

# **Design and Synthesis of high Affinity Ligands for the Asialoglycoprotein Receptor (ASGP-R)**

**Inauguraldissertation**

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

**Oleg Khorev**

**aus Moskau, Russland**

Referent: Prof. Dr. Beat Ernst

Korreferent: Dr. Gerhard Müller

Basel, September 2007

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag  
von

Prof. Dr. Beat Ernst, Institut für Molekulare Pharmazie, Universität Basel,  
Schweiz

Dr. Gerhard Müller, GPC Biotech AG, München, Deutschland

Prof. Dr. Hans-Peter Hauri  
Dekan

Basel, den 19. September 2007

*To my dear grandparents*

*Ludmila and Yuri*

## Acknowledgements

I would like to thank Prof. Dr. Beat Ernst for giving me the opportunity to pursue a PhD in Medicinal Chemistry in a challenging and interesting topic in an excellent institute, and for his help and support.

I would like to sincerely thank Dr. Oliver Schwardt for his help, advice, guidance and support throughout my PhD journey, and for proofreading parts of the thesis.

I am really grateful to Dr. Brian Cutting for performing NMR experiments for my thesis, as well as constantly helping out with NMR problems.

Many thanks go to Daniela Stokmaier and Daniel Ricklin for patiently and painstakingly testing my compounds, and to Morena Spreafico, Martin Smiesko and Markus Lill for doing molecular modeling for me.

I would like to thank my diploma students Matthias Wittwer and Fabienne Böni for making an invaluable contribution to my work, and for their friendship.

I also want to thank all the members of the ASGP-R team (Daniela Stokmaier, Daniel Ricklin, Claudia Riva, Karin Johansson, Rita Born, Dr. Said Rabbani) for interesting discussions and fruitful collaborations.

A special thank you goes to Gabi Pernter for her warm friendship and constant support since day one of my Swiss adventure.

And of course I am lucky to have worked with all of the members of the IMP, past and present.

I would like to kindly thank Cornelia and her entire family for everything they did to support and encourage me in difficult times, and for making me feel at home in Switzerland.

And last but not least, I want to thank all of my close friends and family for always being there for me.

## Abstract

The asialoglycoprotein receptor (ASGP-R) is a carbohydrate-binding protein from the C-type lectin family that is expressed exclusively and in high numbers on mammalian hepatocytes. The human ASGP-R is a transmembrane protein, consisting of two homologous subunits (H1 and H2), that recognizes and binds desialylated glycoproteins with terminal galactose or *N*-acetylgalactosamine residues. The binding process is followed by receptor-mediated endocytosis of the receptor-ligand complex by the parent hepatocyte. The ASGP-R is then recycled back to the surface, whereas the ligand is ferried to the lysosomes for enzymatic degradation. Due to its location and efficient ligand uptake, the ASGP-R has for a long time been a validated target for liver-specific drug delivery. Furthermore, there is substantial evidence that the ASGP-R is involved in hepatitis B and C virus entry into the liver cells.

The focus of this thesis was to design and synthesize various high affinity ligands for the ASGP-R that could be used as (1) drug carriers for liver-specific drug delivery, (2) small molecular weight inhibitors of hepatitis B/C entry, (3) a spin-labeled GalNAc-based molecular probe for second binding site screening by NMR, and (4) a set of trivalent compounds for investigating the local concentration effect on ligand affinity towards the ASGP-R by surface plasmon resonance (BIAcore).

The trivalent drug carrier for liver-specific drug delivery was shown to bind with high affinity and selectivity to the ASGP-R, and is now awaiting the next step, namely, its conjugation to a therapeutic agent and *in vivo* testing.

The TEMPO spin-labeled GalNAc derivative was successfully used as a first-site ligand for second-site screening by NMR, in which imidazole was identified as a potential second-site ligand. Therefore, after the removal of the TEMPO spin label the first-site ligand will be used in further studies, involving "*in situ* click

chemistry”, in order to find the appropriate linker for joining the first- and second-site ligands.

The four trivalent compounds synthesized for investigating the local concentration effect had an identical molecular mass and scaffold, but differed in the ratio of D-galactose to D-glucose moieties per molecule. Since the affinity of glucose towards the ASGP-R is  $> 20$  mM, and that of galactose is 2.2 mM, the affinity was expected to increase with increasing number of galactose moieties. However, the compound bearing two galactose and one glucose residue unexpectedly showed an affinity greater than that for a compound with three galactose residues. The phenomenon is yet to be explained and verified by further experiments. Nevertheless, the results presented in this work did confirm that the statistical local concentration effect has a weaker influence on multivalency than the chelate effect.

## Abbreviations

$[\alpha]_D$	Optical rotation at $\lambda=589$ nm
AIBN	2,2'-Azobisisobutyronitrile
ASGP	Asialoglycoprotein
ASGP-R	Asialoglycoprotein receptor
ASOR	Asialoorosomuroid
ax.	axial
Con A	Concavalin A
CRD	Carbohydrate recognition domain
DCC	<i>N,N'</i> -Dicyclohexylcarbodiimide
DCE	Dichloroethane
DCM	Dichloromethane
DIPEA	Diisopropylethylamine
DMF	<i>N,N</i> -Dimethylformamide
DMSO	Dimethylsulfoxide
EcorL	<i>Erythrina corallodenrum</i> lectin
EDTA	Ethylenediaminetetraacetic acid
eq.	equatorial
ESI-MS	electrospray ionization mass spectrometry
Gal	D-Galactose
Gal-3	Galectin 3
GalNAc	D- <i>N</i> -Acetylgalactosamine
Glc	D-Glucose
H1/H2	Human ASGP-R subunit 1/2
HBV	Hepatitis B virus
HCV	Hepatitis C virus
HepG2	Human hepatocellular carcinoma cell line
HOBt	1-Hydroxybenzotriazole
HSQC	Heteronuclear single quantum coherence
IFN- $\alpha$	Interferon- $\alpha$



K <sub>D</sub>	Equilibrium dissociation constant
LCMS	Liquid chromatography mass spectrometry
NIS	<i>N</i> -iodosuccinimide
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
o.y.	Overall yield
PAA	Polyacrylamide
r.t.	Room temperature
RHL-1	Rat hepatic lectin subunit 1
RP-C18	Reverse phase silica gel
SAR	Structure-activity relationship
SPR	Surface plasmon resonance
TEMPO	2,2,6,6-Tetramethylpiperidine- <i>N</i> -oxyl
THF	Tetrahydrofuran
TLC	Thin layer chromatography
Tris	2-Amino-2-(hydroxymethyl)-1,3-propanediol

# Table of Contents

<b>Chapter 1: General Introduction to the ASGP-R.....</b>	<b>13</b>
1.1 Introduction.....	13
1.1.1 Receptor structure.....	14
1.1.2 Physiological role of the ASGP-R.....	15
1.1.3 Ligand structure: Multivalency.....	17
1.1.4 Receptor-mediated endocytosis.....	20
1.1.5 Endosomal Compartments.....	21
1.1.6 Targeting hepatocytes for gene and drug delivery.....	22
<b>Chapter 2: Trivalent, Gal/GalNAc-containing Ligands designed for the Asialoglycoprotein Receptor.....</b>	<b>24</b>
2.1 Abstract.....	25
2.2 Introduction.....	26
2.3 Results and Discussion.....	31
2.3.1 Synthesis of fluorescent, trivalent ligands <b>6</b> and <b>7</b> and the negative control <b>8</b> .....	32
2.3.2 Biological Evaluation.....	36
2.3.3 Fluorescence Microscopy.....	38
2.3.4 Flow Cytometry.....	39
2.4 Conclusion.....	42
2.5 Acknowledgement.....	44
2.6 Experimental Section.....	44
2.6.1 General Methods.....	44
2.6.2 Abbreviations.....	46
2.6.3 General procedure A: preparation of triacetylated compounds <b>11</b> , <b>13</b> and <b>15</b> .....	46
2.6.4 General procedure B: preparation of triallylated compounds <b>12</b> and <b>14</b> .....	48

2.6.5 Ligand binding and internalization.....	62
2.6.5.1 Fluorescence microscopy.....	62
2.6.5.2 Flow cytometry.....	63
2.7 References.....	65

## **Chapter 3: Directed Library of Ligands for the ASGP-R.....68**

3.1 Introduction.....	68
3.1.1 Hepatitis B.....	68
3.1.2 Hepatitis C.....	69
3.1.3 Design of a small molecular weight ASGP-R ligand.....	71
3.1.4 Binding mode of GalNAc to the ASGP-R.....	72
3.1.5 Structure and topology of the binding site.....	74
3.1.6 The Huisgen 1,3-dipolar cycloaddition.....	75
3.2 Results and Discussion.....	78
3.2.1 Scaffold synthesis.....	78
3.2.2 Synthesis of phenyl propargyl ethers.....	80
3.2.3 Library synthesis.....	81
3.2.4 Competitive binding assay.....	85
3.3 Conclusion.....	88
3.4 Experimental.....	89
3.4.1 Scaffold synthesis: compound <b>4</b> .....	89
3.4.2 Synthesis of substituted phenyl propargyl ethers.....	91
3.4.3 Library synthesis: Huisgen 1,3-dipolar cycloaddition.....	94

## **Chapter 4: Local Concentration.....107**

4.1 Introduction.....	107
4.2 Results and Discussion.....	109
4.2.1 Synthesis of compounds <b>57</b> , <b>58</b> , <b>59</b> and <b>60</b> .....	109
4.2.2 Biological evaluation on BIACORE.....	114
4.3 Conclusion.....	116
4.4 Experimental.....	117

<b>Chapter 5</b>	<b>Synthesis of spin-labeled GalNAc for second-site NMR screening.....</b>	<b>127</b>
5.1	Introduction.....	127
5.1.1	NMR in drug discovery and development.....	128
5.1.2	Paramagnetic relaxation enhancement.....	129
5.1.3	Linking of first- and second-site fragments.....	130
5.1.4	Spin Labels.....	131
5.2	Results and Discussion.....	133
5.2.1	Synthesis of compound <b>75</b> .....	133
5.2.2	Testing of compound <b>75</b> in the competitive binding assay.....	135
5.2.3	Transverse relaxation rate (T1rho) measurements.....	136
5.3	Conclusion.....	138
5.4	Experimental.....	140
<b>Appendix 1:</b>	<b>Polymer assay.....</b>	<b>143</b>
<b>Appendix 2:</b>	<b>BIACORE.....</b>	<b>144</b>
<b>Appendix 3:</b>	<b>Second-site screening by NMR.....</b>	<b>146</b>
<b>Appendix 4:</b>	<b>General methods.....</b>	<b>147</b>
<b>References.....</b>		<b>148</b>
<b>Curriculum Vitae.....</b>		<b>154</b>

# Chapter 1: General introduction to the ASGP-R

## 1.1 Introduction

In spite of their relative weakness, carbohydrate-protein interactions have been shown to be very specific. Nevertheless, the endogenous ligands are often complex carbohydrates or glycoproteins that are unsuitable for therapeutic use. Therefore, it is of extreme importance for medicinal chemistry to design carbohydrate mimics with simplified structures, improved biostabilities and higher affinities towards their targets.

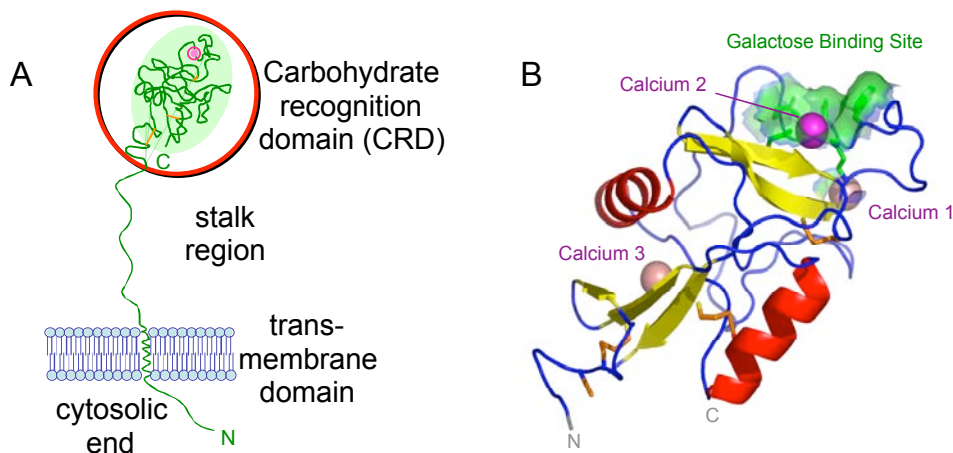
The asialoglycoprotein receptor (ASGP-R) is a carbohydrate-binding protein, or lectin, which recognizes and binds glycoproteins with terminal, non-reducing galactose or *N*-acetylgalactosamine residues. It is located in high numbers on hepatocytes [1], and was originally discovered by Ashwell and Morell [2]. Binding of the ligand to the ASGP-R leads to receptor-mediated endocytosis of the ligand-receptor complex by the hepatocyte. The exact physiological function of the ASGP-R still remains unclear. However, it is definitely involved in clearing desialylated glycoproteins from the blood, thus maintaining serum glycoprotein homeostasis [3].

Due to its high level of expression on the hepatocytes, and its efficient endocytosis of appropriate ligands, the ASGP-R has for long been a validated target in medicinal chemistry for liver-specific drug and gene delivery [4].

### 1.1.1 Receptor structure

The focus of this thesis is on the human ASGP-R, which is an integral transmembrane protein composed of two subunits designated H1 and H2 ( $M_r = 46$  kDa and 50 kDa, respectively), with exoplasmic C-termini and endoplasmic N-termini [1]. The subunits share 57% sequence homology and have the same polypeptide domain construct. Moreover, both are post-translationally modified by the addition of N-linked oligosaccharides and palmitoylation [1]. Each subunit is a type II transmembrane protein, and the subunits oligomerize in the ratio of 1:2-5 (H1:H2) [1]. The ASGP-R is located on the basolateral (circulation facing) membrane of the parenchymal liver cells, and it is estimated that there are approx 500,000 ASGP-R subunits/cell, however the number varies according to cell type and method of estimation [5,6].

Starting from the N-terminus (Figure 1A), the H1 or H2 subunit of the ASGP-R is composed of a cytosolic domain, a trans-membrane domain consisting of approx. 20 hydrophobic amino acids, a stalk region and a C-terminal carbohydrate recognition domain, or CRD (Figure 1B). The stalk region is involved in the oligomerization of the subunits. The X-ray crystal structure of the H1-CRD has recently been published [7].



**Figure 1.** (A) The H1-subunit of the ASGP-R. (B) H1-CRD (Picture courtesy of D. Ricklin).

The ASGP-R belongs to the C-type lectin family, which implies that the ligand binding is calcium-dependent and requires an optimal calcium concentration of 0.1-2 mM [8-10].

The H1-CRD contains three  $\text{Ca}^{2+}$  ions (Figure 1B). One is located in the binding site, and interacts directly with the terminal Gal or GalNAc residue of the ligand, the other two are responsible for structural integrity. The H1-CRD also contains 7 cysteines, 6 of which form 3 disulfide bonds [7].

The binding site of the ASGP-R is specific for D-galactose and D-N-acetylgalactosamine, with a 50-fold higher affinity for the latter [11,12].

Both subunits contain a sugar binding site, however, it is believed that only the H1 subunit is responsible for sugar recognition and high affinity binding [13], whereas the H2 subunit simply serves to generate the functional native receptor since both subunits are necessary for efficient ligand binding and internalization by the hepatocyte [14-16].

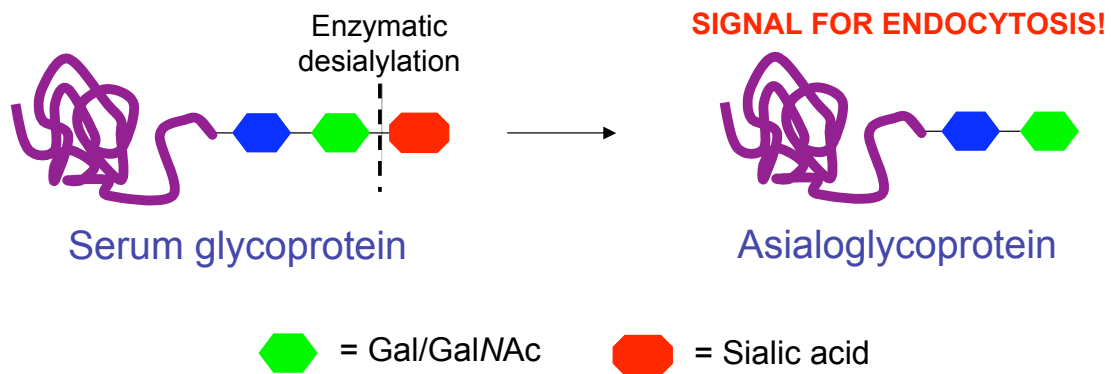
Furthermore, the ASGP-Rs cluster together on the hepatocyte surface to form receptor patches. However, the exact *in vivo* arrangement of the native receptor subunits is not accurately known [17].

### **1.1.2 Physiological role of the ASGP-R**

The exact physiological function of the ASGP-R is not yet fully elucidated. However, it is definitely involved in clearing desialylated glycoproteins from the blood, thus maintaining serum glycoprotein homeostasis [3]. This is supported by the findings that patients with liver diseases like cirrhosis or liver cancer have

elevated levels of asialoglycoproteins, presumably because of impaired liver – and hence ASGP-R – function [18].

The penultimate Gal/GalNAc residue on N-linked oligosaccharides of serum glycoproteins is practically always capped by a sialic acid. When this saccharide is removed by sialidases, an asialoglycoprotein (ASGP) is created which binds to the ASGP-R and gets internalized by the liver (Figure 2).



**Figure 2.** Generation of asialoglycoproteins by the action of sialidases in the serum.

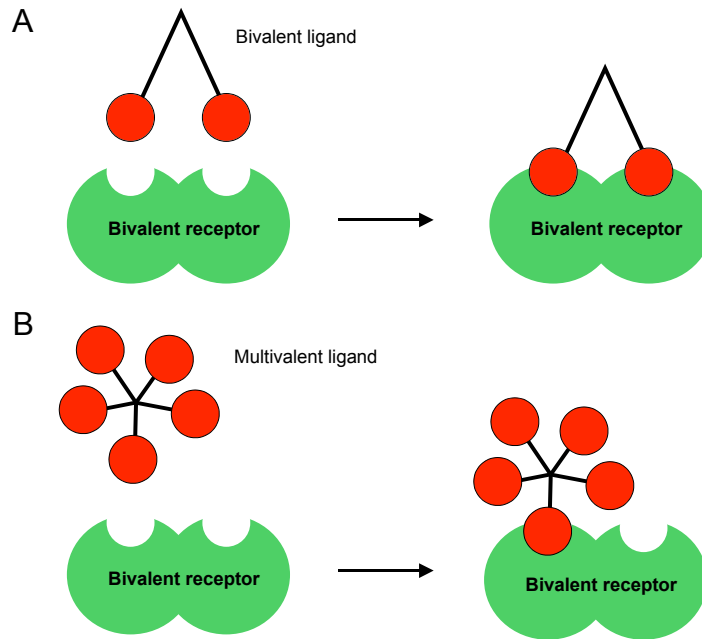
The desialylating activity in the serum is ubiquitous and random, hence there is a steady production of desialylated glycoproteins, which should be degraded/recycled. An example of this is the clearance of remnants of apolipoprotein E, which is secreted in the sialoprotein form, and subsequently desialylated in the serum [19].

Another function of the ASGP-R could be the uptake of glycoproteins essential for the liver, such as immunoglobulin A (IgA), which contains terminal Gal and GalNAc residues on its oligosaccharides [20].



### 1.1.3 Ligand structure: Multivalency

Multivalency consists of two components [21,22] that are illustrated in Figure 3. The first is the *chelate effect*, which leads to binding affinity enhancement due to simultaneous spanning of two or more binding sites by the ligand. The second is the *statistical effect*, which increases the binding due to an increased local concentration of the available ligand or binding motives on one ligand.



**Figure 3.** (A) Multivalency consisting of both the chelate and the statistical effect. (B) The multivalent ligand is unable to bridge two binding sites, and hence only the statistical effect operates.

In cases where the ligand is unable to bridge two binding sites on the receptor, the purely statistical effect operates (Figure 3B).

Since the H1 and H2 subunits oligomerize in the 1:2-5 ratio, respectively, and each subunit contains one CRD, this implies that 3-6 sugar binding sites per receptor are presented on the cell surface.

The dissociation constants ( $K_D$ ) for ligands possessing single Gal or GalNAc residues are low, being in the millimolar range. However, a dramatic increase in affinity is observed for ligands that are oligovalent with respect to the number of terminal Gal or GalNAc residues. The binding hierarchy is: tetraantennary  $\geq$  triantennary  $\gg$  diantennary  $\gg$  monoantennary, with the binding affinities being  $10^{-9}$ ,  $5 \times 10^{-9}$ ,  $1 \times 10^{-6}$  and  $1 \times 10^{-3}$  M, respectively. This phenomenon is known as the *cluster glycoside effect* [23,24].

Many studies using natural and synthetic ligands have illustrated the importance of the spatial arrangement of the terminal Gal/GalNAc residues for the binding affinity to the ASGP-R. Based on affinity studies on several neoglycoproteins with defined sugar arrangements and geometries [25], it was concluded that the terminal sugar residues position themselves at the corners of a triangle, whose sides measure 15, 22 and 25 Å [26]. Hence, structures with shorter intergalactose distances or lower flexibility had a lower affinity than that for compounds in which the spatial arrangement of sugar residues was complementary to the arrangement of the receptor binding sites [26].

A further effect exerted by multivalent ligands on biological systems is the induction of receptor subunit clustering on the cell surface, which in the case of lectins was demonstrated by Kiessling *et al.* [27]. The studies were performed on the soluble periplasmic glucose/galactose binding protein (GGBP) on *E. coli*, which is responsible for recognizing chemoattractants (*i.e.* glucose/galactose) and thus mediating chemotaxis. It was thus shown that galactose-bearing polymers increased bacterial chemotaxis in proportion to the number of galactose residues on the polymers. It was also shown by fluorescence microscopy that the multivalent galactose-bearing polymers did indeed induce chemotactic receptor clustering on the bacterial periplasmic membrane. Since receptor subunit clustering is also involved in the ASGP-R-mediated endocytosis,

it is possible that multivalent ligands induce the endocytic cycle by increasing receptor clustering. However, this has to be further verified experimentally.

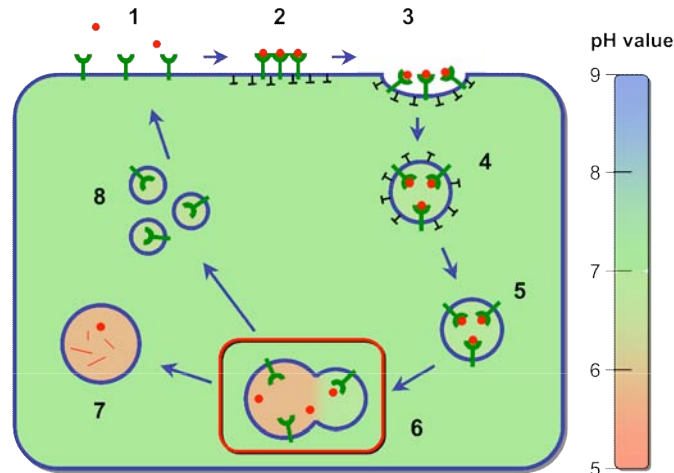
The interaction energies for multivalent ligands with their targets are discussed in detail by Toone *et al.* [24]. The overall entropy of a particle in solution consists of four terms: the translational, rigid-body rotational, conformational and solvation-associated. The translational and rotational terms logarithmically depend on the molecular mass, *i.e.* the greater the overall mass of the particle, the greater the translational and rotational entropies. Thus, upon tethering of two monovalent ligands, the entropy balances out and remains at a value roughly equivalent to that of a monovalent ligand, *i.e.* 15-20 kJ/mol. Attempts to quantify these entropy terms in solution are fraught with uncertainty, especially for cases involving a highly participating solvent like water. The value for translational and rigid-body rotational entropy in solution is often quoted to be around 43 kJ/mol [28], however recent studies have placed the estimate at almost half that value [24].

Conformational entropy tends to decrease upon ligand tethering, with an estimated value of around 5.8 kJ/mol [29], and solvation effects on the entropy of multivalent ligand formation are still poorly understood [24].

The enthalpic component of multivalent binding results mainly from the linker itself. If the linker is able to interact favorably with the protein surface, this leads to favorable changes in the free energy of binding. However, the conformational effects on the linker upon folding are highly influenced by its rotational characteristics, *i.e.* the rotational barrier about the C-C bond for ethane is around 12.5 kJ/mol. Therefore, if the linker is capable of assuming an energetically favorable “relaxed” conformation while at the same time presenting the binding residues in an optimal orientation, this leads to favorable enthalpic consequences, *e.g.* the eclipsed form of butane (about the C<sub>2</sub>-C<sub>3</sub> bond) is 21-25 kJ/mol higher than the lowest energy anti-conformation.

### 1.1.4 Receptor-mediated endocytosis

The ASGP-R-mediated endocytic pathway is schematically summarized in Figure 4, and is reviewed in detail by Spiess *et al.* [3].



**Figure 4.** Diagram illustrating the ASGP-R-mediated endocytic pathway inside the hepatocyte, with associated pH changes. 1) Ligand association; 2) Receptor clustering; 3) Endocytosis; 4) Clathrin-coated vesicle; 5) Endosome; 6) Fusion with lysosome; 7) Ligand degradation; 8) Recycling of receptor. (Picture courtesy of Daniel Ricklin)

The initial step of the ASGP-R-mediated endocytosis involves clustering of the receptors on the hepatocyte cell membrane into clathrin-coated pits, which cover an area of  $\approx 0.1 \mu\text{m}^2$  [30]. Upon ligand binding, the membrane invaginates, and the ligand-receptor complex gets internalized, ending up in a clathrin-coated vesicle, which upon clathrin uncoating fuses with a lysosome.

The pH dependence of ligand binding is an important general feature of most endocytic receptors for it enables ligand release in the acidic environment of the endosomes, caused by the  $\text{H}^+$ -translocating ATPase [31].

The ASGP-R is also internalized via clathrin-coated pits without the presence of the ligand; hence, almost two-thirds of the receptor is located intracellularly. However, binding of the ligand increases the rate of internalization by a factor of two [15,32].

### **1.1.5 Endosomal Compartments**

In the ASGP-R-mediated endocytic cycle, the endosome is a central compartment, since it gives rise to distinct vesicles that either proceed to fuse with the lysosome (degradative pathway) or return to, and fuse with the cell membrane (recycling pathway) [1].

After ligand binding followed by membrane invagination and clathrin-coated vesicle formation, the clathrin coat is eventually removed by uncoating ATPase [33], the vesicles then fuse with endosomal compartments called early endosomes. Ligand binding to the ASGP-R is only effective above pH 6.5, so the lower pH in the early endosomes (pH 6.0) causes an acid-induced conformational change of the protein, which results in the dissociation of the ligand-receptor complex. The ligand is then segregated into the late endosomes (pH = 5.5), which subsequently fuse with the terminal endocytic compartments, the dense lysosomes, where the pH is even lower and the ligand undergoes degradation. The receptor, on the other hand, is rapidly returned from the early endosomes via recycling vesicles to the cell membrane [3].

Kinetic studies [34,35] have shown that an ASGP is internalized within minutes at 37 °C, and that the receptor is recycled back to the surface within a half-time of 5-7 minutes, whereas 50-75% of the internalized ASGPs is retained within the cell [36].

Furthermore, the kinetics of the endocytic cycle were studied in detail by Schwartz *et al.* [37] using HepG2 human hepatoma cells, a reliable model for human hepatocytes [38], in which the entire cycle (ligand binding to ASGP reaching the lysosome) took around 15 minutes. At high ligand concentration, binding to the receptor occurred within 1 minute, internalization was within 2 minutes, and after ligand-receptor complex dissociation, the receptor was recycled back to the surface in 4.2 minutes. The studies measured the linear  $^{125}\text{I}$ -asialoorosomuroid ( $^{125}\text{I}$ -ASOR) uptake at an average rate of 0.02-0.03 pmol/min/ $10^6$ -cells at 37 °C. This value was slightly lower than, but comparable to that of isolated hepatocytes, *i.e.* 0,07-0.1 pmol/min/ $10^6$  cells at 37 °C [39]. Upon binding of  $^{125}\text{I}$ -ASOR at 4 °C, removal of excess ligand and a temperature shift to 37 °C, most of the bound  $^{125}\text{I}$ -ASOR was internalized within 6-8 minutes, in a process that reached a steady state after 30 minutes. The mean lifetime of the receptor ligand complex after internalization was determined to be 2.16 minutes. The main differences between HepG2 cells and normal parenchymal hepatocytes are in the number of receptor subunits on the cell surface, *i.e.* 150,000 on HepG2 vs. 500,000 on hepatocytes, and in the percentage of the receptor found in the cytoplasm, *i.e.* 14% in HepG2 vs. 60% in isolated rat hepatocytes [40]. However, the number of intracellular receptors is strongly influenced by the ligand concentration [37].

### **1.1.6 Targeting hepatocytes for gene and drug delivery**

The liver is a major metabolic organ, which can be damaged by various xenobiotics, by-products of metabolism (*e.g.*, radical species), inflammatory mediators (*e.g.*, cytokines) and microorganisms. Therefore, delivering drugs or genes directly to the liver is a highly promising therapeutic strategy for modifying errors in metabolism, preventing liver damage and inhibiting hepatitis viral replication [4].

An example of ASGP-R-mediated gene delivery *in vivo* was done by Wu *et al.* [41], and involved injecting rats with a poly-L-lysine-DNA complex that was covalently linked to asialoorosomuroid. This resulted in DNA incorporation into, and expression by the liver cells.

The efficiency of the degradative pathway is known to be less than complete. Hence, some substances internalized via the ASGP-R have been shown to escape degradation in the lysosome. This was demonstrated using an asialoglycoprotein-diphtheria toxin A construct that was still lethal to the hepatocyte after being internalized [42,43].

Liver-specific drug delivery was demonstrated by De Vruueh *et al.* [44]. The authors showed that a derivative of the anti-hepatitis B drug 9-(2-phosphonylmethoxyethyl)adenine [45] (PMEA, adefovir), when conjugated to a carrier designed to bind specifically to the ASGP-R, was much more efficiently taken up by the liver in rats (69% of dose vs. <5% free drug) and a lot less by the kidneys (<2% of dose vs. >45% free drug).

## **Chapter 2:**

### **Trivalent, Gal/GalNAc-containing Ligands designed for the Asialoglycoprotein Receptor**

Accepted by Bioorganic & Medicinal Chemistry

Oleg Khorev, Daniela Stokmaier, Oliver Schwardt, Brian Cutting, Beat Ernst\*

Institute of Molecular Pharmacy, Pharmacenter - University of Basel  
Klingelbergstrasse 50, CH-4056 Basel, Switzerland

\* Corresponding author:

Prof. Dr. B. Ernst,

Tel.: +41-61-267 15 51

Fax.: +41-61-267 15 52

E-mail: [beat.ernst@unibas.ch](mailto:beat.ernst@unibas.ch)



## 2.1 Abstract

A series of novel, fluorescent ligands designed to bind with high affinity and specificity to the asialoglycoprotein receptor (ASGP-R) has been synthesized and tested on human liver cells. The compounds bear three non-reducing,  $\beta$ -linked Gal or GalNAc moieties linked to flexible spacers for an optimal spatial interaction with the binding site of the ASGP-R. The final constructs were selectively endocytosed by HepG2 cells derived from parenchymal liver cells - the major human liver cell type - in a process that was visualized with the aid of fluorescence microscopy. Furthermore, the internalization was analyzed with flow cytometry, which showed the process to be receptor-mediated and selective. The compounds described in this work could serve as valuable tools for studying hepatic endocytosis, and are suited as carriers for site-specific drug delivery to the liver.

**KEYWORDS:** asialoglycoprotein receptor (ASGP-R); drug delivery; flow cytometry; fluorescence microscopy; fluorescent probes

## 2.2 Introduction

The asialoglycoprotein receptor (ASGP-R) is located on hepatocytes and is a  $\text{Ca}^{2+}$ -dependent carbohydrate-binding protein, or C-type lectin. It is expressed on mammalian liver cells [1]. Its main function is to maintain serum glycoprotein homeostasis by the recognition, binding and endocytosis of asialoglycoproteins (ASGPs), *i.e.*, desialylated glycoproteins with terminal galactose or GalNAc residues. After internalization via clathrin-coated pits and their fusion with endosomes, the ASGPs are released in the acidic environment of the endosome and transported to lysosomes for degradation, while the receptor is recycled back to the cell surface [2,3].

In addition to the ASGP-R, there are three additional Gal/GalNAc-receptors in the C-type lectin family: the Kupffer cell receptor, the macrophage galactose lectin and the scavenger receptor C-type lectin (SRCL) [4-7]. Their binding properties were recently profiled by Drickamer et al. [8].

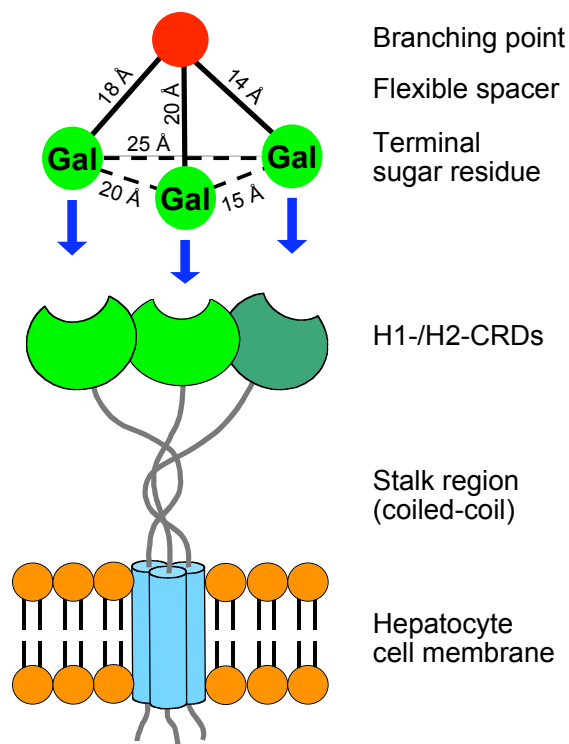
The affinity and specificity of the ASGP-R is a consequence of oligovalent interactions with its physiological ligands, a process termed *cluster glycoside effect* by Lee et al. [9]. The receptor consists of two homologous subunits, designated H1 and H2 in the human system, which form a non-covalent heterooligomeric complex with an estimated ratio of 2-5:1, respectively. Both subunits are single-spanning membrane proteins with a calcium-dependent galactose/*N*-acetylgalactosamine recognition domain [10]. Recently, the X-ray crystal structure of the carbohydrate recognition domain (CRD) of the major subunit H1 was elucidated [11].

Many studies have been performed with both natural and synthetic carbohydrates to establish the structure-affinity relationship for the ASGP-R. Baenzinger et al. [12,13] have shown that the human receptor exhibits specificity for terminal Gal and GalNAc (with an approx. 50-fold higher affinity for the latter)

on desialylated glycoproteins. Triantennary ligands displayed a higher affinity than their mono- and diantennary counterparts. Furthermore, the studies led to the conclusion that only the terminal residues are necessary for specific recognition, and that the binding process proceeds through a simultaneous interaction of 2 to 3 sugar residues with 2 to 3 binding sites of the heterooligomeric receptor. On the native receptor on the hepatocyte surface these binding sites are 25-30 Å apart.

Studies on rabbit hepatocytes by Lee et al. [9,14], using synthetic oligosaccharides, further reinforced the binding hierarchy of polyvalent ligands: tetraantennary > triantennary >> diantennary >> monoantennary. The IC<sub>50</sub>-values for mono-, di-, tri- and tetraantennary oligosaccharides were found to be approx.  $1 \times 10^{-3}$ ,  $1 \times 10^{-6}$ ,  $5 \times 10^{-9}$  and  $10^{-9}$  M, respectively. In other words, although the number of Gal residues/mol of ligand increased only 4-fold, the inhibitory potency increased 1'000'000-fold. Because the fourth Gal moiety present in the tetraantennary ligand does not markedly enhance the affinity, it was assumed that the binding requirements of the cell-surface receptor are largely satisfied by the triantennary structure [15].

The optimal distance of the Gal moieties in these oligosaccharides was determined by binding assays with synthetic carbohydrates representing partial structures of *N*-linked glycans [16], high-resolution NMR and molecular modeling studies [17]. Based on these results, Lee et al. [9,16] presented a model for the optimal spatial arrangement of the terminal sugar residues (Figure 1).

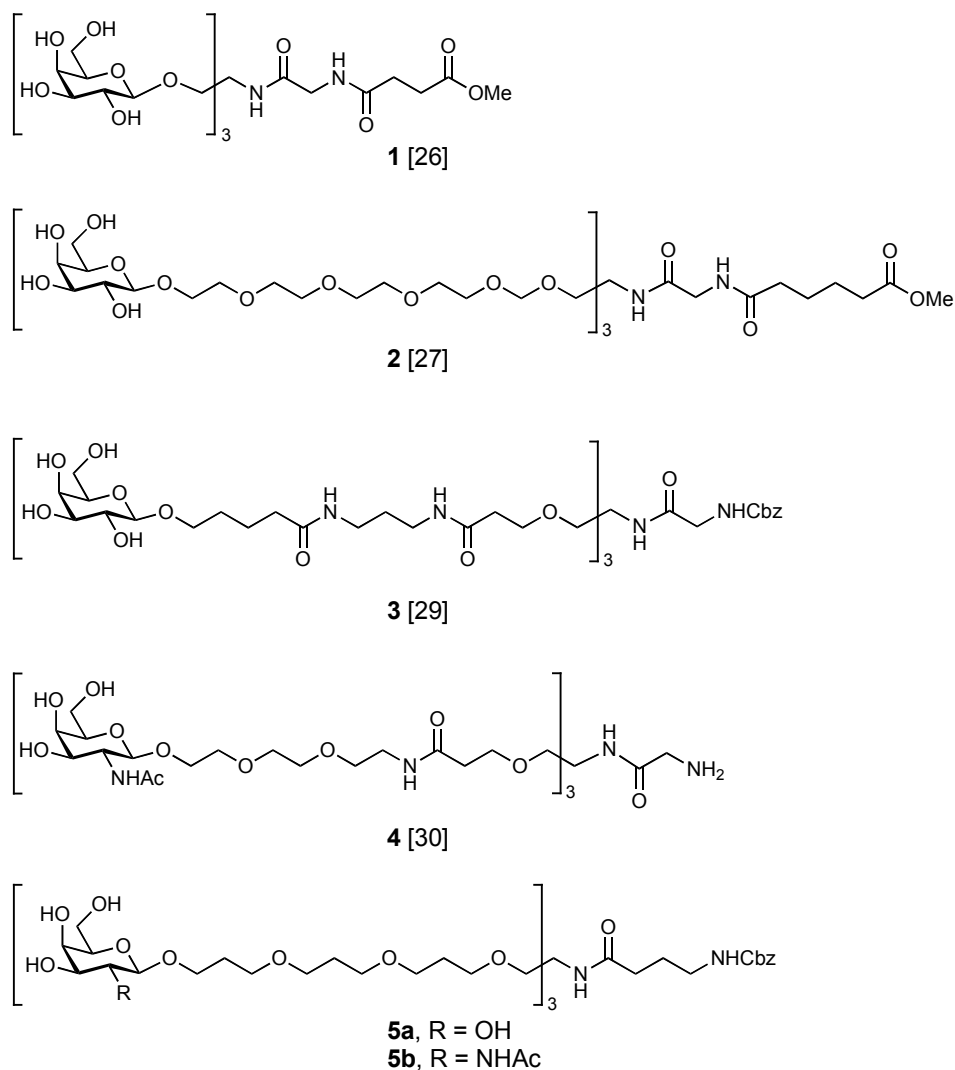


**Figure 1.** Binding model for ASGP-R ligands in an optimal conformation to the heterooligomeric receptor consisting of H1 and H2 subunits. Dashed line indicates the distance between the C-4 of each Gal moiety; filled line represents approximate distance between branching point and C-6 of Gal (14-20 Å). Adapted from Lee et al. [16].

Due to its specificity, predominant expression on hepatocytes and high capacity for receptor-mediated endocytosis, the ASGP-R has been validated as a potential target for drug and gene delivery to the liver [7,8,19]. As an alternative to ex vivo gene transfer to the liver, which requires invasive surgery [20], there is much interest *in vivo* protocols: (i) Wu et al. [21] demonstrated successful *in vivo* gene transfer to hepatocytes with poly-L-lysine linked asialoorosomucoid, (ii) Hara et al. [22-24] showed that asialofetuin-labeled liposomes that encapsulate plasmid DNA cause gene expression and (iii) successful gene transfer to hepatocytes using liposomal gene carriers that possess synthetic galactose residues as a targetable ligand for parenchymal liver cells has been reported [25].

In order to further exploit the ASGP-R for therapeutic purposes, trivalent ligands with pendant Gal or GalNAc residues connected by flexible spacers with appropriate lengths to a common branching point were synthesized. All these ligands incorporate 2-amino-2-hydroxymethyl-1,3-propanediol (Tris) as the branching point (Figure 2). Kempen et al. [26] synthesized the trivalent, Gal-terminated ligand **1**, where the carbohydrate moieties were directly linked to Tris. When **1** was labeled with cholesterol and incorporated into liposomes, they were mainly taken up by the Kupffer cells, via the Gal/Fuc-recognizing receptor, and not by the parenchymal liver cells via the ASGP-R.

Therefore, a new generation of ligands with optimal spacers was created. Biessen et al. [27,28] extended the distance between the Tris branching point and the Gal residues by using tetraethylene glycol spacers approximately 20 Å in length. This indeed led to ligands with improved affinities (see **2**,  $K_i = 0.2 \mu\text{M}$ , Figure 2) determined in a competition assay with  $^{125}\text{I}$ -labeled asialoorosomucoid. In 1999, Sliedregt et al. [29] designed a second generation of cluster glycosides containing an essential modification (see **3**,  $K_i = 93 \text{ nM}$ , Figure 2). To enhance the chemical stability, the methylene acetal groups in **2**, which connect the spacers to Tris, were replaced by acid stable ether bonds. Furthermore, the spacers were no longer based on tetraethylene glycol to achieve the appropriate spacing between the Gal residues, but rather on a twelve atom fragment containing two amide bonds. Finally, Rensen et al. [30] combined the various features from **2** and **3** to generate compound **4** ( $K_i = 2 \text{ nM}$ , Figure 2), which exploited the expected 50-fold higher affinity of GalNAc over Gal towards the ASGP-R [31].



**Figure 2.** Trivalent compounds **1**, **2**, **3** and **4** were specifically designed for, and tested on, the ASGP-R [26-30]. Compounds **5a** and **5b** are the trivalent, Cbz-protected intermediates introduced herein.

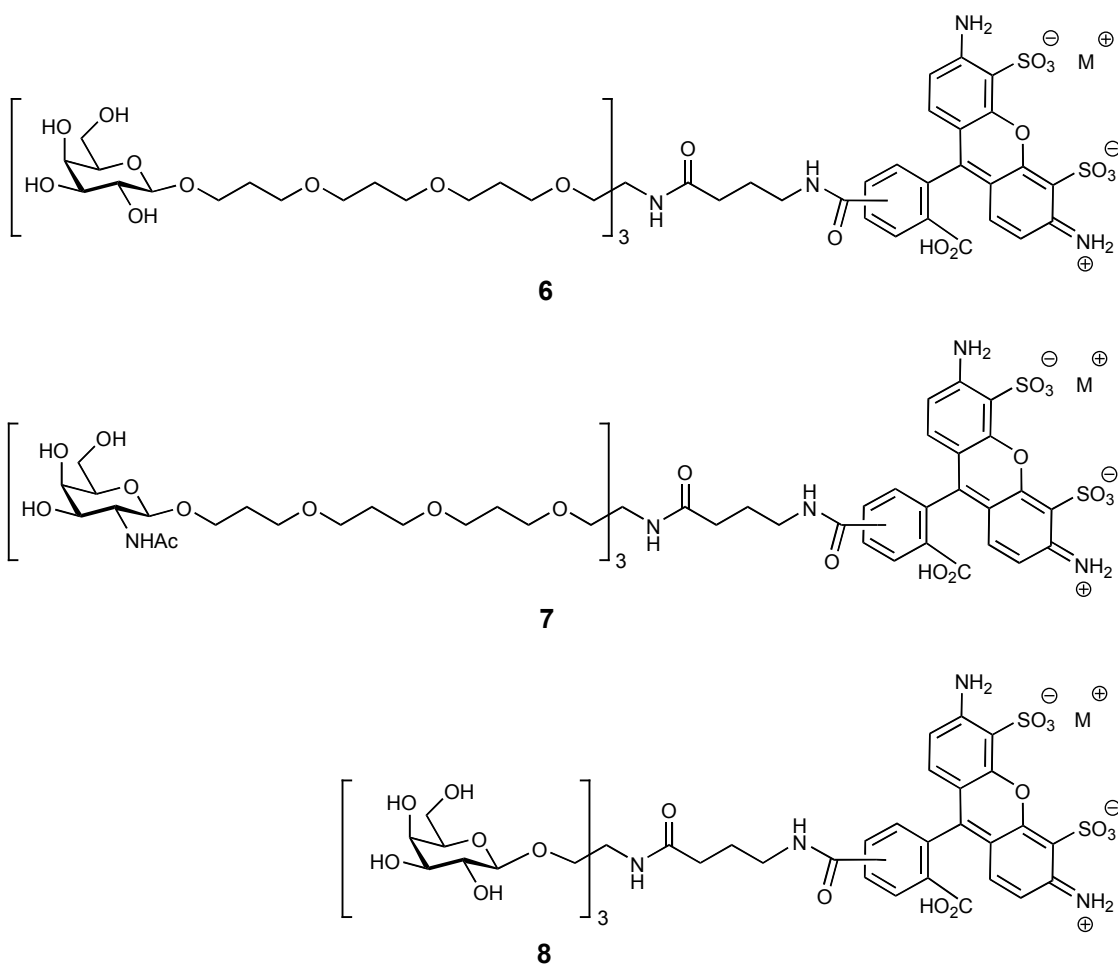
Based on the knowledge gained in previous studies, we set out to synthesize the optimal trivalent carrier (**19**, Scheme 1) with reduced synthetic complexity and high *in vivo* stability. Furthermore, the flexibility and hydrolytic stability of the quintessential spacers was improved without compromising their solubility in water. The resultant intermediates **5a** and **5b** (Figure 2), which possess terminal Gal or GalNAc moieties, respectively, were then fluorescently labeled and tested for selective uptake by hepatocytes using fluorescence microscopy and flow

cytometry. Moreover, since most of the previous research was done on rat [26-30] and mouse [32] liver cells, and the final aim of this research is liver-selective drug delivery in humans, all our biological assays were performed using cell lines of human origin.

## 2.3 Results and Discussion

The main structural features of the trivalent ASGP-R ligands **5a** and **5b** are as follows: (i) Tris is the central branching point, (ii) the spacers are based on polypropylene oxide, which combines flexibility with amphiphilicity, (iii) the linkage between Tris and the spacers is a hydrolytically stable ether bond and (iv) the length of the spacers can be easily varied.

The glycine acylating the amino group of Tris in **4** (Figure 2) has been replaced with Cbz-protected  $\gamma$ -aminobutyric acid, which upon deprotection furnishes a versatile primary amino group for the attachment of fluorescent labels and, at a later stage, therapeutic agents. For our studies, the amino group was coupled to Alexa Fluor<sup>®</sup> 488 fluorescent label [33] ( $\rightarrow$  **6** and **7**, Figure 3), but in theory it could also be coupled to a therapeutic agent. As a negative control for the fluorescence microscopy studies, and especially to demonstrate the significance of the polypropylene oxide spacers featured in our final compounds **6** and **7**, we also synthesized compound **8** (Figure 3) The latter, in contrast to **6** and **7**, has only short spacers, and therefore does not fulfill the spatial requirements for trivalent binding to the ASGP-R.



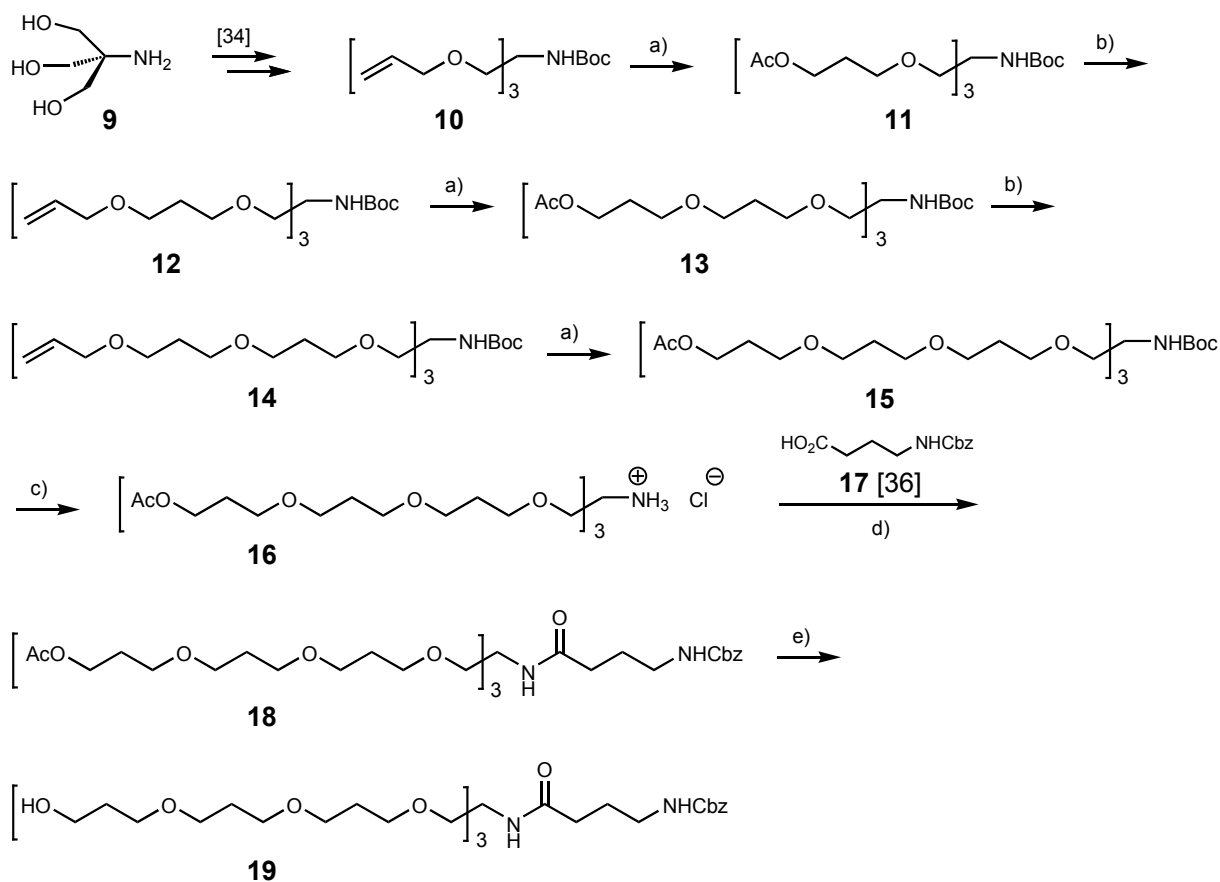
**Figure 3.** Fluorescent, trivalent compounds **6**, **7**, and control **8**;  $M^+$  are variable counterions.

### 2.3.1 Synthesis of fluorescent, trivalent ligands **6** and **7** and the negative control **8**

Starting from 2-amino-2-hydroxymethyl-1,3-propanediol (Tris, **9**), the polypropylene oxide spacers were gradually extended by repetitive allylation-oxidative hydroboration steps using 9-BBN in THF followed by  $H_2O_2$  and aqueous NaOH (Scheme 1). For the synthesis of compounds **12** and **14**, several allylation procedures were examined using NaH, KOH,  $K_2CO_3$  as bases in various solvents (e.g., THF, DMF, dioxane), with and without the addition of crown ethers and quaternary ammonium salts as phase transfer catalysts. All procedures, including the literature procedure used to obtain **10** [34] in 68%, led



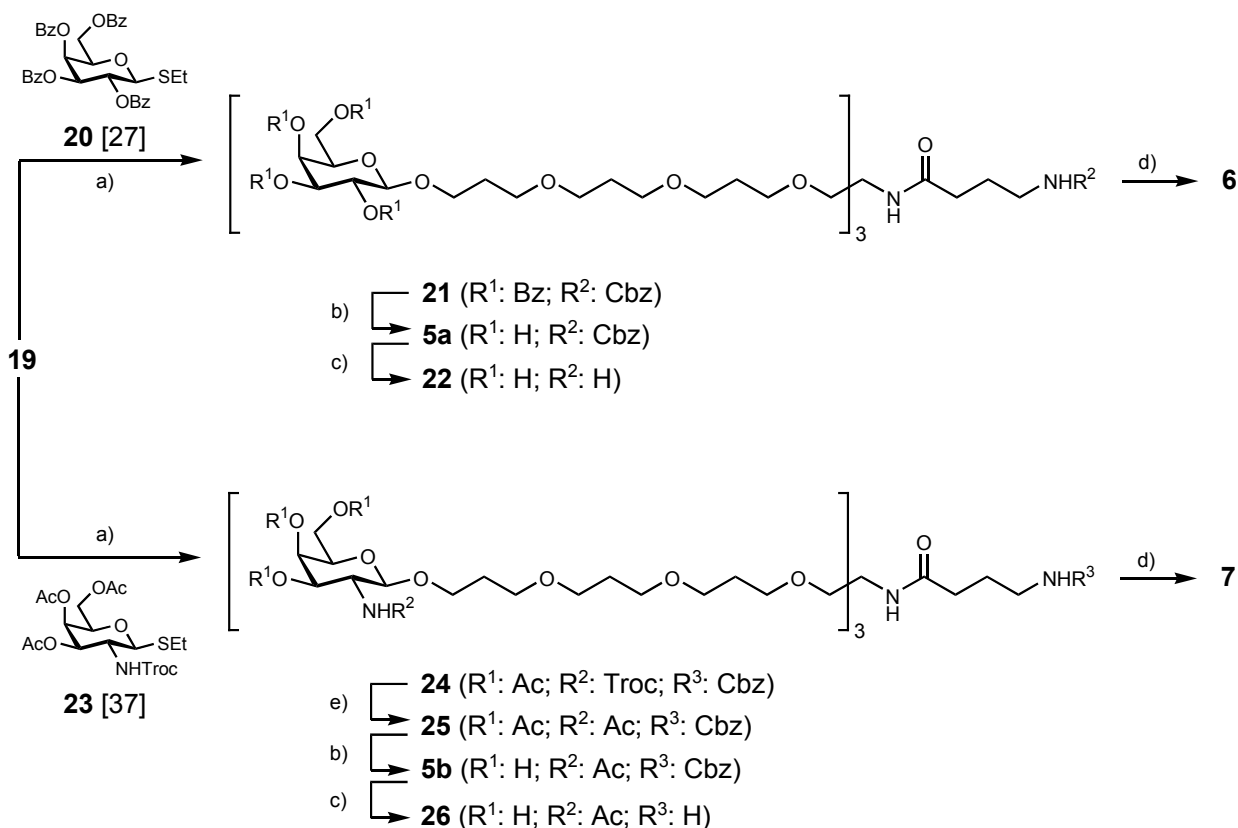
to unacceptably low yields of approx. 40% for **12**, along with a considerable amount of a tetraallylated side product. The desired triallylated compounds could finally be obtained in almost quantitative yields with only traces of *N*-allylation, by employing liquid-liquid phase transfer catalysis [35]. Thus, **12** and **14** were obtained in 95 and 90%, respectively, from the corresponding triols using allyl bromide in refluxing DCM/50% aqueous NaOH (1:1) with a catalytic amount of 15-crown-5. Oxidative hydroboration and acetylation gave **13** and **15** in excellent overall yields. The peracetylation step ( $\rightarrow$  **11**, **13** and **15**) was applied in order to facilitate purification and characterization of the intermediate triols. The subsequent deacetylation of **11** and **13** was achieved under standard Zemplén conditions. For the elaboration at the *N*-terminus of **15**, the Boc protecting group was selectively removed using 4 M HCl in dioxane leading quantitatively to **16**. Subsequent condensation with the *N*-Cbz-protected  $\gamma$ -aminobutyric acid linker **17** [36] using PyBOP in DMF/dioxane (1:3) and DIPEA as base yielded **18**. In the final step, deacetylation under Zemplén conditions furnished the trivalent glycosyl acceptor **19** in an overall yield of 27%, starting from Tris (**9**).



**Scheme 1.** (a) i. 9-BBN, THF, rt, 24 h, then aq. NaOH, H<sub>2</sub>O<sub>2</sub>, 0 °C → rt, 24 h; ii. Ac<sub>2</sub>O, pyridine, rt, 3 h, (**11**: 81%; **13**: 87%; **15**: 88%); (b) i. NaOMe, MeOH, rt, 24 h, quant.; ii. allyl bromide, 15-crown-5, DCM/50% (w/v) aqueous NaOH, reflux, 24 h, (**12**: 95%; **14**: 90%); (c) 4M HCl in dioxane, rt, 30 min, quant.; (d) PyBOP, DIPEA, dioxane/DMF, rt, 24 h, 85%; (e) NaOMe, MeOH, rt, 4 h, 90%.

Galactosylation of **19** with ethyl 2,3,4,6-tetra-*O*-benzoyl-1-thio- $\beta$ -D-galactopyranoside (**20**) [27] using DMTST as promoter furnished the trivalent intermediate **21** in a 68% yield (Scheme 2). Debenzoylation ( $\rightarrow$  **5a**) followed by cleavage of the Cbz protecting group gave **22**, which was coupled to the *N*-hydroxysuccinimidyl (NHS)-activated Alexa Fluor<sup>®</sup> 488 fluorescent label to yield compound **6** in 81% yield. Alexa Fluor<sup>®</sup> 488 was found to be the optimal fluorescent label for our purposes, combining high chemical and photostability with high fluorescence intensity. An analogous sequence of reactions was applied for the synthesis of **7**. First, the *N*-acetylgalactosamine trimer **24**

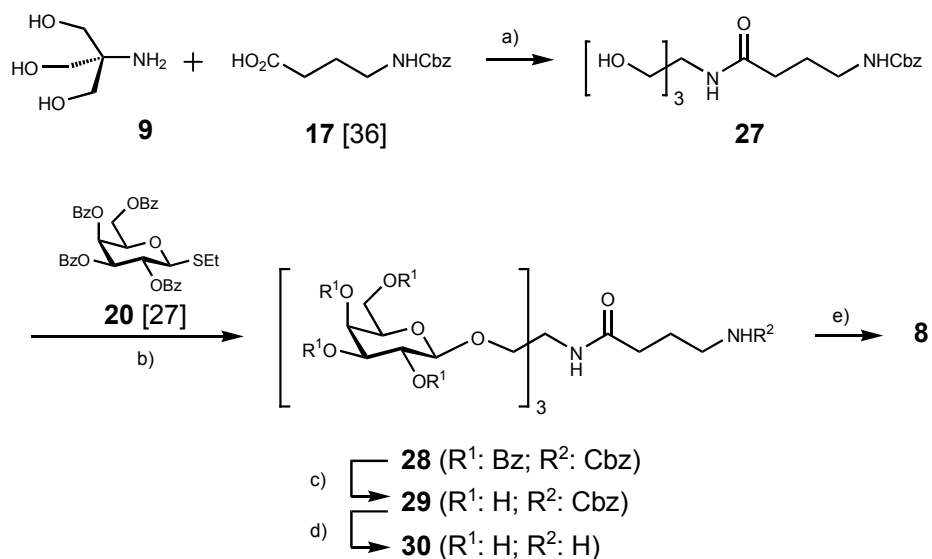
(Scheme 2) was obtained in 91% by glycosylating **19** with ethyl 3,4,6-tri-*O*-acetyl-2-deoxy-1-thio-2-(2,2,2-trichloroethoxycarbonylamino)- $\beta$ -D-galactopyranoside (**23**) [37]. After cleavage of the Troc protecting group, the free amine was directly acetylated to furnish **25**. Upon deprotection of the *N*-acetylgalactosamine moieties ( $\rightarrow$  **5b**), the Cbz group was cleaved yielding compound **26**, which was labeled with Alexa Fluor<sup>®</sup> 488 producing **7** in a 90% yield.



**Scheme 2.** (a) DMTST, 4 Å MS, DCM, 0 °C  $\rightarrow$  10 °C, 48-72 h, 68% for **21**, 91% for **24**; (b) NaOMe, MeOH/dioxane, rt, 4 h, 94% for **5a**, 72% for **5b**; (c) H<sub>2</sub>, Pd/C, EtOH/dioxane, rt, 24 h, 87% for **22**, 95% for **26**; (d) Alexa Fluor<sup>®</sup> 488-NHS, DIPEA, 4 Å MS, DMF/dioxane, rt, 4 d, 81% for **6**, 90% for **7**; (e) Zn dust, Ac<sub>2</sub>O, dioxane, rt, 24 h, 82%.

As a negative control for cellular assays, compound **8** (Scheme 3) was synthesized via acylation of Tris (**9**) with *N*-Cbz-protected  $\gamma$ -aminobutyric acid (**17**) [36] using EEDQ in pyridine [38], yielding compound **27** in a 77% yield. The

latter was then galactosylated with donor **20** [27] using NIS/TfOH as promoter to give **28** in 51%. After debenzoylation ( $\rightarrow$  **29**), the Cbz group was cleaved by hydrogenolysis to furnish compound **30**, which was subsequently coupled to the *N*-hydroxysuccinimidyl (NHS)-activated Alexa Fluor<sup>®</sup> 488 fluorescent label yielding **8** in a 96% yield.



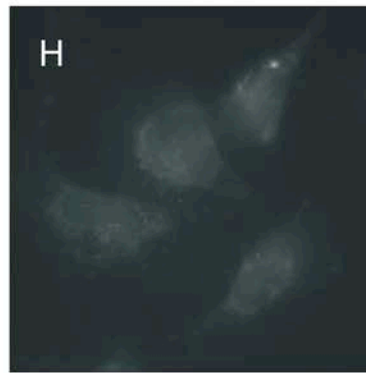
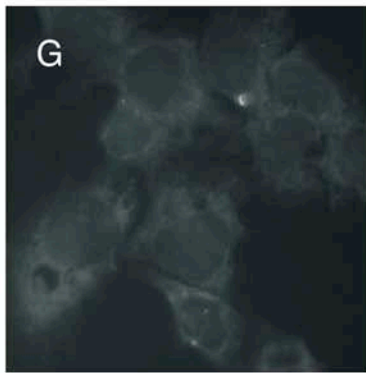
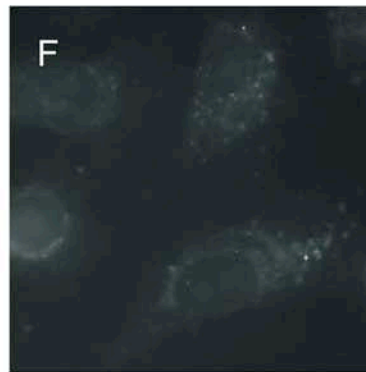
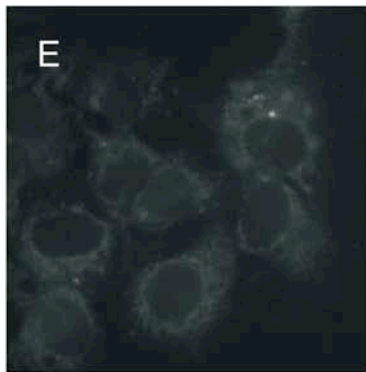
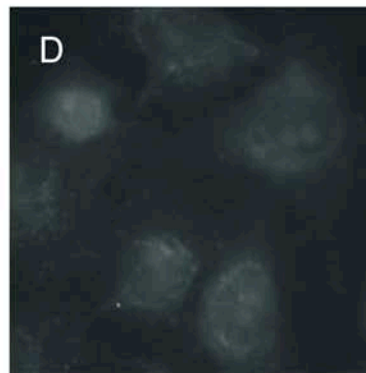
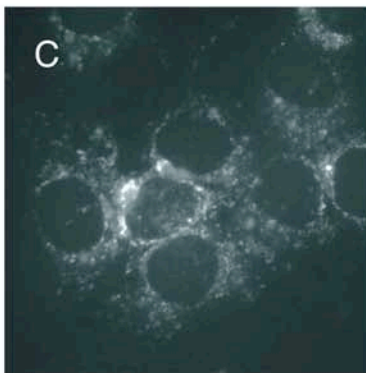
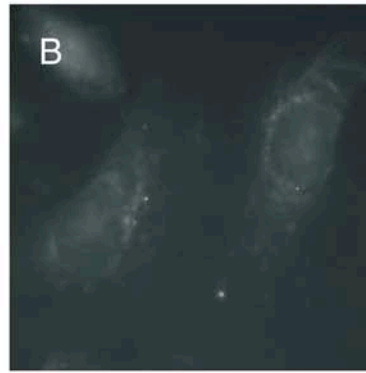
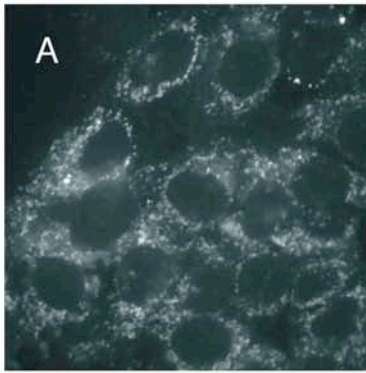
**Scheme 3.** (a) EEDQ, pyridine, 90 °C, 24 h, 77%; (b) NIS, TfOH, 4 Å MS, DCE/Et<sub>2</sub>O, 0 °C, 1 h, 51%; (c) NaOMe, dioxane/MeOH, rt, 6 h, 85%; (d) H<sub>2</sub>, Pd/C, MeOH, rt, 48 h, 87%; (e) Alexa Fluor<sup>®</sup> 488-NHS, DIPEA, 4 Å MS, DMF, rt, 4 d, 96%.

### 2.3.2 Biological Evaluation

The trivalent ligands **6-8** were examined for their selective binding to, and internalization by the ASGP-R applying fluorescence microscopy and flow cytometry. Two different cell lines of hepatic origin were used: HepG2 cells derived from a human hepatocellular carcinoma expressing the ASGP-R [39], and the human more endothelial-like SK-Hep1 cells which lack the receptor [40].

**HepG2 (with ASGP-R)**

**SK-Hep1 (without ASGP-R)**



**Figure 4.** Fluorescence microscopy images depicting the ASGP-R-specific uptake of Alexa Fluor<sup>®</sup> 488-labeled compounds. A) Compound **6** in HepG2 cells; B) Compound **6** with SK-Hep1 cells; C) Compound **7** in HepG2 cells; D) Compound **7** with SK-Hep1 cells; E) Compound **8** with HepG2 cells; F) Compound **8** with SK-Hep1 cells; G) Control HepG2 cells; H) Control SK-Hep1 cells.

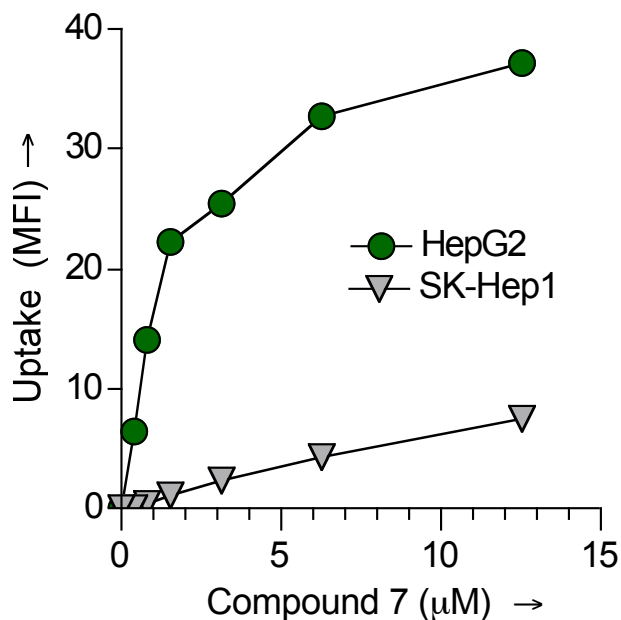
### 2.3.3 Fluorescence Microscopy

The cells were incubated with the Alexa Fluor<sup>®</sup> 488-labeled compounds **6**, **7**, or **8** for 1.5 h on ice to allow binding of the compounds to the receptor while preventing unspecific uptake. In a washing step, unbound ligand was removed, and the cells were incubated for an additional 40 min at 37 °C to allow receptor-mediated endocytosis of bound compounds to take place. The specific uptake led to punctuate staining of the cells representing endosomes containing the ligands, which were visualized by fluorescence microscopy. HepG2 cells showed specific uptake of **6** and **7**, and only negligible uptake of **8**. The fluorescent content of the endosomes can be distinctly seen (Figure 4, panels A and C) for compounds **6** and **7**, respectively. Because the cells were grown and incubated on glass cover slips, which were then mounted upside down for visualization, enriched fluorescence can only be observed in cytosolic areas which are not blocked by the nuclei. Panel E shows little or no such fluorescent vesicles, since control compound **8** was not internalized via the ASGP-R owing to insufficient spacer length. As expected, no internalization into SK-Hep1 cells (which do not express the ASGP-R) could be observed for compounds **6** and **7** (Figure 4, panels B and D). However, compound **8** showed a minor tendency to be internalized by this cell line in an ASGP-R-independent manner (Figure 4, panel F). Panels G and H show the autofluorescence of non-treated HepG2 and SK-Hep1 cells as controls.

### 2.3.4 Flow Cytometry

**Flow Cytometry:** The ASGP-R-mediated uptake of compounds **7** and **8** (negative control) was quantitatively evaluated by flow cytometry (Figures 5-6). Instead of performing the previously described steps (prebinding on ice, removal of the excess and internalization of bound compound), the cells were continuously incubated with the test compounds at 37 °C and analyzed.

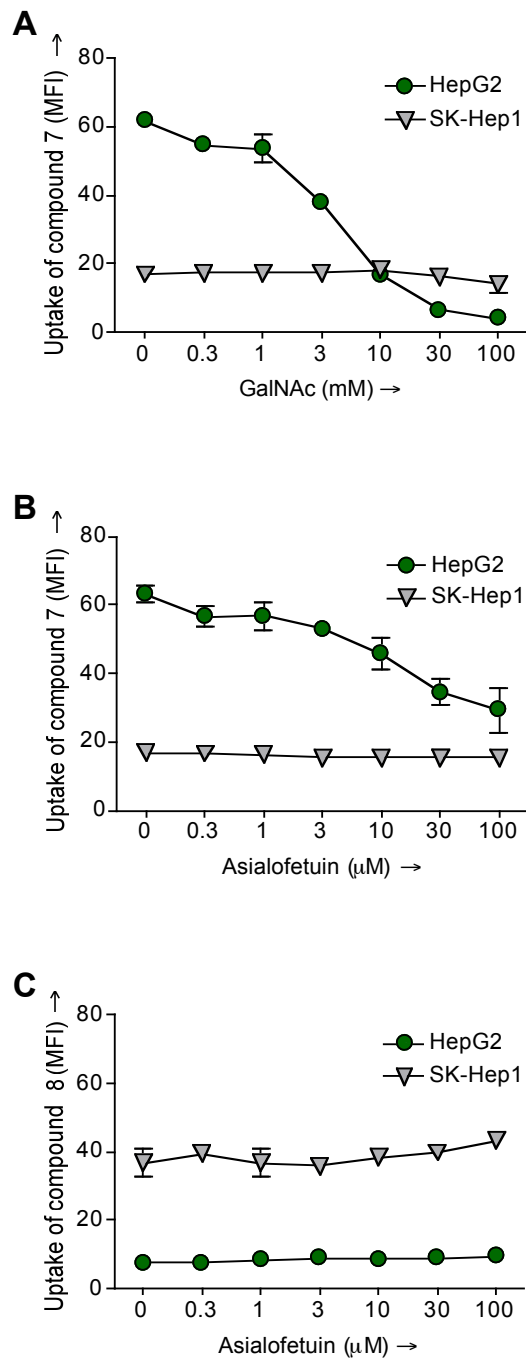
The median fluorescence intensity (MFI) of cells incubated with compound **7** at concentrations ranging from 0.4 to 12.5  $\mu\text{M}$  revealed low uptake of the compound into SK-Hep1 cells compared to HepG2 cells, in which the uptake leads to a saturation hyperbola as it is typical for a receptor-mediated process (Figure 5) [41].



**Figure 5.** Titration of compound **7**: Adherent HepG2 and SK-Hep1 cells were incubated with compound **7** at concentrations ranging from 0.4 to 12.5  $\mu\text{M}$  for 40 min at 37 °C. MFI is the shift in median fluorescence intensity from untreated to treated cells.

Uptake of compound **7** into HepG2 cells via the ASGP-R at a concentration of 10  $\mu\text{M}$  was competitively inhibited by the presence of monosaccharide ligands: GalNAc ( $\text{IC}_{50} = 4.55 \pm 0.32 \text{ mM}$ ) (Figure 6 A) and asialofetuin ( $\text{IC}_{50} = 45.60 \pm 2.70 \mu\text{M}$ ) (Figure 6 B), whereas the uptake into SK-Hep1 was low and not affected by the presence of asialofetuin.





**Figure 6.** Competitive uptake of compound **7** at a concentration of 10  $\mu$ M in the presence of either GalNAc (0.3-100 mM) (A) or asialofetuin (0.3-100  $\mu$ M) (B). The graphs represent the mean of median fluorescence intensity (MFI)  $\pm$  SD of 3 independent experiments. (C) Uptake of control compound **8** at a concentration of 10  $\mu$ M in the presence of asialofetuin (0.3 –100  $\mu$ M) into HepG2 and ASGP-R-negative SK-Hep1 cells.

In ASGP-R-bearing HepG2 cells, uptake of control compound **8** was low and proved to be unspecific as it could not be inhibited by asialofetuin, a natural high affinity ligand of the receptor (Figure 6 C). ASGP-R-negative SK-Hep1 cells, on the other hand, evinced high uptake of compound **8**, unaffected by the presence of asialofetuin (Figure 6 C) which could be explained by their high endocytic activity that is usually associated with endothelial cells.

## 2.4 Conclusion

Studies using fluorescent-labeled ligands for the ASGP-R have been carried out before. Ishihara et al. [42] prepared fluorescein isothiocyanate-labeled, galactosylated polystyrene ligands and analyzed their interaction with the ASGP-R by flow cytometry. Wu et al. [43] introduced a new synthetic route, based on solid phase peptide synthesis, towards fluorescent, synthetic, trivalent, *N*-acetylgalactosamine-terminated glycopeptides [43] as a ligands for the ASGP-R.

However, in this study we have introduced a set of novel, fluorescent, trivalent, simplified oligosaccharide mimics as ligands for the ASGP-R (**6** and **7**, Figure 3). These compounds not only comply with the afore-mentioned optimal ASGP-R ligand criteria, but also are synthetically easily accessible and hydrolytically stable. Both criteria are a prerequisite for a therapeutic application at a later stage.

Moreover, using fluorescence microscopy and flow cytometry, we have shown that compounds **6** and **7** exhibit selective uptake by the ASGP-R on HepG2 cells derived from human parenchymal liver cells – the major liver cell type. The formation of distinct endocytic vesicles could be clearly visualized. Furthermore, competition with asialofetuin, a naturally occurring serum glycoprotein and known ligand of the ASGP-R, and GalNAc confirmed the involvement of the ASGP-R in the uptake of **7**. Experiments using compound **8** have further re-enforced the

generally accepted assumption that the sugar residues have to be in an optimal spatial arrangement in order to interact selectively and with high affinity with the native ASGP-R. In final analysis, we have demonstrated that compound **7** has a high potential for use in site-specific delivery of therapeutic agents (chemotherapeutics, DNA, etc.) to the liver. The follow-up experiments are currently being performed.

## 2.5 Acknowledgement

The authors would like to thank the Swiss National Science Foundation (SNF) for funding the research. We are grateful to R. Sütterlin, Drs. M. Dürrenberger and E. Casanova for the technical support in fluorescence microscopy. Furthermore, we would like to thank M. Cavallari from the Experimental Immunology Group of Prof. Dr. Gennaro De Libero for his assistance with the flow cytometry experiments. We would also like to thank W. Kirsch for high accuracy elemental analysis measurements.

## 2.6 Experimental Section

### 2.6.1 General Methods

NMR spectra were recorded on a Bruker Avance DMX-500 (500 MHz) spectrometer. Assignment of  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra was achieved using 2D methods (COSY, HSQC, TOCSY). Chemical shifts are expressed in ppm using residual  $\text{CHCl}_3$ ,  $\text{CHD}_2\text{OD}$  and HDO as references. Optical rotations were measured using a Perkin-Elmer Polarimeter Model 341. ESI-MS spectra were measured on a Waters Micromass ZQ mass spectrometer. Reactions were monitored by TLC using glass plates coated with silica gel 60 F<sub>254</sub> (Merck) with the following mobile phases: A) petrol ether/EtOAc (4:1); B) petrol ether/EtOAc (1:1); C) petrol ether/EtOAc (3:7); D) EtOAc; E) EtOAc/MeOH (9:1); F) DCM/MeOH/H<sub>2</sub>O (10:4:0.8). Carbohydrate-containing compounds were visualized by charring with a molybdate solution (0.02 M solution of ammonium cerium sulfate dihydrate and ammonium molybdate tetrahydrate in aqueous 10% H<sub>2</sub>SO<sub>4</sub>). Compounds **6**, **7**, and **8** were visualized with UV light. All other compounds were visualized with KMnO<sub>4</sub> solution (2% KMnO<sub>4</sub> and 4% NaHCO<sub>3</sub> in

water). Column chromatography was performed on silica gel 60 (Fluka, 0.040-0.060 mm). Size exclusion chromatography was performed on Sephadex LH-20 and Sephadex G-15 (Pharmacia). Methanol (MeOH) was dried by refluxing with sodium methoxide and distilled immediately before use. Pyridine was freshly distilled under argon over CaH<sub>2</sub>. Dichloromethane (DCM) and dichloroethane (DCE) were dried by filtration over Al<sub>2</sub>O<sub>3</sub> (Fluka, type 5016 A basic). Tetrahydrofuran (THF), dioxane, diethyl ether (Et<sub>2</sub>O) and toluene were dried by refluxing with sodium and benzophenone. Dry DMF was purchased from Fluka (absolute, ≥99.8%) and was further dried over powdered 4 Å molecular sieves. Molecular sieves (4 Å) were activated in vacuo at 500 °C for 2 h immediately before use. Alexa Fluor<sup>®</sup> 488 carboxylic acid succinimidyl ester (A20000, mixture of isomers) was purchased from Molecular Probes, Eugene, Oregon, USA. Zinc dust was activated according to standard procedures [44].

All cell culture media, supplements and phosphate buffered saline (PBS) were purchased from Invitrogen, except collagen type S from rat's tail was obtained from Roche Applied Science. Paraformaldehyde, NaN<sub>3</sub> and *N*-propyl gallate were obtained from Fluka. Bovine serum albumin (BSA) was from Sigma and Mowiol 4-88 from Hoechst. HepG2 (human hepatocellular carcinoma) and SK-Hep1 (human liver adenocarcinoma) cell-lines were obtained from DSMZ (Deutsche Sammlung für Mikroorganismen und Zellkulturen). Both cell lines were propagated in Dulbecco's modified Eagle's medium (DMEM) high-glucose, without phenol red, supplemented with fetal bovine serum (FBS, 10%) 2 mM L-glutamine, 100 U/ml penicillin and 100 µg/ml streptomycin (complete medium). During the incubation steps of the cells outside the incubator, medium with a CO<sub>2</sub>-independent buffer system was used (DMEM high-glucose, without phenol red and FBS, containing 25 mM HEPES).

## 2.6.2 Abbreviations

NIS, *N*-iodosuccinimide; TfOH, trifluoromethanesulfonic acid; NHS, *N*-hydroxysuccinimide; PyBOP, benzotriazol-1-yl-oxytripyrrolidino-phosphonium hexafluorophosphate; Tris, 2-Amino-2-(hydroxymethyl)-1,3-propanediol; EEDQ, ethyl 1,2-dihydro-2-ethoxy-1-quinolinecarboxylate; DMTST, dimethyl(methylthio)sulfonium trifluoromethanesulfonate; DIPEA, diisopropylethylamine; HEPES, 4-(2-hydroxyethyl)-1-piperazine-1-ethanesulfonic acid.

## 2.6.3 General procedure A: preparation of triacetylated compounds **11**, **13** and **15**

To the corresponding triallylated compound (3.54 mmol) was added 9-BBN (0.5 M in THF, 38 ml) dropwise. The solution was then stirred at rt under argon for 24 h. The mixture was cooled to 0°C, and aqueous NaOH (3 M, 39 ml) was added dropwise, followed by the dropwise addition of H<sub>2</sub>O<sub>2</sub> (30%, 8.9 ml). The resultant mixture was stirred vigorously at rt for 24 h. The mixture was saturated with K<sub>2</sub>CO<sub>3</sub>, and the organic layer was separated. The aqueous layer was then extracted with THF (3 × 80 ml), and the combined organic layers were dried (Na<sub>2</sub>SO<sub>4</sub>) and concentrated under reduced pressure. The residue was dissolved in pyridine (33.5 ml), acetic anhydride (33.5 ml) was added, and the mixture was stirred at rt for 3 h. The mixture was co-evaporated with toluene (200 ml), and the resultant syrup was purified by silica gel chromatography to afford compound **11**, **13** or **15** as an oils.

**Tris(5-acetoxy-2-oxapentyl)-*N*-(*tert*-butyloxycarbonyl)-methylamine (11):** According to general procedure A, compound **10** [34] (1.21 g, 3.54 mmol) was reacted with 9-BBN (0.5 M in THF, 38 ml), and then treated with aqueous NaOH

(3 M, 39 ml) and H<sub>2</sub>O<sub>2</sub> (30%, 8.9 ml). After peracetylation, work-up and chromatography on silica gel (petrol ether/EtOAc 8:2 → 7:3; R<sub>f</sub> 0.52 B) 1.49 g (81%) of **11** were obtained. <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ = 1.42 (s, 9H, CMe<sub>3</sub>), 1.88 (m, 6H, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc), 2.04 (s, 9H, 3 × OAc), 3.50 (t, *J* = 6.1 Hz, 6H, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc), 3.63 [s, 6H, C(CH<sub>2</sub>O)<sub>3</sub>], 4.13 (t, *J* = 6.5 Hz, 6H, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc), 4.90 (s, 1H, NH); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ = 21.1 (3C, 3 × CH<sub>3</sub>C=O), 28.4 (3C, CMe<sub>3</sub>), 28.9 (3C, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc), 58.5 [C(CH<sub>2</sub>O)<sub>3</sub>], 61.7 (3C, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc), 67.8 (3C, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc), 69.5 [3C, C(CH<sub>2</sub>O)<sub>3</sub>], 79.1 (CMe<sub>3</sub>), 154.8 [N-(C=O)O], 171.1 (3C, 3 × CH<sub>3</sub>C=O); Anal. Calcd for C<sub>24</sub>H<sub>43</sub>NO<sub>11</sub>: C, 55.26; H, 8.31; N, 2.69. Found: C, 55.20; H, 8.24; N, 2.70.

**Tris(9-acetoxy-2,6-dioxanonyl)-*N*-(*tert*-butyloxycarbonyl)-methylamine (13):**

According to general procedure A, compound **12** (1.82 g, 3.54 mmol) was reacted with 9-BBN (0.5 M in THF, 38 ml), and then treated with aqueous NaOH (3 M, 39 ml) and H<sub>2</sub>O<sub>2</sub> (30%, 8.9 ml). After peracetylation, work-up and chromatography on

silica gel (petrol ether/EtOAc 3:1 → 3:2; R<sub>f</sub> 0.32 B) 2.14 g (87%) of **13** were obtained. <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ = 1.42 (s, 9H, CMe<sub>3</sub>), 1.78-1.91 [m, 12H, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OAc], 2.05 (s, 9H, 3 × OAc), 3.45-3.50 (m, 18H, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc), 3.62 [s, 6H, C(CH<sub>2</sub>O)<sub>3</sub>], 4.15 (t, *J* = 6.5 Hz, 6H, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc), 4.93 (s, 1H, NH); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ = 21.0 (3C, 3 × CH<sub>3</sub>C=O), 28.4 (3C, CMe<sub>3</sub>), 29.0, 29.9 [6C, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OAc], 58.4 [C(CH<sub>2</sub>O)<sub>3</sub>], 61.8 (3C, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc), 67.2, 67.9, 68.3, (9C, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc), 69.5 [3C, C(CH<sub>2</sub>O)<sub>3</sub>], 79.1 (CMe<sub>3</sub>), 154.8 [N-(C=O)O], 171.1 (3C, 3 × CH<sub>3</sub>C=O); Anal. Calcd. for C<sub>33</sub>H<sub>61</sub>NO<sub>14</sub>: C, 56.96; H, 8.84; N, 2.01. Found: C, 57.67; H, 8.80; N, 2.56.

**Tris(13-acetoxy-2,6,10-trioxatridecyl)-*N*-(*tert*-butyloxycarbonyl)-methylamine (15):**

According to general procedure A, compound **14** (2.44 g, 3.54 mmol)

was reacted with 9-BBN (0.5 M in THF, 38 ml), and then treated with aqueous NaOH (3 M, 39 ml) and H<sub>2</sub>O<sub>2</sub> (30%, 8.9 ml). After peracetylation, work-up and chromatography on silica gel (petrol ether/EtOAc 1:1 → 3:7; R<sub>f</sub> 0.1 B) 2.71 g (88%) of **15** were obtained. <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ = 1.39 (s, 9H, CMe<sub>3</sub>), 1.75-1.88 [m, 18H, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OAc], 2.01 (s, 9H, 3 × OAc), 3.41-3.47 (m, 30H, 3 × [(OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc], 3.59 [s, 6H, C(CH<sub>2</sub>O)<sub>3</sub>], 4.12 (t, *J* = 6.5 Hz, 6H, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc), 4.90 (s, 1H, NH); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ = 20.9 (3C, 3 × CH<sub>3</sub>C=O), 28.4 (3C, CMe<sub>3</sub>), 28.9, 29.6, 29.9 [9C, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OAc], 58.4 [C(CH<sub>2</sub>O)<sub>3</sub>], 61.7 (3C, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc), 67.2, 67.7, 67.8, 67.9, 68.3, 68.4 [15C, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc], 69.4 [3C, C(CH<sub>2</sub>O)<sub>3</sub>], 78.8 (CMe<sub>3</sub>), 154.7 [N-(C=O)O], 171.0 (3C, 3 × CH<sub>3</sub>C=O); Anal. Calcd. for C<sub>42</sub>H<sub>79</sub>NO<sub>17</sub>: C, 57.98; H, 9.15; N, 1.61; O, 31.26. Found: C, 58.08; H, 9.17; N, 1.70; O, 31.11.

#### 2.6.4 General procedure B: preparation of triallylated compounds **12** and **14**

The corresponding triacetylated compound (4.6 mmol) was dissolved in a solution of sodium methoxide in MeOH (0.1 M, 40 ml), and the resultant solution was stirred at rt for 4 h under argon. The solution was neutralized with Dowex 50X8 (H<sup>+</sup>-form), and the solvent was removed under reduced pressure to afford the desired product in a quantitative yield as a colorless oil, which was used without further purification.

The corresponding triol (1 mmol) was then dissolved in DCM (5 ml) and the solution was added to a mixture of 50% aqueous NaOH (16 ml, w/v) and 15-crown-5 (19.8 μl, 0.1 mmol). Allyl bromide (1.64 ml, 19.1 mmol) was then added, and the resultant mixture was refluxed with vigorous stirring for 24 h. The mixture was cooled, and the DCM (top) layer was separated, dried with Na<sub>2</sub>SO<sub>4</sub>, and the



solvent evaporated in vacuo. The resultant syrup was purified by silica gel chromatography to yield compound **12** or **14** as a yellow oil.

**Tris(5-allyloxy-2-oxapentyl)-N-(tert-butyloxycarbonyl)-methylamine (12):**

According to general procedure B, compound **11** (2.39 g, 4.6 mmol) was deacetylated under Zemplén conditions, and after work-up, reacted with allyl bromide (7.5 ml, 87.9 mmol) under phase transfer catalysis conditions. After work-up and chromatography on silica gel (petrol ether/EtOAc 19:1 → 9:1 → 4:1; R<sub>f</sub> 0.25 A) 2.25 g (95%) of **12** were obtained. <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ = 1.38 (s, 9H, CMe<sub>3</sub>), 1.79 (quintet, J = 6.3 Hz, 6H, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>O), 3.43-3.48 (m, 12H, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>O), 3.59 [s, 6H, C(CH<sub>2</sub>O)<sub>3</sub>], 3.91 (m, 6H, 3 × CH<sub>2</sub>CH=CH<sub>2</sub>), 4.90 (s, 1H, NH), 5.17 (m, 6H, 3 × CH<sub>2</sub>CH=CH<sub>2</sub>), 5.86 (m, 3H, 3 × CH<sub>2</sub>CH=CH<sub>2</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ = 28.4 (3C, CMe<sub>3</sub>), 30.0 (3C, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>O), 58.5 [C(CH<sub>2</sub>O)<sub>3</sub>], 67.4, 68.2 (6C, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>O), 69.6 [3C, C(CH<sub>2</sub>O)<sub>3</sub>], 71.9 (3C, 3 × CH<sub>2</sub>CH=CH<sub>2</sub>), 78.8 (CMe<sub>3</sub>), 116.7 (3C, 3 × CH<sub>2</sub>CH=CH<sub>2</sub>), 135.0 (3C, 3 × CH<sub>2</sub>CH=CH<sub>2</sub>), 154.8 [N-(C=O)O]; Anal. Calcd. for C<sub>27</sub>H<sub>49</sub>NO<sub>8</sub>: C, 62.89; H, 9.58; N, 2.72; O, 24.82. Found: C, 62.66; H, 9.61; N, 2.68; O, 24.87.

**Tris(9-allyloxy-2,6-dioxanonyl)-N-(tert-butyloxycarbonyl)-methylamine (14):**

According to general procedure B, compound **13** (3.2 g, 4.6 mmol) was deacetylated under Zemplén conditions, and after work-up, reacted with allyl bromide (7.5 ml, 87.9 mmol) under phase transfer catalysis conditions. After work-up and chromatography on silica gel (petrol ether/EtOAc 4:1 → 1:1; R<sub>f</sub> 0.64 B) 2.86 g (90%) of **14** were obtained. <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ = 1.42 (s, 9H, CMe<sub>3</sub>), 1.78-1.87 (m, 12H, 6 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>O), 3.45-3.52 (m, 24H, 6 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>O), 3.62 [s, 6H, C(CH<sub>2</sub>O)<sub>3</sub>], 3.96 (m, 6H, 3 × CH<sub>2</sub>CH=CH<sub>2</sub>), 4.93 (s, 1H, NH), 5.22 (m, 6H, 3 × CH<sub>2</sub>CH=CH<sub>2</sub>), 5.91 (m, 3H, 3 × CH<sub>2</sub>CH=CH<sub>2</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ = 28.4 (3C, CMe<sub>3</sub>), 29.9, 30.1 (6C, 6 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>O), 58.4 [C(CH<sub>2</sub>O)<sub>3</sub>], 67.3, 67.8, 68.4 (12C, 6 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>O),

69.5 [3C, C(CH<sub>2</sub>O)<sub>3</sub>], 71.8 (3C, 3 × CH<sub>2</sub>CH=CH<sub>2</sub>), 79.0 (CMe<sub>3</sub>), 116.7 (3C, 3 × CH<sub>2</sub>CH=CH<sub>2</sub>), 134.9 (3C, 3 × CH<sub>2</sub>CH=CH<sub>2</sub>), 155.0 [N-(C=O)O]; Anal. Calcd. for C<sub>36</sub>H<sub>67</sub>NO<sub>11</sub>: C, 62.67; H, 9.79; N, 2.03; O, 25.51. Found: C, 62.67; H, 9.72; N, 2.10; O, 25.37.

**Tris(13-acetoxy-2,6,10-trioxatridecyl)-methylamine hydrochloride (16):**

Compound **15** (680 mg, 0.781 mmol) was dissolved in 4 M HCl in dioxane (10 ml), and the resultant mixture was stirred at rt under argon for 30 min. The solvent was removed in vacuo to yield **16** (630 mg, quantitative) as an oil. <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ = 1.68-1.77 [m, 18H, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OAc], 1.92 (s, 9H, 3 × OAc), 3.35-3.44 [m, 36H, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc, C(CH<sub>2</sub>O)<sub>3</sub>], 4.01 (t, *J* = 6.5 Hz, 6H, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ = 20.5 (3C, 3 × CH<sub>3</sub>C=O), 28.6, 29.3, 29.6 [9C, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OAc], 59.1 [C(CH<sub>2</sub>O)<sub>3</sub>], 61.6 (3C, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc), 66.7, 67.0, 67.4, 67.6, 68.3, 68.7 [18C, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc, C(CH<sub>2</sub>O)<sub>3</sub>], 171.5 (3C, 3 × CH<sub>3</sub>C=O); ESI-MS: Calcd. for C<sub>37</sub>H<sub>72</sub>NO<sub>15</sub> (M+H)<sup>+</sup>: 770.49 ; Found m/z 770.54.

***N*-{Tris[13-acetoxy-2,6,10-trioxatridecyl]methyl}-4-(benzyloxycarbonyl-**

**amino)-butyramide (18):** Compound **16** (239 mg, 0.297 mmol), **17** [36] (70.4 mg, 0.297 mmol) and PyBOP (186 mg, 0.357 mmol) were dissolved in dioxane/DMF (4 ml, 3:1 v/v), and DIPEA (229 μl, 1.34 mmol) was added. The mixture was stirred at rt under argon for 24 h. The resultant solution was partitioned between DCM (15 ml) and H<sub>2</sub>O (15 ml). The DCM layer was separated, and the aqueous phase was extracted with DCM (25 ml). The DCM fractions were combined, dried (Na<sub>2</sub>SO<sub>4</sub>), and the solvent was removed under reduced pressure. The resultant syrup was purified by silica gel chromatography (petrol ether/EtOAc 1:1→3:7→0:1) to afford **18** (250 mg, 85%, R<sub>f</sub> 0.22 D) as a yellow oil. <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ = 1.76-1.90 (m, 20H, 3 × [(OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OAc, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O], 2.03 (s, 9H, 3 × OAc), 2.07-2.19 (m,

3H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.20-3.24 (m, 3H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.41-3.50 [m, 30H, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc], 3.66 [s, 6H, C(CH<sub>2</sub>O)<sub>3</sub>], 4.13 (t, *J* = 6.5 Hz, 6H, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc), 5.07 (s, 2H, CH<sub>2</sub>Ph), 5.28 (br s, 1H, NH, Cbz), 5.86 (s, 1H, NH, Tris), 7.28-7.34 (m, 5H, C<sub>6</sub>H<sub>5</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ = 20.9 (3C, 3 × CH<sub>3</sub>C=O), 25.8 (1C, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 29.0, 29.8, 30.1 [9C, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OAc], 34.4 (1C, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 40.3 (1C, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 59.8 [C(CH<sub>2</sub>O)<sub>3</sub>], 61.8 (3C, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc), 66.5 (1C, CH<sub>2</sub>Ph), 67.3, 67.7, 67.9, 68.4 [15C, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OAc], 69.1 [3C, C(CH<sub>2</sub>O)<sub>3</sub>], 128.0, 128.1, 128.5, 136.7 (6C, C<sub>6</sub>H<sub>5</sub>), 156.6 [N-(C=O)O], 171.1 (3C, 3 × CH<sub>3</sub>C=O), 172.4 (C=O, amide); Anal. Calcd. for C<sub>49</sub>H<sub>84</sub>N<sub>2</sub>O<sub>18</sub>: C, 59.50; H, 8.56; N, 2.83; O, 29.11. Found: C, 59.12; H, 8.36; N, 2.98; O, 29.56.

**4-(Benzyloxycarbonylamino)-*N*-{tris[13-hydroxy-2,6,10-trioxatridecyl]-methyl}-butyramide (19)**: Compound **18** (231 mg, 0.233 mmol) was dissolved in a solution of sodium methoxide in dry methanol (0.05 M, 20 ml), and the resultant solution was stirred at rt under argon for 4 h. The reaction mixture was neutralized with Dowex 50X8 (H<sup>+</sup>-form), and the solvent was removed in vacuo. The resultant oil was purified by silica gel chromatography (EtOAc/MeOH 95:5→9:1) to afford **19** (181 mg, 90%, R<sub>f</sub> 0.2 E) as an oil. <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ = 1.76-1.83 [m, 20H, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OH, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O], 2.18 (t, *J* = 6.8 Hz, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 2.62 (bs, 3H, 3 × OH), 3.21 (m, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.43-3.51 (m, 24H, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OH), 3.58 [t, *J* = 5.8 Hz, 6H, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>(OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OH], 3.67 [s, 6H, C(CH<sub>2</sub>O)<sub>3</sub>], 3.73 (t, *J* = 5.7 Hz, 6H, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OH), 5.07 (s, 2H, CH<sub>2</sub>Ph), 5.45 (s, 1H, NH, Cbz), 6.00 (s, 1H, NH, Tris), 7.28-7.34 (m, 5H, C<sub>6</sub>H<sub>5</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ = 25.7 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 29.8, 30.0, 32.0 [9C, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OH], 34.3 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 40.3 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 59.9 [C(CH<sub>2</sub>O)<sub>3</sub>], 61.7 (3C, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OH), 66.5 (CH<sub>2</sub>Ph), 67.7, 68.1, 68.3 (12C, 3 ×

OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OH), 69.2 [3C, C(CH<sub>2</sub>O)<sub>3</sub>], 69.7 [3C, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>(OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OH], 128.0, 128.1, 128.5, 136.7 (6C, C<sub>6</sub>H<sub>5</sub>), 156.7 [N-(C=O)O], 173.5 (C=O, amide); ESI-MS: Calcd. for C<sub>43</sub>H<sub>78</sub>N<sub>2</sub>O<sub>15</sub>Na (M+Na)<sup>+</sup>: 885.53; Found m/z 885.68.

***N*-{Tris[13-(2,3,4,6-tetra-*O*-benzoyl-β-*D*-galactopyranosyloxy)-2,6,10-trioxatridecyl]methyl}-(4-benzyloxycarbonylamino)-butyramide (21):** Compounds **19** (151 mg, 0.173 mmol) and ethyl 2,3,4,6-tetra-*O*-benzoyl-1-thio-β-*D*-galactopyranoside (**20**) [27] (670 mg, 1.04 mmol) were dissolved in dry DCM (10 ml), and the mixture was stirred with 4 Å molecular sieves (500 mg) at rt under argon for 2 h. The mixture was cooled to 0 °C, and DMTST (538 mg, 2.08 mmol) was added. The reaction was stirred at 0 °C for 24 h, and then at 10 °C for another 24 h under argon. The mixture was then filtered and extracted with aqueous NaHCO<sub>3</sub> solution (10 ml, 1 M) and brine (10 ml). The organic phase was dried (Na<sub>2</sub>SO<sub>4</sub>), and the solvent was removed under reduced pressure. The resultant syrup was purified by silica gel chromatography (EtOAc/petrol ether 1:1 → 7:3 → 1:0) to afford the desired product **21** (301 mg, 68%, R<sub>f</sub> 0.15 C) as a colorless solid. [α]<sub>D</sub> = +72.9 (c 1, CHCl<sub>3</sub>); <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ = 1.64-1.69, 1.76-1.85 [m, 20H, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OGal, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O], 2.16 (m, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.18 (m, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.22-3.47 [m, 30H, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGal], 3.66, [s, 6H, C(CH<sub>2</sub>O)<sub>3</sub>], 3.68, 4.01-4.06 (m, 6H, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGal), 4.32 (m, 3H, 3 × H5-Gal), 4.40 (dd, J<sub>5,6</sub> = 6.7, J<sub>6,6'</sub> = 11.3 Hz, 3H, 3 × H6-Gal), 4.68 (dd, J<sub>5,6'</sub> = 6.4, J<sub>6,6'</sub> = 11.2 Hz, 3H, 3 × H6'-Gal), 4.81 (d, J<sub>1,2</sub> = 7.9 Hz, 3H, 3 × H1-Gal), 5.07 (s, 2H, CH<sub>2</sub>Ph), 5.27 (bs, 1H, NH, Cbz), 5.61 (m, 3H, 3 × H3-Gal), 5.78 (m, 3H, 3 × H2-Gal), 5.86 (s, 1H, NH, Tris), 5.99 (m, 3H, 3 × H4-Gal), 7.22-7.26, 7.28-7.35, 7.37-8.09 (m, 65H, 13 × C<sub>6</sub>H<sub>5</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ = 26.6 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 29.7, 29.8, 29.9 [9C, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OGal], 35.1 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 42.2 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 59.8 [C(CH<sub>2</sub>O)<sub>3</sub>], 61.9 (3C, 3 × C6-Gal), 67.0 (CH<sub>2</sub>Ph), 67.4, 67.8, 67.9 [12C, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGal], 68.1 (3C, 3 × C4-Gal),

68.4 (3C, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGal), 69.1 [3C, C(CH<sub>2</sub>O)<sub>3</sub>], 69.8 (3C, 3 × C2-Gal), 71.2 (3C, 3 × C3-Gal), 71.7 (3C, 3 × C5-Gal), 101.8 (3C, 3 × C1-Gal), 128.0-136.7 (78C, C<sub>6</sub>H<sub>5</sub>), 165.2-166.0 (13C, 13 C=O); Anal. Calcd. for C<sub>145</sub>H<sub>156</sub>N<sub>2</sub>O<sub>42</sub>: C, 67.01; H, 6.05; N, 1.08. Found: C, 66.51; H, 6.12; N, 1.13.

**4-(Benzyloxycarbonylamino)-N-{tris[13-(β-D-galactopyranosyloxy)-2,6,10-trioxatridecyl]methyl}-butyramide (5a):** Compound **21** (30 mg, 0.015 mmol) was dissolved in dry dioxane (1 ml), and a solution of sodium methoxide in methanol (0.1 M, 1 ml) was added. The resultant mixture was stirred at rt under argon for 4 h, after which it was neutralized with Dowex 50X8 (H<sup>+</sup>-form), filtered, and the solvent was removed under reduced pressure. The residue was purified by silica gel chromatography (DCM/MeOH/H<sub>2</sub>O 10:3:0 → 10:4:0 → 10:4:0.1 → 10:4:0.2 → 10:4:0.4) to afford **5a** (14.6 mg, 94%, R<sub>f</sub> 0.2 F) as a colorless solid. [α]<sub>D</sub> = -5.73 (c 0.96, MeOH); <sup>1</sup>H-NMR (500 MHz, MeOD): δ = 1.73-1.88 [m, 20H, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OGal, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O], 2.19 (t, J = 7.3 Hz, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.15 (t, J = 6.8 Hz, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.44-3.75 [m, 54H, 3 × H2-Gal, 3 × H3-Gal, 3 × H5-Gal, 3 × H6-Gal, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGal, C(CH<sub>2</sub>O)<sub>3</sub>, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGal-H<sub>a</sub>], 3.82 (m, 3H, 3 × H4-Gal), 3.95 (m, 3H, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGal-H<sub>b</sub>), 4.20 (d, J<sub>1,2</sub> = 7.4 Hz, 3H, 3 × H1-Gal), 5.07 (s, 2H, CH<sub>2</sub>Ph), 7.29-7.35 (m, 5H, C<sub>6</sub>H<sub>5</sub>); <sup>13</sup>C NMR (125 MHz, MeOD): δ = 27.5 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 31.0, 31.1 [9C, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OGal], 35.0 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 41.1 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 61.6 [C(CH<sub>2</sub>O)<sub>3</sub>], 62.4 (3C, 3 × C6-Gal), 67.4 (CH<sub>2</sub>Ph), 67.9 (3C, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGal), 68.8 (3C, 3 × C[CH<sub>2</sub>O]<sub>3</sub>), 69.4 (3C, 3 × C4-Gal), 68.8, 70.2, 72.6, 75.0, 76.6 [24C, 3 × C2-Gal, 3 × C3-Gal, 3 × C5-Gal, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGal], 105.1 (3C, 3 × C1-Gal), 128.9-129.5, 138.4 (6C, C<sub>6</sub>H<sub>5</sub>), 158.9 [N-(C=O)O], 175.6 (C=O); ESI-MS: Calcd. for C<sub>61</sub>H<sub>109</sub>N<sub>2</sub>O<sub>30</sub> (M+H)<sup>+</sup>: 1349.71; Found m/z 1349.87.

**4-Amino-N-{tris[13-( $\beta$ -D-galactopyranosyloxy)-2,6,10-trioxatridecyl]methyl}-butyramide (22):** Compound **5a** (25 mg, 18.5  $\mu$ mol) was dissolved in ethanol/dioxane (2 ml, 1:1 v/v), and Pd/C (10% Pd, 20 mg) was added. The mixture was vigorously stirred under a H<sub>2</sub> atmosphere (1 atm) at rt for 24 h. The mixture was then diluted with ethanol, filtered and concentrated in vacuo to yield **22** as a colorless solid (19.5 mg, 87%).  $[\alpha]_D = -5.8$  (c 1, MeOH); <sup>1</sup>H-NMR (500 MHz, MeOD):  $\delta = 1.79$ -1.89 [m, 20H, 3  $\times$  (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OGal, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O], 2.31 (t,  $J = 7.1$  Hz, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 2.87 (t,  $J = 7.3$  Hz, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.44-3.56 [m, 39H, 3  $\times$  H2-Gal, 3  $\times$  H3-Gal, 3  $\times$  H5-Gal, 3  $\times$  (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGal], 3.60-3.75 (m, 15H, 3  $\times$  H6-Gal, C[CH<sub>2</sub>O]<sub>3</sub>, 3  $\times$  OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGal-H<sub>a</sub>), 3.82 (m, 3H, 3  $\times$  H4-Gal), 3.96, (m, 3H, 3  $\times$  OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGal-H<sub>b</sub>), 4.2 (d,  $J_{1,2} = 7.2$  Hz, 3H, 3  $\times$  H1-Gal); <sup>13</sup>C NMR (125 MHz, MeOD):  $\delta = 26.9$  (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 30.7, 31.0, 31.1 [9C, 3  $\times$  (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>Gal], 34.6 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 48.3 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 62.5 (3C, 3  $\times$  C6-Gal), 67.9, 68.8, 68.9, 69.4, 69.7 [21C, 3  $\times$  (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGal, 3  $\times$  OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGal, C(CH<sub>2</sub>O)<sub>3</sub>], 70.2 (3C, 3 $\times$  C4-Gal), 72.6 (3C, 3 $\times$  C2-Gal), 75.0 (3C, 3 $\times$  C3-Gal), 76.6 (3C, 3 $\times$  C5-Gal), 105.1 (3C, 3  $\times$  C1-Gal); ESI-MS: Calcd. for C<sub>53</sub>H<sub>103</sub>N<sub>2</sub>O<sub>28</sub> (M+H)<sup>+</sup>: 1215.67; Found m/z 1215.91.

**Fluorescent-labeled, Gal-terminated compound (6):** A stock solution containing compound **22** (10 mg, 8.23  $\mu$ mol), DIPEA (20  $\mu$ l, 156  $\mu$ mol) and 4 Å molecular sieves (25 mg) in dry DMF/dioxane (1 ml, 1:1) was stirred at rt under argon for 2 h. The solution (500  $\mu$ l) was transferred to a small vial containing Alexa Fluor<sup>®</sup> 488-NHS (1 mg, 1.55  $\mu$ mol), 4Å molecular sieves (25 mg), and a stirring bar. The resultant mixture was stirred in the dark at rt under argon for 4 d. The mixture was then diluted with MeOH, filtered, and the solvents were removed in vacuo. The residue was purified by gel filtration on a Sephadex LH-20 column (2.5  $\times$  35 cm) using MeOH as eluant, then on an RP-18 column (H<sub>2</sub>O/MeOH stepwise gradient 1:0 to 1:1) to yield **6** (2.2 mg, 81%) as a red solid

after final lyophilization from water. ESI-MS: Calcd for  $C_{74}H_{112}N_4O_{38}S_2^{2-}$  (M/2)<sup>-</sup>: 864.32; Found m/z 864.85.

***N*-(Tris{13-[3,4,6-tri-*O*-acetyl-2-(2,2,2-trichloroethoxycarbonylamino)- $\beta$ -D-galactopyranosyloxy]-2,6,10-trioxatridecyl}methyl)-(4-benzyloxycarbonylamino)-butyramide (24)**: Compound **19** (15 mg, 17.3  $\mu$ mol) and ethyl 3,4,6-tri-*O*-acetyl-2-deoxy-1-thio-2-(2,2,2-trichloroethoxycarbonylamino)- $\beta$ -D-galactopyranoside (**23**) [37] (54.6 mg, 104  $\mu$ mol) were dissolved in dry DCM (1 ml) and stirred with 4 Å molecular sieves (50 mg) at rt under argon for 2 h. The suspension was cooled to 0 °C, and DMTST (53.8 mg, 208  $\mu$ mol) was added. The mixture was stirred at 0 °C for 24 h, then at 10 °C for 48 h. The reaction was quenched with triethylamine (50  $\mu$ l), diluted with DCM (5 ml), washed with brine (5 ml), dried ( $Na_2SO_4$ ) and the solvent was removed in vacuo. The residue was purified by silica gel chromatography (petrol ether/EtOAc 1:4  $\rightarrow$  1:9) to afford **24** (35.1 mg, 91%,  $R_f$  0.32 D) as a colorless solid.  $[\alpha]_D = -3.6$  (c 1.76,  $CHCl_3$ );  $^1H$ -NMR (500 MHz,  $CDCl_3$ ):  $\delta = 1.76$ -1.84 [m, 20H, 3  $\times$  (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OGalN, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O], 1.96 (s, 9H, 3  $\times$  OAc), 2.03 (s, 9H, 3  $\times$  OAc), 2.14 (s, 9H, 3  $\times$  OAc), 2.20 (m, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.22 (m, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.43-3.56 [m, 30H, 3  $\times$  (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGalN], 3.60 (m, 3H, 3  $\times$  CH<sub>2</sub>OGalN-H<sub>a</sub>), 3.66, [s, 6H, C(CH<sub>2</sub>O)<sub>3</sub>], 3.85-3.90 (m, 3  $\times$  H2-GalN, 6H, 3  $\times$  H5-GalN), 3.97 (m, 3H, 3  $\times$  CH<sub>2</sub>OGalN-H<sub>b</sub>), 4.05-4.19 (m, 6H, 3  $\times$  H6-GalN, 3  $\times$  H6'-GalN), 4.56 (d,  $J_{1,2} = 7.7$  Hz, 3H, 3  $\times$  H1-GalN), 4.66, 4.75 (A, B of AB,  $J = 11.7$  Hz, 6H, 3 $\times$ CH<sub>2</sub>, Troc), 5.08 (s, 2H, CH<sub>2</sub>Ph), 5.14 (m, 3H, 3  $\times$  H3-GalN), 5.35 (m, 3H, 3  $\times$  H4-GalN), 5.78 (d,  $J = 7.6$  Hz, 3H, N-H, GalN), 5.95 (bs, 1H, NH, Tris), 7.32 (m, 5H, C<sub>6</sub>H<sub>5</sub>);  $^{13}C$  NMR (125 MHz,  $CDCl_3$ ):  $\delta = 20.6$ , 20.7, 20.8 (9C, 9  $\times$  CH<sub>3</sub>, AcO), 25.8 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 29.6, 29.7, 29.8 [9C, 3  $\times$  (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OGalN], 34.8 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 40.7 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 59.8 [C(CH<sub>2</sub>O)<sub>3</sub>], 61.4 (3C, 3 $\times$  C6-GalN), 66.6 (CH<sub>2</sub>Ph), 66.9 (3C, 3  $\times$  C4-GalN), 67.4, 67.5, 67.6 [15C, 3  $\times$  (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGalN], 68.4 (3C, 3  $\times$  OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGalN), 69.1 [3C, C(CH<sub>2</sub>O)<sub>3</sub>], 70.1 (3C, 3  $\times$  C3-GalN), 70.5 (3C, 3

× C5-GalN), 95.6 (3C, 3 × CCl<sub>3</sub>, Troc), 101.6 (3C, 3 × C1-GalN), 128.0-128.5 (6C, C<sub>6</sub>H<sub>5</sub>), 154.3, 154.5 [4C, 4 × N-(C=O)O], 170.3-171.2 (10C, 10 × C=O); Anal. Calcd. for C<sub>88</sub>H<sub>132</sub>Cl<sub>9</sub>N<sub>5</sub>O<sub>42</sub>: C, 46.95; H, 5.91; N, 3.11. Found: C, 47.22; H, 6.00; N, 3.03.

***N*-{Tris[13-(2-acetamido-3,4,6-tri-*O*-acetyl-β-D-galactopyranosyloxy)-2,6,10-trioxatridecyl]methyl}-(4-benzyloxycarbonylamino)-butyramide (25):** Compound **24** (20 mg, 8.88 μmol) was dissolved in dry dioxane (1 ml), and activated Zn dust (55 mg, 84.1 mmol) was added, followed by acetic anhydride (272 μl, 2.66 mmol), and the reaction mixture was stirred at rt under argon for 16 h. The mixture was filtered and the solvents were removed in vacuo. The residue was purified by silica gel chromatography (DCM/MeOH stepwise gradient 99:1→93:7) to afford **25** (13.5 mg, 82 %, R<sub>f</sub> 0.13 E) as a colorless solid. [α]<sub>D</sub> = -12.2 (c 0.5, CHCl<sub>3</sub>); <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ = 1.76-1.85 [m, 20H, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OGalN, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O], 1.95 (s, 9H, 3 × NAc), 1.99, 2.04, 2.14 (s, 27H, 9 × OAc), 2.19 (m, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.22 (m, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.44-3.48 [m, 30H, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGalN], 3.58 (m, 3H, 3 × CH<sub>2</sub>OGalN-H<sub>a</sub>), 3.67 [s, 6H, C(CH<sub>2</sub>O)<sub>3</sub>], 3.89-3.96 (m, 6H, 3 × H5-GalN, 3 × CH<sub>2</sub>OGalN-H<sub>b</sub>), 4.00 (m, 3H, 3 × H2-GalN), 4.09-4.18 (m, 6H, 3 × H6-GalN, 3 × H6'-GalN), 4.64 (d, J<sub>1,2</sub> = 8.3 Hz, 3H, 3 × H1-GalN), 5.08 (s, 2H, CH<sub>2</sub>Ph), 5.23 (m, 3H, 3 × H3-GalN), 5.34 (s, 3H, 3 × H4-GalN), 5.41 (m, 1H, NH, Cbz), 6.05 (d, J = 8.6 Hz, 3H, NHAc), 7.30-7.34 (m, 5H, C<sub>6</sub>H<sub>5</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ = 20.6 (9C, 9 × CH<sub>3</sub>, AcO), 23.3 (3C, 3 × CH<sub>3</sub>, AcHN), 25.9 (1C, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 29.9, 29.8, 29.6 [9C, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OGalN], 34.5 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 40.3 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 51.3 (3C, 3 × C2-GalN), 59.8 [C(CH<sub>2</sub>O)<sub>3</sub>], 61.4 (3C, 3 × C6-GalN), 66.5 (CH<sub>2</sub>Ph), 66.7 (3C, 3 × C4-GalN), 67.0 (3C, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGalN), 67.7 [15C, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGalN], 69.0 [3C, C(CH<sub>2</sub>O)<sub>3</sub>], 70.0 (3C, 3 × C3-GalN), 70.4 (3C, 3 × C5-GalN), 101.2 (3C, 3 × C1-GalN), 127.9, 128.0, 128.5, 136.6 (6C, C<sub>6</sub>H<sub>5</sub>), 156.7 [1C, N-(C=O)O], 170.2-172.5, (13C, 13 × C=O);



Anal. Calcd. for C<sub>85</sub>H<sub>135</sub>N<sub>5</sub>O<sub>39</sub>: C, 55.15; H, 7.35; N, 3.78. Found: C, 54.84; H, 7.34; N, 3.61.

***N*-{Tris[13-(2-acetamido-β-D-galactopyranosyloxy)-2,6,10-**

**trioxatridecyl]methyl}-4-(benzyloxycarbonylamino)-butyramide (5b):**

Compound **25** (35 mg, 18.9 μmol) was dissolved in dry dioxane (1 ml), and a solution of sodium methoxide in methanol (0.1 M, 1 ml) was added. The mixture was stirred at rt under argon for 4 h, after which it was neutralized with Dowex 50X8 (H<sup>+</sup>-form), filtered, and the solvent was removed under reduced pressure. The residue was purified on an RP-18 column (H<sub>2</sub>O/MeOH, stepwise gradient, 1:0→2:3) to afford **5b** (20.1 mg, 72%, R<sub>f</sub> 0.1 F) as a colorless solid. [α]<sub>D</sub> = -1.7 (c 1, MeOH); <sup>1</sup>H-NMR (500 MHz, MeOD): δ = 1.75-1.82 [m, 20H, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OGalN, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O], 1.99 (s, 9H, 3 × NAc), 2.20 (t, *J* = 7.3 Hz 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.16 (t, *J* = 6.8 Hz, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.47-3.49 [m, 33H, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGalN, 3 × H5-GalN], 3.54-3.61 (m, 6H, 3 × CH<sub>2</sub>OGalN-H<sub>a</sub>, 3 × H3-GalN), 3.67 [s, 6H, C(CH<sub>2</sub>O)<sub>3</sub>], 3.72-3.79 (m, 6H, 3 × H6-GalN), 3.83 (m, 3H, 3 × H4-GalN), 3.89-3.95 (m, 6H, 3 × CH<sub>2</sub>OGalN-H<sub>b</sub>, 3 × H2-GalN), 4.36 (d, *J*<sub>1,2</sub> = 8.4 Hz, 3H, 3 × H1-GalN), 5.07 (s, 2H, CH<sub>2</sub>Ph), 7.24-7.35 (m, 5H, C<sub>6</sub>H<sub>5</sub>); <sup>13</sup>C NMR (125 MHz, MeOD): δ = 23.1 (3C, 3 × CH<sub>3</sub>, AcHN), 27.6 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 31.0, 30.1 [9C, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OGalN], 35.0 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 41.1 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 54.3 (3C, 3 × C2-GalN), 61.6 [C(CH<sub>2</sub>O)<sub>3</sub>], 62.5 (3C, 3 × C6-GalN), 67.4 (4C, 3 × OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGalN, CH<sub>2</sub>Ph), 68.9, 68.8, 68.6 [15C, 3 × (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGalN], 66.4 (3C, 3 × C4-GalN), 69.6 [3C, C(CH<sub>2</sub>O)<sub>3</sub>], 73.3 (3C, 3 × C3-GalN), 76.6 (3C, 3 × C5-GalN), 103.1 (3C, 3 × C1-GalN), 128.8, 128.9, 129.5, 138.4 (6C, C<sub>6</sub>H<sub>5</sub>), 158.9 [N-(C=O)O]; ESI-MS: Calcd. for C<sub>67</sub>H<sub>117</sub>N<sub>5</sub>O<sub>30</sub>Na (M+Na)<sup>+</sup>: 1494.77; Found *m/z* 1495.41.

***N*-{Tris[13-(2-acetamido- $\beta$ -D-galactopyranosyloxy)-2,6,10-**

**trioxatridecyl]methyl}-4-amino-butylamide (26):** Compound **5b** (15 mg, 10.2  $\mu$ mol) was dissolved in methanol/dioxane (2 ml, 1:1 v/v), and Pd/C (10% Pd, 15 mg) was added. The resultant suspension was vigorously stirred under a H<sub>2</sub> atmosphere (1 atm) at rt for 3 h. The mixture was then diluted with methanol, filtered and concentrated in vacuo to yield **26** as a colorless solid (13 mg, 95%).  $[\alpha]_D = -1.49$  (c 0.67, MeOH); <sup>1</sup>H-NMR (500 MHz, MeOD):  $\delta = 1.78$ -1.82 [m, 20H, 3  $\times$  (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>O)<sub>3</sub>GalN, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O], 1.99 (s, 9H, 3  $\times$  NAc), 2.29 (t,  $J = 7.2$  Hz, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 2.82 (m, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.48-3.51 [m, 33H, 3  $\times$  (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGalN, 3  $\times$  H5-GalN], 3.55-3.63 (m, 6H, 3  $\times$  CH<sub>2</sub>OGalN-H<sub>a</sub>, 3  $\times$  H3-GalN), 3.67 [s, 6H, C(CH<sub>2</sub>O)<sub>3</sub>], 3.70-3.79 (m, 6H, 3  $\times$  H6-GalN), 3.84 (m, 3H, 3  $\times$  H4-GalN), 3.89-3.95 (m, 6H, 3  $\times$  CH<sub>2</sub>OGalN-H<sub>b</sub>, 3  $\times$  H2-GalN), 4.35 (d,  $J_{1,2} = 8.4$  Hz, 3H, 3  $\times$  H1-GalN); <sup>13</sup>C NMR (125 MHz, MeOD):  $\delta = 23.1$  (3C, 3  $\times$  CH<sub>3</sub>, AcHN), 27.5 (1C, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 31.0, 30.1 [9C, 3  $\times$  (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>3</sub>OGalN], 34.7 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 41.2 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 54.3 (3C, 3  $\times$  C2-GalN), 61.6 [C(CH<sub>2</sub>O)<sub>3</sub>], 62.5 (3C, 3  $\times$  C6-GalN), 67.4 (3C, 3  $\times$  OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGalN), 68.6, 68.8, 68.9 [15C, 3  $\times$  (OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OGalN], 69.4 (3C, 3  $\times$  C4-GalN), 69.7 [3C, C(CH<sub>2</sub>O)<sub>3</sub>], 73.3 (3C, 3  $\times$  C3-GalN), 76.6 (3C, 3  $\times$  C5-GalN), 103.1 (3C, 3  $\times$  C1-GalN), 174.0, 175.2 (4  $\times$  C=O); ESI-MS: Calcd. for C<sub>59</sub>H<sub>112</sub>N<sub>5</sub>O<sub>28</sub> (M+H)<sup>+</sup>: 1338.75; Found m/z 1339.22.

**Fluorescent-labeled, GalNAc-terminated compound (7):** A stock solution containing compound **26** (10 mg, 7.47  $\mu$ mol), DIPEA (50  $\mu$ l, 292  $\mu$ mol) and 4 Å molecular sieves (25 mg) in dry DMF/dioxane (1 ml, 1:1) was stirred at rt under argon for 2 h. The resultant solution (500  $\mu$ l) was transferred to a small vial containing Alexa Fluor<sup>®</sup> 488-NHS (1 mg, 1.55  $\mu$ mol), 4 Å molecular sieves (25 mg), and a stirring bar. The mixture was stirred in the dark at rt under argon for 4 d, then diluted with MeOH, filtered, and the solvents were removed in vacuo. The residue was purified by gel filtration on a Sephadex LH-20 column (2.5  $\times$  35 cm)

using MeOH as eluant, then on an RP-18 column (H<sub>2</sub>O/MeOH stepwise gradient 1:0 to 1:1) to yield **7** (2.6 mg, 90%). ESI-MS: Calcd. for C<sub>80</sub>H<sub>121</sub>N<sub>7</sub>O<sub>38</sub>S<sub>2</sub><sup>2-</sup> (M/2)<sup>-</sup>: 925.86; Found m/z 926.05.

**4-(Benzyloxycarbonylamino)-N-[[tris(hydroxymethyl)]methyl]-butyramide**

**(27)**: Compound **17** [36] (3.01 g, 12.6 mmol) and EEDQ (3.26 g, 13.2 mmol) were stirred in dry pyridine (100 ml) at rt under argon for 1 h. Tris (**9**) (1.33 g, 10.9 mmol) was then added, and the resultant suspension was stirred at 90 °C for 24 h. The solvent was removed in vacuo, and the residue was triturated with EtOAc and a few drops of MeOH. The product was filtered off, washed with EtOAc, then with cold water, and dried under high vacuum to yield **27** (2.88 g, 77%) as a white solid. <sup>1</sup>H-NMR (500 MHz, MeOD): δ = 1.79 (quintet, *J* = 7.0 Hz, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 2.28 (t, *J* = 7.4 Hz, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.16 (t, *J* = 6.8 Hz, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.73 (s, 6H, C[CH<sub>2</sub>O]<sub>3</sub>), 5.09 (m, 2H, CH<sub>2</sub>Ph), 7.28-7.35 (m, 5H, C<sub>6</sub>H<sub>5</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ = 27.1 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 34.5 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 41.0 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 62.5 [3C, C(CH<sub>2</sub>O)<sub>3</sub>], 63.6 [C(CH<sub>2</sub>O)<sub>3</sub>], 67.4 (CH<sub>2</sub>Ph), 128.8, 129.0, 129.4, 130.6, 138.3 (6C, C<sub>6</sub>H<sub>5</sub>), 159.0 [N-(C=O)O], 176.6 (C=O); ESI-MS: Calcd. for C<sub>16</sub>H<sub>24</sub>N<sub>2</sub>O<sub>6</sub>Na (M+Na)<sup>+</sup>: 363.15; Found m/z 363.13.

**N-[[Tris(2,3,4,6-tetra-O-benzoyl-β-D-galactopyranosyloxy)methyl]methyl]-(4-benzyloxycarbonylamino)-butyramide (28)**: Compounds **27** (100 mg, 0.294 mmol) and ethyl 2,3,4,6-tetra-O-benzoyl-1-thio-β-D-galactopyranoside (**20**) [27] (753 mg, 1.17 mmol) were stirred with 4 Å molecular sieves (500 mg) in DCE/Et<sub>2</sub>O (20 ml, 1:1) at rt under argon for 2 h. The mixture was cooled to 0 °C, and a solution of NIS (263 mg, 1.17 mmol) and TfOH (0.117 mmol, 10.2 μl) in DCE/Et<sub>2</sub>O (10 ml, 1:1) was added. The resultant mixture was stirred at 0 °C under argon for 1 h, upon which it turned deep brown. The mixture was diluted with DCM (50 ml), filtered through Celite, washed with Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> (1 M, 25 ml), followed by NaHCO<sub>3</sub> (0.1 M, 25 ml), and dried with Na<sub>2</sub>SO<sub>4</sub>. The solvent was removed under reduced pressure, and the residue was purified by silica gel

chromatography (petrol ether/EtOAc 3:1 → 1:1) to yield **28** (311 mg, 51%,  $R_f$  0.23 B) as a colorless solid.  $[\alpha]_D = +39$  (c 1,  $\text{CHCl}_3$ );  $^1\text{H-NMR}$  (500 MHz,  $\text{CDCl}_3$ ):  $\delta = 1.65$  (m, 2H,  $\text{NCH}_2\text{CH}_2\text{CH}_2\text{C=O}$ ), 1.86 (m, 2H,  $\text{NCH}_2\text{CH}_2\text{CH}_2\text{C=O}$ ), 3.09 (m, 2H,  $\text{NCH}_2\text{CH}_2\text{CH}_2\text{C=O}$ ), 3.52 (d, A of AB,  $J = 10.2$  Hz, 3H, 3 ×  $\text{CH}_2\text{OGal-H}_a$ ), 3.72 (m, 3H, 3 × H5-Gal), 4.10 (d,  $J_{1,2} = 8.0$  Hz, 3H, 3 × H1-Gal), 4.32 (dd,  $J_{5,6'} = 7.4$ ,  $J_{6,6'} = 11.2$  Hz, 3H, 3 × H6'-Gal), 4.35 (d, B of AB,  $J = 10.1$  Hz, 3H, 3 ×  $\text{CH}_2\text{OGal-H}_b$ ), 4.51 (dd,  $J_{5,6} = 6.2$ ,  $J_{6,6'} = 11.2$  Hz, 3H, 3 × H6-Gal), 5.02 (m, 2H,  $\text{CH}_2\text{Ph}$ ), 5.32 (s, 1H, NH, Cbz), 5.38 (dd,  $J_{4,3} = 3.3$ ,  $J_{2,3} = 10.3$  Hz, 3 × H3-Gal), 5.62 (dd,  $J_{1,2} = 8.1$ ,  $J_{2,3} = 10.2$  Hz, 3 × H2-Gal), 5.75 (s, 1H, NH, Tris), 5.84 (d,  $J_{3,4} = 3.2$  Hz, 3H, 3 × H4-Gal), 7.21-8.10 (m, 65H, 13 ×  $\text{C}_6\text{H}_5$ );  $^{13}\text{C NMR}$  (125 MHz,  $\text{CDCl}_3$ ):  $\delta = 25.2$  ( $\text{NCH}_2\text{CH}_2\text{CH}_2\text{C=O}$ ), 33.8 ( $\text{NCH}_2\text{CH}_2\text{CH}_2\text{C=O}$ ), 40.1 ( $\text{NCH}_2\text{CH}_2\text{CH}_2\text{C=O}$ ), 59.4 [ $\text{C}(\text{CH}_2\text{O})_3$ ], 61.4 (3C, 3 × C6-Gal), 66.4 ( $\text{CH}_2\text{Ph}$ ), 67.8, 67.9 [6C,  $\text{C}(\text{CH}_2\text{O})_3$ , 3 × C4-Gal], 69.9 (3C, 3 × C2-Gal), 71.0, 71.1 (6C, 3 × C3-Gal, 3 × C5-Gal), 101.8 (3C, 3 × C1-Gal), 127.9-136.7 (78C, 13 ×  $\text{C}_6\text{H}_5$ ), 156.5 [ $\text{N}(\text{C=O})\text{O}$ ], 164.9-172.6 (13C, 13 ×  $\text{C=O}$ ); Anal. Calcd. for  $\text{C}_{118}\text{H}_{102}\text{N}_2\text{O}_{33}$ : C, 68.27; H, 4.95; N, 1.35. Found: C, 67.81; H, 5.08; N, 1.36.

**4-(Benzyloxycarbonylamino)-N-[[tris( $\beta$ -D-galactopyranosyloxy)methyl]-methyl]-butyramide (29):** Compound **28** (241 mg, 0.115 mmol) was dissolved in dry MeOH/dioxane (35 ml, 2.5:1), and sodium metal (65 mg, 2.83 mmol) was added. The resultant mixture was stirred at rt under argon for 6 h, after which it was neutralized with Dowex 50X8 ( $\text{H}^+$ -form), filtered, and the solvent was removed under reduced pressure. The residue was purified on a Sephadex G-25 column, and, after the removal of water in vacuo, washed with  $\text{Et}_2\text{O}:\text{DCM}$  (ca. 2:1) to afford the desired product **29** (81.5 mg, 85 %,  $R_f$  0.11 F) as a white powder.  $[\alpha]_D = -0.2$  (c 1, MeOH);  $^1\text{H-NMR}$  (500 MHz, MeOD):  $\delta = 1.78$  (m,  $\text{NCH}_2\text{CH}_2\text{CH}_2\text{C=O}$ ), 2.22 (t,  $J = 7.1$  Hz, 2H,  $\text{NCH}_2\text{CH}_2\text{CH}_2\text{C=O}$ ), 3.16 (t,  $J = 6.9$  Hz, 2H,  $\text{NCH}_2\text{CH}_2\text{CH}_2\text{C=O}$ ), 3.45-3.54 (m, 9H, 3 × H2-Gal, 3 × H3-Gal, 3 × H5-Gal), 3.69 (dd,  $J_{5,6'} = 5.2$ ,  $J_{6,6'} = 11.3$  Hz, 3H, 3 × H6'-Gal), 3.76 (dd,  $J_{5,6} = 6.9$ ,  $J_{6,6'} = 11.4$  Hz, 3H, 3 × H6-Gal), 3.81 (m, 3H, 3 × H4-Gal), 3.93 (d, A of AB,  $J =$

10.2 Hz, 3H, 3 × CH<sub>2</sub>OGal-H<sub>a</sub>), 4.27 (d,  $J_{1,2} = 7.6$  Hz, 3H, 3 × H1-Gal), 4.32 (d, B of AB,  $J = 10.2$  Hz, 3H, 3 × CH<sub>2</sub>OGal-H<sub>b</sub>), 5.07 (s, 2H, CH<sub>2</sub>Ph), 7.29-7.35 (m, 5H, C<sub>6</sub>H<sub>5</sub>); <sup>13</sup>C NMR (125 MHz, MeOD):  $\delta = 27.1$  (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O); 35.0 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 41.2 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O); 62.5 (3C, 3 × C6-Gal), 67.4 (CH<sub>2</sub>Ph); 69.3 [3C, C(CH<sub>2</sub>O)<sub>3</sub>], 70.4 (3C, 3 × C4-Gal), 72.6 (3C, 3 × C2-Gal), 74.9 (3C, 3 × C3-Gal), 76.7 (3C, 3 × C5-Gal), 105.5 (3C, 3 × C1-Gal), 127.9-129.5 (6C, C<sub>6</sub>H<sub>5</sub>), 159.0 [N-(C=O)O], 175.9 (C=O); ESI-MS: Calcd. for C<sub>34</sub>H<sub>54</sub>N<sub>2</sub>O<sub>21</sub>Na (M+Na)<sup>+</sup>: 849.31; Found m/z 849.44.

**4-Amino-N-[[tris( $\beta$ -D-galactopyranosyloxy)methyl]methyl]-butyramide (30):**

Compound **29** (27.2 mg, 32.9  $\mu$ mol) was dissolved in methanol (2.5 ml), and Pd/C (10% Pd, 25 mg) was added. The resultant mixture was vigorously stirred under a H<sub>2</sub> atmosphere (1 atm) at rt for 48 h. The mixture was then filtered and concentrated in vacuo to yield **30** as a colorless solid (19.9 mg, 87%). [ $\alpha$ ]<sub>D</sub> = +3.1 (c 1, MeOH); <sup>1</sup>H-NMR (500 MHz, MeOD):  $\delta = 1.89$  (m, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 2.32 (m, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.00 (m, 2H, NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 3.43-3.50 (m, 9H, 3 × H2-Gal, 3 × H3-Gal, 3 × H5-Gal), 3.68 (dd,  $J_{5,6'} = 5.1$ ,  $J_{6,6'} = 11.2$  Hz, 3H, 3 × H6'-Gal), 3.72 (dd,  $J_{5,6} = 6.9$ ,  $J_{6,6'} = 11.3$  Hz, 3H, 3 × H6-Gal), 3.77 (m, 3H, 3 × H4-Gal), 3.87 (d, A of AB,  $J = 10.1$  Hz, 3H, 3 × CH<sub>2</sub>OGal-H<sub>a</sub>), 4.25 (d,  $J_{1,2} = 6.8$  Hz, 3H, 3 × H1-Gal), 4.31 (d, B of AB,  $J = 10.1$  Hz, 3H, 3 × CH<sub>2</sub>OGal-H<sub>b</sub>); <sup>13</sup>C NMR (125 MHz, MeOD):  $\delta = 24.5$  (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 34.7 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 40.3 (NCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>C=O), 61.3 [C(CH<sub>2</sub>O)<sub>3</sub>], 62.6 (3C, 3 × C6-Gal), 69.1 [3C, C(CH<sub>2</sub>O)<sub>3</sub>], 70.8 (3C, 3 × C4-Gal), 73.0 (3C, 3 × C2-Gal), 75.3 (3C, 3 × C3-Gal), 77.1 (3C, 3 × C5-Gal), 105.8 (3C, 3 × C1-Gal), 174.9 (C=O); ESI-MS: Calcd. for C<sub>26</sub>H<sub>49</sub>N<sub>2</sub>O<sub>19</sub> (M+H)<sup>+</sup>: 693.29; Found m/z 693.39.

**Fluorescent-labeled, Gal-terminated control compound (8):** A stock solution containing compound **30** (12 mg, 17.3  $\mu$ mol), DIPEA (50  $\mu$ l, 291  $\mu$ mol) and 4 Å molecular sieves (50 mg) in dry DMF (1 ml) was stirred at rt under argon for 2 h. The resultant mixture (500  $\mu$ l) was transferred to a small vial containing Alexa

Fluor<sup>®</sup> 488-NHS (1 mg, 1.55  $\mu$ mol), 4Å molecular sieves (25 mg), and a stirring bar. The mixture was then stirred in the dark at rt under argon for 4 d. The mixture was then diluted with MeOH, filtered, and the solvents were removed in vacuo. The residue was purified on an RP-18 column (H<sub>2</sub>O/MeOH gradient 1:0 to 19:1) to yield **8** (1.8 mg, 96%) as a red solid after final lyophilization from water. ESI-MS: Calcd. for C<sub>47</sub>H<sub>59</sub>N<sub>4</sub>Na<sub>2</sub>O<sub>29</sub>S<sub>2</sub> (M+2Na+H)<sup>+</sup>: 1253.25; Found m/z 1253.59.

## 2.6.5 Ligand binding and internalization

### 2.6.5.1 Fluorescence microscopy:

One day before the experiments, the cells were seeded at a density of  $2 \times 10^5$  cells/well into 12-well plates containing-collagen coated glass cover slips. The cells were washed once with PBS, and then serum-starved for 30 min on ice in 1 ml of DMEM containing 25 mM HEPES. They were then incubated with 500  $\mu$ l/well of the Alexa Fluor<sup>®</sup> 488-labeled compounds **6-8** (100  $\mu$ M) in the same medium on ice for 1.5 h in the dark. After the binding step, the cells were washed carefully 4 times with cold PBS. Then fresh, prewarmed, complete DMEM medium (1 ml/well) was administered and the cells were incubated for 40 min in an incubator at 37 °C in a humidified CO<sub>2</sub> atmosphere (5 %, v/v), leading to the internalization of the receptor-bound compounds into the cells. After the internalization step, the cells were washed twice with PBS and then fixed with 3% paraformaldehyde (PFA) in PBS for 30 min at 4 °C. After fixation, the cover-slips were washed abundantly with PBS and mounted upside down, in a Mowiol 4-88 mounting buffer containing *N*-propyl gallate, onto glass slides.

Selective cellular uptake of the Alexa Fluor<sup>®</sup> 488-labeled compounds was visualized using a Zeiss Axiovert 135 microscope with a 63 x planapo objective

(numerical aperture = 1.4, oil) with the appropriate filter set (450/490, FT 510, LP 520) equipped with a Zeiss AxioCam MRm CCD camera run by AxioVision 3.1 imaging software.

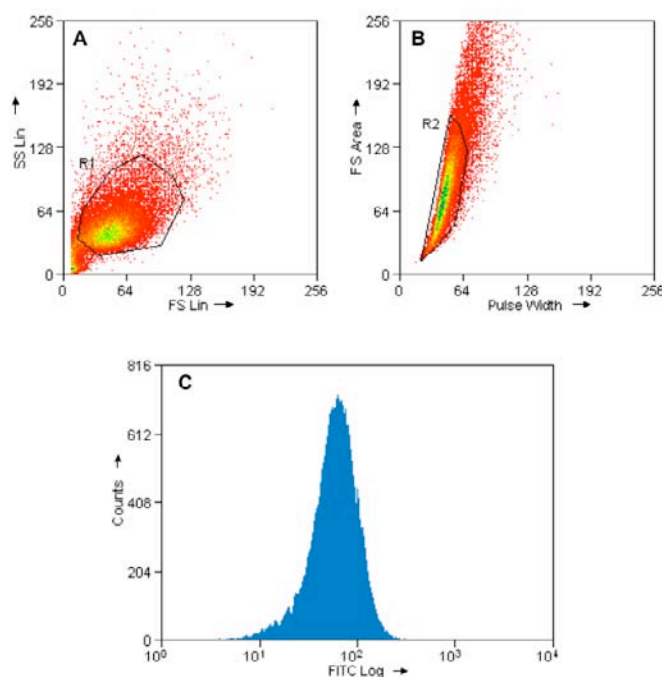
#### **2.6.5.2 Flow cytometry:**

Cells were grown for 24 h in collagen coated (80  $\mu\text{g/ml}$ ) 24-well plates at a density of  $3 \times 10^5$  cells/well or in 96-well plates at  $1.5 \times 10^5$  cells/well.

The titration experiments were performed in 24-well plates, the cell-layers were first washed twice with cold PBS before incubation with compound **7** at concentrations ranging from 0.4 to 12.5  $\mu\text{M}$  (1:2 serial dilutions) in 200  $\mu\text{l}$  of DMEM without FBS for 40 min at 37 °C. Then the cells were washed twice with cold PBS, detached and stripped from surface-bound compound by incubating them in a mixture containing 0.025% trypsin and 5 mM EDTA in PBS for 10 min on ice. Addition of complete medium quenched this process. The detached cells were collected and centrifuged at a speed of 1500 rpm for 3 min. Finally, the cells were fixed in 2% PFA in PBS for 15 min on ice followed by an aldehyde-quenching step with 100 mM lysine in PBS for 10 min. The fixed cells were then washed once with FACS buffer (PBS containing 0.5% BSA and 0.1%  $\text{NaN}_3$ ) and resuspended in 200  $\mu\text{l}$  of the same buffer for measuring.

The competitive uptake experiments were performed with cells grown in 96-well plates. 20  $\mu\text{l}$  of asialofetuin dilutions at concentrations ranging from 0.6 to 200  $\mu\text{M}$  or GalNAc at 0.6 to 200 mM were added directly to the cells, immediately followed by the addition of 20  $\mu\text{l}$  of compound **7** or **8** diluted to 20  $\mu\text{M}$  in DMEM without FBS and then incubated for 60 min at 37 °C in the incubator. The cells were washed, detached and fixed as described above.

Analyses were performed on a CyAN ADP flow cytometer with Summit 4.1 software (Dako Cytomation). An example for the analysis of HepG2 cells is given in Figure 7. The forward and side scatter gate R1 was set to count 30'000 intact cells of each sample (Figure 7, dot plot A). In gate R2, the cells counted in gate R1 were discriminated for doublets (Figure 7, dot plot B). Histogram C depicts an overlay of the log of fluorescence intensity at 488 nm of untreated (grey) and treated HepG2 cells (green) from gate R2. Uptake of a compound into cells was evaluated by comparing the shift in median intensity of fluorescence (MFI) between untreated cells (background fluorescence) and treated cells.



**Figure 7.** Example of flow cytometry analysis showing uptake of compound **7** into HepG2 cells. Dot plot (A) represents the HepG2 cell population gated for analysis (R1=30'000), plotted as a function of forward scatter (FS) and side scatter (SS). Dot blot (B) represents the population of R1 gated for single cells (R2) as a function of pulse-width and FS-area and histogram (C) depicts the log fluorescence intensity at 488 nm of the cells gated in R2. SK-Hep1 cells were analyzed in the same way, with a more compact population of cells and less debris in gate R1 and therefore smaller peaks due to less clumping of the cells (data not shown). Further analysis and IC<sub>50</sub> calculations were done with GraphPad Prism 4 software.



## 2.7 References

- [1] G. Ashwell, J. Harford, *Ann. Rev. Biochem.* **1982**, *51*, 531-554.
- [2] M. Spiess, *Biochemistry* **1990**, *29*, 10009-10018.
- [3] I. Geffen, M. Spiess, *Int. Rev. Cytol.* **1992**, *137B*, 181-219.
- [4] M. Li, H. Kurata, N. Itoh, I. Yamashina, T. Kawasaki, *J. Biol. Chem.* **1990**, *265*, 11295-11298.
- [5] K. Nakamura, H. Funakoshi, K. Miyamoto, F. Tokunaga, T. Nakamura, *Biochem. Biophys. Res. Commun.* **2001**, *280*, 1028-1035.
- [6] G.W. Hoyle, R.L. Hill, *J. Biol. Chem.* **1988**, *263*, 7487-7492.
- [7] J. Wu, M. H. Nantz, M. A. Zern, *Frontiers in Bioscience* **2002**, *7*, d717-725.
- [8] P. J. Coombs, M. E. Taylor, K. Drickamer, *Glycobiology* **2006**, *16*, 1C-7C.
- [9] Y. C. Lee, R. R. Townsend, M. R. Hardy, J. Lönngren, J. Arnarp, M. Haraldsson, H. Lönn, *J. Biol. Chem.* **1983**, *258*, 199-202.
- [10] M. D. Bider, J. M. Wahlberg, R. A. Kammerer, M. Spiess, *J. Biol. Chem.* **1996**, *271*, 31996-32001.
- [11] M. Meier, M.D. Bider, V.N. Malashkevich, M. Spiess, P. Burkhard, *J. Mol. Biol.* **2000**, *300*, 857-865.
- [12] J. U. Baenziger, Y. Maynard, *J. Biol. Chem.* **1980**, *255*, 4607-4613.
- [13] J. U. Baenziger, D. Fiete, *Cell* **1980**, *22*, 611-620.
- [14] D. T. Connolly, R. R. Townsend, K. Kawaguchi, W. R. Bell, Y. C. Lee, *J. Biol. Chem.* **1982**, *257*, 939-945.
- [15] M. Mammen, S. -K. Choi, G.M. Whitesides, *Angew. Chem.* **1998**, *110*, 2908-2953.
- [16] Y. C. Lee, R. T. Lee, in *Carbohydrates in Chemistry and Biology*, B. Ernst, G. W. Hart, P. Sinaÿ, Eds., Wiley-WCH, Weinheim, **2000**, Vol. 4, 549-561; and references cited therein.
- [17] K. Bock, J. Arnarp, J. Lönngren, *Eur. J. Biochem.* **1982**, *129*, 171-178.
- [18] M. Nishikawa, *Biol. Pharm. Bull.* **2005**, *28*, 195-200.
- [19] J.M. Wilson, M. Grossman, C.H. Wu, N.R. Chowdhury, C.Y. Wu, J.R. Chowdhury, *J. Biol. Chem.* **1992**, *267*, 963-967.

- [20] M. Grossman, S.E. Raper, K. Kozarsky, E.A. Stein, J.F. Engelhardt, D. Muller, P.J. Lupien, J.M. Wilson, *Nature Genet.* **1994**, 6, 335-341.
- [21] G.Y. Wu, C.H. Wu, *J. Biol. Chem.* **1991**, 263, 14621-14624.
- [22] T. Hara, Y. Aramaki, S. Takada, K. Koike, S. Tsuchiya, *Gene Ther.* **1995**, 2, 784-788.
- [23] T. Hara, Y. Aramaki, S. Takada, K. Koike, S. Tsuchiya, *Gene* **1995**, 159, 167-174.
- [24] T. Hara, H. Kawasaki, Y. Aramaki, S. Takada, K. Koike, K. Ishidate, H. Kato, S. Tsuchiya, *Biochem. Biophys. Acta* **1996**, 1278, 51-58.
- [25] S. Kawakami, F. Yamashita, M. Nishikawa, Y. Takakura, M. Hashida, *Biochem. Biophys. Res. Commun.* **1998**, 252, 78-83.
- [26] H. J. M. Kempen, C. Hoes, J. H. van Boom, H. H. Spanjer, J. de Lange, A. Langendoen, T. J. C. van Berkel, *J. Med. Chem.* **1984**, 17, 1306-1312.
- [27] E. A. L. Biessen, D. M. Beuting, H. C. P. F. Roelen, G. A. van de Marel, J. H. van Boom, T. J. C. van Berkel, *J. Med. Chem.* **1995**, 38, 1538-1546.
- [28] E. A. L. Biessen, H. Broxtermann, J. H. van Boom, T. J. C. van Berkel, *J. Med. Chem.* **1995**, 38, 1846-1852.
- [29] L. A. J. M. Sliedregt, P. C. N. Rensen, E. T. Rump, P. J. van Santbrink, M. K. Bijsterbosch, A. R. P. M. Valentijn, G. A. van de Marel, J. H. van Boom, T. J. C. van Berkel, E. A. L. Biessen, *J. Med. Chem.* **1999**, 42, 609-618.
- [30] P. C. N. Rensen, S. H. van Leeuwen, L. A. J. M. Sliedregt, T. J. C. van Berkel, E. A. L. Biessen, *J. Med. Chem.* **2004**, 47, 5798-5808.
- [31] R. T. Lee, Y. C. Lee, *Bioconjugate Chem.* **1997**, 8, 762-765.
- [32] P. C. N. Rensen, L. A. J. M. Sliedregt, M. Ferns, E. Kieviet, S. M. W. van Rossenberg, S. H. van Leeuwen, T. J. C. van Berkel, E. A. L. Biessen, *J. Biol. Chem.* **2001**, 276, 37577-37584.
- [33] R. P. Haugland, *Handbook of Fluorescent Probes and Research Products*, Molecular Probes, Inc., **2002**, 9<sup>th</sup> Ed.
- [34] M. Segura, F. Sansone, A. Casnati, R. Ungaro, *Synthesis* **2001**, 14, 2105-2112.

- [35] E. V. Dehmlow, S. S. Dehmlow, *Phase-transfer catalysis*, Verlag-Chemie, Weinheim/Bergstr, West Germany, **1993**, 2<sup>nd</sup> Ed.
- [36] A. P. Fosker, H. D. Law, *J. Chem. Soc.* **1965**, 5, 7305-7307.
- [37] U. Ellervik, G. Magnusson, *Carbohydr. Res.* **1996**, 280, 251-260.
- [38] I. Villanueva, B. Hernandez, V. Chang, M. D. Heagy, *Synthesis* **2000**, 10, 1435-1438.
- [39] A. L. Schwartz, S. E. Fridovich, B. B. Knowles, H. F. Lodish, *J. Biol. Chem.* **1981**, 256, 8878-8881.
- [40] S. C. Heffelfinger, H. H. Hawkins, J. Barrish, L. Taylor, G. J. Darlington, *In Vitro Cell Dev. Biol.* **1992**, 28A, 136-142.
- [41] D. A. Lauffenburger, J. J. Lindermann, *Receptors: Models for binding, trafficking, and signalling*, Oxford University Press, New York, **1993**, 1<sup>st</sup> Ed.
- [42] H. Yura, M. Ishihara, Y. Kanatani, B. Takase, H. Hattori, S. Suzuki, M. Kawakami, T. Matsui, *J. Biochem.* **2006**, 139, 637.
- [43] Y. T. Wu, W. T. Jiaang, K. G. Lin, C. M. Huang, C. H. Chang, Y. L. Sun, K. H. Fan, W. C. Hsu, H. E. Wang, S. B. Lin, S. T. Chen, *Curr. Drug. Deliv.* **2004**, 1, 119.
- [44] *Vogel's "Textbook of Practical Organic Chemistry"*, Prentice Hall, **1989**, 5<sup>th</sup> Ed., p. 467.

# Chapter 3: Directed library of small molecular weight ligands for the ASGP-R

## 3.1 Introduction

In the past, most research efforts in designing high affinity ligands for the ASGP-R were focused on multivalent compounds, with high molecular weights, for use as homing devices in liver-specific drug and gene delivery [4]. However, for a number of years it has also been postulated that the ASGP-R could be involved in hepatitis B [46] and C [47] entry into the liver cells. Therefore, antiviral therapy involving small molecular weight drugs that bind to the ASGP-R is a promising field of research.

### 3.1.1 Hepatitis B

In 2003, more than 350 million people worldwide suffered from hepatitis B virus (HBV) infection [48], with chronic infections resulting in liver cirrhosis and/or hepatocellular carcinoma [49].

Chronic hepatitis B treatment involves antiviral drugs such as lamivudine and adefovir and immune system modulators such as pegylated interferon alpha (IFN- $\alpha$ ) [50].

HBV particles possess three related envelope glycoproteins: S, preS1 and preS2 [51]. Experiments by Treichel *et al.* [46] on ASGP-R-containing human hepatocellular carcinoma HepG2 cells [38] and on the purified receptor have demonstrated that the ASGP-R is able to specifically bind the HBV via the preS1 and preS2 envelope proteins. The study also suggested that this attachment

could lead to viral endocytosis by the hepatocytes, which could be a mechanism for viral entry into its host. Moreover, the binding of HBV to the ASGP-R was inhibited by natural ligands (asialofetuin and ASOR), EDTA (which is known to chelate and sequester  $\text{Ca}^{2+}$ , necessary for the structural integrity of the binding site), anti-ASGP-R antibodies and monoclonal anti-preS1 antibodies. Furthermore, the binding and inhibition data indicated that the interaction occurred via specific, D-galactose-containing glycosides on the HBV preS1 glycoprotein.

In contrast, earlier studies showed that despite being highly glycosylated, the preS2 envelope protein contains no D-galactose residues on its N-linked glycans [52], therefore the role of preS2 in HBV attachment is yet to be clarified.

Further studies on HBV attachment to, and entry into hepatocytes suggested that the endocytosis of HBV particles by hepatocytes is a prerequisite for HBV liver infection [53].

### **3.1.2 Hepatitis C**

According to World Health Organisation estimates [54], around 170 million people around the world are infected with the hepatitis C virus (HCV). More than 70% of those infected fail to clear the virus and have a persistent infection which may also lead to liver cirrhosis and/or hepatocellular carcinoma [55,56].

The exact details regarding the early stages of viral infection remain largely unknown, however it is generally accepted that the HCV envelope proteins (E1 and E2) play a crucial role in binding to the host cell [47].

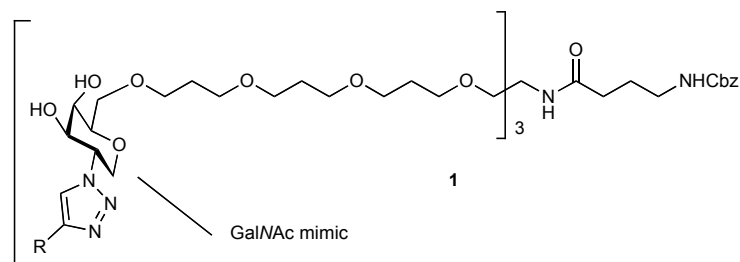
Since hepatocytes are the primary site of HCV infection, it has been suggested that the virus may also utilise the ASGP-R for binding and entry [47]. The

envelope proteins on HCV are glycosylated, and it has been shown that they can bind to, and get internalised by human hepatocellular carcinoma (HepG2) cells [57]. The fact that the HCV envelope proteins could bind to the ASGP-R via a carbohydrate-protein interaction involving the ASGP-R sugar binding site was supported by the following: (1) the binding was inhibited by EDTA, and (2) the binding was also inhibited by asialoorosomucoid, a known natural ligand of the ASGP-R [57, 58].

Treatment of hepatitis C also involves antiviral drugs such as ribavirin, as well as IFN- $\alpha$ . However, even a combination of these leads to a viral clearance in only about 50% of the patients [59].

The facts described above suggest that blocking of HBC/HCV attachment to hepatocytes *in vivo* could be a promising new route to anti-hepatitis therapy. Thus, there is an obvious need for the development of a small molecular weight Gal/GalNAc mimic that would bind with a higher affinity to the ASGP-R than the carbohydrates on the viral envelope proteins. Such a mimic could then be used for anti-hepatitis B/C therapy by inhibiting viral attachment to the hepatocytes via the ASGP-R.

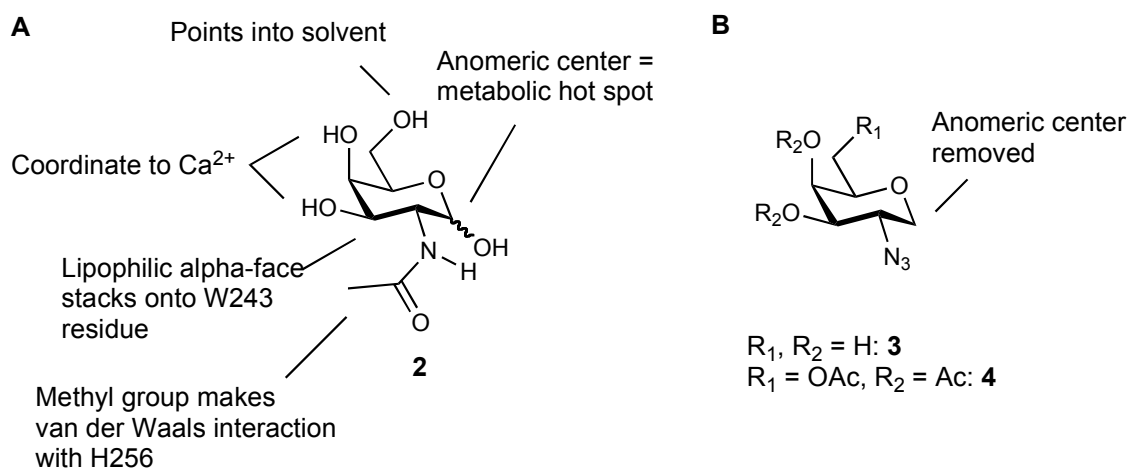
Furthermore, by replacing the Gal/GalNAc residues on multivalent ligands (**1**, Figure 5) with a higher affinity mimic, a compound with an even higher affinity would be generated improving existing multivalent ligands intended for liver-specific drug delivery [60].



**Figure 5.** A hypothetical trivalent ligand (**1**) for the ASGP-R featuring a GalNAc mimic attached via the 6-position for use in liver-specific drug delivery.

### 3.1.3 Design of a small molecular weight ASGP-R ligand

Relatively little research has been done on increasing the affinity of monovalent ligands towards the ASGP-R. Up to now, the highest affinity small molecular weight, monovalent ligand for the ASGP-R has been *N*-acetylgalactosamine (GalNAc, **2**, Figure 6A). The chemical structure and physicochemical properties of GalNAc make it a poor candidate for therapeutic use. Its high polarity, due to the many hydrogen bond donors and acceptors, causes it to violate Lipinski's rules for good absorbance, and makes it prone to fast elimination. It has a highly metabolically labile anomeric center susceptible to oxidation, reduction and metabolism leading to a low plasma half-life.

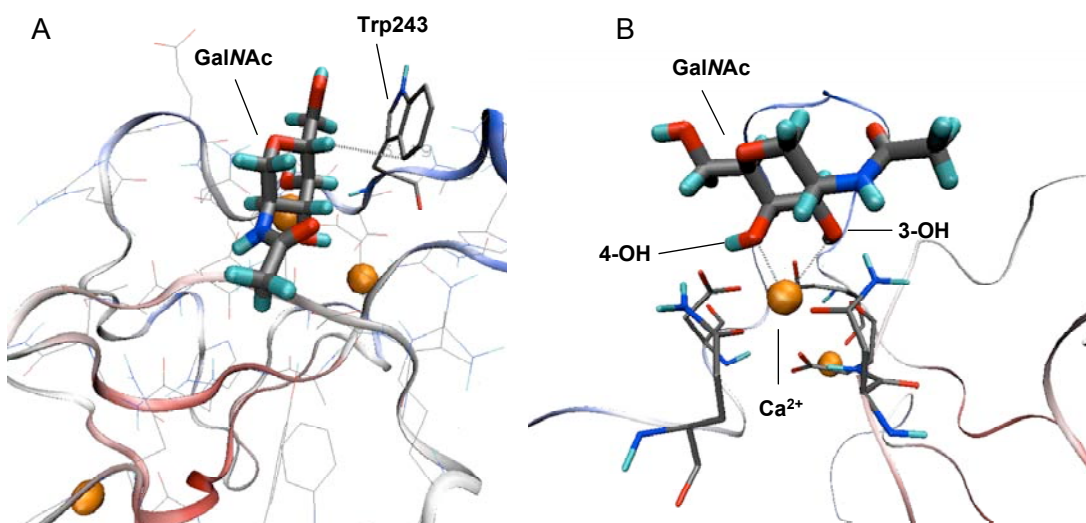


**Figure 6.** (A) Lead compound *N*-acetylgalactosamine (GalNAc, **2**), showing important interactions with the binding site of the ASGP-R H1-CRD; (B) Compound **3** is dehydroxylated at the 6-position. Compound **4** is the protected scaffold used for directed library.

Therefore, the focus of this project was to design a GalNAc mimic, which would be more lipophilic, metabolically stable, synthetically easily accessible and have a higher affinity towards the ASGP-R than GalNAc.

### 3.1.4 Binding mode of GalNAc to the ASGP-R

Taking GalNAc as the lead compound, it was first necessary to examine its interaction with the binding site of the ASGP-R H1-CRD. The model (Figure 7), based on the crystal structure of H1-CRD [7], clearly illustrates the important interactions involved in the binding of GalNAc to the ASGP-R. These interactions are further summarized in Figure 6A. Thus, the 3-OH and the 4-OH of GalNAc must be equatorial and axial [61], respectively, and cannot be modified because they coordinate to the  $\text{Ca}^{2+}$  ion. The hydrophobic patch on the  $\alpha$ -face of GalNAc, formed by the circular arrangement of ring C-H bonds, is involved in a lipophilic interaction by stacking onto the indole side chain of Trp243.



**Figure 7.** A model of GalNAc docked into the sugar binding site of the H1 CRD. (A) Illustration of the hydrophobic interaction between Trp243 and the  $\alpha$ -face of GalNAc. (B) Illustration of the 3- and 4-OH groups coordinating to the calcium ion. (Picture courtesy of M. Spreafico, manual docking trial, MacYeti 7.05)



The anomeric centre is strictly speaking not necessary for binding; and since it acts as a metabolic hotspot, it can be totally removed making the compound more metabolically stable. The 6-OH is not required for binding because it points into the solvent, and has enough space around it to accommodate a wide range of substituents [61]. Therefore, it could always be removed altogether (compound **3**, Figure 6B) or replaced by a more lipophilic substituent in order to improve the Lipinski parameters. Alternatively, the 6-OH could also serve as an attachment point for further elaboration or conjugation to oligovalent carriers, fluorescent labels, etc. (cf. Figure 5). The acetamido substituent in the 2-position must be equatorial, but could in principle be replaced by other substituents.

The theoretical binding mode of Gal/GalNAc to the ASGP-R has been further reinforced by several groups, which performed binding studies using galactose derivatives [62,63]. Studies using the closely related rat hepatic lectin subunit 1 (RHL-1) [62], showed that acylating the amino group of galactosamine with carboxylic acids featuring alkyl chains longer than 2 carbons led to a decrease in affinity. For example, by simply replacing the *N*-acetyl group in GalNAc with a bulkier *N*-benzoyl group, several independent research groups proved that the affinity towards the related rat ASGP-R dropped by up to 9-fold [63-65].

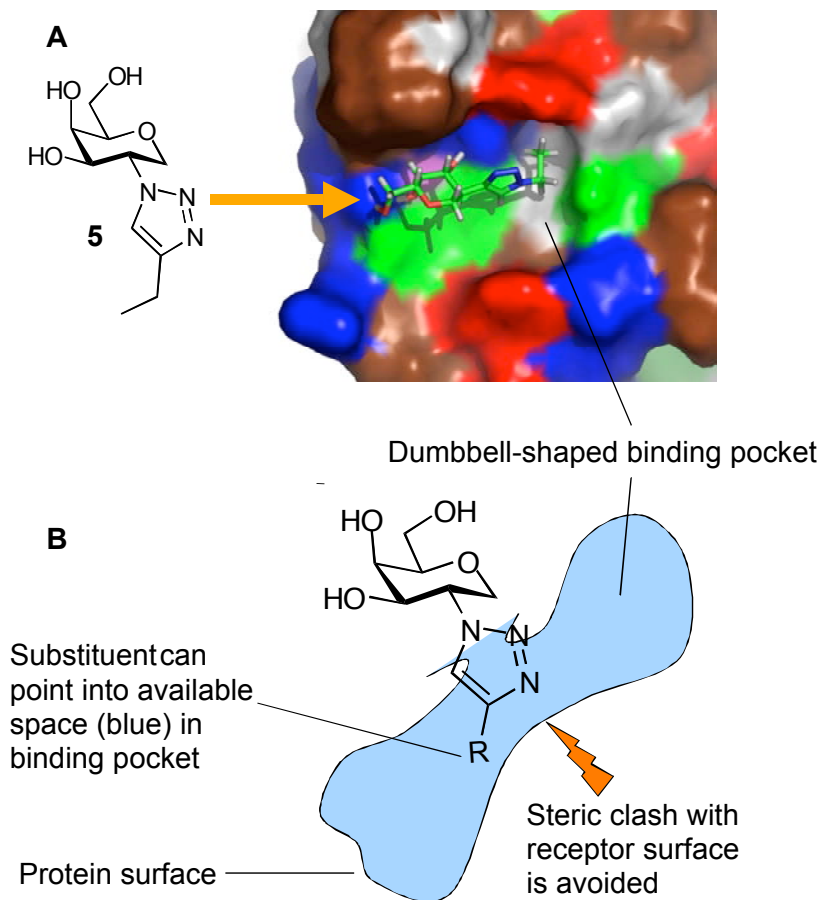
Nevertheless, as it is evident from molecular modeling studies [66] and affinity data for ligands synthesized in this work, the binding pocket has a shape that on the one hand quickly leads to a steric clash with bulky substituents (cf. Figure 8A and B), but on the other hand has a definite capacity to accommodate a wide variety of other substituents in the 2-position provided they satisfy some essential criteria, which will be described further on.

Thus, the possibility to modify the 2-position made GalNAc a good lead compound for the generation of a directed library of small drug-like molecules. By removing the anomeric center, and replacing the acetamido group with an azide, a versatile scaffold (**4**, Figure 6B) was created. It was then used to generate a

library of 1,4-disubstituted-1,2,3-triazoles attached directly to the tetrahydropyran ring via a Huisgen 1,3-dipolar cycloaddition.

### 3.1.5 Structure and topology of the binding site

A closer look at the binding pocket surrounding the 2-position of GalNAc reveals a dumbbell-shaped cavity (Figure 8A).



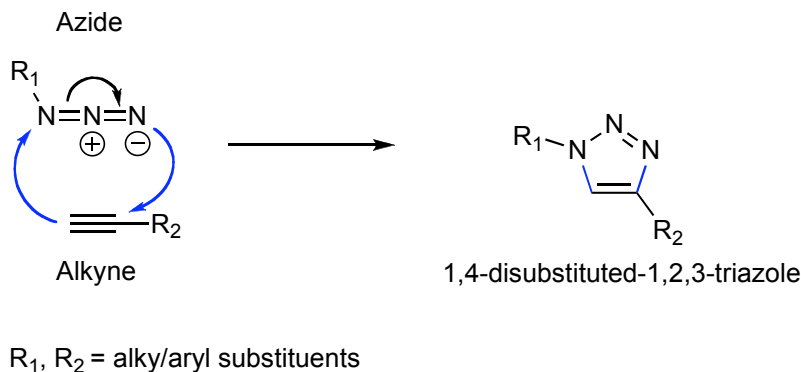
**Figure 8.** (A) A model of a hypothetical substituted triazole compound **5**, similar in structure to the compounds generated in this work, docked with the ASGP-R H1-CRD [66]. Colors: red = positively charged amino acids; blue = negatively charged amino acids; purple = Ca<sup>2+</sup>; green = polar amino acids; brown = hydrophobic amino acids; grey = aromatic amino acids. (B) Schematic representation of the orientation of the 4-substituent of the triazole in the dumbbell binding pocket.

This may explain why bulky and/or long substituents in the 2-position cause a drop in the affinity due to a direct steric clash with the protein surface (Figure 8B) [67]. Nevertheless, the dumbbell-shaped binding pocket offers enough room for the substituents to seek interactions at the sides, by orienting themselves approx. 90° with respect to the scaffold-triazole axis (Figure 8B).

### 3.1.6 The Huisgen 1,3-dipolar cycloaddition

In general, cycloaddition reactions are defined as two components coming together to form two new  $\sigma$ -bonds leading to ring formation. Cycloadditions are pericyclic reactions, proceeding through a characteristic cyclic transition state with a concerted movement of electrons resulting in a simultaneous breaking and forming of bonds. Furthermore, they are usually suprafacial [68].

In the Huisgen 1,3-dipolar cycloaddition of azides and alkynes [69] (Figure 9), the azide is a 1,3-dipole. This group is isoelectronic with an allyl anion, having a conjugated system of three p-orbitals on three atoms forming a four electron conjugated system. Due to its mesomeric structure, both ends of the 1,3-dipole have nucleophilic and electrophilic properties. The other component of the cycloaddition reaction, the alkyne, is a dipolarophile. There is a wide possibility of structures of 1,3-dipoles, with the participating atoms being carbon, nitrogen, oxygen or sulphur in any combination, with double or triple bonds connecting them.

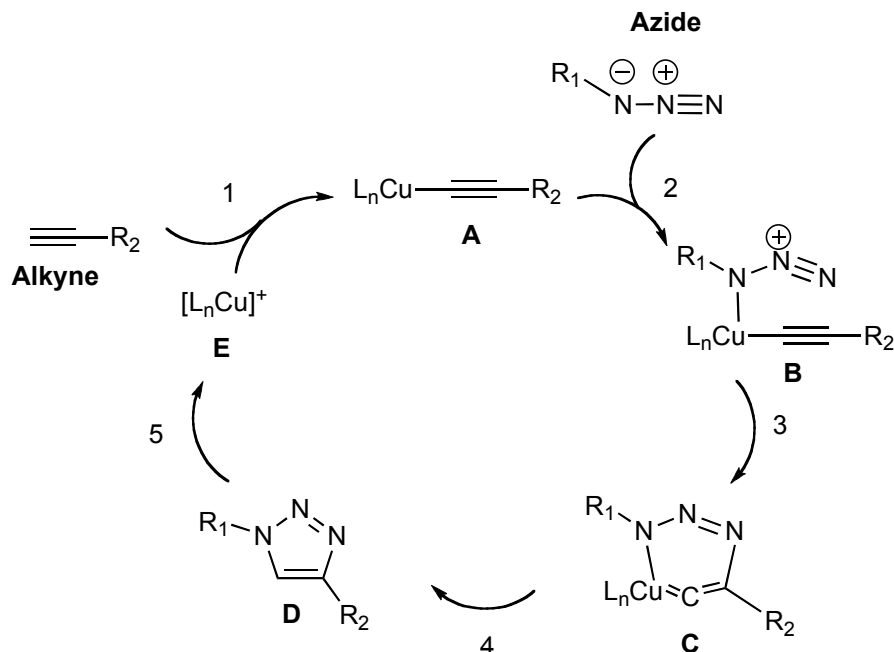


**Figure 9.** Mechanism of the classical Huisgen 1,3-dipolar cycloaddition between an alkyne and an azide.

The mechanism of the reaction is basically an electron shift (Figure 9). Two  $\pi$ -electrons of the dipolarophile and four electrons of the 1,3-dipole are shifted in the concerted and pericyclic manner, as described before. Electron-donating and withdrawing substituents modify the reactivity.

The regioselectivity of the reaction depends on steric and electronic characteristics of the starting materials, and is partially predictable, with the 1,4-disubstituted triazole being favoured on steric grounds. Furthermore, in order for the reaction to go to a fast completion, either the dipolarophile or the dipole must be electron rich, while the other must be electron poor. Therefore, the discovery of copper(I) catalysts for this process [70] which then leads exclusively to the 1,4-product, and increases the rate of the reaction, opened up many possibilities for applying the reaction in organic synthesis [71].

The reaction taking place via a copper(I) catalyst, is a stepwise reaction, and can therefore no longer be called a concerted classical Huisgen 1,3-dipolar cycloaddition [68] (Figure 10).

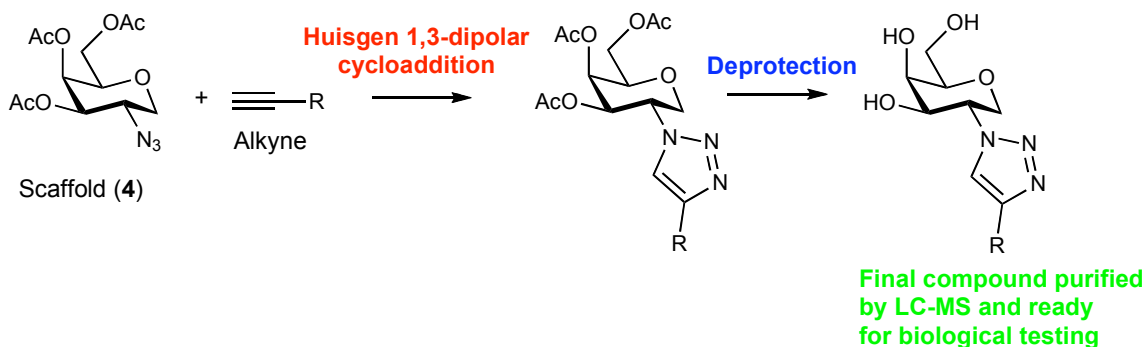


**Figure 10.** Cu(I)-catalyzed Huisgen 1,3-dipolar cycloaddition reaction cycle. Adapted from Rostovtsev *et al.* [72]

In the first step (1), the alkyne is activated and forms a Cu(I)-acetylide complex **A** [72-73]. In the second step (2), the Cu(I)-acetylide complex reacts with the azide leading to the formation of intermediate **B**. Step 3 is called ligation, and proceeds to form a six-membered copper-containing intermediate **C**. Step 4 finally leads to the formation of a thermally and hydrolytically stable triazole (**D**). The copper(I) catalyst **E** is liberated (Step 5) and reacts with the next alkyne (Step 1). A heterocyclic chelate ligand (L, Figure 10) is often added to accelerate the rate of the reaction [74], however, the reaction can still proceed under “ligand-free” conditions [73].

Due to the structure of intermediate **B**, this reaction is highly regioselective, and leads exclusively to the formation of a 1,4-disubstituted 1,2,3-triazole. However, the yield does greatly depend on the nature of both the dipole and the dipolarophile [68].

Thus, due to its high regioselectivity, which allows the substituent in the 4-position of the triazole to point directly into the dumbbell-shaped binding pocket (Figure 8), and the large variety of available alkynes, the copper(I)-catalyzed Huisgen 1,3-dipolar cycloaddition was chosen for the generation of the directed library (Figure 11).



**Figure 11.** General strategy for the synthesis of compounds for the directed ASGP-R ligand library.

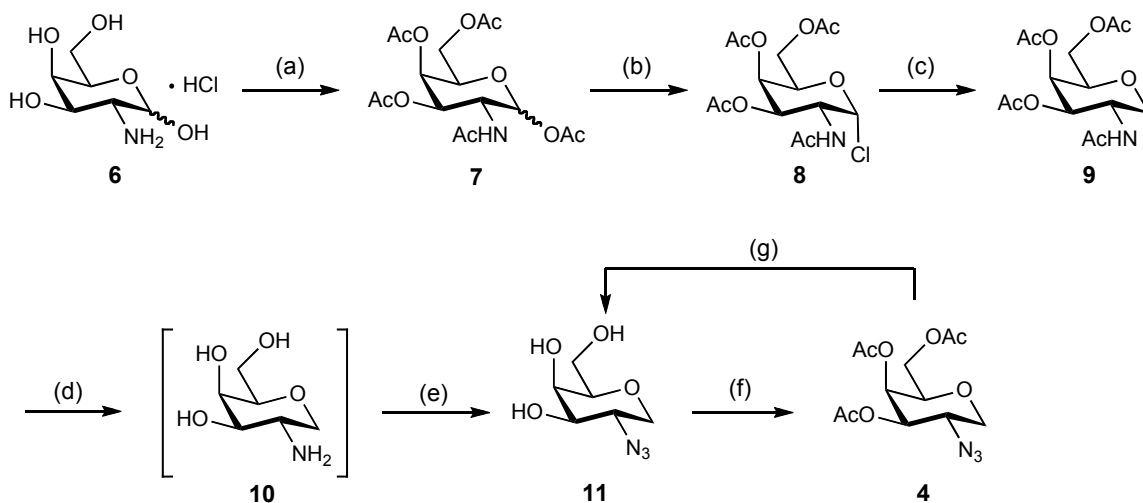
The compounds produced in the directed library were then tested for affinity towards the ASGP-R using a competitive binding assay (Appendix 1).

## 3.2 Results and Discussion

### 3.2.1 Scaffold synthesis

For the synthesis of the scaffold (**4**, Scheme 1), galactosamine hydrochloride (**6**) was peracetylated using acetic anhydride in pyridine ( $\rightarrow$  **7**). Treatment of *N*-acetylgalactosamine with acetyl chloride, a known procedure for synthesizing glycosyl chlorides [75], resulted in poor yields of **8** (< 30%). Therefore, a new procedure was developed, which involved treating 2-acetamido-1,3,4,6-tetra-*O*-acetyl-2-deoxy-*D*-galactopyranoside (**7**) with  $\text{TiCl}_4$  in DCM, giving yields of up to

89% of **8**. The latter was then reductively dehalogenated using  $(\text{Bu})_3\text{SnH}$  and AIBN in refluxing toluene, yielding **9** quantitatively, which upon deacetylation ( $\rightarrow$  **10**), amine-azide exchange [76] ( $\rightarrow$  **11**) and subsequent peracetylation furnished the scaffold **4** in a 76% yield over three steps.

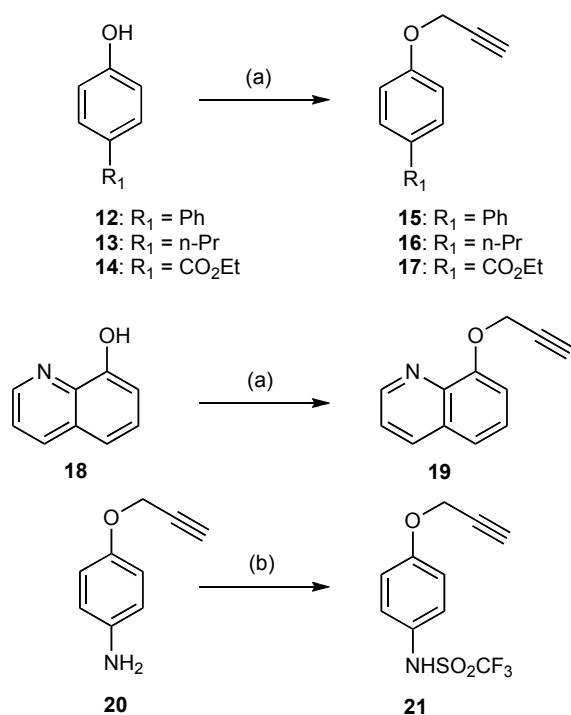


**Scheme 1.** (a)  $\text{Ac}_2\text{O}$ , pyridine,  $0\text{ }^\circ\text{C} \rightarrow \text{r.t.}$ , 24 h, 89%; (b)  $\text{TiCl}_4$ , DCM, r.t., 3 d, 89%; (c)  $(\text{Bu})_3\text{SnH}$ , AIBN, PhMe, reflux, 1 h 20 min, quant.; (d) KOH, 18-crown-6, dioxane/ $\text{H}_2\text{O}$ , reflux, 6.5 h; (e)  $\text{TfN}_3$ ,  $\text{NaHCO}_3$ ,  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ ,  $\text{H}_2\text{O}/\text{PhMe}/\text{MeOH}$ , r.t., 24 h; (f)  $\text{Ac}_2\text{O}$ , pyridine, r.t., 24 h, 76% (3 steps); (g)  $\text{NaOMe}/\text{MeOH}$ , r.t., 3 h, 100%.

Compound **4** was deprotected under standard Zemplén conditions yielding **11** quantitatively. The latter was submitted as a control compound for biological testing.

### 3.2.2 Synthesis of phenyl propargyl ethers

For the synthesis of substituted phenyl propargyl ethers (Scheme 2), the corresponding substituted phenols (**12**, **13** or **14**) were alkylated with propargyl bromide in refluxing acetone, using  $K_2CO_3$  as base ( $\rightarrow$  **15**, **16** or **17**, respectively). Compound **19** was synthesized by an analogous procedure starting from 8-hydroxyquinoline (**18**). Compound **21** was obtained in a 69% yield by treating **20** [77] with triflic anhydride in DCM in the presence of triethylamine.

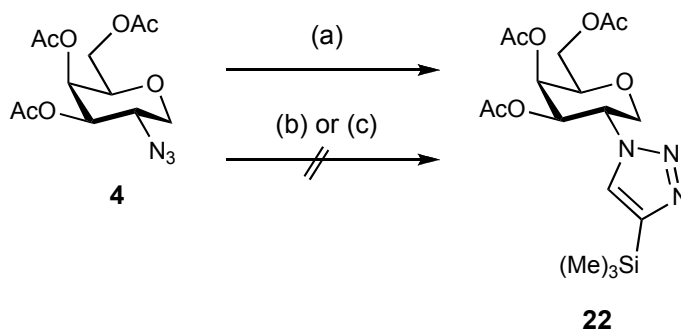


**Scheme 2.** (a) Propargyl bromide,  $K_2CO_3$ , acetone, reflux, 5 h, (**15**: 84%; **16**: 99%; **17**: 73%; **19**: 88%); (b)  $Tf_2O$ ,  $Et_3N$ , DCM, 0 °C  $\rightarrow$  r.t., 24 h, 69%.



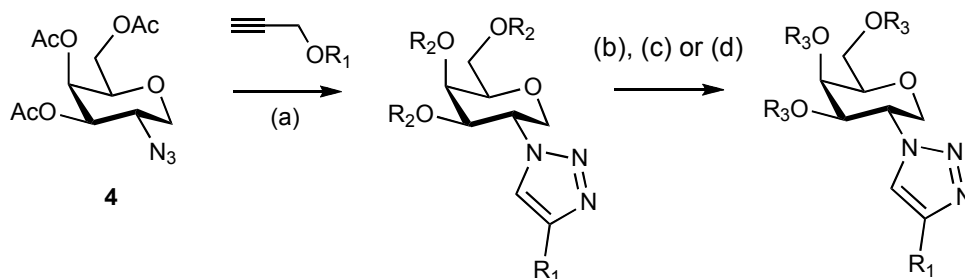
### 3.2.3 Library synthesis

When the Huisgen 1,3-dipolar cycloaddition was done under the relatively popular “click chemistry” conditions [78] (Scheme 3), *i.e.* CuSO<sub>4</sub> as the source of Cu(II), and sodium ascorbate as the reducing agent for the formation of catalytically active Cu(I) in *t*-BuOH/water, the reactions gave only low yields or even failed to give detectable amounts of product. Therefore, a new procedure was devised (Scheme 3), which involved the use of CuCl as the direct source of Cu(I), triethylamine as base, in DCM. The latter proved to be the best solvent for the reactions, with the highest yield of 97% obtained for compound **22**, whereas when the same reaction was done with THF, no product could be detected.



**Scheme 3.** (a) Ethynyltrimethylsilane, CuCl, Et<sub>3</sub>N, DCM, r.t., 24 h, 97%; (b) Ethynyltrimethylsilane, CuCl, Et<sub>3</sub>N, THF, r.t., 24 h; (c) Ethynyltrimethylsilane, CuSO<sub>4</sub>•5H<sub>2</sub>O, sodium ascorbate, *t*-BuOH/H<sub>2</sub>O, r.t., 24 h.

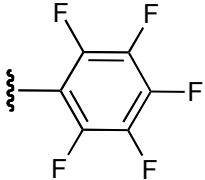
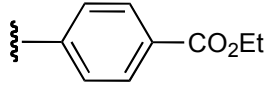
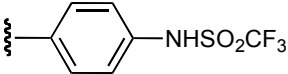
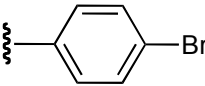
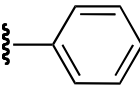
Thus, Scheme 4 illustrates the application of the new synthetic strategy, with the results summarized in Table 1.



**Scheme 4.** (a) CuCl, Et<sub>3</sub>N, DCM, r.t., 24 h; (b) NaOMe/MeOH, 3 h, r.t.; (c) H<sub>2</sub>O/MeOH/Et<sub>3</sub>N (5:5:1), r.t., 24 h; (d) NaOH, H<sub>2</sub>O/MeOH, r.t., 24 h.

**Table 1.** Summary of compounds synthesized by the strategy shown in Scheme 4.

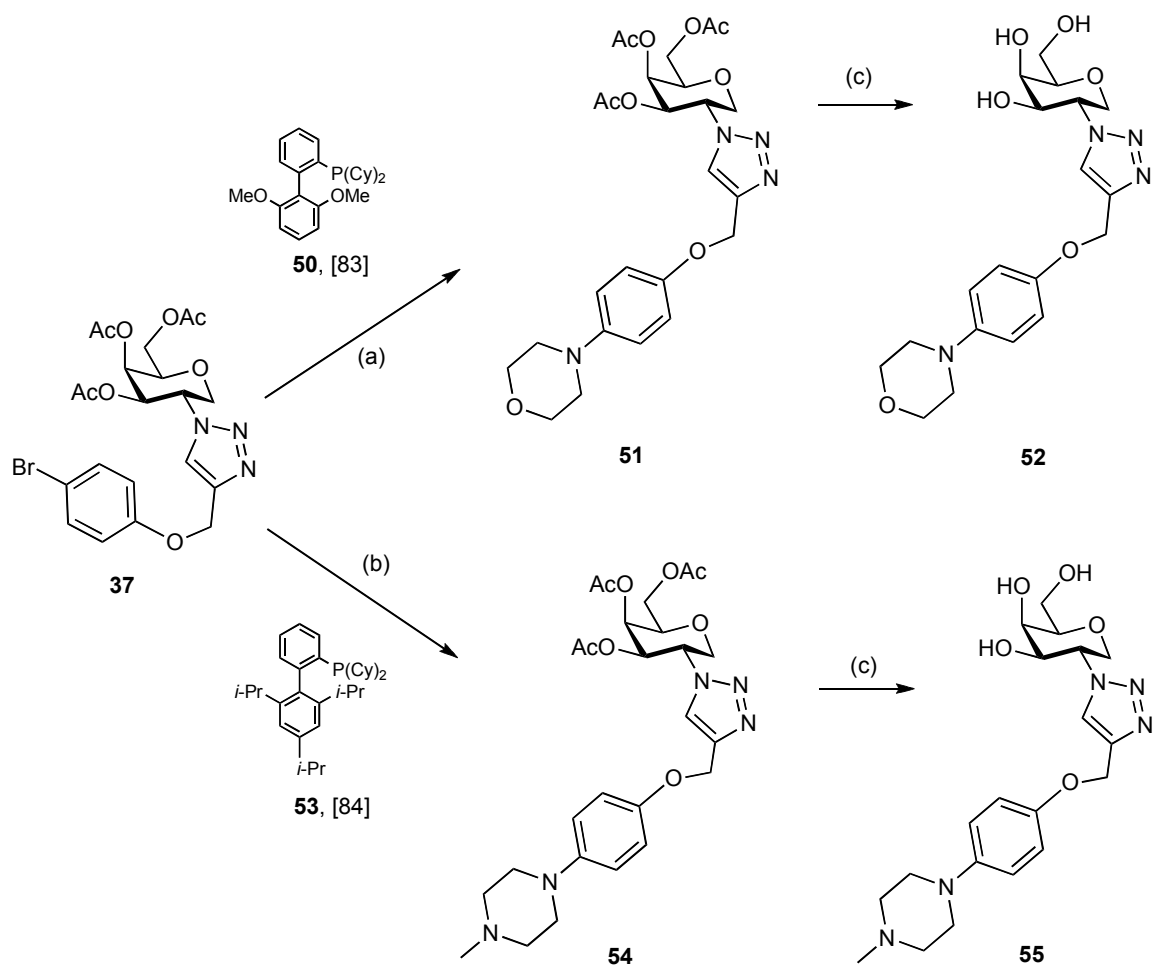
Entry	Alkyne	R <sub>1</sub> =	R <sub>2</sub> = Ac	Yield (%)	R <sub>3</sub> = H	Yield (%)
1	<b>15</b>		<b>28</b>	74	<b>38</b>	93
2	<b>23</b> [79]		<b>29</b>	85	<b>39</b>	34
3	<b>16</b>		<b>30</b>	96	<b>40</b>	41
4	<b>24</b> [77]		<b>31</b>	95	<b>41</b>	33
5	<b>25</b> [80]		<b>32</b>	98	<b>42</b> <sup>(1)</sup>	81
6	<b>19</b>		<b>33</b>	80	<b>43</b>	85

7	<b>26</b> [81]		<b>34</b>	73	<b>44</b>	68
8	<b>17</b>		<b>35</b>	82	<b>45<sup>(1)</sup></b>	74
9	<b>21</b>		<b>36</b>	41	<b>46</b>	73
10	<b>27</b> [79]		<b>37<sup>(1)</sup></b>	84	–	–
11	<b>47</b>		<b>48</b>	93	<b>49</b>	45

(1) Compounds **32** and **35** were deprotected using conditions (c) and (d) (Scheme 4), respectively, while all other compounds were deprotected using conditions (b). Compound **37** was not deprotected since it was intended for use in coupling reactions.

Morpholine and *N*-methylpiperazine often appear as privileged motifs in medicinal chemistry, therefore, organometallic chemistry was employed to couple these to compound **37**. Thus, compound **51** was synthesized via a Buchwald-Hartwig Pd-catalyzed coupling [82] (Scheme 5). The aryl bromide (**37**) was treated with morpholine in the presence of Pd<sub>2</sub>(dba)<sub>3</sub>, ligand **50** [83], Cs<sub>2</sub>CO<sub>3</sub> in toluene, furnishing **51** in a 40% yield. Using THF or dioxane instead of toluene gave no detectable product. Compound **51** was deacetylated under standard Zemplén conditions to give **52** in a 57% yield.

Attempts to synthesize compound **54** by the same procedure failed. Therefore, **54** was synthesized using a different ligand [84] (**53**, Scheme 5) for the catalyst. Zemplén deacetylation gave the desired product (**55**), in a 44% yield over two steps.



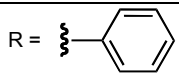
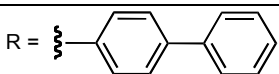
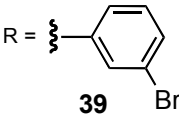
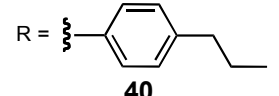
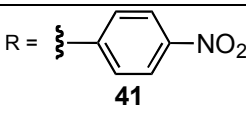
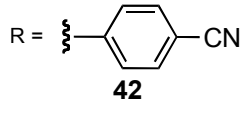
**Scheme 5.** (a) morpholine, Pd<sub>2</sub>(dba)<sub>3</sub>, Cs<sub>2</sub>CO<sub>3</sub>, PhMe, 80 °C, 24 h, 40%; (b) N-methylpiperazine, Pd<sub>2</sub>(dba)<sub>3</sub>, Cs<sub>2</sub>CO<sub>3</sub>, PhMe, 80 °C, 24 h; (c) NaOMe/MeOH, r.t., 24 h, (**52**: 57%; **55**: 44%).

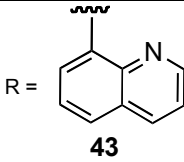
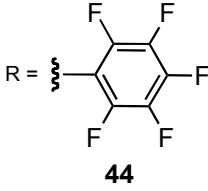
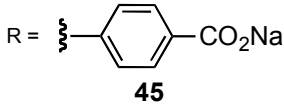
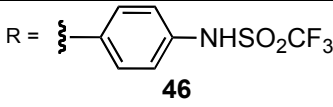
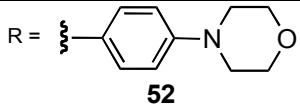
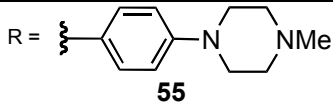
### 3.2.4 Competitive binding assay

(Performed by Daniela Stokmaier, Institute of Molecular Pharmacy, University of Basel)

The final compounds from the directed library were then tested for their affinity towards the ASGP-R H1-CRD using a competitive binding assay (Appendix 1), and the results are summarised in Table 2.

**Table 2.** Competitive binding assay results of directed library compounds.

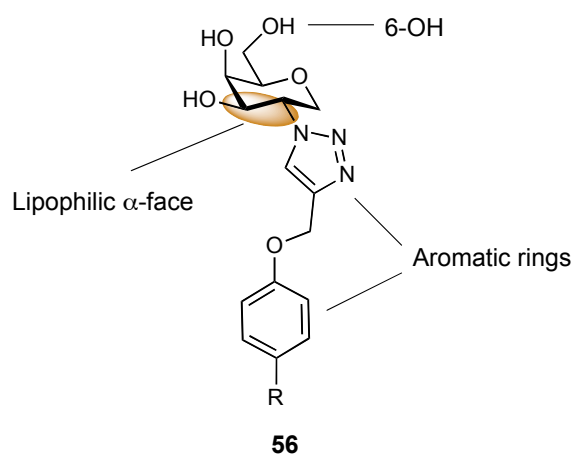
Entry	Compound	$rIC_{50} \pm s. d.$
1	GalNAc	$1.0 \pm 0.3$
2	<b>11</b>	$13.7 \pm 4.3$
3	R =  <b>49</b>	$0.8 \pm 0.2$
4 <sup>(1)</sup>	R =  <b>38</b>	-
5 <sup>(1)</sup>	R =  <b>39</b>	-
6 <sup>(1)</sup>	R =  <b>40</b>	-
7 <sup>(1)</sup>	R =  <b>41</b>	-
8 <sup>(1)</sup>	R =  <b>42</b>	-

9 <sup>(1)</sup>	 <p style="text-align: center;"><b>43</b></p>	-
10	 <p style="text-align: center;"><b>44</b></p>	2.6 ± 0.7
11 <sup>(2)</sup>	 <p style="text-align: center;"><b>45</b></p>	8.7
12 <sup>(1)</sup>	 <p style="text-align: center;"><b>46</b></p>	-
13 <sup>(1)</sup>	 <p style="text-align: center;"><b>52</b></p>	-
14	 <p style="text-align: center;"><b>55</b></p>	1.1

(1) Compounds were not soluble in the 10% DMSO/buffer solution used for the assay.

(2) Compound **45** was turned into a sodium salt by dissolving the acid precursor in MeOH, and adding an equimolar amount of 0.1 M NaOH<sub>(aq)</sub>, followed by solvent removal in vacuo.

Solubility of the library products proved to be a crucial point. As can be seen from the data (Table 2), most compounds were insoluble in the 10% DMSO/buffer solution used in the assay. This is probably due to the presence of two lipophilic aromatic moieties (compound **56**, Figure 12), which are known to  $\pi$ -stack together, and thus hinder solubilization. Furthermore, the lipophilic  $\alpha$ -face of the scaffold, formed by the non-polar C-H groups on the underside of the scaffold ring, further increases the probability of stacking and/or aggregate formation. For the compounds that did dissolve, it seems to be beneficial not to remove the 6-OH group from the original scaffold as it could be vital for aqueous solubility.



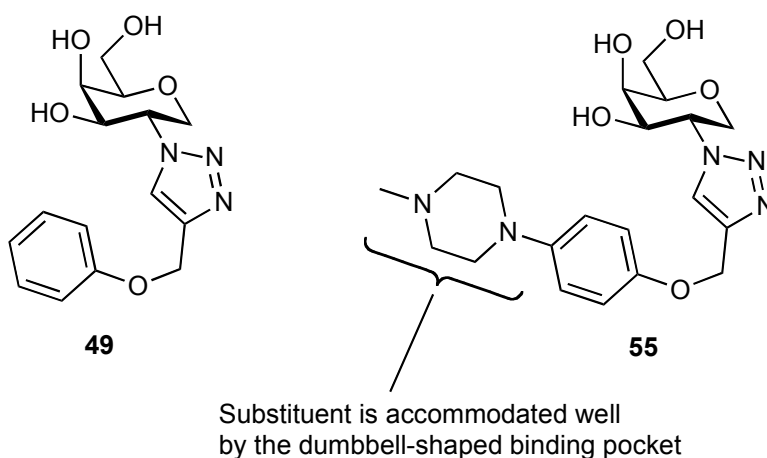
**Figure 12.** Compound **56** represents a general structure of the compounds synthesized in the directed library.

Nevertheless, the biological testing results showed that two compounds (**49** and **55**, Table 2) with quite different structures bound to the H1-CRD with an affinity comparable to that of GalNAc - the best small molecular weight ligand for the ASGP-R so far.

### 3.3 Conclusion

The results presented in this chapter indicate that the strategy for synthesizing the triazole based ligands for the ASGP-R (see Figure 11, p. 78) is efficient, and was successfully used to generate a set of compounds with diverse functional groups. However, the main drawback of the final compounds was their poor aqueous solubility, which could perhaps be overcome by the addition of solubilizing groups.

By comparing the structures of **49** and **55** (Figure 13), and assuming that both compounds have the same binding mode, it is evident that there is ample room in the binding pocket for accommodating substituents in the *para*-position of the phenyl group. This further reinforces the “dumbbell-shaped binding pocket” hypothesis that was central to this strategy. Hence, after further molecular modeling studies, it would be worthwhile to generate compounds with other substituents in the *para*-position, that would interact favourably with the protein and improve the affinity even more.



**Figure 13.** A comparison of the structures of the best binding ligands generated in the directed library described in this thesis.



## 3.4 Experimental

General methods are described in Appendix 4, p. 147.

### 3.4.1 Scaffold synthesis: compound 4

**2-Acetamido-1,3,4,6-tetra-O-acetyl-2-deoxy-D-galactopyranoside (7):** Acetic anhydride (23.7 ml, 230 mmol) was added dropwise to a suspension of galactosamine hydrochloride (**6**, 5.00 g, 23.0 mmol) in dry pyridine (50 ml) at 0 °C under argon. After the addition was complete, the reaction mixture was allowed to reach r.t. and stirred overnight. The solvent was removed in vacuo, and ice was added to the resultant paste. The resultant white precipitate was filtered, washed with H<sub>2</sub>O (2 × 100 ml) and dried (P<sub>2</sub>O<sub>5</sub>) giving compound **7** (8.04 g, 89%) as a white powder. The analytical data were identical to those found in literature [85].

**2-Acetamido-3,4,6-tri-O-acetyl-2-deoxy- $\alpha$ -D-galactopyranosyl chloride (8):** Titanium tetrachloride (185  $\mu$ l, 1.69 mmol) was added to a suspension of **7** (500 mg, 1.28 mmol) in dry DCM (5 ml). The mixture was stirred at r.t. under argon for 3 d. The solvent was removed in vacuo, and the residue was purified by flash chromatography on silica gel (petrol ether/EtOAc 2:3 → 1:4), yielding **8** (419 mg, 89% R<sub>f</sub> 0.27 B). The analytical data were identical to those found in literature [86].

**2-Acetamido-3,4,6-tri-O-acetyl-1,5-anhydro-2-deoxy-D-galactitol (9):** Compound **8** (2.24 g, 6.12 mmol) was dissolved in dry PhMe (33 ml) and the resultant solution was degassed in an ultrasound bath under a steady flow of argon for 30 min. Tributyltin hydride (2.14 g, 7.35 mmol) and AIBN (ca. 25 mg) were added and the resultant mixture was refluxed under argon for 1h 20 min. The solvent was removed in vacuo, and the residue was purified by flash chromatography on

silica gel (PhMe/EtOAc (1:4 → 1:9 → 0:1), yielding **9** (2.03 g, quant.,  $R_f$  0.26 C).  $^1\text{H-NMR}$  (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  1.93 (s, 3H, NAc), 2.02, 2.04, 2.14 (s, 9H, 3  $\times$  OAc), 3.15 (t,  $J$  = 11.1 Hz, 1H, H1 ax.), 3.76 (m, 1H, H5), 4.06 (m, 2H, H6), 4.18 (dd,  $J$  = 5.2, 11.3 Hz, 1H, H1 eq.), 4.40 (m, 1H, H2), 4.92 (dd,  $J$  = 3.3, 11.2 Hz, 1H, H3), 5.35 (m, 1H, H4), 5.69 (d,  $J$  = 8.0 Hz, 1H, NH);  $^{13}\text{C NMR}$  (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  20.6, 20.7, 20.8 (3  $\times$   $\text{CH}_3$ , OAc), 23.2 ( $\text{CH}_3$ , NAc), 46.5 (C2), 62.2 (C6), 67.1 (C4), 68.6 (C1), 71.5 (C3), 75.0 (C5), 170.25, 170.3, 170.5, 171.4 (4  $\times$  C=O); Anal. calcd. for  $\text{C}_{14}\text{H}_{21}\text{NO}_8$ : C, 50.75; H, 6.39; N, 4.23. Found: C, 51.07; H, 6.49; N, 4.12.

**2-Amino-1,5-anhydro-D-galactitol (10)**: Compound **9** (3.17 g, 9.58 mmol) was dissolved in dioxane (8.5 ml). Water (6.5 ml), 18-crown-6 (21.1 mg) and potassium hydroxide (840 mg, 15.0 mmol) were added. The reaction mixture was refluxed for 5 h. Another 5 eq. (2.10 g) of potassium hydroxide was then added, and the mixture was refluxed for another 1.5 h. According to TLC, no starting material was left. The solvent was removed in vacuo and the crude product (**10**, 3.91 g) was used without further purification in the next step.

**3,4,6-Tri-O-acetyl-1,5-anhydro-2-azido-2-deoxy-D-galactitol (5)**: *Triflyl azide stock solution preparation*: sodium azide (3.49 g, 53.7 mmol) was dissolved in water (8.7 ml). Toluene (8.7 ml) was added, and the mixture was cooled down to 0 °C with stirring. Then triflic anhydride (5.8 ml, 34.4 mmol) was added dropwise. The biphasic reaction mixture was stirred vigorously at 0 °C for 1 h and at 10 °C for another 2 h. The reaction mixture was neutralized with saturated aqueous  $\text{NaHCO}_3$ . The phases were separated, and the aqueous phase extracted with toluene (2  $\times$  8 ml). The organic layers were combined to give the triflyl azide stock solution.

*Amine-azide exchange*: compound **10** (1.56 g, 9.56 mmol),  $\text{NaHCO}_3$  (319 mg, 38.0 mmol) and  $\text{CuSO}_4 \cdot \text{H}_2\text{O}$  (95.0 mg, 380  $\mu\text{mol}$ ) were dissolved in water (6.4 ml). The triflyl azide stock solution (21 ml, 9.56 mmol) was added. The biphasic reaction mixture was made homogenous by the dropwise addition of MeOH. The

mixture was stirred at r.t. overnight. The colour of the mixture turned from blue to green. The solvent was removed in vacuo; the residue contained 1,5-anhydro-2-azido-2-deoxy-D-galactitol (**11**).

The residue containing **11** (1.81 g, 9.6 mmol) was dissolved in dry pyridine (36 ml, 447 mmol), and acetic anhydride (8.2 ml, 86.2 mmol) was added. The reaction mixture was stirred at r.t. under argon overnight. The solvent was removed in vacuo. The mixture was purified by flash chromatography (petrol ether/EtOAc 20:1 → 9:1) to give **4** (2.31 g, 76%,  $R_f$  0.59 D).  $[\alpha]_D = + 0.04^\circ$  (c 1, CHCl<sub>3</sub>); <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  2.09, 2.16, 2.22 (s, 9H, 3 × OAc), 3.24 (t,  $J = 11.3$  Hz, 1H, H1 ax.), 3.79 (dt,  $J = 0.9, 6.5$  Hz, 1H, H5), 3.93 (m, 1H, H2), 4.08 (d,  $J = 6.5$  Hz, 2H, H6), 4.12 (dd,  $J = 5.4, 11.6$  Hz, 1H, H1 eq.), 4.91 (dd,  $J = 3.3, 10.4$  Hz, 1H, H3), 5.40 (d,  $J = 0.9$  Hz, 1H, H4); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  20.6, 20.8, 22.1 (3 × CH<sub>3</sub>), 55.9 (C2), 61.9 (C6), 67.1 (C4), 68.1 (C1), 73.4 (C3), 74.9 (C5); Anal. Calcd. for C<sub>12</sub>H<sub>17</sub>N<sub>3</sub>O<sub>7</sub>: C, 45.72; H, 5.43; N, 13.33. Found: C, 45.90; H, 5.44; N, 13.18.

**3,4,6-Tri-O-acetyl-1,5-anhydro-2-azido-2-deoxy-D-galactitol (11)**: Compound **4** (20 mg, 63.4  $\mu$ mol) was dissolved in MeOH (2 ml) and sodium metal (10 mg) was added. The solution was stirred overnight, the solvent was removed in vacuo and the residue was purified on an RP-C18 column (H<sub>2</sub>O/MeOH 20:0 → 9:1, stepwise gradient) yielding **11** (12 mg, quant.,  $R_f$  0.37 E).  $[\alpha]_D = - 0.12^\circ$  (c 1, MeOH); <sup>1</sup>H-NMR (500 MHz, CD<sub>3</sub>OD):  $\delta$  3.07 (t,  $J = 11.1$  Hz, 1H, H1 ax.), 3.38 (m, 1H, H5), 3.50 (dd,  $J = 3.3, 9.9$  Hz, 1H, H3), 3.62-3.72 (m, 3H, H2, H6), 3.83 (d,  $J = 3.0$  Hz, 1H, H4), 3.95 (dd,  $J = 5.4, 11.2$  Hz, 1H, H1 eq.); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  60.6 (C2), 62.9 (C6), 69.1 (C1), 70.6 (C4), 75.3 (C3), 81.0 (C5); ESI-MS: Calcd. for C<sub>6</sub>H<sub>11</sub>N<sub>3</sub>O<sub>4</sub>Na [M+Na]<sup>+</sup>: 212.06; Found m/z 213.94.

### 3.4.2 Synthesis of substituted phenyl propargyl ethers

**1-Phenyl-4-(2-propynyloxy)-benzene (15)**: 1-Hydroxybiphenyl (**12**, 1.19 g, 6.99 mmol) was dissolved in dry acetone (7 ml), and K<sub>2</sub>CO<sub>3</sub> (3.01 g, 21.7 mmol) was

added, followed by propargyl bromide (855  $\mu$ l, 7.92 mmol). The mixture was refluxed for 5 h, after which it was poured onto ice and the resultant precipitate was filtered off, dried and recrystallized from hexane to yield **15** (1.21 g, 84%). The analytical data were identical to those found in literature [87].

**8-(2-Propynyloxy)-quinoline (19)**: Prepared according to the procedure for **15**, from 8-hydroxyquinoline (**17**, 1.01 g, 7.01 mmol), propargyl bromide in PhMe (855  $\mu$ l, 791 mmol) and  $K_2CO_3$  (3.01 g, 21.7 mmol) in refluxing acetone (7 ml). The residue was purified by flash chromatography on silica gel (petrol ether/EtOAc 9:1  $\rightarrow$  7:3), yielding **19** (1.12 g, 88%,  $R_f$  0.20 D).  $^1H$ -NMR (500 MHz,  $CDCl_3$ ):  $\delta$  2.53 (m, 1H, CH), 5.03 (d,  $J$  = 2.2 Hz, 2H,  $OCH_2$ ), 7.26-7.50 (m, 4H, H3, H6, H7, H8 quinoline), 8.12 (m, 1H, H4 quinoline), 8.94 (m, 1H, H2 quinoline);  $^{13}C$  NMR (125 MHz,  $CDCl_3$ ):  $\delta$  56.5 ( $OCH_2$ ), 46.1 (CH), 78.3 ( $C_Q$ ), 109.9 (C8 quinoline), 120.7, 121.7, 126.4 (3C, C3, C6, C7 quinoline), 129.5 (C5 quinoline), 135.9 (C4 quinoline), 140.4 (C10 quinoline), 149.5 (C2 quinoline), 153.1 (C9 quinoline); Anal. calcd. for  $C_{12}H_9NO$ : C, 78.67; H, 4.95; N, 7.65. Found: C, 77.19; H, 5.19; N, 7.39.

**1-Propyl-4-(2-propynyloxy)-benzene (16)**: Prepared according to the procedure for **15**, from *p*-propylphenol (953 mg, 7.01 mmol), propargyl bromide in PhMe (855  $\mu$ l, 791 mmol) and  $K_2CO_3$  (3.01 g, 21.7 mmol) in refluxing acetone (7 ml). The mixture was poured onto ice and extracted with  $Et_2O$  (2  $\times$  20 ml). The combined organic fractions were dried ( $Na_2SO_3$ ), filtered, and the solvents were removed under low vacuum. The residue was purified by flash chromatography on silica gel (petrol ether 100%), to give **16** (1.20 g, 99%,  $R_f$  0.66 F).  $^1H$ -NMR (500 MHz,  $CDCl_3$ ):  $\delta$  0.92 (t,  $J$  = 7.3 Hz, 3H,  $CH_3$ ), 1.60 (m, 2H,  $CH_2CH_2CH_3$ ), 2.50 (m, 3H,  $CH_2CH_2CH_3$ , CH), 4.64 (d,  $J$  = 2.2 Hz, 2H,  $OCH_2$ ), 6.87-6.89, 7.08-7.10 (AA'BB', 4H,  $C_6H_4$ );  $^{13}C$  NMR (125 MHz,  $CDCl_3$ ):  $\delta$  13.7 ( $CH_3$ ), 24.6 ( $CH_2CH_2CH_3$ ), 37.1 ( $CH_2CH_2CH_3$ ), 55.7 ( $OCH_2$ ), 75.2 (CH), 78.7 ( $C_Q$ ), 114.6 (2C, 2  $\times$  C-ortho), 129.2 (2C, 2  $\times$  C-meta), 135.7 (C-para), 155.6 (C-ipso); Anal. calcd. for  $C_{12}H_{14}O$ : C, 82.72; H, 8.10. Found: C, 82.66; H, 8.14.

**Ethyl 4-(2-propynyloxy)-benzoate (17):** Ethyl 4-hydroxybenzoate (500 mg, 3.01 mmol) was dissolved in dry acetone (5 ml), and  $K_2CO_3$  (580 mg, 4.21 mmol) was added, followed by propargyl bromide (651  $\mu$ l, 6.02 mmol), and the mixture was refluxed for 3 h. The mixture was then diluted with DCM (50 ml), washed with  $H_2O$  (25 ml) and brine (25 ml), dried ( $Na_2SO_4$ ), and the solvent was removed in vacuo. The residue was purified by flash chromatography on silica gel (petrol ether/EtOAc 20:1  $\rightarrow$  4:1), yielding **17** (452 mg, 73% R<sub>f</sub> 0.26 F).  $^1H$ -NMR (500 MHz,  $CDCl_3$ ):  $\delta$  1.36 (t,  $J = 7.2$  Hz, 3H,  $OCH_2CH_3$ ), 2.55 (t,  $J = 2.4$  Hz, 1H, CH), 4.33 (q,  $J = 7.2$  Hz, 2H,  $OCH_2CH_3$ ), 4.73 (d,  $J = 2.4$  Hz, 2H,  $CH_2$ , propynyl), 6.98 (AA' of AA'BB',  $J = 9.0$  Hz, 2H, 2  $\times$  CH-*ortho*), 8.00 (BB' of AA'BB',  $J = 9.0$  Hz, 2H, 2  $\times$  CH-*meta*);  $^{13}C$  NMR (125 MHz,  $CDCl_3$ ):  $\delta$  14.3 ( $OCH_2CH_3$ ), 55.7 ( $OCH_2$ ), 60.5 ( $OCH_2CH_3$ ), 76.0 (CH), 77.8 ( $C_Q$ ), 114.4 (2C, 2  $\times$  C-*ortho*), 123.7 (C-*para*), 131.4 (2C, 2  $\times$  C-*meta*), 161.0 (C-*ipso*), 166.2 (C=O); Anal. calcd. for  $C_{12}H_{12}O_3$ : C, 70.58; H, 5.92. Found: C, 70.58; H, 5.93.

**1-Trifluoromethylsulfonamido-4-(2-propynyloxy)benzene (21):** Compound **20** [77] (100 mg, 679  $\mu$ mol) was dissolved in dry DCM (3 ml). Triethylamine (103  $\mu$ l, 747  $\mu$ mol) was added, followed by the dropwise addition of triflic anhydride (123  $\mu$ l, 747  $\mu$ mol) at 0  $^\circ$ C. The resultant solution was allowed to reach r.t. and stirred under argon overnight. The solvent was removed in vacuo and the residue was purified by flash chromatography on silica gel (petrol ether/EtOAc 19:1  $\rightarrow$  9:1), yielding **21** (130 mg, 69% R<sub>f</sub> 0.27 F).  $^1H$ -NMR (500 MHz,  $CDCl_3$ ):  $\delta$  2.54 (t,  $J = 2.4$  Hz, 1H, CH); 4.69 (d,  $J = 2.4$  Hz, 2H,  $OCH_2$ ); 6.97 (AA' of AA'BB',  $J = 8.9$  Hz, 2H, 2  $\times$  CH-*ortho*); 7.23 (BB' of AA'BB',  $J = 8.9$  Hz, 2H, 2  $\times$  CH-*meta*);  $^{13}C$  NMR (125 MHz,  $CDCl_3$ ):  $\delta$  56.1 ( $OCH_2$ ), 76.1 (CH), 77.9 ( $C_Q$ ), 115.8 (2C, 2  $\times$  C-*ortho*), 119.8 ( $J = 322.8$  Hz,  $CF_3$ ), 126.6 (C-*para*), 126.9 (2C, 2  $\times$  C-*meta*), 157.2 (C-*ipso*); Anal. calcd. for  $C_{10}H_8NO_3F_3S_1$ : C, 43.01; H, 2.89; N, 5.02. Found: C, 42.95; H, 3.00; N, 4.97.

### 3.4.3 Library synthesis: Huisgen 1,3-dipolar cycloaddition

**1-(3,4,6-Tri-O-acetyl-1,5-anhydro-2-deoxy-D-galactitol-2-yl)-4-phenoxy-methyl-1,2,3-triazole (48):** Compound **4** (50.0 mg, 158  $\mu$ mol) was dissolved in dry DCM (3 ml). The mixture was degassed in an ultrasound bath under a flow of argon for 20 min. Copper(I) chloride (31.2 mg, 316  $\mu$ mol), DIPEA (54.4  $\mu$ l, 316  $\mu$ mol) and phenyl propargyl ether (**47**, 40.6  $\mu$ l, 316  $\mu$ mol) were added. The mixture was stirred under argon at r.t. for 24 h. The solvent was removed in vacuo, and the crude mixture was purified by flash chromatography (petrol ether/EtOAc 3:2) to yield compound **48** (66.5 mg, 93%,  $R_f$  0.25 D).  $^1\text{H-NMR}$  (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  1.81, 2.08, 2.19 (s, 9H, 3  $\times$   $\text{CH}_3$ ), 4.00-4.07 (m, 2H, H1 ax., H5), 4.15 (d,  $J$  = 6.4 Hz, 1H, H6), 4.33 (dd,  $J$  = 5.1, 11.6 Hz, 1H, H1 eq.), 4.92 (dt,  $J$  = 5.0, 11.0 Hz, 1H, H2), 5.22 (s, 2H,  $\text{CH}_2\text{OPh}$ ), 5.50 (dd,  $J$  = 3.2, 11.0 Hz, 1H, H3), 5.54 (d,  $J$  = 2.5 Hz, 1H, H4), 6.95-6.99, 7.26-7.30 (m, 5H,  $\text{C}_6\text{H}_5$ ), 7.62 (s, 1H, H triazole);  $^{13}\text{C-NMR}$  (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  20.3, 20.7, 20.7 (3  $\times$   $\text{CH}_3$ ), 56.0 (C2), 61.8 (2C,  $\text{CH}_2\text{OPh}$ , C6), 67.0 (C4), 68.7 (C1), 71.5 (C3), 75.3 (C5), 114.7, 121.4, 122.7, 129.6 (6C,  $\text{C}_6\text{H}_5$ ), 118.5 (C5 triazole), 144.3 ( $\text{C}_Q$  triazole), 169.3, 170.0, 170.5 (3  $\times$  C=O); ESI-MS: Calcd. for  $\text{C}_{21}\text{H}_{26}\text{N}_3\text{O}_8$   $[\text{M}+\text{H}]^+$ : 448.17; Found  $m/z$  448.25.

**1-(1,5-Anhydro-2-deoxy-D-galactitol-2-yl)-phenoxy-methyl-1,2,3-triazole (49):** Compound **48** (65.0 mg, 145  $\mu$ mol) was dissolved in dry MeOH (5 ml) and sodium metal (20.0 mg, 869  $\mu$ mol) was added. The solution was stirred at r.t. under argon for 3 h, after which the solvent was removed in vacuo and the residue was purified by LCMS to give compound **49** (21 mg, 45%,  $R_f$  0.52 G).  $[\alpha]_D = +36.4^\circ$  (c 0.17, MeOH);  $^1\text{H-NMR}$  (500 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  3.31 (t,  $J$  = 1.6 Hz, 1H, H5), 3.61-3.81 (m, 2H, H6), 3.84 (t,  $J$  = 11.1 Hz, 1H, H1 ax.), 3.99 (d,  $J$  = 3.0 Hz, 1H, H4), 4.14-4.18 (m, 2H, H1 eq., H3), 4.81 (dt,  $J$  = 5.0, 10.8 Hz, 1H, H2), 5.16 (s, 2H,  $\text{CH}_2$ ), 6.93-7.01 (m, 5H,  $\text{C}_6\text{H}_5$ ), 8.14 (s, 1H, H5 triazole);  $^{13}\text{C-NMR}$  (125 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  60.7 (C2), 62.3 ( $\text{CH}_2\text{OPh}$ ), 62.9 (C6), 69.8 (C1), 70.4 (C4), 73.5 (C3), 81.5 (C5), 122.2, 125.8, 130.6, 159.8 (6C,  $\text{C}_6\text{H}_5$ ), 115.9 (C5 triazole),

144.53 (C<sub>Q</sub> triazole); ESI-MS: Calcd. for C<sub>15</sub>H<sub>20</sub>N<sub>3</sub>O<sub>5</sub> [M+H]<sup>+</sup>: 322.14; Found m/z 322.07.

**1-(3,4,6-Tri-O-acetyl-1,5-anhydro-2-deoxy-D-galactitol-2-yl)-4-(4-phenylphenoxy)methyl-1,2,3-triazole (28)**: Prepared according to the procedure described for **48**, using **4** (50.0 mg, 160 μmol) and 1-phenyl-4-(2-propynyloxy)-benzene (**15**, 65.8 mg, 316 μmol). The compound was purified by flash chromatography (hexane/EtOAc 3:1 → 1:1) to give **28** (61.0 mg, 74%, R<sub>f</sub> 0.28 D). <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ 1.81, 2.07, 2.18 (s, 9H, 3 × CH<sub>3</sub>), 4.00-4.07 (m, 2H, H1 ax., H5), 4.15 (d, *J* = 6.4 Hz, 1H, H6), 4.34 (dd, *J* = 5.0, 11.6 Hz, 1H, H1 eq.), 4.93 (dt, *J* = 5.0, 11.0 Hz, 1H, H2), 5.26 (s, 2H, CH<sub>2</sub>OPh), 5.51 (dd, *J* = 3.2, 10.9 Hz, 1H, H3), 5.54 (d, *J* = 2.6 Hz, 1H, H4), 7.02-7.54 (m, 9H, biphenyl), 7.65 (s, 1H, H5 triazole); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>): δ 20.3, 20.6, 20.7 (3 × CH<sub>3</sub>), 56.0 (C2), 61.8 (C6), 62.0 (CH<sub>2</sub>OPh), 67.0 (C4), 68.7 (C1), 71.5 (C3), 76.3 (C5), 115.0 (C2 biphenyl), 122.8 (C5 triazole), 126.7, 126.8, 128.2, 128.8, 134.5, 140.5, (9C, biphenyl), 144.2 (C<sub>Q</sub> triazole), 157.6 (C1 biphenyl), 169.3, 169.9, 170.5 (3C, 3 × C=O); ESI-MS: Calcd. for C<sub>27</sub>H<sub>30</sub>N<sub>3</sub>O<sub>8</sub> [M+H]<sup>+</sup>: 524.20; Found m/z 524.22.

**1-(1,5-Anhydro-2-deoxy-D-galactitol-2-yl)-4-(4-phenylphenoxy)methyl-1,2,3-triazole (38)**: Compound **28** (51.0 mg, 97.4 μmol) was dissolved in MeOH (5 ml) and sodium metal (20.0 mg, 869 μmol) was added. The solution was stirred at r.t. under argon for 3 h upon which the deacetylated product precipitated. The solvent was removed in vacuo and the residue was washed with MeOH (2 × 5 ml) and dried in vacuo to give **38** (36.1 mg, 93%). <sup>1</sup>H-NMR (500 MHz, DMSO-d<sub>6</sub>): δ 3.42-3.58 (m, 3H, H5, H6), 3.66 (t, *J* = 10.9 Hz, 1H, H1 ax.), 3.86 (s, 1H, H4), 4.00 (m, 2H, H1 eq., H3), 4.73 (m, 1H, H2), 5.17 (s, 2H, CH<sub>2</sub>OPh), 7.14-7.63 (m, 9H, biphenyl), 8.33 (s, 1H, H5 triazole); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>): δ 58.8 (C2), 60.6 (C6), 61.3 (CH<sub>2</sub>OPh), 68.0 (C1), 68.1 (C4), 71.8 (C3), 80.1 (C5), 124.8 (C5 triazole), 115.1, 126.2, 126.8, 127.8, 128.9, 132.9, 139.8 (9C, biphenyl), 142.0 (C<sub>Q</sub> triazole); ESI-MS: Calcd. for C<sub>21</sub>H<sub>23</sub>N<sub>3</sub>O<sub>5</sub>Na [M+Na]<sup>+</sup>: 420.41; Found m/z 420.23.

**1-(3,4,6-Tri-O-acetyl-1,5-anhydro-2-deoxy-D-galactitol-2-yl)-4-(3-bromophenoxy)methyl-1,2,3-triazole (29):** Prepared according to the procedure described for **48**, using **4** (50.0 mg, 160  $\mu$ mol) and 1-bromo-3-(2-propynyloxy)-benzene [79] (**23**, 66.7 mg, 316  $\mu$ mol). The compound was purified by flash chromatography (hexane/EtOAc 7:3  $\rightarrow$  3:2) to give **29** (70.1 mg, 85%,  $R_f$  0.29 D).  $^1\text{H-NMR}$  (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  1.83, 2.08, 2.19 (s, 9H, 3  $\times$   $\text{CH}_3$ ), 4.02-4.08 (m, 2H, H1 ax., H5), 4.16 (d,  $J$  = 6.4 Hz, 1H, H6), 4.33 (dd,  $J$  = 5.0, 11.6 Hz, 1H, H1 eq.), 4.94 (dt,  $J$  = 5.0, 11.0 Hz, 1H, H2), 5.19 (s, 2H,  $\text{CH}_2\text{OPh}$ ), 5.51-5.56 (m, 2H, H3, H4), 6.89-6.91, 7.10-7.16 (m, 5H,  $\text{C}_6\text{H}_5$ ), 7.65 (s, 1H, H triazole);  $^{13}\text{C-NMR}$  (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  20.2, 20.5, 20.6 (3C, 3  $\times$   $\text{CH}_3$ ), 56.0 (C2), 61.8 (C6), 62.0 ( $\text{CH}_2\text{OPh}$ ), 66.9 (C4), 68.6 (C1), 71.4 (C3), 75.2 (C5), 113.4 (2C, 2  $\times$  C-ortho), 118.5 (C5 triazole), 122.8 (C-Br), 124.4 (C-para), 130.6 (2C, 2  $\times$  C-meta), 143.6 ( $\text{C}_Q$  triazole), 158.7 (C-ipso), 169.2, 169.9, 170.4 (3  $\times$  C=O); ESI-MS: Calcd. for  $\text{C}_{21}\text{H}_{25}\text{BrN}_3\text{O}_8$   $[\text{M}+\text{H}]^+$ : 526.08; Found  $m/z$  526.18.

**1-(1,5-Anhydro-2-deoxy-D-galactitol-2-yl)-4-(4-bromophenoxy)methyl-1,2,3-triazole (39):** Compound **29** (70.0 mg, 133  $\mu$ mol) was deacetylated according to the procedure described for **49**. The final product was purified by LCMS to give **39** (18.0 mg, 34%).  $^1\text{H-NMR}$  (500 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  3.62 (m, 1H, H5), 3.72 (dd,  $J$  = 4.9, 11.4 Hz, 1H, H6a), 3.77-3.85 (m, 2H, H1 ax., H6b), 3.98 (d,  $J$  = 2.9 Hz, 1H, H4), 4.13-4.18 (m, 2H, H1 eq., H3), 4.81 (dt,  $J$  = 5.0, 10.8 Hz, 1H, H2), 5.16 (s, 2H,  $\text{CH}_2\text{OPh}$ ), 6.98-7.21 (m, 4H,  $\text{C}_6\text{H}_4$ ), 8.14 (s, 1H, H5 triazole);  $^{13}\text{C-NMR}$  (125 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  60.8 (C2), 62.6 ( $\text{CH}_2\text{OPh}$ ), 62.9 (C6), 69.8 (C1), 70.4 (C4), 73.5 (C3), 81.5 (C5), 119.3 (C5 triazole), 123.8 (C-Br), 114.8, 125.3, 126.0, 132.0 (4C,  $\text{C}_6\text{H}_4$ ), 144.0 ( $\text{C}_Q$  triazole), 160.7 (C-ipso); ESI-MS: Calcd. for  $\text{C}_{15}\text{H}_{19}\text{BrN}_3\text{O}_5$   $[\text{M}+\text{H}]^+$ : 400.05; Found  $m/z$  400.05.

**1-(3,4,6-Tri-O-acetyl-1,5-anhydro-2-deoxy-D-galactitol-2-yl)-4-(4-bromophenoxy)methyl-1,2,3-triazole (37):** Prepared according to the procedure described for **48**, using **4** (200 mg, 634  $\mu$ mol) and 1-bromo-4-(2-propynyloxy)benzene [79]



(**27**, 267 mg, 1.07 mmol). The compound was purified by flash chromatography (hexane/EtOAc 1:1) to give **37** (279 mg, 84%,  $R_f$  0.24 D).  $^1\text{H-NMR}$  (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  1.74, 2.00, 2.11 (s, 9H, 3  $\times$   $\text{CH}_3$ ), 3.93-4.00 (m, 2H, H1 ax., H5), 4.07 (d,  $J = 6.4$  Hz, 1H, H6), 4.24 (dd,  $J = 5.0, 11.6$  Hz, 1H, H1 eq.), 4.86 (dt,  $J = 5.0, 11.0$  Hz, 1H, H2), 5.10 (s, 2H,  $\text{CH}_2\text{OPh}$ ), 5.43-5.47 (m, 2H, H3, H4), 6.77 (AA' of AA'BB',  $J = 8.9$  Hz, 2H, H-ortho), 7.30 (BB' of AA'BB',  $J = 8.9$  Hz, 2H, H-meta), 7.57 (s, 1H, H triazole);  $^{13}\text{C-NMR}$  (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  20.2, 20.5, 20.6 (3  $\times$   $\text{CH}_3$ ) 56.0 (C2), 61.8, 61.9 (C6,  $\text{CH}_2\text{OPh}$ ), 66.9 (C4), 68.6 (C1), 71.4 (C3), 75.2 (C5), 113.5 (C-Br), 116.5 (2C, 2  $\times$  C-ortho), 122.8 (C5 triazole), 132.3 (2C, 2  $\times$  C-meta), 143.7 (C<sub>Q</sub> triazole), 157.0 (C-ipso), 169.2, 169.9, 170.4 (3  $\times$  C=O); ESI-MS: Calcd. for  $\text{C}_{21}\text{H}_{25}\text{BrN}_3\text{O}_8$   $[\text{M}+\text{H}]^+$ : 526.08; Found  $m/z$  526.16.

**1-(3,4,6-Tri-O-acetyl-1,5-anhydro-2-deoxy-D-galactitol-2-yl)-4-(4-propylphenoxy)methyl-1,2,3-triazole (30)**: Prepared according to the procedure described for **48**, using **4** (50.0 mg, 160  $\mu\text{mol}$ ) and **16** (55.0 mg, 316  $\mu\text{mol}$ ). The compound was purified by flash chromatography (hexane/EtOAc 7:3  $\rightarrow$  3:2), yielding **30** (74.2 mg, 96%,  $R_f$  0.33 D).  $^1\text{H-NMR}$  (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  0.92 (t,  $J = 7.4$  Hz, 3H,  $\text{CH}_2\text{CH}_2\text{CH}_3$ ), 1.59 (sextet,  $J = 7.6$  Hz, 2H,  $\text{CH}_2\text{CH}_2\text{CH}_3$ ), 1.81, 2.07, 2.18 (s, 9H, 3  $\times$   $\text{CH}_3$ ), 2.52 (t,  $J = 7.6$  Hz, 2H,  $\text{CH}_2\text{CH}_2\text{CH}_3$ ), 4.02-4.08 (m, 2H, H1 ax., H5), 4.15 (d,  $J = 6.4$  Hz, 1H, H6), 4.32 (dd,  $J = 5.0, 11.6$  Hz, 1H, H1 eq.), 4.94 (dt,  $J = 5.1, 10.9$  Hz, 1H, H2), 5.18 (s, 2H,  $\text{CH}_2\text{OPh}$ ), 5.51-5.55 (m, 2H, H3, H4), 6.87 (AA' of AA'BB',  $J = 8.6$  Hz, 2H, H-ortho), 7.08 (BB' of AA'BB',  $J = 8.6$  Hz, 2H, H-meta), 7.66 (s, 1H, H triazole);  $^{13}\text{C-NMR}$  (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  13.6 ( $\text{CH}_2\text{CH}_2\text{CH}_3$ ), 20.2, 20.5, 20.6 (3  $\times$   $\text{CH}_3$ ), 24.6 ( $\text{CH}_2\text{CH}_2\text{CH}_3$ ), 37.0 ( $\text{CH}_2\text{CH}_2\text{CH}_3$ ), 55.9 (C2), 61.8 (C6), 61.9 ( $\text{CH}_2\text{OPh}$ ), 67.0 (C4), 68.6 (C1), 71.4 (C3), 75.1 (C5), 114.4 (2C, 2  $\times$  C-ortho), 122.7 (C5 triazole), 129.3 (2C, 2  $\times$  C-meta), 135.5 (C-para), 144.4 (C<sub>Q</sub> triazole), 156.0 (C-ipso), 169.2, 169.9, 170.4 (3  $\times$  C=O); ESI-MS: Calcd. for  $\text{C}_{24}\text{H}_{32}\text{N}_3\text{O}_8$   $[\text{M}+\text{H}]^+$ : 490.22; Found  $m/z$  490.24.

**1-(1,5-Anhydro-2-deoxy-D-galactitol-2-yl)-4-(4-propylphenoxy)methyl-1,2,3-triazole (40)**: Compound **30** (42.5 mg, 86.8  $\mu\text{mol}$ ) was deacetylated according to

the procedure described for **49**, and purified by LCMS to yield **40** (13.0 mg, 41%). <sup>1</sup>H-NMR (500 MHz, CD<sub>3</sub>OD): δ 0.92 (t, *J* = 7.4 Hz, 3H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>), 1.61 (sextet, *J* = 7.4 Hz, 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>), 2.54 (t, *J* = 7.5 Hz, 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>), 3.74 (m, 1H, H5), 3.78-3.90 (m, 3H, H1 ax., H6), 4.09 (d, *J* = 3.0 Hz, 1H, H4), 4.20-4.27 (m, 2H, H1 eq., H3), 4.94 (dt, *J* = 5.1, 10.9 Hz, 1H, H2), 4.87 (m, 1H, H2), 5.20 (s, 2H, CH<sub>2</sub>OPh), 6.97 (AA' of AA'BB', *J* = 8.6 Hz, 2H, H-ortho), 7.15 (BB' of AA'BB', *J* = 8.5 Hz, 2H, H-meta), 8.20 (s, 1H, H triazole); <sup>13</sup>C-NMR (125 MHz, CD<sub>3</sub>OD): δ 13.9 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>), 25.6 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>), 37.8 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>), 60.4 (C2), 62.3 (CH<sub>2</sub>OPh), 62.5 (C6), 69.4 (C1), 69.9 (C4), 73.1 (C3), 81.1 (C5), 115.7 (2C, 2 × C-ortho), 126.1 (C5 triazole), 130.4 (2C, 2 × C-meta), 136.9 (C-para), 144.4 (C<sub>Q</sub> triazole), 157.2 (C-*ipso*); ESI-MS: Calcd. for C<sub>18</sub>H<sub>26</sub>N<sub>3</sub>O<sub>5</sub> [M+H]<sup>+</sup>: 364.19; Found m/z 364.18.

**1-(3,4,6-Tri-O-acetyl-1,5-anhydro-2-deoxy-D-galactitol-2-yl)-4-(4-nitrophenyloxy)methyl-1,2,3-triazole (31)**: Prepared according to the procedure described for **48**, using **4** (50.0 mg, 158 μmol) and 1-nitro-4-(2-propynyloxy)-benzene [**77**] (**24**, 49.7 mg, 316 μmol). The compound was purified by flash chromatography (petrol ether/EtOAc 1:1) to give **31** (74.0 mg, 95%, R<sub>f</sub> 0.22 D). <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ 1.80, 2.03, 2.15 (s, 9H, 3 × CH<sub>3</sub>), 3.99-4.04 (m, 2H, H1 ax., H5), 4.12 (d, *J* = 6.4 Hz, 2H, H6), 4.30 (dd, *J* = 5.0, 11.6 Hz, 1H, H1 eq.), 4.93 (m, 1H, H2), 5.26 (s, 2H, CH<sub>2</sub>OPh), 5.49-5.52 (m, 2H, H3, H4), 7.02 (AA' of AA'BB', *J* = 9.3 Hz, 2H, H-ortho), 7.70 (s, 1H, H5 triazole), 8.15 (BB' of AA'BB', *J* = 9.3 Hz, 2H, H-meta); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>): δ 20.3, 20.6, 20.7 (3 × CH<sub>3</sub>), 56.2 (C2), 61.9, 62.3 (C6, CH<sub>2</sub>OPh), 67.0 (C4), 68.6 (C1), 71.4 (C3), 75.3 (C5), 114.8 (2C, 2 × C-ortho), 123.2 (C5 triazole), 125.9 (2C, 2 × C-meta), 141.9 (C-para), 142.8 (C<sub>Q</sub> triazole), 163.0 (C-*ipso*), 169.3, 170.0, 170.5 (3 × C=O); ESI-MS: Calcd. for C<sub>21</sub>H<sub>24</sub>N<sub>4</sub>O<sub>10</sub> [M+H]<sup>+</sup>: 493.16; Found m/z 493.20.

**1-(1,5-Anhydro-2-deoxy-D-galactitol-2-yl)-4-(4-nitrophenoxy)methyl-1,2,3-triazole (41)**: Compound **31** (74.0 mg, 150 μmol) was deacetylated according to the procedure described for **49**. The final product was purified by LCMS to give

**41** (18.0 mg, 33%). <sup>1</sup>H-NMR (500 MHz, CD<sub>3</sub>OD): δ 3.76-3.87 (m, 2H, H5, H6), 3.90 (t, *J* = 11.2 Hz, 1H, H1 ax.), 4.10 (d, *J* = 3.1 Hz, 2H, H4), 4.24 (dd, *J* = 5.1, 11.2 Hz, 1H, H1 eq.), 4.28 (dd, *J* = 3.2, 10.6 Hz, 1H, H3), 4.87 (m, 1H, H2), 5.38 (s, 2H, CH<sub>2</sub>OPh), 7.23 (AA' of AA'BB', *J* = 9.3 Hz, 2H, H-ortho), 8.29 (m, 3H, H5 triazole, H-meta); <sup>13</sup>C-NMR (125 MHz, CD<sub>3</sub>OD): δ 60.3 (C2), 62.5, 62.6 (C6, CH<sub>2</sub>OPh), 69.3 (C1), 69.8 (C4), 72.9 (C3), 81.0 (C5), 116.1 (2C, 2 × C-ortho), 126.5 (C5 triazole), 127.0 (2C, 2 × C-meta), 142.6 (C-para), 143.3 (C<sub>Q</sub> triazole), 164.5 (C-ipso); ESI-MS: Calcd. for C<sub>15</sub>H<sub>19</sub>N<sub>4</sub>O<sub>7</sub> [M+H]<sup>+</sup>: 367.13; Found *m/z* 367.15.

**1-(3,4,6-Tri-O-acetyl-1,5-anhydro-2-deoxy-D-galactitol-2-yl)-4-(4-cyanophenyl-1,2,3-triazole) methyl-1,2,3-triazole (32)**: Prepared according to the procedure described for **48**, using **4** (50.0 mg, 158 μmol) and 1-cyano-4-(2-propynyloxy)-benzene [80] (**25**, 49.7 mg, 316 μmol). The compound was purified by flash chromatography (petrol ether/EtOAc 1:1 → 0:1) to give **32** (73.0 mg, 98%, R<sub>f</sub> 0.33 C). <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ 1.79, 2.04, 2.16 (s, 9H, 3 × CH<sub>3</sub>), 3.99-4.04 (m, 2H, H1 ax., H5), 4.12 (d, *J* = 6.4 Hz, 2H, H6), 4.30 (dd, *J* = 5.0, 11.6 Hz, 1H, H1 eq.), 4.92 (dt, *J* = 4.9, 10.7 Hz, 1H, H2), 5.22 (s, 2H, CH<sub>2</sub>OPh), 5.48-5.52 (m, 2H, H3, H4), 7.01 (m, 2H, H-ortho), 7.56 (m, 2H, H-meta), 7.67 (s, 1H, H5 triazole); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>): δ 20.2, 20.5, 20.6 (3 × CH<sub>3</sub>), 56.1 (C2), 61.7, 61.9 (C6, CH<sub>2</sub>OPh), 66.9 (C4), 68.6 (C1), 71.3 (C3), 75.2 (C5), 115.4 (2C, 2 × C-ortho), 118.8 (C-para), 123.0 (C5 triazole), 134.0 (2C, 2 × C-meta), 142.9 (C<sub>Q</sub> triazole), 161.2 (C-ipso), 169.2, 169.8, 170.4 (3 × C=O); ESI-MS: Calcd. for C<sub>22</sub>H<sub>24</sub>N<sub>4</sub>O<sub>8</sub> [M+H]<sup>+</sup>: 473.17; Found *m/z* 473.06.

**1-(1,5-Anhydro-2-deoxy-D-galactitol-2-yl)-4-(4-cyanophenyl-1,2,3-triazole) methyl-1,2,3-triazole (42)**: Compound **32** (73.0 mg, 155 μmol) was dissolved in H<sub>2</sub>O/MeOH/Et<sub>3</sub>N (5:5:1, 5.5 ml) and stirred at r.t. overnight. The solvent was removed in vacuo and the residue was purified by flash chromatography (DCM/MeOH 9:1) to give **42** (44.0 mg, 81%, R<sub>f</sub> 0.62 G). <sup>1</sup>H-NMR (500 MHz, CD<sub>3</sub>OD): δ 3.62 (m, 1H, H5), 3.73 (dd, *J* = 4.9, 11.4 Hz, 1H, H6a), 3.80 (dd, *J* =

7.1, 11.4 Hz, 1H, H6b), 3.84 (t,  $J = 11.1$  Hz, 1H, H1 ax.), 3.99 (d,  $J = 3.0$  Hz, 1H, H4), 4.14-4.18 (m, 2H, H1 eq., H3), 4.82 (dt,  $J = 5.0, 10.9$  Hz, 1H, H2), 5.36 (s, 2H, CH<sub>2</sub>OPh), 7.17 (AA' of AA'BB',  $J = 8.9$  Hz, 2H, H-ortho), 7.67 (AA' of AA'BB',  $J = 9.2$  Hz, 2H, H-meta), 8.19 (s, 1H, H5 triazole); <sup>13</sup>C-NMR (125 MHz, CD<sub>3</sub>OD):  $\delta$  60.7 (C2), 62.6 (CH<sub>2</sub>OPh), 62.9 (C6), 69.7 (C1), 70.3 (C4), 73.5 (C3), 80.5 (C5), 105.3 (CN), 116.8 (2C, 2  $\times$  C-ortho), 120.0 (C-para), 126.2 (C5 triazole), 135.2 (2C, 2  $\times$  C-meta), 143.5 (C<sub>Q</sub> triazole), 163.2 (C-ipso); ESI-MS: Calcd. for C<sub>16</sub>H<sub>19</sub>N<sub>4</sub>O<sub>5</sub> [M+H]<sup>+</sup>: 347.14; Found m/z 347.12.

**1-(3,4,6-Tri-O-acetyl-1,5-anhydro-2-deoxy-D-galactitol-2-yl)-4-(quinol-9-yloxy)methyl-1,2,3-triazole (33)**: Prepared according to the procedure described for **48**, using **4** (50.0 mg, 160  $\mu$ mol) and **19** (55.0 mg, 316  $\mu$ mol). The compound was purified by flash chromatography (DCM/MeOH/Et<sub>3</sub>N stepwise gradient 95:5:0.1) to give **33** (63 mg, 80%, R<sub>f</sub> 0.43 H). <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  1.70, 2.03, 2.13 (s, 9H, 3  $\times$  CH<sub>3</sub>), 3.95-4.09 (m, 2H, H1 ax., H5), 4.10 (d,  $J = 6.4$  Hz, 1H, H6), 4.28 (dd,  $J = 5.0, 11.6$  Hz, 1H, H1 eq.), 4.85 (dt,  $J = 5.0, 11.0$  Hz, 1H, H2), 5.45-5.55 (m, 4H, H3, H4, CH<sub>2</sub>OPh), 7.21 (m, 1H, H3 quinoline), 7.37-7.42 (m, 3H, H6, H7, H8 quinoline), 7.76 (s, 1H, H5 triazole), 8.09-8.12 (m, 1H, H4 quinoline), 8.89-8.90 (m, 1H, H2 quinoline); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  20.2, 20.5, 20.6 (3  $\times$  CH<sub>3</sub>), 55.9 (C2), 61.8 (C6), 62.7 (CH<sub>2</sub>OPh), 66.9 (C4), 68.6 (C1), 71.4 (C3), 75.1 (C5), 109.7 (C8 quinoline), 120.2, 121.6, 126.6 (C3, C6, C7 quinoline), 123.4 (C5 triazole), 129.4 (C5 quinoline), 135.9 (C4 quinoline), 140.2 (C10 quinoline), 144.0 (C<sub>Q</sub> triazole), 149.3 (C2 quinoline), 153.6 (C9 quinoline), 169.2, 169.9, 170.4 (3  $\times$  C=O); ESI-MS: Calcd. for C<sub>24</sub>H<sub>27</sub>N<sub>4</sub>O<sub>8</sub> [M+H]<sup>+</sup>: 499.18; Found m/z 499.23.

**1-(1,5-Anhydro-2-deoxy-D-galactitol-2-yl)-4-(quinol-9-yloxy)methyl-1,2,3-triazole (43)**: Compound **33** (63.0 mg, 126  $\mu$ mol) was deacetylated according to the procedure described for **49**, and purified by LCMS to yield **43** (40.0 mg, 85%). <sup>1</sup>H-NMR (500 MHz, DMSO-d<sub>6</sub>):  $\delta$  3.50 (m, 1H, H5), 3.55 (m, 2H, H6), 3.70 (t,  $J = 11.0$  Hz, 1H, H1 ax.), 3.84 (m, 1H, H4), 4.01-4.08 (m, 2H, H1 eq., H3),

4.69 (t,  $J = 5.5$  Hz, 1H, 6-OH), 4.76 (dt,  $J = 5.0, 10.8$  Hz, 1H, H2), 4.88 (d,  $J = 4.9$  Hz, 1H, 4-OH), 5.14 (d,  $J = 6.9$  Hz, 1H, 3-OH), 5.34 (s, 2H, CH<sub>2</sub>OPh), 7.44 (m, 1H, H3 quinoline), 7.52-7.56 (m, 3H, H6, H7, H8 quinoline), 8.33 (m, 1H, H4 quinoline), 8.39 (s, 1H, H5 triazole), 8.83 (m, 1H, H2 quinoline); <sup>13</sup>C-NMR (125 MHz, DMSO-d<sub>6</sub>):  $\delta$  55.8 (C2), 60.7 (C6), 61.8 (CH<sub>2</sub>OPh), 68.1 (C1), 68.2 (C4), 71.7 (C3), 80.0 (C5), 109.9 (C8 quinoline), 120.0, 121.9, 126.8 (C3, C6, C7 quinoline), 125.0 (C5 triazole), 129.1 (C5 quinoline), 135.8 (C4 quinoline), 139.7 (C10 quinoline), 141.9 (C<sub>Q</sub> triazole), 149.0 (C2 quinoline), 154.0 (C9 quinoline); ESI-MS: Calcd. for C<sub>18</sub>H<sub>21</sub>N<sub>4</sub>O<sub>5</sub> [M+H]<sup>+</sup>: 373.15; Found m/z 373.20.

**1-(3,4,6-Tri-O-acetyl-1,5-anhydro-2-deoxy-D-galactitol-2-yl)-4-(2,3,4,5,6-pentafluorophenoxy)methyl-1,2,3-triazole (34)**: Prepared according to the procedure described for **48**, using **4** (50.0 mg, 158  $\mu$ mol) and 1,2,3,4,5-pentafluoro-6-(2-propynyloxy)-benzene [81] (**26**, 70.2 mg, 316  $\mu$ mol). The compound was purified by flash chromatography (petrol ether/EtOAc 4:1  $\rightarrow$  1:1) to give **34** (62.0 mg, 73%). <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  1.85, 2.05, 2.18 (s, 9H, 3  $\times$  CH<sub>3</sub>), 3.97-4.02 (m, 2H, H1 ax., H5), 4.13 (d,  $J = 6.4$  Hz, 2H, H6), 4.29 (dd,  $J = 5.0, J = 11.6$  Hz, 1H, H1 eq.), 4.95 (dt,  $J = 5.0, J = J = 11.0$  Hz, 1H, H2), 5.27 (s, 2H, CH<sub>2</sub>OPh), 5.49-5.53 (m, 2H, H3, H4), 7.74 (m, 1H, H5 triazole), <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  20.1, 20.5, 20.6 (3  $\times$  CH<sub>3</sub>), 56.1 (C2), 61.8 (C6), 66.9 (C4), 67.6 (CH<sub>2</sub>OPh), 68.6 (C1), 71.3 (C3), 75.2 (C5), 123.6 (C5 triazole), 136.9, 139.0, 141.1, 142.9 (6C, C<sub>6</sub>F<sub>5</sub>), 142.5 (C<sub>Q</sub> triazole), 169.3, 169.9, 170.4 (3  $\times$  C=O); ESI-MS: Calcd. for C<sub>21</sub>H<sub>21</sub>F<sub>5</sub>N<sub>3</sub>O<sub>8</sub> [M+H]<sup>+</sup>: 538.12; Found m/z 538.09.

**1-(1,5-Anhydro-2-deoxy-D-galactitol-2-yl)-4-(2,3,4,5,6-pentafluorophenoxy)-methyl-1,2,3-triazole (44)**: Compound **34** (60.0 mg, 111  $\mu$ mol) was deacetylated according to the procedure described for **49**. The final product was purified by LCMS to give **44** (31.0 mg, 68%). <sup>1</sup>H-NMR (500 MHz, CD<sub>3</sub>OD):  $\delta$  3.59 (m, 1H, H5), 3.69 (dd,  $J = 4.9, 11.4$  Hz, 1H, H6a), 3.75-3.90 (m, 2H, H1 ax., H6b), 3.97 (m, 1H, H4), 4.10-4.14 (m, 2H, H1 eq., H3), 4.78 (dt,  $J = 5.0, 10.8$  Hz, 1H, H2), 5.24 (s, 2H, CH<sub>2</sub>OPh), 8.17 (s, 1H, H5 triazole); <sup>13</sup>C-NMR (125 MHz,

CD<sub>3</sub>OD):  $\delta$  60.7 (C2), 62.9 (C6), 68.4 (CH<sub>2</sub>OPh), 66.5 (C1), 70.4 (C4), 73.5 (C3), 81.5 (C5), 126.9 (C5 triazole), 142.9 (C<sub>Q</sub> triazole); ESI-MS: Calcd. for C<sub>15</sub>H<sub>15</sub>F<sub>5</sub>N<sub>3</sub>O<sub>5</sub> [M+H]<sup>+</sup>: 412.09; Found m/z 412.10.

**1-(3,4,6-Tri-O-acetyl-1,5-anhydro-2-deoxy-D-galactitol-2-yl)-4-(4-(N-morpholino)phenoxy)methyl-1,2,3-triazole (51)**: Compound **37** (50.0 mg, 95.0  $\mu$ mol) and Cs<sub>2</sub>CO<sub>3</sub> (46.2 mg, 189  $\mu$ mol) were azeotropically dried with PhMe, and Pd<sub>2</sub>dba<sub>3</sub> (2.00 mg, 1.91  $\mu$ mol), 2-dicyclohexylphosphino-2',6'-dimethoxybiphenyl [82] (**50**, 5.00 mg, 7.59  $\mu$ mol), morpholine (16.5 mg, 189  $\mu$ mol) and dry PhMe (2 ml) were added. The resultant mixture was stirred at 80 °C under argon for 24 h. The solvent was then removed in vacuo, and the residue purified by flash chromatography (petrol ether/EtOAc 1:1  $\rightarrow$  1:3) to give **51** (20.0 mg, 40%, R<sub>f</sub> 0.37 C). <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  1.82, 2.07, 2.18 (s, 9H, 3  $\times$  CH<sub>3</sub>), 3.07 (bs, 4H, (CH<sub>2</sub>)N), 3.86 (bs, 4H, (CH<sub>2</sub>)O), 3.99-4.05 (m, 2H, H1 ax., H5), 4.15 (d, *J* = 6.5 Hz, 2H, H6), 4.32 (dd, *J* = 5.0, 11.6 Hz, 1H, H1 eq.), 4.91 (dt, *J* = 5.0, 11.0 Hz, 1H, H2), 5.16 (s, 2H, CH<sub>2</sub>OPh), 5.50 (dd, *J* = 3.2, 10.9 Hz, 1H, H3), 5.14 (d, *J* = 2.6 Hz, 1H, H4), 6.90 (bs, 4H, C<sub>6</sub>H<sub>5</sub>), 7.61 (s, 1H, H5 triazole); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  20.4, 20.7, 20.8 (3  $\times$  CH<sub>3</sub>), 50.7 (2C, (CH<sub>2</sub>)<sub>2</sub>N), 56.0 (C2), 61.9 (C6), 62.4 (CH<sub>2</sub>OPh), 67.0 (3C, (CH<sub>2</sub>)<sub>2</sub>O, C4), 68.8 (C1), 71.5 (C3), 75.3 (C5), 117.7 (6C, C<sub>6</sub>H<sub>4</sub>), 123.4 (C5 triazole), 144.5 (C<sub>Q</sub> triazole), 169.4, 170.0, 170.5 (3  $\times$  C=O); ESI-MS: Calcd. for C<sub>25</sub>H<sub>33</sub>N<sub>4</sub>O<sub>9</sub> [M+H]<sup>+</sup>: 533.55; Found m/z 533.32.

**1-(1,5-Anhydro-2-deoxy-D-galactitol-2-yl)-4-(4-(N-morpholino)phenoxy)methyl-1,2,3-triazole (52)**: Compound **51** (14.0 mg, 26.3  $\mu$ mol) was deacetylated according to the procedure described for **49**. The final product was purified on an RP-C18 column (H<sub>2</sub>O/MeOH stepwise gradient 1:0  $\rightarrow$  4:1) to give **50** (6.00 mg, 57%, R<sub>f</sub> 0.75 G). <sup>1</sup>H-NMR (500 MHz, DMSO-d<sub>6</sub>):  $\delta$  3.07 (t, *J* = 4.7 Hz, 4H, (CH<sub>2</sub>)<sub>2</sub>N), 3.49 (m, 1H, H5), 3.54 (m, 2H, H6), 3.65 (t, *J* = 11.0 Hz, 1H, H1 ax.), 3.73 (t, *J* = 4.7 Hz, 4H, (CH<sub>2</sub>)<sub>2</sub>O), 3.83 (d, *J* = 2.7 Hz, 1H, H4), 3.98 (dd, *J* = 5.0, 10.8 Hz, 1H, H1 eq.), 4.02 (dd, *J* = 3.0, 10.5 Hz, 1H, H3), 4.71 (m, 1H, H2), 5.04 (s, 2H, CH<sub>2</sub>OPh), 6.92 (m, 4H, C<sub>6</sub>H<sub>5</sub>), 8.27 (s, 1H, H5 triazole); <sup>13</sup>C-NMR (125

MHz, DMSO-d<sub>6</sub>):  $\delta$  49.6 (2C, (CH<sub>2</sub>)<sub>2</sub>N), 58.6 (C2), 60.6 (C6), 61.4 (CH<sub>2</sub>OPh), 66.2 (2C, (CH<sub>2</sub>)<sub>2</sub>O), 68.1 (2C, C1, C4), 71.6 (C3), 79.9 (C5), 115.1 (2C, 2  $\times$  C-*ortho*), 116.9 (2C, 2  $\times$  C-*meta*), 124.4 (C5 triazole), 142.3 (C-*para*), 145.6 (C<sub>Q</sub> triazole), 151.8 (C-*ipso*); ESI-MS: Calcd. for C<sub>19</sub>H<sub>27</sub>N<sub>4</sub>O<sub>6</sub> [M+H]<sup>+</sup>: 407.19; Found m/z 407.18.

**1-(3,4,6-Tri-O-acetyl-1,5-anhydro-2-deoxy-D-galactitol-2-yl)-4-(1-N-methyl-4-N-piperazino)phenoxyethyl-1,2,3-triazole (54)**: Compound **37** (75.0 mg, 142  $\mu$ mol) and Cs<sub>2</sub>CO<sub>3</sub> (64.8 mg, 199  $\mu$ mol) were azeotropically dried with PhMe, and Pd<sub>2</sub>dba<sub>3</sub> (3.00 mg, 2.89  $\mu$ mol), 2-dicyclohexylphosphino-2',4',6'-triisopropylbiphenyl [83] (**53**, 1.35 mg, 2.84  $\mu$ mol), N-methylpiperazine (17.0  $\mu$ l mg, 157  $\mu$ mol) and dry PhMe (3 ml) were added. The resultant mixture was stirred at 80 °C under argon for 24 h. The solvent was then removed in vacuo, and the residue purified by flash chromatography (EtOAc/MeOH 9:1  $\rightarrow$  4:1) to give **54**, which was still impure according to NMR. The compound was used without further purification in the next step.

**1-(1,5-Anhydro-2-deoxy-D-galactitol-2-yl)-4-(1-N-methyl-4-N-piperazino)-phenoxyethyl-1,2,3-triazole (55)**: Compound **54** (35.1 mg, 64.3  $\mu$ mol) was deacetylated according to the procedure described for **49**. The final product was purified by flash chromatography (DCM/MeOH 9:1  $\rightarrow$  7:3) to give **55** (12 mg, 44%, R<sub>f</sub> 0.19 I). <sup>1</sup>H-NMR (500 MHz, CD<sub>3</sub>OD):  $\delta$  2.63 (s, 3H, NCH<sub>3</sub>), 3.02 (m, 4H, (CH<sub>2</sub>)<sub>2</sub>NCH<sub>3</sub>), 3.18 (m, 4H, (CH<sub>2</sub>)<sub>2</sub>NPh), 3.59 (m, 1H, H5), 3.69 (dd, *J* = 4.9, 11.4 Hz, 1H, H6a), 3.74-3.82 (m, 2H, H1 ax., H6b), 3.95 (d, *J* = 3.0 Hz, 1H, H4), 4.10-4.13 (m, 2H, H1 eq., H3), 4.02 (dd, *J* = 3.0, 10.5 Hz, 1H, H3), 4.76 (m, 1H, H2), 5.07 (s, 2H, CH<sub>2</sub>OPh), 6.92 (m, 4H, C<sub>6</sub>H<sub>5</sub>), 8.09 (s, 1H, H5 triazole); <sup>13</sup>C-NMR (125 MHz, CD<sub>3</sub>OD):  $\delta$  44.7 (NCH<sub>3</sub>), 49.5 (2C, (CH<sub>2</sub>)<sub>2</sub>NCH<sub>3</sub>), 55.4 (2C, (CH<sub>2</sub>)<sub>2</sub>NPh), 60.7 (C2), 62.8 (2C, C6, CH<sub>2</sub>OPh), 69.7 (C1), 70.4 (C4), 73.4 (C3)

81.4 (C5), 116.7 (2C, 2 × C-ortho), 119.9 (2C, 2 × C-meta), 125.8 (C5 triazole), 144.6 (C-para), 146.4 (C<sub>Q</sub> triazole), 154.7 (C-*ipso*); ESI-MS: Calcd. for C<sub>20</sub>H<sub>30</sub>N<sub>5</sub>O<sub>5</sub> [M+H]<sup>+</sup>: 420.22; Found m/z 420.27.

**1-(3,4,6-Tri-O-acetyl-1,5-anhydro-2-deoxy-D-galactitol-2-yl)-4-(4-ethoxycarbonylphenoxy)methyl-1,2,3-triazole (35)**: Prepared according to the procedure described for **48**, using **4** (50.0 mg, 160 μmol) and **17** (64.5 mg, 316 μmol). The compound was purified by flash chromatography (hexane/EtOAc 3:2 → 2:3) to give **35** (62.0 mg, 82%, R<sub>f</sub> 0.15 D). <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ 1.35 (t, *J* = 7.1 Hz, 3H, CH<sub>2</sub>CH<sub>3</sub>), 1.79, 2.05, 2.16 (s, 9H, 3 × OAc), 3.99-4.10 (m, 2H, H1 ax., H5), 4.13 (d, *J* = 6.4 Hz, 1H, H6), 4.29-4.34 (m, 3H, H1 eq., CH<sub>2</sub>CH<sub>3</sub>), 4.92 (dt, *J* = 4.9, 10.9 Hz, 1H, H2), 5.23 (s, 2H, CH<sub>2</sub>OPh), 5.49-5.53 (m, 2H, H3, H4), 6.95 (AA' of AA'BB', *J* = 8.9 Hz, 2H, H-ortho), 7.66 (s, 1H, H5 triazole), 7.96 (BB' of AA'BB', *J* = 8.9 Hz, 2H, H-meta); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>): δ 14.3 (CH<sub>2</sub>CH<sub>3</sub>), 20.2, 20.5, 20.6 (3 × CH<sub>3</sub>), 56.0 (C2), 60.6 (CH<sub>2</sub>CH<sub>3</sub>), 61.8 (C6, CH<sub>2</sub>OPh), 66.9 (C4), 68.6 (C1), 71.4 (C3), 75.2 (C5), 114.2 (2C, 2 × C-ortho), 122.9 (C5 triazole), 131.5 (2C, 2 × C-meta), 131.6 (C-para), 143.5 (C<sub>Q</sub> triazole), 161.5 (C-*ipso*), 169.2, 169.9, 170.4 (3 × C=O); ESI-MS: Calcd. for C<sub>24</sub>H<sub>30</sub>N<sub>3</sub>O<sub>10</sub> [M+H]<sup>+</sup>: 520.19; Found m/z 520.37.

**1-(1,5-Anhydro-2-deoxy-D-galactitol-2-yl)-4-(4-carboxyphenoxy)methyl-1,2,3-triazole (45)**: Compound **35** (62.0 mg, 119 μmol) was dissolved in MeOH/H<sub>2</sub>O (4 ml, 1:1), and NaOH (170 mg, 4.72 mmol) was added. The solution was stirred at r.t. for 24 h, after which the solvent was removed in vacuo, and the residue was purified by flash chromatography (DCM/MeOH 5:1 → 2:1) to give **45** (32.0 mg, 74%, R<sub>f</sub> 0.47 J). <sup>1</sup>H-NMR (500 MHz, CD<sub>3</sub>OD): δ 3.62 (m, 1H, H5), 3.73 (m, 2H, H6), 3.81 (t, *J* = 11.3 Hz, 1H, H1 ax.), 3.99 (d, *J* = 3.0 Hz, 1H, H4), 4.11-4.17 (m, 2H, H3, H1 eq.), 4.80 (m, 1H, H2), 5.20 (s, 2H, CH<sub>2</sub>OPh), 7.02 (AA' of AA'BB', *J* = 8.8 Hz, 2H, H-ortho), 7.93 (BB' of AA'BB', *J* = 8.8 Hz, 2H, H-meta), 8.19 (s, 1H, H5 triazole); <sup>13</sup>C-NMR (125 MHz, CD<sub>3</sub>OD): δ 60.6 (C2), 62.4



(CH<sub>2</sub>OPh), 62.9 (C<sub>6</sub>), 69.7 (C<sub>4</sub>), 70.5 (C<sub>1</sub>), 73.3 (C<sub>3</sub>), 81.2 (C<sub>5</sub>), 115.1 (2C, 2 × C-*ortho*), 126.0 (C<sub>5</sub> triazole), 127.5 (C-*para*), 132.5 (2C, 2 × C-*meta*), 144.1 (C<sub>Q</sub> triazole), 162.8 (C-*ipso*), 173.7 (C=O); ESI-MS: Calcd. for C<sub>16</sub>H<sub>18</sub>N<sub>3</sub>O<sub>7</sub> [M]<sup>-</sup>: 364.11; Found m/z 364.12.

**1-(3,4,6-Tri-O-acetyl-1,5-anhydro-2-deoxy-D-galactitol-2-yl)-4-(trifluoromethylsulfonamidophenyl)methyl-1,2,3-triazole (36)**: Prepared according to the procedure described for **48**, using **4** (50.0 mg, 158 μmol) and **21** (88.2 mg, 316 μmol). The compound was purified by flash chromatography (petrol ether/EtOAc 3:1 → 1:1) to give **36** (39.0 mg, 41%, R<sub>f</sub> 0.18 D). <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ 1.81, 2.06, 2.18 (s, 9H, 3 × CH<sub>3</sub>), 4.00-4.05 (m, 2H, H1 ax., H5), 4.14 (d, *J* = 6.4 Hz, 2H, H6), 4.32 (dd, *J*<sub>1e,2</sub> = 5.0, *J*<sub>1a,1e</sub> = 11.6 Hz, 1H, H1 eq.), 4.95 (m, 1H, H2), 5.12 (s, 2H, CH<sub>2</sub>OPh), 5.50-5.53 (m, 2H, H3, H4), 6.88 (AA' of AA'BB', *J* = 8.8 Hz, 2H, H-*ortho*), 7.22 (BB' of AA'BB', *J* = 8.8 Hz, 2H, H-*meta*), 7.68 (s, 1H, H5 triazole), 8.56 (s, 1H, NH); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>): δ 20.2, 20.6, 20.7 (3 × CH<sub>3</sub>), 56.3 (C<sub>2</sub>), 61.7, 61.9 (C<sub>6</sub>, CH<sub>2</sub>OPh), 67.0 (C<sub>4</sub>), 68.6 (C<sub>1</sub>), 71.4 (C<sub>3</sub>), 75.2 (C<sub>5</sub>), 115.4 (2C, 2 × C-*ortho*), 119.84 (q, *J* = 323.1 Hz, CF<sub>3</sub>), 123.2 (C<sub>5</sub> triazole), 126.6 (2C, 2 × C-*meta*), 127.1 (C-*para*), 143.5 (C<sub>Q</sub> triazole), 157.3 (C-*ipso*), 169.5, 170.1, 170.6 (3 × C=O); ESI-MS: Calcd. for C<sub>22</sub>CuH<sub>25</sub>F<sub>3</sub>N<sub>4</sub>O<sub>10</sub> [M+Cu]<sup>+</sup>: 657.05; Found m/z 656.99.

**1-(1,5-Anhydro-2-deoxy-D-galactitol-2-yl)-4-(4-trifluoromethylsulfonamidophenyl)methyl-1,2,3-triazole (46)**: Compound **36** (39.0 mg, 61.6 μmol) was deacetylated according to the procedure described for **49**. The final product was purified by flash chromatography on silica gel (DCM/MeOH 10:1) to give **46** (21.2 mg, 73%, R<sub>f</sub> 0.63 I). <sup>1</sup>H-NMR (500 MHz, CD<sub>3</sub>OD): δ 3.58 (m, 1H, H5), 3.68 (dd, *J* = 4.8, 11.4 Hz, 1H, H6a), 3.73-3.81 (m, 2H, H1 ax., H6b), 3.95 (d, *J* = 3.0 Hz, 1H, H4), 4.10-4.13 (m, 2H, H1 eq., H3), 4.77 (dt, *J* = 5.1, 10.8 Hz, 1H, H2), 5.11 (s, 2H, CH<sub>2</sub>OPh), 6.94 (AA' of AA'BB', *J* = 9.0 Hz, 2H, H-*ortho*), 7.13 (BB' of AA'BB', *J* = 9.0 Hz, 2H, H-*meta*), 8.11 (s, 1H, H5 triazole); <sup>13</sup>C-NMR (125 MHz, CD<sub>3</sub>OD): δ 60.7 (C<sub>2</sub>), 62.6 (CH<sub>2</sub>OPh), 62.9 (C<sub>6</sub>), 69.7 (C<sub>1</sub>), 70.4 (C<sub>4</sub>), 73.4 (C<sub>3</sub>), 81.4

(C5), 116.3 (2C, 2 × C-ortho), 121.9 (q,  $J = 323.9$  Hz, CF<sub>3</sub>), 125.9 (C5 triazole), 126.8 (2C, 2 × C-meta), 131.4 (C-para), 144.3 (C<sub>Q</sub> triazole), 158.0 (C-*ipso*); ESI-MS: Calcd. for C<sub>16</sub>H<sub>19</sub>F<sub>3</sub>N<sub>4</sub>O<sub>7</sub>SNa [M+Na]<sup>+</sup>: 491.08; Found m/z 491.10.

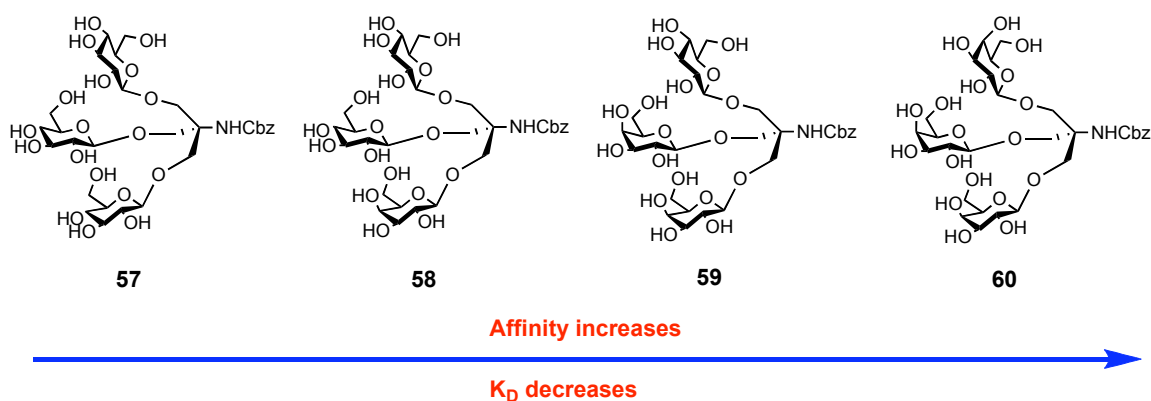
## Chapter 4: Local Concentration

### 4.1 Introduction

The most basic assumption in ligand-receptor interactions, and this includes drug-receptor interactions, is that the ligand has to collide with a small portion of the receptor (the binding site) in the right orientation and stay bound long enough to either stimulate (agonist) or block (antagonist) a biological effect mediated by that receptor. This was first formulated in the late 1800s by Paul Ehrlich, who coined the term “*Corpora non agunt nisi fixata*”, which means “a drug does not work unless physically bound to its target”. The statement was based on his proposed theory that a drug can not only chemically react with its target (*e.g.* alkylating agents) but also interact via a supramolecular or “physical” mechanism, which is now accepted as being the general mechanism by which most drugs interact with their target receptors.

Carbohydrate-protein interactions tend to be a lot weaker than protein-protein interactions, therefore to enhance the affinity of carbohydrate-protein interactions, nature often uses multivalency, *i.e.* linking several sugar units of low affinity together to generate a compound with a higher overall affinity [24]. The resulting increase in affinity is usually attributed to simultaneous interactions of more than one sugar residue with more than one binding site present on the same or different receptor subunits. However, the effect of having several sugar units on one molecule, and hence having a locally higher sugar concentration per molecule, on carbohydrate-protein interactions has not been thoroughly investigated in the case of the ASGP-R.

In this chapter, a series of trivalent compounds **57**, **58**, **59** and **60** (Figure 14) is introduced that were synthesized to investigate the effect of local sugar concentration on the affinity of the ligands towards the ASGP-R H1-CRD using BIACORE. The SPR signal depends on the mass change on the surface of the BIACORE chip. Since Glc and Gal are epimers, compounds **57**, **58**, **59** and **60** have identical molecular masses and scaffolds, which is important when comparing SPR sensograms, but have different numbers of Gal residues (0-3), and hence were expected to have a different affinity towards the H1-CRD.



**Figure 14.** Trivalent compounds **57**, **58**, **59** and **60** synthesized for investigating the local concentration effect on ligand binding to the ASGP-R H1-CRD, showing the expected increase in affinity.

On the molecular level, the interaction on the BIACORE chip between one immobilized H1-CRD protein and one of the trivalent compounds can be assumed to follow a one-to-one binding model. That is: one trivalent ligand binds only to one H1-CRD subunit via one sugar residue, without bridging two or more subunits, and hence the equilibrium dissociation constant  $K_D$  of this binding interaction can be deduced.

Since each ligand possesses three sugar residues, and therefore three possible binding candidates, the probability of a correct collision between a sugar residue and the ASGP-R H1-CRD sugar binding site is three times higher than that for a monovalent ligand, *i.e.* with just one sugar residue. This can also be expressed in

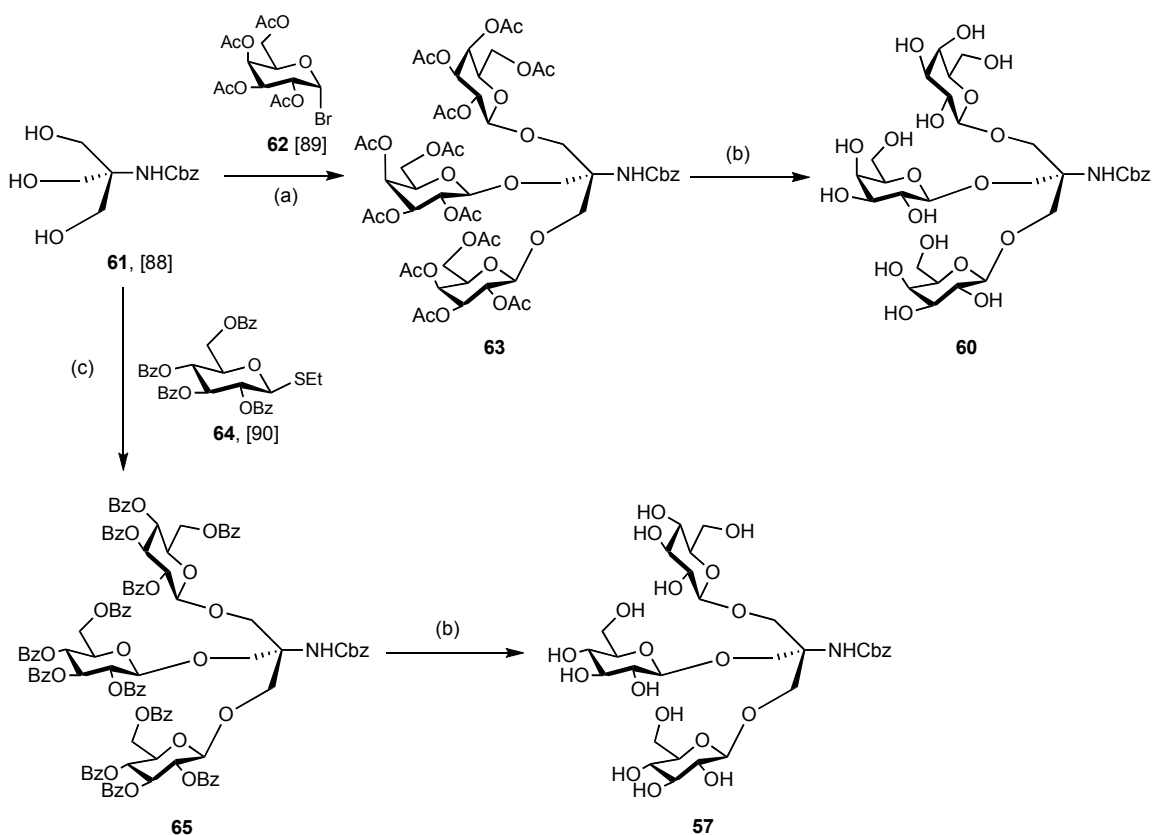
terms of a higher local sugar concentration. Thus, according to our estimates the binding hierarchy of the four compounds (**57**, **58**, **59** and **60**) should be as shown in Figure 14.

Compounds **57**, **58**, **59** and **60** were screened using BIACORE (Appendix 2), and their affinities were measured and compared to standards (Glc, Gal, GalNAc). The underlying assumption in all the BIACORE experiments was that no chelation or bridging of two H1-CRD subunits by one trivalent ligand was possible on the BIACORE chip. This was due to short inter-sugar distances on the ligands, and distances between immobilized H1-CRD proteins that could not be spanned by one molecule.

## 4.2 Results and Discussion

### 4.2.1 Synthesis of compounds **57**, **58**, **59** and **60**

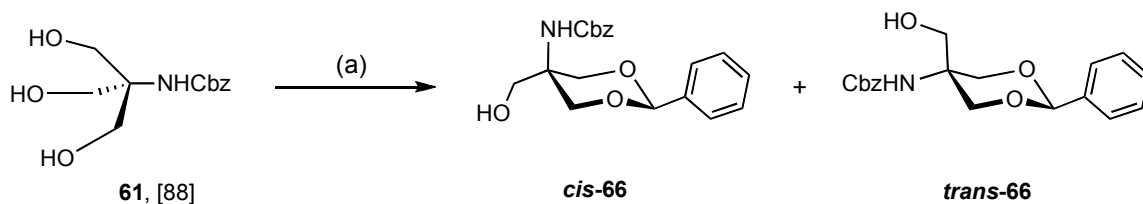
The Gal trimer **60** was prepared by glycosylating **61** [88] with 2,3,4,6-tetra-O-acetyl-1-bromo- $\alpha$ -D-galactopyranoside [89] (**62**, Scheme 6) using  $\text{Hg}(\text{CN})_2/\text{HgBr}_2$  as promoter ( $\rightarrow$  **63**), and was obtained in a 24% o.y. after deacetylation under standard Zemplén conditions.



**Scheme 6.** (a) Hg(CN)<sub>2</sub>, HgBr<sub>2</sub>, CH<sub>3</sub>CN, r.t., 17 h, 43%; (b) NaOMe, MeOH, r.t., 2 h, 56%; (c) NIS, TfOH, DCM/Et<sub>2</sub>O, 4 Å MS, 0 °C, 75%.

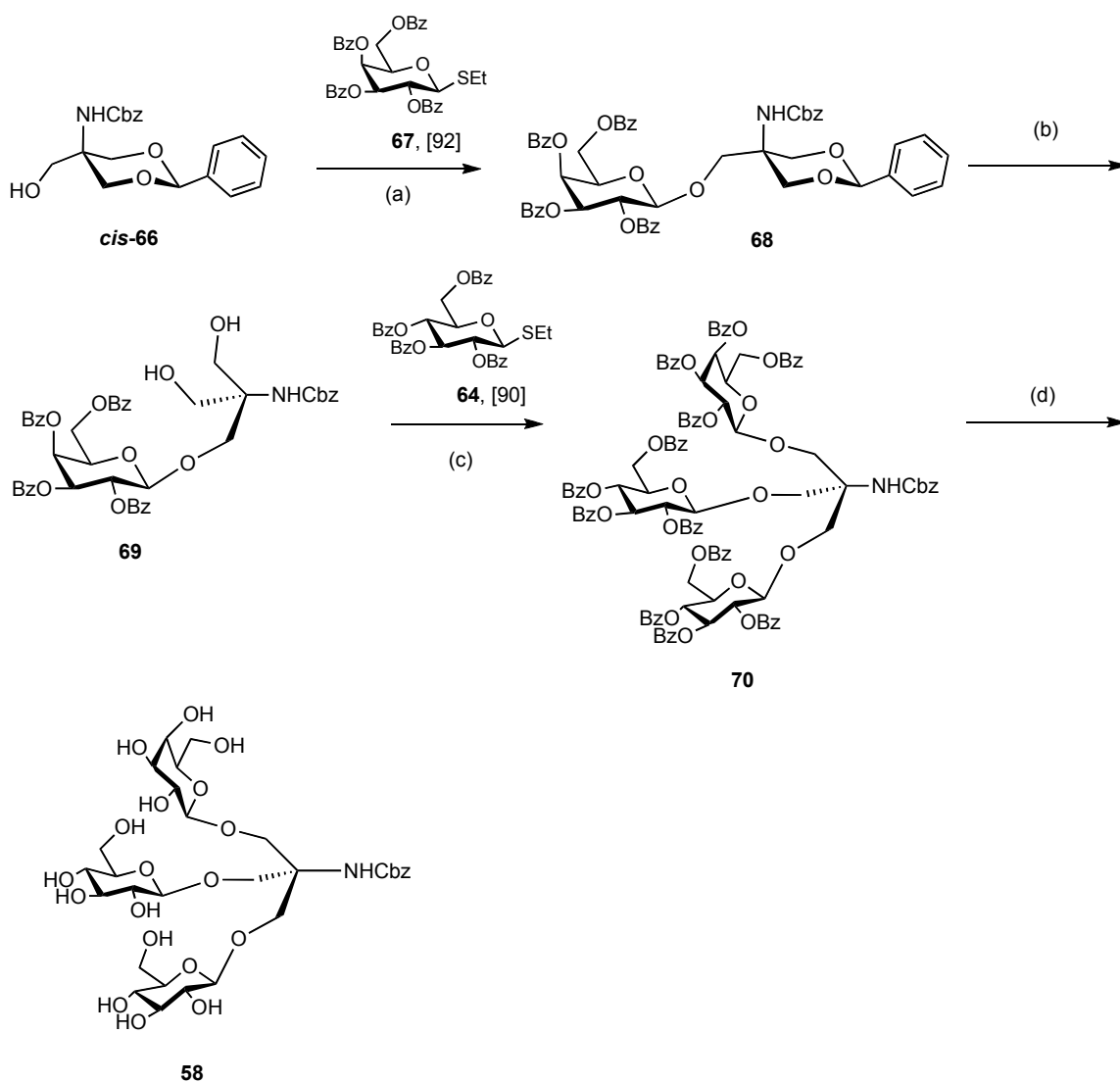
The Glc trimer **57** was prepared by glycosylating **61** [88] with ethyl 2,3,4,6-tetra-O-benzoyl-1-thio-β-D-glucopyranoside [90] (**64**, Scheme 6) using NIS/TfOH as promoter ( $\rightarrow$  **65**), and was obtained in a 69% o.y. after debenzoylation under standard Zemplén conditions.

For the synthesis of mono-Gal-di-Glc (**58**) and mono-Glc-di-Gal (**59**), two of the three hydroxy groups on **61** were protected via a benzylidene acetal using dimethoxy benzaldehyde and catalytic pTfOH•H<sub>2</sub>O to give the corresponding isomers of *cis*-**66** and *trans*-**66** (1.2:1, respectively) in a 99% yield (Scheme 7). The structure of the *cis*-**66** and *trans*-**66** diastereoisomers was determined by NMR according to the procedure described by Eliel *et al.* [91].



**Scheme 7.** (a) DBA, pTsOH•H<sub>2</sub>O, CH<sub>3</sub>CN, r.t., 24 h, 99%, *cis/trans* mixture (1.2:1).

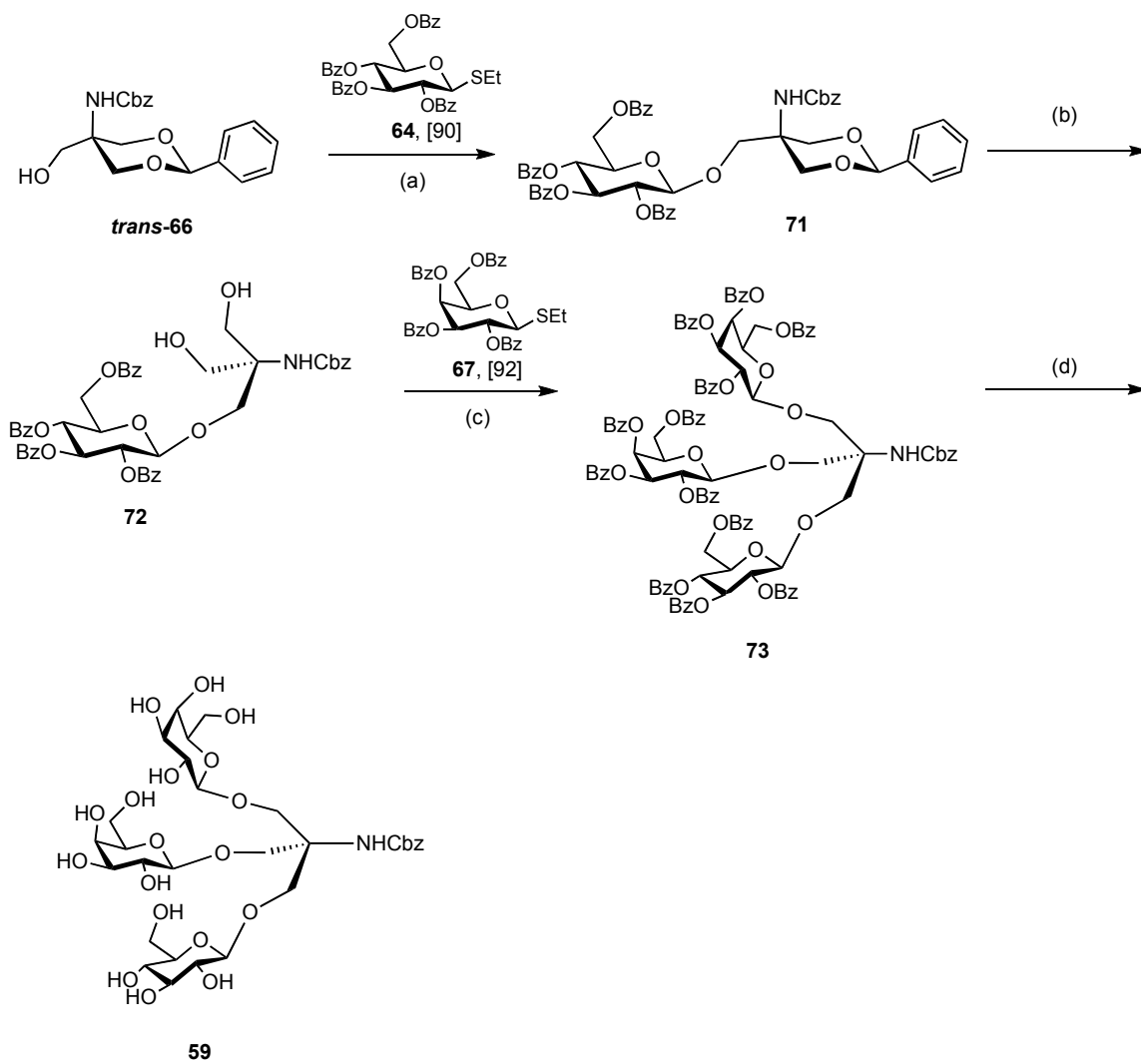
Compound *cis*-**66** was then glycosylated with ethyl 2,3,4,6-tetra-*O*-benzoyl-1-thio-β-*D*-galactopyranoside [92] (**67**, Scheme 8) using NIS/TfOH as promoter (→ **68**), the benzylidene protecting group was cleaved using 80% aqueous acetic acid giving **69**, which was glycosylated with ethyl 2,3,4,6-tetra-*O*-benzoyl-1-thio-β-*D*-glucopyranoside [90] (**64**) using NIS/TfOH in DCM/Et<sub>2</sub>O as promoter yielding **70**. The latter was debenzoylated under standard Zemplén conditions to give the mono-Gal-di-Glc trimer **58** in a 27% o.y. starting from *cis*-**66**.



**Scheme 8.** (a) NIS, TfOH, DCE/Et<sub>2</sub>O, 4 Å MS, -10 °C, 71%; (b) CH<sub>3</sub>CO<sub>2</sub>H (80% in H<sub>2</sub>O), 50 °C, 5 h, 60%; (c) NIS, TfOH, DCM/Et<sub>2</sub>O, 4 Å MS, 0 °C, 71%; (d) NaOMe, MeOH, r.t., 1.5 h, 89%.

By an analogous strategy, mono-Glc-di-Gal (**59**) was prepared (Scheme 9) in a 34% o.y. starting from *cis*-**66**.





**Scheme 9.** (a) NIS, TfOH, DCE/Et<sub>2</sub>O, 4 Å MS, 0 °C, 83%; (b) CH<sub>3</sub>CO<sub>2</sub>H (80% in H<sub>2</sub>O), 50 °C, 5 h, quant.; (c) NIS, TfOH, DCM/Et<sub>2</sub>O, 4 Å MS, 0 °C, 59%; (d) NaOMe, MeOH, r.t., 2 h, 69%.

## 4.2.2 Biological evaluation on BIACORE

(Performed by Daniel Ricklin, Institute of Molecular Pharmacy, University of Basel)

The four final compounds **57**, **58**, **59** and **60** were tested on BIACORE (Appendix 2) for their affinity towards the ASGP-R H1-CRD. The equilibrium dissociation constants ( $K_D$ s) were measured, and the results are summarized in Table 3.

**Table 3.** Summary of the BIACORE results for the trivalent compounds.

Entry	Compound	$K_D$ (mM)
1	GalNAc ( <b>2</b> )	0.12
2	Gal	2.2
3	Glc	>20
4	Gal-Gal-Gal ( <b>60</b> )	2.1
5	Gal-Gal-Glc ( <b>59</b> )	1.3
6	Gal-Glc-Glc ( <b>58</b> )	3.4
7	Glc-Glc-Glc ( <b>57</b> )	12

The results presented in Table 3 show a binding hierarchy that deviates from the expected one in Entries 4 and 5 (compounds **60** and **59**, respectively). Since the affinity of galactose is more than 10-fold higher than that of glucose (Table 3), it was expected that compound **60** (three Gal moieties) would have higher affinity than **59** (two Gal moieties), however even after repeated BIACORE experiments with different batches of protein, the binding hierarchy remained unchanged. Thus, it could be that the one Glc moiety in **59** makes an interaction with a part of the protein other than the binding site. Another possibility is due to the fact that the compounds are relatively compact and possess a restricted flexibility with a concomitant formation a tight network of hydrogen bonds that somehow influences the binding characteristics.

The increase in affinity due to multivalent interactions is a result of two components: statistical binding probability and the chelate effect [22]. The fact that no dramatic increases in affinity are observed for compounds **57**, **58**, **59** and **60** is a proof that the local concentration effect in this system is operating on a purely statistical level, *i.e.* more sugars in the vicinity of the binding site implies a higher probability of collision. The statistical component of multivalent interactions is less potent and influential than the chelate effect [22], and since there is no multivalent binding possible, no substantial increase in affinity is observed. This may also explain why compound **60** showed almost an identical  $K_D$  to that of Gal, in spite of possessing three vs. one Gal moiety.

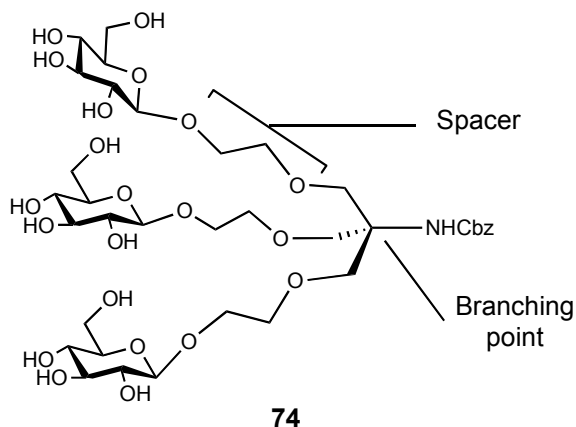
A similar result was observed by Kiessling *et al.* [93], who synthesized lactosylated neoglycopolymers for binding to galectin 3 (Gal-3) and *Erythrina corallodendrum* lectin (EcorL). In spite of the neoglycopolymer presenting many more sugar residues than the monovalent lactose, no exponential increase in relative affinity was observed, rather a small increase for statistical reasons was seen. The explanation for this is that both Gal-3 and EcorL possess sugar binding sites that are located on the opposite faces of the lectins, and hence no chelate effect by simultaneously binding to both sites was possible.

In contrast, the same research group managed to observe exponential enhancements in the binding affinity of mannose-substituted polymers designed to bind to the concavalin A (Con A) lectin [94]. Unlike Gal-3 and EcorL, Con A has the sugar binding sites on the same face of the protein, which does allow chelation to take place.

## 4.3 Conclusion

In this chapter, a set of four trivalent compounds (**57**, **58**, **59** and **60**) with a different affinity for the ASGP-R was introduced. The compounds were synthesized, and their  $K_D$  was measured on BIACORE. In order to understand and explain the surprising discrepancy in the binding hierarchy caused by compounds **59** and **60**, it might be useful to perform some modeling and NMR studies on the conformation of the compounds to determine the actual binding mode.

Furthermore, if the main reason for the observed  $K_D$  values for **59** and **60** is excessive sterical rigidity of the compounds, it might help to increase the spacer length between the trivalent branching point and the sugar residues (*i.e.* compound **74**, Figure 15), which would enable the sugars to rotate more freely for optimal binding.



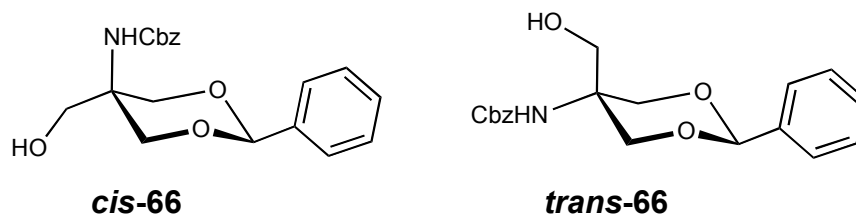
**Figure 15.** Structure of the proposed trivalent compound featuring extended spacers for increased flexibility.

Nevertheless, the results presented in this chapter further reinforce the accepted theory that the statistical local concentration component of multivalency is a lot weaker than the chelate effect [24].

## 4.4 Experimental

General methods are described in Appendix 4, p. 147.

***cis/trans*-5-Benzyloxycarbonylamino-5-hydroxymethyl-2-phenyl-1,3-dioxane (*cis/trans*-66):**



Compound **62** [88] (1.00 g, 3.92 mmol) was dissolved in dry CH<sub>3</sub>CN (30 ml), and benzaldehyde dimethylacetal (901 μl, 5.87 mmol) followed by *p*-toluenesulfonic acid monohydrate (74.1 mg, 0.392 mmol) were added. The solution was stirred at r.t. under argon for 24 h, quenched with triethylamine (2 ml), and the solvent was removed in vacuo. The residue was purified by flash chromatography on silica gel (petrol ether/EtOAc 7:3 → 3:2 → 1:1, with 0.5% triethylamine), yielding the two completely separated diastereoisomers of **66** (1.34 g, 99%, 1.2:1 *cis/trans* ratio, R<sub>f</sub> 0.28 *cis*, 0.36 *trans*, K) as oils, which crystallized upon storage.

***cis*-66:** <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ 3.76 (s, 2H, CH<sub>2</sub>OH), 3.86, 4.24 (A, B of AB, *J* = 11.7 Hz, 4H, H4, H6), 4.29 (bs, 1H, OH), 5.12 (s, 2H, CH<sub>2</sub>Ph), 5.48 (s, 1H, H2), 5.84 (s, 1H, NH), 7.33-7.50 (m, 10H, 2 × C<sub>6</sub>H<sub>5</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ 53.86 (C5), 64.2 (CH<sub>2</sub>OH), 67.2 (CH<sub>2</sub>Ph), 71.5 (2C, C4, C6), 101.9 (C2), 125.9, 128.1, 128.3, 128.6, 129.2, 135.8, 137.2 (12C, 2 × C<sub>6</sub>H<sub>5</sub>), 156.8 (C=O); Anal. Calcd. for C<sub>19</sub>H<sub>21</sub>NO<sub>5</sub>: C, 66.46; H, 6.16; N, 4.08. Found: C, 66.22; H, 6.18; N, 3.91. ***trans*-66:** <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ 2.95 (bs, 1H, OH), 4.06 (s, 2H, CH<sub>2</sub>OH), 4.16 (m, 4H, H4, H6), 5.07 (s, 2H, CH<sub>2</sub>Ph), 5.08 (s, 1H, NH),

5.51 (s, 1H, H2), 7.36-7.46 (m, 10H, 2 × C<sub>6</sub>H<sub>5</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ 52.0 (C5), 62.7 (CH<sub>2</sub>OH), 66.9 (CH<sub>2</sub>Ph), 68.9 (2C, C4, C6), 101.5 (C2), 126.1, 128.1, 128.3, 128.6, 129.1, 135.9, 137.3 (12C, 2 × C<sub>6</sub>H<sub>5</sub>), 155.5 (C=O); Anal. Calcd. for C<sub>19</sub>H<sub>21</sub>NO<sub>5</sub>: C, 66.46; H, 6.16; N, 4.08. Found: C, 65.80; H, 6.18; N, 3.90.

**5-(2,3,4,6-Tetra-O-benzoyl-β-D-galactopyranosyl)oxymethyl-5-benzyloxy-carbonylamino-2-phenyl-1,3-dioxane (68)**: Compound *cis*-**66** (100 mg, 0.291 mmol) and ethyl 2,3,4,6-tetra-O-benzoyl-1-thio-β-D-galactopyranoside [92] (**67**, 373 mg, 0.582 mmol) were dissolved in dry DCE/Et<sub>2</sub>O (20 ml, 1.5:1 v/v) and stirred with activated 4 Å mol sieves (1 g) at r.t. for 1 h under argon. The solution was cooled to -10 °C, and NIS (131 mg, 0.582 mmol) was added. The mixture was stirred for another 20 min before a drop of TfOH was added. The reaction was stirred at -10 °C under argon until completion (TLC), after which it was filtered through Celite diluted with DCM (10 ml) and washed with aqueous Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> (1 M, 25 ml), then aqueous NaHCO<sub>3</sub> (0.1 M, 25 ml). The organic phase was dried with Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated in vacuo. The residue was purified by flash chromatography on silica gel (petrol ether/EtOAc 7:3) obtaining **68** (190 mg, 71%, R<sub>f</sub> 0.22 L) as a colorless solid. [α]<sub>D</sub> = +54.9 (c 1, CHCl<sub>3</sub>); <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ 3.88 (A of AB, J = 11.8 Hz, 1H, CH<sub>2</sub>O, dioxane), 3.98 (A of AB, J = 11.8 Hz, 1H, CH<sub>2</sub>O, dioxane), 4.07 (A of AB, J = 10.7 Hz, 1H, CH<sub>2</sub>OGal), 4.19-4.24 (m, 3H, B of AB CH<sub>2</sub>O, dioxane, B of AB, CH<sub>2</sub>OGal, H5-Gal), 4.30 (B of AB, J = 11.7 Hz, 1H, CH<sub>2</sub>O, dioxane), 4.44 (dd, J = 6.6, 11.2 Hz, 1H, H6a-Gal), 4.66 (dd, J = 6.6, 11.3 Hz, 1H, H6b-Gal), 4.75 (d, J = 7.45 Hz, 1H, H1-Gal), 4.96, 5.07 (A, B of AB, J = 12.2 Hz, 2H, CH<sub>2</sub>Ph), 5.24 (s, 1H, H acetal) 5.32 (s, 1H, N-H), 5.61 (dd, J = 2.2, 10.4 Hz, 1H, H3-Gal), 5.79 (m, 1H, H2-Gal), 6.00 (s, 1H, H4-Gal), 7.23-8.12 (m, 30H, 6 × C<sub>6</sub>H<sub>5</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ 52.6 (C<sub>Q</sub>, dioxane), 62.0 (C6-Gal), 66.4 (CH<sub>2</sub>Ph), 68.1 (C4-Gal), 69.7 (CH<sub>2</sub>OGal), 69.8 (C2-Gal), 70.3 (2C, 2 × CH<sub>2</sub>O, dioxane), 71.2 (C3-Gal), 71.3 (C5-Gal), 101.4 (C acetal), 102.1 (C1-Gal), 125.1, 128.0-130.0, 133.2-134.1,

136.1, 137.3, 137.4 (36C, 6 × C<sub>6</sub>H<sub>5</sub>), 155.3 (C=O, Cbz), 165.2-165.9 (4C, 4 × C=O, OBz); ESI-MS: Calcd. for C<sub>53</sub>H<sub>48</sub>NO<sub>14</sub> [M+H]<sup>+</sup>: 922.31; Found m/z 922.52.

**Tris[(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl)oxymethyl]-N-benzyloxy-carbonylmethylamine (63):** To a solution of **61** [88] (50.1 mg, 0.196 mmol), Hg(CN)<sub>2</sub> (99.1 mg, 0.392 mmol) and HgBr<sub>2</sub> (141 mg, 0.392 mmol) in dry CH<sub>3</sub>CN (3 ml) was added a solution of 2,3,4,6-tetra-O-acetyl-1-bromo-α-D-galactopyranoside [89] (**62**, 322 mg, 0.784 mmol) in dry CH<sub>3</sub>CN (3 ml) dropwise over 30 min at r.t. under argon. After 2.5 h, more donor **62** (80.5 mg, 0.196 mmol) was added and the reaction was stirred for 13.5 h at r.t. under argon. The mixture was then diluted with DCM (50 ml), washed with aqueous KI solution (30%, 20 ml), then aqueous NaHCO<sub>3</sub> solution (0.1 M, 20 ml), followed by H<sub>2</sub>O (20 ml), and dried with Na<sub>2</sub>SO<sub>4</sub>. The solvent was removed in vacuo, and the residue was purified by flash chromatography on silica gel (DCM/acetone 20:1 → 9:1) obtaining **63** (105 mg, 43%, R<sub>f</sub> 0.16 M) as a colorless solid. [α]<sub>D</sub> = -9.2 (c 1, CHCl<sub>3</sub>); <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ 1.96, 1.98, 2.03, 2.13 (s, 36H, 12 × OAc), 3.78 (A of AB, J = 10.3 Hz, 3H, 3 × CH<sub>2</sub>O), 3.85 (m, 3H, 3 × H5-Gal), 4.03 (B of AB, J = 10.3 Hz, 3H, 3 × CH<sub>2</sub>O), 4.09 (m, 6H, 3 × H6-Gal), 4.40 (d, J = 7.5 Hz, 3H, 3 × H1-Gal), 4.96-5.07 (m, 5H, 3 × H3-Gal, CH<sub>2</sub>Ph), 5.12 (dd, J = 8.0, 10.4 Hz, 3H, 3 × H2-Gal), 5.33 (m, 4H, 3 × H4-Gal, N-H), 7.33 (m, 5H, C<sub>6</sub>H<sub>5</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ 20.6 (12C, 12 × CH<sub>3</sub>, OAc), 58.2 (3C, 3 × C6-Gal) 66.8 (4C, 3 × C4-Gal, CH<sub>2</sub>Ph), 68.4 (C<sub>Q</sub>), 68.8 (6C, 3 × C2-Gal, 3 × CH<sub>2</sub>O), 70.4 (6C, 3 × C3-Gal, 3 × C5-Gal), 101.6 (3C, 3 × C1-Gal), 128.2, 128.3, 128.4, 136.0 (6C, C<sub>6</sub>H<sub>5</sub>), 154.7 (C=O, Cbz), 169.2, 169.8, 170.1, 170.2 (12C, 12 × C=O, OAc); ESI-MS: Calcd. for C<sub>54</sub>H<sub>72</sub>NO<sub>32</sub> [M+H]<sup>+</sup>: 1246.4; Found m/z 1246.65; Anal. Calcd. for C<sub>54</sub>H<sub>71</sub>NO<sub>32</sub>: C, 52.05; H, 5.74; N, 1.12. Found: C, 52.68; H, 6.05; N, 1.23.

**Tris[(β-D-galactopyranosyl)oxymethyl]-N-benzyloxycarbonylmethylamine (60):** Compound **63** (125 mg, 0.099 mmol) was dissolved in dry methanol (2.5 ml) and a solution of NaOMe in MeOH (0.1 M, 2.5 ml) was added. The resultant

mixture was stirred at r.t. under argon for 2 h. The solution was neutralized with Dowex 50X8 (H<sup>+</sup>-form), filtered and the solvent removed in vacuo. The residue was purified on an RP-C18 column (H<sub>2</sub>O/MeOH, stepwise gradient 20:0 → 17:3) to give **60** (42 mg, 56%, R<sub>f</sub> 0.25 N). [α]<sub>D</sub> = -0.085 (c 1, MeOH); <sup>1</sup>H-NMR (500 MHz, CD<sub>3</sub>OD): δ 3.45-3.54 (m, 9H, 3 × H2-, 3 × H3-, 3 × H5-Gal), 3.69 (dd, J = 5.4, 11.4 Hz, 3H, 3 × H6a-Gal), 3.75 (dd, J = 6.9, 11.3 Hz, 3H, 3 × H6b-Gal), 3.81 (d, J = 3.2 Hz, 3H, H4-Gal), 3.90 (A of AB, J = 10.3 Hz, 3H, 3 × CH<sub>2</sub>O), 4.24-4.28 (m, 6H, 3 × H1-Gal, 3 × CH<sub>2</sub>O), 4.59 (bs, 1H, N-H), 5.03 (m, 2H, CH<sub>2</sub>Ph), 7.27-7.37 (m, 5H, C<sub>6</sub>H<sub>5</sub>); <sup>13</sup>C NMR (125 MHz, CD<sub>3</sub>OD): δ 60.1 (C<sub>Q</sub>, Tris), 62.4 (3C, 3 × C6-Gal), 67.2 (CH<sub>2</sub>Ph), 69.6 (3C, 3 × CH<sub>2</sub>O), 70.3 (3C, 3 × C4-Gal), 72.4 (3C, 3 × C2-Gal), 74.8 (3C, 3 × C3-Gal), 76.7 (3C, 3 × C5-Gal), 105.6 (3C, 3 × C1-Gal), 129.1, 129.6, 138.1 (6C, C<sub>6</sub>H<sub>5</sub>), 157.4 (C=O, Cbz); ESI-MS: Calcd. for C<sub>30</sub>H<sub>47</sub>NO<sub>20</sub>Na [M+Na]<sup>+</sup>: 764.26; Found m/z 764.40.

**Tris[(2,3,4,6-tetra-O-benzoyl-β-D-glucopyranosyl)oxymethyl]-N-benzyloxy-carbonylmethylamine (65)**: Compound **61** [88] (50.1 mg, 0.196 mmol) and ethyl 2,3,4,6-tetra-O-benzoyl-1-thio-β-D-glucopyranoside [90] (**64**, 500 mg, 0.785 mmol) were dissolved in dry DCM/Et<sub>2</sub>O (15 ml, 2:1 v/v) and stirred with activated 4 Å mol sieves (500 g) at r.t. for 2 h under argon. The solution was cooled to 0 °C, and NIS (176 mg, 0.785 mmol) was added, followed by a drop of TfOH. The reaction was stirred at 0 °C under argon for 1 h, then more TfOH (10 μl) was added, and the reaction was stirred at r.t. under argon until completion (TLC). The mixture was then filtered through Celite, diluted with DCM (10 ml) and washed with aqueous Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> (1 M, 25 ml), then aqueous NaHCO<sub>3</sub> (0.1 M, 25 ml). The organic phase was dried with Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated in vacuo. The residue was purified by flash chromatography on silica gel (petrol ether/EtOAc 7:3 → 3:2) obtaining **65** (293 mg, 75%, R<sub>f</sub> 0.61 A) as a colorless solid. All analytical data were consistent with those published [88].



**5-(2,3,4,6-Tetra-O-benzoyl- $\beta$ -D-glucopyranosyl)oxymethyl-5-benzyloxy-carbonylamino-2-phenyl-1,3-dioxane (71):** Compound *cis*-**66** (100 mg, 0.291 mmol) and ethyl 2,3,4,6-tetra-O-benzoyl-1-thio- $\beta$ -D-glucopyranoside [90] (**64**, 373 mg, 0.582 mmol) were dissolved in dry DCM/Et<sub>2</sub>O (20 ml, 1.5:1 v/v) and stirred with activated 4 Å mol sieves (1 g) at r.t. for 2 h under argon. The solution was cooled to 0 °C, and NIS (131 mg, 0.582 mmol) was added. The mixture was stirred for 1 h before TfOH (15  $\mu$ l, 0.172 mmol) was added. The reaction was stirred at 0 °C under argon until completion (TLC), when it was filtered through Celite, diluted with DCM (10 ml) and washed with aqueous Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> (1 M, 25 ml), then aqueous NaHCO<sub>3</sub> (0.1 M, 25 ml). The organic phase was dried with Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated in vacuo. The residue was purified by flash chromatography on silica gel (petrol ether/EtOAc 4:1  $\rightarrow$  7:3) obtaining **71** (230 mg, 83%, R<sub>f</sub> 0.25 L) as a colorless solid.  $[\alpha]_D = +12.5$  (c 1, CHCl<sub>3</sub>); <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  3.82 (A of AB, *J* = 11.7 Hz, 1H, CH<sub>2</sub>O, dioxane), 3.91 (A of AB, *J* = 11.7 Hz, 1H, CH<sub>2</sub>O, dioxane), 3.95-4.04 (m, 2H, H5-Glc, A of AB, CH<sub>2</sub>OGlc), 4.19-4.24 (m, 3H, 2  $\times$  B of AB CH<sub>2</sub>O, dioxane, B of AB, CH<sub>2</sub>OGlc), 4.45 (dd, *J* = 4.9, 12.2 Hz, 1H, H6a-Glc), 4.65 (dd, *J* = 2.9, 12.1 Hz, 1H, H6b-Glc), 4.75 (d, *J* = 7.6 Hz, 1H, H1-Glc), 4.90, 5.04 (A, B of AB, *J* = 12.3 Hz, 2H, CH<sub>2</sub>Ph), 5.22 (s, 1H, N-H), 5.28 (s, 1H, H acetal), 5.51 (dd, *J* = 8.0, 9.8 Hz, 1H, H2-Glc), 5.68 (t, *J* = 9.7 Hz, 1H, H4-Glc), 5.87 (t, *J* = 9.7 Hz, 1H, H3-Glc), 7.25-8.55, 7.83-8.06 (m, 30H, 6  $\times$  C<sub>6</sub>H<sub>5</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  52.4 (C<sub>O</sub>, dioxane), 62.7 (C6-Glc), 66.4 (CH<sub>2</sub>Ph), 68.8 (CH<sub>2</sub>OGlc), 69.5 (C4-Glc), 69.8, 70.3 (2C, 2  $\times$  CH<sub>2</sub>O, dioxane), 71.8 (C2-Glc), 72.2 (C5-Glc), 72.6 (C3-Glc), 101.5 (C acetal), 101.7 (C1-Glc), 125.9-129.8, 133.1, 133.3, 133.4, 133.5, 137.4 (36C, 6  $\times$  C<sub>6</sub>H<sub>5</sub>), 155.3 (C=O, Cbz), 165.1, 165.7, 166.1 (4C, 4  $\times$  C=O, OBz); ESI-MS: Calcd. for C<sub>53</sub>H<sub>47</sub>NO<sub>14</sub>Na [M+Na]<sup>+</sup>: 944.29; Found *m/z* 944.43; Anal. Calcd. for C<sub>53</sub>H<sub>47</sub>NO<sub>14</sub>: C, 69.05; H, 5.14; N, 1.52. Found: C, 68.60; H, 5.14; N, 1.61.

**2-(2,3,4,6-Tetra-O-benzoyl- $\beta$ -D-glucopyranosyl)oxymethyl-2-benzyloxy-carbonylaminopropan-1,3-diol (72):** Compound **71** (110 mg, 0.119 mmol) was dissolved in 80% aqueous acetic acid (5 ml), and the resultant solution was

heated at 50 °C for 3 h. The solvent was removed in vacuo, and the residue was purified by flash chromatography on silica gel (petrol ether/EtOAc 3:2 → 1:1) giving **72** (99 mg, quant., R<sub>f</sub> 0.17 D) as a solid. [ $\alpha$ ]<sub>D</sub> = +13.7 (c 1, CHCl<sub>3</sub>); <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  3.48 (m, 2H, CH<sub>2</sub>OH), 3.65 (m, 2H, CH<sub>2</sub>OH), 3.87, 4.02 (A, B of AB, *J* = 10.1 Hz, 2H, CH<sub>2</sub>OGlc), 4.09 (m, 1H, H5-Glc), 4.44 (dd, *J* = 5.1, 12.2 Hz, 1H, H6a-Glc), 4.67 (dd, *J* = 2.9, 12.2 Hz, 1H, H6b-Glc), 4.80 (d, *J* = 7.9 Hz, 1H, H1-Glc), 4.89, 4.99 (A, B of AB, *J* = 12.2 Hz, 2H, CH<sub>2</sub>Ph), 5.48 (dd, *J* = 7.9, 9.8 Hz, 1H, H2-Glc), 5.68 (t, *J* = 9.7 Hz, 1H, H4-Glc), 5.74 (s, 1H, N-H), 5.92 (t, *J* = 9.7 Hz, 1H, H3-Glc), 7.25-8.04 (m, 25H, 5 × C<sub>6</sub>H<sub>5</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  59.3 (C<sub>Q</sub>, Tris), 62.7 (C6-Glc), 63.8 (2C, 2 × CH<sub>2</sub>OH), 66.9 (CH<sub>2</sub>Ph), 69.4 (C4-Glc), 70.5 (CH<sub>2</sub>OGlc), 72.0 (C2-Glc), 72.3 (C5-Glc), 72.4 (C3-Glc), 101.4 (C1-Glc), 128.0-128.8, 129.7, 129.8, 133.2, 133.5, 136.0 (30C, 5 × C<sub>6</sub>H<sub>5</sub>), 156.9 (C=O, Cbz), 165.1, 165.7, 165.8, 166.2 (5C, 5 × C=O, OBz); ESI-MS: Calcd. for C<sub>46</sub>H<sub>43</sub>NO<sub>14</sub>Na [M+Na]<sup>+</sup>: 856.26; Found *m/z* 856.23

#### **Tris[( $\beta$ -D-glucopyranosyl)oxymethyl]-*N*-benzyloxycarbonylmethylamine**

**(57)**: Compound **65** (276 mg, 0.139 mmol) was dissolved in dry methanol (2.5 ml), a solution of NaOMe in MeOH (0.2 M, 2.5 ml) was added, and the resultant mixture was stirred at r.t. under argon for 1.5 h. It was then neutralized with Dowex 50X8 (H<sup>+</sup>-form), filtered and the solvent removed in vacuo. The residue was purified on an RP-C18 column (H<sub>2</sub>O/MeOH, stepwise gradient 1:0 → 3:1) to give **57** (95 mg, 92%, R<sub>f</sub> 0.44 N). All analytical data were consistent with those reported [88].

#### **2-(2,3,4,6-Tetra-*O*-benzoyl- $\beta$ -D-galactopyranosyl)oxymethyl-2-benzyloxy-**

**carbonylaminopropan-1,3-diol (69)**: Compound **68** (115 mg, 0.125 mmol) was dissolved in 80% aqueous acetic acid (5 ml), and the resultant solution was heated at 50 °C for 5 h. The solvent was removed in vacuo, and the residue was purified by flash chromatography on silica gel (petrol ether/EtOAc 3:2 → 1:1) giving **69** (62 mg, 60%, R<sub>f</sub> 0.17 D) as a solid. [ $\alpha$ ]<sub>D</sub> = +61.7 (c 1, CHCl<sub>3</sub>); <sup>1</sup>H-NMR

(500 MHz, CDCl<sub>3</sub>):  $\delta$  3.39-3.65 (m, 4H, 2  $\times$  CH<sub>2</sub>OH), 3.81 (A of AB,  $J$  = 10.1 Hz, 1H, CH<sub>2</sub>OGal), 3.98 (B of AB,  $J$  = 10.1 Hz, 1H, CH<sub>2</sub>OGal), 4.21 (t,  $J$  = 6.5 Hz 1H, H5-Gal), 4.35 (dd,  $J$  = 6.6, 11.3 Hz, 1H, H6a-Gal), 4.54 (dd,  $J$  = 6.6, 11.3 Hz, 1H, H6b-Gal), 4.71 (d,  $J$  = 7.7 Hz, 1H, H1-Gal), 4.82, 4.94 (A, B of AB,  $J$  = 12.2 Hz, 2H, CH<sub>2</sub>Ph), 5.58 (dd,  $J$  = 3.2, 10.4 Hz, 1H, H3-Gal), 5.64 (m, 1H, H2-Gal), 5.69 (s, 1H, N-H), 5.91 (d,  $J$  = 3.1 Hz, 1H, H4-Gal), 7.14-8.01 (m, 25H, 5  $\times$  C<sub>6</sub>H<sub>5</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  59.3 (C<sub>Q</sub>, Tris), 61.8 (C6-Gal), 63.7, 64.1 (2C, 2  $\times$  CH<sub>2</sub>OH), 66.9 (CH<sub>2</sub>Ph), 67.9 (C4-Gal), 69.9 (C2-Gal), 70.5 (CH<sub>2</sub>OGal), 71.0 (C3-Gal), 71.5 (C5-Gal), 101.8 (C1-Gal), 128.0-129.9, 133.3-133.6, 136.0 (30C, 5  $\times$  C<sub>6</sub>H<sub>5</sub>), 156.7 (C=O, Cbz), 165.4, 165.5, 165.8, 165.9, 166.0 (5  $\times$  C=O, OBz); ESI-MS: Calcd. for C<sub>46</sub>H<sub>43</sub>NO<sub>14</sub>Na [M+Na]<sup>+</sup>: 856.26; Found m/z 856.37.

**(2,3,4,6-Tetra-O-Benzoyl- $\beta$ -D-galactopyranosyl)oxymethyl-bis[(2,3,4,6-tetra-O-benzoyl- $\beta$ -D-glucopyranosyl)oxymethyl]-N-benzyloxycarbonylmethylamine (70):** Compound **69** (61 mg, 73.1  $\mu$ mol) and ethyl 2,3,4,6-tetra-O-benzoyl-1-thio- $\beta$ -D-glucopyranoside [90] (**64**, 164 mg, 0.256 mmol) were dissolved in dry DCM/Et<sub>2</sub>O (15 ml, 2:1 v/v) and stirred with activated 4 Å mol sieves (500 mg) at r.t. for 2 h under argon. The solution was cooled to 0 °C, and NIS (57.6 mg, 0.256 mmol) was added, followed by TfOH (15  $\mu$ l, 0.175 mmol). The reaction was stirred at 0 °C under argon until completion (TLC), after which it was filtered through Celite, diluted with DCM (10 ml) and washed with aqueous Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> (1 M, 25 ml), then aqueous NaHCO<sub>3</sub> (0.1 M, 25 ml). The organic phase was dried with Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated in vacuo. The residue was purified by flash chromatography on silica gel (petrol ether/EtOAc 7:3) obtaining **70** (104 mg, 71%, R<sub>f</sub> 0.47 D) as a solid.  $[\alpha]_D^{25} = +13.1$  (c 1, CHCl<sub>3</sub>); <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  3.36-3.43 (m, 5H, CH<sub>2</sub>OGal, CH<sub>2</sub>OGlc<sub>a</sub>, CH<sub>2</sub>OGlc<sub>b</sub>, H5-Glc<sub>a</sub>, H5-Glc<sub>b</sub>), 3.66 (t,  $J$  = 6.6 Hz, 1H, H5-Gal), 3.79 (d,  $J$  = 7.9 Hz, 1H, H1-Glc<sub>a</sub>), 3.85 (d,  $J$  = 7.9 Hz, 1H, H1-Glc<sub>b</sub>), 4.01 (d,  $J$  = 7.9 Hz, 1H, H1-Gal), 4.06 (B of AB,  $J$  = 10.6 Hz, 1H, CH<sub>2</sub>OGlc<sub>a</sub>), 4.17-4.28 (m, 5H, H6a-Gal, H6a-Glc<sub>a</sub>, H6a-Glc<sub>b</sub>, CH<sub>2</sub>OGal, CH<sub>2</sub>OGlc<sub>b</sub>), 4.39-4.45 (m, 3H, H6b-Gal, H6b-Glc<sub>a</sub>, H6b-Glc<sub>b</sub>), 4.66 (s, 2H, CH<sub>2</sub>Ph), 5.16-5.21 (m, 2H, H2-Glc<sub>a</sub>, H2-Glc<sub>b</sub>), 5.24 (dd,  $J$  = 3.5, 10.5 Hz, 1H, H3-

Gal), 5.34 (m, 2H, H4-Glc<sub>a</sub>, H4-Glc<sub>b</sub>), 5.45-5.58 (m, 3H, H2-Gal, H3-Glc<sub>a</sub>, H3-Glc<sub>b</sub>), 5.73 (d,  $J = 3.3$  Hz, 1H, H4-Gal), 7.13-7.99 (m, 65H,  $13 \times \text{C}_6\text{H}_5$ );  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  58.7 (C<sub>Q</sub>, Tris), 61.9 (C6-Gal), 62.9, 63.0 (2C, C6-Glc<sub>a</sub>, C6-Glc<sub>b</sub>), 66.1 ( $\text{CH}_2\text{Ph}$ ), 67.9 (C4-Gal), 68.2, 68.5, 68.7 (3C,  $\text{CH}_2\text{OGlc}_a$ , -Glc<sub>b</sub>, -Gal) 69.4 (2C, H4-Glc<sub>a</sub>, H4-Glc<sub>b</sub>), 69.8 (H4-Gal), 70.5, 71.0, 71.2, 71.7, 71.9, 72.4 (C2-Glc<sub>a</sub>, -Glc<sub>b</sub>, C3-Glc<sub>a</sub>, -Glc<sub>b</sub>, -Gal, C5-Glc<sub>a</sub>, -Glc<sub>b</sub>, -Gal), 101.3 (C1-Glc<sub>b</sub>), 101.5 (C1-Glc<sub>a</sub>), 101.9 (C1-Gal), 127.6-130.3, 133.2-133.7, 136.4 (78C,  $13 \times \text{C}_6\text{H}_5$ ) 154.8 (C=O, Cbz), 164.5-166.1 (12C,  $12 \times \text{C}=\text{O}$ , OBz); ESI-MS: Calcd. for  $\text{C}_{114}\text{H}_{95}\text{NO}_{32}\text{Na}$   $[\text{M}+\text{Na}]^+$ : 2012.57; Found  $m/z$  2013.74; Anal. Calcd. for:  $\text{C}_{114}\text{H}_{95}\text{NO}_{32}$  C, 68.77; H, 4.81; N, 0.70. Found: C, 68.61; H, 5.04; N, 0.81.

**( $\beta$ -D-Galactopyranosyl)oxymethyl-bis[( $\beta$ -D-glucopyranosyl)oxymethyl]-N-**

**benzyloxycarbonylmethylamine (58):** Compound **70** (90.1 mg, 45.2  $\mu\text{mol}$ ) was dissolved in dry methanol/dioxane (3 ml, 2:1 v/v) and a solution of NaOMe in MeOH (0.2 M, 1 ml) was added. The resultant mixture was stirred at r.t. under argon for 1.5 h. The solution was neutralized with Dowex 50X8 ( $\text{H}^+$ -form), filtered and the solvent was removed in vacuo. The residue was purified on an RP-C18 column ( $\text{H}_2\text{O}/\text{MeOH}$ , stepwise gradient 1:0  $\rightarrow$  3:1) to give **58** (30 mg, 89%,  $R_f$  0.32 N).  $[\alpha]_D = -11.6$  (c 1, MeOH);  $^1\text{H}$ -NMR (500 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  3.18-3.22 (m, 2H, H2-Glc<sub>a</sub>, H2-Glc<sub>b</sub>), 3.28 (m, 2H, H4-Glc<sub>a</sub>, H4-Glc<sub>b</sub>), 3.31 (m, 2H, H5-Glc<sub>a</sub>, H5-Glc<sub>b</sub>), 3.36 (m, 2H, H3-Glc<sub>a</sub>, H3-Glc<sub>b</sub>), 3.45-3.55 (m, 3H, H2-Gal, H3-Gal, H5-Gal) 3.64 (m, 2H, H6a-Glc<sub>a</sub>, H6a-Glc<sub>b</sub>), 3.70 (dd,  $J = 5.3, 11.4$  Hz, 1H, H6a-Gal), 3.76 (dd,  $J = 6.9, 11.4$  Hz, 1H, H6b-Gal), 3.82-3.93 (m, 6H,  $\text{CH}_2\text{OGlc}_a$ ,  $\text{CH}_2\text{OGlc}_b$ ,  $\text{CH}_2\text{OGal}$ , H6b-Glc<sub>a</sub>, H6b-Glc<sub>b</sub>, H4-Gal), 4.22-4.33 (m, 6H, H1-Glc<sub>a</sub>, H1-Glc<sub>b</sub>, H1-Gal,  $\text{CH}_2\text{OGlc}_a$ ,  $\text{CH}_2\text{OGlc}_b$ ,  $\text{CH}_2\text{OGal}$ ), 5.04 (s, 2H,  $\text{CH}_2\text{Ph}$ ), 7.27-7.37 (m, 5H,  $\text{C}_6\text{H}_5$ );  $^{13}\text{C}$  NMR (125 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  60.1 (C<sub>Q</sub>, Tris), 62.4 (C6-Gal), 62.7 (2C, C6-Glc<sub>a</sub>, C6-Glc<sub>b</sub>), 67.2 ( $\text{CH}_2\text{Ph}$ ), 69.5 (3C,  $\text{CH}_2\text{OGlc}_a$ , -Glc<sub>b</sub>, -Gal), 70.2 (C4-Gal) 71.5 (2C, C4-Glc<sub>a</sub>, C4-Glc<sub>b</sub>), 72.5 (C2-Gal), 74.8 (C3-Gal), 75.0 (2C, C2-Glc<sub>a</sub>, C2-Glc<sub>b</sub>), 76.7 (C5-Gal), 77.9, 78.0 (4C, C3-Glc<sub>a</sub>, C3-Glc<sub>b</sub>, C5-Glc<sub>a</sub>, C5-Glc<sub>b</sub>)

104.9 (2C, C1-Glc<sub>a</sub>, C1-Glc<sub>b</sub>), 105.4 (C1-Gal), 128.9, 129.0, 129.5, 138.2 (6C, C<sub>6</sub>H<sub>5</sub>), 157.5 (1C, C=O, Cbz); ESI-MS: Calcd. for C<sub>30</sub>H<sub>47</sub>NO<sub>20</sub>Na [M+Na]<sup>+</sup>: 764.26; Found m/z 764.34.

**Bis[(2,3,4,6-tetra-O-benzoyl-β-D-galactopyranosyl)oxymethyl]-(2,3,4,6-tetra-O-benzoyl-β-D-glucopyranosyl)oxymethyl-N-benzyloxycarbonylmethylamine (73):** Compound **72** (50.1 mg, 59.9 μmol) and ethyl 2,3,4,6-tetra-O-benzoyl-1-thio-β-D-galactopyranoside [92] (**67**, 134 mg, 0.209 mmol) were dissolved in dry DCM/Et<sub>2</sub>O (15 ml, 2:1 v/v) and stirred with activated 4 Å mol sieves (500 mg) at r.t. for 2 h under argon. The solution was cooled to 0 °C, and NIS (46.9 mg, 0.209 mmol) was added, followed by TfOH (10 μl, 20.9 μmol). The reaction was stirred at 0 °C under argon until completion (TLC), after which it was filtered through Celite, diluted with DCM (10 ml) and washed with aqueous Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> (1 M, 25 ml), then aqueous NaHCO<sub>3</sub> (5% w/v, 25 ml). The organic phase was dried with Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated in vacuo. The residue was purified by flash chromatography on silica gel (petrol ether/EtOAc 7:3 → 3:2) obtaining **73** (70 mg, 59%, R<sub>f</sub> 0.45 D) as a solid. [α]<sub>D</sub> = +17.4 (c 0.5, CHCl<sub>3</sub>); <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ 3.42 (m, 1H, H5-Glc), 3.51-3.57 (m, 3H, CH<sub>2</sub>OGal<sub>a</sub>, CH<sub>2</sub>OGal<sub>b</sub>, CH<sub>2</sub>OGlc), 3.71 (t, *J* = 6.7 Hz, 1H, H5-Gal<sub>b</sub>), 3.76 (t, *J* = 6.6 Hz, 1H, H5-Gal<sub>a</sub>), 3.82 (d, *J* = 7.8 Hz, 1H, H1-Glc), 4.06 (d, *J* = 7.9 Hz, 1H, H1-Gal<sub>a</sub>) 4.10-4.16 (m, 3H, H1-Gal<sub>b</sub>, CH<sub>2</sub>OGal<sub>a</sub>, CH<sub>2</sub>OGal<sub>b</sub>), 4.22-4.37 (m, 4H, CH<sub>2</sub>OGlc, H6b-Gal<sub>a</sub>, H6b-Gal<sub>b</sub>, H6b-Glc), 4.41-4.55 (m, 3H, H6a-Gal<sub>a</sub>, H6a-Gal<sub>b</sub>, H6a-Glc) 4.74 (s, 2H, CH<sub>2</sub>Ph), 5.24-5.38 (m, 3H, H2-Glc, H3-Gal<sub>a</sub>, H3-Gal<sub>b</sub>), 5.42 (t, *J* = 9.7 Hz, 1H, H4-Glc), 5.56 (t, *J* = 9.7 Hz, 1H, H3-Glc), 5.59-5.66 (m, 2H, H2-Gal<sub>a</sub>, H2-Gal<sub>b</sub>), 5.81-5.84 (m, 2H, H4-Gal<sub>a</sub>, H4-Gal<sub>b</sub>), 7.21-8.31 (m, 75H, 13 × C<sub>6</sub>H<sub>5</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ 58.8 (C<sub>Q</sub>, Tris), 61.7, 61.9 (C6-Gal<sub>a</sub>, C6-Gal<sub>b</sub>), 62.9 (C6-Glc), 66.0 (CH<sub>2</sub>Ph), 68.1, 68.5, 68.9 (CH<sub>2</sub>OGal<sub>a</sub>, -Gal<sub>b</sub>, -Glc), 69.1. 69.2 (3C, C2-Gal<sub>a</sub>, C2-Gal<sub>b</sub>, C4Glc), 71.1, 71.2 (4C, C3-Gal<sub>a</sub>, C3-Gal<sub>b</sub>, C5-Gal<sub>a</sub>, C5-Gal<sub>b</sub>), 71.7 (C2-Glc), 72.0 (C5-Glc), 72.3 (C3-Glc), 101.4 (C1-Glc), 101.8 (C1-Gal<sub>b</sub>), 102.0 (C1-Gal<sub>a</sub>), 128.2-129.9, 133.2-133.7, 136.4 (78C, 13 × C<sub>6</sub>H<sub>5</sub>), 154.8

(C=O, Cbz), 164.5-166.0 (12C, 12 × C=O, OBz); ESI-MS: Calcd. for C<sub>114</sub>H<sub>95</sub>NO<sub>32</sub>Na [M+Na]<sup>+</sup>: 2012.57; Found m/z 2014.00; Anal. Calcd. for: C<sub>114</sub>H<sub>95</sub>NO<sub>32</sub> C, 68.77; H, 4.81; N, 0.70. Found: C, 68.55; H, 4.89; N, 0.83.

**Bis[(β-D-galactopyranosyl)oxymethyl]-(β-D-glucopyranosyl)oxymethyl-N-benzyloxycarbonylmethylamine (59)**: Compound **73** (70.1 mg, 35.2 μmol) was dissolved in dry methanol/dioxane (3 ml, 2:1 v/v) and a solution of NaOMe in MeOH (0.2 M, 1 ml) was added. The resultant mixture was stirred at r.t. under argon for 2 h. The solution was neutralized with Dowex 50X8 (H<sup>+</sup>-form), filtered and the solvent removed in vacuo. The residue was purified on an RP-C18 column (H<sub>2</sub>O/MeOH, stepwise gradient 1:0 → 9:1) to give **59** (18 mg, 69%, R<sub>f</sub> 0.30 N). [α]<sub>D</sub> = -6.1 (c 1, MeOH); <sup>1</sup>H-NMR (500 MHz, CD<sub>3</sub>OD): δ 3.19 (m, 1H, H2-Glc), 3.27 (m, 1H, H4-Glc), 3.29 (m, 1H, H5-Glc), 3.35 (t, *J* = 8.9 Hz, 1H, H3-Glc) 3.44-3.54 (m, 6H, H2-Gal<sub>a</sub>, H2-Gal<sub>b</sub>, H3-Gal<sub>a</sub>, H3-Gal<sub>b</sub>, H5-Gal<sub>a</sub>, H5-Gal<sub>b</sub>), 3.63 (m, 1H, H6a-Glc), 3.69 (dd, *J* = 5.4, 11.4 Hz, 2H, H6a-Gal<sub>a</sub>, H6a-Gal<sub>b</sub>), 3.75 (dd, *J* = 6.9, 11.4 Hz, 2H, H6b-Gal<sub>a</sub>, H6b-Gal<sub>b</sub>), 3.81 (d, *J* = 3.14 Hz, 2H, H4-Gal<sub>a</sub>, H4-Gal<sub>b</sub>), 3.83-3.92 (m, 4H, H6b-Glc, CH<sub>2</sub>OGlc, CH<sub>2</sub>OGal<sub>a</sub>, CH<sub>2</sub>OGal<sub>b</sub>), 3.93-4.29 (m, 5H, H1-Gal<sub>a</sub>, H1-Gal<sub>b</sub>, CH<sub>2</sub>OGlc, CH<sub>2</sub>OGal<sub>a</sub>, CH<sub>2</sub>OGal<sub>b</sub>), 4.31 (d, *J* = 7.8 Hz, 1H, H1-Glc), 5.03 (m, 2H, CH<sub>2</sub>Ph), 7.27-7.37 (m, 5H, C<sub>6</sub>H<sub>5</sub>); <sup>13</sup>C NMR (125 MHz, CD<sub>3</sub>OD): δ 60.1 (C<sub>Q</sub>, Tris); 62.4 (2C, C6-Gal<sub>a</sub>, C6-Gal<sub>b</sub>); 62.6 (C6-Glc); 67.3 (CH<sub>2</sub>Ph); 69.6 (3C, CH<sub>2</sub>OGal<sub>a</sub>, -Gal<sub>b</sub>, -Glc); 70.2 (2C, C4-Gal<sub>a</sub>, C4-Gal<sub>b</sub>), 71.5 (C4-Glc), 72.5 (2C, C2-Gal<sub>a</sub>, C2-Gal<sub>b</sub>), 74.8 (2C, C3-Gal<sub>a</sub>, C3-Gal<sub>b</sub>), 75.0 (C2-Glc), 76.6 (2C, C5-Gal<sub>a</sub>, C5-Gal<sub>b</sub>), 77.8 (4C, C3-Gal<sub>a</sub>, C3-Gal<sub>b</sub>, C5-Gal<sub>a</sub>, C5-Gal<sub>b</sub>), 104.8 (C1-Glc), 105.4 (2C, C1-Gal<sub>a</sub>, C1-Gal<sub>b</sub>), 129.0, 129.5, 138.1 (6C, C<sub>6</sub>H<sub>5</sub>), 158.8 (C=O, Cbz); ESI-MS: Calcd. for C<sub>30</sub>H<sub>47</sub>NO<sub>20</sub>Na [M+Na]<sup>+</sup>: 764.26; Found m/z 764.34.

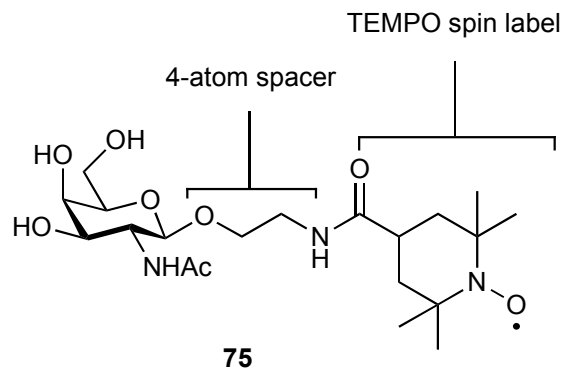
# Chapter 5: Synthesis of spin-labeled GalNAc for second-site screening by NMR

## 5.1 Introduction

Traditionally in the pharmaceutical industry, identification of high affinity ligands for therapeutic targets is done by biochemical/biological screening of vast compound libraries either from corporate collections or those generated by combinatorial chemistry, in hope of finding a “hit” - a compound that binds with a relatively good affinity to a particular target. This “hit” then becomes the lead compound, which is optimized by laborious medicinal chemistry efforts, and enters clinical trials. Thereafter, should the trials be successful, it assumes the status of a drug.

Medicinal chemistry is an ever-changing science, with a major portion of research efforts being focused on the speeding up of the drug discovery and development process. New techniques for the design, synthesis and testing of drug candidates are constantly being introduced.

The focus of this project was to synthesize a spin-labeled compound (termed first-site ligand) that binds to the known binding site on the ASGP-R H1-CRD. The ligand is intended for use in NMR screening for a second-site ligand that might bind to a different and unknown binding site on the same protein. As the first-site ligand, GalNAc (**2**, Figure 6A, p. 71) was chosen, since it is the best known ASGP-R ligand and has a  $K_D$  of 120  $\mu\text{mol}$  [95]. GalNAc was therefore labeled with a TEMPO spin label to generate compound **75** (Figure 16), and used successfully in the screening of a library of second-site candidates.



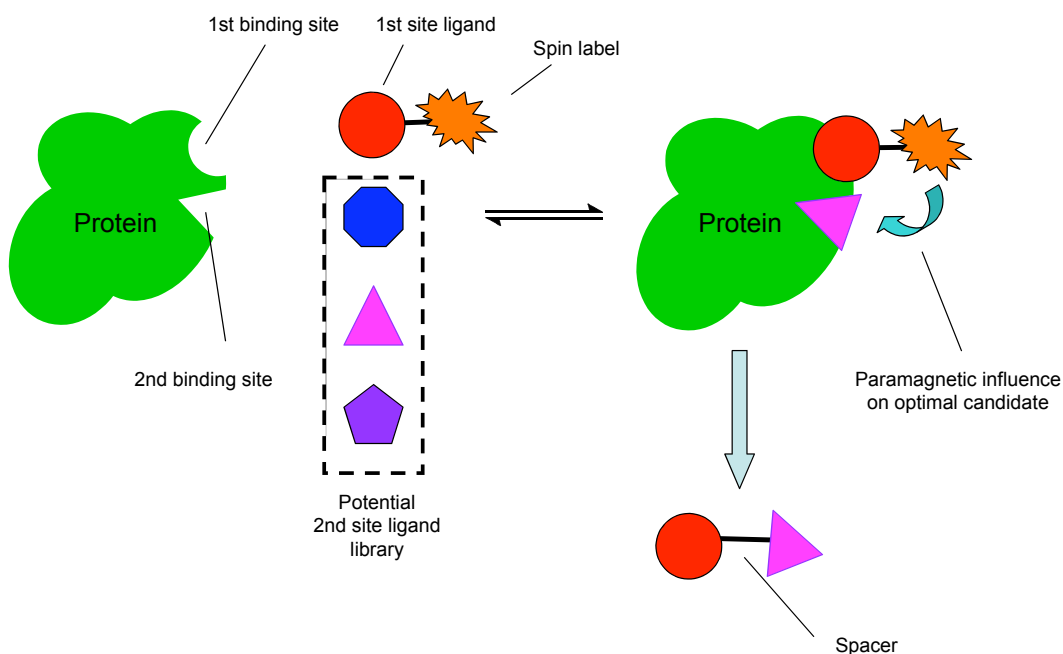
**Figure 16.** GalNAc-TEMPO adduct **75**, the first-site ligand for second-site NMR screening.

### 5.1.1 NMR in drug discovery and development

In recent years, NMR spectroscopy has become an integral part of modern drug discovery and development. A pioneering NMR-based technique called “SAR by NMR” was introduced that greatly narrowed down the number of compounds necessary for first-site screening [96]. SAR by NMR is used for finding a ligand for the first or “already known” binding site, optimizing it, finding a ligand for the second binding site, also optimizing it, and finally linking the two compounds together [97]. The second-site ligand should ideally bind simultaneously to the protein at a different binding site than that of the first ligand, but in the vicinity of it. The technique mainly relies on  $^{15}\text{N}$ -labeled protein, whose residue signals experience discernible chemical shifts in the  $^{15}\text{N}$ -HSQC spectrum upon ligand binding. By knowing the structure of the protein, and using chemical shift analysis and intermolecular protein-ligand NOE, one could then deduce not only the location of the binding site, but also the actual ligand that bound to it as well as its binding mode. However, identification of the second-site ligand requires the saturation of the first binding site with the first-site ligand in order to prevent the potential second-site ligand from also binding there, and thus resulting in a false positive hit [97]. However, saturation is often not possible due to the poor solubility of the first-site ligand.



Therefore another technique, which involves second-site NMR screening with a spin-labeled first ligand, was developed [98]. The method enables the identification of a ligand from a diverse library of compounds, which binds to a second-site on the same target protein (Figure 17). The first ligand without the spin label is then chemically linked to the second-site ligand.



**Figure 17.** Basic illustration of the second-site NMR screening process that literally takes place in the NMR tube.

### 5.1.2 Paramagnetic relaxation enhancement

The physical principle behind second-site NMR screening with a spin-labeled first ligand is called paramagnetic relaxation enhancement effect or  $R2_{para}$ . The spin-spin relaxation rates ( $R2$ ) of two resonating species, *i.e.* nuclei or electrons, are proportional to the product of the squares of their gyromagnetic ratios,  $\gamma$ . Thus,

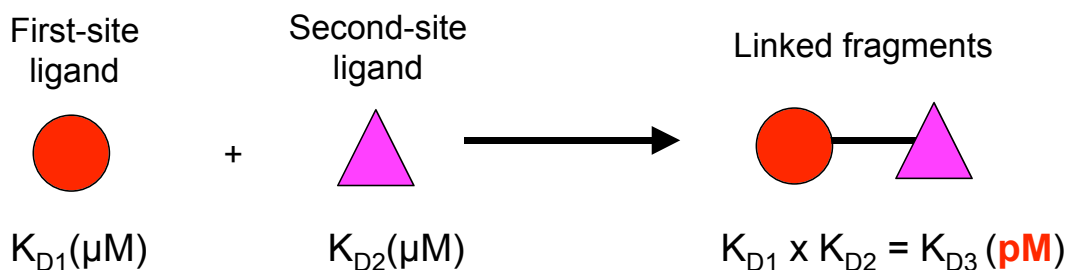
the greater the  $\gamma$  value of one or both species, the greater is the  $R_2$ . The gyromagnetic ratio of an unpaired electron is 658 times greater than that of a proton, hence a proton in the vicinity of such an electron experiences a dramatic increase in its relaxation rate ( $R_2$ ) relative to nuclear-nuclear relaxation. In the NMR spectrum, the  $R_{2\text{para}}$  effect manifests itself in line broadening and occasional small shifts of the signal for the particular proton in the vicinity of the spin label. The effect can only be detected if the first-site spin-labeled compound and the second-site ligand bind simultaneously to the protein, within a maximum distance of 20 Å from each other. Thus, the problem of binding site overlap that can occur in “SAR by NMR” [96] is eliminated.

### 5.1.3 Linking of first- and second-site fragments

The < 20 Å distance also allows the design of a reasonable length spacer for linking the two fragments. Furthermore, the protons on a second-site candidate that are the closest to the spin label experience greater paramagnetic relaxation enhancement than the one on the distant side of the ligand. This gives a clue to the ligand’s relative orientation, and provides valuable information for optimal linker design. Due to its high sensitivity, the technique also requires lower protein amounts per screen than other NMR-based techniques [99].

Once the 1<sup>st</sup> and 2<sup>nd</sup> site ligands are linked together, the association and dissociation of one of the ligands is strongly influenced by the binding event of the other ligand to the neighboring binding site. The first and second-site ligand fragments may have relatively low  $K_D$ s in the milli- to micromolar range. However, due to additive binding energies and favorable entropic terms, upon chemical linkage a compound is generated whose  $K_D$  is the product of both fragment  $K_D$ s, and can be in nano- to picomolar affinity range [100] (Figure 18). The final dissociation constant,  $K_{D3}$  is also influenced by a term that accounts for the changes in the binding affinity caused by linking, and is derived largely from

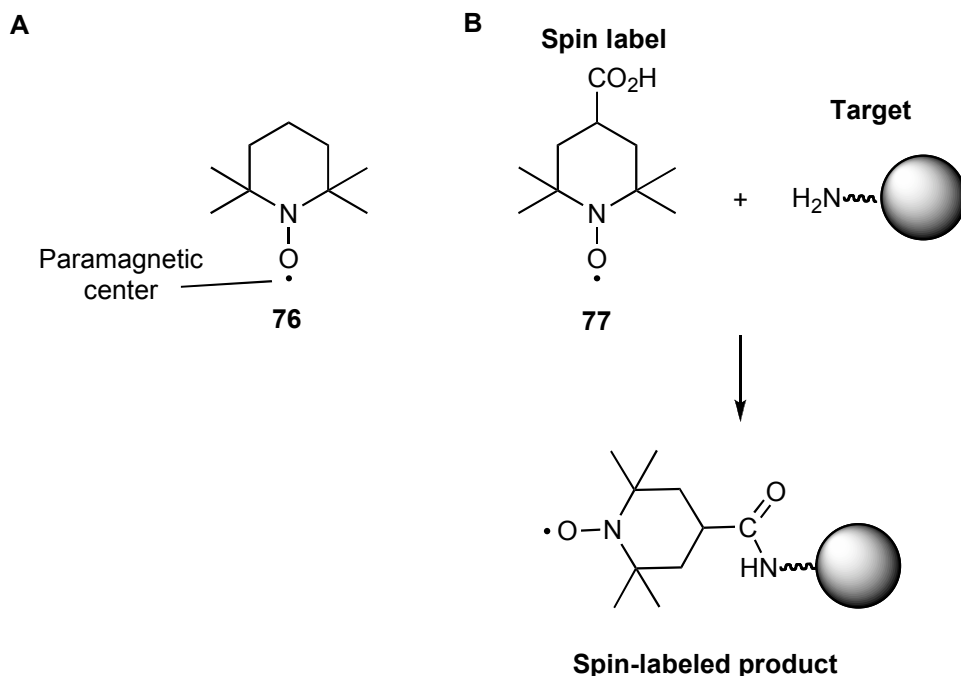
changes in translational and rigid-body rotational entropy of the linked fragments, and can make the Gibbs free energy of binding less favorable [100].



**Figure 18.** Schematic representation of the resultant  $K_D$  after chemical linking of the individual first- and second-site fragments.

#### 5.1.4 Spin Labels

NMR spin labels are generally organic nitroxides, in which the oxygen atom on the nitrogen has an unpaired electron [98]. Such compounds are termed stable radicals, and are paramagnetic by nature. Spin labels are commercially available with many additional functional groups, which allow their facile conjugation to the target protein, nucleic acid, organic molecule, etc. A widely employed spin label is 2,2,6,6-tetramethylpiperidine-*N*-oxyl or TEMPO (**76**, Figure 19A). Figure 19B depicts the 4-carboxy version of TEMPO (**77**) designed for labeling primary amines.



**Figure 19.** (A) 2,2,6,6-tetramethylpiperidine-*N*-oxyl (TEMPO, **76**). (B) A scheme illustrating the use of 4-carboxy-TEMPO (**77**) as a spin label for a target with a primary amine.

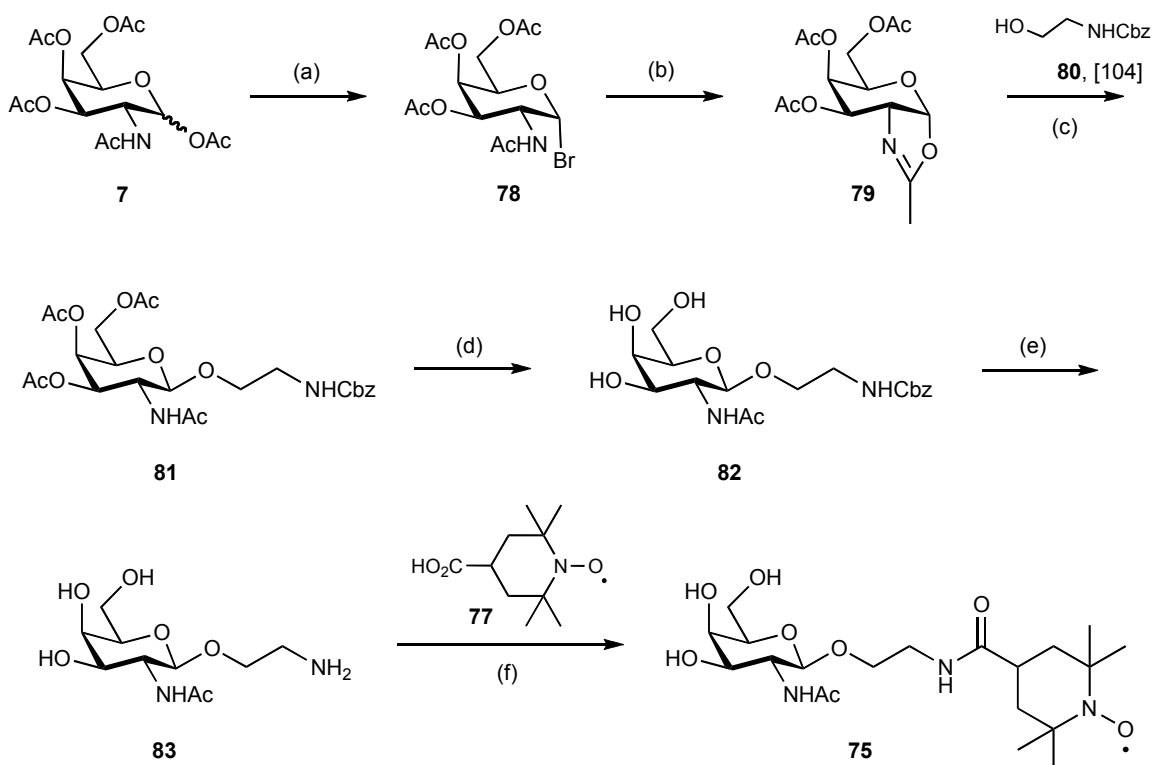
In general, TEMPO (**76**) is a very stable and chemically inert compound. It is compatible with most organic solvents as well as water, and can be used in the pH range of 3-10 [101]. The four methyl groups prevent the paramagnetic oxygen atom from nitroxide dimer formation and disproportionation by acting as steric shields [101].

## 5.2 Results and Discussion

### 5.2.1 Synthesis of compound **75**

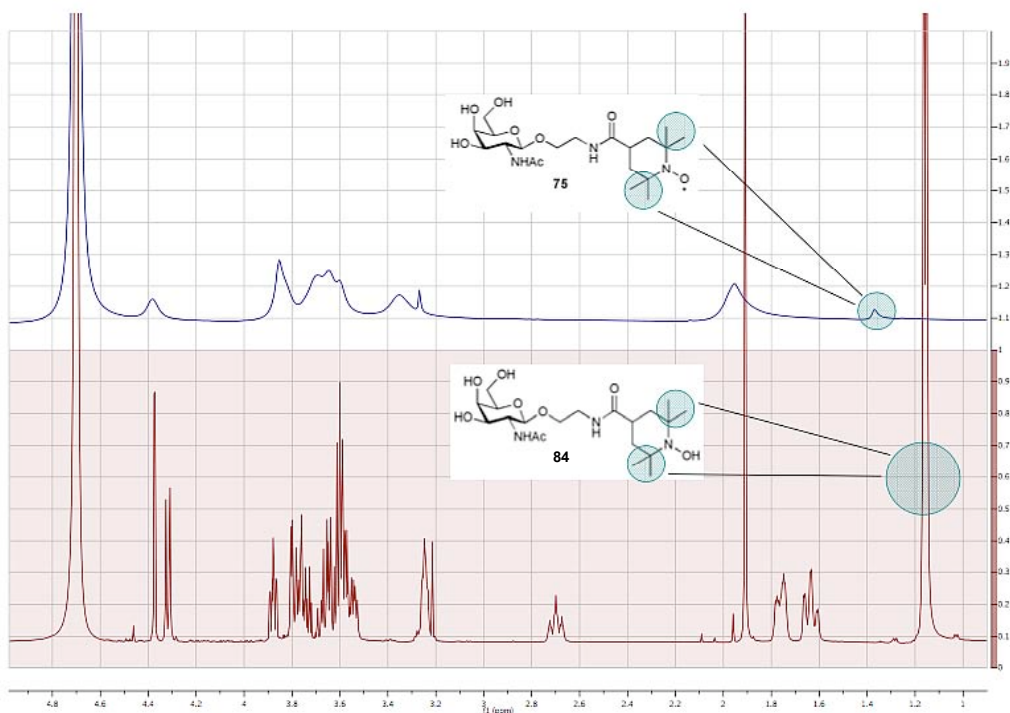
In order to retain the original affinity of GalNAc, the TEMPO spin label was attached via a 4-atom spacer to the 1-position of GalNAc (**75**, Figure 16), which is known to be less important for binding to the ASGP-R [102].

The synthesis of TEMPO-labeled GalNAc (**75**, Scheme 10) began with the formation of the  $\alpha$ -bromide of GalNAc (**78**), which involved treating **7** with HBr in acetic acid. Attempts to synthesize the GalNAc oxazoline (**79**) by the procedure of Jeanloz *et al.* [103] failed to give a product. Therefore, a new procedure was devised, which involved treating crude **78** with silver triflate in the presence of pyridine leading to the formation of the oxazoline **79** in an 89% yield starting from **7**. Compound **79** was then used as the donor for the stereospecific glycosylation of Cbz-protected ethanolamine [104] (**80**), promoted by triflic acid, yielding **81** in an 85% yield. The next two steps gave quantitative yields, and involved deacetylation under Zemplén conditions ( $\rightarrow$  **82**) and Cbz group removal by catalytic hydrogenation ( $\rightarrow$  **83**). The resultant amine (**83**) was conjugated to 4-carboxy-TEMPO (**77**) via an amide bond using DCC and HOBT in DMF to give compound **75** in a 47% yield. For the determination of the NMR spectrum, compound **75** was reduced directly in the NMR tube ( $\rightarrow$  **84**) with a slight excess of sodium ascorbate in D<sub>2</sub>O.



**Scheme 10.** (a) HBr/CH<sub>3</sub>CO<sub>2</sub>H, Ac<sub>2</sub>O, DCM, r.t., 24 h; (b) AgOTf, pyridine, CH<sub>3</sub>CN/PhMe (1:2), 4 Å MS, r.t., 24 h, 89% (2 steps); (c) TfOH, DCM, 4 Å MS, r.t., 24 h, 85%; (d) NaOMe, MeOH, r.t., 24 h, quant.; (e) H<sub>2</sub>, Pd/C, MeOH, r.t., 24 h, quant.; (f) DCC, HOBT, DMF, r.t., 24 h, 47%.

The presence of an unpaired electron in compound **75** had a dramatic effect on its NMR spectrum (Figure 20). As previously mentioned, it could also be seen that, the protons closest to the paramagnetic center were affected the most, *i.e.* the 12 protons of the four TEMPO methyl groups give a prominent signal in the spectrum of the reduced form, whereas this signal is almost completely diminished in the spectrum of the oxidized form (compound **84**, Figure 20).



**Figure 20.** (Top) NMR spectrum of the oxidized, radical form of **75** showing the paramagnetic influence on the four methyl groups adjacent to the oxygen. (Bottom) NMR spectrum of the reduced form (compound **84**).

It can also be seen in Figure 20, that the signals for the protons further away from the oxygen radical are also affected, but to a lesser extent.

### 5.2.2 Testing of compound **75** in the competitive binding assay

(Performed by Daniela Stokmaier, Institute of Molecular Pharmacy, University of Basel)

To verify that as a result of TEMPO addition, GalNAc did not lose its affinity towards the ASGP-R H1-CRD, **75** was tested with an in-house polymer assay (Appendix 1). The  $IC_{50}$  for **75** was determined as 0.17 mM, which was close to the value for GalNAc (0.12 mM) as determined in the same assay. These results

indicate that compound **75** is a suitable first-site ligand for second-site NMR screening.

### 5.2.3 Transverse relaxation rate (T1rho) measurements

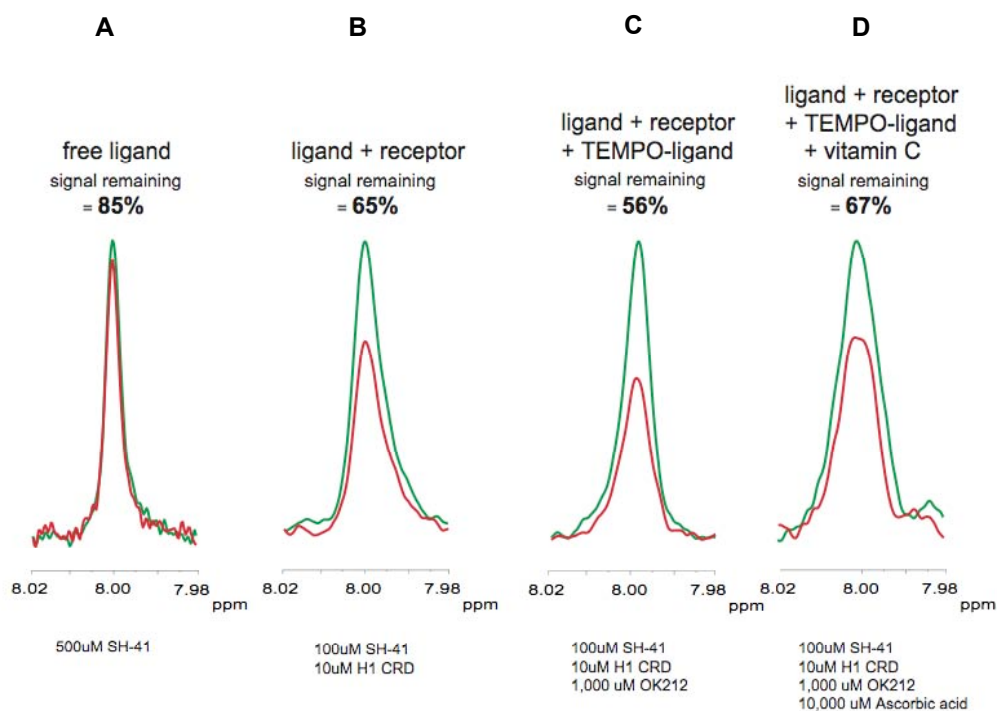
(Performed by Dr. Brian Cutting, Institute of Molecular Pharmacy, University of Basel)

The NMR measurements, to quantitate the interaction of TEMPO-GalNAc (**75**) with the H1-CRD, involved an in-house protocol (Appendix 3) which measures transverse relaxation rates (T1rho), but is less sensitive to magnetic field inhomogeneity, which may be caused by switching from the free ligand to the ligand with the protein. In spite of being a primarily qualitative assay, T1rho measurements can give an accurate indication of protein-ligand binding. Since the rate of transverse relaxation increases with molecular weight, T1rho can be used as a parameter to investigate ligand binding because the protein-ligand complex has a greater molecular weight than the free ligand [105]. In effect, this can be seen as a decrease in specific signal intensities in the NMR spectrum.

Thus, Figure 21 shows the typical changes in the signal pattern of a ligand (from a mixture of potential second-site compounds) that binds to a second-site. Part A shows the signal of a proton of the free ligand in solution. Part B shows the decrease in its signal intensity upon addition of the ASGP-R. As the ligand binds to the protein, the overall mass of the ligand-receptor complex is greater than that of the free ligand, which leads to faster T1rho relaxation times with concomitant signal intensity loss, so only 65% of the original signal remains. Upon addition of the spin-labeled compound **75** (C), the signal experiences paramagnetic relaxation enhancement and diminishes further in intensity to 56%.



Upon addition of sodium ascorbate, which reduces TEMPO, the signal regains its intensity to the same level as it had in the presence of only the protein.



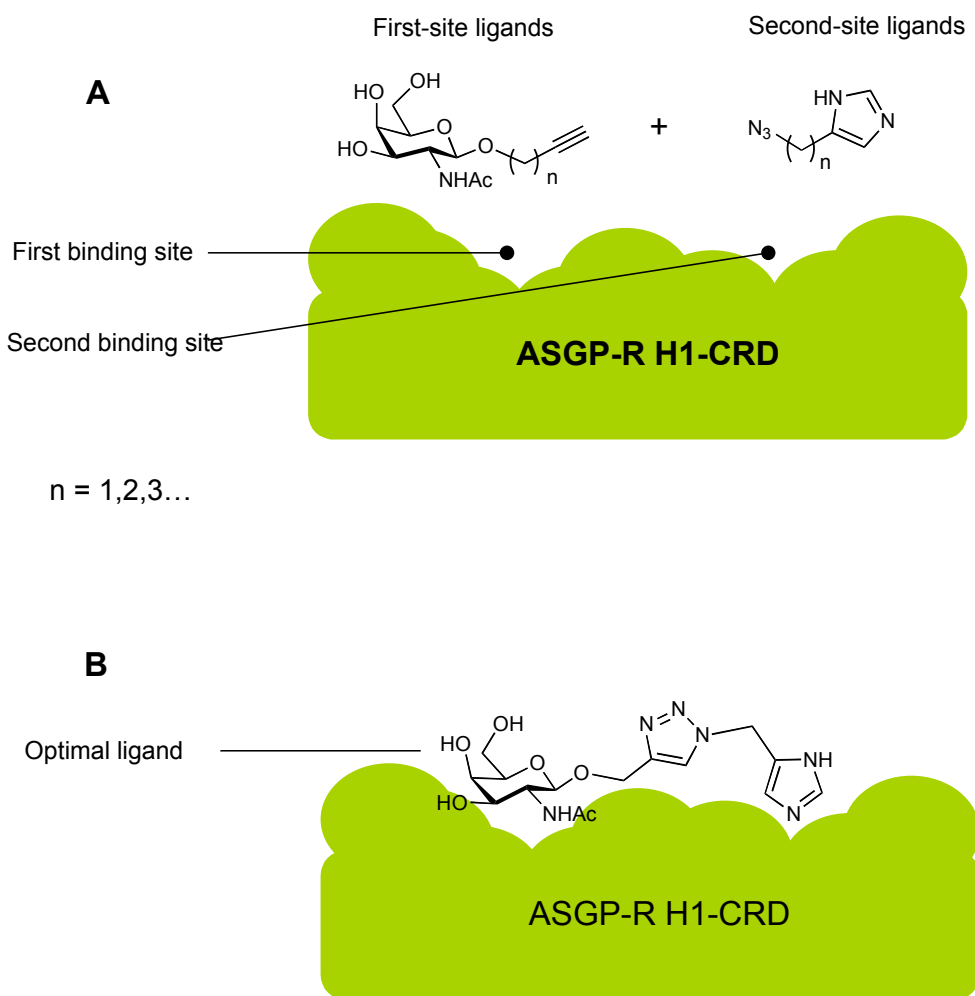
**Figure 21.** A typical NMR signal pattern evinced by a potential 2<sup>nd</sup> site candidate (SH-41 = imidazole; OK 212 = compound **75**).

This change in the second-site ligand's NMR peak pattern indicates that it does indeed bind simultaneously with first-site ligand, but to a different binding site on the ASGP-R.

After deconvolution of the library, the ligand was identified as imidazole. The next step would involve the linking of the imidazole to GalNAc via an optimal linker selected by *in situ* click chemistry [106].

## 5.3 Conclusion

Compound **75** was successfully synthesized and tested by two independent assays, which confirmed that it can indeed be used for second-site screening by NMR. The screening itself identified imidazole as a possible second-site ligand, which should be linked to GalNAc and the resulting adduct tested for its affinity towards the ASGP-R. The linking procedure is termed *in situ* click chemistry [106], and is schematically outlined in Figure 22.



**Figure 22.** An illustration of an *in situ* click chemistry experiment.

The procedure would involve combining homologous series of the first- and second-site ligands featuring alkyne and azide functional groups, respectively, in the presence of the ASGP-R H1-CRD. The protein is intended to act as a template by bringing the appropriate length azide and alkyne into close proximity and in a favorable orientation to react via a Huisgen 1,3-dipolar cycloaddition, thus linking the two components by an optimal length linker. The strength of the approach lies in the fact that the best possible compound is selected from a mixture of several possible compounds by the protein itself, therefore eliminating the need to synthesize a series of compounds with different spacer lengths and testing them individually in a biological assay.

## 5.4 Experimental

General methods are described in Appendix 4, p. 147.

### **2-Methyl-4,5-(3,4,6-tri-O-acetyl-2-deoxy- $\alpha$ -D-galactopyrano)- $\Delta^2$ -oxazoline**

**(79)**: Compound **7** (851 mg, 2.18 mmol) was dissolved in dry DCM (18 ml) in an amber glass flask, and acetic anhydride (100  $\mu$ l) was added. To the resultant solution, HBr in acetic acid (3.4 ml, 19.7 mmol) was added, and the solution was stirred at r.t. for 24 h. The mixture was diluted with DCM (30 ml) and washed with ice cold water (100 ml), followed by saturated NaHCO<sub>3</sub> solution (100 ml) and brine (100 ml), dried (Na<sub>2</sub>SO<sub>4</sub>) and the solvent was removed in vacuo to give a foam containing mainly compound **78**, which was used without further purification.

Compound **78** (895 mg, 2.18 mmol) was dissolved in dry CH<sub>3</sub>CN/toluene (6 ml, 1:2) in an amber glass flask, 4 Å mol sieves (300 mg) were added, and the resultant suspension was stirred at r.t. under argon for 2 h. Dry pyridine (260  $\mu$ l, 3.27 mmol) was then added, followed by AgOTf (840 mg, 3.27 mmol), and the mixture was stirred at r.t. under argon for 24 h. The mixture was then filtered through Celite, and the solvents were removed in vacuo. The residue was purified by flash chromatography on silica gel (toluene/EtOAc/Et<sub>3</sub>N 100:200:1), yielding **79** (640 mg, 89%, R<sub>f</sub> 0.22 C). The analytical data were identical with literature values [107].

**(2-Benzyloxycarbonylamino)ethyl 2-acetamido-3,4,6-tri-O-acetyl-2-deoxy- $\beta$ -D-galactopyranoside (81)**: Compound **79** (240 mg, 729  $\mu$ mol) and **80** [104] (170 mg, 874  $\mu$ mol) were dissolved in dry DCM (3 mL) and stirred with activated mol. sieves (4 Å, 200 mg) for 2 h at r.t. under argon. Triflic acid (64  $\mu$ L, 729  $\mu$ mol) was then added, and the resultant mixture was stirred for 24 h at r.t. under argon. The reaction was quenched with triethylamine (100  $\mu$ L), filtered through Celite, and

the solvent was removed in vacuo. The residue was purified by flash chromatography on silica gel (petrol ether/EtOAc 1:9 → 0:1), yielding **81** (325 mg, 85%,  $R_f$  0.24 C).  $[\alpha]_D = -9.5$  (c 1, MeOH);  $^1\text{H-NMR}$  (500 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  1.86, 1.90, 1.97, 2.07 (s, 12H, 4 ×  $\text{CH}_3\text{CO}$ ), 3.27 (m, 2H,  $\text{OCH}_a\text{H}_b\text{CH}_2\text{N}$ ), 3.59 (m, 1H,  $\text{OCH}_a\text{H}_b\text{CH}_2\text{N}$ ), 3.81 (m, 1H,  $\text{OCH}_a\text{H}_b\text{CH}_2\text{N}$ ), 3.97 (m, 1H, H5), 4.04-4.13 (m, 3H, H2, H6), 4.55 (d,  $J = 8.4$  Hz, 1H, H1), 5.01-5.07 (m, 3H, H3,  $\text{CH}_2\text{Ph}$ ), 5.30 (d,  $J = 3.0$  Hz, 1H, H4), 7.25-7.31 (m, 5H,  $\text{C}_6\text{H}_5$ );  $^{13}\text{C NMR}$  (125 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  20.5, 20.6, 22.9 (4C, 4 ×  $\text{CH}_3\text{CO}$ ), 41.7 ( $\text{OCH}_2\text{CH}_2\text{N}$ ), 51.4 (C2), 62.7 (C6), 67.4 ( $\text{CH}_2\text{Ph}$ ), 68.1 (C4), 69.6 ( $\text{OCH}_2\text{CH}_2\text{N}$ ), 71.7 (C3), 71.9 (C5), 102.6 (C1), 128.8, 128.9, 129.4, 138.1 (6C,  $\text{C}_6\text{H}_5$ ), 158.6 (C=O, Cbz), 171.6, 172.0, 173.8 (4C, 4 × C=O, Ac); Anal. calcd. for  $\text{C}_{24}\text{H}_{32}\text{N}_2\text{O}_{11}$ : C, 54.96; H, 6.15; N, 5.34. Found: C, 54.65; H, 6.09; N, 5.15.

**(2-Benzyloxycarbonylamino)ethyl 2-acetamido-2-deoxy- $\beta$ -D-galactopyranoside (82)**: Compound **81** (180 mg, 343  $\mu\text{mol}$ ) was dissolved in dry MeOH (5 mL) and sodium (30 mg) was added. The resultant mixture was stirred for 24 h at r.t. under argon. The solvent was removed in vacuo, and the residue was purified on a reverse phase RP-C18 column ( $\text{H}_2\text{O}/\text{MeOH}$  1:0 → 4:1, stepwise gradient), yielding **82** (136 mg, quantitative,  $R_f$  0.47 G).  $[\alpha]_D = -3$  (c 2, MeOH);  $^1\text{H-NMR}$  (500 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  1.95 (s, 3H,  $\text{CH}_3\text{CO}$ ); 3.20 (m, 2H,  $\text{OCH}_a\text{H}_b\text{CH}_2\text{N}$ ); 3.41 (t,  $J = 6.0$  Hz, 1H, H5); 3.47-3.53 (m, 2H,  $\text{OCH}_a\text{H}_b\text{CH}_2\text{N}$ , H3); 3.65 (m, 2H, H6); 3.73-3.81 (m, 2H,  $\text{OCH}_a\text{H}_b\text{CH}_2\text{N}$ , H4); 3.88 (m, 1H, H2); 4.28 (d,  $J = 8.4$  Hz, 1H, H1); 4.97 (s, 2H,  $\text{CH}_2\text{Ph}$ ); 7.18-7.25 (m, 5H,  $\text{C}_6\text{H}_5$ );  $^{13}\text{C NMR}$  (125 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  22.3 ( $\text{CH}_3\text{CO}$ ), 42.0 ( $\text{OCH}_2\text{CH}_2\text{N}$ ), 54.9 (C2), 62.5 (C6), 67.5 ( $\text{CH}_2\text{Ph}$ ), 69.6 (2C,  $\text{OCH}_2\text{CH}_2\text{N}$ , C4), 73.0 (C3), 76.8 (C5), 102.9 (C1), 128.9, 129.1, 129.5, 138.3 (6C,  $\text{C}_6\text{H}_5$ ), 158.8 (C=O, Cbz), 175.5 (C=O, Ac); ESI-MS: Calcd. for  $\text{C}_{18}\text{H}_{27}\text{N}_2\text{O}_8$   $[\text{M}+\text{H}]^+$ : 399.18; Found  $m/z$  399.15.

**(2-Amino)ethyl 2-acetamido-2-deoxy- $\beta$ -D-galactopyranoside (83)**: Compound **82** (72.1 mg, 181  $\mu\text{mol}$ ) was dissolved in dry MeOH (2.5 mL) and Pd/C

(10% Pd, 24 mg) was added. The mixture was vigorously stirred under a H<sub>2</sub> atmosphere (1 atm) at r.t. for 24 h. The resultant mixture was filtered through Celite, and the solvent was removed in vacuo yielding **83** (47 mg, quantitative), which was used without further purification.  $[\alpha]_D = -5.3$  (c 1, MeOH); <sup>1</sup>H-NMR (500 MHz, D<sub>2</sub>O):  $\delta$  2.01 (s, 3H, CH<sub>3</sub>CO), 2.75-2.83 (m, 2H, OCH<sub>a</sub>H<sub>b</sub>CH<sub>2</sub>N), 3.60-3.66 (m, 2H, OCH<sub>a</sub>H<sub>b</sub>CH<sub>2</sub>N, H5), 3.68-3.79 (m, 3H, H3, H6), 3.85-3.90 (m, 3H, OCH<sub>a</sub>H<sub>b</sub>CH<sub>2</sub>N, H2, H4), 4.42 (d,  $J = 8.4$  Hz, 1H, H1); <sup>13</sup>C NMR (125 MHz, D<sub>2</sub>O):  $\delta$  22.2 (CH<sub>3</sub>CO), 40.0 (OCH<sub>2</sub>CH<sub>2</sub>N), 52.4 (C2), 60.9 (C6), 67.7 (C4), 70.8 (C3), 71.0 (OCH<sub>2</sub>CH<sub>2</sub>N), 75.0 (C5), 101.8 (C1), 174.8 (C=O); ESI-MS: Calcd. for C<sub>10</sub>H<sub>21</sub>N<sub>2</sub>O<sub>6</sub> [M+H]<sup>+</sup>: 265.14; Found m/z 265.02.

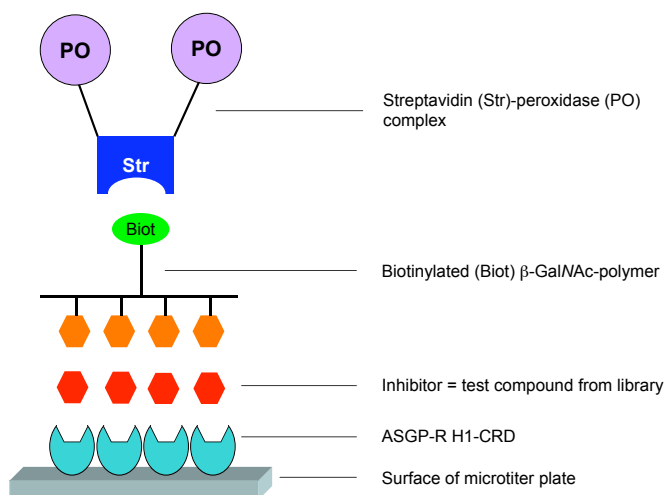
**2-(2,2,6,6-Tetramethyl-N-oxylpiperidyl-4-carbonyl)aminoethyl 2-acetamido-2-deoxy- $\beta$ -D-galactopyranoside (75):** Compound **83** (32.7 mg, 124  $\mu$ mol), 4-carboxy TEMPO (**77**, 25 mg, 124  $\mu$ mol), DCC (56.2 mg, 273  $\mu$ mol) and HOBT (19 mg, 124  $\mu$ mol) were dissolved in dry DMF (2 mL) and the mixture was stirred under argon at r.t. in the dark for 24 h. The solvent was then removed in vacuo and the residue purified on a reverse phase RP-C18 column (H<sub>2</sub>O/MeOH 1:0  $\rightarrow$  9:1, stepwise gradient), yielding **75** (26 mg, 47%, R<sub>f</sub> 0.32 G). For NMR measurements, sodium ascorbate (10 mg) was added to the sample dissolved in D<sub>2</sub>O in the NMR tube. Compound **84**:  $[\alpha]_D = -5.4$  (c 1, MeOH); <sup>1</sup>H-NMR (500 MHz, D<sub>2</sub>O):  $\delta$  1.15 (s, 6H, CH<sub>3</sub>, TEMPO), 1.16 (s, 6H, CH<sub>3</sub>, TEMPO), 1.64 (m, 2H, CH<sub>2</sub>, TEMPO), 1.76 (m, 2H, CH<sub>2</sub>, TEMPO), 1.91 (s, 3H, CH<sub>3</sub>CO), 2.69 (m, 1H, CH, TEMPO), 3.25 (m, 2H, OCH<sub>a</sub>H<sub>b</sub>CH<sub>2</sub>N), 3.53-3.69 (m, 5H, OCH<sub>a</sub>H<sub>b</sub>CH<sub>2</sub>N, H3, H5, H6), 3.72-3.81 (m, 3H, OCH<sub>a</sub>H<sub>b</sub>CH<sub>2</sub>N, H2, H4), 4.31 (d,  $J = 8.5$  Hz, 1H, H1); <sup>13</sup>C NMR (125 MHz, D<sub>2</sub>O):  $\delta$  19.2 (2C, 2  $\times$  CH<sub>3</sub>, TEMPO), 22.3 (CH<sub>3</sub>CO), 28.8 (2C, 2  $\times$  CH<sub>3</sub>, TEMPO), 34.9 (CH, TEMPO), 39.3 (OCH<sub>2</sub>CH<sub>2</sub>N), 40.0 (2C, 2  $\times$  CH<sub>2</sub>, TEMPO), 52.3 (C2), 60.9 (C6), 62.5 (C<sub>Q</sub>, TEMPO), 67.7 (C4), 68.0 (OCH<sub>2</sub>CH<sub>2</sub>N), 70.9 (C3), 75.1 (C5), 101.5 (C1), 174.6 (C=O, Ac), 177.0 (C=O, TEMPO); ESI-MS: Calcd. for C<sub>20</sub>H<sub>37</sub>N<sub>3</sub>O<sub>8</sub> [M+H]<sup>+</sup>: 448.27; Found m/z 448.26.

# Appendix 1

## Polymer assay

(Performed by Daniela Stokmaier, Institute of Molecular Pharmacy, University of Basel)

The competitive target-based assay, or polymer assay, was developed by D. Stokmaier [108] at our institute, and is illustrated in Figure 23. The basic outline of the assay is as follows: flat-bottom 96-well microtiter plates are coated with recombinant human H1-CRD. Then the library compound (inhibitor) is added, directly followed by the addition of the preformed conjugate-complex of biotinylated GalNAc-PAA and streptavidin-peroxidase. The plates are incubated for 2 h at r.t. in a humid chamber on a laboratory shaker. Then ABTS-substrate is then added to each well. The colour is allowed to develop for 2 min, after which the reaction is quenched with 2% aqueous oxalic acid. Bound GalNAc-PAA-complex is measured by determining the optical density of the formed blue-green soluble product at  $\lambda = 415 \text{ nm}$ . The  $IC_{50}$  values for the tested compounds are calculated with the help of GraphPad Prism 4 software.



**Figure 23.** Outline of the competitive binding assay.

## Appendix 2

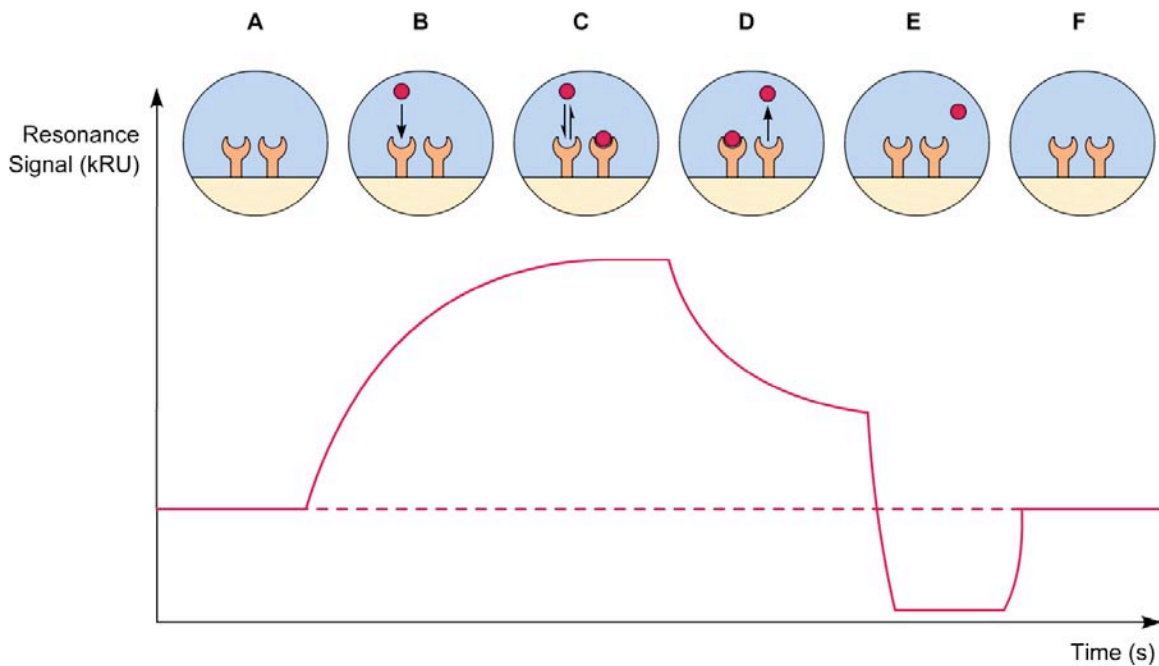
### BIACORE

(Performed by Daniel Ricklin, Institute of Molecular Pharmacy, University of Basel)

BIACORE is a bioanalytical technique that has become an integral part of the modern drug discovery process [109]. It is based on the surface plasmon resonance (SPR) effect, and enables real time and label-free monitoring of molecular interactions, *i.e.* between proteins and their ligands. BIACORE technology not only measures the equilibrium dissociation constants ( $K_D$ s) but can also be used to obtain kinetic data for a particular interaction [110]. During a BIACORE experiment, a sensogram showing response vs. time is obtained (Figure 24).

When performing measurements with compounds presented in this thesis, H1-CRD was immobilized on a sensor chip, and the ligand (analyte) was injected in solution. Ligand binding to the protein caused an increase in mass on the chip, which in turn led to changes in the refractive index of the chip's surface. The changes in the refractive index were detected and plotted as a sensogram.





**Figure 24.** A typical BIACORE sensogram. Steps A-F represent the various phases of the measurement. (Picture courtesy of A. Vöggtli)

In Figure 24, (A) represents the baseline, when only the protein is immobilized on the chip. Analyte injection leads to the association phase (B) when the ligand begins to bind to the protein. A steady state (C) occurs when the interaction reaches equilibrium, which is then followed by the dissociation phase (D). After a regeneration phase (E), the response returns to the baseline level (F).

## Appendix 3

### Second-site screening by NMR

(Performed by Dr. Brian Cutting, Institute of Molecular Pharmacy, University of Basel)

During the actual second-site NMR screening experiment, the TEMPO-labeled first-site ligand, the protein and a set of potential second-site ligands are subjected to a 90° pulse with a spinlock for a variable duration (10-200 ms). If the first- and second-site compounds bind simultaneously and in the vicinity of each other to the protein, then the second-site ligand will experience a paramagnetic relaxation enhancement. The  $R_{2\text{para}}$  effect is not observed after ascorbic acid, which reduces TEMPO, is added. In a control experiment, when both ligands are analyzed in the absence of the protein, no  $R_{2\text{para}}$  effect is observed.

The protein concentrations were in 20  $\mu\text{M}$  range, while a GalNAc-TEMPO (compound **75**) concentration of approximately 200  $\mu\text{M}$  was used. For all  $T_{1\rho}$  experiments, the buffer consisted of 20 mM Tris-d11 (pH 7), 1 mM  $\text{CaCl}_2$  in 99.5%  $\text{D}_2\text{O}$ .

To estimate the transverse relaxation rate, five decay times, beginning from 10 ms and ending at 200 ms, were measured. All  $T_{1\rho}$  measurements were performed on a 500 MHz Bruker DRX NMR spectrometer. The signal loss with increasing relaxation times was quantified and analyzed with the Prism curve-fitting software package from GraphPad Software, Inc.

## Appendix 4

### General Methods.

NMR spectra were recorded on a Bruker Avance DMX-500 (500 MHz) spectrometer. Assignment of  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra was achieved using 2D methods (COSY, HSQC, TOCSY). Chemical shifts are expressed in ppm using residual  $\text{CHCl}_3$ ,  $\text{CHD}_2\text{OD}$  and HDO as references. Optical rotations were measured using a Perkin-Elmer Polarimeter Model 341. ESI-MS spectra were measured on a Waters Micromass ZQ mass spectrometer. Reactions were monitored by TLC using glass plates coated with silica gel 60  $F_{254}$  (Merck) with the following mobile phases: A) toluene/EtOAc 4:1; B) toluene/EtOAc 1:4; C) EtOAc; D) petrol ether/EtOAc 1:1; E) EtOAc/MeOH 9:1; F) petrol ether/EtOAc 9:1; G) DCM/MeOH 5:2; H) DCM/MeOH 9:1; I) DCM/MeOH 3:2; J) DCM/MeOH 2:1; K) petrol ether/EtOAc/ $\text{Et}_3\text{N}$  1:1:0.1; L) petrol ether/EtOAc 7:3; M) DCM/acetone 9:1; N) DCM/MeOH/ $\text{H}_2\text{O}$  6:4:1. Carbohydrate-containing compounds were visualized by charring with a molybdate solution (0.02 M solution of ammonium cerium sulfate dihydrate and ammonium molybdate tetrahydrate in aqueous 10%  $\text{H}_2\text{SO}_4$ ). All other compounds were visualized with  $\text{KMnO}_4$  solution (2%  $\text{KMnO}_4$  and 4%  $\text{NaHCO}_3$  in water). Column chromatography was performed on silica gel 60 (Fluka, 0.040-0.060 mm). Methanol (MeOH) was dried by refluxing with sodium methoxide and distilled immediately before use. Pyridine was freshly distilled under argon over  $\text{CaH}_2$ . Dichloromethane (DCM) and dichloroethane (DCE) were dried by filtration over  $\text{Al}_2\text{O}_3$  (Fluka, type 5016 A basic). Tetrahydrofuran (THF), dioxane, diethyl ether ( $\text{Et}_2\text{O}$ ) and toluene were dried by refluxing with sodium and benzophenone. Dry DMF was purchased from Fluka (absolute,  $\geq 99.8\%$ ) and was further dried over powdered 4 Å molecular sieves. Molecular sieves (4 Å) were activated in vacuo at 500 °C for 2 h immediately before use.

## References

- [1] M. Spiess, *Biochemistry*, **1990**, 29, 10009.
- [2] G. Ashwell, A. G. Morell, *Adv. Enzymol. Relat. Areas Mol. Biol.* **1974**, 41, 99.
- [3] I. Geffen, M. Spiess, *Int. Rev. Cytol.* **1992**, 137B, 181.
- [4] J. Wu, M. H. Nantz, M. A. Zern, *Frontiers in Bioscience* **2002**, 7, d717-725.
- [5] J. U. Baenzinger, D. Fiete, *Cell* **1980**, 22, 611.
- [6] A. L. Schwartz, D. Rup, H. F. Lodish, *J. Biol. Chem.* **1980**, 255, 9033.
- [7] M. Meier, M.D. Bider, V.N. Malashkevich, M. Spiess, P. Burkhard, *J. Mol. Biol.* **2000**, 300, 857.
- [8] K. Drickamer, *J. Biol. Chem.* **1988**, 263, 9557.
- [9] D. C. Kilpatrick, *Biochim. Biophys. Acta* **2002**, 1572, 187.
- [10] A. N. Zelensky, J. E. Gready, *FEBS J.* **2005**, 272, 6179.
- [11] J. U. Baenziger, Y. Maynard, *J. Biol. Chem.* **1980**, 255, 4607.
- [12] P. C. N. Rensen, L. A. J. M. Sliedregt, M. Ferns, E. Kieviet, S. M. W. van Rossenberg, S. H. van Leeuwen, T. J. C van Berkel, E. A. L. Biessen, *J. Biol. Chem.* **2001**, 276, 37577.
- [13] M. McPhaul, P. Berg, *Proc. Natl. Acad. Sci U.S.A.* **1986**, 83, 8863.
- [14] M. R. Hardy, R. R. Townsend, S. M. Parkhurst, Y. C. Lee, *Biochemistry* **1985**, 24, 22.
- [15] A. L. Schwartz, C. J. Steer, E. S. Kempner, *J. Biol. Chem.* **1984**, 259, 12025.
- [16] M. D. Bider, J. M. Wahlberg, R. A. Kammerer, M. Spiess, *J. Biol. Chem.* **1996**, 271, 31996.
- [17] P. H. Weigel, J. H. N. Yik, *Biochim. Biophys. Acta* **2002**, 1572, 341.
- [18] J. S. Marshall, A. M. Green, J. Pensky, S. Williams, A. Zinn, D. M. Carlson, *J. Clin. Invest.* **1974**, 54, 555.

- [19] V. I. Zannis, J. McPherson, G. Goldberger, S. K. Karathanasis, J. L. Breslow, *J. Biol. Chem.* **1984**, *259*, 5495.
- [20] W. R. Brown, T. M. Koppel, *Immunol. Invest.* **1989**, *18*, 269.
- [21] D. Wright, L. Usher, *Curr. Org. Chem.* **2001**, *5*, 1107.
- [22] M. Mammen, S.-K. Choi, G. M. Whitesides, *Angew. Chem. Int. Ed.* **1998**, *37*, 2754.
- [23] Y. C. Lee, R. R. Townsend, M. R. Hardy, J. Lönngren, J. Arnarp, M. Haraldsson, H. Lönn, *J. Biol. Chem.* **1983**, *258*, 199.
- [24] J. J. Lundquist, E. J. Toone, *Chem. Rev.* **2002**, *102*, 555.
- [25] R. R. Townsend, M. R. Hardy, T. C. Wong, Y. C. Lee, *Biochemistry* **1986**, *25*, 5716.
- [26] K. Bock, J. Arnarp, J. Lönngren, *Eur. J. Biochem.* **1982**, *129*, 171.
- [27] J. E. Gestwicki, L. E. Strong, L. L. Kiessling, *Angew. Chem. Int. Ed.* **2000**, *39*, 4567.
- [28] W. P. Jenks, *Adv. Enzymol.* **1975**, *43*, 219.
- [29] M. I. Page, W. P. Jenks, *Proc. Nat. Acad. Sci. U.S.A.* **1971**, *68*, 1678; M. Mammen, E. I. Shakhnovich, G. M. Whitesides, *J. Org. Chem.* **1998**, *63*, 3168.
- [30] R. G. W. Anderson, J. L. Goldstein, M. S. Brown, *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 2434.
- [31] I. Mellman, R. Fuchs, A. Hellenius, *Annu. Rev. Biochem.* **1986**, *55*, 663.
- [32] J. E. Zijderhand-Bleekemolen, A. L. Schwartz, J. W. Slot, G. J. Strous, H. J. Geuze, *J. Cell Biol.* **1987**, *104*, 1647.
- [33] J. E. Rothman, S. L. Schmid, *Cell*, **1986**, *46*, 5.
- [34] J. L. Goldstein, M. S. Brown, R. G. W. Anderson, D. W. Russel, W. J. Schneider, *Annu. Rev. Cell Biol.* **1985**, *1*, 1.
- [35] P. P. Breitfeld, C. F. Simmons, G. J. A. M. Strous, H. J. Geuze, A. L. Schwartz, *Int. Rev. Cytol.* **1985**, *97*, 47.
- [36] C. F. Simmons, A. L. Schwartz, *Mol. Pharmacol.* **1984**, *26*, 509.
- [37] A. L. Schwartz, S. E. Fridovich, H. F. Lodish, *J. Biol. Chem.* **1982**, *257*, 4230.

- [38] A. L. Schwartz, S. E. Fridovich, B. B. Knowles, H. F. Lodish, *J. Biol. Chem.* **1981**, 256, 8878.
- [39] P. Weigel, *J. Biol. Chem.* **1980**, 255, 6111; 6.
- [40] C. J. Steer, G. Ashwell, *J. Biol. Chem.* **1980**, 255, 3008; 5.
- [41] G. Y. Wu, C. H. Wu, *J. Biol. Chem.* **1988**, 263, 14621.
- [42] D. B. Cawley, D. L. Simpson, H. R. Herschman, *Proc. Natl. Acad. Sci. U. S. A.* **1981**, 78, 3383.
- [43] T.-M. Chang, D. W. Kullberg, *J. Biol. Chem.* **1982**, 257, 12563.
- [44] R. L. A. de Vruueh, E. T. Rump, E. van de Bilt, R. van Veghel, J. Balzarini, E. A. L. Biessen, T. J. C. van Berkel, M. K. Bijsterbosch, *Antimicrob. Agents Chemother.* **2000**, 44, 477.
- [45] R. L. De Vruueh, E. T. Rump, L. A. Sliedregt, E. A. Biessen, T. J. Van Berkel, M. K. Bijsterbosch. *Pharm. Res.* **1999**, 16, 1179.
- [46] U. Treichel, K.-H. Meyer zum Büschenfelede, R. J. Stockert, T. Poralla, G. Gerken, *J. Gen. Virol.* **1994**, 75, 3021.
- [47] G. Diedrich, *FEBS J.* **2006**, 273, 3871.
- [48] J. Liu, Y. H. Li, C. F. Xue, *World J. Gastroenterol.* **2003**, 9, 295.
- [49] J. Y. Lau, T. L. Wright, *Lancet* **1993**, 342, 1335.
- [50] G. K. Lau, T. Piratvisuth, K. X. Luo, P. Marcellin, S. Thongsawat, G. Cooksley, E. Gane, M. W. Fried, W. C. Chow, S. W. Paik, W. Y. Chang, T. Berg, R. Flisiak, P. McCloud, N. Pluck, *N. Eng. J. Med.* **2005**, 352, 2682.
- [51] K. H. Heermann, U. Goldmann, W. Schwartz, T. Seyffahrt, H. Baumgarten, W. H. Gerlich *J. Virol.* **1984**, 52, 396.
- [52] C. C. Yu Ip, W. J. Miller, D. J. Kubek *Biochemistry* **1992**, 31, 285.
- [53] U. Treichel, K.-H. Meyer zum Büschenfelede, H.-P. Dienes, G. Gerken, *Arch. Virol.* **1997**, 142, 493.
- [54] Anonymous Consensus Statement, *J. Hepatol.* **1999**, 31, 3.
- [55] J. H. Hoofnagle, *Hepatology* **1997**, 26, 15S.
- [56] G. M. Lauer, B. D. Walker, *N. Engl. J. Med.* **2001**, 345, 41.
- [57] B. Saunier, M. Triyatni, L. Ulianich, P. Maruvada, P. Yen, L. D. Kohn, *J. Virol.* **2003**, 77, 546.

- [58] W. Stoorvogel, H. G. Geuze, G. J. Strous, *J. Cell Biol.* **1987**, *104*, 1261.
- [59] M. P. Manns, J. G. McHutchinson, C. S. Gordon, V. K. Rustgi, M. Shiffman, R. Reindollar, Z. D. Goodman, K. Koury, M. Ling, J. K. Albrecht, *Lancet*, **2001**, *358*, 958.
- [60] O. Khorev, D. Stokmaier, O. Schwardt, B. Cutting, B. Ernst, *Bioorg. Med. Chem.* **2007**, article in press.
- [61] R. T. Lee, *Biochemistry*, **1982**, *21*, 1045.
- [62] S. T. Iobst, K. Drickamer, *J. Biol. Chem.* **1996**, *271*, 6686.
- [63] T. C. Wong, R. R. Townsend, Y. C. Lee, *Carbohydr. Res.* **1987**, *170*, 27.
- [64] Y. C. Lee, R. T. Lee, K. Rice, Y. Ichikawa, T.-C. Wong, *Pure Appl. Chem.* **1991**, *63*, 499.
- [65] D. Yi, R. T. Lee, P. Longo, E. T. Boger, Y. C. Lee, W. A. Petri Jr., R. L. Schnaar, *Glycobiology* **1998**, *8*, 1037.
- [66] Performed by Dr. M. Lill, Biograph Laboratories, Basel.
- [67] A. R. Kolatkar, A. K. Leung, R. Isecke, R. Brossmer, K. Drickamer, W. I. Weis, *J. Biol. Chem.* **1998**, *273*, 19502-19508.
- [68] I. Fleming, *Pericyclic Reactions*, Oxford Chemistry Primers, Oxford University Press Inc., New York, **1999**, 1.
- [69] R. Huisgen in *1,3-Dipolar Cycloaddition Chemistry*, Ed. : A. Padwa, **1984**, Vol 1, p.1, Wiley, New York.
- [70] C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057.
- [71] F. Fazio, M. C. Bryan, O. Blixt, J. C. Paulson, C.-H. Wong, *J. Am. Chem. Soc.* **2002**, *124*, 14397.
- [72] V.V. Rostovtsev, L.G. Green, V.V. Fokin, K.B. Sharpless, *Angew. Chem. Int. Ed.* **2002**, *41*, 2596.
- [73] V. O. Rodionov, V. V. Fokin, M. G. Finn, *Angew. Chem.* **2005**, *117*, 2250.
- [74] T. R. Chan, R. Hilgraf, K. B. Sharpless, V. V. Fokin, *Org. Lett.* **2004**, *6*, 2853.
- [75] D. Horton, *Methods Carbohydr. Chem.* **1972**, *6*, 282.
- [76] A. Titz, Z. Radic, O. Schwardt, B. Ernst, *Tetrahedron Letters* **2006**, *47*, 2383.

- [77] H. Li, F. Cheng, A. M. Duft, A. Adronov, *J. Am. Chem. Soc.* **2005**, *127*, 14518.
- [78] H.C. Kolb, M.G. Finn, K.B. Sharpless, *Angew. Chem. Int. Ed.* **2001**, *40*, 2004; 72.
- [79] H. Kotsuki, T. Araki, A. Miyazaki, M. Iwasaki, P. K. Datta, *Org. Lett.* **1999**, *1*, 499.
- [80] M. Argentini, D. F. Dos Santos, R. Weinreich, H.-J. Hansen, *Inorg. Chem.* **1998**, *37*, 6018.
- [81] G. M. Brooke, *Perkin I*, 1974, 233.
- [82] J. P. Wolfe, S. Wagaw, J.-F. Marcoux, S. L. Buchwald, *Acc. Chem. Res.* **1998**, *31*, 805.
- [83] T. E. Barder, S. D. Walker, J. R. Martinelli, S. L. Buchwald, *J. Am. Chem. Soc.* **2005**, *127*, 4685.
- [84] X. Huang, K. W. Anderson, D. Zim, L. Jiang, A. Klapars, S. L. Buchwald, *J. Am. Chem. Soc.* **2003**, *125*, 6653.
- [85] M. Dowlut, D. G. Hall, O. Hindsgaul, *J. Org. Chem.* 2005, *70*, 9809-9813.
- [86] Z. Tarasiejska, R. W. Jeanloz, *J. Am. Chem. Soc.* 1958, *80*, 6325.
- [87] P. Vitorelli, J. Peter-Katalinic, G. Mukherjee-Mueller, H.-J. Hansen, H. Schmid, *Helv. Chim. Acta* 1975, *58*, 1379-1424.
- [88] P. R. Ashton, S. E. Boyd, C. L. Brown, N. Jayaraman, S. A. Nepogodiev, J. F. Stoddart, *Chem. Eur. J.* **1996**, *2*, 1115.
- [89] H. J. M. Kempen, C. Hoes, J. H. van Boom, H. H. Spanjer, J. de Lange, A. Langendoen, T. J. C. van Berkel, *J. Med. Chem.* **1984**, *17*, 1306.
- [90] Z. Pakulski, D. Peirozynski, A. Zamojski, *Tetrahedron* **1994**, *50*, 2975.
- [91] E. L. Eliel, R. M. Enanoza, *J. Am. Chem. Soc.* **1972**, *94*, 8072.
- [92] E. A. L. Biessen, D. M. Beuting, H. C. P. F. Roelen, G. A. van de Marel, J. H. van Boom, T. J. C. van Berkel, *J. Med. Chem.* **1995**, *38*, 1538.
- [93] J. E. Gestwicki, L. E. Strong, L. L. Kiessling, *Chem. Biol.* **2000**, *7*, 583.
- [94] E. J. Gordon, W. J. Sanders, L. L. Kiessling, *Nature*, **1998**, *392*, 30.
- [95] In-house measurements.



- [96] S. B. Shuker, P. J. Hajduk, R. P. Meadows, S. W. Fesik, *Science* **1996**, 274, 1531.
- [97] P. J. Hajduk, G. Sheppard, D. G. Nettlesheim, E. T. Olejniczak, S. B. Shuker, *J. Am. Chem. Soc.* **1997**, 119, 5818.
- [98] W. Jahnke, in *BioNMR in Drug Research*, Ed.: O. Zerbe, **2002**, p. 431, Wiley-VCH Verlag GmbH.
- [99] W. Jahnke, L. B. Perez, C. G. Paris, A. Strauss, G. Fendrich, C. M. Nalin, *J. Am. Chem. Soc.* **2000**, 122, 7394.
- [100] W. P. Jencks, *Proc. Nat. Acad. Sci. U.S.A.* **1981**, 78, 4046.
- [101] L. J. Berliner, *Prog. Bioorg. Chem.* **1974**, 3, 1.
- [102] R. T. Lee, R. W. Myers, Y. C. Lee, *Biochemistry*, **1982**, 21, 6292.
- [103] S. Nakabayashi, C. D. Warren, R. W. Jeanloz, *Carbohydr. Res.* **1986**, 150, C7.
- [104] D. J. Miller, M. B.-U. Surfraz, M. Akhtar, D. Gani and R. K. Allemann, *Org. Biomol. Chem.* 2004, 2, 671.
- [105] J. W. Peng, C. A. Lepre, J. Fejzo, N. Abdul-Manan, J. M. Moore, *Methods Enzymol.* **2001**, 338, 202.
- [106] R. Manetsch, A. Krasinski, Z. Radic, J. Raushel, P. Taylor, K. B. Sharpless, H. C. Kolb, *J. Am. Chem. Soc.* **2004**, 126, 12809.
- [107] N. Pravdic, T. D. Inch, H. G. Fletcher, *J. Org. Chem.* **1967**, 32, 1815.
- [108] D. Stokmaier, PhD Thesis, University of Basel, Switzerland.
- [109] M. A. Cooper, *Nat. Rev. Drug Discov.* **2002**, 7, 515.
- [110] P.-O. Markgren, W. Schwaal, M. Hamalainen, A. Karlen, A. Hallberg, B. Samuelsson, U. H. Danielson, *J. Med. Chem.* **2002**, 45, 5430.

## Curriculum Vitae

Name: Oleg  
Surname: Khorev  
Date and place of birth: 21<sup>st</sup> of August 1981, Moscow, Russian Federation  
Address: Murifeldweg 59, Bern CH-3006, Switzerland  
Contact details: Tel.: +41 61 267 1557 (work)  
+41 76 547 6909 (mobile)  
Email: oleg.khorev@hotmail.com  
Gender: Male  
Nationality: Russian  
Marital status: Single

### Education

2003-2007: PhD student in Medicinal Chemistry  
Institute of Molecular Pharmacy  
Pharmacenter, University of Basel  
Klingelbergstrasse 50, Basel CH-4056, Switzerland  
**Thesis title:** "Design and synthesis of high affinity ligands for the asialoglycoprotein receptor (ASGP-R)"  
**Research focus:** rational design and synthesis of drugs and site-specific drug delivery systems  
**Supervisor:** Prof. Dr. Beat Ernst  
**Grade:** magna cum lauda

1999-2003: Honors degree in Medicinal Chemistry  
University of Dublin, Trinity College, Dublin 2, Ireland  
**Grade:** I (first class honors)

1993-1999: Templeogue College High School  
Templeville Road, Dublin 6W, Ireland

### **Work experience**

Dec. 2003-Oct. 2007: PhD student in Medicinal Chemistry at the University of Basel, Switzerland

May-Sept. 2006: Diploma work supervisor of F. Böni

December 2005: Assistant scientific editor to Prof. Dr. B. Ernst for the *Chimia* 1/2 2006 – Vol. 60 “Pharmaceutical Sciences at Uni Basel”

May-Sept. 2005: Diploma work supervisor of M. Wittwer

Jul.-Aug. '00, '01, '02: Research assistant to Dr. Yurii Gun'ko  
Department of Chemistry, University of Dublin, Trinity College, Dublin 2, Ireland

## Scientific merits

Publication:

Oleg Khorev, Daniela Stokmaier, Oliver Schwardt, Brian Cutting, Beat Ernst: "Trivalent, Gal/GalNAc-containing Ligands designed for the Asialoglycoprotein Receptor" *Bioorganic & Medicinal Chemistry*, **2007**, accepted for publication.

Presentation given at the SCS Fall Meeting 2005, Lausanne, Switzerland. Abstract 64: "Fluorescent, trivalent, Gal/GalNAc-terminated ligands for the asialoglycoprotein receptor (ASGP-R)".

## Languages

### Level

English	Native speaker
Russian	Native speaker
German	Intermediate
French	Intermediate
Spanish	Basic

## Hobbies and interests

Fishing and tae kwon do.

## Referees

Prof. Dr. Beat Ernst  
Institute of Molecular Pharmacy  
Pharmacenter  
University of Basel  
Klingelbergstrasse 50  
Basel CH-4056  
Switzerland  
Tel.: +41 61 267 1551  
Email: [Beat.Ernst@unibas.ch](mailto:Beat.Ernst@unibas.ch)

Prof. Yurii Gun'ko  
Department of Chemistry  
University of Dublin,  
Trinity College  
Dublin 2  
Ireland  
Tel.: +353 1 608 3543  
Email: [igounko@tcd.ie](mailto:igounko@tcd.ie)

Dr. Oliver Schwardt  
Oberassistent  
Institute of Molecular Pharmacy  
Pharmacenter  
University of Basel  
Klingelbergstrasse 50  
Basel CH-4056  
Switzerland  
Tel.: +41 61 267 1557  
Email: [Oliver.Schwardt@unibas.ch](mailto:Oliver.Schwardt@unibas.ch)