

Investigations into Novel Enzymatic Glycosylation Methods

Thomas A Taylor

Pembroke College



A Thesis Submitted to the Board of the Faculty of Physical Sciences in Partial Fulfilment for the Degree of Doctor of Philosophy in Organic Chemistry at the University of Oxford

Department of Chemistry

University of Oxford

Trinity Term 2015

Investigations into Novel Enzymatic Glycosylation Methods

Thomas A Taylor

Pembroke College

Degree of Doctor of Philosophy in Organic Chemistry

Trinity Term 2015

Glycosylation is one of the most prevalent post-translational modifications found on human proteins. There is a great interest in developing new methodologies for the synthesis of glycoproteins, both to elucidate the functions of glycans on proteins and to exploit the beneficial properties they can confer on biotherapeutics. Here, research was carried out investigating enzymatic techniques to achieve the formation of homogenous glycoproteins.

EndoS was the first enzyme investigated. EndoS natural activity is for the hydrolysis of glycans specifically from IgG, yet there has been research into its use for the opposite reaction and for glycan extension reactions. Here, a number of EndoS mutants were formed and investigated for this reaction, with a number of novel constructs giving a moderate yield of the glycosylated product. Structural investigations of EndoS were also conducted.

In addition, research was conducted into the use of PglB *in vitro* for the direct glycosylation of an asparagine residue on a peptide with GlcNAc. Studies here demonstrate that it is possible to use far simpler glycan donor substrates with PglB. Additional studies showed that it was possible to conduct further enzymatic glycosylation reactions with GalT and EndoA after the initial PglB reaction.

Further research was undertaken with EndoA, with it being desired to investigate whether it could catalyse the formation of a thioglycosidic linkage between two glycans, with the thiol bearing glycan acceptor to be attached to a peptide using PglB. After the successful synthesis of all the substrates, disappointingly, neither of the two enzymatic reactions were successful.

Finally, the use of a glycal donor in an EndoA catalysed glycosylation was investigated. Unfortunately, no consumption of the glycan was observed in this reaction.

For My Family
Past, Present and Future

*Do the difficult things while they are easy and do the great things
while they are small. A journey of a thousand miles must begin
with a single step*

Lao Tzu

The universe is made of stories, not of atoms.

Muriel Rukeyser

Author Declaration

The work presented in this Thesis was conducted at the Chemical Research Laboratory under the supervision of Professor Benjamin G. Davis. All the work is my own, except where otherwise stated, and has not been submitted for any other degree at this or any other university.

In Chapter 2, the mutant EndoS Y305F plasmid was formed by Dr Keisuke Yamamoto. GST-EndoS D233A was also expressed by Dr Keisuke Yamamoto. Testing of the hydrolytic activities of the EndoS constructs was conducted in collaboration with Dr Thomas Parsons. Testing of the glycosylation activities of the EndoS constructs and all densitometry analysis was performed by Dr Thomas Parsons. Wendy Offen (Prof. Gideon Davies lab, University of York) set up all the crystallisation trials for EndoS and mutants thereof. The pHLSec plasmid encoding for IgG1 Fc was a kind gift from the Dr Max Crispin group.

In Chapter 3 and 4, the initial PglB used in some of the glycosylation reactions was a gift from Glycovaxyn.

Where stated that a compound had been previously formed in the group and/or no experimental data is given for a compound, this has been formed by another member of Prof. Ben Davis Group.

Publications

The work in this Thesis has been published in the following journals:

F. Liu, B. Vijayakrishnan, A. Faridmoayer, T. A. Taylor, T. B. Parsons, G. J. L. Bernardes, M. Kowarik and B. G. Davis, *J. Am. Chem. Soc.*, 2014, **136**, 566 – 569.

T. B. Parsons, W. B. Struwe, J. Gault, K. Yamamoto, T. A. Taylor, R. Raj, K. Wals, S. Mohammed, C. V. Robinson, J. L. P. Benesch and B. G. Davis, *Angw. Chem.*, 2016, **128**, 2407 – 2413.

Acknowledgements

First, I wish to thank Prof. Ben Davis for giving me the opportunity to be part of his group for 3.5 years and giving me a large number of different and varied projects, allowing me to learn a range of different techniques and be involved in broad scope of science. I also wish to thank Dr Balakumar Vijayakrishnan and Dr Thomas Parsons, for all their supervision and support. Both of you spent significant amounts of your time answering my questions and giving me helpful suggestions and encouragement when I encountered problems; for this I am extremely grateful. Furthermore, I am grateful to Dr Keisuke Yamamoto for all his patience at the start of my DPhil, teaching me all the biological techniques and coping with my queries whilst finishing his own DPhil.

The NMR and Mass Spec teams have been amazing, insuring all the equipment was always available and assisting in obtaining characterisation data for my compounds. Thanks must also be given to Amy Webb, for ensuring the smooth running of the labs and help fixing HPLC machines, and Wendy Offen (Prof. Gideon Davies lab, University of York) for setting up all the crystallisation trials and regularly updating me on these.

Dr Jens Willwacher, Dr Maria Espling, Dr Robert Vallée, Dr Charles Fehl, Dr Bhaskar Bhushan and Matthew Bilyard all very kindly proof read this Thesis and made it a lot more readable than it otherwise would have been. Thank you.

The most amazing part of my DPhil was all the fantastic people I meet and worked alongside. As well as those mentioned above, almost everyone I have met in the group

has helped me at some point and made my DPhil a very enjoyable experience. Particular thanks has to be given to Sébastien Galan, Ritu Raj and the rest of “Night Watch” and members of LG10. You guys put up with my bad jokes, my moaning and complaining and my appalling attempts at foreign languages. I will never forget the laughter and great times we had.

Finally, the greatest thanks has to be given to my Wife and my Family for giving me love and support during this endeavour. Thank you Emma for putting up with me coming home late at night, disappearing at the weekends and constant talk about reaction yields and *E. coli*. Without your support, none of this would have been possible. This is as much yours as mine.

Contents

Abstract.....	i
Author Declaration	v
Publications	vi
Acknowledgements	vii
Contents.....	ix
Abbreviations, Notations and Nomenclature.....	xiii
Chapter 1 – Introduction	1
1.1 - Glycobiology – General Considerations.....	2
1.2 – Structure Glycans and Glycoproteins in Nature	5
1.2.1 – Introduction	5
1.2.2 - <i>N</i> -Glycans.....	5
1.2.3 - <i>O</i> -Glycans	9
1.3 –Mechanisms for Glycan Processing Enzymes	14
1.3.1- Introduction	14
1.3.2 – Glycosyltransferase	14
1.3.3 - Glycoside Hydrolase	18
1.4 - Synthesis of Homogeneous Glycoproteins	22
1.4.1- Introduction	22
1.4.2 - Chemical Synthesis	22
1.4.3 - Unnatural Amino Acid Incorporation, Chemical Mutagenesis and Post- Translational Modification	25
1.4.4 - Enzymatic Glycoprotein Synthesis	28
1.4.5 - Engineering Glycosylation Pathways.....	33
1.5 - Aims of this Thesis	34
1.5.1 – Introduction	34
1.5.2 - Outline of Thesis	34
1.6 – References.....	36

Chapter 2 – Investigating EndoS;

An Antibody Specific Endo Enzyme.....	43
2.1 – Introduction	44
2.2 – Formation of an Optimised EndoS Expression and Purification System	49
2.3 – Initial Formation and Testing of EndoS Mutants	52
2.4 – Building an EndoS Model and Forming Further Mutants.....	58
2.5 – Attempts to Obtain a Crystal Structure of EndoS	66
2.6 – Conclusions and Future Work	74
2.7 - Biological Experimental.....	76
2.7.1 – General Protocols	76
2.7.2 – Mutagenesis on EndoS Plasmid and Plasmid Preparation	77
2.7.3 – EndoS and EndoS Mutants Expression and Purification.....	80
2.7.4 – IgG1 Fc Expression	83
2.8 – Protein Sequences	86
2.8.1 – EndoS	86
2.8.2 – IgG1 Fc	87
2.8.3 – EndoS/EndoF ₃ Amino Acid Sequence Alignment	88
2.9 – EndoS Mutant Catalysed Reactions	89
2.9.1 – Hydrolytic Activity of EndoS Mutants.....	89
2.9.2 – Glycosylation Activity of EndoS Mutants	89
2.10 – Crystallisation Trials of EndoS	90
2.11 – References.....	92

Chapter 3 – Investigations into the Activity of PglB *In Vitro* for the Synthesis of Glycopeptides

3.1 – Introduction	96
3.2 – PglB Substrate Synthesis	102
3.2.1 – Lipid Synthesis	102
3.2.2 – Coupling of the Lipid to the Sugar, Forming the PglB Substrate.....	106
3.3 – PglB Expression and Initial Testing	110
3.3.1 – PglB Expression	110
3.3.1 – Initial Testing of PglB	111

3.4 – Testing the Newly-Formed PglB Substrates	114
3.5 – Further Enzymatic Glycosylation Reactions After PglB	116
3.5.1 – Formation of the Peptide-GlcNAc Acceptor with PglB and Further Glycosylation with GalT	116
3.5.2 – EndoA Glycosylation	120
3.6 – Conclusions and Future Work	129
3.7 - Biological Experimental.....	130
3.7.1 - PglB Expression and Purification	130
3.7.2 – Formation of EndoA Mutants and Expression and Purification of Wild Type EndoA and Mutants.....	131
3.8 – Protein Sequences	136
3.8.1- PglB.....	136
3.8.2 – EndoA.....	137
3.9 - Chemistry Experimental.....	139
3.9.1 – General	139
3.9.2 – Compounds.....	141
3.10 - Glycosylation Reactions.....	175
3.10.1 - Investigating the Activity of PglB.....	175
3.10.2 – Investigating the Effect of Different Lipid Lengths on the Reactivity of the PglB Substrate	175
3.10.3 - <i>N</i> -Glycosylation of TAMRA-DANYTK with PglB (3.29).....	176
3.10.4 - Glycosylation of TAMRA-GlcNAc Peptide with GalT (3.30).....	177
3.10.5 - Glycosylation of a Model System with GST-EndoA (3.32)	177
3.10.6 - Glycosylation of TAMRA-GlcNAc Peptide with GST-EndoA (3.33).....	178
3.10.7 - Glycosylation of TAMRA-GlcNAc Peptide with GST-EndoA Mutants.....	178
3.11 – References.....	180

Chapter 4 – Synthesis and Testing of PglB Substrates with Variation on the Sugar; Investigating 4-Thiol GlcNAc	183
4.1 – Introduction	184
4.2 – Thiol GlcNAc PglB Substrate Synthesis	190
4.3 – Testing the Thiol-GlcNAc Substrates with PglB	198

4.4 – Investigating EndoA Glycosylation Activity on a Thiol Acceptor	206
4.4.1 – Formation of a Test System.....	206
4.4.2 – Investigation of EndoA Catalysed Glycosylation with the Model System	208
4.5 – Towards the Synthesis of PglB Substrates Modified at the 3-Position.....	213
4.5.1 – Targets and Synthetic Strategy	213
4.5.2 – Work Towards Disaccharide Synthesis	215
4.5.3 – Work Towards the Formation of a Tagged GlcNAc	217
4.6 – Conclusions and Future Work	220
4.7- Chemistry Experimental.....	222
4.8 - Glycosylation Reactions.....	260
4.8.1 – Investigating the Use of WZ ₃ -Thiol-GlcNAc (4.15) as a PglB Substrate	260
4.8.2 – Initial Investigations into the Use of WEZ ₄ -Thiol-GlcNAc (4.17) as a PglB Substrate	260
4.8.3 – Further Investigations into the Use of WEZ ₄ -Thiol-GlcNAc (4.17) as a PglB Substrate	261
4.8.4 - Glycosylation of a Thiol Model System with GST-EndoA (3.32)	261
4.9 – References.....	262

Chapter 5 – Investigating the Use of a Glycal Substrate with Endo Enzymes	267
5.1 – Introduction	268
5.2 - Glycal Synthesis.....	271
5.3 – Formation of EndoA Mutants	274
5.4 – Investigating Endo Catalysed Glycosylation with the Glycal Donor.....	275
5.5 – Conclusions and Future Work	281
5.6 - Biological Experimental.....	286
5.7- Chemistry Experimental.....	287
5.8 – Chemo-Enzymatic Experimental	296
5.8.1 – Investigating Endo Glycosylation Reactions Using a Glycal Donor	296
5.8.2- Checking EndoA Mutants Activities and the Effect of the Glycal on Activity ..	296
5.8.3 – Investigating Endo Catalysed Hydrolysis of the Glycal.....	297
5.9 – References.....	298

Abbreviations, Notations and Nomenclature

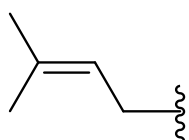
Amino Acids

A	Ala	Alanine
R	Arg	Arginine
N	Asn	Asparagine
D	Asp	Aspartic Acid
C	Cys	Cysteine
E	Glu	Glutamic Acid
Q	Gln	Glutamine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
L	Leu	Leucine
K	Lys	Lysine
M	Met	Methionine
F	Phe	Phenylalanine
P	Pro	Proline
S	Ser	Serine
T	Thr	Threonine
W	Trp	Tryptophan
Y	Tyr	Tyrosine
V	Val	Valine

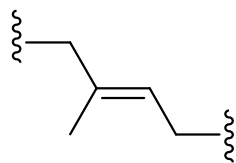
Nucleic Acids

A	Adenine
C	Cytosine
G	Guanine
T	Thymine
U	Uracil

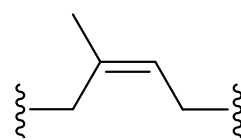
Lipid Structures



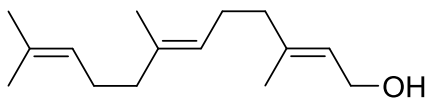
W



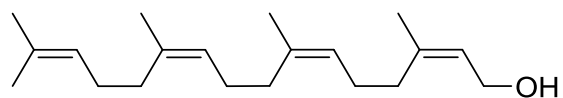
E



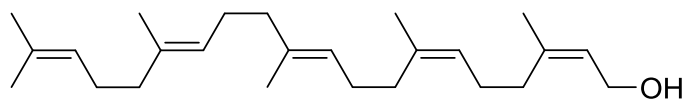
Z



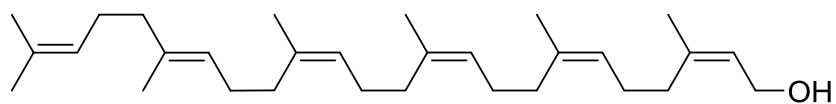
WE₂-OH



WZ₃-OH



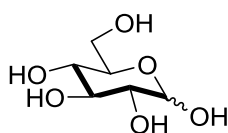
WE₂Z₂-OH



WEZ₄-OH

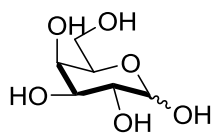
Sugar Notations and Nomenclature

Sugars are symbolised according to the Consortium for Functional Glycomics notations and nomenclatures and structures drawn using this nomenclature with GlycoWorkbench.



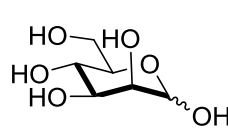
Glucose

Glc



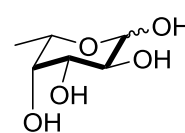
Galactose

Gal



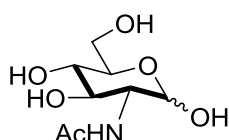
Mannose

Man



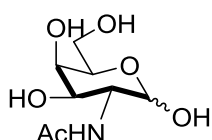
Fucose

Fuc



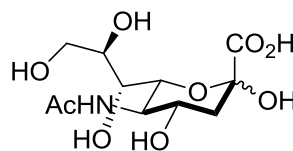
N-Acetylglucosamine

GlcNAc



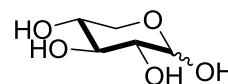
N-Acetylgalactosamine

GalNAc



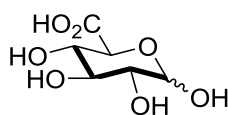
N-Acetylneuraminic Acid

Neu5Ac



Xylose

Xyl



Glucuronic Acid

GlcA

All monosaccharides are assumed to be in the D-configuration except fucose which is in the L-configuration.

All glycosidically-linked monosaccharides are assumed to be in the pyranose form.

All glycosidic linkages are assumed to originate from the 1-position, except sialic acids, which originate from the 2-position.

General Abbreviations

Δ	Reflux/Heating
Ac	Acetyl
ALG	Altered in Glycosylation Enzyme
Amp	Ampicillin
BCA	Bicinchoninic Acid
BCIP/NBT	5-Bromo-4-Chloro-3-Indolyl Phosphate/Nitro Blue Tetrazolium
Bn	Benzyl
BSP	1-Benzenesulfinyl Piperidine
Bu	Butyl
BSA	Bovine Serum Albumin
CDI	Carbonyldiimidazole
CMP	Cytidine Monophosphate
CSA	Camphor Sulfonic Acid
d	doublet
Da	Dalton
DHP	Dihydropyran
DIPEA	<i>N,N</i> -Diisopropylethylamine
DMEM	Dulbecco's Modified Eagle's Medium
DMF	Dimethylformamide
DMPU	Dimethylpropylene Urea/1,3-Dimethyl-3,4,5,6-Tetrahydro-2(1H)- Pyrimidinone
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DNP	Dinitrophenol
dppp	1,3-Bis(diphenylphosphino)propane
DTBMP	2,6-Di-Tert-Butyl-4-Methylpyridine
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EEDQ	<i>N</i> -Ethoxycarbonyl-2-Ethoxy-1,2-Dihydroquinoline
Endo	Endo- β -N-acetylglucosaminidase
EPO	Erythropoietin

ER	Endoplasmic Reticulum
ESI	Electrospray Ionisation
Et	Ethyl
Et ₂ O	Diethyl ether
EtOAc	Ethyl Acetate
Equiv	Equivalents
Fab	Fragment Antigen-Binding
FBS	Fetal Bovine Serum
Fc	Fragment Crystallizable
Fmoc	Fluorenylmethyloxycarbonyl
FPLC	Fast Protein Liquid Chromatography
g	Gravitational Force
GalT	Galactose Transferase
GDP	Guanosine Diphosphate
GH	Glycoside Hydrolase
GST	Glutathione S-Transferase
GT	Glycosyltransferase
HEK	Human Embryonic Kidney
HEPES	2-[4-(2-Hydroxyethyl)Piperazin-1-yl]Ethanesulfonic Acid
HIV	Human Immunodeficiency
HPLC	High Performance Liquid Chromatography
HRMS	High Resonance Mass Spectrometry
Hz	Hertz
IgG	Immunoglobulin G
IPTG	Isopropyl-β-D-Thiogalactopyranoside
IR	Infrared (Spectrometry)
LB	Lysogeny Broth
LCMS	Liquid Chromatography Mass Spectrometry
m	multiplet
Maj	Major
MALDI	Matrix Assisted Laser Desorption/Ionization
mCPBA	<i>meta</i> -Chloroperoxybenzoic Acid

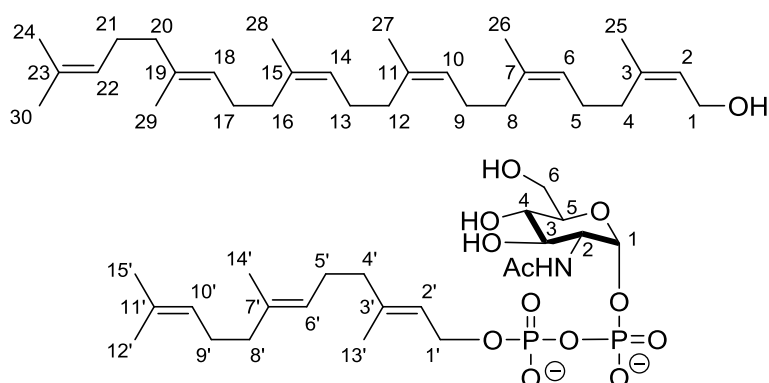
Me	Methyl
MES	2-(N-Morpholino)Ethanesulfonic Acid
Min	Minor
MOPS	3-(N-Morpholino)Propansulfonic Acid
MP	Melting Point
MS	Mass Spectrometry/Molecular Sieves
Ms	Mesyl
Mw	Molecular weight
MWCO	Molecular Weight Cut-Off
<i>m/z</i>	Mass/Charge
NBS	<i>N</i> -Bromosuccinimide
NCL	Native Chemical Ligation
NiNTA	Nickel Nitrilotriacetic Acid
NMR	Nuclear Magnetic Resonance
OD	Optical Density
O-GlcNAcase	O- β - <i>N</i> -Acetylglucosaminidase
OGT	O-GlcNAc Transferase
OST	Oligosaccharyltransferase
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
Pen/Strep	Penicillin-Streptomycin
Pet. ether	Petroleum Ether (40 – 60 °C fraction)
Ph	Phenyl
<i>R_f</i>	Retention Factor
RNA	Ribonucleic Acid
rpm	rotations per minute
s	singlet
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
S _N 2	Substitution nucleophilic bimolecular
S _N i	Substitution nucleophilic internal
SOC	Super Optimal Broth with Catabolite Repression
SPPS	Solid Phase Peptide Synthesis

t	triplet
TAMRA	Carboxytetramethylrhodamine
TB	Terrific Broth
TBAF	Tetra- <i>N</i> -butylammonium Fluoride
TBDMS	<i>tert</i> -Butyldimethylsilyl
TBDPS	<i>tert</i> -Butyldiphenylsilyl
TCEP	<i>tris</i> (2-Carboxyethyl)Phosphine
TEV	Tobacco Etch Virus Protease
Tf	Triflate/Trifluoromethanesulfonate
THF	Tetrahydrofuran
THP	Tetrahydropyranyl
Tr	Triphenylmethyl
tRNA	Transfer Ribonucleic acid
Tris	2-Amino-2-Hydroxymethyl-Propane-1,3-Diol
Ts	Tosyl
UDP	Uridine Diphosphate
UV	Ultraviolet (light)
V	Voltage
$v_{\max}$	Absorbencies
v/v	percentage volume per volume
w/v	percentage weight per volume

NMR Numbering and Assignments

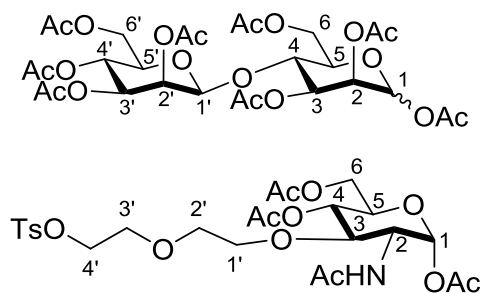
Given below are examples of the numbering system used for the assignment of the ^1H and ^{13}C NMR. In general for the lipids, numbering goes along the chain from right to left, before labelling methyl groups off the chain, again right to left.

Examples



For carbohydrates, numbering starts on the reducing end saccharide and numbered with accordance to standard convention. Hyphens are used on subsequent saccharides or other groups attached, as seen above for the diphosphate bound lipid.

Examples



For further information about NMR assignments, see Chapter 3 – 3.9.1 – Chemistry Experimental – General.

Chapter 1 – Introduction

Chapter 1 - Introduction

1.1 - Glycobiology – General Considerations

Arguably the greatest achievement in chemical biology and molecular biochemistry over the last 20 years was the completion of the Human Genome Project and the complete sequencing of the *Homo sapiens* genome.^{1, 2} The immediate surprising result from this endeavour was how few genes humans have; from the initial prediction of around 2 million, the number of genes actually found in the human genome is approximately 21000, comparable in size to that of the genome of a dog³ (25000 genes) and a mouse⁴ (23000 genes) and about 1.5 times the size of that of a fruit fly⁵ (13,600 genes). Furthermore, it was found that about 60% of the fruit fly genes are shared with the human genome and it is a commonly known fact that humans share $\geq 30\%$ of their DNA with Bananas.⁶

From these findings, the obvious question is, how does this fit with the central paradigm of biochemistry that DNA and genes code for proteins which lead to the complexity of living organisms when the relatively simple fruit fly and banana share so much with humans? A significant part of the answer to this question is that human proteins often contain post-translational modifications, which add another degree of complexity and allows for variation between proteins that contain the same primary structure.

There are a large variety of different post-translational modifications that have been reported including phosphorylation, acetylation and ubiquitylation.⁷ One of the most important post-translational modifications is glycosylation; the addition of sugar residues. Protein glycosylation is one of the most common and a vital post-translational

modification that can occur on a protein, with over 50% of all human proteins being glycosylated.⁸ It has been implicated in a wide range of important biological functions, including but not limited to: membrane structure and stability, intercellular communication, transportation and adhesion, as well as inflammatory responses.⁹⁻¹¹ Glycosylation is also imperative to the functioning of the immune system, with the *N*-glycan attached to the heavy chain of an antibody insuring the Fragment Crystallizable (Fc) region on the antibody is able to bind to Fc receptors found on B lymphocytes, natural killer cells, macrophages and other cells in the immune system, allowing the clearance of pathogens.¹¹

Due to their importance in living systems, glycoproteins have also been implicated in a number of diseases, including cardiovascular disease,^{12, 13} immunology disorders^{11, 14} and Alzheimer's disease^{13, 15} amongst others. A specific example of how protein glycosylation is important in disease is that of rheumatoid arthritis. In patients suffering from rheumatoid arthritis, the IgG have a decreased amount of terminal β -linked galactose with some IgG containing no terminal galactose at all. The severity of disease has a negative correlation with the amount of galactosylation of the IgG, with the disease subsiding during pregnancy due to an increase in galactosylation of the IgG.^{11, 13} Protein glycosylation is also extremely important in oncology.^{13, 16, 17} In cancer cells, the enhanced expression of GlcNAc transferase V results in an increase in the β -1,6 branching of *N*-glycans, leading to the observation that tumour cell derived glycopeptides are larger than normal cell derived glycopeptides. Cells with increased GlcNAc transferase V expression are linked with an increased incident of metastasis.¹⁷ The exact mechanisms that cause this phenomenon are unknown, although it is

proposed that this β -1,6 branching may increase interaction with lectins or otherwise the β -1,6 branching may cause a general biophysical effect due to its conformation.^{13,17}

As glycoproteins are involved in a variety of complex and critical processes in biology, chemical glycobiology is “emerging as one of the most uniquely powerful sub-disciplines of chemical biology”.¹⁸ One of the key challenges in chemical glycobiology, however, is the formation of the glycoproteins and glycans to be studied. It is extremely challenging to purify homogenous glycoproteins to investigate the effect that the glycan and its structure has on the activity of the protein as a whole, meaning that glycoproteins often have to be synthesised in order to be investigated. Unlike peptides and nucleic acids, which have a linear based structure and can easily be formed using a sequential series of reactions to couple the building blocks to give the desired structure, glycans and their attachment to a protein is more challenging due to their non-linear, branching structure and the large variety of different glycan structures present in biology, as well as the possibility of modifications to the glycans, such as sulfation.

1.2 – Structure of Glycans and Glycoproteins in Nature

1.2.1 - Introduction

As stated above, the major difficulty in the synthesis of glycans/oligosaccharides is that glycans have no defined linear structure and often exist as branched structures. The complexity of glycans can therefore be seen to be twofold: first the different saccharide units that constitute the mature glycan and the order of these, and secondly, the position of the linkage between the saccharide units and stereochemistry of these linkages. Moreover, when regarding glycoproteins, it should be considered that the glycan chain could be attached to a number of different amino acids at a number of different positions on the protein, giving another degree of complexity to reflect on when thinking about the synthesis of glycoproteins. It would seem that there are almost infinite possibilities for the glycan structure on a glycoprotein. In reality, however, nature chooses a limited number of structures and positions at which the glycan is attached to the protein. The vast majority of glycans are attached to the side chains of three amino acids: asparagine to give an *N*-glycan, and serine or threonine to give an *O*-glycan. Furthermore, the structures of *N*-glycans and *O*-glycans are very distinctive and these are discussed below.

1.2.2 - *N*-Glycans

For *N*-glycans, all structures have a core pentasaccharide unit, made up of two GlcNAc units and three mannose units ($\text{Man}_3\text{GlcNAc}_2$). The branching off this core structure gives the three types of *N*-glycans; either high mannose, complex or hybrid structures (Figure 1.1). These basic structures can be further modified by the presence of either a bisecting GlcNAc residue and/or a core fucose residue. There are also limits on the

location of *N*-glycans on proteins; a minimum recognition sequence of Asn-X-Ser/Thr is required, where Asn is the amino acid that the glycan is attached to and X is any amino acid except proline.¹¹

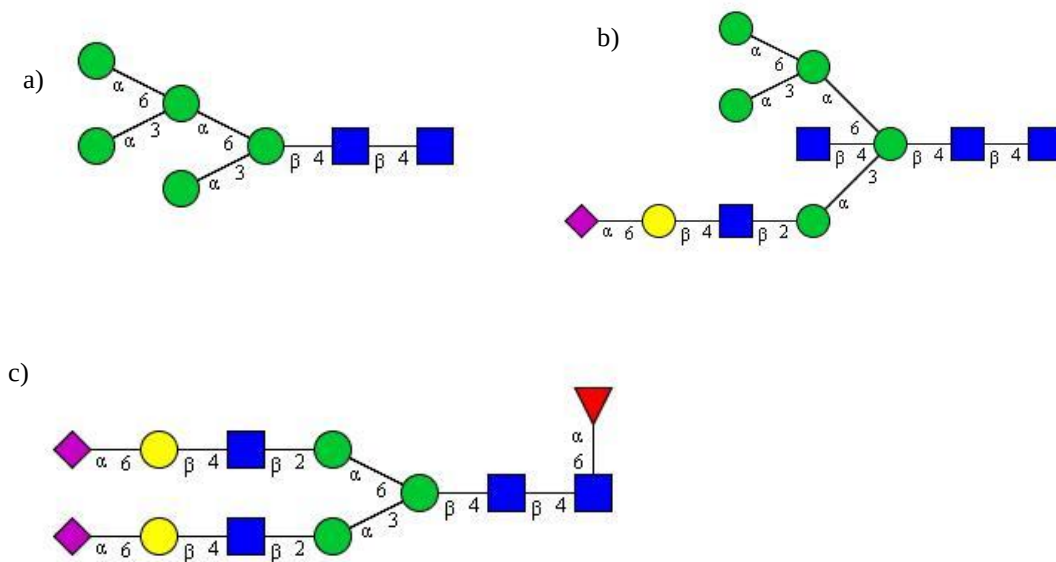
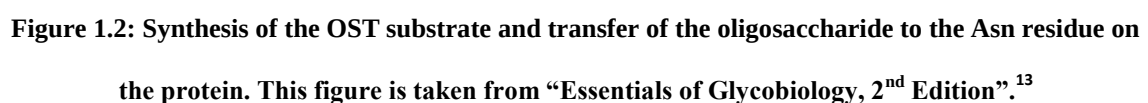


Figure 1.1: Structures of *N*-Glycans, a) High mannose glycan, b) Hybrid glycan with bisecting GlcNAc unit and c) Complex glycan with core fructose.¹³

With regards to the synthesis of *N*-glycans in mammals, and more specifically the synthesis of complex glycans, it can be thought to occur in two stages; first the formation of the oligosaccharyltransferase (OST) substrate and transfer of the glycan to the protein and then the maturation and the derivation of the glycan.^{13, 19} For the first step, the OST substrate that transfers the glycan is a dolichol pyrophosphate *N*-glycan and this is formed initially on the cytoplasm face of the endoplasmic reticulum (ER, Figure 1.2). Dolichol is a polyisoprenol lipid made up of five carbon isoprene units, where the number of units varies depending on the organism involved; the exact form of the dolichol unit will be further discussed in Chapter 3. The synthesis of the OST substrate begins with the transfer of GlcNAc-phosphate on dolichol phosphate to give



Once this transfer has occurred, the second stage, the trimming and derivation of the glycan, takes place (Figure 1.3). This begins with removal of the glucose units through α -glucosidases I and II in the ER. The glycoprotein then folds into its tertiary structure and may lose one mannose unit, via α -mannosidase I and then leaves the ER and enters the Golgi apparatus. Here, more mannose units are lost with hydrolysis which is catalysed by a variety of mannosidase enzymes to give the critical $\text{Man}_5\text{GlcNAc}_2$ glycan intermediate; this is the key building block to form the hybrid and complex glycans. From this point, a variety of processing steps are possible, using both glycoside hydrolase and glycosyltransferase enzymes to form the desired final glycan attached to the glycoprotein. These steps all occur using specific enzymes for a specific step, with each step occurring in a specific location in the Golgi apparatus.

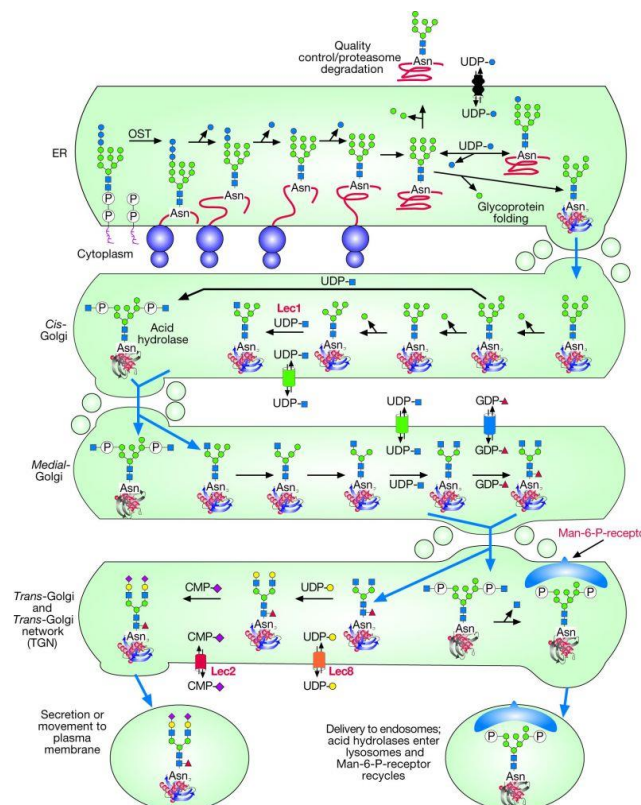


Figure 1.3: Trimming and derivation of the *N*-glycan attached to the protein to form the Complex glycan. This figure is taken from “Essentials of Glycobiology, 2nd Edition”.¹³

The exact details of these glycoprocessing enzymes and the glycan structures that they require is beyond the scope of this thesis. However, the key point, as stated above, is that specific enzymes add specific sugar units and the subsequent actions of many of the enzymes rely on the presence of specific residues and branches on the glycan and the whole synthesis of the glycan is tightly controlled.

N-glycans are a critical feature of biology. The deletion of the gene that encodes for GlcNAcT-I, which is necessary for the synthesis of hybrid and complex *N*-glycans, in mice results in death during the embryonic stage.¹³ *N*-glycans have also been implicated in cell proliferation, cell signalling protein folding, the secretory pathway and cancer.^{13,}

20, 21

1.2.3 - O-Glycans

1.2.3.1 - Introduction

The main focus of this work will be on *N*-glycans. However, the importance of *O*-glycans means that they should not be ignored and some brief features and examples and their synthesis will be described here. In contrast to *N*-glycans, *O*-glycans are much more varied in structure. First, where *N*-glycans only appear on asparagine residues, *O*-glycans occur on either serine or threonine. Also, there is no fixed consensus sequence that is required for *O*-glycosylation and, where *N*-glycans always have a GlcNAc residue attached to asparagine, the sugar unit attached to the serine or threonine is varied. Sugar units that can be attached to a serine or threonine include *O*-GalNAc, *O*-GlcNAc, *O*-xylose, *O*-mannose, *O*-fucose and *O*-glucose amongst others.¹³ The large variety of variation in *O*-glycan structures and functions make it impossible to fully

describe all features here. A brief overview will be given of two of the most important *O*-glycosylations and glycoproteins; mucin and proteoglycans.

1.2.3.2 - Mucin

Mucins are highly *O*-glycosylated proteins that are found in mucous secretions from many epithelial cells, including the cells of the tracheobronchial and gastrointestinal tracts, where they protect against chemical and physical damage and also protect against infection from pathogens.^{13, 22} The *O*-glycans attached to mucins have four common core structures (denoted core structures 1 – 4) and there are another four less common structures (denoted core structures 5 – 8).^{9, 13, 23} All these core structures contain an *N*-acetyl galactosamine (GalNAc) attached to the serine or threonine via an α -linkage. The core structures can then be further derivatised by the addition of a number of different sugar units and other modifications, such as *O*-acetylation of neuraminic 5Ac acid and *O*-sulfation of galactose and GlcNAc.¹³ It can, therefore, be appreciated that the *O*-glycans on mucin are highly heterogeneous.

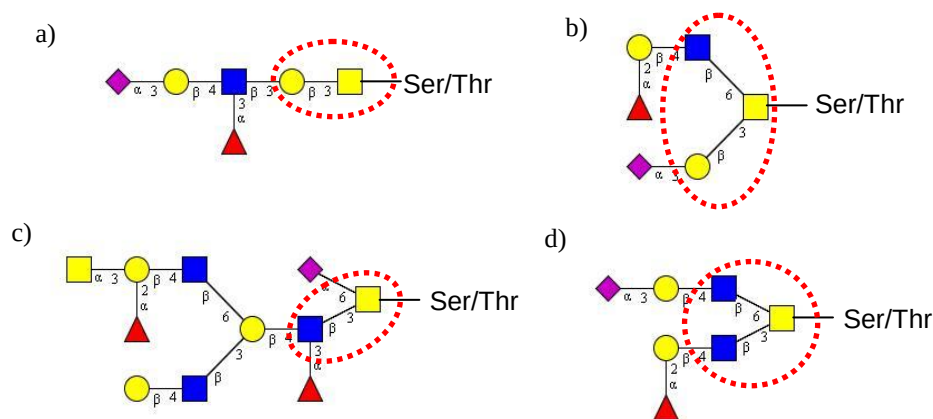


Figure 1.4: Examples of structures of mucin *O*-glycans, a) Core structure 1, b) Core structure 2, c) Core structure 3 and d) Core structure 4. Note: All core structure are circled in red, with further derivation also being shown.¹³

The synthesis of the O-glycans attached to mucin does not occur as with N-glycans via block transfer of an oligosaccharide, followed by modification catalysed by glycoside hydrolase and glycosyltransferase enzymes. Instead, each individual sugar unit is attached to the protein/glycan structure through glycosyltransferase enzymes. As stated above, all O-glycans attached to mucin contain a GalNAc attached to the serine or threonine via an α -linkage; this addition is catalysed by polypeptide-N-galactosaminyltransferase^{13, 24} and occurs in the Golgi. Once the initial GalNAc has been attached, glycosyltransferase enzymes then catalyse the addition of further saccharide units to give the desired structure. The formation of the oligosaccharide is controlled by which glycosyltransferases are present, as well as the specificity of the glycosyltransferase for certain glycan acceptors.

1.2.3.3 - Proteoglycan

Proteoglycans consist of a protein attached to a large glycan formed of repeating saccharide units, leading to glycans of significant size (approx. 80 sugar residue, 20 kDa in weight).^{13, 25, 26} Proteoglycans can have multiple chains attached, meaning that a significant amount of the mass of the protein, if not the majority, is made up of the sugar chains. Examples of these repeating saccharide chains include keratin sulfate, heparan sulfate and chondroitin sulfate^{13, 26} (Figure 1.5). Functions of proteoglycans include their ability to bind water and hence hydrate the intercellular space, the binding of growth factors and mediating cell to cell adhesion, amongst other functions.^{13, 27}

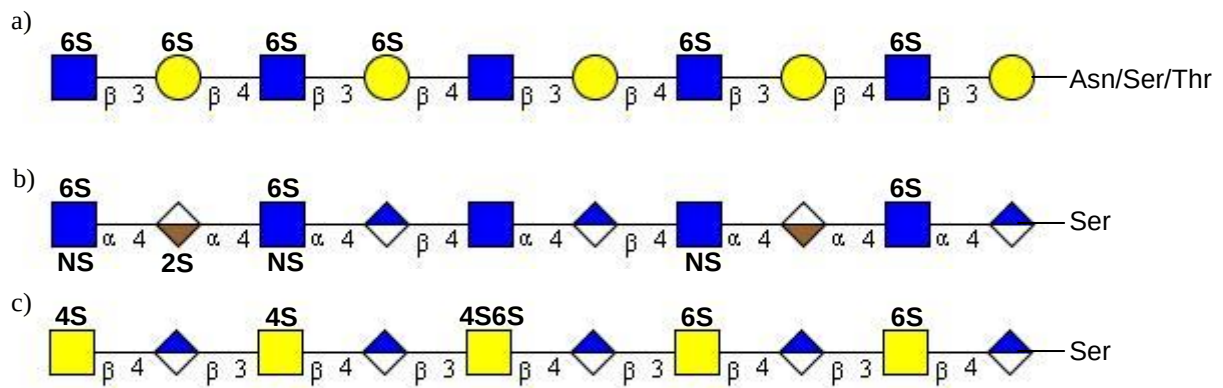


Figure 1.5: Structures of some proteoglycans, a) keratan sulfate, b) heparan sulfate and c) chondroitin sulfate. Note: 4S = sulfate group on the 4-position, 6S = sulfate group on the 6-position and NS = sulfate group attached to the nitrogen atom.¹³

The synthesis of these oligosaccharides is similar as described for the mucin glycans. The keratan sulfate chain is attached to either an *N*-glycan (replacing the sialic acid with the keratan sulfate chain) or the core mucin glycan structure. The *N*-glycan or the *O*-glycan is synthesised as described before, and then glycosyltransferase enzymes catalyse the formation of the oligosaccharide chain.¹³ Sulfotransferases catalyse the attachment of sulfate groups to give the keratan sulfate chain.^{13, 26} For heparan sulfate and chondroitin sulfate, the oligosaccharide chain is attached to a serine residue via a xylose unit, the addition of which is catalysed by xylosyltransferase. Once the xylose is installed, specific glycosyltransferase enzymes again catalyse the formation of the oligosaccharide chain. Whether chondroitin sulfate or heparan sulfate is formed is determined by the amino acids in the proximity of the attachment site. The first GlcNAc residue in heparan sulfate is attached via an enzyme called EXTL3 and this enzyme normally acts on sites where there are groups of acidic amino acids near the attachment site.^{13, 26} As with keratan sulfate, other modifications take place on chondroitin sulfate and heparan sulfate (phosphorylation, sulfation and epimerisation of sugar units) that

are enzymatically catalysed, resulting the formation of the mature chondroitin sulfate and heparan sulfate.^{13, 25, 26}

1.3 –Mechanisms for Glycan Processing Enzymes

1.3.1- Introduction

Glycans attached to proteins can come in many forms and are imperative for the functioning of cell and life itself. The synthesis of these glycans and how each addition and hydrolysis of a sugar unit is catalysed by an enzyme have also been described. These enzymes have been only briefly discussed so far and no information has been given on the enzyme substrate or the mechanism of their activity. Enzymes that modify the glycan (i.e. sulfotransferase, acetyltransferase amongst others) will not be considered here and only the activity of glycosyltransferase (transfer of a saccharide unit) and glycoside hydrolase (the hydrolysis and removal of a saccharide unit) will be explored.

1.3.2 - Glycosyltransferase

As stated above, glycosyltransferases catalyse the addition of saccharide units to either an amino acid/protein backbone, lipid, another saccharide or other aglycone and hence form a glycosidic bond via a glycosylation reaction. For a glycosylation reaction to take place, whether that is enzymatic or chemical, a glycosyl donor (the molecule that is donating the saccharide to be attached) and a glycosyl acceptor (the molecule which the saccharide unit is to be attached to) must be present. Nucleophilic attack of the glycosyl acceptor on the glycosyl donor results in the formation of the glycosidic bond. Glycosyl acceptors, as stated above, can be either sugars or the side chain of an amino acid. Glycosyl donors are most often activated sugar nucleotides, where the saccharide has a mono/diphosphonucleotide attached at the anomeric position⁹ (Figure 1.6). The most common diphosphonucleotide to be attached is uridine diphosphate (UDP), although

other nucleobases are known to be used in glycosyl donors (such as guanosine and cytidine) and sometimes the nucleosugar may be attached via one phosphate group as opposed to two.⁹ The presence of the phospho group on the glycosyl donor is critical; this is a facile leaving group and hence allows the glycosylation to occur. Another important class of glycosyl donors are polyprenol phosphates/pyrophosphates, where a lipid is attached to the saccharide via the phosphate(s).^{28, 29}

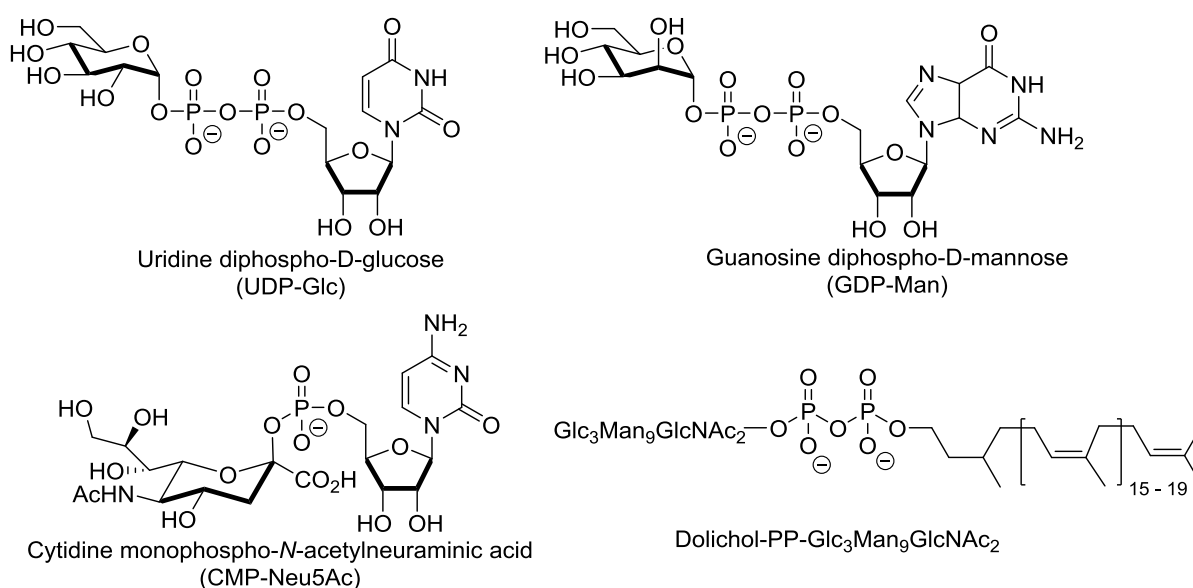
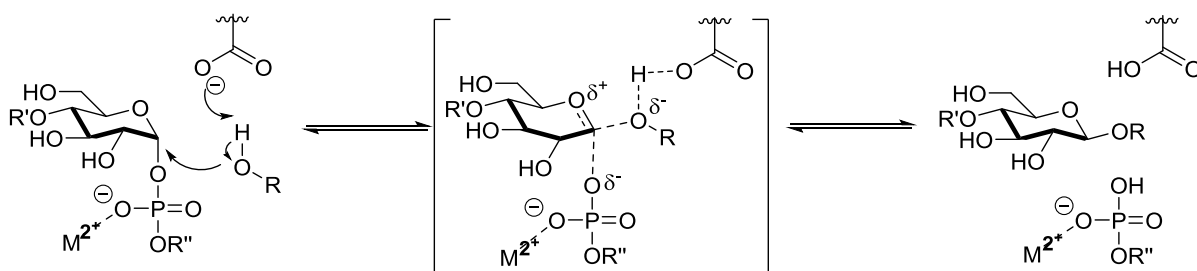


Figure 1.6: Examples of substrates for glycosyltransferase enzymes.

There are two possible outcomes for the glycosyltransferase catalysed glycosylation reaction; either the configuration at the anomeric centre on the glycosyl donor is retained (retaining glycosyltransferase) or the configuration is inverted (inverting glycosyltransferase). Both of these are the result of the different mechanisms that the glycosyltransferase employs.

For inverting glycosyltransferases, structural and kinetic data indicate a single nucleophilic substitution step.³⁰⁻³² The enzyme deprotonates the glycosyl acceptor and

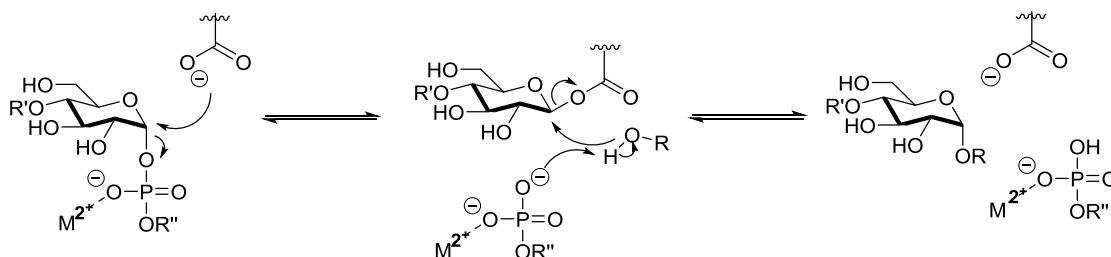
the forming anion then attacks the glycosyl donor, resulting in the loss of the phospho group attached to the donor and the formation of the glycosidic bond³⁰ (Scheme 1.1). The transition state for this reaction is thought to have significant oxocarbenium ion character.



Scheme 1.1: Mechanism for inverting glycosyltransferase enzymes.³⁰

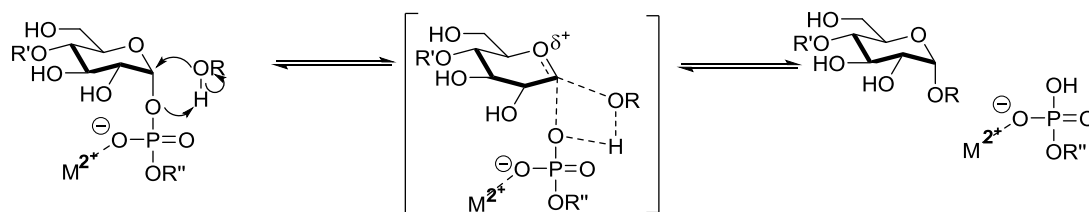
For retaining glycosyltransferase, there is less concrete evidence for the underlying mechanism. One possible mechanism is the Koshland retaining mechanism,^{30, 33} which is also utilised by some glycoside hydrolases (and will be explained further in this chapter). Here, a nucleophilic residue in the enzyme active site attacks the glycosyl donor, resulting in loss of the phospho group and formation of an enzyme-substrate complex, with inversion of configuration at the anomeric centre on the glycosyl donor (Scheme 1.2). This complex is then attacked by the glycosyl acceptor, via a second nucleophilic attack, and hence a second inversion at the anomeric centre of the glycosyl donor, resulting in the formation of the glycosidic bond with retention of configuration at the anomeric centre of the glycosyl donor.³⁰ There is evidence for this mechanism through the identification of possible nucleophilic residues for retaining glycosyltransferase, with mutation of these residues resulting in a loss of activity.^{30, 34} There is also some crystallography evidence,^{35, 36} but as of yet, there is no crystal

structure showing the glycosyltransferase-substrate complex with a covalent bond between enzyme and sugar.



Scheme 1.2: Mechanism for retaining glycosyltransferase enzyme, going via a Koshland retaining like mechanism.³⁰

The alternative mechanism is based on a S_Ni process,^{30, 37, 38} of which there is kinetic data to suggest this is possible³⁹ (Scheme 1.3). In this mechanism, loss of the phospho group and attack of the glycosyl acceptor occur on the same face of the glycosyl donor, with the alternative face blocked by the enzyme.³⁰ It is unknown whether this is a concerted process or goes via an oxocarbenium ion intermediate.

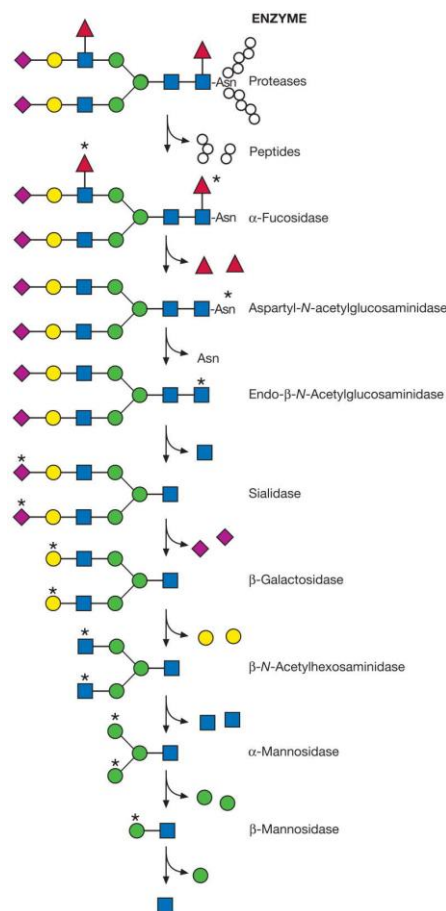


Scheme 1.3: Mechanism for retaining glycosyltransferase enzyme, going via S_Ni mechanism.³⁰

For many but not all glycosyltransferases, a metal co-factor, such as a magnesium or a manganese ion, is required for activity.³⁰ This is mainly for glycosyltransferase that use diphospho type substrates. The metal ion binds to an oxygen on each of the two phosphate groups, as well as side-chain carboxyl groups present in the active site of the glycosyltransferase, presumably to help bind the substrate in the active site of the enzyme.³⁰

1.3.3 - Glycoside Hydrolase

As well as being able to form glycans and glycoproteins, it is important that cells are able to break glycans down too. This is important not only for the formation of *N*-glycans, but also for the recycling of sugars to be used elsewhere in the cell. There are a huge number of different glycoside hydrolases covering a large range of activities and specificities and allowing the complete breakdown of a complex *N*-glycan to its constituent monosaccharides (Scheme 1.4).¹³

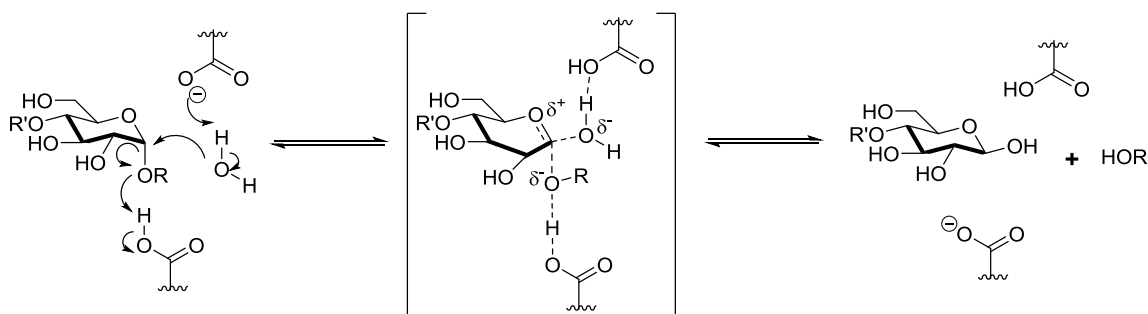


Scheme 1.4: Enzymatic hydrolysis of an *N*-linked complex glycan and glycoside hydrolase enzymes that catalyse this process. This figure is taken from “Essentials of Glycobiology, 2nd Edition”.¹³

The breakdown of the glycan and the hydrolysis of the glycosidic bond between two saccharide units is catalysed by glycoside hydrolases. As with glycosyltransferase, there

are two possible outcomes after the activity of glycoside hydrolases; either the configuration at the anomeric centre on the glycosyl donor is retained (retaining glycoside hydrolase) or the configuration is inverted (inverting glycoside hydrolase). Each of these outcomes results from a different mechanism, with variation possible depending on the configuration at the anomeric centre of the glycosidic bond as well as the neighbouring groups.

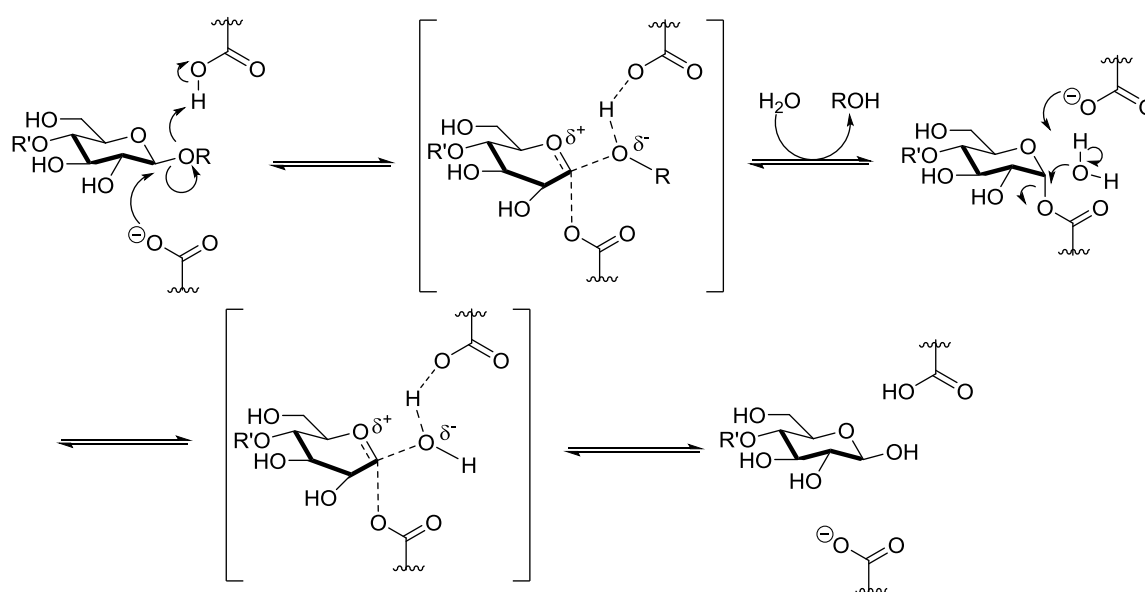
For inverting glycoside hydrolases, the mechanism for hydrolysis occurs very similarly to the mechanism for inverting glycosyltransferase, as a one-step nucleophilic attack, going via an oxocarbenium ion like transition state⁴⁰⁻⁴² (Scheme 1.5). However, instead of a deprotonated glycosyl donor attacking, water acts as the nucleophile. The enzyme deprotonates the water and the forming anion simultaneously attacks the glycosidic bond, resulting in the breaking of this bond. The displaced glycosidic oxygen is then protonated by an acidic residue in the active site.



Scheme 1.5: Mechanism for inverting glycoside hydrolase enzymes.⁴²

For retaining glycoside hydrolases, the classic mechanism is the Koshland mechanism.⁴⁰⁻⁴² This is very similar to what was described for retaining glycosyltransferase. The glycan is initially attacked by a nucleophilic residue in the active site, resulting in the glycosidic bond breaking and the formation of an enzyme-

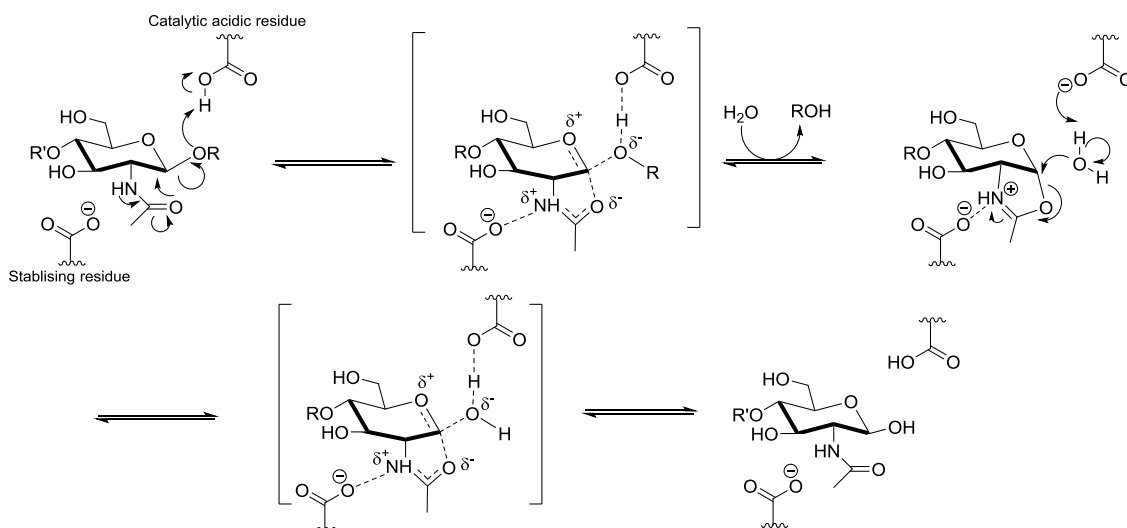
substrate complex (Scheme 1.6). This results in inversion at the anomeric centre. The glycosidic oxygen is protonated by a residue in the active site that is acting as an acid/base catalyst. For the second step of this reaction, this acid/base catalyst residue deprotonates a molecule of water and the forming anion attacks the enzyme-substrate complex, breaking the bond between them and releasing the saccharide. Both steps of this mechanism go via an oxocarbenium-ion like transition state. As with retaining glycosyltransferase that are proposed to go through this mechanism, retention of configuration at the anomeric centre is achieved via double inversion. Unlike for retaining glycosyltransferase, there is mass spectrometric and crystallographic evidence for this mechanism with retaining glycoside hydrolase.⁴³



Scheme 1.6: Mechanism for retaining glycoside hydrolase enzymes.⁴⁰⁻⁴²

An alternative mechanism occurs when hydrolysing an oligosaccharide with a GlcNAc unit at the reducing end. In this case, no nucleophilic residue is required and instead neighbouring group participation is key to the mechanism and hydrolysis.⁴⁴⁻⁴⁶ The enzymes that hydrolyse these types of glycans are termed endo- β -N-acetylglucosaminidase (Endo) and contain a stabilising residue (either a carboxyl or

amide side chain) and a catalytic acid/base residue. The first step of this mechanism is the intramolecular nucleophilic attack of the 2-acetamido group on the GlcNAc, forming an oxazolinium ion intermediate, resulting in the cleavage of the glycosidic bond (Scheme 1.7). The oxazolinium ion intermediate is stabilised by the stabilising residue. The glycosidic oxygen is protonated by the catalytic acid/base residue. For the second step of this reaction, the acid/base residue deprotonates a molecule of water and, as with the retaining mechanism described above, the developing anion attacks the oxazolinium ion intermediate, resulting in the retention of configuration through double inversion.



Scheme 1.7: Mechanism for glycan hydrolysis with endo- β -N-acetylglucosaminidase enzymes.⁴⁴⁻⁴⁶

These are the main mechanism for hydrolysis of glycans with glycoside hydrolases but other mechanisms do exist, such as the mechanism of action of neuraminidase, which goes via an oxocarboanion intermediate⁴⁷. However, these will not be explored here.

1.4 - Synthesis of Homogeneous Glycoproteins

1.4.1- Introduction

The need for the ability to synthesise or otherwise obtain glycoproteins with a homogenous glycan of a known structure is twofold. Throughout this chapter so far, the importance of glycoproteins has been described and how nature attaches glycans to proteins. The complexity of these glycans and some of the different glycan structures have also been explored. However, the effects of certain glycoproteins and, specifically, the glycans attached to the glycoproteins, are unknown. In order to uncover the function of specific glycans on proteins, it is necessary to have a source of homogenous glycoprotein. But glycoproteins are rarely expressed homogeneously within a cell and instead, glycoproteins often consist of heterogeneous mixtures, containing proteins with different glycoforms. This makes it impossible to elucidate which glycan structure shows the desired activity. The second reason why the ability to synthesise homogenous glycoprotein is sought after is that certain glycoforms on certain proteins confer desired characteristics. For example, proteins with *N*-glycans with increased α -2,3 sialylation have increased serum half-lives^{9, 48} and increased antibody-dependent cellular cytotoxicity is observed for monoclonal antibodies lacking core α -1,6 fucosylation.^{9, 49}

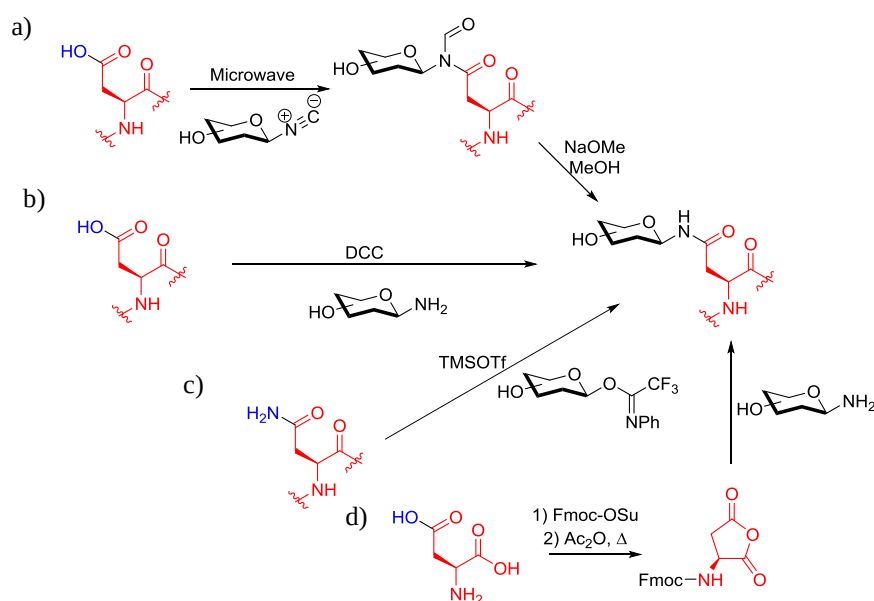
An overview will now be given to some of the methods utilised to form homogenous glycoproteins.

1.4.2 - Chemical Synthesis

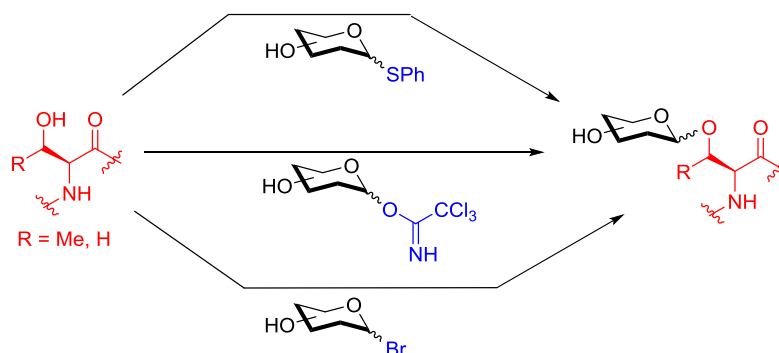
One method is the complete chemical synthesis of the glycoprotein. This is normally achieved using one of two methods; either the glycan is attached to an amino acid,

which can then be utilised in solid phase peptide synthesis (SPPS) and subsequent coupling reactions. Alternatively, a short peptide is formed to which the glycan can be attached and this is subsequently coupled to other peptide fragments to give the glycoprotein.

Considering the coupling of the glycan to amino acids, there are a number of ways that this can be achieved. For *N*-glycans, the classic method is through amide coupling on an activated aspartic acid residue (Scheme 1.8b).⁹ Alternative methods for the formation of *N*-glycans include primary amide glycosylation (Scheme 1.8c), isonitrile-based coupling (Scheme 1.8a) and amidation of aspartic anhydrides⁹ (Scheme 1.8d). For *O*-glycans, the attachment to the serine/threonine is accomplished through standard glycosylation methods⁹ (Scheme 1.9).

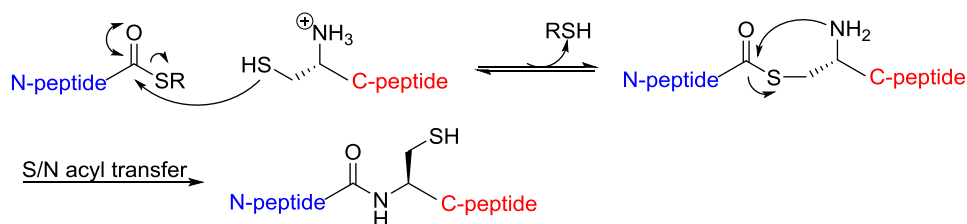


Scheme 1.8: Methods for the chemical formation of *N*-glycans.⁹



Scheme 1.9: Methods for the chemical formation of O-glycans.⁹

As suggested above, glycoprotein synthesis cannot be completed using SPPS alone. Due to the fact that SPPS is not 100% efficient, after each reaction a small amount of material is lost. As most proteins are made up of well over 100 amino acids, this inefficiency becomes a limiting factor. One solution to this is to form a variety of larger peptides and then couple these peptides together. The most common method used to achieve this is Native Chemical Ligation (NCL); the reaction between a cysteine on the N-terminal on one peptide and a thioester on the C-terminal of a second peptide and proceeds via transthioesterification followed by $S \rightarrow N$ acyl shift.^{10, 50, 51} If multiple peptide couplings are required, it is possible to mask a cysteine as a thiazolidine to avoid unwanted homocoupling. An interesting strategy that can be employed with NCL, instead of synthesising all the peptide fragments, is to express peptide fragments in *E. coli* and then employ these fragments in a coupling with a peptide fragment containing the glycan.⁵² Other methods for the coupling of peptides include removable auxiliary-assisted ligation,⁵³ serine/threonine ligation^{54, 55} and serine forming ligations⁵⁶ (although the last of these has not been applied for the synthesis of glycoproteins).



Scheme 1.10: Process and mechanism of Native Chemical Ligation.¹⁰

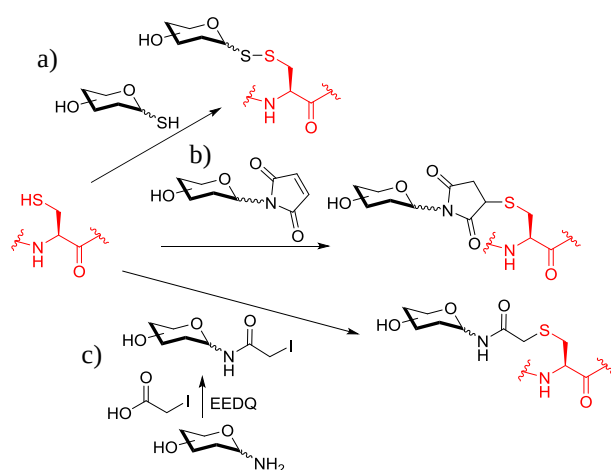
Using this methodology, a number of different glycoproteins have been formed, including, MUC1, a subunit of Human Follicle-Stimulating Hormone,⁵⁷ Glycocin F⁵⁸ and Erythropoietine⁵⁹ amongst others. The weakness of this method is that, even when using coupling methods between large peptides, the synthesis of medium to large proteins is still challenging. Also, after the synthesis of the primary structure of the protein, a folding step has to occur, in order to allow the protein to adopt its 3D shape. It can easily be imagined that it may be challenging to fold some proteins into their correct shape, especially if there are multiple disulfide bonds that are required to form.

1.4.3 - Unnatural Amino Acid Incorporation, Chemical Mutagenesis and Post-Translational Modification

To avoid the challenges with the total synthesis of proteins using synthetic organic chemistry, another method for the synthesis of homogenous glycoprotein is to use biology (and most often *E. coli*) to express and form the protein. In order to attach the glycan, the protein is expressed with a reactive amino acid at the site of glycosylation, which can then be used to react with the glycan to give the desired glycoprotein. This method is termed Chemical Mutagenesis⁶⁰ or “Tag and Modify” approach.⁶¹ The amino acid is usually either a cysteine or an unnatural amino acid bearing a functionality that allows bioorthogonal modification. The disadvantage of this method is that the linkage

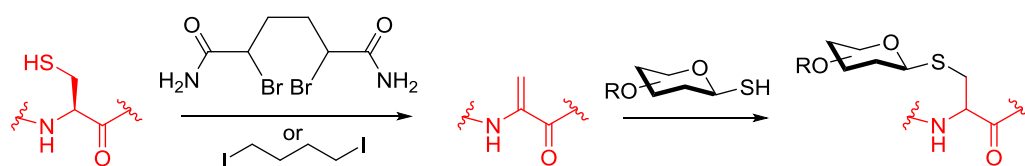
between the protein and amino acid is unnatural and thus, the use of proteins in further experiments may not be reflective of biology in nature, although proteins formed this way are often seen to give accurate information and are good mimics. Other possible disadvantages include that the use of cysteine as a reactive tag may be challenging if other cysteine residues are present and expression levels from unnatural amino incorporation is often poor.

Nevertheless, cysteine often is a good residue for post-translational modification. Due to it being a natural amino acid, cysteine can easily be introduced using site-selective mutagenesis and has the advantage of being the least common amino acid found in proteins, meaning there is a low chance of chemoselectivity issues, although this can be a problem occasionally. Once the cysteine is installed, there are a number of different ways to attach the glycan. These include the formation of a disulfide linkage (Scheme 1.11a), reaction with a maleimide bearing glycan (Scheme 1.11b), reaction of an amine bearing glycan with iodoacetamide, followed by reaction with cysteine (Scheme 1.11c), as well as other methods.¹⁰



Scheme 1.11: Methods for the attachment of glycans to cysteine.¹⁰ Note, EEDQ = *N*-ethoxycarbonyl-2-ethoxy-1,2-dihydroquinoline.

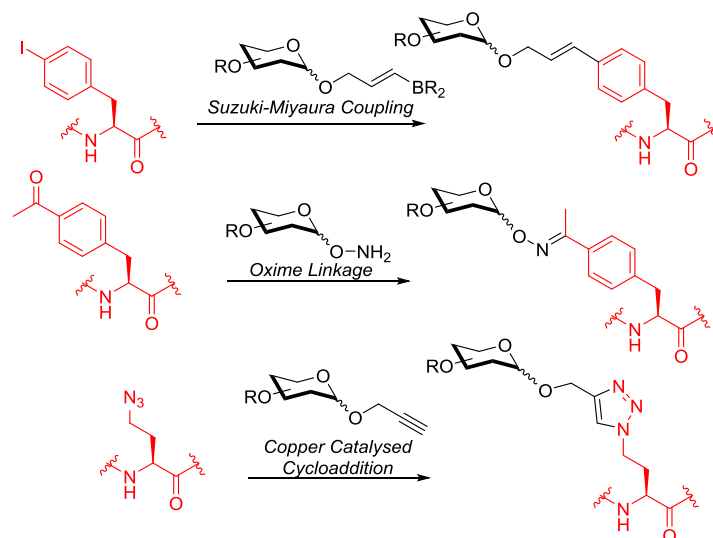
An alternative method for the formation of glycoproteins using cysteine is the formation of dehydroalanine. The reaction of cysteine with 1,4-diiodobutane, 2,5-dibromohexanediamide as well as other reagents,^{10, 62} results in the formation of a double bond on the amino acid. This can then be reacted with a thiol bearing glycan, resulting in the attachment of the sugar (Scheme 1.12). There are a number of reports of this to form glycoprotein analogues.⁶²⁻⁶⁴



Scheme 1.12: Formation of dehydroalanine and addition of glycan.⁶²

The use of unnatural amino acids has the distinct advantage of being unique in the protein and hence cross reactivity is less of an issue, with many of the amino acids and reactions used to install the glycan conferring to Click Chemistry principles.⁶⁵ These unnatural amino acids are incorporated by either amber codon suppression or replacement of methionine during protein expression by the unnatural amino acid. Both of these methods result in lower levels of protein expression compared to wild-type expression, a potential issue with this methodology. Another potential issue for the incorporation of an unnatural amino acid through amber codon suppression is the fact that it requires the development of a tRNA synthetase that is capable of forming the tRNA with the desired amino acid attached for protein translation. Unnatural amino acids that have been successfully incorporated into proteins and used to install a glycan include *p*-iodophenylalanine^{62, 66} (via Suzuki-Miyaura coupling, Scheme 1.13), *p*-acetylphenylalanine⁶⁷ (via the formations of oxime linkage), azidohomoalanine⁶³ and

bicyclononyne⁶⁸ bearing amino acid (both via copper catalysed azide-alkyne cycloaddition).



Scheme 1.13: Additions of glycans through unnatural amino acids.^{10, 63, 68}

1.4.4 - Enzymatic Glycoprotein Synthesis

It has been described how nature synthesises glycoproteins, using glycosyltransferase enzymes to catalyse the addition of an activated glycan donor (normally a diphosphonucleotide activated glycan) to an acceptor, either a glycan acceptor or an amino acid side chain. It is possible to use some of these glycosyltransferases, with the related activated glycan donor, to form desired glycan structures on proteins *in vitro*. For this application, glycosyltransferases are most often used to further derivatize glycans and require the presence of a glycan acceptor already attached to the protein. The result of this is that glycosyltransferases are often used in combination with other techniques, such as using SPPS or post-translational modification, using these methods to form and/or install the initial glycan acceptor followed by the use glycosyltransferases to extend the glycan. A number of glycosyltransferases are commercially available, including galactosyltransferases, fucosyltransferases,

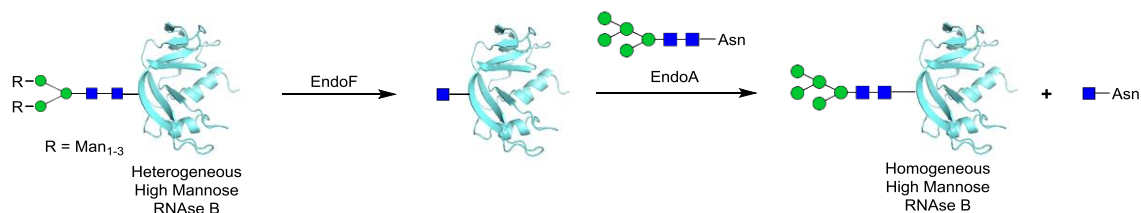
mannosyltransferases and sialyltransferases,⁹ each transferring the saccharide indicated in the name, with different versions of these transferases having different chemo- and regio-specificities. The activated donors are also commercially available. As well as the commercially available glycosyltransferases, a large number of others has been expressed and utilised, allowing a large variety of different glycan structures to be formed on proteins.⁹ The commercial availability, as well as the ease of expression of some of the glycosyltransferases has made this strategy a widely used technique.^{9, 10, 29,}

67, 69

An alternative method for the enzymatic extension of *N*-glycans on proteins involves the use of endo- β -*N*-acetylglucosaminidase (Endo) enzymes and glycan remodelling.^{44-46, 70, 71} This involves the expression of a glycoprotein which contains a heterogeneous *N*-glycan, cleavage of the glycan to leave only the core GlcNAc attached to the protein using an Endo enzyme, such as EndoH, EndoF₃, EndoA amongst others, followed by enzymatic extension to give the desired homogenous glycan.⁹ One approach for this enzymatic extension is the use of glycosyltransferases, as described above

A different strategy is to use an Endo enzyme to catalyse the glycosylation reaction and form the extended glycan, a process termed glycan remodelling. It was originally shown that Endo enzymes were capable of cleaving a glycan off an asparagine residue/peptide and then transfer the released glycan to an acceptor molecule, such as GlcNAc. This was shown to be possible with EndoA⁷² and EndoM.⁷³ Furthermore, this reaction was shown to be applicable to the synthesis of heterogeneous glycoproteins with EndoA

being used for the transfer a high mannose glycan from an asparagine residue to RNase B previously treated with EndoF (Scheme 1.14).⁷⁴

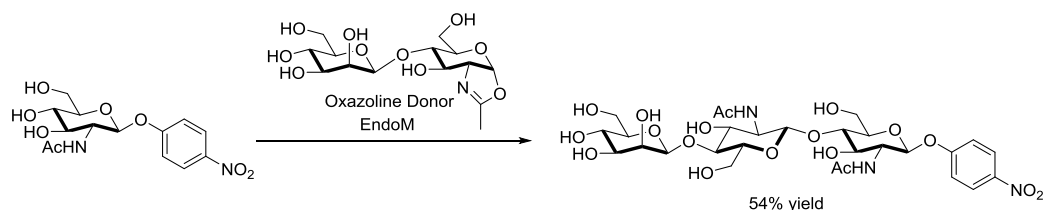


Scheme 1.14: Initial use of Endo enzymes for the formation of homogenous high mannose RNase

B.⁷⁴ Crystal structure of RNase B from *McPherson et al.*⁷⁵

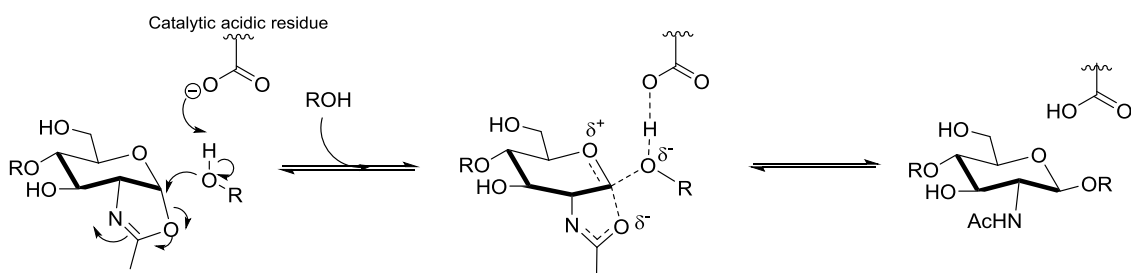
The major weakness of this method was the low yield due to the slow glycosylation reaction and predominant hydrolytic activity of the Endo enzyme, hydrolysing the product back to starting material. Improvement in the yield of the desired transglycosylation product was first achieved by using organic co-solvents in the reaction,⁷⁶ although this is not be compatible with protein chemistry. The breakthrough for the use Endo enzymes for transglycosylation reactions came from *Yamamoto et al.*⁷⁷ They predicted that an oxazoline glycosyl donor would act as a mimic of the oxazolinium-ion intermediate that forms during the Endo catalysed hydrolytic mechanism. Hence, they predicted that the oxazoline donor would have increased affinity in the active site of the Endo enzyme compared to a GlcNAc donor and therefore act as a superior donor substrate for glycosylation with an Endo enzyme in the presence of a GlcNAc acceptor. They formed a novel disaccharide oxazoline donor and showed that this did indeed function as a substrate for EndoA and EndoM catalysed glycosylation when using a GlcNAc-Asn or GlcNAc-pNP acceptor (Scheme 1.15). Moreover, *Yamamoto et al.* reported that no hydrolysis of the trisaccharide

glycosylation product was observed, suggesting that the glycosylation reaction proceeds significantly faster than hydrolysis, compared to when using a GlcNAc donor.



Scheme 1.15: First use of an oxazoline donor in an glycosylation reaction with EndoM.⁷⁷

They predicted that the mechanism of this glycosylation reaction is very similar to the mechanism for hydrolysis after formation of the oxazolinium-ion intermediate, except that the glycan acceptor opens the oxazoline instead of water (Scheme 1.16).⁷⁷ Further studies also suggested this mechanism described by *Yamamoto et al.*^{44,44,45}

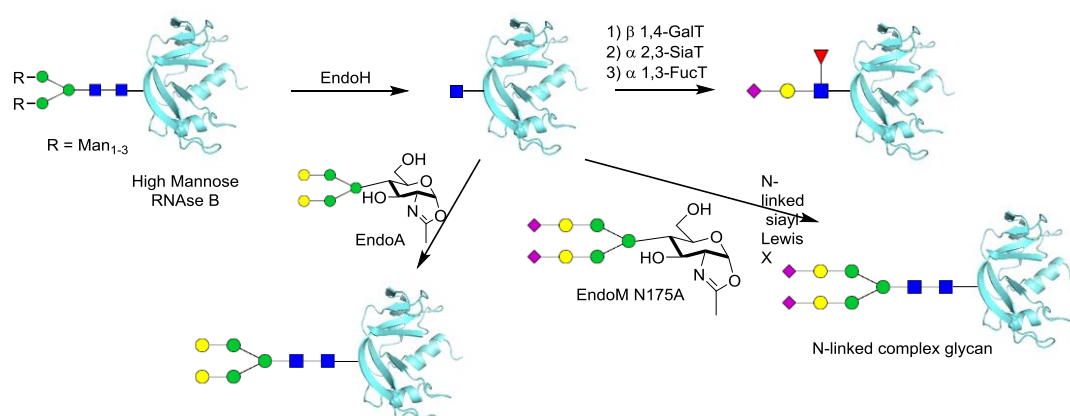


Scheme 1.16: Mechanism for glycosylation with Endo enzymes using an oxazoline donor.^{44, 45, 77}

The major advantage of Endo catalysed glycosylations is that they are carried out in a blockwise manner, where multiple sugar units can be added in one reaction. Disaccharides are the smallest unit added using this method,^{78,79} with other reports describing the addition of tetrasaccharides^{70,80} and oligosaccharides⁸¹

The obvious weakness of this method is that the glycan formed is the starting material for natural hydrolytic activity. As stated above for EndoA and EndoM for certain Endo

enzymes, hydrolysis of the product is slow and does not pose an issue, for other Endo enzymes the product is only fleetingly formed before being hydrolysed. One solution to this problem is the formation of mutants. When discussing the mechanism of glycan hydrolysis using Endo enzymes, it was said that a stabilising residue is required for the formation of the oxazolinium-ion intermediate and hence hydrolytic activity. Therefore, mutation of this stabilising residue results in the oxazolinium-ion intermediate being discouraged from being formed and reduced or abolished hydrolytic activity. However, mutation of the stabilising residue often also results in reduced glycosylation activity, perhaps due to decreased binding of the oxazoline donor. Due to this, a balance often has to be found, so that the mutant enzyme allows increased yields of the desired glycosylated product. Other mutations have been made on some Endo enzymes, aiming for increased yields of the desired glycosylation reaction, guided by the crystal structure of the enzyme.⁴⁵ There are lots of different Endo enzymes with reported glycosylation activity and mutants have been formed of some of them.^{44, 45, 81, 82} Remodelled glycoproteins using this method include RNase⁹ (Scheme 1.15), 34-mer representing a section of the HIV-1 envelope,⁷⁹ the Fc region of IgG⁸⁰ and IgG⁸¹ itself.



Scheme 1.15: Glycan remodelling on RNA using Endo enzymes. Crystal structure of RNase B from

McPherson et al.⁷⁵

1.4.5 - Engineering Glycosylation Pathways

The final method for the synthesis of homogenous glycoproteins to be described here is the engineering of glycosylation pathways into organisms, such as yeast and *E. coli*. There is large number of different methods to achieve this and only a few will be briefly mentioned here. One of the simplest methods to accomplish this is the use of small molecules to modulate the enzymes involved in the formation of the glycan.⁸³ This technique of using small molecule inhibitors was used to investigate the effect that different glycan structures have on the structure of IgG.⁸⁴

More complex techniques involve the re-engineering of the glycosylation pathway into organisms that normally have no ability to glycosylate proteins. This is no trivial task; however there are examples where this has been successful. For example, yeast has been engineered to form complex *N*-linked glycans attached to recombinant monoclonal antibodies,⁸⁵ and EPO, which showed activity similar to the naturally occurring protein.⁸⁶ In *E. coli*, it has been possible to express proteins containing an *N*-glycan in a homogenous manner by inserting the gene to express PglB, an oligosaccharyltransferase found in *C. jejuni* that attaches an oligosaccharide to an asparagine residue, as well as the enzymatic machinery to form the PglB substrate.⁸⁷ As PglB has a relaxed specificity, it is possible to attach a variety of different glycan structures. This methodology has also been combined with the glycan remodelling method described above *in vitro* for the overall attachment of Man₃GlcNAc₂ to the protein.⁸⁷

1.5 - Aims of this Thesis

1.5.1 - Introduction

The overall aim of this thesis was to further explore methods of forming homogenous glycoproteins using enzymatic methods. Broadly, this involved investigating the glycosylation reactivity of Endo enzymes, attempting to solve the critical weakness of using Endo enzymes for glycosylation; the hydrolysis of the formed product via the natural activity of the enzyme. The other protein extensively investigated here is PglB and its use *in vitro* as a glycosylation catalyst, as well as the scope for the addition of different saccharides using PglB and further glycosylation reactions using Endo enzymes and glycosyltransferases.

1.5.2 - Outline of Thesis

Chapter 2 – Investigating EndoS; An Antibody Specific Endo Enzyme

EndoS is an Endo enzyme that has specific activity for the hydrolysis of glycans on IgG. Here, investigations were made into mutating the enzyme to create a catalyst that results in greater yields of the glycosylation product. Also, studies were undertaken on the structure of EndoS, to inform the design of more efficient catalyst for the formation of the glycosylation product and to understand why EndoS is specific for IgG.

Chapter 3 – Investigations into the Activity of PglB *In Vitro* for the Synthesis of Glycopeptides

The use of PglB *in vitro* has recently been shown. Here, the effect of lipid chain length of the polyprenol pyrophosphate substrate on PglB activity is investigated. Expression of PglB and synthesis of these substrates is described. Finally, it is demonstrated that

further glycosylation reactions can be conducted after the PglB reaction using other enzymes.

Chapter 4 – Synthesis and Testing of PglB Substrates with Variation on the Sugar;

Investigating 4-Thiol GlcNAc

Described herein are attempts at forming a thioglycosidic bond on a peptide using an Endo enzyme, with thioglycosides being shown to be resistant to hydrolysis. This was envisaged to be possible by two enzymatic reactions; first using PglB to install a 4-thiol GlcNAc sugar onto a peptide and then the formation of the thioglycosidic bond using EndoA and an oxazoline donor. Investigations into both these reactions and synthesis of the substrates are described. Furthermore, work towards the synthesis of other modified glycans is described, aiming to form novel glycan donors to test with PglB.

Chapter 5 – Investigating the Use of a Glycal Substrate with Endo Enzymes

It was predicted that a 2-deoxy sugar lacking the acetlyamine at the 2-position would be resistant to Endo hydrolysis. In order to form this structure, EndoA glycosylation activity was investigated when using a glycal donor, with it shown by others that some glycosyl hydrolyases are able to accept glycals as substrates for glycosylation reactions. Synthesis of the glycal substrate and investigation of its use with Endo enzymes and mutants in glycosylation reactions is explained here.

1.6 - References

- 1 J. C. Venter, M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton *et al.*, *Science*, 2001, **291**, 1304-1351.
- 2 E. S. Lander, L. M. Linton, B. Birren, N. Chad, M. C. Zody, J. Baldwin *et al.*, *Nature*, 2001, **409**, 860-921.
- 3 K. Lindblad-Toh, C. M. Wade, T. S. Mikkelsen, E. K. Karlsson *et al.*, *Nature*, 2005, **438**, 803-819.
- 4 A. T. Chinwalla, L. L. Cook, K. D. Delehaunty, G. A. Fewell, L. A. Fulton *et al.*, *Nature*, 2002, **420**, 520-562.
- 5 M. D. Adams, S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne *et al.*, *Science*, 2000, **287**, 2185-2195.
- 6 J. Knight, *Nature*, 2002, **417**, 374-376.
- 7 M. Mann and O. N. Jensen, *Nat. Biotechnol.*, 2003, **21**, 255-261.
- 8 C. Unverzagt and Y. Kajihara, *Chem. Soc. Rev.*, 2013, **42**, 4408-4420.
- 9 R. M. Schmaltz, S. R. Hanson and C. Wong, *Chem. Rev.*, 2011, **111**, 4259-4307.
- 10 D. P. Gamblin, E. N. Scanlan and B. G. Davis, *Chem. Rev.*, 2009, **109**, 131-163.
- 11 R. A. Dwek, *Chem. Rev.*, 1996, **96**, 683-720.
- 12 E. Selvin, S. Marinopoulos, G. Berkenblit, T. Rami, F. L. Brancatu, N. R. Powe and S. H. Golden, *Ann. Intern. Med.*, 2004, **141**, 421-431.
- 13 A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart and M. E. Etzler, *Essentials of Glycobiology*, 2nd edition, Cold Spring Harbor, New York, 2009.
- 14 M. Allhorn, J. G. Briceno, L. Baudino, C. Lood, M. L. Olsson, S. Izui and M. Collin, *Blood*, 2010, **115**, 5080-5088.

- 15 J. Wang, I. Grundke-Iqbal and K. Iqbal, *Nat. Med.*, 1996, **2**, 871-875.
- 16 S. Hakomori, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 10231-10233.
- 17 J. W. Dennis, M. Granovsky and C. E. Warren, *Biochimica et Biophysica Acta*, 1999, **1473**, 21-34.
- 18 B. G. Davis and L. Wang, *Chem. Sci.*, 2013, **4**, 3381-3394.
- 19 F. Schwarz and M. Aebi, *Curr. Opin. Struct. Biol.*, 2011, **21**, 576-582.
- 20 Y. Zhao, M. Takahashi, J. Gu, E. Miyoshi, A. Matsumoto, S. Kitazume and N. Taniguchi, *Cancer Sci.*, 2008, **99**, 1304-1310.
- 21 K. Fiedler and K. Simons, *Cell*, 1995, **81**, 309-312.
- 22 J. A. Voynow and B. K. Rubin, *Chest*, 2009, **135**, 505-512.
- 23 K. Marino, J. Bones, J. J. Kattla and P. M. Rudd, *Nat. Chem. Bio.*, 2010, **6**, 713-723.
- 24 T. A. Fritz, J. H. Hurley, L. Trinh, J. Shiloach and L. A. Tabak, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 15307-15312.
- 25 J. D. Esko, *Curr. Opin. Cell Biol.*, 1991, **3**, 805-816.
- 26 K. Prydz and K. T. Dalen, *J. Cell Sci.*, 2000, **113**, 193-205.
- 27 N. Perrimon and M. Bernfield, *Semin. Cell Dev. Biol.*, 2001, **12**, 65-67.
- 28 M. Kowarik, N. M. Young, S. Numao, B. L. Schulz, I. Hug, N. Callewaert, D. C. Mills, D. C. Watson, M. Hernandez, J. F. Kelly, M. Wacker and M. Aebi, *EMBO J.*, 2006, **25**, 1957-1966.
- 29 F. Liu, B. Vijayakrishnan, A. Faridmoayer, T. A. Taylor, T. B. Parsons, G. J. L. Bernardes, M. Kowarik and B. G. Davis, *J. Am. Chem. Soc.*, 2014, **136**, 566-569.
- 30 L. L. Lairson, B. Herissat, G. J. Davies and S. G. Withers, *Annu. Rev. Biochem.*, 2008, **77**, 521-555.

- 31 M. Jinek, Y. W. Chen, H. Clausen, S. M. Cohen and E. Conti, *Nat. Struct. Mol. Biol.*, 2006, **13**, 945-946.
- 32 C. Garinot-Schneider, A. C. Lellouch and R. A. Geremia, *J. Biol. Chem.*, 2000, **275**, 1407-1413.
- 33 D. E. Koshland, *Biol. REv. Camb. Phil. Soc.*, 1953, **28**, 416-436.
- 34 Y. N. Zhang, G. J. Swaminathan, A. Deshpande, E. Boix, R. Natesh, Z. Xie, K. R. Acharya and K. Brew, *Biochemistry*, 2003, **42**, 13512-13521.
- 35 K. Persson, H. D. Ly, M. Dieckelmann, W. W. Wakarchuk, S. G. Withers and N. C. J. Strynadka, *Nat. Struct. Bio.*, 2001, **8**, 167-176.
- 36 G. K. Wagner, T. Pesnot, M. M. Palcic and R. Jogensen, *J. Biol. Chem.*, 2015, **290**, 31162-31172.
- 37 D. C. Phillips, *Proc. Natl. Acad. Sci. U. S. A.*, 1967, **57**, 484-495.
- 38 D. C. Phillips, *Sci. Am.*, 1966, **215**, 78-90.
- 39 S. S. Lee, S. S. Hong, J. C. Errey, A. Izumi, G. J. Davies and B. G. Davis, *Nat. Chem. Bio.*, 2011, **7**, 631-638.
- 40 J. D. McCarter and S. G. Withers, *Curr. Opin. Sturct. Biol.*, 1994, **4**, 885-892.
- 41 C. S. Rye and S. G. Withers, *Curr. Opin. Chem. Biol.*, 2000, **4**, 573-580.
- 42 G. Davies and B. Henrissat, *Stucture*, 1995, **3**, 853-859.
- 43 D. J. Vocadlo, G. J. Davies, R. Laine and S. G. Withers, *Nature*, 2001, **412**, 835-838.
- 44 M. Umekawa, C. Li, T. Higashiyama, W. Huang, H. Ashida, K. Yamamoto and L. Wang, *J. Biol. Chem.*, 2010, **285**, 511-521.
- 45 J. Yin, L. Li, N. Shaw, Y. Li, J. K. Song, W. Zhang, C. Xia, R. Zhang, A. Joachimiak, H. Zhang, L. Wang, Z. Liu and P. Wang, *Plos One*, 2009, **4**, e4658.
- 46 L. Wang and J. V. Lomino, *ACS Chem. Biol.*, 2012, **7**, 110-122.

- 47 N. R. Taylor and M. von Itzstein, *J. Med. Chem.*, 1994, **37**, 616-624.
- 48 G. Walsh and R. Jefferis, *Nat. Biotechnol.*, 2006, **24**, 1241-1252.
- 49 M. Satoh, S. Iida and K. Shitara, *Expert Opin. Biol. Ther.*, 2006, **6**, 1161-1173.
- 50 P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. Kent, *Science*, 1994, **266**, 776-779.
- 51 R. J. Payne and C. Wong, *Chem. Commun.*, 2010, **46**, 21-43.
- 52 P. Siman and A. Brik, *Org. Biomol. Chem.*, 2012, **10**, 5684-5697.
- 53 B. Wu, J. Chen, D. J. Warren, G. Chen, Z. Hua and S. J. Danishefsky, *Angew. Chem. Int. Ed.*, 2006, **45**, 4116-4125.
- 54 Y. Zhang, C. Xu, H. Y. Lam, C. L. Lee and X. Li, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 6657-6662.
- 55 C. Xu, H. Y. Lam, Y. Zhang and X. Li, *Chem. Commun.*, 2013, **49**, 6200-6202.
- 56 I. Pusterla and J. W. Bode, *Nat. Chem.*, 2015, **7**, 668-672.
- 57 B. Aussedat, B. Fasching, E. Johnston, N. Sane, P. Nagorny and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2012, **134**, 3532-3541.
- 58 M. A. Brimble, P. J. Edwards, P. W. R. Harris, G. E. Norris, M. L. Patchett, T. H. Wright, S. Yng and S. E. Carley, *Chem. Eur. J.*, 2015, **21**, 3556-3561.
- 59 R. J. Payne, *Angew. Chem. Int. Ed.*, 2013, **52**, 505-507.
- 60 J. M. Chalker and B. G. Davis, *Curr. Opin. Chem. Biol.*, 2010, **14**, 781-789.
- 61 J. M. Chalker, G. J. L. Bernardes and B. G. Davis, *Accounts Chem. Res.*, 2011, **44**, 730-741.
- 62 J. M. Chalker, S. B. Gunnoo, O. Boutureira, S. C. Gerstberger, Ferandez-Gonzalez, G. J. Bernardes, L. Griffin, H. Hailu, C. J. Schofield and B. G. Davis, *Chem. Sci.*, 2011, **2**, 1666-1676.

- 63 M. Fernandez-Gonzalez, O. Boutureira, G. J. L. Bernardes, J. M. Chalker, M. A. Young, J. C. Errey and B. G. Davis, *Chem. Sci.*, 2010, **1**, 709-715.
- 64 G. J. L. Bernardes, J. M. Chalker, J. C. Errey and B. G. Davis, *J. Am. Chem. Soc.*, 2008, **130**, 5052-5053.
- 65 H. C. Kolb, M. G. Finn and K. B. Sharpless, *Angew. Chem. Int. Ed.*, 2001, **40**, 2004-2021.
- 66 C. D. Spicer and B. G. Davis, *Chem. Commun*, 2013, **49**, 2747-2749.
- 67 H. Liu, L. Wang, A. Brock, C. Wong and P. G. Schultz, *J. Am. Chem. Soc.*, 2003, **125**, 1702-1703.
- 68 T. Machida, K. Lang, L. Xue, J. W. Chin and N. Winssinger, *Bioconjugate Chem.*, 2015, **26**, 802-806.
- 69 D. P. Gamblin, P. Garnier, S. van Kasteren, N. J. Oldham, A. J. Fairbanks and B. G. Davis, *Angew. Chem. Int. Ed.*, 2004, **43**, 828-833.
- 70 J. J. Goodfellow, K. Baruah, K. Yamamoto, C. Bonomelli, B. Krishna, D. J. Harvey, M. Crispin, C. N. Scanlan and B. G. Davis, *J. Am. Chem. Soc.*, 2012, **134**, 8030-8033.
- 71 T. B. Parsons, M. K. Patel, A. B. Boraston, D. J. Vocadlo and A. J. Fairbanks, *Org. Biomol. Chem.*, 2010, **8**, 1861-1869.
- 72 K. Takegawa, S. Yamaguchi, A. Kondo, H. Iwamoto, I. Kato and S. Iwahara, *Biochem. Int.*, 1991, **24**, 849-855.
- 73 K. Yamamoto, S. Kadowaki, J. Watanabe and H. Kumagai, *Biochem. Bioph. Res. Co.*, 1994, **203**, 244-252.
- 74 K. Takegawa, M. Tabuchi, S. Yamaguchi, A. Kondo, I. Kato and S. Iwahara, *J. Biol. Chem.*, 1995, **270**, 3094-3099.
- 75 T. Ko, R. Williams and A. McPherson, *Acta Crystallogr D*, 1996, **52**, 160-164.

- 76 J. Fan, K. Takegawa, S. Iwahara, A. Kondo, I. Kato, C. Abeygunawardana and Y. C. Lee, *J. Biol. Chem.*, 1995, **270**, 17723-17729.
- 77 M. Fujita, S. Shoda, K. Haneda, T. Inazu, K. Takegawa and K. Yamamoto, *Biochimica et Biophysica Acta*, 2001, **1528**, 9-14.
- 78 M. Fernandez-Gonzalez, O. Boutureira, G. J. L. Bernardes, J. M. Chalker, M. A. Young, J. C. Errey and B. G. Davis, *Chem. Sci.*, 2010, **1**, 709-715.
- 79 B. Li, Y. Zeng, S. Hauser, H. Song and L. Wang, *J. Am. Chem. Soc.*, 2005, **127**, 9692-9693.
- 80 G. Zou, H. Ochiai, W. Huang, Q. Yang, C. Li and L. Wang, *J. Am. Chem. Soc.*, 2011, **133**, 18975-18991.
- 81 W. Huang, J. Giddens, S. Fan, C. Toonstra and L. Wang, *J. Am. Chem. Soc.*, 2012, **134**, 12308-12318.
- 82 C. D. Heidecke, Z. Ling, N. C. Bruce, J. W. B. Moir, T. B. Parsons and A. J. Fairbanks, *Chem. Bio. Chem.*, 2008, **9**, 2045-2051.
- 83 W. P. Almeida and F. Coelho, *Tet. Lett.*, 2003, **44**, 937-940.
- 84 T. A. Bowden, K. Baruah, C. H. Coles, D. J. Harvey, X. Yu, B. Song, D. I. Stuart, A. R. Aricescu, C. N. Scanlan, Y. E. Jones and M. Crispin, *J. Am. Chem. Soc.*, 2012, **134**, 17554-17563.
- 85 H. Li, N. Sethuraman, T. A. Stadheim, D. Zha, B. Pronz, N. Ballew, P. Bobrowicz, B. Choi, W. J. Cook, M. Cukan, N. R. Houston-Cummings, R. Davidson, B. Gong, S. R. Hamilton, J. P. Hoopes, Y. Jiang, N. Kim, R. Mansfield, J. H. Nett, S. Rios, R. Strawbridge, S. Wildt and T. U. Gerngross, *Nat. Biotechnol.*, 2006, **24**, 210-215.
- 86 S. R. Hamilton, R. C. Davidson, N. Sethuraman, J. H. Nett, Y. Jiang, S. Rios, P. Bobrowicz, T. A. Stadheim, H. Li, B. Choi, D. Hopkins, H. Wischniewski, J. Roser, T.

Mitchell, R. R. Strawbridge, J. Hoopes, S. Wildt and T. U. Gerngross, *Science*, 2006, **313**, 1441-1443.

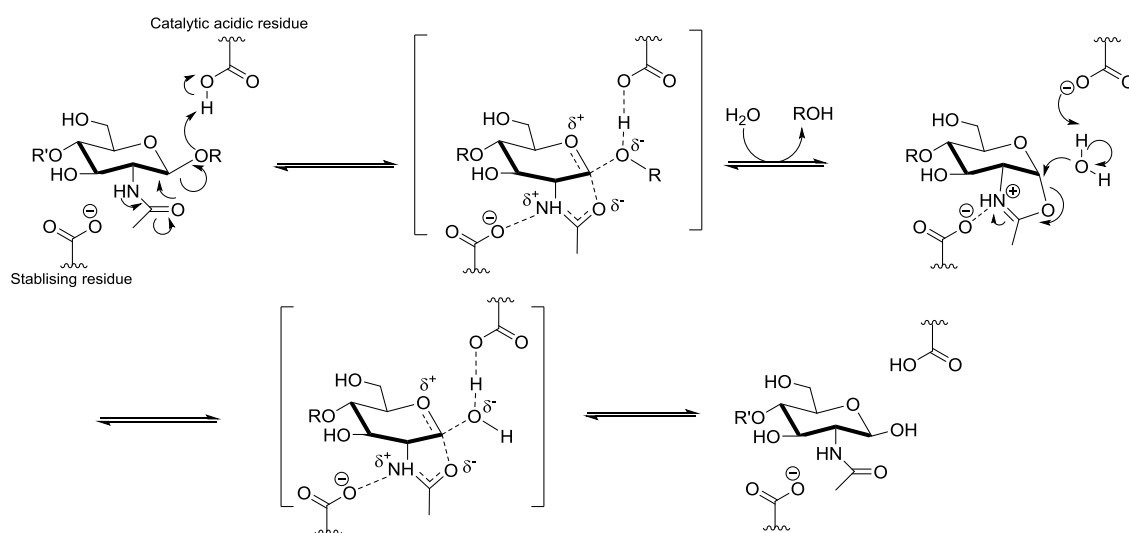
87 F. Schwarz, W. Huang, C. Li, B. L. Schulz, C. Lizak, A. Palumbo, S. Numao, D. Neri, M. Aebi and L. Wang, *Nat. Chem. Bio.*, 2010, **6**, 264-266.

Chapter 2 – Investigating EndoS; An Antibody Specific Endo Enzyme

Chapter 2 – Investigating EndoS; An Antibody Specific Endo Enzyme

2.1 – Introduction

As described in Chapter 1, Endo enzymes are glycoside hydrolyases, where their natural activity is to catalyse the hydrolysis of the $\beta 1 \rightarrow 4$ linkage between two *N*-acetylglucosamines (GlcNAc) units, resulting in only the reducing end GlcNAc unit remaining of the *N*-glycan on the glycoprotein. The mechanism for this reaction goes via an oxazolinium ion intermediate, which is stabilised in the active site of the enzyme (Scheme 2.1).¹



Scheme 2.1: Endo hydrolysis mechanism, going via an oxazoline intermediate.¹

The Endo enzymes are split into two glycoside hydrolase families; GH18 and GH85.² The GH18 family contains bacterial Endo enzymes, such as EndoH,³ EndoF₁, EndoF₂ and EndoF₃.²⁴ The GH85 family contains Endo enzymes from both prokaryotes and eukaryote, examples including EndoA,¹ EndoD⁵ and EndoM.² One of the key differences between these two families is the stabilising residue, where the stabilising residue is an Asn residue for GH85 enzymes and an Asp residue for GH18 enzymes.

Although Endo enzymes belong to two separate glycoside hydrolyase families, the structures of many of them are very similar, either seen by crystallography or predicted by sequence alignment, and contain a catalytic $(\beta/\alpha)_8$ TIM barrel fold (Figure 2.1).

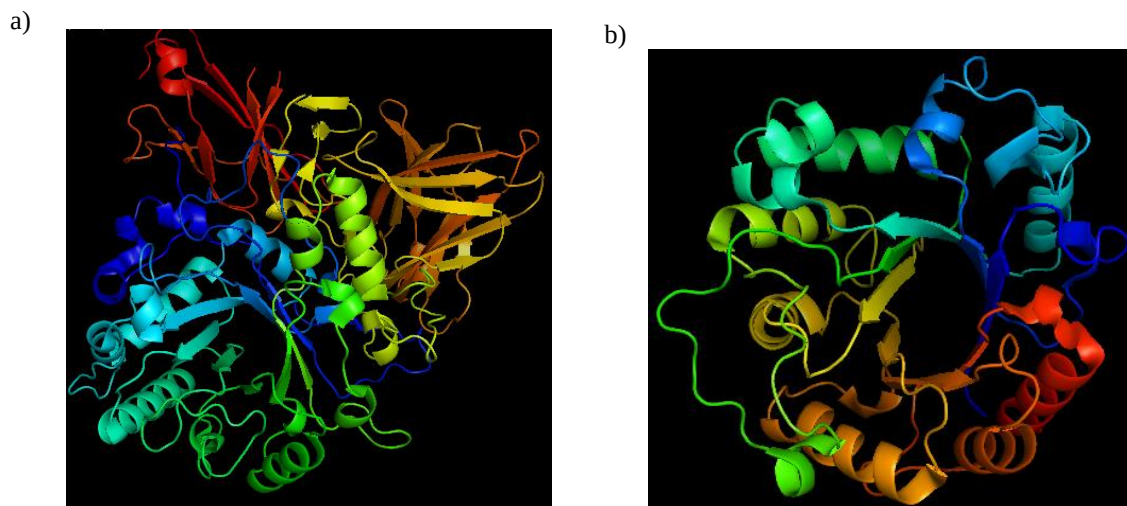
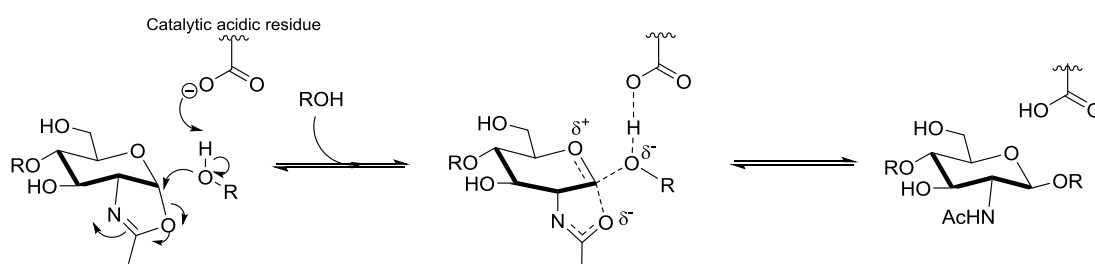


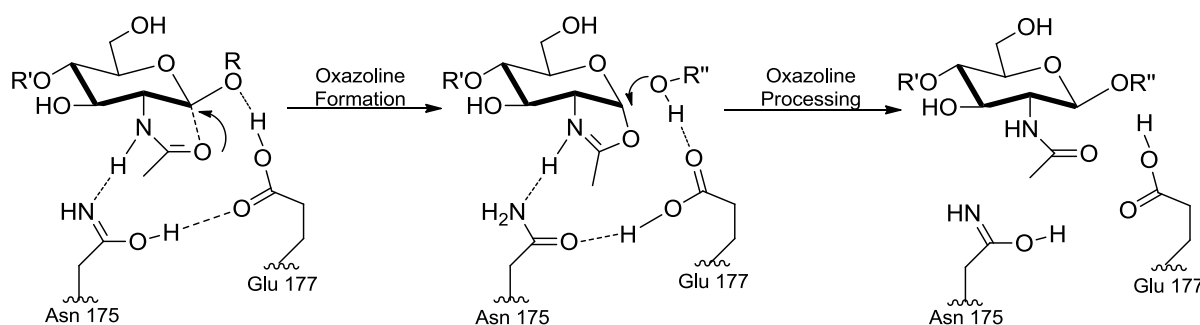
Figure 2.1: Crystal structure of a) EndoA⁶ and b) EndoF_{3.4}.⁴ Both show a catalytic $(\beta/\alpha)_8$ TIM barrel fold.

As stated in Chapter 1, there is a lot of interest in the use of Endo enzymes to catalyse glycosylation reactions and doing the reverse reaction to its natural activity. This is possible through the use of sugar oxazoline donor substrates, which mimic the oxazolinium ion intermediate in the hydrolysis mechanism (Scheme 2.2).¹ This glycosylation activity has been reported for EndoA,^{1, 7, 8} EndoM,^{2, 8, 9} EndoFV,¹⁰ EndoF₃,¹¹ EndoD^{9, 12} and EndoOM¹³ and has been used to form a variety of different glycopeptides and glycoproteins.



Scheme 2.2: Mechanism of Endo catalysed glycosylation using an oxazoline donor.¹

The obvious weakness of this strategy for glycosylation reactions is that the desired glycosylated product is the substrate for the natural hydrolysis activity of the enzyme. To tackle this issue, Endo mutants have been investigated, aiming to form an ENGase; an Endo glycosyltransferase that shows good glycosylation reactivity and little or no hydrolysis activity. These have been formed with varying degrees of success and have been studied with EndoA and EndoM by Wang *et al.*^{1, 2, 14} and Fairbanks *et al.*¹⁵. The main approach for the formation of Endo mutants that could be considered an ENGase is to mutate the stabilising residue (Asn in the case of EndoA and EndoM) so that the oxazolinium ion intermediate cannot form and hence hydrolysis is not possible. EndoA N171A (where N171 is the stabilising residue) has been formed and indeed shows no hydrolysis, but glycosylation reactivity is also greatly diminished.¹ The same is true for EndoM N175A (where N175 is the stabilising residue). However, saturated mutagenesis of N175 on EndoM found that EndoM N175Q only has 1% of the hydrolytic activity of wild type EndoM but 150% of its' glycosylation activity.² It is thought the reason for this activity is due to N175Q not being able to stabilise the oxazolinium ion intermediate (hence low hydrolytic activity) but can still form a hydrogen bond to the catalytic residue and interact with it and so retains glycosylation activity (Scheme 2.3).



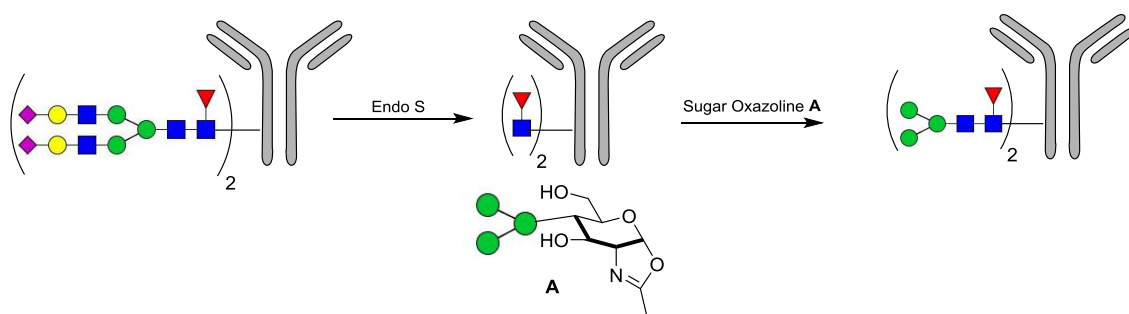
Scheme 2.3: Proposed mechanism of hydrolytic/glycosylation activity of EndoM (GH85 enzyme) by Wang *et al.*, showing the proposed critical activity between N175 and E177.²

In order to inform the mutation studies of Endo enzymes, aiming for the formation of an ENGase, the crystal structure of a number of Endo enzymes have been obtained, including EndoA,¹ EndoF₃⁴ and EndoT¹⁶ amongst others.^{3, 5} From these crystal structures, a number of important amino acids have been found in the active site, aside from the catalytic residue and the stabilising residue. For example, the crystal structure of EndoA showed that Y299 interacts with the substrate in the active site and it was found that EndoA Y299F has increased glycosylation activity.¹ A number of other important amino acids were found in EndoA's active site and mutation of some of these has resulted in different hydrolysis/glycosylation activities.

Endo- β -N-acetylglucosaminidase S (EndoS) was first discovered in *Streptococcus pyogenes*.¹⁷ *Streptococcus pyogenes* is a human pathogen that causes serious infections such as pharyngitis and scarlatina and invasive diseases such as sepsis.^{17, 18} What is unique about EndoS compared to other Endo enzymes is that EndoS has specific activity on Immunoglobulin G (IgG) type antibodies and experimental data has shown EndoS is specific for the 3D structure of IgG.^{17, 18} EndoS belongs to the GH18 family¹⁷, with the catalytic domain of EndoS being most similar to that of EndoF₂ by sequence alignment.^{17, 19} The unique specificity of EndoS has led to investigations of its use to treat autoimmunity diseases.^{18, 20, 21}

It had recently been shown in the group that, like other Endo enzymes, EndoS shows glycosylation activity.²² This activity is seen on IgG, with EndoS being specific to this substrate. IgG, with heterogeneous glycans, was treated with EndoS resulting in the hydrolysis of the glycans, and only the core GlcNAc and fucose unit remained. With

addition of a tetrasaccharide oxazoline donor (**A**), EndoS catalysed glycosylation occurred, resulting in essentially a single glycoform on the IgG (Scheme 2.4).²² However, this glycosylated product was only fleetingly seen and could only be formed under a strict set of conditions; product hydrolysis by the natural activity of EndoS was a major issue.¹⁴



Scheme 2.4: Hydrolysis and glycosylation activity of EndoS on an IgG antibody.²²

Here, it was desired to investigate mutating EndoS, with the aim of forming an EndoS mutant that could be considered an EndoS ENGase and thus allow the formation of homogenous glycans on IgG. The specificity of EndoS to IgG would allow, along with the use of other glycoside hydrolysis and glycosyltransferase enzymes, the formation of glycan structures on IgG with tuned *in vivo* characteristics, ideal for therapeutic IgGs. In order to assist the design of EndoS mutants with ENGase properties, structural investigations were also carried out, aiming to obtain the crystal structure of EndoS, ideally in complex with a substrate.

2.2 – Formation of an Optimised EndoS Expression and Purification System

As this study would require a large amount of protein, studies were undertaken into how the most protein could be obtained in the least amount of time. The EndoS gene had been previously inserted into a pGEX-4T-1 plasmid, resulting in the expression of EndoS with a GST tag for affinity purification, with a thrombin cleavage site between the tag and enzyme. GST-EndoS had previously been expressed using an auto-induction system, where the lactose produced by *E. coli* causes induction; although this resulted in the formation of EndoS, this was an overnight process. It was desired to see if isopropyl- β -D-thiogalactopyranoside (IPTG) also resulted in EndoS expression, as the plasmid should allow for this. Expression trials were undertaken, with the EndoS plasmid transformed into competent *E. coli* cells (BL21 (DE3)) and expressed either in lysogeny broth (LB) media and induced with IPTG or expressed in auto-induction media. Samples were taken across a time course and analysed by SDS-PAGE (Figure 2.2). It was clearly seen that GST-EndoS is expressed with IPTG induction and in similar amounts to when expressed in auto-induction media.

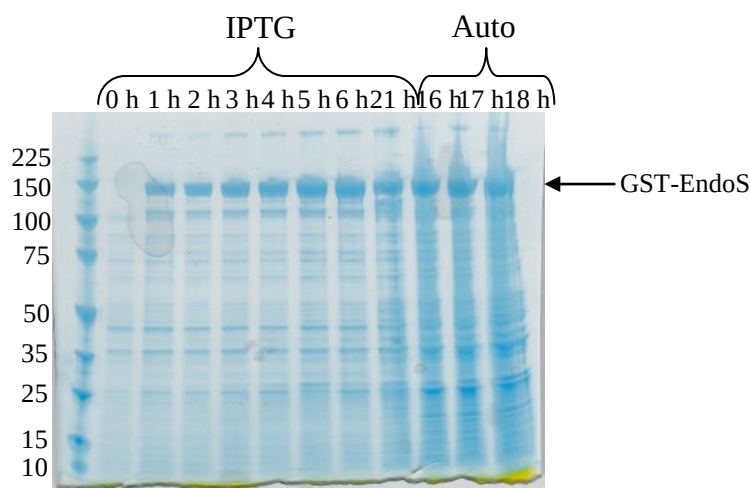


Figure 2.2: Expression trials of EndoS.

The expressed GST-EndoS has a thrombin cleavage site between the GST tag and EndoS. As it was desired to form large amounts of protein, this would mean the use of large amounts of thrombin. To avoid this, a TEV protease cleavage site was inserted between the GST tag and EndoS. TEV cleaves at the recognition site ENLYFQ↓G (P1' can vary).²³ The advantage of the use of TEV compared to thrombin is that TEV had previously been expressed in *E. coli* and large stocks were available, whereas thrombin cannot be expressed in *E. coli* and is expensive to purchase. Furthermore, the smaller size of TEV (~27 kDa) compared to thrombin (~37 kDa) should allow purification of EndoS through just size-exclusion chromatography and without further purification steps (such as ion-exchange). The insertion of the TEV cleave site was done successfully via two rounds of site directed mutagenesis. To ensure this system would work for EndoS purification and that the enzymatic activity would not be affected, a test expression of EndoS was conducted. In short, EndoS was expressed in LB media and protein expression induced by IPTG. The cell pellet was then harvested, the cells lysed by sonication and the soluble fraction collected by centrifugation. This was then purified by GSTrap columns, affording GST-EndoS (Figure 2.3a). Investigating the cleavage of GST tag from EndoS, it was found that incubating overnight at room temperature in a Tris buffer using a ratio of 1 mg TEV to 25 mg of GST-EndoS gave significant cleavage (Figure 2.3b). Some GST-EndoS was still seen under all ratios of TEV to GST-EndoS; this is thought to be due to disulfide bonding between the GST tag and EndoS. Purification by size exclusion chromatograph gave pure EndoS (Figure 2.3c).

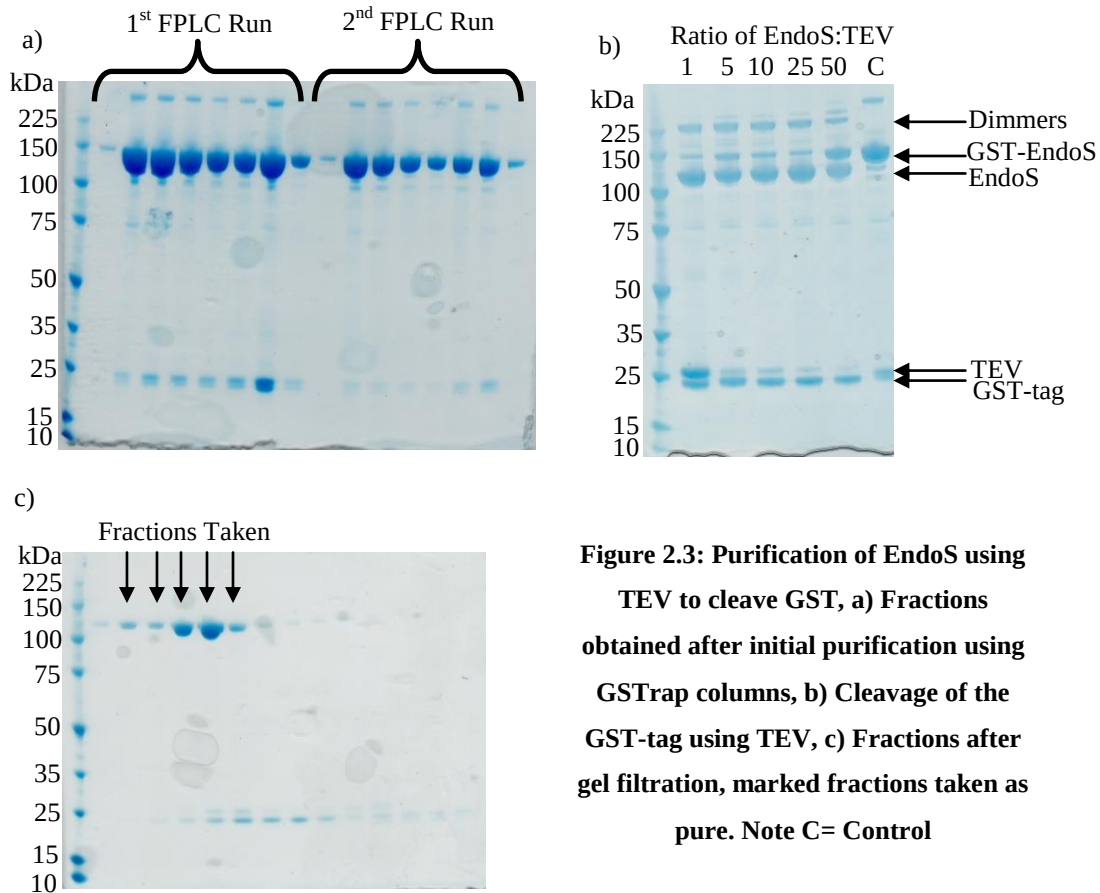


Figure 2.3: Purification of EndoS using TEV to cleave GST, a) Fractions obtained after initial purification using GSTrap columns, b) Cleavage of the GST-tag using TEV, c) Fractions after gel filtration, marked fractions taken as pure. Note C= Control

With pure protein in hand, it was desired to confirm it is active. This was done by treating IgG with the expressed EndoS and seeing a mass change by SDS-PAGE, corresponding to the hydrolysis of the glycans (Figure 2.4). This was indeed observed, confirming the expressed EndoS is active and the modifications to the expression and purification protocols were successful.

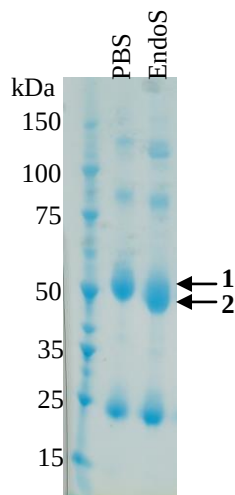


Figure 2.4 – Test EndoS hydrolytic activity. Commercially available IgG was treated with either PBS or EndoS for 1 h, after which the cleavage of the glycans was observed via the mass shift on the SDS-PAGE. 1 = Glycosylated IgG heavy chain, 2 = GlcNAc IgG heavy chain

2.3 – Initial Formation and Testing of EndoS Mutants

From the previous research into the formation of EndoA and EndoM mutants, the obvious starting point to investigate EndoS mutants for increased glycosylation yields was the stabilising residue. It was expected that mutation of this residue would result in significantly reduced hydrolysis activity, but hopefully glycosylation activity would be retained. For EndoS this residue is D233, whereas the catalytic residue is E235. It was decided to investigate EndoS D233A, EndoS D233E and EndoS D233Q. EndoS D233A was chosen as this mutation to alanine would mean that no stabilisation of the critical oxazolinium ion intermediate for hydrolysis would be possible, although this mutation may mean that the glycosylation activity would also be significantly reduced, as seen for EndoA N171A. For EndoS D233E, it was aimed to emulate EndoM N175Q, with little hydrolysis (due to lack of stabilisation of the oxazolinium ion) but retain glycosylation activity due to the hydrogen bond to the catalytic residue being retained. As the difference between Asn and Gln is a methylene unit, as seen for EndoM N175Q, the addition of a methylene unit on the stabilisation residue on EndoS results in EndoS D233E. Shortly after EndoS D233A and EndoS D233E had been expressed, a study was published by Wang *et al.*¹⁴ where EndoS D233Q was used in glycosylation reactions and so this was formed as a reference to compare to the other mutants.

All these mutants were again formed through site-directed mutagenesis, using the pGEX-4T-1 EndoS TEV plasmid formed above, and gene sequencing confirmed the desired mutation had been inserted. These were all expressed as before, using the optimised conditions and purified with a GSTrap column (Figure 2.5). The fractions seen to contain the desired protein by SDS-PAGE were then treated with TEV to cleave

the GST tag and purification by size exclusion chromatography resulted in the desired EndoS mutants (Figure 2.6).

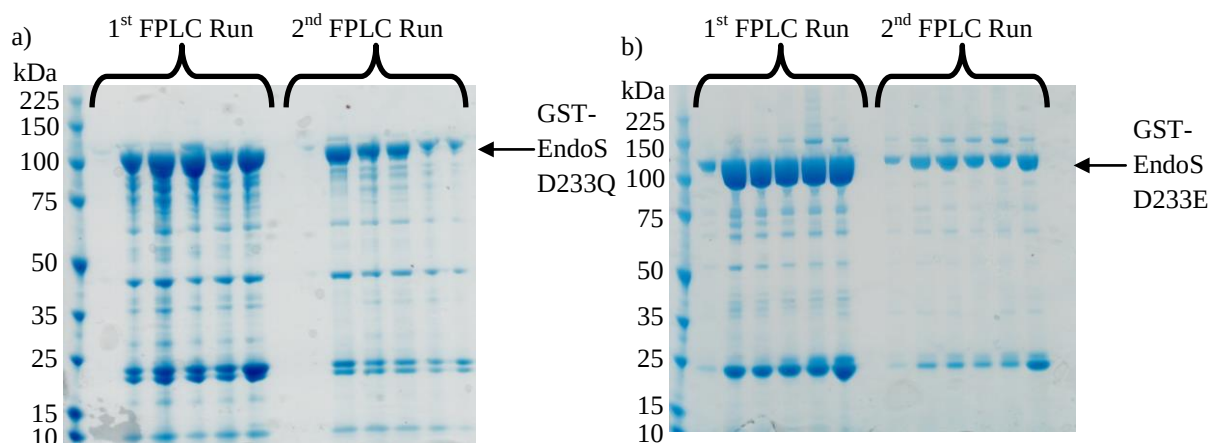


Figure 2.5: SDS-PAGE of the GSTrap purification of EndoS D233Q and EndoS D233E, a) GST-EndoS D233Q and b) GST-EndoS D233E. Note EndoS D233A was expressed and purified by another member of the group.

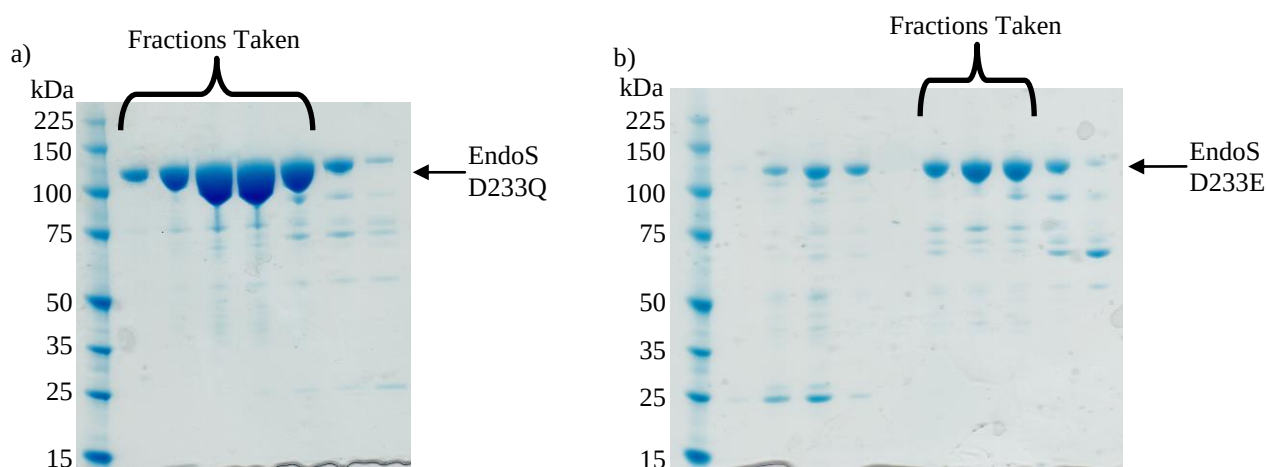


Figure 2.6: SDS-PAGE of the size exclusion purification of EndoS D233Q and EndoS D233E, a) EndoS D233Q and b) EndoS D233E. Note EndoS D233A was expressed and purified by another member of the group.

With the EndoS mutants prepared, their hydrolytic and glycosylation activities could be tested. For hydrolytic activity, Herceptin, a commercially available monoclonal

antibody, was treated with the EndoS mutants. The reactions were incubated at room temperature and samples were taken, denatured, to separate the heavy and light chains, and analysed by SDS-PAGE. A mass shift is observed in the molecular weight of the heavy chain, owing to the cleavage of the glycan off the Fc region. From the SDS-PAGE, densitometry analysis was carried out to obtain a semi-qualitative yield of hydrolysis (Figure 2.7 and Graph 2.1).

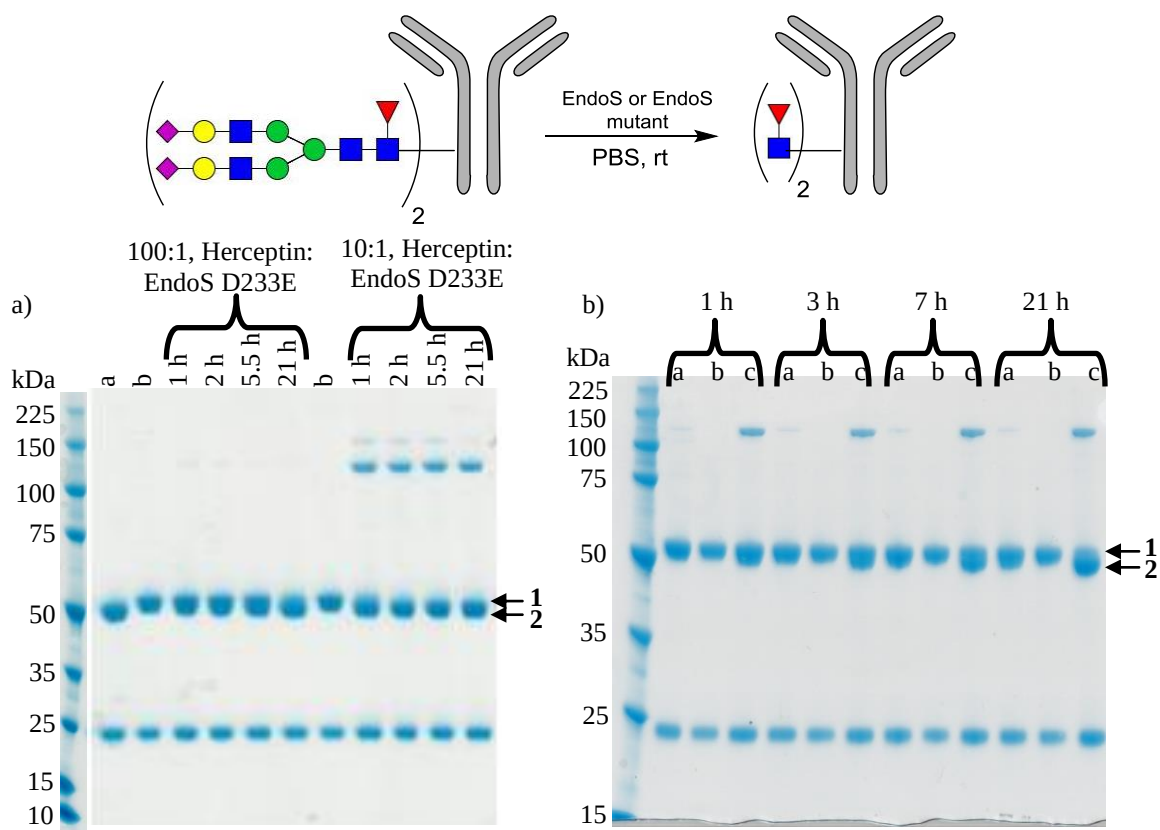
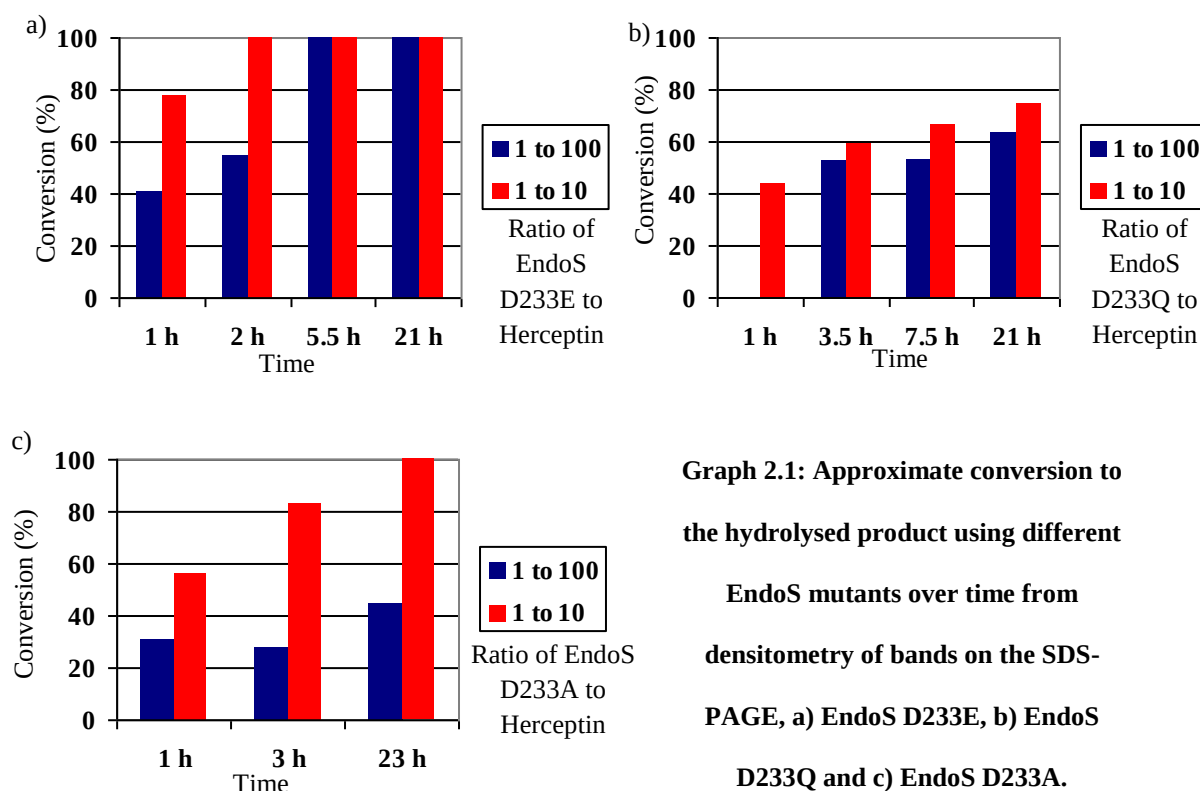


Figure 2.7: SDS-PAGE of testing EndoS mutants for hydrolytic activity, a) EndoS D233E, where a = GlcNAc-Herceptin and b = glycosylated-Herceptin, b) EndoS D233Q, where a = 100:1 Herceptin:EndoS D233Q, b = glycosylated-Herceptin and c = 100:1 Herceptin:EndoS D233Q. For both, 1 = glycosylated-heavy chain and 2 = GlcNAc-heavy chain. Note, the SDS-PAGE for EndoS D233A is shown in Figure 2.12



Although the densitometry analysis only gives semi-quantitative data, a number of important observations can be made. First, all the EndoS mutants still had a degree of hydrolytic activity and none of the mutations made completely abolished this. It can also be seen that EndoS D233E is the most hydrolytically active of these EndoS mutants, with EndoS D233A and EndoS D233Q having similar activity (Graph 2.1).

Next, it was desired to test the glycosylation activities of the EndoS mutants. This was done by treating trimmed Herceptin with a deca-saccharide oxazoline that had been semi-synthesised, via the extraction of the sugar from egg yolks,^{24, 25} and with the EndoS mutants. Again, these reactions were analysed by SDS-PAGE, where this time an increase in the mass of the heavy chain was expected, due to the addition of the

decasaccharide (Figure 2.8). As with the hydrolytic test, densitometry was done on the bands to get a semi-qualitative yield of the glycosylation (Graph 2.2).

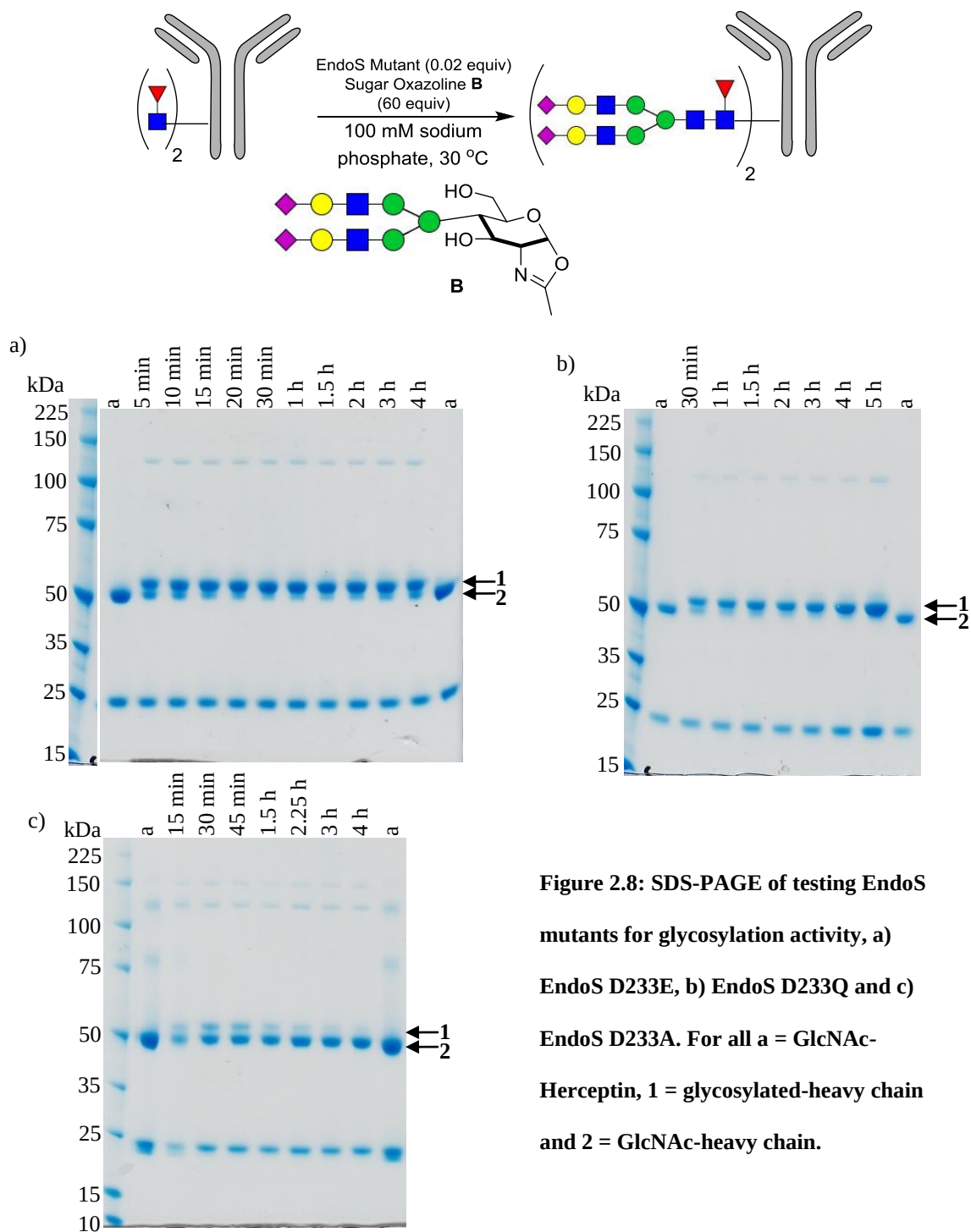
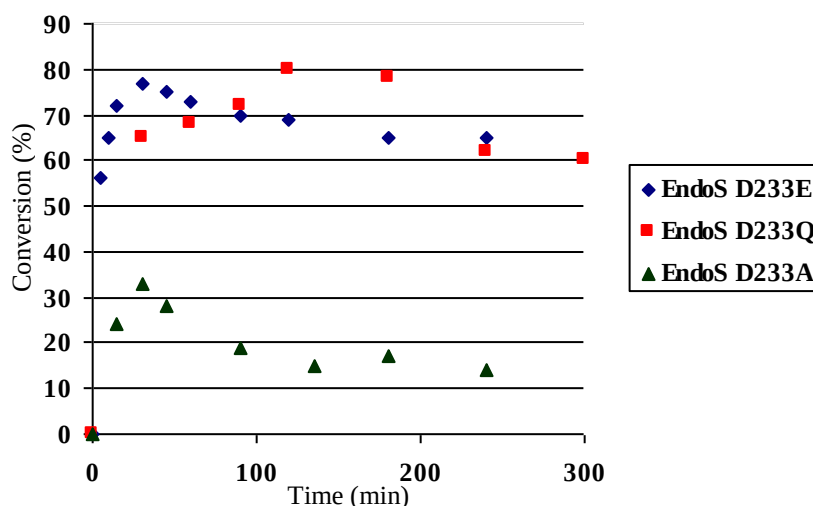


Figure 2.8: SDS-PAGE of testing EndoS mutants for glycosylation activity, a) EndoS D233E, b) EndoS D233Q and c) EndoS D233A. For all a = GlcNAc-Herceptin, 1 = glycosylated-heavy chain and 2 = GlcNAc-heavy chain.



Graph 2.2: Approximate conversion to the glycosylated product using different EndoS mutants over time from densitometry of bands on the SDS-PAGE

From the glycosylation tests, it was found that all the EndoS mutants had glycosylation activity and that the desired product was formed. It was seen that the activity is low for EndoS D233A, whereas both EndoS D233E and EndoS D233Q resulted in a similar conversion to the glycosylated product. It was observed that EndoS D233E was the most active and formed the product faster than EndoS D233Q, yet EndoS D233E also had greater hydrolytic activity and so there is only a small window of time to obtain the greatest yield of the desired product. For EndoS D233Q, hydrolysis activity was lower, as was glycosylation activity. It should be noted that Wang *et al.* reported that EndoS D235Q could be used to form the glycosylated product with no product hydrolysis,¹⁴ yet this is not seen here. Although all these EndoS mutants do show glycosylation activity, none could be considered to be an ENGase, with complete conversion to the glycosylation product not being observed.

2.4 – Building an EndoS Model and Forming Further Mutants

With lack of success in designing an ENGase, a different approach was needed. The critical issue is the same as reported for other Endo enzymes; that when mutations are made to the stabilising residue to decrease hydrolysis, glycosylation activity also decreases. It was therefore decided to investigate mutations of other residues in the active site that might bind to the substrate.

The major difficulty with the strategy of mutating other residues in the active site, in order to obtain the desired activity, is that no crystal structure of EndoS bound to a substrate was available. Due to this, the crystal structure of EndoF₃⁴ was studied instead; EndoF₃ is the closest related protein to EndoS that the crystal structure has been solved bound to a substrate. Looking at the crystal structure of EndoF₃ (Figure 2.9), four residues can be seen interacting with the substrate; D126 (stabilising residue), E128 (catalytic residue), Q211 and Y213. As mutation of the catalytic residue would be expected to abolish all activity and the effect of mutating the stabilising residue already studied, attention was focused on Q211 and Y213. Sequence alignment of the catalytic domain of EndoS (residue 37 – 445) with Endo F₃ using T-Coffee²⁶ showed that Q211 and Y213 are both conserved in EndoS as Q303 and Y305, respectively. A primitive model of EndoS was also formed with this alignment. This was done using the SWISS-MODEL²⁷⁻²⁹ homology-modelling server, using the crystal structure of EndoF₃ as the structure the catalytic domain of EndoS was to be modelled on (Figure 2.10), attempting to repeat the model reported by *Allhorn et al.*¹⁹. Although this model was predicted to contain errors, Q303 and Y305 were both seen to be positioned so that they could

interact with substrate and the catalytic residue in the active site, lending support to the idea that these residues could be important for the activity of EndoS.

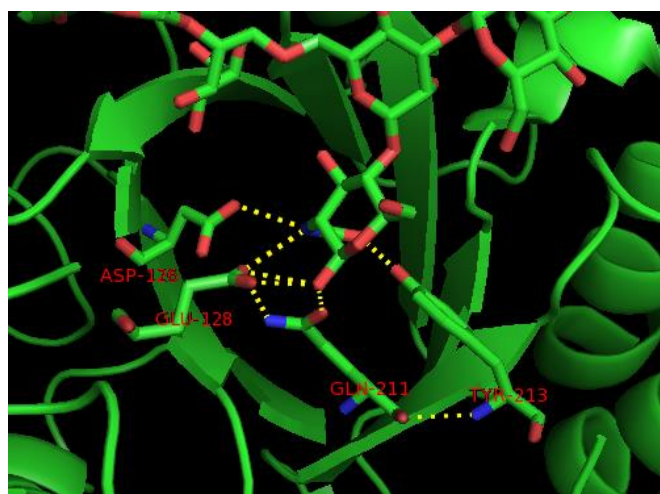
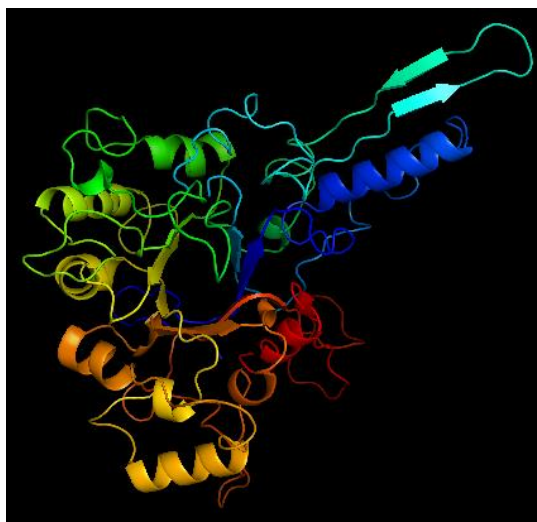


Figure 2.9: Crystal structure of EndoF₃ bound to a substrate, with residues that predicted to bind to the substrate (N126, E128, Q211 and Y213) shown.⁴

a)



b)

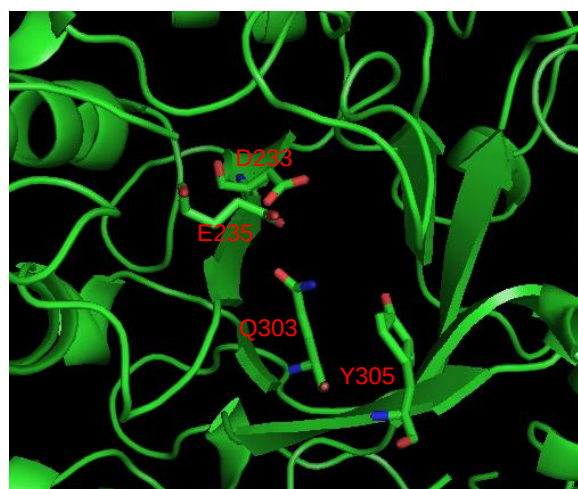


Figure 2.10: Model of the catalytic of EndoS (residue 37 – 445) form through sequence alignment with EndoF₃ and SWISS-MODEL homology-modelling server, a) Overview of the predicted structure of the EndoS catalytic domain, b) Residue predicted to bind to the substrate.

With these two residues being predicted to interact in the active site of EndoS, the decision was made to form EndoS Y305F and EndoS D233A Q303E. For EndoS

Y305F, it can be seen in the crystal structure of EndoF₃ that Y213 (conserved as Y305 in EndoS) interacts with the substrate. Assuming that a similar interaction occurs for EndoS, EndoS Y305F would remove this interaction, whilst maintain any π - π stacking or other interactions this residue is involved in. Removal of the interaction with the substrate was thought to possibly discourage the formation of oxazolinium ion intermediate that occurs during hydrolysis and hence reduce hydrolysis whilst retaining glycosylation activity. Evidence that this strategy may bear fruit was that EndoA Y205F (where Y205 is conserved as Y305F in EndoS) shows 44% of the hydrolytic activity of wild-type EndoA, but still retains 94% of the glycosylation activity.¹

The idea behind EndoS D233A Q303E relates back to EndoM N177Q. This is the only example where a mutation of the stabilising residue has resulted in decreased hydrolysis but retention of glycosylation activity on an Endo enzyme. As described above, it is thought that this is due to this mutation not stabilising the oxazolinium ion intermediate but still being able to interact with the catalytic residue as a hydrogen bond donor/acceptor (see Scheme 2.4). The elimination of this interaction with the catalytic residue may be why for EndoS, and for other Endo enzymes, mutation of the stabilising residue results in decrease of glycosylation activity. Evidence for this explanation comes from the findings so far of mutating the stabilising residue of EndoS, where the order of activity is Wild Type > D233E > D233Q > D233A. It can be envisaged that D233E and D233Q are able to interact with the catalytic residue more than D233A, hence giving weight to the idea that D233 interacts with catalytic residue and that this interaction is critical for both hydrolytic and glycosylation activity. Therefore, in order to form an ENGase EndoS mutant it is necessary to mutate D233 to remove hydrolytic

activity and then either this mutation or a second mutation near the catalytic residue needs to be able to interact with the catalytic residue and thus retain glycosylation activity.

It was thought that mutating Gln303 in EndoS to Glu may result in greater interactions with the catalytic residue and hence greater reactivity. Assuming that the proposed mechanism for EndoM also applies to EndoF₃ and EndoS and that the structure of EndoF₃ and EndoS is comparable, it was predicted that EndoS D233A Q303E should result in an enzyme that has diminished hydrolytic activity (through the D233A mutation) but retain some glycosylation activity (through the Q303E mutation).

Site directed mutagenesis was used to install the desired mutations and expression in component cells, followed by purification by GSTrap column, gave the desired GST-EndoS Y305F and GST-EndoS D233A Q303E, as confirmed by SDS-PAGE (Figure 2.11). The enzymes were then tested for hydrolytic and glycosylation activity as described for the previous EndoS mutants. It should be noted that the GST tag has no effect on the activity of EndoS. For GST-EndoS Y305F, high hydrolytic activity was seen (Figure 2.12 and Graph 2.3), showing that this mutation does not affect the formation of the oxazolinium ion intermediate and EndoS hydrolytic activity. Due to this fact, GST-EndoS Y305F was not tested for its glycosylation activity.

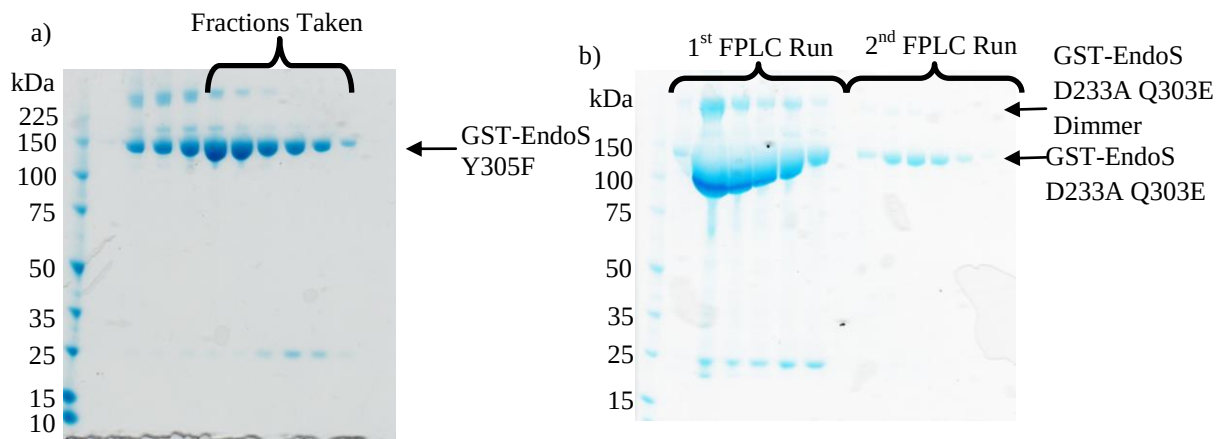


Figure 2.11: SDS-PAGE of the purification of GST-EndoS Y305F and GST-EndoS D233A Q303E, a) size-exclusion chromatograph of GST-EndoS Y305F and b) GSTrap purification of GST-EndoS

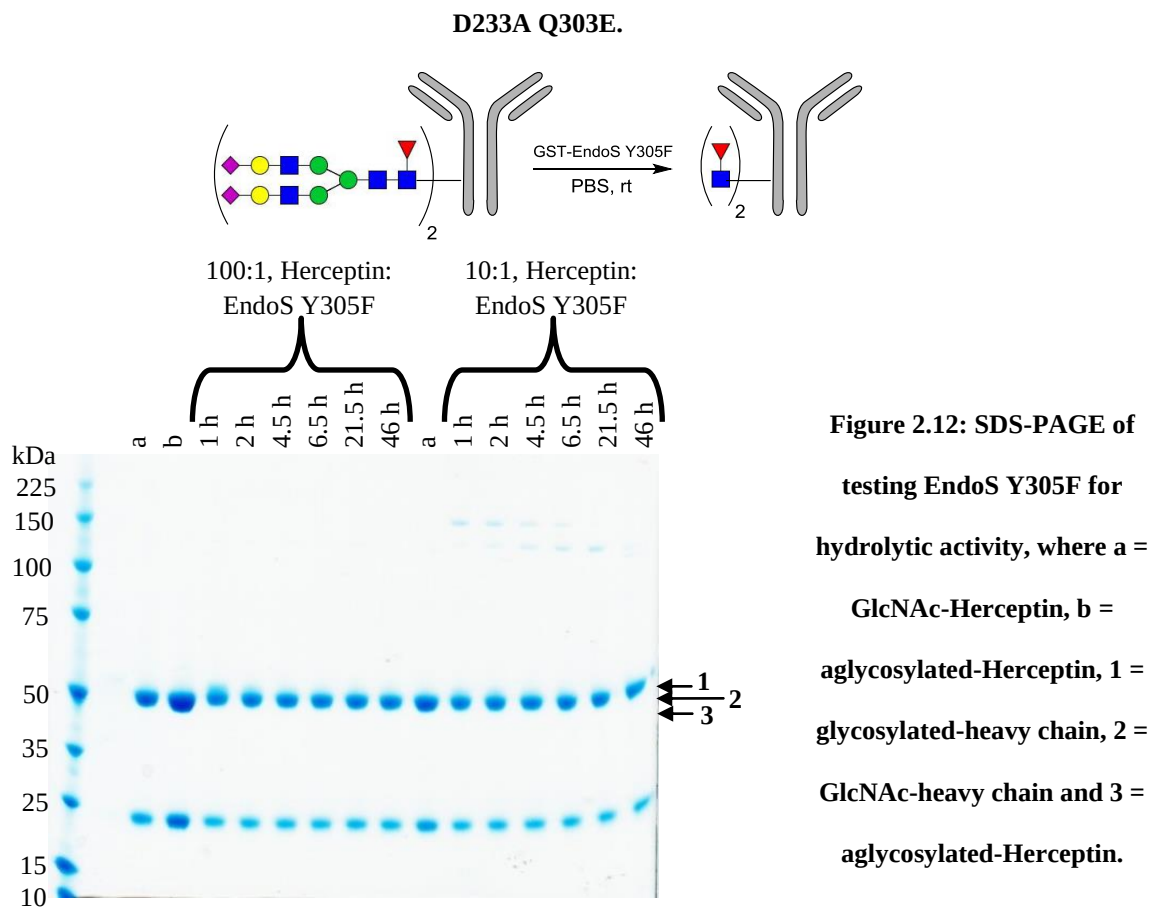
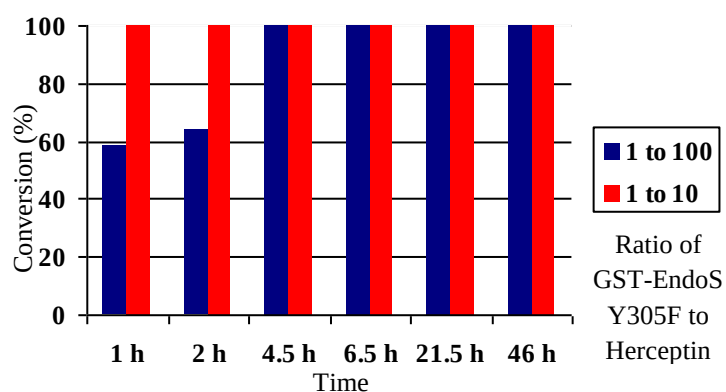


Figure 2.12: SDS-PAGE of testing EndoS Y305F for hydrolytic activity, where a = GlcNAc-Herceptin, b = aglycosylated-Herceptin, 1 = glycosylated-heavy chain, 2 = GlcNAc-heavy chain and 3 = aglycosylated-Herceptin.



Graph 2.3: Approximate conversion to the hydrolysed product using GST-EndoS Y305F over time from densitometry of bands on the SDS-PAGE.

Considering GST-EndoS D233A Q303E, hydrolytic activity against Herceptin was seen be similar to that of EndoS D233A; significantly reduced, although hydrolysis was seen after 1 h with high amounts of EndoS D233A Q303E and after 3 h with low amounts (Figure 2.13 and Graph 2.4). This was as expected. For glycosylation, there was an increase in activity for EndoS D233A Q303E compared to the single mutant EndoS D233A (Figure 2.14 and Graph 2.5), with the caveat that the densitometry data is only semi-quantitative and may not reflect the true enzymatic activity. It would seem that the Q303E mutation had a positive effect on glycosylation activity. However, compared to EndoS D233E and EndoS D233Q, the glycosylation reactivity of EndoS D233A Q303E was low, both in maximum conversion to the desired product and in time taken to reach maximum conversion and thus EndoS D233A Q303E cannot be considered to be a true ENGase. The results here suggest that mutating Q303 has less effect than mutating D233, the stabilising residue, which would be expected due to the implicit role that D233 has in the mechanism of action of EndoS. That said, these findings would also suggest a role for Q303 in the EndoS active site.

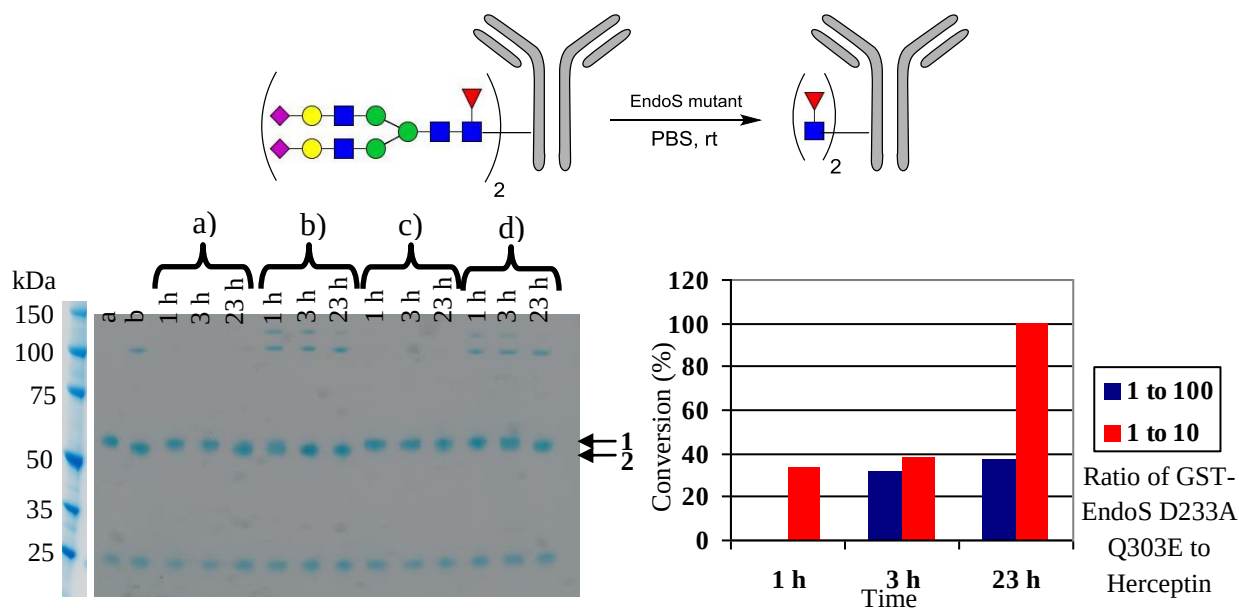


Figure 2.13: SDS-PAGE of testing EndoS D233A and EndoS D233A Q303E for hydrolytic activity, where a = glycosylated-Herceptin, b = GlcNAc-Herceptin. a) 1:100, EndoS D233A:Herceptin, b) 1:10, EndoS D233A:Herceptin, c) 1:100, EndoS D233A Q303E:Herceptin and d) 1:100, EndoS D233A Q303E:Herceptin. 1 = glycosylated-heavy chain, 2 = GlcNAc-heavy chain.

Graph 2.4: Approximate conversion to the hydrolysed product using GST-EndoS D233A Q303E over time from densitometry of bands on the SDS-PAGE.

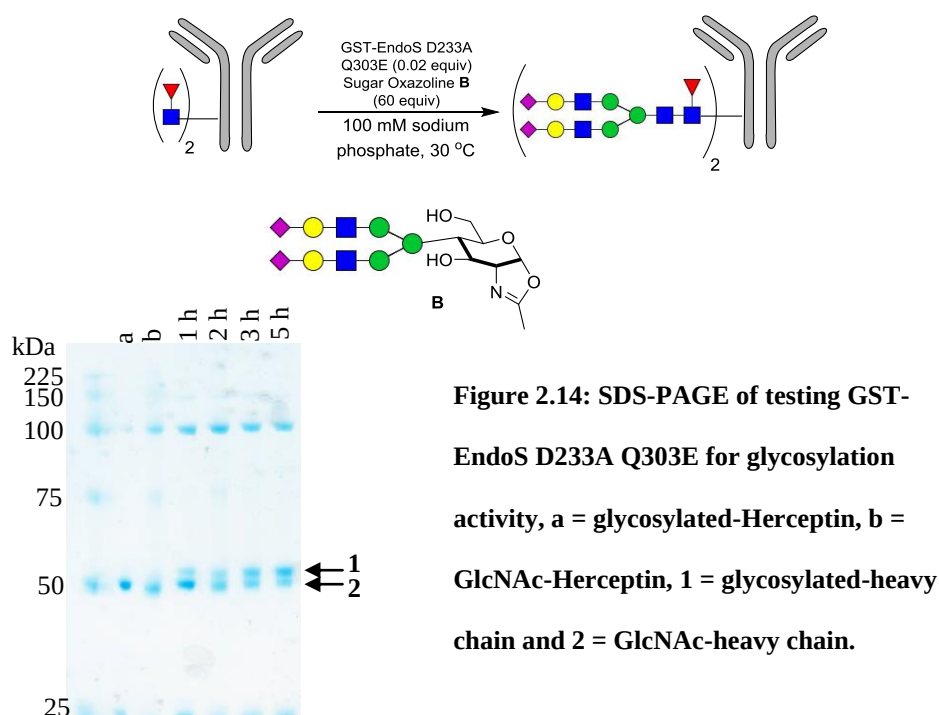
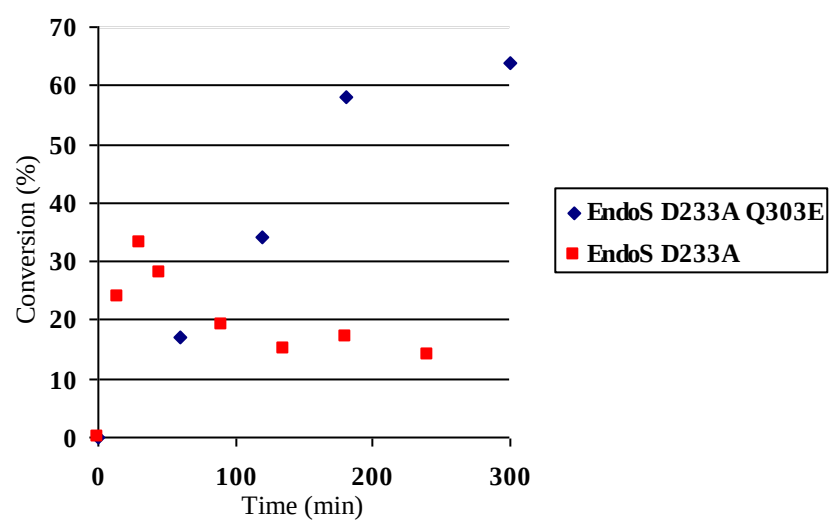


Figure 2.14: SDS-PAGE of testing GST-EndoS D233A Q303E for glycosylation activity, a = glycosylated-Herceptin, b = GlcNAc-Herceptin, 1 = glycosylated-heavy chain and 2 = GlcNAc-heavy chain.



Graph 2.5: Approximate conversion to the glycosylated product using EndoS D233A Q303E over time from densitometry of bands on the SDS-PAGE. Endo D233A is shown for comparison.

2.5 – Attempts to Obtain a Crystal Structure of EndoS

Although a number of the EndoS mutants formed above showed promising glycosylation activity, no true EndoS ENGase had been formed. Therefore, it was decided to try to obtain the crystal structure of EndoS. Having the actual structure of EndoS, as opposed to using models based on EndoF₃, could give critical information into other amino acids in the EndoS active site that bind to the substrate and thus are interesting targets to mutate in order to form an EndoS ENGase.

As stated above, the catalytic domain of EndoS (amino acids 37 – 445) is predicted to have a structure similar to EndoF₃ both by computational studies and amino acid sequence alignment.¹⁹ It was therefore thought that this catalytic domain could be a good initial target to obtain the crystal structure of, as this is smaller than the mature protein, yet is the active fragment. Also, if this fragment could be expressed, its activity could be compared to native EndoS and may suggest which part of EndoS causes the IgG specificity.

To investigate this, site directed mutagenesis was used to form the plasmid encoding EndoS K446Stop. This plasmid was used to attempt to express just the catalytic domain. Unfortunately, very little protein was expressed, especially compared to the previously expressed GST-EndoS, wild type and mutants. A number of different expression conditions were studied, but none of these resulted in a positive effect and the expression of EndoS K446Stop (Table 2.1 and Figure 2.15).

Cells	Induction Method	Expression Temperature	Result
BL21 (DE3)	IPTG	30 °C	No Expression
BL21 (DE3)	IPTG	15 °C	No Expression
BL21 (DE3)	Auto	30 °C	No Expression
BL21 (DE3) pLysS	IPTG	30 °C	No Expression

Table 2.1: Conditions used to try to express EndoS K446Stop.

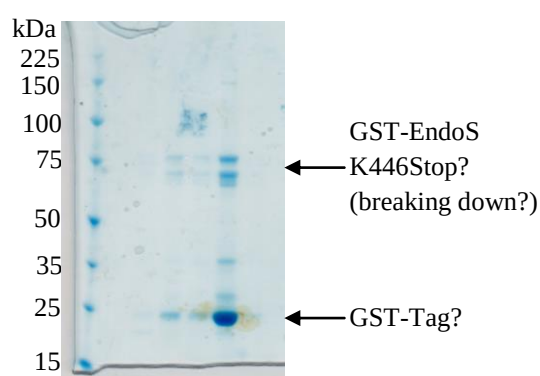


Figure 2.15: Attempt at purifying GST-EndoS K446Stop with GSTrap

With GST-EndoS expressing extremely well, it was thought that the reason why GST-EndoS K446Stop was not being observed was not due to poor expression but instead the protein not being stable and rapidly degrading after expression.

Unable to obtain just the catalytic domain of EndoS, attempts were made at gaining the crystal structure of the whole protein. As stated above, it was desired to obtain the crystal structure of EndoS to see which residues are critical for its activity and binding to the substrate. Therefore, crystal screens were also set up with a thiazoline ligand^{30, 31}, a ligand that is known to bind in the active site of Endo enzymes. A large number of different screens were set up, with and without the thiazoline ligand and at different concentrations. After 7 days a number of interesting precipitates were obtained (Figure

2.16). It was seen that high concentrations and the presence of the ligand resulted in more crystalline precipitates. Attempts made at optimising these conditions resulted in the formation of small crystals. These were too small to use to obtain the X-ray structure, but were used instead of seeds in further crystallisation attempts. Disappointingly, no crystals were obtained that were large enough to be of use.

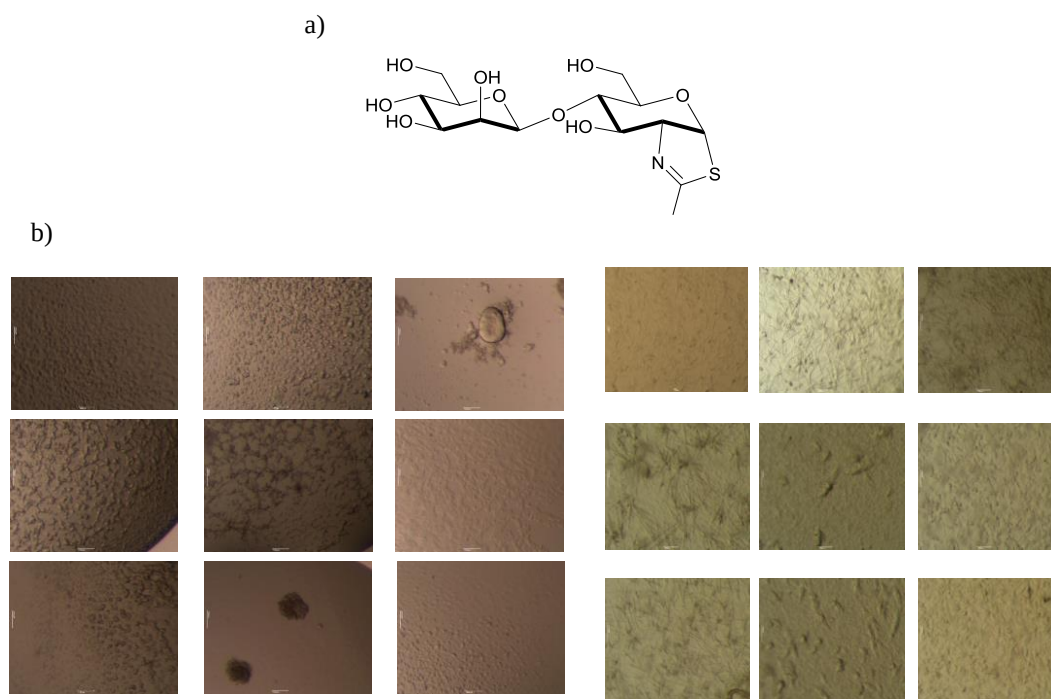


Figure 2.16: Crystallisation trials of EndoS, a) thiazoline ligand used in the crystallisation screen, b) precipitates and small crystals obtained from crystallisation trials with EndoS and EndoS mutants.

Optimisation of these conditions did not yield any improvement.

It was thought that the lack of crystals could be due to the large size of EndoS and that perhaps EndoS could take a number of different conformations, which would make crystallisation difficult. It was believed that if EndoS was bound to IgG, it would be in one conformation and hence avoid this issue. This would be possible with EndoS E235Q, an EndoS mutant that is reported to show no activity but instead binds to IgG.²⁰

However, it would be extremely challenging to obtain a crystal of EndoS bound to IgG, with there being very few reports of crystal structures of IgG. Therefore, instead of using whole IgG, it was decided to use IgG1 Fc (Figure 2.17), the sugar bearing region of the antibody, with previous reports stating that EndoS is active on this fragment of IgG.²² This fragment is significantly smaller than IgG and there are previous reports of its crystal structure.³²

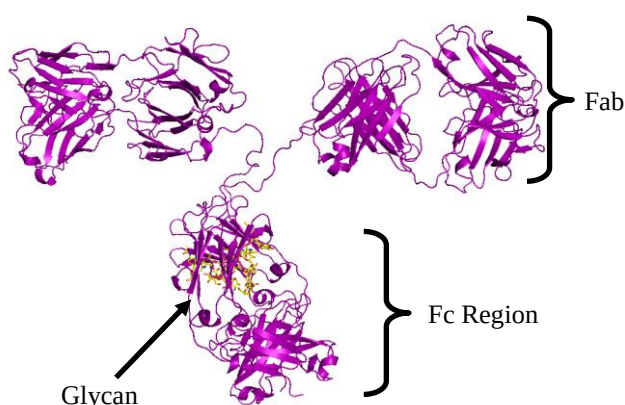


Figure 2.17: Structure of IgG. Fab, Fc and glycan (in yellow) are shown. Crystal structure from McPherson et al.³³

Site directed mutagenesis was used to form a plasmid encoding for GST-EndoS E235Q and this protein was expressed. Cleavage of the GST-tag by TEV and purification by size exclusion chromatography afforded EndoS E235Q (Figure 2.18).

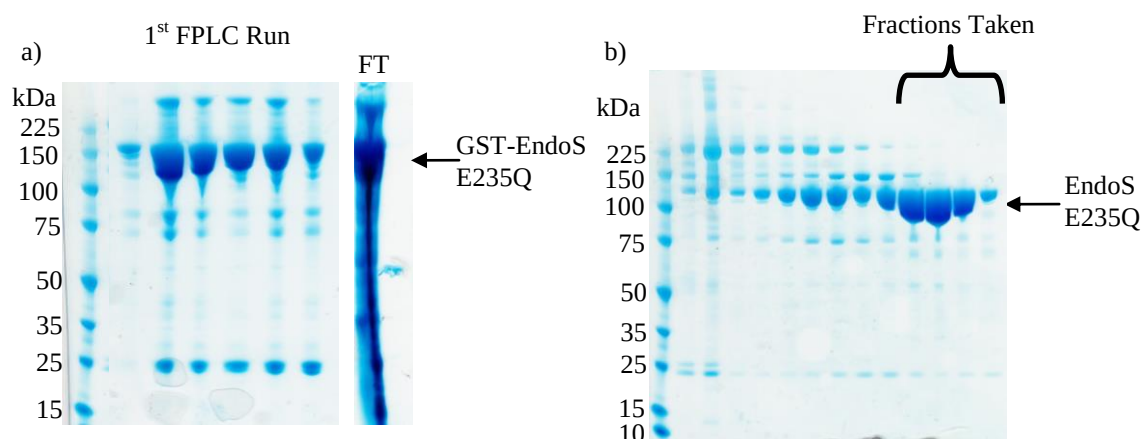


Figure 2.18: SDS-PAGE of purification of EndoS E235Q, a) GSTrap purification of GST-EndoS E235Q (Note, FT = Flow through, which was seen to still contain a large amount of desired proteins), b) Size exclusion chromatography.

With EndoS E235Q in hand, next IgG1 Fc had to be expressed. Being a glycosylated protein, IgG1 Fc could not be expressed in *E. coli* but instead expression was conducted in HEK293T cells.^{32, 34} Briefly, the pHLSec plasmid encoding for IgG1 Fc was transformed into XL-10 gold cells and an overnight culture was prepared from a single colony. The resulting cells were harvested and the plasmid collected via a MegaPrep kit. This plasmid was then transfected into the HEK293T cells that had been grown to confluence in Hyperflasks, using polyethylenimine.hydrogen chloride (Mw ~25000 Da) as the transfection reagent. After being left for a number of days to allow expression, the media, containing the expressed IgG1 Fc, was collected and purification conducted with a metal affinity column (the IgG1 Fc was expressed with a His-tag) and size exclusion. Presence of the protein was confirmed by SDS-PAGE and western blot (Figure 2.18). IgG1 Fc was obtained as a dimer (Sample 1, Figure 2.18c) and a monomer (Sample 2), which were both utilised for crystallisation trials.

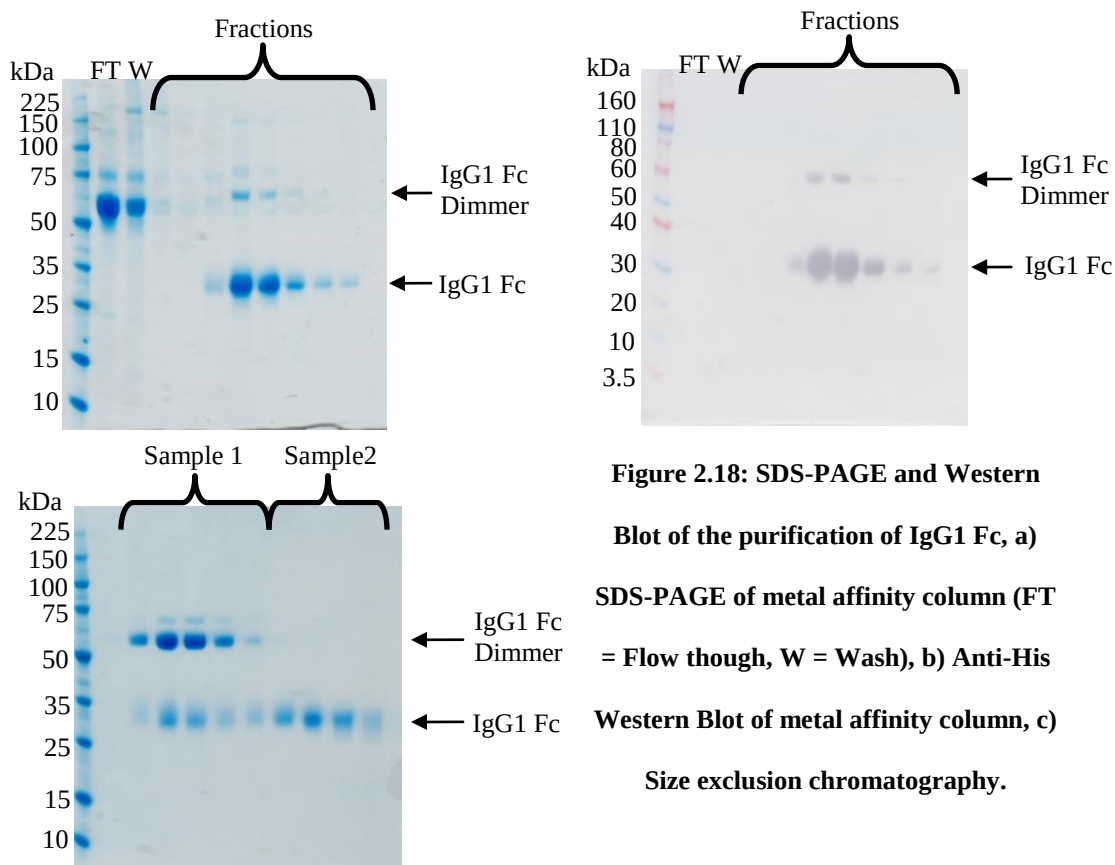


Figure 2.18: SDS-PAGE and Western Blot of the purification of IgG1 Fc, a) SDS-PAGE of metal affinity column (FT = Flow through, W = Wash), b) Anti-His Western Blot of metal affinity column, c) Size exclusion chromatography.

With both IgG1 Fc and EndoS E235Q being expressed and purified, crystallisations screens were set up, again with a variety of different conditions. As before, some of the conditions resulted in the formation of crystalline precipitates, which were optimised and resulted in the formation of small crystals being formed (Figure 2.20).

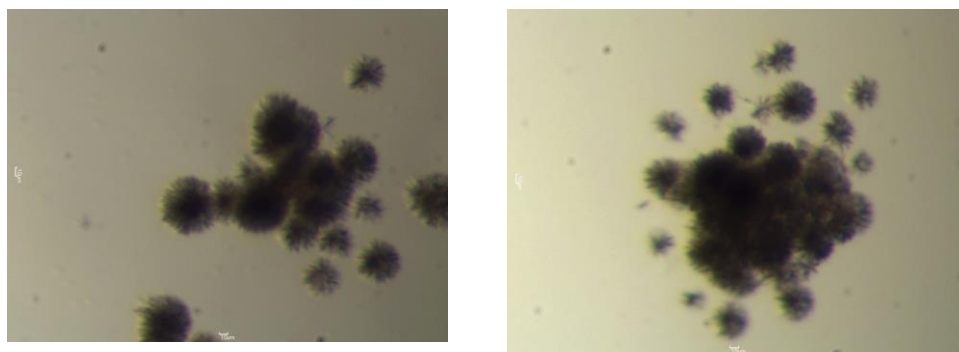


Figure 2.20: Small crystals obtained through optimisation of the crystallisation of EndoS E235Q with IgG1 Fc (0.1 M sodium acetate/bromide, 18% PEG 3550).

Shortly after further optimisation was started on these small crystals, *Sundberg et al.*³⁵ reported that they were able to obtain a crystal of EndoS(98-995) using liquid-liquid diffusion. They stated that residues 1-97 forms the signalling peptide and a putative coiled-coil structure that they predicted would encourage the formation of oligmers that are likely to inhibit crystallisation. They attempted to obtain a crystal of EndoS(98-445), the catalytic domain of EndoS, however, they state that this immediately degraded after purification, agreeing with the findings here when attempting to express GST-EndoS K446Stop. Their initial attempts at obtaining a crystal of EndoS(98-995) with conventional sitting drop vapour diffusion methods, as tried here, did result in the formation of crystals and diffracted at 3.5 Å resolution but the crystals suffered from anisotropy and the structure could be only processed to a 7.5 Å resolution. When using

liquid-liquid diffusion, though, they were able to obtain crystals with greatly diminished anisotropy and could process them to a 2.6 Å resolution.³⁵

From the crystals, they were able to solve the structure of EndoS(98-995).³⁶ They found that EndoS has 5 distinct domains; the catalytic domain (residues 98-445, Figure 2.21), a leucine-rich repeat (residue 446-631), a hybrid Ig domain (residues 632-764), a carbohydrate-binding domain (residues 765-923) and a three-helix bundle (residues 924-995). The structure confirmed that the catalytic domain adopts the $(\beta/\alpha)_8$ barrel fold common to other Endo enzymes and is very similar in structure to EndoF₃, although the loops outside the active site differ between the catalytic domain of EndoS and EndoF₃, maybe explaining the different glycan specificities of the two enzymes (EndoF₃ cleaves tri- and biantennary complex glycans, whereas EndoS only cleaves biantennary complex glycans). Furthermore, there are a number of different tryptophan residues which are thought to be important for enzymatic activity and specificity and are predicted to accommodate the protein and oligosaccharide substrate. These include Trp121, Trp153, Trp161, Trp314 and Trp358.

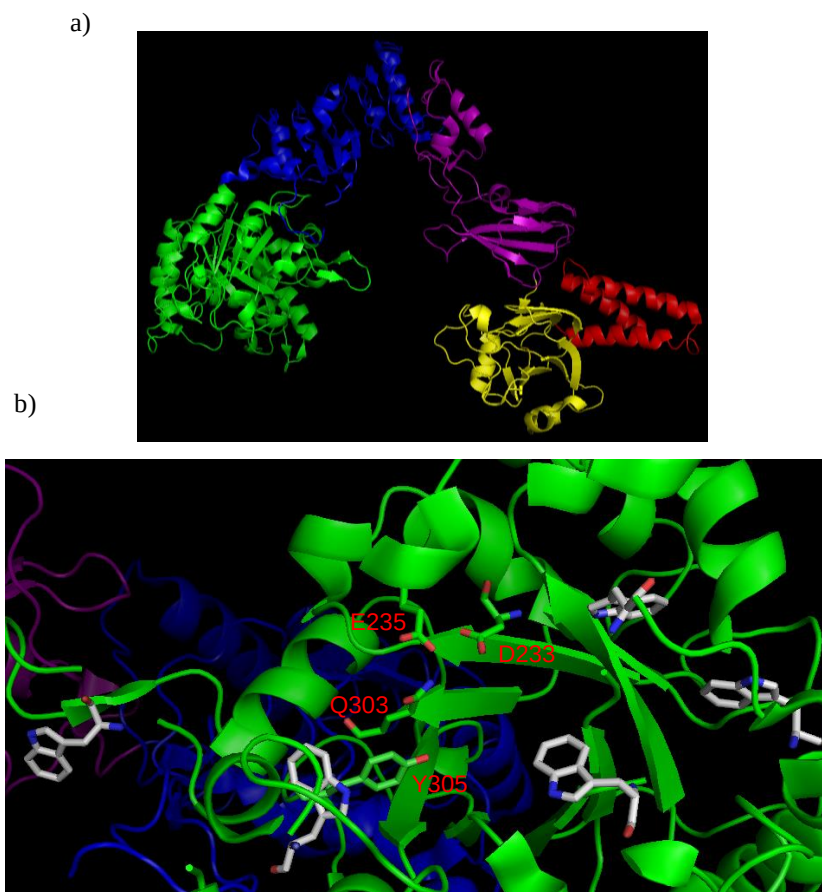


Figure 2.21: Crystal structure of EndoS,³⁶ a) Overall structure of EndoS (green = catalytic domain, blue = leucine-rich repeat, purple = hybrid Ig domain, yellow = carbohydrate-binding domain and red = three-helix bundle), b) EndoS active site, residues predicted to be important here are labelled and the important Trp residues coloured grey.

With regards to the other domains, through a number of different experiments (mutagenesis experiments, surface plasma resonance and computational modelling), they found that the leucine-rich region is critical to protein stability and the Ig region, which is similar in structure to IgG, contributes to the binding to IgG and hence to the IgG specificity. The carbohydrate-binding domain binds to the glycan on IgG and mutations at certain residues in this region (W803, E833 and R908), result in reduced binding to IgG and/or reduced hydrolytic activity.

2.6 – Conclusions and Future Work

A number of different EndoS mutants have been formed, attempting to form an ENGase; an EndoS mutant that is capable of high glycosylation activity and low hydrolytic activity, resulting in high yields of the glycosylation product. Studies here have found that EndoS D233Q best fits this description, although this EndoS mutant still shows some hydrolytic activity and only a moderate amount of glycosylation activity. For EndoS D233E, it was hoped that the extra methylene group on the stabilising residue would discourage the formation of the oxazolinium ion intermediate required for hydrolysis and that the glycosylation activity would be unaffected. Here, EndoS D233E showed significant amount of glycosylation activity, being more active than EndoS D233Q, but also showed high hydrolysis activity and hence was unsuitable for use as an ENGase. The high activity is presumed to be due to the D233E mutation being subtle and stabilisation of oxazolinium ion still feasible. In contrast, EndoS D233A was shown to have very low hydrolytic activity due to the stabilising residue changing from being ionic to hydrophobic and hence unable to stabilise the oxazolinium ion for hydrolysis. But the glycosylation activity of EndoS D233A was seen to be poor. The findings from EndoS D233E/Q/A all suggest that the stabilising residue not only stabilises the oxazolinium ion for hydrolysis but is also involved in the catalysis of glycosylations, where more significant mutations result in a greater decrease in activity. This is most likely due to interaction of the stabilising residue with the catalytic residue, as seen for EndoM (Scheme 2.3).

An attempt at regaining this interaction with EndoS D233A was made, forming EndoS D233A Q303E, predicting that the Q303E residue may replace the interaction with the

catalytic residue lost through the D233A mutation. This mutant showed similar hydrolytic activity to EndoS D233A and, although the densitometry is only semi-quantitative, an increased glycosylation activity was observed. This would strongly suggest the Q303, and conserved residues in other Endo enzymes, has a role in the active site, as suggested from the computational model. Further studies are warranted with EndoS D233A Q303E to confirm the effect of the Q303E mutation and to investigate this novel mutation on other Endo enzymes. Work also continues in the group studying EndoS D233Q, as it has the highest ENGase activity of EndoS mutant to date, and its utility into forming homogeneous glycosylated antibodies that have therapeutic properties.

Attempts to obtain the crystal structure of EndoS were also made here. However the crystal structure was obtained by *Sundberg et al.*^{35, 36} before it could be solved here. But although this crystal structure of EndoS gives a lot of valuable information about residues in the active site, only the crystal structure of the protein on its own has been solved. To confirm the critical residues required for EndoS activity and specificity, it would be more informative to have a crystal structure of EndoS bound to a substrate. Further work could be done to obtain this using the same crystallisation methods used by *Sundberg et al.* to obtain the crystal structure of EndoS bound to a substrate, which could then be used to fully inform EndoS mutation studies.

2.7 - Biological Experimental

2.7.1 – General Protocols

2.7.1.1 - SDS-PAGE

Sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) was carried out to check the purity of all proteins. These were carried out using 4 – 12% bis-tris polyacrylamide gels with either MOPS or MES buffer. To the protein sample (10 µL) was added 2 × SDS loading buffer (10 µL 80 mM Tris-HCl buffer pH 6.8, 8% v/v glycerol, 40 µM bromophenol blue, 80 mM SDS, 2% β-mercapethanol). The solution was heated to 95 °C for 5 min. The sample was applied to a well in the gel. The gel was run at 200 V for 1 h and visualised by the addition of Coomassie InstantBlue dye (Gentauro) and gentle stirring for 2 h at room temperature.

2.7.1.2 - Western Blot

SDS-PAGE was done as before, but using a Pre-Stained protein ladder and the gel was not stained with InstantBlue dye. Instead, the gel was electroblotted onto a membrane for 7 min using an iBlot machine. The membrane was then washed with 50 mM Tris-HCl, 150 mM NaCl and 0.05% Tween 20, pH 7.6 (TBST buffer, 45 mL) containing BSA (2 g) for 1 h. The membrane was then washed with TBST buffer (3 × 20 mL) before the addition of TBST buffer (45 mL) containing BSA (0.5 g) and monoclonal anti-polyhistidine-alkaline phosphatase antibody produced in mouse (Sigma, 16 µL). The membrane was then rocked for 4 h at room temperature. After this time, the solution was removed and the membrane washed with TBST buffer (3 × 20 mL). BCIP/NBT alkaline phosphatase substrate (4 mL) was added and the membrane was

gently rocked until protein bands could be seen (~15 min). The membrane was then washed with water.

2.7.1.3 - Preparation of Lysogeny Broth Agar Plates

Lysogeny broth (LB, 6.25 g) and agar (3.75 g) were both added to water (250 mL) and the mixture was autoclaved. After cooled to ~45 °C, a solution of ampicillin (0.25 mL, 100 mg/mL) was added and the agar was poured into plates and allowed to set. Plates were stored at 4 °C.

2.7.2 – Mutagenesis on EndoS Plasmid and Plasmid Preparation

2.7.2.1 - Formation of EndoS Mutant Plasmids

Site directed mutagenesis was undertaken on the pGEX-4T-1-ndoS plasmid, which codes for GST-Tagged EndoS, using a Stratagene Quikchange II Site-Directed Mutagenesis Kit with the following primers, with the mutated nucleotides highlighted (Primers given 5' to 3'). Note: For the insertion of the TEV cleavage site between EndoS and the GST-tag, two rounds of mutagenesis were conducted using the EndoS TEV1 primers first, followed by a second round using EndoS TEV2 primers, to insert the necessary bases.

Primer	Code (mutated bases highlighted)
EndoS TEV1 Forward	CTGGTTCCGCGT GAGAACCTT GGATCCATGGACAAAC
EndoS TEV1 Reverse	GTTTGTCCATGGATCC AAGGTTCTC ACGCGGAACCAG
EndoS TEV2 Forward	CCGCGTGAGAACCTT TATTTCCAG GGATCCATGGACAAA C
EndoS TEV2 Reverse	GTTTGTCCATGGATCC CTGGAAATA AAGGTTCTCACGCG G
EndoS D233A Forward	CTGGACGGCCTGGATGTGG CT GTTGAACATGATTCAATTC
EndoS D233A Reverse	GAATTGAATCATGTTCAAC AG CCACATCCAGGCCGTCCA G
EndoS D233E Forward	CTGGACGGCCTGGATGTGGA A GTTGAACATGATTCAATT C
EndoS D233E Reverse	GAATTGAATCATGTTCAAC T TCCACATCCAGGCCGTCCAG
EndoS D233Q Forward	CTGGACGGCCTGGATGTGGA A GTTGAACATGATTCAATT C
EndoS D233Q Reverse	GAATTGAATCATGTTCAAC T TCCACATCCAGGCCGTCCAG
EndoS Q303E Forward	CAACCTGCTGCTGGT G CAGGTGTATGG
EndoS Q303E Reverse	CCATACACCTG C ACCAGCAGCAGGTTG
EndoS K446Stop Forward	CGCTGAAAACCGTGATGCTG T AAGATAAATCTTACGATC TG
EndoS K446Stop Reverse	CAGATCGTAAGATTTATCTT A CAGCATCACGGTTTTTCAGC G
EndoS E235Q Forward	CCTGGATGTGGACGTT C AACATGATTCAATTC
EndoS E235Q Reverse	GAATTGAATCATGTT G AACGTCCACATCCAGC

Note, EndoS Y305F plasmid was formed by another member of the group.

The following were all added together: Milli-Q water (35 μ L), 10 $\times$ Reaction Buffer (5 μ L), DMSO (3 μ L), pGEX-4T-1-ndoS plasmid (1 μ L, 50 ng/ μ L), forward primer (2 μ L, 62.5 ng/ μ L), reverse primer (2 μ L, 62.5 ng/ μ L), dNTP mix (1 μ L, 10 mM) and Pfa Ultra HF DNA Polymerase (1 μ L, 2.5 U/mL).

The solutions were then subject to the following PCR cycle:

95 °C – 2 min
 $\left(\begin{array}{l} 95\text{ °C} - 30\text{ sec} \\ 55\text{ °C} - 30\text{ sec} \\ 72\text{ °C} - 10\text{ min} \end{array} \right) \times 25$
10 °C - ∞

After this, Dpn 1 restriction enzyme (2 µL, 10 U/mL) was added and the mixture was incubated at 37 °C for 1 h. Transformation was then carried out as described below.

2.7.2.2 - Transformation of Mutant Plasmids into XL-Gold Ultracompetent Cells and Preparation of Cell Pellet for Miniprep

To a pre-chilled tube, XL-Gold Ultracompetent Cells (50 µL) and XL-Gold β-mercaptoethanol mix (2 µL) were added. After incubating for 10 min on ice, a portion of the PCR mix (2 µL) was added and the mixture was incubated on ice for a further 30 min. After this time, the mixture was subject to a 42 °C heat pulse for 30 sec before incubating on ice for a further 2 min. SOC media (0.5 mL) was added and the mixture incubated at 37 °C and 220 rpm for 1 h. After this time, a portion of the mix (50 µL) was added to an LB-Amp agar plate and was incubated overnight at 37 °C to allow colonies to form. The plates were then stored at 4 °C

A single colony from one of the plates was added to LB-Amp media (8 mL, 8 µL 100 mg/mL Amp) and this mix was incubated overnight at 37 °C and 220 rpm. The mixture was then centrifuged (3000 rpm, 15 min, 4 °C) and the supernatant discarded. The cell pellet was stored at –20 °C.

2.7.2.3 - Preparation of the Plasmid and Sequencing

Preparation of the mutant plasmid for sequencing was carried out using a MiniPrep Kit purchased from QIAGEN and following the included manual. The concentration of the resulting plasmid was measured using a Nanodrop spectrometer. A sample was prepared and submitted to Geneservice at the Department of Biochemistry, University of Oxford for sequencing. The plasmid was sequenced with pGEX 3', pGEX 5' and custom designed primers (given below) to insure that the mutation was in place and no undesired mutations had occurred. EndoS Primer allowed sequencing of the middle part of the EndoS gene, whereas the GST Primer allowed sequencing to confirm the insertion of the gene sequence to give the TEV cleavage site.

Primer	Code
GST Primer	CGTATTGAAGCTATCCCAC
EndoS Primer	GTCCGGAACAATACATGATCG

2.7.3 – EndoS and EndoS Mutants Expression and Purification

2.7.3.1 - Transformation of Plasmids into BL21 (DE3) and BL21 (DE3) pLysS Cells

To a pre-chilled tube, BL21 (DE3) or BL21 (DE3) pLysS Cells (50 μ L) and the desired plasmid (2 μ L, ~100 ng/ μ L) were added and incubated on ice for 30 min. After this time, the mixture was subject to a 42 °C heat pulse for 30 sec before incubating on ice for a further 2 min. SOC media (0.5 mL) was added and the mixture incubated at 37 °C and 220 rpm for 1 h. A portion of the mix (50 μ L) was added to an LB-Amp agar plate and was incubated overnight at 37 °C to allow colonies to form. The plates were then stored at 4 °C.

2.7.3.2 - Over-Expression of GST-EndoS using IPTG Induction

Either a small sample of a glycerol stock (glycerol stocks were formed from 200 μ L 60% glycerol and 600 μ L of an overnight culture) of *E. coli* containing the desired plasmid or a colony from a competent cell line agar plate was added to LB-Amp media (20 mL, 20 μ L 100 mg/mL Amp). After incubating this media at 37 °C overnight, a portion of it (10 mL) was added to LB-Amp media (1 L, 1 mL 100 mg/mL Amp). The media was incubated for 3 h, when the OD was approximately 0.5-1 and the culture was induced by the addition of IPTG (5 mL, 100 mmol). The cultures were incubated for 5 h at 30 °C. After this time, the cultures were combined and centrifuged (8000 rpm, 4 °C, 10 min) and the cell pellet collected. The cell pellet was stored at –80 °C.

2.7.3.3 - Over-Expression of GST-EndoS using Auto-Induction

An overnight culture was set up as described for the over-expression when using IPTG induction. A portion of the overnight culture (5 mL) was added to ZYM-5052 medium. This medium was made by the addition of tryptone (5 g), and yeast extract (2.5 g) to water (477 mL) and, after autoclaving, the addition of 50 $\times$ M (1.25 M Na₂HPO₄, 1.25 M KH₂PO₄, 2.5 M NH₄Cl and 0.25 M Na₂SO₄ 10 mL), 50 $\times$ 5052 (25% glycerol, 2.5% glucose and 10% α -lactose in water, 10 mL), magnesium sulfate solution (1 M, 1 mL), iron(III) chloride solution (50 mM, 1mL) and Amp solution (100 mg/mL, 400 μ L). The ZYM-5052 medium was incubated at 28 °C for 22 h. After this time, the cultures were combined and centrifuged (8000 rpm, 4 °C, 10 min) and the cell pellet collected. The cell pellet was stored at –80 °C.

2.7.3.4 – Lysis of Cells and Purification of GST-EndoS

Having been thawed on ice, the cell pellet was resuspended in PBS buffer (~3 mL per 1 g cell pellet) containing lysozyme (1 mg per 10 mL PBS) and Complete protease inhibitors (1 tablet). After resuspending, the suspension was sonicated (4 × 1 min, 1 sec pulse, 0.5 pause, 1 min intervals between cycles), centrifuged (20000 rpm, 20 min) and the supernatant filtered (0.2 µm membrane). The solution was applied to FPLC using two GStrap HP columns (2 × 5 mL columns) and PBS buffer (buffer A, loading buffer) and 30 mM Tris-HCl, 10 mM reduced L-glutathione (pH 8) buffer (buffer B, elution buffer). The buffer gradient was 5% Buffer B for 2 column volumes (CV) to 100% Buffer B for 2 CV before washing the column with Buffer A with 3 CV. Flow rate was 1 mL/min. The flow through fraction was applied to FPLC a second time on occasion under the same conditions to ensure all desired protein was collected. After analysis by SDS-PAGE, all the fractions containing the desired protein were combined and concentrated by Vivaspin (MWCO 30000). This solution was dialysed into either PBS buffer or 50 mM Tris-HCl (pH 7.4) twice (15 h + 3 h) and subject to Bicinchoninic Acid assay (BCA assay). This solution was stored at -80 °C.

2.7.3.5 - Removal of GST-tag by TEV Cleavage and Purification by Size Exclusion

Chorography

The GST-tagged protein in 50 mM Tris-HCl (pH 7.4) was concentrated (using Vivapin) to 3 mg/mL and TEV was added (1 mg TEV to 25 mg protein). The solution was left to incubate overnight at room temperature. The solution was dialysed into PBS by using Vivaspin (MWCO 30000) to concentrate and then diluted with the buffer. This process was done three times to remove both the GST-tag and the TEV. The sample was subject

to gel filtration using a 320 mL Superdex 200 column on a FPLC with PBS buffer. Flow rate was 2.5 mL/min. Fractions seen to contain clean desired protein by SDS-PAGE were combined and concentrated using Vivaspin (MWCO 30000) and the concentration of the protein solution was found using BCA assay. Protein samples were then stored at -20°C

2.7.3.6 - Protein Yields

Protein	GST-tag Yield	Recombinant Yield	Purification Techniques
EndoS	27 mg/L	3.4 mg/L	GSTrap and Size Exclusion
EndoS D233E	80 mg/L	11 mg/L	GSTrap and Size Exclusion
EndoS D233Q	100 mg/L	35 mg/L	GSTrap and Size Exclusion
EndoS Y305F	300 mg/L	75 mg/L	GSTrap and Size Exclusion
EndoS D233A Q303E	33 mg/L	N/A	GSTrap
EndoS E235Q	70 mg/L	14 mg/L	GSTrap and Size Exclusion

Note, EndoS D233A was expressed by another member of the group.

2.7.4 – IgG1 Fc Expression

2.7.4.1 - DNA Preparation for Transfection

To a pre-chilled tube, XL-Gold Ultracompetent Cells (50 μL) and XL-Gold β -mercaptoethanol mix (2 μL) were added. After incubating for 10 min on ice, pHLSec plasmid encoding for IgG1 Fc (100 ng/ μL , 2 μL) was added and the mixture was incubated on ice for a further 30 min. After this time, the mixture was subject to a 42°C heat pulse for 30 sec before incubating on ice for a further 2 min. SOC media (0.5 mL) was added and the mixture incubated at 37°C and 220 rpm for 1 h. A portion of the mix (50 μL) was added to an LB-Amp agar plate and was incubated overnight at 37°C to allow colonies to form. The plates were stored at 4°C

A single colony from the plate was added to LB-Amp media (3×500 mL, 500 μ L, 100 mg/mL Amp) and this mix was incubated overnight at 37 °C and 220 rpm. The mixture was then centrifuged (8000 rpm, 10 min, 4 °C) and the supernatant discarded. The DNA was then extracted and purified using a MegaPrep kit (Qiagen), following the enclosed instructions. The amount of DNA was found to be 2.6 mg (1.7 mg/L media) by Nanodrop. The solution was sterile filtered (0.2 μ m membrane) and stored at –20 °C

2.7.4.2 - HEK293T Cell Growth

HEK293T cells were grown in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS) and 1% Penicillin-Streptomycin (10000 U/mL, Pen/Strep) at 37 °C, 5% CO₂. A vial of stock cells (1 mL) was first diluted with media (5 mL) and centrifuged (1200 rpm, 5 min). The media was discarded and the cells were resuspended in media (6 mL), added to a T25 flask and grown for 24 h. After this time, the cells were 90% confluent and were passaged (media discarded, cell washed with PBS) and transferred into 6 $\times$ T175 flasks with fresh media (6 $\times$ 50 mL). After a further 4 days of growth, the cells were passaged into Hyper flasks (2 $\times$ 500 mL) with fresh media (2 $\times$ 550 mL). After 3 days the cells were 90% confluent and the media in the flasks was discarded and transfection started.

2.7.4.3 - HEK293T Transfection

To 125 mL DMEM was added the plasmid DNA (0.5 mL 1.3 mg/mL in water, 650 µg) and Polyethylenimine “Max” (230 µL, 10 mg/mL in water, 2.3 mg). This solution was left for 20 min with occasional gentle mixing at room temperature. This solution was added to 500 mL DMEM supplemented with 2% FBS and 1% Pen/Strep (10000 U/mL), which was then added to a Hyper flask containing confluent HEK293T cells. The flasks were incubated at 37 °C, 5% CO₂ for 4 days.

2.7.4.4 - Purification by Metal Affinity Column and Gel Filtration

After 4 days, the media was collected and water (150 mL) was added to dilute. The media was centrifuged (7000 rpm, 4 °C, 30 min) and buffered by the addition of 1 M Tris-HCl pH 8 buffer (50 mL). It was then applied to a freshly recharged Talon Cobalt column (200 mL resin) and allowed to flow through. The column was washed with 10 mM Tris-HCl, 150 mM NaCl pH 8, 10 mM imidazole (300 mL) and the desired protein eluted with 10 mM Tris-HCl, 150 mM NaCl pH 8, 225 mM imidazole (225 mL). The elution wash was collected in 25 mL fractions and fractions that contained the desired protein as analysed by SDS-PAGE were combined and concentrated using Vivaspin (MWCO 5000) to 5 mL. The sample was then subjected to gel filtration using a 320 mL Superdex 200 column on a FPLC with PBS buffer. Flow rate of PBS was 2.5 mL/min. Fractions containing clean desired protein as analysed by SDS-PAGE were combined and concentrated using Vivaspin and the concentration of the protein solution was determined using BCA assay. Protein samples were then stored at –20 °C.

Yield of IgG1 Fc: 18.5 mg/L

2.8 – Protein Sequences

2.8.1 – EndoS

Nucleotide Sequence –

```

ATGGACAAAC ATCTGCTGGT TAAACGCACG CTGGGCTGTG TGTGTGCTGC TACGCTGATG
GGCGCTGCTC TGGCGACGCA CCACGATAGC CTGAACACCG TGAAAGCCGA AGAAAAACG
G TTCAGGTCC AAAAAGGTCT GCCGTCAATC GATTTCGCTGC ATTATCTGAG CGAAAACTCT
AAAAAAGAAT TCAAAGAAGA ACTGAGCAAA GCGGGCCAGG AATCTCAAAA AGTTAAAGAA
ATCCTGGCAA AAGCTCAGCA AGCCGATAAA CAGGCACAAG AACTGGCTAA AATGAAAATT
CCGGAAAAAA TCCCGATGAA ACCGCTGCAT GGTCCGCTGT ACGGCGGTTA TTTCCGTACC
TGGCACGATA AAACGTCAGA CCCGACCGAA AAAGACAAAG TCAACTCGAT GGGCGAACTG
CCGAAAGAAG TGGATCTGGC TTTTATTTTC CATGATTGGA CCAAAGACTA CTCTCTGTTT
TGGAAGAAG TGGCAACGAA ACACGTTCCG AAAGTGAACA AACAGGGTAC GCGTGTCTAT
CGTACCATTG CGTGGCGCTT CCTGGCTGGC GGTGATAATT CAGGCATCGC GGAAGACACC
TCGAAATATC CGAACACGCC GGAAGGTAAT AAAGCGCTGG CCAAAGCAAT CGTCGATGAA
TACGTGTACA AATACAATCT GGACGGCCTG GATGTGGACG TTGAACATGA TTCAATTCCG
AAAGTGGATA AAAAAAGAAG CACCGCCGGC GTGGAACGTT CGATCCAGGT TTTTGAAGAA
ATTGGTAAAC TGATCGGCCC GAAAGGTGTT GATAAAAGCC GTCTGTTCAT CATGGATTCT
ACCTATATGG CCGACAAAAA TCCGCTGATT GAACGCGGTG CACCGTACAT CAACCTGCTG
CTGGTCCAGG TGTATGGCAG CCAAGGTGAA AAAGGCGGTT GGAACCGGT GTCTAACCGT
CCGGAAAAAA CCATGGAAGA ACGCTGGCAG GGCTACTCAA AATATATTCG TCCGGAACAA
TACATGATCG GCTTTTCGTT CTATGAAGAA AACGCGCAGG AAGGTAATCT GTGGTACGAT
ATTAATAGTC GCAAAGATGA AGACAAAGCC AACGGCATTG ATACCGATAT CACGGGTACC
CGTGCGGAAC GCTATGCCCC TTGGCAGCCG AAAACCGGCG GTGTAAAGG CGGTATTTTT
AGCTACGCGA TCGATCGTGA CGGTGTCGCC CATCAGCCGA AAAAATACGC AAAACAAAAA
GAATTCAAAG ATGCTACCGA CAACATCTTC CACAGCGATT ACAGTGTCTC CAAAGCGCTG
AAAACCGTGA TGCTGAAAGA TAAATCTTAC GATCTGATCG ACGAAAAAGA TTTTCCGGAC
AAAGCGCTGC GCGAAGCCGT TATGGCACAG GTCGGCACCC GCAAAGGTGA CCTGGAACGT
TTTAATGGCA CGCTGCGCCT GGATAACCCG GCCATTGAGA GCCTGGAAGG TCTGAATAAA
TTCAAAAAAC TGGCACAAC TGGACCTGATT GGCCTGAGCC GTATCACCAA ACTGGATCGC
TCTGTGCTCG CGGCCAACAT GAAACCGGGT AAAGACACGC TGGAAACCGT TCTGGAAACC
TACAAAAAAG ATAACAAAGA AGAACCGGCA ACGATCCCGC CGGTGTCTCT GAAAGTTTCC
GGCCTGACCG GTCTGAAAGA ACTGGATCTG AGCGGCTTTG ACCGTGAAAC GCTGGCAGGT
CTGGATGCGG CCACGCTGAC CAGTCTGGAA AAAGTTGATA TTTCCGGCAA TAACTGGAC
CTGGCGCCGG GTACCGAAAA CCGCCAGATT TTTGATACGA TGCTGAGTAC CATCTCCAAC
CATGTTGGCA GCAATGAACA GACCGTCAAA TTCGACAAAC AAAAACCGAC GGGCCACTAC
CCGATACGTT ATGGTAAAA CAGCCTGCGT CTGCCGGTCG CCAACGAAAA AGTGGATCTG
CAGTCTCAAC TGCTGTTTGG CACGGTTACC AATCAGGGTA CCCTGATTAA CAGCGAAGCA
GATTACAAAG CTTACCAAAA CCATAAAATC GCGGGTCGCT CATTTGTGGA TTCGAACTAC
CACTACAACA ACTTCAAAGT TAGTTACGAA AACTACACCG TTAAAGTCAC GGATTCCACC
CTGGGCACCA CGACCGATAA AACGCTGGCC ACCGACAAAG AAGAAACCTA CAAAGTCGAT
TTCTTTAGCC CCGCAGACAA AACGAAAGCG GTGCATACCG CCAAAGTGAT TGTGGCGGAT
GAAAAAACCA TGATGGTGAA CCTGGCTGAA GGTGCGACGG TTATCGGCGG TTCCGCAGAC
CCGGTTAACG CTCGCAAAGT CTTTGATGGC CAGCTGGGTA GTGAAACCGA TAATATTTCC
CTGGGTGGG ACTCAAAACA GTCGATTATC TTCAAACCTGA AAGAAGACGG CCTGATCAAA
CACTGGCGTT TCTTTAACGA TAGTGCCCGC AATCCGGAAA CGACCAACAA ACCGATTGAG
GAAGCATCCC TGCAAATCTT CAACATCAAA GATTACAACC TGGACAATCT GCTGGAAAAC
CCGAATAAAT TCGATGACGA AAAATACTGG ATCACGGTGG ATACCTATAG CGCGCAGGGC
GAACGTGATA CGGCGTTTAG TAACACCTTG AACAATATTA CGTCCAAATA CTGGCGTGTG
GTTTTCTGTA CCAAAGGTGA CCGCTATAGC TCTCCGGTCG TGCCGGAACG GCAGATTCTG
GGCTATCCGC TGCCGAATGC TGATACGATC ATGAAAAACCG TGACGACCGC GAAAGAACTG
TCACAGCAAA AAGATAAATT CTCGCAGAAA ATGCTGGACG AACTGAAAAT TAAAGAAATG
GCTCTGAAAA CCAGCCTGAA TTCTAAAAAT TTCGATGTTA CGGCGATCAA TGCTAACGCT
GGTGTGCTGA AAGACTGTAT TGAAAAACGC CAACTGCTGA AAAAATGA

```

Amino Acid Sequence –

MDKHLVLVKRT	LGCVCAATLM	GAALATHHDS	LNTVKAEEKT	VQVQKGLPSI	DSLHYLSENS
KKEFKEELSK	AGQESQKVKE	ILAKAQQADK	QAQELAKMKI	PEKIPMKPLH	GPLYGGYFRT
WHDKTSDPTE	KDKVNSMGEL	PKEVDLAFIF	HDWTKDYSLF	WKELATKHVP	KLNKQGTRVI
RTIPWRFLAG	GDNSGIAEDT	SKYPNTPEGN	KALAKAIVDE	YVYKYNLDGL	DVDVEHDSIP
KVDKKEDTAG	VERSIQVFEE	IGKLIGPKGV	DKSRLFIMDS	TYMADKNPLI	ERGAPYINLL
LVQVYGSQGE	KGGWEPVSNR	PEKTMEERWQ	GYSKYIRPEQ	YMIGFSFYEE	NAQEGNLWYD
INSRKDEDKA	NGINTDITGT	RAERYARWQP	KTGGVKGGIF	SYAIDRDGVA	HQPKKYAKQK
EFKDATDNIF	HSDYSVSKAL	KTVMLKDKSY	DLIDEKDFPD	KALREAVMAQ	VGTRKGDLER
FNGTLRLDNP	AIQSLEGLNK	FKKLAQLDLI	GLSRITKLDR	SVLPANMKPG	KDTLETVLET
YKKDNKEEPA	TIPPVSLKVS	GLTGLKELDL	SGFDRETLAG	LDAATLTSLE	KVDISGNKLD
LAPGTENRQI	FDTMLSTISN	HVGSNEQTVK	FDKQKPTGHY	PDTYGKTSLR	LPVANEKVDL
QSQLLFGTVT	NQGTLINSEA	DYKAYQNHKI	AGRSFVDSNY	HYNNFKVSYE	NYTVKVTDS
LGTTTDKTLA	TDKEETYKVD	FFSPADKTKA	VHTAKVIVGD	EKTMVMNLAE	GATVIGGSAD
PVNARKVFDG	QLGSETDNIS	LGWDSKQSII	FKLKEDGLIK	HWRFFNDSAR	NPETTNKPIQ
EASLQIFNIK	DYNLDNLLN	PNKFDDEKYW	ITVDTYSAQG	ERATAFSNTL	NNITSKYWRV
VFDTKGDRYS	SPVVPQLQIL	GYPLPNADTI	MKTVTTAKEL	SQQKDKFSQK	MLDELKIKEM
ALETSLNSKI	FDVTAINANA	GVLKDCIEKR	QLLKK		

2.8.2 – IgG1 Fc

Nucleotide Sequence –

GCCACCATGG	GGATCCTTCC	CAGCCCTGGG	ATGCCTGCGC	TGCTCTNCNC	GTGAGCCTTC
TCTCCGTGCT	GCTGATGGGT	TGCGTAGCTG	AAACCGGTCA	CACATGCCCA	CCGTGCCCAG
CACCTGAACT	CCTGGGGGGA	CCGTCAGTCT	TCCTCTTCCC	CCCAAAACCC	AAGGACACCC
TCATGATCTC	CCGGACCCCT	GAGGTCACAT	GCGTGGTGGT	GGACGTGAGC	CACGAAGACC
CTGAGGTCAA	GTTCAACTGG	TACGTGGACG	GCGTGGAGGT	GCATAATGCC	AAGACAAAGC
CGCGGGAGGA	GCAGTACAAC	AGCACGTACC	GTGTGGTCAG	CGTCCTCACC	GTCTGCACCC
AGGACTGGCT	GAATGGCAAG	GAGTACAAGT	GCAAGGTCTC	CAACAAAGCC	CTCCCAGCCC
CCATCGAGAA	AACCATCTCC	AAAGCCAAAG	GGCAGCCCCG	AGAACCACAG	GTGTACACCC
TGCCCCCATC	CCGGGATGAG	CTGACCAAGA	ACCAGGTGAG	CCTGACCTGC	CTGGTCAAAG
GCTTCTATCC	CAGCGACATC	GCCGTGGAGT	GGGAGAGCAA	TGGGCAGCCG	GAGAACAAC
ACAAGGCCAC	GCCTCCCGTG	CTGGACTCCG	ACGGCTCCTT	CTTCTCTTAC	AGCAAGCTCA
CCGTGGACAA	GAGCAGGTGG	CAGCAGGGGA	ACGTCTTCTC	ATGCTCCGTG	ATGCATGAGG
CTCTGCACAA	CCACTACACG	CAGAAGAGCC	TCTCCCTGTC	TCCGGGTGGT	ACCAAGCACC
ACCATCACCA	TCATAATGA				

Amino Acid Sequence –

MGILPSPGMP	ALLSLVSLLS	VLLMGCVAET	GHTCPPCPAP	ELLGGPSVFL	FPPKPKDTLM
ISRTPEVTCV	VVDVSHEDPE	VKFNWYVDGV	EVHNAKTKPR	EEQYNSTYRV	VSVLTVLHQD
WLNKEYKCK	VSNKALPAPI	EKTISKAKGQ	PREPQVYTL	PSRDELTKNQ	VSLTCLVKGF
YPSDIAVEWE	SNGQPENNYK	ATPPVLDSDG	SFFLYSKLTV	DKSRWQQGNV	FSCSVMHEAL
HNHYTQKSL	LSPGGTKHHH	HHH			

2.8.3 – EndoS/EndoF₃ Amino Acid Sequence Alignment

EndoS (37-450) amino acid sequence was aligned with the amino acid sequence of EndoF₃ using T-Coffee,²⁶ with the alignment shown below. The key catalytic sequence is highlighted in yellow (residues 126 – 129 in EndoF₃, residues 233 – 236 in EndoS) and the predicted important glutamine and tyrosine are highlighted in green (Q211 and Y213 in EndoF₃ and Q303 and Y305 in EndoS). This alignment was then used to form a model of EndoS using the SWISS-MODEL²⁷⁻²⁹ homology-modelling server and the crystal structure of EndoF₃.

```

EndoF      MKKIFFFAQCSILLMLGSCSKMTEDMTPESVNKEASVKSATALAGSNGVC 50
EndoS      EEKTVQVQKG--LPSIDSLHYLSENSKKEF--KEELSKAGQESQKVKEIL 46
           :* . . * . * :.* :.* : . * ** *:. : :

EndoF      IAYYITDGRNPTFKLKDIPDKVDMVILFGLKYW----SLQDTTKLPGGTG 96
EndoS      AKAQQADKQAQELAKMKIPEKIPMKPLHGPLYGGYFRTWHDKTSDPTEKD 96
           :* : : .**:* : * *. * * : :*. * * ..

EndoF      MMGSFK-----SYKDLDL-T-QIRSLQSRGIKVLQN 124
EndoS      KVNSMGELPKQVDLAFIFHDWTKDYSLFWKELATKHVPKLNKQGTRVIRT 146
           :.* : :.* * * :. :*. * :*:.

EndoF      I-----DDVSWQSSKPGGFASAAAYGDAIKSIVIDKWKLKG 161
EndoS      IPWRFLAGGDNSGIAEDTSKYPTPEGN---KALAKAIVDEYVYKYNLDG 193
           * :.* ...* * * ..** . :*:*:*

EndoF      ISLDIEHSGAKPNPIPTFPGYAATGYNGWYSGSMAATPAFLNVISELTKY 211
EndoS      LDVDVEHDSI-----PK-----VDKKEDTAGVERSIQVFEEIGKL 228
           :.*:*.*. * . . *.. :*:. * *

EndoF      FGTTAPNNKQLQIASGIDVYAWNKMENFRNNFNFIQLQSYGANVSRTQ- 260
EndoS      IGPKGVDKSRLFIMDSTYMADKNPLIERGAPYINLLLVGVYGSQGEKGGW 278
           :*... :.*:* * .. : * :.* :* :* **:. :

EndoF      -----LMMNYATGTNKIPASKMVFGAYAE 284
EndoS      EPVSNRPEKTMEERWQGYSKYIRPEQYMIGFSFYEENAQEGNLWYDINSR 328
           *:. :. :. :. :. :.

EndoF      GGTNQANDV-----EVAKWTP-TQGAKGGMMIYTYN----- 314
EndoS      KDEDKANGINTDITGTRAERYARWQPKTGGVKGGIFSYAIDRDGVAHQPK 378
           . :.*.: . *:* * * *.***: : * :

EndoF      -----SNVSYANAVRDA----VKN 329
EndoS      KYAKQKEFKDATDNIFHSDYSVSKALKTVMLKDKSY 414
           *:* * :.*:. . .

```

2.9 – EndoS Mutant Catalysed Reactions

2.9.1 – Hydrolytic Activity of EndoS Mutants

To a solution of Herceptin (25 μ L, 2.0 mg/mL in PBS, 50 μ g) and PBS (20 μ L) was added wild type EndoS or EndoS mutant (5 μ L, 1 mg/mL or 0.1 mg/mL in PBS, 5 μ g or 0.5 μ g) to give either 1:10 or 1:100, EndoS:Herceptin ratio by mass. The reaction was then incubated at room temperature and samples (5 μ L) were taken over the time course and snap frozen. After all samples had been collected, samples were denatured (95 °C for 10 min, in reducing loading buffer, see 2.7.1.1) to ensure separation of the light and heavy chain and analysed by SDS-PAGE. The gel was then imaged using Gel Doc XR+ molecular imager and densitometry of the bands carried out using Image Lab software. Due to the fact that the product and starting material bands are very close in molecular weight on the gel, this only gives a semi-quantitative conversion.

2.9.2 – Glycosylation Activity of EndoS Mutants

To a solution of GlcNAc Herceptin (5 μ L, 7 mg/mL in PBS, 35 μ g, Herceptin previously treated with EndoS to cleave the glycan and purified by Protein A spin column), decasaccharide oxazoline (3.5 μ L, 10 mg/mL, 35 μ g), sodium phosphate buffer (3.5 μ L, 500 mM) and water (22 μ L) was added an EndoS mutant (0.7 μ L, 1 mg/mL in PBS 0.7 μ g) to give 1:50 EndoS:Herceptin ratio by mass. The reaction was then incubated at 30 °C and samples (1 μ L) were taken over the time course and snap frozen. After all samples had been collected, samples were denatured (95 °C for 10 min, in reducing loading buffer, see 2.7.1.1) to ensure separation of the light and heavy chain and analysed by SDS-PAGE. Densitometry was carried out as described for hydrolysis reactions.

2.10 – Crystallisation Trials of EndoS

All protein solutions were transferred into 20 mM Tris-HCl (pH 7.3) buffer by dialysis (15 h + 3 h) before crystallisation screens were prepared. Protein crystallisation screens were then set up, all incubated at 20 °C.

Typical initial screen with EndoS or EndoS mutant

Screen	Well Solution (nL)	Protein Solution (nL)	Thiazoline Ligand
PACT	150	150	–
	75	150	–
	150	150	+
Stura	150	150	–
	75	150	–
Index	150	150	–
	150	150	+
Newcastle	150	150	–
	150	150	+
MPD	150	150	–
	150	150	+

Initial protein concentrations typically at 20 mg/mL. The best precipitates seemed to occur when using the ligand and 25% PEG 3350 with variety of salts (0.2 M).

Typical initial screen with EndoS E235Q with IgG1 Fc

Screen	Well Solution (nL)	Protein Solution (nL)	IgG1 Fc
PACT	150	150	Monomer
	150	150	Dimer
JCSG+	150	150	Monomer
	150	150	Dimer
Salt Rx	150	150	Monomer
	150	150	Dimer
MPD	150	150	Monomer
	150	150	Dimmer

Initial protein concentrations typically at 20 mg/mL and all initial trial set up with 1:1 EndoS E235Q:IgG1 Fc. Optimisation was done by varying protein to solution ratios as well as the components and concentrations of the well solutions. The best results were obtained using 0.1 M sodium acetate/bromide, 18% PEG 3550.

2.11 – References

- 1 J. Yin, L. Li, N. Shaw, Y. Li, J. K. Song, W. Zhang, C. Xia, R. Zhang, A. Joachimiak, H. Zhang, L. Wang, Z. Liu and P. Wang, *Plos One*, 2009, **4**, e4658.
- 2 M. Umekawa, C. Li, T. Higashiyama, W. Huang, H. Ashida, K. Yamamoto and L. Wang, *J. Biol. Chem.*, 2010, **285**, 511-521.
- 3 V. Rao, P. Van Roey and T. Cue, *Protein Sci.*, 1999, **8**, 2338-2346.
- 4 C. A. Waddling, T. H. Plummer, A. L. Tarentino and P. Van Roey, *Biochemistry*, 2000, **39**, 7878-7885.
- 5 D. W. Abbott, M. S. Macauley, D. J. Vocadlo and A. B. Boraston, *J. Biol. Chem.*, 2009, **284**, 11676-11689.
- 6 Y. Chang, C. Liu, C. Guo, Y. Wang, J. Fang and W. Cheng, *J. Org. Chem.*, 2008, **73**, 7197-7203.
- 7 G. Zou, H. Ochiai, W. Huang, Q. Yang, C. Li and L. Wang, *J. Am. Chem. Soc.*, 2011, **133**, 18975-18991.
- 8 C. D. Heidecke, T. B. Parsons and A. J. Fairbanks, *Carbohydr. Res.*, 2009, **344**, 2433-2438.
- 9 Y. Tomabechi, M. A. Squire and A. J. Fairbanks, *Org. Biomol. Chem.*, 2014, **12**, 942-955.
- 10 T. Hamaguchi, T. Ito, Y. Inoue, T. Limpaseni, P. Pongsawasdi and K. Ito, *Glycobiology*, 2010, **20**, 420-432.
- 11 W. Huang, J. Li and L. Wang, *Chem. Bio. Chem.*, 2011, **12**, 932-941.
- 12 T. B. Parsons, M. K. Patel, A. B. Boraston, D. J. Vocadlo and A. J. Fairbanks, *Org. Biomol. Chem.*, 2010, **8**, 1861-1869.

- 13 S. Murakami, Y. Takaoka, H. Ashida, K. Yamamoto, H. Narimatsu and Y. Chiba, *Glycobiology*, 2013, **23**, 736-744.
- 14 W. Huang, J. Giddens, S. Fan, C. Toonstra and L. Wang, *J. Am. Chem. Soc.*, 2012, **134**, 12308-12318.
- 15 C. D. Heidecke, Z. Ling, N. C. Bruce, J. W. B. Moir, T. B. Parsons and A. J. Fairbanks, *Chem. Bio. Chem.*, 2008, **9**, 2045-2051.
- 16 I. Stals, S. Karkehabadi, S. Kim, S. Ward, M. Ward, A. Van Landschoot, B. Devreese and M. Sandgren, *Plos One*, 2012, **7**, e40854.
- 17 M. Collin and A. Olsen, *EMBO J.*, 2001, **20**, 3046-3055.
- 18 M. Collin and A. Olsen, *Infect. Immun*, 2001, **69**, 7187-7189.
- 19 M. Allhorn, A. Olsen and M. Collin, *BMC Microbiology*, 2008, **8**, 3.
- 20 M. Allhorn, A. I. Olin, F. Nimmerjahn and M. Collin, *Plos One*, 2008, **3**, e1413.
- 21 M. Allhorn, J. G. Briceno, L. Baudino, C. Lood, M. L. Olsson, S. Izui and M. Collin, *Blood*, 2010, **115**, 5080-5088.
- 22 J. J. Goodfellow, K. Baruah, K. Yamamoto, C. Bonomelli, B. Krishna, D. J. Harvey, M. Crispin, C. N. Scanlan and B. G. Davis, *J. Am. Chem. Soc.*, 2012, **134**, 8030-8033.
- 23 D. S. Waugh, *Protein Expres. Purif.*, 2011, **80**, 283-293.
- 24 A. Seko, M. Koketsu, M. Nishizono, Y. Enoki, H. R. Ibrahim, L. R. Juneja, M. Kim and T. Yamamoto, *Biochimica et Biophysica Acta*, 1997, **1335**, 23-32.
- 25 B. Sun, W. Bao, X. Tian, M. Li, H. Liu, J. Dong and W. Huang, *Carbohydr. Res.*, 2104, **396**, 62-69.
- 26 C. Notredame, D. G. Higgins and J. Heringa, *J. Mol. Biol.*, 2000, **302**, 205-217.

- 27 M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T. G. Cassarino, M. Bertoni, L. Bordoli and T. Schwede, *Nucleic Acids Res.*, 2014, **42**, W252-W258.
- 28 K. Arnold, L. Bordoli, J. Kopp and T. Schwede, *Bioinformatics*, 2006, **22**, 195-201.
- 29 F. Kiefer, K. Arnold, M. Kunzli, L. Bordoli and T. Schwede, *Nucleic Acids Res.*, 2009, **37**, D387-D392.
- 30 j. M. Macdonald, C. A. Tarling, E. J. Taylor, R. J. Dennis, D. S. Myers, S. Knapp, G. J. Davies and S. G. Withers, *Angew. Chem. Int. Ed.*, 2010, **49**, 2599-2602.
- 31 B. Li, K. Takegawa, T. Suzuki, K. Yamamoto and L. Wang, *Bioorg. Med. Chem.*, 2008, **16**, 4670-4675.
- 32 T. A. Bowden, K. Baruah, C. H. Coles, D. J. Harvey, X. Yu, B. Song, D. I. Stuart, A. R. Aricescu, C. N. Scanlan, Y. E. Jones and M. Crispin, *J. Am. Chem. Soc.*, 2012, **134**, 17554-17563.
- 33 L. J. Harris, S. B. Larson, K. W. Hasel and A McPherson, *Biochemistry*, 1997, **36**, 1581- 1597.
- 34 A. R. Aricescu, W. Lu and Y. E. Jones, *Acta Cryst.*, 2006, **D62**, 1243-1250.
- 35 B. Trastoy, J. V. Lomino, L. Wang and E. J. Sundberg, *Acta Crystallogr F*, 2013, **69**, 1405-1410.
- 36 B. Trastoy, J. V. Lomino, B. G. Pierce, L. G. Carter, S. Gunther, J. P. Giddens, G. A. Snyder, T. M. Weiss, Z. Weng, L. Wang and E. J. Sundberg, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 6714-6719.

Chapter 3 –

Investigations into the

Activity of PglB *In Vitro*

for the Synthesis of

Glycopeptides

Chapter 3 – Investigations into the Activity of PglB In Vitro for the Synthesis of Glycopeptides

3.1 – Introduction

PglB is an *N*-glycosyltransferase found in *Campylobacter jejuni* that catalyses the transfer of an oligosaccharide donor to the side chain nitrogen on the asparagine of a peptide acceptor.¹ As described in Chapter 1, glycosylation is a common and critical post-translational modification on eukaryotic proteins. However, protein glycosylation is not unique to eukaryotic organisms and does occur within the cells of archaea² and certain proteobacteria,^{3,4} of which *C. jejuni* is an example. The process for the formation of glycoproteins with *N*-linked glycans is the same within these organisms as compared to eukaryotic organisms, beginning with the formation of the oligosaccharide donor, transfer of the oligosaccharide to the asparagine of the protein, followed by further derivation steps.^{1, 4, 5} It should be noted though, that the oligosaccharide being transferred onto the protein has a different composition and structure to that transferred by eukaryotic systems, with one of the key differences being the oligosaccharide is attached to the 3-position on the reducing end sugar unit as opposed to the being linked at the 4-position.¹ For *N*-glycosylation to occur in these organisms, the enzymatic machinery must be present to catalyse this process. As stated above, PglB is the enzyme that catalyses the key step where the oligosaccharide is attached to the asparagine on the proteins in *C.jejuni* and hence having analogous activity to oligosaccharyltransferase (OST) in eukaryotic systems and Stt3p, the major subunit of OST.^{1, 4, 5} PglB orthologues have been found in other non-eukaryotic systems that have *N*-glycosylation.⁶⁻⁸ Like OST, the oligosaccharide donor substrate for PglB is a polyprenol pyrophosphate linked

oligosaccharide; for PglB, this lipid is a C55 lipid⁹ (Figure 3.1). Also like OST, PglB is a membrane protein.

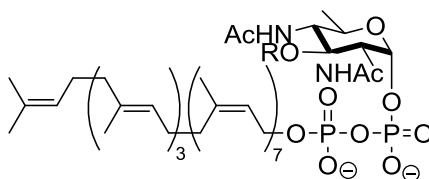


Figure 3.1: The natural PglB substrate, where R = oligosaccharide.

Although PglB has the same activity as OST, there are a number of critical differences between OST and PglB. Perhaps the most important of these is the fact that PglB requires no additional subunits for activity and hence can catalyse *N*-glycosylation alone.¹⁰ The consequences of this are that PglB is much easier to study than OST and easier to use as a glycosylation catalyse. Furthermore, although a membrane protein, it is possible to express PglB in *E.Coli*.^{1, 5, 9, 11}

Regarding the study of PglB, it has been shown that the amino acid recognition sequence for PglB is D/E-X-N-X-T/S,^{1, 12} where X is any amino acid apart from proline. Compared to the recognition sequence of OST (N-X-T/S), it can be seen as more stringent. This recognition sequence was first found using peptide studies,¹² but has been further confirmed and explained by the solving of the crystal structure of PglB, both on its own and in complex with a peptide containing the recognition sequence¹ (Figure 3.2). Further to the information obtained about the amino acid recognition sequence required for glycosylation by PglB, modelling studies were done with the crystal structure (Figure 3.3), predicting how the oligosaccharide donor substrate fits inside the active site of PglB, giving information on the mechanism of the enzyme and possible information on the structure and mechanism of OST.

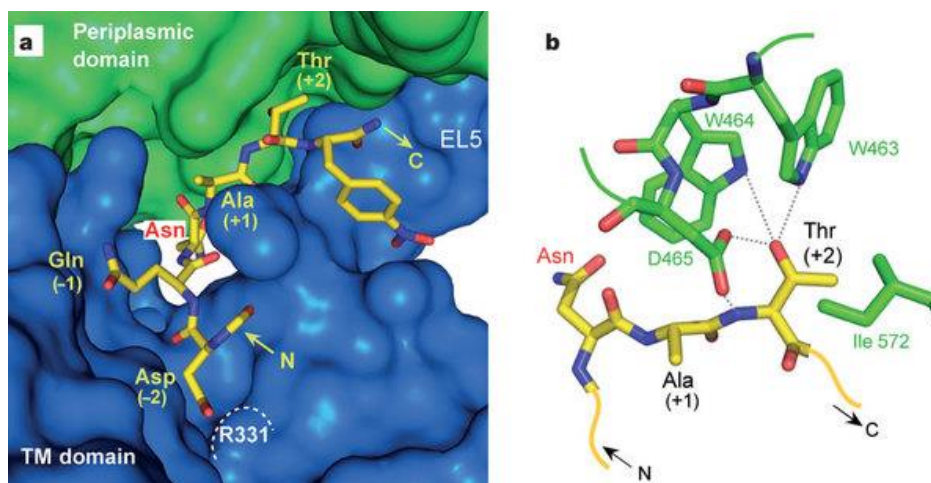


Figure 3.2: Crystal structure of PglB with peptide acceptor bound. This figure is taken from *Lizak et al.*¹

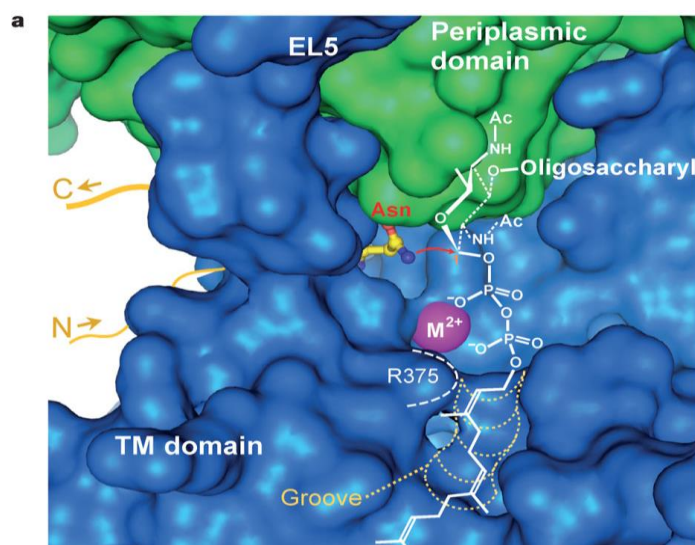


Figure 3.3: Crystal structure of PglB with the oligosaccharide donor modelled into the structure.

This figure is taken from *Lizak et al.*¹.

As well as investigations into the structure and mechanism of PglB, there have been a number of investigations into the use of PglB in glycosylation reactions.¹¹⁻¹⁵ Many of these investigations have been on peptides containing the recognition sequence of PglB and have been conducted *in vitro*. There have also been reports in the use of PglB *in vivo*, achieved by inserting the PglB gene cluster into *E.Coli*, along with the genes to

form the oligosaccharide donor and the gene to express the protein desired to be glycosylated.¹⁴

Although there are reports of the use of PglB for the glycosylation of peptides and proteins, the critical difficulty in this reaction is the use of the C55 polyprenol pyrophosphate-linked oligosaccharide substrate. This is an extremely laborious and challenging synthesis. Although there are ways to avoid this synthesis, as described above, the availability of the complex oligosaccharide substrate is still the limiting factor in the use of PglB to form homogenous glycopeptides and glycoproteins.

Previous research in the group has tried to solve this issue by investigating the effect of shortening the length of the lipid on the oligosaccharide substrate, testing if a substrate with shorter, easier to synthesise lipid would function with PglB. This research showed that shorter lipids on the glycan donor substrate are accepted by PglB, with C20 lipids being shown as the minimum length lipid required to see glycosylation on a fluorescent peptide.⁹ Furthermore, similar reactivity is seen when using a C30 lipid substrate compared to the natural C55 lipid substrate (Figure 3.3).⁹ For these reactions, C30-GlcNAc was used as the glycosyl donor to be attached to the fluorescent peptide in the PglB glycosylation reaction.

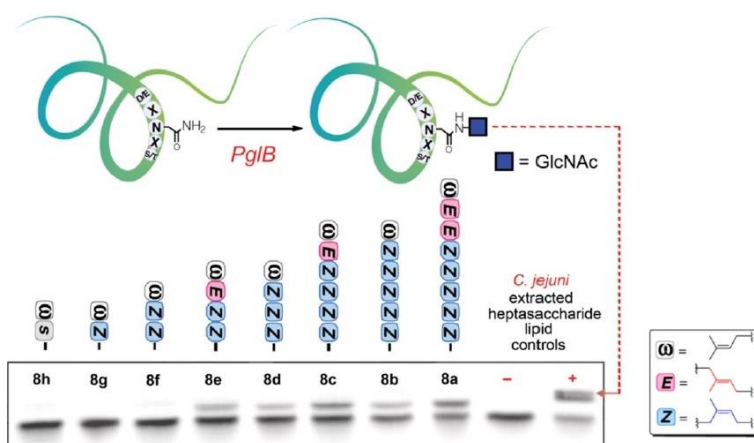


Figure 3.4: Demonstration that glycan donor substrates with shorter lipids than the natural substrate still function with PglB for the transfer of a GlcNAc unit to a peptide.⁹

From this point, it was desired to further investigate the effect of different lipid lengths. With C30 and C20 lipid lengths having been investigated and having fairly different reactivity, investigation of a C25 lipid length was attractive (Figure 3.5), seeing if this PglB substrate had reactivity between that of C30 and C20 lipid length. The advantage of the C25 lipid length, compared to the C30 lipid length is that it is significantly easier to synthesise, as will be shown.

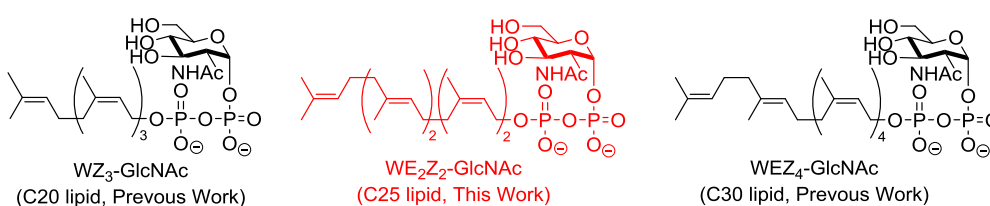


Figure 3.5: Propose C25 lipid PglB substrate to be synthesised, compared to the previously synthesised C20 and C30 lipid PglB substrates.

It was also desired to show that further glycosylation reactions were possible after the PglB reaction. With the addition of a GlcNAc to the peptide using PglB, it should be possible to conduct further enzymatic glycosylation reactions on this acceptor to form

more complex glycol structures. For this, the glycosylation reactivity of GalT and EndoA were investigated on the PglB peptide product.

In this chapter, the investigations of the C25 lipid substrate and further glycosylation reactions after the PglB reaction are described. In order to investigate both of these areas, it was necessary to synthesise the C25 lipid substrate (via synthesis of the lipid, followed by coupling to the GlcNAc unit) as well as express PglB and EndoA for the further glycosylation reaction.

3.2 – PglB Substrate Synthesis

3.2.1 – Lipid Synthesis

As described above, the PglB glycan donor substrate is a polyprenol pyrophosphate-linked glycan. In this case, the glycan is a GlcNAc unit. The synthesis of the substrate can be broken down into three steps; synthesis of the lipid, synthesis of the sugar and the coupling of the two. The GlcNAc phosphate sugar had already been previously synthesised,⁹ meaning only the lipid had to be formed before the coupling reaction.

The lipid attached to the PglB substrate is a polyprenol lipid, formed of C5 prenol units, either in a *cis* (Z) or a *trans* (E) conformation. The C55 lipid on the natural substrate has a WE₃Z₅ confirmation, whereas the C20 and C30 lipid used in the previous study had a WZ₃ and WEZ₄ confirmation, respectively. The C25 lipid target was WE₂Z₂, chosen due to the fact that this can be formed from one coupling between a WE₂ fragment and a Z₂ fragment. This will be described in greater detail below. The WEZ₄ lipid was also synthesised, although not used here; this lipid was utilised in Chapter 4 and the synthesis is described here to contrast the synthesis of this more complex lipid with the relatively easier WE₂Z₂ lipid.

The synthesis of these lipids has been described in a number of different studies.^{9, 16, 17}

The strategy used to form the lipid here was the formation of building blocks containing either a chloride group or a phenylsulfone group. These building blocks were coupled together by the deprotonation of the sulfone and subsequent nucleophilic attack on the chloride building block, forming the desired carbon-carbon bond. For the formation of the WE₂Z₂ lipid, WE₂-SO₂Ph and Z₂-Cl building blocks were required (one coupling

reaction), whereas for the WEZ₄ lipid WE-SO₂Ph, Z₂-Cl and Z₂-SO₂Ph building blocks were needed (two coupling reactions). These building blocks were formed from nerol, geraniol and *trans, trans* farnesol (Figure 3.6).

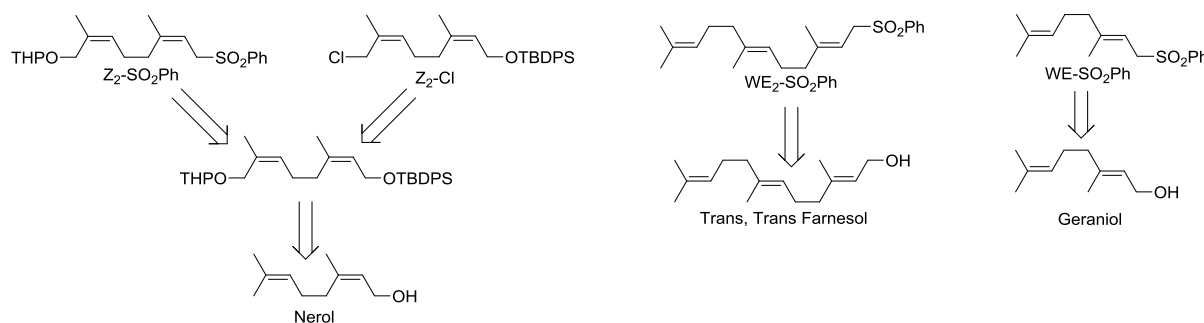
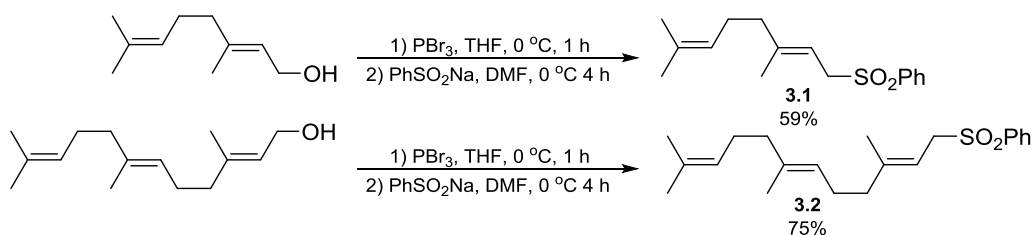


Figure 3.6: The building blocks to give the desired lipids.^{9, 17}

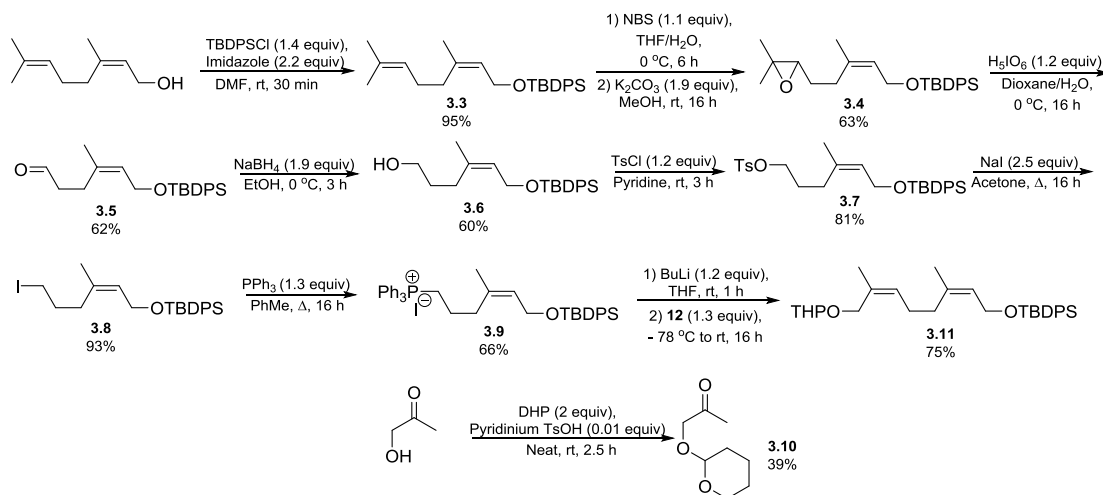
The WE-SO₂Ph (**3.1**) and the WE₂-SO₂Ph (**3.2**) building blocks were readily formed in two steps starting from geraniol and *trans, trans* farnesol respectively (Scheme 3.1).



Scheme 3.1: Formation of WE and WE₂ building blocks.^{9, 16}

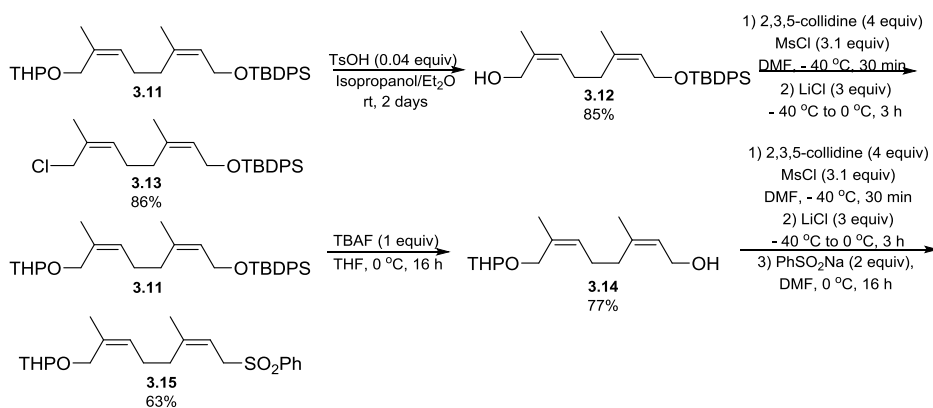
The Z₂-Cl and Z₂-SO₂Ph building blocks are both formed from a Z₂ building block containing complementary protecting groups for the hydroxyl groups at either end of the molecule. The formation of this Z₂ building block was more challenging and requires a number of steps. Following a reported synthesis and starting from nerol, the hydroxyl group was first protected with a silyl group before epoxide formation (**3.4**, Scheme 3.2).^{9, 17} Oxidative cleavage then resulted in the formation of the aldehyde (**3.5**) which was then reduced to give the alcohol (**3.6**). This alcohol is transformed into the

iodide (**3.8**) via the tosylate, which was then used to give a phosphonium ylide (**3.9**). This was then used in a Wittig reaction to the desired Z₂ building block (**3.11**) in multigram scale.



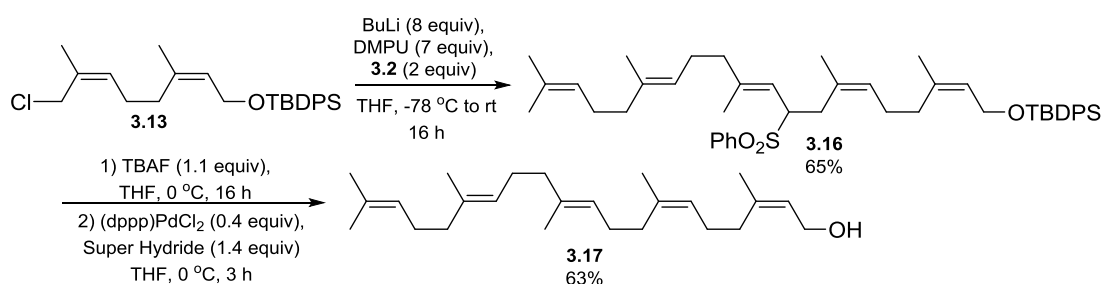
Scheme 3.2: Synthesis of Z₂ building block (**3.11**).^{9,17}

The Z₂ building block material was split into two. Half was converted to the Z₂-Cl (**3.13**) via deprotection of the THP group to give the alcohol followed by conversion to the chloride (Scheme 3.3). The other half was converted to the Z₂-SO₂Ph (**3.15**) by deprotection of the silyl protecting group to give the alcohol, followed by conversion to the sulfone.



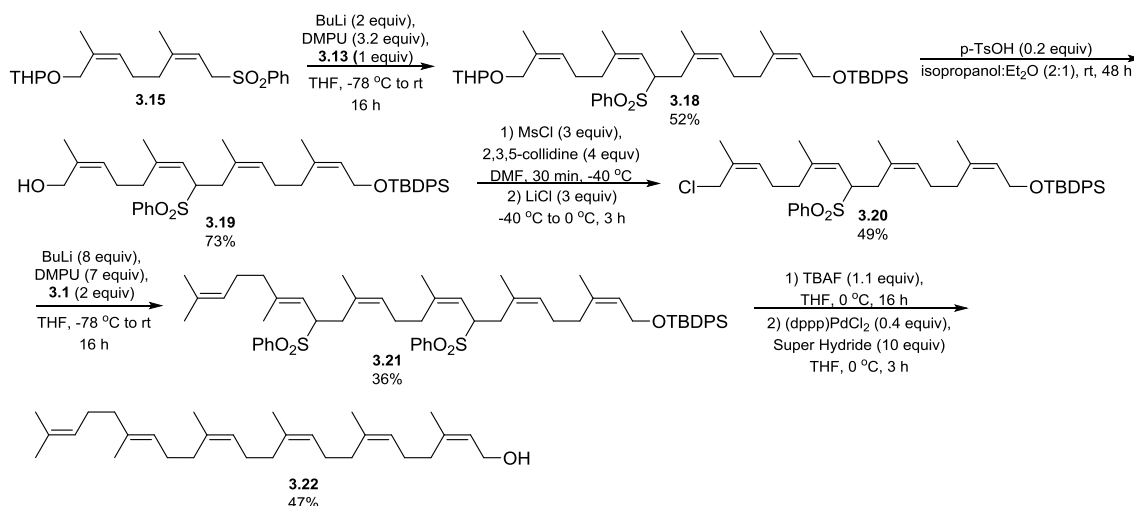
Scheme 3.3: Formation of Z₂ chloro building block (**3.13**) and Z₂ phenylsulfonate building block (**3.15**).^{9,17}

Then, the synthesis of the WE₂Z₂ lipid was completed by coupling the WE₂-SO₂Ph building block (**3.2**) with the Z₂-Cl building block (**3.13**), followed by removal of the silyl protecting group and the sulfone group to give the desired lipid (**3.17**, Scheme 3.4). Preliminary difficulties toward this coupling were overcome by increasing the amount of base (butyl lithium) used compared to previous reports, resulting in the formation of the product in an acceptable yield.



Scheme 3.4: Formation of WE₂Z₂ lipid.^{9, 17}

For the WEZ₄ lipid, first the Z₂-Cl (**3.13**) and the Z₂-SO₂Ph (**3.15**) molecules had to be coupled together to give the Z₄ unit (**3.18**, Scheme 3.5). Deprotection of the THP group gave the alcohol, which was then converted into the chloride to give the Z₄-Cl unit (**3.20**). This could then be coupled to the WE-SO₂Ph (**3.1**) building block to give the desired lipid structure (**3.21**). Again, this coupling was challenging originally, resulting in a poor yield of product, but the use of the base in great excess increased the yield 4-fold. Removal of the silyl protecting group and the sulfone groups resulted in the formation of the desired WEZ₄ lipid (**3.22**). This lipid is used in Chapter 4.



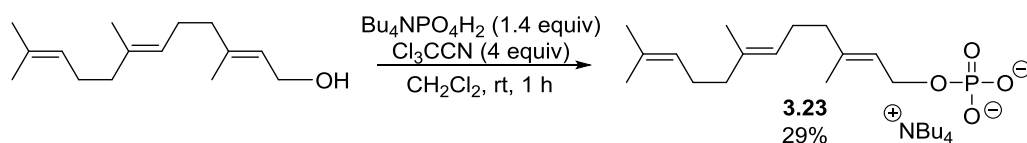
Scheme 3.5: Formation of the WEZ₄ lipid.^{9, 17}

3.2.2 – Coupling of the Lipid to the Sugar, Forming the PglB Substrate

With the WE₂Z₂ lipid in hand, and the GlcNAc phosphate previously synthesised, the coupling between lipid and the sugar could be investigated. This coupling reaction has previously been seen to be a low-yielding reaction, so it was attempted to optimise this reaction. Before this coupling was attempted with the synthesised WE₂Z₂, it was first studied with *trans, trans* farnesol (WE₂). The major advantage of studying WE₂ first was its commercial availability, so this could be used to find conditions to couple the sugar and lipid before moving to the more valuable WE₂Z₂ lipid. The fact that this lipid is commercially available also made it interesting to investigate with PglB; although the minimum length of lipid to see activity with before was C20, it was thought that under optimized conditions, this C15 lipid may function as a substrate and hence avoid the lipid synthesis altogether. This would make the PglB glycosylation system more attractive as a method of forming glycosylated peptides and proteins.

The coupling reaction occurs in three steps; the phosphorylation of the lipid, activation of the phosphate with carbonyldiimidazole (CDI) and the coupling of the phosphate

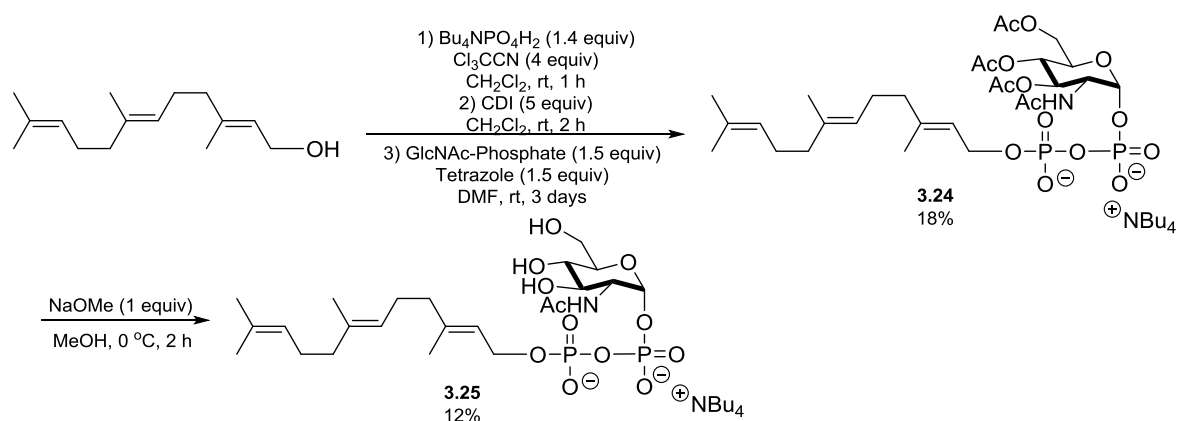
sugar. A final deprotection of the acetyl groups from the sugar results in the formation of the PglB substrate.⁹ Previously, purification was carried out only after the coupling and deprotection reactions, and it was envisaged that purification after the initial phosphorylation reaction and activation by CDI may result in greater yields. It was thought that one of the reasons for the low yield may be double phosphorylation of the lipid. The phosphorylation reaction was carried out as before with four equivalents of Bu₄NPO₄H₂, the phosphating reagent, followed by purification by silica column with 1:2:4, water:isopropanol:ethyl acetate. However, the purification was challenging, with difficulties separating the remaining Bu₄NPO₄H₂ from the product, along with product loss on the silica. In order to combat this, the phosphorylation reaction was carried out with 1.4 equivalents of Bu₄PO₄H₂. This allowed pure phosphorylated *trans, trans* farnesol to be obtained (**3.23**, Scheme 3.6).



Scheme 3.6: Phosphorylation of *trans, trans* farnesol.⁹

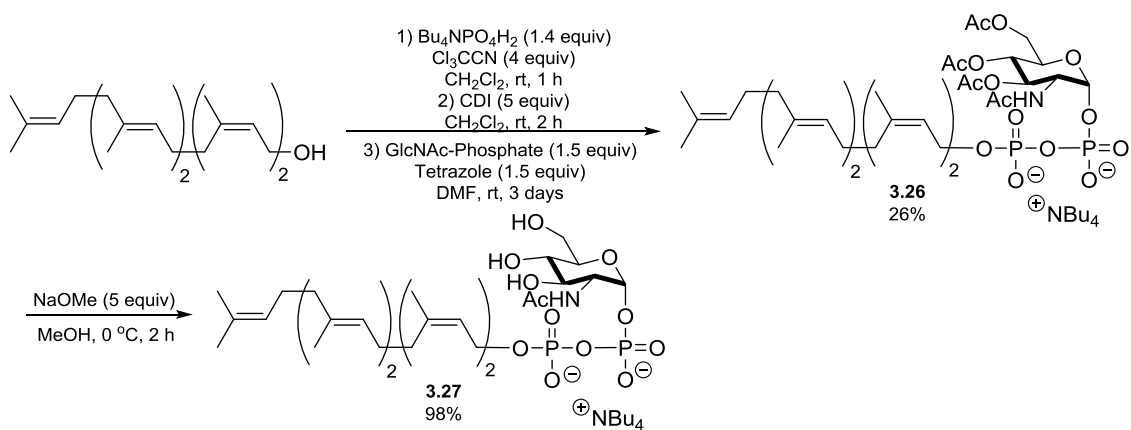
With phosphorylated *trans, trans* farnesol in hand, the activation with CDI was attempted. Again, purification with silica gel was difficult, but this time lowering the equivalents did not result in an easier purification. Co-elution of the product and the imidazole by-product and loss of product due to retention on the silica were the two issues. From the difficulties had with the purification of these phospholipids, it was speculated that the low yields seen before were not due to the reaction being poor, but instead due to the purification method. Therefore, the phosphorylation reaction was carried out again on WE₂ followed by CDI activation, the addition of the phosphorylated GlcNAc, and the coupling reaction with no purification between

reactions. Mass spectrometry was used to follow the reaction and confirm that the desired products had formed. ^{31}P NMR was also done after the initial phosphorylation reaction to ensure double phosphorylation had not occurred. As it was known that the product is very polar and hence prone to stick to the silica, instead of using silica gel, purification was attempted using reverse phase C-18 resin following previous reports of this method to purify this type of compound.¹⁷ This resulted in a greater yield than previously reported (Scheme 3.7). Deprotection of the acetyl groups on the sugar, again purified using reverse phase, resulted in formation of the desired PglB substrate (**3.25**).



Scheme 3.7: Coupling of GlcNAc phosphate to WE_2 lipid and subsequent deprotection.^{9, 17}

With conditions identified for the coupling of the lipid and sugar and purification of the product, these conditions were applied to the WE_2Z_2 lipid. This resulted in the formation of the PglB GlcNAc substrate conjugated to the WE_2Z_2 lipid (**3.27**, Scheme 3.8).



Scheme 3.8: Coupling of GlcNAc phosphate to WE_2Z_2 lipid and subsequent deprotection.^{9, 17}

3.3 – PglB Expression and Initial Testing

3.3.1 – PglB Expression

With the PglB glycan donor substrates now in hand, the next component required in order to investigate the PglB glycosylation reaction was PglB itself. The PglB gene was inserted into a pBAD-myc-HisA vector previously, resulting in a plasmid that would permit the expression of PglB with a His-tag attached, allowing purification by metal affinity chromatography. PglB was expressed in *E.Coli* using standard methods but with a few modifications, due to PglB being a membrane protein.^{1, 9, 11} Briefly, the plasmid was transformed into competent cells and expression cultures were set up using Terrific Broth. The expression cultures were incubated to OD 3, after which they were induced by the addition of arabinose. The resulting cell pellet was resuspended in buffer, lysed by sonication and centrifuged as per standard procedure. Once the inclusion body had been removed, the lysate was centrifuged a second time using an ultracentrifuge to obtain the membrane fraction. This was resuspended in dodecyl- β -maltopyranoside-containing buffer. Purification by NiNTA resin was carried out and the resulting fractions were analysed by SDS-PAGE and anti-His western blot (Figure 3.7).

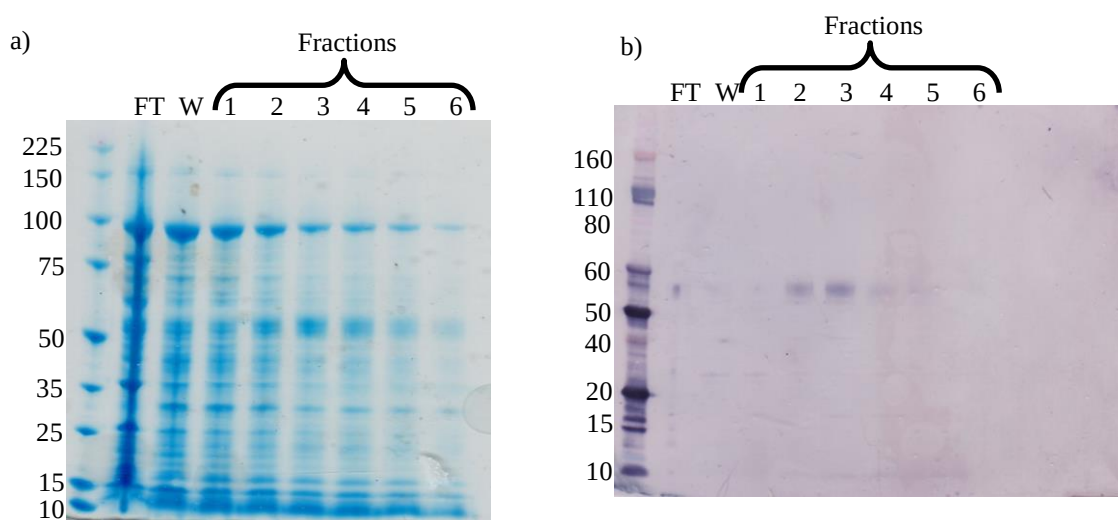
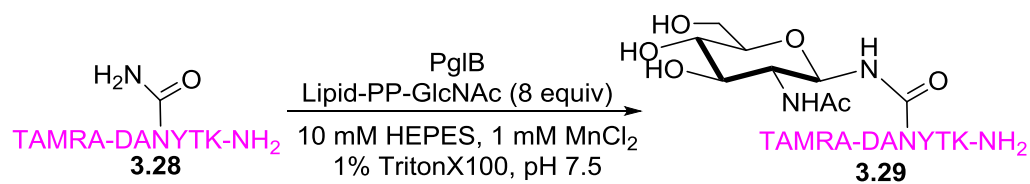


Figure 3.7: Purification of PglB expressed on a large scale, a) SDS-PAGE, b) Western Blot, FT = Flow through, W = Wash.

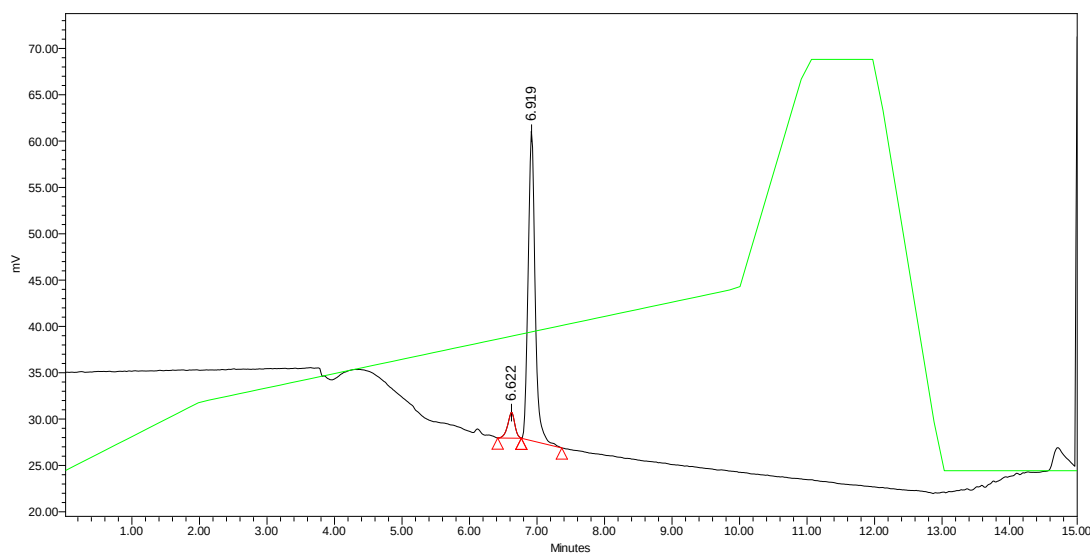
Clear bands can be seen on the western blot, showing the formation of PglB. Notably, PglB does not appear at the expected molecular weight on an SDS-PAGE/western blot due to it being a membrane protein. PglB-containing fractions were combined and concentrated.

3.3.2 – Initial Testing of PglB

Although the western blot showed a strong band for a protein with a His-tag attached, it was necessary to confirm that the protein expressed was PglB and that the protein was active. It was also necessary to find the concentration of the protein solution. In order to achieve this, the activity of the newly expressed PglB was tested using the fluorescent peptide (TAMRA-peptide, **3.28**) system used before in our group. This activity was compared to a PglB preparation of known concentration.⁹ The fluorescent peptide had the sequence DANYTK (thus containing the PglB recognition sequence) and a TAMRA fluorescent tag, allowing the reaction to be monitored by HPLC with a fluorescent detector, with absorption at 542 nm and emission at 568 nm. The reaction was carried out in a buffer containing MnCl₂, the metal co-factor that binds in the active site of PglB, and a detergent (Triton X-100) to ensure PglB remained soluble. The previously synthesised WEZ₄-GlcNAc PlgB substrate was used as the glycosyl donor.⁹

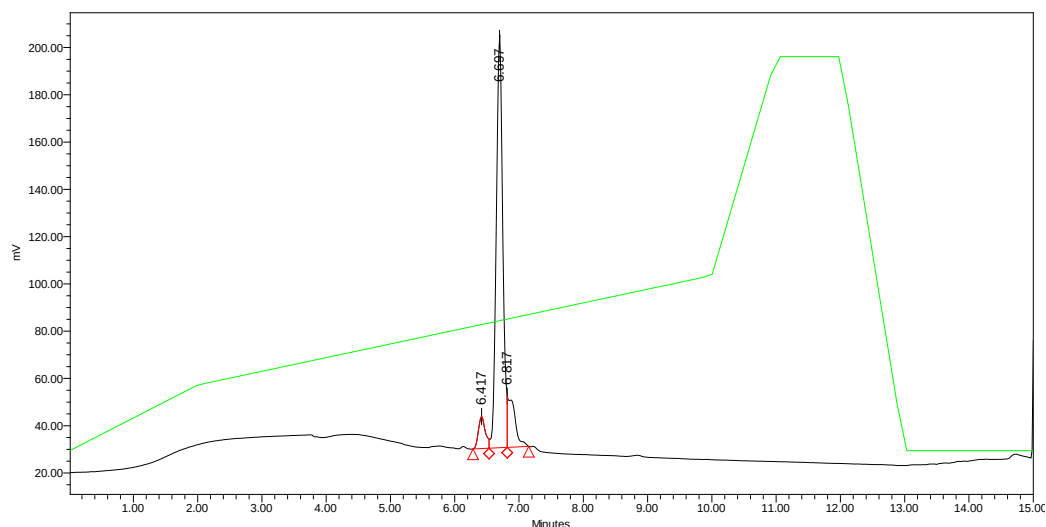


a) Starting Material



	Name	Retention Time	Area	% Area
1	Impurity	6.622	20401	7.84
2	3.28	6.919	239937	92.16

b) Control Reaction (0.088 pmol PglB)



	Name	Retention Time	Area	% Area
1	Impurity	6.417	100546	6.70
2	3.29	6.697	1222429	81.44
3	3.28	6.817	178012	11.86

Figure 3.8: HPLC trace of PglB reaction with TAMRA-peptide (3.28) and WEZ₄-GlcNAc substrates, a) TAMRA-Peptide, b) Control reaction (0.088 pmol PglB). Green line on each trace is the percentage amount of acetonitrile.

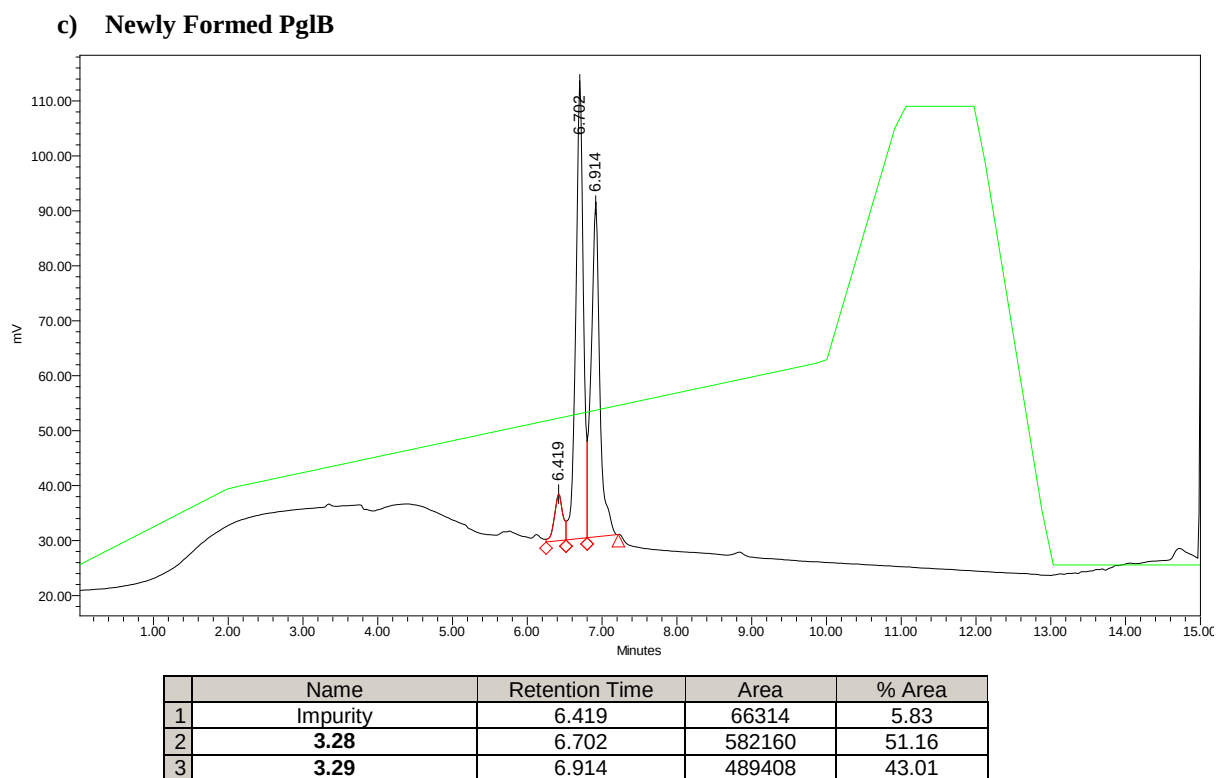
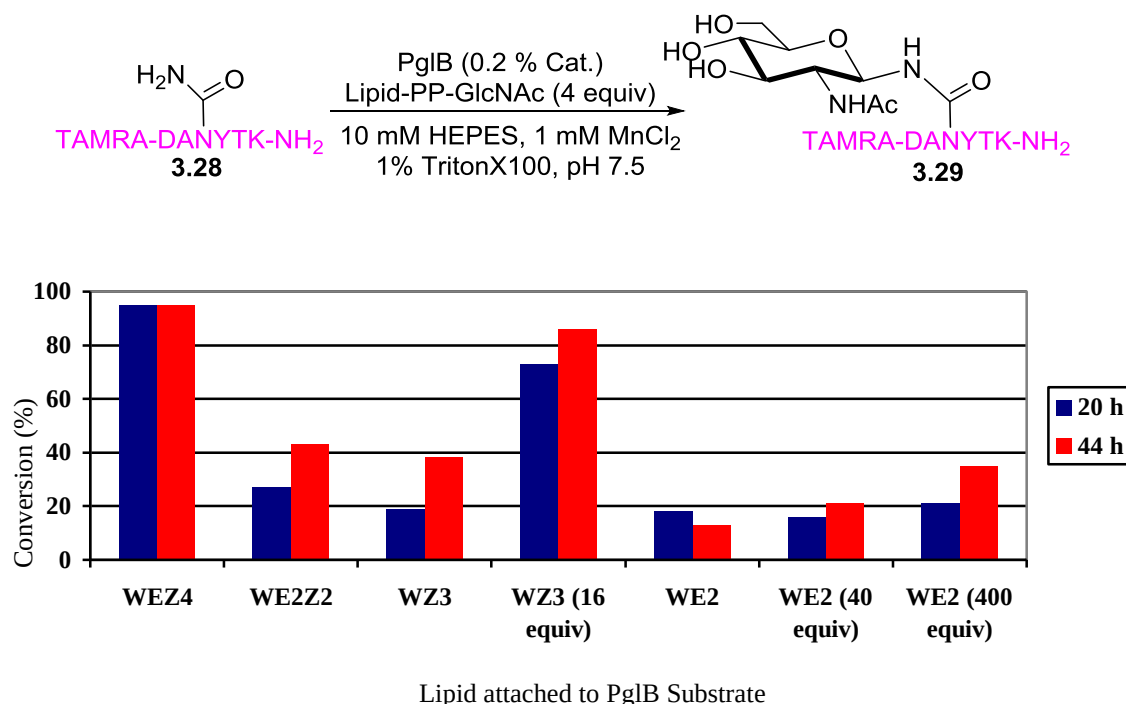


Figure 3.8 cont'd: HPLC trace of PglB reaction with TAMRA-peptide (3.28) and WEZ₄-GlcNAc substrates, c) Reaction with newly formed PglB. Green line on each trace is the percentage amount of acetonitrile.

The HPLC traces (Figure 3.8), showed that the newly formed PglB was active in forming the desired product (3.29). Comparing the amount of product formed of the newly formed PglB with that of the control reaction, approximately 0.055 pmol of PglB was added to the reaction, which indicated that the concentration of the PglB solution was approximately 0.24 μ M. Although the amount of protein was low, this was in part expected due to PglB being a membrane protein and previous reports expressing this protein on a large scale and expressed at a very high OD.¹ Although only a low amount was obtained, there was enough PglB to continue.

3.4 – Testing the Newly-Formed PglB Substrates

With the desired PglB glycan donor substrates formed and active PglB expressed, it was now possible to test what effect the different lipid lengths had on the activity of the substrates. These were compared to the WEZ₄-GlcNAc (C-30) and WZ₃-GlcNAc (C-20) substrates (previously formed⁹). The reactions were set up as described above using the TAMRA-peptide (**3.28**), with the same amount of PglB in each reaction. As well as testing the activity of the two newly-formed substrates (WE₂-GlcNAc (C-15, **3.25**) and WE₂Z₂-GlcNAc (C-25, **3.27**)), it was also desired to investigate whether it was possible to obtain higher conversions with the WZ₃-GlcNAc substrate. Samples were taken after 20 h and 44 h and analysed by fluorescent HPLC and the conversion of each reaction found via integration of the signals. (Graph 3.1).



Graph 3.1: Conversion to the TAMRA-GlcNAc-Peptide (3.28**) product when using different lipids on the substrate. Note, 4 equiv of the lipid-GlcNAc substrate used to TAMRA-Peptide unless indicated otherwise.**

It was expected that the longer the lipid chain, the better the substrate and hence greater conversion to the product. This had been observed from previous experiments.⁹ The order of activity was expected to range from WE₂-GlcNAc with the lowest activity, followed by WZ₃-GlcNAc, then WE₂Z₂-GlcNAc, with WEZ₄-GlcNAc having the greatest activity. This was confirmed by the experimental data. What was slightly unexpected was the fact that the activity of the WZ₃-GlcNAc and the WE₂Z₂-GlcNAc substrate were very similar, with the extra lipid length having very little effect on the activity of the substrate. One possible explanation for this would be the different confirmations of the lipids. The WE₂Z₂-GlcNAc substrate contains two *trans* prenyl units, whereas the lipid on the WZ₃-GlcNAc substrate contains only *cis* prenyl units. The different confirmations of the lipid will quite likely have an effect on the binding of the substrate in PglB. Otherwise, it may be that WEZ₄-GlcNAc has increased activity due to binding interactions unique to its lipid length or confirmation.

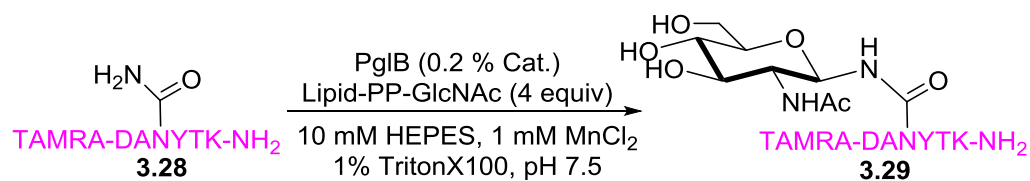
More surprising was that WE₂-GlcNAc functioned as a PglB substrate. It was expected that the WE₂-GlcNAc substrate would not function at all with PglB, due to the short lipid chain. Yet, it was observed with 400 equivalents of the WE₂ lipid substrate that significant conversion occurs (greater than 30%). This was very interesting, as the WE₂ lipid is commercially available and hence it is possible when using increased amounts of PglB that even better conversion could occur. It could also be seen that it was possible to obtain similar conversions to the WEZ₄-GlcNAc substrate using the WZ₃-GlcNAc substrate, using increased equivalence and time. This confirmed that the synthesis of long lipids may not be required for the use of PglB to obtain high yields of glycopeptides and glycoproteins.

3.5 – Further Enzymatic Glycosylation Reactions After PglB

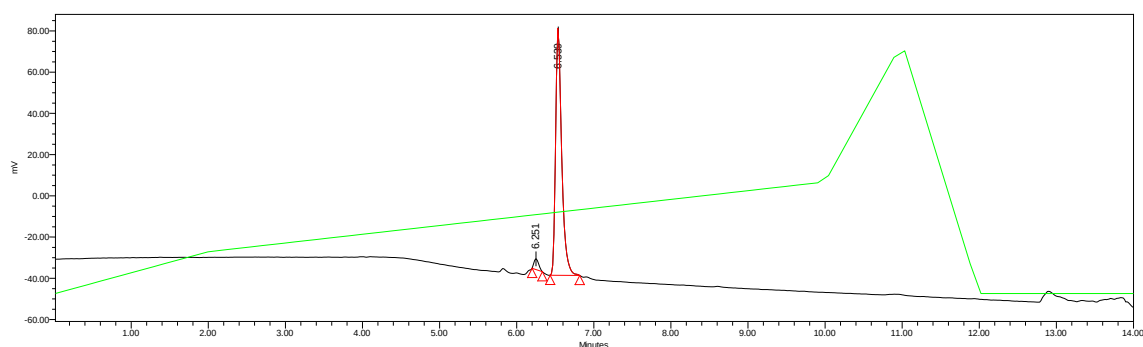
3.5.1 - Formation of the Peptide-GlcNAc Acceptor with PglB and Further Glycosylation with GalT

It has been shown here and elsewhere that it is possible to use PglB as an *N*-glycosyltransferase *in vitro*. The above data confirmed that PglB can add a GlcNAc unit to a peptide. As described in Chapter 1, a single sugar unit is rarely attached to a glycoprotein and the endogenous glycans attached to the protein are normally oligosaccharides. It was therefore desired to demonstrate that after the initial PglB glycosylation it was possible to functionalise the GlcNAc product to form more complex glycan structures using enzymatic reactions. The glycosylation reactions to be studied were that of Galactose Transferase (GalT) and EndoA.

First, the fluorescent TAMRA peptide with the GlcNAc (**3.29**) attached had to be formed using PglB. This was successfully done using the WEZ₄-GlcNAc with a 95% conversion seen after 20 h by HPLC and with the desired product formation being confirmed by MALDI-MS (Figure 3.9 and 3.10).

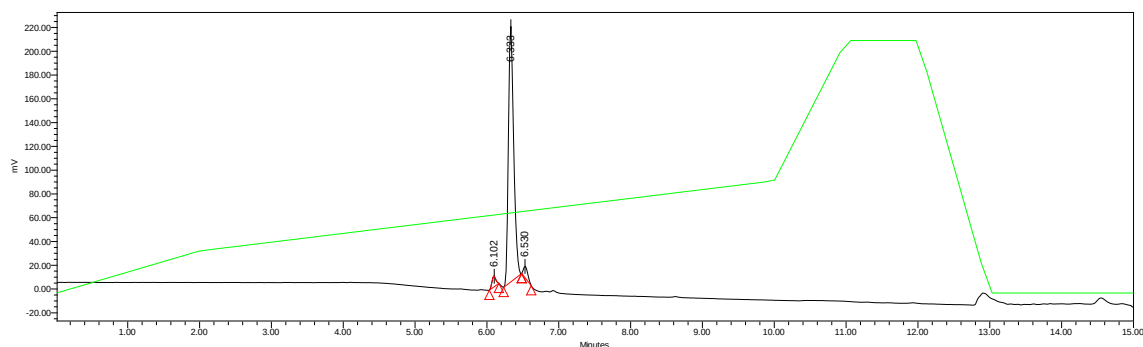


a) Starting Material



	Name	Retention Time	Area	% Area
1	Impurity	6.251	22147	3.26
2	3.28	6.539	656376	96.74

b) PglB Reaction



	Name	Retention Time	Area	% Area
1	Impurity	6.102	34578	2.95
2	3.29	6.333	1094883	93.47
3	3.28	6.530	41861	3.57

Figure 3.9: HPLC trace of a) TAMRA-Peptide strating material and b) glycosylation reaction with PglB.

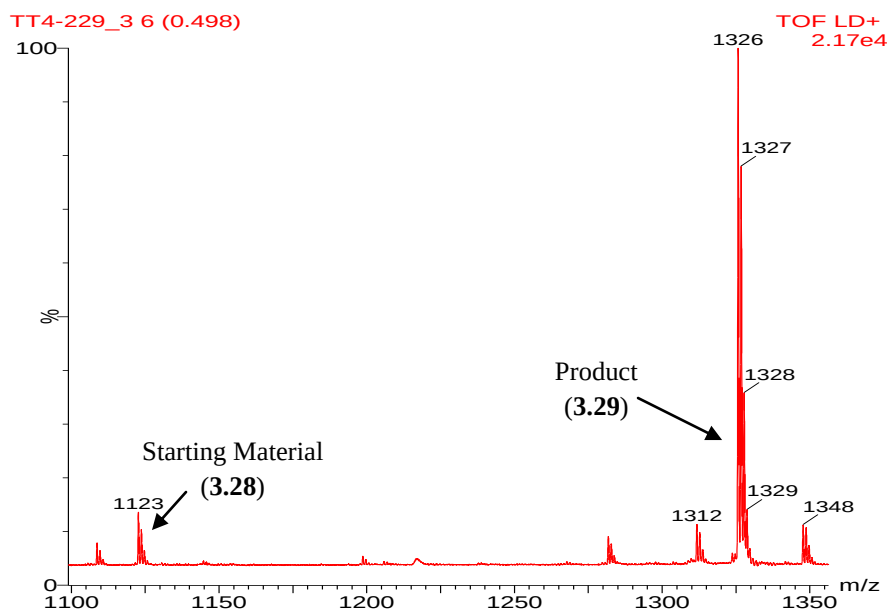
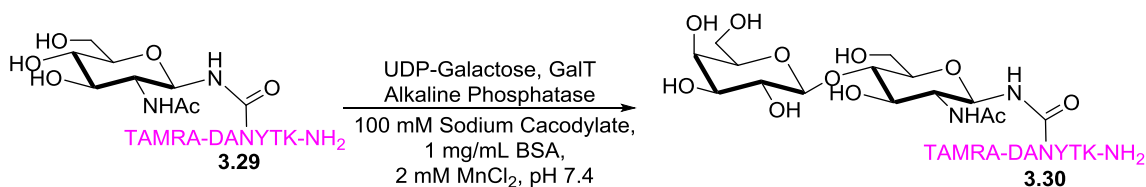


Figure 3.10: MALDI mass spectrum of PglB reaction showing formation of the desired TAMRA-GlcNAc product (3.29).

When sufficient *N*-glycosylated TAMRA-peptide was prepared, investigations could be done toward further glycosylation using GalT.

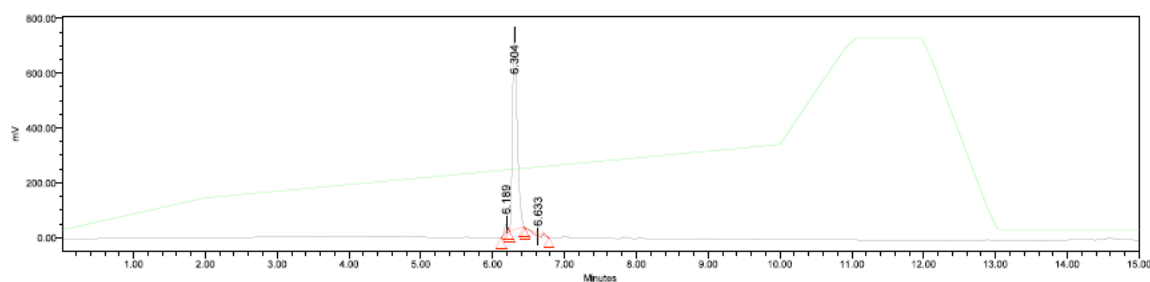
As the name suggests, Galactosyl Transferase (GalT) catalyses the addition of a galactose unit to a GlcNAc unit, in this instance, forming a 1,4- β -glycosidic bond. This glycosylation reaction uses UDP-galactose as the glycosyl donor substrate, with alkaline phosphatase being used as a co-enzyme. Initial attempts at this reaction on the TAMRA-GlcNAc-peptide used a relatively large amount of alkaline phosphatase and a small amount of GalT. Reactions were monitored by LC-MS (Table 3.1, entries 1 and 2) and showed the formation of the desired product (3.30) but also showed the loss of the lysine residue, both on the starting material and the desired glycosylated product. It was predicted that the loss of the lysine was due to the large amount of alkaline phosphatase being present. This was decreased, and an increased amount of GalT was used, aiming for complete conversion to the desired product. Pleasingly, these changes led to the

formation of the desired product (as seen by HRMS) in a high yield, as observed by fluorescent HPLC (Table 3.1, entries 3 and 4, Figure 3.11). Furthermore, the reaction was also seen to be complete overnight as opposed to 48 h.



Entry	GalT	Alkaline Phosphatase	UDP-Galactose	Time	Result
1	12 μ U	20 mU	13.3 equiv	48 h	3% yield + ~20% starting material with Lys hydrolysed
2	54 μ U	100 mU	66.7 equiv	48 h	10% yield + ~20% starting material with Lys hydrolysed. Product with Lys hydrolysed also seen
3	600 μ U	500 μ U	13.3 equiv	18 h	100% yield (94% by HPLC)
4	6 mU	500 mU	66.7 equiv	18 h	100% yield

Table 3.1: Optimisation of GalT glycosylation.⁹



	Name	Retention Time	Area	% Area
1	Impurity	6.189	72095	2.09
2	3.30	6.304	3257899	94.49
3	3.29	6.633	117709	3.41

Figure 3.11: HPLC trace of glycosylation reaction with GalT and TAMRA-GlcNAc-peptide.⁹

3.5.2 – EndoA Glycosylation

3.5.2.1 – EndoA Expression and Initial EndoA Glycosylation Reactions

After finding success with GalT, attention was turned to EndoA glycosylation reactions. As described in both Chapters 1 and 2, Endo enzymes are glycosyl hydrolases and naturally hydrolyse glycans off glycoproteins, but it is possible to use them to catalyse the addition of a glycan to a GlcNAc unit on a protein when using a glycan oxazoline donor.¹⁸¹⁹⁻²² First, EndoA had to be expressed. As with EndoS, described in Chapter 2, EndoA was expressed with a GST tag attached. It is reported that GST-EndoA is still active and cleavage of the tag is not necessary. The protein was expressed in competent cells (BL21(DE3)), with the cells then harvested and lysed using standard techniques. Purification by GST affinity column (GSTrap), followed by size exclusion purification resulted in formation of pure GST-EndoA (Figure 3.12).

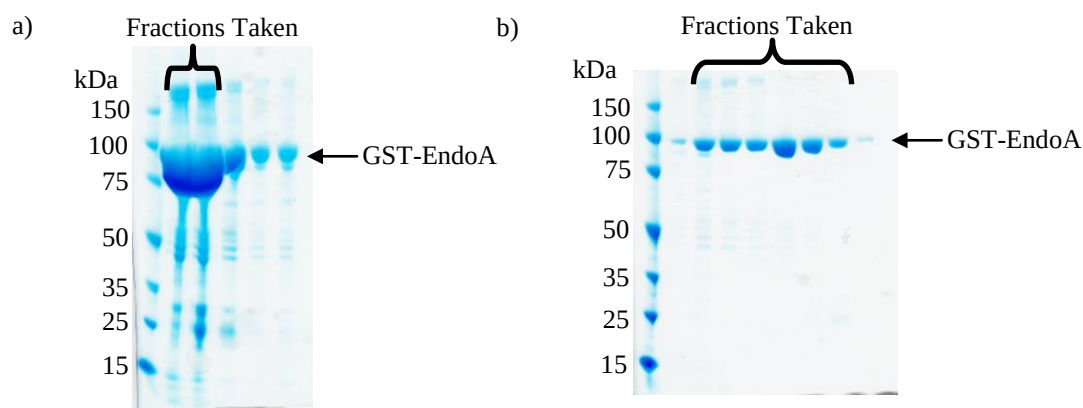
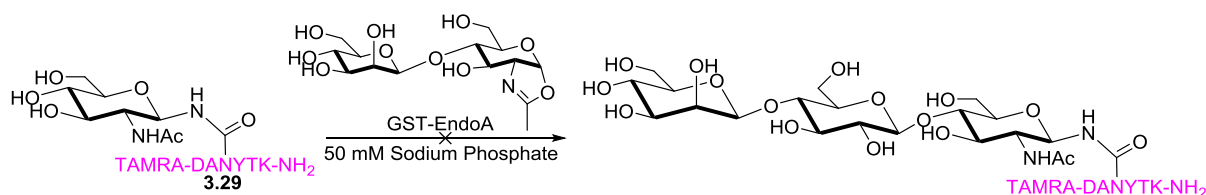


Figure 3.12: Purification of GST-EndoA, a) SDS-PAGE after GSTrap and b) SDS-PAGE after size exclusion.

With GST-EndoA in hand, investigations into the glycosylation reaction could begin, first looking at the reaction using a previously formed disaccharide oxazoline glycan donor.²³ A variety of different conditions were set up and, as with the GalT reactions, monitored by LC-MS. The critical variable in this reaction was the ratio of EndoA to

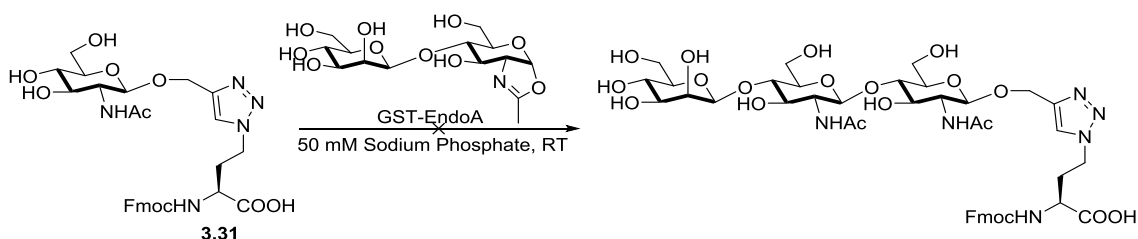
peptide substrate: too little EndoA would result in slow conversion, while too much EndoA would promote EndoA-catalysed hydrolysis of the product and/or disaccharide oxazoline donor. Unfortunately, however, all conditions and ratios of EndoA resulted in only starting material being seen (Table 3.2).



Entry	EndoA	EndoA:Peptide	Disaccharide	pH	Temp	Result
1	GST-EndoA	1:60	10 equiv	6.5	30 °C	No product
2	GST-EndoA	1:12	50 equiv	6.5	30 °C	No product
3	GST-EndoA	1:1500	10 equiv	6.5	25 °C	No product
4	GST-EndoA	1:4000	30 equiv	6	25 °C	No product

Table 3.2: Attempted optimisation of EndoA glycosylation with disaccharide oxazoline.

Due to the lack of any desired product, it was decided to look at a model system for which the glycosylation reaction is known to confirm that the enzyme was active. This would also aid in finding conditions that could be transferred to the TAMRA-GlcNAc peptide. For this purpose, an amino acid with a GlcNAc attached via a triazole (**3.31**) was used as a model substrate, which had been previously synthesised and used as a substrate for the attachment of the disaccharide oxazoline using EndoA.²³ Different amounts of EndoA were used and the reaction was regularly monitored, but the desired glycosylated product again was not detected (Table 3.3).

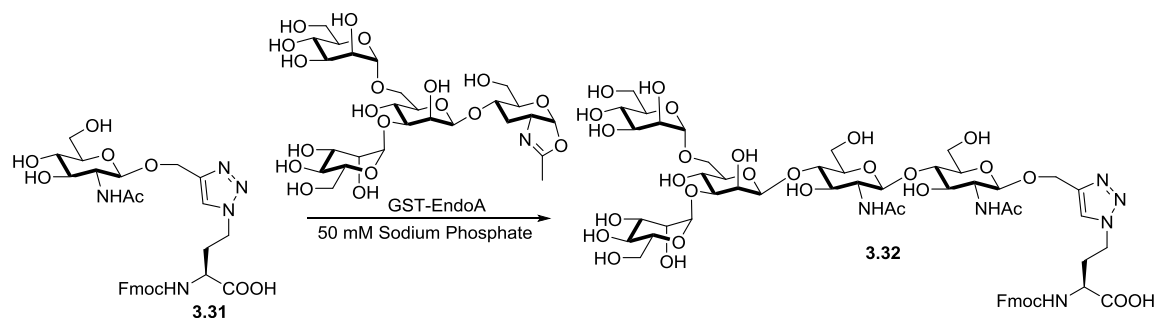


Entry	EndoA:Amino acid	pH	Result
1	1:20000	6.5	No product
2	1:15000	6.5	No Product
3	1:4000	6.5	No Product
4	1:30000	6	No Product
5	1:20000	6	No Product
6	1:4000	6	No Product
7	1:1500	6	No Product
8	1:750	6	No Product

Table 3.3: Attempted optimisation of EndoA glycosylation with disaccharide oxazoline on model amino acid.

It was postulated that the reason why no glycosylation product was seen with EndoA and the disaccharide oxazoline was that the concentration of the reaction was not high enough. The concentration of reagents previously reported for the formation of the glycosylated product using this system was higher than used here.²³ The disaccharide is possibly a poor substrate for glycosylation with EndoA, which requires high concentrations of reagents to form the glycosylated product. Unfortunately, it was not possible to concentrate the reaction further due to lack of material.

With no desired product observed from the disaccharide oxazoline, attention was turned to a tetrasaccharide oxazoline, Man₃GlcNAc. This larger oxazoline donor has been demonstrated to be a better substrate for Endo-catalysed glycosylation reactions, which could function at lower concentrations.²³ The triazole amino acid/GlcNAc substrate (3.31) was used as a model system to screen conditions that could be transferred to the TAMRA-GlcNAc peptide (3.28). Pleasingly, conditions were found that resulted in the desired product, using a ratio of approx. 1:4000, EndoA:substrate (Table 3.4, entry 2), as analysed by LC-MS and HPLC (Figure 3.13). The formation of the desired product (3.32) was confirmed by HRMS.



Entry	EndoA:Amino acid	Result
1	1:22000	No product
2	1:4000	Product seen after 15 min

Table 3.4: Attempted optimisation of EndoA glycosylation with tetrasaccharide oxazoline on model amino acid.

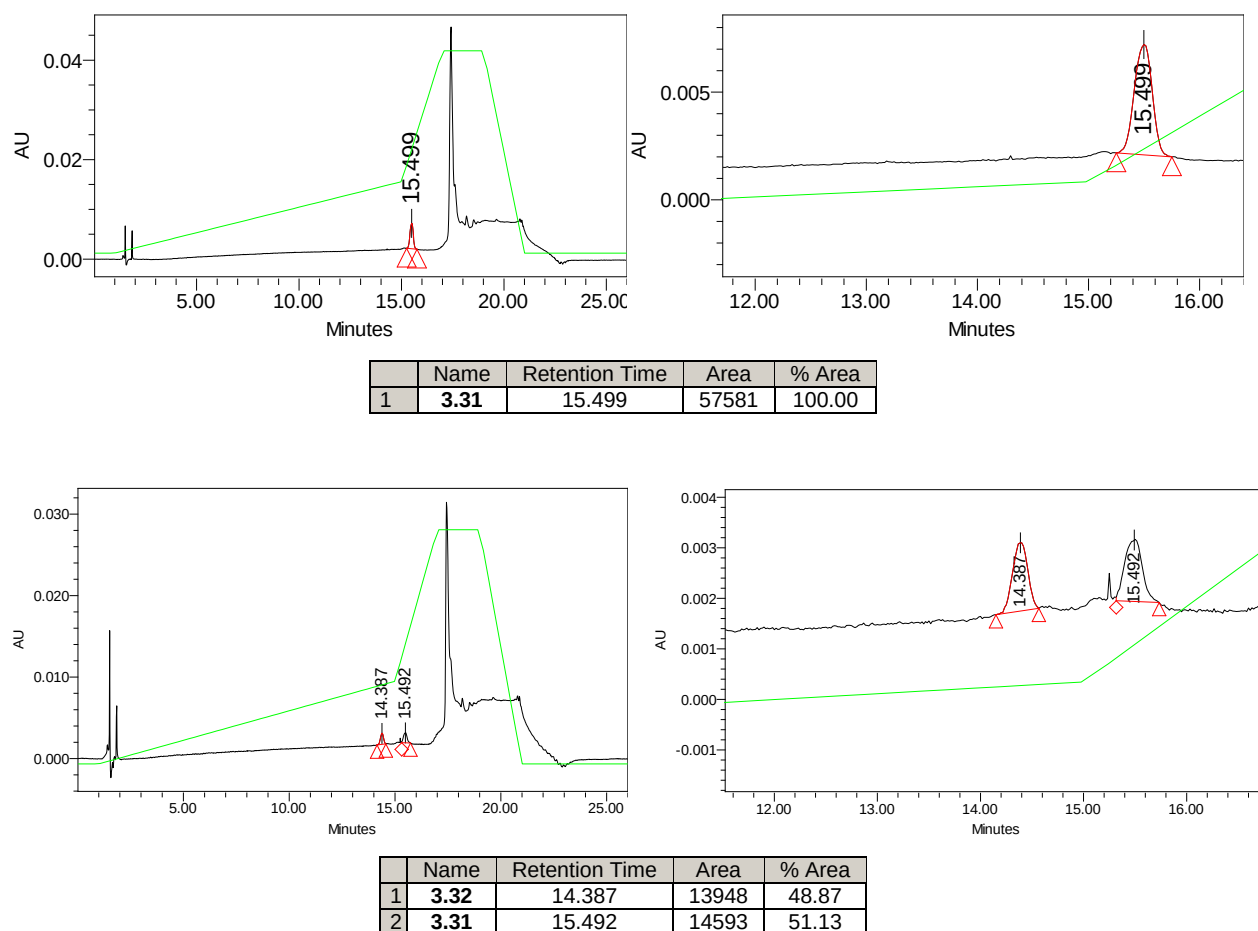
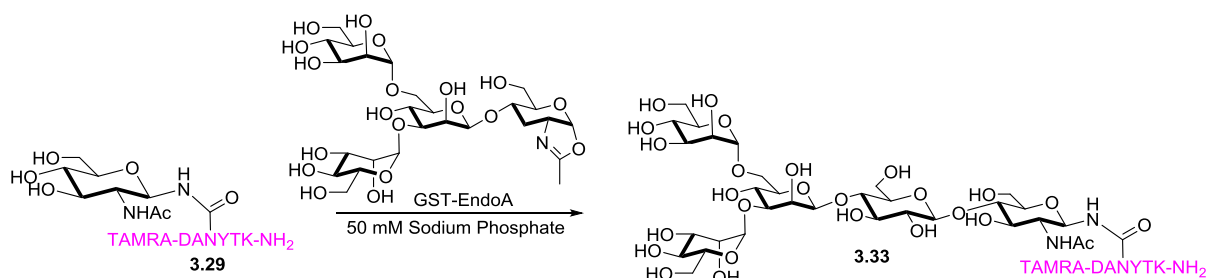


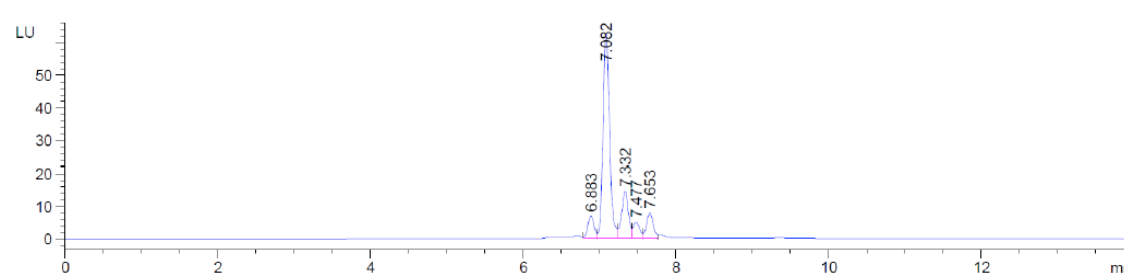
Figure 3.13: HPLC trace of Endo glycosylation on model amino acid, a) GlcNAc-amino acid (3.31) starting material, b) reaction after 30 min, peak at 14.4 min confirmed to be glycosylated product (3.32) by MS. Green line on each trace is the percentage amount of acetonitrile.

Transferring these conditions to the TAMRA-GlcNAc-peptide system and with further optimisation, the desired product (**3.33**) was formed successfully, characterized by LC-MS and fluorescent HPLC. Even though additional amounts of enzyme and tetrasaccharide oxazoline donor were added during the reaction, it was unable to promote the reaction to completion. However, a 64% yield of the desired material was detected by HPLC (Table 3.5, entry 3, Figure 3.14).



Entry	EndoA:Peptide	Disaccharide	pH	Temp	Result
1	1:60	10 equiv	6.5	30 °C	No product
2	1:1500	10 equiv	6.5	25 °C	No product
3	1:2000	8 equiv + 8 equiv + 8 equiv	6	25 °C	64% conversion after 4.5 h by HPLC

Table 3.5: Attempted optimisation of EndoA glycosylation with tetrasaccharide oxazoline.



	Name	Retention Time	Area	% Area
1	Unknown	6.883	39.72	6.81
2	3.33	7.082	376.29	64.47
3	3.29	7.332	89.68	14.44
4	Unknown	7.477	29.42	4.92
5	Unknown	7.653	48.47	7.75

Figure 3.14: HPLC trace of glycosylation reaction with GST-EndoA, tetrasaccharide oxazoline and TAMRA-GlcNAc-peptide.

3.5.2.2 – Further Glycosylation Reactions using EndoA Mutants

Being unable to obtain complete conversion to the glycosylated product using EndoA, investigations were undertaken to identify possible factors behind this using a time course reaction. EndoA mutants were also investigated, since a number of different

EndoA mutants have been reported for the formation of glycosylated products. The mutants investigated here were EndoA N171A, EndoA Y299F and EndoA E173H. EndoA N171A²⁴ and EndoA E173H²⁵ have both been reported to have no hydrolytic activity but retain some, though diminished, glycosylation activity, whereas EndoA Y299F²⁴ retains wild-type hydrolytic activity but has a 3-fold increase in glycosylation activity compared to wild-type. These mutants therefore could shed light on whether the glycosylation activity or the hydrolytic activity (on the product of the reaction) of the enzyme had the greatest effect on the conversion of the reaction.

All these mutants were formed using site-directed mutagenesis and were expressed as GST-tagged proteins, as with the wild-type EndoA. Purification by GST affinity column afforded the desired mutated proteins (Figure 3.15). For these mutants, size-exclusion purification was not carried out and only the most pure fraction from the GST affinity purification was taken.

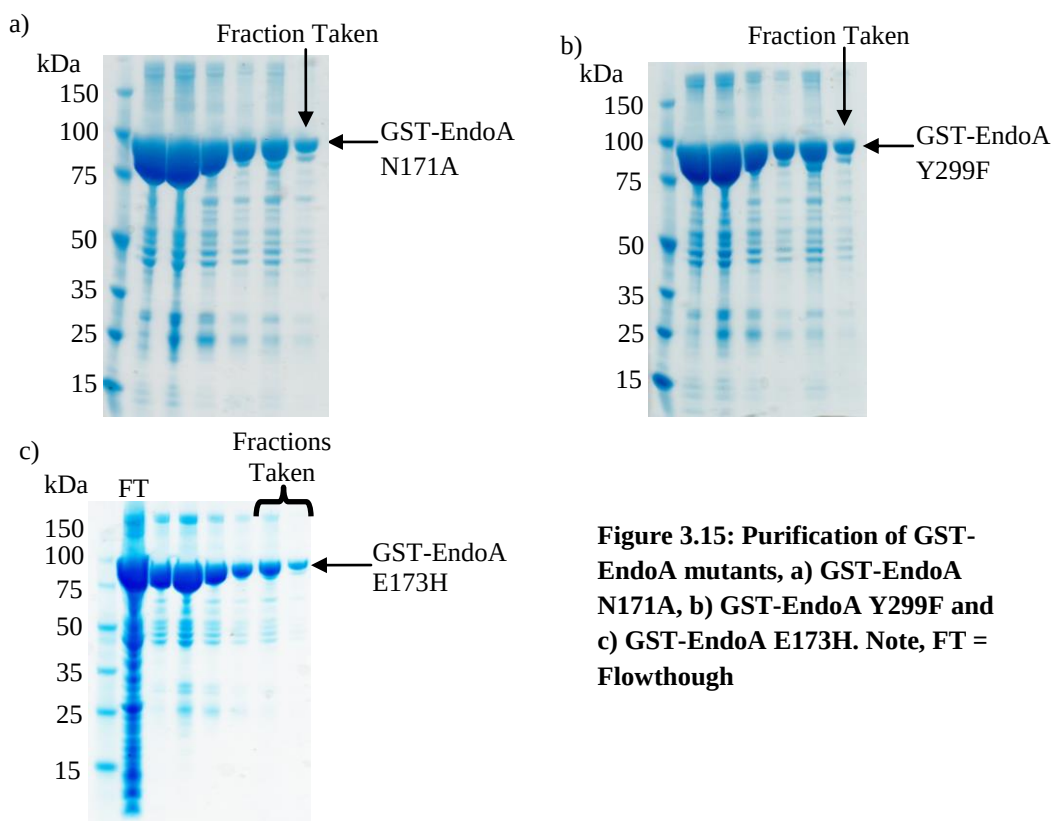


Figure 3.15: Purification of GST-EndoA mutants, a) GST-EndoA N171A, b) GST-EndoA Y299F and c) GST-EndoA E173H. Note, FT = Flowthrough

With the EndoA mutants in hand, first the hydrolysis activity of the enzymes was investigated. This was done by treating RNase B, a protein glycosylated with a high-mannose glycan, with wild type EndoA or EndoA mutants. As expected, EndoA Y299F had similar hydrolytic activity to wild type EndoA, whereas EndoA N171A and EndoA E173H had no hydrolytic activity (Figure 3.16).

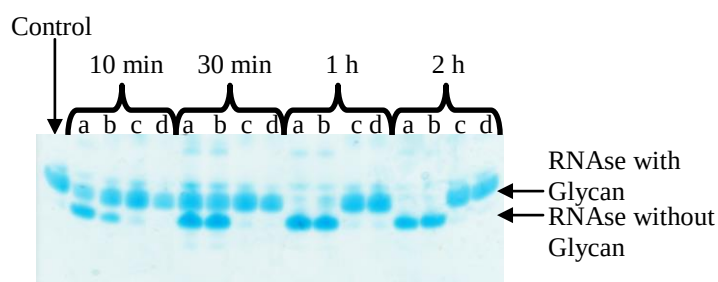


Figure 3.16: Hydrolytic activity of EndoA and mutants on RNase B, a) EndoA, b) EndoA Y299F, c) EndoA N171A and d) EndoA E173H.

Next, the glycosylation activity of the mutant enzymes was studied. This was done as before, using the TAMRA-GlcNAc peptide, glycosylated by PglB, and a single addition of both the tetrasaccharide oxazoline donor and the enzyme. The reactions were monitored by fluorescent HPLC and samples of the reactions were taken after 15 min, 30 min, 1 h, 2 h, 4 h, 24 h and 48 h. Again, integration of the peaks gave the reaction conversion. No product was observed when EndoA N171A or EndoA E173H were used, whereas significant glycosylation activity was seen for EndoS Y299F, greater than that detected for the wild-type enzyme (Figure 3.17). From the time course, it can be observed that maximum yield is obtained around 1 h with EndoA Y299F, whereas maximal yield is obtained after 24 h with EndoA. However, with both enzymes the majority of conversion occurred in the first hour. Interestingly, for both enzymes the rate of hydrolysis of the product was significantly slower than the glycosylation reaction, with 24% and 38% product still present after 48 h with EndoA and EndoA Y299F, respectively.

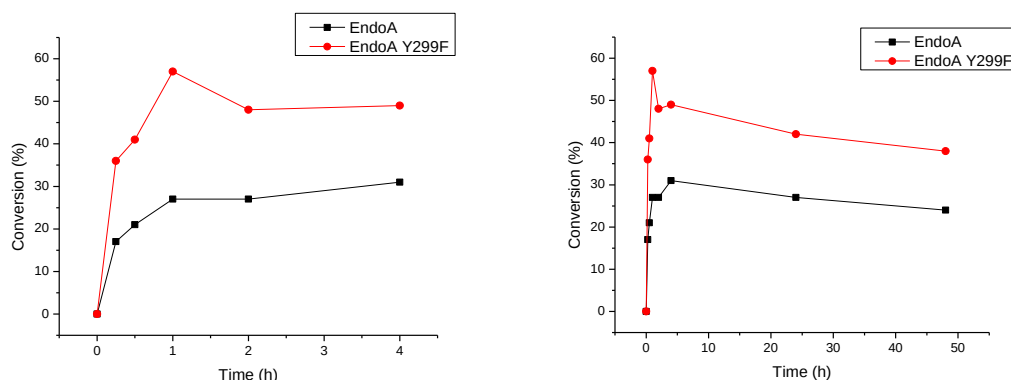


Figure 3.17: Glycosylation activity of EndoA and EndoA Y299F, a) over 4 h and b) over 48 h.

Although the conversion achieved with EndoA Y299F is slightly less than that achieved when using EndoA with multiple additions of tetrasaccharide oxazoline (64% vs 61%), it can reasonably be assumed that greater yields can be achieved with EndoA Y299F if there are multiple additions of tetrasaccharide oxazoline, as this was demonstrated here to be a superior glycosylation enzyme when compared to wild type EndoA under the same conditions. The lack of activity of EndoA N171A and EndoA E173H confirmed that their reported lack of hydrolytic activity comes at the price of significantly suppressed glycosylation activity. Increased loading of these two enzymes may result in some conversion. The lack of hydrolysis of the product was an interesting observation. Previously it had been shown that high loading of EndoA resulted in no conversion, and it was thought that the product was forming but hydrolysing before observation. The fact that hydrolysis of the product was measured to be slow here suggests that the issue may instead be hydrolysis of the tetrasaccharide oxazoline donor, meaning that it was unable to react with the TAMRA-GlcNAc peptide acceptor.

3.6 – Conclusions and Future Work

PglB was demonstrated to be a suitable biocatalyst for the synthesis of defined glycopeptides. Various lipid-sugar donors were characterized, to profile the effect of the substrate on this reaction. It was shown that it is possible to use a commercially available short lipid (C15, WE₂ lipid) to form a functioning PglB glycan donor substrate. Although this study adds two more lipid lengths to those already studied on PglB substrates, a more extensive study is required in order to elucidate if PglB has a conformational preference of the lipid on the PglB substrate and whether the *cis/trans* stereochemistry of the prenyl units in the lipid affect the activity of the substrate.

It has also been shown that it is possible to use PglB catalysis followed by other enzymatic glycosylation reactions to form glycopeptides containing different glycan structures. A clear outlook of this study is to apply the glycosyltransferase activity of PglB to form complex glycoproteins. This has briefly been demonstrated in the group via the glycosylation of AcrA, the natural protein acceptor substrate for PglB.⁹ Attempts have been made to repeat this result, however, there have been difficulties in following the reaction by LC-MS due to the presence of detergent in the reaction mixture. Tests into the removal of the detergent from protein solutions suggest that the dodecyl- β -maltopyranoside used in the PglB purification carried through into the purified protein solutions, affecting the analysis. Further studies should be done to see if other detergents could be successfully used in PglB purification.

3.7 - Biological Experimental

See Chapter 2 for protocols for SDS-PAGE, Western Blot and Preparation of Lysogeny Agar Plates.

3.7.1 - PglB Expression and Purification

3.7.1.1- Transformation of pBAD-myc-HisA PglB Plasmid into BL21 (DE3) RILX Cells

To a pre-chilled tube, BL21 (DE3) RILX cells (50 μ L) and the pBAD-myc-HisA PglB plasmid (2 μ L, $\sim$ 100 ng/ μ L) were added and incubated on ice for 30 min. After this time, the mixture was subjected to a 42 °C heat pulse for 30 sec before incubating on ice for a further 2 min. SOC media (0.5 mL) was added and the mixture incubated at 37 °C and 220 rpm for 1 h. After this time, a portion of the mix (50 μ L) was added to an LB-Amp agar plate and incubated overnight at 37 °C to allow colonies to form. The plates were then stored at 4 °C

3.7.1.2 - Over-Expression of PglB using Arabinose Induction^{1,9}

The expression of PglB was done on a 10 L scale. A single colony from the BL21 (DE3) RILX agar plate was added to LB-Amp media (100 mL, 100 μ L 100 mg/mL Amp). After incubating this media at 37 °C overnight, a portion of it (10 mL) was added to Terrific Broth (1 L, 10 mL glycerol, 1 mL 100 mg/mL Amp). The media was incubated for 4 h, when the OD was approximately 3, and the culture was induced by the addition of arabinose (10% w/v, 10 mL). The cultures were then incubated for 4 h at 28 °C. After this time, the culture was centrifuged (8000 rpm, 4 °C, 10 min) and the cell pellet collected ($\sim$ 50 g from 10 $\times$ 1 L). The cell pellet was stored at -80 °C.

3.7.1.3 - Lysis of Cells and Purification of PglB^{1,9}

Having been thawed on ice, the cell pellet (~50 g from 10 L culture) was resuspended in 25 mM Tris-HCl pH 8, 250 mM NaCl (200 mL) and Complete protease inhibitors (1 tablet) were added. After resuspending, the suspension was sonicated (4 × 1 min, 1 sec pulse, 0.5 pause, 1 min intervals between cycles) and centrifuged (20000 rpm, 20 min). The supernatant was then centrifuged a second time in an ultracentrifuge (30000 rpm, ~100000g) and the membrane fraction was collected. This was resuspended in 25 mM Tris-HCl pH 8, 250 mM NaCl, 10% v/v glycerol and 1% dodecyl- β -maltopyranoside (20 mL, loading buffer). The solution was then applied to NiNTA resin (1 mL). The column was first washed with loading buffer + 60 mM imidazole (3 mL, 3 column volumes) to remove unspecific bound proteins. The protein was eluted with loading buffer + 250 mM imidazole (3 mL, 3 column volumes) and fractions (0.5 mL) were collected. After analysis by SDS-PAGE and western blot, all the fractions containing the desired protein were combined and desalted into 10 mM MES, pH 6.5, 100 mM NaCl, 0.5 mM EDTA and 3% glycerol using a PD MidiTrap G-25 column (GE Life Sciences). The solution was concentrated by Amicon Ultra 0.5 mL concentrator (Merck Millipore). The protein solution was stored at -20 °C.

3.7.2 – Formation of EndoA Mutants and Expression and Purification of Wild Type EndoA and Mutants

3.7.2.1- Formation of EndoA Mutant Plasmids

Site directed mutagenesis was undertaken on the pGEX-4T-EndoA plasmid, which codes for GST-Tagged EndoA with a thrombin cleavage site, using a Stratagene Quikchange II Site-Directed Mutagenesis Kit using the following primers, with the mutated nucleotides highlighted.

Primer	Code (mutated bases highlighted)
EndoA N171A Forward	GACGGCTGGTTTATTGCCCAAGAAACAGAAGGG
EndoA N171A Reverse	CCCTTCTGTTTCTTGGGCAATAAACCAGCCGTC
EndoA E173H Forward	GGCTGGTTTATTAACCAACATACAGAAGGGGGCAGACGA AG
EndoA E173H Reverse	CTTCGTCTGCCCCTTCTGTATGTTGGTTAATAAACCAGC C
EndoA Y299F Forward	CATCACTCGGGTTATTCCGTCCAGATTGGG
EndoA Y299F Reverse	CCCAATCTGGACGGAATAACCCGAGTGATG

The following were all added together: Milli-Q water (35 µL), 10× reaction buffer (5 µL), DMSO (3 µL), pGEX-4T-EndoA plasmid (1 µL, 50 ng/µL), forward primer (2 µL, 62.5 ng/µL), reverse primer (2 µL, 62.5 ng/µL), dNTP mix (1 µL, 10 mM) and Pfa Ultra HF DNA polymerase (1 µL, 2.5 U/mL).

The solutions were then subject to the following PCR cycle:

95 °C – 2 min
 $\left(\begin{array}{l} 95\text{ }^{\circ}\text{C} - 30\text{ sec} \\ 55\text{ }^{\circ}\text{C} - 30\text{ sec} \\ 72\text{ }^{\circ}\text{C} - 10\text{ min} \end{array} \right) \times 25$
 10 °C - ∞

After this, Dpn 1 restriction enzyme (2 µL, 20 U/µL) was added and the mixture was incubated at 37 °C for 1 h. Transformation was then carried out as described below.

3.7.2.2 - Transformation of Mutant Plasmids into XL-Gold Ultracompetent Cells and Preparation of Cell Pellet for Miniprep

To a pre-chilled tube, XL-Gold Ultracompetent Cells (50 μ L) and XL-Gold β -mercaptoethanol mix (2 μ L) were added. After incubating for 10 min on ice, a portion of the PCR mix (2 μ L) was added and the mixture was incubated on ice for a further 30 min. After this time, the mixture was subjected to a 42 °C heat pulse for 30 sec before incubating on ice for a further 2 min. SOC media (0.5 mL) was added and the mixture incubated at 37 °C and 220 rpm for 1 h. After this time, a portion of the mix (50 μ L) was added to an LB-Amp agar plate and was incubated overnight at 37 °C to allow colonies to form. The plates were then stored at 4 °C

A single colony from one of the plates was added to LB-Amp media (8 mL, 8 μ L 100 mg/mL Amp) and this mix was incubated overnight at 37 °C and 220 rpm. The mixture was then centrifuged (3000 rpm, 15 min, 4 °C) and the supernatant discarded.

3.7.2.3 - Preparation of the Plasmid and Sequencing

Preparation of the mutant plasmid for sequencing was carried out using a MiniPrep Kit purchased from QIAGEN and following the included manual. The concentration of the resulting plasmid was measured using a Nanodrop spectrometer. A sample was prepared and submitted to Geneservice at the Department of Biochemistry, University of Oxford, for sequencing. The plasmid was sequenced with pGEX 3' and pGEX 5' to insure that the mutation was in place and no undesired mutations had occurred.

3.7.2.4 - Transformation of Mutant Plasmids into BL21 (DE3) Cells

To a pre-chilled tube, BL21 (DE3) cells (50 μ L) and the desired plasmid (2 μ L, ~100 ng/ μ L) were added and incubated on ice for 30 min. After this time, the mixture was subjected to a 42 °C heat pulse for 30 sec before incubating on ice for a further 2 min. Autoclaved SOC media (0.5 mL) was added and the mixture incubated at 37 °C and 220 rpm for 1 h. After this time, a portion of the mix (50 μ L) was added to an LB-Amp agar plate and was incubated overnight at 37 °C to allow colonies to form. The plates were then stored at 4 °C

3.7.2.5 - Over-expression of GST-EndoA Wild-Type and Mutants using IPTG Induction

The expression of EndoA wild type and mutants was done a 1 L scale. A single colony from the BL21 (DE3) agar plate was added to a LB-Amp media (20 mL, 20 μ L 100 mg/mL Amp). After incubating this media at 37 °C overnight, a portion of it (10 mL) was added to LB-Amp media (1 L, 1 mL 100 mg/mL Amp). The media was incubated for 3 h, when the OD was approximately 0.5-1 and the culture was induced by the addition of IPTG (0.1 M, 5 mL). The cultures were then incubated for 5 h at 30 °C. After this time, the cultures were centrifuged (8000 rpm, 4 °C, 10 min) and the cell pellet collected (~6 g per litre culture). The cell pellet was stored at –80 °C.

3.7.2.6 – Lysis of Cells and Purification of GST-EndoA Wild Type and Mutants

Having been thawed on ice, the cell pellet was resuspended in PBS buffer (~3 mL per 1 g cell pellet) and lysozyme (1 mg per 10 mL PBS) and Complete protease inhibitors (1 tablet) were added. After resuspending, the suspension was sonicated (4 $\times$ 1 min, 1 sec pulse, 0.5 pause, 1 min intervals between cycles), centrifuged (20000 rpm, 20 min) and

the supernatant filtered (0.2 µm membrane). The solution was then applied to FPLC using two GSTrap HP columns (2 × 5 mL columns) and PBS buffer (Buffer A) and 30 mM Tris-HCl, 10 mM reduced L-glutathione (pH 8) buffer (Buffer B). The buffer gradient was 5% Buffer B for 2 column volumes (CV) to 100% Buffer B for 2 CV before washing the column with Buffer A with 3 CV. Flow rate was 1 mL/min. After analysis by SDS-PAGE, all the fractions containing the desired protein were combined and concentrated by Vivaspin. This solution was then dialysed into PBS buffer twice (15 h + 3 h) and subjected to Bicinchoninic Acid assay (BCA assay). The protein solution was stored at –80 °C.

3.7.2.7- Superdex Purification of GST-EndoA

A small sample (7 mg in 3 mL buffer) of GST-EndoA purified by GSTrap column was subjected to gel filtration for further purification. The sample was applied to gel filtration using a 120 mL Superdex 200 column on a FPLC with PBS buffer. Flow rate was 1 mL/min. Fractions that contained clean desired protein as analysed by SDS-PAGE were combined and concentrated using Vivaspin and the concentration of the protein solution was determined by BCA assay. The sample was then stored at –20 °C.

3.7.2.8 - Protein Yields

Protein	Yield
GST-EndoA N171A	1.8 mg/L
GST-EndoA E173H	2.0 mg/L
GST-EndoA Y299F	1.2 mg/L

Note, this is the yield of only the most pure fractions after GST purification, greater yields would have been achieved with further purification.

3.8 – Protein Sequences

3.8.1- PglB

Nucleotide Sequence –

```

ATGGTGAAAA AAGAGTATTT AAAAAACCCT TATTTAGTTT TGTTTGCGAT GATTATATTA
GCTTATGTTT TTAGTGATTT TTGCAGGTTT TATTGGGTTT GGTGGGCAAG TGAGTTTAAT
GAGTATTTTT TCAATAATCA GTTAATGATC ATTTCAAATG ATGGCTATGC TTTTGCTGAG
GGCGCAAGAG ATATGATAGC AGGTTTTTCAT CAGCCTAATG ATTTGAGTTA TTATGGATCT
TCTTTATCCG CGCTTACTTA TTGGCTTTAT AAAATCACAC CTTTTTCTTT TGAAAGTATC
ATTTTATATA TGAGTACTTT TTTATCTTCT TTGGTGGTGA TTCCTACTAT TTTGCTAGCT
AACGAATACA AACGTCCTTT AATGGGCTTT GTAGCTGCTC TTTTAGCAAG TATAGCAAAC
AGTTATTATA ATCGCACTAT GAGTGGGTAT TATGATACGG ATATGCTGGT AATTGTTTTG
CCTATGTTTA TTTTATTTTT TATGGTAAGA ATGATTTTAA AAAAAGACTT TTTTTCATTG
ATTGCCTTGC CGTTATTTAT AGGAATTTAT CTTTGGTGGT ATCCTTCAAG TTATACATTA
AATGTAGCTT TAATTGGACT TTTTAAATAT TATACACTTA TTTTTCATAG AAAAGAAAAG
ATTTTTTATA TAGCTGTGAT TTTGTCTTCT CTTACTCTTT CAAATATAGC ATGGTTTTAT
CAAAGTGCCA TTATAGTAAT ACTTTTTTGCT TTATTCGCCT TAGAGCAAAA ACGCTTAAAT
TTTATGATTA TAGGAATTTT AGGTAGTGCA ACTTTGATAT TTTTGATTTT AAGTGGTGGG
GTTGATCCTA TACTTTATCA GCTTAAATTT TATATTTTGA GAAGTGATGA AAGTGCGAAT
TTAACGCAGG GCTTTATGTA TTTTAATGTC AATCAAACCA TACAAGAAGT TGAAAATGTA
GATCTTAGCG AATTTATGCG AAGAATTAGT GGTAGTGAAA TTGTTTTTTTT GTTTTCTTTG
TTTGGTTTTG TATGGCTTTT GAGAAAACAT AAAAGTATGA TTATGGCTTT ACCTATATTG
GTGCTTGGGT TTTTAGCCTT AAAAGGGGGG CTTAGATTTA CCATTTATTC TGTACCTGTA
ATGGCCTTAG GATTTGTTTT TTTATTGAGC GAGTTTAAGG CTATAATGGT TAAAAAATAT
AGCCAATTAA CTTCAAAATG TTGTATTGTT TTTGCAACTA TTTTGACTTT AGCTCCAGTA
TTTATCCATA TTTACAACTA TAAAGCGCCA ACAGTTTTTT CTCAAAATGA AGCATCATTA
TTAAATCAAT TAAAAAATAT AGCCAATAGA GAAGATTATG TGGTAACTTG GTGGGATTAT
GGTTATCCTG TGCGTTATTA TAGCGATGTG AAAACTTTAG TAGATGGTGG AAAGCATTTA
GGTAAGGATA ATTTTTTCCC TTCTTTTGCT TTAAGCAAAG ATGAACAAGC TGCAGCTAAT
ATGGCAAGAC TTAGTGTAGA ATATACAGAA AAAAGCTTTT ATGCTCCGCA AAATGATATT
TTAAAAACAG ACATTTTGCA AGCCATGATG AAAGATTATA ATCAAAGCAA TGTGGATTTG
TTTCTAGCTT CATTATCAAA ACCTGATTTT AAAATCGATA CGCCAAAAAC TCGTGATATT
TATCTTTATA TGCCCGCTAG AATGTCCTTG ATTTTTTCTA CGGTGGCTAG TTTTCTTTTT
ATTAATTTAG ATACAGGAGT TTTGGATAAA CCTTTTACCT TTAGCACAGC TTATCCACTT
GATGTTAAAA ATGGAGAAAAT TTATCTTAGC AACGGAGTGG TTTTAAGCGA TGATTTTAGA
AGTTTTAAAA TAGGTGATAA TGTGGTTTCT GTAAATAGTA TCGTAGAGAT TAATTCTATT
AAACAAGGTG AATACAAAAT CACTCCAATT GATGATAAGG CTCAGTTTTA TATTTTTTAT
TTAAAGGATA GTGCTATTCC TTACGCACAA TTTATTTTAA TGGATAAAAC CATGTTTAAT
AGTGCTTATG TGCAAATGTT TTTTATAGGA AATTATGATA AGAATTTATT TGAATTGGTG
ATTAATTCTA GAGATGCTAA GGTTTTTAAA CTTAAAATTC ATCATCATCA TCATCATCAT
CATCATCATT GA

```

Amino Acid Sequence –

```

MVKKEYLKNP YLVLFAMIIL AYVFSVFCRF YWVWWASEFN EYFFNNQLMI ISNDGYAF AE
GARDMIAGFH QPNDSLSSYGS SLSALTYWLY KITPFSFESI ILYMSTFLSS LVVIPTILLA
NEYKRPLMGF VAALLASIAN SYNRTMSGY YDTDMLVIVL PMFILFFMVR MILKKDFFSL
IALPLFIGIY LWYWPSSYTL NVALIGLFLI YTILFHRKEK IFYIAVILSS LTLSNIAWFY
QSAIIVILFA LFALEQKRLN FMIIGILGSA TLIFLILSGG VDPILYQLKF YIFRSDSAN
LTQGFMYFNV NQTIQEVENV DLSEFMRRIS GSEIVFLFSL FGFVWLLRKH KSMIMALPIL
VLGFLALKGG LRFTIYSVPV MALGFGFLLS EFKAIMVKKY SQLTSNVCIV FATILT LAPV
FIHIYNYKAP TVFSQNEASL LNQLKNIANR EDYVVTWWDY GYPVRYSDV KTLVDGGKHL
GKDNFFPSFA LSKDEQAAAN MARLSVEYTE KSFYAPQNDI LKTDILQAMM KDYNQSNVDL
FLASLSKPDF KIDTPKTRDI YLYMPARMSL IFSTVASFSF INLDTGVLDK PFTFSTAYPL

```

DVKNGEIYLS NGVVLSDDFR SFKIGDNVVS VNSIVEINSI KQGEYKITPI DDKAQFYIFY
LKDSAIPYAQ FILMDKTMFN SAYVQMFFLG NYDKNLFDLV INSRDAKVFK LKIHNNNNNNH
HHH

3.8.2 – EndoA

Nucleotide Sequence –

TCTACGTACA ACGGCCCCGCT GTCCTCCCAT TGGTTTCCAG AGGAACTTGC CCAATGGGAA
CCAGACAGTG ATCCAGACGC ACCCTTTAAC AGAAGCCATG TTCCGCTGGA ACCAGGCCGC
GTTGCGGATA GGGTAAATGC TAATGCAGAC AAGGACGCAC ACCTTGTTTC GTTGTCGCG
CTAAACAGGC ATACATCAGG TGTTCCATCG CAAGGAGCGC CAGTTTTCTA TGAAAATACG
TTCAGCTATT GGCATTATAC AGATTTGATG GTTTATTGGG CTGGTTCAGC TGGCGAAGGC
ATTATCGTTC CGCCAAGTGC CGATGTCATT GATGCATCGC ACCGAAATGG GGTGCCGATT
TTAGGAAAATG TGTTCTTCCC GCCGACGGTT TATGGAGGGC AGCTAGAGTG GCTAGAACAA
ATGTTAGAGC AAGAGGAGGA CGGTTTCATTC CCCCTTGCTG ACAAATTGCT AGAAGTCGCA
GACTATTATG GGTGTGACGG CTGGTTTATT AACCAAGAAA CAGAAGGGGC AGACGAAGGA
ACAGCCGAAG CCATGCAAGC TTTTCTCGTT TATTTGCAGG AACAAAAGCC AGAAGGCATG
CACATCATGT GGTATGACTC GATGATTGAT ACAGGGGCGA TCGCCTGGCA AAACCATTTA
ACGGATCGAA ATAAAATGTA CTTGCAAAAT GGCTCGACCC GCGTCGCTGA CAGCATGTTT
TTGAACTTTT GGTGGCGTGA CCAGCGCCAA TCGAACGAAT TGGCACAAGC ACTTGGCAGG
TCTCCGTATG ACCTCTATGC CGGAGTGGAT GTGGAAGCAC GAGGGACAAG TACCCCTGTT
CAGTGGGAAG GCCTGTTTCC TGAAGGAGAA AAGGCGCATA CATCACTCGG GTTATACCGT
CCAGATTGGG CATTTTCAGTC AAGTGAAACA ATGGAAGCGT TTTATGAAAA AGAACTACAA
TTTTTGGGTTG GCTCGACAGG AAATCCAGCC GAAACAGACG GCCAGTCAAA TTGGCCTGGC
ATGGCGCACT GGTTTCCCGC GAAAAGCACC GCTACTTCGG TACCCTTTGT GACTCACTTT
AATACGGGCA GCGGCGCTCA GTTTTCGGCA GAAGGCAAAA CTGTGTGCGA ACAGGAATGG
AATAACCGCA GCCTTCAAGA TGTGCTGCCG ACATGGCGCT GGATTCAGCA TGGCGGCGAT
TTAGAGGCAA CATTTTCTTG GGAAGAAGCG TTTGAAGGGG GAAGCTCGTT ACAATGGCAT
GGCTCATTAG CGGAAGGAGA ACACGCCCAA ATCGAGCTCT ATCAAACAGA GTTGCCGATA
AGCGAAGGCA CTTGCTAAC GTGGACATTT AAAAGCGAGC ACGACAACGA TTTAAATGTG
GGCTTCCGTT TAGATGGGGA AGAGGACTTC CGTTATGTGG AAGGAGAACA GCGTGAATCG
ATAAATGGTT GGACGCAAGT GACGTTGCCG CTGGATGCGT TTGCTGGTCA GACGATAACA
GGGCTGGCAT TTGCAGCGGA AGGGAATGAG ACTGGGCTGG CAGAATTCTA TATTGGACAA
CTGGCCGTAG GTGCTGATAG CGAAAAGCCT GCCGCTCCAA ACGTGAACGT ACGCCAGTAC
GACCCAGACC CGAGTGGCAT TCAGCTCGTA TGGGAAAAAC AAAGCAACGT CCACCATTAC
CGCGTTTATA AAGAAACAAA GCACGGCAAA GAGCTAATTG GCACATCTGC TGGAGATCGA
ATTTACCTAG AAGGCCTAGT CGAGGAAAGC AAACAAAACG ACGTGCCTCT GCATATAGAA
GCACTAAGTG AAACATTTGT GCCAAGTGAT GCTCGCATGA TCGACATAAA AAGCGGCTCG
TTTTAG

Amino Acid Sequence –

STYNGPLSSH WFPEELAQWE PDSDPDAPFN RSHVPLEPGR VADRVNANAD KDAHLVLSLA
LNRHTSGVPS QGAPVFYENT FSYWHYTDLM VYWAGSAGEG IIVPPSADVI DASHRNGVPI
LGNVFFPPTV YGGQLEWLEQ MLEQEEDGSF PLADKLLEVA DYYGFDGWF I NQETEGADEG
TAEAMQAFIV YLQEQKPEGM HIMWYDSMID TGAIAWQNH L TDRNKMYLQN GSTRVADSMF
LNFWRDQRR SNELAQALGR SPYDLYAGVD VEARGTSTPV QWEGLFPEGE KAHTSLGLYR
PDWAFQSSET MEAFYEKELQ FWVGSTGNPA ETDGQSNWPG MAHWFPKST ATSVPFVTHF
NTGSGAQFSA EGKTVSEQEW NNRSLQDVLP TWRWQHGGD LEATFSWEEA FEGGSSLQWH

Chapter 3 –Investigations into the Activity of PglB In Vitro for the Synthesis of Glycopeptides

GSLAEGEHAQ IELYQTELPI SEGTSLTWTF KSEHVNDLNV GFRLDGEEDF RYVEGEQRES
INGWTQWTLP LDAFAGQTIT GLAFAAEGNE TGLAEFYIGQ LAVGADSEKP AAPNVNVRQY
DPDPSGIQLV WEKQSNVHHY RVYKETKHGK ELIGTSAGDR IYLEGLVEES KQNDVRLHIE
ALSETFVPSD ARMIDIKSGS F

3.9 - Chemistry Experimental

3.9.1 - General

Reaction Conditions: Where reactions are said to be done using dry solvents, these were carried out under an argon atmosphere, using oven dried glassware and anhydrous solvents obtained either by filtration through drying column or from commercial dry sources.

Flash Column Chromatography: Flash Column Chromatography was performed using 230-400 mesh silica. Pre-coated aluminum backed silica gel plates were used for thin layer chromatography and plates were visualized using UV light (254 nm) and/or Phosphomolybdic acid/Ce (IV) dye.

Infrared Spectra: Absorbencies ($\nu_{\max}$) are reported in wavenumbers (cm^{-1}). Spectra were obtained using a Bruker Tensor 27 Spectrometer.

Nuclear Magnetic Spectra: Data is reported as following: Chemical shift, integration, multiplicity (s=singlet, d=doublet, t=triplet and m=multiplet), coupling constants (if applicable) and assignment. Chemical shift values are given in parts per million from tetramethylsilane for both ^1H and ^{13}C NMR spectra and solvent resonance was used as the internal standard for both types of spectra (^1H NMR spectra: CHCl_3 : $\delta = 7.27$ ppm, MeOD: $\delta = 3.31$ ppm, d_6 -DMSO: $\delta = 2.50$, D_2O : $\delta = 4.79$. ^{13}C NMR spectra: CDCl_3 : $\delta = 77.00$ ppm, MeOD: $\delta = 49.0$ ppm and d_6 -DMSO: $\delta = 39.5$). J coupling values are reported in Hz. Where inseparable mixtures of diastereoisomers were formed, if they had a coincidental chemical shift they are reported with an integration of 1 H. Where diastereoisomers have separate chemical shifts they are reported as either 1 H^{maj} and 1 H^{min} for the major and minor diastereoisomer, respectively, as 1 H^{α} and 1 H^{β} for α and β anomers, respectively, or as 1 H^{a} or 1 H^{b} if there were a 1:1 ratio of the two

diastereoisomers. All ^{13}C NMR spectrums reported are ^1H decoupled. Spectra were recorded on a Bruker AMX-400 or AMX-500.

Mass Spectrometry: Mass spectrometry was performed on an Agilent Technologies 6120 Quadrupole LC/MS or on a Waters LCT Premier spectrometer. High Resolution mass spectra were obtained a Bruker μTOF spectrometer.

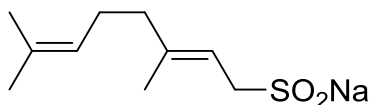
Melting Points: Melting points were obtained on a Leica Galen III hot plate.

Optical Rotations: Optical rotations are reported as measured from the sodium D line ($[\alpha]_{\text{D}}$, 589 nm), with the concentration (with units 10mg/mL) and solvent given. Optical rotations were obtained on a Perkin Elmer Model 341 Polarimeter or a Schmidt Haensch Unipol L Polarimeter at room temperature in a 10 cm, 1 mL cell.

Reagents: Commercial reagents were used as supplied or purified using standard techniques.

3.9.2 – Compounds

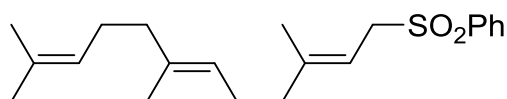
(*E*)-1-Benzenesulfonyl-3,7-dimethylocta-2,6-diene (**3.1**)⁹



A solution of geraniol (1 g, 6.5 mmol) in dry THF (2 mL) was cooled to 0 °C and phosphorus bromide (0.3 mL, 3.25 mmol) was added. After stirring for 1 h at this temperature, the reaction was diluted with diethyl ether (30 mL) and then washed with saturated sodium bicarbonate (20 mL), water (20 mL) and brine (20 mL) and dried with sodium sulfate. The solvent was removed under reduced pressure and the material was re-dissolved in dry DMF (4 mL). The solution was cooled to 0 °C and benzenesulfinic acid sodium salt (1.4 g, 8.7 mmol) was added. The reaction was stirred at 0 °C for 4 h after which the reaction was diluted with ethyl acetate (30 mL). The mixture was washed with 1 M HCl (20 mL), saturated sodium bicarbonate (20 mL), water (20 mL) and brine (20 mL) dried with sodium sulfate. The solvent was removed under reduced pressure and purification by flash chromatography (85:15, pet. ether: ethyl acetate) afforded the lipid product (**3.1**, 1.06 g, 3.8 mmol, 59%) as a colourless oil. *R_f* 0.32 (85:15, pet. ether: ethyl acetate). ¹H NMR (400 MHz, CDCl₃) δ 7.88 – 7.83 (2 H, m, 2 × Ph), 7.65 – 7.59 (1 H, m, 1 × Ph), 7.55 – 7.48 (2 H, m, 2 × Ph), 5.17 (1 H, t, *J* = 7.0, H-2), 5.05 – 4.99 (1 H, m, H-6), 3.80 (2 H, d, *J* = 8.1, 2 × H-1), 2.01 – 1.96 (4 H, m, 2 × H-4 and 2 × H-5), 1.67 (3 H, s, 3 × H-9), 1.57 (3 H, s, 3 × H-10) and 1.29 (3 H, s, 3 × H-8). ¹³C NMR (100 MHz, CDCl₃) δ 146.6 (C-3), 138.6 (Ph), 133.6 (Ph), 132.0 (C-7), 128.9 (2 × Ph), 128.5 (2 × Ph), 123.5 (C-6), 110.3 (C-2), 56.1 (C-1), 39.7 (C-5), 26.2 (C-4), 25.7 (C-8), 17.7 (C-10) and 16.1 (C-9). MS (ESI): *m/z* Calcd for C₁₆H₂₂O₂SNa [M+Na]⁺: 301.1: Found 301.1.

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.⁹

(2E, 6E)-1-Benzenesulfonyl-3,7,11-trimethyldodeca-2,6,10-triene (3.2)¹⁶

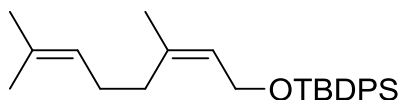


A solution of *trans, trans* farnesol (0.5 g, 2.3 mmol) in dry THF (2 mL) was cooled to 0 °C and phosphorus bromide (0.1 mL, 1.2 mmol) was added. After stirring for 1 h at this temperature, the reaction was diluted with diethyl ether (30 mL) and then washed with saturated sodium bicarbonate (20 mL), water (20 mL) and brine (20 mL) and dried with sodium sulfate. The solvent was removed under reduced pressure and the material was re-dissolved in dry DMF (4 mL). The solution was cooled to 0 °C and benzenesulfinic acid sodium salt (1.4 g, 8.7 mmol) was added. The reaction was stirred at 0 °C for 4 h after which the reaction was diluted with ethyl acetate (30 mL). The mixture was washed with 1 M HCl (20 mL), saturated sodium bicarbonate (20 mL), water (20 mL) and brine (20 mL) dried with sodium sulfate. The solvent was removed under reduced pressure and purification by flash chromatography (85:15, pet. ether: ethyl acetate) afforded the lipid product (**3.2**, 0.59 g, 1.7 mmol, 75%) as a colourless oil. *R_f* 0.30 (85:15, pet. ether: ethyl acetate). ^1H NMR (400 MHz, CDCl_3) δ 7.90 – 7.85 (2 H, m, 2 $\times$ Ph), 7.67 – 7.61 (1 H, m, 1 $\times$ Ph), 7.57 – 7.51 (2 H, m, 2 $\times$ Ph), 5.20 (1 H, t, J = 7.8, H-2), 5.12 – 5.03 (2 H, m, H-6 and H-10), 3.81 (2 H, d, J = 7.9, 2 $\times$ H-1), 2.11 – 1.95 (8 H, m, 2 $\times$ H-4, 2 $\times$ H-5, 2 $\times$ H-8 and 2 $\times$ H-9), 1.68 (3 H, s, 3 $\times$ H-13), 1.60 (3 H, s, 3 $\times$ H-14), 1.59 (3 H, s, 3 $\times$ H-15) and 1.31 (3 H, s, 3 $\times$ H-12). ^{13}C NMR (100 MHz, CDCl_3) δ 146.5 (C-3), 138.6 (Ph), 135.8 (Ph), 133.6 (C-7), 1313.5 (C-11), 129.0 (2 $\times$ Ph), 128.6

(2 × Ph), 124.2 (C-6), 123.3 (C-10), 110.3 (C-2), 56.1 (C-1), 39.7 (C-5 and C-9), 26.7 (C-4), 26.2 (C-8), 25.7 (C-12), 17.7 (C-15) and 16.2 (C-13 and C-14). MS (ESI): m/z Calcd for $C_{21}H_{30}O_2SNa$ $[M+Na]^+$: 369.2: Found 369.2.

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.¹⁶

***t*-Butyldiphenylsilyl neryl ether (3.3)⁹**

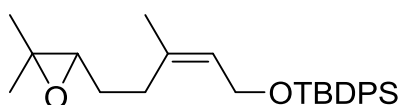


To a solution of imidazole (14.5 g, 213 mmol) in dry DMF (75 mL) was added nerol (17 mL, 97 mmol) and *t*-butyldiphenylsilyl chloride (35.4 mL, 137 mmol). After stirring for 30 min at room temperature, the reaction was poured into stirring hexane (400 mL). The layers were separated and the DMF layer was extracted a second time with hexane (400 mL). The combined organic layers were washed with brine (100 mL) and dried with sodium sulfate. The solvent was removed under reduced pressure and purification by flash chromatography (98:2, pet. ether: diethyl ether) afforded the protected nerol (**3.3**, 36.2 g, 92.1 mmol, 95%) as a colourless oil. R_f 0.32 (85:15, pet. ether: ethyl acetate). ^1H NMR (400 MHz, $CDCl_3$) δ 7.80 – 7.75 (4 H, m, 4 × Ph), 7.52 – 7.52 (6 H, m, 6 × Ph), 5.48 (1 H, t, J = 6.3, H-2), 5.07 (1 H, tt, J = 6.7, 1.4, H-6), 4.28 (2 H, d, J = 6.4, 2 × H-1), 2.08 – 2.00 (2 H, m, 2 × H-4) 1.99 – 1.93 (2 H, m, 2 × H-5), 1.79 (3 H, s, 3 × H-9), 1.69 (3 H, s, 3 × H-10), 1.59 (3 H, s, 3 × H-8) and 1.13 (9 H, s, $C(CH_3)_3$). ^{13}C NMR (100 MHz, $CDCl_3$) δ 137.5 (C-3), 135.7 (4 × Ph), 134.1 (2 × Ph), 131.8 (C-7), 129.6 (2 × Ph), 127.7 (4 × Ph), 125.0 (C-2), 124.0 (C-6), 60.9 (C-1), 32.3 (C-5), 26.9

(C(CH₃)₃), 26.6 (C-4), 25.7 (C-8), 23.5 (C-9), 19.2 (C(CH₃)₃) and 17.7 (C-10). MS (ESI): *m/z* Calcd for C₂₆H₃₆OSiNa [M+Na]⁺: 415.2: Found 415.3.

The observed characterisation data (¹H and ¹³C) was consistent with that previously reported in the literature.⁹

(Z)-1-(*tert*-butyldiphenylsilyloxy)-6,7-epoxy-3,7-dimethylocta-2-ene (3.4)⁹

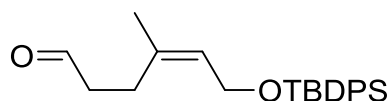


A solution of **3.3** (27 g, 69 mmol) in THF (150 mL) and water (60 mL) was cooled to 0 °C and *N*-bromosuccinimide (13.5 g, 76 mmol) was added over 30 min. The reaction was stirred for a further 6 h. After this time, the solvent was removed under reduced pressure and the resulting residue was extracted with diethyl ether (2 × 200 mL). The combined organic layers were washed with saturated sodium bicarbonate (75 mL), water (75 mL) and brine (75 mL) and dried with sodium sulfate. The solvent was removed under reduced pressure and the crude material was re-dissolved in dry methanol (150 mL) and potassium carbonate (18.1 g, 131 mmol) was added. The reaction was stirred for 16 h at room temperature after which the reaction was filtered and the solvent was removed under reduced pressure. The residue was re-dissolved in pet. ether and filtered a second time and the solvent again removed under reduced pressure. Purification by flash chromatography (98:2, pet. ether: ethyl acetate) afforded the protected nerol epoxide (**3.4**, 17.6 g, 43.1 mmol, 63%) as a colourless oil. *R_f* 0.32 (85:15, pet. ether:ethyl acetate). ¹H NMR (400 MHz, CDCl₃) δ 7.79 – 7.75 (4 H, m, 4 × Ph), 7.50 – 7.42 (6 H, m, 6 × Ph), 5.52 (1 H, t, *J* = 6.1, H-2), 4.28 (2 H, d, *J* = 6.4, 2 × Ph),

H-1), 2.64 (1 H, t, $J = 6.4$, H-6), 2.13 – 2.00 (2 H, m, $2 \times$ H-4), 1.78 (3 H, s, $3 \times$ H-9), 1.70 – 1.60 (1 H, m, H-5), 1.55 – 1.44 (1 H, m, H-5), 1.28 (3 H, s, $3 \times$ H-10), 1.22 (3 H, s, $3 \times$ H-8) and 1.12 (9 H, s, $C(CH_3)_3$). ^{13}C NMR (100 MHz, $CDCl_3$) δ 136.6 (C-3), 135.7 ($4 \times$ Ph), 134.0 ($2 \times$ Ph), 129.6 ($2 \times$ Ph), 127.7 ($4 \times$ Ph), 125.5 (C-2), 63.9 (C-6), 60.7 (C-1), 58.3 (C-7), 28.8 (C-4), 27.5 (C-5), 26.9 ($C(CH_3)_3$), 24.9 (C-10), 23.4 (C-9), 19.2 ($C(CH_3)_3$) and 18.7 (C-8). MS (ESI): m/z Calcd for $C_{26}H_{36}O_2SiNa$ $[M+Na]^+$: 431.2: Found 431.2.

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.⁹

(Z)-1-(tert-butyldiphenylsilyloxy)-3-methyl-2-hexene-6-al (3.5)⁹

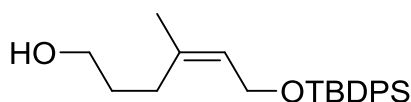


A solution of **3.4** (17.5 g, 42.9 mmol) in dioxane (25 mL) and water (12 mL) was cooled to 0 °C and periodic acid (12.2 g, 53.6 mmol) was added over 10 min. The reaction was then stirred at 0 °C for 16 h, after which it was filtered through Celite. The mixture was extracted with EtOAc (200 mL) and the organic layer was washed with saturated sodium bicarbonate (50 mL) and brine (50 mL) and dried with sodium sulfate. The solvent was removed under reduced pressure and purification by flash chromatography (9:1, pet. ether: ethyl acetate) afforded the aldehyde (**3.5**, 9.7 g, 26.5 mmol, 62%) as a colourless oil. R_f 0.32 (85:15, pet. ether:ethyl acetate). ^1H NMR (400 MHz, $CDCl_3$) δ 9.69 (1 H, s, H-6), 7.81 – 7.73 (4H, m, $4 \times$ Ph), 7.52 – 7.40 (6 H, m, $6 \times$ Ph), 5.51 (1 H, t, $J = 6.3$, H-2), 4.27 (2 H, d, $J = 6.4$, $2 \times$ H-1), 2.41 (2 H, td, $J = 8.0$, 1.4,

2 × H-5), 2.23 (2 H, t, J = 7.9, 2 × H-4), 1.75 (3 H, s, 3 × H-7) and 1.12 (9 H, s, C(CH₃)₃). ¹³C NMR (100 MHz, CDCl₃) δ 201.9 (C-6), 135.7 (4 × Ph), 134.9 (2 × Ph), 133.8 (C-3), 129.7 (2 × Ph), 127.7 (4 × Ph), 126.1 (C-2), 60.5 (C-1), 42.2 (C-5), 26.9 (C(CH₃)₃), 24.5 (C-4), 23.1 (C-7) and 19.2 (C(CH₃)₃). MS (ESI): *m/z* Calcd for C₂₃H₃₀O₂SiNa [M+Na]⁺: 389.2: Found 389.1.

The observed characterisation data (¹H and ¹³C) was consistent with that previously reported in the literature.⁹

(Z)-1-O-*tert*-Butyldiphenylsilyl-3-methylhexa-2-en-1,6-diol (3.6)⁹

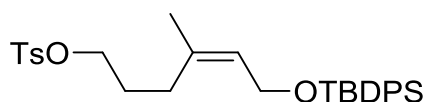


A solution of **3.5** (17.5 g, 47.8 mmol) in ethanol (75 mL) was cooled to 0 °C and sodium borohydride (3.45 g, 91 mmol) was added over 10 min. The reaction was then stirred at 0 °C for 3 h, after which the reaction was quenched by the addition of 1 M HCl until evolution of gas stopped. The mixture was filtered through Celite and the solvent removed under reduced pressure. The resulting residue was re-dissolved in diethyl ether (250 mL) and washed with brine (50 mL) and dried with sodium sulfate. The solvent was removed under reduced pressure and purification by flash chromatography (85:15, pet. ether: ethyl acetate) afforded the alcohol (**3.6**, 10.6 g, 28.5 mmol, 60%) as a yellow oil. *R_f* 0.22 (85:15, pet. ether:ethyl acetate). ¹H NMR (400 MHz, CDCl₃) δ 7.77 – 7.72 (4 H, m, 4 × Ph), 7.50 – 7.40 (6 H, m, 6 × Ph), 5.46 (1 H, t, J = 6.7, H-2), 4.23 (2 H, d, J = 6.7, 2 × H-1), 3.57 (2 H, t, J = 6.3, 2 × H-6), 2.09 (2 H, t, J = 7.5, 2 × H-4), 1.75 (3 H, s, 3 × H-7), 1.63 (2 H, quintet, J = 6.3, H-5) and 1.09 (9 H, s, C(CH₃)₃). ¹³C NMR (100

MHz, CDCl₃) δ 138.4 (C-3), 135.7 (4 $\times$ Ph), 133.8 (2 $\times$ Ph), 129.7 (2 $\times$ Ph), 127.7 (4 $\times$ Ph), 124.8 (C-2), 61.6 (C-6), 60.5 (C-1), 30.2 (C-4), 27.7 (C-5), 26.8 (C(CH₃)₃), 23.2 (C-7) and 19.2 (C(CH₃)₃). MS (ESI): m/z Calcd for C₂₃H₃₂O₂SiNa [M+Na]⁺: 391.2: Found 391.2.

The observed characterisation data (¹H and ¹³C) was consistent with that previously reported in the literature.⁹

(Z)-1-O-*tert*-Butyldiphenylsilyl-3-methyl-6-O-(4-toluenesulfonyl)hexa-2-en-1,6-diol
(3.7)⁹

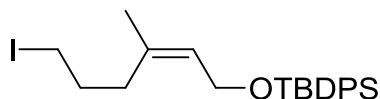


To **3.6** (10.5 g, 28.5 mmol) in dry pyridine (20 mL) was added tosyl chloride (6.5 g, 34.2 mmol) in two portions. After stirring at room temperature for 2 h, the solvent was removed under reduced pressure and the resulting residue re-dissolved in pet. ether (250 mL). This mixture was filtered and the filtrate washed with 50 mM copper sulfate (50 mL), water (50 mL), saturated sodium bicarbonate (50 mL) and brine (50 mL). After drying with sodium sulfate, the solvent was removed under reduced pressure and purification by flash chromatography (7:3, pet. ether:ethyl acetate) afforded the tosyl product (**3.7**, 11.9 g, 23 mmol, 81%) as a yellow oil. R_f 0.68 (7:3, pet. ether:ethyl acetate). ¹H NMR (400 MHz, CDCl₃) δ 7.79 – 7.70 (6 H, m, 6 $\times$ Ph), 7.50 – 7.39 (6 H, m, 6 $\times$ Ph), 7.31 (2 H, d, J = 8.1, 2 $\times$ Ph), 5.44 (1 H, t, J = 6.3, H-2), 4.16 (2 H, d, J = 7.2, 2 $\times$ H-1), 3.93 (2 H, t, J = 6.6, 2 $\times$ H-6), 2.45 (3 H, s, Ts-CH₃), 1.90 (2 H, dd, J = 8.2, 6.5, 2 $\times$ H-4), 1.67 (3 H, d, J = 0.9, 3 $\times$ H-7), 1.65 (2 H, dd, J = 8.7, 6.6, H-5) and

1.08 (9 H, s, C(CH₃)₃). ¹³C NMR (100 MHz, CDCl₃) δ 144.8 (Ts), 135.8 (C-3), 135.6 (4 × Ph), 133.9 (2 × Ph), 133.2 (Ts), 129.9 (2 × Ts), 129.7 (2 × Ph), 127.9 (2 × Ts), 127.7 (4 × Ph), 126.0 (C-2), 70.1 (C-6), 60.5 (C-1), 27.8 (C-4), 27.2 (C-5), 26.9 (C(CH₃)₃), 23.2 (C-7), 21.7 (Ts-CH₃) and 19.2 (C(CH₃)₃). MS (ESI): *m/z* Calcd for C₃₀H₃₈O₄SSiNa [M+Na]⁺: 545.2: Found 545.2.

The observed characterisation data (¹H and ¹³C) was consistent with that previously reported in the literature.⁹

(Z)-1-*O*-*tert*-Butyldiphenylsilyl-3-methyl-6-iodo-hexa-2-en-1-ol (3.8)⁹

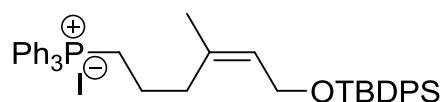


To a solution of 3.7 (11.9 g, 23 mmol) in dry acetone (60 mL) was added sodium iodide (8.55 g, 57 mmol). The reaction was refluxed for 16 h, after which the reaction was cooled to room temperature and a 10% sodium thiosulfate solution (50 mL) was added. The reaction was extracted with pet. ether (3 × 100 mL). The organic layer was dried with sodium sulfate and the solvent removed under reduced pressure to afford the iodo product (**3.8**, 10.2 g, 21.3 mmol, 93%) as a colourless oil, which was used without further purification. *R_f* 0.27 (85:15, pet. ether:ethyl acetate). ¹H NMR (400 MHz, CDCl₃) δ 7.61 (4 H, dd, *J* = 7.6, 1.5, 4 × Ph), 7.37 – 7.27 (6 H, m, 6 × Ph), 5.37 (1 H, t, *J* = 6.4, H-2), 4.13 (2 H, d, *J* = 6.4, 2 × H-1), 2.92 (2 H, t, *J* = 7.0, 2 × H-6), 1.89 - 1.82 (2 H, m, 2 × H-4), 1.76 - 1.67 (2 H, m, 2 × H-5), 1.60 (3 H, s, 3 × H-7) and 0.97 (9 H, s, C(CH₃)₃). ¹³C NMR (100 MHz, CDCl₃) δ 135.77 (C-3), 135.69 (4 × Ph), 134.0 (2 × Ph), 129.6 (2 × Ph), 127.7 (4 × Ph), 126.2 (C-2), 60.7 (C-1), 32.8 (C-4), 32.0 (C-5), 26.9

(C(CH₃)₃), 23.3 (C-7), 19.2 (C(CH₃)₃) and 6.3 (C-6). MS (ESI): *m/z* Calcd for C₂₃H₃₁OISiNa [M+Na]⁺: 501.1: Found 501.1.

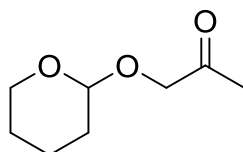
The observed characterisation data (¹H and ¹³C) was consistent with that previously reported in the literature.⁹

[(Z)-1-*O*-*tert*-Butyldiphenylsilyl-3-methyl-hexa-2-en-1-ol][6-triphenylphosphonium iodide] (3.9)⁹



To a solution of **3.8** (10 g, 21 mmol) in dry toluene (45 mL) was added triphenylphosphine (7.4 g, 28 mmol, freshly re-crystallised from ethanol) under argon. The reaction was heated to reflux and was allowed to stir at reflux for 16 h. After this time, the reaction was cooled to room temperature and pet. ether (150 mL) was added. The resulting solid was filtered off and washed with diethyl ether (75 mL) to remove excess triphenylphosphine. The solid was dried under reduced pressure to afford the desired phosphonium iodide salt (**3.9**, 10.2 g, 13.8 mmol, 66%) as white solid, which was used without further purification or characterisation.

1-(Tetrahydropyran-2-yloxy)propan-2-one (3.10)¹⁷

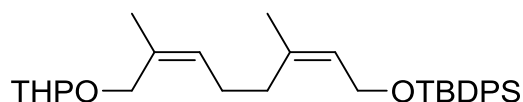


To 3,4-dihydro-2H-pyran (10 mL, 110 mmol) was added hydroxyacetone (3.77 mL, 55 mmol) and pyridium *p*-toluenesulfonate (138 mg, 0.55 mmol) under anhydrous

conditions. The reaction was stirred for 2.5 h at room temperature before being directly purified by flash chromatography (9:1, pet. ether: ethyl acetate) to afford the protected alcohol (**3.10**, 3.43 g, 22 mmol, 39%) as a colourless oil. *R_f* 0.1 (9:1, pet. ether:ethyl acetate). ¹H NMR (400 MHz, CDCl₃) δ 4.61 (1 H, t, *J* = 3.7, H-1), 4.25 (1 H, d, *J* = 17.4, 1 × OCH₂CO), 4.10 (1 H, d, *J* = 17.4, 1 × OCH₂CO), 3.86 – 3.78 (1 H, m, 1 × H-5), 3.53 – 3.47 (1 H, m, 1 × H-5), 2.16 (3 H, s, CH₃) and 1.90 – 1.49 (6 H, m, 2 × H-2, 2 × H-3 and 2 × H-4). ¹³C NMR (100 MHz, CDCl₃) δ 206.7 (CO), 98.7 (C-1), 72.3 (OCH₂CO), 62.3 (C-5), 30.3 (C-2), 26.5 (CH₃), 25.3 (C-4) and 19.2 (C-3). MS (ESI): *m/z* Calcd for C₈H₁₄O₃Na [M+Na]⁺: 181.1: Found 181.1.

The observed characterisation data (¹H and ¹³C) was consistent with that previously reported in the literature.¹⁷

(2Z,6Z)-1-O-(*tert*-butyldiphenylsilyloxy)-3,7-dimethyl-8-(2-tetrahydropyranyloxy)-octa-2,6-diene (3.11)⁹

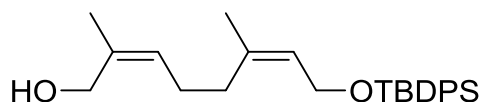


To **3.9** (10.2 g, 13.8 mmol) re-suspended in dry THF (50 mL) was added butyl lithium (10 mL, 1.6 M solution in hexane, 16 mmol) over 5 min at –78 °C. After stirring for 1.5 h at –78 °C, **3.10** (2.92 g, 18.5 mmol) was added in dry THF (12 mL) over 5 min. The solution stirred for 16 h and allowed to warm to room temperature over this time. The reaction was then quenched by the addition of water (50 mL) and extracted with pet. ether (2 × 250 mL). After drying with sodium sulfate, the solvent was removed under reduced pressure and purification by flash chromatography (9:1, pet. ether: ethyl

acetate) afforded the protected diene product (**3.11**, 5.1 g, 10.4 mmol, 75%) as a colourless oil. *R_f* 0.59 (9:1, pet. ether:ethyl acetate). ¹H NMR (400 MHz, CDCl₃) δ 7.74 (4 H, dd, *J* = 7.6, 1.5, 4 × Ph), 7.47 – 7.40 (6 H, m, 6 × Ph), 5.45 (1 H, t, *J* = 6.1, H-2), 5.28 (1 H, t, *J* = 7.2, H-6), 4.57 (1 H, t, *J* = 3.7, H-1'), 4.24 (2 H, d, *J* = 6.3, 2 × H-1), 4.06 (1 H, d, *J* = 11.4, 1 × H-8), 4.04 (1 H, d, *J* = 11.4, 1 × H-8), 3.91 – 3.84 (1 H, m, 1 × H-5'), 3.53 – 3.47 (1 H, m, 1 × H-5'), 2.14 – 2.06 (2 H, m, 2 × H-5), 1.95 (2 H, t, *J* = 7.3, 2 × H-4), 1.88 – 1.81 (1 H, m, 1 × H-3'), 1.75 (6 H, s, 3 × H-9 and 3 × H-10), 1.73 – 1.67 (1 H, m, 1 × H-2'), 1.65 – 1.48 (4 H, m, 1 × H-2', 1 × H-3' and 2 × H-4') and 1.10 (9 H, s, C(CH₃)₃). ¹³C NMR (100 MHz, CDCl₃) δ 137.0 (C-3), 135.6 (4 × Ph), 134.0 (2 × Ph), 132.2 (C-7), 129.6 (2 × Ph), 128.9 (C-6), 127.6 (4 × Ph), 125.3 (C-2), 97.5 (C-1'), 65.2 (C-8), 62.1 (C-5'), 60.8 (C-1), 32.4 (C-4), 30.7 (C-2'), 26.9 (C(CH₃)₃), 26.1 (C-5), 25.5 (C-4'), 23.4, 21.7 (C-9 and C-10), 19.5 (C-3') and 19.2 (C(CH₃)₃). MS (ESI): *m/z* Calcd for C₃₁H₄₄O₃SiNa [M+Na]⁺: 515.3: Found 515.3.

The observed characterisation data (¹H and ¹³C) was consistent with that previously reported in the literature.⁹

(2Z,6Z)-1-O-(*tert*-butyldiphenylsilyloxy)-3,7-dimethylocta-2,6-dien-1,8-diol (3.12)⁹

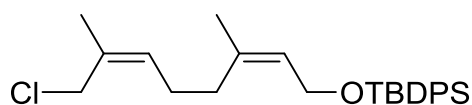


To a solution **3.11** (3g, 6.1 mmol) in isopropanol (20 mL) and ethanol (10 mL) was added *p*-toluenesulfonic acid (46 mg, 0.24 mmol). The reaction was stirred at room temperature for 48 h, after which the reaction was diluted with diethyl ether (150 mL) and washed with saturated sodium bicarbonate (50 mL) and brine (50 mL). After drying

with sodium sulfate, the solvent was removed under reduced pressure and purification by flash chromatography (85:15, pet. ether:ethyl acetate) afforded the protected diene product (**3.12**, 2.1 g, 5.2 mmol, 85%) as a colourless oil. *R_f* 0.36 (85:15, pet. ether:ethyl acetate). ¹H NMR (400 MHz, CDCl₃) δ 7.72 – 7.67 (4 H, m, 4 × Ph), 7.47 – 7.36 (6 H, m, 6 × Ph), 5.43 (1 H, t, *J* = 6.1, H-2), 5.16 (1 H, t, *J* = 7.3, H-6), 4.16 (2 H, d, *J* = 6.1, 2 × H-1), 3.99 (2 H, s, 2 × H-8), 2.08 – 2.00 (2 H, m, 2 × H-5), 1.91 (2 H, t, *J* = 7.5, 2 × H-4), 1.72 (6 H, d, *J* = 0.9, 3 × H-9 and 3 × H-10), 1.35 (1 H, s, -OH) and 1.06 (9 H, s, C(CH₃)₃). ¹³C NMR (100 MHz, CDCl₃) δ 137.1 (C-3), 135.6 (4 × Ph), 135.1 (C-7), 133.9 (2 × Ph), 129.6 (2 × Ph), 127.64 (4 × Ph), 127.35 (C-6), 125.4 (C-2), 61.4 (C-8) 60.8 (C-1), 32.3 (C-4), 26.9 (C(CH₃)₃), 26.0 (C-5), 23.5 (C-9) 21.3 (C-10) and 19.2 (C(CH₃)₃). MS (ESI): *m/z* Calcd for C₂₆H₃₆O₂SiNa [M+Na]⁺: 431.2: Found 431.2.

The observed characterisation data (¹H and ¹³C) was consistent with that previously reported in the literature.⁹

(2Z,6Z)-1-O-(tert-butyldiphenylsilyloxy)-8-chloro-3,7-dimethylocta-2,6-diene (3.13)⁹

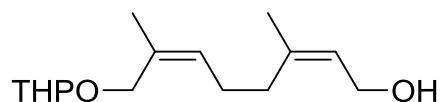


A solution of **3.12** (2 g, 4.9 mmol) in dry DMF (15 mL) was cooled to –40 °C and 2,3,5-collidine (2.6 mL, 20 mmol) and methanesulfonyl chloride (1.2 mL, 15 mmol) was added. After stirring for 30 min, lithium chloride (0.64 g, 15 mmol, dried with toluene) was added in dry DMF (10 mL). The reaction was stirred for 2.5 h whilst slowly warming to 0 °C. The reaction was then stirred at 0 °C for 30 min. After this time, the reaction was quenched with ice-cold saturated sodium bicarbonate (20 mL) and

extracted with 1:1 diethyl ether:hexane (2 × 75 mL). The organic layer was washed with saturated ammonium chloride (50 mL) and brine (50 mL) and dried with sodium sulfate. The solvent was removed under reduced pressure and purification by flash chromatography (9:1, pet. ether:ethyl acetate) afforded the chloro diene product (**3.13**, 1.8 g, 4.2 mmol, 86%) as a colourless oil. *R_f* 0.71 (95:5, pet. ether:ethyl acetate). ¹H NMR (400 MHz, CDCl₃) δ 7.72 – 7.68 (4 H, dd, *J* = 7.6, 1.5, 4 × Ph), 7.46 – 7.36 (6 H, m, 6 × Ph), 5.43 (1 H, t, *J* = 6.0, H-2), 5.25 (1 H, td, *J* = 7.3, 1.0, H-6), 4.19 (2 H, d, *J* = 6.0, 2 × H-1), 3.95 (2 H, s, 2 × H-8), 2.09 – 2.02 (2 H, m, 2 × H-5), 1.93 (2 H, t, *J* = 8.0, 2 × H-4), 1.75 (3 H, d, *J* = 0.9, 3 × H-9), 1.72 (3 H, d, *J* = 1.0, 3 × H-10) and 1.06 (9 H, s, C(CH₃)₃). ¹³C NMR (100 MHz, CDCl₃) δ 136.6 (C-3), 135.6 (4 × Ph), 134.0 (C-7), 130.4 (C-6), 129.6 (2 × Ph), 127.6 (6 × Ph), 125.6 (C-2), 60.8 (C-1), 43.5 (C-8), 31.9 (C-4), 26.9 (C(CH₃)₃), 26.3 (C-5), 23.4 (C-9) 21.5 (C-10) and 19.2 (C(CH₃)₃). MS (ESI): *m/z* Calcd for C₂₆H₃₅OClSiNa [M+Na]⁺: 449.2: Found 449.2.

The observed characterisation data (¹H and ¹³C) was consistent with that previously reported in the literature.⁹

(2Z,6Z)-3,7-dimethyl-8-(2-tetrahydropyranyloxy)-octa-2,6-dien-1,8-diol (3.14**)⁹**



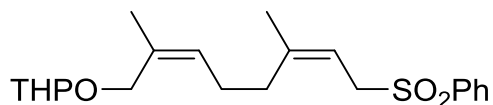
A solution of **3.11** (2 g, 4.1 mmol) in dry THF (10 mL) was cooled to 0 °C and tetra-*N*-butylammonium fluoride (4.1 mL, 1 M solution in THF) was added dropwise. The reaction was then stirred for 16 h at 0 °C. After this time, the mixture was diluted with pet. ether (50 mL), washed with brine (20 mL) and dried with sodium sulfate. The

solvent was removed under reduced pressure and purification by flash chromatography (9:1, CH₂Cl₂:ethyl acetate) afforded the protected diene product (**3.14**, 0.8 g, 3.2 mmol, 77%) as a colourless oil. Note: Rotomers were observed in the ¹³C NMR. *R_f* 0.26 (9:1, CH₂Cl₂:ethyl acetate). ¹H NMR (400 MHz, CDCl₃) δ 5.45 (1 H, t, *J* = 7.2, H-2), 5.39 (1 H, t, *J* = 6.9, H-6), 4.62 (1 H, t, *J* = 3.2, H-1'), 4.13 – 4.08 (4 H, m, 2 × H-1 and 2 × H-8), 3.93 – 3.85 (1 H, m, 1 × H-5'), 3.58 – 3.50 (1 H, m, 1 × H-5'), 2.23 – 2.10 (4 H, m, 2 × H-4 and 2 × H-5), 1.88 – 1.81 (1 H, m, 1 × H-3'), 1.79 (3 H, s, 3 × H-10), 1.76 (3 H, s, 3 × H-9), 1.74 - 1.67 (1 H, m, 1 × H-2') and 1.66 – 1.50 (4 H, m, 1 × H-2', 1 × H-3' and 2 × H-4'). ¹³C NMR (100 MHz, CDCl₃) δ 140.4, 140.0, 139.6, 139.2 (C-3), 135.6, 135.4, 132.5, 132.3 (C-7), 129.0, 128.8, 127.5, 127.4 (C-6), 124.81, 124.75, 121.9, 121.8 (C-2), 97.9, 97.7, 97.5, 97.3 (C-1'), 65.3, 65.2, 63.4, 63.2 (C-8), 62.22, 62.20, 62.1, 62.0 (C-5'), 61.30, 61.29, 58.91, 58.87 (C-1), 32.31, 32.27, 32.1 (C-4), 30.70, 30.66, 30.58 (C-2'), 26.4, 26.3, 26.2, 26.1 (C-5), 25.50, 25.48, 25.47 (C-4'), 23.63, 23.60, 23.5 (C-9), 21.78, 21.76, 21.43, 21.41 (C-10), 19.57, 19.51, 19.50 and 19.3 (C-3'). MS (ESI): *m/z* Calcd for C₁₅H₂₆O₃Na [M+Na]⁺: 277.2: Found 277.2.

The observed characterisation data (¹H and ¹³C) was consistent with that previously reported in the literature.⁹

(2Z,6Z)-1-benzenesulfonyl-3,7-dimethyl-8-(2-tetrahydropyranyloxy)-octa-2,6-diene

(3.15)⁹

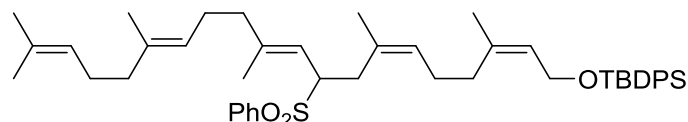


A solution of **3.14** (0.75 g, 3 mmol) in dry DMF (15 mL) was cooled to -40°C and 2,3,5-collidine (1.6 mL, 12 mmol) and methanesulfonyl chloride (0.7 mL, 9.3 mmol) were added. After stirring for 30 min, lithium chloride (0.38 g, 9 mmol, dried with toluene) was added in dry DMF (10 mL). The reaction was stirred for 2.5 h whilst slowly warming to 0°C . The reaction was then stirred at 0°C for 30 min. After this time, the reaction was quenched with ice-cold saturated sodium bicarbonate (20 mL) and extracted with 1:1 diethyl ether:hexane (2×75 mL). The organic layer was washed with saturated ammonium chloride (50 mL) and brine (50 mL) and dried with sodium sulfate. The solvent was removed under reduced pressure and the resulting residue was re-dissolved in dry DMF (25 mL) and cooled to 0°C . Benzenesulfinic acid sodium salt (1 g, 6 mmol) was added and the reaction was stirred for 16 h at 0°C . After this time, the reaction was diluted with ethyl acetate (50 mL). The mixture was washed with 1 M HCl (20 mL), saturated sodium bicarbonate (20 mL), water (20 mL) and brine (20 mL) and dried with sodium sulfate. The solvent was removed under reduced pressure and purification by flash chromatography (85:15, pet. ether: ethyl acetate) afforded the phenylsulfinate diene product (**3.15**, 0.72 g, 1.9 mmol, 63%) as a light orange oil. *R_f* 0.14 (85:15, pet. ether:ethyl acetate). ¹H NMR (400 MHz, CDCl₃) δ 7.90 – 7.85 (2 H, m, 2 $\times$ Ph), 7.64 (1 H, tt, $J = 6.6, 1.0$, 1 $\times$ Ph), 7.54 (2 H, t, $J = 7.9$, 2 $\times$ Ph), 5.27 – 5.17 (2 H, m, H-2 and H-6), 4.55 (1 H, t, $J = 2.8$, H-1'), 4.04 (1 H, d, $J = 11.6$, 1 $\times$ H-8), 3.99 (1 H, d, $J = 11.6$, 1 $\times$ H-8), 3.89 – 3.75 (3 H, m, 2 $\times$ H-1 and 1 $\times$ H-5'), 3.54 – 3.47 (1

H, m, 1 × H-5'), 1.93 (2 H, dd, J = 15.4, 7.8, 2 × H-5), 1.86 – 1.75 (1 H, m, 2 × H-4 and 1 × H-3'), 1.72 (6 H, s, 3 × H-9 and 3 × H-10), 1.71 - 1.67 (1 H, m, 1 × H-2') and 1.63 – 1.49 (4 H, m, 1 × H-2', 1 × H-3' and 2 × H-4'). ¹³C NMR (100 MHz, CDCl₃) δ 145.7 (C-3), 138.8 (Ph), 133.6 (Ph), 132.8 (C-7), 129.0 (2 × Ph), 128.5 (2 × Ph), 128.2 (C-6), 111.2 (C-2), 97.3 (C-1'), 65.1 (C-8), 62.1 (C-5'), 55.8 (C-1), 32.0 (C-4), 30.7 (C-2'), 25.6 (C-5), 25.5 (C-4') 21.7, (C-10), 19.4 (C-3') and 14.2 (C-9). MS (ESI): *m/z* Calcd for C₂₁H₃₀SO₄Na [M+Na]⁺: 401.2: Found 401.2.

The observed characterisation data (¹H and ¹³C) was consistent with that previously reported in the literature.⁹

(2Z,6Z,10E,14E)-1-O-(*tert*-butyldiphenylsilyloxy)-9-benzenesulfonyl-3,7,11,15-tetramethylicosa-2,6,10,14,18-pentaene (3.16)¹⁶

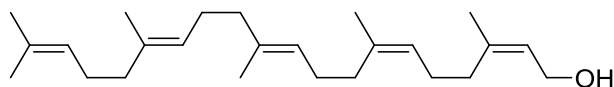


A solution of **3.2** (159 mg, 0.47 mmol) in dry THF (2 mL) was cooled to –78 °C and 1,3-dimethyl-3,4,5,6-tetrahydro-2(1H)-pyrimidinone (0.2 mL, 1.6 mmol) was added, followed by butyl lithium (0.75 mL, 2.5 M in hexane, 1.88 mmol) added dropwise. The mixture was stirred for 15 min, before **3.13** (100 mg, 0.23 mmol) was added in dry THF (2 mL) over 15 min at –78 °C. The reaction was then slowly warmed to room temperature over 16 h. The reaction was quenched with saturated ammonium chloride (5 mL) and extracted with ethyl acetate (2 × 50 mL). The organic layers were dried with sodium sulfate and the solvent was removed under reduced pressure and purification by

flash chromatography (95:5, pet. ether: ethyl acetate) afforded the pentaene product (**3.16**, 110 mg, 0.15 mmol, 65%) as a light yellow oil. *R_f* 0.27 (95:5, pet. ether:ethyl acetate). ¹H NMR (400 MHz, CDCl₃) δ 7.87 – 7.82 (2 H, m, 2 × Ph), 7.71 – 7.66 (4 H, m, 4 × Ph), 7.63 – 7.58 (1 H, m, 1 × Ph), 7.53 – 7.47 (2 H, m, 2 × Ph), 7.46 – 7.36 (6 H, m, 6 × Ph), 5.40 (1 H, t, *J* = 6.3, H-2), 5.12 – 5.00 (3 H, m, H-6, H-14 and H-18), 4.92 (1 H, d, *J* = 10.4, H-10), 4.17 (2 H, d, *J* = 6.6, 2 × H-1), 3.83 (1 H, td, *J* = 11.0, 3.1, H-9), 2.73 (1 H, dd, *J* = 13.3, 2.3, 1 × H-8), 2.46 (1 H, dd, *J* = 13.0, 11.3, 1 × H-8), 2.05 – 1.83 (12 H, m, 2 × H-4, 2 × H-5, 2 × H-12, 2 × H-13, 2 × H-16 and 2 × H-17), 1.69 (6 H, s, 3 × H-24 and 3 × H-25), 1.61 (3 H, s, 3 × H-21), 1.57 (3 H, s, 3 × H-20), 1.54 (3 H, s, 3 × H-22), 1.13 (3 H, s, 3 × H-23) and 1.05 (9 H, s, C(CH₃)₃). ¹³C NMR (100 MHz, CDCl₃) δ 145.3 (C-11), 137.9 (Ph), 136.8 (C-3), 135.7 (C-15), 135.6 (4 × Ph), 134.0 (2 × Ph), 133.4 (Ph), 131.4 (C-19), 130.4 (C-7), 129.6 (2 × Ph), 129.2 (2 × Ph), 128.7 (2 × Ph), 128.1 (C-6), 127.6 (4 × Ph), 125.4 (C-2), 124.2 (C-14), 116.9 (C-10), 63.6 (C-9), 60.7 (C-1), 39.8 (C-12), 39.7 (C-16), 32.0 (C-4), 29.6 (C-8), 26.9 (C(CH₃)₃), 26.7 (C-5), 26.3 (C-13), 26.2 (C-17), 25.7 (C-25), 23.5 (C-22), 23.4 (C-21), 19.2 (C(CH₃)₃), 17.7 (C-20), 16.3 (C-23) and 16.0 (C-24). MS (ESI): *m/z* Calcd for C₄₇H₆₄O₃SSiNa [M+Na]⁺: 759.4: Found 759.4.

The observed characterisation data (¹H and ¹³C) was consistent with that previously reported in the literature.¹⁶

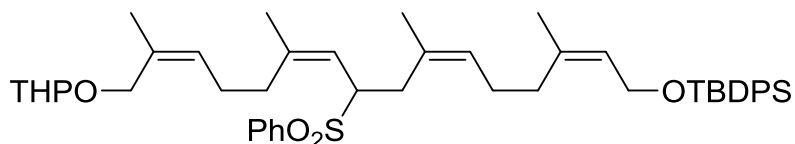
(2Z,6Z,10E,14E)-3,7,11,15-tetramethylicosa-2,6,10,14,18-pentaen-1-ol (3.17)¹⁶



A solution of **3.16** (0.11 g, 0.15 mmol) in dry THF (4 mL) was cooled to 0 °C and tetrabutylammonium fluoride (0.16 mL, 1 M in THF, 0.16 mmol) was added. The reaction was stirred at 0 °C for 16 h before being diluted with diethyl ether (50 mL). The organic layer was washed with brine (2 × 15 mL), dried with sodium sulfate and the solvent removed under reduced pressure. The resulting residue was dissolved in dry THF (2 mL) and (dppp)PdCl₂ (35 mg, 0.06 mmol) were added and the mixture cooled to 0 °C. Super hydride (1.5 mL, 1 M in THF, 1.5 mmol) was added and the reaction was stirred at 0 °C for 3 h. The reaction was quenched with saturated ammonium chloride (1 mL) and extracted with diethyl ether (2 × 25 mL). The organic layer was washed with brine (2 × 15 mL) and dried with sodium sulfate. The solvent was removed under reduced pressure and purification by flash chromatography (95:5 pet. ether:ethyl acetate) afforded the pentaene product (**3.17**, 34 mg, 0.09 mmol, 63%) as a colourless oil. *R_f* 0.45 (CH₂Cl₂). ¹H NMR (400 MHz, CDCl₃) δ 5.45 (1 H, t, *J* = 7.2, H-2), 5.18 – 5.09 (4 H, m, H-6, H-14, H-10 and H-18), 4.10 (2 H, d, *J* = 7.0, 2 × H-1), 2.13 – 1.96 (16 H, m, 2 × H-4, 2 × H-5, 2 × H-8, 2 × H-9, 2 × H-12, 2 × H-13, 2 × H-16 and 2 × H-17), 1.76 (3 H, s, 3 × H-21), 1.71 (3 H, s, 3 × H-22), 1.70 (3 H, s, 3 × H-25), 1.63 (3 H, s, 3 × H-23) and 1.62 (6 H, s, 3 × H-20 and 3 × H-24). ¹³C NMR (100 MHz, CDCl₃) δ 140.0 (C-3), 136.6 (C-7), 135.4 (C-11), 135.3 (C-15), 131.5 (C-19), 124.5, 124.43 (C-2 and C-6), 124.41 (C-18), 124.2 (C-10), 124.1 (C-14), 59.0 (C-1), 39.8 (C-12 and C-16), 32.3 (C-4), 32.0 (C-8), 26.8, 26.64, 26.62, 26.3 (C-5, C-9, C-13 and C-17), 25.8 (C-25), 23.5, 23.4 (C-21 and C-22), 17.7 (C-20), 16.04 and 16.03 (C-23 and C-24). MS (ESI): *m/z* Calcd for C₂₅H₄₂ONa [M+Na]⁺: 381.3: Found 381.4.

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.¹⁶

(2Z,6Z,10Z,14Z)-1-O-(tert-butyldiphenylsilyloxy)-8-benzenesulfonyl-3,7,11,15-tetramethyl-16-(2-tetrahydropyranyloxy)-hexadeca-2,6,10,14-tetraene (3.18)⁹

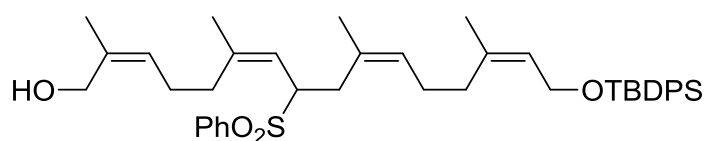


A solution of **3.15** (265 mg, 0.7 mmol, dried in toluene) in dry THF (5 mL) was cooled to $-78\text{ }^{\circ}\text{C}$ and 1,3-dimethyl-3,4,5,6-tetrahydro-2(1H)-pyrimidinone (0.28 mL, 2.3 mmol) was added, followed by butyl lithium (0.56 mL, 2.5 M in hexane, 1.4 mmol) added dropwise. The mixture was stirred for 2 h, before **3.13** (300 mg, 0.7 mmol) was added in dry THF (5 mL) over 90 min. The reaction was slowly warmed to $-30\text{ }^{\circ}\text{C}$ over 90 min, stirred at this temperature for a further 90 min, before being allowed to slowly warm to room temperature over 16 h. The reaction was quenched with saturated ammonium chloride (5 mL) and extracted with ethyl acetate ($2 \times 50\text{ mL}$). The organic layers were dried with sodium sulfate and the solvent was removed under reduced pressure and purification by flash chromatography (85:15, pet. ether: ethyl acetate) afforded the tetraene product (**3.18**, 0.28 g, 0.36 mmol, 52%) as a colourless oil. *R_f* 0.33 (85:15, pet. ether:ethyl acetate). ^1H NMR (400 MHz, CDCl_3) δ 7.83 (2 H, d, $J = 8.1$, $2 \times \text{Ph}$), 7.71 – 7.66 (4 H, m, $4 \times \text{Ph}$), 7.60 (1 H, t, $J = 7.3$, $1 \times \text{Ph}$), 7.50 (2 H, t, $J = 7.5$, $2 \times \text{Ph}$), 7.45 – 7.36 (6 H, m, $6 \times \text{Ph}$), 5.39 (1 H, t, $J = 6.1$, H-2), 5.15 – 5.03 (2 H, m, H-6 and H-14), 4.93 (1 H, d, $J = 10.5$, H-10), 4.54 – 4.52 (1 H, m, H-1'), 4.16 (2 H, d, $J = 6.3$, $2 \times \text{H-1}$), 4.03 (1 H, dd, $J = 11.3$, 4.4 , $1 \times \text{H-16}$), 3.96 (1 H, dd, $J = 11.4$, 3.2 , $1 \times \text{H-16}$), 3.90 – 3.80 (2 H, m, H-9 and $1 \times \text{H-5'}$), 3.56 – 3.48 (1 H, m, $1 \times \text{H-5'}$), 2.64 (1 H,

d, $J = 13.0$, $1 \times \text{H-8}$), 2.44 (1 H , td, $J = 11.4$, 4.7 , $1 \times \text{H-8}$), $1.96 - 1.38$ (14 H , m, $2 \times \text{H-4}$, $2 \times \text{H-5}$, $2 \times \text{H-12}$, $2 \times \text{H-13}$, $2 \times \text{H-2'}$, $2 \times \text{H-3'}$ and $2 \times \text{H-4'}$), 1.71 (3 H , s, $3 \times \text{H-20}$), 1.68 (3 H , s, $3 \times \text{H-17}$), 1.60 (3 H , s, $3 \times \text{H-19}$), 1.55 (3 H , s, $3 \times \text{H-18}$) and 1.04 (9 H , s, $\text{C}(\text{CH}_3)_3$). ^{13}C NMR (100 MHz , CDCl_3) δ 144.8 (C-11), 137.9 (Ph), 136.8 (C-3), 135.6 ($4 \times \text{Ph}$), 134.0 (C-15), 133.5 (Ph), 132.5 (Ph), 130.1 (C-7), 129.6 ($2 \times \text{Ph}$), 129.2 ($2 \times \text{Ph}$), 128.7 ($2 \times \text{Ph}$), 128.4 (C-6), 128.2 (C-14), 127.6 ($4 \times \text{Ph}$), 125.3 (C-2), 117.7 (C-10), 97.3 (C-1'), 65.1 (C-16), 63.2 (C-9), 62.1 (C-5'), 60.7 (C-1), 32.2 (C-4), 32.0 (C-12), 30.7 (C-2'), 30.0 (C-8), 26.9 ($\text{C}(\text{CH}_3)_3$), 26.4 (C-5), 25.5 (C-13 and C-4'), 23.6 (C-18), 23.4 , 23.3 (C-17 and C-19), 21.1 (C-20), 19.5 (C-3') and 19.2 ($\text{C}(\text{CH}_3)_3$). MS (ESI): m/z Calcd for $\text{C}_{47}\text{H}_{64}\text{SSiO}_5\text{Na}$ $[\text{M}+\text{Na}]^+$: 791.4 : Found 791.4 .

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.⁹

(2Z,6Z,10Z,14Z)-1-O-(*tert*-butyldiphenylsilyloxy)-8-benzenesulfonyl-3,7,11,15-tetramethylhexadeca-2,6,10,14-tetraene-1,16-diol (3.19)⁹

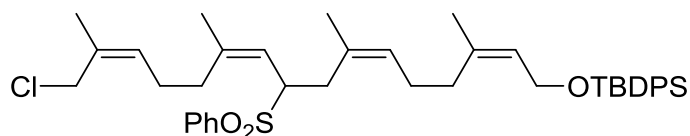


To a solution of **3.18** (340 mg , 0.44 mmol) in isopropanol (10 mL) and diethyl ether (5 mL) was added *p*-toluenesulfonic acid (15 mg , 0.2 equiv) and the reaction was stirred at room temperature for 3 days. After this time the reaction was diluted with diethyl ether (50 mL) and washed with saturated sodium bicarbonate (15 mL) and brine (15 mL). The organic layer was dried with sodium sulfate and the solvent removed under reduced pressure. Purification by flash chromatography ($85:15$, pet. ether:ethyl acetate) afforded

the tetraene product (**3.19**, 220 mg, 0.32 mmol, 73%) as a colourless oil. *R_f* 0.07 (85:15, pet. ether:ethyl acetate). ¹H NMR (400 MHz, CDCl₃) δ 7.86 – 7.81 (2 H, m, 2 × Ph), 7.73 – 7.69 (4 H, m, 4 × Ph), 7.64 – 7.59 (1 H, m, 1 × Ph), 7.52 (2 H, t, *J* = 7.8, 2 × Ph), 7.47 – 7.37 (6 H, m, 6 × Ph), 5.40 (1 H, t, *J* = 5.7, H-2), 5.12 – 5.03 (2 H, m, H-6 and H-14), 4.92 (1 H, d, *J* = 10.5, H-10), 4.16 (2 H, d, *J* = 6.1, 2 × H-1), 4.06 (1 H, d, *J* = 11.9, 1 × H-16), 3.99 (1 H, d, *J* = 11.9, 1 × H-16), 3.85 (1 H, td, *J* = 11.0, 2.9, H-9), 2.61 (1 H, dd, *J* = 13.3, 2.3, 1 × H-8), 2.36 (1 H, dd, *J* = 12.8, 11.3, 1 × H-8), 1.96 – 1.55 (8 H, m, 2 × H-4, 2 × H-5, 2 × H-12 and 2 × H-13), 1.75 (3 H, s, 3 × H-20), 1.67 (3 H, s, 3 × H-17), 1.64 (3 H, s, 3 × H-19), 1.54 (3 H, s, 3 × H-18) and 1.05 (9 H, s, C(CH₃)₃). ¹³C NMR (100 MHz, CDCl₃) δ 145.0 (C-11), 137.7 (Ph), 135.7 (C-3), 135.6 (4 × Ph), 135.1 (C-15), 133.9 (Ph), 133.5 (Ph), 131.0 (Ph), 130.1 (C-7), 129.5 (2 × Ph), 129.1 (2 × Ph), 128.8 (2 × Ph), 128.3 (C-6), 127.6 (4 × Ph), 127.0 (C-14), 125.3 (C-2), 117.5 (C-10), 63.2 (C-9), 61.2 (C-16), 60.6 (C-1), 32.3 (C-4), 31.9 (C-12), 30.3 (C-8), 26.8 (C(CH₃)₃), 26.2 (C-5), 25.5 (C-13), 23.5 (C-18), 23.3 (C-17 and C-19), 21.3 (C-20) and 19.1 (C(CH₃)₃). MS (ESI): *m/z* Calcd for C₄₂H₅₇O₄SSi [M+H]⁺: 685.4; Found 685.4.

The observed characterisation data (¹H and ¹³C) was consistent with that previously reported in the literature.⁹

(2Z,6Z,10Z,14Z)-1-O-(tert-butyldiphenylsilyloxy)-16-chloro-8-benzenesulfonyl-3,7,11,15-tetramethylhexadeca-2,6,10,14-tetraene (3.20)⁹

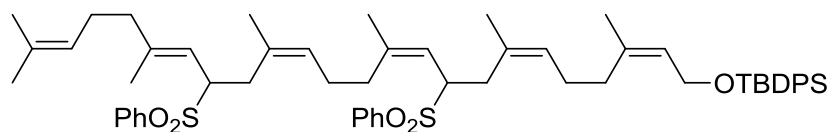


A solution of **3.19** (850 mg, 1.24 mmol) in dry DMF (7 mL) was cooled to $-40\text{ }^{\circ}\text{C}$ and 2,3,5-collidine (0.7 mL, 5.0 mmol) and methanesulfonyl chloride (0.29 mL, 3.7 mmol) were added. After stirring for 30 min, lithium chloride (0.16 g, 3.7 mmol, dried with toluene) was added in dry DMF (6 mL). The reaction was stirred for 2.5 h whilst slowly warming to $0\text{ }^{\circ}\text{C}$. The reaction was stirred at $0\text{ }^{\circ}\text{C}$ for 30 min. After this time, the reaction was quenched with ice-cold saturated sodium bicarbonate (10 mL) and extracted with 1:1 diethyl ether:hexane ($2 \times 30\text{ mL}$). The organic layer was washed with saturated ammonium chloride (10 mL) and brine (10 mL) and dried with sodium sulfate. The solvent was removed under reduced pressure and purification by flash chromatography (85:15, pet. ether:ethyl acetate) afforded the chloro tetraene product (**3.20**, 427 mg, 0.6 mmol, 49%) as a colourless oil. R_f 0.38 (85:15, pet. ether:ethyl acetate). ^1H NMR (400 MHz, CDCl_3) δ 7.85 (2 H, d $J = 7.5$, $2 \times \text{Ph}$), 7.70 (4 H, d, $J = 6.4$, $4 \times \text{Ph}$), 7.62 (1 H, t, $J = 7.3$, $1 \times \text{Ph}$), 7.52 (2 H, t, $J = 7.6$, $2 \times \text{Ph}$), 7.47 – 7.36 (6 H, m, $6 \times \text{Ph}$), 5.40 (1 H, t, $J = 6.1$, H-2), 5.17 (1 H, t, $J = 6.3$, H-14), 5.10 – 5.03 (1 H, m, H-6), 4.98 (1 H, d, $J = 10.5$, H-10), 4.17 (2 H, d, $J = 6.3$, $2 \times \text{H-1}$), 3.98 (1 H, d, $J = 10.8$, $1 \times \text{H-16}$), 3.94 (1 H, d, $J = 10.8$, $1 \times \text{H-16}$), 3.83 (1 H, td, $J = 10.8$, 2.6, H-9), 2.63 (1 H, dd, $J = 13.1$, 2.0, $1 \times \text{H-8}$), 2.41 (1 H, dd, $J = 13.0$, 11.4, $1 \times \text{H-8}$), 1.97 – 1.57 (8 H, m, $2 \times \text{H-4}$, $2 \times \text{H-5}$, $2 \times \text{H-12}$ and $2 \times \text{H-13}$), 1.78 (3 H, s, $3 \times \text{H-20}$), 1.68 (3 H, s, $3 \times \text{H-17}$), 1.65 (3 H, s, $3 \times \text{H-19}$), 1.55 (3 H, s, $3 \times \text{H-18}$) and 1.05 (9 H, s, $\text{C}(\text{CH}_3)_3$). ^{13}C NMR (100 MHz, CDCl_3) δ 144.6 (C-11), 137.9 (Ph), 135.8 (C-3), 135.6 ($4 \times \text{Ph}$), 134.0

(2 × Ph), 133.6 (Ph), 132.1 (C-15), 130.2 (C-7), 129.9 (C-14), 129.6 (2 × Ph), 129.2 (2 × Ph), 128.8 (2 × Ph), 128.4 (C-6), 127.6 (4 × Ph), 125.4 (C-2), 117.5 (C-10), 63.4 (C-9), 60.7 (C-1), 43.5 (C-16), 32.0 (C-4), 31.9 (C-12), 30.3 (C-8), 26.9 (C(CH₃)₃), 26.4 (C-5), 25.9 (C-13), 23.7 (C-18), 23.37 (C-17), 23.35 (C-19), 21.6 (C-20) and 19.1 (C(CH₃)₃). MS (ESI): *m/z* Calcd for C₄₂H₅₅O₃ClSSiNa [M+Na]⁺: 725.3: Found 725.3.

The observed characterisation data (¹H and ¹³C) was consistent with that previously reported in the literature.⁹

(2Z,6Z,10Z,14Z,18E)-1-O-(tert-butylidiphenylsilyloxy)-9,17-di(benzenesulfonyl)-3,7,11,15,19,23-hexamethylicosakaitetra-2,6,10,14,18,22-hexaene (3.21)⁹

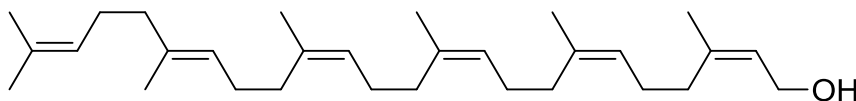


A solution of **3.1** (167 mg, 0.6 mmol) in dry THF (2 mL) was cooled to −78 °C and 1,3-dimethyl-3,4,5,6-tetrahydro-2(1H)-pyrimidinone (0.25 mL, 2.1 mmol) was added, followed by butyl lithium (1.0 mL, 2.5 M in hexane, 2.4 mmol) added dropwise. The mixture was stirred for 15 min, before **3.20** (210 mg, 0.3 mmol) was added in dry THF (2 mL) over 15 min at −78 °C. The reaction was then slowly warmed to room temperature over 16 h. The reaction was quenched with saturated ammonium chloride (1 mL) and extracted with 1:1 hexane:diethyl ether (2 × 10 mL). The organic layers were dried with sodium sulfate and the solvent was removed under reduced pressure and purification by flash chromatography (95:5, pet. ether: ethyl acetate) afforded the hexaene product (**3.21**, 103 mg, 0.11 mmol, 36%) as a light yellow oil. *R_f* 0.49 (7:3, pet. ether:ethyl acetate). ¹H NMR (400 MHz, CDCl₃) δ 7.89 - 7.81 (4 H, m, 4 × Ph), 7.72 -

7.67 (4 H, m, 4 × Ph), 7.65 - 7.58 (2 H, m, 2 × Ph), 7.55 - 7.48 (4 H, m, 2 × Ph), 7.46 - 7.35 (6 H, m, 6 × Ph), 5.40 (1 H, t, J = 6.3, H-2), 5.06 (1 H, t, J = 6.1, H-6), 5.02 - 4.86 (4 H, m, H-10, H-14, H-18 and H-22), 4.17 (2 H, d, J = 6.3, 2 × H-1), 3.83 (2 H, td, J = 11.0, 3.1, H-9 and H-17), 2.77 - 2.61 (2 H, m, 1 × H-8 and 1 × H-16), 2.50 - 2.39 (2 H, m, 1 × H-8 and 1 × H-16), 1.97 - 1.77 (12 H, m, 2 × H-4, 2 × H-5, 2 × H-12, 2 × H-13, 2 × H-20 and 2 × H-21), 1.74 - 1.48 (18 H, 3 × H-24, 3 × H-25, 3 × H-26, 3 × H-27, 3 × H-28 and 3 × H-30), 1.11 - 1.08 (3 H, m, 3 × H-29) and 1.05 (9 H, s, C(CH₃)₃). ¹³C NMR (100 MHz, CDCl₃) δ 145.0 (C-11), 144.6 (C-19), 137.8 (Ph), 137.7 (Ph), 136.7 (C-3), 135.5 (4 × Ph), 133.9 (2 × Ph), 133.5 (Ph), 133.4 (C-23), 130.7 (Ph), 130.6 (C-15), 130.0 (C-7), 129.5 (2 × Ph), 129.1 (4 × Ph), 128.7 (2 × Ph), 128.6 (2 × Ph), 128.3 (C-6), 127.5 (4 × Ph), 127.4 (C-14), 125.2 (C-2), 123.3 (C-22), 117.7, 116.9 (C-10 and C-18), 63.3, 63.1 (C-9 and C-17), 61.6 (C-1), 39.6 (C-20), 31.9, 31.8 (C-4 and C-12), 30.0 (C-16), 29.4 (C-8), 26.8 (C(CH₃)₃), 26.3 (C-5), 26.0 (C-13), 25.6 (C-30), 25.5 (C-21), 23.4, 23.3, 23.2, 23.1 (C-25, C-26, C-27 and C-28) 19.1 (C(CH₃)₃), 17.6 (C-24) and 16.1 (C-29). MS (ESI): *m/z* Calcd for C₅₈H₇₆O₅S₂SiNa [M+Na]⁺: 967.5: Found 967.5.

The observed characterisation data (¹H and ¹³C) was consistent with that previously reported in the literature.⁹

(2Z,6Z,10Z,14Z,18E)-3,7,11,15,19,23-hexamethylicosakaitetra-2,6,10,14,18,22-hexaen-1-ol (3.22)⁹

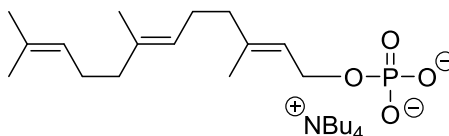


A solution of **3.21** (233 mg, 0.25 mmol) in dry THF (10 mL) was cooled to 0 °C and tetrabutylammonium fluoride (0.27 mL, 1 M in THF, 0.27 mmol) was added. The reaction was stirred at 0 °C for 16 h before being diluted with diethyl ether (50 mL). The organic layer was washed with brine (2 × 15 mL), dried with sodium sulfate and the solvent removed under reduced pressure. The resulting residue was dissolved in dry THF (5 mL) and (dppp)PdCl₂ (59 mg, 0.4 mmol) were added and the mixture cooled to 0 °C. Super hydride (2.5 mL, 1 M in THF, 2.5 mmol) was added and the reaction was stirred at 0 °C for 3 h. The reaction was quenched with saturated ammonium chloride (1 mL) and extracted with diethyl ether (2 × 25 mL). The organic layer was washed with brine (2 × 15 mL) and dried with sodium sulfate. The solvent was removed under reduced pressure and purification by flash chromatography (95:5, pet. ether:ethyl acetate) afforded the hexaene product (**3.22**, 50 mg, 0.11 mmol, 47%) as a colourless oil. *R*_f 0.2 (95:5, pet. ether:ethyl acetate). ¹H NMR (400 MHz, CDCl₃) δ 5.45 (1 H, t, *J* = 7.0 Hz), 5.17 – 5.07 (5 H, m, H-6, H-10, H-14, H-18 and H-22), 4.10 (2 H, d, *J* = 7.2, 2 × H-1), 2.13 – 1.95 (20 H, m, 2 × H-4, 2 × H-5, 2 × H-8, 2 × H-9, 2 × H-12, 2 × H-13, 2 × H-16, 2 × H-17, 2 × H-20 and 2 × H-21), 1.75 (3 H, s, 3 × H-25), 1.70 (12 H, s, 3 × H-26, 3 × H-27, 3 × H-28 and 3 × H-30), 1.61 (3 H, s, 3 × H-24). ¹³C NMR (100 MHz, CDCl₃) δ 139.8 (C-3), 136.0 (C-7), 135.4, 135.1, 134.8 (C-11, C-15 and C-19), 131.3 (C-23), 124.9, 124.8, 124.5, 124.4 (C-6, C-10, C-14 and C-18), 124.3 (C-2), 124.1 (C-22), 59.0 (C-1), 39.7 (C-20), 32.20, 32.18, 32.14 (C-4, C-8 and C-12), 31.9 (C-16), 26.7

(C-21), 26.6, 26.4, 26.33, 26.28 (C-5, C-9, C-13 and C-17), 25.7 (C-30), 23.4, 23.3 (C-25, C-26, C-27 and C-28), 17.6 (C-24) and 15.9 (C-29). MS (ESI): m/z Calcd for $C_{30}H_{50}ONa$ $[M+Na]^+$: 449.4: Found 449.4.

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.⁹

***trans, trans* Farnesol phosphate (3.23)²⁶**

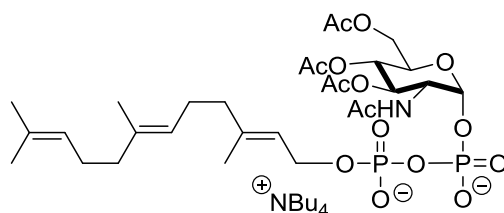


To solution of *trans, trans* farnesol (50 mg, 0.22 mmol) and $Bu_4NPO_4H_2$ (103 mg, 0.30 mmol) in dry dichloromethane (6 mL) was added trichloroacetonitrile (0.11 mL, 1.1 mmol) and reaction was stirred at room temperature for 1 h. The solvent was removed under reduced pressure and the resulting residue was dissolved in THF (3 mL) and ammonia hydroxide (1 mL, 35% NH_3 in water). This was stirred at room temperature for 30 min before 1:1 methanol:toluene (28 mL) was added and the mixture was stirred for a further 20 min. The mixture was filtered and the solvent removed under reduced pressure. The resulting residue was washed with pet. ether (10 mL) and re-dissolved in methanol (10 mL) and filtered. Dowex 50WX8 (NH_4^+ form, 100 mg) was added and the mixture stirred for 30 min, after which the pH was approx. 7. The resin was filtered off and solvent removed under reduced pressure. Purification by flash chromatography (1:2:4, water:isopropanol:ethyl acetate) afforded the phosphate product (**3.23**, 36 mg, approx 1.3:1 Bu_4N^+ :Product by NMR, 0.06 mmol, 29%) as a white solid. R_f 0.29 (1:2:4,

water:isopropanol:ethyl acetate). ^1H NMR (400 MHz, MeOD) δ 5.37 (1 H, td, J = 6.6, 1.0, H-2), 5.12 – 5.03 (2 H, m, H-6 and H-10), 4.37 (1 H, t, J = 6.7, H-1), 2.13 – 1.90 (8 H, m, $2 \times$ H-4, $2 \times$ H-5, $2 \times$ H-8 and $2 \times$ H-9), 1.65 (3 H, s, $3 \times$ H-12), 1.63 (6 H, s, $3 \times$ H-13 and $3 \times$ 14) and 1.56 (3 H, s, $3 \times$ H-15). NBu_4^+ : 3.26 (8 H, m, $8 \times$ NCH_2), 2.13 – 2.02 (8 H, m, $8 \times$ NCH_2CH_2), 1.42 (8 H, quad, J = 7.4, $8 \times$ CH_2CH_3) and 1.03 (12 H, t, J = 7.4 $12 \times$ CH_3). ^{13}C NMR (100 MHz, MeOD) δ 140.1 (C-3), 136.3, 132.2 (C-7 and C-11), 125.5, 125.3 (C-6 and C-10), 123.1 (C-2), 63.1 (C-1), 41.0, 40.8 (C-8), 27.9, 27.5 (C-5 and C-9), 26.1 (C-15), 18.0 (C-12), 16.6 (C-13) and 16.2 (C-14). NBu_4^+ : 50.0 (NCH_2), 24.9 (NCH_2CH_2), 20.9 (CH_2CH_3) and 14.1 (CH_3). ^{31}P NMR (162 MHz, MeOD) δ 5.3. MS (ESI): m/z Calcd for $\text{C}_{15}\text{H}_{26}\text{O}_4\text{P}$ $[\text{M}-\text{H}]^-$: 301.2: Found 301.1.

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.²⁶

P^1 -2-*N*-acetamido-2-deoxy-3,4,6-tri-*O*-acetyl- α -D-glucopyranosyl P^2 -(2*E*,6*E*)-3,7,11-trimethyldodac-2,6,10-trien-1-yl diphosphate (3.24)



To a solution of *trans, trans* Farnesol (50 mg, 0.22 mmol) and $\text{Bu}_4\text{NPO}_4\text{H}_2$ (120 mg, 0.36 mmol) in dry dichloromethane (6 mL) was added trichloroacetonitrile (0.1 mL, 0.67 mmol). After stirring at room temperature for 1 h, the solvent was removed under reduced pressure and the resulting residue was re-dissolved in THF (3 mL) and

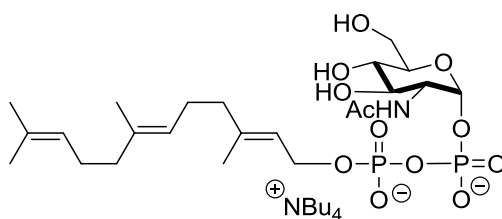
ammonium hydroxide (1 mL, 35% NH₃ in water). This was stirred at room temperature for 30 min before 1:1 methanol:toluene (28 mL) was added and the mixture was stirred for a further 20 min. The mixture was filtered and the solvent removed under reduced pressure. The resulting residue was washed with pet. ether (10 mL) and re-dissolved in methanol (10 mL) and filtered. Dowex 50WX8 (NH₄⁺ form, 100 mg) was added and the mixture stirred for 30 min, after which the pH was approx. 7. The resin was filtered off and solvent removed under reduced pressure to afford the crude phosphate.

Half the crude phosphate was taken through to the 2nd stage.

To a solution of the crude phosphate in dry dichloromethane (2 mL) was added carbonyldiimidazole (60 mg, 0.37 mmol) and the reaction was stirred at room temperature for 2 h. After this time, methanol (0.2 mL) was added and the reaction was stirred for a further 30 min. The solvent was removed under reduced pressure and 2-acetamido-3,4,6-O-acetyl- α -D-glucopyranosyl-1-phosphate (48 mg, 0.11 mmol) and tetrazole (0.26 mL, 0.45 M in MeCN, 0.15 mmol) were added in dry DMF (2 mL). The reaction was stirred at room temperature for 3 days before the solvent was removed under reduced pressure. Purification by reverse phase C-18 resin (95:5, water:MeCN to 75:20:5, water:isopropanol:MeCN) afforded the pyrophosphatelinked lipid sugar product (**3.24**, 28 mg, approx 3:1 Bu₄N⁺:Product by NMR, 0.02 mmol, 18%) as a colourless foam. *R_f* 0.34 (75:20:5, water:isopropanol:MeCN, C-18 resin). ¹H NMR (400 MHz, MeOD) δ 5.61 (1 H, dd, *J* = 7.4, 3.1, H-1), 5.47 (1 H, t, *J* = 5.5, H-2'), 5.35 (1 H, t, *J* = 9.5, H-3), 5.15 – 5.07 (3 H, m, H-5, H-6' and H-10'), 4.60 (2 H, t, *J* = 6.3, 2 $\times$ H-1'), 4.42 (1 H, dt, *J* = 10.3, 2.5, H-4), 4.37 (1 H, dd, *J* 12.5, 3.3, 1 $\times$ H-6), 4.31 (1 H, dd,

$J = 10.5, 3.0, \text{H-2}), 4.13 (1 \text{ H, dd, } J = 12.5, 2.2, 1 \times \text{H-6}), 2.14 - 1.93 (8 \text{ H, m, } 2 \times \text{H-4}', 2 \times \text{H-5}', 2 \times \text{H-8}', 2 \times \text{H-9}'), 2.04 (3 \text{ H, s, OAc}), 2.00 (3 \text{ H, s, NHAc}), 1.98 (3 \text{ H, s, OAc}), 1.92 (3 \text{ H, s, OAc}), 1.66 (3 \text{ H, s, } 3 \times \text{H-14}'), 1.65 (3 \text{ H, s, } 3 \times \text{H-13}'), 1.60 (3 \text{ H, s, } 3 \times \text{H-15}') \text{ and } 1.59 (3 \text{ H, s, } 3 \times \text{H-12}'). \text{NBu}_4^+$: $3.26 (8 \text{ H, m, } 8 \times \text{NCH}_2), 2.13 - 2.02 (8 \text{ H, m, } 8 \times \text{NCH}_2\text{CH}_2), 1.42 (8 \text{ H, quad, } J = 7.4, 8 \times \text{CH}_2\text{CH}_3) \text{ and } 1.03 (12 \text{ H, t, } J = 7.4, 12 \times \text{CH}_3).$ ^{13}C NMR (100 MHz, MeOD) δ 174.5 (CO), 172.7 (CO), 171.8 (CO), 171.6 (NCO), 140.2 (C-3'), 136.2, 132.2 (C-7' and C-11'), 125.6, 125.5 (C-6' and C-10'), 123.2 (C-2'), 95.9 (C-1), 73.5 (C-3), 70.0 (C-5), 69.8 (C-4), 64.0 (C-1'), 63.1 (C-6), 53.4 (C-2), 41.0, 40.9 (C-4' and C-8'), 27.8, 27.7 (C-5' and C-9'), 26.1 (C-15'), 23.0 (NHAc), 20.8, (3 $\times$ OAc), 17.9 (C-12'), 16.8 (C-13') and 16.2 (C-14'). NBu_4^+ : 59.6 (NCH₂), 24.9 (NCH₂CH₂), 20.9 (CH₂CH₃) and 14.1 (CH₃). ^{31}P NMR (162 MHz, MeOD) δ -10.7 and -13.7. HRMS (ESI): m/z Calcd for C₂₉H₄₆NO₁₅P₂ [M-H]⁻: 710.2348; Found 710.2357.

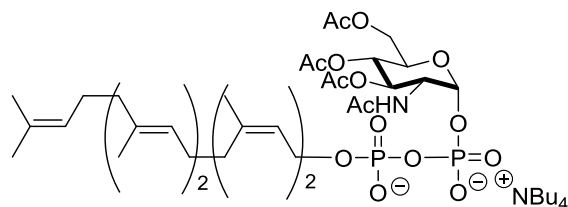
***P*¹-2-*N*-acetamido-2-deoxy- α -D-glucopyranosyl *P*²-(2*E*,6*E*)-3,7,11-trimethyldodac-2,6,10-trien-1-yl diphosphate (3.25)**



A solution of **3.24** (28 mg, 3:1, Bu₄N⁺:**3.24**, 0.02 mmol) in dry methanol (2 mL) was cooled to 0 °C and freshly prepared in NaOMe (1.2 mL, 50 mM, 0.06 mmol) was added and the reaction was stirred at 0 °C for 2 h. After this time, the reaction was quenched by the addition of Dowex 50WX8 (NH₄⁺ form, 100 mg) and the reaction was stirred for 15 min. The Dowex was removed by filtration and the solvent removed under reduced

pressure. Purification by reverse phase C-18 resin (95:5, 2% NH_4HCO_3 in water:MeCN to 1:1, 2% NH_4HCO_3 in water:MeCN) afforded the pyrophosphatelinked lipid sugar product (**3.24**, 12 mg, approx 2:1 Bu_4N^+ :Product by NMR, 2.4 μmol , 12%) as an orange foam. *Rf* 0.54 (1:1, 2% NH_4HCO_3 in water:MeCN, C18-resin). ^1H NMR (400 MHz, MeOD) δ 5.51 (1 H, dd, $J = 7.4$, 3,4, H-1), 5.42 (1 H, t, $J = 6.9$, H-2'), 5.14 – 5.04 (2 H, m, H-6' and H-10'), 4.53 (2 H, t, $J = 6.4$, $2 \times \text{H-1'}$), 3.99 – 3.93 (2 H, m, H-2 and H-5), 3.85 (1 H dd, $J = 12.0$, 2.5, H-6a), 3.75 (1 H, t, $J = 9.6$, H-3), 3.64 (1 H, dd, $J = 11.7$, 5.6, H-6b), 3.33 (1 H, d, $J = 9.8$, H-4), 2.13 – 1.92 (8 H, m, $2 \times \text{H-4'}$, $2 \times \text{H-5'}$, $2 \times \text{H-8'}$, $2 \times \text{H-9'}$), 2.03 (3 H, s, NHAc), 1.68 (3 H, s, $3 \times \text{H-14'}$), 1.62 (3 H, s, $3 \times \text{H-13'}$) and 1.58 (6 H, s, $3 \times \text{H-12'}$ and $3 \times \text{H-15'}$). NBu_4^+ : 3.26 (8 H, m, $8 \times \text{NCH}_2$), 2.13 – 2.02 (8 H, m, $8 \times \text{NCH}_2\text{CH}_2$), 1.42 (8 H, quad, $J = 7.4$, $8 \times \text{CH}_2\text{CH}_3$) and 1.03 (12 H, t, $J = 7.4$, $12 \times \text{CH}_3$). ^{13}C NMR (100 MHz, MeOD) δ 173.1 (NCO), 140.7 (C-3'), 136.3, 132.2 (C-7' and C-11'), 125.6, 125.4 (C-6' and C-10'), 122.8 (C-2'), 96.3 (C-1), 75.1 (C-5), 73.3 (C-3), 72.5 (C-4), 64.0 (C-1'), 63.1 (C-6), 55.6 (C-2), 41.0, 40.8 (C-4' and C-8'), 28.0, 27.7 (C-5' and C-9'), 26.1 (C-15'), 23.2 (NHAc), 17.9 (C-12'), 16.8 (C-13') and 16.2 (C-14'). NBu_4^+ : 59.6 (NCH_2), 24.9 (NCH_2CH_2), 20.9 (CH_2CH_3) and 14.1 (CH_3). ^{31}P NMR (162 MHz, MeOD) δ -10.3 and -12.6. HRMS (ESI): *m/z* Calcd for $\text{C}_{23}\text{H}_{40}\text{NO}_{12}\text{P}_2$ $[\text{M-H}]^-$: 584.2020: Found 584.2029.

***P*¹-2-*N*-acetamido-2-deoxy-3,4,6-tri-*O*-acetyl- α -D-glucopyranosyl (2*Z*,6*Z*,10*E*,14*E*)-3,7,11,15-tetramethylicos-2,6,10,14,18-pentaen-1-yl diphosphate (3.26)**

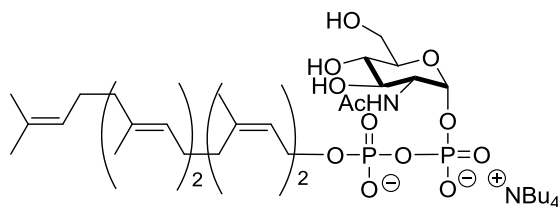


To a solution of **3.17** (10 mg, 0.028 mmol) and Bu₄NPO₄H₂ (15 mg, 0.04 mmol) in dry dichloromethane (1 mL) was added trichloroacetonitrile (0.01 mL, 0.08 mmol). After stirring at room temperature for 1 h, the solvent was removed under reduced pressure and the resulting residue was re-dissolved in THF (1 mL) and ammonium hydroxide (1.5 mL, 35% NH₃ in water). This was stirred at room temperature for 30 min before 1:1 methanol:toluene (4 mL) was added and the mixture was stirred for a further 20 min. The mixture was filtered and the solvent removed under reduced pressure. The resulting residue was washed with pet. ether (2 mL) and re-dissolved in methanol (5 mL) and filtered. Dowex 50WX8 (NH₄⁺ form, 100 mg) was added and the mixture stirred for 30 min, after which the pH was approx. 7. The resin was filtered off and solvent removed under reduced pressure to afford the crude phosphate.

To a solution of the crude phosphate in dry dichloromethane (1 mL) was added carbonyldiimidazole (23 mg, 0.14 mmol) and the reaction was stirred at room temperature for 2 h. After this time, methanol (0.1 mL) was added and the reaction was stirred for a further 30 min. The solvent was removed under reduced pressure and 2-acetamido-3,4,6-*O*-acetyl- α -D-glucopyranosyl-1-phosphate (17 mg, 0.04 mmol) and tetrazole (0.1 mL, 0.45 M in MeCN, 0.05 mmol) were added in dry DMF (0.5 mL). The reaction was stirred at room temperature for 3 days before the solvent was removed

under reduced pressure. Purification by reverse phase C-18 resin (95:5, water:MeCN to 75:20:5, water:isopropanol:MeCN) afforded the pyrophosphate linked lipid sugar product (**3.26**, 7 mg, approx 1:2 Bu₄N⁺:Product by NMR, 7 μmol, 26%) as a colourless foam. *R_f* 0.15 (75:20:5, water:isopropanol:MeCN, C-18 resin). ¹H NMR (500 MHz, MeOD) δ 5.64 – 5.60 (1 H, m, H-1), 5.54 (1 H, t, *J* = 6.2, H-2'), 5.32 (1 H, t, *J* = 9.6, H-3), 5.16 – 5.06 (5 H, m, H-5, H-6', H-10', H-14' and H-18'), 4.55 – 4.49 (3 H, m, H-4 and 2 × H-1'), 4.40 - 4.28 (3 H, m, H-2, H-4 and 1 × H-6), 4.16 (1 H, dd, *J* = 12.3, 1.6, 1 × H-6), 2.15 – 1.94 (16 H, m, 2 × H-4', 2 × H-5', 2 × H-8', 2 × H-9', 2 × H-12', 2 × H-13', 2 × H-16' and 2 × H-17'), 2.05 (3 H, s, OAc), 2.00 (3 H, s, NHAc), 1.98 (3 H, s, OAc), 1.95 (3 H, s, OAc), 1.74 (3 H, s, 3 × H-21'), 1.68 (3 H, s, 3 × H-22'), 1.67 (3 H, s, 3 × H-25'), 1.62 (3 H, s, 3 × H-23') and 1.60 (6 H, s, 3 × H-20' and 3 × H-24'). NBu₄⁺: 3.26 (8 H, m, 8 × NCH₂), 2.13 – 2.02 (8 H, m, 8 × NCH₂CH₂), 1.42 (8 H, quad, *J* = 7.4, 8 × CH₂CH₃) and 1.03 (12 H, t, *J* = 7.4 12 × CH₃). ¹³C NMR (125 MHz, MeOD) δ 174.2 (CO), 172.7 (CO), 172.0 (CO), 171.5 (NCO), 141.2 (C-3'), 136.9 (C-7'), 136.1, 136.0, 131.0 (C-11', C-15' and C-19'), 126.9, 125.9, 125.6 (C-6', C-10', C-14' and C-18'), 123.3 (C-2'), 96.0 (C-1), 72.9 (C-3), 70.1 (C-5), 69.8 (C-4), 64.0 (C-1'), 63.0 (C-6), 53.3 (C-2), 44.0 (C-4', C-8', C-12' and C-16'), 28.0, 27.8, 27.5 (C-5', C-9' and C-13'), 26.1 (C-25'), 22.9 (NHAc), 23.9 (C-21', C-22' and C-23') 20.8, (3 × OAc), 17.9 (C-20') and 16.3 (C-24'). NBu₄⁺: 59.6 (NCH₂), 24.9 (NCH₂CH₂), 20.9 (CH₂CH₃) and 14.1 (CH₃). ³¹P NMR (162 MHz, MeOD) δ –9.1 and –11.9. HRMS (ESI): *m/z* Calcd for C₃₉H₆₂NO₁₅P₂ [M–H][–]: 846.3600: Found 846.3594.

***P*¹-2-*N*-acetamido-2-deoxy- α -D-glucopyranosyl (2*Z*,6*Z*,10*E*,14*E*)-3,7,11,15-tetramethylcosa-2,6,10,14,18-pentaen-1-yl diphosphate (3.27)**



A solution of **3.26** (7 mg, 1:2, Bu₄N⁺:**3.26**, 7 μ mol) in dry methanol (1 mL) was cooled to 0 °C and freshly prepared in NaOMe (1 mL, 220 mM, 0.22 mmol) was added and the reaction was stirred at 0 °C for 1.5 h. After this time, the reaction was quenched by the addition of Dowex 50WX8 (NH₄⁺ form, 100 mg) and the reaction was stirred for 15 min. The Dowex was removed by filtration and the solvent removed under reduced pressure. Purification by reverse phase C-18 resin (95:5, 2% NH₄HCO₃ in water:MeCN to 50:50, 2% NH₄HCO₃ in water:MeCN) afforded the pyrophosphate linked lipid sugar product (**3.27**, 0.7 mg, approx 1:1.6 Bu₄N⁺:Product by NMR, 6.8 μ mol, 98%) as an orange foam. *R_f* 0.43 (1:1, 2% NH₄HCO₃ in water:MeCN, C18-resin). ¹H NMR (500 MHz, MeOD) δ 5.54 (1 H, dd, *J* = 6.6, 2.7, H-1), 5.46 (1 H, t, *J* = 6.3, H-2'), 5.17 – 5.07 (4 H, m, H-6', H-10', H-14' and H-18'), 4.54 (2 H, t, *J* = 6.2, 2 $\times$ H-1'), 4.01 – 3.96 (2 H, m, H-2 and H-5), 3.89 (1 H, dd, *J* = 12.0, 2.1, 1 $\times$ H-6), 3.77 (1 H, t, *J* = 9.3, H-3), 3.66 (1 H, dd, *J* = 12.0, 5.8, 1 $\times$ H-6), 3.36 (1 H, d, *J* = 9.3, H-4), 2.16 – 1.94 (16 H, m, 2 $\times$ H-4', 2 $\times$ H-5', 2 $\times$ H-8', 2 $\times$ H-9', 2 $\times$ H-12', 2 $\times$ H-13', 2 $\times$ H-16' and 2 $\times$ H-17'), 2.06 (3 H, s, NHAc), 1.74 (3 H, s, 3 $\times$ H-21'), 1.68 (3 H, s, 3 $\times$ H-22'), 1.67 (3 H, s, 3 $\times$ H-25'), 1.62 (3 H, s, 3 $\times$ H-23') and 1.60 (6 H, s, 3 $\times$ H-20' and 3 $\times$ H-24'). NBu₄⁺: **3.26** (8 H, m, 8 $\times$ NCH₂), 2.13 – 2.02 (8 H, m, 8 $\times$ NCH₂CH₂), 1.42 (8 H, quad, *J* = 7.4, 8 $\times$ CH₂CH₃) and 1.03 (12 H, t, *J* = 7.4 12 $\times$ CH₃). ¹³C NMR (125 MHz, MeOD) δ 174.5 (NCO), 140.3 (C-3'), 136.5 (C-7'), 136.0, 135.9, 132.0 (C-11', C-15' and C-19'), 125.9,

125.5 (C-6', C-10', C-14' and C-18'), 123.7 (C-2'), 96.2 (C-1), 74.9 (C-5), 73.2 (C-3), 72.2 (C-4), 63.6 (C-1'), 62.9 (C-6), 55.5 (C-2), 40.9 (C-4', C-8', C-12' and C-16'), 27.8, 27.61, 27.56 (C-5', C-9' and C-13'), 25.9 (C-25'), 23.72, 23.69, 23.60 (C-21', C-22' and C-23'), 23.0 (NHAc), 17.7 (C-20') and 16.1 (C-24'). NBu₄⁺: 59.6 (NCH₂), 24.9 (NCH₂CH₂), 20.9 (CH₂CH₃) and 14.1 (CH₃). ³¹P NMR (162 MHz, MeOD) δ -9.5 and -12.5. HRMS (ESI): *m/z* Calcd for C₃₃H₅₅NO₁₂P₂ [M-H]⁻: 720.3280: Found 720.3285.

3.10 - Glycosylation Reactions

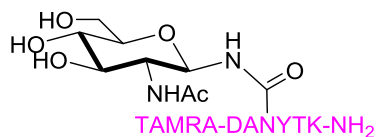
3.10.1 - Investigating the Activity of PglB

To the TAMRA-peptide (**3.28**, 0.5 μ L 1 mM, 0.5 nmol) in 10 mM HEPES, 1 mM MnCl_2 , 1% Triton X100, pH 7.5 (16 μ L) was added WEZ₄-GlcNAc (2 μ L 2 mM, 4 nmol) and either the newly expressed PglB (0.2 μ L) or PglB of a known concentration (0.44 μ M, 0.2 μ L, 0.088 pmol). The reaction was shaken at room temperature for 16 h before analysis by HPLC (Waters 2795 Separations Module using a Jupiter 4U Proteo 90 Å 250 $\times$ 4.60 mm 4 micron column and using a Jasco FP-920 Intelligent Fluorescent Detector, 542 nm excitation, 568 emission).

3.10.2 – Investigating the Effect of Different Lipid Lengths on the Reactivity of the PglB Substrate

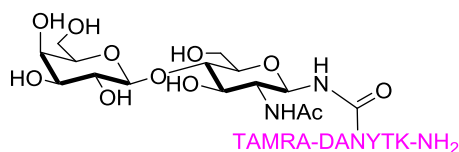
To the TAMRA-peptide (**3.28**, 1 μ L 1 mM, 0.5 nmol) in 10 mM HEPES, 1 mM MnCl_2 , 1% Triton X100, pH 7.5 (33 μ L) was added PglB (0.44 μ M, 4 μ L, 1.76 pmol) and either WEZ₄-GlcNAc (2 μ L, 2 mM, 4 nmol), WE₂Z₂-GlcNAc (**3.27**, 2 μ L, 2 mM, 4 nmol), WZ₃-GlcNAc (1 μ L, 4 mM, 4 nmol or 4 μ L, 4 mM, 16 nmol) or WE₂-GlcNAc (2 μ L, 2 mM, 4 nmol or 2 μ L, 20 mM, 40 nmol or 2 μ L, 200 mM, 400 nmol). The reactions were shaken at room temperature and samples removed at 20 h and 44 h and analysed by HPLC (Waters 2795 Separations Module using a Jupiter 4U Proteo 90 Å 250 $\times$ 4.60 mm 4 micron column and using a Jasco FP-920 Intelligent Fluorescent Detector, 542 nm excitation, 568 emission).

3.10.3 - N-Glycosylation of TAMRA-DANYTK with PglB (3.29)



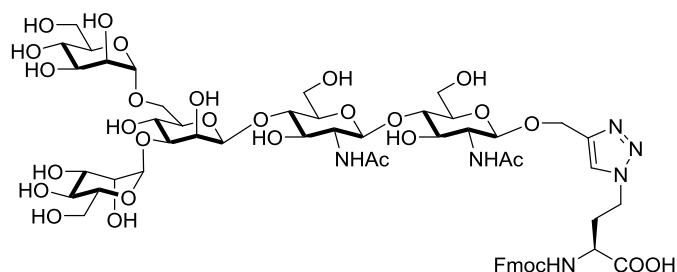
To the TAMRA-peptide (**3.28**, 25 μ L 1 mM, 25 nmol) in 10 mM HEPES, 1 mM MnCl₂, 1% Triton X100, pH 7.5 (325 μ L) was added PglB (100 μ L 0.44 μ M, 44 pmol) and WEZ₄-GlcNAc (50 μ L 2 mM, 100 nmol). The reaction was shaken at room temperature overnight. After this time the solution was diluted with water (1.5 mL) and PglB was removed by Vivaspin 10000 MWCO. This process was repeated twice more to ensure complete removal of the PglB. The resulting 5 mL solution was split into 5 $\times$ 1 mL solutions. To each 1 mL solution was added 10 mL 3:2 CHCl₃:MeOH and 2 mL 4 mM MgCl₂. The aqueous layer was extracted and the organic layer was washed with 6 mL of the top layer of 6:4:2 CHCl₃:MeOH:4 mM MgCl₂ to remove any remaining lipid starting material. This was then repeated. The combined aqueous layers were subject to a stream of the nitrogen to remove any MeOH and the aqueous was removed by lyophilisation. Conversion of the reaction was quantified by HPLC (Waters 2795 Separations Module using a Jupiter 4U Proteo 90 Å 250 $\times$ 4.60 mm 4 micron column and using a Jasco FP-920 Intelligent Fluorescent Detector, 542 nm excitation, 568 emission) at 95%. HRMS (ESI): m/z Calcd for C₆₃H₈₂N₁₂O₂₀ [M+2H]²⁺: 663.2872: Found 663.2869.

3.10.4 - Glycosylation of TAMRA-GlcNAc Peptide with GalT (3.30)



To the TAMRA-GlcNAc peptide (**3.29**, 1 μ g, 0.75 nmol) in 100 mM sodium cacodylate, 1 mg/mL BSA, 2 mM MnCl₂ pH 7.4 (21 μ L) was added UDP-Galactose (1 μ L 10 mM, 10 nmol), alkaline phosphate (200 μ U) and GalT (600 μ U). The reaction was shaken at 37 °C overnight. Conversion of the reaction was quantified by HPLC (Waters 2795 Separations Module using a Jupiter 4U Proteo 90 Å 250 $\times$ 4.60 mm 4 micron column and using a Jasco FP-920 Intelligent Fluorescent Detector, 542 nm excitation, 568 emission) at 94%. HRMS (ESI): m/z Calcd for C₆₉H₉₂N₁₂O₂₅ [M+2H]²⁺: 744.3143; Found 744.3121.

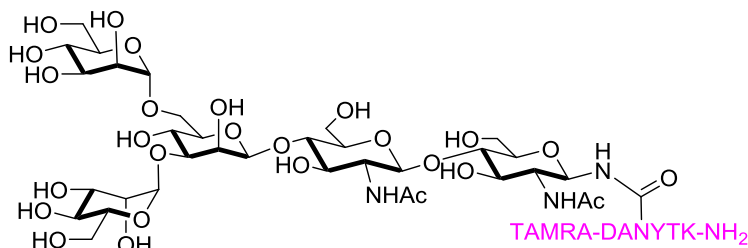
3.10.5 - Glycosylation of a Model System with GST-EndoA (3.32)



To the GlcNAc-triazole-amino acid (**3.31**, 3 μ L, 4.8 mM, 14.4 nmol) in 50 mM sodium phosphate buffer pH 6.5 (34 μ L) was added Man₃GlcNAc oxazoline (16 μ L 5 mM, 80 nmol), and GST-EndoA (9.6 μ L, 28 μ g/ml, 0.3 μ g, 3.3 pmol). The reaction was shaken at room temperature. Samples (10 μ L) were taken at 15 min, 30 min and 1 h and quenched (1 μ L 1% formic acid). Conversion of the reaction was quantified by HPLC (Waters 2795 Separations Module using a Synergi 4U Hydro-RP 80 Å 100 $\times$ 4.6 mm 4

micron column and using a Waters 2996 Photodiode Array Detector, detecting at 260 nm) at 49% after 30 min. HRMS (ESI): m/z Calcd for $C_{56}H_{78}N_6O_{30}Na$ $[M+Na]^+$: 1337.4655: Found 1337.4663.

3.10.6 - Glycosylation of TAMRA-GlcNAc Peptide with GST-EndoA (3.33)



To the TAMRA-GlcNAc peptide (**3.29**, 1 μ g, 0.75 nmol) in 50 mM sodium phosphate buffer pH 6 (20 μ L) was added Man₃GlcNAc oxazoline (1 μ L 6 mM, 6 nmol), and GST-EndoA (15 ng, 0.19 pmol). The reaction was shaken at room temperature and additional Man₃GlcNAc oxazoline (1 μ L 6 mM, 6 nmol) was added after 1.5 h and 2.5 h and additional GST-EndoA (15 ng, 0.19 pmol) added after 1.5 h. Conversion of the reaction was quantified by HPLC (Waters 2795 Separations Module using a Jupiter 4U Proteo 90 Å 250 $\times$ 4.60 mm 4 micron column and using a Jasco FP-920 Intelligent Fluorescent Detector, 542 nm excitation, 568 emission) at 64% after 3.5 h. HRMS (ESI): m/z Calcd for $C_{89}H_{125}N_{13}O_{40}$ $[M+2H]^{2+}$: 1007.9068: Found 1007.9054.

3.10.7 - Glycosylation of TAMRA-GlcNAc Peptide with GST-EndoA Mutants

To the TAMRA-GlcNAc peptide (**3.29**, 1 μ g, 0.75 nmol) in 50 mM sodium phosphate buffer pH 6 (16 μ L) was added Man₃GlcNAc oxazoline (2 μ L 6 mM, 12 nmol), and GST-EndoA or GST-EndoA mutant (28 ng, 0.38 pmol). The reaction was shaken at room temperature and conversion of the reaction was quantified by HPLC (Waters 2795

Separations Module using a Jupiter 4U Proteo 90 Å 250 × 4.60 mm 4 micron column and using a Jasco FP-920 Intelligent Fluorescent Detector, 542 nm excitation, 568 emission) after 15 min, 30 min, 1 h, 2 h, 4 h, 24 h and 48 h.

3.11 - References

- 1 C. Lizak, S. Gerber, S. Numao, M. Aebi and K. P. Locher, *Nature*, 2011, **474**, 350-356.
- 2 D. Calo, L. Kaminski and J. Eichler, *Glycobiology*, 2010, **20**, 1065-1076.
- 3 H. Nothaft and C. M. Szymanski, *J. Biol. Chem.*, 2013, **288**, 6912-6920.
- 4 R. H. Langdon, J. Cuccui and B. W. Wren, *Future Microbiol*, 2009, **4**, 401-412.
- 5 M. Wacker, D. Linto, P. G. Hitchen, M. Nita-Lazar, S. M. Haslam, S. J. North, M. Panico, H. R. Morris, A. Dell, B. W. Wren and M. Aebi, *Science*, 2002, **29**, 1790-1793.
- 6 F. Schwarz, C. Lizak, Y. Fan, S. Fleurken, M. Kowarik and M. Aebi, *Glycobiology*, 2011, **21**, 45-54.
- 7 A. Jervis, R. H. Langdon, P. G. Hitchen, A. J. Lawson, A. Wood, J. L. Fothergill, H. R. Morris, A. Dell, B. W. Wren and D. J. Linton, *J. Bacteriol*, 2010, **192**, 5228-5236.
- 8 M. V. Ielmini and M. F. Feldman, *Glycobiology*, 2011, **21**, 734-742.
- 9 F. Liu, B. Vijayakrishnan, A. Faridmoayer, T. A. Taylor, T. B. Parsons, G. J. L. Bernardes, M. Kowarik and B. G. Davis, *J. Am. Chem. Soc.*, 2014, **136**, 566-569.
- 10 K. J. Glover, E. Weerapana, S. Numao and B. Imperiali, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **12**, 1311-1316.
- 11 L. Li, R. Woodward, Y. Ding, X. Liu, W. Yi, V. S. Bhatt, M. Chen, L. Zhang and P. G. Wang, *Biochem. Biophys. Res. Co.*, 2010, **394**, 1069-1074.
- 12 M. Kowarik, N. M. Young, S. Numao, B. L. Schulz, I. Hug, N. Callewaert, D. C. Mills, D. C. Watson, M. Hernandez, J. F. Kelly, M. Wacker and M. Aebi, *EMBO J.*, 2006, **25**, 1957-1966.
- 13 K. J. Glover, E. Weerapana, S. Numao and B. Imperiali, *Chem. Biol*, 2005, **12**, 1311-1316.

- 14 F. Schwarz, W. Huang, C. Li, B. L. Schulz, C. Lizak, A. Palumbo, S. Numao, D. Neri, M. Aebi and L. Wang, *Nat. Chem. Bio.*, 2010, **6**, 264-266.
- 15 M. M. Chen, E. Weerapana, E. Ciepichal, J. Stupak, C. W. Reid, E. Swiezewska and B. Imperiali, *Biochemistry*, 2007, **46**, 14342-14348.
- 16 Y. Chang, C. Liu, C. Guo, Y. Wang, J. Fang and W. Cheng, *J. Org. Chem.*, 2008, **73**, 7197-7203.
- 17 B. Wu, R. Woodward, L. Wen, X. Wang, G. Zhao and P. G. Wang, *Eur. J. Org. Chem.*, 2013, 8162-8173.
- 18 M. Umekawa, C. Li, T. Higashiyama, W. Huang, H. Ashida, K. Yamamoto and L. Wang, *J. Biol. Chem.*, 2010, **285**, 511-521.
- 19 B. Li, Y. Zeng, S. Hauser, H. Song and L. Wang, *J. Am. Chem. Soc.*, 2005, **127**, 9692-9693.
- 20 G. Zou, H. Ochiai, W. Huang, Q. Yang, C. Li and L. Wang, *J. Am. Chem. Soc.*, 2011, **133**, 18975-18991.
- 21 W. Huang, J. Li and L. Wang, *Chem. Bio. Chem.*, 2011, **12**, 932-941.
- 22 J. J. Goodfellow, K. Baruah, K. Yamamoto, C. Bonomelli, B. Krishna, D. J. Harvey, M. Crispin, C. N. Scanlan and B. G. Davis, *J. Am. Chem. Soc.*, 2012, **134**, 8030-8033.
- 23 M. Ferandez-Gonzalez, O. Boutureira, G. J. L. Bernardes, J. M. Chalker, M. A. Young, J. C. Errey and B. G. Davis, *Chem. Sci.*, 2010, **1**, 709-715.
- 24 J. Yin, L. Li, N. Shaw, Y. Li, J. K. Song, W. Zhang, C. Xia, R. Zhang, A. Joachimiak, H. Zhang, L. Wang, Z. Liu and P. Wang, *Plos One*, 2009, **4**, e4658.
- 25 C. D. Heidecke, Z. Ling, N. C. Bruce, J. W. B. Moir, T. B. Parsons and A. J. Fairbanks, *Chem. Bio. Chem.*, 2008, **9**, 2045-2051.

26 L. M. Lira, D. Vasilev, R. A. Pilli and C. A. Wessjahann, *Tet. Lett.*, 2013, **54**, 1690-1692.

Chapter 4 – Synthesis and Testing of PglB Substrates with Variation on the Sugar; Investigating 4-Thiol GlcNAc

Chapter 4 – Synthesis and Testing of PglB Substrates with Variation on the Sugar; Investigating 4-Thiol GlcNAc

4.1 – Introduction

In the previous chapter it was demonstrated that it is possible to use PglB to catalyse the addition of GlcNAc to a specific consensus sequence. The effect of changing the lipid length on PglB substrate was also studied. It was desired next to investigate the effect of modifications on the sugar unit. More specifically, it was desired to investigate whether it was possible to use PglB to attach a GlcNAc unit with a thiol at the 4-position, to allow the synthesis of thioglycosides.

Thioglycosides are glycosides where the glycosidic bond, instead of being formed of a carbon-oxygen bond, is formed of a carbon-sulfur bond (Figure 4.1). The rate of enzymatic hydrolysis of S-glycosidic bonds is several times slower than that of O-glycosidic bonds, making them more metabolically stable.^{1, 2} Moreover thioglycosides adopt conformations similar to that of the naturally occurring O-glycosides.¹

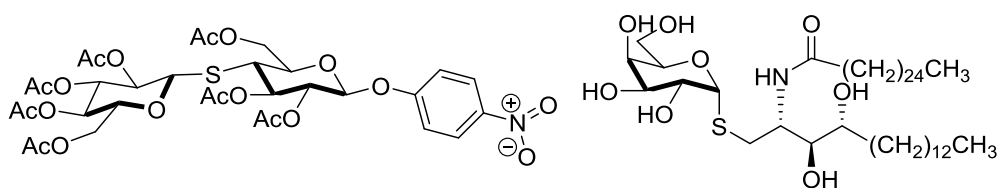
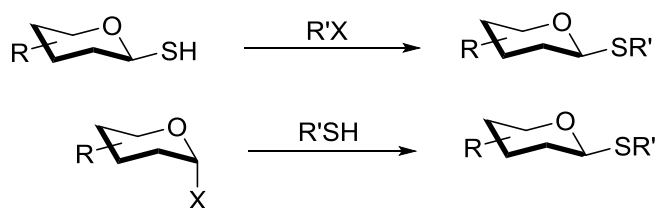


Figure 4.1: Two examples of thioglycosides, where a glycosidic linkage is formed of a sulfur atom.^{3,4}

Due to these two facts, there is a great interest in forming thioglycoside mimics of naturally occurring glycosides. They have been used as competitive inhibitors for glycosidases, including α -L-fucosidases⁵ and pancreatic α -amylase,⁶ and for the formation of stable complexes with proteins, allowing for the X-ray crystal structure of

proteins bound to substrate-like molecules to be obtained. Examples of the latter use of thioglycosides include the structure of endoglucanase Cel7B⁷ and maize β -glucosidase ZMGlu1.⁸ Thioglycosides are also gaining attention for their therapeutic use, including studies into their use for activation of invariant natural killer T-cells and for the control of hyperglycemia in diabetes.^{3,9}

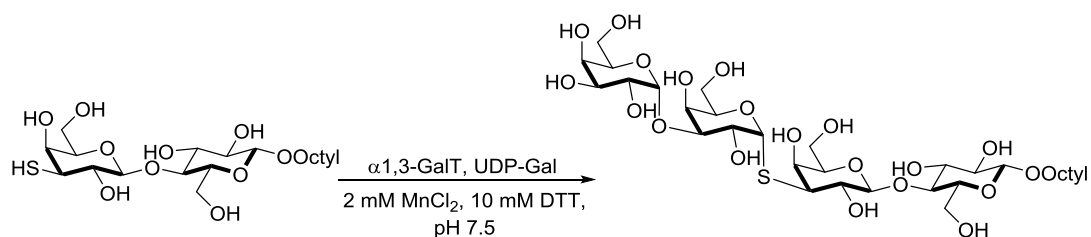
The chemical synthesis of thioglycosides is most commonly achieved through an S_N2 displacement reaction, with a thiol group attacking and displacing a halogen, triflate or other leaving group² (Scheme 4.1). This method has been used for the formation of disaccharides with an S-glycosidic linkage, as well as for the attachment and formation of a carbon-sulfur bond between the anomeric centre of the glycan and other chemical moieties. Alternative methods to S_N2 reactions include Michael addition of the thiol and Mitsunobu condensation.² Using these chemical methods, a variety of different thioglycosides have been formed, including thioglycopeptides where a glycan is attached to a peptide via an S-glycosidic linkage.¹⁰



Scheme 4.1: Chemical synthesis of thioglycosides via an S_N2 reactions, where X = halogen, triflate or another leaving group.²

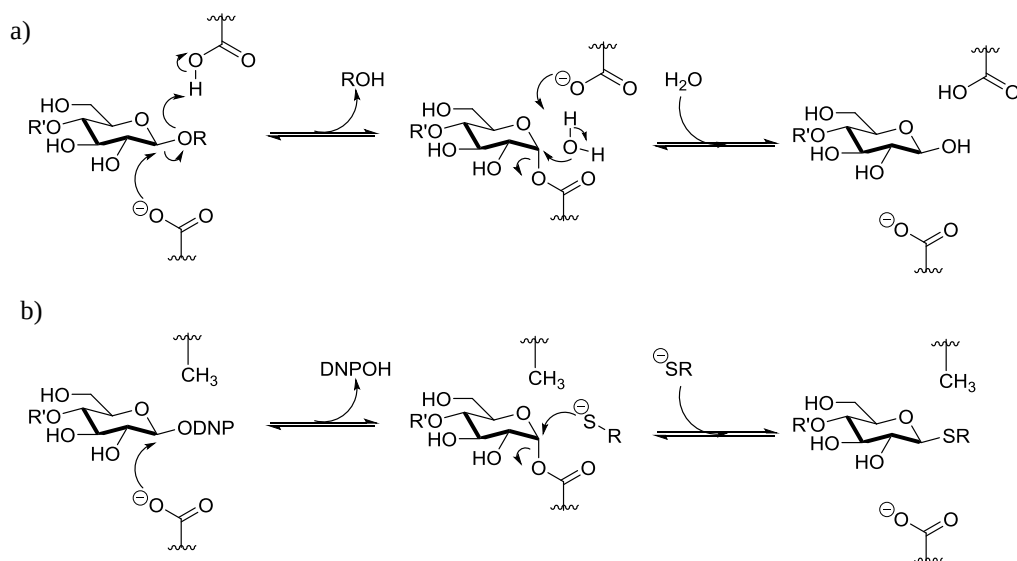
As well as the formation of thioglycosides chemically, research has been carried out into forming these enzymatically. One such example is work by *Rich et al.*¹, who

demonstrated that the glycosyltransferase α 1,3-GalT is able to attach a galactose unit to a thiol on a glycan acceptor using UDP-galactose as a donor (Scheme 4.2).



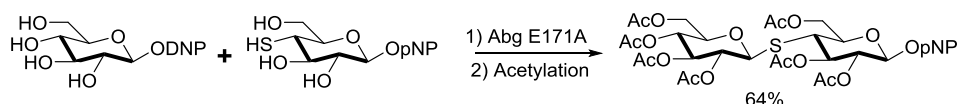
Scheme 4.2: Formation of a thioglycoside using a glycosyltransferase. Note, GalT attached two galactose units in this reaction.¹

Further research into enzymatic formation of thioglycosides was conducted by *Withers et al.*^{4, 11, 12} Instead of the use of glycosyltransferases, they used mutant glycosidase that have a retaining mechanism (see Chapter 1). The mechanism of these enzymes involves two acid/base catalytic residues, where one acts as a nucleophile, forming an enzyme-substrate complex and the other protonates the leaving group, acting as an acid. This residue subsequently acts as a base, deprotonating water, which breaks the bond between the enzyme and substrate, giving the hydrolysed product (Scheme 4.3). *Withers et al.*⁴ predicted that if an activated glycan donor was used instead (having dinitrophenyl attached at the anomeric position, for example), then the acid/base catalytic residue that does not bind to the substrate would not be required and could be removed through mutagenesis, with the enzyme-substrate complex still able to form. This complex could then be attacked by a thiol attached to a glycan acceptor, acting as a nucleophile, to give the thioglycoside (Scheme 4.3).



Scheme 4.3: Use of glycosidase mutants to form thioglycosides, a) mechanism of the natural activity of glycosidases, b) predicted mechanism for the formation of thioglycoside using mutant glycosidase.⁴ Note, DNP = dinitrophenol.

This reaction was investigated with β -glucosidase Abg E171A and β -mannosidase Man2A E429A mutants, using an appropriate donor and an acceptor with a thiol group. The desired thioglycoside was observed in both cases (Scheme 4.4).⁴

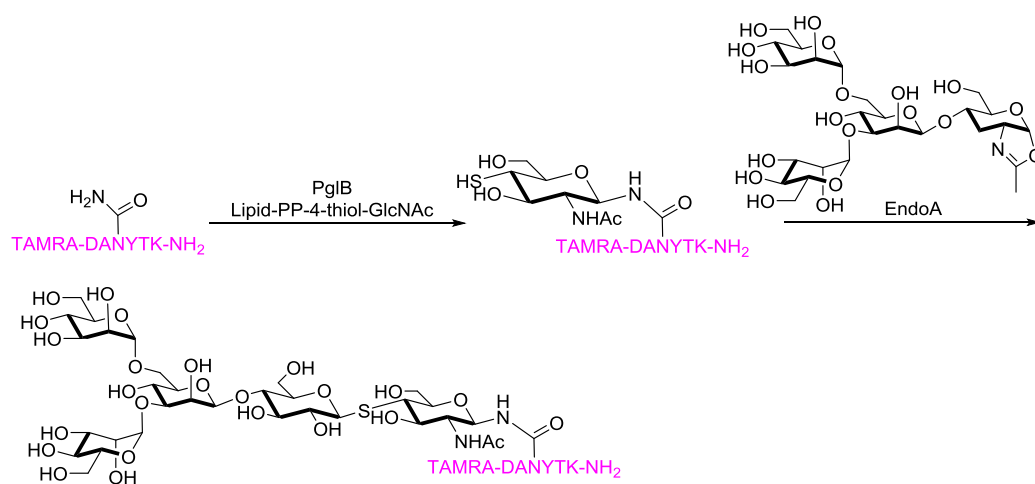


Scheme 4.4: Use of Abg E171 glycosidase for the formation of a thioglycoside. Note, DNP = dinitrophenol and pNP = *para*-nitrophenyl.⁴

Further research on Abg showed that the E171Q mutation resulted in a 5-fold enhancement in the reaction rate compared to Abg E171A, with the E171Q being presumed to provide catalytic assistance.¹¹ The E171G mutation was also seen to be an effective mutation on Abg, predicted to be due to increased flexibility in the active site. It was Abg E171G that *Withers et al.* took on for further study and demonstrated that

not only could this enzyme be used for the formation of disaccharide thioglycosides, but could also be used for the formation of thioglycosides where the glycan acceptor is attached to a protein, creating a thioglycoprotein.¹³

Here, it was wished to investigate thioglycoside enzymatic synthesis with Endo enzymes by combining the PglB and Endo catalysis, as demonstrated in Chapter 3. This strategy had two novel steps. The first was the use of PglB to attach a GlcNAc with a thiol at the 4-position that could then be utilised in the second novel step, the formation of the thioglycoside through an Endo catalysed glycosylation reaction using an oxazoline donor (Scheme 4.5). The resulting S-glycosidic linkage would be resistant to hydrolysis and hence would solve the major disadvantage to Endo catalysed glycosylation reactions where the product is a substrate for natural hydrolytic activity, as observed in Chapter 2 and, in part, Chapter 3. As the product would be resistant to hydrolysis, complete conversion to the glycosylation product should be possible. With thioglycosides having similar conformations to natural glycosides, these would make good minics of naturally occurring glycoproteins.

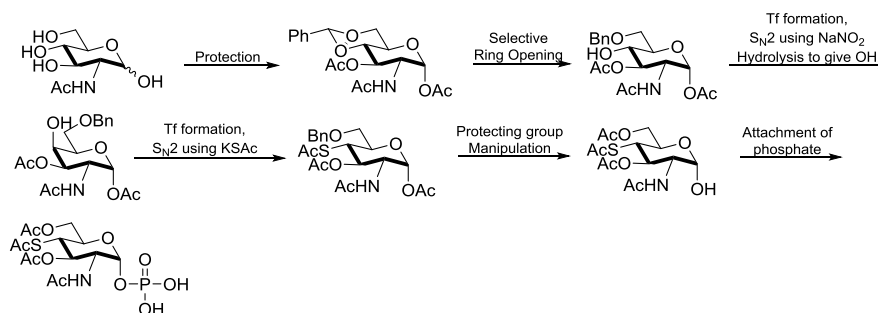


Scheme 4.5: Strategy for the formation of a thioglycoside attached to a peptide, via a PglB catalysed reaction and an EndoA catalysed reaction, as investigated here.

As stated at the beginning of this chapter, the second reason why it was desired to investigate the attachment of a GlcNAc unit with a thiol at the 4-position was to investigate the substrate specificity of PglB, more specifically, whether PglB can attach unnatural sugars. As part of the study into this, work towards the synthesis of other interesting sugars was undertaken, aiming to form novel polyprenol pyrophosphate glycan substrates to be tested with PglB. This work, as well as the synthesis of the polyprenol pyrophosphate thiol-GlcNAc substrate, the testing of this with PglB and studies into thioglycoside formation using EndoA, is described.

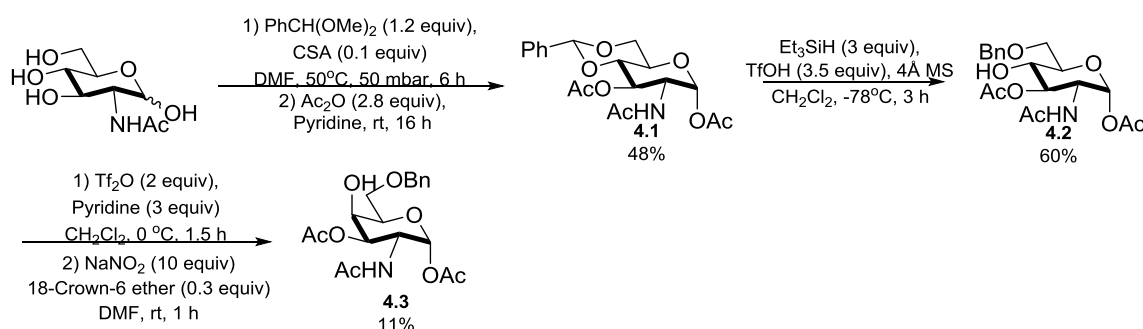
4.2 – Thiol GlcNAc PglB Substrate Synthesis

To investigate whether PglB can accept a substrate where there is a thiol group at the 4-position and the subsequent glycosylation with EndoA, it was first necessary to synthesise a PglB substrate with this modification. As stated in Chapter 3, it can be envisaged that the synthesis can be broken down into three steps; sugar synthesis, lipid synthesis and coupling of the two. The synthesis of the lipid has been described in Chapter 3. Concerning the synthesis of the sugar, the major challenge foreseen was the insertion of the thiol group in the correct configuration. The initial proposed synthesis of the sugar started from *N*-acetylglucosamine and began with the protection of all positions except the 4-position (Scheme 4.6). This was possible by benzylidene formation, acetylation and regioselective ring opening, leaving the 4-position unprotected. Then, formation of a triflate leaving group on the 4-position would allow an S_N2 attack to occur with sodium nitrite that, after hydrolysis, would result in a hydroxyl group in an axial configuration at this position, as reported by *Chen* and *Withers*.¹⁴ This process, forming a triflate, followed by an S_N2 attack, is repeated but using potassium thioacetate to give a thioacetate group in an equatorial configuration at the 4-position. Protecting group manipulation followed by deprotection of the anomeric centre would then allow the phosphate group to be installed resulting in the desired sugar.



Scheme 4.6: Proposed synthesis of the thiol GlcNAc unit.¹⁴

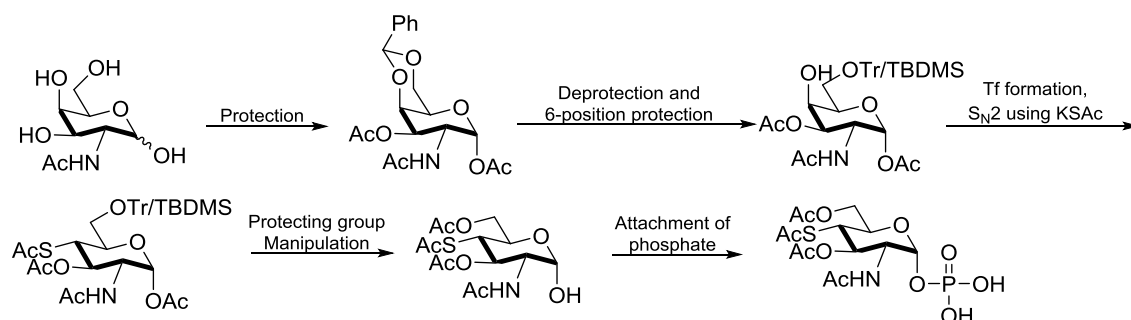
The initial protecting group manipulations were all successfully carried out. The benzylidene formation was carried out with benzaldehyde dimethyl acetal using an acid catalyst as reported for glucose.^{15, 16} Initial use of *p*-toluenesulfonic acid as the catalyst resulted in a poor yield (16% after acetylation); replacing *p*-toluenesulfonic acid with camphor sulfonic acid, resulted in a 3-fold increase in yield (48%, **4.1**, Scheme 4.7). Triethylsilane with trifluoromethanesulfonic acid at -78°C was then used to selectively open the benzylidene ring¹⁷ (**4.2**); other methods used either resulted in no reaction (Sodium cyanoborohydride with methanesulfonic acid¹⁸ or hydrochloric acid¹⁹) or reduction of the acetyl groups (triethylsilane with boron trifluoride diethyl etherate²⁰). It should be noted here that although triethylsilane with trifluoromethanesulfonic acid gave the desired product, yields were inconsistent over several reaction runs (13% to 60%). Next, the first $\text{S}_{\text{N}}2$ reaction and inversion of the hydroxyl group at the 4-position was attempted. The formation of the triflate was seen to be facile and subsequent displacement with a nitrite group and hydrolysis resulted in the formation of the desired product. However, the purification of this compound was challenging, resulting in a disappointing yield (**4.3**, 11% yield).



Scheme 4.7: Initial attempt at synthesis of the thiol sugar, via a double inversion.

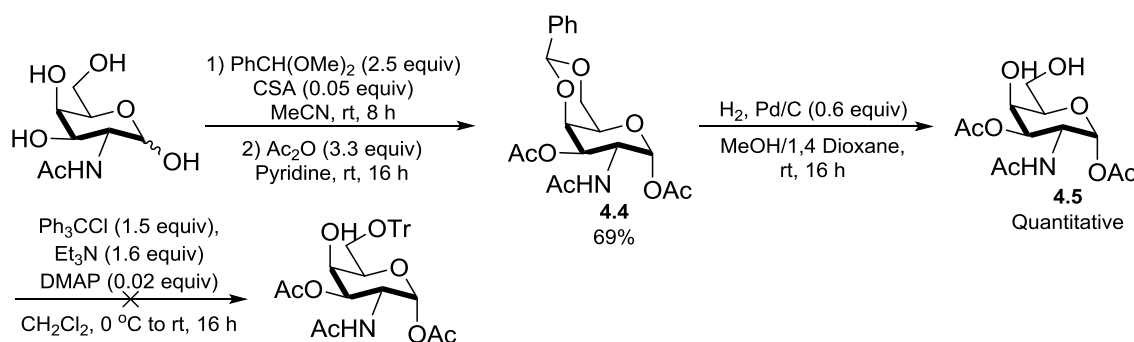
With a low yield obtained from this reaction, as well as difficulties with the selective benzylidene ring opening, an alternative method to obtain the sugar was sought. Based

on the work of *Nishimura et al.*,²¹ it was decided to start the synthesis from *N*-2-acetylgalactosamine, instead of *N*-2-acetylglucosamine, eliminating the need for the sodium nitrite induced inversion at the 4-position. Furthermore, the specific benzylidene ring opening step was removed by instead deprotecting the 4- and 6-positions by hydrogenation of the benzylidene ring, followed by protection of the 6-position with a triphenylmethyl group (Scheme 4.8).



Scheme 4.8: Alternative strategy for the synthesis of the thiol sugar based on the work of *Nishimura et al.*²¹

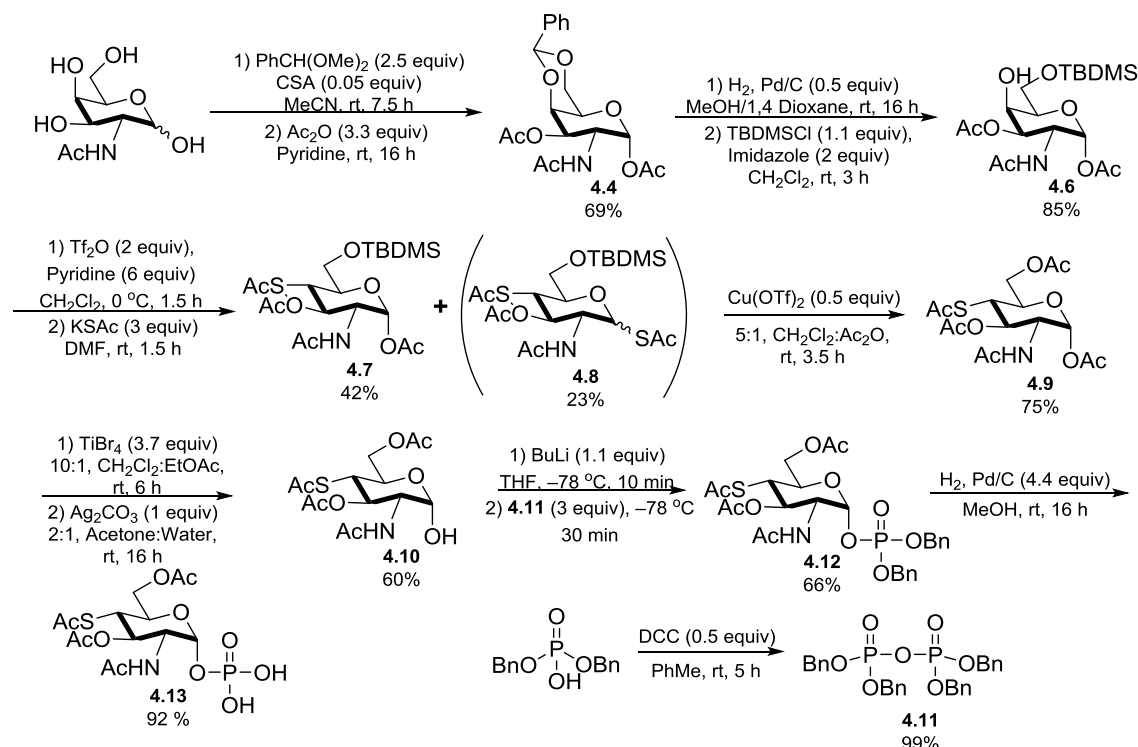
The initial protection and subsequent hydrogenation of the benzylidene both proceeded as desired (**4.4** and **4.5**, Scheme 4.9), but the triphenylmethyl protection failed using the condition described by *Nishimura et al.*²¹



Scheme 4.9: Initial attempt at the protection of GalNAc to leave position-4 unprotected.

With the triphenylmethyl protection failing, the 6-position was protected with *tert*-butyldimethylsilyl instead (**4.6**, Scheme 4.10), with the hydrogenation and silyl protection done in one pot. The S_N2 reaction with thioacetate was successfully conducted; the triflate was formed and converted to the thioacetate in one pot, using conditions similar to those previously reported in the literature^{14, 19} (**4.7**). The yield from this reaction was lower than expected, with the displacement of the anomeric acetyloxy group with a thioacetate group the major side product formed (**4.8**). With the thioacetate group installed, the next step was the replacement of the *tert*-butyldimethylsilyl group with an acetyl group. This was first attempted using tetrabutylammonium fluoride, followed by acetylation with acetic anhydride and pyridine. However, this resulted in deprotection of the thioacetate group and formation of disulfide dimers. It was believed that this was the result of deprotection of the *tert*-butyldimethylsilyl followed by migration of the acetyl group on the thioacetate to the 6-position followed by dimerisation. Due to this, an alternative approach was sought, where the *tert*-butyldimethylsilyl was not removed, but directly substituted with an acetyl group. Chandra *et al.*²² reported the direct conversion of silyl ethers into acetate groups using acetic anhydride with a catalytic amount of copper triflate (5 mol%), demonstrating this reaction on a variety of molecules obtaining good yields (over 65 %) in all cases. This reaction was attempted on the thiol-sugar and pleasingly gave the desired product in a good yield (**4.9**), although 50% mol catalyst was required. The final steps to the desired product were performed as described by Elhalabi and Rice¹⁹, namely the anomeric acetate group was converted to a bromide before being hydrolysed to give an unprotected hydroxyl at the anomeric position (when the deprotection of the anomeric position was attempted with hydrazine acetate, deprotection of the thioacetate was

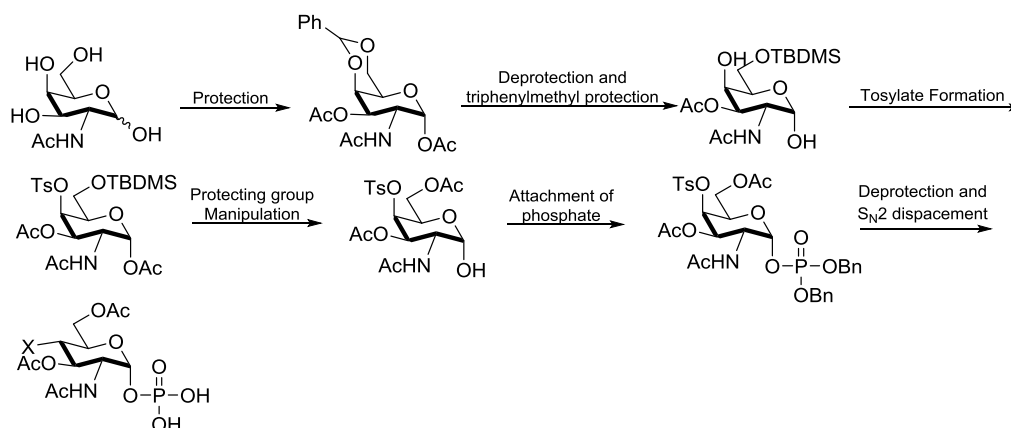
seen), to which the phosphate group was attached using tetrabenzylpyrophosphate, followed by hydrogenation to give the desired protected thiol GlcNAc sugar (**4.13**).



Scheme 4.10: Synthesis of the 4-thio-GlcNAc phosphate sugar.

As well as the insertion of a thiol group on the 4-position of GlcNAc, it was desired to also insert other functionalities, creating a number of different sugars with variations at the 4-position, that could then be tested as PglB substrates once the phospholipid was attached. Depending on the functionality, this would allow further derivatisation on the peptide or protein after the PglB reaction. To achieve this, it was thought that it would be possible to take the protected GalNAc sugar with the 4-position unprotected from the thiol GlcNAc synthesis (**4.6**) and then form a tosylate at this position (Scheme 4.11). Protecting group manipulations, attachment of the phosphate group at the anomeric position, followed by late stage derivatization could be used to form a number of

different phosphosugars with modifications at the 4-position. Coupling to a phospholipid would then give the PglB substrates.



Scheme 4.11: Planned synthesis of GlcNAc phosphate with variation at the 4-position, where X = functionality that can be installed via an S_N2 reaction.

A variety of different conditions were attempted for the tosylation of the 4-position, using a base to deprotonate the hydroxyl group followed by the addition of tosyl chloride, but unfortunately the desired product was not observed under any conditions (Table 4.1). It was thought that the problem is most likely the attachment of the tosyl group; the resulting tosylate group would have been in an axial position and next to the *tert*-butyldimethylsilyl group, causing a steric clash.

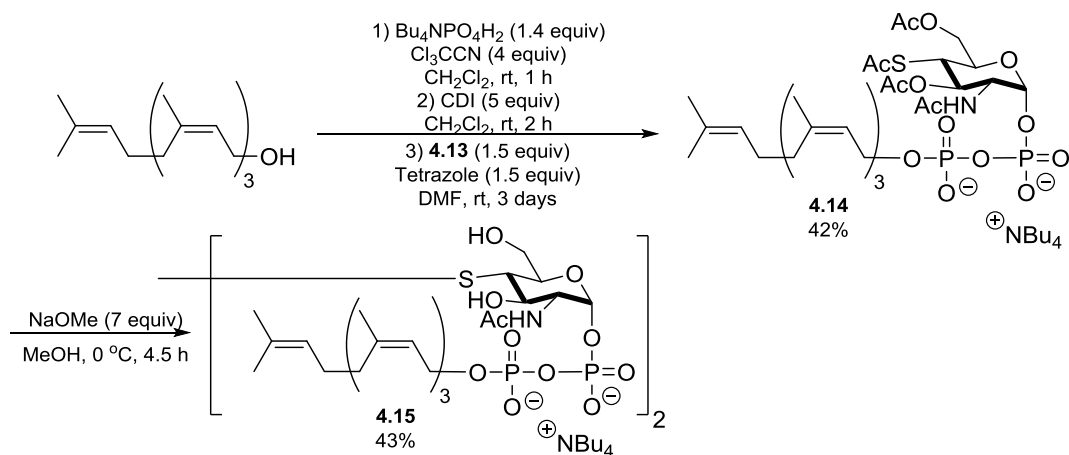


Conditions	Result
TsCl (2 equiv), Pyridine (6 equiv), DIPEA (6 equiv), CH ₂ Cl ₂ , rt, 16 h	No reaction
TsCl (1.2 equiv), NaH (2.5 equiv), CH ₂ Cl ₂ , Δ, 16 h	Decomposition
TsCl (1.2 equiv), NaH (2.5 equiv), DMF, rt, 16 h	No reaction

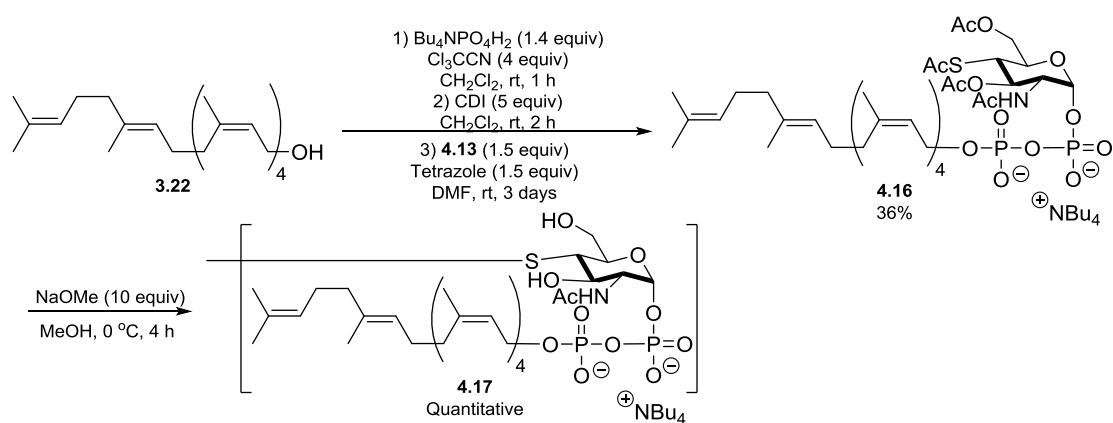
Table 4.1: Attempts at forming the tosylate group at 4-position.

With the protected thiol GlcNAc in hand, coupling to the lipid was attempted. This was done as before (see Chapter 3). The lipids used were a previously synthesised WZ₃ lipid²³ and the WEZ₄ lipid (**3.22**, see Chapter 3 for the synthesis). The coupling of the lipids to the thiol-GlcNAc and subsequent deprotection successfully gave WZ₃-thiol-GlcNAc and WEZ₄-thiol-GlcNAc (Schemes 4.12 and 4.13). An important factor to note here is that after deprotection both the WZ₃-thiol-GlcNAc and WEZ₄-thiol-GlcNAc formed disulfide dimers. Also, it should be stated that WEZ₄-thiol-GlcNAc was not purified after final deprotection, resulting in the greater yield compared to WZ₃-thiol-GlcNAc.

Chapter 4 – Synthesis and Testing of PglB Substrates with Variation on the Sugar; Investigating 4-Thiol GlcNAc



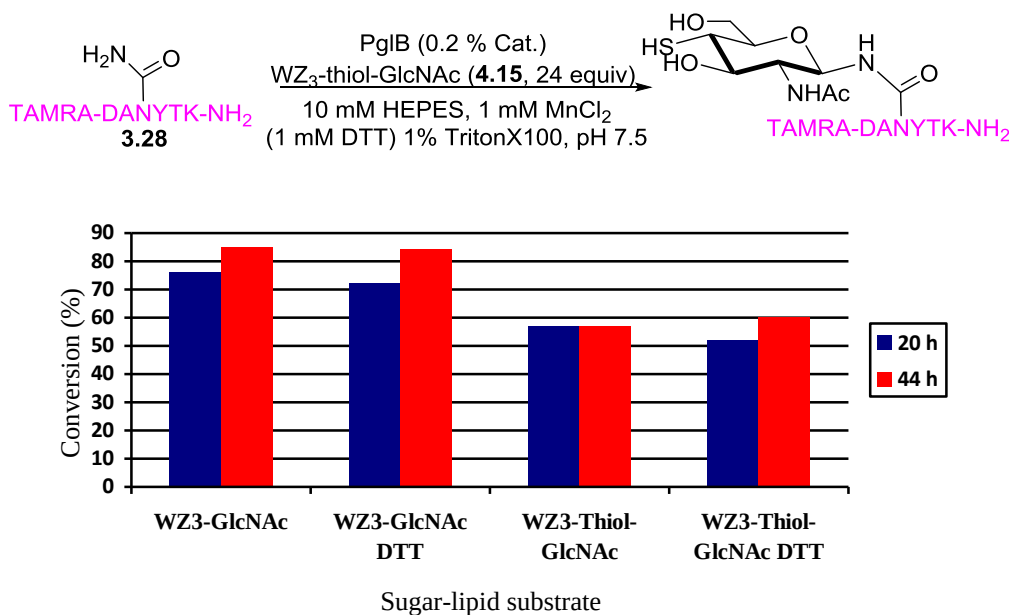
Scheme 4.12: Coupling of thiol-GlcNAc phosphate (4.14) to WZ₃ lipid and subsequent deprotection.



Scheme 4.13: Coupling of thiol-GlcNAc phosphate (4.14) to WEZ₄ lipid (3.22) and subsequent deprotection.

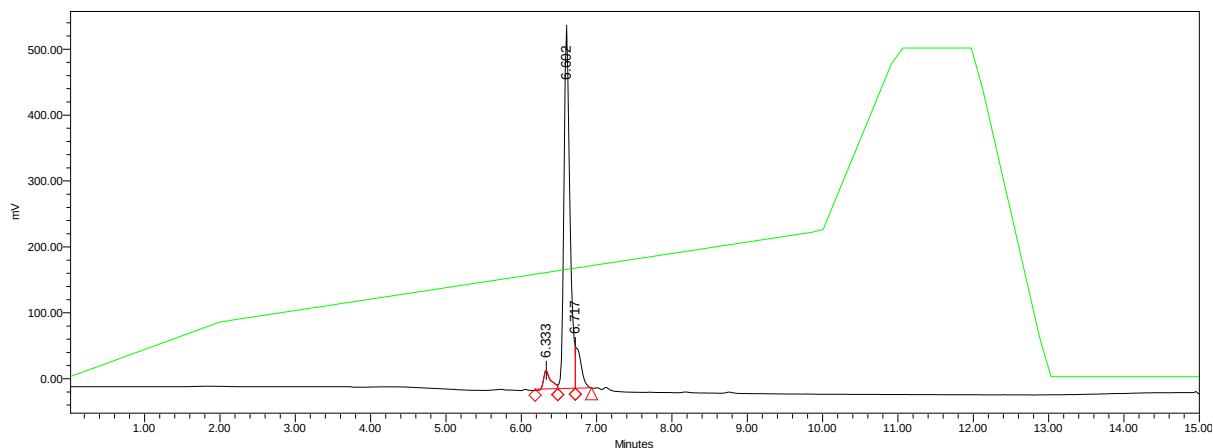
4.3 – Testing the Thiol-GlcNAc Substrates with PglB

With both thiol-GlcNAc substrates in hand, testing with PglB could commence. The WZ₃-thiol-GlcNAc was tested first. Again, the TAMRA-peptide (**3.28**) was used as the acceptor and WZ₃-GlcNAc, which had been previously formed, used as a positive control. Reactions were also set up where the buffer contained 1 mM dithiothreitol (DTT); it was known that the thiol-GlcNAc substrate formed disulfide dimers and it was expected that these needed to be broken in order for it to be accepted as a substrate, although it was unknown what effect DTT would have on PglB. Conversions of the reactions were found after 20 h and 44 h by HPLC, using fluorescent detection (Graph 4.1, Figure 4.2).



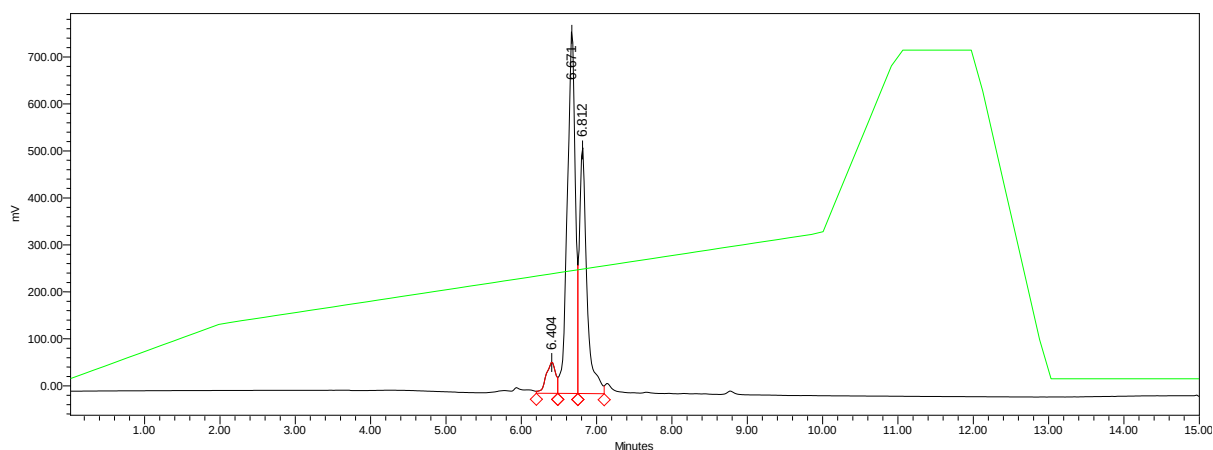
Graph 4.1: Conversion of the TAMRA-Peptide when using either WZ₃-GlcNAc or WZ₃.thiol-GlcNAc (**4.15**).

a) WZ₃-GlcNAc substrate after 20 h



	Name	Retention Time	Area	% Area
1	Impurity	6.333	189540	5.36
2	TAMRA-GlcNAc- Peptide (3.29)	6.602	3014370	85.18
3	TAMRA-Peptide (3.28)	6.717	334887	9.46

b) WZ₃-thiol GlcNAc substrate after 20 h



	Name	Retention Time	Area	% Area
1	Impurity	6.404	577063	5.84
2	Product	6.671	5636654	57.00
3	TAMRA-Peptide (3.28)	6.812	3675413	37.17

Figure 4.2: HPLC trace of PglB reaction with TAMRA-peptide and a) WZ₃-GlcNAc or b) WZ₃-Thiol-GlcNAc. Green line on each trace is the percentage amount of acetonitrile. All reactions done with 0.44 pmol PglB.

These experiments seemed to indicate that WZ₃-thiol-GlcNAc did function as a substrate with PglB, with a product being formed with a similar retention time to that

when using the standard GlcNAc substrate. It was also seen that the DTT in the buffer had little effect on any of the reactions. This was rather unexpected in the case of the WZ₃-thiol-GlcNAc substrate, where it was expected that the use of the DTT was to be critical, due to the formation of disulfide dimers.

Although the retention time of the product was as expected, it was necessary to confirm that the desired product had successfully formed. The sample was therefore analysed by LC-MS and MALDI mass spectrometry (Figure 4.3).

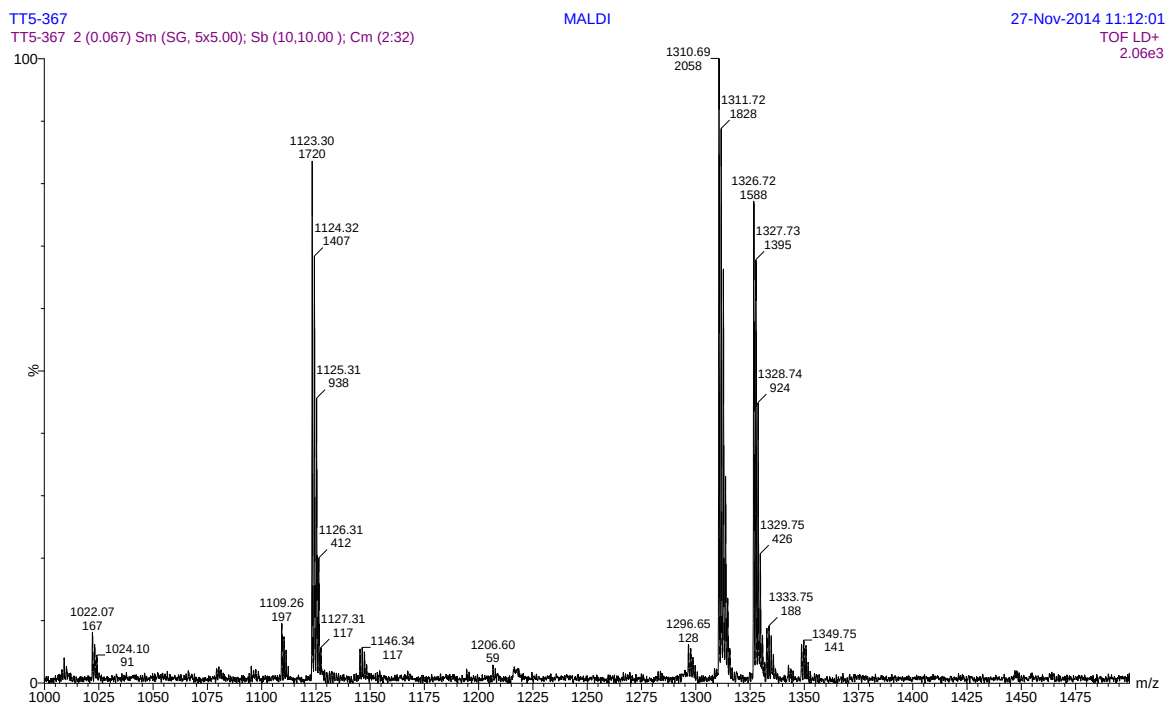


Figure 4.3: MALDI MS spectrum of the Thiol-GlcNAc-WZ₃ reaction. The starting material (TAMRA-Peptide) peak can be seen (1123.5 Da) as well as two product peaks (1310.5 Da and 1326.5 Da).

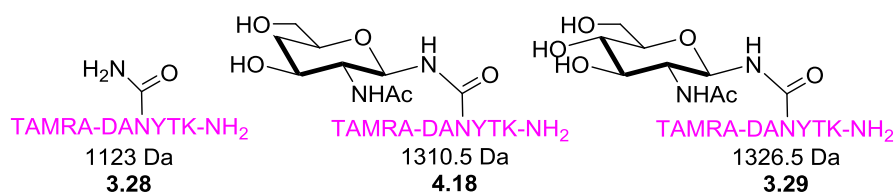
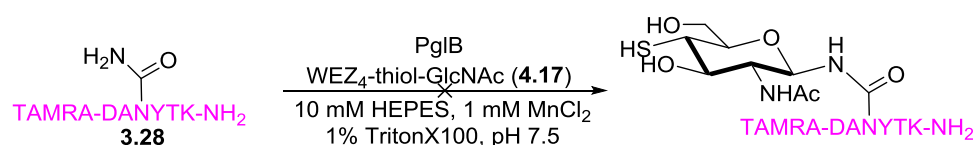


Figure 4.4: Predicted structures of compounds seen in the MALDI MS of the WZ₃-thiol GlcNAc reaction and their corresponding protonated masses.

Unfortunately, the desired mass (1341.5 Da) could not be observed and instead two peaks were observed at 1310.5 Da and 1326.5 Da. A peak for the starting material could also be seen, due to the reaction not going to completion. The proposed products relating to these two masses was the attachment of a GlcNAc to the TAMRA-peptide (**3.29**, 1326.5 Da, Figure 4.4) and the attachment of 4-deoxy GlcNAc to the TAMRA-peptide (**4.18**, 1310.5 Da). Returning to the mass spectrum of the WZ₃-thiol-GlcNAc substrate, the mass of WZ₃-GlcNAc and WZ₃-4-deoxy-GlcNAc could be seen, explaining the mass peaks observed. This also clarifies why no difference was seen when not using DTT with the thiol-GlcNAc-WZ₃, where the reduction of the disulfide was thought to be key. WZ₃-GlcNAc and WZ₃-4-deoxy-GlcNAc are thought to have formed either in the coupling reaction or the final deprotection, with the NMR of the 4-thio-sugar (**4.13**) being very clean. It is unknown how these are formed. Although it is interesting that PglB can accept 4-deoxy GlcNAc, the desired product was not perceived and the presences of impurities meant that this substrate was not investigated further.

Next the WEZ₄-thiol-GlcNAc (**4.17**) substrate was investigated with PglB and the TAMRA-peptide, hoping the longer lipid would be an advantage (compared to WZ₃-thiol-GlcNAc, **4.15**) and would allow the desired product to be formed.

As it had no effect on the control reactions, greater amounts of DTT were investigated, to ensure that all disulfide dimers were reduced, to rule out their possible effect on reaction yields. The use of *tris*(2-carboxyethyl)phosphine (TCEP) as a reducing agent was also studied. A variety of reactions were set up, using differing amounts of DTT and TCEP (Table 4.2), and the appropriate control reactions with WEZ₄-GlcNAc. The conversion of the reactions were found by HPLC with a fluorescent detector after 20 h (Figure 4.5).

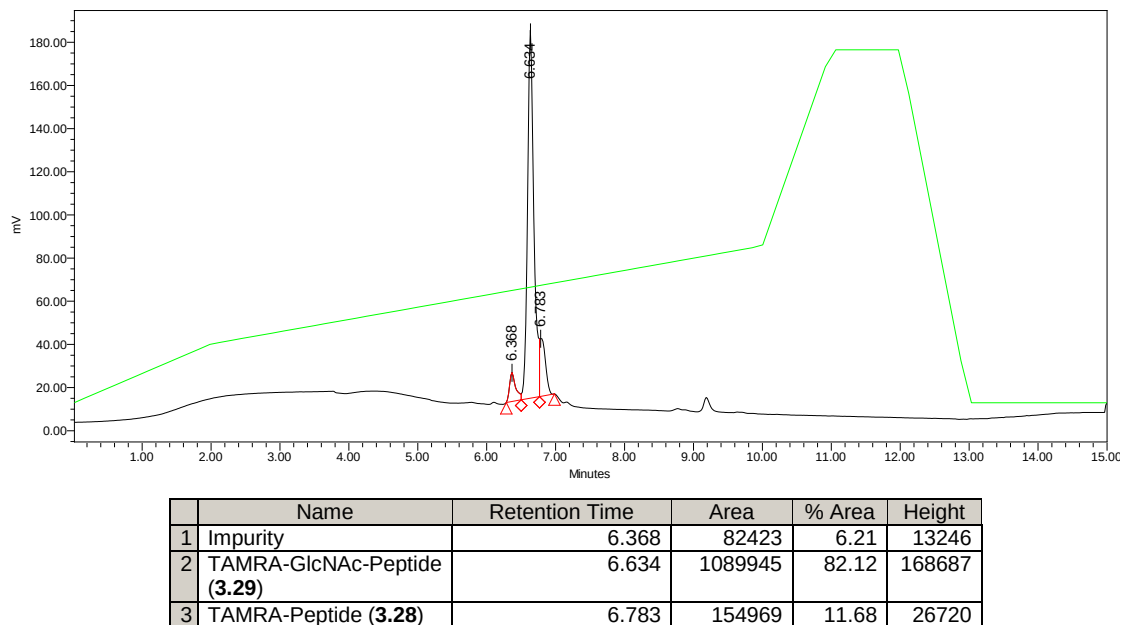


Entry	WEZ ₄ -thiol-GlcNAc (4.17) Equiv	Reducing agent	PglB	Result
1	8	2 mM DTT	0.05%	No reaction
2	8	8 mM DTT	0.05%	No reaction
3	16	50 mM DTT	0.2%	27% conversion to 3.29
4	16	25 mM TCEP	0.2%	68% conversion to 3.29

Table 4.2: Condition attempted when using the WEZ₄-thiol-GlcNAc in PglB glycosylation reaction.

Chapter 4 – Synthesis and Testing of PglB Substrates with Variation on the Sugar; Investigating 4-Thiol GlcNAc

a) Control Reaction with WEZ₄-GlcNAc, 0.05% PglB and 50 mM DTT



b) Reaction with WEZ₄-thiol-GlcNAc, 0.2% PglB and 50 mM DTT

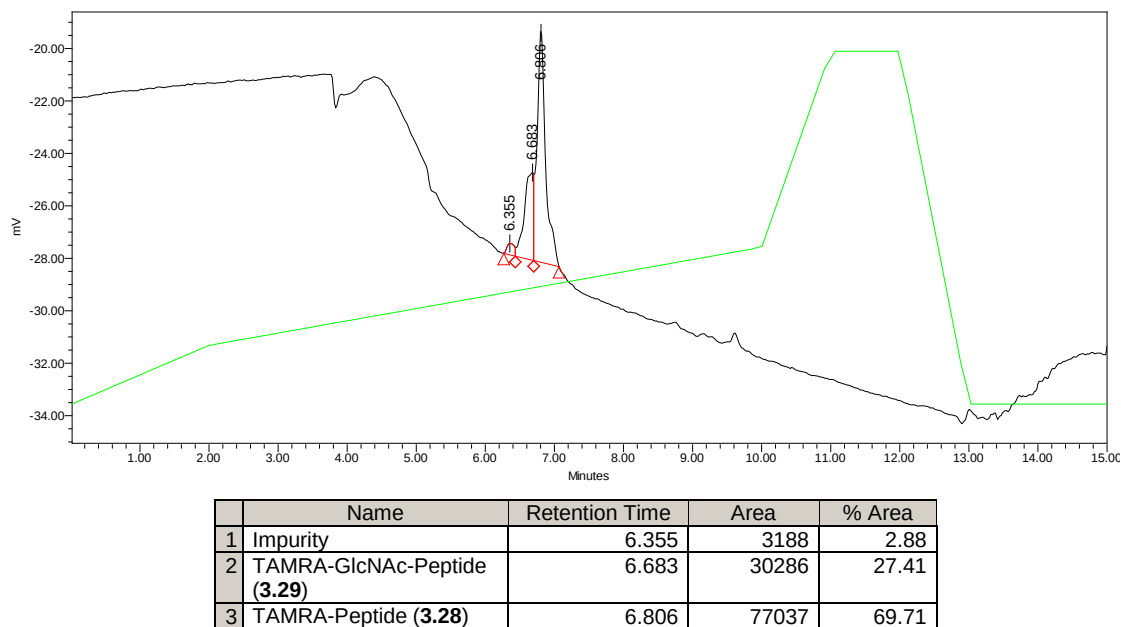
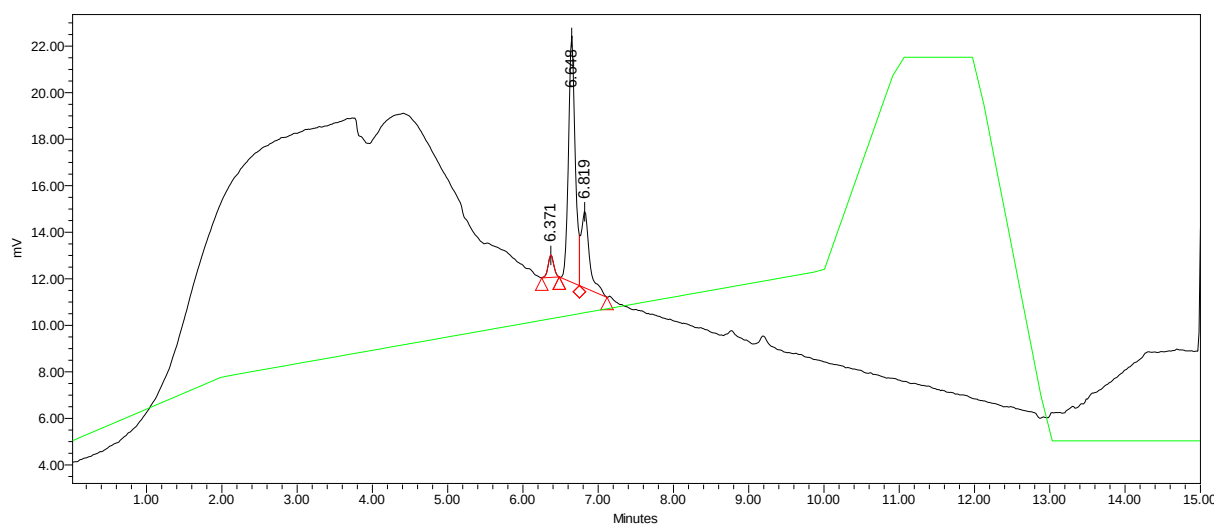


Figure 4.5: HPLC traces of PglB reaction with TAMRA-peptide and a) WEZ₄-GlcNAc with 0.05% PglB and 50 mM DTT and b) WEZ₄-thiol-GlcNAc with 0.2% PglB and 50 mM DTT. Green line on each trace is the percentage amount of acetonitrile.

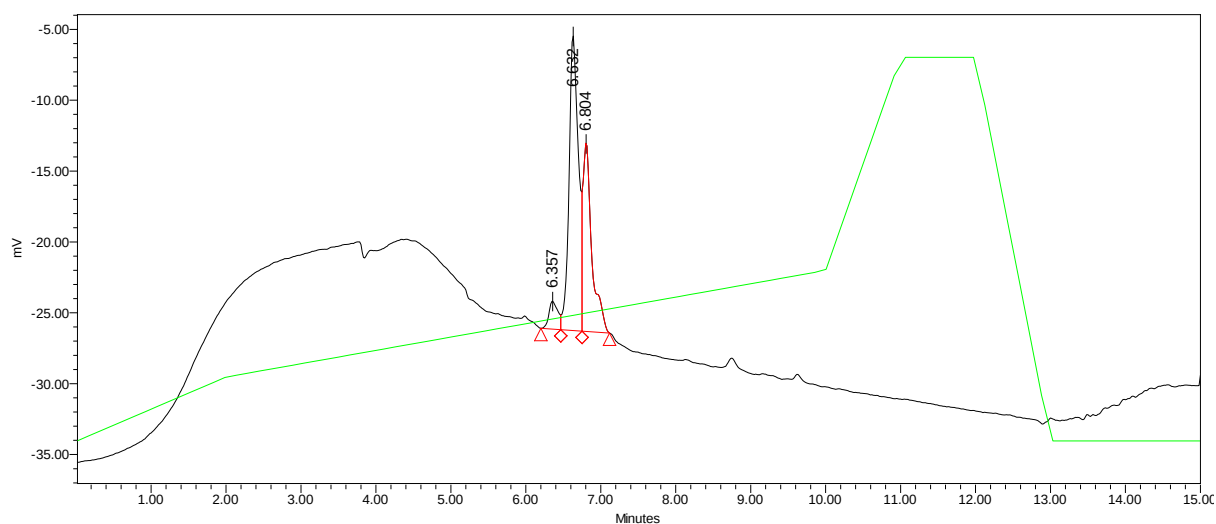
Chapter 4 – Synthesis and Testing of PglB Substrates with Variation on the Sugar; Investigating 4-Thiol GlcNAc

c) Control Reaction with WEZ₄-GlcNAc, 0.05% PglB and 25 mM TCEP



	Name	Retention Time	Area	% Area
1	Impurity	6.371	5085	5.25
2	TAMRA-GlcNAc-Peptide (3.29)	6.648	65691	67.84
3	TAMRA-Peptide (3.28)	6.819	26059	26.91

d) Reaction with WEZ₄ thiol GlcNAc, 0.2% PglB and 25 mM TCEP



	Name	Retention Time	Area	% Area
1	Impurity	6.357	16465	5.32
2	TAMRA-GlcNAc-Peptide (3.29)	6.632	182620	59.06
3	TAMRA-Peptide (3.28)	6.804	110136	35.62

Figure 4.5 cont'd: HPLC traces of PglB reaction with TAMRA-peptide c) WEZ₄-GlcNAc with 0.05% PglB and 25 mM TCEP and d) WEZ₄-thiol-GlcNAc with, 0.2% PglB and 25 mM TCEP, Green line on each trace is the percentage amount of acetonitrile.

All control reactions occurred as expected, with formation of the TAMRA-GlcNAc-peptide (**3.29**) as observed by HPLC. It should be noted, that some of the control reactions did not go to completion (Figure 4.5a and 4.5c), suggesting that the high amount of reducing agent diminishes PglB activity, most likely due to the effect that the reducing agent has on the pH of the reaction buffer. Most reactions with WEZ₄-thiol-GlcNAc returned starting material. Only in the reactions where the substrate and PglB were used in greater amounts (reactions with 50 mM DTT and 25 mM TCEP, Table 4.1, entries 3 and 4) was consumption of the starting material observed. MS analysis was conducted on both these reactions to confirm the identity of the product. Disappointingly, MS suggested the formation of TAMRA-GlcNAc-peptide (**3.29**), presumably due to the presence of a low amounts of WEZ₄-GlcNAc in the reaction. As with the reaction using the WZ₃-thiol-GlcNAc substrate, how WEZ₄-GlcNAc came to be in the reaction with the WEZ₄-thiol-GlcNAc substrate is unknown. From these experiments, it was concluded that PglB is unable to catalyse the addition of 4-thiol-GlcNAc onto a peptide or protein and is unable to accept a glycan donor with this sugar unit attached.

One observation of note is the increased conversion of the TAMRA-peptide starting material to TAMRA-GlcNAc-peptide in the reaction with TCEP, compared to the reaction using DTT. It is possible that TCEP is reacting with the WEZ₄-thiol-GlcNAc substrate to form WEZ₄-GlcNAc which then forms the observed product. The effect of TCEP on the PglB reaction warrants further study.

4.4 – Investigating EndoA Glycosylation Activity on a Thiol Acceptor

4.4.1 – Formation of a Test System

Whilst investigating the activity of PglB with 4-thiol-GlcNAc glycan substrates, studies were conducted into the second part for the synthesis of a glycopeptide with a hydrolytic resistant glycan; the glycosylation reaction with EndoA using a thiol acceptor.

First, a thiol acceptor substrate had to be synthesised. With it previously seen that a GlcNAc attached to an amino acid via a triazole could be used as the glycan acceptor in an EndoA catalysed glycosylation (**3.31**, Figure 4.6, see Chapter 3 and previously published work²⁴), the thiol analogue of this was formed (**4.19**).

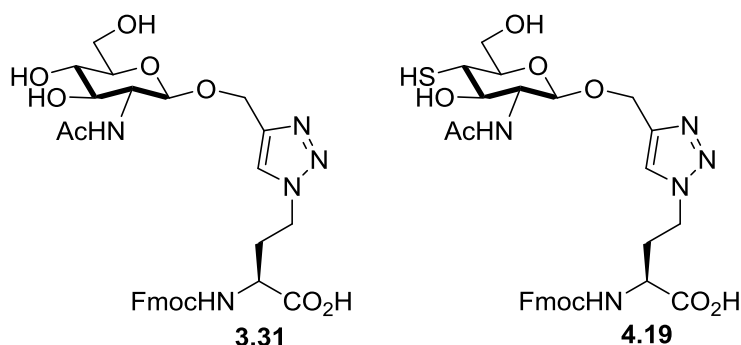
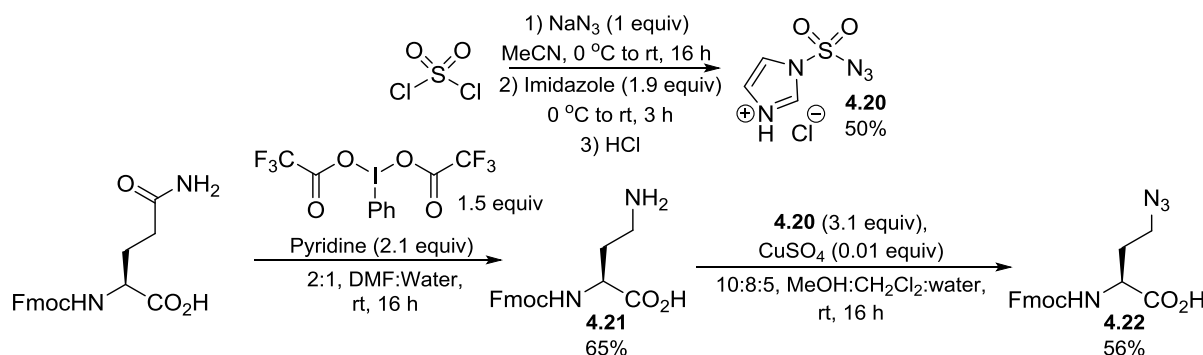


Figure 4.6: Model systems that can be used to investigate EndoA glycosylation to form a thioglycoside.

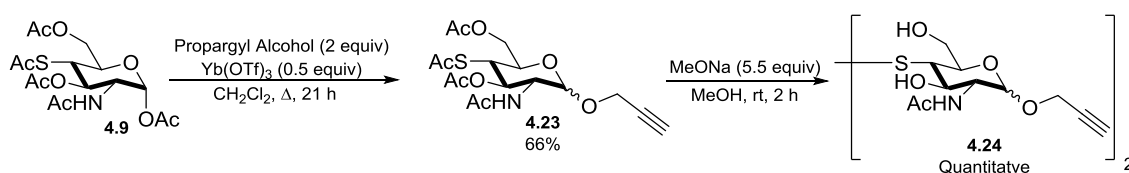
The synthesis of this substrate can be broken into two parts; the amino acid (Fmoc azidohomoalanine) and the thiol GlcNAc with propargyl group attached at the anomeric position. These two could then be coupled via a copper catalysed cycloaddition reaction.

Fmoc azidohomoalanine was successfully formed in 2 steps starting from Fmoc Glutamine; a Hoffmann rearrangement followed by a diazo transfer,²⁵ with a third synthetic step being required to form the diazo transfer reagent²⁶ (Scheme 4.14).



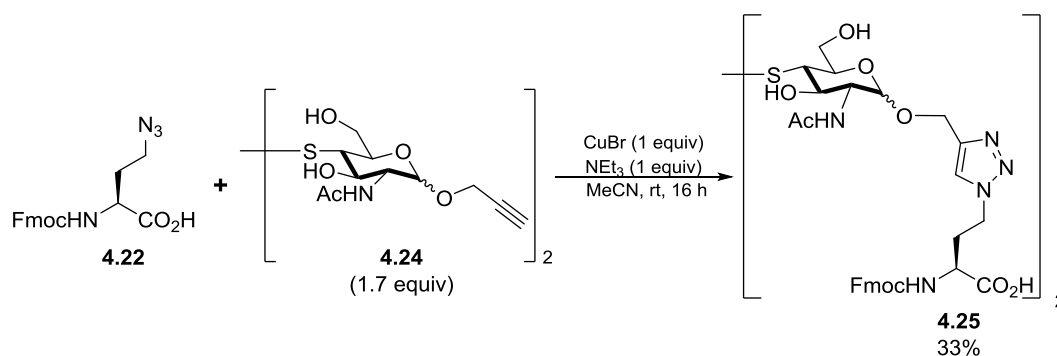
Scheme 4.14: Synthesis of Fmoc-azidohomoalanine (4.22).^{25, 26}

For the synthesis of the sugar, it was envisaged that it would be possible to start from the acetylated thiol GlcNAc (**4.9**) and substitute the anomeric acetyl group for a propargyl group, followed by deprotection. The initial attempt at this substitution used boron trifluoride diethyl etherate, as described for this transformation on glucose.²⁷ However, this returned just starting material. An alternative method was found using Yb(OTf)₃²⁸, giving the desired product (**4.23**, Scheme 4.15). Deprotection with sodium methoxide afforded the desired sugar unit (**4.24**).



Scheme 4.15: Formation of the thiol-GlcNAc alkyne (4.24).²⁸

The amino acid and the sugar unit were successfully coupled using a copper-catalysed cycloaddition reaction and purified using HPLC (Scheme 4.16 and Figure 4.7), resulting in the formation of the desired thiol glycan acceptor.



Scheme 4.16: Coupling of the amino acid and sugar to give the model system.²⁴

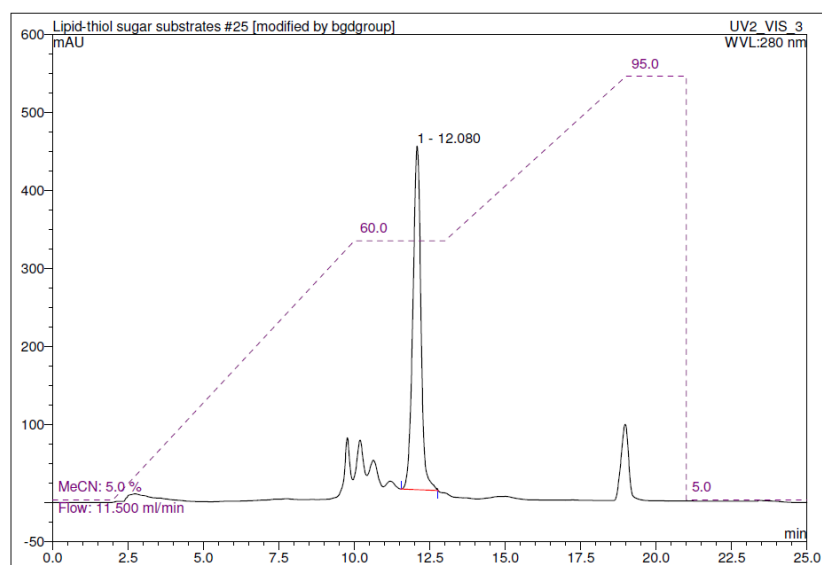


Figure 4.7: HPLC trace at 280 nm of the purification of **4.25**. The retention time of the product is 12 min.

4.4.2 – Investigation of EndoA Catalysed Glycosylation with the Model System

With the model system now in hand, the glycosylation reaction with EndoA could be investigated. Again, a reducing agent had to be used, as the model system formed disulfide dimers. TCEP was used, due to its activity over the same pH range as the activity of EndoA. EndoA is most active at pH 6 - 6.5, whereas DTT requires pH above 7 to be active for disulfide reduction. First, a control reaction was set up using the same conditions as optimised previously for EndoA glycosylation using GlcNAc-triazole-amino acid acceptor (**3.31**), adding TCEP to the buffer to confirm that it did not affect

the activity of EndoA. Monitoring by HPLC, the formation of the desired glycosylation product was observed and TCEP had very little effect on the activity of EndoA (Figure 4.8).

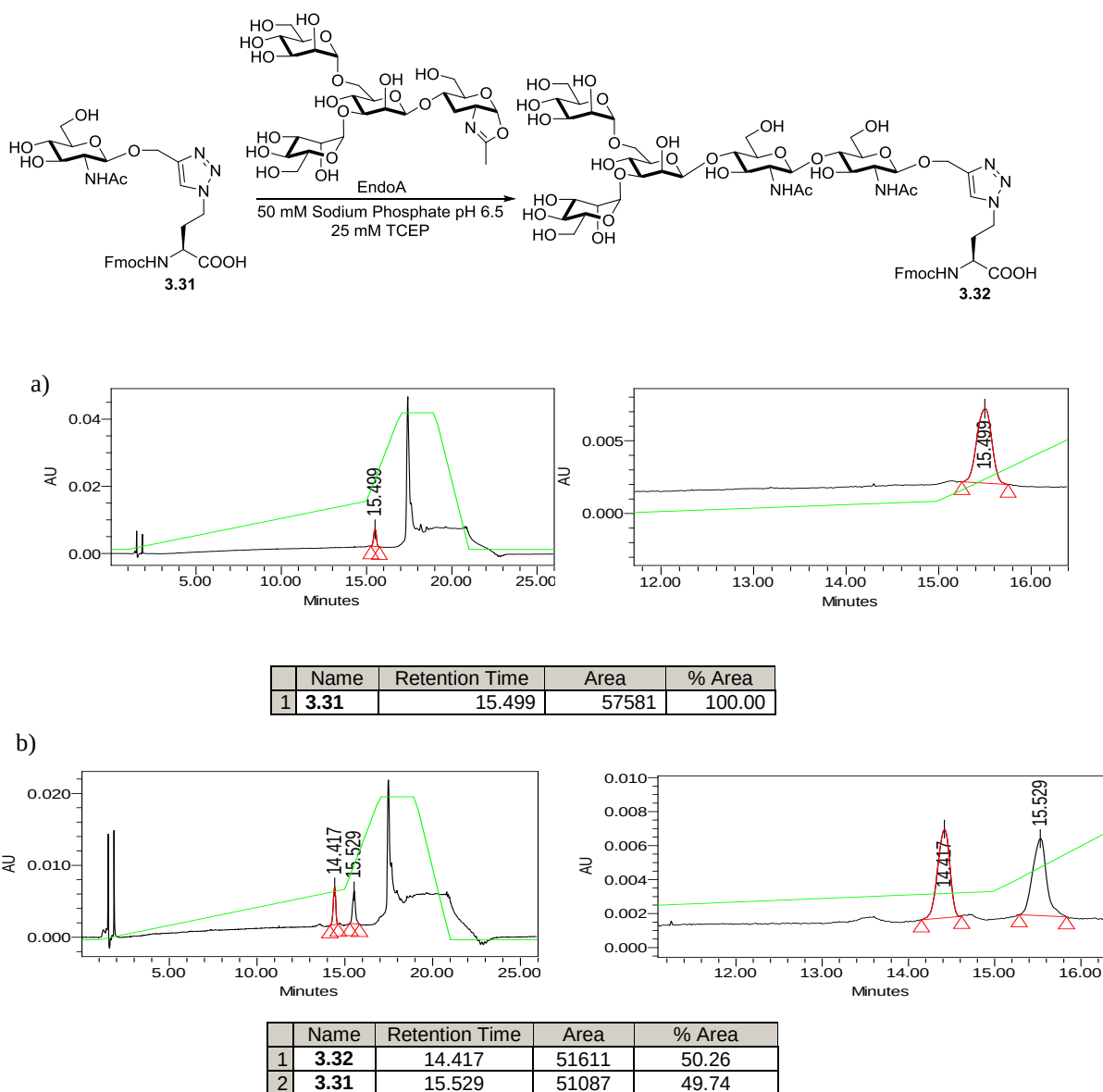


Figure 4.8: HPLC traces at 260 nm of the glycosylation reaction with **3.31** and EndoA, a) HPLC trace of **3.31**, b) Glycosylation reaction after 2 h, showing formation of the desired glycosylated product (**3.32**). A negative reaction without EndoA showed no reaction. The green line on the trace is the percentage amount of acetonitrile.

Confirming that TCEP had no effect on EndoA activity, next the thiol-GlcNAc substrate (4.25) was tested for glycosylation, using the same conditions as above. Although the synthesised thiol glycan acceptor was a mix of anomers, it was expected that reactivity would be seen if EndoA could glycosylate the thiol acceptor. This reaction was again monitored by HPLC (Figure 4.9).

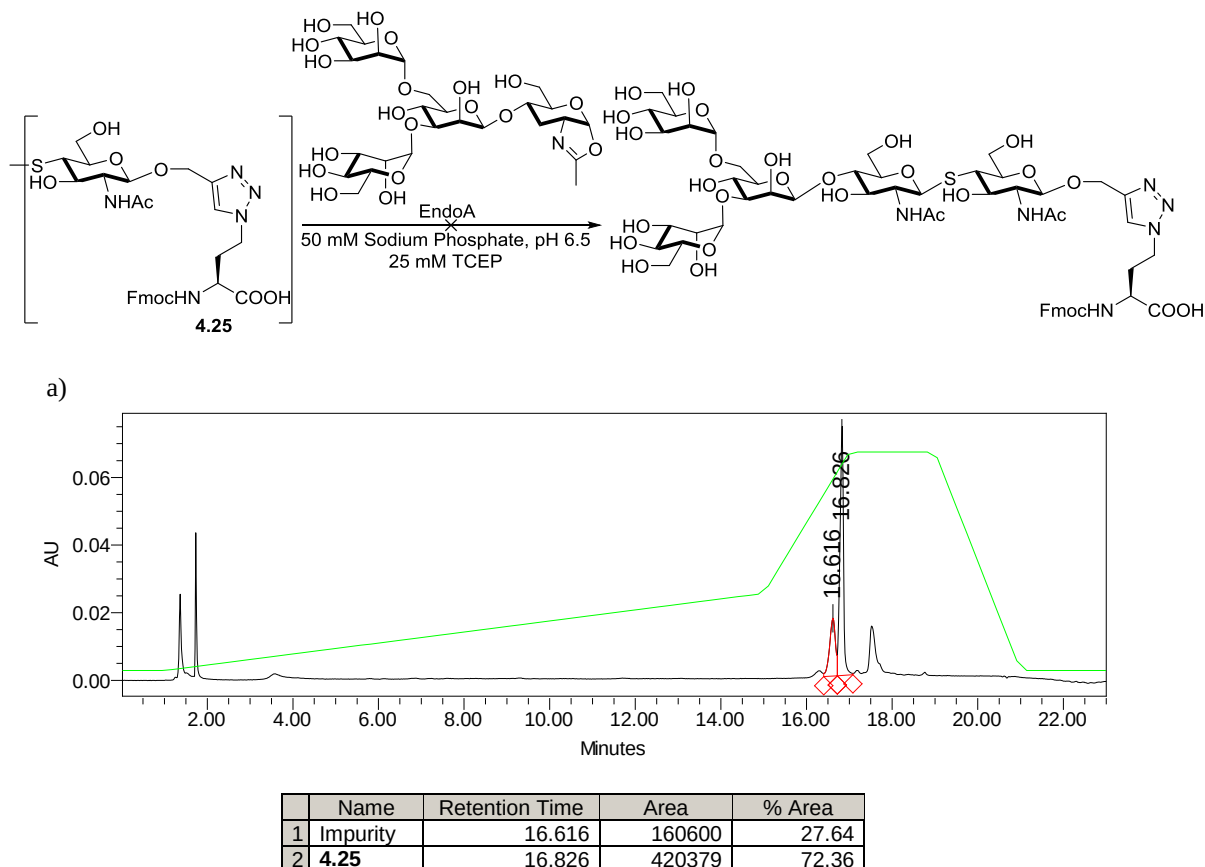


Figure 4.9: HPLC traces at 260 nm of the glycosylation reaction with 4.25 and EndoA, a) HPLC trace of 4.25 (high concentration). The green line on each trace is the percentage amount of acetonitrile.

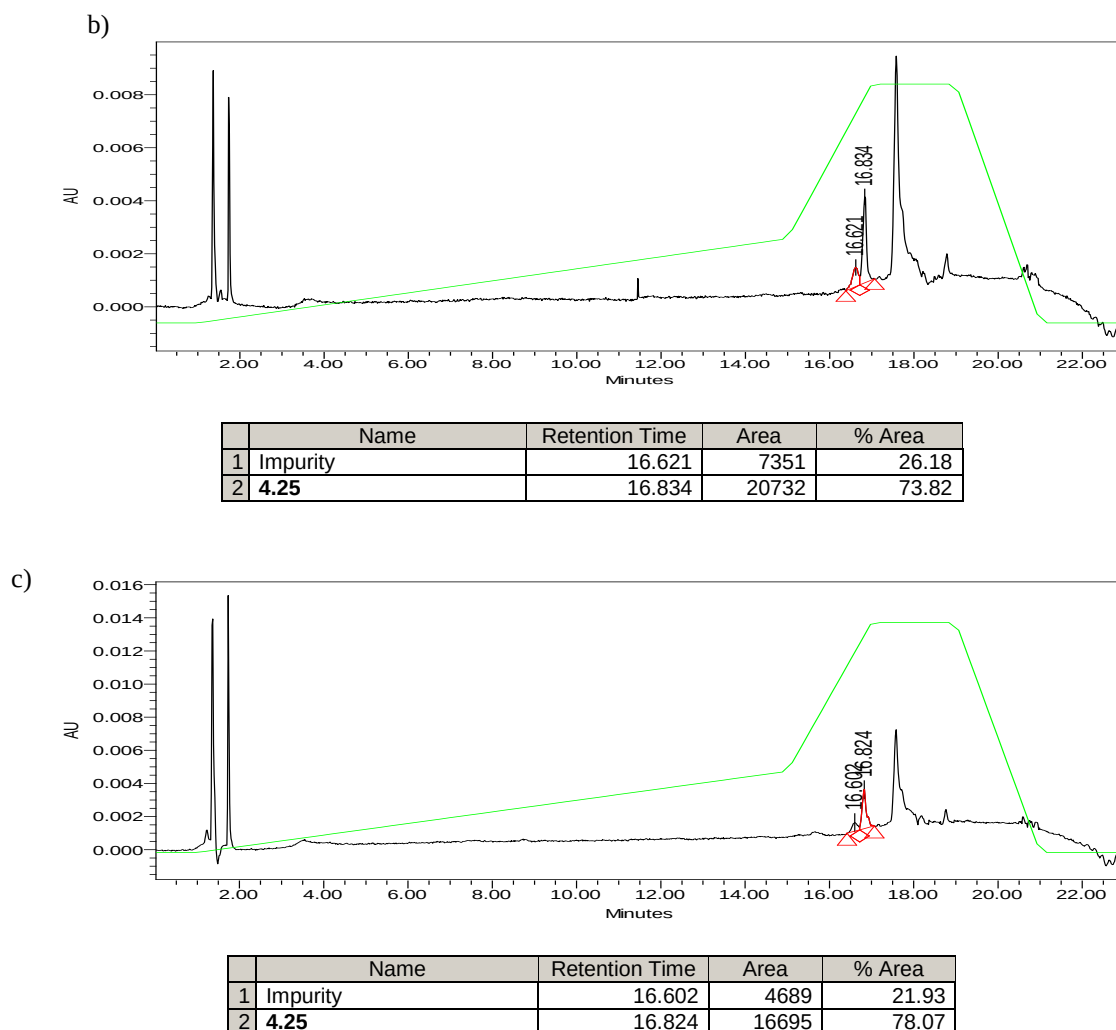


Figure 4.9 cont'd: HPLC traces at 260 nm of the glycosylation reaction with 4.25 and EndoA, b) Glycosylation reaction after 2 h and c) Glycosylation reaction after 24 h. A negative reaction without EndoA also showed no reaction. The green line on the trace is the percentage amount of acetonitrile.

Unfortunately, no consumption of the starting material was observed after 2 h or 24 h. Considering that these conditions had been previously optimised (see Chapter 3) and EndoA glycosylation reactions are typically conducted at pH 6 – 7,^{24, 29-32} there was little scope for optimisation of this reaction. Hence, results here suggest that EndoA is

unable to process the thiol acceptor and thus unable to catalysis the formation of a thioglycosidic bond.

4.5 – Towards the Synthesis of PglB Substrates Modified at the 3-Position

4.5.1 – Targets and Synthetic Strategy

The difficulties in the synthesis of GlcNAc sugars modified at the 4-position (Scheme 4.12) and their use as PglB substrates has been described above. Previous research has also shown that PglB substrates modified at the 2-position and the 6-position of the GlcNAc do not function as PglB substrates.²³ As the anomeric position has the pyrophosphate lipid attached, the only location not yet explored for modification was the 3-position.

The 3-position on the GlcNAc of the PglB substrates would seem a sensible choice for modification. On the natural substrate for PglB, the oligosaccharide is attached to the 3-position on the reducing end GlcNAc and the crystal structure of PglB shows that steric hindrance is not an issue for modification at the 3-position, compared to the 2- and 6-position, where these positions seem to tightly fit in the PglB active site.³³

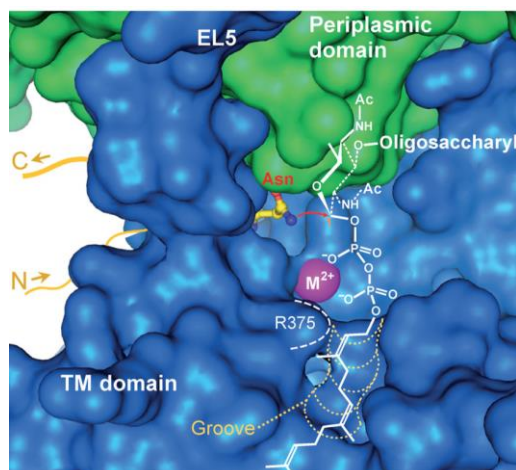


Figure 4.10: Crystal structure of PglB with the glycan pyrophosphate lipid substrate bound (through crystal structure of PglB with pyrophosphate bound and computational modelling). Note the sterics around the 2- and 6-positions and the space around the 3-position. This figure is taken from Lizak *et al.*.³³

The 3-position modified glycan targets chosen to study were a disaccharide and a GlcNAc unit with a chemical tag attached via a linkage chain (Figure 4.11). The disaccharide was chosen as the previous PglB catalysed reaction have all involved the addition of a monosaccharide (GlcNAc). The investigation of the addition of a disaccharide would be the next logical step in the evolution of the PglB catalysed glycosylation reaction. The 3-position tagged glycan, on the other hand, would complement previous studies with the azide tagged 2- and 6-position glycan PglB non-substrates. Also, if it was shown to be possible to attach a glycan with a tag attached using PglB, this would allow the selective modification of proteins without the insertion of unnatural amino acids, something which remains a challenge. The azide group was chosen as the tag due to its reaction with alkynes adhering to “Click Chemistry” principles³⁴ and previous reports of the copper catalysed azide-alkyne cycloaddition on proteins.

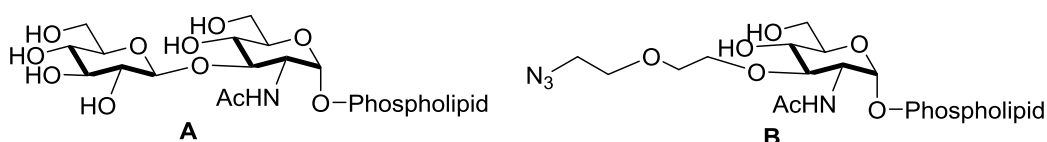
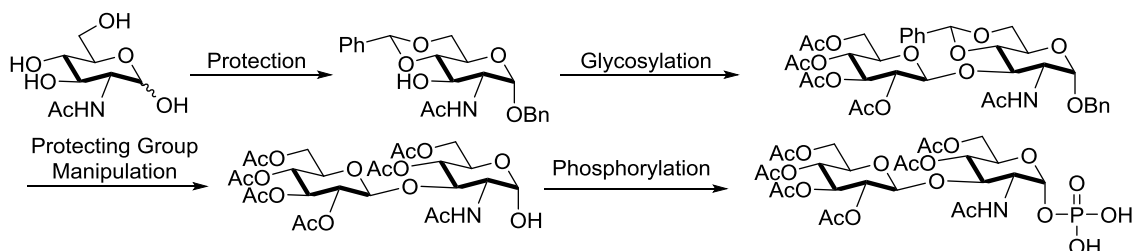


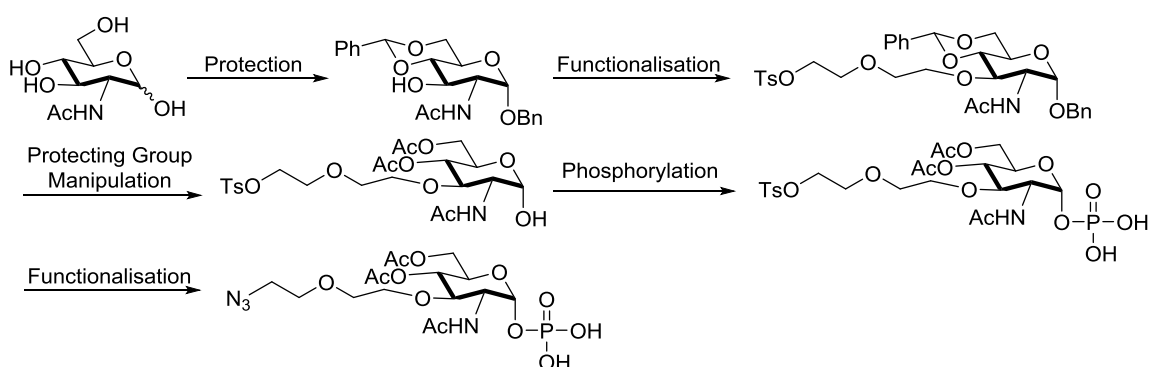
Figure 4.11: Synthetic glycan targets, modified at the 3-position, A = Disaccharide, B = Tagged glycan.

The strategy for the synthesis of both these glycans was similar to that described for the synthesis of the thiol-GlcNAc given above; protect all positions bar the 3-position, functionalisation (glycosylation to give the disaccharide or addition of a reactive short chain to give the tagged glycan), protecting group manipulation to leave the anomeric position unprotected, followed by phosphorylation (Scheme 4.17). For the tagged glycan, a final reaction is required to install the azide (Scheme 4.18). Once both the

desired glycans have been formed, coupling to the phospholipid can be done as before to give the PglB substrate.



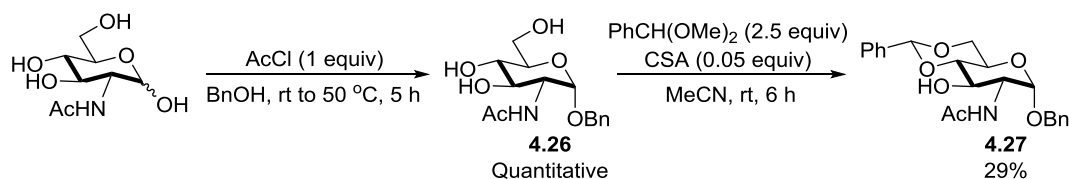
Scheme 4.17: Strategy for the synthesis of the 1→3 disaccharide phosphate.



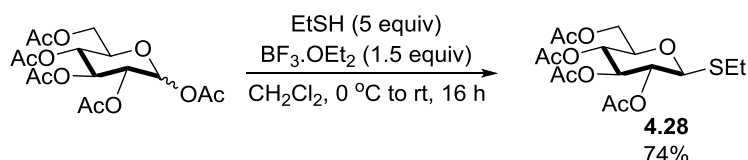
Scheme 4.18: Strategy for the synthesis of a glycan with an azide tag at the 3-position.

4.5.2 – Work Towards Disaccharide Synthesis

Both glycans required the synthesis of a protected GlcNAc unit, with all positions protected, bar the 3-position. Following work by Yamaguchi *et al.*,³⁵ this was done by first selectively benzylating the anomeric position of GlcNAc (**4.26**, Scheme 4.19), followed by benzylidene formation, leaving the 3-position free for the glycosylation reaction (**4.27**). The thiol donor for the glycosylation reaction was formed in one step from peracetylated glucose³⁶ (**4.28**, Scheme 4.20).

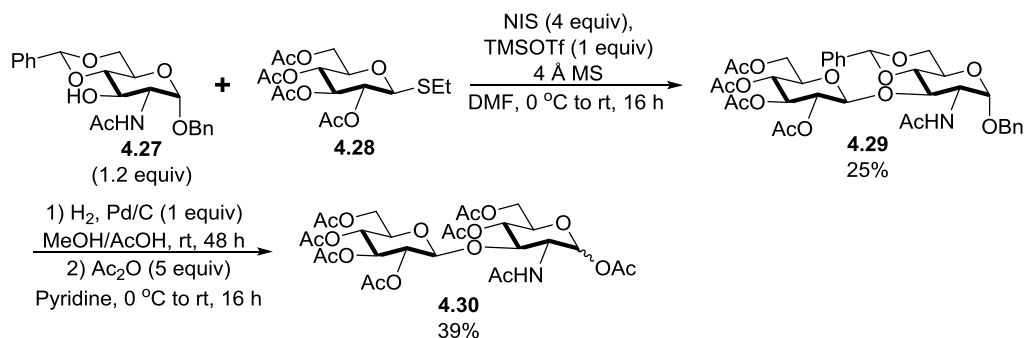


Scheme 4.19: Formation of the protected GlcNAc unit (4.27).^{16, 35}

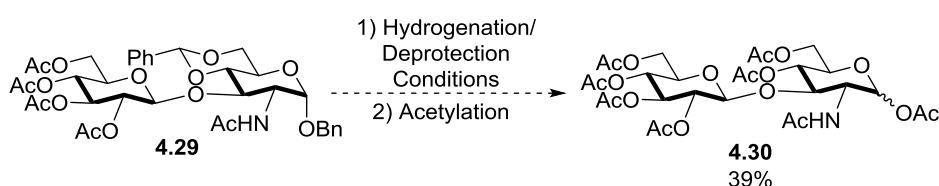


Scheme 4.20: Formation of the thiol donor (4.28).³⁶

With acceptor and donor in hand, the glycosylation reaction was attempted, following conditions reported by *Wen et al.*,³⁷ using the *N*-iodosuccinimide and trimethylsilyl trifluoromethanesulfonate activation system. This gave the desired disaccharide, although in a low yield (**4.29**, Scheme 4.21). Slight variations in the reaction were investigated but no great increase in yield was observed. The major by-product of the reaction was peracetylated disaccharide (as observed by mass spectroscopy), suggesting that greater yield maybe achieved using different protecting groups on the donor. The next step, the removal of the benzyl and benzylidene, proved to more difficult than expected; no product was seen when using standard hydrogenation conditions. A variety of different methods were tried (Table 4.3), but almost all conditions returned starting material or partial deprotection. The only conditions that gave the desired product was Pd/C and H₂ gas in 50% acetic acid in methanol for 48 h (**4.30**). The yield of this reaction was low (after acetylation) due to cleavage of the glycosidic bond. Due to this and the fact that the glycosylation reaction was also low yielding, it was decided to move away from the synthesis of this disaccharide.



Scheme 4.21: Glycosylation to 4.29 and hydrogenation and acetylation to give 4.30. Both reactions were poor yielding.³⁷



Hydrogenation Conditions	Result
Pd/C (2 equiv), H ₂ , MeOH, rt, 16 h	No reaction
Pd/C (1 equiv), H ₂ , 9:1, MeOH:AcOH, rt, 48 h	Starting material and cleavage of the glycosidic bond
Pd/C (1 equiv), H ₂ , 1:1, MeOH:AcOH, rt, 48 h ³⁷	Product (39% yield)
Pd/C (2 equiv), Et ₃ SiH (10 equiv), MeOH, rt, 2 h ³⁸	No reaction
Pd(OH) ₂ /C (0.5 equiv), H ₂ , EtOAc, rt, 48 h	No reaction
PtO ₂ (0.1 equiv), H ₂ , THF, 16 h	No reaction
FeCl ₃ (15 equiv), CH ₂ Cl ₂ , rt, 2 h ³⁹	Only deprotection of the benzyl group observed

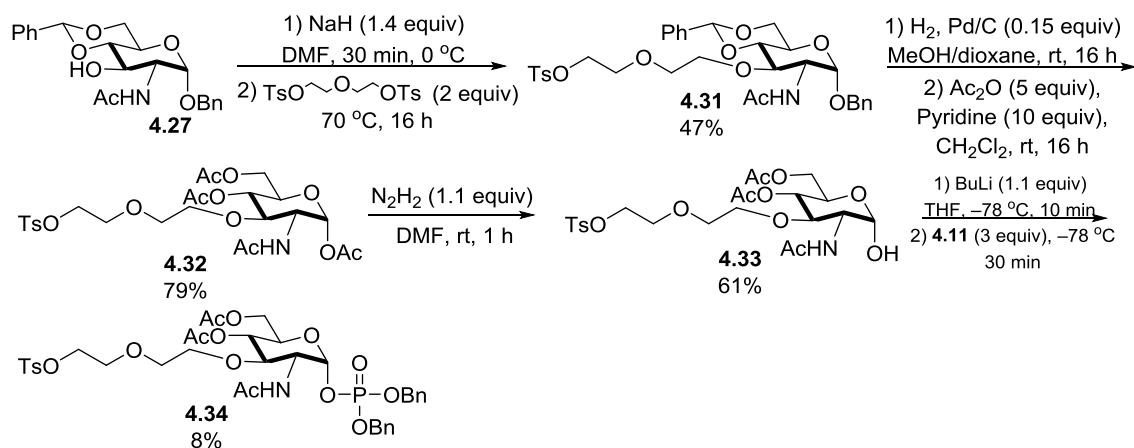
Table 4.3: Conditions investigated for the hydrogenation of 4.29

4.5.3 – Work Towards the Formation of a Tagged GlcNAc

As described above, the strategy for the formation of the tagged GlcNAc was similar to that of the formation of the disaccharide, relying on the formation of the protected GlcNAc first (4.27), which was done as before. For the attachment of an azido group via a short linker, it was envisaged that unprotected 3-position could be deprotonated,

followed by an S_N2 reaction with one end of diethylene glycol di(*p*-tosylate). The remaining tosyl group could then be substituted in a second subsequent S_N2 reaction to install the azide group at the end of the synthesis. Pleasingly, the installation of the diethylene tosyl was successful when using sodium hydride and heating at 70 °C for 16 h, resulting in a moderate yield of the desired product (**4.31**, Scheme 4.22). The following deprotection manipulations, hydrogenation to remove the benzyl and benzyldiene groups followed by acetylation (**4.32**) and then hydrazine deprotection of the anomeric acetyl group²³ (**4.33**), were successfully carried out. It should be noted that the acetylation had to be carried out in dichloromethane; using pyridine as the solvent resulted in degradation of the starting material. The phosphorylation was attempted as before, using the butyl lithium/tetrabenzylpyrophosphate methodology. This lead to the desired product (**4.34**), but, unfortunately, in a low yield. This was thought to be due to the use of butyl lithium, with the compound possibly being unstable under strongly basic conditions, as seen with the compounds precursors. An alternative method using dibenzyl *N,N*-di-*iso*-propylphosphoramidite and mCPBA was investigated, using conditions previously described for the phosphorlyation of similar compounds.²³ Disappointingly, this resulted in degradation. Moving forward, the tetrabenzylpyrophosphate methodology should be attempted again, but using a different base, such as pyridine or trimethylamine.

Chapter 4 – Synthesis and Testing of PglB Substrates with Variation on the Sugar; Investigating 4-Thiol GlcNAc



Scheme 4.22: Attachment of linker with tosylate group on (4.31) and subsequent protecting group manipulations (4.32 and 4.33) and phosphate group attachment (4.34).

4.6 – Conclusions and Future Work

At the beginning of the chapter, it was described how it was desired to form glycoproteins where the glycan is resistant to hydrolysis, via the formation of thioglycosidic linkage. It was thought that this could be achieved through two novel enzymatic reactions; the use of PglB to attach a GlcNAc unit with a thiol at the 4-position, followed by an Endo catalysed glycosylation using an oxazoline donor and the thiol acceptor (Scheme 4.5). With regards to the first PglB reaction, two novel substrates with 4-thiol GlcNAc were successfully synthesised. Disappointingly, the desired product was not observed in PglB catalysed glycosylation reaction. The reason for this lack of reaction could be due to the increase in the atomic radius at the 4-position, going from 60 pm for oxygen (GlcNAc) or 65 pm for nitrogen (the natural PglB substrate) to 100 pm for sulfur. This possibly causes a steric clash in the active site of PglB and thus the substrate is unable to fit. One interesting result of note here was the seen addition of 4-deoxy-GlcNAc to the peptide catalysed by PglB, demonstrating that certain modifications at the 4-position may be tolerated and that the 4-hydroxyl group on the glycan is not critical for the substrate to be accepted by PglB.

With regards to the second enzymatic reaction, a model thiol acceptor was formed and tested with EndoA and an oxazoline donor. Disappointingly, the desired product was not observed and no reaction detected. Once more, the larger atomic radius of the sulfur atom on the acceptor could be causing a steric clash inside the active site of EndoA, and hence EndoA is not able to catalyse the reaction. Furthermore, in previously reported reactions when a glycoside hydrolyase was used to catalyse the formation of a thioglycoside, the glycan donor had a leaving/activating group attached and the

glycoside hydrolyase used proceeds via a different mechanism compared to EndoA (via an enzyme substrate-complex instead of an oxazolinium ion intermediate). These two factors maybe critical for the desired reaction to occur.

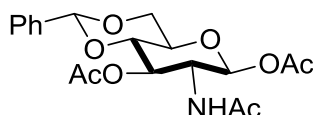
For the synthesis of a glycan modified with a tag (azide) at the 3-position for the formation of a PglB substrate, the addition of the phosphate group should be attempted again using tetrabenzylpyrophosphate but with a weaker base, such as triethylamine or pyridine. This should hopefully result in a greater yield of the desired phosphated product, which can then be used to complete the synthesis of the tagged PglB substrate.

4.7- Chemistry Experimental

See Chapter 3 for General Chemistry Experimental.

2-*N*-acetamido-2-deoxy-1,3-di-*O*-acetyl-4,6-*O*-benzylidene- β -D-glucopyranose

(4.1)⁴⁰



A solution *N*-acetylglucosamine (17.7 g, 80 mmol), benzaldehyde dimethyl acetal (7.1 mL, 48 mmol) and camphorsulfonic acid (1.83 g, 8 mmol) in dry DMF (100 mL) was stirred on a rotary evaporator at 50 mbar, 50 °C for 6 h. After this time, the solvent was removed under reduced pressure and pyridine (120 mL) and acetic anhydride (20.8 mL, 224 mmol) were both added. The reaction was stirred at room temperature for 16 h before the solvent was removed under reduced pressure. Ethanol (100 mL) was added and the resulting solid was filtered off and washed a second time with ethanol (100 mL) to afford the protected sugar product (**4.1**, 15.1 g, 48%) as a white solid, exclusively as the β anomer, and was used without further purification. *R_f* 0.39 (1:1, pet. ether:ethyl acetate). $[\alpha]_D -36.4$ (c 0.25, CHCl₃). ¹H NMR (400 MHz, CDCl₃) δ 7.48 – 7.42 (2 H, m, 2 $\times$ Ph), 7.39 – 7.34 (3 H, m, 3 $\times$ Ph), 5.77 (1 H, d, *J* = 9.3, H-1), 5.70 (1 H, d, *J* = 8.7, CHPh), 5.22 (1 H, t, *J* = 9.9, H-3), 4.38 – 4.31 (2 H, m, H-2 and 1 $\times$ H-6), 3.83 – 3.73 (2 H, m, H-5 and 1 $\times$ H-6), 3.61 (1 H, ddd, *J* = 14.5, 9.6, 4.9, H-4), 2.13 (3 H, s, OAc), 2.10 (3 H, s, OAc) and 1.95 (3 H, s, NHAc). ¹³C NMR (100 MHz, CDCl₃) δ 171.5 (CO), 170.4 (CO), 169.5 (CO), 136.7 (Ph), 129.3 (Ph), 128.3 (2 $\times$ Ph), 126.2 (Ph), 101.5 (CHPh), 93.3 (C-1), 78.1 (C-4), 71.8 (C-3), 68.3 (C-5), 67.4 (C-6), 53.6 (C-2), 23.3

(NHAc), 20.90 (OAc) and 20.89 (OAc). MS (ESI): m/z Calcd for $C_{19}H_{23}NO_8Na$ $[M+Na]^+$: 416.1; Found 416.1. MP = 246 – 248 °C.

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.⁴⁰

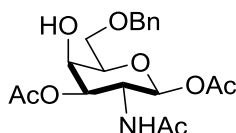
2-*N*-acetamido-2-deoxy-1,3-di-*O*-acetyl-6-*O*-benzyl- β -D-glucopyranose (**4.2**)



A solution of **4.1** (0.2 g, 0.5 mmol) and 4 Å molecular sieves (50 mg) in dry CH_2Cl_2 (8 mL) was stirred at room temperature for 1 h before being cooled to -78 °C and triethylsilane (0.24 mL, 1.5 mmol) was added, followed by trifluoromethane sulfonic acid (0.15 mL, 1.75 mmol) added dropwise. The reaction was stirred at -78 °C for 3 h before being quenched by the addition of a mix of methanol (1.5 mL) and triethylamine (1.5 mL). The mixture was filtered through Celite and diluted with chloroform (50 mL). The mixture was then washed with saturated sodium bicarbonate (20 mL) and brine (20 mL) and dried with sodium sulfate. The solvent was removed under reduced pressure and purification by flash chromatography (9:1, chloroform:methanol) afforded the protected sugar (**4.2**, 0.12 g, 0.3 mmol, 60%) as a white solid and exclusively as the β anomer. R_f 0.25 (9:1, chloroform:methanol). $[\alpha]_D -15.6$ (c 0.5, $CHCl_3$). ν_{max} (film)/ cm^{-1} 3293 (O-H), 3089, 2919, 2870, 1746 (C=O), 1663, 1548, 1453, 1432, 1371, 1314, 1222, 1072, 1036, 914, 736 and 700. ^1H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.28 (5 H, m, 5 $\times$ Ph), 6.04 (1 H, d, J = 9.6, H-1), 5.68 (1 H, d, J = 8.7, NH), 5.06 (1 H, dd, J = 10.7 and 9.2, H-3), 4.57 (2 H, quartet, J = 12.1, CH_2Ph), 4.17 (1 H, quartet, J = 9.5, H-2), 3.84 –

3.63 (4 H, m, H-4, H-5 and 2 × H-6), 2.10 (3 H, s, OAc), 2.09 (3 H, s, OAc) and 1.91 (3 H, s, NHAc). ^{13}C NMR (100 MHz, CDCl_3) δ 172.1 (CO), 170.5 (CO), 169.8 (CO), 137.5 (Ph), 128.5 (2 × Ph), 128.0 (Ph), 127.8 (2 × Ph), 92.8 (C-1), 75.3 (CH_2Ph), 75.1 (C-5), 73.8 (C-3), 69.8 (C-4), 69.6 (C-6), 52.9 (C-2), 23.3 (NHAc), 21.0 (2 × OAc). HRMS (ESI): m/z Calcd for $\text{C}_{19}\text{H}_{25}\text{NO}_8\text{Na}$ $[\text{M}+\text{Na}]^+$: 418.1472; Found 418.1467. MP = 169 – 171 °C.

2-N-acetamido-2-deoxy-1,3-di-O-acetyl-6-O-benzyl- β -D-galactopyranoside (4.3)

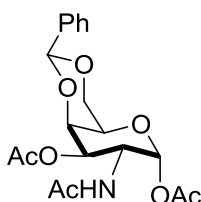


A solution of **4.2** (1.5 g, 3.75 mmol) in dichloromethane (75 mL) was cooled to 0 °C and pyridine (1.8 mL, 11.3 mmol) and trifluoromethanesulfonic anhydride (1.3 mL, 7.5 mmol) were added dropwise. After stirring at 0 °C for 1 h, the reaction was diluted with dichloromethane (75 mL) and washed with ice-cold 1 M HCl (30 mL), saturated sodium bicarbonate (50 mL) and brine (50 mL). The organic layer was dried with sodium sulfate, solvent removed under reduced pressure and the resulting residue re-dissolved in dry DMF (50 mL). 18-Crown-6 (0.3 g, 1.1 mmol) and sodium nitrite (2.6 g, 37.5 mmol) were added and the reaction was stirred at room temperature for 1 h. After this time, the reaction was diluted with ethyl acetate (100 mL) and washed with 1 M HCl (30 mL), saturated sodium bicarbonate (50 mL) and brine (50 mL). The organic layer was dried with sodium sulfate, solvent removed under reduced pressure and purification by flash chromatography (9:1, chloroform:methanol) followed by recrystallisation from methanol afforded the protected sugar (**4.3**, 175 mg, 0.44 mmol, 12%) as a white solid and exclusively as the β anomer. R_f 0.41 (9:1, chloroform:methanol). $[\alpha]_D^{+15.4}$ (c 0.5,

CHCl_3). ν_{max} (film)/ cm^{-1} 3323 (O-H), 2870, 2360, 2342, 1745 (C=O), 1660, 1555, 1454, 1369, 1309, 1222, 1077, 1039, 913, 766 and 734. ^1H NMR (400 MHz, CDCl_3) δ 7.38 – 7.29 (5 H, m, $5 \times \text{Ph}$), 5.68 (1 H, d, $J = 8.9$, H-1), 5.49 (1 H, d, $J = 9.5$, NH), 5.01 (1 H, dd, $J = 11.0, 2.9$, H-3), 4.62 – 4.50 (3 H, m, H-2 and CH_2Ph), 4.16 – 4.13 (1 H, m, H-4), 3.84 – 3.72 (3 H, m, H-5 and $2 \times \text{H-6}$), 2.91 (1 H, d, $J = 2.9$, OH), 2.13 (3 H, s, OAc), 2.12 (3 H, s, OAc) and 1.93 (3 H, s, NHAc). ^{13}C NMR (100 MHz, CDCl_3) δ 170.2 ($2 \times \text{CO}$), 169.7 (NHCO), 137.3 (Ph), 128.5 ($2 \times \text{Ph}$), 128.0 (Ph), 127.8 ($2 \times \text{Ph}$), 93.3 (C-1), 75.8 (C-5 and CH_2Ph), 72.8 (C-3), 68.9 (C-4), 67.4 (C-6), 49.7 (C-2), 23.3 (NHAc), 21.0 (OAc) and 20.9 (OAc). HRMS (ESI): m/z Calcd for $\text{C}_{19}\text{H}_{25}\text{NO}_8\text{Na}$ $[\text{M}+\text{Na}]^+$: 418.1472; Found 418.1468. MP = 170 – 173 $^\circ\text{C}$.

2-N-acetamido-2-deoxy-1,3-di-O-acetyl-4,6-O-benzylidene- α -D-galactopyranose

(4.4)²¹

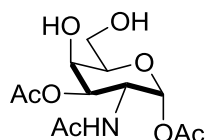


A solution of *N*-acetylgalactosamine (3 g, 12.9 mmol), benzaldehyde dimethyl acetal (4.9 mL, 32.2 mmol) and camphorsulfonic acid (150 mg, 0.65 mmol) in dry acetonitrile (40 mL) was stirred at room temperature for 8 h. After this time the solvent was removed under reduced pressure and pyridine (30 mL) and acetic anhydride (3.9 mL, 41.9 mmol) were added. The reaction was stirred at room temperature for 16 h before the solvent was removed under reduced pressure. The residue was dissolved in chloroform (100 mL) and washed with 1 M HCl (20 mL), saturated sodium bicarbonate (20 mL) and brine (20 mL). The organic layer was dried with sodium sulfate and the

solvent removed under reduced pressure. The residue was washed with ethanol (100 mL) and the resulting solid was filtered off and washed a second time with ethanol (100 mL) to afford the protected sugar product (**4.4**, 3.5 g 8.9 mmol, 69%) as a white solid, exclusively as the α anomer, that was used without further purification. *R_f* 0.48 (20:1, chloroform:methanol). $[\alpha]_D +76.4$ (c 0.25, CHCl₃). ¹H NMR (400 MHz, CDCl₃) δ 7.56 – 7.52 (2 H, m, 2 $\times$ Ph), 7.42 – 7.36 (3 H, m, 3 $\times$ Ph), 6.34 (1 H, d, *J* = 3.5, H-1), 5.56 (1 H, d, *J* = 8.7, CHPh), 5.47 (1 H, d, *J* = 9.0, NH), 5.27 (1 H, dd, *J* = 11.6, 3.2, H-3), 4.92 (1 H, ddd, *J* = 11.8, 9.2, 3.5, H-2), 4.35 (1 H, d, *J* = 2.8, H-4), 4.29 (1 H, dd, *J* = 12.7, 1.4, 1 $\times$ H-6), 4.05 (1 H, dd, *J* = 12.7, 1.7, 1 $\times$ H-6), 3.82 (1 H, s, H-5) 2.18 (3 H, s, OAc), 2.13 (3 H, s, OAc) and 1.95 (3 H, s, NHAc). ¹³C NMR (100 MHz, CDCl₃) δ 172.1 (CO), 170.0 (CO), 169.0 (NCO), 137.3 (Ph), 129.2 (Ph), 128.3 (2 $\times$ Ph), 126.4 (2 $\times$ Ph), 101.0 (CHPh), 92.2 (C-1), 73.3 (C-4), 69.0 (C-6), 68.5 (C-3), 64.5 (C-5), 46.9 (C-2), 23.3 (NHAc) and 21.1 (2 $\times$ OAc). MS (ESI): *m/z* Calcd for C₁₉H₂₃NO₈Na [M+Na]⁺: 416.1: Found 416.1. MP = 227 – 229 °C.

The observed characterisation data (¹H and ¹³C) was consistent with that previously reported in the literature.²¹

2-*N*-acetamido-2-deoxy-1,3-di-*O*-acetyl- α -D-glucopyranose (**4.5**)²¹

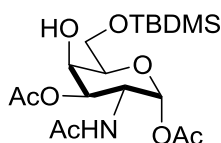


To **4.4** (3.5 g, 8.9 mmol) and Pd/C (600 mg, 10% loading, 5.7 mmol) was added dry methanol (20 mL) and dry 1,4-dioxane (20 mL). The reaction was put under an atmosphere of hydrogen and stirred at room temperature for 16 h. The reaction was then

filtered through Celite and the solvent removed under reduced pressure to afford the protected sugar (**4.5**, 0.42 g, 1.3 mmol, 55%) as a colourless oil, exclusively as the α anomer, that was used without further purification. R_f 0.23 (7:1, chloroform:methanol). $[\alpha]_D^{+126.0}$ (c 1.0, MeOH). ^1H NMR (400 MHz, CDCl_3) δ 6.19 (1 H, d, $J = 3.7$, H-1), 5.76 (1 H, d, $J = 9.2$, NH), 5.14 (1 H, dd, $J = 11.3, 2.8$, H-3), 4.82 (1 H, ddd, $J = 12.8, 9.3, 3.8$, H-2), 4.19 (1 H, d, $J = 2.6$, H-4), 3.92 – 3.80 (3 H, m, H-5 and $2 \times$ H-6), 2.15 (3 H, s, OAc), 2.13 (3 H, s, OAc) and 1.94 (3 H, s, NHAc). ^{13}C NMR 171.7 (CO), 170.6 (CO), 169.4 (NCO), 91.7 (C-1), 72.0 (C-6), 70.5 (C-3), 68.2 (C-4), 62.5 (C-5), 46.9 (C-2), 23.2 (NHAc), 21.1 (OAc) and 21.0 (OAc). MS (ESI): m/z Calcd for $\text{C}_{12}\text{H}_{19}\text{NO}_8\text{Na}$ $[\text{M}+\text{Na}]^+$: 328.1: Found 328.1.

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.²¹

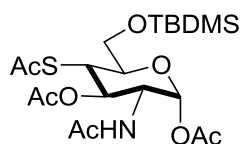
2-*N*-acetamido-2-deoxy-1,3-di-*O*-acetyl-6-*O*-*tert*-butyldimethylsilyl- α -D-galactopyranose (4.6**)**



To **4.4** (3 g, 7.6 mmol) and Pd/C (400 mg, 10% loading, 3.8 mmol) was added dry methanol (20 mL) and dry 1,4-dioxane (20 mL). The reaction was put under an atmosphere of hydrogen and was stirred at room temperature for 16 h. The reaction was then filtered through Celite and the solvent removed under reduced pressure. Imidazole (1 g, 15.2 mmol), dry dichloromethane (25 mL) and *tert*-butyldimethylsilyl chloride were all added and the reaction was stirred at room temperature for 3 h. The reaction

was then washed with brine (50 mL) and the aqueous layer was extracted with chloroform (75 mL). The combined organic layers were dried with sodium sulfate and the solvent removed under reduced pressure. Purification by flash chromatography (EtOAc) afforded the protected sugar (**4.6**, 2.7 g, 6.4 mmol, 85%) as a white solid and exclusively as the α anomer. R_f 0.23 (EtOAc). $[\alpha]_D^{+25} +94.5$ (c 1.0, CHCl_3). ν_{max} (film)/ cm^{-1} 3286 (O-H), 1743 (C=O), 1659, 1545, 1375, 1226, 1182, 1131, 1038, 1012, 791, 735 and 700. ^1H NMR (400 MHz, CDCl_3) δ 6.22 (1 H, d, J = 3.7, H-1), 5.54 (1 H, d, J = 9.3, NH), 5.16 (1 H, dd, J = 11.3, 2.9, H-3), 4.85 (1 H, ddd, J = 11.3, 7.0, 3.8, H-2), 4.23 (1 H, d, J = 2.6, H-4), 3.94 (1 H, ddd, J = 12.2, 7.2, 2.9, H-5), 3.86 (1 H, d, J = 3.8, 1 $\times$ H-6), 3.83 (1 H, d, J = 3.8, 1 $\times$ H-6), 2.16 (3 H, s, OAc), 2.15 (3 H, s, OAc), 1.94 (3 H, s, NHAc), 0.89 (9 H, s, $\text{C}(\text{CH}_3)_3$), 0.09 (3 H, s, $\text{Si}(\text{CH}_3)_2$) and 0.08 (3 H, s, $\text{Si}(\text{CH}_3)_2$). ^{13}C NMR (100 MHz, CDCl_3) δ 171.8 (CO), 170.1 (CO), 169.1 (NCO), 91.8 (C-1), 71.3 (C-6), 70.7 (C-3), 68.2 (C-4), 63.2 (C-5), 47.0 (C-2), 25.8 ($\text{C}(\text{CH}_3)_3$), 23.2 (NHAc), 21.1 (2 $\times$ OAc), 18.2 ($\text{C}(\text{CH}_3)_3$) and -5.6 (2 $\times$ $\text{Si}(\text{CH}_3)_2$). HRMS (ESI): m/z Calcd for $\text{C}_{18}\text{H}_{33}\text{NO}_8\text{SiNa}$ $[\text{M}+\text{Na}]^+$: 442.1868; Found 442.1854. MP = 85 – 87 $^\circ\text{C}$.

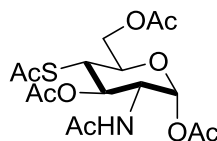
2-N-acetamido-2-deoxy-1,3-di-O-acetyl-4-thio-4-S-acetyl--4-deoxy-6-O-tert-butyltrimethylsilyl- α -D-glucopyranose (4.7**)**



A solution of **4.6** (0.4 g, 0.95 mmol) in dry dichloromethane (5 mL) and dry pyridine (0.46 mL, 5.73 mmol) was cooled to 0 $^\circ\text{C}$ and triflic anhydride (0.32 mL, 1.91 mmol) was added dropwise. The reaction was stirred at room temperature for 1.5 h before being diluted with dichloromethane (20 mL) and washed with ice cold 1 M HCl (5 mL),

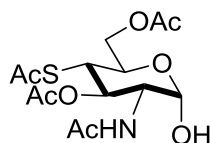
saturated sodium bicarbonate (5 mL) and brine (5 mL). The organic layer was then dried with sodium sulfate and the solvent removed under reduced pressure. The resulting residue was dissolved in dry DMF (4 mL) and potassium thioacetate (326 mg, 2.86 mmol) was added and the reaction was stirred at room temperature for 1.5 h. The reaction was then diluted with dichloromethane (50 mL) and washed with brine (10 mL). The organic layer was dried with sodium sulfate and the solvent removed under reduced pressure. Purification by flash chromatography (1:1, pet. ether:ethyl acetate to 100% ethyl acetate) afforded the thio sugar (**4.7**, 0.19 g, 0.4 mmol, 42%) as an orange foam and exclusively as the α anomer. *R_f* 0.41 (Ethyl acetate). $[\alpha]_D^{+25} +54.4$ (c 1.0, CHCl₃). $\nu_{\max}$ (film)/cm⁻¹ 2930, 1749 (C=O), 1661, 1541, 1372, 1222, 1124, 1011, 940, 836, 779, 732 and 623. ¹H NMR (400 MHz, CDCl₃) δ 6.20 (1 H, d, *J* = 3.7, H-1), 5.56 (1 H, d, *J* = 9.2, NH), 5.51 (1 H, t, *J* = 10.8, H-3), 4.39 (1 H, ddd, *J* = 10.5, 9.3, 3.7, H-2), 4.00 (1 H, ddd, *J* = 11.3, 4.3, 1.8, 1 $\times$ H-6), 3.86 – 3.68 (3 H, m, H-4, H-5 and 1 $\times$ H-6), 2.33 (3 H, s, SAc), 2.18 (3 H, s, OAc), 2.05 (3 H, s, OAc), 1.93 (3 H, s, NHAc), 0.87 (9 H, s, C(CH₃)₃), 0.04 (3 H, s, Si(CH₃)₂) and 0.02 (3 H, s, Si(CH₃)₂). ¹³C NMR (100 MHz, CDCl₃) δ 192.6 (SCO) 171.7 (CO), 171.0 (CO), 168.9 (NCO), 91.2 (C-1), 72.8 (C-5), 69.8 (C-3), 62.8 (C-6), 52.2 (C-2), 43.6 (C-4), 30.8 (SAc), 25.8 (C(CH₃)₃), 21.1 (NHAc), 21.0 (OAc), 20.7 (OAc), 18.2 (C(CH₃)₃), -5.3 (Si(CH₃)₂) and -5.4 (Si(CH₃)₂). HRMS (ESI): *m/z* Calcd for C₂₀H₃₅NSSiO₈Na [M+Na]⁺: 500.1745: Found 500.1739.

2-N-acetamido-2-deoxy-1,3,6-tri-O-acetyl-4-thio-4-S-acetyl--4-deoxy- α -D-glucopyranose (4.9)



To **4.7** (0.67 g, 1.4 mmol) in dry dichloromethane (25 mL) was added acetic anhydride (5 mL) and copper triflate (0.25 g, 0.7 mmol). The reaction was stirred at room temperature for 3.5 h after which it was diluted with dichloromethane (50 mL) and washed with saturated sodium bicarbonate (30 mL). The aqueous layer was extracted with chloroform (30 mL) and the combined organic layers were washed with brine (30 mL). The organic layer was then dried with sodium sulfate and the solvent removed under reduced pressure. Purification by flash chromatography (ethyl acetate) afforded the thio sugar (**4.9**, 0.42 g, 1 mmol, 75%) as a light brown foam and exclusively as the α anomer. R_f 0.31 (Ethyl acetate). $[\alpha]_D^{+25} +93.6$ (c 1.0, CHCl_3). ν_{max} (film)/ cm^{-1} 1745 (C=O), 1704, 1665, 1539, 1433, 1372, 1223, 1120, 1073, 1031, 1013, 942, 733 and 623. ^1H NMR (400 MHz, CDCl_3) δ 6.18 (1 H, d, $J = 3.7$, H-1), 5.72 (1 H, d, $J = 9.2$, NH), 5.22 (1 H, t, $J = 11.1$, H-3), 4.46 (1 H, dt, $J = 9.3, 3.7$, H-2), 4.37 (1 H, dd, $J = 12.5, 4.3$, 1 $\times$ H-6), 4.13 – 4.06 (2 H, m, H-5 and 1 $\times$ H-6), 3.80 (1 H, t, $J = 11.3$, H-4), 2.32 (3 H, s, SAc), 2.18 (3 H, s, OAc), 2.07 (3 H, s, OAc), 2.03 (3 H, s, OAc) and 1.92 (3 H, s, NHAc). ^{13}C NMR (100 MHz, CDCl_3) δ 192.1 (SCO) 171.5 (CO), 170.8 (CO), 170.1 (CO) 168.8 (NCO), 91.0 (C-1), 70.5 (C-5), 69.2 (C-3), 62.7 (C-6), 51.9 (C-2), 43.2 (C-4), 30.7 (SAc), 23.0 (NHAc), 20.9 (OAc), 20.7 (OAc) and 20.6 (OAc). HRMS (ESI): m/z Calcd for $\text{C}_{16}\text{H}_{23}\text{NSO}_9\text{Na}$ $[\text{M}+\text{Na}]^+$: 428.0986 Found 428.0982. MP = 58 – 60 $^\circ\text{C}$.

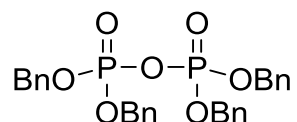
2-N-acetamido-2-deoxy-3,6-di-O-acetyl-4-thio-4-S-acetyl-4-deoxy- α -D-glucopyranose (4.10)



To a solution **4.9** (0.42 g, 1 mmol) in dry dichloromethane (10 mL) and dry ethyl acetate (1 mL) was added titanium tetrabromide (1.42 g, 3.9 mmol) and the reaction was stirred at room temperature for 6 h. After this time, the reaction was filtered through Celite and washed with saturated sodium bicarbonate (10 mL). The aqueous layer was extracted with chloroform (10 mL) and the combined organic layers were dried with sodium sulfate and the solvent removed under reduced pressure. The resulting residue was dissolved in acetone (15 mL) and water (7.5 mL) and silver carbonate (0.29 g, 1 mmol) was added and the reaction was stirred for 16 h at room temperature in the dark. The reaction was filtered through Celite and the solvent removed under reduced pressure. Purification by flash chromatography (1:1, pet. ether:ethyl acetate to 100% ethyl acetate) afforded the thio sugar (**4.10**, 0.23 g, 0.62 mmol, 60%) as a colourless oil and exclusively as the α anomer. R_f 0.25 (Ethyl acetate). $[\alpha]_D^{+67.7}$ (c 1.0, CHCl_3). ν_{max} (film)/ cm^{-1} 3660 (O-H), 1738 (C=O), 1703, 1657, 1537, 1432, 1371, 1313, 1228, 1037, 966, 916, 820, 732 and 624. ^1H NMR (400 MHz, CDCl_3) δ 6.01 (1 H, d, J = 9.5, NH), 5.28 (1 H, d, J = 3.8, H-1), 5.25 (1 H, t, J = 10.8, H-3), 4.34 (1 H, dd, J = 12.2, 4.4, 1 $\times$ H-6), 4.31 – 4.24 (2 H, m, H-2 and H-5), 4.17 (1 H, dd, J = 12.1, 1.8, 1 $\times$ H-6), 3.79 (1 H, t, J = 11.3, H-4), 2.33 (3 H, s, SAc), 2.09 (3 H, s, OAc), 2.01 (3 H, s, OAc) and 1.95 (3 H, s, NHAc). ^{13}C NMR (100 MHz, CDCl_3) δ 192.5 (SCO) 171.3 (CO), 170.0 (CO), 170.5 (NCO), 91.7 (C-1), 69.6 (C-3), 68.5 (C-5), 63.3 (C-6), 53.3 (C-2), 43.6 (C-4), 30.7 (SAc), 23.1 (NHAc), 20.8 (OAc) and 20.7 (OAc). HRMS (ESI): m/z Calcd for

$C_{14}H_{21}NSO_8Na$ $[M+Na]^+$: 386.0880: Found 386.0868.

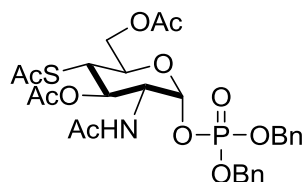
Tetrabenzylpyrophosphate (4.11)⁴¹



To a solution of dibenzyl phosphate (2 g, 7.2 mmol) in dry toluene (20 mL) was added *N,N'*-dicyclohexylcarbodiimide (0.74 g, 3.6 mmol) and the reaction was stirred at room temperature for 5 h. After this time, the reaction was filtered and the solvent removed under reduced pressure to afford tetrabenzylpyrophosphate (**4.11**, 1.9 g, 0.36 mmol, 99%) which became a white solid upon freezing and was used without further purification. ^1H NMR (400 MHz, $CDCl_3$) δ 7.30 (20 H, s, $20 \times Ph$) and 5.08 (8 H, dd, $J = 5.4, 3.2$, CH_2Ph). ^{13}C NMR (100 MHz, $CDCl_3$) δ 134.9 ($4 \times Ph$), 128.6 ($4 \times Ph$), 128.5 ($8 \times Ph$), 128.0 ($8 \times Ph$) and 70.3 ($4 \times CH_2Ph$). ^{31}P NMR (162 MHz, $CDCl_3$) δ -13.0. MS (ESI): m/z Calcd for $C_{28}H_{28}P_2O_7Na$ $[M+Na]^+$: 561.1 Found 561.2. MP = 58 – 59 °C

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.⁴¹

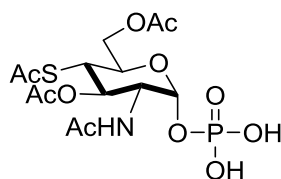
2-N-acetamido-2-deoxy-3,6-di-O-acetyl-4-thio-4-S-acetyl-4-deoxy-- α -D-glucopyranosyl-1-dibenzylphosphate (4.12)



A solution of **4.10** (0.29 g, 0.8 mmol) in dry THF (5 mL) was cooled to -78°C and butyl lithium (0.35 mL, 2.5 M in hexane, 0.88 mmol) was added dropwise. The reaction was stirred at -78°C for 5 min before tetrabenzylpyrophosphate (**4.11**, 1.36 g, 2.4 mmol) was added in dry THF (2.5 mL) over 5 min and the reaction was stirred at -78°C for a further 30 min. The reaction was then diluted with diethyl ether (20 mL) and washed with saturated sodium bicarbonate (10 mL) and water (10 mL). The organic layer was dried with sodium sulfate and the solvent removed under reduced pressure. Purification by flash chromatography (6:4, pet. ether:ethyl acetate, eluting with 99:1, ethyl acetate: NEt_3) afforded the dibenzylphosphate sugar (**4.12**, 0.33 g, 0.53 mmol, 66%) as a white foam and exclusively as the α anomer. R_f 0.21 (99:1, ethyl acetate: NEt_3). $[\alpha]_D^{+52.1}$ (c 1.0, CHCl_3). ν_{max} (film)/ cm^{-1} 1743 (C=O), 1704, 1542, 1456, 1372, 126, 1125, 1012, 952, 744, 699 and 623. ^1H NMR (400 MHz, CDCl_3) δ 7.44 – 7.31 (10 H, m, $10 \times \text{Ph}$), 5.71 (1 H, dd, $J = 6.1, 3.8$, H-1), 5.66 (1 H, d, $J = 9.5$, NH), 5.17 – 5.01 (5 H, m, H-3 and $4 \times \text{CH}_2$), 4.41 – 4.32 (1 H, m, H-2), 4.28 (1 H, dd, $J = 12.2, 3.9$, $1 \times \text{H-6}$), 4.02 (1 H, ddd, $J = 11.7, 3.9, 2.0$, H-5), 3.96 (1 H, dd, $J = 12.5, 2.0$, $1 \times \text{H-6}$), 3.80 (1 H, t, $J = 11.3$, H-4), 2.33 (3 H, s, SAc), 2.014 (3 H, s, OAc), 2.008 (3 H, s, OAc) and 1.69 (3 H, s, NHAc). ^{13}C NMR (100 MHz, CDCl_3) δ 192.0 (SCO) 171.0 (CO), 170.6 (CO), 170.1 (NCO), 135.3 (Ph), 135.2 (Ph), 129.0 ($2 \times \text{Ph}$), 128.80 ($2 \times \text{Ph}$), 128.75 ($2 \times \text{Ph}$), 128.1 ($4 \times \text{Ph}$), 96.7 (C-1), 70.7 (C-5), 70.0 (CH_2Ph), 69.9 (CH_2Ph), 68.7 (C-3), 62.5 (C-6),

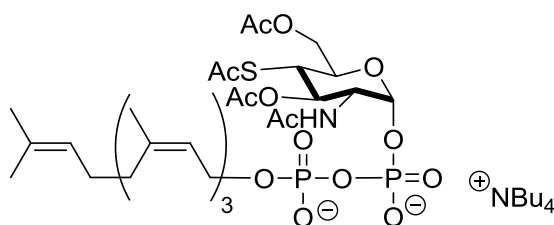
52.7 (C-2), 43.0 (C-4), 30.6 (SAC), 22.7 (NHAc), 20.6 (OAc) and 20.5 (OAc). ^{31}P NMR (162 MHz, CDCl_3) δ – 2.3. MS (ESI): m/z Calcd for $\text{C}_{28}\text{H}_{34}\text{NSPO}_{11}\text{Na}$ $[\text{M}+\text{Na}]^+$: 646.1482; Found 646.1482.

2-N-acetamido-3,6-di-O-acetyl-4-thio-4-S-acetyl-4-deoxy- α -D-glucopyranosyl-1-phosphate (4.13)



To **4.12** (0.33g, 0.53 mmol) and Pd/C (250 mg, 10% loading, 0.39 mmol) was added dry methanol (15 mL). The reaction was then put under a hydrogen atmosphere and stirred at room temperature for 16 h. After this time, the reaction was filtered through Celite and the solvent removed under reduced pressure to afford the phosphate sugar (**4.13**, 0.22 g, 0.49 mmol, 92%) as a light brown foam, exclusively as the α anomer, that was used without further purification. $[\alpha]_{\text{D}} +26.4$ (c 1.0, MeOH). ν_{max} (film)/ cm^{-1} 3280 (O-H), 2938, 1739 (C=O), 1656, 1545, 1435, 1374, 1229, 1118, 1035 and 943. ^1H NMR (400 MHz, MeOD) δ 5.56 (1 H, s br, H-1), 5.25 (1 H, t, $J = 10.8$, H-3), 4.38 (1 H, dd, $J = 12.0, 2.9$, H-2), 4.33 – 4.21 (2 H, m, H-5 and $1 \times$ H-6), 4.11 (1 H, d, $J = 12.0$, $1 \times$ H-6), 3.78 (1 H, t, $J = 11.3$, H-4), 2.29 (3 H, s, SAc), 2.02 (3 H, s, OAc), 1.94 (3 H, s, OAc) and 1.91 (3 H, s, NHAc). ^{13}C NMR (100 MHz, CDCl_3) δ 194.4 (SCO) 173.8 (CO), 172.7 (CO), 172.0 (NCO), 96.4 (C-1), 71.5 (C-5), 70.6 (C-3), 64.2 (C-6), 54.5 (C-2), 45.1 (C-4), 30.7 (SAC), 22.6 (NHAc), 20.9 (OAc) and 20.7 (OAc). ^{31}P NMR (162 MHz, MeOD) δ –1.9. HRMS (ESI): m/z Calcd for $\text{C}_{14}\text{H}_{21}\text{NSPO}_{11}\text{Na}$ $[\text{M}-\text{H}]^-$: 442.0567; Found 442.0576.

***P*¹-2-*N*-acetamido-2-deoxy-3,6-di-*O*-acetyl-4-thio-4-*S*-acetyl-4-deoxy- α -D-glucopyranosyl (2*Z*,6*Z*,10*E*)-3,7,11-trimethylhexadeca-2,6,10,14-tetraen-1-yl diphosphate (4.14)**

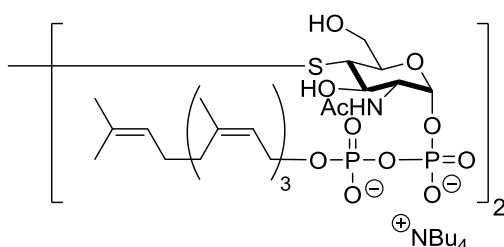


To a solution of WZ₃-OH (10 mg, 0.03 mmol) and Bu₄NPO₄H₂ (21 mg, 0.06 mmol) in dry dichloromethane (1 mL) was added trichloroacetonitrile (10 μ L, 0.1 mmol). After stirring at room temperature for 1 h, the solvent was removed under reduced pressure and the resulting residue was re-dissolved in THF (1.5 mL) and ammonium hydroxide (1.5 mL, 35% NH₃ in water). After stirring at room temperature for 30 min, 1:1 methanol:toluene (4 mL) was added and the mixture was stirred for a further 20 min. The mixture was filtered and the solvent removed under reduced pressure. The resulting residue was washed with pet. ether (1 mL) and re-dissolved in methanol (5 mL) and filtered. Dowex 50WX8 (NH₄⁺ form, 100 mg) was added and the mixture stirred for 30 min, after which the pH was approx. 7. The resin was filtered off and the solvent was removed under reduced pressure to afford the crude phosphate.

To a solution of the crude phosphate in dry dichloromethane (1 mL) was added carbonyldiimidazole (28 mg, 0.17 mmol) and the reaction was stirred at room temperature for 2 h. After this time, methanol (0.1 mL) was added and the reaction was stirred for a further 30 min. The solvent was removed under reduced pressure and 2-acetamido-3,6-*O*-acetyl-4-*S*-acetyl-4-thio- α -D-glucopyranosyl-1-phosphate (**4.13**, 23 mg, 0.05 mmol) and tetrazole (0.1 mL, 0.45 M in MeCN, 0.05 mmol) were added in dry

DMF (1 mL). The reaction was stirred at room temperature for 3 days before the solvent was removed under reduced pressure. Purification by reverse phase C-18 resin (95:5, water:MeCN to 75:20:5, water:isopropanol:MeCN) afforded the pyrophosphatelinked lipid sugar product (**4.14**, 16 mg, approx 2:1 Bu₄N⁺:Product by NMR, 0.013 mmol, 42%) as a yellow foam. *R_f* 0.54 (50:45:5, water:isopropanol:MeCN, C-18 resin). ¹H NMR (400 MHz, MeOD) δ 5.67 – 5.64 (1 H, m, H-1), 5.50 – 5.47 (1 H, m, H-2'), 5.28 (1 H, t, *J* = 10.5, H-3), 5.18 – 5.07 (4 H, m, H-5, H-6', H-10' and H-14'), 4.61 – 4.51 (3 H, m, 1 × H-6 and 2 × H-1'), 4.43 – 4.34 (1 H, m, 1 × H-6), 4.28 – 4.21 (1 H, m, H-2), 4.10 (1 H, dd, *J* = 12.4, 1.4, H-5), 3.90 (1 H, t, *J* = 11.4, H-4), 2.30 (3 H, s, SAc), 2.14 – 1.90 (12 H, m, 2 × H-4', 2 × H-5', 2 × H-8', 2 × H-9', 2 × H-12' and 2 × H-13'), 2.08 (3 H, s, OAc), 2.05 (3 H, s, OAc), 1.95 (3 H, s, NHAc), 1.68 (12 H, s, 3 × H-17', 3 × H-18', 3 × H-19' and 3 × H-20') and 1.61 (3 H, s, H-16'). NBu₄⁺: 3.26 (8 H, m, 8 × NCH₂), 2.13 – 2.02 (8 H, m, 8 × NCH₂CH₂), 1.42 (8 H, quad, *J* = 7.4, 8 × CH₂CH₃) and 1.03 (12 H, t, *J* = 7.4 12 × CH₃). ¹³C NMR (125 MHz, MeOD) δ 194.4 (SCO), 174.4 (CO), 172.8 (CO), 171.6 (NCO), 140.0 (C-3'), 136.4 (C-7'), 132.5 (C-11'), 129.4 (C-15'), 128.7, 128.6, 126.3 (C-6', C-10' and C-14'), 125.5 (C-2'), 96.2 (C-1), 72.1 (C-3), 71.1 (C-5), 64.3 (C-1'), 63.8 (C-6), 55.0 (C-2), 44.7 (C-4), 33.5, 33.4 (C-4' and C-8'), 33.0 (C-12'), 27.9, 27.7 (C-5' and C-9'), 26.1 (C-20'), 23.9 (C-17', C-18' and C-19'), 23.0 (NHAc), 20.9 (2 × OAc) and 17.9 (C-16'). NBu₄⁺: 59.6 (NCH₂), 24.9 (NCH₂CH₂), 20.9 (CH₂CH₃) and 14.1 (CH₃). ³¹P NMR (162 MHz, MeOD) δ –9.3 and –12.3. HRMS (ESI): *m/z* Calcd for C₃₄H₅₃NO₁₄P₂S [M–H][–]: 794.2746; Found 794.2743.

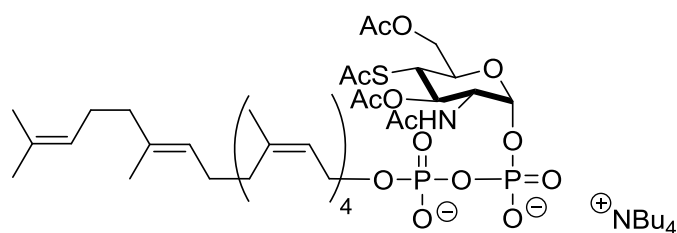
Dis [P¹-2-N-acetamido-2-deoxy-4-thio 4-deoxy- α -D-glucopyranosyl (2Z,6Z,10E)-3,7,11-trimethylhexadeca-2,6,10,14-tetraen-1-yl diphosphate] 4'-disulfide (4.15)



A solution of **4.14** (16 mg, 2:1, Bu₄N⁺:**4.14**, 0.013 mmol) in dry methanol (1 mL) was cooled to 0 °C and freshly prepared in NaOMe (0.4 mL, 200 mM, 0.08 mmol) was added and the reaction was stirred at 0 °C for 4.5 h. After this time, the reaction was quenched by the addition of Dowex 50WX8 (NH₄⁺ form, 100 mg) and the reaction was stirred for 15 min. The Dowex was removed by filtration and the solvent removed under reduced pressure. Purification by reverse phase C-18 resin (95:5, 2% NH₄HCO₃ in water:MeCN to 50:50, 2% NH₄HCO₃ in water:MeCN) afforded the disulfide pyrophosphate linked lipid sugar product (**4.15**, 6 mg, approx 0.8:1 Bu₄N⁺:Product by NMR, 3 μmol, 43%) as a light yellow foam. *R_f* 0.10 (1:1, 2% NH₄HCO₃ in water:MeCN, C-18 resin). ¹H NMR (400 MHz, MeOD) δ 5.64 – 5.55 (1 H, m, H-1), 5.48 – 5.42 (1 H, m, H-2'), 5.17 – 5.10 (3 H, m, H-6', H-10' and H-14'), 4.58 – 4.49 (2 H, m, 2 × H-1'), 4.42 – 3.80 (6 H, m, H-2, H-3, H-4, H-5, 2 × H-6), 2.16 – 1.98 (15 H, m, 2 × H-4', 2 × H-5', 2 × H-8', 2 × H-9', 2 × H-12', 2 × H-13' and NHAc), 1.73 (3 H, s, H-17'), 1.68 (9 H, s, 3 × H-18', 3 × H-19' and 3 × H-20') and 1.61 (3 H, s, H-16'). NBu₄⁺: 3.26 (8 H, m, 8 × NCH₂), 2.13 – 2.02 (8 H, m, 8 × NCH₂CH₂), 1.42 (8 H, quad, *J* = 7.4, 8 × CH₂CH₃) and 1.03 (12 H, t, *J* = 7.4 12 × CH₃). ¹³C NMR (125 MHz, MeOD) δ 174.7 (NCO), 140.0 (C-3'), 136.4 (C-7'), 132.5 (C-11'), 129.4 (C-15'), 128.64, 128.58, 126.3 (C-6', C-10' and C-14'), 125.5 (C-2'), 95.3 (C-1), 68.38 and

68.33 (C-3 and C-5), 63.78 and 63.76 (C-6 and C-1'), 57.5 (C-2), 46.7 (C-4), 33.5, 33.4 (C-4' and C-8'), 33.0 (C-12'), 27.9, 27.7 (C-5' and C-9'), 26.1 (C-20'), 23.9 (C-17', C-18' and C-19'), 21.1 (NHAc) and 17.9 (C-16'). NBu_4^+ : 59.6 (NCH_2), 24.9 (NCH_2CH_2), 20.9 (CH_2CH_3) and 14.1 (CH_3). ^{31}P NMR (162 MHz, MeOD) δ -9.2 and -11.6. HRMS (ESI): m/z Calcd for $\text{C}_{56}\text{H}_{94}\text{N}_2\text{O}_{22}\text{P}_4\text{S}_2$ $[\text{M}-2\text{H}]^{2-}$: 667.2351; Found 667.2371.

***P*¹-2-*N*-acetamido-2-deoxy-3,6-di-*O*-acetyl-4-thio-4-*S*-acetyl-4-deoxy- α -D-glucopyranosyl ((2*Z*,6*Z*,10*Z*,14*Z*,18*E*)-3,7,11,15,19,23-hexamethylicosakaitetra-2,6,10,14,18,22-hexaen-1-yl diphosphate (4.16)**

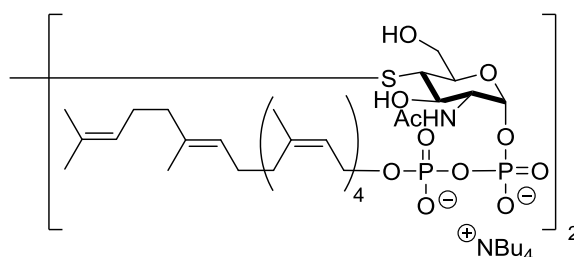


To a solution of WEZ₄-OH (**3.22**, 10 mg, 0.023 mmol) and Bu₄NPO₄H₂ (13 mg, 0.038 mmol) in dry dichloromethane (1 mL) was added trichloroacetonitrile (10 μL , 0.07 mmol). After stirring at room temperature for 1 h, the solvent was removed under reduced pressure and the resulting residue was re-dissolved in THF (1.5 mL) and ammonium hydroxide (1.5 mL, 35% NH₃ in water). After stirring at room temperature for 30 min, 1:1 methanol:toluene (4 mL) was added and the mixture was stirred for a further 20 min. The mixture was filtered and the solvent removed under reduced pressure. The resulting residue was washed with pet. ether (1 mL) and re-dissolved in methanol (5 mL) and filtered. Dowex 50WX8 (NH₄⁺ form, 100 mg) was added and the mixture stirred for 30 min, after which the pH was approx. 7. The resin was filtered off and solvent removed under reduced pressure to afford the crude phosphate.

To a solution of the crude phosphate in dry dichloromethane (1 mL) was added carbonyldiimidazole (19 mg, 0.12 mmol) and the reaction was stirred at room temperature for 2.5 h. After this time, methanol (0.1 mL) was added and the reaction was stirred for a further 30 min. The solvent was removed under reduced pressure and 2-acetamido-3,6-*O*-acetyl-4-*S*-acetyl-4-thio- α -D-glucopyranosyl-1-phosphate (**4.13**, 15 mg, 0.035 mmol) and tetrazole (0.1 mL, 0.45 M in MeCN, 0.05 mmol) were added in dry DMF (1 mL). The reaction was stirred at room temperature for 3 days before the solvent was removed under reduced pressure. Purification by reverse phase C-18 resin (95:5, water:MeCN to 75:20:5, water:isopropanol:MeCN) afforded the pyrophosphate linked lipid sugar product (**4.16**, 11 mg, approx 1.7:1 Bu₄N⁺:Product by NMR, 0.008 mmol, 36%) as a yellow foam. *R*_f 0.36 (C-18 resin, 50:45:5, water:isopropanol:MeCN). ¹H NMR (400 MHz, MeOD) δ 5.68 – 5.56 (1 H, m, H-1), 5.47 – 5.41 (1 H, m, H-2'), 5.25 (1 H, t, *J* = 11.0, H-3), 5.18 – 5.12 (4 H, m, H-6', H-10', H-14' and H-18'), 5.09 (1 H, dt, *J* = 6.9, 1.3, H-22'), 4.61 – 4.51 (2 H, m, 2 $\times$ H-1'), 4.49 (1 H, dd, *J* = 12.5, 3.0, 1 $\times$ H-6), 4.44 – 4.38 (1 H, m, 1 $\times$ H-6), 4.27 (1 H, dt, *J* = 10.3, 2.7, H-2), 4.12 (1 H, dd, *J* = 12.3, 1.7, H-5), 3.88 (1 H, t, *J* = 11.2, H-4), 2.30 (3 H, s, SAc), 2.15 – 2.02 (18 H, m, 2 $\times$ H-4', 2 $\times$ H-5', 2 $\times$ H-8', 2 $\times$ H-9', 2 $\times$ H-12', 2 $\times$ H-13', 2 $\times$ H-16', 2 $\times$ H-17' and 2 $\times$ H-21'), 2.00 – 1.97 (8 H, m, 3 $\times$ OAc and 2 $\times$ H-20'), 1.93 (3 H, s, 3 $\times$ NHAc), 1.73 (3 H, s, 3 $\times$ H-25'), 1.70 – 1.64 (12 H, m, 3 $\times$ H-26', 3 $\times$ H-27', 3 $\times$ H-28' and 3 $\times$ H-30'), 1.61 (3 H, s, 3 $\times$ H-29') and 1.60 (3 H, s, 3 $\times$ H-24'). NBu₄⁺: 3.26 (8 H, m, 8 $\times$ NCH₂), 2.13 – 2.02 (8 H, m, 8 $\times$ NCH₂CH₂), 1.42 (8 H, quad, *J* = 7.4, 8 $\times$ CH₂CH₃) and 1.03 (12 H, t, *J* = 7.4 12 $\times$ CH₃). ¹³C NMR (100 MHz, MeOD) δ 194.4 (SCO), 174.3 (CO), 172.8 (CO), 171.7 (NCO), 140.7 (C-3'), 136.4, 136.3, 136.2, 135.7 (C-7', C-11', C-15' and C-19'), 132.2 (C-23'), 126.4, 126.3, 126.1, 125.6, 125.5 (C-6', C-10', C-14',

C-18' and C-22'), 120.6 (C-2'), 96.3 (C-1), 71.9 (C-5), 71.3 (C-3), 64.2 (C-1'), 63.9 (C-6), 54.7 (C-2), 44.7 (C-4), 41.0 (C-20'), 33.5, 33.4 (C-4', C-8' and C-12'), 33.0 (C-16'), 30.7 (SCOCH₃), 27.9 (C-21'), 27.8, 27.71, 27.67 (C-5', C-9', C-13' and C-17'), 26.1 (C-30'), 23.97, 23.93, 23.91, 23.88 (C-25', C-26', C-27' and C-28'), 23.0 (NHAc), 20.93 (OCOCH₃), 20.7 (OCOCH₃), 18.0 (C-24') and 16.3 (C-29'). NBu₄⁺: 59.6 (NCH₂), 24.9 (NCH₂CH₂), 20.9 (CH₂CH₃) and 14.1 (CH₃). ³¹P NMR (202 MHz, MeOD) δ -9.1 and -12.1. HRMS (ESI): *m/z* Calcd for C₄₄H₇₀NO₁₄P₂S [M-H]⁻: 930.3998: Found 930.4017.

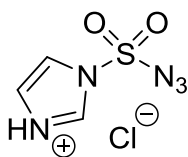
Dis [P¹-2-*N*-acetamido-2-deoxy--4-thio--4-deoxy- α -D-glucopyranosyl (2*Z*,6*Z*,10*E*)-3,7,11-trimethylhexadeca-2,6,10,14-tetraen-1-yl diphosphate] 4'-disulfide (4.17)



A solution of **4.16** (11 mg, 1.7:1, Bu₄N⁺:**4.16**, 0.008 mmol) in dry methanol (1 mL) was cooled to 0 °C and freshly prepared in NaOMe (0.6 mL, 200 mM, 0.12 mmol) was added and the reaction was stirred at 0 °C for 4 h. After this time, the reaction was quenched by the addition of Dowex 50WX8 (NH₄⁺ form, 100 mg) and the reaction was stirred for 15 min. The Dowex was removed by filtration and removal of the solvent under reduced pressure afforded the disulfide pyrophosphate linked lipid sugar product (**4**, 8 mg, approx 0.8:1 Bu₄N⁺:Product by NMR, 0.008 mmol, quantitative) as a light yellow foam, which was used without further purification. *R_f* 0.50 (3:1, 2% NH₄HCO₃ in water:MeCN, C-18 resin). ¹H NMR (400 MHz, MeOD) δ 5.63 – 5.56 (1 H, m, H-1),

5.47 – 5.40 (1 H, m, H-2'), 5.17 – 5.12 (4 H, m, H-6', H-10', H-14' and H-18'), 5.09 (1 H, tt, $J = 7.1, 1.4$, H-22'), 4.52 (2 H, t, $J = 6.0$, $2 \times$ H-1'), 4.42 – 3.68 (6 H, m, H-2, H-3, H-4, H-5, $2 \times$ H-6), 2.13 – 2.02 (21 H, m, $2 \times$ H-4', $2 \times$ H-5', $2 \times$ H-8', $2 \times$ H-9', $2 \times$ H-12', $2 \times$ H-13', $2 \times$ H-16', $2 \times$ H-17', $2 \times$ H-21' and NHAc), 2.00 – 1.97 (2 H, m, $2 \times$ H-20'), 1.73 (3 H, s, $3 \times$ H-25'), 1.69 – 1.66 (12 H, m, $3 \times$ H-26', $3 \times$ H-27', $3 \times$ H-28' and $3 \times$ H-30'), 1.61 (3 H, s, $3 \times$ H-29') and 1.60 (3 H, s, $3 \times$ H-24'). NBu_4^+ : 3.26 (8 H, m, $8 \times$ NCH_2), 2.13 – 2.02 (8 H, m, $8 \times$ NCH_2CH_2), 1.42 (8 H, quad, $J = 7.4$, $8 \times$ CH_2CH_3) and 1.03 (12 H, t, $J = 7.4$, $12 \times$ CH_3). ^{13}C NMR (100 MHz, MeOD) δ 175.8 (NHAc), 140.2 (C-3'), 136.5, 136.4, 136.3, 136.1 (C-7', C-11', C-15' and C-19'), 132.3 (C-23'), 126.4, 125.6 (C-6', C-10', C-14', C-18' and C-22'), 121.0 (C-2'), 96.1 (C-1), 71.6 (C-3 and C-5), 68.8, 62.9 (C-6 and C-1'), 57.6 (C-2), 41.0 (C-4 and C-20'), 33.4, 33.2 (C-4', C-8' and C-12'), 33.0 (C-16'), 27.9 (C-21'), 27.8, 27.72, 27.68 (C-5', C-9', C-13' and C-17'), 26.1 (C-30'), 23.9 (C-25', C-26', C-27' and C-28'), 20.9 (NHAc), 17.9 (C-24') and 16.3 (C-29'). NBu_4^+ : 59.6 (NCH_2), 24.9 (NCH_2CH_2), 20.9 (CH_2CH_3) and 14.1 (CH_3). ^{31}P NMR (202 MHz, MeOD) δ -8.9 and -12.2. HRMS (ESI): m/z Calcd for $\text{C}_{76}\text{H}_{126}\text{N}_2\text{O}_{22}\text{P}_4\text{S}_2$ $[\text{M}-2\text{H}]^{2-}$: 803.3603; Found 803.3610.

Imidazole-1-sulfonyl Azide Hydrochloride (4.20)²⁶

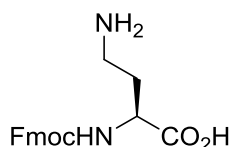


A solution of sodium azide (0.65 g, 10 mmol) in acetonitrile (10 mL) was cooled to 0 °C and sulfonyl chloride (0.8 mL, 10 mmol) was added dropwise. The solution was allowed to stir for 16 h at room temperature. After this time, the reaction cooled to 0 °C and

imidazole (1.3 g, 19 mmol) was added slowly in portions. The reaction was allowed to stir at room temperature for 3 h before being diluted with ethyl acetate (30 mL). After this time, the solution was washed with water (2 × 30 mL) and saturated sodium bicarbonate solution (2 × 30 mL) before being dried with magnesium sulfate. The solution was cooled and hydrogen chloride (formed by the slow addition of acetyl chloride (1 mL, 15 mmol) to ice-cooled dry ethanol (3.8 mL)) was added dropwise and the product precipitated out of solution. The solid was filtered off and washed with ethyl acetate (20 mL) to afford the hydrogen chloride salt of the sulfonyl (**4.20**, 1.06 g, 5 mmol, 50%) as a white solid, which was used without further purification. ¹H NMR (200 MHz, D₂O) δ 8.91 (1 H, t, J = 1.3, H-2), 7.81 (1 H, dd, J = 2.0, 1.6, H-5) and 7.37 (1 H, dd, J = 2.1, 1.1 H-4). ¹³C NMR (100 MHz, D₂O) δ 138.1 (C-2), 123.9 (C-4) and 120.4 (C-5). MS (ESI): *m/z* Calcd for C₃H₄N₅O₂S [M-Cl]⁺: 174.0: Found 174.0.

The observed characterisation data (¹H and ¹³C) was consistent with that previously reported in the literature.²⁶

***N*-Fmoc-diaminobutyric acid (**4.21**)²⁵**

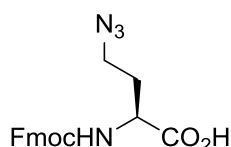


To [bis(trifluoroacetoxy)iodo] benzene (1.7 g, 4.1 mmol) in DMF (13 mL) and water (7 mL) was added *N*-Fmoc glutamine (1 g, 2.7 mmol). The reaction was stirred at room temperature for 15 min before pyridine (0.46 mL, 5.2 mmol) was added and the reaction was stirred at room temperature for 16 h. The solvent was then removed under reduced

pressure and the resulting residue dissolved in water (20 mL). Concentrated HCl (0.5 mL) was added and the aqueous layer was washed with diethyl ether (2×15 mL). The pH of the aqueous layer was adjusted to 6 with NaOH and the resulting solid was filtered off and washed with water (10 mL), ice-cold ethanol (10 mL) and diethyl ether (2×10 mL). The solid was dried under reduced pressure to give the desired amine amino acid as a cream solid (**4.21**, 600 mg, 1.7 mmol, 65%) which was used without further purification. $[\alpha]_D +18.0$ (c 0.25, DMSO). ^1H NMR (400 MHz, d_6 -DMSO) δ 7.89 (2 H, d, $J = 7.5$, $2 \times \text{Ph}$), 7.68 (2 H, d, $J = 6.3$, $2 \times \text{Ph}$), 7.41 (2 H, t, $J = 7.3$, $2 \times \text{Ph}$), 7.33 (2 H, t, $J = 7.3$, $2 \times \text{Ph}$), 6.69 (1 H, d, $J = 5.0$, CONH), 4.30 – 4.15 (3 H, m, Fmoc-CH₂ and Fmoc-CH), 3.69 – 3.59 (1 H, m, NCH), 2.93 – 2.81 (2 H, m, $1 \times \text{CH}_2\text{CH}_2\text{NH}_2$), 1.93 – 1.83 (1 H, m, $1 \times \text{CH}_2\text{NH}_2$), and 1.81 – 1.70 (1 H, m, $1 \times \text{CH}_2\text{NH}_2$). ^{13}C NMR (100 MHz, d_6 -DMSO) δ 172.7 (CO₂H), 155.2 (CONH), 144.0 (Ph), 143.9 (Ph), 140.0 ($2 \times \text{Ph}$), 127.6 ($2 \times \text{Ph}$), 127.1 ($2 \times \text{Ph}$), 125.2 ($2 \times \text{Ph}$), 120.1 ($2 \times \text{Ph}$), 65.5 (Fmoc-CH₂), 54.9 (NCH), 46.7 (Fmoc-CH), 37.3 (CH₂CH₂NH₂) and 30.5 (CH₂NH₂). MS (ESI): m/z Calcd for C₁₉H₂₀N₂O₄Na $[\text{M}+\text{H}]^+$: 341.1: Found 341.1. MP = 149 – 150 °C.

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.²⁵

N-Fmoc-azidohomoalanine (4.22)²⁵

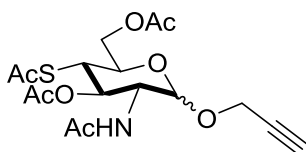


To **4.21** (0.5 g, 1.6 mmol) in methanol (15 mL), dichloromethane (12 mL) and water (7.5 mL) was added **4.20** (1 g, 4.9 mmol) and copper sulfate (3 mg, 0.016 mmol). The

pH of the reaction was adjusted to 9 with aqueous potassium carbonate and the reaction was stirred at room temperature for 16 h. After this time, the reaction was diluted with dichloromethane (20 mL) and the organic layer was extracted with saturated sodium bicarbonate (2×20 mL). The combined aqueous layers were washed with diethyl ether (2×10 mL). The pH of the aqueous layer was adjusted to 2 using concentrated HCl and the aqueous layer was extracted with diethyl ether (2×20 mL). The organic layer was dried with sodium sulfate and the solvent was removed under reduced pressure to afford the azido amino acid as a white solid (**4.22**, 326 mg, 0.9 mmol, 56%) which was used without further purification. $[\alpha]_D +14.0$ (c 1.0, CHCl_3). ^1H NMR (400 MHz, CDCl_3) δ 7.77 (2 H, d, $J = 7.3$, $2 \times \text{Ph}$), 7.63 – 7.52 (2 H, m, $2 \times \text{Ph}$), 7.41 (2 H, t, $J = 7.1$, $2 \times \text{Ph}$), 7.33 (2 H, tt, $J = 1.2$, 7.6, $2 \times \text{Ph}$), 5.50 (1 H, d, $J = 8.1$, CONH), 4.65 – 4.40 (3 H, m, Fmoc- CH_2 and NCH), 4.23 (1 H, t, $J = 6.4$, Fmoc-CH), 3.43 (2 H, t, $J = 6.4$, CH_2N_3), 2.24 – 2.14 (1 H, m, $1 \times \text{CH}_2\text{CH}_2\text{N}_3$) and 2.06 – 1.97 (1 H, m, $1 \times \text{CH}_2\text{CH}_2\text{N}_3$). ^{13}C NMR (100 MHz, CDCl_3) δ 175.9 (CO_2H), 156.1 (CONH), 143.7 (Ph), 143.5 (Ph), 141.3 ($2 \times \text{Ph}$), 127.8 ($2 \times \text{Ph}$), 127.1 ($2 \times \text{Ph}$), 125.0 ($2 \times \text{Ph}$), 120.0 ($2 \times \text{Ph}$), 67.2 (Fmoc- CH_2), 51.6 (NCH), 47.6 (CH_2N_3), 47.1 (Fmoc-CH) and 31.3 ($\text{CH}_2\text{CH}_2\text{N}_3$). MS (ESI): m/z Calcd for $\text{C}_{19}\text{H}_{18}\text{N}_4\text{O}_4\text{Na}$ $[\text{M}+\text{Na}]^+$: 389.1: Found 389.1.

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.²⁵

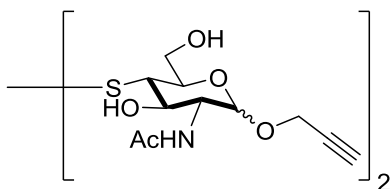
Propargyl 2-N-acetamido-2-deoxy-3,6-di-O-acetyl-4-thio-4-S-acetyl-4-deoxy-D-glucopyranoside (4.23)



To **4.9** (0.25 g, 0.6 mmol) in dry dichloromethane (20 mL) was added propargyl alcohol (0.07 mL, 1.2 mmol) and Yb(OTf)₃ (191 mg, 0.3 mmol). The reaction was heated to reflux (55 °C) and stirred for 20 h. After this time, the reaction was cool to room temperature and diluted with dichloromethane (20 mL) and washed with water (20 mL). The aqueous layer was extracted with dichloromethane (20 mL) and the combined organic layers were dried with sodium sulfate and the solvent removed under reduced pressure. Purification by flash chromatography (ethyl acetate) afforded the thio alkyne sugar (**4.23**, 48 mg, 0.12 mmol, 20%) as a colourless oil and a 55:45 mix of α : β anomers by NMR. The diastereomers are denoted α and β in the ¹H NMR but are unassigned in the ¹³C NMR and are denoted a and b. *R*_f 0.33 (ethyl acetate). ν_{max} (film)/cm⁻¹ 1740, 1702, 1664 (all C=O), 1539, 1430, 1371, 1311, 1224, 1033, 959, 913, 814, 730 and 623. ¹H NMR (400 MHz, CDCl₃) δ 5.73 (1 H ^{α} , d, *J* = 9.5, NH ^{α}), 5.64 (1 H ^{β} , d, *J* = 9.1, NH ^{β}), 5.24 (1 H ^{α} , t, *J* = 10.3, H-3 ^{α}), 5.16 (1 H ^{β} , t, *J* = 10.3, H-3 ^{β}), 5.07 (1 H, d, *J* = 3.4, H-1 ^{α}), 4.79 (1 H, d, *J* = 8.6, H-1 ^{β}), 4.43 – 4.33 (2 H + 1 H ^{α} , m, 2 $\times$ CH₂CCH and 1 $\times$ H-6 ^{α}), 4.26 (1 H ^{α} , d, *J* = 2.2, 1 $\times$ H-6 ^{α}), 4.25 (1 H ^{β} , d, *J* = 2.2, 1 $\times$ H-6 ^{β}), 4.22 (1 H ^{β} , dd, *J* = 2.2, 6.6, 1 $\times$ H-6 ^{β}), 4.15 (1 H ^{α} , dd, *J* = 2.2, 10.3, H-2 ^{α}), 4.03 (1 H ^{α} , ddd, *J* = 2.0, 4.4, 11.3, H-5 ^{α}), 4.00 – 3.93 (1 H ^{β} , m, H-2 ^{β}), 3.87 (1 H ^{β} , ddd, *J* 2.0, 4.9, 11.0, H-5 ^{β}), 3.82 (1 H ^{α} , t, *J* = 11.3, H-4 ^{α}), 3.65 (1 H ^{β} , t, *J* = 11.0, H-4 ^{β}), 2.49 – 2.45 (1 H, m, CCH), 2.324 (3 H ^{α} , s, SAc ^{α}), 2.319 (3 H ^{β} , s, SAc ^{β}), 2.10 (3 H ^{α} , s, OAc ^{α}), 2.08 (3 H ^{β} , s, OAc ^{β}), 2.03 (3 H ^{β} , s, OAc ^{β}), 2.01 (3 H ^{α} , s, OAc ^{α}) and 1.95 (3 H, s, NHAc). ¹³C NMR (100 MHz, CDCl₃) δ

192.8 (SCO^a), 192.2 (SCO^b), 171.1 (CO^a), 170.9 (CO^b), 170.5 (CO), 170.3 (NCO^a), 170.0 (NCO^b), 98.2 (C-1^a), 96.4 (C-1^b), 78.6 (C^aCH), 78.2 (C^bCH), 75.4 (CC^aH), 75.3 (CC^bH), 69.6 (C-5^a), 69.4 (C-5^b), 69.6 (C-3^a), 69.4 (C-3^b), 63.1 (C-6^a), 63.0 (C-6^b), 55.6 (C^aH₂CCH), 55.18 (C^bH₂CCH), 55.15 (C-2^a), 52.5 (C-2^b), 44.1 (C-4^a), 43.4 (C-4^b), 30.7 (SAC^a), 30.6 (SAC^b), 23.3 (NHAc^a), 23.2 (NHAc^b), 20.8 (OAc) and 20.6 (OAc). HRMS (ESI): *m/z* Calcd for C₁₇H₂₃NO₈SNa [M+H]⁺: 424.1037: Found 424.1027.

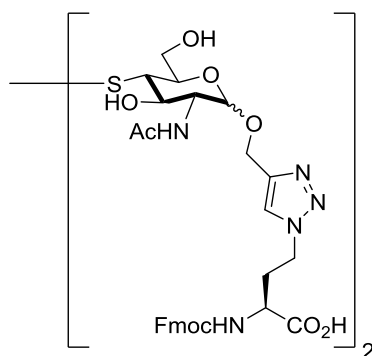
Dis (Propargyl 2-N-acetamido-2-deoxy-4-thio-4-deoxy-D-glucopyranoside) 4-disulfide (4.24)



To **4.23** (0.11 g, 0.27 mmol) in dry methanol (3 mL) was added freshly prepared sodium methoxide (3 mL, 500 mM, 1.5 mmol) and the reaction was stirred at room temperature for 2 h. After this time, the reaction was quenched by the addition of Dowex 50WX8 (H⁺ form, 0.2 g) and the reaction was stirred for 45 min. The Dowex was removed by filtration and removal of the solvent under reduced pressure afforded the deprotected thiol alkyne sugar (**4.24**, 80 mg, 0.14 mmol, quantitative) as an orange foam and a 6:4 mix of α : β anomers by NMR, which was used without further purification. The diastereomers are denoted α and β in the ¹H and ¹³C NMR. ν_{max} (film)/cm⁻¹ 3279 (O-H, br), 2924, 1640 (C=O), 1555, 1449, 1376, 1316, 1106, 1034, and 639. ¹H NMR (400 MHz, MeOD) δ 6.00 (1 H ^{α} , d, *J* = 3.5, H-1 ^{α}), 5.57 (1 H ^{β} , d, *J* = 8.4, H-1 ^{β}), 5.32 (1 H, d, *J* = 2.1, 1 $\times$ CH₂CCH), 4.23 (1 H, d, *J* = 4.0, 1 $\times$ CH₂CCH), 4.00 (1 H ^{β} , t, *J* = 3.1, H-2 ^{β}), 3.97 (1 H ^{α} , d, *J* = 2.8, H-2 ^{α}), 3.94 – 3.52 (4 H, m, H-3, H-5 and 2 $\times$ H-6), 2.86 – 2.79 (1

H, m, CCH), 1.96 (3 H, s, NHAc) and 1.24 (1 H, m, H-4). ^{13}C NMR (100 MHz, MeOD) δ 174.0 (NCO $^{\beta}$), 173.9 (NCO $^{\alpha}$), 100.2 (C-1 $^{\beta}$), 97.3 (C-1 $^{\alpha}$), 80.1 (CCH), 76.45 (CC $^{\beta}$ H), 76.43 (CC $^{\alpha}$ H), 74.0 (C-5 $^{\beta}$), 73.8 (C-5 $^{\alpha}$), 69.0 (C-3 $^{\beta}$), 68.9 (C-3 $^{\alpha}$), 63.1 (C-6 $^{\beta}$), 63.0 (C-6 $^{\alpha}$), 56.7 (C $^{\alpha}$ H₂CCH), 56.4 (C $^{\beta}$ H₂CCH), 56.0 (C-2 $^{\beta}$), 55.6 (C-2 $^{\alpha}$), 30.8 (C-4), 23.2 (NHAc $^{\beta}$) and 23.2 (NHAc $^{\alpha}$). HRMS (ESI): m/z Calcd for C₂₂H₃₂N₂O₁₀S₂Na [M+Na] $^{+}$: 571.1391; Found 571.1387.

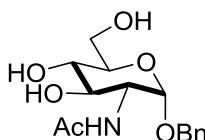
Dis {(S)-2-(9H-fluoren-9-ylmethoxycarbonylamino)-4-{4-[(2-N-acetamido-2-deoxy-4-thio-4-deoxy-D-glucopyranosyloxy)methyl]-[1,2,3]-triazole-1-yl}}butanoic acid} 4 disulfide (4.25)



To a deoxygenated (argon bubbled though for 5 min) solution of **4.24** (30 mg, 0.1 mmol) in acetonitrile was added water (0.3 mL), copper(I) bromide (10 mg, 0.07 mmol), triethylamine (10 μL , 0.7 mmol) and Fmoc-Aha (**4.22**, 24 mg, 0.7 mmol). The reaction was stirred at room temperature for 16 h. After this time, the solvent was removed under reduced pressure and the residue re-dissolved in DMSO. Purification by HPLC (Dionex Ultimate 3000 pump and variable wavelength detector with a Synergi 4U Hydro-RP 80 Å column, 100 $\times$ 21.20 mm, 4 micro) afford the desired glycosylated amino acid mimic (**4.25**, 14 mg, 0.02 mmol 33%) as a yellow foam and a 6:4 mix of α : β anomers by NMR. The diastereomers are denoted α and β in the ^1H and ^{13}C NMR. ^1H

NMR (500 MHz, MeOD) δ 8.00 – 7.96 (1 H ^{α} , m, CH ^{α} _{Triazole}) 7.94 – 7.90 (1 H ^{β} , m, CH ^{β} _{Triazole}) 7.78 (2 H, d, J = 7.1, 2 $\times$ Ph), 7.68 (2 H, t, J = 7.3 2 $\times$ Ph), 7.40 – 7.34 (2 H, m, 2 $\times$ Ph), 7.30 (2 H, quad, J = 7.1, 2 $\times$ Ph), 4.94 – 4.89 (1 H + 1 H ^{α} , 1 $\times$ OCH₂ and H-1 ^{α}), 4.80 – 4.71 (1 H + 1 H ^{β} , 1 $\times$ OCH₂ and H-1 ^{β}), 4.56 (1 H ^{β} , d J = 12.3, 1 $\times$ H-6 ^{β}), 4.52 – 4.44 (2 H + 1 H ^{β} , m, 2 $\times$ CH₂N and 1 $\times$ H-6 ^{β}), 4.42 – 4.35 (2 H, m, Fmoc-CH₂), 4.23 (1 H, t, J = 6.5, Fmoc-CH), 4.14 – 4.07 (1 H + 1 H ^{α} , m, NCH and H-2 ^{α}), 4.03 – 3.83 (2 H + 2 H ^{α} , H-3, H-5 and 2 $\times$ H-6 ^{α}), 3.00 (1 H, s, H-4), 2.83 – 2.73 (1 H ^{β} , m, H-2 ^{β}), 2.53 – 2.46 (1 H, m, 1 $\times$ CH₂CH₂N), 2.28 – 2.18 (1 H, m, 1 $\times$ CH₂CH₂N), 1.94 (3 H ^{α} , s, NHAc ^{α}) and 1.91 (3 H ^{β} , s, NHAc ^{β}) ¹³C NMR (100 MHz, MeOD) δ 174.9 (NCO ^{α}), 174.0 (NCO ^{β}), 173.8 (CO₂H), 158.8 (OC(O)N), 145.5, 145.3, 142.8 (4 $\times$ Ph and CH_{Triazole}), 128.9 (2 $\times$ Ph), 128.4 (2 $\times$ Ph), 126.4 (2 $\times$ Ph), 125.8 (C_{Triazole}) 121.1 (2 $\times$ Ph), 100.9 (C-1 ^{β}), 98.6 (C-1 ^{α}), 73.6 (C-5 ^{α}), 72.7 (C-5 ^{β}), 69.2 (C-3 ^{α}), 69.0 (C-3 ^{β}), 63.1 (C-6 ^{α}), 62.5 (OCH₂) 61.6 (C-6 ^{β}), 59.1 (Fmoc-CH₂), 56.6 (C-2 ^{α}), 56.0 (C-2 ^{β}), 52.9 (NCH), 48.4 (CH₂N and Fmoc-CH), 42.6 (C-4), 33.4 (CH₂CH₂N), 23.3 (NHAc ^{β}) and 22.9 (NHAc ^{α}). HRMS (ESI): *m/z* Calcd for C₆₀H₆₈N₁₀O₁₈S₂Na [M+Na]⁺: 1303.4047: Found 1303.4014.

Benzyl 2-*N*-acetamido-2-deoxy- α -D-glucopyranoside (4.26)⁴²

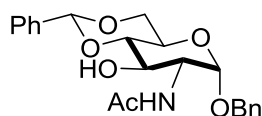


To a solution of *N*-acetylglucosamine (5 g, 22.6 mmol) in dry benzyl alcohol (17.5 mL) was added acetyl chloride (1.6 mL, 22.6 mmol). The reaction was stirred at room temperature for 1 h, before being heated to 50 °C and the reaction was stirred at this temperature for 4 h. After this time, the reaction mixture was slowly added to ice-cold

diethyl ether (175 mL) and was stirred at 0 °C for 1.5 h. The resulting solid was filtered off and dried under reduced pressure to afford the benzyl compound (**4.26**, 7.1 g, 22.6 mmol, quantitative) as a white solid and exclusively as the α anomer, which was used without further purification. $[\alpha]_D +137.6$ (c 1.0, MeOH). ^1H NMR (400 MHz, d_6 -DMSO) δ 7.86 (1 H, d, $J = 8.1$ NH), 7.40 – 7.25 (5 H, m, $5 \times \text{Ph}$), 4.82 (2 H, s br, 3-OH and 4-OH), 4.70 (1 H, d, $J = 3.4$, H-1), 4.66 (1 H, d, $J = 12.5$, $1 \times \text{CH}_2\text{Ph}$), 4.49 (1 H, s, 6-OH), 4.42 (1 H, d, $J = 12.5$, $1 \times \text{CH}_2\text{Ph}$), 3.73 – 3.63 (2 H, m, H-2 and $1 \times \text{H-6}$), 3.57 – 3.42 (3 H, m, H-3, H-5 and $1 \times \text{H-6}$), 3.16 (1 H, t, $J = 9.1$, H-4) and 1.83 (3 H, s, COCH_3). ^{13}C NMR (100 MHz, d_6 -DMSO) δ 169.4 (CO), 138.0 (Ph), 128.2 ($2 \times \text{Ph}$), 127.5 ($2 \times \text{Ph}$), 126.4 (Ph), 95.9 (C-1), 73.1 (C-5), 70.9 (C-3), 70.6 (CH_2Ph), 67.7 (C-4), 60.9 (C-6), 53.8 (C-2) and 22.6 (COCH_3). MS (ESI): m/z Calcd for $\text{C}_{15}\text{H}_{22}\text{NO}_6$ $[\text{M}+\text{H}]^+$: 312.1; Found 312.1. MP = 167 – 170 °C.

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.⁴²

Benzyl 4,6-O-benzylidene-2-N-acetamido-2-deoxy- α -D-glucopyranoside (**4.27**)³⁵

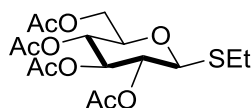


To a solution of **4.26** (6 g, 19 mmol) in dry acetonitrile (80 mL) was added benzaldehyde dimethyl acetal (7.3 mL, 48 mmol) and camphorsulfonic acid (0.23 g, 1 mmol). The reaction was stirred at room temperature for 6 h, before being quenched by the addition of triethylamine (0.4 mL). After stirring for a further 30 min, the solvent was removed under reduced pressure and the resulting solid washed with methanol ($2 \times$

100 mL). The solid was dried under reduced pressure to afford the protected sugar (**4.27**, 2.2 g, 5.5 mmol, 29%) as a white solid and exclusively the α anomer, which was used without further purification. *R_f* 0.25 (EtOAc). $[\alpha]_D +85.5$ (c 0.2, MeOH). ^1H NMR (400 MHz, d_6 -DMSO) δ 7.50 – 7.26 (10 H, m, 10 $\times$ Ph), 5.62 (1 H, s, CHPh), 4.80 (1 H, d, J = 3.5, H-1), 4.71 (1 H, d, J = 12.7, 1 $\times$ CH₂Ph), 4.49 (1 H, d, J = 12.7, 1 $\times$ CH₂Ph), 4.15 (1 H, dd, J = 9.0, 3.8, 1 $\times$ H-6), 3.86 (1 H, ddd, J = 12.2, 8.4, 3.7, H-2), 3.78 – 3.67 (3 H, m, H-3, H-5 and 1 $\times$ H-6), 3.51 (1 H, t, J = 8.9, H-4) and 1.85 (3 H, s, COCH₃). ^{13}C NMR (100 MHz, d_6 -DMSO) δ 169.5 (CO), 137.74 (Ph), 137.71 (Ph), 128.3 (2 $\times$ Ph), 128.1 (Ph), 128.0 (2 $\times$ Ph), 127.64 (2 $\times$ Ph), 127.52 (Ph), 126.4 (2 $\times$ Ph), 100.9 (CHPh), 97.0 (C-1), 82.1 (C-4), 68.6 (C-3), 68.0 (C-5), 67.2 (CH₂Ph), 62.9 (C-6), 54.2 (C-2) and 22.5 (COCH₃). MS (ESI): *m/z* Calcd for C₂₂H₂₆NO₆ [M+H]⁺: 400.2: Found 400.2. MP = 218 – 220 °C.

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.³⁵

Ethyl 2,3,4,6-tetra-*O*-acetyl-1-thio- β -D-glucopyranoside (**4.28**)^{36, 43}

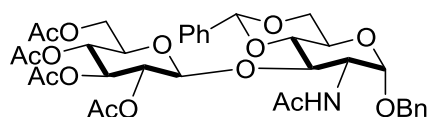


To a solution of penta-*O*-acetyl glucopyranoside (1.5 g, 3.8 mmol) in dry dichloromethane (20 mL) was added ethanethiol (1.4 mL, 19.2 mmol). The solution was cooled to 0 °C and BF₃·OEt₂ (0.67 mL, 5.7 mmol) was added dropwise and the reaction was stirred for 16 h, slowly warming to room temperature. The reaction was quenched by the addition of triethylamine (1 mL) and directly purified by flash chromatography

(7:3, hexane:ethyl acetate) to afford the thiol sugar (**4.28**, 1.1 g, 2.8 mmol, 74%) as a white solid and exclusively as the β anomer. R_f 0.2 (7:3, hexane:ethyl acetate). $[\alpha]_D -7.2$ (c 1.0, CHCl_3). ^1H NMR (400 MHz, CDCl_3) δ 5.22 (1 H, t, $J = 9.5$, H-4), 5.08 (1 H, t, $J = 9.8$, H-3), 5.03 (1 H, t, $J = 9.3$, H-2), 4.24 (1 H, t, $J = 10.1$, H-1), 4.24 (1 H, dd, $J = 12.4$, 5.0 $1 \times$ H-6), 4.13 (1 H, dd, $J = 12.4$, 2.4, $1 \times$ H-6), 3.71 (1 H, ddd, $J = 7.2$, 4.8, 2.3, H-5), 2.76 – 2.63 (2 H, m, SCH_2), 2.07 (3 H, s, COCH_3), 2.05 (3 H, s, COCH_3), 2.02 (3 H, s, COCH_3), 2.00 (3 H, s, COCH_3) and 1.26 (3 H, t, $J = 7.5$, SCH_2CH_3). ^{13}C NMR (100 MHz, CDCl_3) δ 170.6 (CO), 170.2 (CO), 169.4 ($2 \times$ CO), 83.5 (C-1), 75.8 (C-5), 73.8 (C-2), 69.8 (C-3), 68.3 (C-4), 62.1 (C-6), 24.1 (SCH_2), 20.7 ($2 \times$ COCH_3), 20.58 (COCH_3), 20.55 (COCH_3) and 14.8 (SCH_2CH_3). MS (ESI): m/z Calcd for $\text{C}_{16}\text{H}_{24}\text{O}_9\text{Na}$ $[\text{M}+\text{Na}]^+$: 415.1; Found 415.1. MP = 67 – 69 °C.

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.⁴³

2,3,4,6-Tetra-*O*-acetyl- β -D-glucopyranosyl-(1 $\rightarrow$ 3)-benzyl-4,6-*O*-benzylidene-2-*N*-acetamido-2-deoxy- α -D-glucopyranoside (4.29**)**

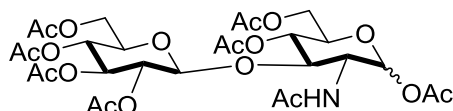


A solution of **4.28** (250 mg, 0.65 mmol), **4.27** (310 mg, 0.8 mmol) and 4 Å molecular sieves (200 mg) in dry dichloromethane (5 mL) was stirred for 1 h. After this time, the solution was cooled to 0 °C and *N*-iodosuccinimide (586 mg, 2.6 mmol) and

trimethylsilyl trifluoromethanesulfonate (0.12 mL, 0.65 mmol) were added and the reaction was stirred for 16 h and slowly warmed to room temperature. The reaction was then diluted with dichloromethane (50 mL) and filtered through Celite. The resulting solution was washed with 5% sodium thiosulfate (20 mL), saturated sodium bicarbonate (20 mL), water (20 mL) and brine (20 mL). The organic layer was dried with sodium sulfate and the solvent removed under reduced pressure. Purification by flash chromatography (1:1, pet. ether:ethyl acetate, elute with ethyl acetate) afforded the protected disaccharide (**4.29**, 116 mg, 0.16 mmol, 25%) as a colourless oil and exclusively as the α anomer. *R*_f 0.41 (ethyl acetate). $[\alpha]_{\text{D}} +26.0$ (c 1.0, CHCl₃). ν_{max} (film)/cm⁻¹ 1748 (C=O), 1654, 1545, 1454, 1372, 1217, 1173, 1119, 1089, 1034, 911, 732 and 699. ¹H NMR (400 MHz, CDCl₃) δ 7.50 – 7.45 (2 H, m, 2 $\times$ Ph), 7.40 – 7.30 (8 H, m, 8 $\times$ Ph), 5.66 (1 H, d, *J* = 9.2, NH), 5.56 (1 H, s, CHPh), 5.12 (1 H, t, *J* = 9.3, H-2'), 5.04 (1 H, t, *J* = 9.6, H-4'), 4.97 (1 H, d, *J* = 7.9, H-3'), 4.94 (1 H, d, *J* = 3.5, H-1), 4.72 (1 H, d, *J* = 3.8, 1 $\times$ CH₂Ph), 4.70 (1 H, d, *J* = 7.6, H-1'), 4.46 (1 H, d, *J* = 11.6, 1 $\times$ CH₂Ph), 4.32 (1 H, td, *J* = 9.9, 3.7, H-2), 4.22 (1 H, dd, *J* = 10.1, 4.4, 1 $\times$ H-6), 4.15 – 4.09 (1 H, m, 1 $\times$ H-6'), 3.99 – 3.93 (2 H, m, H-5' and 1 $\times$ H-6'), 3.87 (1 H, dd, *J* = 9.8, 4.4, H-5), 3.82 – 3.69 (2 H, m, H-4 and 1 $\times$ H-6), 3.49 (1 H, dquad, *J* = 9.6, 4.4, H-3), 1.99 (3 H, s, OAc), 1.97 (3 H, s, OAc), 1.962 (6 H, s, 2 $\times$ OAc) and 1.957 (3 H, s, NHAc). ¹³C NMR (100 MHz, CDCl₃) δ 170.6 (CO), 170.3 (CO), 169.7 (CO), 169.6 (CO), 169.3 (CO), 137.2 (Ph), 136.7 (Ph), 129.0 (Ph), 128.7 (2 $\times$ Ph), 128.4 (Ph), 128.3 (2 $\times$ Ph), 128.2 (2 $\times$ Ph), 126.0 (2 $\times$ Ph), 101.2 (C-1'), 100.3 (CHPh), 97.5 (C-1), 80.1 (C-4), 76.3 (C-5'), 72.8 (C-2'), 71.9 (C-3'), 71.6 (C-3), 70.2 (CH₂Ph), 68.8 (C-6), 68.1 (C-4'), 63.2 (C-5), 61.8 (C-6'), 52.5 (C-2), 23.2 (NHCOCH₃), 20.6 (OCOCH₃), 20.55

(OCOCH₃) and 20.52 (2 × OCOCH₃). HRMS (ESI): *m/z* Calcd for C₃₆H₄₄NO₁₅Na [M+Na]⁺: 752.2536; Found 752.2517.

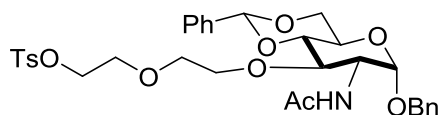
2,3,4,6-Tetra-O-acetyl-β-D-glucopyranosyl-(1→3)--1,3,4,6-tetra-O-acetyl-2-N-acetamido-2-deoxy-D-glucopyranose (4.30)



To **4.29** (180 mg, 0.25 mmol) was added Pd/C (10% w/w, 90 mg, 0.25 mmol) and the mixture was placed under an argon atmosphere. Methanol (3 mL) and acetic acid (0.3 mL) were added and the reaction was placed under a hydrogen atmosphere and stirred at room temperature for 48 h. After this time, the reaction was filtered through Celite and the solvent removed under reduced pressure. To the resulting residue was added dry pyridine (3 mL) and the reaction was cooled to 0 °C. Acetic anhydride (0.11 mL, 1.25 mmol) was added and the reaction was stirred for 16 h, slowly warming to room temperature. The solvent was removed under reduced pressure and the resulting residue was dissolved in chloroform (50 mL) and washed with 1 M HCl (10 mL), saturated sodium bicarbonate (10 mL) and brine (10 mL). The organic layer was dried with sodium sulfate and the solvent removed under reduced pressure. Purification by flash chromatography (1:1, pet. ether:ethyl acetate, elute with ethyl acetate) afforded the protected disaccharide (**4.30**, 66 mg, 0.10 mmol, 39%) as a colourless oil and 7:3 mix of α/β anomers. NMR data given only for α anomer. *R_f* 0.06 (99:1, ethyl acetate:NEt₃). *v*_{max} (film)/cm⁻¹ 1744 (C=O), 1663, 1541, 1432, 1368, 1213, 1131, 1034, 913, 730 and 648. ¹H NMR (400 MHz, CDCl₃) δ 6.05 (1 H, d, *J* = 3.7, H-1), 5.58 (1 H, d, *J* = 9.5, NH), 5.24 – 5.13 (1 H, m, H-2'), 5.09 – 4.97 (2 H, m, H-4 and H-4'), 4.89 – 4.79 (1 H,

m, H-3'), 4.68 (1 H, d, $J = 7.8$, H-1'), 4.51 (1 H, td, $J = 10.0$, 3.7, H-2), 4.39 (1 H, dd, $J = 12.55$, 4.4, $1 \times$ H-6), 4.27 – 4.15 (2 H, m, $1 \times$ H-6 and $1 \times$ H-6'), 4.07 – 4.02 (1 H, m, $1 \times$ H-6b), 4.01 – 3.95 (1 H, m, H-5), 3.94 – 3.86 (1 H, m, H-5'), 3.74 – 3.66 (1 H, m, H-3), 2.07 (6 H, s, $2 \times$ OAc), 2.04 (3 H, s, OAc), 2.03 (6 H, s, $2 \times$ OAc), 2.01 (3 H, s, OAc), 2.00 (3 H, s, OAc) and 1.97 (3 H, s, OAc). ^{13}C NMR (100 MHz, CDCl_3) δ 170.8 (CO), 170.5 (CO), 170.3 (CO), 169.62 (CO), 169.59 (CO), 169.3 (CO), 169.0 (CO), 168.6 (CO), 100.5 (C-1'), 91.2 (C-1), 76.2 (C-5'), 72.6 (C-2'), 71.54 (C-3), 71.49 (C-3'), 69.7 (C-5), 68.1 (C-4), 68.0 (C-4'), 61.72 (C-6) 61.68 (C-6'), 51.2 (C-2), 21.1 (NHCOCH_3), 21.0 (OCOCH_3), 20.8 (OCOCH_3), 20.7 (OCOCH_3) and 20.6 ($4 \times$ OCOCH_3). HRMS (ESI): m/z Calcd for $\text{C}_{28}\text{H}_{39}\text{NO}_{18}\text{Na}$ $[\text{M}+\text{Na}]^+$: 700.2059: Found 700.2052.

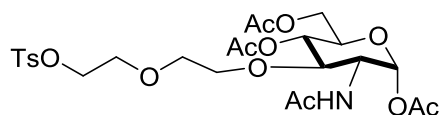
Benzyl 4,6-O-benzylidene-3-diethylene-glycol-tosyl-2-N-acetamido-2-deoxy- α -D-glucopyranoside (4.31)



A solution of **4.27** (0.8 g, 2 mmol) in dry DMF (15 mL) was cooled to 0 °C and sodium hydride (0.11 g, 60% dispersion in oil, 2.8 mmol) was added slowly. The reaction was stirred at 0 °C for 30 min before diethylene glycol di(p-tosylate) (1.66 g, 4 mmol) was added and the reaction was heated to 70 °C. After stirring at 70 °C for 16 h, the reaction was quenched by the addition of saturated ammonium chloride (0.5 mL). The solvent was removed under reduced pressure and the resulting residue was re-dissolved in dichloromethane (50 mL) and washed with water (10 mL). The organic layer was dried with sodium sulfate and the solvent removed under reduced pressure. Purification by

flash chromatography (dichloromethane) afforded the tosylated sugar (**4.31**, 0.6 g, 0.94 mmol, 47%) as a white solid and exclusively as the α anomer. *R_f* 0.44 (ethyl acetate). $[\alpha]_D^{+45.8}$ (c 1.0, CHCl₃). $\nu_{\max}$ (film)/cm⁻¹ 2872, 1650 (C=O), 1553, 1451, 1369, 1175, 1130, 1087, 1052, 1002, 922, 817, 733, 695 and 663. ¹H NMR (400 MHz, CDCl₃) δ 7.76 (2 H, d, *J* = 8.2, 2 $\times$ Ph), 7.51 – 7.45 (2 H, m, 2 $\times$ Ph), 7.40 – 7.28 (10 H, m, 10 $\times$ Ph), 6.03 (1 H, d, *J* = 8.6, NH), 5.56 (1 H, s, CHPh), 5.00 (1 H, d, *J* = 3.7, H-1), 4.72 (1 H, d, *J* = 11.9, 1 $\times$ CH₂Ph), 4.52 (1 H, d, *J* = 11.9, 1 $\times$ CH₂Ph), 4.28 – 4.18 (3 H, m, H-2 and 2 $\times$ H-4') 3.93 (1 H, ddd, *J* = 11.6, 4.1, 2.9, 1 $\times$ H-3'), 3.87 (1 H, dd, *J* = 9.6, 4.6, 1 $\times$ H-6), 3.81 – 3.51 (8 H, m, H-3, H-4, H-5, 1 $\times$ H-6, 2 $\times$ H-1', 1 $\times$ H-2' and 1 $\times$ H-3') 3.47 (1 H, ddd, *J* = 11.0, 4.3, 2.9, 1 $\times$ H-2'), 2.42 (3 H, s, Ph-CH₃) and 1.97 (3 H, s, COCH₃). ¹³C NMR (100 MHz, CDCl₃) δ 170.1 (CO), 144.8 (Ph), 137.3 (Ph), 137.0 (Ph), 132.8 (Ph), 129.8 (2 $\times$ Ph), 128.9 (Ph), 128.5 (2 $\times$ Ph), 128.2 (2 $\times$ Ph), 128.13 (2 $\times$ Ph), 128.07 (Ph), 127.9 (2 $\times$ Ph), 125.9 (2 $\times$ Ph), 101.2 (CHPh), 97.5 (C-1), 82.6 (C-4), 77.4 (C-3), 71.53 (C-2'), 71.50 (C-3'), 70.0 (CH₂Ph), 69.3 (C-1'), 68.9 (C-4'), 68.6 (C-5), 63.0 (C-6), 52.9 (C-2), 23.2 (COCH₃) and 21.6 (PhCH₃). HRMS (ESI): *m/z* Calcd for C₃₃H₄₀NO₁₀SNa [M+Na]⁺: 642.2367; Found 642.2362. MP = 128 – 130 °C.

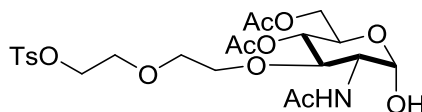
1,4,6-tri-O-acetyl-3-diethylene-glycol-tosyl-2-N-acetamido-2-deoxy- α -D-glucopyranose (4.32)



To **4.31** (600 mg, 0.94 mmol) was added Pd/C (10% w/w, 150 mg, 0.14 mmol) and the mixture was placed under an argon atmosphere. Methanol (1.5 mL) and 1,4 dioxane (1.5 mL) were added and the reaction was placed under a hydrogen atmosphere and

stirred at room temperature for 16 h. After this time, the reaction was filtered through Celite and the solvent removed under reduced pressure. The resulting residue was dissolved in dry dichloromethane (5 mL) and dry pyridine (0.75 mL, 9.4 mmol) and acetic anhydride (0.43 mL, 4.7 mmol) was added and the reaction was stirred for 16 h. The solvent was removed under reduced pressure and purification by flash chromatography (95:5, dichloromethane:methanol) afforded the protected sugar (**4.32**, 440 mg, 0.75 mmol, 79%) as a colourless oil and exclusively as the α anomer. *R_f* 0.45 (95:5, dichloromethane:methanol). $[\alpha]_D +20.1$ (c 1.0, CHCl₃). $\nu_{\max}$ (film)/cm⁻¹ 2954, 1741 (C=O), 1679, 1360, 1233, 1177, 1136, 1044, 1016, 923, 818, 776 and 664. ¹H NMR (400 MHz, CDCl₃) δ 7.78 (2 H, d, *J* = 8.4, 2 $\times$ Ph), 7.38 (2 H, d, *J* = 8.1, 2 $\times$ Ph), 6.24 (1 H, d, *J* = 3.9, H-1), 6.15 (1 H, d, *J* = 8.1, NH), 5.10 (1 H, t, *J* = 9.6, H-4), 4.36 (1 H, ddd, 11.1, 8.1, 3.5, H-2), 4.28 – 4.18 (2 H, m, 1 $\times$ H-6 and 1 $\times$ H-4'), 4.16 – 4.10 (1 H, m, 1 $\times$ H-4'), 4.04 (1 H, dd, *J* = 12.4, 2.1, 1 $\times$ H-6), 3.97 – 3.91 (1 H, m, H-5), 3.85 (1 H, dd, *J* = 11.0, 9.5, 1 $\times$ H-3'), 3.82 – 3.62 (5 H, m, H-3, 2 $\times$ H-1', 1 $\times$ H-2' and 1 $\times$ H-3'), 3.52 – 3.43 (1 H, m, 1 $\times$ H-2'), 2.47 (3 H, s, Ph-CH₃), 2.17 (3 H, s, COCH₃), 2.11 (3 H, s, COCH₃), 2.08 (3 H, s, COCH₃) and 1.98 (3 H, s, COCH₃). ¹³C NMR (100 MHz, CDCl₃) δ 170.8 (CO), 170.4 (CO), 169.2 (CO), 169.0 (CO), 145.3 (Ph), 132.7 (Ph), 130.0 (2 $\times$ Ph), 127.8 (2 $\times$ Ph), 90.9 (C-1), 77.8 (C-3), 72.1 (C-2'), 70.11 (C-3'), 70.06 (C-5), 69.9 (C-4), 69.2 (C-1'), 68.9 (C-4'), 61.8 (C-6), 51.3 (C-2), 23.0 (COCH₃), 21.6 (PhCH₃), 20.93 (COCH₃), 20.87 (COCH₃) and 20.7 (COCH₃). HRMS (ESI): *m/z* Calcd for C₂₅H₃₅NO₁₃SNa [M+Na]⁺: 612.1721; Found 612.1720.

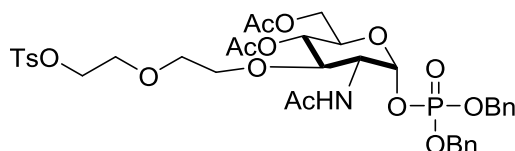
4,6-di-O-acetyl-3-diethylene-glycol-tosyl-2-N-acetamido-2-deoxy- α -D-glucopyranose (4.33)



To a solution of **4.32** (56 mg, 0.1 mmol) in dry DMF (1 mL) was added hydrazine acetate (11 mg, 0.11 mmol). The reaction was stirred at room temperature for 1.5 h, after which the solvent was removed under reduced pressure. The resulting residue was re-dissolved in chloroform (20 mL) and washed with saturated sodium bicarbonate (2 × 10 mL). The aqueous layer was extracted with chloroform (20 mL). The combined organic layers were dried with sodium sulfate and the solvent removed under reduced pressure. Purification by flash chromatography (dichloromethane to 95:5, dichloromethane:methanol) afforded the hemiacetal (**4.33**, 32 mg, 0.06 mmol, 61%) as a colourless foam exclusively as the α anomer. *R_f* 0.35 (9:1, dichloromethane:methanol). $[\alpha]_D^{+10.2}$ (c 1.0, CHCl₃). $\nu_{\max}$ (film)/cm⁻¹ 3298 (O-H, br), 2918, 1744 (C=O), 1657 1544, 1372, 1353, 1242, 1176, 1123, 1040, 924, 819, 777 and 664. ¹H NMR (400 MHz, CDCl₃) δ 7.79 (2 H, d, *J* = 8.2, 2 × Ph), 7.36 (2 H, d, *J* = 7.6, 2 × Ph), 6.42 (1 H, d, *J* = 7.8, NH), 5.30 (1 H, d, *J* = 4.7, H-1), 5.02 (1 H, d, *J* = 9.3, H-4), 4.28 – 4.01 (5 H, m, H-2, 2 × H-6 and 2 × H-4'), 3.86 (1 H, t, *J* = 10.7, H-3), 3.79 – 3.53 (6 H, m, H-5, 2 × H-1', 1 × H-2' and 2 × H-3'), 3.50 – 3.42 (1 H, m, 1 × H-2'), 2.45 (3 H, s, Ph-CH₃), 2.09 (3 H, s, COCH₃), 2.07 (3 H, s, COCH₃) and 2.01 (3 H, s, COCH₃). ¹³C NMR (100 MHz, CDCl₃) δ 170.9 (CO), 170.7 (CO), 169.5 (CO), 145.1 (Ph), 132.6 (Ph), 129.9 (2 × Ph), 127.9 (2 × Ph), 91.6 (C-1), 77.2 (C-3), 71.2 (C-2'), 70.4 (C-3'), 70.2 (C-5), 69.2 (C-4) 68.6 (C-1'), 67.8 (C-4'), 62.8 (C-6), 53.6 (C-2), 23.3 (COCH₃), 21.6 (PhCH₃)

20.9 (COCH₃) and 20.8 (COCH₃). HRMS (ESI): *m/z* Calcd for C₂₃H₃₃NO₁₂SNa [M+Na]⁺: 570.1616; Found 570.1614.

4,6-di-O-acetyl-3-diethylene-glycol-tosyl-2-N-acetamido-2-deoxy- α -D-glucopyranosyl-1-dibenzylphosphate (4.34)



A solution of **4.33** (70 mg, 0.13 mmol) in dry THF (2 mL) was cooled to -78°C and butyl lithium (0.06 mL, 2.5 M in hexane, 0.14 mmol) was added dropwise. The reaction was stirred at -78°C for 5 min before tetrabenzylpyrophosphate (**4.34**, 222 mg, 0.39 mmol) was added in dry THF (1 mL) over 5 min and the reaction was stirred at -78°C for a further 30 min. The reaction was quenched with water (0.5 mL) before being diluted with diethyl ether (20 mL). The mixture was washed with saturated sodium bicarbonate (10 mL) and water (10 mL) and the organic layer was dried with sodium sulfate. The solvent was removed under reduced pressure and purification by flash chromatography (1:1, pet. ether:ethyl acetate, eluting with 99:1, ethyl acetate:NEt₃) afforded the dibenzylphosphate sugar (**4.34**, 8 mg, 1 μmol , 8%) as a white foam exclusively as the α anomer. *R_f* 0.18 (ethyl acetate). $[\alpha]_{\text{D}} +44.8$ (c 0.5, CHCl₃). ν_{max} (film)/cm⁻¹ 1745 (C=O), 1656, 1598, 1542, 1433, 1240, 1134, 1015, 955, 748, 699 and 664. ¹H NMR (400 MHz, CDCl₃) δ 7.79 (2 H, d, *J* = 8.3, 2 $\times$ Ph), 7.41 – 7.32 (12 H, m, 12 $\times$ Ph), 6.17 (1 H, d, *J* = 8.1, NH), 5.75 (1 H, ddd, *J* = 13.7, 5.6, 3.2, H-1), 5.12 – 5.02 (5 H, m, H-4 and 2 $\times$ CH₂Ph), 4.28 – 4.19 (1 H, m, H-2), 4.17 – 4.11 (2 H, m, 2 $\times$ H-4'), 4.10 – 4.07 (1 H, m, 1 $\times$ H-6), 3.99 – 3.89 (2 H, m, H-5 and 1 $\times$ H-6), 3.77 – 3.61

(6 H, m, H-3, 2 × H-1', 1 × H-2' and 2 × H-3'), 3.48 – 3.43 (1 H, m, 1 × H-2'), 2.45 (3 H, s, Ph-CH₃), 2.07 (3 H, s, COCH₃), 1.99 (3 H, s, COCH₃) and 1.89 (3 H, s, COCH₃). ¹³C NMR (100 MHz, CDCl₃) δ 170.7 (CO), 170.4 (CO), 169.2 (CO), 144.9 (Ph), 135.8 (Ph), 135.5 (Ph), 135.3 (Ph), 132.9 (Ph), 129.9 (Ph), 128.9 (Ph), 128.8 (Ph), 128.7 (2 × Ph), 128.5 (2 × Ph), 128.4 (Ph), 128.13 (Ph), 128.10 (Ph), 128.08 (Ph), 128.0 (Ph), 127.8 (Ph) 96.7 (C-1), 79.3 (C-3), 72.3 (C-2'), 71.3 (C-3'), 70.58 (CH₂Ph), 70.54 (CH₂Ph), 70.0 (C-5), 69.9 (C-4) 69.3 (C-1'), 69.0 (C-4'), 61.6 (C-6), 52.3 (C-2), 23.1 (COCH₃), 21.7 (PhCH₃) 20.9 (COCH₃) and 20.6 (COCH₃). ³¹P NMR (162 MHz, MeOD) δ -2.57. HRMS (ESI): *m/z* Calcd for C₃₇H₄₆NO₁₅PSNa [M+Na]⁺: 830.2218: Found 830.2213.

4.8 - Glycosylation Reactions

4.8.1 – Investigating the Use of WZ₃-Thiol-GlcNAc (4.15) as a PglB Substrate

To the TAMRA-peptide (**3.28**, 1 μ L 1 mM solution, 1 nmol) in 10 mM HEPES, 1 mM MnCl₂, 1% Triton X100, pH 7.5 (29 μ L) were added PglB (0.44 μ M, 4 μ L, 1.76 pmol), DTT (40 mM, 1 μ L) and WZ₃-thiol-GlcNAc (**4.17**, 6 μ L, 4 mM, 24 nmol). Control reactions were set up using WZ₃-GlcNAc and without the use of DTT. The reactions were shaken at room temperature and analysed by HPLC (Waters 2795 Separations Module using a Jupiter 4U Proteo 90 Å 250 × 4.60 mm 4 micron column and using a Jasco FP-920 Intelligent Fluorescent Detector, 542 nm excitation, 568 emission), after 20 h.

4.8.2 – Initial Investigations into the Use of WEZ₄-Thiol-GlcNAc (4.17) as a PglB Substrate

To the TAMRA-peptide (**3.28**, 0.5 μ L 1 mM solution, 0.5 nmol) in 10 mM HEPES, 1 mM MnCl₂, 1% Triton X100, pH 7.5 (16 μ L) were added PglB (0.28 μ M, 1 μ L, 0.28 pmol), DTT (40 mM, 1 μ L or 4 μ L) and WEZ₄-thiol-GlcNAc (**4.17**, 2 μ L, 2 mM, 4 nmol). Control reactions were set up using WEZ₄-GlcNAc. The reactions were shaken at room temperature and analysed by HPLC (Waters 2795 Separations Module using a Jupiter 4U Proteo 90 Å 250 × 4.60 mm 4 micron column and using a Jasco FP-920 Intelligent Fluorescent Detector, 542 nm excitation, 568 emission), after 20 h.

4.8.3 – Further Investigations into the Use of WEZ₄-Thiol-GlcNAc (**4.17**) as a PglB Substrate

To the TAMRA-peptide (**3.28**, 0.5 μ L 1 mM solution, 0.5 nmol) in 10 mM HEPES, 1 mM MnCl₂, 1% Triton X100, pH 7.5 (16 μ L) were added PglB (0.28 μ M, 1 μ L, 0.28 pmol), WEZ₄-thiol-GlcNAc (**4.17**, 2 μ L, 4 mM, 8 nmol) and either DTT (0.5 M, 2 μ L) or TCEP (0.5 M, 1 μ L). Control reactions were set up using WEZ₄-GlcNAc. The reactions were shaken at room temperature and analysed by HPLC (Waters 2795 Separations Module using a Jupiter 4U Proteo 90 Å 250 × 4.60 mm 4 micron column and using a Jasco FP-920 Intelligent Fluorescent Detector, 542 nm excitation, 568 emission), after 20 h.

4.8.4 - Glycosylation of a Thiol Model System with GST-EndoA (**3.32**)

To the thiol-GlcNAc-triazole-amino acid (**4.25**, 4.6 mM, 3 μ L, 14.4 nmol) in 50 mM sodium phosphate buffer pH 6 (28 μ L) were added Man₃GlcNAc oxazoline (16 μ L 5 mM solution, 80 nmol), TCEP (0.5 M, 3 μ L) and GST-EndoA (9.6 μ L, 28 μ g/ml, 0.3 μ g, 3.3 pmol). Control reactions were set up with GlcNAc-triazole-amino acid (**3.31**, positive control) and no GST-EndoA (negative control). The reactions were shaken at room temperature. Samples (10 μ L) were taken at 2 h and 24 h, quenched (1 μ L 1% formic acid) and analysed by HPLC (Waters 2795 Separations Module using a Synergi 4U Hydro-RP 80 Å 100 × 4.6 mm 4 micron column and using a Waters 2996 Photodiode Array Detector, detecting at 260 nm).

4.9 – References

- 1 J. R. Rich, A. Szpacenko, M. M. Palcic and D. R. Bundle, *Angew. Chem. Int. Ed.*, 2004, **43**, 613-615.
- 2 K. Pachamuthu and T. Schmidt, *Chem. Rev.*, 2006, **106**, 160-187.
- 3 A. E. Hogan, V. O'Reilly, M. R. Dunne, R. T. Dere, S. G. Zeng, C. O'Brien, S. Amu, P. G. Fallon, M. A. Exley, C. O'Farrelly, X. Zhu and D. G. Doherty, *Clinical Immunology*, 2011, **140**, 196-207.
- 4 M. Jahn, J. Marles, R. A. J. Warren and S. G. Withers, *Angew. Chem. Int. Ed.*, 2003, **42**, 352-354.
- 5 Z. J. Witczak and D. Boryczewski, *Bioorg. Med. Chem. Lett.*, 1998, **8**, 3265-3268.
- 6 M. Blanc-Musser, L. Vigne, H. Driguez, J. Lehmann, J. Steck and K. Urbahns, *Carbohydr. Res.*, 1992, **224**, 59-71.
- 7 G. Sulzenbacher, H. Driguez, B. Henrissat, M. Schulein and G. J. Davies, *Biochemistry*, 1996, **35**, 15280-15287.
- 8 M. Czjke, M. Cicek, V. Zamboni, W. P. Burmeister, D. R. Bevan, B. Henrissat and A. Esen, *Biochem. J.*, 2001, **354**, 37-46.
- 9 F. Castaneda, A. Burse, W. Boland and R. K. H. Kinne, *Int. J. Med. Sci.*, 2007, **4**, 131-139.
- 10 X. Zhu, T. Haaq and R. R. Schmidt, *Org. Biomol. Chem.*, 2004, **2**, 31-33.
- 11 J. Mullegger, M. Jahn, H. M. Chen, R. A. J. Warren and S. G. Withers, *Protein. Eng. Des. Sel.*, 2005, **18**, 33-40.
- 12 M. Jahn, H. Chen, J. Mullegger, J. Marles, R. A. J. Warren and S. G. Withers, *Chem. Commun.*, 2004, , 274-275.

- 13 J. Mullegger, H. M. Chen, R. A. J. Warren and S. G. Withers, *Angew. Chem. Int. Ed.*, 2006, **45**, 2585-2588.
- 14 H. Chen and S. G. Withers, *Carbohydr. Res.*, 2007, **342**, 2212-2222.
- 15 M. O. Lederer, C. M. Dreisbusch and R. M. Bundschuh, *Carbohydr. Res.*, 1997, **301**, 111-121.
- 16 H. Driguez, J. C. McAuliffe, R. V. Stick, M. G. Tilbrook and S. J. Williams, *Aust. J. Chem.*, 1996, **49**, 343-348.
- 17 A. H. Daranas, H. Shimizu and S. W. Homans, *J. Am. Chem. Soc.*, 2004, **126**, 11870-11876.
- 18 M. K. Singh, R. Xu, S. Moebs, A. Kumar, Y. Queneau, S. J. Cowling and J. W. Goodby, *Chem. Eur. J.*, 2013, **19**, 5041-5049.
- 19 J. Elhalabi and K. G. Rice, *Carbohydr. Res.*, 2001, **335**, 159-165.
- 20 S. D. Debenham and E. J. Toone, *Tetrahedron: Asymmetry*, 2000, **11**, 385-387.
- 21 S. Nishimura, M. Hato, S. Hyugaji, F. Feng and M. Amano, *Angew. Chem. Int. Ed.*, 2012, **51**, 3386-3390.
- 22 K. L. Chandra, P. Saravanan and V. K. Singh, *Tet. Lett.*, 2001, **42**, 5309-5311.
- 23 F. Liu, B. Vijayakrishnan, A. Faridmoayer, T. A. Taylor, T. B. Parsons, G. J. L. Bernardes, M. Kowarik and B. G. Davis, *J. Am. Chem. Soc.*, 2014, **136**, 566-569.
- 24 M. Fernandez-Gonzalez, O. Boutureira, G. J. L. Bernardes, J. M. Chalker, M. A. Young, J. C. Errey and B. G. Davis, *Chem. Sci.*, 2010, **1**, 709-715.
- 25 Y. H. Lau and D. R. Spring, *Synlett*, 2011, , 1917-1919.
- 26 E. D. Goddard-Borger and R. V. Stick, *Org. Lett.*, 2007, **9**, 3797-3800.
- 27 C. D. Spicer and B. G. Davis, *Chem. Commun*, 2013, **49**, 2747-2749.

- 28 S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony and B. G. Davis, *Nature*, 2007, **446**, 1105-1109.
- 29 G. Zou, H. Ochiai, W. Huang, Q. Yang, C. Li and L. Wang, *J. Am. Chem. Soc.*, 2011, **133**, 18975-18991.
- 30 B. Li, Y. Zeng, S. Hauser, H. Song and L. Wang, *J. Am. Chem. Soc.*, 2005, **127**, 9692-9693.
- 31 J. Yin, L. Li, N. Shaw, Y. Li, J. K. Song, W. Zhang, C. Xia, R. Zhang, A. Joachimiak, H. Zhang, L. Wang, Z. Liu and P. Wang, *Plos One*, 2009, **4**, e4658.
- 32 C. D. Heidecke, Z. Ling, N. C. Bruce, J. W. B. Moir, T. B. Parsons and A. J. Fairbanks, *Chem. Bio. Chem.*, 2008, **9**, 2045-2051.
- 33 C. Lizak, S. Gerber, S. Numao, M. Aebi and K. P. Locher, *Nature*, 2011, **474**, 350-356.
- 34 H. C. Kolb, M. G. Finn and K. B. Sharpless, *Angew. Chem. Int. Ed.*, 2001, **40**, 2004-2021.
- 35 T. Yamaguchi, D. Hesek, M. Lee, A. G. Oliver and S. Mobashery, *J. Org. Chem.*, 2010, **75**, 3515-3517.
- 36 D. Cancogni and L. Lay, *Synlett*, 2014, **25**, 2873-2878.
- 37 X. Wen, D. C. Crick, P. J. Brennan and P. G. Hultin, *Bioorg. Med. Chem.*, 2003, **11**, 3579-3587.
- 38 P. K. Mandal and J. S. McMurray, *J. Org. Chem.*, 2007, **72**, 6599-6601.
- 39 J. I. Padron and J. T. Vazquez, *Tetrahedron: Asymmetry*, 1995, **6**, 857-858.
- 40 D. Falguni and A. Laurens, *Carbohydr. Res.*, 1990, **202**, 239-255.
- 41 Y. Zhou, S. Yin, Y. Geo, Y. Zhao, M. Goto and L. Han, *Angew. Chem. Int. Ed.*, 2010, **49**, 6852-6855.

42 A. Babic and S. Pecar, *Tetrahedron: Asymmetry*, 2008, **19**, 2265-2271.

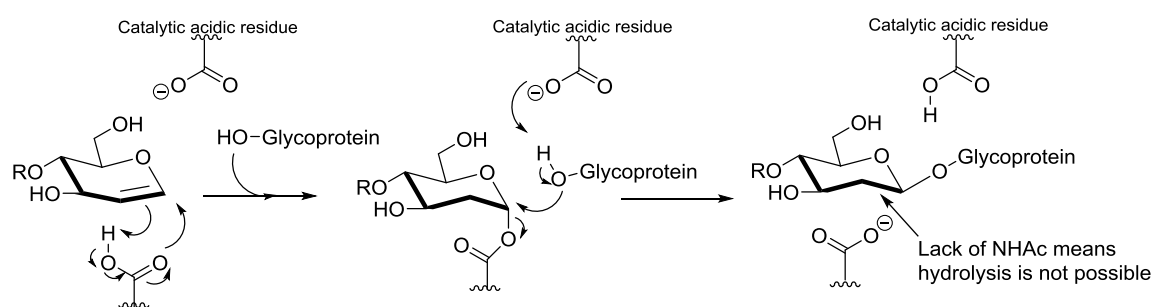
43 C. S. Barry, E. J. Cocinero, P. Carcabal, D. P. Gamblin, E. C. Stanca-Kaposta, S. M. Remmert, M. Fernandez-Gonzalez, S. Rudic, J. P. Simons and B. G. Davis, *J. Am. Chem. Soc.*, 2013, **135**, 16895-16903.

Chapter 5 – Investigating the Use of a Glycal Substrate with Endo Enzymes

Chapter 5 – Investigating the Use of a Glycal Substrate with Endo Enzymes

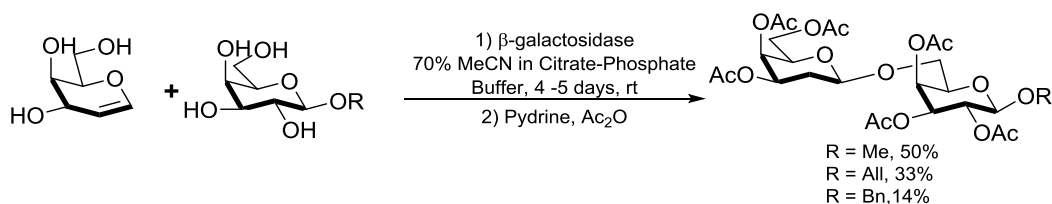
5.1 – Introduction

As seen in Chapters 1 - 3, the main difficulty in the use of Endo enzymes as synthetic tools to attach glycans to proteins is the fact that the product of the glycosylation reaction is the substrate for the natural hydrolysis activity. It is proposed that one solution to this problem maybe to use a glycal as the substrate donor for glycosylation. The advantage of this is that the product of this reaction will not be a substrate for hydrolysis; the lack of an acetylamino group means that the oxazolinium ion intermediate can not form. This type of reaction is known for GH1 enzymes¹⁻⁵ and it is proposed that the mechanism for this reaction starts with the protonation and nucleophilic attack of the glycan double bond from an acidic residue, forming a covalent bond with the enzyme. It is then proposed that the incoming acceptor is deprotonated through the catalytic ionic residue such that it can then attack C-1, forming the desired glycosidic bond and breaking the covalent bond between enzyme and substrate.^{6,7} (Scheme 5.1).



Scheme 5.1: Proposed mechanism for the glycosylation reaction with a glycal and a GH1 enzyme. It is thought that that mechanism may also be applicable to GH18 (EndoS) and GH85 (EndoA) enzymes.^{6,7}

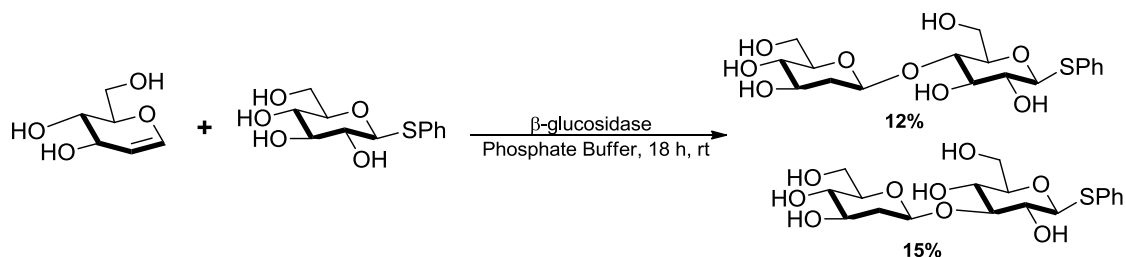
As stated above, this type of reaction has been reported for glycosidases (GH1 enzymes, β -Glucosidase and β -Galactosidase), being used to form disaccharides, from glucal and galactal. Original reports of this reaction used a large amount of organic co-solvent; this was done to limit addition of water across the double bond of the glycal. These reactions also required a low pH buffer, presumably to protonate the glycal. For example, *Petit et al.*³ were able to form acetylated disaccharides from galactal and protected galactose using β -Galactosidase in 70% MeCN, followed by acetylation (Scheme 5.2). Yields of these reactions, however, were low to moderate, with addition of water across the glycal double bond observed for the β -Glucosidase reaction, even though high amounts of organic solvent were being used to avoid this. Similar work has been reported by *Bay and Cantacuzene*¹.



Scheme 5.2: Formation of disaccharides using β -galactosidase.³

With it being desired to use this glycosylation method with glycals to form glycopeptides and glycoproteins, the use of high amounts of organic solvent is not ideal, as organic solvents are often incompatible with protein chemistry. However, *Prade et al.*² showed that it was possible to utilise β -Glucosidase to catalyze the formation of 2-deoxy-D-glucopyranosides at pH 7 and without using organic solvents. The yield of this reaction was only moderate and a 1:1 mix of (1 $\rightarrow$ 3)- β and (1 $\rightarrow$ 4)- β -glycosidic linkages was observed (Scheme 5.3), but this reaction does show that it is possible to use enzymes to catalyse the glycosylation between a glycal and an acceptor under aqueous

conditions. Kröger and Thiem⁵ have also reported similar enzymatic glycosylations with GH1 enzymes in aqueous buffers using a glycal donor.



Scheme 5.3: Glycosylation reaction with glucal and β -glucosidase without organic co-solvents.²

The use of glycals as glycosyl donors in enzymatic glycosylation reactions has only been reported with GH1 enzymes, not for other glycoside hydrolase or glycosyltransferase enzymes. Nor has there been any investigation where the glycan acceptor is bound to an amino acid, peptide or protein, for the formation of glycopeptides and glycoproteins. Here, initial studies into the feasibility of the strategy of using glycals as glycan donors with Endo enzymes in glycosylation reactions are described, with the aim of forming a system in which the Endo enzyme partakes in the glycosylation but then is unable to hydrolyse the product.

5.2 - Glycal Synthesis

Before being able to investigate their use in Endo catalysed glycosylation reactions, it was necessary to synthesise a glycal. It was decided to synthesise a disaccharide glycal, with it being previously shown that the disaccharide is the smallest oxazoline that will function as a glycosylation substrate with an Endo enzyme. The disaccharide to be formed is Man β (1 $\rightarrow$ 4)Glucal (**A**, Figure 5.1), based on the structure of the disaccharide oxazoline (**B**), which in turn has structure based on Man β (1 $\rightarrow$ 4)GlcNAc.

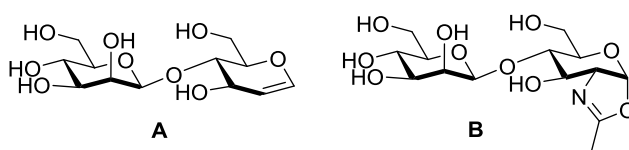
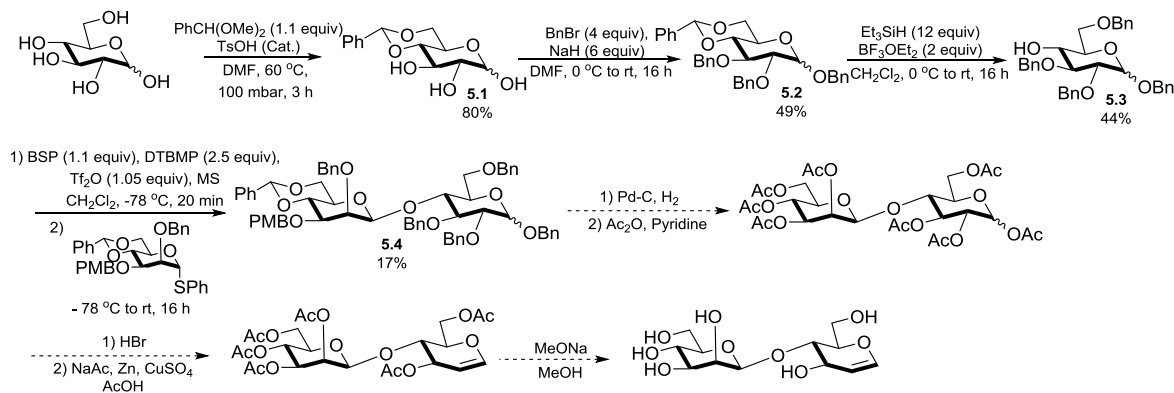


Figure 5.1: Structure of the proposed disaccharide glycal donor substrate (A) compared to the structure of the smallest oxazoline donor (B) with reported activity with EndoA.⁸

The initial planned synthesis involved the chemical glycosylation between a protected glucose unit and a protected mannose unit that had been formed previously.⁹ Once the disaccharide with the desired 1 $\rightarrow$ 4 linkage had been formed, protecting group manipulations, followed by installation of the double bond and subsequent deprotection was expected to yield the desired disaccharide glycal (Scheme 5.4). Starting from glucose, the 4- and 6-positions were protected via benzylidene formation^{10, 11} (**5.1**), with the remaining positions protected with benzyl groups^{11, 12} (**5.2**). Selective ring opening of the benzylidene gave the desired protected glucose unit¹³ (**5.3**). The glycosylation reaction was then attempted, using the previously formed protected mannose unit and previously reported conditions.^{9, 14, 15} This did result in the formation of the desired disaccharide (**5.4**), but the yield was disappointing. Due to both this and the fact that

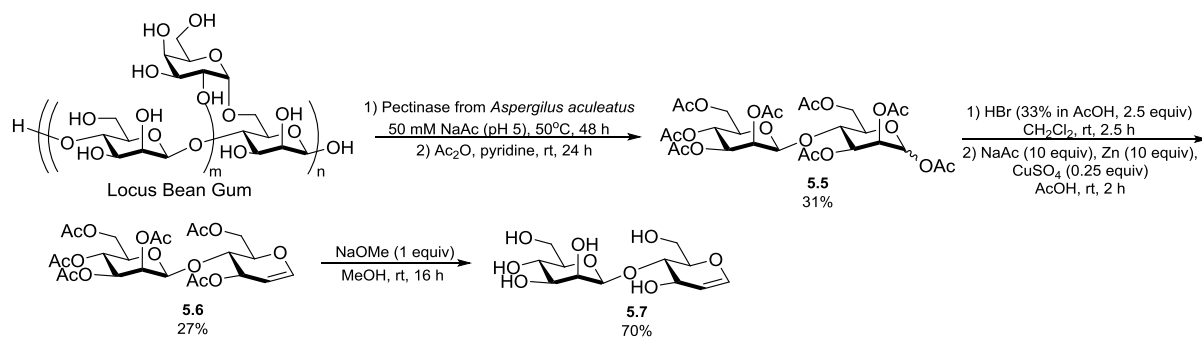
some of steps to form the protected glucose unit were only moderate yielding, this route was not investigated further.



Scheme 5.4: Initial proposed synthesis of glycal 5.1 and yield of initial steps.^{9-11, 15}

With the initial synthesis not yielding the desired disaccharide, an alternative approach was taken. Kumar *et al.*¹⁶ reported that it was possible to extract Man β (1 $\rightarrow$ 4)Man disaccharide from locust bean gum using Pectinase and, through further chemical modifications, they were able to form a protected disaccharide glycal. Given that this method allows direct access to the disaccharide and thus removes a number of chemical steps, this semi-synthetic method was an attractive route to the target molecule.

The first step in this route was extraction of the disaccharide and subsequent acetylation to give peracetylated Man β (1 $\rightarrow$ 4)Man. The yield of this step was relatively low (5.7, Scheme 5.5), but the fact that locust bean gum is inexpensive meant that this reaction could be performed on a scale that gave enough material for the subsequent steps. The double bond was successfully installed (5.8, via the bromide)¹⁶ and a final deprotection step resulted in the formation of the desired disaccharide glycal (5.7).



Scheme 5.5: Synthesis of glycal 2, starting from the extraction of the disaccharide from locus bean gum.¹⁶

5.3 – Formation of EndoA mutants

With the glycal substrate in hand, attention was turned to the next component required to investigate Endo catalysed glycosylation with a glycal donor, an Endo enzyme. It was chosen to investigate EndoA (GH85) due to its previously reported glycosylation activity with a disaccharide acceptor.^{8, 15} However, the proposed mechanism for glycal glycosylation with GH1 enzymes suggests the requirement of two acidic catalytic residues, one amino acid apart (i.e. D/E-X-D/E, where X is any amino acid apart from proline). EndoA contains a catalytic acidic residue (E173) but what would be proposed as the second catalytic residue, the stabilising residue (due to its location and known interaction with the glycan during hydrolysis), is instead an amide (N171). It was therefore required to mutate the stabilising residue to an acidic residue. With it being unknown which acidic residue would be optimal for this proposed reaction, both EndoA N171D and EndoA N171E were targeted.

Site-selective mutagenesis was carried out and the mutant plasmids were successfully formed. The plasmids were then transformed into component cells and over-expressed. Purification by GSTrap columns afforded both GST-EndoA N171D and GST-EndoA N171E (Figure 5.2). With it known that the GST tag has no effect on the activity of other Endo constructs, these mutants were used without further purification.

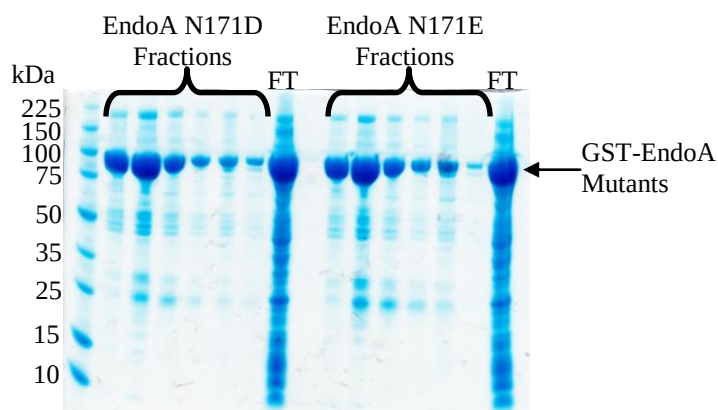
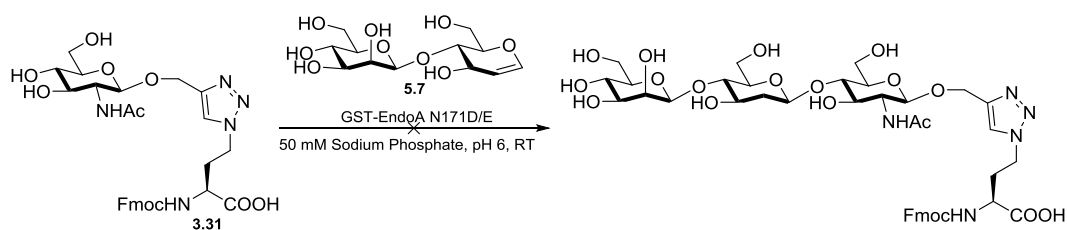


Figure 5.2: Purification of GST-EndoA N171D and GST-EndoA N171E by GSTrap column. FT = Flowthrough

5.4 – Investigating Endo Catalysed Glycosylation with the Glycal Donor

With both the EndoA mutants and the disaccharide glycal donor in hand, investigations could be started into the possible glycosylation reaction using these. To do this, the amino acid with a GlcNAc attached via a triazole⁸ (**3.31**) was used as a model substrate acceptor again, given the previous experience with its use in EndoA glycosylations (see Chapter 3 and 4) and previous reports showing that it can be used as substrate for the attachment of the disaccharide oxazoline using EndoA.⁸

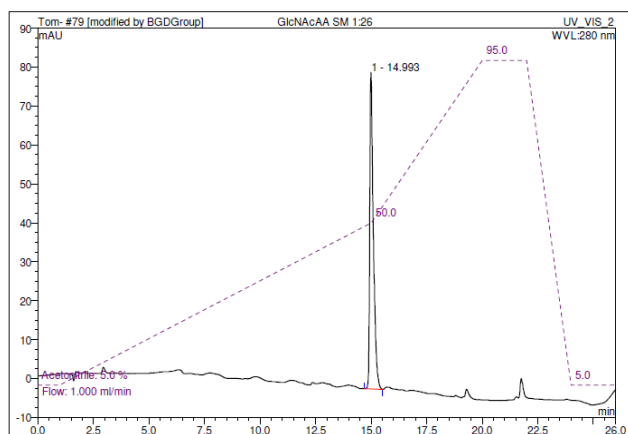
A variety of different reactions were set up using the glycal donor and GlcNAc-amino acid acceptor, varying the EndoA mutants used and their amount (Table 5.1). All these reactions were analysed by HPLC (280 nm) and Mass Spectroscopy (MS). Although it initially seemed that a reaction may have occurred, due to the shift in the retention time of the HPLC peak, disappointingly, MS showed no consumption of starting material in any of the reactions after 48 h (Figure 5.3).



Entry	Enzyme	Enzyme: Acceptor	Glycal	Acceptor Concentration	Result
1	GST-EndoA N171D	1:2000	8 equiv	0.25 mM	No Reaction
2	GST-EndoA N171D	1:200	8 equiv	0.25 mM	No Reaction
3	GST-EndoA N171D	1:20	8 equiv	0.6 mM	No Reaction
4	GST-EndoA N171E	1:2000	8 equiv	0.25 mM	No Reaction
5	GST-EndoA N171E	1:200	8 equiv	0.25 mM	No Reaction
6	GST-EndoA N171E	1:20	8 equiv	0.6 mM	No Reaction

Table 5.1: Conditions attempted for the glycosylation with the glycal donor (5.7) on the GlcNAc-triazole-amino acid acceptor using EndoA

a)



b)

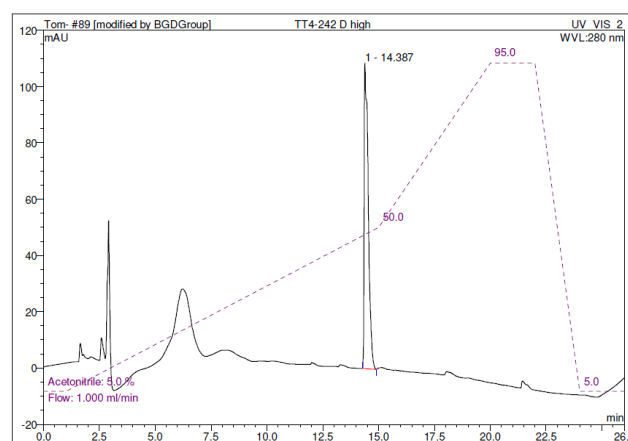


Figure 5.3: HPLC traces of initial glycosylation reactions with EndoA, glycal donor (5.7) and GlcNAc-amino acid (3.31) acceptor, a) HPLC trace of 3.31, b) HPLC trace of entry 2 in Table 5.1 after 1 h. Although a shift can be seen in the retention time of the peak, MS showed only starting material. HPLC traces at 280 nm. All reactions showed a similar HPLC trace.

There are a number of possible different reasons as to why the desired glycosylation reaction may not be occurring. One possibility is that the EndoA mutants are inactive or that the mutation resulted in significantly reduced activity. Alternatively, the glycal may have been acting as an inhibitor for the enzyme, with the second stage of the mechanism, the attack of the glycosyl acceptor, not occurring owing to the stability of the substrate-enzyme complex. Glycals inhibiting enzymes including glycoside hydrolase enzyme has been reported.^{6, 17-19} To confirm whether either of these two potential issues explained why the desired reaction was not taking place, RNase B was treated with EndoA N171D and EndoA N171E both in the presence and the absence of the glycal, to see if the presence of the glycal would affect hydrolysis of the glycan from RNase B (Figure 5.4). This experiment was also conducted with wild-type EndoA to allow comparison between the mutant and the natural hydrolytic activity.

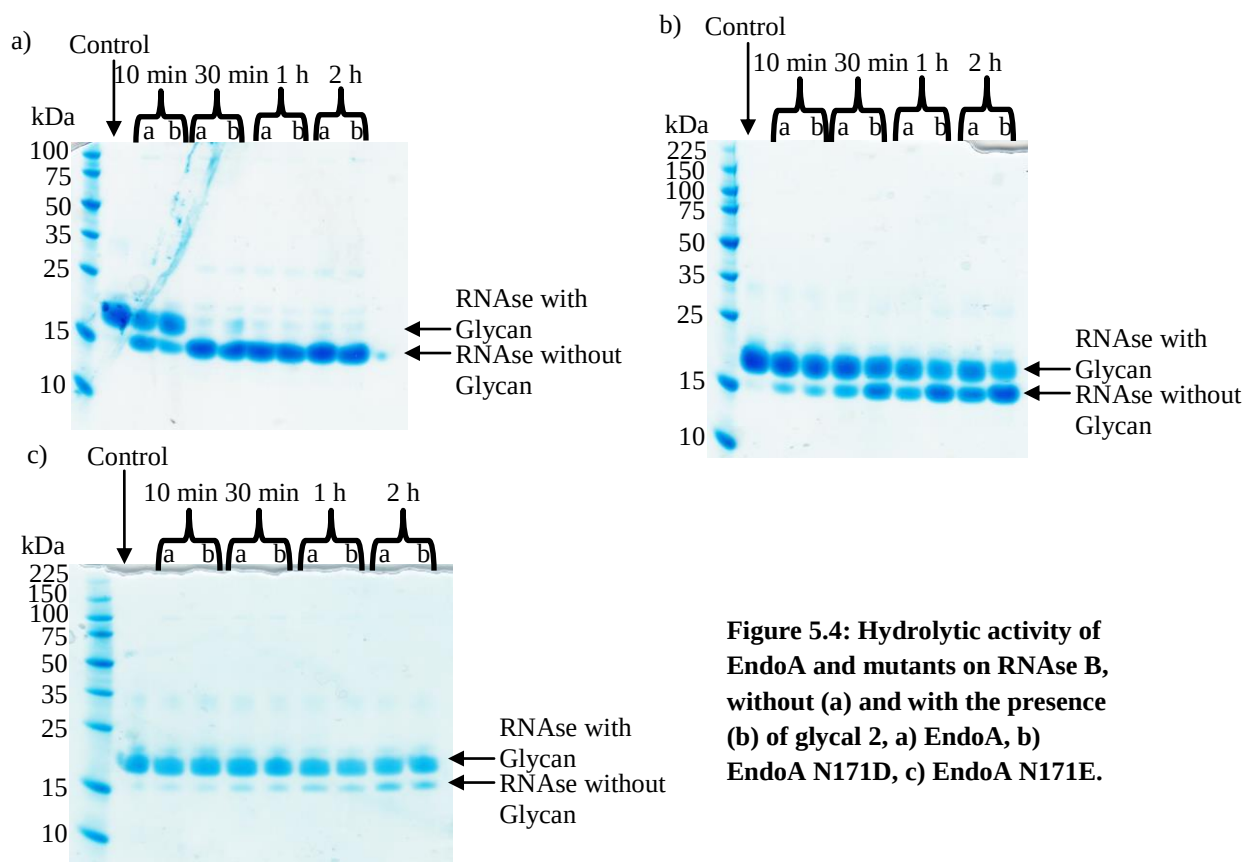


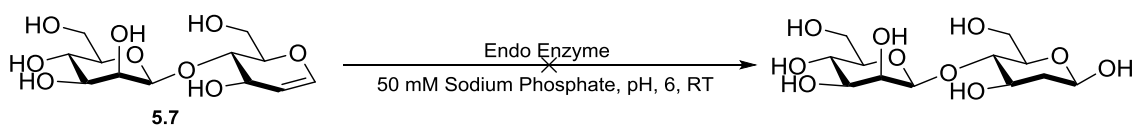
Figure 5.4: Hydrolytic activity of EndoA and mutants on RNase B, without (a) and with the presence (b) of glycal 2, a) EndoA, b) EndoA N171D, c) EndoA N171E.

Two conclusions could be drawn from this experiment. Firstly, the presence of the glycal had no effect on the activities of any of the enzymes, meaning that the glycal was not getting irreversibly covalently bound to the enzyme and was not inhibiting EndoA activity. Second, the activity of the EndoA mutants was significantly less than wild-type EndoA, particularly for EndoA N171E. Although both mutants showed less activity than wild-type enzyme, both EndoA N171D and EndoA N171E were seen to show some activity.

At this point, based on results described thus far, it was decided to take a step back. In Chapter 3 it was seen that the main side reaction that occurs when doing the standard glycosylation reaction with EndoA and an oxazoline donor is the hydrolysis of the oxazoline (i.e. water acts as the acceptor in the EndoA catalysed reaction). Moreover, it has previously been reported that addition of water across the glycal double bond occurs as a side reaction in the glycosyl hydrolyase catalysed glycosylation using a glycal acceptor.³ Knowing this, it would be expected that addition of water across the double bond of the glycal would also occur here, if the glycal could be used as a glycosyl donor in an EndoA catalysed reaction.

To study this, the EndoA mutants were added to the disaccharide glycal (5.7). It was also decided to investigate EndoS^{9, 20, 21} and mutants thereof, to see if they could hydrolyse the glycal. EndoS belongs to a different glycoside hydrolyase family to EndoA (GH18 and GH 85 respectively) and so may show different activity. EndoS and EndoS D233E were used; as described above for EndoA, the mechanism requires the stabilising residue in the Endo enzyme to be acidic, which is true for wild type EndoS.

Since it had been previously formed (see Chapter 3), EndoS D233E was tested as well. The catalytic residue in EndoS is E235, and hence is acidic and fulfils the requirement of the need of two active acidic residues in the active site. All the reactions were carried out on a large scale (2 mg of glycal), allowing the reaction to be followed by both MS and NMR. Reactions with all four enzymes were set up and were conducted at a high concentration (Table 5.2). Samples were taken after 24 h and 48 h.

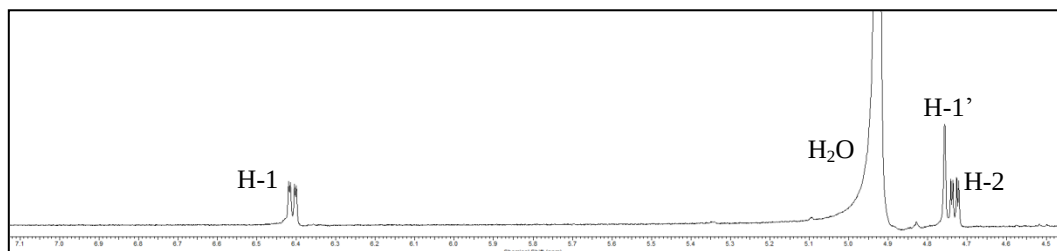


Entry	Enzyme	Enzyme:Glycal	Glycal Concentration	Result
1	GST-EndoA N171D	1:100	32 mM	No Reaction
2	GST-EndoA N171E	1:100	32 mM	No Reaction
3	GST-EndoS	1:100	32 mM	No Reaction
4	GST-EndoS D233E	1:100	32 mM	No Reaction

Table 5.2: Conditions attempted for the Endo catalysed hydrolysis of the disaccharide glycal (5.7).

Disappointingly, only starting material was seen again in all reactions, with no change in the position of H-1 or H-2 observed in the NMR after 48 h (Figure 5.5). This would suggest the enzymes tested here are unable to process the disaccharide glycal.

a)



b)

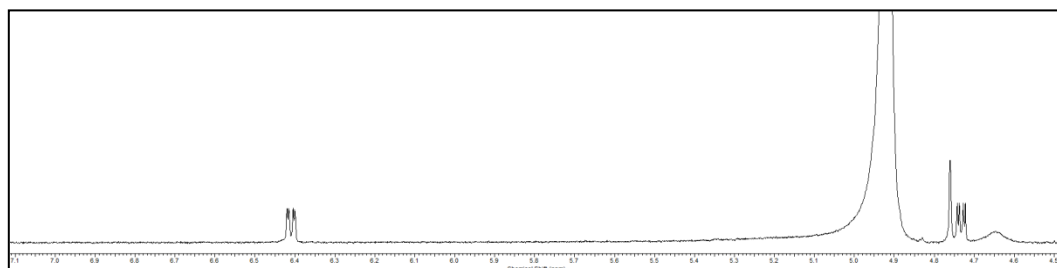


Figure 5.5: NMR spectra when investigating Endo catalysed hydrolysis of the disaccharide glycal (5.7), a) NMR of the 5.7, b) NMR of entry 1, Table 5.2.

5.5 – Conclusions and Future Work

From these initial studies here, it would seem that the EndoA mutants and EndoS and mutants thereof are unable to process the disaccharide glycal. There are two potential reasons for why this maybe the case. Firstly, it is possible that the glycal substrate is too small to be an effective substrate. It has been seen previously in this work that the analogous disaccharide oxazoline substrate was not an effective substrate for the glycosylation reaction with EndoA (Chapter 3), whereas the tetrasaccharide oxazoline gave the glycosylation product. To fully investigate whether Endo enzymes and Endo mutants can process a glycal substrate it may be necessary to synthesise the tetrasaccharide glycal, an analogue of the tetrasaccharide oxazoline donor for which the EndoA glycosylation reaction is known to be facile. However, such a synthesis would not be trivial. It should be noted, however, that there are reports of the use of the disaccharide oxazoline with EndoA for the glycosylation reaction.^{15, 22} For EndoS, the smallest glycan used as the donor for glycosylation is a tetrasaccharide, so it is possible that the disaccharide glycal donor is too small to function as an effective substrate with this enzyme.

The second potential reason is that EndoA, EndoS and mutants are simply unable to function like a GH1 enzyme. The natural hydrolysis mechanism of GH1 enzymes goes via a different mechanism and intermediate compared to the mechanism of EndoS and EndoA. The mechanism of EndoS and EndoA involves the formation of the oxazolinium-ion intermediate,²³ whereas that of GH1 enzymes goes via an enzyme-substrate complex²⁴ (for full details of the mechanisms, see Chapter 1). The mechanism for glycosylation with a glycal donor is proposed to be very similar to the GH1

hydrolysis mechanism.^{2, 6} The fact that Endo enzymes' hydrolysis mechanism is different to that of GH1 enzymes, for an Endo enzyme to process the glycal it would be necessary to adopt a GH1-like hydrolysis mechanism, something which the Endo enzyme will not be optimised for.

Evidence for this comes from the crystal structures of the enzymes. Looking at the crystal structure of SsβG²⁵ (Figure 5.6), a GH1 enzyme, in complex with a glycan (blue), it can be seen that one of the catalytic acidic residue (Glu387, yellow) is bound to the glycan, with the glycan and residue at the correct distance and angle to make this possible. The other catalytic acidic residue, which protonates the out-going acceptor (Glu206, also yellow), is above the glycan, in such a position as to allow catalytic activity.

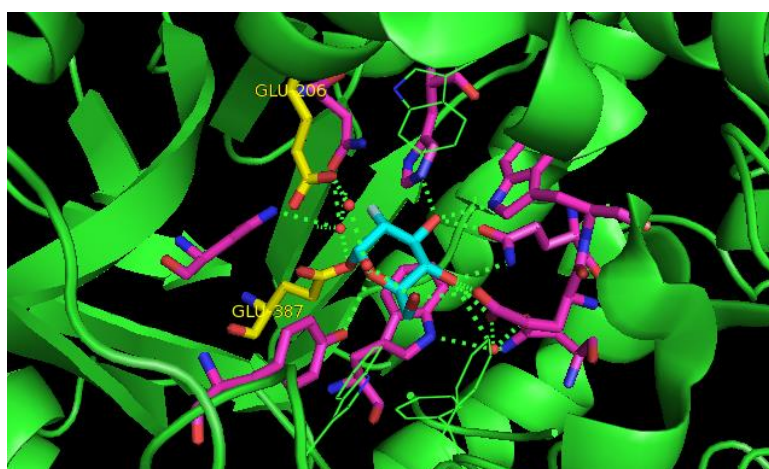


Figure 5.6: Crystal structure of SsβG in complex with a glycal.²⁵

Comparing the structure of SsβG to EndoA²³ (Figure 5.7), it can be seen that the residue that stabilises the oxazolinium ion intermediate (Gln171, yellow) and is analogous to the residue that forms the bond between the GH1 enzyme and substrate in SsβG, is both at a different angle to and further away from the glycan substrate (blue). Such, it would be

very difficult/impossible for EndoA N171D/E to process the glycal via a substrate-enzyme complex as proposed in the mechanism. It should also be noted that the acidic catalytic residue (Glu 171, yellow) that protonates the out-going acceptor is in a similar position to the analogous residue in Ss β G (Glu206). A caveat is that the crystal structure of any enzyme is only a snapshot of a single confirmation; in solution the enzyme structure will be dynamic and the enzyme may adopt other confirmations. Hence, it is possible that EndoA does adopt a confirmation closer to that of Ss β G.

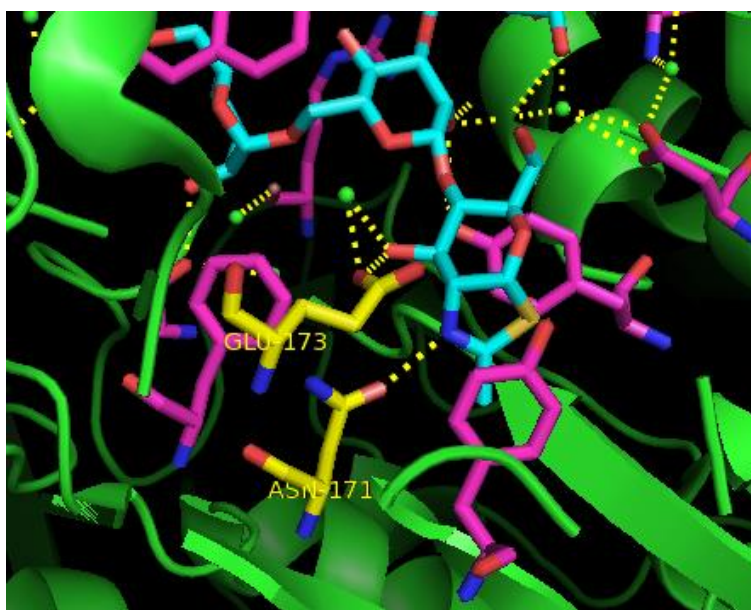
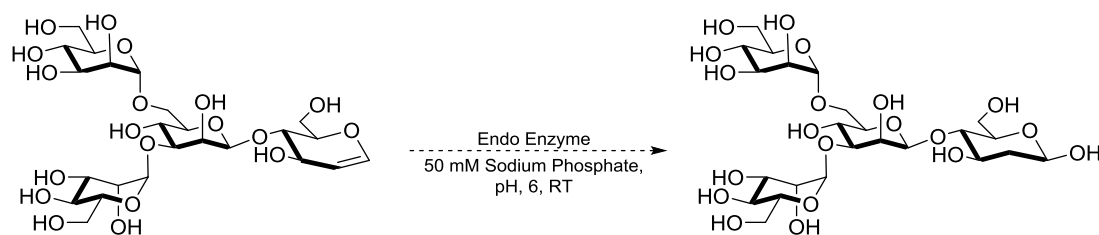


Figure 5.7: Crystal structure of EndoA in complex with a tetrasaccharide oxazoline.²³

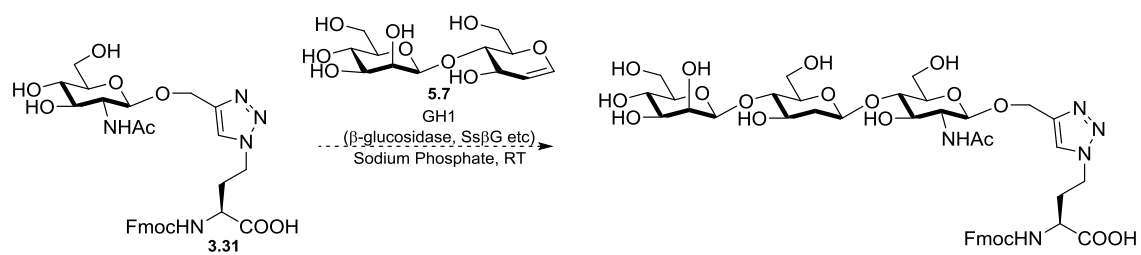
There are three potential paths going forward for the use of glycal donors with Endo enzymes to form glycopeptides and glycoproteins. First would be the synthesis and testing of a larger glycal donor (tetrasaccharide) as previously mentioned (Scheme 5.6).



Scheme 5.6: Potential tetrasaccharide glycal to be formed and potential reaction to test reactivity with an Endo enzyme.

The second possible way forward would be to do a large mutagenesis study to try to form an Endo enzyme that is able to process the glycal donor. This would most likely involve directed evolution mutagenesis,^{26, 27} mutating the active site of the Endo enzyme so the mechanism of hydrolysis is similar to GH1 enzymes.

Finally, a step back could be taken and work focused on GH1 enzymes and their application to form glycopeptide and glycoproteins with glycal donors. As stated at the beginning of this chapter, the use of glycals as a glycan donor to glycosylate a glycan acceptor attached to a peptide or protein is unknown. Work with GH1 enzymes and glycal donors for glycosylation thus far have only been concerned with the formation of di- or trisaccharides unattached to amino acids or peptides. With it known that GH1 enzymes are able to glycosylate with a monosaccharide glycal donor, it would perhaps be more sensible to evolve this work further and investigate whether it is possible to apply this reaction with a disaccharide glycal donor and if it is possible to use a glycan acceptor that is bound to an amino acid (Scheme 5.7). Should both of these reactions yield the desired glycosylation products, further research could focus on the use of glycan acceptors attached to peptides and proteins and using GH1 enzymes with a glycal donor to form glycopeptides and glycoproteins.



Scheme 5.7: Potential glycosylation reaction with a GH1 enzyme.

5.6 - Biological Experimental

See Chapter 2 for protocols for SDS-PAGE and Preparation of Lysogeny Agar Plates and see Chapter 3 for protocols for EndoA Plasmid Mutagenesis, Expression of GST-EndoA mutants and Purification of GST-EndoA mutants.

EndoA Mutagenesis Primers

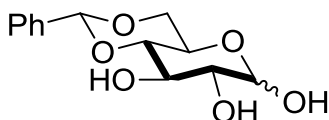
Primer	Code (mutated bases highlighted)
EndoA N171D Forward	GACGGCTGGTTTATTGACCAAGAAACAGAAGG
EndoA N171D Reverse	CCTTCTGTTTCTTGGTCAATAAACCAGCCGTC
EndoA N171E Forward	GGTTTGACGGCTGGTTTATTGAACAAGAAACAGAAGG GGCAG
EndoA N171E Reverse	CTGCCCCTTCTGTTTCTTGTTCAATAAACCAGCCGTCA AACC

Protein Yields

Protein	Yield
GST-EndoA N171D	36 mg/L
GST-EndoA N171E	36 mg/L

5.7- Chemistry Experimental

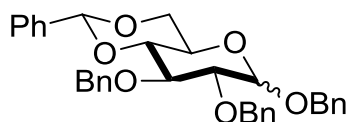
See Chapter 3 for General Chemistry Experimental.

4,6-O-benzylidene-D-glucopyranose (5.1)²⁸

To a solution of D-glucose (10 g, 55.6 mmol) in DMF (40 mL) was added benzaldehyde dimethyl acetal (9.3 mL, 61.2 mmol) and *p*-toluenesulfonic acid (80 mg, 0.42 mmol). The flask was attached to a rotary evaporator and stirred at 50 °C, 100 mbar for 3 hours. After this time, the solvent was removed under reduced pressure. Purification by flash chromatography (9:1, CHCl₃:methanol) afforded the protected sugar (**5.1**, 11.98 g, 44.7 mmol, 80%) as a white solid and a 2:8 mix α : β anomers by NMR. *R_f* 0.15 (9:1, CHCl₃:methanol). ¹H NMR (400 MHz, *d*₆-DMSO) δ 7.46 (2 H, m, 2 $\times$ Ph), 7.39 (3 H, m, 3 $\times$ Ph), 6.86 (1 H ^{β} , d, *J* = 6.7, OH ^{β} -1), 6.58 (1 H ^{α} , d, *J* = 4.7, OH ^{α} -1), 5.59 (1 H ^{β} , s, PhCH ^{β}), 5.58 (1 H ^{α} , s, PhCH ^{α}), 5.28 (1 H ^{β} , d, *J* = 4.6, OH ^{β} -3), 5.20 (1 H ^{β} , d, *J* = 5.0, OH ^{β} -2), 5.13 (1 H ^{α} , d, *J* = 5.0, OH ^{α} -3), 4.80 (1 H ^{α} , d, *J* = 7.6, OH ^{α} -2), 4.48 (1 H ^{β} , t, *J* = 7.2, H ^{β} -1), 4.39 (1 H ^{α} , t, *J* = 7.2, H ^{α} -1), 4.18 (1 H ^{β} , dd, *J* = 10.4, 4.4, 1 $\times$ H ^{β} -6), 4.11 (1 H ^{α} , dd, *J* = 9.9, 4.6, 1 $\times$ H ^{α} -6), 3.86 (1 H ^{α} , m, H ^{α} -5), 3.72 – 3.62 (2 H ^{α} + 1 H ^{β} , m, 1 $\times$ H ^{β} -6, 1 $\times$ H ^{α} -6 and H ^{α} -3), 3.47 – 3.62 (2 H ^{α} + 3 H ^{β} , m, H ^{α} -2, H ^{α} -4, H ^{β} -3, H ^{β} -4 and H ^{β} -5) and 3.09 – 3.01 (1 H ^{β} , ddd, *J* = 13.0, 7.8, 5.2, H ^{β} -2) ¹³C NMR (100 MHz, *d*₆-DMSO) δ 134.1 (Ph), 133.4 (Ph), 133.2 (2 $\times$ Ph), 131.63 (Ph), 131.59 (Ph), 106.1 (C ^{β} HPh), 105.9 (C ^{α} HPh), 102.8 (C ^{β} -1), 98.4 (C ^{α} -1), 87.0 (C ^{α} -4), 86.1 (C ^{β} -4), 81.0 (C ^{β} -2), 78.15 (C ^{α} -2), 78.11 (C ^{β} -3), 74.9 (C ^{α} -3), 73.7 (C ^{α} -6), 73.3 (C ^{β} -6), 71.0 (C ^{β} -5) and 67.2 (C ^{α} -5). MS (ESI): *m/z* Calcd for C₁₃H₁₆O₆Na [M+Na]⁺: 291.1: Found 291.1. MP = 164 – 167 °C.

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.²⁸

Benzyl 4,6-O-benzylidene-2,3-di-O-benzyl-D-glucocopyranoside (5.2)¹¹

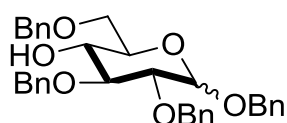


To a solution of **5.1** (4 g, 15.2 mmol) in dry DMF (60 mL) at 0 °C was added sodium hydride (3.65g 60% dispersion in oil, 91.2 mmol) in portions. The reaction was stirred at 0 °C for 15 min before benzyl bromide (6.6 ml, 60.8 mmol) was added over 5 min. The reaction was allowed to warm slowly to room temperature. After 1 h, an addition portion of dry DMF (15 mL) was added to solubilise the resulting solid. The reaction was allowed to stir at room temperature for 16 h before being quenched by the addition of methanol (10 mL). Water (100 mL) was added and the reaction was extracted with diethyl ether (3 × 200 mL). The organic layer was dried with sodium sulfate and the solvent removed under reduced pressure. Purification by flash chromatography (9:1, pet. ether:ethyl acetate) afforded the protected sugar (**5.2**, 4.00 g, 7.4 mmol, 49%) as a off white solid and a 4:6 mix α : β anomers by NMR. R_f 0.31 (9:1, pet. ether:ethyl acetate). ^1H NMR (400 MHz, CDCl_3) δ 7.59 (2 H, d, J = 6.6, 2 × Ph), 7.53 – 7.30 (18 H, m, 18 × Ph), 5.67 (1 H^β , s, H^β -1), 5.64 (1 H^α , s, H^α -1), 5.08 – 4.65 (7 H, m, 6 × CH_2Ph and CHPh), 4.47 (1 H^β , dd, J = 10.5, 5.0, 1 × H^β -6), 4.30 (1 H^α , dd, J = 10.2, 4.9, 1 × H^β -6) and 4.05 – 3.47 (5 H, m, H-2, H-3, H-4, H-5 and 1 × H-6). ^{13}C NMR (100 MHz, CDCl_3) δ 138.9, 138.6, 138.4, 138.3, 137.5, 137.4, 137.2, 137.1 (4 × Ph, α/β), 129.1 – 125.4 (20 × Ph), 103.2 (C^β -1), 101.3 (C^βHPh), 101.2 ($\text{C}^\alpha\text{HPh}$), 96.6 (C^α -1), 82.3 (C^α -4), 82.2 (C^β -2), 81.6 (C^β -4), 81.0 (C^β -3), 79.3 (C^α -2), 78.8 (C^α -3), 75.5, 75.4, 75.2, 73.5,

71.7, 69.4 ($3 \times \text{CH}_2\text{Ph}$, α/β), 69.1 ($\text{C}^\alpha\text{-6}$), 68.9 ($\text{C}^\beta\text{-6}$), 66.1 ($\text{C}^\beta\text{-5}$) and 62.7 ($\text{C}^\alpha\text{-5}$). MS (ESI): m/z Calcd for $\text{C}_{34}\text{H}_{34}\text{O}_6\text{Na}$ $[\text{M}+\text{Na}]^+$: 561.2; Found 561.3. MP = 132 – 134 °C.

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.¹¹

Benzyl 2,3,6-tri-*O*-benzyl-D-glucocopyranoside (5.3)²⁹

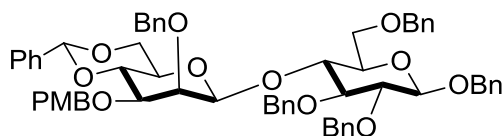


A solution of **5.2** (3.9 g, 7.3 mmol) in dry CH_2Cl_2 (80 mL) was cooled to 0 °C and boron trifluoride-diethyl ether (1.8 mL, 14.5 mmol) and triethylsilane (13.9 mL, 87 mmol) were both added. After stirring for 1 h at 0 °C, the reaction was warmed to room temperature and stirred for a further 16 h. After this time, the reaction was quenched by the addition saturated sodium bicarbonate (30 mL). The reaction was extracted with CH_2Cl_2 (2×50 mL) and the combined organic layers were washed with brine (50 mL) and dried with sodium sulfate. The solvent was removed under reduced pressure and purification by flash chromatography (7:3, pet. ether:ethyl acetate) afforded the protected sugar (**5.3**, 1.71 g, 3.17 mmol, 44%) as a light orange oil and a 4:6 mix of α/β anomers. R_f 0.34 (7:3, pet. ether:ethyl acetate). ^1H NMR (400 MHz, CDCl_3) δ 7.49 – 7.29 (20 H, m, $20 \times \text{Ph}$), 5.10 – 4.95 (2 H, m, CH_2Ph), 4.92 (1 H^α , d, $J = 3.5$, $\text{H}^\alpha\text{-1}$) 4.84 – 4.27 (6 H + 1H^β , m, $6 \times \text{CH}_2\text{Ph}$ and $\text{H}^\beta\text{-1}$), 3.96 – 3.48 (6 H, m, H-2, H-3, H-4, H-5, $2 \times \text{H-6}$). ^{13}C NMR (100 MHz, CDCl_3) δ 138.7, 138.4, 138.12, 138.10, 138.07, 138.04, 137.4, 137.2 ($4 \times \text{Ph}$, α/β), 128.7 – 127.0 ($20 \times \text{Ph}$), 102.6 ($\text{C}^\beta\text{-1}$), 95.5 ($\text{C}^\alpha\text{-1}$), 84.2 ($\text{C}^\beta\text{-2}$), 81.8 ($\text{C}^\beta\text{-3}$), 81.6 ($\text{C}^\alpha\text{-2}$), 79.6 ($\text{C}^\alpha\text{-3}$), 75.5, 75.4, 74.9 ($1.5 \times \text{CH}_2\text{Ph}$, α/β), 74.2 ($\text{C}^\alpha\text{-$

4) 73.7, 73.6, 72.3 ($1.5 \times \text{CH}_2\text{Ph}$, α/β), 71.5 ($\text{C}^\beta\text{-4}$), 71.2 ($0.5 \times \text{CH}_2\text{Ph}$, α/β), 70.8 ($\text{C}^\beta\text{-5}$), 70.29 ($\text{C}^\alpha\text{-5}$), 70.25 ($0.5 \times \text{CH}_2\text{Ph}$, α/β) 69.4 ($\text{C}^\alpha\text{-6}$) and 69.1 ($\text{C}^\beta\text{-6}$). MS (ESI): m/z Calcd for $\text{C}_{34}\text{H}_{36}\text{O}_6\text{Na}$ $[\text{M}+\text{Na}]^+$: 563.2: Found 563.3.

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.²⁹

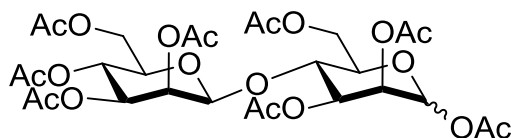
2-O-Benzyl-3-O-p-methoxybenzyl-4,6-O-benzylidene- β -D-mannopyranosyl-(1 $\rightarrow$ 4)-benzyl-2,3,6,tri-O-benzyl- β -D-glucopyranoside (5.4)



To **5.3** (500 mg, 0.88 mmol), 1-benzenesulfinyl piperidine (200 mg, 0.96 mmol) and 2,6-di-tert-butyl-4-methylpyridine (450 mg, 2.2 mmol) was added toluene (7 mL). The solvent was removed under reduced pressure and the solids dried by high vacuum. To the mixture was added 4 Å molecular sieves (300 mg) and dry CH_2Cl_2 (7 mL). After stirring at room temperature for 30 min, the solution was cooled to -78°C and trifluoromethanesulfonic anhydride (0.15 mL, 0.92 mmol) was added. The solution was stirred at -78°C for 20 min and then the chemical glycan donor (475 mg, 0.88 mmol) was added in dry CH_2Cl_2 (3 mL). The reaction was stirred at -78°C for a further 1 h before being warmed -40°C . After stirring at this temperature for 4 h, the reaction was allowed to slowly warm to room temperature overnight. The reaction was quenched by the addition of triethylamine (0.6 mL, 4.4 mmol) and then filtered through a pad of Celite. The mixture was diluted with CH_2Cl_2 (50 mL) and washed with saturated sodium bicarbonate (30 mL) and brine (30 mL) and the organic layer was dried with

sodium sulfate. The solvent was removed under reduced pressure and purification by flash chromatography (8:1, toluene:ethyl acetate) afforded the protected disaccharide (**5.4**, 146 mg, 0.15 mmol, 17%) as a colourless oil exclusively as the β anomer R_f 0.38 (8:1, toluene:ethyl acetate). $[\alpha]_D -4.7$ (c 2.0, CHCl_3). ν_{max} (film)/ cm^{-1} 3063, 3031, 2868, 1721, 1610, 1513, 1497, 1454, 1363, 1304, 1249, 1210, 1173, 1094, 1052, 1028, 914, 821, 737 and 698. ^1H NMR (400 MHz, CDCl_3) δ 7.47-7.05 (32 H, m, $32 \times \text{Ph}$), 6.76 (2 H, d, $J = 8.5$, $2 \times \text{Ph}$), 5.44 (1 H, s, CHPh), 5.17 (1 H, s, H-1), 4.97 (1 H, t, $J = 9.9$, $1 \times \text{CH}_2\text{Ph}$), 4.87 (1 H, t, $J = 9.9$, $1 \times \text{CH}_2\text{Ph}$), 4.79 – 4.58 (6 H, m, $6 \times \text{CH}_2\text{Ph}$), 4.56 (1 H, s, H-1'), 4.51 (1 H, dd, $J = 13.1$, 4.3, $1 \times \text{CH}_2\text{Ph}$), 4.48 – 4.40 (3 H, m, $3 \times \text{CH}_2\text{Ph}$), 4.32 (1 H, d, $J = 11.8$, H-3), 4.19 (1 H, d, $J = 12.0$, $1 \times \text{H-6'}$), 4.03 (3 H, m, H-2, H-4 and H-4'), 3.69 (3 H, s, OMe), 3.64 – 3.52 (2 H, m, H-2' and $1 \times \text{H-6'}$), 3.51 (2 H, m, $2 \times \text{H-6}$), 3.38 – 3.29 (1 H, m, H-5), 3.28 – 3.23 (1 H, m, H-3) and 3.05 (1 H, m, H-5'). ^{13}C NMR (100 MHz, CDCl_3) δ 138.67, 138.65, 137.98, 137.91, 137.84, 137.7, 137.5 ($7 \times \text{Ph}$), 129.5 – 125.4 ($34 \times \text{Ph}$), 113.8 (Ph), 102.6 (C-1), 101.7 (CHPh), 101.4 (C-1'), 83.1 (C-2), 81.8 (C-3), 80.3 (C-4), 79.1 (C-5), 78.8 (C-4'), 78.1 (C-3'), 77.0 (C-2'), 75.1, 75.0, 73.6, 73.4, 72.4, 71.1 ($6 \times \text{CH}_2\text{Ph}$), 69.3 (C-6), 68.7 (C-6'), 67.4 (C-5') and 55.3 (OMe). MS (ESI): m/z Calcd for $\text{C}_{62}\text{H}_{64}\text{O}_{12}\text{Na}$ $[\text{M}+\text{Na}]^+$: 1023.4290: Found: 1023.4264.

2,3,4,6-tetra-O-acetyl- β -D-mannopyranosyl-(1 $\rightarrow$ 4)-1,2,3,6-tetra-O-acetyl-D-mannopyranose (5.5**)¹⁶**



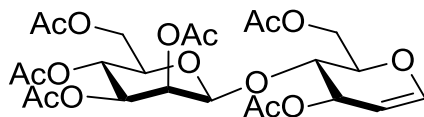
To Locust bean gum (2 g) in 50 mM sodium acetate buffer (pH 5, 50 mL) was added Pectinase (*Aspergillus aculeatus*, 3800 U/mL, 0.2 mL) and this mixture was stirred at

50 °C for 48 h. After this time, the reaction was quenched by refluxing for 15 min. Ethyl acetate (70 mL) was added and the reaction was then filtered. After the addition of toluene (70 mL), the solvent was removed under reduced pressure. The resulting residue was cooled to 0 °C and pyridine (20 mL) and acetic anhydride (20 mL) were added. This was stirred at room temperature for 16 h before the reaction was quenched by pouring into ice water (100 mL). The solution was extracted with CHCl₃ (2 × 100 mL) and the organic layer washed with water (50 mL), 1 M HCl (50 mL), saturated sodium bicarbonate (50 mL) and brine (50 mL). The solvent was removed under reduced pressure and purification by flash chromatography (1:1, pet. ether:ethyl acetate) afforded the protected disaccharide (**5.5**, 1 g, 1.5 mmol, 31%) as a light yellow oil and a 65:35 mix α : β anomers by NMR. *R_f* 0.20 (1:1, pet. ether:ethyl acetate). ¹H NMR (400 MHz, CDCl₃) δ 5.96 (1 H ^{α} , d, *J* = 1.8, H-1 ^{α}), 5.75 (1 H ^{β} , s, H-1 ^{β}), 5.39 – 5.36 (1 H, m, H-2'), 5.34 (1 H ^{α} , d, *J* = 3.4, H-3 ^{α}), 5.32 (1 H ^{β} , d, *J* = 3.7, H-2 ^{β}), 5.18 – 5.16 (1 H, m, H-3 ^{β}), 5.15 – 5.12 (1 H + 1 H ^{α} , m, H-4' and H-2 ^{α}), 5.00 – 4.94 (1 H, m, H-3), 4.69 (1 H ^{α} , s, H-1' ^{α}), 4.67 (1 H ^{β} , s, H-1' ^{β}), 4.30 – 4.15 (3 H, m, 2 × H-6 and 1 × H-6'), 4.02 – 3.87 (2 H + 1 H ^{α} , m, H-4, 1 × H-6' and H-5 ^{α}), 3.73 – 3.68 (1 H ^{β} , m, H-5 ^{β}), 3.61 – 3.55 (1 H, m, H-5') and 2.13 – 1.91 (24 H, m, 8 × OAc). ¹³C NMR (100 MHz, CDCl₃) δ 170.6, 170.5, 170.4, 170.3, 170.1, 170.0, 169.6, 169.5, 168.3 (8 × CO), 98.1 (C-1 ^{α}), 97.7 (C-1 ^{β}), 90.4 (C-1'), 73.7 (C-5 ^{β}), 72.5 (C-4 ^{α} and C-5 ^{α}), 72.3 (C-4 ^{β}), 70.7 (C-5' ^{α}), 70.6 (C-3' ^{α} and C-5' ^{β}), 70.3 (C-3' ^{β}), 69.0 (C-3 ^{β}), 68.6, 68.4, 68.3 (C-2, C-2' and C-3 ^{α}), 65.8 (C-4'), 62.6, 62.4, 62.3, 60.4 (C-6 and C-6') and 20.9 (8 × OAc). MS (ESI): *m/z* Calcd for C₂₈H₃₈O₁₉Na [M+Na]⁺: 701.2: Found 701.2.

The observed characterisation data (¹H and ¹³C) was consistent with that previously

reported in the literature.¹⁶

2,3,4,6-*tert*-O-acetyl- β -D-mannosyl-(1 $\rightarrow$ 4)-3,6-di-O-acetyl-D-glucal (5.6)¹⁶

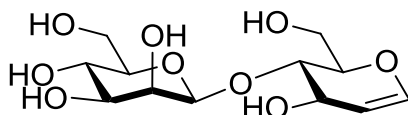


To a solution of **5.5** (500 mg, 0.75 mmol) in dry CH_2Cl_2 (3 mL) was added HBr (0.34 mL, 33% in acetic acid, 1.88 mmol). After stirring at room temperature for 2.5 h, toluene (5 mL) was added and the solvent was removed under reduced pressure. Ethyl acetate (0.3 mL), acetic acid (3 mL), sodium acetate (615 mg, 7.5 mmol), zinc (492 mg, 7.5 mmol) and $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ (47 mg, 0.19 mmol) were all added to the reaction, which was stirred at room temperature for 2 h. The reaction was then filtered through Celite and ethyl acetate (30 mL) was added. The mixture was washed with water (30 mL), saturate sodium bicarbonate (30 mL) and brine (30 mL) and the solvent was removed under reduced pressure. Purification by flash chromatography (1:1, pet. ether:ethyl acetate) afforded the protected disaccharide glycal (**5.6**, 115 mg, 0.2 mmol, 27%) as white solid. *R_f* 0.16 (6:4, pet. ether:ethyl acetate). ^1H NMR (400 MHz, CDCl_3) δ 6.34 (1 H, d, $J = 6.1$, H-1), 5.44 (1 H, dd, $J = 5.0, 3.1$, H-3), 5.38 (1 H, d, $J = 3.1$, H-2'), 5.16 (1 H, t, $J = 9.9$, H-4'), 4.99 (1 H, dd, $J = 9.9, 3.2$, H-3'), 4.81 (1 H, s, H-1'), 4.72 (1 H, dd, $J = 6.1, 3.1$, H-2), 4.35 (1 H, dd, $J = 12.0, 2.6$, 1 $\times$ H-6), 4.23 (1 H, dd, $J = 12.2, 5.8$, 1 $\times$ H-6'), 4.16 (1 H, dd, $J = 12.2, 5.3$, 1 $\times$ H-6), 4.11 – 3.96 (3 H, m, H-4, H-5, 1 $\times$ H-6'), 3.62 – 3.57 (1 H, m, H-5'), 2.11 (3 H, s, OAc), 2.05 (3 H, s, OAc), 2.03 (3 H, s, OAc), 2.02 (3 H, s, OAc), 1.98 (3 H, s, OAc) and 1.93 (3 H, s, OAc). ^{13}C NMR (100 MHz, CDCl_3) δ 170.7 (CO), 170.4 (CO), 170.3 (CO), 170.1 (CO), 170.0 (CO), 169.6 (CO),

145.6 (C-1), 99.0 (C-2), 97.9 (C-1'), 74.4 (C-5), 74.0 (C-4), 72.6 (C-5'), 70.7 (C-3'), 68.60 (C-3), 68.55 (C-2'), 65.8 (C-4'), 62.5 (C-6), 61.8 (C-6'), 21.1 (OAc), 20.8 (OAc), 20.71 (OAc). 20.68 (2 × OAc) and 20.56 (OAc). MS (ESI): m/z Calcd for $C_{24}H_{34}N_2O_6Na$ $[M+Na]^+$: 583.2: Found 583.2. MP = 39 – 41 °C.

The observed characterisation data (^1H and ^{13}C) was consistent with that previously reported in the literature.¹⁶

β -D-mannosyl-(1 $\rightarrow$ 4)-D-glucal (5.7)



To a solution of **5.6** (115 mg, 0.2 mmol) in methanol (5 mL) was added sodium methoxide (10 mg, 0.2 mmol) and the reaction was stirred at room temperature for 16 h. The solvent was removed under reduced pressure to afford the disaccharide glycal (**2**, 44 mg, 0.14 mmol, 70%) as a white solid that was used without further purification. $[\alpha]_D +0.24$ (c 1.0, $CHCl_3$). ν_{max} (film)/ cm^{-1} 3332 (O-H), 2928, 2360, 1651, 1594, 1379, 1351, 1237, 1158, 1070, 1031, 866, 823, 795, 751, 719 and 668. ^1H NMR (400 MHz, MeOD) δ 6.29 (1 H, d, J = 6.1, H-1), 4.65 (1 H, s, H-1'), 4.65 (1 H, dd, J = 6.0, 2.0), 4.25 (1 H, d, J = 6.7, H-3), 3.86 (1 H, d, J = 2.8, H-4), 3.83, (1 H, d, J = 2.1, H-2'), 3.80 (1 H, d, J = 2.0, H-4'), 3.76 – 3.70 (3 H, m, H-3', and 2 × H-6), 3.60 (1 H, dd, J = 11.8, 6.6, 1 × H-6'), 3.44 (1 H, d, J = 9.5, 1 × H-6'), 3.39 (1 H, dd, J = 9.3, 2.9, H-5) and 3.23 – 3.19 (1 H, m, H-5'). ^{13}C NMR (100 MHz, $CDCl_3$) δ 143.9 (C-1), 102.3 (C-2), 100.3 (C-1'), 77.5 (C-5'), 77.3 (C-4), 77.1 (C-5), 73.7 (C-2'), 71.0 (C-3), 68.1 (C-4'), 67.1 (C-3'), 61.2 (C-6) and 59.9 (C-6'). HRMS (ESI): m/z Calcd for $C_{12}H_{20}O_9Na$ $[M+Na]^+$:

331.1000: Found 331.0991. MP = 183 – 186 °C

5.8 – Chemo-Enzymatic Experimental

5.8.1 – Investigating Endo Glycosylation Reactions Using a Glycal Donor

To the GlcNAc-triazole-amino acid (**3.31**, 19 μ L 5.4 mM, 0.1 μ mol) and the disaccharide glycal (**5.7**, 100 μ L, 8 mM, 0.8 μ mol) in 50 mM sodium phosphate (pH 6, 380 μ L) was added either GST-EndoA N171D or N171E (1.5 μ L 3.6 mg/mL or 36 mg/mL, 5.4 μ g or 54 μ g) to give an approximate 1:2000 or 1:200, acceptor:EndoA molar ratio. More concentrated reactions were run with less 50 mM sodium phosphate (pH 6, 150 μ L) and more enzyme (15 μ L, 36 mg/mL, 0.5 mg, approximate 1:20 acceptor:Endo molar ratio). Reactions were shaken at room temperature. Samples (20 μ L) were taken at 30 min, 1 h, 2 h, 4 h and 16 h, quenched (1 μ L, 1% formic acid) and analysed by HPLC (Dionex Ultimate 3000 pump and variable wavelength detector with a Synergi 4U Hydro-RP 80 Å 100 $\times$ 4.6 mm 4 micron column, detecting at 280 nm) and MS.

5.8.2- Checking EndoA Mutants Activities and the Effect of the Glycal on Activity

To RNase B (5 μ L, 10 mg/ mL, 50 μ g) in 50 mM sodium phosphate (pH 6, 90 μ L) was added wild type GST-EndoA, GST-EndoA N171D or GST-EndoA N171E (5 μ L, 0.1 mg/mL, 0.5 μ g) and either additional 50 mM sodium phosphate (pH 6, 1 μ L) or the disaccharide glycal (**5.7**, 1 μ L, 2.6 mg/mL, 2.6 μ g, 8.4 nmol). Reactions were incubated at room temperature. Samples (10 μ L) were taken at 10 min, 30 min, 1 h and 2 h, quenched (1 μ L 1% formic acid) and analysed by SDS-PAGE.

5.8.3 – Investigating Endo Catalysed Hydrolysis of the Glycal

To the disaccharide glycal (**5.7**, 2 mg, 6.4 μmol) in 50 mM sodium phosphate (pH 6, 200 μL) was added either GST-EndoA N171D, GST-Endo N171E (100 μL , 10 mg/mL, 1 mg), GST-EndoS or GST-EndoS D233E (100 μL , 16 mg/mL, 1.6 mg) to give an approximate 1:500, glycal:enzyme molar ratio. Reactions were incubated at room temperature for 48 h. After this time the reactions were diluted (300 μL water) and the enzymes were removed by Vivaspin (MWCO 30000), with the flowthrough being collected. This process was repeated a further two times. The collected flowthrough was combined and the lyophilized to remove the solvent. The resulting residue was dissolved in deuterated methanol and analysed by NMR.

5.9 – References

- 1 S. Bay and D. Cantacuzene, *Bioorg. Med. Chem. Lett.*, 1992, **2**, 423-426.
- 2 H. Prade, L. F. Mackenzie and S. G. Withers, *Carbohydr. Res.*, 1998, **305**, 371-381.
- 3 J. Petit, F. Paquet and J. Beau, *Tet. Lett.*, 1991, **32**, 6125-6128.
- 4 A. Trincone, E. Pagnotta, M. Rossi, M. Mazzone and M. Moracci, *Tetrahedron: Asymmetry*, 2001, **12**, 2783-2787.
- 5 L. Kroger and J. Thiem, *Carbohydr. Res.*, 2014, **398**, 8-12.
- 6 E. C. K. Lai, S. A. Morris, I. P. Street and S. G. Withers, *Bioorg. Med. Chem.*, 1996, **4**, 1929-1937.
- 7 J. B. Kempton and S. G. Withers, *Biochemistry*, 1992, **31**, 9961-9969.
- 8 M. Fernandez-Gonzalez, O. Boutureira, G. J. L. Bernardes, J. M. Chalker, M. A. Young, J. C. Errey and B. G. Davis, *Chem. Sci.*, 2010, **1**, 709-715.
- 9 J. J. Goodfellow, K. Baruah, K. Yamamoto, C. Bonomelli, B. Krishna, D. J. Harvey, M. Crispin, C. N. Scanlan and B. G. Davis, *J. Am. Chem. Soc.*, 2012, **134**, 8030-8033.
- 10 J. U. Rhee, B. I. Bliss and T. V. Rajanbabu, *J. Am. Chem. Soc.*, 2003, **125**, 1492-1493.
- 11 N. K. Jalsa, *Tet. Lett.*, 2011, **52**, 6587-6590.
- 12 A. H. Daranas, H. Shimizu and S. W. Homans, *J. Am. Chem. Soc.*, 2004, **126**, 11870-11876.
- 13 S. Bera and R. J. Linhardt, *J. Org. Chem.*, 2011, **76**, 3181-3193.
- 14 G. Zou, H. Ochiai, W. Huang, Q. Yang, C. Li and L. Wang, *J. Am. Chem. Soc.*, 2011, **133**, 18975-18991.
- 15 B. Li, Y. Zeng, S. Hauser, H. Song and L. Wang, *J. Am. Chem. Soc.*, 2005, **127**, 9692-9693.

- 16 R. H. V. Kumar, K. Naruchi, R. Miyoshi, H. Hinou and S. Nishimura, *Org. Lett.*, 2013, **15**, 6278-6281.
- 17 J. Zhang, L. Han, S. Yan and C. Liu, *BMC Microbiology*, 2014, **14**, 95.
- 18 G. Legler, K. Roeser and H. Illig, *Eur. J. Biochem.*, 1979, **101**, 85-92.
- 19 M. J. Rogers and K. G. Brandt, *Biochemistry*, 1971, **10**, 4624-4630.
- 20 M. Collin and A. Olsen, *EMBO J.*, 2001, **20**, 3046-3055.
- 21 W. Huang, J. Giddens, S. Fan, C. Toonstra and L. Wang, *J. Am. Chem. Soc*, 2012, **134**, 12308-12318.
- 22 B. G. Davis and L. Wang, *Chem. Sci.*, 2013, **4**, 3381-3394.
- 23 J. Yin, L. Li, N. Shaw, Y. Li, J. K. Song, W. Zhang, C. Xia, R. Zhang, A. Joachimiak, H. Zhang, L. Wang, Z. Liu and P. Wang, *Plos One*, 2009, **4**, e4658.
- 24 D. J. Vocadlo, G. J. Davies, R. Laine and S. G. Withers, *Nature*, 2001, **412**, 835-838.
- 25 T. M. Gloster, S. Roberts, V. M. Ducros, G. Perugino, M. Rossi, R. Hoos, M. Moracci, A. Vasella and G. J. Davies, *Biochemistry*, 2004, **43**, 6101-6109.
- 26 P. A. Dalby, *Curr. Opin. Struct. Biol*, 2011, **21**, 473-480.
- 27 S. Lutz, *Curr. Opin. Biotechnol*, 2010, **21**, 734-743.
- 28 M. O. Lederer, C. M. Dreisbusch and R. M. Bundschuh, *Carbohydr. Res.*, 1997, **301**, 111-121.
- 29 V. Balakumar, A. Aravind and S. Baskaran, *Synlett*, 2004, **4**, 647-650.

You can't always get what you want

But if you try sometime you find

You get what you need

The Rolling Stones

You Can't Always Get What You Want (1969)

NATURE PUBLISHING GROUP LICENSE TERMS AND CONDITIONS

Aug 28, 2016

This Agreement between Thomas Taylor ("You") and Nature Publishing Group ("Nature Publishing Group") consists of your license details and the terms and conditions provided by Nature Publishing Group and Copyright Clearance Center.

License Number	3916681421214
License date	Jul 26, 2016
Licensed Content Publisher	Nature Publishing Group
Licensed Content Publication	Nature
Licensed Content Title	X-ray structure of a bacterial oligosaccharyltransferase
Licensed Content Author	Christian Lizak, Sabina Gerber, Shin Numao, Markus Aebi, Kaspar P. Locher
Licensed Content Date	Jun 15, 2011
Licensed Content Volume Number	474
Licensed Content Issue Number	7351
Type of Use	reuse in a dissertation / thesis
Requestor type	non-commercial (non-profit)
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
High-res required	no
Figures	Figure 3 and Figure 5 from the article
Author of this NPG article	no
Your reference number	
Title of your thesis / dissertation	Investigations into Novel Enzymatic Glycosylation Methods
Expected completion date	Jul 2016
Estimated size (number of pages)	320
Requestor Location	Thomas Taylor Oxfordshire
Billing Type	Invoice
Billing Address	Thomas Taylor Oxfordshire Attn: Thomas Taylor

Total

0.00 GBP

Terms and Conditions

Terms and Conditions for Permissions

Nature Publishing Group hereby grants you a non-exclusive license to reproduce this material for this purpose, and for no other use, subject to the conditions below:

1. NPG warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to Nature Publishing Group and does not carry the copyright of another entity (as credited in the published version). If the credit line on any part of the material you have requested indicates that it was reprinted or adapted by NPG with permission from another source, then you should also seek permission from that source to reuse the material.
2. Permission granted free of charge for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to the work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version. Where print permission has been granted for a fee, separate permission must be obtained for any additional, electronic re-use (unless, as in the case of a full paper, this has already been accounted for during your initial request in the calculation of a print run). NB: In all cases, web-based use of full-text articles must be authorized separately through the 'Use on a Web Site' option when requesting permission.
3. Permission granted for a first edition does not apply to second and subsequent editions and for editions in other languages (except for signatories to the STM Permissions Guidelines, or where the first edition permission was granted for free).
4. Nature Publishing Group's permission must be acknowledged next to the figure, table or abstract in print. In electronic form, this acknowledgement must be visible at the same time as the figure/table/abstract, and must be hyperlinked to the journal's homepage.
5. The credit line should read:
Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)
For AOP papers, the credit line should read:
Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

Note: For republication from the *British Journal of Cancer*, the following credit lines apply.

Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication) For AOP papers, the credit line should read:

Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

6. Adaptations of single figures do not require NPG approval. However, the adaptation should be credited as follows:

Adapted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)

Note: For adaptation from the *British Journal of Cancer*, the following credit

line applies.

Adapted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK:
[JOURNAL NAME] (reference citation), copyright (year of publication)

7. Translations of 401 words up to a whole article require NPG approval. Please visit <http://www.macmillanmedicalcommunications.com> for more information. Translations of up to a 400 words do not require NPG approval. The translation should be credited as follows:

Translated by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication).

Note: For translation from the *British Journal of Cancer*, the following credit line applies.

Translated by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

We are certain that all parties will benefit from this agreement and wish you the best in the use of this material. Thank you.

Special Terms:

v1.1

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

JOHN WILEY AND SONS LICENSE TERMS AND CONDITIONS

Aug 28, 2016

This Agreement between Thomas Taylor ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	3916680771668
License date	Jul 26, 2016
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Angewandte Chemie
Licensed Content Title	Optimal Synthetic Glycosylation of a Therapeutic Antibody
Licensed Content Author	Thomas B. Parsons, Weston B. Struwe, Joseph Gault, Keisuke Yamamoto, Thomas A. Taylor, Ritu Raj, Kim Wals, Shabaz Mohammed, Carol V. Robinson, Justin L. P. Benesch, Benjamin G. Davis
Licensed Content Date	Jan 12, 2016
Licensed Content Pages	7
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	4
Original Wiley figure/table number(s)	Figures s1a, s2a, s3a and s4a from supporting information
Will you be translating?	No
Title of your thesis / dissertation	Investigations into Novel Enzymatic Glycosylation Methods
Expected completion date	Jul 2016
Expected size (number of pages)	320
Requestor Location	Thomas Taylor Oxfordshire Attn: Thomas Taylor
Publisher Tax ID	EU826007151
Billing Type	Invoice
Billing Address	Thomas Taylor

Oxfordshire

Attn: Thomas Taylor

Total

0.00 GBP

[Terms and Conditions](#)

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your RightsLink account (these are available at any time at <http://myaccount.copyright.com>).

Terms and Conditions

- The materials you have requested permission to reproduce or reuse (the "Wiley Materials") are protected by copyright.
- You are hereby granted a personal, non-exclusive, non-sub licensable (on a stand-alone basis), non-transferable, worldwide, limited license to reproduce the Wiley Materials for the purpose specified in the licensing process. This license, **and any CONTENT (PDF or image file) purchased as part of your order**, is for a one-time use only and limited to any maximum distribution number specified in the license. The first instance of republication or reuse granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before the end date may be distributed thereafter). The Wiley Materials shall not be used in any other manner or for any other purpose, beyond what is granted in the license. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Wiley Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Wiley Material. Any third party content is expressly excluded from this permission.
- With respect to the Wiley Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Wiley Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Wiley Materials without the prior permission of the respective copyright owner. **For STM Signatory Publishers clearing permission under the terms of the [STM Permissions Guidelines](#) only, the terms of the license are extended to include subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and does not involve the separate exploitation of the permitted figures or extracts.** You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Wiley

Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.

- The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto
- NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.
- WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.
- You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.
- IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.
- Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

- The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.
- This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.
- Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.
- These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.
- In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.
- WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
- This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.
- This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

WILEY OPEN ACCESS TERMS AND CONDITIONS

Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

The Creative Commons Attribution License

The [Creative Commons Attribution License \(CC-BY\)](#) allows users to copy, distribute and

transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-

Creative Commons Attribution Non-Commercial License

The [Creative Commons Attribution Non-Commercial \(CC-BY-NC\) License](#) permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.(see below)

Creative Commons Attribution-Non-Commercial-NoDerivs License

The [Creative Commons Attribution Non-Commercial-NoDerivs License](#) (CC-BY-NC-ND) permits use, distribution and reproduction in any medium, provided the original work is properly cited, is not used for commercial purposes and no modifications or adaptations are made. (see below)

Use by commercial "for-profit" organizations

Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee.

Further details can be found on Wiley Online

Library<http://olabout.wiley.com/WileyCDA/Section/id-410895.html>

Other Terms and Conditions:

v1.10 Last updated September 2015

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.