Glycoconjugates

Solid-phase synthesis and biological applications

Fredrik Wallner
Title
Glycoconjugates – Solid-phase synthesis and biological applications

Author
Fredrik Wallner, Department of Chemistry, Organic Chemistry, Umeå University, SE-901 87 Umeå, Sweden.

Abstract
Glycoconjugates are biologically important molecules with diverse functions. They consist of carbohydrates of varying size and complexity, attached to a non-sugar moiety as a lipid or a protein. Glycoconjugate structures are often very complex and their intricate biosynthetic pathways makes overexpression difficult. This renders the isolation of pure, structurally defined compounds from natural sources cumbersome. Therefore, to better address questions in glycobiology, synthetic glycoconjugates are an appealing alternative. In addition, synthetic methods allow for the preparation of non-natural glycoconjugates that can enhance the understanding of the influence of structural features on the biological responses.

In this thesis, synthetic methods for the preparation of glycoconjugates, especially glycolipid analogues, have been developed. These methods make use of solid-phase chemistry and are amenable to library synthesis of series of similar compounds. Solid-phase synthesis is a technique where the starting material of the reaction is attached to small plastic beads through a linker. This allows large excess of reagents to speed up the reactions and the sometimes difficult purifications of intermediate products are reduced to simple washings of the beads.

One problem with solid-phase synthesis is the difficulties to monitor the reactions and characterize the intermediate products. Gel-phase $^{19}$F-NMR spectroscopy, using fluorinated linkers and protecting groups, is an excellent tool to overcome this problem and to monitor solid-phase synthesis of e.g. glycoconjugates. Two novel fluorinated linkers for the attachment of carboxylic acids have been developed and are presented in the thesis. These linkers can be cleaved with both acids of varying strengths and nucleophiles like hydroxide ions, and they are stable to glycosylation conditions. In addition, a novel filter reactor for solid-phase synthesis was designed. The reactor fits into an ordinary NMR spectrometer to facilitate the reaction monitoring with gel-phase $^{19}$F-NMR spectroscopy.

The biological applications of the synthesized glycolipids were demonstrated in two different settings. The CD1d restricted binding of glycolipids carrying the monosaccharide $\alpha$-GalNAc as carbohydrate could be detected on viable cells of mouse origin. CD1d is one of several antigen presenting molecules (the CD1 proteins) that presents lipids and glycolipids to circulating T-cells that in turn can initiate an immune response. The CD1 molecules are relatively sparsely investigated, and the method to measure glycolipid binding on viable cells, as described in the thesis, has the possibility to greatly enhance the knowledge of the structural requirements for CD1-binding.

Serine-based neoglycolipids with terminal carboxylic acids were used to prepare glycoconjugate arrays with covalent bonds to secondary amines on microtiter plates. Carbohydrate arrays have great possibilities to simplify the study of interactions between carbohydrates and e.g. proteins and microbes. The usefulness of the glycolipid arrays constructed in the thesis was illustrated with two lectins, RCA$_{120}$ from *Ricinus communis* and BS-1 from *Bandeiraea simplicifolia*. Both lectins bound to the array of neoglycolipids in agreement with their respective specificity for galactosides.

Glycobiology is a large area of great interest and the methods described in this thesis can be used to answer a variety of glycoconjugate-related biological questions.

Keywords
Glycoconjugate, Glycolipid, Solid-phase synthesis, Glycosylation, Gel-phase $^{19}$F-NMR spectroscopy, Fluorinated linker, Carbohydrate array, CD1

ISBN: 91-7305-984-6
To Anna and Thea...
List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


Reprints were made with kind permission from the publishers.
Solid-phase synthesis and biological applications

List of abbreviations

aq. aqueous
atm atmosphere
BH$_3$·DMS borane dimethyl sulfide complex
BS-1 Bandeiraea simplicifolia lectin
BSA bovine serum albumin
CD1 cluster of differentiation 1
cDNA complementary deoxyribonucleic acid
DBU 1,8-diazabicyclo[5.4.0]undec-7-ene
DC dendritic cell
DIC $N,N'$-diisopropyl carbodiimide
DMAP $N,N$-dimethylaminopyridine
DMF $N,N$-dimethylformamide
DMSO dimethylsulfoxide
EDC 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride
FACS fluorescent-activated cell sorting
FITC fluorescein isothiocyanate
Fmoc 9-fluorenylmethoxycarbonyl
Fmoc-Gly-OH $N$-Fmoc-glycine
Fmoc-$p$-F-Phe-OH $N$-Fmoc-$para$-fluoro-phenylalanine

- vii -
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gal</td>
<td>galactose</td>
</tr>
<tr>
<td>GalNAc</td>
<td>$N$-acetyl-galactosamine</td>
</tr>
<tr>
<td>HIV</td>
<td>human immunodeficiency virus</td>
</tr>
<tr>
<td>HOAt</td>
<td>1-hydroxy-7-azabenzotriazole</td>
</tr>
<tr>
<td>HOBt</td>
<td>1-hydroxybenzotriazole</td>
</tr>
<tr>
<td>HRP</td>
<td>horseradish peroxidase</td>
</tr>
<tr>
<td>LC/MS</td>
<td>liquid chromatography, mass spectrometry</td>
</tr>
<tr>
<td>MeIm</td>
<td>$N$-methyl imidazole</td>
</tr>
<tr>
<td>MFI</td>
<td>mean fluorescence intensity</td>
</tr>
<tr>
<td>mFPh</td>
<td>meta-fluorophenyl</td>
</tr>
<tr>
<td>MHC</td>
<td>Major Histocompatibility Complex</td>
</tr>
<tr>
<td>MSNT</td>
<td>1-(2-mesitylenesulfonyl)-3-nitro-$1H$-$1,2,4$-triazole</td>
</tr>
<tr>
<td>NHS</td>
<td>$N$-hydroxysuccinimide</td>
</tr>
<tr>
<td>NIS</td>
<td>$N$-iodosuccinimide</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>oFBrn</td>
<td>orto-fluorobenzyl</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
</tr>
<tr>
<td>pFBz</td>
<td>para-fluorobenzoyl</td>
</tr>
<tr>
<td>RCA$_{120}$</td>
<td><em>Ricinus communis</em> agglutinin</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>TFA</td>
<td>trifluoroacetic acid</td>
</tr>
<tr>
<td>TfOH</td>
<td>trifluoromethane sulphonic acid</td>
</tr>
<tr>
<td>THF</td>
<td>tetrahydrofuran</td>
</tr>
<tr>
<td>TLC</td>
<td>thin-layer chromatography</td>
</tr>
</tbody>
</table>
## Contents

1 Glycoconjugates ........................................ 1
   1.1 Glycolipids ....................................... 3
       1.1.1 CD1 ........................................ 4
   1.2 Glycoconjugate synthesis ......................... 6
2 Solid-phase synthesis ................................. 9
   2.1 Gel-phase $^{19}$F-NMR spectroscopy .............. 9
   2.2 Linkers .......................................... 11
3 Carbohydrate arrays .................................. 13
4 Aims of the thesis .................................. 17
5 Glycolipid synthesis ................................ 19
   5.1 Synthesis of $\alpha$-N-acetyl-galactosamine glycolipids (paper I) ......................... 19
   5.2 Solid-phase glycosylations in glycolipid synthesis (paper II) ............................. 22
6 Glycolipid loading of CD1d (paper I) ............... 27
7 Preparation of glycolipid arrays (paper II) ....... 33
8 Design, synthesis, and evaluation of fluorinated linkers ... 37
   8.1 Synthesis of a new fluorinated monoalkoxy linker (paper II) ............................... 37
   8.2 Design, synthesis, and evaluation of a fluorinated dialkox y linker (paper III) ............... 38
9 Design of an NMR tube filter reactor for solid-phase synthesis and gel-phase $^{19}$F-NMR spectroscopy (paper IV) .......... 47
10 Concluding remarks and future perspectives ........ 51
11 Acknowledgments .................................... 55
References .............................................. 57
1 Glycoconjugates

Glycoconjugates are carbohydrates covalently linked to a non-sugar moiety such as a lipid or protein. They are often found on the outside of cell membranes and are for example involved in specific recognition events between cells.

Oligosaccharides, together with nucleic acids and proteins constitute the three major biopolymers. While nucleic acids and proteins are linear polymers, oligosaccharides are often branched in their structure. Each monosaccharide has several hydroxyl groups that other monosaccharides can be coupled to and the glycosidic linkage can have two different stereoisomers, the $\alpha$ and $\beta$ configurations. In addition, each monosaccharide can exist in two different ring forms, the furanose and the pyranose, and have further substituents on any of the hydroxyls. Together, this means that the possible ways for two monosaccharides to connect are numerous (see figure 1.1) and, consequently, large amounts of biological information can be encoded in a rather short saccharide. For example, over 38,000 different trisaccharides can be formed out of three D-hexoses. Despite this capacity for information storage, oligosaccharides were long thought to act only as a physical defense, energy storage, or to change the physical properties of proteins and lipids. Today, it is known that carbohydrate-protein and carbohydrate-carbohydrate interactions are important in many biological processes like development, immune responses, and microbial infections as well as protein folding.

While the biosynthesis of proteins and nucleic acids follows a template, the assembly of the oligosaccharides is more complex and depends on which glycosylating enzymes that are active in the cell and if they are capable of glycosylating the substrate. These enzyme reactions do not always reach completion, resulting in a set of oligosaccharides instead of a single isomer. Two copies of the same protein, synthesized in the same cell, may have somewhat different glycosylation patterns. These proteins are referred to as two glycoforms.
and even though there are several glycoforms of a protein in a given cell, the set of glycoforms, or the glycan profile, is often very characteristic for a cell type and/or condition.\(^7\) It is well established that the glycosylation pattern of a cell changes with the development of cancer and inflammation, with the first examples as early as 1968.\(^9\) These changes can be used to diagnose or treat the disease. Several examples exist where immunizations with glycoconjugates are used in tumor immunotherapy and different vaccines are in phase I to III clinical trials.\(^7\) Here, one of the large obstacles is that the individual glycans are endogenous and often exist also in embryonal stages and/or in low amounts in normal tissue. However, the specific glycosylation patterns with elevated or reduced levels of certain glycans are different between healthy cells and cancer cells.\(^7\) On the other hand, microbial pathogens often have completely different glycans than humans and carbohydrate based vaccines are used against sev-
eral pathogens and much efforts are made on many more.\textsuperscript{10} As a result of the difference in carbohydrate structures of humans and pathogens, an immune response raised by a vaccine against a microbial pathogen is more likely to be specific than a cancer vaccine.

The information in oligosaccharides is often decoded by lectins, carbohydrate recognizing proteins other than antibodies and enzymes,\textsuperscript{11} but also antibodies\textsuperscript{4} and T-cells\textsuperscript{12} can recognize specific carbohydrate structures. The interaction between proteins and carbohydrates is mostly very week and multiple simultaneous binding events are required for a biological response.\textsuperscript{13} This phenomenon, called multivalency, has a number of physiological advantages such as a graded response,\textsuperscript{14} context-dependent interactions,\textsuperscript{4} fast on/off rates,\textsuperscript{15} and resistance to shear stress.\textsuperscript{13} At the same time, it complicates the use of small molecules to interfere with, and study, the interactions and therefore multivalent ligands are increasingly used.\textsuperscript{13}

\subsection{1.1 Glycolipids}

Glycoconjugates where the carbohydrate is attached to a lipid are called glycolipids. The combination of a polar carbohydrate and a non-polar lipid makes glycolipids very amphiphilic and they are almost always found in cell membranes, preferably on the outside.\textsuperscript{16} In animals, the majority of glycolipids are glycosphingolipids, where the carbohydrate is attached to a sphingoid base that is acylated with a fatty acid\textsuperscript{16} (see figure 1.2a). The carbohydrate parts of the glycosphingolipids are however very diverse, and more than 200 natural glycosphingolipids have been identified.\textsuperscript{17} Lower organisms, e.g. bacteria and protozoa, have a more heterogeneous collection of glycolipids, unique both in the carbohydrate and the lipid parts (figure 1.2b),\textsuperscript{16,18} although the total amounts are not as high.

The glycolipids have similar biological responses as the other glycoconjugates. The placement of their carbohydrate in close proximity to the cell membrane has led to a common use as attachment points for microbes such as uropathogenic \textit{E. coli}\textsuperscript{19} and HIV.\textsuperscript{20}
While the presentation of protein and peptide antigens to T-cells through Major Histocompatibility Complex (MHC) molecules is well studied, the T-cell recognition of lipids and glycolipids was recently discovered. This recognition is moderated by the CD1 molecules, a class of antigen-presenting molecules with structural similarities to MHC class I molecules. The CD1 proteins have a hydrophobic groove with deep pockets well suited to bind lipids. The polar heads of the lipids protrude from the protein to be recognized by a T-cell (figure 1.3). The human CD1 proteins can be divided into two groups depending on the sequence similarities. Group 1 consists of CD1a, b, and c, while CD1d is the only member of group and CD1e seem to be intermediate or possibly a third group. Both group 1 and 2 CD molecules participate in both the innate or natural, and the adaptive or acquired, immune systems. Group 2 is however reported to be involved in an earlier stage of the immune response.
The different CD1 molecules have somewhat different binding motifs, regarding the lengths and number of the lipid chains. Although all four studied CD1 proteins (a-d) can bind the self glycolipid sulfatide with two lipid chains of about 15-30 carbons each,\textsuperscript{37,38} CD1c also presents single chained glycolipids\textsuperscript{39} and CD1b is capable of binding very long-chained lipids like mycolic acid with up to 80 carbons in one chain\textsuperscript{27} and possibly also lipids with more than two chains.\textsuperscript{25} In addition to their somewhat different binding motifs, the CD1 molecules have diverse trafficking pathways in the cell to survey a larger variety of lipids.\textsuperscript{40}

In recent years, much knowledge has been gained about the role of CD1 molecules in the immune system. Still, our understanding of CD1-restricted immunity is far from complete and considerable efforts are made and needed in this area.

Figure 1.3: Schematic drawing of CD1 presentation of a glycolipid to a T-cell.

- 5 -
Glycoconjugates

1.2 Glycoconjugate synthesis

The presence of several glycoforms and glycolipids of similar, but not identical, composition in biological samples makes isolation of pure, structurally defined glycoconjugates cumbersome. This imposes a need for synthetic glycoconjugates and neoglycoconjugates. The large variability of glycans also makes their synthesis complex and steps need to be taken to ensure the regio- and stereoselectivity of the reactions. In chemical synthesis of saccharides the selectivity is mostly controlled via protecting groups. Masking of the hydroxyls that should not react results in regioselectivity and the correct protecting group can also give stereoselectivity in the glycoside formation. As an example, the two galactose units 1 and 2 will form the β-disaccharide 3 while galactoside 4, where the ester at O-2 in 1 is changed to an ether, will mainly react with 2 to form α-disaccharide 5 (see scheme 1.1). The protecting groups will also have an effect on the reactivity of the saccharides, where electron withdrawing groups deactivate both donors and acceptors, making 4 more reactive than 1. In addition, the stereoselectivity and the reactivity can be influenced by changes in temperature and solvents.

Carbohydrate chemistry has received interest for over 100 years but, in spite of recent advances, glycoconjugate synthesis is still a dif-

Scheme 1.1: Protecting groups in glycosylations
Solid-phase synthesis and biological applications

Difficult and time-consuming process. The preparation of starting materials often require numerous steps and considerable efforts have to be made on optimizations of glycosylation conditions.
2 Solid-phase synthesis

Since the pioneering work of Merrifield, solid-phase organic synthesis has become a subject of great interest. By attaching the starting material onto small plastic beads, two big advantages are obtained. The reactions can be sped up and brought to completion with the use of relatively large excess of reactants and the purification of intermediates is reduced to a simple filtering and washing procedure, where unreacted reactants and reagents can be removed from the resin-bound product. However, the attachment to the beads also makes it difficult to analyze the intermediates and monitor the reactions. For peptides, simple methods have been devised to monitor the progress of the reaction, and their relatively pure reaction chemistry, with few byproducts, diminish the need to analyze resin-bound intermediates. This has made solid-phase synthesis a standard method for peptides as well as oligonucleotides and automated synthesizers have been developed. Commercially available building blocks can thus be combined in a routine fashion, and synthetic peptides and oligonucleotides have been important in answering a wide range of biological questions. For more complicated molecules, like oligosaccharides, the lack of reaction monitoring and analysis techniques has hampered the development. Lately, some techniques for on-resin analysis have been developed and efficient methods for glycosylations have even made automated solid-phase synthesis of carbohydrates possible. However, since different glycosylations behave differently, detailed optimizations of reaction conditions, often in solution, are still required, and solid-phase oligosaccharide synthesis is far from routine (see section 1.2).

2.1 Gel-phase $^{19}$F-NMR spectroscopy

One good method to monitor solid-phase reactions is gel-phase $^{19}$F-NMR spectroscopy. $^{19}$F has several desirable properties as a nucleus for gel-phase NMR spectroscopy. It has a high sensitivity and 100%
natural abundancy, leading to a good signal to noise ratio in a short acquisition time. The large polarizability of the fluorine atom leads to a large distribution of the chemical shifts reducing peak overlapping. Gel-phase NMR spectra suffer from rather broad signals due to highly hindered motion and dipolar couplings and the large chemical shift dispersion is essential to receive detailed information from the spectra. In addition, fluorine is not present in solid-phase resins for supported synthesis and hence no interfering signals will arrive from the resin.

These properties have led to the use of $^{19}$F-NMR spectroscopy to monitor the progress of solid-phase reactions with the beginning in 1980. Later it was found that TentaGel, with its flexible poly(ethylene glycol) spacers, gives rise to spectra of high resolution. This meant that also small chemical shift changes can be monitored, and byproducts such as diastereomers can be quantified. The influence of flexible spacers on the spectra quality was recently confirmed by a thorough study of the influence of resins and solvents in gel-phase $^{19}$F-NMR spectroscopy.

To be able to use $^{19}$F-NMR spectroscopy, fluorine atoms must be present in the reactions. This can be accomplished with a fluorinated linker, protecting groups, reactants or a combination of the three. The strategy based on fluorinated protecting groups is especially well suited for carbohydrate synthesis where protecting groups are essential (see section 1.2). Fluorinated versions of several common protecting groups are commercially available and their use in solid-phase carbohydrate synthesis in combination with fluorinated linkers have been demonstrated.

In addition, fluorinated molecules have received an increased interest in medicinal chemistry during the last two decades. The insertion of a fluorine atom into a molecule influences its reactivity, metabolic stability, physiochemical and conformational properties, and protein binding. The presence of a fluorine atom also permits $^{19}$F-NMR spectroscopy to be used to study binding interactions of small molecules to proteins. The combined use of $^{19}$F-NMR spectroscopy in solid-phase synthesis of libraries of fluorinated molecules and screening has however not been well explored.
2.2 Linkers

Solid-phase synthesis typically requires the use of a linker through which the starting material can be attached to the resin and from which the product later can be cleaved. Such a linker needs to be stable during the reaction conditions and cleavable under conditions that does not affect the product. A large variety of linkers have been devised throughout the years since Merrifield's chloromethylated nitro-benzyl linker (for examples, see figure 2.1). Essentially all protecting groups can be adopted to be a linker, and linkers exist that are cleaved by for example acid, base, light, reduction, and oxidation, some during very mild conditions while others require a more harsh treatment. Among the acid sensitive linkers, it is common to make use of the resonance stabilization of benzylic cations. The original papers by Merrifield discussed this point and also described how to increase the acid stability through electron withdrawing groups. Wang later took advantage of the opposite when he created a more acid labile linker by including a phenol ether in the Merrifield linker (see figure 2.1, compound 7). This linker, later named the Wang linker, has become the standard linker for the synthesis of peptide acids and is also used in a large variety of other solid supported organic reactions. Addition of more electron donating or resonance stabilizing groups can in a similar way further increase the acid lability, as in the SASRIN linker (figure 2.1).

![Figure 2.1: Three linkers used to immobilize carboxylic acids and examples of their respective cleavage methods.](image-url)

Incorporation of a fluorine atom into the linker makes it useful for $^{19}$F-NMR spectroscopy. It can be used as an internal standard and is often sensitive enough to give direct information about subse-
sequent reactions, such as linker loading and product cleavage. Fluori-
nated linkers described in the literature include fluorinated versions
of several commonly used linkers that can be cleaved with acid,\textsuperscript{71–74}
nucleophiles,\textsuperscript{57–60, 71–76} oxidation,\textsuperscript{61} and metal assisted chemistry,\textsuperscript{77, 78}
as well as a linker that relies on more product specific cleavage condi-
tions\textsuperscript{79} (for examples, see figure 2.2). Since fluorine is electron with-

![Diagram of fluorinated linkers](image)

**Figure 2.2:** Three fluorinated linkers used to immobilize carboxylic acids and examples of their respective cleavage methods.\textsuperscript{57, 72, 73}

drawing, the fluorinated versions of the Wang linker require more
harsh conditions for acidic cleavage than the original.\textsuperscript{72}

Gel-phase $^{19}\text{F}$-NMR spectroscopy has been demonstrated as a use-
ful technique to monitor solid-phase synthesis on fluorinated linkers.
Several such linkers, with a broad panel of cleavage conditions, have
been devised, derivatized, and used, e.g. in the synthesis of small
molecule libraries and oligosaccharides.
Nucleotide microarrays are important tools in the post-genomic process to analyze gene transcription. On one glass plate, the transcripts from the whole human genome can be analyzed from a small sample.\textsuperscript{80} In addition, protein microarrays are increasingly used to study protein interactions with other proteins, nucleic acids and small molecules.\textsuperscript{81} The large structural variations in oligosaccharides and glycoconjugates (see chapter 1), together with the relative difficulty to access large amounts of pure glycans make carbohydrate microarrays a promising technique in glycobiology.\textsuperscript{82}

Increasing efforts are made to construct carbohydrate arrays and both commercial\textsuperscript{83–85} and non-commercial\textsuperscript{86} arrays are now available. In general, a carbohydrate array is constructed by attaching oligosaccharides or glycoconjugates in an ordered fashion to a solid surface (see figure 3.1). To this carbohydrate array another substance, such as a lectin or an antibody, can be added and its binding can be measured, usually by a fluorescent marker or a colorimetric method. By indirect measurement, the effect of enzymes or small molecule inhibition of a binding can also be studied. In addition, viable cells or bacteria can bind to the array for analysis.\textsuperscript{87}

To demonstrate the usefulness of carbohydrate arrays several groups have prepared arrays with a multitude of attachment techniques. In 1992, Shao biotinylated glycans derived from ovalbumin and added them to microtiter plates, precoated with streptavidin.\textsuperscript{88} The formed glycan array was used to study binding specificities of lectins. Streptavidin coated microtiter plates and biotin conjugated carbohydrates have later for example been used to investigate the influence of the linker on carbohydrate-protein interactions\textsuperscript{89} and the specificity of cell receptors related to HIV infections.\textsuperscript{86} In addition, other non-covalent linkages have been demonstrated. Neoglycolipids, glycolipids, and polysaccharides have been adsorbed on microtiter plates\textsuperscript{90,91} nitrocellulose,\textsuperscript{92–94} and modified plastic\textsuperscript{95} and glass slides.\textsuperscript{96} A variation with synthetic
Figure 3.1: Carbohydrate array construction and applications.
fluorous carbohydrates on modified glass slides was recently reported.\textsuperscript{97} As an interesting extension of the non-covalent methods, several reports have been made on the use of simple chemistry to create neoglycolipids directly in microtiter plates from modified saccharides and reactive lipids coated in the wells.\textsuperscript{98–101}

The exclusivity of the glycans makes reusable arrays desirable. To accomplish that, and to increase the general stability of the array, covalent attachment of the carbohydrate is an interesting option. In addition, almost all covalent methods result in a site-specific binding of the glycan to the solid surface.\textsuperscript{87} This gives a higher control over the actual saccharide epitope that is presented by the array and a higher signal to noise ratio. An increasing proportion of the prepared carbohydrate arrays uses a covalent attachment to microtiter plates,\textsuperscript{102–106} modified glass slides,\textsuperscript{83, 84, 107–121} or fiber optic microspheres.\textsuperscript{122} Among the bond forming reactions are amide formations, reductive aminations, Diels-Alder reactions, 1,3-dipolar cycloadditions and more. Most of the reactions rely on specific groups present on the glycan, imposing the need of synthetic compounds or derivatization of natural carbohydrates, often with destruction of the monosaccharide at the reducing end. An interesting exception is the use of hydrazide or aminoxy derivatized glass slides to form glycosylamines directly from reducing saccharides.\textsuperscript{121}

In the construction of carbohydrate arrays, the glycan density is of importance. It should be high enough to allow multivalent protein-carbohydrate interactions. It is however also reported that a too high density can diminish binding, probably due to steric interactions.\textsuperscript{108} Low non-specific binding of proteins to the array is also of great significance to decrease the noise in the measurements.\textsuperscript{123}

Even though carbohydrate arrays have the possibility to display up to tens of thousands of glycans\textsuperscript{110} most reports so far use considerably fewer different glycans to demonstrate the utility of the array format and the immobilization method. Ligand profiling of several proteins capable of binding the HIV protein gp120,\textsuperscript{86, 106, 114} as well as other proteins involved in the immune system,\textsuperscript{94} has been reported by several groups. One interesting report found that the ligand specificity of a protein changed with the ligand density.\textsuperscript{107} At low surface density the lectin from Bauhinia purpurea bound preferably to a thiol conjugate of the disaccharide β-galactosyl-(1-3)-β-N-acetyl-2-amino-2-deoxy-galactose (12 figure 3.2) while at high densi-
ties, it preferred α-galactosyl-(1-3)-α-N-isovaleroyl-2-amino-2-deoxy-glucose (13). Angeloni et al. made a glycoprofiling of model cell extracts of Caco-2 intestinal cells that were arrayed and profiled with a panel of lectins and could show changes in the glycan expression depending on differentiation and treatment with 3′-sialyllactose.\textsuperscript{119} Both prokaryotic\textsuperscript{116} and eukaryotic\textsuperscript{83} cells have been shown to adhere to carbohydrate arrays. Detection and profiling of bacterial carbohydrate binding could be useful as a mean to diagnose diseases.\textsuperscript{116} The interaction between carbohydrate arrays and enzymes has been used to, among others, study the substrate specificity of β-1,4-galactosyltransferase,\textsuperscript{109} synthesize a tetrasaccharide on an array\textsuperscript{117} and screen for inhibitors to an α-1,3-fucosyltransferase.\textsuperscript{99} Also inhibitors to lectin carbohydrate interactions have been studied, e.g. by Houseman and Mrksich.\textsuperscript{109} A few studies are reported with more than 100 carbohydrates and in one as much as 200 different glycans were arrayed and used for ligand profiling of nine proteins as well as human serum and intact virus.\textsuperscript{115}

Carbohydrate arrays have great potential answering questions in glycobiology. In spite of the public availability of such arrays, they have however not yet reached their full potential and the applications are relatively few. Streamlined synthesis of both natural and non-natural glycoconjugates and array preparation will improve the utility with arrays custom made to answer a specific question.
4 Aims of the thesis

Initially, we set out to study the CD1 binding of glycolipids, and their use in immunizations. The first goal was to establish synthetic methods to use for preparation of glycolipid libraries and a method to measure glycolipid binding to CD1d molecules on viable mouse cells. With these methods at hand, the next steps would be to prepare the libraries, and develop cell free methods to measure CD1d binding. Immunizations using glycolipids that bind well to CD1 and carry different immunologically relevant carbohydrates would be an interesting continuation.

However, after the first paper was finished, we changed focus towards preparation of glycoconjugate arrays to obtain a streamlined process including synthesis, reaction monitoring, purification, and array manufacturing. The goal was to develop synthetic methods and equipment that allow synthesis of libraries of glycolipid analogues. The preparation and use of glycoconjugate arrays formed from these glycolipid analogues were also to be demonstrated.
5 Glycolipid synthesis

5.1 Synthesis of $\alpha$-N-acetyl-galactosamine glycolipids (paper I)

In order to study glycolipid binding to CD1d molecules on viable murine cells (see chapter 6), glycolipids with specific carbohydrates were needed. The glycan should not be normally expressed on the cells in question, to allow detection with antibodies, and should preferably be simple and easy to work with. In this light, we chose to work with $\alpha$-N-acetyl-galactosamine ($\alpha$-GalNAc, see figure 5.1). $\alpha$-GalNAc is a common monosaccharide, present e.g. in mucins, where it is linked to serine or threonine. However, apart from the blood group determinant A, terminal $\alpha$-GalNAcs are rare in healthy humans. Some cancers are associated with an increase in the levels of terminal $\alpha$-GalNAc, and terminal $\alpha$-GalNAc connected to a serine or threonine is referred to as the Tn antigen. To confer

![Figure 5.1: $\alpha$-GalNAc containing glycolipids.](image_url)
CD1d-binding, the commercial lipids 2-(n-hexadecyl)-stearic acid and palmitic acid were chosen to be attached to the α-GalNAc through a spacer (see figure 5.1). 2-(n-hexadecyl)-stearic acid is an analogue to the mycolic acids (see figure 1.2), that are known to bind to CD1 proteins. Two different lengths of the spacer were used to ensure that the carbohydrate protruded enough from the CD1d molecules to be recognized by monoclonal antibodies.

In 1893, Emil Fischer reported a way to synthesize glycosides of alcohols by boiling the sugar, together with an acid, in the alcohol that works as both acceptor and solvent.\(^{124}\) This process leads to the thermodynamically most stable glycoside, most often the α-pyranose. The protocol works very well for simple, liquid, low-boiling, and inexpensive alcohols. In the case of α-GalNAc, more complex alcohols are frequently glycosylated with 2-azido-2-deoxy-galactose donors, giving the appropriate α-glycoside.\(^{125}\) Later, the azide is reduced to an amine, which in turn is acetylated. An example of a synthetic route is shown in scheme 5.1.

\[ \text{Scheme 5.1: Example of a synthesis of α-GalNAc glycosides.} \]

To simplify this synthesis, we aimed to use a modified Fischer synthesis, using an aprotic solvent and Fmoc-protected amino-alcohols in modest excess. It was found that refluxing GalNAc with five equivalents of \(N\)-Fmoc-2-amino-ethanol and four equivalents of the Lewis acid boron trifluoride etherate in THF for 19 hours gave the wanted α-glycoside 22 in 34% yield (see scheme 5.2). The reaction could also be performed with four equivalents \(N\)-Fmoc-6-amino-hexanol...
Solid-phase synthesis and biological applications

and 10% p-toluene sulfonic acid in THF/toluene, resulting in somewhat lower yields of 23. Experiments were also made using microwave assisted heating of the reactions and a small experimental design resulted in conditions that looked promising according to analytical LC/MS on crude products. However, the isolated yields were only ~15% and a hard, glassy solid was formed around the walls of the reactors. Most probably the starting material decomposed or polymerized under the high temperatures and acidic conditions. This method was therefore abandoned. However, a recent report demonstrates successful microwave assisted Fischer glycosylations using simple alcohols.126

Scheme 5.2: Synthesis of GalNAc-bearing lipids. a) Fmoc-ω-aminoalcohol, Lewis acid, aprotic solvent. b) 2-chlorotrityl chloride polystyrene resin, DMAP, CH2Cl2. c) i. Piperidine, 20% in DMF. ii. Palmitic acid or 2-(n-hexadecyl)-stearic acid, DIC, HOAt. d) 5% TFA in CH2Cl2.

The formed α-GalNAc glycosides were then to be attached to the
Glycoconjugates

fatty acids 2-(n-hexadecyl)-stearic acid and palmitic acid. Acylation of the deprotected amines in solution turned out to be very difficult, likely due to the different solubilities of the sugar and the acid. Instead the GalNAc derivatives 22 and 23 were attached to a polystyrene resin, derivatized with 2-chlorotrityl chloride. The resin, glycoside and $N,N$-dimethylaminopyridine (DMAP) were dissolved/suspended in CH$_2$Cl$_2$ and the mixture was stirred slowly at reflux for 24 hours. This yielded resins 24 and 25 with approximately 0.2 mmol glycoside per gram according to Fmoc analysis. The glycoside was most likely attached to the resin through the primary 6-hydroxyl group. Deprotection of the amine using piperidine (20% in DMF) and acylation with palmitic acid or 2-(n-hexadecyl)-stearic acid yielded the resin-bound glycolipids 26 and 27 in a straightforward manner. The formed glycolipids were cleaved from the solid support using 5% TFA in CH$_2$Cl$_2$ and, after chromatography, the pure lipids 14a, 14b, and 15a were obtained in 66%, 75%, and 30% yield respectively, based on Fmoc determination on resins 24 and 25.

Three glycolipids containing the monosaccharide $\alpha$-N-acetylgalactosamine were prepared in a short synthesis. The synthetic route, based on the immobilized spacer glycosides 24 and 25, is well suited for library synthesis where the spacer and lipid parts can be varied more extensively.

5.2 Solid-phase glycosylations in glycolipid synthesis (paper II)

Serine-based neoglycolipids (see 29-31, figure 5.2) are analogues of glycosphingolipids like 28 and have been used in several studies.$^{127-150}$ Glycosphingolipids are normally present in cell membranes (see section 1.1) where the lipid tails are inserted into the membrane and the polar head is exposed at the surface for recognition events. The serine and its adjacent amide bonds mimic the polar head of the ceramide and the serine-based neoglycolipids have been shown to mimic the sphingolipid e.g. in inhibition of HIV invasion,$^{140,141,143}$ ceramide glycanase transglycosylations,$^{132}$ and activation of $\beta$-glucocerebrosidase.$^{146}$ There are also indications that a serine-based glycolipid have similar bioactivities, although to a lesser degree, as the CD1d-binding $\alpha$-galactosylceramide KR7000.$^{149}$
Solid-phase synthesis and biological applications

![Figure 5.2: A β-galactosyl ceramide (28) and serine-based analogues 29-31.](image)

To explore the possibilities for solid-phase glycosylations of serine based ceramide analogues and preparation of glycoconjugate arrays of the formed neoglycolipids three different glycolipids were synthesized (29-31, scheme 5.3). Linker 32 was attached on ArgoGel®-NH₂ with N,N'-diisopropyl carbodiimide (DIC) and 1-hydroxy-7-azabenzotriazole (HOAt) in DMF. The free alcohol of the linker resin 33 was esterified with an Fmoc-ω-amino acid with the help of 1-(2-mesitylenesulfonyl)-3-nitro-1H-1,2,4-triazole (MSNT), and N-methyl imidazole (MeIm). Continued amide formations with Fmoc-serine and a single stranded fatty acid using DIC and 1-hydroxybenzotriazole (HOBr) gave the solid supported lipids 36 and 37. The amide formations were monitored with bromophenol blue to ensure near quantitative yields and the esterification was confirmed to be quantitative by gel-phase ¹⁹F-NMR spectroscopy.

Investigations on the solid-phase glycosylation of the lipids 36 and 37 were made with both α- and β-selective glycosylations using the previously described thioglycoside donors 41 and 42. The donor 41 has a benzoate ester at the 2-hydroxyl group, enabling neighboring group participation (see section 1.2) which in turn leads to good β-selectivity. The resins 36 and 37 were thus subjected to four equivalents of 41 together with N-iodosuccinimide (NIS) and
Scheme 5.3: Synthesis of serine-based glycolipids. a) ArgoGel®-NH$_2$ resin, DIC, HOAt. b) Fmoc-$\omega$-amino acid, MSNT, MeIm. c) i. Piperidine 20% in DMF. ii. Fmoc-Ser-OH, DIC, HOBt. iii. Piperidine 20% in DMF. iv. Fatty acid, DIC, HOBt. d) 2 x 4 eq. 41 or 5 eq. 42, NIS, TfOH. e) i. TFA/water 9:1, 60 °C. ii. NaOMe or i. LiOH. ii. H$_2$ (g), Pd/C.
Solid-phase synthesis and biological applications

<table>
<thead>
<tr>
<th>Entry</th>
<th>Solvent</th>
<th>Temperature (°C)</th>
<th>Yield (%)</th>
<th>α/β-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CH₂Cl₂</td>
<td>-40</td>
<td>53</td>
<td>2:1</td>
</tr>
<tr>
<td>2</td>
<td>CH₂Cl₂</td>
<td>room temperature</td>
<td>quant.</td>
<td>2:1</td>
</tr>
<tr>
<td>3</td>
<td>CH₂Cl₂/THF</td>
<td>-40</td>
<td>45</td>
<td>2:1</td>
</tr>
<tr>
<td>4</td>
<td>CH₂Cl₂/THF</td>
<td>room temperature</td>
<td>quant.</td>
<td>4:1</td>
</tr>
</tbody>
</table>

Table 5.1: Investigation on α-selective glycosylations. Reactions were performed with 5 eq. of donor 42.

trifluoromethane sulfonic acid (TfOH). The mixture was shaken at room temperature, protected from light, for 3.5 hours to give the glycolipids 38 and 39 in ~50% yield. Repeating the reaction once increased the yield to ~75%. The yields were determined from ¹⁹F-NMR spectra that also showed β/α-selectivities greater than 5:1. The donor 42 lacks the participating group and should react to form mostly α-glycosides (see section 1.2). To improve the selectivity, α-glycosylations are often performed at low temperatures. Initially, this strategy was used for the formation of 40 using five equivalents of donor 42, NIS, and TfOH. The reaction was performed in CH₂Cl₂ at -40 °C with a yield of ~50% that increased to ~85% upon repetition. The selectivity was rather modest (2:1) with a third, unknown product at equimolar ratios to the β-anomer. Solid-phase reactions are greatly simplified if performed at room temperature and therefore the possibilities for α-selective glycosylations at room temperature were investigated (see table 5.1). Repeating the reaction at room temperature made the yield quantitative after one single glycosylation but the selectivity remained the same. Ether solvents are known to have an α-directing effect and running the reaction at room temperature in CH₂Cl₂/THF (1:1) increased the selectivity to 4:1. The third, unknown product was still formed in equimolar ratios to the β-anomer. At low temperature the effect of the ether solvent could not be detected. This investigation of different reaction conditions could be performed quickly and easily, thanks to the fluorinated linker and protecting groups. The higher yields of the α-glycolipid 40, compared to 38 and 39, are most likely due to the higher reactivity of the more electron-rich donor 42 compared to 41.

The β-glycolipids were cleaved from the resin using trifluoroacetic
acid (TFA) and water (9:1) at 60 °C. Debenzoylation produced the
glycolipids 29 and 30 in 11% and 26% total yields respectively, based
on the initial loading capacity of the ArgoGel® resin. One difference
between the on-resin and the cleaved yields could be the presence of
low reactivity amino groups on the resin, lowering the effective load-
ing. The yields are comparable to those obtained in other solid-phase
glycosylations of serine.\textsuperscript{58} Subjecting the resin bound glycolipid 40 to
the same cleavage conditions resulted in a complex product mixture,
where the benzylidene group was cleaved and most likely the ben-
zyls had migrated. Instead, the glycolipid 31 was formed from basic
ester hydrolysis with LiOH, followed by deprotection with catalytic
hydrogenation, in a total yield of 22%. Also the β-anomer of the
protected glycolipid was purified and its structure was confirmed by
NMR spectroscopy. Unfortunately, the unknown product could not
be isolated from the cleavage mixture and remain elusive.

Three different serine-based glycolipids, analogues to glycosphin-
golipids, were synthesized on solid-phase. Fluorinated protecting
groups and a fluorinated linker made reaction monitoring and on-
resin analysis of diastereomeric ratios with \textsuperscript{19}F-NMR spectroscopy
possible. Thanks to the on-resin analysis, the α-selectivity of one re-
action could easily be improved, without the need for solution-phase
optimizations or cleavage from the resin.
The investigations of lipid binding to CD1 molecules (see section 1.1.1) has so far mostly been based on T-cell recognition of glycolipid-CD1 complexes. In those methods, the obtained signal depends on both the CD1 binding of the lipid and the T-cell response, leading to a complex interpretation of the data. In addition they are badly suited to measure non-specific binding of glycolipids in the cell membrane. Some studies have been made on lipid binding to soluble or surface-bound CD1 molecules using surface plasmon resonance, however predominantly with heavily modified lipids.151–153

As an alternative method to measure glycolipid binding to CD1 molecules on viable cells the use of carbohydrate specific antibodies, selective for the glycolipid, was imagined. Five different glycolipids, 14a, 14b, and 15a, prepared as described in section 5.1, and the previously synthesized 43154 and 44,155 carrying three different carbohydrates were used to develop this method (see figure 6.1). The study was performed with CD1d-bearing mouse cells of different origins. The surfaces of these cells normally lack all the carbohydrate structures used in the glycolipids, enabling the specific recognition with monoclonal antibodies (data not shown). Viable cells were incubated with the glycolipids in PBS buffer containing 5% DMSO and, after washing, with monoclonal antibodies specific for the carbohydrate or CD1d. Secondary antibodies labeled with fluorescein isothiocyanate (FITC) and fluorescent-activated cell sorting (FACS) were used to quantify the monoclonal antibodies. A dose-response relationship for the glycolipid loading of mouse EL-4 cells, carrying CD1d proteins, was established using the galabiose lipid 43 (see figure 6.2). The CD1d selectivity of the binding was investigated using all five glycolipids and wild-type as well as β2-microglobulin-knockout EL-4 cells. The knockout cells lack all β2-microglobulin-associated proteins, including CD1d. Significantly higher responses were obtained for the wild-type cells with all glycolipids (figure 6.3) indicating a CD1d-restricted loading. A monoclonal CD1d-specific antibody con-
Figure 6.1: Glycolipids used in CD1d loading experiments.
Figure 6.2: Loading of glycolipid 43 to EL-4 cells in a dose-dependent manner. The bars represent the mean fluorescence intensity (MFI) for EL-4 cells incubated with glycolipid 43 and a monoclonal antibody. For more information, see experimental section of paper I. Figure adapted from paper I.

Confirmed that the β2-microglobulin-knockout EL-4 cells indeed lacked CD1d. The ceramide-based GM3-lactam 44 showed a higher unspecific binding, probably due to the ability of the ceramide lipid to bind directly in the cell membrane, where glycosphingolipids are normally found (see section 1.1). The single stranded glycolipid 14b showed high loading, even though the CD1d binding motif is mostly double-chained.24 There have been reports, however, on CD1d binding and T-cell activation by single-chained glycolipids.156 On the contrary, an analogue of the lipids 14, where the fatty acid is replaced with an acetyl group, did not bind (data not shown). Importantly, the experiment showed that the carbohydrates of the glycolipids are exposed to the antibodies, and not buried in the CD1 molecule. In addition, the glycolipids 14a and 15a gave similar responses in spite of the difference in length of the spacer between the lipid and carbohydrate.

The fact that the β2-microglobulin-knockout cells lack all β2-m associated proteins makes the CD1-specificity of the binding uncertain. To further look into this specificity, EL-4 cells were preincubated with different concentrations of the known CD1d ligand α-galactosyl ceramide (α-GalCer, 45, figure 6.4).157 Subsequent incubation with glycolipid 14a showed that the binding of the latter was inhibited in a dose-dependent manner (see figure 6.5). This indicate that α-GalCer and 14a bind to the cells in the same way, namely by CD1d-specific interactions.
Glycoconjugates

**Figure 6.3:** Selective binding of glycolipids to CD1d positive EL-4 cells. The bars represent the mean fluorescence intensity (MFI) for EL-4 (black bars) and β2m −/− EL-4 cells (grey bars) incubated with a glycolipid and a monoclonal antibody. For more information, see experimental section of paper I. Figure adapted from paper I.

**Figure 6.4:** The known CD1d-binding glycolipid KRN7000, an α-galactosylceramide (α-GalCer)

Dendritic cells (DCs) are specialized antigen-presenting molecules, and preloaded with a glycolipid antigen they are capable of inducing a strong immune response upon injection into mice.\textsuperscript{158} DCs from B6 and CD1d-negative B6 mice were loaded with glycolipid 14a as before. Again, a dose-dependent loading, specific for
Figure 6.5: $\alpha$-GalCer (45) inhibits binding of glycolipid 14a to CD1d positive EL-4 cells. The bars represent the mean fluorescence intensity (MFI) for EL-4 cells incubated with different concentrations of $\alpha$-GalCer 45 followed by glycolipid 14a and a monoclonal antibody. For more information, see experimental section of paper I. Figure adapted from paper I.

The wild-type DCs, was noted (figure 6.6). Since CD1d-negative B6 mice are true CD1d-knockouts, the possible interaction with other $\beta_2$-microglobulin-associated proteins was ruled out. Also thymocytes from B6 mice could be loaded with the $\alpha$-GalNAc containing lipids 14a, b, and 15a (data not shown).

Figure 6.6: Glycolipid 14a binds selectively to CD1d positive DCs in a dose-dependent manner. The bars represent the mean fluorescence intensity (MFI) for DCs from B6 (black bars) and CD1d $-/-$ B6 (grey bars) mice incubated with glycolipid 14a and a monoclonal antibody. For more information, see experimental section of paper I. Figure adapted from paper I.

Using synthetic glycolipids and monoclonal antibodies, CD1d specific and dose-dependent glycolipid loading of viable cells was confirmed. This FACS based method is rapid and estimates CD1 binding
independent of T-cell activation. The method is therefore suitable in screening for lipids capable of binding CD1. The glycolipids used in the study all have immunologically relevant saccharides and can be used to investigate the capacity of these conjugates to evoke a CD1d restricted immune response. In addition, three new CD1d-binding lipid motives were discovered.
To start exploring the possibilities of glycolipid microarrays, glycolipids prepared through solid-phase synthesis were covalently bound to amine groups in microtiter plates. The water-soluble neoglycolipids 29 and 31 (figure 7.1, see also section 5.2), with terminal carboxylic acid functionalities, and commercial microtiter plates with secondary amines tethered to the plastic surface (CovaLink™) were used in the study. The glycolipids were serially diluted in the wells and coupled to the amines with an aqueous solution of N-hydroxysuccinimide (NHS) and 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC). After extensive washing and blocking of the wells with bovine serum albumin (BSA) the carbohydrates were detected with labeled lectins. Biotin-labeled agglutinin from *Ricinus communis* (RCA<sub>120</sub>) has affinity for α- and β-Gal and horseradish peroxidase (HRP)-conjugated lectin from *Bandeiraea simplicifolia* (BS-1) for
α-Gal and α-GalNAc. To the RCA<sub>120</sub> lectin, HRP-conjugated avidin was added and both lectins were then detected through HRP-catalysed oxidation of o-phenylenediamine and absorbance measurements of the resulting 2,3-diaminophenazine. The RCA<sub>120</sub> recognized both glycolipids 29 and 31 with similar response curves (see figure 7.2), even though its affinity for β-Gal is four times higher than for α-Gal. This indicates that the measurements of

![Figure 7.2: Binding of lectin RCA<sub>120</sub> to a carbohydrate array with different concentrations of 29 (black circles) and 31 (grey squares). The graph shows the mean and standard deviations from duplicate columns, with the lipid concentration in the coupling solutions on the x-axis. The mean from eight blank wells were subtracted from all points. For more information, see the experimental section of paper II. Figure adapted from paper II.](image)

lectin binding to the carbohydrate array are qualitative rather than quantitative. The α-selective BS-1 only recognized the α-galactosyl lipid 31 as expected (see figure 7.3). The overall signal for BS-1 is lower, probably due to the signal-amplifying properties of the biotin-avidin construct for RCA<sub>120</sub>. In a previous study, using the lactose derivative of 29 and 31 and a variant of the microtiter plates with primary amines, the lactose-binding lectin from *Erythrina coralloendron* also bound in a similar fashion to the lectins in the present study.<sup>127</sup> Importantly, the plates could be regenerated through washing with a 10% SDS solution, followed by reblocking with BSA, as shown in figure 7.4. Since the synthetic glycolipids are precious, and the preparation of the plates rather time-consuming, the possibility to regenerate plates is valuable.

The neoglycolipids 29 and 31, prepared by solid-phase synthesis, were covalently attached to microtiter plates to form arrays of neoglycolipids of varying density. The protein-binding properties of the
arrays were tested with two different lectins, RCA120, specific for α- and β-galactose, and BS-1, specific for α-galactose and α-GalNAc. Both lectins bound to the array of neo-glycolipids in agreement with their respective specificity.

**Figure 7.3:** Binding of lectin BS-1 to a carbohydrate array with different concentrations of 29 (black circles) and 31 (grey squares). The graph shows the mean and standard deviations from duplicate columns, with the lipid concentration in the coupling solutions on the x-axis. The mean from eight blank wells were subtracted from all points. For more information, see the experimental section of paper II. Figure adapted from paper II.

**Figure 7.4:** Regeneration of CovaLink™ wells containing 29. a) Detection with RCA120 before regeneration. b) Detection without lectin after stripping with 10% SDS. c) Renewed detection with RCA120. The bars show the mean and standard deviations from duplicate wells, with the mean from two blank wells subtracted from all points. For more information, see the experimental section of paper II. Figure adapted from paper II.
8 Design, synthesis, and evaluation of fluorinated linkers

For solid-phase reaction monitoring with gel-phase $^{19}\text{F-NMR}$ spectroscopy, fluorinated linkers are crucial. So far, relatively few fluorinated linkers have been described in the literature (see section 2.2) and no such linkers are commercially available. This chapter describes the synthesis and application of two novel linkers for immobilization of carboxylic acids. Both linkers can be cleaved with either acids of varying strength or nucleophiles, e.g. hydroxide.

8.1 Synthesis of a new fluorinated monoalkoxy linker (paper II)

The linker 46 (see figure 8.1) is a fluorinated analogue to the classic Wang linker and has been applied in solid-phase synthesis monitored with gel-phase $^{19}\text{F-NMR}$ spectroscopy.\textsuperscript{71,72} When the starting material for the synthesis of that linker no longer was commercially available, a new synthetic route towards such a linker was needed. The availability of 3-fluoro-4-hydroxy-benzoic acid (47) prompted synthesis of (2-fluoro-4-(hydroxymethyl)-phenoxy)-acetic acid (32) in three steps (see scheme 8.1). Reduction of the acid 47 using trimethyl borate and borane dimethyl sulfide complex (BH\textsubscript{3},DMS) proceeded well, in nearly quantitative yields according to thin-layer
chromatography (TLC). The crude alcohol 48 was alkylated with ethyl bromoacetate and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU). After chromatographic purification the ester 49 was obtained in 57% yield over two steps. Finally, the ester was hydrolyzed with LiOH to give the linker 32 in 90% yield.

Scheme 8.1: Synthesis of a fluorinated monoalkoxy benzyl alcohol linker. a) B(OMe)₃, BH₃-DMS, THF. b) ethyl bromoacetate, DBU, MeCN, reflux. c) LiOH, THF/MeOH/water 3:1:1.

The linker 32 was prepared in 51% total yield in a three step synthesis, amenable to gram scale preparation. The formed linker is stable to glycosylating conditions and can be cleaved under both acidic (TFA/water 9:1, 60 °C) and basic (LiOH, 20–33 mM, THF/water 1:2) conditions as described in section 5.2.

8.2 Design, synthesis, and evaluation of a fluorinated dialkoxy linker (paper III)

The acidic conditions used to cleave a carboxylic acid from the linker 32 are rather harsh, as seen in section 5.2, and can pose problems when the product is not sufficiently stable. On the other hand, also the basic conditions can be troublesome, e.g. serine glycosides can undergo β-elimination. Therefore, a fluorinated linker from which carboxylic acids can be cleaved using milder
Solid-phase synthesis and biological applications

acid is desired. Addition of a second alkoxy-group to the linker 32 should make it more acid labile due to stabilization of the benzylic cation (see section 2.2). A retrosynthetic analysis indicated that (6-fluoro-4-(hydroxymethyl)-3-methoxy-phenoxy)-acetic acid (58) could be synthesized from a trifluorobenzoic acid derivative with nucleophilic aromatic substitutions as key steps. Both tert-butyl 2,4,5-trifluorobenzoate\textsuperscript{161,162} (50, scheme 8.2) and 2,4,5-trifluorobenzonitrile\textsuperscript{161,163–167} (51, scheme 8.3) have been reported as electrophiles in such reactions. Based on this, a strategy using the benzoate 50 and methyl glycolate was initially explored (see scheme 8.2). Unfortunately, the nucleophilic substitutions, especially with methoxide, were low yielding due to trans esterifications and decomposition. Instead, the use of benzyl alcohol as nucleophile for introduction of the oxygen at the para-position was investigated. Again, trans esterifications caused problems with the benzoate 50. Benzonitrile 51 on the other hand, gave good yields in the nucleophilic aromatic substitutions, using potassium tert-butoxide (t-BuOK) as base, and was therefore used in the continued synthesis (scheme 8.3). The dialkoxy benzonitrile 53 was thus obtained in 67% yield over two steps. Hydrolysis of the nitrile in refluxing NaOH in ethanol/water 4:3 and catalytic hydrogenolysis of the benzylether furnished the benzoic acid 55 in

\[
\begin{align*}
\text{Scheme 8.2: Attempted synthesis of linker 58 from benzoate 50 and methyl glycolate. a) Methyl glycolate, } & \text{t-BuOK, THF, 0 °C. b) MeOH, } \text{t-BuOK, THF, 0 °C - room temperature.}
\end{align*}
\]
Scheme 8.3: Synthesis of linker 58 from benzonitrile 51. a) BnOH, \( t \)-BuOK, THF, -78 °C – room temperature. b) MeOH, \( t \)-BuOK, THF, -50 °C – 0 °C. c) NaOH (aq.), EtOH, reflux. d) \( \text{H}_2 \) (g, 1 atm.), Pd/C, AcOH. e) \( \text{B(OH)}_3 \), \( \text{BH}_3\cdot\text{DMS}, \) THF. f) ethyl bromoacetate, DBU, MeCN, reflux. g) \( \text{LiOH}, \) THF/MeOH/water 3:1:1.

53% total yield from benzonitrile 51. The acid 55 is an analogue to 47, the starting material in the synthesis of the monoalkoxy linker 32 (see section 8.1). Unfortunately, the extra methoxy group disturbed the continued synthesis. The reduction to the benzyl alcohol could be performed in good yields (~90%) but the product was very sensitive, probably towards oxidation, and was only briefly purified and characterized before continuation of the synthesis. Alkylation of the phenolic oxygen with ethyl bromoacetate and DBU in the same conditions as for 48 surprisingly yielded the dimer 59 as the main product (see figure 8.2). Since dimer 59 and its monomer 57 have very similar spectroscopic properties, and both gave the benzylic cations as main peaks in LC/MS analysis,
Solid-phase synthesis and biological applications

Figure 8.2: The dimer formed in alkylation of 56.

x-ray crystallography was used to verify the structures (data not shown). Reduction of the reaction time diminished the formation of the dimer and the monomer 57 could be isolated in 66% yield over two steps. Finally, the linker 58 was received in 81% yield by basic ester hydrolysis. In total, 58 was synthesized in seven steps from trifluorobenzonitrile 51 with an overall yield of 29%.

To test the cleavage conditions for linker 58, a benzoate-terminated dipeptide was synthesized and cleaved (scheme 8.4). The linker 58 and its analogue 32 were attached to TentaGel HL-NH₂ through amide bonds formed with DIC and HOBr. Capping of any remaining amines with acetic anhydride and pyridine followed by deesterification with NaOMe in MeOH gave the linker-resins 60 and 61. To these, Fmoc-glycine (Fmoc-Gly-OH) was esterified using 1-(2-mesitylenesulfonyl)-3-nitro-1H-1,2,4-triazole (MSNT) and N-methyl imidazole (MeIm) in quantitative yields according to ¹⁹F-NMR spectroscopy. Continued amide formations using DIC and HOBr attached Fmoc-para-fluoro-phenylalanine (Fmoc-p-F-Phe-OH) and, after deprotection, 2,4-difluorobenzoic acid. All couplings proceeded quantitatively according to ¹⁹F-NMR spectroscopy. However, both linkers were cleaved to a small extent during the reactions and when the solid-supported peptides 64 and 65 were ready, 11% and 8% free linker was present on the respective resins. The non-fluorinated versions of 58 (such as 8, figure 2.1) are considered very acid labile and carboxylic acids are cleaved from them using as little as 1% TFA in CH₂Cl₂. The electron withdrawing properties of the fluorine, and the fact that fluorinated versions of the Wang-resin require more harsh cleaving conditions than the original, prompted the examination of 5% TFA in CH₂Cl₂ as cleaving solution for 58. The resins were submitted to the cleaving solution for 30 minutes at room temperature. After extensive
Scheme 8.4: Synthesis of a small peptide for investigation of cleavage conditions. a) TentaGel HL-NH₂, DIC, HOBt. b) Fmoc-Gly-OH, MSNT, MeIm. c) i. Piperidine, 20% in DMF. ii. Fmoc-\textit{p}-F-Phe-OH, DIC, HOBt. iii. Piperidine, 20% in DMF. iv. 2,4-difluorobenzoic acid, DIC, HOBt.
washings, $^{19}$F-NMR spectroscopy showed $\sim 25\%$ cleavage from the dialkoxy linker (i.e resin 65) and none from 64. Repeated cleavage for 1.5, 2, and 3 hours in sequence showed continued cleavage from resin 65 while resin 64 remained unaffected (see figure 8.3). After the full 7 hours, $\sim 85\%$ of the peptide had been cleaved from resin 65. Preliminary studies indicated that the acid lability of the linker 46 (see sections 8.1 and 2.2) is intermediate to 32 and 58.

![Figure 8.3: Investigation of peptide acid cleavage from resins 64 (black circles) and 65 (grey squares). The resins were subjected to 5% TFA in CH$_2$Cl$_2$ for 0.5, 1.5, 2, and 3 hours in sequence. After each period, the resins were washed extensively and $^{19}$F-NMR spectra were recorded. The ratio between the linker and the peptide was so measured after a total of 0.5, 2, 4, and 7 hours.](image)

To further examine the usefulness of the linker 58, its stability to solid-phase glycosylation with thioglycosides was studied. The peptide 66 was synthesized in a manner analogous to 65. The serine residue in the peptide was then glycosylated using the previously described glycosyl donor 67, activated with NIS and TfOH (see scheme 8.5). Repeated glycosylations (2 x 2 equivalents) furnished the resin-bound glycopeptide 68 in 73% yield according to $^{19}$F-NMR spectroscopy. This is a rather high yield considering the relatively few equivalents of donor used. Many solid-phase glycosylations are routinely carried out with 2 x 5 equivalents donor. No cleavage of the peptide could be detected under these conditions. The glycopeptide was cleaved from the solid-phase with concurrent deprotection of the fluoro-benzylidene group by using TFA/water 9:1 at room temperature for 2 hours. After preparative LC/MS the partially protected glycopeptide 69 was received in 13% based on the initial loading capacity of TentaGel HL-NH$_2$. Debenzoylation was performed using
LiOH in MeOH/water 4:1 to give the deprotected glycopeptide in 55% yield. Unfortunately the $^1$H NMR spectra of 25 were slightly impure and the anomeric proton could not be assigned. However, a $^1$H-$^{13}$C correlation spectrum showed one major product with the anomeric $^{13}$C shift at 103.8 ppm, indicating the β-anomer.

In conclusion the linker 58 was prepared in 29% yield from trifluorobenzonitrile 51 using nucleophilic aromatic substitutions. The linker can be cleaved with mild acid (5% TFA in CH$_2$Cl$_2$) and is sufficiently stable for glycosylations. The linker 58 therefore is an important complement to the less acid sensitive monoalkoxy linker 32.
Scheme 8.5: Solid-phase glycosylation of a peptide bound to linker 58 and its cleavage and deprotection. a) NIS, TfOH. b) TFA/water 9:1, 2 h. c) LiOH (20 mM in MeOH/water 4:1).

= TentaGel HL-NH₂
9 Design of an NMR tube filter reactor for solid-phase synthesis and gel-phase $^{19}$F-NMR spectroscopy (paper IV)

When monitoring solid-phase reactions with $^{19}$F-NMR spectroscopy the transfer of resin between the reaction vessel and an NMR tube with the following preparation of the gel-phase is a both impractical and time-consuming step. The modification of an NMR tube to a solid-phase reactor was anticipated to address these problems.

The reactor was constructed out of an ordinary medium-thick (0.8 mm) walled borosilicate glass tube with 5 mm outer diameter (i, figure 9.1). Since gel-phase NMR spectra have rather large line widths, the high precision of specialized NMR-tubes was not needed and the cost was thereby reduced. A hollow teflon plug (ii) with a teflon filter (iii) was inserted into the glass tube, making the reactor. For reactions, this reactor was attached to a circulating pump through two teflon adapters (figure 9.1a and f). At the bottom of the reactor, a teflon Luer adapter (iv) covered the plug ii and the base of the tube i. This Luer adapter was connected to the pump via an ordinary Luer syringe needle carrying a teflon hose. At the top of the tube, another teflon adapter (v) was attached, with an inserted steel tube carrying a teflon hose. By using this setup, reaction solutions could be circulated through resin placed in the reactor (figure 9.1f). For washings, the top adapter (v) was instead connected to a container for washing solvents and the pump outlet was diverted to a waste (see figure 9.1g). At the time for NMR spectroscopy, CDCl$_3$ was added to form a gel-phase, the two adapters iv and v were removed and the hole in plug ii was closed by the teflon stopper vi (see figure 9.1e). The reactor was now ready to be inserted into an NMR spectrometer and gel-phase spectra showed similar line-width to those recorded in normal NMR tubes.$^{58}$

To test the applicability of the reactor, the N-benzoylated dipeptide 71 was synthesized. The synthesis was performed on TentaGel
Figure 9.1: The NMR tube filter reactor. a) The reactor during a reaction. b) The reactor ready for NMR. c) The teflon adapters. d) The teflon filter, plug, and stopper. e) Schematic drawing of the reactor ready for NMR. f) Schematic drawing of the reactor during synthesis. g) Schematic drawing of the reactor during washing. Figure adapted from manuscript IV
HL-NH$_2$ with the linker 32 (scheme 9.1), essentially as for the test of cleavage reactions of linker 58 (see section 8.2). The same building blocks and reagents led up to resin-bound peptide 64. However, during this full synthesis, including gel-phase NMR spectroscopy, the resin was never removed from the reactor. All reactions were quantitative according to $^{19}$F-NMR spectroscopy, but as for the synthesis in section 8.2 a small amount of the peptide was cleaved from the linker during the synthesis. When the peptide was ready, the resin was removed from the reactor to allow for heating and the peptide was cleaved with TFA/water 9:1 at 60 °C. Purification with preparative LC/MS gave the peptide in 37% yield based on the initial loading capacity of TentaGel HL-NH$_2$.

A solid-phase filter reactor that can be inserted into an NMR spectrometer was constructed and its usefulness was demonstrated with the synthesis of a small peptide. The resin-bound peptide was constructed without removal of the resin from the reactor during the full synthesis, including NMR spectroscopy. The gel-phase $^{19}$F-NMR spectra were of similar quality to those obtained in standard NMR tubes. Gel-phase $^{19}$F-NMR spectroscopy is a good method to monitor solid-phase synthesis and with the described reactor the resin handling is simplified and the synthesis and analysis process is thus streamlined.
Scheme 9.1: Synthesis and cleavage of a small peptide. a) TentaGel HL-NH₂, DIC, HOBt. b) Fmoc-Gly-OH, MSNT, Melm. c) i. Piperidine, 20% in DMF. ii. Fmoc-p-F-Phe-OH, DIC, HOBt. iii. Piperidine, 20% in DMF. iv. 2,4-difluorobenzoic acid, DIC, HOBt. d) TFA/water 9:1, 60 °C.
10 Concluding remarks and future perspectives

Glycoconjugates are important molecules with diverse biological functions. The complex structure and biosynthesis of glycoconjugates make the isolation of pure, structurally defined compounds from natural sources cumbersome. Therefore, to better address questions in glycobiology, the use of synthetic glycoconjugates is an appealing alternative.

In this thesis, synthetic methods for the preparation of glycolipid analogues have been developed. These methods are amenable to library synthesis and make use of solid-phase chemistry. The biological applications of the glycolipids were demonstrated in two different settings. The CD1 restricted binding of glycolipids carrying α-GalNAc as carbohydrate was detected on viable cells of mouse origin and glycoconjugate arrays was prepared using serine-based neoglycolipids. The usefulness of the arrays was illustrated with two lectins, RCA$_{120}$, specific for α- and β-galactose, and BS-1, specific for α-galactose and α-GalNAc. Both lectins bound to the array of neoglycolipids in agreement with their respective specificity.

Gel-phase $^{19}$F-NMR spectroscopy, using fluorinated linkers and protecting groups, is an excellent tool for monitoring solid-phase synthesis of e.g. glycoconjugates. Two novel fluorinated linkers for the attachment of carboxylic acids have been developed and are presented in the thesis. These linkers can be cleaved with both acids of varying strengths and nucleophiles like hydroxide ions, and they are stable to glycosylation conditions. In addition, a novel filter reactor for solid-phase synthesis was designed. The reactor fits into an ordinary NMR spectrometer to facilitate the reaction monitoring with gel-phase $^{19}$F-NMR spectroscopy.

The projects in the thesis have several interesting openings for continued research. In the CD1-related project (section 5.1 and chapter 6) the first thing to do would be to synthesize a library of α-GalNAc containing glycolipids with different lipid composition.
library can be evaluated for CD1 binding using the method described in chapter 6 with the possible extension of competitive experiments to obtain more quantitative responses. Also the serine-based glycolipids from section 5.2 would be exciting to test for CD1-binding. To follow up, immunization studies on mouse and preferably also guinea pigs with some immunologically relevant glycolipids is of interest. Guinea pigs are of special interest in such experiments because of their full CD1 repertoire in contrast to the mouse that only have CD1d molecules.

The synthetic methods developed in section 5.2 and chapters 8 and 9 can be used to prepare serine-based glycolipid libraries suitable for array preparation. In addition to the microtiter approach described in chapter 7, also glass slides with an amine surface can be used for attachment and hence additional miniaturizations can be achieved through the use of equipment and techniques developed for oligonucleotide and cDNA microarrays. To synthesize more complex glycolipids, enzymatic modifications of array-bound glycolipids are an appealing alternative.

The fluorinated dialkoxy linker 58 have, apart from the library synthesis mentioned above, a number of possible applications. One would be to investigate direct glycosylations on the linker and further oligosaccharide synthesis. Acidic cleavage of the linker-bound glycoside should give the hemiacetal that could be further derivatized, e.g. to give a glycosyl donor for glycoconjugate synthesis. The attachment and cleavage of alcohols would also be interesting to investigate. Non-fluorinated benzaldehyde linkers can be used to immobilize amines through reductive amination and the formed amine can be acylated and cleaved. The oxidation of the alcohol in resin-bound 58 will result in a resin that may be used in a similar way. This would enable e.g. synthesis of serine-based glycolipids without the terminal carboxylic acid.

Regarding the NMR tube filter reactor, more reactions need to be tested. The first thing to investigate would be glycosylation reactions. Also other, more harsh reaction conditions, e.g. the acidic cleavage from linker 58, are of interest. Finally, more automated methods for reactions and washings would be appealing.

In summary, the methods and equipment described in this thesis can be used to answer a variety of questions in glycobiology such as structural requirements for CD1 binding and specificity of carbohy-
Solid-phase synthesis and biological applications

drake recognizing proteins from bacteria and virus.
Acknowledgments

Det finns väldigt många att tacka här i världen. Här tänkte jag begränsa mig till de som har någonting med avhandlingsarbetet att göra men det är fortfarande alltför många för att jag ska komma ihåg alla. Så ett alldeles särskilt tack till de som blir bortglömda!

Till att börja med vill jag tacka Mikael! Det har varit både roliga och spännande år och jag har särskilt uppskattat att du har haft tid när jag har kommit med frågor. Du har också haft en förmåga att entusiasmera och pigga upp när data inte tyckt som jag...

Jag vill också tacka Jan och Björn för introduktionen till kolhydratkemi. Alla sockerbagare, Brodde, Petter, Lotta, Moge, Sussi, Tomas?, Majecz, Shaji, Baquer, Anna L, Sara, socker är gott!


Mikael, Liying och Annalena på KI, tack för ett gott samarbete och för att jag fick komma och prova på själv!

Tack Sara, som tar över, du fick slita i början...

Mikael, Moge, Sussi, Maciej och Anna, tack för korrekturläsningen. Ni hittade mycket, men jag är säker på att en del återstår att se och jag har nog stoppat in fler fel sen ni fick läsa.

Pelle, Bert, Carina, och Ann-Helén för all hjälp med alla praktiska bestyr.

Gamla och nuvarande rums- och labkamrater, Björn, Lotta, Shaji, Hasse, Neas, Henke, Anna, Veronica, Sussi, Maciej och alla exjobbare
Glycoconjugates

för prat och tjat både nu och då.

Alla andra på avdelningen, för att det är lätt att fika lite för länge...

Göstasgången, med ingredienser som Trulsson, Koffe, Björn, Lotta, Tomas, Mattias, PJ, och många fler. En gång i tiden var det nästan gott!

Alla i kp96 med omnejd för att studietiden kändes så rätt.

Övriga vänner som förgyller min vardag!

Mamma, pappa, Jon, Ida, Noah och Gustav för att ni trott på mig även om ni inte har förstått vad jag säger.

Familjen Bergstrand för att ni tagit emot mig och blivit familj ni med.

Tack Gud, för inspiration, peppning och för att jag hållit mig frisk under skrivandet! Stort tack till alla som bett och ber för mig, Anna och Thea.

Anna och Thea, mina tjejer! Utan er vore det inget värt! Tack för att ni stått ut de senaste veckorna.
Solid-phase synthesis and biological applications

References


Glycoconjugates


Solid-phase synthesis and biological applications


Solid-phase synthesis and biological applications


Solid-phase synthesis and biological applications


Glycoconjugates


[119] S. Angeloni, J. L. Ridet, N. Kusy, H. Gao, F. Crevoisier, S. Guinchard,
Solid-phase synthesis and biological applications


[127] M. Elofsson, J. Broddefalk, T. Ekberg, and J. Kihlberg, Synthesis of a watersoluble serine-based neoglycolipid which can be covalently linked to solid phases. Carbohydrate Research, 1994, 258, 123–133.


[134] K. Yamada, E. Fujita, and S.-I. Nishimura, High performance polymer sup-


