Freie Universität Berlin

Convergent Syntheses of Multivalent Binding Glycoconjugates

Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry and Pharmacy
of Freie Universität Berlin

by
Lukas Michael Artner
from Mödling, Austria

2014

The work presented herein was carried out under supervision of Prof. Dr. Christian P. R. Hackenberger at the Institute of Chemistry and Biochemistry in the Department of Biology, Chemistry and Pharmacy of the Freie Universität Berlin, and at the Leibniz-Insitut für molekulare Pharmakologie im Forschungsverbund Berlin e.V. (FMP) from October 2009 to August 2014.

1st Reviewer: Prof. Dr. Christian P. R. Hackenberger
2nd Reviewer: Prof. Dr. Rainer Haag

Date of defense: December 17th, 2014

"An expert is a man who has made all the mistakes which can be made in a very narrow field."

– Niels Bohr

Acknowledgments

Since a doctoral thesis is not an achievement you do all by yourself, I kindly want to give a "thank you" note to the people that helped my in its realization.

First and foremost I want to thank my supervisor, Prof. Dr. Christian Hackenberger, for giving me the opportunity of proceeding my thesis within his laboratory. I appreciate the scientific challenges and discussions, and the given freedom to further develop my projects.

In addition I want to thank Prof. Dr. Rainer Haag for being my reviewer and his encouraging way of supporting young scientists.

I also want to thank my collaborators Prof. Dr. Nediljko Budisa and Dr. Jens Dernedde, as well as their co-workers for scientific advice and discussions along the way. Here I want to mention especially Nina Bohlke, who dealt with the molecular biology and the biosynthesis of proteins. To her I owe a big "thank you".

Another one goes out to Dr. Chris Weise, as well as the members and alumni of the Hackenberger group. I especially want to mention to the "veterans", Da'san, Ina, Michaela, Paul, Robert, Verena, as well as Max, for helping me out in the very beginning of my thesis.

At this point I also want to thank the people that became dear friends: My former and fellow lab mates, Nicole and Jordi, and the most precious people outside of my lab, Emanuel and Katja.

Last but not least I want to thank my family, Ernst, Brigitte and Markus, for being a constant support the whole time.

Table of Contents

Figures

Schemes

Abbreviations

α	1,2-*cis* glycosidic bond (in D-hexoses)	CuAAC	copper-catalyzed alkyne azide cycloaddition
Å	angstrom(s)	δ	chemical shift in parts per million
A	adenine (nucleobase), adenosine (ribonucleoside)	d	doublet (spectral); deci
		DCE	1,2-dichloroethane
A	absorbance	DCM	dichloromethane
AA	(canonical) amino acid	DMAP	4-(N,N-dimethylamino)pyridine
aaRS	aminoacyl tRNA synthetase	DMSO	dimethyl sulfoxide
Ac	acetyl	DNA	deoxyribonucleic acid
ADP	adenosine 5′-diphosphate	dr	diastereomer ratio
AMP	adenosine 5′-monophosphate	DTT	dithiothreitol
Aha	L-azidohomoalanine	ε	extinction coefficient
anhyd	anhydrous	E1	unimolecular elimination
aq	aqueous	E2	bimolecular elimination
ASGP-R	asialoglycoprotein receptor	EDTA	ethylenediaminetetraacetic acid
Ar	aryl	eq	equation
atm	atmosphere(s)	equiv	equivalent
ATP	adenosine 5′-triphosphate	er	enantiomer ratio
β	1,2-*trans* glycosidic bond (in D-hexoses)	ESI	electrospray ionization
		Et	ethyl
b*	barstar, barnase inhibitor protein	EtOAc	Ethyl acetate
Bn, Bzl	benzyl	FACS	fluorescence-activated cell sorting
BOC	Boc tert-butoxycarbonyl	FT	Fourier transform
br	broad (spectral)	g	gram(s)GTP guanosine 5′-triphosphate
BSA	bovine serum albumin		
Bu, n-Bu	normal (primary) butyl	*G*	Gibb's free energy
t-Bu	tert-butyl	G	guanine (nucleobase), guanosine (ribonucleoside)
Bz	benzoyl (not benzyl)		
°C	degrees Celsius	Gal	D-galactose
c	concentration	GalNAc	*N*-acetylgalactosamine, 2-(acetylamino)-2-deoxy-D-galactose
C	cytosine (nucleobase), cytodine (ribonucleoside)		
		GFP	green fluorescent protein
calcd	calculated	Glc	D-glucose
cAMP	adenosine cyclic 3′,5′-phosphate	GlcNAc	*N*-acetylglucosamine, 2-(acetylamino)-2-deoxy-D-glucose
CRD	carbohydrate recognition domain		
cat	catalytic	GPI	Glycophosphatidylinositol
CD	circular dichroism	h	hour(s)
cDNA	complementary deoxyribonucleic acid	*H*	enthalpy
		HIV	Human immunodeficiency virus
cm	centimeter(s)	Hpg	L-homopropargylglycine
cm⁻¹	wavenumber(s)	HMBC	heteronuclear multiple bond correlation
COSY	correlation spectroscopy		

HMQC	heteronuclear multiple quantum correlation	m/z	mass-to-charge ratio
HOMO	highest occupied molecular orbital	N	normal (equivalents per liter)
HPLC	high-performance liquid chromatography	Neu5Ac	*N*-acetylneuraminic acid
		ncAA	non-canonical amino acid
HRMS	high-resolution mass spectrometry	NIS	*N*-iodosuccinimide
HSQC	heteronuclear single quantum correlation	nm	nanometer(s)
		NMR	nuclear magnetic resonance
Hz	hertz	Nu	nucleophile
IC_{50}	dose that inhibits 50% of test subjects	PAA	polyacrylamide
		PAGE	polyacrylamide-gel electrophoresis
		PBS	phosphate-buffered saline
IR	infrared	PDB	RCSB protein data bank, (www.rcsb.org/pdb)
ITC	isothermal titration calorimetry		
J	coupling constant (in NMR spectrometry)	PEG	polyethylenglycol
		PG	protecting group; polyglycerol
k	kilo	dPG	dendritic polyglycerol
K	kelvin(s) (absolute temperature)	hPG	hyperbranched polyglycerol
K	rate constant	PNA	peanut agglutinin
K_d	dissociation constant	ppm	part(s) per million
L	liter(s)	Pr	propyl
λ	wavelength	iPr	isopropyl
Lac	D-lactose	PTM	post-translational modification
LG	leaving group	py	pyridine
lit.	literature value	q	quartet (spectral)
LUMO	lowest unoccupied molecular orbital	rel	relative
μ	micro	Rf	retention factor (in chromatography)
m	multiplet (spectral); meter(s); milli		
M	molar (moles per liter)	ROS	reactive oxygen species
M^+	parent molecular ion	rRNA	ribosomal ribonucleic acid
MALDI	matrix-assisted laser desorption ionization	rt	room temperature
		s	singlet (spectral); second(s)
Man	D-mannose	*S*	entropy
max	maximum	SEC	size-exclusion chromatography
MBP	mannose-binding lectin, mannan-binding protein	SBA	soybean agglutinin
		SDS	sodium dodecylsulfate
Me	methyl	S_N1	unimolecular nucleophilic substitution
MHz	megahertz		
min	minute(s); minimum	S_N2	bimolecular nucleophilic substitution
mM	millimolar (millimoles per liter)		
MMR	mannose macrophage receptor	SOMO	single-occupied molecular orbital
MO	molecular orbital	SPI	selective pressure incorporation
mol	mole(s); molecular (as in mol wt)	SPR	surface plasmon resonance, SPR spectroscopy
mp	melting point		
mRNA	messenger ribonucleic acid	SPtR	Staudinger phosphite reaction
Ms	methylsulfonyl (mesyl)	SPnR	Staudinger phosphonite reaction
MS	mass spectrometry	t	triplet (spectral)
MW	molecular weight	t	time; temperature in units of

ix

	degrees Celsius (ºC)
T	absolute temperature in units of kelvins (K)
T	thymine (nucleobase), 5-methyluridine (ribonucleoside)
TBAB	tetrabutylammonium bromide
TBAHS	tetrabutylammonium hydrogen sulfate
TBTA	tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine
TCA	trichloroacetic acid
THPTA	tris(3-hydroxypropyltriazolyl methyl)amine
Tf	trifluoromethanesulfonyl (triflyl)
TF-antigen	Thomsen-Friedenreich antigen
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin-layer chromatography
TMS	trimethylsilyl; tetramethylsilane
TOF	time-of-flight
tRNA	transfer ribonucleic acid
Ts	para-toluenesulfonyl (tosyl)
U	uracil (nucleobase), uridine (ribonucleoside)
UV	ultraviolet
V, vol	volume
v/v	volume per unit volume (volume-to-volume ratio)
w/w	weight per unit weight (weight-to-weight ratio)
wt-%	weight in gramm (g) per 100 milliliter (mL)
WGA	wheat germ agglutinin

1. General Introduction

1.1. About this chapter

This chapter should give a brief description and general background information on postransational modifications and their necessity in biology, focusing especially on carbohydrates and glycans. Following this, the reader will be introduced to carbohydrate recognizing proteins, the so called lectins.

1.2. Fundamentals

Biochemistry, biological chemistry and chemical biology all thrive to understand the nature of living organism through chemistry. Therefore they are interdisciplinary bridges between basic chemistry – the science of interactions and structures of atoms and molecules – and biology – the study of interactions and structures of cells and living organisms. Biology recognizes three major classes of macromolecules that contribute to most events in cells through supramolecular assembly with themselves or between each other. They are proteins, polysaccharides and nucleic acids, and, in addition to those, lipids. Each macromolecular structure consists of smaller monomers, as amino acids for proteins, carbohydrates for polysaccharides and nucleotides for nucleic acids. All the structures are interconnected with each other. For example, all the information for a cell to "live" is stored in its genetic code as deoxyribonucleic acid (DNA). From that it can produce, or *translate* it to functional proteins. The proteins themselves will further contribute to the variation found within the cell. [1]–[3]

1.2.1. Translation of the genetic code

As known from the "central dogma of molecular biology" (**Figure 1-1**, p. 2),[2] postulated by Crick[4], the genetic information is encoded in bases of nucleic acids, either deoxyribonucleic acid (DNA) or ribonucleic acid (RNA). From this point on, it can be either stored in those two forms, conserved as transfer-RNA (tRNA), ribosomal RNA (rRNA), and many more, or translated *via* messenger-RNA (mRNA) to proteins. In addition the genetic information can be reversibly transcribed from messenger-RNA
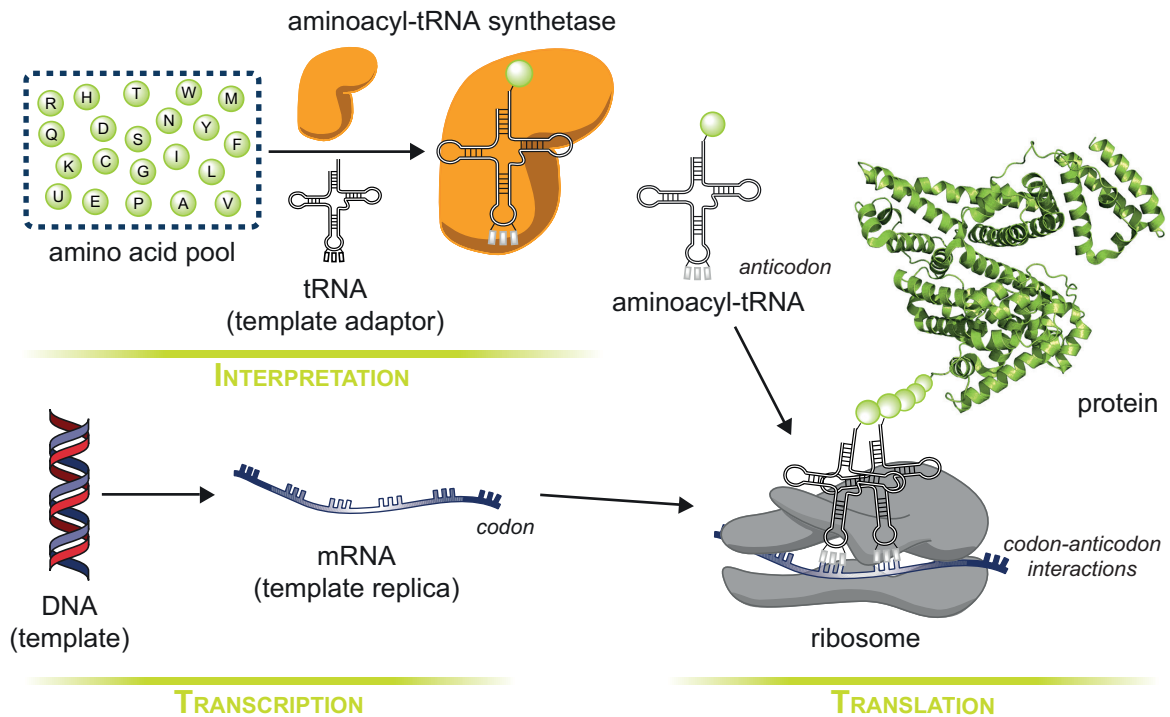
Figure 1-1: **The transmission of the genetic code as a flow chart according to the "central dogma".**

For successful translation of genetic information into proteins, the following events have to occur accurately: 1) the tRNA has to be aminoacylated with its corresponding amino acid by an aminoacyl-tRNA-synthetase. 2) Proper transcription from DNA to mRNA provides the right template for the aminoacylated tRNA to interact. Finally, 3) mRNA-tRNA codon-anticodon interactions at the ribosome lead to transfer of the activated amino acid at the tRNA to the growing polypeptide chain, thus translating into a protein.

(mRNA) by reverse transcriptase proteins. This postulate was pointing towards the direction of genetic information to proteins in living organisms. It also emphasizes that there is transfer of information possible between nucleic acids (DNA, RNA), one-directional from DNA and RNA to proteins, but neither between proteins nor from proteins back to the level of nucleic acids.

The genetic code is basically written in four different nucleic acids, the letters. Every triplet of three subsequent letters – the so-called codon – stands for either an amino acid or termination during the ribosomal protein synthesis. The precise series of triplets form a polynucleotide that is subsequently "translated" to a sequence of amino acids, to form a polypeptide. After the deciphering of the genetic code, 20 so-called canonical amino acids have been assigned to 61 coding and three terminating triplets. **Figure 1-2**, p. 3 depicts this schematically as a wheel. This emphasizes the combinatorial possibilities of a three letter code. [5]
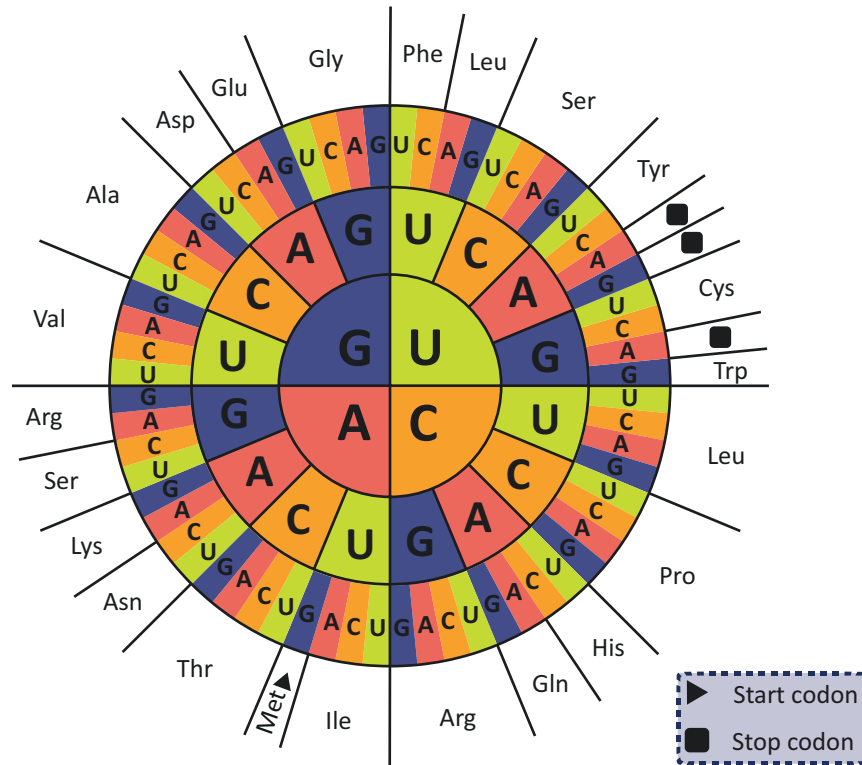
Figure 1-2:   **The genetic code.**

Every triplet of nucleotides forms a coding unit, or codon. 61 of them are "sense" codons and encode for twenty canonical amino acids. The remainder of the 64 codons signal termination of protein translation. They are "stop" or "nonsense" codons. AUG is the one triplet the encodes for both start of translation and the amino acid methionine.

The graphic was redrawn and adapted from references [2] and [172].

1.2.2. Post-translational modifications

The combinatorial possibilities of the 20 encoded canonical amino acids can already give rise to a huge variety of peptides and proteins. In organisms even more complexity is applied through tissue-specific, alternative mRNA splicing and by introducing post-translational modifications (PTM) of amino acid side-chains in proteins. For PTMs this usually happens after the translation and is often mediated by specific enzymes. A lot of different functionalities are known to be introduced, including phosphorylation, acylation, methylation, prenylation, glycosylation, hydroxylation, biotinylation, ubiquitinylation and lipoylation to name some (**Figure 1-3**, p. 4). However, phosphorylation and glycosylation seem to be the most abundant ones, with glycosylation having the most side-chain variety by far. [6]–[8]

The abundance of glycosylation as PTMs in mammalian cells is widely known.[9] If one is taking a closer look at this side-chain modification by carbohydrates, it seems that this PTM was created by nature as an additional code to apply even more variety to proteins.
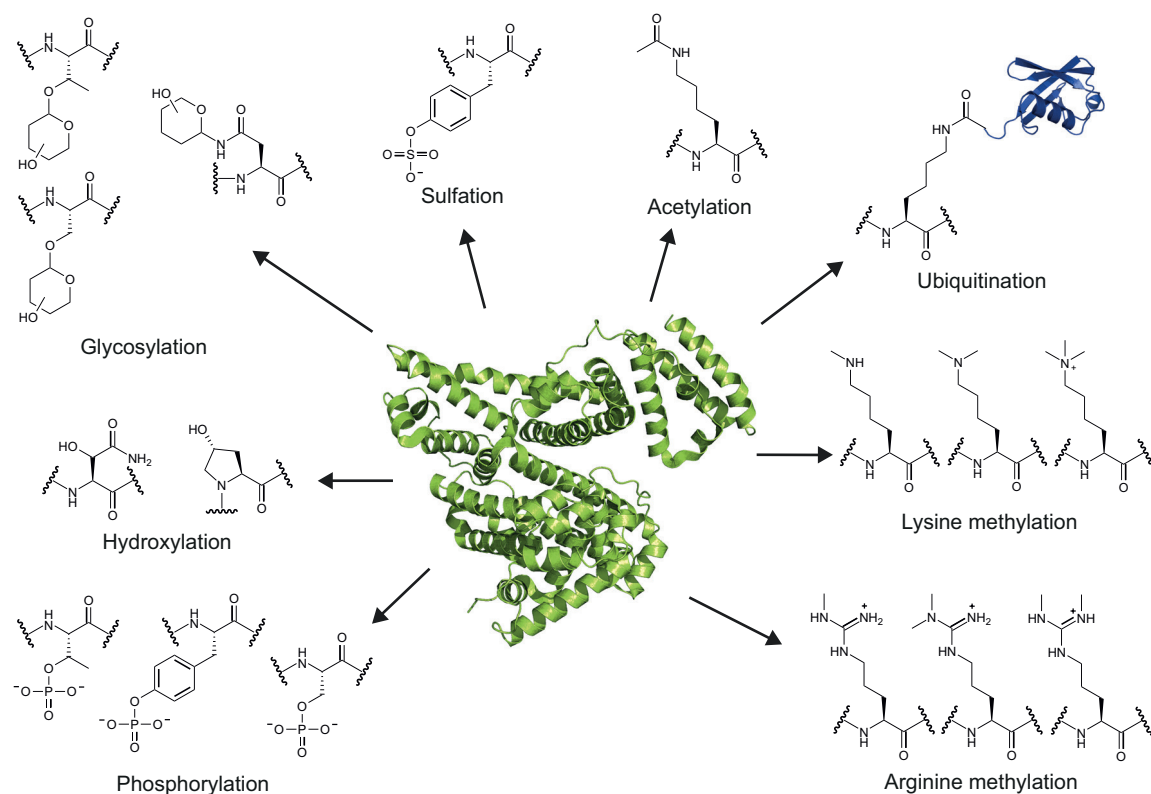
Figure 1-3: **A selection of some of the most common post-translational modifications of proteins.**

Post-translational modifications of proteins extend the combinatorial possibility and variation even more. In addition it expands the complexity of biological interactions and phenomena. The modifications depicted herein represent just a selection of the wide array of discovered PTMs.[6]–[8]

Taking into consideration that the major classes of glycosylations in mammalian cells are either *O*- or *N*-glycans, mainly consisting of seven different carbohydrates, the combinatorial possibilities of assembly already result in a myriad of different forms. In addition to this structural variety, modifications like sulfatation and phosphorylation are also common in glycans.

1.3.   The biological role of carbohydrates and glycans

Despite the fact that the whole genetic information is stored in the DNA, molecular biology cannot alone provide answers to the complexity of cells and their hierarchical assemblies to tissues, organs over to whole organisms. As explained in the previous passages, post-translational modifications that happen after the protein biosynthesis do their part as well. Major contributors to this puzzle are carbohydrates – or sugars – in the form of cell recognition and signaling, or as structural components and energy source. Carbohydrate structure are especially important, since they contribute in protein

4

folding and stability, but also take an active part in the interaction between cells and the surrounding extracellular matrix, which is crucial for multicellular environments and communication, such as in organs and living organisms. [9], [10], [11]

The surface of all cells and a large number of macromolecules inside and outside of them are decorated with various forms of carbohydrates. They can be conjugated to single sugars (mono-saccharides) or sugar chains (oligo-saccharides) – which are generally referred to as glycans – and contribute to various classes of glycoproteins, glycolipids, GPI anchors, or simply, glycoconjugates. The complexity and heterogeneity of the outer carbohydrate environment makes glycans an ideal mediator for cell–cell, cell–matrix and other cell–molecule interactions. This "sweet husk" – or glycocalyx – is unique and characteristic for individual cell and tissue types, their state of development or even a marker for their health status.[12] The latter part is from significant importance for todays medicine, since it contributes not only to interaction of different organisms with each other (host–pathogen and host–symbiont interactions), but also marks malignant and uncontrolled cell developments like cancer. [11]–[13]

The relevance of carbohydrates for biology and medicine is there. Since the biochemistry and biophysics of carbohydrate-mediated communications are yet not fully understood, further research has to be conducted. Better understanding of those ligand–receptor interactions open up further possibilities for modern medicine and therapeutics, as did the discovery of blood groups by Dr. Karl Landsteiner in 1900. [14]

1.3.1. Lectins and carbohydrate-binding proteins

The interactions of carbohydrates with proteins occur regularly in nature and often in a non-covalent way. There are carbohydrate-specific enzymes – like glycosyltransferases, glycosidehydrolases and transglycosylases to name a few families – and there are antibodies which innate a immune response by recognizing specific carbohydrate antigens. Both of this classes play an important role especially in higher organisms. Within the context of cell interactions and adhesion processes, a there is also a third class of proteins involved that deals with carbohydrate specificity – lectins. They contain a "carbohydrate recognition domain" (CRD) which sole purpose is the effective differentiation between various oligo-saccharides. The ligands and receptors are presented multivalently to enhance avidity and specificity. Details about the phenomenon of multivalency will be given later.[15]

Lectins recognize and reversibly bind mono– and oligosaccharides, even partial structures of glycans, without having a catalytic activity or being the product of the immune system like antibodies are. They "choose", "pick" or "select" their respective carbohydrates and this behavior is reflected in the latin translation of those words where the name lectin is deducted from – *legere*. Their capability of binding to saccharides is

possible because of carbohydrate recognition domains which is specific for each family of lectins.[15], [16]

Historically the first lectins were discovered over more than 100 years ago and isolated from plants. Later they were also found in other organisms, like bacteria, eukaryotes and more complex organism. Even viruses were found to express lectin-like proteins. In regard to their active site, the CRD, many lectins bind in a multivalent fashion and are therefore represented as di– or oligomers. Due to their ability to strongly bind carbohydrate epitopes, lectins can cross-link cells and glycoproteins non-covalently to form precipitates. Those lectins were also referred to as "agglutinates". A special form of agglutination is the clotting of highly glycosylated erythrocytes (red blood cells) by lectins, the so-called "hemagglutination". This capability of lectins is still often used for their characterization and detection itself, but also a useful medicinal tool for the distinction of blood types. [17]

Other lectins are often anchored to the lipid bilayer as membrane proteins. Their CRDs can either be pointing towards the cell exterior or the cytosol. If represented on cell organelles the interior side would be their lumen. Soluble lectins on the other side are usually found in the serum or the extracellular matrix. All together, lectins widely vary in their size, structure, assembly and specificity towards carbohydrates, making them a rather heterogenous group of protein oligomers. It is therefore not surprising that their classification was in the beginning rather difficult. Starting out by dividing them according to their binding behavior, for example mannoside-binding *vs.* galactoside-binding, the advances in molecular biology led to classification based on sequence and structural homology of their CRDs. For example, "C-type" lectins bind different carbohydrates in dependence of calcium, whereas calnexin, calreticulin and various L-type lectins also show dependency on calcium for binding, but have a different structure of the CRD and specificity for carbohydrates.

As of today, lectins that have evolutionary conserved or related CRDs are classified in an increasing number of groups. Drickamer *et al.* presented a classification of them.[18], [19]

Since a detailed view of the vast amount of medicinally and biologically relevant lectin groups would be beyond the scope of this introduction, the following section will only cover those relevant to conducted work that is represented in the text. The chosen lectins will be explained chronically of the appearance in the thesis, beginning with lectins that are derived from plants, over C-type lectins – the largest group of animal lectins – to hemagglutinin as a medically relevant viral glycan-binding protein.

1.3.1.1. Plant lectins

The first lectins ever discovered were plant lectins. Since than, biological and chemical research aided in their characterization and therefore made plant lectins useful tools in glycobiology and related fields dealing with carbohydrates. Due to their variety of specificities to saccharides, immobilization of plant lectins enabled additional means for the purification of oligosaccharides and glycoproteins through affinity chromatography. In addition, they proved to be ideal for model studies for basic research, since they are readily available and mostly well-characterized.

Since most of plant lectins are assembled to oligomers, they usually bind in a multivalent fashion to their respective carbohydrate epitope. With multiple binding sites for sugars, large clusters can be formed. This effective multivalent clustering of carbohydrate bearing conjugates by plant lectins is the reason behind their ability to efficiently mediate hemagglutination. Legume lectins – plant lectins that are isolated from the seeds of leguminous plants – are currently the best studied carbohydrate-binding proteins of their kind. *Concanavalin A*, a tetrameric plant lectin that was isolated from jack beans, as one of the first ones described.[20] It binds strongly to mannose-containing glycans and especially branched mannotrioside structures. Another well-studied plant lectin that also happens to be a soluble tetramer is peanut agglutinin (PNA) from *Arachis hypogaea*. Although crystallization experiments were reported back in the 1980's, a refined structure was obtained only more than ten years later. Nevertheless it proves as a solid model for a soluble multivalent binding plant lectin. [21]–[23]

The reason for the existence of plant lectins is still dazzling researchers. However, it is proposed that the physiological role of plant lectins may be to protect the host plant from various sources, such as phytopathogenic microorganisms and insects as well as herbivores. Another hypothesis assumes that plant lectins should help in maintaining the symbiotic bacteria. [24]

In 1976 the Liener lab came up with probably the earliest publication about insecticidal lectins derived from legumes. They were feeding black bean lectins to bruchid beetles, which resulted in the death of their larvae. This finding in addition to other studies supported the insecticidal action of various lectins.

The claim of the protective ability of lectins against pathogenic microorganisms arose from an observation whereas wheat germ agglutinin (WGA), peanut agglutinin (PNA) and soybean agglutinin (SBA) prevented certain fungi, among them *Trichoderma viride*, *Penilicium notatum* and *Aspergillus niger*,[25] from sporulation and therefore growth. Other studies supported the conclusions drawn from the previous findings, showing that an even wider range of lectins disrupt germination of spores in various fungi.[26]

However, the hypothesis that lectins mediate between plant and symbiotic bacteria maintains an uncertainty. Early findings showed that some nitrogen-fixing rhizobia and

legumes do share a carbohydrate-specific association behavior, which was not found with bacteria that are symbionts to other leguminous plants. The suggestions in that matter where for instance based on SBA binding to lipopolysaccharides of *Bradyrhizobium japonicum*, but not to nonnodulating *Bradyrhizobium* strains. Nevertheless, such a lectin mediated interface could not be identified for most host–symbiont systems, which sparked controversy and doubt regarding this suggestion. In addition, numerous lines of soybeans are in symbiosis with corresponding rhizobial strains, although no lectins were found in their seeds or vegetative tissue. The debate could not be fully resolved, but results of molecular genetics[27], [28] as well as the discovery of nodulation factors as a special type of plant lectins are bolstering the recognition hypothesis.[29]

1.3.1.2. C-type lectins

The largest constitutes of animal lectins known, are the Ca^{2+}-depended C-type lectins. This carbohydrate-binding proteins take part in a wide array of biological functions, usually related to the 'innate' immunity of the organism. Since the group of C-type lectins is containing a lot of members, it is further divided into the sub-groups of *endocytotic lectins*, *collectins*, and *selectins*. The latter consist of only three known representatives, L-, E-, and P-selectin. They take part in the removal of leukocytes from the blood circulation by selectively binding to them and mediate their migration into tissues where their action is required.[30], [31]

The membrane-bound endocytotic lectins are receptors with different specificities to carbohydrates. The mammalian hepatic asialoglycoprotein receptor (ASGP-R, **Figure 1-4**, p. 9), which was discovered by Ashwell and Morell in the seventies, is one of the prominent representatives of those lectins.[32], [37] It facilitates clearance of serum glycoproteins, by specifically binding to terminal galactose and *N*-acetylgalatosamine (*vide infra*, **4.2**, pp. 74ff.). In contrast to the regulatory activities of ASGP-R, the mannose macrophage receptor (MMR, **Figure 1-4**, p. 9) has the purpose of defense. It is a endocytotic, mannose-binding lectin expressed on macrophages. Macrophages are part of the organisms innate immunity, which is not dependent on T cells or antibodies. Their purpose is to target infectious microorganisms and destroy them after successful receptor binding. MMR is unusual in its constitution since it contains eight C-type lectin domains in one protein strand and a cystein-rich region. It recognizes a wide range of different carbohydrates, such as mannose, fucose, *N*-acetylglucosamine, but no glactose epitopes. Hereby should be noted that mannoses and glucoses are frequently found on the surface of infectious microorganism, but not galactoses. Nevertheless it was shown, that MMR can still bind GalNAc-4-SO$_4$ residues which is mainly represented in glycoprotein hormones. However, further studies revealed that it is the cystein-rich region of the receptor that mediates the interaction. [10]
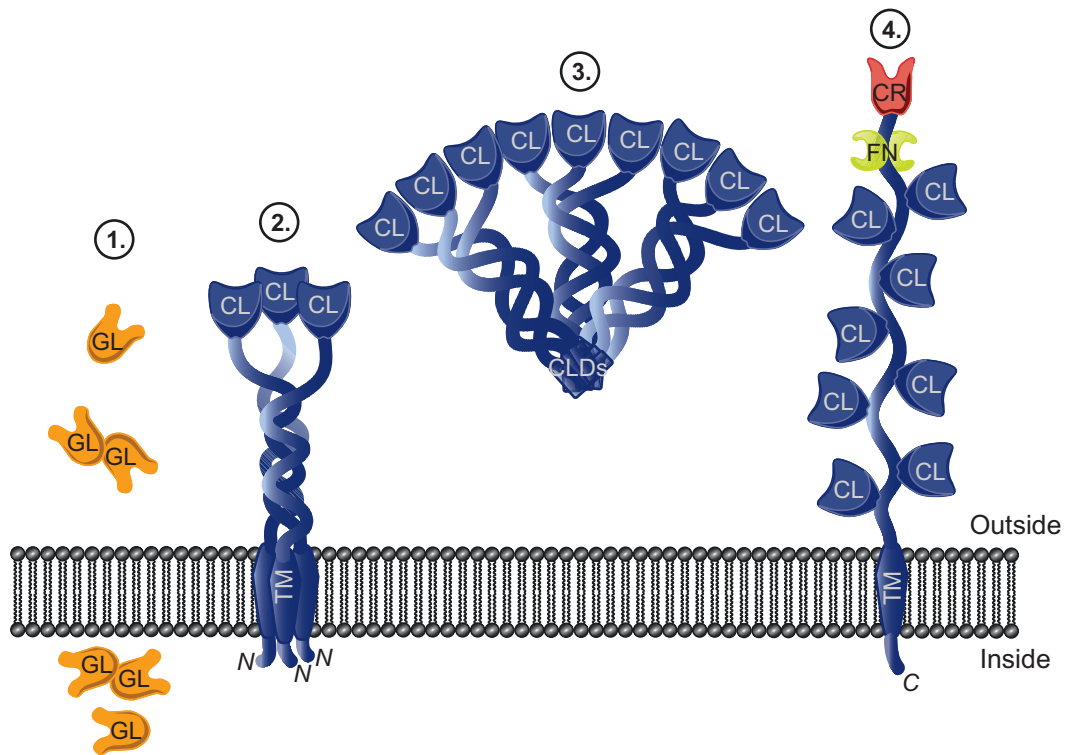
Figure 1-4: **Some examples of animal lectins.**

The examples are based on protein structures according to K. Drickamer, Ref. [18], [19].

Various forms of animal lectins exist. They usually hybridize to multimers and are either membrane-bound or soluble. Galectin-1 (1.) is a soluble lectin. The monomer consist to a major part of a carbohydrate recognition domain (GL) that resembles that of legume lectins of plants. Membrane-bound endocytotic C-type lectins (2.) like DC-SIGN or ASGP-R usually form multimers. In the case of ASGP-R the C-type lectin-like carbohydrate recognition domain (CL) sits on top of a coiled-coil stalk region that hybridizes with other monomers. It is bound to the cell-membrane by transmembrane domain (TM). MBP (3.) usually forms even bigger multimers. They are soluble and bound together by a collagen-like domain (CLD). Finally, MMR (4.) is another example of a membrane-bound animal lectin. It comprises of multiple CLs, as well as a ricin-like carbohydrate recognition domain (CR) and a fibronectin domain (FN). All of them can bind various terminal carbohydrates.

Figure adapted and redrawn from: Essentials of Glycobiology. 2nd edition. Varki A, Cummings RD, Esko JD, et al., editors. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 2009; Copyright © 2009, The Consortium of Glycobiology Editors, La Jolla, California.

Collectins got their name from the collagen-like domain they are containing. That part of the protein is necessary to assemble them in large oligomers, usually consisting of 9 to 27 subunits. Of the nine collectins discovered up to now, mannose-binding protein (MBP, **Figure 1-4**), sometimes also mannose-binding lectin or mannan-binding lectin/ protein, is probably the most prominent representative of them. It has a wide range of specificity, namely for terminal carbohydrates with an equatorial orientation of C3–OH and C4–OH groups. Such an orientation of hydroxy–groups is of course found in

mannose, but also in glucose and *N*-acetylglucosamine. Interestingly, the glycocalix of Gram-positive and Gram-negative bacteria, as well as the cell wars of yeasts and certain parasites are heavily constituted by those hexoses. In this context it should be noted, that terminating sialic acid, galactose, and *N*-acetylgalactosamine are not subject of the MBP. The aforementioned specificities makes MBP an ideal contributor to innate immunity of an organism. This kind of self-defense against potential pathogenic organisms is executed by mannose-binding protein associated serine proteases MASP-1 and MASP-2. Structurally, MBP consists of an initial assembly of a trimeric helical structure, mediated by the collagen-like domain, which can further aggregate to a large macromolecular architecture with three to six trimers arranged in a "bouquet-like" structure. Within the trimer each CRD is approximately 53 Å away from each other. Since a singular binding event of the MBP CRD with a terminal mannose in an oligosaccharide is very weak, the spacing and number of mannose-bearing ligands is from uttermost importance for effective binding. This means, that multiple MBP CRDs have to interact with multiple, appropriately-spaced mannans to achieve high avidity in the range of nanomolar dissociation constants. Interestingly, such extended mannose-containing glycoconjugates are especially found in the cell wall of bacteria, yeast, and certain parasites. At this point it should also be stated that MBP not only plays a role in the defense against microorganisms, but that there is also evidence that MBP may be beneficial against HIV. So far, conducted studies have yield controversial results. Apparently MBP binds to carbohydrates represented on gp120 by a wide array of HIV strains. However, whether this binding results in effective neutralization of HIV remains unclear. It also could not be fully resolved in what way MBP contributes to HIV clearance. Nevertheless, MBP's ability to intercept interactions between HIV and cell-surface lectins such as DC-SIGN is opening a new frontier for scientific exploration. [9], [10], [33]–[36]

DC-SIGN (dendritic cell-specific intercellular adhesion molecule-grabbing non-integrin) is another important example for a calcium-dependent carbohydrate-binding protein that is relevant for immunity. It is a mannose-specific C-type lectin expressed by dendritic cells and binds to the intercellular adhesion molecule (ICAM) embodied on the T-lymphocyte surface. For this reason it is playing a crucial role in the activation of T-lymphocytes by dendritic cells. Aside from those characteristics, DC-SIGN is also a strong binder to HIV, the human immunodeficiency virus. Due to this interaction HIV can remain bound to DC-SIGN for extended periods and therefore be efficiently transported throughout the organism, which results in highly effective infection. A similar C-type lectin that is expressed in specific endothelial cells has also been reported to have specificities analogous to the ones described for DC-SIGN. Due to its close relationship to DC-SIGN it was named DC-SIGN-related molecule (DC-SIGNR).[9], [10], [33]–[36]

Thus, one can see the importance of C-type lectins for the innate immunity of an organism. Another interesting feature is that of endocytosis which is mediated by both ASGP-R and MMR. The reasons for this are different for both receptors: Primarily it seems the reasons are self-defense and homeostasis of serum glycoproteins for MMR and ASGP-R, respectively.

β–Linked galactose or *N*-acetylgalactosamine often is the last hexose before the terminal sialic acids in animal glycans. Erythrocytes are no exception and also bear such pattern of glycosylation. However, errors in glycosylation of proteins and aging of the blood cells can expose galactose or *N*-acetylgalactosamine as the terminal residue. ASGP-R, an animal C-type lectin expressed on the surface of hepatic cells is recognizing those carbohydrate moieties. If an asialylated glycoprotein – one that is lacking terminal sialic acid – is bound by that hepatic receptor, it is internalized *via* clathrin-coated pits. Inside the hepatocytes, the glycoprotein will be transported to the lysosomes for degradation, whereas ASGP-R will be uncoupled from the endosomes and recycled (**Figure 1-5**, p. 12).[38] Aside from the clearance of glycoproteins from the serum, recent findings suggest that the role of asialoglyoprotein receptor is much more complex than that. It has been reported that oligosaccharides that have either an αNeu5Ac-(2→6)-β-D-Gal*p*-(1→4)-D-GlcNAc*p*[39] or αNeu5Ac-(2→6)-β-D-GalNAc*p*-(1→4)-D-GlcNAc*p*[40] but not an αNeu5Ac-(2→3)-β-D-Gal*p*-(1→4)-D-GlcNAc*p* motif as their terminal residue are also subject of the asiaologlycoprotein receptor. As a consequence of those observation it was suggested that ASGP-R has an additionally function as a regulator of the relative concentrations of serum glycoproteins. Therefore it can be stated that it maintains the homeostasis of serum glyoproteins in a thorough way.

ASGP-R was the first animal lectin discovered and due to its function and structure it still is governed as a prototype for C-type lectins. Although it has not been possible to crystallize a full receptor subunit up to now, 3D models and structures have been proposed due to experimental evidence, partial crystallizations,[19],[41]–[45] and homology modeling.[46] Basically, the human asialoglycoprotein receptor consists of two transmembrane protein subunits of type II, H1 and H2. Both have an exoplasmic C-terminus, an endoplasmic N-terminus and a molecular weight of approximately 46 kDa (H1) and 50 kDa (H2). They share a strong resemblance with each other, having a sequence homology of 57%, as well as the same domain structure. In addition, both are post-translationally palmitoylated and *N*-glycosylated. The active receptor is usually constituted by an oligomerization of both subunits in a ratio of 1:2 to 1:5 (H1:H2), forming heterotrimers or double heterotrimers. To effectively assume its role as a regulator of serum glycoprotein homeostasis, the receptor is facing the circulatory system on the basolateral membrane of hepatic cells.

Each subunit starts with an *N*-terminal cytosolic domain, followed by a transmembrane domain. The stalk region finally connects the C-terminal carbohydrate recognition
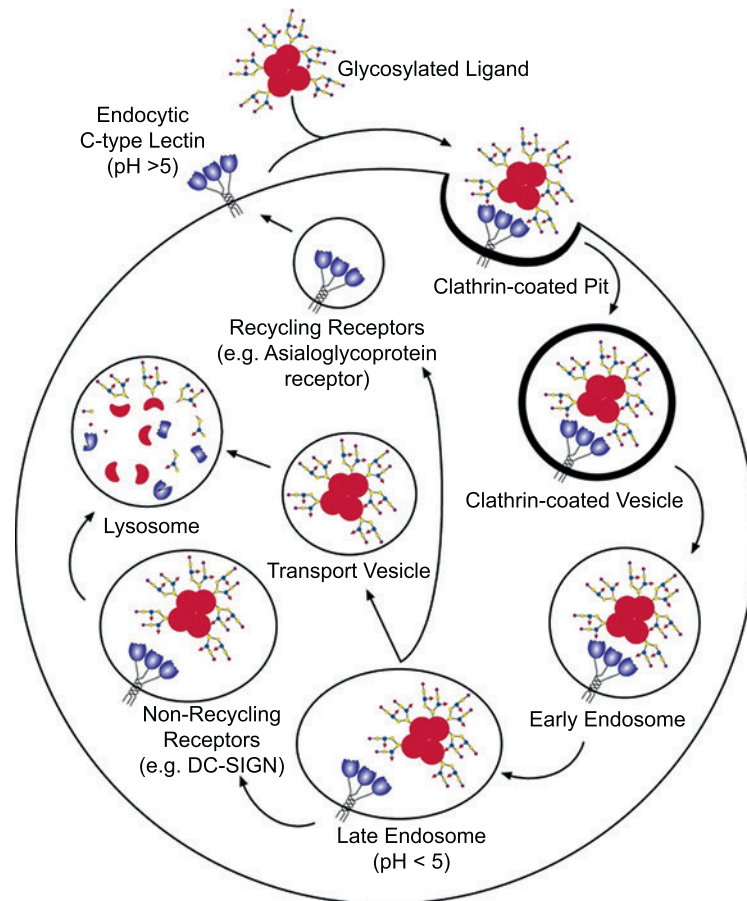
Figure 1-5:   Endocytosis mediated by C-type lectins.

Glycosylated ligands are recognized by specific lectins and are subsequently internalized by clathrin-coated pits. When the vesicle reaches the state of late endosomes, the receptors may either be recycled or degraded. This is usually dependenant on the endocytotic receptor itself and its ligand. Both, asialoglycoprotein receptor (ASGP-R) and the mannose macrophage receptor (MMR) are recycled to the cell surface after endocytosis.

From: Essentials of Glycobiology. 2nd edition. Varki A, Cummings RD, Esko JD, et al., editors. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 2009; Copyright © 2009, The Consortium of Glycobiology Editors, La Jolla, California.

domain with the membrane bound part of the receptor and is responsible for the oligomerization with other subunits. As a member of the C-type lectin family, ASGP-R is calcium dependent. From the crystallization of the H1-CRD it is known that three Ca^{2+}-atoms are located in the binding site, mediating the interaction with galactose or GalNAc residues. Interestingly the assembly of the receptor is in a way to fit the triantennary structure of *N*-glycans. Also, it encourages ligand and receptor clustering due to the fact that only one sugar can be bound by one CRD. This implies that three to six Gal/GalNAc residues per receptor-oligomer are needed for an efficient binding event. This phenomenon was extensively investigated by Y. C. Lee for over two decades, using synthetic glycoconjugates bearing one to three galactose moieties. Coining the term
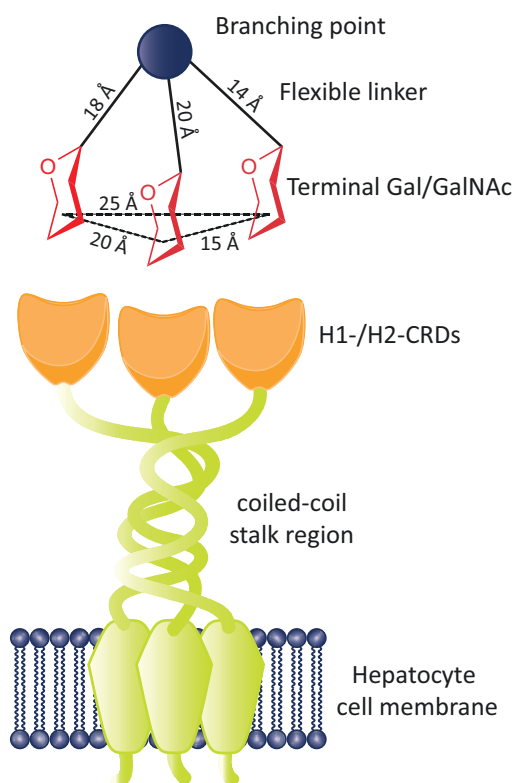
Figure 1-6: **The ideal ASGP-R ligand conformation.**
Binding model for ASGP-R ligands in an optimal conformation to the heterooligomeric receptor consisting of H1 and H2 subunits. Dashed line indicates the distance between the C-4 of each Gal moieties; filled line represents approximate distance between branching point and C-6 of Gal (14–20 Å). Redrawn and adapted from references [37] and [48].

"carbohydrate cluster effect", he described that when ligand presentation is increased linearly, the affinity to the receptor increased logarithmically. This also marked the concept of multivalent binding glycoconjugates (*vide infra,* **3**, pp. 60ff., **6**, pp. 101ff. **4**, pp. 74ff.). Furthermore, due to clever design of glycoconjugates, evaluation of ligand affinities and avidity, as well as known structural considerations, Lee could suggest a solid model for the actual physiological display of ASGP-R. It was therefore postulated that the ideal ligand design should aim for a single or double heterotrimer of subunits. In addition, the ASGP-R specific glycoconjugate should have the possibility to bridge the distance of each subunit, which is 15, 20, and 25 Å (**Figure 1-6**). Due to the physiological and medicinal relevance of the mammalian asialoglycoprotein receptor, as well as the ability of being a perfect model for endocytosis and drug delivery, numerous attempts have been conducted regarding optimal ligand design for multivalent ASGP-R targeting glycoconjugates.[47]–[50]

1.3.1.3. Hemagglutinin

As with plants, animals and bacteria, also viruses tend to express carbohydrate-binding proteins on their surface. The first viral lectin to be discovered was hemagglutinin (HA), which was identified by Gottschalk in the 1950s. He found out that HA from influenza virus strongly binds to eryhtrocytes and therefore induce hemagglutination. Also other cells bearing sialic acids residues on their cell surface were subject of binding to influenza HA. Approximately thirty years later, in 1981, Wiley and co-workers were able to crystallize a viral hemagglutinin with its ligand sialyllactose and solve its structure.[51] From this point on, numerous hemagglutinins have been crystallized and characterized.[52]

Thus, influenza virus hemagglutinin is still the best-studied viral carbohydrate-binding protein. Like with other lectins, the principal affinity of the receptor to its natural ligand – sialylated glycans – is rather low. Here the same principal applies as with other lectins: Oligomerization of the receptor and dense glyco-clustering on the cell surface of the host create an ideal multivalent system for interaction with high avidity. For the virus, strong surface binding is a necessary prerequisite for infection. Only if it binds tightly, the viral envelope and the plasma membrane of the host can fuse, which leads to uptake of the virus into the host cell, thus infecting it.

Viral hemagglutinins come in many variations and are tailored for the specific host the strain is targeting. For example, human influenza A and B viruses have a high specificity to epitopes containing an Human strains of influenza-A and -B viruses bind primarily to cells containing an αNeu5Ac-(2→6)-β-D-Gal*p* motif, whereas avian influenza viruses tend to bind to αNeu5Ac-(2→3)-β-D-Gal*p* residues. Interestingly, porcine borne influenza bind to both αNeu5Ac-(2→3)-β-D-Gal*p* and αNeu5Ac-(2→6)-β-D-Gal*p* containing glycans. This specificity is due to certain changes of amino acids inside of hemagglutinin's carbohydrate-recognition domain. A special case is the influenza-C virus, that exclusively binds to 9-O-acetylated N-acetylneuraminic acid motifs that are found in some glycoproteins and glycolipids.

The previously mentioned specificities of different hemagglutinins with corresponding sialyl-galactose residues are strongly correlating with the availability of glycans on the target epithelial cells of the host. Interestingly, the αNeu5Ac-(2→6)-β-D-Gal*p* motif is very much present in human epithelial cells of the trachea, but other tissue predominantly contains αNeu5Ac-(2→3)-β-D-Gal*p* residues. For this reason hemagglutinin is determining the species and cell targeting of the virus. Vice versa, hemagglutinin is often targeted neutralizing antibodies of the host cells. It is a major antigen and due to periodical mutations of this protein, viral outbreaks are often the consequence. Physiologically it is important not only for the binding event, but especially after the internalization, since it is facilitating the fusion of the endosomal

membrane and the viral envelope.

Structurally, hemagglutinin consists of two subunits, HA_1 and HA_2, which are derived by proteolytic cleavage of a parent polypeptide, HA_0. This is usually done by a trypsin-like serine endoprotease that typically cleaves after a specific arginine in the sequence of HA_0, thus generating the C-terminus of HA_1 and the N-terminus of HA_2. The residual arginine is subsequently removed from the sequence by a carboxypeptidase. The separation of HA_0 into its subunits HA_1 and HA_2 is required for its proper functioning in the event of membrane fusion and general infectivity. When the cleavage is done, the two disulfide-bonded subunits form that active matured protein monomer, which further oligomerizes into the trimeric hemagglutinin. Each monomer by itself starts with a C-terminal domain inside of the viral envelope and continues by spanning the membrane with an hydrophobic domain. The viral receptor is further elongated by a stem region that – after 135 Å past the membrane – is topped by globularly shaped carbohydrate-recognition domain. In the crystal structure of the CRD two binding pockets were identified: A primary for sialosides and a secondary that seemed more suited for hydrophobic molecules. Whereas the biological relevance of the primary site is clear due to its specificity to either αNeu5Ac-(2→6)-β-D-Gal*p*, αNeu5Ac-(2→3)-β-D-Gal*p*, or both, for the second binding site it is not. Although recent findings of the involvement of a second binding site of hemagglutinin-neuraminidase of parainfluenza viruses in membrane fusion have been reported,[53], [54], [55] similar evidence for influenza hemagglutinin has yet to be found.

1.4. Multivalency

Weak interactions are commonly found in nature. There is hydrogen-bonding, hydrophobic interactions, electrostatic, and dipole-dipole-interactions, to name some of them. However, multiple interactions of the same kind have been found to be stronger than one. On a visible level this can be seen with the burdock: It has many hooks presented on its surface that may interact with multiple loops and loop-like structures of – for example – the fur of an animal. In the early 1940s this ability of the burdock even inspired the Swiss engineer George de Mestral to invent the Velcro hook-and-loop fastener, a strong and reversible closing mechanism that is made up of single weak interactions.[56], [57]

The same principle is true for the concept of multivalency that is used in biology and chemistry: The combination of low binding affinities lead to high binding avidity. Where the burdock or Velcro has hooks and loops, multivalency in biology has ligands and receptors. One of the most important examples for multivalency are carbohydrate–protein interactions, which are fundamental for a vast number of biological processes,

including fertilization, cell–cell and pathogen–cell interactions, and inflammation-related ones. However, in nearly every case of carbohydrate–protein interactions, single binding events are extremely weak, having dissociation constants in the millimolar range. For this reason it has to be assumed that the specificity and strength of many observed carbohydrate–protein interactions is coming from or is significantly improved by multivalency.

In fact, previously mentioned carbohydrate-binding proteins (*vide supra*) have already given a strong indication of the positive effect of multivalent receptor display. Since such receptors are of interest to the biomedical and life sciences, deciphering the molecular action principle that is underlying multivalency would strongly contribute to an improvement in the treatment and understanding of diseases. Initial attempts and major contributions to the case were achieved by the groups of Lee and Whitesides. Lee *et al.*, did not only coin the term "carbohydrate cluster effect", but also designed a multivalently binding glycoconjugate for asialoglycoprotein receptor and therefore contributed to the elucidation of its structure. [37], [58]–[60]The Whiteside lab on the other hand showed that multivalent ligand design can help in the development of antiviral agents.[59] The following section should be a brief introduction into to principles and theories of multivalency.

1.4.1. Types of protein–carbohydrate interactions

Monovalent interactions of carbohydrates with proteins are only of minor relevance because they often have low affinities. However, in combination with powerful structural elucidation tools, such as X-ray crystallography and NMR, even those weak interaction can help illuminating the underlying principles of carbohydrate–protein binding, either directly or by guiding . Fortunately, solvent exposed, extracellular CRDs are often lacking deep binding pockets, hence limiting the direct contact points of ligand glycoproteins and carbohydrate-binding receptor.[61] The following part will provide a brief overview of the different types of observed interactions.

1.4.1.1. Hydrogen bonding

When looking at the general structure of a carbohydrate, one may notice the high number of hydroxy group. Consequently it is expected that they take part in the recognition event. Although it is still a matter of debate whether hydrogen bonds between carbohydrates and proteins have a n energetic contribution to the binding event, it is clear that the specificity is connected to hydrogen bonding. Due to the arrangement of basic amino acids like arginine, lysine, asparagine, and histidine in the binding pocket, multiple direct interaction points can be established by the CRD with its ligand *via* hydrogen bonding. Evidence for that phenomenon was found int the crystal structure
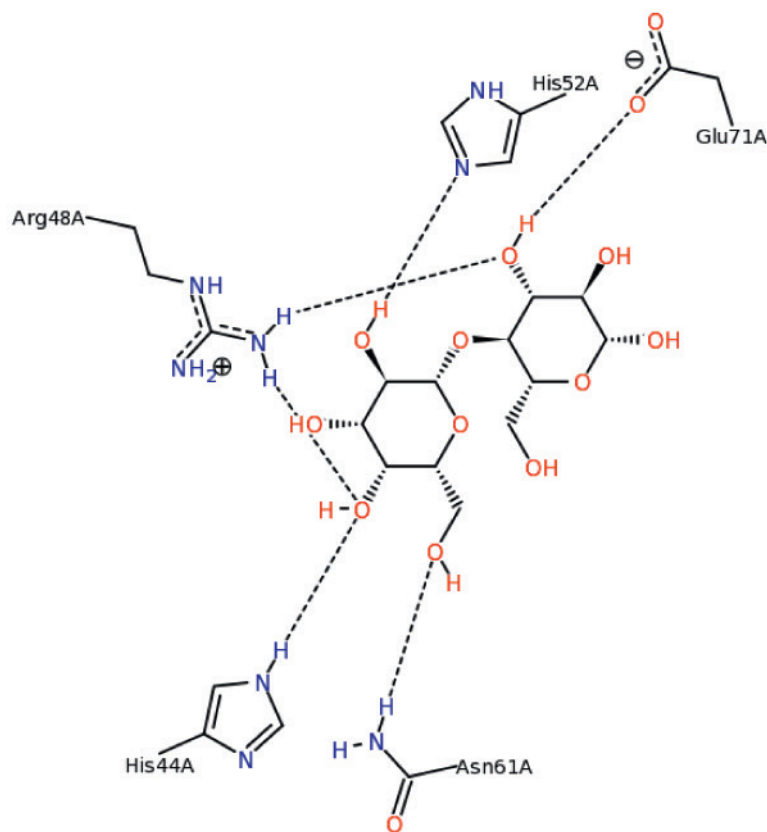
Figure 1-7: **Lactose complexed in the binding pocket of galectin-1.**

The hydroxyl groups of β-lactose interact primarily with basic amino acids by hydrogen bonding (black, dotted line). However, also the acidic amino acid (Glu71) is acting as an hydrogen acceptor for a hydroxy proton of lactose.

The image was generated by PoseView, Ref. [64]–[67], from PDB-entry 3M2M.

of galactin-1 complexed with its ligand lactose (**Figure 1-7**).[62], [63] Because of this, it is possible for the lectin to differentiate between certain carbohydrates that are often similar in structure. Often the sugar-binding protein uses multiple interactions of one hydroxyl group with a single amino acid to tell apart saccharides that are only distinct in the stereochemical orientation of one hydroxy group. Especially configurations at the C-3 and C-4, whether the hydroxy groups are equatorial–equatorial or equatorial–axial, were shown to be discriminated by a single amino acid. But also hydrogen bonds mediated by water (*vide infra*) are common next to the direct interaction of amino acid side chains with carbohydrates. Interestingly, if structural features are not of importance to the receptor, they are usually exposed to the solvent.

1.4.1.2. Hydrophobic interactions

The fact that saccharides are literally covered in hydroxy groups often leads to overseeing their nature as amphiphiles. A significant apolar surface area can therefore

interact with hydrophobic amino acids. In the binding pocket of the CRD it is the hydrogen bonds that arrange the saccharide in a way, that hydrophobic surfaces are aligned with each other and result in slightly tighter binding of the carbohydrate and the corresponding CRD.[68] Several observations have been made that conclude that carbohydrates functionalized with hydrophobic aglycones as well as small molecules that do not resemble carbohydrates at all can often bind to the CRD with higher affinity. This favorable interactions can be explained by the partial positivation of the aliphatic protons of carbohydrates, which can interact with the π-cloud of aromatic amino acid side chains.[69] In addition, a stabilizing influence of aromatic residues to the favorable conformation of a carbohydrate was been discovered, which explains their abundance in the binding pockets of lectin.[70]

1.4.1.3. Coulombic interactions

Within the group of biologically relevant carbohydrates and glycans, such as with Neu5Ac-bearing ones and heparin, anionic variants are often found. Thus it is no surprise that carbohydrate-binding proteins often differentiate between ligands based on the occurrence of charge. The binding can hereby either be mediated by direct hydrogen bonding, or through electrostatics of certain amino acids, like arginine. Examples for this would be P- and L-selectin binding. [61]

1.4.1.4. Calcium ions

For legume and mammalian C-type lectins, Ca^{2+} are a requirement for their activity. Concerning this, calcium often serve a double purpose: First and most important, the ion is orientating amino acid side chains within the binding cavity for optimal ligand presentation. In addition to this, calcium also assists in the binding itself, since it can be coordinated by C-3 and C-4 hydroxy groups of the carbohydrate ligand. For example, the complexing with calcium can be observed with mannose binding protein.[71], [72]

1.4.1.5. Water

Although the influence of water to the binding of carbohydrates and lectins has long been recognized, its main role and energetic contributions to binding are still not clear. X-ray crystallographic data clearly shows that water is mediating due to being both hydrogen bond acceptor and donor between carbohydrates and amino acid residues within the binding pockets of certain lectins. Especially the specificity of lectins to carbohydrates seems to rely on water molecules.[73] In those cases water often stabilizes favorable conformations of the ligand and the binding pocket by hydrogen bonding. [23], [74]–[76]

However, crystallography can not divulge energetic contributions of water molecules

to the binding itself. Whilst complexion was observed between water, lectin and carbohydrate, its sole energetic benefaction is a matter of debate. It has been argued that the reorganization of water is the main contributor to carbohydrate complexation. [59] One proposed model is based on enthalpic and entropic compensation. Thus, the hydrophobic surface of the carbohydrate is increasing the energy of water at the surface. Subsequently, the release of disordered water will be an decrease in entropy, but on the same side favorable for the enthalpy of the system. On the other hand, water molecules attached on the hydrophilic surface of the carbohydrate will be reorganized and disrupted, leading to an increase in entropy and a decrease in enthalpy. Therefore it was rationalized that the driving force for carbohydrate–protein complexation is the rearrangement of the solvation hull of water. However, the details of this binding interactions remains elusive.[77], [78]

1.4.2. Contributors to multivalent ligand binding

Multivalency is a complex biological phenomenon wherein multiple factors contribute to the increased binding strength displayed in multivalent systems. Although there are other multivalent interactions in nature, carbohydrate-protein interactions are the most prominent ones. Many lectins aggregate to oligomeric quaternary structures, thus are able to bind multiple carbohydrates (**1.3.1**, pp. 5ff., and chapter **4**, pp. 74ff.) at once in a multivalent fashion. Observations on single protein-carbohydrate interactions are helping, but they are not sufficient to fully understand the complexity of a biological system, wherein multiple receptor-ligand interactions are present. A lot of different physical-chemical mechanisms take part in the total increase in binding affinity of "multivalent" and "cluster" effects. Understanding of those individual parameters can help in the rational design of multivalent ligands for carbohydrate-protein recognition and in the process of elucidating biologically present multivalency.

1.4.2.1. Entropy

Entropy usually plays a crucial part in binding events. Thus, it is no surprise that it is also considered a necessary factor for multivalent interactions. Ligands that are presented in a polyvalent fashion will lose entropy with the first binding event. However, subsequent interactions will not have to pay an additional entropy penalty[79], [80] since it is also intramolecular rather than intermolecular.[81] Against this favorable entropic contribution stands the conformational entropy of the scaffold system. Different features are contributing to the actual conformational energy cost, such as rigidity of the scaffold and spatial orientation of the ligands.[80], [82]–[84] In addition, the entropic penalty for translation in the event of binding to a lectin is also related to its whereabouts. Membrane-bound receptors can only diffuse in two dimensions and

therefore would incur less entropic penalty than soluble lectins.

1.4.2.2. Chelate effect

Originally taken from organometallic chemistry, the "chelate effect" refers to the enhanced coordination of an electrophilic metal by multiple electron donors that were covalently fused to a scaffolding group. As described by Schwarzenbach in the 1950's, chelate means – as with multivalency itself – the cooperative interaction of several groups during the event of binding (**Figure 1-8**, p. 21).[85]–[87] Usually, chelation refers to single groups and a centrally coordinated molecule or atom, whereas multivalency in general refers to more complex interactions, such as carbohydrate-protein interactions in biology.[57] However, both phenomena rely on the principle of positive thermodynamic additivity.[79], [80] This means, that after the first binding is established, subsequent contacts are intramolecular rather than intermolecular considering entropic penalties and enthalpic gains (*vide supra*). Therefore the change in Gibb's free energy that is observed upon binding is in relation to the sum of all free energies for each contributing ligand-receptor interaction. Thus, considering this model, each additional contact would give an increase in the overall avidity of the system. However, this assumption takes not into account potential entropic penalties due to conformational changes of the linker structure and the ligands itself. Nevertheless it should be noted, that the energy cost of conformational entropy changes may not be as high [88]–[90] as proposed earlier. [80] Additionally it has been shown in the case of a rather "simple" system like EDTA and calcium, that enthalpy is the driving factor in complexation, rather than entropy.[91], [92]

1.4.2.3. Clustering of receptors

Another phenomenon in regards to multivalency is receptor clustering (**Figure 1-8**, p. 21). Since a lot of carbohydrate-binding proteins are actually membrane-bound receptors, they are anchored within the lipid bilayer or cell membrane. This location renders them incapable of using all three dimensions of spatial movement, but allows free two-dimensional diffusion within the cellular surface, which may lead to a clustering of receptors. Although there is an entropic penalty implied, receptor clustering helps binding of antennary gaped saccharides and glycoconjugates to receptors, without requiring exact receptor-ligand display on the cell membrane.[84]

1.4.2.4. Concentration

As stated before, a high local concentration of receptors can increase the observed avidity of a given system (**Figure 1-8**, p. 21). In particular this effect may not be a result of chelation mechanisms. As with receptor clustering, that also leads to improved
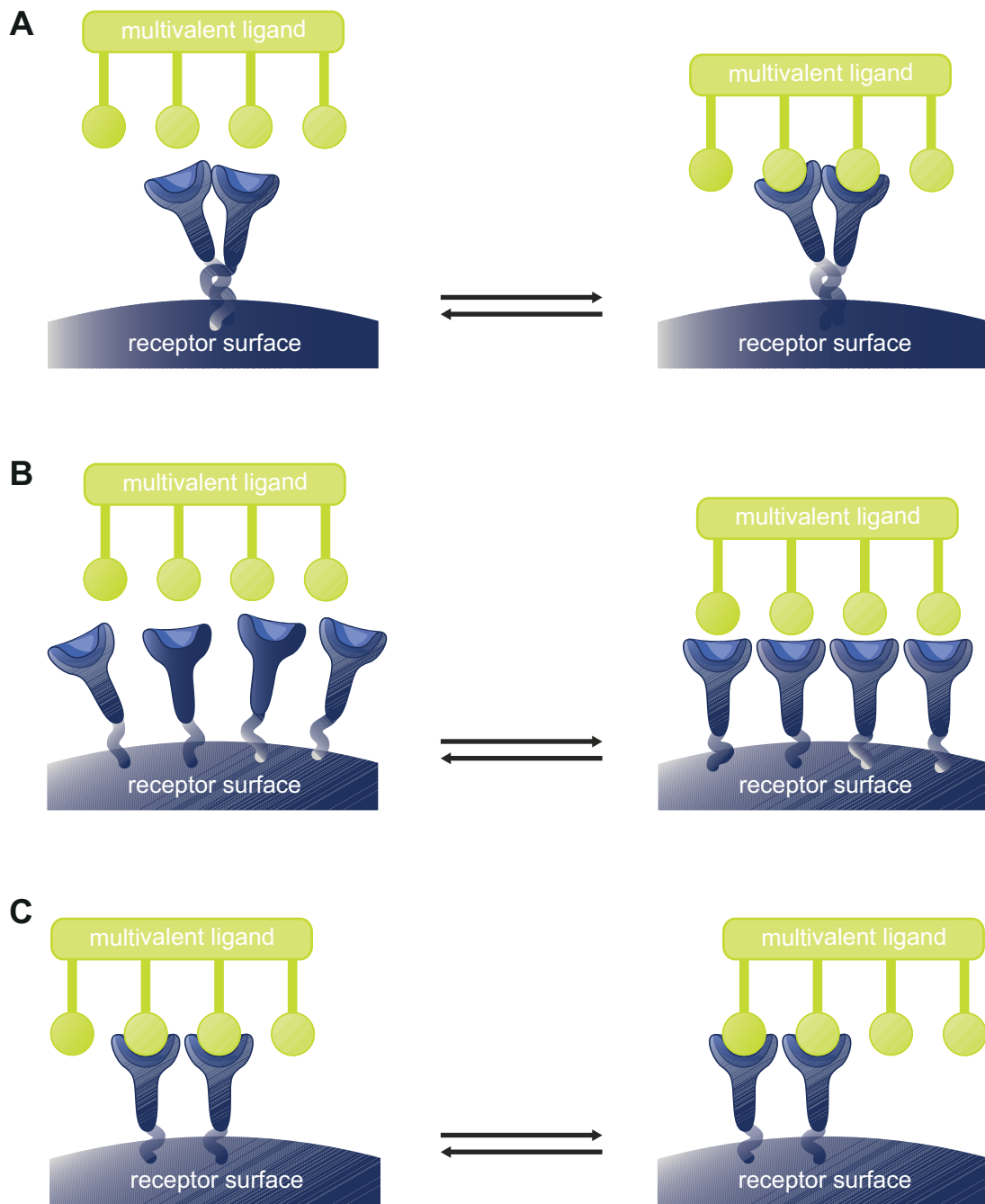
Figure 1-8:   Modes of multivalent binding.
Possible molecuar mechanisms for increases in functional affinities for multivalent ligands. (a)
Interaction of a multivalent ligand with a multivalent receptor (chelate effect). (b) Clustering of
receptors by a multivalent ligand. (c) Increased local ligand concentration (slow off-rates due to
statistical effects).
Adapted and redrawn from reference [61].

binding with monovalent ligands, a multivalent ligand also tend have increased avidity.
The reason for this can be depicted in the statistics of a two-step binding event. The
first stage of the binding would be the formation of an intermediate pre-binding

complex. From that on, the ligand can either form a stable receptor-ligand complex, or dissociate again to an unbound state. Considering a multivalently presented epitope, the probability to form a stable bound complex is higher than for the monovalent epitope. Thus, an increase in local concentration of a ligand presented to a receptor contributes to an increase in complex fromation, and therefore binding. A system, in which a combination of multivalent-binding receptor and a suitable multivalent ligand is present, may profit from even higher avidity. [93], [94]

Various papers on multivalent architectures describe increased binding per active ligand in different systems.[82], [84], [95]–[102] It is reasonable that to the increased local concentration of active ligands due to their high representation positively contribute to the observed avidity. However, influences of the spacer between ligand and scaffold on the multivalent binding could yet not be fully elucidated.

1.4.3. Methods to determine binding constants and events

To quantify and elucidate multivalent binding events, several experimental techniques have been applied. Each of those methods can determine different variables and features of the interaction, such as binding constants and enthalpies. It should also be noted, that those techniques usually are meant for complementary purpose, since they have unique strengths and limitations depending on the system they are applied to. The following section will briefly comment on the most popular techniques regarding elucidation of multivalent interactions.[103], [104]

1.4.3.1. Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) is a powerful technique with a wide applicability for determining thermodynamical parameters of a binding event.[105] As implied by its name, ITC measures the change in temperature upon complexation of a ligand titrated to a receptor, or vice versa (**Figure 1-9**, p. 23). All the changes are compared to a reference cell that needs to maintain equal temperature, hence the word "isothermal". Plotting the gained information about the change in temperature as a function of the concentration of the ligand relative to that of the receptor can then be analyzed to determine the enthalpy of binding, ΔH^0, and the dissociation constant, K_d. Furthermore, the Gibbs energy (ΔG^0), as well as the entropy of binding, ΔS^0 (through $T\Delta S^0 = \Delta H^0 - \Delta G^0$), can be deducted from those values. In principle, ITC is ideal for protein measurements since it is carried out under stable temperature and therefore able to neglect temperature-induced conformational changes. It is also very reliable with regards to the accuracy of data it is producing. Thus, it is no surprise that is a very common technique for elucidating thermodynamic contributions of multivalent binding phenomena. Unfortunately, the Achilles heel of this method is the material it is
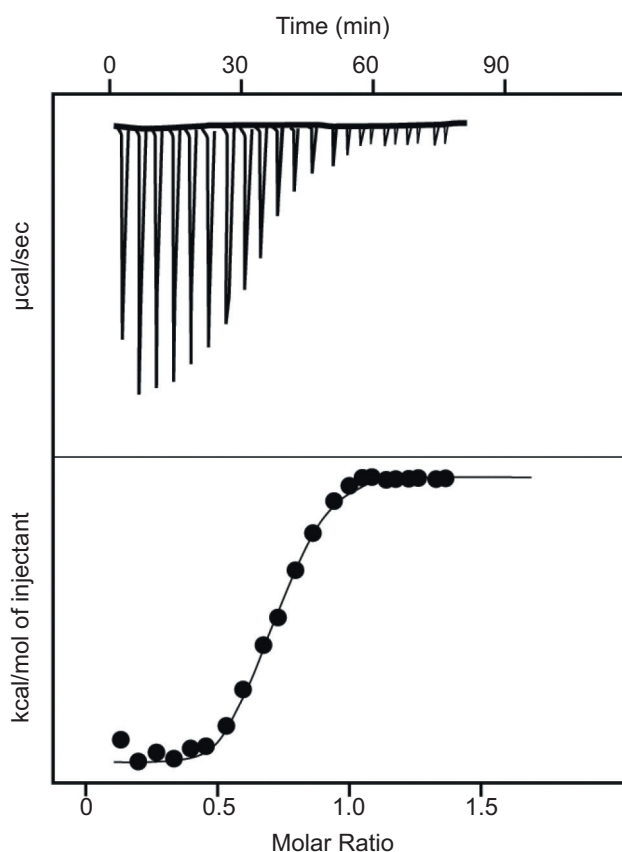
Figure 1-9: **The principle of isothermal titration microcalorimetry (ITC) explained on the example of galactose binding protein**

The ligand (e.g., a glycan) is injected with increasing amounts to a fixed value of receptor (e.g., galactose-binding protein, GBP) inside the measuring cell. Upon binding, heat is produced that is measured as µcal/sec. The obtained value of total kcal/mole of injected ligand relative to the molar ratio is plotted. Thermodynamic parameters of the binding – such as the K_d of interaction between the glycan and the GBP – can then be directly obtained from the plotted data .

From: Essentials of Glycobiology. 2nd edition. Varki A, Cummings RD, Esko JD, et al., editors. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 2009; Copyright © 2009, The Consortium of Glycobiology Editors, La Jolla, California.

consuming for a solid analysis, especially in the case of generally weak interactions,[106]-[108] such as carbohydrate-protein binding.

1.4.3.2. Nuclear magnetic resonance techniques

Another valuable and powerful tool to study protein-ligand interactions is nuclear magnetic resonance (NMR) spectroscopy. Depending on the experimental setup, either receptor or ligand signals will be the focus of observation. The shape – frequency and width – of obtained signals are also in strong dependency of the binding equilibria between ligand and receptor. When assuming a two-state equilibrium, only a free and

a complexed state are possible. Intrinsic NMR characteristics, such as chemical shifts, relaxation, translation and diffusion coefficients, of either receptor or ligand remain the same in an unbound, state. However, in the presence of each other and especially in the event of binding, the ligand molecule's NMR parameters are changing in favor to the – usually bigger – receptor. The receptor on the other hand, can be subject to conformational and therefore environmental chances, thus modulating its parameters as well. Since the event of binding is a thermodynamical phenomenon, ΔG, ΔS, and ΔH can be determined, as well as ligand-receptor ratios. [109]–[111], [112], [113]

1.4.3.3. Surface plasmon resonance

Surface plasmon resonance (SPR) is a sensitive technique that enables the measurement of kinetic and thermodynamic parameters of a binding event. The principle of this method is based upon the transfer of elecromagnetic energy (light) to electrons of a metal surface, usually gold, overflown by a solution. As the light is passing through a prism, it gets monochromatically polarized and subsequently partially reflected from the gold surface. A part of the light is penetrating the gold surface and forming evanescent waves. Those waves can then excite the plasma waves on the surface of the gold layer, thus creating "surface plasmons" that can travel along the metal surface. For the maximum absorption, the transfer of momentum would match those of the plasmons, thus creating a resonance condition. At this point, the reflected light is strongly losing intensity. Since the evanescent waves decay exponentially with distance, extremely thin layers of gold are directly evaporated on a glass surface, enabling a dependency of the resonance condition not only on the gold layer but also on its surroundings. Thus, SPR is highly sensitive to changes of the gold-solution interface. To make use of this sensitivity for the research on biological phenomena – including multivalency – the gold surface of the SPR "chip" is coated with a dextran matrix. The matrix is suitable functionalized to allow immobilization of a receptor on the gold-dextran surface. Subsequently an analyte within flow of the solution is passed over the surface, changing the refractive index upon interacting with the surface. This process is monitored in real-time and plotted into a "sensogram" (**Figure 1-10**, p. 25). The intensities obtained from measured responses are proportional to the mass that is bound to the surface.[103], [104], [114]–[116]

With SPR kinetic rate constants of the binding event can be determined, thus dissociation constants and affinity can also be calculated. In addition, assays that can provide empirical parameters, like IC_{50}, are well suited and quite reliable for SPR measurements. The applicability of IC_{50} values determined by SPR for multivalent systems is especially useful to compare different degrees of functionalization or complete systems with each other. If the number of presented epitopes on a multivalent scaffold is standardized to the monovalent ligand, the enhancement factor β can be determined. IC_{50} and
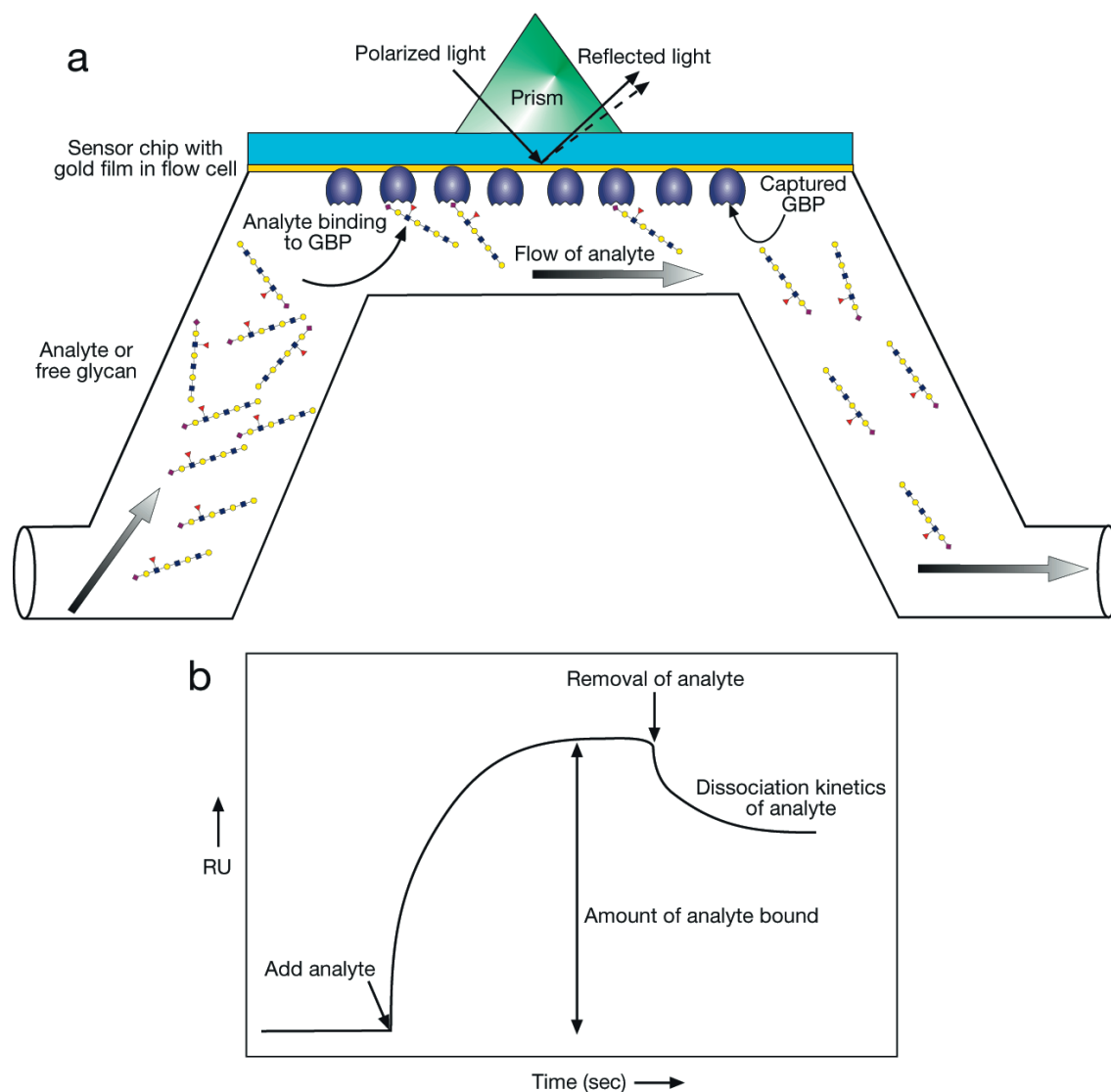
Figure 1-10: The principle of surface plasmon resonance (SPR) explained on on the example of galactose binding protein.

(a) In SPR, the alteration of reflected light on a metal surface is measured. This response changes upon the binding of the ligand (analyte) in the flow cell to the immobilized receptor. The data is then transforred into a sensogram (b), showing the binding of the analyte to the surface. Thereby the kinetics of binding and dissociation can be assessed. RU indicates resonance units.

From: Essentials of Glycobiology. 2nd edition. Varki A, Cummings RD, Esko JD, et al., editors. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 2009; Copyright © 2009, The Consortium of Glycobiology Editors, La Jolla, California.

competitive binding assays are therefore a good tool for assessing the effectiveness of a presented multivalent system relative to another. [117]–[120]

1.4.3.4. Fluorescence-activated cell sorting

From all the techniques described in here (*vide supra*), fluorescent-activated cell sorting (FACS) can give the least direct information on thermodynamics of a multivalent binding event. However, it is a solid, well-established technique and especially useful when dealing with endocytotic receptors and cell membrane-associated lectins. It is not only a tool for quantification of events, but also for the statistical evaluation of many biological processes.[121], [122] Taking this into account, FACS can be a valuable tool for the evaluation of multivalent binding events and for the analysis and rationalization of data assessed by biophysical methods like SPR and ITC.[123], [124] There are numerous examples where FACS was used to quantify multivalent binding. Because of the good relation of fluorescent intensities to binding events, methods are described for the deduction of thermodynamic parameters from FACS data.[125]

In principle, the functioning of FACS relies on the different scattering of light and the emission of fluorescent active dyes and compounds. It is widely applicable, demanding only the prerequisite of cells that are fluorescently labeled in one way or the other. The FACS analyzer usually dilutes cells with a buffer solution and let them flow through a detector one by one, where a laser is set on them. The light emitted by the laser is reflected and scattered and subsequently analyzed by various detectors, deducting, for example, cell size, granulation and specifically fluorescent emissions. The cells are than sorted and counted, revealing statistical data about the experiment (**Figure 1-11**, p. 27).[126]

1.5. Chemistry of carbohydrates and glycoconjugates

Carbohydrates represent one of the most abundant and important class of molecules for biochemistry, biology, and related disciplines. Apart from being an energy source for heterotroph organisms, they play a crucial role in many physiological and biological events (*vide supra*, **1.3**, pp. 4ff.). To address questions that arose in glycoscience – like multivalent interactions between carbohydrate and proteins – synthetic organic chemists, and later biologists through emerging chemoenzymatic techniques, developed various strategies to generate complex oligosaccharides and glycoconjugates.[9] The main problem was encountered in the regio- and stereoselective formation of glycosidic bonds, either between mono- and oligosaccharides, or aglycones.[9], [127], [128] Over the time different methods have emerged to solve this problems. Many of those methods dealt with the functionalization of the anomeric center by varying leaving groups and activating reagents. In addition extensive research was conducted on the influence of protecting groups, resulting in the armed-disarmed approach, and on the participation
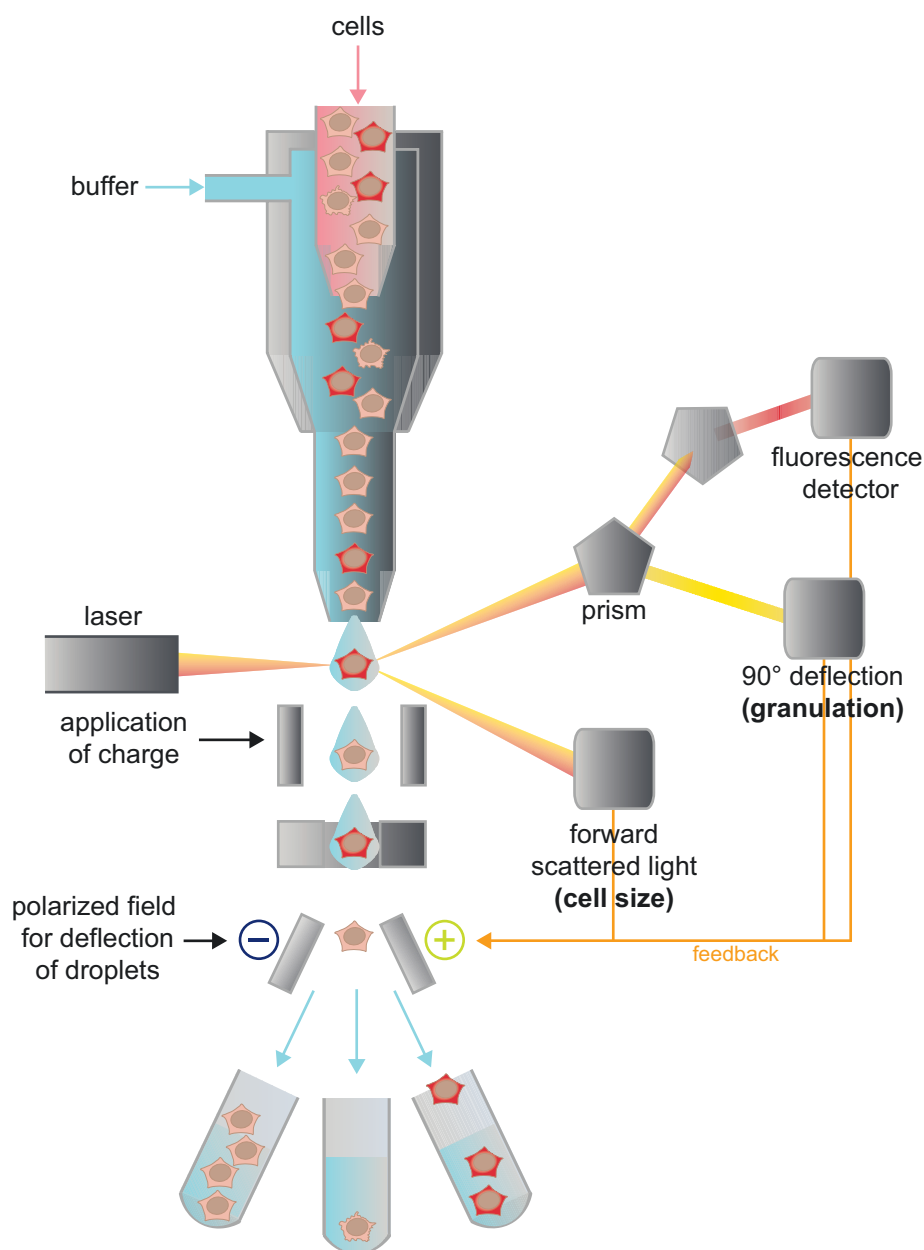
Figure 1-11: The principle and function of a fluorescence-activated cell sorter (FACS).
A mixture of fluorescent-labeled and normal cells is analyzed by this devices through a constant, droplet-forming flow. When passing though the system, a charge is applied to every cell depending on their fluorescence read-out. Subsequently the charge is used to sort and count the cells. Ideally only one cell should be detected per droplet and therefore giving valid statistical data.

of the solvent during glycosylation reactions. The following section gives a brief overview about common approaches for the synthesis of simple glycoconjugates. It will mainly deal with the stereoselective functionalization of readily available monosaccharides in the form of pyranosides and *O*-glycosidic linkage with aglycones. [9], [127], [128]
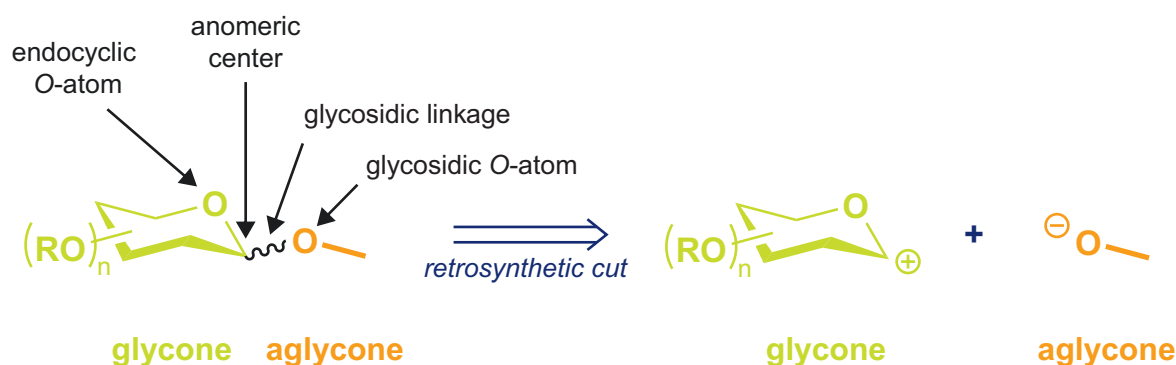
Scheme 1-1: Minimalistic schematic of an *O*-glycoside and its synthons.

1.5.1. Glycosylations at the anomeric center

As stated before, the synthesis of glycosides and glycoconjugates plays a crucial role in answering and addressing questions of glycoscience. Although carbohydrate chemistry as a discipline of organic synthesis deals also with the generation of *C*-, *N*-, and *S*-glycosides, those classes of glycosides will not be the focus of this text. However, the synthesis of *O*-glycosides and glycoconjugates with an *O*-glycosidic bond will be discussed briefly.

Basically, an *O*-gycoside is a mixed acetal with a cyclic "glycone" on one side of the glycosidic linkage and an "aglycone" on the other (**Scheme 1-1**). Usually the glycone of an *O*-glycoside is either a pyranoside – a 6-membered ring – or a furanoside – a 5-membered ring. It bears the anomeric center, which is the carbon of the acetal, whereas the aglycone can be derived from any hydroxylated compound. [9], [127], [128]

From a retrosynthetic point of view, the glycosidic bond can be cleaved to give rise to two synthons: the electrophilic "glycosyl donor" and the nucleophilic "glycosyl acceptor" (**Scheme 1-1**). When the glycosylation reaction happens and the glycosidic bond is formed, the hemiacetal – the glycone part – is converted to the acetal, whereas the alcohol – the aglycone – is glycosylated in the same process. It should be noted, that the nomenclature and configuration of the glycosidic bond is always indicated by the glycosyl donor, and therefore the anomeric center. [9], [127],

Although the glycosylation is formally a condensation reaction, the biological and chemical process to regio- and stereoselectively form a glycosidic bond is far more complex. Emil Fischer, the pioneer of biological chemistry,[130] invented a very straightforward method, nowadays coined Fischer glycosylation, where he heated carbohydrates together with excess of alcohol under acidic conditions to form simple glycosides. Since the remaining alcohol had to be removed after the reaction, this strategy is limited to alcohols with a low boiling point. In addition, the anomeric configuration of the glycoside product is only thermodynamically controlled, yielding a mixture of the most stable anomers.

Those features of the Fischer glycosylation are severely limiting the applicability of the reaction to few glycosyl acceptors, hence different methods for the synthesis of more complex glycosides, such as disaccharides, and glycoconjugates are required. To meet the requirements of a more controlled approach, only one hydroxyl group has to be accessible for the glycosylation, and the cyclic hemiacetal of the glycone moiety has to have a selective reactivity. [9], [127]

Therefore synthesis of more complex glycosides and glycoconjugates demands an adequately protected glycosyl donor equipped with a suitable leaving group at the anomeric position, that can be selectively activated by a specific promoter to allow participation in the stereoselective formation of a glycosidic bond with a glycosyl acceptor. Depending on the nature of the glycosyl acceptor, it might also be partially protected to lead to the desired product. 9], [127]–[129][128]

As of today, many different protocols have been published and their numbers are still rising. However, the outcome of a glycosylation reaction is often dependent on the glycosidic bond that is formed and the most suitable method is usually hard to predict. Hans Paulsen commented this phenomenon with the words: "Each oligosaccharide synthesis remains an independent problem which resolution requires considerable systematic research and a good deal of know-how. There are no universal reaction conditions for oligosaccharide synthesis." It is therefore no surprise that their is an increasing number of glycosyl donors equipped with all kinds of leaving groups, ranging from halides and acetates, over trichloroacetimidates and thioethers, to pentenyl and vinyl.

Regardless of the sheer amount of different glycosylation methods that have been applied until now, most syntheses of *O*-glycosides and glycoconjugates, however, can be realized with the three most prominent donor systems in carbohydrate synthesis:

- *Glycosyl halides* in **Koenigs-Knorr** type reactions

- *Glycosyl thrichloroacetimidate* or "**Schmidt** donors"

- *Thioglycosides*, or other stable glycosyl donors such as *n*-pentenyl

A comprehensive list of the leaving groups and examples for promoters of these common glycosylation reactions are summarized in **Table 1-1**, p. 30. Here, it should be noted that thioglycosides and *n*-pentenyl donors are ideal donor systems for consecutive glycosylation reactions since they are stable towards many reaction conditions until activated by a suitable promoter. [9], [127]–[129]

As mentioned before, the biggest challenge in *O*-glycoside and glycoconjugate synthesis is the stereochemical outcome of the reaction. This means one has to control the regiochemistry and the configuration of the newly formed glycosidic bond. The first part is usually achieved by a proper protecting group strategy, the latter is dependent on applied reaction conditions, like solvent, participating protecting groups, and the

Table 1-1: Common leaving groups and promoter for *O*-glycosylations

Glycosyl donor	Promoter
LG = OAc	$BF_3 \cdot OEt_2$, $SnCl_4$, TMSOTf
LG = Br	$AgCO_4$, AgOTF, $Hg(CN)_2$
LG = Cl	AgOTf, $Hg(CN)_2$, $HgBr_2$
LG = F	$SnCl_2$–AgOTf
LG = $OC(NH)CCl_3$	$BF_3 \cdot OEt_2$, TMSOTf
LG = SR	TfOH–NIS, DMTST, IDCP, etc.
LG = $O(CH_2)_3CH=CH_2$	NBS/NIS–TESOTf, IDCP, etc.

choice of donor-promoter method. If the reaction is proceeding with a neighboring participating protecting group at the C-2 position of a pyranoside, usually the 1,2-*trans* glycoside is formed stereospecifically. This is in contrast to glycosidations where there is non-participating group installed. In such cases both 1,2-*trans* and 1,2-*cis* glycosides can be formed, usually with lesser stereoselectivity. This phenomenon as well as more details about the previously mentioned common glycosylation methodologies are explained briefly in the following sub-sections.

1.5.1.1. General aspects of *O*-glycosidations

Although all details of glycosylation reactions have yet to be elucidated, the proposed principle has widely been accepted by the research community. Basically, glycosidations are nucleophilic displacements on the anomeric center of the glycone. Since the reaction takes place at a secondary carbon applying weak nucleophiles (alcohols and or other carbohydrates), it is often an unimolecular mechanism, hence S_N1. Most of the herein described methods utilize a glycosyl donor that provides an intermediate oxocarbenium ion or a related species. Depending on the nature of the C-2 protecting group, either the 1,2-*trans* (α in D-hexoses) or 1,2-*cis* glycoside (β in D-hexoses) is preferentially formed (**Scheme 1-2**, p. 31). [9], [127]–[129], [131], [132]

1,2-*trans* Glycosidation usually applies participating neighboring groups, such as acyl moieties, for the protection of the C-2 hydroxyl or amine. The intermediate oxocarbenium ion that is formed after the activation is then leading to an acyloxonium ion intermediate due to an anchimeric effect of the neighboring group. Since the bottom side of the pyranoside is now shielded due to the dioxolane ring, a nucleophilic attack from the top face is strongly favored, leading to *trans*-cleavage of the fused ring and a 1,2-*trans* glycosidic bond (**Scheme 1-2**, p. 31). However, due to reaction equilibria and low reactivity of nucleophiles, also 1,2-*cis* glycosides can be formed as a minor product. In addition orthoester formation can be observed when the nucleophilic attack occurs on the dioxolane carbon. In this case it either isomerizes to the respective
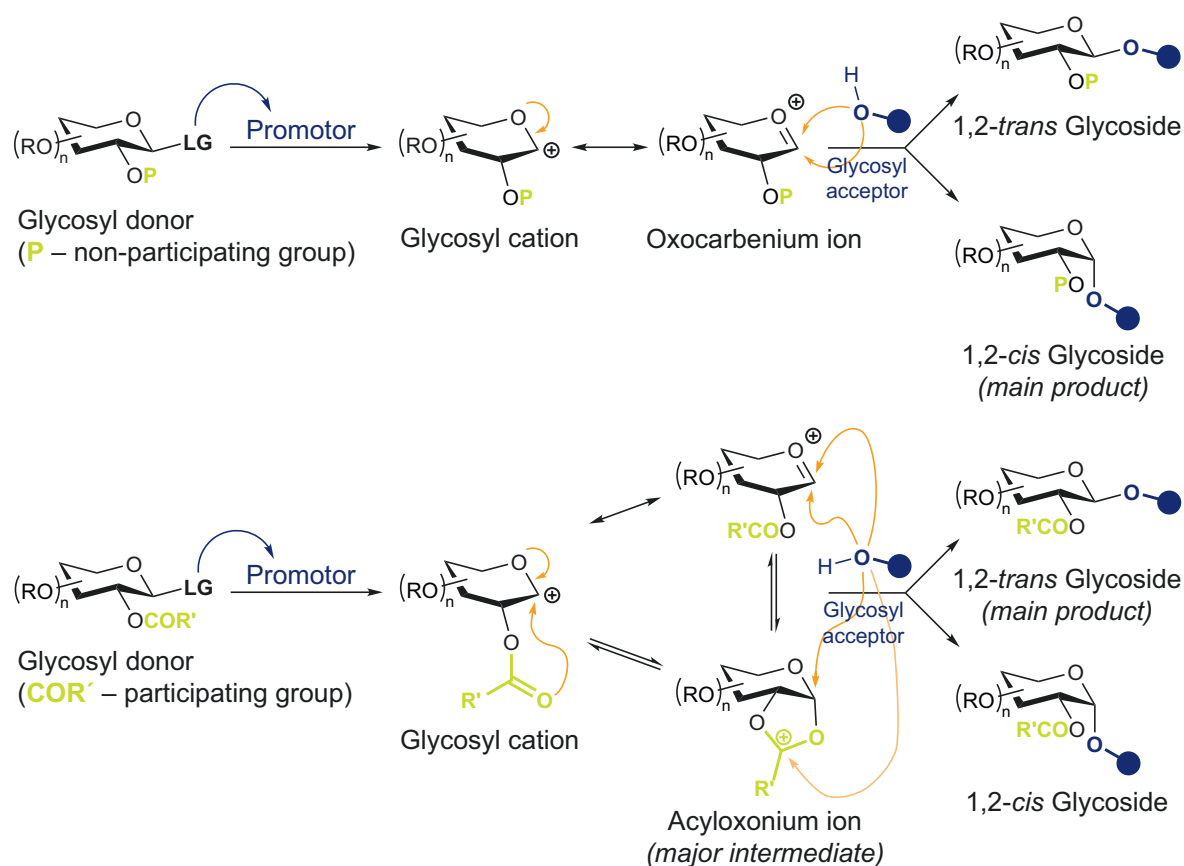
Scheme 1-2: General considerations on *O*-glycosidic bond formations to form 1,2-trans or 1,2-cis glycosides.

1,2-*trans* glycoside, or the orthoester stays as a stable reaction product. The latter route occurs when under neutral and basic conditions. However, 1,2-*trans* glycosides are often obtained *via* acyloxonium intermediates in a straightforward way as the major product.[127]–[129]

1,2-*cis* Glycosidation is somewhat harder to achieve in a controlled, stereospecific manner. Since a participating neighboring group would mainly lead to the 1,2-*trans* product, it has to be omitted. Potential non-participating groups are therefore ethers for hydroxyl and azides for amino moieties. In addition to this, the reaction at the anomeric center should ideally occur in a S_N2-type reaction *via* a tight ion pair complex. Lemieux and co-workers[133] were able to realize this challenge by applying tetrabutylammonium bromide to an -glycosyl bromide. *In situ* anomerization of the more stable α-halide to the more reactive β-halide then provided the final α-glycoside in an acceptable proportion. In spite of this achievement it should be noted that this approach is not widely applicable and non-participating C-2 protecting groups tend to give rise to anomeric mixtures that often can be hard to separate. Nevertheless, the ratio of formed products can often be influenced by the used activator. For example, silver triflate proved to be beneficial in the case of glycosyl bromides, whereas silver oxide could effectively be employed with

the synthesis of 2-amino-2-deoxy-α-ᴅ-glycosides.[127]–[129]

In additon, **solvent effects**, can have a dramatic impact on the anomeric ratio of a glycosidation. In carbohydrate chemistry, the stereodirecting effects of some solvents can be exploited to improve the outcome of glycosylations that proceed without participating neighboring groups, thus under S_N1-type conditions. Especially the influence of ethers and acetonitrile was studied extensively. [134], [135] Depending on the choice of solvents, glycosylation reactions in absence of a participating protecting group at C-2 can either favor the formation of 1,2-*cis* glycosides (ethers) or 1,2-*trans* glycosides (acetonitrile). The reason for this phenomenon lies within the different reaction pathways the solvents are providing. Ethers preferentially interact with the oxonium ion from the α-face, presumingly forming another intermediate, a β-ether-glycosyl-oxonium ion, due to the anomeric effect[133]. Thus, a following attack of the nucleophile is more likely to happen from the bottom side of the ring, resulting in a 1,2-*cis* glycoside. [134] Acetonitrile on the other side favors the formation of the 1,2-*trans* glycoside. It is believed that this is due to the so-called "nitrile effect" where an axial α-acetonitrilium ion is formed as an intermediate. Since it is shielding the bottom side of the ring, a nucleophile can attack from the top, yielding the 1,2-*trans* glycoside. This early hypothesis has been backed up with experimental evidence by trapping the intermediate α-acetonitrilium ion with 2-cholobenzoic acid to form the corresponding amide with a 1,2-*cis* configuration. [9], [127]–[129]

Interestingly, this early finding is still the only validation of this theory, since research could yet not provide solid experimental data by NMR.[135] Even so, recent theoretical studies and calculations rationalized the phenomenon of solvent effects with differences in conformations and counterion distributions of the oxonium ion intermediates.[136]

Although the applicability of acetonitrile as preferable solvent to form β-anomers was demonstrated with a variety of different glycosyl donors, exploiting the effect for β-mannosides or β-ʟ-rhamnoses[137] was not successful.

Concluding from the literature, [9], [127]–[129], [138] a lot of factors have influence on the stereochemical outcome of an *O*-glycosidation, including:

- the configuration of the glycosyl donor
- the choice of protecting groupson every protected hydroxyl or amino group of the glycosyl donor
- the selection of the leaving group
- the applied promoter system to activate the glycosyl donor
- effects of the solvents used in the glycosidation
- the type of glycosyl acceptor
- other reaction conditions, such as temperature and pressure.

In the case of simple glycosylation reactions, this list can often partially be omitted. This is because peracetylated carbohydrates, for example penta-O-acetyl-β-D-galactopyranose, work well as a easily accessible glycosyl donor that can be activated by Lewis acids, like BF$_3$·OEt$_2$, to yield the desired O-glycoside or glycoconjugate stereoselectively in β-configuration (*vide infra,* **10.2.2.3**, pp. 122ff.). However, the limitations of this methodology to certain types of carbohydrates and to simple acceptors, makes other glycosyl donors more desirable. Therefore the carbohydrate chemist of today is in most cases advised to use glycosyl donors like *halides* (Koenigs-Knorr type reactions), *trichloroacetimidates*, and *thioglycosides*. [9], [127], [128]

1.5.1.2. The Koenigs-Knorr method

The application of glycosyl halides for the synthesis of O-glycosides and glycoconjugates is long known. Back in 1879, the famous Arthur Michael reported a glycosylation were he used "acetochlorhydrose" (tetra-O-acetyl-α-D-glucosyl chloride) to synthesize a phenol glycoside.[139] However, it took an additional two decades until Wilhelm Koenig and Eduard Knorr came up with their methodology. By applying the more reactive tetra-O-acetyl-α-D-glucosyl bromide, they were able to glycosylate it with methanol with the addition of silver salts.[140] Over the decades numerous variations and improvements have been made to this reaction. The silver salts used in the reaction can either be insoluble like silver oxide (Ag$_2$O) and silver carbonate (Ag$_2$CO$_3$), or soluble like silver triflate (AgOTf) and silver percholrate (AgClO$_4$). The addition of desiccants or molecular sieves is also quite common. The most influential improvements were made by Helferich *et al.*[141], [142] with the combination of mercury salts as promoters, such as HgBr$_2$/HgO or Hg(CN)$_2$, good selectivities and yields can be achieved. The applicability of the Koenigs-Knorr reaction in general has also been demonstrated with the challenging synthesis of β-mannosides. By applying insoluble silver silicate or silver zeolite as a promoter to the reaction, a 2-O-benzyl-α-D-mannopyranosyl halide could be converted to the corresponding β-anomeric glycoside.[143] Astonishing is the simplicity of this variation of the Koenigs-Knorr reaction: A presumably simple "push-pull" mechanism effectively shields the bottom of the glycosyl donor, guiding the nucleophile to produce the β-mannoside.

Although Koenigs-Knorr type reactions are a solid choice for glycosidations, they suffer from some basic drawbacks. The fact that glycosyl halides are pretty labile donors, occasionally leading to glycal formation or hydrolysis, and with β-glycosyl bromides as the most unstable example due to the anomeric effect can be considered a limiting factor. In addition, the use of heavy metal salts as promoter in equimolar amounts is not desired in a standard reaction either. However, the ease of accessibility of glycosyl halides as donor molecules for glycosylation proved to be a strong advantage of the

reaction. Especially when considering that they can sometimes even be generated directly from the unprotected carbohydrate.

1.5.1.3. Glycosyl trichloroacetimidates

Described first by Richard R. Schmidt[144] for the use in carbohydrate chemistry, trichloroacetimidate donors are a common choice for the synthesis of *O*-glycosides, simple and complex, as well as glycoconjugates. They have a somewhat different advantages and disadvantages compared to the previously mentioned Koenigs-Knorr method and its glycosyl halides. One of the biggest drawbacks is probably the accessibility of the donors themselves. Whereas glycosyl halides are converted in one-step, the synthesis of trichloroacetimidates involves several steps, including deprotection of the anomeric center, followed by base-treatment and trichloroacetonitrile to yield the corresponding glycosyl trichloroacetimidates. However, since both anomers can be obtained by applying different reactions and are fairly stable at low temperatures, the disadvantage of a longer synthesis seems negligible. Also the fact that glycosyl trichloroacetimidates can generally be activated very mildly with $BF_3 \cdot OEt_2$ and TMSOTf proved advantageous. Another strength of these "Schmidt donors" are their versatility. Depending on the Lewis acid applied, glycosidations result in different sterochemistry. Briefly, trichloroacetimidate donors with a non-participating group at C-2 usually yield glycosides with inversion of configuration at the anomeric position when treated with a mild promoter such as $BF_3 \cdot OEt_2$. By application of a stronger promoter, like TMSOTf, solvent effects and reaction temperature have an increasing influence of the stereochemical reaction outcome. Whereas in ether and dichloromethane the α-D-glycoside is formed predominantly, nitrile solvents like propionitrile and acetonitrile strongly favor the formation of the β-D-glycoside. However, with acyl-protecting groups on the C-2 position, the observed product is the β-D-glycoside.

Glycosyl trichloroacetimidate glycosidations are often superior to Koenigs-Knorr type reactions. They often work with only catalytic amounts of Lewis acid and resulting in good yields. The general applicability for the synthesis of oligosaccharides, simple *O*-glycosides, as well as glycoconjugates and other carbohydrate based derivatives, makes them a proper, highly versatile choice for modern carbohydrate chemists. The only glycosyl donors that are probably even more flexible in usage are thioglycosides.

1.5.1.4. Thioglycosides

Thioglycosides are derivatives of carbohydrates where the linking oxygen between glycone and aglycone is replaced by a sulfur atom. They are fairly stable and are commonly prepared directly from peracetylated saccharides with $BF_3 \cdot OEt_2$ as promoter to give predominantly rise to the 1,2-*trans* thioglycoside. The aglycone part of this glycosyl

donor is either an alkyl or aryl moiety and is often dependent on the glycosylation strategy. Probably the greatest strength of thioglycosides is their stability in absence of a thiophilic promoter. Whereas other glycosyl donors are not stable towards protecting group exchange (deprotection–protection), acetylated thioglycosides survive deacetylation conditions with ease. Thus thioglycosides can both function as efficient leaving groups and temporary protecting groups at the anomeric carbon, rendering them very useful for more demanding carbohydrate syntheses, for instance in the generation of oligosaccharides, amino sugars, sialosides and other challenging hexoses. [9], [127], [128], [145]–[149]

In addition, thioglycosides can be converted to other glycosides with ease and great variety. It has even been shown, that they can directly be used for Koenigs-Knorr or halide catalysis through *in situ* generation of the corresponding halides. However, thioglycosides are not only percursors for halide-employing reactions, but can be activated with lots of different thiophilic promoter systems. Hereby, a common approach is the addition of *N*-iodosuccinimide to generate an iodonium ion under acidic conditions to interact with the sulfur to furnish an intermediate sulfonium ion. This intermediate is parent to the active carbocationic glycoside which can further react with a nucleophile to form a glycosidic bond. [127], [145]–[149]

The glycosylations of thiogylcosides follow similar principles regarding the stereoselective outcome as in other glycosidation reactions, such as transformations with thrichloroacetimidates. Thus the reaction outcome is more predictable and controllable. In short, thioglycosides with non-participating groups at the C-2 position will result in 1,2-*cis* glycosides, whereas 1,2-*trans* glycosides are the product of participating groups. Solvent effects may influence the stereochemical outcome of the glycosylation reaction. In contrast to previously mentioned glycosyl donors, the combination of temperature and promoter has a stronger influence on the reaction outcome, since it has been suggested that intermediate glycosyl triflates may be formed during the course of the reaction. At low temperatures they are presumingly more stable and tend to anomerize at elevated temperatures, thus changing the stereochemical outcome of the formed glycosidic bond.

In addition to the various types of available thio-based leaving group–promoter systems, changes in the reactivity of the thioglycosides depending on the C-2 substituent gave rise to interesting concepts. Especially the "armed/disarmed" approach proved to be valuable for the synthesis of complex oligosaccharides by sequential and one-pot synthesis. [145]–[149]

Nevertheless, as with all strategies in carbohydrate chemistry, individual optimizations for specific carbohydrate reactions are also necessary for this donor system. Especially in the synthesis of so-called challenging carbohydrates, such as β-mannosides and α-sialosides, thioglycosides are often applied.
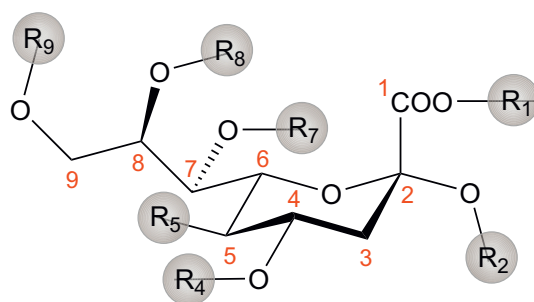
Figure 1-12:   The structural diversity of sialic acids.
The 9-carbon-chain is displayed in a chair-form and as the α-anomer. The residues (R1-R9) can be different moieties, e.g. acetylation (all), sulfation (C-8), methylation (C-8) and phosphorylation (C-9).

1.5.2. Sialic acids – challenging carbohydrate synthesis

Sialic acids (Sias) are 9-carbon-chain carbohydrates that are mostly found in organisms of the deuterostoma lineage, which includes vertebrates and partially invertebrates, but also in *Drosphila* and some bacterial strains. [150]–[153] Structure-wise, Sias are a family of very diverse 2-keto-3-deoxy-nononic acids. Probably the most prominent representatives are the C-5-amino modified neuraminic acids, either *N*-acetylated (Neu5Ac) or *N*-glycolylated (Neu5Gc), as well as the non-modified 2-keto-3-deoxy-nononic acid (Kdn). Nature often introduces even more variety on the glycerol chain of C-7 to C-9, such as acetylation (all), sulfation (C-8), methylation (C-8) and phosphorylation (C-9) (**Figure 1-12**).[10] Besides those modifications, various glycosylations can also be found. Their biological relevance is also accounted due to their presentation in *N*-glycans and *O*-glycans whereas Sias are often attached to the terminal, non-reducing end of those complex carbohydrates. Usually those linkages are either α(2→3), α(2→6), α(2→8) or α(2→9). As can be seen from the combinatorial possibilities of Sias, nature reserved a place of biological importance for them. This is even more underlined by the exploitation of receptor–sialic acid interaction of pathogens like viruses and bacteria. [154]–[156]Due to their biological and medicinal importance, obtaining highly defined α-sialic acid derivatives that resemble the natural linkage, is from major importance to the field since their discovery approximately 70 years ago.

Carbohydrate chemistry tries to mimic natural occurring glycans or part of it by chemical means. This is no different with sialic acid chemistry. Formally, a condensation reaction between the leaving group bearing anomeric carbon of the glycosyl donor and the alcohol moiety of the glycosyl acceptor occurs (**Scheme 1-1**, p. 28). This usually occurs under highly-controlled stereo selectivity with monosaccharides bearing neighbouring group that can affect the reaction outcome (**Scheme 1-2**, p. 31). Sialic acids are lacking those auxiliaries. Therefore stereo control has to be achieved by different

means. In addition to the challenge of selectively obtaining α-sialosides, elimination reactions that form the glycal also have to be avoided (**5.3.1**, pp. 95ff.). For 40 years carbohydrate chemists are tackling those problems by various means. Around the year 2000 two principal approaches arose.[157] The direct chemical sialylation primarily deals with probing different leaving groups, to achieve the final glycosylation in one step. In contrast, the indirect approach introduces an auxiliary group on the C-3 position to use neighbouring effects to guide the stereo selective outcome. A lot of conclusions were drawn from those investigations, leading to a huge variety of leaving groups for the C-2 position that include amongst others halogens, thioalkanes, thioaryls, xanthates, phosphites and phosphates. In the past years the attention shifted towards the C-5 position of Neu5Ac. The groups of Takahashi, De Meo and Crich conducted extensive research on that,[158]–[163] applying a new viewing angle on stereo control in sialic acid chemistry. For a more detailed and historical view on this topic, the reader is advised to reviews[157], [158] and recent literature examples of modern sialic acid donors. [164]

1.6. Unnatural protein expression and incorporation of non-canonical amino acids

1.6.1. The need of homogeneous post-translational modifications

As stated previously, post-translational modifications are ubiquitous and strongly vary in mammalian cells at given times.[6] To better understand and investigate the role and function of those modifications, in a cellular environment, the chemical biologist of today is in strong need of methods to accurately address this challenge.[8] To obtain homogeneously modified proteins to study interactions mediated by PTMs, different approaches from chemistry[165] and molecular biology need to applied that enable the selective and site-specific functionalization of amino acid side chains within the fully expressed protein. Over the years different strategies have emerged to address this matter. Modifications can either be introduced directly by targeting canonical amino acids[166] or by "chemical mutagenesis"[167] prior to a subsequent tagging reaction. Since those strategies tend to be not fully selective and therefore site-specific, the most obvious choice to counter this would be the direct incorporation of a modified amino acid or a mimetic on the translational level. With the incorporation of a non-canonical amino acid (ncAA), bearing a functional group in its side-chain it can subsequently be uniquely addressed by specific chemistry. To really ensure site-specific and selective

modification of proteins, incorporate non-canonical amino acids is advantageous. The common methods for the incorporation of non-canonical amino acids – suppression-based and the selective pressure incorporation (SPI) method – will be described in detail below, with a stronger focus on SPI.

1.6.2. Suppression-based incorporation

It is often coined as site-specific incorporation method, since it requires a single mutation of an amino acid codon to a *nonsense* or a *missense* sequence of nucleic acids on the DNA or mRNA level. A *nonsense* mutation is usually implying a change to one of the three *stop* codons: *amber*, *ochre* and *opal*. They stand for either UAG (*amber*), UAA (*ochre*), and UGA (*opal*) on the mRNA level, or TAG (*amber*), TAA (*ochre*), and TGA (*opal*) on the DNA level. As implied by its name, a *stop* codon will end the translation process, which means that the protein sequence will be terminated by binding of a "release factor". However, it was discovered out that specific *nonsense* codons actually encode for non-canonical amino acids. To be more specific, the UGA codon was found to code additionally for selenocystein, pyrrolysin as well as cystein and tryptophan. *Missense* means the mutation of nucleic acids in a codon to encode for another. For example, UCA would encode for serine, but when its mutated to UUA, leucine will be incorporated instead. If the mutation converts UCA to UAA it would lead to the termination of the protein synthesis because UUA encodes for *stop*. [168] – [172]

As can be seen, *nonsense* and *missense* mutations may lead to inactive or fragmented protein chains which could in the end harm the longevity of an organism. To counteract this, numerous triplets code for the same amino acids. In addition to this, nature evolved suppressor tRNAs, that would allow a "read-through" of the mutation, inserting another amino acid in its place and restoring activity and function of an otherwise truncated oligopeptide. Especially the existence of *nonsense* suppressor tRNA could raise the question of "side-effects" like super-long proteins that could again be harmful for the cell. However, as of now, no clear evidence for this was found. In general it can be concluded from other findings, like different efficiency of the suppressor depending on the site in the sequence, that there is an additional "context effect", supporting the proper usage of this phenomenon, that is rather exceptional in nature. [168] – [172]

The underlying potential of this concept was pioneered by various groups, starting from *in vitro* systems and chemically misacylated amber suppressor tRNA. In those cases suitable suppressor tRNA was designed and misacylated with a non-canonical amino acid of interested. This was paired with an *in vitro* protein expression system, to translate the previously point-mutated gene to the modified protein of interest. Although it was possible to charge the suppressor tRNA with various ncAAs, the protein recovery was often very low. However, this was not because of the generally lower expression yields

of *in vitro* systems but rather the inefficient "read-through" of the suppressed codon, due the previously mentioned "context effect" and the very nature of the incorporated non-canonical amino acid.[172]

To further improve this methodology, an *in vivo* system was proposed. For the achievement of this idea some requirements had to be fulfilled. First, a codon must be programmed to solely code for the non-canonical amino acid. Secondly there has to be a specific tRNA that is capable of deciphering this code and translate it to the right non-canonical amino acid. As third requirement, an aminoacyl tRNA synthetase (aaRS) has to be available to charge this tRNA with the non-canonical amino acid. In addition to all those previous prerequisites, they all should be able to freely interact but not cross-react with the existing physiology and translational machinery of the host. Basically this means that there should be an evolved orthogonal tRNA/aaRS pair available that is not native to the expressing cell, but still able to be used and translated on the protein level (**Figure 1-13**, p. 40).

The major breakthrough in this research was achieved with yeast suppressor tRNA in *E. coli*. Later a suppressor tRNA from *M. jannaschii* was used, also in *E. coli* as a host. Finally the advantages of the latter system – nearly full orthogonality and a broad substrate spectrum – made it one of the suppression-based systems of choice nowadays (**Figure 1-13**, p. 40). [168] – [172]

Using modern site-specific, suppression-based incorporation methods proved to be a powerful tool for chemical biologists, biochemists, protein engineers and biophysicists. It helped installing reactive side chains and probes that revealed more details on protein structures, functions and interactions. Surely, this technology still suffers from limitations, like low expression yields and competition with release factors, but it is still a valuable technology for protein science. [172]

1.6.3. Selective pressure incorporation

Aside from the previously mentioned suppression-based methodologies, residue-specific replacements of canonical amino acids are another way to introduce noncanonical ones.[172],[174] To make this kind of incorporation possible, an organism must be able to accept a non-native substitute of one amino acid for the other. In the laboratory, a special class of mutant bacteria or eukaryotic cells have been revoked their ability to provide their own amino acid supply. This so-called auxotroph strains are ideal for the task to incorporate ncAAs, since their selection of amino acids for protein biosynthesis can be controlled by the scientist. Due to the high specificity of synthetases for their native substrate, having an error rate of misacylation of approximately 10^{-4}, the expansion of their amino acid repertoire can only be achieved by lowering the availability of the natural amino acid and supplement it by its noncanonical analog. By
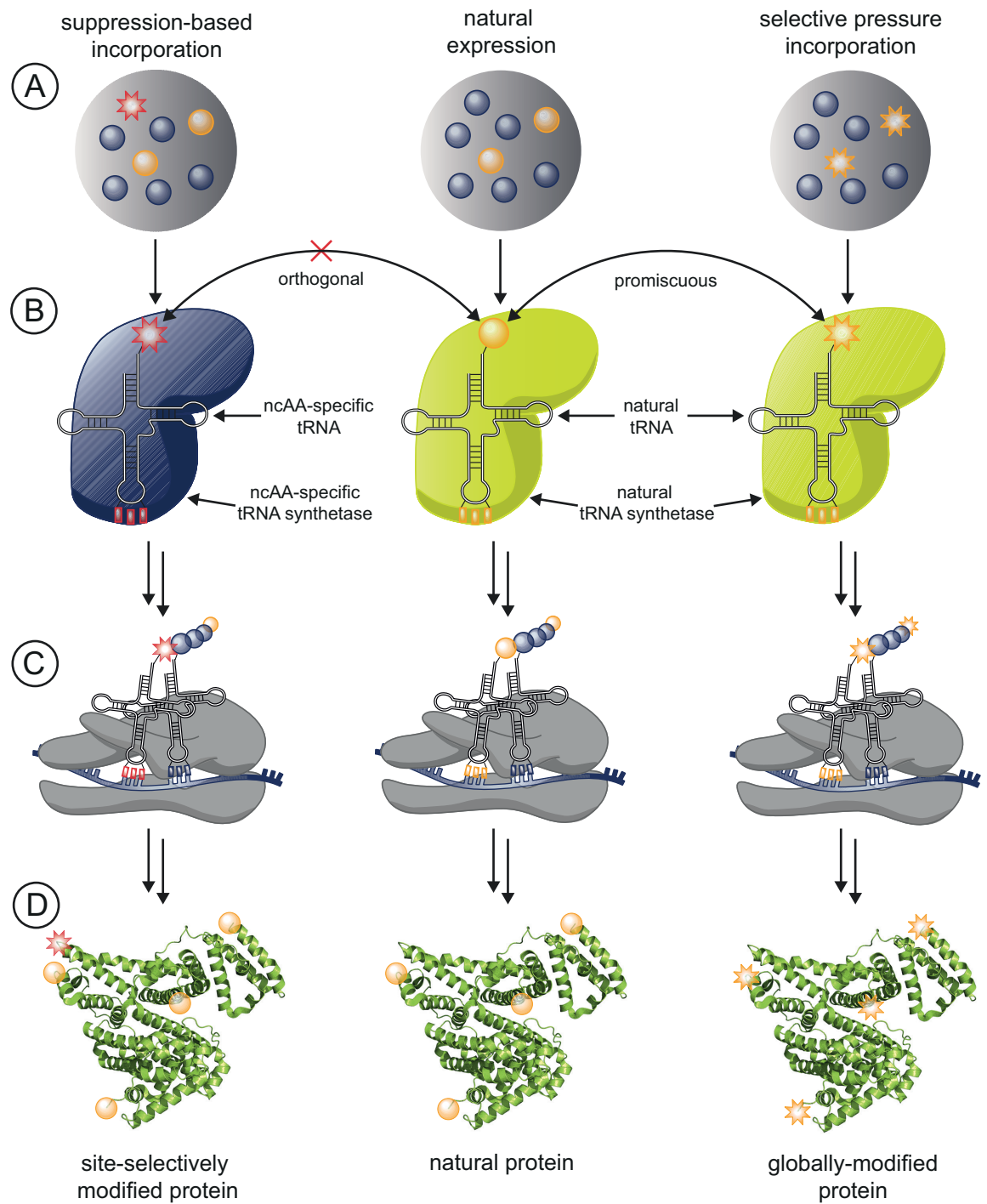
Figure 1-13: Incorporation of canonical and non-canonical amino acids by applying different techniques.

(A) Additional canonical (cAA) and non-canonical amino acids (ncAA) are supplied by the media.
(B) cAA and ncAA are accepted by different tRNA synthetases to charge the corresponding tRNA.
(C) Ribosomal protein synthesis incorporates ncAAs site-specifically or globally, (D) furnishing the respective modified proteins.

Adapted and redrawn from references [172] and [174].

limiting and replacing the amino acid of interest with a noncanonical one, a wide range of different ncAAs could be incorporated by this residue-specific technique. [172]–[175]

It should be noted that the selective pressure for the mutant organism has to be specifically designed for a product of interest. For an successful, proteome wide residue-specific replacement, the organism needs to be able to take up the noncanonical amino acid, attach it to its corresponding tRNA and incorporate it into the growing polypeptide chain instead of its native counter-part. In other words, the sense-codon of an organism has to be successfully reprogrammed to allow a canonical→noncanonical amino acid substitution (**Figure 1-13**, p. 40). [172]–[175]

As can be seen, the strong point in this methodology is the efficient, proteome-wide replacement of one kind of natural amino acid to an unnatural one (**Figure 1-13**, p. 40). Since protein expression is still quite high in this approach and the level of incorporation can easily be tuned by controlling the concentrations of substituted amino acids. Especially the methionine substitution proved quite valuable, since the synthetase is accepting an array of noncanonical amino acids, including alkyne and azide bearing side-chains. This reactive chemical moieties enable bio-orthogonal non-canonical amino acid tagging, or BONCAT as described by the Tirrell laboroatory.[176] This method was used to specifically address proteins, pulse-labeled with azidohomoalanine.

The whole approach relies on the general promiscuity of the used synthetases. Nowadays there are many different noncanonical amino acids available that are generally well accepted. However some substitutes are not as efficiently incorporated as others. To counteract this phenomenon, overexpression of the corresponding aaRS can lead to higher efficiency of incorporation and expression levels. [172]–[175]

In combination with site-specific mutations, the selective pressure incorporation methodology can yield proteins of interest that can be addressed residue-specifically. Although the variety of supplemental amino acids is more limited than with amber suppression, multiple incorporations of chemical and bio-orthogonal reacting side-chains at higher expression rates in one protein of interest can be advantageous for certain experimental setups and biologically relevant questions (*see* **chapter 3**, pp. 60ff., and **chapter 4**, pp. 74ff.). This holds true especially in the case of neoglycoconjugate synthesis on the basis of proteins. To efficiently generate such neoglycoproteins, the protein-based starting material should be homogeneously expressed with a certain number of chemoselectively addressable handles or should bear the ability to be transformed efficiently by other reagents. After this first step, high-yielding reactions that can be performed in aqueous media can follow to install the bio-active moiety on that linker (*see* **3**, pp. 60ff., **4**, pp. 74ff.).[177]

1.7. The copper-catalyzed alkyne-azide cycloaddition as a paragon for bioconjugation reactions

Previous sections of this introduction already mentioned the interest in modification reactions on proteins. Herein, the focus will be laid on chemoselective and bioorthogonal strategies. To classify a reaction as bioorthogonal, it has to hold true to the following premises: selective, stable, fast, low toxicity, and accessible. In this context, selectivity mainly means chemoselectivity – the ability of a reactant to solely react with another specific moiety under the given conditions of the reaction. In a biological environment that is full of reactive groups like thiols, alcohols, amines, *et cetera*, this specific group therefore needs to be stable to any other biological or chemical condition that is not required for the reaction of interest. In addition the ideal bioorthogonal reaction needs to have fast kinetics that lead to a stable covalent bond formation to investigate biological phenomena on a timescale suitable for the observed organism. The necessity of a high turn-over and conversion rate lies especially in "snap-shotting" the stages of development of an organism and its metabolic state at the time of the reaction. Another advantage of fast and high-yielding reactions is the possibility to lower the amount of reactants and therefore ensure proper labeling and low toxicity in biological experiments. The challenge herein lies in the capability of the reaction to still be efficient also under physiologically relevant ranges of pH and temperature. As if the demands already made are not enough, optimal bioorthogonal reactions should require at least one of the reactants to have a relatively small reactive moiety. The size of the functional group is relevant for its availability and accessibility in regards of its introduction into the protein or proteome of the observed organism.[173]–[176], [178]–[181],

All of the considerations mentioned above together make the requirements for a bioorthogonal reaction quite challenging. One might argue that an ideal approach has yet to be found. Nevertheless, tremendous advancements towards those requirements have been made since Carolyn Bertozzi coined this term in 2003.[182]

The concept mentioned above also holds true for site-selective protein modification. Dating back to the 1960's where Koshland was able to convert an active serine to a cystein on the protein level,[183] site-selective alteration and functionalization of proteins have been from great interest and utility. Applications of covalently modified proteins range from tracking when an analytical probe is attached, to improved and novel activities, and possible therapeutics. In regards to chemical biology and glycobiology the site-selective modification with bioactive moieties, like carbohydrates or glycomimetics, enables the synthesis of defined glycoprotein analogues and neoglycoproteins. Such biomacromolecules can be valuable tools to shed light on carbohydrate-protein and

(neo)glycoprotein-lectin interactions, respectively (*see* **3**, pp. 60ff., **4**, pp. 74ff.).

Nowadays a lot of different reactions are known that can be themed as bioorthogonal reactions and are readily used for the aforementioned applications. For the purpose of this thesis, only some chemoselective bioconjugation reactions that were also used in the generation of glycoconjugates will be described exemplary. That usually includes the use of azides and/or alkynes as reactive moiety, such as in the ligand-assisted copper-catalyzed alkyne-azide cycloaddition (CuAAC) where a special focus will be set on.

Both alkyne and azido moieties can be incorporated into proteins and the surface of cells with readiliy available techniques like stop-codon supression (**1.6.2**, pp. 38ff.), selective pressure incorporation (**1.6.3**, pp. 39ff.), metabolic incorporation, and chemically through tag and ligation chemistries. For the synthesis of homogeneously functionalized proteins the most convenient way is by incorporation of either alkyne- or azide-bearing amino acids into the protein. Azidohomoalanine (Aha) and homopropargylglycine (Hpg) can be introduced into the polypeptide chain as discussed before (**1.6**, pp. 37ff.). Considering their range of chemoselective reactions they can undergo, both azide[184], [185] and alkynes are very suitable moieties for bioorthogonal modification strategies. They also react with each other under copper catalysis in a [3+2] cycloaddition. This copper-catalyzed alkyne-azide cycloadditon (CuAAC) was found independently by the groups of Meldal[186] and Sharpless.[187] Soon after they observed the tremendous increase in reaction speed of this 1,3-dipolar cycloaddition between alkynes and azides – initially described by Huisgen – when Cu(I) is applied, the first proposals on the mechanism of the CuAAC were claimed. Initially believed to have only one mononuclear Cu-acetylide intermediate[188], experimental findings and calculations suggested otherwise.[189]–[191] This kinetic studies pointed more towards the involvement of two copper species, interacting with both alkyne and azide moieties. A recent publication by Fokin *et al.*[192] reveals more evidence for the dinuclear copper hypothesis, contributing a major breakthrough into finally unraveling the mechanism of the copper-catalyzed alkyne-azide cycloaddition (**Scheme 1-3**, p. 44). Following the reaction by heat-flow calorimetry, they observed that initially formed mononuclear copper acetylide is not reactive until additional copper is added. Further evidence was provided by cross-over experiments with isotopically enriched copper that were analyzed by electrospray ionization time-of-flight mass spectrometry (ESI TOF-MS). Thereby Fokin *et al.*[192] observed that the two copper centers are equally participating in the reaction and that the formation of nitrogen-carbon bonds happened step-wise.

CuAAC is an efficient reaction that can be carried out in water using easily accessible reagents. It is no surprise that it is often used synonymously to "click reaction" – a definition coined by B.G. Sharpless. [193] Since its discovery it found numerous improvements and applications, including bioconjugation. The latter was demonstrated
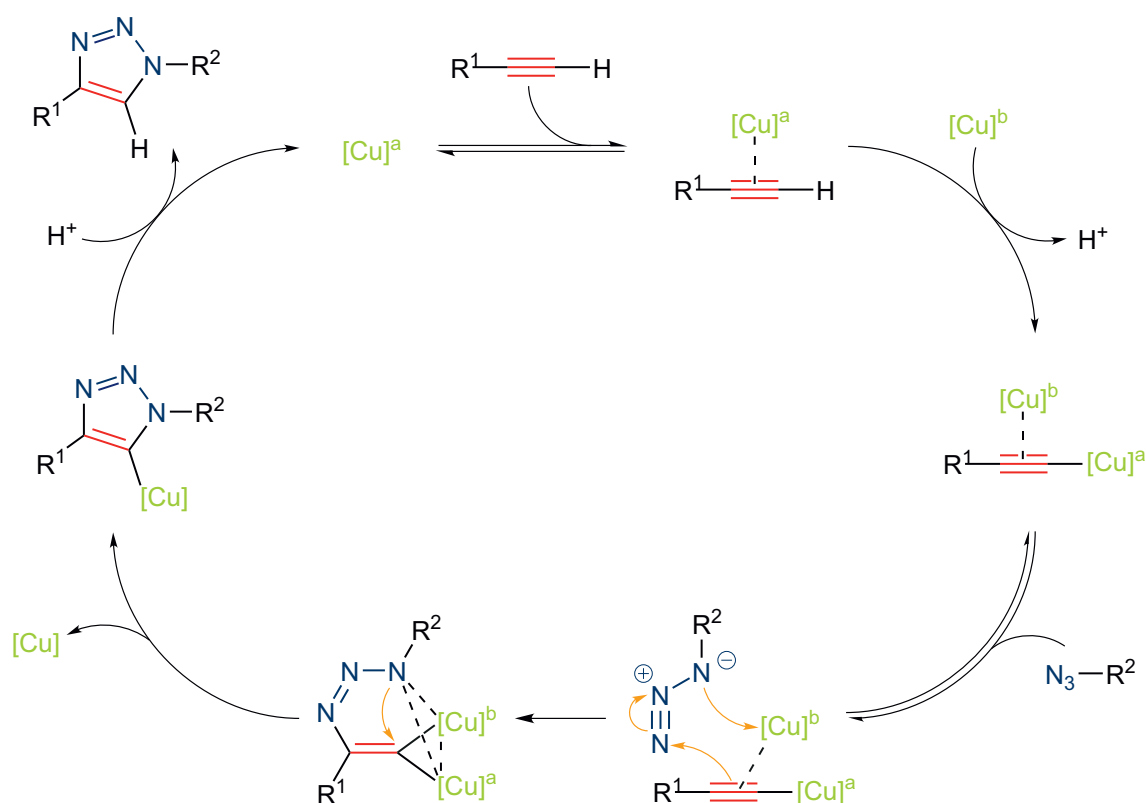
Scheme 1-3: Refined proposed mechanism of the copper-catalyzed alkyne-azide cycloaddition involving two copper atoms.
Adapted and redrawn from reference [192].

only little after its discovery in by a combined effort of the groups of Sharpless, Fokin and Finn.[194] Further applications in the field of chemical biology and the synthesis of neoglycoproteins were natural to follow, since CuAAC is a chemoselective reaction with a regioselective outcome and high reaction kinetics if a suitable ligand is used. One severe limitation for *in vivo* application of CuAAC is the cytotoxicity of the copper(I) catalyst.[195] Another side effect originates from the redox system that is commonly used to activate copper. The combination of sodium ascorbate and copper generates other reactive species including reactive oxygen species (ROS) that can interact with either biological or polymeric scaffold material.[196]–[199]

Generally speaking, the mentioned limitations can be overcome by the use of a suitable improved catalytic system. Soon after its discovery, Fokin *et al.*[200] described the use of a copper(I)-chelating agent that accelerated the CuAAC. Further development of such CuAAC ligands addressed the challenges of efficient ligation through

- stabilization of the reactive Cu(I),
- acceleration of the alkyne-azide cycloaddition,
- suppression of side-product formation, and
- binding of copper to minimize damage on biomaterial and ease subsequent removal.

TBTA

THPTA

BTTES

BTTAA
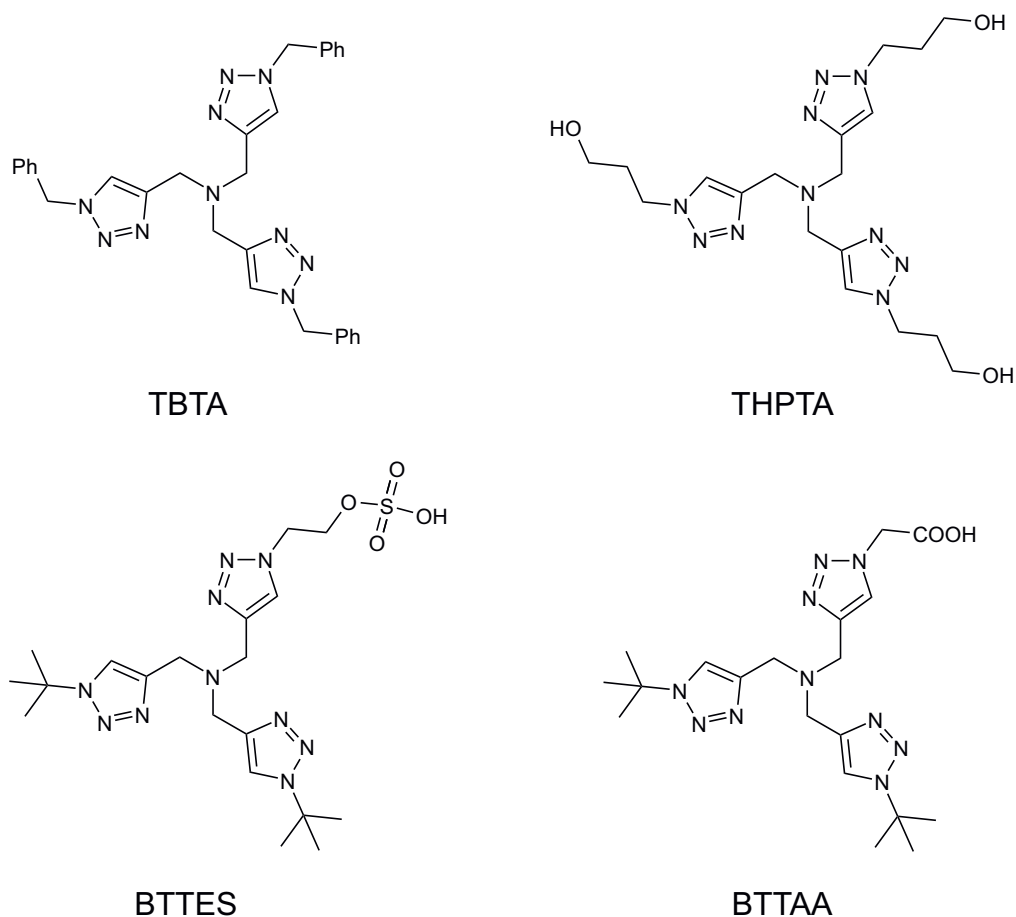
Scheme 1-4: Common tris-triazolyl based copper(I)-stabilizing ligands.
Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA) and the water-soluble tris(3-hydroxypropyltriazolyl methyl)amine (THPTA) are first and second generation ligands, respectively. Third generation ligands 2-[4-{(bis[(1-tert-butyl-1H-1,2,3-triazol-4-yl)methyl]amino)methyl}-1H-1,2,3-triazol-1-yl]ethyl hydrogen sulfate (BTTES) and 2-[4-{(bis[(1-tert-butyl-1H-1,2,3-triazol-4-yl) methyl]amino)methyl}-1H-1,2,3triazol-1-yl]acetic acid (BTTAA) showed improved reaction kinetics.

In the last decade different ligands were promoted in the literature. Based on the initial tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA)[200], the water-soluble tris(3-hydroxypropyltriazolyl methyl)amine (THPTA)[201] was developed. Latter was equally efficient in promoting functionalization of biomolecules compared to bathophenanthroline disulfonate (BPS),[202], [203] but easier to handle under air. The applicability of THPTA for the functionalization of biomolecules[204], [205] and even live cells[206] made it the most valuable choice of CuAAC-ligand.

The next generation of Cu(I)-stabilizing ligands was introduced by Wu and co-workers. First they used the asymmetric 2-[4-{(bis[(1-tert-butyl-1*H*-1,2,3-triazol-4-yl)methyl] amino)methyl}-1*H*-1,2,3-triazol-1-yl]ethyl hydrogen sulfate (BTTES) and showed efficient labeling with zebrafish embryos.[207] Although the potential of asymmetric ligands was investigated before[208], the results shown by Wu *et al.* were astonishing.

Further screening of ligand with improved beneficial abilities finally lead to 2-[4-{(bis[(1-tert-butyl-1H-1,2,3-triazol-4-yl)methyl]amino)methyl}-1H-1,2,3triazol-1-yl]acetic acid (BTTAA).[209] This ligand exceeds all previously described ligands in terms of CuAAC acceleration, bioconjugation and labeling, making it the probably best ligand available of today.[210]

As stated before, the accelerating and protecting abilities of CuAAC additives like THPTA and BTTAA renders them ideal for the conjugation of sensitive bioactive ligands to protein-based scaffolds and labeling experiments. This has been demonstrated by numerous reports on this topic, ranging from DNA, over proteins to cell-surface labeling. [211]–[215]

1.8. References

1. D. Voet, J. G. Voet, *Biochemistry*, John Wiley & Sons, Hoboken, NJ, **2011.**

2. B. Alberts, *Molecular biology of the cell*, Garland Science, New York, **2008.**

3. D. L. Van Vranken, G. A. Weiss, *Introduction to bioorganic chemistry and chemical biology*, Garland Science, New York, **2013.**

4. F. H. C. Crick, *Symp. Soc. Exp. Biol.* **1957**, *12*, 138–163.

5. M. Nirenberg, *Trends Biochem. Sci.* **2004**, *29*, 46–54.

6. C. T. Walsh, S. Garneau-Tsodikova, G. J. Gatto, *Angew. Chem. Int. Ed.* **2005**, *44*, 7342–7372.

7. D. G. Knorre, N. V. Kudryashova, T. S. Godovikova, *Acta Naturae* **2009**, *1*, 29–51.

8. W. R. Liu, Y.-S. Wang, W. Wan, *Mol. BioSyst.* **2011**, *7*, 38–47.

9. T. K. Lindhorst, *Essential of carbohydrate chemistry and biochemistry*, Wiley-VCH, Weinheim, **2007.**

10. A. Varki, *Essentials of glycobiology*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y, **2009.**

11. R. S. Haltiwanger, J. B. Lowe, *Annu. Rev. Biochem.* **2004**, *73*, 491–537.

12. K. Ohtsubo, J. D. Marth, *Cell* **2006**, *126*, 855–867.

13. M. M. Fuster, J. D. Esko, *Nat. Rev. Cancer* **2005**, *5*, 526–542.

14. K. Landsteiner, *Zentralbl. Bakteriol.* **1900**, *27*, 357–362.

15. H. Lis, N. Sharon, *Chem. Rev.* **1998**, *98*, 637–674.

16. Y. C. Lee, R. T. Lee, *Acc. Chem. Res.* **1995**, *28*, 321–327.

17. W. M. Watkins, *Transfus. Med.* **2001**, *11*, 243–265.

18. R. B. Dodd, K. Drickamer, *Glycobiology* **2001**, *11*, 71R–79R.

19. K. Drickamer, *J. Biol. Chem.* **1988**, *263*, 9557–9650.

20. J. B. Sumner, *J. Biol. Chem.* **1919**, *37*, 137–143.

21. D. M. Salunke, M. J. Swamy, M. I. Khan, S. C. Mande, A. Surolia, M. Vijayan, *J. Biol. Chem.* **1985**, *260*, 13576–13579.

22. R. Banerjee, S. C. Mande, V. Ganesh, K. Das, V. Dhanaraj, S. K. Mahanta, K. Suguna, A. Surolia, M. Vijayan, *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 227–231.

23. R. Ravishankar, M. Ravindran, K. Suguna, A. Surolia, M. Vijayan, *Curr. Sci.* **1997**, *72*, 855–861.

24. N. Sharon, H. Lis, *Glycobiology* **2004**, *14*, 53R–62R.

25. R. Barkai-Golan, D. Mirelman, N. Sharon, *Arch. Microbiol.* **1978**, *116*, 119–121.

26. R. Brambl, W. Gade, *Physiol. Plant* **1985**, *64*, 402–408.

27. J. W. Kijne, *Chemtracts Biochem. Mol. Biol.* **1996**, *6*, 180–187.

28. A. M. Hirsch, *Curr. Opin. Plant Biol.* **1999**, *2*, 320–326.

29. G. Kalsi, M. E. Etzler, *Plant Physiol.* **2000**, *124*, 1039–1048.

30. S. Wedepohl, F. Beceren-Braun, S. Riese, K. Buscher, S. Enders, G. Bernhard, K. Kilian, V. Blanchard, J. Dernedde, R. Tauber, *Europ. J. Cell Biol.* **2012**, *91*, 257–264.

31. S. Schmidt, M. Moser, M. Sperandio, *Mol. Immunol.* **2013**, *55*, 49–58.

32. A. G. Morell, C. J. A. Van Den Hamer, I. H. Scheinberg, G. Ashwell, *J. Biol. Chem.* **1966**, *241*, 3745–3749.

33. W. I. Weis, K. Drickamer, *Structure* **1994**, *2*, 1227–1240.

34. H.-J. Hoppe, K. B. M. Reid, *Structure* **1994**, *2*, 1129–1133.

35. D. C. Kilpatrick, *Biochim. Biophys. Acta* **2002**, *1572*, 401–413.

36. D. C. Kilpatrick, *Biochim. Biophys. Acta* **2002**, *1572*, 187–197.

37. Y. C. Lee, R. T. Lee in *Carbohydrates in Chemistry and Biology* (Eds.: B. Ernst, G. W. Hart, P. Sina), WILEY-VCH Verlag GmbH, Weinheim, Germany, **2000.**

38. L. M. Bareford, P. W. Swaan, *Adv. Drug Deliv. Rev.* **2007**, *59*, 748–758.

39. E. I. Park, Y. Mi, C. Unverzagt, H.-J. Gabius, J. U. Baenziger, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17125–17129.

40. L. M. Steirer, E. I. Park, R. R. Townsend, J. U. Baenziger, *J. Biol. Chem.* **2009**, *284*, 3777–3783.

41. J. U. Baenziger, D. Fiete, *Cell* **1980**, *22*, 611-620.

42. A. L. Schwartz, D. Rup, H. F. Lodish, *J. Biol. Chem.* **1980**, *255*, 9033–9036.

43. M. Meier, M.D. Bider, V.N. Malashkevich, M. Spiess, P. Burkhard, *J. Mol. Biol.* **2000**, *300*, 857–365.

44. A. N. Zelensky, J. E. Gready, *FEBS J.* **2005**, *272*, 6179–6217.

45. J. U. Baenziger, Y. Maynard, *J. Biol. Chem.* **1980**, *255*, 4607–4613.

46. A. M. Bianucci, F. A. Chiellini, *J. Biomol. Struct. Dyn.* **2000**, *18*, 435-451.

47. Bernardes, Gonçalo J. L., R. Kikkeri, M. Maglinao, P. Laurino, M. Collot, S. Y. Hong, B. Lepenies, P. H. Seeberger, *Org. Biomol. Chem.* **2010**, *8*, 4987-4996.

48. O. Khorev, D. Stokmaier, O. Schwardt, B. Cutting, B. Ernst, *Bioorg. Med. Chem.* **2008**, *16*, 5216-5231.

49. S. H. Medina, V. Tekumalla, M. V. Chevliakov, D. S. Shewach, W. D. Ensminger, M. E. El-Sayed, *Biomaterials* **2011**, *32*, 4118-4129.

50. U. Westerlind, J. Westman, E. Törnquist, Smith, C. I. Edvard, S. Oscarson, M. Lahmann, T. Norberg, *Glycoconjugate J.* **2004**, *21*, 227-241.

51. I. A. Wilson, J. J. Skehel, D. C. Wiley, *Nature* **1981**, *289*, 366-373.

52. M. B. Eisen, S. Sabesan, J. J. Skehel, D. C. Wiley, *Virology* **1997**, *232*, 19–31.

53. T. L. Bousse, G. Taylor, S. Krishnamurthy, A. Portner, S. K. Samal, T. Takimoto, *J. Virol.* **2004**, *78*, 13351–13355.

54. M. Porotto, M. Fornabaio, G. E. Kellogg, A. Moscona, *J. Virol.* **2007**, *81*, 3216–3228.

55. V. Zaitsev, M. von Itzstein, D. Groves, M. Kiefel, T. Takimoto, A. Portner, G. Taylor, *J. Virol.* **2004**, *78*, 3733–3741.

56. http://www.velcro.de/index.php?id=6 – *Velcro Website*

57. C. Fasting, C. A. Schalley, M. Weber, O. Seitz, S. Hecht, B. Koksch, J. Dernedde, C. Graf, E.-W. Knapp, R. Haag, *Angew. Chem. Int. Ed.* **2012**, *51*, 10472–10498.

58. R. T. Lee, Y. C. Lee, *Glycoconjugate J.* **1987**, *4*, 317-328.

59. M. Mammen, S.-K. Choi, G. M. Whitesides, *Angew. Chem. Int. Ed.* **1998**, *37*, 2754–2794.

60. J. J. Lundquist, E. J. Toone, *Chem. Rev.* **2002**, *102*, 555–578.

61. L. L. Kiessling, T. Young, T. D. Gruber, K. H. Mortell in *Glycoscience* (Eds.: B. O. Fraser-Reid, K. Tatsuta, J. Thiem), Springer, **2008**.

62. D.-I. Liao, G. Kapadia, H. Ahmed, G. R. Vasta, O. Herzberg, *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1428–1432.

63. E. Salomonsson, A. Larumbe, J. Tejler, E. Tullberg, H. Rydberg, A. Sundin, A. Khabut, T. Frejd, Y. D. Lobsanov, J. M. Rini, U. F. Nilsson, H. Leffler, *Biochemistry* **2010**, *49*, 9518–9532.

64. K. Stierand, M. Rarey, *Med. Chem. Lett.* **2010**, *1*, 540–545.

65. K. Stierand, M. Rarey, *ChemMedChem* **2007**, *2*, 853–860.

66. P. Fricker, M. Gastreich, M. Rarey, *J. Chem. Inf. Comp. Sci.* **2004**, *44*, 1065–1078.

67. K. Stierand, P. Maaß, M. Rarey, *Bioinformatics* **2006**, *22*, 1710–1716.

68. F. A. Qiocho, D. K. Wilson, N. K. Vyas, *Nature* **1989**, *340*, 404–407.

69. W. I. Weis, K. Drickamer, *Annu. Rev. Biochem.* **1996**, *65*, 441–473.

70. M. J. Borrok, L. L. Kiessling, K. T. Forest, *Protein Sci.* **2007**, *16*, 1032–1041.

71. W. I. Weis, R. Kahn, R. Fourme, K. Drickamer, W. A. Hendrickson., *Science* **1991**, *254*, 1608–1615.

72. W. I. Weis, K. Drickamer, W. A. Hendrickson, *Nature* **1992**, *360*, 127–134.

73. M. Vijayan, N. Chandra, *Curr. Op. Struct. Biol.* **1999**, *9*, 707–714.

74. C. P. Swaminathan, N. Surolia, A. Surolia, *J. Am. Chem. Soc.* **1998**, *120*, 5153–5159.

75. A. Nurisso, B. Blanchard, A. Audfray, L. Rydner, S. Oscarson, A. Varrot, A. Imberty, *J. Biol. Chem.* **2010**, *285*, 20316–20327.

76. F. P. C. Binder, K. Lemme, R. C. Preston, B. Ernst, *Angew. Chem. Int. Ed.* **2012**, *51*, 7327–7331.

77. R. U. Lemieux, *Acc. Chem. Res.* **1996**, *29*, 373–380.

78. K. D. Puri, A. Surolia, *Pure Appl. Chem.* **1994**, *68*, 497–502.

79. W. P. Jencks, *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 4046–4050.

80. M. Mammen, E. I. Shakhnovich, G. M. Whitesides, *J. Org. Chem.* **1998**, *63*, 3168–3175.

81. J. Rao, J. Lahiri, L. Isaacs, R. M. Weis, G. M. Whitesides *Science* **1998**, *280*, 708–711.

82. J. E. Gestwicki, C. W. Cairo, L. E. Strong, K. A. Oetjen, L. L. Kiessling, *J. Am. Chem. Soc.* **2002**, *124*, 14922–14933.

83. L. L. Kiessling, L. E. Strong, J. E. Gestwicki in *Annual reports in medicinal chemistry* (Eds.: A. M. Doherty, W. F. Greenlee), Academic Press, San Diego, **2000.**

84. C. W. Cairo, J. E. Gestwicki, M. Kanai, L. L. Kiessling, *J. Am. Chem. Soc.* **2002**, *124*, 1615–1619.

85. G. T. Morgan, H. D. K. Drew, *J. Chem. Soc., Trans.* **1920**, *117*, 1456–1465.

86. G. Schwarzenbach, *Helv. Chim. Acta* **1952**, *35*, 2344–2359.

87. G. Schwarzenbach, *Anal. Chim. Acta* **1952**, *7*, 141–155.

88. M. I. Page, W. P. Jencks, *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 1678–1683.

89. K. S. Pitzer, W. D. Gwinn *J. Chem. Phys.* **1942**, *10*, 428–440.

90. D. G. Truhler, *J. Comp. Chem.* **1991**, *12*, 266–270.

91. T. Christensen, D. M. Gooden, J. E. Kung, E. J. Toone, *J. Am. Chem. Soc.* **2003**, *125*, 7357–7366.

92. V. Vallet, U. Wahlgren, I. Grenthe, *J. Am. Chem. Soc.* **2003**, *125*, 14941–14950.

93. N. L. Pohl, L. L. Kiessling, *Synthesis* **1999**, *SI*, 1515–1519.

94. P. I. Kitov, D. R. Bundle, *J. Am. Chem. Soc.* **2003**, *125*, 16271–16284.

95. V. Böhrsch, T. Mathew, M. Zieringer, Vallée, M. Robert J., L. M. Artner, J. Dernedde, R. Haag, C. P. Hackenberger, *Org. Biomol. Chem.* **2012**, *10*, 6211–6216.

96. J. Dernedde, A. Rausch, M. Weinhart, S. Enders, R. Tauber, K. Licha, M. Schirner, U. Zugel, A. von Bonin, R. Haag, *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 19679–19684.

97. S. Espuelas, C. Thumann, B. Heurtault, F. Schuber, B. Frisch, *Bioconjugate Chem.* **2008**, *19*, 2385–2393.

98. J. A. F. Joosten, V. Loimaranta, C. C. M. Appeldoorn, S. Haataja, F. Ait El Maate, R. M. J. Liskamp, J. Finne, R. J. Pieters, *J. Med. Chem.* **2004**, *47*, 6499–6508.

99. K. Matsuura, M. Hibino, T. Ikeda, Y. Yamada, K. Kobayashi, *Chem. Eur. J.* **2004**, *10*, 352–359.

100. L. Gu, P. G. Luo, H. Wang, M. J. Meziani, Y. Lin, L. M. Veca, L. Cao, F. Lu, X. Wang, R. A. Quinn et al., *Biomacromolecules* **2008**, *9*, 2408–2418.

101. E. Mahon, T. Aastrup, M. Barboiu, *Chem. Commun.* **2010**, *46*, 5491–5493.

102. J.-F. Nierengarten, J. Iehl, V. Oerthel, M. Holler, B. M. Illescas, A. Muñoz, N. Martín, J. Rojo, M. Sánchez-Navarro, S. Cecioni et al., *Chem. Commun.* **2010**, *46*, 3860–3862.

103. A. Ciulli in *Methods in Molecular Biology* (Eds.: M. A. Williams, T. Daviter), Humana Press, Totowa, NJ, **2013**.

104. V. M. Krishnamurthy, L. A. Estroff, G. M. Whitesides in Methods and *Principles in Medicinal Chemistry* (Eds.: W. Jahnke, D. A. Erlanson), Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, FRG, **2006**.

105. T. Wiseman, S. Williston, J. F. Brandts, L.-N. Lin, *Anal. Biochem.* **1989**, *179*, 131–137.

106. W. B. Turnbull, A. H. Daranas, *J. Am. Chem. Soc.* **2003**, *125*, 14859–14866.

107. T. K. Dam, C. F. Brewer, *Chem. Rev.* **2002**, *102*, 387-430.

108. D. K. Mandal, N. Kishore, C. F. Brewer, *Biochemistry* **1994**, *33*, 1149-1156.

109. N. K. Sauter, M. D. Bednarski, B. A. Wurzburg, J. E. Hanson, G. M. Whitesides, J. J. Skehel, D. C. Wiley, *Biochemistry* **1989**, *28*, 8388–8396.

110. J. E. Hanson, N. K. Sauter, J. J. Skehel, D. C. Wiley, *Virology* **1992**, *189*, 525–533.

111. N. K. Sauter, J. E. Hanson, G. D. Glick, J. H. Brown, R. L. Crowther, S. J. Park, J. J. Skehel, D. C. Wiley, *Biochemistry* **1992**, *31*, 9609–9621.

112. S. R. Shenoy, L. G. Barrientos, D. M. Ratner, B. R. O'Keefe, P. H. Seeberger, A. M. Gronenborn, M. R. Boyd, C*hem. Biol.* **2002**, *9*, 1109–1118.

113. T. Onizuka, H. Shimizu, Y. Moriwaki, T. Nakano, S. Kanai, I. Shimada, H. Takahashi, *FEBS J.* **2012**, *279*, 2645–2656.

114. J. Homola, S. S. Yee, G. Gauglitz, *Sens. Actuators B Chem.* **1999**, *54*, 3–15.

115. R. P. Kooyman in *Handbook of Surface Plasmon Resonance* (Eds.: R. B. M. Schasfoort, A. J. Tudos), Royal Society of Chemistry, Cambridge, **2008**.

116. P. Pattnaik, *Appl. Biochem. Biotechnol.* **2005**, *126*, 79–92.

117. K. H. Schlick, M. J. Cloninger, *Tetrahedron* **2010**, *66*, 5305–5310.

118. I. Papp, J. Dernedde, S. Enders, R. Haag, *Chem. Commun.* **2008**, 5851–5853.

119. I. Papp, J. Dernedde, S. Enders, S. B. Riese, T. C. Shiao, R. Roy, R. Haag, *Chembiochem* **2011**, *12*, 1075–1083.

120. I. Papp, C. Sieben, A. L. Sisson, J. Kostka, C. Böttcher, K. Ludwig, A. Herrmann, R. Haag, *Chembiochem* **2011**, *12*, 887–895.

121. T. S. Hawley, R. G. Hawley, *Flow cytometry protocols*, Humana Press, Totowa, N.J, **2004**.

122. A. L. Givan, *Flow Cytometry: First Principles*, John Wiley & Sons, Inc., New York, USA, **2001**.

123. S. Hong, P. R. Leroueil, I. J. Majoros, B. G. Orr, J. R. Baker, M. M. Banaszak Holl, *Chem. Biol.* **2007**, *14*, 107–115.

124. T. van den Beucken, H. Pieters, M. Steukers, M. van der Vaart, R. C. Ladner, H. R. Hoogenboom, S. E. Hufton, *FEBS Letters* **2003**, *546*, 288–294.

125. K. Ho, Y. Lapitsky, M. Shi, M. S. Shoichet, *Soft Matter* **2009**, *5*, 1074–1080.

126. F. H. Kayser, *Medical Microbiology*, Georg Thieme Verlag, Stuttgart, New York, NY, **2005**.

127. R. V. Stick, S. J. Williams, *Carbohydrates. The essential molecules of life*, Elsevier, London, **2009**.

128. A. V. Demchenko, *Handbook of Chemical Glycosylation*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, **2008**.

129. L. Bohé, D. Crich, *C. R. Chim.* **2011**, *14*, 3–16.

130. H. Kunz, *Angew. Chem. Int. Ed.* **2002**, *41*, 4439–4451.

131. S. Chandrasekhar, *ARKIVOC* **2005**, 37–66.

132. S. C. Ranade, A. V. Demchenko, *J. Carbohydr. Chem.* **2013**, *32*, 1–43.

133. R. U. Lemieux, *Pure Appl. Chem.* **1971**, *25*, 527–548.

134. R. R. Schmidt, M. Behrendt, A. Toepfer, *Synlett* **1990**, 694–696.

135. A. Demchenko, T. Stauch, G.-J. Boons, *Synlett* **1997**, 818–820L. Bohé, D. Crich, *C. R. Chim.* **2011**, *14*, 3–16.

136. H. Satoh, H. S. Hansen, S. Manabe, van Gunsteren, Wilfred F., P. H. Hünenberger, *J. Chem. Theory Comput.* **2010**, *6*, 1783–1797.

137. D. Crich, M. Patel, *Carbohydr. Res.* **2006**, *341*, 1467–1475.

138. J. Guo, X.-S. Ye, *Molecules* **2010**, *15*, 7235–7265.

139. A. Michael, *Am. Chem. J.* **1879**, *1*, 305–312.

140. W. Koenigs, E. Knorr, *Ber. Dtsch. Chem. Ges.* **1901**, *34*, 957–981.

141. B. Helferich, K.-F. Wedemeyer, *Justus Liebigs Ann. Chem.* **1949**, *563*, 139–145.

142. B. Helferich, J. Zirner, *Chem. Ber.* **1962**, *95*, 2604–2611.

143. H. Paulsen, O. Lockhoff, *Chem. Ber.* **1981**, *114*, 3102–3114.

144. R. R. Schmidt, J. Michel, *Angew. Chem. Int. Ed. Engl.* **1980**, *19*, 731–732.

145. J. D. C. Codée, R. E. J. N. Litjens, L. J. van den Bos, H. S. Overkleeft, G. A. van der Marel, *Chem. Soc. Rev.* **2005**, *34*, 769–782.

146. B. G. Davis, *J. Chem. Soc., Perkin Trans. 1* **2000**, 2137–2160.

147. X. Zhu, R. R. Schmidt, *Angew. Chem. Int. Ed.* **2009**, *48*, 1900–1934.

148. B. Fraser-Reid, C. J. López (Eds.) *Topics in Current Chemistry*, Springer Berlin Heidelberg, Berlin, Heidelberg, **2011**.

149. L. K. Mydock, A. V. Demchenko, *Org. Biomol. Chem.* **2010**, *8*, 497–510.

150. A. Varki, *Glycobiology* **1992**, *2*, 25–40.

151. T. Angata, A. Varki, *Chem. Rev.* **2002**, *102*, 439–470.

152. R. Schauer, *Curr. Opin. Struct. Biol.* **2009**, *19*, 507–514.

153. X. Chen, A. Varki, *ACS Chem. Biol.* **2010**, *5*, 163–176.

154. M. J. Kiefel, M. von Itzstein, *Chem. Rev.* **2002**, *102*, 471–490.

155. B. Byrne, G. G. Donohoe, R. O'Kennedy, *Drug Discov. Today* **2007**, *12*, 319–326.

156. E. R. Vimr, *ISRN Microbiol.* **2013**, 1–26.

157. G.-J. Boons, A. V. Demchenko, *Chem. Rev.* **2000**, *100*, 4539–4566.

158. C. de Meo, U. Priyadarshani, *Carbohydr. Res.* **2008**, *343*, 1540–1552.

159. M. D. Farris, C. de Meo, *Tetrahedron Lett.* **2007**, *48*, 1225–1227.

160. D. Crich, W. Li, *J. Org. Chem.* **2007**, *72*, 2387–2391.

161. D. Crich, W. Li, *J. Org. Chem.* **2007**, *72*, 7794–7797.

162. H. Tanaka, Y. Nishiura, T. Takahashi, *J. Am. Chem. Soc.* **2008**, *130*, 17244–17245.

163. H. Tanaka, Y. Tateno, Y. Nishiura, T. Takahashi, *Org. Lett.* **2008**, *10*, 5597–5600.

164. C.-H. Hsu, K.-C. Chu, Y.-S. Lin, J.-L. Han, Y.-S. Peng, C.-T. Ren, C.-Y. Wu, C.-H. Wong, *Chem. Eur. J.* **2010**, *16*, 1754–1760.

165. C. P. R. Hackenberger, D. Schwarzer, *Angew. Chem. Int. Ed.* **2008**, *47*, 10030–10074.

166. J. M. Chalker, G J. L. Bernardes, Y. A. Lin, B. G. Davis, *Chem. Asian J.* **2009**, *4*, 630–640.

167. J. M. Chalker, B. G. Davis, *Curr. Opin. Chem. Biol.* **2010**, *14*, 781–789.

168. D. Belin, *Genetics* **2003**, 165, 455–456.

169. L. Wang, P. G. Schultz, *Angew. Chem. Int. Ed.* **2005**, *44*, 34–66.

170. T. S. Young, P. G. Schultz, *J. Biol. Chem.* **2010**, *285*, 11039–11044.

171. G. Bertram, S. Innes, O. Minella, J. P. Richardson, I. Stansfield, *Microbiology* **2001**, *147*, 255–269.

172. N. Budisa, *Angew. Chem. Int. Ed.* **2004**, *43*, 6426–6463.

173. A. J. Link, M. L. Mock, D. A. Tirrell, *Curr. Opin. Biotech.* **2003**, *14*, 603–609.

174. J. A. Johnson, Y. Y. Lu, J. A. Van Deventer, D. A. Tirrell, *Curr. Opin. Chem. Biol.* **2010**, *14*, 774–780.

175. J. T. Ngo, D. A. Tirrell, *Acc. Chem. Res.* **2011**, *44*, 677–685.

176. D. C. Dieterich, A. J. Link, J. Graumann, D. A. Tirrell, E. M. Schuman, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 9482–9487.

177. J. M. Chalker, G. J. L. Bernardes, B. G. Davis, *Acc. Chem. Res.* **2011**, *44*, 730–741.

178. J. A. Prescher, C. R. Bertozzi, *Nat. Chem. Biol.* **2005**, *1*, 13–21.

179. N. J. Agard, C. R. Bertozzi, *Acc. Chem. Res.* **2009**, *42*, 788–797.

180. E. M. Sletten, C. R. Bertozzi, *Acc. Chem. Res.* **2011**, *44*, 666–676.

181. K. E. Beatty, *Mol. BioSyst.* **2011**, *7*, 2360–2367.

182. H. C. Hang, C. Yu, D. L. Kato, C. R. Bertozzi, *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 14846–14851.

183. K. E. Neet, D. E. Koshland, *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *56*, 1606–1611.

184. S. Bräse, C. Gil, K. Knepper, V. Zimmermann, *Angew. Chem. Int. Ed.* **2005**, *44*, 5188–5240.

185. M. F. Debets, C. W. J. van der Doelen, F. P. J. T. Rutjes, F. L. van Delft, *ChemBioChem* **2010**, *11*, 1168–1184.

186. C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057–3064.

187. V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem.* **2002**, *114*, 2708–2711.

188. F. Himo, T. Lovell, R. Hilgraf, V. V. Rostovtsev, L. Noodleman, K. B. Sharpless, V. V. Fokin, *J. Am. Chem. Soc.* **2005**, *127*, 210–216.

189. V. O. Rodionov, V. V. Fokin, M. G. Finn, *Angew. Chem. Int. Ed.* **2005**, *44*, 2210–2215.

190. M. Ahlquist, V. V. Fokin, *Organometallics* **2007**, *26*, 4389–4391.

191. D. Cantillo, M. Ávalos, R. Babiano, P. Cintas, J. L. Jiménez, J. C. Palacios, *Org. Biomol. Chem.* **2011**, *9*, 2952–2958.

192. B. T. Worrell, J. A. Malik, V. V. Fokin, *Science* **2013**, *340*, 457–460.

193. H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2021.

194. Q. Wang, T. R. Chan, R. Hilgraf, V. V. Fokin, K. B. Sharpless, M. G. Finn, *J. Am. Chem. Soc.* **2003**, *125*, 3192–3193.

195. D. C. Kennedy, C. S. McKay, M. C. B. Legault, D. C. Danielson, J. A. Blake, A. F. Pegoraro, A. Stolow, Z. Mester, J. P. Pezacki, *J. Am. Chem. Soc.* **2011**, *133*, 17993–18001.

196. P.-Y. Liu, N. Jiang, J. Zhang, X. Wei, H.-H. Lin, X.-Q. Yu, *Chem. Biodivers.* **2006**, *3*, 958–966.

197. S C. Fry, *Biochem. J.* **1998**, *332*, 507–515.

198. A. Kumar, K. Li, C. Cai, *Chem. Commun.* **2011**, *47*, 3186–3188.

199. E. Lallana, R. Riguera, E. Fernandez-Megia, *Angew. Chem. Int. Ed.* **2011**, *50*, 8794–8804.

200. T. R. Chan, R. Hilgraf, K. B. Sharpless, V. V. Fokin, *Org. Lett.* **2004,** *6*, 2853–2855.

201. V. Hong, S. I. Presolski, C. Ma, M. G. Finn, *Angew. Chem. Int. Ed.* **2009**, *48*, 9879–9883.

202. S. Sen Gupta, J. Kuzelka, P. Singh, W. G. Lewis, M. Manchester, M. G. Finn, *Bioconjug. Chem.* **2005**, *16,* 1572–1579.

203. V. Hong, A. K. Udit, R. A. Evans, M. G. Finn, *ChemBioChem* **2008**, *9,* 1481–1486.

204. L. M. Artner, L. Merkel, N. Bohlke, F. Beceren-Braun, C. Weise, J. Dernedde, N. Budisa, Hackenberger, C. P. R., *Chem. Commun.* **2011**, *48*, 522–524.

205. D. Banerjee, A. P. Liu, N. R. Voss, S. L. Schmid, M. G. Finn, *ChemBioChem* **2010**, *11*, 1273–1279.

206. V. Hong, N. F. Steinmetz, M. Manchester, M. G. Finn, *Bioconjugate Chem.* **2010**, *21*, 1912–1916.

207. D. Soriano del Amo, W. Wang, H. Jiang, C. Besanceney, A. C. Yan, M. Levy, Y. Liu, F. L. Marlow, P. Wu, *J. Am. Chem. Soc.* **2010**, *132*, 16893–16899.

208. S. I. Presolski, V. Hong, S.-H. Cho, M. G. Finn, *J. Am. Chem. Soc.* **2010**, *132*, 14570–14576.

209. C. Besanceney-Webler, H. Jiang, T. Zheng, L. Feng, D. Soriano del Amo, W. Wang, L. M. Klivansky, F. L. Marlow, Y. Liu, P. Wu, *Angew. Chem. Int. Ed.* **2011**, *50*, 8051–8056.

210. H. Jiang, T. Zheng, A. Lopez-Aguilar, L. Feng, F. Kopp, F. L. Marlow, P. Wu, *Bioconjug. Chem.* **2014**, DOI: 10.1021/bc400502d.

211. P. V. Chang, X. Chen, C. Smyrniotis, A. Xenakis, T. Hu, C. R. Bertozzi, P. Wu, *Angew. Chem. Int. Ed.* **2009**, *48*, 4030–4033

212. J.-K. Rhee, M. Hovlid, J. D. Fiedler, S. D. Brown, F. Manzenrieder, H. Kitagishi, C. Nycholat, J. C. Paulson, M. G. Finn, *Biomacromolecules* **2011**, *12*, 3977–3981.

213. J.-K. Rhee, M. Baksh, C. Nycholat, J. C. Paulson, H. Kitagishi, M. G. Finn, *Biomacromolecules* **2012**, *13*, 2333–2338.

214. R. Ribeiro-Viana, M. Sánchez-Navarro, J. Luczkowiak, J. R. Koeppe, R. Delgado, J. Rojo, B. G. Davis, *Nat. Comms.* **2012**, *3*, 1303.

215. C. Uttamapinant, A. Tangpeerachaikul, S. Grecian, S. Clarke, U. Singh, P. Slade, K. R. Gee, A. Y. Ting, *Angew. Chem. Int. Ed.* **2012**, *51*, 5852–5856.

2. Aim of the study

2.1. Overview

Over the last decades glycobiology gained more and more interest from academic and industrial research. Especially the phenomenon of multivalent binding events between carbohydrates and proteins seemed from common interest.

To elucidate multivalent binding behavior the primary focus was set on the chemoselective modification of proteins to generate precise scaffolds. Briefly, the applied work-flow was 1) to obtain non-canonic amino acid bearing proteins, 2) modify them by copper-catalyzed alkyne azide cycloaddition, and 3) biologically evaluate the outcome (**Figure 2-14**). In addition to this, polydisperse, multivalent carbohydrate-modified polymers should be evaluated as well.

Each project within the thesis can basically be broken down in the following categories:

- Synthesis of biologically active ligands, basically carbohydrate azides;

- Chemoselective reaction of the synthesized sugars with specific templates to present them in multivalent fashion;

- Evaluation of the created scaffolds with biologically relevant assays.
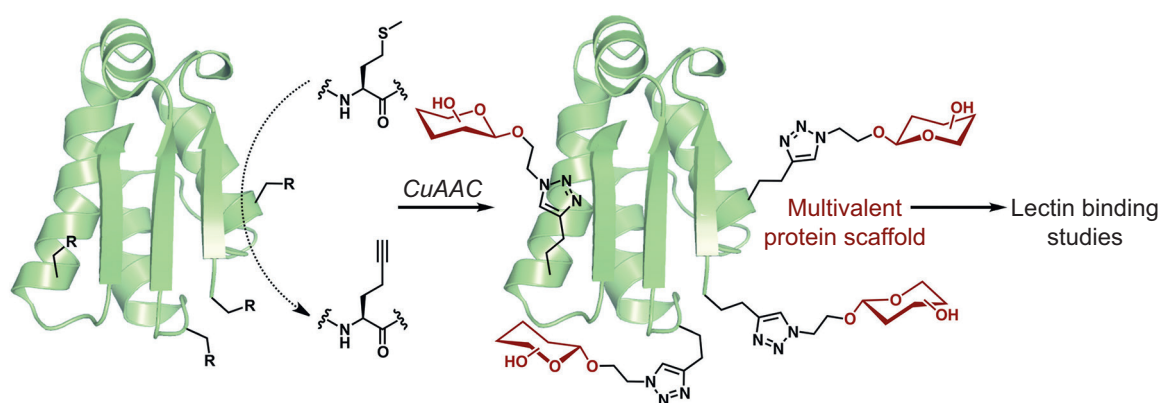
Figure 2-14: Working model of multivalent protein scaffolds.
Proteins, bearing non-canonical amino acids (ncAAs) are obtained by selective pressure incorporation (SPI). Subsequently they are functionalized by CuAAC with carbohydrate azides. Finally, the neoglycoproteins are subjected to biological assays.
Reproduced from Artner et al., *Chem. Commun.* **2012**, *48*, 522–524 with permission of The Royal Society of Chemistry.

For every different target receptor or lectin, a different combination of carbohydrate azides and scaffolds was chosen. The following sub-sections provide more detail on that matter.

2.1.1. Targets

For biological evaluation of applied scaffolds, the following carbohydrate binding proteins should be addressed:

- Peanut agglutinin (PNA),

- Asialoglycoprotein receptor (ASGP-R), and

- Hemagglutinin.

Peanut agglutinin (PNA) is a well-researched, soluble lectin derived from peanut and was chosen as a model study. Follow-up experiments should then be conducted on ASGP-R and hemagglutinin with rationally designed protein probes.

2.1.2. Synthesis of biologically active carbohydrate azides

To address the different targets specifically, various carbohydrates had to be synthesized. For the binding to PNA, β-linked galactose and lactose series were accessed, whereas ASGP-R demanded β-linked *N*-acetylgalactosamine as positive and α-mannose conjugates as negative control. Hemagglutinin's specificity is for α-linked *N*-acetylneuraminic acids, or sialic acids, which tend to be one of the most challenging carbohydrate conjugates from a synthetic point of view.

Each series of sugar azides – galactose, lactose, N-acetylgalactosamine, mannose and N-acetylneuraminic acid – consisted of an azide moiety directly attached to the anomeric center, indicating no linker, an azido ethoxy linker, or an 2-[2-(2-azidoethoxy)ethoxy] ethyl spacer. To achieve the synthesis of the fifteen carbohydrates, different glycosylation and transformation strategies were applied, from simple halogenation followed by nucleophilic substitutions, to methods from carbohydrate chemistry, such as Koenigs-Knorr or Schmidt glycosylations, as well as thioglycoside donors for stereospecific functionalization.

2.1.3. Functionalization of scaffolds

The method of choice for the functionalization of protein-based scaffolds was CuAAC. It was reported to have a wide range of applicability, from small molecules to macromolecules like polymers and proteins, as well as surfaces. Because of this variety, different protocols have already been published. Therefore one of the key elements in successful protein functionalization with CuAAC was the adaptation of previous protocols to an

extent were they are optimized for the specific protein scaffold.

2.1.4. Evaluation of active (bio-)polymers

To properly estimate the success and to draw conclusions from the various scaffolds tested for PNA and ASGP-R, different biological assays and methods should be applied. With PNA, the focus was strongly on surface plasmon resonance. This was justified by the fact that PNA is a soluble lectin, that can perfectly interact with its binding partners in a flow-based binding assay. Additionally one of the known biologically active ligands of PNA, Thomsen-Friedenreich-antigen, can easily be used to functionalize the surface plasmon resonance sensor chips for the assay, since its quite common and commercially available. PNA is an ideal candidate for the evaluation of binding but has no true biological relevance.

In contrast to PNA, ASGP-R is a very good model for endocytosis and also candidate for drug delivery approaches. ASGP-R is also a membrane protein, expressed in hepatic cell lines, and therefore suitable for quantification by fluorescence-activated cell sorting (FACS) to evaluate the efficacy of neoglycoproteins.

3. Barnase-Inhibitor Protein (barstar) – a Model Study†

3.1. Contributions

L. M. Artner, L. Merkel, N. Bohlke, F. Beceren-Braun, C. Weise, J. Dernedde, N. Budisa, and C. P. R. Hackenberger designed the experiments and the research project. L. Merkel and N. Bohlke expressed, purified and characterized ψ-barstar proteins prior to modification. F. Beceren-Braun and L. M. Artner performed surface plasmon resonance of modified proteins and controls. L. Merkel and L. M. Artner did SDS-PAGE and recorded fluorescence and circular dichroism spectra of given proteins. L. Merkel processed and analyzed those spectra. C. Weise and L. M. Artner obtained MALDI-TOF-MS data from given proteins. L. M. Artner synthesized carbohydrate azides and the THPTA ligand, and performed CuAAC on ψ-barstar proteins and their purification.

3.2. Introduction

A lot of different cellular and biological interactions are mediated through post-translational modifications (PTMs) and understanding them seems crucial.[1] For mammalians, one of the most common but also quite complex PTM is glycosylation.[2, [3] As described previously, they are important for various biological phenomena, like cellular interactions and development of diseases (**1.3**, pp. 4ff.), and often occur in a multivalent fashion (**1.4**, pp. 15ff.).[4]–[14] Although proteins are long known as bio-compatible scaffolds for various research topics, the term "rational protein design" seemed to be limited to enzyme engineering. Taking advantage of readily available protein crystal structures and rationally changing the design of the scaffold for other reasons than enzymatic activity was not pursued. For the purpose of elucidating

multivalent modes of interaction this methodology was fitting naturally. Especially when taking into consideration that a lot of known multivalent interactions happen between protein-based receptors and ligands. Glycosylated proteins often are the acting part in those interactions. Therefore artificially generating glycosylated proteins with a defined number of known ligands at previously chosen sites of the protein backbone seemed to be a logical step to shed light on multivalent interactions.

In this first study, a cystein-free mutant of barnase-inhibitor protein, was taken as a model substrate. The "pseudo-wild type" barstar, ψ-barstar, or ψ-b*, as it was termed, had three mutations: Pro27Ala, Cys40Ala, and Cys82Ala. It consists of a mixture of three parallel strands of β-sheets and four α-helices and is relatively small with a molecular weight of around 10 kDa. Since the supplementation incorporation method (SPI) is proteome-wide replacing L-methionine with the supplemented non-canonical amino acid of choice, Met-sites had to be cloned into the protein. For this reason the following mutations were performed: Lys23Met, Glu47Met and Lys79Met. If a careful look is taken on the crystal structure, one could see, that the positions are randomly distributed, but still focusing on facing only one side. [15]–[17]

Since this project was designed as a model study, peanut agglutinin (PNA) was used as the lectin of interest. Its crystal structure, assembly and substrate specificity is already known to the community and therefore ideal for initial prove-of-principle experiments. [18]–[20]

To generate an initial step towards the rational design of multivalent binding proteins, ψ-b* was obtained from the workgroup of Prof. Nediljko Budisa at the Technische Universität Berlin. This protein was expressed in an auxotrophic bacterial strain to incorporate homopropargylglycine *via* selective pressure incorporation (SPI, **1.6.3**, pp. 39ff.). [3], [21]–[24] Copper-catalazyed alkyne-azide cycloaddition (CuAAC)[24]–[27] of alkyne-proteins and carbohydrate azides yielded neoglycoproteins for evaluation by surface plasmon resonance (SPR).

3.3. Results

Terminal β-D-galactose is known to bind well to PNA. Therefore a series of β-linked galactosyl azides (**1**–**3, Scheme 3-5**, p. 62) as well as β-linked lactosyl azides (**4**–**6**, **Scheme 3-6**, p. 63) were prepared for the functionalization of homopropargylglycine bearing ψ-b* variants. To prove a multivalent binding effect, ψ-barstar proteins bearing four (ψ-b*-4M[Hpg]) and one non-canonical amino acid (ψ-b*-1M[Hpg]) were functionalized by an optimized CuAAC. The so acquired neoglycoproteins beared lactose and galactose residues on their surface and were subsequently evaluated by a SPR-based competitive lectin binding assay against PNA.
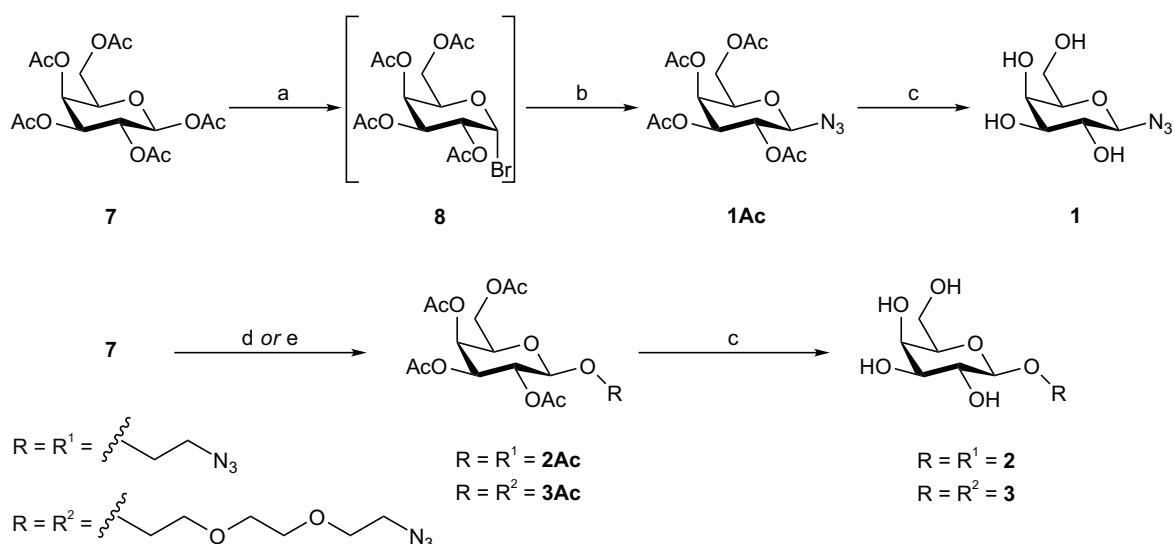
Scheme 3-5: Synthesis of galactose azides.

a) AcOH, HBr/AcOH, CH$_2$Cl$_2$, 0 °C → r.t.; b) NaN$_3$, TBAS, 1 M K$_2$CO$_3$/CH$_2$Cl$_2$, r.t.; c) NaOMe, MeOH, r.t.; d) BF$_3$·OEt$_2$, N$_3$EtOH, CH$_2$Cl$_2$, 0 °C → r.t.; e) TMSOTf, N$_3$(EtO)$_2$EtOH, CH$_2$Cl$_2$, 0 °C.

3.3.1. Synthesis of carbohydrate azides

The target lectin of this study – PNA – is binding terminal galactose in glycans. Therefore carbohydrate azides with a variation of three different linker lengths, directly attached, ethoxy-spaced, and triethylenglycol-spaced, were synthesized from D-galactose and D-lactose. To synthetically obtain those carbohydrate derivatives, both series were started out from the per-acetylated sugars (**Scheme 3-5**, **Scheme 3-6**, p. 63). In the case of D-lactose this could easily be achieved by adding it to a heated mixture of sodium acetate in acetic anhydride. Although the yields of the acetylation were not high and only approximately 50% of the β-acetate was obtained, the procedure was very simple and easily up-scalable to a hundred gram scale.[28] This let to enough starting material for the glycosylations of lactoses derivatives.

To obtain an azide moiety directly attached to the anomeric center in β-postion of galactose, a pseudo-one-pot approach was used.[29] For this the unprotected D-galactose was acetylated under acidic conditions with acetic anhydride and hydrobromic acid. *In-situ* transformation to the bromide **8** was then achieved by increasing the amount of hydrobromic acid in the reaction mixture. Subsequently bromide **8** was substituted to the azide in a two-phase reaction, to yield the desired acetylated galactose azide **1Ac** (**Scheme 3-5**).

Obtaining the other galactose azides should be achieved by a multi-step route. Since acetylation with pyridine and acetic anhydride was mainly leading to the α-anomer, which is less reactive than its β-counterpart, it was decided to go for a Schmidt-donor system. To transform the galactose into this more reactive glycosyl-donor, the anomeric
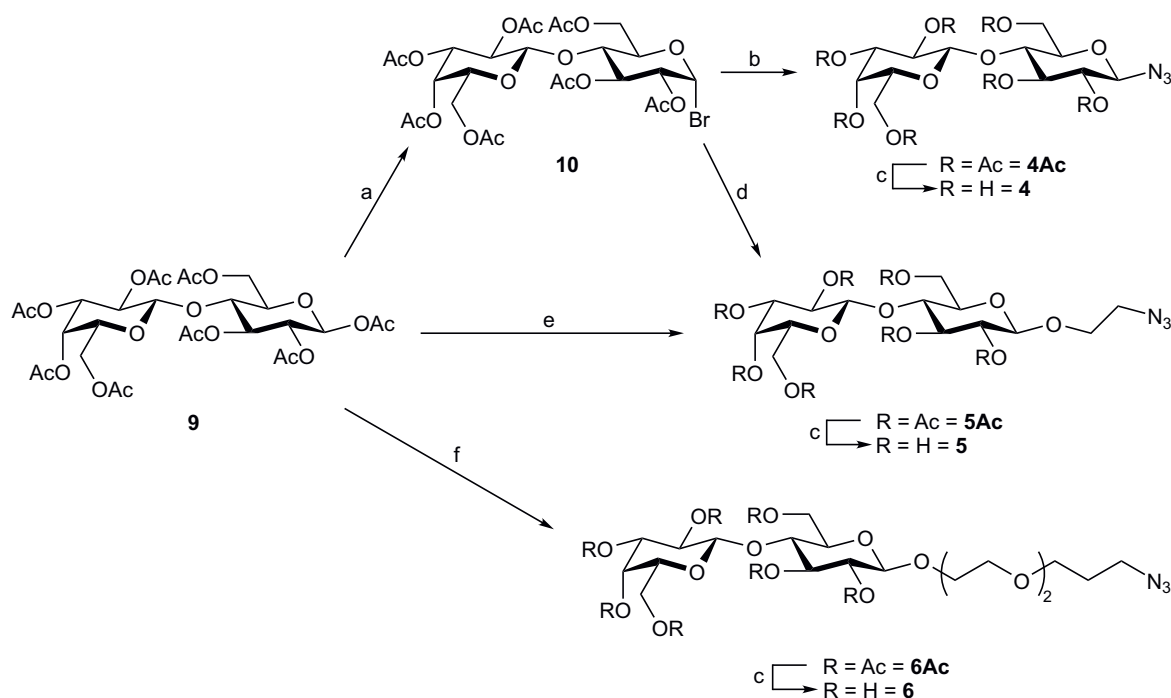
Scheme 3-6: Synthesis of lactose azides.

a) AcOH, HBr/AcOH, CH$_2$Cl$_2$, 0 °C → r.t.; b) NaN$_3$, TBAS, 1 м K$_2$CO$_3$/CH$_2$Cl$_2$, r.t.; c) NaOMe, MeOH, r.t.; d) HgBr$_2$/Hg(CN)$_2$, N$_3$EtOH, MeCN, r.t.; e) BF$_3$·OEt$_2$, N$_3$EtOH, CH$_2$Cl$_2$, 0 °C → r.t.; f) TMSOTf, N$_3$(EtO)$_2$EtOH, CH$_2$Cl$_2$, 0 °C.

oxygen first had to be deprotected. Here it was found out that the deacetylation of the C-1 position is not selective per se, but rather based on the relative reactivity of the other positions. It was found out that longer reaction times can even lead to deprotection of other positions. Nevertheless galactose acetate with a free hydroxy group on the anomeric center for further reactions could be obtained after chromatography on silica gel. The final installment of the Schmidt-donor group, a tricholoracetimidate, was achieved by reacting the previously obtained galactose with trichloroacetonitrile in the presence of base. Both potassium carbonate and the non-nucleophilic triethylamine performed well in this case. The Schmidt-donor is highly reactive and can be activated with a catalytic amount of trimethylsilyl triflate at low temperatures to give the desired β-anomer in good yields. (*For details about glycosyl thrichloroacetimidates, see* **1.5.1.3**, pp. 34ff.)

Since realization of this route was rather time consuming, a more straight forward way was introduced by reacting commercially available 1,2,3,4,6-penta-*O*-acetyl-β-D-galacto-pyranoside (**7**) under Lewis acid conditions with 2-azidoethanol or 2-[2-(2-azido-ethoxy)ethoxy]ethanol. Following this, both acetylated galactose azides **2Ac** and **3Ac** were obtained in moderate yields.[30], [31]

With the β-anomer of the lactose-octa-*O*-acetate **9** in hand,[28] two principle methodologies proofed to be straight-forward to obtain the desired lactose derivatives,

63

4, **5**, **6** (**Scheme 3-6**, p. 63). For the first two azides, it seemed convenient to go *via* the α-lactosyl bromide **10**.[32] The reaction with hydrobromic acid in glacial acetic acid was quantitative and yielded the carbohydrate in high purity without the necessity of column chromatography. Synthesizing acetylated lactose **4Ac** from the bromide was achieved by a nucleophilic substitution in DMSO with sodium azide at room temperature. Subsequent deprotection afforded the lactose azide **4** in acceptable yields. At this point it should also be noted, that the synthesis could also be achieved in a one-pot-like fashion. Therefore the β-lactose-octa-*O*-acetate **9** was transformed to the bromide and – without purification – directly reacted to the azide in a two-phase reaction (**Scheme 3-6**, p. 63).[29] Acetylated lactosyl azide **5Ac** was synthesized by glycosylation of bromide **10** with 2-azidoethanol under Koenigs-Knorr-Helferich conditions.[32], [33] This methodology proved to give good yields and stereo-selectivity. Especially purification was easier due to the stronger difference in retention times of the bromide and the product. However, also direct glycosidation of acetate **9** with 2-azidoethanol under BF$_3$·OEt$_2$ promotion proved reliable for the synthesis of **5Ac**.[34] The last carbohydrate azide from the lactose series, **6Ac**, could be realized by using Lewis acid catalyzed glycosylation with TMSOTf (**Scheme 3-6**, p. 63). [34]

To finally yield the unprotected carbohydrate azides[30]–[38] from both the lactose and the galactose series, the precursors were deacetylated under Zemplén conditions[39] in good yields.

3.3.2. Modification of ψ-barstar proteins by copper-catalyzed alkyne-azide cycloaddition

In collaboration with the work group of Prof. Dr. Nediljko Budisa, the mutants of ψ-b* should be used as a model scaffold for a multivalent presentation of carbohydrates for lectin binding studies. Since modification of the proteins should be achieved by the well-known copper-catalyzed alkyne azide cycloaddition (CuAAC),[40]–[42] the alkyne moiety needed to be incorporated into the protein. For this the variants ψ-b*-1M and ψ-b*-4M were expressed in a L-methionine auxotrophic strain of *Escherichia coli*, and Met was substituted for the unnatural amino acid homopropargylglycine (Hpg). The so-obtained proteins were called ψ-b*-1M[Hpg] and ψ-b*-4M[Hpg]. Although modification of single alkyne bearing proteins were reported to be straightforward and high-yielding, it was found out, that complete functionalization of multiple groups was not. Because of this, copper(I)-stabilizing ligands were thought of. As proposed by the work groups of M.G. Finn and others, such ligands protect the copper(I) state from oxidation and stabilize it in aqueous media. This proves especially valuable for prolonged reaction times and when one is limited by the availability of the starting materials, which is common when using proteins. The ligands are readily accessible by organic synthesis.
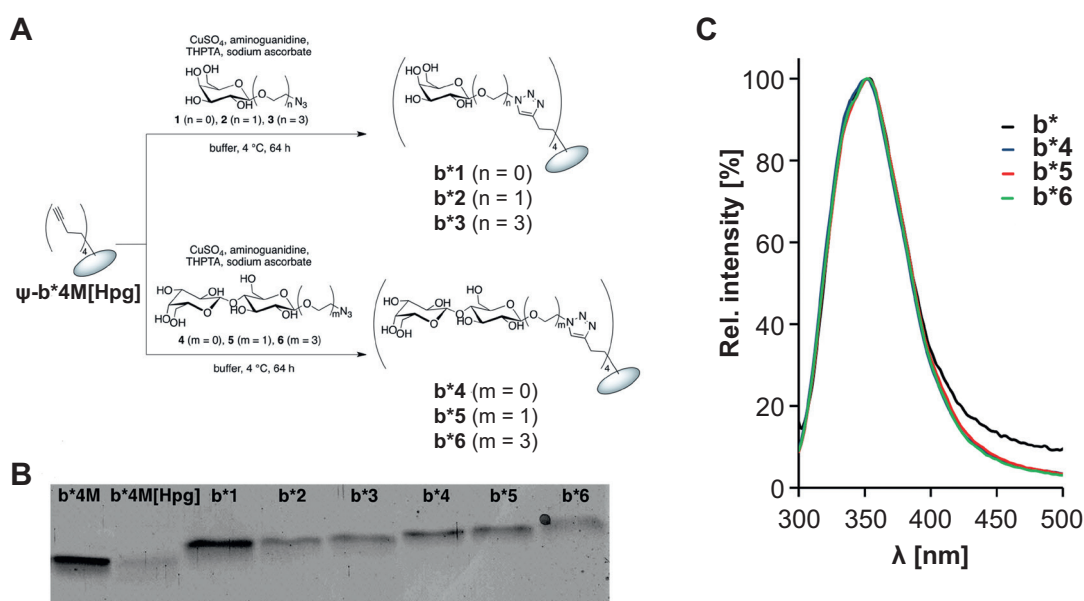
Figure 3-15: Functionalization of ψ-b* proteins by CuAAC.
A) Auxotrophically expressed ψ-b*-4M[Hpg] were functionalized by CuAAC with either galactose azides (**1**, **2**, **3**) or lactoses azides (**4**, **5**, **6**). B) A visible shift in SDS-PAGE and increase in molecular mass pointed towards good functionalization. C) To prove structural integrity of modified ψ-b* neoglycoproteins, fluorescence was measured and compared to the native ψ-b*.

For initial experiments the first generation ligand TBTA was used. Unfortunately it is badly soluble in water and must therefore be provided in DMSO.[25] This issue was limiting its usage for the modification of proteins, since precipitation was encountered on a regular basis. A major advantage was the introduction of the water-soluble second generation ligand, THPTA. Extensive studies on this matter led to a nearly perfect protocol, which was used as a starting point for optimizing CuAAC for ψ-b*-1M[Hpg] and ψ-b*-4M[Hpg].[43]–[45], [46] A crucial drawback from modifying ψ-b*-4M[Hpg] under the published conditions was the instability of the protein. This means, that over time, the protein degraded. Finally this could be prevented by lowering the reaction temperature to 4 °C. The therefrom arising problem of slower proceeding of the CuAAC was immediately countered by prolonging the reaction time to approximately 60 hours (**Figure 3-15, A**).[46]

When the conjugation to the carbohydrates was accomplished, the neoglycoproteins were purified by gel filtration using a FPLC, followed by concentration with centrifugal filter tubes. SDS-PAGE of the proteins was performed to visualize the increase in mass due to the attached carbohydrates. In addition to this, MALDI-MS spectra were taken to confirm the reaction outcome. It can be easily seen from the spectra, that the most abundant peak is the fully functionalized protein, accompanied from matrix adduct peaks (see Supporting Information of reference [46]). Both the increased m/z values,
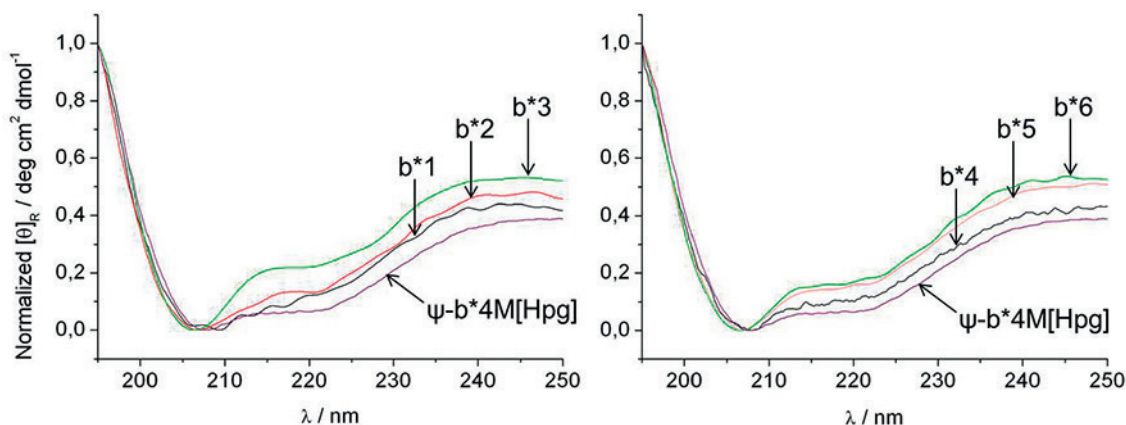
Figure 3-16: Modified ψ-b*-4M[Hpg] proteins were analyzed by circular dichroism.
The melting curves were normalized for better visualization and comparison of structural features.

as well as the visible shift in the coomassie-stained SDS-PAGE gel pointed towards full functionalization as the major product of the CuAAC between carbohydrate azides and ψ-b*-4M[Hpg] (**Figure 3-15**, p. 65**, B,** and Supporting Information of reference [46]). [46]

Additionally the structural integrity of the protein after the conjugation had to be investigated. For this, a close look on Trp53 was taken. Since tryptophan is known as a fluorescent amino acid which emission spectra is changing upon changes of its environment, it was used as very sensitive, internal read-out for the shape of the protein.[17], [47] The obtained fluorescent spectra of the modified barstar-proteins all matched the unreacted starting material in shape. This was an indication for the structural integrity of the neoglycoprotein (**Figure 3-15**, p. 65**, C**). In addition to this, the modified proteins were analyzed by circular dichroism (**Figure 3-16**). Because of sensitivity issues, the melting curves had to be normalized. However, the shape of the curves were similar, which was an additional evidence for structural integrity.

3.3.3. Competitive lectin binding assay with peanut agglutinin

(Performed with Dr. Figen Beceren-Braun, Charité Berlin)

As stated above, a structurally well-defined neoglycoprotein was obtained. To test the applicability of the approach, an assay for surface plasmon resonance (SPR) was envisioned. Peanut agglutinin (PNA) was chosen as the lectin of interest. In the last decades, PNA was extensively researched. Therefore its crystal structure and its carbohydrate preference is well documented in the literature. Briefly, native PNA is non-glycosylated, legume-type lectin, assembled as a homotetramer. The binding pockets of each subunit point into different directions. Due to its structure, a multivalent binding behavior can only come from clustering with its ligands.
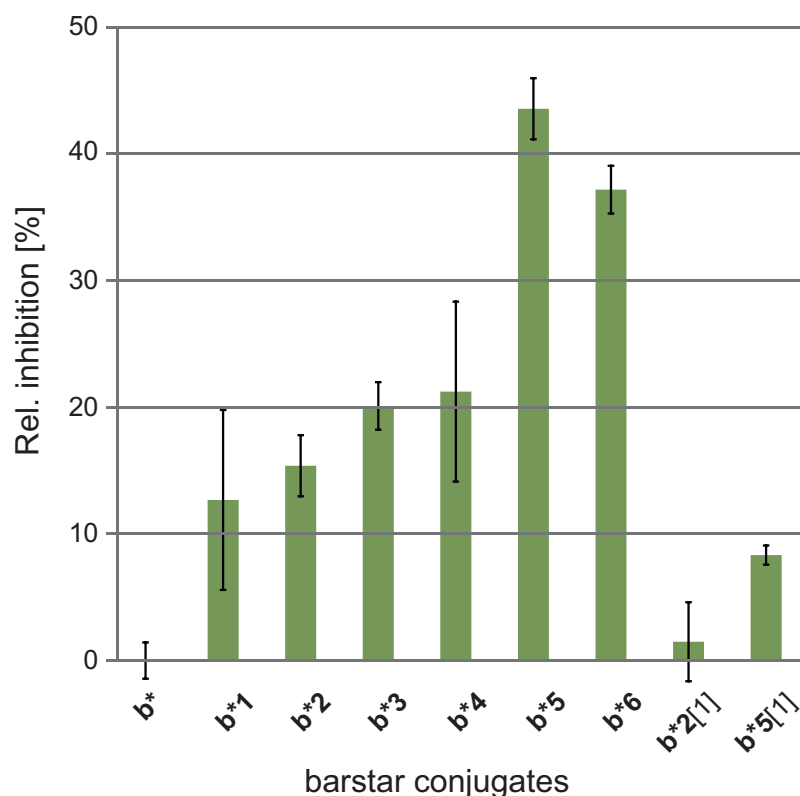
Figure 3-17: **Results of a competitive SPR-based PNA binding assay with barstar proteins.**
The competitive binding assay was carried out with newly functionalized ψ-b* neoglycoproteins.
Whereas less to none inhibition of PNA was observed with native ψ-b* (**b***) and monovalent **b*2**[1],
up to approximately 40% could be inhibited at a final neoglycoprotein concentration of 10 μM with
b*5.
Reproduced from reference [46] with permission of The Royal Society of Chemistry.

For the SPR assay, the natural binder of PNA, Thomsen-Friedenreich antigen (TFα),
a Gal-GalNAc disaccharide (Galβ1-3GalNAcα-O-Ser), should be used as a competitive
binder on the surface of the chip. Therefore Thomsen-Friedenreich (TF) antigen linked
to a biotinylated polyacrylamide (PAA) carrier (Galβ1-3GalNAcα-$(CH_2)_3$-PAA-$(CH_2)_6$-
biotin) was immobilized on the chip surface through biotin-streptavidin binding.
Subsequently PNA was incubated with different neoglycoproteins, unfunctionalized
protein and buffer. Those mixtures were then injected to flow over the functionalized
SPR chip. The observed reduction in the response signal was compared with
data obtained from runs with unmodified barstar and without (only buffer). All
neoglycoproteins showed an inhibitory effect due to stronger binding to PNA than to
the TFα-functionalized surface (**Figure 3-17**).

For evaluation of this observation, experiments had to be conducted to elucidate if the
binding is due to clustering or because of increased ligand concentrations. Therefore
monovalent and unbound ligand at similar concentrations was analyzed (**Figure
3-18**, p. 68). In addition monovalent b*1M was also tested in the aforementioned
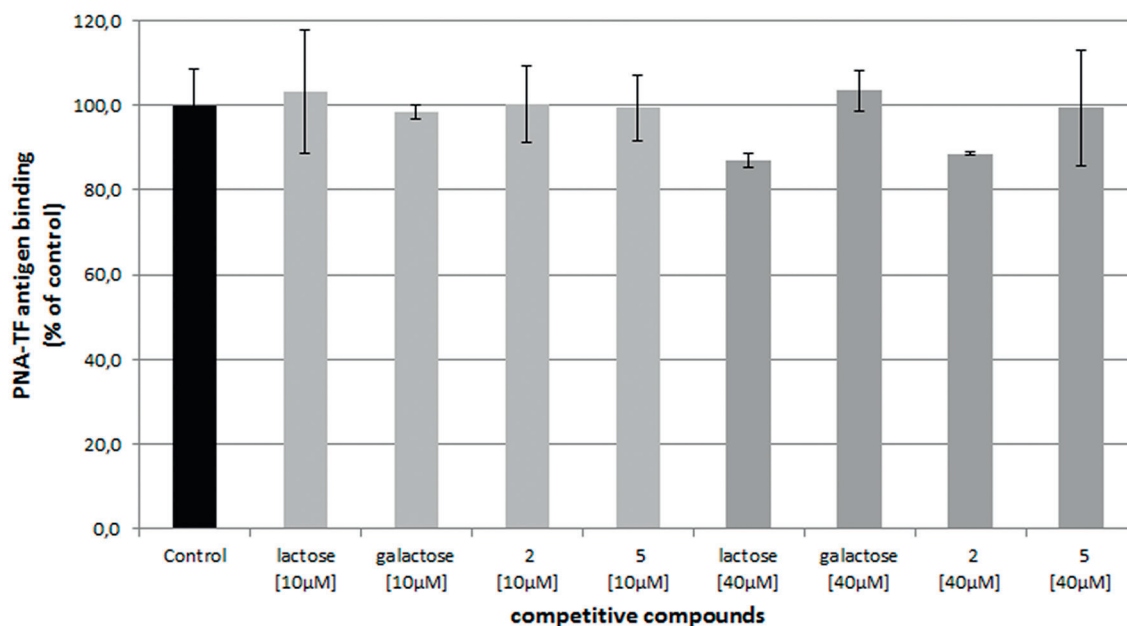
Figure 3-18:   Results of a competitive SPR-based PNA binding assay with simple carbohydrates and carbohydrate azides.

Reproduced from reference [46] with permission of The Royal Society of Chemistry.

competitive binding assay and measured by SPR. Unbound galactose and lactose, as well as there azide-containing counterparts **2** and **5** were showing a lower inhibitory effect at the same nominal and effective concentrations in comparison to the multivalent scaffolds **b*1** to **b*6**. The monovalent versions of the best-binding conjugates, **b*2[1]** and **b*5[1]** were – as expected – weaker binders.

3.4. Discussion

The synthesis of carbohydrate azides – namely lactoses and galactoses – yielded satisfactory yields and purity. Because both saccharides are quite common for applications and modifications on the aglycon part, previously established synthetic routes could easily be applied. As soon as the carbohydrate synthesis was complete, thorough investigations on the optimization of copper-catalyzed azide-alkyne cyloaddition on proteins was conducted. Although there are a lot of examples within the literature, protocols for complete conversions of all participating alkyne within the proteins are hardly found. Initially the conditions reported from the Sharpless lab were applied, but with barstar being a very sensitive protein, soon the limitations of the classic CuAAC were discovered. High amounts of copper sulfate, sodium ascorbate and prolonged reaction times lead to the degradation of His-tagged barstar protein.

MALDI-ToF mass spectrometry revealed that mainly unspecific cleavage of the poly-His-tag was the reason for this. However, other work-groups also encountered limitations of the standard Sharpless conditions and therefore added tris-triazole-based ligands to the reaction. Solubility issues with TBTA, a first-generation Cu-stabilizing ligand, were soon encountered, but thorough literature research lead to the application of THPTA, a water-soluble second generation tris-triazole-based ligand. This ligand worked fairly well, and, in combination with prolonged reaction times at lower temperature, degradation could be prevented and very good conversation achieved. Although purification was not optimized, dialysis followed by spin-filtration for concentration, still yielded sufficient amounts of the proteins to conduct further experiments. It should be noted, that dialysis and concentration was performed subsequently with EDTA-containing buffer and Millipore water to remove copper from the protein solution. From the analysis of the whole protein by MALDI-ToF mass spectrometry and fluorescence measurements, it was concluded that the protein is still intact and in its native structure after CuAAC and purification. This assumption was especially based on the emission spectra obtained from lactose-functionalized barstar. Proteins in their native, tertiary structure yield a specific fluorescence spectra. [[47] Changes in the environment of contributing amino acids would easily manipulate those emission spectra. Therefore, one might say that fluorescence emission spectra of proteins present a fingerprint of their tertiary structure. With barstar there is no difference concerning this concept. It contains three tryptophans, one buried (Trp53) and two on solvent-exposed sites (Trp38, Trp44).[17] The Trp-residues of barstar are a major contributor to the structural intergrity of the protein,[48]in addition to being relevant to the fluorescence of the protein. Within the context of structural integrity, Trp53 is from uttermost importance. Its side chain is in a network with Phe56 and Phe74, and changes within this environment would change the emission profile of the protein. Essentially, the emission spectra of the lactose-modified proteins were the same as those from unreacted barstar. Since neither a hypsochromic, nor a bathochromic shift was observed, it can reasonably be agreed that the native structure of the protein is not significantly impaired.

Assuming that the modified protein still resembles the its crystal structure,[17] the multivalent binding of barstar to PNA can be explained by clustering of multiple receptors and ligands. As with experiments on ConA with multivalent small molecular weight ligands,[49] it is reasonable that the modified barstar variants cross-link and aggregate with PNA in a similar way. The reduced binding of singly modified barstar in comparison to the quadruply functionalized protein is also explained since monovalent barstar cannot interact with more than one PNA. A simple increase in concentration could not make up for the effectiveness of scaffold-bound, multivalently presented carbohydrates (**Figure 3-17**, p. 67, and **Figure 3-18**, p. 68).

3.5. Summary and conclusions

In the light of recent advances in copper-catalyzed alkyne-azide cycloaddition (CuAAC) for bioconjugations, an artificial, multivalently binding protein scaffold was envisioned. By combining selective pressure induced incorporation methodology with bioorthogonal conjugation chemistry, multivalent neoglycoproteins were engineered. The application of well-known synthetic routes towards galactosyl and lactosyl azides lead to the desired carbohydrates in acceptable yields and good purity. Well-defined glycoconjugates were obtained after the optimization of the CuAAC between the alkyne-bearing protein and the azide-containing carbohydrates. Although the pre-selected sites for the installation of unnatural amino acids that bear bioorthogonal, chemically reactive side-chains, were randomly chosen, the advantage of this strategy towards highly modular multivalent protein scaffolds was made clear. The synthesized tetravalent barstar variants proved superior in potency of inhibition compared to monovalent proteins, the bare ligands, and simple carbohydrates. In addition, the results of fluorescent measurements suggested structural integrity of the proteins after CuAAC and purification *via* dialysis. It was therefore possible to rationalize the increased affinity with cluster-induced aggregation of the soluble PNA and the tetravalent neoglycoconjugates. Based on this observations, the herein presented model system provides a good starting point for further advances. In combination with the availability of three-dimensional structural information of proteins that can be used as scaffolds, but also for the assessment of new lectin-based targets, rationally designed multivalent probes made from proteins can be envisioned. This glycoconjugates are multivalent, but – in comparison to polymers – also well-characterized, mono-disperse and flexible in design.

The conclusions drawn from this model study can be applied to a variety of different protein-based systems and carbohydrate-binding lectins (*vide infra,* chapter **4**, pp. 74ff.).

3.6. References

1. C. T. Walsh, S. Garneau-Tsodikova, G. J. Gatto, Jr., *Angew. Chem. Int. Ed.* **2005**, *44*, 7342–7372.

2. T. K. Lindhorst, *Essential of carbohydrate chemistry and biochemistry*, Wiley-VCH, Weinheim, **2007**.

3. D. P. Gamblin, E. M. Scanlan, B. G. Davis, *Chem. Rev.* **2009**, *109*, 131–163.

4. D. A. Tirrell, *Nature* **2004**, *430*, 837.

5.   H. Lis, N. Sharon, *Chem. Rev.* **1998**, *98*, 637–674.

6.   Y. C. Lee, R. T. Lee, *Acc. Chem. Res.* **1995**, *28*, 321–327.

7.   N. Sharon, *J. Biol. Chem.* **2007**, *282*, 2753–2764.

8.   C. R. Bertozzi, *Chem. Biol.* **1995**, *2*, 703–708.

9.   L. L. Kiessling, J. E. Gestwicki, L. E. Strong, *Angew. Chem. Int. Ed.* **2006**, *45*, 2348–2368.

10.   M. Mammen, S. K. Choi, G. M. Whitesides, *Angew. Chem. Int. Ed.* **1998**, *37*, 2754–2794.

11.   C. R. Bertozzi, L. L. Kiessling, *Science* **2001**, *291*, 2357–2364.

12.   D. Deniaud, K. Julienne, S. G. Gouin, *Org. Biomol. Chem.* **2011**, *9*, 966–979.

13.   B. E. Collins, J. C. Paulson, *Curr. Opin. Chem. Biol.* **2004**, *8*, 617–625.

14.   D. H. Dube, C. R. Bertozzi, *Nat. Rev. Drug Disc.* **2005**, *4*, 477–488.

15.   L. Merkel, H. S. G. Beckmann, V. Wittmann, N. Budisa, *ChemBioChem* **2008**, *9*, 1220–1224.

16.   S. Lepthien, L. Merkel, N. Budisa, *Angew. Chem. Int. Ed.* **2010**, *49*, 5446–5450.

17.   M. Rubini, S. Lepthien, R. Golbik, N. Budisa, *Biochim. Biophys. Acta* **2006**, *1764*, 1147–1158.

18.   D. M. Salunke, M. J. Swamy, M. I. Khan, S. C. Mande, A. Surolia, M. Vijayan, *J. Biol. Chem.* **1985**, *260*, 13576–13579.

19.   R. Banerjee, S. C. Mande, V. Ganesh, K. Das, V. Dhanaraj, S. K. Mahanta, K. Suguna, A. Surolia, M. Vijayan, *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 227–231.

20.   R. Ravishankar, M. Ravindran, K. Suguna, A. Surolia, M. Vijayan, *Curr. Sci.* **1997**, *72*, 855–861.

21.   N. Budisa, *Angew. Chem. Int. Ed.* **2004**, *43*, 6426–6463.

22.   L. Wang, P. G. Schultz, *Angew. Chem. Int. Ed.* **2005**, *44*, 34–66.

23.   M. G. Hoesl, N. Budisa, *ChemBioChem* **2011**, *12*, 552–555.

24.   S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony, B. G. Davis, *Nature* **2007**, *446*, 1105–1109.

25.   E. Kaya, K. Gutsmiedl, M. Vrabel, M. Müller, P. Thumbs, T. Carell, *ChemBioChem* **2009**, *10*, 2858–2861.

26.   H. Lin, C. T. Walsh, *J. Am. Chem. Soc.* **2004**, *126*, 13998–14003.

27. V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem.* **2002**, *114*, 2708–2711.

28. C.S. Hudson, J. M. Johnson, *J. Am. Chem. Soc.* **1915**, *37*, 1270–1275.

29. R. Kumar, P. Tiwari, P. R. Maulik, A. K. Misra, *Eur. J. Org. Chem.* **2006**, 74–79.

30. F. Fazio, M. C. Bryan, O. Blixt, J. C. Paulson, C.-H. Wong, *J. Am. Chem. Soc.* **2002**, *124*, 14397–14402.

31. C. Bouillon, A. Meyer, S. Vidal, A. Jochum, Y. Chevolot, J. P. Cloarec, J. P. Praly, J. J. Vasseur, F. Morvan, *J. Org. Chem.* **2006**, *71*, 4700–4702.

32. R. Šardzík, G. T. Noble, M. J. Weissenborn, A. Martin, S. J. Webb and S. L. Flitsch, *Beilstein J. Org. Chem.* **2010**, *6*, 699–703.

33. Y. Gao, A. Eguchi, K. Kakehi, Y. C. Lee, *Bioorg. Med. Chem.* **2005**, *13*, 6151–6157.

34. Y. Koshi, E. Nakata, M. Miyagawa, S. Tsukiji, T. Ogawa, I. Hamachi, *J. Am. Chem. Soc.* **2008**, *130*, 245–251.

35. J. Geng, J. Lindqvist, G. Mantovani, G. Chen, C. T. Sayers, G. J. Clarkson, D. M. Haddleton, *QSAR Comb. Sci.* **2007**, *26*, 1220–1228.

36. V. Vicente, J. Martin, J. Jimÿnez-Barbero, J. L. Chiara and C. Vicent, *Chem. Eur. J.* **2004**, *10*, 4240–4251.

37. J.-F. Nierengarten, J. Iehl, V. Oerthel, M. Holler, B. M. Illescas, A. Muñoz, N. Martín, J. Rojo, M. Sánchez-Navarro, S. Cecioni, S. Vidal, K. Buffet, M. Durka, S. P. Vincent, *Chem. Commun.* **2010**, *46*, 3860–3862.

38. Y. Zhang, S. Luo, Y. Tang, L. Yu, K.-Y. Hou, J.-P. Cheng, X. Zeng, P. G. Wang, *Anal. Chem.* **2006**, *78*, 2001–2008.

39. G. Zemplén, E. Pacsu, *Ber. dtsch. Chem. Ges. A/B* **1929**, *62*, 1613–1614.

40. C. W. Tornøe, C. Christensen, M. Meldal, J. Org. Chem. 2002, 67, 3057–3064.

41. H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2021.

42. S. K. Mamidyala, M. G. Finn, *Chem. Soc. Rev.* **2010**, *39*, 1252–1261.

43. V. Hong, A. K. Udit, R. A. Evans, M. G. Finn, *ChemBioChem* **2008**, *9*, 1481–1486.

44. V. Hong, S. I. Presolski, C. Ma, M. G. Finn, *Angew. Chem. Int. Ed.* **2009**, *48*, 9879–9883.

45. S. I. Presolski, V. Hong, S.-H. Cho, M.G. Finn, *J. Am. Chem. Soc.* **2010**, *132*, 14570–14576.

46. L. M. Artner, L. Merkel, N. Bohlke, F. Beceren-Braun, C. Weise, J. Dernedde, N. Budisa, C. P. R. Hackenberger, *Chem. Commun.* **2012**, *48*, 522–524.

47. J. R. Lakowitz in *Protein fluorescence*, 2nd edn, Kluwer Academic/Plenum Publisher, Dordrecht, **1999**.

48. N. Budisa, S. Alefelder, J. H. Bae, R. Golbik, C. Minks, R. Huber, L. Moroder, *Protein Sci.* **2001**, *10*, 1281–1292.

49. S. M. Dimick, S. C. Powell, S. A. McMahon, D. N. Moothoo, J. H. Naismith, E. J. Toone, *J. Am. Chem. Soc.* **1999**, *121*, 10286–10296.

4. Neoglycoconjugates for Multivalent Binding to Asialoglycoprotein Receptor

4.1. Contributions

L. M. Artner, N. Bohlke, J. Hütter, B. Lepenies, N. Budisa, and C. P. R. Hackenberger designed the experiments and the research project. N. Bohlke cloned, expressed, purified and characterized green fluorescent proteins (GFP). M. Schneider was responsible for the computational experiments on GFP mutant stability. J. Hütter performed fluorescent-activated cell sorting (FACS) of modified and unreacted GFPs. L .M. Artner synthesized carbohydrate azides and the THPTA ligand, and performed CuAAC on GFP proteins, as well as their purification and characterization.

4.2. Introduction

Asialoglycoprotein-receptor (ASGP-R) is a C-type lectin, mainly located in the plasma membrane and internal membranes of mammalian hepatic cells. For this reason it was historically called mammalian hepatic lectin (HL).[1], [2] It was first described by Ashwell and Morell, who performed a series of experiments, initially to track serum glycoproteins. They observed that their artificially de-sialylated and labeled glycoproteins are readily cleared from the serum, ending up in parenchymal liver cells. This lead to the hypothesis of liver cells specifically targeting asialolyated glycoproteins.[3]–[6] Later they were able to isolate and characterize the receptor.[7], [8]

ASGP-R is presented as a heterooligomer of two subtypes, H1 and H2, each containing a cytosolic domain, a trans-membrane domain, a stalk reagion and a Ca^{2+}-dependant C-type lectin carbohydrate binding domain at the C-terminus. Primarly it is maintaining the homeostasis of serum glycoproteins through recognition and endocytosis of asialoglycoproteins, thus clearing them from the serum.[1], [9], [10] ASGP-R is internalizing

the recognized glycoproteins *via* clathrin-coated pits. After endosomal fusion, the glycoproteins are transported to lysosomes where they are degraded, whereas the receptor is recycled. [1] This endocytosis-mediating lectin has a very high specificity to glycans terminated with galactose (Gal) and *N*-acetylgalactosamine (GalNAc) residues. [9] Additionally, Baenziger and co-workers found that ASGP-R also binds to sialylated glycoproteins, namely αNeu5Ac-(2→6)-β-D-Gal*p*-(1→4)-D-GlcNAc*p*[11] and αNeu5Ac-(2→6)-β-D-GalNAc*p*-(1→4)-D-GlcNAc*p*,[12] marking its influence on serum glycoprotein regulation.

The exact biological assembly and high-resolved structural models are still not available of ASGP-R. Although the CRD of this hepatic lectin could be crystallized, the remainder of structure and assembly is still not fully elucidated. It is therefore no surprise that various studies were conducted to substantiate a structure-activity relationship with natural or synthetic ligands. Early investigations by Baenziger *et al.*[9] revealed a stronger binding to GalNAc as opposed to Gal. In addition, tri-antennary structures showed higher affinity than their di-antennary and linear counterparts. It was also found that only terminal carbohydrates matter for specific recognition. This lead to the conclusion, that the two to three subunits of the receptor interact with two to three carbohydrates.

Lee *et al.* intensified the research on ASGP-R and expanded the proposed binding hierarchy of antennary glycostructures through experiments with synthetic oligosaccharides: tetra > tri ≫ di ≫ mono. Interestingly, with linear increase in the number of ligands, an exponential increase in the inhibitory potency was observed, except for the change of three to four displayed ligands. It was therefore concluded that the fourth galactose arm is not extensively contributing to the binding of the ligand. Therefore it was assumed that a triantennary ligand architecture suffice for asialoglycoprotein receptor.

Another focus of this research on ASGP-R was the determination of an optimal carbohydrate display. Lee *et al.* therefore applied different techniques, from high-resolution NMR to fluorescence-based assays to supplement the information already known about the biological assembly of membrane-bound ASGP-R. The came up with a proposal for the optimal distance and spatial arrangement of carbohydrate ligands for binding (*vide supra*, **Figure 1-6**, p. 13, and **Figure 4-19**, p. 76) based on their observations with synthetic oligosaccharides and glycoconjugates.[2], [15]–[23]

ASGP-R has been considered as a potential and valuable target for drug delivery. [24]–[27] The high specificity and the abundant localization in hepatocytes, as well as a very high endocytotic turn-over rate, makes this hepatic lectin an ideal candidate. As of now, various scaffolds[28]–[31] were described for efficient drug and gene delivery in hepatocytes. Neoglycoproteins were also found to be useful as conjugates for drug delivery in many cases, although with some encountered drawbacks. [32], [33]

Nevertheless, the groups of Lee and others[34], [35] showed that carbohydrate-protein
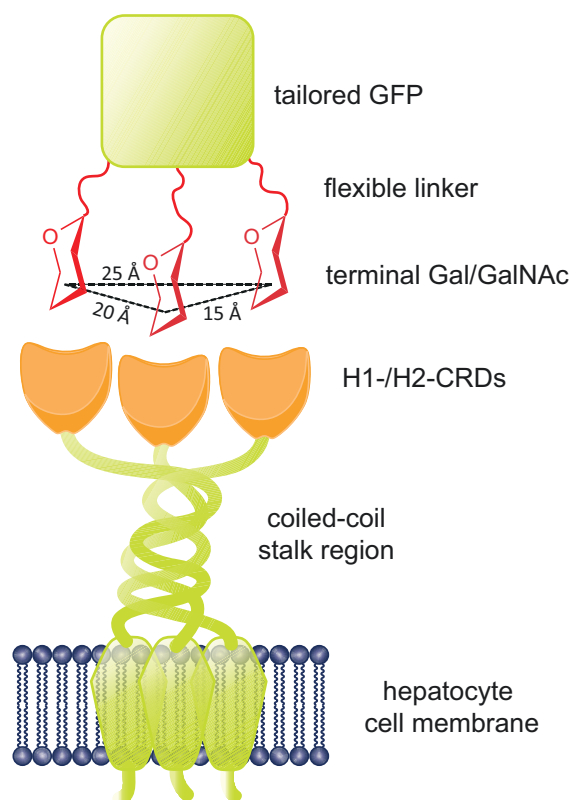
Figure 4-19: Proposal for a rationally designed GFP that can bind to ASGP-R.
Based on observations by the groups of Lee and others, a multivalent-binding neoglycoprotein based on GFP should be designed to perfectly fit the suggested physiological arrangement of ASGP-R.

conjugates are a valuable tool for the characterization of a receptor. Many others described various strategies for the synthesis and application of statistically and site-specifically glycosylated neoglycoproteins. Latter is from interest for defined, homogeneous proteins and is usually achieved by native chemical ligation or a with a combination of site-directed mutagenesis, incorporation of noncanonical amino acids, and bioorthogonal conjugation strategies. Well-characterized glycoproteins also build the experimental foundation for potential "glyco vaccines", carbohydrate therapeutics and probes, as well as model systems to study interactions of glycosylated proteins *in vitro* and *in vivo*.

To take this concept of neoglycoproteins one step in the direction of rational protein design, functional proteins have to be considered more as architectures than biologically active material. A protein-based scaffold with an unique biological read-out and a high specificity for asialoglycoprotein receptor was envisioned. For this reason, the well known green fluorescent protein (GFP), an auto-fluorescent protein initially derived from *Aequorea victoria*[36]–[39], was observed more closely in regards to information already known about asialoglycoprotein receptor. Judged from readily available crystal structures, GFP resembles a barrel, with C- and N-terminus located at the bottom, and

various loop regions at the top of it. The optimal spatial arrangement within the binding Gal and GalNAc as known from the literature was applied to the crystal structure of GFP. Molecular modeling and calculations on the relative protein stability provided the e information for the most stable mutations within the loop regions. Selective pressure incorporation of non-canonical amino acid homopropargylglycine (Hpg) then furnished the fluorescent protein scaffold. Subsequently a series of GalNAc and Man azides, varying in the length of the aglycon linker, were conjugated to ncAA-bearing GFP variants to generate neoglycoproteins that vary in linker-length and the presented number of carbohydrates. Finally, fluorescence-activated cell sorting provided data for the effectiveness of those fluorescent protein scaffolds, and additional rationale for the biological assembly and spatial arrangement of ASGP-R and its ligands, respectively.

4.3. Results

Rational design of a methionine-free superfolder variant of green fluorescent protein (GFP) provided the basis for a fluorescent probe. Incorporation of ncAAs in the superloop-regions at one side of the GFP-barrel structure enables the installation of carbohydrate moieties by copper-catalyzed alkyne-azide cycloaddition (CuAAC), thus creating a perfectly matched protein-based ligand for asialoglycoprotein receptor. This endocytotic receptor was binding neoglycoproteins with different efficiency based on the nature of their functionalization and degree of valency.

4.3.1. Synthesis of carbohydrate azides

Since asiologlycoprotein receptor is known for it's strong binding to terminal galactose (Gal) and *N*-acetylgalactosamine (GalNAc) bearing glycans, a series of glycosyl azides were envisioned. As a strong binding control, GalNAc was used, whereas a mannose-based series should provide a weak binding control. Inside those series, a variation of linker length should be established, to investigate its relevance for optimal binding.

Starting out from commercially available D-galactosamine hydrochloride, the first step was acetylation, using acetic anhydride and dry pyridine. Usually fully acetylated GalNAc, 2-acetamido-1, 3, 4, 6-tetra-*O*-acetyl-2-deoxy-β-D-galactopyranoside (**11**), was obtained in good yields and purity. The first carbohydrate azide in the GalNAC series (**Scheme 4-7**, p. 78), **G1**, was obtained *via* the chloride intermediate, 1-chloro-2-acetamido-3, 4, 6-tetra-*O*-acetyl-2-deoxy-β-D-galactopyranoside (**12**), that could be obtained easily by reacting **11** with $TiCl_4$. Subsequent substitution of the chloride to the azide with NaN_3 resulted in **G1Ac**. The carbohydrate acetate was deacetylated with catalytic amounts of sodium methoxide in methanol (*Zemplén conditions*), yielding the
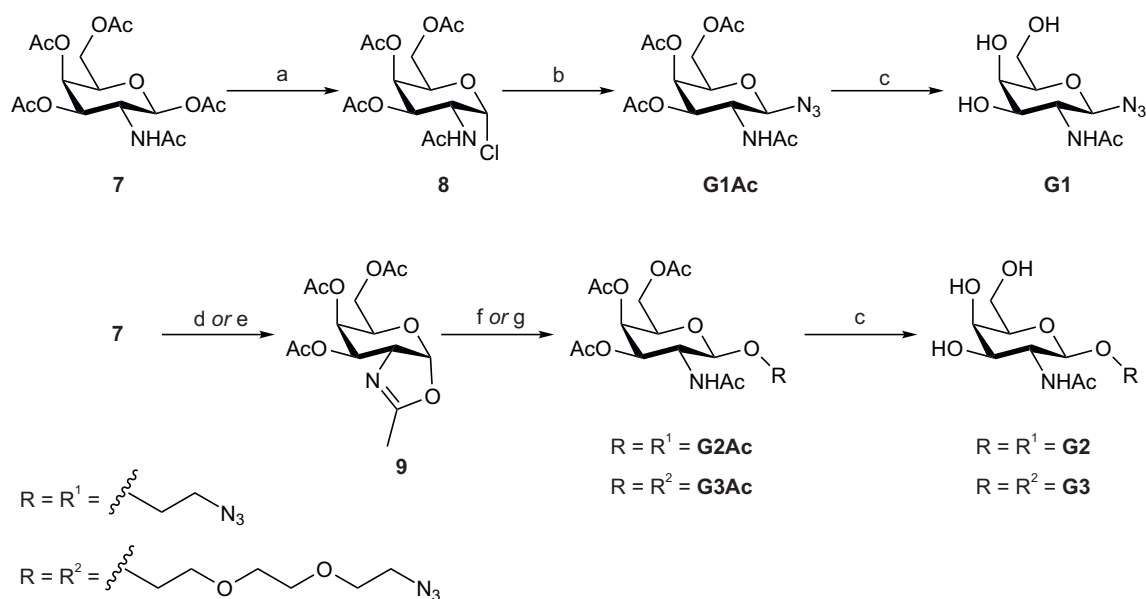
Scheme 4-7: Synthesis of *N*-acetyl galactosamine azides G1–G3.
a) TiCl$_4$, CH$_2$Cl$_2$, r.t.; b) NaN$_3$, DMSO, r.t.; c) NaOMe, MeOH, r.t.; d) FeCl$_3$, CH$_2$Cl$_2$, r.t.; e) TMSOTf, C$_2$H$_4$Cl$_2$, 50 °C; f) cat. H$_2$SO$_4$, N$_3$EtOH, CH$_2$Cl$_2$, r.t.; g) TMSOTf, N$_3$(EtO)$_2$EtOH, CH$_2$Cl$_2$, 0 °C;

first GalNAc azide of this series, **G1** (**Scheme 4-7**).

For both the 2-azidoethanol and 2-(2-[2-azidoethoxy]ethoxy)ethanol bearing GalNAc a different strategy was applied (**Scheme 4-7**). Here oxazoline **13** was first formed as an intermediate. Although different procedures reported equally good yields for the formation of **13** with FeCl$_3$ in dichloromethane or TMSOTf in 1,2-dichloroethane, only the latter method resulted in good yields steadily. It is suspected that the air and moisture sensitivity of FeCl$_3$ is strongly influencing the reaction outcome, since good yields could only be achieved by drying FeCl$_3$ overnight. Therefore the usage of TMSOTf was more straight forward in obtaining reasonable amounts of oxazoline **9**. For the glycosylation, simply oxazoline **9** was reacted with the glycosyl-acceptor of choice under a catalytic amount of acid – either H$_2$SO$_4$ or TMSOTf – to GalNAc azides **G2Ac** and **G3Ac**, respectively, in acceptable yields. After the removal of *O*-acetyl protecting groups with sodium methoxide, the desired carbohydrate azides, **G2** and **G3**, could be obtained.

For the synthesis of mannose azides, acetylation of the unprotected sugar was also the first starting point (**Scheme 4-8**, p. 79). Although **M2** could also be directly generated from unprotected D-mannose by reacting it without solvent directly in 2-bromoethanol with acidic ion-exchange resin, the route *via* 1, 2, 3, 4, 5-tetra-*O*-acetyl-D-mannopyranoside (**14**) is more convenient due to easier purification. Subsequent glycosylation with 2-azidoethanol and 2-(2-[2-azidoethoxy]ethoxy)ethanol under Lewis acid conditions with either BF$_3$·OEt$_2$ or TMSOTf, following deacetylation, yielded
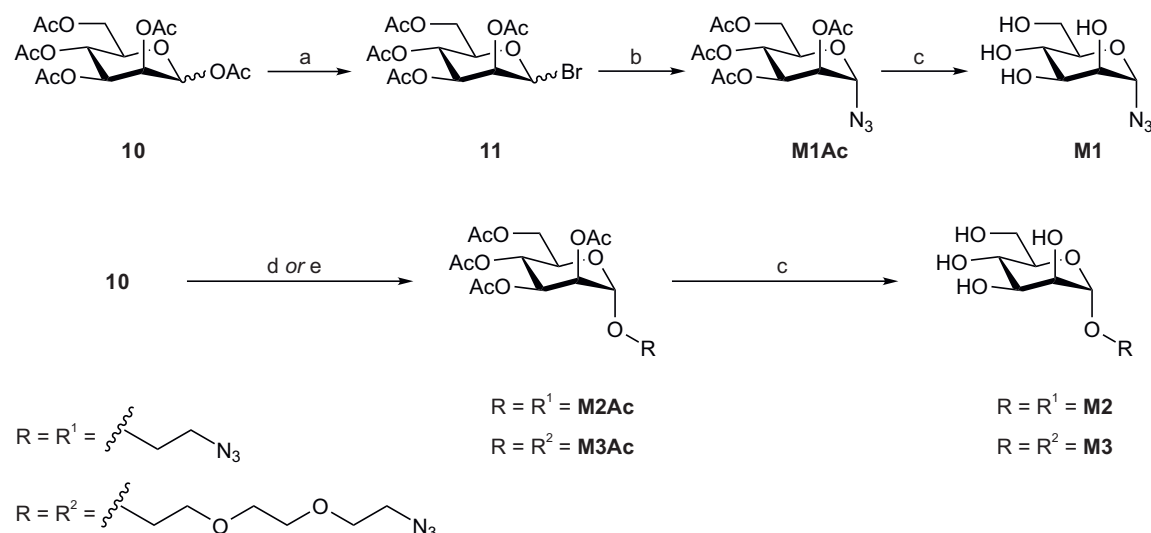
Scheme 4-8: **Synthesis of mannose azides M1–M3.**
a) AcOH, HBr/AcOH, CH_2Cl_2, 0 °C → r.t.; b) NaN_3, TBAS, 1 M K_2CO_3/CH_2Cl_2, r.t.; c) NaOMe, MeOH, r.t.;
d) $BF_3·OEt_2$, N_3EtOH, CH_2Cl_2, 0 °C → r.t.; e) TMSOTf, $N_3(EtO)_2EtOH$, CH_2Cl_2, r.t.

the sought mannose azides **M2** and **M3** in acceptable yields (**Scheme 4-8**). **M1** was obtained from bromide **11** followed by deacetylation Zemplén conditions of **M1Ac** (**Scheme 4-8**).
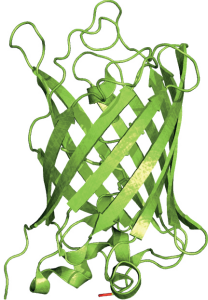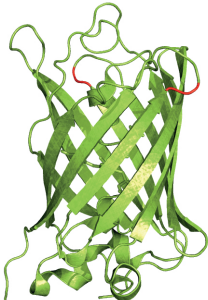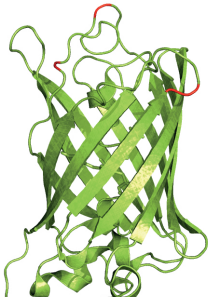
4.3.2. Rational design of a green fluorescent protein for binding to asialoglycoprotein receptor

(Performed by Nina Bohlke, Technische Universität Berlin, and Michael Schneider, Technische Universität Berlin)

Because of the proposed geometry and assembly by Lee *et al.* (**Figure 4-19**, p. 76), the crystal structure of a superfolder GFP (PDB-ID 2B3P) was examined for applying the geometry of the binding triangle at the superloops of GFP. Nina Bohlke (Technische Universität Berlin) designed variants of trivalent GFP based on known mutations (**Table 4-2)**. Subsequently Michael Schneider (Technische Universität Berlin) applied *in silico* methods to predict the influence of the chosen sites of mutations on the stability of the protein. Briefly, the crystal structure of a superfolder GFP (PDP ID: 2B3P)[40] was taken as the basis of the calculations. After various calculations, like minimization of the energy of the structure by a ROSETTA all-atom force-field[41], and optimal placement of the side chains with a Monte-Carlo algorithm and the rotamer library of Dunbruck[42], a stable mutant was identified. Applied to the methionine-free variant of a superfolder GFP, GFP-hs1RM[43], a tailored GFP for ASGP-R could be provided by Nina Bohlke (Technische Universität Berlin). In addition to methionine mutations at T50M, D134M, and E143M,

a bivalent variant GFP (T50M, K102M) was provided for experimental comparison. Due the inefficient cleavage of homopropargylglycine at the N-terminus,[44], [45] which leads to inhomogeneous proteins, another series of mutants were prepared. N. Bohlke designed the new series of GFP to be equipped with a cleavable poly-histidine tag at the N-terminus for purification. For this a tobacco etch virus (TEV) protease cleavage sequence[46] was inserted between the N-terminal affinity-tag and the genuine protein. Since the monovalent control was solely based on the methionine-free GFP without mutations in the loop regions and only bearing the non-canonical amino acid at the N-terminus, it was not subject to this matter and had the cleavable purification tag fused to the C-terminus. Finally, three different green fluorescent proteins, GFP-hs1RM (**GFPx1**), GFP(T50M, K102M) (**GFPx2**), and GFP(T50M, D134M, E143M) (**GFPx3**)

Table 4-2: Tailored green fluorescent proteins

Entry	Mutations to GFP-hs1RM	Approx. distances [a]	Abbreviated	3D-outline [b]
1	- (1 Hpg at the *N*-terminus)	-	**GFPx1**	
2	T50M, K102M (2 Hpg in the loop regions)	28 Å	**GFPx2**	
3	T50M, D134M, E143M (3 Hpg in the loop regions)	25 Å, 15 Å, 22 Å	**GFPx3**	

[a] The distances were measured with PyMol Measurement Wizard from each Cα to the next of the mutated amino acids.

[b] Sites of incorporated Hpg are marked red. The figures were created with PyMol using PDB-entry 2B3P.

could be provided by the Budisa lab (**Table 4-2**, p. 80). This series have different valencies, numbers of incorporated Hpg, and spatial distribution of ncAAs, ideal for functionalization and biological evaluation (**Table 4-2**, p. 80).

4.3.3. Modification of ASGP-R-binding green fluorescent protein

Initial experiments for decorating GFP with carbohydrate azides *via* CuAAC were performed on heterogenous variants of GFP(T50M, D134M, E143M), which had the His-tag and TEV cleavage site at the C-terminus, were promising. Nevertheless modification of the homogeneous GFP(T50M, D134M, E143M; M1ins*TEV*) had to be optimized additionally. During this it was found out that final concentrations of protein inside the reaction mixture should not exceed 40 μM. Interestingly, higher concentrated protein solutions were only recovered in lower amounts. Since longer reaction times alone did not yield fully functionalized **GFPx3**, sequential addition of the copper reducing agent, sodium ascorbate, was tried. With this the reaction outcome and the recovery of the protein improved a lot. The final optimization was achieved by combining the sequential addition with in-between micro-workups. In general, this meant adding sodium ascorbate in two equal portions over a duration of 4 hours, followed by rebuffering of the protein by diafiltration with centrifugal filter tubes with a molecular weight cut-off of 10 kDa. Subsequently the reaction was re-initiated by adding the same amount of reagents again. The optimized procedure allowed extended reaction times of approximately 28 hours and a subsequent high protein recovery of 60–90% (based on determined concentrations) after desalting with an ÄKTA FPLC system and concentration with centrifugal spin filter tubes (A, **Figure 4-20**, p. 82). A similar procedure to ensure high rates of functionalization was published before by the group of B.G. Davis *et al.*[47] They described that full functionalization of bacteriophage Qβ with carbohydrate dendrimers could be achieved by additionally adding a reactive Cu(I) after a given amount of time. Those findings support the herein described approach as a valuable improvement of the initially described procedure by M.G. Finn *et al.*[48]

4.3.3.1. Mass spectrometry of neoglycoproteins

For the evaluation of the reaction outcome, previously modified GFP were analyzed by mass spectrometry. For this, matrix-assisted laser desorption ionization (MALDI) mass spectrometry was chosen. As stated before, the high tolerance of MALDI towards salt decontamination and the primarily generation of single and double charged m/z values makes it an ideal method for protein analysis. In our approach three different instruments were used to analyze previously synthesized neoglycoconjugates: a) Bruker Ultraflex II, b) Applied Biosystems 4700 Proteomics Analyzer, and c) Applied Biosystems Sciex
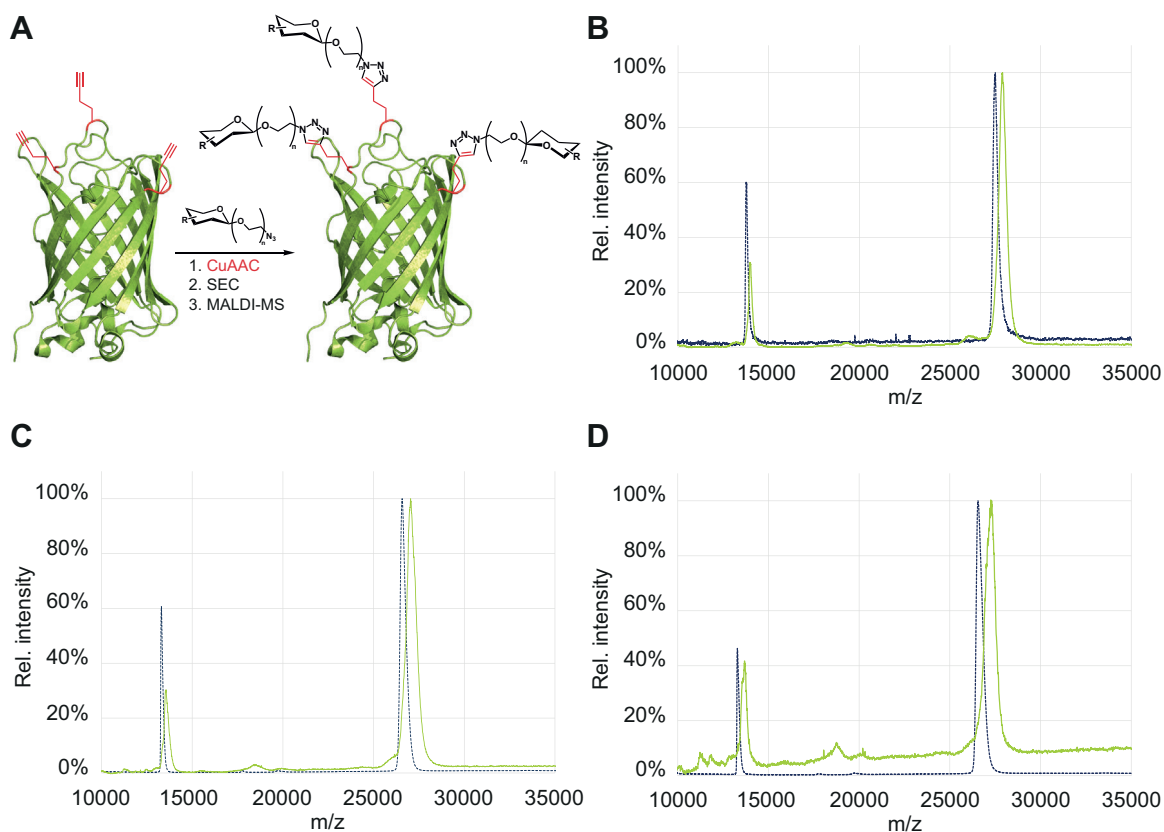
Figure 4-20: Modification of Hpg-bearing GFP.

A) Genetically modified GFPs bearing homopropargylglycine were functionalized with carbohydrate azides *via* an optimized CuAAC protocol, purified by SEC and analyzed by MALDI-TOF-MS. B) MALDI-TOF mass spectrometry spectrum of **GFPx1** before (dotted blue line) and after (solid green line) CuAAC with azide **G3**. C) MALDI-TOF mass spectrometry spectrum of **GFPx2** before (dotted blue line) and after (solid green line) CuAAC with azide **G3**. D) MALDI-TOF mass spectrometry spectrum of **GFPx3** before (dotted blue line) and after (solid green line) CuAAC with azide **G3**.

5800. To obtain high quality spectra, sample preparation procedures had to be carried that help removing disturbing components of the purified neoglycoprotein solutions. Therefore the proteins were precipitated in ice-cold acetone and again solubilized in water. With this procedure a majority of the salt was removed.

Acidifying the samples with trifluroacetic acid or sodium acetate gave the best result when samples were spotted with 2,5-dihydroxy acetophenone (DHAP) as matrix. DHAP was chosen as a matrix because of its ability to prevent in-source decay and suppression of low molecular weight compounds. In addition it gives apart from the usual single charged m/z also the double and sometimes even the triple charged masses. Incorporating those m/z values for the calculation of the observed molecular weight of the uncharged protein, higher mass accuracy can be obtained. This methodology is comparable to electro-spray ionization, were peaks of highly charged ions are detected. However, the MALDI-TOF mass spectra obtained by this technique still suffered from
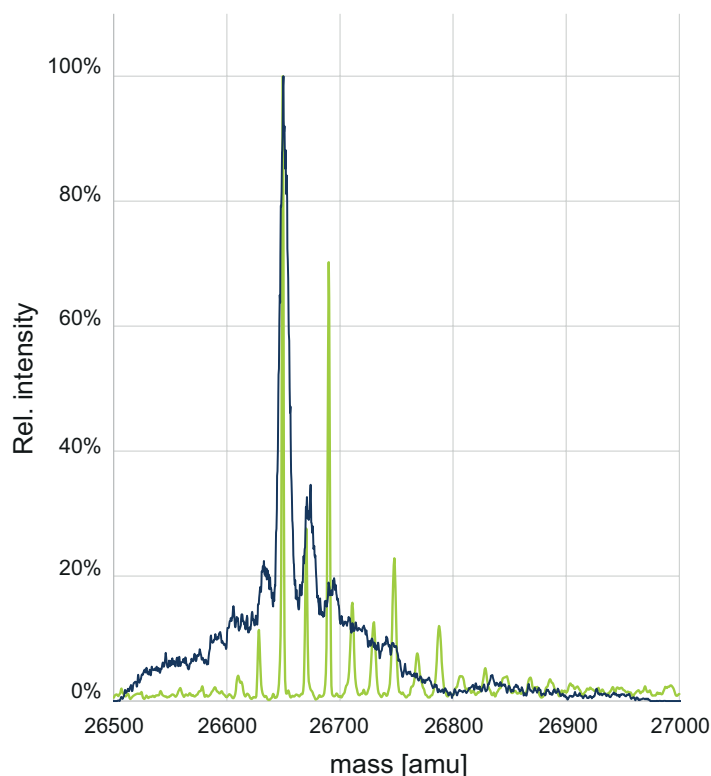
Figure 4-21: Deconvoluted ESI-MS spectra of unreacted, GFPx3.

The deconvoluted electro-spray mass spectra of Orbitrap- (dotted blue line) and TOF-analyzers (solid green line) show major abundance of GFPx3 bearing three Hpg in the loop region. However, higher m/z values correspond to incomplete incorporation of Hpg, resulting of a shift of approximately +22 Da per Met. Due to lower resolution, Na⁺ adduct peaks partially overlap in the deconvoluted ESI-TOF spectrum.

peak broadening, which made accurate determination of molecular weight and product distribution impossible (B–D, **Figure 4-20**, p. 82). This observation can be explained by considering the inhomogeneity of the samples. When taking a look at well resolved, deconvoluted ESI mass spectra (**Figure 4-21**), the expected mass of the corresponding protein can be observed with the highest intensity. However, additional masses with shifts of +22 Da are also present. This masses fit to GFP where Met is incorporated instead of Hpg. Consequently each undesired amino acid results in an increased mass of +22 Da. In addition, common adducts that can stick to the analyte, like sodium, potassium and partially lithium ions can also be detected. This is readily observed in MALDI-TOF mass spectra which adds to peak broadening due to sample inhomogeneity.

4.3.3.2. Fluorescence measurements of neoglycoproteins

Previously it was demonstrated that fluorescence can be used to prove structural integrity of a protein after modification (**3.3.2**, pp. 64ff., and references therein). In addition to this, GFP has a specific excitation/emission profile, that adds another
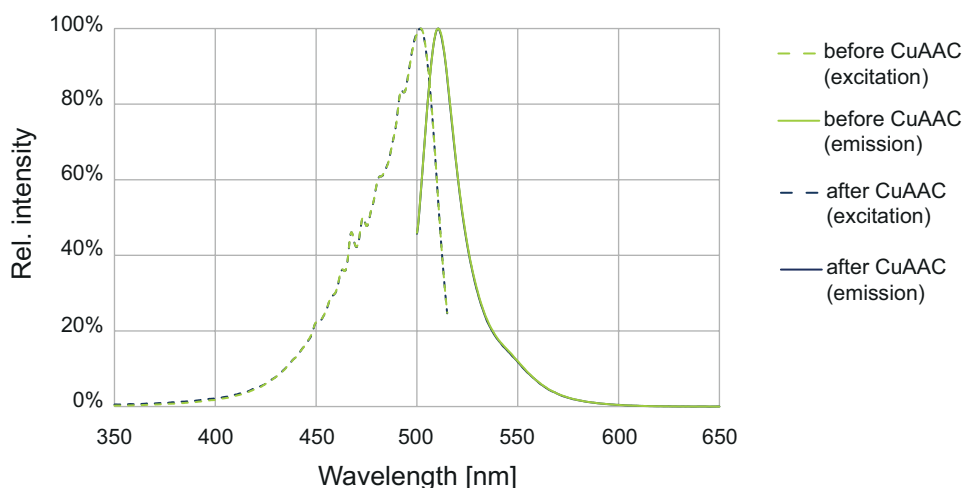
Figure 4-22: Excitation/emission spectra of green fluorescent proteins.
The trivalent, Hpg-bearing mutant GFP(T50M, D134M, E143M) (green lines) and its modified counterpart (blue lines) were analyzed with a fluorescence spectrophotometer. For better visualization and evaluation, obtained excitation/emission spectra were normalized for their intensities. The curves show identical shape and essentially are overlapping.

read-out to this experiment. Therefore excitation/emission spectra of trivalent GFP before and after CuAAC were recorded. The range of the wavelength was set to 350–515 nm for excitation and 500–650 nm for emission with a bandpass of 3 nm Due to the sensitivity of those measurements, obtained spectra were normalized for the measured intensity for better comparison. When looking at the shape of the obtained curves, no difference is visible (**Figure 4-22**). This confirmed the structural integrity of GFP before and after CuAAC.

4.3.4. ASGP-R-mediated cell uptake of neoglycoproteins

(Performed by Julia Hütter, Max-Planck-Insitute for Colloids and Interfaces, Berlin)

To prove the applicability of the newly generated artificial glycoproteins, previously generated glycoconjugates were tested for their specificity and efficiency with fluorescence-activated cell sorting (FACS) in the lab of Dr. Bernd Lepenies by Julia Hütter. For this, two series of experiments were conducted. The first initial experiments were focused solely on the mention uptake efficiency and specificity. This was probed by varying the length of the linker that separates the carbohydrate moiety from the protein scaffold. Within those experiments an interesting phenomenon could be observed: Compared to the unmodified, homopropargylglycine containing GFP, uptake of both mannose- an *N*-acetylgalactosamine-functionalized GFP was in range of each other with ethoxy-spaced conjugates (**Figure 4-23**, p. 85). For triethylenglycol-linked neoglycoproteins the specificity and the uptake efficiency changed dramatically (**Figure**
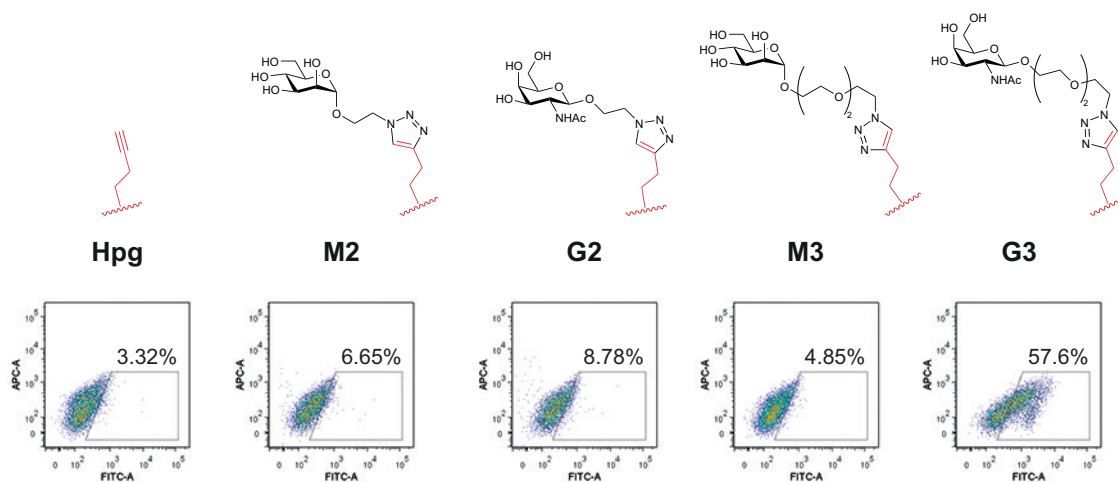
Figure 4-23: FACS analysis of neoglyco-GFPs with different ligands and linker length.
"ASGP-R-GFP" was modified *via* CuAAC with carbohydrate azides **M2**, **M3**, **G2** or **G3**. Subsequently, functionalized and unfunctionalized GFPs were subjected to fluorescence activated cell sorting (FACS) at a 10 µm concentration of neoglycoprotein. All conjugates were taken up between approximately 3 and 9%. Only with the longest linker, a strong increase for GalNAc-bearing neoglycoproteins could be observed.

4-23, p. 85). Now low µm concentrations were enough to ensure internalization of trivalent GFP conjugates – but only when GalNAc was presented. For Man-bearing GFP this could not be observed. To further ensure the specificity of the uptake, **GFP-G3** conjugates were incubated with known asialoglycoprotein-receptor substrates. The results of this competitive uptake study not only revealed specific uptake-inhibition of **GFP-G3**, it also pointed towards a different uptake of **GFP-M3** since those values were not influenced (*data not shown*).

The second experimental series should shed light on the influence of multivalent binding phenomenon and ligand presentation. For this, **G3**-conjugated monovalent and bivalent GFP were also tested and their level of internalization was compared to obtained values of the trivalent **GFP-G3**. As expected, trivalent presentation of carbohydrates was superior to mono- and bivalent one (**Figure 4-24**, p. 86). Additionally, when the effective concentration of carbohydrates was compared between the glycoconjugates, it was found that trivalently presented GalNAc was taken up more strongly than its bi- and monovalent counterparts (**Figure 4-24**, p. 86). It should be noted here, that – judged from mass analysis – the proteins present a heterogenous mixture between all variations of functionalization. However, based on MALDI-TOF mass spectrometry, a similar distribution of bi- and trivalently functionalized GFP was assumed.
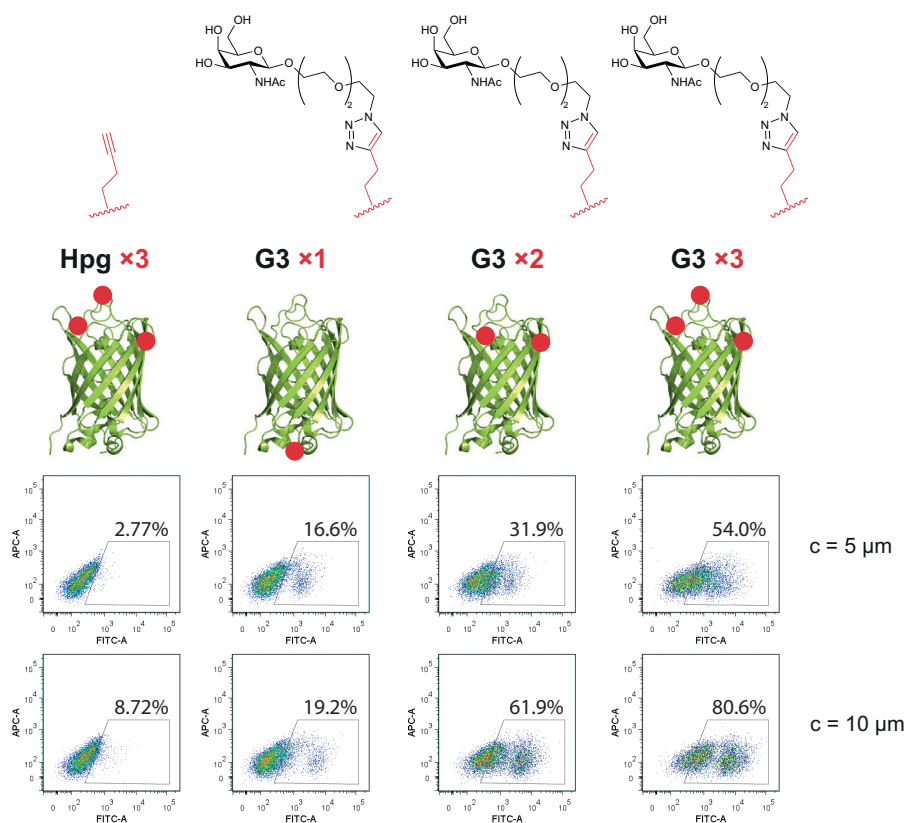
Figure 4-24: FACS analysis of neoglyco-GFPs with different degrees of valency.
Functionalized and unfunctionalized GFP were subjected to fluorescence activated cell sorting (FACS) to evaluate their specificity towards ASGP-R and their potency regarding endocytosis. The upper row shows experiments with a concentration of 5 μM and the bottom row at 10 μM of neoglycoprotein. With an increasing number of presented ligands, the uptake was stronger.

4.4. Discussion

Mannose and *N*-acetylgalactoseamine azides were obtained in moderate yields but very good stereoselectivity. Derivatives from the Man-series were α-anomers, whereas the GalNAc-series was comprised solely of β-anomers. The deprotected carbohydrate azides were coupled to green fluorescent proteins, either bearing one, two or three alkynes. Although CuAAC was often reported as an easily applicable bioconjugation reaction, limitations were found for the herein described system. Inactivation of the copper species usually is the primary reason for bad conversions, accompanied with oxidative damage to biomaterial.[49]–[52] The use of THPTA as a stabilizing ligand, should counter those side reactions and ensure efficient proceeding of the reaction. However, as also reported by Wu *et al.*[53] THPTA has its limitations as well. Switching to an approach were the reagents are exchanged and renewed after a given time still improved the outcome of the reaction. Nevertheless, full conversion to one species of fully doubly or triply modified GFP could not be verified by MALDI-TOF mass

86
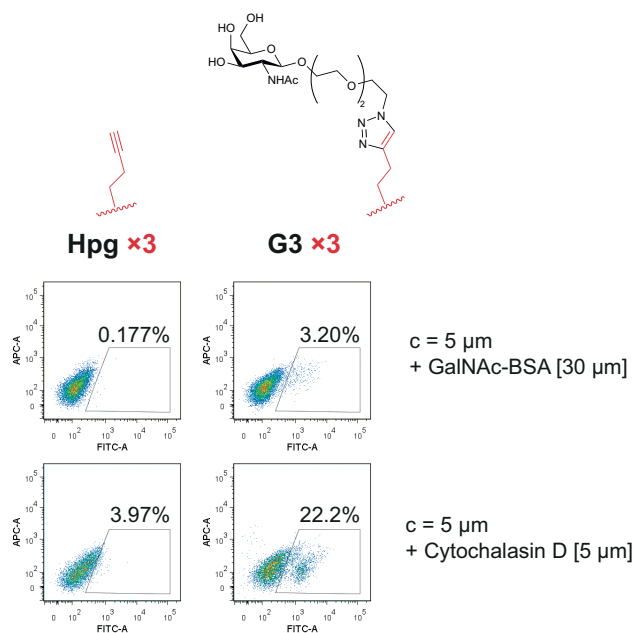
Figure 4-25: Inhibition of endocytosis.
The neoglycoprotein GFP×3-**G3** (5 µM) was incubated together with either GalNAc-BSA conjugated to bovine serum albumine (GalNAc-BSA, 30 µM) or endocytosis inhibitor cytochalasin D (5 µM), resulting in reduced uptake in both experiments.

spectrometry. Interestingly, strong peak broadening was observed, which indicated multiple species of similar masses. The reasons for this are probably adducts of multiple ions in addition to a mixture of differently modified proteins. The tendency to cluster multiple adducts and ions and therefore generating broad peaks in the positive linear mode of the MS is especially pronounced with glycoproteins. The higher the number of unprotected carbohydrates attached, the more possibilities of adducts are given.[54] A solution to that challenge usually involves suppression or removal of adduct ions. In the case of multivalent ASGP-R binding GFPs, 2,5-DHAP and fairly strong acidic conditions were used to suppress alkali-ion addition. Indeed, higher accuracies could be achieved, especially because multiple charged species could also be factored into the calculations of the actual mass. However, only moderate suppression of adduct ions was observed, still resulting in broadened peaks. When the samples were spiked with sodium acetate – a common technique for glycan analysis – only sensitivity changed noticeable. This observation pointed more towards actual product heterogeneity, rather than just adduct formation. The tendency of the protein to bind adduct ions, even without carbohydrates was further visualized by total entropy deconvolution of UPLC-ESI-TOF mass spectrometry spectra of unmodified trivalent GFP (**Figure 4-21**, p. 83). In addition, incomplete incorporation of homopropargylglycine contributed to initial heterogeneity of the starting material (**Figure 4-21**, p. 83). LC-ESI-MS was sufficient for unreacted proteins, however, for functionalized samples, only strongly

suppressed spectra could be obtained. Reasons for this are speculative, but it is possible that carbohydrate entities in combination with adduct ions lead to decreased signal intensities.

However, proper sample preparation still lead to acceptable MALDI-TOF mass spectra. Masses between total and incomplete functionalization were obtained and the purified neoglycoproteins were evaluated by FACS. Interestingly, the linker length strongly contributed to the efficiency of internalization. Initial screenings revealed unspecific uptake of Man- and GalNAc-conjugated GFP if the spacer length was less than PEG-3 (**M1**, **M2**, and **G1**, **G2**, respectively). Not only did the uptake increase significantly when PEG-3 (**G3** and **M3**) was used, but also the specificity for this conjugate. Incubation of ASGP-R expressing HepG2 cells with GalNAc-BSA[55] or cytochalasin D[56] together with neoglycoproteins resulted in tremendously reduced uptake (**Figure 4-25**, p. 87). This observation suggests specific, ASGP receptor mediated uptake of **G3**-conjugated neoglyco-GFP. When asioalofetuin was added to neoglycoproteins with a shorter linker, endocytosis was inhibited equally for both series (*data not shown*).

The results on the influence of the linker length were not all that surprising when previously reported studies of different glycoconjugates were considered. Looking at the structures of those small molecule probes[57], [58], [19]–[22], there seemed to be a requirement of a minimal distance for specific binding of ASGP-R. Interestingly, even longer linker than triethylenglykol[57] resulted in high uptake.

4.5. Conclusion

Starting from what was previously known and reported about the biological and physiological assembly of multimeric asialoglycoprotein receptor, a rational approach towards a specific and multivalent binding probe was envisioned. *In silico* design and stability predictions based on a crystal structure of GFP yielded candidates for expression and incorporation of non-canonical amino acids. An improved CuAAC protocol for the conjugation of those alkyne-bearing green fluorescent proteins with mannose and *N*-acetylgalactosamine azides resulted in fluorescent neoglycoproteins with good protein recovery (60-90%).

Those various neoglycoproteins were all unspecifically endocytosed as FACS analysis revealed. However, under certain conditions – namely, trivalent configuration and a linker length of PEG-3 – a tremendous increase in uptake by HepG2 cells was noticed. Competitive inhibition of the uptake with GalNAc-BSA and cytochalasin D, as well as asialofetuin further implied specific uptake of PEG-3-spaced neoglycoproteins. This could not be shown with shorter linkers.

With this, the principal applicability of the concept "rational protein design" was

shown, opening up the possibility of further advancements and improvements of this methodology.

However, the heterogeneity of the analyzed samples present a strong influence of the reproducibility of the experiments: Batch to batch discrepancies lead to stronger deviations in cellular uptake studies. Aside from those findings, a solid trend could still be deducted from this experiments. Future studies and improved resolution of mass spectra to determine the ratio between the different stages of functionalization of bi- and trivalent GFP would lead to better understanding and rationalization of those initial results.

4.6. References

1. M. Spiess, *Biochemistry* **1990**, *29*, 10009–10018.

2. Y. C. Lee, R. T. Lee in *Carbohydrates in Chemistry and Biology* (Eds.: B. Ernst, G. W. Hart, P. Sina), WILEY-VCH Verlag GmbH, Weinheim, Germany, **2000**.

3. A. G. Morell, C. J. A. Van Den Hamer, I. H. Scheinberg, G. Ashwell, *J. Biol. Chem.* **1966**, *241*, 3745–3749.

4. A. G. Morell, R. A. Irvine, I. Sternlieb, I. H. Scheinberg, G. Ashwell, J. Biol. Chem. **1968**, 243, 155–159.

5. A. G. Morell, G. Gregoriadis, I. A. Scheinberg, J. Hickman, G. Ashwell, *J. Biol. Chem.* **1971**, *246*, 1461–1467.

6. W. E. Pricer, G. Ashwell, *J. Biol. Chem.* **1971**, *246*, 4825–4833.

7. R. L. Hudgin, W. E. Pricer, G. Ashwell, R. J. Stockert, A. G. Morell, *J. Biol. Chem.* **1974**, *249*, 5536–5543.

8. T. Kawasaki, G. Ashwell, *J. Biol. Chem.* **1977**, *252*, 6536–6543.

9. J. U. Baenziger, D. Fiete, *Cell* **1980**, *22*, 611–620.

10. G. Ashwell, J. Harford, *Annu. Rev. Biochem.* **1982**, *51*, 531–554.

11. E. I. Park, Y. Mi, C. Unverzagt, H.-J. Gabius, J. U. Baenziger, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 17125–17129.

12. L. M. Steirer, E. I. Park, R. R. Townsend, J. U. Baenziger, *J. Biol. Chem.* **2009**, *284*, 3777–3783.

13. D. T. Connolly, R. R. Townsend, K. Kawaguchi, W. R. Bell, Y. C. Lee, *J. Biol. Chem.* **1982**, *257*, 939–945.

14. Y. C. Lee, R. R. Townsend, M. R. Hardy, J. Lönngren, J. Arnarp, M. Haraldsson, H. Lönn, *J. Biol. Chem.* **1983**, *258*, 199–202.

15. M. R. Hardy, R. R. Townsend, S. M. Parkhurst, Y. C. Lee, *Biochemistry* **1985**, *24*, 22–28.

16. R. T. Lee, Y. C. Lee, *Biochemistry* **1986**, *25*, 6835–6841.

17. R. R. Townsend, M. R. Hardy, T. C. Wong, Y. C. Lee, *Biochemistry* **1986**, *25*, 5716–5725.

18. R. T. Lee, P. Lin, Y. C. Lee, *Biochemistry* **1984**, *23*, 4255–4261.

19. R. T. Lee, Y. C. Lee, *Glycoconjugate J.* **1987**, *4*, 317-328.

20. R. T. Lee, K. G. Rice, N. B. N. Rao, Y. Ichikawa, T. Barthel, V. E. Piskarev, Y. C. Lee, *Biochemistry* **1989**, *28*, 8351–8358.

21. K. G. Rice, O. A. Weisz, T. Barthel, R. T. Lee, Y. C. Lee, *J. Biol. Chem.* **1990**, *265*, 18429–18434.

22. R. T. Lee, Y. C. Lee, *Bioconjugate Chem.* **1997**, *8*, 762–765.

23. Y. C. Lee, *J. Biochem.* **1997**, *121*, 818–825.

24. M. Nishikawa, *Biol. Pharm. Bull.* **2005**, *28*, 195–200.

25. L. M. Bareford, P. W. Swaan, *Adv. Drug Deliv. Rev.* **2007**, *59*, 748–758.

26. Y. Malam, M. Loizidou, A. M. Seifalian, *Trends Pharmacol. Sci.* **2009**, *30*, 592–599.

27. N. Mishra, N. P. Yadav, V. K. Rai, P. Sinha, K. S. Yadav, S. Jain, S. Arora, *Biomed. Res. Int.* **2013**, *2013*, 1–20.

28. G. J. L. Bernardes, R. Kikkeri, M. Maglinao, P. Laurino, M. Collot, S. Y. Hong, B. Lepenies, P. H. Seeberger, *Org. Biomol. Chem.* **2010**, *8*, 4987–4996.

29. S. Hong, P. R. Leroueil, I. J. Majoros, B. G. Orr, J. R. Baker, M. M. Banaszak Holl, *Chem. Biol.* **2007**, *14*, 107–115.

30. C. Managit, S. Kawakami, F. Yamashita, M. Hashida, *J. Pharm. Sci.* **2005**, *94*, 2266–2275.

31. J. Wu, T.-M. Sun, X.-Z. Yang, J. Zhu, X.-J. Du, Y.-D. Yao, M.-H. Xiong, H.-X. Wang, Y.-C. Wang, J. Wang, *Biomater. Sci.* **2013**, *1*, 1143–1150.

32. D. K. F. Meijer, P. van der Sluijs, *Pharm. Res.* **1989**, *6*, 105–118.

33. V. P. Torchilin, A. N. Lukyanov, *Drug Discov. Today* **2003**, *8*, 259–266.

34. M. J. Krantz, N. A. Holtzman, C. P. Stowell, Y. C. Lee, J. W. Weiner, H. H. Liu, *Biochemistry* **1976**, *15*, 3963–3968.

35. C. P. Stowell, R. T. Lee, Y. C. Lee, *Biochemistry* **1980**, *19*, 4904–4908.

36. O. Shimomura, F. H. Johnson, Y. Saiga, *J. Cell. Comp. Physiol.* **1962**, *59*, 223–239.

37. R. Y. Tsien, *Annu. Rev. Biochem.* **1998**, *67*, 509–544.

38. M. Chalfie, S. Kain, *Green fluorescent protein. Properties, applications, and protocols*, Wiley-Interscience, Hoboken, N.J, **2006**.

39. O. Shimomura, *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 5590–5602.

40. J.-D. Pédelacq, S. Cabantous, T. Tran, T. C. Terwilliger, G. S. Waldo, *Nat. Biotechnol.* **2005**, *24*, 79–88.

41. R. Das, D. Baker, *Annu. Rev. Biochem.* **2008**, *77*, 363–382.

42. R. L. Dunbrack, *Curr. Opin. Struct. Biol.* **2002**, *12*, 431–440.

43. N. Soundrarajan, S. Sokalingam, G. Raghunathan, N. Budisa, H.-J. Paik, T. H. Yoo, S.-G. Lee, D. Jones, *PLoS ONE* **2012**, *7*, e46741.

44. L. Merkel, Y. Cheburkin, B. Wiltschi, N. Budisa, *ChemBioChem* **2007**, *8*, 2227–2232.

45. B. Wiltschi, L. Merkel, N. Budisa, *ChemBioChem* **2009**, *10*, 217–220.

46. J. C. Carrington, W. G. Dougherty, *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 3391–3395.

47. R. Ribeiro-Viana, M. Sánchez-Navarro, J. Luczkowiak, J. R. Koeppe, R. Delgado, J. Rojo, B. G. Davis, *Nat. Commun.* **2012**, *3*, 1303.

48. V. Hong, S. I. Presolski, C. Ma, M. G. Finn, *Angew. Chem. Int. Ed.* **2009**, *48*, 9879–9883.

49. P.-Y. Liu, N. Jiang, J. Zhang, X. Wei, H.-H. Lin, X.-Q. Yu, *Chem. Biodivers.* **2006**, *3*, 958–966.

50. S C. Fry, *Biochem. J.* **1998**, *332*, 507–515.

51. A. Kumar, K. Li, C. Cai, *Chem. Commun.* **2011**, *47*, 3186–3188.

52. E. Lallana, R. Riguera, E. Fernandez-Megia, *Angew. Chem. Int. Ed.* **2011**, *50*, 8794–8804.

53. C. Besanceney-Webler, H. Jiang, T. Zheng, L. Feng, D. Soriano del Amo, W. Wang, L. M. Klivansky, F. L. Marlow, Y. Liu, P. Wu, *Angew. Chem. Int. Ed.* **2011**, *50*, 8051–8056.

54. M. K. Patel, B. Vijayakrishnan, J. R. Koeppe, J. M. Chalker, K. J. Doores, B. G. Davis, *Chem. Commun.* **2010**, *46*, 9119–9121.

55. C. P. Stowell, R. T. Lee, Y. C. Lee, *Biochemistry* **1980**, *19*, 4904–4908.

56. S. S. Kaufman, P. L. Blain, J. H. Park, D. J. Tuma, *Am. J. hys.* **1990**, *259*, G639–G645.

57. O. Khorev, D. Stokmaier, O. Schwardt, B. Cutting, B. Ernst, *Bioorg. Med. Chem.* **2008**, *16*, 5216–5231.

58. U. Westerlind, J. Westman, E. Törnquist, Smith, C. I. Edvard, S. Oscarson, M. Lahmann, T. Norberg, *Glycoconjugate J.* **2004**, *21*, 227–241.

5. Syntheses of Sialic Acid Derivatives for Bioconjugation

5.1. Contributions

L. M. Artner, N. Bohlke, C. Sieben, A. Herrmann, N. Budisa, and C. P. R. Hackenberger designed the experiments and the research project. N. Bohlke cloned, expressed, purified and characterized green fluorescent proteins (GFP). Viral capsids of bacteriophage Qβ were provided by A. K. Udit. K. Märker, W.-O. Luthardt, E. Croset, T. Sauer and M. Menger contributed equally by synthesizing sialyl donors, azides and corresponding precursors during their internships, bachelor's, or master's theses. L. M. Artner synthesized sialyl donors, neuraminic acid azides and the THPTA ligand, and functionalized, purified and characterized modified proteins.

5.2. Introduction

Based on the previous success on multivalent binding protein scaffolds[1], similar approaches were sought for the binding to the trimeric hemagglutinin receptor of influenza. Although it has been shown that influenza A virus and H3N2 preferably bind to αNeu5Ac-(2→6)-β-D-Gal*p* motifs, recent multivalent targeting approaches show that only sialic acid moiety is enough for sufficient binding.[2], [3]

For further investigation of this phenomenon, sialic acid azides were synthesized for the conjugation to different protein scaffolds. One of those is the icosahedral capsid of bacteriophage Qβ, which should serve as an ideal protein based template for a highly mutlivalent scaffold. This self-assembled macromolecule consists of 180 copies of a coat protein that proved to be ideal for conjugation strategies. The flexibility and applicability of this scaffold was demonstrated in many cases, especially by the groups of Finn, Tirrell and Davis.[4]–[6] Modification of the Qβ capsid with bioactive ligands can be established either by 1) targeting surface exposed lysines, through 2) selective pressure induced incorporation of non-canonical amino acids (ncAAs) with a chemoselectively reacting moiety, or by 3) mixed assemblies.
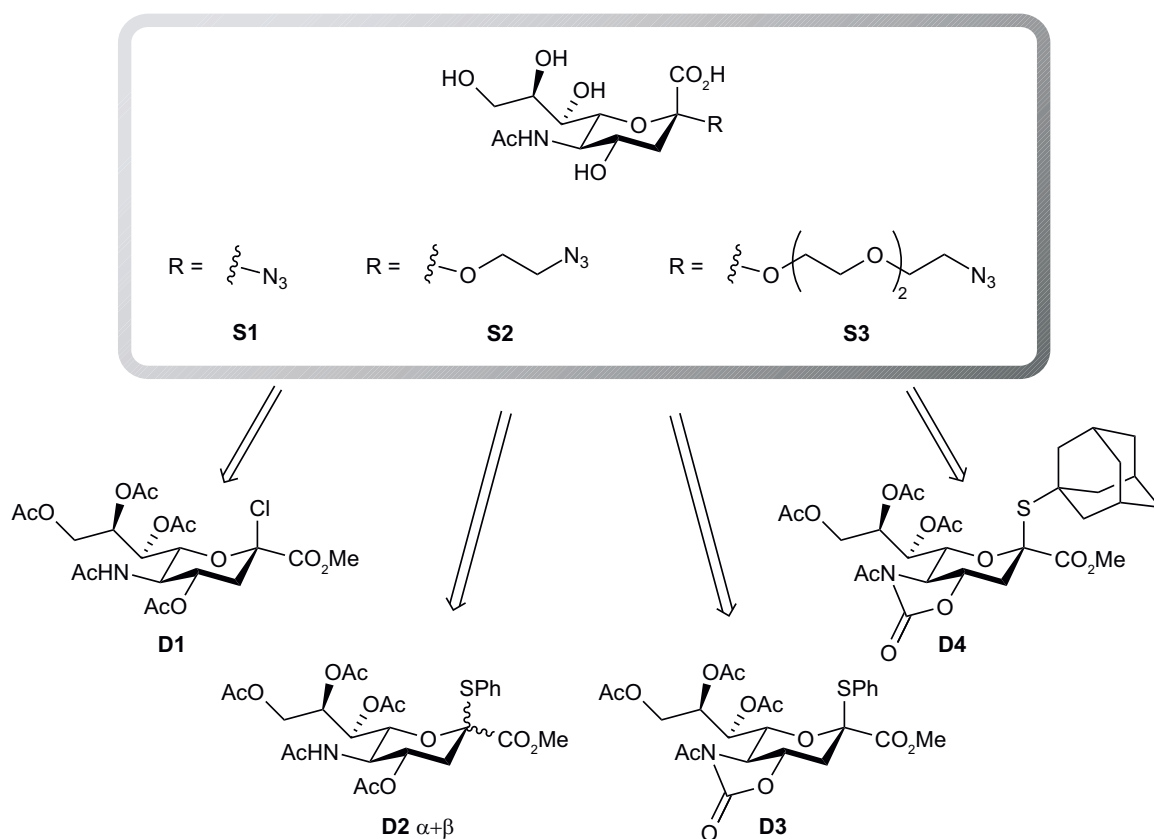
Scheme 5-9: Target compounds and α-selective sialyl donors.

If a high degree of multivalency can be omitted, green fluorescent protein can provide the basis for a functional scaffold. The possibility to incorporate ncAAs for chemoselective conjugation strategies (*vide supra,* **1.6**, pp. 37ff., **4**, pp. 74ff.), its high stability and its intrinsic fluorescence all provide good features for a hemagglutinin targeting probe. Based on the initial efforts in the group of James C. Paulson, αNeu5Ac-(2→6)-β-D-Gal*p*-(1→4)-β-D-Glc*p* (sialyl lactose) and αNeu5Ac-(2→6)-β-D-Gal*p*-(1→4)-β-D-GlcNAc*p* (sialyl lactosamine), both bearing a 2-azidoethyl moiety on the anomeric center of the reductive end, were chemoenzymatically synthesized as before. This ligands have a higher specificity and affinity to hemagglutinin in comparison to Neu5Ac and should therefore be conjugated to a bivalent green fluorescent protein.

The major focus of the following study relates to the stereoselective synthesis of sialic acid azides for the follow-up conjugation to protein-based scaffolds.

5.3. Results & discussion

Various strategies for the generation of hemagglutinin binding scaffolds were followed. Central to two of those were the stereoselective synthesis of αNeu5Ac-conjugates.
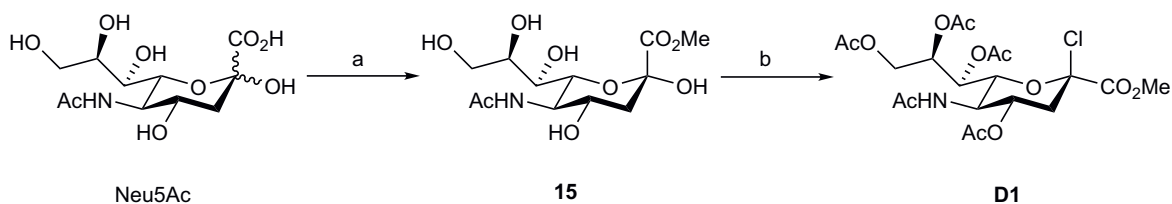
Scheme 5-10: Synthesis of sialyl donor D1.
a) DOWEX, MeOH, r.t., b) AcOCl, cat. MeOH, r.t.

Basically three azido-derivatives (**Scheme 5-9**, p. 94) should be obtained for functionalization of either GFP or Qβ capsid. To realize this, various glycosyl donor systems were synthesized and evaluated based on the outcome of the glycosidation reaction. Although the glycosyl acceptors used herein seemed rather trivial compared to protected mono- and oligosaccharides, the synthesis was challenging and demanding. As it was stated before by Prof. David Crich , the organic synthesis of α-anomeric sialic acid derivatives is coined "challenging" due to lot of reasons. Interestingly, the establishment of a versatile route towards azide analogues of Neu5Ac followed the historical development of the different donor systems. After the evaluation of several of these literature known sialyl donors, Crich's thioadamanatyl sialoside proved to be the most feasible, yielding **S2** and **S3** in high α-selectivity and good yields.

5.3.1. Syntheses of sialic acid derivatives

For the synthesis of a *N*-acetyl neuraminic acid (Neu5Ac) derivative with an azide directly bound to the anomeric center (**S1**), a previously established synthetic route was applied. For this the corresponding methyl ester (**15**) was obtained by esterification of Neu5Ac in methanol, catalyzed by IR-120 (H⁺) ion exchange resin.[8] Subsequently chloride **D1** was obtained in a straightforward way by reaction of acetyl chloride with **15** (**Scheme 5-10**). The yields of this reaction where almost quantitative, especially when AcOCl was distilled prior to use, and the intermediate product was used without further purification. Nucleophilic substitution of the chloride to the azide proceeded smoothly and by-products from previous steps can easily be removed by column chromatography after the reaction. Subsequent Zemplén deacetylation and saponification with sodium hydroxide provided sialic acid azide **S1** in acceptable over yields (**Scheme 5-11**, p. 96).[8]

Unfortunately glycosylation reactions towards **S2** and **S3** did not go as smoothly. Early and recent literature references suggested Koenigs-Knorr type reactions.[9] According to examples with other simple glycosyl acceptors, namely alcohols, this seemed reasonable. However, considerable amounts of a *glycal* and anomeric mixtures of the acetylated precursor **S2Ac** were observed. When silver salicylate, a heterogenous
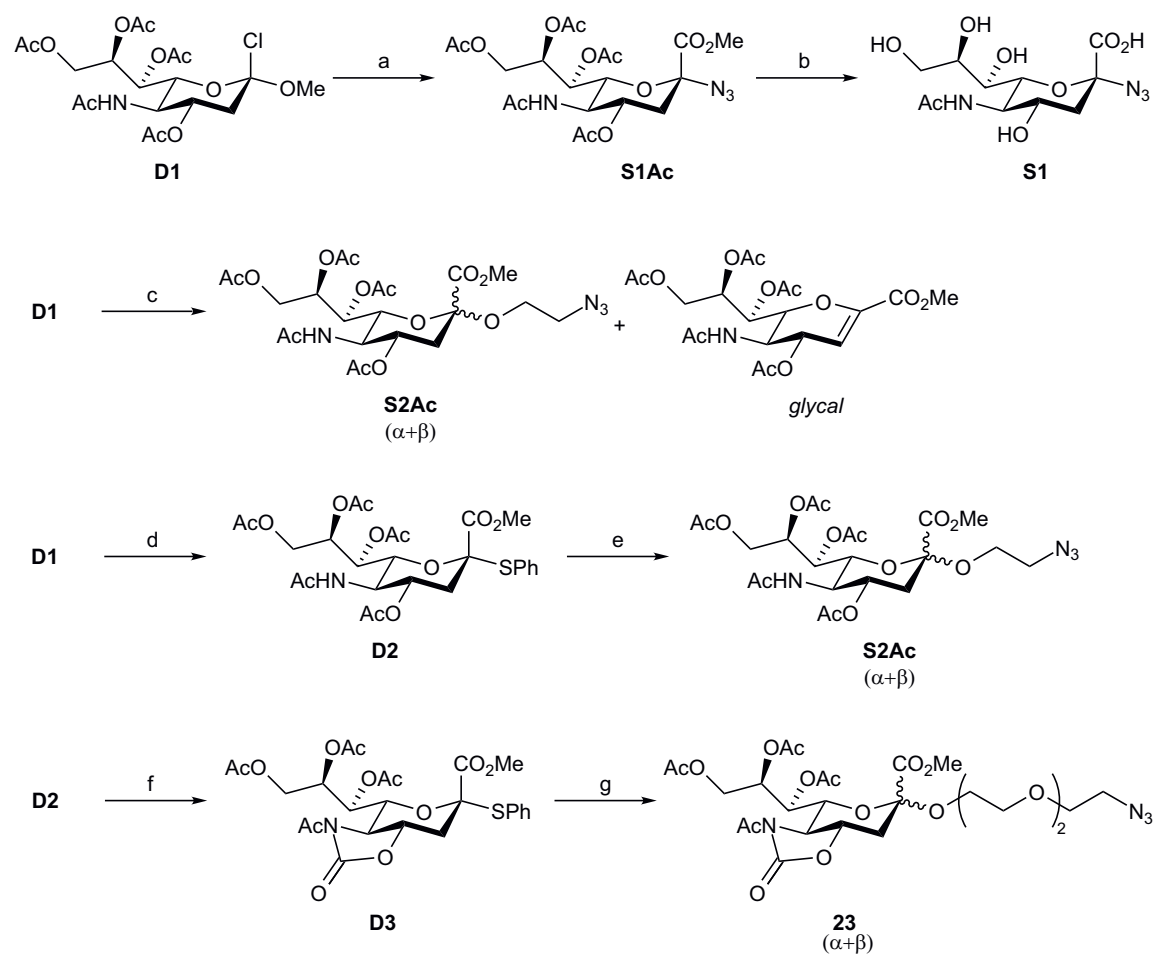
Scheme 5-11: Syntheses from sialyl donors D1, D2, and D3.

a) NaN₃, r.t.; b) NaOMe, MeOH, then NaOH/H₂O; c) N₃EtOH, AgOTf, MeCN, 0 °C; d) PhSH, TBAHS, 1 M K₂CO₃/EtOAc, r.t; e) NIS/TfOH, N₃EtOH, CH₂Cl₂/THF, 40 °C → r.t.; f) MsOH, MeOH, rf.; 4-O₂N-C₆H₄OCl, NaHCO₃, H₂O/MeCN, 0 °C; Ac₂O, pyridine, 0 °C → r.t.; AcOCl, DIPEA, CH₂Cl₂, 0 °C; g) NIS/TfOH, CH₂Cl₂, –40 °C.

silver salt, was applied as a promoter for this Koenigs-Knorr glycosylation, sialylated salicylate was formed as the major product.[10] Apart from the bad stereoselectivity, it was the formation of the *glycal* that seriously limited the Koenigs-Knorr methodology for the sialylation of 2-azidoethanol and 2-[2-(2-azidoethoxy)ethoxy]ethanol (**Scheme 5-11**). Even different promoter did not solve this problem.[2] Thus thiosialosides were chosen as another donor type.[11]–[17]

Thiophenol sialoside **D2** was prepared from chloride **D1** following an improved protocol. Judged from ¹H-NMR, the α-anomer was the main product formed.[12], [13] Subsequently 2-azidoethanol was sialylated with **D2** under *N*-iodosuccinimide/ triflic acid activation. As suggested from the literature, the reaction was carried out in acetonitrile and CH₂Cl₂/THF (5:1), respectively, to test the influence of the solvent on the reaction outcome. Although both systems were reported to promote α-selectivity,
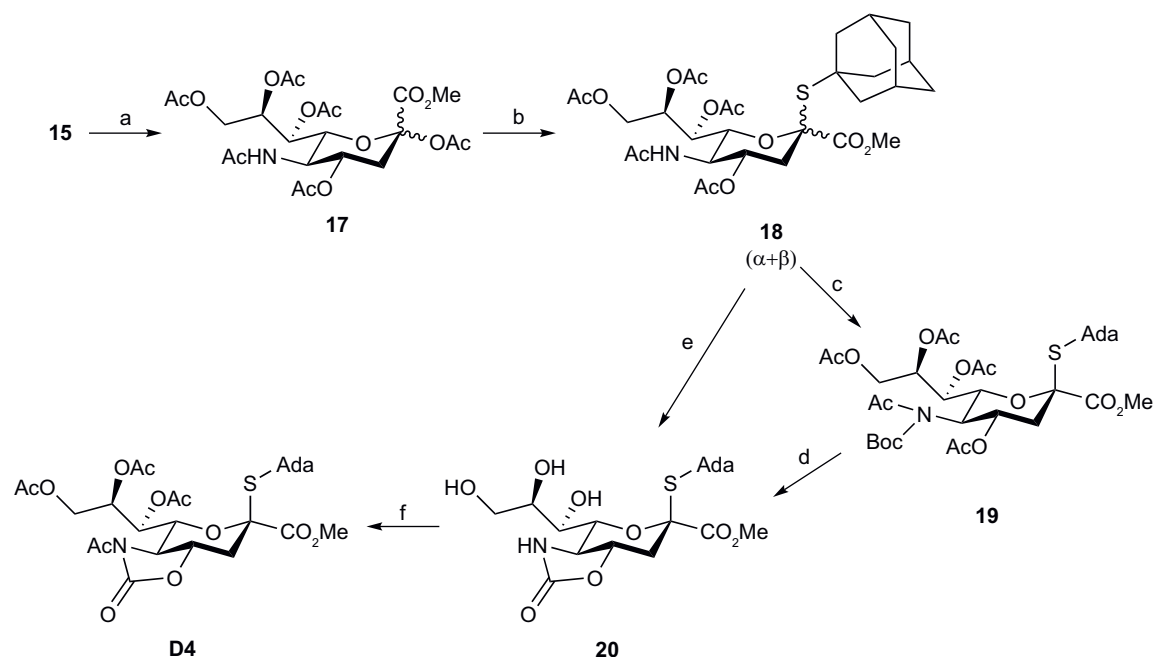
Scheme 5-12: Syntheses from sialyl donor D4.

a) Ac$_2$O, pyridine, 0 °C → r.t.; b) Ada-SH, BF$_3$·OEt$_2$, CH$_2$Cl$_2$, 0 °C → r.t.; c) Boc$_2$O, DMAP, THF, 60 °C; d) NaOMe, MeOH, r.t., then TFA; d) 4-O$_2$N-C$_6$H$_4$OCl, NaHCO$_3$, H$_2$O/MeCN, 0 °C; e) MsOH, MeOH, rf.; 4-O$_2$N-C$_6$H$_4$OCl, NaHCO$_3$, H$_2$O/MeCN, 0 °C; f) Ac$_2$O, pyridine, 0 °C → r.t., then AcOCl, DIPEA, CH$_2$Cl$_2$, 0 °C.

only anomeric mixtures could be obtained from this sialylations.

Due to the low selectivities given by donor **D3** improvements were made. The groups of De Meo and Crich reported sialyl donors with increased α-selectivities when the hydroxyl group at C-4 and the amino group at C-5 are fused by an 4-*O*, 5-*N*-oxazolidinone or an 4-*O*, 5-*N*-acetyloxazolidinone.[14]–[17] Based on their observations, **D3** was synthesized first to evaluate the reported improvements. For this, **D2** was completely deprotected with methanesulfonic acid and subsequently the 4-*O*, 5-*N*-oxazolidinone was formed by addition of 4-nitro chloroformate. Acetylation of the amide finally gave **D3** (**Scheme 5-11**, p. 96). Analogously, a similar procedure[16] was applied for the synthesis of **D4**. Starting from unprotected Neu5Ac, peracetylation after esterification afforded sialoside **17** in an anomeric mixture. Subsequently glycosylation with 1-adamantanethiol under BF$_3$·OEt$_2$ promotion afforded thioglycoside **18**. The β-anomer was completely deprotected to install the 4-*O*, 5-*N*-oxazolidinone protecting group. This was either done directly from the acetate **18** by methanesulfonic acid, or subsequently via deprotection of the Boc-protected sialoside **19**. Total acetylation of the compound resulted in 4-*O*, 5-*N*-acetyloxazolidinone **D4** (**Scheme 5-12**)

Compared to donor **D3**, the thiosialoside **D4** could be activated under lower temperatures with *N*-iodosuccinimide and triflic acid. When donor **D4** was reacted
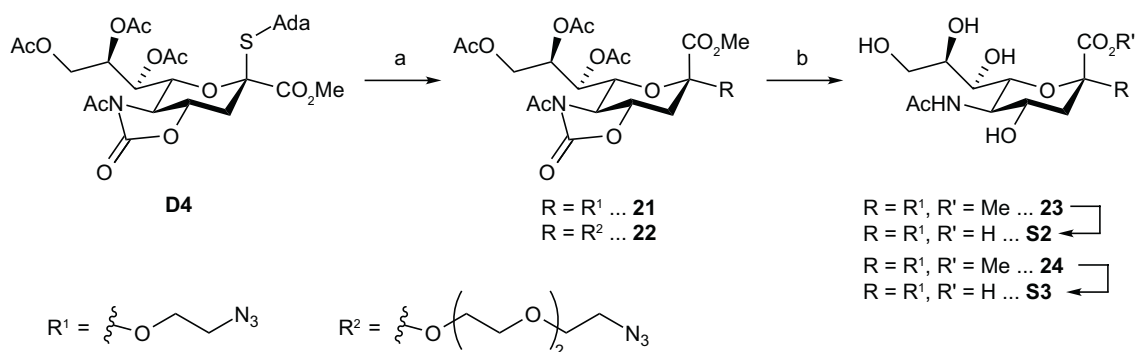
Scheme 5-13: Synthesis of sialic acid azides S2 and S3.
a) N₃EtOH or N₃(EtO)₂EtOH, NIS/TfOH, CH₂Cl₂/MeCN, –78 °C; b) NaOMe, MeOH, r.t.; c) LiOH, MeOH/H₂O, r.t.

with either 2-azidoethanol or 2-[2-(2-azidoethoxy)ethoxy]ethanol, both precursor sialosides **21** and **22** could be obtained in good yields and excellent stereoselectivity. To furnish the unprotected derivatives **S2** and **S3**, a two-step deprotection protocol was applied. The acetyl groups were removed first by applying Zemplén conditions,[19] followed by saponification with lithium hydroxide to remove the methyl ester (**Scheme 5-13**).

5.3.2. Preliminary studies on modification of bacteriophage Qβ capsids

Assembled viral capsids of bacteriophage Qβ were obtained in different variants. A. K. Udit, Occidental College, Los Angeles, provided mutants bearing only a single surface exposed lysine, LysKO2, the precursor for selecective pressure incorporation, K16M, and a pseudo wildtyp variant. In addition, mixed viral capsids were prepared, containing a statistical distribution of the Matsubara-peptide with an peptidic linker in between. Since the polydisperse peptide-bearing capsids do not need additional functionalization, they could be directly applied to hemagglutination and binding assays, but did not show specificity (*data not shown*).

For the LysKO2 variant a protocol was envisioned, following a strategy applied in the group of M. G. Finn.[20] Therefore a simple pentynoic acid succinimide ester was synthesized and subsequently reacted with LysKO2. Although the capsids could subsequently be purified in a straightforward way by size-exclusion and ultra-diafiltration with vivaspin-tubes, analysis by MALDI-TOF mass spectrometry was tedious. No reliable denaturation protocol could be established, leaving acquired data of modified capsids in question. As reported by the Davis lab,[6] ESI-TOF was tried as an alternative, but total entropy deconvolution of spectra only showed unmodified

capsids. The same was true when lactosyl azide was incubated with the supposingly alkynylated viral capsid under CuAAC conditions.

5.3.3. Preliminary studies on sialidation of green fluorescent protein

A green fluorescent protein (GFP) based on the GFPhs1-RM was mutated to bear two homopropargylglycines (Hpg) in the loop regions with an approximated distance of 23Å. Compared to the crystal structure of a hemagglutinin trimer the space between the binding site and the pseudo-binding site could be bridged with such a distance. Chemoenzymatic sialylation of lactosyl and *N*-acetyl lactosamine azides was performed in the Paulson lab (Dr. F. Pfrengle) and furnished αNeu5Ac-(2→6)-β-ᴅ-Galp-(1→4)-β-ᴅ-Glc*p* (**SL**) and αNeu5Ac-(2→6)-β-ᴅ-Galp-(1→4)-β-ᴅ-GlcNAc*p* ethyl azides (**SNL**).

To evaluate the most optimal conditions for the modification of GFP by CuAAC, previously synthesized lactose ethyl azide **5** [1] was used as model substrate. Applying a protocol modified from Finn *et al.*[20] provided a mixture of singly and doubly modified GFP with the majority being on the double side (*data not shown*). Unfortunately, using this optimized procedure did not result in complete conjugation of **SL** or **SNL** to GFP according to MALDI-TOF mass spectrometry (*data not shown*).

5.4. Conclusions

After the initial approach to obtain sialoside **S2** by Koenigs-Knorr type glycosylation failed, more developed sialyl donor systems were investigated. According to the literature, mainly publications by Crich *et al.*, thiophenols and thioadamantanyls were used as leaving groups for sialylation reactions. Especially the application of 4-*O*, 5-*N*-acetyloxazolidinone to yield donor **D4** proved advantageous for follow-up sialidations. Compared to previously synthesized **D3**, its selectivity for the α-anomer was better. In addition to the increased stereoselectivity, higher yields reflected the effectiveness of sialoside **D4**. Zemplén deacetylation and saponification of sialidation products finally afforded α-anomers **S2** and **S3** in acceptable overall yields.

Modification procedures to furnish fully functionalized multivalent biomacromolecules were applied to both GFP and Qβ scaffolds. Unfortunately product formation could not be detected. The reasons for this are currently investigated, but it is proposed that both, the CuAAC reaction and the mass spectrometric detection, has yet to be fully optimized for sialic acid bearing conjugates.

5.5. References

1. L. M. Artner, L. Merkel, N. Bohlke, F. Beceren-Braun, C. Weise, J. Dernedde, N. Budisa, Hackenberger, C. P. R., *Chem. Commun.* **2011**, *48*, 522–524.

2. I. Papp, C. Sieben, A. L. Sisson, J. Kostka, C. Böttcher, K. Ludwig, A. Herrmann, R. Haag, *ChemBioChem* **2011**, *12*, 887–895.

3. I. Papp, C. Sieben, K. Ludwig, M. Roskamp, C. Böttcher, S. Schlecht, A. Herrmann, R. Haag, *Small* **2010**, *6*, 2900–2906.

4. E. Strable, D. E. Prasuhn, A. K. Udit, S. Brown, A. J. Link, J. T. Ngo, G. Lander, J. Quispe, C. S. Potter, B.Carragher et al., *Bioconjugate Chem.* **2008**, *19*, 866–875.

5. J.-K. Rhee, M. Hovlid, J. D. Fiedler, S. D. Brown, F. Manzenrieder, H. Kitagishi, C. Nycholat, J. C. Paulson, M. G. Finn, *Biomacromolecules* **2011**, *12*, 3977–3981.

6. R. Ribeiro-Viana, M. Sánchez-Navarro, J. Luczkowiak, J. R. Koeppe, R. Delgado, J. Rojo, B. G. Davis, *Nat. Commun.* **2012**, *3*, 1303.

7. D. Crich, *J. Org. Chem.* **2011**, *76*, 9193–9209.

8. R. Roy, C.A. Laferriere, *Can. J. Chem.* **1990**, *68*, 2045-2054.

9. R. Šardzík, G. T. Noble, M. J. Weissenborn, A. Martin, S. J. Webb, S. L. Flitsch, *Beilstein J. Org. Chem.* **2010**, *6*, 699–703.

10. K. Märker, Bachelor thesis, Freie Universität Berlin, Berlin, **2013**.

11. F. González Núñez, M. T. E. Campos Valdes, E. Aruca, R. R. Schmidt, V. Verez Bencomo, *J. Carbohydr. Chem.* **2003**, *22*, 395–406.

12. A. Marra, P. Sinaÿ, *Carbohydr. Res.* **1989**, *187*, 35–42.

13. A. V. Orlova, A. M. Shpirt, N. Y. Kulikova, L. O. Kononov, *Carbohydr. Res.* **2010**, *345*, 721–730.

14. C. de Meo, U. Priyadarshani, *Carbohydr. Res.* **2008**, *343*, 1540–1552.

15. D. Crich, W. Li, *J. Org. Chem.* **2007**, *72*, 2387–2391.

16. D. Crich, W. Li, *J. Org. Chem.* **2007**, *72*, 7794–7797.

17. D. Crich, B. Wu, *Tetrahedron* **2008**, *64*, 2042–2047.

18. P. Fügedi, P. Garegg, H. Lönn, T. Norberg, *Glycoconjugate J.* **1987**, *4*, 97–108.

19. G. Zemplén, E. Pacsu, *Ber. dtsch. Chem. Ges. A/B* **1929**, *62*, 1613–1614.

20. V. Hong, S. I. Presolski, C. Ma, M. G. Finn, *Angew. Chem. Int. Ed.* **2009**, *48*, 9879–9883.

6. Surface Plasmon Resonance of Multivalent Binding Carbohydrate-Polyglycerol-Conjugates[†]

6.1. Contributions

V. Böhrsch, T. Mathew, M. R. J. Vallée, J. Dernedde, R. Haag and C. P. R. Hackenberger designed the experiments and the research project. M. R. J. Vallée synthesized alkynyl-phosphonites, performed CuAAC and applied Staudinger phosphonite reactions on polymer scaffolds. M. Zieringer and R. Haag provided polyglycerols for functionalization. T. Mathew synthesized carbohydrate phosphites. V. Böhrsch carried out the Staudinger phosphite reaction and deprotection on polyglycerols. L. M. Artner synthesized carbohydrate azides and performed surface plasmon resonance on functionalized and unreacted polyglycerols.

6.2. Introduction

Carbohydrates are essential molecules of life that mediate various biological interactions (*vide supra*). Since the (bio-)synthesis of glycoproteins and neoglycoproteins has the drawback of either inhomogeneous functionalization, low yields, or both, artificial glycoconjugates are often used to study multivalent carbohydrate-mediated interactions and phenomena. Typically synthetic dendrimers or polymers are used as scaffolds.

†   *The majority of this chapter was already published in:*

V. Böhrsch, T. Mathew, M. Zieringer, M. R. J. Vallée, L. M. Artner, J. Dernedde, R. Haag, C. P. R. Hackenberger, *Org. Biomol. Chem.* **2012**, *10*, 6211–6216; DOI: 10.1039/c2ob25207d

M. R. J. Vallée, L. M. Artner, J. Dernedde, C. P. R. Hackenberger, *Angew. Chem. Int. Ed.* **2013**, *52*, 9504–9508; DOI: 10.1002/anie.201302462

M. R. J. Vallée, L. M. Artner, J. Dernedde, C. P. R. Hackenberger, *Angew. Chem.* **2013**, *125*, 9682–9686; DOI: 10.1002/ange.201302462
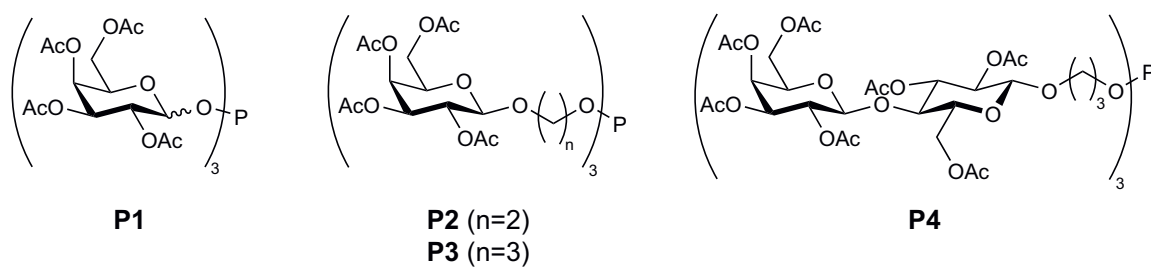
Scheme 6-14: Glycosyl phosphites for functionalization of azido-polymers by Staudinger phosphite reaction.
Acetyl protected carbohydrate phosphites were synthesized by Dr. Mathew, using Koenigs-Knorr glycosylation (see reference [1] for more details).

They can be obtained in good yields with a varying degree of dispersity and different chemical backbones, ranging from acrylamides to glycerols. The most obvious reasons for the use of polymers is their ease of accessibility, accompanied by the variation of given features and the chemistry that can applied to them.

In this chapter, two chemoselective functionalization strategies[1], [2] based on the Staudinger reaction were investigated and their applicability evaluated through binding assays measured by surface plasmon resonance. The group of Hackenberger recently demonstrated the applicability of Staudinger-type reactions for various biological systems.[3]–[9]

Based on previously successful applications of carbohydrate-functionalized polyglycerols, newly developed strategies' were applied to functionalize azide-bearing polyglycerols. In contrast to earlier published work by Haag *et al.*,[10]–[12] herein copper-free chemoselective conjugation strategies were applied. Although both, the Staudinger phosphite (SPtR) as well as the Staudinger phosphonite reaction (SPnR), were used before for chemoselective functionalization, the approaches herein highlight the applicability of the reaction as a tool for chemoselective functionalization of polymers.

6.3. Results

Azido-polyglycerols of different molecular weight and with varying degree of functionalization built the foundation of the glycopolymers. The polyglycerol core was subsequently modified with either symmetrical carbohydrate phosphites on one hand or lactose triazolyl phosphonites on the other hand. Since the last conjugation step is performed copper-free, dialysis of the resulting product provides glycopolymers in high purity, suitable for surface plasmon resonance (SPR).
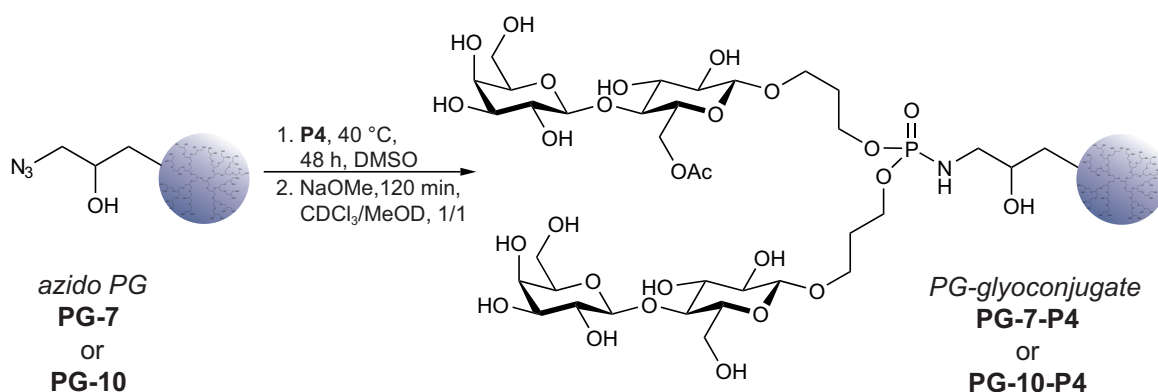
Scheme 6-15: Functionalization of azido-polyglycerols by Staudinger phospite reaction.
Azide-bearing polyglycerols **PG-7** and **PG-10** were functionalized with lactosyl phosphite **P4** *via* the Staudinger phosphite reaction. After deacetylation, PG-glycoconjugates **PG-7-P4** and **PG10-P4** were obtained. **PG-7**: core 7.7 kDa, DF = 30%; **PG-10**: 10.6 kDa, DF = 32%; **PG-7-P4**: core 7.7 kDa, DF = 14%; **PG-10-P4**: 10.6 kDa, DF = 19%; degree of functionalization (DF) based on ^{1}H-NMR data. Adapted from reference [1] with permission of The Royal Society of Chemistry.

6.3.1. Functionalization of polyglycerols by Staudinger phosphite reaction

(Exclusively performed by Dr. Verena Böhrsch, Dr. Thresen Mathew, and Dr. Maximilian Zieringer, all Freie Universität Berlin)

Various symmetrical galactosyl and lactosyl phosphites have been synthesized by Dr. T. Mathew (**Scheme 6-14**, p. 102) applying Koenigs-Knorr type glycosylation, followed by reaction with PCl_3 and a non-nucleophilic base. Subsequently Dr. V. Böhrsch performed SPtR of those glycosyl phosphites to functionalize polyglycerols, kindly provided by Dr. M. Zieringer of the group of Prof. R. Haag. Deprotection of the newly-generated glyco-polymers under Zemplén conditions[13] resulted in approximately 2 mg unprotected polyglycerols each.[1]

To evaluate the applicability of glycosyl phosphites, water-soluble, lectin-binding glyco-polymers were synthesized by Dr. V. Böhrsch, using polyglycerols with a lower degree of functionalization (DF). Therefore **PG-7** (core 7.7 kDa, DF = 30%) and **PG-10** (core 10.6 kDa, DF = 32%) were reacted with lactosyl phosphite **P4**. By applying the same reaction sequence, polyglycerol-glycoconjugates **PG-7-P4** (core 7.7 kDa, total DF = 14%) and **PG-10-P4** (core 10.6 kDa, total DF = 19%) were obtained (**Scheme 6-15**). After purification, the glyco-polymers were subjected to a competitive binding assay with peanut agglutinin (PNA) and Thomsen-Friedenreich (TF) antigen and analyzed by surface plasmon resonance (*vide infra,* **3.3.3**, pp. 66ff.).
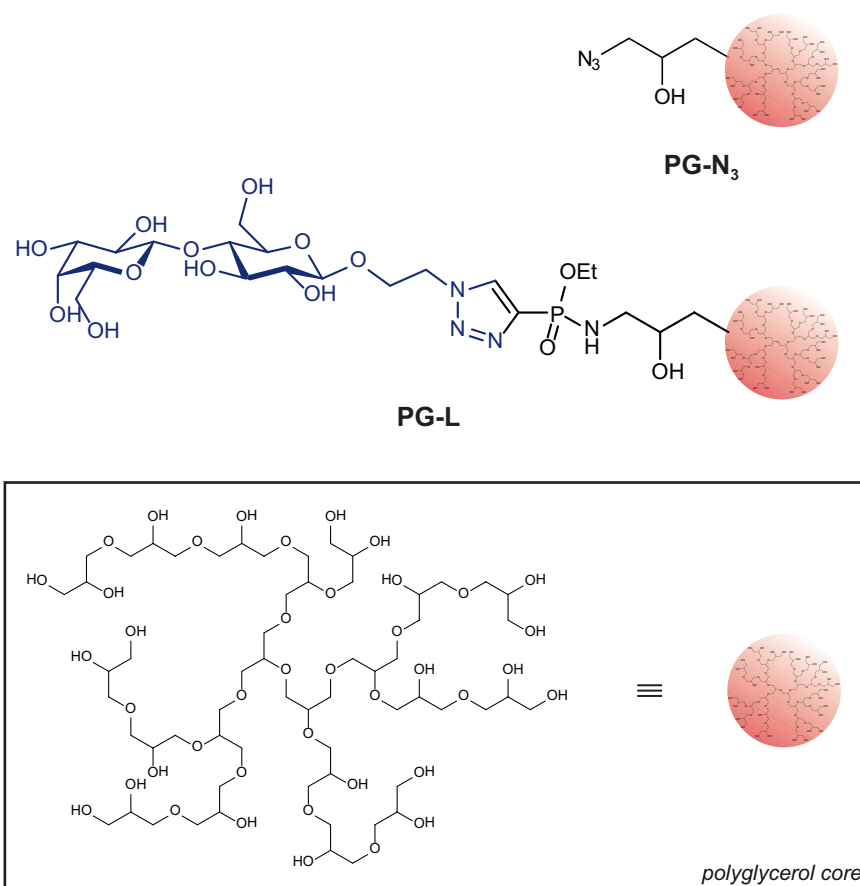
Figure 6-26: Chemical structures of polyglycerol-azide PG-N$_3$ and its functionalized counterpart, PG-L.

PG-N$_3$ was provided by the lab of Prof. Dr. R. Haag[2] and **PG-L** was synthesized by Dr. M. R. J. Valée.[2] Adapted from reference [2] with permission of Wiley-VCH.

6.3.2. Functionalization of polyglycerols by subsequent azide-azide coupling

(Performed by Dr. M. Robert J. Vallée, Freie Universität Berlin)

For the generation of carbohydrate-bearing Staudinger phosphonite reagents, peracetylated lactose was glycosylated with 2-azidoethanol by Koenigs-Knorr type reactions *via* the bromide, or directly by Lewis acid promotion to yield 1-azidoethoxy-(2,3,4,6,-tetra-*O*-acetyl-β-ᴅ-galactopyranosyl)-(1→4)-2,3,6-tri-O-acetyl-β-ᴅ-glucopyranoside (**5Ac**, **3.3.1**, pp. 62ff., **10.2.3.5**, p. 126).[14] Deacetylation under Zemplén conditions afforded the analogous free lactosyl azide (**5**, **3.3.1**, pp. 62ff., **10.2.3.6**, p. 126).[14]

Subsequently Dr. M. R. J. Vallée synthesized phosphonite reagents[2], [3], [14], [15] from lactosyl azides *via* CuAAC. After deprotection of the borane group with 1,4-diazabicyclo[2.2.2] octane (DABCO) in DMSO, the phosphonite was reacted with azide-containing
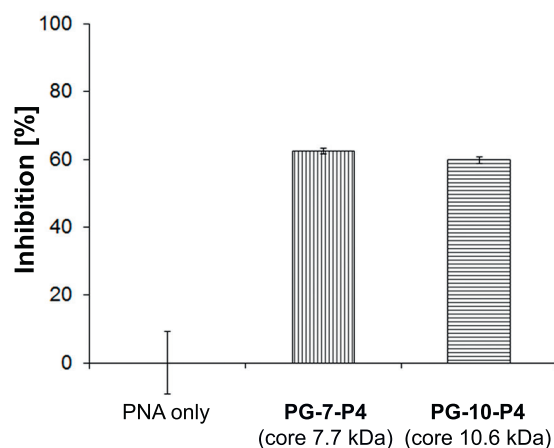
Figure 6-27: SPR-measurements of polylactose(poly)glycerols PG-7-P4 and PG-10-P4 at 100µM.
Previously obtained polyglycerol-glycoconjugates **PG-7-P4** and **PG-10-P4** were subjected to a competitive PNA binding assay. Lactose-functionalized **PG-7-P4** and **PG-10-P4** inhibited binding of PNA to TF-antigen for approximately 62% and 60%, respectively.
Adapted from reference [1] with permission of The Royal Society of Chemistry.

polyglycerols to furnish lactose-functionalized glycopolymer **PG-L** (**Figure 6-26**, p. 104). Finally, dialysis of this polyglycerol resulted in highly pure product (>94%, according to ¹H-NMR) suitable for lectin binding studies.[2], [3], [15]

6.3.3. Surface plasmon resonance of functionalized polyglycerols

For the analysis of obtained carbohydrate-polymer conjugates, two different binding assays were tested to determine the efficiency of binding to a lectin. For this purpose, peanut agglutinin (PNA), a soluble plant lectin that binds to terminal galactose moieties was used as a model protein. This well-characterized and studied lectin was also used before for multivalent scaffolds within the Hackenberger group.[14] Therefore solid experience with the assays and comparability was known.

For the preparation of th competitive binding assay, the chip was functionalized with a well-binding ligand of PNA, Thomsen-Friedenreich (TF) antigen. This disaccharide was bound to a polyacrylamide-based polymeric support, equipped with biotin. The sensor chip was already functionalized with dextran and streptavidin. Incubation of the ligand with the chip resulted in biotin-streptavidin-mediated[16] conjugation.

For the assay itself, the glyco-polymers prepared by Staudinger phosphite reaction were incubated with PNA for 20 minutes and subsequently submitted to SPR. Signal intensities were compared to incubation of PNA in sole buffer to determine the efficiency of inhibition (**Figure 6-27**). The analysis of both lactosyl glycopolymers resulted in a
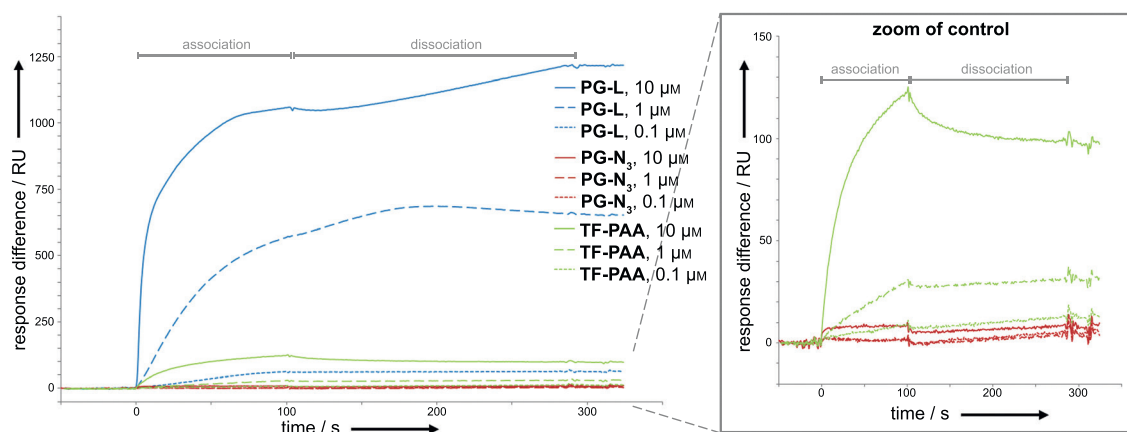
Figure 6-28:   Processed sensogram of the direct binding assay.

Direct binding assay of peanut agglutinin (PNA) and lactose-functionalized **PG-L**, **TF-PAA**, 30% **PG-N₃** (left) as well as **TF-PAA** and 30% **PG-N₃** (control experiments, right). Final RU binding signals (response difference/RU) were obtained by subtraction of the RU values of the reference lane from the ligand lane.

Adapted from reference [2] with permission from Wiley-VCH.

reduction of binding intensities to PNA of roughly 60% at 100 µM.

Glycopolymers synthesized by Staudinger phosphonite reaction were evaluated differently. To functionalize the surface of the sensor chip to be suitable for direct binding measurements, the carboxymethylated dextran-coated chip (CM5) was activated with EDC and NHS. Subsequently PNA was coupled to the surface and unreacted activated esters were quenched with 2-aminoethanol. Direct binding measurements were then performed in three different concentrations of 0.1 µM, 1 µM, and 10 µM with the functionalized glycopolymer (**PG-L**, an unreacted azido-polymer (**PG-N₃**), and a polyacrylamide-bound TF-antigen (**TF-PAA**). Nearly no response could be observed in the case of **PG-N₃** at any concentrations, whereas **TF-PAA** and **PG-L** both interacted with the surface bound PNA (**Figure 6-28**). Interestingly, the response of **PG-L** at a concentration of 10 µM was approximately nine times higher than the signal observed by **TF-PAA** at the same concentration.

6.4.   Discussion

Comparing the strategies applied for both Staudinger phosphite and phosphonite "glycosylation", they are somewhat different from their methodology. Both approaches have advantages of their own. As can be seen from lactosyl phosphite **P4**, a low degree of functionalization of the polyglycerol still resulted in good inhibition of PNA-TF-binding with approximately 60% at a concentration of 100 µM. This phenomenon
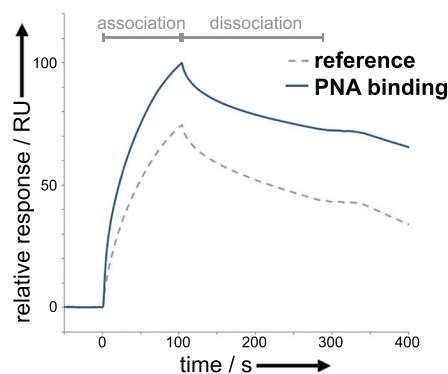
Figure 6-29: Processed, relative sensogram of the individual flow chambers.
Separate RU binding signals (relative response/RU) of lactose-functionalized **PG-L** for the ligand lane containing immobilized PNA (solid line) and the reference lane (dotted line).
Adapted from reference [2] with permission of Wiley-VCH.

is explained with the nature of the Staudinger phosphite reaction of symmetrical phosphites. This transformation effectively yields two carbohydrate moieties directly from the reacting entity. Taking into account that it only takes up to three steps from a commercially available starting material to the Staudinger phosphite reagent it is a straightforward alternative to artificial dendrimer synthesis for creating branched glycoconjugates. This also means, that the calculated functionalization can be counted as double the modification with other chemoselective approaches like CuAAC. Since further addition did not increase the degree of functionalization[1], there is a strong argument for steric reasons. This also makes sense when imagining at the structure of an unreacted glycosyl phosphite. Basically it consists of three carbohydrate moieties fused to a central phosphorous atom. The resulting tetrahedron is a large structure with moleculare weights in the kDa-range. Therefore it is reasonable to assume that steric repulsion prevent additional conjugation to an already densely functionalized glycopolymer. This is all in favor for efficient, copper-free modification of polyglycerols and similar polymeric scaffolds. Only a slight drawback is the deacetylation after the reaction, because it requires strong basic conditions and thus hinders applicability on the protein level.

The combination of CuAAC and Staudinger phosphite reaction on the other hand provides a flexible tool to create numerous bioactive ligands for copper-free conjugations in a straightforward way. The fact that copper can easily be removed after attaching the azide moiety to the alkynyl phosphonite at the stage of the borane-protected triazolyl phosphonite underlines the applicability of the approach. Especially for the functionalization of polymers this is advantages, since they tend to chelate copper due to their side-chain moieties, making it difficult to get rid of this metal.

Unfortunately, carbohydrate triazolyl-phosphonites have to stay protected until prior to the reaction, since they suffer from low storage stability. However, deprotection with DABCO in either benzene, toluene or DMSO for approximately 20 hours provide a solid approach to freshly prepared Staudinger phosphonite reagents. Without further purification, those phosphonites can directly be reacted with azide-bearing scaffolds. As demonstrated with azide-containing polyglycerols, this strategy yields highly-functionalized polymeric scaffolds *via* a copper-free reaction that can easily be purified by dialysis.

Comparison of the copper-free glycopolymer **PG-L** with other polymers (**PG-N$_3$** and **TF-PAA**) by a direct-binding SPR assay revealed strong interaction with surface-bound PNA. When looking at the dissociation phase of the measurement (**Figure 6-28**, p. 106), the strong binding can be deducted from the drift of the sensogram to higher RU values. Normally, the sensogram should decrease because of the release of the analyte from the surface. However, when taking a look at the sensograms of each flow-chamber separately, slower signal decay is observed in the functionalized lane compared to the unfunctionalized reference lane (**Figure 6-28**, p. 106, **Figure 6-29**, p. 107). Since final RU values are derived from the difference in the measured RU values of each lane, the graph of the dissociation phase can result in an increase. The results form comparative studies with unmodified **PG-N$_3$** rule out the possibility that this binding phenomenon is due to unspecific interaction of the polyglycerol scaffold (**Figure 6-28**, p. 106). In that case only a very weak response in the sensogram was observed at high concentrations. In addition, the polyacrylamide TF-antigen **TF-PAA** did not show increased RU values in the dissociation phase even at high concentrations (**Figure 6-28**, p. 106). The latter result makes it clear that the contribution of the carbohydrate ligand is also not alone responsible for the phenomenon observed with **PG-L**. This leaves the shape of the scaffold and the linker as contributors for an increased signal during the dissociation phase.

6.5. Conclusion

In summary, two different copper-free functionalization strategies were used to generate multivalent binding glycopolymers. The Staudinger phosphite reaction proves advantageous when the polymer should be decorated with branched saccharides in one step. Since symmetrical carbohydrate phosphites result in two carbohydrate moieties per reacted azide moiety, functionalization presumingly occurs until there is no space left on the surface of the polyglycerol. Therefore multivalent glycopolymers could be generated and their applicability was shown by a competitive lectin binding assay assessed with SPR. An inhibition of 60% for binding between PNA

and surface-bound TF antigen was observed when polyglycerols functionalized with lactosyl phosphite **P4** was applied. Other workgroups already demonstrated the high applicability of carbohydrate-decorated polyglycerols.[10]–[12] However, for syntheses of such neoglycoconjugates usually CuAAC is applied for efficient functionalization. The herein presented approach completely avoids copper. Still, it is suitable to deliver a high degree of functionalization that resulted in comparable inhibition to previously reported systems.[11]

Th other strategy was based on the Staudinger phosphonite reaction. For a polyglycerol that was reacted with lactose triazolyl phosphonite, strong interaction with PNA was measured by SPR. In that case a direct binding assay was applied and comparison to **TF-PAA** and **PG-N$_3$** revealed that glycopolymer **PG-L** still binds well, even in the dissociation phase. Although some potential explanations were ruled out, the influence of scaffold shape and linker design could not be deducted from the experiments presented herein. To fully understand the binding mode of **PG-L** to surface-bound PNA, further studies have to be concluded. Nevertheless, the flexibility provided by having a alkyne moiety to be functionalized by CuAAC prior to a Staudinger reaction with an azide-bearing scaffold proves advantageous. It simplifies synthesis of ligands, due to azides often being chemically more accessible. This holds also true for applications such as combinatorial approaches where a formal azide-azide coupling eases up the screening of ligands. Since deprotection of the borane protecting group can also be carried out in DMSO, functionalization of proteins can be a potential future employment of the strategy presented herein.

6.6. References

1. V. Böhrsch, T. Mathew, M. Zieringer, Vallée, M. Robert J., L. M. Artner, J. Dernedde, R. Haag, C. P. R. Hackenberger, *Org. Biomol. Chem.* **2012**, *10*, 6211–6216, *and references therein.*

2. M. R. J. Vallée, L. M. Artner, J. Dernedde, C. P. R. Hackenberger, *Angew. Chem. Int. Ed.* **2013**, *52*, 9504–9508, *and references therein.*

3. M. R. J. Vallée, P. Majkut, I. Wilkening, C. Weise, G. Müller, C. P. R. Hackenberger, *Org. Lett.* **2011**, *13*, 5440-5443.

4. N. Nischan, A. Chakrabarti, R.A. Serwa, P.H.M. Bovee-Geurts, R. Brock, C.P.R. Hackenberger, *Angew. Chem. Int. Ed.* **2013**, *52*, 11920–11924.

5. I. Wilkening, G. del Signore, C.P.R. Hackenberger, *Chem. Commun.* **2011**, *47*, 349–351.

6. R. Serwa, B. Horstmann, P. Majkut, J.-M. Swiecicki, M. Gerrits, E. Krause, C.P.R. Hackenberger, *Chemical Science* **2010**, 1, 596-602.

7. R. Serwa, J.-M. Swiecicki, D. Homann, C. P. R. Hackenberger, *J. Pep. Sci.* **2010**, 16, 563–567.

8. D. M. M. Jaradat, H. Hamouda, C. P. R. Hackenberger, *Eur. J. Org. Chem.* **2010**, 26, 5004–5009.

9. V. Böhrsch, R. Serwa, P. Majkut, E. Krause, C. P. R. Hackenberger, *Chem. Commun.* **2010**, 46, 3176–3178.

10. I. Papp, J. Dernedde, S. Enders, R. Haag, *Chem. Commun.* **2008**, 5851–5853.

11. I. Papp, J. Dernedde, S. Enders, S. B. Riese, T. C. Shiao, R. Roy, R. Haag, *Chembiochem* **2011**, *12*, 1075–1083.

12. I. Papp, C. Sieben, A. L. Sisson, J. Kostka, C. Böttcher, K. Ludwig, A. Herrmann, R. Haag, *Chembiochem* **2011**, *12*, 887–895.

13. G. Zemplén, E. Pacsu, *Ber. dtsch. Chem. Ges. A/B* **1929**, *62*, 1613–1614.

14. L. M. Artner, L. Merkel, N. Bohlke, F. Beceren-Braun, C. Weise, J. Dernedde, N. Budisa, C. P. R. Hackenberger, *Chem. Commun.* **2011**, *48*, 522–524.

15. M. R. J. Vallée, *PhD thesis*, Freie Universität Berlin, Berlin, **2013**.

16. N. M. Green, *Biochem. J.* **1963**, *89*, 599–609.

7. Outlook

7.1. Rational neoglycoprotein design for ASGP-R

After initial studies with barnase inhibitor protein barstar as a protein-based scaffold, the concept was broadened and intensified by using GFP for multivalent presentation of biologically active carbohydrates. For the target, the endocytotic asialoglycoprotein receptor, different variants of ncAA-bearing GFPs were designed and expressed. The herein presented investigations further supported previously reported specificities of this receptor. Aside from this discovery, an influence of linker on carbohydrate specificity could be observed. Although the exact reasons for this are at this time unknown, it can be argued that a minimal linker length is required to reach the active carbohydrate recognition domain of ASGP-R. Homology modeling and other *in silico* methods may provide an answer to this phenomenon and are under consideration for future experiments.

In addition the degree of multivalency and the importance of spatial presentation of carbohydrate ligands were investigated. In the course of the studies, first experimental evidence for a multivalent binding event for neoglyco-GFP and ASGP-R could be provided by FACS analysis. This findings were reasonable and could be rationalized by previously reported models from the groups of Kitov and Whitesides. Nevertheless, to fully understand the extent and necessity of multivalency in the binding of ASGP-R to glycoproteins, various controls have to be investigated. It has to be addressed if the increased binding of the trivalently modified GFP is an effect due to the rational design or because of increased probability. Further experiments to thoroughly investigate this multivalent binding model are currently performed in a collaborative effort of the groups of B. Lepenies, N. Budisa and C. P. R. Hackenberger.

7.2. Polymer-based glycoconjugates

The combination of carbohydrate ligands and polymers is a well established way to generate highly multivalent, polydisperse glycoconjugates. Within the context of this thesis, new chemical strategies and methods were applied for the generation of

polymer-based glycoconjugates. Both Staudinger phosphite reaction with glycosyl phosphites and azides, and subsequent azide–azide coupling using copper-catalyzed alkyne azide cycloaddition (CuAAC) and the Staudinger phosphonite reaction on azides, provided highly multivalent polyglycerols. The applicability of this conjugates for biological questions and to address the concept of multivalency could be shown by surface plasmon resonance assays. Especially subsequent azide-azide coupling has a vast potential. Due to the CuAAC, different carbohydrate ligands or other bio-active compounds can be efficiently introduced and subsequently reacted in a copper-free way. The fact that the synthesis of both, carbohydrate azides and azido-polymer is usually straight-forward and high-yielding makes this approach ideal for ligand and target screenings.

7.3. Glycoconjugates for the inhibition of hemagglutinin

Several literature known sialyl donors have been investigated for efficient and α-selective glycosylation reactions for N-acetyl neuraminic acid derivatives. Finally, the donor system developed in D. Crich's group – an N,O-oxazolidinone protected S-adamantyl sialoside – proved to be superior to the other systems tested. The synthetic potential of this compound should be exploited much further. Due to the high yields and stereoselectivity that could be achieved with this ligand, one should aim to use it for the synthesis of biologically relevant di- or trisaccharides. Potential future compounds could therefore be sialylated galactoses, lactoses and N-acetyl lactosamines. The mentioned oligo-saccharides are potentially strong binders to influenza's hemagglutinin. If presented in a multivalent fashion, avidity would be increased and detection of hemagglutinin would be facilitated. Either GFP or assembled viral capsids would be a suitable platform to decorate. GFP with its intrinsic fluorescence would be perfect for ELISA-based assays, whereas viral capsids would be able to help in electron microscopy due to their high symmetry. In addition, viral capsids are suitable for the synthesis of monodisperse, highly multivalent, globular glycoconjugates. By introducing glycodendrimers on the surface of viral capsids, an even higher degree of multivalency would be achieved.

8. Abstract

Numerous cellular and biological processes are regulated and mediated by post-translational modifications (PTMs) of proteins, such as glycosylation. The binding events of glycoproteins are often mediated by lectins and typically occur in cooperative or multivalent fashion. In this thesis, different strategies have been applied to study this multivalent binding events.

8.1. Rational design of multivalent binding neoglycoproteins

Initially a pseudo-wildtype variant of *E. coli* barnase inhibitor protein (ψ-b*) containing four alkyne moieties was provided by combining site-directed mutagenesis and selective pressure incorporation (SPI). Subsequently monosaccharide azides were conjugated to the protein by copper-catalyzed alkyne-azide cycloaddition (CuAAC), thus generating different mono- and multivalent neoglycoproteins. The applicability of this approach was further assessed by a competitive lectin-binding surface plasmon resonance (SPR) assay with peanut agglutinin (PNA). Experimental data pointed towards an increased binding due to the multivalent architecture.

As a next step, a rational approach was envisioned for the targeting of asialoglycoprotein receptor (ASGPR), an endocytic lectin. Therefore a green fluorescent protein (GFP) with a defined spatial presentation of carbohydrates was generated. By taking a look at the crystal structure of GFP and the proposed geometry of ASGPR, several positions in the loop region were chosen for non-canonical amino acid (ncAA) incorporation. Subsequently the sites were modified by CuAAC to introduce *N*-acetylgalactosamine and mannose conjugates with different linker length. The resulting neoglycoproteins were incubated with ASGPR expressing cells and analyzed by fluorescence-activated cell sorting (FACS). A series of experiments were conducted with a different neoglyco-GFPs and correlation of linker length and carbohydrate specificity was observed. Competitive binding against GalNAc conjugated to bovine serum albumin, and inhibition of endocytosis with cytochalasin D resulted in reduced binding in both cases, providing further evidence for their specificity. In addition, enhancement of binding to ASGPR could be observed by comparing trivalent with mono- and bivalent neoglyco-GFPs. All observations combined lead to the conclusions, that there is a positive contribution for

binding efficacy when ligands are presented the right way. Those findings support the concept of rationally designed neoglycoproteins as structurally defined scaffolds for the elucidation of multivalent binding events.

8.2. Syntheses of sialic acid derivatives for bio-conjugations

N-Acetyl neuraminic acid (Neu5Ac) is an important carbohydrate and often terminates mammalian glycans. Due to its biological relevance, glycoconjugates of Neu5Ac are desirable. In the context of generating multivalent binding neoglycoproteins, the chemical synthesis of Neu5Ac azide-derivatives was explored. Different sialyl-donors were prepared and tested for their applicability in generating Neu5Ac conjugates α-selectively. After unsatisfactory initial experiments with Koenigs-Knorr donors, thiophenol donors were investigated. Unfortunately the observed α-selectivity was not sufficient for our purpose. Finally the best outcome was achieved by applying the thioadamantane donor of Crich *et al*. With this donor–promoter system sialyl azides could be obtained with very good α-selectivity and high yields of the glycosylation reaction.

However, the bioconjugation to protein-scaffolds could not be realized, probably because of insufficient optimization of CuAAC and mass analysis.

8.3. Surface plasmon resonance of carbohydrate-polymer conjugates

The efficacy the Staudinger phosphite (SPtR) and Staudinger phosphonite reactions (SPnR) as conjugation reactions for the functionalization of azide-bearing polyglycerols was assessed. Previously generated neoglycopolymers were therefore subjected to SPR. Polyglycerols that were functionalized with glycosyl phosphites were applied to a competitive lectin binding assay against PNA. In this study, inhibition of approximately 60% was achieved. For polyglycerols functionalized with lactose triazolyl phosphonites a different assay was conducted. For this, PNA was immobilized on the SPR chip. The binding of functionalized polyglycerols was measured and compared to azido-polyglycerol, as well as polyacrylamide-conjugated Thomsen-Friedenreich antigen. The SPR experiments showed a clear and strong interaction of the neoglyco-polyglycerol with surface-bound PNA. It therefore could be concluded that both conjugation

reactions are suitable as a functionalization strategy for azido-polymers. The fact that those alternative approaches are copper-free, can be seen as an additional advantage regarding the generation of carbohydrate-bearing polyglycerols.

9. Kurzzusammenfassung

Eine Vielzahl zellulärer und biologischer Prozesse werden durch post-translationale Modifikationen (PTMs), wie Glykosylierungen, an Proteinen reguliert und vermittelt Eine der wohl komplexesten PTMs sind Glykosylierungen. Die Bindungsereignisse von Glykoproteinen werden oft durch Lektine vermittelt und sind typischerweise kooperativ oder multivalent. In dieser Doktorarbeit wurden unterschiedliche Strategien angewandt um diese multivalenten Bindungsereignisse zu untersuchen.

9.1. Rationales Design multivalent-bindender Neoglykoproteine

Zu Beginn wurde eine Pseudo-Wildtyp-Variante des Barnase-Inhibitor-Proteins (ψ-b*) von *E. coli* mit vier Alkinen ausgestattet. Dies wurde durch eine Kombination von ortsspezifischen Mutationen und auxotrophem Einbau von unnatürlichen Aminosäuren (suppression based incorporation, SPI) verwirklicht. Anschließend wurden Monosaccharid-Azide an diese Proteine mittels Kupfer-katalysierter Alkin-Azid Cycloaddition (copper-catalyzed alkyne-azide cycloaddition, CuAAC) konjugiert. Somit wurden verschiedene mono- und multivalente Neoglykoproteine hergestellt. Diese künstlich modifizierten Proteine wurden anschließend anhand einer kompetitiven Lektin-Bindungsstudie mittels Oberflächenplasmonenresonanz (surface plasmon resonance, SPR) untersucht. Hierfür wurde Erdnuss-Agglutinin (peanut agglutinin, PNA) verwendet, da es für seine hohe Spezifität für Galaktose-terminierte Saccharide bekannt ist. Die experimentell erhaltenen Daten deuteten schließlich auf erhöhte Bindung an PNA aufgrund der multivalenten Architektur hin.

Im nächsten Schritt sollte ein rationaler Ansatz für die Bindung an Asialoglykoprotein Rezeptor (ASGPR), ein endozytierendes Lektin, entwickelt werden. Dafür wurde ein grün-fluoreszierendes Protein (GFP) mit einer definierten räumlichen Ausrichtung der Zucker hergestellt. Durch einen Vergleich der Kristallstruktur von GFP mit der vermuteten Geometrie des ASGPR konnten mehrere Positionen in den Schleifenregionen des GFP für den Einbau unnatürlicher Aminosäuren (non-canonical amino acid, ncAA) ausgemacht werden. Die ausgewählten Stellen wurden anschließend mit CuAAC modifiziert um *N*-Acetylgalaktosamin und Mannose Konjugate mit unterschiedlicher Länge an Verbindungselementen zu präsentieren. Die somit erhaltenen Neoglyko-

proteine wurden zusammen mit ASGP-R-exprimierenden Zellen inkubiert und mittels fluoreszenzbasierter Durchflusszytometrie (fluorescence-activated cell sorting, FACS) analysiert. Eine Reihe von Experimenten mit unterschiedlichen Neoglyko-GFP wurde durchgeführt, wobei eine Korrelation der Länge an Verbindungselementen mit der Zucker-Spezifität zu beobachten war. Kompetitive Bindung gegen an bovines Serumalbumin konjugiertes GalNAc und die Inhibition von Endocytose mit Cytochalasin D, resultierten in beiden Fällen in reduzierte Bindung, was einen weiteren Beweis für die Spezifität erbringt. Zusätzlich konnte erhöhte Bindung an ASGP-R beobachtet werden, indem trivalente mit mono- und bivalenten Neoglyko-GFPs verglichen wurden. Zusammengefasst: Diese Beobachtungen führten zu dem Schluss, dass es einen positiven Einfluss auf die effektive Bindung hat, wenn die Liganden in der richtigen Art präsentiert werden. Diese Entdeckungen unterstreichen das Konzept von rational entworfenen Neoglykoproteinen als strukturell-definierte Plattformen für die Aufklärung multivalenter Bindungsereignisse.

9.2. Synthese von Sialinsäure-Aziden für Bio-konjugationen

N-Acetyneuraminsäure (Neu5Ac) ist ein wichtiger Zucker, der oft am Ende von Säugetier-Glykanen zu finden ist. Aufgrund ihrer biologischen Relevanz sind Glykokonjugate von Neu5Ac wünschenswert. In diesem Zusammenhang wurde die chemische Synthese von Neu5Ac-Aziden für die Herstellung von multivalent-bindenden Glykoproteinen untersucht. Unterschiedliche Sialyl-Donoren wurden dargestellt und in Bezug auf ihre α-Selektivität untersucht. Nach nicht zufriedenstellenden anfänglichen Versuchen mit Koenigs-Knorr Donoren wurden Thiophenol-Donoren geprüft. Bedauerlicherweise war deren α-Selektivität nicht ausreichend für diesen Zweck. Schlussendlich wurden gute Ergebnisse mit Thioadamantan-Donoren nach Crich *et al.* erzielt. Mit diesem Donor-Promotor-System konnten Sialinsäureazide mit sehr guter α-Selektivität und hohen Ausbeuten erhalten werden.

Die Konjugation an Protein-Oberflächen konnte aufgrund fehlender Optimierung der CuAAC und Massenanalyse jedoch nicht realisiert werden.

9.3. Oberflächenplasmonenresonanz von Zucker-Polymer-Konjugaten

Die Effektivität der Staudinger-Phosphit-Reaktion (SPtR) und der Staudinger-Phosphonit-Reaktion (SPnR) als Konjugationsreaktion für die Funktionalisierung von Polyglycerol-Aziden wurde evaluiert. Für diesen Zweck wurden zuvor dargestellte Neoglykopolymere der Oberflächenplasmonenresonanz-Spektroskopie unterzogen.

Mit Glycosylphosphiten umgesetzte Polyglycerole wurden mit einem kompetitiven Lektin Bindungs-Assay gegen PNA untersucht. Bei dieser Studie wurde eine Inhibition von circa 60% erzielt. Ein anderer Versuch wurde bei Polyglycerolen angewandt, welche mit Laktose-Triazolyl-Phosphoniten funktionalisiert wurden. Dafür wurde PNA auf dem SPR Chip immobilisiert. Die Bindung der funktionalisierten Polyglycerole wurde gemessen und mit Azido-Polyglycerol, sowie mit an Polyacrylamid-konjugiertem Thomsen-Friedenreich-Antigen verglichen. Das SPR Experiment zeigte eine klare und starke Interaktion des Neoglyko-Polyglycerols mit oberflächengebundenem PNA.

Somit kann geschlussfolgert werden, dass sich beide Konjugationsreaktionen für die Funktionalisierung von Polymeraziden eignen. Die Tatsache, dass es sich hierbei um kupferfreie Ansätze handelt, bietet einen zusätzlichen Vorteil im Bezug auf die Herstellung von Zucker-präsentierenden Polyglycerolen.

10. Experimental Section

10.1. General

10.1.1.　NMR

Nuclear magnetic resonance spectra of protons (^{1}H NMR) and carbons (^{13}C NMR) were recorded on a Bruker ECX 400 (400 MHz for ^{1}H and 101 MHz for ^{13}C), on a Delta JEOL Eclipse 500 (500 MHz for ^{1}H and 176 MHz for ^{13}C), on a Bruker AV300 (300 MHz for ^{1}H and 75 MHz for ^{13}C), as well as on a Bruker AC250 (250 MHz for ^{1}H) spectrometer. Assignments of ^{1}H and ^{13}C NMR were achieved by using 2D methods (COSY, HMQC, HMBC). All chemical shifts (δ) are reported in parts per million (ppm) relative to the residual solvent peaks. [1] Coupling constants J were reported in Hz and observed multiplicities were annotated as s = singlet, br s = broad singlet, d = doublet, t = triplet, q = quadruplet, m = Multiplet, and combinations of them.

10.1.2.　Mass Spectrometry

10.1.2.1.　MALDI-TOF

Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) was carried out on a Ultraflex-II TOF/TOF instrument (Bruker Daltonics, Bremen, Germany) equipped with a 200 Hz solid-state Smart beam™ laser, an AB SCIEX 4700 MALDI TOF/TOF Proteomics Analyzer (Applied Biosystems Deutschland GmbH, Darmstadt, Germany) equipped with a 200 Hz Nd-YAG laser, or an AB SCIEX 5800 MALDI TOF/TOF Analyzer (Applied Biosystems Deutschland GmbH, Darmstadt, Germany) equipped with a 1000 Hz Nd-YAG laser. MALDI-TOF mass spectrometer were operated in the positive linear mode. Mass spectra were acquired over an m/z range of 4,000–15,000 for barstar proeins and over an m/z range of 5,000–40,000 for green fluorescent proteins. The obtained data were analyzed with the provided vendor software flexAnalysis® (Bruker Daltonics) and DataExplorer® (Applied Biosystems), as well as Microsoft Excel®.

Protein samples were desalted prior to the analysis, either by spin-filter tubes (Amicon®,

Millipore, or Vivaspin®, Sartorius) With an appropriate molecular-weight cut-off, or by precipitation in ice-cold acetone, followed by solubilization in ultrapure water. Sinapinic acid (SA) and 2,5-dihydroxy acetophenone (DHAP)[2] were used as matrices. Preparation and application of SA and DHAP was as reported before.[2], [3]

10.1.2.2.   ESI-TOF

Low resolution mass spectra (LRMS) and medium resolution mass spectra (MRMS) were recorded on a Micromass LCT Premier TOF spectrometer (Waters GmbH, Eschborn, Germany) either by direct injection over a syringe pump or by chromatography with a Acquity Ultra Performance LC® System (Waters GmbH, Eschborn, Germany). High resolution mass spectra (HRMS) of small molecules were obtained by direct injection or *via* HPLC to an Agilent 6210 ESI-TOF MS system (Agilent Technologies, Santa Clara, CA, U.S.A.).

LC systems and ESI-TOF mass spectrometer were operated as instructed or by qualified persons. The obtained spectra were analyzed with the provided vendor software, MassLynx® (Waters) and Mass Hunter® (Agilent), as well as Microsoft Excel®.

10.1.3.   Materials and Methods

Infrared (IR) spectra were recorded on a Bruker Tensor 27 FT-IR.

Thin-layer chromatography (TLC) was carried out on aluminium-backed silica gel plates (60 F_{254}, 0.2 mm, Merck, Darmstadt, Germany) and detected by UV, charring with 10% H_2SO_4 in ethanol, or staining with 4-anisaldehyde solution (16 mL 4-anisaldehyde, 2 mL H_2SO_4, 182 mL ethanol).

For column chromatography, silica gel 60 Å (230–400 mesh) or silica gel for flash chromatography (J.T. Baker) was used.

Acidic ion-exchange resins (DOWEX®-HCR-W2, DOWEX®-50WX4, Amberlite®-IR120) were washed extensively with methanol prior to use.

Reactions sensitive to moist or air were performed under an argon atmosphere in flame-dried glass vessels (Schlenk-type conditions). Dry solvents were obtained by Acros Organics or Sigma-Aldrich. Tetrahydrofuran (THF) was distilled prior to use in column chromatography. All other solvents and reagents were purchased from commercial suppliers as reagent grade and used without further purification.

For protein reactions, only ultra pure water was used. Buffers were always prepared freshly from either 10× stocks, stock solutions, or directly from the inorganic salts. Dulbecco's phosphate buffered saline was prepared as Ca^{2+}/Mg^{2+} free variant according to the literature.[4]

10.2. Carbohydrate Synthesis

10.2.1. General procedures for carbohydrate synthesis

10.2.1.1. Deacetylation under Zemplén conditions

Carbohydrate azides were deprotected using Zemplén conditions.[5] For this the carbohydrate azide (1 equiv.) was suspended in MeOH (3 mL/equiv.) and stirred at room temperature for 10 minutes. Subsequently a solution of NaOMe in MeOH (30 wt.%, 150 μL/equiv.) was added and the reaction was stirred at room temperature until completion. The mixture as then neutralized with methanol-washed ion exchange resin (H+) and filtered. Removal of solvents under reduced pressure is usually yielding unprotected carbohydrate azides without further purification.

10.2.1.2. Saponification of sialic acid methyl esters derivatives

The methyl ester (0.06 mmol) was dissolved in methanol (3 mL) and stirred at room temperature. Subsequently $LiOH \cdot H_2O$ (11 mg, 0.26 mmol, 4 equiv.) in H_2O (1 mL) was added and the reaction was stirred for 14 hours. Then it was cooled to 0 °C and acidified with DOWEX®-50WX4 resin for a few minutes. Subsequently the resin was removed by filtration and the free acid was obtained by lyophilization.

10.2.2. Galactose derivatives

10.2.2.1. Synthesis of 2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranose azide (1Ac)

The title compound was obtained as described before.[6] Briefly, D-(+)-galactose (1.8 g, 10.0 mmol) was dissolved in acetic anhydride (5 mL) and cooled with an ice bath. HBr/AcOH (30%, 2.8 mL, 10.4 mmol) was added and the reaction mixture was allowed to stir at room temperature for 45 minutes. Subsequently it was cooled again to 0 °C and HBr/AcOH (30%, 5.5 mL, 20.4 mmol) was added. The reaction was monitored by TLC (EtOAc/cyclo-hexane 1:1) and reagents were removed under reduced pressure after 2.5 hours. The brown syrup was dissolved in dry CH_2Cl_2 (55 mL) and sodium azide (1.3 g, 20.0 mmol) and tetrabutylammonium hydrogen sulfate (TBAHS) (515 mg, 1.5 mmol) and aqueous K_2CO_3 (1 M, 70 mL) were added consecutively. The 2-phase reaction mixture was stirred vigorously at room temperature over night. After diluting with CH_2Cl_2 the organic layer was separated, washed with water, aq. sat. $NaHCO_3$ and

brine. Subsequently it was dried (MgSO$_4$), filtered and evaporated to dryness. The crude product was purified by reduced pressure column chromatography (EtOAc/ cyclo-hexane 3:1). The obtained product was then further purified by recrystallization (cyclo-hexane/diethylether 1:1), yielding 1.49 g (4.0 mmol, 40%) of the title compound. The obtained spectra were in accordance with the literature.[6]

^{1}H NMR (400 MHz, CDCl$_3$): δ = 5.42 (dd, J = 3.3, 1.1 Hz, 1H), 5.16 (dd, J = 10.3, 8.7 Hz, 1H), 5.03 (dd, J = 10.4, 3.4 Hz, 1H), 4.59 (d, J = 8.7 Hz, 1H), 4.19 (dd, J = 10.0, 5.5 Hz, 1H), 4.14 (dd, J = 10.1, 4.9 Hz, 1H), 4.01 (td, J = 6.5, 1.2 Hz, 1H), 2.17 (s, 3H), 2.09 (s, 3H), 2.06 (s, 3H), 1.99 (s, 3H).

10.2.2.2.   Synthesis of β-D-galactopyranose azide (1)

Azide **1Ac** (2 mmol, 753 mg) was deacetylated according to the general procedure for deprotection under Zemplén conditions, yielding 371 mg (1.81 mmol, 90%) of the title compound. The obtained spectra were in accordance with the literature.[7]–[9]

^{1}H NMR (400 MHz, D$_2$O): δ = 4.63 (d, J = 8.7 Hz, 1H), 3.93 (d, J = 3.4 Hz, 1H), 3.80 – 3.67 (m, 3H), 3.65 (dd, J = 9.9, 3.4 Hz, 1H), 3.47 (dd, J = 9.5, 9.0 Hz, 1H).

^{13}C NMR (101 MHz, D$_2$O): δ = 90.6, 77.3, 72.7, 70.4, 68.6, 61.0.

HR-MS (ESI-TOF): m/z calcd. for C$_6$H$_{11}$N$_3$O$_5$Na$^+$ [M+Na]$^+$: 228.0591, found: 228.0605.

10.2.2.3.   Synthesis of 1-azidoethoxy-2,3,4,6-tera-*O*-acetyl-β-D-galactopyranose (2Ac)

As described before,[10] 1,2,3,4,6-penta-*O*-acetyl-β-D-galactopyranoside **7** (0.3 mmol, 117 mg) was dissolved in dry CH$_2$Cl$_2$ (600 μL). After addition of 2-azidoethanol (0.6 mmol, 52 mg) the mixture was cooled to 0 °C. Subsequently BF$_3$·OEt$_2$ (0.39 mmol, 50 μL) was added drop-wise to the solution. The reaction was then stirred for 21 hours and was allowed to reach room temperature. The mixture was diluted with CH$_2$Cl$_2$ and washed with aq. sat. NaHCO$_3$, water and brine. The organic layer was then dried over MgSO$_4$ and concentrated. The resultant product mixture was purified by column chromatography (EtOAc/cyclo-hexane 45:55), yielding 42 mg (0.1 mmol, 33%) of the title compound. The obtained spectra were in accordance with the literature.[10]

^{1}H NMR (400 MHz, CDCl$_3$): δ = 5.39 (dd, J = 3.4, 1.1 Hz, 1H), 5.23 (dd, J = 10.4, 8.0 Hz, 1H), 5.01 (dd, J = 10.5, 3.4 Hz, 1H), 4.55 (d, J = 8.0 Hz, 1H), 4.20-4.08 (m, 3H), 4.04 (ddd, J = 10.7, 4.6, 3.6 Hz, 1H), 3.92 (td, J = 6.7, 0.9 Hz, 1H), 3.68 (ddd, J = 10.7, 8.4, 3.3 Hz, 1H), 3.50 (ddd, J = 13.2, 8.6, 3.5 Hz, 1H), 3.29 (ddd, J = 13.4, 4.6, 3.4 Hz, 1H), 2.14 (s, 3H), 2.05 (s, 3H), 2.04 (s, 3H), 1.97 (s, 3H).

10.2.2.4. Synthesis of 1-azidoethoxy-β-ᴅ-galactopyranose (2)

Azide **2Ac** (0.21 mmol, 87 mg) was treated as described in the general procedure, yielding 44 mg (0.18 mmol, 84%) of unprotected azide **2**. The obtained spectra were in accordance with the literature.[10]

^{1}H-NMR (400 MHz, CD$_3$OD): δ = 4.27 (d, *J* = 7.5 Hz, 1H), 4.02 (dt, *J* = 10.8, 5.3 Hz, 1H), 3.84 (dd, *J* = 3.2, 0.9 Hz, 1H), 3.74 (dtd, *J* = 8.9, 6.2, 2.8 Hz, 3H), 3.54-3.50 (m, 2H), 3.49-3.46 (m, 3H).

^{13}C NMR (176 MHz, D$_2$O): δ = 103.0, 75.2, 72.8, 70.8, 68.7, 68.4, 61.0, 50.6.

HR-MS (ESI-TOF): m/z calcd. for C$_8$H$_{15}$N$_3$O$_6$Na$^+$ [M+Na]$^+$: 272.0853, found: 272.0866.

10.2.2.5. Synthesis of 2-[2-(2-azidoethoxy)ethoxy]ethoxy-2,3,4,6-tetra-*O*-acetyl-β-ᴅ-galacto-pyranoside (3Ac)

To 1,2,3,4,6-penta-O-acetyl-β-D-galactopyranoside **7** (12.8 mmol, 5 g) in dry CH$_2$Cl$_2$ (60 mL), 2-[2-(2-chloro-ethoxy)ethoxy]ethanol (25.6 mmol, 4.3 g) was added and the mixture was stirred for 20 minutes at −4 °C. Subsequently TMSOTf (17.5 mmol, 3.9 g, 3 mL) was added drop-wise over 30 minutes. After stirring for 4 hours at −4 °C, the reaction mixture was diluted with CH$_2$Cl$_2$ (40 mL) and consecutively washed with aq. sat. NaHCO$_3$, water and brine. The organic layer was then dried (MgSO$_4$), filtered and concentrated *in vacuo*. 2.8 g of the purified intermediate (column chromatography: EtOAc/cyclo hexane gradient 1:3→1:1) were dissolved in DMF (15 mL) and NaN$_3$ (11 mmol, 715 mg) was added. It was then stirred for 20 hours at 90 °C. The reaction mixture was subsequently diluted with EtOAc (85 mL) and washed with water and brine. After drying (MgSO$_4$), filtration and concentration, 1.7 g (3.4 mmol, 27% over two steps) of **3Ac** were obtained. The obtained spectra were in accordance with the literature.[11]

^{1}H NMR (400 MHz, CDCl$_3$): δ = 5.38 (dd, *J* = 3.4, 1.0 Hz, 1H), 5.20 (dd, *J* = 10.5, 8.0 Hz, 1H), 5.01 (dd, *J* = 10.5, 3.4 Hz, 1H), 4.57 (d, *J* = 8.0 Hz, 1H), 4.17 (dd, *J* = 11.2, 6.6 Hz, 1H), 4.12 (dd, *J* = 11.2, 6.8 Hz, 1H), 3.96 (dt, *J* = 11.0, 4.2 Hz, 1H), 3.90 (td, *J* = 6.7, 1.0 Hz, 1H), 3.75 (ddd, *J* = 11.0, 6.9, 4.0 Hz, 1H), 3.70 – 3.61 (m, 8H), 3.39 (t, *J* = 5.2, 2H), 2.14 (s, 3H), 2.06 (s, 3H), 2.04 (s, 3H), 1.98 (s, 3H).

10.2.2.6. Synthesis of 2-[2-(2-azidoethoxy)ethoxy]ethoxy-β-ᴅ-galactopyranoside (3)

The acetylated precursor **3Ac** (25 mg, 50 μmol) was deprotected as described in the general procedure, yielding 14 mg (0.04 mmol, 80%) of azide **3**. The obtained spectra were in accordance with the literature. [12]

^{1}H NMR (400 MHz, CD$_3$OD): δ = 4.26 (d, *J* = 7.5 Hz, 1H), 4.05 – 3.99 (m, 1H), 3.82 (dd,

J = 3.3, 1.0 Hz, 1H), 3.76 – 3.64 (m, 11H), 3.56 – 3.49 (m, 2H), 3.47 (dd, *J* = 9.7, 3.3 Hz, 1H), 3.39 (t, *J* = 5.0, 2H).

^{13}C NMR (101 MHz, D$_2$O): δ = 102.9, 75.2, 72.8, 70.8, 69.8, 69.7, 69.5, 69.3, 68.7, 61.0, 50.2.

HR-MS (ESI-TOF): m/z calcd. for C$_{12}$H$_{23}$N$_3$O$_8$Na$^+$ [M+Na]$^+$: 360.1377, found: 360.1391.

10.2.3.   Lactose derivatives

10.2.3.1.   Synthesis of (2,3,4,6,-tetra-*O*-acetyl-β-ᴅ-galacto-pyranosyl)-(1→4)-1,2,3,6-tetra-*O*-acetyl-β-ᴅ-glucopyranoside (9)

Starting from ᴅ-lactose, the title compound was synthesized according to a procedure by Hudson and Johnson.[13]

^{1}H NMR (250 MHz,CDCl$_3$): δ = 5.65 (d, *J* = 8.2 Hz, 1H), 5.33 (d, *J* = 3.3 Hz, 1H), 5.22 (t, *J* = 9.0 Hz, 1H), 5.05 (ddd, *J* = 16.1, 9.9, 8.0 Hz, 2H), 4.92 (dd, *J* = 10.4, 3.4 Hz, 1H), 4.49 – 4.37 (m, 2H), 4.17 – 3.97 (m, 3H), 3.91 – 3.67 (m, 3H), 2.13 (s, 3H), 2.10 (s, 3H), 2.07 (s, 3H), 2.05 (s, 3H), 2.03 (s, 3H), 2.02 (s, 3H), 2.01 (s, 3H), 1.94 (s, 3H).

10.2.3.2.   Synthesis of 1-bromo-(2,3,4,6-tetra-*O*-acetyl-β-ᴅ-galacto-pyranosyl)-(1→4)-2,3,6-tri-*O*-acetyl-α-ᴅ-glucopyranoside (10)

As described previously,[14] β-ᴅ-lactosyl octa-*O*-acetate **9** (6.8 g, 10.0 mmol) was dissolved in dry CH$_2$Cl$_2$ (28 mL). Acetic anhydride (1 mL) was added and the mixture was stirred 15 minutes on ice. Then HBr (33 wt.-% in acetic acid) was transferred to the mixture at 0 °C. After 30 minutes the ice-bath was removed and the reaction was left at room temperature (≈22 °C) for 2 hours. The reaction was therefore diluted with CH$_2$Cl$_2$ (90 mL) and washed consecutively with aqueous sat. NaHCO$_3$, water and brine. The organic layer was dried over MgSO$_4$ and concentrated under reduced pressure. Subsequently the obtained white foam was recrystallized from CHCl$_3$/cyclo-hexane, yielding bromide **10** quantatively and was used without further purification. The obtained spectra was in accordance with the literature.[14]

^{1}H NMR (400 MHz, CDCl$_3$): δ = 6.52 (d, *J* = 4.0 Hz, 1H), 5.55 (dd, *J* = 9.9, 9.4 Hz, 1H), 5.35 (dd, *J* = 3.5, 1.2 Hz, 1H), 5.12 (dd, *J* = 10.4, 7.9 Hz, 1H), 4.95 (dd, *J* = 10.3, 3.4 Hz, 1H), 4.75 (dd, *J* = 10.0, 4.1 Hz, 1H), 4.55 – 4.45 (m, 2H), 4.24 – 4.02 (m, 4H), 3.93 – 3.80 (m, 2H), 2.15 (s, 3H), 2.13 (s, 3H), 2.09 (s, 3H), 2.06 (s, 3H), 2.06 (s, 3H), 2.05 (s, 3H), 1.96 (s, 3H).

10.2.3.3. Synthesis of 1-azido-(2,3,4,6,-tetra-*O*-acetyl-β-D-galactopyranosyl)-(1→4)-2,3,6-tri-*O*-acetyl-β-D-glucopyranoside (4Ac)

Unprotected D-lactose (2.22 g, 6.5 mmol) was stirred in acetic anhydride (5 mL) at 0 °C. HBr (33 wt-% in AcOH, 1.6 mL, 6.5 mmol) was added in one portion and the warmed reaction was further stirred at room temperature for 30 minutes. Subsequently HBr (33 wt-% in AcOH, 3.2 mL, 13.0 mmol) was added dropwise over 30 minutes to the reaction under ice-bath cooling. After an additional hour at room temperature, the reaction mixture was diluted with toluene and volatile reagents and solvents were removed by evaporation. The yellowish residue was then taken up in anhydrous CH_2Cl_2 (10 mL) and consecutively sodium azide (845 mg, 13 mmol) and tetrabutylammonium bisulfate (331 mg, 975 μmol) were added. The mixture was overlaid with K_2CO_3 solution (1 M in H_2O, 10 mL) and the reaction was started by vigorous stirring. After 4 h it was diluted with CH_2Cl_2 and the phases were separated. The organic layer was then washed with aqueous sat. $NaHCO_3$, water and brine. Following drying over $MgSO_4$ and filtration, the solvents were removed under reduced pressure, affording **4Ac** as a yellowish syrup (2.8 g, 4.23 mmol, 65%). The obtained spectra were in accordance with the literature.[6]

^{1}H NMR (400 MHz, CDCl$_3$): δ 5.43 (d, *J* = 3.4 Hz, 1H), 5.29 (t, *J* = 9.2 Hz, 1H), 5.18 (dd, *J* = 10.2, 8.0 Hz, 1H), 5.04 (dd, *J* = 10.3, 3.3 Hz, 1H), 4.94 (t, *J* = 9.1 Hz, 1H), 4.71 (d, *J* = 8.5 Hz, 1H), 4.61-4.56 (m, 2H), 4.22-4.16 (m, 3H), 3.96 (t, *J* = 6.7 Hz, 1H), 3.90 (t, *J* = 9.5 Hz, 1H), 3.78 (ddd, *J* = 10.0, 5.0, 1.8 Hz, 1H), 2.23 (s, 3H), 2.21 (s, 3H), 2.15 (2x s, 6H), 2.13 (s, 6H), 2.04 (s, 3H).

10.2.3.4. Synthesis of 2-azido-(β-D galactopyranosyl)-(1→4)-β-D-glucopyranoside (4)

Previously obtained **4Ac** (3.93 mmol, 2.6 g) were dissolved in MeOH (40 mL). To this a solution of NaOMe in MeOH (0.1 M, 800 μL) was added and the reaction was stirred at room temperature for 3 hours. Subsequently another portion of NaOMe sol. (0.1 M, 1.2 mL) was added and the reaction was further stirred over night at room temperature. White crystals formed and were filtered off.

^{1}H NMR (400 MHz, D$_2$O): δ = 4.76 (d, *J* = 8.8 Hz, 2H), 4.44 (d, J = 7.8 Hz, 1H), 3.97 (d, *J* = 12.0 Hz, 1H), 3.91 (d, *J* = 3.4 Hz, 1H), 3.84-3.80 (m, 1H), 3.79-3.74 (m, 2H), 3.72-3.69 (m, 1H), 3.65 (dd, *J* = 2.5, 0.9 Hz, 1H), 3.64 (d, *J* = 3.3 Hz, 1H), 3.55-3.51 (m, 1H), 3.33 (s, 2H), 3.32-3.30 (m, 1H).

^{13}C NMR (101 MHz, D$_2$O): δ = 102.9, 90.0, 77.8, 76.7, 75.4, 74.4, 72.6, 72.6, 71.0, 68.6, 61.1, 59.9.

HR-MS (ESI-TOF): m/z calcd. for $C_{12}H_{21}N_3O_{10}Na^+$ [M+Na]$^+$: 390.1119, found: 390.1135.

10.2.3.5. Synthesis of 1-azidoethoxy-(2,3,4,6,-tetra-*O*-acetyl-β-ᴅ-galactopyranosyl)-(1→4)-2,3,6-tri-*O*-acetyl-β-ᴅ-glucopyranoside (5Ac) *via* 10

As previously described,[14] β-ᴅ-lactosyl bromide **10** (784 mg, 1.12 mmol) was dissolved in dry acetonitrile (3 mL) and added to a mixture of HgBr$_2$ (202 mg, 0.56 mmol, 0.5 equiv.), Hg(CN)$_2$ (141 mg, 0.56 mmol, 0.5 equiv.) and 2-azidoethanol (107 mg, 1.23 mmol, 1.1 equiv.) in dry acetonitrile (5 mL). The reaction was stirred overnight (16 h) at room temperature (≈22 °C). Then it was cooled to 4 °C, diluted with CH$_2$Cl$_2$ (50 mL) and consecutively washed with aqueous sat. NaHCO$_3$, water and brine. The organic phase was dried over MgSO$_4$, filtered and evaporated to dryness. The residue was subjected to column chromatography (EtOAc/cyclo-hexane gradient, 25:75 → 75:25), resulting in 480 mg (0.68 mmol, 60%) of azide **5Ac**. The obtained spectra was in accordance with the literature.[14], [15]

^{1}H NMR (400 MHz, CDCl$_3$): δ = 5.34 (dd, *J* = 3.4, 1.0 Hz, 1H), 5.20 (t, *J* = 9.2 Hz, 1H), 5.11 (dd, *J* = 10.4, 7.9 Hz, 1H), 4.94 (ddd, *J* = 17.3, 9.9, 5.6 Hz, 2H), 4.56 (d, *J* = 7.9 Hz, 1H), 4.53 (dd, *J* = 12.2, 2.3 Hz, 1H), 4.49 (d, *J* = 7.9 Hz, 1H), 4.16 – 4.05 (m, 3H), 3.99 (ddd, *J* = 10.7, 5.0, 3.5 Hz, 1H), 3.90 – 3.85 (m, 1H), 3.85 – 3.79 (m, 1H), 3.71 – 3.65 (m, 1H), 3.62 (ddd, *J* = 9.9, 4.9, 2.1 Hz, 1H), 3.51 – 3.44 (m, 1H), 3.27 (ddd, *J* = 13.5, 5.0, 3.4 Hz, 1H), 2.15 (s, 3H), 2.12 (s, 3H), 2.06 (s, 3H), 2.04 (3x s, 9H), 1.96 (s, 3H).

10.2.3.6. Synthesis of 2-azidoethoxy-(β-ᴅ-galactopyranosyl)-(1→4)-β-ᴅ-glucopyranoside (5)

The acetylated lactosyl azide **5Ac** (0.25 mmol, 176 mg) was deprotected according to the general procedure, yielding 83 mg (0.20 mmol, ≈80%) of the title compound as a white solid. The obtained spectra was in accordance with the literature.[14], [15]

^{1}H NMR (400 MHz, CD$_3$OD): δ = 4.36 (dd, *J* = 7.7, 4.4 Hz, 2H), 4.04 – 3.98 (m, 1H), 3.92 (dd, *J* = 12.1, 2.5 Hz, 1H), 3.86 – 3.81 (m, 2H), 3.81 – 3.75 (m, 2H), 3.75 – 3.68 (m, 2H), 3.61 – 3.45 (m, 8H), 3.44 – 3.40 (m, 1H), 3.27 (dd, *J* = 8.7, 8.1 Hz, 1H).

^{13}C NMR (101 MHz, D$_2$O): δ 103.0, 102.2, 78.4, 75.4, 74.9, 74.4, 72.8, 72.6, 71.0, 68.7, 68.6, 61.1, 60.1, 50.6.

HR-MS (ESI-TOF): m/z calcd. for $C_{14}H_{25}N_3O_{11}Na^+$ [M+Na]$^+$: 434.1381, found: 434.1397.

10.2.3.7. Synthesis of 2-[2-(2-azidoethoxy)ethoxy]ethoxy-(2,3,4,6,-tetra-*O*-acetyl-β-D-galactopyranosyl)-(1→4)-2,3,6-tri-*O*-acetyl-β-D-glucopyranoside (6Ac)

β-D-Lactosyl octa-*O*-acetate **9** (1.5 g, 2.21 mmol) was transferred into a round-bottom flask containing dry CH_2Cl_2 (15 mL) and molecular sieves (4Å, 1 g). 2-[2-(2-chloroethoxy)ethoxy]ethanol (963 μL, 6.63 mmol, 3.0 eq.) was added and the solution was stirred for 1 hour at room temperature (≈22 °C). It was then cooled to 0 °C with an ice-salt bath and TMSOTf (975 μL, 4.41 mmol, 2 eq.) was added drop wise over 5 minutes. The mixture was left at 0 °C for approximately 6 hours, followed by dilution with CH_2Cl_2 (90 mL) and consecutively washing with aqueous sat. $NaHCO_3$ and brine. The organic layer was dried over $MgSO_4$, filtered and concentrated in vacuo, yielding slightly yellow syrup. Excess of 2-[2-(2-chloroethoxy)ethoxy]ethanol was removed by decanting with water and column chromatography (EtOAc/cyclo-hexane gradient, 50:50 → 66:33). The intermediate product was dissolved in DMF (10 mL) and reacted with sodium azide (198 mg, 3.05 mmol, 1.38 eq.) at 80 °C for 18 hours. After cooling to room temperature the reaction mixture was diluted with EtOAc and washed with water and brine. The organic phase was then dried over $MgSO_4$, filtered and evaporated to dryness, yielding clear syrup (300 mg, 0.38 mmol, 17%). The obtained spectra was in accordance with the literature.[15]

^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.34 (dd, *J* = 3.5, 1.1 Hz, 1H), 5.23 – 5.15 (m, 1H), 5.10 (dd, *J* = 10.4, 7.9 Hz, 1H), 4.95 (dd, *J* = 10.4, 3.5 Hz, 1H), 4.89 (dd, *J* = 9.6, 8.0 Hz, 1H), 4.56 (d, *J* = 7.9 Hz, 1H), 4.49 – 4.46 (m, 1H), 4.09 (dt, *J* = 11.1, 7.1 Hz, 3H), 3.94 – 3.84 (m, 2H), 3.82 – 3.75 (m, 1H), 3.75 – 3.57 (m, 11H), 3.41 – 3.36 (m, 2H), 2.14 (s, 3H), 2.11 (s, 3H), 2.05 (s, 3H), 2.05(3x s, 9H), 1.96 (s, 3H).

10.2.3.8. Synthesis of 2-[2-(2-azidoethoxy)ethoxy]ethoxy-(β-D-galactopyanosyl)-(1→4)-β-D-glucopyranoside (6)

The acetylated lactosyl azide **6Ac** (0.25 mmol, 198 mg) was deprotected according to the general procedure, yielding the title compound as syrup (97 mg, 0.19 mmol, ≈79%). The obtained spectra were in accordance with the literature.[16]

^{1}H NMR (500 MHz, CD_3OD): δ = 4.35 (t, *J* = 8.1 Hz, 2H), 4.05 – 3.98 (m, 1H), 3.90 (dd, *J* = 12.1, 2.4 Hz, 1H), 3.83 (dt, *J* = 11.3, 6.1 Hz, 2H), 3.79 – 3.73 (m, 2H), 3.71 (dd, *J* = 5.6, 4.0 Hz, 3H), 3.70 – 3.66 (m, 7H), 3.61 – 3.55 (m, 3H), 3.52 (d, *J* = 8.6 Hz, 1H), 3.49 (dd, *J* = 9.7, 3.2 Hz, 1H), 3.44 – 3.40 (m, 1H), 3.40 – 3.37 (m, 2H), 3.27 (t, *J* = 8.4 Hz, 1H).

^{13}C NMR (176 MHz, D_2O): δ = 103.0, 102.1, 78.5, 75.4, 74.8, 74.37, 72.9, 72.6, 71.0, 69.7, 69.6, 69.5, 69.2, 68.8, 68.6, 61.1, 60.2, 50.2.

HR-MS (ESI-TOF): m/z calcd. for $C_{18}H_{33}N_3O_{13}Na^+$ [M+Na]$^+$: 522.1906, found: 522.1936.

10.2.4.   N-Acetyl galactosamine derivatives

10.2.4.1.   Synthesis of 2-acetamido-1, 3, 4, 6-tetra-*O*-acetyl-2-deoxy-β-ᴅ-galactopyranoside (11)

ᴅ-Galactosamine hydrochloride (323 mg, 1.5 mmol) was dissolved in pyridine (9 mL) and cooled to 0 °C when Ac_2O (6 mL) was added. The reaction mixture was then stirred for 16 hours and was allowed to reach room temperature. Subsequently fully acetylated **11** (520 mg, 1.34 mmol, 89%) was obtained by co-evaporation with toluene and recrystallized from MeOH. The obtained spectra were in accordance with the literature. [17]

^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.70 (d, *J* = 8.8 Hz, 1H), 5.54 (d, *J* = 9.5 Hz, 1H), 5.37 (dd, *J* = 3.3, 1.1 Hz, 1H), 5.09 (ddd, *J* = 11.3, 3.4, 0.9 Hz, 1H), 4.44 (dt, *J* = 11.3, 8.3 Hz, 1H), 4.17 (dd, *J* = 11.3, 6.7 Hz, 1H), 4.11 (dd, *J* = 11.3, 6.4 Hz, 1H), 4.02 (td, *J* = 6.5, 6.5, 1.2 Hz, 1H), 2.17 (s, 3H), 2.13 (s, 3H), 2.04 (s, 3H), 2.02 (s, 3H), 1.94 (s, 3H).

10.2.4.2.   Synthesis of 1-chloro-2-acetamido-3, 4, 6-tri-*O*-acetyl-2-deoxy-β-ᴅ-galactopyranoside (12)

Chloride **12** was prepared according to Ref. [18]. Briefly, galactosamine **11** (500 mg, 1.28 mmol) dissolved in CH_2Cl_2 (dry, 5 mL). To this was added $TiCl_4$ (1.69 mmol, 185 μL) and the mixture was stirred under argon at room temperature for 72 hours. Subsequently the solvents were removed under reduced pressure and the yellowish crude was purified by column chromatography (cyclo-hexane/EtOAc gradient, 60:40 → 20:80), yielding 328 mg (0.897 mmol, 70%) of the title compound.

^{1}H NMR (400 MHz, $CDCl_3$): δ = 6.27 (d, *J* = 3.7 Hz, 1H), 5.71 (d, *J* = 8.9 Hz, 1H), 5.46 (dd, *J* = 3.3, 1.3 Hz, 1H), 5.28 (dd, *J* = 11.3, 3.3 Hz, 1H), 4.79 (ddd, *J* = 11.2, 8.9, 3.7 Hz, 1H), 4.52 – 4.45 (m, 1H), 4.18 (dd, *J* = 11.4, 6.3 Hz, 1H), 4.08 (dd, *J* = 10.8, 6.1 Hz, 1H), 2.16 (s, 3H), 2.06 (s, 3H), 2.02 (s, 3H), 2.00 (s, 3H).

10.2.4.3.   Synthesis of 1-azido-2-acetamido-3, 4, 6-tri-*O*-acetyl-2-deoxy-β-ᴅ-galactopyranoside (G1Ac) *via* 11

Fully acetylated galactosamine **11** (38.9 mg, 0.10 mmol) was added to a pre-cooled solution of HBr in glacial acetic acid (33 wt-%, 2 mL). After 3 hours at 0 °C, the reaction mixture was diluted with CH_2Cl_2 and consecutively washed with cold water and aqueous sat. $NaHCO_3$. The organic layer was then dried over $MgSO_4$, filtered and concentrated under reduced pressure. The residue was dissolved in anhydrous CH_2Cl_2 (2 mL) and sodium azide (585 mg, 90.0 mmol) in water (2 mL) as well as TBAHS

(33.9 mg, 0.10 mmol) were added. Subsequently the mixture was vigorously stirred at room temperature for 4 hours, at which point the reaction was diluted with CH_2Cl_2, followed by separation of the phases. The aqueous layer was extracted with CH_2Cl_2 and the combined organic layers were dried over $MgSO_4$, filtered and concentrated. Column chromatography (cyclo-hexane/EtOAc, 20:80 then 10:90) of the residue yielded 9 mg (0.024 mmol, 24%) of the protected galactosamine azide **G1Ac**. The obtained spectra were in accordance with the literature.[19]–[21]

^{1}H NMR (400 MHz, CDCl$_3$): δ = 5.56 (d, J = 8.8 Hz, 1H), 5.38 (dd, J = 3.3, 1.2 Hz, 1H), 5.24 (dd, J = 11.2, 3.3 Hz, 1H), 4.79 (d, J = 9.2 Hz, 1H), 4.19 – 4.09 (m, 2H), 4.07 – 3.97 (m, 2H), 2.16 (s, 3H), 2.06 (s, 3H), 2.01 (s, 3H), 1.99 (s, 3H).

10.2.4.4. Synthesis of 1-azido-2-acetamido-3, 4, 6-tri-*O*-acetyl-2-deoxy-β-ᴅ-galactopyranoside (G1Ac) *via* 12

Galactosamine chloride **12** (153.6 mg, 0.42 mmol) was dissolved in DMSO (3 mL) and sodium azide (162.6 mg, 2.50 mmol, 5.95 eq.) was added. The reaction mixture was then stirred at room temperature until completion (TLC). Subsequently it was diluted with water (60 mL) and extracted with CH_2Cl_2. The combined organic layers were then washed with brine, dried over Na_2SO_4, filtered and evaporated to dryness. The residue was purified by column chromatography on silica (cyclo-hexane/EtOAc, 20:80 then 10:90), yielding 74 mg (0.20 mmol, 47%) of the acetylated precursor **G1Ac**.

The obtained spectra were in accordance with the literature.[19]–[21]

^{1}H NMR (400 MHz, CDCl$_3$) δ = 5.56 (d, J = 8.8 Hz, 1H), 5.38 (dd, J = 3.3, 1.2 Hz, 1H), 5.24 (dd, J = 11.2, 3.3 Hz, 1H), 4.79 (d, J = 9.2 Hz, 1H), 4.19 – 4.09 (m, 2H), 4.07 – 3.97 (m, 2H), 2.16 (s, 3H), 2.06 (s, 3H), 2.01 (s, 3H), 1.99 (s, 3H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 170.6, 170.6, 170.5, 170.2, 88.8, 72.9, 69.8, 66.6, 61.5, 50.9, 23.5, 20.8, 20.8, 20.7.

10.2.4.5. Synthesis of 1-azido-2-acetamido-2-deoxy-β-ᴅ-galacto-pyranoside (G1)

Acetylated azide **G1Ac** (0.20 mmol, 74 mg) were deprotected under Zemplén conditions according to the general procedure. Recrystallization of the residue from ethanol–*n*-pentane afforded the title compound in 80% (0.16 mmol, 39 mg) yield. Analytical data was in accordance with the literature. [20], [21]

^{1}H NMR (400 MHz, CD$_3$OD): δ = 4.45 (d, J = 9.3 Hz, 1H), 3.98 (dd, J = 10.6, 9.3 Hz, 1H), 3.87 (dd, J = 3.4, 0.9 Hz, 1H), 3.81 (dd, J = 11.4, 6.9 Hz, 1H), 3.74 (dd, J = 11.5, 5.1 Hz, 1H), 3.64 – 3.58 (m, 2H), 1.99 (s, 3H).

^{13}C NMR (101 MHz, CD$_3$OD): δ = 174.1, 90.5, 79.1, 72.9, 69.5, 62.5, 53.4, 22.9.

10.2.4.6. Synthesis of 2-methyl-(3,4,6-tri-*O*-acetyl-1,2-dideoxy-α-ᴅ-galactopyrano)[1,2-d]-oxazoline (13) – FeCl₃ procedure

Analogues to reference [22], anhydrous FeCl₃ (811 mg, 5.0 mmol) was dried under vacuum and put under argon atmosphere. To this a solution of **11** (779 mg, 2.0 mmol) in anhydrous CH_2Cl_2 (15 mL) was added and the mixture was stirred under argon at room temperature for 1.5 hours. Subsequently ice-cold water was poured into it and the layers were separated. The aqueous phase was back-extracted with CH_2Cl_2 and the combined organic layers were consecutively washed with aqueous sat. $NaHCO_3$ and brine. Drying over $MgSO_4$, filtering and evaporation of the solvents yielded a residue that was further subjected to column chromatography on silica gel (EtOAc) for purification. In the end, 433 mg (1.31 mmol, 66%) of oxazoline **13** could be obtained. The obtained spectra were in accordance with the literature. [22],[23]

^{1}H NMR (400 MHz, CDCl₃): δ = 5.99 (d, *J* = 6.8 Hz, 1H), 5.46 (t, *J* = 3.0 Hz, 1H), 4.90 (dd, *J* = 7.5, 3.3 Hz, 1H), 4.28 – 4.15 (m, 2H), 4.10 (dd, *J* = 11.0, 5.7 Hz, 1H), 3.99 (td, *J* = 7.3, 1.4 Hz, 1H), 2.12 (s, 3H), 2.06 (s, 6H), 2.05 (d, *J* = 1.4 Hz, 3H).

10.2.4.7. Synthesis of 2-methyl-(3,4,6-tri-*O*-acetyl-1,2-dideoxy-α-ᴅ-galactopyrano)[1,2-d]-oxazoline (13) – TMSOTf procedure

Analogous to reference [23], per-acetylated galactosamine **11** (2.0 g, 5.14 mmol) was dissolved in 1,2-dichloroethane (20 mL) and stirred at 50 °C. Subsequently TMSOTf (1.1 mL, 6.08 mmol, 1.18 eq.) was added to the solution and the reaction was further conducted at 50 °C for 17 hours. The mixture was then cooled to 0 °C and quenched with excess triethylamine. Subsequently it was concentrated under reduced pressure and purified by column chromatography on silica gel (EtOAc), yielding 1.2 g (3.64 mmol, 71%) of oxazoline **13**. The obtained spectra were in accordance with the literature. [22],[23]

^{1}H NMR (400 MHz, CDCl₃): δ = 5.93 (d, *J* = 6.8 Hz, 1H), 5.39 (t, *J* = 3.0 Hz, 1H), 4.85 (dd, *J* = 7.4, 3.4 Hz, 1H), 4.19 (ddd, *J* = 6.9, 5.8, 2.7 Hz, 1H), 4.13 (dd, *J* = 11.2, 6.9 Hz, 1H), 4.05 (dd, *J* = 11.1, 5.8 Hz, 1H), 3.94 (ddd, *J* = 7.6, 6.9, 1.4 Hz, 1H), 2.06 (s, 3H), 2.01 (s, 6H), 1.99 (d, *J* = 1.4 Hz, 3H).

10.2.4.8. Synthesis of 2-azidoethoxy-3,4,6-tetra-*O*-acetyl-2-*N*-acetylamido-2-deoxy-β-ᴅ-galactopyranoside (G2Ac)

Oxazoline **13** (351 mg, 1.07 mmol), 2-azidoethanol (232 mg, 2.17 mmol) and powdered molecular sieve 3Å (300 mg) were stirred in anhydrous CH_2Cl_2 (4 mL) for 1 hour under argon atmosphere. Subsequently a catalytic amount of sulfuric acid (20 µL) was added drop-wise and the reaction was stirred for additional 18 hours at room temperature. After dilution with CH_2Cl_2 and quenching with excess triethylamine, the mixture was

filtered and concentrated under reduced pressure. The residue was purified by column chromatography on silica gel (cyclo-hexane/EtOAc gradient, 80:20 → 20:80), affording carbohydrate azide **G2Ac** in 52% yield (232 mg, 0.56 mmol). The obtained spectra were in accordance with the literature. [22], [23]

^{1}H NMR (400 MHz, CDCl$_3$): δ = 5.56 (d, J = 8.5 Hz, 1H), 5.41 – 5.35 (m, 2H), 4.87 (d, J = 8.4 Hz, 1H), 4.17 (dd, J = 11.2, 6.7 Hz, 1H), 4.12 (dd, J = 11.2, 6.6 Hz, 1H), 4.07 (ddd, J = 10.8, 4.6, 3.3 Hz, 1H), 3.95 (td, J = 6.6, 1.1 Hz, 1H), 3.93 – 3.87 (m, 1H), 3.71 (ddd, J = 10.8, 8.7, 3.1 Hz, 1H), 3.53 (ddd, J = 13.4, 8.7, 3.2 Hz, 1H), 3.27 (ddd, J = 13.5, 4.6, 3.0 Hz, 1H), 2.14 (s, 3H), 2.05 (s, 3H), 2.00 (s, 3H), 1.96 (s, 3H).

10.2.4.9. Synthesis of 2-azidoethoxy-2-*N*-acetylamido-2-deoxy-β-ᴅ-galactopyranoside (G2)

Galactosamine **G2Ac** (216 mg, 0.51 mmol) were deprotected under Zemplén conditions according to the general procedure. Recrystallization of the residue from ethanol–Et$_2$O afforded the title compound in 79% (119 mg, 0.41 mmol) yield. Analytical data were in accordance with the literature. [23]

^{1}H NMR (500 MHz, CD$_3$OD): δ = 4.46 (d, J = 8.4 Hz, 1H), 4.04 (ddd, J = 10.9, 5.3, 3.3 Hz, 1H), 3.95 (dd, J = 10.7, 8.4 Hz, 1H), 3.83 (dd, J = 3.3, 1.0 Hz, 1H), 3.78 (dd, J = 11.3, 6.9 Hz, 1H), 3.74 (dd, J = 11.3, 5.2 Hz, 1H), 3.68 (ddd, J = 11.1, 8.0, 3.3 Hz, 1H), 3.60 (dd, J = 10.7, 3.3 Hz, 1H), 3.50 (ddd, J = 6.7, 5.3, 1.1 Hz, 1H), 3.44 (ddd, J = 13.4, 8.0, 3.3 Hz, 1H), 1.98 (s, 3H).

HR-MS (ESI): m/z calculated for C$_{10}$H$_{19}$N$_4$O$_6$$^+$ [M+H]$^+$: 291.1299, found: 291.1282.

10.2.4.10. Synthesis of 2-[2-(2-azidoethoxy)ethoxy]ethoxy-3,4,6-tetra-*O*-acetyl-2-*N*-acetylamido-2-deoxy-β-ᴅ-galactopyranoside (G3Ac)

Oxazoline **13** (296 mg, 0.90 mmol) and 2-[2-(2-azidoethoxy)ethoxy]ethanol (175 mg, 1.0 mmol, 1.1 eq.) were stirred in anhydrous CH$_2$Cl$_2$ (6 mL), containing powdered molecular sieves (4Å, 300 mg), under argon at room temperature for 1.5 hours. Then TMSOTf (95 mg, 0.43 mL) was added drop wise. After 20 hours at room temperature, the mixture was cooled to 0 °C and quenched with excess triethylamine. Following removal of the volatile under reduced pressure, column chromatography on silica gel (CH$_2$Cl$_2$/MeOH gradient, 100:0 → 96:4, then cyclo hexane/EtOAc 20:80) was carried out. Finally, 160 mg (0.32 mmol, 35%) of galactosamine azide **G3Ac** could be obtained. Spectroscopical data were in accordance with the literature.[24]

^{1}H NMR (300 MHz, CDCl$_3$): δ = 6.16 (d, J = 9.2 Hz, 1H), 5.30 (d, J = 3.3 Hz, 1H), 5.05 (dd, J = 11.1, 3.3 Hz, 1H), 4.75 (d, J = 8.5 Hz, 1H), 4.26 – 4.03 (m, 3H), 3.93 – 3.78 (m, 3H),

3.75 – 3.58 (m, 7H), 3.44 (t, J = 5.0 Hz, 2H), 2.13 (s, 3H), 2.02 (s, 3H), 1.97 (s, 3H), 1.96 (s, 3H).

^{13}C NMR (75 MHz, CDCl$_3$): δ = 170.7, 170.7, 170.6, 170.5, 102.3, 71.7, 71.0, 70.8, 70.7, 70.5, 69.9, 68.7, 66.8, 61.7, 50.8, 50.6, 23.3, 20.8, 20.8, 20.8.

10.2.4.11. Synthesis of 2-[2-(2-azidoethoxy)ethoxy]ethoxy-2-*N*-acetylamido-2-deoxy-β-ᴅ-galactopyranoside (G3)

Acetylated azide **G3Ac** (0.20 mmol, 74 mg) were deprotected under Zemplén conditions according to the general procedure. Recrystallization of the residue from ethanol–*n*-pentane afforded the title compound in 80% (0.16 mmol, 39 mg) yield.

^{1}H NMR (500 MHz, CD$_3$OD): δ = 4.44 (d, J = 8.5 Hz, 1H), 3.98 – 3.90 (m, 2H), 3.82 (dd, J = 3.3, 1.0 Hz, 1H), 3.80 – 3.71 (m, 3H), 3.71 – 3.61 (m, 8H), 3.58 (dd, J = 10.7, 3.2 Hz, 1H), 3.49 (ddd, J = 6.7, 5.4, 1.1 Hz, 1H), 3.39 (t, J = 5.0 Hz, 2H), 1.99 (s, 3H).

^{1}H NMR (300 MHz, D$_2$O): δ = 4.44 (d, J = 8.4 Hz, 1H), 4.01 – 3.89 (m, 1H), 3.87 (d, J = 2.7 Hz, 1H), 3.82 (d, J = 8.2 Hz, 1H), 3.73 (t, J = 6.0 Hz, 2H), 3.70 – 3.59 (m, 11H), 3.44 (t, J = 4.8 Hz, 2H), 1.98 (s, 3H).

^{13}C NMR (75 MHz, D$_2$O): δ = 174.6, 101.5, 75.0, 71.0, 69.6, 69.6, 69.5, 69.1, 68.8, 67.76, 60.9, 52.3, 50.1, 22.2.

HR-MS (ESI): m/z calcd. for C$_{14}$H$_{26}$N$_4$O$_8{}^+$ [M+H]$^+$: 379.1823, found: 379.1816.

10.2.5.   Mannose derivatives

10.2.5.1.   Synthesis of 1, 2, 3, 4, 6-penta-*O*-acetyl-ᴅ-mannopyranoside (14)

Mannose acetate **14** was provided by Dr. D. M. M. Jaradat, Department of Organic Chemistry, Freie Universität Berlin.[25]

10.2.5.2.   Synthesis of 1-azido-2, 3, 4, 6-tetra-*O*-acetyl-α-ᴅ-mannopyranoside (M1Ac)

Mannose azide **M1Ac** was prepared from mannose acetate **14** by Dr. D. M. M. Jaradat, Department of Organic Chemistry, Freie Universität Berlin.[25]

10.2.5.3.   Synthesis of 1-azido-α-ᴅ-mannopyranoside (M1)

The protected mannose azide **M1Ac** (500 mg, 1.34 mmol) was deacetylated under Zemplén conditions according to the general procedure. After evaporation of solvents,

275 mg (1.34 mmol, quant.) of the title compound were obtained. Analytical data were in accordance with the literature.[7]

^{1}H NMR (300 MHz, CD$_3$OD): δ = 5.36 (d, J = 1.9 Hz, 1H), 3.86 (dt, J = 12.0, 2.0 Hz, 1H), 3.73 (ddd, J = 12.7, 2.1 Hz, 2H), 3.67 – 3.57 (m, 3H).

^{13}C NMR (75 MHz, CD$_3$OD): δ = 90.5, 75.5, 70.4, 70.3, 66.5, 61.2.

HR-MS (ESI): m/z calcd. for C$_6$H$_{11}$N$_3$O$_5$Na$^+$ [M+Na]$^+$:228.0591, found: 228.0588.

10.2.5.4. Synthesis of 1-azidoethyl-2, 3, 4, 6-tetra-*O*-acetyl-α-ᴅ-mannopyranoside (M2Ac)

To a solution of **14** (1.95 g, 5.0 mmol) in anhydrous CH$_2$Cl$_2$ (20 mL) was added 2-azidoethanol (653 mg, 7.5 mmol) and the mixture was stirred under argon at 0 °C for 1 hour. Subsequently BF$_3$·OEt$_2$ (3.55 g, 25.0 mmol) was added drop wise at 0 °C and stirred for 24 hours at room temperature. The reaction was quenched with excess triethylamine at 0 °C, diluted with CH$_2$Cl$_2$, and washed consecutively with aqueous sat. NaHCO$_3$, water and brine. The organic layer was then dried over MgSO$_4$, filtered and concentrated under reduced pressure. The residue was then purified by column chromatography on silica gel (*n*-hexanes/EtOAc, gradient 75:25 → 50:50, Rf≈0.68 at 40:60), yielding carbohydrate azide **M2Ac** in a yield of 43% (905 mg, 2.17 mmol). The obtained spectra was in accordance with the literature.[7], [26]

^{1}H NMR (300 MHz, CDCl$_3$): δ = 5.39 – 5.24 (m, 3H), 4.87 (d, J = 1.7 Hz, 1H), 4.29 (dd, J = 12.2, 5.3 Hz, 1H), 4.12 (dd, J = 12.3, 2.3 Hz, 1H), 4.04 (ddd, J = 9.1, 5.3, 2.4 Hz, 1H), 3.87 (ddd, J = 10.7, 6.8, 4.1 Hz, 1H), 3.72 – 3.63 (m, 1H), 3.55 – 3.39 (m, 2H), 2.16 (s, 3H), 2.10 (s, 3H), 2.05 (s, 3H), 1.99 (s, 3H).

^{13}C NMR (75 MHz, CDCl$_3$): δ = 170.8, 170.2, 170.0, 169.9, 97.9, 69.5, 69.0, 67.2, 66.1, 62.6, 50.5, 21.0, 20.9, 20.9, 20.8.

10.2.5.5. Synthesis of 1-azidoethyl-α-ᴅ-mannopyranoside (M2)

The acetylated mannose azide **M1Ac** (500 mg, 1.34 mmol) was deprotected under Zemplén conditions according to the general procedure. After evaporation of solvents, 275 mg (1.34 mmol, quant.) of the title compound were obtained. Analytical data were in accordance with the literature.[7], [26]

^{1}H NMR (300 MHz, D$_2$O): = δ 4.89 (d, J = 1.8 Hz, 1H), 3.95 (dd, J = 3.6, 1.7 Hz, 1H), 3.89 (ddd, J = 12.1, 5.9, 2.3 Hz, 2H), 3.81 (dd, J = 9.4, 3.3 Hz, 1H), 3.80 – 3.66 (m, 2H), 3.65 (td, J = 6.3, 3.2 Hz, 2H), 3.50 (tdt, J = 10.1, 7.3, 3.5 Hz, 2H).

^{13}C NMR (75 MHz, D$_2$O) δ 99.6, 72.7, 70.2, 69.8, 66.5, 66.2, 60.8, 50.0.

HR-MS (ESI): m/z calcd. for C$_8$H$_{15}$N$_3$O$_6$Na$^+$ [M+Na]$^+$: 272.0853, found: 272.0848.

10.2.5.6. Synthesis of 2-[2-(2-azidoethoxy)ethoxy]ethyl-2, 3, 4, 6-tetra-*O*-acetyl-α-ᴅ-mannopyranoside (M3Ac)

Mannose acetate **14** (976 mg, 2.5 mmol) was dissolved in anhydrous CH_2Cl_2 (5 mL) containing activated, powdered molecular sieves 4Å (1 g). Then 2-[2-(2-azidoethoxy) ethoxy]ethanol (6.8 mmol, 1 mL) was added and the mixture was stirred for 1 hour under argon. Subsequently TMSOTf (4.5 mmol, 1 mL) was added drop wise and the reaction was conducted for 5 hours at room temperature. After dilution with CH_2Cl_2, the mixture was filtered and washed with aqueous sat. $NaHCO_3$ and brine. The organic layer was dried over $MgSO_4$, filtered and concentrated under reduced pressure. After subjection to column chromatography on silica gel (n-hexanes/EtOAc 60:40, then 40:60), **M3Ac** was obtained in a yield of 32% (410 mg, 0.81 mmol). The obtained spectra were in accordance with the literature.[27]

^{1}H NMR (300 MHz, $CDCl_3$): δ = 5.37 (dd, J = 10.1, 3.3 Hz, 1H), 5.34 – 5.26 (m, 2H), 4.88 (d, J = 1.6 Hz, 1H), 4.30 (dd, J = 12.4, 5.1 Hz, 1H), 4.14 – 4.04 (m, 2H), 3.89 – 3.77 (m, 1H), 3.76 – 3.63 (m, 9H), 3.41 (t, J = 5.1 Hz, 2H), 2.17 (s, 3H), 2.11 (s, 3H), 2.05 (s, 3H), 2.00 (s, 3H).

^{13}C NMR (75 MHz, $CDCl_3$): δ = 170.8, 170.2, 170.0, 169.9, 97.8, 70.9, 70.8, 70.2, 70.2, 69.7, 69.2, 68.5, 67.5, 66.3, 62.5, 50.8, 21.1, 20.9, 20.9, 20.8.

10.2.5.7. Synthesis of 2-[2-(2-azidoethoxy)ethoxy]ethyl-α-ᴅ-mannopyranoside (M3)

According to the general procedure, mannose azide **M3Ac** (410 mg, 0.811 mmol) was deacetylated under Zemplén conditions. After evaporation of the solvents, 270 mg (0.8 mmol, 98%) of the title compound were obtained. The analytical data was in accordance with the literature.[16]

^{1}H NMR (300 MHz, CD_3OD): δ = 4.79 (d, J = 2.4 Hz, 1H), 3.91 – 3.76 (m, 3H), 3.74 – 3.63 (m, 14H), 3.59 (dd, J = 8.1, 2.7 Hz, 3H), 3.38 (t, J = 4.8 Hz, 2H).

^{13}C NMR (75 MHz, CD_3OD): δ = 101.7, 74.6, 72.5, 72.1, 71.7, 71.5, 71.4, 71.2, 68.6, 67.8, 62.9, 51.7.

HR-MS (ESI): m/z calcd. for $C_{12}H_{23}N_3O_8Na^+$ [M+Na]$^+$: 360.1377, found: 360.1367.

10.2.6.   N-Acetylneuraminic acid derivatives

10.2.6.1.   Synthesis of methyl 5-acetamido-3,5-dideoxy-D-glycero-D-galacto-2-nonulopyranosonate (15)

N-Acetyl neuraminic acid (5 mmol, 1.546 g) was suspended in MeOH (125 mL) and DOWEX® HCR-W2 ion-exchange resin (2 g) was added. The resulting mixture was stirred at room temperature for 96 hours. Since TLC showed full conversion, the reaction was filtered and concentrated in vacuo, yielding the title product quantatively. Analytically pure samples were obtained by recrystallization from MeOH/Et$_2$O. The obtained spectra were in accordance with the literature.[28]

^{1}H NMR (400 MHz, D$_2$O): δ = 3.96 – 3.88 (m, 2H), 3.79 (d, J = 10.3 Hz, 1H), 3.69 (dd, J = 11.3, 3.2, 1H), 3.69 (s, 3H), 3.59 (ddd, J =9.1, 6.3, 2.6, 1H), 3.47 (dd, J = 11.8, 6.3 Hz, 1H), 3.40 (dd, J = 9.3, 1.2 Hz, 1H), 3.20 (s, 1H), 2.17 (dd, J = 13.1, 4.8 Hz, 1H), 1.91 (s, 3H), 1.77 (dd, J = 13.1, 11.5 Hz, 1H).

^{13}C NMR (101 MHz, D$_2$O): δ = 174.9, 171.5, 95.4, 70.4, 70.2, 68.3, 66.7, 63.2, 53.6, 52.1, 38.7, 22.1.

10.2.6.2.   Synthesis of methyl (5-acetamido-4,7,8,9-tetra-O-acetyl-2,3,5-tridesoxy-2-chloro-D-glycero-β-D-galacto-non-2-ulopyranoside)onate (D1)

Methyl ester **15** (5.9 mmol, 1.9 g) was mixed with freshly distilled acetyl chloride (40 mL). While cooled with an ice-bath, glacial acetic acid (5 mL) and methanol (1.5 mL) for activation were added. The reaction vessel was sealed and stirred under argon for 2 days at room temperature. Subsequently the mixture was evaporated to dryness, yielding sialyl donor **D1** quantitatively. The obtained spectra were in accordance with the literature.[28]

^{1}H-NMR (400 MHz, CDCl$_3$): δ = 5.47 (dd, J = 7.0, 2.4 Hz, 1H), 5.39 (ddd, J = 11.2, 10.2, 4.8 Hz, 1H), 5.17 (ddd, J = 6.9, 5.8, 2.7 Hz, 1H), 4.42 (dd, J = 12.5, 2.7 Hz, 1H), 4.38 – 4.31 (m, 1H), 4.20 (q, J = 10.4 Hz, 1H), 4.06 (dd, J = 12.5, 5.8 Hz, 1H), 3.87 (s, 3H), 2.78 (dd, J = 13.9, 4.8 Hz, 1H, H3$_{eq}$), 2.27 (dd, J =13.9, 11.2 Hz, 1H, H3$_{ax}$), 2.12 (s, 3H), 2.07 (s, 3H), 2.04 (s, 3H), 2.04 (s, 3H), 1.90 (s, 3H).

^{13}C NMR (101 MHz, CDCl$_3$): δ = 171.1, 170.7, 170.5, 170.0, 169.8, 165.7, 96.7, 74.0, 70.0, 68.8, 66.9, 62.2, 53.9, 48.8, 40.7, 32.2, 23.2, 21.0, 20.9, 20.9, 20.8.

10.2.6.3.   Synthesis of Methyl [phenyl 5-acetylamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-2-thio-D-glycero-D-galacto-non-2-ulopyranoside]onate (D2)

Previously prepared sialyl chloride **D1** ($\approx$ 3.00 mmol, 1.53 g) was used without purification and dissolved in EtOAc (20 mL). K_2CO_3 (1 M, 50 mL) and TBAHS (3.00 mmol, 680 mg) were added. The two-phase mixture was then vigorously stirred at room temperature. Subsequently thiophenol (12 mmol, 1.2 mL) was added and the reaction was stirred for 18 hours at ambient temperature. It was then diluted with EtOAc (40 mL) and the phases were separated. The organic layer was consecutively washed with aq. sat. $NaHCO_3$ (2x 25 mL) and brine (3x 25 mL). Drying over $MgSO_4$, filtration and concentration in vacuo afforded the crude product. Further purification by column chromatography on silica gel (EtOAc/cyclo hexane gradient, 25:75 $\rightarrow$ 100:0) afforded the title compound in a yield of 78% (2.25 mmol, 1.31 g). The obtained spectra were in accordance with the literature.[29]

^{1}H NMR (400 MHz, CDCl$_3$): δ = 7.54 – 7.49 (m, 2H), 7.43 – 7.31 (m, 3H), 5.32 – 5.24 (m, 2H), 5.16 (d, *J* = 9.9 Hz, 1H), 4.85 (ddd, *J* = 11.7, 10.1, 4.7 Hz, 1H), 4.39 (dd, *J* = 12.4, 2.5 Hz, 1H), 4.20 (dd, *J* = 12.4, 5.1 Hz, 1H), 3.98 (q, *J* = 10.2 Hz, 1H), 3.89 (dd, *J* = 10.7, 1.8 Hz, 1H), 3.57 (s, 3H), 2.82 (dd, *J* = 12.8, 4.7 Hz, 1H), 2.14 (s, 3H), 2.07 – 2.01 (3x s, 9H), 1.86 (s, 3H).

10.2.6.4.   Synthesis of methyl (phenyl 5-acetamido-7,8,9-tri-*O*-acetyl-5-*N*,4-*O*-carbonyl-3,5-dideoxy-2-thio-D-glycero-β-D-galacto-non-2-ulopyranoside)onate (D3)

Sialyl donor **D2** (0.476 mmol, 278 mg) was dissolved in MeOH (4.5 mL) and methanesulfonic acid (93 µL, 3 equiv.) was added at room temperature. The mixture was then stirred for 24 hours at 60 °C, and refluxed for 4 hours. After cooling to room temperature, the mixture was quenched with excess triethylamine (15 mL). The whole solution was then concentrated in vacuo. To the residue was added NaHCO$_3$ (2.38 mmol, 200 mg, 5 equiv.), MeCN (3 mL) and water (1.5 mL). The mixture was then cooled to 0 °C at which point 4-nitrophenyl chloroformate (1.191 mmol, 240 mg, 2.5 equiv.) in MeCN (1.5 mL) was added through a dropping funnel for 1 hour. The reaction was then further stirred at 0 °C for 4 hours. After extraction with EtOAc (3x 5 mL), the combined organic layers were washed with brine, dried over $MgSO_4$, filtered and concentrated under reduced pressure. Column chromatography of the residue (EtOAc/MeOH gradient, 100:0 $\rightarrow$ 90:10) afforded intermediate methyl (phenyl 5-*N*,4-*O*-carbonyl-3,5-dideoxy-2-thio-D-glycero-β-D-galacto-non-2-ulopyranoside)onate (**16**) in acceptable purity.

Sialoside **16** was then dissolved in pyridine (1 mL) and cooled to 0 °C. Ac_2O (1.5 mL) was added and the mixture was stirred over night and was allowed to reach room temperature. Co-evaporation of the reaction mixture with toluene yielded the crude product (~70 mg, 0.133 mmol) which was dried over night on the high-vac. It was then dissolved in anh. CH_2Cl_2 (1 mL). DIPEA (1.3 mmol, 168 mg) was added and the mixture was stirred for 30 minutes at room temperature and then cooled to 0 °C. AcOCl (1.064 mmol, 77 μL) was added slowly and the reaction was stirred for 1 hour at 0 °C. The mixture was then poured into aq. sat. $NaHCO_3$ and extracted two times with CH_2Cl_2. The combined organic layers were dried over $MgSO_4$, filtered and concentrated in vacuo. The residue was then subjected to column chromatography (EtOAc/cyclohexane gradient, 25:75 → 66:34), yielding the title compound in 28% (0.133 mmol, 70 mg). The obtained spectra were in accordance with the literature.[30]

16: ^1H NMR (400 MHz, CD_3OD): δ = 7.61 – 7.50 (m, 1H), 7.46 – 7.32 (m, 3H), 3.84 – 3.76 (m, 3H), 3.67 – 3.59 (m, 2H), 3.58 (s, 3H), 3.54 (d, J = 2.1 Hz, 1H), 3.11 (dd, J = 11.6, 3.8 Hz, 1H), 2.21 (dd, J = 12.6, 11.7 Hz, 1H).

D3: ^1H NMR (400 MHz, $CDCl_3$): δ = 7.58 – 7.49 (m, 2H), 7.49 – 7.28 (m, 3H), 5.54 (dd, J = 6.3, 1.6 Hz, 1H), 5.33 (dt, J = 6.5, 6.5, 2.9 Hz, 1H), 4.41 (dd, J = 12.2, 2.9 Hz, 1H), 4.33 (dd, J = 9.5, 1.6 Hz, 1H), 4.18 (dd, J = 12.2, 6.6 Hz, 1H), 3.94 (ddd, J = 12.7, 11.2, 3.6 Hz, 1H), 3.60 (d, J = 9.4 Hz, 1H), 3.58 (s, 3H), 3.09 (dd, J = 12.1, 3.7 Hz, 1H), 2.44 (s, 3H), 2.15 (s, 3H), 2.10 (t, J = 12.5, 1H), 2.06 (s, 6H).

10.2.6.5. Synthesis of methyl (5-acetamido-2,4,7,8,9-penta-*O*-acetyl-3,5-dideoxy-ᴅ-glycero-ᴅ-galacto-non-2-ulopyranoside)onate (17)

Methyl ester **15** (1.547 mmol, 500 mg) was dissolved in pyridine (6 mL) and acetic anhydride (7 mL) was added at 0 °C. The mixture was stirred for 4 hours at 0 °C and was then stirred for additional 96 hours at ambient temperature. Subsequently the solution was evaporated to dryness, using toluene as an auxiliary. An anomeric mixture was obtained, which could be separated by column chromatography (toluene/MeOH 40:1), yielding the 57.1 mg (0.10 mmol, 10%) of the α- and 79% (1.23 mmol, 656.0 mg) of the β-anomer. The reported spectra for the β-anomer were accordance with the literature. [31]

^1H NMR (400 MHz, $CDCl_3$): δ = 5.70 (d, J = 8.8 Hz, 1H), 5.54 (d, J = 9.5 Hz, 1H), 5.37 (dd, J =3.3, 1.1 Hz, 1H), 5.09 (ddd, J = 11.3, 3.4, 0.9 Hz, 1H), 4.44 (dt, J = 11.3, 8.3 Hz, 1H), 4.17 (dd, J = 11.3, 6.7 Hz, 1H), 4.11 (dd, J = 11.3, 6.4 Hz, 1H), 4.02 (td, J = 6.5, 6.5, 1.2 Hz, 1H), 2.17 (s, 3H), 2.13 (s, 3H), 2.04 (s, 3H), 2.02 (s, 3H), 1.94 (s, 3H).

^{13}C NMR (75 MHz, $CDCl_3$): δ = 171.2, 170.8, 170.5, 170.5, 170.4, 168.4, 166.5, 97.6, 95.5, 73.0, 71.6, 68.4, 67.9, 62.3, 53.4, 49.4, 36.0, 23.3, 21.1, 21.0, 21.0, 20.9, 20.9

10.2.6.6. Synthesis of methyl (1-adamantanyl 5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-2-thio-D-glycero-β-D-galacto-non-2-ulopyranoside)onate (18)

N-Acetyl neuraminic acid acetate **17** (4.50 mmol, 2.40 g) was dissolved in anhydrous CH_2Cl_2 (25 mL). To this 1-adamantanthiol (4.95 mmol, 833 mg, 1.1 equiv.) was added and the mixture was cooled to 0 °C. Subsequently $BF_3 \cdot OEt_2$ (10.8 mmol, 1.4 mL, 2.4 equiv.) was added drop wise. The reaction was stirred for 20 hours under argon and was allowed to reach room temperature. The mixture was diluted to 500 mL with CH_2Cl_2 and consecutively washed with aqueous sat. $NaHCO_3$ and brine. The organic layer was dried over Na_2SO_4, filtered and concentrated under reduced pressure. Column chromatography on silica gel (EtOAc/cyclohexane/propan-2-ol 50:50:2, Rf≈0.09) afforded the title compound in 82% yield (2.38 g, 3.71 mmol). The obtained spectra were in accordance with the literature.[32]

^{1}H NMR (400 MHz,CDCl$_3$): δ= 5.54 (d, *J* = 10.3 Hz, 1H), 5.46 (dd, *J* = 2.8, 1.4 Hz, 1H), 5.26 (ddd, *J* = 11.9, 10.4, 4.7 Hz, 1H), 5.14 (dt, *J* = 9.1, 1.3 Hz, 1H), 5.01 (dd, *J* = 12.3, 1.7 Hz, 1H), 4.54 (dd, *J* = 10.5, 2.7 Hz, 1H), 4.19 (dd, *J* = 12.3, 9.0 Hz, 1H), 4.13 – 3.98 (m, 1H), 3.81 (s, 3H), 2.52 (dd, *J* = 13.5, 4.7 Hz, 1H), 2.11 (s, 3H), 2.06 (s, 3H), 2.06 – 1.98 (m, 1H), 2.02 – 1.92 (m, 12H), 1.85 (s, broad, 6H), 1.64 (s, broad, 6H).

^{13}C NMR (75 MHz, CDCl$_3$): δ = 171.3, 170.8, 170.3, 170.1, 170.1, 169.8, 86.1, 73.9, 72.7, 69.2, 68.9, 63.3, 52.7, 50.4, 49.5, 43.4, 39.9, 35.8, 29.7, 23.1, 21.0, 20.9, 20.7, 20.6.

10.2.6.7. Synthesis of methyl (1-adamantanyl 5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-5-*N*-(1,1-dimethylethoxy)carbonyl-2-thio-D-glycero-β-D-galacto-non-2-ulopyranoside)onate (19)

Di-*tert*-butyl dicarbonate (5.5 mmol, 1.2 g) and DMAP (0.22 mmol, 27 mg) were added to a solution of **17** (0.55 mmol, 353 mg) in anhydrous THF (20 mL) at room temperature. The mixture was stirred at 60 °C over night under argon. After cooling to room temperature the reaction was concentrated under reduced pressure and the residue was purified by silica gel column chromatography (c-hexane/EtOAc 2:1). The title compound was obtained in a yield of 87% (0.48 mmol, 355 mg). The obtained spectra were in accordance with the literature.[32]

^{1}H NMR (400 MHz, CDCl$_3$): δ = 5.65 (td, *J* = 11.0, 11.0, 4.6 Hz, 1H), 5.37 (dd, *J* = 10.1, 2.3 Hz, 1H), 5.31 (s, broad, 1H), 5.13 (d, *J* = 9.1 Hz, 1H), 4.98 (dd, *J* = 12.5, 1.7 Hz, 1H), 4.76 (t, *J* = 10.5, 10.5 Hz, 1H), 4.23 (dd, J = 12.5, 8.9 Hz, 1H), 3.83 (s, 3H), 2.63 (dd, J = 13.5, 4.7 Hz, 1H), 2.33 (s, 3H), 2.07 (s, 3H), 2.05 (s, 3H), 2.00 (s, 3H), 2.00 – 1.94 (m, 6H), 1.92 (s, 3H), 1.91 – 1.84 (m, 3H), 1.71 – 1.63 (m, 15H).

10.2.6.8. Synthesis of methyl (1-adamantanyl 5-*N*,4-*O*-carbonyl-3,5-dideoxy-2-thio-D-glycero-β-D-galacto-non-2-ulopyranoside)onate (20) *via* 18

Sialoside **18** (2.93 mmol, 1.88 g) was dissolved in dry MeOH (30 mL). After methanesulfonic acid (9 mmol, 0.65 mL) was added drop-wise, the resulting mixture was refluxed for 19 hours under argon atmosphere. When the reaction was cooled to ambient temperature, it was quenched with DIPEA (3 mL) and concentrated under reduced pressure. The crude residue was subsequently dissolved in acetonitrile (15 mL) and water (20 mL) containing $NaHCO_3$ (24 mmol, 2 g) and cooled to 0 °C. Then 4-nitrophenyl chloroformate (7.5 mmol, 1.62 g) in acetonitrile (5 mL) was added drop-wise. The temperature of the mixture was maintained around 0°C and the reaction was carried out for additional 3 hours. After completion of the reaction, it was extracted with EtOAc and the combined organic layers were washed with brine and dried over Na_2SO_4. After concentration of the filtrate under reduced pressure, the residue was purified by column chromatography on silica gel (EtOAc/MeOH gradient, 100:0 → 24:1), yielding the title compound with a yield of 57% (1.66 mmol, 760 mg). The obtained analytical data were in accordance with the literature.[32]

^{1}H NMR (400 MHz, CD$_3$OD): δ = 4.55 (ddd, *J* = 12.6, 11.3, 3.8 Hz, 1H), 4.48 (dd, *J* = 10.0, 2.1 Hz, 1H), 3.84 (s, 3H), 3.81 (d, *J* = 8.0 Hz, 1H), 3.75 – 3.65 (m, 2H), 3.54 (dd, *J* = 8.7, 2.1 Hz, 1H), 3.52 (dd, *J* = 11.4, 10.0 Hz, 1H), 2.66 (dd, *J* = 12.6, 3.8 Hz, 1H), 2.24 (t, *J* = 12.6, 1H), 2.05 – 1.90 (m, 9H), 1.69 (s, broad, 6H).

^{13}C NMR (75 MHz, CD$_3$OD): δ = 173.3, 162.4, 87.5, 79.2, 75.3, 72.3, 71.6, 64.9, 59.8, 53.7, 48.2, 44.6, 40.5, 37.1, 31.4.

10.2.6.9. Synthesis of methyl (1-adamantanyl 5-*N*,4-*O*-carbonyl-3,5-dideoxy-2-thio-D-glycero-β-D-galacto-non-2-ulopyranoside)onate (20) *via* 19

Sialoside **19** (0.46 mmol, 341 mg) was dissolved in anhydrous MeOH (2 mL) and a catalytic amount of sodium methoxide solution (5.55 M NaOMe in MeOH) was added. The deacetylation was stirred at room temperature for approximately 3 hours. The reaction was then diluted with MeOH and neutralized with DOWEX® HCR-W2 ion-exchange resin. Filtration and concentration under reduced pressure afforded a residue to which was added trifluoroacetic acid (2 mL). After 1 hour at room temperature, the mixture was dried with an argon stream. The crude and $NaHCO_3$ (2.75 mmol, 231 mg, 6 equiv.) were dissolved in MeCN (4 mL) and water (8 mL) and cooled to 0 °C. Subsequently a solution of 4-nitrophenyl chloroformate (1.38 mmol, 147 mg) in MeCN (4 mL) was added slowly via a dropping funnel. After stirring for approximately 4 hours at 0 °C,

the mixture was further diluted with MeCN/water and extracted with ethyl acetate. The combined organic layers were washed with brine and dried over Na_2SO_4. After filtration and concentration under reduced pressure, the obtained residue was purified by column chromatography (EtOAc/MeOH gradient, 100:0 → 24:1). Unfortunately the product yield was only 19% (41 mg, 0.09 mmol). The obtained spectra were in accordance with the literature.[32]

^{1}H NMR (400 MHz, CD_3OD): δ = 4.56 (ddd, J = 12.6, 11.3, 3.8 Hz, 1H), 4.50 (dd, J = 9.9, 2.1 Hz, 1H), 3.85 (s, 3H), 3.83 (d, J = 8.2 Hz, 1H), 3.77 – 3.67 (m, 2H), 3.58 – 3.50 (m, 2H), 2.68 (dd, J = 12.6, 3.9 Hz, 1H), 2.26 (t, J = 12.6, 12.6 Hz, 1H), 2.06 – 1.95 (m, 9H), 1.71 (s, broad, 6H).

10.2.6.10. Synthesis of methyl (1-adamantanyl 5-acetamido-*N*,4-*O*-carbonyl-7,8,9-tetra-*O*-acetyl-3,5-dideoxy-2-thio-D-glycero-β-D-galacto-non-2-ulopyranoside)onate (D4)

Oxazolidinon **20** (0.29 mmol, 133 mg) was dissolved in pyridine (2.6 mL) and acetic anhydride (3.2 mL) was added slowly under cooling. The mixture was stirred for approximately 72 hours at ambient temperature before it was diluted with toluene and concentrated under reduced pressure. Subsequently the crude reaction intermediate was dissolved in anhydrous CH_2Cl_2 (5 mL) and DIPEA (2.49 mmol, 0.5 mL) was added. The mixture was cooled to 0 °C and acetyl chloride (2.52 mmol, 0.18 mL) was added slowly. The reaction was stirred for 2 hours and was allowed to reach room temperature. Subsequently the mixture was diluted with CH_2Cl_2 (30 mL) and was quenched by pouring into aqueous sat. $NaHCO_3$ (50 mL). The layers were separated and the aqueous phase was extracted with CH_2Cl_2. The combined organic fraction was consecutively washed with brine, dried over Na_2SO_4, filtered, and concentrated under reduced pressure. Column chromatography on silica gel (EtOAc/*n*-hexanes 3:2) afforded 117 mg (0.19 mmol, 66%) of donor **D4** as a white foam. The obtained spectra were in accordance with the literature.[32]

^{1}H NMR (300 MHz, $CDCl_3$): δ = 5.71 (t, J = 2.9 Hz, 1H), 5.31 (dt, J = 8.0, 2.4 Hz, 1H), 4.82 – 4.62 (m, 3H), 4.15 (dd, J = 12.1, 8.0 Hz, 1H), 3.84 (s, 3H), 3.67 (dd, J = 11.3, 9.2 Hz, 1H), 2.80 (dd, J = 12.8, 3.6, 1H), 2.48 (s, 3H), 2.17 (t, J = 12.8, 1H), 2.12 (s, 3H), 2.11 (s, 3H), 2.02 (s, 3H), 2.04 – 1.95 (m, broad, 3H), 1.94 – 1.83 (m, broad, 3H), 1.67 (m, broad, 6H).

^{13}C NMR (75 MHz, $CDCl_3$): δ = 172.4, 171.2, 170.7, 169.8, 169.5, 153.8, 85.8, 75.3, 74.6, 73.5, 72.5, 63.4, 60.4, 53.1, 51.5, 43.7, 38.8, 36.1, 30.0, 24.9, 21.3, 20.9, 20.9.

10.2.6.11. Synthesis of methyl (2-azidoethyl 5-acetamido-*N*,4-*O*-carbonyl-7,8,9-tetra-*O*-acetyl-3,5-dideoxy-ᴅ-glycero-α-ᴅ-galacto-non-2-ulopyranoside)onate (21) *via* D4

2-Azidoethanol (0.40 mmol, 31 mg) and sialyl donor **D4** (0.24 mmol, 148 mg) were stirred in a mixture of dry CH_2Cl_2–acetonitrile (2.2/2.6 mL), containing activated, powdered molecular sieves (4 Å), under argon at room temperature for 1 hour. Subsequently the solution was cooled to –78 °C and *N*-iodo succinimide (0.57 mmol, 128 mg) was added. The mixture was treated with triflic acid (0.24 mmol, 33.7 μL) and stirred for approximately 1 hour. Then the glycosylation reaction was quenched under cooling with excess of triethyl amine. After dilution with CH_2Cl_2 and filtration over Celite®, the filtrate was washed with 20% aqueous $Na_2S_2O_3$ and dried over Na_2SO_4. Subsequently the solvents were removed under reduced pressure and the resultant residue was purified by column chromatography on silica gel (*n*-hexanes/tetrahydrofuran, 5:1). Finally, 108 mg (0.20 mmol, 83%) of sialoside **22** could be obtained.

^1H NMR (300 MHz, $CDCl_3$): δ = 5.61 (dd, *J* = 8.8, 1.6 Hz, 1H, 7-H), 5.44 (ddd, *J* = 8.8, 6.8, 2.9 Hz, 1H, 8-H), 4.65 (dd, *J* = 9.5, 1.6 Hz, 1H, 6-H), 4.38 (dd, *J* = 12.2, 2.9 Hz, 1H, 9a-H), 4.11 – 3.97 (m, 3H, 4-H, 9b-H, 1′-H), 3.82 (s, 3H, CO_2Me), 3.73 (dd, *J* = 11.3, 9.5 Hz, 1H, 5-H), 3.56 (dt, *J* = 10.5, 4.9 Hz, 1H, 1′-H), 3.36 (t, *J* = 4.9 Hz, 2H, 2′-H), 2.88 (dd, *J* = 12.2, 3.6 Hz, 1H, 3_{eq}-H), 2.49 (s, 3H, NAc), 2.14 (s, 3H), 2.13 (s, 3H), 2.13 – 2.12 (m, 1H, 3_{ax}-H), 2.03 (s, 3H).

^{13}C NMR (75 MHz, $CDCl_3$): δ = 172.1 (CO_2CH_3), 170.8, 170.4, 170.2 (3x *C*O), 168.8 (N*C*OCH_3), 153.8 (N*C*OO), 99.1 (C-2), 75.8 (C-6), 75.0 (C-4), 71.7 (C-7), 69.0 (C-8), 65.0 (C-1′), 63.3 (C-9), 59.2 (C-5), 53.0 (CO_2*C*H_3), 50.7 (C-2′), 36.7 (C-3), 24.8 (NCO*C*H_3), 21.3, 21.0, 20.9 (3x O*C*H_3).

IR (ATR): ṽ = 2960 (CH_2), 2360 (OCO), 2110 (N_3), 1740 (COOR), 1650 (NAc), 1460 (CH_2), 1370 (OAc), 1030 (CO) cm^{-1}.

HR-MS (ESI-TOF): m/z calcd. for $C_{21}H_{28}N_4O_{13}Na^+$ $[M+Na]^+$ 567.1527, found 567.1545.

10.2.6.12. Synthesis of methyl (2-(2-(2-azidoethoxy)ethoxy)ethyl 5-acetamido-*N*,4-*O*-carbonyl-7,8,9-tetra-*O*-acetyl-3,5-dideoxy-ᴅ-glycero-α-ᴅ-galacto-non-2-ulopyranoside)onate (22) *via* D4

To dry CH_2Cl_2–acetonitrile (2.7/3.2 mL), containing activated, powdered molecular sieves (4 Å), was added sialyl donor **D4** (0.3 mmol, 187 mg) and 2-[2-(2-azidoethoxy)ethoxy]ethanol (0.48 mmol, 84 mg). The mixture was stirred under argon at room temperature for 1 hour and then cooled to –78 °C. Then the cooled solution was consecutively treated with *N*-iodo succinimide (0.72 mmol, 161 mg) and triflic acid (0.30 mmol, 42.3 μL). After approximately 1 hour the reaction was quenched under

cooling with excess of triethyl amine. The mixture was then diluted with CH_2Cl_2 and filtered over Celite®. Subsequently the filtrate was washed with 20% aqueous $Na_2S_2O_3$ and dried over Na_2SO_4. The solvents were removed under reduced pressure and the resultant residue was purified by column chromatography on silica gel (*n*-hexanes/ tetrahydrofuran, 4:1). Finally, 108 mg (0.20 mmol, 83%) of sialoside **23** could be obtained.

^{1}H NMR (300 MHz, CDCl$_3$): δ = 5.58 (d, *J* = 7.8 Hz, 1H, 7-H), 5.40 (td, *J* = 7.5, 2.8 Hz, 1H), 4.60 (d, *J* = 9.1 Hz, 1H), 4.37 (dd, *J* = 12.3, 3.1 Hz, 1H, 9a-H), 4.05 (dd, *J* = 12.2, 6.6 Hz, 1H, 9b-H), 4.01 – 3.88 (m, 2H, 4-H, 5´-H), 3.80 (s, 3H, CO$_2$Me), 3.75 – 3.62 (m, 9H, 5-H, 1´–4´-H), 3.50 (ddd, *J* = 10.2, 5.4, 3.9 Hz, 1H, 5´-H), 3.38 (t, *J* = 5.1 Hz, 2H, 6´-H), 2.89 (dd, *J* = 2.2, 3.6 Hz, 1H, 3$_{eq}$-H), 2.48 (s, 3H, NAc), 2.16 (s, 3H), 2.13 (s, 3H, OAc), 2.12 (s, 3H, OAc), 2.08 (t, *J* = 3.4 Hz, 1H, 3$_{ax}$-H), 2.03 (s, 3H, OAc).

^{13}C NMR (75 MHz, CDCl$_3$): δ = 172.0 (*C*O$_2$CH$_3$), 170.6, 170.2, 170.0 (3x s, *C*OCH$_3$), 168.7 (N*C*OCH$_3$), 153.7 (N*C*OO), 99.3 (C-2), 75.5 (C-6), 75.1 (C-4), 71.8 (C-7), 70.7, 70.5, 70.2, 70.1(C-1´-4´), 69.1 (C-8), 65.1 (C-5´), 63.1 (C-9), 59.0 (C-5), 53.0 (CO$_2$*C*H$_3$), 50.7 (C-2´), 36.5 (C-3), 24.7 (NCO*C*H$_3$), 21.1, 20.9, 20.8 (3x CO*C*H$_3$).

IR (ATR): ṽ = 2930 (CH$_2$), 2360 (OCO), 2110 (N$_3$), 1740 (COOR), 1640 (NAc), 1440 (CH$_2$), 1370 (OAc), 1030 (CO) cm^{-1}.

HR-MS (ESI-TOF): m/z calcd. for C$_{25}$H$_{36}$N$_4$O$_{15}$Na$^+$ [M+Na]$^+$ 655.2068, found 655.2069.

10.2.6.13. Synthesis of methyl (2-azidoethyl 5-acetamido-3,5-dideoxy-ᴅ-glycero-α-ᴅ-galacto-non-2-ulopyranoside)onate (23)

Sialoside **22** (0.13 mmol, 72 mg) was deacetylated according to the general procedure, yielding methyl ester **24** quantitatively as a pale yellow foam.

^{1}H NMR (300 MHz, CD$_3$OD): δ = 4.13 – 3.91 (m, 2H), 3.90 – 3.78 (m, 3H), 3.77 – 3.63 (m, 5H), 3.62 – 3.45 (m, 4H), 2.86 (dd, *J* = 12.8, 3.5 Hz, 1H, 3$_{eq}$-H), 2.01 (s, 3H), 1.60 (t, *J* = 11.2 Hz, 1H, 3$_{ax}$-H).

^{13}C NMR (151 MHz, CD$_3$OD): δ = 175.5, 174.1, 101.9, 74.5, 72.96, 70.4, 69.5, 64.3, 64.2, 54.2, 52.1, 42.6, 22.6.

IR (ATR): ṽ = 3360 (OH/NH), 2950 (CH$_2$), 2360 (OCO), 1730 (COOR), 1560 (NHAc), 1460 (CH$_2$), 1020 (CO) cm^{-1}.

10.2.6.14. Synthesis of methyl (2-(2-(2-azidoethoxy)ethoxy) ethyl 5-acetamido-3,5-dideoxy-ᴅ-glycero-α-ᴅ-galacto-non-2-ulopyranoside)onate (24)

Sialyl azide **23** (0.31 mmol, 195 mg) was deprotected under Zemplén conditions according to the general procedure. After evaporation of the solvents, sialoside **24** was

obtained quantitatively (0.31 mmol, 152 mg).

^{1}H NMR (300 MHz, D$_2$O): δ = 3.99 (dd, J = 10.1, 3.6 Hz, 1H), 3.91 (s, 3H), 3.87 (m, 2H), 3.83 – 3.69 (m, 11H), 3.69 – 3.56 (m, 3H), 3.54 (t, J = 4.8 Hz, 2H), 2.76 (dd, J = 13.1, 4.8 Hz, 1H, 3$_{eq}$-H), 2.07 (s, 3H), 1.86 (t, J = 12.3 Hz, 1H, 3$_{ax}$-H).

^{13}C NMR (151 MHz, D$_2$O): δ = 174.5, 169.4, 98.5, 72.4, 70.1, 69.2, 69.1, 69.0, 68.9, 68.8, 67.8, 66.7, 63.0, 62.6, 52.9, 51.3, 49.7, 38.7, 21.6.

IR (ATR): ṽ = 3340 (OH/NH), 2880 (CH$_2$), 2360 (OCO), 1730 (COOR), 1560 (NHAc), 1460 (CH$_2$), 1030 (CO).

HR-MS (ESI-TOF): m/z calcd. for C$_{18}$H$_{32}$N$_4$O$_{11}$Na$^+$[M+Na]$^+$ 503.1960, found 503.1942.

10.2.6.15. Synthesis of 2-azidoethyl 5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranoside)onate (S2)

Sialoside **24** (0.06 mmol, 23 mg) was saponified according to the general procedure, yielding the title compound quantitatively (0.06 mmol, 24 mg).

^{1}H NMR (300 MHz, D$_2$O): δ = 4.09 – 3.89 (m, 1H), 3.88 – 3.60 (m, 7H), 3.60 – 3.48 (m, 2H), 3.45 (t, J = 4.9 Hz, 2H), 2.71 (dd, J = 12.5, 4.5 Hz, 1H, 3$_{eq}$-H), 2.00 (s, 3H), 1.71 (t, J = 12.1 Hz, 1H, 3$_{ax}$-H).

^{13}C NMR (75 MHz, D$_2$O): δ = 174.9, 174.6, 95.7, 70.1, 70.0, 68.2, 66.9, 63.0, 52.0, 51.6, 50.2, 38.9, 21.9.

HR-MS (ESI): m/z calcd. for C$_{13}$H$_{22}$N$_4$O$_9$Na$^+$ [M+Na]$^+$ 401.1279, found: 401.1234.

10.2.6.16. Synthesis of (2-(2-(2-azidoethoxy)ethoxy)ethyl 5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranoside)onate (S3)

Sialoside **25** (0.06 mmol, 30 mg) was saponified according to the general procedure, yielding the title compound quantitatively (0.06mmol, 28 mg).

^{1}H NMR (300 MHz, D$_2$O): δ = 3.87 – 3.77 (m, 1H), 3.72 (dd, J = 10.1, 5.7 Hz, 3H), 3.67 – 3.56 (m, 11H), 3.53 (d, J = 4.2 Hz, 2H), 3.46 (dd, J = 18.6, 8.1 Hz, 1H), 3.37 (t, J = 4.8 Hz, 2H), 2.58 (dd, J = 12.7, 4.4 Hz, 1H, 3$_{eq}$-H), 1.90 (s, 3H), 1.65 (t, J = 12.1 Hz, 1H, 3$_{ax}$-H).

^{13}C NMR (75 MHz, D$_2$O): δ = 174.7, 173.4, 95.2, 71.5, 70.2, 70.0, 69.4, 69.3, 69.0, 68.1, 66.6, 63.0, 60.2, 51.9, 50.0, 38.7, 21.9.

HR-MS (ESI-TOF): m/z calcd. for C$_{17}$H$_{30}$N$_4$O$_{11}$Na$^+$ [M+Na]$^+$ 489.1803, found: 489.1799.

10.3. Chemical Synthesis

10.3.1. Synthesis of azido-linkers

10.3.1.1. Synthesis of 2-azidoethanol

To 2-bromoethanol (14.62 g, 117 mmol) was added sodium azide (13.0 g, 200 mmol, 1.7 equiv.) and tetra-N-butyl ammonium bromide (1.0 g, 3.1 mmol, 0.03 equiv.).Then diethyl ether (2 mL) was added and the reaction mixture was stirred for 24 hours at 75 °C. After cooling to room temperature it was diluted further with diethyl ether. Insoluble solids were filtered off and the solvent was removed under reduced pressure. 2-Azidoethanol was then obtained in a yield of 98% (10.0 g, 115 mmol). Analytical data were in accordance with the literature.[33]

^{1}H NMR (300 MHz, CDCl$_3$): δ = 3.78 (t, J = 4.9 Hz, 3H), 3.44 (t, J = 5.0 Hz, 3H), 2.00 (s, 1H). ^{13}C NMR (75 MHz, CDCl$_3$): δ = 61.6, 53.7.

10.3.1.2. Synthesis of 2-[2-(2-azidoethoxy)ethoxy]ethanol

A mixture of 2-[2-(2-chloroethoxy)ethoxy]ethanol (1.69 g, 10.0 mmol) and NaN$_3$ (1.30 g, 20.0 mmol, 2 equiv.) in water (10 mL) was heated to 75 °C and stirred for 72 hours. Subsequently the solvent was evaporated and the residue was take up in diethyl ether. The solid was removed by filtration over a cotton plug with Celite®. Evaporation of the solvent afforded 2-[2-(2-azidoethoxy)ethoxy]ethanol as a clear oil (1.49 g, 8.5 mmol, 85%). The obtained spectra were in accordance with the literature.[34]

^{1}H NMR (300 MHz, CDCl$_3$): δ = 3.77 – 3.71 (m, 2H), 3.70 – 3.66 (m, 6H), 3.64 – 3.59 (m, 2H), 3.40 (t, J = 5.0 Hz, 2H). ^{13}C NMR (75 MHz, CDCl$_3$): δ = 72.6, 70.8, 70.6, 70.2, 61.9, 50.8.

10.3.2. Synthesis of CuAAC ligands

10.3.2.1. Synthesis of TBTA

Tripropargylamine (20 mmol, 2.62 g) was dissolved in acetonitrile (30 mL) and 2,6-lutidine (20 mmol, 2.14 g, 1.0 equiv.) as well as previously synthesized benzyl azide (90 mmol, 11.98 g, 4.5 equiv.) were added consecutively. The mixture was put on an ice-bath and Cu(MeCN)$_4$PF$_6$ (0.78 mmol, 290 mg, 0.04 equiv.) was added to start the CuAAC. After 3 days a white precipitate formed, which was filtered off and washed with cold acetonitrile. It was then dried under vacuum overnight, yielding the title compound

in 85% (17 mol, 9.02 g). The obtained spectra were in accordance with the literature.[35]

^{1}H-NMR (400 MHz, CDCl$_3$): δ = 7.59 (s, 3H), 7.30-7.25 (m, 9H), 7.19-7.16 (m, 6H), 5.42 (s, 6H), 3.62 (s, 6H)

^{13}C-NMR (101 MHz, CDCl$_3$): δ = 134.8, 129.2, 128.8, 128.1, 123.9, 54.2.

10.3.2.2. Synthesis of 3-bromo-propyl acetate intermediate for THPTA

A mixture of acetic anhydride (72 mmol, 7.4 g, 1.0 eq.) and triethylamine (72 mmol, 7.3 g, 1.0 equiv.) was prepared and added to a solution of 3-bromopropan-1-ol (72 mmol, 10 g) in CH$_2$Cl$_2$ (60 mL). After 4 hours at room temperature, aqueous sat. NaHCO$_3$ (30 mL) was added. The phases were separated and the organic layer was consecutively washed with aqueous sat. NaHCO$_3$ (30 mL) and brine (2x 30 mL). The solvents were removed under reduced pressure, yielding 3-bromo-propyl acetate (8.7 g, 48 mmol, 68%) as an oil.[36]

^{1}H-NMR (400 MHz, CDCl$_3$): δ = 4.18 (t, J = 6.1 Hz, 2H), 3.45 (t, J = 6.5 Hz, 2H), 2.16 (t, J = 6.3 Hz, 2H), 2.04 (s, 3H).

10.3.2.3. Synthesis 3-azido-propyl acetate intermediate for THPTA

It was then dissolved in H$_2$O (35 mL) and stirred with NaN$_3$ (100 mmol, 6.5 g) at 90–95 °C overnight. The reaction mixture was then cooled to room temperature and extracted with CH$_2$Cl$_2$ (3x25 mL). The combined organic layers were washed with water and brine, dried over MgSO$_4$, filtered, and concentrated under reduced pressure. 3-azidopropyl acetate could be obtained in a yield of 79% (38 mmol, 5.45 g).[36]

^{1}H NMR (400 MHz, CDCl$_3$): δ = 4.15 (t, J = 6.2 Hz, 2H), 3.39 (t, J = 6.7 Hz, 2H), 2.06 (s, 3H), 1.90 (quintet, J = 6.4 Hz, 2H).

10.3.2.4. Synthesis of THPTA

3-Azidopropyl acetate (4 mmol, 573 mg), tripropargylamine (1 mmol, 132 mg) and a catalytic amount of Cu(MeCN)$_4$PF$_6$ were dissolved in dry tetrahydrofurane (10 mL) and refluxed overnight under argon atmosphere. Subsequently the solvent was removed and the reaction crude was purified by column chromatography (chloroform/methanol, gradient 100:0 → 94:6), yielding 370 mg (0.66 mmol, 66%) of the desired intermediate, tris [(3 hydroxypropyltriazol-4-yl)methyl]amine acetate. This intermediate (280 mg, 0.50 mmol) was then dissolved in 2 м NH$_3$ in MeOH (7 mL) and stirred at 40 °C for 16 hours. The reaction mixture was then cooled to room temperature, followed by concentration *in vacuo*. The white solid was then dispersed in MeCN, filtered and dried under vacuum, yielding tris [(3 hydroxypropyltriazol-4-yl)methyl]amine (141 mg,

0.33 mmol, 65%) in high purity. Spectra of the final product were in accordance with the literature.[36]

^{1}H-NMR (400 MHz, DMSO-d_6): δ = 8.03 (s, 3H), 4.67 (t, J = 5.0 Hz, 3H), 4.41 (t, J = 7.1 Hz, 6H), 3.62 (s, 6H), 3.40 (q, J = 5.6 Hz, 6H), 1.96 (quintet, J = 6.7 Hz, 6H).

^{13}C-NMR (101 MHz, DMSO-d_6): δ = 143.4, 124.0, 57.5, 47.1, 46.6, 33.0.

10.3.3. Miscellaneous Compounds

10.3.3.1. Synthesis of 4-pentynoic acid NHS ester

4-Pentynoic acid (5.10 mmol, 500 mg) was dissolved in anhydrous CH_2Cl_2 (25 mL). Consecutively N-hydroxy succinimide (5.85 mmol, 670 mg) and EDC·HCl (10.20 mmol, 1.955 g) were added and the reaction mixture was stirred overnight at room temperature. Subsequently it was washed with 2.5% aqueous $NaHSO_4$, dried over Na_2SO_4 and filtered. After evaporation to dryness, 734 mg (3.76 mmol, 74%) of the 4-pentynoic acid NHS ester were obtained. The analytical data were in accordance with the literature.[37]

^{1}H NMR (300 MHz, CDCl$_3$): δ = 2.88 (t, J = 7.4 Hz, 2H), 2.84 (s, 4H), 2.61 (td, J = 7.7, 7.1, 2.7 Hz, 2H), 2.05 (t, J = 2.6 Hz, 1H).

^{13}C NMR (75 MHz, CDCl$_3$): δ = 169.0, 167.2, 81.0, 70.2, 30.4, 25.7, 14.2.

10.4. Protein chemistry and analytics

ψ-barstar variants were provided by Dr. Lars Merkel (Department of Chemistry, Technische Universität Berlin) and Nina Bohlke (Department of Chemistry, Technische Universität Berlin).

Non-canonical amino acid bearing green fluorescent proteins (GFP) were obtained from Nina Bohlke (Department of Chemistry, Technische Universität Berlin).

Protein concentrations were determined using the molar extinction coefficient (ε), caclulated with ExPasy's ProtParam tool,[38] and the Beert-Lambert law (A = ε·c·l, whereas A is the absorbance, c is the concentration and l is the path length). The absorbance was measured with an UV-spektrophotometer or a NanoDrop device.

10.4.1. Optimized protocol for CuAAC of ψ-barstar

For CuAAC between given carbohydrate azides and alkyne-bearing ψ-b*, the following reagents and reactants were prepared as stock solutions in ultra pure water: carbohydrate azides **1–6** (250 mM), CuSO$_4$ (20 mM), THPTA, (50 mM), aminoguanidine hydrochloride (100 mM), and sodium ascorbate (100 mM). The stocks were stored at

–20 °C, except for sodium ascorbate solution, which was always made freshly prior to the reactions. ψ-b* proteins were used in concentrations of 0.07–0.2 mg/mL in PBS (pH 8.0, 10% glycerol).

Typically, the reactions were carried out by subsequently adding the indicated volumes of the stock solutions to an Eppendorf tube in the following order:

1. 150 µL of ψ-b*
2. 32.5 µL of Dulbecco's PBS, Ca^{2+} and Mg^{2+} free, pH 7.3
3. 10 µL of carbohydrate azide solution
4. 7.5 µL of a Cu-THPTA solution, premixed from 2.5 µL $CuSO_4$ and 5 µL THPTA stock.
5. 25 µL of aminoguanidine
6. 25 µL of sodium ascorbate

Then the reaction vessel was equipped with a magnetic stir-bar, mixed well and sealed. The CuAAC was carried out at 4 °C for approximately 64 hours before the reaction was quenched with Dulbecco's PBS containing EDTA (2 mM, pH 7.4). Subsequently the samples were dialyzed for 22 hours against Dulbecco's PBS containing EDTA (2 mM, pH 7.4) and Dulbecco's PBS. Modified ψ-b* proteins were then concentrated with spin-filter tubes (MWCO 3 kDa). The reaction outcome was assessed by MALDI-TOF-MS.[3]

10.4.2. Optimized protocol for CuAAC of green fluorescent protein

For the modification of alkyne-bearing GFPs with carbohydrate azides by CuAAC, the following reagents and reactants were used as stock solutions in ultra pure water: carbohydrate azides **G1**–**G3** (250 mM) and **M1**–**M3** (250 mM), $CuSO_4$ (20 mM), THPTA, (50 mM), aminoguanidine hydrochloride (100 mM), and sodium ascorbate (100 mM). Sodium ascorbate solution was always prepared freshly prior to the reactions. The other stocks were stored at –20 °C. GFPs were rebuffered to PBS (50 mM Na_2HPO_4, 150 mM NaCl, pH 7.8) and concentrated with spin-filter tubes (MWCO 10 kDa) to approximately 100 µM.

The CuAAC reactions were carried out by consecutively adding the indicated volumes of the stock solutions to an Eppendorf tube:

1. 15 µL of alkyne-GFP
2. 85 µL of PBS (50 mM Na_2HPO_4, 150 mM NaCl, pH 7.8)
3. 2.5 µL of carbohydrate azide solution
4. 7.5 µL of a Cu-THPTA solution, premixed from 2.5 µL $CuSO_4$ and 5 µL THPTA stock.
5. 20 µL of aminoguanidine
6. 20 µL of sodium ascorbate

Subsequently the reaction vessel was sealed and vortexed. The CuAAC was shaken at 15 °C and 1000 rpm for approximately 4 hours. Then the reaction was rebuffered to

PBS (50 mM Na_2HPO_4, 150 mM NaCl, pH 7.8) and concentrated with spin-filter tubes (MWCO 10 kDa). The reaction mixture was brought back to a volume of 100 µL with PBS and steps 3.–6. were repeated. The CuAAC was then carried out for 18 hours at 15 °C, followed by rebuffering and concentration with spin-filter tubes (MWCO 10 kDa). The volume of the concentrate was brought back to 100 µL and steps 3.–6. were repeated again. After 4 hours at 15 °C the CuAAC was subjected to an ÄKTA FPLC (GE Healthcare, Freiburg, Germany) equipped with a HiTrap Desalting column and rebuffered to PBS (50 mM Na_2HPO_4, 150 mM NaCl, pH 7.8). Relevant fractions were pooled and subsequently concentrated with spin-filter tubes (MWCO 10 kDa). The reaction outcome was assessed by MALDI-TOF-MS.

10.5. Surface Plasmon Resonance (SPR)

(Performed together with Dr. Figen Beceren-Braun, Christian Kühne, and Dr. Jens Dernedde, Clinical Chemistry and Pathobiochemistry, Charité, Berlin)

10.5.1.   Competitive binding assay for functionalized ψ-barstar[3]

SPR measurements were carried out at 25 °C on a Biacore X instrument (GE Healthcare, Freiburg, Germany). Binding probe was the commercial available Thomsen-Friedenreich (TF) antigen linked to a biotinylated polyacrylamide (PAA) carrier: Galβ1→3GalNAcα-$(CH_2)_3$-PAA-$(CH_2)_6$-biotin (Lectinity Holdings Inc., Moscow, Russia). The probe had a molecular weight of ~30 kDa and contained 20 mol% TF antigen and 5 mol% biotin. Via the strong biotin-streptavidin interaction the probe was coupled to 671 resonance units (RU) on a streptavidin functionalized sensor chip SA (GE Healthcare, Freiburg, Germany). Before immobilization, the sensor chip was conditioned with three consecutive 1 min injections of 1 M NaCl and 50 mM NaOH. The TF antigen probe was diluted to 4.2 µg/ml in HBS-EP buffer (GE Healthcare) and passed over one lane of the chip surface, the second lane remained untreated and served as a reference. After the immobilization procedure, the chip surface was equilibrated with three consecutive 1 min injections of running buffer, containing 20 mM HEPES, pH 7.4; 150 mM NaCl and 1 mM $CaCl_2$. A 35 µL sample volume was injected over both lanes, whereas the final binding signals were obtained by subtraction of data from the free reference lane. The association phase was set to 105 s followed by a 180 s dissociation phase. The response values were calculated by subtraction of the report point at the beginning of the sample injections (0 s) from the report point at the end of the dissociation phase (285 s). Regeneration of the chip was performed after each run, with a 60 s flow of 4 M $MgCl_2$.

Binding analyses were performed with running buffer at a flow rate of 20 µl/min. To measure peanut agglutinin (PNA) interaction to immobilized TF antigen, a 800 nM solution of PNA (Axxora GmbH, Loerrach, Germany) was used and the resulting RU value was set to 100% binding (positive control). For all competitive measurements PNA was preincubated for 18 min at room temperature with barstar-conjugates at a final concentration of 10 µM protein before injection. The resulting RU values were calculated as X% binding of the control and converted to % inhibition.

10.5.2.   Competitive binding assay for functionalized polyglycerols[39]

SPR measurements were carried out as indicated in **10.5.1**, p. 148.

10.5.3.   Direct binding assay for functionalized polyglycerols[40]

Binding of peanut agglutinin (PNA) to sugar probes was recorded on a Biacore X instrument (GE Healthcare, Freiburg, Germany). Covalent amine coupling of PNA (Axxora GmbH, Loerrach, Germany) was performed on a carboxymethyl-dextran (CM5) sensor chip according to the manufacturer's instructions by EDC/NHS chemistry. Protein was dissolved in 10 mM sodium acetate, pH 4.5 and injected immediately after surface activation. Unreacted carboxyl groups were quenched by the addition of 1 M methanolamine. Surface was washed with 100 µL 4 M $MgCl_2$ and equilibrated with 3 consecutive injections of 100 µl running buffer (20 mM HEPES, pH 7.4; 150 mM NaCl; 1 mM $CaCl_2$; sterile filtrated; degassed; RT). Coupling yielded in a resonance shift of 3277 resonance units (RU) for the ligand lane and 9 RU for the reference lane which was treated equally except PNA injection. A 35 µL sample volume was injected over both lanes, whereas the final binding signals were obtained by subtraction of data from the reference lane. The association phase was set to 105 s followed by a 180 s dissociation phase. Regeneration of the chip was performed after each run, with 4 M $MgCl_2$ at flow rates of 100 µL/min for 60 s. Obtained data curves were aligned to zero on both axes, using sample inject start as time point zero and the average the of data before sample inject start for zero of the response difference scale (manufacturers software, BIAevaluation).

Binding analyses were performed with running buffer at a flow rate of 20 µl/min. Analytes PG 26 and as positive control the Thomsen-Friedenreich (TF) antigen linked to a polyacrylamide (PAA) carrier: Galβ1→3GalNAcβ-$(CH_2)_3$-PAA (Lectinity Holdings Inc., Moscow, Russia) were injected at final concentrations of 0.1 µM, 1 µM and 10 µM. To rule out possible influences by the polyglycerol core, an unglycosylated polyglycerol

with an azide functionalization of 30% and an approximated mass of 13 kDa (**PG-N₃**) was measured at respective concentrations. Measurements were performed at room temperature (25 °C).

10.6. References

1. H. E. Gottlieb, V. Kotlyar, A. Nudelman, *J. Org. Chem.* **1997**, *62*, 7512–7515.

2. T. Wenzel, K. Sparbier, T. Mieruch, M. Kostrzewa, *Rapid Commun. Mass Spectrom.* **2006**, *20*, 785–789.

3. L. M. Artner, L. Merkel, N. Bohlke, F. Beceren-Braun, C. Weise, J. Dernedde, N. Budisa, C. P. R.Hackenberger, *Chem. Commun.* **2011**, *48*, 522–524.

4. R. Dulbecco, M. Vogt, *J. Exp. Med.* **1954**, *99*, 167–182.

5. G. Zemplén, E. Pacsu, *Ber. dtsch. Chem. Ges. A/B* **1929**, *62*, 1613–1614.

6. R. Kumar, P. Tiwari, P. R. Maulik, A. K. Misra, *Eur. J. Org. Chem.* **2006**, 74–79.

7. J. Geng, J. Lindqvist, G. Mantovani, G. Chen, C. T. Sayers, G. J. Clarkson, D. M. Haddleton, *QSAR Comb. Sci.* **2007**, *26*, 1220–1228.

8. V. Vicente, J. Martin, J. Jiménez-Barbero, J. L. Chiara, C. Vicent, *Chem. Eur. J.* **2004**, *10*, 4240–4251.

9. T. Tanaka, H. Nagai, M. Noguchi, A. Kobayashi, S.-I. Shoda, *Chem. Commun.* **2009**, 3378–3379.

10. F. Fazio, M. C. Bryan, O. Blixt, J. C. Paulson, C.-H. Wong, *J. Am. Chem. Soc.* **2002**, *124*, 14397–14402.

11. C. Bouillon, A. Meyer, S. Vidal, A. Jochum, Y. Chevolot, J.-P. Cloarec, J.-P. Praly, J.-J. Vasseur, F. Morvan, *J. Org. Chem.* **2006**, *71*, 4700–4702.

12. J.-F. Nierengarten, J. Iehl, V. Oerthel, M. Holler, B. M. Illescas, A. Muñoz, N. Martín, J. Rojo, M. Sánchez-Navarro, S. Cecioni, S. Vidal, K. Buffet, M. Durka, S. P. Vincent, *Chem. Commun.* **2010**, *46*, 3860–3862.

13. C. S. Hudson, J. M. Johnson, *J. Am. Chem. Soc.* **1915**, *37*, 1270–1275.

14. R. Šardzík, G. T. Noble, M. J. Weissenborn, A. Martin, S. J. Webb, S. L. Flitsch, *Beilstein J. Org. Chem.* **2010**, *6*, 699–703.

15. Y. Koshi, E. Nakata, M. Miyagawa, S. Tsukiji, T. Ogawa, I. Hamachi, *J. Am. Chem. Soc.* **2008**, *130*, 245–251.

16. Y. Zhang, S. Luo, Y. Tang, L. Yu, K.-Y. Hou, J.-P. Cheng, X. Zeng, P. G. Wang, *Anal. Chem.* **2006**, *78*, 2001–2008.

17. J. Wu, Z. Guo, *Bioconjugate Chem.* **2006**, *17*, 1537–1544.

18. D. Stokmaier, O. Khorev, B. Cutting, R. Born, D. Ricklin, T. O. Ernst, F. Böni, K. Schwingruber, M. Gentner, M. Wittwer, M. Gentner, M. Wittwer, M. Spreafico, A. Vedani, S. Rabbani, O. Schwardt, B. Ernst, *Bioorg. Med. Chem.* **2009**, *17*, 7254–7264.

19. A. W. Harrison, J. F. Fisher, D. M. Guido, S. J. Couch, J. A. Lawson, D. M. Sutter, M. V. Williams, G. L. DeGraaf, J. E. Rogers, D. T. Pals, D.W. DuCharme, *Bioorg. Med. Chem.* **1994**, *2*, 1339–1361.

20. L. Szilágyi, Z. Györgydeák, *Carbohydr. Res.* **1985**, *143*, 21-41.

21. D. J. Lee, K. Mandal, P. W. R. Harris, M. A. Brimble, S. B. H. Kent, *Org. Lett.* **2009**, *11*, 5270–5273.

22. G. J. L. Bernardes, R. Kikkeri, M. Maglinao, P. Laurino, M. Collot, S. Y. Hong, B. Lepenies, P. H. Seeberger, *Org. Biomol. Chem.* **2010**, *8*, 4987–4996.

23. A. Hasegawa, T. Terada, H. Ogawa, M. Kiso, *J. Carbohydr. Chem.* **1992**, *11*, 319–331.

24. P. C. N. Rensen, S. H. van Leeuwen, L. A. J. Sliedregt, T. J. C. van Berkel, E. A. L. Biessen, *J. Med. Chem.* **2004**, *47*, 5798–5808.

25. D. M. M. Jaradat, *PhD thesis*, Freie Universität Berlin, Berlin, **2010**.

26. M. Dowlut, D. G. Hall, O. Hindsgaul, *J. Org. Chem.* **2005**, *70*, 9809–9813.

27. Y. Chevolot, C. Bouillon, S. Vidal, F. Morvan, A. Meyer, J.-P. Cloarec, A. Jochum, J.-P. Praly, J.-J. Vasseur, E. Souteyrand, *Angew. Chem. Int. Ed.* **2007**, *46*, 2398–2402.

28. R. Roy, C. A. Laferrière, *Can. J. Chem.* **1990**, *68*, 2045–2054.

29. A. Marra, P. Sinaÿ, *Carbohydr. Res.* **1989**, *187*, 35–42.

30. D. Crich, W. Li, *J. Org. Chem.* **2007**, *72*, 2387–2391.

31. A. Marra, P. Sinaÿ, *Carbohydr. Res.* **1989**, *190*, 317–322.

32. D. Crich, W. Li, *J. Org. Chem.* **2007**, *72*, 7794–7797.

33. H. R. Pfaendler, V. Weimar, *Synthesis* **1996**, 1345–1349.

34. S. P. Amaral, M. Fernandez-Villamarin, J. Correa, R. Riguera, E. Fernandez-Megia, *Org. Lett.* **2011**, *13*, 4522–4525.

35. T. R. Chan, R. Hilgraf, K. B. Sharpless, V. V. Fokin, *Org. Lett.* **2004**, *6*, 2853–2855.

36. V. Hong, S. I. Presolski, C. Ma, M. G. Finn, *Angew. Chem. Int. Ed.* **2009**, *48*, 9879–9883.

37. M. Slater, M. Snauko, F. Svec, J. M. J. Fréchet, *Anal. Chem.* **2006**, *78*, 4969–4975.

38. http://web.expasy.org/protparam/

39. V. Böhrsch, T. Mathew, M. Zieringer, M. R. J. Vallée, L. M. Artner, J. Dernedde, R. Haag, C. P. R. Hackenberger, *Org. Biomol. Chem.* **2012**, *10*, 6211–6216.

40. M. R. J. Vallée, L. M. Artner, J. Dernedde, C. P. R. Hackenberger, *Angew. Chem. Int. Ed.* **2013**, *52*, 9504–9508.

11. List of Publications

The thesis resulted so far in the following publications:

- <u>L. M. Artner</u>, L. Merkel, N. Bohlke, F. Beceren-Braun, C. Weise, J. Dernedde, N. Budisa, C. P. R. Hackenberger, *Chem. Commun.* **2012**, *48*, 522–524; DOI: 10.1039/ C1CC16039G.

- V. Böhrsch, T. Mathew, M. Zieringer, M. R. J. Vallée, <u>L. M. Artner</u>, J. Dernedde, R. Haag, C. P. R. Hackenberger, *Org. Biomol. Chem.* **2012**, *10*, 6211–6216; DOI: 10.1039/ c2ob25207d.

- M. R. J. Vallée, <u>L. M. Artner</u>, J. Dernedde, C. P. R. Hackenberger, *Angew. Chem. Int. Ed.* **2013**, *52*, 9504–9508; DOI: 10.1002/anie.201302462, and

  M. R. J. Vallée, <u>L. M. Artner</u>, J. Dernedde, C. P. R. Hackenberger, *Angew. Chem.* **2013**, *125*, 9682–9686; DOI: 10.1002/ange.201302462.