Synthesis of glycopeptides and novel glycopeptide mimics

Javier Blanc-Meliá, BSc. MSc.

A Thesis submitted for the degree of
Doctor of Philosophy
The University of Edinburgh
2007
To my family and friends.
DECLARATION

I hereby declare that this thesis has been entirely composed by me and that the work described herein is my own except where clearly mentioned either in acknowledgement, reference or text. It has not been submitted, in whole or in part, for any other degree, diploma or other qualification.
ABSTRACT

Glycoproteins are involved in important biological processes. Their role can be understood by introducing native structures or mimetics at predetermined positions in proteins, and may lead to the discovery of new drugs. A scalable synthesis of core-I disaccharide (galactose(\(\beta-1, 3\))\(\alpha-N\)-acetylgalactosamine), \(O\)-linked to threonine was accomplished. This building block was then incorporated into a peptide using solid phase peptide synthesis.

Additionally, a novel neo-glycopeptide linkage was created using the Cu(I)-catalysed modification of Huisgen cycloaddition (click chemistry), between glycosyl azides and peptides modified with acetylene moieties. This linkage can be formed either on solid support or in aqueous solution. Moreover, the synthesis was investigated on solid phase of a tripeptide scaffold (Ac-\(L\)-Cys-\(L\)-Lys-\(L\)-Cys-NH\(_2\)) that displays \(N\)-acetylglactosamine through the thiol groups of the two cysteine residues as a potential “oligosaccharide mimic”. This structure can be linked to a peptide via the glycosyl azide of \(N\)-acetylglucosamine appended to the \(\epsilon\)-amino group of the central lysine residue. Two such oligosaccharide mimics were incorporated into an erythropoietin (EPO) derived peptide fragment in 69 % yield using the novel neo-glycopeptide linkage in an aqueous solution.

To demonstrate the tripeptide scaffold displays carbohydrates and can be recognised in a biological environment, the EPO peptide fragment bearing two oligosaccharide mimics was joined to a Green Fluorescent Protein C-terminal thioester using native chemical ligation. Afterwards, the protein-carbohydrate interaction was observed between this glycoprotein mimic and lectin Ricinus communis RCA\(_{120}\) immobilised on agarose beads.
CONTENTS

1 The biological role of glycoproteins

1.1 Structure and biosynthesis of glycoproteins
   1.1.1 N-Linked oligosaccharides
   1.1.2 O-Linked oligosaccharides

1.2 Recombinant Human Erythropoietin (rhEPO) and its biological role

1.3 GlyCAM-1 and its biological role

1.4 Conclusion

2 Synthesis of natural glycopeptides and glycopeptide mimics

2.1 Solution Phase Oligosaccharide Synthesis
   2.1.1 Protecting groups in oligosaccharide synthesis
   2.1.2 Glycoside bond formation
   2.1.3 Stereocontrol in glycosidation
      2.1.3.1 Mannosylation by intramolecular aglycon delivery (IAD)

2.2 Carbohydrate mimics

2.3 Conclusions and aims for the project

3 Core-1 O-linked to threonine, a glycobuilding block for SPPS

3.1 Introduction of amino acid glycobuilding blocks to SPPS

3.2 Objective: synthesis of core-1 O-linked to threonine for SPPS

3.3 Design and retro-synthetic analysis of core-1 glycobuilding block O-linked to threonine for SPPS

3.4 Synthesis of core-1 glycobuilding block O-linked to threonine for SPPS
   3.4.1 Synthesis of glycosyl acceptor
   3.4.2 Synthesis of glycosyl donor
3.4.3 Synthesis of $N^{\alpha}$-(fluorene-9-nylmethoxycarbonyl)-O-[2, 3, 4, 6-tetra-O-acetyl-D-galactopyranosyl-(\(\beta\)-1, 3)-4, 6-diacyetyl-2-acetamido-2-deoxy-\(\alpha\)-D-galactopyranosyl]-L-threonine ......................................................... 49

3.5 Conclusions and future work ............................................................................. 61

4 Synthesis of a novel unnatural peptide scaffold to display carbohydrates (PSDC) ................................................................................................................................. 63

4.1 Introduction ......................................................................................................... 63

4.2 Objective: synthesis of a Peptide Scaffold Displaying Carbohydrates (PSDC). 67

4.3 Design and retro-synthetic analysis of the PSDC ............................................. 68

4.4 Synthesis of the glycobuilding blocks ............................................................... 70

4.4.1 Synthesis of 2, 3, 4, 6-tetra-O-acetyl-D-galactopyranosyl-(\(\beta\)-1, 4)-2-acetamido-2-deoxy-3, 6-di-O-acetyl-1-N-[1-(2-bromo)acetyl]-\(\beta\)-D-glucopyranose building block, 90 ................................................................. 70

4.4.2 Synthesis of 2-acetamido-2-deoxy-3,6-di-O-acetyl-4-O-carboxymethyl-\(\beta\)-D-glucopyranosyl azide, 100 ................................................................. 73

4.5 Synthesis of a peptide scaffold to display carbohydrates on solid phase ........ 77

4.6 Conclusions and future work ............................................................................. 86

5 Novel neo-glycopeptide linkage using Cu(I) catalysed 1, 2, 3-triazole formation ................................................................................................................................. 87

5.1 Introduction ......................................................................................................... 87

5.2 “Click chemistry” ............................................................................................... 90

5.3 Objective: Development of a novel neo-glycopeptide linkage compatible with native chemical ligation ................................................................. 91

5.3.1 Preliminary studies in solution: model experiments. 171 ............................. 92
5.3.2 Studies on solid phase: novel neo-glycopeptide linkage using “click chemistry” \(^\text{171}\) ........................................................................................................................................ 96
5.3.3 Native Chemical Ligation (NCL). Compatibility of the novel neo-glycopeptide linkage with NCL ....................................................................................................................... 99

5.4 Introduction of two PSDC on an EPO peptide fragment applying the novel neo-glycopeptide linkage using Cu(I) catalysed 1, 2, 3 triazole formation .......... 101

5.5 NMR observations ........................................................................................................ 109

5.6 Conclusions and future work .................................................................................... 112

6 Protein-carbohydrate interactions between PSDC and lectins visualised using modified GFP .................................................................................................................. 115

6.1 Introduction .................................................................................................................. 115
6.2 Objective: development of a methodology to visualise protein-carbohydrate interactions .................................................................................................................. 120

6.3 Green Fluorescent Protein (GFP) ................................................................................. 121
6.4 Expressed Protein Ligation (EPL) ................................................................................ 121

6.5 Development of methodology to visualise carbohydrate-protein interaction .. 123

6.5.1 Synthesis of N-(iodoacetyl)-p-aminophenyl mannopyranoside .................. 123
6.5.2 Semi-synthesis of GFP modified with an unnatural glycopeptide displaying three mannoses, for protein immobilisation ........................................... 124
6.5.3 Protein-carbohydrate interaction test on agarose beads, visualised with modified GFP ................................................................................................................. 126

6.6 Recognition of the PSDC in a biologically relevant environment .................. 128

6.6.1 Semi-synthesis of GFP modified with an EPO peptide fragment containing two PSDC, for protein immobilisation ...................................................... 128
6.6.2 Protein-carbohydrate interaction test on agarose beads, visualised with modified GFP ................................................................................................................. 129
6.7 Conclusions and future work ................................................................. 131

Overall conclusions .................................................................................. 135

7 Experimental section .............................................................................. 138

7.1 General Experimental Details .............................................................. 138

7.1.1 Instrumentation ............................................................................. 138

7.1.2 Chromatography .......................................................................... 138

7.1.3 Solvents and reagents ................................................................. 139

7.2 Experimental Procedures ................................................................. 140

8 Acknowledgements .................................................................................. 196

Annexe 1: Published work ........................................................................ 198

Annexe 2: References .............................................................................. 202
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D-NMR</td>
<td>two-dimensional nuclear magnetic resonance</td>
</tr>
<tr>
<td>Ac₂O</td>
<td>acetic anhydride</td>
</tr>
<tr>
<td>Acm</td>
<td>S-acetamido methyl</td>
</tr>
<tr>
<td>AFGP</td>
<td>antifreeze glycoprotein</td>
</tr>
<tr>
<td>Aha</td>
<td>azidohomoalanine</td>
</tr>
<tr>
<td>Ara</td>
<td>arabinose</td>
</tr>
<tr>
<td>Asn</td>
<td>asparagine</td>
</tr>
<tr>
<td>Bn</td>
<td>Benzyl</td>
</tr>
<tr>
<td>Boc</td>
<td>tert-butoxy carbonyl</td>
</tr>
<tr>
<td>CAN</td>
<td>Cerium Ammonium Nitrate</td>
</tr>
<tr>
<td>CBD</td>
<td>Chitin Binding Domain</td>
</tr>
<tr>
<td>ConA</td>
<td>Concanavalin A</td>
</tr>
<tr>
<td>Cu</td>
<td>copper</td>
</tr>
<tr>
<td>Cys</td>
<td>cysteine</td>
</tr>
<tr>
<td>d</td>
<td>doublet</td>
</tr>
<tr>
<td>DAST</td>
<td>(diethylamino) sulphur trifluoride</td>
</tr>
<tr>
<td>DBU</td>
<td>1,8-diazabicyclo[5.4.0]undec-7-ene</td>
</tr>
<tr>
<td>DCC</td>
<td>1,3-dicyclohexyl carbodiimide</td>
</tr>
<tr>
<td>DCM</td>
<td>dichloromethane</td>
</tr>
<tr>
<td>DDQ</td>
<td>dichloro dicyano quinone</td>
</tr>
<tr>
<td>DIPEA</td>
<td>diisopropylethylamine</td>
</tr>
<tr>
<td>DMAP</td>
<td>4-(N,N-dimethylamino) pyridine</td>
</tr>
<tr>
<td>DMDO</td>
<td>dimethyldioxirane</td>
</tr>
<tr>
<td>DMF</td>
<td>N,N-dimethylformamide</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethylsulphoxide</td>
</tr>
<tr>
<td>DPPA</td>
<td>diphenylphosphoryl azide</td>
</tr>
<tr>
<td>DTBMP</td>
<td>2,6-di-tert-butyl-4-methylpyridine</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>ECM</td>
<td>extracellular matrix</td>
</tr>
<tr>
<td>EDT</td>
<td>ethane dithiol</td>
</tr>
<tr>
<td>EGF</td>
<td>epidermal growth factor</td>
</tr>
<tr>
<td>EGFP</td>
<td>enhanced green fluorescent protein</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme-linked Immunosorbent Assay</td>
</tr>
<tr>
<td>EPL</td>
<td>Expressed Protein Ligation</td>
</tr>
<tr>
<td>equiv.</td>
<td>equivalents</td>
</tr>
<tr>
<td>ER</td>
<td>endoplasmatic reticulum</td>
</tr>
<tr>
<td>FITC</td>
<td>Fluorescein Isothiocyanate-Conjugated</td>
</tr>
<tr>
<td>FITC-AA</td>
<td>Fluorescein Isothiocyanate-Conjugated-\textit{A. aurantia}</td>
</tr>
<tr>
<td>FITC-ConA</td>
<td>Fluorescein Isothiocyanate-Conjugated-\textit{Concanavalin A}</td>
</tr>
<tr>
<td>FITC-EC</td>
<td>Fluorescein Isothiocyanate-Conjugated-\textit{E. cristagalli}</td>
</tr>
<tr>
<td>FITC-NPA</td>
<td>Fluorescein Isothiocyanate-Conjugated-\textit{N. pseudonarcissus}</td>
</tr>
<tr>
<td>FITC-TV</td>
<td>Fluorescein Isothiocyanate-Conjugated-\textit{T. vulgaris}</td>
</tr>
<tr>
<td>Fmoc</td>
<td>9-fluorenlymethoxycarbonyl</td>
</tr>
<tr>
<td>Fuc</td>
<td>fucose</td>
</tr>
<tr>
<td>Gal</td>
<td>galactose</td>
</tr>
<tr>
<td>GalNAc</td>
<td>N-acetylgalactosamine</td>
</tr>
<tr>
<td>GalNAcT</td>
<td>N-acetylgalactosaminyl transferase</td>
</tr>
<tr>
<td>GFP</td>
<td>Green Fluorescent Protein</td>
</tr>
<tr>
<td>Glc</td>
<td>glucose</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>GlcNAc</td>
<td>N-acetylg glucosamine</td>
</tr>
<tr>
<td>GlcNAcT</td>
<td>N-acetylg glucosaminyl transferase</td>
</tr>
<tr>
<td>GlyCAM-1</td>
<td>Glycosylation-dependent Cell Adhesion Molecule-1</td>
</tr>
<tr>
<td>GPI</td>
<td>Glycosylphosphatidylinositol</td>
</tr>
<tr>
<td>h</td>
<td>hours</td>
</tr>
<tr>
<td>HEV</td>
<td>High Endothelial Venules</td>
</tr>
<tr>
<td>HPLC</td>
<td>High Performance Liquid Chromatography/High Pressure Liquid Chromatography</td>
</tr>
<tr>
<td>IAD</td>
<td>intramolecular aglycon delivery</td>
</tr>
<tr>
<td>ICAM-1</td>
<td>Intercellular Adhesion Molecule-1</td>
</tr>
<tr>
<td>m</td>
<td>multiplet</td>
</tr>
<tr>
<td>MAdCAM-1</td>
<td>Mucosal Addressin Cell Adhesion Molecule-1</td>
</tr>
<tr>
<td>Man</td>
<td>mannose</td>
</tr>
<tr>
<td>MBP</td>
<td>maltose binding protein</td>
</tr>
<tr>
<td>MeOH</td>
<td>methanol</td>
</tr>
<tr>
<td>MeOTf</td>
<td>methyl triflate</td>
</tr>
<tr>
<td>MESNA</td>
<td>mercaptoethanesulphonic acid</td>
</tr>
<tr>
<td>MHz</td>
<td>mega hertz</td>
</tr>
<tr>
<td>MS</td>
<td>Mass Spectroscopy</td>
</tr>
<tr>
<td>MS</td>
<td>molecular sieves</td>
</tr>
<tr>
<td>MW</td>
<td>molecular weight</td>
</tr>
<tr>
<td>NBS</td>
<td>N-bromosuccinimide</td>
</tr>
<tr>
<td>NCL</td>
<td>Native Chemical Ligation</td>
</tr>
<tr>
<td>NEB</td>
<td>New England Biolabs</td>
</tr>
<tr>
<td>NeuAc</td>
<td>N-acetyl neuraminic acid (sialic acid)</td>
</tr>
<tr>
<td>NIS</td>
<td>N-iodo succinimide</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>NMM</td>
<td>N-methyl morpholine</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>Pfp</td>
<td>pentafluorophenyl</td>
</tr>
<tr>
<td>Ph</td>
<td>phenyl</td>
</tr>
<tr>
<td>Phth</td>
<td>phthalamide</td>
</tr>
<tr>
<td>Piv</td>
<td>pivaloyl</td>
</tr>
<tr>
<td>PMB</td>
<td>para-methoxy benzyl</td>
</tr>
<tr>
<td>PSDC</td>
<td>Peptide Scaffold Displaying Carbohydrates</td>
</tr>
<tr>
<td>PSGL-1</td>
<td>P-selectin-glycoprotein-ligand-1</td>
</tr>
<tr>
<td>p-TsOH</td>
<td>para-toluene sulphonic acid</td>
</tr>
<tr>
<td>RER</td>
<td>rough endoplasmatic reticulum</td>
</tr>
<tr>
<td>rhEPO</td>
<td>recombinant human erythropoietin</td>
</tr>
<tr>
<td>rpm</td>
<td>revolutions per minute</td>
</tr>
<tr>
<td>s</td>
<td>singlet</td>
</tr>
<tr>
<td>Ser</td>
<td>serine</td>
</tr>
<tr>
<td>sLe\textsuperscript{x}</td>
<td>Sialyl Lewis x</td>
</tr>
<tr>
<td>SPOS</td>
<td>Solid Phase Oligosaccharide Synthesis</td>
</tr>
<tr>
<td>SPPS</td>
<td>Solid Phase Peptide Synthesis</td>
</tr>
<tr>
<td>SPR</td>
<td>Surface Plasmon Resonance</td>
</tr>
<tr>
<td>t</td>
<td>triplet</td>
</tr>
<tr>
<td>t\textsubscript{1/2}</td>
<td>half life</td>
</tr>
<tr>
<td>TBAF</td>
<td>tetrabutylammonium fluoride</td>
</tr>
<tr>
<td>TBAHS</td>
<td>tetrabutylammonium hydrogen sulphate</td>
</tr>
<tr>
<td>TBAI</td>
<td>tetrabutylammonium iodide</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>TBDMS</td>
<td>tert-butyldimethylsilyl</td>
</tr>
<tr>
<td>TBDPS</td>
<td>tert-butyldiphenylsilyl</td>
</tr>
<tr>
<td>TBTA</td>
<td>tris-benzyltriazolylmethylamine</td>
</tr>
<tr>
<td>TCEP-HCl</td>
<td>tris-(2-carboxyethyl)phosphine hydrochloride</td>
</tr>
<tr>
<td>Tf₂O</td>
<td>trifluoromethanesulphonic anhydride</td>
</tr>
<tr>
<td>TFA</td>
<td>trifluoro acetic acid</td>
</tr>
<tr>
<td>THF</td>
<td>tetrahydrofuran</td>
</tr>
<tr>
<td>Thr</td>
<td>threonine</td>
</tr>
<tr>
<td>TLC</td>
<td>Thin Layer Chromatography</td>
</tr>
<tr>
<td>TMS-OTf</td>
<td>trimethylsilyl trifluoromethanesulphonate</td>
</tr>
<tr>
<td>UDP</td>
<td>uridine diphosphate</td>
</tr>
<tr>
<td>UV</td>
<td>ultra violet</td>
</tr>
<tr>
<td>Xyl</td>
<td>xylose</td>
</tr>
</tbody>
</table>
1 The biological role of glycoproteins

Carbohydrates are often found in nature associated with proteins or lipids, to form glycoconjugates such as glycoproteins, proteoglycans, glycolipids and GPI-anchors. These structures are very important and perform multiple roles in cells. Consequently, these molecules have captivated the attention of the chemical and biochemical community.\(^1\)

Glycoproteins are the product of a post- or co-translational modification of a protein, called glycosylation, which is not under direct genetic control. The outcome of this process is a heterogeneous mixture of glycosylated forms, termed "glycoforms".\(^2\) Such a procedure confers an additional level of complexity upon the recipient protein.\(^3\) This may explain, for example, why such a large combination of human protein phenotypes can derive from 30,000 genes.\(^4,5\) O-Methylation, O-acetylation, O-sulphation, O-phosphorylation or oxidation also amplifies the protein phenotype.

The biological properties of a protein can be influenced by the degree of glycosylation. This modification can change its biological and physicochemical properties, eg. protection from proteolysis, correct protein folding, enhancement of solubility, and confer antifreezing properties.\(^6\) Moreover, these structures mediate cellular adhesion, cell growth regulation, affect intracellular transport of enzymes to lysosomes, determine human blood groups (ABO antigens),\(^7\) and regulate leukocyte trafficking to sites of inflammation.\(^8\) Conversely, aberrant glycosylation of protein, due to incomplete synthesis of carbohydrate chains and accumulation of their precursors on the cell surface of cancer cells, is associated with
1. The biological role of glycoproteins

oncogenesis and has been implicated in metastasis. Also glycoproteins are involved in chronic inflammation, autoimmune diseases, xenograft rejection, virus replication; virus, bacterial and parasitic infection.

1.1 Structure and biosynthesis of glycoproteins

The extent to which proteins are glycosylated in nature can differ enormously from collagen (1% w/w) to glycogen (99% w/w). The modified protein may contain one or several oligosaccharide side chains. The saccharide residues are most frequently covalently linked to the protein backbone through N- or O-glycosidic linkages. The N-glycosidic bond always uses the side chain amide nitrogen of the asparagine (Asn) residue in the consensus sequence Asn-Xxx-Ser/Thr, whereas most of α-O-glycosidic bonds usually require the hydroxyl group of serine (Ser) or threonine (Thr), introduced by a peptidic region rich in proline. C-, S- and P-dependent glycosidic linkages are found less frequently in nature. GPI-Anchors use an ethanolamine phosphate linkage type between the protein chain and carbohydrate structure.

1.1.1 N-Linked oligosaccharides

N-Linked glycosylation is the most common form of glycosylation in eukaryotic cells. This modification involves a co-translational transfer of a saccharide from a dolichol-linked pyrophosphate donor to an asparagine in the consensus sequence Asn-Xxx-Ser/Thr. Although all N-glycoproteins share a peptide-linked pentasaccharide motif, called the core region, in the biosynthesis of these structures, the activity of different types of enzymes is responsible for the high diversity associated with mature N-linked glycoproteins in eukaryotic cells.
The biosynthesis of N-linked oligosaccharides starts on the cytosolic surface of the endoplasmic reticulum (ER) membrane by adding sugars in a stepwise fashion to dolichylphosphate referred to as the phosphodolichol pathway or dolichol cycle (Scheme 1.1). Two UDP-GlcNAc residues are transferred to phosphodolichol by a 1-phosphotransferase and GlcNAc-transferase: this is followed by the addition of five mannose residues catalysed by different mannosyltransferases. After mannosylation, the glycolipid is flipped from the cytoplasmatic face of the ER to the lumen. The growing glycolipid receives another four mannose and three glucose units from dolichol-phosphate linked glycosyl donors. Then, an oligosaccharyltransferase enzyme complex catalyses the transfer of the core tetradecasaccharide \((\text{Glc}_3 \text{Man}_9 \text{GlcNAc}_2)\) to the asparagine acceptor in the growing polypeptide with a consensus sequence Asn-X-Thr/Ser, where X can be any amino acid except proline or aspartic acid. However, not every asparagine in a consensus sequence is glycosylated: steric factors can prevent glycosylation.

Afterwards, the tetradecasaccharide structure is “trimmed”. Three glucose and one terminal mannose are removed by ER \(\alpha\)-1,2-glucosidase I, \(\alpha\)-1,3-glucosidase II, and \(\alpha\)-1,2-mannosidase respectively (Scheme 1.1). This deglycosylation process leads to an oligosaccharide structure \((\text{Man}_n \text{GlcNAc}_2)\) that signals the nascent glycoprotein to be transported to the \emph{cis}-Golgi compartment for further processing. After being “trimmed” of the first two glucose residues, the new glycopeptide is released from the ribosome of the ER to the lumen, and then associates with membrane bound calnexin or soluble calreticulin, chaperones which recognise the monoglucosylated oligosaccharide, and assist the glycoprotein to fold correctly. It is at this stage where disulphide bridges are formed. If, after the removal of the third glucose, correct folding fails, the chaperone retains the glycoprotein in the ER for further re-glycosylations to readdress the improper folding.
1. The biological role of glycoproteins

The high-mannose type glycoprotein structure is transported to the "cis cisternae" of the Golgi apparatus by vesicular transport. Depending on the final destination of the mature glycoproteins, two possible biosynthetic pathways are believed to exist (Scheme 1.1):²

- a phospho-N-GlcNAc unit is added in the position 6 of a mannose, and then a phosphodiesterase removes the N-GlcNAc to form a mannose-6-phosphate. This structure is finally transported to lysosomes.
- three units of mannose are cleaved by mannosidases IA, IB and IC (Man₃GlcNAc₂).

This glycoprotein is finally matured in the "medial cisternae".

Diversification of N-glycans occurs in the medial cisternae generating different subtypes of N-linked glycoproteins: high mannose, hybrid or complex N-oligosaccharides (Scheme 1.2). Two mannose units are removed by mannosidase II, and afterwards, two GlcNAc and a fucose residue are added by GlcNAcT II and fucosyltransferase respectively. Finally, in the "trans cisternae" Golgi, the oligosaccharide chain is elongated by additional
glycosyltransferases that add: GlcNAc, galactose, and/or sialic acid residues. This final elongation step contributes greatly to the highly diverse and complex nature of N-linked glycoproteins.

Scheme 1.2: different subtypes of N-linked glycoproteins: high mannose, hybrid or complex N-oligosaccharides.

1.1.2 O-Linked oligosaccharides

In the biosynthesis of O-glycan, glycosyltransferases act directly to glycosylate proteins in the Golgi. As a result of this direct glycosylation, O-glycans have fewer carbohydrate branches than most N-glycans, and are commonly found as biantennary structures.

O-Linked oligosaccharides differ in a manner dependent upon how the first sugar molecule is attached to the recipient protein. Linkage can occur through serine or threonine residues (Scheme 1.3). Mucin-type O-glycoproteins, which are highly glycosylated, are commonly found on the surface of epithelial cells. These structures exert a number of biological effects including structural and functional properties for the carrier protein and are involved in molecular recognition and adhesion events. Mucin-type O-linked glycoproteins are initiated by the addition of the monosaccharide N-acetylgalactosamine, GalNAc. The addition of GalNAc in the protein backbone proceeds through the hydroxyl group of Ser/Thr,
and is catalysed by polypeptide GalNAc UDP-transferase, ppGalNAcT using UDP-GalNAc in the Golgi membrane. A consensus sequence for ppGalNAcT has not been determined, although predictive algorithms have been proposed. Glycosidation normally occurs where proline is at the -1 and +3 position. Prolines appear to influence protein conformation by breaking helix formation and promoting the formation of β turns and β sheets. Alanine, serine and threonine residues are commonly found adjacent to the glycosylated residue.

Scheme 1.3: Common core structures of O-linked glycoproteins.

Following the addition of the first monosaccharide to the polypeptide, most O-glycans are observed to contain the core-1 subtype structure formed by the addition of galactose in a β-1, 3 linkage to the GalNAc (Scheme 1.3). In mammalian cells, this biotransformation is performed by core 1 β-1, 3 galactosyltransferase (core 1 GalT).

Core 2-type O-linked glycoproteins can be generated by the addition of GlcNAc to the core GalNAc in a β-1, 6 linkage. The production of core 2 O-glycans requires the presence of core 1 structure as a substrate. This reaction is catalysed by core 2 β6-N-acetylgalactosaminy-
transferase (core 2 $\beta 6$-GlcNAcT). The enzyme recognises most of the substituents of core 1: the 4- and 6-hydroxyls of both Gal and GalNAc, and the 2-$\text{N}$-acetyl group of the glycosyl acceptor.

Glycosyl addition on core 1 and core 2 structures generates other types of core structures (Scheme 1.3). The product of these glycosyltransferases can be sulphated by sulphotransferases. Elongation proceeds by:

- Repeating Gal$\beta$-1, 3/4GalNAc sequences (poly-$\text{N}$-acetyllactosamine chains).
- Branching GlcNAc$\beta$-1, 6 residues.

The termination process of the elaboration of carbohydrate structure involves the action of a fucosyltransferase and a sialyltransferase. Sialylation imparts a negative charge to the final carbohydrate, and may protect the $\text{O}$-glycoprotein from being recognised by lectins or the immune system. $\text{O}$-Acetylation of sialic acid, performed by acetyl-CoA, protects the sugar from degradation by intestinal bacteria. Moreover, Gal or GlcNAc residues of mucins are often sulphated, conferring an additional negative charge to $\text{O}$-glycan chains. Sulphated $\text{O}$-glycans may play a role in cell adhesion through selectin binding, and bacterial binding. Sulphated ester groups also control biosynthetic pathways of $\text{O}$-glycans by blocking specific glycosylation steps.

Other types of glycosylation afford: $\text{O}$-linked $\text{N}$-acetylglucosamine (GlcNAc), found in the cytoplasm and nucleus of cells; $\text{O}$-linked fucose, found in the epidermal growth factor (EGF) domains; and $\text{O}$-linked xylose, abundant on the cell surface of the extracellular matrix (ECM).

Different types of $\text{N}$- and $\text{O}$-linked glycoproteins with interesting biological activity exist in nature. The most relevant for this thesis are erythropoietin and GlyCAM-1.
1.2 Recombinant Human Erythropoietin (rhEPO) and its biological role

Recombinant Human Erythropoietin (rhEPO) is a 166 amino acid residue glycoprotein that regulates the proliferation, differentiation and maturation of red blood cells, and consequently the haematocrit. Although in the early stages of foetal life, the liver is responsible for the production of this secreted hormone, suprarenal glands are responsible for EPO generation in adults. At high altitudes, the production of the glycoprotein is increased to counter the effects of hypoxia. Conversely, during hyperoxia, there is a reduction of the hormone.33

EPO is commonly used as a therapeutic for the treatment of anaemia associated with: chronic kidney disease, chemotherapy treatment for cancer; HIV infection, and during surgery to reduce allogenic blood transfusion requirements.34,35

The secreted glycoprotein contains approximately 40% carbohydrate and has an approximate molecular weight of 30 kDa. rhEPO has four cysteines linked by disulphide bonds between Cys 7-161 and 29-33. Reduction of disulphide bonds results in loss of biological activity, that can be restored by reoxidation.36 rhEPO is glycosylated at N24, 38, 83 and S126 (Scheme 1.4).35

The carbohydrate attached to asparagine has the typical structure of N-linked glycoproteins (Scheme 1.4). The carbohydrate attachments influence protein stability and solubility but are not required for in vitro activity. Thus the active site is retained in the protein portion of the molecule.34 Sialylation of the carbohydrate is essential for in vivo activity. Desialylation exposes galactose residues to hepatic metabolism. Desialylated molecules are sequestered in the liver and metabolised, increasing the clearance and consequently decreasing the half life of EPO.
1. The biological role of glycoproteins

1.3 GlyCAM-1 and its biological role

The presence of glycoproteins on the cell surface facilitates the lymphocyte recirculation from the intravascular compartment to the secondary lymphoid organs and also mediates immune responses. This recirculation increases the chances of the lymphocyte encountering foreign antigens.\textsuperscript{1,2}

L-Selectins, glycoproteins from the family of the lectins, present on the surface of the leukocytes have different types of biological ligands: GlyCAM-1, CD34, MA\textsuperscript{d}CAM-1 and PCLP. Glycosidation-dependent cell adhesion molecule-1 (GlyCAM-1), is a secreted mucin-type O-glycoprotein of 132 amino acids, with a MW of 50 kDa, that lacks a transmembrane domain and contains numerous serine and threonine residues. GlyCAM-1 is present on the surface of high endothelied venules (HEV) in secondary lymphoid organs in mice.\textsuperscript{3} This counter-receptor requires carbohydrate-based posttranslational modification for recognition by L-selectins.\textsuperscript{37} Sialylation and fucosylation are necessary for L-selectin recognition.\textsuperscript{38} Sulphation of carbohydrates increase binding to L-selectin compared to non sulphated molecules (Scheme 1.5).\textsuperscript{39}
During inflammation, in zones of injury, there is a decrease in circulation, allowing the interaction between blood cells and the vessel endothelium. It is at this stage where L-selectin and its ligands make low affinity interactions (glycoprotein-lectin), slowing the circulation of leukocytes in a process called rolling (Scheme 1.5).^2

![Scheme 1.5: Process of leukocytes extravasation from the blood stream.](image)

Afterwards, this leads to the activation of integrins present on the surface of the leukocytes,\(^4^1\) that facilitate strong protein-protein interactions with intercellular adhesion molecule-1 (ICAM-1) present on the endothelial surface, leading to the firm adhesion of the leukocytes on the surface of the blood vessel. Finally, there is a transmigration of leukocytes through the endothelial layer, a process called vascular extravasation (Scheme 1.5).

### 1.4 Conclusion

Carbohydrates represent a new level of complexity and diversity above gene expression. These molecules, introduced to proteins by nature, confer new properties to the host protein
acceptor. Properties that are hidden by its heterogeneity can only be specifically decoded by the synthesis and biosynthesis of pure and homogeneous native or unnatural glycoproteins. The understanding of the role of glycans is already providing new therapies and treatments, and will enlighten science in the coming years.
2 Synthesis of natural glycopeptides and glycopeptide mimics

Polysaccharides are highly functionalised biopolymers with a large number of stereocentres. These molecules have attracted the attention of chemists because of their complexity. The main problems of the synthesis of these compounds are:

- Stereoselective glycoside bond formation between the glycosyl donor and the acceptor (generation of $\alpha$ or $\beta$ anomer).
- Regioselective glycosylation of the acceptor.
- Selective protection and deprotection of different building blocks before and after each glycosylation step.

To combat these challenging problems, scientists have engineered different methods, either by chemical or enzyme-catalysed glycosylation, to overcome these difficulties (Scheme 2.1). The routes are complimentary.

Enzymes are used to accomplish glycosylation with total regio- and stereo-control, in non-organic solvents. If the desired enzyme is commercially available or if it can be expressed in cells, glycosylation is performed with high efficiency and without any protecting groups. On the other hand, in chemical synthesis, natural and non-natural saccharide building blocks can be assembled with natural or non-natural linkages to generate any oligosaccharide, analogue or glycoconjugate. This last approach involves the use of suitable protecting group strategies and glycosyl activators.
2. Synthesis of natural glycopeptides and glycopeptide mimics

### Chemical glycosidation

- **Protecting group (P):** Prevents undesired glycosidation.
- **Leaving group (L):** Leaves the molecule after reaction.

#### i. Promoter (P':Ac)

1. **Reaction Step:** 
   - 
   - 
   - 

2. **Deprotection:** 
   - 
   - 
   - 

#### ii. Deprotection

1. **Reaction Step:** 
   - 
   - 
   - 

2. **Deprotection:** 
   - 
   - 
   - 

**Scheme 2.1:** General examples of glycosidation reactions.

### Enzyme-catalysed glycosidation

- **UDP:** Uridine diphosphate

#### β-1,3-GaIT

1. **Reaction Step:** 
   - 
   - 
   - 

2. **β-1,4-GalT:** 
   - 
   - 
   - 

**Scheme 2.1:** General examples of glycosidation reactions.

2.1 Solution Phase Oligosaccharide Synthesis

2.1.1 Protecting groups in oligosaccharide synthesis

As has been discussed earlier, carbohydrates are highly hydroxylated branched molecules. These hydroxyl groups share similar polarity and nucleophilic properties. Moreover, the reactivity of these functional groups can vary depending on their environment, suggesting that carbohydrate synthesis relies upon a good protecting group strategy to satisfy the stereo- and regioselectivity synthetic demand.

This protecting group strategy has to meet different requirements: a given hydroxyl group must be protected selectively; a protecting group has to withstand the conditions of the intended reactions; and finally, selective deprotection must be achieved at the end of the sequence in high yield.
The literature describes a variety of different protecting groups for either hydroxyl or amino functional groups.\textsuperscript{43} Despite this large number, only a fraction of them have emerged as suitable protecting groups for carbohydrate chemistry.\textsuperscript{42,44} With modern protecting group strategies, there is potential for creating every possible carbohydrate structure, either by modifying the chemical properties of the molecule, by controlling the stereochemistry of the glycosylation bond formation, or by changing the intrinsic reactivity of the carbohydrate. In fact, applying this final property, Fraser-Reid and co-workers,\textsuperscript{45,46} developed a sequential synthesis from more reactive \textit{n}-pentenyl glycosides (\textit{armed} glycosyl donor), furnished with electron-donating hydroxyl protecting groups; to less reactive (\textit{disarmed}) acceptors, protected with acyl electron-withdrawing groups. Wong and co-workers have applied this difference in reactivity of glycosyl donors modulated with protecting groups to develop one-pot oligosaccharide synthesis.\textsuperscript{47}

One of the major drawbacks of carbohydrate chemistry is the long and laborious synthesis to furnish the sugar unit with the adequate protecting groups, leaving free the desired hydroxyl group. Very recently, Hung and co-workers have performed a highly-regioselective one-pot protection of carbohydrates (Scheme 2.2).\textsuperscript{48} With this methodology, it is possible to install different protecting groups on a monosaccharide in a one-pot fashion, removing the need to carry out multiple synthetic steps, and repeated isolation and purification of intermediates.

First, they treated the compound $1a$ or $1b$ with an aryl aldehyde in the presence of trimethylsilyl triflate (TMSOTf) as a catalyst for the formation of compound $8$ (Scheme 2.2). Addition of another aryl aldehyde, followed by treatment with triethylsilane (Et$_3$SiH), allowed the generation of $10$. TBAF deprotection removed the TMS group at the C-2 hydroxyl position to afford $2$. In addition, $4$ was prepared in a one-pot fashion when $2$ was treated with an appropriate electrophile under basic conditions, followed by the addition of
DDQ. 4 was also obtained when the C-2 hydroxyl of 10 was acetylated, followed by DDQ treatment.

Scheme 2.2: Diagram for the TMSOTf-catalysed one-pot protection of carbohydrates.

Once the monosaccharide 3 was obtained, selective deprotection of the hydroxyl at C-4 or C-6, was achieved either by treating the molecule with HCl and Na(CN)BH3, or with BH3/THF, to obtain 5 and 6, respectively (Scheme 2.2).

Applying this methodology, the authors claimed to produce “hundreds” of glycobuilding blocks. To exemplify their strategy, they generated a library of the influenza virus-binding trisaccharide (NeuAc-α-2, 6-Gal-β-1, 4-GlcNAc).
2. Synthesis of natural glycopeptides and glycopeptide mimics

2.1.2 Glycoside bond formation

Glycoside bond formation generally refers to a process where a carbohydrate is added to an acceptor molecule, either a further carbohydrate or another molecule entity. While in nature, enzymes catalyse this reaction by bringing together both donor and acceptor, chemists have developed different procedures to reproduce this natural process.

In chemical glycoside bond formation (Scheme 2.3), there is most commonly an activation and elimination of a leaving group at the anomeric position of the carbohydrate, called the glycosyl donor, to generate an oxo-carbenium species. Stabilisation of this cationic intermediate occurs by donation of the oxygen lone pair of electrons, participation of active counter ions, such as triflate; or in some cases, neighbouring group participation (C-2 substituent participation). Final nucleophilic attack of a free hydroxyl group of another carbohydrate unit, the glycosyl acceptor, generates the desired glycosyl bond. Glycosidation is performed under strictly anhydrous conditions.

Scheme 2.3: Conventional and new approaches for stereoselective glycosylation: A) classical neighbouring group participation by C-2 ester leading to 1,2-trans-glycosides; B) neighbouring group participation by C-2 S auxiliary to afford 1,2-cis-glycosides.49
Since Knoenigs-Knorr reported the first efficient glycosylation with glycosyl bromides activated with equimolar amounts of silver salts, a wide variety of procedures and reviews have been described in the literature.

2.1.3 Stereocontrol in glycosidation

The stereochemistry of the glycosidic bond can be controlled using different protecting groups in adjacent position to the anomeric centre (C-2): neighbouring-group participation. Traditionally, the presence of an ester group, such as acetate or benzoate, on the 2-hydroxyl group, stabilises the intermediate glycosyl cation by cyclisation. This cyclic oxonium ion can then be opened in an SN2 fashion by the external nucleophile, with inversion of configuration (participation of C-2 substituent, Scheme 2.3.A).

Boons and co-workers developed a general strategy for stereoselective glycosylations. The authors introduced on the C-2 of the glycosyl donor a (1S)-phenyl-2-(phenylsulphanyl)ethyl moiety, in the presence of a Lewis acid (BF$_3$OEt$_2$), before introducing the trichloroacetimidate leaving group at the anomeric position. Due to steric and electronic factors, the sulphonium ion is formed as a trans-decalin ring system. Displacement of the sulphonium ion by a nucleophilic hydroxyl leads to the stereoselective formation of $\alpha$-glycosides. Finally, the moiety was removed by treating the oligosaccharide with a Lewis acid (BF$_3$OEt$_2$) in the presence of acetic anhydride to achieve the acetylation of the 2-hydroxyl group.

While C-2 substituent neighbouring participating group fills the synthetic tool-kit for glycosylation, $\beta$-mannoside formation frequently requires more sophisticated procedures. Since $\beta$-mannosylation is disfavoured under the above glycoside bond formation conditions, different conditions have been developed to overcome the difficulties (Scheme 2.4 and 2.5).
Garegg and Paulsen developed a methodology to perform β-mannosylation with mannosyl bromides, using insoluble silver salts (Scheme 2.4.A). Unfortunately, the degree of stereoselectivity was quite variable, and therefore, more stereoselective methodologies have been developed.

A) Insoluble silver salt

B) Mannosyl sulfoxide

C) Glycosylation-inversion protocol

Scheme 2.4: Different types of strategies to achieve a β-mannosylation: a) insoluble silver salt; b) mannosyl sulfoxides; c) glycosylation-inversion protocol, either by oxidation-reduction or by SN2 displacement.

Crich and co-workers have developed a successful direct β-mannosylation using α-glycosyl sulfoxides. High β-stereoselectivity could be achieved when the sulfoxide was preactivated at -78 °C with triflic anhydride and 2, 6-di-tert-butyl-4-methylpyridine (DTBMP). The formation of an α-glycosyl triflate intermediate allowed inversion of configuration in an SN2 reaction to finally obtain the desired compound (Scheme 2.4.B).
Danishefsky and co-workers reported a more elaborate procedure. In a 3 step synthesis, they performed a \( \beta \)-mannosylation from glucal.\(^6\) After treating the molecule with DMDO in the presence of a glycosyl acceptor, they subjected the product of the glycosidation to a sequential oxidation (DMSO/\( \text{Ac}_2\text{O} \) (2:1)) and reduction (\( \text{NaBH}_4 \)), to achieve inversion from the \( \beta \)-glucoside to the \( \beta \)-mannoside (Scheme 2.4.C). On the other hand Lichtenthaler and co-workers\(^{61,62} \) oxidized the C-2 position of 2, 3, 4, 6-tetra-\( \text{O} \)-acetyl-1-bromo-1-deoxy-\( \beta \)-D-glucopyranose to achieve the cyclic ketone derivative in four synthetic steps (Scheme 2.4.C). After glycosidation, they reduced the carbonyl using \( \text{NaBH}_4 \).

Applying glycosyl-inversion protocols, Kunz and Unverzagt addressed the introduction of \( \beta \)-mannosides by an \( \text{Sn2} \)-type displacement.\(^{63,65} \) The authors reacted an unprotected 2-hydroxyl of a \( \beta \)-glucose residue with triflic anhydride. The inversion of the C-2 position was performed by intramolecular nucleophilic attack of a 3-\( \text{O} \)-carbamate, affording the desired \( \beta \)-mannose after Zémplen deprotection\(^{66} \) (Scheme 2.4.C).

### 2.1.3.1 Mannosylation by intramolecular aglycon delivery (IAD)

IAD methodology was described first by Baressi and Hindsgaul.\(^{67} \) The glycosyl acceptor is introduced using an aglycon structure at the C-2-axial position, forcing the glycosyl acceptor to form exclusively the expected \( \beta \)-stereoisomer (Scheme 2.5). Different types of aglycon structures are described for IAD methodology: isopropenyl ether,\(^{67} \) silylketal structures\(^{68} \) and \( p \)-methoxybenzyl (PMB) (Scheme 2.5).\(^{69} \)

The PMB aglycon has become the most widely used IAD aglycon structure for \( \beta \)-mannosylation (Scheme 2.5), because it is easy to introduce and the conditions required for the formation of the mixed acetal are mild. Methyltrifluoromethanesulfonate (MeOTf) is usually the promoter for the \( \beta \)-mannosylation, once the glycosyl acceptor is appended.\(^{70} \)
Different examples are found in the literature using the PMB aglycon structure for \(\beta\)-mannosylation.\(^{71,72}\)

With the development of these technologies, chemists have the potential to synthesise any type of complex carbohydrate structure, overcoming the most challenging difficulties. For the production of these oligosaccharides, solution phase, solid phase, and also enzymatic glycosidation can be applied.

**Scheme 2.5:** Intramolecular aglycon delivery for \(\beta\)-mannosylation. Different types of aglycon structures: a) isopropenyl ether aglycon;\(^{67}\) b) silylketal aglycon;\(^{68}\) c) PMB aglycon.\(^{70}\)
For fast synthesis of polysaccharides, two major advances have been developed which potentially streamline chemical synthesis of oligosaccharides: one-pot reactions and polymer supported synthesis. In one-pot synthesis (Scheme 14), glycosidation occurs sequentially in a single reaction vessel: the most reactive glycosyl donor (armed glycosyl donor), i.e. the molecule that donates the anomeric centre, is triggered first and the least (disarmed) is coerced to engage in the final reaction. The reactivity of the sugar is highly dependent on the protecting groups and the anomeric activating group used. By adding substrates in sequence from the most reactive to the least reactive, one can assure the predominance of a desired target compound. These types of reactions are typically performed in solution, but in order to avoid work-up and purification steps, the final acceptor, i.e. the molecule that accepts the anomeric centre, may be attached to a solid support. Compared with stepwise solid-phase synthesis, the one-pot approach requires protecting-group manipulation only at the stage of building block synthesis and thus holds greater potential for automation, where a computer selects the building blocks for sequential one-pot oligosaccharide synthesis, giving a huge diversity of oligosaccharide structures.

Automated solid phase oligosaccharide synthesis (SPOS) (Scheme 2.6), the reaction is performed on a polymeric support, allowing rapid removal of reactants and easy purification. Encoding the product can be done either by position or, in mix and split type library construction, by an accessory encoding reaction, in which labels are added to the solid support as the chain is extended. It can also be encoded by radio frequency-encoded combinatorial chemistry technology, allowing the synthesis of oligosaccharide libraries. Although it is possible to access many carbohydrate structures on solid-phase, there are sequences that remain difficult to make and not all kinds of oligosaccharides can be assembled on a polymer matrix. Moreover, the most time consuming step is the synthesis...
of sufficient quantities of building blocks. These will become commercially available, facilitating the work of carbohydrate chemists.\footnote{80}

**Solid Phase Oligosaccharide Synthesis**

Fast and easy accessibility to diverse and complicated glycosyl structures is likely to impact significantly on medicine and glycomics. Within all of the advances in carbohydrate chemistry, the synthesis of glycoproteins is far from easy and demands huge effort. On the other hand, mimicking these molecules permits fast access to elaborate structures and allows the study of saccharide interactions that could provide better understanding of biological activities and indentify potential targets for development as therapeutic agents.

### 2.2 Carbohydrate mimetics

The term “carbohydrate mimic” is frequently used to refer to any carbohydrate derivative or other compound that has multiple hydroxyl groups and thus looks like a sugar.\footnote{81}

Carbohydrate mimetics have a number of advantages over their parent structures as therapeutic agents. They can be designed such that they are more stable toward endogenous
2. Synthesis of natural glycopeptides and glycopeptide mimics
degradative enzymes, have improved bioavailability and reduced clearance rates, and have a
higher affinity and selectivity for their cognate receptors by taking advantage of interactions
that the natural saccharide does not. By constructing polymers or oligomers with multiple
copies of the mimetic to allow for polyvalent interactions, the affinity can be increased
further. In the case of inhibitors of glycosyltransferases, glycosidase and carbohydrate-
modifying enzymes (e.g. sulphotransferase), mimetics can be designed that imitate the
transition states of these reactions, rather than the ground states, and thus can inhibit the
enzymes better than simple substrate analogues.

Different examples can be found in the literature for carbohydrate mimetics. Wong and co-
workers developed novel fucopeptides as sLe\textsuperscript{x} mimetics. They chose \( O-\alpha\)-fucosyl-L-
threonine and \( O-\alpha\)-fucosyl-(1\text{R},2\text{R})-2-aminocyclohexanol as a template for the incorporation
of a hydroxyacid (\( P_1 \)) to emulate galactose residues, and a carboxylate or sulphate group (\( P_2 \))
to replace the sialic acid residue (Scheme 2.7).

These authors determined the activity of these fucopeptides derivatives with the IC\textsubscript{50} values
inhibiting the sLe\textsuperscript{x} glycoconjugate binding to E-selectin (Scheme 2.7). According to the
results, all the functional groups required for E-selectin binding in sLe\textsuperscript{x} existed in 11, and as
expected it exhibited the same activity as sLe\textsuperscript{x} (IC\textsubscript{50} = 0.5 mM). Compounds 12 and 13 were
less active than sLe\textsuperscript{x}, probably due to the presence of a free carboxyl group in the threonine
moiety. Compound 14 was surprisingly active as the OH group of Tyr in 14 is not exactly a
spacial mimic of the essential OH groups of the galactose residue in sLe\textsuperscript{x}. Whether the
hydrophobic nature of the aromatic group or the additional amino group contributes to the
binding was unclear. Compounds 15 and 16 were more active than sLe\textsuperscript{x}, perhaps due to the
increasing constraint of the hydroxyproline-containing peptide bonds. Replacement of the
threonine component with (1\text{R},2\text{R})-2-aminocyclohexanol resulted in a significant decrease in
activity.
Meldal and co-workers introduced glycopeptides as mimetics of complex oligosaccharide structures (a branched mannose pentasaccharide containing two terminal Man-6-P units). The idea was that the carbohydrate would provide the specificity of the binding by directing the ligand to the oligosaccharide binding site (mannose 6-phosphate receptors), while the peptide (Thr-Lys-Thr) would function as a scaffold to provide optimal orientation of the glycan portion (Scheme 2.8). In addition, since peptide ligands generally bind with high affinity to peptide receptors, it was also expected that the glycopeptide could furthermore interact favourably with the carbohydrate-binding receptor through the peptide scaffold, thus leading to increased binding affinity.
2. Synthesis of natural glycopeptides and glycopeptide mimics

Scheme 2.8: Models of a phosphorylated high-mannose M7 and a glycopeptide mimic which show similar conformations. They are seen from the point of interaction with mannose-6-phosphate receptor (P. M. St. Hilarire and M. Meldal: Glycopeptide and oligosaccharide libraries. Angew. Chem. Int. Ed. 2000, 39,1162-1179. Copyright Wiley-VCH Verlag GmbH & Co.KGaA. Reproduced with permission).

The source of this enhanced affinity of peptides may lie in their ability to perform an induced fit because the relatively limited flexibility around peptide bonds leads to fast on-rates for the
binding. In contrast to this glycopeptide, highly flexible neo-glycoconjugates do not show high affinity due to the large entropic penalty.\textsuperscript{84}

The principle of glycopeptides mimicking oligosaccharides was consolidated through binding studies with an array of phosphorylated glycopeptides and the divalent mannose 6-phosphate receptor. The structural similarity between the oligosaccharide and the most active glycopeptide ligand was supported by molecular dynamics.\textsuperscript{83}

Also based on glycopeptide scaffolds, Reymond and co-workers,\textsuperscript{85} have recently synthesised a 15,625 membered peptide dendrimer combinatorial library, furnished with an $\alpha$-C-fucosyl residue, at its four N-termini on solid support (Scheme 2.10). This was screened for binding to fucose-specific lectins: UEA-I from \textit{Ulex europaeus} and PA-III form \textit{Pseudomonas aeruginosa}.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{Scheme_2.10.png}
\caption{Library of neo-glycopeptide dendrimers acetylated with fucose at the N-termini. The red highlighted amino acids in predetermined positions, formed part of the most active peptide against fucose-specific lectins: UEA-I from \textit{Ulex europaeus} ($IC_{50}=11\ \mu M$) and PA-III form \textit{Pseudomonas aeruginosa} ($IC_{50}=0.14\ \mu M$).}
\end{figure}
Bertozzi and co-workers have synthesised mucin mimetics compromising a long, synthetically tractable polymer backbone decorated with α-, β-linked GalNAc residues and LacNAc. They introduced the carbohydrates through an oxime linkage from a ketone-functionalised polymer with an aminoxyl GalNAc derivative, and LacNAc. Afterwards, the authors created a model cell surface, mimicking the membrane-bound mucin glycoproteins found in living cells (Scheme 2.11).

Scheme 2.11: Modification of the polymer with aminoxyl carbohydrate derivatives. Schematic of supported lipid bilayer formation from small liposomes and polymer incorporation in a supported lipid bilayers.

With this system they studied the interaction of the membrane-associated mucin mimetics with proteins (lectins). The study was performed with Helix pomatia agglutinin (HPA), with high specificity for α-linked GalNAc, and Bauhinia purpurea agglutinin (BPA) that binds specifically to β-linked GalNAc. Both lectins were available in fluorescein isothiocyanate (FITC)-labelled form. Moreover, the LacNAc modified polymer was also studied for lectin
binding activity using FITC-conjugated *Erythrina cristagalli* agglutinin (ECA). In all cases, these structures showed predicted interaction with their counter ligand. Thus, the authors created a simple recognition model where they mimic cell-surface interactions, relevant to cell differentiation or metastasis, by introducing polymeric chains displaying sugars.

Recently, Davis and co-workers have introduced carbohydrate structures to a leukocyte transmembrane human protein, P-selectin-glycoprotein-ligand-1 (PSGL-1) using "click chemistry" (Scheme 2.12).[^1]

![Scheme 2.12](image)

**Scheme 2.12:** reconstitution of biological activity of PSGL derivatives by the addition of siaLacNAc-alkyne or sLe^5^-alkyne, applying "click chemistry".87

Applying the technology developed by David Tirrell for incorporation of unnatural amino acids into recombinant proteins for chemoselective modifications,88 the authors introduced azidohomosalanine (Aha) and a cystiene chemical-tag to the positions 43 and 439 respectively (Aha43-Cys439), of PSGL-1, using protein expression of methionine-auxotrophic *E. coli*. Afterwards, they incorporated a sulphotyrosine mimic group (Tys), by a disulphide ligation, followed by the attachment of a siaLacNAc-alkyne or sLe^5^-alkyne
derivative applying Cu(I)-catalysed modification of the Huisgen cycloaddition (Scheme 2.12). With these unnatural modifications, the authors recovered the biological activity of PSGL-1, involved in the initial rolling-adhesion events of inflammation response.

2.3 Conclusions and aims for the project

Technology related to the semi-synthesis of pure glycoproteins may allow scientists to understand the biological role of such structures and will lead the development of new drugs and treatments. Unfortunately, many obstacles still have to be overcome. While this technology develops, carbohydrate mimetics, due to their simplicity and easy accessibility, should help to answer these challenging questions, and therefore, fill the gaps of knowledge in the biological role of glycoproteins.

The Macmillan group already has experience in the synthesis of natural and unnatural glycoproteins. In the semi-synthesis of GlyCAM-1, an O-linked glycoprotein (Chapter 1), 13 units of GalNAc were introduced in predetermined positions.\(^9\) In addition, applying auxiliaries for cysteine-free native glycopeptide ligation, GlyCAM-1 peptide fragments were synthesised containing GalNAc O-linked to threonine.\(^9\) On the other hand, the research group also has experience with N-linked glycoproteins: erythropoietin (EPO) (Chapter 1). To this protein a site directed mutagenesis/glycosylation approach has been applied to E. coli-expressed rhEPO protein to introduce GlcNAc selectively through the thiol of a cysteine without any observable side reaction with any of the other four cysteine residues involved in disulphide bonds.\(^9\) Moreover, different sugars have been introduced onto EPO peptide fragments on solid phase using iodoacetamide carbohydrate derivatives. This technology was compatible with native chemical ligation (NCL).\(^9\) Finally, EPO protein backbone was selectively cleaved with cyanogen bromide. Afterwards, and applying expressed protein
ligation (EPL), an unnatural glycopeptide was introduced on the N-terminus containing GlcNAc. While in all the previous examples a carbohydrate was incorporated either on a peptide or a protein through a native or unnatural glycosidic bond, in all cases, the sugar incorporated was not the physiologically relevant structure found either in GlyCAM-1 or EPO.

Having available previous knowledge accumulated in the research group, the aim this thesis project was to produce either natural or unnatural glycopeptides that can be recognised in a biological environment. During the development of this research I hoped to discover novel methodology that could allow to obtain complex carbohydrate structures, and technology to recognise protein-carbohydrate interactions.

The aims for this project were:

- **Synthesis of core-1 disaccharide O-linked to threonine building block ready to incorporate on a peptide using SPPS (Chapter 3).** This structure can be more readily elaborated to more relevant structure using enzymes.

- **Development of a peptide scaffold displaying carbohydrates (PSDC) based on the work of Meldal (Scheme 2.8),** which could potentially mimic the typical core pentasaccharide found in N-linked glycoproteins. This structure should be equipped with the proper functional groups to attach to a peptide/protein backbone (Chapter 4).

- **Invention of a new methodology to incorporate carbohydrate structures on peptides (Chapter 5).** This new neo-glycopeptide linker should facilitate the incorporation of complex carbohydrate structures such as the PSDC.

- **Development of a technology to visualise protein-carbohydrate interactions using GFP (Chapter 6),** and apply this technology to observe the interactions between the carbohydrate structures developed in the thesis and lectins.
3. Core-1 \(O\)-linked to threonine, a glycodebuilding block for SPPS

3.1 Introduction of amino acid glycodebuilding blocks to SPPS

As discussed in Chapter 1, glycoproteins are the product of co- and post-translational modifications of proteins. This process provides the host protein molecule a new level of complexity and new properties. Moreover, glycosylation of proteins allows biological systems to generate a combinatorial library of different glycoproteins, and therefore, amplifies the genetic code, by modifying the same core protein structure. Unfortunately, this procedure proceeds without gene control, and therefore, a wide variety of heterogeneous mixtures are produced. To understand the role and relevance of these compounds, there is a need for the production of homogeneous and specific glycoproteins. For this purpose, the most convenient pathway to obtain these molecules is through a chemical or enzymatic approach.\(^9^4\)

Synthesis of glycopeptides can be achieved either in solution or on solid phase. In both cases, the most common synthetic route is the introduction of glyco-amino acid building blocks in the stepwise synthesis of the glycopeptide.\(^9^5,9^6\) Alternatively, in \(N\)-linked oligosaccharides, the carbohydrate can be conjugated directly to a full-length peptide.\(^9^7\)

For the synthesis of glyco-amino acid building blocks, a good protecting group strategy must be applied. Protecting groups have to be compatible and robust during the synthesis of the required molecule, the introduction of the glyco-amino acid building block to the peptide, cleavage from the resin in SPPS, and also be easily removed in the final deprotection stage.\(^9^6\)
Synthesis of N-linked glycosyl amino acid building blocks for the production of glycopeptides on solid phase have been widely described in the literature: bis-glycosylated interleukin 8 (IL8) receptor fragment or RNase B fragments. On the other hand, synthesis of the glycobuilding block as glycosyl amines, followed by direct aspartylation in solution with the peptide can be achieved under classical peptide coupling conditions for the synthesis of N-linked glycopeptides.

For the semi-synthesis of O-linked glycopeptides and glycoproteins, a general synthetic approach is the synthesis of glycosyl amino acid building blocks ready for SPPS. In O-linked glycopeptides and glycoproteins, as described earlier (Chapter 1), the most common monosaccharide found is N-acetylgalactosamine (GalNAc), α-O-linked to the amino acid: serine or threonine.

Lemieux and Ratcliffe used 2-azidogalactose donors to ensure α-selectivity in the glycosylation with the amino acid (Scheme 3.1). Final reduction and acetylation of the azide is performed to obtain GalNAc. On the other hand, Schmidt and co-workers used the presence of a “nitroglycal” to perform a Michael addition of the free hydroxyl of the amino acid to the monosaccharide. Subsequent reduction of the 2-nitro group and acetylation of the primary amine provided GalNAc-Ser/Thr for SPPS (Scheme 3.1).

Scheme 3.1: Synthesis of α-O-linked glycosylated amino acid building blocks ready for SPPS.
Wong and co-workers developed a method to convert unprotected amino sugars to azides with retention of stereochemistry by diazo transfer with triflyl azide in the presence of catalytic amounts of copper (II), nickel (II) or zinc (II) (Scheme 3.2). Moreover, this procedure is scalable without using anhydrous conditions. The authors proposed a mechanism where the amino compound complexes with the metal catalyst under basic conditions. Due to the extreme electrophilicity of triflyl azide, nucleophilic attack by the amine to the triflyl azide occurs, followed by deprotonation to form a metal-stabilised tetrazene. Decomposition of the intermediate would provide the azide product and the metal-triflyl imido complex (Scheme 3.2).

Scheme 3.2: Schematic procedure for the synthesis of azido sugars from their amino derivatives. Possible mechanism for the transition metal-catalysed diazotransfer reaction.
Once the carbohydrate is furnished with the required functionalities, it can react with the amino acid to obtain the desired glycobuilding block for SPPS. Although glycosylation of the protected threonine or serine can be achieved successfully, and with high $\alpha$-selectivity using trichloroacetimidate and phosphate methods, normally, glycosylation of the amino acid is achieved employing the Konigs-Knorr procedure.\(^\text{36}\)\(^{109}\)

Two convergent synthetic routes can be followed to synthesise highly elaborate glyco-amino acid building blocks:

- Threonine or serine is glycosylated with 3, 4, 6-tri-$\text{O}$-acetyl-2-azido-2-deoxy-$\alpha$-$\text{D}$-galactopyranosyl bromide, and afterwards, the sugar is deacetylated for further glycosidation of the hydroxyls in order of reactivity: 6-$\text{OH}$>3-$\text{OH}$>4-$\text{OH}$.\(^\text{110}\)\(^{111}\) This synthetic route is called the “cassette approach”.\(^\text{112}\)\(^{114}\) This methodology partially facilitates difficult glycosylation of amino acids, and does not require protecting group manipulation.

- Threonine or serine is glycosylated with a complex oligosaccharide. This strategy is called the “classical approach”.\(^\text{109}\)

Because of the biological implications of $\text{O}$-linked glycoproteins in processes such as antifreezing glycoproteins, inflammation,\(^\text{37}\)\(^\text{115}\) and cancer,\(^\text{116}\) there is huge interest in the area, and different glyco-amino acid building blocks have been developed. Interestingly, the core-$\text{I}$ carbohydrate structure (Scheme 1.3, Chapter 1) forms part of most of $\text{O}$-linked glycoproteins. Therefore, this glycobuilding block is a starting point in many synthetic routes.

Meldal and co-workers used the “cassette approach”, for the production of glycopeptides displaying core-$\text{I}$, core-$\text{2}$, core-$\text{3}$, core-$\text{4}$ and core-$\text{6}$ (Scheme 3.3).\(^\text{112}\)\(^\text{113}\) Moreover, the authors have applied the “classical approach”\(^\text{117}\)\(^\text{118}\) for the development of libraries of glycopeptides displaying T$^\text{a}$ and T antigenic structures,\(^\text{119}\) and glycopeptides as potential
cancer vaccines.\textsuperscript{117} One common feature of these "glycosyl cassettes" is the fact that the carboxylic group is protected either with a tert-butyl or activated pentafluorophenyl ester.

Scheme 3.3: Reagents: i) TMSOTf; ii) Zn, Ac$_2$O, AcOH; iii) Ac$_2$O, pyridine; iv) AcOH (80\% \textsuperscript{\textdegree}C; v) TFA.\textsuperscript{113}

Meldal and co-workers managed to synthesise four different core carbohydrate structures found in mucin domains (Scheme 3.3), iteratively by: i) glycosylation using trichloroacetimidate donors;\textsuperscript{120,121} ii) concurrent reduction of the azide and cleavage of the Teoc- protecting group, followed by acetylation of the free primary amino group;\textsuperscript{122} iii)
acetylation of free hydroxyl groups; iv) deprotection of benzylidene acetal in the presence of 80 % aqueous solution of acetic acid at 80 °C; and finally v) deprotection of tert-butyl ester in the presence of TFA.

Also using the same “cassette approach”, Pratt and Bertozzi performed a semi-synthesis of two different glyco-amino acid building blocks related to 6-sulpho sialyl Lewis X, corresponding to the core-1 and core-6 branches of the L-selectin ligand (Scheme 3.4). Although they did not apply SPPS to these molecules, they demonstrated the immense power of using glycosyl transferases to introduce the remaining monosaccharide units in the later stages of the synthesis, and as a source for creating diversity, to achieve the desired molecules.

Scheme 3.4: Structure of sulphoadhesin with core-1 and core-2 branches highlighted. Synthetic targets corresponding to: i) core-1, and ii) core-6.
In a “classical approach”, Nishimura and co-workers synthesised an antifreeze glycoprotein (AFGPs), by introducing the core-1 structure into a tripeptide using glycosyl fluorides (Scheme 3.5). Further modification of the glycopeptide, and polymerisation of the monomer (H₂N-Ala-Ala-Thr(core-1)-CO₂H), employing diphenylphosphoryl azide (DPPA) as an efficient promoter in the presence of triethylamine allowed the authors to obtain an artificial AFGP between 10-12 glycopeptide units (Scheme 3.5). Physical studies of a glycopeptide library of different lengths of peptide backbone, and displaying either: N-acetyl galactosamine, N-acetyl lactosamine, galactose, lactose or core-1 structure O-linked to threonine or serine demonstrated that analogous glycopeptides displaying the same disaccharide as the native AFGP, regardless of the peptide backbone length, have more antifreeze activity than those which do not display the core-1 structure.

Danishefsky and co-workers, in a “classical approach”, developed a convergent synthesis of a tumour-associated mucin motif (Scheme 3.6). They joined 2,6-sialyl T₅ antigen to...
threonine and serine using a wide variety of glycosyl activators, such as: trichloroacetimidates, phosphates and bromides. Interestingly, the glycosyl bromide was elected as the best donor to glycosylate the protected amino acid with the free hydroxyl group. Afterwards, this glyco amino acid building block was incorporated into a peptide for the construction of a tumour associated glycopeptide.

Scheme 3.6: convergent coupling between 2,6-sialyl T₅ antigen and threonine or serine.¹⁰⁹

3.2 Objective: synthesis of core-1 O-linked to threonine for SPPS

Macmillan research group achieved the semi-synthesis of GlyCAM-1, by combining SPPS and EPL. In this work, 13 units of GalNAc were introduced at predetermined positions to the mucin domains using native glycosidic bonds. Finally the glycoprotein was assembled by EPL.⁸⁹,¹²⁸ Despite developing new technology to produce native glycoproteins, the carbohydrates attached to the protein backbone, GalNAc, were not the physiologically relevant structures found in the native GlyCAM-1. In fact, sialyl Lewis X (sLe⁺) is the oligosaccharide found in this L-selectin counter-receptor (Scheme 3.7).

Based on previous work, the preparative synthesis of core-1 structure molecule (Gal-(β-1,3)-GalNAc) was decided, which could be the starting point for enzymatic elaboration to different carbohydrate structures. To achieve these goals, the synthetic route must be
scalable and high-yielding. Application of this structure to SPPS, and modification of the glycopeptide by either glycosyl transferases, NCL or EPL should lead to a new level of complexity. Having available homogeneous and predefined glycoproteins should allow better understanding of the biological role of the glycoproteins and their derivatives.

Scheme 3.7: Sulphated O-linked chains of GlyCAM-1.

### 3.3 Design and retro-synthetic analysis of core-1 glycobuilding block O-linked to threonine for SPPS

The synthetic strategy taken here was to introduce core-1 disaccharide structure into a peptide as a glycoamino acid using SPPS. For this goal, the target molecule needed a $\beta$-glycosidic bond between the two monosaccharide units; subsequently, this disaccharide was attached to the amino acid, threonine, through an $\alpha$-linkage.

In order to perform the synthesis of the glycopeptide on solid phase using Fmoc-chemistry, it was required to choose the appropriate protecting groups that were stable under the basic conditions of Fmoc deprotection (20%-25% piperidine or 50% morpholine).
in DMF or NMP), compatible with the peptide coupling reactions, robust to the harshly acid conditions of the peptide cleavage, and removable under mild conditions.\textsuperscript{43}

For the protection of the 4- and 6- hydroxyl groups of proximal GalNAc subunit, a level of flexibility and versatility was introduced as this could allow, in the future, the introduction of N-acetyl glucosamine (GlcNAc) to enable the synthesis of the core-2 carbohydrate structure or even more ambitious structures. Bearing these precepts in mind, the benzylidene acetal protecting group for 4- and 6- hydroxyls of GalNAc, emerged as one of the potential possibilities: the benzylidene acetal can be selectively deprotected with Na(BH\textsubscript{3})CN in TFA to afford the free 4-hydroxyl group. Conversely, selective deprotection to afford 6-hydroxyl occurs in the presence of LiAlH\textsubscript{4} and AlCl\textsubscript{3} in ether.\textsuperscript{42} On the other hand, acetylation of the hydroxyl groups could fulfil the requirements needed for the synthesis of the target molecule. The acetyl protecting group is robust under a broad set of conditions, including Fmoc- chemistry; and it is easy to deprotect by Zemplen deprotection.\textsuperscript{66} Moreover, due to the nature of the $\beta$-linkage between the two monosaccharides of core-1 structure, neighbouring group participation of the acetyl group in the 2-hydroxyl group of galactose (Gal) subunit would be necessary to control the stereoselectivity of the glycosyl linkage.\textsuperscript{42,130}

The natural requirement of having an N-acetyl group of GalNAc in the position 2' of core-1 structure $\alpha$-O-linked to an amino acid (serine or threonine), required the use of the anomeric effect rather than the neighbouring group effect to perform the glycosylation of the amino acid. The presence of an azide in the position 2' of the disaccharide could fulfil these natural requirements (donor 2, Scheme 3.9). Moreover, mild reduction conditions followed by acetylation of the resulting primary amine should lead to the desired final functionality.

A wide variety of protecting groups are available for the protection of the carboxylic group of the amino acid. There are precedents in the literature for the use of ester groups, such as: benzyl or tert-butyl ester; or even pentafluorophenyl esters (Pfp).\textsuperscript{95,96} Unfortunately, the
requirement to perform glycosylation of the protected amino acid under mild acid conditions, discarded the use of tert-butyl ester, due to its acid-lability.

In the retro-synthetic analysis (Scheme 3.8), the glyco-amino acid building block is assembled by reacting the fully protected disaccharide equipped with a good leaving group such as fluoride, bromide or trichloroacetimidate; and Fmoc threonine, in which the carboxylic acid could be masked either with a benzyl or Pfp-ester. Obviously, the protected threonine would come from the natural unprotected amino acid 70. Formation of an α-linkage between the carbohydrate and the free hydroxyl group of the amino acid, would be facilitated by non-participating azido group at C-2.

Scheme 3.8: retro-synthetic analysis for the amino acid glycobuilding block for SPPS.

The disaccharide would come from a glycosyl acceptor with a free C-3-hydroxyl group, 56, and a peracetylated donor with trichloroacetimidate as a good leaving group, 23, that would
3. Core-1 O-linked to threonine, a glycobuilding block for SPPS

provide a \(\beta\)-linkage (Scheme 3.8). Both donor and acceptor units would come from commercially available D-galactose, 48.

3.4 Synthesis of core-1 glycobuilding block O-linked to threonine for SPPS

3.4.1 Synthesis of glycosyl acceptor

The synthesis of the glycosyl acceptor 56 was successfully accomplished on a multigram scale (Scheme 3.9), starting with commercially available D-galactose 48.

Galactose 48 was peracetylated with acetic anhydride, in the presence of a catalytic amount of perchloric acid, affording a quantitative reaction within 3 hours (Scheme 3.9). Five singlets between 2.18 and 1.97 ppm integrating for 15 protons revealed the peracetylation of all hydroxyl groups by \(^1\)H-NMR. Moreover, a small coupling constant for H1 denoted the generation of the thermodynamic \(\alpha\)-product \((J_{1,2} = 1.7\) Hz).

Halides are commonly used as leaving groups in nucleophilic substitution. Bromides, chlorides and fluorides are the most commonly used as anomeric halides. They are particularly useful as glycosyl leaving groups. Moreover, glycosyl bromides could be used for the generation of glycals. Synthesis of the glycosyl bromide 50 was initially attempted using titanium tetrabromide, TiBr\(_4\). Unfortunately, a poor yield of only 37 % was achieved. On the other hand, \(\alpha\)-bromination was obtained quantitatively when the peracetylated galactose was treated with HBr in acetic acid (Scheme 3.9). The presence of two isotopic species was observed in FAB-MS (found 409.0 and 411.0), confirming the presence of bromine in the molecule.
Scheme 3.9: Synthesis of glycosyl acceptor.

For the production of the galactal 51 through reductive elimination, there are several precedents in the literature. The reaction was accomplished in high yield by treating with zinc in acetic acid. In this reaction there was a simultaneous elimination of the bromide and 2-acetyl group, producing a cyclic carbohydrate enol ether, termed a glycal.

As discussed earlier in this chapter (Section 3.4), it is essential to have and N-acetyl group at the C-2 position of the galactopyranose in core-1. Moreover, this monosaccharide has to be α-O-linked to the peptide/protein. Therefore, to achieve this target molecule, the most common procedure is the introduction of a non-participating azido group at C-2 by azidonitration. Later on, it can be easily converted to an N-acetyl group by consecutive reduction and acetylation.
The azidonitration of protected glycals disclosed by Lemieux and Ratcliffe in 1979\textsuperscript{104} is one of the most common routes for the generation of 2-azido sugars, although other successful procedures have been described.\textsuperscript{107,108,134-136} In this reaction (Scheme 3.9), the crude glycal 51 was dissolved in anhydrous acetonitrile, and then, cerium ammonium nitrate (CAN) and sodium azide were added. The oxidation of the anomeric position of the glycal with CAN facilitated the regioselective incorporation of the azide group at C-2 and the nitrate group, to produce the desired molecule (Scheme 3.10). The regiospecificity of this procedure could be due to the stability of the radical formed at the anomeric carbon, which was oxidised again with CAN to obtain a carbocation that was consequently attacked by the nitrate. The formation of the N-acetyl group at the anomeric position of one of the side products 57 could be related to a competitive nucleophilic attack of the solvent, acetonitrile, to the carbocation intermediate, decreasing the overall yield of the reaction to 44 %.

Scheme 3.10: plausible mechanism for the introduction of a non-participating azido group at C-2 by azidonitration.

Anomeric denitration is often achieved with sodium nitrite in aqueous dioxane at 80 °C for 3 days, achieving the resulting hemiacetal in a good yield. On the other hand, and according to
a shorter procedure developed by Sina$^{137,138}$ and co-workers, the reaction was carried out with thiophenol in the presence of DIPEA in dry acetonitrile. The hemiacetal 53 was formed in 100% yield and was isolated as an anomeric mixture ($\alpha/\beta$; 1:4) (Scheme 3.9).

Silyl ethers are commonly used for the protection of alcohols due to their facile introduction, relatively good stability and easy deprotection. This type of protecting group is selectively removed by treatment with a source of organic soluble fluoride, such as tetrabutylammonium fluoride (TBAF) or mixtures of HF/pyridine.$^{43}$ Therefore, for the masking of the anomeric hydroxyl of the hemiacetal 53, the reaction was carried out with tert-butyldimethylsilyl chloride in the presence of imidazole, which also participates by activating the silyl chloride, achieving a 88% yield for the protected molecule 54 (Scheme 3.9). Due to the bulkiness of the silyl group, the only product obtained was the $\beta$-anomer, and its structure was confirmed by $^1$H-NMR: $J_{\alpha,\beta} = 7.6$ Hz. This result suggested that during the course of the reaction anomerisation occurred.

Deacetylation of the monosaccharide was accomplished by dissolving the acetylated sugar 54 in methanol, and treating the mixture with a base, sodium methoxide, for 20 hours (pH below 9).$^{66}$ When the reaction was completed, the crude mixture was neutralised with Amberlyst-15 acid ion exchange resin (Scheme 3.9). Special care must be taken when adding the base in the reaction (pH below 8.5), to avoid intermolecular migration of the silyl group from the anomeric position to the deacetylated 6-hydroxyl group.$^{43}$

Protection of 55 with benzaldehyde dimethylacetal was performed with a catalytic amount of $p$-tosic acid at 50 °C to obtain the desired product 56 in 89% yield (Scheme 3.9).

The acceptor molecule 56 possessed all the requirements defined for the glycosidation with the donor 23. While all the hydroxyl groups of the molecule were strategically protected, the
C-3-hydroxyl group was free. Therefore, the next aim was the synthesis of the glycosyl donor 23.

3.4.2 Synthesis of glycosyl donor

Synthesis of the glycosyl donor was accomplished in three steps (Scheme 3.11). The synthetic route also started with the commercially available D-galactose, which was peracetylated with acetic anhydride and catalytic perchloric acid.

Anomeric deacetylation was pursued for the introduction of a trichloroacetimidate as a leaving group for glycosidation of the glycosyl acceptor 56. Different approaches were studied, either by direct or indirect hemiacetal formation.

Previous work carried out in the Macmillan research group achieved the generation of the hemiacetal 58 in a preparative two steps synthesis (Scheme 3.11). 49 was converted to a thioglycoside 59, and this was followed by the conversion of the molecule into the desired hemiacetal 58 with NBS, acetone and water in half an hour.

Although the overall yield over two steps was 68 %, direct hemiacetal formation was studied to reduce the number of the synthetic and chromatographic steps, avoid the use of large
amounts of toxic and odorous thiophenol, and to increase the yield for the anomeric deprotection.

\[
\begin{align*}
\text{AcO} & \quad \text{OAc} \\
\text{OAc} & \quad \text{OAc} \\
\text{AcO} & \quad \text{AcO} \\
\text{AcO} & \quad \text{OAc}
\end{align*}
\]

\[\text{PhSH, SnCl}_4, \text{DCM} \quad 87\% \]

\[
\begin{align*}
\text{AcO} & \quad \text{OAc} \\
\text{OAc} & \quad \text{OAc} \\
\text{AcO} & \quad \text{AcO}
\end{align*}
\]

\[\text{NBS, acetone, water} \quad 78\% \]

**Scheme 3.12:** synthesis of the hemiacetal through a thiophenol intermediate.

There are different procedures for selective anomeric deprotection in the literature. In the first attempt, the protected monosaccharide 49 was treated with ammonium carbonate in methanol to achieve just 16% yield of the hemiacetal 58, and other by-products as a result of non-selective deacetylation of the peracetylated molecule (Scheme 3.13). When the molecule was subjected to hydrazine and acetic acid, \( \text{N}_2\text{H}_4, \text{AcOH, DMF} \) 58 was isolated in 53% yield, and 47% yield of the starting material 49. Finally, 95% yield of the hemiacetal 58 was efficiently obtained when the peracetylated galactose 49 was treated with benzylamine \( \text{Bn-NH}_2 \) in THF at 50°C for 24 hours (Scheme 3.11 and 3.13).

\[
\begin{align*}
\text{AcO} & \quad \text{OAc} \\
\text{OAc} & \quad \text{OAc} \\
\text{AcO} & \quad \text{AcO} \\
\text{AcO} & \quad \text{OAc}
\end{align*}
\]

\[\text{(NH}_4\text{)CO}_3, \text{MeOH, THF} \quad 16\% \]

\[
\begin{align*}
\text{AcO} & \quad \text{OAc} \\
\text{OAc} & \quad \text{OAc} \\
\text{AcO} & \quad \text{AcO} \\
\text{AcO} & \quad \text{OAc}
\end{align*}
\]

\[\text{NH}_2\text{--NH}_2, \text{AcOH, DMF} \quad 53\% \]

\[
\begin{align*}
\text{AcO} & \quad \text{OAc} \\
\text{OAc} & \quad \text{OAc} \\
\text{AcO} & \quad \text{AcO} \\
\text{AcO} & \quad \text{OH}
\end{align*}
\]

\[\text{Bn-NH}_2, \text{THF} \quad 95\% \]

**Scheme 3.13:** different approaches for anomeric deprotection.
As described earlier (Chapter 2, section 2.1.2.), there is a wide variety of leaving groups for glycosidation. Between all these groups, trichloroacetimidates, introduced by Schmidt and co-workers\textsuperscript{120} using as an activator the Lewis acid BF\textsubscript{3}-OEt\textsubscript{2}, allows the formation of the oxo-carbenium ion, and consequently, this can be trapped by a nucleophile such as a free hydroxyl (Scheme 3.14).\textsuperscript{121} Introduction of trichloroacetimidates to a glycosyl donor is achieved by reacting the hemiacetal with trichloroacetonitrile under basic conditions (NaH, DBU or K\textsubscript{2}CO\textsubscript{3}) in dry DCM.\textsuperscript{54} The generation of an amide is the driving force for the glycosidation reaction. Addition of activated molecular sieves avoids the formation of the hydrolysed adduct. In this reaction, different types of Lewis acids can be used as a catalyst to generate mild acid conditions to activate the glycosyl donor: BF\textsubscript{3}-OEt\textsubscript{2},\textsuperscript{120} TMS-OTf\textsuperscript{121} or Ag-OTf. Glycosyl trichloroacetimidates can be isolated and stored. This elegant methodology has been demonstrated to be robust and efficient in different glycosidation contexts.\textsuperscript{96}

Scheme 3.14: Introduction of the trichloroacetimidate in the glycosyl donor 58 under basic conditions, and glycosidation of a glycosyl acceptor 60 in the presence of a Lewis acid catalyst to produce a more complex carbohydrate structure 61.

The introduction of trichloroacetimidate into the hemiacetal 58 was initially conducted with K\textsubscript{2}CO\textsubscript{3} in dry DCM, resulting in a 71 % yield of the anomeric mixture of $\alpha$ and $\beta$ (2:3),
separately isolated by silica flash chromatography. A considerable increase in the yield was achieved by treating the hemiacetal with trichloroacetonitrile in the presence of 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU), producing exclusively, the \( \alpha \)-anomer product, \( \alpha-23 \), in 90% yield (Scheme 3.15).

**Scheme 3.15:** Introduction of trichloroacetimidate to the hemiacetal 23.

3.4.3 **Synthesis of \( N^\alpha-(\text{fluorene-9-nylmethoxycarbonyl})-O-[2, 3, 4, 6\text{-tetra-}O\text{-acetyl-D-galactopyranosyl-(1,3)}]-4, 6\text{-diacetyl-2-acetamido-2-deoxy-}\alpha\text{-D-galactopyranosyl]-L-threonine}.**

Upon formation of the synthesis of the glycosyl donor 23 with the trichloroacetimidate as an activator, and the glycosyl acceptor 56 with the free 3-hydroxyl group, the glycosidation reaction between these two monosaccharides was performed (Scheme 3.16).

The \( \alpha \) and \( \beta \) epimeric mixture of the glycosyl donors 23 was activated with TMS-OTf under strictly anhydrous conditions; and reacted with the free 3-hydroxyl group of the glycosyl acceptor (Scheme 3.16). Despite the lability of the benzylidene acetal group in the acceptor monosaccharide under the acidic conditions of the glycosidation, the reaction afforded the product in 80% yield. The \( \beta \)-stereochemistry of the glycosidated product at the anomeric centre was confirmed by \(^1\text{H-NMR} \,(J_{1,2} = 8.0 \text{ Hz}).}
After the glycosidation between the glycosyl donor 23 and the acceptor 56, the potential routes for coupling the core-1 disaccharide to threonine to obtain the desired product were investigated. To achieve this goal, it was planned to introduce a fluoride, bromide or trichloroacetimidate as a leaving group for glycosylation (Scheme 3.17). Moreover, it was investigated which protecting groups within the disaccharide and the amino acid were the most appropriate for the glycosylation between the two molecules.

Silyl deprotection of disaccharide 63 was performed using Takano’s procedure, with tetrabutyl ammonium fluoride (TBAF), buffered with acetic acid in THF at room temperature overnight (Scheme 3.17), obtaining a 76 % yield for the reaction.

The first reasonable approach to obtain the desired disaccharide equipped with the leaving group on the anomeric position to glycosylate the amino acid, based on previous experience, was the direct introduction of the trichloroacetimidate. This was performed following the procedure of Meldal and co-workers: trichloroacetonitrile, DBU in dry DCM at 0 °C for 2 hours, achieved a 100 % yield of the disaccharide as anomeric mixture, with the desired activator 64 (Scheme 3.17).
3. Core-1 \(O\)-linked to threonine, a glycobuilding block for SPPS

**Scheme 3.17:** synthesis of disaccharide donor equipped with different leaving groups (trichloroacetimidate, bromide or fluoride), and protecting groups.

Synthesis of glycosyl fluorides can be achieved by reacting the hemiacetal with DAST in THF\(^{142}\) or by reacting a thioglycoside with DAST and NBS\(^{143}\). The reactivity of the glycosyl fluorides is quite similar to the bromides and chlorides, although several advantages make them more attractive: easy synthesis under mild conditions, relatively high stability towards manipulation and storage, and wide range of methods for glycosylation\(^ {52}\). One of the disadvantages observed was that during activation, the generation of fluoride ion could be inconvenient if the donor or acceptor is protected with silyl groups, leading to the undesired deprotection and possible undesired glycosylation (Chapter 4). Nevertheless, having available the hemiacetal 63, the fluoride was introduced at the anomeric position using DAST in THF to achieve a poor yield (23\%) of the product 65\(^ {142,144}\). The presence of the fluoride was detected by \(^{19}\)F-NMR spectroscopy, -146 ppm, while \(\alpha\)-fluorination of the structure was confirmed by \(^1\)H-NMR spectroscopy, which showed a small constant between H-1' and H-2' \((J_{1',2'} = 2.5\text{ Hz})\).
On the other hand, the possibility of changing the benzylidene acetal for acetyl groups was envisaged, as this could decrease steric hindrance and rigidity of the ring structure during the formation of the oxo-carbenium intermediate in the glycosylation process, and therefore, increase the final yield of the process. To change the protecting groups, an extensive study to find suitable conditions for this purpose was performed. Firstly, the hemiacetal 63 was treated with catalytic amounts of perchloric acid in acetic anhydride to achieve a 52 % yield of the purified product 66. Afterwards, adapting a procedure from Meldal and co-workers, the disaccharide 63 was treated with acetic anhydride and sulphuric acid for 15 hours at 0 °C to achieve 88 % of 66 on a 0.32 mmol scale (200 mg). Unfortunately, this reaction was temperamental when the reaction was scaled up, promoting the decomposition of the components under such harshly acidic conditions. By changing sulphuric acid for trifluoroacetic acid, acetylation of the hemiacetal was achieved without damaging the benzylidene acetal group. Finally, in order to achieve the peracetylation of the molecule, the disaccharide 63 was treated first with an aqueous solution of acetic acid 80 % for 2 hours at 80 °C, and then with acetic anhydride/pyridine (1:2) to achieve a 100 % yield of the peracetylated molecule 66. Moreover, this procedure proved to be more robust when scaled up to preparative quantities (1.54 g, 2.47 mmol scale).

Based on previous experience, the anomeric deacetylation was pursued for the formation of the hemiacetal with benzylamine in THF, to achieve a 68 % of the desired product 67. Introduction of the trichloroacetimidate was conducted using the same conditions previously described to achieve 68 in 58 % yield. On the other hand, introduction of the bromide at the anomeric position of the disaccharide proved to be simpler, by treating the peracetylated molecule 66 with titanium tetrabromide to produce the α-bromide 69 in 91 % yield (Scheme 3.17). The α-stereochemistry of the brominated product was confirmed by 1H-NMR ($J_{1,2}$ = 3.8 Hz). No column chromatography purification was required to isolate the pure product.
Having a small panel of disaccharides with different glycosyl activators (Scheme 3.17), the requirements for the amino acid component were addressed. As mentioned before (Section 3.4), the acceptor molecule must have an Fmoc- group to protect the primary amine; and an ester resistant to potentially acidic conditions of glycosylation, but quantitatively removable when required, to protect the carboxylic acid. Benzyl and pentafluorophenyl esters emerged as potential protecting groups. Obviously, the hydroxyl group of the threonine must be accessible for glycosylation.

FmocThr(OBn) 46 was already available in the Macmillan research group. Generation of FmocThr(OPfp) 72 was achieved following literature procedures. The primary amine of the threonine 70 was protected with the fluorenyl protecting group through a reaction with FmocOSu in water and acetonitrile, achieving 100 % yield. Afterwards, the formation of the Pfp ester was achieved by treatment with pentafluorophenol and DCC in ethyl acetate to achieve 75 % yield (Scheme 3.18).

Scheme 3.18: Synthesis of FmocThr(OPfp) 72.

Having access to both glycosyl donor and the amino acid acceptor, in suitably protected forms, trial glycosylation conditions were studied.

In preliminary studies, suitable conditions for the glycosylation of threonine were explored. Model experiments were carried out between peracetylated galactose trichloroacetimidate, and Fmoc- threonine, either with benzyl 46 or Pfp ester 72 protecting the carboxylic acid (Scheme 3.19). Similar conditions were applied for the formation of the disaccharide 62. In the presence of TMS-OTf in dry DCM at -50 °C, the benzyl ester amino acid 46 tended to be
hydrolysed faster than the PfP derivative 72 consequently having a dramatic effect on the final yield of the product: 30 % and 70 % respectively. These results were in accordance with Meldal's research work on glycosylation of PfP ester amino acids, and pivotal, because it drew the project towards the decision to choose the activated ester as a protecting group for the amino acid, when glycosylation with the core-1 structure was going to be realised in the presence of TMS-OTf or AgOTf.

Applying this knowledge, it was decided to glycosylate the PfP ester derivative of threonine 72 with the disaccharide 64. Using the same conditions previously described, an epimeric mixture of the α- and β-anomer 75 in 32 % and 12 % yield respectively was obtained. Unreacted amino acid 72 was also recovered (Scheme 3.20).

Scheme 3.19: Model experiments for glycosylation on different protected threonine acceptors.

Scheme 3.20: Glycosylation between glycosyl donor and threonine PfP ester acceptor.
The crude mixture was difficult to purify by flash chromatography due to the large amount of undetermined side products. Moreover, TLC demonstrated that some decomposition of the Pfp ester occurred on silica gel. Similar results with 64 and 72 were obtained when AgOTf was used as activator. As predicted (Section 3.3),47 better conversion was obtained under these same conditions when the peracetylated disaccharide 68 was employed to afford 76: α-anomer: 36 %; β-anomer: 18 % (Scheme 3.20). Purification of the crude mixture and stability of the products under the chromatography conditions, seemed to be challenging.

In an attempt to increase the yield of the glycosylation, reactivity of the glycosyl fluoride 65 (Scheme 3.17) with threonine was studied. Glycosyl fluorides were introduced by Mukaiyama.151 Glycosylation was achieved using AgClO4 and SnCl2 as promoters. This methodology has been expanded to a repertoire of other promoters: BF3·Et2O,152 TMSOTf,153 Cp2HfCl2,154 Cp2ZrCl2·AgBF4 and Cp2HfCl2·AgTfO/AgClO4126 and Cp2ZrCl2·AgClO4.125

Having all these precedents in the literature, and based on Nakahara’s work155 glycosylation of threonine derivative 46 was pursued using the modified core-1 structure 65 in the presence of Cp2ZrCl2 and AgClO4 was carried out. Surprisingly, the benzylidene acetal protecting group did not withstand the acidic conditions of the reaction, and was cleaved to provide the partially deprotected adduct as an epimeric mixture in 35 % yield.

To increase the yield for the glycosylation of threonine, use of bromide as a leaving group for glycosylation of threonine was investigated.

Glycosyl bromides were the first type of glycosyl donors to be used for disaccharide formation more than a century ago by Knoenigs-Knorr.50 This reaction involves the coupling of a glycosyl bromide (or chloride) with a hydroxyl component in the presence of a halophilic promoter, typically a heavy metal ion (silver or mercury). The original activators used were insoluble silver oxide or carbonate (Ag2O or Ag2CO3) but these have been superseded by the use of soluble silver triflate or perchlorate (AgOTf or AgClO4), or by
other types of soluble mercury salts (e.g. HgBr₂ or Hg(CN)₂).¹⁵⁷ The general mechanism of this reaction (Scheme 3.21) involves the \textit{in situ} generation of the more reactive β-halide, followed by glycosylation with inversion to achieve the α-glycoside. Construction of the β-glycoside may be accomplished by neighbouring group participation of 2-acyl substituents.

![Scheme 3.21: Proposed mechanism of the Koenigs-Knorr glycosylation reaction.][1]

Glycosyl bromides are not particularly stable and normally they are synthesised immediately prior to being used for glycosylation. Moreover they need at least a stoichiometric amount, or even more, of a heavy metal promoter making the disposal of the metal a problem with the methodology.

Using the reaction conditions for glycosylation of amino acids with glycosyl bromides developed by Danishefsky and co-workers,¹⁰⁹ the disaccharide 69 reacted with the benzyl ester threonine derivative 46 in the presence of AgClO₄ in dry DCM at room temperature to obtain 34% of the α-anomer 77 and 18% of the β-anomer. Interestingly, apart from the unreacted threonine derivative 46, the presence of the hydrolysed disaccharide 78 was also observed (44% yield). This result could give the potential to recycle the disaccharide by: i) acetylation the hemiacetal, and ii) re-bromination of the donor. On the other hand, by simply adding 2,4,6-collidine in the reaction (2.5 equiv), a 45% of the desired α-anomer 77 was achieved without any trace of the β-derivative. A 22% yield of the hydrolysed disaccharide 78 was recovered, and this was acetylated successfully with acetic anhydride and pyridine.
3. Core-1 O-linked to threonine, a glycobuilding block for SPPS

(1:2) in 90 % yield to afford the peracetilated molecule 66, which was recycled to increase the overall yield of the process (Scheme 3.22).

Scheme 3.22: Glycosylation of the threonine derivative 46 with α-bromide, 69.

In this glycosylation process it was observed that the yields were lower than predicted. This result could be due to an intramolecular hydrogen bonding interaction between the hydrogen of the amino group and the pair of electrons of the oxygen of the hydroxyl of the acceptor 46, decreasing dramatically the nucleophilicity of the amino acid. Polt and co-workers\textsuperscript{159,160} by-passed this drawback by transforming the primary amine into an imine. Therefore, benzyl-N-(diphenylmethylene)-L-threonine 79 was synthesised by treating the benzyl ester amino acid with diphenylmethylene imine in DCM for 24 h to achieve a 89 % yield.\textsuperscript{161} Following the Polt procedure,\textsuperscript{166} 69 reacted with 79 in the presence of silver perchlorate and collidine in dry DCM, to achieve the desired product 80 in 44 % yield (Scheme 3.23). This low yield is partially in accordance with the results that Polt and co-workers obtained for 3,4,6-tri-O-acetyl-2-deoxy-2-azido-1-bromogalactose.\textsuperscript{160}

Having the previous results, and therefore, consequent deprotection of the amino group of the glyco-amino acid, and re-protection with Fmoc- group, made this final approach nonviable for the preparative intentions of the present work.
3. Core-1 O-linked to threonine, a glycobuilding block for SPPS

Scheme 3.23: Glycosylation of the threonine derivative 79, by the brominated core-1 structure 69.

In conclusion, the overall study for glycosylation of threonine led to two different approaches: i) glycosylation of the Pfp ester threonine derivative 72 with the trichloroacetimidate derivative, either 64 or 68, lead to a epimeric mixture of 75 and 76 respectively (Scheme 3.20); but conversely, ii) glycosylation of threonine benzyl ester derivative 46 with brominated disaccharide 69 in the presence of 2,4,6-collidine provides only, the α-anomer product 77 (Scheme 3.22).

Once the core-1 structure 77 was formed (Scheme 3.24), the reduction of the azide at the C-2 position, followed by subsequent acetylation of the amino group, was achieved successfully in a one-pot fashion, affording the N-acetylated product 81 in 95% yield.

Scheme 3.24: Synthesis of core-1 structure α-O-linked to threonine 82 for SPPS.

Finally, the removal of the benzyl ester protecting group was performed by catalytic hydrogenation using 10% palladium on charcoal in methanol, in 3 hours. The
hydrogenolysis afforded the generation of the desired glyco-amino acid building block 82 in 95\% yield (Scheme 3.24).

Having available the different glyco-amino acid building blocks ready for SPPS: α-75, α-76 and 82, it was intended to assess the insertion of these structures into a peptide, using standard procedures of SPPS (Scheme 3.25). The testing glycopeptide sequence for these molecules was: Cys·Ser·Ser·Glu·Leu·Glc·Thr(core-1)·Ser, molecule 83. The synthesis of the glycopeptide was carried out on a pre-loaded Fmoc-Ser(tBu)-NovaSyn TGT resin (loading = 0.22 mmol/g). The scale of the synthesis was 0.10 mmol. For each peptide coupling it was used 10 equivalents of each amino acid with HBTU/HOBt and DIPEA, for 3 hours. On the other hand, 1.5 equivalents of the glyco-amino acid building block reacted with the peptide on the solid support, in the presence of 1.5 equivalents of DCC/HOBt, for 24 hours in DMF (see Experimental section).

Scheme 3.25: glyco-amino acid building blocks α-75 and 82 ready for SPPS for the synthesis of the glycopeptide sequence: Cys·Ser·Ser·Glu·Leu·Glc·Thr(core-1)·Ser.

Purification of α-75 and α-76 tended to be very difficult for small amounts of the crude reaction. But when the production of these compounds was scaled up, the purification proved impossible. In addition, the purification of α-75 from the Pfp ester linked to threonine 72 was unsuccessful. This contamination of the compound α-75 was a limitation of this synthesis of the desired glycopeptide. When it was applied on SPPS, a mixture of two
products was obtained: the desired glycopeptide and the unglycosylated peptide. An obvious explanation could justify this result: traces of the contaminant, Pfp ester threonine 72 reacts faster with the growing peptide than the desired glycosylated reagent α-75. This could be due to the fact that the impurity has lower molecular weight than the fully elaborated glyco-amino acid building block. Moreover, the presence of a disaccharide on the activated Pfp-amino acid could reduce the accessibility of the primary amine of the growing peptide, due to steric hindrance. This produced an inevitable loss of yield. This drawback obligated to obtain highly purified glyco-amino acid building blocks to react on SPPS.

The molecule 82 seemed to meet all the previous requirements of purity and scalable synthesis. When this component was applied to SPPS, applying the same reaction conditions (see Experimental section), the desired glycopeptide was obtained after cleavage from the resin. Finally, the disaccharide of the glycopeptide was deacetylated in phosphate buffer 0.1 M, pH 8 solution, with 5 % hydrazine hydrate and 10 % DTT; obtaining the desired glycopeptide 83 in 29 % yield (Scheme 3.26).

Scheme 3.26: LC-MS of the purified desired glycopeptide 83; MWcalc: 1347.52, MWobs: 1348.
This last successful approach for the synthesis of the molecule 83, established the glyco-amino acid building block 82 as the chosen structure for SPPS, for the synthesis of glycopeptides.

### 3.5 Conclusions and future work

The successful synthesis of the glyco-amino acid building block 82 opened a wide variety of different research opportunities, using the molecule as starting material for SPPS.

As part of Macmillan research program, development of technology for the production of pure glycopeptides and glycoproteins is a priority. Consequently, bacterial expression of active human core-2 $\beta$-GlcNAc transferase is part of the interest of the Macmillan research group. This enzyme is a mammalian glycosyl transferase that transfers a single unit of GlcNAc to core 1, by a $\beta$-1,6 linkage, employing $\alpha$-UDP-GlcNAc. The glycopeptide 83 could help to evaluate the activity of this new active human core-2 $\beta$-GlcNAc transferase. In the presence of this bacterial expressed enzyme, and $\alpha$-UDP-GlcNAc; and in the appropriate medium, an increase in the molecular weight of the glycopeptide substrate should be observed by LC-MS, related to the transfer of a unit of GlcNAc to 83. However, it has not been possible to overexpress soluble and active core-2 $\beta$-GlcNAc transferase from E. coli and consequently this approach has had to be abandoned.

Regardless of the successful synthesis of core-1 structure $O$-linked to threonine ready for SPPS 82; the synthetic route is notoriously laborious and time consuming. Moreover, although the synthesis of this structure is convergent, through a "classical approach", the glycosylation of threonine requires very strict conditions of dryness, and still, the yield is low for such a valuable disaccharide.
Owing to the complexity of the synthesis of the core-1 glyco-building block, and lack of available enzymes, it was decided to explore more productive methodologies to elaborate complex carbohydrate structures based on Meldal's research work (Chapter 4). In addition, preliminary investigations using triazole formation to make modified glycoproteins were proving more successful and this avenue was also explored (Chapter 5).
4 Synthesis of a novel unnatural peptide scaffold to display carbohydrates (PSDC)

4.1 Introduction

Production of native glycoproteins is a challenging task. While the total synthesis of these structures has been achieved by the Macmillan research group and by others,\textsuperscript{89,162-164} the methodology remains far from simple. The difficulties stem from long and laborious synthesis of complex carbohydrates; high demand for stereo- and regiocontrol of each glycosidation and low efficiency of protein glycosidation.

As previously discussed, a number of attempts have been made to address these issues. Promising strategies include improved methods of glycosidation/ glycosylation,\textsuperscript{51} solid phase oligosaccharide synthesis,\textsuperscript{165} or enzymatic carbohydrate synthesis.\textsuperscript{166} Moreover, NCL and EPL\textsuperscript{167} have provided the opportunity to assemble building blocks for the production of glycoproteins. However, with all these advances, production of glycoprotein and glycopeptides is still difficult.

One option to tackle these difficulties is the synthesis of glycoprotein/glycopeptide mimics. These compounds have been created to study saccharide function, modification of the targeting properties, improvement of the biological activity, and ease of synthesis. Different strategies used for preparation of glycoproteins mimics exist (Scheme 4.1). In the first approach, the protein backbone remains, but the saccharide is substituted. In another approach, the protein is unchanged, the saccharide may or not may be replaced, but the
4. Synthesis of a novel unnatural peptide scaffold to display carbohydrates

linkages between the carbohydrate and protein and the location of glycosylation are altered. In the most radical approach, the protein scaffold is either altered or completely abandoned.\textsuperscript{81}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{scheme41.png}
\caption{Different approaches for designing glycoprotein mimetics.\textsuperscript{81}}
\end{figure}

Introduction of neo-glycopeptide linkages, through chemoselective ligation,\textsuperscript{168} using eg. acetamides,\textsuperscript{169} disulphide bond linkages,\textsuperscript{170} or “click chemistry” glycosydic bonds,\textsuperscript{171} has facilitated the glycosylation of peptides or proteins (Chapter 5). In contrast, the construction of complex carbohydrate structures remains challenging.

Kent and co-workers\textsuperscript{172,173} accomplished the total synthesis of an homogeneous polymer-modified erythropoiesis protein (SEP). An unnatural tetra-branched polymer was introduced into predetermined positions (Lys24 and Lys126) using a chemical ligation\textsuperscript{168} (Scheme 4.2, A). One of the aims of this work was to avoid the synthesis of the complex polysaccharide found in EPO (Scheme 1.4). Introduction and display of negative charges at physiological pH was achieved through appending carboxylic acids to the polymeric chain (Scheme 4.2, B, carboxylic acids highlighted in red). This structure was intended to mimic the native tetraentanery polysaccharide that displays negative charges through the presence of sialic acid residues at the ends of the sugar (Scheme 1.4).

The polymeric chain to the peptide fragments (SEP 1-32 and SEP 117-166) was introduced by oxime-forming ligation at predetermined positions. Ordered sequential ligations were
carried out to connect the different building blocks. Final folding of the molecule provided the desired unnatural glycoprotein (Scheme 4.2, C).

### Scheme 4.2: Synthesis of Synthetic Erythropoiesis Protein (SEP).

A) Diagram of SEP indicating the amino acid sequence, disulphide bonds and the modified lysines (Lys24 and Lys126), and cysteines (Cys89 and Cys117). B) Structure of branched, negatively charged, polymer moiety. C) Scheme for the synthesis of SEP by chemical ligation.  

This unnatural glycoprotein exhibited specific activity in vitro and extended half life in vivo, relative to human erythropoietin (EPO). This was attributed to a decrease of metabolism due to the protective effect of the negative charges of the carboxylic acids of the polymeric chain. Unfortunately, this molecule showed an immunological response in humans during clinical studies (information provided by Prof. Stephen B. H. Kent, December 2005).

Nishimura and co-workers attempted to increase the solubility and decrease hepatic decomposition of Insulin (a non-glycosylated pancreatic peptide hormone that controls the glucose levels in blood, and is therapeutically used for the treatment of diabetes). The authors employed a transglutaminase (TGase) to selectively introduce a lactose at the
terminal glutamine residue of a mutagenic insulin: Ins(B-F1Q). Sia-2,6-Lac was subsequently obtained through the use of a recombinant α-2,6-sialyltransferase (Scheme 4.3).

This insulin-mimic demonstrated a longer half life in vivo than either the wild type or the lactose modified peptide (Lac-Ins(B-F1Q)). Extending this principle with mono-, di- or trisialyloligosaccharide dendrimers, it was demonstrated that the in vivo half life of the modified peptide increases with the number of sialic acids displayed.175

Scheme 4.3: Synthesis of glycosylated insulin (Sia2,6-Lac-Ins(B-F1Q)).174
4. Synthesis of a novel unnatural peptide scaffold to display carbohydrates

4.2 Objective: synthesis of a Peptide Scaffold Displaying Carbohydrates (PSDC)

Based upon literature precedents concerning the introduction of unnatural substrates that mimic sugars to peptides or proteins, possibility to synthesise a PSDC through the residues of the amino acids was envisioned. The resulting structure could then be incorporated in a final polypeptide fragment or protein.

The architecture of the PSDC was based on Meldal and co-worker’s research. As described earlier (Chapter 2, Section 2.2), the authors introduced glycopeptides as mimetics of complex oligosaccharide structures (a branched mannose pentasaccharide containing two terminal Man-6-P units). The idea was that the carbohydrate introduced as a building block could provide the specificity of the binding by directing the ligand to the oligosaccharide binding site (mannose 6-phosphate receptors), while the peptide (L-Thr-L-Lys-L-Thr) would function as a scaffold for optimal orientation of the glycan portion. The principle of glycopeptides mimicking oligosaccharides was consolidated through binding studies with an array of phosphorylated glycopeptides and the divalent mannose 6-phosphate receptor.

With the PSDC it is possible to produce complex unnatural carbohydrate structures, avoiding long and difficult synthesis, bypassing the need of regio- and stereoselective control of each glycosylation (Chapter 2 and 3). In the case of N-linked glycoproteins, the requirement of performing β-mannosylation greatly increased the complexity of the oligosaccharide synthesis (Chapter 2, Section 2.1.3).

Therefore, the aim was to generate a PSDC, as a mimic of the N-linked diantennary nonasaccharide found in EPO (Chapter 1, Scheme 1.4). This structure could further be introduced into an EPO peptide fragment, applying the technology that was developed for producing neoglycopeptides (Chapter 5). The synthesis of the scaffold should be
4. Synthesis of a novel unnatural peptide scaffold to display carbohydrates straightforward employing available building blocks. Finally, incorporation of the final structure on the EPO peptide fragment should also be readily available.

4.3 Design and retro-synthetic analysis of the PSDC

As discussed earlier, the PSDC is inspired on Meldal research work (Scheme 2.8). Based on the Macmillan research group experience, it was decided to develop the peptide structure and the introduction of the carbohydrates on solid phase. The solid-phase synthesis of the modified peptide could provide easier access to new levels of complexity. This approach avoids purification of intermediates and allows libraries to be constructed from varied amino acids and oligosaccharide building blocks. Finally, this structure could be attached to a peptide or protein backbone through a monosaccharide (Scheme 4.4).

Scheme 4.4: Structure of diantennary nonasaccharide structure found in EPO and PSDC, 104.
4. Synthesis of a novel unnatural peptide scaffold to display carbohydrates

Based on Macmillan research\(^2\) for the introduction of carbohydrates on a solid support to a peptide, cysteine was chosen as the amino acid to form part of the peptide scaffold backbone. Moreover, based on Meldal's work, lysine was chosen to be the central amino acid for the peptide scaffold (Scheme 4.4). The primary amine of the lysine can be protected with a wide range of protecting groups. Interestingly, allyloxy carbonyl (Alloc\(^-\)) has been used in the past on solid support without damaging the integrity of either the peptide or the linker. This is accomplished with palladium tetraakis(triphenyl phosphine) (Pd(PPh\(_3\))\(_4\)) and a nucleophile.\(^{176-179}\) Using this protecting group, it may also be possible to generate a library of peptide scaffolds on the solid support, by changing the number of carbons of the amino acid, and consequently control the distance between the carbohydrates presented through the peptide scaffold and the peptide/protein backbone (Scheme 4.4).

At the final it appeared desirable to "seal" the tripeptide core scaffold with an amide bond at the C-terminus, and an acetyl group on the N-termini (Scheme 4.4), to make the molecule less reactive and resistant to carboxypeptidases, increasing the \textit{in vivo} half life.\(^{180}\)

To demonstrate that the pre-designed PSDC could mimic the diantennary polysaccharide of EPO, \textit{N}-acetyl lactosamine attached through cysteines, and \textit{N}-acetyl glucosamine (GlcNAc) attached through lysine were the chosen carbohydrate structures for the study. Therefore, a carboxymethyl chain should be inserted in the 4-hydroxyl group of GlcNAc (Scheme 4.4).

Finally, this PSDC would be attached to the peptide/protein backbone. The fact that either aspatilation,\(^{97}\) acetamide glycosylation\(^{169}\) or \textit{"click chemistry"}\(\textit{ glycosylation}\(^{71}\) require at some stage in the synthesis the presence of an azide group at the anomeric position, suggested the need to insert the mentioned functional group into GlcNAc.

Therefore, having determined the final architecture of the PSDC, the retro-synthetic analysis suggested a possible route for 103 or 104 (Scheme 4.5). The 2-acetamido-2-deoxy-3, 6-di-
4. Synthesis of a novel unnatural peptide scaffold to display carbohydrates

O-acetyl-4-O-carboxymethyl-β-D-glucopyranosyl azide 100 could couple with the lysine of the peptide on the solid support. The intermediate comes from glycosylation between LacNAc bromoacetamides 90, and the thiols of two cysteines of the tripeptide. Altering the order of the reactions could lead to formation of an epimeric mixture due to the incompatibility of the dithiothreitol in the presence of azide.

Scheme 4.5: retrosynthetic analysis of 103 and 104.

4.4 Synthesis of the glycobuilding blocks

4.4.1 Synthesis of 2, 3, 4, 6-tetra-O-acetyl-D-galactopyranosyl-(β-1, 4)-2-acetamido-2-deoxy-3, 6-di-O-acetyl-1-N-[1-(2-bromo)acetyl]β-D-glucopyranose building block, 90

For the synthesis of LacNAc bromoacetamide 90, the procedure developed by Bertozzi and co-workers was adopted. With some modifications, the synthesis was accomplished successfully in 6 steps (Scheme 4.6).
The synthesis of the bromoacetamide glycobuilding block began with the commercially available glucosamine hydrochloride $84$. This was pre-treated with a stoichiometric solution of NaOMe in dry methanol, filtered and treated with phthalic anhydride. Next and finally, the molecule was peracetylated with acetic anhydride in pyridine, to obtain the fully protected intermediate $85$ in 30% yield. The generation of the beta product was demonstrated by $^1$H-NMR.

Scheme 4.6: Synthesis of 90 and 91.

The introduction of an azide group at the anomeric position was performed following the procedure of Szilagyi and co-workers. The protected monosaccharide $85$ was treated with...
TMS-\( \text{N}_3 \) in the presence of SnCl\(_4\) as a Lewis acid, to obtain, exclusively, the beta product \( \text{86} \) in a yield of 73 %.

In the subsequent sequence of reactions, the protected azido monosaccharide \( \text{86} \) was treated first with catalytic amounts of sodium methoxide in dry methanol, to perform a Zemplén deprotection (Scheme 4.6).\(^{66}\) This was followed by a regiospecific protection of the 6-hydroxyl group of the sugar with tert-butyl diphenyilsilyl chloride, to obtain the desired product \( \text{87} \) with the 3 and 4 hydroxyl groups free, in 93 % yield.

Following the procedure of Bertozzi and co-workers,\(^{181}\) the protected glycosyl acceptor \( \text{87} \) reacted with 2,3,4,6-tetra-O-acetyl-\( \alpha \)-D-galactopyranosyl fluoride \( \text{92} \),\(^{134}\) silver triflate, tin (II) dichloride and collidine in a mixture of DCM/toluene (5:1). Unfortunately, a yield of only 40 % was obtained rather than the claimed 84 %. Moreover, when the reaction was scaled up (3.00 g, 5.24 mmol), a 33 % yield of the desired product was obtained, and 29 % of a disaccharide lacking the silyl protecting group. The regioselectivity of the glycosylation was not determined.

Having available 2,3,4,6-tetra-O-acetyl-\( \alpha \)-D-O-galactopyranosyl trichloroacetimidate \( \text{23} \), the same reaction conditions were applied for the synthesis of core-1 disaccharide \( \text{62} \). One equivalent of the protected glycosyl acceptor \( \text{87} \) reacted with two equivalents of peracetylated galactose \( \text{23} \) with catalytic amounts of TMS-OTf, in DCM, to obtain a 92 % yield of the required disaccharide \( \text{88} \) (Scheme 4.6). The regioselectivity of the glycosidation was determined to be as desired by comparison with the spectroscopic data provided by Bertozzi and co-workers.\(^{181}\) The regioselectivity of the glycosylation could be due to steric factors, namely the bulk of the silyl and phthalic protecting groups directing the donor towards the 4-hydroxyl group of the acceptor molecule.
Having generated the disaccharide 88, a sequence of reactions was performed to obtain the peracetylated N-acetyl lactosamine 89. Silyl deprotection in the presence of TBAF and acetic acid in THF was followed by phthalamide deprotection. Per-O-acetylation of the disaccharide was achieved by treatment with a mixture of acetic anhydride and pyridine (1:2).

Finally, the azide group of the peracetylated disaccharide was reduced by catalytic hydrogenation. The desired glycobuilding block 90 was isolated in a yield of 95 %, following treatment with bromoacetic anhydride (Scheme 4.6). In parallel, the iodoacetamide derivative 91 was also produced in 32 % yield, by treating the reduced intermediate with 2-idoacetyl chloride.

Efforts were then directed towards the glycobuilding block 100, designed to couple with the lysine residue (Scheme 4.7).

4.4.2 Synthesis of 2-acetamido-2-deoxy-3,6-di-O-acetyl-4-O-carboxymethyl-β-D-glucopyranosyl azide, 100

The synthesis of glycobuilding block 100 was started from commercially available N-acetyl-D-glucosamine, 93 (Scheme 4.7). In a domino reaction, the molecule was peracetylated and a chloride was introduced at the anomeric position, to obtain the desired thermodynamic molecule 94, in 34 % yield.

As discussed earlier in this chapter, the insertion of an azide group at the anomeric position was considered critical to access the desired peptide scaffold, 103 or 104 (Scheme 4.5), either by acetamide linkage or "click chemistry" glycosylation. Therefore, this functional group was introduced by displacement of the chloride with sodium azide in the presence of tetrabutylammonium hydrogensulphate (TBAHS), to obtain the adduct 95, in 82 % yield. Afterwards, in a sequence of reactions, Zemplén deprotection was performed.
followed by protection of the hydroxyls in the position 4 and 6 with p-anisaldehyde dimethylacetal, and final acetylation of the free 3-hydroxyl group, furnished the intermediate 96 in 71 % yield. Finally, regioselective acetal ring opening with Na(CN)BH₃ and TFA in DMF, generated the desired monosaccharide 97 in 66 % yield, with the 4-hydroxyl group free (Scheme 4.7).

Scheme 4.7: Synthesis of 2-acetamido-2-deoxy-3,6-di-O-acetyl-4-O-carboxymethyl-β-D-glucopyranosyl azide 100.

The preparation of the tert-butyl acetate ether of the 4-hydroxyl group of the monosaccharide 97 was not a straightforward procedure (Scheme 4.7). Model experiments were carried out with derivatives of 87 (Scheme 4.6), where the 3-hydroxyl was acetylated: 101 (Scheme 4.8). Treatment of the monosaccharide 101 with tert-butylbromo acetate in the presence of
sodium hydride failed to generate the desired product. Instead, a complex mixture of undefined degradation products was obtained (Table 4.1). Addition of the catalyst TBAI did not improve results. It was suspected that deacetylation reactions were occurring under such harsh basic conditions. Consequently, milder conditions were required for the success of the reaction.

**Scheme 4.8:** Model experiments for the introduction to the 4-hydroxyl position of tert-butylbromoacetate.

Applying the protocol of Giese and co-workers, the monosaccharide 101 was treated with 5 equivalents of tert-butylbromoacetate and silver (I) oxide in DMF. Unfortunately, no reaction was observed after 24 hours and heating to 50 °C had no effect. Surprisingly, after the addition of pyridine 28 % of the desired product 102 was obtained. Similar results were obtained when the order of addition of reagents was switched: no reaction was observed when the monosaccharide 101 was treated with 5 equivalents of tert-butylbromoacetate, and 10 equivalents of pyridine; only when 2.5 equivalents of silver oxide were added a 46 % yield of the desired product 102 was achieved (Scheme 4.8). Optimisation of the reaction conditions afforded a 59 % yield when pyridine was replaced with triethylamine in the presence of silver oxide.

Application of the latter conditions to the intermediate 97 generated modest results. Further optimisation led to 57% yield, when the monosaccharide 97 reacted in the presence of 3.6
4. Synthesis of a novel unnatural peptide scaffold to display carbohydrates equivalents of tert-butylbromo acetate, 1.8 equivalents of silver (I) oxide, and 2 equivalents of triethyl amine, in dry DMF with exclusion of light for 16 hours (Scheme 4.7).

<table>
<thead>
<tr>
<th>Monosaccharide</th>
<th>tert-butylbromo acetate</th>
<th>Ag₂O</th>
<th>Base</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>101</td>
<td>2 equiv.</td>
<td>-</td>
<td>2 equiv. NaH</td>
<td>Decomposition</td>
</tr>
<tr>
<td>101</td>
<td>2 equiv.</td>
<td>-</td>
<td>2 equiv. NaH and 2 equiv. TBAI</td>
<td>Decomposition</td>
</tr>
<tr>
<td>101</td>
<td>2 equiv.</td>
<td>-</td>
<td>2 equiv. NaH and 0.1 equiv. TBAI</td>
<td>Decomposition</td>
</tr>
<tr>
<td>101</td>
<td>5 equiv.</td>
<td>1.5 equiv.</td>
<td>-</td>
<td>No reaction</td>
</tr>
<tr>
<td>101</td>
<td>5 equiv.</td>
<td>1.5 equiv.</td>
<td>3 equiv. pyridine</td>
<td>28 %</td>
</tr>
<tr>
<td>101</td>
<td>5 equiv.</td>
<td>-</td>
<td>10 equiv. pyridine</td>
<td>No reaction</td>
</tr>
<tr>
<td>101</td>
<td>5 equiv.</td>
<td>2.5 equiv.</td>
<td>10 equiv. pyridine</td>
<td>46 %</td>
</tr>
<tr>
<td>101</td>
<td>5 equiv.</td>
<td>2.5 equiv.</td>
<td>1.2 equiv. DIPEA</td>
<td>24 %</td>
</tr>
<tr>
<td>101</td>
<td>2 equiv.</td>
<td>1.2 equiv.</td>
<td>1.5 equiv. NEt₃</td>
<td>18 %</td>
</tr>
<tr>
<td>101</td>
<td>5 equiv.</td>
<td>2.5 equiv.</td>
<td>1.2 equiv. NEt₃</td>
<td>49 %</td>
</tr>
<tr>
<td>101</td>
<td>5 equiv.</td>
<td>2.5 equiv.</td>
<td>1.5 equiv. NEt₃</td>
<td>59 %</td>
</tr>
<tr>
<td>97</td>
<td>5 equiv.</td>
<td>2.5 equiv.</td>
<td>1.5 equiv. NEt₃</td>
<td>42 %</td>
</tr>
<tr>
<td>97</td>
<td>5 equiv.</td>
<td>2.5 equiv.</td>
<td>2 equiv. NEt₃</td>
<td>49 %</td>
</tr>
<tr>
<td>97</td>
<td>3.6 equiv.</td>
<td>1.8 equiv.</td>
<td>2 equiv. NEt₃</td>
<td>57 %</td>
</tr>
</tbody>
</table>

Table 4.1: Different sets of conditions for the formation of tert-butyl acetate ether with the 4-hydroxyl group of the monosaccharide.

During the development of this reaction, it became clear that the reaction did not proceed without the presence of silver oxide and a low pKa base (pyridine or triethylamine). The coupling of a hard, hindered nucleophile, such as a secondary hydroxyl group, with a soft
4. Synthesis of a novel unnatural peptide scaffold to display carbohydrates
electrophile, might be expected to be challenging, however, the requirement for silver (I)
oxide was unexpected.

Having successfully obtained 98, PMB deprotection was carried out under standard
oxidative conditions. This was followed immediately by peracetylation with acetic
anhydride to produce the intermediate 99 in quantitate yield. Tert-butyl ester hydrolysis
yielded the glycobuilding block 100, ready for SPPS (Scheme 4.7).

Once the different glycobuilding blocks: 90 and 100 were available, it was decided to
introduce them to the solid support, by simple coupling with unprotected cysteine or lysine,
respectively.

4.5 Synthesis of a peptide scaffold to display carbohydrates on solid phase

As stated previously (Section 4.3), the assembly of the peptide core scaffold was to be
carried out on a solid support, using SPPS. The introduction of the glycobuilding blocks 90
and 100 was planned after orthogonal deprotection of the relevant cysteine and lysine
protecting groups.

Rink amide MBHA resin was selected as a solid support for the synthesis.5 When the
peptide is cleaved from the resin under acid conditions it produces the desired amide bond at
the C-termini. The synthesis of the peptide core scaffold was straightforward, using 20 %
piperidine in DMF for Fmoc- deprotection, and standard peptide coupling conditions
(HBTU/HOBt and DIPEA for three hours) to couple the next amino acid. The sequence of
amino acids introduced to the synthesis of the peptide scaffold core was: L-Cys-L-Lys-L-
Cys. The thiols of the cysteines were protected with tert-butyldithiol groups,92 and the
primary amine of the lysine was masked with an Alloc- protecting group (Scheme 4.9).
Finally, after the coupling of the last amino acid, cysteine, *Fmoc*- deprotection was carried out, followed by acetylation of the amino group with acetic anhydride and pyridine (1:2). After each step, small aliquots of cleaved resin were qualitative monitored by MS and Kaiser test (Scheme 4.10).

**Scheme 4.9**: Synthesis of the peptide core structure on solid support: i) *tert*-butylthiol deprotection with DTT and ii) introduction of LacNAc bromoacetamide, 90, into the peptide scaffold.

From previous work, it was known that orthogonal deprotection of the disulphides of a peptide on solid support could be achieved by treatment of the resin with dithiothreitol (DTT). Therefore, treatment of the core peptide structure with DTT in dry DMF, and 2.5% v/v DIPEA, for 16 hours, generated the desired intermediate, as verified by MS (Scheme 4.10.i). Exhaustive washings of the resin with DMF, mixtures of DMF/H_{2}O (1:1), DMF and DCM were performed to remove any residual DTT that could interfere in the following steps.
4. Synthesis of a novel unnatural peptide scaffold to display carbohydrates

Scheme 4.10: Mass spectra of cleaved samples during the synthesis of the peptide scaffold after: i) tert-butyldisulphide deprotection, ii) coupling with 90, and iii) Alloc-deprotection.
Treatment of the resulting resin, twice, with 3 equivalents of LacNAc iodoacetamide 91 per thiol (section 4.4.1), 2.5 % v/v pyridine in DMF, failed to achieve complete coupling between the peptide on the resin and the disaccharide. However, when the resin was treated with 2 equivalents of LacNAc bromoacetamide 90 per thiol, and an excess of triethylamine, in dry DMF for 24 hours (Scheme 4.9), MS analysis of the cleaved PSDC demonstrated the acetylation of the peptide cysteines and no presence of the starting material (Scheme 4.10.ii).

Once LacNAc bromoacetamide 90 was mounted on to the peptide scaffold, the next step was Alloc-deprotection. A number of examples of this reaction on solid support had been reported in the literature. Common to each was the need for palladium (0), and an allyl scavenger, such as PhSiH₃ or N-methylmorpholine. The procedure of Albericio and co-workers was applied wherein palladium tetrakisphenyolphosphine (Pd(PPh₃)₄) was the source of palladium (0), and phenylsilane, the allyl scavenger. A ninhydrin test and MS analysis confirmed the deprotection of the primary amine of lysine (Scheme 4.10.iii).

The glycobuilding block 100 reacted specifically with the free lysine of the peptide scaffold, under standard coupling conditions, in the presence of 5 equivalents of DCC and HOBt in DMF for 24 hours. Although this procedure was performed twice, the ninhydrin test and MS analysis revealed the presence of uncoupled peptide scaffold.

Finally, the resin was treated with a cleavage cocktail (95 % TFA, 2.5 % ethane dithiol, 2.5 % water) to liberate the peptide from the solid support. After precipitating the peptide with cold ether, lyophilising the precipitate, and purifying it by semi-preparative HPLC in reverse phase, a 5 % of yield of 103 was obtained over 10 steps, in the best of the cases. Due to this adverse result obtained, it was studied where the peptide scaffold has gone. The resin was treated a further two times with the cleavage cocktail to isolate additional material, unfortunately not enough material was obtained to address this question.
Based upon previous experience of both the SPPS and bromoacetamide methodologies, it appeared unlikely that either of these steps had been particularly problematic. It was decided to review the Alloc-deprotection protocols. Re-synthesis of the PSDC on a 0.25 mmol scale revealed that the resin began to shrink unexpectedly upon lysine deprotection, and seemed likely that by changing this conditions this effect might be avoided.

Applying Griffin and co-workers conditions, the resin was subjected to 3 equivalents of Pd(PPh₃)₄, in chloroform, N-methylmorpholine and acetic acid (37.5:2:1), for 4 and then 12 hours. Despite repeated washings of the resin with DCM and DMF after the reaction, the solid support was still stained black, due to the persistent presence of palladium. After coupling the glycobuilding block and the peptide cleavage, was obtained in a yield of 8%, over 10 steps. At this stage it was clear that the limiting step was the Alloc-deprotection. It was noted that the use of N-methylmorpholine in place of phenylsilane prevented loss of material from the resin. Finally, applying the procedure of Tabor and co-workers, the resin was treated with 2 equivalents of Pd(PPh₃)₄ in a specific solvent system: DMF/CHCl₃/AcOH/NMM (18.5:18.5:2:1). It was shaken for 2 hours with exclusion of air and light. Finally, the resin was washed sequentially with: i) 0.5% DIPEA in DMF (v/v); ii) DMF; iii) 0.5% sodium diethyldithiocarbamate trihydrate in DMF (w/v), and iv) DMF (Scheme 4.11). Applying this sequential washing procedure, most of the residual palladium was eliminated from the resin. The natural colour and the state of the resin were re-established. Moreover, MS analysis of the cleaved peptide (Scheme 4.10) and a ninhydrin test indicated deprotection of the lysine and no presence of unreacted glycopeptide. Finally, after coupling 5 equivalents of the glycobuilding block with the resin and peptide cleavage, was generated in a 17% yield over 10 steps.

During development of these sets of reactions, a number of additional peaks were noted in the HPLC trace. These peaks exhibited retention times similar to that of the peracetylated
PSDC **103** \( \left( t_R = 30.3 \text{ minutes} \right) \). Closer inspection of the relevant fractions by MS revealed a difference in molecular mass of “42 x n” less than **103**. This range of compounds could be related to deacetylations of the carbohydrate produced during the course of the synthesis (molecular weight of the fragment \( \text{CH}_3\text{CO}= 42 \)). Therefore, it was decided to treat the resin bound mixture with a solution of pyridine and acetic anhydride \( (2:1) \) for 16 hours to re-acetylate the deacetylated hydroxyl groups. After exhaustive washing of the polymer with DMF and DCM, the yield of synthesis increased to 30% over 10 steps (Scheme 4.11).

Scheme 4.11: (Continued from scheme 4.9) Synthesis of the peptide core structure on solid support: i) Alloc- deprotection, ii) coupling of **100** with the peptide scaffold and iii) peracetylation of the molecule, and iv) final cleavage of the PSDC, **103**, from the resin.

Peracetylation of the whole molecule with pyridine and acetic anhydride \( (2:1) \) allowed identification of further by-products of the synthesis of the peptide scaffold: the Alloc-protected lysine adduct \( \left[ M \right]^+; 1826.6, \left[ M \right]^{2+}; 913.8 \); and the peracetylated lysine adduct \( \left[ M \right]^+; 1784.60, \left[ M \right]^{2+}; 892.8 \), that corresponded to the uncoupled peptide scaffold with the glycobuilding block **100** (Scheme 4.12).
4. Synthesis of a novel unnatural peptide scaffold to display carbohydrates

Scheme 4.12: HPLC trace of the crude of the peptide scaffold displaying carbohydrates. Mass spectra of 103, MW: 2113.40 (MW_{calc}: 2113.05); and related by-products: MW: 1785.00 and 1825.20 (MW_{calc}: 1783.60 and 1825.61).
These results are in accordance with previous highlighted observations. Therefore double Alloc- deprotection and more reactive coupling conditions between the peptide scaffold and the glycobuilding block 100 should increase the overall yield of the total synthesis on solid phase. In the process, a better HPLC chromatogram would be obtained, and consequently, the yield of the desired product would be improved.

Therefore, the Alloc- deprotection was performed twice, followed by coupling of the peptide scaffold with 2.5 equivalents of the glycobuilding block 100, 2.5 equivalents PyBOP, 2.5 equivalents HOBt, and 5 equivalents DIPEA in DMF for 16 hours, twice, to ensure total conversion of in the reaction. Final deacetylation of the product cleaved from the resin with a water/methanol solution (10:1) of 5 % hydrazine (6 mL), yielded the desired product in 47 %, over 10 steps (Scheme 4.13).

Scheme 4.13: (Continued from scheme 4.9) Synthesis of the peptide core structure on solid support: i) Alloc- deprotection, ii) coupling of 100 with the peptide scaffold, iii) cleavage of the PSDC from the resin and iv) final deacetylation of the molecule.
Elution of the deacetylated product 104 was more rapid than the peracetylated PSDC ($t_R = 14.2$ minutes), but did not add any further complications to the purification process (Scheme 4.14).

Scheme 4.14: HPLC trace of the crude of the peptide scaffold displaying carbohydrates. Mass spectra of 104, MW: 1525.00 (MW: 1524.54).

With both peptide scaffolds available, 103 and 104, they could be inserted in a peptide, either by Lansbury aspartilation\(^97\) or bromoacetamides after reduction of the azide. Unfortunately, in both cases, the reduction of the azide could produce epimerisation in the anomeric centre. On the other hand, Cu(I)-catalysed modification of the Huisgen cycloaddition could be applied to directly introduce 103 or 104 to a modified peptide fragment without any extra synthetic step.\(^{171}\) The development of this novel neo-glycopeptide linkage is explained in the following chapter.
4.6 Conclusions and future work

During this work, a peptide scaffold bearing carbohydrates on solid phase that could potentially mimic the diantennary carbohydrate structure found in N-linked glycoproteins has been synthesised. The PSDC allows the construction of highly complex unnatural carbohydrate structures from pre-defined glycobuilding blocks, avoiding control of the regio- and stereochemistry, and therefore simplifies the total synthesis of the desired molecule (Chapter 3). This concept has been exemplified by avoiding the complication of the β-
mannosylation.

In addition, having available this new methodology to obtain PSDC on solid phase, a library of these molecules could be prepared by increasing the number of cysteines, and therefore changing the number of carbohydrate side chains displayed. Moreover, the lysine residue could be altered by changing the number of carbons of the alkylc side chain. This could provide the opportunity to identify peptide scaffolds showing better biological activity.

One of the main problems found during the construction of core-1 O-linked to threonine glycobuilding block for SPPS, lay in the glycosylation of 46 with the glycosyl bromide 69 (Scheme 3.23). This process proved to be low yielding, and encouraged investigation of new pathways to introduce, in a convergent synthesis, carbohydrates on peptides. Cu(I)-catalysed modification of Huisgen cycloaddition seemed to fill the requirements of orthogonal reactivity with control of regioselectivity and the development of the synthesis of a novel neo-glycopeptide linkage using “click chemistry”, described in the next chapter.
5 Novel neo-glycopeptide linkage using Cu(I) catalysed 1, 2, 3-triazole formation

5.1 Introduction

As discussed earlier, protein glycosylation can be incredibly challenging and difficult. Parallel to the research and development of advanced methodologies to overcome these technical difficulties of glycosylation, there is increasing interest in the development of novel unnatural glycosidic bonds. The primary goals of these type of linkages are to perform chemoselective and site-specific glycosylations for the synthesis of glycoprotein analogues.

Various examples of non-native glycosidic bonds are described in the literature (Scheme 5.1), with chemoselective ligation, acetamides and disulphide bond linkages being the most relevant examples to this thesis. These linkages aim to meet the need for regio- and glycan-specific glycosylation on proteins. Chemoselective ligation is based on the introduction of two mutually and uniquely reactive functional groups onto unprotected fragments (ketone or aldehyde, and aminoxy sugars), and the convergent coupling of the fragments for the construction of glycopeptide mimetics. The acetamide linkage approach and disulphide bond linkage methodologies are based on the selective linkage of the carbohydrate to the peptide/protein backbone, through a free cysteine group, either through a bromo/iodo acetamide or a thiosugar, respectively (Scheme 5.1).
The Macmillan research group has applied the “site directed mutagenesis glycosylation approach” to *E. coli*-expressed rhEPO protein,91 and on SPPS for the synthesis of unnatural glycopeptides.92 The native glycosylated asparagine residues were changed for cysteines in three mutant rhEPO proteins (His10-Asn24Cys, His10-Asn38Cys, His10-Asn83Cys). A single unit of GlcNAc iodoacetamide was introduced selectively through the thiol of a cysteine without any observable side reaction with any of the other four cysteine residues involved in disulphide bonds (C29-C33 and C7-C161). In a later example, iodoacetamide carbohydrate derivatives (GlcNAc, LacNAc, Glc, Lac, and chitobiose) were selectively introduced into a peptide on solid phase, through a selectively deprotected thiol of a cysteine (Scheme 5.2). Moreover, it was also demonstrated that this technology was compatible with the generation of thioesters using the “safety-catch” linkers, Native Chemical Ligation (NCL) and Expressed Protein Ligation (EPL).
Scheme 5.2: Synthesis of a novel glycopeptide produced by: i) orthogonal deprotection of tert-butyl dithiol, followed by ii) introduction of carbohydrate derivatives on cysteines, iii) activation of “safety catch” resin and production of a thioester, and finally iv) trityl deprotection of the N-terminal cysteine.

Although this type of technology facilitates the chemoselective and site-specific ligation of carbohydrates on a peptide or protein, it is not exempt from extra modifications of the reactive functional groups and it employs unstable glycosamines, which are prone to hydrolysis and anomeration.

“Click chemistry” has emerged as a potent chemoselective and regioselective reaction that facilitates the chemistry of carbon-heteroatom bond formation, allowing easier and faster access to a generation of new compounds. Consequently, it could have huge potential for the synthesis of glycopeptides and glycoproteins.
5.2 "Click chemistry"

The term "click chemistry" was introduced by Barry K. Sharpless and co-workers to refer to a group of reactions that must be modular, of wide scope, produce high yields, generate only inoffensive by-products that may be removed by non-chromatographic, and be stereospecific (but not necessarily enantioselective). Required process characteristics include simple reaction conditions (the process should be insensitive to oxygen and water), readily available starting materials and reagents, the use of no solvent or solvent that is benign (such as water) or easily removed, and simple product isolation. Purification, if required, must be by non-chromatographic methods, such as crystallization or distillation, and the product must be stable under physiological conditions.\(^9\)

Nevertheless, in chemistry jargon, "click chemistry" normally refers to the modification of the Huisgen 1,3-dipolar cycloaddition between an alkyne and azide, in the presence of copper (I) (Cu(I)) as a metal catalyst. This reaction was modified simultaneously and independently by the research groups of Meldal\(^{10}\) and Sharpless\(^{11}\) to ligate two types of modular molecules: peptides and alkyl building blocks. In a postulated mechanism (Scheme 5.3),\(^{12}\) the catalytic cycle begins with the metal insertion of Cu(I) into the alkyne. This slow process can occur at pH values ranging from 4 to 12.\(^{13}\) They subsequently obtained the thermally and hydrolytically stable 1, 4-disubstituted 1, 2, 3-triazole with a complete regioselectivity,\(^{14}\) in an irreversible fast stepwise sequence via a six-membered copper-containing ring intermediate.

One of the major drawbacks of using Cu(I) as a catalyst is its low thermodynamic stability which results in its oxidation to Cu(II) and/or disproportionation to Cu(0) and Cu(II), compromising the scope of conditions for this type of reaction.\(^{15}\) A breakthrough in overcoming this problem came with the introduction of the ligand tris-(benzyltriazolylmethyl)amine (TBTA). This molecule has the ability to envelope the Cu(I)
centre tightly, providing additional electron density through the tertiary amine, so that the ligand protects the metal from water and oxygen, and accelerates Cu(I)-catalysed transformations.\textsuperscript{196}

The orthogonal reaction between an azide and an alkyne has found a wide range of uses in bioconjugation,\textsuperscript{197-201} organic synthesis,\textsuperscript{202,203} combinatorial chemistry,\textsuperscript{204} and carbohydrate chemistry.\textsuperscript{205-210}

Although most of the applications of "click chemistry" in the field of carbohydrate research have been to study the reactivity and synthesis of large building block; some authors have envisioned the potential to bypass some of the biggest problems this area of research presents: the stereoselectivity of glycosidation of unprotected sugars,\textsuperscript{211} and direct glycosylation of amino acids.\textsuperscript{211,212}

5.3 Objective: Development of a novel neo-glycopeptide linkage compatible with native chemical ligation

Based on the potential of Cu(I) catalysed 1,2,3-triazole formation, and previous literature precedent of this type of reaction with carbohydrates, the possibility to introduce azido
carbohydrates directly on n acetylene modified peptide, either on solid phase or in solution can be envisaged. To achieve this goal, preliminary studies in solution between simple structures should be carried out to understand the outcome of these reactions. Afterwards, the knowledge generated could be used directly to modify a peptide. This new technology should facilitate and accelerate the production of glycopeptide mimetics.

For example, having available this novel neo-glycopeptide linkage, it could be applied to introduce two PSDC (Chapter 4) on an EPO peptide fragment, avoiding extra synthetic steps, and consequently, increasing the overall yield of the total synthesis of a glycopeptide mimetic.

5.3.1 Preliminary studies in solution: model experiments

In preliminary work, the type of linker to use for the introduction of carbohydrates on peptides and the best conditions suitable for the “click chemistry” reaction were investigated. The reaction was performed between an azido sugar 86, and an alkyne, N-(propargyl) acetamide 111 with catalytic amounts of copper sulphate pentahydrate as a source of Cu(II), with the same catalytic amounts of sodium ascorbate, to reduce Cu(II) to Cu(I); in a biphasic system: CHCl₃/EtOH/H₂O (9:1:1) (Scheme 5.4). Interestingly, this reaction was so slow that after 48 hours at 37 °C that the starting materials was recovered. A similar result was obtained under the same reaction conditions for 95. Wang and co-workers reported a one-pot synthesis of triazole-linked glycoconjugates to develop neo-glycoconjugates derived from unprotected and peracetylated azido oligosaccharides. Interestingly, the authors had reacted peracetylated D-mannose with N-(propargyl) bromoacetamide 112, a compound that was generated in 76 % yield, from the reaction between propargyl amine 115 and bromoacetic anhydride 116 in a saturate aqueous solution of NaHCO₃ (experimental section). Using N-(propargyl) bromoacetamide 112, and applying identical reaction
conditions, satisfactory results were obtained in preparing 114 and 113 in 66% and 100% yield, respectively (Scheme 5.4).

The previous unsuccessful attempts for the formation of the triazole ring appeared to be in accordance with observations that Thorson and Bertozzi reported, where they noted that the presence of an electro-withdrawing group proximal to the alkyne accelerates the rate of the cycloaddition. In the present case, the presence of a halogen, namely bromide, δ to the alkyne 112, seems to have a decisive effect on the kinetics of the reaction. This acceleration of the process could be due to an electronic effect of the electronegative atom on the alkyne facilitating the limiting step of the “click chemistry”, the metal insertion of the copper into the triple bond.

Scheme 5.4: “Click chemistry” between azido sugars and propargyl amide derivatives: CuSO₄·5H₂O, sodium ascorbate, CHCl₃/EtOH/H₂O (9:1:1).

To understand the chemistry that was going to be developed later for solid phase, model experiments in solution phase were necessary. 2-Acetamido-2-deoxy-3,4,6-tri-O-acetyl-β-D-glucopyranosyl azide 95 was used as the source of the azido sugar, N-(propargyl)-bromoacetamide 112 was the source of the alkyne, and benzylmercaptan 117 was used as the model thiol (in place of cysteine). Since the alkyne motif 112 could be used as an acceptor for “click chemistry” and S₂2 reactions, it was interesting to investigate the versatility of this methodology and, in particular, whether the order of the reactions alter the outcome (Scheme 5.5).
Given previous work on “click chemistry” with azido sugars and N-(propargyl)-bromoacetamide 112, the conditions described earlier were applied to reproduce the synthesis of the compound 114 (Scheme 5.5). Based on the experience with bromoacetamides, 114 was reacted with benzylmercaptan 117 and triethylamine in DMF at room temperature for 16 hours, to obtain the thioether product 118 in 75 % yield. Given these encouraging results, the versatility of this method was investigated, where an SN2 reaction between 112 and benzylmercaptan 117 was performed first, and then the product reacted with 95 to obtain 118. This could potentially be of great benefit for the synthesis of glycopeptides and glycoproteins because the whole synthetic process would be convergent, that it is by first introducing the acetylenes on a peptide/protein and subsequently performing the “click chemistry” reaction with the azido sugars; it would increase the overall yield of these structures and reduce amounts of starting materials required. When 112 reacted with benzylmercaptan 117 with triethylamine in DMF at room temperature for 16 hours in an SN2 fashion, 119 was obtained in 84 % yield. Subsequent application of triazole formation between 119 and 95, under identical conditions as described previously, allowed the formation of 118 in a 91 % yield.

To establish whether the products might be stable to the usual peptide cleavage conditions, 118 was treated with 95 % TFA, 2.5 % H2O and 2.5 % 1,2-ethanedithiol for 3 hours. 1H-NMR analysis of the crude material after evaporation revealed no decomposition of the starting material. This result confirmed the triazole structure is stable under the peptide cleavage conditions.

Finally, the acetyl esters of 118 were cleanly removed upon exposure to 2 % v/v hydrazine hydrate in ethanol for 72 hours, and the fully deprotected compound 120 was obtained in 66 % yield (Scheme 5.5). The low yield and the long reaction time for the reaction was related to the insolubility of the starting material under the reaction conditions. Alternative
conditions could be used to accelerate and increase the reaction yield, but the aim of these experiments was to demonstrate the compatibility of these reactions that were subsequently going to be applied to the synthesis of glycopeptides. Encouraged by the ability to demonstrate the facility and the flexibility of these sets of reactions, it was decided to apply this methodology to solid phase.

**Scheme 5.5:** Model reactions between 95, 112 and benzylmercaptan 117, to demonstrate that the order in which the reactions took place had no effect upon the resulting product.

Parallel to these studies, it was decided to investigate if unprotected azido carbohydrates could react with acetylenes in these systems. This could have a huge effect in present research because they could open up the possibility of performing a final direct orthogonal coupling between peptides or proteins furnished with acetylenes and azido polysaccharides, in water, or even in buffer solutions.
5. Novel neo-glycopeptide linkage using Cu(I) catalysed 1, 2, 3-triazole formation

Preliminary studies were performed using 121 and 112. When these compounds reacted in the presence of catalytic amounts of copper sulphate pentahydrate and sodium ascorbate in water/tert-butanol (1:1), the desired product 122 was obtained in 35 % yield (Scheme 5.6). Due to the insolubility of 112 in water, it was necessary to use a binary solvent system to solubilise all the components of the reaction. Significant progress was made when the ligand tris-(benzytriazolylmethyl)amine (TBTA) was introduced into the reaction. Under the same set of conditions, the desired product 122 was obtained in 71 % yield (Scheme 5.6). However, by changing the source of copper to (Cu(CNCH₃)PF₆) and carrying out the reaction in water/tert-butanol (1:2), resulted in a decrease in yield (64 %).

Scheme 5.6: Model reactions of “click chemistry”, between an azido sugar and acetylene.

These encouraging results could allow the possibility to perform couplings between different building blocks, either in water or in buffer solutions.

5.3.2 Studies on solid phase: novel neo-glycopeptide linkage using “click chemistry”

The technology was applied to solid phase (experiments carried out by Dr. D. Macmillan). The peptide fragment 123, similar in sequence to human EPO (residues 21-32) was assembled with an N-terminal cysteine residue, and furnished with two disulphide bond protected cysteine residues at pre-determined positions. The peptide was assembled using
standard protocols for Fmoc- solid phase peptide synthesis and in an automated fashion. The cysteine residues were deprotected on solid phase by exposure to 10 % w/v dithiothreitol (DTT) containing 2.5 % v/v DIPEA to expose the thiol functional groups (Scheme 5.7).92

N-Acetylglucosamine and the disaccharide chitobiose were then incorporated by exposure of the resin to bromoacetamides 114 or 124, employing three equivalents of the modified sugars per thiol in each reaction. After 16 hours reaction at room temperature, cleavage of a small resin sample indicated that the reaction was complete as it was unable to detect any starting material by MS. After cleavage from the solid support by treatment of the resin with peptide cleavage cocktail for 4 hours, the crude products were purified by semi-preparative HPLC, lyophilised, and treated with 2 % v/v aqueous hydrazine hydrate containing 5 % w/v DTT to obtain the fully deprotected products 125 and 126 in quantitative yield as, determined by HPLC (Scheme 5.7).

Scheme 5.7: Reaction and conditions for the synthesis of glycopeptides 125 and 126 (Results provided by Dr. D. Macmillan).
Applying the same methodology developed previously in solution, it was attempted to perform the “click chemistry” reaction on solid phase (Scheme 5.8). First, tert-butyl dithiol deprotection was accomplished with 10 % w/v DTT containing 2.5 % v/v DIPEA, to expose the thiol functional groups. Next, the free cysteines of the peptide were subjected to alkylation with 112 and triethylamine in DMF for 16 hours. Finally, the modified peptide on the resin 127 was resuspended in CHCl₃/EtOH/50mM sodium phosphate buffer (9:1:1), with the azido sugar 95 in the presence of catalytic amounts of sodium ascorbate and copper sulphate pentahydrate, and incubated at 37 °C at 500 rpm, for 24 hours, to obtain the desired neo-glycopeptide 128 as evidenced by HPLC-MS.

Scheme 5.8: Application of “click chemistry” on solid phase to obtain a novel neo-glycopeptide 128 (Results provided by Dr. D. Macmillan).

Once the reliability of formation of this neo-glycopeptide linkage on solid phase was established, it was decided to study the compatibility of this technology with native chemical ligation (NCL).
5.3.3 Native Chemical Ligation (NCL). Compatibility of the novel neo-glycopeptide linkage with NCL

NCL involves the coupling of two unprotected synthetic peptide fragments, where the first step is a reversible chemoselective reaction at physiological pH of an unprotected synthetic peptide-α-thioester in the C-terminal, while the other segment contains an N-terminal cysteine residue to give a thioester-linked intermediate as the initial covalent product. This initial transthioesterification reaction is then followed by a rapid intramolecular S→N acyl shift via a 5 membered ring transition state to generate an amide bond at the ligation junction (Scheme 5.9).²¹⁴,²¹⁵

\[
\text{Scheme 5.9: Scheme of Native Chemical Ligation.}
\]

This reaction has been widely used for ligating two synthetic peptide fragments together, where the necessary functionalities can be incorporated into the fragments during SPPS. Incorporation in the ligation reaction of guanidinium chloride or urea can avoid peptide aggregation. Addition of exogenous thiols can modulate the reaction and avoid oxidation of cysteines.
5. Novel neo-glycopeptide linkage using Cu(I) catalysed 1, 2, 3-triazole formation

It was aimed to demonstrate the compatibility of the novel neo-glycopeptide linkage using Cu(I) catalysed 1,2,3-triazole formation with NCL (experiments carried out by Dr. D. Macmillan). It was decided to couple the fragment 126 with a peptide thioester (EPO residues 1-19) (Scheme 5.10).214

![Scheme 5.10](image)

Scheme 5.10: The neo-glycopeptide is compatible with NCL (Results provided by Dr. D. Macmillan).

The construction of the peptide thioester, corresponding to human EPO residues 1-19, and its release from the solid support were monitored using the dual-linker approach described by Unverzagt and co-workers.100 In the ligation reaction equimolar quantities of each peptide were combined in 0.25 mL of 6 M guanidine hydrochloride containing 300 mM sodium phosphate buffer; pH 8, 0.1 % w/v mercaptoethanesulfonic acid (MESNA) and 10 mM tris-carboxyethylphosphine (TCEP) for 36 hours with shaking at room temperature. After this time the reaction mixture was purified by directly loading it onto a semi-preparative HPLC
column. The ligated product 129 was the only species observed by HPLC and certified by MS, demonstrating the robustness and versatility of this methodology.

Having demonstrated that the novel neo-glycopeptide linkage using Cu(I) catalysed 1,2,3-triazole formation is compatible with NCL, the possibility to use this methodology to introduce a PSDC on an EPO peptide fragment was investigated.

### 5.4 Introduction of two PSDC on an EPO peptide fragment applying the novel neo-glycopeptide linkage using Cu(I) catalysed 1,2,3 triazole formation.

The capacity of this methodology to introduce two PSDC units in an EPO peptide fragment was investigated. The peracetylated PSDC 103 reacted with 112, under the same conditions employed to couple azido carbohydrates with acetylenes (Scheme 5.5). The reaction proceeded without any complication and after 16 hours, the presence of 130 was detected by MS (Chapter 7, Experimental). Unfortunately, the work-up proved to be difficult and was complicated by the formation of an orange solid. As a result, the isolated yield of the desired product varied significantly. A possible explanation could derive from the insolvabilities of either the reactant 103 or the product 130. Due to the limited solubility and the limited availability of the PSDC, replacement of 103 with the deacetylated derivative 104 might provide a more efficient alternative synthesis, based in previous results obtained with 121, in a mixture of water and tert-butanol (Scheme 5.6).

Exposure of the compound 104 to N-propargylbromoacetamide 112, in the presence of catalytic amounts of sodium ascorbate, tris-benzyltriazolymethylamine (TBTA), and CuSO₄·5H₂O, in water/tert-butanol (1:1) generated the desired molecule 131 (MW: 1700.55), in 64%. Moreover, work-up was simple, and during the purification by HPLC, the product was readily isolated from the excess of acetylene 112 (Scheme 5.11).
5. Novel neo-glycopeptide linkage using Cu(I) catalysed 1, 2, 3-triazole formation

![Scheme 5.11: HPLC trace of the crude reaction between 104 and 112. Mass spectra data for 131, MW: 1701.40 (MW \text{calc}: 1700.55).]

Having obtained the PSDC with the bromoacetamide moiety 131, it was decided to introduce this structure to a peptide fragment of EPO applying the technology described above.\textsuperscript{92,171}

The modified PSDC 131 and an EPO peptide fragment (21-32) 132 with the two free thiols were reacted in a buffered solution of sodium phosphate, pH 7.4, to afford the desired unnatural glycopeptide 133 in 29 % yield.\textsuperscript{92,171} Quantitative deprotection of the Acm- group
was subsequently achieved after dissolving the modified peptide in acetic acid in the presence of mercury (II) acetate. To this was added dithiothreitol. MS analysis of the HPLC fractions determined the presence of the desired product 134 (Scheme 5.12).

Scheme 5.12: HPLC trace of the crude reaction after removing Acm-protecting groups from the EPO peptide fragment (21-32), 133, that displays two peptide scaffolds furnished with carbohydrates. Mass spectra data for 134, MW: 4401.20 (MW_{calc}: 4399.65).

Encouraged by the results obtained for attaching the peptide scaffold displaying carbohydrates to an EPO peptide fragment, it was decided to further investigate the technology developed in-house (Chapter 5). The possibility to modify peptides with N-(propargyl)-bromoacetamide 112 would allow the incorporation of complex carbohydrate structures to the core peptide, avoiding elaborated synthetic routes. The resulting increase in the efficiency of the procedure would minimise loss of precious material.
An initial proposal was to introduce the peptide scaffold to an EPO peptide on solid support through the thiols of two cysteines. Previous experience demonstrated that tert-butyl disulphide deprotection is useful for the modification of small peptides on solid support. However, parallel research demonstrated that when this technology was applied to larger peptides (greater than 10 amino acids), or depending on some amino acid sequence, complications could arise. Total deprotection, on solid support, of two cysteine residues of EPO fragment (1-32) had proved difficult (Dr. S. Mezzato personal communication). The major product observed was the unreduced substrate, followed by the monodeprotected material. The desired double deprotected peptide was identified only in small amounts (Scheme 5.13). Moreover, studies of Cu(I)-catalysed modification of Huisgen cycloaddition performed on solid support for the synthesis of cyclic peptides suggested that the rate of the reaction is slower when the acetylene, rather than the azide, is on the solid support. As discussed earlier (Chapter 5, Section 5.2), the copper (I) insertion in the acetylene is the limiting step in the cycloaddition reaction. When the acetylene is in solution, the kinetics of the process should be fast enough to avoid generation of side products. In such a case the performance of the reaction would be improved. Consequently, having all these precedents, incorporation of different units of the PSDC to an EPO peptide fragment would be preferred to be performed in solution.

To demonstrate this hypothesis and to evaluate the potential of the work, 112 was introduced to an EPO peptide fragment (22-32) 135, on solid support, through the thiols of two cysteine residues (Cys24 and Cys 30). Cleavage of the molecule from the resin, was followed by purification by semi-preparative HPLC and freeze drying. At this point it was decided to perform the "click chemistry" reaction between the peptide fragment 136, and the un-natural carbohydrate scaffold 104.
5. Novel neo-glycopeptide linkage using Cu(I) catalysed 1, 2, 3-triazole formation

Lin and co-workers had modified the expressed proteins, maltose binding protein (MBP) and enhanced green fluorescent protein (EGFP). In both cases an unnatural amino acid with an alkyne chain was introduced at the C-terminus. These compounds were then conjugated with azido molecules, either containing GlcNAc, biotin or fluorescein isothiocyanate (FITC). Finally, the modified alkyne MBP was mounted on an azido-modified glass slide efficiently using "click chemistry". Site-specific ligations were performed with biotinylated maltose to corroborate optimal modifications. Applying the protocols of Lin and co-workers, the alkyne modified EPO peptide fragment (22-32) was treated with 7 equivalents of the (3.5 equivalents of peptide scaffold displaying carbohydrates per acetylene), 36 equivalents of tris-(2-carboxyethyl)phosphine hydrochloride (TCEP), tris-benzyltriazolylmethylamine (TBTA) and 11 equivalents of buffered copper (II) sulphate. However, after 24 hours, all PSDC equipped with an azido group appeared to have been fully reduced to the amine. MS-analysis of the HPLC fractions confirmed the reduction of the azide at the anomeric position, and the HPLC trace showed the presence of an

Scheme 5.13: HPLC trace of EPO (1-32) fragment, after tert-butyl dithiol deprotection: 5 % DTT and 2.5 % DIPEA in DMF for 24 hours (picture donated by Dr. S. Mezzato).
5. Novel neo-glycopeptide linkage using Cu(I) catalysed 1, 2, 3-triazole formation

epimeric mixture (Scheme 5.14). At this stage it was necessary to consider a possible explanation for such a result.

Scheme 5.14: HPLC trace of the crude from direct glycosylation of the EPO peptide fragment (22-32) by “click chemistry” in the presence of TCEP. Mass spectra data from the reduced PSDC, 137: MW: 1500.00 (MW\text{calc}: 1498.54).

Conceivably, the alkyl chains were shielded from reaction by the steric bulk of the surrounding peptide scaffold. This could diminish the rate of the reaction significantly. Alternatively, the combination of excess of TCEP and reducing phosphine additive could favour a competitive process such as the Staudinger reaction. This, in turn, would generate the mixture of epimers observed, 137.

It was still considered valuable to perform the Cu(I)-catalysed modification of the Huisgen cycloaddition on the modified peptide scaffold in buffered solution. As a result, efforts were directed towards optimising this transformation. In model studies, the modified EPO peptide
5. Novel neo-glycopeptide linkage using Cu(I) catalysed 1, 2, 3-triazole formation

Fragment (22-32) would remain the source of alkynes, while N-acetylglucosamine azide would be used to replace the key PSDC (Table 5.1).

<table>
<thead>
<tr>
<th>Cu</th>
<th>Reducing agent</th>
<th>TBTA ligand</th>
<th>Solvent</th>
<th>Time</th>
<th>Temperature</th>
<th>Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2 equiv CuSO4</td>
<td>0.4 equiv. TCEP</td>
<td>0.2 equiv.</td>
<td>PBS buffer, 0.1 M, pH 8</td>
<td>16 h</td>
<td>40 °C</td>
<td>18 %</td>
</tr>
<tr>
<td>0.2 equiv CuSO4</td>
<td>0.4 equiv. sodium ascorbate</td>
<td>0.2 equiv.</td>
<td>PBS buffer, 0.1 M, pH 8</td>
<td>16 h</td>
<td>40 °C</td>
<td>15 %</td>
</tr>
<tr>
<td>0.2 equiv CuSO4</td>
<td>0.4 equiv. sodium ascorbate</td>
<td>0.2 equiv.</td>
<td>H2O</td>
<td>16 h</td>
<td>40 °C</td>
<td>NO</td>
</tr>
<tr>
<td>0.2 equiv CuSO4.H2O</td>
<td>0.4 equiv. sodium ascorbate</td>
<td>0.2 equiv.</td>
<td>H2O/2BuOH (1:1)</td>
<td>16 h</td>
<td>40 °C</td>
<td>NO</td>
</tr>
<tr>
<td>0.2 equiv CuSO4.H2O</td>
<td>1.8 equiv. TCEP</td>
<td>0.4 equiv.</td>
<td>PBS buffer, 0.1 M, pH 8</td>
<td>48 h</td>
<td>40 °C</td>
<td>23 %</td>
</tr>
<tr>
<td>0.2 equiv CuSO4.H2O</td>
<td>1.8 equiv. sodium ascorbate</td>
<td>0.4 equiv.</td>
<td>PBS buffer, 0.1 M, pH 8</td>
<td>48 h</td>
<td>40 °C</td>
<td>26 %</td>
</tr>
<tr>
<td>0.2 equiv CuSO4.H2O</td>
<td>1.8 equiv. sodium ascorbate</td>
<td>0.4 equiv.</td>
<td>H2O</td>
<td>48 h</td>
<td>40 °C</td>
<td>NO</td>
</tr>
<tr>
<td>1.0 equiv CuSO4.H2O</td>
<td>10.0 equiv. sodium ascorbate</td>
<td>10.0 equiv.</td>
<td>PBS buffer, 0.1 M, pH 8</td>
<td>48 h</td>
<td>20 °C</td>
<td>61 %</td>
</tr>
</tbody>
</table>

Table 5.1: Different sets of conditions applied during the development of the “click chemistry” reaction in solution, between a modified peptide and an azido sugar.

The Cu(I)-catalysed modification of Huisgen cycloaddition between 136 and 121 was attempted under a range of conditions (Table 5.1). In each case the reaction was performed on a micromolar scale (1.04x10^{-3} mmol), with respect to the peptide, and the reaction was monitored by MS-analysis of the HPLC fractions.

From these experiments, a number of conclusions could be drawn: i) increased amounts of reducing agent avoided the formation of peptide dimers through the N-terminal cysteine residues; ii) the presence of buffer also minimised dimer formation; and iii) increased amounts of the TBTA additive generally led to cleaner reactions with higher yields of the desired products.
Having reviewed the literature and analysed the outcome of the model experiments, it was possible to devise an optimal protocol for this transformation. The reaction conditions used were: PBS buffered (0.1 M, pH 8), 1 equivalent of CuSO₄·5H₂O, and 25 equivalents of sodium ascorbate and TBTA ligand. No TCEP was employed and the reaction time would be reduced, because it did not affect to the final yield after 24 hours (Scheme 5.15).

**Scheme 5.15:** i) reaction conditions between an azido sugar 121, and an acetylene modified EPO peptide fragment (22-32) 136, to obtain an unnatural peptide, 138. ii) HPLC trace of the crude of previous reaction. iii) Mass spectra of the fraction containing the unnatural peptide 138, MW: 1844.20 (MW calc: 1842.66).

The next step was to apply this protocol to introduce the PSDC 104 to the alkyne modified EPO peptide fragment (22-32) 136. Therefore, 104 and 136 were dissolved together in a 0.1 M PBS buffer solution pH 8.0, in the presence of 25 equivalents of sodium ascorbate and TBTA, and 1 equivalent of CuSO₄·5H₂O; and the mixture was stirred for 24 hours at 20 °C and 1200 rpm. After HPLC purification, the desired modified EPO peptide fragment (22-32)
with 2 PSDC attached 139 was obtained in 69 % yield, and 18 % of recovery of unreacted 104 (Scheme 5.16).

Scheme 5.16: HPLC trace of the reaction between the PSDC 104, and the EPO peptide fragment (22-32) 136, to produce an unnatural glycopeptide 139, MW: 4402.80 (MW_{calc}: 4400.60).

This final result demonstrated the potential of the Cu(I)-catalysed modification of Huisgen cycloaddition for introducing carbohydrates to peptides, regardless of the size of the sugar, either on solid phase or in buffer solution. Moreover, the complex reactions involved in β-mannosylation and glycosyl elaboration, were avoided (Chapter 4).

5.5 NMR observations

During the development of the neo-glycopeptide linkage compatible with NCL,\textsuperscript{171} interesting effects were observed by \textsuperscript{1}H-NMR spectroscopy during the formation of the triazole ring.
This phenomenon was also observed in the case of the PSDC, either linked with the acetylene motif or with the EPO peptide fragment.

When the peracetylated azido GlcNAc \(95\) reacted with \(N\)-(propargyl)-bromoacetamide \(112\) in the presence of copper (I), a triazole ring was generated. Formation of the aromatic ring was indicated in the \(^1\)H-NMR spectrum, by the appearance of an apparent singlet signal at 7.72 ppm. This singlet couples with another apparent singlet at 4.19 ppm that corresponds to the two protons \(\alpha\) to the triazole (coupling determined by COSY experiments). Furthermore, the methyl component of the N-acetyl moiety is shifted to lower frequency, from 1.96 to 1.52 ppm on the formation of the aromatic ring (see experimental part, Chapter 7). A similar observation was made in the case of deacetylated azido GlcNAc molecule: from 2.16 (121) to 1.71 ppm (122) (Scheme 5.17). A possible explanation for this process could be related to the fact that the protons of the N-acetyl group were located perpendicular to the aromatic ring. In this position, they would be aligned opposite to the orientation of the magnetic field of the triazole. Consequently, the singlet that corresponds to 3 protons, was shifted to lower frequency. Theoretical model structures analysed by ChemDraw 3D\(^\circledR\), confirmed the proposed stereoconformation previously described for 122.

\[\text{Scheme 5.17: two different conformational angles for 122, visualised by ChemDraw 3D\(^\circledR\).}\]
These effects were also observed once the PSDC, 104 (Scheme 5.18), was attached to the EPO peptide fragment (22-32), 136 (Scheme 5.19). In this case, the shift of the methyl group in 104 was 2.05 ppm, while in 136 it was 1.82 ppm.

To verify the previous hypothesis of stereoconformation of the methyl of the N-acetyl group, it would be interesting to obtain a crystal structure of derivatives from 122.

**Scheme 5.18:** $^1$H-NMR spectrum of 104, with amplifications of the aromatic and the aliphatic regions.
5. Novel neo-glycopeptide linkage using Cu(I) catalysed 1, 2, 3-triazole formation

Scheme 5.19: $^1$H-NMR spectrum of 136, with amplifications of the aromatic and the aliphatic regions.

5.6 Conclusions and future work

A new methodology to facilitate the incorporation of carbohydrates (azido sugars) on modified peptides with acetylenes has been developed. This new technology can be applied without any restriction in the presence of previous orthogonal functional groups, avoiding harsh and complicated sets of conditions for glycosylation (Chapter 3). In addition, this methodology is also versatile and flexible because: i) it can be applied either on solution or solid phase; ii) the order of the reactions do not alter the outcome of the reaction; iii) the chemistry is compatible with all the peptide components and functional groups; and iv) the product is also compatible with harsh acid conditions and ligations, such as NCL. ¹⁷¹
Applying this orthogonal reaction it should be possible to incorporate more elaborated and ambitious carbohydrate structures on a peptide.

In addition, two PSDC units (Chapter 4) have been attached to an EPO peptide fragment in aqueous solution, both by an S<sub>2</sub>2 reaction between PSDC bromoacetamide and cysteine thiol group, or by Cu(I)-catalysed modification of Huisgen cycloaddition. Finally, diamagnetic effects on the protons of the methyl group of the N-acetyl close to the anomeric centre, after the formation of the triazole ring have been observed. Crystal structure determination should confirm the hypothetical conformation of the methyl group.

Contemporary to this work, Danishefsky and co-workers also reported a similar methodology to prepare neo-glycopeptides using "click chemistry" for the production of anticancer vaccines<sup>143</sup> (Scheme 5.20)<sup>217</sup>. In this work, they attached to a peptide backbone through modified lysines with acetylenes<sup>141</sup>; either an amino acid<sup>142</sup>, or a tripeptide scaffold furnished with Tn and STn antigens; and equipped with a 3-aminopropyl azide residue. Applying "click chemistry", they manage to ligate these two building blocks in the presence of copper. This neo-glycopeptide linkage facilitates the insertion of elaborated structures to a peptide increasing the yields and avoiding extra-modifications compared with aspartilation carried out in their research group.

At this stage, it remained to demonstrate that, although these structures are unnatural and have a neo-glycopeptide linkage, they behave and can be recognised as the native oligosaccharide found in N-linked glycoproteins. Therefore, a qualitative method to demonstrate this behaviour was required before attempting more challenging projects (Chapter 6).
5. Novel neo-glycopeptide linkage using Cu(I) catalysed 1, 2, 3-triazole formation

Scheme 5.20: Carbohydrate-based anticancer vaccines 143 through “click chemistry” \(^\text{217}\)
6 Protein-carbohydrate interactions between PSDC and lectins visualised using modified GFP

6.1 Introduction

Supramolecular interactions occur between the three classes of natural biopolymers: nucleic acids, proteins and carbohydrates. These forces are responsible for significant biological effects. Although the study of nucleic acid and protein interactions are well known through genomics and proteomics respectively (Scheme 6.1), glycomics, the study of carbohydrate interactions with the two other biopolymers and the biological consequences, is an emerging scientific discipline in its early stages.

Scheme 6.1: Interactions of the three main biopolymers.\textsuperscript{218}

Since carbohydrate interactions are involved in cell recognition and adhesion, glycomics promises to be a fertile ground for glycobiology.\textsuperscript{219} It can be anticipated that modulating,
amplifying or suppressing these supramolecular interactions will allow the understanding of their biological roles and may lead to the discovery and availability of new therapeutic drugs and treatments.

Chemists have developed many methodologies to generate pure saccharides (described in Chapter 2). Now, with the development of high-throughput screening, such as microarray-based methods, glycochips, surface plasmon resonance spectroscopy (SPR), mass spectrometric methods, NMR or ELISA-type methods, it is possible to quantify the interactions of a large number of carbohydrates with their biological counter ligand. 219-221

Microarray-based methods or glycochips, is a technique analogous to those developed for DNA and proteins. Carbohydrates are covalently or non-covalently immobilized on a surface (polystyrene, thiol-derivatised glass slides, or gold slides). 228 This immobilization has emerged as a new tool to facilitate the study of carbohydrate-protein interactions and for the identification of optimal sugar ligands and inhibitors. 208 Covalent immobilization can be achieved through a Diels-Alder, hetero-Michael addition or 1,3-dipolar cycloaddition reaction between an azide and alkyne group, in the presence of catalytic amounts of copper (I). 229

Shin and co-workers 222 developed carbohydrate chips for studying high-throughput carbohydrate-protein interactions. They prepared carbohydrate microarrays by immobilizing maleimide-linked carbohydrate on thiol-derivatised glass slides (Scheme 6.2.A). In addition, sLe\(^x\) was prepared from GlcNAc attached to the glass slide, by three consecutive glycosyl transferase-catalysed reactions to obtain the desired final oligosaccharide structure (Scheme 6.2.C). These glycochips allowed the authors to quantify the binding affinities between the carbohydrates and lectins. Moreover, binding affinities of lectins to carbohydrates were also quantitatively analysed by determining IC\(_{50}\) values of soluble carbohydrates with the carbohydrate array.
6. Protein-carbohydrate interactions between PSDC and lectins visualised using modified GFP


Globo H tumour antigen hexasaccharide is an epitope found on the cell surface in breast, prostate and ovarian cancers cells. Wong and co-workers\(^{223}\) developed a carbohydrate microarray for profiling antibody interaction with this hexasaccharide tumour antigen (Scheme 6.3). The authors arrayed covalently, on glass slides, a library of different length Globo H polysaccharides obtained by one-pot oligosaccharide synthesis.\(^{47}\) Thereafter, they used two monoclonal antibodies (MRrl and VK-9) against the hexasaccharide and serum from breast cancer patients using fluorescence-based binding analysis. Analysis with the
monoclonal antibodies showed that the terminal tetrasaccharide and the native hexasaccharide bind equally well (Scheme 6.3.A., 144a, 144b, 145a and 145b). In addition, the fucose residue was found to be required for effective binding. On the other hand, serum bound both defucosylated pentasaccharide and fucosylated hexasaccharide without a significant difference (Scheme 6.3.B., 145a, 145b and 146). This final result could be caused either due to a possible polyclonal nature of the antibodies from the patients against Globo H, or the antibodies were generated at different stages to recognise the defucosylated pentasaccharide and fucosylated hexasaccharide.

Seeberger and co-workers used protein-carbohydrate interactions to develop a carbohydrate-functionalsed soluble poly(p-phenylene ethylene) (PPE) that can be used as a fast method (10 and 15 minutes) for the detection of *Escherichia coli* (Scheme 6.4). They functionalised the polymer after polymerisation with mannose and galactose. The molecule derivatised with mannose showed strong affinity to *E. coli*. This interaction and cell aggregation is due to multivalent interactions between the mannosylated polymer and mannose receptors located on the bacterial pili.

**Scheme 6.4:** Laser scanning confocal microscopy image of A) mutant *E. coli* that does not bind to 174a (mannose derivative polymer). B) fluorescent bacterial aggregate due to multivalent interactions between the mannose-binding bacteria pili and 174a. C) fluorescence microscopy image of a large fluorescent bacterial cluster. D) Structure of the polymers used in the study (Reprinted with permission from *J. Am. Chem. Soc.* 2004, 126, 13343-13346. 10.1021/ja047936i CCC: $27.50 © 2004 American Chemical Society).
6.2 Objective: development of a methodology to visualise protein-carbohydrate interactions

Based upon literature precedent for detection of protein-carbohydrate interactions, a new simple methodology to visualise these interactions using fluorescence was set out by attaching three carbohydrate tags linked through cysteines of a peptide to Green Fluorescent Protein (GFP). The rationale behind using GFP is because it is easy to produce by bacteria expression. Moreover, the product would be an unnatural glycoprotein, a molecule that is more relevant to the biological studies than attaching the unnatural glycopeptide to a dansyl or fluorescein tag.

The introduction of the sugar molecules was carried out applying the technology described previously. The introduction of the modified peptide with carbohydrates to GFP would be achieved by expressed protein ligation (EPL).

To visualise protein-carbohydrate interactions, it was envisaged that the modified GFP would be dissolved in an aqueous solution. In the presence of agarose beads coated with lectins, the carbohydrates present on the modified GFP should interact specifically with their counter ligand-lectin, staining the agarose beads green. After several washes, this type of interaction should remain intact, with the green colour staining the agarose beads.

Once the viability of the methodology to visualise protein-carbohydrate interaction was demonstrated, it could be used to demonstrate that the PSDC which was developed (Chapter 4) can be recognised by a lectin once it is on a protein, and therefore corroborate the hypothesis that these structures, although unnatural, behave similarly to typical native polysaccharide structures found in N-linked glycoproteins.
6.3 Green Fluorescent Protein (GFP)

GFP is a protein produced by the Pacific North West jellyfish *Aequoria Victoria*. Wild-type GFP is a stable proteolysis-resistant single chain of 238 residues. The protein has a fluorescent activity at 394 nm, with an emission maximum of 508 nm. This fluorophore results from the autocatalytic cyclization of the polypeptide backbone between residues Ser$^{65}$ and Gly$^{67}$ and oxidation of the $\alpha$-$\beta$ bond of Tyr$^{66}$.

Expression of the protein in *E. coli* and crystal structure determination has been crucial for the widespread use of GFP as a reporter for gene expression and protein localization. The Macmillan research group has developed the technology to express on *E. coli* GFP thioester, allowing the modification of the protein by expressed protein ligation (EPL).

6.4 Expressed Protein Ligation (EPL)

EPL is a procedure to synthesise recombinant proteins based on the natural posttranslational phenomenon of protein splicing. This is a process in which a single protein sequence undergoes an intramolecular rearrangement resulting in the extrusion of an internal sequence, *intein*, and the joining of lateral sequences, *exteins* (Scheme 6.5). In EPL, the expressed protein can be bound to an affinity resin, where the incipient thioester can be trapped by thiolysis or with an N-terminal cysteine peptide fragment, thus eluting the target protein as a C-terminal thioester or the desired modified protein; and leaving the remainder of the fusion protein bound to the affinity resin.

EPL is a very powerful technique for the modification of proteins. Molecular biology and chemistry come together, providing the opportunity to generate large proteins or introduce unnatural substrates to the protein, providing to the acceptor molecule new physical and chemical properties.
6. Protein-carbohydrate interactions between PSDC and lectins visualised using modified GFP

Expressed Protein Ligation

Protein Splicing

Scheme 6.5: Scheme of Expressed Protein Ligation and Protein Splicing.

It has been possible to perform intein-mediated protein splicing in living cells, with the horizon of new biological applications. Different strategies are described in the literature:

- Non-covalent association of 2 intein fragments to generate an active molecule that performs the splicing event to produce a functional protein.\(^ {232-234} \)
- Binding of a small molecule to activate an intein that can then activate a target molecule.\(^ {235} \)

Muir and co-workers\(^ {229} \) synthesised a pair of biologically active proteins, C-terminal Src kinase, Csk, molecule of 50 kDa. Wong,\(^ {236} \) et. al., combined EPL with cysteine derivatives, like H-Cys-Asn(GlcNAc)-OH, and later modified this by galactosyl transferase. As mentioned earlier, Macmillan and Bertozzi,\(^ {89,128} \) were also able to synthesise a physiological glycoprotein using this technique. A larger number of other proteins have been engineered using EPL.\(^ {237} \)
Applying EPL it would be possible to modify GFP thioester available in the Macmillan research group, to introduce unnatural glycopeptides, such as 134 and 139.

6.5 Development of methodology to visualise carbohydrate-protein interaction

6.5.1 Synthesis of N-(iodoacetyl)-p-aminophenyl mannopyranoside

For the successful synthesis of N-(2-iodoacetyl)-p-aminophenyl mannopyranoside 155, four synthetic steps were necessary (Scheme 6.6). This molecule was synthesised as an iodoacetamide derivative with a view to being incorporated into peptides through cysteine residues, and then to be attached to a protein by EPL. Therefore, the synthesis of the desired mannose derivative 155 began from the commercially available p-nitrophenyl mannopyranoside 150.

Scheme 6.6: synthesis of N-(2-iodoacetyl)p-aminophenyl mannopyranoside 155.

p-Nitrophenyl mannopyranose 150 was quantitatively peracetylated with a solution of acetic anhydride and pyridine (1:2) to obtain the desired product 151 (Scheme 6.6). Afterwards, the nitro group was reduced by catalytic transfer hydrogenation (10 % Pd/C, \( \text{NH}_4 \text{HCO}_3 \)),\(^{238,239}\) to obtain p-aminophenyl-2,3,4,6-tetra-O-acetyl-\( \alpha \)-D-mannopyranoside in
100 % yield. The primary amino group was then acetylated with 2-iodoacetyl chloride with stoichiometric amounts of pyridine, in dry DCM, at -30 °C, to achieve the product in 69 % yield. Finally, the compound was deacetylated, by subjecting the molecule to Zemplén deprotection, to afford \( N\)-(2-iodoacetyl)-\( p\)-aminophenyl mannopyranoside in 92 % yield.

6.5.2 Semi-synthesis of GFP modified with an unnatural glycopeptide displaying three mannoses, for protein immobilisation

Once the desired mannose derivative was obtained, this structure was introduced to a peptide, on solid phase. First a peptide fragment was synthesised using Fmoc-SPPS. The predefined sequence was: L-Cys-L-Ser-L-Cys-L-Cys-L-Cys-L-Ser. The N-terminus was protected with a Boc- group, the hydroxyls of the serines were protected with the tert-butyl protecting group, the thiol of the terminal cysteine was protected with trityl group, and the other thiols of the remaining cysteines, were protected with tert-butyl disulphide protecting group (Scheme 6.7). The disulphide bond protected cysteine derivatives were reduced in the presence of 10 % DTT and 2.5 % DIPEA in DMF. Afterwards, was incorporated into the peptide with a mixture of 2.5 % pyridine in DMF. After TFA cleavage, and purification, the desired hexamer was obtained with three mannose derivatives incorporated. Finally, the modified glycopeptides were incorporated on Green Fluorescent Protein (GFP), using (EPL).

GFP Chitin Binding Domain (CBD) Fusion Protein was overexpressed in Escherichia coli strain BL21-DE3 transformed with pTYB1-GFP. The protein was loaded on chitin beads New England Biolabs (NEB) and equilibrated with column buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM EDTA, 1 mM PMSF). The cell-free extract loaded was provided as a gift by Dr. D. Macmillan.
Scheme 6.7: semi-synthesis of modified Green Fluorescent Protein (GFP) in the C-terminus 161, with a peptide displaying mannose derivatives through the thiols of 3 cysteines.

Afterwards, the chitin beads (NEB) were washed with 200 mM sodium phosphate buffered solution, pH 8.0, 100 mM sodium chloride, 2 % w/v mercaptoethanesulfonic acid (MESNA), 5 mM \( tris \)-\(2\)-carboxyethyl\)phosphine hydrochloride (TCEP-HCl) solution (5.0 mL), and the column was allowed to drain by gravity flow.

The modified peptide was dissolved in aqueous solution, and the contents were mixed gently for 48 hours. After eluting and washing the chitin beads (NEB) with buffer, the samples were concentrated in a Vivaspin \(^\text{®} \) 2 mL concentrator, 10,000 \( \mu \)m MWCO PES, to remove the excess of unreacted glycopeptide (Scheme 6.8).
6. Protein-carbohydrate interactions between PSiDC and lectins visualised using modified GFP

Scheme 6.8: a) structure of the mannose derivative incorporated into the peptide through cysteines to synthesise peptide tags; b) ESI-MS of the peptide tag: MW: 1539.00 (MW calc: 1537.41); c) ESI-MS of the peptide tag incorporated into GFP; d) photograph of the fluorescence of GFP coated on Concanavalin A beads.

6.5.3 Protein-carbohydrate interaction test on agarose beads, visualised with modified GFP

Once the modified GFP with an unnatural glycopeptide displaying 3 mannose derivatives was obtained, it was decided to perform interaction protein-carbohydrate studies between the unnatural glycoprotein and lectin-coated agarose beads.

In these interaction studies, native GFP served as a control test, while modified GFP protein with a peptide displaying 3 units of mannose was the test experiment. It was decided to use agarose beads coated with Concanavalin A, as a solid support to visualise protein-carbohydrate interactions between these two molecules. Concanavalin A is a legume lectin
6. Protein-carbohydrate interactions between PSDC and lectins visualised using modified GFP form *Canavalia ensiformis* that specifically recognises mannose residues of many complex glycans.\(^{240}\)

Consequently, both molecules were in PBS solution and dispatched in separate 1.5 mL eppendorfs tubes. Afterwards, small amounts of agarose beads coated with a lectin, *Concanavalin A*, were added to each eppendorf, and the mixture was shaken gently. After 30 minutes, any protein-carbohydrate interaction was observed in the test experiment eppendorf (Scheme 6.9: + ConA-agarose beads). Finally, after washing the agarose beads coated with the lectin with PBS solution, the agarose beads were still stained with GFP visually indicating the predicted protein-carbohydrate interaction between the lectin and the mannose tags (Scheme 6.9).

**Scheme 6.9:** Interaction studies between GFP labelled peptide displaying mannose, and lectin (Con A)-coated agarose beads.
6.6 Recognition of the PSDC in a biologically relevant environment

Once a method to visualise specific protein-carbohydrate interaction was developed, this technology was applied to verify if the PSDC, once attached to the EPO peptide fragments 134 and 139 displayed carbohydrates in such a way that can be recognised by a lectin, and therefore, certify that the structure mimics the natural nonasaccharide structure found in N-linked glycoproteins (Scheme 4.4).

6.6.1 Semi-synthesis of GFP modified with an EPO peptide fragment containing two PSDC, for protein immobilisation

GFP was overexpressed as a fusion intein-CBD in E. coli. Subsequently, the protein was loaded on chitin beads. The suspension of chitin beads coated with GFP was provided, as a gift, by Dr. J. P. Richardson (Macmillan research group). Afterwards, the beads were washed with 200 mM sodium phosphate buffer solution, pH 8.0, 100 mM NaCl, 2 % w/v MESNA, 5 mM TCEP solution (5.0 mL). The column was allowed to drain by gravity flow. Finally, the modified EPO fragments with the peptide scaffold displaying carbohydrates 134 and 139 were dissolved in the column buffer, and the contents were mixed gently, for 48 hours. SDS-PAGE analysis of samples from the reaction indicated partial success in the reaction (Scheme 6.10).

Afterwards, the reaction mixture was eluted with PBS buffer, and the eluted samples were concentrated and dialyzed with Dulbecco’s PBS solution, in a Vivaspin® 2 mL concentrator, 10,000 μm MWCO PES, a membrane that dialyses all particles with molecular weight below 10 kDa. Finally, the green dialysed protein was resuspended again with Dulbecco’s PBS solution.
6.6.2 **Protein-carbohydrate interaction test on agarose beads, visualised with modified GFP**

Having the methodology to qualify and visualise protein-carbohydrate interactions (Scheme 6.9), it was decided to apply it to the new GFP protein that was modified with an EPO peptide fragment containing two PSDC. In this case, the test to visualise the protein-carbohydrate interaction was carried out with agarose beads coated with *Ricinus Communis* RCA120, a lectin that has specific avidity for galactose (the terminal monosaccharide of the PSDC is galactose, see Chapter 4).

Consequently, the modified proteins with the peptide 134 and 139 were dissolved in Dulbecco’s PBS solution and dispatched separately in eppendorfs. Native GFP was used as a control test in a separated eppendorf tube (Scheme 6.11). Afterwards, agarose beads coated with *Ricinus Communis* RCA120, were added to each eppendorf, and the mixture was shaken gently for 30 minutes. Finally, after washing the agarose beads coated with the lectin with Dulbecco’s PBS solution, no green colouration of the agarose beads was observed. This unfortunate result appeared to reveal that no interaction occurred between the lectin and...
the GFP modified with the EPO peptide fragment furnished with two PSDC molecules. Studying carefully this result, it was realised that during the ligation between GFP-intein chitin binding and the unnatural glycopeptides, 134 and 139 (Scheme 6.10), an undetermined excess of the unnatural glycopeptides were dissolved in the reaction mixture, and therefore, the remaining excess would be in the crude product. Therefore, the presence of excess 134 and 139 in the test could compete and interfere with the interaction between the lectin (RCA120) and the GFP modified with the EPO fragment containing two PSDC. As a consequence, excess of 134 and 139 could inhibit the protein-carbohydrate interaction.

A possible reason for the presence of excess of 134 and 139 could have been due to the obstruction of the pores of the dialysis membrane when the crude of ligation reaction was dialyzed in the Vivaspin® 2 mL concentrator, 10,000 μm MWCO PES. Consequently, the dialysis was repeated using a porous filter diameter of 30,000 μm, assuming that better dialysis efficiency could be achieved (dialyzed all compounds below MW: 30 kDa), and eliminating the excess of unreacted unnatural glycopeptide (134, MW: 4401.20 (MWcalc: 4399.65); 139, MW: 4402.80 (MWcalc: 4400.60)).

Once again, the interaction test was repeated between unnatural glycoprotein and agarose beads coated with *Ricinus Communis* RCA120. When the agarose beads were washed with Dulbecco’s PBS solution, a green coloration of the agarose beads was observed, demonstrating interaction between the galactose present in the PSDC and the lectin *Ricinus Communis* RCA120 (Scheme 6.11). This coloration of the solid support was not as intense as the previous experiments when GFP labelled peptide displaying mannose was studied (Scheme 6.9). This could be due to a remaining presence of excess of non-ligated EPO peptide fragment displaying two PSDC (134 and 139) and consequently, partially inhibiting the interaction of the modified GFP with the lectin.
can be recognised in a biological environment. Having previous positive results, it is possible to extend this technology for the semi-synthesis of non-natural erythropoietin, where the terminal galactose could be enzymatically modified by the addition of sialic acid.

For the semi-synthesis of non-natural EPO, protein fragment should be expressed on *E. coli*. Afterwards, and applying the technology that the Macmillan research group developed, the expressed protein could be selectively cleaved with cyanogen bromide, introducing an internal cysteine to the *N-terminus* (Scheme 6.12).

On the other hand, an EPO peptide fragment (1-32) has been synthesised using *Fmoc*-SPPS, using a “safety-catch” linker to generate, afterwards, a thioester. In this peptide sequence, in positions 24 and 30, a cysteine modified with an alkyne chain through the thiol has been introduced. Finally, the unnatural EPO peptide fragment has been cleaved by activating the “safety-catch” linker and generate a thioester in the presence of a thiol (Scheme 6.12).

\[
\text{Scheme 6.12: Possible semi-synthesis of an alkyne-modified mutant EPO (Asn24Cys and Asn30Cys).}
\]
Having available the two different EPO polypeptide fragments, and applying expressed protein ligation, the full length EPO protein (1-166) could be generated, where, as mentioned before, in the positions 24 and 30, the protein would have an alkyne substituent (Scheme 6.12).

Afterwards, and applying the technology described in Chapter 5, the "click chemistry" reaction could be performed, in buffered solution at pH 8, between the modified EPO protein, and at least two PSDC furnished with an azide group (Scheme 6.13).

![Scheme 6.13: Introduction of two PSDC through the alkynes of the modified mutant EPO by click chemistry, addition of sialic acid with α-2,3-sialyltransferase, and final folding to achieve a therapeutic unnatural EPO.](image)

Finally, and using a commercially available sialyltransferase, a sialic acid could be inserted on the 3-hydroxy group of the galactose to obtain the desired structure. Final protein
6. Protein-carbohydrate interactions between PSDC and lectins visualised using modified GFP folding should provide an unnatural therapeutic glycoprotein with 4 to 8 sialic acid residues, responsible for protecting hrEPO from hepatic metabolism, and therefore increasing the serum half life \( (t_{1/2}) \).
Overall conclusions

Carbohydrates are very important and perform multiple roles in biology. They can influence and modify the biological and physicochemical properties of proteins. Consequently, these molecules have captivated the attention of the chemist and biochemist community. Unfortunately, glycosylation is not under genetic control, and a heterogeneous mixture of glycosylated forms is generated. Therefore, semi-synthesis of homogeneous glycoproteins should allow the understanding of the biological role of the structures.

Synthesis of natural glycoproteins can potentially help to comprehend the different functions of these molecules in biology. Unfortunately, the semi-synthesis of glycoproteins is far from easy and demands a large number of steps. On the other hand, mimicking these molecules may permit more rapid access to complex structures and allow the study of saccharide interactions that could provide better understanding of the biological activity and potential target for therapeutic agents provided that the unnatural elements of the structure are tolerated.

Based on the knowledge accumulated in the Macmillan research group, the semi-synthesis of a natural O-linked glycopeptide was accomplished, and unnatural N-linked glycoprotein. These molecules can be recognised in a biological environment.

A successful preparative synthesis of the glyco-amino acid building block has been accomplished. The molecule could allow different research opportunities, using the molecule as starting material for SPPS. Moreover, most of the reactions developed for the
synthesis of core-1 building block, could be used for the synthesis of the core-2 structure. On the other hand, this synthetic route was laborious and time consuming, and the final glycosylation step to attach the disaccharide to threonine appeared to be unproductive for such a valuable molecule. In the end, the glyco-amino acid building block 82 was prepared in 6.83 % yield after 15 synthetic steps.

Consequently, a PSDC on solid phase has been developed. Based on the work of Meldal (Scheme 2.8),14 103 and 104 were designed to mimic the diantennary carbohydrate structure found in N-linked glycoproteins. The PSDC allows the construction of highly complex unnatural carbohydrate structures from only two predefined glycobuilding blocks, avoiding the need to control the regio and stereochemistry in many instances, and therefore simplifying the total synthesis of the desired molecule. This concept is exemplified by avoiding the difficult β-mannosylation step. The deacetylated PSDC 104 was ultimately prepared in 10 steps in 47 % yield.

In addition, having available the methodology to obtain PSDC on solid phase, a library of these molecules could be prepared by increasing the number of cysteines, and therefore changing the number of carbohydrates displayed. Moreover, the lysine residue could be altered by, for example, changing the number of carbons of the alkyl chain. This should provide the opportunity to identify which peptide scaffold shows better biological activity. The PSDC synthesis does however still need optimization and detailed conformational studies before it can be truly described as an oligosaccharide mimic.

A novel neo-glycopeptide linkage using Cu(I) catalysed 1,2,3-triazole formation has been developed. With this new technology it is possible to introduce carbohydrates to peptides. Moreover, it can be applied in the presence of various orthogonal functional groups, avoiding harsh and problematic set of conditions for glycosylation. In addition, the novel neo-glycopeptide linkage is also versatile and flexible because: i) it can be applied either in
solution or on solid phase; ii) the order in which the reactions are performed does not appear to alter the outcome of the reaction; iii) the chemistry is compatible with all the peptide components and functional groups; and iv) and the product is also compatible with harsh acid conditions and ligations, such as NCL. Applying this new linkage, two PSDC have been attached to a EPO peptide fragment in aqueous solution, either by an S_N2 reaction between PSDC bromoacetamide and cysteine thiol group (19 % yield in 2 steps), or by Cu(I)-catalysed modification of Huisgen cycloaddition (69 % yield in 1 step).

Finally, a method to visualise protein-carbohydrate interactions has been developed, where GFP modified with a peptide containing carbohydrates (mannose), interacts with agarose beads coated with lectins (Concanavalin A). This interaction is visualised by the green colour of the modified GFP that stains the agarose beads, and the colouration remains after several buffer washes. Applying this methodology to visualise protein-carbohydrate interactions demonstrated that the PSDC attached to an EPO peptide fragment presents correctly galactose, and so, they can be recognised in a biological environment by a lectin (in this case RCA_{120}). Therefore, this technology can be extended for the semi-synthesis of non-natural erythropoietin, where the terminal galactose could be enzymatically modified by the addition of sialic acid.
7 Experimental section

7.1 General Experimental Details

7.1.1 Instrumentation

$^1$H NMR spectra were recorded at 250, 300 and 500 MHz, $^{13}$C NMR spectra were recorded at 63 and 75 MHz and $^{19}$F NMR spectra were recorded at 235 MHz on a Bruker 250Y instrument. Chemical shifts ($\delta$) were reported in ppm and coupling constants ($J$) in Hz, signals were sharp unless stated as broad (br), s: singlet, d: doublet, t: triplet, m: multiplet and q: quartet. Residual protic solvents: CDCl$_3$ ($\delta_{H}$: 7.26, s), CD$_3$OD ($\delta_{H}$: 3.31, p), D$_2$O (4.79, s), were used as the internal calibrant in $^1$H-NMR spectra unless otherwise stated. Electrospray mass spectrometry was carried out on a Micromass Quattro LC electrospray with an applied voltage of 25-60V.

Automated peptide synthesis was conducted using a peptide synthesizer: Applied Biosystems 433A Peptide Synthesizer, employing the Fastmoc™ protocol.

7.1.2 Chromatography

Analytical TLC was carried out on Merck aluminium backed plates coated with silica gel 60F$_{254}$. Flash chromatography was carried out over Fisher silica gel 60 Å particle size 35-70 micron. Components were visualized using UV light (254 nm) followed by p-anisaldehyde dip.
Purification of peptides on a semi-preparative scale was carried in a reverse phase gradient (water/acetonitrile (5-95%), 0.1 % TFA, over 45 minutes). The HPLC was equipped with a binary pump (P680 HPLC pump Dionex), a semi-preparative column (Phenomenex Luna, 10u, C18, 100A, 250 x 10), UV detector (UVD 170U/340U Dionex), and a fraction collector.

7.1.3 Solvents and reagents

All reagents and solvents were standard laboratory grade and used as supplied unless otherwise stated. Where a solvent was described as dry it was purchased as anhydrous grade. All organic extracts were dried over anhydrous magnesium sulphate prior to evaporation under reduced pressure.
7. Experimental section

7.2 Experimental Procedures

1, 2, 3, 4, 6-Penta-O-acetyl-α-D-galactopyranose, [49]

HClO₄ (100 μL) was added dropwise to a stirring solution of acetic anhydride (225 mL) at 0 °C. Galactose 48 (25.0 g, 0.13 mol) was added in small portions to the reaction mixture during 1 h. After 3 h, TLC (7:3 ethyl acetate/hexane) indicated that the reaction was complete. The reaction mixture was concentrated to one third of its original volume. The mixture was diluted with chloroform (600 mL) and washed with saturated aqueous NaHCO₃ (4 x 150 mL), water (150 mL) and brine (150 mL). The organic phase was dried with MgSO₄, filtered and evaporated under reduced pressure to afford 49 (54.0 g; 100 %) as a pale brown solid. Rf= 0.42 (7:3 ethyl acetate/hexane). ¹H-NMR (250 MHz, CDCl₃) δ (ppm): 6.34 (1H, d, J= 1.7 Hz, H1); 5.46 (1H, d, J= 4.5 Hz, H5); 5.27-5.23 (2H, m, H2, H3); 4.31 (1H, dd, J= 6.5 Hz, J= 4.5 Hz, H4); 4.06-4.04 (2H, m, H6a, H6b); 2.18, 2.12, 2.00, 1.98 and 1.97 (15H, 5 x s, CH₃CO). ¹³C-NMR (63 MHz, CDCl₃) δ (ppm): 170.3, 170.0, 168.9, 166.3 (CO); 89.6, 67.3, 67.2, 66.4 (CH); 61.2 (CH₂); 20.5 (CH₃). FAB-MS calculated for C₁₆H₂₂O₁₁ [MH]+ 391.1162, found 391.12404.

2, 3, 4, 6-Tetra-O-acetyl-α-D-galactopyranosyl bromide, [50]
The crude peracetylated galactose 49 (54.2 g, 0.14 mol) was dissolved in chloroform (75 mL) and the solution was stirred at 0 °C. A solution of HBr 30 % w/v in acetic acid (170 mL) was added and stirring was continued at 0 °C. After 2 h, TLC (1:1 ethyl acetate/hexane) indicated that the reaction was complete. The reaction mixture was diluted with chloroform (250 mL) and washed with water (3 x 150 mL) and saturated aqueous NaHCO₃ (4 x 150 mL). The organic phase was dried with MgSO₄, filtered and evaporated under reduced pressure to afford 50 (57.13 g; 100 %) as a yellow syrup. Rf = 0.43 (1:1 ethyl acetate/hexane). ¹H-NMR (250 MHz, CDCl₃) δ (ppm): 6.67 (1H, d, J₁₂ = 3.9 Hz, H₁); 5.50 (1H, dd, J₃₄ = 3.3 Hz, J₄₅ = 1.1 Hz, H₄); 5.39 (1H, dd, J₂₃ = 10.6 Hz, J₃₄ = 3.3 Hz, H₃); 5.03 (1H, dd, J₂₃ = 10.6 Hz, J₁₂ = 3.9 Hz, H₂); 4.47 (1H, ddd, J₂₃ = 10.6 Hz, J₃₄ = 6.8 Hz, J₅₆α = 6.3 Hz, J₄₅ = 1.1 Hz, H₅); 4.18 (1H, dd, J₆α-₆β = 11.42, J₅₆α = 6.3 Hz, H₆α); 4.08 (1H, dd, J₆α-₆β = 11.42, J₅₆β = 6.8 Hz, H₆β); 2.13, 2.09, 2.05 and 1.99 (12H, 4 x s, CH₃CO). ¹³C-NMR (63 MHz, CDCl₃) δ (ppm): 170.3, 170.0, 169.9, 169.7 (CO); 88.0, 71.0, 67.7, 66.9 (CH); 60.8 (CH₂); 20.6 (CH₃). FAB-MS calculated for C₁₄H₁₉BrO₉ [MH]⁺ 410.0212 (based on ⁷⁹Br) and 412.0212 (based on ⁸¹Br), found 409.01342 and 411.01150.

3, 4, 6-Tri-O-acetyl-D-galactal, [51]

![Chemical Structure](image)

A suspension of zinc dust (64.0 g, 0.97 mol) in water (210 mL) was stirred mechanically at 0 °C. Acetic acid (210 mL) was added to the stirring mixture. A solution of the crude galactosyl bromide 50 (48.6 g, 0.12 mol) in anhydrous ether (120 mL) was added dropwise to the stirring reaction at 0 °C. After the addition was complete, the reaction was allowed to warm to room temperature overnight. After 13 h, the reaction mixture was filtered with
7. Experimental section

The filtrate was evaporated under reduced pressure. The residue was dissolved in ethyl acetate (400 mL) and the organic phase was washed with water (3 x 100 mL) and saturated aqueous NaHCO₃ (3 x 100 mL). The organic phase was dried with MgSO₄, filtered and evaporated under reduced pressure to afford 51 (30.1 g; 94 %) as a brown oil. R_f = 0.25 (1:1 ethyl acetate/petroleum ether 40-60). ¹H-NMR (250 MHz, CDCl₃) δ (ppm): 6.44 (1H, dd, J₁,₂ = 6.3 Hz, J₁,₃ = 1.8 Hz, H1); 5.62-5.60 (1H, m, J₂,₄ = 4.6 Hz, J₁,₃ = 1.8 Hz, J₂,₃ = 1.3 Hz, H3); 5.39 (1H, dd, J₃,₄ = 4.6 Hz, J₄,₅ = 1.6 Hz, H4); 4.80-4.78 (1H, m, J₁,₂ = 6.3 Hz, J₂,₃ = 1.3 Hz, H2); 4.30-4.04 (3H, m, H5, H6a, H6b); 2.19, 2.15 and 2.09 (9H, 3 x s, CH₃CO).

¹³C-NMR (63 MHz, CDCl₃) δ (ppm): 170.5, 170.2, 170.1 (CO); 145.4, 98.8, 72.7, 64.2, 64.1 (CH); 61.9 (CH₂); 20.7 (CH₃). FAB-MS calculated for C₁₂H₁₆O₇ [MH]⁺ 273.0896, found 273.09743.

3, 4, 6-Tri-O-acetyl-2-azido-2-deoxy-D-galactopyranosyl nitrate, ¹⁰⁴,¹⁶⁰ [52]

The crude glycal 51 (36.0 g, 0.13 mol), CAN (219.3 g, 0.40 mol) and NaN₃ (13.0 g, 0.20 mol) were dissolved in anhydrous acetonitrile (870 mL) and reacted at -15 °C under nitrogen, in the presence of molecular sieves. After 3 h, TLC (6:4 petroleum ether 40-60/ethyl acetate) indicated that the reaction was complete. The crude reaction mixture was diluted with cold ether (200 mL) and cold water (200 mL), then filtered through Celite and the aqueous phase was washed with diethyl ether (3 x 100 mL). The organic phase was dried with MgSO₄, filtered and evaporated under reduced pressure. The crude product was purified by flash chromatography over silica (8:2 hexane/ethyl acetate) to afford a mixture of anomers (ratio of anomic α:β mixture (7:1)) 52 (21.5 g; 44 %) as a white resin. R_f = 0.38
7. Experimental section

(6:4 petroleum ether 40-60/ethyl acetate). \( ^1 \text{H-NMR} \) (250 MHz, CDCl\(_3\)) \( \delta \) (ppm) \( \alpha\)-anomer:

6.32 (1H, d, \( J_{1-2} : 4.2 \text{ Hz}, \text{H}1 \)); 5.50-5.46 (1H, m, H4); 4.35-4.33 (1H, m, H5); 4.23 (1H, dd, \( J_{2-3} : 11.3 \text{ Hz}, J_{3-4} : 3.2 \text{ Hz}, \text{H}3 \)); 4.18-4.03 (3H, m, H2, H6\(_b\), H6\(_b\)); 2.15, 2.05, 2.02 (9H, 3 x s, CH\(_3\)CO).

\( \beta\)-anomer:

5.64 (1H, d, \( J_{1-2} : 8.8 \text{ Hz}, \text{H}1 \)); 5.41 (1H, m, H4); 5.05 (1H, d, \( J_{2-3} : 10.6 \text{ Hz}, J_{3-4} : 3.26, \text{H}3 \)); 3.85 (1H, dd, \( J_{2-3} : 10.6 \text{ Hz}, J_{1-2} : 8.8, \text{H}2 \)); 2.18, 2.10, 2.03 (9H, 3 x s, CH\(_3\)CO).

\( ^{13}\text{C-NMR} \) (63 MHz, CDCl\(_3\)) \( \delta \) (ppm) \( \alpha\)-anomer:

170.2, 169.7, 169.4 (CO); 96.8, 69.6, 68.5, 66.6 (CH); 60.9 (CH\(_2\)); 55.9 (CH); 20.5 (CH\(_3\)). \( \beta\)-anomer:

170.2, 169.7, 169.4 (CO); 97.9, 71.7, 66.8, 65.7 (CH); 60.9 (CH\(_2\)); 57.3 (CH); 20.5 (CH\(_3\)).

FAB-MS calculated for C_{12}H_{16}N_{4}O_{10} \[\text{MH}]^+ 377.0866, \text{found} 377.09447.

3, 4, 6-Tri-O-acetyl-2-azido-2-deoxy-D-galactopyranose,\(^{137,138}\) [53]

A solution of the azidonitrate 52 (6.2 g, 16.50 mmol) and thiophenol (5.2 mL, 49.40 mmol) in dry acetonitrile (100 mL) was treated with DIPEA (2.9 mL, 16.50 mmol) under nitrogen at room temperature. After 5 minutes, TLC (1:1 hexane/ethyl acetate) indicated that the reaction was complete. The reaction mixture was concentrated and purified by flash chromatography over silica (1:1 hexane/ethyl acetate) to afford the hemiacetal as a yellow oil. The azido deoxy galactopyranose intermediate was dissolved in DMF (50.00 mL) and cooled to 0 °C. TBDMS-Cl (4.98 g, 33.0 mmol) and imidazole (4.50 g, 66.0 mmol) were added to the reaction mixture, which was allowed to warm at room temperature. After 4 h, TLC (7:3 petroleum ether 40-60/ethyl acetate) indicated that the reaction was complete. The reaction mixture was diluted with water (40 mL) and the product was extracted with DCM (3
7. Experimental section

x 35 mL). The organic phase was washed with water (30 mL), dried with MgSO₄, filtered and evaporated under reduced pressure. The crude product was purified by flash chromatography over silica (7:3 petroleum ether 40-60/ethyl acetate) to afford 54 (6.48 g; 88 %) as a transparent oil. Rf = 0.42 (7:3 petroleum ether 40-60/ethyl acetate). ¹H-NMR (250 MHz, CDCl₃) δ (ppm): 5.29 (1 H; dd; J₃ = 3.4 Hz, J₄ = 1.1 Hz; H4); 4.73 (1H; dd; J₂,₃ = 10.9 Hz, J₃,₄ = 3.4 Hz; H3); 4.56 (1H; d; J₁,₂ = 7.6 Hz; H1); 4.16-4.02 (2H; m; H6 a, H6 b); 3.82 (1H; dd; J₅,₆ a = J₅,₆ b = 7.2 Hz, J₆ a, b = 1.1 Hz; H5); 3.58 (1H; dd; J₂,₃ = 10.9 Hz, J₃,₄ = 7.6 Hz; H2); 2.13, 2.02, 2.01 (9H; 3 x s; 3 x CH₃CO); 0.92 (9H; s; (CH₃)₃C-Si); 0.15 (6H; s; CH₃-Si).

¹³C-NMR (63 MHz, CDCl₃) δ (ppm): 170.4, 169.8 (CO); 97.4, 70.7, 70.6, 66.4, 63.0 (CH); 60.3 (CH₂); 25.5 (CH₃); 20.6 (CH₃, C(CH₃)₂); 18.0 (qC, SiC(CH₃)₃); -4.5 (CH₃, SiCH₃).

FAB-MS calculated for C₈H₃N₃O₇Si [MH]+ 446.1880, found 446.19587.

**Tert-butyldimethylsilyl-2-azido-2-deoxy-β-D-galactopyranoside,¹¹¹ [55]**

![Chemical Structure](image)

The acetyl protected glycoside 54 (4.8 g, 10.80 mmol) was dissolved in dry methanol (80 mL) to which 0.5 M NaOMe solution in methanol (560 µL, 0.28 mmol) was added. The reaction was allowed to stir for 20 h at room temperature, after which time, TLC (7:3 petroleum ether 40-60/ethyl acetate) indicated that the reaction was complete. The reaction mixture was neutralized using ion exchange resin, Amberlyst-15, filtered and evaporated under reduced pressure. The crude product was purified by flash chromatography over silica (the column was packed with DCM; the mobile phase was ethyl acetate) to afford 55 (3.3 g; 96 %) as a white solid. Rf = 0.35 (7:3 petroleum ether 40-60/ethyl acetate). ¹H-NMR (250 MHz, CDCl₃) δ (ppm): 4.51 (1H; d; J₁,₂ = 7.2 Hz; H1); 4.11 (1H; dd; J₂,₃ = 14.5 Hz, J₁,₂ = 7.2
7. Experimental section

Hz; H2); 3.98 (1H; d; J1,4= 2.8 Hz; H4); 3.91-3.83 (1H; m; H5); 3.47 (1H; dd; J2,3= 14.5 Hz, J3,4= 2.8 Hz; H3); 3.44-3.39 (2H; m; H6a, H6b); 0.92 (9H; s; (CH3)3C-Si); 0.15 (6H; s; CH3-Si). 13C-NMR (63 MHz, CDCl3) δ (ppm): 97.4, 74.0, 71.8, 68.5, 66.3 (CH2); 62.5 (CH2); 25.5 (CH3, C(CH3)3); 17.8 (qC, SiC(CH3)3); -4.4 (CH3, SiCH3). FAB-MS calculated for C18H31N4O5Si [MH]+ 320.1563, found 320.16418.

**Tert-butyldimethylsilyl-2-azido-2-deoxy-4, 6-O-benzylidene-β-D-galactopyranoside,21 [56]**

\[ \begin{align*}
55 & \quad \text{p-Toluene sulphonate acid (0.17 g; 0.91 mmol) was added to a stirred solution of the silyl derivative 55 (11.62 g, 36.43 mmol) and benzaldehyde dimethylacetal (15 mL, 102.0 mmol) in DMF (300 mL). The reaction mixture was allowed to stir at 50 °C for 4 h. After this time, TLC (1:1 petroleum ether 40-60/ethyl acetate) indicated that the reaction was complete. The reaction was allowed to cool and triethylamine (880 µL) was added. The solvent was evaporated under reduced pressure. The crude product was purified by flash chromatography over silica (1:1 petroleum ether 40-60/ethyl acetate) to afford 56 (13.2 g; 89 %) as a transparent oil. R= 0.50 (1:1 petroleum ether 40-60/ethyl acetate).} \\
56 & \quad \text{1H-NMR (250 MHz, CDCl3) δ (ppm): 7.43-7.20 (5H; m; ArH); 5.37 (1H; s; Ph-CH); 4.37 (1H; d; J1,4= 7.4 Hz; H1); 4.10 (1H; dd; J6a-6b= 12.4 Hz, J5,6a= 1.5 Hz; H6a); 3.99-3.95 (1H; m; H4); 3.89 (1H; dd; J6a-6b= 12.4 Hz, J5,6a= 2.0 Hz; H6a); 3.33-3.23 (3H; m; H2, H3,H5); 0.74 (9H; s; (CH3)3C-Si); 0.01 (6H; s; CH3-Si).} \\
57 & \quad \text{13C-NMR (63 MHz, CDCl3) δ (ppm): 137.3 (qC, Ar); 129.2,}
\end{align*} \]
7. Experimental section

128.2, 126.3 (Ar-H); 101.3, 97.1, 74.4 (CH); 69.0 (C\textsubscript{11}); 66.4, 66.3 (CH); 25.6 (\textit{CH}_{3}, C(CH\textsubscript{3})\textsubscript{3}); 17.9 (qC, SiC(CH\textsubscript{3})\textsubscript{3}); -4.2 (CH\textsubscript{3}, SiCH\textsubscript{3}). FAB-MS calculated for C\textsubscript{19}H\textsubscript{29}N\textsubscript{3}O\textsubscript{5}Si [MH\textsuperscript{+}] 408.1876, found 408.19548.

2, 3, 4, 6-Tetra-O-acetyl-\alpha-D-O-galactopyranosyl trichloroacetimidade\textsuperscript{120,243} [\alpha-23] (via 58)

The peracetylated galactose 49 (7.8 g, 20.0 mmol) was dissolved in anhydrous THF (120 mL) and benzylamine (2.6 mL, 24.0 mmol) was added under nitrogen. The reaction mixture was then stirred at 50 °C for 24 h under nitrogen. After that time TLC (4:1 ethyl acetate/petroleum ether 40-60) indicated the presence of both hemi-acetal and a little unreacted starting material. Then the reaction mixture was concentrated under reduced pressure. The crude product was purified by flash chromatography over silica (4:1 ethyl acetate/petroleum ether 40-60) to afford the anomeric deacetylated molecule 58 (6.2 g; 95 %) as a yellow oil. Rf= 0.46 and 0.30 (4:1 ethyl acetate/petroleum ether 40-60). ESI-MS: \(m/z=371.3\ [\text{MNa}]; \text{C}_{14}\text{H}_{20}\text{O}_{10}\) requires \(m/z=348.3\). A solution of 58 was dissolved in DCM (125 mL) and cooled to 0 °C, and trichloroacetonitrile (7 mL, 60 mmol) and DBU (3 mL, 20.0 mmol) were added. After 5 h stirring at 0 °C under N\textsubscript{2}, the reaction was concentrated to afford a solid. The crude product was purified by flash chromatography over silica (1:1 ethyl acetate/petroleum ether 40-60) to afford \(\alpha\text{-23}\) (8.5 g; 86 %) as a dark brown oil. Rf= 0.34 (2:1 ethyl acetate/petroleum ether 40-60). \textsuperscript{1}H-NMR (250 MHz, CDCl\textsubscript{3}) \(\delta\) (ppm): 8.65 (1H, s, NH); 6.59 (1H, d, \(J_{1,2}=3.2\ Hz, H1\)); 5.55 (1H, dd, \(J_{3,4}=2.9\ Hz, J_{4,5}=1.0\ Hz, H4\));
5.42 (1H, dd, J$_{2,3}$ = 10.8 Hz, J$_{3,4}$ = 2.9 Hz, H3); 5.35 (1H, dd, J$_{3,5}$ = 10.8 Hz, J$_{1,2}$ = 3.2 Hz, H2); 4.43 (1H, dt, J$_{6,7}$ = 7.2 Hz, J$_{5,6}$ = 1.0 Hz, H5); 4.16 (1H, dd, J$_{6a,6b}$ = 11.2 Hz, J$_{5,6a}$ = 7.2 Hz, H$_{6a}$); 4.07 (1H, dd, J$_{6a,6b}$ = 11.2 Hz, J$_{5,6b}$ = 6.8 Hz, H$_{6b}$); 2.29, 2.15, 2.14, 2.13 (12H, s, CH$_3$CO); 13C-NMR (63 MHz, CDCl$_3$) δ (ppm): 170.7, 170.5, 170.4, 161.3 (qC)); 93.9, 69.4, 67.9, 67.7, 67.3 (CH); 61.6 (CH$_2$); 21.0 (CH$_3$). ESI-MS: m/z = 515.7 [MNa$^+$]; C$_{16}$H$_{20}$NO$_{10}$Cl$_3$ requires m/z (35.5C$_{1}$) = 492.0.

2, 3, 4, 6-Tetra-O-acetyl-D-galactopyranosyl-(β-1, 3)- tert-butyldimethylsilyl-2-azido-2-deoxy-4, 6-O-benzylidene-β-D-galactopyranoside, $^{24}$[62]

The acceptor 56 (3.0 g, 7.30 mmol) and the trichloroacetimidate derivative 23 (7.25 g, 14.7 mmol) were dissolved in dry DCM (90 mL) containing molecular sieves 4Å, and were stirred at room temperature for 1 h under nitrogen. The mixture was then cooled to -18 °C and a solution of TMS-OTf (162 μL, 0.90 mmol) in DCM (12 mL) was added. After 1 h, the reaction mixture was allowed to warm to room temperature. After 1 h, TLC (1:1 petroleum ether 40-60/ethyl acetate) indicated that the reaction was complete. The reaction mixture was diluted with DCM (100 mL) and neutralized with solid NaHCO$_3$, then filtered, and the solvent was removed under reduced pressure. The crude product was purified by flash
chromatography over silica (1:1 petroleum ether 40-60/ethyl acetate) to afford 62 (4.29 g; 80 %) as an orange oil. Rf = 0.27 (1:1 petroleum ether 40-60/ethyl acetate). 1H-NMR (250 MHz, CDCl3) δ (ppm): 7.53-7.35 (5H; m; ArH); 5.53 (1H; s; Ph-CH); 5.38 (1H; dd; J1,2 = 3.5 Hz, J4,5 = 1.1 Hz; H4); 5.24 (1H; dd; J2,3 = 10.4 Hz, J1,2 = 8.0 Hz; H2); 5.01 (1H; dd; J2,3 = 10.4 Hz, J3,4 = 3.5 Hz; H3); 4.79 (1H; d; J1,2 = 8.0 Hz; H1); 4.53 (1H; dd; J1,2 = 7.6 Hz; H1'); 4.16-4.12 (6H; m; H5, H6a, H6b, H4', H6'a, H6'b); 3.88 (1H; dd; J5,6 = 6.5 Hz, J4,5 = 1.0 Hz; H5'); 3.71 (1H; dd; J2,3 = 10.7 Hz, J1,2 = 7.6 Hz; H2'); 3.42 (1H; dd; J2,3 = 10.7 Hz, J3,4 = 3.4 Hz; H3'); 1.99, 1.90, 1.87 and 1.81 (12H, s, CH3CO); 0.93 (9H; s; (CH3)3C-Si); 0.15 (6H; s; CH3-Si). 13C-NMR (63 MHz, CDCl3) δ (ppm): 170.2 (GO); 137.3 (qC, Ar); 128.8, 128.0, 126.1 (Ar-H); 102.2, 100.6, 97.4, 78.2, 74.9, 70.8, 70.6, 66.8, 66.5 (CH); 61.2 (CH2); 60.3 (CH); 25.5 (CH3, C(CH3)3); 20.6 (CH2); 17.9 (qC, SiC(CH3)3); 14.1 (CH3). ESI-MS: m/z = 760.8 [MNa]+; C33H47N3O14Si requires m/z = 737.2.

2, 3, 4, 6-Tetra-O-acetyl-D-galactopyranosyl-β-1, 3)-2-azido-4,6-O-benzylidene-2-azido-2-deoxy-D-galactopyranose,245 [63]

To a mixture of the protected disaccharide 62 (1.59 g, 2.16 mmol) and acetic acid (1.2 mL, 21.60 mmol) in anhydrous THF (65 mL) was added TBAF (1M in THF) (8.6 mL, 8.63 mmol). The mixture was stirred at room temperature. After 15 h, TLC (ethyl acetate) indicated that the reaction was complete. The solvent was removed under reduced pressure. The crude was diluted with ethyl acetate (100 mL) and washed with water (100 mL) and brine (100 mL); the organic phase was dried with MgSO4, filtered and the solvent was
removed under vacuum. The crude product was purified by flash chromatography over silica (the column was packed with 6:4 ethyl acetate/petroleum ether 40-60; the mobile phase was ethyl acetate) to afford a white solid 63 (1.03 g; 76 %). $R_f = 0.15$ and 0.07 (6:4 ethyl acetate/petroleum ether 40-60). $^1$H-NMR (250 MHz, CDCl$_3$), anomeric $\alpha/\beta$ mixture. ESI-MS: $m/z = 641.1$ [M$\text{NH}_4^+$]; $m/z = 646.6$ [M$\text{Na}^+$]; C$_{27}$H$_{33}$N$_3$O$_{14}$ requires $m/z = 623.2$.

2, 3, 4, 6-Tetra-$O$-acetyl-$D$-galactopyranosyl-($\beta$-1, 3)-1, 4, 6-tri-$O$-acetyl-2-azido-2-deoxy-$D$-galactopyranose,$^{246}$ [66]

![Chemical structure](image)

The disaccharide 63 (1.54 g, 2.47 mmol), was dissolved in 80 % acetic acid aqueous solution (250 mL) and the solution was heated at 80 °C for 2 hours. Then the reaction was concentrated, and the resulting product was dissolved in a solution of pyridine/acetic anhydride (2:1) (250 mL). After stirring for 16 h the reaction was concentrated to afford a brown oil (2.3 g). The crude product was purified by flash chromatography over silica (the column was packed with 3:1 petroleum ether 40-60/ethyl acetate; the mobile phase was gradually changed to 1:3 petroleum ether 40-60/ethyl acetate) to afford the purified $\beta$-product $\beta$-66 ($R_f = 0.50$) and $\alpha$-product $\alpha$-66 ($R_f = 0.42$), (ratio of anomeric $\alpha/\beta$ mixture (7:3), (total 1.6 g; 100 %), (3:1 ethyl acetate/petroleum ether 40-60). $^1$H-NMR (250 MHz, CDCl$_3$) $\delta$ (ppm): $\beta$-anomer: 5.40 (1H, d, $J_{1:2} = 8.3$ Hz, H1'); 5.34 (1H, d, $J_{3:4} = 3.3$ Hz, H4'); 5.30 (1H, d, $J_{3:4} = 2.2$ Hz, H4); 5.11 (1H, dd, $J_{2:3} = 10.5$ Hz, $J_{1:2} = 7.7$ Hz, H2); 4.96 (1H, dd, $J_{2:3} = 10.5$ Hz, $J_{3:4} = 2.2$ Hz, H3); 4.65 (1H, d, $J_{1:2} = 7.7$ Hz, H1); 4.20-3.83 (6H, m, H5, H5', H6b, H6a, H6b', H6a'); 3.66 (1H, dd, $J_{2:3} = 10.5$ Hz, $J_{1:2} = 8.3$ Hz, $J_{2:4} = 0.9$ Hz, H2'); 3.58 (1H, dd,
7. Experimental section

$J_{2,3} = 10.3$ Hz, $J_{3,4} = 3.4$ Hz, H3'); 2.14, 2.02, 1.99, 1.97, 1.95, 1.93, 1.92 (21H, s, CH3CO).

$^{13}$C-NMR (63 MHz, CDCl3) δ (ppm): β-anomer: 169.8, 169.6, 169.5, 169.3, 168.8, 168.6, 168.0 (CO); 100.7, 92.0, 76.9, 71.7, 70.1, 69.9, 68.0, 67.1, 66.0, 61.2 (CH); 60.3 (CH2); 20.2, 19.9, 19.8 (CH3). ¹H-NMR (250 MHz, CDCl3) δ (ppm): α-anomer: 6.29 (1H, d, $J_{1,2} = 3.7$ Hz, H1'); 5.49 (1H, d, $J_{3,4} = 2.5$ Hz, H4'); 5.35 (1H, dd, $J_{3,4} = 3.4$ Hz, $J_{4,5} = 1.1$ Hz, H4); 5.17 (1H, dd, $J_{2,3} = 10.5$ Hz, $J_{1,2} = 7.7$ Hz, H2); 4.99 (1H, dd, $J_{3,4} = 10.5$ Hz, $J_{4,5} = 3.4$ Hz, H3); 4.72 (1H, d, $J_{2,3} = 7.8$ Hz, H1); 4.23-3.84 (6H; m; H5, H5', H6, H6', H6'a, H6'b); 3.98 (1H, dd, $J_{2,3} = 10.6$ Hz, $J_{3,4} = 2.5$ Hz, H3'); 3.87 (1H, dd, $J_{2,3} = 10.6$ Hz, $J_{1,2} = 3.7$ Hz, H2'); 2.16, 2.14, 2.13, 2.05, 2.02, 1.99, 1.97, (21H, s, CH3CO). ¹³C-NMR (63 MHz, CDCl3) δ (ppm): α-anomer: 170.4, 170.3, 170.1, 170.0, 169.4, 169.3, 168.5 (CO); 101.2, 90.5, 74.7, 70.6, 69.2, 68.6, 66.5 (CH); 62.0, 60.8 (CH2); 58.8 (CH); 20.8, 20.5 (CH3). ESI-MS: $m/z = 679.2$ [MNH₄]⁺; $m/z = 684.2$ [MNa]⁺; C₂₆H₃₅N₃O₁₇ requires $m/z = 661.2$.

2, 3, 4, 6-Tetra-O-acetyl-D-galactopyranosyl-(β-1, 3)-4, 6-di-O-acetyl-2-azido-1-bromo-1, 2-dideoxy-α-D-galactopyranose,²⁴⁶ [69]

A solution of the acetylated disaccharide 66 (1.73 g, 2.62 mmol) in anhydrous DCM (35 mL) was stirred for 12 h in the presence of TiBr₄ (2.39 g, 6.50 mmol). Then the mixture was diluted with DCM (150 mL), and washed with ice water (2 x 200 mL), then was dried with MgSO₄, filtered and the solvent was removed under vacuum to afford a brown syrup product 69 (1.62 g; 91 %). Rf = 0.51 (1:2 petroleum ether 40-60/ethyl acetate). ¹H-NMR (250 MHz, CDCl₃) δ (ppm): 6.46 (1H, d, $J_{1,2} = 3.8$ Hz, H1'); 5.52 (1H, dd, $J_{3,4} = 3.1$ Hz, $J_{4,5} = 0.91$ Hz, 150
7. Experimental section

H4'); 5.35 (1H, dd, J3,4 = 3.3 Hz, J4,5 = 1.0 Hz, H4); 5.18 (1H, dd, J2,3 = 10.5 Hz, J1,2 = 7.7 Hz, H2); 4.97 (1H, dd, J1,2 = 10.5 Hz, J3,4 = 3.4 Hz, H3); 4.72 (1H, d, J1,2 = 7.8 Hz, H1); 4.38-4.05 (6H; m; H5, H5', H6a, H6b, H6'a, H6'b); 4.00 (1H, dd, J2,3 = 10.2 Hz, J3,4 = 3.1 Hz, H3'); 3.89 (1H, dd, J2,3 = 10.2 Hz, J1,2 = 3.8 Hz, H2'); 2.14, 2.12, 2.04, 2.03, 1.96 (18H, s, CH3CO).

13C-NMR (63 MHz, CDCl3) δ (ppm): 170.4, 170.4, 170.2, 170.0, 169.5, 169.3 (CO); 101.2, 90.0, 75.7, 72.4, 71.0, 70.7, 68.8, 68.3, 66.8, 61.7 (CH); 61.1, 60.8 (CH2); 20.6, 20.5 (CH3).

ESI-MS: m/z = 700.4 [MNH4]+; m/z = 705.4 [MNa]+; C24H32BrN3O15 requires m/z = 682.4.

A solution of the protected amino acid 46 (1.53 g, 3.55 mmol) in anhydrous DCM (18 mL) containing 4Å molecular sieves, 2,4,6-collidine (1.2 mL, 8.88 mmol) and AgClO4 (1.84 g, 8.88 mmol), was stirred under argon at room temperature. The disaccharide 69 (2.42 g, 3.55 mmol) was dissolved in anhydrous DCM (18 mL) under argon, and added dropwise to the stirring mixture. The reaction was allowed to stir overnight, after which time TLC (6:4 hexane/ethylacetate) indicated that the reaction was complete. The mixture was filtered through Celite and the solvent was removed under vacuum. The crude product was purified by flash chromatography over silica (the column was packed with 2:1 hexane/ethyl acetate; the mobile phase increased polarity through the chromatography) to afford the purified
product 77 (1.66 g; 45%), $R_f = 0.6$ (3:1 ethyl acetate/hexane); and 2,3,4,6-tetra-O-acetyl-D-galactopyranosyl-(β1,3)-4,6-di-O-acetyl-2-azido-2-deoxy-D-galactopyranose 78 (0.48 g; 22 %), $R_f = 0.15$ (3:1 ethyl acetate/hexane). $^1$H-NMR (250 MHz, CDCl$_3$) $\delta$ (ppm) 77: 7.75 (2H, d, $J_{A-B} = 7.5$ Hz, Ar-H$_{Fme}$); 7.58 (2H, d, $J_{C-D} = 7.5$ Hz, Ar-H$_{Fme}$); 7.41-7.27 (9H, m, Ar-H$_{Fme}$ and Ar-H$_{phenyl}$); 5.65 (1H, d, $J = 9.5$ Hz, NH); 5.45 (1H, d, $J_{J_{3,4}} = 2.1$ Hz, H$^4$); 5.35 (1H, dd, $J_{3,4} = 3.4$ Hz, $J_{4,5} = 1.0$ Hz, H$^5$); 5.21 (2H, s, -CO$_2$-CH$_2$-Ph); 5.16 (1H, dd, $J_{2,3} = 10.4$ Hz, $J_{3,4} = 7.9$ Hz, H$^2$); 5.00 (1H, dd, $J_{2,3} = 10.4$ Hz, $J_{3,4} = 3.4$ Hz, H$^3$); 4.83 (1H, d, $J = 3.8$ Hz, H$^1$); 4.68 (1H, d, $J = 7.9$ Hz, H$^1$); 4.51-3.88 (12 H, m, H$^3$, H$^5$, H$^5'$, H$^6$, H$^6'$, H$^6''$, H$^6'''$, H$^6''''$, CH$_3$CO); 1.30 (3H, d, $J = 6.2$ Hz, CH$_3$). $^{13}$C-NMR (63 MHz, CDCl$_3$) $\delta$ (ppm) 77: 170.3, 170.0, 169.5 (CO); 157.1 (CONH); 156.7 (COO); 143.8, 143.6, 141.2 (Ar); 128.7, 128.5, 127.7, 125.0, 121.3, 120.0, 119.9 (Ar-H); 101.4, 99.2, 74.8, 70.7, 69.3, 68.7, 67.8 (CH); 67.3 (CH$_2$); 66.7 (CH); 62.8, 60.9 (CH$_2$); 59.6 (CH); 58.7, 47.0 (CH); 23.9 (CH$_3$); 20.6, 20.5, 18.4 (CH$_3$). ESI-MS 77: $m/z = 1050.0$ [MNH$_4^+$], $m/z = 1055.0$ [MNa$^+$]; C$_{50}$H$_{56}$N$_4$O$_{20}$ requires $m/z = 1033.0$.

2, 3, 4, 6-Tetra-O-acetyl-D-galactopyranosyl-(β1, 3)-1, 4, 6-tri-O-acetyl-2-azido-2-deoxy-D-galactopyranose,$^{246}$ [66]

The disaccharide 78 (0.68 g, 1.10 mmol) was dissolved in pyridine (10 mL) and the mixture was cooled down to 0 °C. Acetic anhydride (5 mL) was added dropwise. After the addition was complete the addition, the mixture was allowed to warm to room temperature and stirred continuously overnight. After stirring for 16 h the reaction was concentrated under vacuum.
7. Experimental section

The crude was purified by flash chromatography over silica (packed with 3:1 petroleum ether 40-60/ethyl acetate; the mobile phase was gradually changed to 1:3 petroleum ether 40-60/ethyl acetate) to afford the epimeric mixture (ratio of anomeric α:β mixture (7:3) 66 (0.63 g; 90%). $^1$H-NMR (250 MHz, CDCl$_3$) δ (ppm): β-anomer: 5.40 (1H, d, $J_{1,2}$ = 8.3 Hz, H1');
5.34 (1H, d, $J_{3,4}$ = 3.3 Hz, H4'); 5.30 (1H, d, $J_{3,4}$ = 2.2 Hz, H4); 5.11 (1H, dd, $J_{2,3}$ = 10.5 Hz, $J_{1,2}$ = 7.7 Hz, H2); 4.96 (1H, dd, $J_{2,3}$ = 10.5 Hz, $J_{3,4}$ = 2.2 Hz, H3); 4.65 (1H, d, $J_{1,2}$ = 7.7 Hz, H1); 4.20-3.83 (6H, m, H5, H5', H6a, H6b, H6'a, H6'b); 3.66 (1H, dd, $J_{2,3}$ = 10.5 Hz, $J_{1,2}$ = 8.3 Hz, $J_{2,3}$ = 0.9 Hz, H2'); 3.58 (1H, dd, $J_{2,3}$ = 10.3 Hz, $J_{3,4}$ = 3.4 Hz, H3'); 2.14, 2.02, 1.99, 1.97, 1.95, 1.93, 1.92 (21H, s, CH$_3$CO). 13C-NMR (63 MHz, CDCl$_3$) δ (ppm): β-anomer:
169.8, 169.6, 169.5, 169.3, 168.8, 168.6, 168.0 (CO); 100.7, 92.0, 76.9, 71.7, 70.1, 69.9, 68.0, 67.1, 66.0, 61.2 (CH); 60.3 (CH$_2$); 20.2, 19.9, 19.8 (CH$_3$). $^1$H-NMR (250 MHz, CDCl$_3$) δ (ppm): α-anomer: 6.29 (1H, d, $J_{1,2}$ = 3.7 Hz, H1'); 5.49 (1H, d, $J_{3,4}$ = 2.5 Hz, H4');
5.35 (1H, dd, $J_{3,4}$ = 3.4 Hz, $J_{4,5}$ = 1.1 Hz, H4); 5.17 (1H, dd, $J_{2,3}$ = 10.5 Hz, $J_{1,2}$ = 7.7 Hz, H2); 4.99 (1H, dd, $J_{2,3}$ = 10.5 Hz, $J_{3,4}$ = 3.4 Hz, H3); 4.72 (1H, d, $J_{1,2}$ = 7.8 Hz, H1); 4.23-3.84 (6H; m; H5, H5', H6a, H6b, H6'a, H6'b); 3.98 (1H, dd, $J_{2,3}$ = 10.6 Hz, $J_{3,4}$ = 2.5 Hz, H3'); 3.87 (1H, dd, $J_{2,3}$ = 10.6 Hz, $J_{1,2}$ = 3.7 Hz, H2'); 2.16, 2.14, 2.13, 2.05, 2.02, 1.99, 1.97, 21H, s, CH$_3$CO). 13C-NMR (63 MHz, CDCl$_3$) δ (ppm): α-anomer: 170.4, 170.3, 170.1, 170.0, 169.4, 169.3, 168.5 (CO); 101.2, 90.5, 74.7, 70.6, 69.2, 68.6, 66.5 (CH); 62.0, 60.8 (CH$_2$); 58.8 (CH); 20.8, 20.5 (CH$_3$). ESI-MS: $m/z$=679.2 [MNH$_4$]$^+$; $m/z$ = 684.2 [MNa]$^+$; C$_{26}$H$_{35}$N$_3$O$_{17}$ requires $m/z$ = 661.2.
N\textsuperscript{\alpha}-(Fluorene-9-nylmethoxycarbonyl)-O-[2, 3, 4, 6-tetra-O-acetyl-D-galactopyranosyl-(\beta-1,3)-4,6-di-O-acetyl-2-acetamido-2-deoxy-\alpha-D-galactopyranosyl]-L-threonine benzylester, [81]

The glycoamino acid 77 (0.22 g, 0.23 mmol) was dissolved in THF/acetone/acetic acid (3:2:1) (2.6 mL). Zinc dust (0.2 g) was added followed by a saturated aqueous solution of CuSO\textsubscript{4} (38 μL). After 3 h, the reaction mixture was filtered through Celite and the Celite pad was washed with THF. The filtrate was reduced under vacuum to afford the crude product as a yellow oil. The crude product was purified by flash chromatography over silica (3:1 ethyl acetate/petroleum ether 40-60) to afford the purified product 81 (0.2 g; 95%) as an ambar oil, R\textsubscript{f} = 0.26 (3:1 ethyl acetate/petroleum ether 40-60).

\textsuperscript{1}H-NMR (250 MHz, CDCl\textsubscript{3}) δ (ppm): 7.71 (2H, d, J\textsubscript{AB} = 7.1 Hz, Ar-H\textsubscript{Fmoc}); 7.54 (2H, d, J\textsubscript{CD} = 7.1 Hz, Ar-H\textsubscript{Fmoc}); 7.29-7.25 (9H, m, Ar-H\textsubscript{Fmoc} and Ar-H\textsubscript{phenyl}); 5.95 (1H, d, J = 9.2 Hz, AcNH); 5.80 (1H, d, J = 9.3 Hz, NH); 5.32-5.11 (4H, m, H\textsubscript{4}, H\textsubscript{4'}, -CO\textsubscript{2}-CH\textsubscript{2}-Ph); 5.07 (1H, dd, J\textsubscript{2,3} = 10.4 Hz, J\textsubscript{1,2} = 7.9 Hz, H2); 4.91 (1H, dd, J\textsubscript{2,3} = 10.4 Hz, J\textsubscript{1,2} = 3.2 Hz, H3); 4.79 (1H, d, J\textsubscript{1,2} = 3.6 Hz, H1\textsuperscript{1}); 4.55 (1H, d, J\textsubscript{1,2} = 7.8 Hz, H1); 4.47 (2H, d, J = 6.7 Hz, CH\textsubscript{2}CH\textsubscript{2}O); 4.39 (1H, dd, J\textsubscript{2,3} = 11.0 Hz, J\textsubscript{1,2} = 3.6 Hz, H1\textsuperscript{1}); 4.30 (1H, t, J = 6.7 Hz, CH\textsubscript{2}CH\textsubscript{2}O); 4.24-3.83 (8H, m, H\textsubscript{5}, H\textsubscript{5'}, H\textsubscript{6}, H\textsubscript{6'}, H\textsubscript{6a}, H\textsubscript{6b}, H\textsubscript{6c}, H\textsubscript{6d}, H\textsubscript{6e}, 2H\textsubscript{3H}); 3.79 (1H, dd, J\textsubscript{2,3} = 11.0 Hz, J\textsubscript{1,2} = 2.7 Hz, H3\textsuperscript{1}); 2.35, 2.32, 2.27, 2.24, 2.23, 2.19 (21 H, s, CH\textsubscript{3}CO); 1.45 (3H, d, J = 6.2 Hz, CH\textsubscript{3}). \textsuperscript{13}C-NMR (63 MHz, CDCl\textsubscript{3}) δ (ppm): 170.5, 170.3, 170.2, 170.0, 169.9, 169.8, 169.4 (CO); 156.3 (CONH); 143.5, 143.4, 141.2 (Ar); 134.2 (COO); 128.8, 128.7, 128.4, 127.7, 127.0, 124.7, 120.0 (Ar-H); 100.4, 99.8, 77.1, 76.7, 72.9, 70.5, 68.5, 68.4, 67.7 (CH); 67.6 (CH\textsubscript{2}); 66.5 (CH); 62.8, 60.9 (CH\textsubscript{3});
7. Experimental section

58.4, 48.4, 47.0 (CH); 23.2 (CH$_3$); 20.5, 18.2 (CH$_2$). ESI-MS: $m/z = 1066.4$ [MNH$_4$]$^+$, $m/z = 1071.4$ [MNa]$^+$; C$_{52}$H$_{60}$N$_2$O$_{21}$ requires $m/z = 1048.4$

$N^\alpha$-(Fluorene-9-nymethoxycarbonyl)-O-[2, 3, 4, 6-tetra-O-acetyl-D-galactopyranosyl-(β-1,3)-4, 6-diacetyl-2-acetamido-2-deoxy-α-D-galactopyranosyl]-L-threonine$^{113}$[82]

The glycoamino acid 81 (1.8 g, 1.72 mmol) and 10 % Pd/C (0.1 g) were taken up in anhydrous methanol (50 mL). The reaction mixture was then catalytically hydrogenated under a hydrogen atmosphere for 3 h at room temperature. The catalyst was then removed by filtration through Celite and the Celite pad was washed with methanol. The filtrate was evaporated under vacuum to afford 82 (1.57 g; 95 %) as a white solid. $^1$H-NMR (250 MHz, CDCl$_3$, TFA 1 %) δ (ppm): 7.75 (2H, d, $J_{A,B}$ = 7.4 Hz, Ar-H$_{Fmoc}$); 7.48 (2H, dd, $J_{C,D}$ = 7.14 Hz, Ar-H$_{Fmoc}$); 7.41 (2H, dd, $J_{C,D}$ = 7.5 Hz, Ar-H$_{Fmoc}$); 7.32 (2H, dd, $J_{A,B}$ = 7.4 Hz, Ar-H$_{Fmoc}$); 7.15 (1H, d, $J$ = 10.2, AcNH); 6.80 (1H, d, $J$ = 9.7 Hz, NH); 5.46 (1H, d, $J_{1,2}$ = 2.6 Hz, H1'); 5.40 (1H, d, $J_{3,4}$ = 3.0 Hz, H4); 5.12 (1H, d, $J_{3,4}$ = 3.0 Hz, H4'); 5.09 (1H, dd, $J_{2,3}$ = 10.5 Hz, $J_{1,2}$ = 7.9 Hz, H2); 4.95 (1H, dd, $J_{2,3}$ = 10.5 Hz, $J_{3,4}$ = 3.0 Hz, H3); 4.68 (2H, m, CHCH$_2$O); 4.56 (1H, d, $J_{1,2}$ = 7.6 Hz, H1); 4.36-4.01 (8H, m, 2 x CH-Thr, H3', H6a, H6b, H6b', and CHCH$_2$O); 3.90 (1H, dd, H2'); 3.86 (1H, dd, H5'); 3.82 (1H, dd, H5); 2.17, 2.16, 2.10, 2.08, 2.05 and 2.00 (21H, s, CH$_3$CO); 1.09 (3H, d, $J$ = 6.3 Hz, CH$_3$). $^{13}$C-NMR (63 MHz, CDCl$_3$, 1 % TFA) δ (ppm): 171.3, 170.3, 170.0, 169.5, 169.0 (CO); 157.1 (CO$_2$H); 143.3, 143.2, 140.6 (Ar); 126.9, 126.3, 124.2, 124.1, 119.1 (Ar-H); 100.4, 98.9, 75.3, 73.1,
70.1, 69.6, 69.1, 68.1, 66.9, 66.5 (CH); 65.7, 62.4, 60.3 (CH$_2$); 57.8, 48.0, 46.5 (CH); 21.3, 18.9, 18.8, 18.6, 17.2 (CH$_3$). ESI-MS: $m/z = 976.9$ [MNH$_4^+$], $m/z = 981.9$ [MNa$^+$]; $C_{45}H_{44}N_{21}O_{21}$ requires $m/z = 958.9$.

**Synthesis of compound 83**

![Compound 83](image)

*Solid phase glycopeptide synthesis.*

Solid-phase peptide synthesis was carried out manually using a plastic syringe fitted with a Teflon filter and connected to a vacuum waste chamber via a Teflon valve. Synthesis of the glycopeptide (CSSQLET*$_2$S) was carried out on pre-loaded Fmoc-Ser(tBu)-NovaSyn TGT resin (loading = 0.22 mmol/g). The scale of the synthesis of the glycopeptide was 0.10 mmol. The Fmoc group was removed by treatment with 20 % piperidine in DMF (5 and 15 min) between each coupling and deprotection step. Coupling reactions with Fmoc amino acids were achieved using 1.0 mmol (10 equiv.) of each of HBTU/HOBt and DIPEA in DMF, for 3 h. The coupling reaction with Fmoc-(Core-1)-Thr, 82, was achieved using 0.15 mmol (1.5 equiv.) of the glyco-building block and DCC/HOBt for 24 h in DMF. The reaction progress was monitored using LCMS and the Kaiser ninhydrin test. Between each coupling and deprotection step, the resin was washed with DCM and DMF (5 min each).
The resin was washed with DCM and DMF five times (5 min each wash). The glycopeptide was cleaved from the resin by incubation with the cleavage cocktail (95 % TFA, 2.5 % water and 2.5 % EDT) for 3 h. The resin was subsequently washed twice with the cleavage cocktail (10 min each wash). The glycopeptide was precipitated with cold diethyl ether and centrifuged for 10 min. The precipitate was dried and redissolved in a solution 1:1 water/acetonitrile and lyophilised to afford a fluffy white solid.

Deprotection of the acetyl groups.

The fluffy white solid was dissolved in a 1 mL phosphate buffer (0.1M, pH8 solution); with 5 % hydrazine hydrate and 10 % MESNA. The solution was stirred gently for 16 h at room temperature. The crude product was purified by reverse phase semi preparative HPLC (water-acetonitrile 5 %-95 % in 40 minutes, 0.1 % TFA). The fractions were freeze-dried to afford 83 (28 mg; 29 %) as a fluffy white solid.

Analytical data for compound 83: C_{51}H_{85}N_{11}O_{29}S. MW: 1347.52. Found: [M]^+ 1348.00.
2-Phthalimido-2-deoxy-1, 3, 4, 6-tetra-0-acetyl-β-D-glucopyranose, 92 [85]

To a stirred solution of sodium methoxide (3.0 g, 46.30 mmol) in anhydrous methanol (75 mL) was added glucosamine hydrochloride 84 (10.0 g, 46.30 mmol). The reaction mixture was stirred for 30 min at room temperature and then filtered with suction. Phthalic anhydride (3.5 g, 23.0 mmol) was then added to the filtrate and stirring was continued for a further 20 min. A further portion of phthalic anhydride (3.5 g, 23.00 mmol) was then added followed by triethylamine (7.6 mL, 55.60 mmol). The reaction mixture was stirred at room temperature for 10 min and then cooled for 1 h in an ice bath and then filtered with suction. The precipitate was washed with cold methanol (2 x 20 mL) and dried under high vacuum. The dry white solid was then suspended in acetic anhydride (45 mL) and cooled down to 0 °C, and then pyridine (22.7 mL) was added carefully with stirring. The reaction was then stirred at room temperature for 16 h. After this time, the reaction mixture was poured into ice/water (200 mL) and extracted with chloroform (3 x 200 mL). The combined organic extracts were washed with 5 % aqueous HCl (1 x 120 mL), saturated aqueous NaHCO₃ (120 mL), water (120 mL) and brine (100 mL). The organic phase was dried with MgSO₄, filtered and concentrated under reduced pressure to afford an orange oil. The crude product was purified by flash chromatography over silica (1:1 hexane/ethyl acetate) to afford the pure product 85 (6.6 g; 30%) as a white foam. Rf = 0.56 (4:1 ethyl acetate/petroleum ether 40-60). ¹H-NMR (250 MHz, CDCl₃); δ (ppm): 7.74-7.62 (4H, m, ArH); 6.35 (1H, d, J₁,₂ = 8.9 Hz, H₁); 5.73 (1H, dd, J₂,₃ = J₃,₄ = 9.8 Hz, H₃); 5.05 (1H, dd, J₄,₅ = J₅,₆ = 9.8 Hz, H₄); 4.34 (1H, dd, J₆ₐ,₆ₐ = 11.6 Hz, J₅,₆ₐ = 3.3 Hz, H₆ₐ); 4.30 (1H, dd, J₁,₂ = 8.9 Hz, J₂,₃ = 9.8 Hz, H₂); 3.98 (1H, dd, J₆ₐ,₆ₐ = 11.6 Hz, J₅,₆ₐ = 4.2 Hz, H₆ₐ); 3.91 (1H, ddd, J₄,₅ = 9.8 Hz, J₅,₆ₐ = 3.3 Hz,
7. Experimental section

$J_{5,6b} = 4.2$ Hz, H5); 1.92, 1.88, 1.81 and 1.70 (12H, 4 x s, CH$_3$CO). $^{13}$C-NMR (63 MHz, CDCl$_3$) $\delta$ (ppm): 169.8, 169.3, 168.8, 167.9, 166.7 (CO); 134.0 (Ar-H); 130.6 (Ar); 123.2 (Ar-H); 89.1, 72.0, 69.8, 67.7 (CH); 60.9 (CH$_3$); 52.9 (CH); 20.0 (CH$_3$). FAB-MS calculated for C$_{22}$H$_{23}$NO$_{11}$: 476.9, found 494.9 [M+NH$_4$]$^+$.  

3, 4, 6-Tri-O-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl azide, $^{183}$ [86]

![Chemical diagram]

To a suspension of 1, 3, 4, 6-tetra-O-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranose 85 (2.36 g, 4.82 mmol) in DCM (25 mL) were added trimethylsilyl azide (770 µL, 5.78 mmol) and tin tetrachloride (280 µL, 2.41 mmol) and the mixture was stirred for 24 h. After that time TLC (3:1 ethyl acetate/petroleum ether 40-60) indicated that the reaction was complete. The reaction was diluted with DCM (50 mL) and washed with saturated aqueous NaHCO$_3$ (40 mL) and water (40 mL). The organic phase was dried over MgSO$_4$, filtered and concentrated under reduced pressure. The crude product was dissolved in the minimum volume of DCM, and precipitation was initiated by the addition of methanol to afford 86 (1.61 g; 73 %) as a white crystalline solid. Rf = 0.36 (3:1 ethyl acetate/petroleum ether 40-60). $^1$H-NMR (250 MHz, CDCl$_3$): $\delta$ (ppm): 7.88-7.84 (2H, q, $J_{4,5} = 5.5$ Hz, 2H); 7.76-7.73 (2H, q, $J_{5,6} = 5.4$ Hz, 2H); 5.79 (1H, dd, $J_{2,3} = 10.6$ Hz, $J_{3,4} = 9.2$ Hz, 1H); 5.64 (1H, d, 2, 9.5 Hz, 1H); 5.18 (1H, dd, $J_{2,3} = 10.6$ Hz, $J_{3,4} = 9.2$ Hz, 1H); 4.38-4.16 (3H, m, H-2, 2H); 3.97 (1H, m, H5); 2.12, 2.03 and 1.85 (9H, 3 x s, CH$_3$CO). $^{13}$C-NMR (63 MHz, CDCl$_3$) $\delta$ (ppm): 170.6, 170.0 and 169.4 (CO); 134.5 (Ar-H);
131.2 (Ar); 123.7 (Ar-H); 85.5, 73.9, 70.3, 68.4 (CH); 61.7 (CH₂); 53.9 (CH); 20.7, 20.5, 20.3 (CH₃). FAB-MS calculated for C₂₀H₂₀N₄O₉ [MNH₄]⁺ 478.1230, found 478.1348.

6-O-Tert-butyldiphenylsilyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl azide,¹⁸³

The glycosyl azide 86 (1.61 g, 3.50 mmol) was dissolved in methanol (24 mL). A 0.5 M solution of NaOMe in methanol was added until a pH of 10 was obtained. After the reaction was stirred for 16 h, it was neutralized by addition of acetic acid. The mixture was then concentrated and azeotroped with toluene before being placed under high vacuum for 1 h. The resulting white solid was dissolved in DCM (12 mL) to which was added DIPEA (1.2 mL, 7.0 mmol) and DMAP (43 mg, 0.35 mmol). TBDPSCI (1 mL, 3.85 mmol) was then added, and the reaction was stirred for 16 h. After that time TLC (4:1 ethyl acetate/petroleum ether 40-60) indicated that the reaction was complete. The reaction mixture was washed with water (2 x 20 mL) and brine (2 x 20 mL). The organic phase was dried with MgSO₄, filtered and concentrated under reduced pressure. The crude product was purified by flash chromatography over silica (5 % methanol in DCM) to afford 87 (1.87 g; 93 %) as a transparent oil. Rf = 0.56 (4:1 ethyl acetate/petroleum ether 40-60). ¹H-NMR (250 MHz, CDCl₃): δ (ppm): 7.76-7.61 (9 H, m, ArH); 7.43 (5 H, m, ArH); 5.35 (1 H, d, J₁,₂ = 9.5 Hz, H₁); 4.40 (1 H, dd, J₁,₄ = 10.5 Hz, J₂,₃ = 7.8 Hz, H₃); 4.05 (1 H, dd, J₁,₂ = 9.5 Hz, J₂,₃ = 7.8 Hz, H₂); 3.97 (1 H, m, H₅); 3.70 (1 H, m, H₄); 3.65 (2 H, m, H₆a, H₆b); 1.09 (9 H, s, C(CH₃)₃). ¹³C-NMR (63 MHz, CDCl₃) δ (ppm): 168.2 (CO); 135.5 (Ar-H); 134.2 (Ar);
132.8 (Ar-H); 131.3 (Ar); 129.8, 127.7, 123.5 (Ar-H); 85.3, 77.2, 72.5, 71.2 (CH); 63.8 (CH$_2$); 55.8 (CH); 26.7 (CH$_3$, C(CH$_3$)$_3$); 19.1 (qC, C(CH$_3$)$_3$). FAB-MS calculated for C$_{30}$H$_{32}$N$_4$O$_6$Si [MH]$^+$ 573.2091, found 573.2169.

$O$-(2, 3, 4, 6-Tetra-$O$-acetyl-D-galactopyranosyl)-(β-1, 4)-6-$O$-tert-butyldiphenylsilyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl azide,$^{181}$ [88]

![Chemical Structure](image)

The trichloroacetimidate derivative 23 (4.69 g, 9.50 mmol) and the silylated sugar 87 (2.72 g, 4.75 mmol) were dissolved in dry DCM (25 mL), containing 4Å molecular sieves, and the solution was cooled down to -20 °C. A solution of 0.1 M TMS-OTf in dry DCM (4.8 mL, 0.48 mmol) was added. After 2 h, the reaction was warmed to room temperature. After 1 h, TLC (8:2 toluene/ethyl acetate) indicated that the reaction was finished. The reaction mixture was diluted with DCM (50 mL) and neutralized with solid NaHCO$_3$, then was filtered and the solvent was removed under reduced pressure. The crude product was purified by flash chromatography over silica (4:1 toluene/ethyl acetate) to afford 88 (3.93 g; 92 %) as a white foam. $R_f$ = 0.26 (4:1 toluene/ethyl acetate). $^1$H-NMR (250 MHz, CDCl$_3$): δ (ppm): 7.66-7.35 (10H, m, 2 x C$_6$H$_5$); 7.16-7.06 (4H, m, Pht-H); 5.36 (1H, d, $J_{1-2}$ = 9.5 Hz, H1'); 5.27 (1H, d, $J_{3-4}$ = 3.3 Hz, H4); 5.15 (1H, dd, $J_{23}$ = 10.4 Hz, $J_{1-2}$ = 8.1 Hz, H2); 4.90 (1H, dd, $J_{23}$ = 10.4 Hz, $J_{1-2}$ = 8.1 Hz, H1'); 4.65 (1H, d, $J_{1-2}$ = 8.1 Hz, H1); 4.42 (1H, dd, $J_{23}$ = 10.4 Hz, $J_{x-4}$ = 8.5 Hz, H3'); 4.06 (1H, dd, $J_{23}$ = 10.4 Hz, $J_{1-2}$ = 9.5, H2'); 4.00-3.45 (7H, m, H5, H6$_a$, H6$_b$, H4', H5', H6$_a$, H6$_b$); 2.26, 2.04, 1.90, 1.65 (12 H, 4 x s, COCH$_3$); 1.04 (9H, s, C(CH$_3$)$_3$). $^{13}$C-NMR (63 MHz, CDCl$_3$) δ (ppm): 170.4, 170.0, 169.9, 169.2, 168.1, 167.5

161
7. Experimental section

(CO); 135.4 (Ar-H); 134.0 (Ar); 132.8 (Ar-H); 131.3 (Ar); 129.7, 127.7, 123.4 (Ar-H); 100.9, 85.2, 80.0, 77.2, 71.9, 71.1, 70.6, 68.6, 66.8 (CH); 63.5, 61.2 (CH₂); 55.8 (CH); 26.6 (CH₃, C(CH₃)₁); 20.3 (COCH₃); 19.1 (qC, C(CH₃)₃). FAB-MS calculated for C₄₃H₉₈N₄O₁₅Si [MH]⁺ 903.3042, found 903.311.

2, 3, 4, 6-Tetra-O-acetyl-D-galactopyranosyl-(β-1, 4)-3, 6-di-O-acetyl-2-acetamido-2-deoxy-β-D-glucopyranosyl azide,¹⁸¹ [89]

To a solution of the disaccharide 88 (946 mg, 1.05 mmol) and acetic acid (600 µL, 10.50 mmol) at 0 °C in THF (31.5 mL) was added a 1 M solution of TBAF in THF (4.2 mL, 4.20 mmol). The reaction mixture was stirred for 15 h at room temperature. The mixture was subsequently concentrated to a yellow oil. This oil was dissolved in n-BuOH (20 mL), and ethylenediamine (5 mL) was added. The reaction was heated to 90 °C and stirred for 16 h. The mixture was then concentrated under high vacuum and dissolved in a solution of 2:1 pyridine/Ac₂O (100 mL) and the reaction was stirred for 8 hours. The reaction was again concentrated to afford a yellow oil which was purified by silica gel chromatography (the column was packed with 1 % methanol in DCM; the mobile phase was 3 % methanol in DCM) to afford the purified product 89 (1.2 g; 76 %) as a white foam. Rᶠ= 0.23 (3 % methanol in DCM). ¹H-NMR (360 MHz, CDCl₃) δ (ppm): 5.96 (1H, d, J₈H₂= 9.5 Hz, NHAc); 5.32 (1H, dd, J₃,₄= 3.3 Hz, H₄); 5.10-5.04 (2H, m, H₂, H₁'); 4.96 (1H, dd, J₃,₅= 10.2 Hz, J₅,₆= 3.2 Hz, H₃); 4.53-4.48 (2H, m, H₁, H₃'); 4.12-3.71 (7H, m, H₅, H₆a, H₆b, H₂', H₄', H₆a', H₆b'); 3.68 (1H, m, H₅'); 2.26, 2.13, 2.10, 2.08, 2.04, 2.00, 1.95 (21 H, 7 x s, COCH₃).

¹³C-NMR (63 MHz, CDCl₃) δ (ppm): 171.0, 170.4, 170.3, 170.1, 170.0, 169.3 (CO); 101.1,
7. Experimental section

88.3, 75.6, 74.5, 72.6, 70.6, 68.9, 66.4 (CH); 61.7, 60.6 (CH₂); 52.9 (CH); 23.0, 20.7, 20.5, 20.4 (CH₃). FAB-MS calculated for C₂₆H₃₆N₄O₁₆ [MH]⁺ 661.2126, found 661.1630.

2, 3, 4, 6-Tetra-O-acetyl-D-galactopyranosyl-[(β-1, 4)-2-acetamido-2-deoxy-3, 6-di-O-acetyl-1-N-[1-(2-bromo)acetyll-5-D-glucopyranose, [90]

![Chemical Structure](image)

The peracetylated disaccharide 89 (1.98 g, 3.0 mmol) and 10 % Pd/C (205 mg), were dissolved in anhydrous methanol (90 mL), and stirred at room temperature under an atmosphere of hydrogen for 2 h. The catalyst was then removed by filtration through Celite, and the filter pad was washed with methanol. The solvent was removed under vacuum. The crude product was redissolved in anhydrous DMF (20 mL), and bromoacetic anhydride (1.01 g, 3.90 mmol) was added. The reaction was stirred at room temperature overnight. The reaction was diluted with chloroform (200 mL), and washed with 5 % HCl (200 mL), saturated aqueous NaHCO₃ (200 mL), and water (200 mL). The organic phase was dried over MgSO₄, filtered and concentrated under reduced pressure. The crude product was purified by flash chromatography over silica (2.5 % methanol in DCM) to afford a brown solid foam 90 (2.19 g; 95 %). \( R_f = 0.25 \) (2 % methanol in DCM). \[^1\text{H-NMR (250 MHz, CDCl}_3\text{) }\delta \text{(ppm): } 7.65 \text{ (1H, d, } J_{NH} = 8.1 \text{ Hz, } C1'-\text{NHCO}); 7.04 \text{ (1H, d, } J_{NH} = 8.3 \text{ Hz, NHAc}); 5.27 \text{ (1H, dd, } J_{4} = 3.0 \text{ Hz, } J_{4,5} = 0.6 \text{ Hz, H4}); 5.05 \text{ (1H, m, H3'); 4.97 \text{ (1H, dd, } J_{1,2} = J_{2,3} = 10.1 \text{ Hz, H2}); 4.90 \text{ (1H, dd, } J_{2,3} = 10.1 \text{ Hz, } J_{3,4} = 3.2 \text{ Hz, H3}); 4.42 \text{ (1H, d, } J_{1,2} = 10.1 \text{ Hz, H1}); 4.35 \text{ (1H, d, } J_{1,2} = 11.61 \text{ Hz, H1'}); 4.06-3.66 \text{ (8H, m, H2', H4', H5, H5', H6a, H6b, H6a', H6b'); 3.75 \text{ (2H, s, COCH}_2\text{Br}); 2.11, 2.09, 2.08, 2.05, 2.04, 1.98, 1.94 \text{ (21H, 7 x s, CH}_3\text{CO).}

[^13]C-NMR (63 MHz, CDCl₃) \( \delta \text{(ppm): } 172.2, 171.0, 170.3, 170.1, 169.2, 167.4 \text{ (CO); 101.0,} \)
7. Experimental section

80.0, 76.8, 74.4, 72.7, 70.9, 70.6, 68.9, 66.9 (CH); 62.0, 60.9 (CH$_2$); 52.7 (CH); 28.2 (CH$_2$); 22.9, 20.9, 20.6, 20.5 (CH$_3$). IR (CHCl$_3$) cm$^{-1}$: 3018.4, 1751.2 and 1670.2. FAB-MS calculated for C$_{26}$H$_{39}$BrN$_2$O$_7$ [MNa]$^+$ 777.13297 (based on $^{79}$Br), found: 777.13018.

2-Acetamido-2-deoxy-3, 4, 6-tri-O-acetyl-ß-D-glucopyranosyl chloride,$^9$ [94]

\[ \begin{array}{c}
\text{OH} & \text{OAc} \\
\text{HO} & \text{AcCl} \\
\text{HO} & \text{AcO} \\
\end{array} \]

\[ \begin{array}{c}
\text{HO} & \text{AcN} \\
\end{array} \rightarrow \begin{array}{c}
\text{OAc} & \text{OAc} \\
\text{AcN} & \text{Cl} \\
\end{array} \]

N-Acetyl glucosamine 93 (20.0 g, 90.46 mmol) was added to stirring acetyl chloride (30 mL). The suspension was stirred magnetically for 20 h. Afterwards the reaction was diluted with CHCl$_3$ (50 mL) and the resulting solution was washed with ice-water (50 mL) and ice-saturated aqueous NaHCO$_3$ (50 mL). The organic phase was dried over MgSO$_4$, filtered and concentrated under reduced pressure. The crude product was purified by silica gel chromatography (the column was packed with 1:1 petroleum ether 40-60/ethyl acetate; mobile phase 1:2) to afford a transparent solid 94 (11.36g; 34%). R$_f$ = 0.40 (Ethyl acetate).

$^1$H-NMR (300 MHz, CDCl$_3$) $\delta$ (ppm): 6.15 (1H, d, $J_{1-2}$ = 3.7 Hz, H1); 6.03 (1H, d, $J_{4-5}$ = 8.7 Hz, NH); 5.30 (1H, dd, $J_{2-3}$ = $J_{4-5}$ = 9.9 Hz, H2); 5.17 (1H, $J_{3-4}$ = $J_{4-5}$ = 9.9 Hz, H4); 4.51 (1H, $J_{3-4}$ = 9.9 Hz, $J_{5-6}$ = 8.7 Hz, $J_{1-2}$ = 3.7 Hz, H2); 4.28-4.20 (2H, m, H5, H6a), 4.11-4.07 (1H, m, H6b); 2.06, 2.01 and 1.95 (9H, s, CH$_3$CO). $^{13}$C-NMR (75 MHz, CDCl$_3$) $\delta$ (ppm): 171.4, 170.6, 170.2, 169.1 (CO); 93.7, 70.9, 70.1, 67.0 (CH); 61.1 (CH$_2$); 53.4 (CH); 23.0, 20.7, 20.5 (CH$_3$). FAB-MS calculated for C$_{14}$H$_{30}$NO$_4$Cl [MH]$^+$ 366, found 366.
2-Acetamido-2-deoxy-3, 4, 6-tri-O-acetyl-β-D-glucopyranosyl azide,\(^92\) [95]

![Chemical structure](image)

To a solution of the glucosyl chloride 94 (11.36 g, 31.12 mmol), TBAHS (10.57 g, 31.12 mmol) and NaN\(_3\) (6.07 g, 93.36 mmol) in DCM (110 mL), was added saturated aqueous NaHCO\(_3\) (110 mL). The resulting biphasic solution was stirred vigorously at room temperature for 1 h. Ethyl acetate (200 mL) was then added and the organic layer was separated and washed with saturated aqueous NaHCO\(_3\) (100 mL), and water (2 x 100 mL). The organic phase was dried over MgSO\(_4\), filtered and concentrated under reduced pressure, to afford a white solid 95 (9.54 g; 82%). \(R_f = 0.30\) (Ethyl acetate). \(^1\)H-NMR (300 MHz, CDCl\(_3\)) \(\delta\) (ppm): 6.02 (1 H, d, \(J_{NH-2}= 8.9\) Hz, NH); 5.26 (1 H, \(J_{3,4}= 9.8\) Hz, \(J_{5,3}= 9.6\) Hz, H3); 5.08 (1 H, \(J_{3,4}= J_{4,5}= 9.8\) Hz, H4); 4.78 (1 H, d, \(J_{4,5}= 9.3\) Hz, H1); 4.25 (1 H, dd, \(J_{6a,6b}= 12.4\) Hz, \(J_{3,6a}= 4.8\) Hz, H6a); 4.14 (1 H, dd, \(J_{6a,6b}= 12.4\) Hz, \(J_{5,6b}= 2.3\) Hz, H6b); 3.91 (1 H, dd, \(J_{5,6b}= 9.6\) Hz, \(J_{1,2}= 9.3\) Hz, \(J_{NH-2}= 8.9\) Hz, H2); 3.80 (1 H, dd, \(J_{4,5}= 9.8\) Hz, \(J_{3,6a}= 4.8\) Hz, \(J_{5,6b}= 2.3\) Hz, H5); 2.10, 2.02, 2.01 and 1.96 (12H, s, CH\(_3\)CO). \(^1\)C-NMR (75 MHz, CDCl\(_3\)) \(\delta\) (ppm): 170.9, 170.7, 170.6, 169.3 (CO); 88.4, 73.9, 72.2, 68.2 (CH); 61.9 (CH\(_2\)); 54.1 (CH); 23.2, 20.7, 20.6, 20.6 (CH\(_3\)). FAB-MS calculated for C\(_{141-120}\)N\(_{40}\)O\(_{8}\) [MH\(^+\)] 372.3, found 373.0.

2-Acetamido-2-deoxy-3-O-acetyl-4, 6-O-p-methoxybenzylidene-β-D-glucopyranosyl azide,\(^92\) [96]

![Chemical structure](image)
The glycosyl azide 95 (3.0 g, 8.07 mmol), was dissolved in dry MeOH (15 mL), sodium methoxide (200 µL of a 0.5 M solution in methanol) was added and the reaction mixture was stirred for 2 h at room temperature. The reaction mixture was neutralized by adding acetic acid (10 µL), and concentrated under reduced pressure to afford the crude deacetylated product as a pale yellow foam. The crude azide was dissolved in anhydrous DMF (10 mL) and p-anisaldehyde dimethylacetal (3.37 g, 18.54 mmol) and p-Tosic acid (0.28 g, 1.61 mmol) were added. After stirring for 1.5 h at 50 °C, the reaction mixture was concentrated under vacuum. The residue was poured into a cold mixture of saturated aqueous NaHCO₃ (30 mL) and DCM (30 mL) and cooled for 10 min at 4 °C. The precipitate was crystallized from ethyl acetate (30 mL). The product was filtered, dried under vacuum and isolated as a white solid. The resulting product was dissolved in pyridine (20 mL) and acetic anhydride (10 mL). The mixture was stirred for 24 h at room temperature and then concentrated under vacuum. Dichloromethane (150 mL) was then added and the organic phase was washed with water (20 mL), and saturated aqueous NaHCO₃ (20 mL). The organic layer was dried over MgSO₄, and the solvent was removed under vacuum to afford the crude product as a white solid which was crystallised from ethyl acetate to afford the pure product 96 (1.19 g; 70 %) as a white solid. Rₜ= 0.30 (Ethyl acetate). ¹H-NMR (300 MHz, CDCl₃) δ (ppm): 7.34 (2H, d, J = 8.7 Hz, ArH); 6.88 (2H, d, J = 8.8 Hz, ArH); 5.88 (1H, d, J₁₂= 9.5 Hz, NHAc); 5.48 (1H, s, CHPh); 5.24 (1H, dd, J₁₂= 9.9 Hz, H3); 4.48 (1H, d, J₅-₆₂= 9.2 Hz, H1); 4.32 (1H, dd, J₅₆₂= 10.5 Hz, J₅₆₆₂= 4.6 Hz, H6₂); 4.11 (1H, ddd, J₁₂= 9.9 Hz, J₁₂= 9.2 Hz, H2); 3.79 (3H, s, CH₃O); 3.77-3.61 (3H, m, H4, HS, H6b); 2.09 and 2.00 (6H, s, CH₃CO). ¹³C-NMR (75 MHz, CDCl₃) δ (ppm): 171.4, 170.9 (CO); 159.8, 128.1 (Ar); 127.1, 113.2 (Ar-H); 101.2, 88.8, 78.1, 71.4, 68.1 (CH); 67.9 (CH₂); 54.9 (CH₃); 53.3 (CH); 22.2, 20.3 (CH₃). FAB-MS calculated for C₁₈H₂₂N₄O₇ [MNa]⁺ 429.0, found 429.0.
2-Acetamido-2-deoxy-3-O-acetyl-6-O-p-methoxybenzyl-β-D-glucopyranosyl azide, [97]

TFA (6.8 mL, 88.7 mmol) in dry DMF (55 mL) was cooled down to 0 °C, and was added dropwise to a stirring solution of the sugar azide 96 (3.6 g, 8.87 mmol), sodium cyanoborohydride (2.79 g, 144.35 mmol) and 4Å molecular sieves in dry DMF (70 mL) at 0 °C in an ice bath. After the addition was complete, the ice bath was removed and the reaction mixture was stirred at room temperature for 16 h. Next, the reaction was filtered with suction. The filtrate was poured into ice-cold saturated aqueous NaHCO₃ and the product was extracted with dichloromethane (5 x 60 mL). The combined organic extracts were washed with a saturated aqueous solution of NH₄Cl (100 mL) and water (100 mL). The organic phase was dried over MgSO₄, filtered and concentrated under reduced pressure. The crude product was purified by silica gel chromatography (the column was packed with 1:1 petroleum ether 40-60/ethyl acetate; mobile phase: ethyl acetate) to afford 97 (1.0 g; 66 %) as a white solid. Rf= 0.20 (Ethyl acetate). ¹H-NMR (300 MHz, CDCl₃) δ (ppm): 7.28 (2H, d, J= 8.7, ArH); 6.91 (2H, d, J= 8.7, ArH); 6.02 (1H, d, J₁₂-NHAc= 9.2 Hz, NHAc); 5.08 (1H, dd, J₂,₃= 10.5 Hz, J₃,₄= 9.2 Hz, H3); 4.59 (1H, d, J₁,₂= 9.3 Hz, H1); 4.54 (2H, q, OCH₂Ph); 4.09 (1H, ddd, J₂,₃= 10.5 Hz, J₁,₂= 9.3 Hz, J₂-H₂NAC= 9.2 Hz; H2); 3.78 (3H, s, CH₃O); 3.77-3.46 (4H, m, H4, F1,5, H6 a, I -I60; 3.26 (1H, bp, OH); 2.09 and 1.96 (6H, s, CH₃CO). ¹³C-NMR (75 MHz, CDCl₃) δ(ppm): 171.8, 171.4 (CO); 159.3, 129.7 (Ar); 129.4, 113.8 (Ar-H); 88.8, 77.4, 75.4 (CH); 73.4, 68.9 (CH₂); 68.6 (CH); 55.1 (CH₃); 53.1 (CH); 22.7, 20.7 (CH₃). FAB-MS calculated for C₁₈H₂₄N₄O₇ [MH]+ 431.4, found 431.4.
2-Acetamido-2-deoxy-3-O-acetyl-4-O-tert-butoxycarboxymethyl-6-O-p-methoxybenzyl-β-D-glucopyranosyl azide, [98]

To a solution of the sugar derivative 97 (2.28 g, 5.60 mmol) in DMF (15 mL) at 0 °C, NEt₃ (1.6 mL, 11.20 mmol) was added and the solution was stirred for 30 min. Then, tert-butylbromo acetate (3 mL, 20.30 mmol) and silver oxide (2.3 g, 10.20 mmol) were added. The reaction was stirred overnight at room temperature under nitrogen gas with exclusion of light. The crude reaction was diluted with DCM (50 mL) and filtered through Celite. The organic phase was washed with saturated aqueous NaHCO₃ (4 x 50 mL) and water. The organic phase was dried MgSO₄, filtered and concentrated under reduced pressure. The crude product was purified by silica gel chromatography (the column was packed with 1:1 petroleum ether 40-60/ethyl acetate; mobile phase 1:2) to afford 98 (1.67 g, 57%) as a brown oil. Rf = 0.5 (ethyl acetate). ¹H-NMR (300 MHz, CDCl₃) δ (ppm): 7.21 (2H, d, J = 8.7, ArH); 6.85 (2H, d, J = 8.7, ArH); 6.35 (1H, d, J_H₂,NHAc = 9.6 Hz, NHAc); 5.14 (1H, dd, J₁₂,NHAc = 10.9 Hz, J₁₂, H₁); 4.57 (1H, dd, J₃₄,H₄ = 8.4 Hz, H₄); 4.43 (1H, d, J₅, H₅ = 11.5 Hz, H₅); 3.98 (2H, s, COCH₂Br); 3.77 (3H, s, CH₃O); 3.75-3.58 (3H, m, H₅, H₆₆₈, H₆₀); 2.07 and 1.95 (6H, s, CH₃CO); 1.41 (9H, s, (CH₃)₃C). ¹³C-NMR (75 MHz, CDCl₃) δ (ppm): 171.4, 170.5, 168.6 (CO); 159.3, 130.1 (Ar); 129.6, 129.4, 113.8 (ArH); 88.7, 81.8, 76.7, 75.1 (CH); 73.2, 70.3, 67.9 (CH₂); 55.3 (CH₃); 53.4 (CH); 28.1 (CH₃); 22.7, 20.7 (CH₃). FAB-MS calculated for C₃₅H₄₅N₄O₉ [MNa⁺] 545.5, found 545.5.
The azido sugar 98 (145 mg, 0.28 mmol) was dissolved in 9:1 MeCN/H₂O (1.5 mL) and CAN (307 mg, 0.56 mmol) was added. After 1 h, TLC (ethyl acetate) indicated that the reaction was complete. The crude reaction was diluted with DCM (25 mL) and washed with saturated aqueous NaHCO₃ (10 mL). The organic phase was dried over MgSO₄, filtered and concentrated under reduced pressure. The residue was dissolved in a solution of 2:1 pyridine/Ac₂O (6 mL), and the reaction was stirred for 16 h. The reaction was concentrated to afford a yellow oil. The crude product was purified by silica gel chromatography (the column was packed with 1:1 petroleum ether 40-60/ethyl acetate; mobile phase 1:2) to afford the purified product 99 (95 mg; 85 %) as a colourless oil. Rf = 0.5 (ethyl acetate) ¹H-NMR (300 MHz, CDCl₃) δ (ppm): 6.32 (1H, d, J = 9.6 Hz, NHAc); 5.17 (1H, dd, J₂,₄ = 10.4 Hz, J₂,₃ = 9.2 Hz, H3); 4.65 (1H, d, J₁,₂ = 9.2 Hz, H1); 4.43 (1H, dd, J₆₋₆b = 12.2 Hz, J₅₋₆a = 2.1 Hz, H6b); 4.29 (1H, dd, J₆₋₆b = 12.2 Hz, J₅₋₆a = 4.7 Hz, H6a); 4.12-3.96 (3H, m, H4, CH₂); 3.78 (1H, m, H5); 3.53 (1H, dd, J₁₋₂ = 9.2 Hz, H2); 2.10, 2.08, 1.96 (9H, s, CH₃CO); 1.42 (9H, s, (CH₃)₃C). ¹³C-NMR (75 MHz, CDCl₃) δ (ppm): 171.3, 170.6, 170.4, 168.4 (CO); 88.5 (CH); 82.1 (qC, C(CH₃)₃); 77.5, 74.4 (CH); 74.4 (CH₂); 70.2 (CH); 62.7 (CH₂); 53.3 (CH); 28.0 (CH₂, C(CH₃)₃); 23.1, 21.1, 20.8 (CH₃). FAB-MS calculated for C₁₈H₂₈N₄O₉ [MNa]⁺ 467.4, found 467.4.
2-Acetamido-2-deoxy-3, 6-di-O-acetyl-4-O-carboxymethyl-β-D-glucopyranosyl azide, [100]

A solution of the monosaccharide derivative 99 (422 mg, 0.95 mmol) in TFA/H₂O (95:5, 10 mL) was stirred at room temperature for 2 h. Afterwards, the solution was concentrated under vacuum to give the desired compound 100 (382 mg; 98%) as a brown oily wax. ¹H-NMR (300 MHz, CDCl₃) δ (ppm): 6.67 (1H, d, J = 9.4 Hz, NHAc); 5.21 (1H, dd, J₂,₃ = J₃,₄ = 9.5 Hz, H3); 4.68 (1H, d, J₁,₂ = 9.4 Hz, H1); 4.49 (dd, J₆₋₆b = 11.6 Hz, H₆₆); 4.31 (dd, J₆₋₆b = 11.6 Hz, J₅₋₆b = 9.5 Hz, H₅); 4.24 (2H, d, J = 3.7 Hz, CH₂); 4.02-3.99 (1H, m, H₂); 3.77 (1H, m, CH₂); 3.61 (1H, dd, J₃₋₄ = J₄₋₅ = 9.5 Hz, H₄); 2.12, 2.09 and 2.03 (9H, s, CH₃CO). ¹³C-NMR (75 MHz, CDCl₃) δ (ppm): 171.2, 170.4, 170.1, 168.0 (CO); 87.7 (CH); 75.2 (CH₂); 75.1, 74.2, 70.1 (CH); 63.4 (CH₂); 54.3 (CH); 22.2, 20.7 (CH₃). IR (CHCl₃) cm⁻¹: 3300.0, 3020.3, 2119.6, 1733.9, 1683.7 and 1217.0. FAB-MS calculated for C₁₄H₂₇N₄O₉ [MH⁺]: 389.13085, found: 389.13041.

Synthesis of compound [103]
General experimental details for manual PSDC 103 synthesis.

Solid-phase peptide synthesis was carried out manually using a plastic syringe fitter with a Teflon frit and connected to a vacuum waste chamber via a Teflon valve. Synthesis of the PSDC 103 was carried out on MBHA Rink amide resin (loading = 0.58 mmol/g).


The synthesis of the compound 103 was conducted on 0.05 mmol scale. The Fmoc group was removed by treatment with 20 % piperidine in DMF (1 x 5 and 1 x 15 min) between each coupling step. Coupling reactions with Fmoc amino acids were conducted using 0.25 mmol (5 equiv.) of each of HBTU/HOBt and DIPEA for 3 h. The reaction progress was monitored using the Kaiser ninhydrin test. Between each coupling and deprotection step, the resin was washed with DCM and DMF (5 min each).

The resin was treated with a mixture (Ac₂O/pyridine, 2:1) for 16 h. Finally the resin was washed exhaustively with DMF and DCM.

General S'Bu deprotection.

DTT (100 mg) was dissolved in dry DMF (900 µL) and 2.5 % v/v DIPEA (25 µL) was added. After stirring for 5 minutes the solution was transferred to a peptide synthesis vessel containing resin-bound S'Bu protected peptide. After 16 h the resin was filtered and washed exhaustively with DMF, DMF/H₂O (1:1), DMF and finally DCM. The S'Bu procedure was repeated twice.

General bromoacetamide couplings.

Bromoacetamide 90 (2 equiv. per thiol: 0.05 mmol x 2 thiols x 2 equiv.= 151 mg, 0.20 mmol) was dissolved in DMF (400 µL) and NEt₃ (42 µL, 0.3 mmol) and transferred to a
peptide synthesis vessel containing resin-bound S'Bu deprotected peptide. The reaction was allowed to proceed for 24 h. After this time, the resin was filtered and washed exhaustively with DMF, DMF/H₂O (1:1), DMF and finally DCM.

**General Alloc deprotection.**

The resin was washed for 5 min with a solution of DMF/CHCl₃/AcOH/N-methyl morpholine (NMM) (18.5:18.5:2:1). Pd(PPh₃)₄ (115 mg, 0.1 mmol) was dissolved in a solution of DMF/CHCl₃/AcOH/NMM (18.5:18.5:2:1) (0.05 M Pd(PPh₃)₄). The mixture was stirred under nitrogen for 2 h with exclusion of light. The process was repeated again. Finally, the resin was then washed sequentially with the following solutions: 0.5 % DIPEA in DMF (v/v) (4 x 2 mL); DMF (6 x 2 mL); 0.5 % (w/v) sodium diethyldithiocarbamate trihydrate in DMF (4 x 2 mL); followed by a final wash with DMF (6 x 2 mL).

**Coupling of 2-acetamido-2-deoxy-3,6-di-O-acetyl-4-carboxymethyl-β-D-glucopyranosyl azide 100 and peracetylation of the molecule.**

The resin was washed with DMF (2 mL). Coupling reactions with 2-acetamido-2-deoxy-3,6-diacetyl-4-carboxymethyl-β-D-glucopyranosyl azide 100 were conducted using 0.25 mmol (5 equiv.) of the sugar, HBTU/HOBt and DIPEA for 3 h.

Afterwards, the resin was treated with a mixture of pyridine and acetic anhydride (2:1) for 16 h. Finally the resin was washed exhaustively with DMF and DCM.

**Cleavage of compound 103.**

The resin was washed with DCM and DMF five times (5 min each wash). The compound 103 was cleaved from the resin by incubation with the cleavage cocktail (95 % TFA, 2.5 % water and 2.5 % EDT) for 3 h. The resin was treated again with the same cleavage cocktail for another 3 h. The compound 103 was precipitated with cold diethyl ether and centrifuged.
for 10 min. The precipitate was dried and redissolved in a 1:1 water/acetonitrile solution and lyophilised.

**Compound 103 purification.**

The crude product was purified by reverse phase semi preparative HPLC (water-acetonitrile 5 %–95 % in 40 minutes, 0.1 % TFA), \( t_R = 30.3 \) min. The corresponding fractions containing 103 were pooled and freeze dried to afford a fluffy white solid (32 mg, 30 % overall yield of 10 steps).

Analytical data for compound 103: \( \text{C}_{54}\text{H}_{121}\text{N}_{13}\text{O}_{46}\text{S}_{2} \). MW: 2113.05. Found: \([M]^+ 2113.05, [M]^+ 2 1057.52

**Synthesis of compound [104]**

![Chemical structure of compound 104]

**General experimental details for manual PSDC 104 synthesis.**

Solid-phase peptide synthesis was carried out manually using a plastic syringe fitter with a Teflon filter and connected to a vacuum waste chamber via a Teflon valve. Synthesis of the PSDC 104 was carried out on MBHA Rink amide resin (loading = 0.58 mmol/g).
General Solid-phase peptide synthesis: Ac-L-Cys-L-Lys-L-Cys-NH₂

The synthesis of compound 104 was conducted on 0.10 mmol scale. The Fmoc group was removed by treatment with 20 % piperidine in DMF (5 and 15 min) between each coupling and deprotection step. Coupling reactions with Fmoc amino acids were conducted using 0.50 mmol (5 equiv.) of each of HBTU/HOBt and DIPEA for 3 h. The reaction progress was monitored using the Kaiser ninhydrin test. Between each coupling and deprotection step, the resin was washed with DCM and DMF (5 min each).

The resin was treated with a mixture (Ac₂O/pyridine, 2:1) for 16 h. Finally the resin was washed exhaustively with DMF and DCM.

General S'Bu deprotection.

DTT (200 mg) was dissolved in dry DMF (1.8 mL) and 2.5 % v/v DIPEA (50 μL) was added. After stirring for 5 minutes the solution was transferred to a peptide synthesis vessel containing resin-bound S'Bu protected peptide. After 16 h the resin was filtered and washed exhaustively with DMF, DMF/H₂O (1:1), DMF and finally DCM. The S'Bu deprotection procedure was repeated twice.

General bromoacetamide couplings.

Bromoacetamide 90 (2 equiv. per thiol: 0.10 mmol x 2 thiols x 2 equiv.= 302 mg, 0.40 mmol) was dissolved in DMF (1 mL) and NEt₃ (84 μL, 0.6 mmol) and transferred to a peptide synthesis vessel containing resin-bound S'Bu deprotected peptide. The reaction was allowed to proceed for 24 h. After this time, the resin was filtered and washed exhaustively with DMF and the DCM.
7. Experimental section

**General Alloc deprotection.**

The resin was washed for 5 min with a solution of DMF/CHCl₃/AcOH/N-methyl morpholine (NMM) (18.5:18.5:2:1). Pd(PPh₃)₄ (230 mg, 0.2 mmol) was dissolved in a solution of DMF/CHCl₃/AcOH/NMM (18.5:18.5:2:1) (0.05 M Pd(PPh₃)₄). The mixture was under stirred nitrogen for 2 h with exclusion of light. The process was repeated again. Finally, the resin was washed sequentially with the following solutions: 0.5 % DIPEA in DMF (v/v) (4 x 2 mL); DMF (6 x 2 mL); 0.5 % (w/v) sodium diethyldithiocarbamate trihydrate in DMF (4 x 2 mL); followed by a final wash with DMF (6 x 2 mL).

**Coupling of 2-acetamido-2-deoxy-3,6-di-O-acetyl-4-carboxymethyl-β-D-glucopyranosyl azide 100 and peracetylation of the molecule.**

The resin was washed with DMF (2 mL). Coupling reactions with 2-acetamido-2-deoxy-3,6-diacetyl-4-carboxymethyl-β-D-glucopyranosyl azide, 100 (97 mg, 0.25 mmol, 2.5 equiv.), 2.5 equiv. PyBOP/HOBt and 5 equiv. DIPEA for 16 h. The process was repeated. Afterwards the resin was exhaustively washed with DCM and DMF.

**Cleavage of compound 104.**

The resin was washed with DCM and DMF five times (5 min each wash). The compound 104 was cleaved from the resin by incubation with the cleavage cocktail (95 % TFA, 2.5 % water and 2.5 % EDT) for 3 h. The resin was treated again with the same cleavage cocktail for another 3 h. The compound 104 was precipitated with cold diethyl ether and centrifuged for 10 min. The precipitate was dried by flushing with nitrogen for 30 min.

**Acetyl deprotection of the carbohydrates.**

The precipitate was resuspended with a water/methanol solution (10:1) of 5 % hydrazine (6 mL). The reaction was stirred at room temperature for 24 h.
Compound 104 purification.

The crude product was purified by reverse phase semi preparative HPLC (water-acetonitrile 5 \(\%\)-95 \(\%\) in 40 minutes, 0.1 \(\%\) TFA), \(\tau_g=14.2\) min. The corresponding fractions containing 104 were pooled and freeze dried to afford a fluffy white solid (72 mg; 47 \(\%\) overall yield of 10 steps).

Analytical data for compound 104: \(C_{56}H_{93}N_{13}O_{32}S_2\). MW: 1524.79. Found: [M]: 1525.79; [M]+2: 763.48; [M+\(\text{NH}_2\text{NH}_2\)]+2:779.40; [M+2\(\text{NH}_2\text{NH}_2\)]+2: 795.63.

2-Bromo-N-prop-2-ynyl-acetamide,\textsuperscript{209} [112]

\[
\begin{align*}
\text{Br}\text{-..\text{-..---}} & \text{Br} \\
\text{H} & \text{-..\text{-..---}} \\
\text{N} & \text{H} \\
\text{O} & \text{O} \\
\text{N} & \text{H} \\
\text{Br} & \text{Br}
\end{align*}
\]

Propargylamine 115 (200 \(\mu\)L, 2.90 mmol) was dissolved in water (30 mL) and this was followed immediately by the addition of bromoacetic anhydride 116 (3.7 g, 14.50 mmol) and \(\text{NaHCO}_3\) (5.0 g). The reaction was stirred at room temperature for 16 h. The reaction was quenched with 5 \(\%\) v/v aqueous \(\text{HCl}\) solution (50 mL), and the product was extracted with ethylacetate (3 x 50 mL). The organic phase was washed with \(\text{NaOH} 1 \text{ N}\) (5 x 100 mL) and water (2 x 100 mL). The organic phase was dried with \(\text{MgSO}_4\), filtered and concentrated under reduced pressure to afford 112 (0.4 g; 76 \(\%\)) as a crystalline solid. \(R_e=0.33\) (petroleum ether 40-60/ethyl acetate 1:1). \(^1\text{H}-\text{NMR}\) (250 MHz, \text{CDCl}_3) \(\delta\) (\(\text{ppm}\)): 6.78 (1H, s, \(\text{NH}\)); 4.07 (2H, q, \(J=5.4\) Hz, \(J=2.6\) Hz, \(\text{CH}_2\)); 3.88 (2H, s, CO\(\text{CH}_2\text{Br}\)); 2.27 (1H, t, \(J=2.6\) Hz, \(\text{CH}\)). \(^{13}\text{C}-\text{NMR}\) (63 MHz, \text{CDCl}_3) \(\delta\) (\(\text{ppm}\)): 165.3 (CO); 78.5 (qC, alkyne); 72.2 (CH), 29.9, 28.6 (\(\text{CH}_2\)). \(\text{IR}\) (\text{CHCl}_3) \text{cm}^{-1}: 3307.7, 3018.4, and 1670.2. FAB-MS calculated for
Experimental section

177

C_{12}H_{13}BrNO [MH]^+ 174.96, found: 197.85 (based on ^{79}\text{Br}), and 199.85 [MNa]^+ (based on ^{81}\text{Br}).

2-Benzylsulfanyl-N-prop-2-ynyl-acetamide, [119]

![Chemical structure](image)

2-Bromo-N-prop-2-ynyl-acetamide 112 (100 mg, 0.57 mmol) was dissolved in DMF (4 mL). Benzylmercaptan 117 (700 \mu L, 5.77 mmol) and triethylamine (885 \mu L, 6.35 mmol) were added. The reaction was stirred for 16 h. The crude of the reaction was diluted with chloroform (10 mL), and washed with NaOH (1M, 10 mL), 5% HCl (10 mL), saturated aqueous NaHCO₃ (10 mL) and water (10 mL). The organic phase was dried with MgSO₄, filtered and concentrated under reduced pressure. The crude product was purified by flash chromatography over silica (petroleum ether 40-60/ethyl acetate, 1:1) to afford 119 (105 mg; 84\%) as a brown solid. R_f= 0.38 (petroleum ether 40-60/ethyl acetate 1:1).^1\text{H}-NMR (300 MHz, CDCl₃) \delta (ppm): 7.34-7.22 (5H, m, Ph); 6.91 (1H, bp, NH); 3.95 (2H, q, J= 5.4 Hz, J= 2.5 Hz, CH₂); 3.72 (2H, s, COCH₂S); 3.12 (2H, s, PhCH₂S); 2.24 (1H, t, J= 2.5 Hz, CH).^13\text{C}-NMR (75 MHz, CDCl₃) \delta (ppm): 168.4 (CO); 137.0 (Ar); 129.0, 128.8, 127.5 (Ar-H) 79.3 (qC, alkyne); 71.8 (CH), 37.1, 35.0, 29.4 (CH₂). IR (CHCl₃) cm⁻¹: 3300.0, 3018.4, and 1670.2. FAB-MS calculated for C_{12}H_{13}NOS [MH]^+ 220.07906 (based on ^{79}\text{Br}), found 220.07910.
1-N-(3, 4, 6-Tri-O-acetyl-2-deoxy-2-N-acetyl-β-D-glucopyranosyl)-4-(N'-methylidene-2'-bromoacetamido)-1, 2, 3-triazole, [114]

The azido sugar 95 (100 mg, 0.27 mmol) and N-propargyl-bromoacetamide 112 (47 mg, 0.27 mmol) were dissolved in a biphasic solution of CHCl₃/EtOH/H₂O (9:1:1) (1.1 mL). Sodium ascorbate (54 mg, 0.27 mmol) and CuSO₄·5H₂O (2 mg, 0.007 mmol) were added. The reaction was shaken at 600 rpm, 50 °C overnight. Afterwards, it was diluted with CHCl₃ (10 mL) and washed with saturated aqueous NaHCO₃ (3 x 20 mL), and the organic phase was dried with MgSO₄, filtered and concentrated under reduced pressure to afford 114 (98 trig; 66 %) as a brown solid. Rₛ = 0.40 (ethyl acetate). ¹H-NMR (300 MHz, CDCl₃ and 5 % v/v CD₃OD) δ (ppm): 7.72 (1H, bp, CH-triazole); 5.76 (1H, d, J₁₂= 9.7 Hz, H1); 5.21 (1H, dd, J₂₃= J₃₄=9.8 Hz, H3); 4.99 (1H, dd, J₃₄= J₄₅=9.8 Hz, H4); 4.28 (1H, bp, H2); 4.19 (2H, s, CH₂NH); 4.08 (2H, dd, J₆₄₋₆₅= 12.7 Hz, J₅₋₆₅= 4.8 Hz, H₅); 3.92 (1H, dd, J₆₋₆₅= 12.7 Hz, J₅₋₆₅= 1.8 Hz, H₆); 3.87-3.83 (1H, in, H₅); 3.63 (2H, s, COCH₂Br); 1.86, 1.85, 1.82 (9H, 3 x s, CH₃CO); 1.52 (3H, s, CH₃CONH). ¹³C-NMR (75 MHz, CDCl₃ and 5 % of CD₃OD) δ (ppm): 171.7, 170.9, 170.4, 169.6 (CO); 82.0, 74.5, 72.0, 68.0 (CH); 61.7 (CH₂); 53.1 (CH); 35.0 (CH₂, Ar-CH₂-NH); 28.0 (CH₂, CO-CH₂-Br); 21.8, 20.2, 20.1, 20.1 (CH₃). FAB-MS calculated for C₁₉H₂₉BrN₅O₉ [MH]⁺: 548.09867 (based on ⁷⁹Br), found: 548.10034.
7. Experimental section

1-N-(3, 4, 6-Tri-O-acetyl-2-deoxy-2-N-acetyl-β-D-glucopyranosyl)-4-(N'-methylidene-2'-thiobenzylacetamido)-1, 2, 3-triazole, [118]

The azido sugar 95 (100 mg, 0.27 mmol) and propargyl derivative 119 (59 mg, 0.27 mmol) were dissolved in a biphasic solution of CHCl₃/EtOH/H₂O (9:1:1) (1.1 mL). Sodium ascorbate (54 mg, 0.27 mmol) and CuSO₄5H₂O (2 mg, 0.01 mmol) were added. The reaction was shaken at 600 rpm, 50 °C overnight. Afterwards, it was diluted with CHCl₃ (10 mL) and washed with saturated aqueous NaHCO₃ (3 x 20 mL), and the organic phase was dried with MgSO₄, filtered and concentrated under reduced pressure to afford 118 (146 mg; 91 %) as a brown solid. Rf = 0.33 (ethyl acetate). ¹H-NMR (300 MHz, CDCl₃ and 5 % v/v CD₃OD) δ (ppm): 7.75 (1H, s, CH-triazole); 7.24-7.18 (5H, m, Ph); 5.86 (1H, d, J₁₂ = 9.9 Hz, H1); 5.33 (1H, dd, J₂₃ = J₃₄ = 9.9 Hz, H3); 5.12 (1H, dd, J₃₄ = J₄₅ = 9.9 Hz, H4); 4.38 (1H, dd, J₃₄ = J₅₆ = 9.9 Hz, H2); 4.35 (2H, s, CH₂NH); 4.19 (1H, dd, J₆₅ = J₅₆ = 12.6 Hz, J₅₆ = 4.8 Hz, H₆); 4.03 (1H, dd, J₆₇ = 12.6 Hz, J₅₆ = 1.9 Hz, H₆₇); 3.94 (1H, ddd, J₆₇ = 9.9 Hz, J₅₆ = 4.8 Hz, H₆₇); 3.66 (2H, s, COCH₃); 3.03 (2H, s, SCH₂Ph); 1.97, 1.95, 1.93 (9H, 3 x s, CH₃CO); 1.62 (3H, s, CH₃CONH). ¹³C-NMR (75 MHz, CDCl₃ and 5 % v/v CD₃OD) δ (ppm): 171.4, 170.9, 170.6, 169.7, 169.6 (CO); 137.1 (Ar); 128.9, 128.5, 127.2 (Ar-H); 121.2 (Ar); 85.9, 74.7, 72.2, 68.0 (CH); 61.7 (CH₂); 53.3 (CH); 36.8, 34.8, 34.7 (CH₂); 23.6, 22.2, 20.5, 20.4 (CH₃). FAB-MS calculated for C₂₈H₃₃N₅O₉S [MH⁺]: 592.20717, found: 592.20858.
1-N-(3, 4, 6-Tri-O-acetyl-2-deoxy-2-N-acetyl-β-D-glucopyranosyl)-4-(N'-
methylideneyl-2'-thiobenzylacetamido)-1, 2, 3-triazole, [118]

The sugar derivative 114 (50 mg, 0.09 mmol) was dissolved in DMF (615 µL). Benzomercaptan 117 (106 µL, 0.90 mmol) and triethylamine (138 µL, 0.73 mmol) were added. The reaction was stirred for 16 h. The reaction was diluted with chloroform (10 mL), and washed with NaOH (1M, 10 mL), 5 % HCl (10 mL), saturated aqueous NaHCO₃ (10 mL) and water (10 mL). The organic phase was dried with MgSO₄, filtered and concentrated under reduced pressure. The crude product was purified by flash chromatography over silica (petroleum ether 40-60/ethyl acetate, 9:1 to 1:9) to afford 118 (40 mg; 75 %) as a brown solid. Rf = 0.33 (ethyl acetate). ¹H-NMR (300 MHz, CDCl₃ and 5 % v/v CD₂OD) δ (ppm): 7.75 (1H, s, CH-triazole); 7.24-7.18 (5H, m, Ph); 5.86 (1H, d, J₁,₂ = 9.9 Hz, H₁); 5.33 (1H, dd, J₂,₃ = J₂,₄ = 9.9 Hz, H₃); 5.12 (1H, dd, J₃,₄ = J₃,₅ = 9.9 Hz, H₄); 4.38 (1H, dd, J₁,₂ = J₂,₃ = 9.9 Hz, H₂); 4.35 (2H, s, CH₃-NH); 4.19 (1H, dd, J₃,₄ = 12.6 Hz, J₃,₅ = 4.8 Hz, H₃a); 4.03 (1H, dd, J₃,₅ = 12.6 Hz, J₃,₆b = 1.9 Hz, H₅b); 3.94 (1H, dd, J₅,₆b = 1.9 Hz, J₅,₆b = 1.9 Hz, H₆b); 3.66 (2H, s, COCH₃-S); 3.03 (2H, s, SCH₂-Ph); 1.97, 1.95, 1.93 (9H, 3 x s, CH₃CO); 1.62 (3H, s, CH₂CONH). ¹³C-NMR (75 MHz, CDCl₃ and 5 % of CD₂OD) δ (ppm): 171.4, 170.9, 170.6, 169.7, 169.6 (CO); 137.1 (Ar); 128.9, 128.5, 127.2 (Ar-H); 121.2 (Ar); 85.9, 74.7, 72.2, 68.0 (CH); 61.7 (CH₂); 53.3 (CH); 36.8, 34.8, 34.7 (CH₃); 23.6, 22.2, 20.5, 20.4 (CH₃). FAB-MS calculated for C₂₆H₃₃N₅O₉S [MH]⁺: 592.20717, found: 592.20858.
Experimental section

1-N-(2-Deoxy-2-N-acetyl-β-D-glucopyranosyl)-4-(N'-methylene-2'-thiobenzylacetamido)-1, 2, 3-triazole, [120]

The modified sugar 118 (137 mg, 0.23 mmol) was dissolved in a 2 % solution of hydrazine monohydrate in ethanol (5 mL). After 3 days TLC (10 % methanol in DCM) indicated that the reaction was complete. The solvent was removed under high vacuum, and the crude product was purified by flash chromatography over silica (10 % methanol in DCM) to afford the pure product 120 (71 mg; 66 %) as a light brown wax. Rf 0.01 (10 % methanol in DCM).

$^1$H-NMR (300 MHz, D$_2$O/CD$_3$OD) $\delta$ (ppm): 8.04 (1H, s, CH-triazole); 7.30-7.24 (5H, m, Ph); 5.76 (1H, d, $J_{1,2}$ = 9.8 Hz, H1); 4.40 (2H, s, triazole-CH$_2$-NH); 4.20 (1H, dd, $J_{1,2}$ = $J_{2,3}$ = 9.8 Hz, H2); 3.90-3.54 (5H, m, H3, H4, H5, H6a, H6b); 3.78 (2H, s, CO-CH$_2$-S); 3.12 (2H, s, S-CH$_2$-Ph); 1.76 (3H, s, COCH$_3$). $^{13}$C-NMR (75 MHz, D$_2$O/CD$_3$OD) $\delta$ (ppm): 173.5, 172.2 (CO); 146.0, 139.0 (Ar); 130.2, 129.5, 128.2, 123.0 (Ar-H); 88.2, 81.2, 75.6, 71.4 (CH); 62.3 (CH$_2$); 56.8 (CH); 37.5, 35.8, 35.5 (CH$_2$); 22.6 (CH$_3$). FAB-MS calculated for C$_{20}$H$_{27}$N$_5$O$_6$S [MH]: 466.17548, found: 466.17335.
To a vigorously stirring suspension of the unprotected sugar \(121\) (10 mg, 0.04 mmol) in tert-butanol (87 \(\mu\)L), was added propargylbromoacetamide \(112\) (30 mg, 0.17 mmol). The reaction was initiated by the addition of a solution of \(\text{CuSO}_4 \cdot 5\text{H}_2\text{O}\) (2 mg, 0.02 mmol), sodium ascorbate (3 mg, 0.02 mmol) and trisbenzyltriazolylmethylamine (TBTA) (4 mg, 0.08 mmol) in water (87 \(\mu\)L). The reaction was shaken at 37 °C, 1200 rpm, for 5 h. Afterwards, the reaction was diluted with water (325 \(\mu\)L), centrifuged at 1400 rpm for 5 minutes, and filtered through celite and cotton wool in a Pasteur pipette. The crude product was purified by reverse phase semi preparative HPLC (water-acetonitrile 5%-30% in 40 minutes, 0.1% TFA). The corresponding fractions were freeze dried to afford \(122\) (12 mg; 71 %) as a foamy white solid. \(^1\)H-NMR (300 MHz, D\(_2\)O) \(\delta\) (ppm): 8.10 (1H, s, CH-triazole); 5.80 (1H, d, \(J_{1,2} = 9.7\) Hz, H1); 4.49 (2H, s, C\(=\)N); 4.22 (1H, dd, \(J_{1,2} = J_{2,3} = 9.8\) Hz, H2); 3.94-3.50 (7H, m, H3, H4, H5, H6\(_a\), H6\(_b\), COCH\(_2\)Br); 1.71 (3H, s, CH\(_3\)CONH). \(^13\)C-NMR (75 MHz, D\(_2\)O) \(\delta\) (ppm): 174.5, 170.2 (CO); 86.8, 79.3, 73.9, 69.7 (CH); 60.8 (CHI); 55.7 (CH); 35.3 (CH\(_2\), Ar-CH\(_2\)-NH); 28.3 (CH\(_2\), CO-CH\(_2\)-Br); 22.0 (CH\(_3\)). FAB-MS calculated for C\(_{13}\)H\(_{20}\)BrN\(_3\)O\(_5\): 421.33 (based on \(^{79}\)Br), found 422.33 and 424.33[M\(^+\)].
7. Experimental section

**Synthesis of compound [130]**

The compound 103 (5 mg, 0.0023 mmol) and propargyl bromoacetamide 112 (1 mg, 0.0023 mmol) were dissolved in a biphasic solution of CHCl₃/EtOH/H₂O (9:1:1) (220 μL). Sodium ascorbate (1 mg, 0.0023 mmol) and CuSO₄·5H₂O (0.1 mg, 0.0002 mmol) were added. The reaction was shaken at 600 rpm, 37 °C overnight. Afterwards, the reaction was diluted with ethyl acetate and washed with saturated aqueous NaHCO₃ (3 x 10 mL), and the organic phase was dried with MgSO₄, filtered and concentrated under reduced pressure to afford the pure product 130 (2.5 mg; 50 %) as a fair brown wax.

Analytical data: C₉₉H₁₂₇BrN₁₄O₄₇S₂. MW: 2289.06. Found: [M]⁺ 1146.15.

**Synthesis of compound [131]**
The compound 104 (10 mg, 6.60x10^{-3} mmol) and propargylbromoacetamide 112 (5 mg, 2.60x10^{-2} mmol) were dissolved in an aqueous solution of sodium ascorbate (2.60x10^{-3} mmol, 52.00 mM, 50 μL) and a tert-butanol solution of tris-benzyltriazolylmethylamine (TBTA) (1.30x10^{-3} mmol, 13.10 mM, 100 μL). Afterwards, an aqueous solution of CuSO₄·5H₂O (1.30x10^{-3} mmol, 26.20 mM, 50 μL) was added. The reaction was shaken for 16 hours at 37 °C, 1200 rpm. Then, it was diluted with water (800 μL), centrifuged at 1400 rpm for 5 minutes, and filtered through celite plug in a Pasteur pipette. The crude product was purified by reverse phase semi preparative HPLC (water-acetonitrile 5 %-50 % in 40 minutes, 0.1 % TFA), t_R = 16.6 min. The corresponding fractions containing 131 were pooled and freeze dried to afford a fluffy white solid (7 mg; 64 %).


**Synthesis of compound [137]**

The desired EPO peptide fragment (residues 21-32), 132 (2 mg, 1.15x10^{-3} mmol) and the 131 (10 mg, 5.53x10^{-3} mmol) were dissolved in a 0.1 M sodium phosphate buffer solution,
pH 7.4 (400 μL). The mixture was incubated at 600 rpm, 37 °C, overnight. The crude product was purified by reverse phase semi preparative HPLC (water-acetonitrile 5%-95% in 40 minutes, 0.1% TFA), \( t_R = 16.2 \text{ min} \). The corresponding fractions containing 133 were pooled and freeze-dried to afford a fluffy white solid (2 mg; 29%).


**Synthesis of compound [134]**

133 (2 mg, \( 3.00 \times 10^{-4} \text{ mmol} \)) was dissolved in 10% AcOH (340 μL) at a conc. of 5 mg/mL, and mercury (II) acetate (3 mg, \( 9.00 \times 10^{-3} \text{ mmol} \)), was added followed 90 min later by dithiothreitol (3 mg, \( 2.00 \times 10^{-2} \text{ mmol} \)), to a final concentration of 50 mM. The mixture was allowed to stand at room temperature for 5 hours. The white precipitate that formed was removed by centrifugation at 13000 rpm for 5 min. The remaining solution was purified directly by reverse phase semi preparative HPLC (water-acetonitrile 5%-95% in 40
minutes, 0.1 % TFA), \( t_R = 16.3 \text{ min} \). The corresponding fractions containing 134 were pooled and freeze dried to afford a fluffy white solid (2 mg; 100 %).


**Synthesis of compound 136: (EPO derived peptide fragment (residues 22-32))**

![Chemical structure of compound 136](image)

*Solid phase peptide synthesis.*

Solid-phase peptide synthesis was carried out automatically in a peptide synthesizer (Applied Biosystems peptide synthesizer model 433A). Synthesis of the peptide sequence: \((\text{CQC(SS}^\text{Bu})\text{ITTGC(Acm)C(SS}^\text{Bu})\text{AS})\) was carried out on PEGA Rink amide resin (loading = 0.26 mmol/g). The scale of the synthesis of the glycopeptide was 0.05 mmol. The \( Fmoc \) group was removed by treatment with 20 % piperidine in NMP (5 and 15 min) between each coupling and deprotection step. Coupling reactions with \( Fmoc \) amino acids were done using 1.0 mmol (10 equiv.) of each of HBTU/HOBt and DIPEA in \( DMF \), for 30 min (Fastmoc\textsuperscript{TM} protocol). Between each couplings and deprotection, the resin was washed with DCM and DMF (5 min each).
7. Experimental section

General StBu deprotection.

DTT (100 mg) was dissolved in dry DMF (0.9 mL) and 2.5 % v/v DIPEA was added. After stirring for 5 minutes the solution was transferred to a peptide synthesis vessel containing resin-bound S'Bu protected peptide. After 48 h the resin was filtered and washed exhaustively with DMF, mixtures of DMF/H₂O (1:1), DMF and finally DCM. This procedure was then repeated once again.

General bromoacetamide couplings.

N-(Propargyl)-bromacetamide 112 (2 equiv. per thiol: 0.05 mmol x 2 thiols x 2 equiv.= 35 mg, 0.20 mmol) was dissolved in DMF (1 mL) and NEt₃ (42 μL, 0.6 mmol) and transferred to a peptide synthesis vessel containing resin-bound S'Bu deprotected peptide. The reaction was allowed to proceed for 24 h. After this time, the resin was filtered and washed exhaustively with DMF and the DCM.

Cleavage of compound 136.

The resin was washed with DCM and DMF five times (5 min each wash). The compound 136 was cleaved from the resin by incubation with the cleavage cocktail (95 % TFA, 2.5 % water and 2.5 % EDT) for 3 h. The compound 136 was precipitated with cold diethyl ether and centrifuged for 10 min.

Compound 136 purification.

The crude product was purified by reverse phase semi preparative HPLC (water-acetonitrile 5 %-95 % in 40 minutes, 0.1 % TFA), tᵣ = 25.5 min. The corresponding fractions containing 136 were pooled and freeze dried to afford 136 as a fluffy white solid (13 mg; 19 % yield).
Analytical data for compound 136: C$_{52}$H$_{82}$N$_{14}$O$_{20}$S$_{5}$. MW: 1350.47. Found: [M]$^+$: 1352.42; [M]$^{12}$: 676.74.

Synthesis of compound [138]

To a solution of the modified peptide 136 (1.5 mg, 1.04x10$^{-3}$ mmol) in a 0.1 M PBS buffer solution (pH 8.0, 800 µL), a PBS solution of the azido sugar 121 (1.0 mg, 4.15x10$^{-3}$ mmol, 41.00 mM, 100 µL) was added. Afterwards, sodium ascorbate (5.3 mg, 2.70x10$^{-2}$ mmol) and tris-benzyltriazolylmethyamine (TBTA) (14.3 mg, 2.70x10$^{-2}$ mmol) were added to the solution. The final suspension was sonicated for 5 minutes in a cold bath. Finally, a PBS solution of CuSO$_4$·5H$_2$O (0.3 mg, 1.04x10$^{-3}$ mmol, 10.40 mM, 100 µL) was added. The reaction was finally shaken for 24 hours at 20 °C and 1200 rpm. Then, the crude of the reaction was centrifuged at 14000 rpm for 5 minutes, and filtered through a celite plug in a Pasteur pipette. The crude product was purified by reverse phase semi preparative HPLC (water-acetonitrile 5 %-95 % in 40 minutes, 0.1 % TFA), t$_R$= 15.6 min. The corresponding fractions containing 138 were pooled and freeze dried to afford 138 as a fluffy white solid (1.3 mg; 65 %).

Analytical data: C$_{98}$H$_{110}$N$_{22}$O$_{36}$S$_{4}$. MW: 1842.66. Found: [M]$^+$ 1844.20, [M]$^{12}$ 923.10.
Synthesis of compound [139]

The modified cysteine alkyne peptide 136 (3.6 mg, 2.67x10^{-3} mmol) and 104 (14.3 mg, 9.38x10^{-3} mmol) were dissolved in a 0.1 M PBS buffer solution (pH 8.0, 1.9 mL). Sodium ascorbate (13.0 mg, 6.66x10^{-2} mmol) and tris-benzyltriazolylmethylamine (TBTA) (35.0 mg, 6.66x10^{-2} mmol) were added to the solution. The final suspension was sonicated for 5 minutes in a cold bath. Finally, a PBS solution of CuSO_{4}·5H_{2}O (0.7 mg, 2.67x10^{-3} mmol, 26.70 mM, 100.0 µL) was added. The reaction was finally shaken for 24 hours at 20 °C and 1200 rpm. Then, the crude of the reaction was centrifuged at 14000 rpm for 5 minutes, and filtered through a celite plug in a Pasteur pipette. The crude product was purified by reverse phase semi preparative HPLC (water-acetonitrile 5 %-95 % in 40 minutes, 0.1 % TFA), t_R=16.3 min. The fractions containing 139 were pooled and freeze dried to afford 139 as a fluffy white solid (t_R=16.1 min) (8.1 mg; 69 %); and a white fluffy white solid 104 (t_R=14.2 min) (5.1 mg; 59 % of recovered starting material).

Analytical data: C_{166}H_{260}N_{40}O_{44}S_{8}. MW: 4400.6. Found: [M]^{+3}: 1468.6, [M+K]^{+4}: 1111.5, [M]^{+4}: 1101.7, [M]^{+5}: 881.7, [M(scaffold-N3)]^{+2}: 741.9.
7. Experimental section

*p-Nitrophenyl-2, 3, 4, 6-tetra-O-acetyl-α-D-mannopyranoside, [151]*

\[ \text{p-Nitrophenyl mannopyranoside } 150 \text{ (200 mg, 0.66 mmol) was dissolved in acetic anhydride (1.3 mL) at 0 °C. Pyridine (650 μL) was added dropwise. The mixture was stirred overnight. After 16 h, TLC (1:1 ethyl acetate/chloroform) indicated that the reaction was complete. The solvent was removed under reduced pressure. The crude product was diluted with dichloromethane (50 mL) and washed with 1 M HCl (50 mL), saturated aqueous NaHCO₃ (50 mL) and water (50 mL); then was dried with MgSO₄, filtered and the solvent was removed under vacuum to afford } 151 \text{ (308 mg; 100 %) as a yellow solid. } R_f= 0.46 \text{ (1:1 ethyl acetate/chloroform). } \]

\[ \text{H-NMR (250 MHz, CDCl₃) } \delta \text{ (ppm): } 8.21 \text{ (2H, d, } J_{\text{ortho}} = 7.1 \text{ Hz, H-meta); 7.19 } \text{ (2H, d, } J_{\text{ortho}} = 7.1 \text{ Hz, H-ortho); 5.61 } \text{ (1H, d, } J_{\text{1-2}} = 1.8 \text{ Hz, H1); 5.55 } \text{ (1H, dd, } J_{\text{3-4}} = 10.0 \text{ Hz, } J_{\text{2-3}} = 3.5 \text{ Hz, H3); 5.45 } \text{ (1H, dd, } J_{\text{2-3}} = 3.5 \text{ Hz, } J_{\text{1-2}} = 1.8 \text{ Hz, H2); 5.37 } \text{ (1H, dd, } J_{\text{4-5}} = 10.0 \text{ Hz, H4); 4.26 } \text{ (1H, dd, } J_{\text{6a-6b}} = 12.0 \text{ Hz, } J_{\text{5-6a}} = 5.6 \text{ Hz, H6a); 4.05 } \text{ (1H, dd, } J_{\text{6a-6b}} = 12.0 \text{ Hz, } J_{\text{5-6b}} = 2.5 \text{ Hz, H6b); 3.98 } \text{ (1H, m, H5); 2.20, 2.04, 2.03, 2.01 } \text{ (12H, 4 x s, CH₃CO).} \]

\[ \text{C-NMR (63 MHz, CDCl₃) } \delta \text{ (ppm): } 170.3, 169.9, 169.6 \text{ (CO); 160.1, 143.1 } \text{ (Ar); 125.8, 116.4 } \text{ (Ar-H); 95.7, 69.8, 68.9, 68.5, 65.6 } \text{ (CH); 61.9 } \text{ (CH₂); 20.7 } \text{ (CH₃). ESI-MS: } m/z= 487.12 \text{ [M NH₄]^+}, m/z= 492.12 \text{ [MNa]^+}, C_{20}H_{25}NO_{12} \text{ requires } m/z= 469.12. \]
p-Aminophenyl-2, 3, 4, 6-tetra-O-acetyl-α-D-mannopyranoside, [152]

\[
\text{HCO}_2\text{NH}_4, \text{MeOH, Pd/H}_2
\]

\[
\text{151} \rightarrow \text{152}
\]

\[\text{p-Aminophenyl-2,3,4,6-tetra-O-acetyl-α-D-mannopyranoside 151 (308 mg, 0.66 mmol) was dissolved in 25 mL of MeOH containing ammonium formate (832 mg, 13.22 mmol) and 10% Pd/C (62 mg). The mixture was stirred at room temperature under nitrogen. After 1 h, TLC (1:1 ethyl acetate/chloroform) indicated that the reaction was complete. The crude product was filtered through celite and the solvent was removed under reduced pressure. The crude was diluted with dichloromethane (25 mL) and washed with water (25 mL); then dried with MgSO}_4, filtered and the solvent was removed under vacuum to afford 152 (289 mg; 100%) as a solid. Rf = 0.22 (1:1 ethyl acetate/chloroform).} \]

\[^1\text{H-NMR (250 MHz, CDCl}_3\text{) } \delta\text{ (ppm): 6.88 (2H, d, J}\text{-ortho} = 8.9\text{ Hz, H-ortho); 6.60 (2H, d, J}\text{-meta} = 8.9\text{ Hz, H-meta); 5.52 (1H, dd, J}_{3,4} = 10.0\text{ Hz, J}_{2,3} = 3.5\text{ Hz, H3); 5.40 (1H, dd, J}_{2,3} = 3.5\text{ Hz, J}_{1,3} = 1.8\text{ Hz, H2); 5.32 (1H, t, J}_{5,6} = J_{1,2} = 10.0\text{ Hz, H4); 5.30 (1H, d, J}_{5,6} = 1.8\text{ Hz, H1); 4.26 (1H, dd, J}_{6a,6b} = 11.8\text{ Hz, J}_{5,6a} = 5.0\text{ Hz, H6a); 4.12-4.08 (2H, m, H-5, H6b); 3.26 (2H, s, NH}_2; 2.16, 2.03, 2.01 (12H, s, COCH}_3).} \]

\[^13\text{C-NMR (63 MHz, CDCl}_3\text{) } \delta\text{ (ppm): 170.6, 170.0, 169.9, 169.7 (CO); 148.6, 142.0 (Ar); 118.0, 116.0 (Ar-H); 96.7, 69.5, 68.9, 68.9, 66.0 (CH); 62.1 (CH}_2; 20.8, 20.7 (CH}_3).} \]

ESI-MS: \(m/z = 462.15\) [MNa]+, \(C_{20}H_{23}NO_{12}\) requires \(m/z = 439.15\).
7. Experimental section

2-Iodoacetyl chloride,

![Chemical structure](image)

Thionyl chloride (79 μL, 1.08 mmol) was added dropwise with stirring to iodoacetic acid (167 mg, 0.90 mmol) and the resulting suspension was stirred at room temperature with exclusion of moisture and light. After 24 h the excess thionyl chloride was removed under reduced pressure to afford the acid chloride 153 (141 mg; 70 %) as an orange oil, which was used without further purification.

N-(2-Iodoacetyl)-p-aminophenyl-2, 3, 4, 6-tetra-O-acetyl mannopyranoside,

![Chemical structure](image)

A solution of 152 (289 mg, 0.66 mmol) in dry DCM (4 mL) was treated with pyridine (54 μL, 0.66 mmol) and cooled to -30 °C with exclusion of light and under nitrogen. Then, this solution was treated dropwise with 2-iodoacetyl chloride 153 (141 mg, 0.66 mmol) dissolved in anhydrous DCM (4 mL). After 0.5 h, the reaction was allowed to warm to room temperature. After a further hour, TLC (1:1 ethyl acetate/chloroform) indicated that the reaction was complete. The reaction was quenched with chloroform (50 mL) and the resulting solution was washed with 1 M HCl (40 mL), saturated aqueous NaHCO₃ (40 mL) and water (40 mL). The organic phase was dried over MgSO₄, filtered and concentrated.
under reduced pressure. The crude product was purified by flash chromatography over silica (1:1 ethyl acetate/chloroform) to afford 154 (276 mg; 69 %) as an orange solid. Rf = 0.36 (1:1 ethyl acetate/chloroform). $^1$H-NMR (250 MHz, CDCl$_3$) $\delta$ (ppm): 8.35 (1H, s, NH); 7.43 (2H, d, $J_{o-m}$ = 9.0 Hz, H-meta); 6.73 (2H, d, $J_{o-m}$ = 9.0 Hz, H-ortho); 5.50 (1H, dd, $J_{3,4}$ = 10.2 Hz, $J_{2,3}$ = 3.4 Hz, H3); 5.44 (1H, d, $J_{1,2}$ = 1.9 Hz, H1); 5.39 (1H, dd, $J_{5,6a}$ = 3.4 Hz, $J_{6a-6b}$ = 1.9 Hz, H2); 5.32 (1H, dd, $J_{3,4}$ = 10.0 Hz, H4); 4.24 (1H, dd, $J_{6a-6b}$ = 12.3 Hz, $J_{6a-6b}$ = 5.5 Hz, H6b); 4.07 (2H, m, H-5, H6b); 3.80 (2H, s, COCH$_2$I); 2.16, 2.02, 2.01, 2.00 (12H, 4 x s, COCH$_3$). $^{13}$C-NMR (63 MHz, CDCl$_3$) $\delta$ (ppm): 170.4, 169.9, 169.5, 165.5 (CO); 152.3, 132.7 (Ar); 121.5, 116.8 (Ar-H); 95.8, 69.1, 68.9, 68.7, 65.7 (CH); 61.9 (CH$_3$); 20.7, 20.5 (CH$_3$); -0.32 (COCH$_2$I). ESI-MS: $m/z$ = 625.06 [MNH$_4$]$^+$, $m/z$ = 630.06 [MNa]$^+$, C$_{22}$H$_{21}$NO$_{11}$ requires $m/z$ = 607.06.

$N$-(2-Iodoacetyl)-$p$-aminophenyl mannopyranoside, [155]

![Chemical structure](image)

The iodoacetamide 154 (276 mg, 0.46 mmol) as a solution in anhydrous methanol (36 mL) was treated with a solution of sodium methoxide (0.5 M in methanol (276 µL, 0.14 mmol)) under nitrogen with the exclusion of light and the resulting solution was stirred at room temperature for 3 h. After 3 h the reaction was quenched with a solution of acetic acid 0.1 M (140 µL) and concentrated under reduced pressure. The solid was redissolved in water and washed with ethyl acetate (2 x 10 mL) and finally the aqueous layers were freeze dried to afford 155 (184 mg; 91 %) as a pale brown solid. $^1$H-NMR (250 MHz, D$_2$O) $\delta$ (ppm): 7.33 (2H, d, $J_{o-m}$ = 9.0 Hz, H-meta); 7.10 (2H, d, $J_{o-m}$ = 9.0 Hz, H-ortho); 5.54 (1H, d, $J_{1,2}$ = 1.7 Hz,
7. Experimental section

H3); 4.13 (1H, dd, J1,2 = 3.3 Hz, J1,2 = 1.7 Hz, H2); 4.01 (1H, dd, J1,2 = 8.4 Hz, J2,3 = 3.3 Hz, H3); 3.85 (2H, s, COCH3); 3.77-3.67 (4H, m, H4, H5, H6a, H6b).

\[ ^{13} \text{C-NMR (63 MHz, D}_2\text{O)} \]

δ (ppm): 171.9 (CO); 154.5, 133.0 (Ar); 124.8, 119.1 (Ar-H); 99.9, 74.9, 71.9, 71.4, 68.1 (CH); 62.2 (CH2); 0.00 (COCH3). FAB-MS calculated for C14H18INO7 [MNa]+ 462.00257, found: 462.00351.

Ligations between Green Fluorescent Protein (GFP)-intein chitin binding domain to agarose beads and unnatural glycopeptides B and D

GFP-Sci Intein Binding Domain (CBD) Fusion Protein was overexpressed in Escherichia coli strain BL21-(DE3) transformed with pTYB1-GFP. The cell free extract was loaded on chitin beads, New England Biolabs (NEB), and equilibrated with column buffer (10 mL, 20 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM EDTA, 1 mM PMSF). The cell-free extract loaded was provided as a gift by Dr. D. J. P. Richardson.

The beads were washed with 200 mM sodium phosphate solution, pH 8, 100.0 mM sodium chloride, 2 % w/v MESNA, 5 mM TCEP solution (5 mL). The column was allowed to run by gravity and then plugged. Finally, the desired unnatural glycopeptide (165: 4.2 mg, 9.54×10⁻⁴ mmol; 169: 1.6 mg, 3.64×10⁻⁴ mmol) was dissolved in the previous sodium phosphate buffer solution (100.0 μL), and the contents were mixed gently for 48 h. After 2 days of reaction, the crude was eluted with 200 mM sodium phosphate solution, pH8, 100 mM sodium chloride, 5 mM TCEP solution (2 mL). The eluted samples were concentrated in a Vivaspin® 2 mL concentrator, 30,000 molecular weight cut-off (MWCO). Afterwards, the green precipitate was dialysed with 4 mL of Dulbecco’s PBS solution, pH 7.4. Finally, the product was resuspended with 200 μL of Dulbecco’s PBS solution, pH 7.4.
Dulbecco’s PBS solution, pH 7.4.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NaCl</td>
<td>8 g</td>
</tr>
<tr>
<td>KCl</td>
<td>0.2 g</td>
</tr>
<tr>
<td>Na₂HPO₄</td>
<td>1.15 g</td>
</tr>
<tr>
<td>KH₂PO₄</td>
<td>0.2 g</td>
</tr>
<tr>
<td>CaCl₂</td>
<td>0.1 g</td>
</tr>
<tr>
<td>MgCl₂·6H₂O</td>
<td>0.1 g</td>
</tr>
<tr>
<td>H₂O</td>
<td>1000 mL</td>
</tr>
</tbody>
</table>

All the salts were mixed together and made up to 1000 mL of water.

Interaction tests between unnatural glycoprotein and agarose beads coated with *Ricinus Communis RCA₁₂₀*

50 µL of the re-suspended glycoprotein in 200 µL of Dulbecco’s PBS solution, pH 7.4, was dispatched in a 1.5 mL eppendorf tube. Parallel to this experiment was dispatched the same amount in the same buffer system GFP. A digital photograph was taken under UV light (λ:365 nm). A small amount of agarose beads coated with *Ricinus Communis RCA₁₂₀* was added to each eppendorf, and it was shaken gently for 30 min. Afterwards, a second digital photograph was taken under UV light (λ:365 nm). Finally, the agarose beads coated with *Ricinus Communis RCA₁₂₀* were washed twice with 50 µL of Dulbecco’s PBS solution, pH 7.4. Then, a third digital photograph was taken under UV light (λ:365 nm).
8 Acknowledgements

My thanks to Dr. Derek Macmillan for offering me the chance to work in his research group and for his guidance during my PhD. Moreover, I am grateful to him for performing the last experiments on solid phase of the neoglycopeptide linkage (Chapter 5).

My thanks to the University of Edinburgh for funding my PhD studies, and to University College London (UCL), for providing me extra-funding during my period in London.

My thanks to Dr. Lubna Sadiq and Dr. David W. Anderson for sharing good moments in the lab 25D, at the University of Edinburgh. Moreover, my thanks to Dr. Jonathan P. Richardson for expressing GFP, helping me with EPL and test interaction studies with lectin beads and proof reading the introduction; Mr. Jin Massania for the production of TBTA for “click chemistry”; and the Italian connection: Dr. Steffano Mezzato, for his sympathy, and specially, Miss Silvia Marchesan, for being very helpful, and for her unlimited support.

My thanks to Mr. John Millar and Dr. Abil Aliev; and Mr. Robert Smith, Dr. Lisa Harris and Mr. John Hill, for all the help with the spectral analysis.

My thanks for Pepe, Tere, Marivi, Isabel, Alejandro, Ana Pizarro, José, Vicente, Jone, Fred, Manu, Mike, Stefanie, Michele, Emiliano, Emilio, Iria, Jesús, Ana, Inma, Rafa, Andriy, Larisa, Shirin, Sofia, George Lucas, Nerea, Quim and Laura, Rob, Graeme and Stephen, etc..., and Pablito “Metroooo”, for their help and the good moments we had together.
8. Acknowledgements

My thanks to my friends from Spain: Ignacio, Gloria, Juan, Sonia, Jaime, Queralt, Albert, Laura, Paco, Marina, Nacho, Javi, Mamen, Natàlia, Jose "Pinoporro", "Xevi", Joan Josep and Eva; and my new little friends: Alicia, Laia, Ainara and Nil; for giving me a great and unconditional friendship. "¡Por las buenas risas!".

And the most important of all, my special thanks to my parents, my two sisters: Begoña and Beatriz, my grandmother, my brother in law, David, my uncle Joaquin, my family from Barcelona, Valencia and Madrid, and my little niece: Asuntita; for the help, advice, patience and the never ending support.
A novel neoglycopeptide linkage compatible with native chemical ligation†

Derek Macmillan* and Javier Blanc

Received 22nd May 2006, Accepted 25th May 2006
First published as an Advance Article on the web 3rd July 2006
DOI: 10.1039/b607200c

The straightforward synthesis of a novel class of neoglycopeptide and its fusion with a larger peptide chimera using sequential chemo-selective ligations is described.

The synthesis of glycosylated proteins is a formidable challenge and it is fair to say that synthetic glycoproteins have had modest impact, thus far, on the study of glycobiology when compared to genetic methods. This is largely due to the well documented difficulties associated with synthesis of oligosaccharides and glycopeptides such as the requirement for extensive protection and deprotection strategies and the stereoselective formation of glycosidic linkages. We, and others, have pursued alternative methodologies for the assembly of modified peptides and proteins.1

Using the click "ligation"2 for bioconjugation was of particular interest due to its reliability and compatibility with the aqueous milieu.3 We had previously shown that glycosyl iodoacetamides can react with the sulfhydryl groups of cysteine on solid-phase to afford glycopeptide mimetics and that this methodology was compatible with useful protein assembly methods such as native chemical ligation (NCL).4 We were keen to investigate whether the coupling of glycosyl azides with peptides displaying acetylene dipolarophiles may also be useful for this purpose since the glycosyl azides are more stable and can be prepared in fewer steps than the corresponding glycosyl iodoacetamides (Scheme I).

Initially we prepared the peracetylated glycopyranosylazides of N-acetylglucosamine (1), N-acetyllactosamine (2) and chitobiose (3) which are all constituents of the N-linked class of glycoproteins. We investigated conditions for their union with the heterobifunctional adaptor 2-bromoacetyl propargylamide (4) and the reaction of these saccharides with 4 proceeded smoothly under conditions reported in the recent literature:5 the 1,4-addition product being favoured in the presence of a Cu(I)catalyst (Table I). The crude products did not require purification by column chromatography.

Pleased with the facility with which such glycoconjugates could be prepared we next aimed to expose the bromoacetamides to conditions typically encountered in peptide chemistry and native chemical ligation. We were particularly interested in the potential to modify the thiol groups of cysteine residues. Additionally, to demonstrate the versatility of the approach, we explored the possibility of reacting thiol groups directly with the bromoacetamide products 5-7 on solid phase and the reaction of cysteine thioesters with 4 such that click chemistry could subsequently be investigated in solution or on solid-phase with peptides displaying acetylenes

Scheme 1 Construction of novel neoglycopeptides.

Table 1 Synthesis of neoglycopeptide precursors bromoacetamides

<table>
<thead>
<tr>
<th>Saccharide</th>
<th>Catalyst</th>
<th>Product</th>
<th>Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. R = Ac</td>
<td>Cu(I) (5 eq.)</td>
<td>5</td>
<td>100</td>
</tr>
<tr>
<td>1. R = Ac</td>
<td>Cu(I)SO, (0.1 eq.)</td>
<td>5</td>
<td>97</td>
</tr>
<tr>
<td>2. R = (D alternate Glc-Ac- Cu(1)SO, (0.1 eq.)</td>
<td>6</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td>3. R = (D alternate Glc-Ac- Cu(I)SO, (0.1 eq.)</td>
<td>7</td>
<td>87</td>
<td></td>
</tr>
</tbody>
</table>

*Methanol as solvent. † Active copper species generated in the presence of 1.1 equiv. of sodium ascorbate in H-O or H2O as solvent. ‡ Isolated yield.

(Scheme 2). As anticipated, 4 and 5 reacted cleanly with benzyl mercaptoaniline forming model thioureas 8 and 9 in 84% and 75% yield respectively. 8 also reacted cleanly with peracetylated 2-acetamido-2-deoxy-d-glucopyranosyl azide, affording 9 in 91% yield.
yield. To establish whether the products might be stable to the usual acidic peptide cleavage conditions 9 was subjected to 95% aqueous TFA for 3 h. NMR analysis of the crude material after evaporation showed no decomposition had taken place. Finally the acetyl esters were cleanly removed upon exposure to 2% v/v hydrazine hydrate in EtOH for 72 h and the fully deprotected compound 10 was obtained. Encouraged by the preliminary results we assembled a peptide fragment (11), similar in sequence to human erythropoietin (residues 21–32), plus an N-terminal cysteine residue, and furnished with two disulfide bond protected cysteine residues at predetermined positions (Scheme 3). The peptide was assembled using standard protocols for Fmoc solid-phase peptide synthesis and in an automated fashion. The cysteine residues were deprotected on solid-phase by exposure to 10% w/v dithiothreitol (DTT) containing 2.5% v/v DIPEA to expose the thiol functional groups. N-Acetylglucosamine and the disaccharide chitobiose were then incorporated by exposure of the resin to bromoacetamide 5 or 7, employing three equivalents 5 or 7 per thiol in each reaction. After 16 h reaction at room temperature, cleavage of a small resin sample indicated that the reaction was complete as the starting material was not observed. After cleavage from the solid support by treatment of the resin with 95% TFA, 2.5% ethanediol and 2.5% H2O for four hours the crude products were purified by semi-preparative HPLC, lyophilized, and treated with 2% v/v aqueous hydrazine hydrate containing 5% w/v DTT to obtain the fully deprotected products 12 and 13 in quantitative yield (determined by HPLC). Bromoacetamide 6 and acetylenic bromoacetamide 4 could also be incorporated into synthetic peptides in an identical fashion.

Scheme 2 Model reactions with benzylisocyanate demonstrate that the order in which the reactions take place appears unimportant. Reagents and conditions: i) 10% w/v DTT, 2.5% DIPEA, DMF, 16 h, ii) 5 or 7 (3 eq. per thiol), 2.5% w/v DIPEA, 2.5% ethanediol, 2.5% H2O, 4 h, iii) 2% v/v aqueous hydrazine monohydrate, 1 h.

Scheme 3 Reagents and conditions: i) 10% w/v DTT, 2.5% DIPEA, DMF, 16 h, ii) 5 or 7 (3 eq. per thiol), 2.5% w/v Et3N, DMF, 16 h, iii) 95% TFA, 2.5% ethanediol, 2.5% H2O, 4 h, iv) 2% v/v aqueous hydrazine monohydrate, 1 h.
Fragment 13 was then coupled to a peptide thioester in a native chemical ligation reaction. The construction of the peptide thioester, corresponding to human erythropoietin residues 1-19, and its release from the solid support were monitored using the dual-linker approach recently described by Unger and coworkers. In the ligation reaction equimolar quantities of each peptide were combined in 0.25 ml of 6 M guanidine hydrochloride containing 80 mM sodium phosphate buffer, pH 8.0, 1% w/v mercaptoethanesulfonic acid (MESNA) and 10 mM tris(carboxymethyl)phosphine (TCEP) for 36 h with shaking at room temperature (Scheme 4). After this time the reaction mixture was purified by directly loading it onto a semi-preparative HPLC column. The ligated product (14) was the only species observed by HPLC.

In summary we have developed a novel class of neoglycopeptide that is compatible with modification of cysteine mutant proteins, with synthetic peptides, and native chemical ligation. Furthermore the fusion of glycosyl azides with peptides displaying acetylenes may be of particular interest since NCL has been shown to fail when large oligosaccharide appendages are located proximal to the ligation site. Therefore, if acetylenes can be installed so as to "encode" for glycosylation then bulky saccharide motifs may be installed after ligation using click chemistry in solution. We have already observed that the reaction between peptides displaying acetylenes derived from 4 and glycosyl azides proceeds on solid-support. Although the linkage between the carbohydrate and peptide moiety is unnatural, difficulties associated with the synthesis of native glycopeptides and glycoproteins dictates that new modes of presentation of carbohydrates (or other appendages) should be explored concurrently. Such neoglycoconjugates may improve the pharmacokinetic profile of therapeutic glycoproteins. The fact that the neoglycopeptides described are simple to prepare, homogeneous, and are also compatible with native chemical ligation, may render them attractive building blocks for neoglycoprotein assembly.

Acknowledgements
The authors would like to acknowledge The Royal Society, The BBSRC, The University of Edinburgh and EUC for financial support.

Notes and references
Annexe 1: Published work


7 See supporting information for HPLC and mass spectrometry data for products 12 and 13 and NCL product 14.


Annexe 2: References

(1) Varki, A. Glycobiology 1993, 3, 97-130.


(3) Davis, B. G. Science 2004, 303, 480-482.


(9) Fukuda, M. Biochimica et Biophysica Acta 2002, 1573, 394-405.


(40) www.bve.virginia.edu/lev/images/map.gif.
(50) Koenigs, W.; Knorr, E. Berichter 1901, 34, 957-981.
(64) Unverzagt, C. Angew. Chem. 1994, 106, 1170-1173 (See also Angew Chem, Int Ed Engl, 1994, 33(10), 1102-1104).
Annexe 2: References

(121) Grundler, G.; Schmidt, R. R. Liebigs 1894, 1826-1847.


Schon, I.; Kisfaludy, L. *Synthesis* 1986, 303-305.


Debehan, J.; Rodebaugh, R.; Fraser-Reid, B. Liebigs 1997, 791-802.


Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. Angew. Chem. Int. Ed. 2002, 41, 2596-2599.


