SYNTHESIS OF TOOLS FOR GLYCOPROTEIN REMODELLING

Jinit Masania

University College London

PhD (Chemistry)

Supervisor: Dr Derek Macmillan

2010

Acknowledgements

I would like to express my thanks to Dr Derek Macmillan who has supervised my research and encouraged me throughout my time at UCL. I would also like to thank Dr Jonathan Richardson who gave me a lot of advice and was very supportive throughout my research and all the past and present members of the Macmillan lab. I would like to thank my parents for their support and to dedicate the thesis to my nephew Vivek, who was born during this period of research. I also thank BBSRC who funded this research.

Declaration

‘ I, Jinit Masania confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.'

Abstract

The production of therapeutic glycoproteins typically relies on the purification of the desired glycoprotein from engineered tissue culture systems. This process is costly and inefficient, and the isolated glycoproteins are adorned with an array of heterogeneous sugars. Expression of glycoproteins in cheaper non-mammalian host cells such as yeast produces large quantities of folded glycoproteins though the sugar chains are immunogenic to humans. This project initially aimed to develop a general strategy for the humanisation of yeast glycoproteins using the copper (I) mediated ligation between azide and acetylene functional groups, commonly known as “click” chemistry. It required that azide and acetylene groups could be efficiently incorporated into (glyco)peptides and sugars respectively, and to this end propargylic glycosides of mannose, *N*-acetyl glucosamine, and *N*-acetyl lactosamine were successfully prepared.

Two different glycoprotein remodelling methods were ultimately investigated. The first method explored a chemoenzymatic strategy developed in our group. A target glycoprotein, erythropoietin (EPO), was expressed in the presence of azide bearing mannose sugars in the hope that azide groups would be incorporated during glycoprotein biosynthesis. Subsequent elaboration of the azide functionality with propargyl *N*-acetyl lactosamine followed by enzymatic sialylation was expected to yield a glycoprotein with human-like sialyl lactosamine antennae. However, upon purification from *Pichia pastoris* incorporation of azide could not be detected, limiting the strategy to *in vitro* glycoprotein remodelling. Several technological advances were made including optimisation of the final “click” reaction between the propargyl glycoside of *N*-acetyllactosamine and 4-azidomannose followed by enzymatic transfer of sialic acid. While considerable effort was directed towards the key transfer of unnatural azidosugars, from nucleotide donors to potential substrates using an α-1,2-mannosyltransferase, this step proved unreliable.

The second method introduced azide groups into fully synthetic peptides using the amino acid azidohomoalanine (Aha). Click chemistry with synthetic propargyl glycosides allowed further modification to homogeneous glycopeptide analogues which were shown to be compatible with native chemical ligation, a proven tool for glycoprotein synthesis and semi-synthesis. Ultimately both methods may eventually facilitate glycoprotein synthesis and remodelling such that the biological activity and immunogenicity may be modulated to suit future therapeutic requirements.

Table of Contents

Page number

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of tables

List of figures

Figure 5.18 *Regenerating the nucleotide sugar GDP Mannose for mannose transfer.*

Figure 5.19 *Modified CD52 peptide **167** as a potential substrate for endo-NAGs.*

Abbreviations

6AzFuc = 6-azido fucose
ADG = Alpha-dystroglycan
Alg = Asparagine-linked glycosylation
Aha = Azidohomoalanine
β3GnT1 = β-3-*N*-acetylglucosaminyltransferase
BHK = Baby hamster kidney
Boc = Tertiary-butyloxycarbonyl Group
CAN = Cerium ammonium nitrate
CFE = Cell-free extract
CHO = Chinese hamster ovary cell
C. jejuni = *Campylobacter jejuni*
CMD = Congenital muscular dystrophy
CMP = Cytidine 5'-monophosphate
CuAAC = copper (I) catalysed azide-alkyne 1,3-dipolar cycloaddition
DBU = Diaza(1, 3) bicyclo[5.4.0]undecane
DCM = Dichloromethane
DIFO = Difluorinated cyclooctyne
DMF = Dimethylformamide
Dol-P-Man = Dolichol-phosphate-mannose synthase
Dol-P = Dolichyl-phosphate
ECM = Extracellular matrix
E.coli = *Escherichia coli*
EGF = Epidermal growth factor
Endo-NAG = Endo-β-*N*-acetylglucosaminidases
EPO = Erythropoietin
ER = Endoplasmic reticulum
FKRP = Fukutin related protein
Fmoc = Fluorenylmethyloxycarbonyl
GAGs = Glycosaminoglycans
GalNAc = *N*-acetylgalactosamine
GalNAz = *N*-azidoacetylgalactosamine
GDP = Guanosine 5'-diphosphate
GlcNAc = *N*-acetylglucosamine
GlcNAz = *N*-azidoacetylglucosamine

GPI = Glycosylphosphatidylinositol

GTs = Glycosyltransferases

HEK 293 = Human embryonic kidney cell line 293

HMPA = Hexamethylphosphoramide

HRP = Horseradish peroxidase

IDCP = Iodonium dicollidine perchlorate

IPTG = Isopropyl *β*-*D*-1-thiogalactopyranoside

LacNAc = *N*-Acetyl-*D*-lactosamine

LARGE = Like-glycosyltransferase

LC-MS = Liquid chromatography mass spectrometry

ManNAz = *N*-azidoacetylmannosamine

Man T = Mannosyltransferase

MESNa = sodium 2-mercaptoethane sulfonate

NCL = Native chemical ligation

NBS = *N*-bromosuccinimide

NDP = Nucleotide diphosphate

NIS = *N*-iodosuccinimide

Neu5Ac = *N*-acetylneuraminic acid

OmpA = outer membrane protein A

PBS = Phosphate buffered saline

PCC = Pyridinium chlorochromate

PDI = Protein disulphide isomerise

P. pastoris = *Picha pastoris*

Propargyl LacNAc = Propargyl *N*-acetyllactosamine

POFUT1 = Protein *O*-fucosyltransferase 1

POMGnT1 = Protein *O*-linked mannose beta1,2-*N*-acetylglucosaminyltransferase

POMT 1 = Protein *O*-Mannosyltransferase-1

POMT2 = Protein *O*-Mannosyltransferase-2

RuAAC = Ruthenium catalysed 1,3-dipolar azide-alkyne cycloaddition

S. *cerevisiae* = *Saccharomyces cerevisiae*

SiaNAz = *N*-azidoacetyl sialic acid

SIBLINGs = Small integrin-binding ligand *N*-linked glycoproteins

TBAF = Tetrabutylammonium fluoride

TBDPSCl = Tertiary-butyldiphenylsilyl chloride

TBTA = Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl] amine

TfOH = Triflic acid

TLC = Thin layer chromatography

TMR = Tetramethylrhodamine

TMSC = Trimethylsilyl chloride

UDP = Uridine 5'-diphosphate

Chapter 1

Introduction

For many years glycoproteins have been a subject of interest for biochemists and biologists, partly due to the fact that glycoproteins were discovered to be abundant in living organisms and by the diverse functions of glycoproteins; glycoproteins appear in many biological process [1]. Glycoproteins are proteins that contain carbohydrates which are covalently attached to a polypeptide side-chain. The carbohydrate may be in the form of monosaccharides, disaccharides, oligosaccharides, or polysaccharides, and are sometimes referred to as glycans. The carbohydrates are attached to the protein in a co-translational or post-translational modification. This process is known as glycosylation [1 - 6].

Many glycoproteins have different structural functions [1, 2]. For instance, in proteins that have segments extending extracellularly, this segment is often glycosylated and is known to play a role in cell-cell interaction. Glycoproteins are also found in gastrointestinal mucus secretions where they are used as protective agents and lubricants [3]. They also form connective tissues such as collagen and are abundantly found in the blood plasma where they serve many functions [4].

The carbohydrate content in glycoproteins can vary markedly; collagens contain 1% carbohydrate content, while certain mucins have 60% carbohydrate content [3, 5]. 50% of human proteins are glycosylated [7] therefore it is not surprising to envisage the importance of glycoproteins in the regulation of the human body. A few examples of these glycoproteins include antibodies (immunoglobulins) that interact directly with antigens, major histocompatibility complex molecules that interact with the T-cells as part of an adaptive immune response [8] and components of the zona pellucida which are important for sperm-egg interaction [9]. The prevalence of glycoproteins in disease processes can be harnessed to address diagnostic requirements. For instance the presence of small integrin-binding ligand *N*-linked glycoproteins (SIBLINGs) is now a recognised biomarker used to detect malignant transformation and cancer progression [10]. Glycoproteins are also important in blood to carry blood group determinants. The carbohydrate portion of this glycoprotein is usually a small sugar component made up of individual monosaccharide units. This carbohydrate component may be a combination of up to seven of the many naturally occurring sugar molecules in mammals including glucose, glucosamine, galactose, galactosamine, mannose, fucose, and sialic acid [5]. Another important glycoprotein related

to blood is erythropoietin (EPO). Produced by the peritubular capillary endothelial cells in the kidney, EPO is a hormone that regulates red blood cell production [5, 11].

The importance of glycoproteins for the regulation of human body to function properly is evident; therefore glycosylation defects may result to a whole spectrum of diseases. Congenital muscular dystrophy (CMD) is a general muscle weakness disease and is caused by various different genetic errors [12]. Reduced glycosylation of alpha-dystroglycan (ADG) has been found to relate to CMD. ADG is a peripheral membrane protein that undergoes multiple and complex glycosylation steps to regulate its ability to effectively interact with extracellular matrix proteins, such as laminin and agrin [12]. In fact, reduced glycosylation of ADG is now referred to as dystroglycanopathies and mutations in six genes (POMT1, POMT2, POMGnT1, Fukutin, FKRP and LARGE) have been identified in patients with a dystroglycanopathy [12].

A further role for glycosylation of proteins may lie in protecting the associated protein and aid the glycoprotein to function. An example of this is where heavily glycosylated mucins give them considerable water holding capacity and make them resistant to proteolysis, which may be important in maintaining mucosal barriers [3, 5]. Glycosylation in general aids the functionality of the protein and the biological processes that it is involved in.

1.1 Structure and linkages of glycoproteins

The biosynthesis of glycoproteins involves the addition of sugars to dolichol which are then transferred to a protein or by directly transferring sugars to a protein. The covalent bond between carbohydrate and protein is dependent on the amino acid which is glycosylated. Glycoproteins can be classified into a series of glycosylated forms.

***N*-linked** glycoproteins, by far one of the most common forms, is when the saccharide is linked to the nitrogen atom of the side chain amide group of an asparagine (Asn). This residue must be embedded in the consensus sequence Asn-X-Ser/Thr though this sequence, in itself, is not sufficient for glycosylation; [13, 14] where the X represents any amino acid except proline.

***O*-linked** glycoproteins are when the hydroxyl group of a serine or threonine, residue in a proline-rich region are involved in the linkage; [13, 15, 16]. This type of glycosylation is one of the most common forms.

***C*-linked** glycoproteins are defined by a C-C linkage between the carbohydrate moiety and the peptide; examples include tryptophan indole bound to monosaccharides [14, 17] and of the amino terminal isoleucine backbone bound to methyl deoxy-fructose [13, 17]. C-glycosylation is not a typical process therefore is redundant.

***P*-linked** glycoproteins are when either a phosphonamide or a phosphodiester hinge is joined to various amino acids or glycopeptides; [14, 18]. An example where P-glycosylation is found is in the glycosylphosphatidylinositol (GPI) anchor.

***S*-linked** glycoproteins is when the carbohydrate is joined through a thioglycosidic linkage to the side chain of a cysteine (Cys), [14, 19-21] or through a thiazolidine linkage [20, 21].

The most common glycoproteins belong to the first two groups, the biosynthesis of which will be the focus of this chapter.

1.2 Biosynthesis of *N*-linked glycoproteins

N-glycosylation is a co-translational process which occurs in the endoplasmic reticulum (ER) an important organelle in mammalian cells. The oligosaccharide chain starts to assemble by the addition of two *N*-acetylglucosamine (GlcNAc) and five mannose residues, one by one on to a dolichol phosphate lipid anchor at the cytosolic surface of the rough (ER) [22, 23]. Next, the glycolipid flips into the lumenal side of the ER (**Figure 1.0**) and another four mannose and three glucose units are transferred before the polypeptide is conjugated. Three glucose residues and one terminal mannose are trimmed off by three different enzyme α-1,2-glucosidase I, α-1,3-glucosidase II, and α-1,2-mannosidase (**Figure 1.1**) [23]. The removal of the two glucose molecules by the sequential action of glucosidase I and glucosidase II produces a monoglucosylated species that is recognised by the chaperones calnexin and calreticulin [24, 25]. Calnexin and calreticulin are lectins that form complexes with the glycoprotein and ensure correct folding of the glycoprotein [24, 25]. This is controlled upon removal of the last glucose by glucosidase II, and reformed by glucosyltransferases activity. Calnexin and calreticulin each consist of a globular domain, which contains the sugar-binding site, and an extended arm (the P-domain), the tip of which weakly associates with ERp57 (oxidoreductase), a member of the protein disulphide isomerase (PDI) superfamily. Once the glycoprotein has acquired its native conformation, glucosidase II hydrolyses the remaining glucose residue and the glycoprotein releases from the lectin anchor for transport to the Golgi [26].

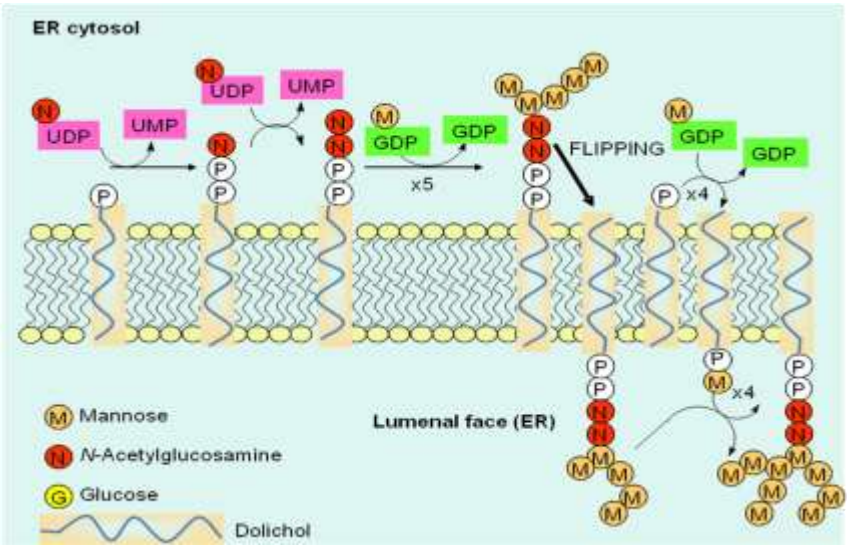


Figure 1.0 *Biosynthesis of N-linked glycoproteins [22, 23].*

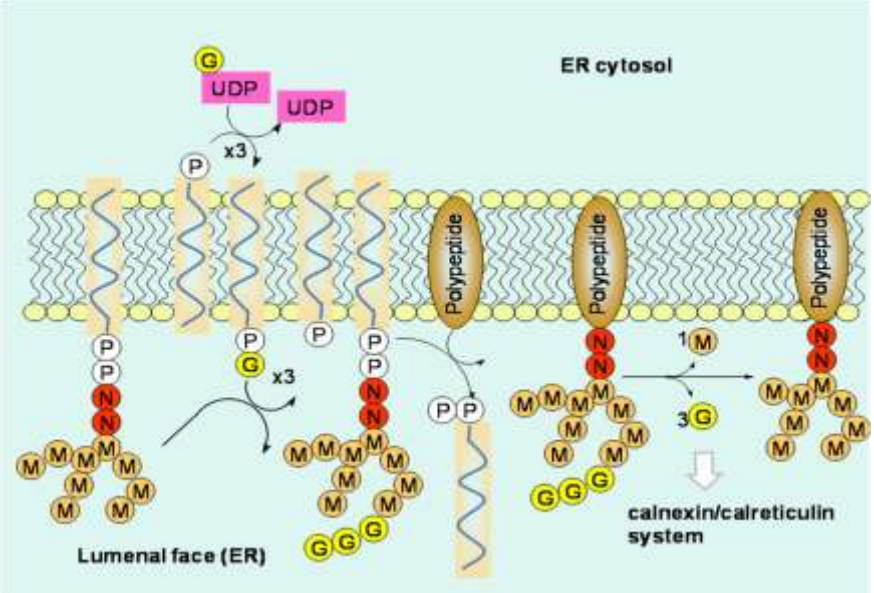


Figure 1.1 *Continued biosynthesis of N-linked glycoproteins: three glucose units are received by the glycolipid before the glycan is conjugated to the polypeptide by oligosaccharyltransferase. The removal of one mannose and two glucose results in a monoglucosylated species that is recognised by the chaperones calnexin and calreticulin. Hydrolysis of the remaining glucose residue releases the glycoprotein from the calnexin and calreticulin cycle and ready for transport to the Golgi [22-26].*

Transport of the glycoproteins *via* membranous vesicles to the Golgi apparatus allows further glycosylation/deglycosylation. A variety of processing enzymes in the cisternae of the Golgi stack are responsible for this process. The Golgi stack can be divided into the *cis*, *medial* and *trans* and each plays an important role in glycoprotein maturation. The *cis* apparatus cleaves three mannose units with mannosidase, and *N*-acetylglucosamine transferase 1 (GlcNAcT I) adds an *N*-acetylglucosamine (GlcNAc) unit. In the *medial* cisternae, two mannose units are removed by mannosidase II, and two GlcNAc are added by GlcNAcT II, and one fucose residue is added by fucosyltransferase [27]. Finally, in the *medial* and *trans* Golgi, other glycosyltransferases add GlcNAc, galactose, and/or sialic acid residues [28]. This results in four groups of heterogeneous glycoprotein, the carbohydrate component is classified as either: high-mannose, complex, hybrid or poly-*N*-acetyllactosamine *N*-linked oligosaccharides (**Figure 1.2**).

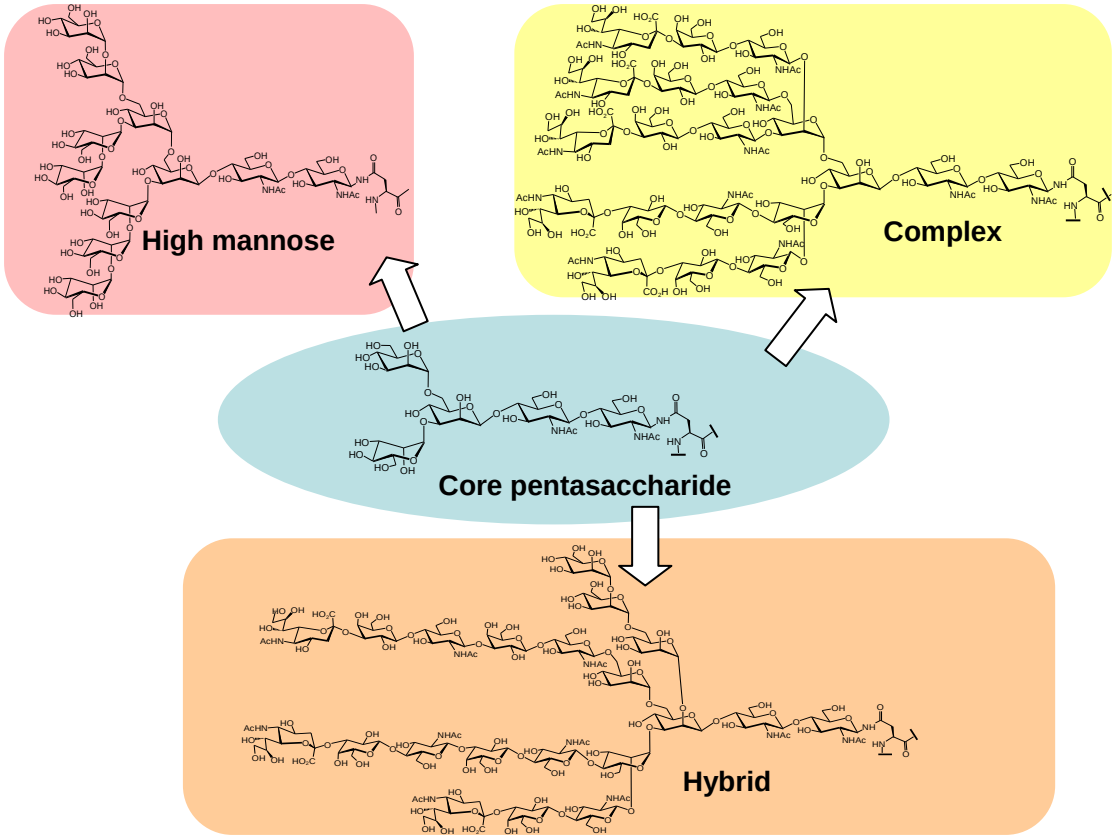


Figure 1.2 *Example of high mannose, complex and hybrid oligosaccharides containing the conserved core pentasaccharide.*

It is important to note that all of the *N*-linked oligosaccharides feature the same common core pentasaccharide [29]. However it is the different arrangement of monosaccharide units on the core pentasaccharide which determines the characteristic of each species [30, 31].

1.3 *O*-linked glycoproteins

O-linked glycoproteins occur when the carbohydrates are attached to a hydroxyl group of serine, threonine, or hydroxylysine (hLys) of a polypeptide backbone [32]. Unlike *N*-linked glycosylation, *O*-linked glycosylation is a 'true post-translational' event in that it does not require an oligosaccharide precursor which is transferred to a protein. The most common type of *O*-linked glycosylation is the attachment of GalNAc to the protein backbone [33]. This is commonly known as mucin-type *O*-glycosylation and generally produces highly heterogeneous glycans, usually shorter in length compared to *N*-linked glycans [33].

The *β*-*O*-GlcNAc modification has been recently shown to be related to phosphorylation states and dynamic processes related to cell signalling events. Furthermore β-*O*-GlcNAc

modification has been linked to several pathological states, including cancer, diabetes and Alzheimer's disease [34, 35].

One notable *O*-linked structure is the Gal-β(1-3)GalNAc sequence that has antigenic properties. Termination of *O*-linked glycans usually includes Gal, GlcNAc, GalNAc, fucose, or sialic acid. By far the most common modification of the core Gal-β(1-3)-GalNAc is mono-, di-, or trisialylation (**Figure 1.3**). A less common, but widely distributed *O*-linked hexasaccharide structure contains β(1-4)-linked Gal and β(1-6)-linked GlcNAc as well as sialic acid [33].

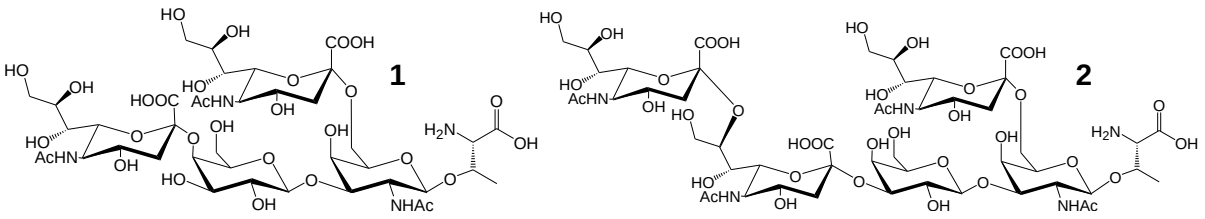


Figure 1.3 *Di- and Trisialylated O-linked Core.*

O-galactosylation is found on hydroxylysine residues of collagens, and the extent of this modification influences the three-dimensional structure of the protein, which is important to exert its biological function in tissue regeneration [36].

Proteins with epidermal growth factor (EGF) modules posses less common *O*-fucosylation and *O*-glucosylation repeats. *O*-fucosylation was first evident through the identification of Glcβ1-3Fucα-Thr, in human urine [33]. More interest grew when *O*-fucosylation was identified in human urokinase, and other clotting proteins such as tissue plasminogen activator and clotting factor VII. Initially fucose was considered to be a terminal sugar, so there was no inclination to search for glycosyltransferase enzyme that use fucose as an acceptor. However, clotting factor VII has a glycan structure Sia-α(2,6)-Gal-β(1,4)-GlcNAc-β-(1,3)Fuc-α-, where the formation of GlcNAc-β-(1,3)Fuc-α- bond suggest that a transferase unique to this pathway is required [33]. Recently, new *O*-fucosylated patterns have been identified within thrombospondin type-1 repeats, suggesting a functional role in cell adhesion [37].

The Notch signaling pathway demonstrates the importance of *O*-fucosylation and *O*-glucosylation for cell signaling systems in most multicellular organisms. The Notch protein is

a transmembrane receptor protein where ligand proteins binding to the extracellular domain induce proteolytic cleavage and release of the intracellular domain. The cleaved intracellular domain migrates to the cell nucleus to alter gene expression [38]. The Notch 1 extracellular domain is composed primarily of small cysteine knot motifs called EGF-like repeats [39]. Each EGF-like repeat is composed of approximately 40 amino acids, and its structure is defined largely by six conserved cysteine residues. The EGF-like repeat can be modified by *O*-linked glycans at specific sites [40]. An *O*-glucose sugar may be added between the first and second conserved cysteines, and an *O*-fucose may be added between the second and third conserved cysteines. The addition of *O*-fucose by GDP-fucose and protein *O*-fucosyltransferase 1 (POFUT1). POFUT is absolutely necessary for Notch function, and, without *O*-fucose, the Notch protein fails to function properly [39]. The *O*-fucose on Notch can be further modified by an *N*-acetylglucosaminyltransferase called fringe, followed by the addition of galactose by a galactosyltransferase, and sialic acid by a sialyltransferase. The addition of these sugars may inhibit or potentiate signaling through ligand proteins binding to the extracellular domain although the full mechanism is not completely understood [41].

O-mannosylation was discovered in yeast cells in the 1950's and only recently has been found among mammals, where it occurs especially in the brain, nerves, and skeletal muscles [33, 42]. Particular interest has been directed towards a tetrasaccharide derived from α-dystroglycan (ADG), a component of the dystrophin-glycoprotein complex in skeletal muscle, where abnormal glycosylation is associated with neuromuscular diseases [42, 43]. ADG is a peripheral membrane protein that undergoes multiple and complex glycosylation steps to regulate its interaction with extracellular matrix proteins such as laminin, agrin and perlecan [44]. The mucin type *O*-glycans are clustered in a mucin-like domain near the *N*-terminus of mature ADG, which includes unique *O*-mannosyl glycans and sialic acid: Neu5Ac-α-2,3Galβ-1,4GlcNAcβ-1,2Manα-1Ser/Thr [12]. Defects in glycosylation of the *O*-mannosyl glycans have been shown to cause muscular dystrophy [12]. Seven glycosyltransferases or glycosyltransferase-like genes, including POMT1, POMT2, POMGnT1, Fukutin, Fukutin-related protein, LARGE, and LARGE2, have been found to be involved in ADG functional glycosylation [12]. In particular, overexpression of LARGE has restored dystroglycan glycosylation and laminin binding properties in primary cell cultures from patients affected by different genetically defined dystroglycanopathy variants [44]. LARGE displays 2 distinct structural domains homologous to UDP-glucose protein glucosyltransferase and β-3-*N*-acetylglucosaminyltransferase (β3GnT1) [44]. It is β3GnT1 that plays a critical role in the synthesis of the laminin-binding glycans. Furthermore, several reports have shown that defects in ADG are associated with breast, colon, oral, and prostate carcinomas but only recently it has been demonstrated that reduced expression of β3GnT1 leads to diminished

synthesis of laminin-binding glycans, increased metastasis of cancer cells, and increased tumour formation [44].

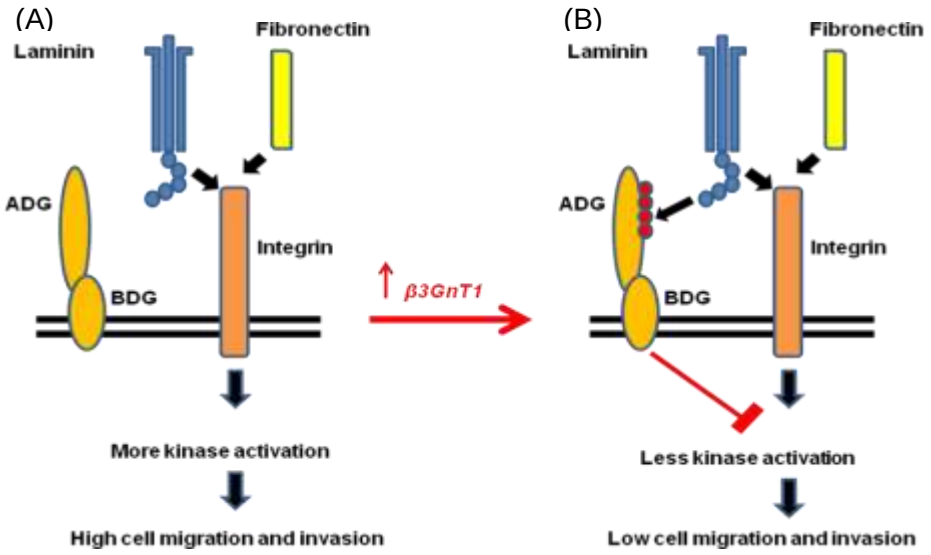


Figure 1.4 *The effect of ADG glycosylation on extracellular matrix proteins. (A) Interaction of intact ECM molecules such as laminin and fibronectin with integrin initiates kinase activation and promotes cancer cell migration and proliferation. (B) Up-regulation of β3GnT1 results in ADG playing an increased role in counteracting integrin-mediated signalling, resulting in a low migratory cell phenotype [44].*

In mammals, *O*-xylose is found as the first sugar residue of the tetrasaccharide GlcAβ1-3Galβ1-3Galβ1-4Xylβ1-*O*-Ser, initiating the formation of the glycosaminoglycans (GAGs). These highly negatively charged molecules are primarily located on the surface of cells or in the extracellular matrix, where they act as lubricating fluids and provide passageways between cells, allowing for cell migration. GAGs are long unbranched polysaccharides containing repeating disaccharide units. The disaccharide units contain either of two modified sugars, *N*-acetylgalactosamine (GalNAc) or *N*-acetylglucosamine (GlcNAc), and an uronic acid such as glucuronate or iduronate [45, 46].

Some important GAGs include:

- Hyaluronates composed of *D*-glucuronate-β(1,3)-GlcNAc linkages; found in synovial fluid and the ECM of loose connective tissue [46].
- Dermatan sulphates composed of *L*-iduronate-α(1,3)-GalNAc-4-sulphate linkages; found in skin, blood vessels and heart valves [46].

- Chondroitin 4- and 6-sulphates composed of *D*-glucuronate-β(1,3)-GalNAc-4- or 6-sulphate linkages; found in cartilage, bone, heart valves and is the most abundant GAG [46].
- Heparin and heparan sulphates composed of iduronate-2-sulphate-α(1,4)-*N*-sulfo-*D*-glucosamine-6-sulphate linkages. These sulphates are a component of intracellular granules of mast cells lining the arteries of the lungs, liver and skin [46].
- Keratan sulphates composed of galactose-β(1,4)GlcNAc-6-sulphate linkages and are found in cornea, bone and cartilage aggregated with chondroitin sulphates [46].

1.4 Glycosyltransferases

Glycosyltransferases (GTs) are enzymes that catalyse the transfer of a monosaccharide unit from an activated sugar phosphate (known as the "glycosyl donor") to an acceptor molecule, usually an alcohol. GTs usually transfer sugar molecules to the amino acid tyrosine, serine or threonine to give *O*-linked glycoproteins, or to asparagine to give *N*-linked glycoproteins. The commercial availability of GTs are limited and the nucleotide diphosphate-monosaccharide donors are expensive. Despite this drawback, GTs still remains attractive in synthesising oligosaccharides over traditional chemical methods. The main advantages of GTs over chemical synthesis are the absence of protecting group chemistry and the ability of GTs to easily control regioselectivity and stereoselectivity.

New methods to adapt the role of GTs have been investigated for potential biomedical applications [47-51]. For example, changing the glycosyl acceptor, the substrate to which a GT transfers its sugar molecule, allows the synthesis of a variety of oligosaccharides with different acceptor molecules. However this ultimately relies on the tolerance of the GTs to the glycosyl acceptor [52-57]. Less commonly GTs also catalysed reactions, where the glycosyl donor, a modified nucleotide diphosphate is used instead of its natural nucleotide diphosphate. This comprises of either a chemoenzymatic or a synthetic approach to produce a modified glycosyl donors where the modification is directly on the sugar which will be transferred by GTs.

The reaction mechanism of GTs can either proceed with inversion or retention of configuration at the anomeric centre. The mechanism of inversion has been elucidated (**Figure 1.5**), but the mechanism by retention is still under debate.

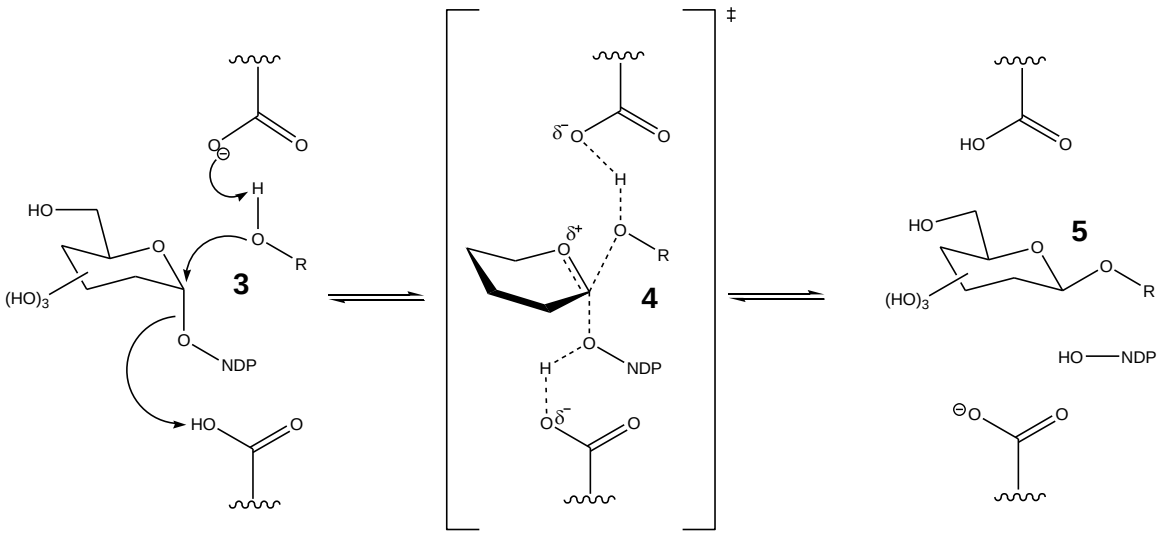


Figure 1.5 *Reaction mechanism of an inverting GT [57, 58].*

The two most widely accepted hypotheses that rationalise the retention of configuration involve either a double displacement dependent on acidic/basic residues on the enzyme (**Figure 1.6**) or through a direct attack of the aglycon with concomitant release of the nucleotide diphosphate (NDP) in a S_{Ni} mechanism. This involves an “internal return” intermediate (**Figure 1.7**) [57, 58].

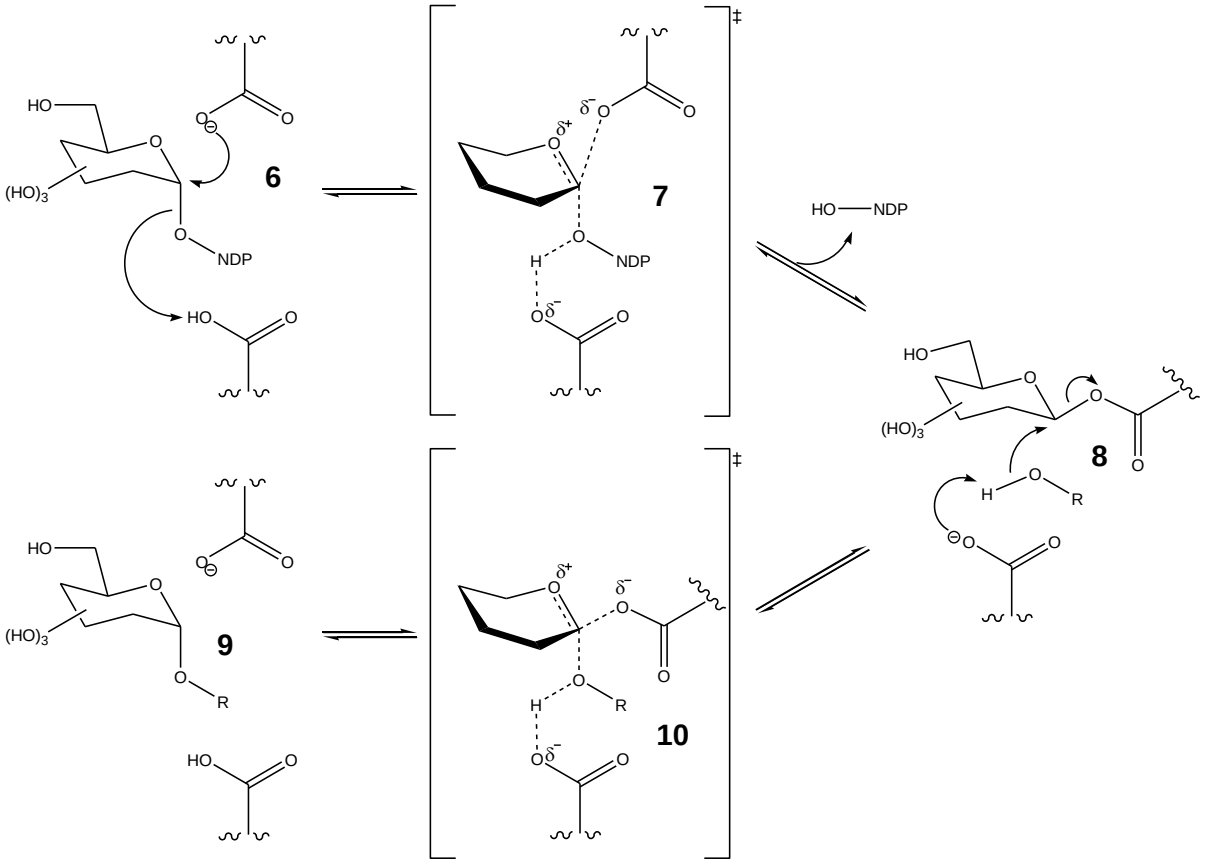


Figure 1.6 *Reaction mechanism of a retaining GT* [57, 58].

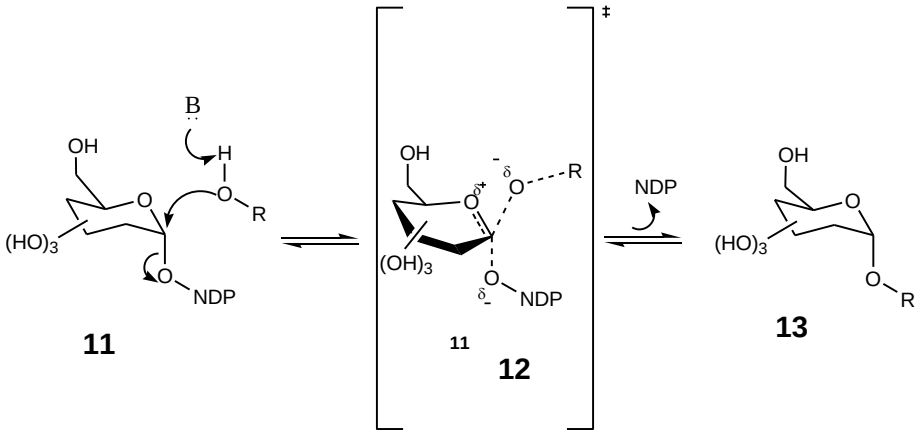


Figure 1.7 *Proposed S_{Ni} mechanism of a retaining GT* [57, 58].

1.5 Glycosidase and glycosynthase

Glycosidases are enzymes which hydrolyse the glycosidic bonds between sugar molecules, and in so doing, they catalyse the reverse process to that of GTs. Withers *et al*. have shown that glycosynthase functionality can be conferred upon a subset of glycosidase enzymes by mutagenising the active site. He demonstrated that the mutagenised enzymes form glycoconjugates while losing the ability to hydrolyse the product [59].

Glycosynthases were originally formed from glycosidases by mutating the active site nucleophilic amino acid (usually an aspartate or glutamate) to a small non-nucleophilic amino acid (usually alanine or glycine) [60]. More modern approaches use directed evolution to screen for amino acid substitutions that enhance glycosynthase activity.

Two discoveries led to the development of glycosynthase enzymes. The first was changing of the active site nucleophile of a glycosidase from a carboxylate to another amino acid prevented the formation of a hydrolysed product. The second discovery was that some glycosidase enzymes were able to catalyse the hydrolysis of glycosyl fluorides with inversion of configuration (**Figure 1.8 A**) [61, 62]. The enzymes underwent a transglycosidation reaction to form a disaccharide, which was then hydrolysed. The removal of the carboxylate group in glycosynthase prevents the formation of a covalent acyl enzyme intermediate. An activated glycosyl donor with a good anomeric-leaving group (often fluorine) is required. The leaving group is displaced by an alcohol of the acceptor sugar aided by the enzyme active site (**Figure 1.8 B**) [61, 62].

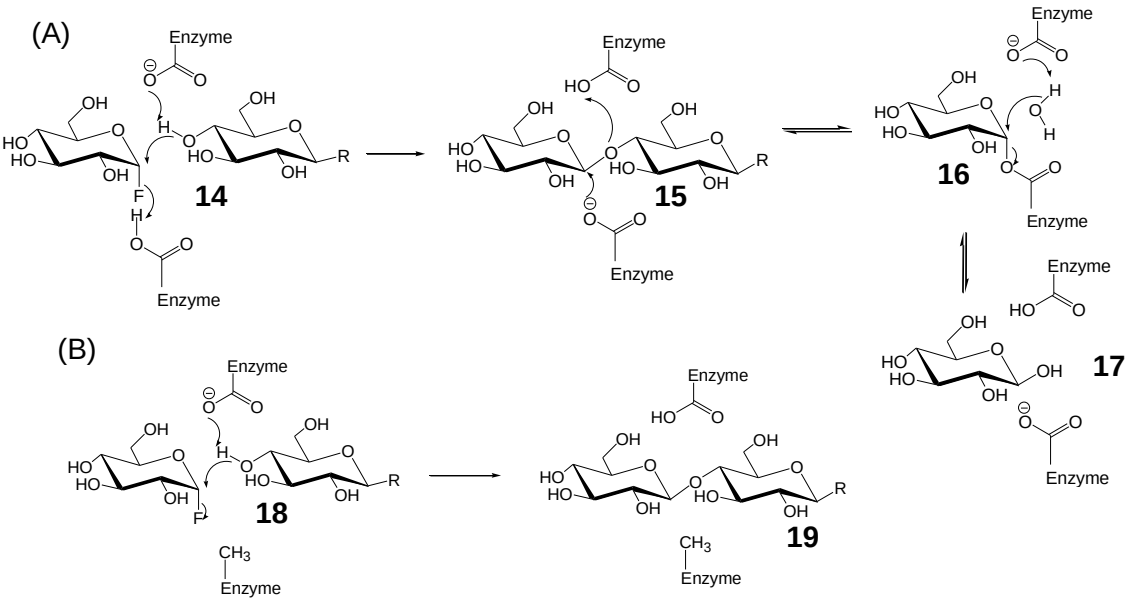


Figure 1.8 *Mechanism of glycosidase and glycosynthase enzymes. (A) Transglycosidation followed by hydrolysis with retention. (B)* A glycosidic linkage is formed where a non-covalent enzyme intermediate is formed [61, 62].

1.6 Mannosyltransferase (Man T)

Mannose is present in a wide range of glycoproteins and exhibits *O*-linkages to other sugars or to the hydroxyl groups of Ser or Thr residues of peptides [63]. Mannose forms *C*-linkages

to Trp residues of protein chains and *P*-linkages through monophosphate units to other carbohydrates or polyisoprenoids (the dolichol lipid, through β linkage in humans) [64].

In eukaryotic organisms the occurrence of mannosylation in the *N*-glycosylation pathway is best understood. This is governed primarily by a class of enzymes called mannosyltransferases which catalyse the transfer of mannose from a donor to an acceptor molecule which is frequently another carbohydrate. There are two different types of glycosyl donors in which mannose is linked to, guanosine diphosphate (GDP) or dolichol-phosphate. Many ManTs are strongly discriminating regarding the nature of the glycosyl donor. However, changing the acceptor may be viable, for instance dolichol-phosphate mannose synthase transferred mannose to a dolichyl-phosphate (Dol-P) and synthetic Dol-P analogues containing fluorescent probes [65].

Many conserved eukaryotic ManT genes were originally identified in *S. cerevisiae* (asparagine-linked glycosylation (*alg*) genes), so it is not surprising that yeast nomenclature is mainly used. ManTs can be classified according to the pathway they are involved in (*N*-glycosylation, *O*-glycosylation, or glycosylphosphatidylinositol (GPI) anchoring), the type of linkage formed (either α or β) and by position of the glycosidic bond. This can be shown by the common core *N*-linked glycoproteins where eight different ManTs are used respectively β-1,4(alg1p), α-1,3(alg2p), α-1,6(alg2p), α-1,2(alg11p), α-1,6(alg12p), α-1,3(alg3p), α-1,2(alg9p) and dolichol-phosphate-mannose synthase (Dol-P-Man). Dol-P-Man is used to transfer mannose from GDP-Man to dolichol-P which in turn serves as a donor and transfers mannose to glycan-PP-dolichol. The exact pathway position and number of mannose introduced by these mannosyltransferase is shown (**Figure 1.0**) but can be summarised below:

β-1,4-ManT: transfers mannose from GDP-man to dolichol-PP-$GlcNAc_2$ at the cytoplasmic face of the ER membrane (**Figure 1.9**)[66].

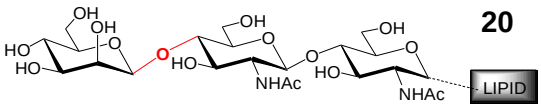


Figure 1.9 *The structure of lipid-linked trisaccharide after the addition of mannose by β-1,4-ManT.*

α-1,3-/α-1,6-ManT: adds one α-1,3- and one α-1,6-linked mannose to the trisaccharide-PP-dolichol, at the cytosolic face of the ER membrane (**Figure 1.10**) [67].

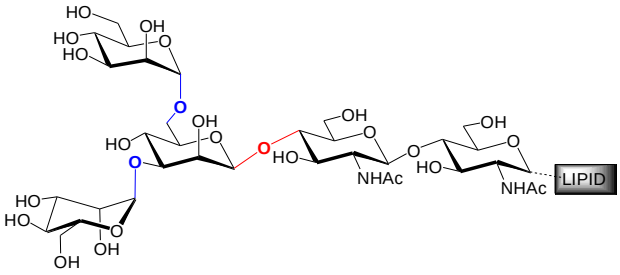


Figure 1.10 *The structure of lipid-linked pentasaccharide after the addition of mannose by α-1,3-/α-1,6-ManT.*

α-1,2-ManT: uses GDP-mannose to add the next two mannose units to form the heptasaccharide-PP-dolichol, which is then flipped to the other side of the ER (**Figure 1.11**) [68].

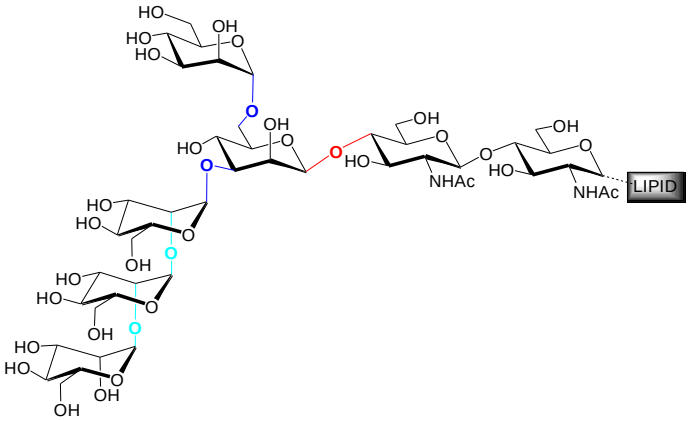


Figure 1.11 *The structure of lipid-linked oligosaccharide after the addition of mannose by α-1,2-ManT.*

α-1,3-ManT: is the first ManT to add a mannose residue from the dolichol-P donor to the heptasaccharide-PP-dolichol at the lumenal face of the ER membrane in the *N*-glycosylation pathway [69] (**Figure 1.12**).

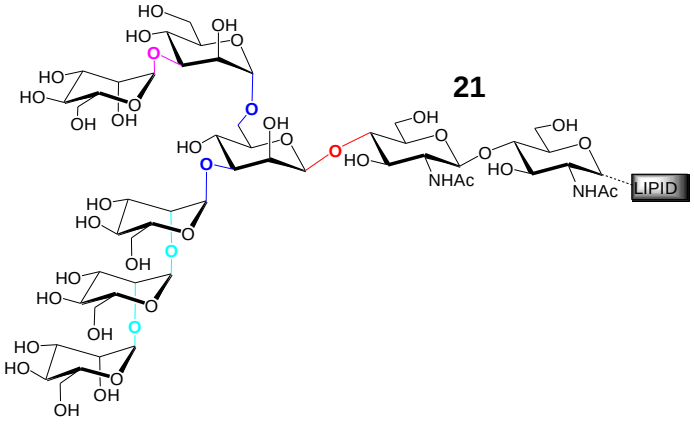


Figure 1.12 *The structure of lipid-linked oligosaccharide after the addition of mannose by α-1,3-ManT.*

α-1,2-ManT: adds two mannose units to the growing oligosaccharide by using the dolichol-P-man as a donor [70] (**Figure 1.13**). This protein has high homology with GPI-MT-I (PIG-Mp), GPI-MT-II(PIG-Vp), GPI-MT-III (PIG-Bp), and GPI-MT-IV (SMP3p) which sequentially adds the four mannose units to the GPI anchor [71]. They are all localised on the lumenal side of the ER membrane.

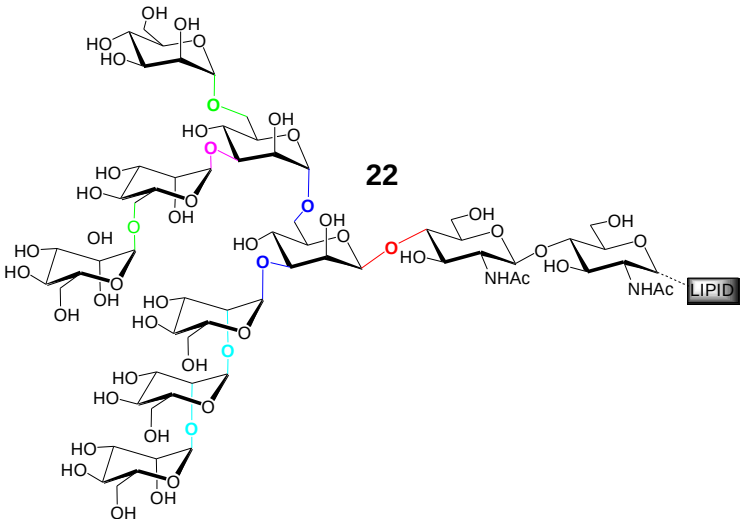


Figure 1.13 *The structure of lipid-linked oligosaccharide after the addition of mannose by α-1,2-ManT.*

α-1,6-ManT: adds another mannose to the lipid-linked oligosaccharide, and shares significant homology with α-1,2-ManT [72].

Dol-P-Man synthase : transfers mannose from GDP to dolichol-phosphate which in turn will serve as a donor in GPI anchoringand for *O*- and *C*-glycosylation pathways. The enzyme is a β-inverting transferase that is highly conserved among eukaryotes and in humans. It forms a complex with two other membrane proteins Dpm2p and Dpm3p [73]. Another significant mannosyltransferase is protein *O*-mannosyl-transferase 1 which is encoded by the POMT1 gene in humans. It is a member of the dolichyl-phosphate-mannose-protein mannosyltransferases family. POMT1 and POMT2 both form a heterocomplex which *O*-mannosylates α-dystroglycan in the ER, by using dolichol-P-man [74].

The catalytic transfer of mannose units onto acceptor substrates is not restricted to a single ManT enzyme, but despite this apparent redundancy of function the ManT's exhibit limited sequence homology. Christine and coworkers [75], compared unrelated families of yeast Golgi ManT's and showed a short motif containing two aspartate residues (DXD motif) conserved in both groups of proteins. Mutagenesis of one of the members of these families,

the α-1,3-ManT, showed that alteration of either of these aspartates eliminated all enzymatic activity [76]. It is further believed that this DXD motif is involved in binding a metal ion (usually magnesium), which is used to coordinate the phosphate of the glycosyl donor in the active site of the protein.

Whilst the properties of a small number of ManTs have been well documented [76], the majority remain poorly characterised. For instance, only one eukaryotic ManT crystal structure has been reported to date, and online database entries show sequences of potential ManTs with unknown substrates [76]. Also, the ManT responsible for C-mannosylation in humans has still not been identified. One of the reasons for this may be the difficulty in handling these proteins, as they are transmembrane proteins containing hydrophobic transmembrane domains which makes them difficult to isolate. Despite these difficulties, ManTs are attracting more interest in the scientific community for their involvement in diverse human disease states [77-84]. The interaction of mannose with its receptor on macrophages and endothelial cells is involved in the abnormal immunological response, which includes failed clearance of pathogens and autoimmune reactions [85]. The receptor is also involved in tumour proliferation, and in the uptake of vaccines and drugs. Great potential exists for ManTs as novel drug targets for treating or understanding some of these pathological conditions.

So far, we have looked at the biosynthesis and role of glycoproteins in biology and concentrated on glycosyltransferases and, in particular, mannosyltransferases. Next we will look at the engineering of glycoproteins as drugs and the different methods currently employed.

1.7 Glycoprotein drugs/engineering

Glycoproteins play a key role in a variety of biological functions from cell communication to immune response and inflammation. The carbohydrate moiety has been found to be essential for correct folding, delivery, and solubility of many glycoproteins [23, 86]. The ability to harness these functions by engineering glycoproteins to be non-immunogenic for therapeutic use, at industrial scale and cost poses a huge challenge but is nevertheless highly desirable. To control glycosylation by manual or automated means presents a complex challenge compared to other biological processes such as nucleotide assembly. The number of monosaccharides and the possible ways of linking them together leads to an immense diversity in the glycan structure.

The production of glycoproteins can be divided into three major routes using biological, synthetic and semi-synthetic methods. Controlling the glycosylation of proteins by bioengineering yeast, bacteria, plant cells, insect cells, and mammalian cells have been investigated [87, 88]. The use of glycosylation enzymes to transfer sugars that can then be chemically modified is another approach. Total synthesis and enzyme protein ligation adds to the vast number of methods currently being used.

More than 50% of human proteins are glycoproteins, which tend to be heterogeneously glycosylated [89]. This means that the sugars are attached at different sites along the protein (macroheterogeneity) and in different arrangements to form an array of glycan structures (microheterogeneity). Heterogeneous glycosylation is also found in bioactive natural products such as avermectin, antiparasitics, enediyne antibiotics, macrolides such as erythromycin and vancomycin, a natural glycopeptide antibiotic [89].

The ability to control the type of glycosylation and produce homogeneous glycoproteins will allow researchers to relate glycan structures to biological activity and potentially tailor glycoprotein drugs to suit stability, uptake, immunogenicity and many other pharmacokinetic properties [89]. One example of an *N*-linked glycoprotein that has been studied extensively is the hormone erythropoietin (EPO) involved in the production of red blood cells (erythrocytes) [90]. Glycosylation of EPO is critical for its biological longevity and activity, but deducing the favoured glycosylation for optimum pharmacokinetic activity still poses a problem. EPO, like many other glycoproteins, is currently produced commercially in mammalian cells, which express it as a mixture of glycoforms. Hence controlling the glycosylation is beneficial to obtain biologically uniform molecules which are tailored for optimum therapeutic use.

1.8 Commercial production of *N*-linked glycoproteins

There are 140 therapeutic proteins approved in the US and Europe [91], with many more in clinical and preclinical trials. The therapeutic protein market can be divided into proteins which are non-glycosylated and those that are mostly *N*-glycosylated. Non-glycosylated proteins constitute for around 40% of the therapeutic market and are normally expressed in *E. coli* or *S. cerevisiae* [91]. The use of mammalian cell lines (for example baby hamster kidney [BHK] or Chinese hamster ovary cells [CHO]) is currently the expression system of choice for many approved therapeutic glycoproteins, including recombinant interferon-β, which is used in the treatment of multiple sclerosis and EPO, which is used to treat certain anaemias [92]. Although human type glycosylation is observed in different glycoforms,

culturing mammalian cells is a lengthy and very expensive process and handling mammalian cells can be difficult as they can be sensitive to environmental and process changes. So other expression systems may be attractive alternatives.

1.9 Engineering bacteria to produce *N*-linked glycoproteins

It was initially thought that expressing proteins in prokaryotic hosts such as *E. coli* did not glycosylate the protein. However, Wacker *et al* reported the presence of *N*-glycoproteins in the gram negative bacterium *Campylobacter jejuni* (C. *jejuni*) [93]. The glycan structure GalNAc-α4-GalNAc-α4-[Glcβ3-]GalNAc-α4-GalNAc-α4-GalNAc-α4-GalNAc-α-3-Bac-β-Asn (Bac = bacillosamine) was found to be vastly different to eukaryote glycan $Glc_3Man_9GlcNAc_2$. Furthermore, the synthesis of $Glc_3Man_9GlcNAc_2$ normally requires several enzymes and a "flippase" which are encoded by the *pgl* gene clusters. Studies have shown that the *pglB* gene was found in C. *jejuni* and researchers succeeded in reconstituting the C. *jejuni* glycosylation machinery in *E. coli*, by introducing the *pgl* gene cluster. The results showed that the bacteria *N*-glycosylation site is more specific than in eukaryotic systems, as the peptide acceptor sequence which *pglB* protein recognises is extended in bacteria: Asp/Gly-Yaa-Asn-Xaa-Ser/Thr (Xaa and Yaa are not proline). However, *in vitro* assay of *pglB* protein showed that *pglB* can transfer *N*-glycans to completely folded proteins, which indicates that *N*-glycosylation apparently occurs without certain organelles such as ER which are not present in bacteria. Using *pglB* for the production of human type *N*-glycans is still at an early stage and further developments for the uptake of specific donor substrates, such as UDP-GlcNAc and UDP-GalNAc or GlcNAc-α-3-Bac-undecaprenylpyrosphates will be a good starting point to engineer *N*-linked glycoproteins in bacteria.

1.10 Engineering yeast to produce mammalian glycoproteins

Due to low cost and ease of handling, yeast and bacteria are often favoured when manufacturing proteins and enzymes on an industrial scale [94, 95]. However, the use of yeast, while convenient, is not without a significant caveat: glycosylation conferred upon the recipient protein is of the high mannose type and is thus immunogenic to humans.

The *N*-glycosylation pathways in humans and yeast is initiated in the ER, where the core oligosaccharide $Glc_3Man_9GlcNAc_2$ is transferred onto the nascent polypeptide. In humans, glucosidase I, glucosidase II and ER-residing α-(1,2)-mannosidases trim the glycan to $Man_8GlcNAc_2$ for processing in the Golgi apparatus. The situation in yeast however is markedly different. *Schizosaccharomyces pombe* lacks ER α-(1,2)-mannosidases and

therefore transport glycoproteins containing $Man_9GlcNAc_2$ to the Golgi apparatus [92]. In *S. cerevisiae*, the Golgi apparatus contains enzymes α-(1,2), α-(1,3), and α-(1,6)-mannosyltransferases which transfer mannose to the *N*-glycan. This results in the formation of hypermannosylated glycans, with terminal mannose residues that are immunogenic in humans, and are recognised for rapid clearance from the circulation by the mannose macrophage receptor. Hypermannosylation thereby presents a barrier to the use of yeast glycoproteins as therapeutics. Furthermore the yeast Golgi is devoid of GlcNAc transferases, and *N*-acetylneuraminic acid (sialic acid) transferase, which are present in humans (**Figure 1.14**) [92]. If yeast is to be exploited to produce glycoproteins fit for therapeutic applications it must first be engineered or "humanised". Humanisation strategies are not trivial, and frequently rely upon the elimination of endogenous hypermannosylation pathways and the introduction of non-native enzymes by genetic means.

Recently, a company called GlycoFi has successfully generated human-like proteins from yeast [89]. Genetically engineering yeast, they produced a library of yeast strains that have been engineered to perform specific human glycosylation at high fidelity. They engineered yeast *P. pastoris* to secrete human glycoproteins with fully complex terminally sialylated *N*-glycans. By the knockout of four genes to eliminate yeast-specific glycosylation, and introduction of 14 heterologous genes, allowed step wise replication of human glycosylation. The reported cell lines produced complex glycoproteins with greater than 90% terminal sialylation. The utility of these yeast strains were shown by the production of functional recombinant erythropoietin [95].

Of all of the expression systems specifically engineered to produce therapeutic glycoproteins for use in humans, yeast is the most advanced. Further work will need to concentrate on genetically incorporating specialised sugars to allow complex glycopeptide synthesis with a variety of glycans.

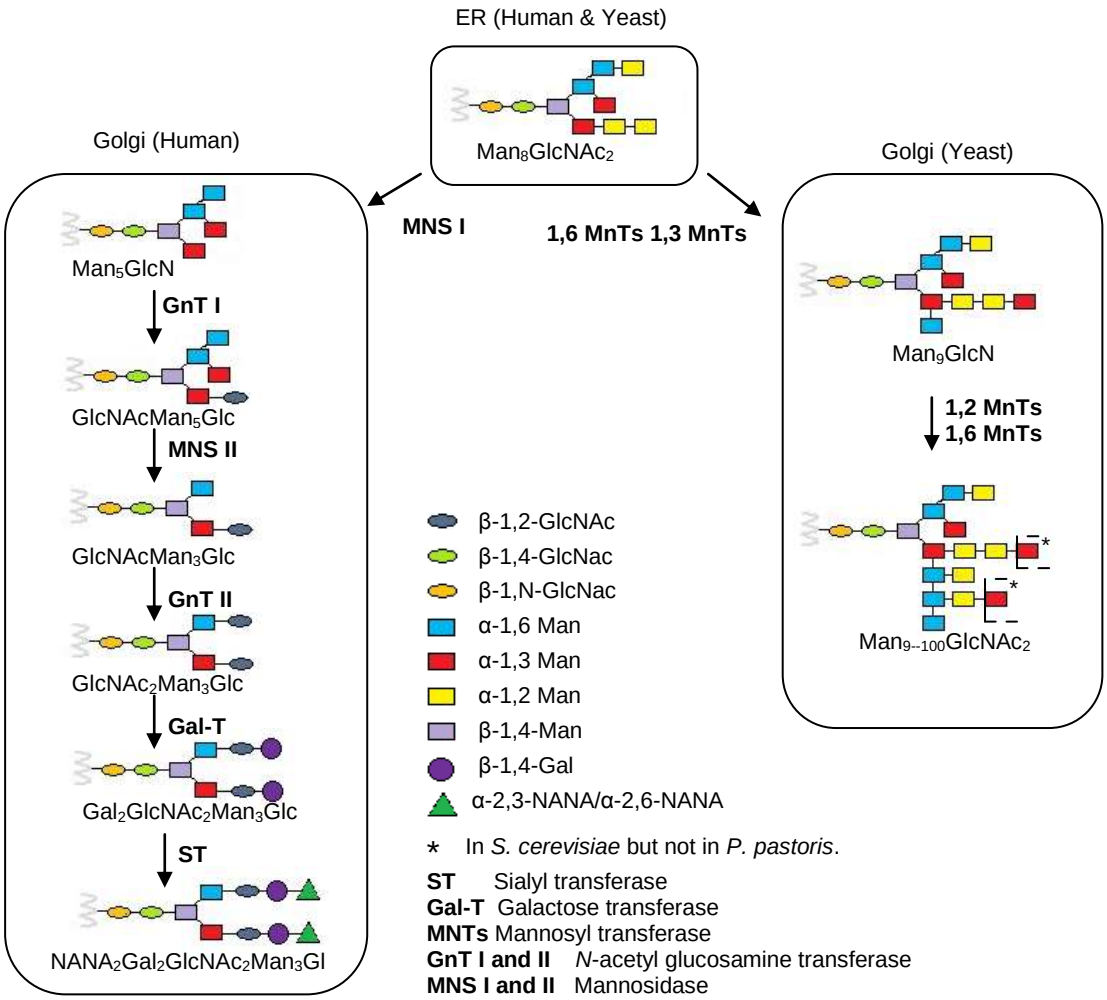


Figure 1.14 *Major N-glycosylation pathways in humans and yeast.*

1.11 Engineering plant and insect cells to produce glycoproteins

In contrast to yeast and bacteria, plant cells contain organelles capable of producing complex *N*-glycans. However, the presence of immunogenic fucose and xylose sugars render the associated glycoproteins unsuitable for use as therapeutic agents [89]. This apparent limitation was overcome by Stross and co-workers [89] who engineered plants to eliminate fucose and xylose and later developed a process to express galactosyltransferase in plants, which adds galactose to glycoproteins (this had never previously been achieved in plants [89]). Recently Steinkellner constructed a functional CMP-sialic acid biosynthesis pathway in *Arabidopsis thaliana*. The expression of key enzymes for mammalian *N*-acetylneuraminic acid (Neu5Ac) biosynthesis was achieved. These enzymes UDP-*N*-acetylglucosamine 2-epimerase / *N*-acetylmannosamine kinase, *N*-acetylneuraminic acid phosphate synthase, and CMP-*N*-acetylneuraminic acid synthetase resulted in the generation of Neu5Ac. Which was further converted to cytidine monophospho-*N*-

acetylneuramic acid by co-expression of CMP-*N*-acetylneuraminic acid synthetase [96]. These results suggested that it is only a matter of time before the first fully humanised *N*-linked-glycoprotein will be produced in genetically engineered plants.

Genetically engineered insect cells are another method which has been used to produce glycoproteins. Betenbaugh and coworkers have humanised glycoproteins produced in insect cells by genetically engineering the cells to express metabolic enzymes and glycosyltransferases required to add GlcNAc, galactose, and sialic acid to insect glycans [97]. The main advantage of this system compared to mammalian systems is the ability to provide more recombinant glycoprotein in a biologically system that is less influenced by environmental changes.

1.12 Total synthesis of glycoproteins and glycans

So far we have looked some biological systems focussed upon genetically engineering cells to perform human-type glycosylation and with natural sugar-protein linkages. Engineering *N*-linked glycopeptides can also be achieved synthetically by total synthesis or semi-synthetically using a combination of biologically/chemically modified enzymes and proteins. Researchers believe that to chemically synthesise large proteins from the start is both costly and inefficient with only small quantities produced. Ligating two or more separate peptides together to create a full length protein can overcome this problem as synthesis of smaller peptides can be readily automated. Native chemical ligation (NCL) is rapidly becoming established as a robust method for the production of full-length proteins from peptide components [98]. NCL is the chemoselective coupling of two protein fragments, one containing a N-terminal cysteine residue and the other bearing a C-terminal thioester, where a native peptide (amide) bond is formed at the point of ligation [99]. There are many adaptions of NCL where development of thiol auxiliaries and sugar assisted ligation has overcome the limitation of the need for *N*-terminal cysteine residues [100,101] (**Figure 1.15**).

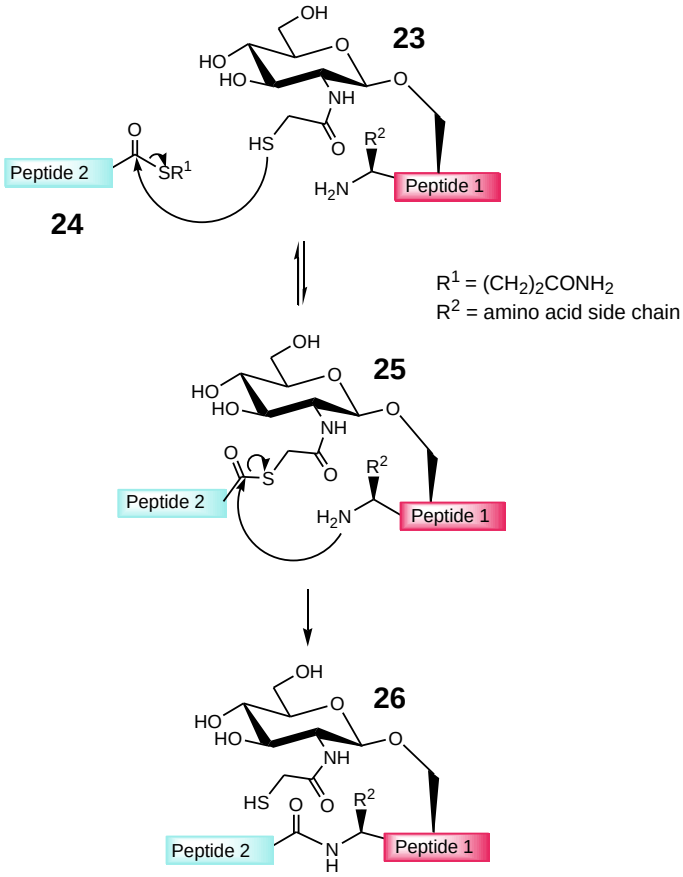


Figure 1.15 *Sugar-assisted ligation (SAL)*.

NCL is an attractive approach for the synthesis of *N*-linked glycoproteins, as it allows the construction of completely native structures by total chemical synthesis in aqueous solutions in the absence of protecting groups. This is particularly important for glycoproteins, as control can be exerted upon the position, number and type of oligosaccharides adorning the protein sequence. Additionally, controlling the synthesis of complex homogenous glycoproteins or chemically functionalising glycoproteins can be useful for therapeutic use and monitoring. Danishefsky and co-workers demonstrated the power of NCL by synthesising homogeneous EPO-like glycopeptides using cysteine-free ligations [102]. The method involved conjugating trifunctional glycopeptide using chemical ligation. Firstly the attachment of chitobiose and then exposing the glycopeptide to tris-carboxyethyl phosphine (TCEP), which reduced the disulphide bonds in both the acyl transfer donor and auxiliary-linked acceptor. Following the O→S acyl shift, the thioester was formed (**Figure 1.16**) which participates in an NCL reaction.

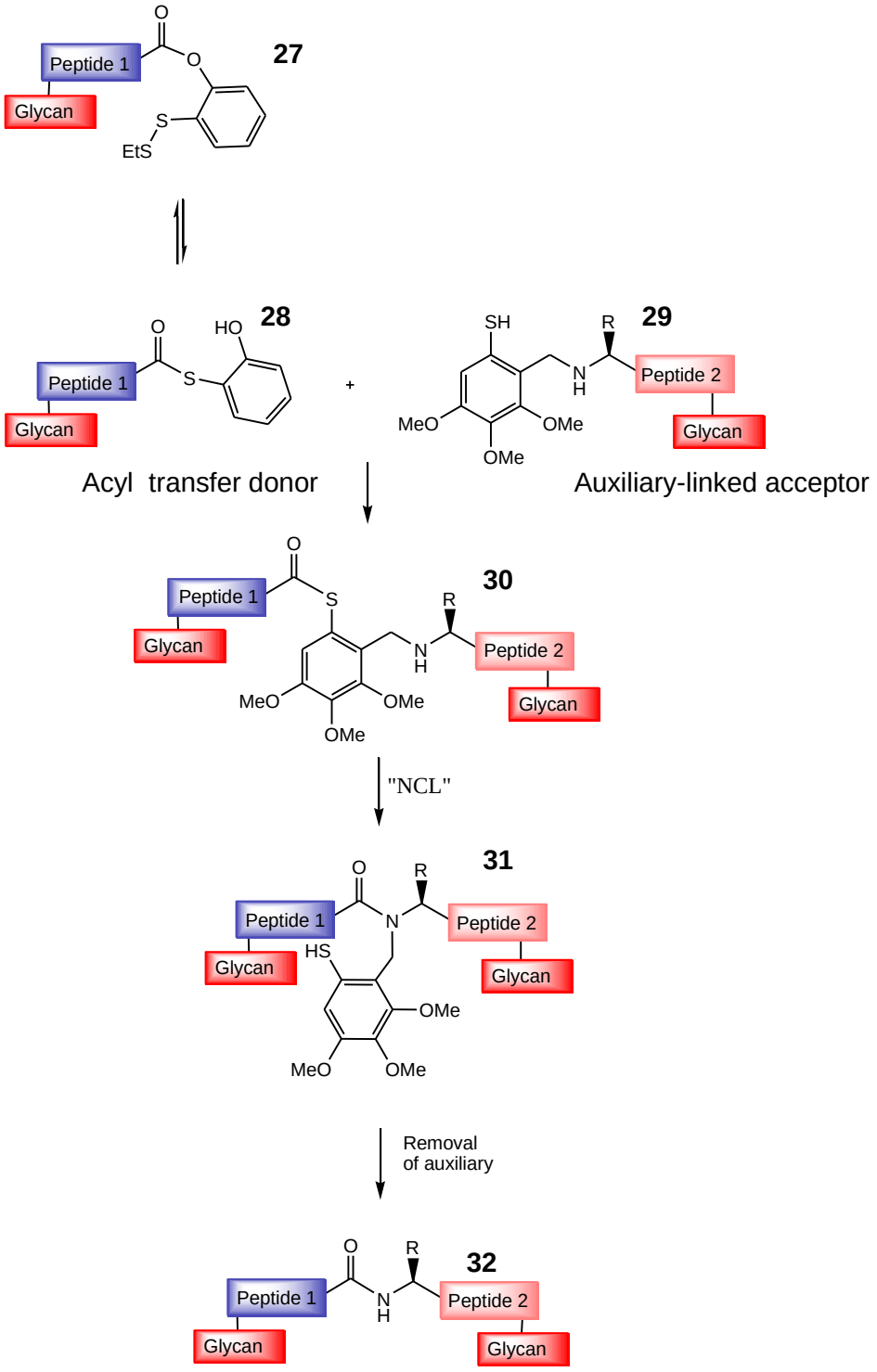


Figure 1.16 *Danishefsky's cysteine-free ligation. Reduction of the disulphide bond and an O→S acyl shift regenerates the thioester, which undergoes auxiliary assisted NCL.*

The Kent laboratory further examined the kinetics of aryl and alkyl thioester participation in NCL reactions and found that aryl thioesters were more reactive, but were produced *in situ* with the addition of thiophenol to an alkyl thioester [103].

Total chemical synthesis of glycans remains one of the oldest and most challenging strategies to produce glycans. Unverzagt and coworkers have pursued the total synthesis of complex oligosaccharides, including bisected *N*-glycans [104]. Their recent work showed a double regio- and stereoselective glycosylation strategy for the synthesis of *N*-glycans. They showed using a benzylidene acetal protected β-mannoside, the selective α-glycosylation at OH-3 of the 2,3-diol with GlcNβ1,2Man trichloroacetimidate donors [105]. In addition, this selectivity was increased with a phthalimido-protected donor. The total synthesis of glycans remains both laborious and time consuming due to its protecting group strategy. However it will always be very important, especially when complex homogenous glycoproteins will need to be synthesised with one point ligation to proteins.

1.13 Semi-synthesis of glycoproteins

Apart from the use of NCL to produce glycoproteins, glycoproteins produced by recombinant methods are easily accessible, and it is usually the glycan segment that is immunogenic to humans. Therefore to utilise the protein part and remodel the glycan *in vitro*, by semi-synthetic means, could pose a good method to produced homogeneous glycoproteins.

One approach in this semi-synthesis category involves isolation of a mixture of glycoforms of a single protein followed by exposure to a glycoside hydrolase, which mediates glycan degradation to the point of homogeneity. The uniformly glycosylated protein can then be remodelled *in vitro* using enzymes. An exo-glycoside hydrolase acts by sequential removal of non-reducing terminal monosaccharides, while an endo-glycoside hydrolyses the internal glycosidic bond of glycans [106,107]. A good example of an endo-glycoside hydrolase are endo-β-*N*-actylglucosaminidases (endo-NAGs), which hydrolyses the β(1-4) linkage of the *N*,*N*'-diacetylchitobiose unit found in *N*-linked glycans. The resulting product by digestion with endo-NAGs will be free from microheterogeneity and further remodelling with enzymes such as glycosyltransferases may allow non immunogenic uniformly glycosylated proteins. Some endo-NAGs can also be used to catalyse transglycosylation of oligosaccharides from glycopeptides to βGlcNAc-Asn [107,108]. Two widely used endo-NAGs are bacterial enzyme Endo-A *(Arthobacter)* and the fungal enzyme Endo-M *(Mucor hiemalis*). These two *endo*-enzymes have distinct substrate specificity. Endo-A is specific for high-mannose type *N*-glycans, whereas Endo-M can act on *N*-glycans of the three major types (high-mannose type, hybrid type, and complex type), with preference to the complex type sugars. Furthermore, endo-NAGs can also use GlcNAc-based oxazolines as substrates for glycoprotein remodelling [107,108]. The nature of the glycan substrates that are transferred determines

whether endo-NAGs hydrolyse the product formed. For example disaccharide oxazoline, Glc-β-(1,4)-GlcNAc-oxazoline, was a substrate of Endo-M for transglycosylation but was not a substrate for Endo-A. However low reaction rates with Endo-M were observed with only a 5% yield due to hydrolysis of the product. When the substrate was extended to the trisaccharide Man-α-(1,3)-Glc-β-(1,4)-GlcNAc-oxazoline, it became an excellent substrate for Endo-M, and the transglycosylation proceeded with 91% yield without hydrolysis.

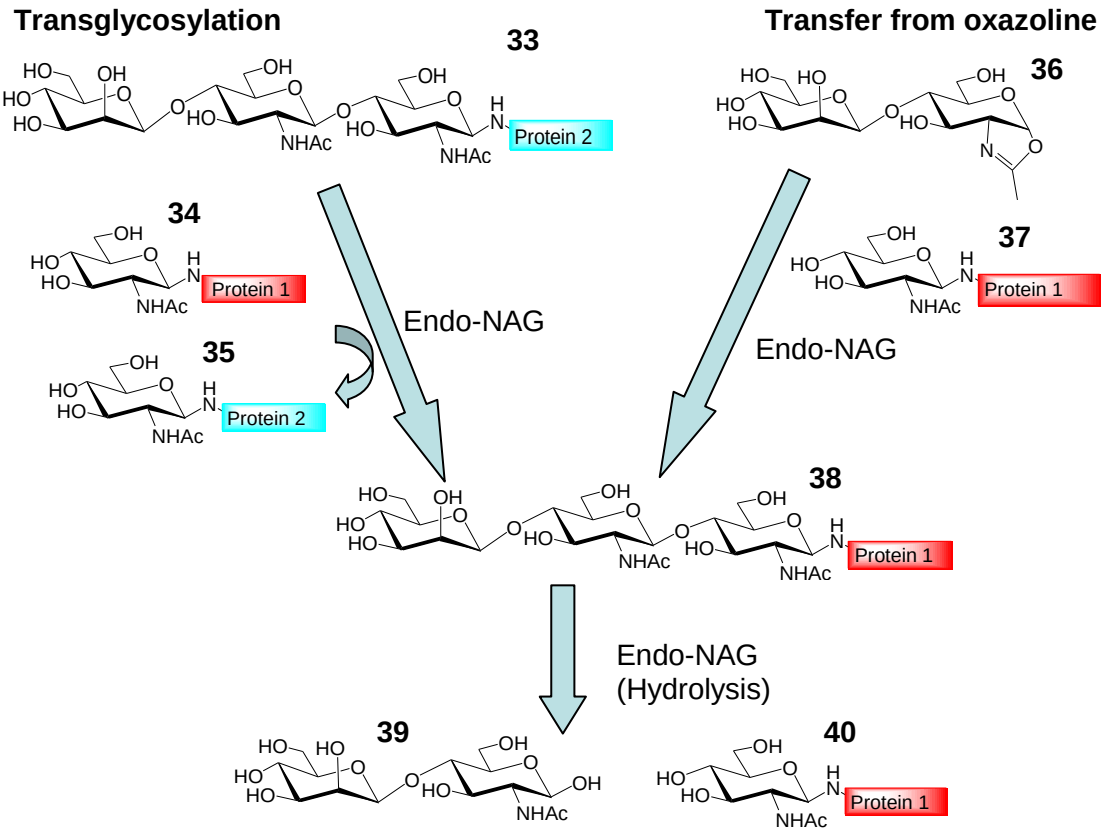


Figure 1.17 *The uses of endo-NAGs.*

Further developments of endoglycosidase-catalysed transglycosylation of triazole-linked glucose and *N*-acetylglucosamine-containing di-peptides and polypeptides were achieved using sugar oxazoline as donor substrates. Endo-A was effective in transglycosylation to both *N*- and *C*-linked Glc/GlcNAc-containing triazole derivatives [107-109].

These recent developments of endo-NAG-catalysed glycosyl transfer for sugar oxazolines will be imperative in the preparation of homogeneous *N*-linked glycoproteins. In fact a new method of producing homogeneous *N*-linked glycoproteins from eukaryotic cells has been demonstrated. This involves engineering the functional transfer of the *C. jejuni* glycosylation machinery in *E. coil* to express glycosylated proteins containing GlcNAc-Asn linkage. The *N*-glycan is trimmed by α-*N*-acetylgalactosaminidase to give a GlcNAc-tagged glycoprotein.

Finally, endoglycosidase-catalysed transglycosylation using pre-assembled *N*-glycan species as substrates gives rise to a pool of homogenous glycoproteins [110].

Many ligation methods have been developed to attach glycans to recombinant proteins. One method utilises the properties of selenium and sulphur compounds to ligate sugars to proteins through disulphide bonds [111]. Another method uses “click chemistry” (see below) by using the heterobifunctional adaptor 2-bromoacetyl propargylamide to ligated azido sugars and cysteine containing peptides. An unnatural azido amino acid azidohomoalanine (Aha) has also been incorporated into a protein where click chemistry with propargyl sugars ligates the sugars to the protein [112]. Further work will be required to test the durability of one point ligation of sugars to proteins and assess the toxicity of the chemical modification of the glycoprotein for human therapeutic use.

1.14 Click chemistry

In 2002 the groups of Meldal and Sharpless independently reported that catalytic amounts of copper (I) catalysed azide-alkyne 1,3-dipolar cycloaddition (CuAAC) and improved its regioselectivity, yielding almost exclusively 1,4-substituted triazoles (**Figure 1.18**) [113, 114].

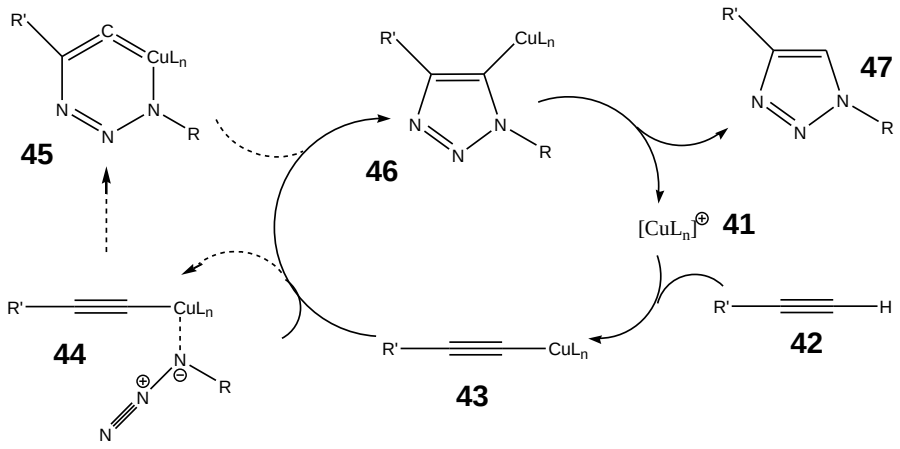


Figure 1.18 *Proposed mechanism for the Cu (I) catalysed 1,3-dipolar cycloaddition between a terminal alkyne and an azide* (*CuAAC).*

Nevertheless, the reaction is not flawless: the Cu (I) is thermodynamically unstable and is easily oxidised to Cu (II) or can disproportionate to Cu (II) and Cu (0), compromising the application of the reaction. For this reason, especially in aqueous environments, it is vital to add either anti-oxidant agents (e.g. ascorbic acid), or ligands typically Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl] amine (TBTA) that can preserve the Cu (I) species [113]. Since the conditions have been optimised for biological applications, this reaction has been widely

used in the field of carbohydrates [115].The terminology click chemistry was introduced by Sharpless, although originally to indicate a general group of efficient reactions with broad scope, and that are particularly convenient, fast and reliable [116]. Further developments of click chemistry using ruthenium catalysed 1,3-dipolar azide-alkyne cycloaddition (RuAAC) gave the 1,5-triazole [117, 118]. Unlike CuAAC in which only terminal alkynes reacted, in RuAAC both, terminal and internal alkynes can participate in the reaction [117, 118]. This suggests that ruthenium acetylides are not involved in the catalytic cycle.

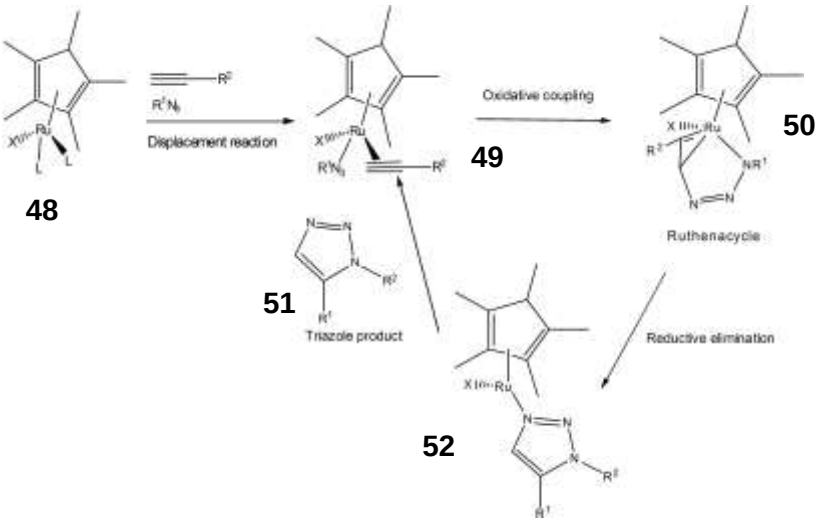


Figure 1.19*: Proposed mechanism for the Ru catalysed 1,3-dipolar cycloaddition between an alkyne and an azide. L= spectator ligand and X= Cl.*

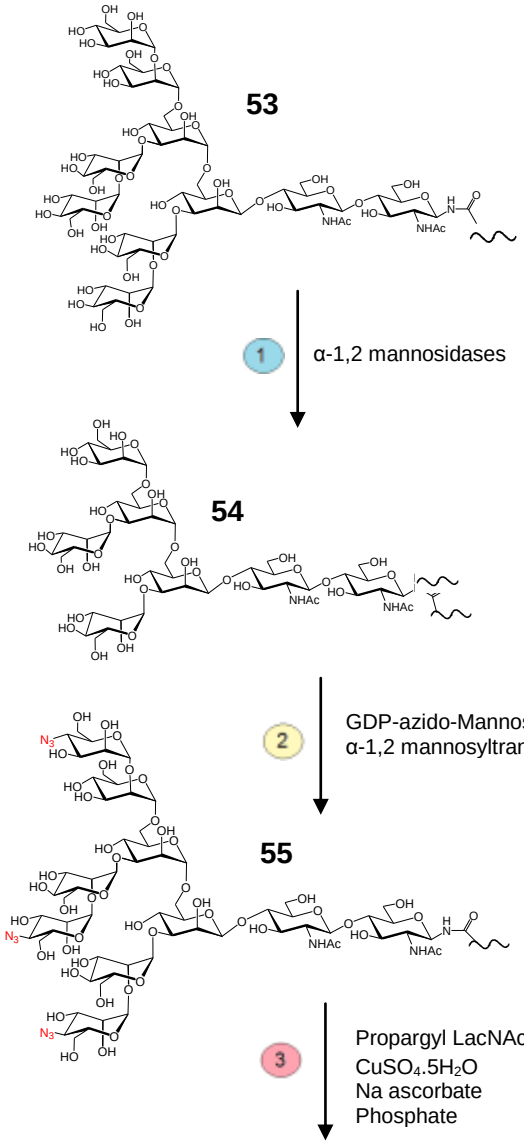
The proposed mechanism (**Figure 1.19**) begins where the spectator ligands undergo a displacement reaction to produce an activated complex which is converted, *via* oxidative coupling of an alkyne and an azide to the ruthenium containing metallocycle (ruthenacycle) [117, 118]. A new C-N bond is formed between the less sterically-demanding and more electronegative carbon of the alkyne and the terminal nitrogen of the azide. The metallocycle intermediate then undergoes reductive elimination to produce the aromatic triazole product and regenerate the catalyst for further reaction cycles.

1.15 Project aim

The overall aim of the project was to synthesise homogenous glycoproteins by three different methods. Initially the synthesis of propargyl glycosides of mannose, *N*-acetyl glucosamine, and *N*-acetyl lactosamine was investigated. These propargyl glycosides were used to remodel glycoproteins via click chemistry using the acetylene group of the glycosides and the azide group of the glycoprotein.

Aim 1: The first method involved expressing the glycoprotein in *Pichia pastoris* in the presence of azido sugars, in an attempt to incorporate the azido sugars into the yeast glycan. The isolated azide containing glycoprotein was then remodelled using our propargyl glycosides to produce a "humanised" chemically homogenous molecule.

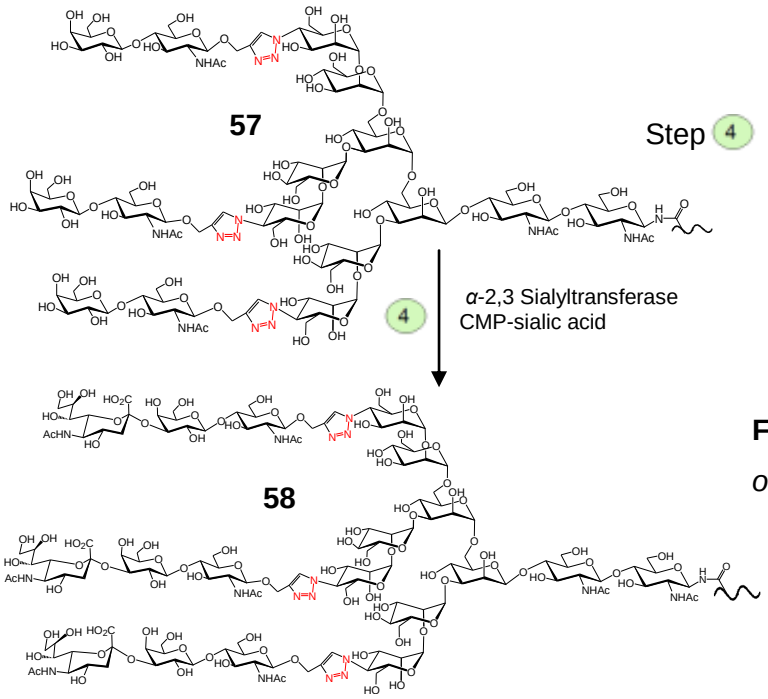
Aim 2: The second method uses a novel chemoenzymatic strategy which is summarised below:



Step 1 The glycosylation of the expressed protein will be trimmed by subjecting the purified hyperglycosylated protein to glycosidase enzymes (α-1,2 mannosidases).

Step 2 We will next investigate the transfer of an unnatural azidosugar using α-1,2 mannosyltransferase. We have synthesised GDP-azidosugars [119], and these will be used to transfer the azido sugar to the trimmed glycoprotein.

Step 3 Next the azide functionality of the transferred azidosugar will undergo click chemistry with a synthetic



propargyl glycoside of *N*-acetyllactosamine.

Step 4 Enzymatic transfer of sialic acid will finally complete the remodelling of the glycoprotein.

Figure 1.20 *Chemoenzymatic strategy of glycoprotein remodelling.*

The resulting engineered glycoprotein containing a triazole linked *N*-acetyllactosamine and sialic acid may mimic human type glycosylation and therefore be less immunogenic to humans. The trimming of yeast glycans and the specific transfer of sugars will produce a homogeneous modified glycoprotein.

Towards this aim a range of GDP-azidomannoses have been synthesised in our group where the key phosphate coupling of GTP to azidomannose-1-phosphate was performed enzymatically with GDP-mannose-pyrophosphorylase (**Figure 1.20**) [119]. The use of these GDP-azidomannose derivatives to transfer an azidomannose onto a yeast expressed and trimmed glycoprotein will be the basis of our remodelling technique.

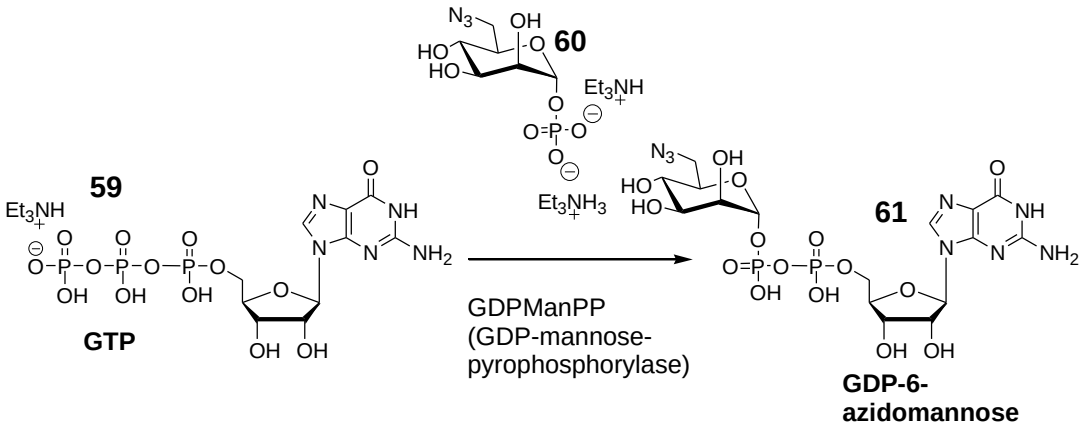


Figure 1.21 *Synthesis of GDP-azidomannose. (The figure shows the synthesis of 6-azidomannose but 2, 3, 4, and 6-azidomannose have been synthesised within our group).*

Aim 3: The third method involved chemically synthesising peptides using azidohomoalanine (Aha). Click chemistry with synthetic propargyl glycosides would allow further modification to the homogeneous glycopeptide analogues, which were tested for their compatibility with native chemical ligation.

References:

[1] L. W. Ruddock, M. Molinari, *J. Cell Sci.* 2006, 119, 4373.

[2] S. L. Flitsch, *Nature,* 2005, 437, 201.

[3] G. Patsos, A. Corfield, *Biol. Chem*. 2009, 390, 581.

[4] D. J. Hulmes, *J. Struct. Biol*. 2002, 137, 2.

[5] S. P. Parker, McGraw-Hill Encyclopedia of Science and Technology, 5th ed**,** The McGraw-Hill Companies, Inc., 2005.

[6] https://www.sigmaaldrich.com/img/assets/15880/glycan_classification.pdf (25/01/09)

[7] C. H. Wong, *J. Org. Chem*. 2005, 11, 4219.

[8] A. Grigorian, S. Torossian, M. Demetriou, *Immunol. Rev*. 2009, 230, 232.

[9] A. J. Chapman, M. I. Chuah, W. G. Breed, *Histol. Histopathol*. 2010, 1, 121.

[10] A. Bellahcène, V. Castronovo, K. U. E. Ogbureke, L. W. Fisher, N. S. Fedarko, *Nat. Rev. Cancer. 2008,* 8, 212.

[11] J. Lefkovits, E. F. Plow, E. J. Topol, *N. Engl. J. Med.* 1995, 23, 1553.

[12] F. Muntoni, S, Torelli, M. Brockington, *Neurotherapeutics*, 2008, 4, 627.

[13] C. M. Taylor, *Tetrahedron Lett.* 1998, 54, 11317.

[14] A. J. Petrescu, M. R. Wormald, R. A. Dwek, *Curr. Opin. Struct. Biol.* 2006, 16, 600.

[15] S. Wopereis, D. J. Lefeber, E. Morava, R. A. Wevers, *Clin. Chem*. 2006, 52, 574.

[16] M. Goto, *Biosci. Biotechnol. Biochem.* 2007, 71, 1415.

[17] T. Nishikawa, Y. Koide, S. Kajii, K. Wada, M. Ishikawa, M. Isobe, *Org. Biomol. Chem.* 2005, 3, 687.

[18] A. V. Nikolaev, I. V. Botvinko, A. J. Ross, *Carbohydrate. Res.* 2007, 342, 297.

[19] L. J. McDonald, L. A. Wainschel, N. J. Oppenheimer, J. Moss, *Biochem.* 1992, 31, 11881.

[20] L. J. McDonald, J. Moss, *Mol. Cell. Biochem.* 1994, 138, 221.

[21]D.Jorcke, M. Ziegler, A. Herrero-Yraola, M. Schweiger, *Biochem. J.* 1998, 332, 189.

[22] A. Helenius, M. Aebi, *Science*, 2001, 291, 2364.

[23] C. R. Bertozzi, L. L. Kiessling, *Science,* 2001, 291, 2357.

[24] D. B. Williams, *J. Cell Sci.* 2006, 119, 615.

[25] Y. Takeda, K. Totani, I. Matsuol, Y. Ito, *Curr. Opin. Chem. Biol*. 2009, 582.

[26] M. Fukuda, *Biochem. Biophys. Acta*. 2003, 346. 394.

[27] M. Crispin, D. J. Harvey, V. T. Chang, C. Yu, A. R. Aricescu, E. Y. Jones, S. J. Davis, R. A. Dwek, P. M. Rudd, *Glycobiology,* 2006, 16, 748.

[28] J. Roth, *Chem. Rev.* 2002, 102, 285.

[29] R. A. Dwek, *Chem. Rev.* 1996, 96, 683.

[30] D. Voet, J. G. Voet, *Biochemistry*, 3rd ed., John Wiley and Sons, Inc., 2004.

[31] L. Stryer, *Biochemistry,* 4th ed., Freeman., 1996.
[32] S. Wopereis, D. J. Lefeber, E. Morava, R. A. Wevers, *Clin. Chem.* 2006, 52, 574.
[33] A. Varki, R. Cummings, J. Esko, M. Chrispeels, H. Freeze, G. Hart, J. Marth, *Essentials of glycobiology*. 1st ed., Cold Spring Harbour Laboratory Press., 1999.
[34] H. C. Hang, C. R. Bertozzi, *Bioorg. Med. Chem.* 2005, 13, 5021.
[35] G. W. Hart, M. P. Housley, C. Slawson, *Nature,* 2007, 446, 1017.
[36] D. E. Birk, P. Bruckner, *Top. Curr. Chem*. 2005, 247, 185.
[37] L. Shao, R. S. Haltiwanger, *Cell. Mol. Life Sci.* 2003, 60, 241.
[38] F. Oswald, B Täuber, T Dobner, S Bourteele, U, Kostezka, G Adler, S Liptay, R. M. Schmid, *Mol. Cell. Biol.* 2001, 22, 7761.
[39] B. Ma, J. L. S. Grant, D. E. Taylor, *Glycobiology,* 2006, 12, 158.
[40] L, Shao, Y, Luo, D. J. Moloney, R. Haltiwanger, *Glycobiology,* 2002*,* 763.
[41] L. Lu, P. Stanley, *Methods Enzymol.* 2006*,* 417, 127.
[42] T. Willer, M. C. Valero, W. Tanner, J. Cruces, S. Strahl, *Curr. Opin. Struct. Biol.* 2003, *13*, 621.
[43] T. Endo, *Biochim. Biophys. Acta.* 1999, *1473*, 237.
[44] X. Bao, M. Kobayashi, S. Hatakeyama, K. Angata, D. Gullberg, J. Nakayama, M. N. Fukuda, M. Fukuda, *Proc. Natl. Acad. Sci. USA*. 2009, 106, 12109.
[45] T. Koike, T. Izumikawa, J. I. Tamura, H. Kitagawa, *Biochem. J*. 2009, 421, 157.
[46] http://themedicalbiochemistrypage.org/glycans.html (25/01/09)
[47] L. Lairson, B. Henrissat, G. J. Davies, S. G. Withers, *Ann. Rev. Biochem.* 2008, 77, 521.
[48] S. A. Borisova, L. Zhao, C. E. Melancon, C. L. Kao, H. W. Liu, *J. Am. Chem. Soc.* 2004, 126, 6534.
[49] M. C. T. Hartman, S. Jiang, J. S. Rush, C. J. Waechter, J. K. Coward, *Biochemistry,* 2007, 46, 11630.
[50] M. C. Galan, C. S. Dodson, A. P. Venot, G. J. Boons, *Bioorg. Med. Chem. Lett.* 2004, 14, 2205.
[51] I. Baig, M. Kharel, A. Kobylyanskyy, L. Zhu, Y. Rebets, B. Ostash, A. Luzhetskyy, A. Bechthold, V. A. Fedorenko, J. Rohr, *Angew. Chem. Int. Ed.* 2006, 45, 7842.
[52] S. A. Borisova, C. Zhang, H. Takahashi, H. Zhang, A. W. Wong, J. S. Thorson, H. W. Liu, *Angew. Chem. Int. Ed.* 2006, 45, 2748.
[53] J. R. Rich, A. Szpacenko, M. M. Palcic, D. R. Bundle, *Angew. Chem. Int. Ed.* 2004, 43, 613.
[54] A. Minami, R. Uchida, T. Eguchi, K. Kakinuma, *J. Am. Chem. Soc.* 2005, 127, 6148.
[55] M. Kopp, C. Rupprath, H. Irschik, A. Bechthold, L. Elling, R. Mueller, *ChemBioChem.* 2007, 8, 813.

[56] K. Qasba Pradman, B. Ramakrishnan, E. Boeggeman, *Trends Biochem. Sci.* 2005, 30, 53.
[57] L. L. Lairson, W. W. Wakarchuk, S. G. Withers, *Chem. Commun.* 2007, 365.
[58] L. F. Mackenzie, Q. P. Wang, S. G. Withers, *J. Am. Chem. Soc.* 1998, 120, 5583.
[59] S. M. Hancock, M. D. Vaughan, S. G. Withers, Curr. *Opin. Chem. Biol*. 2006, 10, 509.
[60] C. Mayer, D. L. Jakeman, M. Mah, G. Karjala, L. Gal, R. A. J. Warren, S. G. Withers, *Chem. Biol.* 2001, 8, 437.
[61] S. G. Withers, K. Rupitz, D. Trimbur, R. A. Warren, *J. Biochemistry*, 1992, 31, 9979.
[62] S. J. Williams, S. G.Withers, *Carbohydr. Res.* 2000, 327, 27.
[63] T. Willer, M. C. Valero, W. Tanner, J. Cruces, S. Strahl, *Curr. Opin. Struct. Biol.* 2003, 13, 621.
[64] I. Sprung, A. Ziegler, S. L. Flitsch, *Chem. Commun.* 2002, 32, 2676.
[65] E. Lamani, R. B. Mewbourne, D. S. Fletcher, S. D. Maltsev, L. L. Danilov, V. V. Veselovsky, A. V. Lozanova, N. Y. Grigorieva, O. A. Pinsker, J. Xing, W. T. Forsee, H. C. Cheung, J. S. Schutzbach, V. N. Shibaev, M. J. Jedrzejas, *Glycobiology,* 2006, 16, 666.
[66] C. F. Albright, P. W. Robbins, *J. Biol. Chem.* 1990, *265*, 7042.
[67] M. K. O'Reilly, G. Zhang, B. Imperiali, *Biochemistry.* 2006, 45, 9593.
[68] C. B. Sharma, R. Knauer, L. Lehle, *J. Biol. Chem.* 2001, 382, 321.
[69] C. G. Frank, M. Aebi, *Glycobiology,* 2005, 15, 1156.
[70] J. Y. Kang, Y. Hong, H. Ashida, N. Shishioh, Y. Murakami, Y. S. Morita, Y. Maeda, T. Kinoshita, *J. Biol. Chem.* 2005, 280, 9489.
[71] J. F. Cipollo, R. B. Trimble, *Glycobiology,* 2002, 12, 749.
[72] H. Ashida, Y. Maeda, T. Kinoshita, *J. Biol. Chem.* 2006, 281, 896.
[73] F. Muntoni, S. Torelli, M. Brockington, *Neurotherapeutics*, 2008, 4, 627.
[74] H. Manya, A. Chiba, R. U. Margolis, T. Endo, *Glycobiology,* 2006, 16, 863.
[75] C. A. R. Wiggins, S. Munro, *Proc. Natl. Acad. Sci. USA.* 1998, 95, 7945.
[76] http://www.rcsb.org/pdb/home/home.do (17/7/2009)
[77] A. M. Almeida, Y. Murakami, D. M. Layton, P. Hillmen, G. S. Sellick, Y. Maeda, S. Richards, S. Patterson, I. Kotsianidis, L. Mollica, D. H. Crawford, A. Baker, M. Ferguson, I. Roberts, R. Houlston, T. Kinoshita, A. Karadimitris, *Nature Med.* 2006, 12, 846.
[78] C. Thiel, M. Schwarz, M. Hasilik, U. Grieben, F. Hanefeld, L. Lehle, K. von Figura, C. Koerner, *Biochem. J.* 2002, 367, 195.
[79] M. Schwarz, C. Thiel, J. Lubbehusen, B. Dorland, T. de Koning, K. von Figura, L. Lehle, C. Korner, *Am. J. Hum. Genet.* 2004, 74, 472.
[80] T. Kawaguchi, H. Nakakuma, *Int. J. Hematol.* 2007, 86, 27.

[81]C. Thiel, M. Schwarz, J. H. Peng, M. Grzmil, M. Hasilik, T. Braulke, A. Kohlschutter, K. von Figura, L. Lehle, C. Korner, *J. Biol. Chem.* 2003, 278, 22498.

[82] T. Imbach, B. Schenk, E. Schollen, P. Burda, A. Stutz, S. Grunewald, N. M. Bailie, M. D. King, J. Jaeken, G. Matthijs, E. G. Berger, M. Aebi, T. Hennet, *J. Clin. Invest.* 2000, 105, 233.

[83] L. W. Sun, E. A. Eklund, W. K. Chung, C. Wang, J. Cohen, H. H. Freeze, *J. Clin. Endocr. Metab.* 2005, 90, 4371.

[84] C. G. Frank, C. E. Grubenmann, W. Eyaid, E. G. Berger, M. Aebi, T. Hennet, *Am. J. Hum. Genet.* 2004, 75, 146.

[85] P. R. Taylor, S. Gordon, L. Martinez-Pomares, *Trends Immunol.* 2005, 26, 104.

[86] P. M. Rudd, T. Elliott, P. Cresswell, I. A. Wilson, R. A. Dwek, *Science,* 2001, 291, 2370.

[87] N. Jenkins, *Cytotechnolog,* 2007, 53, 121.

[88] Y. Chiba, Y. Jigami, *Curr. Opin. Chem. Biol.* 2007, 11, 670.

[89] http://pubs.acs.org/cen/coverstory/84/8436glycosylation.html (17/7/2009)

[90] J. Tóvári, R. Pirker, J. Tímár, G. Ostoros, G. Kovács, B.Döme, Curr. Mol. Med. 2008, 6, 481.

[91] T. U.Gerngross, *Nature Biotec*. 2004, 22, 1409.

[92] S. Wildt, T. U. Gerngross, *Nat .Rev. Microbiol.* 2005, 3, 119.

[93] M. Wacker, D, Linton, P. G. Hitchen, M. Niya-Lazar, S. M. Halslam, S. J. North, M. Panico, H. R. Morris, A. Dell, B. W. Wren, Aebi, *Science,* 2002, 298, 1790.

[94] P. P. Jacobs, N. Callewaert, Curr. Mol. Med. 2009, 9, 774.

[95] L. Huijuan, N. Sethuraman, T. A. Stadheim, D. Zha, B. Prinz, N. Ballew, P. Bobrowicz, B. Kwon, W. J. Cook, M. Cukan, N. Eawa H. 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, T. U. Gerngross, *Nat. Biotechnol.* 2006, 24, 210.

[96] A. Castilo, M. Pabst, R. Leonard, C. Veit, F. Altmann, L. Mach, J. Glossl. R. Strasser, and H. Steinkellner, *Plant physioiogy*. 2008, 147, 331.

[97] K. Viswanathan, N. Tomiya, J. Park, S. Singh, Y. C. Lee, K. Palter, M. J. Betenbaugh, *J. Biol. Chem*. 2006, 281, 15929.

[98] P. E. Dawson, T. W. Muir, I. C. lewis, S. B. H. Kent, *Science*, 1994, 266, 776.

[99] D. Macmillan, *Angew. chem. Int. Ed*. 2006, 45, 7668.

[100] P. Botti, M. R. Carrasco, S. B. H. Kent, *Tetrahedron Lett*. 2001, 42, 1831.

[101] A. Brik, Y. Y. Tang, S. Ficht, C. H. Wong, *J. Am. Chem. Soc*. 2006, 128, 562.

[102] B. Wu, J. Chen, J. D. Warran, G. Chen, Z. Hua, S. J. Danishefsky, *Angew. chem. Int .Ed.* 2006, 45, 4116.

[103] D. Bang, B. L. Pentelute, S. B. H. Kent, *Angew. chem. Int. Ed*. 2007, 46, 4173.

[104] S. Eller, R. Schuberth, G. Gundel, J. Seifert, C. Unverzagt, *Angew. Chem. Int. Ed*. 2007, 46, 4173.

[105] C. Unverzagt, S. Eller, S. Mezzato, R. Schuberth, *Chem. Eur. J*. 2008, 14, 1304.

[106] J. R. Rich, S. G. Withers, *Nat. chem. Biol*. 2009, 5, 206.

[107] L. X. Wang, Carbohydr. Res. 2008, 242, 1509.

[108] D. P. Gamblin, P. Garnier, S. J. Ward, N. J. Oldham, A. J. Fairbanks, B. G. Davis, *Org. Biomol. Chem.* 2003, 1, 3642.

[109] D. P. Gamblin, P. Garnier, S. v. Kasteren, N. J. Oldham, A. J. Fairbanks, B. G. Davis, *Angew. Chem*. 2004, 116, 846.

[110] F. Schwarz, W. Huand, C. Li, B. L. Schulz, C. Lizak, A. Palumbo, S. Numao, D. Neri, M. Aebi, L. X. Wang, *Nat. Chem. Bio.* 2010, 314, 1.

[111] G. J. L. Bernardes, E. J. Grason, S. Thompson, J. M. Chalker, J. C. Errey, F. E. Oualid, T. D. W. Claridge, B. G. Davis. *Angew. Chem. Int. Ed*. 2008, 47, 2244.

[112] S. I. V. Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony, B. G. Davis, *Nature*, 2007, 446, 993.

[113] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem., Int. Ed.* 2002, 41, 2596.

[114] T. R. Chan, R. Hilgraf, K. B. Sharpless, V. V. Fokin, *Org. Lett.* 2004, 6, 2853.

[115] S. Dedola, S. A. Nepogodiev, R. A. Field, *Org. Biomol. Chem.* 2007, 5, 1006.

[116] H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed.* 2001, 40, 2004.

[117] L. Zhang, X. Chen, P. Xue, H. H. Y. Sun, I. D. Williams, K. B. Sharpless, V. V. Fokin, G. Jia *J. Am. Chem. Soc*. 2005**,** 46, 15998.

[118] B. C. Boren, S. Narayan, L. K. Rasmussen, L. Zhang, H. Zhao, Z. Lin, G. Jia, V. V. Fokin *J. Am. Chem. Soc*. 2008, 28, 8923.

[119] S. Marchesan, D. Macmillan, *Chem Commum.* 2008, 36, 4321.

Chapter 2

The synthesis of propargylic glycoside tools for protein remodelling

To test our glycoprotein remodelling strategies we need to synthesise tools that would allow the introduction of sugars to our modified glycoprotein. The synthesis of propargylic glycosides such as propargyl *N*-acetyl-*D*-lactosamine, propargyl mannose and propargyl *N*-acetyl-*D*-glucosamine will provide useful tools for incorporation into the azide modified glycoproteins. As discussed in chapter one, click chemistry will be the method used to introduce the propargylic glycosides to the azide bearing glycopeptides.

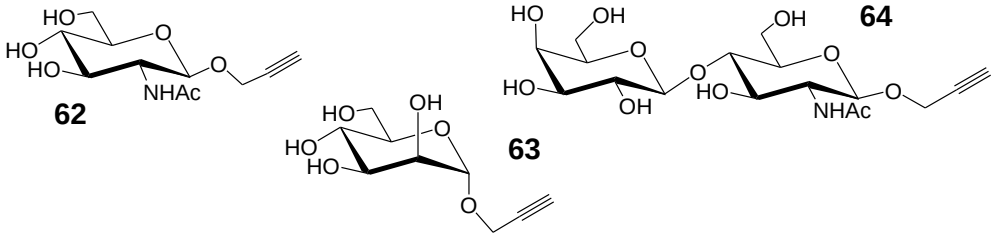


Figure 2.0 *The propargylic glycosides which will be synthesised in this chapter.*

2.1 *Results*

2.2 The synthesis of propargyl *N*-acetyl-*D*-lactosamine

The synthesis of propargyl *N*-acetyl-*D*-lactosamine has been demonstrated using ytterbium trifluoromethanesulfonate to introduce the propargyl group to a fully acetylated *N*-acetyl glucosamine [1] in the first instance. A round of deprotection and selective protection of the 6-hydroxy to form a glycosyl acceptor. Trichloroacetimidate activated galactose is used to form the glycosidic bond with the acceptor.

We will explore this chemistry further by attempting to synthesise this disaccharide from simple starting materials. Commercially available glucosamine hydrochloride and galactose will be initially used. The propargyl group will be introduced at the end of the synthesis. The retro-synthesis of propargyl LacNAc that we will undertake is illustrated in **Figure 2.1**.

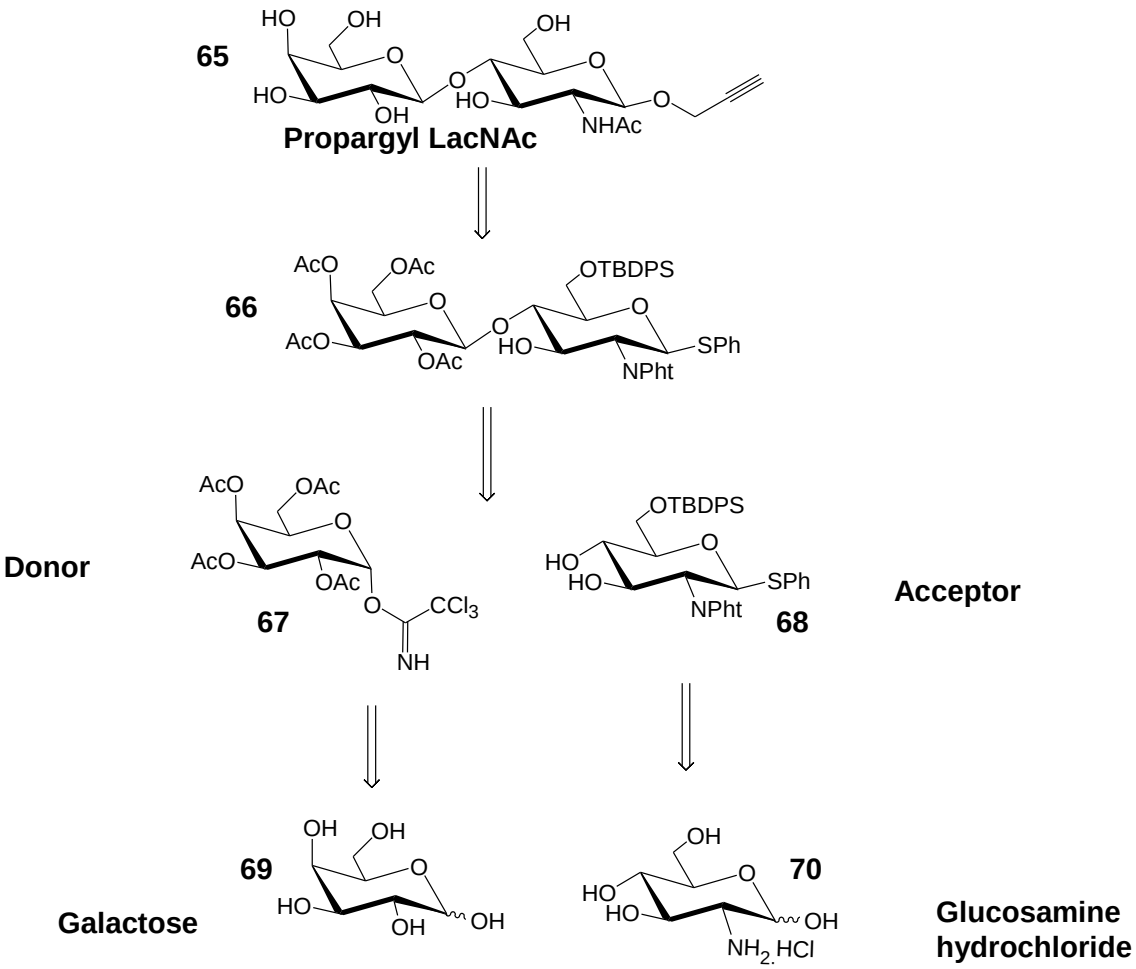


Figure 2.1 *Proposed retro-synthesis of Propargyl LacNAc.*

Glucosamine hydrochloride **69** (**Figure 2.2**) was converted to the free base with sodium methoxide in methanol and protected with phthalic anhydride. Apart from neighbouring group participation, the phthalimide at C-2 can also facilitate selective glycosylation of the hydroxyl at C-4 rather than hydroxyl at C-3, by sterically hindering the hydroxyl at C-3. The hydroxyl groups of the phthalimide protected sugar were acetylated to produce the fully protected glycoside **71** (**Figure 2.2**). It is important to note that before acetylating the product, the partially protected phthalimide sugar should be completely dried under high vacuum to obtain higher yields of **71**. Next the activation of the anomeric position with tin tetrachloride and treatment with thiophenol resulted in yields of up to 79% of the thiophenyl glycoside **72**. Deprotection of the acetates and selective protection of the less hindered primary hydroxyl group at position 6, with the bulky silylating reagent TBDPSCI, gave the partially protected thiophenyl glycoside acceptor **68**.

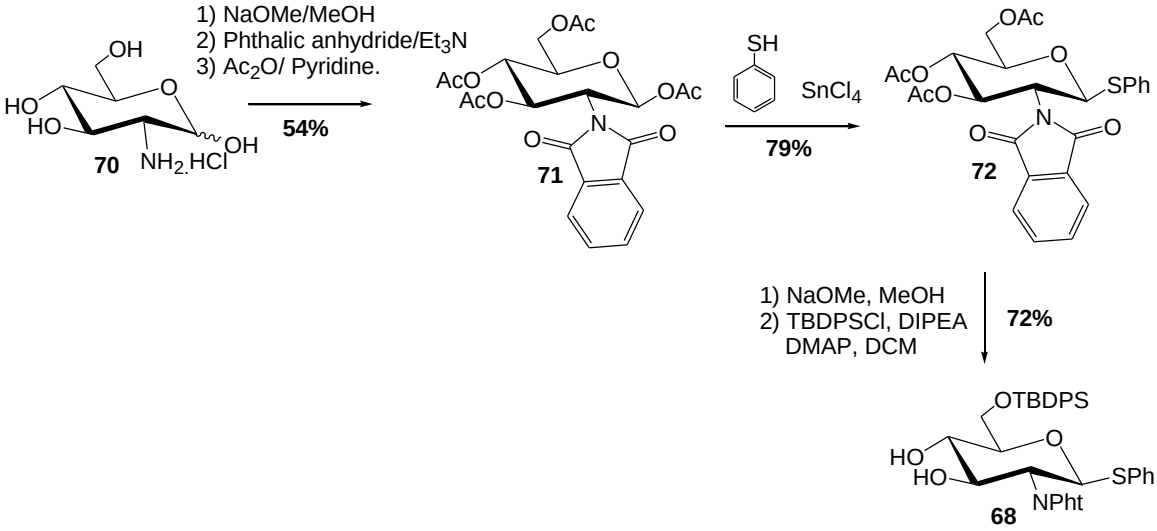


Figure 2.2 *Synthetic route for the thiophenyl glycoside acceptor **68** [2].*

The synthesis of the trichloroacetimidate donor was achieved from galactose **69**, which was fully acetylated with acetic anhydride and perchloric acid to give a good yield (98%) of the peracetylated galactose **73** (**Figure 2.12**). This was then treated with benzylamine at 50 ^{o}C, to selectively deprotect the anomeric position to give the hemi- acetal **74** in 82% yield.

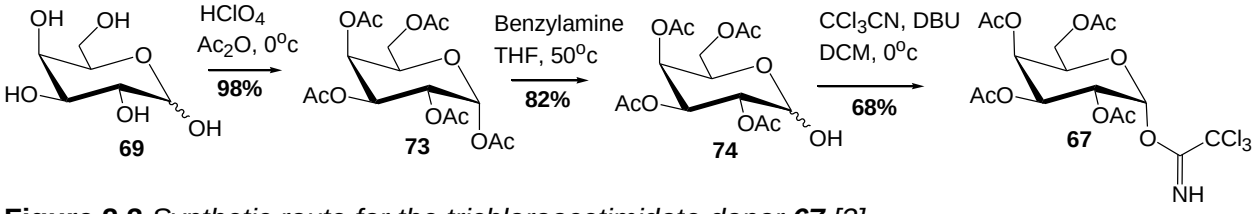


Figure 2.3 *Synthetic route for the trichloroacetimidate donor **67** [3].*

Next the anomeric trichloroacetimidate was produced by treatment of the hemi-acetal **74** with trichloroacetonitrile and diaza(1, 3) bicyclo[5.4.0]undecane (DBU). It is important to note that depending on the type of base used either the α- or β trichloroacetimidate may be obtained. With a weak base such as potassium carbonate, mutarotation occurs faster than the nucleophilic attack onto trichloroacetonitrile by the anomeric hydroxyl, and so the β product is predominantly formed. However for our chemistry we used the strong base DBU, which quantitatively deprotonates the anomeric OH and the rate of nucleophlic attack is faster than the rate of mutarotation. This means the thermodynamically more stable α-isomer **67** is formed (**Figure 2.4**).

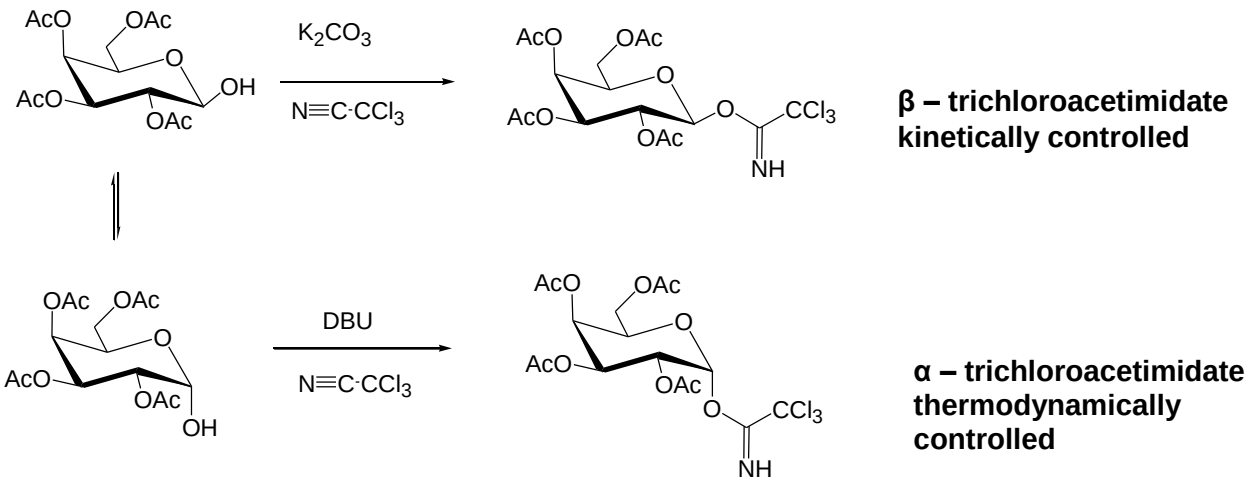


Figure 2.4 *Thermodynamically and kinetically controlled stereochemistry of trichloroacetimidates [4].*

Next, the glycosidic linkage was formed with the activation of the trichloroacetimidate with TMS-OTf *via* neighbouring group participation (**Figure 2.5**). The S_{N1} type mechanism selectively forms a β- 1-4 linkage, where the 4-hydroxyl group is the favoured nucleophile. The reason for this is predominately due to the presence of the bulky phthalimide which sterically hinders the 3-hydroxyl group. This would favour the β-linkage with the more accessible 4-hydroxyl group. This type of regioselectivity has been shown to be used in different glycosyl donors such as thioglycosides for the synthesis of sialyl-Lewis^x [5]. The protected thiophenyl disaccharide **66** was obtained in 84% yield. A series of selective deprotections with TBAF to remove the TBDPS protecting group, and ethylenediamine to remove the phthalimide produced the fully unprotected disaccharide which was acetylated with pyridine and acetic anhydride to give **75**. The final step of this synthesis was to install the propargyl group. This was investigated with the use of propargyl alcohol and the thiophenyl glycoside **75** in the presence of a soft electrophile. Initially *N*-iodosuccinimide (NIS) was used with catalytic TfOH. The activation of the anomeric sulphide to form a sulphonium leaving group which then undergoes a S_{N1} type of pathway with neighbouring group participation was envisaged to afford the desired β-propargyl glycoside. However this final step was problematic since a mixture of starting thiophenyl glycoside and its hemi-acetal **76** was mainly obtained. Mass Spectrometry analysis indicated that there were peaks representing the formation of oxazolines **77**.

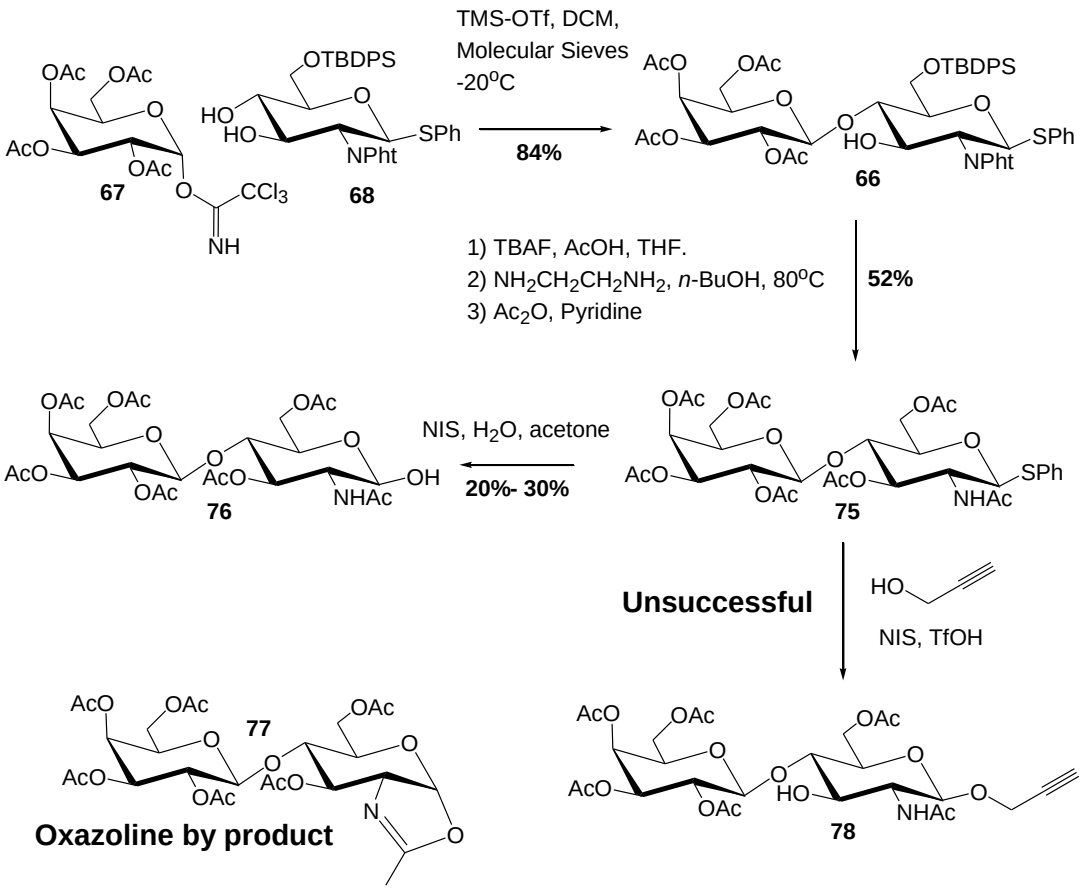


Figure 2.5 *Towards the synthesis of propargyl LacNAc.*

Alternative soft electrophiles such as *N*-bromosuccinimide (NBS) and iodonium dicollidine perchlorate (IDCP) were investigated, both being unsuccessful. With unsuccessful introduction of the propargyl group we decided to convert the thiophenyl glycoside into the hemi-acetal **76** and then to the trichloroacetimidate. The hemi-acetal was formed in poor yield so this route was not feasible. Losing material during the later stages of the synthesis was very inefficient and time consuming. An alternative route to introduce the propargyl group before the glycosidic linkage was undertaken. Starting with the thio-glycoside **68**, the 3- and 4-hydroxyl groups were acetylated with acetic anhydride and pyridine to give the fully protected glycoside **79** (**Figure 2.15**). This was then treated with NBS and water in acetone to form the hemi-acetal **80**. Furthermore the reaction with potassium carbonate and trichloroacetonitrile formed the kinetically controlled β-trichloroacetimidate **81**. Next the propargyl group was introduced with the lewis acid; boron trifluoride diethyl etherate and propargyl alcohol. With the presence of neighbouring group participation an S_{N1} type mechanism operates to yield the propargyl sugar **82**. Sodium methoxide and methanol was used to deprotect the 3- and 4-hydroxyl groups and to produce the new glycosyl acceptor **83** (**Figure 2.6**).

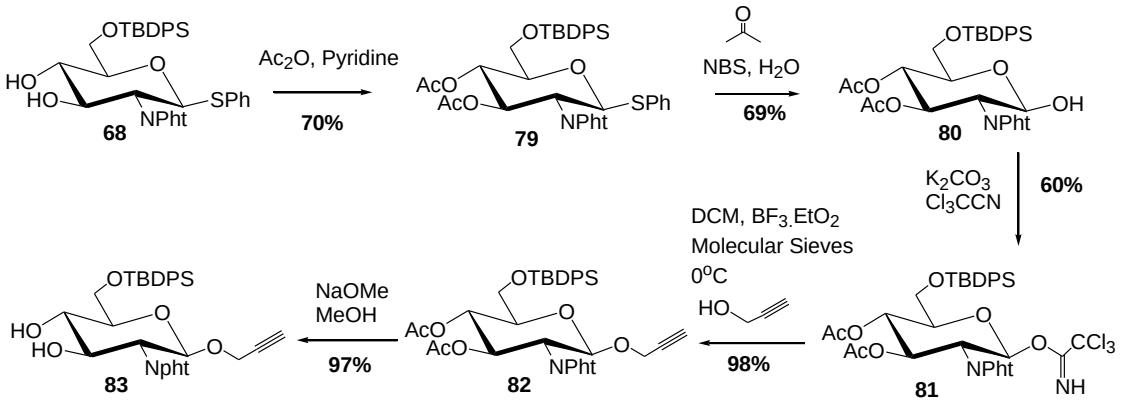


Figure 2.6 *Synthesis of building block **83**.*

This same glycosyl acceptor was more simply produced from the fully protected glycoside **71**, directly installing the propargyl group with boron trifluoride diethyl etherate (**Figure 2.7**). This reaction was difficult to monitor by thin layer chromatography (TLC), as the Rf values were identical in all the solvent systems tested. Therefore the reaction was monitored by LC-MS. The crude product was directly deprotected with sodium methoxide, then selective protection of the 6-hydroxy group with TBDPSCI to produce the glycosyl acceptor **83** (**Figure 2.7**).

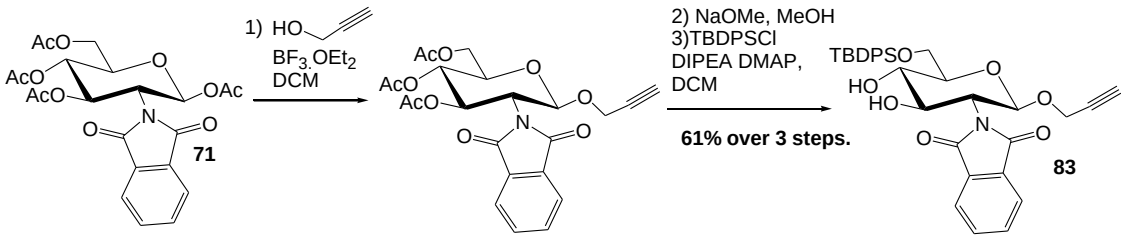


Figure 2.7 *Improved synthesis of acceptor **83**.*

The donor **67** was then used to selectively form the β-1,4 glycosidic linkage with acceptor **83**. The resulting fully protected propargyl *N*-acetyllactosamine **84** was deprotected with TBAF and ethylenediamine, and then acetylated to produce the fully acetylated propargyl *N*-acetyllactosamine **85** (**Figure 2.8**).

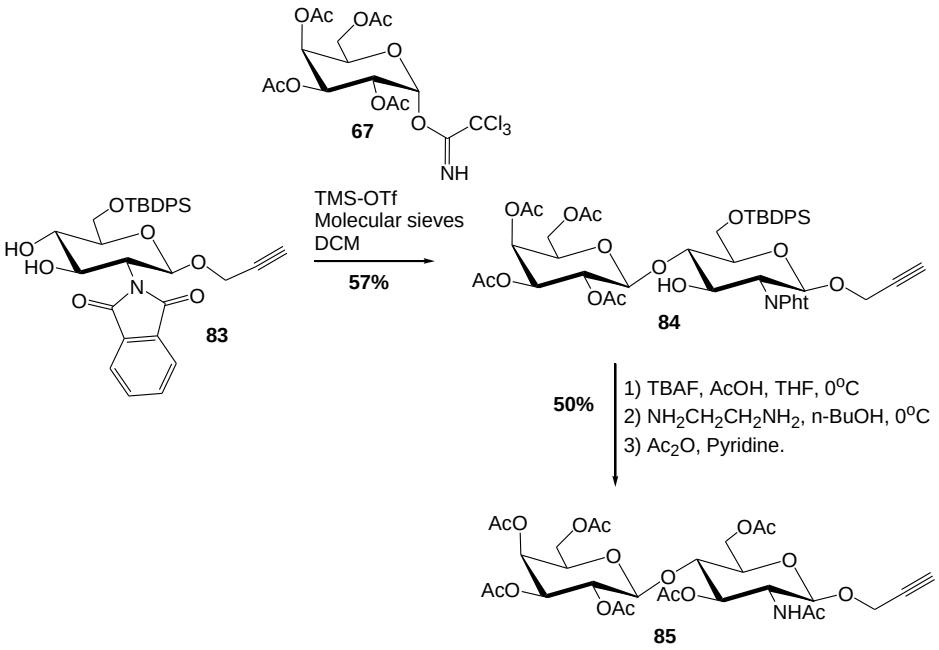


Figure 2.8 *The synthesis of peracetyl-propargyl LacNAc **85**.*

The product was confirmed by NMR and LC-MS, which gave a single peak with the mass of the desired product (**Figure 2.9**). Note that the end product is fully acetylated; this is common throughout the synthesis, and allows simple handling and purification of intermediates and products. Zèmplen deacetylation in sodium methoxide in methanol can conveniently free the hydroxyl groups.

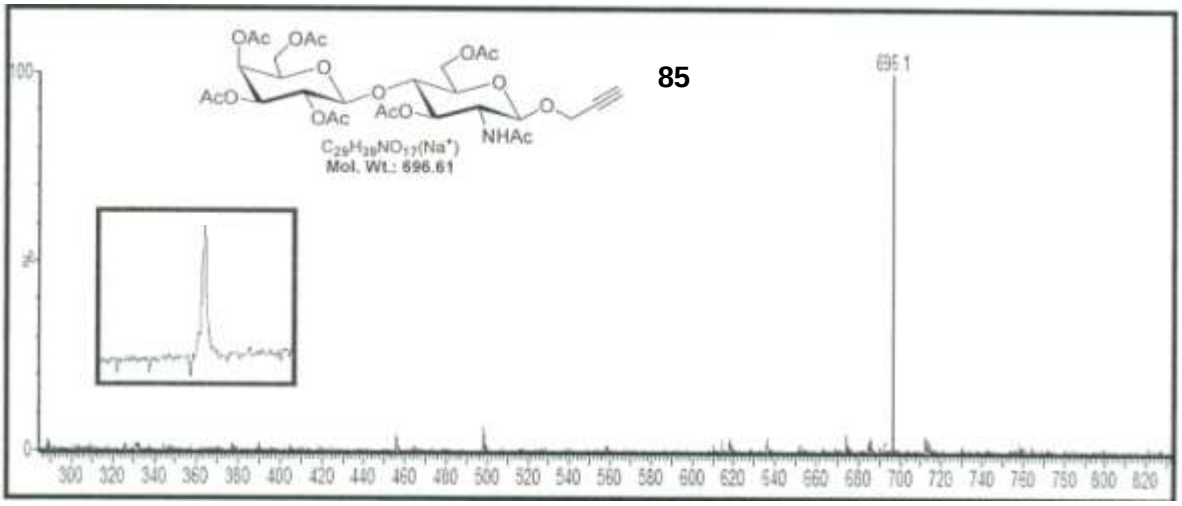


Figure 2.9 *LC-MS trace of peracetyl-propargyl N-acetyl-lactosamine **85**.*

2.3 The synthesis of propargyl *N*-acetyl-*D*-glucosamine

Propargyl *N*-acetyl-*D*-glucosamine (propargyl GlcNAc) was produced by starting with fully protected glycoside **71**, and directly installing the propargyl group with boron trifluoride diethyl etherate. The crude β-propargyl glycoside was deprotected with ethylenediamine and

re-protected with acetic anhydride and pyridine to produce the desired propargyl peracetylated *N*-Acetyl-*D*-glucosamine **86** (**Figure 2.10**).

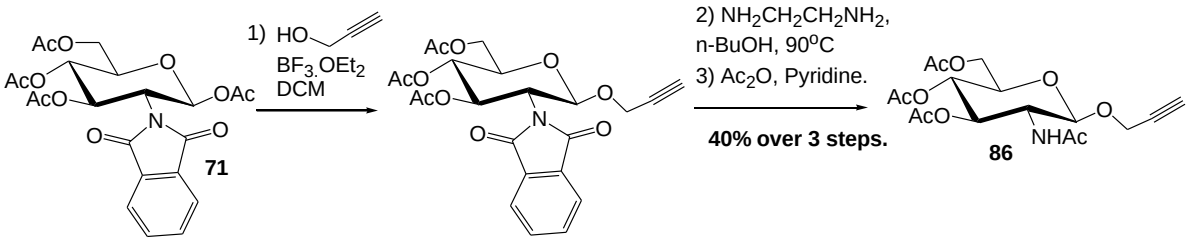


Figure 2.10 *Synthesis of propargyl GlcNAc [2,4].*

2.4 The synthesis of propargyl mannose

The synthesis of propargyl mannose was achieved by starting with α-*D*-mannose, the sugar was fully acetylated and then selectively deprotected at the anomeric position with benzyl amine. The resulting hemi-acetal **88** was activated to the trichloroacetimidate **89** which underwent glycosidation with neighbouring group participation, to yield fully protected α-propargyl mannose **90** (**Figure 2.11**).

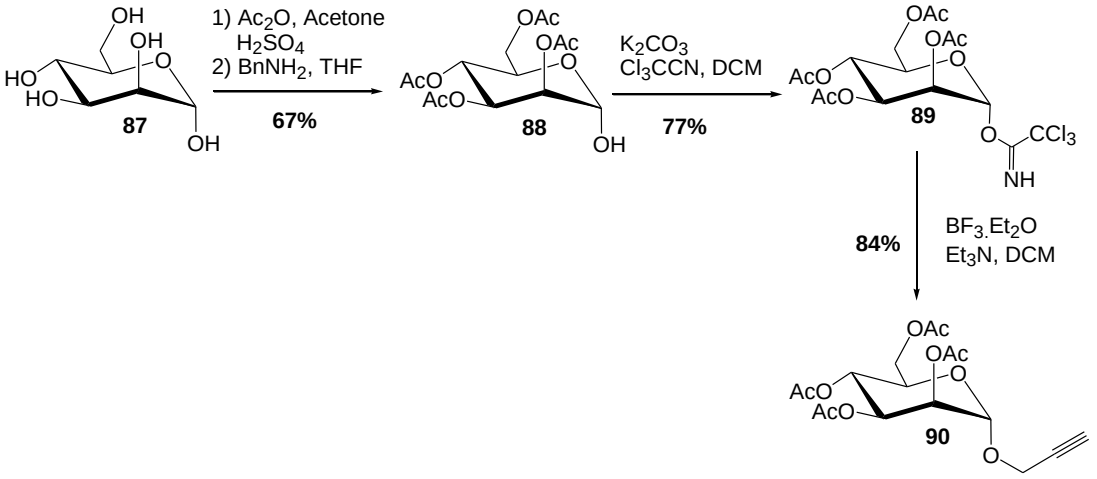


Figure 2.11 *Synthesis of propargyl mannose.*

2.5 Conclusion

The synthesis of three sugars; propargyl *N*-acetyl-*D*-lactosamine, propargyl *N*-acetyl-*D*-glucosamine and propargyl α-*D*-mannoside was completed. We have synthesised fully acetylated propargyl LacNAc by two different routes, both routes producing the novel glycosidic acceptor **83**. The route starting with the fully protected glycoside **71**, and directly installing the propargyl group, is by far the best route as it requires fewer synthetic steps. Trying to install the propargyl group on to the thiophenyl disaccharide became problematic due to the formation of the hemi-acetal and oxazolines. Propargyl *N*-acetyl-*D*-glucosamine

has been synthesised from glucosamine hydrochloride and propargyl mannose from α-*D*-mannose, both from readily available starting materials. These propargyl sugar derivatives will provide useful tools to incorporate into the azido-modified glycoproteins.

References:

[1] S. I. V. Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony, B. G. Davis, *Nature,* 2007, 446, 993.

[2] D. Macmillan, A. M. Daines, M. Bayrhuber, S. L. Flitsch, *Org. Lett*. 2002, 4, 1467.

[3] D. Macmillan, J. Blanc, *Org. Biomol. Chem*. 2006. 15, 2847.

[4] K. K. Yeoh, D. T. Butters, B. L. Wilkinson, A. J. Fairbanks, *Carbohydr. Res.* 2009, 334, 586.

[5] U. Ellervik, G. Magnusson, *J. Org. Chem*. 1998, 63, 9314.

[6] M. L. Phillips, E. Nudelman, F. C. A. Gaeta, M. Perez, A. K. Singhal, S. I. Hakomori, J. C. Paulson, *Science*, 1990, 250, 1130.

Chapter 3

The synthesis and metabolic incorporation of unnatural sugars into proteins

With the synthesis of propargyl glycosides established, the next step was to incorporate the azide functionality into our glycoprotein so that we could perform click chemistry with the propargylic glycosides. There were three methods investigated when incorporating the azide functionality into our glycoprotein. This chapter will focus on the *in vivo* metabolic incorporation, where azidosugars will be incorporated during expression of our glycoprotein in yeast.

Incorporating chemical handles in biological systems allows scientists to monitor and reveal novel biological functions. In particular, metabolic labelling of glycans enables their visualisation in cells and organisms as well as the enrichment of specific glycoprotein types for proteomic analysis [1]. However, it is important that the metabolic label does not interfere with or terminate the biological process that is being monitored. Azides and alkynes are good candidates for metabolic labelling, in that they are inert to biological modification but react specifically in chemical reactions with each other allowing them to be detected [2]. The detection of azides can be achieved by click chemistry (discussed in chapter 1) or by Staudinger ligation.

3.1 The Staudinger ligation

The Staudinger ligation (sometimes called the Bertozzi-Staudinger ligation) is an adaptation of the original Staudinger reaction discovered in 1919 by Staudinger and Meyer, who reported the formation of iminophosphoranes from azides and triaryl phosphines [3]. The reaction is conducted under mild conditions and is often quantitative. First, the phosphorus is believed to attack the azide to give a phosphazide; then, an intramolecular rearrangement *via* a four-membered ring transition state, yields an aza-ylide, with loss of nitrogen (**Figure 3.0**). The iminophosphorane obtained offers a wide variety of uses in organic synthesis: not only can the highly nucleophilic nitrogen potentially react with a range of electrophiles, but also it is an amine precursor [3].

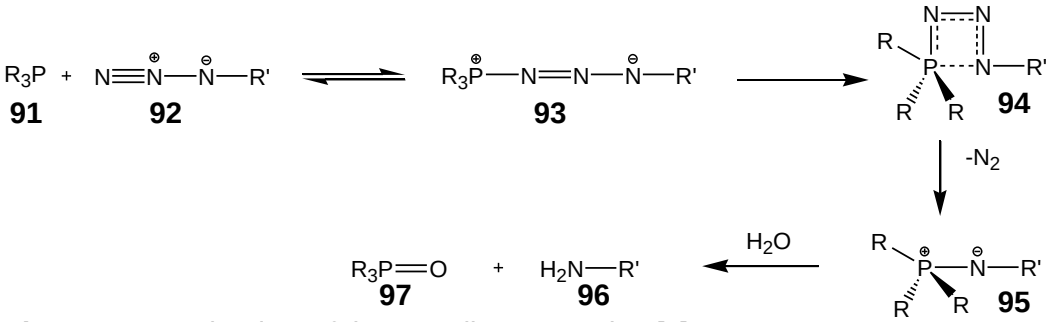


Figure 3.0 *Mechanism of the Staudinger reaction [3].*

In 2000, Saxon and Bertozzi [4] designed a variant of the original reaction (**Figure 3.1**). When an electrophillic trap, which is an ester, is introduced *ortho* to the phosphine, the aza-ylide intermediate is captured by intramolecular cyclisation [5]. This process ultimately produces a stable amide bond. Furthermore if the phosphine is directly attached to a biological probe, then detection of azides by Staudinger ligation can be monitored *via* the probe.

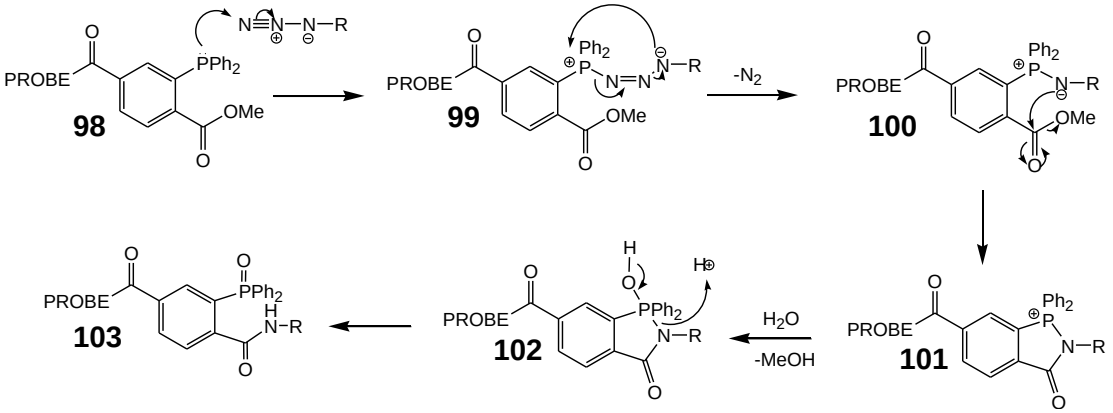


Figure 3.1 *Mechanism of the Bertozzi-Staudinger ligation [4 ,5]*.

One useful probe is the FLAG (**Figure 3.2**) which is selectively recognised by a specific antibody (*α*-FLAG) conjugated to horseradish peroxidase (HRP). The subsequent addition of the peroxidase substrate, together with a dye/chemoluminescent precursor, leads to a colour precipitate/light emission that serves for detection.

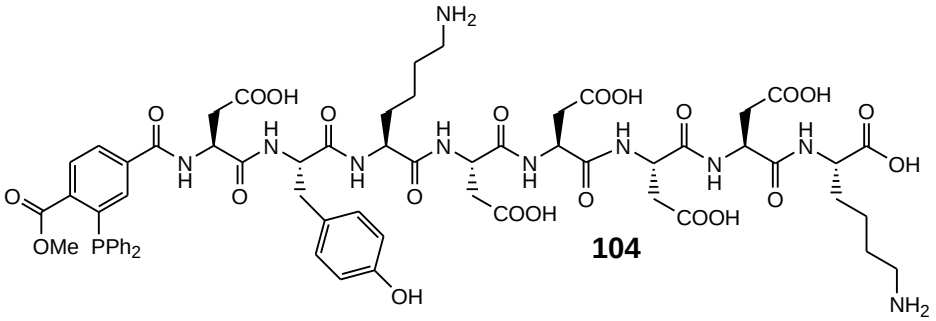


Figure 3.2 *FLAG peptide phosphine probe [4,5].*

Another probe which can be useful is a biotin tag. Biotinylation of a Staudinger probe will allow detection of the biotin with anti-biotin antibodies or avidin/streptavidin tagged detectors like HRP or fluorescent dyes. Protein biotinylation can be *via* amine, thiol or aldehyde groups with sulfo-NHS biotin**,** EZ-Link biotin-HPDP (*N*-(6-(Biotinamido)hexyl)-3'-(2'-pyridyldithio)-propionamide and biotin hydrazide repectively (**Figure 3.3**).

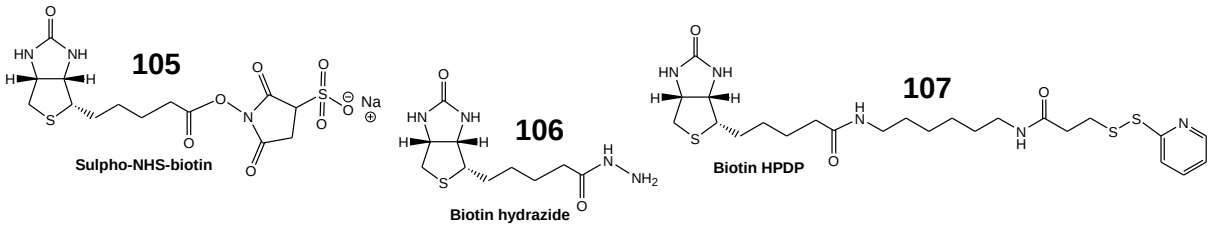


Figure 3.3 *Biotin structures [5].*

3.2 Metabolic labelling of glycans

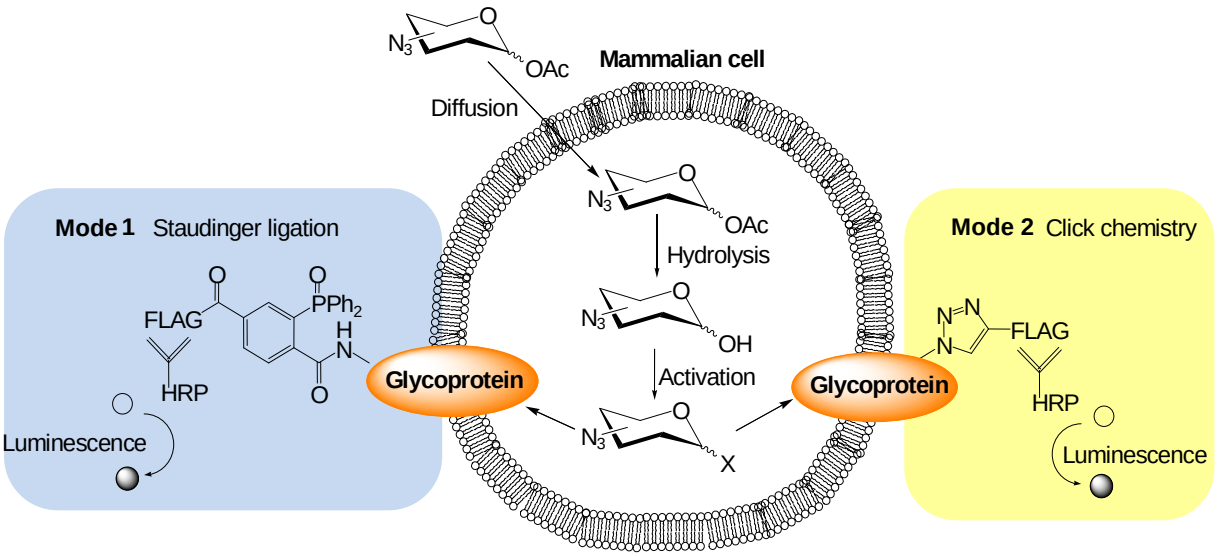


Figure 3.4 *Different methods for labelling glycans in mammalian cells. Both methods begin with incubation of mammalian cells with azido sugar. These are hydrolysed and incorporated* into *the glycoprotein where different modes of detection are used. Mode one uses Staudinger ligation while mode two uses click chemistry.*

Bertozzi and co-workers first devised a method that involves metabolic labelling of glycans with a monosaccharide precursor that is functionalised with an azide group (**Figure 3.4**) [1]. Bertozzi synthesised five azido sugars, respectively; *N*-azidoacetylgalactosamine (GalNAz), *N*-azidoacetyl sialic acid (SiaNAz), *N*-azidoacetylmannosamine (ManNAz), *N*-azidoacetylglucosamine (GlcNAz) and 6-azido fucose (6AzFuc). These sugar derivatives were incubated with mammalian cells and installed within cellular glycans by the biosynthetic machinery. With the successful incorporation of these azidosugars, they were poised to

react selectively with functionalised probes. Bertozzi showed that Staudinger ligation with phosphine-FLAG or phosphine-FLAG-His_6 probe were successful for the identification of the incorporated azidosugars. Furthermore the Staudinger ligation was believed to be a better option for *in vivo* glycan labelling as it avoided the cytotoxcity of copper, which is associated with click chemistry. The result indicated that metabolic labelling is highly dependent on the cell type used. For instance, COS-7 [**C**V-1 (simian) in **O**rigin, and carrying the **S**V40 genetic material] cells treated with SiaNAz and Ac_4ManNAz constituted 5% of the total sialic acid. However, the same treatment in *Jurkat* and *CHO* cells, showed 20-40% of the total sialylic acid content labelled with SiaNAz and Ac_4ManNAz [1].

As mentioned copper mediated click chemistry requires a toxic heavy metal catalyst, while the Staudinger ligation is limited by relatively slow kinetics but difluorinated cyclooctyne (DIFO) for ring strain-promoted cycloaddition, may overcome these problems. The rapid kinetics of DIFO to react with azides is demonstrated with the monitoring of cellular glycosylation of zebrafish embryos [7]. It has been found that the ring-strain promoted 1,3-dipolar cycloaddition of (DIFO) with azides (**Figure 3.5**) is a useful alternative to the Staudinger ligation [6].

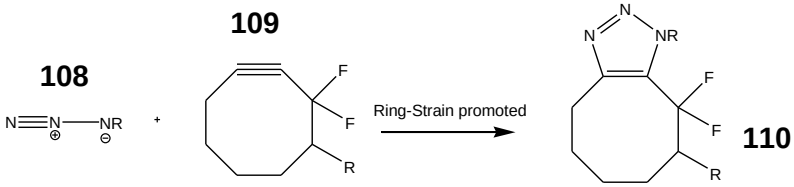


Figure 3.5 *Ring-strain promoted 1,3-dipolar cycloaddition of difluorinated cyclooctyne with azide[6].*

Wong and co-workers [7] used a slightly different approach to incorporate sugars in which alkynyl sugar analogues were used to label and visualise glycoconjugates in cancer cell lines. Alkynyl fucose and *N*-acetylmannosamine derivatives were synthesised and incorporated into fucosylated and sialylated glycans in several cancer cell lines [7]. Visualisation of the cell surface, intracellular and individual glycoproteins was achieved, with a biotin/fluroescein-conjugated streptavidin probe and on the cell surface with 3-azido-7-hydoxycoumarin probe (**Figure 3.6**).

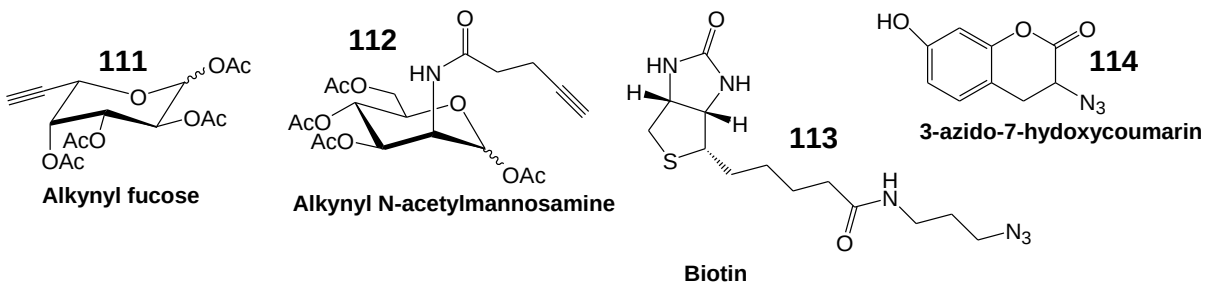


Figure 3.6 *Alkynl sugar derivatives and their visualisation tools [7].*

We have seen a few examples where azides are used in biological systems for labelling and tagging. Such systems can be used to monitor *in vivo* animal modelling and allow potential methodology of glycan remodelling [8,7]. For our project azidosugars will be fed to *Pichia pastoris* at the same time as EPO expression is induced in the hope that the azido sugars will be incorporated into the expressed glycoprotein.

3.3 *Results*

3.4 The use of unnatural sugars and feeding experiments

Prior to the metabolic incorporation of azidosugars, the synthesis of azidomannose must first be investigated. The synthesis of 6-azidomannose and 4-azidomannose will be performed and these sugars will be incubated with yeast expressing EPO. If the azidosugars have been incorporated then they will be detected with Staudinger ligation with a FLAG phosphine probe and itself be detected with anti-FLAG alkaline phosphatase and phosphatase substrate blue. If successful, the incorporated azido sugars can be used for further glycoprotein remodelling by click chemistry with synthetic propargyl glycosides following glycoprotein isolation.

Introducing azido sugars into yeast glycans by feeding azidosugars to yeast expressing EPO will not allow complete control in which the azido sugars will be incorporated. The position and the type of glycosidic linkage that the azido sugars may form will be dependent on the glycoprotein biosynthesis. This may result in different glycosylation patterns and different glycan structure. Therefore if the unnatural sugars are incorporated, then the glycoprotein mass should be different to the wild type glycoprotein. Also the metabolic incorporation of unnatural sugars will not control the homogeneity of the expressed glycoprotein but will be a good starting point to see if these azido sugars can be tolerated by the natural cell machinery.

In addition to synthesising 6- and 4-azidomannose, the synthesis of 6- and 4- thio mannose will also be undertaken. The thio sugars will also be incorporated into yeast expressed EPO and the glycosylation patterns will be examined. It is important to note that the Staudinger ligations cannot be performed on the thioacetate derivatives and we will only examine the glycosylation patterns by Western blot. If the thio sugars are incorporated it will potentially demonstrate another route to for glycoprotein remodelling using *N*-(propargyl)-bromoacetamide and azido sugars (**Figure 3.7**).

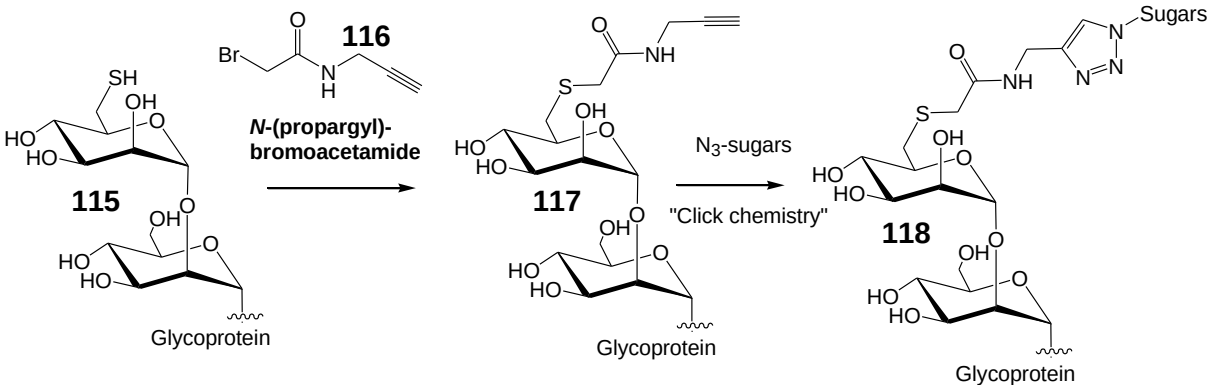


Figure 3.7 *The metabolic incorporation of the thiosugars will allow potential glycan remodelling with N-(propargyl)-bromoacetamide and azidosugars.*

All the unnatural sugar analogues will be synthesised in the fully acetylated form, as this modification is known to increase their cellular uptake [1, 3]. It is envisaged, that the acetate esters are subsequently hydrolysed in the cytosol to give the free hydroxyl sugars and the free thiol [1].

3.5 Synthesis of peracetyl-6-azidomannose 122.

The preparation of the first azidosugar bearing an azide group at C-6 was based on the protocol described by Bertozzi [3]. Initially, *p*-toluenesulphonyl chloride selectively reacted with the primary alcohol function of α-methyl-*D*-mannopyranoside to yield **120** (**Figure 3.8**). Next **120** was treated with sodium azide in DMF, in order to substitute the tosyl group at carbon 6 leading to the formation of the azide **121**. The remaining hydroxyl groups were then protected as acetate esters by a mixture of acetic anhydride and acetic acid in the presence of sulphuric acid as a catalyst. In the same reaction, an acetate group substituted the methyl glycoside at the anomeric position. The fully acetylated 6-azido-6-deoxy-α-*D*-mannopyranose **122** was obtained in 72% yield over the two steps.

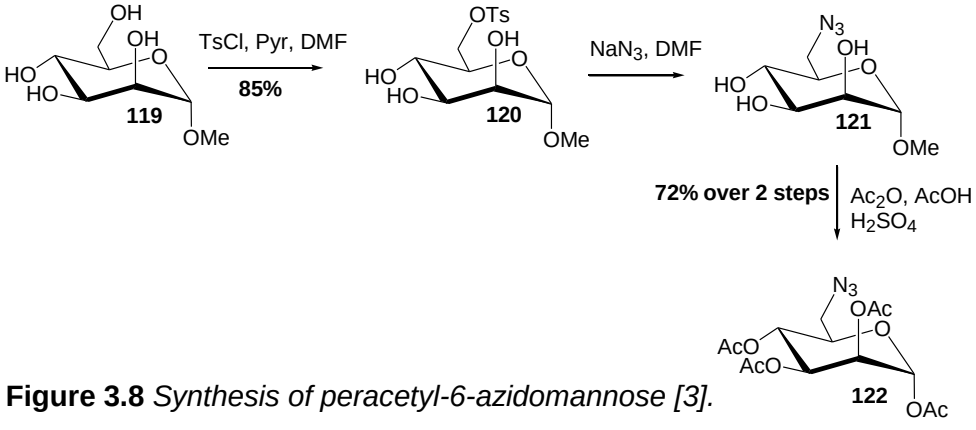


Figure 3.8 *Synthesis of peracetyl-6-azidomannose [3].*

Fully acetylated 6-thioacetate-6-deoxy-α-*D*-mannopyranose (**124**) was produced in a similar manner to that of the azide derivative. Although thio-glycosides are commonly used to activate the anomeric position in carbohydrate chemistry, there is little research on the replacement of hydroxyl groups by thioacetates, primarily due to thioacetates being very reactive electrophiles [9]. However, in the published synthesis of 3- and 4-thio analogues of 4-nitrophenyl 2-acetamido-2-deoxy-*β*-*D*-gluco- and galactopyranosides [10], thioacetates where used. The key step involved activation of the 4-hydroxyl with a triflate and displacement using potassium thioacetate. We will adopt this approach to install the thioaceate to the mannose derivatives. To introduce the thioacetate the aprotic solvent hexamethylphosphoramide (HMPA) was used. After successful introduction of the thioacetate at C-6, treatment with acetic anhydride and acetic acid in the presence of sulphuric acid produced the fully acetylated 6-thioacetate mannose **124** (**Figure 3.9**).

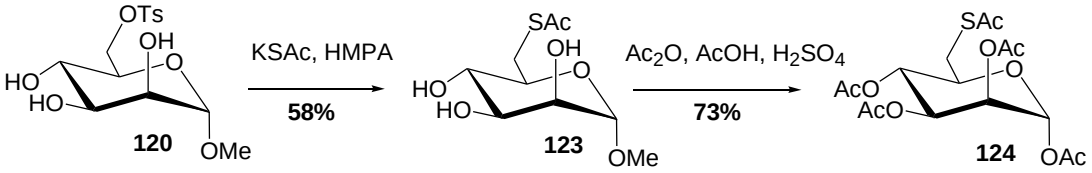


Figure 3.9 *Synthesis of peracetyl-6-thioacetate-mannose.*

3.6 Synthesis of peracetyl-4-azidomannose 32.

The synthesis of peracetylated 4-azidomannose was achieved from the cheap and readily available methyl-α-mannoside in seven steps (**Figure 3.10**): mannose was firstly protected by standard procedures to **125**, leaving access to the hydroxyl at carbon-4; PCC oxidation, followed by stereoselective $NaBH_4$ reduction inverted the configuration at C-4 to **126** in 90% yield. Triflation and substitution with sodium azide gave access to the fully protected 4-azido-4-deoxymannose **127**. Finally, subsequent stepwise deprotection and acetyl protection led to the desired azido monosaccharide **128** in 76% yield (**Figure 3.10**).

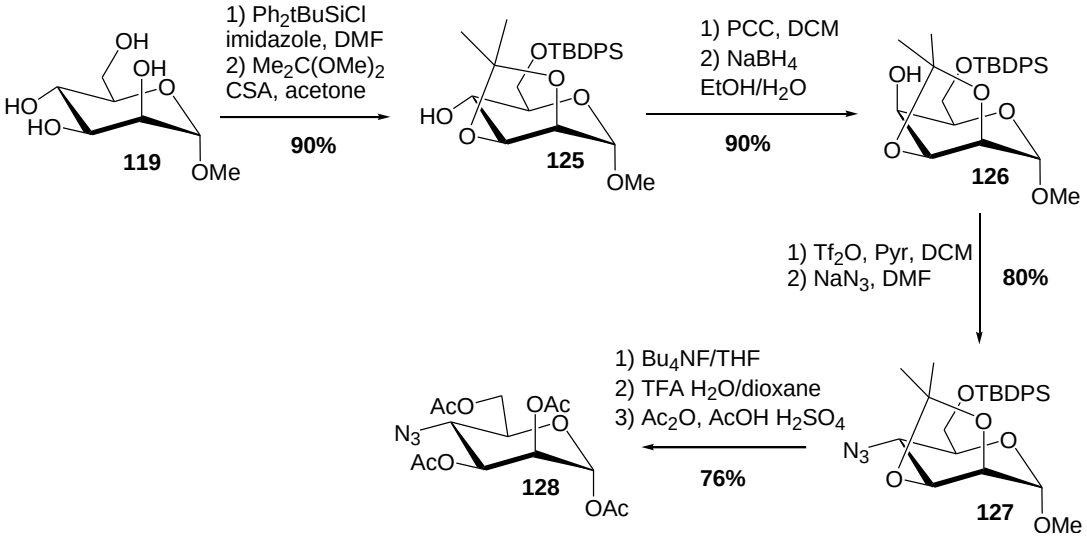


Figure 3.10 *Synthetic route to peracetyl-4-azidomannose [11].*

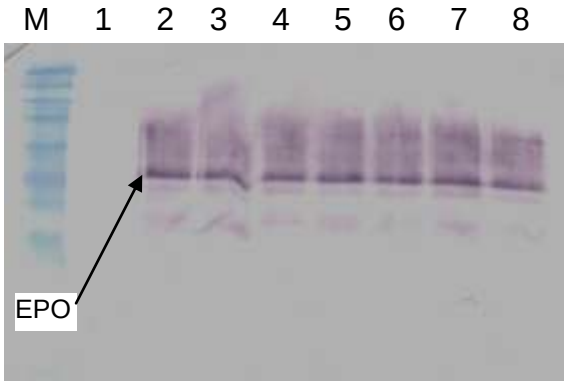
When we tried to reproduce the published method [11], we were pleased by the simplicity of the reaction procedures which indeed, when dry conditions were maintained, could be nicely followed by TLC to nearly quantitative yield in most cases. The only relatively low yielding step (for which Davis reported 64% yield), was the nucleophilic substitution by sodium azide. This was improved to 80% yield over two steps. Furthermore, we found that the deprotection steps with TBAF and 1,4-dioxane and subsequent protection with acetic anhydride and concentrated H_2SO_4 can all be performed without purification. This vastly improved the overall yield of the desired peracetylated 4-azidomannose. With the synthesis of the peracetylated 4-azidomannose completed, we repeated the same route for the synthesis of the thioacetate derivative. However, introducing the thioacetate at carbon-4 was problematic. Initially trying HMPA as the solvent and then changing to DMF both was unsucessful, as there was no sign of the desired product.

With the synthesis of 4-, 6- azidomannose, and 6-thioacetate-mannose completed, we decided to perform the feeding experiments to see if any of these sugars would be incorporated into *Pichia* expressed EPO.

3.7 Feeding experiments on *Pichia* expressed EPO.

Wild type *Pichia pastoris* yeast was genetically engineered to express and secrete human erythropoietin following addition of methanol to the culture medium (*Pichia* expressing strain donated by Dr Jonathan Richardson). After inducing EPO expression, cultures were supplemented with either 4- or 6-azidomannose or 6-thioacetate-mannnose and further incubated to allow uptake and assimilation of unnatural sugars. After 48 hours, yeast were removed from the growth medium by centrifugation. The growth medium was then

concentrated and analysed by Westen blot using Anti-EPO monoclonal antibodies (**Figure 3.11**). The results indicated that the presence of the azido- and thioacetate-sugars does not prevent EPO from being expressed or secreted. However, it appears that none of the unnatural sugars were been incorporated into the glycosylation that adorns the secreted EPO molecule. This was initially evident as there were no gross alterations in the banding pattern when compared with the cells cultured in the absence of the unnatural sugar. One reason why this may be that only a single unnatural sugar may had been incorporated and the sugar chain extended. Then perhaps Western blotting alone would not reveal the presence of a single azido group. If an azido group had been incorporated it could be detected by using the Staudinger ligation with a FLAG phosphine probe.



Key:
M: Pre –stained molecular weight markers
1: Empty vector (-ve control)
2: *Picha* induced with methanol
3: *Picha* + 4-azidomannose (50 mM)
4: *Picha* + 4-azidomannose (100 mM)
5: *Picha* + 6-azidomannose (50 mM)
6: *Picha* + 6-azidomannose (100 mM)
7: *Picha* + 6-thioacetate-mannose (50 mM)
8: *Picha* + 6-thioacetate-mannose (100 mM)

Figure 3.11 *Westen blot of EPO-expressing Pichia in the presence of different concentrations of unnatural sugars.*

The Staudinger ligation was performed on the secreted EPO samples that were incubated with 4- and 6-azidomannose. The membrane was probed with 250 mM FLAG-phosphine in phosphate buffered saline (PBS) for 3 hours and then with anti-FLAG alkaline phosphatase conjugate, alkaline phosphatase and phosphatase substrate blue (**Figure 3.12**).

There was no signal from the Staudinger reaction product indicating that none of the azido sugars were incorporated. Further work will need to be carried out by repeating the feeding assays, together with a control for the antibody specificity. However, these preliminary results suggest that the transfer of these azidomannose derivatives were unsuccessful.

M 1 2 3 4 5

Key:
M: Markers (pre-stained molecular weight markers)
1 *Pichia* induced with methanol
2: *Pichia* + 4-azidomannose (50 mM)
3: *Pichia* + 4-azidomannose (100 mM)
4: *Pichia* + 6-azidomannose (50 mM)
5: *Pichia* + 6-azidomannose (100 mM)

Figure 3.12 *Western blot of expressed EPO in the presence of azidomannose. To examine the incorporation of azidosugars the membrane was treated* with FLAG-phosphine*, anti-FLAG, alkaline phosphatase and phosphatase substrate blue.*

3.8 Conclusion.

One explanation why the azidosugars may have not been incorporated into the expressed protein might be the position of the modification. Directly replacing a hydroxyl group with a chemical precursor such as an azide or thioacetate on a sugar, then feeding to yeast cells has not to date been investigated. In fact as we have seen, analogues like ManNAz, where the azide functionality is not directly replacing a hydroxyl group, are the only existing molecules that have been successfully incorporated metabolically. Ultimately the presence of free hydroxyl groups in mannose might be important for hydrogen bonding with the active site of the mannosyltransferase enzyme. Directly replacing the hydroxyl with an azide may prevent these intermolecular interactions from occurring and hence disable azidomannose transfer. Furthermore the unnatural mannose may not be able to compete with endogenous mannose production therefore *in vitro* methods will need to be investigated.

References:

[1] J. A. Prescher, C. R. Bertozzi, *Nat. Chem. Biol*. 2005, 1, 13.

[2] H .C. Hang. C. Yu, M. R. Pratt, C. R. Bertozzi, *J. Am. Chem. Soc*. 2004, 126, 6.

[3] H. Staudinger, J. Meyer, *Helv. Chim. Acta*. 1919, 2, 635.

[4] E. Saxon, C R. Bertozzi, *Science,* 2000, 287, 2007.

[5] F. L. Lin, H. M. Hoyt, H. V. Halbeek, R. G. Bergman, C. R. Bertozzi, *J. Am. Chem. Soc*. 2005, 127, 2686.

[6] N. J. Agard, C. R. Bertozzi, *Acc. Chem. Res*. 2009, 6, 788.

[7] T. L. Hsu, S. R. Hanson, K. Kishikawa, S. K Wang, M. Sawa, C. H. Wong, *Proc. Natl. Acad. Sci. USA*. 2007, 104, 2614.

[8] F. Liu, A. J. Aubry, I. C. Schoenhofen, S. M. Logan, M. E. Tanner, *Chembiochem*. 2009, 10. 1317.

[9] R. T. Dere, X. Zhu, *Org. Lett*. 2008, 10, 4641.

[10] H. M. Chen, S. G. Withers, *Carbohydr. Res*. 2007, 342, 2212.

[11] B. G. Davis, R. J. Nash, A. A. Watson, C. Smith, G. W. J. Fleet, *Tetrahedron Lett*. 1999, 55, 4501.

Chapter 4

Enzymatic assays to test the transfer of mannose and azidomannose

With the unsuccessful incorporation of azidosugars by *in vivo* metabolic pathway, we next examined *in vitro* test conditions. The transfer of azidomannose from GDP-azidomannose to different acceptor substrates was investigated. This involved some proof of principle experiments, initially we sought to optimise the transfer of mannose from GDP-mannose to different substrates. Then we changed the donor to GDP-azidomannose and examined the transfer of azidomannose to the substrates. The synthesis of GDP-azidomannose analogues has already been established in our group [1]. However, the transfer of the azidomannose has not been investigated in detail. We will explore this further by trying to quantify the amount of α-1,2 mannosyltransferase required and develop a method for isolating the products.

4.1 α-1,2-mannosyltransferase

The commercial availability of mannosyltransferases (ManTs) is very limiting. Although there are many different types of ManTs, only one member of the family is commercially available this is the soluble fragment of Kre2p/Mnt1p, an *α*-1,2 mannosyltransferase from yeast *S. cerevisiae*. Wong and co-workers have shown that the use of the plasmid pManFLAG20 for protein expression in *E. coli*, will encode for the soluble fragment (amino acid 31-442) of Kre2p/Mnt1p α-1,2-mannosyltransferase [2, 3]. The pManFLAG expression vector has two N-terminal tags, an OmpA (outer membrane protein A) tag which is a peptide that directs the recombinant peptide to the periplasmic space. This is beneficial for purification, as very few proteins are present in this cellular compartment. The other tag is the immunogenic FLAG peptide which can be used for identification and purification of the recombinant protein. The soluble fragment 31-442 of Kre2p/Mnt1p pManFLAG20, shows catalytic activity, but does not encode for the full enzyme leaving out its transmembrane domain. Kre2p/Mnt1p is found in the golgi of *S. cerevisiae*, where it is used to sequentially add mannose units to the growing oligosaccharide chain. The natural substrates for Kre2p/Mnt1p have been unambiguously defined, as *in vivo* studies show that it is responsible for the addition of mannose to serine or thereonine-linked mannobiose units [3], while *in vitro* examples show that the *N*-linked glycans ($Man_{15\text{-}30}$-GlcNAc) are natural substrates. Another example indicates *P. pastoris* expressed α-1,2,-ManT fragment, can transfer mannose from GDP-mannose to a variety of substrates including methyl-α-1,2-mannobiose and methyl-α-1,-

mannoside [4]. Wong and co-workers showed that the fragment Kre2p/Mnt1p expressed in *E. coli* was able to transfer mannose to a peptide linked mannose substrate but *para*-nitro phenyl substrate showed no relative activity (**Figure 4.0**)

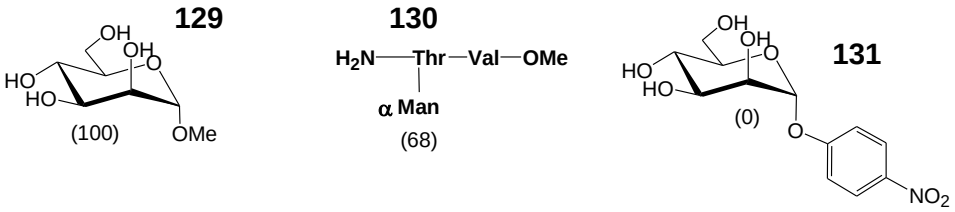


Figure 4.0 *Unnatural acceptors tested on the soluble fragment of Kre2p/Mnt1p expressed in E. coli and the relative activities are indicated in brackets [3, 5].*

4.2 The crystal structure of α- 1,2-mannosyltransferase (Kre2p/Mnt1p)

The crystal structure of Kre2p/Mnt1p is shown bellow (**Figure 4.1**). Its structure has two protein domains, each with an *N*-linked glycan attached. There are three disulphide bonds, a mixture of α/β folds, one Rossmann fold which is common to class-A glucosyltransferases. The cofactor magnesium is coordinated by a modified DXD motif and molecular modelling studies of glycosyltransferases suggest that the mannose donor is within distance for Van der Waals interactions with Ala 362, Trp 325, Met 223, and Cys 224. Hydrogen bonding is suggested with the oxygen of at Ser 326, while the side-chain amine of Arg 245, the carboxylic oxygen of Asp 361 nitrogen of Asn 327, all form a binding pocket in the active site of the enzyme. However it is only the Asp 361 which has been shown critical for activity by mutagensis [6]. There is little known about the interactions between the mannose donor and the enzyme.

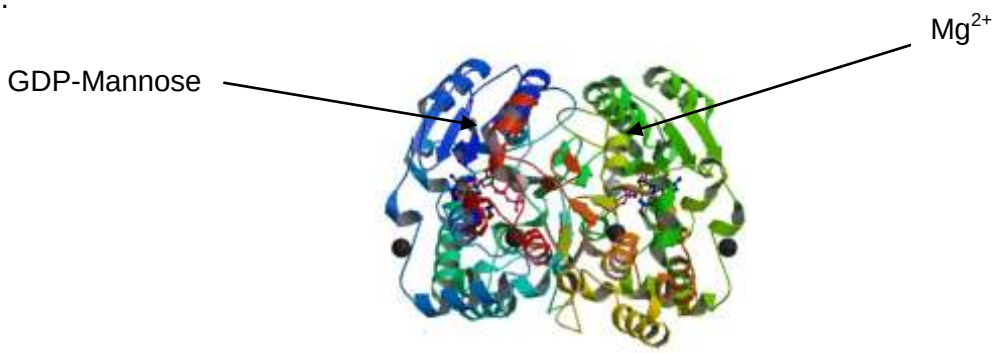


Figure 4.1 *Crystal structure of Kre2p/Mnt1p, displaying magnesium, and GDP-mannose in the catalytic site (PDB codes 1S4N, 1S4O, 1S4P) [6].*

4.3 Modified nucleotide-sugars as glycosyl donors

Modification of nucleotide-sugars are less commonly studied compared to the acceptor molecules. There is an instance where two thiopyranoside analogues of GDP-sugars were

synthesised and tested for transfer [7]. The synthetic route to GDP-5-thio-*D*-mannose and GDP-5-thio-*L*-fucose included the phosphorylation of tetra-*O*-acetyl-5-thio-*D*-mannosyl bromide and tri-*O*-benzoyl-*L*-fucosyl bromide with silver dibenzyl phosphate. Then deprotection of the phosphate groups, and condensation of the deprotected phosphates with GMP-imidazolidate in the presence of $MgCl_2$. These GDP-sugar analogues were found to be donor substrates for α-1,2-mannosyltransferase and α-1,3-fucosyltransferase, affording a 5-thiomannose containing disaccharide (**Figure 4.2**) and a 5-thiofucose-containing trisaccharide [7].

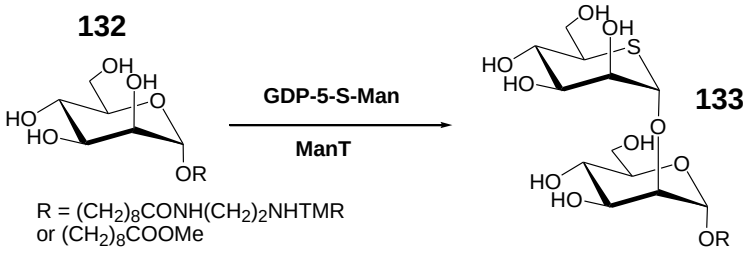


Figure 4.2 *Unnatural GDP-5-S-mannose as a glycosyl donor [7].*

The reaction of the tetramethylrhodamine (TMR)-labelled glycosyl acceptor was analysed by capillary electrophoresis and MALDI-TOF MS to reveal the corresponding product of the enzymatic 5-S-saccharide transfer [7]. The conditions for the transfer were not especially practical since the reactions were left for 44 days and high quantities of enzyme (280 mU) were used. This raises questions as to the efficiency and cost effectiveness of the reaction and furthermore illustrates the difficulty of trying to transfer unnatural sugars onto glycosyl acceptors.

4.4 *Results*

4.5 Initial transfer assays and testing

GDP-mannose was initially used as a positive control for the transfer of mannose unit to its known acceptor; methyl-α-*D*-mannoside. The conditions employed were those developed by our group previously (S. Marchesan personal communication). However, the only difference was that the α-1,2-mannosyltransferase was expressed and purified previously whereas this study used the commercially available enzyme. Using the commercial enzyme would allow quantification (in millunits) of how much enzyme was added in each reaction. The assay consisted of ammonium formate buffer (50 mM) at pH 7.5, manganese chloride (10 mM), GDP-mannose (1 mM), methyl-α-*D*-mannoside (OMe) (10 mM) and various quantities of α-1,2-mannosyltransferase (ManT). The assays were conducted for 24 hours at 30 ^{o}C, and then analysed by mass spectrometry. Unfortunately, there was no sign of the desired

product by ESI-MS, only that of the starting materials were shown. Changing the ratio of GDP-mannose and methyl-α-D-mannoside from 1:1 and 2:1, still indicated no formation of products (**Table 1**).

Entry	*[GDP-man]/mM*	*[Manα-OMe] /mM*	*ManT/mU*	*% Yield*
1	***1***	***10***	***5***	***0***
2	***1***	***10***	***10***	***0***
3	***1***	***1***	***10***	***0***
4	***2***	***1***	***10***	***0***

Table 1 *Transfer assay of mannose from GDP-mannose to methyl-α-D-mannoside. All reactions conducted in ammonium formate buffer (50 mM) at pH 7.5, with $MnCl_2$ (10 mM).*

Further literature searching revealed procedures where Triton X-100 had been added to assays as a source of detergent. Given these findings the assay was repeated in the presence of 0.1% v/v Triton X-100. A mini-extraction in an eppendorf tube was performed, with water and ethyl acetate to remove the Triton X-100. The extraction procedure was required otherwise Triton X-100 dominates the MS analysis. However, small peaks of the desired product were now visible. In attempting to further optimise this assay we next investigated different buffers, and the results showed that both ammonium formate and (4-(2-hydroxyethyl)-1-piperazineethanesulphonic acid (HEPES) buffers yielded encouraging results (**Table 2**).

Entry	*[GDP-man]/mM*	*[Manα-OMe] /mM*	*ManT/mU*	*Buffer/50 mM pH 7.5*	*% Yield*
5	***10***	***100***	***3***	***Tris buffer***	***0***
6	***10***	***100***	***3***	***Ammonium formate***	***5***
7	***10***	***100***	***3***	***HEPES***	***2***
8	***10***	***100***	***3***	***Ammonium Acetate***	***0***

Table 2 *Transfer assays in different buffers. All reactions conducted at pH 7.5 with $MnCl_2$ (10 mM) and Triton X-100 (0.1%). Yields estimated by relative intensity of MS peak.*

Establishing that some transfer of mannose was achieved with these assays, we next investigated another substrate. We sought a substrate that with a UV chromophore so the reaction could be monitored and the product isolated by HPLC. To this end, benzyl-α-mannopyranoside (OBn) was used instead of methyl-α-*D*-mannoside. Repeating the assays with the new acceptor in both HEPES and ammonium formate buffers, then extracting with DCM prior MS analysis, give no sign of the desired product. Furthermore, increasing the amount of enzyme from 3 mU to 18.3 mU showed no improvement in the results (**Table 3**).

Entry	*[GDP-man]/mM*	*[Manα-OBn] /mM*	*ManT/mU*	*Buffer/50 mM pH 7.2*	*% Yield*
6	***10***	***100***	***3***	***Ammonium acetate***	***0***
7	***10***	***100***	***3***	***HEPES***	***0***
8	***10***	***100***	***3***	***Ammonium formate***	***0***
9	***10***	***100***	***18.3***	***Ammonium formate***	***0***
10	***10***	***100***	***18.3***	***HEPES***	***0***

Table 3 *Transfer assays with benzyl-α-mannopyranoside as the acceptor. All reactions conducted at pH 7.2 with $MnCl_2$ (10 mM) and Triton X-100 (0.1%).*

Extensive literature searches revealed instances where a cocktail of phosphatase inhibitors were used to prevent breakdown of GDP-sugars [8]. We first decided to test the addition of phosphatase inhibitor (P5726, Sigma) with the methyl-α-*D*-mannoside substrate as we had already observed successful transfer previously. We found that the addition of phosphatase inhibitors greatly improved the rate of formation (**Table 4**) of the disaccharide product as observed by MS. While the MS analysis is not quantitative, there was an obvious improvement in the ionisation of the desired product. This can be seen by comparing the mass spectrum of the assay, with- and without phosphatase inhibitors (**Figure 4.3** and **4.4**). Note that a 10 fold excess of methyl-α-*D*-mannoside was used so this would be the predominant peak. Furthermore, looking at the reaction in ES –ve mode, it was apparent that

the GDP-mannose was abundantly present and did not degrade with the addition of phosphatase inhibitors.

Entry	*[GDP-man]/mM*	*[Manα-OMe] /mM*	*ManT/mU*	*Phosphatase inhibitors (4 µl)*	*% Yield*
11	***10***	***100***	***18.3***	**-**	***12***
12	***10***	***100***	***18.3***	**+**	***35***

Table 4 *Transfer assay conditions with- and without phosphatase inhibitors with methyl-α-D-mannoside as the acceptor. All reactions conducted in ammonium formate (50 mM) at pH 7.2 with $MnCl_2$ (10 mM) and Triton X-100 (0.1%).*

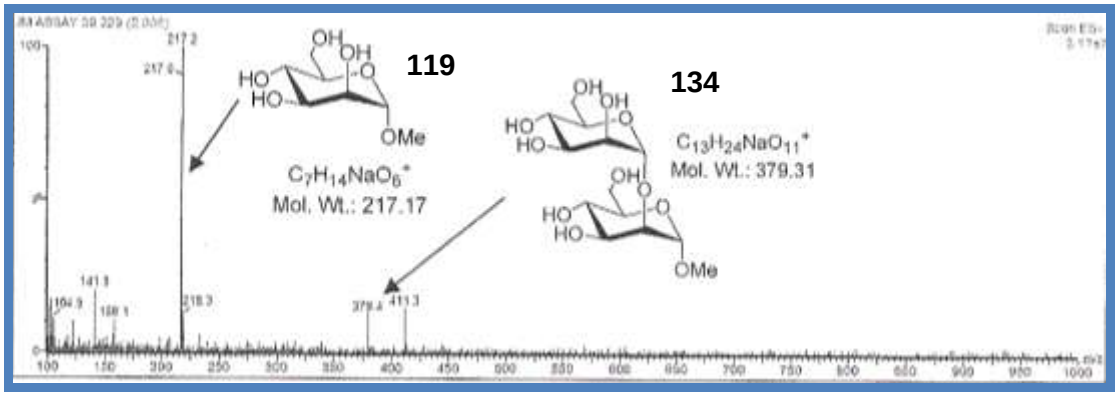


Figure 4.3 *MS of transfer assay without phosphatase inhibitors after 24 hours.*

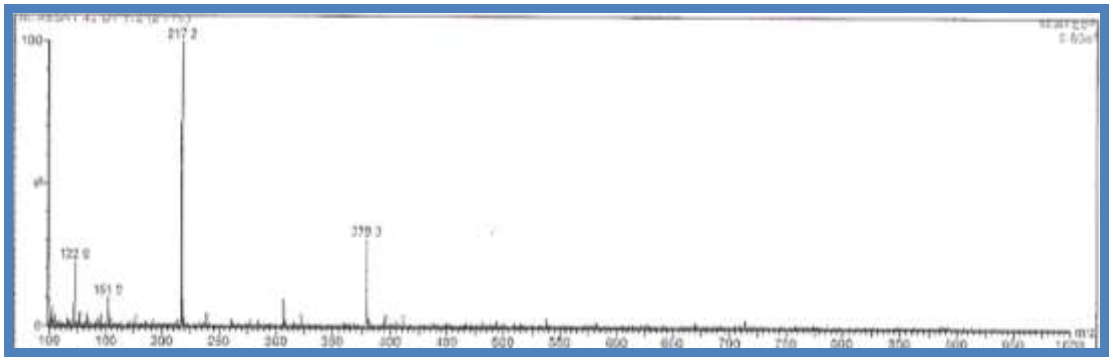


Figure 4.4 *MS of transfer assay with phosphatase inhibitors after 24 hours.*

Establishing that there was some improvement with the assay we decided to perform the assay with benzyl-α-mannopyranoside acceptor (**Table 5**). Again, a successful transfer was observed as the desired product was produced ~15 minutes after addition of the reactants (**Figure 4.5**).

Entry	*[GDP-man]/mM*	*[Manα-OBn] /mM*	*ManT/mU*	*Phosphatase inhibitors (4 µl)*	*% Yield*
13	***10***	***100***	***18.3***	**+**	***40***

Table 5 *Assay with benzyl-α-mannopyranoside with phosphatase inhibitors. The reaction was conducted in ammonium formate (50 mM) at pH 7.2 with $MnCl_2$ (10 mM) and Triton X-100 (0.1%).*

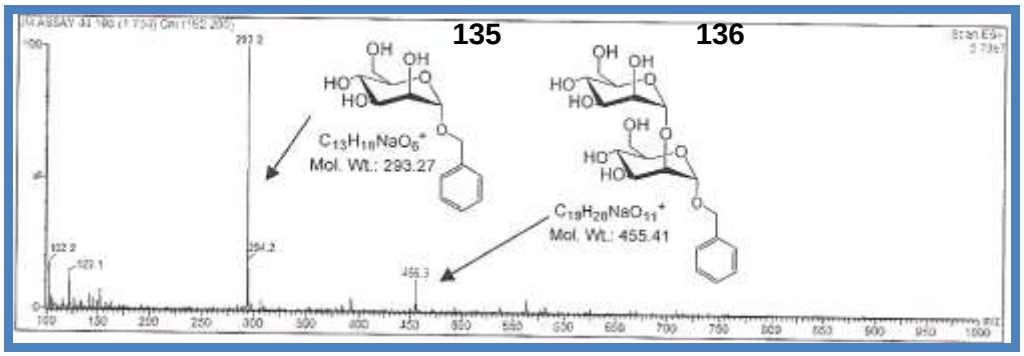


Figure 4.5 *MS of the transfer of mannose from GDP-mannose to benzyl-α-mannopyranoside ~15 minutes after the addition of the reactant.*

Having demonstrated a successful transfer of mannose to benzyl-α-mannopyranoside the next step was to isolate the product in good yield for structural analysis. We increased the amount of donor; GDP-mannose and the acceptor; benzyl-α-mannopyranoside by 10 fold. The enzyme α-1,2-mannosyltransferase was increased by 5-fold and the reaction left to proceed for 48 hours at 30 ^{o}C, followed by a mini-extraction in an Eppendorf tube, with DCM. The aqueous phase was collected and purified by reverse phase HPLC using a C18 column at 245nm, 5-95% (water/acetonitrile) for 30 minutes (**Figure 4.6**). The purified disaccharide **136** was lyophilised to afford a white foam (95% yield), which could be fully characterised using standard techniques such as MS (**Figure 4.7**), ^{1}H and ^{13}C NMR spectroscopy.

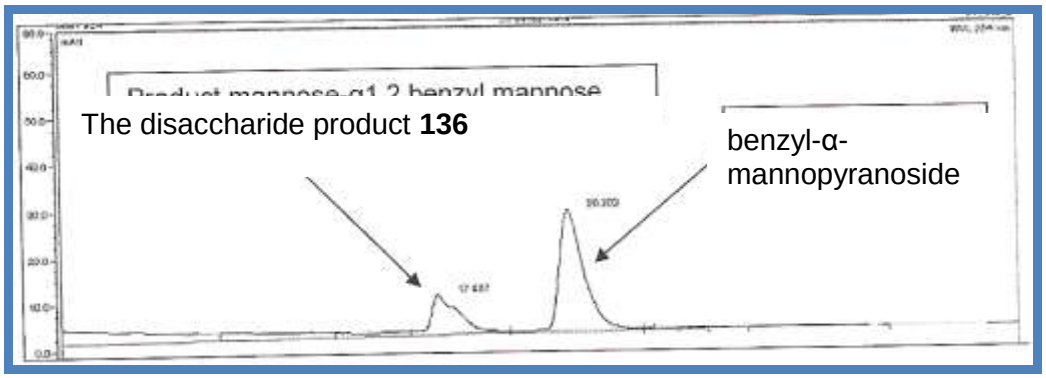


Figure 4.6 *LC trace of the reaction mixture containing the disaccharide **136**.*

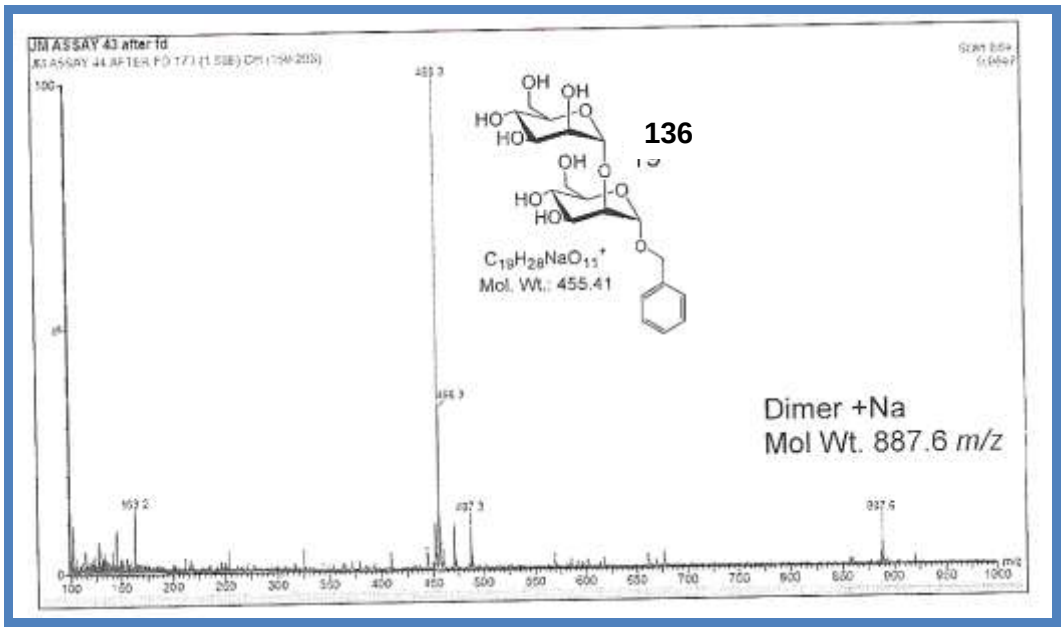


Figure 4.7 *MS of the purified disaccharide **136**.*

With a quantifiable transfer assay in place, the next step was to test the azido transfer from GDP-azidomannose derivatives to the benzyl acceptor. The GDP-azidomannose derivatives were donated by the previous student working on this project. We used the same successful transfer assay (entry 13) but replaced the GDP-mannose with GDP-azidomannose (**Figure 4.8**).

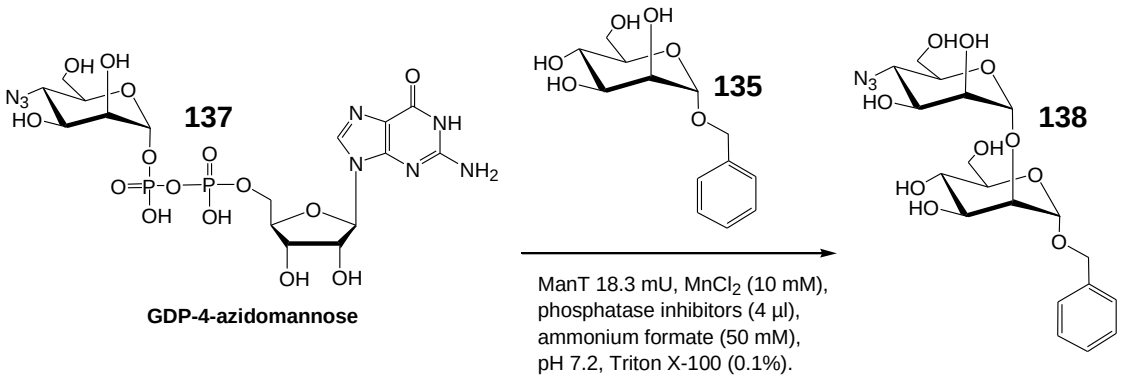


Figure 4.8 *Assay of GDP-4-azidomannose with benzyl mannose acceptor.*

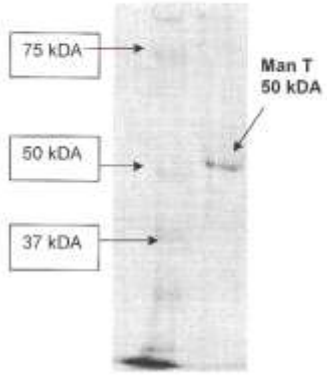
The transfer of azido sugars onto the benzyl substrate was not observed. All four different GDP-azidomannose derivatives were tested; GDP 2-azido-, 3-azido-, 4-azido- and 6-azidomannose. None were successful, and the control of GDP-mannose showed a positive result. To further check that the assay was working we added GDP-mannose in the same assay that contained GDP-azidomannose and again observed the successful transfer of mannose to benzyl mannose.

These results suggest that the transfer of azidosugars was unsuccessful, although intriguingly the transfer of 4-azidomannose on to methyl-α-*D*-mannoside was observed previously (S. Marchesan personal communication). Previously the enzyme α-1,2-mannosyltransferase was expressed and tested for transfer while our assays used the commercially available α-1,2-mannosyltransferase. Perhaps expressing and isolating the α-1,2-mannosyltransferase may give a better indication whether there is a difference between the expressed and commercial enzyme.

4.6 Expression of soluble Kre2p/Mnt1p in *E. coli*

Wong and co-workers reported the successful expression of the recombinant soluble Kre2p/Mnt1p in *E. coli* by induction with low-levels of isopropyl *β-D*-1-thiogalactopyranoside IPTG (5 µM), and use of minimum medium enriched with 2% casaminoacids (M9-CA), at 30 ^{o}C for 12 hours. The enzymatic tests were carried out on a partially purified protein batch, obtained after isolation of the periplasmic fraction by osmotic shock [2, 5].

Soluble Man T was produced previously by our group (S. Marchesan personal communication) using a modified version to Wongs original procedure in which glycerol was used as a carbon source, and magnesium as a co-factor. Interestingly, and in contrast to the findings published by Wong [2, 5], the expressed Man T enzyme was not found in the periplasmic space of the cell, but appeared to be located in the cytoplasm.



However, a continued analysis of Man T expression in this study demonstrated that Man T was present in the periplasmic space which agrees with Wongs original method and this was confimed by SDS PAGE analysis (**Figure 4.9**)

Figure 4.9 *Analysis of the periplasmic fraction by SDS PAGE.*

The purified enzyme was observed to transfer mannose from GDP-mannose to benzyl-α-mannopyranoside. Having thus demonstrated the transfer of the natural sugar mannose (from GDP-mannose). The assay was performed with GDP-azidomannose molecules. Under all conditions tested the transfer of azido sugars was not observed. In an attempt to eliminate any batch specific anomalies associated with our protein preparation, the transfer assays were repeated using the commercially available Man T enzyme. An efficient transfer of mannose was observed, although definitive evidence for azidomannose transfer remained elusive.

4.7 Conclusion

The observations made in this study are in contrast to those reported previously (S. Marchesan personal communication), in which purified α-1,2-mannosyltransferase was observed to catalyse the transfer of 4-azidomannose to methyl-α-*D*-mannoside (demonstrated by MS analysis). One possible explanation for this may lie in the method of purification of Man T. We found the enzyme in the periplasmic space which was in contrast to the previous study but agreed with Wongs original protocol. This difference in purification may be linked to the amount of enzyme used in the transfer assays. Whilst (55 µL, 18.3 mU) of commercial enzyme and (55 µL, concentration unknown) of purified enzyme showed the transfer of mannose from GDP-mannose to benzyl-α-mannopyranoside in a 95% yield (for commercial). No azidomannose transfer was observed using either the commercial or purified enzyme. An unspecified amount (107.2 µL, concentration unknown) of enzyme was reportedly used in the study previously undertaken. This previous study successfully transferred 4-azidomannose and when we repeated the same procedure, albeit with our purified enzyme transfer was not observed.

Given this observation, it may be reasonable to assume that the amount of enzyme used in the previous study may have been in vast excess compared with the amount used here, which may serve to explain the azido transfer observed previously or the enzyme used previously may have contained other proteins as there was no clear SDS PAGE data indicating the purity of the previously expressed enzyme.

References:

[1] S. Marchesan, D. Macmillan, *Chem Commun*. 2008, 36, 4321.

[2] P. Wang, G. J. Shen, Y. F. Wang, Y. Ichikawa, C. H. Wong, *J. Org. Chem*. 1993, 58, 3985.

[3] H. A. D. Boer, L. J. Comstock, M. Vasser, *Proceed. Natl. Acad. Sci. USA.* 1983, 80, 21.

[4] P. A. Romero, M. Lussier, A. M. Sdicu, H. Bussey, A. Herscovics, *Biochem. J*. 1997, 321, 289.

[5] G. F. Hermann, P. Wang, G. J. Shen, E. G. Sunceda, S. H. Khan, K. L. Matta, C. H. Wong, *J. Org. Chem.* 1994, 59, 6356.

[6] Y. D. Lobsanov, P. A. Romero, B. Sleno, B. Yu, P. Yip, A. Herscovics, P. L. Howell, *J. Biol. Chem*. 2004, 279, 17921.

[7] O. Tsuruta, H. Yuasa, H. Hashimoto, K. Sujino, A. Otter, H. Li, M. M. Palcic, *J. Org. Chem*. 2003, 68. 6400.

[8] C. Unverzagt, H. Kunz, J. C. Paulson, *J. Am. Chem. Soc.* 1990,112, 9309.

Chapter 5

Azidohomoalanine for glycoprotein remodelling

5.1 Synthesis of tools to expand glycoprotein remodelling

The third method to introduce azide groups to proteins was investigated in this study. This method introduced azide groups into fully synthetic peptides using the amino acid azidohomoalanine (Aha). Click chemistry with synthetic propargyl glycosides would allow further modification to the homogeneous glycopeptide analogues, which were tested for their compatibility with native chemical ligation. Although this synthetic method differs from the *in vitro* strategy of using GDP-azidosugars for glycoprotein remodelling, it nevertheless creates a homogeneous glycoprotein using click chemistry.

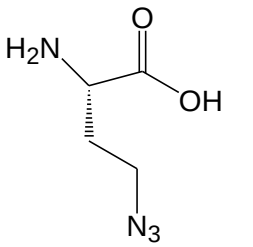
Apart from incorporation of Aha into peptides we have also performed test reactions with our synthetic propargyl sugar analogues. Click chemistry of propargyl LacNAc to an azidosugar was performed and the resulting trisaccharide was enzymatically modified with sialyltransferase. These series of experiments validates the use of propargyl sugars for glycoprotein remodelling technique where chemical clicked sugars bearing a triazole can be tolerated enzymatically by sialyltransferase.

In addition synthesis of potential tools that will help expand the use of our GDP-azidomannose analogues where investigated in this chapter. The synthesis of Chitobiosyl azide would potentially act as a different substrate, for the transfer of azidosugars using the alternative enzyme β-1,4-mannosyltransferase. This may give us another route to explore our azidosugar transfer. Another tool that we have investigate is a peptide-linked mannose. The peptide will include a His_6 tag, which will enable easy purification and isolation of the transferred azido product. Hopefully, these tools will be useful as intermediate molecules to test the transfer of azidosugars and further aid the ongoing research of the use GDP-azidomannose for glycoprotein remodelling.

5.2 Azidohomoalanine (Aha) in protein engineering

The introduction of non-natural amino acids with a chemically unique group has numerous applications in protein engineering and functional studies. In 2002, Bertozzi demonstrated that a methionine surrogate, azidohomoalanine (**Figure 5.0**), is activated by the methionyl-tRNA synthetase of *E. coli* and replaces methionine in proteins expressed in methionine-depleted bacterial cultures [1]. Furthermore, proteins containing azidohomoalanine were selectively modified in the presence of other cellular proteins by means of Staudinger ligation

with triarylphosphine reagents [1]. In fact, a number of different azide–bearing amino acids were incorporated but Aha was by far the most translationally active azide-bearing analogue of methionine. The yield of purified Aha-substituted recombinant protein obtained in standard protein expression experiments was comparable to yields of wild-type, methionine-substituted protein [1].



139

Figure 5.0 *Structure of L-azido*homoalanine (Aha)[1].

We will incorporate Aha into synthetic peptides and perform click chemistry with our propargyl glycosides. These synthetic propargyl glycosides will be tested for their acceptance of sialylic acid by sialyltransferase.

5.3 *Results*

5.4 Chitobiosyl azide synthesis

The synthesis of chitobiosyl azide has already been published by Flitsch and co-workers [2]. We will replicate its synthesis and use it as a potential tool for our glycoprotein remodeling method. The anomeric azide can also be used to attach to a peptide *via* an *N*-(propargyl)-bromoacetamide. This will produce a peptide-linked chitobiose (**Figure 5.1**).

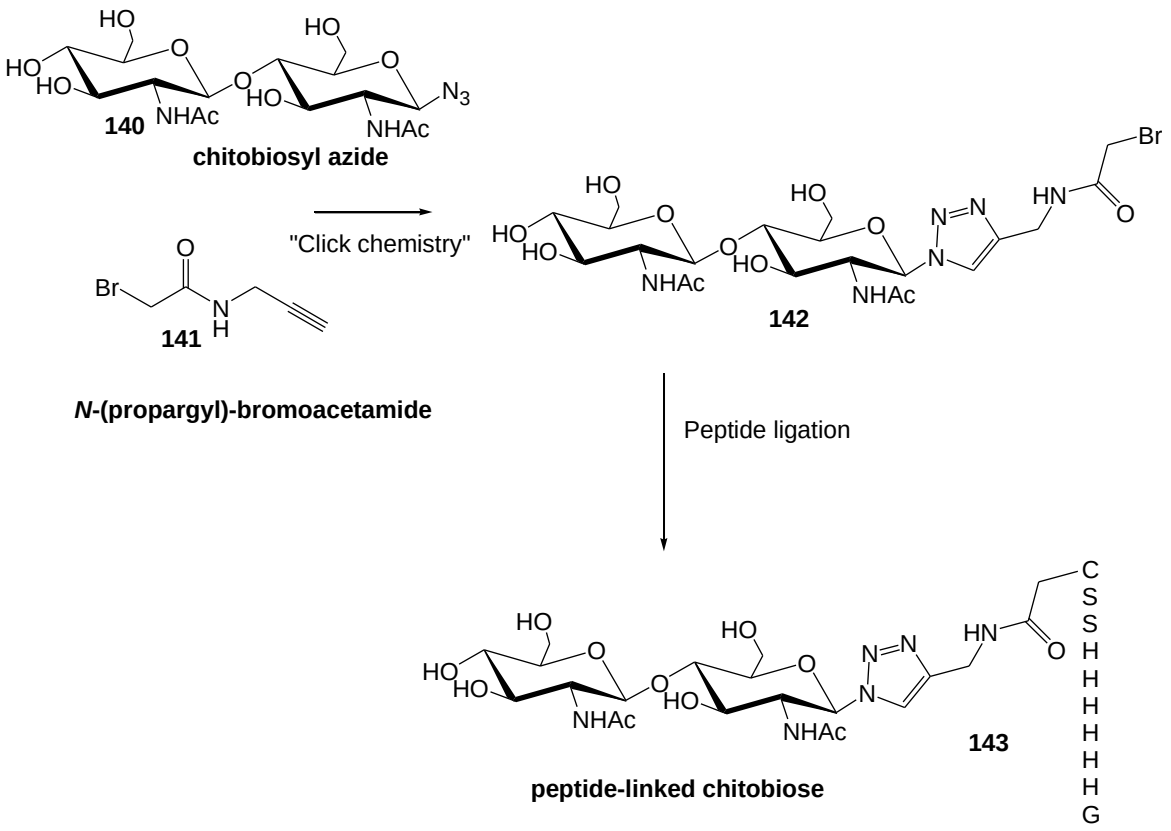


Figure 5.1 *The use of chitobiosyl azide as a peptide-linked precursor. The attached peptide tool **143** contains a His_6 tag which would allow easy purification of the product, after testing of the potential azido transfer.*

The chitobiose-linked peptide would be a useful tool to investigate the transfer of our azido mannose by the alternative glycosyltransferase enzyme, α-1-4 mannosyltransferase. Chitobiosyl azide was synthesised *via* the glycosidic azide acceptor **149** and the trichloroacetimidate donor **151**.

The glycosidic azide acceptor was synthesised in a five-step process where readily available *N*-acetyl glucosamine reacted with acetyl chloride to produce the fully protected glycosyl chloride **145** (**Figure 5.2**). Next, sodium azide and tetrabutylammonium hydrogen sulphate (TBAHS) introduced the azide at the anomeric position, in a phase transfer catalysis reaction to produce the β-glycosyl azide **146**. Hydroxyl deprotection and selective protection of the free 4- and 6-hydroxyl with *p*-anisaldehyde dimethylacetal, produced the *p*-methoxybenzylidine protected azido glycoside **147**. Further protection of **147** with acetic anhydride in pyridine yielded **148**. Reductive regeoselective ring opening of **148** with sodium cyanoborohydride in the presence of trifluoroacetic acid resulted in the formation of the azide acceptor **149** (**Figure 5.2**).

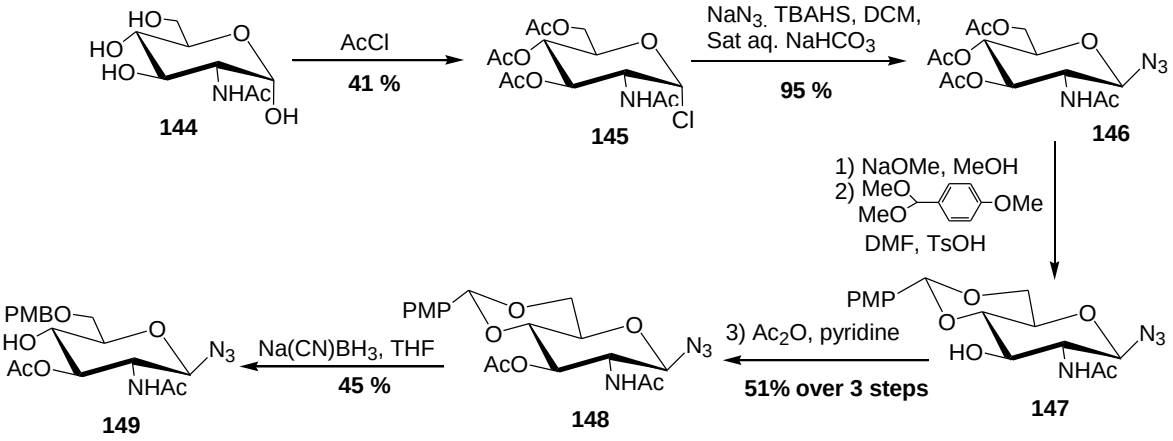


Figure 5.2 *Synthesis of the azide acceptor **149** [2]*.

The synthesis of glycosyl donor started with glucosamine hydrochloride which was converted to the free base with sodium methoxide and then the 2-amino group was protected, to produce the phthalimido sugar (**Figure 5.3**). It was found that drying the *N*-phthalimido sugar overnight under high vacuum, before completing the protection with acetic anhydride and pyridine, resulted in a greater yield of the acetylated *N*-phthalimido sugar **71**.

Next, the activation of the anomeric position with tin tetrachloride and thiophenol resulted in yields of up to 79% of the thiophenyl glycoside **72**. Selective deprotection of the anomeric position with *N*-bromosuccinimide (NBS) proved to be difficult. It was found that when monitoring the reaction by TLC, deprotection of not only the thioglycoside group, but also the acetyl groups, which was evident by the presence of four distinct spots when stained with *p*-anisaldehyde dip.

This could be controlled by recrystallising the NBS from boiling water and carefully optimising the reaction times. The optimum time of four hours, allowed a reasonable yield of the hemi-acetal product **150** and the starting thio-glycoside could be easily recovered by flash chromatography. The glycosyl donor was finally produced as the β-trichloroacetimidate by reaction with trichloroacetonitrile with **150** in potassium carbonate to give **151** (**Figure 5.3**).

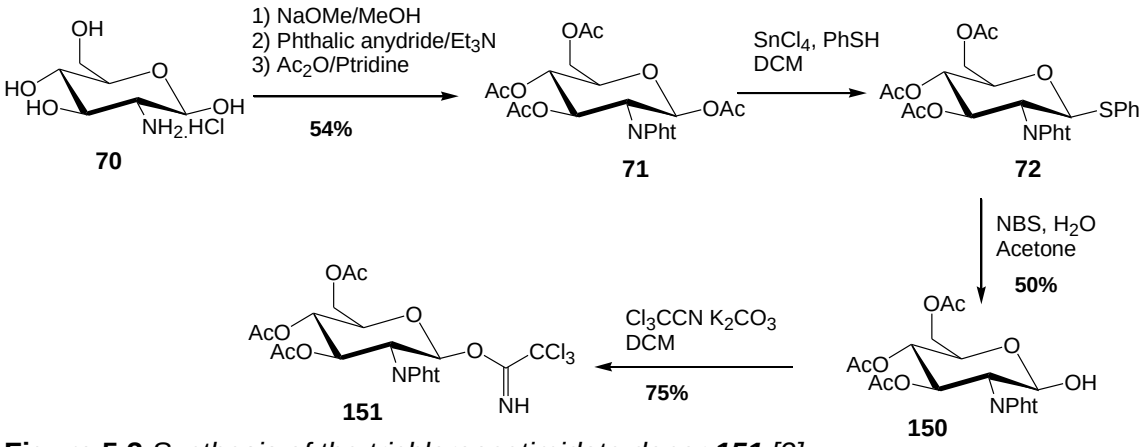


Figure 5.3 *Synthesis of the trichloroacetimidate donor **151** [2].*

Coupling of the trichloroacetimidate donor **151** to the azide acceptor **149** was a poor yielding step. It was vital that completely dry conditions were employed, where both the acceptor and the donor were dried under high vacuum for 16 hours before the reaction was performed. Activation of the anomeric trichloroacetimidate with a Lewis acid BF_3,Et_2O *via* S_N1-type mechanism with neighboring group participation yielded the fully protected disaccharide **152** in only 27% yield (**Figure 5.4**). The yield obtained was comparable to that reported [2], however by mass analysis it was apparent that a mixture of by-products were also formed. This included starting materials, the hemi-acetal form where the trichloroacetimidate donor was hydrolysed and a double *N*-phthalimido sugar disaccharide where the trichloroacetimidate donor reacted with the hemi-acetal by-product. ^{1}H NMR analysis of these by-products was complicated and so further purification was required.

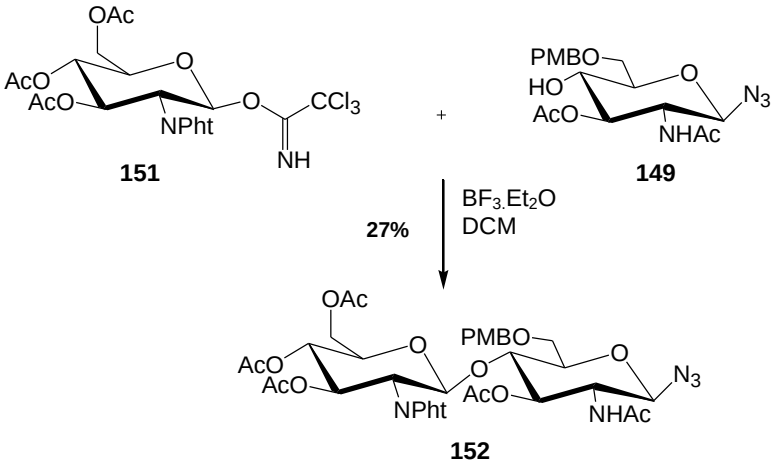


Figure 5.4 *The glycosidic linkage between the trichloroacetimidate donor and the azide acceptor for the synthesis of chitobiosyl azide **152** [2].*

Furthermore, we also tried to perform the β-1,4-glycosidic linkage using thioglycoside **72** as the donor. The use of thioglycosides has its advantages, as they are more chemically stable

than the trichloroacetimidates and will not react with the vast majority of nucleophiles unless specificlly activated with a soft electrophile. *N*-iodosuccinimide (NIS) is most frequently used along with a catalytic amount of trifluoromethanesulfonic acid (trific acid, TfOH). The activation of the anomeric sulphur occurs by the nucleophilic attack of the sulphur on NIS with the loss of succinimide. This forms the sulphonium leaving group and the azide acceptor is trapped *via* an S_N1 type pathway. In a first attempt we used NBS in the presence of TfOH to form the β-1,4-glycosidic linkage using thioglycoside **72** and acceptor **149**. However, no sign of the product was found, changing the soft electrophile to NIS, still resulted in gave a negative result, but only the loss of the *p*-methoxybenzyl protecting groups were observed. The experiments discussed above demonstrated that the formation of a new glycosidic linkage between thioglycoside **72** and the azide **149** containing a *p*-methoxybenzyl protecting group was unsuccessful. However, with the disaccharide **152** from the original methodology *(***Figure 5.4***)* the final deprotection steps with cerrium ammonium nitrate and hydrazine hydrate and then complete protection with acetic anhydride and pyridine were performed resulting in the fully protected chitobiosyl azide **153** (**Figure 5.5**).

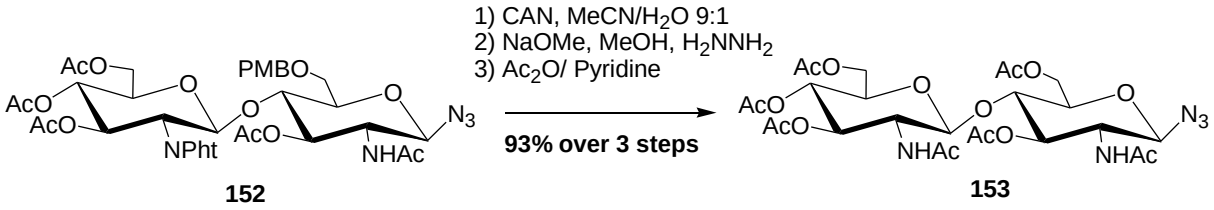


Figure 5.5 *Synthesis of chitobiosyl azide **153** [2].*

5.5 The synthesis of mannose-linked peptide

We envisaged that having another tool such as a mannose-linked peptide, would be useful to see the transfer of azidomannose onto a peptide precursor after the initial transfer of azidosugars have been rectified. A mannose containing an iodoacetimide group **154** was ligated to a cysteine residue of a His_6 tag peptide **155** (**Figure 5.6**). The His_6 tag would enable quick and easy purification after azido transfer using a nickel column.

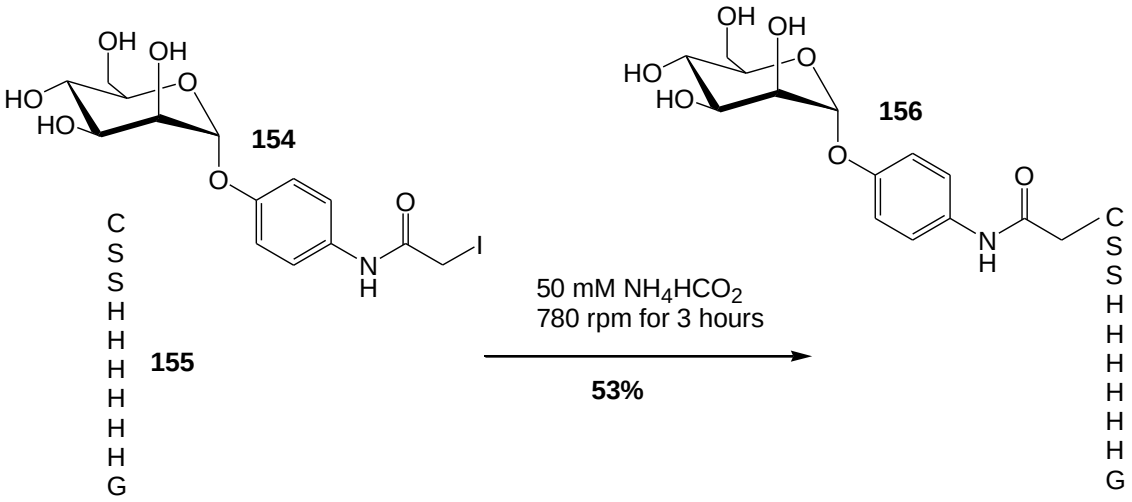


Figure 5.6 *Ligation of iodoacetamide mannose with the His_6 tag.*

The peptide chosen contained a cysteine residue, which was required for the sugar ligation, two serine amino acids to allow some flexibility of the peptide, six histidine molecules and a glycine residue to contain the His_6 tag and perhaps prevent terminal damage. Cysteine was loaded onto a rink amide resin manually, and the remaining peptide was synthesised in automated fashion. The rink amide resin was used due to its availability and its high loading capability also it is important to note that the rink amide resin is cleaved to give a C-terminal amide.

The Kaiser test was performed to check efficiency of loading. The Kaiser test involves ninhydrin (2,2-Dihydroxyindane-1,3-dione) a chemical used to detect ammonia or primary and secondary amines. When reacting with these free amines, a deep blue or purple colour known as Ruhemann's purple is evolved [3]. After the first amino acid was coupled to the resin the Kaiser test showed a colorless/yellow. This indicated that the resin was fully coupled with the first amino acid as there were no free amines on the resin. Fluorenylmethyloxycarbonyl (Fmoc) deprotection and spectrophotometric methodology were used to calculate the level of first residue attachment. The actual loading of 0.225 mmol/g was comparable with the theoretical amount, which was 0.226 mmol/g. After automated synthesis of the remaining peptide, the purity of the peptide was examined by HPLC and MS, a distinct single peak at retention time: 11.06 minutes represented the His_6 tagged peptide **155** (**Figure 5.7**).

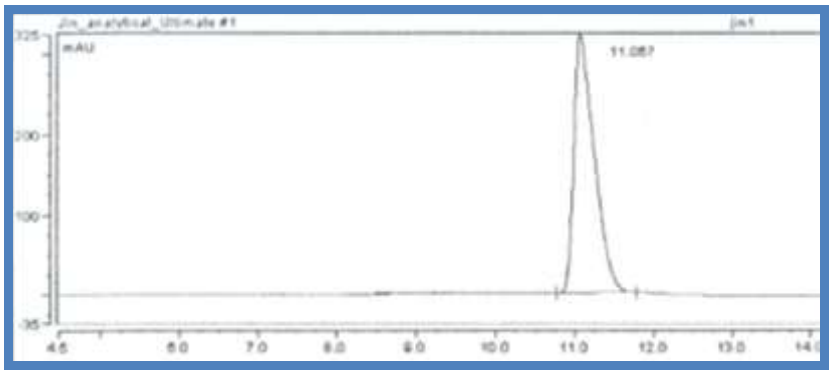


Figure 5.7 *HPLC trace* of *His-tag peptide, performed with* (5-95%) *acetonitrile/water (0.1* % *TFA), with* C18 *column (Luna* C18 *100A Size: 250* x *10.00 mm,* 10 *micron).*

Next with the iodoacetamide sugar (donated by Dr Derek Macmillan) **154** was ligated to the cysteine-terminated His_6 peptide **155** in 50 mM ammonium formate buffer and purification by HPLC, resulted in a the mannose-linked peptide **156** as a white foam (**Figure 5.6**). MS showed the molecular ion as 1485.5, as well as 1509.76 (M+Na) and 744.45 (the 2+ ion). This peptide-linked mannose may be an important tool that will enable subsequent testing of azidosugar transfer. This tool will eventually be used to test the transfer of azidosugars, with both α-1,2 mannosyltransferase and β-1,4 mannosyltransferase enzymes. The His_6 tag will be the main advantage of testing with this tool, as it will allow easy isolation of the product by nickel affinity chromatography.

5.6 Click chemistry with propargyl LacNAc

To test the final step of our glycoprotein remodeling technique we will need to demonstrate click chemistry of synthetic propagyl LacNAc to an azidosugar analogue and then test the resulting product, to see whether it is subsequently recognised by the enzyme sialyltransferase. These experiments would provide a good model for the final stages of our glycoprotein remodelling technique where a transferred azidosugar can be clicked with propagyl LacNAc and enzymatically modified.

Click chemistry was performed between the fully protected propargyl LacNAc and the fully protected 4-azidomannose. Initially the reaction was performed in water/methanol/phosphate buffer (50 mM)/CI_3CH (1:1:1:9), sodium ascorbate, and $CuSO_4.5H_2O$. However the reaction seemed to be very slow and only 15-20% yield was obtained after flash column chromatography. Cu(I) catalyses the formation of 1,2,3-triazoles from azides and alkynes. However it is thermodynamically unstable, resulting in easy oxidation to Cu(II) and/or disproportionation to Cu(0) and Cu(II) [14]. Tris-(benzyltriazolylmethyl)amine (TBTA), is a

stabilising ligand for Cu(I) developed by the Sharpless group [4]. TBTA protects Cu(I) from oxidation and disproportionation, while enhancing its catalytic activity. The ligand is a C3-symmetric derivative, and is useful for azide-alkyne cycloaddition with upto 10^6-fold of rate acceleration [4].

The synthesis of TBTA **160** was achieved by producing benzyl azide **158** from benzyl chloride **157** and then performing click chemistry with tripropargylamine **159** (**Figure 5.8**).

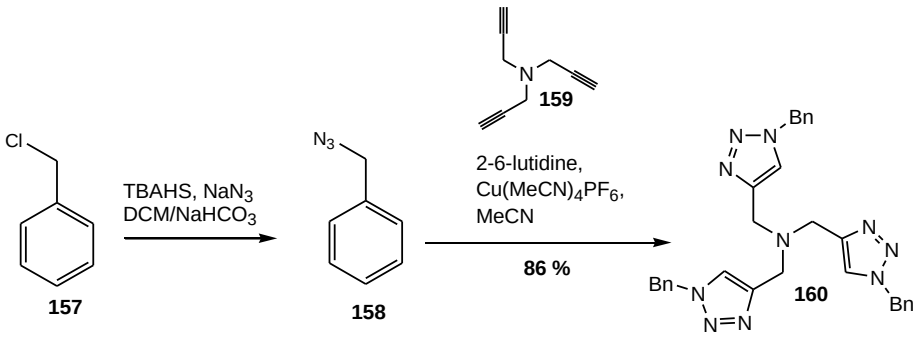


Figure 5.8 *The synthesis of TBTA ligand **160** [4].*

The ligand was then used in the click reaction between propargyl LacNAc and 4-azidomannose (**Figure 5.9**). The reaction proceeded for 5 hours and after flash column chromatography a 67% yield of the desired trisaccharide **161** was obtained (**Figure 5.10**).

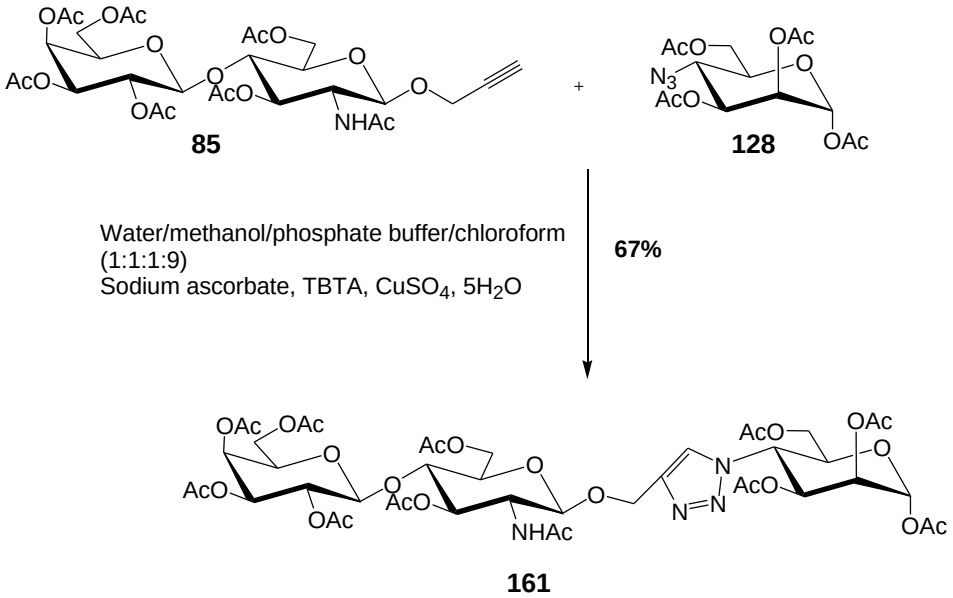


Figure 5.9 *Click chemistry between propargyl LacNAc and 4–azidomannose in the presence of TBTA.*

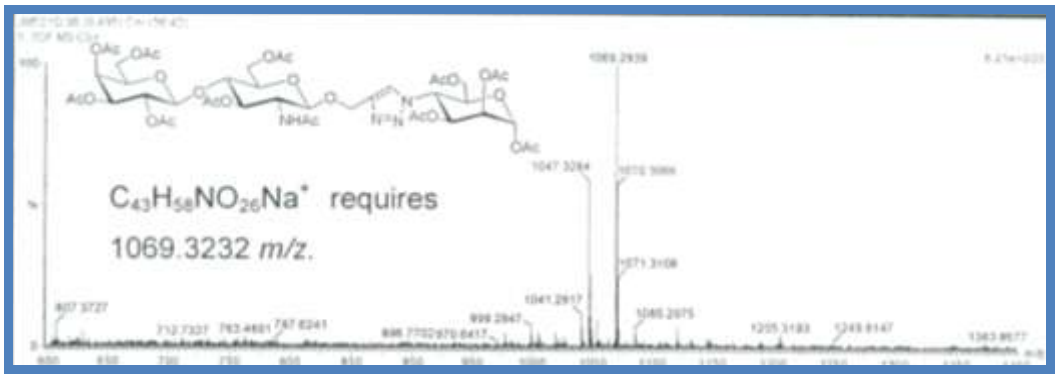
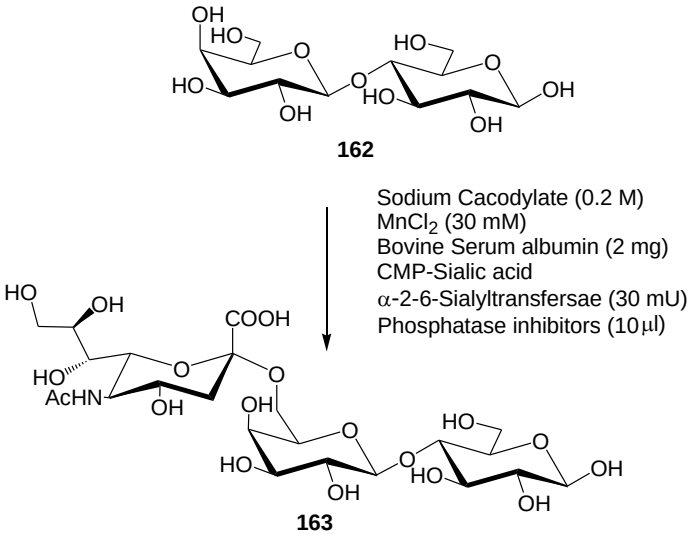


Figure 5.10 *MS of the clicked trisaccharide product **161**.*

The next step was to test the transfer of sialic acid onto our trisaccharide **161**. Initially the transfer conditions were optimised on the natural substrate lactose, which is commercially available. The reaction mixture was adjusted to pH 7.1, prior to addition of CMP-sialic acid, α-2-6-sialyltransferase and lactose. The product α-2-6-sialyllactose **163** was purified by HPLC and MS showed the desired product.



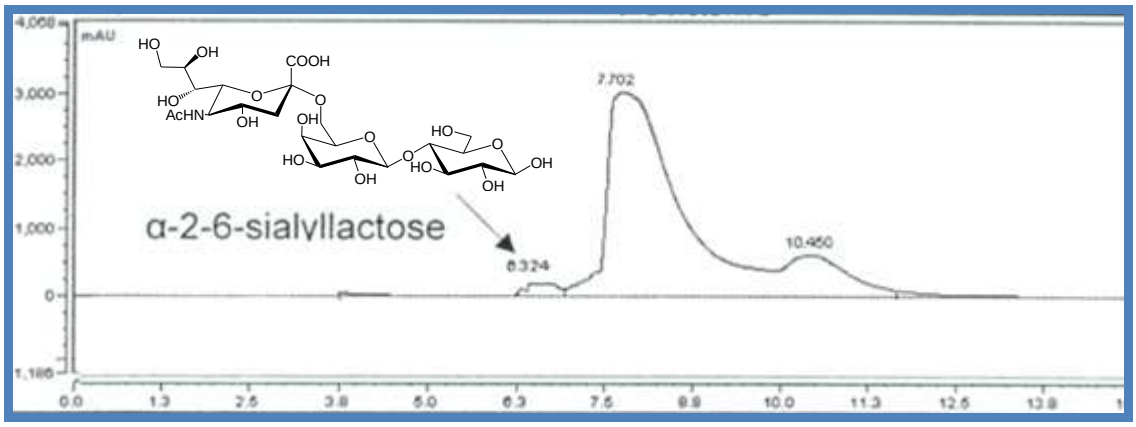


Figure 5.11 *The reaction conditions for the synthesis of α-2-6- sialylactose and the HPLC trace of the reaction mixture containing α-2-6-sialyllactose (RT: 6.324) Sphere Clone 5u SAX Size: 250 x 10.00 mm, 5 micron.*

Deprotection of trisaccharide **161** was achieved with NaOMe in anhydrous MeOH and after neutralising with acetic acid and removal of solvent. The transfer of sialic acid was performed with our optimised conditions. After HPLC purification, the transfer of sialic acid was successful shown by MS analysis to form the tetrasaccharide product **164** (**Figure 5.12**).

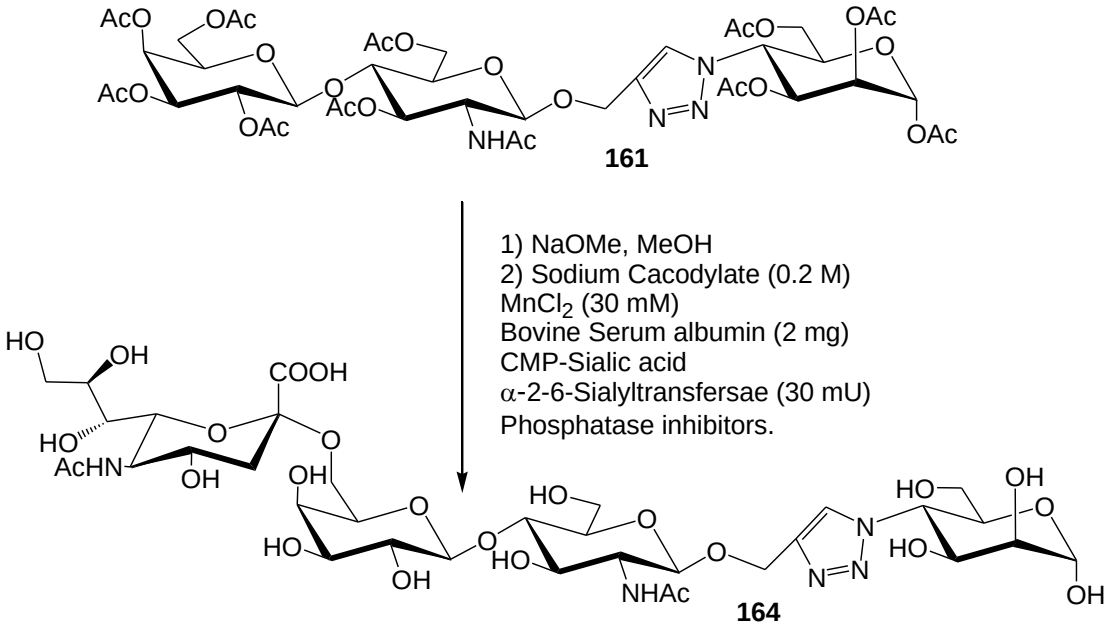


Figure 5.12 *Sialic acid transfer to form the product **164**.*

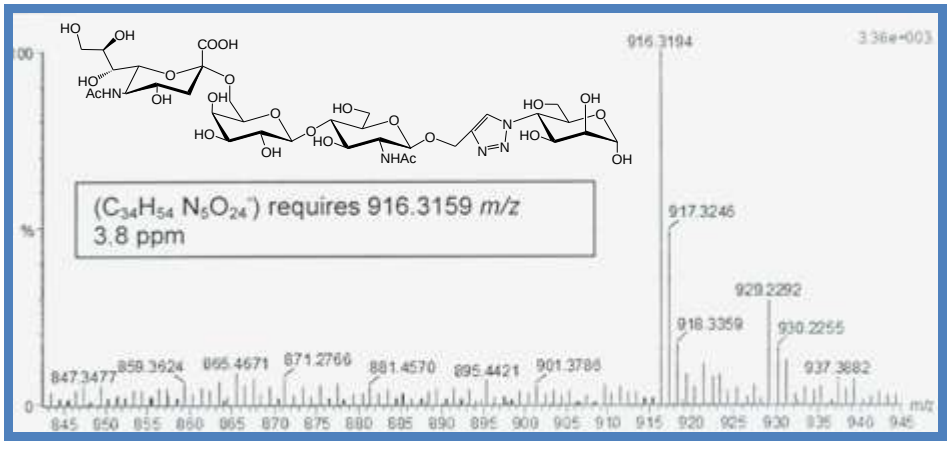


Figure 5.13 *MS of tetrasaccharide **164**.*

These results show that synthetic propargyl LacNAc can be “clicked “ onto an azido sugar and can be tolerated by the α-2,6-sialyltransferase for the transfer of sialic acid. Therefore these results give a good indication that the last stages of our overall remodelling technique, where glycoprotein incorporated azido sugars are clicked with propargyl LacNAc and

modified by sialyltransferase is a feasible process.

5.7 Synthetic incorporation of azidohomoalanine for glycoprotein remodelling

The introduction of azide groups into a fully synthetic peptide using the amino acid azidohomoalanine (Aha) was an alternative method we investigated for glycoprotein remodelling. Click chemistry with synthetic propargyl glycosides would allow further modification to homogeneous glycopeptide analogues which were tested for their compatibility with native chemical ligation. The glycoproteins CD52 and α-dystroglycan (ADG) were used as our model proteins. CD52 is a glycosylphosphatidylinositol (GPI)-anchored cell surface antigen expressed on almost all human lymphocytes and sperm cells [15]. It is one of the smallest natural glycoproteins consisting of only twelve amino acids residues but carries a large complex- type *N*- glycan at the Asn-3 residue [5]. Our peptide was based on the sequence of CD52 with the following exceptions: Asn will be substituted with Aha to allow click chemistry with propargyl LacNAc and propargyl GlcNAc also, Asp-4 will be replaced with Glu as Asp residues are known to be unstable during thioesterification. Furthermore, an extra Gly and Cys residue will be added to allow thioesterification to produce a thioester which can be used for native chemical ligation (**Figure 5.14**).

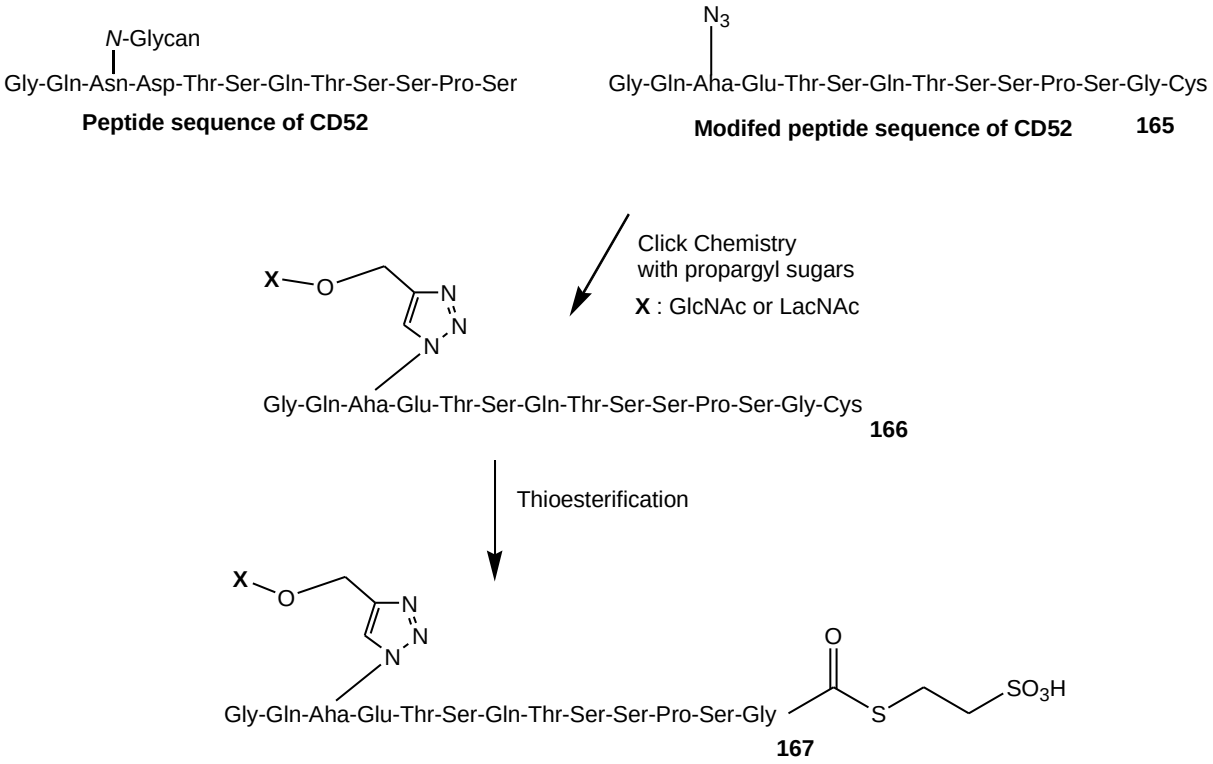


Figure 5.14 *The synthetic route to produce a chemically modified peptide of CD52 containing clicked sugars and terminal thioester.*

ADG is an extra-cellular protein containing both *N*-linked and *O*-linked glycosylation sites and its defective glycosylation is associated with muscular dystrophy (Chapter 1). The mucin-like *O*-mannosylation domain in ADG spans from amino acid 316 to 489 [6]. Two similar peptide sequences, corresponding to residues 401–420 and 336–355, respectively, were shown to be strongly mannosylated by POMT, whereas other peptides from ADG and peptides of various mucin tandem repeat regions were poorly mannosylated [6]. Furthermore substitution of Ala residues for Ser or Thr residues have shown that Thr-414 of peptide 401–420 and Thr-351 of peptide 336–355 were prominently modified by *O*-mannosylation [6]. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry and Edman degradation analysis of the mannosylated peptide 401–420 indicated that Thr-414 was the residue that was most prominently modified by *O*-mannosylation [6]. With these results in mind the synthesis of peptide 401-420 was undertaken. The peptide was contructed from two fragments where Thr 404, 406 and 414 were replaced by Aha. The first fragment (peptide 401-409) contained an extra Cys residue could be thioesterified for native chemical ligation. The second fragment (peptide 410-420) replaced Tyr 410 with Cys facilitate native chemical ligation. These Aha bearing fragments will then undergo click chemistry with our propargyl mannose and then native chemical ligation to produce the chemical modified glycopeptide 401-420 of the mucin- like domain of ADG (**Figure 5.15**). If successful then this chemically modified glycopeptide can be potentially used as test substrates for the transfer of mannose by POMT, to see if the enzyme can tolerate the unnatural triazole linkage.

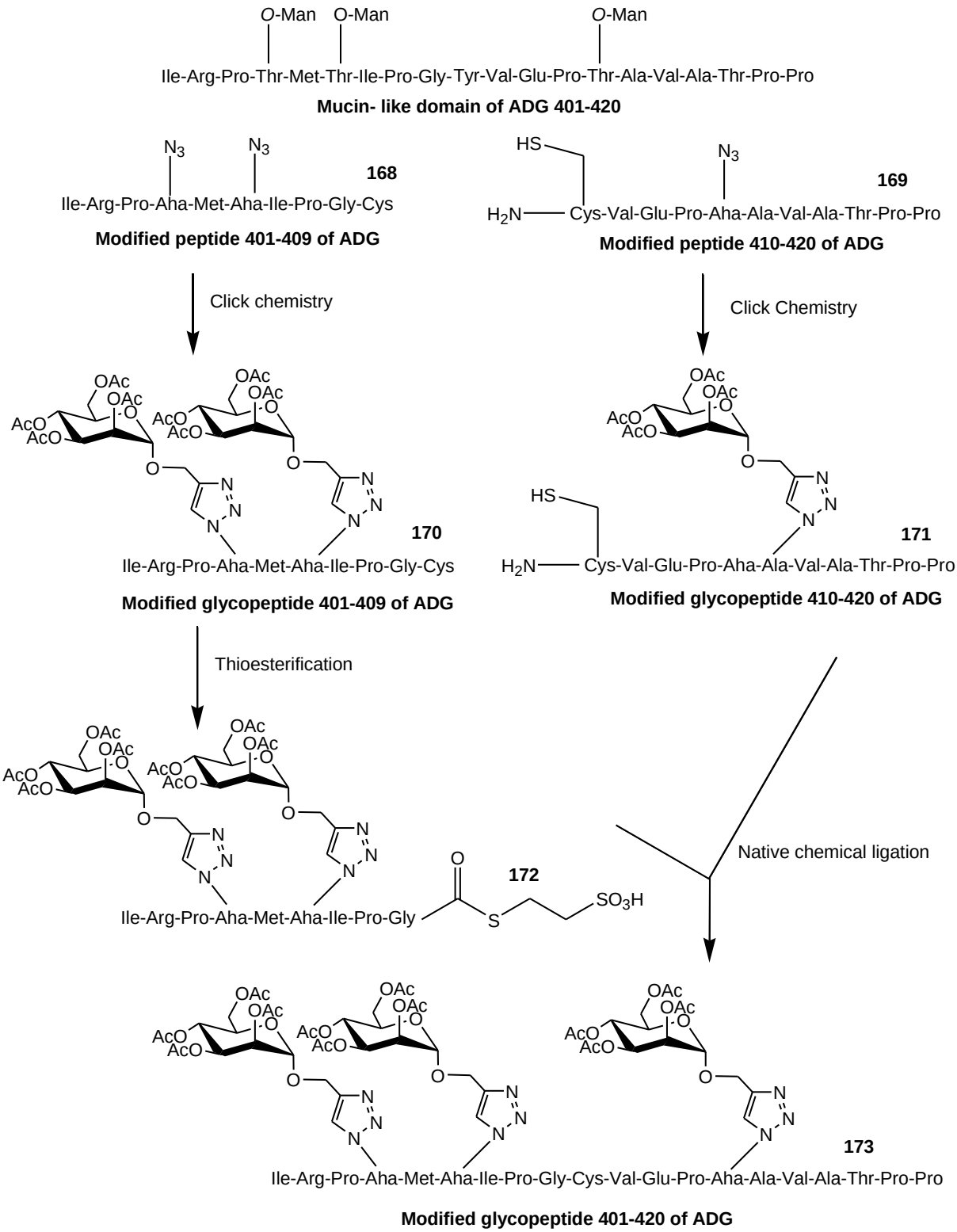


Figure 5.15 *The synthetic route to produce the chemically modified glycopeptide 401-420 of the mucin- like domain of ADG. Click chemistry, thioesterification and native chemical ligation methodologies are used to produce the modified peptide.*

The modified peptides 401-409 and 410-420 of ADG and CD52 peptide **59** were

synthesised in 0.01 mmol scale using the peptide synthesiser. Mass analysis showed the expected mass for all three peptides (**168**, **169** and **165)** with the respective mass charge ratio values 1111.2, 1095.3 and 1359.1. With the correct sequence of peptides at hand, we performed click chemistry with our propargyl glycosides and then HPLC purification using C18 column (100Å Size: 250 x 10.00 mm, 10 micron). The resulting glycopeptides were examined using mass analysis (**Table 6**).

Propargyl glycoside	Azido-modified peptide (number)	Expected mass of glycoprotein product *m/z*	Mass Found *m/z*	% yield
GlcNAc	CD52 (**165**)	1779.75	1778.50	15-20
LacNAc	CD52 (**165**)	2068.00	-	-
Mannose	401-409 ADG (**168**)	1911.07	1911.08	15-20
Mannose	410-420 ADG (**169**)	1495.61	1496.20	15-20

Table 6 *Mass analysis of click chemistry reactions between the azide bearing peptides CD52* ***165****, ADG 401-409* ***168****, and ADG 410-420* ***169*** *and fully acetylated propargyl glycosides. Click chemistry was performed in sodium ascorbate, Cu(II)SO_4, (9:1:1) $CHCl_3$: EtOH: 50 mM sodium phosphate.*

Click chemistry of propargyl GlcNAc to the modified CD52 peptide **165** was successful however, the same reaction with propargyl LacNAc was unsuccessful. This may be due to the time given for the reaction to proceed, perhaps repeating the reaction with the presence of TBTA stabilising ligand may improve the reaction. On the other hand the click reaction between propargyl mannose and peptides **168** and **169** were both successful showing not only the correct masses by mass analysis but also the plus two ions.

The next step was the thioesterification of the modified glycopeptides 401-410 of ADG **168** and the GlcNAc clicked CD52 peptide **166**. The thioesterification reaction involves an N→S acyl shift with the terminal cystine residue, under acidic conditions to produce unstable thioester intermediate which reacts with MESNa to give a stable thioester (**Figure 5.16**) [7, 8].

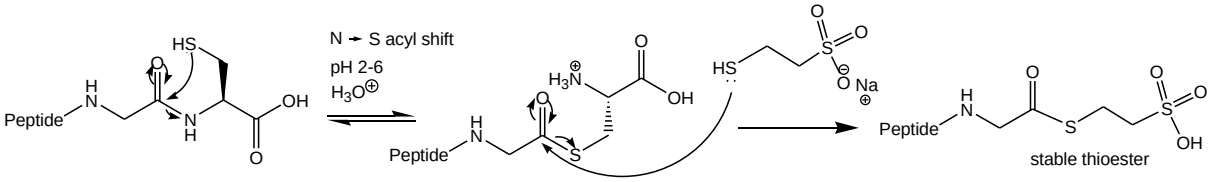


Figure 5.16 *Thioesterification reaction with MESNa [7].*

Initial results showed that modified glycopeptide 401-410 of ADG **168** did not undergo thioesterification. However, glycopeptide 16**6** underwent thioesterification and after HPLC purification, mass analysis showed the major product as the desired glycopeptide **167** (**Figure 5.17**).

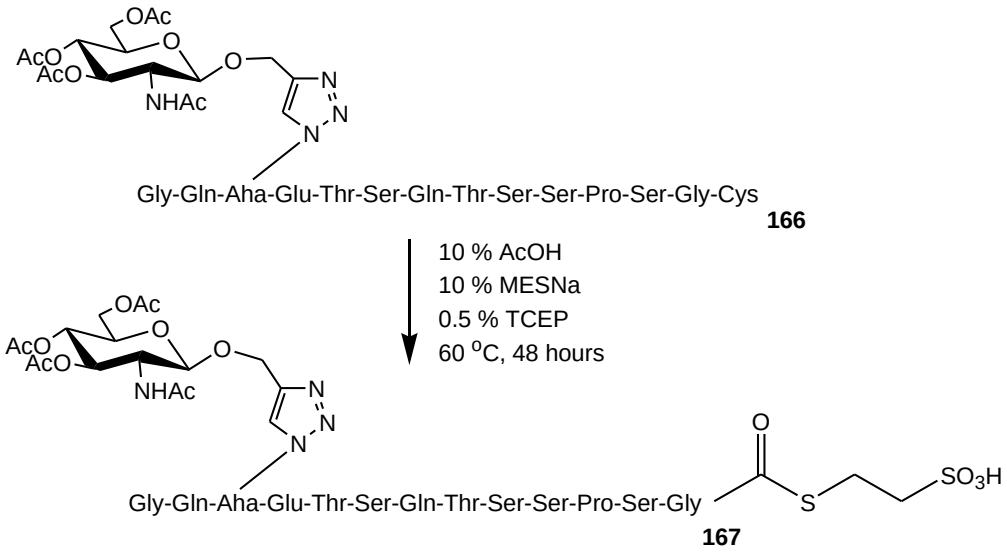


Figure 5.17 *Thioesterification of glycopeptide **166**. *The reaction was also repeated in 1M sodium phosphate pH 5.8 without AcOH and also showed the product **167**. Mass analysis showed the mass of 1801.8 m/z the calculated mass of **167** is 1801.5.*

The incorporation of azide groups into a fully synthetic peptide using the amino acid azidohomoalanine was successful. Click chemistry with synthetic propargyl glycosides of GlcNAc and mannose was also achieved. Thioesterification of modified peptide ADG 401-410 was problematic showing no sign of the desired thioester product. However GlcNAc clicked CD52 peptide **166** was successful resulting in the desired thioester containing glycopeptide **167**. This is the first time that an azidohomoalanine incorporated glycopeptide has tolerated our thioesterification procedure. Further work will need to test the compatibility of the thioester **167** with native chemical ligation.

5.8 Conclusion

This project aimed to investigate and synthesise tools that would aid in a novel glycoprotein remodeling technique using GDP-azidomannose analogues. We have synthesised one of the key components, propargyl LacNAc and have shown that it can be chemically clicked to an azido sugar and enzymatically modified by α-2-6-sialyltransferase which we hope would suitably mimic terminal sugars found on *N*-linked glycoproteins. We have also used α-1,2-mannosyltransferase to test the transfer of azidomannose from GDP-azidomannose, to different substrates. Although this was unsuccessful we have optimised the transfer reaction using GDP-mannose and developed a method to isolate the product using α-benzyl mannoside as the acceptor substrate.

We have synthesised azido- and thioacetate-sugars and administered the sugars to yeast expressing EPO in an attempt to incorporate the synthetic sugars into yeast glycoprotein. Unfortunately, preliminary results show that the incorporation was unsuccessful. As an ongoing research for the use of GDP-azidomannose for glycoprotein remodeling, we have synthesised acceptor molecules including mannose linked His_6 tagged peptide and chitobiosyl azide. These tools can potentially allow easy purification and further testing, to transfer azido mannose from GDP-azidomannose.

One of the major reasons why we have not tried our glycoprotein remodeling technique on an actual glycoprotein is that the azido-transfer from GDP-azidomannose has been unsuccessful. One reason for this may be that GDP-azidomannose may be broken down to the nucleotide diphosphate. The nucleotide diphosphate GDP is known to inhibit the initial transfer of the sugar from its sugar nucleotide [9] and therefore may prevent our azido-sugars from being transferred.

A procedure to regenerate the nucleotide sugar and remove nucleotide diphosphates may help push our unsuccessful azido-transfer reaction. This regeneration of nucleotide sugars can be seen in **Figure 5.18**. The cycle starts with the transfer of the mannose from the GDP-mannose to the acceptor (*O*-methyl mannose). The by-product is a nucleotide diphosphate (GDP), which is known to inhibit the initial transfer of the sugar from its sugar nucleotide [9]. GDP is therefore recycled, by conversion to its corresponding nucleotide triphosphate (GTP) with phosphoenol pyruvate. The product GTP together with mannose-1-phosphate are substrates for the enzyme GDP-mannose-pyrophosphorylase (GDP-ManPP), which regenerates the sugar nucleotide GDP-mannose.

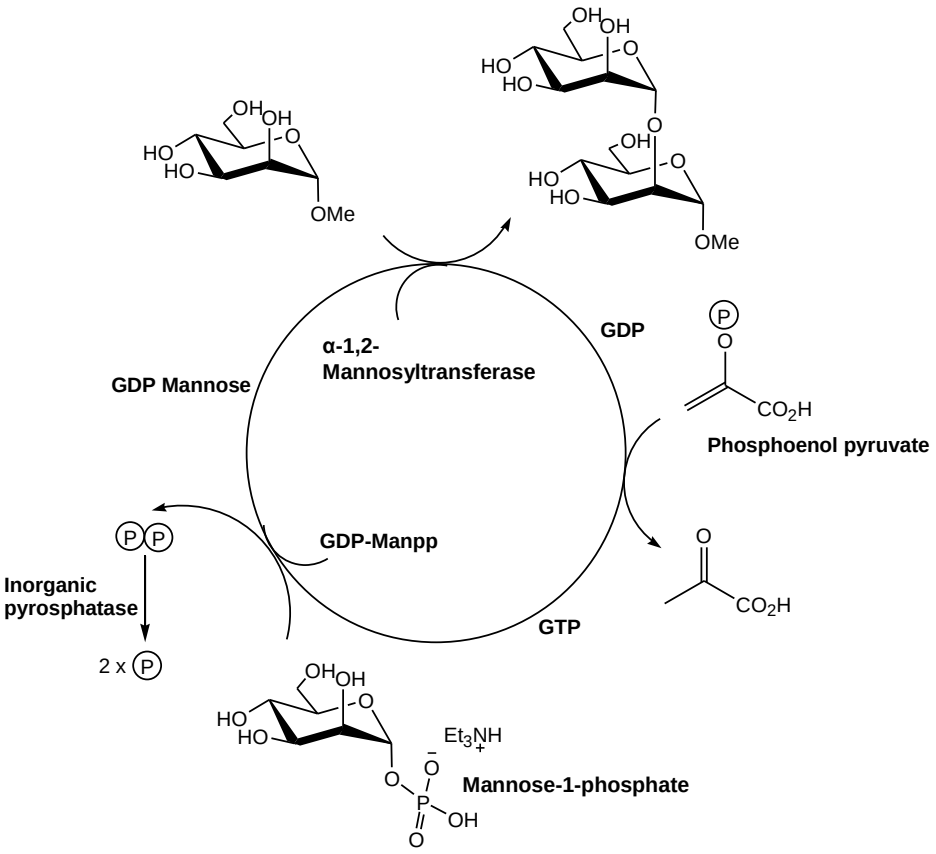


Figure 5.18 *Regenerating the nucleotide sugar GDP mannose for mannose transfer [9].*

A major by-product from the cycle is the inorganic diphosphates produced when forming the sugar nucleotide (GDP-mannose). These diphosphates can also potentially inhibit the cycle, but if broken down with inorganic phophatase to produce free phosphates, then they are rendered inactive. If this cycle becomes successful then we will apply the principle to our GDP-azidomannose derivatives.

The purification of GDP-ManPP, an enzyme required for this regenerating cycle has been achieved (see experimental section). GDP-ManPP is only one component, within the regenerating cycle. The other components, such as phosphoenol pyruvate, inorganic pyrophosphatase and GDP mannose, are commercially available. While α-1,2-mannosyltransferase has been expressed and is also commercially available. The only component missing to test the cycle is the mannose-one-phosphate. Further work will be required to test the cycle and apply it to our azido-mannose derivatives.

During this project we have demonstrated the synthesis of propargyl glycosides and have successfully shown that these propargyl sugars can be ligated to azide bearing peptides using click chemistry. This can be seen with the synthesis of CD52 and ADG analogues which were modified using click chemistry. The modified CD52 peptide has also

successfully undergone thioesterification to produce a thioester which is poised for native chemical ligation. We have demonstrated that remodelling glycoproteins enzymatically using α-1,2 mannosyltransferase and GDP-azidomannose has been difficult but synthetic methods using Aha and propargyl glycosides to produce homogeneous glycopeptides analogues is less problematic.

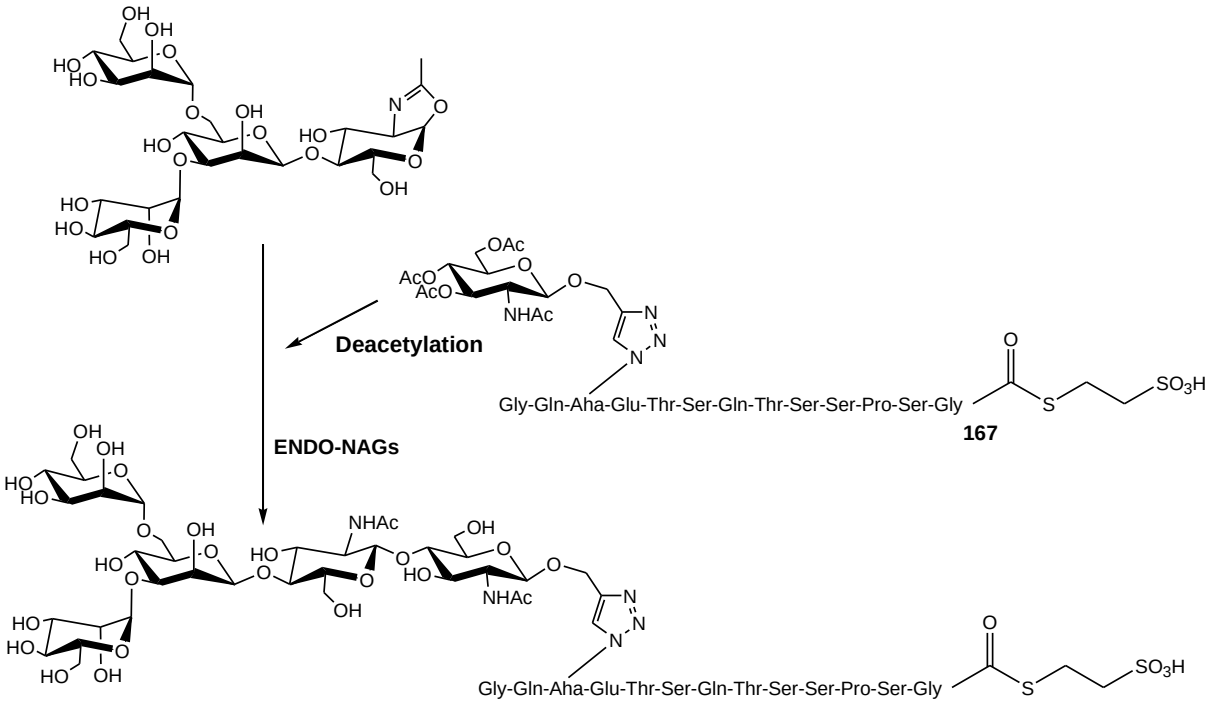


Figure 5.19 *Modified CD52 peptide **167** as a potential substrate for endo-NAGs.*

These synthetic homogeneous glycoprotein mimics may ultimately act as substrates for further methods of glycoprotein remodelling using enzymes such as endo-NAGs (**Figure 5.19**) to ultimately produce homogeneous glycoproteins to suit therapeutic use.

References:

[1] K. L. Kiick, E. Saxon, D. A. Tirrell, C. R. Bertozzi, Proc. *Natl. Acad. Sci. USA*. 2002, 99, 19.

[2] D. Macmillan, A. M. Daines, M. Bayrhuber, S. L. Flitsch, *Org. Left.* 2002, 4, 1467.

[3] E. Kaiser, R. L. Colescott, C. D. Bossinger, P. I. Cook, *Anal. Biochem.* 1970, 34, 595.

[4] T. R. Chan, R. Hilgraf, K. B. Sharpless, V. V. Fokin, *Org. Left.* 2004, 17, 2853.

[5] M. Umekawa, C. L. T. Higashiyama, W. Huang, H. Ashida, K. Yamamoto, L. X. Wang, *J. Biol. Chem*. 2010, 285, 511.

[6] H. Manya, T. Suzuki, K. A. Manya, H. K. Ishida, M. Mizuno, Y. Suzuki, T. Inazu, N. Dohmae, T. Endo, *J. Biol. Chem*. 2007, 282, 20200.

[7] J. Kang, J. P. Richardson, D. Macmillan, *Chem. Commun*. 2009, 407.

[8] J. Kang, N. L. Reynolds, C. Tyrrell, J. R. Dorin, D. Macmillan, *Org. Biomol. Chem.* 2009, 7, 4918.

[9] C. Unverzagt, H. Kunz, J. C. Paulson, *J. Am. Chem. Soc.* 1990,112, 9309.

Chapter 6

Experimental section

6.1 Organic chemistry

6.2 General techniques and reagents

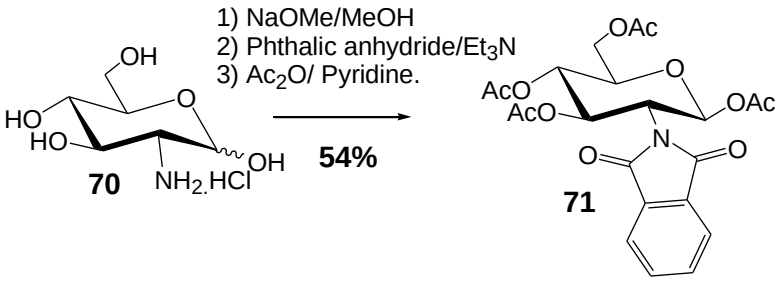
All solvents and reagents were standard laboratory grade purchased from Sigma Aldrich, Fisher or Acros and used as received unless otherwise stated. Where a solvent was described as dry it was purchased as anhydrous grade or from the University College London in-house anhydrous solvent system.

6.3 Instrumentation

^{1}H NMR spectra were recorded at 250, 300, 400, or 500 MHz and ^{13}C NMR spectra were recorded at 75, 100 or 125 MHz on a Bruker 250Y instrument. Chemical shifts (δ) are reported in *ppm* relative to $Si(CH_3)4$ (δ=0) and coupling constants *(J)* in Hz, signals are sharp unless stated as broad (br), s: singlet, d: doublet, t: triplet, m: multiplet and q: quaternary. Residual protic solvent, $CDCl_3$ (δH: 7.26, s) was used as the internal standard in ^{1}H-NMR spectra unless otherwise stated. Low resolution electrospray mass spectroscopy was carried out on a Micromass Quattro LC electrospray mass spectrometer with an applied voltage of 35-50 V. High resolution electrospray ionisation (ESI) mass spectra were obtained at the mass spectrometry facility at University College London. UV spectroscopy was carried out in an 8543 UV/Vis spectrophotometer (Hewlett Packard) using disposable cuvettes. Infrared spectra were obtained using a Perkin-Elmer 1600. Paragon Series FT-IR spectrometer as liquid thin films. Peptides were synthesised on a rink amide resin (unless otherwise stated) using standard Fmoc chemistry with a peptide synthesiser (Applied Biosystems 433A).

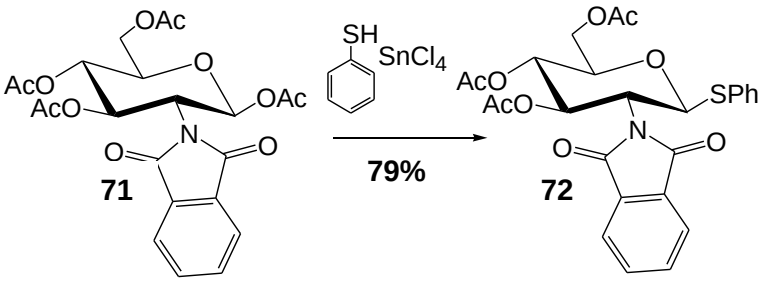
Analytical thin layer chromatography (TLC was carried performed using Merck aluminium backed plates coated with silica gel 60F254. Flash chromatography was carried out over Fisher silica gel 60 Å particle size 35-70 micron. Components were visualised using *p*-anisaldehyde dip and UV light (254 nm). The HPLC was equipped with binary pump (P680 HPLC pump Dionex), a semipreparative column (SAX, 10 μ, 100 A, 250 x 10 mm) or (carbon 18), UV detector (UVD 170U/340U Dionex).

2-*N*-phthalimido-2-deoxy-1,3,4,6-tetra-O-acetyl-β-D-glucopyranose [1].



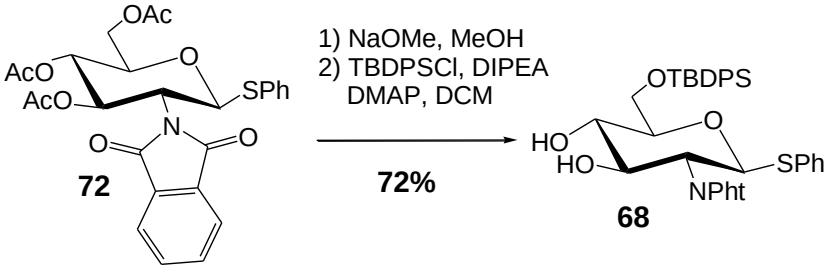
To a stirred solution of glucosamine hydrochloride (10.0 g, 46.3 mmol) and anhydrous methanol (75 ml), sodium methoxide (3.0 g, 46.3 mmol) was added. The reaction mixture was stirred for 10 minutes at room temperature and then filtered with suction. Phthalic anhydride (3.5 g, 23.0 mmol) was added to the filtrate and stirring was continued for a further 20 minutes. A further portion of phthalic anhydride (3.5 g, 23.0 mmol) was then added followed by triethylamine (7.6 ml, 55.6 mmol). The reaction mixture was stirred at room temperature for 10 minutes and then at 50°C for 30 minutes. The resulting mixture (containing a thick white precipitate) was cooled for 1 hour in an ice bath and filtered with suction. The precipitate was washed with cold methanol (2 x 20 ml) and dried under high vacuum. The dry white solid was suspended in acetic anhydride (45 ml), cooled in an ice bath and pyridine (23 ml) was added carefully with stirring. The reaction mixture was stirred at room temperature for 16 hours and then poured into ice/water (200 ml) and extracted with chloroform (3 x 200 ml). The combined organic extracts were washed with 5% HCl (100 ml), saturated $NaHCO_3$ (120 ml), water (120 ml) and brine (100 ml). The organic phase was then dried with (Na_2SO_4), filtered, and the solvent removed under vacuum to afford the crude product as an orange oil which was purified by flash column chromatography over silica (hexane/ ethyl acetate 1: 1 R_f 0.4) to afford 2-*N*-Phthalimido-2-deoxy-1,3,4,6-tetra-*O*-acetyl-β-*D*-glucopyranoside (11.97 g, 54%) as a white foam. ^{1}H NMR (300 MHz, $CDCl_3$):δ ppm = 7.68-7.75 (4H, m, ArH), 6.45 (1H, d, J = 7, H1), 5.80 (1H, dd, J = 8, H3), 5.13 (1H, dd, J = 8, H4), 4.38 (1 H, dd, J = 8, H2), 4.28 (1 H, dd. H6a), 4.05 (1 H, dd, H6b), 3.96 (1 H, m, H5), 2.23 - 2.20 (12H, 4 x s, $COCH_3$) ^{13}C NMR (62 MHz, $CDCl_3$): δ (ppm) = 170.5, 170.0, 169.3, 168.5 and 167.2 (qC), 134.4 (2 x ArCH), 131.1 (C), 123.7 (2 x ArCH), 89.6 (CH), 72.5 (CH), 70.4 (CH), 68.1 (CH), 61.4 (CH_2), 53.3 (CH), 20.6 (2 x OCH_3), 20.4 (OCH_3) and 20.3 ($COCH_3$) [1]. Found m/z (ESI): 500.30 (M+Na), $C_{22}H_{23}NO_{11}Na^+$ requires m/z 500.40.

2-*N*-phthalimido-2-deoxy-3,4,6,-tri-*O*-acetyl-*S*-thiophenyl-β-*D*-glucopyranoside [1].



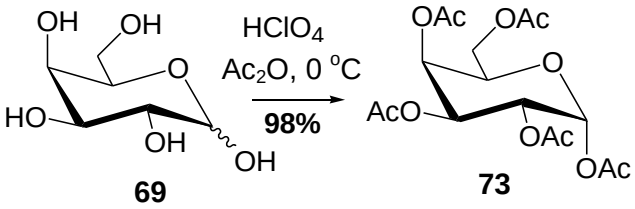
A solution of the acetylated *N*-phthalimido sugar (10.0 g, 20.9 mmol) and tin tetrachloride (7.31 ml of a 1 M solution in DCM, 42.6 mmol) in anhydrous DCM (75.0 ml) under nitrogen was treated with thiophenol (3.06 g, 27.83 mmol). The reaction mixture was stirred at room temperature for 16 hours and then diluted with DCM (100 ml). The solution was washed with water (100 ml), saturated bicarbonate (100 ml), then water (100 ml). The organic phase was dried with $MgSO_4$ and the solvent was removed under reduced pressure to afford the crude product, which was purified by flash chromatography over silica (ethyl acetate/petroleum ether 1: 1 R_f 0.4) to isolate a white solid (8.71 g, 79%) ^{1}H NMR (300 MHz, $CDCl_3$):δ ppm = 7.87-7.73 (4H, m, ArH), 7.41-7.24 (5H, m, ArH), 5.78 (1H, dd, *J* = 8, H3), 5.72 (1H, d, *J* = 8, H1), 5.13 (1H, dd, *J* = 8, H4), 4.34 (1H, dd, *J* = 8, H2), 4.27-4.17 (2H, m, H6b, H6b), 3.92-3.86 (1 H, m, H5), 2.09, 2.21, and 1.82 (9H, 3 x S, $COCH_3$). ^{13}C NMR (62 MHz, $CDCl_3$):δ (ppm) = 170.5, 170.0, 169.3, 167.7, 166.8 (5 x qC), 134.3 (2 x ArCH), 123.6 (2 x ArCH), 134.3 (ArCH), 133.1 (2 x ArCH), 130.8 (qC), 128.7 (2 x ArCH), 128.3 (ArCH), 123.6 (2 x ArCH) (5C, ArH), 82.9 (CH), 75.8 (CH), 71.5 (CH), 68.6 (CH), 62.1 (CH_2), 53.4 (CH), 20.6, 20.5 and 20.3 (3 x CH_3) Found *m/z* (ESI): 551.01 (M+Na), $C_{26}H_2SNO_9Na^+$ requires *m/z* 550.53.

2-*N*-phthalimido-2-deoxy-6-*tert*-butyl-diphenyl-silanyl- *S*-thiophenyl-β-D -glucopyranoside.



The fully protected thioglycoside (1.10 g, 2.08 mmol) was dissolved in methanol and a 0.5 M solution of NaOMe in MeOH was added until a pH of 10 was obtained. The reaction was stirred for 5 hours and monitored by TLC (1:1 EtOAc/hexane) then neutralised with acetic acid. The mixture was then concentrated and azeotroped with toluene, before being dried under high vacuum for 1 hour to yield the crude partial protected sugar (0.81 g, 1.99 mmol). Next this partially protected thioglycoside was dissolved in anhydrous DCM (30 ml), to which TBDPSCl (0.52 ml, 1.99 mmol), DIPEA (0.70 ml, 3.98 mmol), DMAP (24.0 mg, 0.19 mmol), was added, and the reaction stirred under nitrogen for 3.5 hours. The reaction was monitored by TLC (100%, EtOAc, R_f 0.3) to completion and diluted with (20 ml) DCM, washed with $NaHCO_3$ (20 ml), and water (20 ml). The organic phase was dried with $MgSO_4$, filtered and concentrated under reduced pressure to afford a yellow oil. The product was purified by flash chromatography over silica (100%, EtOAc, R_f 0.3) to afford the product as a yellow oil (0.96 g, 72%). ^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) = 7.73-7.17 (19H, m, 19 x ArH), 5.64 (1H, d, *J* = 10, H1), 4.34 (1 H, dd, H3), 4.24 (1 H, dd, H4), 4.10 (2H, dd, H6a, H6b), 3.72 (2H, m, H2, H5), 3.20 (1 H, s, OH), 2.80 (1 H, s, OH), 1.08 (12H, S, CH_3 x 3). ^{13}C NMR (75 MHz, CDCl3): δ (ppm) = 135.7, 134.8, 134.2, 133.1, 132.7,1 32.3, 131.6, 130.0, 129.0, 128.9, 127.9, 127.8 (4 x NPht + 8 x ArH), 83.5,78.7,73.6,72.9,69.0,55.2, (C1-C6), 26.8 (3 x CH_3), 19.2 ($\underline{C}CH_3$). MS (ESI): *m/z* 662.28 (M+Na), $C_{36}H_{37}N0_6SSiNa^+$ requires 662.82.

1,2,3,4,6-penta-*O*-acetyl-*α*-*D*-galactopyranose [2].

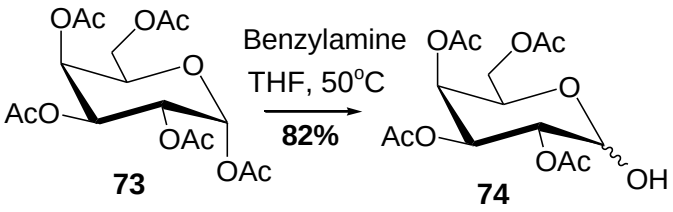


69 **73**

To a 0°C solution of *D*-galactose (25.0 g, 0.13 mol) in acetic anhydride (223 ml), perchloric acid (0.1 ml) was added dropwise. The reaction mixture was stirred for 3 hours, and monitored by TLC to completion. The crude product was washed with chloroform (300 ml), brine (200 ml), $NaHCO_3$ (200 ml) and water (200 ml). The organic phase was dried with $MgSO_4$, filtered, and the solvent removed under reduced pressure to afford a brown oil (50.1 g, 98%) that was used without further purification.^{1}H NMR (300 MH*z*, $CDCl_3$): δ (ppm) = 6.34 (1 H, d, *J* = 2, H1), 5.41 (1 H, dd, H4), 5.33-5.29 (2H, m, H2, H3), 4.34-4.29 (1 H, m, H5), 4.10-4.29 (2H, m, H6a, H6b), 2.19, 2.13, 1.99, 1.97,1.96, (15H, s x 5, $COCH_3$ X 5). ^{13}C NMR (75 MHz, $CDCl_3$): δ (ppm) = 170.3, 170.1, 169.9, 168.9, 166.4 (5 x RCOR), 89.7, 68.7, 67.4,

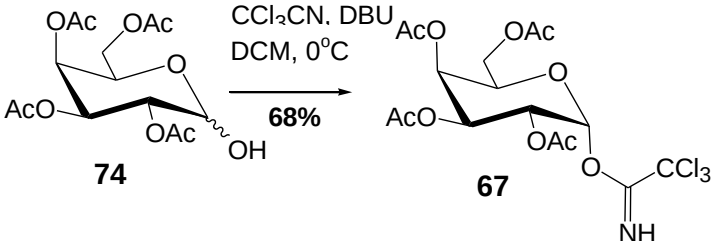
67.3, 66.4, 61.2 (C1-C6), 20.1, 20.9, 20.6, 20.5, 20.5 (5 x CO<u>C</u>H$_3$). MS (ESI): *m/z* 413.15 (M+Na), $C_{16}H_{22}NaO_{11}^+$ requires 413.33.

2,3,4,6-tetra-*O*-acetyl-*D*-galactose [2].



The peracetylated galactoside (7.80 g, 20.00 mmol) was dissolved in anhydrous THF (120 ml) under nitrogen at 50°C. Benzylamine (2.6 ml, 24.00 mmol) was added to the reaction mixture. The reaction was monitored by TLC (4:1 petroleum ether/ ethyl acetate) and when the reaction was almost complete the solvent was removed under reduced pressure and the resulting mixture purified by flash chromatography over silica (4/1 petroleum ether/ ethyl acetate, R_f 0.2) to afford a yellow oil (5.71 g, 82%).^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) = 6.34 (1H, d, *J* = 2, H1), 5.41 (1H, dd, H4), 5.20-5.07 (2H, m, H2, H3), 4.47-4.45 (1 H, m, H5), 4.10-4.29 (2H, m, H6a, H6b), 2.15, 2.10, 2.06, 1.99, (12H, s x 4, $COCH_3$ X 4). ^{13}C NMR (75 MHz, $CDCl_3$): δ (ppm) = 170.3, 170.1, 169.9, 166.4 (4 x RCOR), 89.7, 68.7, 67.4, 67.3, 66.4, 61.2 (C1-C6), 20.1, 20.9, 20.6, 20.5 (4 x CO<u>C</u>H$_3$). MS (ESI): *m/z* 371.01 (M+Na), $C_{14}H_{20}NaO_{10}^+$ requires 371.29.

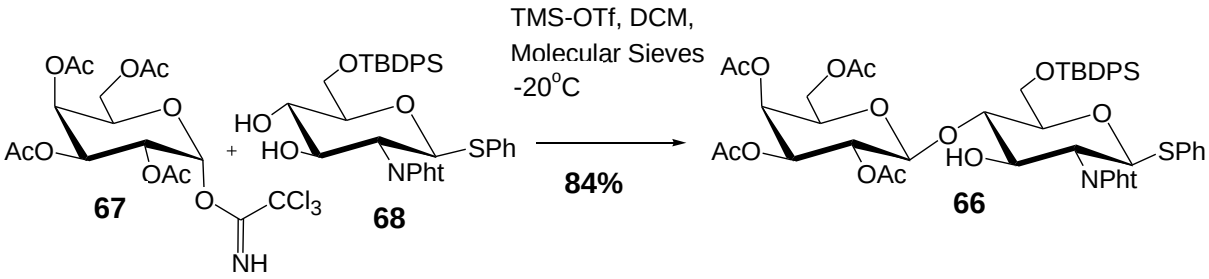
2,3,4,6-tetra-*O*-acetyl-*α*-*D*-*O*-galacopyranosyl trichloroacetimidate [2].



A solution of the hemiacetal (3.66 g, 10.50 mmol) in anhydrous DCM (100 ml) was cooled to 0°C. Trichloroacetonitrile (10.62 ml, 105 mmol) and DBU (1.57 ml, 10.50 mmol) were added to the mixture which was stirred 3 hours until TLC showed that the reaction was complete. The reaction mixture was concentrated under reduced pressure, and purified by flash chromatography over silica (1:1 petroleum ether/ ethyl acetate, R_f 0.5) to afford the product as a yellow oil (3.63 g, 68%).^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) = 8.67 (1H, s, NH) 6.34

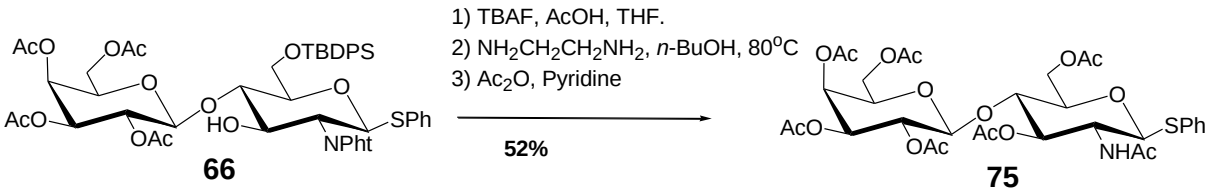
(1H, d, *J* = 3, H1), 5.53 (1H, dd, H4), 5.41-5.29 (2H, m, H2, H3), 4.43-4.38 (1H, m, H5), 4.16-4.01 (2H, m, H6a, H6b), 2.06, 1.99, 1.98, 1.98 (12H, s x 4, $COCH_3$ X 4).^{13}C NMR (75 MHz, $COCl_3$): δ (ppm) = 170.1 (4 x RCOR), 93.6, 69.0, 67.5, 67.4, 66.9, 61.3 (C1-C6), 20.6 (4 x $CO\underline{C}H_3$). MS (ESI): *m/z* 514.06 (M+Na), $C_{16}H_2OCl_3NNa0_{10}^+$ requires 514.01.

Disaccharide 66.



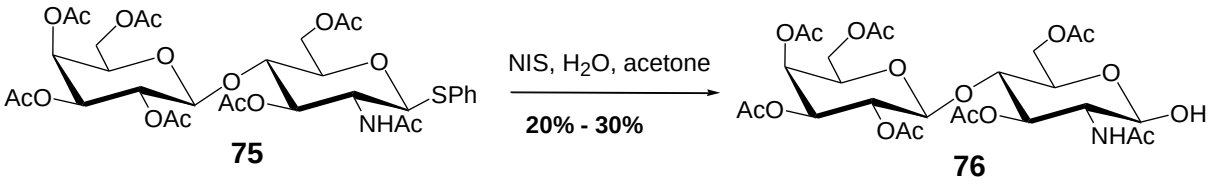
The trichloroacetimidate donor (0.92 g, 1.19 mmol) and the silylated acceptor (0.60 g, 0.94 mmol) were dissolved in anhydrous DCM (20 ml), containing 4 Å molecular sieves at -20°C. A solution of 0.1 M TMS-OTF in anhydrous DCM (4.80 ml, 0.48 mmol) was added to the reaction mixture at -20°C and stirred at -20°C for 2 hours. The reaction was allowed to warm to room temperature then stirred at room temperature until TLC indicated that the reaction was complete. The reaction mixture was diluted with DCM (50 ml) and neutralised with solid $NaHCO_3$, filtered, and solvent removed under reduced pressure. The crude product was purified by flash chromatography over silica (4:1 toluene / ethyl acetate, R_f 0.3) to afford the product as a white foam (0.77 g, 84%) ^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) = 8.14-7.18 (19H, m, ArH), 5.69 (1H, d, *J* = 11, H1), 5.21 (1H, d, *J* = 3, H4), 5.18 (1H, dd, H2), 4.96 (1H, dd, H3), 4.74 (1H, d, *J* = 8, H1'), 4.42 (1H, dd, H3'), 4.12 (1H, dd, H2'), 4.12-3.87 (7H, m, H6a', H6b', H4', H6b, H6b, H5, H5'), 3.61 (1 H, d, OH), 2.16, 2.03, 2.01, 1.69, (12H, s, $COCH_3$). ^{13}C NMR (100 MHz, $CDCl_3$): δ (ppm) = 170.4, 170.0, 169.8, 169.0, 168.1, 197.4, (6 x RCOR), 137.8, 135.9, 135.5, 134.0, 133.5, 132.6, 132.4, 132.1, 131.7,131.7,129.9, 129.8, 129.0, 128.8, 128.2, 127.9, 127.7, 127.6, 125.2, 123.6, 123.2 (4 x NPht, 5 x SPh, 12 x ArH) 101.1, 83.5, 80.4, 78.8, 71.2, 70.7, 70.5, 68.8, 66.8, 61.9, 61.1, 55.3, (C1-C6', C1-C6), 26.9 (3 x CH_3), 21.4, 20.5, 20.3, 20.2 (4 x $CO\underline{C}H_3$), 19.3 ($\underline{C}CH_3$). MS (ESI): *m/z* (M+Na) 992.3202 $C_{50}H_{55}NO_{15}SSiNa^+$ requires 992.3197.

Peracetylated-*N*-acetyl-Lactosamine thioglycoside.



A solution of the phthalimide protected disaccharide **66** (950 mg, 0.98 mmol), and acetic acid (0.59 ml, 9.18 mmol) was cooled at 0°C in THF (20 ml). A solution of 1 M TBAF in THF (1.14 ml, 3.92 mmol) was added to the reaction mixture and stirred at room temperature for 15 hours. The mixture was subsequently concentrated to a yellow oil. The oil was dissolved in *n*-BuOH and ethylendiamine (0.29 ml, 3.92 mmol) and the reaction stirred at 90°C under reflux for 15 hours. The resulting yellow solution was reduced to an oil and dissolved in (2:1) pyridine, acetic anhydride (90 ml) and stirred for 24 hours. The reaction was finally concentrated to give a brown oil and purified by flash chromatography over silica (100% EtOAc, R_f 0.4) to afford the product as a white foam (350 mg, 52%). ^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.73-7.21 (5H, m, ArH), 5.74 (1 H, d, *J* = 10, NH), 5.21 (1 H, d, *J* = 4, H4), 5.09-5.05 (2H, m, H2, H3'), 4.96 (1H, dd, H3), 4.63 (1H, d, *J* = 10, H1'), 4.49-4.47 (2H, m, H6a', H1), 4.14-4.06 (4H, m, H2', H6b', H6a, H6b), 3.89-3.83 (1H, m, H5), 3.73 (1H, dd, *J* = 9, H4'), 3.61 (1 H, m, H5'), 2.13, 2.08, 2.06, 2.04, 2.03, 1.98, 1.95 (21 H, S, $COCH_3$). ^{13}C NMR (75 MHz, $CDCl_3$): δ (ppm) = 170.9, 170.4, 170.3, 170.1, 170.1, 170.0, 169.3 (7 x RCOR), 132.7, 132.3, 128.9, 128.7, 127.9 (5 x ArH), 101.1, 86.8, 77.2, 76.1, 74.2, 70.9, 70.7, 69.1, 66.7, 62.4, 60.8, 52.9, (C1-C6', C1-C6), 23.3 ($NHCO\underline{C}H_3$) 20.6 (6 x $CO\underline{C}H_3$). MS (ESI): *m/z* (M+Na) 750.2001 $C_{32}H_{41}NO_{16}SNa^+$ requires 750.2044.

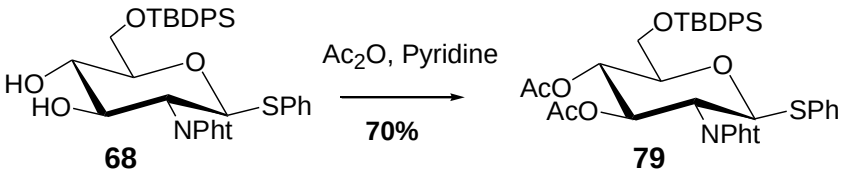
Molecule 76.



The thiophenyl glycoside **75** (200 mg, 0.27 mmol) and *N*-bromosuccinimide (185 mg. 0.82 mmol) were dissolved in acetone (20 ml) and water (1 ml). The mixture was stirred at room temperature for 3.5 hours and monitored by TLC (100% ethyl acetate, R_f 0.2). The reaction mixture was diluted with DCM (50 ml), and washed with water (20 ml), sodium thiosulphate (20 ml), and again with water (20 ml). The organic fraction was dried with $MgSO_4$, filtered

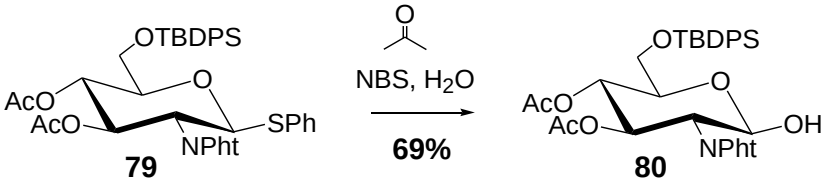
and solvent removed under reduced pressure. The mixture was purified by flash chromatography over silica (100%, ethyl acetate, R_f 0.2) to afford a white foam (38 mg, 20%).^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) = 6.63 (1 H, d, *J* = 10, NH), 5.31 (1 H, dd, H1 '), 5.21 (1 H, d, *J* = 3, H4), 5.18 (1 H, dd, H2'), 4.97 (2H, m, H3. H2), 4.41 (2H, m, H1, OH), 4.08 (5H, m, H6a', H6b', H6a, H6b, H2), 3.87 (1 H, dd, H4'), 3.76 (1 H, m, H5'), 2.14, 2.11, 2.07, 2.06, 2.05, 2.00, 1.99 (21 H, s, $COCH_3$). 13C NMR (75 MHz, $CDCl_3$): δ (ppm) = 170.9, 170.4, 170.3, 170.1, 170.1, 170.0, 169.3 (7 x RCOR), 101.1, 90.1, 77.2, 75.2, 71.2, 70.7, 70.9, 69.0, 68.0, 66.8, 61.2, 52.9, (C1-C6', C1-C6), 23.0 ($NHCO\underline{C}H_3$) 20.6 (6 x $CO\underline{C}H_3$). MS (ESI): *m/z* (M+Na) 658.1926 $C_{26}H_{37}N0_{17}Na^+$ requires 658.0.

3,4-di-*O*-acetyl-2-deoxy-2-*N*-phthalimido-6-*tert*-butyl-diphenylsilyl-β-*S*-thiophenyl-glucopyranoside.



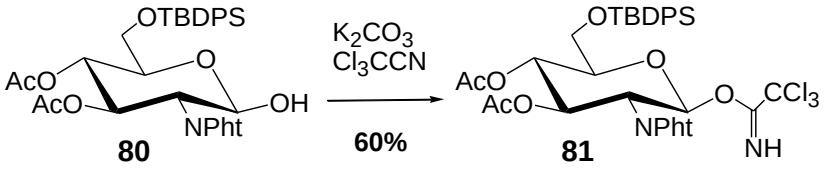
The partially protected sugar **4** (1.12 g, 1.75 mmol) was fully protected with (2/1 pyridine/acetic anhydride 40 ml). The reaction mixture was stirred for 5 hours and concentrated under reduced pressure to afford a brown oil. The resulting oil was purified by flash chromatography over silica (3: 1 petroleum ether/ ethyl acetate, R_t 0.3) to afford a white foam (0.87 g, 70%). ^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) = 7.77-7.21 (19H, m, 19 x ArH), 5.83 (1H, dd, H3) 5.76 (1H, d, *J* = 11, H1), 5.21 (1H, dd, H4), 4.38 (1H, dd, H2), 3.82-3.73 (3H, m, H5, H6a, H6b), 1.88, 1.83 (6H, s x 2, $COCH_3$), 1.08 (12H, S, CH_3 x 3). ^{13}C NMR (75 MHz, $CDCl_3$): δ (ppm) = 171.2, 170.3, 169.3, 167.9 (4 x RCOR) 135.7, 135.7, 134.4, 133.1, 133.0, 132.8, 131.6, 129.8, 129.0, 128.1, 127.8, 123.7 (4 x NPht + 8 x ArH), 83.1, 79.0, 72.1,69.0,62.7, 53.7, (C1-C6), 26.7 (3 x CH_3), 21.0, 20.5 (2 x $CO\underline{C}H_3$), 19.2 ($\underline{C}CH_3$). MS (ESI): *m/z* 746.30 (M+Na), $C_{40}H_{41}NO_8SSiNa^+$ requires 746.90.

3,4-di-*O*-acetyl-2-deoxy-2-*N*-phthalimido-6-*tert*-butyl-diphenylsilyl-β-D-glucopyranose.



The thioglycoside **12** (700 mg, 0.95 mmol) and *N*-bromosuccinimide (516 mg, 2.90 mmol) were dissolved in acetone (20 ml) and water (20 ml). The reaction mixture was stirred at room temperature for 3 hours, then diluted with DCM (100 ml) and subsequently washed with water (100 ml), sodium thiosulphate (100 ml) and water (100 ml). The organic phase was dried with $MgSO_4$, filtered, and the solvent removed under reduced pressure. The mixture was purified by flash chromatography over silica (1: 1 petroleum ether/ ethyl acetate, R_f 0.4) to afford a white foam (69% yield). ^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) = 7.83-7.38 (14H, m, 14 x ArH), 5.83 (1 H, dd, H3) 5.48 (1 H,br s, H1), 5.20 (1 H, dd, H4), 4.24 (1 H, dd, H2), 3.82-3.71 (3H, m, H5, H6a, H6b), 3.16 (1 H, d, OH) 1.90, 1.85 (6H, s x 2, $COCH_3$), 1.07 (12H, S, CH_3 x 3). ^{13}C NMR (75 MHz, $CDCl_3$): δ (ppm) = 171.2, 170.3, 169.3, 167.9 (4 x RCOR) 135.7, 135.7, 134.4, 133.1, 133.0, 132.8, 131.6, 129.8, 129.0, 128.1, 127.8, 123.7 (4 x NPht + 8 x ArH), 83.1, 79.0, 72.1, 69.0, 62.7, 53.7, (C1-C6), 26.7 (3 x CH_3), 21.0, 20.5 (2 x $CO\underline{C}H_3$), 19.2 ($\underline{C}CH_3$). MS (ESI): *m/z* 654.21(M+Na), $C_{34}H_{37}NO_9SiNa^+$ requires 654.73.

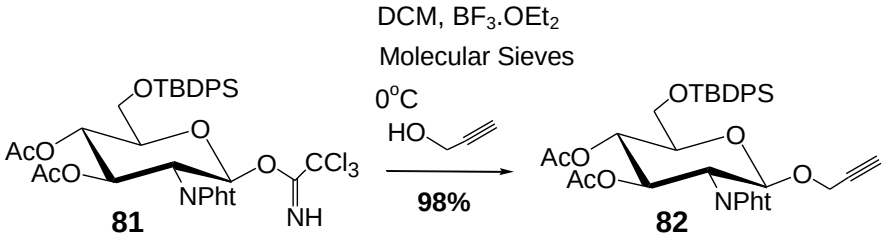
3,4-di-*O*-acetyl-2-deoxy-2-*N*-phthalimido-6-*tert*-butyl-diphenylsilyl-β-trichloroacetimidate-glucopyranoside.



The hemi-acetal **13** (390 mg, 0.61 mmol) was dissolved in anhydrous DCM (10 ml), together with trichloroacetonitrile (892 mg, 6.10 mmol), and potassium carbonate (256 mg, 1.85 mmol). The reaction was stirred under nitrogen for 5 hours, then filtered under suction and the solvent removed under reduced pressure. The reaction mixture was purified by flash chromatography over silica (3:1 petroleum ether/EtOAc, R_f 0.6) to afford a white foam (288 mg, 60%). ^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) = 8.56 (1 H, s, NH), 7.83-7.69 (4H, m,

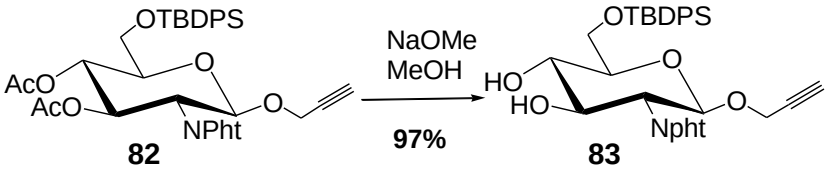
4 x ArH), 7.68-7.36 (10H, m, 10 x ArH), 6.69 (1H, d, *J* = 9 , H1), 5.69 (1H, dd, H3), 5.42 (1 H, dd, H4), 4.67 (1 H, dd, H2), 3.96-3.91 (1 H, m, H5) 3.84-3.76 (2H, m, H6a, H6b), 1.96,1.91 (6H, s x 2, $COCH_3$ X 2), 1.07 (9H, S, CH_3 x 3). ^{13}C NMR (75 MHz, $CDCl_3$): δ (ppm) = 170.3, 169.3, 167.9, 160.5 (4 x R<u>C</u>OR), 135.8, 135.7, 134.4, 133.2, 131.3, 129.7, 129.6, 127.7, 127.6, 123.7 (6 x NPht + 4 x ArH), 93.4, 90.4, 75.6, 71.0, 98.7, 62.3, 53.8 (C1-C6, <u>C</u>=NH), 26.7 (3 x CH_3), 20.6, 20.5 (2 x CO<u>C</u>H_3), 19.25 (<u>C</u>CH_3). MS (ESI): *m/z* 799.21 (M+Na), $C_{36}H_{37}Cl_3N_2O_9SiNa^+$ requires 799.79.

3,4-di-*O*-acetyl-2-deoxy-2-*N*-phthalimido-6-*tert*-butyl-diphenylsilyl-β-*O*-propargyl-glycopyranoside.



A solution of the trichloroacetimidate sugar **14** (176 mg, 0.23 mmol), propargyl alcohol (12 mg, 0.23 mmol) and activated 3 Å molecular sieves was stirred in anhydrous DCM (10 ml). The mixture was cooled to 0°C under nitrogen, and borontrifluoride diethyl etherate (7 µl, 0.06 mmol) was added. The ice bath was removed and the reaction mixture stirred for 16 hours. The reaction was monitored to completion by TLC (1/1 petroleum ether/ethyl acetate). The reaction was neutralised by dropwise addition of triethylamine, filtered and the solvent removed under reduced pressure. The product was purified by flash chromatography over silica (1:1 petroleum ether/EtOAc, R_f 0.5) to afford a yellow oil (172 mg, 98%). ^1H NMR (300 MHz, $CDCl_3$): δ (ppm) = 7.87-7.71 (4H, m, 4 x ArH), 7.70-7.38 (10H, m, 10 x ArH), 5.85 (1H, dd, H3), 5.58 (1H, d, *J* = 9, H1), 5.19 (1 H, dd, H4), 4.36-4.29 (3H, m, H2, CH_z), 3.86-3.74 (3H, m, H6a, H6b, H5), 2.31 (1 H, t, CH_2CC-H), 1.92, 1,84 (6H, s x 2, $COCH_3$ X 2) 1.07 (9H, s, CH_3 x 3). ^{13}C NMR (75 MHz, $CDCl_3$):δ (ppm) = 170.3, 169.4, 167.9, 163.5 (4 x R<u>C</u>OR), 135.8, 135.6, 134.2, 133.3, 133.1, 131.5, 129.8, 127.8, 127.7, 123.6 (6 x NPht + 4 x ArH), 95.7, 78.2, 75.4,74.7, 71.0, 69.3, 62.7 (C1-C6, -CH_2<u>C</u>CH), 55.7 (-<u>C</u>H_2CCH), 54.6 (CH), 26.7 (3 x CH_3), 20.6, 20.5 (2 x CO<u>C</u>H_3), 19.25 (<u>C</u>CH_3). MS (ESI): *m/z* 692.21 (M+Na), $C_{37}H_{39}NO_9SiNa$+ requires 692.22

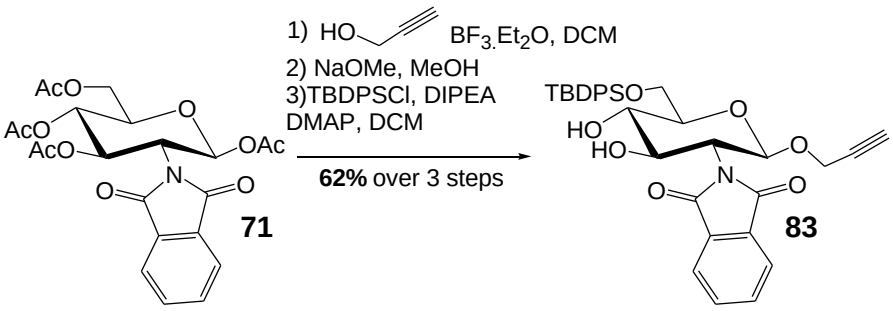
2-*N*-phthalimido-2-deoxy-6-*tert*-butyl-diphenylsilyl-β-*O*-propargyl-glucopyranoside.



82 **83**

The fully protected sugar **15** (164 mg, 0.25 mmol) was dissolved in methanol (10 ml) and NaOMe (0.5 M in MeOH, 0.3 ml) was added. The reaction was monitored by TLC (1/1 petroleum ether/EtOAc, R_f 0.3) to completion and diluted with DCM (20 ml), washed with $NaHCO_3$ (10 ml), and water (10 ml). The organic phase was collected and dried with $MgSO_4$, filtered and the solvent removed under reduced pressure to afford a white foam (140 mg, 97%). ^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) = 7.85-7.70 (4H, m, 4 x ArH), 7.70-7.40 (10H, m, 10 x ArH), 5.40 (1 H, d, *J* = 9, H1), 4.41 (1H, dd, H3), 4.23 (2H, s, CH_2), 4.14 (1H, dd, H4), 3.98-3.95 (2H, m, H2, H5), 3.71-3.61 (2H, m, H6a, H6b), 2.26 (1 H, s, CH_2CC-H), 1.07 (9H, S, CH_3 x 3). ^{13}C NMR (75 MHz, $CDCl_3$): *β* (ppm) = 168.4, 163.5 (2 x RCOR), 135.7, 135.6, 134.1, 132.7, 132.7, 131.8, 130.0, 127.9, 127.7, 123.4 (6 x NPht + 4 x ArH), 96.0, 78.5, 75.1, 74.5, 74.0, 71.5, 64.8 (C1-C6, -CH_2CCH), 56.1 (-CH_2CCH), 55.6 (CH), 26.8 (3 x CH_3), 19.2 (CCH_3). MS (ESI): *m/z* 608.25 (M+Na)+ , $C_{33}H_{35}NO_7SiNa^+$ requires 608.71.

2-*N*-phthalimido-2-deoxy-6-*tert*-butyl-diphenylsilyl-β-*O*-propargyl-glucopyranoside.

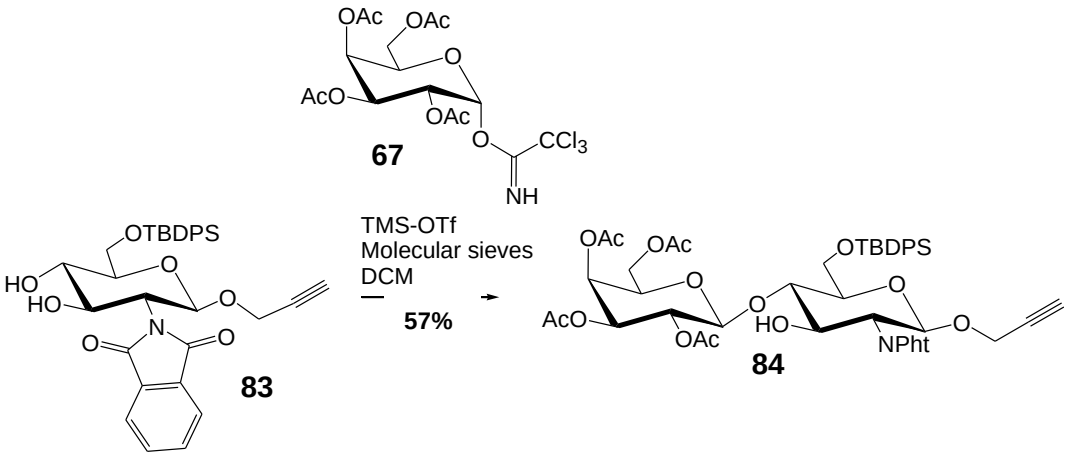


71 **83**

The fully protected sugar **2** (1.00 g, 2.09 mmol), propargyl alcohol (0.12 g, 2.09 mmol) and activated 3 Å molecular sieves were stirred in anhydrous DCM (30 ml). The mixture was cooled to 0^{o}C under nitrogen, and borontrifluoride diethyl etherate (64 µl, 0.52 mmol) was added. The ice bath was removed and the reaction mixture was stirred for 16 hours at room temperature. The reaction was monitored by HPLC, (C18 column) and MS. The reaction mixture was neutralised dropwise with triethylamine, and then filtered and solvent removed under reduced pressure. The crude mixture was dissolved in methanol (25 ml) and 0.5 M NaOMe in MeOH was added until pH 10 was obtained. The reaction was stirred for 16 hours and then neutralised with acetic acid. The mixture was then concentrated under reduced

pressure, azeotroped with toluene, and dried under high vacuum for 1 hour. TBDPSCl (768 µl, 2.91 mmol) DIPEA (0.50 g, 3.89 mmol), DMAP (0.02 g, 0.19 mmol), anhydrous DCM (10 ml), were added to the dried mixture under nitrogen. The reaction was stirred for 3.5 hours. The reaction was monitored by TLC (1:1 petroleum ether/EtOAc, R_f 0.3) to completion and diluted with DCM (20 ml), washed with $NaHCO_3$ (10 ml), and water (10 ml). The organic phase was dried with $MgSO_4$, filtered and the solvent removed under reduced pressure to afford a yellow oil. The oil was purified by flash chromatography over silica (1:1 petroleum ether/EtOAc, R_f 0.3) to afford the product a white foam (749 mg, 62%).^{1}H NMR (300 MHz, $CDCl_3$:δ (ppm) = 7.85-7.70 (4H, m, 4 x ArH), 7.70-7.40 (10H, m, 10 x ArH), 5.40 (1H, d, *J* = 9, H1), 4.41 (1H, dd, H3), 4.23 (2H, s, CH_2), 4.14 (1H, dd, H4), 3.98-3.95 (2H, m, H2, H5), 3.71-3.61 (2H, m, H6a, H6b), 2.26 (1H, s, CH_2CC-H), 1.07 (9H, S, CH_3 x 3).^{13}C NMR (75 MHz, $CDCl_3$): δ (ppm) = 168.4, 163.5 (2 x RCOR), 135.7, 135.6, 134.1, 132.7, 132.7, 131.8,130.0, 127.9, 127.7, 123.4 (6 x NPht + 4 x ArH), 96.0, 78.5, 75.1, 74.5, 74.0, 71.5, 64.8 (7C, C1-C6, -CH_2CCH), 56.1 (1 C, -CH_2CCH), 55.6 (C1, CH), 26.8 (3C, CH_3), 19.2 (1C, CCH_3). MS (ESI): *m/z* 608.25 (M+Na), $C_{33}H_{35}NO_7SiNa^+$ requires 608.71.

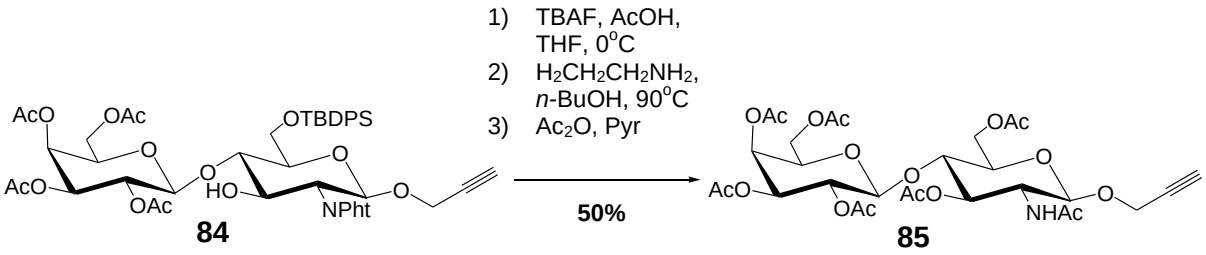
Disaccharide 84.



The trichloroacetimidate **8** (3.02 g, 6.15 mmol) and the glycoside acceptor **16** (1.8 g, 3.07 mmol) were dissolved in anhydrous DCM (100 ml), containing 4 Å molecular sieves at -20^{o}C, under nitrogen. TMS-OTF (0.68 g, 3.07 mmol) was added to the cooled reaction mixture at -20°C and stirred at -20°C for 2 hours. The reaction was then stirred at room temperature until TLC indicated that the reaction was complete. The mixture was diluted with DCM (50 ml) and neutralised with solid $NaHCO_3$, filtered and the solvent removed under reduced pressure. The mixture was purified by flash chromatography over silica (4: 1 toluene / ethyl acetate, R_t 0.3) to afford the product as a white foam (1.60 g, 57%).^{1}H NMR (400 MHz,

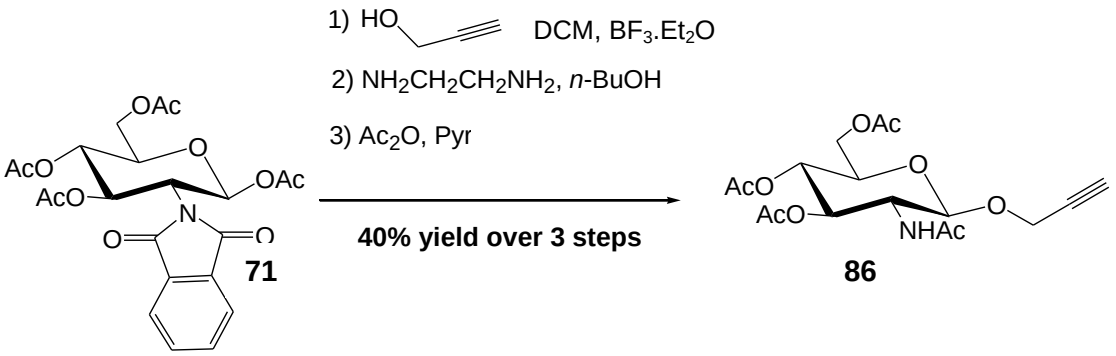
$CDCl_3$): δ (ppm) = 7.87-7.27 (14H, m, ArH), 5.46 (1 H, d, *J* = 9, H1), 5.37 (1 H, br s, H4), 5.20 (1 H, t, H2), 4.97 (1 H, dd, H3), 4.73 (1 H, d, *J* = 8, H1'), 4.50 (1 H, dd, H3'), 4.30 (2H, s, CH_2), 4.18 (1 H, dd, H2'), 4.00-3.83 (7H, m, H6a', H6b', H4', H6a, H6b, H5, H5'), 3.60 (1H, d, OH), 2.33 (1H, t, CH), 2.14, 2.04, 2.00, 1.72, (12H, s, $COCH_3$), 1.12 (12H, S, CH_3 x 3).^{13}C NMR (100 MHz, $CDCl_3$): δ (ppm) = 170.5, 170.1, 170.0, 169.1 (4 x RCOR), 136.0, 135.6, 134.0, 133.5, 132.6, 131.9, 130.0, 129.9, 127.9, 127.7 (6 x NPht + 4 x ArH) 101.3, 95.4, 81.0, 78.3, 75.2, 75.0, 71.2, 70.8, 69.3, 68.7, 66.8, 61.7, 61.3, 56.0, 55.3 (C1-C6', C1-C6, $\underline{C}H_2\underline{C}\underline{C}H$), 26.9 (3 x CH_3), 21.0, 20.5, 20.4, 20.3 (4 x $CO\underline{C}H_3$), 19.4 ($\underline{C}CH_3$). MS (ESI): *m/z* (M+Na) 938.9961 $C_{50}H_{55}NO_{15}SSiNa^+$ requires 938.9973.

Peracetylated *N*-acetyl lactosamine *β*-*O*-propargyl-glycoside.



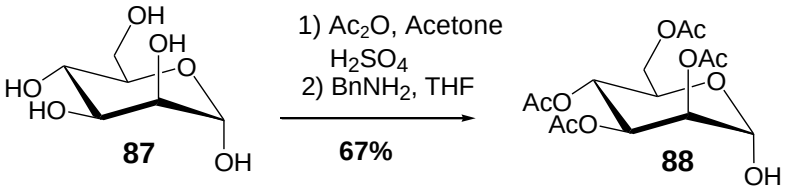
To a solution of the phthalimide protected disaccharide (827 mg, 0.90 mmol), in THF (20 ml) at 0°C , acetic acid (0.52 ml, 9.02 mmol) was added. A solution of 1 M tetrabutylammonium fluoride (TBAF) (1.05 ml, 3.61 mmol) in THF was added and the reaction mixture stirred at room temperature for 15 hours. The mixture was subsequently concentrated to a yellow oil. The oil was dissolved in *n*-BuOH and ethylenediamine (0.27 ml, 3.61 mmol) and the reaction stirred at 90°C under reflux for 15 hours. The resulting orange solution was reduced to an oil and dissolved in 2:1 pyridine/acetic anhydride (45 ml) and stirred for 24 hours. The reaction was finally concentrated to give an orange oil and purified by flash chromatography over silica (100%, EtOAc, R_t 3.2) to afford the product as an orange solid (300 mg, 50%). ^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) = 5.74 (1 H, d, *J* = 10, NH), 5.34 (1 H, d, *J* = 3, H4), 5.12-5.05 (2H, m, H2, H3'), 4.92 (1H, dd, H3), 4.67 (1H, d, *J* = 8, H1'), 4.50-4.48 (2H, m, H6a', H1), 4.33 (2H, s, CH_2), 4.13-4.05 (4H, m, H2', H6b', H6a, H6b), 3.92-3.84 (1 H, m, H5), 3.79 (1 H, d, *J* = 9, H4'), 3.64 (1 H, m, H5'), 2.44 (1 H, t, CCH), 2.13, 2.10, 2.06, 2.04, 2.03, 1.98, 1.97 (21 H, s, $COCH_3$) ^{13}C NMR (500 MHz, $CDCl_3$): δ (ppm) = 170.8, 170.5, 170.4, 170.4, 170.2, 170.1, 168.4 (7 x RCOR), 101.1, 98.4, 78.5, 75.7, 75.4, 72.9, 72.5, 70.9, 70.8, 69.2, 66.7, 62.2, 60.9, 55.7, 53.0 (C1'-C6', C1-C6, -CH_2CCH), 23.3 ($NHCOCH_3$), 20.7 (6 x $COCH_3$). MS (ESI): m/z 696.30 (M+Na), $C_{29}H_{39}NO_{17}Na^+$ requires 696.60.

3,4,6-tri-*O*-acetyl-2-deoxy-2-*N*-acetyl-β-*O*-propargyl-glucopyranoside..



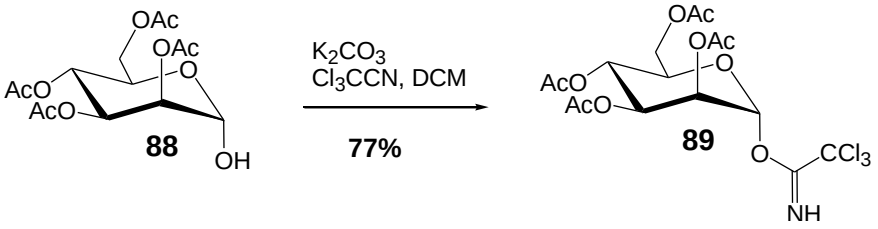
The fully protected sugar (0.25 g, 0.53 mmol), propargyl alcohol (32 µl, 0.53 mmol) and activated 3 Å molecular sieves were stirred in anhydrous DCM (5 ml). The mixture was cooled to 0^{o}C under nitrogen, and borontrifluoride diethyl etherate (64 µl, 0.53 mmol) was added. The ice bath was removed and the reaction mixture stirred for 16 hours at room temperature. The reaction mixture was monitored by LC-MS to completion and then diluted with DCM (10 ml), washed with water (10 ml), $NaHCO_3$ (10 ml), NaCl (10 ml) and water (10 ml). The organic phase was dried with $MgSO_4$, filtered and the solvent removed under reduced pressure to afford a yellow oil. The oil was dissolved in *n*-BuOH (10 ml) and ethylenediamine (116 µl, 1.57 mmol) and the reaction stirred at 90^{o}C under reflux for 16 hours. The resulting orange solution was reduced to an oil and dissolved in 2:1 pyrindine/acetic anhydride (30 ml) and stirred for 5 hours. The reaction mixture was diluted with ethyl acetate (40 ml), washed with water (30 ml), $NaHCO_3$ (30 ml), NaCl (30 ml) and water (30 ml). The organic phase was dried with $MgSO_4$, filtered and the solvent removed under reduced pressure to afford an orange oil. The oil was purified by flash chromatography over silica (100% EtOAc, R_f 0.4) to afford the product a white solid (80 mg, 40% yield over 3 steps).^{1}H NMR (300 MHz, $CDCl_3$:δ (ppm) = 5.62 (1H, d, *J* = 9, HN), 5.30 (1H, dd, H3), 5.11 (2H, dd, H4), 4.86 (1H, d, *J* = 8, H1), 4.37 (2H, s, CH_2), 4.29-4.24 (1H, dd, H6a), 4.15-4.11 (1H, dd, H6b), 3.99 (1H, dd, H2), 3.75-3.69 (1H, m, H5) 2.27 (1H, s, CH_2CC-H), 2.08, 2.02, 2.02, 1.95 (12H, S, CH_3 x 3).^{13}C NMR (75 MHz, $CDCl_3$): δ (ppm) = 170.9, 170.7 170.3, 169.4 (4 x RCOR), 98.3, 78.5, 75.4, 72.4, 72.0, 68.5, 61.9 (7C, C1-C6, -CH_2C$\equiv$CH), 55.9 (1C, -C$\equiv$CH_2CCH), 54.3 (C1, CH), 23.4, 20.7, 20.7, 20.6 (4C, CH_3). MS (ESI): *m/z* 408.30 (M+Na), $C_{17}H_{23}NO_9Na^+$ requires 408.36.

2,3,4,6-tetra-*O*-acetyl-α-*D*-mannopyranose [3].



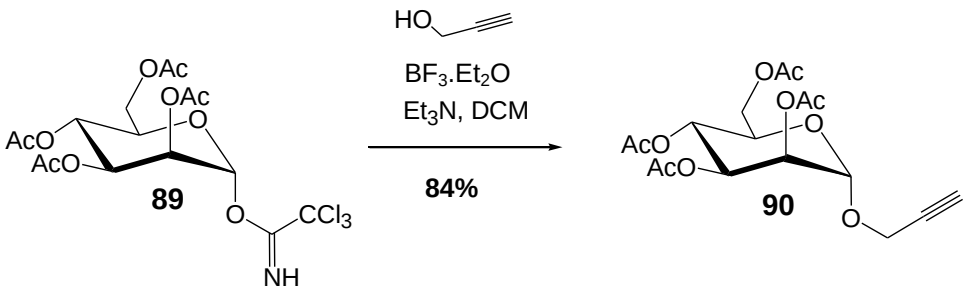
Mannose (3.0 g, 16.64 mmol) was dissolved in 1:1 acetic anhydride/acetone (50 ml) and cooled to 0°C. Concentrated H_2SO_4 (3.2 ml) was added dropwise over 30 minutes. The reaction was allowed to come to room temperature by removal of the ice bath and stirring was continued overnight. After 16 hours the reaction mixture was poured into ice cold water (250 ml) and the aqueous phase extracted with DCM (3 x 100 ml), washed with $NaHCO_3$ (200 ml), water (200 ml) and dried with $MgSO_4$. The organic phase was then filtered, the solvent removed under reduced pressure and purified by flash chromatography over silica (1/1 petroleum ether/EtOAc, R_f 0.3) to afford the paracetylated mannose as a white foam (3.00 g, 70%). This was forwarded to the next step. The paracetylated sugar (3.00 g, 7.68 mmol) in anhydrous THF (40 ml) was treated with benzyl amine (1.09 ml, 9.99 mmol) under nitrogen. The reaction mixture was warmed to 50°C and stirring was continued for 16 hours under nitrogen. The solvent was then removed under reduced pressure and the crude product purified by flash chromatography over silica (1: 1 hexane: diethyl ether) to afford a yellow oil (1.68 g, 67%),^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) = 6.17 (1 H, d, *J* = 4, H1), 5.82 (1H, d, *J* = 9, OH), 5.24 (1H, dd, H3), 5.18 (1H, dd, H4), 4.50 (1H, dd, H2), 4.31-4.10 (3H, m, H6a, H6b, H5), 2.10, 2.03, 2.02, 1.96, (12H, s x 4, $COCH_3$ X 4). ^{13}C NMR (75 MHz, $CDCl_3$): δ (ppm) = 170.3, 170.1, 169.9, 168.9, 166.4 (5C, 5 x RCOR), 89.7, 68.7, 67.4, 67.3, 66.4, 61.2 (6C, C1-C6), 20.1, 20.9, 20.6, 20.5, (4C, $CO\underline{C}H_3$). MS (ESI): m/z 371.29 (M+Na), $C_{14}H_{20}NaO_{1O}^+$ requires 371.29.

2,3,4,6,-tetra-*O*-acetyl-1-trichloroacetimidate-α-*D*-mannopyranoside [3].



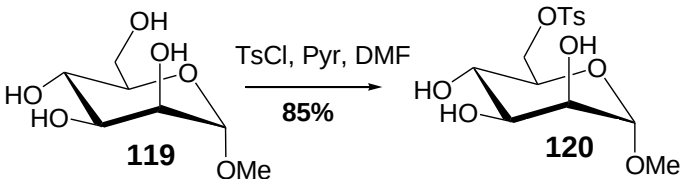
To a solution of the hemi-acetal (1.6 g, 4.59 mmol) in anhydrous DCM containing 3 Å molecular sieves, anhydrous potassium carbonate (1.91 g, 13.7 mmol) and trichloroacetonitrile (4.65 ml, 45.96 mmol) was added. The reaction was stirred under nitrogen for 5 hours and filtered with suction. The solvent was removed under reduced pressure and purified by flash chromatography over silica (3/1 EtOAc/petroleum ether, R_t 0.3) to afford a the product as a white foam (1.6 g, 77%). ^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) = 8.66 (1H, s, NH), 6.55 (1H, d, *J* = 4, H1), 5.52 (1H, d, *J* = 2, H4), 5.36 (1 H, dd, H3), 5.23 (1 H, dd, H2), 4.38 (1 H, m, H5), 4.16-4.01 (2H, m, H6a, H6b), 2.13, 2.02, 2.00,1.98, (12H, s x 4, $COCH_3$ x 4). ^{13}C NMR (75 MHz, $CDCl_3$): δ (ppm) = 170.1 (4C, 4 x RCOR), 93.6, 69.0, 67.5, 67.4, 66.9, 61.3 (C1-C6), 20.6, ($4CH_3$, $CO\underline{C}H_3$). MS (ESI): *m/z* 492.62 (M+Na), $C_{16}H_{20}Cl_3O_{10}Na^+$ requires 492.69.

2,3,4,6-tetra-*O*-acetyl-α-*O*-propargyl-mannopyranoside [3].



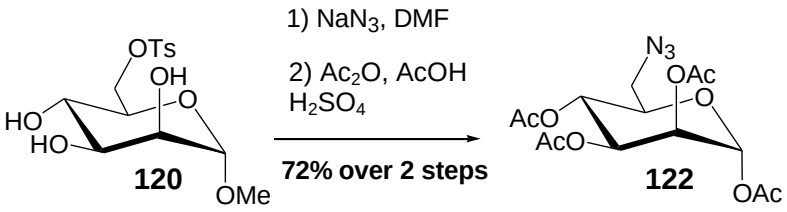
To a solution of the trichloroacetimidate (1.17 g, 3.50 mmol) in anhydrous DCM (10 ml) under nitrogen, powdered 4 Å molecular sieves and propargyl alcohol (0.20 ml, 350 mmol) were added. The resulting mixture was cooled to 0^{o}c, and a solution of $BF_3.OEt_2$ (0.11 ml, 0.87 mmol) in anhydrous DCM (1.6 ml) was added dropwise for 10 minutes. The stirring was continued at room temperature overnight. Then the reaction mixture was neutralised, by dropwise addition of Et_3N, and the reaction filtered through celite. The filtrate was concentrated under reduced pressure and purified by flash chromatography over silica (1: 1 ethyl acetate/ petroleum ether R_f:0.3) to afford the product as a white foam (1.3 g, 84%). ^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) = 5.29 - 5.21 (3H, m, H4, H2, H3), 4.98 (1H, d, *J* = 2, H1), 4.25- 4,21 (3H, m, CH_2, H6a), 4.04 (1 H, dd, H6b), 4.96 (1 H, m, H5), 2.45 (1 H, t, CH), 2.09, 2.04, 1.99, 1.92 (12H, s x 4, $COCH_3$ x4). ^{13}C NMR (125 MHz, $CDCl_3$): δ (ppm) = 170.7, 170.0, 169.9, 169.7 (4qC, 4 x RCOR), 96.3, 78.0, 76.1, 75.7, 69.3, 69.0, 98.7, 62.3, 55.0 (C1-C5, CH_2 x 2, CH), 20.9, 20.8, 20.7, 20.7 ($4CH_3$, $CO\underline{C}H_3$). MS (ESI): *mlz 409.22* (M+Na), $C_{17}H_{22}NaO_{10}$+ requires 409.11.

6-*O*-Toluensulphonyl-α-methyl-*D*-mannopyranoside [4].



A solution of the α-methyl-*O*-mannopyranoside (3.36 g, 17.3 mmol) in anhydrous pyridine (20.00 ml) was cooled to 0°C in an ice bath with stirring. A solution of *p*-toluenesulphonyl chloride (3.61 g, 19.0 mmol) in anhydrous pyridine (15.00 ml) was then added dropwise over 30 minutes and stirring was continued at 0°C until the reaction was complete by TLC. The solvent was then removed under reduced pressure to afford a crude product, which was purified by flash chromatography over silica (15:2 DCM/MeOH) to afford 6-*O*-toluensulphonyl-α-methyl-*D*-mannopyranoside (4.99 g, 85%) as a colourless oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) = 7.77 (2H, d, *J* = 9, 2 x ArH), 7.29 (2H, d, *J* = 8 Hz, 2 x ArH), 4.62 (1H, br s, H1), 4.28 (2H, br s, H4, H5), 3.88 (1 H, br s, H2), 3.68 (3H, m, H3, CH_2), 3.24 (3H, s, OCH_3), and 2.38 (3H, S, CH_3). ^{13}C NMR (63 MHz, $CDCI_3$): δ (ppm) = 144.6 (qC, Ts), 132.3, 129.6, 127.7 (3CH, Ts), 100.7 (CH, C1), 71.3, 70.1, 69.8, 69.6, 66.7, 54.5 (2CH, C2-C5), 30.4 (OCH_3), and 21.3 (CH_3). MS (ESI): *m/z* 370.80 (M+Na), $C1_8H_{20}O_8NSNa^+$ requires 371.40.

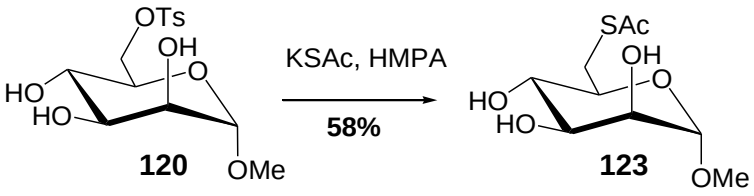
1,2,3,4-tetra-*O*-acetyl-6-azido-6-deoxy-α-*D*-mannopyranoside [5].



To a solution of anhydrous DMF (20 ml) and NaN_3 (4.50 g, 69.0 mmol), was added to a solution of the tosylate (3.45 g, 9.90 mmol) in anhydrous DMF. The solvent had been dried over molecular sieves under nitrogen for 1 hour prior to use. The reaction was heated to 100°C with stirring behind a blast shield and heating continued at 100°C for 6 hours. Upon cooling to room temperature, the colourless precipitate was removed by filtration through a plug of silica under suction and the solvent was evaporated under reduced presure. The solid was dissolved in DCM, filtered and evaporated under vacuum. The product was

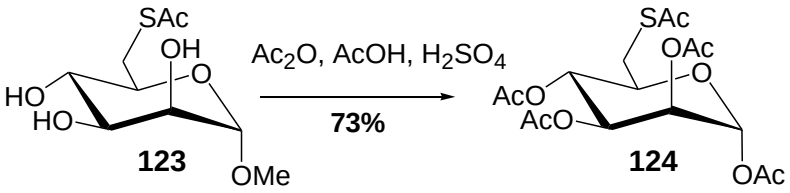
purified by flash chromatography over silica (15:2 DCM/MeOH) to afford 6-azido-6-deoxy-α-methyl-*D*-mannopyranoside as a colourless syrup. This methyl glycoside (2.45 g, 11.00 mmol) was dissolved in (1: 1 AC_2O/AcOH) (206 ml) and cooled to 0°C. Concentrated H_2SO_4 (2.20 ml) was added dropwise over 30 minutes, with stirring. The reaction was then allowed to come to room temperature by removal of the ice bath and stirring was continued overnight. After 26 hours the reaction mixture was poured into ice water (250 ml) and the aqueous phase was extracted with DCM (3 x 100 ml). The combined organic extracts were washed with cold water (200 ml), sat. aq. $NaHCO_3$ (200 ml) and dried with $MgSO_4$. The solvent was removed under reduced pressure to afford the crude product which was purified by flash chromatography over silica (60:40 pet.ether/EtOAc) to afford the peracetylated product 1,2,3,4-tetra-*O*-acetyl-6-azido-6-deoxy-α-*D*-mannopyranoside as a colourless oil (2.93 g, 72% over 2 steps).^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) = 6.04 (1 H, d, *J* = 2, H1), 5.22-5.29 (2H, m, H3, H4), 5.23 (1 H, m, H2), 3.94 (1 H, m, H5), 3.36 (1 H, dd, H6a), 3.25 (1 H, dd, H6b), 2.14, 2.13, 2.02, 1.99 (12H, s x 4, $COCH_3$). ^{13}C NMR (75 MHz, $CDCl_3$): δ (ppm) = 170.5, 169.6, 169.5, 168.0 (qC, 4 x $\underline{C}OCH_3$), 90.6, 70.8, 70.2, 67.6, 62.9 (5CH, C1-C5), 56.6 (CH_2), 20.7 (4CH_3, $CO\underline{C}H_3$). MS (ESI): *m/z* 395.90 (M+Na), $C_{14}H_{19}N_3O_9Na^+$ requires 396.30.

6-thioacetate-6-deoxy-α-methyl-*D*-mannopyranoside.



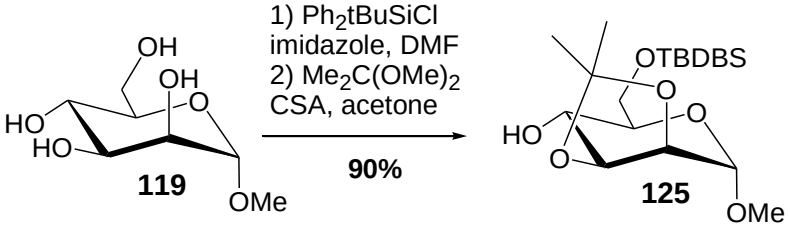
To a solution of HMPA (20.0 ml) and potassium thioacetate (0.98 g, 8.61 mmol), a solution of the tosylate (1.00 g, 2.87 mmol) was added at room temperature. The reaction was monitored by TLC (7:1 DCM/MeOH). The mixture was diluted with DCM (50 ml) washed with water (10 ml), HCl (10 ml) and water (10 ml), and dried with $MgSO_4$. The organic phase was then filtered, concentrated under reduced pressure and purified by flash chromatography over silica (7:1, DCM/MeOH, R_f 0.3) to afford the product as a white foam (0.41 g, 58%).^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) = 4.67 (1H, br s, H1), 3.94 (1H, br s, H2), 3.62-3.53 (2H, m, H4, H5), 3.32 (4H, m, H6a, OCH_3), 3.19 (1H **,** dd, H6b), 2.38, (3H, S, CH_3).^{13}C NMR (75 MHz, $CDCl_3$): δ (ppm) = 198.3 (1C, SCOR), 100.1, 77.2, 71.1, 70.5, 69.8, 55.0 (6C, C1-C6), 31.0, 30.6 (2C, OCH_3, $SOCH_3$). MS (ESI): *m/z* 275.11 (M+Na), $C_9H_{16}SNaO_6^+$ requires 275.27.

1,2,3,4,-tetra-*O*-acetyl-6-thioacetate-6-deoxy-α-*D*-mannopyranoside.



The methyl glycoside (1.80 g, 7.14 mmol) was dissolved in 1:1 Ac_2O/AcOH (20 ml) and cooled to 0°C. Concentrated H_2SO_4 (2.2 ml) was added dropwise over 30 minutes with stirring. The reaction was then allowed to come to room temperature by removal of the ice bath and stirring was continued overnight. After 26 hours the reaction mixture was poured into ice water (250 ml) and the aqueous phase was extracted with DCM (3 x 100 ml). The combined organic extracts were washed with cold water (200 ml), sat. aq. $NaHCO_3$ (200 ml) and dried with $MgSO_4$. The solvent was removed under reduced pressure to afford the crude product which was purified by flash chromatography over silica (1: 1 petroleum ether/EtOAc R_f: 0.2) to afford the peracetylated product as a yellow oil (2.12 g, 73%). ^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) = 6.22 (1 H, d, *J* = 3, H1), 5.22 (3H, m, H2, H3, H4), 3.95 (1 H, m, H5), 3.19 (1H, dd, H6a), 3.06 (1H, dd, H6b), 2.33, 2.17, 2.15, 2.11, 1.98, (15H, s x 5, $ROCH_3$ X 5).^{13}C NMR (75 MHz, $CDCl_3$): δ (ppm) = 194.8 (qC, S<u>C</u>OCH_3) 170.0, 169.8, 169.8, 168.1, (4qC, 4 x <u>C</u>OCH_3), 90.5, 77.2, 71.8, 68.7, 68.4, 67.7 (5CH, CH_2, C1C6), 30.4, 30.1, 20.8, 20.8, 20.6, (5CH_3, RO<u>C</u>H_3). MS (ESI): *m/z* 429.10 (M+Na), $C_{16}H_{22}SNaO_{10}^+$ requires 429.08.

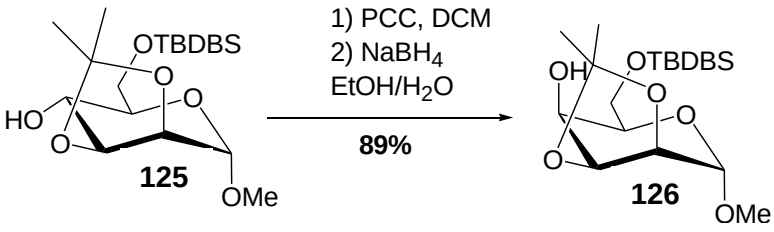
Methyl *6-O-tert*-butyl-diphenylsilyl-2,3-*O*-isopropylidene-α-*D*-mannopyranoside [6].



Tert-butylchlorodiphenylsilane (10.43 g, 37.9 mmol) was added dropwise to a stirred solution of the methyl α-*D*-mannopyranoside (6.70 g, 34.5 mmol) and imidazole (5.17 g, 75.9 mmol) in DMF (140 ml) under nitrogen at room temperature. After 2 hours, TLC (1:9 methanol/EtOAc) showed the loss of starting material (R_f 0.1) and the formation of a major product (R_f 0.6). The solvent was removed under reduced pressure and the residue

dissolved in chloroform (100 ml), and washed with HCl 0.1 M (100 ml), then water (100 ml). The aqueous fractions were re-extracted with chloroform (2 x 30 ml). The organic fractions were combined, dried over $MgSO_4$, filtered, and the solvent was removed under reduced pressure. The residue (12.0 g, 27.8 mmol) was dissolved in a solution of (1:9 2,2-dimethoxypropane in acetone) (125 ml) and the resulting solution adjusted to pH 4 using camphorsulfonic acid. The reaction solution was stirred under nitrogen at room temperature. After 2 hours, TLC (1:3 hexane/EtOAc) showed the loss of starting material and the formation of a major product. The solution was neutralised using aqueous ammonia solution and the solvent removed under reduced pressure, the residue dissolved in chloroform (100 ml), and washed with water (2 x 100 ml). The aqueous fractions were re-extracted with chloroform (2 x 30 ml). The organic fractions were combined, dried over $MgSO_4$, filtered, and the solvent was removed under reduced pressure. The residue was purified by flash chromatography (1:3 hexane/EtOAc) to give the desired product (11.5 g, 90%) as a colourless oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) = 7.69 (4H, m, ArH), 7.26 (6H, m, ArH), 4.88 (1 H, s, H1), 4.10 (2H, m, H2, H3), 3.89 (2H, m, H6a, H6b), 3.78 (1 H, m, H4), 3.59 (1 H, m, H5), 3.34 (3H, s, OCH_3), 2.75 (1 H, d, *J* = 4, OH), 1.51, 1.35 (6H, s x 2, CH_3), 1.07 (9H, s, CH_3). ^{13}C NMR (75 MHz, $CDCl_3$): δ (ppm) = 135.7, 135.6, 134.8, 133.1, 133.0, 129.9, 127.8,127.8 (CH, Ar), 109.5 (qC, $\underline{C}(CH_3)_2$), 98.3, 78.2, 75.3, 70.6, 69.5 (5CH, C1-C5), 64.7 (CH_2), 54.9 (OCH_3), 27.9 (CH_3, C($\underline{C}H_3$)), 26.8 ($3CH_3$, SiC($\underline{C}H_3$)), 26.1 (CH_3, C($\underline{C}H_3$)), and 19.2 (qC, Si$\underline{C}$(CH_3h). MS (ESI): *m/z* 495.5 (M+Na), $C_{26}H_{36}O_6SiNa^+$ requires 495.6.

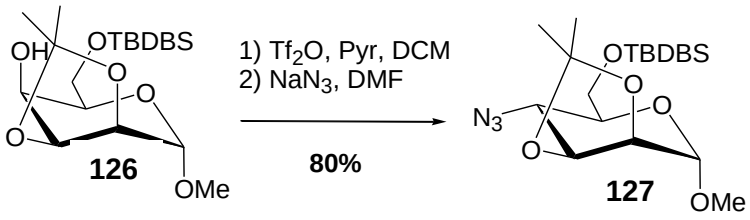
Methyl 6-*O*-*tert*-butyl-diphenyl-silyl-2,3-*O*-isopropylidene-α-talopyranoside [6].



The methyl α-*O*-mannopyranoside (3.88 g, 8.21 mmol), powdered dried molecular sieves (5.61 g) and pyridinium chlorochromate (5.42 g, 25.2 mmol) were stirred in anhydrous DCM (27.00 ml) under nitrogen at room temperature for 5 hours. At this point TLC (1:2 ether/hexane) showed complete conversion of the starting material (R_f 0.3) to a single product (R_f 0.5). The reaction mixture was triturated with ether (50 ml), then filtered through a plug of silica topped with celite (ether eluant). The solvent was removed under reduced pressure, the residue dissolved in (9:1 EtOH/water), and cooled to 0^{o}C. Sodium borohydride (570 mg) was added and the solution allowed to warm slowly to room temperature. After 3 hours TLC (1:3 EtOAc/hexane) showed the formation of a single product (R_f 0.4) and the

complete consumption of the starting material (R_f 0.5). The reaction was quenched with ammonium chloride until effervescence ceased. The solvent was removed under reduced pressure and the residue dissolved in chloroform (40 ml), and washed with water (2 x 40 ml). The organic fraction was dried over $MgSO_4$, filtered, and the solvent was removed under reduced pressure. The residue was purified by flash chromatography (1:3 EtOAc/hexane) to give the desired product (3.47 g, 90%) as a colourless oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) = 7.72 (4H, m, Ar), 7.37 (6H, m, Ar), 5.08 (1H, d, H1), 4.21 (1H, dd, H3), 4.08 (1H, dd, H2), 3.95 (1H, dd, *J* = 7, H6a), 3.94 (1H, dd, *J* = 7, H6b) 3.84 (2H, m, H4, H5), 3.41 (3H, s, OCH_3), 2.36 (1 H, d, *J* = 7, OH), 1.59, 1.36, (6H, s, 2 x $C(CH_3)$), and 1.10 (9H, S, $SiC(CH_3)$). ^{13}C NMR (75 MHz, $CDCl_3$): δ (ppm) = 135.7 (CH, Ar), 135.6 (CH, Ar), 134.9 (qC, Ar), 129.8 (CH, Ar), 127.8 (CH, Ar), 109.4 (qC, $\underline{C}(CH_3h)$, 98.5, 73.8, 72.8 ,69.5, 64.1 (5CH, C1-C5), 63.4 (CH_2), 54.0 (OCH_3), 26.9 ($3CH_3$,$SiC(\underline{C}H_3)$), 25.9, 25.2 ($2CH_3$, $C(\underline{C}H_3)_2$, and 19.2 (qC, $Si\underline{C}(CH_3)_3$. MS (ESI): *m/z* 495.5 (M+Na), $C_{26}H_{36}O_6SiNa^+$ requires 495.6.

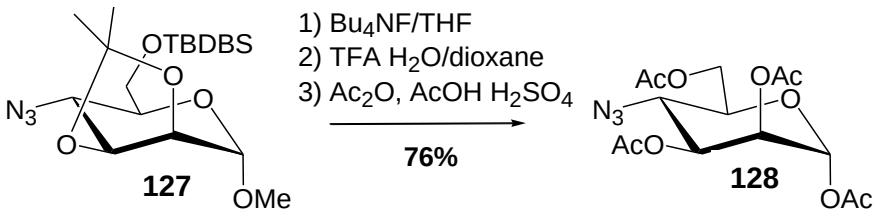
Methyl 4-azido-6-*O*-*tert*-butyl-diphenyl-silyl-4-deoxy-2,3-*O*-isopropylidene-*α*-*D*-mannopyranoside [6].



The methyl-α-*D*-mannopyranoside (3.94 g, 8.32 mmol) was dissolved in anhydrous DCM (80 ml) under nitrogen. Pyridine (2.7 ml, 33.28 mmol) was added and the solution was stirred at – 35°C for 30 minutes. Triflic anhydride (1.62 ml, 9.65 mmol) was added dropwise with stirring and the mixture allowed to reach -10°C. After 2 hours TLC (1:3 ether/hexane) showed complete conversion of the starting material (R_f 0.4) to a single product (R_f 0.5). Methanol (1 ml) was added to quench the reaction and the resulting solution warmed to room temperature. The solution was washed with water (100 ml), pH 7 PBS solution (100 ml), and the aqueous layer re-extracted with chloroform (30 ml x 2). The organic fractions were combined, dried with $MgSO_4$, filtered, and the solvent was evaporated under reduced pressure. The residue was dissolved in DMF (31 ml) under nitrogen and sodium azide (2.16 g, 33.28 mmol) was added. The resulting mixture was heated at 90°C for 3 hours with stirring. The mixture was then allowed to cool down to room temperature, filtered through a plug of silica (DCM as eluant), and the solvent was removed under reduced pressure. The residue was dissolved in chloroform (100 ml), washed with distilled water (100 ml x 2), and the aqueous layers were re-extracted with chloroform (30 ml x 2). The organic fractions were

combined, dried with $MgSO_4$, filtered and the solvent removed under reduced pressure. The crude product was purified by flash chromatography (1:3 ethyl acetate/hexane, R_f 0.6) to give the desired product (3.30 g, 80%) as a colourless oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm)= 7.79-7.72 (4H, m, Ar), 7.46-7.93 (6H, m, Ar), 5.01 (1 H, d, H1), 4.21 (1 H, dd, *J* = 8, *J* = 5, H3), 4.19 (1 H, dd, *J* = 5, H2), 3.88 (2H, m, H6a, H6b), 3.75 (1H, dd, *J* = 11, *J* = 9, H4), 3.41 (1 H, m, H5) 3.36 (3H, s, OCH_3), 1.60, 1.41 (6H, s, $C(CH_3)_2$), and 1.12 (9H, s, $SiC(CH_3)_3$). ^{13}C NMR (75 MHz, $CDCl_3$): δ (ppm) = 135.9, 135.7 (CH x 2, Ar), 133.4, 133.2 (2x qC, Ar), 129.7, 129.7 127.8, 127.7 (4CH, Ar), 110.0 (qC, $\underline{C}(CH_3)_2$), 98.2, 76.7, 75.0, 68.9, (CH, C1, C2, C3, C5), 63.3 (CH_2), 60.7 (CH, C4), 54.9 (OCH_3), 28.2 ($2CH_3$, $C(\underline{C}H_3)_2$), 26.8 ($3CH_3$, $SiC(\underline{C}H_3)_3$), and 19.4 (qC, $Si\underline{C}(CH_3)_3$). MS (ESI): *m/z* 520.23 (M+Na), $C_{26}H_{35}N_3O_5SiNa^+$ requires 520.22.

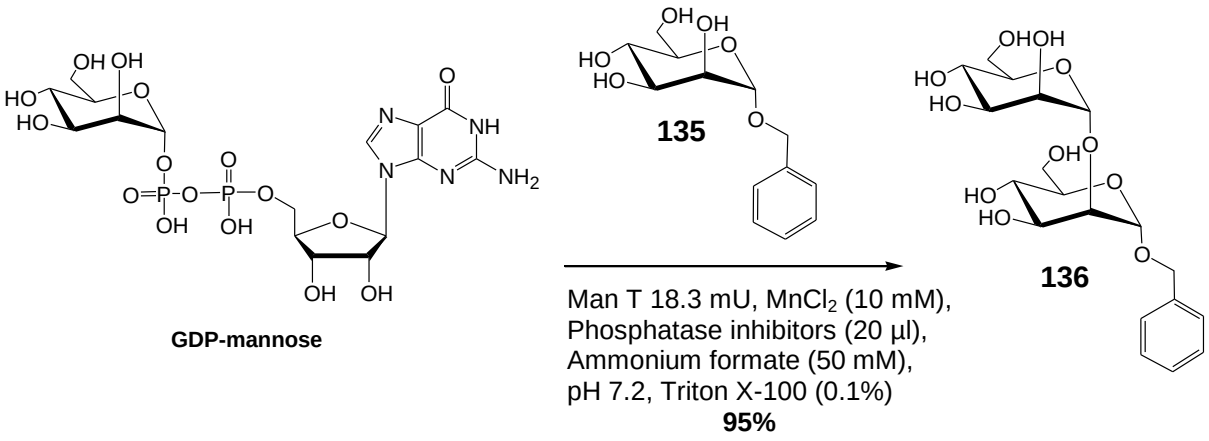
1,2,3,6-tetra-*O*-acetyl-4-azido-4-deoxy-α-*D*-Mannopyranoside [6].



The methyl α-*D*-mannopyranoside **31** (1.12 g, 2.25 mmol) was dissolved in anhydrous THF (12.5 ml) under nitrogen. A 10 M solution of tetrabutylammonium fluoride (TBAF) in THF was slowly added (4.4 ml), and stirring was continued under nitrogen. After 1 hour TLC (3:1 ether/hexane) showed complete conversion of the starting material (R_f 0.5) to a single product (R_f 0.1). The solvent was evaporated under reduced pressure. The residue was dissolved in chloroform (20 ml), washed with distilled water (20 ml x 2), and the aqueous layers were re-extracted with chloroform (10 ml x 2). The organic fractions were combined, dried with $MgSO_4$, filtered and the solvent was evaporated under reduced pressure. The crude product (750 mg, 2.90 mmol) was dissolved in 0.3 ml 1,4-dioxane and aqueous TFA (1:1 TFA/water, 10.00 ml), and stirring was continued at room temperature. After 1 hour TLC (3:1 ether/hexane) showed complete conversion of the starting material (R_f 0.1) to a single product (R_f ~0). The solvent was co-evaporated with toluene (5 ml x 2) to dryness under reduced pressure. The residue was dissolved in acetic anhydride (16.50 ml), and the solution was cooled down to 0°C. Concentrated H_2SO_4 (0.50 ml) was added dropwise and the reaction was allowed to come to room temperature by removal of the ice bath. Stirring was continued overnight. The reaction mixture was then diluted with chloroform (10.00 ml), cooled to 0°C with an ice bath, and neutralised with sat. aq. $NaHCO_3$ which was slowly

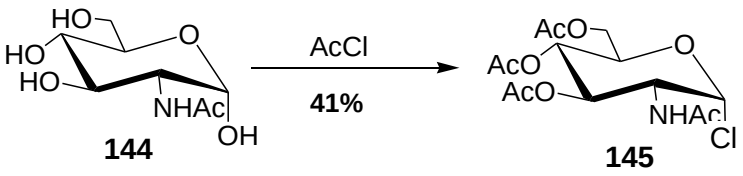
added. The organic layer was washed with distilled water (20 ml x 2), dried $MgSO_4$ and concentrated under reduced pressure. The resulting mixture was purified by flash chromatography over silica (3:2 petroleum ether/EtOAc, R_f 0.3) to afford a pale yellow oil (820 mg, 76% over three steps). ^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) = 6.09 (1H, d, *J* = 1.9 Hz, H1), 5.19 (1H, dd, H3), 5.16 (1H, dd, H2), 4.28 (1H, dd, H6a), 4.23 (1 H, dd, H6b), 3.84 (1 H, dd, H4), 3.75 (1 H, m, H5), 2.10, 2.08, 2.05, and 2.02 (12H, S, $COCH_3$). ^{13}C NMR (75 MHz, $CDCl_3$): δ (ppm) = 170.5, 169.6, 169.5, 168.0 (4qC, $\underline{C}OCH_3$), 90.6, 70.8, 70.2, 67.5 (3 CH, C1, C2, C3, C5), 62.9 (CH_2), 56.6 (CH, C4), 20.7 ($4CH_3$, $CO\underline{C}H_3$). (ESI): *m/z* 396.10 (M+Na), $C_{14}H_{19}N_30_9Na^+$ requires 396.10.

Molecule 136.



To a 1.5 ml Eppendorf tube, GDP-mannose (3.2 mg, 100 mM), mannose α-benzyl glycoside (1000 mM), α-1,2 mannosyltransferase (91.5 mU), ammonium formate buffer (50 mM) at pH 7.2, Triton X-100 (0.1%), $MnCl_2$ (10 mM) and phosphatase inhibitors (20 µl) were added to make up a total volume of 500 µl. The reaction mixture was shaken at 30^{o}C for 48 hours then diluted with water (200 µl) and extracted with DCM (100 µl). The aqueous phase was collected and purified by reverse phase HPLC using C18 column at 254nm (RT: 17.9, gradient 5-95% (water/acetonitrile) over 30 minutes). The purified disaccharide was lyophilised to give the product as a white foam (2 mg, 95%, R_f 0.6, 12:12:4 MeOH:$CHCl_3$:H_2O). ^{1}H NMR (500 MHz, D_2O): δ (ppm) = 7.32 (5H, m, ArH), 5.07 (1, s, H1 '),4.91 (1H, s, H1), 4.32 (1H, d, *J* = 12, CH_2), 3.98 (1H, dd, H2), 3.88 (1H, dd, H2'), 3.81 (1H, dd, H3'), 3.78 (2H, dd, H6a, H6b), 3.73 (1H, dd, H3), 3.68 (2H, dd, H6a, H6b), 3.65-3.51 (4H, m, H4', H5', H4, H5). ^{13}C NMR (125.77 MHz, D_2O): δ (ppm) = 134.0, 129.2, 129.0, 128.8, 170.1 (6 x ArH), 100.6, 100.3, 73.2, 73.3, 70.9, 70.8, 70.4, 70.3, 69.7, 67.2, 67.1, 64.5, 61.2, (C1-C6, C1'-C6', CH_2) MS (ESI): *m/z* (M+Na) 455.1526, $C_{19}H_2SO_{11}Na^+$ requires 455.1529.

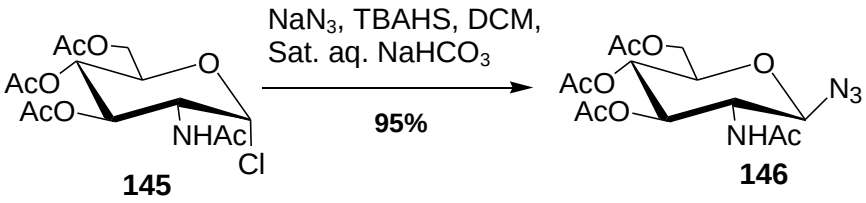
2-Acetamido-2-deoxy-3,4,6-tri-*O*-acetyl-*D*-glucopyranoside chloride [1].



144 **145**

N-acetyl glucosamine (10.10 g, 43.25 mmol) was added to acetyl chloride (20 ml) with stirring and the mixture stirred overnight at room temperature. The mixture was then diluted with DCM (80 ml) and poured into ice-water (40 ml) with stirring. The organic layer was added to saturated sodium bicarbonate solution and ice (80 ml) with stirring until neutral (gas evolution stops). The organic phase was extracted, dried over $MgSO_4$ and evaporated to give a yellow oil. The crude product was purified by column chromatography on silica (ethyl acetate 100% R_f 0.2) to afford the desired product (6.70 g, 41%). 1H NMR (300 MHz, $CDCl_3$):δ ppm = 6.19 (1H, d, *J* = 3, H1), 5.77 (1 H, d, *J* = 9, NH), 5.35-5.19 (2H, m, H3, H4), 4.54 (1 H, m. H2), 4.23 (2H, m, H6a, H5), 4.10 (1 H, dd, H6b), 2.24-1.98 (12H, 3s, 4 x CH_3). Found m/z (ESI): 365.9. $C_{14}H_{21}NO_8Cl$ requires *m/z 366.1.*

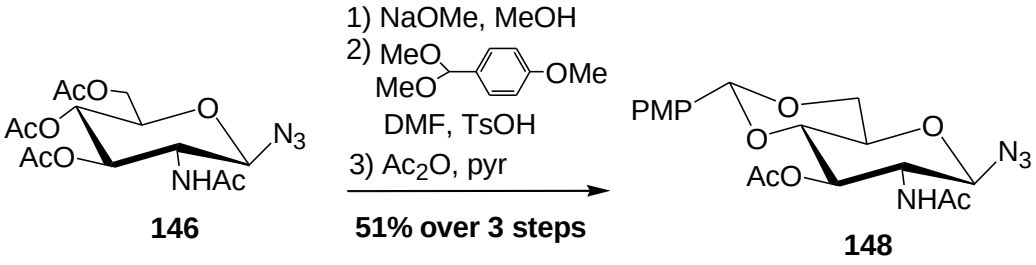
2-Acetamido-2-deoxy-3,4,6-tri-*O*-acetyl-β-*D*-glucopyranosyl azide [1].



145 **146**

2-Acetamido-2-deoxy-3,4,6-tri-*O*-acetyl-α-*O*-glucopyranosyl chloride (6.54 g, 17.66 mmol), sodium azide (3.61 g, 55.55 mmol) and tetrabutylammonium hydrogen sulphate (6.14 g, 18.08 mmol) were dissolved in DCM (70 ml). Saturated sodium bicarbonate was then added and the reaction mixture stirred vigorously for 1.5 hours. The reaction mixture was then diluted with ethyl acetate (120 ml) and extracted a further two times with ethyl acetate (100 ml), brine (100 ml) and water (100 ml), then dried over $MgSO_4$ and the solvent removed under reduced pressure to give 2-acetamido-2-deoxy-3,4,6-tri-*O*-acetyl-β-*D*-glucopyranosyl azide as a white solid (6.21 g, 95%).^{1}H NMR (300 MHz, $CDCl_3$):δ ppm = 5.57 (1 H, d, *J* = 9, NH), 5.24 (1 H, d, *J* = 9, H3), 5.10 (1 H, dd, H4), 4.76 (1 H, d, H1), 4.26 (1 H, dd. H6a), 4.18 (1 H, dd, H6b), 3.90 (1 H, dd, H2), 3.80 (1 H, m, H5), 2.08-1.98 (12H, 4 x s, 4 X CH_3). Found *m/z* (ESI): 395.10 (M+Na), $C_{14}H_{21}N_4O_8Na^+$ requires *m/z 395.11.*

2-Acetamido-2-deoxy-3-*O*-acetyl-4-6-*O*-*p*-methoxybenzylidene-β-*D*-glucopyranosyl azide [1].

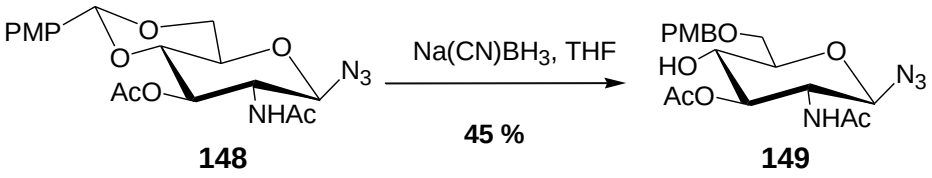


146 **51% over 3 steps** **148**

2-Acetamido-2-deoxy-3,4,6-tri-*O*-acetyl-β-*D*-glucopyranosyl azide **35** (5.01 g, 13.44 mmol) was dissolved in anhydrous MeOH (25 ml) and sodium methoxide (300 µl of 0.5 M solution in MeOH) was added and the reaction mixture was stirred for 2 hours at room temperature. The reaction mixture was the neutralised by the addition of acetic acid (10 µl and then concentrated under reduced pressure to afford 2-acetamido-2-deoxy-3,4,6-tri-*O*-hydroxy-β-*D*-glucopyranosyl azide as a crude pale yellow foam (3.19, 96%), The crude was forwarded on to the next step without further analysis and purification. Found m/z (ESI): 246.00 (M+Na), $C_6H_{14}N_4O_5$ requires 246.22. The crude azide was dissolved in anhydrous DMF (20 ml) and *p*-anisaldehyde dimethyacetal (5.64 g, 30.96 mmol) and *p*-tosic acid (0.46 g, 2,42 mmol) were added. After stirring for 1.5 hours at 50°C the reaction mixture was concentrated under reduced pressure and the residue poured into a cold mixture of saturated sodium bicarbonate (50 ml) and DCM (50 ml). The mixture was cooled to 4°C for 10 minutes and the remaining precipitate was filtered off and then crystallised with ethyl acetate. The solid was dried under vacuum to afford the crude 2-acetamido-2-deoxy-4,6-*O*-*p*-methoxybenzylidene-β-*D*-glucopyranosyl azide as a white solid. Found *mlz* (ESI): 407.40 (M+Na), $C_{17}H_{22}N_4O_6Na^+$ requires 407.38. The crude was forwarded on to the next step without further analysis and purification. Due to its very poor solubility the crude acetal (3.11 g, 8.53 mmol) was dissolved in pyridine (30 ml) and acetic anhydride (15 ml) was added. The mixture was stirred for 24 hours at room temperature then concentrated under vacuum. DCM (150 ml) was then added and the organic phase was washed with water (20 ml), sat. aq. $NaHCO_3$ (20 ml) and water (20 ml). The organic layer was dried with Na_2SO_4 and the solvent removed under vacuum and crystallised from ethyl acetate to afford 2-acetamido-2-deoxy-4,6-*O*-*p*-methoxybenzylidene-β-*D*-glucopyranosyl azide as a white solid (2.76 g, 51% over three steps). ^{1}H NMR (300 MHz, $CDCl_3$): δ ppm = 7.35 (2H, d, J = 8.7, ArH), 6.89 (2H,d, *J* = 9, ArH), 5.68 (1 H, d, *J* = 9, NH), 5.47 (1 H, s. ArCH), 5.24 (1H, dd, H3), 4.57 (1H, d, *J* = 9.2, H1), 4.37 (1H, dd. H6a), 4.11 (1H, dd, H2), 3.80 (3H, s, OCH_3), 3.80 - 3.55 (3H, m, H6a, H4, H5), 2.08 (3H, s, CH_3), 1.92 (3H, s, CH_3) Found *mlz* (ESI): 387.10 (M+Na),

$C_{14}H_{21}N_4O_8Na^+$ requires *mlz 387.12*

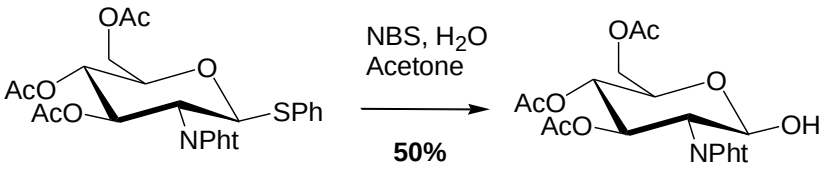
2-Acetamido-2-deoxy-3-*O*-acetyl-6-*O*-*p*-methoxybenzyl-β-*D*-glucopyranosyl azide [1].



148 **149**

Trifluoroacetic acid (5.73 g, 50.25 mmol) in anhydrous DMF (30 ml) was cooled to 0^{o}C and added dropwise to a mixture of *p*-methoxybenzylidene acetal (2.01 g, 4.90 mmol), sodium cyanoborohydride (1.66 g, 26.40 mmol) and 4 Å molecular sieves in anhydrous DMF (40 ml) at 0°C. After the addition was complete, the ice bath was removed and the reaction mixture stirred at room temperature for 16 hours. The reaction mixture was then filtered with suction and the filtrate poured into ice cold sat. aq. $NaHCO_3$. The mixture was extracted with DCM (6 x 60 ml), and the organic layers combined and washed with sat. aq. $NaHCO_3$ (80 ml). The organic layer was dried with $MgSO_4$ and the solvent removed under reduced pressure. The crude product was purified by flash chromatography over silica (100% etheyl acetate) and 2-acetamido-2-deoxy-3-*O*-acetyl-6-*O*-*p*-methoxybenzyl-β-*D*-glucopyranosyl azide was isolated as a white solid (0.90 g, 45%) ^{1}H NMR (300 MHz, $CDCl_3$):δ ppm = 7.25 (2H, d, *J* = 8.6, ArH), 6.88 (2H, d, *J* = 9, ArH), 6.05 (1 H, d, *J* = 9, NH), 5.08 (1H, dd, H3), 4.57 (1H, d, *J* = 9, H1), 4.55 (2H, d. CH_2), 3.94 (1H, m, H2), 3.76 (3H, s, OCH_3), 3.71-3.69 (3H, m, H6a, H6b, H4), 3.60 (1 H, m, H5), 2.16 (3H, s, CH_3), 1.96 (3H, s, CH_3) Found m/z (ESI): 389.53 (M+Na), $C_{14}H_{21}N_4O_8Na^+$ requires m/z 389.14.

2-Phthalimido-2-deoxy-3,4,6-tri-*O*-acetyl-β-*D*-glucopyranose [1].

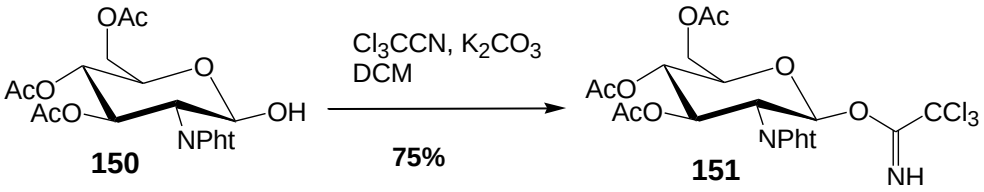


72 **150**

The thiophenyl glycoside **3** (4.5 g, 8.54 mmol) and *N*-iodosuccinimide (3.19 g, 14.2 mmol) was dissolved in acetone (20 ml) and water (2 ml). The mixture was stirred for 3.5 hours at room temperature, then diluted with DCM (100 ml) and washed with water (100 ml), saturated sodium thiosulphate (100 ml) and again with water (100 ml). The organic layer was dried over Na_2SO_4 and the solvent removed under reduced pressure to afford the hemi-

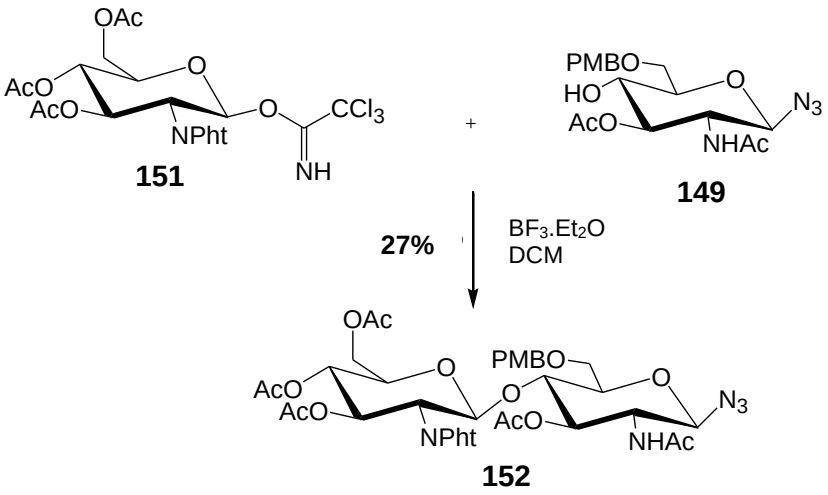
acetal as a crude yellow oil. The product was purified by flash chromatography over silica (ethyl acete/ petroleum ether 1: 1) to afford a white solid (50%).^{1}H NMR (300 MHz, $CDCl_3$):δ ppm = 7.87-7.81 (4H, m, ArH), 5.82 (1 H, d, H1), 5.60 (1 H, dd, H2), 5.16 (1 H, dd, H4), 4.30-4.11 (3H, m, H2, H6b, H6b), 3.96-3.90 (1 H, m, H5), 2.12, 2.03, and 1.82 (9H, 3 x s, $COCH_3$) Found *m/z* (ESI): 458.30 (M+Na), $C_{20}H_{21}NO1_0Na^+$ requires m/z 438.37.

_O_-(2-phthalimido-2-deoxy-3,4,6-tri-_O_-acetyl-β-_D_-glucopyranosyl)-trichloroacetimidate[1].



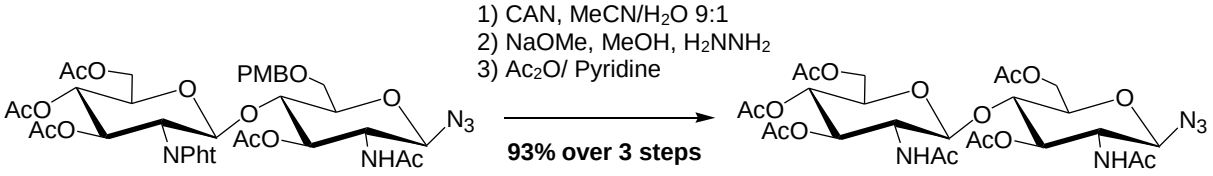
A solution of hemi-acetal **39** (0.7 g, 1.61 mmol) and trichloroacetimidate (3.3 ml, 16.07 mmol) in anhydrous DCM containing 3 Å molecular sieves was added to a solution of potassium carbonate (0.64 g, 4.82 mmol). The reaction mixture was stirred under nitrogen at room temperature for 5 hours, filtered with suction and the solvent removed under reduced pressure. The product was purified by flash chromatography over silica (3: 1 ethyl acetate/ petroleum ether, R_f 0.3) and was isolated as a white solid (0.71 g, 75%). ^{1}H NMR (300 MHz, $CDCl_3$): δ ppm = 7.87-7.73 (4H, m, ArH), 6.64 (1H, d, *J* = 9, H1), 5.95 (1 H, dd, *J* = 9, H3), 5.32 (1 H, dd, *J* = 10, H4), 4.64 (1 H, dd, J = 9, H2), 4.43 (1 H, dd, H6b), 4.09 (1 H, dd, H6a) 4.06 (1 H, m, H5), 2.17, 2.05, and 1.87 (9H, 3 x s, $COCH_3$) Found *m/z* (ESI): 601.07 (M+Na), $C_{22}H_{21}C_{13}N_2O_{10}Na^+$ requires *m/z* 601.01.

2-*N*-Phthalimido-2-deoxy-3,4,6-tri-*O*-acetyl-β-*D*-glucopyranosyl-(1-4)-2-acetamido-2-deoxy-3-*O*-acetyl-6-*O*-*p*-methoxybenzyl-β-*D*-glucopyranosyl azide [1].



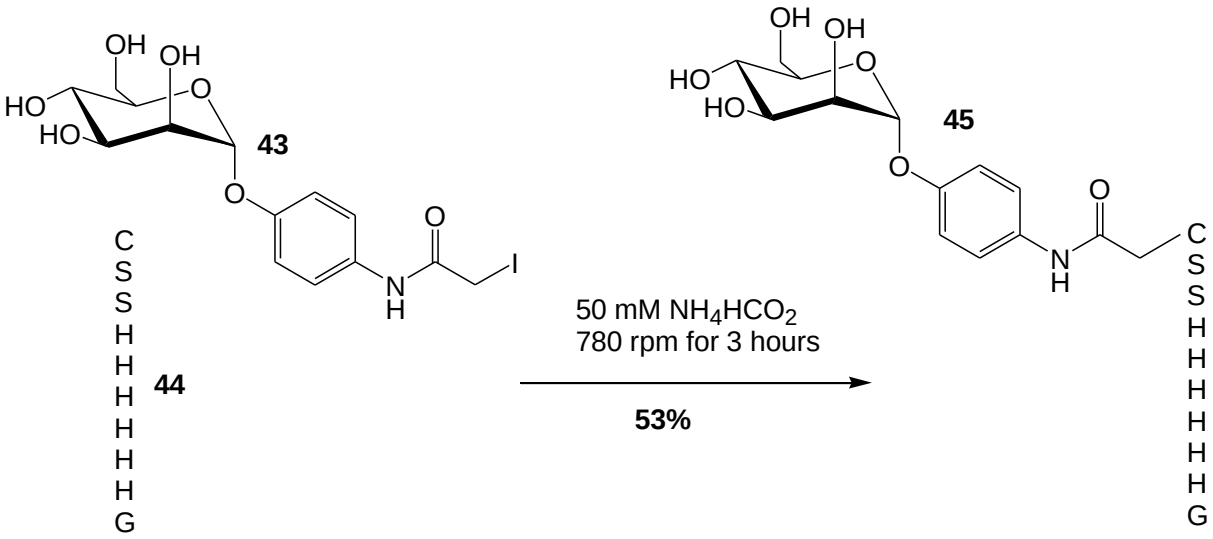
A solution of the azide **149** (0.28 g, 0.71 mmol), trichloroacetimidate **151** (0.40 g, 0.71 mmol) and 3 Å molecular sieves in anhydrous DCM (10 ml) was cooled down to 0°C in an ice bath, under nitrogen. Borontrifluoride diethyletherate (20.0 µl, 0.18 mmol) was added. The reaction was stirred for 16 hours and further borontrifluoride diethyletherate (20.0 µl, 0.18 mmol) was added. After 4 hours the reaction was diluted with DCM (50 ml) and filtered with suction. The filtrate was washed with sat. aq. $NaHCO_3$ (25 ml) and water (25 ml). The organic phase was dried over sodium sulphate, the drying agent was filtered off and the solvent was removed under reduced pressure. The resulting white solid was purified by flash chromatography over slica (100% ethyl acetate, R_f 0.36) and the product was isolated as a white solid (0.16 g, 28%). ^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) = 7.94-7.69 (4H, m, ArH), 7.18 (2H, d, ArH), 6.83 (2H, d, ArH), 6.16 (1 H, d, *J* = 10, NH), 5.71 (1 H, dd, H3), 5.39 (1 H, d, *J* = 8, H1), 5.09 (1 H, dd, H4), 4.99 (1 H, dd. H3) 4.39-4.32 (3H, m, Hi, CH_2OPMB), 4.15-3.99 (3H, m, 3 x CH), 3,79 (3H, s, OCH_3), 3.61-3.29 (4H, m, 4 x CH), 2.07, 2.02, 1.97, 1.93 and 1.79 (H15, 5 x s, $ROCH_3$). ^{13}C NMR (100 MHz, $CDCl_3$): δ (ppm) = 171.2, 170.3, 169.9, 169.3, 158.9 (qC x 5), 134.5 (ArCH), 1.9 (qC), 129.8 (qC), 129.3 (2 x ArCH(PMB)), 123.6 (ArCH), 113.7 (2 x ArCH(PMB)), 97.1 (CH), 88.6 (CH), 76.6 (CH), 73.1 (CH), 72.7(CH), 72.5 (OCH_2Ar), 71.7(CH), 70.5 (CH), 68.4 (CH), 66.9 (CH_2OPMB), 61.5 (CH_2), 55.3 (OCH_3), 54.9 (CH), 53.4 (CH), 23.2, 20.8, 20.7, 20.6 and 20.4 (5C CH_3). MS (ESI): *m/z* 848.70 (M+Na), $C_{38}H_{43}N_5O_{16}Na^+$ requires 848.76,

2-Acetamido-2-deoxy-3 ,4, 6-tri-*O*-acetyl-β-*D*-glucopyranosyl-(1→4)-2-acetamido-2-deoxy-3, 6-di -*O*-acetyl-β-*D*-glucopyranosylazide [1].



152 **153**

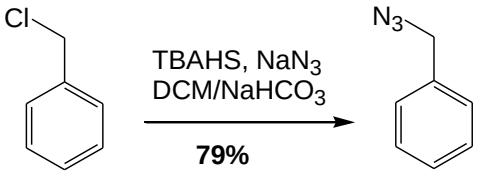
The phthalimide protected disaccharide **41** (134 mg, 0.19 mmol), was dissolved in a solution of (1:9 acetonitrile/water) (1 ml). Ceric ammonium nitrate (212 mg, 0.38 mmol) was added and the reaction mixture was stirred at room temperature for 3 hours until TLC showed that the reaction was complete. The mixture was diluted with DCM (25 ml) and washed with sat. aq. $NaHCO_3$ (10 ml). The organic phase was dried over $MgSO_4$ and the solvent removed under reduced pressure to produce a yellow gum which was purified by flash chromatography over silica (100% EtOAc, R_f 0.3) to afford a white foam. The white foam was dissolved in anhydrous NaOMe in MeOH (0.5 M, 30 µl) and the reaction was stirred for 4 hours. The reaction was then neutralised with acetic acid (3 µl) and the solvent removed under reduced pressure. The residue was dissolved in ethanol (200 µl) and hydrazine hydrate (92 µl, 1.90 mmol) was added and the reaction monitored by TLC to completion. The resulting fully deprotected sugar was suspended in pyridine, acetic anhydride (2:1) total of (6 ml) and the mixture stirred for 16 hours. The solvent was removed under reduced pressure and the mixture was dissolved in DCM (50 ml), washed with 2 M HCl (25 ml), saturated sodium bicarbonate (25 ml) and water (25 ml). The organic phase was dried over sodium sulphate, the drying agent was filtered off and the solvent was removed under reduced pressure. The resulting yellow oil was purified by flash chromatography over slica (95: 1 ethyl acetate/MeOH, R_f 0.2) and the product was isolated as a white solid (107 mg, 93%), ^1H (400 MHz; $CDCl_3$): δ (ppm) = 6.24 (1H, d, *J* = 9, NH), 6.49 (1H, d, *J* = 10, NH), 5.23-5.02 (3H, m, 3 x CH), 460-4.55 (2H, m, 2 x CH), 4.41-4.36 (2H, m, 2 x CH), 3.86-3.84 (1H, m, CH), 3.75-3.64 (3H, m, 3 x CH), 2.17-1.94 (14H, m, 7 x CH_3). ^{13}C NMR (100 MHz, $CDCl_3$): δ (ppm) = 171.2, 171.0, 170.8, 170.6, 170.6, 170.5, 169.4 (7qC, $\underline{C}OCH_3$), 101.3, 88.6, 75.7, 75.0, 72.4, 72.4, 71.9, 68.0, 54.7, 53.1 (10CH, C1 -C5), 62.1, 61.7 ($2CH_2$), 20.7 ($7CH_3$, $RO\underline{C}H_3$). Found *mlz* (ESI): 682.58 (M+Na) $C_{26}H_{37}N_5O_{15}Na^+$ requires 682.59. Ir δ max = 2112.1 (N_3).



Molecule 156.

The His_6-tag peptide **155** (16.10 mg, 0.014 mmol) was measured into a 2.5 ml Eppendoff tube and 50 mM N **154**)$_2$ (2 ml) was added. The iodoacetamide sug**a** **156** (7.2 mg, 0.016 mmol) was added and the reaction mixture was shaken at 780 rpm for 2 hours. The reaction mixture was monitored by MS, and after 1 hour the mixture was concentrated under reduced pressure. The mixture was lyophilised and placed under high vacuum for 2 days. HPLC was performed with (5-95%) acetonitrile/water (0.1% TFA), with C18 column (Luna 10µ C18 100A Size: **155** 10.00 mm, 10 micron). The sample was dissolved in 100 µl of water prior to injection, the product eluted at 13.10 RT at 280 nm, while the iodoacetamide sugar eluted at 14.64 RT. The product was lyophilised and placed under high vacuum for 1 day to a yield a white foam (10.8 mg, 53%). MS (ESI): *m/z* 1486.82 (M+H), $C_6,H_{80}N_{24}0_{19}S^+$ requires 1485.50.

Benzyl azide.

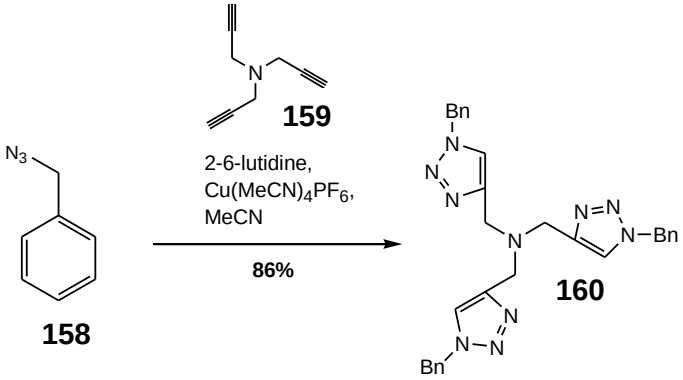


157 **158**

To a solution of benzyl chloride (1.80 ml, 15.80 mmol), TBAHS (5.36 g, 15.80 mmol) and sodium azide (3.08 g, 47.48 mmol) in DCM (40 ml) and $NaHCO_3$ (40 ml) was added. The resulting biphasic solution was stirred vigorously at room temperature for 16 hours. Ethyl acetate (100 ml) was added and the organic layer was separated, washed with sat. aq. $NaHCO_3$ (2 x 40 ml) and NaCl (2 x 10 ml). The organic phase was dried over $MgSO_4$ and

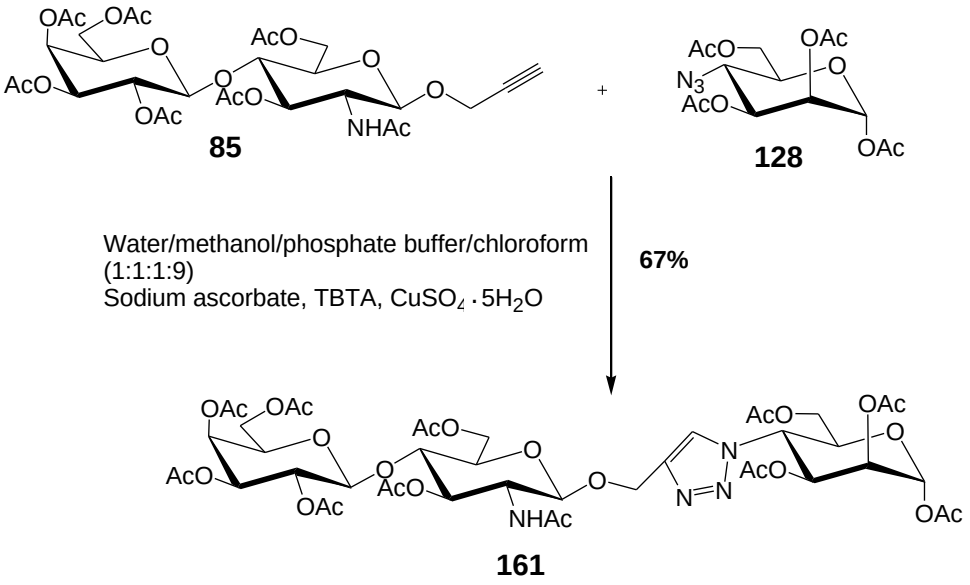
the solvent removed under reduced pressure to produce a yellow oil that was used without further purification (1.65 g, 79%).^{1}H NMR (300 MHz, $CDCl_3$): δ ppm = 7.44 (5H, m, ArH), 4.35 (2H, s ,CH_2), Found *m/z* (ESI): 157.50 (M+Na). $C_7H_7N_3Na^+$ requires *m/z* 156.48. Ir δ max = 2096^{-1} (N_3).

Tris((1-benzyl-1-H-1,2,3-triazol-4-yl)methyl)amine [7].



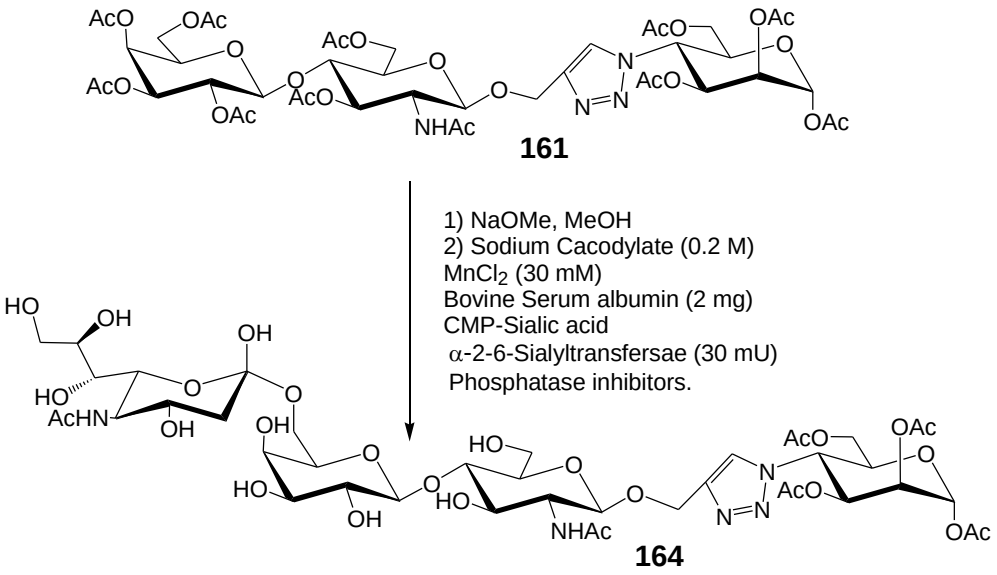
Tripropargylamine (0.23 g, 1.67 mmol) was dissolved in acetonitrile (4 ml). Benzyl azide (1.00 g, 7.51 mmol), 2-6-lutidine (0.18 g, 1.68 mmol) was added together with $Cu(MeCN)_4PF_6$ (0.02 g). Upon addition of the salt the reaction was cooled in an ice bath. The brown solution was stirred for 3 days. A pale white precipitate formed which was filtered off and washed with cold acetonitrile. This crude product (0.7 g) was recrystalised from hot 1:1 *tert*-butyl alcohol/water (30 ml). After filtration and washing with water (2 x 10 ml), the pale pink solid was dried under high vacuum overnight (86%). ^{1}H NMR (300 MHz, $CDCl_3$) δ ppm = 7.68 (3H, s, 3 x NCH=C), 7.36-7.27 (15h, m, 3 x Bn), 5.50 (6H, s, 3 x $BnCH_2N$), 3.72 (6H, s, CCH_2N). MS (ESI): *m/z* 532.69 (M+), $C_{30}H_{31}N_{10}^+$ requires 531.63.

Molecule 161.



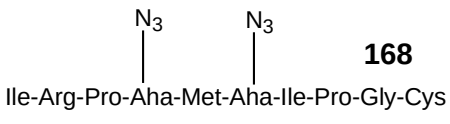
In a 1.5 ml Eppendorf tube propargyl LacNAc **85** (20 mg, 0.03 mmol) and peracetylated 4-azidomannose **128** (22 mg, 0.06 mmol) were added to a solution of water/methanol/phosphate buffer (50 mM)/CI_3CH (1: 1: 1 :9). Sodium ascorbate (5.8 mg, 0.03 mmol), TBTA (1 mg) and $CuSO_4.5H_2O$ (1 mg) was added to the reaction mixture and the reaction was stirred for 5 hours at 30^oC. Next the mixture was diluted with CI_3CH (5 ml) and washed with water (2 ml). The organic phase was collected and the solvent removed under reduced pressure. The crude product was purified by flash chromatography over silica (1:1 ethyl acetate/petroleum ether, then 10% methanol) to afford the product as a white foam (21 mg, 67%).^{1}H NMR (500 MHz, $CDCI_3$): δ (ppm) = 7.73 (1H, br s, C<u>H</u>N), 6.18 (1H, S, H1''), 6.10 (1H, d, *J* = 8, H1'), 5.74 (1H, d, *J* = 10, NH), 5.34 (2H, dd, H4, H2''), 5.08 (1H, dd, H3''), 4.91 (5H, m, H2, H3, H3'', CH_2), 4.56 (1H, br s, H1), 4.49 (5H, m, H4'', H3, H6a', H6a'', H6b''), 4.05 (5H, m, H2', H6b',H6a, H6b, H5''), 3.82 (2H, m, H5, H4'), 3.82 (1H, br s, H5'), 2.22, 2.15, 2.13, 2.13, 2.08, 2.05, 2.04, 1.99, 1.98, 1.89, 1.85 (33H, s x 11, $COCH_3$). ^{13}C NMR (500 MHz, $CDCI_3$): δ (ppm) = 171.0, 170.6 170.5, 170.4, 170.3, 170.3, 170.2, 169.5, 169.5, 169.2, 168.1 (11 x RCOR), 144.8, 125.1, 101.1, 99.9, 90.7, 77.1, 75.6, 72.9, 72.5, 70.9, 70.8, 69.1 ,68.1, 67.5, 66.7, 62.4, 62.3, 62.2, 60.8, 56.7, 53.0 (C1-C6', C1-C6, C1-C6'', -<u>C</u>H$_2$<u>CC</u>HN), 23.3 (NH<u>C</u>OCH$_3$), 23.3 (NH<u>C</u>OCH$_3$), 20.7 (6 x CO<u>C</u>H$_3$). MS (ESI): *m/z* 1069.2939 (M+Na), $C_{43}H_{58}NO_{26}Na^+$ requires 1069.3232.

Molecule 164.

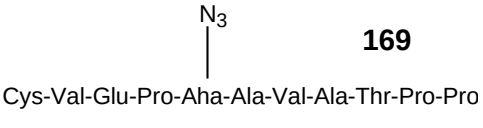


The trisaccharide **161** (21 mg) was deprotected with (0.5 M) NaOMe in anhydrous MeOH. The reaction was then neutralised with acetic acid and the solvent removed under reduced pressure. The residue was dissolved in a total of 1 ml buffer solution at pH 7.1 containing sodium cacodylate (0.2 M), manganese chloride tetrahydrate (30 mM), Bovine Serum albumin (2 mg), phosphatase inhibitors (10 µl, Sigma Aldrich P5726), CMP-sialic acid (10 mg), and α-2-6-sialyltransferase (30 mU). The reaction mixture was shaken for 70 hours at 37 °C. Next the mixture was diluted with water, and ion exchange HPLC using a Sax column (5 µ, 100 A, 250 x 10 mm) was performed (0.6 M ammonium formate / water) gradient 10 minutes at 10%. UV at 254nm was measured and the fractions containing the product (retention time 7.395 minutes) were collected and lyophilised before analysis by MS. MS (ESI): *m/z* 916.3194 (M^-), ($C_{34}H_{54}N_5O_{24}^-$) requires 916.3159.

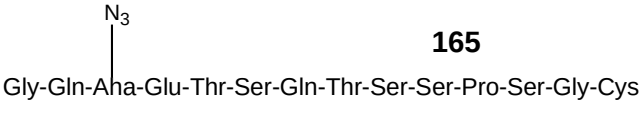
6.4 Aha modified peptides and click chemistry



Modified peptide 401-409 of ADG



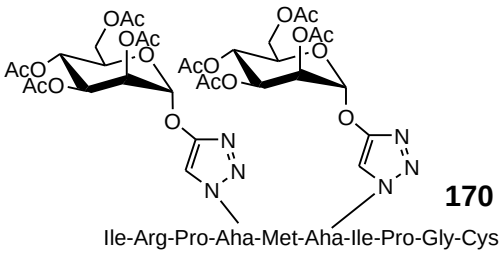
Modified peptide 410-420 of ADG



Modifed peptide sequence of CD52

The modified peptides 401-409 (**168**), 410-420 (**169**) of ADG and CD52 peptide (**165**) were

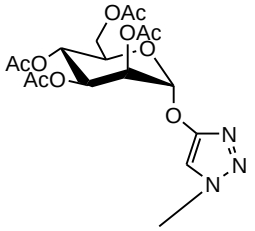
synthesised on 0.01 mmol scale using the peptide synthesiser on a preloaded Fmoc-Cys(Trt)-NovaSyn-TGT resin (Merck Biosciences). Peptide elongation employed ten equivalents of each amino acid, HBTU/HOBt as coupling reagents and *N,N*-diisopropylethylamine as base. Fmoc-Aha-OH (Chiralix) was incorporated manually with five equivalents at 2 hours. Following the synthesis a small amount of the resin was cleaved with trifluoroacetic acid: ethanedithiol: water (95: 2.5: 2.5 = 4.0 ml) for 5 hours. The resin filtered off and the filtrate poured into cold diethyl ether (20 ml), to produce a white solid. The ether layer was decanted and the white solid lyophilised. Mass analysis showed the correct mass for all three peptides. The calculated mass of peptides **168**, **169** and **165** were 1110.3, 1095.2, and 1394.4. The respective mass charge ratio values found where 1111.2, 1095.3 and 1395.1.



170

Ile-Arg-Pro-Aha-Met-Aha-Ile-Pro-Gly-Cys

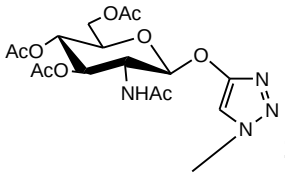
Modified glycopeptide 401-409 of ADG



171

Cys-Val-Glu-Pro-Aha-Ala-Val-Ala-Thr-Pro-Pro

Modified glycopeptide 410-420 of ADG



166

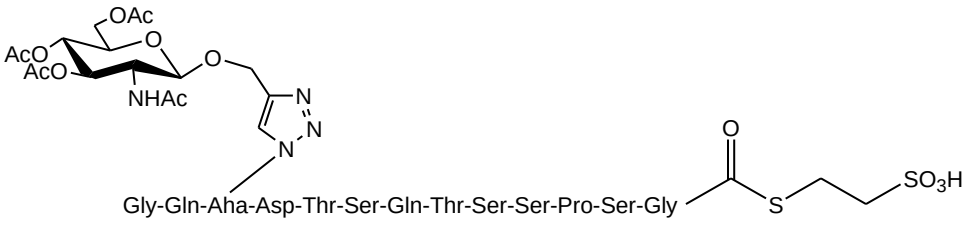
Gly-Gln-Aha-Asp-Thr-Ser-Gln-Thr-Ser-Ser-Pro-Ser-Gly-Cys

Modifed glycopeptide of CD52

Click chemistry was performed between the azide bearing peptides and the propargyl glycosides. The general conditions for click chemistry were: 1 equivalent (eq) resin bound peptide, 5 eq propargyl glycoside, 5 eq sodium ascorbate, Cu (II) SO_4 (1 mg), (9:1:1) $CHCl_3$ EtOH: 50 mM sodium phosphate pH 7.0 in a total volume of 1.1 ml. The reactions were shaken overnight at 40^{o}C and after cleavage from the resin the resulting glycopeptides were purified by HPLC using C18 column (100Å Size: 250 x 10.00 mm, 10 micron) (5-50% acetonitrile 45 minutes). The resulting glycopeptides were examined by mass analysis. The calculated mass of modified glycopeptides **166**, **170**, and **171** were 1911.07, 1779.75, and 1495.61. The respective mass charge ratio values found were 1911.08, 1780.50, and 1496.20 in 15-20% yield.

Thioesterification of the modified glycopeptides **170** and **166** was performed in 2.5 ml Eppendoff tubes each containing (0.8 ml) 6 M guanidine hydrochloride, (0.1 ml) AcOH 10%

(w/v), (0.1g) MESNa 10% (w/v), (5 mg) TCEP 0.5% (w/v) at 60°C for 48 hours. After HPLC purification using C18 column (100Å Size: 250 x 10.00 mm, 10 micron) (5-50% acetonitrile 45 minutes) mass analysis showed successful thioesterification of glycopeptide **166** to form **167.** The calculated mass of **167**: 1801.5 m/z, mass found 1801.8 m/z.



167

6.5 Molecular biology.

6.6 Instruments and general techniques.

Centrifugation of small samples in Eppendorf tubes was carried out at maximum speed in an Eppendorf centrifuge 5410. Larger samples (up to 50 ml) were centrifuged in an Eppendorf 5410R. Large scale (15 ml - 0.3 L) samples were centrifuged in a Beckmann Avanti J-25 centrifuge. Sonication was performed using a Sanyo Soniprep 150 at 50% of maximum power. Cells were sonicated on ice for 15 periods of 30 seconds separated by 30 second intervals.

Protein electrophoresis was carried out using Bio-Rad Mini-PROTEAN III vertical gel electrophoresis apparatus. Polyacrylamide resolving gels were run for approximately 1 hour at 200 V. SDS PAGE gels were stained (0.1% (w/v) Coomassie blue R-250, 40% (v/v) MeOH, 10% v/v AcOH) for 1 hour and destained with several washings of 40% (v/v) MeOH, 10% (v/v) AcOH.

Media, glassware, and equipment were sterilised by autoclaving at 121°C for 15 minutes in an Astell autoclave. Plates were prepared and media inoculated either on ethanol swabbed surfaces or in a Class II environmental cabinet. Ampicillin antibiotic stock solutions were stored at -20°C and were used at a concentration of 100 µg/ml. Shake flask fermentations were carried out in a New Brunswick Scientific INNOVA 4000 incubator shaker (37°C, 250 rpm). Large scale fermentations (10 L) were carried out in a New Brunswick Scientific BioFlow 1000 fermenter at 37°C, stirred at 250 rpm with an air flow of 2 L min^{-1}.

6.7 Growth Media.

Luria-Bertani (L.B) medium. Per litre: Bacto Tryptone (10.5 g), Bacto yeast extract (5 g), NaCl (10 g). Adjust pH to 7.5 with NaOH. Autoclave.

Minimal salts (M9) medium. Dissolve 5.64 g of 5 x M9 salts (Sigma-Aldrich) in 500 ml of deionised water. Autoclave.

6.8 ManT expression [10].

XL-1 Blue cells transformed with pManFlag20 was donated by Dr Derek Macmillan. A single transformed colony was used to seed 10 ml of fresh L.B medium containing ampicillin (100 µg/ml). The culture was incubated at 37°C overnight with shaking at 250 rpm. The overnight culture was then centrifuged (3000 rpm, 4°C, 15 min), the supernatant was discarded, and the cell pellet was resuspended in 10 ml of fresh M9-CA modified medium (M9 salts with 20 g/l casamino acids, 0.4 % glycerol, 2 mM $MgSO_4$) containing ampicillin (100 µg/ml). This was cultured as above until A_{600}=0.5, then in two large conical flasks, 5 ml of the culture was resuspended in 500 ml fresh M9-CA medium containing ampicillin (100 µg/ml). This was cultured at 37^{o}C at 250 rpm until A_{600}=0.5, then protein expression was induced with 5 µM Isopropyl β-D-1-thiogalactopyranoside (IPTG) at 30°C for 12 hours. Cells were harvested by centrifugation at 8000 rpm (10,000 x g) at 4°C for 15 minutes. The pellet could then be stored at – 20°C prior to purification.

6.9 ManT purification.

The cell pellet from 1 L culture was thawed on ice and resuspended in 20 ml of ice-cold 20 % sucrose, 10 mM tris-HCl pH 7.6, containing 2 mM phenylmethylsulphonyl fluoride (PMSF), 0.4 ml of 0.5 M ethylenediaminetetraacetic acid (EDTA) was slowly added. Cells were incubated on ice for 30 minutes, and then collected by centrifugation at 8000 rpm for 20 minutes. The supernatant was removed and the pellet resuspended in 20 ml cold distilled water. The suspension was incubated for 30 minutes on ice and then centrifugation at 8000 rpm. The supernatant was dialysed against 0.1 M ammonium formate, pH 7.5, and then concentrated to 1 ml using a Vivaspin centrifugal concentrator (MWCO 10 kDa).

6.10 Western Blot Buffers.

Western blot transfer buffer: 192 mM glycine, 25 mM Tis-HCl and 20% methanol in a total volume on 100 ml.

Blocking solution: 16 mM disodium hydrogen orthophosphate, 4 mM sodium di hydrogen orthophosphate, 100 mM NaCl, 0.1% tween.

6.11 Western Blot.

A membrane and filter paper was cut to the same dimensions of the SDS-PAGE. The filter paper, membrane and fiber pads were soaked in transfer buffer for 1 hour prior to the preparation of the gel sandwich. In the trans-blot assembly cell the gel sandwich was performed by adding the pre-wetted fiber pad onto the cassette, followed by a sheet of filter paper then the equilibrated gel. Next the pre-wetted membrane was placed on the gel and then the sandwich was completed by adding a piece of filter paper on the membrane followed by the last fiber pad. The cassette was carefully closed, making sure that the gel, filter paper sandwich was undisturbed and placed into the cassette tank. The tank was filled with transfer buffer and the transfer blot was run.

6.12 Preparation of a cell-free extract of His_{10}-GDPMP (CFE) [10].

His_{10}-GDPMP protein expression pellet was donated by Dr Derek Macmillan.
A pellet derived from a 1 L induced culture was thawed on ice and resuspended in 35 ml binding buffer (5 mM imidazole, 500 mM NaCI, 20 mM Tris-HCI, pH 8) containing 10 mM PMSF. The cells were sonicated on ice for 15 periods of 30 seconds separated by 30 seconds intervals. Triton X-100 was added to a final concentration of 0.1% (v/v) and the resulting solution was rocked on ice for 10 minutes. The cells were then sonicated again on ice for 15 periods of 30 seconds separated by 30 second intervals. Inclusion bodies and cellular debris were collected by centrifugation at 10,000 rpm for 15 minutes at 4°C. The resulting pellet was discarded and the CFE was stored at -20°C prior to affinity chromatography.

6.13 Affinity chromatography of His_{10}-GDPMP.

His_{10}-GDPMP was purified from the CFE under non-denaturing conditions using Ni^{2+} immobilised on metal chelation resin as described in the manufacturer's instructions (Sigma-Aldrich). The column (2.5 ml bed volume) was charged with 7 ml aqueous $NiSO_4$ and equilibrated with 12.5 ml binding buffer. Extracts were applied to the column and allowed to drain by gravity flow. Unbound proteins were washed with binding buffer containing 1 mM PMSF, and then wash buffer (50 mM imidazole, 500 mM NaCI, 20 mM Tris-HCI, 1 mM

PMSF, pH 7.8) until A_{280} < 0.1. The target protein was then eluted in 1.5 ml fractions using eluting buffer (500 mM imidazole, 500 mM NaCl, 20 mM Tris-HCl, pH 7) until A_{280} < 0.1. The fractions were then analysed by SDS PAGE. The fractions containing the desired protein were combined, exchanged 3 times against fresh buffer (50 mM Tris pH 7.6), and concentrated to a final volume of 1 ml using a Vivaspin (MWCO 30 kDa) concentrator, which was previously passivated with 5% Tween 20 according to the manufacturer's instructions.

References:

[1] D. Macmillan, A. M. Daines, M. Bayrhuber, S. L. Flitsch, *Org. Lett*. 2002, 4, 1467.

[2] D. Macmillan, J. Blanc, *Org. Biomol. Chem*. 2006. 15, 2847.

[3] K. K. Yeoh, D. T. Butters, B. L. Wilkinson, A. J. Fairbanks, *Carbohydr. Res.* 2009, 334, 586.

[4] X. Wu, X. Tang, M. Xian, P. G. Brauschweiger, P. G. Wang, *Bioorg. Med. Cherm.* 2002, 10, 2303.

[5] S. Cottaz, J. S. Brimacombe, M. A. J. Ferguson, *Carbohydr. Res.* 1993, 247, 341.

[6] B. G. Davis, R. J. Nash, A. A. Watson, C. Smith, G. W. J. Fleet, *Tetrahedron Lett.* 1999, 55, 4501.

[7] T. R. Chan, R. Hilgraf, K. B. Shapless, V. V. Fokin, *Org.Lett*. 2004, 17, 2853.

[8] J. Kang, J. P. Richardson, D. Macmillan, *Chem. Commun*., 2009, 407.

[9] J. Kang, N. L. Reynolds, C. Tyrrell, J. R. Dorin, D. Macmillan, *Org. Biomol. Chem. 2009*, 7, 4918.

[10] S. Marchesan, D. Macmillan, *Chem. Commun.* 2008, 36, 4321.