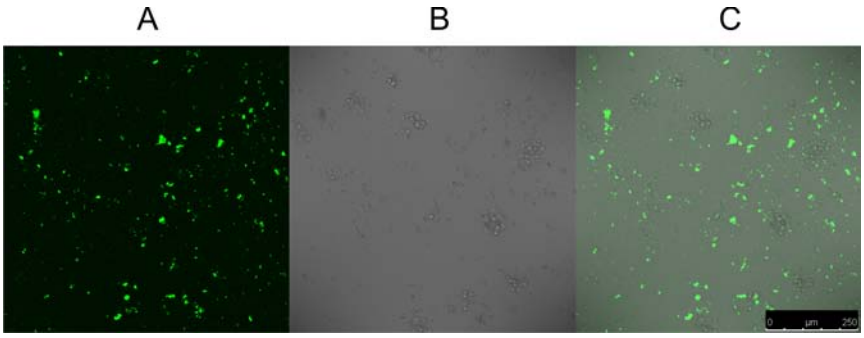


Figure 5.2. *NR8383 cells incubated with* ***5.9*** *for 1 hour. (A) Quantum dot fluorescence (B) brightfield image (C) overlay of A and B. Scale bar = 250µm. Data show uptake of quantum dots into cells.*

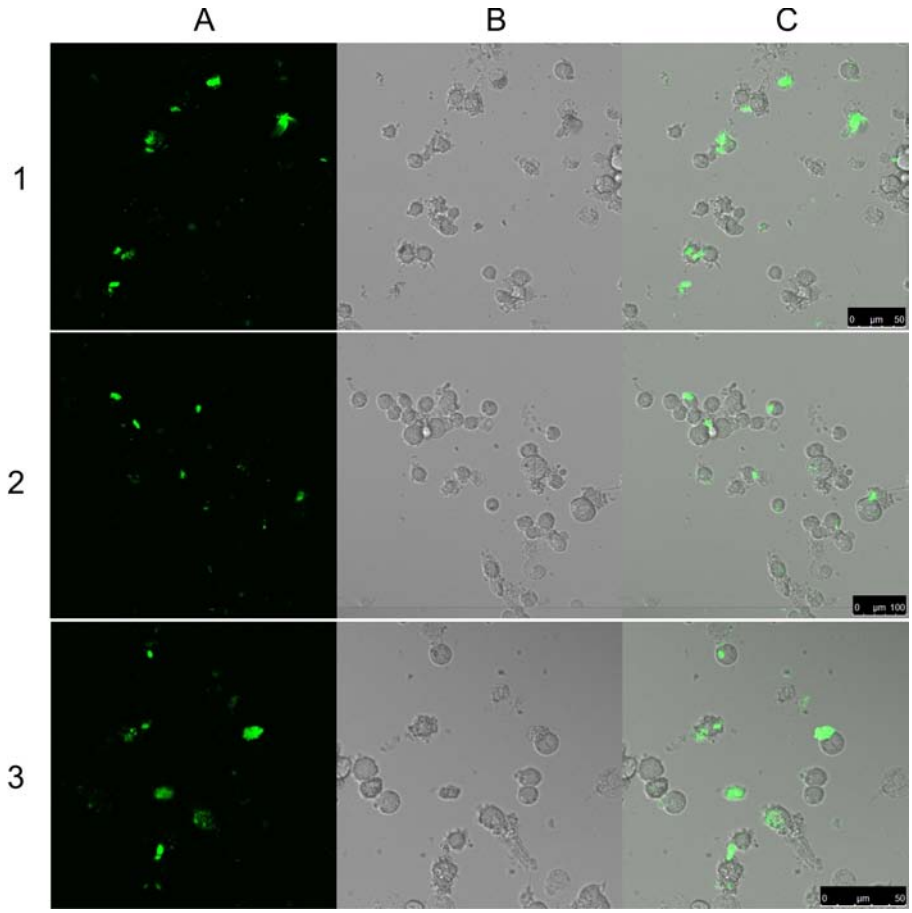


Figure 5.3. *Magnified images of NR8383 cells with* ***5.9*** *after 1 hour. (A) Quantum dot fluorescence (B) brightfield image (C) overlay of A and B. Scale bar = 50µm. Data shows labelling of cells with quantum dots.*

To demonstrate that the uptake of quantum dots was through active transportation rather than diffusion into dead cells, the trypan blue exclusion method was used to selectively label non-viable cells, whose cell membranes are no longer intact and thus allow unrestricted entry of the dye into the cell (Figure 5.4).[27] The lack of co-localization between trypan blue and **5.9** indicates that quantum dot uptake does not generally occur with non-viable cells. We note that there are more modern stains for determining cell viability such as ethidium homodimer III, the eFluor® dyes, or the Live/Dead labelling kits from molecular probes® which have brighter and more stable fluorescence compared to trypan blue. Use of such dyes would be preferable in future experiments.

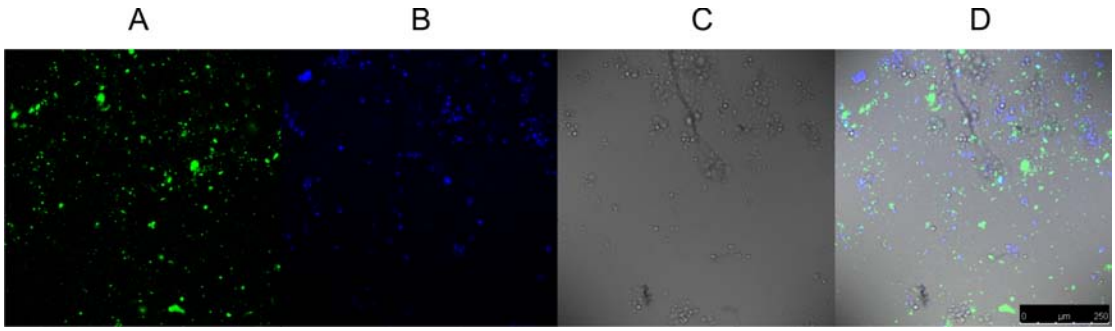
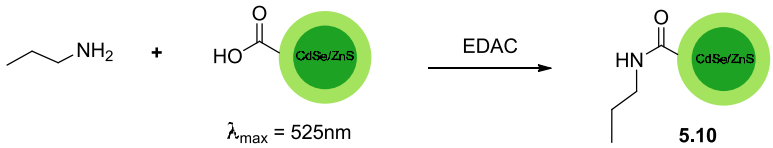


Figure 5.4. *NR8383 cell line with* ***5.9*** *after 2 hours co-stained with trypan blue. (A) Quantum dot fluorescence (B) trypan blue fluorescence (C) brightfield image (D) overlay of A, B and C. Scale bar = 250µm. Data shows quantum dot uptake with only viable cells.*

However, there are several possible mechanisms for the cellular uptake of quantum dots that allow even unmodified nanoparticles to be internalized (such as lipid raft transport).[28] This is especially true for negatively charged nanoparticles such as the carboxylate capped dots used in this work.[28] To demonstrate that cellular uptake was due to the mannose moiety rather than non-specific uptake, a preparation of quantum dots capped with propylamine was synthesized as a control (Scheme 5.5).



Scheme 5.5. *Conditions: EDAC, 20mM sodium phosphate buffer pH 7.5, 1h, RT.*

NR8383 cultured cells did not show any binding or uptake of **5.10**, even after 1 hour incubation (Figure 5.5). Hence, we believe uptake of **5.9** requires mannose for endocytosis rather that through non-specific mechanisms. However, alternative control experiments can also be envisaged and should be performed to fully validate this result. For example, the experiment could be repeated with different concentrations of **5.1** incubated with the cells and **5.9**. If endocytosis occurs due to interaction with the sugar, increased concentrations of free **5.1** should compete with binding to the receptor thus slowing the rate of uptake. A further control that should be attempted involves capping the quantum dots with a carbohydrate that does not bind to the CRDs (for example ribose). The propylamine cap used in **5.11** has very different chemical properties from the mannose used in **5.9**, so an alternative inert carbohydrate would provide a more robust control experiment.

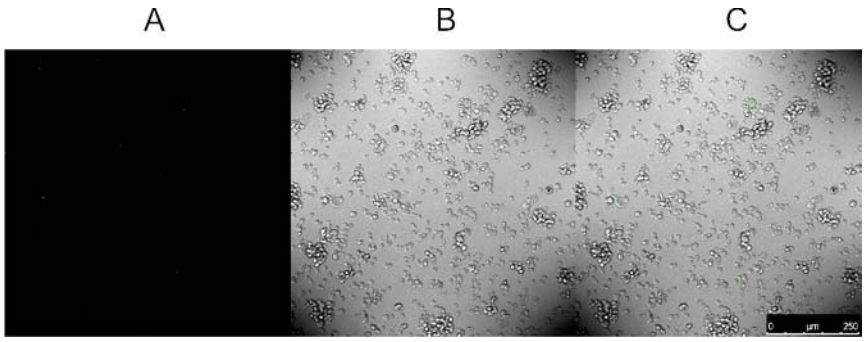


Figure 5.5. *NR8383 cells with* ***5.10*** *after 1 hour. (A) Quantum dot fluorescence (B) brightfield image (C) overlay of A and B. Scale bar = 250µm. Data shows quantum dots capped with propylamine functionality are not taken up into cells.*

Despite successful labelling with **5.9**, the quantum dots were found to be toxic to the cells. Over 2-3 hours, significant cell death had occurred and even cells that survived showed deformities. Time lapse images showed the gradual degradation of the cell structure, which in extreme cases, led to cell explosion (Figure 5.6). Quantum dots are known to be toxic, but the exact mechanism is not fully understood on a cellular level.[29] Various factors, such as heavy metal poisoning or light induced generation of free radicals, may be responsible for the observed apoptosis.

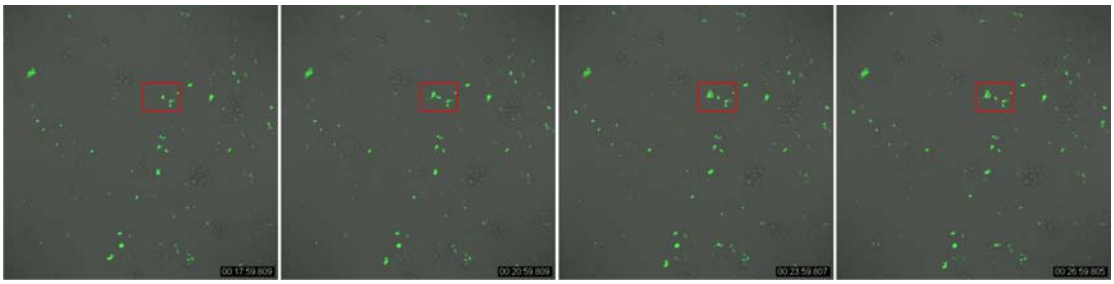


Figure 5.6. *A time lapse of* ***5.9*** *with NR8383. The highlighted region shows the explosive death of a cell after quantum dot internalization. Scale bar = 250µm.*

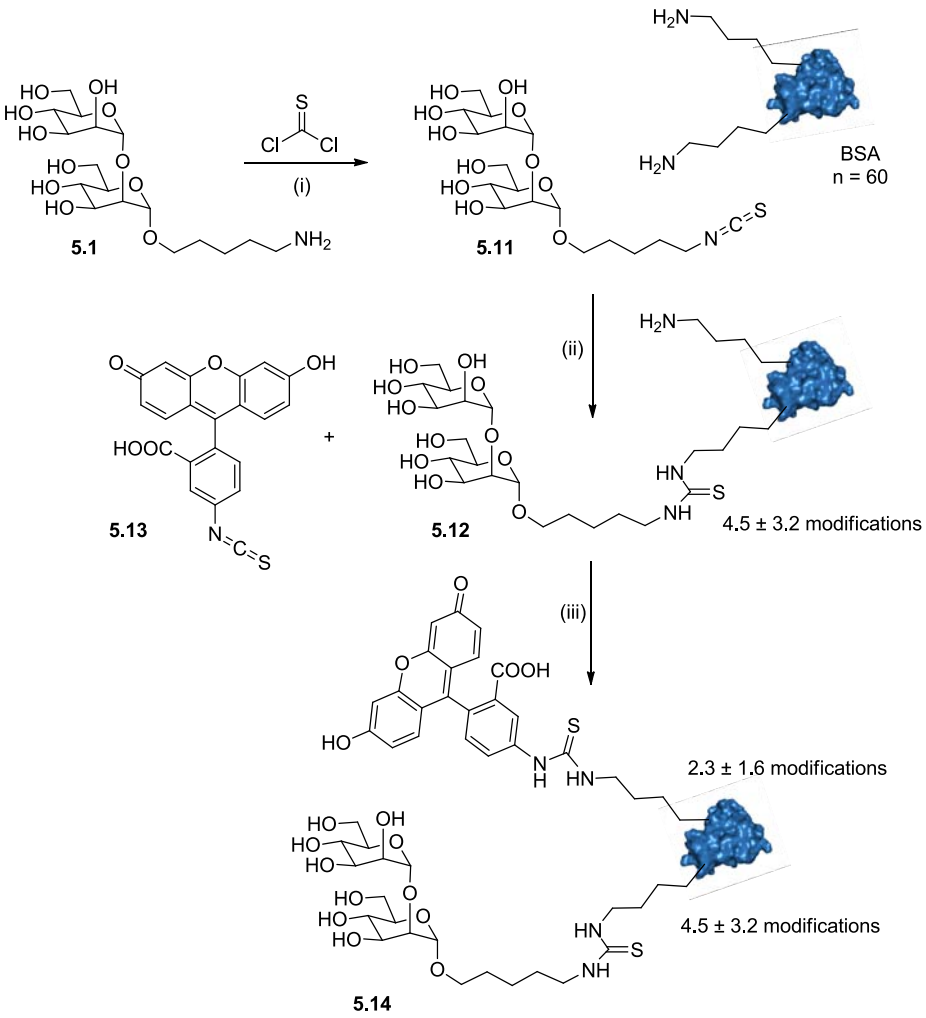
5.4 Fluorescent mannosylated proteins

The toxicity of quantum dots meant that an alternative, bio-compatible probe was needed. Proteins can provide the perfect scaffold for the surface presentation of multiple ligands and several conjugation methodologies have been developed that exploit the chemical reactivity of amino acid side groups.* In particular, thiophosgene coupling chemistry is a facile method for the conjugation of ligands bearing a primary amine with lysine residues, allowing high levels of protein attachment with a stable thiourea linkage.[30] This strategy was used for the synthesis of a mannose-tagged

* For a detailed discussion on protein modification, the reader is directed to Chapter 6, page 273.

fluorescent protein (Scheme 5.6). The reaction of **5.1** with thiophosgene under basic conditions produces the isothiocyanate **5.12**, which is used without further purification for protein modification. BSA is a readily available, stable and non-toxic protein that offers 60 possible ligation sites (59 lysines and the *N*-terminus), so is ideally suited for live cell studies and multiple modifications. Conjugation with **5.12** allowed installation of an average of 4.5 copies of **5.1** per protein (with a dispersity of 3.2 meaning that 95% of the sample contained between 1.3 – 7.7 modifications per protein).[†] This conjugate (**5.13**) was further modified with fluorescein isothiocyanate **5.14** to give a mannosylated fluorescent protein **5.15** tagged with an average of 2.3 dye molecules (with a dispersity of 1.6 modifications meaning that 95% of the sample contained between 0.7 – 3.9 fluorescein molecules per protein).[†] We note that such fluorescent proteins can also be analysed by UV-Visible spectroscopy to determine the extent of fluorescent labelling which would provide further data to support the MALDI characterization. Such an analysis should be considered in future experiments.

[†] The extent of modification and sample dispersity was determined using MALDI mass spectrometry and analysis of the peak width. For a full discussion, the reader is directed to Chapter 6, page 273.



Scheme 5.6. *Conditions: (i) $CSCl_2$, $CHCl_3$, H_2O, 0.3M Na_2CO_3, 0.3M $NaHCO_3$, RT 3h; (ii) 0.3M Na_2CO_3, 0.3M $NaHCO_3$, 12h, RT; (iii) 0.3M Na_2CO_3, 0.3M $NaHCO_3$, 12h, RT. The range in copy numbers was determined from MALDI MS using dispersity analysis (see Chapter 6).*

Pleasingly, uptake of **5.15** by the cultured rat macrophage cell line was detected within 15 minutes of incubation. After 1 hour, there was considerable labelling of the cell membrane and subsequent internalization into intracellular endosomes was observed (Figure 5.7 and Figure 5.8).

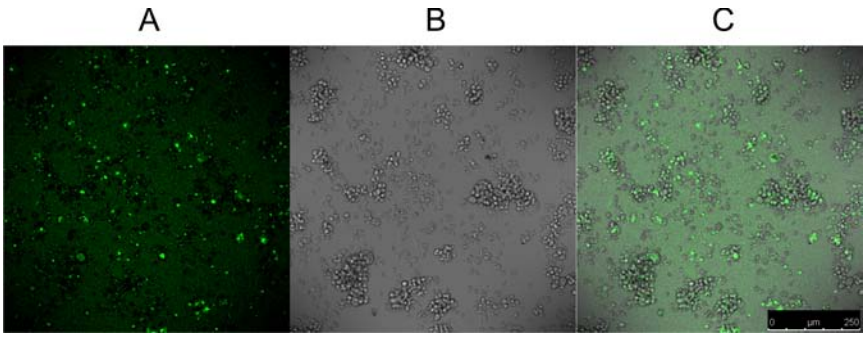


Figure 5.7. *NR8383 cells with **5.15** after 1 hour. (A) Fluorescein fluorescence (B) brightfield image (C) overlay of A and B. Scale bar = 250µm. Data show labelling of cells with fluorescent protein.*

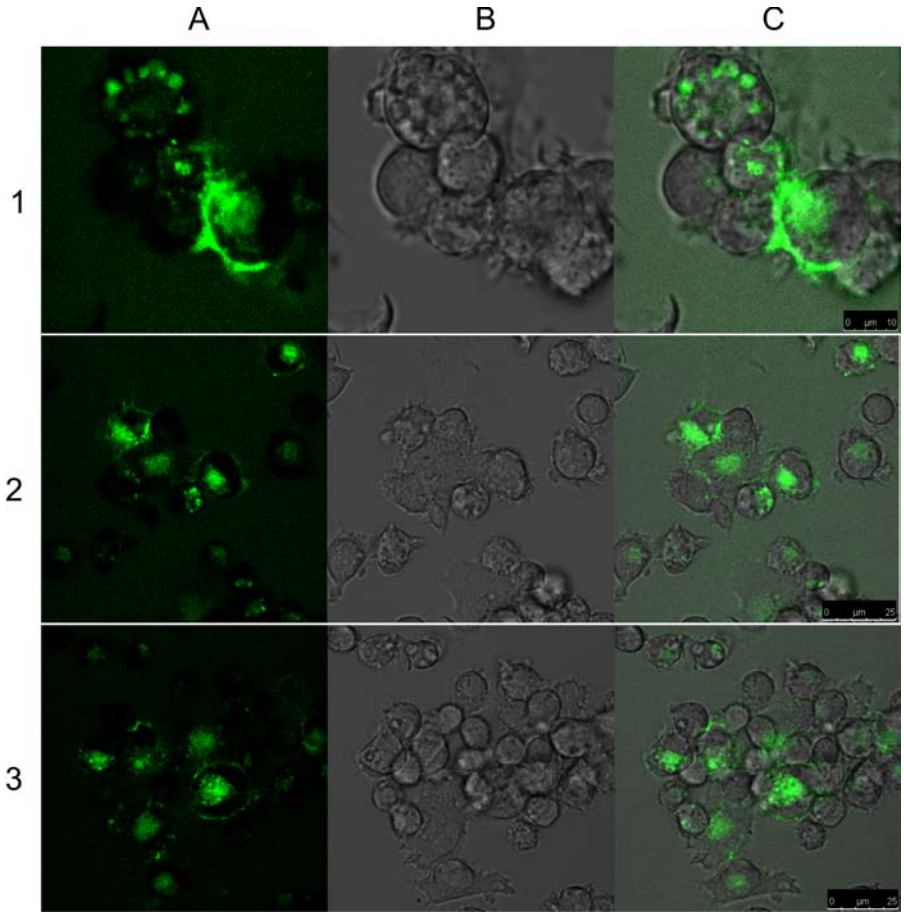


Figure 5.8. *Magnified images of NR8383 cells with* ***5.15*** *after 1 hour. (A) Fluorescein fluorescence (B) brightfield image (C) overlay of A and B. Scale bar = 25µm. Data show internalization of fluorescent protein into cells. Labelling of the cell membrane can be seen and individual endosomes can be observed.*

In contrast to the quantum dots, the protein conjugate was found to be essentially non-toxic even after 2 hours of incubation. Trypan blue was completely excluded by the cells, demonstrating long term cell viability and suggestive of the active transport of **5.15** (Figure 5.9 and Figure 5.10). It is worth noting that microscopy experiments with BSA tend to have a higher background of fluorescein fluorescence compared to the quantum dot experiments. BSA can adhere to the glass slide which may explain these high background readings.

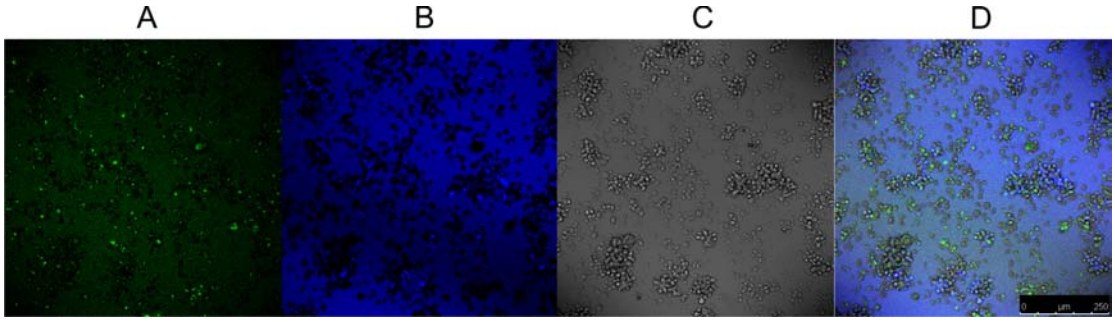


Figure 5.9. *NR8383 cells with* ***5.15*** *after 2 hours co-stained with trypan blue. (A) Fluorescein fluorescence (B) trypan blue fluorescence (C) brightfield image (D) overlay of A, B and C. Scale bar = 250µm. Data shows the exclusion of trypan blue dye indicating the viability of the vast majority of cells.*

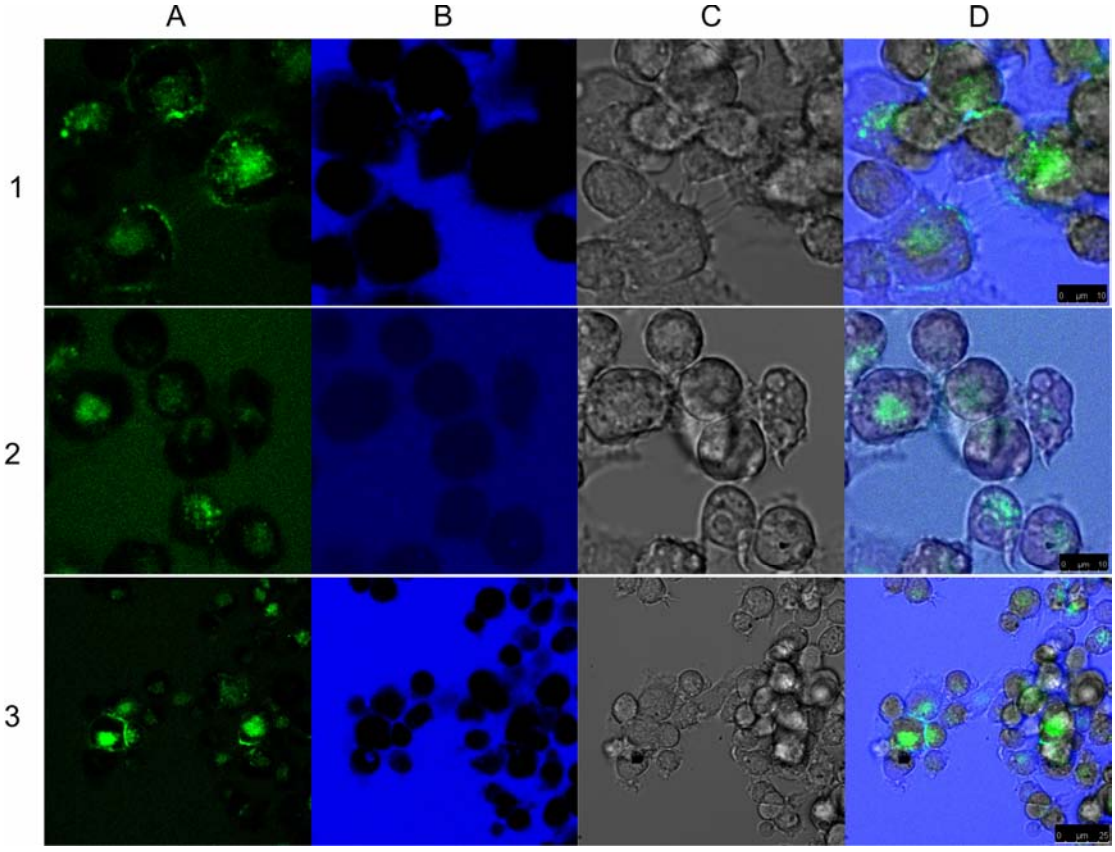


Figure 5.10. *Magnified images of NR8383 cells with* ***5.15*** *after 2 hours co-stained with trypan blue. (A) Fluorescein fluorescence (B) trypan blue fluorescence (C) brightfield image (D) overlay of A, B and C. Scale bar = 25µm. Data shows internalization of fluorescent protein into viable cells which exclude the trypan blue dye. Labelling of the cell membrane can be seen and individual endosomes can be observed.*

For sufficiently magnified images, individual endosomes could be resolved, which provided evidence for the receptor mediated endocytosis of **5.15**. Additionally, neutral red, a fluorescent dye that selectively stains all endosomes and lysosomes,[31] was co-localized with **5.15** for the vast majority of endosomes which suggested an increase in endocytotic activity upon addition of the probe (Figure 5.11). There are more modern alternatives for tracking lysosomes in live cell fluorescent imaging such as

LysoTracker® which have better fluorescent properties compared to neutral red. For future experiments, it would be preferable to use these alternative dyes.

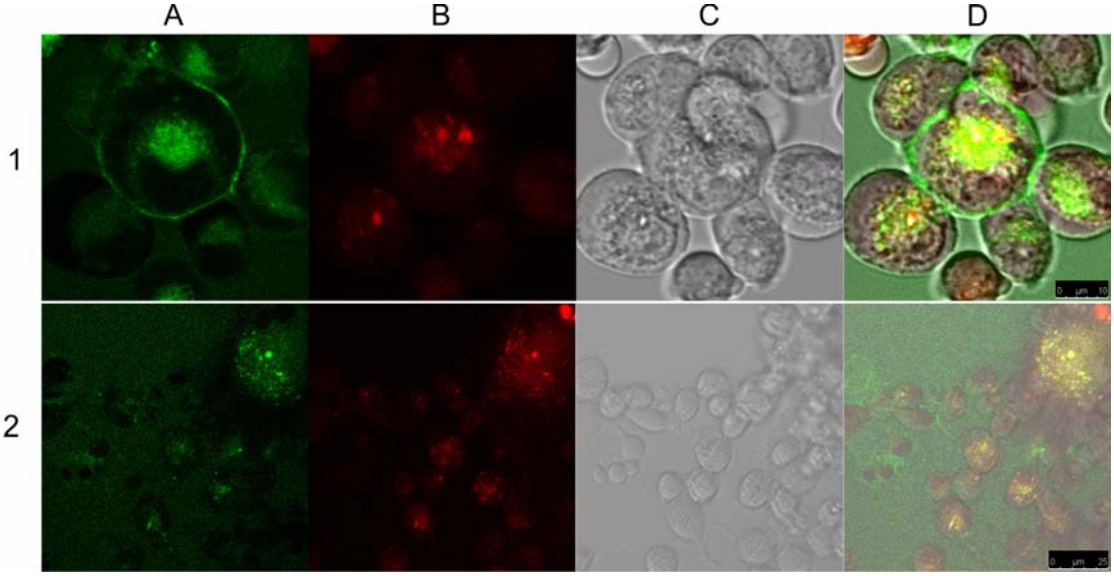
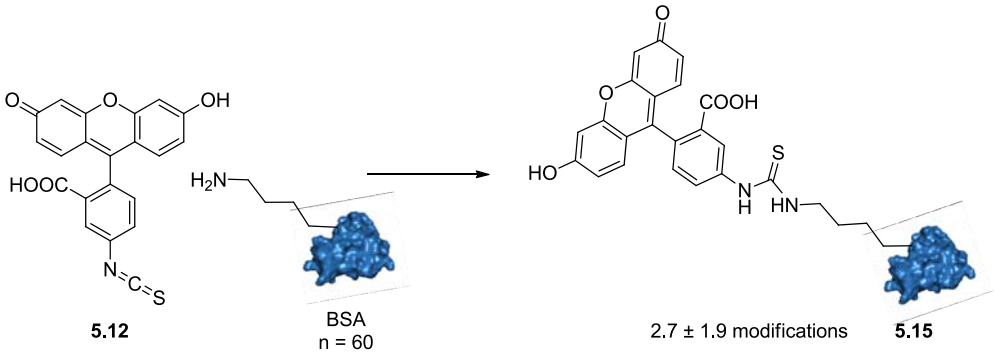


Figure 5.11. *Magnified images of* ***5.15*** *with NR8383 cells after 2 hours co-stained with neutral red. (A) Fluorescein fluorescence (B) neutral red fluorescence (C) brightfield image (D) overlay of A, B and C. Scale bar = 25µm. Neutral red is an indicator for endosomes and lysosomes. Data shows colocalisation of fluorescent protein with neutral red indicating uptake of the protein into endosomes via endocytosis.*

Fluorescent protein lacking the critical mannose moiety (**5.16**) was also synthesized as a control to demonstrate the necessity of the sugar (Scheme 5.7). Incubation with NR8383 cells failed to produce any internalization, even after 1 hour, ruling out alternative non-specific uptake mechanisms such as pinocytosis.[32] Thus, the internalization of **5.15** likely proceeds through endocytosis. As with the quantum dot experiments described above, alternative controls are needed to verify this assertion. Competition experiments with **5.1** would determine whether the sugar is responsible for cellular uptake as described in section **5.3**.



Scheme 5.7. *Conditions: 0.3M Na_2CO_3, 0.3M $NaHCO_3$, 12h, RT. The range in copy numbers was determined from MALDI MS using dispersity analysis (see Chapter 6).*

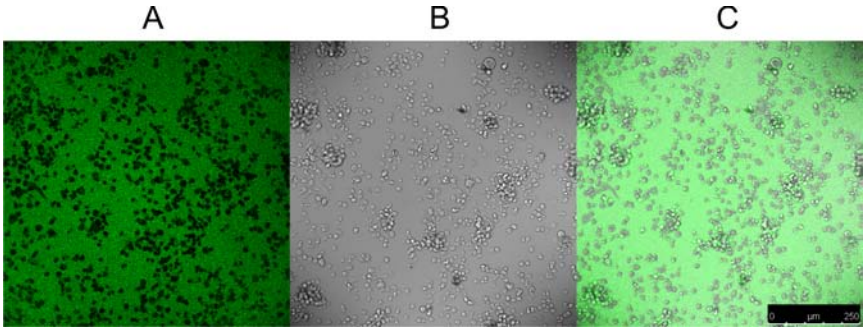


Figure 5.12. *NR8383 cells with* ***5.16*** *after 1 hour. (A)Fluorescein fluorescence (B) brightfield image (C) overlay of A and B. Scale bar = 250µm. Data shows that fluorescent protein without the mannose tag is not taken up by cells.*

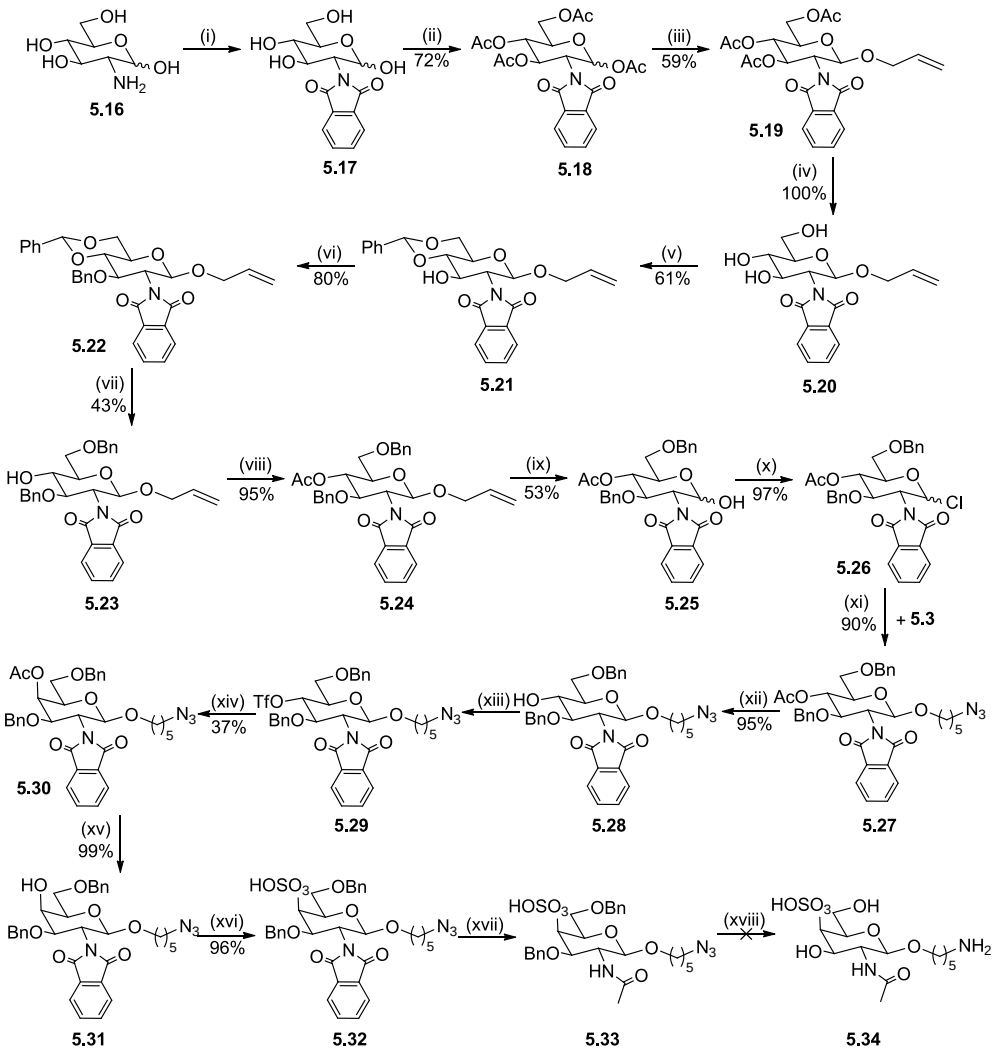
In all the above experiments, we have demonstrated that uptake of our mannose probes is possible, but conclusive evidence for binding to the MR is still required. Here, we suggest a variety of potential experiments that could be performed to accomplish this. To demonstrate the expression of the MR on NR8383 cells, a fluorescently labelled antibody specific to the receptor could be used with successful labelling of the cell surface providing conclusive evidence of expression of the protein. To prove that uptake of our mannose probes was due to binding with this receptor, a control experiment should be conducted with macrophages derived from a $MR^{-/-}$ knockout rat in which uptake would not be observed. Alternatively, a cell line that does not normally express the MR could be transfected with the gene for the receptor. Thus by comparing the uptake between the wild type and transfected cells, the uptake attributable to the MR can be directly determined.

5.5 Attempted synthesis of an *N*-acetyl-galactose-4-sulfate probe

With an effective probe for the CRDs in hand, attention was then turned towards the CRR that is involved in the binding of sulfated carbohydrates. GalNAc-4-SO_4 is known to be a potent natural ligand, so was chosen as the basis of our initial synthetic target. Our synthetic route was based off a previous synthesis of the core GalNAc-4-SO_4 sugar,[33] but slightly modified to allow incorporation of an amine terminated linker at the anomeric position. Unfortunately, this synthesis failed at the final deprotection step and was unsuccessful (Scheme 5.8).

Glucosamine, **5.17**, was protected at the amine with a phthalic group (**5.18**) and subsequently per-*O*-acetylated to give **5.19**. Glycosylation with allyl alcohol furnished **5.20** in 59%, where the allyl functionality served as a protecting group. Deacetylation

gave **5.21** in quantitative yield, which was protected with a 4,6-benzylidene in 61% yield to give selective access to the 3-position (**5.22**). Benzylation of this position with sodium hydride and benzyl bromide proceeded smoothly in 80% to give **5.23**. Ring opening of the benzylide gave selective access to the 4-position **5.24** in a reasonable 43%, which was acetylated in near quantitative yield to furnish **5.25**. The allyl group was isomerized with $(PPh_3)_3RhCl$, allowing hydrolysis with mercuric (II) oxide to free the anomeric position (**5.26**) in 53% yield. Treatment with oxallyl chloride smoothly gave the anomeric chloride **5.27**, which allowed glycosylation with **5.3** in 90% yield to furnish **5.28**. Treatment with sodium methoxide gave selective access to the 4-position **5.29**, which was converted to the triflate **5.30** using triflic anhydride in pyridine and DCM. The crude product was reacted with *tert*-butylammonium acetate (as a soluble source of the acetate ion) which allowed inversion of stereochemistry by S_N2 to the desired galactose configuration. Thus, after HPLC, **5.31** was produced in 37% yield. Once again, deacetylation gave selective access to the 4-position **5.32**, which was sulfated with a sulphur trioxide pyridinium complex to afford **5.33** in an excellent 96% yield. Deprotection of the phthalic group with hydrazine acetate proceeded smoothly to give the free amine, which was acetylated to furnish **5.34**. This highly polar compound was used without further purification. At this stage, the synthesis completely failed. The final benzyl and azide deprotection was first attempted with hydrogenolysis over palladium on charcoal, but the reaction was unsuccessful and neither the starting material nor the product could be recovered. Several different solvent combinations were screened, as well as the use of acid catalysis, but all were ineffective. A frustrating aspect of this reaction was the complete loss of carbohydrate material, which could not be detected by TLC or even crude ^1H NMR.



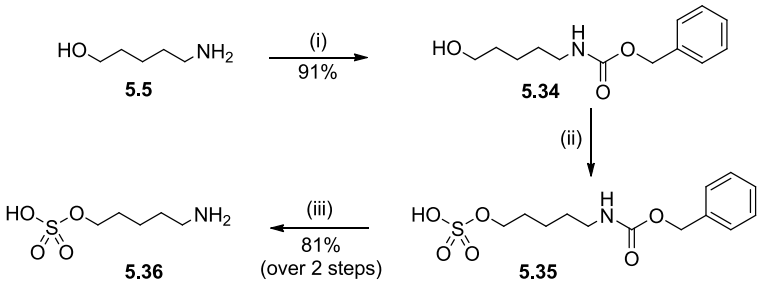
Scheme 5.8. *Conditions: (i) Phthalic anhydride, NaOH, H_2O; (ii) Ac_2O, Py; (iii) allyl alcohol, $SnCl_4$, DCM; (iv) NaOMe, MeOH; (v) dimethyl acetal, TsOH, CH_3CN, reflux; (vi) BnBr, NaH, DMF, 0°C to RT; (vii) Et_3SiH, MsOH, DCM, -78˚C to RT; (viii) Ac_2O, Py; (ix) $(PPh_3)_3RhCl$, DABCO, EtOH, toluene, H_2O, 100°C then HgO, $HgCl_2$, acetone, H_2O;(x) oxallyl chloride, DMF, DCM; (xi) AgOTf, collidine, DCM, 3Å molecular sieves, -40°C; (xii) NaOMe, MeOH; (xiii) Tf_2O, Py, DCM; (xiv) Bu_4NOAc, DMF, RT; (xv) NaOMe, MeOH; (xvi) Py·SO_3, DMF, 60°C; (xvii) NH_2NH_2·AcOH, MeOH, reflux, then Ac_2O, Py; (xviii) H_2, Pd/C, MeOH. H_2O.*

5.6 Synthesis of a generic sulfate probe

The crystal structures of the MR bound to various sulfonated ligands at the CRR have been solved.[11,12] From these structures, it can be seen that the majority of strong hydrogen bonds solely involve the sulfonate group, with the carbohydrate moiety being mostly involved in other peripheral van der Waals interactions. In fact, out of the eight direct H-bonds between the CRR and GalNAc-4-SO_4, six involve the sulfonate group. While the other weaker hydrophobic interactions with the galactose ring are important in fine tuning the selectivity of the CRR, we hypothesized that the

carbohydrate component may not be entirely necessary for binding. Indeed, a crystal structure has been solved with the CRR binding to a molecule of HEPES buffer,[12] and surface plasmon resonance studies have shown that the compound has a reasonable binding affinity for the receptor,[5] demonstrating that the carbohydrate moiety is dispensible. Furthermore, it was found that HEPES binds to the receptor with a K_D of 3.9mM compared to 0.11mM for the GalNAc-4-SO_4.[12] Considering the structural simplicity of a HEPES molecule compared to the natural ligand, the drop in binding affinity is surprisingly small.

For these reasons, we rationalized that a simple sulfonate group tethered to a C_5 amine-terminated linker would provide enough binding for our CRR probes. Whereas HEPES bears a sulfonic acid group, the CRR crystal structures show that a sulfonate probe would benefit from an extra water-mediated hydrogen bond at the oxygen linkage of the sulfate ester, and thus should have an improved avidity. While this was not an ideal solution, the difficulty in synthesizing the desired GalNAc-4-SO_4 scaffold meant that a suitable alternative was needed, at least for these preliminary studies.



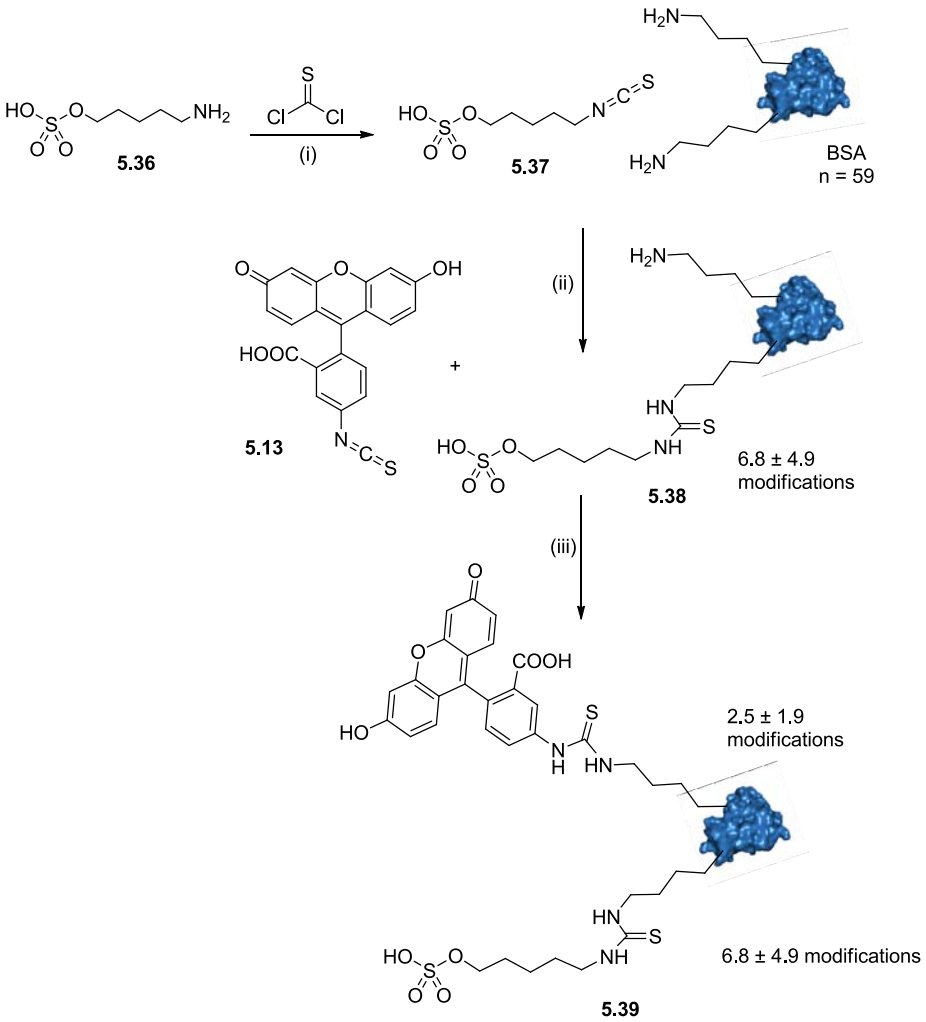
Scheme 5.9. *Conditions: (i) benzyloxycarbonyl chloride, Na_2CO_3, DCM, H_2O, RT; (ii) sulfonic acid, Py, DCM, RT; (iii) H_2, Pd/C, MeOH.*

A simple three step route was required for the synthesis (Scheme 5.9). Carboxybenzyl protection of the amine on compound **5.5** afforded **5.35** in 91% yield. With selective access the hydroxyl functionality, sulfonation was effected with chlorosulfonic acid to produce **5.36**, which was subjected to hydrogenolysis without further purification to furnish the free amine and the target compound **5.37** in 81% yield over two steps.

5.7 Fluorescent sulfonated proteins

The strategy for building a fluorescent probe mirrored that of mannose. BSA was used as a scaffold for initial attachment of **5.37** using isothiocyanate chemistry (Scheme 5.10). This allowed the installation of an average of 6.8 units of **5.37** per protein (with

a dispersity of 4.9 meaning that 95% of the sample contained between 1.9 – 11.7 modifications per protein).[‡] The higher yield of attachment compared to **5.1** is likely due to the lower steric bulk for **5.37**, allowing higher rates of reaction.[§] This was followed by fluorescein ligation to furnish a fluorescent sulfonated protein **5.40** modified with an average of 2.5 dye molecules per BSA (with a dispersity of 1.9 meaning that 95% of the sample contained between 0.6 – 4.4 modifications per protein).[‡]



Scheme 5.10. *Conditions: (i) $CSCl_2$, $CHCl_3$, H_2O, 0.3M Na_2CO_3, 0.3M $NaHCO_3$, RT 3h; (ii) 0.3M Na_2CO_3, 0.3M $NaHCO_3$, 12h, RT; (iii) 0.3M Na_2CO_3, 0.3M $NaHCO_3$, 12h, RT. The range in copy numbers was determined from MALDI MS using dispersity analysis (see Chapter 6).*

[‡] The extent of modification and sample dispersity was determined using MALDI mass spectrometry and analysis of the peak width. For a full discussion, the reader is directed to Chapter 6, page 273.

[§] For a detailed discussion on the trends in the extent of protein modification, the reader is directed to Chapter 6, page 273.

The sulfonated protein **5.40** was incubated with cultured cells from the NR8383 cell line, which pleasingly, showed internalization within 15 minutes of addition. After 1 hour, significant uptake of **5.40** had occurred (Figure 5.13) and with sufficient magnification, individual endosomes could be resolved (Figure 5.14).

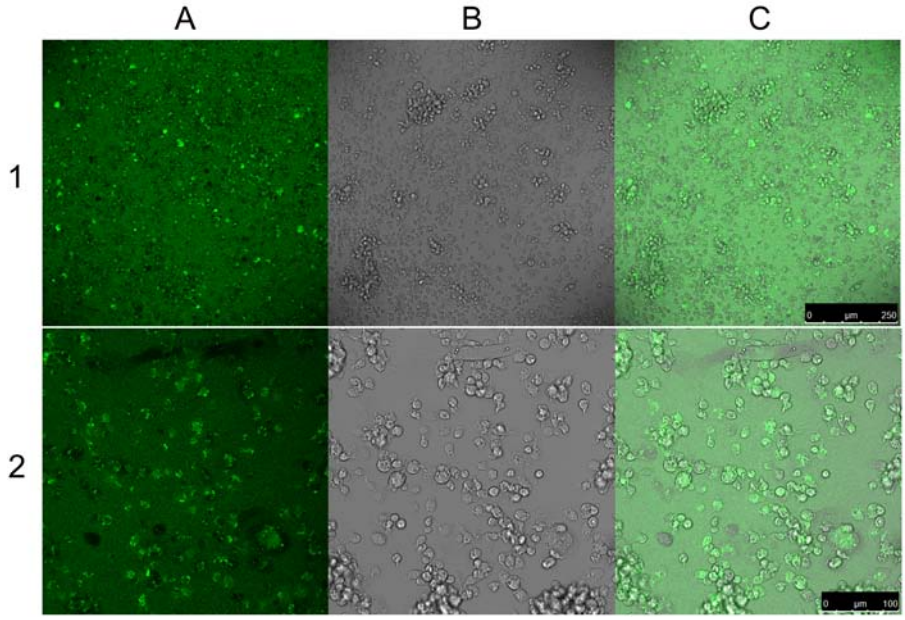


Figure 5.13. *NR8383 cells with after 1 hour. (A) Fluorescein fluorescence (B) brightfield image (C) overlay of A and B. Scale bar = 100µm. Data shows labelling of cells with fluorescent protein.*

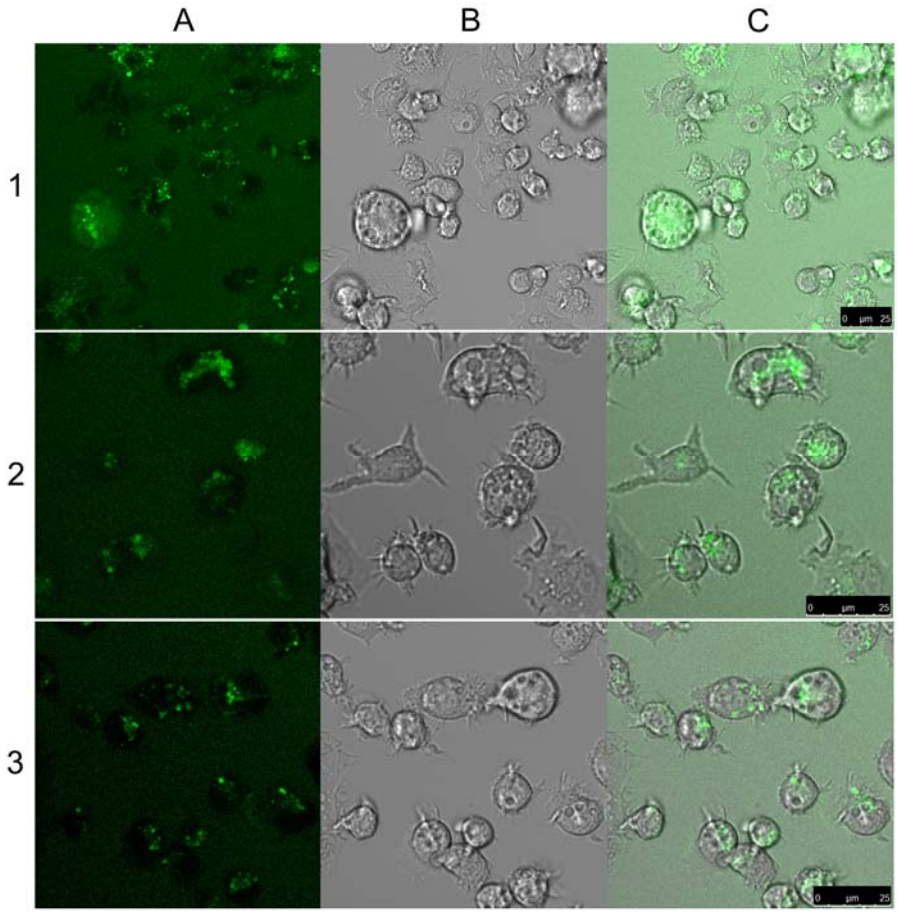


Figure 5.14. *Magnified images of NR8383 cells with after 1 hour. (A) Fluorescein fluorescence (B) brightfield image (C) overlay of A and B. Scale bar = 25µm. Data shows internalization of fluorescent protein into endosomes.*

As with **5.15**, long-term cell viability and the active transport of **5.40** into the cells were demonstrated with the trypan blue exclusion method, with almost all cells still being viable after 2 hours of incubation.

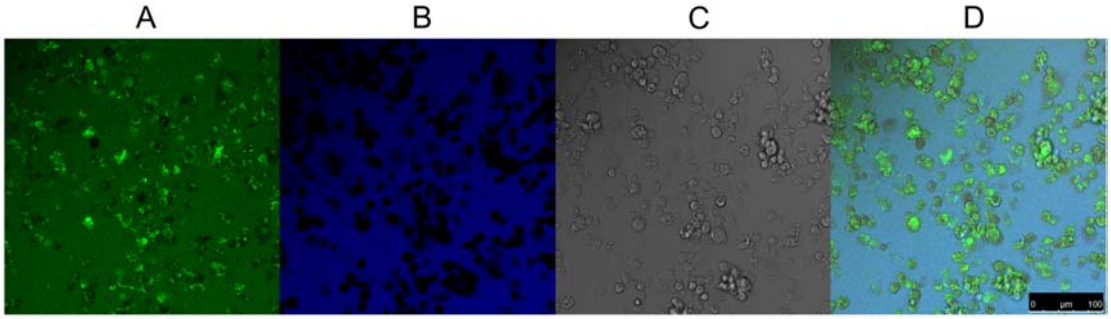


Figure 5.15. *NR8383 cells with after 2 hours co-stained with trypan blue. (A) Fluorescein fluorescence (B) trypan blue fluorescence (C) brightfield image (D) overlay of A, B and C. Scale bar = 100µm. Data shows the exclusion of trypan blue dye indicating the viability of the vast majority of cells.*

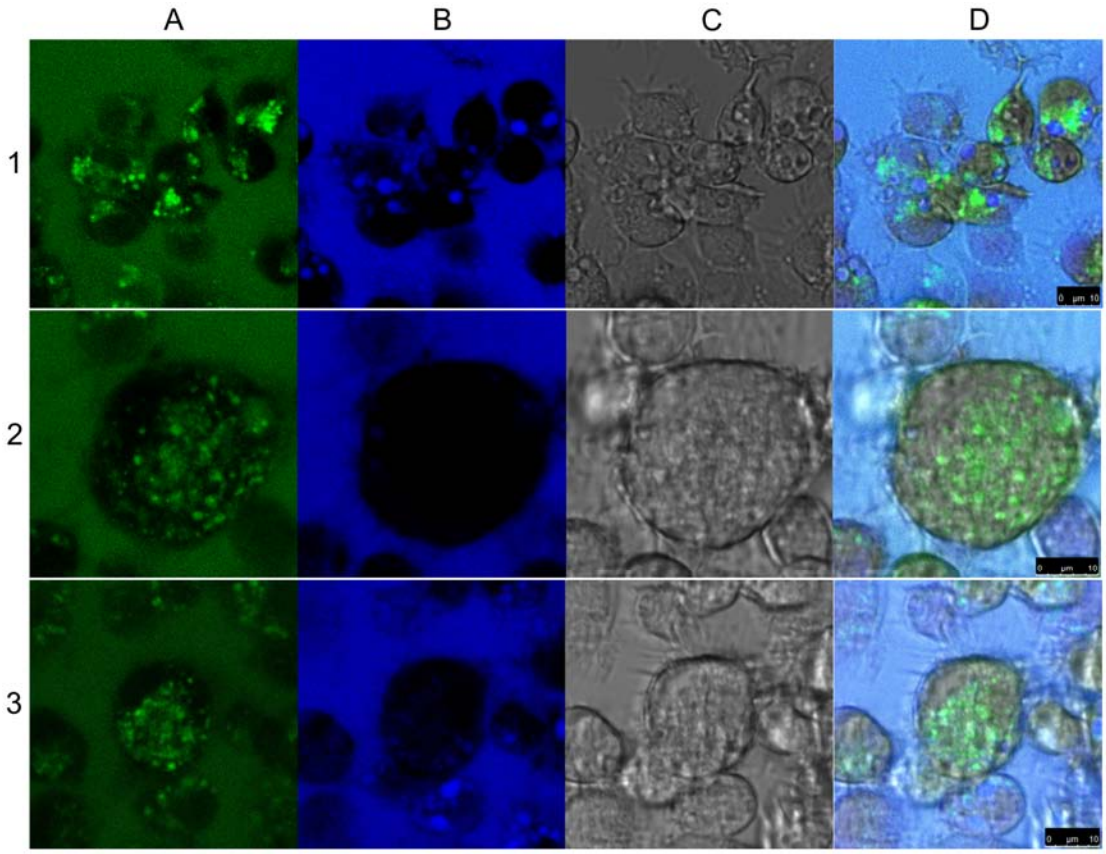


Figure 5.16. *Magnified images of NR8383 cells with after 2 hours co-stained with trypan blue. (A) Fluorescein fluorescence (B) trypan blue fluorescence (C) brightfield image (D) overlay of A, B and C. Scale bar = 50µm. Data shows internalization of fluorescent protein into viable cells which exclude the trypan blue dye. Labelling of the cell membrane can be seen and individual endosomes can be observed.*

Neutral red was used as a co-stain to visualize endosomes and lysosomes. As before, there was a large amount of co-localization between fluorescein and neutral red fluorescence (Figure 5.17). Thus, a large proportion of the resolved endosomes also contained **5.40**, which reveals an increase in endocytosis upon addition of the probe. As with the experiments involving **5.1**, we recommend that future experiments be conducted with more modern alternatives to trypan blue and neutral red.

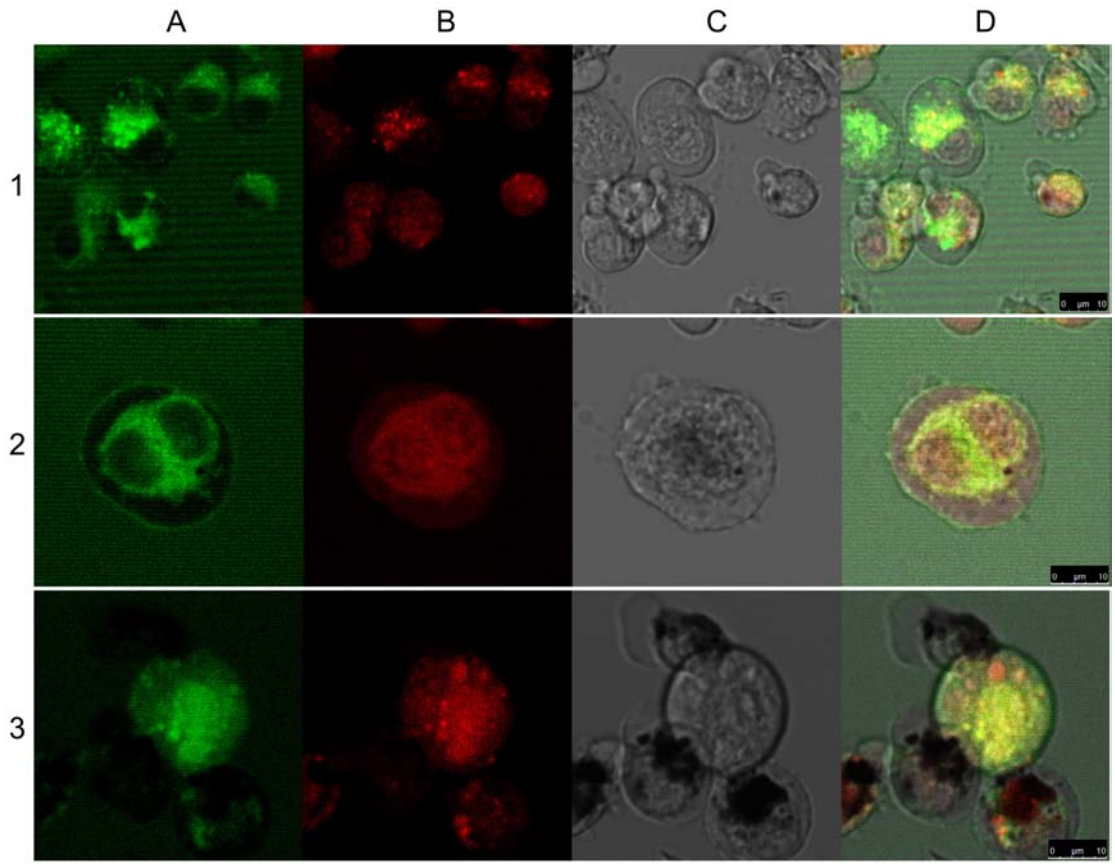


Figure 5.17. *Magnified images of with NR8383 cells after 2 hours co-stained with neutral red. (A) Fluorescein fluorescence (B) neutral red fluorescence (C) brightfield image (D) overlay of A, B and C. Scale bar = 50µm. Neutral red is an indicator for endosomes and lysosomes. Data shows colocalisation of fluorescent protein with neutral red indicating uptake of the protein into endosomes via endocytosis.*

While these experiments demonstrate that **5.39** is taken up by cells, binding to the MRR still remains to be conclusively proven. In future experiments, the desired compound **5.34** would ideally be synthesized. Furthermore, interaction of these probes with the CRR should be experimentally verified using ITC or surface plasmon resonance on recombinant fragment proteins of the MR. As with the mannose probes, further experiments using $MR^{-/-}$ derived macrophages, or cell lines transfected with the MR gene should be used to conclusively demonstrate MR mediated endocytosis.

5.8 Probing receptor dimerization with FRET microscopy

Even though our previous experiments with BSA based fluorophores still require rigorous validation, we decided to conduct some initial testing on FRET microscopy to establish whether such a technique would be viable for determining receptor dimerization.

The measurement of FRET signals can be technically very difficult.[17] Fluorophores are generally chosen to maximize the spectral overlap between donor emission and acceptor absorbance to improve the FRET efficiency. However, this requirement creates problems during image capture due to spectral crosstalk, where the emission of one fluorophore bleeds into the acquisition channel of the other, and cross-excitation where the donor frequency also partially excites the acceptor fluorescence. Of the various methods available for FRET detection, sensitized emission (which simply involves monitoring of both channels after donor excitation) is the most suitable for dynamic systems because both channels can be acquired simultaneously, but crosstalk is still a major problem.[17,34] To correct for this behaviour, donor-only and acceptor-only samples are measured as controls, from which correction factors can be determined. Several algorithms have been developed for the implementation of these techniques.[35] A widely used algorithm developed by van Rheenen *et al*. allows determination of an apparent FRET efficiency (E_A) and is used herein.[36]

For our FRET fluorophores, we used the commercially available ATTO dyes from ATTO-TEC, which are available as *N*-hydrosuccinimide esters that allow facile modification with amines such as **5.1** and **5.37** (section 5.11, page 251). We chose three different dyes (ATTO 520, 565, and 633)[**] that could be used in two different FRET pair combinations with high overlap integrals (Figure 5.18).

[**] The structures of all ATTO dyes are not available due to proprietary reasons and are therefore referred to by name only.

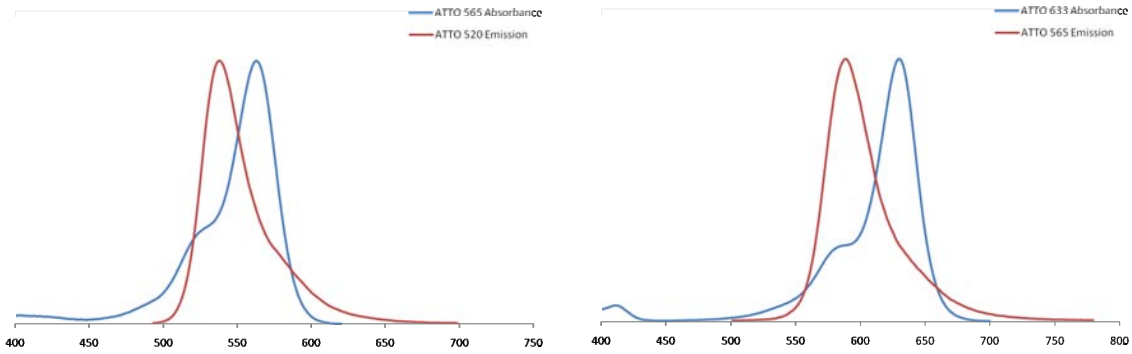


Figure 5.18. *Donor emission and acceptor absorbance spectral profiles showing FRET overlap integrals for the ATTO 520 and ATTO 565 pair (left) and the ATTO 565 and ATTO 633 pair (right).*

These dyes were modified with **5.1** to produce CRD specific FRET probes (section 5.11.6, page 274). These probes were individually incubated with NR8383 cells to determine correction factors for sensitized emission analysis and to ensure adequate localization at the cell membrane (Figure 5.19).

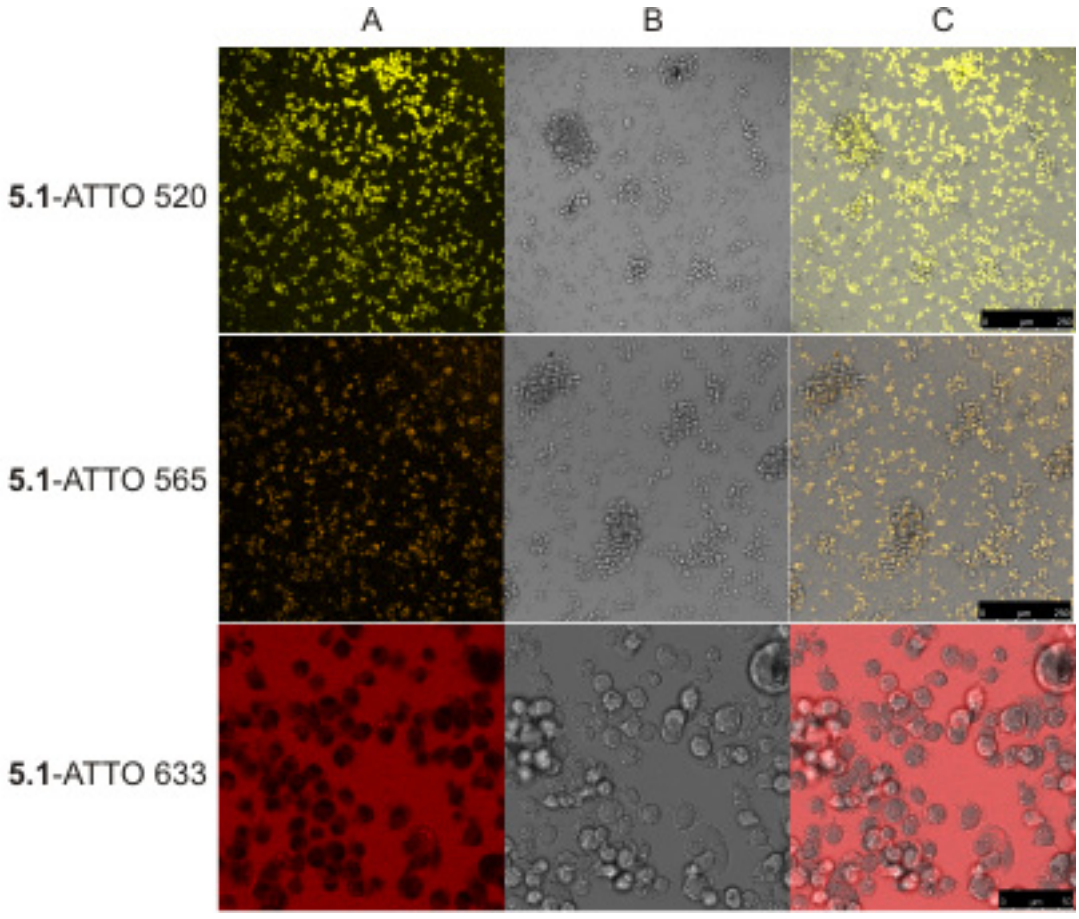


Figure 5.19. *NR8383 cells with ATTO dyes modified with **5.1** after 5 minutes incubation. (A) ATTO fluorescence (B) brightfield image (C) overlay of A and B. Scale bar = 50µm. Data shows that ATTO 520 and ATTO 565 are readily incorporated into the cell membrane. ATTO 633 is more hydrophilic and less absorbed into the membrane.*

5.1-ATTO 520 and 565 were immediately absorbed into the cell membrane. Note that this localisation is not due to interactions with the MR, but is instead attributed to the lipophilic nature of the dyes. In contrast, **5.1**-ATTO 633 was far more hydrophilic and consequently exhibited much lower incorporation into the cell membrane, requiring up to 30 minutes of incubation for any significant levels of absorption.

FRET experiments were performed on NR8383 cells with the ATTO 520 and 565 pair (Figure 5.20, rows 1-3) and the **5.1**-ATTO 565 and 633 pair (Figure 5.20, rows 4-6). The samples were irradiated with the donor excitation wavelength and images were acquired of the donor fluorescence (channel A) and acceptor fluorescence (channel B), which was immediately followed by image capture of acceptor fluorescence under acceptor excitation (channel C). These images were corrected for crosstalk and subsequently used to determine apparent FRET efficiencies,[36] which were expressed using a tricolour red-green-blue (RGB) look-up table (LUT) that provides a readily interpretable representation on the extent of energy transfer.

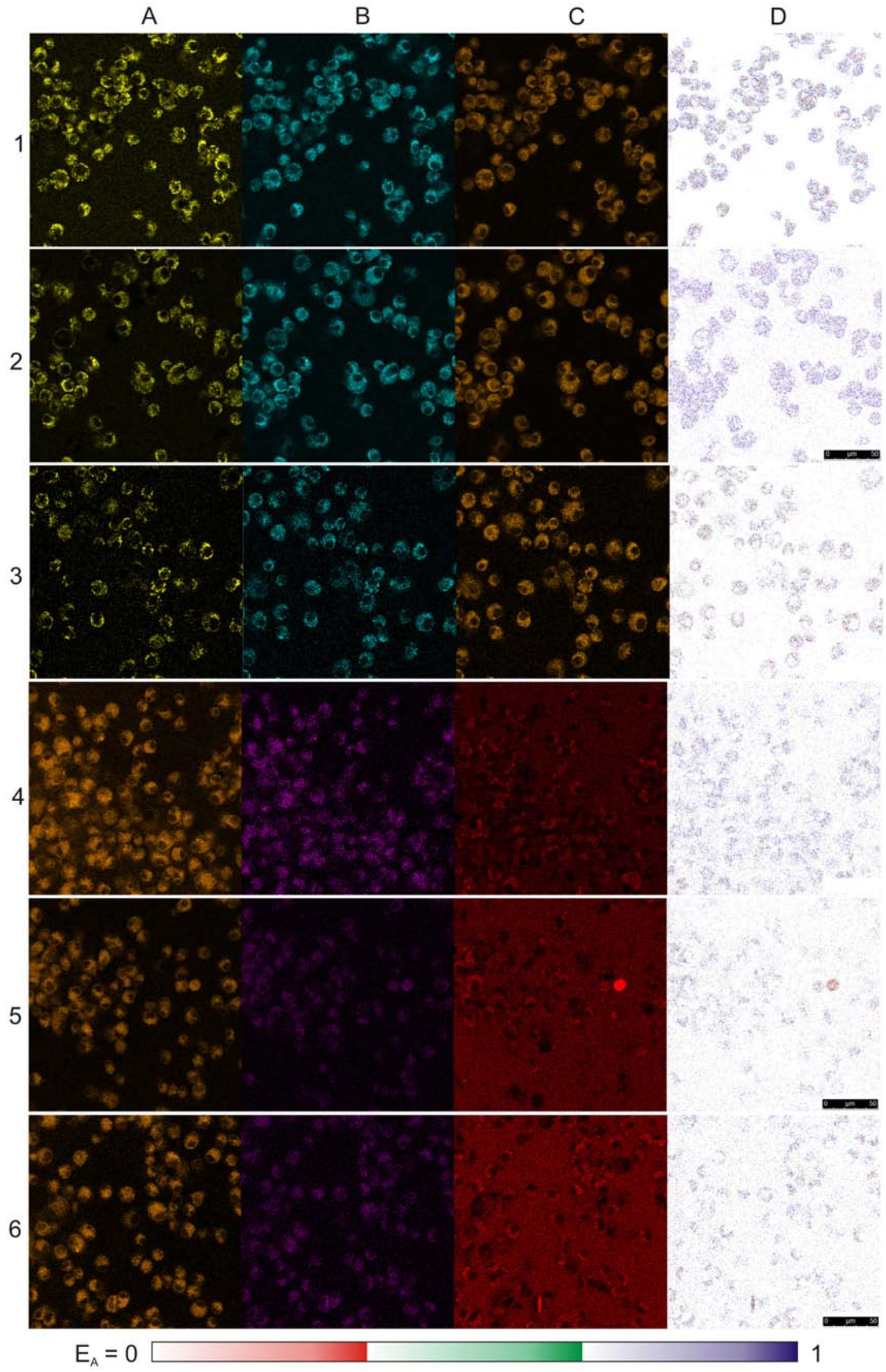


Figure 5.20. *FRET experiments on dyes modified with* ***5.1*** *on NR8383 cells;* ***5.1****-ATTO 520+565 (1-3);* ***5.1****-ATTO 565+633 (4-6). (A) Donor excited, donor emission; (B) donor excited, acceptor emission; (C) acceptor excited, acceptor emission; (D)* E_A *expressed through the tricolour RGB LUT. Scale bar = 250µm. Data shows a high degree of FRET enhancement at the cell membrane for both dye pairs.*

High FRET enhancements (blue – E_A between 66-100%) were detected at the cell surfaces even for the latter FRET pair, despite the lower membrane incorporation of ATTO 633. These FRET data provide complimentary information to the standard fluorescence images. At the resolution of standard confocal microscopy, the acceptor and donor dyes have a high degree of co-localization. However, FRET enhancement requires much more intimate contact between the two fluorophores, and can therefore differentiate between occupation of the same pixel (~200 nm resolution) versus binding to the same protein (~10 nm distances). This difference is exemplified for non-viable cells that are flooded with both ATTO dyes (row 5, Figure 5.20 and Figure 5.21). At the confocal level of resolution, there is a high degree of co-localization between the two dyes, but the FRET image shows very little enhancement (red – E_A between 0-33%) indicating that although the two fluorophores occupy the same space, there are no short range interactions between the two. In comparison, the high FRET signals at the cell membrane for viable cells indicates proximal binding, consistent with the multivalent CRD binding modes in the MR.

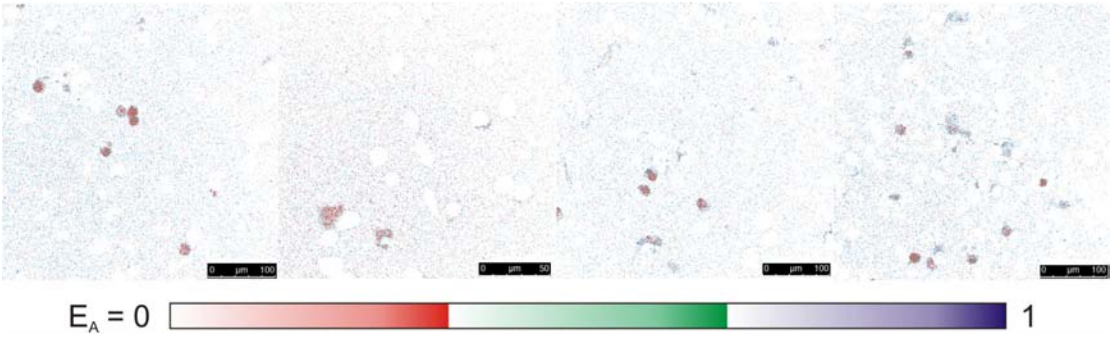


Figure 5.21. *FRET images of non-viable cells with* ***5.1****-ATTO 565+633. Scale bar = 50µm Images were acquired immediately after addition of dyes to minimize uptake of* ***5.1****-ATTO 633 into the cell membranes of viable cells. Non-viable cells are immediately flooded with both ATTO dyes, but despite this intracellular co-localization, there is insignificant FRET enhancement due to lack of sub-10nm proximity.*

However, there are alternative explanations for the FRET signals that should be rigorously explored in future experiments. For instance, the close proximity of the dyes may arise through non-specific binding to the cell membrane resulting in a high surface concentration. Alternatively, endocytosis of the probes would result in a high concentration within a single endosome which could explain the observed FRET enhancements. To elucidate whether either of these mechanisms is responsible for the observed FRET, experiments should be conducted with different concentrations of ATTO dyes or with use of sodium azide to restrict endocytosis. Additionally,

experiments with $MR^{-/-}$ knockout macrophages or MR transfected cell lines could provide more rigorous evidence for MR mediated FRET.

We then turned our attention to the CRR specific FRET probes. As before, the ATTO dyes were modified with the sulfated ligand **5.37** (section 5.11.6, page 274) and individually incubated with NR8383 cells to determine correction factors for sensitized emission analysis (Figure 5.22). **5.37**-ATTO 520 and 565 were quickly incorporated into the cell membrane, but once again **5.37**-ATTO 633 was poorly taken up. Hence, the **5.37**-ATTO 565 and 633 pair was deemed unsuitable for FRET studies and was not used. Instead, FRET images were obtained for the **5.37**-ATTO 520 and 565 pair by sequential irradiation and subsequent analysis of the raw data as above (Figure 5.23).

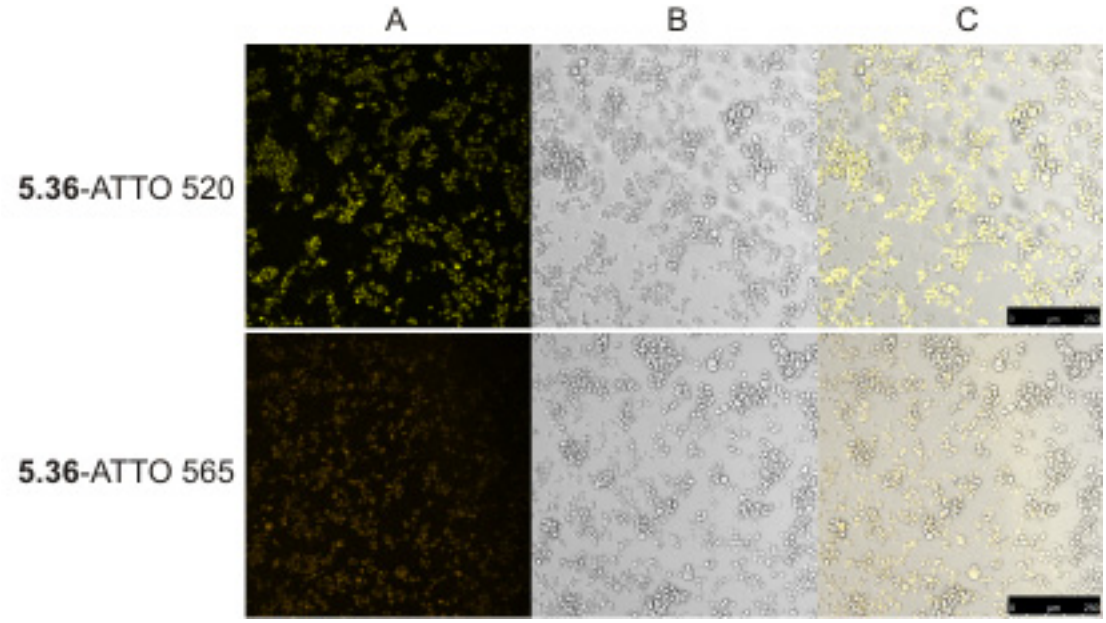


Figure 5.22. *NR8383 cells with ATTO dyes modified with* ***5.37*** *after 5 minutes incubation. (A) ATTO fluorescence (B) brightfield image (C) overlay of A and B. Scale bar = 250µm. Data shows that both ATTO dyes are readily absorbed into the cell membrane.*

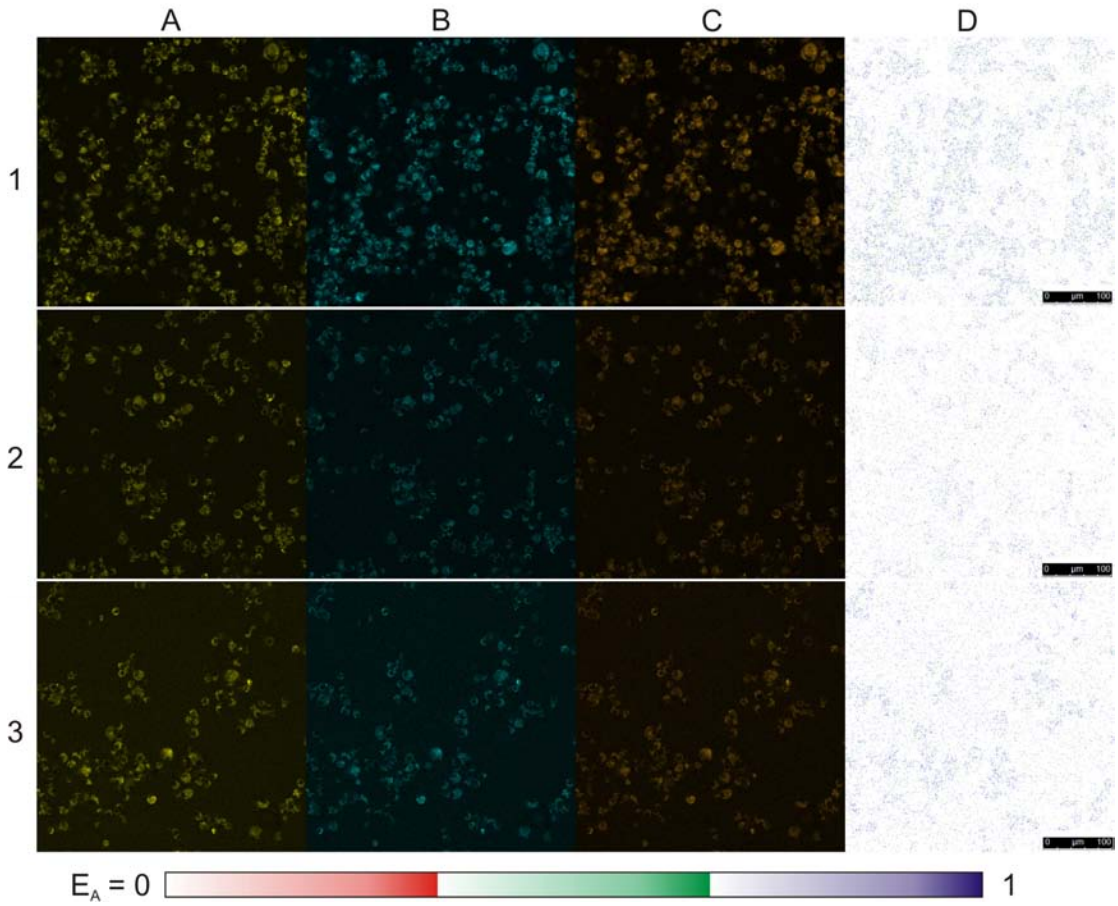


Figure 5.23. *FRET experiments on ATTO 520+565 modified with* ***5.37*** *on NR8383 cells. (A) Donor excited, donor emission; (B) donor excited, acceptor emission; (C) acceptor excited, acceptor emission; (D)* E_A *expressed through the tricolour RGB LUT. Scal bar = 100µm. Data shows FRET enhancement at the cell surface, but is lower than the corresponding enhancements observed for* ***5.1.***

Significant FRET enhancements were detected at the cell surface (blue – E_A between 66-100%). Such a result may be suggestive of MR dimers and would be consistent with the rapid endocytosis of **5.40**, which is reliant upon bivalent binding with the receptor.[2-5] However, as with the FRET experiments on the CRD specific probes, further control experiments would be needed to fully verify this hypothesis. Firstly, the binding of **5.37** to the CRR should be confirmed using recombinant protein as discussed above. FRET experiment. As with the CRD FRET pair, microscopy could be performed at differing concentrations of ATTO dyes or with sodium azide to elucidate the role of non-specific binding or endosome formation in the observed enhancements. Once again, experiments with transfected cell lines or $MR^{-/-}$ would provide more rigorous evidence.

While further experiments are clearly needed to fully establish the origin of our FRET enhancements, our early studies indicate that further studies may be worthwhile. If the

methodology can be rigorously demonstrated, a noteworthy advantage of these FRET images would be the prospect of determining more quantitative information on the extent of MR dimerization. The RGB representation of the FRET data allows for a straightforward histogram analysis of the images, where simple counting of blue pixels that relate to high E_A facilitates a numerical comparison of the close proximity interactions for the CRD and CRR specific probes. By accounting for the differential binding modes of the two probes (Figure 5.1) and their differences in binding affinity, the relative proportions of monomer and dimer may be reasonably approximated.

5.9 Endothelial liver derived cell line

Having studied the FRET microscopy technique on a cell line derived from rat macrophages, we decided to explore the methodology on other cell types. Endothelial liver cells also have a good expression of the MR where a high level of dimerization has been detected, so would make a good system for further studies.[2-5] For experimental convenience, we chose to perform investigations with the immortal cell line SK-HEP-1 which is derived from human endothelial liver cells, rather than using a live source. Our workflow paralleled the previous experiments with NR8383 cells. Initially, the fluorescent mannosylated protein **5.15** and the fluorescent sulfated protein **5.40** were incubated with cultured SK-HEP-1 cells, which were monitored for internalization (Figure 5.24 and Figure 5.25 respectively).

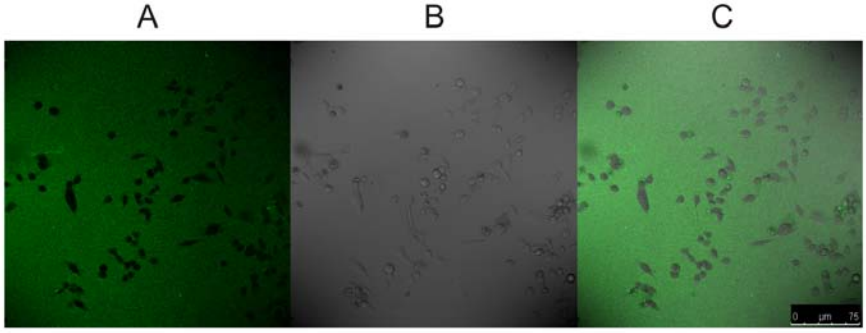


Figure 5.24. *SK-HEP-1 cells with* ***5.15*** *after 3 hours. (A) Fluorescein fluorescence (B) brightfield image (C) overlay of A and B. Scale bar = 75µm. Data shows poor uptake of mannosylated fluorescent protein even after incubation for 3h.*

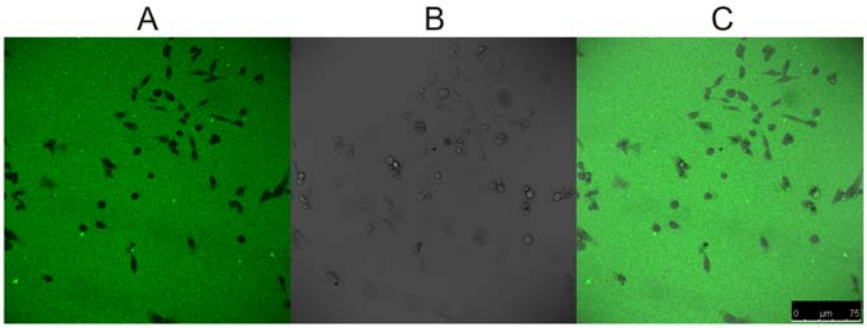


Figure 5.25. *SK-HEP-1 cells with after 3 hours. (A) Fluorescein fluorescence (B) brightfield image (C) overlay of A and B. Scale bar = 75µm. Data shows poor uptake of sulfated fluorescent protein even after incubation for 3h.*

Unlike the NR8383 cells, the endothelial liver derived cells showed much less uptake of **5.15** and **5.40**, even after 3 hours of incubation. Small amounts of internalization could be detected when individual cells were magnified (Figure 5.26 and Figure 5.27; rows 1 – 3). The cells were washed with phosphate buffered saline (PBS) to remove unbound residual protein, which eliminated the background fluorescence and dramatically improved contrast allowing individual endosomes to be visualized (Figure 5.26 and Figure 5.27; rows 4 – 6). Endosomes were stained with neutral red and the vast majority showed co-localization with fluorescein fluorescence, indicating an increase in endocytosis upon addition of the probes. Furthermore, to prove that internalization occurred through receptor mediated endocytosis rather than pinocytosis,[32] the cells were incubated with the control fluorescent protein **5.16** where, even after 3 hours, no uptake was observed (Figure 5.28).

Visual inspection of these images demonstrated that the sulfated probe **5.40** showed greater uptake than the mannosylated analogue **5.15**, although both created substantially lower endocytosis than with NR8383 cells. Hence, it appeared that the cultured SK-HEP-1 cells suffered from low expression of the MR. In future studies, the expression levels of MR in different cells lines could be determined using a fluorescent labelled antibody as discussed above for the NR8383 cells.

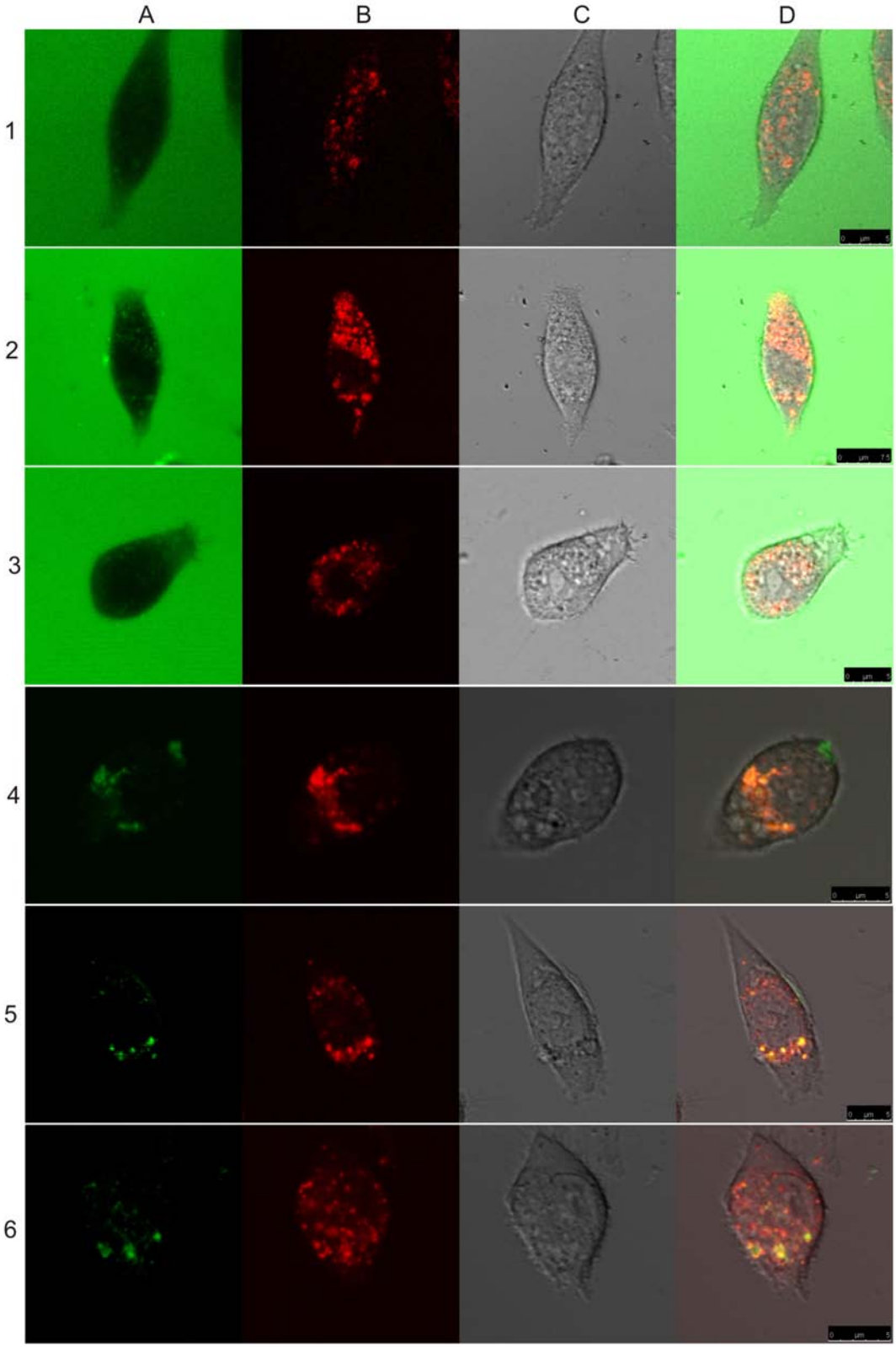


Figure 5.26. *Magnified images of* ***5.15*** *with SK-HEP-1 cells after 3 hours co-stained with neutral red (1-3) and after washing with PBS to remove unbound protein (4-6). (A) Fluorescein fluorescence (B) neutral red fluorescence (C) brightfield image (D) overlay of A, B and C. Scale bar = 5µm. Data shows uptake of mannosylated fluorescent protein into endosomes, which are stained with neutral red.*

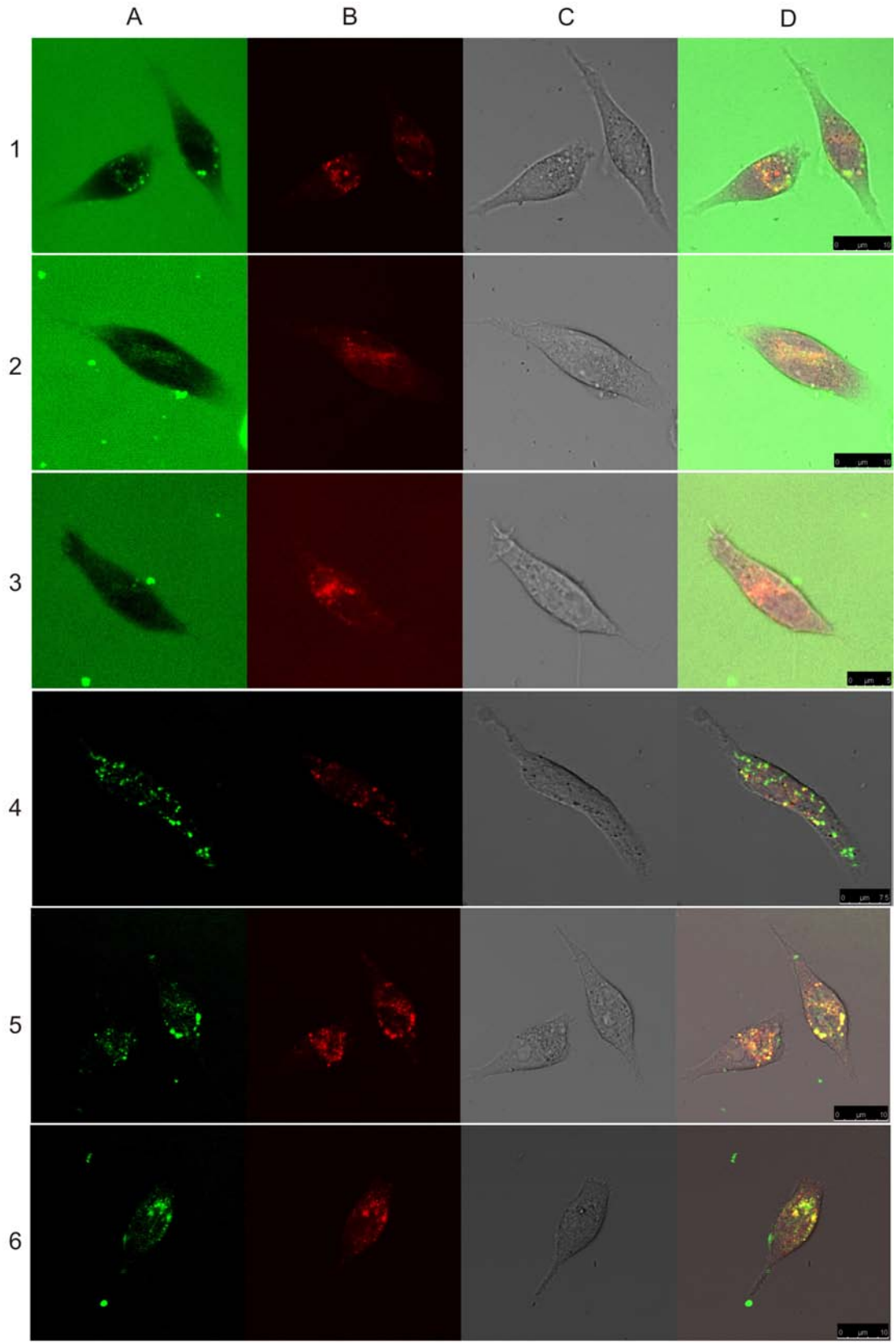


Figure 5.27. *Magnified images of with SK-HEP-1 cells after 3 hours co-stained with neutral red (1-3) and after washing with PBS to remove unbound protein (4-6). (A) Fluorescein fluorescence (B) neutral red fluorescence (C) brightfield image (D) overlay of A, B and C. Scale bar = 10µm. Data shows uptake of sulfated fluorescent protein into endosomes, which are stained with neutral red.*

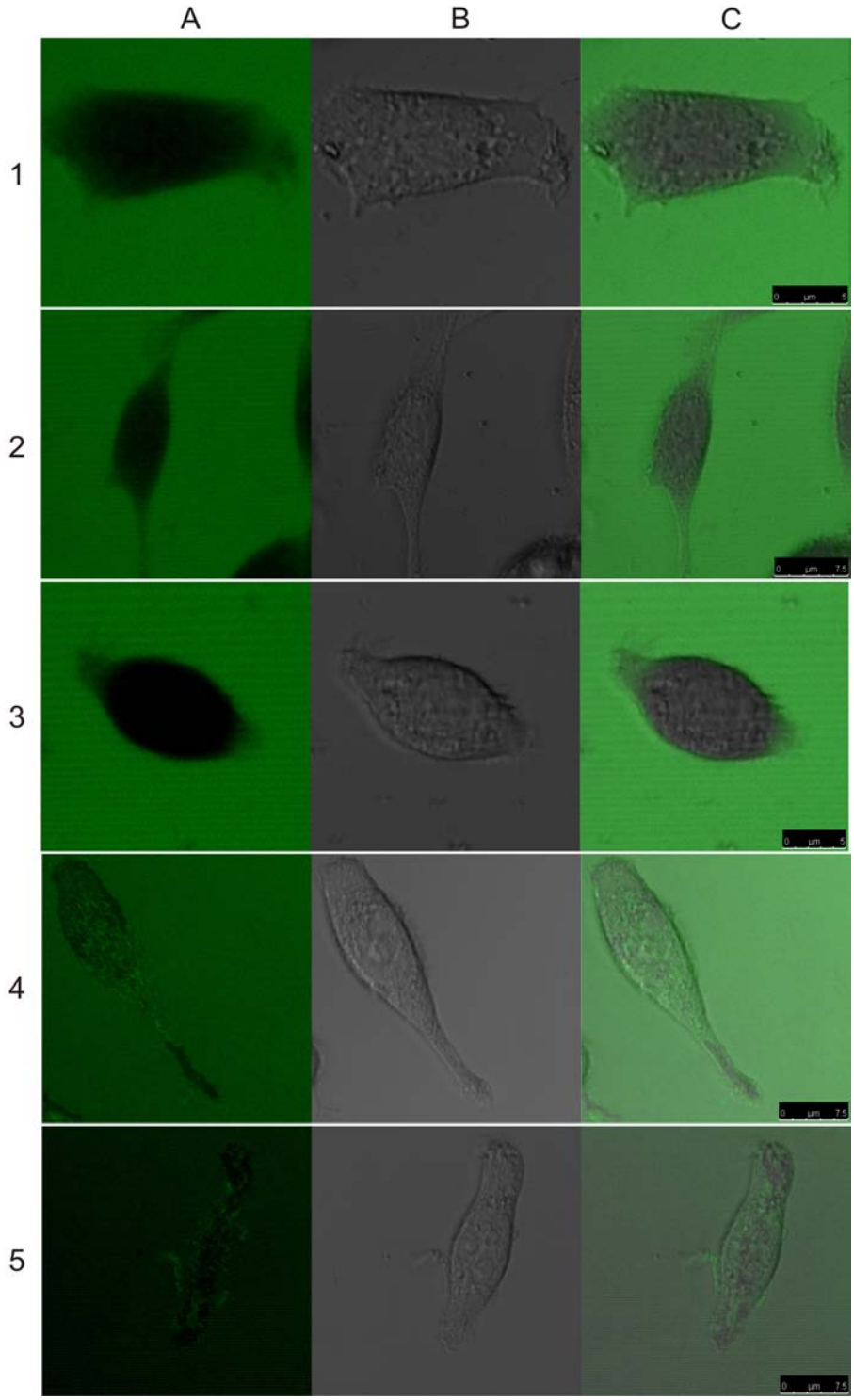


Figure 5.28. *Magnified images of SK-HEP-1 cells with* ***5.16*** *after 3 hours (1-3) and after washing with PBS to remove unbound protein (4-5). (A) Fluorescein fluorescence (B) brightfield image (C) overlay of A and B. Scale bar = 35µm. Data shows that fluorescent protein without the mannose or sulfate tag is not taken up by cells.*

The SK-HEP-1 cells were then investigated with our CRD and CRR specific FRET probes. As before, absorption into the cell membrane was validated with the modified

ATTO dyes, and the images were used to generate correction factors for sensitized emission analysis (Figure 5.29 and Figure 5.30).

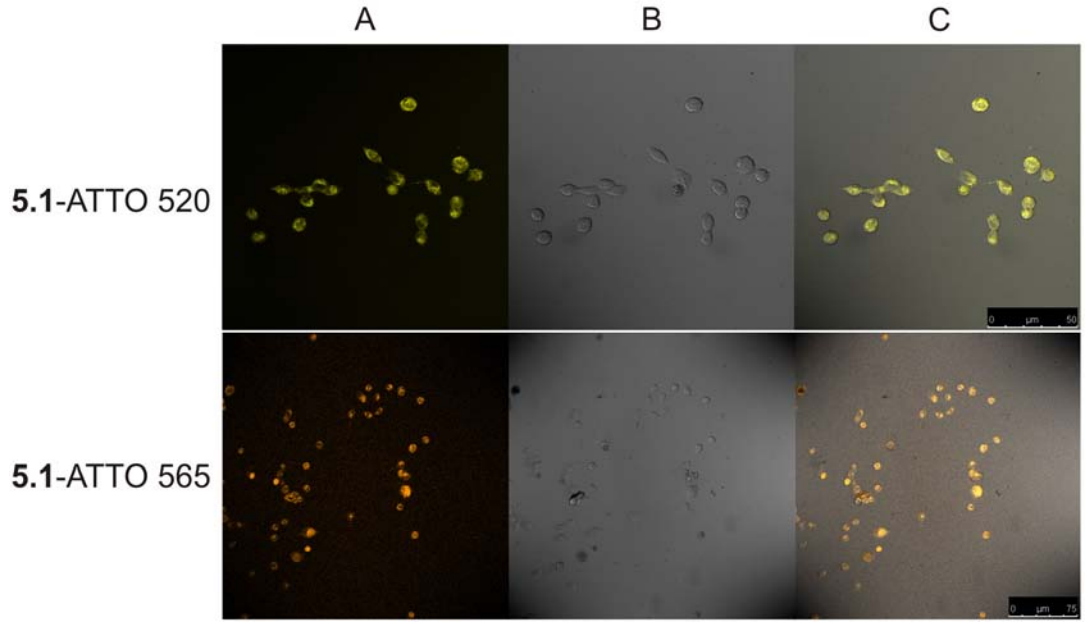


Figure 5.29. *SK-HEP-1 cells with ATTO dyes modified with* ***5.1*** *after 5 minutes incubation. (A) ATTO fluorescence (B) brightfield image (C) overlay of A and B. Sclae bar = 75µm. Data shows that mannosylated ATTO 520 and ATTO 565 are readily incorporated into the cell membrane.*

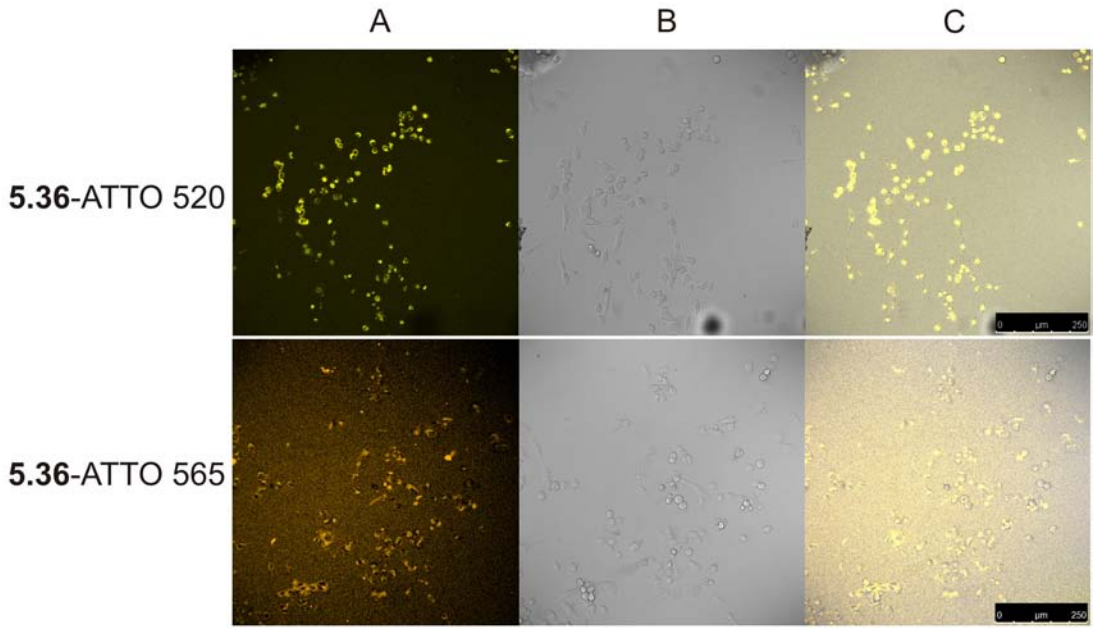


Figure 5.30. *SK-HEP-1 cells with ATTO dyes modified with* ***5.37*** *after 5 minutes incubation. (A) ATTO fluorescence (B) brightfield image (C) overlay of A and B. Sclae bar = 75µm. Data shows that mannosylated ATTO 520 and ATTO 565 are readily incorporated into the cell membrane.*

FRET images were acquired using the previous protocols with the CRD specific (Figure 5.31, rows 1 – 3) and CRR specific (Figure 5.31, rows 4 – 6) ATTO dyes.

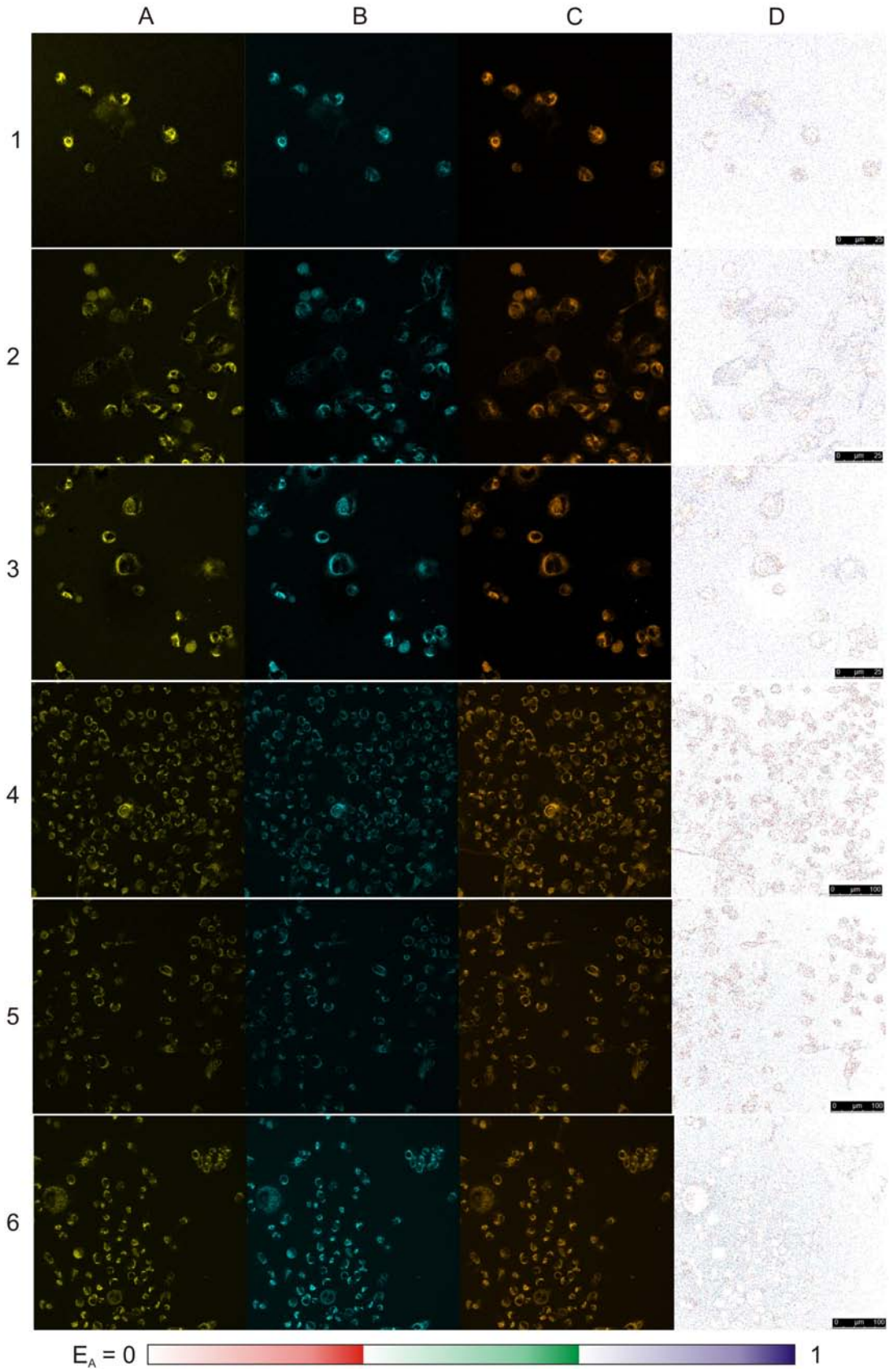


Figure 5.31. *FRET experiments on ATTO 520+565 modified with* ***5.1*** *(1-3) or* ***5.37*** *(4-6) on SK-HEP-1 cells. (A) Donor excited, donor emission; (B) donor excited, acceptor emission; (C) acceptor excited, acceptor emission; (D) E_A expressed through the tricolour RGB LUT. Sclae bar = 100µm. Data shows only a low level of FRET enhancement at the cell surface compared to NR8383 cells*

As expected for cells with poor expression of the MR, very little FRET enhancement was detected (red – E_A between 0-33%). These results, while slightly disappointing, serve as a good negative control for our FRET methodology and clearly demonstrate the need for close-range interactions to observe FRET enhancements.

The lack of MR expression for these cultured cells compared to endothelial liver cells from a live source is most likely attributable to inadequacies with the SK-HEP-1 cell line. Such immortal cell lines, especially those of a cancerous origin, as is the case here, often have an altered physiology from the natural system and it appears that the ability to express the MR has been severely weakened in the SK-HEP-1 system. The immortal endothelial cell line TMNK-1,[37] or primary cells may be suitable alternatives. Alternatively, the SK-HEP-1 cell line could be transfected with the MR gene to provide rigorous evidence for binding of our probes to the receptor.

5.10 Conclusions

The dimerization of cell surface receptors is common across the natural world.[22,38] The study of this phenomenon is vital due to the profound effect it can have on downstream biochemical processes. This is especially true for the MR, where dimerization is known to be essential for the binding of sulfated carbohydrates, yet the biological implications of this are still poorly understood. Our investigations into probing this dimerization are still in their infancy and several control experiments are required to fully validate our methodology. Most importantly, our labelling experiments should be performed on a simplified system to isolate the role of the MR from other effects. As discussed above, $MR^{-/-}$ derived cells, or transfected cell lines would be an ideal method to achieve this. However, if this methodology could be successfully developed, we believe it would be a powerful tool for studying the involvement of MR dimerization in the immune system about which very little is presently known.

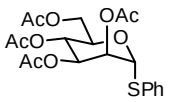
Current hypotheses invoke a role for the MR in linking innate and adaptive immunity, where the MR is involved in both antigen uptake through the CRDs, and subsequent sulfate mediated targeting to secondary lymphoid organs via the CRR, which likely involves receptor dimerization (section 1.3.2, page 14).[4,5,10] Not surprisingly, the expression of the receptor is heavily regulated. An experiment can be envisaged

where machinery of the immune system, such as macrophages, lymphocytes and dendritic cells, are monitored with FRET microscopy at various points during the immune response to quantify and track changes in the extent of MR dimerization. In this way, a truly comprehensive study on MR dimerization could be conducted. Given the potential importance of such a pathway, further experiments to develop our microscopy technique are warranted.

5.11 Experimental

For general experimental conditions see section 2.8.1 page 44.

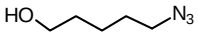
5.11.1 Synthesis of the mannose ligand



α-thiophenyl-2,3,4,6-*O*-acetyl-mannopyranside[39] (5.2)

To a solution of **5.4** (12.0g, 390mmol, 1eq) in anhydrous DCM (50ml) was added $BF_3 \cdot OEt_2$ (7.6ml, 62mmol, 2eq) at room temperature. After 5 minutes of stirring, PhSH (6.3ml, 62mmol, 2eq) was added dropwise. The reaction mixture was stirred at room temperature for 15 hours, after which TLC (3 ethyl acetate : 2 petrol) showed complete consumption of starting material. The reaction was quenched with ice (~100ml). The layers were allowed to separate and the organic layer was washed with saturated $NaHCO_3$ solution (100ml), brine (100ml) and dried over $MgSO_4$. The product was purified by automated silica gel chromatography (10% ethyl acetate in petrol to 100% ethyl acetate over 11 column volumes) to afford the title compound as a colourless oil (10.9g, 80%).

R_f 0.43 (1 petrol : 1 ethyl acetate); $[\alpha]_D^{25}$ +109.5 (c = 1.0, $CHCl_3$), lit.[39] $[\alpha]_D^{27}$ +107.2 (c = 1.0, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.01, 2.05, 2.07, 2.15 (4 x 3H, 4 x s, 4 x OAc), 4.10 (1H, dd, $J_{H6a\text{-}H6b}$ 12.1Hz, $J_{H5\text{-}H6b}$ 2.4Hz, H6b), 4.30 (1H, dd, $J_{H6a\text{-}H6b}$ 12.3Hz, $J_{H5\text{-}H6a}$ 6.0Hz, H6a), 4.54 (1H, ddd, $J_{H4\text{-}H5}$ 8.9Hz, $J_{H5\text{-}H6a}$ 6.0Hz, $J_{H5\text{-}H6b}$ 2.2Hz, H5), 5.31-5.34 (2H, m, H3, H4), 5.48-5.51 (2H, m, H1, H2), 7.28-7.34 (3H, m, ArH_{ortho}, ArH_{para}), 7.48 (2H, m, ArH_{meta}); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.6, 20.7, 20.7, 20.8 (4 x 1C, 4 x CO<u>C</u>H_3), 62.4 (1C, C6), 66.3 (1C, C4), 69.3 (1C, C3), 69.5 (1C, C5), 70.8 (1C, C2), 85.6 (1C, C1), 128.1 (1C, ArC_{para}), 129.1 (2C, ArC_{ortho}), 129.2 (1C, ArC_{meta}), 132.0 (1C, ArC_{ipso}), 169.7, 169.8, 169.9, 170.5 (4 x 1C, 4 x C=O); m/z (ES^+) 458.2 $[M + NH_4]^+$, 463.1 $[M + Na]^+$.

 **5-azidopentan-1-ol (5.3)**

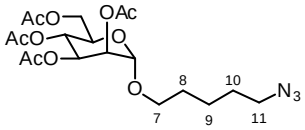
To a suspension of NaN_3 (0.95g, 14.5mmol,1.5eq) in acetonitrile (10ml) at 0°C was added dropwise Tf_2O (1.96ml, 11.6mmol, 1.2eq). The reaction mixture was stirred at this temperature for 2 hours after which **5.5** (50% weight solution in water, 2.0ml, 9.69mmol, 1eq) was added dropwise, followed by $CuSO_4$ (15mg, 0.1mmol, 0.01eq), triethylamine (3ml) and water (1ml). The reaction mixture was allowed to warm to room temperature and stirred for 16 hours. The solvent was removed *in vacuo* and the product purified by silica gel chromatography (3 ethyl acetate : 7 petrol) as a colourless liquid (920mg, 74%).

R_f 0.25 (ethyl acetate); ^{1}H NMR (400MHz, $CDCl_3$) δ ppm 1.45 (2H, m, H3), 1.55-1.71 (4H, m, H2, H4), 3.31 (2H, t, $J_{H1\text{-}H2}$ 5.9Hz, H1), 3.68 (2H, t, $J_{H4\text{-}H5}$ 6.1Hz, H5); ^{13}C NMR (101MHz, $CDCl_3$) δ ppm 23.0 (1C, C3), 28.6 (1C, C2), 32.2 (1C, C4), 51.4 (1C, C1), 62.6 (1C, C5); m/z (ES^+) 405.2 $[M + NH_4]^+$, 410.2 $[3M + Na]^+$; HRMS (FI^+) 130.0910 $[M + H]^+$ required 130.0926.

1-azido-5-pentyl-α-D-2,3,4,6-*O*-acetyl-mannopyranoside (5.6)

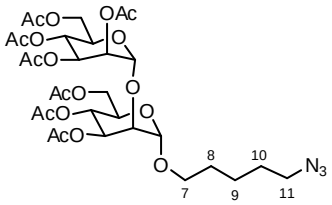
α-D-2,3,4,6-*O*-acetyl-mannopyranosyl-(1→2)-1-azido-5-pentyl-α-D-2,3,4,6-*O*-acetyl-mannopyranoside (5.7)

A solution of **5.4** (1.48g, 3.36mmol, 1eq) and **5.3** (457mg, 3.53mmol, 1.05eq) was stirred with pre-activated 4Å molecular sieves in anhydrous DCM (50ml) at -2°C for 30 minutes. To this mixture was added *N*-iodosuccinimide (907mg, 4.03mmol, 1.2eq), which had been freshly crystallized from 1,4-dioxane and dried under high vacuum for at least 24 hours, and the mixture stirred for 10 minutes. The reaction was initiated by addition of triflic acid (89µl, 1.01mmol, 0.3eq) and the mixture stirred at -2°C for 2 hours after which TLC (3 ethyl acetate : 2 petrol) showed the complete consumption of starting material. The reaction mixture was filtered through Celite® and washed with 10% $Na_2S_2O_3$ (100ml), saturated $NaHCO_3$ (100ml), brine (100ml) and dried over $MgSO_4$. The products were separated by automated silica gel chromatography (10% ethyl acetate in petrol to 100% ethyl acetate over 12 column volumes) to give the monosaccharide **5.6** (62mg, 4%) and disaccharide **5.7** (400mg, 16%). Trisaccharide and tetrasaccharide were also detected but were not isolated with high purity.



*Compound **5.6***

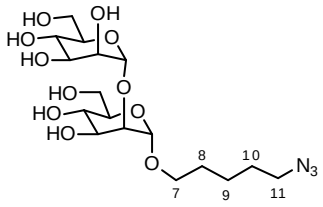
R_f 0.36 (1 petrol : 1 ethyl acetate); $[\alpha]_D^{25}$ +33.3 (c = 1.0, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 1.46 (2H, aquin, $J_{H8\text{-}H9}$ 6.8Hz, $J_{H9\text{-}H10}$ 6.8Hz, H9), 1.55-1.69 (4H, m, H8, H10), 2.00, 2.05, 2.11, 2.16 (4 x 3H, 4 x s, 4 x OAc), 3.46 (1H, dt, $J_{H7\text{-}H7'}$ 9.7Hz, $J_{H7'\text{-}H8}$ 6.3Hz, H7'), 3.67 (2H, t, $J_{H10\text{-}H11}$ 6.6Hz, H11), 3.70 (2H, dt, $J_{H7\text{-}H7'}$ 9.2Hz, $J_{H7\text{-}H8}$ 6.5Hz, H7), 3.98 (1H, ddd, $J_{H4\text{-}H5}$ 9.8Hz, $J_{H5\text{-}H6}$ 5.4Hz, $J_{H5\text{-}H6'}$ 2.5Hz, H5), 4.11 (1H, dd, $J_{H6\text{-}H6'}$ 12.1Hz, $J_{H5\text{-}H6'}$ 2.4Hz, H6'), 4.28 (1H, dd, $J_{H6\text{-}H6'}$ 12.3Hz, $J_{H5\text{-}H6}$ 5.4Hz, H6), 4.80 (1H, d, $J_{H1\text{-}H2}$ 1.6Hz, H1), 5.23 (1H, dd, $J_{H2\text{-}H3}$ 3.5Hz, $J_{H1\text{-}H2}$ 1.9Hz, H2), 5.28 (1H, t, $J_{H3\text{-}H4}$ 9.9Hz, $J_{H4\text{-}H5}$ 9.9Hz, H4), 5.34 (1H, dd, $J_{H3\text{-}H4}$ 9.9Hz, $J_{H2\text{-}H3}$ 3.5Hz, H3); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.7, 20.7, 20.7, 20.9 (4 x 1C, 4 x CO<u>C</u>H_3), 23.3 (1C, C9), 28.6 (1C, C10), 28.8 (1C, C8), 62.5 (1C, C6), 62.6 (1C, C11), 66.2 (1C, C4), 68.1 (1C, C7), 68.4 (1C, C5), 69.1 (1C, C3), 69.6 (1C, C2), 97.6 (1C, C1), 169.7, 169.9, 170.1, 170.6 (4 x 1C, 4 x C=O); FT-IR (thin film) υ 1784 (C=O), 2099 (N_3), 2924 (C-H); HRMS m/z (ES^+) 482.1743 $[M + Na]^+$ required 482.1745.



*Compound **5.7***

R_f 0.3 (1 petrol : 1 ethyl acetate); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 1.45 (2H, quin, $J_{H8\text{-}H9}$ 8.2Hz, $J_{H9\text{-}H10}$ 8.2Hz, H9), 1.63 (4H, quin, $J_{H7\text{-}H8}$ 6.9Hz, $J_{H8\text{-}H9}$ 6.9Hz, $J_{H9\text{-}H10}$ 6.9Hz, $J_{H10\text{-}H11}$ 6.9Hz, H8, H10), 2.01, 2.03, 2.04 (3 x 3H, 3 x s, 3 x OAc), 2.08 (2 x 3H, s, 2 x OAc), 2.14, 2.15 (2 x 3H, 2 x s, 2 x OAc), 3.30 (2H, t, $J_{H10\text{-}H11}$ 6.8Hz, H11), 3.44 (1H, dt, $J_{H7a\text{-}H7b}$ 9.5Hz, 2 x $J_{H7b\text{-}H8}$ 6.3Hz, H7b), 3.71 (1H, dt, $J_{H7a\text{-}H7b}$ 9.5Hz, 2 x $J_{H7a\text{-}H8}$ 6.6Hz, H7a), 3.91 (1H, ddd, $J_{H4\text{-}H5}$ 9.5Hz, $J_{H5\text{-}H6a}$ 4.4Hz, $J_{H5\text{-}H6b}$ 2.5Hz, H5), 4.02 (1H, dd, $J_{H2\text{-}H3}$ 2.8Hz, $J_{H1\text{-}H2}$ 2.2Hz, H2), 4.12 (1H, dd, $J_{H6a\text{-}H6b}$ 12.0Hz, $J_{H5\text{-}H6b}$ 2.2Hz, H6b), 4.13 (1H, dd, $J_{H6'a\text{-}H6'b}$ 12.0Hz, $J_{H5'\text{-}H6'b}$ 3.2Hz, H6'b), 4.16 (1H, ddd, $J_{H4'\text{-}H5'}$ 9.9Hz, $J_{H5'\text{-}H6'b}$ 5.4Hz, $J_{H5'\text{-}H6'a}$ 2.2Hz, H5'), 4.22 (2H, dd, $J_{H6a\text{-}H6b}$ 12.3Hz, $J_{H5\text{-}H6a}$ 5.0Hz, H6a, H6'a), 4.92 (1H, d, $J_{H1\text{-}H2}$ 1.3Hz, H1), 4.92 (1H, d, $J_{H1'\text{-}H2'}$ 1.2Hz, H1'), 5.26 (1H, dd, $J_{H2'\text{-}H3'}$ 3.5Hz, $J_{H1'\text{-}H2'}$ 1.3Hz, H2'), 5.28 (1H, t, $J_{H3'\text{-}H4'}$ 10.1Hz, $J_{H4'\text{-}H5'}$ 10.1Hz, H4'), 5.28 (1H, dd, $J_{H3\text{-}H4}$ 10.4Hz, $J_{H2\text{-}H3}$ 2.5Hz, H3), 5.33 (1H, t, $J_{H3\text{-}H4}$ 10.1Hz, $J_{H4\text{-}H5}$ 10.1Hz, H4), 5.41 (1H, dd, $J_{H3'\text{-}H4'}$ 9.9Hz, $J_{H2'\text{-}H3'}$ 3.3Hz, H3'); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.5, 20.6 (2 x 1C, 2 x CO<u>C</u>H_3), 20.7 (2C, 2 x CO<u>C</u>H_3), 20.7, 20.7, 20.9 (3 x 1C, 3 x CO<u>C</u>H_3), 23.4 (1C, C9), 28.6 (1C, C8), 28.9 (1C, C10), 51.2 (1C, C11), 62.1 (1C, C6'), 62.2 (1C, C6), 66.2 (1C, C4), 66.4 (1C,

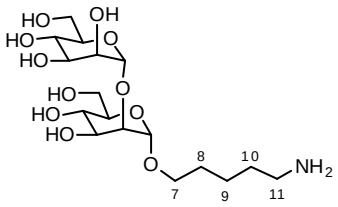
C4'), 68.1 (1C, C7), 68.4 (1C, C3'), 68.5 (1C, C5), 69.1 (1C, C5'), 69.7 (1C, C2'), 70.3 (1C, C3), 77.1 (1C, C2), 98.3 (1C, C1'), 99.2 (1C, C1), 169.4, 169.4, 169.7, 169.8, 170.4, 170.4, 170.9 (7 x 1C, 7 x C=O); FT-IR (ATR) υ 1743 (C=O), 2098 (N_3); HRMS (ES^+) m/z 770.2597 $[M + Na]^+$ (required 770.2596).



α-D-mannopyranosyl-(1→2)-1-azido-5-pentyl-α-D-2,3,4,6-*O*-acetyl-mannopyranoside (5.8)

To a solution of **5.7** (0.40g, 0.54mmol, 1eq) in anhydrous methanol (20ml) was added NaOMe (29mg, 0.54mmol, 1eq). The reaction mixture was stirred at room temperature for 12 hours, after which TLC (1 petrol : 2ethyl acetate) showed complete consumption of starting material. The solvents were removed under reduced pressure and the product purified by automated silica gel chromatography (1 water : 3 isopropanol : 8 ethyl acetate) to afford the title compound as a white amorphous solid (180mg, 73%)

R_f 0.25 (1 water : 3 isopropanol : 8 ethyl acetate), ^{1}H NMR (500MHz, D_2O) δ ppm 1.37 (2H, quin, $J_{H8\text{-}H9}$ 6.6Hz, $J_{H9\text{-}H10}$ 6.6Hz, H9), 1.55 (2H, quin, $J_{H7\text{-}H8}$ 6.9Hz, $J_{H8\text{-}H9}$ 6.9Hz, $J_{H9\text{-}H10}$ 6.9Hz, $J_{H10\text{-}H11}$ 6.9Hz, H8, H10), 3.26 (2H, t, $J_{H10\text{-}H11}$ 6.8Hz, H11), 3.47 (1H, dt, $J_{H7a\text{-}H7b}$ 10.0Hz, 2 x $J_{H7b\text{-}H8}$ 6.0Hz, H7b), 3.53 (1H, t, $J_{H3\text{-}H4}$ 9.8Hz, $J_{H4\text{-}H5}$ 9.8Hz, H4), 3.54 (1H, ddd, $J_{H4\text{-}H5}$ 9.5Hz, $J_{H5\text{-}H6b}$ 6.0Hz, $J_{H5\text{-}H6a}$ 1.9Hz, H5), 3.60 (1H, t, $J_{H3'\text{-}H4'}$ 9.8Hz, $J_{H4'\text{-}H5'}$ 9.8Hz, H4'), 3.64 (1H, dd, $J_{H6a\text{-}H6b}$ 12.0Hz, $J_{H5\text{-}H6b}$ 6.6Hz, H6b), 3.66-3.71 (2H, m, H5, H7a), 3.68 (2H, dd, $J_{H6a\text{-}H6b}$ 12.1Hz, $J_{H5\text{-}H6a}$ 5.8Hz, H6a, H6'a), 3.76 (1H, dd, $J_{H3'\text{-}H4'}$ 9.8Hz, $J_{H2'\text{-}H3'}$ 3.2Hz, H3'), 3.80 (1H, dd, $J_{H6a\text{-}H6b}$ 12.0Hz, $J_{H5\text{-}H6a}$ 2.2Hz, H6a), 3.81 (1H, dd, $J_{H6'a\text{-}H6'b}$ 12.0Hz, $J_{H5'\text{-}H6'a}$ 2.3Hz, H6'a), 3.82 (1H, dd, $J_{H3\text{-}H4}$ 9.1Hz, $J_{H2\text{-}H3}$ 3.8Hz, H3), 3.87 (1H, dd, $J_{H2\text{-}H3}$ 3.3Hz, $J_{H1\text{-}H2}$ 1.7Hz, H2), 3.99 (1H, dd, $J_{H2'\text{-}H3'}$ 3.2Hz, $J_{H1'\text{-}H2'}$ 1.9Hz, H2'), 4.94 (1H, d, $J_{H1'\text{-}H2'}$ 1.6Hz, H1'), 5.02 (1H, d, $J_{H1\text{-}H2}$ 1.6Hz, H1); ^{13}C NMR (126MHz, D_2O) δ ppm 22.7 (1C, C9), 27.7 (1C, C10), 28.0 (1C, C8), 51.0 (1C, C11), 60.9 (1C, C6'), 61.1 (1C, C6), 66.9 (1C, C4'), 66.9 (1C, C4), 67.7 (1C, C7), 69.9 (1C, C2'), 70.3 (2C, C3, C3'), 72.7 (1C, C5), 73.3 (1C, C5'), 78.7 (1C, C2), 98.1 (1C, C1), 102.3 (1C, C1'); FT-IR (ATR) υ 2115 (N_3), 3156 (OH); HRMS (ES^-) m/z 452.1889 $[M - H]^-$ (required 452.1886).

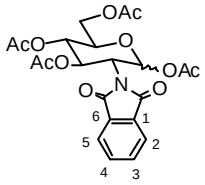


α-D-mannopyranosyl-(1→2)-1-amino-5-pentyl-α-D-2,3,4,6-*O*-acetyl-mannopyranoside (5.1)[40]

A suspension of 5% Pd/C (85mg, 0.040mmol, 0.1eq) and **5.8** (180mg, 0.40mmol, 1eq) in methanol (10ml) and water (2ml) was repeatedly degassed under vacuum and charged with hydrogen. The reaction was vigorously stirred under an atmosphere of hydrogen for 28 hours, after which TLC (1 water : 3 isopropanol : 8 ethyl acetate) showed complete consumption of the starting material. The reaction mixture was filtered through Celite® and the solvent removed *in vacuo*. The product was dissolved in water and lyophilized to give the target compound as an amorphous white solid (167mg, 98%).

R_f 0.4 (1 water : 3 NH_4OH : 5 EtOH), $[\alpha]_D^{25}$ +45.1 (c = 1.0, H_2O), lit.[40] +48.3 (c = 1.0, H_2O); ^{1}H NMR (500MHz, D_2O) δ ppm 1.32 (2H, quin, $J_{H8\text{-}H9}$ 6.9Hz, $J_{H9\text{-}H10}$ 6.9Hz, H9), 1.52 (2H, quin, $J_{H7\text{-}H8}$ 6.3Hz, $J_{H8\text{-}H9}$ 6.3Hz, H8), 1.55 (2H, quin, $J_{H9\text{-}H10}$ 7.9Hz, $J_{H10\text{-}H11}$ 7.9Hz, H10), 2.85 (2H, t, $J_{H10\text{-}H11}$ 7.7Hz, H11), 3.41 (1H, dt, $J_{H7a\text{-}H7b}$ 9.9Hz, 2 x $J_{H7b\text{-}H8}$ 6.0Hz, H7b), 3.46 (1H, ddd, $J_{H4\text{-}H5}$ 9.7Hz , $J_{H5\text{-}H6b}$ 5.8Hz, $J_{H5\text{-}H6a}$ 2.6Hz, H5), 3.47 (1H, t, $J_{H3\text{-}H4}$ 9.7Hz, $J_{H4\text{-}H5}$ 9.7Hz, H4), 3.55 (1H, t, $J_{H3'\text{-}H4'}$ 9.4Hz, $J_{H4'\text{-}H5'}$ 9.4Hz, H4'), 3.58-3.67 (2H, m, H5', H7a), 3.62 (2H, dd, $J_{H6a\text{-}H6b}$ 12.3Hz, $J_{H5\text{-}H6b}$ 6.0Hz, H6b, H6'b), 3.71 (1H, dd, $J_{H3'\text{-}H4'}$ 9.6Hz, $J_{H2'\text{-}H3'}$ 3.3Hz, H3'), 3.76 (2H, dd, $J_{H6a\text{-}H6b}$ 12.8Hz, $J_{H5\text{-}H6a}$ 2.6Hz, H6a), 3.76 (1H, dd, $J_{H3\text{-}H4}$ 9.4Hz, $J_{H2\text{-}H3}$ 3.2Hz, H3), 3.82 (1H, dd, $J_{H2\text{-}H3}$ 3.2Hz, $J_{H1\text{-}H2}$ 1.6Hz, H2), 3.94 (1H, dd, $J_{H2'\text{-}H3'}$ 3.2Hz, $J_{H1'\text{-}H2'}$ 1.9Hz, H2'), 4.89 (1H, d, $J_{H1'\text{-}H2'}$ 1.4Hz, H1'), 4.97 (1H, d, $J_{H1\text{-}H2}$ 1.2Hz, H1); ^{13}C NMR (101MHz, D_2O) δ ppm 22.8 (1C, C9), 27.1 (1C, C10), 28.3 (1C, C8), 39.8 (1C, C11), 61.3 (1C, C6), 61.5 (1C, C6'), 67.3 (1C, C4), 67.3 (1C, C4'), 67.9 (1C, C7), 70.3 (1C, C2'), 70.6 (1C, C3'), 70.7 (1C, C3), 73.1 (1C, C5), 73.6 (1C, C5'), 79.1 (1C, C2), 98.4 (1C, C1), 102.7 (1C, C1'); FT-IR (ATR) υ 2894 (NH_2 bend overtone), 3348 (OH); m/z (ES^+) 428.2 $[M + H]^+$.

5.11.2 Attempted synthesis of GalNAc-4-SO_4 ligand

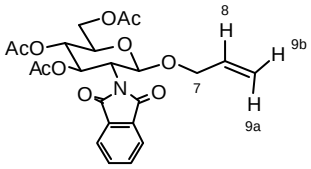


_N_-phthalamido-1,3,4,6-O-acetyl-D-glucose (5.19)

Compound **5.17** (20.0g, 92.7mmol, 1eq) and NaOH (4.1g, 102mmol, 1.1eq) were dissolved in water (100ml) and allowed to stir at room temperature for 10 minutes, after which phthalic anhydride (14.0g, 94.6mmol, 1.02mmol) was added. The reaction mixture was stirred at room temperature for 5 hours, during which time a pale yellow precipitate had formed. The reaction mixture was lyophilized and re-dissolved in anhydrous pyridine (50ml) and cooled to -78°C. Acetic anhydride (70ml, 742mmol, 8eq) was added dropwise over 10 minutes. The reaction was allowed to warm to room temperature and stirred for 16 hours after which TLC (2 petrol : 1 ethyl acetate) showed the formation of a single product. The reaction was quenched by pouring onto ice (~200ml) and the mixture was extracted with ethyl acetate (2 x 150ml). The combined organics were washed with 2M HCl (5 x 100ml), saturated $NaHCO_3$ (2 x 100ml), 1M NaOH (50ml), brine (2 x 100ml) and dried over $MgSO_4$. The solvent was removed *in vavuo* to yield the desired product as a pale yellow solid as a mixture of anomers (32.1g, 72%, α:β 1:1.05).

R_f 0.1 (2 petrol : 1 ethyl acetate), $[\alpha]_D^{25}$ +75.9 (c = 2.0, $CHCl_3$); ^{1}H NMR (400MHz, $CDCl_3$) δ ppm 1.85, 1.86, 1.98, 2.03, 2.05, 2.07, 2.10, 2.11 (8 x 3H, 8 x s, 8 x OAc (4 x α, 4 x β)), 4.02 (1H, ddd, $J_{H4\beta\text{-}H5\beta}$ 10.1Hz, $J_{H5\beta\text{-}H6a\beta}$ 4.0Hz, $J_{H5\beta\text{-}H6b\beta}$ 1.8Hz, H5β), 4.13 (2H, dd, $J_{H6a\text{-}H6b}$ 12.4Hz, $J_{H5\text{-}H6b}$ 1.5Hz, H6bα, H6bβ), 4.31 (1H, dt, $J_{H4\alpha\text{-}H5\alpha}$ 11.9Hz, $J_{H5\alpha\text{-}H6a\alpha}$ 3.4Hz, $J_{H5\alpha\text{-}H6b\alpha}$ 3.4Hz, H5α), 4.36 (2H, dd, $J_{H6a\text{-}H6b}$ 12.3Hz, $J_{H5\text{-}H6a}$ 4.3Hz, H6aα, H6bβ), 4.46 (1H, dd, $J_{H2\beta\text{-}H3\beta}$ 10.4Hz, $J_{H1\beta\text{-}H2\beta}$ 9.1Hz, H2β), 4.71 (1H, dd, $J_{H2\alpha\text{-}H3\alpha}$ 11.6Hz, $J_{H1\alpha\text{-}H2\alpha}$ 3.3Hz, H2α), 5.15 (1H, t, $J_{H3\alpha\text{-}H4\alpha}$ 9.3Hz, $J_{H4\alpha\text{-}H5\alpha}$ 9.3Hz, H4α), 5.20 (1H, t, $J_{H3\beta\text{-}H4\beta}$ 9.8Hz, $J_{H4\beta\text{-}H5\beta}$ 9.8Hz, H4β), 5.87 (1H, t, $J_{H2\beta\text{-}H3\beta}$ 9.9Hz, $J_{H3\beta\text{-}H4\beta}$ 9.9Hz, H3β), 6.27 (1H, d, $J_{H1\alpha\text{-}H2\alpha}$ 3.3Hz, H1α), 6.50 (1H, d, $J_{H1\beta\text{-}H2\beta}$ 8.8Hz, H1β), 6.55 (1H, dd, $J_{H2\alpha\text{-}H3\alpha}$ 11.4Hz, $J_{H3\alpha\text{-}H4\alpha}$ 9.3Hz, H3α), 7.71-7.79 (4H, m, ArH_3, ArH_4 (2 x α, 2 x β)), 7.81-7.88 (4H, m, ArH_2, ArH_5 (2 x α, 2 x β)); ^{13}C NMR (101MHz, $CDCl_3$) δ ppm 20.4, 20.6, 20.7, 21.0 (4 x 1C, 4 x CO$\underline{C}H_3$), 20.6, 20.7 (2 x 2C, 4 x CO$\underline{C}H_3$), 52.8 (1C, C2α), 53.5 (1C, C2β), 61.5, 61.5 (2C, C6α, C6β), 67.0 (1C, C3α), 68.3 (1C, C4β), 69.3 (1C, C4α), 70.1 (1C, C5α), 70.5 (1C, C3β), 72.6 (1C, C5β), 89.7 (1C, C1β), 90.5 (1C, C1α), 123.7, 123.8 (2 x 2C, ArC_2, ArC_5), 131.1, 131.2 (2 x 2C, ArC_1, ArC_6), 134.4, 134.5 (2 x 2C, ArC_3, ArC_4), 167.4, 167.4 (2 x 1C,

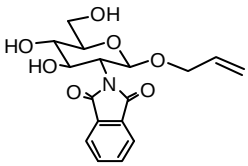
2 x NCO), 168.6, 169.3, 169.4, 169.5, 169.7, 170.0 (6 x 1C, 6 x $\underline{C}OCH_3$), 170.6 (2C, 2 x $\underline{C}OCH_3$); m/z (ES^+) 500.1 $[M + Na]^+$, 977.1 $[2M + Na]^+$.



β-*O*-allyl-*N*-phthalamido-3,4,6-*O*-acetyl-D-glucose[41] (5.20)

To a solution of **5.19** (48.6g, 102mmol, 1eq) in anhydrous DCM (100ml) at room temperature was added dropwise over 20 minutes $SnCl_4$ (23.9ml, 204mmol, 2eq). After 30 minutes of stirring at room temperature, allyl alcohol (10.4ml, 153mmol, 1.5eq) was added dropwise over 10 minutes. After stirring at this temperature for 33 hours, TLC (1 petrol : 1 ethyl acetate) showed complete consumption of starting material. The reaction was quenched by pouring onto ice (400ml) and the mixture was extracted with DCM (500ml). The organic layer was washed with saturated $NaHCO_3$ (2 x 200ml), brine (200ml) and dried over $MgSO_4$. The product was purified by silica gel chromatography (1 petrol : 1 ethyl acetate) as a white amorphous solid (28.7g, 59%, contaminated with a small amount of the α anomer).

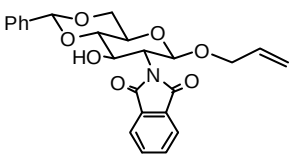
R_f 0.39 (1 petrol : 1 ethyl acetate), $[\alpha]_D^{25}$ +52.0 (c = 1.0, $CHCl_3$); ^1H NMR (500MHz, $CDCl_3$) δ ppm 1.86, 2.03, 2.12 (3 x 3H, 3 x s, 3 x OAc), 3.87 (1H, ddd, $J_{H4\text{-}H5}$ 10.1Hz, $J_{H5\text{-}H6a}$ 4.7Hz, $J_{H5\text{-}H6b}$ 2.5Hz, H5), 4.06 (1H, ddt, $J_{H7a\text{-}H7b}$ 12.9Hz, $J_{H7b\text{-}H8}$ 6.6Hz, $J_{H7b\text{-}H9a}$ 1.1Hz, $J_{H7b\text{-}H9b}$ 1.1Hz, H7b), 4.18 (1H, dd, $J_{H6a\text{-}H6b}$ 12.1Hz, $J_{H5\text{-}H6b}$ 2.4Hz, H6b), 4.28 (1H, ddt, $J_{H7a\text{-}H7b}$ 12.9Hz, $J_{H7a\text{-}H8}$ 5.0Hz, $J_{H7a\text{-}H9a}$ 1.4Hz, $J_{H7a\text{-}H9b}$ 1.4Hz, H7a), 4.33 (1H, dd, $J_{H6a\text{-}H6b}$ 11.5Hz, $J_{H5\text{-}H6a}$ 5.5Hz, H6a), 4.35 (1H, dd, $J_{H2\text{-}H3}$ 10.7Hz, $J_{H3\text{-}H4}$ 8.5Hz, H2), 5.07 (3H, dd, $J_{H8\text{-}H9b}$ 10.6Hz, $J_{H9a\text{-}H9b}$ 1.4Hz, H9b), 5.14 (1H, dd, $J_{H8\text{-}H9a}$ 17.3Hz, $J_{H9a\text{-}H9b}$ 1.6Hz, H9a), 5.18 (1H, dd, $J_{H4\text{-}H5}$ 10.1Hz, $J_{H3\text{-}H4}$ 9.1Hz, H4), 5.41 (1H, d, $J_{H1\text{-}H2}$ 8.5Hz, H1), 5.71 (1H, dddd, $J_{H8\text{-}H9a}$ 17.4Hz, $J_{H8\text{-}H9b}$ 10.4Hz, $J_{H7b\text{-}H8}$ 6.3Hz, $J_{H7a\text{-}H8}$ 5.4Hz, H8), 5.80 (1H, dd, $J_{H2\text{-}H3}$ 10.7Hz, $J_{H3\text{-}H4}$ 9.1Hz, H3), 7.74 (2H, dd, ^3J 5.5Hz, ^4J 3.0Hz, ArH_3, ArH_4), 7.86 (2H, dd, ^3J 5.5Hz, ^4J 3.0Hz, ArH_2, ArH_5); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.4, 20.6, 20.8 (3 x 1C, 3 x $CO\underline{C}H_3$), 54.6 (1C, C2), 62.0 (1C, C6), 69.0 (1C, C4), 70.2 (1C, C7), 70.8 (1C, C3), 71.8 (1C, C5), 97.1 (1C, C1), 117.9 (1C, C9), 123.6 (2C, ArC_2, ArC_5), 131.4 (2C, ArC_1, ArC_6), 133.2 (1C, C8), 134.3 (2C, ArC_3, ArC_4), 167.6 (2C, NCO), 169.5, 170.2, 170.7 (3 x 1C, 3 x $\underline{C}OCH_3$); FT-IR (KBr disc) υ 1227 (O-C), 1718 (N-C=O), 1749 (O-C=O); m/z (ES^+) 498.1 $[M + Na]^+$, 536.2 $[M + {}^i PrOH + H]^+$, 973.3 $[2M + Na]^+$.



β-*O*-allyl-*N*-phthalamido-D-glucose[42] (5.21)

Compound **5.20** (28.6g, 60.2mmol, 1eq) and NaOMe (325mg, 6.02mmol, 0.1eq) were dissolved in anhydrous methanol (75ml) and stirred at room temperature for 7 hours, after which TLC (ethyl acetate) showed complete consumption of starting materials. The reaction was quenched with Dowex 50WX8 (H^+ form) until neutral (pH paper). The solvent was removed under reduced pressure to give the desired compound as a pale yellow amorphous solid (21.0g, 100%).

R_f 0.28 (ethyl acetate), $[\alpha]_D^{25}$ 12.2 (c = 1.0, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 3.43 (1H, dt, $J_{H4\text{-}H5}$ 9.5Hz, $J_{H5\text{-}H6a}$ 2.8Hz, $J_{H5\text{-}H6b}$ 2.8Hz, H5), 3.71 (1H, t, $J_{H3\text{-}H4}$ 9.1Hz, $J_{H4\text{-}H5}$ 9.1Hz, H4), 3.84-3.89 (2H, m, H6a, H6b), 4.01 (1H, ddt, $J_{H7a\text{-}H7b}$ 13.2Hz, $J_{H7b\text{-}H8}$ 6.0Hz, $J_{H7b\text{-}H9a}$ 1.2Hz, $J_{H7b\text{-}H9b}$ 1.2Hz, H7b), 4.13 (1H, dd, $J_{H2\text{-}H3}$ 10.7Hz, $J_{H1\text{-}H2}$ 8.5Hz, H2), 4.24 (1H, ddt, $J_{H7a\text{-}H7b}$ 13.2Hz, $J_{H7a\text{-}H8}$ 5.0Hz, $J_{H7a\text{-}H9a}$ 1.3Hz, $J_{H7a\text{-}H9b}$ 1.3Hz, H7a), 4.30 (1H, dd, $J_{H2\text{-}H3}$ 10.6Hz, $J_{H3\text{-}H4}$ 9.0Hz, H3), 5.01 (1H, dd, $J_{H8\text{-}H9b}$ 10.6Hz, $J_{H9a\text{-}H9b}$ 1.4Hz, H9b), 5.09 (1H, dd, $J_{H8\text{-}H9a}$ 17.0Hz, $J_{H9a\text{-}H9b}$ 1.6Hz, H9a), 5.23 (1H, d, $J_{H1\text{-}H2}$ 8.5Hz, H1), 5.68 (1H, dddd, $J_{H8\text{-}H9a}$ 17.0Hz, $J_{H8\text{-}H9b}$ 10.7Hz, $J_{H7b\text{-}H8}$ 6.0Hz, $J_{H7a\text{-}H8}$ 5.0 Hz, H8), 7.69 (2H, dd, ^3J 5.5Hz, ^4J 3.0Hz, ArH_3, ArH_4), 7.80 (2H, dd, ^2J 5.5Hz, ^3J 3.0Hz, ArH_2, ArH_5); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 56.6 (1C, C2), 61.6 (1C, C6), 70.0 (1C, C7), 71.2 (1C, C4), 71.5 (1C, C3), 75.5 (1C, C5), 97.5 (1C, C1), 117.4 (1C, C9), 123.3 (2C, ArC_2, ArC_5), 131.7 (2C, ArC_1, ArC_6), 133.6 (1C, C9), 134.0 (2C, ArC_3, ArC_4), 168.4 (2C, C=O); FT-IR (KBr disc) υ 1705 (C=O); m/z (ES^+) 372.1 $[M + Na]^+$, 404.1 $[M + MeOH + Na]^+$, 721.2 $[2M + Na]^+$, 753.3 $[2M + MeOH + Na]^+$.

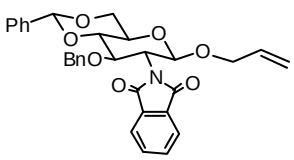


β-*O*-allyl-*N*-phthalamido-4,6-benzylidene-D-glucose[43] (5.22)

A solution of **5.21** (21.7g, 62.1mmol, 1eq), benzyl dimethylacetal (9.41ml, 62.7mmol, 1.01eq) and *p*-toluene sulfonic acid monohydrate (0.59g, 3.10mmol, 0.05eq) in acetonitrile (75ml) was heated to 50°C. After 28 hours, TLC (ethyl acetate) showed the formation of a product but with incomplete consumption of starting material. Extra benzyl dimethylacetal (9.41ml, 62.7mmol, 1.01eq) was added and the reaction mixture heated under reflux for 56 hours after which time TLC (ethyl acetate) showed complete consumption of starting material. The solvent was removed *in vacuo* and the crude residue partitioned between in ethyl

acetate (300ml) and water (200ml). The organic layer was washed with saturated $NaHCO_3$ (200ml), brine (200ml), dried over $MgSO_4$ and concentrated under reduced pressure. The desired compound was purified by silica gel chromatography (7% to 60% ethyl acetate in petrol over 10 column volumes) as a pale yellow foam (16.45g, 61%).

R_f 0.3 (7 petrol : 3 ethyl acetate), $[\alpha]_D^{25}$ -35.1 (c = 1.0, $CHCl_3$), lit.[43] $[\alpha]_D$ -39 (c = 1.0, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.54 (1H, br. s, OH), 3.63 (1H, t, $J_{H3\text{-}H4}$ 9.1Hz, $J_{H4\text{-}H5}$ 9.1Hz, H4), 3.66 (1H, td, $J_{H4\text{-}H5}$ 9.5Hz, $J_{H5\text{-}H6b}$ 9.5Hz, $J_{H5\text{-}H6a}$ 4.7Hz, H5), 3.85 (1H, t, $J_{H5\text{-}H6b}$ 9.8Hz, $J_{H6a\text{-}H6b}$ 9.8Hz, H6b), 4.05 (1H, ddt, $J_{H7a\text{-}H7b}$ 13.0Hz, $J_{H7b\text{-}H8}$ 6.3Hz, $J_{H7b\text{-}H9a}$ 1.4Hz, $J_{H7b\text{-}H9b}$ 1.4Hz, H7b), 4.29 (1H, m, $J_{H7a\text{-}H8}$ 7.9Hz, $J_{H7a\text{-}H9a}$ 1.6Hz, $J_{H7a\text{-}H9b}$ 1.6Hz, H7a), 4.29 (1H, dd, $J_{H2\text{-}H3}$ 10.7Hz, $J_{H1\text{-}H2}$ 8.5Hz, H2), 4.40 (1H, dd, $J_{H6a\text{-}H6b}$ 10.6Hz, $J_{H5\text{-}H6a}$ 4.6Hz, H6a), 4.66 (1H, dd, $J_{H2\text{-}H3}$ 10.6Hz, $J_{H3\text{-}H4}$ 8.7Hz, H3), 5.06 (1H, ddt, $J_{H8\text{-}H9b}$ 10.4Hz, $J_{H9a\text{-}H9b}$ 1.6Hz, $J_{H7a\text{-}H9b}$ 1.3Hz, $J_{H7b\text{-}H9b}$ 1.3Hz, H9b), 5.15 (1H, dq, $J_{H8\text{-}H9a}$ 17.0Hz, $J_{H7a\text{-}H9a}$ 1.6Hz, $J_{H7b\text{-}H9a}$ 1.6Hz, $J_{H9a\text{-}H9b}$ 1.6Hz, H9a), 5.32 (1H, d, $J_{H1\text{-}H2}$ 8.5Hz, H1), 5.58 (1H, s, PhC<u>H</u>), 5.70 (1H, dddd, $J_{H8\text{-}H9a}$ 17.0Hz, $J_{H8\text{-}H9b}$ 10.4Hz, $J_{H7a\text{-}H8}$ 6.3Hz, $J_{H7b\text{-}H8}$ 5.0Hz, H8), 7.38-7.39 (3H, m, Bd ArH_{ortho}, Bd ArH_{para}), 7.50-7.52 (2H, m, Bd ArH_{meta}), 7.74 (2H, dd, ^3J 5.7Hz, ^4J 3.2Hz, ArH_3, ArH_4), 7.84-7.89 (2H, m, ArH_2, ArH_5); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 56.5 (1C, C2), 60.4 (1C, C5), 66.1 (1C, C3), 68.6 (1C, C6), 70.1 (1C, C7), 82.2 (1C, C4), 97.9 (1C, C1), 102.0 (1C, Ph<u>C</u>H), 117.7 (1C, C9), 123.5 (2C, ArC_2, ArC_5), 126.3 (2C, Bd ArC_{meta}), 128.4 (3C, Bd ArC_{ortho}, Bd ArC_{para}), 131.7 (2C, ArC_1, ArC_6), 133.3 (1C, C8), 134.1 (1C, ArC_3, ArC_4), 136.9 (1C, Bd ArC_{ipso}), 168.1 (2C, NCO); m/z (ES^+) 460.1 $[M + Na]^+$, 492.2 $[M + MeOH + Na]^+$, 897.3 $[2M + Na]^+$, 929.3 $[2M + MeOH + Na]^+$.

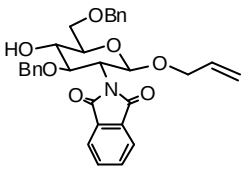


β-*O*-allyl-*N*-phthalamido-3-benzyl-4,6-benzylidene-D-glucose[44] (5.23)

To a solution of **5.22** (16.45g, 37.6mmol, 1eq) in anhydrous DMF (100ml) cooled to 0°C was added NaH (60% dispersion in mineral oil, 1.66g, 41.4mmol, 1.1eq) portionwise over 5 minutes. The reaction mixture was stirred at this temperature for 5 minutes, after which benzyl bromide (4.93ml, 41.4mmol, 1.1eq) was added dropwise. The reaction mixture was allowed to warm to room temperature and stirred for 12 hours after which TLC (7 petrol : 3 ethyl acetate) showed complete consumption of starting materials. The reaction mixture was quenched with isopropanol (10ml) and the solvent removed *in vacuo*. The crude residue was

partitioned between ethyl acetate (500ml) and water (250ml). The organic layer was washed with brine (3 x 250ml), dried over $MgSO_4$ and concentrated *in vacuo*. Silica gel chromatography (5% to 50% ethyl acetate in petrol over 10 column volumes) afforded the desired compound as a white foam (15.90g, 80%).

R_f 0.47 (7 petrol : 3 ethyl acetate), $[\alpha]_D^{25}$ +44.0 (c = 0.45, $CHCl_3$), lit.[44] $[\alpha]_D^{20}$ -12.8 (c = 1.2, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 3.65 (1H, td, $J_{H4\text{-}H5}$ 9.6Hz, $J_{H5\text{-}H6b}$ 9.6Hz, $J_{H5\text{-}H6a}$ 5.0Hz, H5), 3.83 (1H, t, $J_{H3\text{-}H4}$ 9.1Hz, $J_{H4\text{-}H5}$ 9.1Hz, H4), 3.88 (1H, t, $J_{H5\text{-}H6b}$ 10.2Hz, $J_{H6a\text{-}H6b}$ 10.2Hz, H6b), 4.00 (1H, ddt, $J_{H7a\text{-}H7b}$ 12.9Hz, $J_{H7b\text{-}H8}$ 6.3Hz, $J_{H7b\text{-}H9a}$ 1.3Hz, $J_{H7b\text{-}H9a}$ 1.3Hz, H7b), 4.25 (1H, ddt, $J_{H7a\text{-}H7b}$ 11.5Hz, $J_{H7a\text{-}H8}$ 6.3Hz, $J_{H7a\text{-}H9a}$ 1.6Hz, $J_{H7a\text{-}H9b}$ 1.6Hz, H7a), 4.26 (1H, dd, $J_{H2\text{-}H3}$ 10.4Hz, $J_{H1\text{-}H2}$ 8.5Hz, H2), 4.41 (1H, dd, $J_{H6a\text{-}H6b}$ 10.4Hz, $J_{H5\text{-}H6a}$ 5.0Hz, H6a), 4.44 (1H, dd, $J_{H2\text{-}H3}$ 10.4Hz, $J_{H3\text{-}H4}$ 9.1Hz, H3), 4.51 (1H, d, ^2J 12.3Hz, OC$\underline{H}_2$Ph), 4.80 (1H, d, ^2J 12.3Hz, OC$\underline{H}_2$Ph), 5.02 (1H, dq, $J_{H8\text{-}H9b}$ 10.4Hz, $J_{H7a\text{-}H9b}$ 1.3Hz, $J_{H7b\text{-}H9b}$ 1.3Hz, $J_{H8\text{-}H9b}$ 1.3Hz, H9b), 5.11 (1H, dq, $J_{H8\text{-}H9a}$ 17.0Hz, $J_{H9a\text{-}H9b}$ 1.6Hz, $J_{H7a\text{-}H9a}$ 1.6Hz, $J_{H7b\text{-}H9a}$ 1.6Hz, H9a), 5.25 (1H, d, $J_{H1\text{-}H2}$ 8.5Hz, H1), 5.64 (1H, s, PhC$\underline{H}$), 5.66 (1H, dddd, $J_{H8\text{-}H9a}$ 17.0Hz, $J_{H8\text{-}H9b}$ 10.4Hz, $J_{H7a\text{-}H8}$ 6.3Hz, $J_{H7b\text{-}H8}$ 5.0Hz, H8), 6.87 - 6.95 (3H, m, Bn ArH_{meta}, Bn ArH_{para}), 7.00 (2H, dd, $J_{ortho\text{-}meta}$ 6.9Hz, $J_{ortho\text{-}para}$ 1.3Hz, Bn ArH_{para}), 7.36-7.44 (3H, m, Bd ArH_{ortho}, Bd ArH_{para}), 7.51-7.56 (2H, m, Bd ArH_{meta}), 7.64-7.84 (4H, m, NPhth); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 55.8 (1C, C2), 66.0 (1C, C5), 68.8 (1C, C6), 70.1 (1C, C7), 74.1 (1C, Ph$\underline{C}H_2$), 74.5 (1C, C3), 83.1 (1C, C4), 98.0 (1C, C1), 101.3 (1C, Ph$\underline{C}$H), 117.6 (1C, C9), 123.3 (2C, ArC_2, ArC_5), 126.0 (2C, Bd ArC_{meta}), 127.4, 128.0, 128.0, 128.3, 129.0 (10C, Bd ArC, Bn ArC), 131.6 (2C, ArC_1, ArC_6), 133.3 (1C, C8), 133.8 (2C, ArC_3, ArC_4), 137.3, 137.9 (2x 1C, Bd ArC_{ipso}, Bn ArC_{ipso}), 167.6, 167.9 (2 x 1C, NCO); m/z (ES^+) 549.7 $[M + Na]^+$.

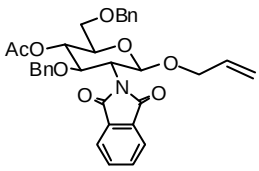


β-*O*-allyl-*N*-phthalamido-3,6-dibenzyl-D-glucose[33] (5.24)

To a solution of **5.23** (15.9g, 30.2mmol, 1eq) in anhydrous DCM (100ml) cooled to -78°C was added Et_3SiH (7.23ml, 45.3mmol, 1.5eq) followed by MsOH (2.35ml, 36.2mmol, 1.2eq). The reaction mixture was maintained at -78°C for 8 hours and then allowed to warm to room temperature over a further 11 hours, after which TLC (8 petrol : 2 ethyl acetate) showed complete cnsumption of starting materials. The reaction mixture was diluted with DCM (200ml) and washed with water (200ml), saturated $NaHCO_3$ (200ml), brine (200ml), dried over $MgSO_4$ and concentrated *in vacuo*. The products were

separated by silica gel chromatography (7% to 60% ethyl acetate in petrol over 9 column volumes) to afford the title compound (6.89g, 43%) as a white foam.

R_f 0.3 (7 petrol : 3 ethyl acetate), $[\alpha]_D^{25}$ +28.7 (c = 1.0, $CHCl_3$), lit.[33] $[\alpha]_D$ +31.4 (c = 1.0, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 2.94 (1H, br. s, 4-OH), 3.65 (1H, dt, $J_{H4\text{-}H5}$ 9.5Hz, $J_{H5\text{-}H6a}$ 5.0Hz, $J_{H5\text{-}H6b}$ 5.0Hz, H5), 3.80 (1H, dd, $J_{H6a\text{-}H6b}$ 10.4Hz, $J_{H5\text{-}H6b}$ 5.4Hz, H6b), 3.83 (1H, t, $J_{H3\text{-}H4}$ 9.0Hz, $J_{H4\text{-}H5}$ 9.0Hz, H4), 3.85 (1H, dd, $J_{H6a\text{-}H6b}$ 10.1Hz, $J_{H5\text{-}H6a}$ 5.0Hz, H6a), 3.99 (1H, ddt, $J_{H7a\text{-}H7b}$ 13.2Hz, $J_{H7b\text{-}H8}$ 6.0Hz, $J_{H7b\text{-}H9a}$ 0.9Hz, $J_{H7b\text{-}H9b}$ 0.9Hz, H7b), 4.23 (1H, ddt, $J_{H7a\text{-}H7b}$ 13.2Hz, $J_{H7a\text{-}H8}$ 4.7Hz, $J_{H7a\text{-}H9a}$ 0.9Hz, $J_{H7a\text{-}H9b}$ 0.9Hz, H7a), 4.20 (1H, dd, $J_{H2\text{-}H3}$ 10.7Hz, $J_{H1\text{-}H2}$ 8.2Hz, H2), 4.26 (1H, dd, $J_{H2\text{-}H3}$ 10.7Hz, $J_{H3\text{-}H4}$ 8.2Hz, H3), 4.55 (1H, d, ^2J 12.3Hz, OC$\underline{H}_2$Ph ^{1}), 4.60 (1H, d, ^2J 12.0Hz, OC$\underline{H}_2$Ph ^{2}), 4.66 (1H, d, ^2J 12.0Hz, OC$\underline{H}_2$Ph ^{2}), 4.75 (1H, d, ^2J 12.3Hz, OC$\underline{H}_2$Ph ^{1}), 5.01 (1H, dq, $J_{H8\text{-}H9b}$ 10.4Hz, $J_{H7a\text{-}H9b}$ 1.3Hz, $J_{H7b\text{-}H9b}$ 1.3Hz, $J_{H9a\text{-}H9b}$ 1.3Hz, H9b), 5.09 (1H, dq, $J_{H8\text{-}H9a}$ 17.3Hz, $J_{H7a\text{-}H9a}$ 1.6Hz, $J_{H7b\text{-}H9a}$ 1.6Hz, $J_{H9a\text{-}H9b}$ 1.6Hz, H9a), 5.19 (1H, d, $J_{H1\text{-}H2}$ 7.9Hz, H1), 5.68 (1H, dddd, $J_{H8\text{-}H9a}$ 17.3Hz, $J_{H8\text{-}H9b}$ 10.4Hz, $J_{H7b\text{-}H8}$ 6.3Hz, $J_{H7a\text{-}H8}$ 5.0Hz, H8), 6.94-6.99 (3H, m, 3-OBn ArH_{meta}, 3-OBn ArH_{para}), 7.06 (2H, dd, $J_{ortho\text{-}meta}$ 7.1Hz, $J_{ortho\text{-}para}$ 2.4Hz, 3-OBn ArH_{ortho}), 7.30-7.39 (5H, m, 6-OBn ArH), 7.70 (3H, br. s., NPhth), 7.81 (1H, br. s., NPhth); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 55.3 (1C, C2), 69.7 (1C, C7), 70.7 (1C, C6), 73.5 (1C, C5), 73.8 (1C, Ph$\underline{C}H_2$), 74.3 (1C, Ph$\underline{C}H_2$), 74.5 (1C, C4), 78.6 (1C, C3), 97.3 (1C, C1), 117.3 (1C, C9), 123.2, 123.4 (2 x 1C, ArC_2, ArC_5), 127.4, 127.8, 127.9, 127.9, 128.1, 128.5 (4 x 2C, 2 x 1C, OBn ArC_{ortho}, OBn ArC_{meta}, OBn ArC_{para}), 131.7 (2C, ArC_1, ArC_6), 133.6 (1C, C8), 133.8 (2C, ArC_3, ArC_4), 137.6, 138.2 (2 x 1C, OBn ArC_{ipso}), 167.6, 168.3 (2 x 1C, 2 x NCO); m/z (ES^+) 552.2 $[M + Na]^+$, (ES^-) 528.2 $[M - H]^-$, 564.2 $[M + Cl]^-$.

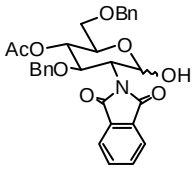


β-*O*-allyl-*N*-phthalamido-3,6-dibenzyl-4-acetyl-D-glucose[33] (5.25)

To a solution of **5.24** (3.69g, 6.98mmol, 1eq) in anhydrous DCM (20ml) and pyridine (1ml) at -5°C was added dropwise triflic anhydride (1.47ml, 8.73mmol, 1.25eq). The reaction mixture was stirred at this temperature for 30 minutes after which TLC (6 ethyl acetate : 4 petrol) showed complete consumption of starting material and the formation of a single product. The reaction mixture was quenced with ice water (20ml) and diluted with DMF (20ml). DCM and water was removed under reduced pressure at 0°C to give a solution of the

crude triflate in DMF. To this solution was added Bu_4NOAc (10.5g, 34.9mmol, 5eq) and the mixture stirred at room temperature for 18 hours. TLC (6 ethyl acetate : 4 petrol) showed the formation of two products- acetyl and hydroxyl subsitution products. Anhydrous pyridine (10ml) and acetic anhydride (5ml) were added to the reaction mixture and stirred at room temperature for 5 hours, after which TLC showed the presence of a single product. The reaction mixture was partitoned between ethyl acetate (500ml) and water (200ml). The organic layer was washed with brine (100ml), 1M HCl (5 x 100ml), brine (5 x 100ml), dried over $MgSO_4$ and the solvent removed *in vacuo*. The product was purified by silica gel chromatography (10% to 60% ethyl acetate in petrol over 10 column volumes) to give the title compound as a pale yellow oil (3.75g, 95%).

R_f 0.75 (6 ethyl acetate : 4 petrol), $[\alpha]_D^{25}$ +60.1 (c = 1.0, $CHCl_3$), lit.[33] +65.4 (c = 0.37, $CHCl_3$); ^{1}H NMR (500MHz, $CCCl_3$) δ ppm 2.15 (3H, s, OAc), 3.59-3.69 (2H, m, H6a, H6b), 3.76 (1H, dt, $J_{H4\text{-}H5}$ 10.1Hz, $J_{H5\text{-}H6a}$ 5.0Hz, $J_{H5\text{-}H6b}$ 5.0Hz, H5), 4.02 (1H, ddt, $J_{H7a\text{-}H7b}$ 12.9Hz, $J_{H7b\text{-}H8}$ 6.0Hz, $J_{H7b\text{-}H9a}$ 1.6Hz, $J_{H7b\text{-}H9b}$ 1.6Hz, H7b), 4.26 (1H, ddt, $J_{H7a\text{-}H7b}$ 12.9Hz, $J_{H7a\text{-}H8}$ 5.0Hz, $J_{H7a\text{-}H9a}$ 2.2Hz, $J_{H7a\text{-}H9b}$ 2.2Hz, H7a), 4.28 (1H, dd, $J_{H2\text{-}H3}$ 10.7Hz, $J_{H1\text{-}H2}$ 8.5Hz, H2), 4.33 (1H, d, ^2J 12.0Hz, PhC$\underline{H}_2{}^1$), 4.45 (1H, dd, $J_{H2\text{-}H3}$ 10.9Hz, $J_{H3\text{-}H4}$ 9.0Hz, H3), 4.57 (1H, d, ^2J 12.0Hz, PhC$\underline{H}_2{}^2$), 4.61 (1H, d, ^2J 12.0Hz, PhC$\underline{H}_2{}^2$), 4.60 (1H, d, ^2J 12.0Hz, PhC$H_2{}^1$), 5.02 (1H, dq, $J_{H8\text{-}H9b}$ 10.4Hz, $J_{H7a\text{-}H9b}$ 1.9Hz, $J_{H7b\text{-}H9b}$ 1.9Hz, $J_{H9a\text{-}H9b}$ 1.9Hz, H9b), 5.10 (1H, dq, $J_{H8\text{-}H9a}$ 17.4Hz, $J_{H7a\text{-}H9a}$ 1.6Hz, $J_{H7b\text{-}H9a}$ 1.6Hz, $J_{H9a\text{-}H9b}$ 1.6Hz, H9a), 5.14 (1H, t, $J_{H3\text{-}H4}$ 9.1Hz, $J_{H4\text{-}H5}$ 9.1Hz, H4), 5.19 (1H, d, $J_{H1\text{-}H2}$ 8.5Hz, H1), 5.68 (1H, dddd, $J_{H8\text{-}H9a}$ 17.5Hz, $J_{H8\text{-}H9b}$ 10.4Hz, $J_{H7b\text{-}H8}$ 6.3Hz, $J_{H7a\text{-}H8}$ 5.0Hz,H8), 6.87-7.07 (5H, m, 3-OBn ArH), 7.28-7.40 (3H, m, 6-OBn ArH), 7.57-7.88 (4H, m, NPhth); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.9 (1C, CO$\underline{C}H_3$), 55.4 (1C, C2), 69.7 (1C, C7), 69.8 (1C, C6), 72.5 (1C, C4), 73.4 (1C, C5), 73.6 (1C, Ph$\underline{C}H_2$), 73.6 (1C, Ph$\underline{C}H_2$), 73.8 (1C, C3), 97.2 (1C, C1), 117.5 (1C, C9), 123.4 (2C, ArC_2, ArC_5), 127.4, 127.5, 127.7, 127.8, 127.8, 128.0, 128.0, 128.1, 128.1, 128.3, 128.4, 128.5 (10C, OBn ArC_{ortho}, OBn ArC_{meta}, OBn ArC_{para}), 131.6 (2C, ArC_1, ArC_6), 133.5 (1C, C8), 133.8 (2C, ArC_3, ArC_4), 137.7, 137.9 (2 x 1C, OBn ArC_{ipso}), 167.6, 168.2 (2 x 1C, 2 x NCO), 169.6 (1C, $\underline{C}OCH_3$); m/z (ES^+) 572.1 $[M + H]^+$, 594.2 $[M + Na]^+$.



_N_-phthalamido-3,6-dibenzyl-4-acetyl-D-glucose[33] (5.26)

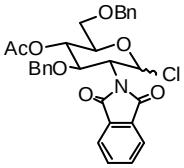
A mixture of **5.25** (3.75g, 6.57mmol, 1eq), $(Ph_3P)_3RhCl$ (910mg, 0.99mmol, 0.15eq) and DABCO (295mg, 2.63mmol, 0.4eq) in 7 ethanol : 3 toluene : 1 water (100ml) was heated at 100°C for 22 hours. The solvent was removed under reduced pressure and the crude mixture re-dissolved in 9 acetone : 1 water. To this solution was added HgO (30mg, 0.138mmol, 0.02eq) and $HgCl_2$ (9.0g, 33.1mmol, 5eq), and the reaction mixture stirred at room temperature for 3 hours after which TLC (6 ethyl acetate : 4 petrol) showed complete consumption of starting material. The solvent was removed *in vacuo* and the crude product partitioned between ethyl acetate (300ml) and 30% NaBr solution (300ml). The organic layer was washed with 10% $Na_2S_2O_3$ (300ml), brine (200ml), dried over $MgSO_4$ and concentrated *in vacuo*. The products were separated by silica gel chromatography (10% to 80% ethyl acetate in petrol over 9 column volumes) to yield the title compound (1.85g, 53%) as a white foam.

R_f 0.27 (6 petrol : 4 ethyl acetate), $[\alpha]_D^{25}$ +74.8 (c = 1.0, $CHCl_3$), m/z (ES^+) 554.2 [M + Na]$^+$; α:β 0.38:1

β anomer: ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 1.94 (3H, s, Oac), 3.56-3.60 (2H, m, H6a, H6b), 3.80 (2 H, ddd, $J_{H4\text{-}H5}$ 9.9Hz, $J_{H5\text{-}H6a}$ 5.4Hz, $J_{H5\text{-}H6b}$ 3.6Hz, H5), 4.18 (1H, dd, $J_{H2\text{-}H3}$ 10.7Hz, $J_{H1\text{-}H2}$ 8.5Hz, H2), 4.34 (1H, d, ^2J 12.0Hz, PhC$\underline{H}_2^1$), 4.52 (1H, dd, $J_{H2\text{-}H3}$ 10.7Hz, $J_{H3\text{-}H4}$ 8.8Hz, H3), 4.50-4.57 (1H, m, PhC$\underline{H}_2^2$), 4.61 (1H, d, ^2J 12.0Hz, PhC$\underline{H}_2^2$), 4.61 (1H, d, ^2J 12.0Hz, PhC$\underline{H}_2^1$), 5.15 (1H, dd, $J_{H4\text{-}H5}$ 10.1Hz, $J_{H3\text{-}H4}$ 9.1Hz, H4), 5.36 (1H, br. S., H1), 6.88-7.11 (5H, m, 3-Obn ArH), 7.28-7.40 (5H, m, 6-Obn ArH), 7.58-7.91 (4H, m, NPhth); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.8 (1C, CO$\underline{C}H_3$), 57.1 (1C, C2), 69.2 (1C, C6), 72.2 (1C, C4), 73.6 (1C, C5), 73.6 (1C, Ph$\underline{C}H_2$), 73.9 (1C, Ph$\underline{C}H_2$), 76.7 (1C, C3), 92.9 (1C, C1), 123.4 (2C, ArC_2, ArC_5), 127.4, 127.7, 127.8, 128.0, 128.0, 128.1, 128.1, 128.3, 128.4, 128.5 (Obn ArC_{ortho}, Obn ArC_{meta}, Obn ArC_{para} α and β), 131.5 (2C, ArC_1, ArC_6), 133.9 (2C, ArC_3, ArC_4), 137.6, 137.7 (2 x 1C, Obn ArC_{ipso}), 168.0 (2C, NCO), 169.6 (1C, $\underline{C}OCH_3$).

α anomer: ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 1.92 (3H, s, OAc), 3.55 (1H, m, H6b), 3.64 (1H, dd, $J_{H6a\text{-}H6b}$ 9.5Hz, $J_{H5\text{-}H6a}$ 6.3Hz, H6a), 4.26 (1H, d, ^2J 12.3Hz, PhC$\underline{H}_2^1$), 4.41 (1H, m, $J_{H5\text{-}H6b}$ 4.3Hz, H5), 4.43 (1H, d, ^2J 12.3Hz, PhC$\underline{H}_2^2$), 4.50-4.57 (2H, m, H2, PhC$\underline{H}_2^2$), 4.62 (1H, d, ^2J 12.0Hz, PhC$\underline{H}_2^1$), 4.93 (1H, dd, $J_{H2\text{-}H3}$ 11.3Hz, $J_{H3\text{-}H4}$ 9.1Hz, H3), 5.23 (1H, dd, $J_{H4\text{-}H5}$ 10.1Hz, $J_{H3\text{-}H4}$ 9.1Hz, H4), 5.38 (1H, br. s, H1), 6.88-

7.11 (5H, m, 3-OBn ArH), 7.28-7.40 (5H, m, 6-OBn ArH), 7.58-7.91 (4H, m, NPhth); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.9 (1C, CO<u>C</u>H$_3$), 55.6 (1C, C2), 68.9 (1C, C6), 71.1 (1C, Ph<u>C</u>H$_2$), 72.5 (1C, C4), 73.5 (1C, C5), 74.0 (1C, Ph<u>C</u>H$_2$), 74.2 (1C, C3), 92.7 (1C, C1), 123.6 (2C, ArC_2, ArC_5), 127.4, 127.7, 127.8, 128.0, 128.0, 128.1, 128.1, 128.3, 128.4, 128.5 (OBn ArC_{ortho}, OBn ArC_{meta}, OBn ArC_{para} α and β), 131.5 (2C, ArC_1, ArC_6), 134.2 (1C, ArC_3, ArC_4), 137.7, 137.9 (2 x 1C, OBn ArC_{ipso}), 168.2 (1C, NCO), 169.6 (1C, <u>C</u>OCH$_3$).

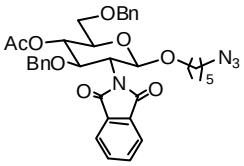


***N*-phthalamido-3,6-dibenzyl-4-acetyl-D-glucose[33] (5.27)**

To a solution of **5.26** (0.60g, 1.13mmol, 1eq) and DMF (26µl, 0.34mmol, 0.3eq) in anhydrous DCM (10ml) was added dropwise oxallyl chloride (0.29ml, 3.39mmol, 3eq). The reaction mixture was stired at room temperature 5 hours, after which the mixture was diluted with DCM (50ml) and quenched with ice water (~50ml). The organic layer was retained, dried over $MgSO_4$ and the solvent removed *in vacuo* to yield the desired compound as a pale yellow foam (0.60g, 97%).

R_f 0.48 (1 petrol : 1 ethyl acetate), $[\alpha]_D^{25}$ +81.1 (c = 0.31, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 1.90 (3H, s, OAcα), 1.95 (3H, s, OAcβ), 3.58 (2H, dd, $J_{H6a\alpha\text{-}H6b\alpha}$ 9.8Hz, $J_{H5\alpha\text{-}H6b\alpha}$ 7.3Hz, H6bα), 3.61 (1H, dd, $J_{H6a\beta\text{-}H6b\beta}$ 11.3Hz, $J_{H5\beta\text{-}H6b\beta}$ 3.8Hz, H6bβ), 3.62 (1H, dd, $J_{H6a\beta\text{-}H6b\beta}$ 11.3Hz, $J_{H5\beta\text{-}H6a\beta}$ 5.0Hz, H6aβ), 3.67 (1H, dd, $J_{H6a\alpha\text{-}H6b\alpha}$ 9.5Hz, $J_{H5\alpha\text{-}H6a\alpha}$ 6.0Hz, H6aα), 3.85 (1H, ddd, $J_{H4\beta\text{-}H5\beta}$ 9.8Hz, $J_{H5\beta\text{-}H6a\beta}$ 4.7Hz, $J_{H5\beta\text{-}H6b\beta}$ 3.8Hz, H5β), 4.25 (1H, d, ^2J 12.6Hz, PhC<u>H</u>$_2$α), 4.32 (2H, m, H3α H5α), 4.32 (1H, d, ^2J 12.0Hz, PhC<u>H</u>$_2$β), 4.41-4.46 (2H, m, H2β, H3β), 4.51-4.67 (3H, m, PhC<u>H</u>$_2$α, PhC<u>H</u>$_2$β), 4.61 (1H, dd, $J_{H2\alpha\text{-}H3\alpha}$ 11.7Hz, $J_{H1\alpha\text{-}H2\alpha}$ 2.8Hz, H2α), 5.22 (1H, dd, $J_{H4\beta\text{-}H5\beta}$ 10.1Hz, $J_{H3\beta\text{-}H4\beta}$ 8.8Hz, H4β), 5.42 (1H, dd, $J_{H4\alpha\text{-}H5\alpha}$ 11.0Hz, $J_{H3\alpha\text{-}H4\alpha}$ 9.1Hz, H4α), 5.99 (1H, d, $J_{H1\beta\text{-}H2\beta}$ 8.8Hz, H1β), 6.23 (1H, d, $J_{H1\alpha\text{-}H2\alpha}$ 3.8Hz, H1α), 6.86-7.13 (5H, m, 3-OBn ArH), 7.28-7.43 (5H, m, 6-OBn ArH), 7.59-7.92 (4H, m, NPhth); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.8 (1C, CO<u>C</u>H$_3$β), 20.9 (1C, CO<u>C</u>H$_3$α), 57.2 (1C, C2α), 58.4 (1C, C2β), 68.1 (1C, C6α), 69.0 (1C, C6β), 71.2 (1C, Ph<u>C</u>H$_2$α), 71.7 (1C, C4β), 72.3 (1C, C3α), 72.8 (1C, C5α), 73.5 (1C, Ph<u>C</u>H$_2$α), 73.6 (1C, Ph<u>C</u>H$_2$α), 73.7 (1C, Ph<u>C</u>H$_2$β), 73.9 (1C, C4α), 74.1 (1C, Ph<u>C</u>H$_2$β), 76.1 (1C, C3), 77.2 (1C, C5, obscured by $CDCl_3$), 86.1 (1C, C1β), 92.4 (1C, C1α), 123.6 (2C, ArC_2β, ArC_5β), 123.6 (1C, ArC_2α, ArC_5α), 127.0, 127.4, 127.6, 127.7, 127.8, 127.8, 128.0, 128.1, 128.1, 128.2, 128.4, 128.5, 128.7 (ArC), 131.3 (2C, ArC_2β, ArC_5β), 131.5 (2C, ArC_2α, ArC_5α),

134.1 (2C, $ArC_1\beta$, $ArC_6\beta$), 134.2 (2C, $ArC_1\alpha$, $ArC_6\alpha$), 137.0, 137.5 (2 x 1C, OBn $ArC_{ipso}\beta$), 137.3, 137.6 (2 x 1C, OBn $ArC_{ipso}\alpha$), 169.4 (1C, $\underline{C}OCH_3\beta$), 169.5 (1C, $\underline{C}OCH_3\alpha$); m/z (ES^+) 568.2 $[M + NH_4]^+$, 572.2 $[M + Na]^+$.

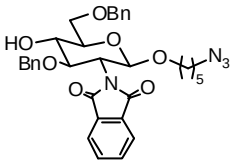


_N_-phthalamido-1-(5-azido-pentan-1-ol)-3,6-dibenzyl-4-acetyl-D-glucose (5.28)

A solution of **5.27** (0.60g, 1.09mmol, 1eq), **5.3** (155mg, 1.20mmol, 1.1eq) and collidine (106mg, 0.87mmol, 0.8eq) in anhydrous DCM (20ml) was stirred with 3Å molecular sieves for 1 hour. The reaction mixture was cooled to -40°C and AgOTf (350mg, 1.36mmol, 1.25eq) was added. The reaction mixture was stirred at this temperature for 7 hours and then allowed to warm to room temperature. After 15 hours, TLC (2 petrol : 1 ethyl acetate) showed complete consumption of starting material. The reaction mixture was diluted with DCM (50ml) and filtered through Celite®. The filtrate was washed with water (2 x 50ml), 1M HCl (100ml), saturated $NaHCO_3$ (100ml), brine (100ml) and dried over $MgSO_4$. The desired product was purified by silica gel chromatography (6% to 50% ethyl acetate in petrol over 9 column volumes) to give the desired compound as a colourless oil (632mg, 90%).

R_f 0.25 (3 petrol : 1 ethyl acetate), $[\alpha]_D^{25}$ +35.8 (c = 0.83, $CHCl_3$); ^1H NMR (500MHz, $CDCl_3$) δ ppm 1.13 (2H, quin., $J_{H8\text{-}H9}$ 6.6Hz, $J_{H9\text{-}H10}$ 6.6Hz, H9), 1.26-1.51 (4H, m, H8, H10), 1.95 (3H, s, OAc), 2.90 (1H, dt, $J_{H11a\text{-}H11b}$ 12.3Hz, $J_{H10a\text{-}H11b}$ 6.9Hz, $J_{H10b\text{-}H11b}$ 6.9Hz, H11b), 2.95 (1H, dt, $J_{H11a\text{-}H11b}$ 12.3Hz, $J_{H10a\text{-}H11a}$ 6.9Hz, $J_{H10b\text{-}H11a}$ 6.9Hz, H11a), 3.40 (1H, ddd, $J_{H7a\text{-}H7b}$ 10.1Hz, $J_{H7b\text{-}H8a}$ 7.3Hz, $J_{H7b\text{-}H8b}$ 5.7Hz, H7b), 3.62 (2H, d, $J_{H5\text{-}H6}$ 4.7Hz, H6a, H6b), 3.76 (1H, dt, $J_{H4\text{-}H5}$ 10.1Hz, $J_{H5\text{-}H6a}$ 4.4Hz, $J_{H5\text{-}H6b}$ 4.4Hz, H5), 3.82 (1H, dt, $J_{H7a\text{-}H7b}$ 9.9Hz, $J_{H7a\text{-}H8a}$ 6.1Hz, $J_{H7a\text{-}H8b}$ 6.1Hz, H7a), 4.24 (1H, dd, $J_{H2\text{-}H3}$ 11.0Hz, $J_{H1\text{-}H2}$ 8.5Hz, H2), 4.33 (1H, d, ^2J 12.0Hz, $OC\underline{H}_2Ph^1$), 4.43 (1H, dd, $J_{H2\text{-}H3}$ 10.9Hz, $J_{H3\text{-}H4}$ 9.0Hz, H3), 4.56 (2H, s, $OC\underline{H}_2Ph^2$), 4.61 (1H, d, ^2J 12.0Hz, $OC\underline{H}_2Ph^1$), 5.13 (1H, dd, $J_{H4\text{-}H5}$ 10.1Hz, $J_{H3\text{-}H4}$ 9.1Hz, H4), 5.14 (1H, d, $J_{H1\text{-}H2}$ 8.2Hz, H1), 6.87-7.03 (5H, m, 3-OBn, ArH), 7.28-7.38 (5H, m, 6-OBn, ArH), 7.61-7.87 (4H, m, NPhth); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.9 (1C, $CO\underline{C}H_3$), 23.0 (1C, C9), 28.3 (1C, C10), 28.7 (1C, C8), 51.1 (1C, C11), 55.5 (1C, C2), 69.2 (1C, C7), 69.7 (1C, C6), 72.6 (1C, C4), 73.5 (1C, C5), 73.6 (1C, $O\underline{C}H_2Ph$), 73.8 (1C, $O\underline{C}H_2Ph$), 77.0 (1C, C3), 98.2 (1C, C1), 123.3 (2C, ArC_2, ArC_5), 127.4, 127.7, 127.8, 127.8, 128.1, 128.4 (10C, 6 x $OCH_2\underline{Ph}$ ArC), 131.5 (2C, ArC_1, ArC_6) 133.9 (2C,

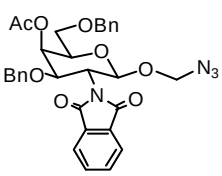
ArC_3, ArC_4), 137.7, 137.9 (2 x 1C, 2 x OBn ArC_{ipso}), 167.3, 168.3 (2 x C, 2 x NPhth C=O) 169.7 (1C, $\underline{C}OCH_3$); FT-IR (thin film) υ 1715 (C=O, NPhth), 1747 (C=O, OAc), 2096 (N_3); HRMS (ES^+) m/z 665.2587 $[M + Na]^+$ (required 665.2582).



_N_-phthalamido-1-(5-azido-pentan-1-ol)-3,6-dibenzyl-D-glucose (5.29)

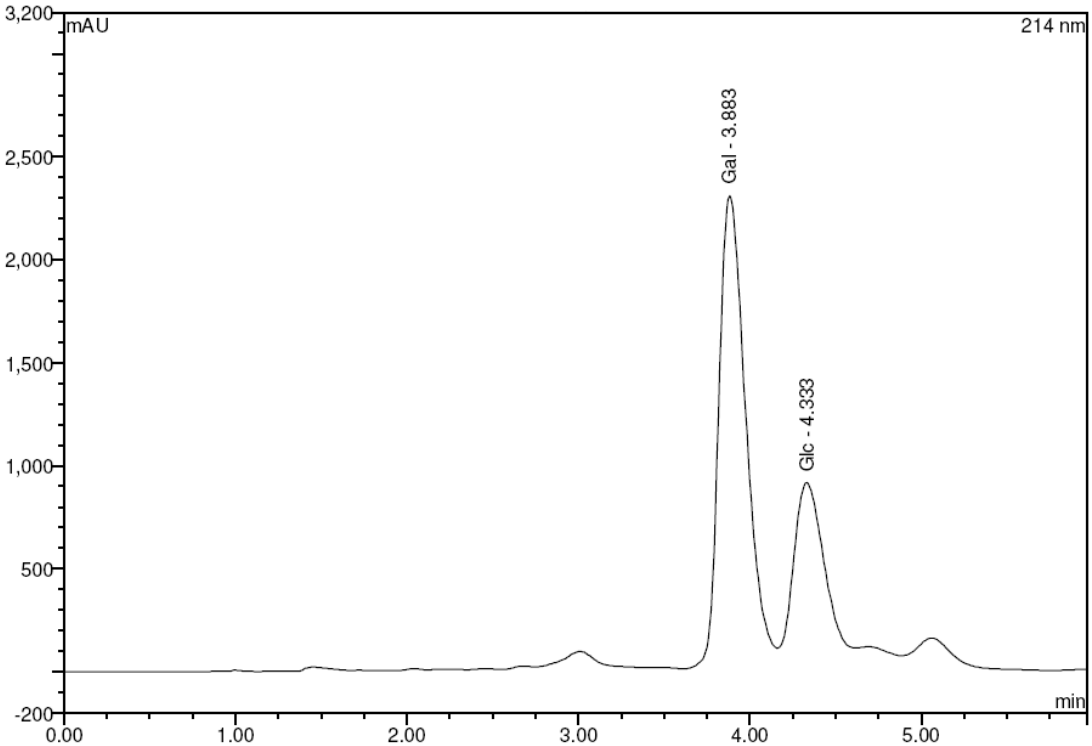
A solution of **5.28** (630mg, 0.98mmol, 1eq) and NaOMe (5mg, 0.1mmol, 0.1eq) in methanol (10ml) was stirred at room temperature for 6 hours, after which TLC (3 petrol : 1 ethyl acetate) showed complete consumption of starting material. The reaction was quenched by addition of Dowex 50WX8 (H^+) until neutral. The reaction mixture was filtered and the solvent removed *in vacuo* to yield the desired compound as a colourless oil (560mg, 95%).

R_f 0.22 (3 petrol : 1 ethyl acetate), $[\alpha]_D^{25}$ +26.8 (c = 0.23, $CHCl_3$); ^1H NMR (500MHz, $CDCl_3$) δ ppm 1.13 (2H, quin., $J_{H8\text{-}H9}$ 6.6Hz, $J_{H9\text{-}H10}$ 6.6Hz, H9), 1.30-1.51 (4H, m, H8, H10), 2.63 (1H, br. s, OH), 2.85-2.97 (4H, m, H11), 3.37 (1H, ddd, $J_{H7a\text{-}H7b}$ 9.8Hz, $J_{H7b\text{-}H8a}$ 6.9Hz, $J_{H7b\text{-}H8b}$ 5.7Hz, H7b), 3.65 (1H, dt, $J_{H4\text{-}H5}$ 9.8Hz, $J_{H5\text{-}H6a}$ 5.0Hz, $J_{H5\text{-}H6b}$ 5.0Hz), 3.76-3.89 (4H, m, H4, H6a, H6b, H7a), 4.15 (1H, dd, $J_{H2\text{-}H3}$ 10.4Hz, $J_{H1\text{-}H2}$ 8.5Hz, H2), 4.23 (1H, dd, $J_{H2\text{-}H3}$ 10.7Hz, $J_{H3\text{-}H4}$ 8.5Hz, H3), 4.55 (1H, d, ^2J 12.3Hz, $OC\underline{H}_2Ph^1$), 4.60 (1H, d, ^2J 12.0Hz, $OC\underline{H}_2Ph^2$), 4.66 (1H, d, ^2J 12.3Hz, $OC\underline{H}_2Ph^2$), 4.76 (1H, d, ^2J 12.3Hz, $OC\underline{H}_2Ph^1$), 5.14 (1H, d, $J_{H1\text{-}H2}$ 8.5Hz, H1), 6.91-7.09 (5H, m, 3-OBn ArH), 7.29-7.41 (5H, m, 6-OBn ArH), 7.63-7.93 (4H, m, NPhth ArH); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 23.0 (1C, C9), 28.3 (1C, C10), 28.7 (1C, C8), 51.1 (1C, C11), 55.4 (1C, C2), 69.1 (1C, C7), 70.8 (1C, C6), 73.5 (1C, C5), 73.8 (1C, $O\underline{C}H_2Ph$), 74.3 (1C, $O\underline{C}H_2Ph$), 74.6 (1C, C4), 78.7 (1C, C3), 98.3 (1C, C1), 123.1, 123.4 (2 x 1C, ArC_2, ArC_5), 127.4, 127.8, 127.8, 127.9, 127.9, 127.9, 128.2, 128.5 (10C, OBn ArC), 131.6, 131.7 (2 x 1C, ArC_1, ArC_6), 133.8 (2C, ArC_3, ArC_4), 137.2, 137.6 (2 x 1C, 2 x 1C, 2 x $OCH_2\underline{P}h$ ArC_{ipso}), 167.7, 168.3 (2 x 1C, 2 x C=O); FT-IR (thin film) υ 1713 (C=O), 2096 (N_3), 2929 (OH); HRMS (ES^+) m/z 623.2468 $[M + Na]^+$ (required 623.2476).

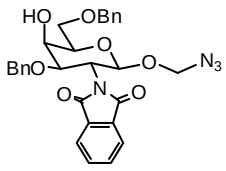


_N_-phthalamido-1-(5-azido-pentan-1-ol)-3,6-dibenzyl-4-acetyl-D-galactose (5.30)

To a solution of **5.29** (325mg, 0.54mmol, 1eq) in anhydrous DCM (10ml) and pyridine (0.5ml) at 0°C was added triflic anhydride (140µl, 0.81mmol, 1.5eq) dropwise over 2 minutes. The reaction mixture was stired at this temperature for 1 hour after which TLC (3 petrol : 1 ethyl acetate) showed complete consumption of starting material. The reaction was quenched by addition of ice (~50ml) and allowed to warm to temperature. The organic layer was washed with ice cold 1M HCl (4 x 50ml), saturated $NaHCO_3$ (50ml), brine (50ml) and dried over $MgSO_4$. The solvent was removed under reduced pressure and the crude triflate product **5.30** dried under high vacuum for 1 hour. The triflate **5.30** was dissolved in dry DMF (10ml) and Bu_4NOAc (814mg, 2.70mmol, 5eq) added. The reaction mixture was stirred under argon at room temperature for 19 hours, after which TLC (3 petrol : 1 ethyl acetate) showed complete consumption of the triflate. The reaction mixture was diluted with ethyl acetate (100ml) and washed with brine (5 x 100ml) and dried over $MgSO_4$. The solvent was removed *in vacuo* and the two products separated by HPLC using a Phenomenex Synergi Hydro (100 x 21.2mm, 4µm) with 65% acetonitrile in water as the mobile phase at a flow rate of 1.0ml/min. Eluants were detected with an in-line UV absorbance detector at 214nm. The desired compound was purified over multiple injections as a colourless oil (129mg, 37%).



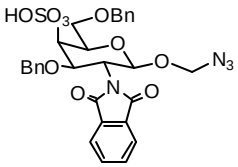
R_f 0.29 (3 petrol : 1 ethyl acetate), $[\alpha]_D^{25}$ +36.3 (c = 0.44, $CHCl_3$); ^1H NMR (500MHz, $CDCl_3$) δ ppm 1.14 (2H, quin., $J_{H8\text{-}H9}$ 6.3Hz, $J_{H9\text{-}H10}$ 6.3Hz, H9), 1.28-1.52 (4H, m, H8, H10), 2.15 (3H, s, OAc), 2.89 (1H, dt, $J_{H11a\text{-}H11b}$ 12.3Hz, $J_{H10a\text{-}H11b}$ 6.9Hz, $J_{H10b\text{-}H11b}$ 6.9Hz, H11b), 2.95 (1H, dt, $J_{H11a\text{-}H11b}$ 12.0Hz, $J_{H10a\text{-}H11a}$ 6.9Hz, $J_{H10b\text{-}H11a}$ 6.9Hz, H11a), 3.38 (1H, ddd, $J_{H7a\text{-}H7b}$ 9.8Hz, $J_{H7b\text{-}H8a}$ 7.3Hz, $J_{H7b\text{-}H8b}$ 6.0Hz, H7b), 3.59 (1H, dd, $J_{H6a\text{-}H6b}$ 9.5Hz, $J_{H5\text{-}H6b}$ 6.9Hz, H6b), 3.65 (1H, dd, $J_{H6a\text{-}H6b}$ 9.5Hz, $J_{H5\text{-}H6a}$ 6.0Hz, H6a), 3.83 (1H, dt, $J_{H7a\text{-}H7b}$ 9.8Hz, $J_{H7a\text{-}H8a}$ 6.0Hz, $J_{H7a\text{-}H8b}$ 6.0Hz, H7a), 3.93 (1H, ddd, $J_{H5\text{-}H6b}$ 7.3Hz, $J_{H5\text{-}H6a}$ 6.0Hz, $J_{H4\text{-}H5}$ 0.9Hz, H5), 4.27 (1H, d, ^2J 12.3Hz, $OC\underline{H}_2Ph^1$), 4.28 (1H, dd, $J_{H2\text{-}H3}$ 11.2Hz, $J_{H3\text{-}H4}$ 3.3Hz, H3), 4.39 (1H, dd, $J_{H2\text{-}H3}$ 11.0Hz, $J_{H1\text{-}H2}$ 8.5Hz, H2), 4.51 (1H, d, ^2J 12.6Hz, $OC\underline{H}_2Ph^2$), 4.61 (1H, d, ^2J 12.6Hz, $OC\underline{H}_2Ph^2$), 4.61 (1H, d, ^2J 12.0Hz, $OC\underline{H}_2Ph^1$), 5.13 (1H, d, $J_{H1\text{-}H2}$ 8.5Hz, H1), 5.67 (1H, dd, $J_{H3\text{-}H4}$ 3.5Hz, $J_{H4\text{-}H5}$ 0.6Hz, H4), 6.93 (2H, t, $J_{ortho\text{-}meta}$ 7.6Hz, $J_{meta\text{-}para}$ 7.6Hz, 3-OBn ArH_{meta}), 6.98 (2H, dd, $J_{ortho\text{-}meta}$ 7.6Hz, $J_{ortho\text{-}para}$ 1.6Hz, 3-OBn ArH_{ortho}), 7.04 (1H, tt, $J_{meta\text{-}para}$ 7.3Hz, $J_{ortho\text{-}para}$1.6Hz, 3-OBn ArH_{para}), 7.28-7.41 (5H, m, 6-OBn ArH), 7.67 (1H, dd, $J_{ArH2\text{-}ArH3}$ 5.7Hz, $J_{ArH2\text{-}ArH4}$ 1.9Hz, ArH_2), 7.73 (1H, td, $J_{ArH2\text{-}ArH3}$ 5.7Hz, $J_{ArH3\text{-}ArH4}$ 5.7Hz, $J_{ArH3\text{-}ArH5}$ 1.3Hz, ArH3), 7.75 (1H, $J_{ArH3\text{-}ArH4}$ 7.6Hz, $J_{ArH4\text{-}ArH5}$ 7.6Hz, $J_{ArH2\text{-}ArH4}$ 1.3Hz, ArH4), 7.86 (1H, dd, $J_{ArH4\text{-}ArH5}$ 6.0Hz, $J_{ArH3\text{-}ArH5}$ 1.9Hz, ArH5); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 20.9 (1C, $CO\underline{C}H_3$), 23.0 (1C, C9), 28.3 (1C, C10), 28.7 (1C, C8), 51.1 (1C, C11), 53.0 (1C, C2), 65.9 (1C, C4), 68.0 (1C, C6), 69.3 (1C, C7), 71.1 (1C, $O\underline{C}H_2Ph$), 72.3 (1C, C5), 73.0 (1C, C3), 73.7 (1C, $O\underline{C}H_2Ph$), 98.6 (1C, C1), 123.1, 123.4 (2 x 1C, ArC_2, ArC_5), 127.6, 127.9, 128.0, 28.1, 128.5 (10C, $OCH_2\underline{Ph}$), 131.6, 131.6 (2 x 1C, ArC_1, ArC_6), 133.8, 134.0 (2 x 1C, ArC_3, ArC_4), 137.4, 137.6 (2 x 1C, 2 x $OCH_2\underline{P}h$ ArC_{ipso}), 167.6, 168.2 (2 x 1C, 2 x NPhth C=O), 170.5 (1C, $\underline{C}OCH_3$); FT-IR (thin film) υ 1713 (C=O, NPhth), 1744 (C=O, OAc), 2096 (N_3); HRMS (ES^+) m/z 665.2580 $[M + Na]^+$ (required 665.2582).



_N_-phthalamido-1-(5-azido-pentan-1-ol)-3,6-dibenzyl-D-galactose (5.31)

A solution of **5.30** (0.10g, 0.155mmol, 1eq) and NaOMe (8mg, 0.155mmol, 1eq) in methanol (10ml) was stirred at room temperature for 5 hours, after which TLC (3 petrol : 1 ethyl acetate) showed complete consumption of starting material. The reaction was quenched by addition of Dowex 50WX8 (H^+) until neutral. The reaction mixture was filtered and the solvent removed *in vacuo* to yield the desired compound as a colourless oil (92mg, 99%).

R_f 0.15 (3 petrol : 1 ethyl acetate), $[\alpha]_D^{22}$ +26.5 (c = 0.43, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 1.14 (2H, quin., $J_{H8\text{-}H9}$ 6.3Hz, $J_{H9\text{-}H10}$ 6.3Hz, H9), 1.28-1.51 (4H, m, H8, H10), 2.62 (1H, br. s, OH), 2.90 (1 H, dt, $J_{H11a\text{-}H11b}$ 12.3Hz, $J_{H10a\text{-}H11b}$ 6.9Hz, $J_{H10b\text{-}H11b}$ 6.9Hz, H11b), 2.96 (1H, dt, $J_{H11a\text{-}H11b}$ 12.0Hz, $J_{H10a\text{-}H11a}$ 6.9Hz, $J_{H10b\text{-}H11a}$ 6.9Hz, H11a), 3.38 (1H, ddd, $J_{H7a\text{-}H7b}$ 10.1Hz, $J_{H8a\text{-}H7b}$ 7.3Hz, $J_{H8b\text{-}H7b}$ 5.7Hz, H7b), 3.76-3.84 (3H, m, H5, H6a, H6b), 3.88 (1H, dt, $J_{H7a\text{-}H7b}$ 8.5Hz, $J_{H7a\text{-}H8a}$ 6.3Hz, $J_{H7a\text{-}H8b}$ 6.3Hz, H7a), 4.17 (1H, d, $J_{H3\text{-}H4}$ 3.2Hz, H4), 4.25 (1H, dd, $J_{H2\text{-}H3}$ 10.9Hz, $J_{H3\text{-}H4}$ 3.3Hz, H3), 4.34 (1H, d, ^2J 12.3Hz, O$\underline{C}H_2Ph^1$), 4.45 (1H, dd, $J_{H2\text{-}H3}$ 11.0Hz, $J_{H1\text{-}H2}$ 8.5Hz, H2), 4.63 (2H, s, O$\underline{C}H_2Ph^2$), 4.64 (1H, d, ^2J 12.3Hz, O$\underline{C}H_2Ph^1$), 5.10 (1H, d, $J_{H1\text{-}H2}$ 8.5Hz, H1), 6.96-7.07 (5H, m, 3-OBn ArH), 7.34-7.40 (5H, m, 6-OBn ArH), 7.63-7.78 (3H, m, NPhth ArH), 7.87 (1H, d, $J_{ArH2\text{-}ArH3}$ 6.6Hz, NPhth $ArH_{2\ or\ 5}$); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 23.0 (1C, C9), 28.3 (1C, C10), 28.7 (1C, C8), 51.1 (1C, C11), 52.4 (1C, C2), 65.6 (1C, C4), 68.9 (1C, C6), 69.1 (1C, C7), 71.3 (1C, O$\underline{C}H_2$Ph), 73.4 (1C, C5), 73.7 (1C, O$\underline{C}H_2$Ph), 75.0 (1C, C3), 98.3 (1C, C1), 123.0, 123.4 (2 x 1C, ArC_2, ArC_5), 127.8, 127.8, 127.8, 127.9, 128.3, 128.4 (10C, OBn ArC), 131.6, 131.7 (2 x 1C, ArC_1, ArC_6), 133.8, 134.0 (2 x 1C, ArC_3, ArC_4), 137.2, 138.0 (2 x 1C, 2 x 1C, 2 x $OCH_2\underline{P}h$ ArC_{ipso}), 167.7, 168.4 (2 x 1C, 2 x C=O); FT-IR (thin film) υ 1711 (C=O), 2095 (N_3), 2924 (OH); HRMS (ES^+) m/z 623.2485 [M + $Na]^+$ (required 623.2476).



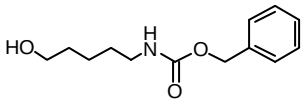
_N_-phthalamido-1-(5-azido-pentan-1-ol)-3,6-dibenzyl-D-galactose-4-sulfate (5.32)

A solution of **5.31** (98mg, 0.163mmol, 1eq) and pyridinium sulphur trioxide complex in anhydrous DMF (10ml) was heated at 60°C for 5 hours. After this time, TLC (2 petrol : 1 ethyl acetate) showed complete consumption of starting material. The solvent was removed *in vacuo* and the crude residue taken up in DCM (100ml) and washed with water (2 x 50ml), brine (50ml) and dried over $MgSO_4$. Removal of solvent under reduced pressure gave the title compound as a pale yellow amorphous solid (107mg, 96%).

R_f 0 (ethyl acetate), $[\alpha]_D^{25}$ +31.1 (c = 1.0, $CHCl_3$); ^{1}H NMR (500MHz, $CDCl_3$) δ ppm 1.10-1.19 (2H, m, H9), 1.33-1.47 (4H, m, H8, H10), 2.91-2.99 (2H, m, H11), 3.43 (1H, dt, $J_{H7a\text{-}H7b}$ 9.9Hz, $J_{H7b\text{-}H8a}$ 6.5Hz, $J_{H7b\text{-}H8b}$ 6.5Hz, H7b), 3.77-3.89 (2H, m, H5, H7a), 3.92 (1H, dd, $J_{H6a\text{-}H6b}$ 9.8Hz, $J_{H5\text{-}H6b}$ 6.6Hz, H6b), 4.16 (1H, dd, $J_{H6a\text{-}H6b}$ 10.2Hz, $J_{H5\text{-}H6a}$ 4.3Hz, H6a), 4.27-4.30 (2H, m, H2, H3), 4.59-4.63 (2H, m, O$\underline{C}H_2Ph^1$,

OC<u>H</u>$_2$Ph^2), 4.70 (2H, d, ^2J 11.6Hz, OC<u>H</u>$_2$Ph^1), 4.91 (1H, d, ^2J 11.3Hz, OC<u>H</u>$_2$Ph^2), 5.16 (1H, d, $J_{H4\text{-}H5}$ 1.9Hz, H4), 5.20 (1H, d, $J_{H1\text{-}H2}$ 8.2Hz, H1), 6.84-7.08 (5H, m, 3-OBn ArH), 7.29-7.44 (6H, m, 6-OH ArH), 7.69-7.88 (4H, m, NPhth ArH); ^{13}C NMR (126MHz, $CDCl_3$) δ ppm 28.3 (1C, C9), 28.8 (1C, C8), 29.7 (1C, C10), 51.1 (1C, C11), 53.2 (1C, C2), 69.0 (1C, C7), 70.0 (1C, C6), 71.2 (1C, O<u>C</u>H$_2$Ph), 71.5 (1C, C4), 73.5 (1C, O<u>C</u>H$_2$Ph), 74.3 (1C, C5), 74.8 (1C, C3), 98.1 (1C, C1), 123.2, 123.3 (2 x 1C, ArC_2, ArC_5), 126.8, 127.5, 127.9, 128.1, 128.3, 128.4 (10C, OBn ArC), 131.5, 131.5 (2 x 1C, ArC_1, ArC_6), 134.1, 134.2 (2 x 1C, ArC_3, ArC_4), 137.4 (2C, 2 x OBn ArC_{ipso}), 168.1, 168.3 (2 x 1C, 2 x C=O); FT-IR (KBr disc) υ 1391 (S=O), 1713 (C=O), 2096 (N_3), 2926 (OH); HRMS (ES^-) m/z 679.2083 $[M - H]^-$ (required 679.2079).

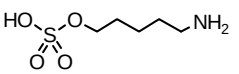
5.11.3 Synthesis of the sulfate ligand



5-Aminobenzyloxycarbonyl pentanol (5.35)

To a solution of **5.5** (6.20 g, 60.0 mmol, 1 eq) in DCM (50ml) and water (50ml) was added benzyloxycarbonyl chloride (12.80 g, 75 mmol, 1.25 eq) and sodium carbonate (15.9 g, 150 mmol, 2.5 eq). After stirring at RT for 4 hours, the reaction mixture was diluted with DCM (250ml) and H_2O (250ml). The organic layer was separated and washed with brine (250ml), dried over $MgSO_4$ and the crude compound purified by automated silica gel chromatography (1 petrol : 1 ethyl acetate) to yield the titled compound (13.0g, 91%) as a white solid.

R_f 0.29 (50% EtOAc/Petrol); ^{1}H NMR (400MHz, $CDCl_3$) δ ppm 1.40 (2H, m, CH_2), 1.65-1.48 (4H, m, OCH$_2$C<u>H</u>$_2$, C<u>H</u>$_2$CH$_2$NHCbz), 3.21 (2H, dd, $J_{CH2\text{-}NH}$ 13.1Hz, $J_{CH2\text{-}CH2}$ 6.7 Hz, CH$_2$NHCbz), 3.64 (2H, br. s, OCH$_2$), 4.84 (1H, s, NH), 5.10 (2H, s, C<u>H</u>$_2$Ph), 7.34-7.27 (5H, m, Ph); ^{13}C NMR (101MHz, $CDCl_3$) δ ppm 22.8 (1C, CH_2), 29.7 (1C, <u>C</u>H$_2$CH$_2$NHCbz), 32.1 (1C, OCH$_2$<u>C</u>H$_2$), 40.8 (1C, <u>C</u>H$_2$NHCbz), 62.5 (1C, OCH$_2$), 66.5 (1C, <u>C</u>H$_2$Ph), 128.0, 128.4, (5C, ArC Ph), 136.5 (1C, ArC_{ipso}), 156.4 (1C, NH<u>C</u>O); HRMS (ES^+) m/z 260.1258 $[M + Na]^+$ (required 260.1257).



5-aminopentan-1-sulfonate (5.37)

To a solution of **5.35** (0.50g, 2.11mmol, 1eq) in anhydrous DCM (2ml) and pyridine (0.1ml) was added dropwise chlorosulfonic acid (154µl, 2.32mmol, 1.1eq). The reaction was stirred at room temperature for 10 minutes, after which TLC (4 DCM : 1 methanol) showed complete consumption of starting material

and the formation of a single product. The reaction was quenched with water (1ml) and the solvents removed *in vacuo*. The crude compound was dissolved in water (0.25ml) and passed through a C_{18} SepPak® cartridge (Waters). The cartridge was eluted with methanol to yield the desired compound **5.36**. The effluent was transferred to a round bottomed flask and made up to 10ml with methanol. Palladium on charcoal (5%, 200mg) was added to this solution, which was repeatedly degassed under vacuum and charged with hydrogen (x3). The mixture was vigorously stirred under hydrogen at room temperature for 25 hours after which TLC (4 DCM : 1 methanol) indicated the complete consumption of the intermediate. The reaction mixture was filtered through Celite® and the filtrand washed with water (20ml). The organic solvent was removed *in vacuo* and the solution lyophilized to give the target compound as a white amorphous solid (312mg, 81%).

R_f 0 (4 DCM : 1 methanol); ^{1}H NMR (400MHz, D_2O) δ ppm 1.30 (2H, quin, *J* 7.6Hz, CH_2), 1.53 (2H, quin, *J* 7.6Hz, $C\underline{H}_2CH_2NH_2$), 1.55 (2H, quin, *J* 7.3Hz, $C\underline{H}_2CH_2OSO_4H$), 2.84 (2H, t, *J* 7.6Hz, $C\underline{H}_2NH_2$), 3.90 (2H, t, *J* 6.2Hz, $C\underline{H}_2OSO_4H$); ^{13}C NMR (101MHz, D_2O) δ ppm 22.3 (1C, CH_2), 26.5 (1C, $\underline{C}H_2CH_2NH_2$), 28.1 (1C, $\underline{C}H_2CH_2OSO_4H$), 39.7 (1C, CH_2NH_2), 69.3 (1C, CH_2OSO_4H); FT-IR (ATR) υ 1161 (S=O), 1611 (NH_2 scissoring), 3083 (NH_2, u), 3191 (NH_2, g); HRMS (ES^-) m/z 182.0488 $[M - H]^-$ (required 182.0492).

5.11.4 Modification of quantum dots

A solution of CdSe-ZnS core-shell quantum dots (emission λ_{max} 525nm, Invitrogen, 50µl of an 8µM solution in borate buffer pH 8, 4×10^{-4} µmol, 1eq), **5.1** (0.34mg, 0.8µmol, 2000eq) and EDAC (0.77mg, 4.0µmol, 10000eq) in 20mM sodium phosphate buffer at pH 7.5 (1ml) was stirred at room temperature for 1 hour. The reaction mixture was repeatedly (x6) concentrated by centrifugal filtration through a 10kDa cut-off membrane and reconstituted in PBS buffer. The mixture was passed through a 0.2µm membrane filter to remove aggregates. The quantum dot solution was made up to 1ml with PBS and the concentration determined using previously reported procedures[45] to be 0.34µM (ε_{350} = 710000 $M^{-1}cm^{-1}$). An agarose gel was used to conclusively demonstrate quantum dot modification, and visual inspection showed a lack of aggregation in the colloidal phase (Figure 5.32). The carbohydrate loading on the quantum dots was determined using the phenol sulphuric acid

method.[46,47] An aliquot of the quantum dot solution (50µl) was treated with concentrated sulphuric acid (75µl) and aqueous phenol (5% w/w, 10µl) and heated to 90°C. After 5 minutes the sample was cooled to room temperature and A_{490} measured, referenced to a solution of carbohydrate modified quantum dots and phenol. The concentration of mannose was determined by comparison to a standardised curve. The carbohydrate content per dot was calculated from the ratio of mannose concentration to the concentration of ZnS-CdSe quantum dots and was found to be ~50 sugars/dot.

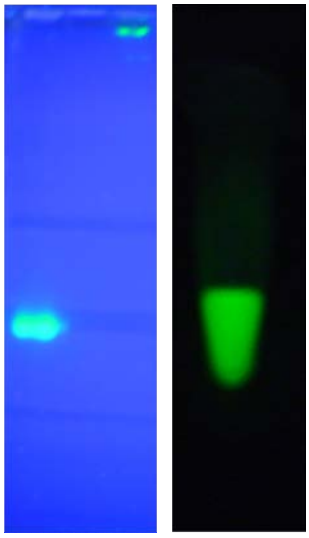


Figure 5.32. *Left: fluorescent image of 1% w/w agarose gel run at 100V for 60min. Lane 1- unmodified quantum dot solution. Lane 2-mannose modified quantum dots. Right: Sample of the carbohydrate modified CdSe-ZnS quantum dot solution under UV irradiation showing a fully homogenous solution.*

A control batch of quantum dots was prepared using propylamine instead of **5.1**, which was prepared as detailed above. The concentration was adjusted to be 0.34µM with PBS buffer.

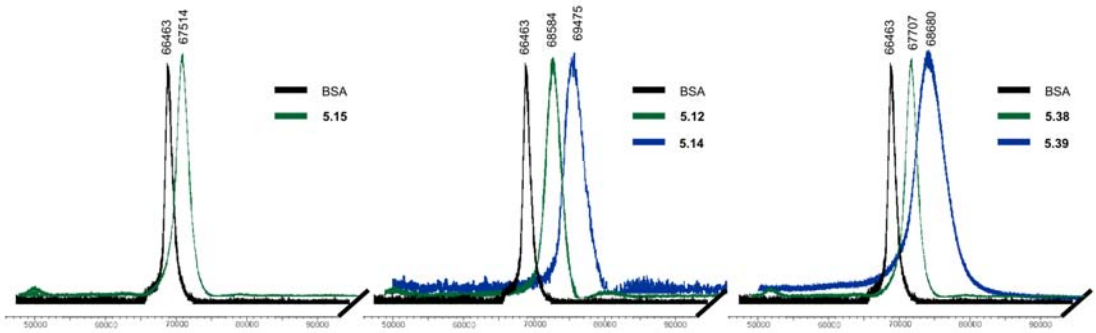
5.11.5 Modification of BSA

To a biphasic mixture of **5.1** or **5.37** (10 x 60 eq) in 0.3 M $NaHCO_3$ and 0.3 M Na_2CO_3 at pH 9 (1ml, pH paper) and chloroform (1ml) was added a solution of thiophosgene (6 x 10 x 60 eq) to the organic layer. The biphasic mixture was vigorously stirred at room temperature and monitored by mass spectrometry (ES^+ for **5.1**, ES^- for **5.37**). After 3 hours, complete consumption of the starting material and the formation of single isothiocyante products were detected. The aqueous layer was retained and washed with chloroform (5 x 1ml). Reactions were performed with 0.25-1mg of BSA (Sigma Aldrich, 60 reactive amines). The solution of isothiocyanate was added to a solution of BSA in 0.3 M $NaHCO_3$ and 0.3 M Na_2CO_3 at pH 9 (pH paper)

to a final concentration of 1mg/ml. The reaction mixture was gently shaken at room temperature for 12 hours after which, excess reagent was removed by repeated (x3) centrifugal filtration though a 10kDa cut-off membrane and the protein re-dissolved in 0.3 M $NaHCO_3$ and 0.3 M Na_2CO_3 at pH 9 (pH paper). The solution was made up to 1ml and **5.14** (0.2 x 10 x 60eq) was added. The reaction mixture was shaken for a further 12 hours, after which, the mixture was concentrated by centrifugal filtration through a 10kDa cut-off membrane. The proteins were purified and buffer exchanged into PBS with a PD G-25 MiniTrap™ size exclusion column (GE Healthcare). The protein concentrations were measured using the BCA method and adjusted to 1.0mg/ml with PBS. The protein solutions were stored at 4˚C until use.

For the synthesis of **5.16**, the initial ligand conjugations were not performed and the synthesis was initiated by reaction with **5.14** (0.2 x 10 x 60eq), using the above conditions.

The extent of modification was determined by MALDI mass spectrometry as follows. 5 μL of the reaction mixture was diluted to 250 μL with 0.1 % TFA in water. This solution was mixed in a 1 : 1 ratio (v/v) with a solution of mass spectrometry grade sinapinic acid (10 mg/mL in 1 : 1 water : acetonitrile with 0.1% TFA). From this combined matrix/ sample solution, 2 μL was spotted onto a steel target and allowed to co-crystallize at room temperature over 3 hours. MALDI analysis was conducted on a Water MALDI Micro MX spectrometer with TOF detection, in positive linear mode. Unmodified protein was used as a lock mass calibrant. For each comparative set of experiments, laser energy and pulse width were optimised for a modified protein sample, and kept identical for the remainder of the sample set. In general, multiple laser shots were combined to produce the final spectrum. Data were further processed using Mass Lynx 4.1. Mass spectra were smoothed with Savitzky-Golay smoothing prior to dispersity analysis.



5.11.6 Modification of ATTO dyes

Reactions were performed with ATTO 520, 565 and 633 NHS esters (Sigma Aldrich) using between 0.25-1mg of the dye. ATTO dyes (1eq) were dissolved in 100µl of solvent (ATTO 520 was dissolved in DMF while ATTO 565 and 633 were dissolved in 1 water : 1 methanol) and added to a solution of **5.1** or **5.37** (1.5eq) in 100mM $NaHCO_3$ buffer at pH 8.3 (400µl). The reactions were stirred at room temperature for ~13 hours and were monitored by silica gel TLC and C_{18} TLC. The solvents were removed under reduced pressure and the crude dyes were purified by automated reverse phase chromatography (Biotage C_{18} 12g SNAP cartridge, 5% acetonitrile in water to 100% water over 10 column volumes). The fluorescent fractions were pooled and the solvent removed *in vacuo* to give the modified dyes. The dyes were dissolved in PBS buffer for microscopy experiments.

5.11.7 Growth of cells

General protocol for cell culture

All cell culture experiments were performed in a laminar flow hood under strict aseptic conditions. Dulbecco's PBS refers to a solution of phosphate buffered saline without Ca^{2+} or Mg^{2+} salts. Cell counting was performed manually using ISL cell counting slides. Viability was confirmed with the Trypan Blue dye exclusion method.

Cell lines were purchase from the American Type Culture Collection and stored in vapour phase of liquid nitrogen until use. On first use, DMSO was removed by centrifugation at 100*g* for 5 minutes and the cells were then grown as described below. Stock suspensions for cryopreservation were made in the growth medium supplemented with DMSO to a 10% final concentration. Cells were harvested in the log phase of growth and suspended at ~3-4 × 10^5 cells/ml in the cryopreservation medium. Cells were frozen using the Nalgene™ Cryo 1˚C Freezing Container and transferred to liquid nitrogen for long term storage.

Growth of NR8383 cells from *Rattus norvegicus*

Alveolar macrophage derived cells from a Sprague-Dawley rat (cell line designation NR8383) were grown in F-12K medium with fetal bovine serum (Sigma Aldrich) to a final concentration of 15%. Cells were grown in T-25 tissue culture flasks for

adherent cell lines (Greiner Bio-One) with 4-5ml of growth medium at 37˚C in a humified atmosphere of 5% CO_2.

The growth medium was renewed every 3-4 days. Non-adherent cells were decanted into a centrifuge tube and pelleted by centrifugation at 200*g* for 5 minutes. The supernatant was discarded and the cell pellet re-suspended in 4-5ml of growth medium and transferred to the original cell culture flask.

The culture was split when adherent cells had reached 60-70% confluence (~ 3-4 × 10^6 total cells). Non-adherent cells were decanted into a centrifuge tube and retained. Adherent cells were washed with Dulbecco's PBS (~3ml) and the washings transferred to the centrifuge tube. Adherent cells were harvested by gentle scraping, re-suspended in Dubecco's PBS (~3ml) and transferred to the centrifuge tube. The tissue culture flask was washed with Dulbecco's PBS (~4ml) and the washings retained. The combined washings were centrifuged at 200*g* for 5 minutes. The supernatant was discarded and the cell pellet re-suspended in the growth medium. Cells were transferred to a fresh T-25 flask containing 3-4ml of growth medium, preheated to 37˚C and exposed to 5% CO_2 for ~ 5 minutes. Cells were split in a ratio between 1:3 to 1:2 to give a cell concentration of ~2-4 × 10^5 viable cells/ml, in a total volume of 4-5ml.

Growth of HK-SEP-1 cells

Endothelial liver derived cells (cell line designation SK-HEP-1) were grown in Modified Eagle's Medium with fetal bovine serum (Sigma Aldrich) to a final concentration of 10%. Cells were grown in T-25 tissue culture flasks for adherent cell lines (Greiner Bio-One) with 4-5ml of growth medium at 37˚C in a humid atmosphere of 5% CO_2.

The growth medium was renewed every 2-3 days. The spent medium was discarded by decantation and replaced with fresh medium.

The culture was split when cells had reached 80-100% confluence (~ 7-8 × 10^5 total cells). The spent medium was discarded by decantation and the adherent cells washed with Dulbecco's PBS (3 x 4ml). Cells were detached incubating at 37˚C for 1-2 minutes with a solution of Trypsin-EDTA (1ml, 1x concentration, Sigma Aldrich). Cells were transferred to a fresh T-25 flask containing 4ml of growth medium,

preheated to 37˚C and exposed to 5% CO_2 for ~ 5 minutes. Cells were split in a ratio between 1:4 to 1:3 to give a cell concentration of ~2-4 × 10^5 viable cells/ml, in a total volume of 4-5ml.

Growth of cells for microscopy

Cells were grown as above on 8-well tissue culture treated chambered coverglass (Nunc). Cells were grown in 0.3ml of growth medium at 37˚C in a humid atmosphere of 5% CO_2 for at least 3 days to allow full recovery. The final cell concentration was ~ 2-4 x 10^5 cells/ml.

For SK-HEP-1 cells, which were harvested by trypsinization, the Trypsin-EDTA solution containing suspended cells was added to growth medium (~5ml) and centrifuged at 200*g* for 5 minutes to pellet cells. The supernatant was discarded. The cell pellet was re-suspended in growth medium and then treated as above.

5.11.8 Confocal microscopy

Microscopy experiments were performed on a Leica Microsystems SP5 Inverted Confocal Microscope. All experiments were performed with the pinhole set at 1 Airy diameter. All images were captured at 512x512 resolution at 400Hz capture rates. All experiments were performed with the 20× dry objective and further magnification was achieved with the optical zoom. Fluorescein was excited at 488nm (Ar laser) and emission collected between 495-640nm. Trypan blue was excited at 633nm (He/Ne laser) and emission collected between 645-790nm. Neutral red was excited at 543nm (He/Ne laser) and emission collected between 565-790nm. QDot™ 525 was excited at 458nm (Ar laser) and emission collected between 465-620nm. Experiments with multiple dyes were performed using sequential excitation to avoid optical cross-talk. Images requiring comparison between experiments were run with identical laser irradiance and photomultiplier tube gain settings.

Generally, neutral red was added to a final concentration of 5μM and trypan blue at 150μM. BSA-fluorescein conjugates were used at a final concentration of 1.7μM. Modified quantum dots were used at a final concentration of 40nM. ATTO dyes for FRET experiments were added to a final concentration of 2μM. Cells were maintained at 37˚C and under an atmosphere of 5% CO_2 during imaging.

Images were processed with the Leica LAS or ImageJ. Deconvolution, where used, was performed with the Biochemical Image Group DeconvolutionLab plugin for ImageJ,[48] using theoretical point spread functions determined using the Biochemical Image Group PSF Generator plugin for ImageJ.[49]

5.11.9 FRET microscopy

FRET imaging was preformed using the Leica FRET Sensitized Emission wizard. For ATTO 520/ ATTO 565 FRET experiments, ATTO 520 was excited at 514nm (Ar laser) and emission collected between 523-572nm, while ATTO 565 was excited at 543nm (He/Ne laser) and emission collected between 573-710nm. For ATTO 565/ ATTO 633 FRET experiments, ATTO 565 was excited at 543nm (He/Ne laser) and emission collected between 547-637nm, while ATTO 633 was excited at 633nm (He/Ne laser) and emission collected between 638-800nm. The donor and acceptor were sequentially excited in a line-by-line mode, as controlled by the wizard. Donor-only and acceptor-only images were taken to generate correction factors accounting for cross-excitation and cross-talk. FRET images were gathered under an identical optical set-up as for the control images. Apparent FRET efficiencies were determined using these correction factors according to previously reported procedures.[36] Images were displayed in an RGB tricolour LUT with Leica LAS imaging software. Histogram analysis and pixel counting was performed with ImageJ.

5.12 References

1. (a) N. Jayaraman, *Chem. Soc. Rev.*, 2009, **38**, 3463. (b) J. J. Lundquist, E. J. Toone, *Chem. Rev.*, 2002, **102**, 555. (c) M. Mammen, S. Choi, G. M. Whitesides, *Angew. Chem. Int. Ed.*, **1998**, 37, 2754.
2. D. S. Roseman, J. U. Baenziger, *Proc. Natl. Acad. Sci.*, **2000**, 97,9949.
3. D. S. Roseman, J. U. Baenziger, *J. Biol. Chem.*, **2001,** 276, 17052.
4. (a) P. R. Taylor, S. Gordon, L. Matinez-Pomares, *Trends Immunol.*, **2005**, 26, 104. (b) L. Matinez-Pomares, S. A. Linehan, P. R. Taylor, S. Gordon, *Immunobiol.*, **2001**, 204, 527.
5. D. S. Roseman, J. U. Baenziger, *Methods Enzymol.*, **2003**, 363, 121.
6. P. D. Stahl, R. A. B. Ezekowitz, *Curr. Opin. Immunol.*, **1998**, 10, 50.

7. (a) H. Feinberg, S. Park-Snyder, A. R. Kolatkar, C. T. Heisse, M. E. Taylor, W. I. Weis, *J. Biol. Chem.*, **2000**, 275, 21539. (b) B. L. Largent, K. M. Walton, C. A. Hoppe, Y. C. Lee, R. L. Schnaar, *J. Biol. Chem.*, **1984**, 259, 1764. (c) M. E. Taylor, J. T. Conary, M. R. Lennarz, M. B. Lennartz, P. D. Stahl, K. Drickamer, *J. Biol. Chem.*, **1990**, 265, 12156. (d) R. A. B. Ezekowitz, K. Sastry, P. Bailly, A. Warner, *J. Exp. Med.*, **1990**, 172, 1785.
8. (a) M. E. Taylor, *Glycobiology*, **1997**, 7, v. (b) K. Drickamer, *Curr. Opin. Struct. Biol.*, **1993**, 3, 393.
9. M. E. Taylor, K. Drickamer, *J. Biol. Chem.*, **1993**, 268, 399.
10. C. Leteux, W. Chai, R. W. Loveless, C. Yuen, L. Uhlin-Hansen, Y. Combarnous, M. Jankovic, S. C. Maric, Z. Misulovin, M. C. Nussenzweig, T. Feizi, *J. Exp. Med.*, **2000**, 191, 1117.
11. Y. Liu, A. J. Chirino, Z. Misulovin, C. Leteux, T. Feizi, M. C. Nussenzweig, P. J. Bjorkman, *J. Exp. Med.*, **2000**, 191,1105.
12. Y. Liu, Z. Misulovin, P. J. Bjorkman, *J. Mol. Biol.*, **2001**, 305, 481.
13. J. Lai, O. K. Bernhard, S. G. Turville, A. N. Harman, J. Wilkinson, A. L. Cunningham, *J. Biol. Chem.*, **2009**, 284, 11027.
14. (a) X. Michalet, A. N. Kapanidis, T. Laurence, F. Pinaud, S. Doose, M. Pflughoefft, S. Weiss, *Annu. Rev. Biophys. Biomol. Struct.*, **2003**, 32, 161. (b) J. W. Lichtman, J. Conchello, *Nature Methods*, **2005**, 2, 910. (c) R. Pepperkok, J. Ellenberg, *Mol. Cell. Bio.*, **2006**, 7, 690.
15. (a) D. M. Shotton, *J. Cell Sci.*, **1989**, 94, 175. (b) D. Axelrod, *Traffic*, **2001**, 2, 764.
16. (a) P. Tinnefeld, M. Sauer, *Angew. Chem. Int. Ed.*, **2005**, 44, 2642. (b) C. Joo, H. Balci, Y. Ishitsuka, C. Buranachai, T. Ha, *Annu. Rev. Biochem.*, **2008**, 77, 51. (c) D. Pappas, S. M. Burrows, R. D. Reif, *Trend. Anal. Chem.*, **2007**, 26, 884.
17. (a) D. W. Piston, G. Kremers, *Trends Biochem. Sci.*, **2007**, 32, 407. (b) J. Lippincott-Schwartz, E. Snapp, A. Kenworthy, *Mol. Cell. Bio.*, **2001**, 2, 444. (c) F. S. Wouters, P. J. Verveer, P. I. H. Bastiaens, *Trends Cell. Biol.*, **2001**, 11, 203. (d) P. I. H. Bastiaens, R. Pepperkok, *Trends Biochem. Sci.*, **2000**, 25, 631.
18. R. B. Sekar, A. Periasamy, *J. Cell. Biol.*, **2003**, 160, 629.

19. D. M. Ojcius, F. Niedergang, A. Subtil, R. Hellio, A. Daustry-Varsat, *Res. Immunol.*, **1996**, 147, 175.
20. (a) A. K. Kenworthy, *Methods*, **2001**, 24, 289. (b) R. M. Murphy, H. Slayter, P. Schurtenburger, R. A. Chamberlin, C. K. Colton, M. L. Yarmush, *Biophys. J.*, **1988**, 54, 45.
21. (a) T. W. Gadella, T. M. Jovin, *J. Cell. Biol.*, **1995**, 129, 1543. (b) Y. Sako, S. Minoghchi, T. Yanagida, *Nature Cell. Biol.*, **2000**, 2, 168.
22. S. Angers, A. Salahpour, M. Bouvier, *Annu. Rev. Pharmacol. Toxicol.*, **2002**, 42, 409.
23. D. Z. Simpson, P. G. Hitchen, E. L. Elmhirst, M. E. Taylor, *Biochem. J.*, **1999**, 343, 403.
24. H. J. Schuster, B. Vijayakrishnan, B. G. Davis, *unpublished result*.
25. (a) A. P. Alivisatos, W. Gu, C. Larabell, *Annu. Rev. Biomed. Eng.*, **2005**, 7, 55. (b) J. M. Klostranec, W. C. W. Chan, *Adv. Mater.*, **2006**, 18, 1953. (c) A. M. Smith, H. Duan, A. M. Mohs, S. Nie, *Adv. Drug Del. Rev.*, **2008**, 60, 1226. (d) M. De, P. S. Ghosh, V. M. Rotello, *Adv. Mater.*, **2008**, 20, 4225. (e) I. L Medintz, H. T. Uyeda, E. R. Goldman, H.Mattoussi, *Nature Mater.*, **2005**, 4, 435. (f) S. J. Rosenthal, J. C. Chang, O. Kovtun, J. R. McBride, I. D. Tomlinson, *Chem. Biol.*, **2011**, 18, 10.
26. (a) T. Masuko, A. Minami, N. Iwasaki, T. Majima, S. Nishimura, Y. C. Lee, *Anal. Biochem.*, **2005**, 339, 69. (b) P. Rao, T. N. Pattabiraman, *Anal. Biochem.*, **1989**, 181. 18.
27. S. A. Altman, L. Randers, G. Rao, *Biotechnol. Prog.*, **1993**, 9, 671.
28. L. W. Zhang, N. A. Monteiro-Riviere, *Toxicol. Sci.*, **2009**, 110, 138.
29. B. A. Rzigalinski, J. S. Strobl, *Toxicol. Appl. Pharmacol.*, **2009**, 238, 280.
30. (a) G. T. Hermanson, *Bioconjugate Techniques*, Academic Press, 2nd edn., 1996. (b) M. Izumi, S. Okumura, H. Yuasa and H. Hashimoto, *J. Carbohydr. Chem.*, 2003, **22**, 317. (c) S. I. Van Kasteren, P. Garnier, B. G. Davis, *Chemical Methods for Mimicking Post-Translational Modifications*, in Nucleic Acids and Molecular Biology, **2009**, Springer.
31. C. Deduve, R. Wattiaux, *Ann. Rev. Physiol.*, **1966**, 28, 435.
32. S. C. Silverstein, R. M. Steinman, Z. A. Cohn, *Ann. Rev. Biochem.*, **1977**, 46, 669.

33. Vandana, O. Hindsgaul, J. U. Baenziger, *Can. J. Chem.*, **1987**, 65, 1645.
34. (a) Y. Sun, H. Wallrabe, S. Seo, A. Periasamy, *ChemPhysChem*, **2011**, 12, 462. (b) E. A. Jares-Erijman, T. M. Jovin, *Nat. Biotechnol.*, **2003**, 21, 1387.
35. (a) C. Berney, G. Danuser, *Biophys. J.*, **2003**, 84, 3992. (b) J. N. Feige, D. Sage, W. Wahli, B. Desvergne, L. Gelman, *Microsc. Res. Techniq.*, **2005**, 68, 51.
36. J. van Rheenen, M. Langeslag, K. Jalink, *Biophys. J.*, **2004**, 86, 2517.
37. T Matsumura, M Takesue, K. A. Westerman, T. Okitsu, M. Sakaguchi, T. Fukuzawa, T. Totsugawa, H. Nogughi, S. Yamamoto, D. B. Stolz, N. Tanaka, P. Leboulch, N. Kobayashi, *Transplantation*, **2004**, 77, 1357.
38. (a) A. Weiss, J. Schlessinger, *Cell*, **1998**, 94, 277. (b) C. Heldin, *Cell*, **1995**, 80, 213.
39. S. K. Maity, S. K. Dutta, A. K. Bannerjee, B. Acharee, S. M. Basudeb, *Tetrahedron*, **1994**, 50, 6965.
40. M. K. Patel, B. Vijayakrishnan, J. R. Koeppe, J. M. Chalker, K. J. Doores, B. G. Davis, *Chem. Commun*, **2010**, 46, 9119.
41. M. Kiso, L. Anderson, *Carbohydr. Res.*, **1985**, 136, 309.
42. R. I. El-Sokkary, B. A. Silwanis, M. A. Nashed, H. Paulsen, *Carbohydr. Res.*, **1990**, 203, 319.
43. R. I. El-Sokkary, B. A. Silwanis, M. A. Nashed, H. Paulsen, *Carbohydr. Res.*, **1990**, 203, 319.
44. L. Wang, C. Li, Q. Wang, Y. Hui, *J. Chem. Soc., Perkin Trans. 1*, **1994**, 6, 621.
45. C. A. Leatherdale, W. K. Woo, F. V. Mikulec, M. G. Bawendi, *J. Phys. Chem. B*, **2002**, 106, 7619.
46. T. Masuko, A. Minami, N. Iwasaki, T. Majima, S. Nishimura, Y. C. Lee, *Anal. Biochem.*, 2005, 339, 69.
47. P. Rao, T. N. Pattabiraman, *Anal. Biochem.*, **1989**, 181, 18.
48. C. Vonesch, M. Unser, *IEEE Trans. Image Process.*, **2008**, 17, 539.
49. A. Griffa, N. Garin, D. Sage, *G.I.T. Imaging & Microscopy*, **2010**, 1, 43.

Chapter VI – Characterization of synthetic glycoproteins

6.1 Introduction

The presentation of carbohydrates on proteins and cell surfaces has far reaching biological significance.[1] Carbohydrates act as markers for critical signalling events and trafficking pathways, which are mediated by interaction with sugar binding proteins known as lectins. While lectins can have high specificity for their ligands, their binding affinity is low. Nature utilizes a high surface density of carbohydrates for multivalent interactions to improve binding.[2] Synthetic glycoproteins that mimic this high surface coverage can interact with these lectins and are useful for clinical applications.[3] Synthetic glycoconjugates have been used as vaccines,[4] for treatment of lysosomal storage diseases,[5] for targeted drug delivery[6] and as vehicles for gene therapy.[7]

The synthesis of pure, well defined glycoproteins is ideal. Introduction of a reactive tag, followed by chemoselective conjugation would allow complete control of ligand placement.[8] Multivalency can be obtained by using either dendrimeric scaffolds[9] or multiple tags.[10] However, a more widespread and pragmatically useful alternative is the non site-selective modification of many functional groups, such as at the nucleophilic primary amine on native lysine residues and the *N*-terminus. Numerous linkage methodologies based on this principle have been reported.[11] Since lysine has a high natural abundance in proteins, multiple modifications occur. This procedure is easier to implement since it intrinsically produces multivalent conjugates while avoiding the need for time-consuming mutagenesis.

A striking drawback to this technique is the formation of a statistical mixture of products and the experimental challenge of both measuring and controlling this product dispersity. Conjugated proteins are usually characterized with matrix-assisted laser desorption/ ionization (MALDI) mass spectrometry, but for large molecular weights the resolving power is insufficient to distinguish between the protein products.[12] Hence, a broad peak is observed, the centre of which is the mean copy number of ligands per protein. Information on the dispersity of the sample is typically neglected. This is surprising since these conjugated proteins are often used directly in pre-clinical and clinical trials and it would seem vital to fully characterize the final

products, not least to ensure reproducibility between batches,[13] but also to fully understand both protein chemical reactivity and therapeutic structure-activity relationships.

At present, there are no simple methods to determine sample dispersity. Top-down proteomics may reveal possible modification sites, but the product distribution is not readily apparent and not yet amenable to high throughput analysis.[14] Fourier transform- ion cyclotron resonance[15] offers high resolution and may be capable of directly measuring product distributions.[16] While this technique is promising, the deconvolution of data is complicated and often requires extensive computation.[17] In addition, the expensive instrumentation and necessity for optimization of detector conditions restricts widespread use. A method for extracting dispersity information form MALDI data would be more straightforward, since such spectra are already gathered for the basic characterization of glycoconjugates.

In this chapter, methodology for determining the dispersity of a synthetic glycoprotein sample from MALDI data is developed. Initially, we explore the expected product distributions of indiscriminate protein modifications from a theoretical standpoint. This allows the development of a simplified method to extract dispersity information from the broadened MALDI peaks of synthetic glycoproteins. The assumptions for this methodology are tested on computed and experimental models and applied for the analysis of mannose-bearing protein conjugates. These dispersity data are analyzed to reveal the key trends in the reactions of carbohydrates with proteins.

6.2 The Gaussian approximation

Initially, protein modification reactions were explored theoretically to determine the expected product distributions. Such reactions can be modelled as a series of competitive consecutive second order reactions of the type:

$$\begin{aligned} A + R &\xrightarrow{k_1} B \\ B + R &\xrightarrow{k_2} C \\ C + R &\xrightarrow{k_3} D \\ \vdots \quad &\vdots \qquad\quad \vdots \end{aligned}$$

The associated rate equations for these reactions can be computationally solved to obtain the expected product distribution. For a protein, the consecutive reaction rate constants would be expected to fall in proportion with the number of available reactive sites. Under these ideal conditions, the product distribution was found to be exactly reproduced by a binomial probability distribution function* (Figure 6.1).

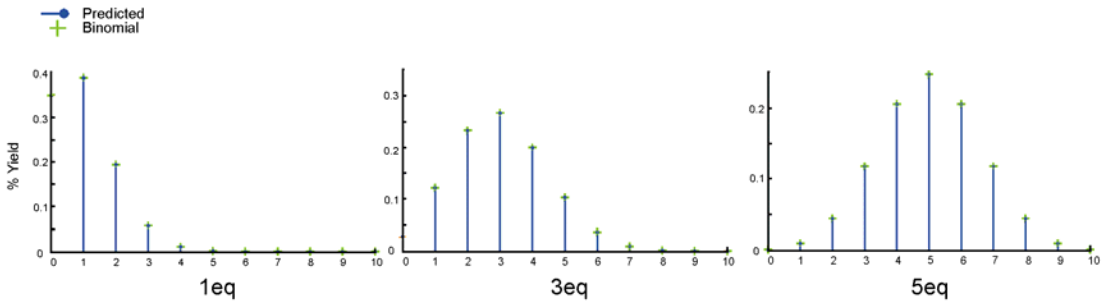


Figure 6.1. *The computed product distribution of a hypothetical protein with 10 reactive sites treated with varying equivalents of a reagent, compared against binomial distributions where n =10 and p = the number of equivalents of reagent/ 10. The data shows that the predicted distribution is exactly given by a binomial distribution.*

Real proteins, however, do not necessarily follow this ideal model which neglects structural effects and assumes the rate constants are solely dependent on the number of free lysines. For the ideal model, the binomial distribution offers an exact analytical solution, but this is not the case when the rate constants are not purely defined by statistics. For instance, a reaction on BSA, which contains 60 reactive amine sites (59 lysine residues and the *N*-terminus), where the consecutive rate constants are hypothesized to be identical would give a hypergeometrical distribution of protein products for which the binomial distribution is a poor match (Figure 6.2).

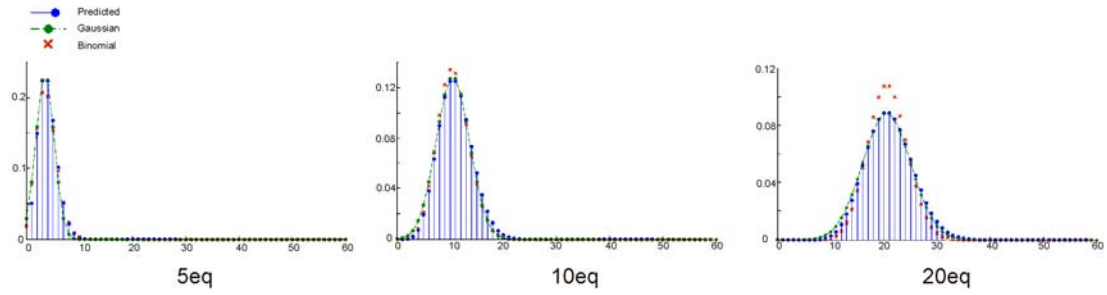


Figure 6.2. *The product distribution for a hypothetical reaction of BSA with varying equivalents of a reagent, compared against best-fit Gaussian and binomial distributions. Data shows that the Gaussian distribution is a good fit under all conditions, but the binomial distribution provides a poor fit at higher equivalents,*

* Binomial distribution $P = \binom{n}{k} p^k (1-p)^{n-k}$

In contrast, a Gaussian distribution[†] offers a very good approximation to the product distribution under a wide range of criteria for the rate constants. This arises because of the mathematical flexibility offered by the Gaussian function – both the mean and peak width can be independently set through μ (the mean) and σ (the standard deviation), and the function can be easily modified to represent the underlying distribution. Thus, even if the actual product distribution for a protein modification is unknown, the Gaussian function could be adapted to provide an excellent approximation. Crucially, the intuitive description of sample spread through σ makes the Gaussian distribution an ideal choice for analyzing the dispersity of glycoproteins.

6.3 Analysis of MALDI peak broadening

The creation of multiple protein adducts results in a MALDI mass spectrum peak with an increased average mass and a wider width. These broadened peaks can, in principle, reveal valuable information on the underlying distribution of the sample (Figure 6.3). The MALDI mass spectrum of an unmodified protein can be approximated as a Gaussian peak which can be mathematically expressed from its full width at half maximum (FWHM). Thus the overall shape is given by:

$$\frac{1}{\sqrt{2\pi\delta^2}} e^{\frac{-x^2}{2\delta^2}} \qquad (1)$$

where *x* is the measured mass and δ is the standard deviation of the peak for the unmodified protein, obtained from the FWHM.[‡] The broadened peak of the protein after modification is also approximately Gaussian and can be similarly represented as:

$$f(x) = \frac{1}{\sqrt{2\pi\lambda^2}} e^{\frac{-(x-\mu)^2}{2\lambda^2}} \qquad (2)$$

Where μ is the relative mean mass of the conjugate proteins and λ is the standard deviation of the peak for the conjugates.[‡]

Indiscriminate modification produces a mixture of protein products, each of which generates its own Gaussian "peaklet". Inadequate resolution leads to peaklet

[†] Gaussian distribution $P = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

[‡] The FWHM and standard deviation are related as: $FWHM = \text{standard deviation} \times 2\sqrt{2\ln 2}$

coalescence producing this widened peak, and therefore the mass spectrum of the modified protein can also be represented as a sum of smaller peaklets. However, the intensity of each peaklet is not identical, but should parallel the distribution of the sample.

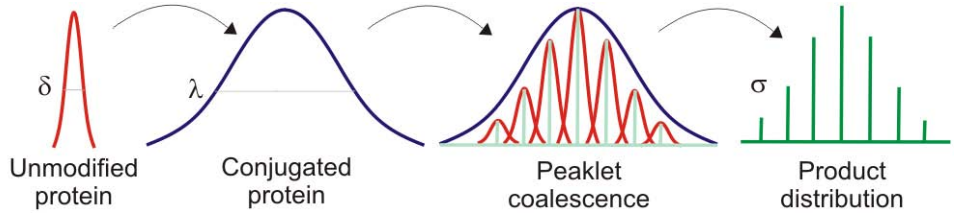


Figure 6.3. *General principle for determining the dispersity of the underlying product distribution.*

Hence, the peaklets themselves should be weighted by a probability distribution function that is representative of the underlying product mixture. As discussed above, the Gaussian distribution offers an excellent approximation and is ideally suited for this purpose (section 6.2, page 282). Hence, assuming the standard deviation of each peaklet stays constant as δ and the ionization efficiency is unaltered by successive modifications, we can write:

$$f(x) = \sum_{0}^{n}\left[\frac{1}{\sqrt{2\pi\delta^2}} e^{\frac{-(x-mn)^2}{2\delta^2}} \int_{n-0.5}^{n+0.5} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx\right] \qquad (3)$$

where *m* is the mass change per modification, *n* is the total number of reactive sites per protein, and σ is the standard distribution of the product distribution. The integration from n – 0.5 to n + 0.5 is a continuity correction to account for discrepancy between the continuous Gaussian distribution and the discrete nature of protein modifications.

Equations (2) and (3) represent different analyses of the same function: that of the broadened peak of the synthetic conjugate mixture. By equating the two formulae, the value of σ can be computationally solved. Thus, by simply measuring the FWHM of modified and unmodified protein, it is possible to calculate the standard deviation of the underlying product distribution. Conveniently, the area of a Gaussian function between $\pm 2\sigma$ equates to 95% of the total, and so $\mu \pm 2\sigma$ represents the range within which 95% of the protein adducts exist thus giving a direct and physically intuitive measure of the sample dispersity.

Alternatively, deconvolution could be used to determine the dispersity of the sample. Deconvolution is widely used in the fields of spectroscopy and microscopy to recover information that is lost during data acquisition. For example, in microscopy, deconvolution is used to sharpen images that are blurred due to inadequacies is the optical setup. Similarly, spectral deconvolution can be used to restore absorption patterns from multiple overlapping spectra. An analogous method could be applied to the above MALDI spectra. The spectra of the unmodified protein essentially provides a function representing the maximum resolving power of the spectrometer for a single pure protein. The conjugated protein spectrum can be viewed as arsing from the convolution of this function with the product distribution. Hence, with the spectra for the conjugated and pure proteins, deconvolution could be used to determine the exact product distribution. Our technique is more approximate than a deconvolution, but only requires the FWHMs as inputs (rather than the entire spectrum) and is arguably a simpler method to implement.

6.4 Validation of the methodology

This methodology is underpinned by various assumptions whose validity required testing prior to real-world application. Crucially, our model assumes that the ionization efficiency is unaltered upon single stepwise modifications, and that the MALDI peak for the unmodified protein is of Gaussian shape and is a suitable representation for all other peaklets in the protein mixture.

To validate these assumptions, we analysed previous MALDI data from synthetic Qβ glycoproteins.[18] In these cases, since the Qβ monomer is a low molecular weight protein (14 kDa, 8 reactive sites), the individual peaklets are well resolved providing a useful model for testing our assumptions (Figure 6.4).

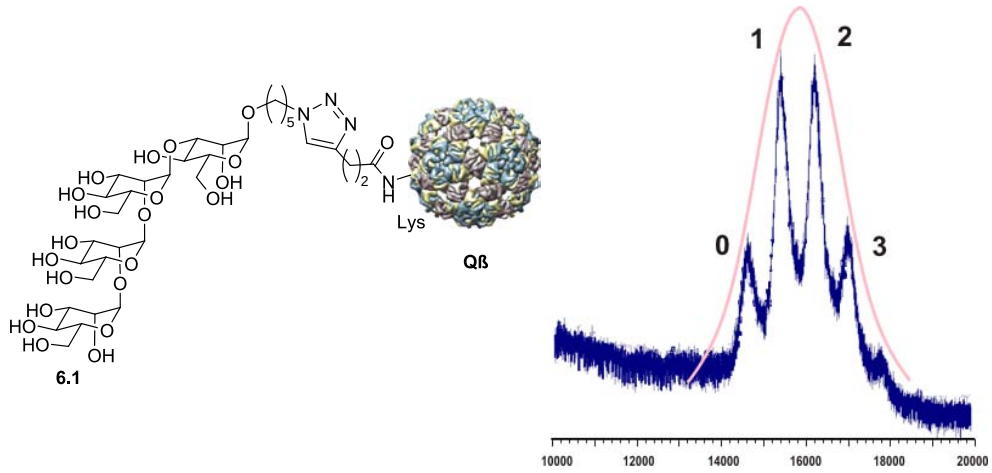


Figure 6.4. *Left: structure of Qβ glycoprotein; right: MALDI mass spectrum of the singly charged adduct. Individual peaklets can be resolved and demonstrate that the variance is broadly unchanged upon modifications and that the product distribution is approximated by a Gaussian function,*

This data shows that individual peaklets clearly have an overall Gaussian shape with identical FWHMs, lending support to our use of just the unmodified protein peak to encapsulate the characteristics of all peaklets. Furthermore, the distribution profile exactly matches the theoretical expectation indicating similar ionisation of the different protein products. This identical ionizabilty of differentially glycosylated proteins has also been witnesses by other researchers,[19] further demonstrating the validity of our assumptions.

To confirm the overall effectiveness of our dispersity analysis, we conducted a computational simulation of a protein modification reaction on BSA using Gaussian shaped peaklets to compare the expected dispersity with that derived from our calculation. Furthermore, real MALDI spectra often exhibit some high mass peak tailing caused by in source decay-desolvation of protein-matrix cluster ions, as well as the formation of cation and matrix photochemical adducts.[20] To probe the effect of this peak tailing on the dispersity calculation, a simulated reaction was performed with a Gumbel function to replicate peaklet asymmetry (Figure 6.5).

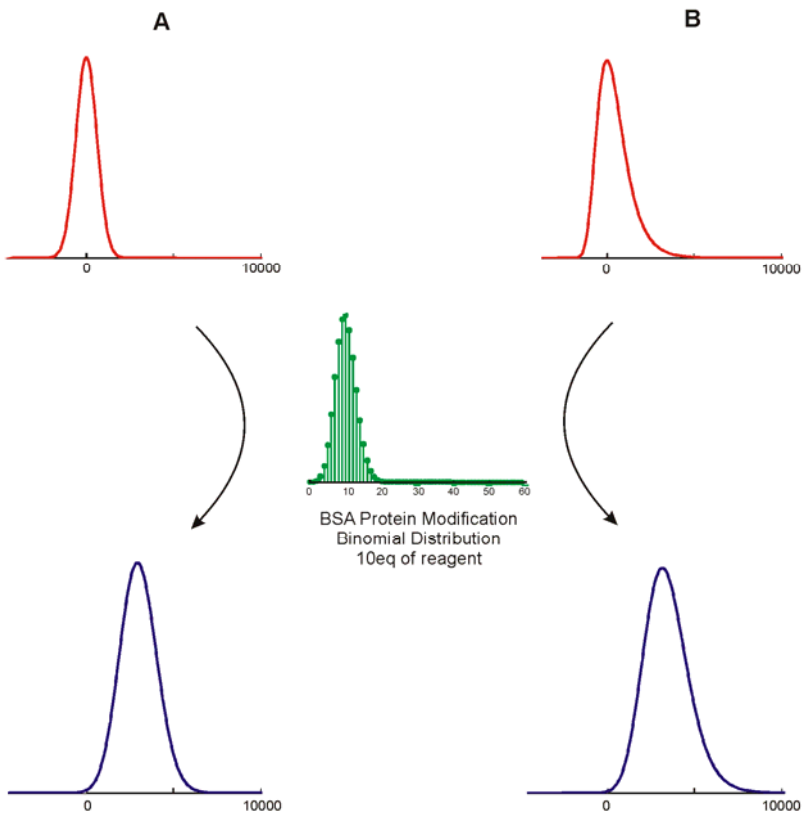


Figure 6.5. *Hypothetical reaction on BSA with 10eq of reagent simulated to give a binomial distribution of products with Gaussian (A) or Gumbel (B) peaklet shapes.The simulated product MALDI peaks (below) were used to test effect of peak tailing on the dispersity calculation.*

In both cases, the protein modification reaction is simulated to produce a binomial distribution of products. Hence, we can say the true dispersity (2σ) of the products is 5.77,[§] which can be compared to the calculated dispersity from the broadened MALDI peaks of the modified protein. For the Gaussian peaklet function, which represents a theoretical ideal, the analysis gives a dispersity of 5.74 thus reproducing the true dispersity very well and validating our core model. For the second, more realistic case, the peak tailing for the unmodified protein induces some asymmetry in the modified protein peak. If the FWHMs are determined as normal, our analysis gives a dispersity of 5.86. Hence, even with significantly skewed peaks, the dispersity analysis produces only a small deviation from the true value. This is not too surprising since the peak tailing serves to increase the FWHM of *both* the unmodified and the modified protein peaks, and is therefore, to an extent, self compensating.

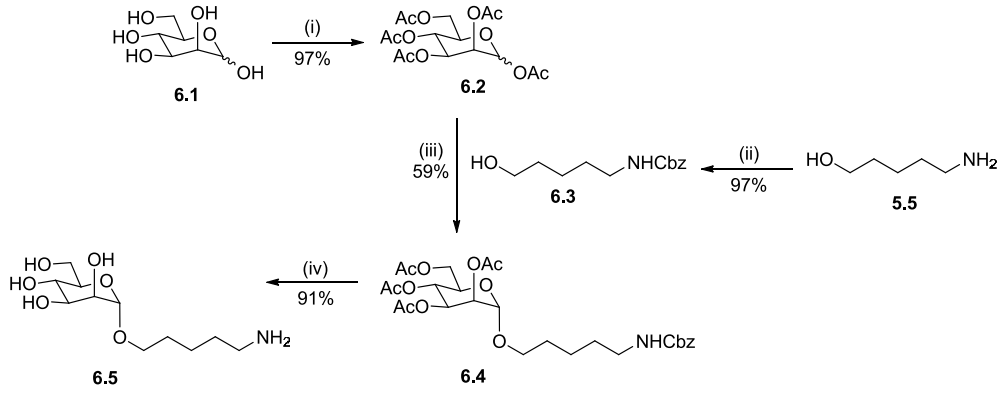
[§] The standard deviation of the binomial distribution is given by $\sqrt{np(1-p)}$

For these reasons, we believe the procedure is robust and likely to perform well under a wide range of experimental condirions. While MALDI conditions should be screened for the best possible peak shapes, a small amount of tailing can be tolerated.

6.5 Dispersity analysis on synthetic glycoproteins

To demonstrate the versatility of this analysis, various glycoconjugates were synthesized and their dispersity determined using MALDI with time of flight (TOF) detection. Mannose-bearing synthetic glycoproteins can lead to potential HIV vaccines due to mimicry of the glycosylation on envelope protein gp120.[3,4] As a result, the conjugation of mannose residues onto immunogenic carrier proteins is an area of active current research and would be well served by our dispersity analysis.[4,18,21,22] Thus efforts were focused on analogous model systems.

Carbohydrates bearing anomerically linked amine terminated linkers were synthesized for protein attachment using thiophosgene conjugation chemistry; a strategy previously explored for synthetic vaccine constructs.[23] Initial attention was directed towards a simplified monosaccharide system (Scheme 6.1).



Scheme 6.1. *Synthesis of a mannose linker for modification of proteins. Conditions: (i) Ac_2O, Py; (ii) CbzCl, Na_2CO_3, DCM, H_2O, RT; (iii) $SnCl_4$, CH_3CN, RT, 19h; (iv) NaOMe, MeOH then H_2, Pd/C, CF_3CH_2OH, H_2O + 1% HCOOH.*

The per-*O*-acetylation of mannose **6.1** with acetic anhydride in pyridine gave **6.2** in an excellent 97% yield. This was glycosylated with **6.3**, which was synthesized by the carboxybenzyl protection of **5.5** in 97% yield, with $SnCl_4$ in acetonitrile to afford the protected monosaccharide **6.4** in a reasonable 59% yield. Deprotection with sodium methoxide followed by hydrogenation gave the target compound **6.5** in 91% yield.

BSA is commonly used for the development of vaccine candidates, but with 60 reactive sites (59 lysine residues and the N-terminus), non site-selective modifications produce a wide mixture of products. Therefore synthetic BSA glycoproteins are ideally suited for analysis with our dispersity calculation and were chosen as a model platform for testing our methodology. Initially, BSA was conjugated with increasing concentrations of **6.5** per lysine residue and the products analysed by MALDI mass spectrometry. Detector settings, laser irradiance and matrix preparation were kept identical for each experiment to ensure consistency (Figure 6.6)

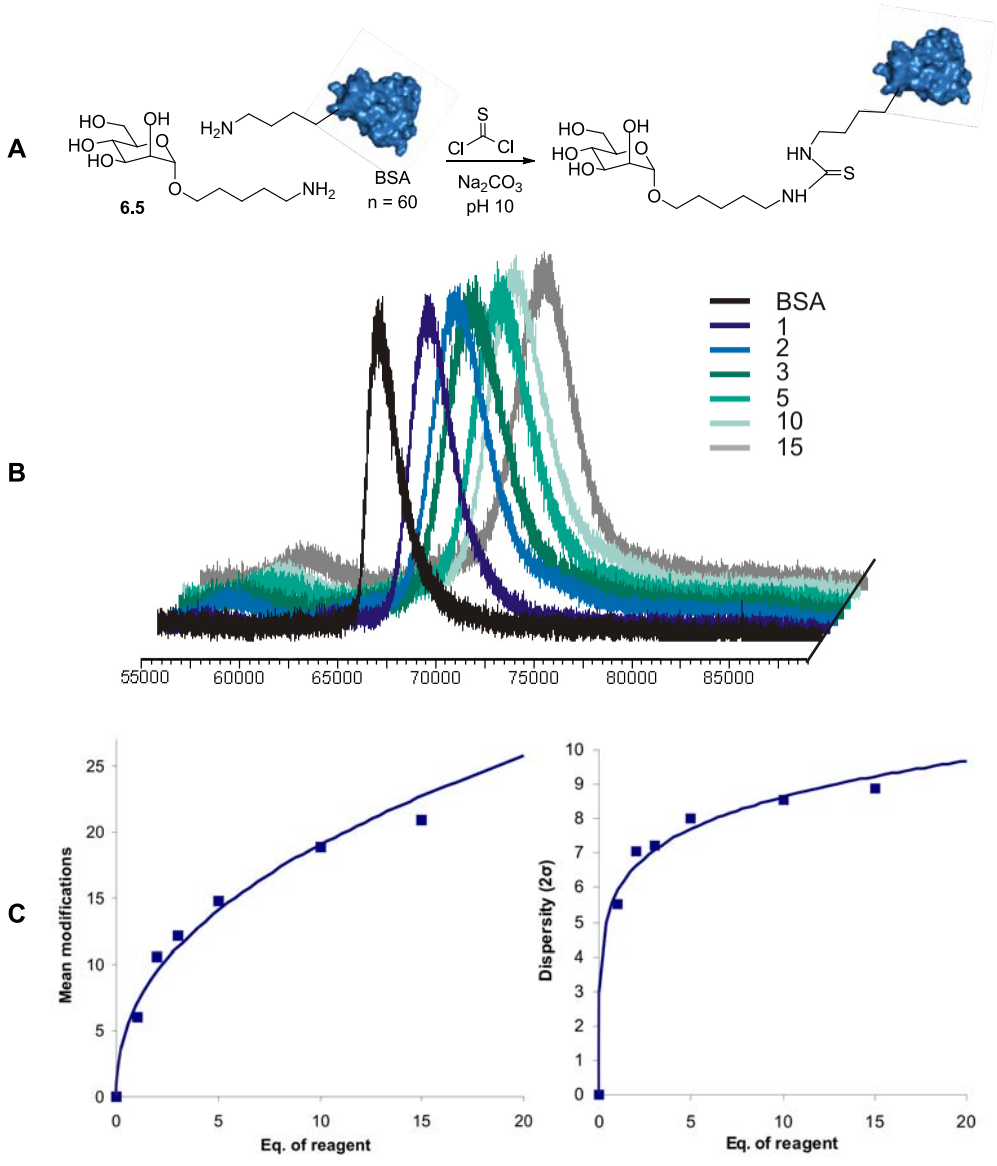


Figure 6.6. *(A) Modification reaction on BSA with increasing concentrations of* ***6.5*** *per lysine. (B) Raw MALDI data highlighting the singly charged protein peak. (C) The extent of modification and dispersity revealed from broadened MALDI peaks. The data were well represented as power curves.*

The analysis of MALDI peak broadening revealed the range of carbohydrate copy numbers present within the sample. For instance, BSA modified with a mean of 15 copies of **6.5** has a dispersity (2σ) of 8. Thus, 95% of the protein products incorporate between 7 to 23 copies of **6.5**. This is a more informative description of the actual product composition than simply stating the mean. As expected from a statistical rationale, the sample dispersity increased with the extent of protein modification.

We then investigated conjugations of BSA with the previously synthesized mannose disaccharide **5.1** as a more realistic example of a vaccine candidate.[**] As before, reactions were performed with varying concentrations of sugar and analysed by MALDI mass spectrometry (Figure 6.7).

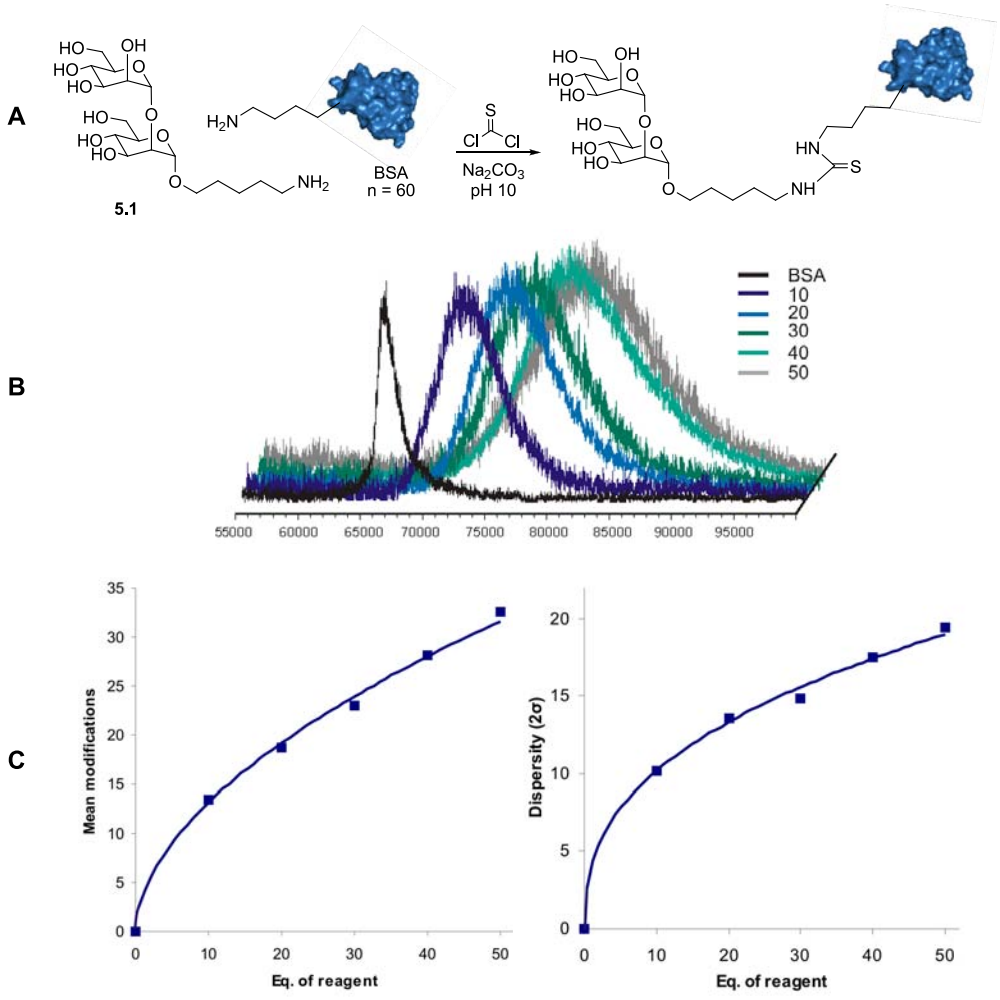


Figure 6.7. *(A) Modification reaction on BSA with increasing concentrations of* ***5.1*** *per lysine. (B) Raw MALDI data highlighting the singly charged protein peak. (C) The extent of modification and dispersity revealed from broadened MALDI peaks. The data were well represented as power curves.*

[**] For the work conducted in this chapter, compound **5.1** was synthesized by Dr Balakumar Vijayakrishnan, University of Oxford, UK. For fulI details, see M. K. Patel, B. Vijayakrishnan, J. R. Koeppe, J. M. Chalker, K. J. Doores, B. G. Davis, *Chem. Commun.*, **2010**, 46, 9119.

The greater steric demand of **5.1** required a larger excess of reagent to achieve the same levels of modification as **6.5**. Interestingly, the FWHM analysis showed an unexpected trend in the dispersity of the two different conjugates.

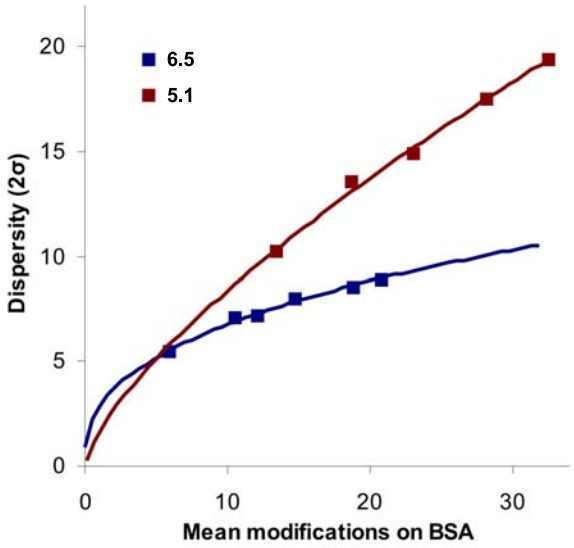


Figure 6.8. *Comparison of the dispersity between BSA conjugates with* ***6.5*** *and* ***5.1****. Data shows that modification with the disaccharide has a higher dispersity than with the monosaccharide for the same levels of modification on BSA.*

The dispersity is determined by the ratios of the consecutive rates constants (the so-called distribution constants).[24] The average rate constant of each successive modification is expected to decrease due to the reducing number of available lysines and increasing steric crowding from the protein structure. Thus, the rate of this decrease determines the dispersity of the sample, with a sharper fall resulting in a less dispersed product distribution. Since **5.1** is bulkier than **6.5**, a slower reaction rate would be expected. However, the ratios between successive rate constants are the critical determinants in dictating product dispersity, and these may not be affected to the same extent. If earlier rates in the sequence are retarded to a greater extent than later ones, the progressive decrease in reaction rates is much slower leading to a more dispersed product distribution. The overall outcome revealed that the use of reagent **6.5** compared to **5.1** appears to have a greater effect on earlier lysine couplings than on those with later, less solvent exposed residues which are already less reactive.

We then sought to expand the methodology to other protein conjugates. The cross reacting material fragment diphtheria toxin (CRM-197) is commonly used in vaccine constructs and was chosen for modification with **5.1**. Significantly, the protein would

give insights into the effect of the number of reactive sites on product dispersity. Thus, CRM[††] was modified with **5.1** and analysed as before (Figure 6.9).

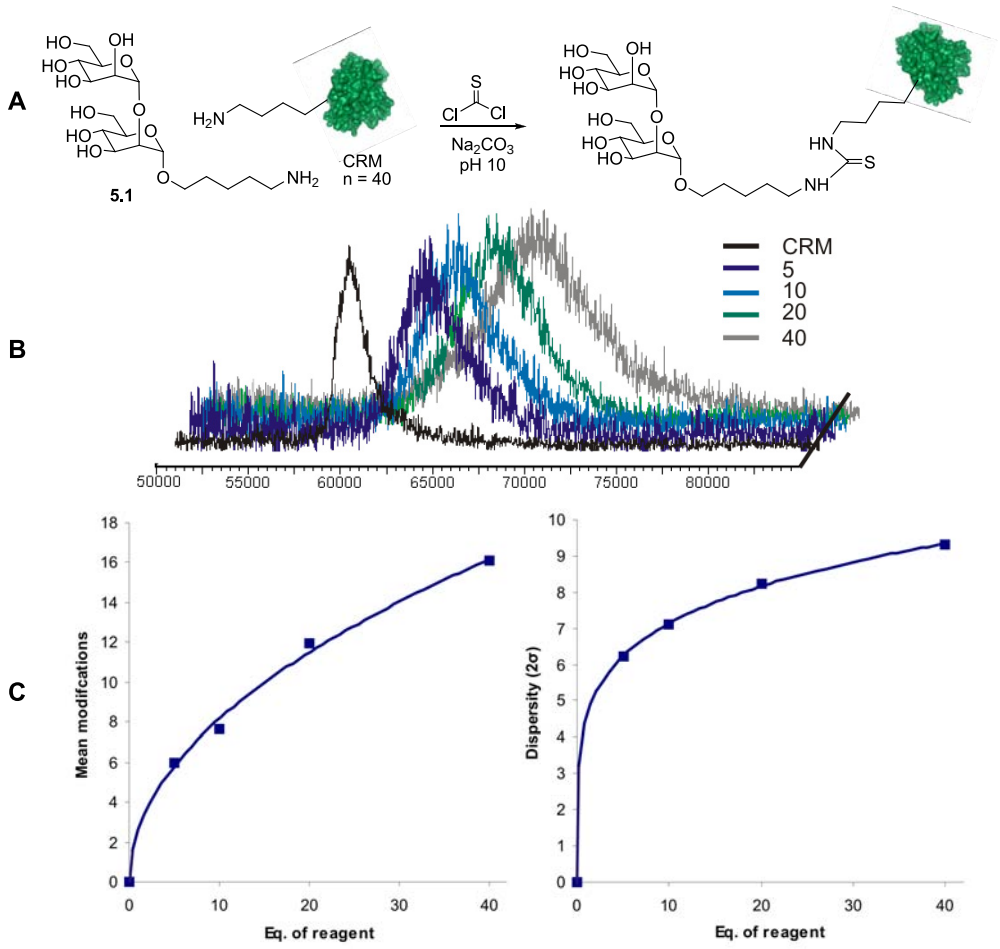


Figure 6.9. *(A) Modification reaction on CRM with increasing concentrations of 5.1 per lysine. (B) Raw MALDI data highlighting the singly charged protein peak. (C) The extent of modification and dispersity revealed from broadened MALDI peaks. The data were well represented as power curves.*

Compared to BSA, the lower number of reactive sites on CRM (39 lysines and the *N*-terminus) meant that the extent of modification was reduced for the latter as would be expected due to the decreased likelihood for a coupling reaction. Furthermore, the dispersity of CRM conjugates is also lower compared with BSA, since the statistical fall in sequential rate constants is exaggerated when there are fewer reactive groups. For instance, after one modification, the expected decrease in rate for CRM would be 97.5% (39/40), whereas for BSA it is only 98.3% (59/60). This steeper fall in rate constants manifests itself as a lower dispersity for reactions on CRM.

[††] CRM was expressed and purified by Dr Julia R. Koeppe, University of Oxford, UK. For full details, see M. K. Patel, B. Vijayakrishnan, J. R. Koeppe, J. M. Chalker, K. J. Doores, B. G. Davis, *Chem. Commun.*, **2010**, 46, 9119.

6.6 Conclusions

Indiscriminate protein modification methodologies are useful for the creation of surface-decorated bioconjugates. Due to their practical simplicity, such methods have found particular use in vaccinology. However, non site-selective conjugation produces a mixture of products and as a result, the purity and characterization of many vaccine candidates falls far short of the standards required for traditional small molecule drugs. The simple mathematical method that we have described here allows dispersity in ligand copy-number to be determined and expressed intuitively, which would allow conjugates to be better characterized without the need for any additional experiments than is currently routine. Our analysis is by no means limited to thiophosgene coupling chemistry and is easily extended to other conjugation methodologies. Given that very few synthetic vaccines in mainstream use are well defined, this should ensure much better reproducibility and allow more logical comparisons to be made between different glycoconjugates.

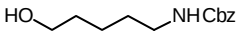
Moreover, comparisons can easily be made between different coupling reagents providing novel insights into protein modification reactions that have not been previously possible. We have shown here that bulkier sugars can increase dispersity, while proteins with fewer reactive groups cause a decrease. Knowledge of the coupling kinetics may allow more predictable conjugations and better control of reactions with proteins.

We propose that our dispersity measure may find use akin to the polydispersity index commonly used in polymer science. Polymers also exist as a mixture of different entities, but unlike bioconjugates, are more rigorously characterized in terms of their underlying distribution.[25] In fact, such measurements are often conducted with mass spectrometry, thus showing many similarities with our dispersity analysis.[26] Other applications of therapeutic relevance can also be envisaged. For example, indiscriminate *in vivo* lysine modification also results in a distribution of metabolic protein products. This "protein glycation" is implicated in several diseases and measuring the dispersity of glycated proteins would assist in disease diagnostics.[19,27]

6.7 Experimental

For general experimental conditions see section 2.8.1 page 44.

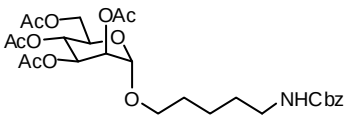
6.7.1 Synthetic protocols



6.3. 5-Aminobenzyloxycarbonyl pentanol

To a solution of **5.5** (3.10g, 30.0mmo1, 1eq) in 1 DCM: 1 H_2O (l00ml) was added benzyloxycarbonyl chloride (6.40g, 37.5mmo1, 1.25eq) and sodium carbonate (7.95g, 75mmol, 2.5eq). After stirring at room temperature for 4 hours, the reaction mixture was diluted with DCM and H_2O and the organic layer was separated and washed with brine, and dried over Na2SOa. The residue was purified by silica gel chromatography (40 to 60% ethyl acetate in petrol over 10 column volumes) to yield the title compound as a white solid (6.93g, 97%).

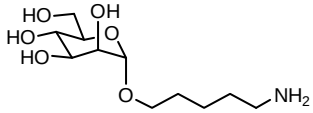
R_f 0.29 (1 petrol : 1 ethyl acetate); ^{1}H NMR (400MHz, $CDCl_3$) δ ppm 1.35-1.45 (2H, m, $CH_2C\underline{H}_2CH_2$), 1.48-1.65 (4H, m, $C\underline{H}_2CH_2NHCbz$, $C\underline{H}_2CH_2OH$), 3.21 (2H, dd, $J_{CH\text{-}NH}$ 13.1Hz, $J_{CH\text{-}CH2}$ 6.7Hz, CH_2NHCbz), 3.64 (2H, t, $J_{CH\text{-}CH2}$ 6.8Hz, $HOC\underline{H}_2$) 4.84 (1H, s, NH), 5.10 (2H, s, $C\underline{H}_2Ph$), 7.34-7.27 (5H, m, ArH); ^{13}C NMR (50MHz, $CDCl_3$) δ ppm 22.8 (CH_2), 29.7 ($\underline{C}H_2CH_2NHCbz$), 32.1 ($\underline{C}H_2CH_2OH$), 40.8 ($\underline{C}H_2NHCbz$), 62.5 ($CH_2OH$), 66.5 ($\underline{C}H_2Ph$), 128.0, 128.4, 136.5 (5 x ArC), 156.4 (NHCO); HRMS (ES^+) m/z 260.1258 $[M + Na]^+$ (required 260.1257).



6.4. 5-(Benzyloxycarbonylamino)pentyl-2,3,4,6-tetra-O-acetyl-a-D-mannopyranoside

To a solution of **6.2** (0.80g, 2.lmmol, leq) in CH_3CN (15ml) was added $SnCl_4$ (0.69g, 2.67mmol, 1.3eq) and the mixture stirred for 20 minutes. To the reaction mixture, **6.3** (0.535g, 2.25mmo1, 1.1eq) dissolved in dry CH_3CN (5ml) was added dropwise and the resultant mixture was stirred for an additional 19 hours at room temperature. The reaction mixture was quenched by addition of saturated $NaHCO_3$ solution and then filtered over a pad of Celite® and washed with hot $CHCl_3$. The filtrate was concentrated and the residue re-dissolved in DCM and then washed with brine and dried over Na_2SO_4. The residue was purified by silica gel chromatography (40 to 60% ethyl acetate in petrol over 10 column volumes) to yield the title compound as a colourless oil (0.68g, 59%).

R_f 0.43 (1 petrol : 1 ethyl acetate); $[\alpha]_D^{25}$ +32.2 (c = 1.03, $CHCl_3$); ^{1}H NMR (400MHz, $CDCl_3$) δ ppm 1.33-1.45 (2H, m, $CH_2C\underline{H}_2CH_2$), 1.47-1.75 (4H, m, $C\underline{H}_2CH_2NHCbz$, $C\underline{H}_2CH_2OH$), 1.99, 2.04, 2.10, 2.16 (4 x 3H, 4 x s, 4 x OAc), 3.21 (2H, dd, $J_{CH\text{-}NH}$ 13.2Hz, $J_{CH\text{-}CH2}$ 6.7Hz, CH_2NHCbz), 3.45 (1H, td, ^2J 9.4Hz, $J_{CH\text{-}CH}$ 9.4Hz, $J_{CH\text{-}CH}$ 6.3Hz, $OCH\underline{H}CH_2$), 3.69 (1H, td, ^2J 9.2Hz, $J_{CH\text{-}CH}$ 9.2Hz, $J_{CH\text{-}CH}$ 6.5Hz, $OC\underline{H}HCH_2$), (1H, ddd, $J_{H4\text{-}H5}$ 9.4Hz, $J_{H5\text{-}H6a}$ 5.2Hz, $J_{H5\text{-}H6b}$ 2.3Hz, H5), 4.10 (1H, dd, $J_{H6a\text{-}H6b}$ 10.2Hz, $J_{H5\text{-}H6b}$ 1.9Hz, H6b), 4.29 (1H, dd, $J_{H6a\text{-}H6b}$ 12.2Hz, $J_{H5\text{-}H6a}$ 5.3Hz, H6a), 4.80 (1H, d, $J_{H1\text{-}H2}$, 1.3Hz, H1), 4.86 (1H, s, NH), 5.10 (2H, s, $C\underline{H}_2Ph$), 5.23 (1H, dd, $J_{H2\text{-}H3}$ 3.3Hz, $J_{H1\text{-}H2}$ 1.7Hz, H2), 5.27 (1H, t, $J_{H3\text{-}H4}$ 9.9Hz, $J_{H4\text{-}H5}$ 9.9Hz, H4), 5.34 (1H, dd, $J_{H3\text{-}H4}$ 10.0Hz, $J_{H2\text{-}H3}$ 3.3Hz, H3), 7.34-7.26 (5H, m, ArH); ^{13}C NMR (50MHz, $CDCl_3$) δ ppm 20.7 (2 x $\underline{C}H_3CO$), 20.8, 20.9 (2 x 1C, $\underline{C}H_3CO$), 23.38 (CH_2), 28.8 ($\underline{C}H_2CH_2NHCbz$), 29.7 ($\underline{C}H_2CH_2OH$), 40.8 ($\underline{C}H_2NHCbz$), 62.5 (C6), 66.2 (C4), 66.5 ($\underline{C}H_2Ph$), 68.2 (OCH_2), 68.3 (C5), 69.0 (C3), 69.6 (C2), 97.4 (C1) 128.0, 128.4, 136.5 (5 x ArC), 156.3 (NHCO), 169.7, 169.9, 170.0, 170.6 (4 x $\underline{C}OCH_3$); HRMS (ES^+) m/z 590.2208 $[M + Na]^+$ (required 590.2197).



6.5. 5-(Amino)pentyl-D-mannopyranoside

To a solution of **6.4** (0.30g, 0.53mmol, l.0eq) in dry MeOH (4ml) and DCM (4ml) was added sodium methoxide (25% wt. solution, 63µl,0.27mmol, 0.5eq) and stirred at room temperature for 3 hours. The reaction mixture was neutralized with Dowex 50WX8 (H^+ form), filtered and the solvent removed in vacuo. The residue was dissolved in 9 CF_3CH_2OH : 1 H_2O with l% HCOOH (3ml) with Pd/C (10% ,0.11g) and stirred continuously under a hydrogen atmosphere for 24 hours. The reaction mixture filtered over Celite® and lyophilized to yield the title compound as an amorphous white solid (0.13g, 91%).

R_f 0.5 (5 ethanol : 3 NH_4OH : 1 water); $[\alpha]_D^{25}$ +44.9 (c = 1.0, MeOH); ^{1}H NMR (500MHz, MeOD) δ ppm 1.45 (2H, quin, $J_{H8\text{-}H9}$ 6.9Hz, $J_{H9\text{-}H10}$ 6.9Hz, H9), 1.55 (2H, quin, $J_{H9\text{-}H10}$ 7.1Hz, $J_{H10\text{-}H11}$ 7.1Hz, H10), 1.63 (2H, dq, $J_{H8\text{-}H8'}$ 13.5Hz, $J_{H7\text{-}H8}$ 6.6Hz, $J_{H8\text{-}H9}$ 6.6Hz, H8, H8'), 2.70 (2H, t, $J_{H10\text{-}H11}$ 7.3Hz, H11), 3.45 (1H, dt, $J_{H7\text{-}H7'}$ 9.8, $J_{H7'\text{-}H8}$ 6.1Hz, H7'), 3.54 (1H, ddd, $J_{H4\text{-}H5}$ 9.7Hz, $J_{H5\text{-}H6'}$ 6.0Hz, $J_{H5\text{-}H6}$ 1.9Hz, H5), 3.61 (1H, t, $J_{H3\text{-}H4}$ 9.5Hz, $J_{H4\text{-}H5}$ 9.5Hz, H4), 3.70 (1H, dd, $J_{H3\text{-}H4}$ 9.6Hz, $J_{H2\text{-}H3}$ 3.9Hz, H3), 3.72 (1H, dd, $J_{H6\text{-}H6'}$ 11.2Hz, $J_{H5\text{-}H6'}$ 5.2Hz, H6'), 3.78 (1H, dt, $J_{H7\text{-}H7'}$ 9.8Hz, $J_{H7\text{-}H8}$ 6.6Hz, H7), 3.80 (1H, dd, $J_{H2\text{-}H3}$ 3.2Hz, $J_{H1\text{-}H2}$ 1.6Hz, H2), 3.85 (1H, dd, $J_{H6\text{-}H6'}$ 11.8Hz, $J_{H5\text{-}H6}$ 2.0Hz, H6), 4.76 (1H, as, H1); ^{13}C NMR (126MHz, MeOD) δ ppm

24.7 (1C, C9), 30.4 (1C, C8), 32.9 (1C, C10), 42.3 (1C, C11), 63.0 (1C, C6), 68.4 (1C, C7), 68.7 (1C, C4), 72.3 (1C, C2), 72.7 (1C, C3), 74.7 (1C, C5), 101.4 (1C, C1); FT-IR (thin film) υ 2968, 2931 (NH_2), 3355 br (OH); HRMS m/z (ES^+) 266.1602 [M + H]$^+$ required 266.1598.

Preparation of CRM

CRM was prepared by Dr Julia R Koeppe (University of Oxford) using the protocol shown below.

Expression of CRM

A 10 mL starter culture of LB containing 0.1 mg/mL ampicillin was inoculated with a single colony from a fresh transformation of CRM in *E. coli* BL21(DE3) cells and grown at 37 ºC overnight with shaking.

One litre of LB containing 0.1 mg/mL ampicillin was inoculated with the 10 mL overnight culture. This was grown at 37 ºC with shaking to an OD_{600} of 0.6. Protein production was induced with the addition of 1 mM IPTG. The culture was allowed to continue shaking at 25 ºC for 16 hours.

The cells were harvested by spinning for 20 minutes at 13000g. The cell pellets were collected and stored at -80 ºC until next use.

Purification of CRM

Cell lysis: The cell pellets collected from 2 L of growth media were lysed in 40 mL of a buffer containing 50 mM TEA at pH 7.8, 300 mM NaCl, 15 mM imidazole, 1 mM BME, 1 mg/mL lysozyme and 0.1 mg/mL DNAse. The pellet and buffer were allowed to incubate on ice for 30 minutes.

The cell pellets were then sonicated at 150 W in bursts of 30 seconds with a wait time of 1 minute between each burst. A total of 4 bursts were used to break the cells. The sonication of the cells was carried out in a plastic beaker placed on ice.

The cell debris was removed from the lysis mixture by centrifugation at 13000g for 30 minutes. Samples of the pellet and soluble fractions were saved for SDS-PAGE evaluation.

NiNTA affinity purification: The soluble portion of the lysed cells was loaded onto a 5 mL pre-packed NiNTA column (GE Healthcare) which had been equilibrated in buffer containing 50 mM TEA at pH 7.8, 300 mM NaCl, 1 mM BME and 15 mM imidazole. CRM was eluted using a linear gradient to the same buffer containing 300 mM imidazole. Protein-containing fractions (as determined by UV absorbance) were analyzed by SDS-PAGE.

Anion exchange purification: Protein-containing fractions from the NiNTA column were pooled and then loaded onto a 1 mL MonoQ 5/50 GL anion exchange column (GE healthcare) which had been equilibrated in 20 mM Tris, pH 7.8, 150 mM NaCl. The protein was eluted using a linear gradient to the same buffer containing 1 M NaCl. Protein-containing fractions were then analyzed by SDS-PAGE.

Size exclusion chromatography: Protein-containing fractions from the MonoQ column were concentrated to a volume < 10 mL. The protein was then loaded onto a S75 size exclusion column (HiLoad 16/60 Superdex 75, GE Healthcare) which had been equilibrated in 20 mM Tris, pH 7.8, 150 mM NaCl. Only 2mL of protein solution was added at any one time. Protein-containing fractions (as determined by UV absorbance) were analyzed by SDS-PAGE and combined and concentrated and stored at -20 ºC until further use.

For glycoconjugation experiments, the protein was exchanged into 300 mM $NaHCO_3$ buffer at pH 9.0 using a PD10 column (GE Healthcare). The protein concentration was checked by BCA assay (Pierce), and the protein was diluted to 3.0 mg/mL with the same $NaHCO_3$ buffer prior to the experiment.

6.7.2 Protein sequences

BSA

```
DTHKSEIAHR FKDLGEEHFK GLVLIAFSQY LQQCPFDEHV KLVNELTEFA KTCVADESHA  60
GCEKSLHTLF GDELCKVASL RETYGDMADC CEKQEPERNE CFLSHKDDSP DLPKLKPDPN  120
TLCDEFKADE KKFWGKYLYE IARRHPYFYA PELLYYANKY NGVFQECCQA EDKGACLLPK  180
IETMREKVLT SSARQRLRCA SIQKFGERAL KAWSVARLSQ KFPKAEFVEV TKLVTDLTKV  240
HKECCHGDLL ECADDRADLA KYICDNQDTI SSKLKECCDK PLLEKSHCIA EVEKDAIPEN  300
LPPLTADFAE DKDVCKNYQE AKDAFLGSFL YEYSRRHPEY AVSVLLRLAK EYEATLEECC  360
AKDDPHACYS TVFDKLKHLV DEPQNLIKQN CDQFEKLGEY GFQNALIVRY TRKVPQVSTP  420
TLVEVSRSLG KVGTRCCTKP ESERMPCTED YLSLILNRLC VLHEKTPVSE KVTKCCTESL  480
VNRRPCFSAL TPDETYVPKA FDEKLFTFHA DICTLPDTEK QIKKQTALVE LLKHKPKATE  540
EQLKTVMENF VAFVDKCCAA DDKEACFAVE GPKLVVSTQT ALA                    583
```

Molecular weight: 66463

CRM

```
GADDVVDSSK SFVMENFSSY HGTKPGYVDS IQKGIQKPKS GTQGNYDDDW KGFYSTDNKY  60
DAAGYSVDNE NPLSGKAGGV VKVTYPGLTK VLALKVDNAE TIKKELGLSL TEPLMEQVGT  120
EEFIKRFGDG ASRVVLSLPF AEGSSSVEYI NNWEQAKALS VELEINFETR GKRGQDAMYE  180
YMAQACAGNR VRRSVGSSLS CINLDWDVIR DKTKTKIESL KEHGPIKNKM SESPNKTVSE  240
EKAKQYLEEF HQTALEHPEL SELKTVTGTN PVFAGANYAA WAVNVAQVID SETADNLEKT  300
TAALSILPGI GSVMGIADGA VHHNTEEIVA QSIALSSLMV AQAIPLVGEL VDIGFAAYNF  360
VESIINLFQV VHNSYNRPAY SPGHKTQPFL HDGYAVSWNT VEDSIIRTGF QGESGHDIKI  420
TAENTPLPIA GVLLPTIPGK LDVNKSKTHI SVNGRKIRMR CRAIDGDVTF CRPKSPVYVG  480
NGVHANLHVA FHRSSSEKIH SNEISSDSIG VLGYQKTVDH TKVNSKLSLF FEIKSLEHHH  540
HHH                                                                543
```

Molecular weight: 59406

6.7.3 Protein modification

General protocol for preparation of isothiocyanates

To a solution of **6.5** or **5.1** (1 eq) in 0.3 M $NaHCO_3$ (6 eq) and 0.3 M Na_2CO_3 (3 eq) at pH 9 (pH paper) was added a solution of thiophosgene (6 eq) dissolved in chloroform. The biphasic mixture was vigorously stirred at room temperature and monitored by TLC (1 water : 4 isopropanol : 4 ethyl acetate) and mass spectrometry (ES^+). After ~ 3 hours, complete consumption of the starting sugar and the formation of a single product was detected. Solvents and excess thiophosgene were removed under reduced pressure. The isothiocyanate of **6.5** was desalted for characterization purposes by silica gel chromatography (9 ethyl acetate : 1 methanol) to afford the desired compound as a pale yellow oil (24 mg, 84 %). The isothiocyanate of **5.1** was used without further purification.

1-Isothiocyano-5-pentyl-*O*-α-D-mannoopyranoside

R_f 0.23 (9 ethyl acetate : 1 methanol), $[\alpha]_D^{18}$ +33.2 (c = 1.2, MeOH); ^{1}H NMR (500 MHz, MeOD) δ ppm 4.77 (1d, *J* = 1.6 Hz, 1H, H-1), 3.83 (m, 1H, H-6a), 3.83 (t, *J* = 2.5 Hz, 1H, H-2), 3.76 (ad, *J* 5.7 Hz, 1H, OC*H*HCH_2), 3.75 (m, 1H, H-6b), 3.73 (dd, *J* = 10.1, 2.5 Hz, 1H, H-3), 3.65 (t, *J* = 9.6 Hz, 1H, H-4), 3.60 (t, *J* = 6.6 Hz, 1H, C*H*HNCS), 3.54 (ddd, *J* = 9.5, 5.5, 2.4 Hz, 1H, H-5), 3.48 (t, *J* = 6.3 Hz, 1H, CH*H*NCS), 3.47 (ddd, *J* = 14.3, 8.8, 2.2 Hz, 1H, OCH*H*CH_2), 1.76 (quin, *J* = 6.9 Hz, 1H, OCH_2C*H*HCH_2), 1.60-1.70 (m, 3H, OCH_2CH*H*, C*H*$_2CH_2$NCS), 1.55 (m, 1H, C*H*HCH_2), 1.45 (m, 1H, CH*H*CH_2); ^{13}C NMR (126 MHz, MeOD) δ ppm 101.6 (C-1), 74.6 (C-5), 72.6 (C-3), 72.2 (C-2), 68.6 (C-4), 68.3 (OCH_2), 62.7 (C-6), 45.9 (*C*H_2NCS), 30.9 (OCH_2*C*H_2), 29.8 (*C*H_2CH_2NCS), 24.5 (CH_2), low intensity NCS

peak not observed; FT-IR (thin film) υ 2102 (N=C=S), 2180 (N=C=S), 2936 (C-H), 3354 br (OH); HRMS m/z (ES^+) 330.0970 $[M + Na]^+$ (required 330.0982).

General protocol for reactions

Reactions were performed using between 0.1 – 1.0 mg of protein at a concentration of 1.5mg/ml. The reactions were conducted in 0.3 M $NaHCO_3$ and 0.3 M Na_2CO_3. A solution of the isothiocyanate was added and the reaction mixture was gently shaken at room temperature for 24 hours. Increasing ratios of isothiocyanate per lysine residue were used per reaction set.

6.7.4 MALDI mass spectrometric analysis

5 μL of the reaction mixture was diluted to 250 μL with 0.1 % TFA in water. This solution was mixed in a 1 : 1 ratio (v/v) with a solution of mass spectrometry grade sinapinic acid (10 mg/mL in 1 : 1 water : acetonitrile with 0.1% TFA). From this combined matrix/ sample solution, 2 μL was spotted onto a steel target and allowed to co-crystallize at room temperature over 3 hours. MALDI analysis was conducted on a Water MALDI Micro MX spectrometer with TOF detection, in positive linear mode. Unmodified protein was used as a lock mass calibrant. For each comparative set of experiments, laser energy and pulse width were optimised for a modified protein sample, and kept identical for the remainder of the sample set. In general, multiple laser shots were combined to produce the final spectrum. Data were further processed using Mass Lynx 4.1. Mass spectra were smoothed with Savitzky-Golay smoothing prior to dispersity analysis.

6.7.5 MATLAB M-file scripts

Example M-file code for BSA (n = 60) modified with **1** (m = 307), where the measured FWHM for unmodified BSA = 1298:

```
function [estimates, model] = Dispersity(xdata, ydata)
start_point = [0 500];
model = @MALDIfit;
options = optimset('Display','iter');
estimates = fminsearch(model, start_point, options);
    function [sse, FittedCurve] = MALDIfit(params)
        mu = params(1);
        sigma = params(2);
        m = 307;
        % m = the mass change per modification
```

```
        delta = 1298/2.35;
        % delta = standard deviation of MALDI peak for unmodified protein
        % 1298 is the measured FWHM
        x = (-10000:20000);
        FittedCurve = 0;
            for n=0:60
            % n = the number of lysines + N-terminus
            FittedCurve = FittedCurve + (normcdf((n+0.5)*m,mu,sigma)-
normcdf((n-0.5)*m,mu,sigma))*normpdf(x, n*m, delta);
            end
        ErrorVector = FittedCurve - ydata;
        sse = sum(ErrorVector .^ 2);
    end
end
```

Command line entry for a MALDI peak of modified protein of average mass relative to unmodified BSA = 3247, and measured FWHM = 3277:

```
xdata = (-10000:20000);
ydata = normpdf(xdata, 3247, 3277/2.35);
[estimates, model] = Dispersity(xdata,ydata)
```

6.8 References

Figures in this chapter were reproduced by permission from The Royal Society of Chemistry from: M. K. Patel, B. Vijayakrishnan, J. R. Koeppe, J. M. Chalker, K. J. Doores, B. G. Davis, *Chem. Commun.*, **2010**, 46, 9119.

1. R. A. Dwek, *Chem. Rev.*, 1996, **96**, 683.
2. (a) N. Jayaraman, *Chem. Soc. Rev.*, 2009, **38**, 3463. (b) J. J. Lundquist, E. J. Toone, *Chem. Rev.*, 2002, **102**, 555. (c) M. Mammen, S. Choi, G. M. Whitesides, *Angew. Chem. Int. Ed.*, **1998**, 37, 2754.
3. K. J. Doores, D. P. Gamblin, B. G. Davis, *Chem. Eur. J.*, 2006, **12**, 656.
4. (a) R. Roy, *Carbohydrate Based Vaccines*, ACS Symposium Series, **2008**,. (b) M. Hecht, P. Stallforth, D. V. Silva, A. Adibekian, P. H. Seeberger, 989*Curr. Opin. Chem. Biol.*, **2009**, 13, 354. (c) R. Roy, *Drug Discovery Today: Technologies*, 2004, **1**, 327. (d) S. B. Oppenheimer, M. Alvarez, J. Nnoli, *Acta Histochemica*, **2008**, 110, 6.
5. Y. Zhu, X. Li, A. McVie-Wylie, C. Jiang, B. L. Thurberg, N. Raben, R. J. Mattaliano, S. H. Cheng, *Biochem. J.*, **2005**, 389, 619.
6. (a) M. A. Robinson, S. T. Charlton, P. Garnier, X.-T. Wang, S. S. Davis, A. C. Perkins, M. Frier, R. Duncan, T. J. Savage, D. A. Wyatt, S. A.Watson, B. G.

Davis, *Proc. Natl. Acad. Sci. USA* , **2004**, 101, 14527. (b) B. G. Davis, M. A. Robinson, *Curr. Opin. Drug Discovery Dev.*, **2002**, 5, 279.

7. O. M. T. Pearce, K. D. Fisher, J. Humphries, L. W. Seymour, A. Smith, B. G. Davis, *Angew. Chem. Int. Ed..*, **2005**, 44, 1057.
8. (a) B. G. Davis, *Chem. Rev.*, **2002**, 102, 579. (b) D. P. Gamblin, E. M. Scanlan, B. G. Davis, *Chem Rev.*, **2009**, 109, 131. (c) B. G. Davis, *Pure Appl. Chem.*, **2009**, 81, 285.
9. P. M. Rendle, A. Seger, J. Rodrigues, N. J. Oldham, R. R. Bott, J. B. Jones, M. M. Cowan, B. G. Davis, *J. Am. Chem. Soc.*, **2004**, 126, 4750.
10. (a) S. I. van Kasteren, H. B. Kramer, D. P. Gamblin, *Nat. Protoc.*, **2007**, 2, 3185. (b) S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony, B. G. Davis, *Nature*, **2007**, 446, 1105.
11. (a) G. T. Hermanson, *Bioconjugate Techniques*, Academic Press, 2nd edn., **1996**. (b) M. Izumi, S. Okumura, H. Yuasa, H. Hashimoto, *J. Carbohydr. Chem.*, **2003**, 22, 317.
12. (a) D. J. Harvey, *Int. J. Mass Spectrom.*, **2003**, 226, 1. (b) D. J. Harvey, *Mass Spectrom. Rev.*, **2009**, 28, 273. (c) D. J. Harvey, *Mass Spectrom. Rev.*, **2008**, 27, 125. (d) D. J. Harvey, *Mass Spectrom. Rev.*, **2006**, 25, 595.
13. C. Jones, *Carbohydrate Based Vaccines*, ACS Symposium Series, **2008**, 989, ch. 2.
14. (a) B. A. Garcia, *J. Am. Soc. Mass Spectrom.*, **2010**, 21, 193. (b) K.Breuker, M. Jin, X. Han, H. Jiang, F. W. Mclafferty, *J. Am. Soc. Mass Spectrom.*, **2008**, 19, 1045. (c) S. M. Patrie, D. E. Robinson, F. Meng, Y. Du, N. L. Kelleher, *Int. J. Mass Spectrom.*, **2004**, 234, 17 5.
15. (a) A. G. Marshall, C. L. Hendrickson, G. S. Jackson, *Mass Spectrom Rev.*, **1998**, 17, 1. (b) M.P. Barrow, W. I. Burkitt, P. J. Derrick, *Analyst*, **2005**, 130, 18. (c) A. G. Marshall, C. L. Hendrickson, *Int. J. Mass Spectrom.*, **2002**, 215, 59. (d) L-K. Dong, D. Rempel, B. M. Pramanik, M. L. Gross, *Mass Spectrom. Rev.*, **2005**, 24, 286.
16. A. G. Marshall, C. L. Hendrickson, G. S. Jackson, *Mass Spectrom. Rev.*, **1998**, 17, 1. (b) N. L. Kelleher, M. W. Senko, M. M. Siegel, F. W. McLafferty, *J. Am. Soc. Mass Spectrom.*, **1997**, 8, 380. (c) K. Nagy, K. Vékey, T. Imre, K.

Lundányi, M. P. Barrow, P. J. Derrick, *Anal. Chem.*, **2004**, 76, 4998. (d) Y. Liu, Y. Pan, *Appl. Spectrosc. Rev.*, **2009**, 44, 231. (e) A. V. Tolmachev, E. W. Robinson, Wu, L. Pasa-Tolic, R. D. Smith, *Int. J. Mass Spectrom.*, **2009**, 287, 32.

17. D. M. Horn, R. A. Zubarev, F. W. McLafferty, *J. Am. Soc Mass Spectrom.*, **2000**, 11, 320.
18. K. J. Doores, Z. Fulton, V. Hong, M. K. Patel, C. N. Scanlan, M. R. Wormald, M. G. Finn, D. R. Burton, I. A. Wilson, B. G. Davis, *Proc. Natl. Acad. Sci. USA*, **2010**, 107, 17107.
19. F. K. Yeboah , V. A. Yaylayan, *Nahrung*, 2001, **45**, 164.
20. E. Sachon, G. Clodic, T. Blasco and G. Bolbach, *J. Am. Soc. Mass Spectrom.*, **2007**, 18, 1880.
21. (a) J. G. Joyce, I. J. Krauss, H. C. Song, D. W. Opalka, K. M. Grimm, D. D. Nahas, M. T. Esser, R. Hrin, M. Feng, V. Y. Dudkin, M. Chastain, J. W. Shiver, S. J. Danishefsky, *Proc. Natl. Acad. Sci. USA*, **2008**, 105, 15684. (b) J. Ni, H. Song, Y. Wang, N. M. Stamatos, L. Wang, *Bioconjugate Chem.*, **2006**, 17,493.
22. R. D. Astronomo, H. Lee, C. N. Scanlan, R. Pantophlet, C. Huang, I. A. Wilson, O. Blixt, R. A. Dwek, C. Wong, D. R. Burton, *J. Virol.*, **2008**, 82, 6359; D. H. Buss, I. J. Goldstein, *J. Chem. Soc. C*, 1968, 1457.
23. (a) G. T. Hermanson, *Bioconjugate Techniques*, Academic Press, 2nd edn., **1996**. (b) M. Izumi, S. Okumura, H. Yuasa, H. Hashimoto, *J. Carbohydr. Chem.*, **2003**, 22, 317. (c) S. I. van Kasteren, P. Garnier, B. G. Davis, *Nucleic Acids and Molecular Biology*, Protein Engineering, **2009**, Springer-Verlag.
24. H. J. R. Maget, *J. Polym. Sci., Part A: Polym. Chem.*, **1964**, 2, 1281.
25. S. A. Schmid, R. Abbel, A. P. H. Schenning, E. W. Meijer, R. P. Sijbesma, L. M. Herz, *J. Am. Chem. Soc.*, **2009**, l3l, 17696.
26. (a) C.A. Jackson, W. J. Simonsick, *Curr. Opin. Solid. St. M.*,**1997**, 2, 661. (b) H.Zhu, T. Yalcin, L.Li, *J. Am. Soc. Mass Spectrom.*, **1998**, 9, 275. (c) P. B. O'Connor, F. W. Mclafferty, *J. Am. Chem. Soc.*, **1995**, 117 , 12826.
27. (a) T. Niwa, *Mass Spectrom. Rev.*, **2006**, 25, 713. (b) A. Lapolla, D. Fedele, P. Traldi, *Mass Spectrom. Rev.*,**2000**, 19, 279.

Chapter VII – Conclusions and Future Directions

In this thesis, we have used the tools of chemistry to attempt to understand the biochemistry of two different carbohydrate systems involving charged functionality. The major achievements of this research will be highlighted in this chapter together with the limitations of our work and possible directions for future research.

Significant progress has been made in elucidating the role of Tre-6-P in plant signalling. Prior to our work, a variety of carbohydrate phosphate esters (Tre-6-P, Glc-1-P, Glc-6-P and Rib-5-P) had been proposed as potential inhibitors of the protein kinase SnRK1. The nature of this interaction is poorly understood and is hypothesized to occur via an intermediary factor. The work described in Chapters II – IV has produced three main contributions to this field. Firstly, by studying the stability of these sugars under assay conditions, we have conclusively demonstrated that Tre-6-P, Glc-1-P and Glc-6-P are true inhibitors of SnRK1. In contrast, the reduction in SnRK1 activity by Rib-5-P was shown to actually occur through metabolism of the sugar resulting in consumption of ATP giving an apparent but not a true inhibition. Thus, we can now be confident that Tre-6-P (the most potent of the three carbohydrates) is a real inhibitor of SnRK1 and attention can be focused on elucidating its mechanism of action.

A second accomplishment of our work is the development of methodology for the synthesis of analogues of Tre-6-P. While various analogues of trehalose have been previously synthesized, similar synthetic variants of Tre-6-P are much rarer. The problems with the synthesis of such compounds originate from (i) the desymmetrization of trehalose; and (ii) the installation of the phosphate group. Initially, we attempted to develop methodology for the efficient desymmetrization of trehalose using a flow reactor. Unfortunately, we could not develop this method for large scales due to blockage of the reactor channel. The use of biphasic or slug flow may be potential workarounds that could be explored in the future. While our large scale synthesis was unsuccessful, our studies furnished kinetic data that provided insight into these reactions, and in particular, explained the particularly low yields that have been observed by other researchers. Our studies on the installation and manipulation of the phosphate group proved more fruitful. We discovered that phosphate migrations could be controlled during transesterfications by using CsF.

This allows facile exchange of phosphoryl protecting groups allowing for greater flexibility in the design of synthetic routes. This methodology was extended to a range of carbohydrate scaffolds and used for the synthesis of a library of Tre-6-P analogues, which formed the basis of understanding structure-activity relationships between Tre-6-P and the intermediary factor. In this way, we found that while the phosphate group was essential for binding, one of the P-OH groups was dispensable. Crucially, we found that bulky groups could be tolerated at the C6' centre of Tre-6-P which could therefore serve as a site of attachment for tethering the molecule onto a solid phase.

Along these lines, a third achievement of our work was the synthesis of a Tre-6-P affinity column using our CsF transesterfication methodology and the data from Tre-6-P structure-activity relationships. This column was used to extract potential Tre-6-P binding proteins from a partially purified extract of SnRK1, which could then be identified by proteomics analyses. In this way, the 14-3-3 class of proteins was identified as potential intermediary factors, but unfortunately, binding to Tre-6-P could not be demonstrated *in vitro* and further work is needed in this direction. The lack of observed sugar binding to the recombinant 14-3-3 may be due to oversimplification of the *in vitro* system. For example, post-translational modification of the 14-3-3 may be required for binding to Tre-6-P. These intricacies would be difficult to replicate for a recombinant protein and alternative experiments may be required. Alternatively, the 14-3-3s may be a rogue result from non-specific binding to the column. To verify if this is the case, a control could be conducted with the trehalose affinity column. Various other proteins were also detected in the extracted fractions, but in too small a quantity for successful identification. Ideally, the extraction should be repeated on a larger scale to isolate greater quantities of these proteins that would allow for cleaner identifications. Additionally, in our studies, the column was eluted with an increasing salt gradient to release bound proteins. Strongly adhered proteins may not be released in this way, and elution with Tre-6-P would be preferable. While there are still many questions that remain to be answered, we have shown that the Tre-6-P affinity column can, in principle, be used to identify the intermediary factor.

Our studies on the binding of GalNAc-4-SO_4 to the MR are clearly still at a very early stage and significant work is required in this area. On the chemistry side, we have

developed syntheses for probes which can allow attachment onto quantum dots, small molecule dyes and proteins. For the latter case, we have developed a technique that can reveal the dispersity in the distribution of labelling using MALDI mass spectroscopy, which would allow better monitoring of reaction products and more consistent comparison between experiments involving such chemically modified proteins. This approach was used to study various protein modifications which have furnished insights into the kinetics of such reactions.

The synthesis of a mannose based probe for the CRDs proved to be successful and can be readily used for further studies. The synthesis for our GalNAc-4-SO_4 probe was more challenging. The reaction scheme failed at the last step and alternative methods for this final deprotection should be investigated. Furthermore, the binding of these probes to the MR should be experimentally investigated. The recombinant expression of various CRD and CRR fragments of the MR would give access to pure protein to which binding studies could be performed using isothermal titration calorimetry or surface plasmon resonance. The generic sulfate probe that we have used in the place of GalNAc-4-SO_4 should also be rigorously studied for interaction with the CRR. Our experiments showed a surprisingly fast cellular uptake for this probe and if significant binding to the CRR could be demonstrated, such a result may point to a wider role for sulfate biochemistry than is currently understood.

Our microscopy studies using the NR8383 cell line have demonstrated that internalization of our probes is possible. While some of our experiments using unlabelled fluorescent probes and with the SK-HEP-1 cell line have hinted at receptor mediated endocytosis, more controls are needed to fully establish the role of the MR. The most rigorous experiments would involve genetic manipulation of cells. For example, macrophages derived from a $MR^{-/-}$ knockout rat would be a good negative control for our NR8383 experiments. An even more rigorous approach would use a cell line that does not normally express MR, which could then be transfected with the gene for the MR. Comparison between the wild type and transfected lines would provide direct evidence for the role of the MR.

Similarly, the FRET microscopy studies have demonstrated signal enhancements at the cell surface for the NR8383 cell line, but the involvement of the MR still needs to be rigorously proved. As well as the general experiments mentioned above, further

studies with differing concentrations of the dye pair should be conducted to rule out non-specific interactions at the cell surface. Further experiments with sodium azide to inhibit endosome formation could also be performed.

While much work is clearly still needed in this area, our early studies have hinted that recognition of our probes may be possible. Moreover, given the potential importance of MR dimerization in the immune response, such studies would be a worthwhile exercise.