The Development of *S*-Glycosylcysteine Derivatives for use in Glycan-Binding Assays

By

Matthew Williams

A dissertation submitted in fulfilment of the requirement for the degree of

Master of Science

In the Department of Chemistry

University of Cape Town

February 2017

Supervisor: Assoc. Prof. D. W. Gammon

Declaration

I, Matthew Williams, know the meaning of plagiarism and declare that all of the work presented in this document "**The Development of *S*-Glycosylcysteine Derivatives for use in Glycan-Binding Assays**" is my own work and to the best of my knowledge has never been submitted for examination for any degree at any university. All sources of information used herein are cited, acknowledged and fully referenced at the end of the document.

Matthew Williams: ..... Date: .....

To my parents Blendynn Peter Williams and Tracey Jill Williams for their love, support and countless opportunities afforded to me throughout my life

Acknowledgements

First and foremost, I would like to thank my supervisor, Assoc. Prof David. W. Gammon for this opportunity which has allowed me to better my understanding of synthetic organic chemistry and allowed me to grow as person. I also thank him for his patience, knowledge and guidance throughout the duration of this project. The most memorable lesson that he taught me is that it is often the things that we are not looking for that are the most interesting.

I would also like to thank everyone who has helped me throughout the course of this project and in particular the following people:

Dr. Sophie Reese-Jones and Deirdre Brooks for their constant support, help and advice.

Arthur Joseph, Yaasien Ely, Andre de Jager and Pieter de Kock for their assistance in the maintenance of lab equipment and overall upkeep of the lab. Peter Roberts, Gianpiero Benincasa and members of the CAF, at Stellenbosch University, for their analytical services.

Dr. Aloysius Nchinda, Dr. Taigh Anderson, Dr. Malkeet Kumar, Lutete Khonde, Shakeela Sayed, Marwaan Rylands, Shankari Nair, Daniel Kusza, Fabrizio L'abbate, Tsekiso Ranqhai, Zaheer Timol and Kirstin Anderson for their assistance in the lab and many helpful discussions.

My friends who decided to take this journey with me; Chad Wilkinson, Muhammad Motala, Chyanne Abbot, Jasmin Ferreira, Linley Barnard, Jessica Akester, Irwin Cassells and Hendrik Els for their friendship. Rukaya Mansoor, Scott de Beer, Francisco Pêra and Niccolò Passerin for their unfailing support and motivation during the tough times.

My chemistry brother-in-arms, Stefan Benjamin, who deserves a special mention for his friendship inside and outside the lab over the past six years. His constant support, humour, goodwill nature, and sound chemistry has made this journey all that much sweeter, and I am a better man for knowing him.

My brothers Dayne Williams and Jonathan Williams for their hard work ethic, which has always pushed me to work harder in the lab, and for always motivating to pursue my further studies.

"A smooth sea never made a skilled sailor" - Franklin D. Roosevelt

Abstract

This dissertation concerns the development of a synthetic route towards novel cysteine-based glycan-binding probes, for incorporation into glycoarrays and or similar applications used in assays of glycan-recognition phenomena.

The need to systematically characterize the glycome and decipher the range of glycosylation patterns found in living cells, has prompted the development of molecular tools such as glycoarrays and related systems for immobilizing defined carbohydrate structures. The preparation of these probes requires access to building blocks where the core structure has defined glycans together with appropriate linkers, and the amino acid cysteine is explored here as one such structure.

In particular, this dissertation describes the synthesis of a *S*-glucosylcysteine derivative SGC, or methyl *N*-(6-aminohexanoyl)-*S*-(β-D-glucopyranosyl)-L-cysteinate trifluoroacetate **67**, as well as its 2-acetamido analogue SAGC, or methyl *N*-(6-aminohexanoyl)-*S*-(2-acetamido-2-deoxy-β-D-glucopyranosyl)-L-cysteinate trifluoroacetate **74**.

The first approach involved initial preparation of *N*-(4-azidobutanoyl)-L-cysteine **12** and attempted reaction of this with 1,2,3,4,6-penta-*O*-acetyl-β-D-glucopyranose **3** to form the initial target of this dissertation, bis-glucoside **13**. This was not successful, but repetition of the reported reaction involving the use of *N*-acetyl-L-cysteine **4** provided a modest yield of partially purified bis-glucosyl cysteine (BGC, **1**). A mechanism for this one-pot, sequential bis-glucosylation is proposed.

The limitations of the one-pot procedure led to investigation of alternative methods for the step-wise introduction of sugar units to the cysteine core. For this purpose the cysteine derivative, methyl *N*-(6-((*tert*-butoxycarbonyl)amino)hexanoyl)-L-cysteinate **40**, was prepared and reacted with **3** to obtain a fully protected precursor of the target SGC. However, inefficiencies in this procedure led to investigation of an alternative strategy for preparation of SGC.

This involved initial preparation of 2,3,4,6-tetra-*O*-acetyl-1-thio-β-D-glucopyranose **21**, reaction of this with methyl (*R*)-2-((*tert*-butoxycarbonyl)amino)-3-iodopropanoate **60**, followed by *N*-Boc deprotection and then formation of an *N*-acyl derivative by reaction with 2,5-dioxopyrrolidin-1-yl 6-((*tert*-butoxycarbonyl)amino)hexanoate **45**. A final de-*O*-acetylation and *N*-Boc deprotection gave the target SGC **67**. A similar sequence starting with 2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-1-thio-β-D-glucopyranose **71** gave the SAGC analogue **74**.

Table of Contents

Abbreviations

Ac	acetyl
BAGC	bis-(*N*-acetylglucosaminyl)-L-cysteine
BGC	bis-glucosyl cysteine
Bn	benzyl
Boc	*tert*-butoxycarbonyl
bs	broad singlet
Bu	butyl
CAF	Central Analytical Facilities
CAS	ceric ammonium sulfate
cat.	catalytic
CBP	carbohydrate-binding protein(s)
CERMAV	Centre de Recherches sur les Macromolecules Végétales
conc.	concentrated
COSY	correlation spectroscopy
CPE	carbohydrate-processing enzyme
CPTH	catalytic phase transfer hydrogenation
Cy	cyclohexyl
d	doublet
DCC	N,N'-dicyclohexylcarbodiimide
DCM	dichloromethane
dd	doublet of doublets
ddd	doublet of doublet of doublets
DMAP	4-dimethylaminopyridine
DMF	dimethylformamide
DMSO	dimethyl sulfoxide

DMTST	dimethyl(methylthio)sulfonium triflate
DNA	deoxyribonucleic acid
dt	doublet of triplets
dq	doublet of quartets
EDC.HCl	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride
eq.	equivalents
ESI	electron spray ionization
Et	ethyl
FGI	functional group interconversion
GA	glycosyl acceptor
GalCer	galactosylceramide
h	hour(s)
Hex	hexane
HIV	human immunodeficiency virus
HOBt	1-hydroxybenzotriazole
HRMS	high-resolution mass spectrometry
HSQC	heteronuclear single quantum correlation
Hz	hertz
iNKTC	invariant natural killer T cells
IR	infrared
IUPAC	International Union of Pure and Applied Chemistry
lit.	literature
m	multiplet
Me	methyl
min	minute(s)
m.p.	melting point
MS	mass spectrometry

NCL	native chemical ligation
NHS	*N*-hydroxysuccinimide
NMR	nuclear magnetic resonance
Ph	phenyl
ppm	parts per million
q	quartet
r.t.	room temperature
s	singlet
SAGC	*S*-(*N*-acetylglucosaminyl)-L-cysteinate
sat.	saturated
SGC	*S*-(glucosyl)-L-cysteinate
SPR	surface plasmon resonance
SPRi	surface plasmon resonance imaging
t	triplet
TBAHS	tetrabutylammonium hydrogen sulfate
tert	tertiary
Tf	trifluoromethanesulfonyl
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin layer chromatography
TMS	trimethylsilane
TMSOTf	trimethylsilyl trifluoromethanesulfonate
ToF-SIMS	time-of-flight secondary ion mass spectrometry
TS	transistion state
TUT	Tshwane University of Technology
UATR	universal attenuated total reflectance
UV	ultraviolet

Chapter One

1. Introduction

1.1. The importance of glycoscience

In the biological scientific community a great emphasis has been placed on the fields of genomics and proteomics to better understand the intricacies of a living organism. In essence they hope to achieve a better understanding of the central dogma, which is the flow of genetic information from nucleic acids to fully functional proteins. Despite this, more than 50 % of all proteins are found as glycoproteins (a generic term for carbohydrate-protein conjugates).[1] This is manifest in a range of glycan structures such as proteoglycans, glycolipids and glycoproteins on the surfaces of or associated with living cells.[2,3] In turn these carbohydrate structures, or glycans, affect various facets of their conjugated proteins character such as folding, stability, trafficking and secretion, all of which contribute to a biologically viable protein.[1] Furthermore specific glycosylation patterns on the surfaces of cells can serve as biomarkers to identify either the early onset or stage of certain diseases. This is exemplified in diseases such as cancer, diabetes and neurodegenerative disorders, all of which feature unnatural glycosylation in the disease state.[2,4] Hence, if the structural identities of these carbohydrate molecules are known, then their role in either the promotion or inhibition of protein function and disease can be better assessed.

The role of these glycans is the domain of glycoscience; this has emerged as a versatile field that strives to better understand the structural and functional properties of glycans in biological systems. Furthermore, this research field can be broken down into the two subdisciplines of glycobiology and glycochemistry, where the former, despite enormous progress, is limited by the current molecular tools available to it. This together with difficulties encountered in the characterization of complex glycan structures and the availability of pure carbohydrates is currently hindering the development of this field.[2,5] Overcoming these challenges will require new and refined molecular tools, structures designed for specific tasks and ones that are robust enough to address current and future challenges in the field of glycoscience.

1.2. A brief account of the human glycome

The structural diversity of the cellular glycome dwarfs both the genome and proteome and is estimated to consist of approximately 100,000-500,000 unique glycan structures.[3] The existence of this diversity of structures is primarily attributed to two factors. Firstly, the cell specific expression of different glycosyltransferases, enzymes which mediate the non-template biosynthesis of glycans, results in different glycans being expressed on the cellular surface of specific cells within an organism.[3,6] Secondly, nucleic acid and protein biomolecules only occur as linear structures, whereas carbohydrate biomolecules can occur as both linear and branched structures, with monomers linked in either α- or β-configurations and thus drastically increasing the number of possible glycan structures. One such structure is depicted (Figure 1) as an illustrative example.[6] In addition to the above, the biosynthesized glycans will often suffer from incomplete glycosylation and undergo post-biosynthetic modifications such as acetylation, sulfation and phosphorylation at the terminal ends, further complicating the number of glycan structures.[3,6] Thankfully, the relevant interactions that take place between complex glycan structures and their targeted antigens often only take place at the terminal ends of the unique glycan structures, thus greatly reducing the number of structural variations that need now be considered. More, specifically, these binding sites are referred to as the glycan determinant.[3,6]

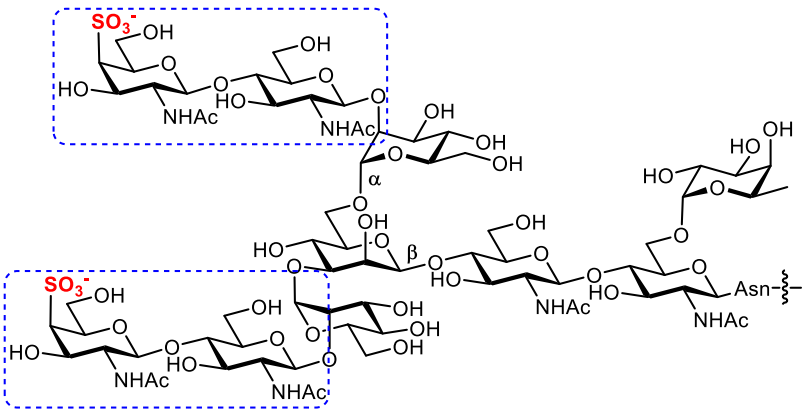
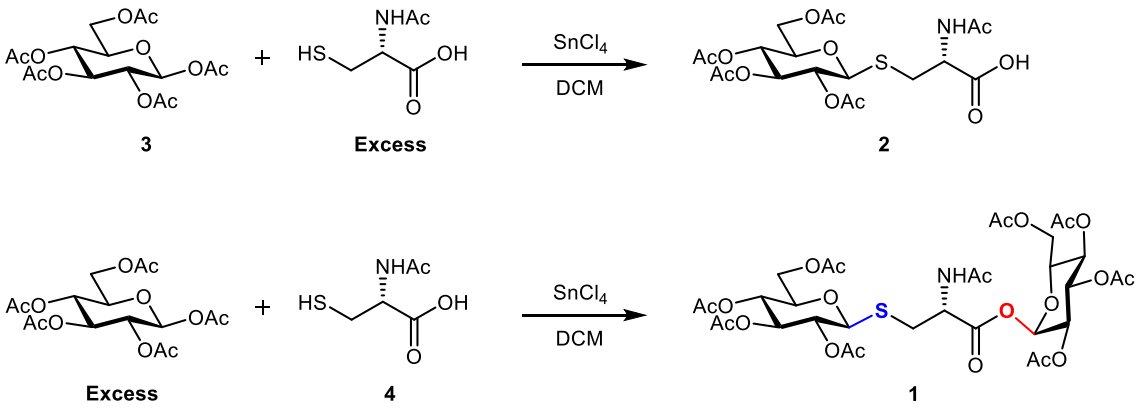


Figure 1: The 4-sulfated LDN biantennary *N*-glycan used as an illustrative example to highlight the complexity of glycans found in the human glycome. The sites at which post-sulfation took place (highlighted in red) and the glycan determinant (encircled in blue) of the biantennary *N*-glycan are also shown.[6]

In a recent review by *Cummings* (2009)[6] the human glycome (which encompasses both glycoproteins and glycolipids) has been estimated to consist of approximately 7,000 unique glycan determinants. A more in depth account of the above glycans and their subclasses is beyond the scope of this project, which has its basis on the serendipitous discovery of a structure which has potential for a range of specific uses in glycoscience.

1.3. A serendipitous discovery

As stated earlier, the advent of new molecular tools is often spurred on by a particular challenge faced in biology. However, at times empirical findings tend to spur on the development of molecular tools that may or may not address a particular issue, such as disease. Such is the case for the focus of this project. A serendipitous discovery by researchers at the Tshwane University of Technology (TUT) gave rise to a novel molecule termed bis-glucosyl cysteine* (BGC, **1**) depicted in Scheme 1.[7] This discovery took place during work conducted by *Nokwequ et al.* in which they were preparing analogues of mycothiol, a metabolite implicated in the defensive strategies of *Mycobacteria*. More specifically, they were attempting to optimize the yield of *N*-acetyl-*S*-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-L-cysteine **2** by reacting an excess 1,2,3,4,6-penta-*O*-acetyl-β-D-glucopyranose **3** with *N*-acetyl-L-cysteine **4** in the presence of $SnCl_4$ as shown in Scheme 1.



Scheme 1: The reaction conditions required to form **1** and **2** at room temperature. The common β-thioglucoside linkage (highlighted in blue) and unusual α-glucosyl ester linkage (highlighted in red) of **1** are also shown.[7]

*Note in instances where a starting reagent or targeted molecule (in this study) is not referred to by their IUPAC name when first introduced, in text, then their IUPAC name will be stated here in the footer.
IUPAC name: *N*-acetyl-*S*-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranosyl)-L-cysteinate (**1**)

Rather than obtaining the desired product **2** in higher yields, it was found that these conditions lead to the serendipitous formation of **1** instead. The structure of **1** seems unique, in that it incorporates both a β-thioglucoside linkage and a rather unusual α-glucosyl ester linkage. This is of interest as the stereoselective formation of a 1,2-*cis*-α-glucosyl ester linkage remains a significant challenge in glycochemistry and the one-step installation of two sugar units in the same molecule in two different configurations is unprecedented. Thus in the ensuing discussion, an argument is developed for the desirability of embarking on an in depth investigation for the use of the BGC scaffold as an enabling tool in the field of glycoscience.

1.4. Potential for BGC as an enabling tool in the field of glycoscience.

BGC, or more specifically its de-*O*-acetylated analogue, is a novel structure in a biologically important class of molecules, the glycopeptides. To date relatively little interest has been shown in the use of cysteine as a scaffold in synthetic glycopeptides such as glycoclusters, molecular tools used to study the multivalent effects of carbohydrate-protein interactions.[8] Two such clusters are the Lindhorst **6** and Dondoni **7** glycoclusters, shown alongside a potential BGC glycocluster **5** in Figure 2.[8,9]

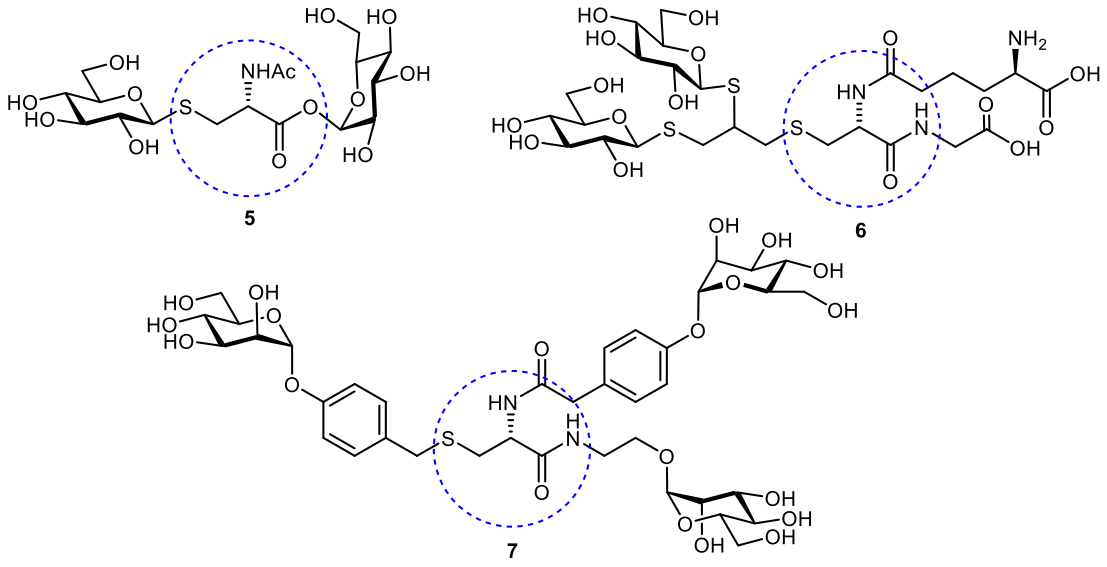


Figure 2: A comparison of a potential de-*O*-acetylated BGC glycocluster **5** with selected, cysteine-based glycoclusters **6** and **7**. The common cysteine core (encircled in blue) is also shown.[8,9]

In addition to its potential use as a glycocluster, BGC also bears a striking resemblance to other important glycans, prompting speculation that BGC could be a template for new bioactive analogues. In particular, it shows similarities to a family of glycolipids termed glycosylceramides that *inter alia* help regulate the immune system *via* their interaction with invariant natural killer T cells (iNKTC).[10] The basic structure of a glycosylceramide consists of a phytosphingosine ceramide backbone that is linked to a glycan in either an α- or β-configuration. The most extensively studied of these glycosylceramides is α-galactosylceramide (α-GalCer, **8**) depicted in Figure 3.[10] This is partly due to its potent anti-tumour activity but primarily due its recognition by iNKTC only, as it is a CD1d-restricted antigen. These iNKTC and **8** have since been implicated in a variety of pathogenic infections and auto-immune diseases; such as lupus, multiple sclerosis, asthma and type I diabetes. In addition, structure activity relationships conducted on **8** have highlighted that the presence of a hydroxyl group at C-3' and long acyl lipid tail to be crucial for its anti-tumour activity.[10] With this in mind, a modified version of BGC is re-imagined (Figure 3), in which global deacetylation and the addition of a lipophilic tail to the nitrogen of BGC has taken place to afford BGC ceramide **9**. This rendered derivative **9** bears a resemblance to **8**, making it an ideal candidate to be probed for anti-tumour activity and stimulation of iNKTC.

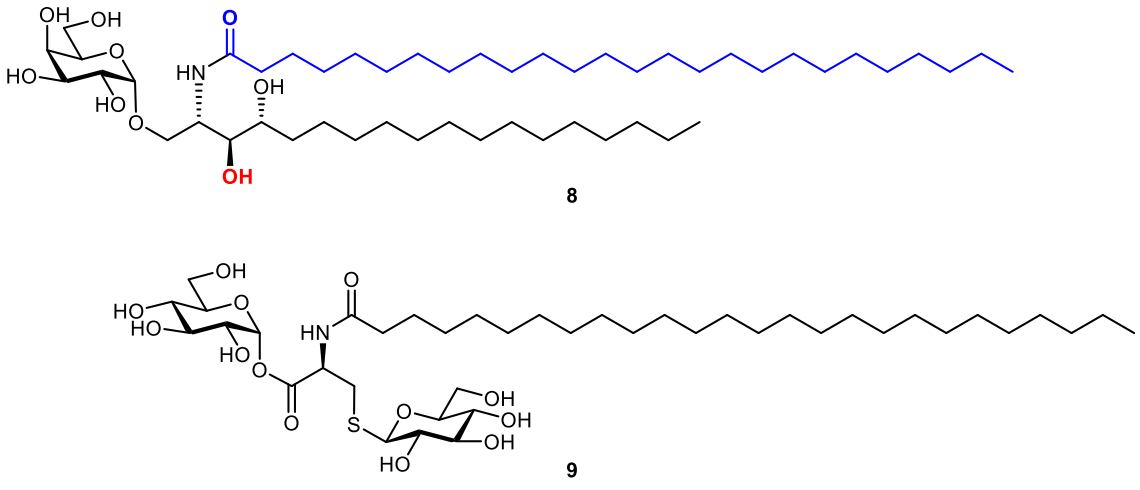
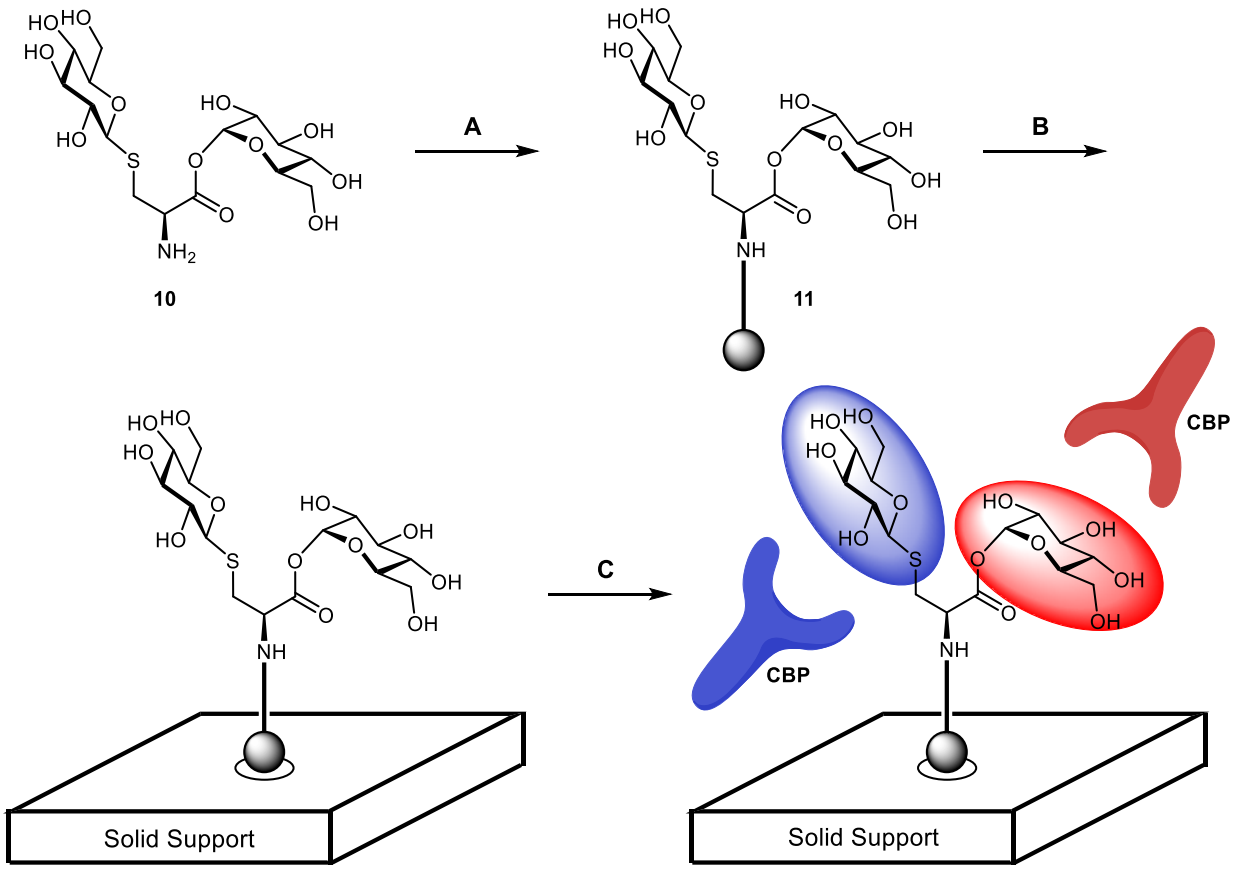


Figure 3: A re-imagined form of BGC as a BGC ceramide **9**; depicted together with the foremost studied glycosylceramide, α-GalCer **8**. The C-3' hydroxyl group (highlighted in red) and the acyl lipid tail (highlighted in blue) play a crucial role in the efficacy of α-GalCer **8** as an anti-tumour agent.[10]

One of the most well established tools currently used for mapping glycan-binding specificity are glycoarrays. They are generally prepared by taking defined glycans bearing suitable terminally-functionalized tethers and binding them onto an active surface or biopolymer.[2–4] This is followed by overlaying them with specific glycan-binding entities and the subsequent visualization of binding and recognition phenomena using a host of spectroscopic techniques. An illustration of this is provided (Scheme 2), which depicts the possibility of using deacetylated BGC derivatives in glycoarray fabrication. The first process (**A**) involves taking a globally deacetylated form of BGC **10** and attaching a suitable tether to the free amino group to afford the glycan-binding probe **11**. The second process (**B**) then involves the covalent attachment of **11**, having an appropriate functional group at the tether terminus, to an appropriate surface. The final process (**C**) will then involve screening a variety of carbohydrate-binding proteins (CBPs) and assaying any binding specificities to either one of the sugar units.



Scheme 2: Processes **A**, **B** and **C** required for the conversion of a modified form of BGC **10** into a novel glycoarray.

With all of the above in mind, the potential scope of the BGC scaffold as a new and innovative enabling tool in glycoscience is indeed a desirable one. A common feature in the above-mentioned examples is the need for the attachment of the tether to the amine functionality in cysteine, such as through an acyl linkage. Thus this project will focus on the establishment of efficient methods for the preparation of *N*-acyl cysteine derivatives such as **11** depicted above in Scheme 2. Once this has been achieved, and the possibility of attachment to surfaces has been demonstrated, it will be possible to extend the work to the development of novel glycoarrays and testing these for use in carbohydrate-recognition assays, in collaboration with research groups at the Alberta Glycomics Centre in Canada and the Centre de Recherches sur les Macromolecules Végétales (CERMAV) in France. This would of course require some significant further developments, including establishing the possibility of preparing analogues of **11** which incorporate a range of sugar units other than glucopyranose.

1.5. Glycoarrays an indispensable tool in glycoscience

Glycoarrays are formally defined as two-dimensional microarrays that are composed of structurally different glycans adhered to a solid support surface.[2] As stated earlier, the enormity of the glycome poses a massive obstacle in systematically deciphering the abundance of glycosylation patterns found on or associated with living cells. Up until the early 21st century, glycoscientists relied heavily upon biochemical tools in order to study theses glycosylation patterns; as the genetic tools available at the time gave limited information in glycomics.[2,4] However, these tools were often time and material consuming, limiting the rate at which patterns could be deciphered.[4] With no immediate solution, at hand, glycoscientists turned their attention to developing a glycan-based microarray (glycoarray) technology; due to the success of DNA and recombinant-protein microarray platforms in genomics and proteomics respectively.[3] This choice in technology paid dividends, as glycoarrays provided a new high-throughput method that allowed for the systematic screening of carbohydrate-processing enzyme (CPE) activities and CBP interactions both quantitatively and qualitatively.[4] In addition to this, glycoarrays also circumvented the need for copious amount of glycan material, thus effectively allowing for the demands of glycobiologists to be met by synthetic glycochemists.[2,4]

Since their inception glycoarrays have found key applications in areas such as lectin binding, antibody binding, cell adhesion, enzyme specificity and pathogenic identification.[2–4] A more detailed account on the applications of glycoarrays is beyond the scope of the project. However a key feature that needs to be taken into account when designing a glycan-binding probe is the immobilization strategy to be used.

1.5.1. Glycoarray immobilization

As stated earlier a key pre-requisite to the design of glycoarrays is the selection of an appropriate functional group on the glycan probe being compatible with the solid array surface. To date, two broad strategies for the immobilization of glycan entities to a solid support surface have been developed in order to achieve this, namely non-covalent and covalent immobilization. In either strategy the actual immobilization can be achieved *via* an automated printing approach or the more common manual binding approach. It must be re-iterated, however, that the focus of this project pertains to the synthesis of the glycan probe and not its actual immobilization to the array surface. Hence, only an informative account of the covalent and non-covalent immobilization strategies used in glycoarray fabrication is given below.[2–4]

1.5.1.1. Non-covalent immobilization strategies

Over the years a large number of covalent and non-covalent methods have been developed in order to attain the desired glycoarray. Some of the first non-covalent methods reported, involved the immobilization of polysaccharides, proteoglycans and neoglycoproteins to either nitrocellulose or oxidized polystyrene surfaces *via* mixed hydrophobic and hydrophilic interactions. This method was later expanded to the use of electrostatic (**A**), fluorophilic (**B**), biotin-streptavidin and recently complementary DNA hybridization (**C**) interactions as shown below in Figure 4.[3,4] Each subsequent strategy overcame the limitations of their respective predecessor. Albeit a useful strategy, a common pitfall to the non-covalent approach is that the glycan probes are randomly orientated on the array surface, a subtle discrepancy that ultimately hampers the efficiency to which the recognition event under investigation can take place.[2–4]

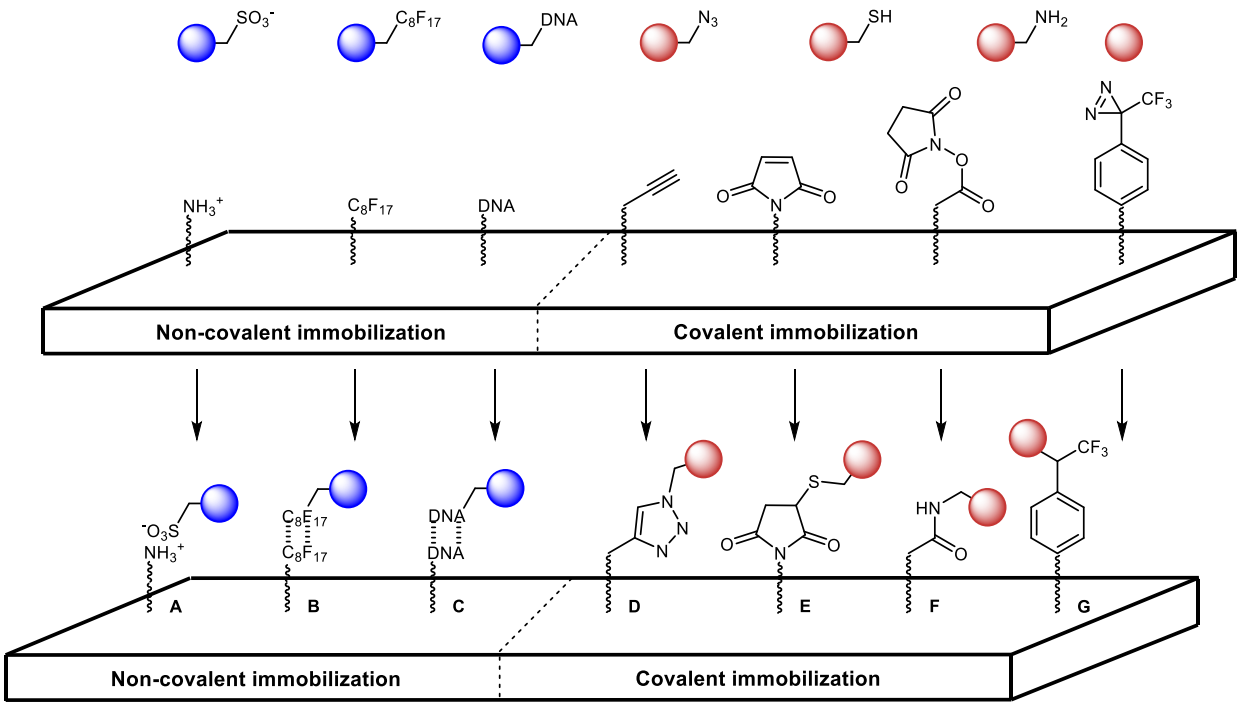


Figure 4: A comparison of non-covalent and covalent immobilization strategies used to array glycan probes (blue and red spheres) on a solid support surface.[3,4]

Strategies for the covalent immobilization of glycan probes to an array surface emerged soon after the first reports of non-covalent strategies and inadvertently provided a solution to this unfortunate limitation.

1.5.1.2. Covalent immobilization strategies

Glycoarrays which are assembled *via* covalent immobilization are more uniformly orientated due to being specifically bound to the solid support surface. This subtle difference allows for the glycan probe to better recognize the CBP under investigation.[2–4] This method requires only that compatible functional groups be present on the solid support surface and linker molecule, prior to immobilization. Currently the foremost covalent methods exploit the use of; Huisgen-type 1,3-dipolar chemistry (**D**), thiol-malemide ligation chemistry (**E**) and peptide ligation chemistry (**F**) as shown in Figure 4.[3,4]

Some of the early work conducted in covalent glycoarray fabrication, made use of thiol-malemide ligation chemistry (**E**) in order to help elucidate the binding epitope of cyanovirin-N.[3] This virucidal protein is produced by the cyanobacterium *Nostoc ellipsosporum* and is a powerful inhibitor of numerous viruses including HIV.[11] Although the use of thiol chemistry remains a popular choice in glycoarray fabrication, the robust nature of amine chemistry has resulted in it being the more widely adopted method due to its compatibility with a plethora of array surfaces which can be extended to affinity chromatography. However, a major limitation to covalent immobilization is the actual synthesis of the glycan probe since this will often require the introduction of a suitably functionalized tether molecule to the glycan of interest, often requiring several steps together with difficulties encountered during purification.[4] In addition to this, the selected functional group needs to be robust enough to undergo further derivatization in the event that access to an appropriate array surface is limited. An alternative approach to the above approaches, which seeks to circumvent this obstacle, is photoimmobilization chemistry. This technique makes use of a photoreactive functionalized surface, to which untethered glycan probes are then covalently attached to by means of a photochemical reaction. A representative example of this is shown above (Figure 4) which makes use of aryl-3-(trifluoromethyl)diazirine photochemistry (**G**) in order to covalently attach the glycan probe to the array surface.[3]

Once the glycan probe has been selectively attached to the array surface, the glycoarray can then be used to probe for CBP and CPE phenomena using a variety of spectroscopic and spectrometric techniques. Over the past decade, there has been a considerable advancement in the field of analytical glycoscience.[12] This advancement, aided in part by the development of glycoarrays, has provided a number of high-throughput analytical screening processes.

1.5.2. Analytical techniques

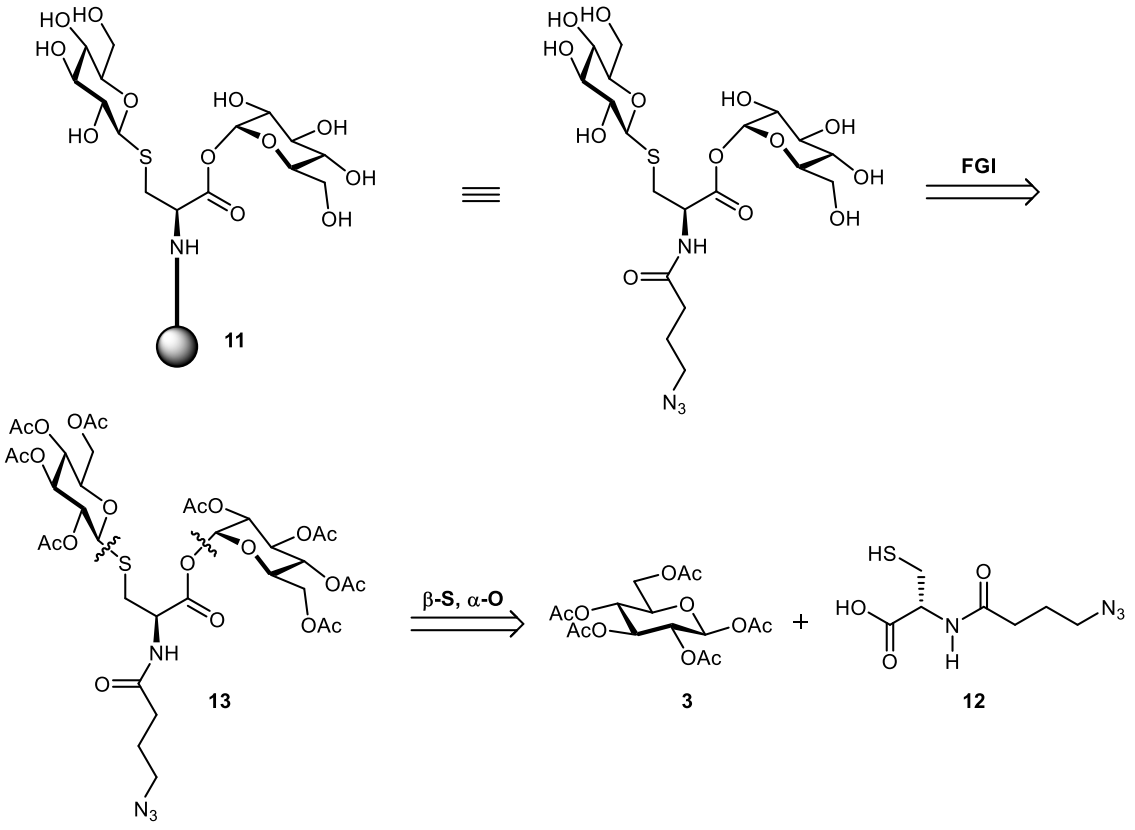
Fluorescence spectroscopy was the initial analytical method of choice used to study CBP interactions. This technique involved screening a panel of CBPs such as lectins (coupled to a fluorescent probe such as rhodamine) with known binding specificities, against the glycan probes.[4] However; due to poor data quantification and interpretation, the need of a label-free technique and the advent of new complex array platforms an alternative method was required.[5]

Mass spectrometry (MS) and surface plasmon resonance (SPR) spectroscopy have since emerged as two prominent methods to tackle the above issues.[4,5,12] More specifically, time-of-flight secondary ion mass spectrometry (ToF-SIMS) and surface plasmon resonance imaging (SPRi) are the two complementary techniques that are used to study the CBP events. ToF-SIMS is a mass spectrometry technique in which the surface interface (≈ 1.5-2.0 nm) is showered with a primary ion beam in order to generate detectable secondary ions. Quantification of these detected ions reveals the complex chemical composition of the glycoarrays, giving information pertaining to the chemical species surface distribution.[5] In contrast, SPRi is a technique based on refractive index, in which the surface is irradiated with ρ-polarized light and the reflected light is detected. This allows for the observation of complex biomolecular interactions that occur at the surface interface, giving information related to the bioactivity of the glycoarray.[5] In addition to this, both techniques have allowed for the surface chemistry of the arrayed glycans to be studied, revealing information pertaining to their neighbouring glycan probes and their molecular conformations. This is of the utmost importance as the surface chemistry plays a crucial role in both the bioavailability and function of the arrayed glycans; meaning that the quality of the printed array can now be quantified. In turn this allows for the fabrication and function of various glycoarrays to be optimized, greatly aiding in the actual application of the glycoarray.[5]

With all of the above in mind a retrosynthetic analysis of the glycan-binding probe **11** from available starting materials was conducted, and a synthetic route towards **11** proposed.

1.6. Retrosynthetic analysis

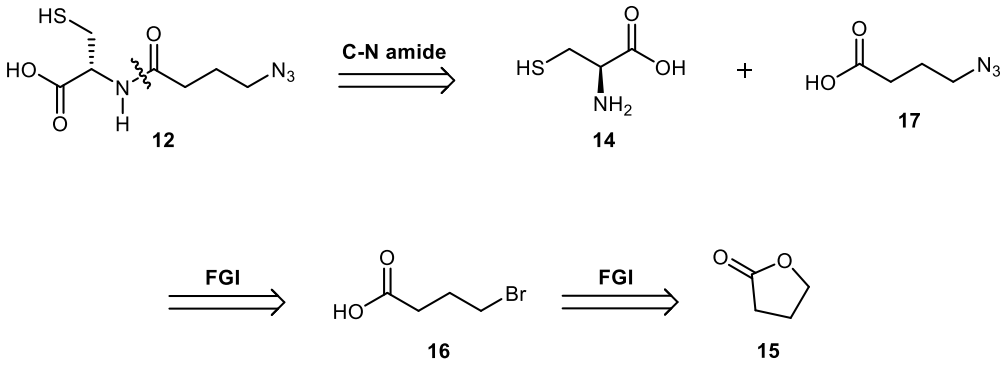
As stated earlier, target bis-glucoside **11** was identified as the initial focus of this project. A retrosynthetic analysis of **11** is presented below (Schemes 3 and 4), incorporating the one-pot bis-glucosylation reaction. Initially it reveals both penta-acetyl glucopyranose **3** and *N*-acyl cysteine **12**, to be the primary starting materials of **11**, with the obvious challenge being whether **3** and **12** will undergo the serendipitously discovered bis-glucosylation reaction to afford the tethered BGC derivative **13**. Deprotection of **13** would then afford target molecule **11**, although it is recognized that selective de-*O*-acetylation without disrupting the α-glucosyl ester linkage would be a challenge.



Scheme 3: Retrosynthesis of bis-glucoside **11** towards primary starting reagents' **3** and **12**.

*IUPAC name: *N*-(4-azidobutanoyl)-L-cysteine (**12**)

While penta-acetyl glucopyranose **3** was commercially available, *N*-acyl cysteine **12** was not and would need to be synthesized. A retrosynthetic analysis of **12** (Scheme 4) suggests L-cysteine **14** and γ-butyrolactone **15** as precursors, both of which were available in house. The formation of **12** would involve ring opening of the lactone **15** together with formation of a terminal bromide to afford 4-bromobutanoic acid **16**. Bimolecular nucleophilic substitution (S_N2) of the bromo acid **16** would then afford 4-azidobutanoic acid **17**, which in turn would undergo a peptide bond coupling reaction with **14** to afford **12**. Compounds **16** and **17** were both commercially available although **17** is considerably more expensive than **16** and the chemistry of the starting cyclic lactone **15** had other advantages.



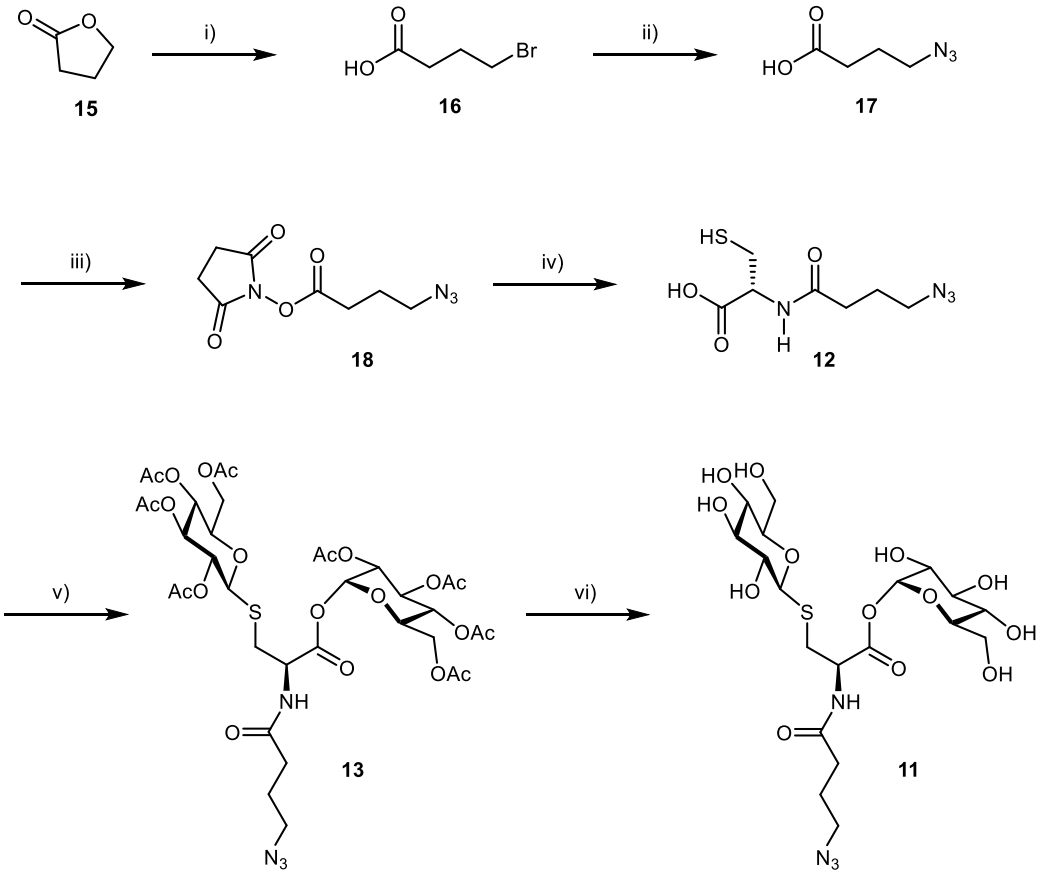
Scheme 4: Retrosynthesis of *N*-acyl cysteine **12** towards secondary starting reagents' **14** and **15**.

Having settled on the use of penta-acetyl glucopyranose **3**, L-cysteine **14** and lactone **15** in order to achieve the formation of bis-glucoside **11**, a synthetic route was proposed together with recommendations on alternative approaches in case of potential setbacks.

*IUPAC name: dihydro-2(3*H*)-furanone (**15**)

1.7. Proposed synthetic route

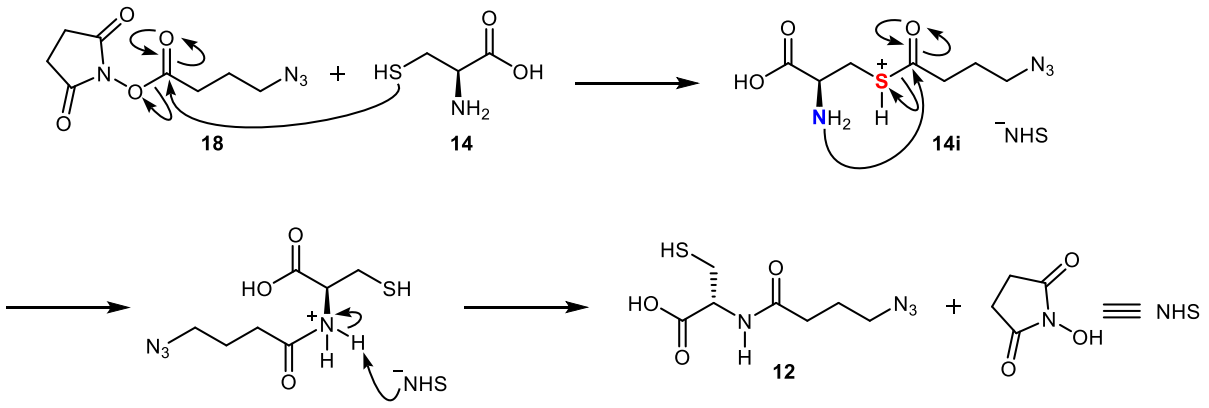
The proposed synthetic route (Scheme 5) sought to achieve the formation of target bis-glucoside **11** starting from the inexpensive lactone **15**, using methodology from previously published work.[7,13–17]



Scheme 5: Reagents and conditions; i) 48 % HBr in H_2O, conc. H_2SO_4, reflux; ii) NaN_3, MeCN, reflux; iii) NHS, EDC.HCl, DCM, r.t.; iv) **14**, NEt_3, DMF, r.t.; v) **3**, $SnCl_4$, DCM, r.t. ; vi) NaOMe (0.1 M), MeOH, r.t.

It is important to highlight six key features present in the proposed synthetic sequence. Firstly, the bromo acid **16** presented an affordable alternative to the starting material **15**, in the event that access to **15** was limited. This is due to **15** being a border-controlled substance in certain countries as it is a prodrug of γ-hydroxybutanoate, a sedative used recreationally with alcohol.

Secondly, the presence of an azido functionality in **17** will ultimately, allow for at least two possible strategies for attaching the glycan-binding probe **11** to a surface or other macromolecule, such as a Huisgen-type 1,3-dipolar cycloaddition reaction with an alkyne or reduction of the azide to an amine followed by peptide coupling as shown earlier in Figure 4, Page 10. Thirdly, formation of the succinimidyl ester **18** will make peptide coupling with L-cysteine **14**, in the subsequent step more likely due to increased electrophilicity at the acyl carbon site favouring the formation of *N*-acyl cysteine **12**. Fourthly, the peptide coupling will be conducted in the absence of thiol protection as it is widely accepted that peptide bond formation, in which **14** is involved, occurs *via* a native chemical ligation (NCL) mechanism involving an *S*- to -*N* acyl transfer as shown in Scheme 6.[18,19] Fifthly, there is a potential risk that the final deprotection step may result in cleavage of the α-glucosyl ester linkage, and this will have to be carefully investigated, including evaluating alternative protecting group strategies if necessary. Lastly, in the event that the bis-glucosylation reaction provides undesired results; the BGC scaffold or an alternative analogue bearing just one sugar unit will need to be prepared *via* alternative means. This alternative analogue would make use of the methyl ester derivative of *S*-glucosylcysteine **2**, depicted earlier (Scheme 1, Page 4), as the formation of such molecules have been well documented.



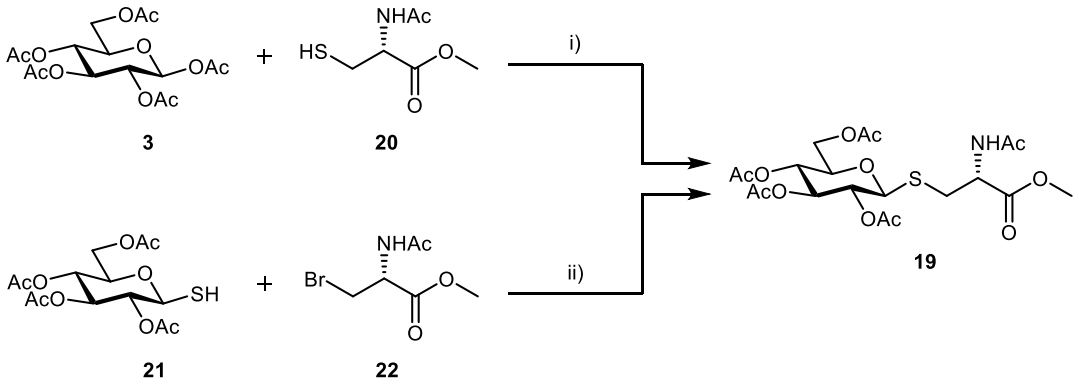
Scheme 6: The proposed NCL mechanism for the reaction between **14** and **18** leading to the formation of **12** *via* **14i**. The key step involves an *S*- to -*N* acyl transfer from the sulfur atom (highlighted in red) to the nitrogen atom (highlighted in blue) of intermediate **14i**.[18,19]

*IUPAC name: 2,5-dioxopyrrolidin-1-yl 4-azidobutanoate

A brief account of the key steps articulated above in the proposed synthesis is discussed below, together with recommendations for alternative approaches.

1.7.1. The formation of *S*-glycosylcysteine derivatives from glycosyl acetates

To date a handful of methods have been developed towards obtaining *S*-glycosylcysteine derivatives such as **19** depicted in Scheme 7.[20–24] Arguably, the most used method to achieve this makes use of a suitable glycosyl donor and Lewis acid promotion. A typical reaction using the glycosyl acetate is promoted with an excess of either $BF_3.OEt_2$ or $SnCl_4$.[20,25] The excess promoter is required as the Lewis acid remains strongly associated to the eliminated anomeric acetate leading to Brønsted acid formation, in addition to being quenched with residual traces of water if present. Despite these set-backs this method allows for the direct conversion of commercially available, inexpensive starting material such as penta-acetyl glucopyranose **3** and methyl *N*-acetyl-L-cysteinate **20** to *S*-glucosylcysteine derivative **19** in contrast to the more indirect route *via* the 1-thioglucopyranose **21** and 3-bromopropanoate **22** (Scheme 7), both of which are typically prepared *via* several steps.[26,27] However, a major advantage of the latter approach is that it is non-moisture sensitive and produces higher reported yields.[24] Nonetheless a Lewis acid approach involving glycosyl acetates would be pursued initially on the basis of the precedent for formation of BGC **1** under these conditions.



Scheme 7: Reagents and conditions: i) $SnCl_4$, DCM, r.t., (lit. 65 %); ii) TBAHS, aq. $NaHCO_3$ (0.6 M), EtOAc, 40 °C. Note this reaction (ii) was performed with a differently protected analogue of **22** with a reported yield of 86 % being obtained for their target *S*-glucosylcysteine derivative).[20,24]

*IUPAC names: methyl *N*-acetyl-*S*-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-L-cysteinate (**19**)
2,3,4,6-tetra-*O*-acetyl-1-thio-β-D-glucopyranose (**21**)
methyl (*R*)-2-acetamido-3-bromopropanoate (**22**)

1.7.2 Peptide bond formation

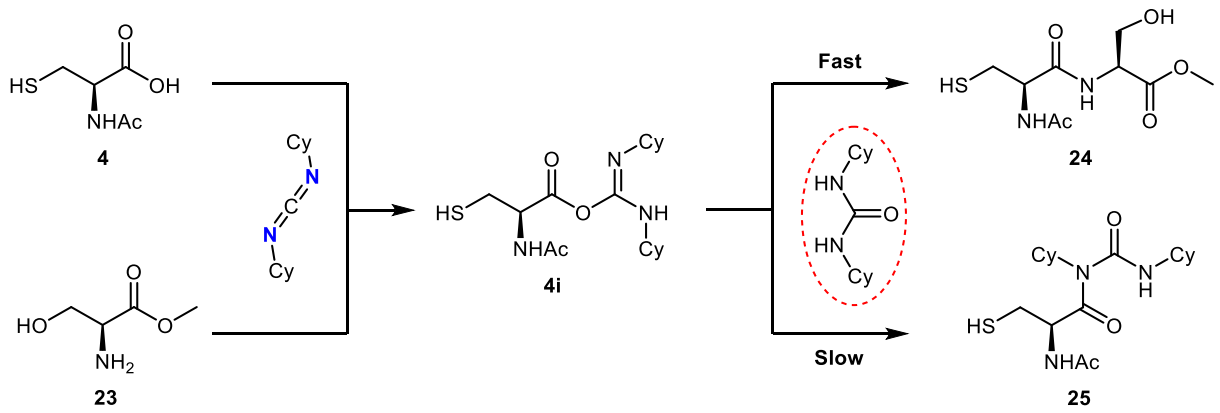
At the turn of the 19th century, Emil Fischer, proposed the use of acid chlorides in the synthesis of peptides. The success of this method in addition to the limited availability of complex peptides, spurred on the development of a several other peptide coupling strategies in the past century.[28] Thus, in addition to acid chlorides; the two other prevalent strategies employed make use of carbodiimides and active ester chemistry. The choice of strategy employed is often dependent on the quantity of coupling reagent to be used, neighbouring groups present, chemistry in either solution or solid phase, and whether automated or manual synthesis is used.[28,29] The latter two variables were not relevant in this study, since the emphasis would be on step-wise, solution-based chemistry.

1.7.2.1. Acid chlorides

The inception of acid chlorides, by Fischer, in peptide chemistry was initially met with great interest. However, this slowly dwindled over time primarily due to the high reactivity of acid chlorides, making them susceptible to a range of possible side reactions including hydrolysis. In addition to this their preparation often involves the use of hazardous chlorinating reagents such as thionyl chloride, phthaloyl dichloride, and oxalyl chloride in excess.[28,29] The use of excess reagent complicates matters even further, as its removal is usually achieved *via* evaporation under vacuum, meaning that frequent preparation of acid chlorides can lead to deterioration of lab apparatus over time.

1.7.2.2. Carbodiimides

The carbodiimides are the reagents of choice in peptide coupling reactions. These carbodiimides contain two weakly basic nitrogen atoms that can initiate a reaction between the carbodiimide and carboxylic acid, such as *N*-acetyl-L-cysteine **4**, leading to the formation of a highly reactive *O*-acylurea species (Scheme 8) such as **4i**. This species can then undergo rapid aminolysis with an amine such as methyl L-serinate **23** to afford the desired peptide product **24**. However, a possible side reaction is the irreversible rearrangement of the *O*-acylurea species **4i** to the more stable *N*-acylurea species **25**. This undesirable side reaction proceeds more rapidly in dimethylformamide than in dichloromethane. In order to circumvent this undesired result, additives bearing a triazole scaffold such as 1-hydroxybenzotriazole (HOBt) are often used in conjunction with the carbodiimide reagent. Firstly, these additives serve to protonate the *O*-acyl urea species, preventing an intramolecular attack from occurring and secondly, they drive the desired reaction forward *via* the formation of an *in situ* active ester. Two commonly employed carbodiimide reagents, are 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC.HCl) and N,N'-dicyclohexylcarbodiimide (DCC) the former being the reagent of choice as it overcomes the shortcomings of its predecessor DCC. This is as a result of DCC forming a urea salt by-product which is not amenable to simple removal *via* aqueous work-up.[28,29]



Scheme 8: A hypothetical carbodiimide reaction between **4** and **23**, promoted by DCC (containing weakly basic nitrogen atoms highlighted in blue) to afford the peptide product **24** and urea salt by-product (encircled in red) *via* the *O*-acylurea species **4i**. The *O*-acylurea species **4i** can also undergo an irreversible rearrangement to afford the undesired *N*-acylurea species **25**.[28,29]

1.7.2.3. Active esters

Thus it is no surprise that the arrival of active esters to the field of peptide chemistry was met with great enthusiasm, as they addressed the drawbacks associated with acid chlorides and carbodiimides. Active esters are in fact mixed anhydrides, being stable enough to be isolated, yet just reactive enough to undergo only the desired nucleophilic acyl substitution reaction.[28,29] They are usually prepared, under standard carbodiimide-ester conditions, from a carboxylic acid such as **4** and either a phenyl or hydroxamic alcohol. In the former approach the introduction of a *p*-nitrophenyl (**pNO$_2$**), 2,4,5-trichlorophenyl (**pCl**) or pentafluorophenyl (**pF**) group (Figure 5) primes the acyl carbon towards nucleophilic attack, due to the excessive electron-withdrawing nature of the substituted aromatic rings. However, the major limitation of phenyl active ester derivatives is the formation of a water insoluble by-product complicating the work-up. With regard to hydroxamic active esters (Figure 5) *o*-phthalimdo active esters were initially used, but they too formed a water insoluble by-product.[28,29] The solution to this came in the form of succinimidyl active esters, as the *N*-hydroxysuccinimide (NHS) released during aminolysis is water soluble. Despite this advantage, NHS active esters are significantly less reactive than the latter four examples.

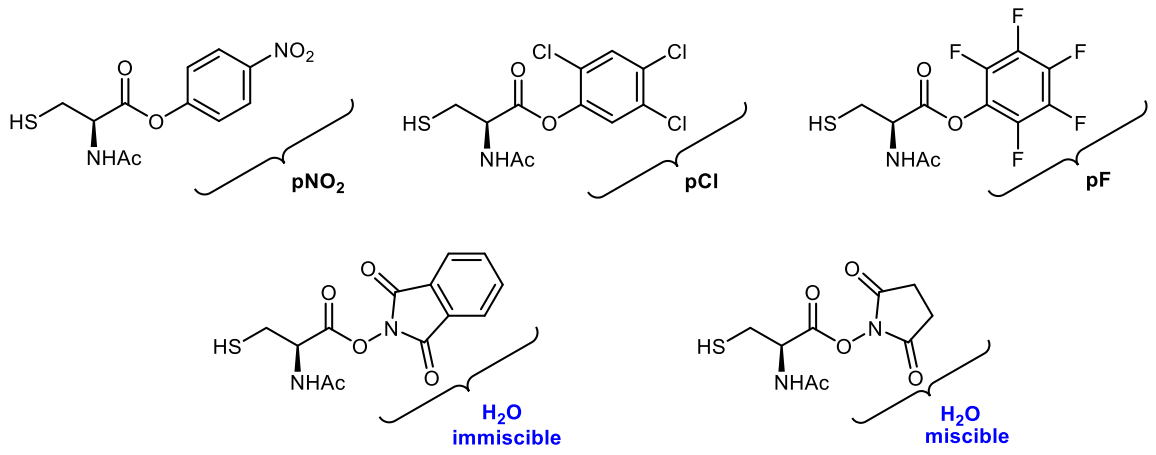


Figure 5: Hypothetical phenyl (top row) and hydroxamic (bottom row) active esters of **4**.[28,29]

It was hoped that if the preparation of glycan-binding probe **11**, as outlined earlier (Scheme 5, Page 15), was successful that the process could be repeated with biologically relevant sugar units other than glucopyranose. In addition to this, two other important variables that needed to be taken into account were the linker length and CBP-specificity to either one or both of the glucosyl units found in **11**. The former factor is of the utmost importance as it plays a significant role during the recognition event, which can be hindered by the probe being in close proximity to the surface. With regards to the latter factor, it is unknown whether the targeted CBP will recognize and bind to the individual glucosyl units of **11**, be unable to distinguish the individual glucosyl units (multivalent presentation), or recognize the entire de-*O*-acetylated BGC scaffold as a single receptor, and this would remain to be established experimentally.[30]

With the above two factors in mind, two additional glycan-binding probes designated as bis-(*N*-acetylglucosaminyl)-L-cysteine (BAGC) and *S*-(glucosyl)-L-cysteinate (SGC), are shown alongside the BGC probe in Figure 6. Both of these probes are of significant interest, together, with the BGC probe as they would arm the glycobiologist with an effective molecular toolset having subtle differences, allowing for the CBP or CPE event under investigation to be better scrutinized.

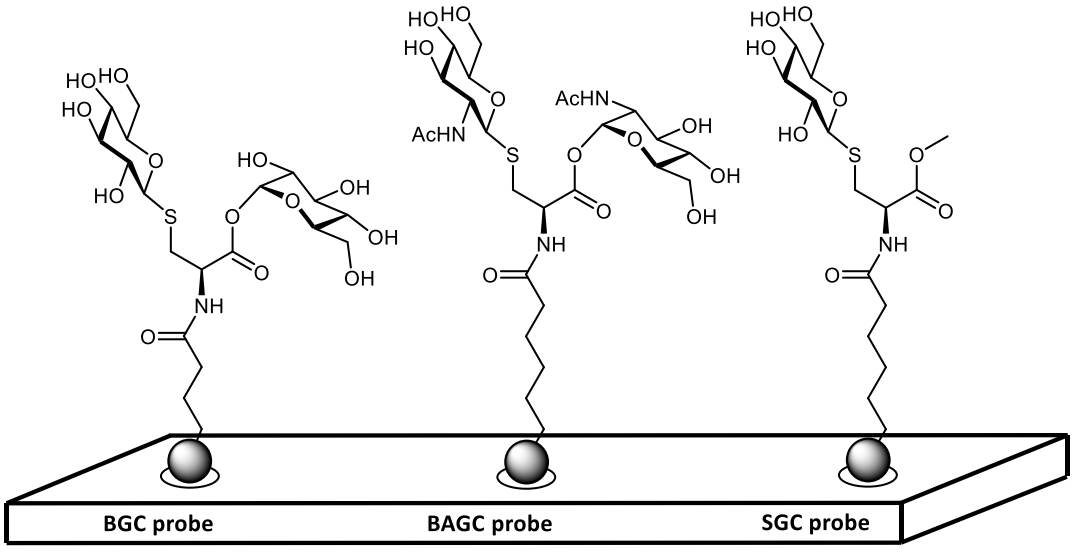


Figure 6: Three potential components of a glycoarray conceptualized from the serendipitous discovery of BGC **1**.

Chapter Two

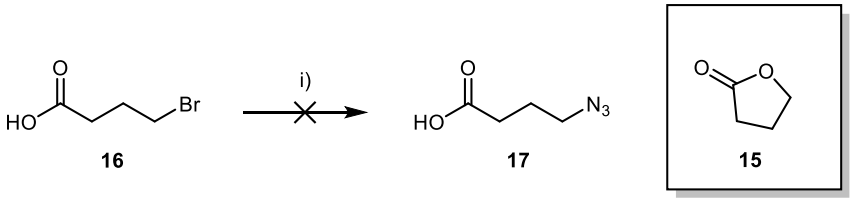
2. Results and Discussion

2.1. Preparation of *N*-(4-azidobutanoyl)-L-cysteine

As outlined earlier (Scheme 5, Page 15), the preparation of key synthetic intermediate **12** was proposed *via* a four-step synthesis starting with γ-butyrolactone **15**. The synthesis, as described below, was eventually achieved in six steps.

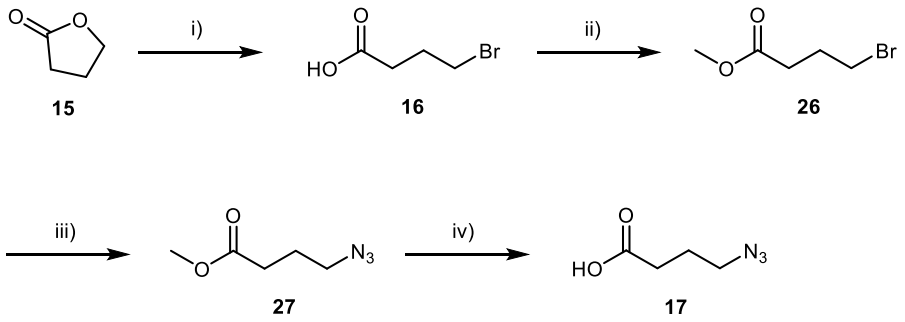
The first step involved refluxing **15** with a solution of 48 % HBr in water and concentrated H_2SO_4 for 15 h, and a single product was isolated as a yellow solid in a 72 % yield, whose m.p. of 32 ^oC agreed with the reported literature value.[13,31] The ^1H NMR spectrum of the solid included a two-proton triplet at 3.48 ppm consistent with the presence of the alkyl bromide, confirming the formation of the bromo acid **16**.

The next step required the S_N2 substitution of the bromo acid **16** with sodium azide in either acetonitrile or dimethylformamide at 85 ^oC.[14,15] Attempts in either solvent were, however, unsuccessful as in both cases the sodium azide remained visibly insoluble. This was confirmed by an IR-spectrum of the crude product isolated from the reaction conducted in dimethylformamide, which showed the absence of a diagnostic azide stretch at ≈ 2100 cm^{-1}. Furthermore the ^1H NMR spectrum of the crude product revealed that **16** had undergone lactonization to form **15** (Scheme 9), rather than the desired S_N2 substitution.



Scheme 9: Undesired lactonization of **16**, under the reaction conditions; i) NaN_3, MeCN or DMF, 85 ^oC, 3 h.

In light of this result it seemed appropriate to attempt the formation of the azido acid **17** *via* an indirect route, involving conversion of the bromo acid **16** to the corresponding methyl ester **26** prior to introduction of the azide as outlined in Scheme 10.[13,32–35]



Scheme 10: Reagents and conditions: i) 48 % HBr in H_2O, conc. H_2SO_4, 0 °C for 1 h and then reflux for 14 h, 72 %; ii) conc. H_2SO_4 (cat.), MeOH, reflux, 3 h, 84 %; iii) NaN_3, DMSO, 50 °C, 4 h, 48 %; iv) aq. NaOH (1.0 M), MeOH, 4 h, r.t., 65 %.

The bromo acid **16** was therefore refluxed in methanol in the presence of a catalytic amount of H_2SO_4 for 3 h to give the methyl ester **26** which was isolated as a colourless, viscous liquid in good yield.[32] This was confirmed, *inter alia*, by the ^{13}C NMR spectrum of the viscous liquid which showed a one-carbon singlet at 51.8 ppm consistent with the presence of the methyl ester.

The methyl ester **26** was then treated with a slight excess of sodium azide in dimethyl sulfoxide at 50 ^{o}C for 4 h to form the 4-azidobutanoate **27** which was isolated as an off-yellow, viscous liquid.[33] The ^{1}H and ^{13}C NMR data for the viscous liquid were in agreement with literature data for this compound.[35] Furthermore the IR-spectrum showed a diagnostic azide stretch at 2095 cm^{-1} consistent with the presence of the azide, confirming the formation of the alkyl azide. Although contamination with dimethyl sulfoxide was evident from the presence of a singlet resonating at δ_H 2.62 ppm, a pure product was eventually obtained in moderate yield when several cold aqueous washes were applied during the work-up.

*IUPAC names: methyl 4-bromobutanoate (**26**)
methyl 4-azidobutanoate (**27**)

The ^{1}H NMR spectra (Figure 7) of isolated compounds **16**, **26** and **27** are compared to highlight the incorporation of the methyl ester and shielding effect of the azide relative to bromine on the terminal methylene group.

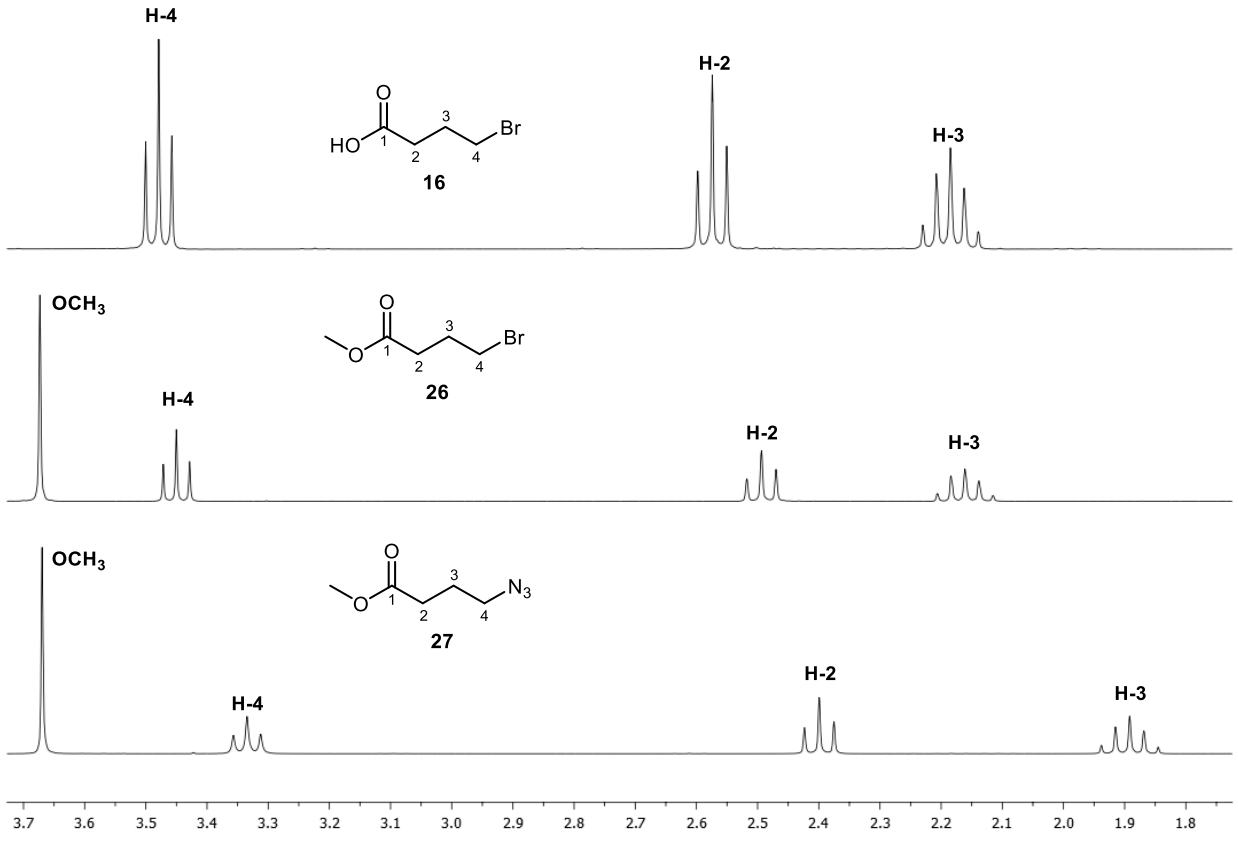


Figure 7: A comparison of the ^{1}H NMR spectra (1.75 - 3.70 ppm) for isolated compounds **16**, **26** and **27**.

Hydrolysis of the 4-azidobutanoate **27** was then accomplished by using an aqueous NaOH solution and a minimal amount of methanol, which was required together with vigorous stirring in order to obtain a homogeneous solution.[34] After 4 h the azido acid **17** was isolated in good yield as an off-yellow, viscous liquid. This was confirmed by the ^{1}H NMR spectrum of the viscous liquid which showed the disappearance of the three-proton singlet at 3.67 ppm for the methyl ester of **27**, consistent with the formation of the carboxylic acid.

With the desired azido acid **17** now in hand, the proposed synthesis of *N*-acyl cysteine **12** could be resumed, with the next step involving esterification with NHS.

The azido acid **17** was thus treated with NHS in the presence of EDC.HCl overnight, yielding a single, less polar product which was isolated as a colourless oil in an 82 % yield.[14,36] The ^{1}H NMR spectrum of the oil showed a four-proton singlet at 2.82 ppm consistent with the presence of the isochronous protons of the succinimidyl group, confirming the formation of the succinimidyl ester **18**. Interestingly a comparison of the ^{13}C NMR spectra (Figure 8) obtained for isolated compounds **17** and **18**, showed a noticeable shielding of C-1 from 179.1 ppm to 169.0 ppm in going from compound **17** to **18**.

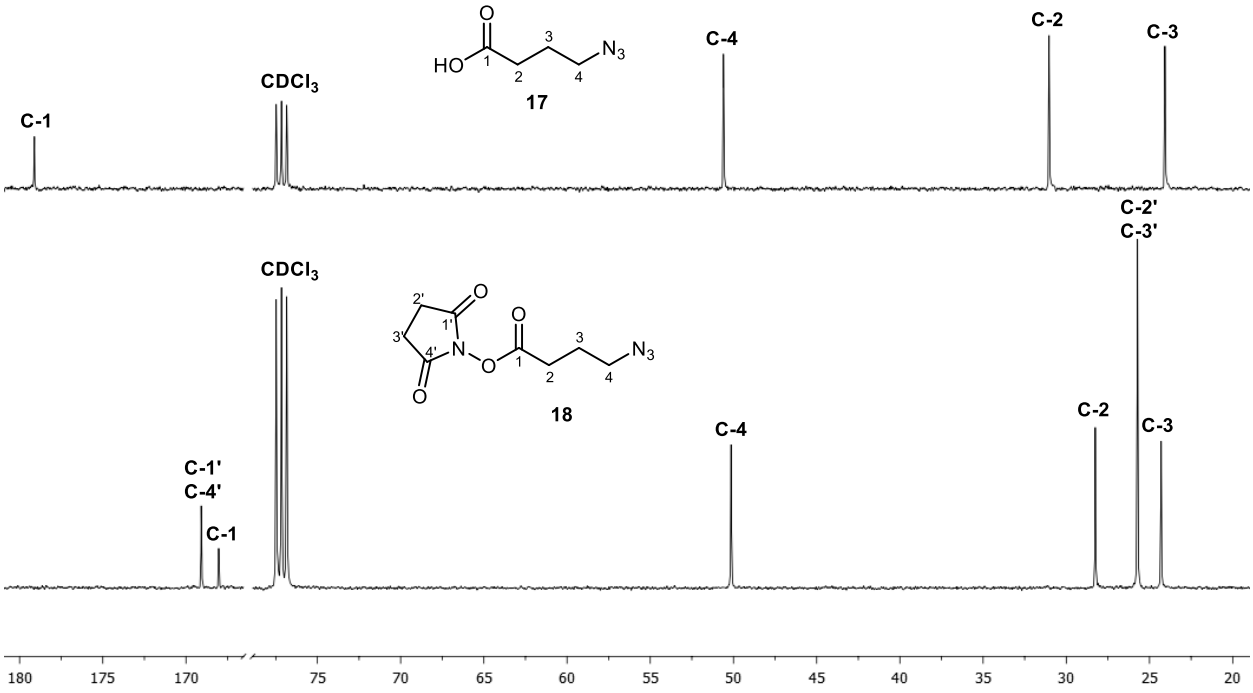


Figure 8: A comparison of the ^{13}C NMR spectra (20.0 - 180.0 ppm) for isolated compounds **17** and **18**. The region (80.0 - 165.0 ppm) has been omitted for clarity.

The succinimidyl ester **18** was then reacted with L-cysteine hydrochloride **14** under basic conditions in dimethylformamide.[15] After 14 h, TLC showed the disappearance of **18**; together with the presence of a highly polar product, having a similar streaking pattern to **14** and giving the same colour as *N*-acetyl-L-cysteine **4** upon heating with a ninhydrin solution. After an extensive aqueous work-up the highly polar product was isolated as an off-yellow oil in a 45 % crude yield

The ^{1}H NMR spectrum of the oil (Figure 9) showed the presence of a one-proton doublet at 6.67 ppm and a one-proton triplet at 1.49 ppm, corresponding to an amide proton and thiol proton respectively, confirming the formation of target *N*-acyl cysteine **12**.

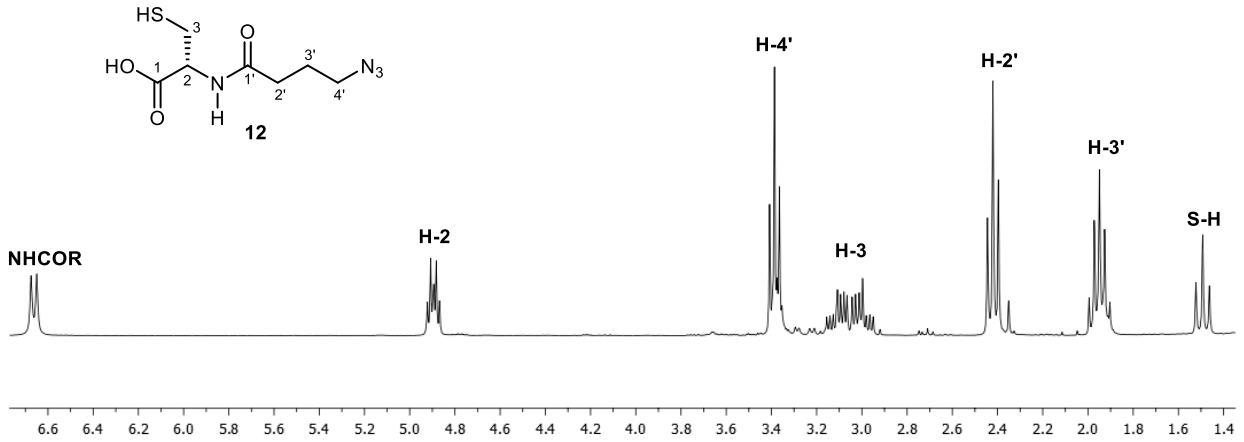
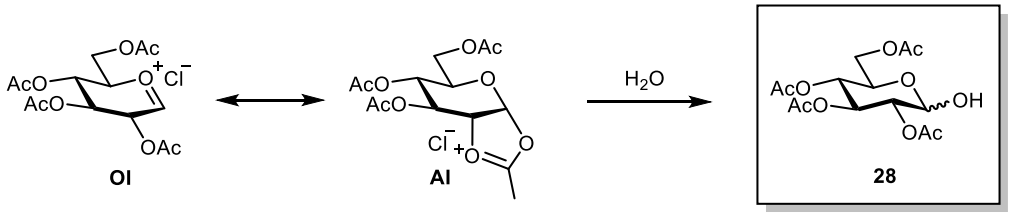


Figure 9: The ^{1}H NMR spectrum of the crude reaction product containing compound **12**. Only the region (1.40 - 6.70 ppm) has been shown due to the presence of some contamination.

At this point the possibility of bis-glucosylation of the partially purified sample of **12** was investigated by treating it with penta-acetyl glucopyranose **3** and $SnCl_4$ in dry dichloromethane.[7] The reaction did not go to completion; however, two more polar products had formed, with the least polar of the two being isolated as a colourless oil after careful column chromatography. This product was identified as the tetra-acetyl glucopyranose **28** according to the ^{1}H NMR spectrum of the oil,[37] indicating that hydrolysis of **3** had taken place, presumably *via* an oxocarbenium ion **OI** or 1,2-acetoxonium ion **AI** as shown in Scheme 11.



Scheme 11: The hydrolysis of penta-acetyl glucopyranose **3** (not shown here) in the presence of $SnCl_4$ *via* **OI** or **AI**, to afford the undesired product **28**.

*IUPAC name: 2,3,4,6-tetra-*O*-acetyl-D-glucopyranose (**28**)

This outcome prompted a detailed re-examination of the reaction conditions required for the originally reported bis-glucosylation reaction.

2.2. Reproducing the bis-glucosylation of *N*-acetyl-L-cysteine

In order to investigate in detail the optimal reaction conditions for the bis-glucosylation of *N*-acetyl-L-cysteine **4** with penta-acetyl glucopyranose **3**, these reactants were initially combined in dry dichloromethane and the promotor $SnCl_4$ added according to the reported literature procedure.[7] After 3 h as judged by TLC, two polar products had formed and a significant amount of starting material **3** still remained. In addition, a small amount of **4** was also visibly present in the reaction vessel, and the reaction also appeared not to progress any further after 3 h. Both of the products and the starting material **3** were isolated as colourless oils after extensive column chromatography. The product presumed to be the starting material **3** was shown by ^{1}H NMR (Figure 10) and comparison with literature data[38] to be a mixture of **3**, its α-anomer **29** and the glucosyl chloride **30** in a ratio of 1:6:8 as judged from integration of their respective anomeric protons. Also shown (Figure 10) is the relatively deshielded three-proton singlet at 2.17 ppm consistent with axially-orientated anomeric acetate of **29**.

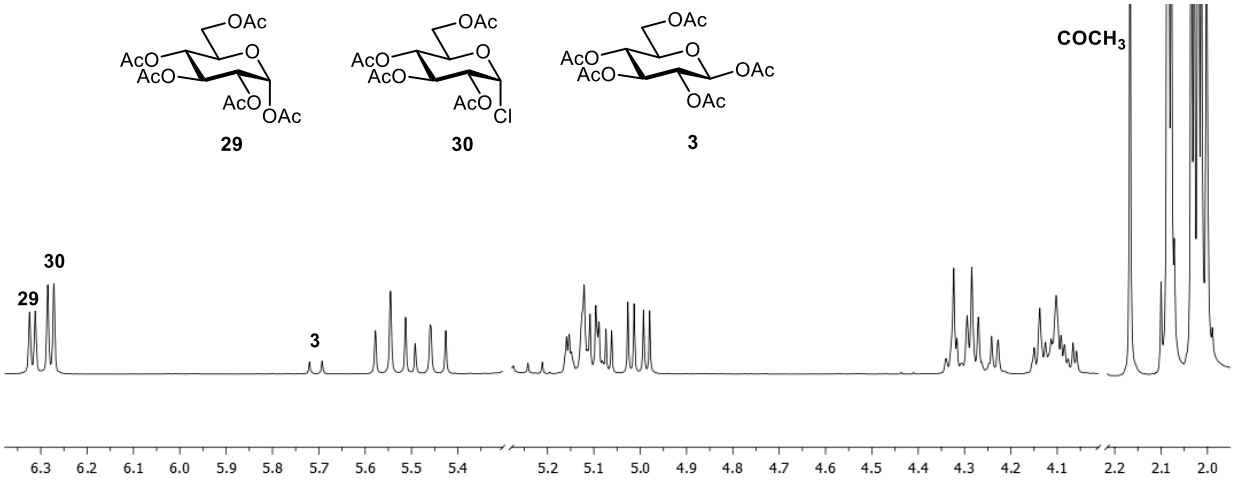


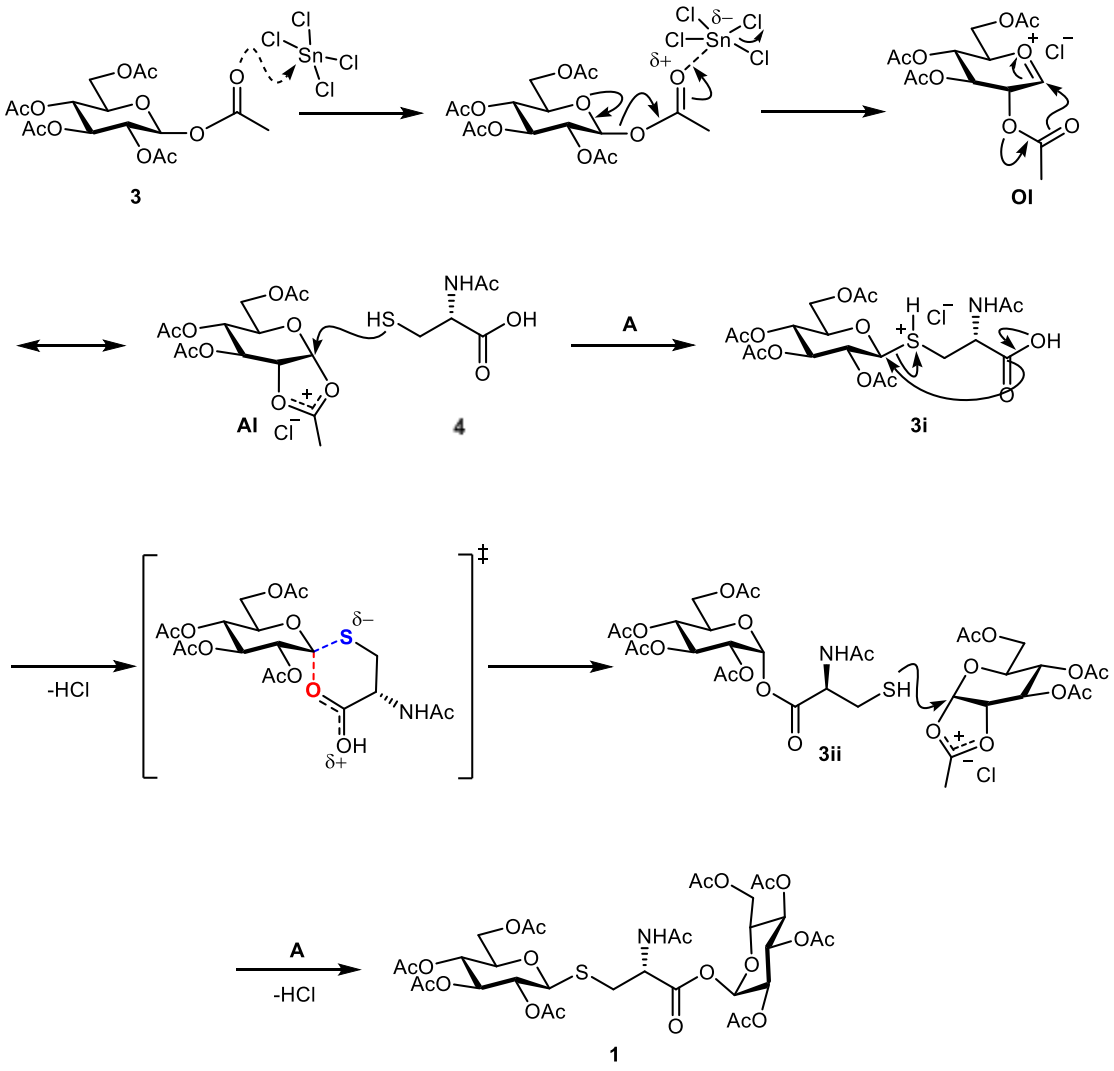
Figure 10: The ^{1}H NMR spectrum (1.95 - 6.35 ppm) of the presumed starting material **3**. Anomeric protons for compounds **3**, **29** and **30** are designated by numbered peaks in the region (5.65 - 6.35 ppm).

*IUPAC name: 1,2,3,4,6-penta-*O*-acetyl-α-D-glucopyranose (**29**)
2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranosyl chloride (**30**)

These observations are consistent with work conducted by *Konstantinović et al.* in which they observed the formation of the glucosyl chloride **30** as a by-product when treating **3** with various alkyl alcohols in the presence of $SnCl_4$ in dry dichloromethane.[39] Tetra-acetyl glucopyranose **28** was again identified as a minor product, and the most polar product identified as the *S*-glucosylcysteine derivative **2** according to their respective ^{1}H NMR spectra.[7,37]

Similar results were obtained on repeating this reaction on a small scale (300-500 mg) over prolonged reaction times or at temperatures up to 35 ^{o}C, and with rigorous exclusion of water by using dry solvents and activated molecular sieves. The use of alternative Lewis acids was also investigated, with trimethylsilyl trifluoromethanesulfonate (TMSOTf) and $BF_3.OEt_2$ providing a similar product distribution, despite no visible presence of **4** in the reaction vessel after 3 h, unlike the case when $SnCl_4$ was used. This time, however, the ^{1}H NMR spectrum of the product contained signals at 3.00-3.20 ppm consistent with the diastereotopic methylene protons in cysteine. Despite this the ^{1}H NMR spectra lacked the required diagnostic signals correlating to BGC **1**. In a final attempt at forming **1**, a $SnCl_4$ solution (1.0 M in DCM) was used but the reaction mixture gave the same TLC profile after 3 h. However, one of the isolated products contained all the diagnostic signals expected for **1**, confirming its formation in a 12 % crude yield. The ^{1}H NMR spectrum of this partially purified sample of **1** can be found in the Appendix (Page 105) of this dissertation. Further purification of **1** was not attempted in light of its potential degradation.

The exact mechanism leading to the formation of **1** is not known. However, based on the configurational outcomes of **1** and **2**, a postulated mechanism leading to the formation of **1** is put forth below in Scheme 12. It must be mentioned that the postulated mechanism bears an intriguing resemblance to some of the key mechanistic steps catalyzed by glycosyltransferases and glycosidases.[40]

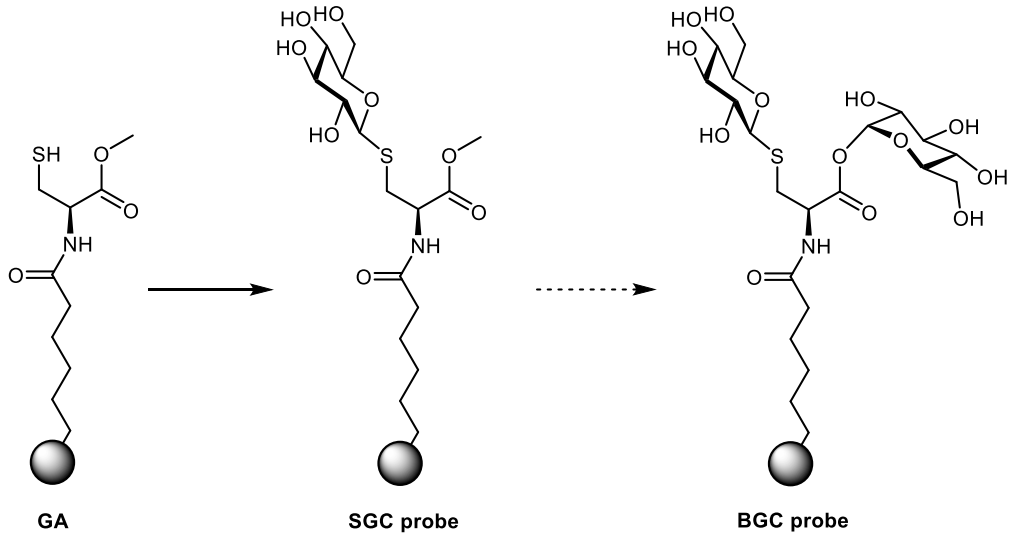


Scheme 12: A postulated mechanism for the formation of BGC **1**. During the key TS‡, the α-glucosyl ester linkage (highlighted in red) is formed and the initial β-thioglucoside linkage (highlighted in blue) is broken leading to **3ii**.

Mechanistically the formation of **1** is envisioned to proceed *via* oxocarbenium ion **OI**, prior to an intramolecular attack at C-1 by the acetyl oxygen to afford intermediate **AI**. This is then followed by attack of the nucleophilic sulfur of **4** at the anomeric position of **AI**, to afford **3i** having a β-configuration. The carboxylic acid of **3i** is then positioned to displace the primed anomeric sulfur *via* a cyclic 6-membered-ring transition state (TS), with inversion of configuration, leading to formation of the α-acylglucoside **3ii**. The α-acylglucoside **3ii** then undergoes process **A** once more to afford **1**.

From these results, it appeared that the bis-glucosylation reaction was being hindered due to the anomerization of **3**, poor solubility of **4** (when using $SnCl_4$) and possible residual traces of water. It was also possible that the poor solubility of **3** was preventing the sulfhydryl group from intercepting the **AI** intermediate, effectively halting the progress of the desired reaction. In an effort to investigate this, a series of additional experiments were set up, where it was hoped that the yield of **1** could be optimized. These experiments involved altering the order in which the reactants were combined, decreasing the rate at which the promotor $SnCl_4$ was added, increasing the equivalents of **3** used, and using dichloromethane that had been distilled and stored over 3 Å molecular sieves for 2 days.[41] Unfortunately they were all met with limited success, as according to TLC the dominating products formed were the glucosyl derivatives **2**, **28** and **30** once again.

With failure to improve on the efficiency of the one-pot formation of **1**, the attention now turned to the selective synthesis of a SGC probe to study carbohydrate-binding phenomena, and as a potential precursor to the step-wise preparation of the bis-glucosylated analogue either chemically or enzymatically, *via* a suitable glycosyltransferase, as depicted in Scheme 13. This would firstly require the preparation of an alternative glycosyl acceptor (GA).

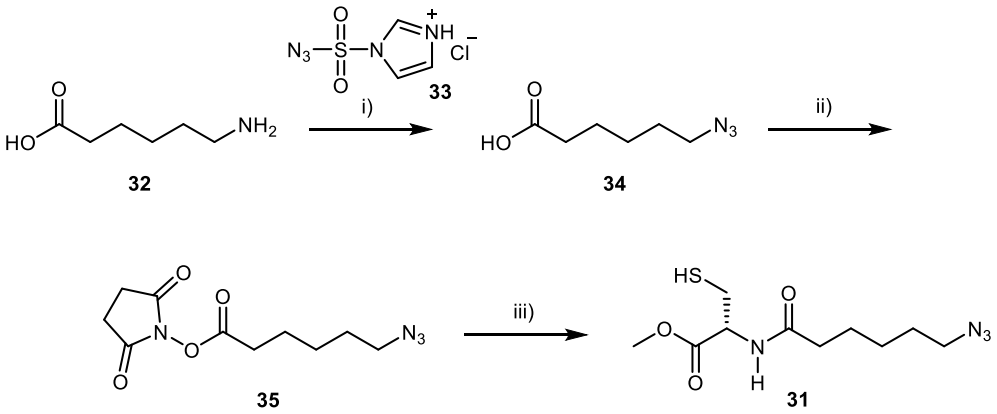


Scheme 13: A revised concept towards the synthesis of a SGC probe *via* an alternative GA and its potential extension towards the BGC analogue.

2.3. Preparation of an alternative glycosyl acceptor

2.3.1. Route *via* an azido acid and methyl L-cysteinate

The first objective in this phase of the work was to prepare adequate quantities of *N*-acyl cysteine **12** or its analogues. The first synthesis of **12**, reported earlier, had led to an impure product in a rather low yield. An alternative was therefore considered in preparing the azido-GA **31** starting with readily available 6-aminohexanoic acid **32** as outlined in Scheme 14.[36,42,43]



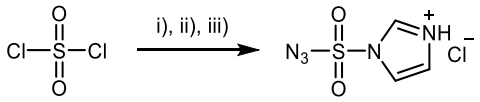
Scheme 14: Reagents and conditions: i) K_2CO_3, $CuSO_4.5H_2O$, **33**, MeOH, 16 h, r.t., 35 %; ii) NHS, EDC.HCl, DCM, 0 ^oC - r.t., overnight, 75 %; iii) methyl L-cysteinate**.**HCl **36**, NEt_3, DCM, 0 ^oC - r.t. over 1 h then r.t. for 13 h, (yield not determined due to contamination).

It was envisaged that the use of the diazotransfer reagent, 1*H*-imidazole-1-sulfonyl azide hydrochloride **33** would allow for the direct conversion of the amino acid **32** to 6-azidohexanoic acid **34**, in a more efficient route than the four-step route to the azido acid **17** as described earlier. Often such conversions involve the use of unstable and explosive trifluoromethanesulfonyl azide (TfN_3) as the diazotransfer reagent whereas the sulfonyl azide **33** is a safer alternative, prepared from inexpensive and readily available materials.[44–46]

*IUPAC name: methyl *N*-(6-azidohexanoyl)-L-cysteinate (**31**)

Secondly, methyl L-cysteinate hydrochloride **36** was chosen instead of L-cysteine **14**. It was hoped that the presence of the methyl ester would make isolation of the azido-GA **31** less troublesome and circumvent the need to use dimethylformamide due to improved solubility.

In order to form the sulfonyl azide **33**, sulfuryl chloride **37** was initially allowed to react with sodium azide in acetonitrile overnight, before the addition of imidazole.[42] After 3 h the product was isolated as a white crystalline solid in a 72 % yield, whose m.p. of 100-101 ^{o}C agreed with the reported literature value.[42] Furthermore the IR-spectrum showed a diagnostic azide stretch at 2169 cm^{-1} consistent with the presence of the azide, confirming the formation of the sulfonyl azide **33**, depicted in Scheme 15. With the bench-stable diazotransfer reagent at hand, the desired diazotransfer reaction could be attempted.



37 **33**

Scheme 15: Reagents and conditions; i) NaN_3, MeCN, 0 ^{o}C - r.t., overnight; ii) imidazole, 0 ^{o}C - r.t. over 30 min then r.t. for 3 h; iii) AcCl, EtOH, 0 ^{o}C, 72 %.

The amino acid **32** was then treated with **33** under basic conditions in the presence of a catalytic amount of $CuSO_4.5H_2O$ in methanol for 16 h, to form the desired azido acid **34** which was isolated as an off-yellow oil in low yield.[42,43] This was confirmed, *inter alia*, by the IR-spectrum of the oil which showed a diagnostic azide stretch at 2090 cm^{-1} consistent with the presence of the alkyl azide.

The azido acid **34** was then subjected to the same reaction conditions that led to the formation of the four-carbon succinimidyl ester **18**, with the desired six-carbon succinimidyl ester **35** being isolated as an off-yellow oil in good yield.[14,36] This was confirmed by the ^{1}H NMR spectrum of the oil which showed a four-proton singlet at 2.82 ppm consistent with the presence of the isochronous protons of the succinimidyl group.

*IUPAC name: 2,5-dioxopyrrolidin-1-yl 6-azidohexanoate (**35**)

Unfortunately attempts at coupling the succinimidyl ester **35** to the amine **36** under basic conditions in dry dichloromethane met with limited success, with TLC revealing the formation of several products in the reaction.[47] This was despite **36** having improved solubility in dichloromethane as was desired. Nonetheless the presumed product was isolated as a crude oil after extensive column chromatography. The ^{1}H NMR spectrum of the oil showed the appearance of signals analogous to those seen for *N*-acyl cysteine **12**, yet consistent with the azido-GA **31** having two additional saturated carbons. However, the ^{1}H NMR spectrum of the oil also revealed the presence of several impurities and attempts to purify this mixture further only led to a substantial loss of product, concomitant with impurities still present.

The lack of selectivity in this reaction could be due to a combination of the thiol group of **31** undergoing oxidation to the disulfide and or that *S*- to *N*-acyl transfer was not going to completion, leading to the potential formation of undesired by-products **38** and **39**, depicted in Figure 11. The possibility of oxidation is supported by work conducted by *Dumitriu et al.* involving treatment of **36** with a phenothiazine succinimidyl ester derivative under the exact same reaction conditions.[48] However, in their case they were able to isolate both the monomeric and dimerized product in yields of 43 % and 53 % respectively.[48]

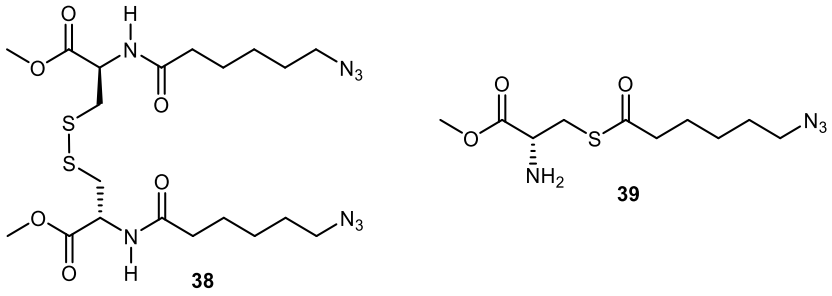
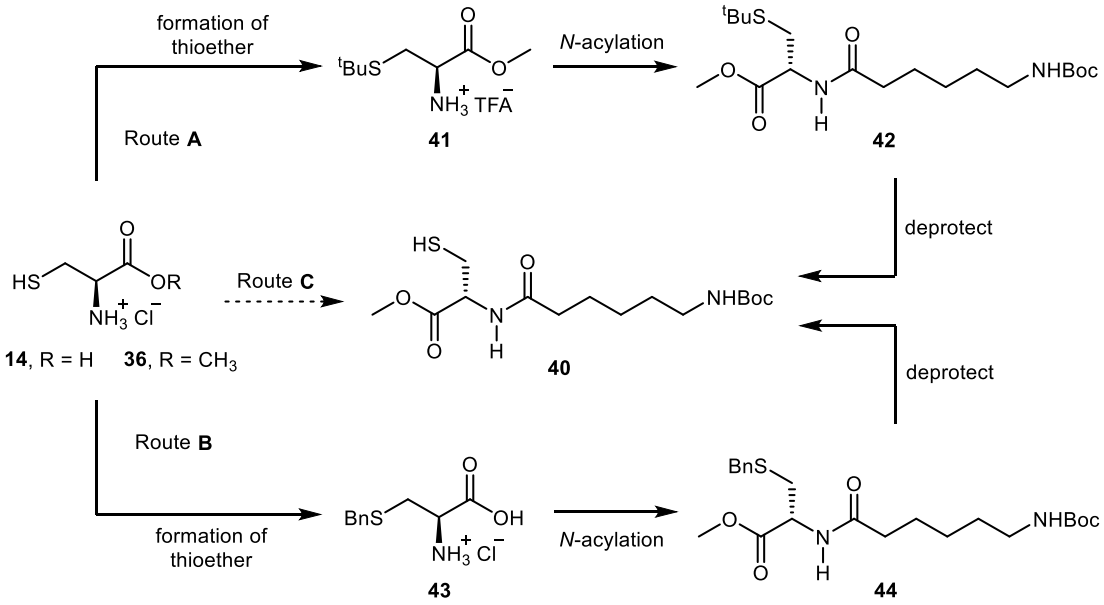


Figure 11: Two potential by-products **38** and **39**, formed in the amide coupling reaction between **35** and **36**.

These observations suggested the need for an alternative route towards the envisioned GA using an *S*-protected derivative of cysteine and a succinimidyl ester with a terminal group other than the azide.

2.3.2. Alternative synthetic route involving *S*-protected cysteine derivatives

After consulting the literature two alternative routes (Route **A** and Route **B**) towards obtaining the *N*-Boc-GA **40** (Scheme 16) were identified, involving protection of the thiol as the *tert*-butyl and benzyl thioethers respectively.[47,49–54]

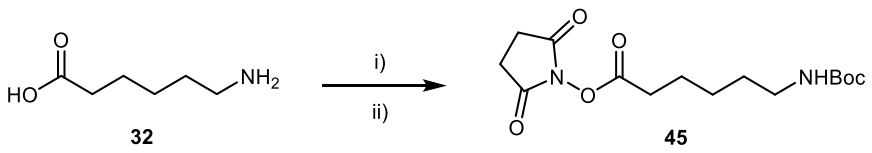


Scheme 16: Illustrates two indirect routes (Route **A** and Route **B**) and one direct route (Route **C**) towards obtaining the target *N*-Boc-GA **40**.

In addition, the low yields encountered in the key azidation step in the earlier sequences led to the incorporation of an *N*-Boc protecting group in the proposed synthetic route. Although this would preclude the use of azide-alkyne "click" chemistry in the subsequent attachment to solid supports, removal of the *N*-Boc group would allow for peptide-coupling, as illustrated earlier (Figure 4, Page 10) in the formation of glycoarrays or similar applications

The first objective was therefore to prepare the succinimidyl *N*-Boc analogue **45**, shown below (Scheme 17), of the earlier succinimidyl azide target **35** which proved successful.[55,56]

*IUPAC names: methyl *N*-(6-((*tert*-butoxycarbonyl)amino)hexanoyl)-L-cysteinate (**40**)
2,5-dioxopyrrolidin-1-yl 6-((*tert*-butoxycarbonyl)amino)hexanoate (**45**)



Scheme 17: Reagents and conditions; i) Boc_2O, NEt_3, MeOH, reflux, 24 h, 88 %; ii) NHS, EDC.HCl, DCM, 0 ^{o}C - r.t., overnight, 86 %.

To this end, the amino acid **32** was treated with Boc-anhydride and refluxed in methanol under basic conditions for 24 h, to give the *N*-Boc acid **46** as a single less polar product which was isolated as a white solid in high yield.[55] This was confirmed by the ^{1}H NMR spectrum of the solid which showed the appearance of a nine-proton singlet at 1.43 ppm consistent with the formation of the carbamate.

Conversion of the *N*-Boc acid **46** to the succinimidyl ester **45** proceeded smoothly in the same manner as described earlier for its azido acid counterparts, with the desired product being isolated as a white solid in high yield.[14,36] This was confirmed by the ^{1}H NMR (Figure 12) and ^{13}C spectra of the solid being consistent with previously reported literature data.[56]

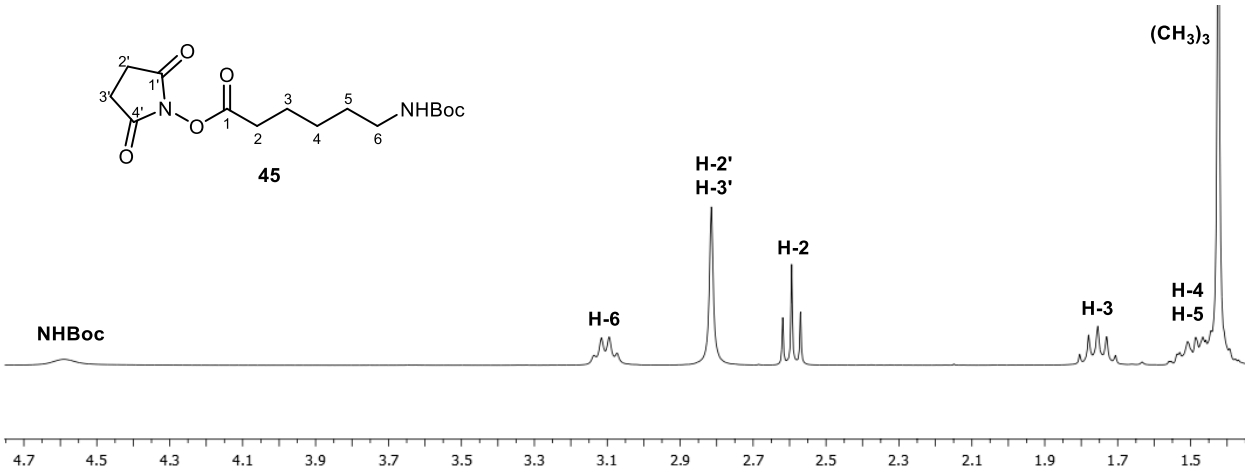


Figure 12: The ^{1}H NMR spectrum (1.35 - 4.75 ppm) of the succinimidyl ester **45**.

With the succinimidyl ester **45** now in hand the preparation of the *N*-Boc-GA **40** *via* Route **A** was pursued.

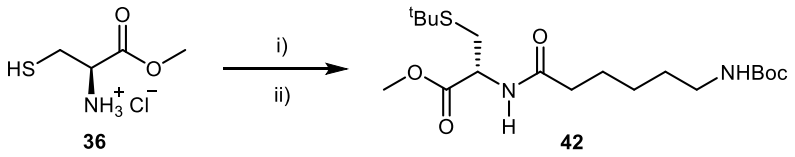
*IUPAC name: 6-((*tert*-butoxycarbonyl)amino)hexanoic acid (**46**)

2.3.2.1. Route A: *S*-*tert*-butyl protection

Route **A** was selected above that of Route **B** initially as the removal of an *S*-benzyl protecting group *via* hydrogenation is known to be problematic due to sulfur poisoning of the catalyst.

Therefore, the amine **36** was allowed to react with *tert*-butanol in trifluoroacetic acid (TFA) for 18 h.[49] After the removal of excess TFA a white wax-like solid was isolated in a 63 % yield. The ^{1}H NMR spectrum of the wax showed the presence of a nine-proton singlet at 1.33 ppm consistent with the *tert*-butyl thiol, confirming the formation of the *tert*-butyl thioether **41**. The presence of the TFA counterion was evident from the three-fluorine singlet at -75.7 ppm in the ^{19}F NMR spectrum.

The succinimidyl ester **45** was then reacted with a slight excess of the *tert*-butyl thioether **41** under basic conditions in dichloromethane.[47] After 24 h TLC showed the disappearance of **45** together with the formation of a slightly less polar product which was isolated as a colourless oil in a 64 % yield. The ^{1}H NMR spectrum of the oil showed a characteristic one-proton doublet at 6.23 ppm consistent with the presence of the amide bond, confirming the formation of the protected *N*-acyl cysteine **42**, depicted in Scheme 18. The formation of **42** was further corroborated by the ^{13}C NMR spectrum having 15 signals, which correlated to the required 19 carbon atoms taking into account signals which coincide due to symmetry, and the detection in the HRMS spectrum of a molecular ion $[M + H]^{+}$ of the required m/z 405.2419.

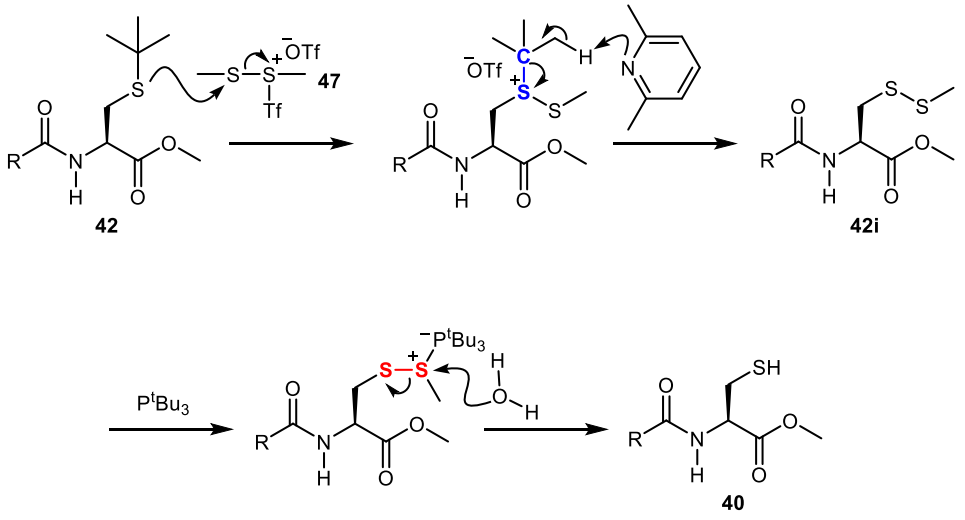


Scheme 18: Reagents and conditions; i) TFA, ^{t}BuOH, r.t., 18 h, 63 %; ii) **45**, NEt_3, DCM, 0 ^{o}C for 1 h then r.t. for 23 h, 64 %.

An attempt was then made to deprotect the thioether *via* disulfide formation, using dimethyl(methylthio)sulfonium triflate (DMTST) followed by reduction.

*IUPAC names: methyl *S*-(*tert*-butyl)-L-cysteinate trifluoroacetate (**41**)
methyl *N*-(6-((*tert*-butoxycarbonyl)amino)hexanoyl)-*S*-(*tert*-butyl)-L-cysteinate (**42**)

Unfortunately at the time the reagents required to prepare DMTST were not at hand.[50] In light of this an alternative to DMTST was prepared by treating dimethyl sulfide with triflic anhydride (Tf_2O) in dichloromethane to afford an equivalent reagent **47** (Scheme 19) as a stock solution (1.0 M in DCM).[57,58] The protected *N*-acyl cysteine **42** was then treated with the freshly prepared solution of **47** under basic conditions in tetrahydrofuran for 6 h followed by addition of a small amount of water and tributylphosphine in order to reduce the presumed disulfide intermediate **42i**, depicted in Scheme 19.[50] Unfortunately after a further 3 h TLC showed the presence of only starting material **42**, and this was confirmed by ^{1}H NMR of the isolated product. The failure of this reaction was not investigated further, although a potential reason is that the thiolating reagent **47** may have been quenched by trace amounts of water. From TLC it was evident that a significant amount of starting material **42** was still present after 6 h.



Scheme 19: A proposed mechanism for the conversion of the protected *N*-acyl cysteine **42** to the desired *N*-Boc-GA **40**, *via* the intermediate disulfide **42i**. The key steps involve S-C bond cleavage (highlighted in blue) and S-S bond cleavage (highlighted in red).[50] Note that R = $C_5H_{10}NHBoc$.

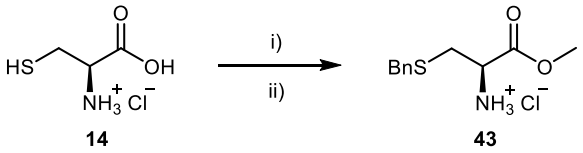
Further attempts at this reaction were not made on the basis of the postulated side reactions taking place, the appropriate reagents not being at hand at the time and the strong residual odour left on the rotary evaporator, despite the reaction being done on a small scale. In light of this Route **A** was abandoned and Route **B** was pursued instead.

2.3.2.2. Route B: *S*-benzyl protection

As an alternative, Route **B** involving preparation of a benzyl thioether, was then investigated.

Therefore, the amine **14** was allowed to react with benzyl bromide under basic conditions in ethanol for 1 h.[51] Acidification of the reaction mixture in ice afforded a white crystalline solid in a 71 % yield, whose m.p. of 201-204 ^{o}C agreed with the reported literature value for *S*-benzyl-L-cysteine hydrochloride **48**.[51] Furthermore both the ^{1}H and ^{13}C NMR data for the solid were in agreement with literature data for this compound, confirming the formation of **48**.

The benzyl thioether **48** was then treated with a methanolic hydrogen chloride solution, generated *in situ* from the addition of excess acetyl chloride to methanol at 0 ^{o}C, and left to stir for 15 h.[52] After standard work-up a white crystalline solid was isolated in a 67 % yield. The ^{1}H NMR spectrum of the solid included a three-proton singlet at 3.74 ppm consistent with the presence of the methyl ester, confirming the formation of the *S*-benzylcysteine methyl ester **43** depicted in Scheme 20.



14 **43**

Scheme 20: Reagents and conditions; i) aq. NaOH (2.0 M), BnBr, EtOH, r.t., 1 h, 71 %; ii) AcCl, MeOH, 0 ^{o}C for 1 h then r.t. for 14 h, 67 %.

The ester **43** was then combined with the succinimidyl ester **45** under basic conditions in dichloromethane and gave, after 24 h a less polar, UV-active product which was isolated as a white solid in a 72 % yield.[47] The ^{1}H NMR spectrum, depicted below (Figure 13) showed a characteristic one-proton doublet at 6.12 ppm consistent with the presence of the amide bond, confirming the formation of the protected *N*-acyl cysteine **44**. The formation of **42** was further corroborated by the ^{13}C NMR spectrum having 18 signals, which correlated to the required 22 carbon atoms taking into account signals which coincide due to symmetry, and the detection in the HRMS spectrum of a molecular ion $[M + H]^{+}$ of the required m/z 439.2278.

*IUPAC names: methyl *S*-benzyl-L-cysteinate hydrochloride (**43**)
methyl *S*-benzyl-*N*-(6-((*tert*-butoxycarbonyl)amino)hexanoyl)-L-cysteinate (**44**)

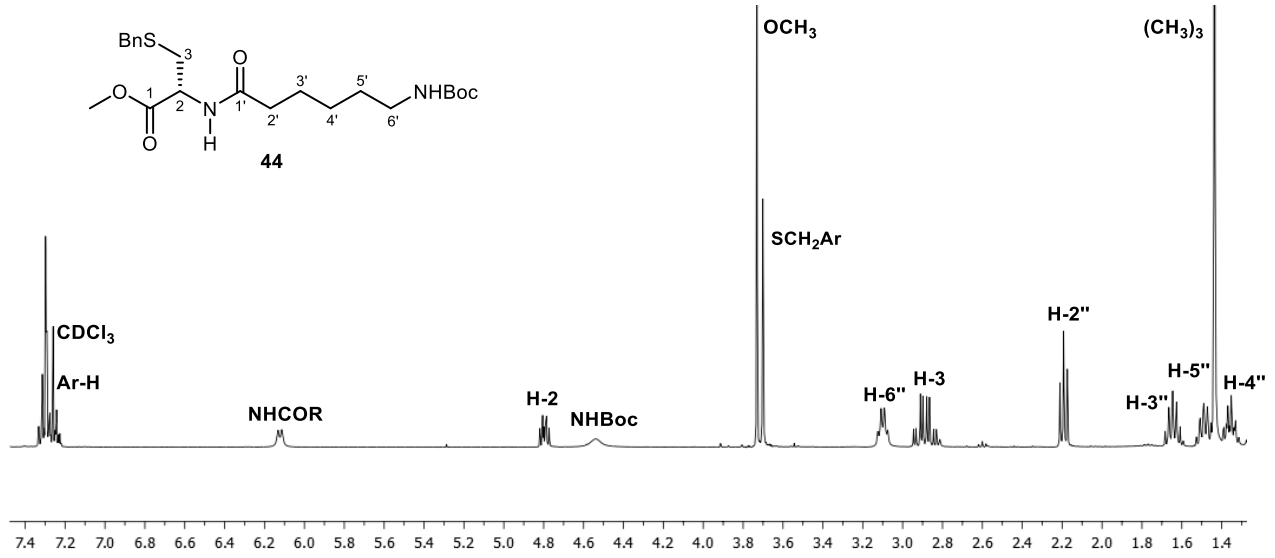


Figure 13: The ^{1}H NMR spectrum (1.35 - 7.40 ppm) of the protected *N*-acyl cysteine **44**.

The next step involved the removal of the *S*-benzyl protecting group of **44**. Arguably the most used method to achieve this makes use of the reductive power of sodium in liquid ammonia, developed by *du Vigneaud et al.* in the 1930's.[53,54] However, with this not being readily available, catalytic hydrogenation of the benzyl group was attempted as the alternative, although reports of the limitations of this approach due to poisoning of the Pd catalyst by sulfur were noted.[59]

However, when the protected *N*-acyl cysteine **44** in the presence of Pd/C (10 %, 10.0 eq.) in dry methanol, was treated under a balloon of H_2 gas, there was no evidence for conversion even after 48 h, and only starting material **44** was recovered. The recovered starting material **44** was therefore subjected to catalytic phase transfer hydrogenation (CPTH) involving the use of a zinc dust as a catalyst and ammonium formate in dry methanol, but this also failed to remove the benzyl group, even after extending the duration of the reaction and adding excess ammonium formate.[60,61] It was thus clear that the use of either the *S*-benzyl or *S*-*tert*-butyl protecting groups was not appropriate. Use of alternative thiol protecting groups, such as *p*-nitrobenzyl, *p*-methoxybenzyl and triphenylmethyl were not therefore not considered as they would be associated with similar limitations.[62]

This led to a return at attempting the preparation of **40** *via* Route **C**, depicted earlier (Scheme 16, Page 35), utilizing the amine **36** and taking into account the possibility of reducing any disulfide formed in the process due to the presence of the free thiol group.

2.3.3. Route *via* an *N*-Boc acid and methyl L-cysteinate

In order to prepare the *N*-Boc-GA **40**, the amine **36** was treated with the succinimidyl ester **45** under basic conditions in a mixture of acetonitrile and water (1:1) instead of dichloromethane in an attempt to deal with the poor solubility of **36** during the initial stages of the reaction.[63] After 18 h, two products with similar polarity were isolated as white solids after repeated column chromatography. The ^{1}H NMR spectra (Figure 14) of both solids are compared and show broad similarities but key differences in the region 3.00-3.20 ppm as highlighted.

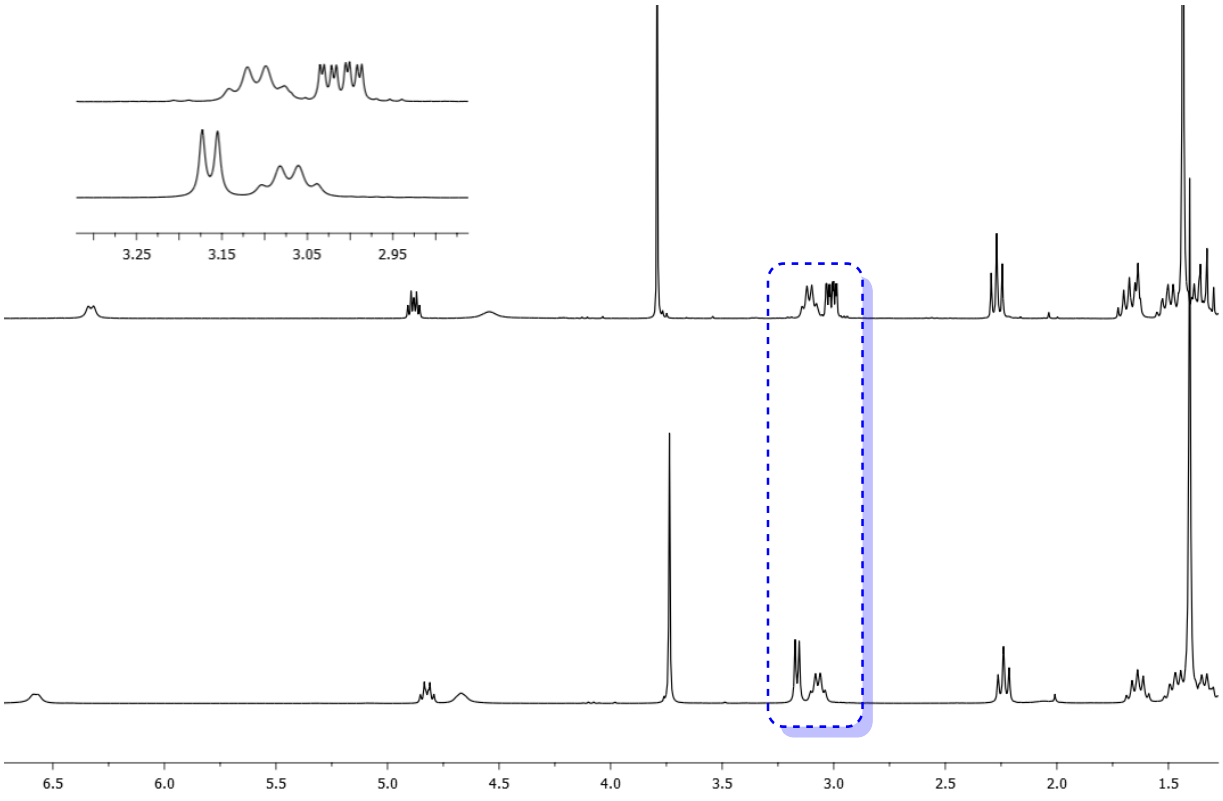
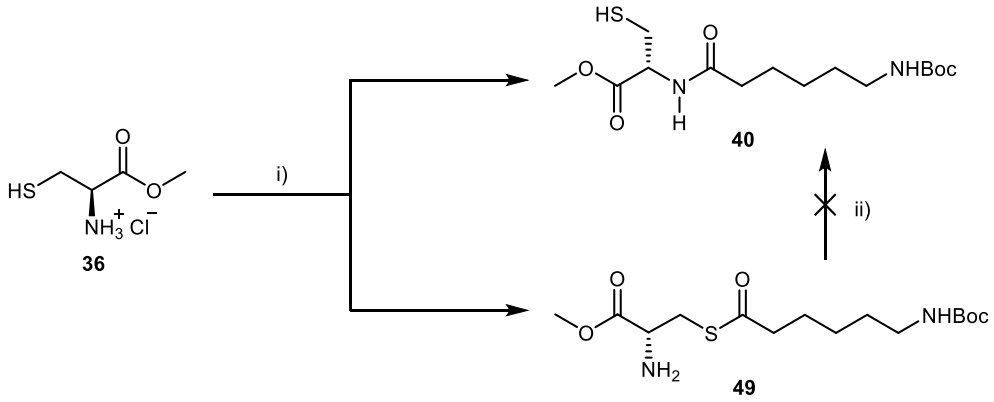


Figure 14: A comparison of the ^{1}H NMR spectra (1.25 - 6.75 ppm) of the presumed *N*-acyl and *S*-acyl products. Despite sharing broad similarities a key difference occurs in the region (3.00 - 3.20 ppm, encircled in blue). A zoomed in section of this region is also shown in the top left corner.

In addition the solids had different melting points yet the same molecular ions $[M + H]^+$ of m/z 349.1798 and m/z 349.1783, the difference being accepted within the given margin of error, in their respective HRMS spectra; suggesting that they could be the isomeric *N*-acyl product **40** and *S*-acyl product **49** shown in Scheme 21.



Scheme 21: Reagents and conditions; i) NEt_3, H_2O:MeCN (1:1), 0 °C for 1 h then r.t. for 17 h, (**40**, 76 %) and (**49**, 16 %); ii) NEt_3, DCM, r.t., 6 h.

This was confirmed by 2D NMR (HSQC, COSY). From their respective COSY spectra the less polar product was identified as the desired *N*-acyl product **40**, with the more polar product being identified as the intermediate *S*-acyl product **49**. The key evidence was the signal in the COSY spectrum of **40**, shown below (Figure 15), corresponding to the homonuclear coupling between the diastereotopic protons (H-3) and the thiol proton (S-H). The *S*-acyl derivative **49** was isolated as the minor product and attempts were made to convert it to the target molecule **40** *via* *S*- to *N*-acyl transfer as illustrated earlier (Scheme 6, Page 16) by stirring **49** in a solution of dry dichloromethane containing triethylamine for several hours. Unfortunately no conversion of **49** to **40** was observed. This could indicate that an inter- rather than intra-molecular *S*- to *N*-acyl transfer takes place in going from **49** to **40** or that the steric bulk of the acyl group hinders the formation of the intermolecular 5-membered-ring transition state as suggested by *El-Gendy et al.*[64] However this was not explored any further as a substantial amount of **40** (≈ 5.00 g) had been prepared at the time and the project focus pertained to the synthesis of the SGC probe.

*IUPAC name: methyl *S*-(6-((*tert*-butoxycarbonyl)amino)hexanoyl)-L-cysteinate (**49**)

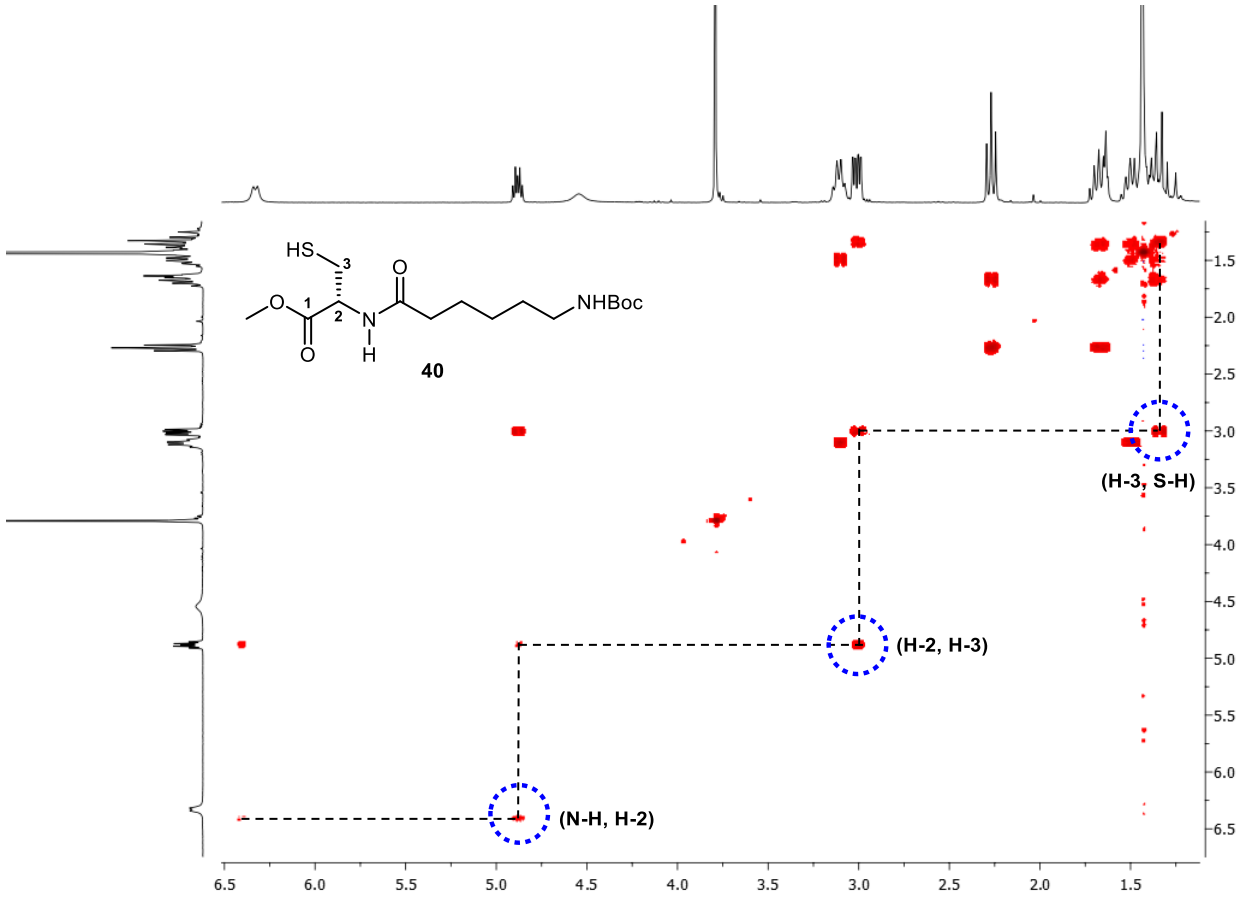


Figure 15: The COSY NMR spectrum (1.25 - 6.50 ppm) of the *N*-acyl product **40**. The key homonuclear couplings (encircled in blue) have been assigned.

With the *N*-Boc-GA **40** now in hand the mono-glucosylation with penta-acetyl glucopyranose **3** could be attempted, under conditions similar to that used in the attempted formation of BGC **1**. Unfortunately attempts at coupling **40** with **3** using $SnCl_4$ or $BF_3.OEt_2$ as promoters met with limited success once again, with a significant amount of unreacted **3** and **40** detected by TLC. Increasing the quantity of the Lewis acid or use of extended reaction times did not improve matters.

In light of this a decision was made to identify ideal reaction conditions leading to *S*-glycosylcysteine formation, under Lewis acid promotion, rather than to waste further precious material.

2.4. Glycosylation of cysteine derivatives with glycosyl acetates

The target *S*-glycosylcysteine derivatives selected for this study were; the 2-*O*-acetyl derivatives **2** and **19**, and their respective 2-acetamido analogues **50** and **51** illustrated in Figure 16.

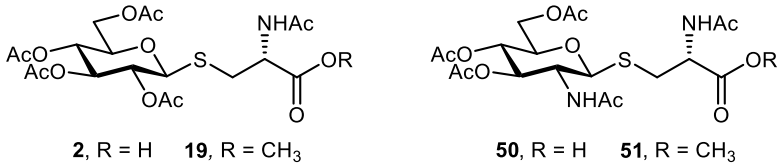
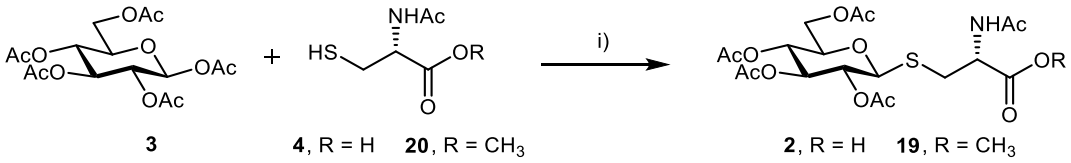


Figure 16: The target 2-*O*-acetyl (left) and 2-acetamido (right) *S*-glycosylcysteine derivatives.

They were chosen with a view to investigate the influence of the free carboxylic acid on the formation and stability of the β-thioglycoside, and to incorporate a biologically relevant sugar unit such as the 2-acetamido sugar. Targets **2** and **19** were prepared using the methods established by *Nokwequ et al.*, although no attempts were made to optimize the yields obtained thus far.[7,20] Initial attempts at preparing **50** by analogous routes were unsuccessful, leading to development of a new strategy. The detailed report of this synthesis follows.

Mono-glucosylation of *N*-acetyl-L-cysteine **4** with penta-acetyl glucopyranose **3** in dry dichloromethane using $SnCl_4$ was performed first, according to the literature procedure to form the desired product **2** which was isolated as a white foam in low yield (Scheme 22) after extensive column chromatography. The low yield obtained is consistent with the work of *Nokwequ et al.* where the instability of *S*-glycosylcysteine derivatives having a free carboxylic acid group was observed.[20]

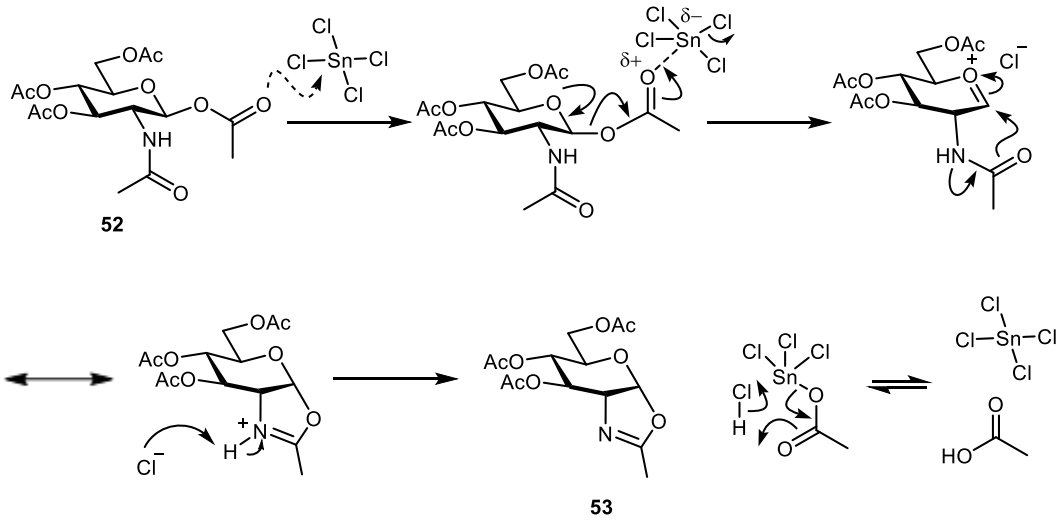


Scheme 22: Reagents and conditions; i) $SnCl_4$, DCM, r.t., 3 h, (**2**, 34 %) or (**19**, 64 %).

*IUPAC names: *N*-acetyl-*S*-(2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl)-L-cysteine (**50**)
methyl *N*-acetyl-*S*-(2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucospyranosyl)-L-cysteinate (**51**)

When the above procedure was repeated with methyl *N*-acetyl-L-cysteinate **20** there was noticeable absence of any hydrolyzed tetra-acetyl glucopyranose **28** and only a minimal amount of unreacted starting material **3** after 3 h. The polar product was isolated as a white solid in a 64 % yield and the ^{1}H NMR spectrum for the solid was in agreement with literature data for this compound, confirming the formation of **19** shown above in Scheme 22.[7,20,23] It was thus confirmed that the methyl ester allowed for a more selective reaction, made the isolation of the product less troublesome and that the presence of a carboxylic acid in the reaction plays a role in facilitating the formation of **28** under Lewis acid promoted conditions. The possible interaction of the carboxylic acid with $SnCl_4$ warrants further investigation.

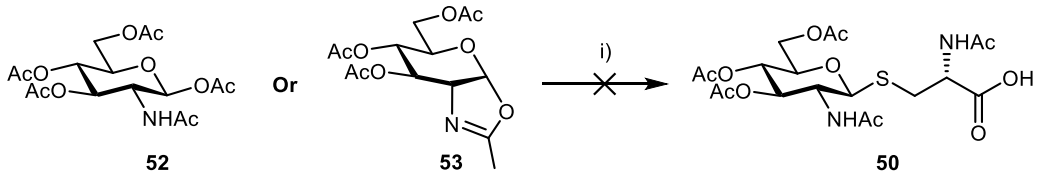
Unfortunately attempts at repeating the mono-glycosylation of **4** with the 2-acetamido glucopyranose **52** were unsuccessful. However, a single product was isolated as an amber oil in a low yield of 11 % and identified by ^{1}H NMR and comparison with literature data as the oxazoline **53**.[65] The low yield obtained is attributed to the poor solubility of the reactants and work-up used. The formation of the oxazoline **53** is explained mechanistically in Scheme 23.[66]



Scheme 23: The mechanism for the formation of **53** from **52**.[65] The regeneration of the catalyst is supported by the work of *Srivastava* who performed the same conversion using only a catalytic amount of $SnCl_4$.[66]

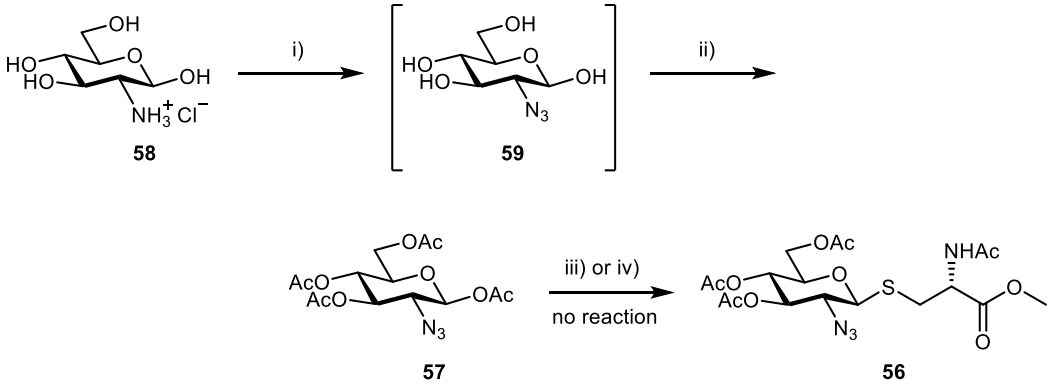
*IUPAC names: 2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranose (**52**)
2-methyl-(3,4,6-tri-*O*-acetyl-1,2-dideoxy-α-D-glucopyrano)-[2,1-*d*]-2-oxazoline (**53**)

Oxazoline **53** is a potential glycosyl donor in its own right, but it was apparent from the reaction outcome that it was not activated by $SnCl_4$. To verify this, additional oxazoline **53** was prepared in a simple two-step process by first acetylating 2-acetamido-2-deoxy-D-glucopyranose **54** with an excess of acetic anhydride in pyridine to afford an anomeric mixture of **52** and its α-anomer **55** in a combined yield of 91 % and 9:1 ratio (α:β). The mixture of anomeric acetates were then refluxed in the presence of excess TMSOTf to yield the oxazoline **53** in a 76 % yield.[67] Treatment of **53** with **4** in the presence of $SnCl_4$ gave the same result as before (no reaction) confirming that both **52** and **53** are not sufficiently reactive under these conditions as shown in Scheme 24.



Scheme 24: The attempted conversion of the 2-acetamido glucopyranose **52** and oxazoline **53** to the desired *S*-glycosylcysteine derivative **50**, under the reaction conditions; i) **4**, $SnCl_4$, DCM, r.t., 3-24 h.

Due to the results illustrated (Scheme 24) a revised approach was considered, using a 2-azido glycosyl derivative as a glycosyl donor. It was thus decided to attempt formation of the *S*-glycosylcysteine derivative **56** using the 2-azido glucopyranose **57** which could be readily obtained from the 2-amino glucopyranose **58**, as outlined in Scheme 25.[20,42,68]



Scheme 25: Reagents and conditions; i) K_2CO_3, $CuSO_4.5H_2O$, **33**, MeOH, r.t., 3 h; ii) Ac_2O, pyridine, 24 h, r.t., 60 % over two steps; iii) $SnCl_4$, DCM, r.t., 3 h; iv) $BF_3.OEt_3$, DCM, 0 °C for 1 h then r.t. for 23 h.

*IUPAC names: methyl *N*-acetyl-*S*-(2-azido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl)-L-cysteinate (**56**)
2-azido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranose (**57**)
2-amino-2-deoxy-D-glucopyranose hydrochloride (**58**)

Briefly, **58** was reacted with the sulfonyl azide **33** under typical diazotransfer conditions as described before and after 2 h a distinct colour change indicated the formation of intermediate **59** shown above in Scheme 25.[42] Co-evaporation of residual solvent from the reaction mixture with toluene was followed by the addition of excess acetic anhydride and pyridine to the crude residue containg **59** (not isolated). After 24 h a single less polar product was isolated as a white crystalline solid in good yield after standard work-up. The ^{1}H and ^{13}C NMR data for the solid were in agreement with literature data for this compound, confirming the formation of the 2-azido glucopyranose **57**.[69,70] This was further corroborated by the IR-spectrum of the solid showing a characteristic azido band at 2113 cm^{-1}. Unfortunately attempts at converting **57** to **56** with either $SnCl_4$ or $BF_3.OEt_2$ were unsuccessful as TLC revealed no product formation, and other Lewis acids were not considered at this time.

At this point it had become evident that the glycosylation of cysteine derivatives with glycosyl acetates under Lewis acidic conditions were giving undesired and or ambiguous results. It must be noted that glycosyl acetates are not regarded as the best glycosyl donors, but were explored here in order to follow up on the original work: it is quite possible that alternative glycosyl donors, such as glycosylimidates, halides, sulfides, etc could be more effective in combination with appropriate promotors,[71,72] but a decision was made to rather pursue the synthesis of the desired SGC probe *via* a nucleophilic glycosyl thiol and a cysteine precursor with a suitable leaving group in the side chain.

*IUPAC name: 2-azido-2-deoxy-β-D-glucopyranose (**59**)

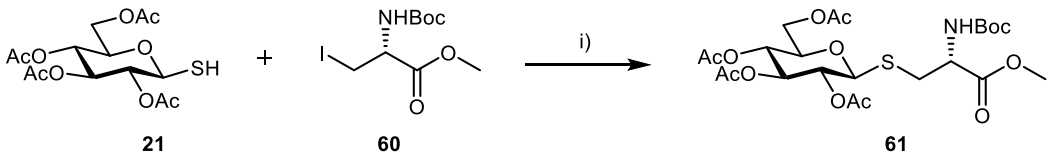
2.5. Alternative synthesis of the SGC and SAGC probes

As was highlighted earlier (Scheme 7, Page 17), an alternative approach to making *S*-glycosylcysteine derivatives involves a straightforward S_N2 substitution in which the anomeric thiol of the desired sugar displaces an appropriate leaving group present on the electrophilic substrate. It was hoped that this alteration in strategy would circumvent any of the undesired reactions from taking place that had been plaguing the anomeric sites of both **3** and **52**, as was discussed in detail earlier.

This alteration in strategy paid dividends in the end and was extended to the successful synthesis of the desired SGC probe *via* a convergent synthesis. In addition to this, repetition of this approach with the 2-acetamido sugar derivative, pleasingly, also led to the successful synthesis of an *S*-(*N*-acetylglucosaminyl)-L-cysteinate (SAGC) probe. A detailed account of this synthesis follows.

2.5.1. Glycosylation of a masked cysteine derivative with 1-thioglucopyranose

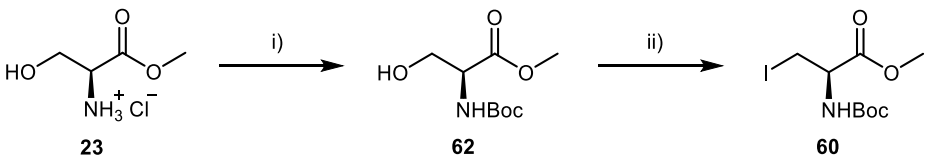
In order to implement this alternative strategy, 1-thioglucopyranose **21** and 3-iodopropanoate **60** were identified as key synthetic intermediates, to allow for preparation of *S*-glucosylcysteine derivative **61** as depicted in Scheme 26.[24,27]



Scheme 26: The revised synthetic approach towards *S*-glucosylcysteine derivative **61** based on literature.

Compounds **21** and **60** were both successfully prepared according to literature procedures and coupled to give **61** in good yield. The detailed report of this synthesis follows, starting with the preparation of **60** as outlined below in Scheme 27.[73,74]

*IUPAC names: methyl (*R*)-2-((*tert*-butoxycarbonyl)amino)-3-iodopropanoate (**60**)
methyl *N*-(*tert*-butoxycarbonyl)-*S*-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-L-cysteinate (**61**)



Scheme 27: Reagents and conditions: i) aq. $NaHCO_3$ (2.3 M), Boc_2O, THF, 0 °C - r.t. over 1 h then r.t. for 6 h, 90 %; ii) I_2, PPh_3, imidazole, DCM, 0 °C for 2 h then 0 °C - r.t. over 4 h, 78 %.

The first step involved treating the amine **23** with Boc-anhydride in methanol under mild basic conditions for 6 h to form the *N*-Boc protected alcohol **62** which was isolated as a colourless oil in high yield.[73] This was confirmed by the ^{1}H NMR spectrum of the oil which showed a nine-proton singlet at 1.43 ppm, consistent with the formation of the carbamate.

The *N*-Boc protected alcohol **62** was then subjected to the Appel reaction to form the 3-iodopropionate **60** which was isolated as a white solid in good yield, despite a lengthy work-up required to remove the triphenylphosphine oxide.[74] The formation of **60** was confirmed by ^{13}C NMR (Figure 17) which showed a diagnostic one-carbon singlet at 7.9 ppm, consistent with the presence of an alkyl iodide.

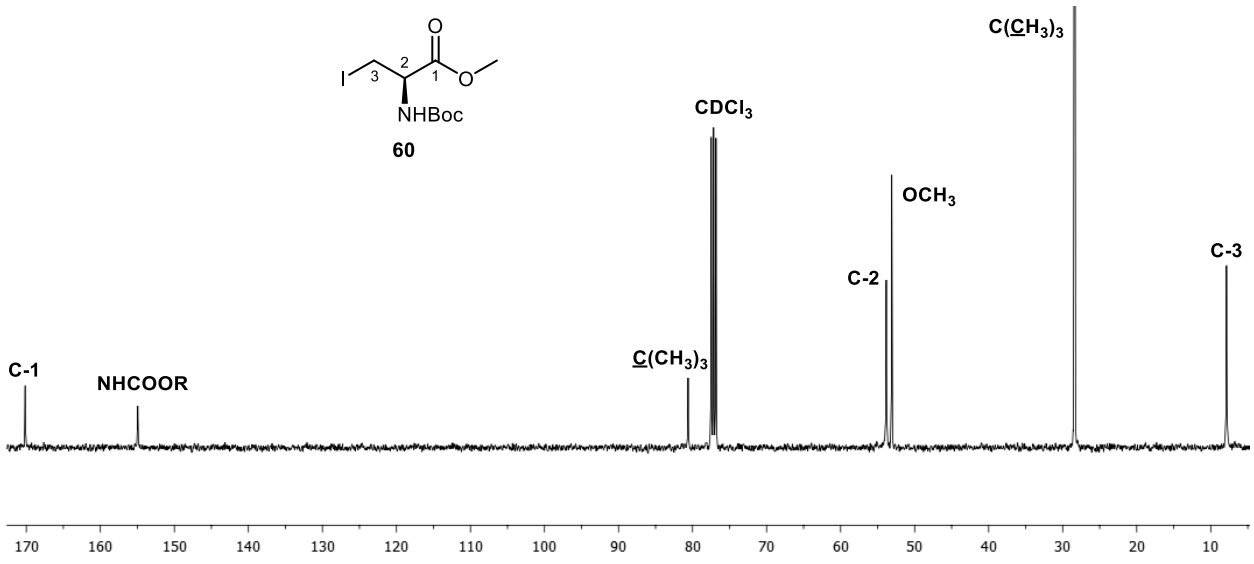
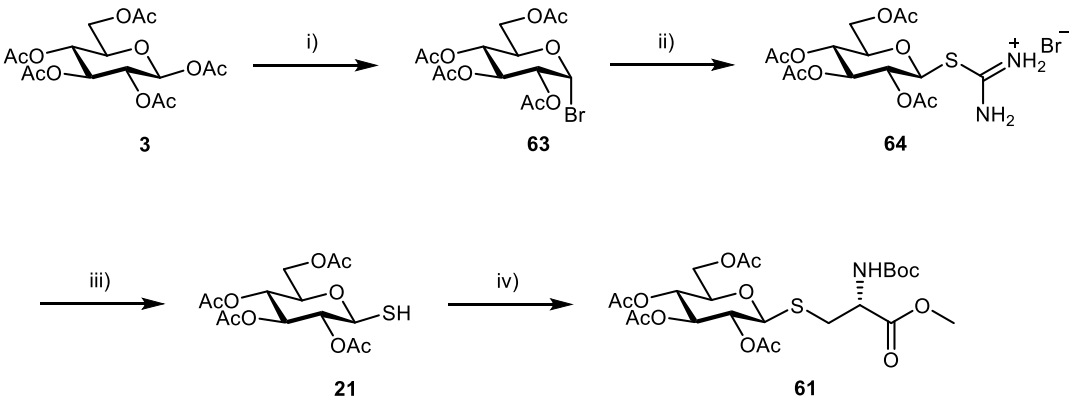


Figure 17: The ^{13}C NMR spectrum (5.0 - 172.0 ppm) of the 3-iodopropionate **60.**

*IUPAC name: methyl *N*-(*tert*-butoxycarbonyl)-L-serinate (**62**)

The attention then turned to preparing the 1-thioglucopyranose **21** and coupling this with the 3-iodopropionate **60**, as outlined in Scheme 28.[24,26,27]



Scheme 28: Reagents and conditions: i) 33 % HBr in AcOH, DCM, 0 °C for 1 h then r.t. for 2 h, 76 %; ii) thiourea, acetone, reflux, 2 h, 60 %; iii) $Na_2S_2O_5$, DCM:H_2O (2:1), 50 °C, 3 h, 95 %; iv) **60**, TBAHS, aq. $NaHCO_3$ (0.6 M), EtOAc, 40 °C, 4 h, 75 %.

The first step involved refluxing penta-acetyl glucopyranose **3** with a solution of 33 % HBr in acetic acid for 4 h to form the glucosyl bromide **63** which was isolated as a white solid in good yield.[26] The ^{1}H and ^{13}C NMR spectra, as well as m.p. for the solid were in agreement with literature data for this compound, confirming the formation of **63**.[26]

The glucosyl bromide **63** was then reacted with thiourea in freshly dried acetone at reflux to form the isothiouronium bromide salt **64**, which had precipitated from solution after 2 h, and was obtained as a white solid in moderate yield.[26] This was confirmed by the ^{1}H NMR spectrum of the solid which showed two-proton singlets at 9.23 and 9.07 ppm, corresponding to the amino groups of the thiourea moiety.

Subsequent treatment of the isothiouronium bromide salt **64** with sodium metabisulphite in a mixture of dichloromethane and water at 50 °C gave the 1-thioglucopyranose **21**, isolated as a white solid in high yield.[26] The ^{1}H NMR spectrum, shown below (Figure 18), of the solid showed the presence of a one-proton doublet at 2.30 ppm consistent with the presence of the free thiol at the anomeric position, having a β-configuration, confirming the formation of **21**.

*IUPAC names: 2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranosyl bromide (**63**)
2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranosyl-1-isothiouronium bromide (**64**)

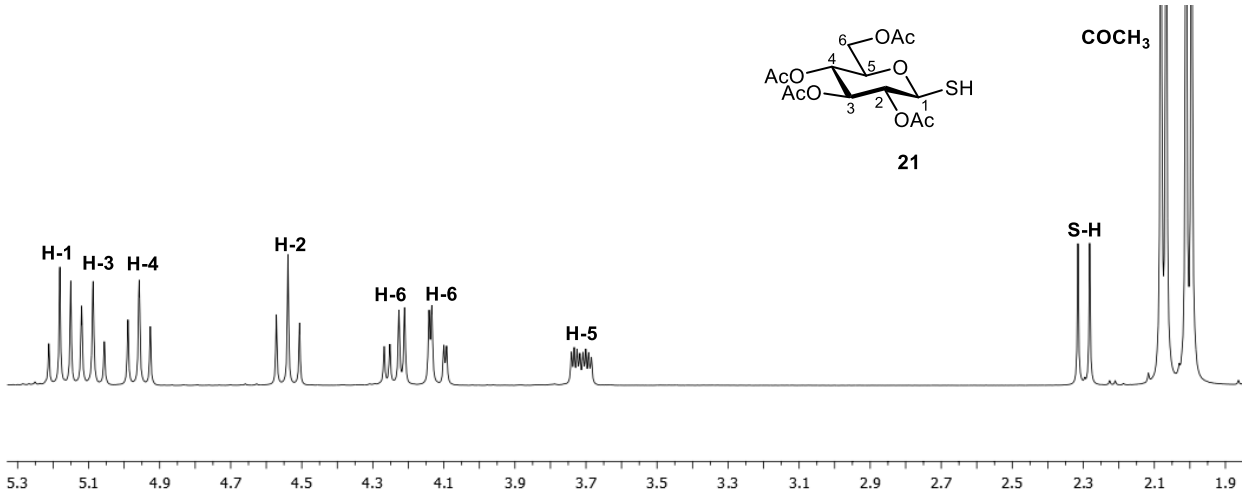


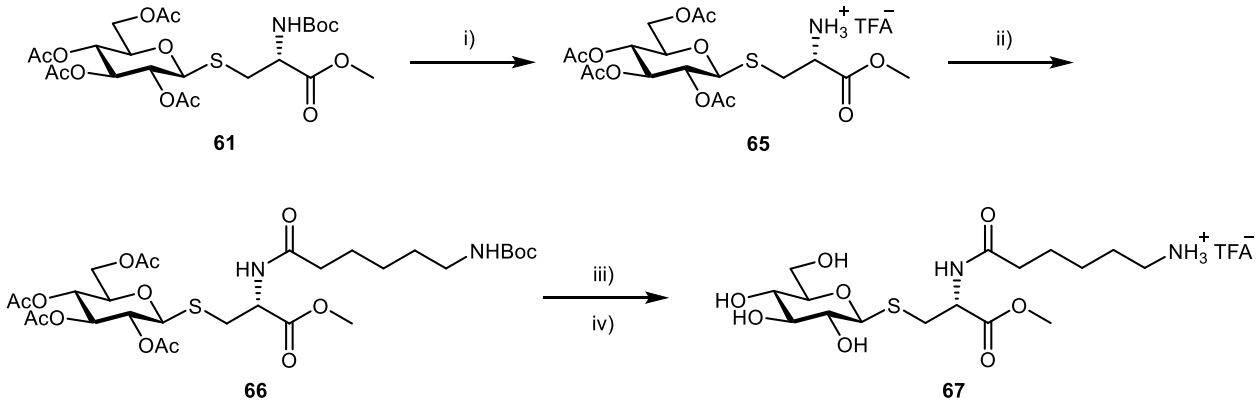
Figure 18: The ^{1}H NMR spectrum (1.90 - 5.30 ppm) of the 1-thioglucopyranose **21**.

The final step then required the S_N2 substitution of **60** (in a slight excess) with 1-thioglucopyranose **21**, and this was carried out under basic conditions in a mixture of freshly distilled ethyl acetate and water under phase-transfer catalyzed conditions using tetrabutylammonium hydrogen sulfate (TBAHS), and at 40 ^{o}C in order to ensure complete solubility of the reagents.[27] After 4 h TLC showed the complete consumption of **21** and formation of a single less polar product which was isolated as a white solid in good yield. The ^{1}H NMR spectrum of the solid was consistent with those reported in literature, confirming the formation of **61**.[21] It was noted that the reported procedure made use of an excess of **21** and a differently protected 3-bromopropionate derivative, but in our procedure an excess of the 3-iodopropionate **60** made chromatographic purification less troublesome, despite a relative reduction in yield of approximately 10 %.[27]

With the desired *S*-glucosylcysteine derivative **61** in hand the attention now shifted to resuming the synthesis of the SGC probe.

2.5.2. Synthesis of the target SGC probe

The final steps of the sequence towards the SGC probe; involved *N*-deprotection, followed by *N*-acylation with the required tether and then final global deprotection as outlined in Scheme 29.[47,72,75]



Scheme 29: Reagents and conditions; i) TFA, DCM, 0 °C for 1 h then r.t. for 2 h; ii) **45**, NEt_3, DCM, 0 °C for 1 h then r.t. for 35 h, 69 % over two steps; iii) NaOMe (0.1 M), MeOH, r.t., 1 h; iv) TFA, DCM, 0 °C for 1 h then r.t. for 2 h, 63 % over two steps.

In the first step the *N*-Boc group was removed under standard conditions (TFA/DCM) to give the amine **65** which was isolated as a colourless oil.[72] This was confirmed by the ^{1}H NMR spectrum, depicted below (Figure 19), of the oil which showed the disappearance of the nine-proton singlet at 1.44 ppm for the carbamate of **61**, consistent with the formation of the amine. The presence of the TFA counterion was evident from the three-fluorine singlet at -75.5 ppm in the ^{19}F NMR spectrum. It must be noted, however, that when this reaction was conducted on a large scale (≈ 1.00 g) the removal of residual solvent proved rather time consuming. In light of this the amine **65** was taken forward to the next step without further drying.

The amine **65** was then subjected to the same *N*-acylation method that led to the formation of protected *N*-acyl cysteine derivatives **42** and **44** in dichloromethane, as described earlier.[47] After 36 h TLC revealed the disappearance of the succinimidyl ester **45** together with the formation of a less polar product, which was duly isolated as a colourless oil in good yield.

*IUPAC name: methyl *S*-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-L-cysteinate trifluoroacetate (**65**)

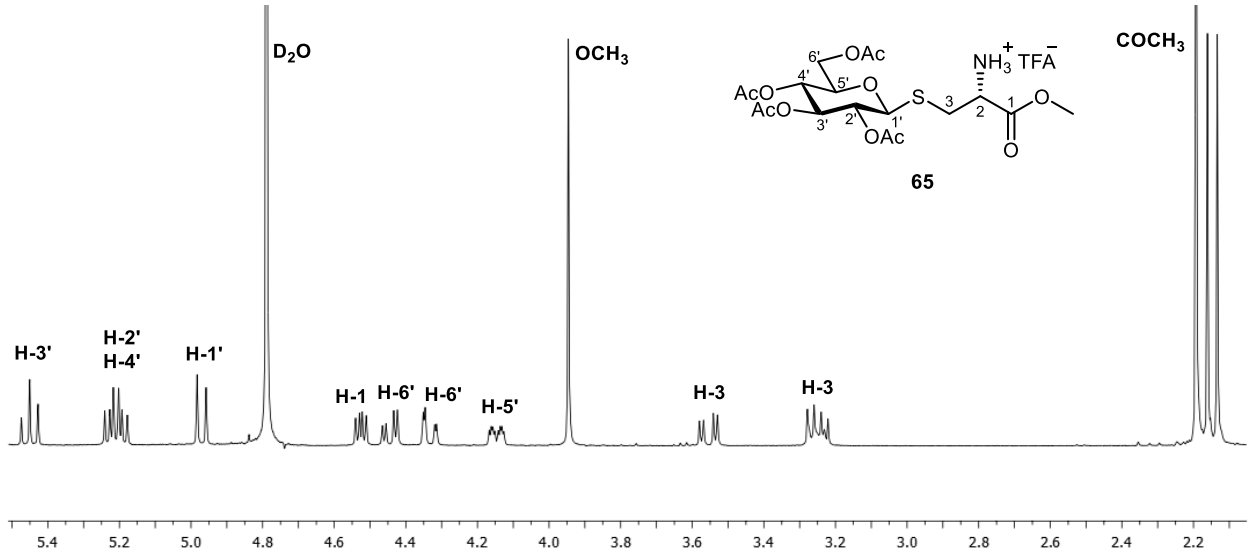


Figure 19: The ^{1}H NMR spectrum (2.00 - 5.60 ppm) of the amine product **65**, obtained from the *N*-Boc deprotection of **61** conducted on a 200 mg scale.

The ^{1}H NMR spectrum of the oil included a one-proton doublet at 6.41 ppm consistent with the presence of the amide bond, confirming the formation of tethered cysteinyl glucoside **66**. The formation of **66** was further corroborated by the ^{13}C NMR spectrum having 25 signals, which correlated to the required 29 carbon atoms taking into account signals which coincide due to symmetry and signal overlap, and the detection in the HRMS spectrum of a molecular ion $[M + H]^{+}$ of the required m/z 679.2758.

Global de-*O*-acetylation of **66** was then achieved under typical Zemplén conditions (cat. NaOMe in MeOH)[75] and subsequent *N*-Boc deprotection, again under standard conditions (TFA/DCM), gave the deprotected cysteinyl glucoside **67** which was isolated as an off-yellow oil in moderate yield. This was confirmed by the ^{1}H NMR spectrum of the oil which showed the disappearance of the five singlets for the *tert*-butyl and acetyl protons, of **66**, consistent with the formation of deprotected glucoside **67**. The formation of **67** was further corroborated by the ^{13}C NMR spectrum having the required 17 signals, and the detection in the HRMS spectrum of a molecular ion $[M]^{+}$ of the required m/z 411.1800.

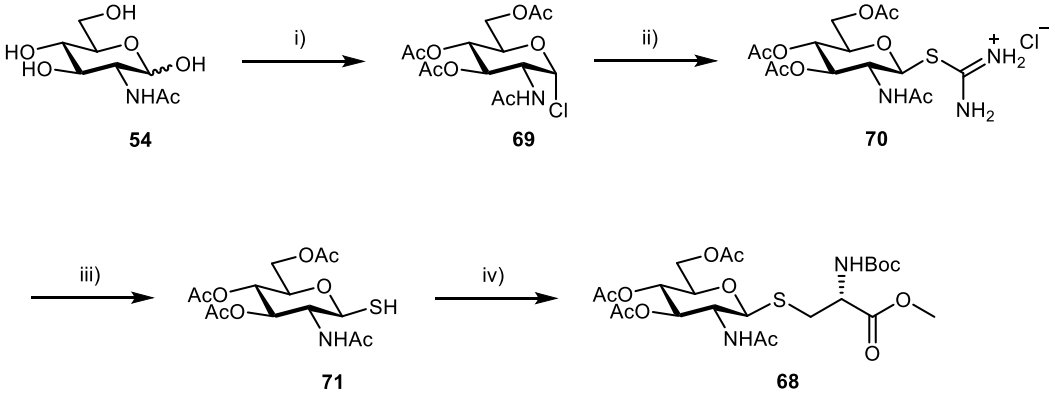
*IUPAC names: methyl *N*-(6-((*tert*-butoxycarbonyl)amino)hexanoyl)-*S*-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-L-cysteinate (**66**) methyl *N*-(6-aminohexanoyl)-*S*-(β-D-glucopyranosyl)-L-cysteinate trifluoroacetate (**67**)

With the successful preparation of the SGC probe in a moderate yield; the attention turned establishing whether the methodology could be used to introduce an alternative sugar unit.

2.5.3. Synthesis of the target SAGC probe

The incorporation of an alternative sugar unit to that of glucopyranose was pursued for three prominent reasons. Firstly it would afford an opportunity to test the generality of the new synthetic approach to glycosyl cysteine probes. Secondly it would allow for the incorporation of sugar units more commonly found in biologically relevant glycans. Thirdly it would be the starting point for the generation of a library of glycosyl cysteine probes.

For this investigation mannopyranose, galactopyranose and the 2-acetamido glucopyranose **54** were all considered as viable sugar units in the first instance, with the incorporation of **54** being pursued as a proof-of-concept. The synthesis of the SAGC probe, pleasingly, was also successful and began with the preparation of *S*-glycosylcysteine derivative **68** as shown in Scheme 30.

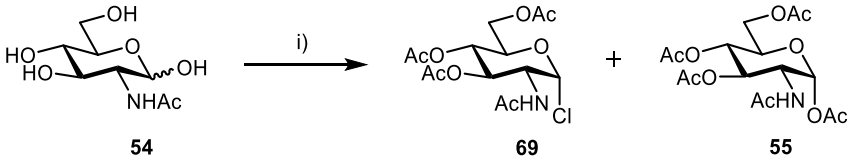


Scheme 30: Reagents and conditions; i) AcCl, DCM, 30 °C, 14 h, ≈ 57 %; ii) thiourea, acetone, reflux, 2 h, ≈ 64 %; iii) $Na_2S_2O_5$, DCM:H_2O (2:1), 50 °C, 3 h, 88 %; iv) **60**, TBAHS, aq. $NaHCO_3$ (0.6 M), EtOAc, 40 °C, 4 h, 75 %.

As illustrated (Scheme 30) the key *S*-glycosylcysteine derivative **68** was prepared from the isothiouronium chloride salt **70**, as described earlier for its 2-*O*-acetyl analogue **61**. Details for the synthesis and characterization of compounds **68**, **70** and **71** are equivalent to those described earlier for their 2-*O*-acetyl analogues.[24,26,27,76,77]

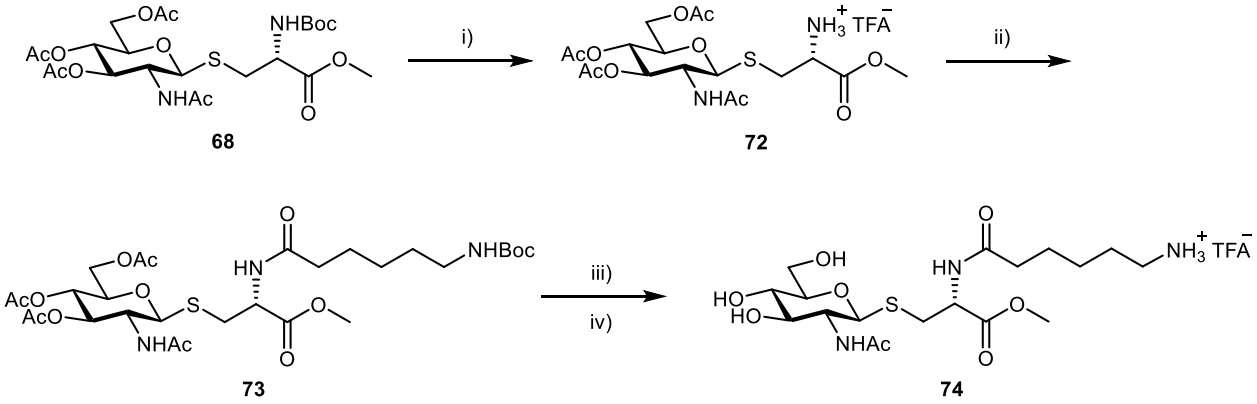
*IUPAC names: methyl *N*-(*tert*-butoxycarbonyl)-*S*-(2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl)-L-cysteinate (**68**)
2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-glucopyranosyl-1-isothiouronium chloride (**70**)
2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-1-thio-β-D-glucopyranose (**71**)

The only discrepancy in the above sequence (Scheme 30) took place during the first step in which the glycosyl chloride **69** was isolated with a small amount of the α-glycosyl acetate **55** during attempts to convert **54** exclusively to **69**, through treatment with acetyl chloride for 14 h as shown in Scheme 31.[76] They were obtained in a ratio of 9:1 (**69**:**55**) as judged from integration of their respective anomeric proton signals in the ^{1}H NMR spectrum. The products were not separated from each other, on the basis that the subsequent step would afford the precipitated isothiouronium chloride salt **70**, leaving the undesired α-glycosyl acetate **55** in solution, which was found to be the case experimentally.



Scheme 31: Reagents and conditions: i) AcCl, DCM, 30 °C, 14 h, (**69**, ≈ 57 %) and (**55**, ≈ 6 %).

With the desired *S*-glycosylcysteine derivative **68** in hand the final steps of the sequence towards the SAGC probe were equivalent to those that led to the formation of the SGC probe and are discussed briefly and summarized in Scheme 32.[47,72,75]



Scheme 32: Reagents and conditions; i) TFA, DCM, 0 °C for 1 h then r.t. for 2 h; ii) **45**, NEt_3, DCM, 0 °C for 1 h then r.t. for 35 h, 78 % over two steps; iii) NaOMe (0.1 M), MeOH, r.t., 1 h; iv) TFA, DCM, 0 °C for 1 h then r.t. for 2 h, 69 % over two steps.

*IUPAC name: 2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-glucopyranosyl chloride (**69**)

Briefly, the *N*-Boc group of **68** was removed under standard *N*-Boc deprotection conditions (TFA/DCM) to form the amine **72** which was isolated as a colourless oil.[72] The amine **72** too was taken forward without extensive removal of stubborn residual solvent. The formation of **72** was confirmed for the same reasons as described earlier for its 2-*O*-acetyl counterpart **65**.

The amine **72** was then subjected to the same *N*-acylation method that led to the formation of the tethered cysteinyl glucoside **66** in dichloromethane, with the desired amide product **73** being isolated as a white solid in good yield.[47] A comparison of the ^{1}H NMR spectra (Figure 20) for isolated compounds **66** and **73** are compared to highlight the shielding of H-2' from 5.00 ppm to 4.00 ppm in going from compound **66** to **73**. Shifts of this magnitude were found to be common when comparing the ^{1}H NMR spectra of the 2-*O*-acetyl derivatives to their 2-acetamido analogues. The formation of **73** was further corroborated by the detection in the HRMS spectrum of a molecular ion $[M + H]^{+}$ of the required m/z 678.2921.

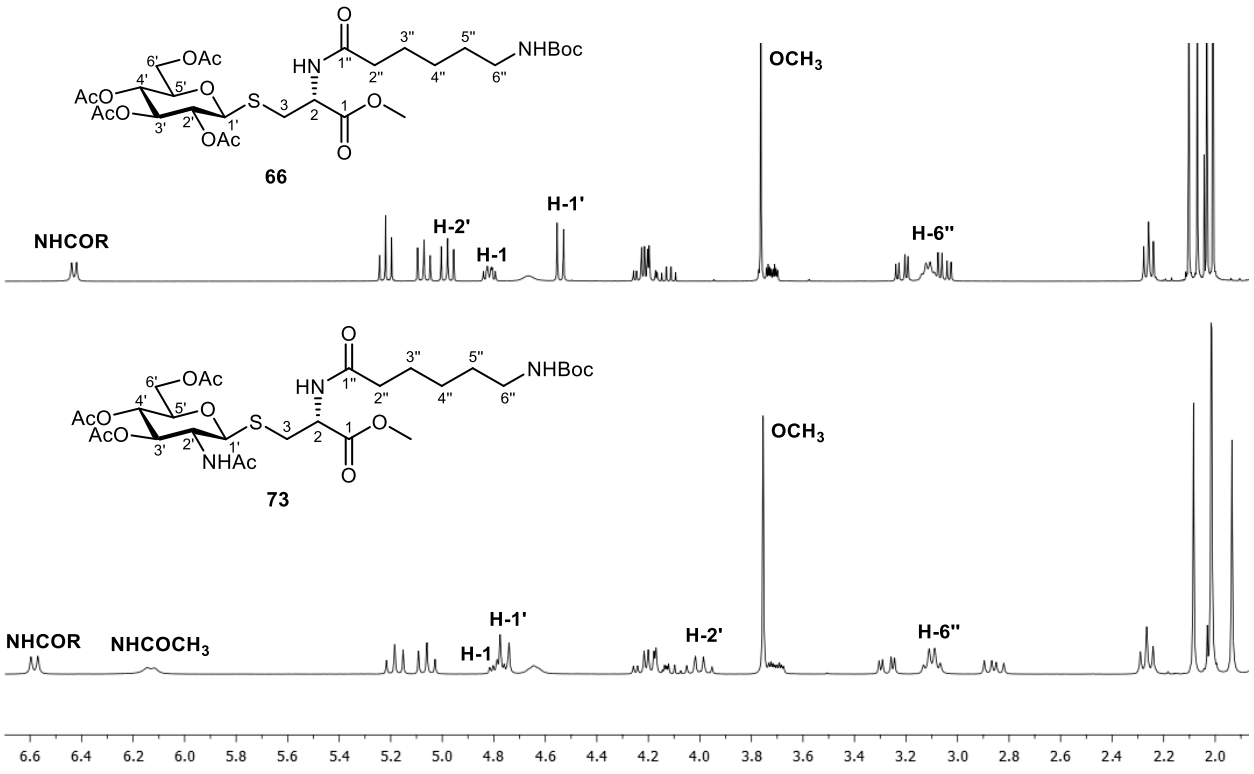


Figure 20: A comparison of the ^{1}H NMR spectra (1.90 - 6.70 ppm) of tethered cysteinyl glycosides **66** and **73**, with key signals assigned.

*IUPAC names: methyl *S*-(2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl)-L-cysteinate trifluoroacetate (**72**) methyl *N*-(6-((*tert*-butoxycarbonyl)amino)hexanoyl)-*S*-(2-acetamido-tri-3,4,6-tetra-*O*-acetyl-2-deoxy-β-D-glucopyranosyl)-L-cysteinate (**73**)

Global de-*O*-acetylation of **73** and subsequent *N*-Boc deprotection was achieved, again under the same conditions[72,75] that led to the formation of **67**, and gave the deprotected cysteinyl glycoside **74**, being isolated as a cream coloured oil in moderate yield. For confirmatory purposes a section of the HSQC spectrum (Figure 21) of the isolated target SAGC **74** is given, with all heteronuclear couplings assigned. The formation of **74** was further corroborated by the detection in the HRMS spectrum of a molecular ion $[M]^+$ of the required m/z 452.2062.

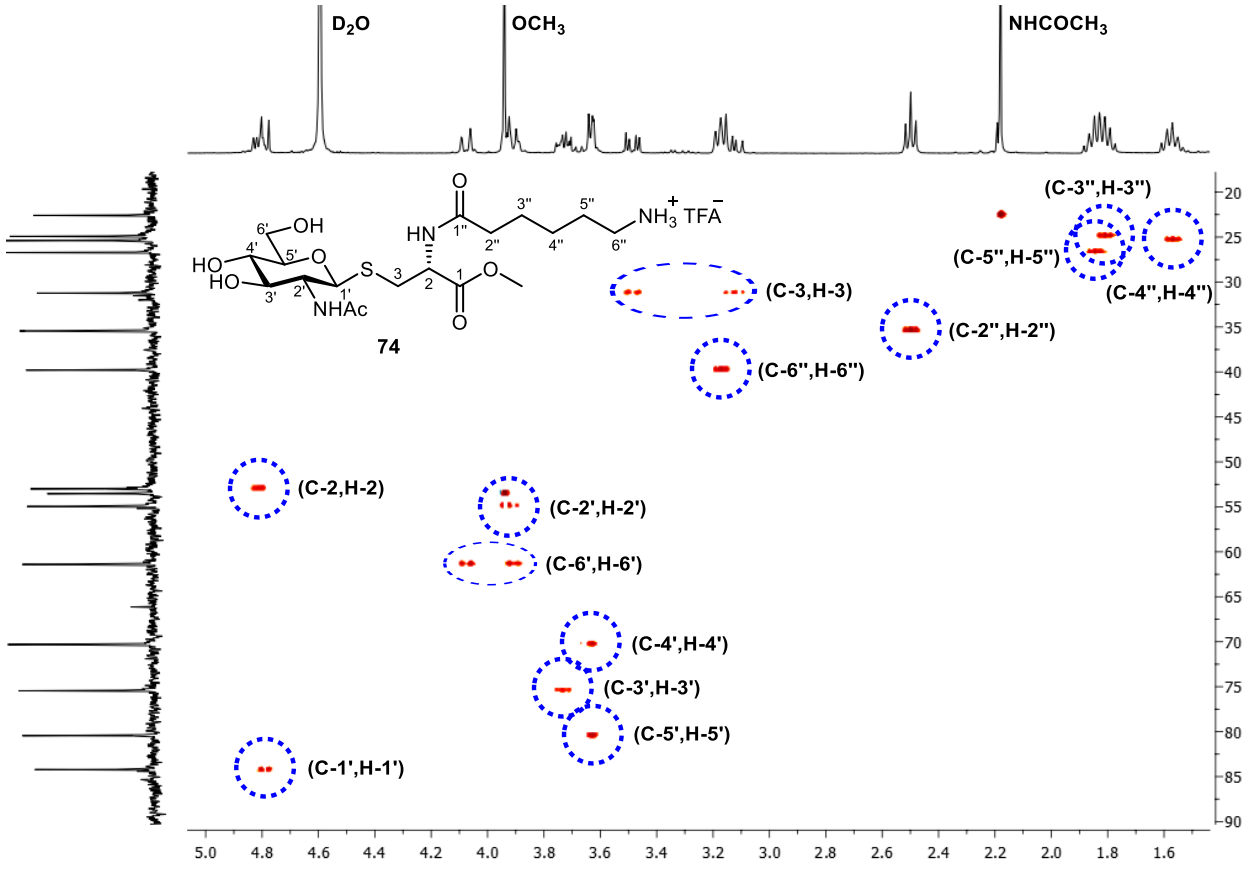


Figure 21: The HSQC spectrum (δ_H: 1.50 - 5.00 ppm/ δ_C: 20.0 - 90.0 ppm), of target SAGC **74**, referenced with respect to the OCH_3 signal at 3.94 ppm. All the heteronuclear couplings (encircled in blue) have been assigned.

*IUPAC name: methyl *N*-(6-aminohexanoyl)-*S*-(2-acetamido-2-deoxy-β-D-glucopyranosyl)-L-cysteinate trifluoroacetate (**74**)

Chapter Three

3. Conclusion and Future Work

3.1. Conclusion

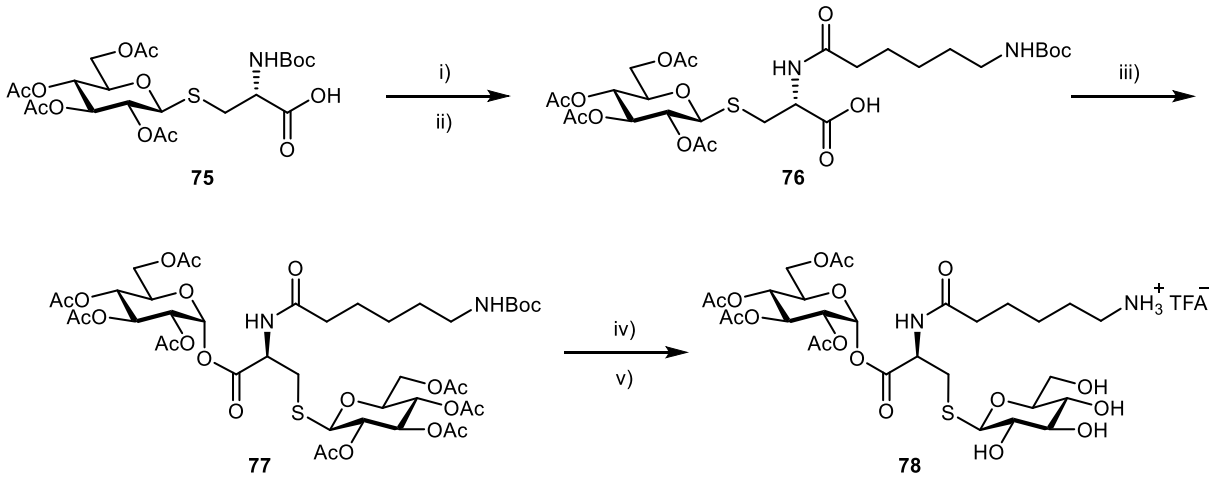
During the course of this project directed at the synthesis of cysteine-based glycan-binding probes it was found that:

The novel SGC probe **67** can be readily prepared in good yields from the *S*-glucosylcysteine derivative **61**, in a short, four-step sequence involving; *N*-Boc deprotection, *N*-acylation with the succinimidyl ester **45**, and finally a combination of de-*O*-acetylation and *N*-Boc deprotection. An alternative and more direct method, *via* treatment of penta-acetyl glucopyranose **3** with the *N*-Boc-GA **40** under Lewis acid promotion followed by *N*-Boc deprotection is still a worthwhile option, but requires further careful evaluation and optimization of reaction conditions. The S-glucosylcysteine derivative **61** in turn was prepared in good yields from **3**, in a short, four-step sequence starting with formation of a glucosyl halide, conversion of this to an isothiouronium salt, mild hydrolysis to generate the glucosyl thiol, and finally reaction of this thiol with the 3-iodopropanoate **60**. The above 8-step sequence was successfully replicated to afford the SAGC probe **74**, starting with 2-acetamido-2-deoxy-D-glucopyranose **54**.

Attempts at identifying efficient reaction conditions for a one-pot, bis-glucosylation of *N*-acetyl-L-cysteine **4** with **3** met with limited success: as in almost all cases instances of anomerization, hydrolysis, chloridation and mono-glucosylation of **3** were observed. An attempt at similar bis-glucosylation of *N*-acyl cysteine **12** with **3** was also unsuccessful. In addition, attempts at the mono-glycosylation of methyl *N*-acetyl-L-cysteinate **20** with the 2-deoxy-glycosyl acetates **52**, **53** and **57**, under $SnCl_4$ promotion, did not lead to the formation of the desired *S*-glycosylcysteine products, suggesting that they are not sufficiently reactive under the conditions used. This was an unexpected result, but does highlight that glycosyl acetates are not regarded as the best glycosyl donors. The selective removal of *S*-protecting groups *via* base-catalyzed elimination followed by disulfide bond cleavage (*tert*-butyl of **42**) or hydrogenation (benzyl of **44**), did not take place under the conditions used.

3.2. Future work

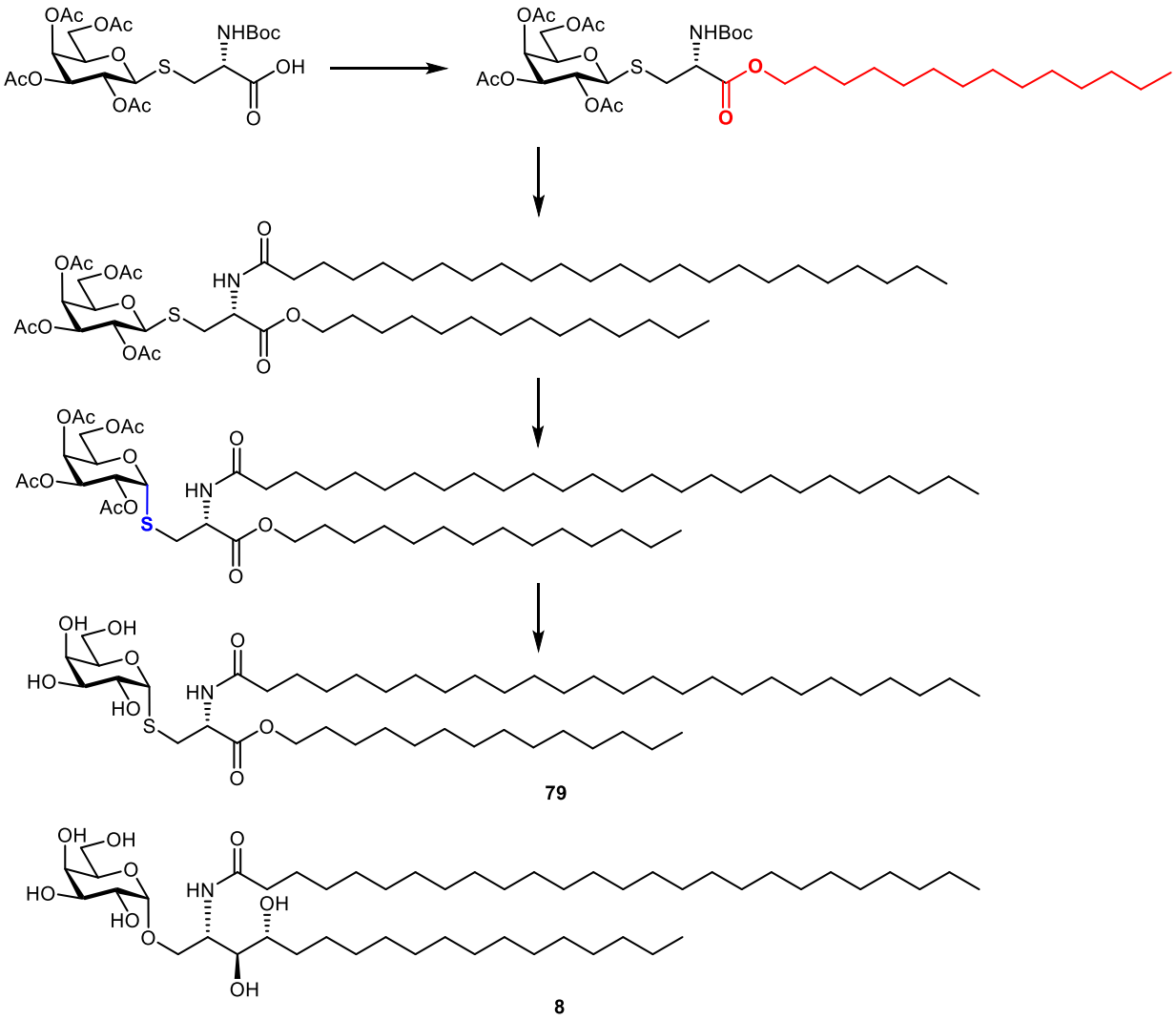
Despite initial target bis-glucoside **11** not being obtained its unusual pattern of bis-glucosylation still holds great promise for the incorporation into a variety of molecular tools. The major challenge faced during its synthesis was the bis-glucosylation of *N*-acyl cysteine **12** with **3**. This suggests a new step-wise synthetic route, which makes use of the methodology described earlier, as outlined in Scheme 33.[78] The key step in this alternative route is the stereoselective glucosylation of the carboxylic acid group in **76** to afford the α-acylglucoside **77**. This will need to be carefully investigated, as the formation of an α-acylglucoside in the presence neighbouring group participation is one of the most challenging reactions in glycochemistry. Nonetheless, if difficulties are encountered during this step and the β-acylglucoside is obtained in tandem with **77**, chromatographic separation of the two diastereoisomers should be possible.



Scheme 33: Reagents and conditions; i) TFA, DCM, 0 °C for 1 h then r.t. for 2 h; ii) **45**, NEt_3, DCM, 0 °C for 1 h then r.t. for 35 h; iii) **28** (α-anomer), DMAP (cat.), EDC.HCl, DCM, r.t.; iv) NaOMe (0.1 M), MeOH, r.t., 1 h; v) TFA, DCM, 0 °C for 1 h then r.t. for 2 h.

If successful this would be the starting point for the generation of a library of potential bis-glycosyl cysteine probes, as was demonstrated earlier for the mono-glycosyl cysteine derivatives. Once this has been achieved, it would then be possible to extend this work to the development of novel glycoarrays in which a targeted CBP could be probed for.

In addition to the above, this methodology could be easily adapted to produce *S*-glycosylceramides such as **79**, shown in Scheme 34.[79,80] These *O*-glycosylceramide analogues are of particular interest, as they aid in the generation and biological evaluation of new glycolipid antigens, and have been found to be more stable during *in vivo* studies.[79] More specifically, ceramide **79** is an analogue of α-GalCer **8** and could be used to study the anti-tumour effects of **8**.



Scheme 34: An envisioned approach towards the *S*-galactosylceramide **79**, an analogue of α-GalCer **8**. Its synthesis would involve use of the methodology established in the synthesis of SGC **67** and SAGC **74**. The incorporation of the tetradecyl ester (red) and the anomerization of the β-thiogalactoside linkage to an α-thiogalactoside linkage (blue) are the only two key adaptations to the established methodology to be implemented.

Chapter Four

4. Experimental

4.1. General methods

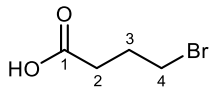
All commercial reagents were purchased from Sigma-Aldrich and used as received. All solvents were purchased either from Sigma-Aldrich or Kimix Chemical And Lab Supply Cc. Dry or freshly distilled solvents were used when reactions were carried out under an anhydrous, inert, atmosphere under a continuous flow of argon gas. Dichloromethane was dried over calcium hydride, and distilled under argon gas.

All reactions were monitored by TLC carried out on pre-coated Merck 60 F_{254} silica gel alumina plates using an ascending technique. The plates were visualized with UV-light (254 nm) or by spraying with solutions of either ceric ammonium sulfate (5.00 g CAS, 50 ml ethanol, 50 mL 2.0 M H_2SO_4), ninhydrin (200 mg ninhydrin, 97 mL ethanol, 3 mL glacial acetic acid) or bromocresol green (100 mg bromocresol green, 500 ml ethanol, 5 ml 0.1 M NaOH). Gravity column chromatography was performed using Merck silica gel 60 (70-230 mesh). All compounds were dried *in vacuo* before yields were determined.

Melting points were determined using a Reichert-Jung Thermovar hot-stage microscope and are uncorrected. NMR spectra (^{1}H, ^{13}C and ^{19}F NMR) were recorded on either a Bruker (400/101/377 MHz) or Varian Unity (300 MHz) instrument at 30 °C and are referenced with respect to the deuterated solvent peak, unless stated otherwise. All chemical shifts (δ) are recorded in ppm with respect to TMS and coupling constants (*J*-values) reported in Hertz (Hz); with signals assigned according to the numbering designated, (*H*-1 = *H* on *C*-1). The numbering scheme given for each compound is for assignment purposes only and not necessarily consistent with the IUPAC naming convention. IR spectra were recorded on a Perkin-Elmer UATR Two spectrophotometer. HRMS spectra were recorded on a Waters Synapt G2 machine in ESI^{+} mode, which was performed at the Central Analytical Facilities (CAF), School of Chemistry, Stellenbosch University.

4.2. Experimental methods

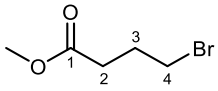
4-bromobutanoic acid (16)[13,31]



To an aqueous solution of HBr (48 % in H_2O, 29 mL) and conc. H_2SO_4 (7 mL) at 0 °C, was added γ-butyrolactone **15** (4.00 mL, 52.0 mmol) and the stirred reaction mixture allowed to warm to room temperature over 1 h. The reaction mixture was then heated at reflux for 14 h, cooled to room temperature and diluted with water (135 mL). The resulting solution was extracted with diethyl ether (4 x 40 mL), and the combined organic extracts washed with sat. brine (40 mL), dried over Na_2SO_4 and filtered. The filtrate was concentrated under reduced pressure and the resulting residue dried *in vacuo* (at 1.0 mbar) to afford **16** as a yellow solid (6.25 g, 37.4 mmol, 72 %).

M.p. 29-30 °C, (lit.[31] 32 °C). ^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 3.48 (2H, t, *J* = 6.4 Hz, *H*-4), 2.57 (2H, t, *J* = 7.1 Hz, *H*-2), 2.18 (2H, p, *J* = 6.7 Hz, *H*-3). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) 179.0 (*C*-1), 32.4 (*C*-2, *C*-4), 27.6 (*C*-3).[13]

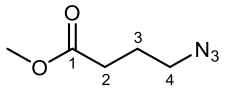
methyl 4-bromobutanoate (26)[32,81]



To a solution of **16** (2.00 g, 12.0 mmol) in methanol (12 mL) at room temperature, was added conc. H_2SO_4 (4 drops) and the reaction mixture heated at reflux for 3 h. The reaction mixture was then cooled to room temperature, diluted with dichloromethane (20 mL) and washed sequentially with sat. $NaHCO_3$ (20 mL), water (20 mL) and sat. brine (20 mL). The extracted organic layer was dried over Na_2SO_4, filtered and concentrated under reduced pressure to afford **26** as a colourless, viscous liquid (1.82 g, 10.1 mmol, 84 %).

^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 3.67 (3H, s, OC*H*$_3$) 3.45 (2H, t, *J* = 6.5 Hz, *H*-4), 2.49 (2H, t, *J* = 7.2 Hz, *H*-2), 2.16 (2H, p, *J* = 6.9 Hz, *H*-3). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) 173.0 (*C*-1), 51.8 (O*C*H$_3$), 32.7, 32.4, 27.9 (*C*-3).[81]

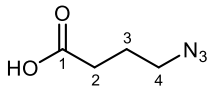
methyl 4-azidobutanoate (27)[33,35,36]



To a solution of **26** (1.74 g, 9.61 mmol) in dimethyl sulfoxide (12 mL) at room temperature, was added sodium azide (0.938 g, 14.4 mmol) and the reaction mixture heated at 50 ^{o}C for 4 h. The reaction mixture was then cooled to room temperature, diluted with diethyl ether (50 mL) and the resulting precipitate recovered by filtration. The filtrate was washed sequentially with sat. $NaHCO_3$ (3 x 20 mL), water (3 x 20 mL) and sat. brine (3 x 20 mL). The extracted organic layer was dried over $MgSO_4$ for 24 h, filtered and concentrated under reduced pressure to afford **27** as a off-yellow, viscous liquid (0.662 g, 4.62 mmol, 48 %).

IR: ν_{max} (ATR, cm^{-1}) 2095 (N=N=N), 1733 (C=O). ^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 3.67 (3H, s, OC*H*$_3$) 3.33 (2H, t, *J* = 6.7 Hz, *H*-4), 2.40 (2H, t, *J* = 7.2 Hz, *H*-2), 1.89 (2H, p, *J* = 6.9 Hz, *H*-3). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) 173.2 (*C*-1), 51.8 (O*C*H$_3$), 50.7 (*C*-4), 31.0 (*C*-2), 24.3 (*C*-3).[35]

4-azidobutanoic acid (17)[34-36]



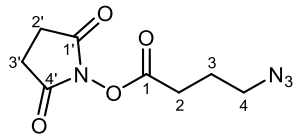
A suspension of **27** (0.604 g, 4.22 mmol) in aq. NaOH (1.0 M, 5 mL) at room temperature, was treated with a small amount of methanol (0.15 mL) and stirred vigorously to obtain a homogeneous solution after 1 h. After a further 3 h the solution was diluted with diethyl ether (20 mL) and the extracted aqueous layer acidified (to pH = 1) by dropwise addition of conc. HCl before extracting further with diethyl ether (4 x 20 mL). The combined organic extracts were dried over Na_2SO_4, filtered and concentrated under reduced pressure to afford **17** as an off-yellow, viscous liquid (0.355 g, 2.75 mmol, 65 %).

^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 9.86 (1H, bs, O*H*), 3.37 (2H, t, *J* = 6.7 Hz, *H*-4), 2.47 (2H, t, *J* = 7.2 Hz, *H*-2), 1.91 (2H, p, *J* = 6.9 Hz, *H*-3). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) 179.1 (*C*-1), 50.6 (*C*-4), 31.0 (*C*-2), 24.1 (*C*-3).[34-36]

General procedure used for the conversion of carboxylic acids to succinimidyl esters

Procedure (A): To a suspension consisting of the carboxylic acid substrate (1.0 eq.) and NHS (1.1 eq.) in dichloromethane (20 mL) at 0 ^{o}C, was added EDC.HCl (1.2 eq.) and the stirred reaction mixture allowed to warm to room temperature overnight. The reaction mixture was then diluted with dichloromethane (20 mL) and washed sequentially with aq. HCl (0.2 M, 3 × 20 mL) and sat. brine (30 mL). The extracted organic layer was dried over Na_2SO_4, filtered and concentrated under reduced pressure to afford the desired succinimidyl ester product.[14,36]

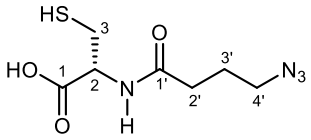
2,5-dioxopyrrolidin-1-yl 4-azidobutanoate (18)[36]



Esterification of **17** (0.270 g, 2.09 mmol) was achieved using *Procedure A* and afforded **18** as a colourless oil (0.391 g, 1.72 mmol, 82 %).

^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 3.44 (2H, t, *J* = 6.6 Hz, *H*-4), 2.82 (4H, s, *H*-2', *H*-3'), 2.72 (2H, t, *J* = 7.2 Hz, *H*-2), 2.01 (2H, p, *J* = 7.0 Hz, *H*-3). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) 169.1 (*C*-1', *C*-4'), 168.0 (*C*-1), 50.1 (*C*-4), 28.3 (*C*-2), 25.7 (*C*-2', *C*-3'), 24.3 (*C*-3).[36]

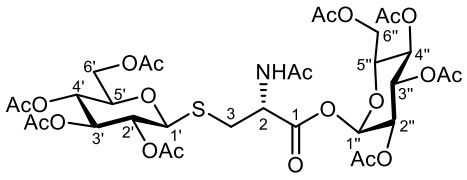
***N*-(4-azidobutanoyl)-L-cysteine (12)**



To a suspension of L-cysteine hydrochloride **14** (0.132 g, 0.838 mmol) and triethylamine (0.12 mL) in dry dimethylformamide (3 mL) at 0 ^{o}C, was added **18** (0.183 g, 0.806 mmol) and the stirred reaction mixture allowed to warm to room temperature over 1 h. After a further 13 h the reaction mixture was then quenched with cold water (40 mL) and diluted with ethyl acetate (30 mL). The separated aqueous layer was acidified (to pH = 2) by dropwise addition of conc. HCl, before extracting further with cold ethyl acetate (3 x 20 mL) and washing the combined organic layers repeatedly with aq. LiCl (10 % w/v, 5 x 10 mL). The combined organic extracts were dried over Na_2SO_4, filtered and solvents removed by azeotropic distillation after addition of toluene, to afford **12** as a crude, off-yellow oil (0.084 g, 0.362 mmol, 45 %).

^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 6.67 (1H, d, *J* = 7.5 Hz, N*H*COR), 4.89 (1H, dt, *J* = 7.5, 4.3 Hz, *H*-2), 3.39 (2H, t, *J* = 6.5 Hz, *H*-4'), 3.18-2.94 (2H, m, *H*-3), 2.42 (2H, t, *J* = 7.2 Hz, *H*-2'), 1.95 (2H, p, *J* = 6.8 Hz, *H*-3'), 1.49 (1H, t, *J* = 9.0 Hz, S*H*). Further analysis was not conducted on the basis of the compound being impure, and hence the above data provides tentative confirmation for **12**.

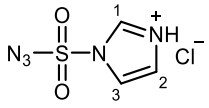
_N_-acetyl-_S_-(2,3,4,6-tetra-_O_-acetyl-β-D-glucopyranosyl)-_O_-(2,3,4,6-tetra-_O_-acetyl-α-D-glucopyranosyl)-L-cysteinate (1)[7]



To a mixture of penta-acetyl glucopyranose **3** (0.504 g, 1.29 mmol) and *N*-acetyl-L-cysteine **4** (0.098 g, 0.601 mmol) in dry dichloromethane (5 mL) at room temperature, was added $SnCl_4$ (1.0 M in DCM, 1.50 mL, 1.50 mmol) dropwise and the reaction mixture left to stir for 3 h. The reaction mixture was then diluted with dichloromethane (15 mL) and washed sequentially with aq. HCl (1.0 M, 2 × 10 mL). The extracted organic layer was dried over Na_2SO_4, filtered and concentrated under reduced pressure to give the crude product. This was then purified *via* gradient column chromatography (1-5 % methanol in dichloromethane) and the pooled fractions concentrated under reduced pressure to afford **1** as a colourless oil (0.058 g, 0.072 mmol, 12 %).

^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 6.53 (1H, d, *J* = 7.3 Hz, N*H*COCH$_3$), 6.33 (1H, d, *J* = 3.6 Hz, *H*-1''), 5.46 (1H, t, *J* = 9.8 Hz), 5.24 (1H, t, *J* = 9.6 Hz), 5.04-5.17 (3H, m), 5.00 (1H, t, *J* = 9.5 Hz), 4.90-4.84 (1H, m, *H*-2), 4.56 (1H, d, *J* = 10.1 Hz, *H*-1'), 4.25-4.05 (5H, m), 3.75 (1H, m), 3.23 (1H, dd, *J* = 14.4, 4.7 Hz, *H*-3), 3.16 (1H, dd, *J* = 14.4, 6.2 Hz, *H*-3), 2.04-1.95 (27H, m, 8 × COC*H*$_3$, NHCOC*H*$_3$). Further analysis was not conducted on the basis of the compound being impure; however the above experimental data is in good agreement with the reported literature data.[7]

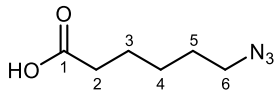
1*H*-imidazole-1-sulfonyl azide hydrochloride (33)[42]



To a suspension of sodium azide (4.00 g, 61.5 mmol) in acetonitrile (60 mL) at 0 °C, was added sulfuryl chloride **37** (4.95 mL, 61.3 mmol) and the stirred reaction mixture allowed to warm to room temperature overnight. The reaction mixture was then cooled to 0 °C again before adding imidazole (7.96 g, 117 mmol) and then allowing the reaction mixture to warm to room temperature over 30 min. After a further 3 h the reaction mixture was diluted with ethyl acetate (60 mL) and washed sequentially with water (2 × 150 mL) and sat. $NaHCO_3$ (2 × 150 mL). The extracted organic layer was dried over Na_2SO_4, filtered and cooled to 0 °C. A solution of HCl in ethanol [generated *in situ* from the dropwise addition of excess acetyl chloride (6 mL) to ethanol (24 mL) at 0 °C] was then added dropwise to the filtrate together with vigorous stirring. The resulting precipitate was recovered by filtration and washed with small amounts of cold ethyl acetate to afford **33** as a white solid (9.23 g, 44.0 mmol, 72 %).

M.p. 100-101 °C, (lit.[42] 100-102 °C). IR: ν_{max} (ATR, cm^{-1}) 2169 (N=N=N), 1425 (S=O), 1160 (S=O). ^{1}H NMR: δ_H (D_2O, 300 MHz) 9.38 (1H, dd, *J* = 1.6, 1.2 Hz, *H*-1), 8.08 (1H, dd, *J* = 2.1, 1.6 Hz, *H*-3), 7.67 (1H, dd, *J* = 2.1, 1.2 Hz, *H*-2). ^{13}C NMR: δ_C (D_2O, 101 MHz) 137.8 (*C*-1), 123.8, 120.1.[42]

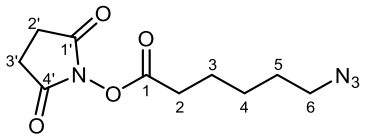
6-azidohexanoic acid (34)[43]



To a suspension of 6-aminohexanoic acid **32** (2.00 g, 15.2 mmol), K_2CO_3 (4.64 g, 33.6 mmol) and $CuSO_4.5H_2O$ (2.7 mol %) in methanol (40 mL) at room temperature, was added the sulfonyl azide **33** (3.84 g, 18.3 mmol) and the reaction mixture left to stir for 16 h. The reaction mixture was then concentrated under reduced pressure and the resulting residue re-dissolved in water (30 mL). The aqueous solution was washed repeatedly with dichloromethane (6 x 30 mL) and the extracted aqueous layer acidified (to pH = 2) by dropwise addition of aq. HCl (1.0 M) before extracting further with diethyl ether (6 x 30 mL). The combined organic extracts were dried over Na_2SO_4, filtered and concentrated under reduced pressure to give the crude product. This was then purified *via* gradient column chromatography (20-80 % ethyl acetate in hexane) and the pooled fractions concentrated under reduced pressure to afford **34** as an off-yellow oil (0.832 g, 5.29 mmol, 35 %).

IR: ν_{max} (ATR, cm^{-1}) 3120 (O-H), 2090 (N=N=N), 1704 (C=O). ^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 9.11 (1H, bs, O*H*), 3.28 (2H, t, *J* = 6.8 Hz, *H*-6), 2.38 (2H, t, *J* = 7.4 Hz, *H*-2), 1.73-1.58 (4H, m, *H*-3, *H*-5), 1.49-1.38 (2H, m, *H*-4). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) 179.9 (*C*-1), 51.3 (*C*-6), 34.0 (*C*-2), 28.7 (*C*-5), 26.3 (*C*-4), 24.3 (*C*-3).[43]

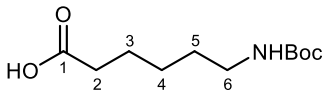
2,5-dioxopyrrolidin-1-yl 6-azidohexanoate (35)[14]



Esterification of **34** (0.546 g, 3.47 mmol) was achieved using *Procedure A* and afforded **35** as an off-yellow oil (0.661 g, 2.60 mmol, 75 %).

^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 3.29 (2H, t, *J* = 6.7 Hz, *H*-6), 2.82 (4H, s, *H*-2', *H*-3') 2.62 (2H, t, *J* = 7.3 Hz, *H*-2), 1.78 (2H, p, *J* = 7.4 Hz, *H*-3), 1.68-1.59 (2H, m, *H*-5), 1.55-1.45 (2H, m, *H*-4). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) 169.2 (*C*-1', *C*-4'), 168.5 (*C*-1), 51.3 (*C*-6), 30.9 (*C*-2), 28.5 (*C*-5), 26.0 (*C*-4), 25.7 (*C*-2', *C*-3'), 24.3 (*C*-3).[14]

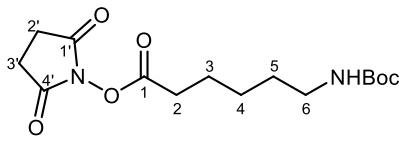
6-((*tert*-butyloxycarbonyl)amino)hexanoic acid (46)[55,56,82]



To a mixture of 6-aminohexanoic acid **32** (2.00 g, 15.2 mmol) and Boc-anhydride (3.32 g, 15.2 mmol) in methanol (40 mL) at room temperature, was added triethylamine (2.60 mL, 18.7 mmol) and the reaction mixture heated at reflux for 24 h. The reaction mixture was then cooled to room temperature, concentrated under reduced pressure and the resulting residue re-dissolved in ethyl acetate (40 mL). The organic solution was washed with aq. HCl (0.2 M, 2 × 50 mL), dried over Na_2SO_4 and filtered. The filtrate was then concentrated under reduced pressure to give a colourless oil which solidified upon storage at -18 ^{o}C. The resulting crude solid was dried *in vacuo* and then purified *via* recrystallization from hexane to afford **46** as a white solid (3.09 g, 13.4 mmol, 88 %).

M.p. 37-39 °C, (lit.[82] 38-39 °C). ^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 4.57 (1H, bs, N*H*COOR), 3.10 (2H, q, *J* = 6.3 Hz, *H*-6), 2.34 (2H, t, *J* = 7.4 Hz, *H*-2), 1.64 (2H, p, *J* = 7.4 Hz, *H*-3), 1.53-1.32 (4H, m, *H*-4, *H*-5), 1.43 (9H, s, 3 × C(C*H*$_3$)). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) 179.0 (*C*-1), 156.2 (NH*C*OOR), 79.3 (*C*(CH_3)$_3$), 40.5 (*C*-6), 34.0 (*C*-2), 29.8 (*C*-5), 28.6 (3 × C(*C*H_3)), 26.4 (*C*-4), 24.5 (*C*-3).[56,82]

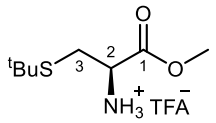
2,5-dioxopyrrolidin-1-yl 6-((*tert*-butoxycarbonyl)amino)hexanoate (45)[56,83]



Esterification of **46** (3.00 g, 13.0 mmol) was achieved using *Procedure A* and afforded **45** as a white solid (3.68 g, 11.2 mmol, 86 %).

M.p. 88-91 °C, (lit.[83] 87-89 °C). ^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 4.59 (1H, bs, N*H*COOR), 3.11 (2H, q, *J* = 6.3 Hz, *H*-6), 2.81 (4H, s, *H*-2', *H*-3'), 2.59 (2H, t, *J* = 7.3 Hz, *H*-2), 1.76 (2H, p, *J* = 7.4 Hz, *H*-3), 1.56-1.37 (4H, m, *H*-4, *H*-5), 1.43 (9H, s, 3 × C(C*H*$_3$)). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) 169.2 (*C*-1', *C*-4'), 168.6 (*C*-1), 156.1 (NH*C*OOR), 79.2 (*C*(CH_3)$_3$), 40.4 (*C*-6), 31.0 (*C*-2), 29.7 (*C*-5), 28.5 (3 × C(*C*H_3)), 26.0 (*C*-4), 25.7 (*C*-2', *C*-3'), 24.4 (*C*-3).[56,83] Note that compound **45** was prepared several times during the project and batches were combined.

methyl *S*-(*tert*-butyl)-L-cysteinate trifluoroacetate (41)[84]



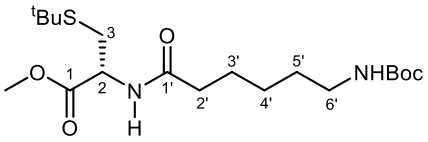
To a solution of methyl L-cysteinate hydrochloride **36** (1.50 g, 8.74 mmol) in TFA (10 mL) at room temperature, was added *tert*-butanol (0.85 mL, 8.89 mmol) and the reaction mixture left to stir for 18 h. Excess TFA was then evaporated off under a steady-stream of air and the resulting solid-like residue suspended in a minimal amount of cold water. The residue was quickly recovered by filtration and washed with small amounts of cold water to afford **41** as a white wax-like solid (1.68 g, 5.50 mmol, 63 %).

^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 7.61 (2H, bs, N*H*$_2$), 4.22 (1H, dd, *J* = 6.5, 5.4 Hz, *H*-2), 3.82 (3H, s, OC*H*$_3$), 3.18 (1H, dd, *J* = 13.7, 5.4 Hz, *H*-3), 3.11 (1H, dd, *J* = 13.7, 6.5 Hz, *H*-3), 1.33 (9H, s, 3 × C(C*H*$_3$)). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) 168.5 (*C*-1), 53.6, 53.1, 44.0 (*C*$(CH_3)_3$), 30.8 (3 × C(*C*H_3)), 28.6 (*C*-3). ^{19}F NMR: δ_C ($CDCl_3$, 377 MHz) -75.7 (3F, s, COC*F*$_3$).

General procedure used for the conversion of amines to amides

Procedure (B): To a solution of the *S*-protected/glycosyl cysteine substrate (1.1-1.3 eq.) in dichloromethane (10 mL) at 0 ^{o}C, was added triethylamine (1.4 eq.) and the reaction mixture left to stir for 30 min, at which point the succinimidyl ester **45** (1.0 eq.) was added. The stirred reaction mixture was then allowed to warm to room temperature over 30 min and after a further 23-35 h the reaction mixture was diluted with dichloromethane (20 mL) and washed sequentially with aq. HCl (0.2 M, 5 × 20 mL) and sat. brine (30 mL). The extracted organic layer was dried over Na_2SO_4, filtered and concentrated under reduced pressure to give the crude product. This was then purified *via* column chromatography and the pooled fractions concentrated under reduced pressure to afford the desired amide product.[47]

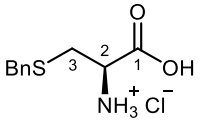
methyl *N*-(6-((*tert*-butoxycarbonyl)amino)hexanoyl)-*S*-(*tert*-butyl)-L-cysteinate (42)



Amide coupling of **41** (0.610 g, 2.00 mmol) with the succinimidyl ester **45** (0.496 g, 1.51 mmol) was achieved using *Procedure B* and afforded **42** as a colourless oil (0.393 g, 0.971 mmol, 64 %) after purification *via* gradient column chromatography (50-100 % ethyl acetate in hexane).

^{1}H NMR: δ_H ($CDCl_3$, 400 MHz) 6.23 (1H, d, *J* = 7.5 Hz, N*H*COR), 4.85 (1H, dt, *J* = 7.5, 4.9 Hz, *H*-2), 4.57 (1H, bs, N*H*COOR), 3.75 (3H, s, OC*H*$_3$), 3.09 (2H, q, *J* = 6.4 Hz, *H*-6'), 3.03-2.95 (2H, m, *H*-3), 2.23 (2H, t, *J* = 7.5 Hz, *H*-2'), 1.65 (2H, p, *J* = 7.4 Hz, *H*-3'), 1.52-1.32 (4H, m, *H*-4', *H*-5'), 1.43 (9H, s, 3 × OC(C*H*$_3$)), 1.29 (9H, s, 3 × SC(C*H*$_3$)). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) 172.7, 171.4, 156.1 (NH*C*OOR), 79.2 (O*C*$(CH_3)_3$), 52.7 (O*C*H_3), 51.9 (*C*-2), 42.8 (S*C*$(CH_3)_3$), 40.5 (*C*-6'), 36.4 (*C*-2'), 31.0 (3 × SC(*C*H_3)), 30.6 (*C*-3), 29.9 (C-5'), 28.6 (3 × OC(*C*H_3)), 26.4 (*C*-4'), 25.2 (*C*-3'). HRMS (ESI$^+$): *m/z* calculated for $C_{19}H_{36}N_2O_5S$, 404.2345; found 405.2419 $[M + H]^+$.

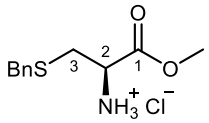
***S*-benzyl-L-cysteine hydrochloride (48)**[51]



To a rapidly stirring solution of aq. NaOH (2.0 M, 11 mL) and ethanol (25 mL) at room temperature, was added L-cysteine hydrochloride **14** (0.904 g, 5.74 mmol) followed by dropwise addition of benzyl bromide (0.70 mL, 5.89 mmol). After 1 h, the reaction mixture was acidified (to pH = 2) by addition of conc. HCl and the resulting precipitate recovered by filtration. The precipitate was then washed sequentially with small amounts of cold water, ethanol and dichloromethane to afford **48** as a white crystalline solid (1.01 g, 4.08 mmol, 71 %).

M.p. 201-204 °C, (lit.[51] 210-211 °C). ^{1}H NMR: δ_H (D_2O, 300 MHz) 7.29-7.09 (5H, m, 5 × Ar-*H*), 3.95 (1H, dd, *J* = 7.6, 4.6 Hz, *H*-2), 3.64 (2H, s, SC*H*$_2$Ar), 2.87 (1H, dd, *J* = 15.0, 4.6 Hz, *H*-3), 2.76 (1H, dd, *J* = 15.0, 7.6 Hz, *H*-3). ^{13}C NMR: δ_C (D_2O, 101 MHz) 170.2 (*C*-1), 137.5 (Ar-*C*), 129.0 (2 × Ar-*C*), 128.9 (2 × Ar-*C*), 127.6 (Ar-*C*), 52.0 (*C*-2), 35.4 (S*C*H$_2$Ar), 30.5 (*C*-3).[51]

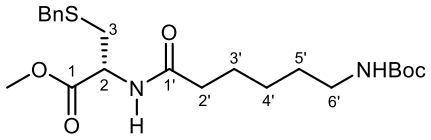
methyl *S*-benzyl-L-cysteinate hydrochloride (43)[51]



To a solution of HCl in methanol [generated *in situ* from the dropwise addition of excess acetyl chloride (1 mL) to methanol (3 mL) at 0 °C], was added **48** (0.769 g, 3.10 mmol) and the stirred reaction mixture allowed to warm to room temperature over 1 h. After a further 14 h the reaction mixture was poured into an evaporating dish and excess volatiles allowed to evaporate overnight. The resulting solid was recovered by filtration and washed with small amounts of cold diethyl ether to afford **43** as white crystalline solid (0.544 g, 2.08 mmol, 67 %).

M.p. 142-144 °C, (lit.[51] 148-150 °C). ^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 8.90 (2H, bs, N*H*$_2$), 7.37-7.18 (5H, m, 5 × Ar-*H*), 4.49-4.40 (1H, m, *H*-2), 3.83 (2H, s, SC*H*$_2$Ar), 3.74 (3H, s, OC*H*$_3$), 3.18 (2H, d, *J* = 5.6 Hz, *H*-3). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) 168.7 (*C*-1), 137.5 (Ar-*C*), 129.3 (2 × Ar-*C*), 128.8 (2 × Ar-*C*), 127.4 (Ar-*C*), 53.6, 52.9, 36.8 (S*C*H$_2$Ar), 31.5 (*C*-3).[51]

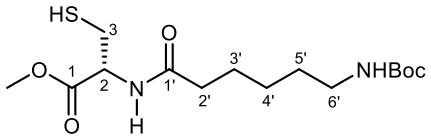
methyl *S*-benzyl-*N*-(6-((*tert*-butoxycarbonyl)amino)hexanoyl)-L-cysteinate (44)



Amide coupling of **43** (0.273 g, 1.04 mmol) with the succinimidyl ester **45** (0.251 g, 0.764 mmol) was achieved using *Procedure B* and afforded **44** as a white solid (0.240 g, 0.547 mmol, 72 %) after purification *via* gradient column chromatography (50-100 % ethyl acetate in hexane).

M.p. 83-84 °C. ^{1}H NMR: δ_H ($CDCl_3$, 400 MHz) 7.34-7.22 (5H, m, 5 × Ar-*H*), 6.12 (1H, d, *J* = 7.8 Hz, N*H*COR), 4.80 (1H, dt, *J* = 7.8, 5.4 Hz, *H*-2), 4.54 (1H, bs, N*H*COOR), 3.73 (3H, s, OC*H*$_3$), 3.70 (2H, s, SC*H*$_2$Ar), 3.10 (2H, q, *J* = 6.5 Hz, *H*-6'), 2.92 (1H, dd, *J* = 13.8, 5.0 Hz, *H*-3), 2.85 (1H, dd, *J* = 13.8, 5.7 Hz, *H*-3), 2.19 (2H, t, *J* = 7.5 Hz, *H*-2'), 1.65 (2H, p, *J* = 7.4 Hz, *H*-3'), 1.48 (2H, p, *J* = 7.4 Hz, *H*-5'), 1.43 (9H, s, 3 × C(C*H*$_3$)), 1.40-1.31 (2H, m, *H*-4'). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) 172.6, 171.4, 156.0 (NH*C*OOR), 137.7 (Ar-*C*), 128.9 (2 × Ar-*C*), 128.6 (2 × Ar-*C*), 127.3 (Ar-*C*), 79.1 (*C*(CH_3)$_3$), 52.5 (O*C*H$_3$), 51.5 (*C*-2), 40.4 (*C*-6'), 36.7 (S*C*H$_2$Ar), 36.2 (*C*-2'), 33.6 (*C*-3), 29.7 (*C*-5'), 28.4 (3 × C(*C*H$_3$)), 26.3 (*C*-4'), 25.0 (*C*-3'). HRMS (ESI$^+$): *m/z* calculated for $C_{22}H_{34}N_2O_5S$, 438.2188; found 439.2278 $[M + H]^+$.

methyl *N*-(6-((*tert*-butoxycarbonyl)amino)hexanoyl)-L-cysteinate (40)



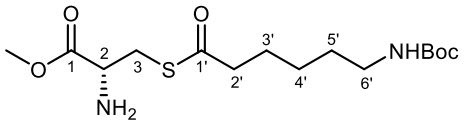
To a solution of methyl L-cysteine hydrochloride **36** (2.87 g, 16.7 mmol) in a H_2O:MeCN (1:1, 50 mL) mixture at 0 °C, was added triethylamine until a basic solution (pH = 9) was obtained and the reaction mixture left to stir for 30 min, at which point the succinimidyl ester **45** (5.00 g, 15.2 mmol) was added. The stirred reaction mixture was then allowed to warm to room temperature over 30 min and after a further 17 h the reaction mixture was diluted with ethyl acetate (50 mL) and washed sequentially with aq. HCl (0.2 M, 5 × 40 mL) and sat. brine (40 mL). The extracted organic layer was dried over Na_2SO_4, filtered and concentrated under reduced pressure to afford the crude product. This was then purified twice *via* flash column chromatography (100 % ethyl acetate) and the pooled fractions concentrated under pressure to afford **40** as a white a solid (4.03 g, 11.6 mmol, 76 %).

M.p. 53-56 °C. ^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 6.33 (1H, d, *J* = 7.4 Hz, N*H*COR), 4.88 (1H, dt, *J* = 7.4, 4.2 Hz, *H*-2), 4.54 (1H, bs, N*H*COOR), 3.79 (3H, s, OC*H*$_3$), 3.11 (2H, q, *J* = 6.4 Hz, *H*-6'), 3.01 (2H, ddd, *J* = 8.9, 4.3, 4.0 Hz, *H*-3), 2.27 (2H, t, *J* = 7.5 Hz, *H*-2'), 1.67 (2H, p, *J* = 7.4 Hz, *H*-3'), 1.55-1.32 (4H, m, *H*-4', *H*-5'), 1.43 (9H, s, 3 × C(C*H*$_3$)), 1.33 (1H, t, *J* = 8.9 Hz, S*H*). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) 172.8, 170.8, 156.1 (NH*C*OOR), 79.1 (*C*(CH_3)$_3$), 53.5 (*C*-2), 52.8 (O*C*H$_3$), 40.4 (*C*-6'), 36.3 (*C*-2'), 29.8 (*C*-5'), 28.5 (3 × C(*C*H$_3$)), 26.9 (*C*-3), 26.4 (*C*-4'), 25.2 (*C*-3'). HRMS (ESI$^+$): *m/z* calculated for $C_{15}H_{28}N_2O_5S$, 348.1719; found 349.1798 $[M + H]^+$.

.

methyl *S*-(6-((*tert*-butoxycarbonyl)amino)hexanoyl)-L-cysteinate (49)

Compound **49** was isolated as a white solid (0.883 g, 2.53 mmol, 17 %) being the minor product in the same reaction and isolation procedure that lead to the formation of **40**.

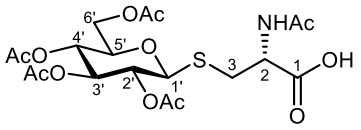


M.p. 110-114 °C. ^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 6.58 (2H, bs, N*H*$_2$), 4.82 (1H, dt, *J* = 7.4, 5.4 Hz, *H*-2), 4.67 (1H, bs, N*H*COOR), 3.75 (3H, s, OC*H*$_3$), 3.16 (2H, d, *J* = 5.4 Hz, *H*-3), 3.11 (2H, q, *J* = 6.5 Hz, *H*-6'), 2.24 (2H, t, *J* = 7.4 Hz, *H*-2'), 1.64 (2H, p, *J* = 7.4 Hz, *H*-3'), 1.52-1.30 (4H, m, *H*-4', *H*-5'), 1.41 (9H, s, 3 × C(C*H*$_3$)). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) 173.0, 171.1, 156.2 (NH*C*OOR), 79.1 (*C*(CH_3)$_3$), 52.8 (O*C*H$_3$), 51.8 (*C*-2), 40.8 (*C*-3), 40.5 (*C*-6') 36.2 (*C*-2'), 29.8 (*C*-5'), 28.5 (3 × C(*C*H$_3$)), 26.4 (*C*-4'), 25.1 (*C*-3'). HRMS (ESI$^+$): *m/z* calculated for $C_{15}H_{28}N_2O_5S$, 348.1719; found 349.1783 $[M + H]^+$.

General procedure used for the formation of thioglucosides from glycosyl acetates under Lewis acid promotion

Procedure (C): To a mixture of penta-acetyl glucopyranose **3** (1.0 eq.) and the cysteine substrate (1.5 eq.) in dry dichloromethane (10 mL) at room temperature, was added $SnCl_4$ (2.0 eq.) dropwise and the reaction mixture left to stir for 3 h. The reaction mixture was then diluted with dichloromethane (20 mL) and washed sequentially with aq. HCl (1.0 M, 2 × 20 mL) and sat. brine (30 mL). The extracted organic layer was dried over Na_2SO_4, filtered and concentrated under reduced pressure to give the crude product. This was then purified *via* column chromatography and the pooled fractions concentrated under reduced pressure to afford the desired *S*-glycosylcysteine product.[20]

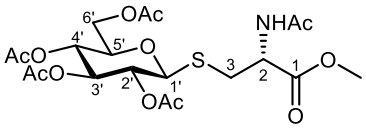
_N_-acetyl-_S_-(2,3,4,6-tetra-_O_-acetyl-β-D-glucopyranosyl)-L-cysteine (2)[20]



Glycosylation of *N*-acetyl-L-cysteine **4** (0.312 g, 1.91 mmol) with penta-acetyl glucopyranose **3** (0.506 g, 1.30 mmol) was achieved using *Procedure C* and afforded **2** as a white foam (0.215 g, 0.436 mmol, 34 %) after purification *via* gradient column chromatography (1-10 % methanol in dichloromethane).

^{1}H NMR: δ_H (CD_3OD, 400 MHz) 5.28 (1H, t, *J* = 9.3 Hz, *H*-3'), 5.04 (1H, t, *J* = 9.8 Hz, *H*-4'), 4.93 (1H, t, *J* = 9.7 Hz, *H*-2'), 4.78 (1H, d, *J* = 10.1 Hz, *H*-1'), 4.50 (1H, dd, *J* = 8.2, 4.5 Hz, *H*-2), 4.27 (1H, dd, *J* = 12.5, 4.7 Hz, *H*-6'), 4.18 (1H, dd, *J* = 12.5, 2.3 Hz, *H*-6'), 3.89 (1H, ddd, *J* = 10.1, 4.7, 2.3 Hz, *H*-5'), 3.26 (1H, dd, *J* = 13.8, 4.5 Hz, *H*-3), 2.94 (1H, dd, *J* = 13.8, 8.2 Hz, *H*-3), [2.07, 2.03, 2.01, 1.97 (15H, 4s, 4 × COC*H*$_3$, NHCOC*H*$_3$)]. ^{13}C NMR: δ_C (CD_3OD, 101 MHz) [176.3, 173.1, 172.4, 171.6, 171.3, 171.2 (4 × *C*OCH$_3$, NH*C*OCH$_3$, *C*-1)], 84.1 (*C*-1'), 76.9 (*C*-5'), 75.2 (*C*-3'), 71.6 (*C*-2'), 69.8 (*C*-4'), 63.3 (*C*-6'), 55.1 (*C*-2), 33.0 (*C*-3), 22.7 (NHCO*C*H$_3$), 20.7 (2 × CO*C*H$_3$), 20.5 (2 × CO*C*H$_3$).[20]

methyl *N*-acetyl-*S*-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-L-cysteinate (19)[20,23]



Glycosylation of methyl *N*-acetyl-L-cysteinate **20** (0.338 g, 1.91 mmol) with penta-acetyl glucopyranose **3** (0.500 g, 1.28 mmol) was achieved using *Procedure C* and afforded **19** as a white solid (0.416 g, 0.820 mmol, 64 %) after purification *via* gradient column chromatography (20-80 % ethyl acetate in hexane).

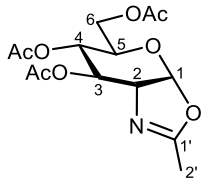
M.p. 100-103 ^{o}C. ^{1}H NMR: δ_H ($CDCl_3$, 400 MHz) 6.41 (1H, d, *J* = 7.5 Hz, N*H*$COCH_3$), 5.21 (1H, t, *J* = 9.4 Hz, *H*-3'), 5.05 (1H, t, *J* = 9.6 Hz, *H*-4'), 4.96 (1H, dd, *J* = 10.1, 9.3 Hz, *H*-2'), 4.81 (1H, dt, *J* = 7.5, 5.4 Hz, *H*-2), 4.53 (1H, d, *J* = 10.1 Hz, *H*-1'), 4.23 (1H, dd, *J* = 12.6, 4.7 Hz, *H*-6'), 4.18 (1H, dd, *J* = 12.6, 2.7 Hz, *H*-6'), 3.75 (3H, s, OC*H*$_3$), 3.71 (1H, ddd, *J* = 10.1, 4.7, 2.7 Hz, *H*-5'), 3.20 (1H, dd, *J* = 14.1, 4.9 Hz, *H*-3), 3.03 (1H, dd, *J* = 14.1, 6.0 Hz, *H*-3), [2.09, 2.05, 2.04, 2.01, 1.99 (15H, 5s, 4 × COC*H*$_3$, NHCOC*H*$_3$)]. ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) [171.1, 170.7, 170.1, 170.0, 169.6, 169.5 (4 × *C*OCH_3, NH*C*OCH_3, *C*-1)], 83.5 (*C*-1'), 76.4 (*C*-5'), 73.8 (*C*-3'), 70.1 (*C*-2'), 68.4 (*C*-4'), 62.1, (*C*-6'), 52.8 (O*C*H_3), 52.0 (*C*-2), 31.9 (*C*-3), 23.0 (NHCO*C*H_3), 20.8 (2 × CO*C*H_3), 20.6 (2 × CO*C*H_3).[20,23]

2-acetamido-1,3,4,6-tetra-*O*-acetyl-2-deoxy-α-D-glucopyranose (55) and 2-acetamido-1,3,4,6-tetra-*O*-acetyl-2-deoxy-β-D-glucopyranose (52)[66,85,86]

A suspension of 2-acetamido-2-deoxy-D-glucopyranose **54** (2.20 g, 9.95 mmol) in pyridine (14 mL) at room temperature, was treated with excess acetic anhydride (9.50 mL) and the reaction mixture left to stir for 25 h. The reaction mixture was then diluted with dichloromethane (150 mL) and washed with aq. HCl (1.0 M, 300 mL). The extracted organic layer was dried over Na_2SO_4, filtered and solvents removed by azeotropic distillation after addition of toluene, to afford a white solid being composed of 2-acetamido-tetra-acetates **55** and **52** in a combined yield of (3.53 g, 9.07 mmol, 91 %) and 9:1 ratio (α:β).

α-anomer (**55**): M.p. 128-132 °C (lit.[86] 128-130 °C). ^{1}H NMR: δ_H ($CDCl_3$, 400 MHz) 6.16 (1H, d, *J* = 3.7 Hz, *H*-1), 5.61 (1H, d, *J* = 9.0 Hz, N*H*COCH$_3$), 5.25-5.16 (2H, m, *H*-3, *H*-4), 4.49-4.43 (1H, m, *H*-2), 4.23 (1H, dd, *J* = 12.4, 4.1 Hz, *H*-6), 4.05 (1H, dd, *J* = 12.4, 2.3 Hz, *H*-6), 3.98 (1H, ddd, *J* = 10.0, 4.1, 2.3 Hz, *H*-5), [2.17, 2.07, 2.03, 2.02 (12H, 4s, 4 × COC*H*$_3$)], 1.91 (3H, s, NHCOC*H*$_3$). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) [171.8, 170.8, 170.1, 169.2, 168.7 (4 × *C*OCH$_3$, NH*C*OCH$_3$)], 90.8 (*C*-1), 70.8 (*C*-3), 69.8 (*C*-5), 67.7 (*C*-4), 61.7 (*C*-6), 51.2 (*C*-2), 23.1 (NHCO*C*H$_3$), [21.0, 20.8, 20.7, 20.6 (4 × CO*C*H$_3$)].[66,85,86]

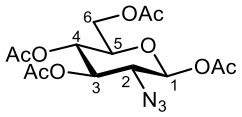
2-methyl-(3,4,6-tri-*O*-acetyl-1,2-dideoxy-α-D-glucopyrano)-[2,1-*d*]-2-oxazoline (53)[66,67]



To a solution of the anomeric mixture composed of **52** and **55** (2.00 g, 5.14 mmol) in dichloromethane (30 mL) at 55 °C, was added TMSOTf (1.15 mL, 6.34 mmol) and the reaction mixture heated at reflux for 17 h. The reaction mixture was then cooled to room temperature, quenched (to pH = 8) by dropwise addition of triethylamine and concentrated under reduced pressure to give the crude product. This was then purified *via* gradient column chromatography (1-5 % methanol in dichloromethane) and the pooled fractions concentrated under reduced pressure to afford **53** as a viscous, amber oil (1.28 g, 3.89 mmol, 76 %).

^{1}H NMR: δ_H ($CDCl_3$, 400 MHz) 5.96 (1H, d, *J* = 7.4 Hz, *H*-1), 5.26 (1H, t, *J* = 2.5 Hz, *H*-3), 4.93 (1H, dq, *J* = 9.3, 1.2 Hz, *H*-4), 4.18 (1H, s, *H*-6), 4.17 (1H, d, *J* = 1.2 Hz, *H*-6), 4.15-4.10 (1H, m, *H*-2), 3.65-3.57 (1H, m, *H*-5), [2.10, 2.09 (6H, 2s, 2 × COCH_3)], 2.08 (3H, d, *J* = 1.9 Hz, *H*-2'), 2.07 (3H, s, COCH_3). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) [170.6, 169.6, 169.2 (3 × *C*OCH_3)], 166.7 (*C*-1'), 99.5 (*C*-1), 70.5 (*C*-4), 68.5 (*C*-3), 67.6 (*C*-5), 65.1 (*C*-2), 63.4 (*C*-6), [20.9, 20.8, 20.7 (3 × CO*C*H_3)], 14.0 (*C*-2').[66,67]

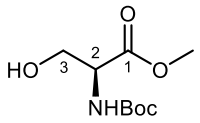
1,3,4,6-tetra-*O*-acetyl-2-azido-2-deoxy-β-D-glucopyranose (57)[69,70]



To a solution of 2-amino-2-deoxy-D-glucopyranose hydrochloride **58** (1.00 g, 4.64 mmol), K_2CO_3 (1.40 g, 10.1 mmol) and $CuSO_4.5H_2O$ (2.7 mol %) in methanol (25 mL) at room temperature, was added the sulfonyl azide **33** (1.16 g, 5.53 mmol) and the reaction mixture left to stir for 3 h. The reaction mixture was then concentrated under reduced pressure and the resulting residue re-dissolved in pyridine (8 mL) at room temperature, and treated with acetic anhydride (4.50 mL, 47.6 mmol). After stirring for 24 h the reaction mixture was quenched with crushed ice, extracted with ethyl acetate (4 x 40 mL) and the combined organic extracts washed sequentially with aq. HCl (1.0 M, 30 mL) and sat. brine (30 mL). The extracted organic layer was dried over Na_2SO_4, filtered and solvents removed by azeotropic distillation after addition of toluene to give the crude product. This was then purified *via* gradient column chromatography (30-80 % ethyl acetate in hexane) and the pooled fractions concentrated under reduced pressure to afford a colourless oil, which solidified upon storage at -18.0 ^{o}C to afford **57** as a white solid (1.04 g, 2.79 mmol, 60 % over two steps).

M.p. 95-96 ^{o}C, (lit.[70] 96-97 ^{o}C). IR: ν_{max} (ATR, cm^{-1}) 2113 (N=N=N), 1761 (C=O), 1747 (C=O), 1737 (C=O), 1733 (C=O). ^{1}H NMR: δ_H ($CDCl_3$, 400 MHz) 5.55 (1H, d, *J* = 8.6 Hz, *H*-1), 5.09 (1H, t, *J* = 9.4 Hz, *H*-3), 5.03 (1H, t, *J* = 9.4, *H*-4), 4.29 (1H, dd, *J* = 12.5, 4.6 Hz, *H*-6), 4.10 (1H, dd, *J* = 12.5, 2.2 Hz, *H*-6), 3.80 (1H, ddd, *J* = 9.7, 4.6, 2.2 Hz, *H*-5), 3.66 (1H, dd, *J* = 9.9, 8.6 Hz, *H*-2), [2.19, 2.09, 2.07, 2.02 (12H, 4s, 4 × COC*H*$_3$)]. ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) [170.6, 169.9, 169.7, 168.7 (4 × *C*OCH$_3$)], 92.8 (*C*-1), 73.0, 72.9, 68.0, 62.8, 61.6 (*C*-6), [21.0, 20.8, 20.7, 20.6 (4 × CO*C*H$_3$)].[69]

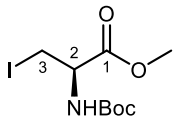
methyl *N*-(*tert*-butoxycarbonyl)-L-serinate (62)[87]



To a mixture of methyl L-serinate hydrochloride **23** (2.00 g, 12.9 mmol) in tetrahydrofuran (15 mL) and aq. $NaHCO_3$ (2.3 M, 15 mL) at 0 ^oC, was added Boc-anhydride (2.81 g, 12.9 mmol) and the reaction mixture allowed to warm to room temperature over 1 h. After a further 6 h the reaction mixture was concentrated under reduced pressure and the remaining aqueous solution acidified (to pH = 3) by dropwise addition of aq. HCl (0.2 M) before extracting further with dichloromethane (3 x 30 mL). The combined organic extracts were washed with sat. brine (50 mL), dried over Na_2SO_4 and filtered. The filtrate was concentrated under reduced pressure to afford **62** as a colourless oil (2.55 g, 11.6 mmol, 90 %).

^1H NMR: δ_H ($CDCl_3$, 300 MHz) 5.51 (1H, d, *J* = 6.5, N*H*COOR), 4.34 (1H, bs, *H*-2), 3.93 (1H, dd, *J* = 11.2, 4.0 Hz, *H*-3), 3.85 (1H, dd, *J* = 11.2, 3.7 Hz, *H*-3), 3.75 (3H, s, OCH_3), 2.67 (1H, s, O*H*), 1.43 (9H, s, 3 × C(CH_3)). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) 171.5 (*C*-1), 155.9 (NH*C*OOR), 80.4 (*C*$(CH_3)_3$), 63.5 (*C*-3), 55.9, 52.7, 28.4 (3 × C(*C*H_3)).[87]

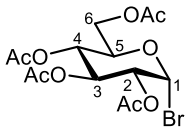
methyl (*R*)-2-((*tert*-butoxycarbonyl)amino)-3-iodopropanoate (60)[74,88]



To a solution of imidazole (1.00 g, 14.7 mmol) and triphenylphosphine (3.86 g, 14.7 mmol) in dichloromethane (50 mL) at 0 °C, was added iodine (3.73 g, 14.7 mmol) over a period of 20 min in the dark. After a further 10 min the stirred reaction mixture was allowed to warm to room temperature over 20 min. The reaction mixture was then cooled to 0 °C again before adding a solution of **62** (2.48 g, 11.3 mmol) in dichloromethane (10 mL) at 0 °C, over a period of 1 h and then allowing the reaction mixture to warm to room temperature over 4 h. The reaction mixture was then diluted with a mixture of Et_2O:Hex (1:1, 300 mL), filtered through a plug of silica and concentrated under reduced pressure to give a solid residue. The residue was re-suspended in diethyl ether (100 mL), filtered through a pad of Celite and concentrated under reduced pressure to give the crude product. This was then purified *via* flash column chromatography (100 % ethyl acetate) and the pooled fractions concentrated under pressure to give a colourless oil, which solidified upon storage at -18.0 °C to afford **60** as a white solid (2.89 g, 8.78 mmol, 78 %).

M.p. 44-45 °C, (lit.[88] 45-47 °C). ^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 5.34 (1H, d, *J* = 6.6 Hz, N*H*COOR), 4.53-4.46 (1H, m, *H*-2), 3.79 (3H, s, OC*H*$_3$), 3.60-3.51 (2H, m, *H*-3), 1.45 (9H, s, 3 × C(C*H*$_3$)). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) 170.2 (*C*-1), 155.0 (NH*C*OOR), 80.6 (*C*(CH_3)$_3$), 53.8, 53.1, 28.4 (3 × C(*C*H$_3$)), 7.9 (*C*-3).[74,88]

2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranosyl bromide (63)[26]



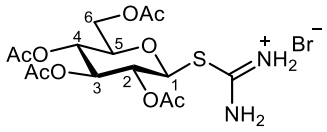
To a solution of penta-acetyl glucopyranose **3** (5.00 g, 12.8 mmol) in dichloromethane (40 mL) at 0 °C, was added HBr (33 % in AcOH, 30 mL) and the stirred reaction mixture allowed to warm to room temperature over 1 h. After a further 2 h the reaction mixture was diluted with water (150 mL), extracted with dichloromethane (4 x 150 mL) and the combined organic extracts basified (to pH = 7) through successive washes with sat. $NaHCO_3$ (2 x 125 mL). The extracted organic layer was washed with sat. brine (200 mL), dried over Na_2SO_4 and filtered. The filtrate was then concentrated under reduced pressure and the resulting crude solid purified *via* recrystallization from ethyl acetate and hexane to afford **63** as a white solid (4.02 g, 9.78 mmol, 76 %).

M.p. 80-82 °C, (lit.[26] 81-83 °C). ^{1}H NMR: δ_H ($CDCl_3$, 400 MHz) 6.59 (1H, d, *J* = 4.1 Hz, *H*-1), 5.54 (1H, t, *J* = 9.7 Hz, *H*-3), 5.14 (1H, t, *J* = 9.8 Hz, *H*-4), 4.82 (1H, dd, *J* = 9.8, 4.1 Hz, *H*-2), 4.33-4.25 (2H, m, *H*-5, *H*-6), 4.15-4.07 (1H, m, *H*-6), [2.08, 2.07, 2.03, 2.01 (12H, 4s, 4 × COC*H*$_3$)]. ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) [170.6, 169.9, 169.8, 169.5 (4 × *C*OCH$_3$)], 86.7 (*C*-1), 72.3 (*C*-5), 70.7 (*C*-2), 70.3 (*C*-3), 67.3 (*C*-4), 61.1 (*C*-6), 20.7 (2 × CO*C*H$_3$), 20.6 (2 × CO*C*H$_3$).[26]

General procedure used for the conversion of glycosyl halides to isothiouronium salts

Procedure (D): To a solution of the glycosyl halide (1.0 eq.) in dry acetone (20 mL) at room temperature, was added thiourea (1.5 eq.) and the reaction mixture heated at reflux for 2 h. The resulting precipitate was recovered by filtration and washed with small amounts of cold acetone to afford the desired isothiouronium salt product.[26]

2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranosyl-1-isothiouronium bromide (64)[26,89]



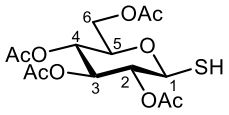
Glycosylation of thiourea (0.790 g, 10.4 mmol) with **63** (2.84 g, 6.91 mmol) was achieved using *Procedure D* and afforded **64** as a white solid (2.02 g, 4.15 mmol, 60 %).

M.p. 186-190 °C, (lit.[89] 191 °C). ^{1}H NMR: δ_H (DMSO-d_6, 400 MHz) 9.23 (2H, bs, NH_2), 9.07 (2H, bs, NH_2), 5.68 (1H, d, *J* = 10.0 Hz, *H*-1), 5.31 (1H, t, *J* = 9.4 Hz, *H*-3), 5.16-5.06 (2H, m, *H*-2, *H*-4), 4.24-4.15 (2H, m, *H*-5, *H*-6), 4.12-4.06 (1H, m, *H*-6), [2.05, 2.02, 2.00, 1.98 (12H, 4s, 4 × COCH_3)]. ^{13}C NMR: δ_C (DMSO-d_6, 101 MHz) [170.0, 169.5, 169.3, 169.2 (4 × *C*OCH$_3$)], 166.3 (*C*NH$_2$NH$_2$), 79.8 (*C*-1), 75.3 (*C*-5), 72.4 (*C*-3), 68.7 (*C*-2), 67.4 (*C*-4), 61.6 (*C*-6), [20.5, 20.4, 20.3 20.2 (4 × CO*C*H$_3$)].[26]

General procedure used for the conversion of isothiouronium salts to 1-thioglycopyranoses

Procedure (E): To a suspension of the isothiouronium salt (1.0 eq.) in a DCM:H_2O (2:1, 30 mL) mixture at room temperature, was added sodium metabisulfite (1.4 eq.) and the reaction mixture heated at 50 °C for 3 h. The reaction mixture was then cooled to room temperature and the extracted aqueous layer re-extracted with dichloromethane (2 × 50 mL). The combined organic extracts were washed with sat. brine (100 mL), dried over Na_2SO_4 and filtered. The filtrate was then concentrated under reduced pressure to give a colourless oil, which solidified upon storage at -18 °C to afford the desired 1-thioglycopyranose product.[26]

2,3,4,6-tetra-*O*-acetyl-1-thio-β-D-glucopyranose (21)[26]



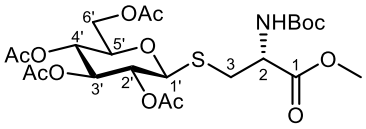
The hydrolysis of **64** (2.00 g, 4.10 mmol) was achieved using *Procedure E* and afforded **21** as a white solid (1.42 g, 3.90 mmol, 95 %).

M.p. 72-74 °C, (lit.[26] 69-70 °C). ^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 5.18 (1H, t, *J* = 9.3 Hz, *H*-3), 5.09 (1H, t, *J* = 9.6 Hz, *H*-4), 4.96 (1H, t, *J* = 9.5 Hz, *H*-2), 4.54 (1H, t, *J* = 9.8 Hz, *H*-1), 4.24 (1H, dd, *J* = 12.5, 4.8 Hz, *H*-6), 4.12 (1H, dd, *J* = 12.5, 2.3 Hz, *H*-6), 3.71 (1H, ddd, *J* = 9.8, 4.8, 2.3 Hz, *H*-5), 2.30 (1H, d, *J* = 9.9 Hz, S*H*), [2.08, 2.07, 2.01, 2.00 (12H, 4s, 4 × COCH_3)]. ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) [170.8, 170.2, 169.7, 169.5 (4 × *C*OCH_3)], 78.9 (*C*-1), 76.5 (*C*-5), 73.7 (*C*-2, *C*-3), 68.3 (*C*-4), 62.2 (*C*-6), 20.8 (2 × CO*C*H_3), 20.7 (2 × CO*C*H_3).[26]

General procedure used for the conversion of 1-thioglycopyranoses to *S*-glycosylcysteines

Procedure (F): To a solution of the 3-iodoproprionate **60** (1.1 eq.) and 1-thioglycopyranose derivative (1.0 eq.) in distilled ethyl acetate (20 mL) at room temperature, was added aq. $NaHCO_3$ (0.6 M, 20 mL) followed by TBAHS (4.0 eq.). The vigorously stirred reaction mixture was then heated at 40 °C for 4 h, cooled to room temperature and the extracted aqueous layer re-extracted with ethyl acetate (2 × 40 mL). The combined organic extracts were washed sequentially with sat. $NaHCO_3$ (50 mL) and sat. brine (50 mL), dried over Na_2SO_4 and filtered. The filtrate was concentrated under reduced pressure to give the crude product which was purified *via* column chromatography. The pooled fractions were then concentrated under reduced pressure to afford the desired *S*-glycosylcysteine product.[24,27]

methyl *N*-(*tert*-butoxycarbonyl)-*S*-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-L-cysteinate (61)[21]



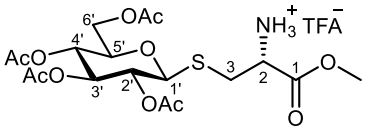
Glycosylation of the 3-iodoproprionate **60** (0.583 g, 1.77 mmol) with **21** (0.607 g, 1.67 mmol) was achieved using *Procedure F* and afforded **61** as a white solid (0.711 g, 1.26 mmol, 75 %), after purification *via* gradient column chromatography (60-100 % ethyl acetate in hexane). Note that this preparation of **61** was repeated and the two batches combined as a basis for the next step.

M.p. 138-139 °C. ^{1}H NMR: δ_H ($CDCl_3$, 400 MHz) 5.65 (1H, d, *J* = 6.9 Hz, N*H*COOR), 5.21 (1H, t, *J* = 9.3 Hz, *H*-3'), 5.09 (1H, t, *J* = 9.7 Hz, *H*-4'), 5.00 (1H, dd, *J* = 10.1, 9.3 Hz, *H*-2'), 4.53 (1H, bs, *H*-2), 4.50 (1H, d, *J* = 10.1 Hz, *H*-1'), 4.23 (1H, dd, *J* = 12.4, 4.8 Hz, *H*-6'), 4.16 (1H, dd, *J* = 12.4, 2.4 Hz, *H*-6'), 3.73 (3H, s, OC*H*$_3$), 3.71 (1H, ddd, *J* = 9.9, 4.8, 2.4 Hz, *H*-5'), 3.22 (1H, dd, *J* = 14.4, 4.2 Hz, *H*-3), 2.97 (1H, dd, *J* = 14.4, 6.1 Hz, *H*-3), [2.10, 2.05, 2.02, 1.99 (12H, 4s, 4 × COC*H*$_3$)], 1.45 (9H, s, 3 × C(C*H*$_3$)). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) [171.1, 170.8, 170.2, 169.5 (4 × *C*OCH$_3$, *C*-1)], 155.3 (NH*C*OOR), 83.2 (*C*-1'), 80.4 (*C*(CH$_3$)$_3$), 76.2 (*C*-5'), 73.9 (*C*-3'), 69.6 (*C*-2'), 68.3 (*C*-4'), 62.1 (*C*-6'), 53.6 (*C*-2), 52.7 (O*C*H$_3$), 32.3 (*C*-3), 28.4 (3 × C(*C*H$_3$)), 20.8 (2 × CO*C*H$_3$), 20.7 (2 × CO*C*H$_3$).[21]

General procedure used for the removal of the *N*-Boc protecting group

Procedure (G): To a solution of the *N*-Boc protected substrate in dichloromethane (10 mL) at 0 °C, was added TFA (2 mL) dropwise over a period of 30 min. The stirred reaction mixture was then allowed to warm to room temperature over 30 min and after a further 2 h excess TFA was evaporated off under a steady-stream of air. The resulting residue was re-dissolved in methanol (10 mL) and solvents removed by azeotropic distillation after addition of toluene, to afford the desired amine product.[72]

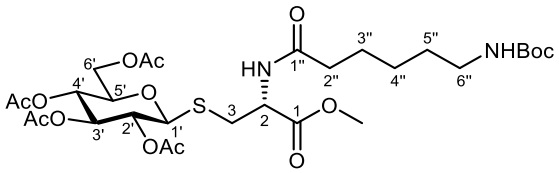
methyl *S*-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-L-cysteinate trifluoroacetate (65)



N-Boc deprotection of **61** (0.946 g, 1.67 mmol) was achieved using *Procedure G* and afforded **65** as a colourless oil (1.08 g), which was taken forward to the next step without further drying.

^{1}H NMR: δ_H (D_2O, 400 MHz) 5.44 (1H, t, *J* = 9.3 Hz, *H*-3'), 5.23-5.17 (2H, m, *H*-2', *H*-4'), 4.96 (1H, d, *J* = 10.1 Hz, *H*-1'), 4.52 (1H, dd, *J* = 7.5, 4.5 Hz, *H*-2), 4.44 (1H, dd, *J* = 12.8, 4.0 Hz, *H*-6'), 4.32 (1H, dd, *J* = 12.8, 2.2 Hz, *H*-6'), 4.14 (2H, ddd, *J* = 10.1, 4.0, 2.2 Hz, *H*-5'), 3.94 (3H, s, OC*H*$_3$), 3.55 (1H, dd, *J* = 15.5, 4.5 Hz, *H*-3), 3.24 (1H, dd, *J* = 15.5, 7.5 Hz, *H*-3), [2.19, 2.18, 2.15, 2.13 (12H, 4s, 4 × COC*H*$_3$)]. ^{13}C NMR: δ_C (D_2O, 101 MHz) [173.6, 173.1, 172.8, 172.7, 168.7 (4 × *C*OCH$_3$, *C*-1], 82.1 (*C*-1'), 75.5 (*C*-5'), 73.8 (*C*-3'), 69.4 (*C*-2'), 68.0 (*C*-4'), 61.9 (*C*-6'), 53.9 (O*C*H$_3$), 53.1 (*C*-2), 29.2 (*C*-3), 20.2 (2 × CO*C*H$_3$), 20.0 (2 × CO*C*H$_3$). ^{19}F NMR: δ_F (D_2O, 377 MHz) -75.5 (3F, s, COC*F*$_3$). HRMS (ESI$^+$): *m/z* calculated for $C_{18}H_{28}NO_{11}S^+$, 466.1378; found 466.1385 $[M]^+$. Note the above analytical data was obtained from a separate experiment, conducted on a 200 mg scale, where the removal of residual solvent proved satisfactory.

methyl *N*-(6-((*tert*-butoxycarbonyl)amino)hexanoyl)-*S*-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-L-cysteinate (66)



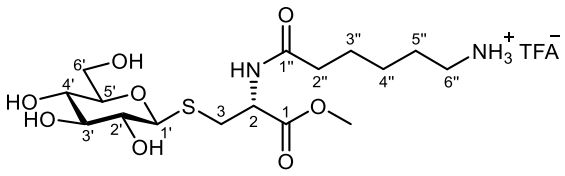
Amide coupling of **65** (1.08 g) with succinimidyl ester **45** (0.558 g, 1.70 mmol) was achieved using *Procedure B* and afforded **66** as a colourless oil (0.784 g, 1.16 mmol, 69 % over two steps) after purification *via* column chromatography (30-100 % hexane in ethyl acetate).

^{1}H NMR: δ_H ($CDCl_3$, 400 MHz) 6.41 (1H, d, *J* = 7.5 Hz, N*H*COR), 5.19 (1H, t, *J* = 9.4 Hz, *H*-3'), 5.05 (1H, t, *J* = 9.6 Hz, *H*-4'), 4.96 (1H, dd, *J* = 10.1, 9.4 Hz, *H*-2'), 4.79 (1H, ddd, *J* = 7.5, 6.2, 4.8 Hz, *H*-2), 4.64 (1H, bs, N*H*COOR), 4.52 (1H, d, *J* = 10.1 Hz, *H*-1'), 4.21 (2H, dd, *J* = 12.5, 4.7 Hz, *H*-6'), 4.16 (1H, dd, *J* = 12.5, 2.5 Hz, *H*-6'), 3.74 (3H, s, OC*H*$_3$), 3.70 (1H, ddd, *J* = 9.9, 4.7, 2.5 Hz, *H*-5'), 3.19 (1H, dd, *J* = 14.0, 4.8 Hz, *H*-3), 3.09 (1H, q, *J* = 6.4 Hz, *H*-6''), 3.03 (1H, dd, *J* = 14.0, 6.2 Hz, *H*-3), 2.23 (2H, t, *J* = 7.5 Hz, *H*-2''), [2.08, 2.04, 2.01, 1.99 (12H, 4s, 4 × COC*H*$_3$)], 1.65 (1H, p, *J* = 7.4 Hz, *H*-3''), 1.49 (1H, p, *J* = 7.5 Hz, *H*-5''), 1.41 (9H, s, 3 × C(C*H*$_3$)), 1.41-1.31 (1H, m, *H*-4''). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) [172.9, 171.1, 170.8, 170.2, 169.6, 169.5 (4 × *C*OCH$_3$, N*H*COR, *C*-1], 156.2 (NH*C*OOR), 83.5 (*C*-1'), 79.2 (*C*(CH$_3$)$_3$), 76.3 (*C*-5'), 73.7 (*C*-3'), 69.9 (*C*-2'), 68.3 (*C*-4'), 62.0 (*C*-6'), 52.9 (O*C*H$_3$), 51.9 (*C*-2), 40.5 (*C*-6''), 36.2 (*C*-2''), 31.9 (*C*-3), 29.9 (*C*-5''), 28.6 (3 × C(*C*H$_3$)), 26.5 (*C*-4''), 25.2 (*C*-3''), 20.8 (2 × CO*C*H$_3$), 20.7 (2 × CO*C*H$_3$). HRMS (ESI$^+$): *m/z* calculated for $C_{29}H_{46}N_2O_{14}S$, 678.2670; found 679.2758 $[M + H]^+$.

General procedure used for the Zemplén deacetylation of glycosyl acetates

Procedure (H): To a solution of the glycosyl acetate substrate (1.0 eq.) in methanol (8 mL) at room temperature, was added a methanolic solution of NaOMe (0.1 M, 1.0 eq.) and the reaction mixture left to stir for 1 h. The reaction mixture was then diluted with methanol, filtered through a pad of Celite and concentrated under reduced pressure to give a crude product. This was then purified *via* flash column chromatography (10 % methanol in ethyl acetate) and the pooled fractions concentrated under reduced pressure to afford the desired de-*O*-acetylated product which was taken directly forward to the next step.[75]

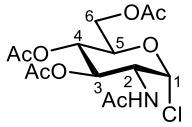
methyl *N*-(6-aminohexanoyl)-*S*-(β-D-glucopyranosyl)-L-cysteinate trifluoroacetate (67)

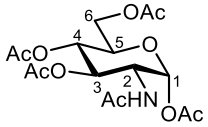


Global deprotection of **66** (0.564 g, 0.831 mmol) was achieved using *Procedure H*, followed by Procedure *E* to afford **67** as an off-yellow oil (0.274 g, 0.522 mmol, 63 % over two steps).

^{1}H NMR: δ_H (D_2O, 300 MHz) 4.85 (1H, o*, *H*-2), 4.70 (1H, d, *J* = 9.8 Hz, *H*-1'), 4.06 (1H, dd, *J* = 12.4, 1.7 Hz, *H*-6'), 3.94 (3H, s, OC*H*$_3$), 3.88 (1H, dd, *J* = 12.4, 5.0 Hz, *H*-6'), 3.69-3.57 (3H, m, *H*-3', *H*-4', *H*-5'), 3.55-3.45 (2H, m, *H*-3, *H*-2'), 3.53-3.39 (3H, m, *H*-3, *H*-6''), 2.52 (2H, t, *J* = 7.3 Hz, *H*-2''), 1.90-1.77 (4H, m, *H*-3'', *H*-5''), 1.62-1.52 (2H, m, *H*-4''). ^{13}C NMR: δ_C (D_2O, 101 MHz) 177.2, 172.9, 85.3 (*C*-1'), 80.5 (*C*-5') 77.7 (*C*-3'), 72.7 (*C*-2'), 70.1 (*C*-4'), 61.5 (*C*-6'), 53.6 (O*C*H$_3$), 53.2 (*C*-2), 39.9 (*C*-6''), 35.5 (*C*-2''), 30.9 (*C*-3), 26.8 (*C*-5''), 25.5 (*C*-4''), 24.9 (*C*-3''). ^{19}F NMR: δ_F (D_2O, 377 MHz) -75.1 (3F, s, COC*F*$_3$). HRMS (ESI$^+$): *m/z* calculated for $C_{16}H_{31}N_2O_8S^+$, 411.1796; found 411.1800 $[M]^+$. *Note that the multiplicity of this signal could not be assigned due to it being obscured by the residual D_2O signal. This was due to the D_2O signal shifting over time, evident in two separate ^{1}H NMR spectra obtained for the same prepared sample. The ^{1}H NMR spectra were instead referenced with respect to the OC*H*$_3$ signal at 3.94 ppm.

2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-glucopyranosyl chloride (69) and 2-acetamido-1,3,4,6-tetra-*O*-acetyl-2-deoxy-α-D-glucopyranose (55)[26,76]

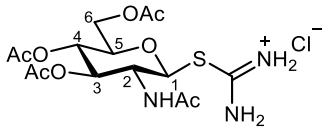




To a solution of excess acetyl chloride (15 mL), fitted with a condenser at room temperature, was added 2-acetamido-2-deoxy-D-glucopyranose **54** (4.00 g, 18.1 mmol) and the reaction mixture heated at 30 ^{o}C for 14 h. The reaction mixture was then cooled to 0 ^{o}C, diluted with ice cold dichloromethane (150 mL) and quenched with ice cold water (100 mL). The extracted organic layer was washed sequentially with cold sat. $NaHCO_3$ (3 × 100 mL) and cold sat. brine (150 mL), dried over Na_2SO_4 and filtered. The filtrate was then concentrated under reduced pressure to afford a white solid being composed of tri-*O*-acetate **69** and tetra-*O*-acetate **55** in approximated yields of (3.79 g, 10.4 mmol, 57 %) and (0.423 g, 1.09 mmol, 6 %) respectively, as judged from integration of their respective anomeric protons.

Glycosyl chloride (**69**); ^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 6.18 (1H, d, *J* = 3.7 Hz, *H*-1), 5.82 (1H, d, *J* = 8.7 Hz, N*H*COCH$_3$), 5.35-5.17 (2H, m, *H*-3, *H*-4), 4.53 (1H, ddd, *J* = 10.5, 8.7, 3.7 Hz, *H*-2), 4.31-4.21 (2H, m, *H*-5, *H*-6), 4.15-4.08 (1H, m, *H*-6), [2.10, 2.05, 2.04 (9H, 3s, 3 × COC*H*$_3$)], 1.98 (3H, s, NHCOC*H*$_3$). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) [171.6, 170.7, 170.2, 169.3 [(3 × *C*OCH$_3$, NH*C*OC*H*$_3$)], 93.8 (*C*-1), 71.1 (*C*-5), 70.3 (*C*-3), 67.1 (*C*-4), 61.3 (*C*-6), 53.7 (*C*-2), 23.2 (NHCO*C*H$_3$), [20.8, 20.7, 20.6 (3 × CO*C*H$_3$)].[26]

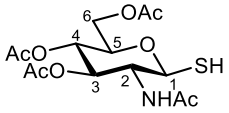
2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-glucopyranosyl-1-isothiouronium chloride (70)[26]



Glycosylation of thiourea (1.04 g, 13.7 mmol) with 3.71 g of the mixture composed of **55** and **69** (≈ 3.34 g, 9.13 mmol) was achieved using Procedure *D* and afforded **70** as a white solid (2.58 g, 5.84 mmol, ≈ 64 %).

M.p. 171-172 °C, (lit.[26] 177-178 °C). ^{1}H NMR: δ_H (DMSO-d_6, 300 MHz) 9.29 (4H, bs, 2 × N*H*$_2$), 8.39 (1H, d, *J* = 9.2 Hz, N*H*COCH$_3$), 5.67 (1H, d, *J* = 10.3 Hz, *H*-1), 5.14 (1H, t, *J* = 9.5 Hz, *H*-3), 4.93 (1H, t, *J* = 9.6 Hz, *H*-4), 4.28-3.95 (4H, m, *H*-2, *H*-5, 2 × *H*-6), [2.01, 1.98, 1.94 (9H, 3s, 3 × COC*H*$_3$)], 1.81 (3H, s, NHCOC*H*$_3$). ^{13}C NMR: δ_C (DMSO-d_6, 101 MHz) 170.0, 169.9, 169.6, 169.3 [(3 × *C*OCH$_3$, NH*C*OC*H*$_3$)], 167.3 (*C*NH$_2$NH$_2$), 80.7 (*C*-1), 74.8 (*C*-5), 72.7 (*C*-3), 67.9 (*C*-4), 61.5 (*C*-6), 51.3 (*C*-2), 22.5 (NHCO*C*H$_3$), [20.6, 20.4, 20.3 (3 × CO*C*H$_3$)].[26]

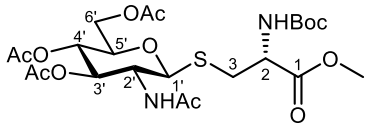
2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-1-thio-β-D-glucopyranose (71)[26]



The hydrolysis of **70** (2.42 g, 5.48 mmol) was achieved using *Procedure E* and afforded **71** as a white solid (1.75 g, 4.82 mmol, 88 %).

M.p. 143-146 °C, (lit.[26] 145-146 °C). ^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 5.97 (1H, d, *J* = 9.4 Hz, N*H*COCH$_3$), 5.14-5.04 (2H, m, *H*-3, *H*-4), 4.60 (1H, t, *J* = 9.7 Hz, *H*-1), 4.21 (1H, dd, *J* = 12.4, 4.8 Hz, *H*-6), 4.14-4.03 (2H, m, *H*-2, *H*-6), 3.69 (1H, ddd, *J* = 9.7, 4.8, 2.4 Hz, *H*-5), 2.53 (1H, d, *J* = 9.3 Hz, S*H*), [2.06, 2.01, 2.00 (9H, 3s, 3 × COC*H*$_3$)], 1.96 (3H, s, NHCOC*H*$_3$). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) [171.2, 170.8, 170.6, 169.4 (3 × *C*OCH$_3$, NH*C*OC*H*$_3$)], 80.3 (*C*-1), 76.3 (*C*-5), 73.6 (*C*-3), 68.4 (*C*-4), 62.3 (*C*-6), 56.9 (*C*-2), 23.4 (NHCO*C*H$_3$), [20.9, 20.8, 20.7 (3 × CO*C*H$_3$)].[26]

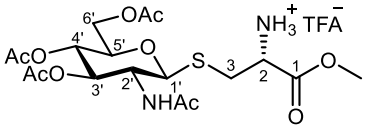
methyl *N*-(*tert*-butoxycarbonyl)-*S*-(2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl)-L-cysteinate (68)[77]



Glycosylation of the 3-iodoproprionoate **60** (1.37 g, 4.16 mmol) with **71** (1.40 g, 3.85 mmol) was achieved using *Procedure F* and afforded **68** as a white solid (1.63 g, 2.89 mmol, 75 %), after purification *via* gradient column chromatography (70-100 % ethyl acetate in hexane).

M.p. 140-141 ^{o}C. ^{1}H NMR: δ_H ($CDCl_3$, 400 MHz) 5.74 (1H, d, *J* = 8.8 Hz, N*H*COCH$_3$), 5.69 (1H, d, *J* = 7.0 Hz, N*H*COOR), 5.16-5.07 (2H, m, *H*-3', *H*-4'), 4.59 (1H, d, *J* = 10.4 Hz, *H*-1'), 4.49 (1H, bs, *H*-2), 4.22 (1H, dd, *J* = 12.3, 5.0 Hz, H-6'), 4.18-4.10 (2H, m, H-2', H-6'), 3.74 (3H, s, OC*H*$_3$), 3.70 (1H, ddd, *J* = 9.6, 5.0, 2.5 Hz, *H*-5'), 3.28 (1H, dd, *J* = 14.3, 3.8 Hz, *H*-3), 2.90 (1H, dd, *J* = 14.3, 7.3 Hz, *H*-3), [2.10, 2.02, (9H, 2s, 3 × COC*H*$_3$)], 1.94 (3H, s, NHCOC*H*$_3$), 1.45 (9H, s, 3 × C(C*H*$_3$)). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) [171.3, 171.2, 170.8, 170.3, 169.4 (3 × *C*OCH$_3$, NH*C*OCH$_3$, *C*-1)], 155.5 (NH*C*OOR), 84.0 (*C*-1'), 80.4 (*C*(CH$_3$)$_3$), 76.3 (*C*-5'), 73.9 (*C*-3'), 68.4 (*C*-4'), 62.4 (*C*-6'), 53.4 (*C*-2), 52.9 (*C*-2'), 52.7 (O*C*H$_3$), 32.3 (*C*-3), 28.5 (3 × C(*C*H$_3$)), 23.3 (NHCO*C*H$_3$), [20.8, 20.7 (3 × CO*C*H$_3$)].

methyl *S*-(2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl)-L-cysteinate trifluoroacetate (72)

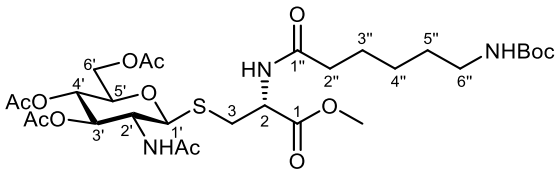


N-Boc deprotection of **68** (1.20 g, 2.13 mmol) was achieved using Procedure *F* and afforded **72** as a colourless oil (1.42 g) which was taken forward to the next step without further drying.

^{1}H NMR: δ_H (D_2O, 400 MHz) 5.32 (1H, t, *J* = 9.4 Hz, *H*-3'), 5.15 (1H, t, *J* = 9.6 Hz, *H*-4'), 4.91 (1H, d, *J* = 10.5 Hz, *H*-1'), 4.51 (1H, dd, *J* = 7.7, 4.4 Hz, *H*-2), 4.44 (1H, dd, *J* = 12.8, 4.0 Hz, *H*-6'), 4.31 (1H, dd, *J* = 12.8, 2.2 Hz, *H*-6'), 4.20 (1H, t, *J* = 10.3 Hz, *H*-2'), 4.09 (1H, ddd, *J* = 10.1, 4.0, 2.2 Hz, *H*-5'), 3.94 (3H, s, OCH_3), 3.56 (1H, dd, *J* = 15.5, 4.4 Hz, *H*-3), 3.22 (1H, dd, *J* = 15.5, 7.7 Hz, *H*-3), [2.18, 2.14, 2.12 (9H, 3s, 3 × COCH_3)], 2.04 (3H, s, NHCOCH_3). ^{13}C NMR: δ_C (D_2O, 101 MHz) 174.5, 173.6, 173.2, 172.8, 168.7 (3 × *C*OCH$_3$, NH*C*OCH$_3$, *C*-1), 83.1 (*C*-1'), 75.5 (*C*-5'), 73.7 (*C*-3'), 68.3 (*C*-4'), 62.0 (*C*-6'), 53.9 (O*C*H$_3$), 53.1 (*C*-2), 52.0 (*C*-2'), 29.3 (*C*-3), 21.9 (NHCO*C*H$_3$), [20.2, 20.1, 20.0 (3 × CO*C*H$_3$)]. ^{19}F NMR: δ_F (D_2O, 400 MHz) -75.5 (3F, s, COCF_3). HRMS (ESI$^+$): *m/z* calculated for $C_{18}H_{28}NO_{11}S^+$, 465.1537; found 465.1537 $[M]^+$. Note the above analytical data was obtained from a separate experiment, conducted on a 200 mg scale, where the removal of excess solvent proved satisfactory.

.
.

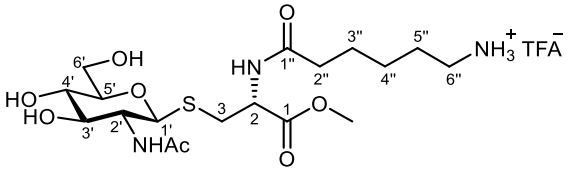
methyl *N*-(6-((*tert*-butoxycarbonyl)amino)hexanoyl)-*S*-(2-acetamido-tri-3,4,6-tetra-*O*-acetyl-2-deoxy-β-D-glucopyranosyl)-L-cysteinate (73)



Amide coupling of **72** (1.42 g) with succinimidyl ester **45** (0.732 g, 2.23 mmol) was achieved using *Procedure B* and afforded **73** as a white solid (1.13 g, 1.67 mmol, 78 % over two steps) after purification *via* column chromatography (70-100 % ethyl acetate in hexane).

M.p. 162-164 °C. ^{1}H NMR: δ_H ($CDCl_3$, 300 MHz) 6.58 (1H, d, *J* = 8.0 Hz, N*H*COCH$_3$), 6.13 (1H, d, *J* = 8.0 Hz, N*H*COR), 5.18 (1H, t, *J* = 9.7 Hz, *H*-3'), 5.06 (1H, t, *J* = 9.7 Hz, *H*-4'), 4.82-4.76 (1H, m, *H*-2), 4.76 (1H, *J* = 10.1 Hz, *H*-1'), 4.65 (1H, bs, N*H*COOR), 4.23 (1H, dd, *J* = 12.4, 4.9 Hz, *H*-6'), 4.15 (1H, dd, *J* = 12.4, 2.5 Hz, *H*-6'), 4.05-3.97 (1H, m, *H*-2'), 3.75 (3H, s, OC*H$_3$*), 3.70 (1H, ddd, *J* = 9.9, 4.9, 2.5 Hz, *H*-5'), 3.27 (1H, dd, *J* = 13.9, 4.0 Hz, *H*-3), 3.10 (1H, q, *J* = 6.6 Hz, *H*-6''), 2.86 (1H, dd, *J* = 13.9, 8.7 Hz, *H*-3), 2.27 (2H, t, *J* = 7.4 Hz, *H*-2''), [2.08, 2.02, 2.00 (9H, 2s, 3 × COC*H$_3$*)], 1.93 (3H, s, NHCOC*H$_3$*), 1.66 (2H, *J* = 7.4 Hz, *H*-3''), 1.49-1.31 (4H, m, *H*-4'', *H*-5''), 1.43 (9H, s, 3 × C(C*H$_3$*)). ^{13}C NMR: δ_C ($CDCl_3$, 101 MHz) [173.5, 171.2, 171.0, 170.8, 170.7, 169.5 (3 × *C*OCH$_3$, NH*C*OCH$_3$, NH*C*OR, *C*-1)], 156.3 (NH*C*OOR), 83.6 (*C*-1'), 79.4 (*C*(CH$_3$)$_3$), 76.3 (*C*-5'), 73.9 (*C*-3'), 68.5 (*C*-4'), 62.3 (*C*-6'), 53.3 (*C*-2'), 52.9 (O*C*H$_3$), 51.1 (*C*-2), 40.7 (*C*-6''), 36.3 (*C*-2''), 32.2 (*C*-3), 29.8 (*C*-5''), 28.6 (3 × C(*C*H$_3$)), 26.4 (*C*-4''), 25.3 (*C*-3''), 23.4 (NHCO*C*H$_3$), [20.9, 20.8, 20.7 (3 × CO*C*H$_3$)]. HRMS (ESI$^+$) *m/z* calculated for $C_{29}H_{47}N_3O_{13}S$, 677.2830; found 678.2921 $[M + H]^+$.

methyl *N*-(6-aminohexanoyl)-*S*-(2-acetamido-2-deoxy-β-D-glucopyranosyl)-L-cysteinate trifluoroacetate (74)



Global deprotection of **73** (0.404 g, 0.596 mmol) was achieved using *Procedure H*, followed by Procedure *E* to afford **74** as a cream coloured oil (0.232 g, 0.410 mmol, 69 % over two steps).

^{1}H NMR: δ_H (D_2O, 400 MHz) 4.82 (1H, dd, *J* = 8.9, 4.7 Hz, *H*-2), 4.80 (1H, *J* = 10.3, *H*-1'), 4.07 (1H, dd, *J* = 12.4, 1.6 Hz, *H*-6'), 3.95-3.88 (2H, m, *H*-2', *H*-6'), 3.94 (3H, s, OC*H*$_3$), 3.76-3.69 (1H, m, *H*-3'), 3.67-3.61 (2H, m, *H*-4', *H*-5'), 3.50 (1H, dd, *J* = 14.2, 4.7 Hz, *H*-3), 3.19-3.10 (3H, m, *H*-3, *H*-6''), 2.50 (2H, t, *J* = 7.4 Hz, *H*-2''), 2.18 (3H, s, NHCOC*H*$_3$), 1.88-1.77 (4H, m, *H*-3'', *H*-5''), 1.58 (2H, p, *J* = 7.4 Hz, *H*-4''). ^{13}C NMR: δ_C (D_2O, 101 MHz) 177.1, 174.7, 172.9, 84.2 (*C*-1'), 80.5 (*C*-5'), 75.5 (*C*-3'), 70.3 (*C*-4'), 61.4 (*C*-6'), 55.0 (*C*-2'), 53.5 (O*C*H$_3$), 53.0 (*C*-2), 39.8 (*C*-6''), 35.4 (*C*-2''), 31.2 (*C*-3), 26.7 (*C*-5''), 25.4 (*C*-4''), 24.9 (*C*-3''), 22.6 (NHCO*C*H$_3$). ^{19}F NMR: δ_C (D_2O, 377 MHz) -74.9 (3F, s, COC*F*$_3$). HRMS (ESI$^+$): *m/z* calculated for $C_{16}H_{31}N_2O_8S^+$, 452.2061; found 452.2062 $[M]^+$. Note the ^{1}H NMR spectrum was referenced with respect to the OC*H*$_3$ signal at 3.94 ppm, for reasons discussed earlier.

References

1. M. M. L. Zulueta, Y.-Q. Zhong and S.-C. Hung, *Chem. Commun.*, 2013, **49**, 3275–3287.

2. B. Donczo, J. Kerekgyarto, Z. Szurmai and A. Guttman, *Analyst*, 2014, **139**, 2650–2657.

3. C. D. Rillahan and J. C. Paulson, *Annu. Rev. Biochem.*, 2011, **80**, 797–823.

4. N. Laurent, J. Voglmeir and S. L. Flitsch, *Chem. Commun.*, 2008, 4400–4412.

5. K. M. Bolles, F. Cheng, J. Burk-Rafel, M. Dubey and D. M. Ratner, *Materials*, 2010, **3**, 3948–3964.

6. R. D. Cummings, *Mol. BioSyst.*, 2009, **5**, 1087–1104.

7. M. G. Nokwequ, C. M. Nkambule and D. W. Gammon, *S. Afr. J. Chem.*, 2014, **67**, 180–183.

8. A. Schierholt, M. Hartmann, K. Schwekendiek and T. K. Lindhorst, *Eur. J. Org. Chem.*, 2010, **2010**, 3120–3128.

9. M. Lo Conte, S. Pacifico, A. Chambery, A. Marra and A. Dondoni, *J. Org. Chem.*, 2010, **75**, 4644–4647.

10. B. L. Anderson, L. Teyton, A. Bendelac and P. B. Savage, *Molecules*, 2013, **18**, 15662–15688.

11. H. Zappe, M. E. Snell and M. J. Bossard, *Adv. Drug Deliv. Rev.*, 2008, **60**, 79–87.

12. M. Fais, R. Karamanska, S. Allman, S. A. Fairhurst, P. Innocenti, A. J. Fairbanks, T. J. Donohoe, B. G. Davis, D. A. Russell and R. A. Field, *Chem. Sci.*, 2011, **2**, 1952–1959.

13. A. A. Wube, A. Hüfner, C. Thomaschitz, M. Blunder, M. Kollroser, R. Bauer and F. Bucar, *Bioorg. Med. Chem.*, 2011, **19**, 567–579.

14. C. Grandjean, A. Boutonnier, C. Guerreiro, J.-M. Fournier and L. A. Mulard, *J. Org. Chem.*, 2005, **70**, 7123–7132.

15. Y.-C. Su, Y.-L. Lo, C.-C. Hwang, L.-F. Wang, M. H. Wu, E.-C. Wang, Y.-M. Wang and T.-P. Wang, *Org. Biomol. Chem.*, 2014, **12**, 6624–6633.

16. T. Nawrath, K. Gerth, R. Müller and S. Schulz, *ChemBioChem*, 2010, **11**, 1914–1919.

17. D. W. Gammon, D. J. Steenkamp, V. Mavumengwana, M. J. Marakalala, T. T. Mudzunga, R. Hunter and M. Munyololo, *Bioorg. Med. Chem.*, 2010, **18**, 2501–2514.

18. Q. Wan and S. J. Danishefsky, *Angew. Chem. Int. Ed.*, 2007, **46**, 9248–9252.

19. Q. Wan, J. Chen, Y. Yuan and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2008, **130**, 15814–15816.

20. M. G. Nokwequ, C. M. Nkambule and D. W. Gammon, *Carbohydr. Res.*, 2012, **359**, 18–23.

21. S. Mandal and U. J. Nilsson, *Org. Biomol. Chem.*, 2014, **12**, 4816–4819.

22. D. A. Thayer, H. N. Yu and M. C. Galan, C.-H. Wang, *Angew. Chem. Int. Ed.*, 2005, **44**, 4596–4599.

23. G. J. L. Bernardes, E. J. Grayson, S. Thompson, J. M. Chalker, J. C. Errey, F. El Oualid, T. D. W. Claridge and B. G. Davis, *Angew. Chem. Int. Ed.*, 2008, **47**, 2244–2247.

24. Y. S. Y. Hsieh, B. L. Wilkinson, M. R. O'Connell, J. P. Mackay, J. M. Matthews and R. J. Payne, *Org. Lett.*, 2012, **14**, 1910–1913.

25. L. A. Salvador, M. Elofsson and J. Kihlberg, *Tetrahedron*, 1995, **51**, 5643–5656.

26. N. Floyd, B. Vijayakrishnan, J. R. Koeppe and B. G. Davis, *Angew. Chem. Int. Ed.*, 2009, **48**, 7798–7802.

27. X. Zhu and R. R. Schmidt, *Chem. Eur. J.*, 2004, **10**, 875–887.

28. A. El-Faham and F. Albericio, *Chem. Rev.*, 2011, **111**, 6557–6602.

29. C. A. G. N. Montalbetti and V. Falque, *Tetrahedron*, 2005, **61**, 10827–10852.

30. J. W. Wehner, M. Hartmann and T. K. Lindhorst, *Carbohydr. Res.*, 2013, **371**, 22–31.

31. G. Aksnes, J. E. Prue, *J. Chem. Soc.*, 1959, 103–107.

32. *WO Pat.* 2000007979, 2000.

33. *WO Pat.* 2010082050, 2010.

34. K. Öberg, J. Ropponen, J. Kelly, P. Löwenhielm, M. Berglin and M. Malkoch, *Langmuir*, 2013, **29**, 456–465.

35. T.-B. Yu, J. Z. Bai and Z. Guan, *Angew. Chem. Int. Ed.*, 2009, **48**, 1097–1101.

36. H. Sun and X. Peng, *Bioconjugate Chem*. 2013, **24**, 1226−1234.

37. S. M. Andersen, M. Heuckendorff and H. H. Jensen, *Org. Lett.*, 2012, **17**, 944–947.

38. T. Rodríguez-pérez, I. Lavandera, S. Fernández, Y. S. Sanghvi, M. Ferrero and V. Gotor, *Eur. J. Org. Chem.* 2007, **2007**, 2769–2778.

39. S. Konstantinović, J. Predjoveć, B. Mojsilović, B. Dimitrijević and G. Milošević, *Indian. J. Chem.*, *Sec. B*, 2001, **40B**, 796–801.

40. T. Desmet and W. Soetaert, *Biocatal. Biotransforn.*, 2011, **29**, 1–18.

41. D. B. G. Williams and M. Lawton, *J. Org. Chem.*, 2010, **75**, 8351–8354.

42. E. D. Goddard-Borger and R. V. Stick, *Org. Lett.*, 2007, **9**, 3797–3800.

43. A. E. Voelker and R. Viswanathan, *J. Org. Chem.*, 2013, **78**, 9647–9658.

44. O. David, W. J. N. Meester, H. Bieräugel, H. E. Schoemaker, H. Hiemstra and J. H. van Maarseveen, *Angew. Chem. Int. Ed.*, 2003, **42**, 4373–4375.

45. J. Cavender and V. J. Shiner Jr., *J. Org. Chem.*, 1972, **37**, 3567–3569.

46. A. P. Kumar, S. P. Sarma and A. G. Samuelson, *Tetrahedron Lett.*, 2014, **55**, 2917–2920.

47. M. D. Threadgill and A. P. Gledhill, *J. Org. Chem.*, 1989, **54**, 2940–2949.

48. G.-M. Dumitriu, E. Bîcu, D. Belei, B. Rigo, J. Dubois, A. Farce and A. Ghinet, *Bioorg. Med. Chem. Lett.*, 2015, **25**, 4447–4452.

49. *WO Pat.* 2012065102, 2012.

50. E. J. Kragulj, J. L. Gustafson, D. B. Grotjahn, *Synlett*, 2007, **2007**, 2851–2584.

51. Y. Jia, X. Dong, P. Zhou, X. Liu, L. Pan, H. Xin, Y. Z. Zhu and Y. Wang, *Eur. J. Med. Chem.*, 2012, **55**, 176–187.

52. J. M. Chalker, S. B. Gunnoo, O. Boutureira, S. C. Gerstberger, M. Fernández-González, G. J. L. Bernardes, L. Griffin, H. Hailu, C. J. Schofield and B. G. Davis, *Chem. Sci.*, 2011, **2**, 1666–1676.

53. R. H. Sifferd and V. du Vigneaud, *J. Biol. Chem.*, 1935, **108**, 753–761.

54. W. I. Patterson and V. du Vigneaud, *J. Biol. Chem.*, 1935, **111**, 393–398.

55. C. Mckeever, M. Kaiser and I. Rozas, *J. Med. Chem.*, 2013, **56**, 700–711.

56. N. Gavande, H.-L. Kim, M. R. Doddareddy, G. A. R. Johnston, M. Chebib and J. R. Hanrahan, *ACS Med. Chem. Lett.*, 2013, **4**, 402–407.

57. C.-H. Wong, X.-S. Ye, Z. Zhang, *J. Am. Chem. Soc.*, 1998, **120**, 7137–7138.

58. J. Tatai and P. Fügedi, *Org. Lett.*, 2007, **9**, 4647–4650.

59. C. Berse, R. Boucher and L. Piché, *J. Org. Chem.*, 1957, **22**, 805–808.

60. D. C. Gowda, B. Rajesh and S. Gowda, *Indian. J. Chem.*, *Sec. B*, 2000, **39B**, 504–508.

61. G. R. Srinivasa, S. N. N. Babu, C. Lakshmi and D. C. Gowda, *Synth. Commun.* 2004, **34**, 1831–1837.

62. A. Isidro-llobet, M. Álvarez and F. Albericio, *Chem. Rev.*, 2009, **109**, 2455–2504.

63. O. Hernout, K. Berthoin, I. Delattre, P. M. Tulkens, S. Carryn and J. Marchand-Brynaert, *Bioorg. Med. Chem. Lett.*, 2007, **17**, 5758–5762.

64. B. E.-D. M. El-Gendy, E. H. G. Zadeh, A. C. Sotuyo, G. G. Pillai and A. R. Katritzky, *Chem. Biol. Drug Des.*, 2013, **81**, 577–582.

65. S. Nakabayashi, C. D. Warren and R. W. Jeanloz, *Carbohydr. Res.*, 1986, **150**, C7–C10.

66. V. K. Srivastava, *Carbohydr. Res.*, 1982, **103**, 286–292.

67. R. Haddoub, N. Laurent, M. M. Meloni and S. L. Flitsch, *Synlett*, 2009, **2009**, 3328–3332.

68. E. C. Lourenço and M. R. Ventura, *Carbohydr. Res.*, 2016, **426**, 33–39.

69. F. Dasgupta and P. J. Garegg, *J. Chem. Soc.*, *Chem. Commun.*, 1989, 1640–1641.

70. V. Pavliak and P. Kováč, *Carbohydrate Res.*, 1991, **210**, 333–337.

71. L. Käsbeck and H. Kessler, *Liebigs Ann./Recueil*, 1997, 165–167.

72. X. Zhu, K. Pachamuthu and R. R. Schmidt, *J. Org. Chem.*, 2003, **68**, 5641–5651.

73. J. Chen, W. Wu and A. J. McNeil, *Chem. Commun.*, 2012, **48**, 7310–7312.

74. N. D. P. Atmuri and W. D. Lubell, *Org. Synth.*, 2015, **92**, 103–116.

75. J. Ramos-Soriano, U. Niss, J. Angulo, M. Angulo, A. J, Moreno-Vargas, A. T. Carmona, S. Ohlson and I. Robina, *Chem. Eur. J.*, 2013, **19**, 17989–18003.

76. D. Horton, *Org. Synth.*, 1966, **46**, 1–5.

77. S. Knapp and D. S. Myers, *J. Org. Chem.*, 2001, **66**, 3636–3638.

78. J. Y. Baek, T. J. Choi, H. B. Jeon and K. S. Kim, *Angew. Chem. Int. Ed.*, 2006, **45**, 7436–7440.

79. C. O'Reilly and P. V. Murphy, *Org. Lett.*, 2011, **13**, 5168–5171.

80. W. Pilgrim, C. O'Reilly and P. V. Murphy, *Molecules*, 2013, **18**, 11198–11218.

81. J. Royes, C. Provenzano, P. Pagliusi, R. M. Tejedor, M. Piñol and L. Oriol, *Macromol. Rapid Commun.*, 2014, **35**, 1890–1895.

82. C. M. Jakobsen, S. R. Denmeade, J. T. Isaacs, A. Gady, C. E. Olsen and S. B. Christensen, *J. Med. Chem.*, 2001, **44**, 4696–4703.

83. M. B. Doughty, C. S. Chaurasia and K. Li, *J. Med. Chem.*, 1993, **36**, 272–279.

84. J. R. Silvius and F. l'Heureux, *Biochemistry*, 1994, **33**, 3014–3022.

85. B. Coxon, *Carbohydr. Res.*, 2005, **340**, 1714–1721.

86. M. L. Wolfrom, R. U. Lemieux and S. M. Olin, *J. Am. Chem. Soc.*, 1949, **71**, 2870–2873.

87. M. Ikubo, A. Inoue, S. Nakamura, S. Jung, M. Sayama, Y. Otani, A. Uwamizu, K. Suzuki, T. Kishi, A. Shuto, J. Ishiguro, M. Okudaira, K. Kano, K. Makide, J. Aoki and T. Ohwada, *J. Med. Chem.*, 2015, **58**, 4204–4219.

88. R. W. Jackson and M. Perez-Gonzalez, *Org. Synth.*, 2005, **81**, 77–88.

89. H. Beyer and U. Schultz, *Chem. Ber.*, 1954, **87**, 78–81.

Appendix

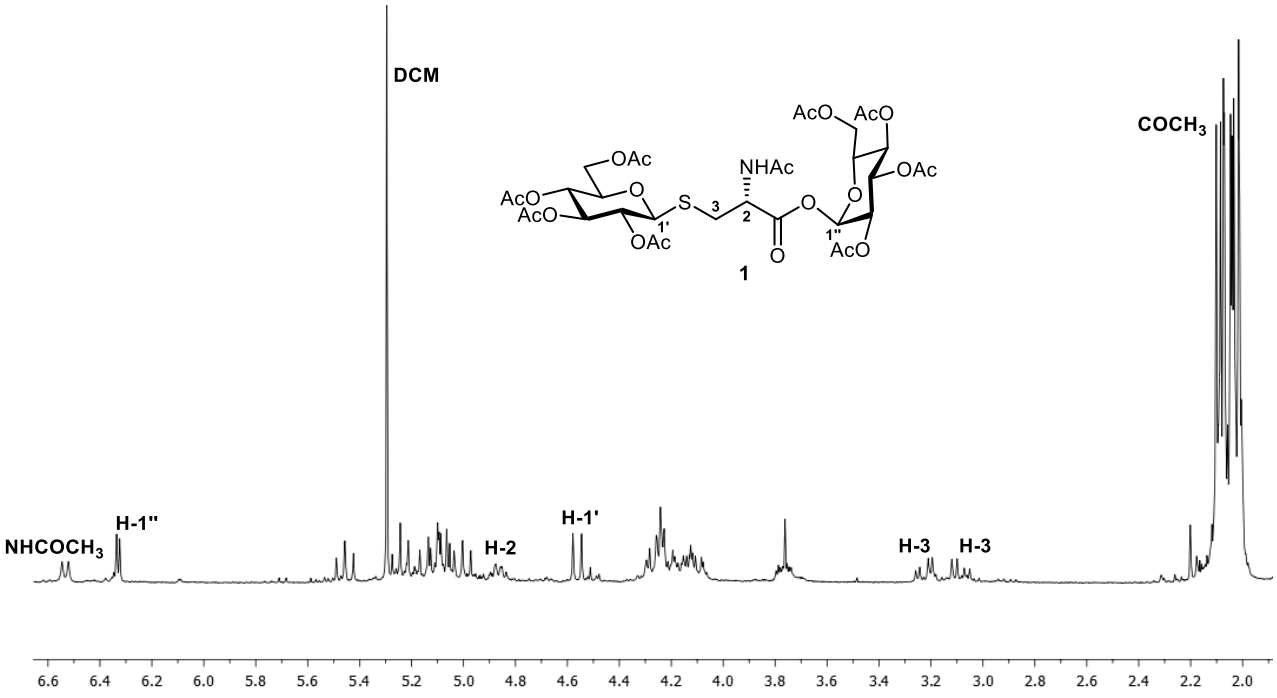


Figure 22: The ^{1}H NMR spectrum (1.90 - 6.60 ppm) of a partially purified sample of BGC **1**. Key signals advocating for the presence of **1** have been assigned and are in good agreement with the reported literature values.[7]