universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Studies of the Binding of Glycosylated Peptides and Influenza Glycoproteins to Recombinantly Expressed Cyanovirin-N"

verfasst von / submitted by

Philipp Schilling BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2018 / Vienna 2018

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 862
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Chemie
Betreut von / Supervisor:	Univ.-Prof. Dr. Christian F. W. Becker
Mitbetreut von / Co-Supervisor:	Mag. Dr. Irene Maier

Table of contents

Abstract

Cyanovirin-N (CV-N), a member of the lectin family found in bacteria, fungi and plants, is active against several viruses, including HIV and influenza and shows binding to Ebola glycoprotein at nanomolar concentrations. The neutralization process is mediated by recognizing and binding to *N*-linked high-mannose glycans located on the virus surface. Furthermore, it has been demonstrated that CV-N specifically binds Man*α*(1–2)Man*α* units located at the termini of the D1/D3 arms of high-mannose oligosaccharides, thereby preventing viral entry into the cell. While these essential properties for binding recognition were identified *via* glycoarrays and most studies focus on binding HIV glycoprotein gp120, hemagglutinin (HA) and its molecular interaction regions remain less covered. The present thesis is investigating the binding affinity of CV-N mutants with varying numbers of disulfide bridges to HA and related glycopeptides.

We synthesized homogenously mannosylated peptides covering a partial sequence of the surface glycoprotein HA, thereby mimicking the direct binding partner of CV-N, by combining solid-phase peptide synthesis (SPPS) and copper(I)-catalyzed azide alkyne cycloaddition (CuAAC). A new synthesis route for azido di-mannose building blocks was developed and both azido mannose and di-mannose were linked to the peptide backbone using click chemistry. CV-N variants were recombinantly expressed to investigate the influence of disulfide bonds on binding affinity. With both binding partners in hand, affinity studies using surface plasmon resonance (SPR) were carried out. The main goal of this project was to understand the essential binding properties of CV-N's high-affinity binding sites with high-mannose ligands (HA) and low molecular weight glycopeptides.

Zusammenfassung

Cyanovirin-N (CV-N), ein Protein aus der Familie der Lektine, ist in Bakterien, Pilzen sowie Pflanzen zu finden und besitzt antivirale Eigenschaften gegenüber HIV, Influenza und bindet Ebola Glykoprotein in nanomolaren Konzentrationen. Dieser Prozess basiert auf der spezifischen Erkennung und Bindung von posttranslationalen *N*-Glykosylierungen, welche unter anderem an der Oberfläche von Viren zu finden sind. Insbesondere terminale Man*α*(1–2)Man*α*-Einheiten, welche die Enden der D1/D3 Verzweigungen von High-mannose Strukturen zieren, bindet CV-N spezifisch. Durch diesen Vorgang wird die Zellpermeabilität des Virus unterbunden. Während diese spezifischen Eigenschaften mittels Bindung an immobilisierten Zuckerstrukturen gezeigt wurden und die Bindung Cyanovirins an HIV Glykoprotein gp120 erforscht wurde, bleiben einige Fragen bezüglich der Binding an Hämagglutinin noch ungeklärt. Da die vorhandenen Disulfidbrücken einen weiteren Stabilitäts- und Affinitätsfaktor der gebildeten CV-N Dimere ausmachen können, stellt die Untersuchung der minimal benötigten Anzahl eine weitere Forschungsfrage dar.

In dieser Arbeit präsentieren wir die Synthese von homogen-mannosylierten Peptiden, welche einer Teilsequenz des Oberflächenglykoproteins Hämagglutinin entsprechen und damit einen direkten Bindungspartner Cyanovirins darstellen sollen. Dies wurde durch die Kombination von Festphasenpeptidsynthese und Kupfer(I)-katalysierter Azid-Alkin-Cycloaddition erreicht. Für die Herstellung der benötigten Kohlenhydratstrukturen wurde eine neue Syntheseroute entwickelt, welche die Isolation der Azid-modifizierten Man*α*(1–2)Man*α*-Einheit ermöglicht. Darüber hinaus wurden Varianten eines CV-N Dimers mit reduzierter Anzahl an Disulfidbücken exprimiert und isoliert. Darauffolgende Bindungsstudien wurden mittels Oberflächenplasmonenresonanzspektroskopie durchgeführt, wobei die synthetisierten Glykopeptide als Ligand und die entsprechenden Proteine als Analyte herangezogen wurden. Das Ziel dieser Arbeit war somit die essentiellen Strukturen viraler Glykoproteine für eine erfolgreiche Bindung an Lektine auf molekularer Ebene näher zu untersuchen.

Abbreviations

6xHis	polyhistidine-tag for protein purification
Ac	acetyl
AGE	agarose gel electrophoreses
APS	ammonium persulfate
Ar	argon atmosphere
aq.	aqueous
bp	base pairs
Boc	*tert*-Butyloxycarbonyl (protecting group)
CAM	cerium ammonium molybdate (stain)
CD	circular dichroism
conc.	concentrated
CuAAC	copper(I)-catalyzed azide alkyne cycloaddition
CV	column volume
CV-N	cyanovirin-N
d	doublet
DCM	dichloromethane
dd	doublet of doublets
DIC	*N,N'*-Diisopropylcarbodiimide
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDC	*N*-(3-Dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride
EDT	1,2-ethanedithiol
ER	endoplasmatic reticulum
ESI	electrospray ionization
EtOAc	ethyl acetate
equiv.	equivalent
Fmoc	fluorenylmethyloxycarbonyl (protecting group)
FT	flow through
HA	hemagglutinin
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
IPTG	isopropyl β-D-1-thiogalactopyranoside
KAN	kanamycin (A)
LB	lysogeny broth

LC-MS	liquid chromatography mass spectrometry
m	multiplet
M	molecular weight marker
MALDI	matrix-assisted laser desorption/ionization
Man	mannosyl
MW	molecular weight
NHS	*N*-Hydroxysuccinimide
NMP	*N*-Methylpyrrolidine
NMR	nuclear magnetic resonance
OD_{600}	optical density at a wavelength of 600 nm
P	pellet
PAGE	polyacrylamide gelectrophoresis
PBS	phosphate-buffered saline (phosphate buffer)
PE	petroleum ether
pET27b(+)	vector with kanamycin resistance for protein expression
PTM	posttranslational modification
r.t.	room temperature
s	singlet
SDS	sodium dodecylsulfate
SI	supporting information
SN	supernatant
SOC	super optimal broth with catabolite repression
SPPS	solid phase peptide synthesis
SPR	surface plasmon resonance (spectroscopy)
t	triplet
TAE	tris, acetic acid, EDTA buffer; for agarose gel electrophoresis
TBS	tris-buffered saline
TBST	tris-buffered saline with Tween20
TBTA	tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine
TCEP	tris(2-carboxyethyl)phosphine
TEMED	tetramethylethylenediamine
TFA	trifluoroacetic acid
TLC	thin layer chromatography
Tris	tris(hydroxymethyl)aminomethane
Vx	variant x (of $(CV\text{-}N)_2L_0$)

1 Introduction

1.1 Cyanovirin-N (CV-N)

Lectins are macromolecules capable of specifically recognizing and binding carbohydrates in a non-covalent manner and with a high degree of stereospecificity.[1] These proteins have been found in several organisms, such as algae, bacteria, fungi and plants.[2] Since cyanobacterial lectins bind directly to viral envelope gp120 and inhibit HIV entry into the host cell, they are more potent anti-HIV agents than plant lectins.[3] Their multivalent nature allows reversible binding with glycoproteins linked to the cell membrane.[4]

Among other cyanobacterial lectins, cyanovirin-N (CV-N) is considered a potential glycan-targeting agent for preventing viral diseases.[5–7] This 11 kDa protein, originated from *Nostoc ellipsosporum* (blue-green algae), mediates its binding activity with glycoproteins found on viral cell surface envelope through specific and high affinity interactions.[8] Cyanovirin-N consists of 101 amino acids with two repeat units (residues 1-50 and 51-101), having 32 % sequence identity and bearing four cysteine residues that allow the formation of two intramolecular disulfide bonds.[9] Yang *et al.* revealed that two monomers, which are symmetrically related, undergo domain swapping to assemble as a dimer.[10] This has been confirmed using X-ray crystallography, while in solution wild-type CV-N occurs mainly as a monomer. The introduction of a flexible linker or hinge, built from residues 49-54, enables monomer-dimer transition through increased flexibility of the two domains, forming either a monomer or a stabilized dimer.[10] High concentration and elevated temperature are critical conditions for domain swapping since the domain swapped dimer is a trapped folding intermediate.[11] The linker region of Pro51 and Ser52 is essential for the determination of cyanovirin-N's oligomeric properties.[12] Computational studies indicate the occurrence of a rare cis-peptide bond between these residues, allowing the formation of domain-swapped dimers under specific conditions.[13] The characteristics and activity of CV-N can vary among mutants and other homolog constructs, thus presenting an interesting topic of research. Gronenborn *et al.* solved the crystal structure of a CV-N homolog from *Cyanothece* (Cyt-CVNH) and presented two tandem repeats of the sequence with 43 % identity compared to CV-N and monomers within the asymmetric unit. Their titration experiments, based on the chemical shift changes upon sugar addition to uniformly ^{15}N-labeled Cyt-CVNH (observed with ^{1}H-^{15}N NMR), showed that both sides on the protein bind Man-9 through its D1 arm (Figure 2).[14,15]

Over the last years, CV-N complexes with various carbohydrate ligands have been crystallized, such as di-mannose Manα(1-2)Manα, hexamannosides or Man_9GlcAc_2 as well as CV-N mutants with dimers of mannose. The results confirmed two carbohydrate binding sites

in CV-N.[16–20] In addition, these primary and secondary sites showed varying affinity towards oligosaccharides.[17] While the primary binding site, which has found to be located near the linker region, forms a deep pocket capable of incorporating the disaccharide, the secondary binding site, located far from the linker region, forms a semicircular cleft that surrounds the carbohydrate moiety. As shown in Figure 1, each of the two domains A and B (symmetrical) of the monomer possesses a glycan binding site, specifically recognizing terminal Manα(1-2)Manα units (Figure 1, orange units) of D1 and D3 arms of high mannose structures (Man-8 and Man-9).[17,18,20–22]

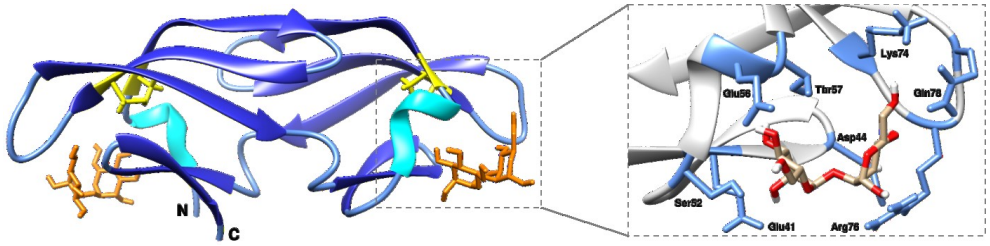


Figure 1 Solution NMR structure of complex of CV-N (blue variations) and Manα(1-2)Manα (orange); Cysteine residues highlighted in yellow; Representation of binding pocket with residues capable of forming H-bonds or water-mediated interactions with mannobiose (beige and red) highlighted in blue. Image rendered using Chimera[23], PDB file: 1IIY.[17]

Lower affinity of cyanovirin-N towards hexamannoside as compared with binding to Man_9GlcAc_2 was revealed using crystallographic studies. Direct binding of three mannose rings of Man-9, while hexamannose lacks one α(1-2) linked mannose ring at its terminal D1 arm (Figure 2), has been shown. As a result, the lectin can only bind two mannose rings of hexamannoside, resulting in lower affinity.[13,18] While domain B shows tighter binding of disaccharides (α-(1-2)-linked mannobiose) than domain A, tri-mannose with two α-(1-2)-linkages binds by equal amount.[17,24] Computational studies suggested recognition of some carbohydrate units (α-(1-2)-linked tri-mannose) with varied binding mode, even for two highly homologous CV-N domains. Further computational calculations regarding binding free energy indicated affinity with both sites through backbone contacts with core α(1-2)-linked mannobiose. The molecular interactions between disaccharide and CV-N were stabilized by two side chains in domain A (Gln6 and Thr7) and domain B (Glu56 and Thr57), respectively.[13]

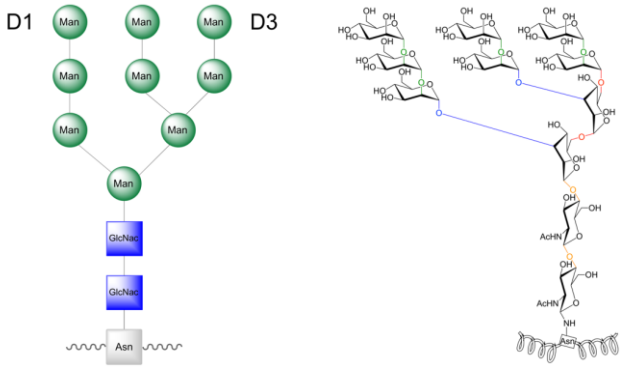


Figure 2 High mannose glycan structure with D1, D2 and D3 branches (green) as schematic (left) and structural formula representation (right); bond legend (right): green are α(1-2), blue are α(1-3), red are α(1-6) and orange are β(1-4) linkages, respectively.

Mayo *et al*. presented crystal structures of designed oligomers with enhanced HIV-1 neutralization activity. The designed oligomers (CVN_2s) contained various lengths of linkers within the hinge region, thus providing higher flexibility and a larger span between the four binding sites. The possible crosslinking of glycosylation sites on gp120 across further distances was enabled, preventing a large number of gp120 subunits from binding to CD4 which in turn increases neutralization activity. However, binding properties of the designed oligomers have not been studied in terms of actual binding affinity to glycoproteins, a prerequisite for estimating its anti-viral activity as a potential therapeutic.[20,25]

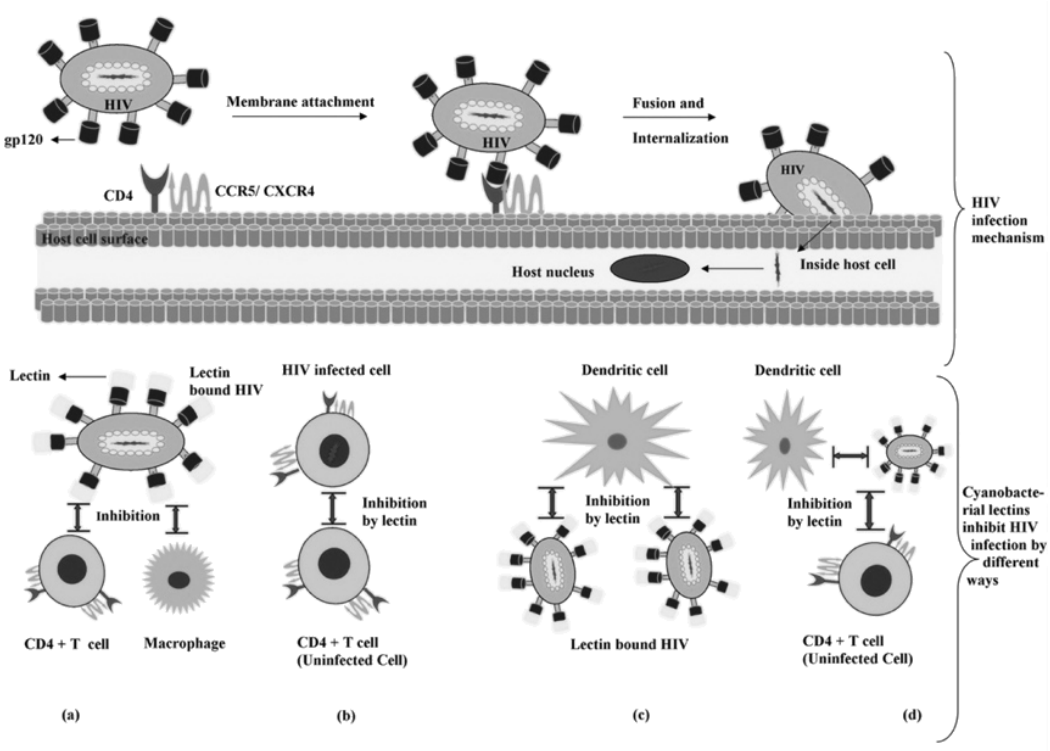


Figure 3 Illustrated HIV infection mechanism and its inhibition by cyanobacterial lectins. a) presents the inhibition of T cell and macrophage infection through lectin binding to viral glycoprotein; b) inhibition of syncytia formation; c) inhibition of HIV capture and d) transmission by dendritic cells.[20] Figure taken from [20].

In conclusion, cyanobacterial lectins possess high specificity towards α(1-2)-linked mannobiose moieties and possibly other complex mannose-containing carbohydrates. Since different epitopes have various carbohydrate recognition sites, these lectins show large potential as anti-HIV agents by interacting with gp120 of viral envelope. Furthermore, CV-N shows various positive attributes, such as anti-HIV activity at low nanomolar concentration, lack of cytotoxicity even at high concentrations and broad activity against human and primate immunodeficiency viruses. Apart from its anti-HIV activity, cyanobacterial lectins are equipped with anti-viral progression potential for also other viruses (ebola, hepatitis, influenza, herpes, etc.). Due to their unique glycan-protein interactions, lectins possess the potential for development of therapeutics through suppressing enveloped viruses and preventing various infectious diseases (Figure 3).[15,26]

1.2 Protein glycosylation

One of the most common posttranslational modifications in nature is glycosylation, taking into account that 50 % of human proteins are either *N*- or *O*-glycosylated. Phosphoglycans, glypiation (GPI-anchored), *C*- and *S*-glycosylation occur as well, however, they are much less abundant.[27] The process of glycosylation is enzymatically regulated and the addition of carbohydrates to proteins or lipids is substrate specific, thereby allowing various forms of glycoconjugates. Besides intrinsic properties, such as folding promotion, solubility and intracellular trafficking of proteins, glycosylation plays an essential role in biological recognition processes, such as signaling, cell adhesion, immune response and host-pathogen interaction.[28]

The most ubiquitous glycopeptide bond is the *β*-glycosidic linkage of *N*-acetyl glucosamine (GlcNAc) to Asn in a consensus sequence of Asn-Xaa-Ser/Thr (with Xaa ≠ Pro, Figure 4), and rarely occurring Asn-Xaa-Cys, Asn-Gly and Asn-Xaa-Val, transferred by oligosaccharyl transferase (OST). While sharing a common core sugar structure of two *N*-acetyl glucosamines, GlcNac-Asn *N*-glycans can be categorized into oligomannose, complex and hybrid types.[29–31]

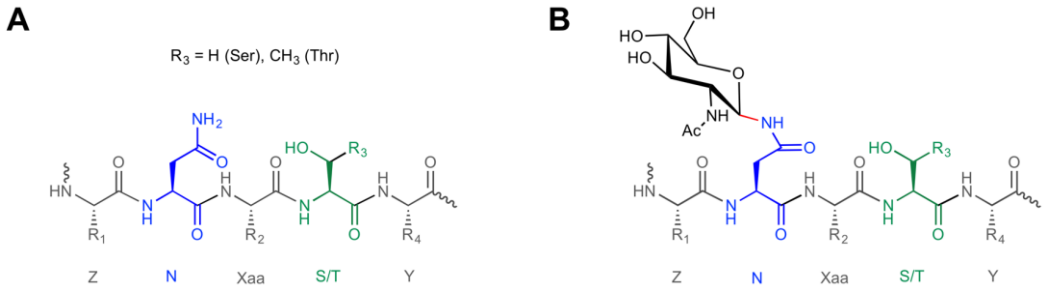


Figure 4 (A) Peptide with consensus sequence for *N*-glycosylation, R_1, R_2, R_4 = amino acid side chains, R_3 = H (Ser) or CH_3 (Thr). **(B)** Peptide with *N*-glycosylation (GlcNac). Xaa (≠ Pro), Y and Z represent amino acids.

Due to the chemical properties of carbohydrates and their heterogeneous processing within a cell, different glycoforms of proteins are produced. The glycosylation pattern is cell type dependent and can vary during cell stress. In addition, the generation of these glycoforms is dependent on the expression, substrate availability as well as the activity and specificity of the corresponding enzymes.[32–36] The role of these different glycoforms in cells and organisms requires experiments, where the homogeneously glycosylated proteins are investigated. However, structural information of complex glycoproteins is not stored within DNA/RNA and can therefore not be obtained from recombinant expression.[35] A viable option for the preparation of these glycoproteins is a synthetic/semisynthetic approach, where homogeneously glycosylated peptide fragments are linked together. The available chemical and chemoenzymatic strategies involve preparation of peptides *via* solid-phase peptide synthesis (SPPS), followed by chemoselective glycosylation prior to deprotection. Hereby, a mono-glycosylated amino acid building block, such as Asn-GlcNAc, is coupled to the

assembling peptide chain, followed by enzymatic elongation of the glycan. The Lansbury Aspartylation couples a glycosylamine with a free aspartic acid residue of a peptide by using standard peptide coupling reagents and bases, resulting in an Asn-*N*-glycoside. However, the peptide backbone needs to be protected extensively to avoid unwanted side reactions.[37,38] Glycans can be introduced by this method as well, however, due to the acid-labile *O*-glycosidic bonds of oligosaccharides, peptide deprotection and cleavage should be performed under neutral, basic or mild acidic conditions.[39] The restriction to only *N*-linked glycopeptides can be expanded to unnatural glycan linkages when using Cu(I) catalyzed azide-alkyne cycloaddition, yielding to 1,4-disubstituted triazoles.[40–42]

An impressive amount of research has already been done in this field.[43,44] Due to the enormous structural complexity of glycans, however, many fundamental questions remain open and hardly any knowledge is obtained about binding affinities of proteins to varying glycopeptides and glycoproteins because of the challenges in the development of proper binding assays as well as the availability of homogeneously glycosylated proteins.

1.3 Solid Phase Peptide Synthesis (SPPS)

In 1963, Merrifield revolutionized the chemical synthesis of peptides by introducing solid phase peptide synthesis (SPPS).[45] The inert and insoluble carrier beads or resin, used in this method, mark the starting point of a growing peptide chain by coupling selectively deprotected amino acids with each other. While ribosomal protein synthesis proceeds from *N*- to *C*-terminus, solid phase peptide synthesis is usually carried out from *C*- to *N*-terminus (Figure 5), but can be performed in reverse as well.[45]

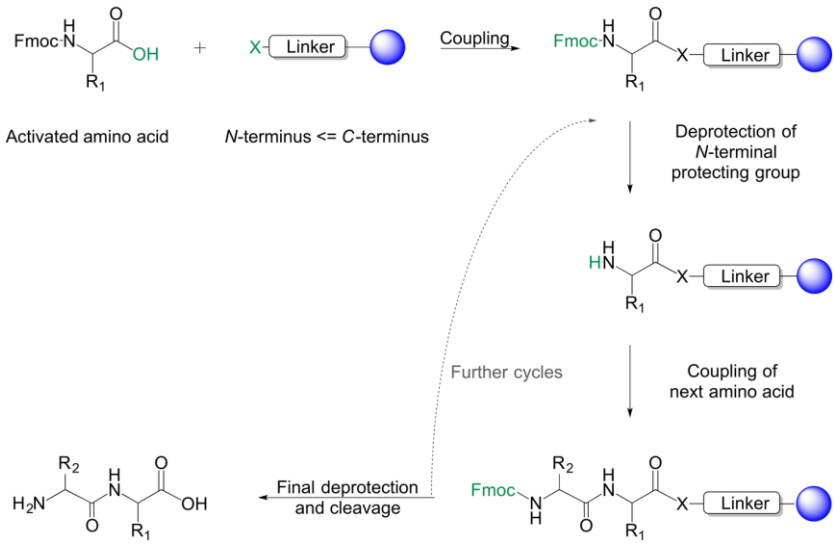


Figure 5 General procedure for solid phase peptide synthesis (SPPS), where X can be chosen according to the desired *C*-terminus (e.g. amide resin for *C*-terminal amide or, as shown in this figure, *C*-terminal carboxylic acid); R_1, R_2 = amino acid side chains.

The general procedure of SPPS is based on repetitive cycles of deprotecting the *N*-terminal amine, representing the non-resin-bound end of the synthesized peptide, and its

coupling to another activated, side-chain protected amino acid. Regarding the protection at the *N*-terminus of the assembling peptide, either base-sensitive 9-fluorenylmethyloxycarbonyl (Fmoc) or acid-sensitive tert-butyloxycarbonyl (Boc), can be utilized. After complete synthesis, the resin-bound and fully protected peptide chain can be cleaved off from the solid support, bringing the crude peptide in solution. The two approaches for protecting the *N*-terminus require different cleavage solutions. While the Boc group is deprotected by TFA during SPPS, the side chain protecting groups and the peptide-resin linkage require anhydrous hydrofluoric acid (HF) to release the crude peptide. The safety requirements present a major drawback, even though this method allows efficient peptide synthesis. The Fmoc-based approach on the other hand, is less hazardous to work with. The Fmoc protecting group can be easily removed by 20 % piperidine (in DMF) and TFA, although dangerous as well, is used for the global deprotection and cleavage from the resin, resulting in an orthogonal synthesis procedure. In addition, the use of piperidine for deprotection of Fmoc does not require neutralization. However, the risk of aggregation is increased by the lack of electrostatic repulsion.[45–47]

In conclusion, the milder cleavage conditions and straightforward handling have introduced the Fmoc-based strategy as the most common method for routine peptide synthesis. In general, the major advantages of solid-phase peptide synthesis are the specificity of sequence progression due to protected side chains and specific deprotection of the *N*-terminus, while allowing easy cleavage from the resin. In addition, the repetitive coupling cycles of this method can be easily automated. The use of insoluble resin support allows the removal of excess reagent, thereby limiting unwanted side reactions. Although the overall yield of longer peptides is relatively low from a synthetic point of view, the yields generated in each step are excellent and viewing peptides as the product of a stepwise synthesis puts the yield into perspective. Peptides and proteins with up to 90 amino acids have been successfully synthesized using SPPS. However, physical properties of the peptides or even proteins, can make synthesis challenging. Ligation methods allow the preparation of larger peptides or proteins of interest by coupling two or more fragments with each other.[45,46,48]

1.4 Copper-catalyzed azide-alkyne cycloadditions (CuAAC)

The copper(I)-catalyzed azide–alkyne cycloaddition (CuAAC) is an excellent method for the synthesis of triazoles, with tremendous impact on bioconjugation reactions. Its predecessor, the Huisgen 1,3-dipolar cycloaddition of organic azides and alkynes is an uncatalyzed and slow reaction. It requires high temperatures and produces 1,4- and 1,5-disubstituted 1,2,3-triazole regioisomers in an equimolar mixture (Figure 6, A).[49–51]

In contrast, Meldal and Sharpless independently presented a regioselective version of this method using catalytic amounts of copper(I) species.[52,53] The reaction has shown to be accelerated by up to seven orders of magnitude compared to the uncatalyzed one. Furthermore, the tremendous improvement of regioselectivity allowed the formation of 1,4-

disubstitued species without any side product (Figure 6, B).[51–54] This significant improvement of the CuAAC reaction with its unique properties and broad applications, resulted in a common high interest among the scientific community.[51]

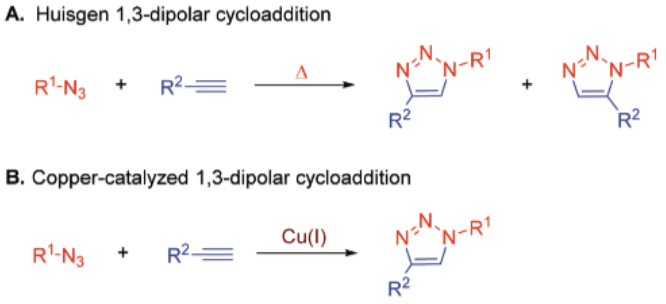


Figure 6 Comparison of thermally (A) and catalytically (B) driven synthesis of 1,2,3-triazoles *via* 1,3-dipolar cycloaddition of azides and terminal alkynes (Figure taken from [51]).

Although the 1,4-disubstituted triazole does not occur in nature, synthetic molecules bearing this species show interesting biological properties.[55] Due to its chemical inertness against hydrolysis under acidic or basic conditions, oxidation, reduction as well as its potential for hydrogen bond formation and dipole-dipole interactions, the 1,2,3-triazole is a good mimetic for the amide bond.[56–59] Unsurprisingly, these properties and the possibility of expanding the biological toolbox, has found valuable applications in various fields of research, such as biochemistry, drug development, materials sciences as well as dendrimer and polymer chemistry.[55,60–64]

The copper(I)-catalyzed azide–alkyne cycloaddition is the premier and most widely known example of ‘click chemistry’. This term, introduced by Sharpless in 2001, comprises reactions with highly efficient bond-forming and their application for the rapid synthesis of molecules with desired function.[65] In addition to allowing the preparation of desired products in excellent yields with little or no side products, these click transformations are straight forward to perform and tolerate many different conditions. The highly energetic and selective organic azides and their dipolar cycloadditions with alkynes or olefins fulfill the criteria of click chemistry.[66]

1.5 Surface Plasmon Resonance (SPR) Spectroscopy

Surface plasmon resonance (SPR) is an optical technique for the qualitative and quantitative characterization of specific binding, allowing for real-time and label-free detection of biomolecular interactions.[67]

The binding of a mobile molecule (analyte), such as antibodies or nucleic acids among others, to a ligand, immobilized on the thin metal film of a chip, changes the refractive index in the evanescent field of this reflective metal surface. When polarized light strikes this electrically conductive surface, electron charge density waves (plasmons) are generated and are reducing

the intensity of reflected light at a certain angle (resonance angle, Figure 7, left). This reduction occurs in proportion to the mass on a sensor surface.[67,70]

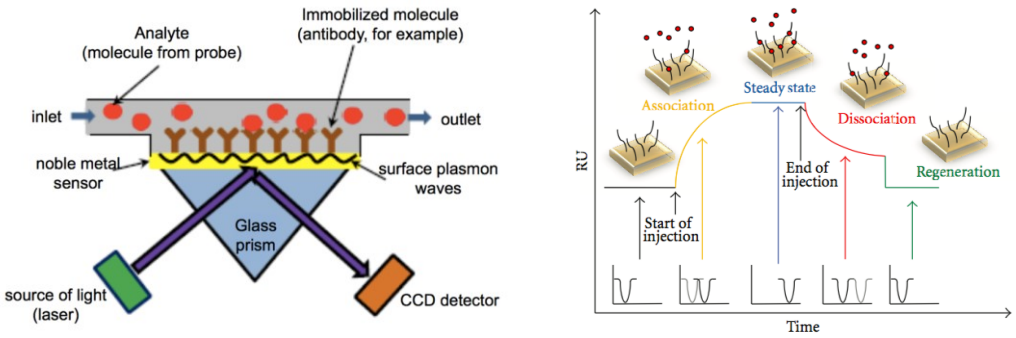


Figure 7 General principle of an instrument for surface plasmon resonance (left, Figure taken from [68]). Characteristic shape (sensorgram) and occurring phases of an SPR-measurement (right, Figure taken from [69]).

The change in angle of extinction of light, which occurs after polarized light impinges upon the film and gets reflected, can be monitored as a change in detector position for the dip in reflected intensity, which is called the plasmon resonance phenomenon. This method can be used without labelling the molecule to be bound to a ligand. For minimum diffusion rates (avoiding mass-transport limited binding of the analytes) it allows for real-time measuring and calculating affinity and binding kinetics. Possible changes of the binding partners' molecular properties due to labelling can be avoided. By determining the ratio of on- and off-rate constants, in association and dissociating phase, binding affinities can be calculated, thereby providing an easy and fast characterization of protein-protein interactions (Figure 7, right).[67]

2 Objectives

The aim of this project was to investigate the binding of anti-viral cyanovirin-N to one of its natural binding partners, namely highly glycosylated hemagglutinin (derived from insect cells), and its interaction with synthetically isolated glycopeptides. While binding studies of CV-N to several viruses have been reported, its affinity to hemagglutinin remains less studied. In order to investigate the role of sugar moieties within this interaction, homogeneously mannosylated peptides covering a partial sequence of the surface glycoprotein HA needed to be synthesized. The preparation of bioisosteric glycopeptides, imitating the natural Asn-carbohydrate structure, was approached by combining solid-phase peptide synthesis (SPPS) and copper(I)-catalyzed azide alkyne cycloaddition (CuAAC). Since not all of the required azido mannose building blocks were available, a new synthesis route needed to be developed, allowing the isolation and application of di-mannose azide in the field of click chemistry. CV-N variants carrying reduced numbers of disulfide bonds were recombinantly expressed to perform affinity studies using surface plasmon resonance (SPR). The main goal of this project was to investigate the influence of disulfide bonds in CV-N wild type and its variants as well as their essential binding properties to interaction sites on influenza hemagglutinin and synthesized glycopeptides (Figure 8).

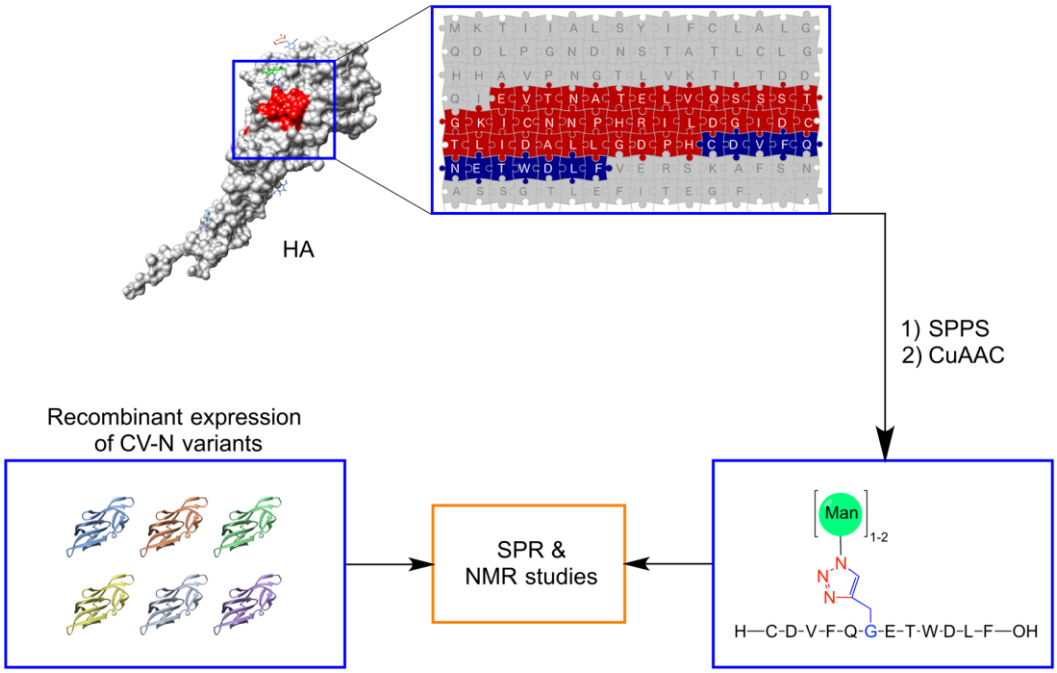


Figure 8 General strategy for binding studies of CV-N and glycopeptides.

3 Materials and Methods

Unless otherwise indicated, 'water' or 'ddH_2O' refers to deionized and filtered (0.22 µm) water provided from a Milli-Q® Integral system (MerckMillipore, Darmstadt, Germany). Sample weights were determined on a Sartorius precision balance (Göttingen, Germany). Inert atmospheres were obtained using either argon or nitrogen gas. 'Buffer A' refers to filtered and degassed MQ water and 'buffer B' to degassed acetonitrile, both treated with varying concentrations of TFA (see below).

Buffer composition for used systems:

Waters LC-MS:	buffer A and B with 0.05 % TFA each
Thermo Fisher LC-MS:	buffer A a fnd B with 0.05 % TFA each
Dionex RP-HPLC:	buffer A with 0.1 % TFA, buffer B with 0.08 % TFA
Varian Pro Star:	buffer A and B with 0.08 % TFA each

3.1 Solid-phase peptide synthesis (SPPS)

Peptides were synthesized using Fmoc-based, automated and microwave-assisted solid-phase peptide synthesis (Liberty Blue™, CEM, Kamp-Lintfort, Germany). Syntheses were performed under a nitrogen atmosphere. Deprotecting and coupling cycles were performed at 90°C. All stock solutions were freshly prepared prior to usage. *N*,*N*-Dimethylformamide (DMF) was used as main solvent, 20 % piperidine (v/v, in DMF) as deprotecting solution and DCM/DMF (1:1, v/v) as resin transfer solution. Fmoc-Phe-TentaGel® R PHB resin with loading of 0.2 mmol/g was used as solid phase and swelled for 1 h (in DMF) prior to the start of the synthesis. Oxyma (1 M in DMF) and DIC (0.5 M in DMF) were used as activators. Fully protected amino acids were prepared as 0.2 M stock solutions in DMF. TentaGel® resin was purchased from Rapp polymere (Tübingen, Germany), Fmoc-protected amino acids and oxyma from Merck (Darmstadt, Germany), DIC from Sigma-Aldrich (Darmstadt, Germany), DCM, DMF and MeOH from VWR (Vienna, Austria). After complete synthesis and final deprotection of the *N*-terminal Fmoc group, the resin was washed with DMF, DCM, MeOH and dried under vacuum. Global deprotection of side chain protecting groups and cleavage of the peptide from the solid support were performed using 'cleavage cocktail K' (3.5 h, under rotation, room temperature), referring to a mixture of TFA/phenol/water/thioanisole/EDT (82.5/5/5/5/2.5, v/m/v/v). The obtained solution was treated with ice-cold diethyl ether to precipitate the crude peptide while removing TFA (washing with diethyl ether was repeated three times), followed by resolubilization, lyophilization and purification (see 3.4).

3.2 Carbohydrate synthesis

Unless otherwise indicated, all solvents and other chemicals were obtained from Sigma-Aldrich (Darmstadt, Germany), TCI (Zwijndrecht, Belgium), Carbosynth (Compton, UK) or Synthose (Ontario, Canada) at the highest purity available and used as received, without further purification. Under standard Schlenk conditions, oven-dried glassware, dry solvents and reagents as well as inert gas (argon) were used. Column chromatography was performed on silica gel 60 (230–400 mesh, 0.04-0.063 nm; Macherey-Nagel, Düren, Germany). Reaction progress was monitored using thin-layer chromatography on aluminium sheets, coated with silica gel 60 with 0.2 mm thickness (Pre-coated TLC-sheets ALUGRAM® Xtra SIL G/UV_{254}; Macherey-Nagel). Compounds were visualized using ceric ammonium molybdate (CAM) staining and heating. Mass spectra were recorded either on a BRUKER maXis spectrometer using ESI (Billerica, USA), or a FINNIGAN MAT-95S spectrometer using EI as ion source (Thermo Fisher Scientific, Waltham, USA). HRMS spectra were recorded on a BRUKER maXis ESI-Qq-oaTOF system in positive or negative ion mode. IR spectra were recorded on a FT-IR-BRUKER Vertex 70-spectrometer at room temperature and neat conditions. Identified peaks are reported as wavenumber $\tilde{\nu}$ [1/λ]. ^{1}H-, and ^{13}C-NMR spectra were either recorded on BRUKER AVIII400 (400 MHz) or BRUKER AVIII600 (600 MHz), using $CDCl_3$, D_2O or deuterated DMSO as solvents. Temperature was set to 298.2 K. 2D NMR spectra were obtained using homonuclear (COSY, TOCSY), heteronuclear through-bond (HSQC, HMBC) and through-space correlation methods (NOESY). Chemical shift (δ) values are reported in ppm ('parts per million'). Spectra recorded with $CDCl_3$ as solvent were referenced to 7.26 (^{1}H) or 77.16 (^{13}C) ppm, with DMSO to 3.33 ppm (^{1}H) and with D_2O to 4.79 (^{1}H), respectively. NMR-spectroscopy splitting patterns were designated as singlet (s), doublet (d), doublet-doublet (dd), doublet of doublet of doublets (ddd) and triplet (t). Splitting patterns that could not be interpreted or easily visualized were designated as multiplet (m). Coupling constants are reported in Hertz (Hz).

3.3 Copper-catalyzed azide-alkyne cycloaddition (CuAAC)

3.3.1 In-solution reaction

Under standard CuAAC reaction conditions, Eppendorf tubes were first flushed with argon. In the next step, the inert gas was bubbled through all solvents (H_2O, DMF) to remove oxygen. Stock solutions of $CuSO_4$ (400 mM in ddH_2O), sodium ascorbate (100 mM in ddH_2O) and TBTA (400 mM in anhydrous DMF) were prepared freshly and used subsequently. Aliquots of these stock solutions were combined in a separate Eppendorf tube, where $CuSO_4$ was added first (blue solution), followed by the addition of the ligand TBTA (dark blue solution) and sodium ascorbate to reduce Cu(II) to Cu(I) (yellow solution). The obtained mixture was homogenized using a pipette. The solution was combined with the mixture of peptide (1 equiv.)

and azide-functionalized sugar (2-5 equiv.), and the reaction vessel was flushed with argon, followed by closing the lid. The reaction mixture (final concentrations: 2-4 mM peptide, 5-18 mM azide-functionalized sugar, 10-30 mM $CuSO_4$, 65-130 mM sodium ascorbate, 14-30 mM TBTA and DMF/H_2O = 4:1 (v/v)) was stirred at room temperature. To monitor the reaction progress *via* LC-MS (C4 or C18 column), aliquots of 1-2 µl were taken out under an argon flow and diluted with 19-28 µl H_2O/ACN/TFA (50:50 + 0.1 %) and TCEP was added. After the reaction was completed, the mixture was diluted up to 5 times of its initial volume with H_2O, thereby enhancing precipitation of TBTA, and purified as described in 3.4.

3.3.2 On-resin reaction

Resin-bound propargylglycine-modified peptide **11b** was dried after automated solid-phase peptide synthesis and stored in a desiccator. For the on-resin CuAAC reaction, resin was weighed out and swelled in DMF for 1 hour. Meanwhile, all reagents were prepared and added to a new Eppendorf tube as described in 3.3.1. Due to lower accessibility of the alkyne-modification, increased equivalents of corresponding azido sugar (4-5 equiv.) were used. The reaction tube was slightly rotated or shaken on an Eppendorf thermomixer comfort, to increase diffusion while preventing resin damage (in case of using a stirring bar). Reaction progress was monitored *via* standard washing and cleavage conditions. Hence, small resin aliquots were transferred into a syringe, washed extensively with DMF, DMF/water (90/10, v/v), DCM and MeOH, followed by drying in vacuum. In the next step, the crude peptide was cleaved from the solid support, deprotected, precipitated with ice-cold diethyl ether, resolubilized, lyophilized and purified as described in 3.4.

3.4 Peptide purification and analysis

Crude peptides were dissolved in 30/70 or 50/50 ddH_2O/ACN with 0.1 % TFA and reducing agent TCEP was added. Samples were first analyzed using liquid chromatography-mass spectrometry (LC-MS) on a Waters AutoPurification HPLC/MS system (Milford, USA). The detected product mass was set as target mass for the subsequent purification step *via* automated fraction collection. Sample separation was accomplished either on an analytical 300-5-C4 or -C18 Kromasil column (4.6 x 50 mm, Sigma-Aldrich) at a flow rate of 1 ml/min and a linear gradient from 5-65 % ACN + 0.1 % TFA (‚buffer B') in ddH_2O with 0.1 % TFA (‚buffer A') over 20 minutes. Mass spectra were obtained by electrospray ionization mass spectrometry (ESI-MS), operating in positive or negative ion mode. In addition, a Dionex UltiMate 3000 (Thermo Fisher Scientific, Waltham, USA) HPLC/MS systems was used for monitoring reactions as well as for analysis of crude and final products. The separation was accomplished with a Kromasil 300-5-C4 column (4.6 x 150 mm) at a flow rate of 1 ml/min using a linear gradient from 5-65 % ACN with 0.1 % TFA in ddH_2O with 0.05 % TFA over 30 minutes.

Purification was performed on the Waters system or on a Varian Pro Star system (Palo Alto, USA) using a Prep (250 x 21.2 mm) or Semiprep (250 x 10 mm) Kromasil 300-10-C18 column at a flow rate of 10 ml/min or 3 ml/min, respectively. The gradient was set to increase from 5-20 % of buffer B over 5 min and 20-50 % of buffer B over 30 min. Crude petpides were dissolved in 10 ml 6 M GuaHCl (pH 4.7), sterile filtered (0.2 µm) and injected for purification. Fractions were automatically collected corresponding to the selected masses when using the Waters AutoPurification HPLC/MS system. Regarding the Varian Pro Star system, peptides were purified using a gradient from 5-25 % ACN (5 min) to 25-55 % ACN (or 25-65 % ACN) over 60 minutes. Buffer systems of A (water) and B (ACN), both containing 0.08 % TFA were used. Fractions were collected as 1 ml, while running a flow rate of 3 ml/min (20 seconds per fraction) and subsequently analyzed using direct injection mass analysis (Waters ESI-MS). Combined fractions were lyophilized and after final analysis used in further steps.

For analytical RP-HPLC measurements and purification, a Dionex UltiMate 3000 instrument at a flow rate of 1 mL/min with a linear gradient from 5 to 65 % buffer B in buffer A over 30 min (buffer A: 0.1 % (v/v) TFA in ddH_2O, buffer B: 0.08 % (v/v) TFA in ACN) was used. Crude peptides were analyzed and purified on a Kromasil 300-5-C18 (150 x 4.6 mm) or 300-5-C4 (150 x 4.6 mm) column.

HRMS spectra were recorded using a BRUKER maXis ESI-Qq-oaTOF system in positive and negative ion mode, respectively. Lyophilization of peptide solutions was performed on a Christ Alpha 2-4 plus system (Martin Christ GmbH, Osterode am Harz, Germany), using a cold trap (-85°C) and operating at a pressure of < 0.1 mbar.

3.5 Recombinant protein expression

Unless otherwise stated, all procedures from transformation to cell harvest (after protein expression) were performed under sterile conditions, either near a Bunsen burner flame or a laminar air flow. For each of the six (CV-N)2L0 variants, lyophilized plasmid DNA was obtained from GenScript® (pET-27b(+) vector, Figure 9; Piscataway, USA) and dissolved in sterile ddH_2O to a final concentration of 100 ng/µl.

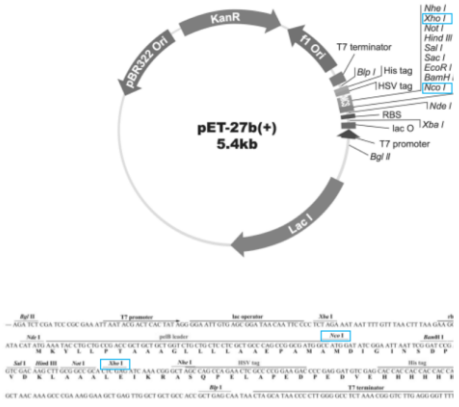


Figure 9 pET27b(+) vector with highlighted restriction enzymes for insert. Image obtained using SnapGene (GSL Biotech LLC, Chicago, USA).

Culture medium LB-Lennox was prepared as follows: 10 g/L peptone, 5 g/L yeast extract and 5 g/L NaCl were dissolved in ddH_2O and the pH was adjusted to 7.4 using 4 M NaOH. LB agar medium was prepared equally, apart from adding 20 g/L agar. The solution was transferred into a Schott bottle and autoclaved (20 minutes at 120°C). LB-Lennox was stored at room temperature; LB agar medium was cooled to 55-60°C, followed by the addition of 50 mg kanamycine per L medium. Agar plates were prepared as follows (laminar air flow): A thin layer of LB agar was poured into each sterile Petri plate and swirled in a circular motion. The plates were cooled with open lid until complete solidification (~ 45 minutes). Plates were sealed, labelled and stored in plastic bags at 4°C.

Restriction digest of the obtained plasmids was performed using *Nco*I and *Xho*I as restriction enzymes. For each variant, 10 µl reaction mixture, consisting of 3 µl plasmid (100 ng/µl), 1 µl CutSmart buffer (10x), 5 µl RO H_2O and 0.5 µl of each restriction enzyme, were prepared. The mixture was incubated at 37°C for 1 hour. TAE buffer containing 40 mM Tris, 20 mM acetic acid and 1 mM EDTA was prepared and used for agarose gel electrophoresis, where 1 % agarose and GelRed (1:20.000 dilution) were added. Each of the restriction digest samples (10 µl) was mixed with 2 µl DNA loading buffer (6x) and pipetted into the pockets. 1 kb DNA ladder (6 µl, 500 µg/ml; NEB, Ipswich, USA) was used as a reference. Electrophoresis was performed at 120 V for 25 minutes. The gel was then illuminated under UV radiation.

Competent BL21(DE3) *E. coli* were prepared as follows: 20 ml of LB broth (LB-Lennox) were transferred into a sterile Erlenmeyer flask, inoculated from a BL21 cell cryo stock (stored at -80°C) and shaken at 200 RPM and 37°C overnight. The culture was then diluted 1:50 with LB broth and shaken at 200 RPM and 37°C for 3 hours, until an OD_{600} of ~ 0.6 was reached. After cooling on ice for 30 minutes, the culture was centrifuged for 15 min. at 4.000 x g and 4°C. The supernatant was discarded, the remaining pellet redissolved in 50 ml 0.1 M $CaCl_2$ and centrifuged again under the same conditions. After discarding the supernatant, 3 ml 0.1 M $CaCl_2$ and 0.6 ml glycerol (87 %) were used to resuspend the pellet, followed by storage on ice. Aliquots of 200 µl were transferred into sterile Eppendorf tubes, immediately frozen in liquid nitrogen and stored at -80°C.

The following transformation was performed equally for each variant (V1-V6). An aliquot of the prepared competent cells was taken out of -80°C and thawed on ice for 30 minutes. In a sterile Eppendorf tube, 2 µl of the plasmid DNA solution (100 ng/µl) were mixed with 80 µl of competent cell solution and stored on ice for 30 minutes. Heat shock was performed for 90 s at 42°C on a shaker, followed by cooling on ice for 3 minutes. Preheated SOC medium (37°C) was added to a final volume of 1 ml, followed by incubation at 37°C for 45 minutes. The solution was split (900 and 100 µl, respectively), transferred on agar plates and gently plated using a sterile metal cell spreader. The plates were incubated overnight at 37°C. Grown colonies were identified and used for protein expression.

Initial protein expression was carried out on a small scale (2 ml inoculated LB-Lennox medium) and successful expression of the desired protein analyzed using SDS-PAGE and Westen Blot. Procedures were carried out equivalently for big scale protein expression (0.5-1 L medium), as described below. Autoclaved LB-Lennox was mixed with kanamycine (50 µg/ml). 50 ml of LB-Lennox were transferred into a sterile Erlenmeyer flask, inoculated from the corresponding colony identified on the agar plate and shaken at 180 RPM and 37°C overnight. After determination of OD_{600} (1 ml aliquot diluted 1:10 with LB broth), the culture was transferred into a 3 L Erlenmeyer flask and diluted with LB broth to an OD_{600} of 0.05-0.1 (approximately 1 L). The culture was then shaken at 180 RPM and 37°C for 3 hours, until an OD_{600} of ~ 0.6 was reached. For induction, the culture was cooled down to 18°C and IPTG was added to a final concentration of 500 µM. The culture was shaken overnight at 180 RPM and 18°C.

Regarding cell lysis, two approaches were performed. Cell lysis and following protein isolation from the soluble fraction was carried out as follows: solution was centrifuged at 4.000 x g and 4°C for 15 minutes, the supernatant was discarded. The obtained pellet was resuspended in PBS buffer and centrifuged again under same conditions as mentioned before. The supernatant was discarded and the remaining pellet was resuspended in 10 ml lysis buffer (50 mM Tris-HCl, 200 mM NaCl, 5 % glycerol, 500 ng/ml lysozyme, 1 mM PMSF, 1 mM DTT, 1 mM $MgCl_2$) and incubated for 1 hour at 37°C. The mixture was then completely frozen (-80°C) and thawn again. After repeating this procedure, the solution was ultrasonificated for a few seconds. The mixture was then centrifuged at 4.000 x g and 4°C for 15 minutes. The soluble and insoluble fractions were separated and analyzed using SDS-PAGE and Western Blot. Cell lysis and following protein purification from the insoluble fraction (inclusion bodies) was performed as follows: cell pellet was resuspended in 40 ml lysis buffer (50 mM Tris-HCl, 200 mM NaCl, 5 % glycerol, 1 mM PMSF, 1 mM DTT, 1 mM $MgCl_2$ and 1 % Triton™ X-100), followed by mechanical cell lysis using a IKA® T18 basic ULTRA-TURRAX® and Constant Systems LTD cell disrupter (1.8 kbar, cleaned with 0.1 M NaOH and 80 % EtOH) attached to a recirculating chiller (VWR RC-10 Digital Chiller). The obtained mixture was centrifuged for 30 min at 4°C and 16.000 RPM, followed by separating supernatant (soluble) and insoluble (inclusion bodies) fractions. The insoluble fraction was resuspended in 40 ml 8 M GuaHCl / 50 mM Tris (pH = 8) by stirring for 24 hours until complete solubilization. Afterwards, the suspended cells were centrifuged (30 min, 16.000 RPM, 4°C), the supernatant was decanted, immediately stored on ice and then used for Ni-NTA chromatography. The pellet was discarded.

3.6 Protein purification and analysis

Cloning of a His-tag into a vector and following expression leads to polyhistidine-tagged recombinant proteins. After cell lysis, this modification allows the separation of the His-tagged

recombinant protein from other proteins in the bacterial cells. Since the polyhistidine-tag coordinates with high affinity to the Nickel-Nitrilotriacetic acid (Ni-NTA), the desired protein remains on the column, while other proteins flow-through. After the washing steps, the polyhistidine-tagged recombinant protein can be eluted using imidazole. In addition, the presence of His-tags can be visualized using Western Blot. After cell lysis, the sample was separated into soluble and insoluble fraction. An aliquot of both was taken and analyzed *via* SDS-PAGE.

The soluble fraction was purified using Ni-NTA column chromatography, either manually or automatically (ÄKTA prime plus) performed. Regarding manual Ni-NTA purification, the obtained soluble fraction was diluted 1:1 with purification buffer (50 mM NaH_2PO_4, 500 mM NaCl, pH 8.0) and the pH was controlled to be about 8. A slurry of Ni-NTA bead suspension in purification buffer was added and the mixture was rotated at 15 RPM and 4°C for 1 hour. The column was packed by adding the suspension and removing the excess liquid, but preventing the column from running dry. The packed column was washed five times with purification buffer (column volume), where fractions were already collected. The same procedure was performed for the washing step, where purification buffer with 10 mM imidazole was used. Finally, elution buffer containing 300 mM imidazole was added and fractions collected. The obtained fractions were analyzed using SDS-PAGE, clean fractions pooled and concentrated using 10 kDa cut-off spin-columns. Buffer exchange was performed in spin-columns, diluting (1:10 with SPR running buffer) and reconcentrating the solution four times. Regarding automatically performed Ni-NTA purification, the soluble fraction was loaded onto a regenerated Ni-NTA bead column (1 ml/min). The system was then washed using TBS buffer (50 mM Tris, 150 mM NaCl, pH 7.5), until no protein signal was observed *via* UV. Afterwards, the gradient (0-100 % 500 mM imidazole in PBS buffer over 60 minutes) was started and fractions were collected (1 ml/min).

The solution obtained after solubilization of the insoluble fraction with 8 M GuaHCl / 50 mM Tris (pH = 8) was purified using Ni-NTA column chromatography. The column was washed 6 M GuaHCl / 150 mM NaCl / 50 mM Tris / pH = 8 until baseline level was reached. Then, the sample was loaded onto the column (1 ml/min), followed by washing with 6 M GuaHCl / 150 mM NaCl / 50 mM Tris / pH = 8 until baseline level was reached (UV). Then a gradient elution was performed and fractions were collected. The used elution buffer contained 6 M GuaHCl / 150 mM NaCl / 50 mM Tris / pH = 8 and 500 mM imidazole. The gradient was set up as a mixture of washing and elution buffer, reaching 100 % elution buffer after 60 minutes. The fractions were collected automatically (1 ml/min tube). Aliquots (50 µl) of the GuaHCl containing fractions were treated with 15 µL of 40 % TCA to precipitating the protein. The mixtures were cooled on ice for 10 minutes, followed by centrifugation (10 minutes, 15.000 RPM, room temperature). The supernatant was discarded and the pellets were resuspended

using 90 % EtOH (200 µL each). After another centrifugation step (10 minutes, 15.000 RPM, r.t.), the supernatants were discarded and ddH_2O as well as SDS buffer (10 µL each) were added. The samples were denatured for 5 minutes at 95°C and centrifuged briefly, followed by SDS-PAGE analysis.

SDS-PAGE running buffer (Laemmli buffer) was prepared from a 10 x stock solution, consisting of 250 mM tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), 2 M glycine and 1 % (w/v) SDS in ddH_2O. Separating gels were prepared as follows: 12.5 mL ddH_2O, 13.5 mL separating gel buffer (1.5 M Tris-HCl, 0.4 % (w/v) SDS, pH 8.8), 26.0 mL 30 % (w/v) acrylamide, 0.55 mL 10 % (w/v) SDS, 0.55 mL 10 % (w/v) APS, 18 µL TEMED. Stacking gels were prepared as follows: 12.5 mL ddH_2O, 2.5 mL stacking gel buffer (0.5 M Tris-HCl, 0.4 % (w/v) SDS, pH 6.8), 3.5 mL 30 % (w/v) acrylamide, 185 µL 10 % (w/v) SDS, 185 µl 10 % (w/v) APS, 18 µL TEMED. SDS-PAGE sample buffer (2 x) was prepared using 500 mM Tris, 35 % (v/v) glycerin, 6 % (w/v) SDS, 3.55 % (v/v) 2-mercaptoethanol, 0.05 % (w/v) bromophenol blue and pH was adjusted to 6.8. Samples were mixed 1:1 with sample buffer and denatured for 5-10 min at 95°C. After spinning down, 15 µl of the corresponding samples were pipetted into the pockets. GE HC Amersham low molecular weight marker (6 µl) was used as a standard. Electrophoresis was performed at 250 V (const.) for 35 minutes. Afterwards, gels were rinsed with water and stained in Coomassie brilliant-blue R250 solution (0.1 % (w/v), 10 % (v/v) acetic acid and 45 % (v/v) methanol in water) for 25 minutes under shaking. After washing twice with water, either destaining solution (10 % (v/v) acetic acid and 40 % (v/v) methanol in water) was added for 1-2 hours or water, followed by shaking overnight. Obtained gels were imaged using a Bio-Rad ChemiDoc™ MP imaging system (Hercules, USA).

Western Blot was performed for crude protein solutions obtained from small scale expressions as well as for isolated fractions after Ni-NTA chromatography. After SDS-PAGE, the gel, Whatman® gel blotting papers and the membrane were soaked in Towbin transfer buffer (25 mM Tris-HCl, 192 mM glycine, pH 8.3, 20 % (v/v) methanol) and then transferred onto the Western Blot device (paper, membrane, gel, paper; bottom to top). The Blot was performed at 20 V for 45 minutes. After using Ponceau S stain (0.2 % (w/v) Ponceau S, 5 % glacial acetic acid) for quick identification of obtained bands, the membrane was washed with water to remove the stain again. In the next step, blocking solution TBST (20 mM Tris, pH 7.5, 150 mM NaCl, 0.1 % Tween 20) with 0.5 % BSA was added and stored at 4°C overnight. The solution was discarded and the first antibody (Anti-His, dissolved in TBST solution with 0.5 % BSA) added, followed by shaking at room temperature for 1 hour. The solution was discarded and the membrane washed twice with TBST solution under 5 minutes of shaking. The second antibody (Anti-Mouse, in TBST + 0.5 % BSA) was added and incubated while shaking at room temperature for 1 hour. The solution was discarded and washed twice with TBST solution (5

min each). The membrane was stained with SIGMAFAST™ BCIP®/NBT and quenched by adding 5 % acetic acid.

Proteins were dialyzed against 100 mM PBS buffer (pH 7.0). Protein solution (5-10 ml) was transferred into a dialysis tube with a 6 kDa pore filter. The casket was then placed in a beaker (3 L) and 1 L of PBS buffer was added. The solution was slightly stirred overnight at 4°C, followed by replacing the PBS buffer with a new batch and further dialysis for 4 hours (repeated twice). The protein solution was then removed from the casket, transferred to Amicon filters with 10 kDa cut-off and further concentrated using centrifugation. For centrifugation, a Beckmann Coulter Cooler Allegra X-30R centrifuge was used at 4.500 x g and 4°C.

Protein concentration was determined using a NanoDrop UV-Vis2000c spectrophotometer. Absorption at wavelengths of 260 and 280 nm were measured and based on the molecular weight and the calculated extinction coefficient (23474 Da for M-(CV-N)2L0 and 20.440 M^{-1} cm^{-1}, see supplementary information). PBS buffer (100 mM, pH 7) was used as blank and protein concentration was determined using 3 measurements of each dilution (1:10-1:100).

Proteins in 100 mM PBS buffer (pH 7.0) were analyzed using liquid chromatography-mass spectrometry (LC-MS) on a Waters HPLC/MS system. Sample separation was accomplished on an analytical 300-5-C4 Kromasil column (4.6 x 50 mm) at a flow rate of 1 ml/min and a linear gradient from 5-65 % ACN with 0.1 % TFA in ddH_2O with 0.1 % TFA over 20 minutes. Mass spectra were obtained by electrospray ionization (ESI-MS), operating in positive or negative ion mode. Samples were measured in continuous mode and deconvoluted afterwards. MagTran was used to identifiy the corresponding molecular peak. For analytical RP-HPLC measurements, Dionex UltiMate 3000 instrument at a flow rate of 1 mL/min with a linear gradient from 5 to 65 % buffer B in buffer A over 30 min (buffer A: 0.1 % (v/v) TFA in ddH_2O, buffer B: 0.08 % (v/v) TFA in ACN) was used. Proteins were analyzed on a Kromasil 300-5-C4 (150 x 4.6 mm) column.

Circular dichroism measurements of the protein solutions were performed using Chirascan-Plus Applied Photophysics. The channel was equilibrated for 2 hours under a N_2 flow. The lamp was heated for 30 minutes and the temperature was set to 25°C. PBS buffer (100 mM NaPi, pH 7) was used as blank and samples (20-30 µM) were measured 5 times. The signals were averaged and manually subtracted from the blank signal.

3.7 Surface plasmon resonance (SPR)

Affinity and kinetics studies were performed on Reichert's SR7500DC (Reichert, Buffalo, USA), a two channel surface plasmon resonance instrument. CMD500M SPR sensorchips were purchased from Xantec (Düsseldorf, Germany), representing

carboxymethyldextran hydrogel coated chips for immobilization of HA and corresponding glycopeptides. Measurements were performed at 25°C and at a flow rate of 25-30 µl/min.

Chips were immobilized using single-channel amine coupling, where 0.4 M EDC/HCl and 0.1 M NHS were used to activate the carboxymethyldextran chip surface. The ligand (HA or glycopeptides) was diluted to 0.02 mg/ml in 10 mM sodium acetate buffer (pH 5) and then transferred to the activated surface. Recombinant influenza A virus hemagglutinin H3 protein was obtained from abcam (Cambridge, UK). The source of origin were baculovirus infected insect cells and the sequence of this protein is according to H3N2 A/Wisconsin/67/05, glycosylated with *N*-linked sugars. Glycopeptides were synthesized and functionalized using SPPS and CuAAC (see results and discussion). Following the immobilization of the ligand, all remaining reactive groups were deactivated using 1 M ethanolamine/HCl (pH 8.5). The second channel, not floated with ligand, served as a reference. Running buffer (HBS-EB (+)) contained 10 mM HEPES, 150 mM NaCl, 3 mM EDTA and 0.05 % Tween at a pH of 7.4. All solutions were degassed and filtered (0.2 µm) before injection into the system. Kinetic studies were performed using analyte dilutions in the range of 10^{-5}-10^{-8} M, a regeneration step of the free ligand after each injection of proteins at various concentrations (8 cycles in total), and blank measurements after different proteins. The chips were cleaned with 10 mM glycine/HCl solution (pH 1.5) and 10 % SDS, taken out of the system and stored in running buffer at 4°C. Channels and needles of the Reichert system were washed with iso-propanol/ACN/MeOH/H_2O + 1% FA (1:1:1:1, v/v/v/v) and 20 % EtOH.

4 Results and Discussion

4.1 Peptide synthesis

We started by synthesizing a peptide consisting of 12 amino acids, which represents a partial sequence of the HA top region. The protein is shown in Figure 10, where the HA top region is partially highlighted in red and the chosen peptide sequence in blue.

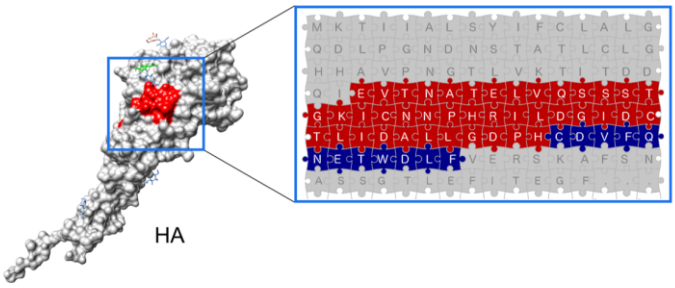


Figure 10 Overview of HA (surface representation) with its top region partially highlighted in red and the sequence of the model peptide in blue.

The peptides were synthesized using automated solid phase peptide synthesis (LibertyBlue, CEM), where peptide **10** (Figure 11) represents the unmodified sequence and peptide **11a** the propargylglycine-modified version (replacing asparagine). This modification was used to introduce the triazole-linked *N*-glycosylation, thereby mimicking the naturally occurring *N*-glycosylation of asparagine.[71]

10 **11a**

Figure 11 Peptides corresponding to a partial sequence of HA, synthesized by automated solid phase peptide synthesis.

The peptides were synthesized on a scale of 0.07 mmol each, using TentaGel-R-PHB resin with loading of 0.2 mmol/g. After cleaving the crude peptide from the resin using ‘solution K’ (see methods), ether precipitation as well as purification on preparative LC-MS were performed. Both peptides were obtained in good purity and yield (26 % for peptide **10**, 14.5 mg, 9.6 µmol, from 53 % resin; 26 % for peptide **11a**, 13.4 mg, 9 µmol, from 50 % resin). The products were analyzed using mass spectrometry as well as on a C18 column RP-HPLC equipped with a UV detector (Figure 12, 13).

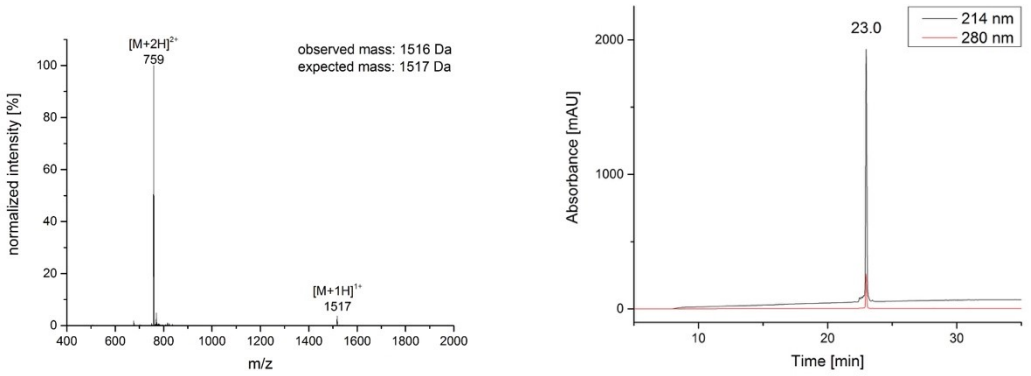


Figure 12 Mass spectrum (left) and RP-HPLC chromatogram (right) for peptide **10**; $[M+H]^{+}_{calc}$ = 1517.6 Da, found 1517.5 Da; $[M+2H]^{2+}_{calc}$ = 759.3 Da, found: 759.1 Da; Retention time (C18 column): 23.0 min.

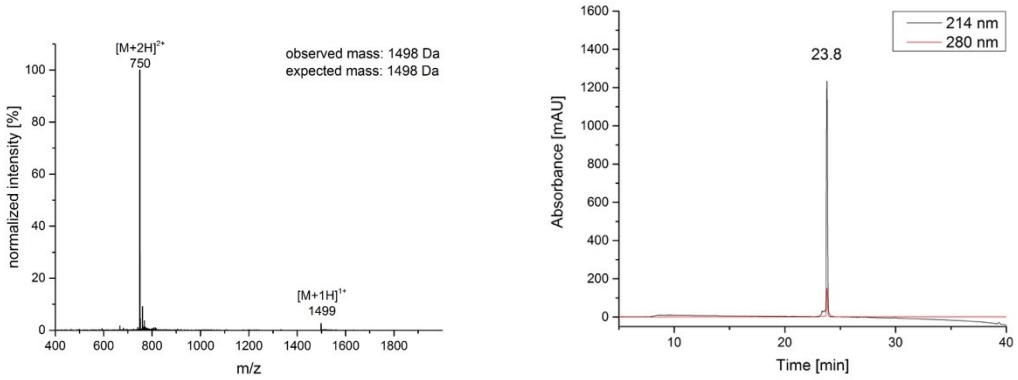


Figure 13 Mass spectrum (left) and RP-HPLC chromatogram (right) for peptide **11a**; $[M+H]^{+}_{calc}$ = 1498.6 Da, found 1498.6 Da; $[M+2H]^{2+}_{calc}$ = 749.6 Da, found: 749.8 Da; Retention time (C18 column): 23.8 min.

While the mass spectra of both crude peptides have shown side products, the designed peptides of interest were isolated without considerable impurities. Unmodified peptide **10** was stored as a reference for the following SPR measurements (later on replaced by comparison of mono- and di-glycosylated peptide), while the modified peptide **11a** was used for CuAAC reactions to introduce the triazole-linked *N*-glycosylation for binding studies.

4.2 Preparation of azido-mannose building block (2)

Besides the available azido building block **1**, deprotected azido-mannose **2** was required for the sequentially applied copper-catalyzed azide-alkyne cycloaddition (CuAAC). Therefore, several deprotection approaches were examined (Figure 14 and Table 1).

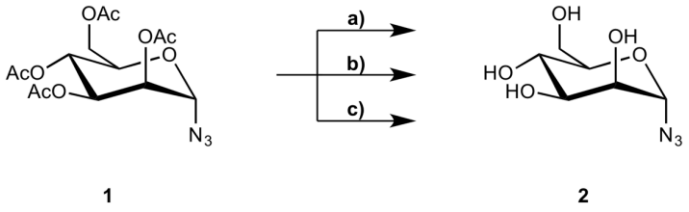


Figure 14 Deprotection approaches for the preparation of building block **2**.

entry	deprotect. method	equiv. base	reaction time [h]	yield [%]
a)	NaOMe/MeOH	cat.	24	92
b)	Hydrazine/MeOH	9	40	99
c)	NMP/MeOH	8	18	97

Table 1 Methods used for deprotection of tetraacetylated azido-mannose **1**.

The common usage of catalytic amounts of sodium methoxide in methanol (MeOH) showed full conversion after 24 hours (entry a)).[72] A yield of 92 % was obtained after filtering the product over Dowex 50 H^+ resin. Drying the product by lyophilization on the other hand, extended the preparation time. The second approach, using hydrazine in MeOH, showed quantitative yields after 40 hours (entry b)).[73] In contrast, the relatively high boiling point of hydrazine (113.5°C at standard conditions) requires long drying procedures. Fortunately, the third approach using NMP (*N*-Methylpyrrolidine) in MeOH led to shorter reaction times and an excellent yield of 97 % as well as easy removal of solvents under vacuum (entry c)).[74] NMR analysis revealed that all methods allowed the preparation of a clean product and no column chromatography was required. The characteristic doublet of C_1-*H* at 5.41 ppm with a coupling constant of 1.8 Hz (^{1}H NMR, α-anomer) and the C_1 signal at 89.7 ppm (^{13}C NMR) are in accordance with literature data.[75–77]

Therefore, this convenient and efficient approach was used for a larger scale transformation of acetylated azido-mannose **1** into fully deprotected azido-mannose **2**. The low boiling points of NMP (80-81°C) and methanol (65°C) allowed fast isolation and drying of azido-mannose **2**. The obtained product was then used for CuAAC reactions.

4.3 Synthesis of di-mannose azide (9)

While mannose building blocks **1** and **2** provided an initial starting point of our CuAAC experiments, the synthesis of carbohydrates needed to be expanded to further include di-mannose due to its biological relevance. Recent studies concluded that CV-N binds strongest to the defined Manα(1-2)Manα unit in glycoarray screenings and NMR sugar titration studies.[14,15] We therefore aimed for the synthesis of di-mannose, modified with an azide on *C1*, and using it for the following click reactions. Fully deprotected di-mannose azide **9** has not been reported in the literature and experimental data for acetylated di-mannose azide **8** are only barely reported (Figure 15).

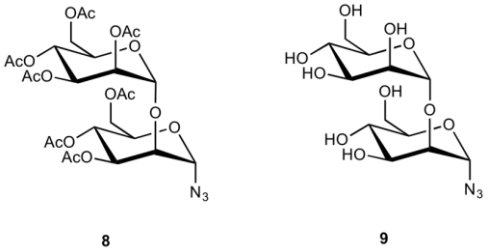


Figure 15 Azide-functionalized di-mannose moieties **8** (acetylated) and **9** (deprotected).

We set up a synthesis route for the desired compounds (Figure 16). For this approach, recently published techniques were combined. In the first few steps, tetra-acetylated bromo-mannose **3** was transformed to methylated di-mannose **5** according to Fraser-Reid and co-workers.[78] As presented by Szurmai *et al.*, the transformation from disaccharide **5** to fully acetylated product **7** was performed.[79] The azide-modification for the preparation of product **8** was based on a reported method, presented for other saccharides.[77] Final deprotection to obtain azido di-mannose **9** was carried out with commonly used sodium methoxide.[72]

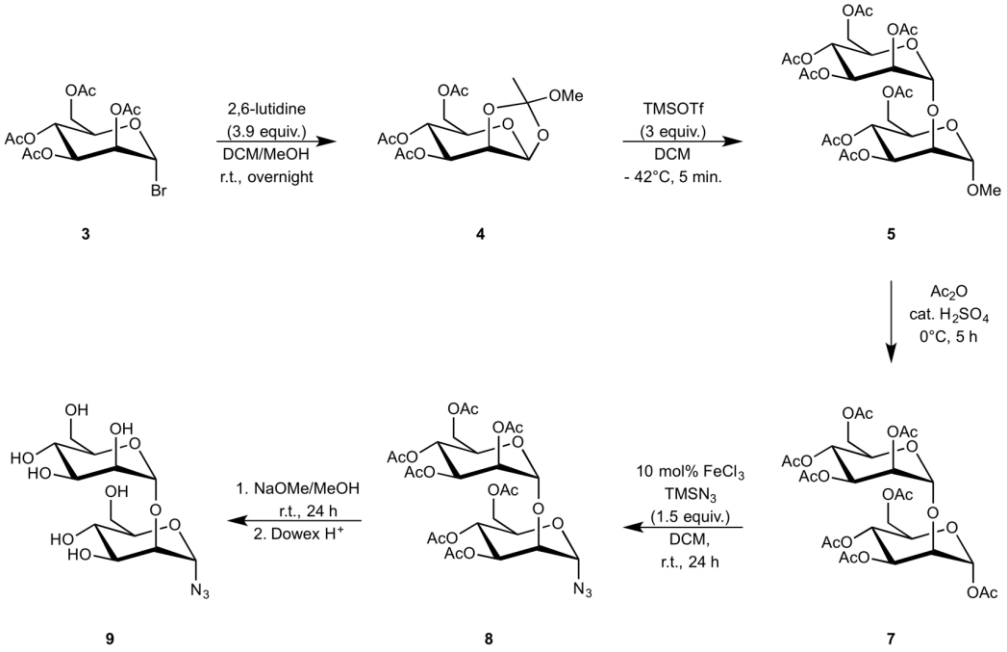


Figure 16 Concept for the synthesis of building block **9**. Methyl tri-mannose **6** (see also Figure 18) was observed as a side product.

4.3.1 Preparation of methyl-1,2-orthoacetate mannose (4)

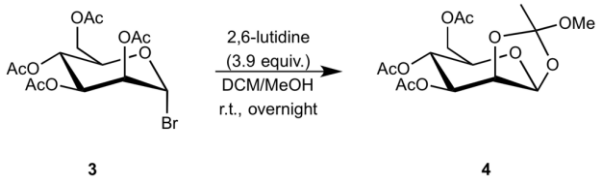


Figure 17 Preparation of methyl-1,2-orthoacetate mannose **4**.

The preparation of methyl-1,2-orthoacetate mannose **4** was achieved using the sterically hindered base 2,6-lutidine in methanol and dichloromethane, as recently reported.[80,81] The susceptibility of bromo-functionalized mannose **3** towards hydrolysis and its high viscosity complicated the setup of this reaction. While the preparation was performed under inert atmosphere (argon), weighing out of the starting material was not carried out under inert conditions. As a result, partial hydrolysis has likely occurred prior to the reaction. The crude product was purified using column chromatography and obtained in repeatedly modest yields (22-31 %). Thin-layer chromatography, subsequently performed after addition of the reagents, revealed a mixture of several compounds. This may indicate early hydrolysis and could be an explanation for the low yields of the desired product **4** while complicating the performed purification. According to the literature, an alternative procedure would be the transformation of D-Mannose to compound **3** using hydrobromic acid and acetic anhydride, followed by subsequent preparation of methyl-1,2-orthoacetate mannose **4** from unpurified compound **3**, thereby limiting the extent of hydrolysis.[81]

4.3.2 Synthesis of methyl di-mannose (5)

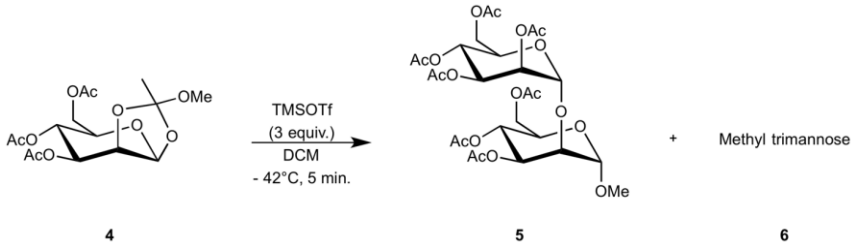


Figure 18 Preparation of methyl di-mannose **5**.

After the isolation of compound **4**, preparation of methyl disaccharide **5** was performed by using TMSOTf at -42°C and under inert conditions. While literature reported reaction conditions of -30°C, we chose a cooling bath consisting of dry ice and acetonitrile due to hazard-related reasons.[80,81] The following purification step revealed a mixture of several products, out of which the desired methyl di-mannose **5** was isolated at low yields (22 %). The altered reaction conditions or wet dichloromethane could have led to a decrease in yield. In addition, the side product methyl tri-mannose **6** was formed, but not isolated from the crude due to the complex mixture with many side products. The formation of compound **6** was revealed by the characteristic peak pattern for the three C_1-*H* atoms obtained from ^{1}H NMR

spectroscopy. Recently, Hamon *et al.* observed the formation of methyl tetra-mannose under the reported conditions.[82] These species and other side products complicated the isolation of methyl di-mannose **5**, thereby preventing higher yields. Nevertheless, pure methyl di-mannose **5** was obtained and the characteristic signals of the two anomeric hydrogens (doublets at 4.84 and 4.91 ppm with coupling constants of 1.8 Hz) were observed in ^{1}H NMR (Figure 19).

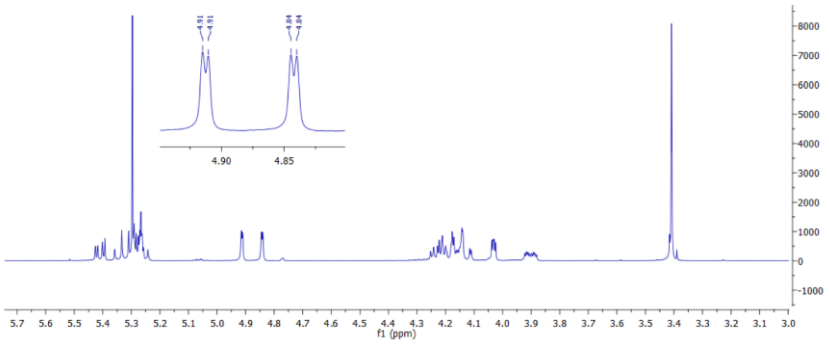
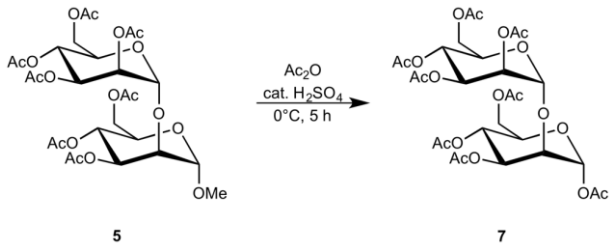


Figure 19 Partial ^{1}H-NMR spectrum of isolated methyl di-mannose **5**. Doublets at 4.91 and 4.84 ppm corresponding to the H_1 and $H_{1'}$ signals, respectively.

Since the focus was on the isolation of methyl di-mannose **5**, the reaction as well as purification were not optimized and the isolation of methyl tri-mannose **6** was not performed either. Nevertheless, this compound is still of interest, since its following azide functionalization (see 4.3.4) has not been reported yet and would open the possibility for its use in CuAAC reactions for the glycosylation of peptides without the necessity of a linker.

4.3.3 Preparation of acetylated di-mannose (7)



5 7

Figure 20: Acetylation of methyl-di-mannose **5**.

Before introducing the desired azide functionality on C_1, methyl-di-mannose **5** was completely acetylated to compound **6** using acetic anhydride and catalytic amounts of sulfuric acid at 0°C (Figure 20).[79] This step allows the transformation of the methoxy group into a better leaving group (acetoxy), which has shown to be critical for the introduction of the azide. The reaction was monitored using TLC and the crude product purified using column chromatography, obtaining compound **7** in a good yield of 82 %. The formation of product **7** was confirmed by ^{1}H NMR spectroscopy, where the disappearance of the characteristic peak of the methoxy group (δ 3.41 ppm (s)) and the signal of an additional acetyl peak (δ 2.0-2.2 ppm (s)) were observed.

4.3.4 Synthesis of acetylated di-mannose azide (8)

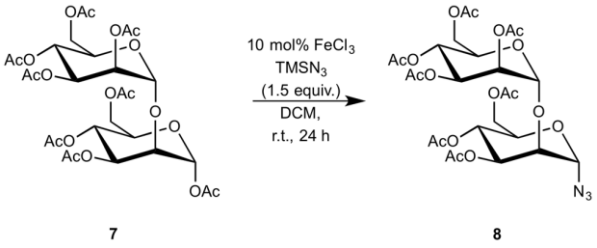


Figure 21 Azide-functionalization of acetyl-di-mannose **7**.

While Salunke *et al.* presented an iron(III) chloride catalyzed azide-introduction to several mono-, di- and trisaccharides, no azide functionalization of di-mannose has previously been reported.[77] We used slightly higher catalyst loading (10 mol%) while working under an inert atmosphere and monitored the reaction *via* TLC, followed by purification using column chromatography (Figure 21). Product **8** was obtained as a white solid in 55 % yield. Attempts to shorten the synthesis by one step, directly using methyl-di-mannose **5** for the azide transfer, were not successful. However, it has not been investigated in detail if water containing DCM (and therefore inactivation of the iron catalyst) inhibited the direct introduction of the azide group or if methoxy as a worse leaving group caused the unsuccessful transformation. In future, the direct transformation from compound **5** to **8** using dry DCM or higher catalyst loading could have potential to shorten the synthesis and thereby increasing the overall yield. Regarding the stereochemistry, the formation of the α-azido-glycoside is favored by the anomeric effect, driven by hyperconjugation or dipole minimization (Figure 22).

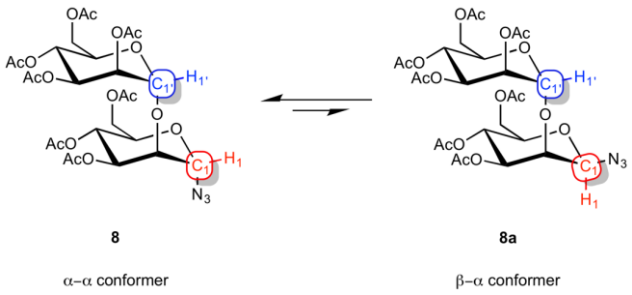


Figure 22 α,α- and β,α-conformer of di-mannose azide **8**.

4.3.5 Preparation of di-mannose azide (9)

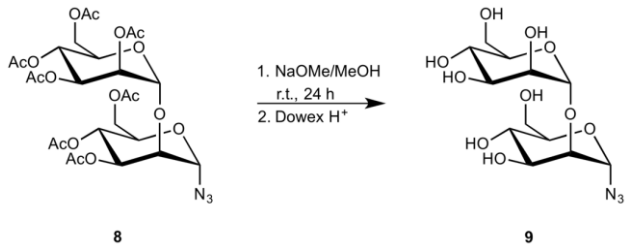


Figure 23 Deprotection of azido-mannose **8**.

In the last step of the synthesis of di-mannose azide **9**, compound **8** was fully deprotected using catalytic amounts of sodium methoxide in methanol, followed by ion

exchange using Dowex 50 H^+ (Figure 23).[72] The desired compound was synthesized in a yield of 93 %. Preliminary crystallization experiments were set up with water, methanol, tetrahydrofuran and dioxane mixtures. However, no single crystals were obtained. To our best knowledge, the preparation of di-mannose azide **9** has not been reported in the literature and our obtained data from NMR, HRMS and IR show the first characterization of this compound (see experimental section). As shown in chapter 4.4, di-mannose azide **9** was used for CuAAC reactions to prepare di-glycosylated peptides (partially mimicking HA), which in turn were used for the following binding studies on CV-N.

In conclusion, the presented synthesis route has shown to provide access to azido di-mannose **9**. Due to several challenges in this first attempt, such as temperature regulation, hydrolysis and product purification, the overall yield was very low (3 % over 5 steps). However, the main losses arise from the preparation of **4** from **3**. This process should be significantly improved by synthesizing compound **3** from D-mannose and subsequent preparation of compound **4** to prevent hydrolysis. In addition, by reproducing literature results reported for the formation of methyl di-mannose **5** and optimizing conditions for the azide transfer to obtain product **8**, it should be possible to increase the overall yield to > 10 %.

4.4 Copper-catalyzed azide-alkyne cycloaddition (CuAAC)

After azido di-mannose building block **9** and peptide **11a** were successfully synthesized and purified, the preparation of the glycopeptides was performed using copper-catalyzed azide-alkyne cycloadditon (CuAAC). The reactions were carried out according to Conibear *et al.*, without any further optimization.[83] However, two different approaches were investigated: performing the click reaction in solution with the purified peptide as well as on resin with the crude peptide. The desired products were mono- and di-mannosylated peptides, which were then used as binding partners in subsequent SPR studies with cyanovirin-N.

4.4.1 In-solution CuAAC

4.4.1.1 Acetylated azido-mannose (1)

As a proof of concept, we started our investigation with CuAAC reactions of purified peptide **11a** and azido mannose building block **1** (Figure 24). The peptide was dissolved in DMF, mixed with the sugar, followed by the addition of the CuAAC cocktail, containing Cu^{2+}, which was *in-situ* reduced to Cu^{+} using sodium ascorbate, as well as the stabilizing ligand TBTA under inert atmosphere to prevent reoxidation of the copper. The reduction and generation of the Cu(I) species allows regioselectivity according to Sharpless and Meldal.[52,53]

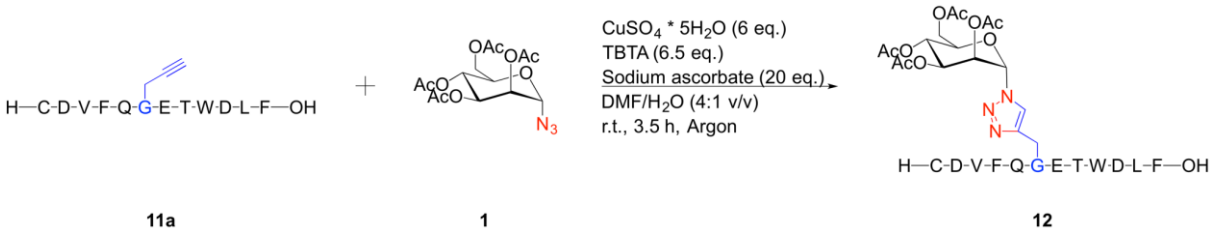


Figure 24 CuAAC of peptide **11a** and acetylated azido mannose **1**, yielding glycopeptide **12**.

The reaction was monitored *via* LC-MS, where full conversion was observed after 3.5 hours. Purification was challenging since TBTA has shown to be difficult to separate from the product. As a result, one preparative LC-MS run, followed by another HPLC purification were necessary to remove TBTA and obtain glycopeptide **12** in good purity (Figure 25).

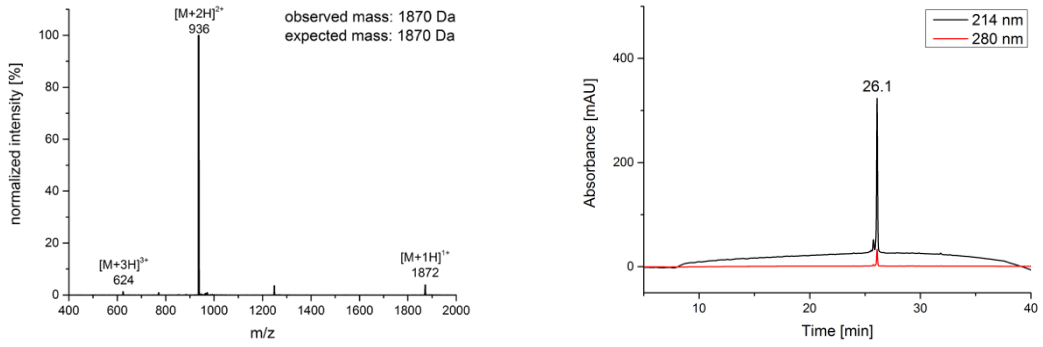


Figure 25 Mass spectrum (left) and HPLC chromatogram (right) of protected mono-mannosylated peptide **12**. $[M+H]^{+}_{calc}$ = 1871.9 Da, found 1871.6 Da; $[M+2H]^{2+}_{calc}$ = 936.5 Da, found 936.1 Da; $[M+3H]^{3+}_{calc}$ = 624.7 Da, found 624.4 Da. Retention time (C4 column): 26.1 min.

Initial reactions were performed on small scales (0.5-2 mg) and scaled up afterwards. The larger scale reaction using 11 mg of peptide **11a** (7.34 µmol, 1 equiv.) and 2.5 equiv. of sugar **1** led to 2.7 mg glycosylated peptide **12** (1.44 µmol, 20 % reaction yield and 5 % overall yield, Figure 25) after 2 purification runs (preparative LC-MS and HPLC). Considering the challenging purification, the product was isolated in good yield and purity.

After the successful glycosylation, the sugar needed to be deprotected since the peptide should mimic the natural binding partner. Therefore, peptide **12** was deprotected using two approaches: hydrazine in methanol as well as sodium methoxide in methanol and DMF (Figure 26). Monitoring the reaction *via* LC-MS revealed that the deprotection of glycopeptide **12** required 3 hours when using hydrazine, while complete deprotection of glycopeptide **12** was obtained after only 15 minutes when using sodium methoxide in methanol and DMF.

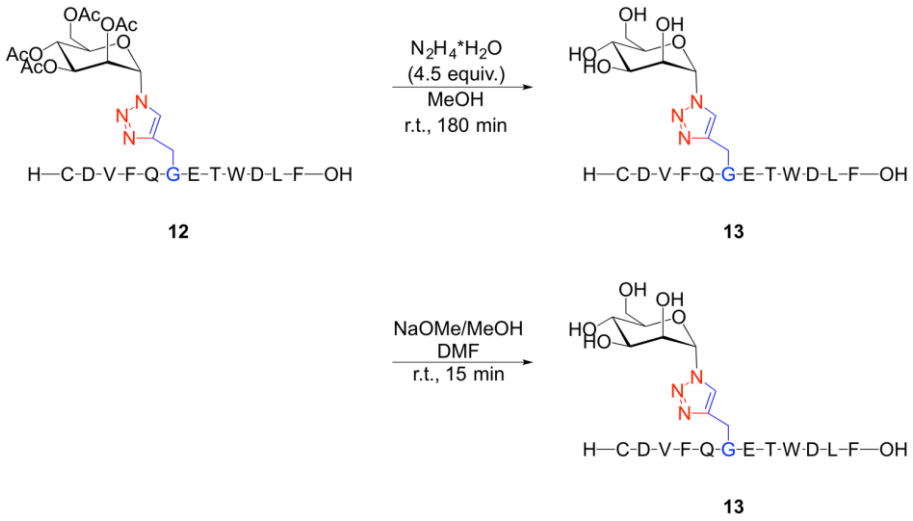


Figure 26 Deprotection approaches for the preparation of glycopeptide **13**.

Interestingly, an additional peak at 770 Da appeared in the MS spectra, but only after deprotection (Figure 27). The observed mass would fit the $[M+2H]^{2+}$ peak of the peptide, where the sugar is cleaved off. However, HPLC showed only one peak and should be shifted to slightly higher retention times if the carbohydrate would be missing.

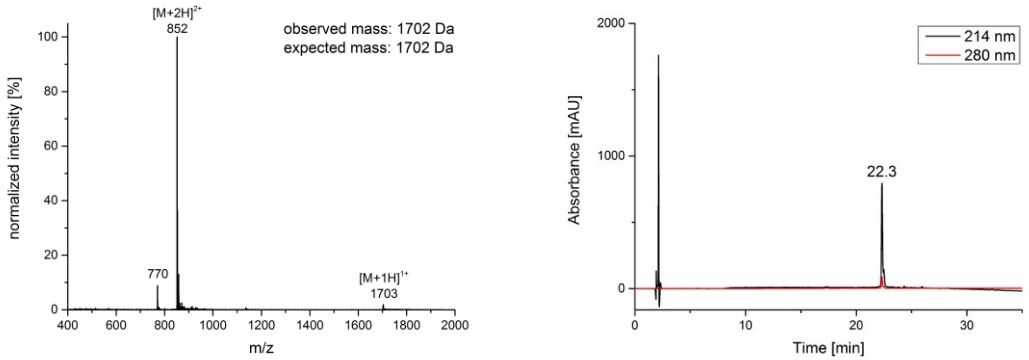
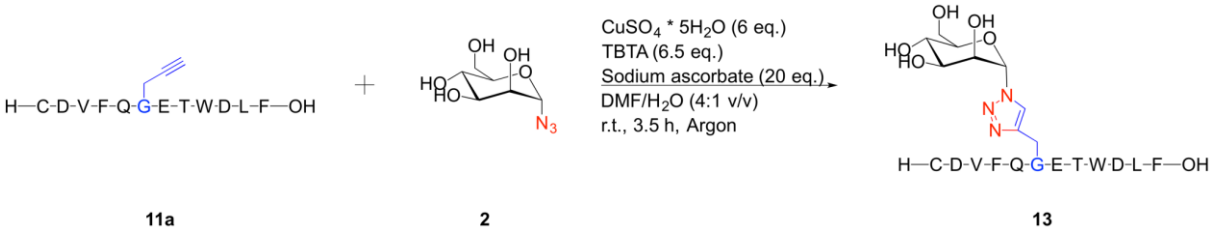


Figure 27 Mass spectra (left) and HPLC chromatogram (right) of deprotected glycopeptide **13**. $[M+H]^{+}_{calc}$ = 1703.8 Da, found 1703.5 Da; $[M+2H]^{2+}_{calc}$ = 852.4 Da, found 852.2 Da; $[M+3H]^{3+}_{calc}$ = 568.6 Da, found 568.7 Da. Retention time (C18 column): 22.3 min.

Therefore, the peptide was first analyzed using MALDI-MS, where no significant peak at 770 Da was detected and by our mass spectrometry core facility, where no notable peak in this area was observed as well. The appearance of this 770 Da peak using the Thermo Fisher Scientific LC-MS device is therefore probably caused by fragmentation during ionization.

The reactions were carried out on a very small scale as a proof of concept and besides detection of the final product, no yield was determined. Since three purification steps would be necessary to isolate pure glycopeptide **13**, the next approach using already deprotected azido mannose **2** for the CuAAC reaction was chosen (4.4.1.2).

4.4.1.2 Deprotected azido-mannose (2)



11a **2** **13**

Figure 28 CuAAC of peptide **11a** and deprotected azido-mannose **2**, yielding glycopeptide **13**.

Based on the conclusions above, deprotected azido sugar **2** was used for the CuAAC reaction to shorten the synthesis by one purification step compared to 4.4.1.1 (Figure 28). Reaction conditions were kept identical as for the protected sugar **1**. An initial small scale reaction showed similar reaction progress as for the CuAAC reaction with sugar **1**. Therefore, the reaction was scaled up using 11 mg of peptide **11a** (7.34 µmol, 1 equiv.) with 2.5 equiv. of sugar **2** (3.8 mg, 18.4 µmol).

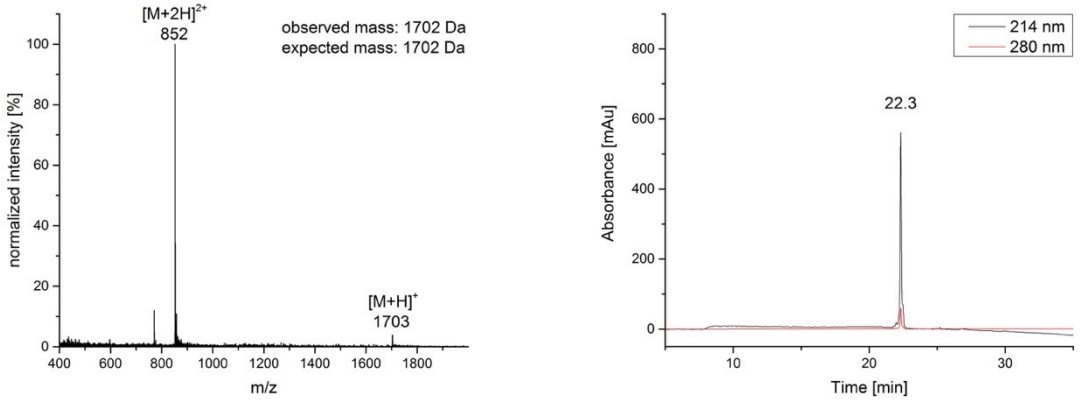


Figure 29 Mass spectra (left) and HPLC chromatogram (right) of deprotected glycopeptide **13**. $[M+H]^{+}_{calc}$ = 1703.8 Da, found 1703.5 Da; $[M+2H]^{2+}_{calc}$ = 852.4 Da, found 852.2 Da; $[M+3H]^{3+}_{calc}$ = 568.6 Da, found 568.7 Da. Retention time (C18 column): 22.3 min.

However, single-step LC-MS purification of the quenched reaction yielded 2.3 mg of TBTA containing product **13** and further side pools, mainly containing TBTA. The following HPLC purification allowed the isolation of 1.2 mg pure mono-glycosylated peptide **13** (0.7 µmol, 2 % overall yield, Figure 29). While the preparation and isolation were successful, the obtained

yield and purification were unsatisfying and led to the idea of performing CuAAC reactions directly on resin (4.4.2).

4.4.2 On-resin CuAAC

Driven by the challenging separation of our initial CuAAC reaction products from the ligand TBTA, in addition to the fact that the reaction solution (75-80 % DMF) required dilution with water to prevent damages of the purification column, we aimed for a higher yielding purification that limited the significant loss of product. By performing the CuAAC reaction with the fully protected and uncleaved peptide, obtained after automated solid-phase peptide synthesis (see chapter 4.1) and a corresponding sugar azide, the reaction solution could be easily removed *via* filtration (Figure 30).

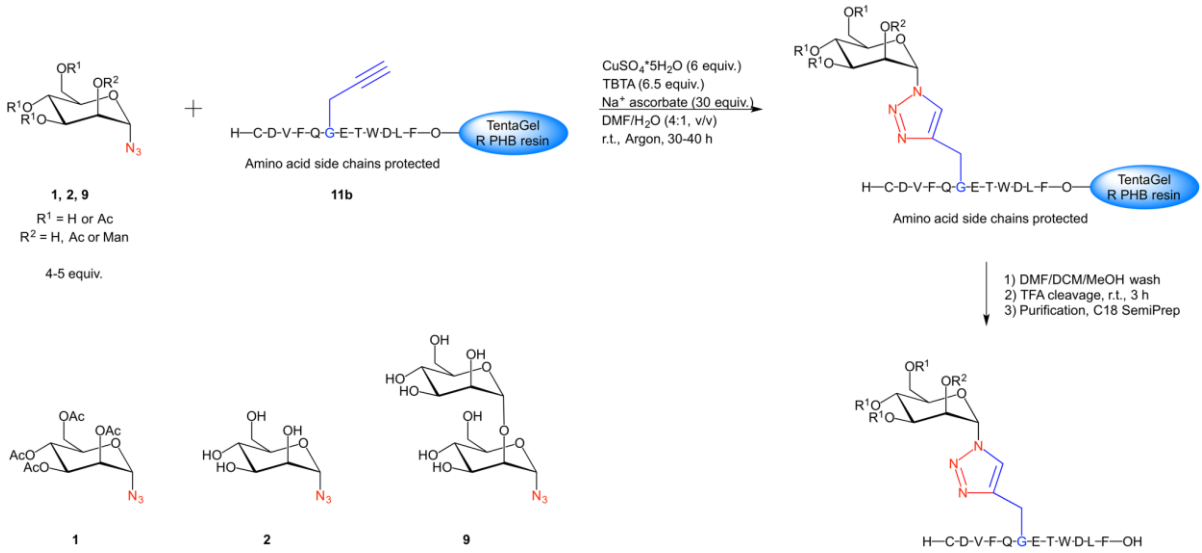


Figure 30 General procedure for on-resin CuAAC using peptide **11b** and several azido sugars.

Due to lower accessibility of the reactive alkyne moiety in peptide **11b**, caused by protected side chains and the sterically hindered nature of resin-bound peptides, the reaction time increased. Monitoring the progress required test cleavages of resin aliquots, taken out of the reaction suspension. Following filtration, additional washing steps using DMF, DMF/H_2O (9/1, v/v), DCM and MeOH were performed to remove any TBTA and salts. The peptide was completely cleaved off the dried resin using a standard TFA cleavage cocktail (see methods) and the crude peptide purified *via* preparative LC-MS. This approach required only one purification step starting from the synthesized peptide (SSPS) until the isolation of purified glycopeptides. Cleavage conditions were compatible with protected as well as unprotected azido mannose derivatives (**1, 2** and **9**) since no significant deprotection or sugar cleavage was observed *via* LC-MS. While this approach increased reaction times and equivalents of required azido sugar (4-5 equiv.), the purification was simplified, overall production time of glycopeptides decreased and yield was increased significantly.

4.4.2.1 Deprotected azido-mannose (2)

Using the described on-resin approach for the synthesis of glycopeptide **13**, small scale reactions revealed the necessity of increased reaction times (30 h) and higher loading of azido sugar (5 equivalents, Figure 31). Nevertheless, the proposed simplification of purification due to washing off TBTA and DMF allowed straightforward access to the isolated glycopeptide **13**. After only one preparative LC-MS run, the product was obtained in good purity (Figure 32 and 33) and yields (3.5 mg, 2.1 µmol, 19 % overall yield, based on resin (68 mg)). Comparing the yield of isolated product obtained from in-solution and on-resin CuAAC, where 1.5 mg pure product (0.9 µmol) was obtained from 226 mg resin with the in-solution approach, the on-resin CuAAC approach is significantly more efficient. As already discussed in chapter 4.4.1, an additional peak at 770 Da was identified using the LC-MS system.

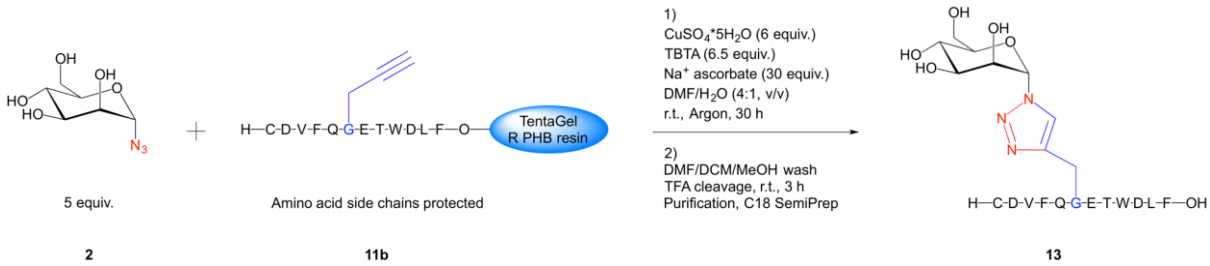


Figure 31 On-resin CuAAC approach for the synthesis of glycopeptide **13**.

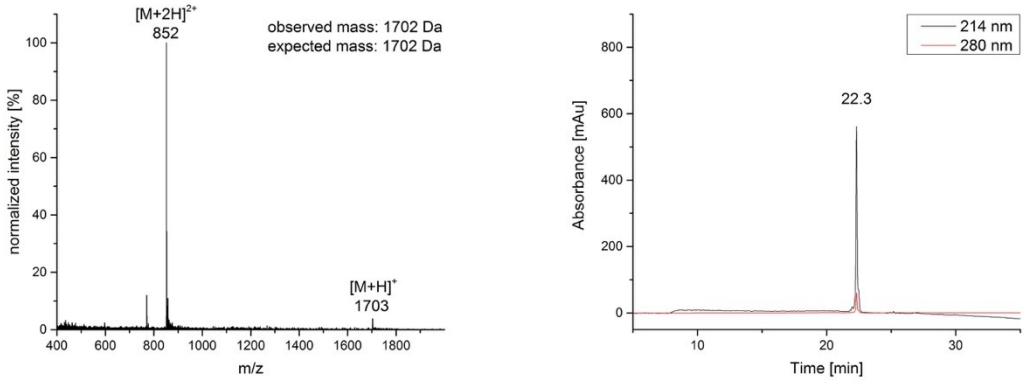


Figure 32 Mass spectra (left) and HPLC chromatogram (right) of deprotected glycopeptide **13**. $[M+H]^{+}_{calc}$ = 1703.8 Da, found 1703.5 Da; $[M+2H]^{2+}_{calc}$ = 852.4 Da, found 852.2 Da; $[M+3H]^{3+}_{calc}$ = 568.6 Da, found 568.7 Da. Retention time (C18 column): 22.3 min.

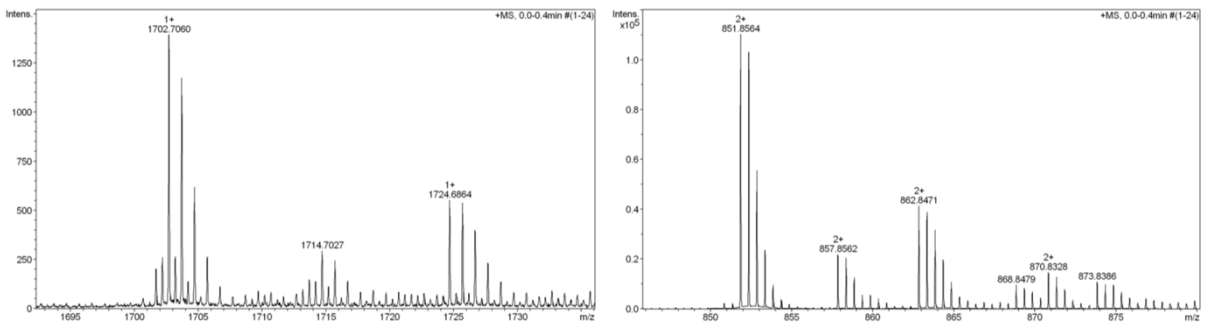


Figure 33 HRMS spectra of deprotected glycopeptide **13**. $[M+H]^{+}_{calc}$ = 1702.7059 Da, found 1702.7060 Da; $[M+Na]^{+}_{calc}$ = 1724.6879 Da, found 1724.6864 Da; $[M+2H]^{2+}_{calc}$ = 851.8569 Da, found 851.8564 Da; $[M+H+Na]^{2+}_{calc}$ = 862.8479 Da, found 862.8471 Da.

4.4.2.2 Deprotected di-mannose azide (9)

Preparation of di-mannosylated peptide was achieved using di-mannose azide **9** and on-resin CuAAC (Figure 34). The equivalents of sugar were reduced due to its limited amount in stock, which could be an explanation for the increased reaction time (40 hours). Nevertheless, after only one preparative LC-MS, the desired product was isolated in good purity (Figure 35 and 36) and yields (2 mg, 1.1 µmol, 10 % overall yield, from 68 mg resin). As described, the obtained yield for glycopeptide **14** was significantly higher than for the in-solution approach.

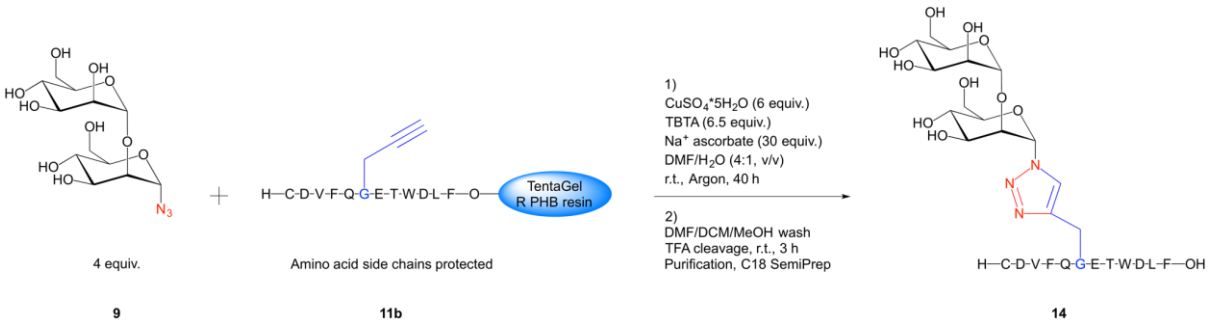


Figure 34 On-resin CuAAC approach for the synthesis of glycopeptide **14**.

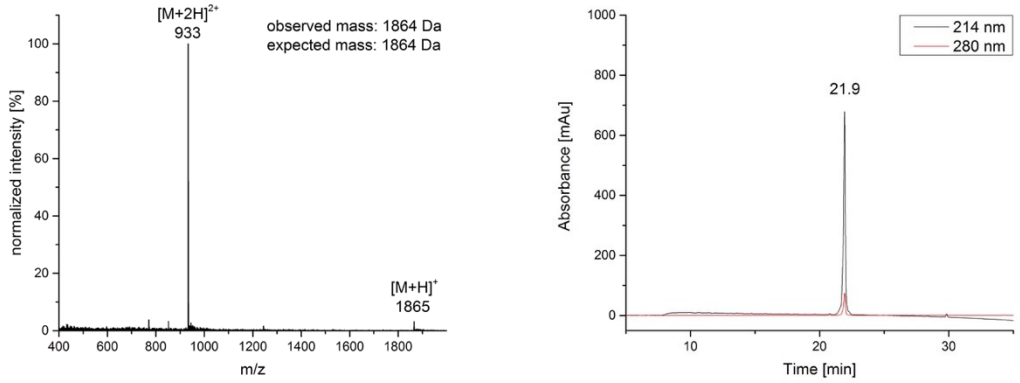


Figure 35 Mass spectra (left) and HPLC chromatogram (right) of deprotected glycopeptide **14**. $[M+H]^{+}_{calc}$ = 1865.9 Da, found 1865.8 Da; $[M+2H]^{2+}_{calc}$ = 933.5 Da, found 933.1 Da; $[M+3H]^{3+}_{calc}$ = 622.7 Da, found 622.4 Da; $[2M+3H]^{3+}_{calc}$ = 1244.3 Da, found 1243.8 Da. Retention time (C18 column): 21.9 min.

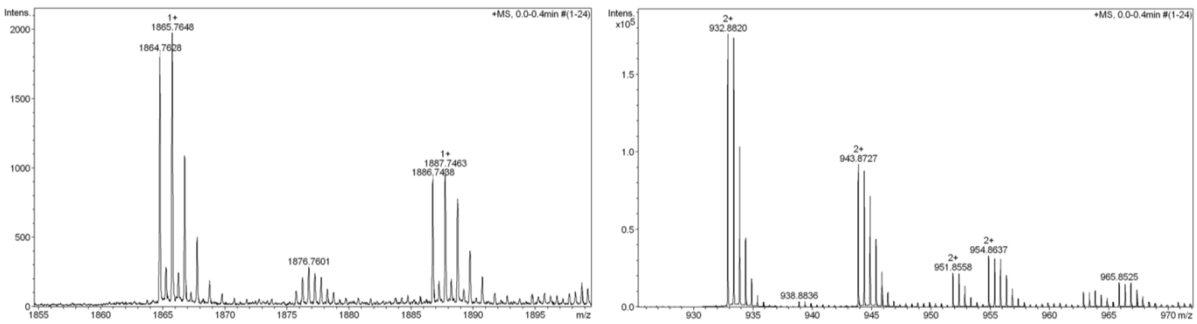


Figure 36 HRMS spectra of di-mannosylated peptide **14**. $[M+H]^{+}_{calc}$ = 1864.7587 Da, found 1864.7628 Da; $[M+Na]^{+}_{calc}$ = 1886.7407 Da, found 1886.7488 Da; $[M+2H]^{2+}_{calc}$ = 932.8833 Da, found 932.8820 Da; $[M+H+Na]^{2+}_{calc}$ = 943.8743 Da, found 943.8727 Da.

4.4.3 Comparison of in-solution and on-resin approach

While mono- and di-glycosylated peptides **13** and **14** were isolated using both in-solution and on-resin CuAAC approaches, the latter one has proven to be more convenient and to provide cleaner products at higher yields. On the other hand, more equivalents of azido sugars were required, which turned out to be critical for di-mannose azide **9** due to its limited availability. Furthermore, the reaction time increased for the on-resin approach. However, taking the intermediate isolation of peptide **11a** into account, the overall production time was quite similar. The removal of TBTA and DMF has shown to be highly efficient using the on-resin approach, while the purification step of the in-solution approach resulted in contaminated fractions and possible TBTA precipitation onto the column as well as potential damages caused by injecting DMF.

In conclusion, the desired glycopeptides **13** (mono-mannosylated) and **14** (di-glycosylated) were prepared in good purity and yield by the on-resin approach. Nevertheless, both strategies allowed the isolation of the desired products.

4.5 Protein expression

For the recombinant expression of (CV-N)2L0 (designed dimer according to PDB-ID: 3S3Y, no linker (L) region in between) and its variants, pET27b+ vectors containing the inserts were produced by GenScript and kindly provided by Dr. Irene Maier. Transformation of these vectors was performed in *E. coli* (BL21 competent cells), followed by protein expression, purification and analysis. The isolated proteins were then used as analytes for studies on binding of CV-N and disulfide-bond variants to the immobilized glycopeptides (as described in 4.4.) and HA (from insect cells). In this context, the term 'wild type' refers to (CV-N)2L0 without any modification of the disulfide bonds. Cysteine residues were mutated to reduce the numbers of disulfide bonds. For example, the sequence of variant 2 was modified to allow only 3 and the ones for V3-5 only 2 disulfide bonds, respectively.

4.5.1 Transformation

For incorporation of the vectors into *E. coli*, chemical transformation (heat shock) was performed. The obtained vectors from GenScript were re-checked using restriction enzymes, and agarose gel electrophoreses, verifying the length of the expected insert and the remaining vector (see SI). After transformation, colonies were selected and transferred onto a master plate. While variant 1 had shown only minor colony growth, the other variants showed acceptable amounts of colonies. After the transformation was accomplished, single colonies were used for inoculation and protein expression.

4.5.2 Expression and purification

Preliminary results have shown that protein expression was most efficient using overnight cultures (37°C, 200 RPM), which were then diluted to an OD_{600} of 0.05-0.10, followed by cell growing to an OD_{600} of about 0.5 and cooling to 18°C. The cultures were then treated with 500 µM IPTG to induce protein expression, which was carried out overnight. After cell lysis, Western blots targeting 6xhis-tagged proteins showed the expression of the desired proteins. However, positive Western blot signals were only obtained for variants 2-5 and (CV-N)2L0. Therefore, the focus was on the large scale expression and purification of these variants. Unfortunately, the majority of protein was found in the pellet (inclusion bodies) and only weak signals were detected in the supernatant. Nevertheless, the proteins were first purified from the soluble fraction using Ni-NTA (see methods) due to the easier handling. The proteins were purified in good purity, although isolation yields were low (SDS-PAGE shown in Figure 38, right). The wild type ((CV-N)2L0) was obtained at 0.5 mg per L culture, V2 at 0.2, V3 at 0.6, and V5 at 0.6 mg per L culture, respectively. The isolated amounts of protein were sufficient for preliminary SPR measurements using immobilized HA on the sensor chip and the expressed proteins as mobile binding partners (analytes). However, aspired NMR studies

investigating the interaction of the synthesized glycopeptides and the expressed proteins required the production of more protein from these variants.

We decided to purify the proteins of interest from inclusion bodies. The insoluble fraction was denatured using guanidinium buffer. The purification was performed *via* computer-assisted Ni-NTA affinity chromatography. Fractions were collected with increasing concentrations of imidazole and then analyzed using SDS-PAGE. Prior to that, guanidinium needed to be removed due to its interfering effect during electrophoresis. Therefore, the protein was precipitated using TCA or EtOH, followed by discarding the supernatant, washing of the pellet and resolubilization of the protein in water and SDS-buffer. An example for the purification of the (CV-N)2L0 is shown on the SDS-PAGE gel in Figure 37. Gels of the other 4 proteins showed similar results as the desired band (23 kDa) was observed.

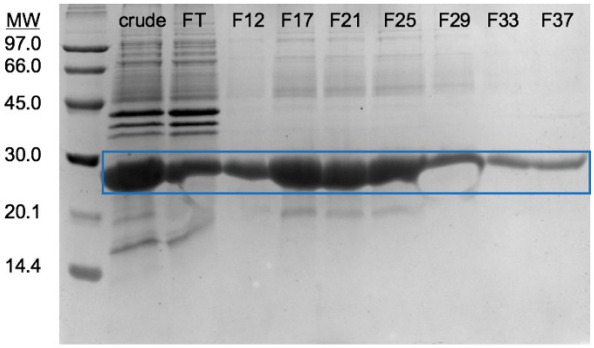


Figure 37 SDS-PAGE gel of (CV-N)2L0 Ni-NTA fractions, desired protein band highlighted in blue.

Following the Ni-NTA affinity chromatography, proteins were dialysed against 100 mM PBS buffer (pH 7), renewing the dialysis solution three times in order to obtain a sufficient buffer exchange. While some product was lost due to protein precipitation, sufficient amounts were remaining. The occurred precipitation was probably caused by the strong gradient between the protein solution (6 M GuaHCl) and the dialysis buffer, resulting in improperly folded protein. Diluting the protein solution to lower GuaHCl concentration could reduce the risk of precipitation.

4.5.3 Analysis

After dialysis, the obtained protein solutions were analyzed *via* SDS-PAGE (Figure 38, left), showing that the higher amounts of proteins which were purified from inclusion bodies were of similar purity as the ones obtained from the soluble fraction (Figure 38, right).

In addition, purification from inclusion bodies allowed the isolation of V4. While other bands are slightly visible, the CV-N band shows a much stronger signal. Determination of protein concentration using NanoDrop showed very high yields for the proteins obtained from inclusion bodies. This confirms the observations obtained during Western blotting of the test expressions, which indicated that most of the protein is located in the insoluble fraction and

only partially in the soluble fraction. While expression conditions can be further optimized, even using 1 % Triton-100x did not lead to a notable transfer of protein into the soluble fractions.

Protein	Yield from pellet (per L culture)	Yield from supernatant (per L culture)
(CV-N)2L0	15 mg/L	0.5 mg/L
V2	14 mg/L	0.2 mg/L
V3	2.6 mg/L	0.6 mg/L
V4	2.3 mg/L	n.d.
V5	19 mg/L	0.6 mg/L

Table 2 Comparison of isolated yields obtained from purification of insoluble and soluble fraction, respectively.

Although purification of the soluble fraction would be preferred due to fewer steps, the desired proteins were isolated in high yields from the insoluble fraction. As stated in Table 2, this represents much higher yields than the ones obtained from the soluble fractions. SDS-PAGE of the obtained proteins showed sufficient purity for the following SPR studies (Figure 38).

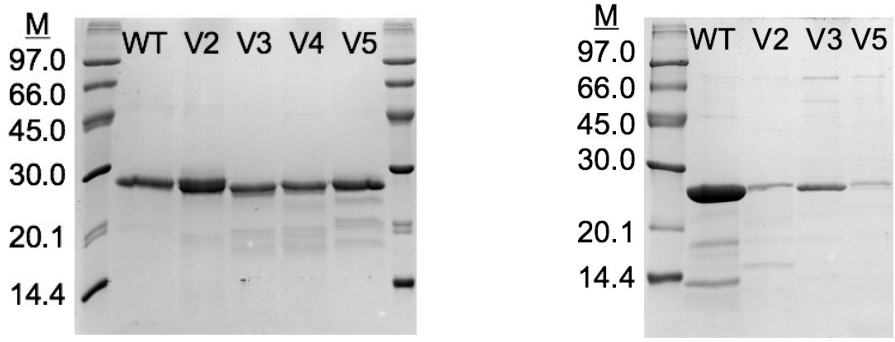


Figure 38 SDS-PAGE of purified CV-N variants and (CV-N)2L0 from inclusion bodies (left) and soluble fraction (right).

For final analysis, the isolated proteins were further analyzed using MS and HPLC (Figure 39-43). (CV-N)2L0 shows a sharp peak and good purity in the HPLC chromatogram (Figure 39), while the calculated mass obtained from the charge state distribution is slightly lower than expected. It could result from a cleaved methionine found at the *N*-terminus, which can occur if the second residue in the sequence is uncharged and small (Met-Gly).[84,85] The isolated proteins (V3-5), obtained from the plasmids provided by our collaboration partners, show even sharper peaks in the HPLC chromatograms and good charge state distributions (Figure 41-43), indicating the isolation of the desired proteins.

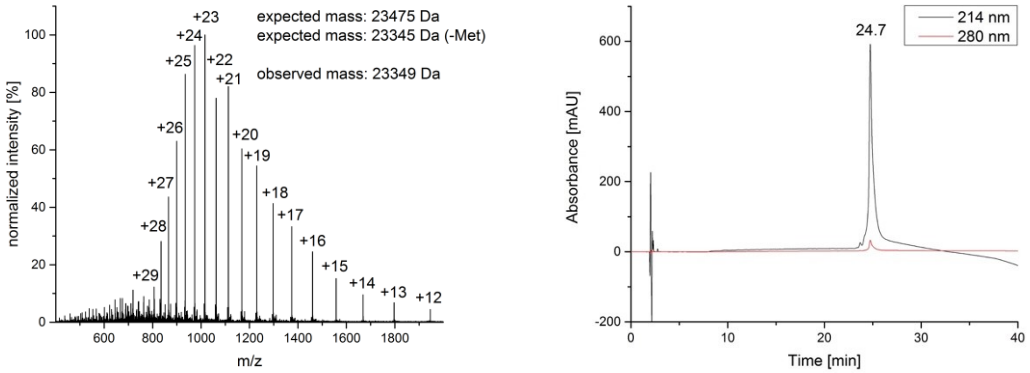


Figure 39 Mass spectrum (left) and HPLC chromatogram (right) of (CV-N)2L0.

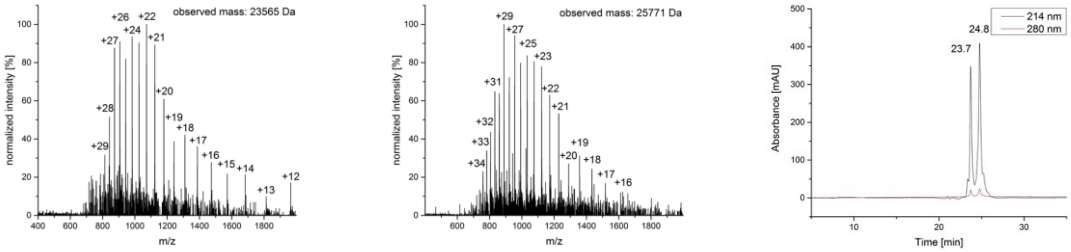


Figure 40 MS spectra of the LC-MS peaks at 23.7 min (left) and 24.8 min (middle) as well as HPLC chromatogram (right) of V2.

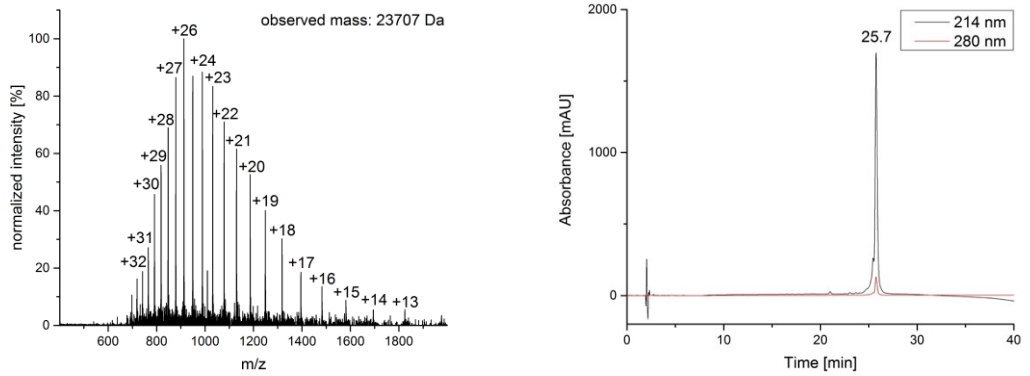


Figure 41 Mass spectrum (left) and HPLC chromatogram (right) of V3.

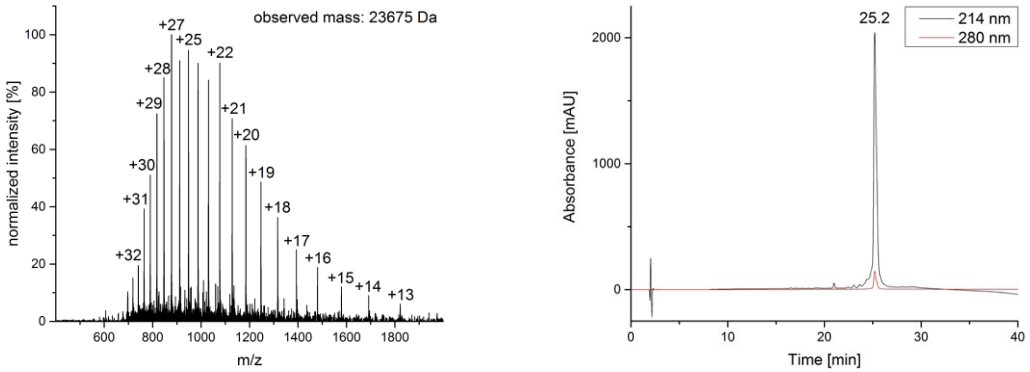


Figure 42 Mass spectrum (left) and HPLC chromatogram (right) of V4.

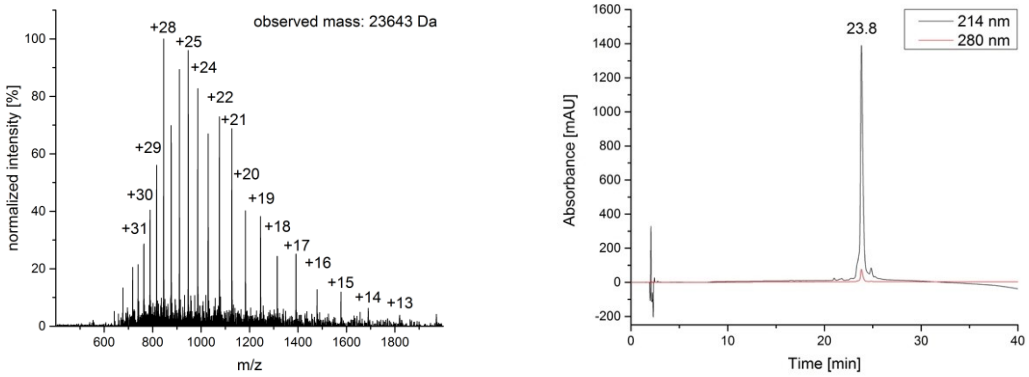


Figure 43 Mass spectrum (left) and HPLC chromatogram (right) of V5.

In addition, circular dichroism measurements were carried out for (CV-N)2L0 and variants V2-4. As shown in Figure 44, V3 and V4 show similar secondary structures to (CV-N)2L0, which in turn was compared to the PDB-ID: 3S3Y (Table 3). The main secondary structures are beta-strands with some influence of alpha-helices (Figure 44 and Table 3). The percentages of helical component were lower for V3 and V4 as compared with (CV-N)2L0 (analyzed in this study) and the published protein from the PDB data.

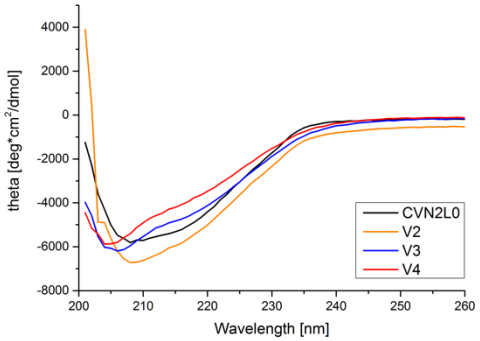


Figure 44 CD spectrum of (CV-N)2L0 and its variants at concentrations of 20 µM ((CV-N)2L0, V3-4) and 25 µM (V2), respectively.

Protein	% helix	% strand	% turn + others	No. –S-S– bonds
3S3Y (PDB-ID)*	6.1	27.2	57.7	4
(CV-N)2L0	7.5	37.4	51.1	4
V2§	20.2	25.9	54.0	3
V3	4.6	33.7	51.7	2
V4	5.0	33.2	51.8	2

Table 3 Composition of secondary structure elements of reference protein 3S3Y and isolated variants. Obtained CD data were analyzed from 200-250 nm using the BeStSel programme.[86] * Sequence and secondary structure data were taken from the PDB file (3S3Y). § Data may differ due to the inhomogeneity of the isolated protein species.

Variant 2, however, has shown to produce slightly different signals. This can result from another protein species present in solution (see Figure 40). Since this variant is supposed to

form 3 disulfide bridges, it may be a mixture of proteins with 2 and 3 disulfide bonds. Together, as observed *via* SDS-PAGE and CD, (CV-N)2L0 and variants V3-4 appear to be sufficiently clean products. While the SDS-PAGE analysis of V2 appeared to show a clean product, CD, HPLC and MS analysis have shown the presence of another protein with a mass about 2 kDa higher than the expected one (Figure 40). At this point, the additional protein cannot be sufficiently explained, it could be another protein with affinity towards the Ni-NTA column. This needs to be further studied and no final conclusions can be drawn from the current data. Variant 5 was also obtained as a clean product according to HPLC and MS (Figure 43) but not further investigated in this study.

4.6 Surface Plasmon Resonance

After expression and purification of the desired proteins as well as after synthesis and isolation of the glycopeptides (mono-mannosylated and di-glycosylated), the final SPR studies were performed. Therefore, HA (derived from insect cells) and the two glycopeptides were immobilized on different chips. The expressed and isolated (CV-N)2L0 and its variants were then analyzed as mobile binding molecule (analytes), assuming a 1:1 binding model. Binding to HA was used as a reference for the interpretation of binding to the peptide chips and for comparison of (CV-N)2L0 with its isolated variants.

4.6.1 HA immobilized, CV-N and variants as analytes

After the immobilization of HA, (CV-N)2L0 and variants V2-V4 were tested as potential binder molecules. After injection of the analyte, concentration-dependent association to the ligand (HA), followed by equilibrium (i.e. for (CV-N)2L0) and dissociation were observed. The obtained sensorgrams are shown below (Figures 45-46 and Figures 77-78) and the kinetics data presented in Table 4. K_D was calculated as the quotient of k_{off}/k_{on} (k_{off}: rate of dissociation, k_{on}: rate of association). Affinity was calculated at equilibrium, assuming a 1:1 binding model, and the binding isotherm (Figure 45-46, right) was calculated through the fitted data points, based on R_{max} and K_D.

(CV-N)2L0

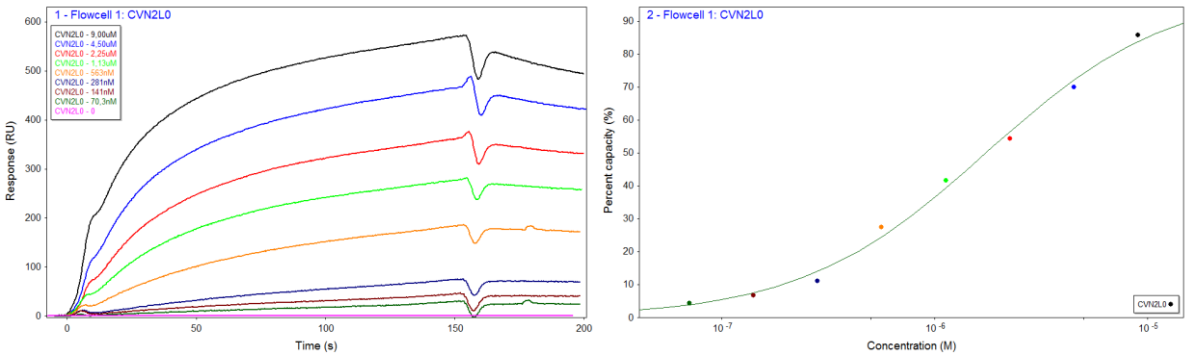


Figure 45 SPR sensorgram (left) and binding capacity (right) of (CV-N)2L0 as analyte and HA as immobilized binding partner. Affinity from equilibrium binding responses (right).

V2

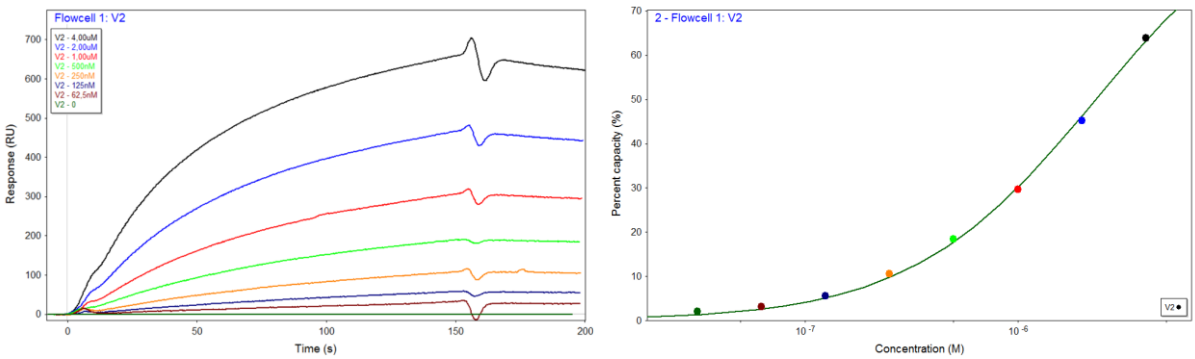


Figure 46 SPR sensorgram (left) and binding capacity (right) of V2 as analyte and HA as immobilized binding partner. Affinity from equilibrium binding responses (right).

sample	Kinetics		
	k_{on} [M^{-1} s^{-1}]	k_{off} [s^{-1}]	K_D [nM]
CVN2L0	4.74 e3	2.8 e-3	591
V2	4.42 e3	2.3 e-3	520
V3	1.20	0.625	2.4 e6
V4	5.97 e3	0.711	119 e3

Table 4 Kinetics data obtained for (CV-N)2L0 and variants as analytes against immobilized HA as binding partner. $K_D = k_{off}/k_{on}$

Comparing the kinetics data of the expressed and isolated proteins, (CV-N)2L0 and V2 show nanomolar affinity against the immobilized HA, which is in a similar range as CV-N activity measured against HIV in neutralization assays.[25] Variant 4, however, shows only micromolar activity against HA and V3 even only millimolar. Since (CV-N)2L0 possesses 4 disulfide bonds, V2 possesses 3 and V3-4 possess only 2 disulfide bonds, these results could indicate the importance of disulfide bonds for activity. Regarding the interaction of the expressed proteins with the synthesized and immobilized di-glycosylated peptide, binding of all isolated proteins was shown in micromolar affinity (see 4.6.2, Table 5).

4.6.2 Di-mannosylated peptide immobilized, CV-N and variants as analytes

After immobilization of the di-mannosylated peptide, (CV-N)2L0 and variants 2 and 4 were tested as potential binder molecules. Sensorgrams of (CV-N)2L0 and V2 (Figure 47 and 48) are shown and the kinetics data presented in Table 5. Analysis of the obtained data is based on a 1:1 binding model.

(CV-N)2L0

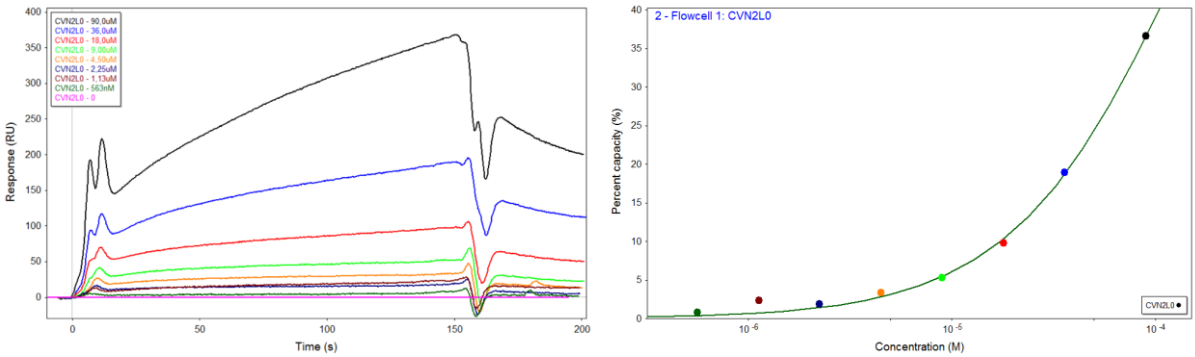


Figure 47 SPR sensorgram (left) and binding capacity (right) of (CV-N)2L0 as analyte and di-glycosylated peptide as immobilized binding partner. Affinity from equilibrium binding responses (right).

V2

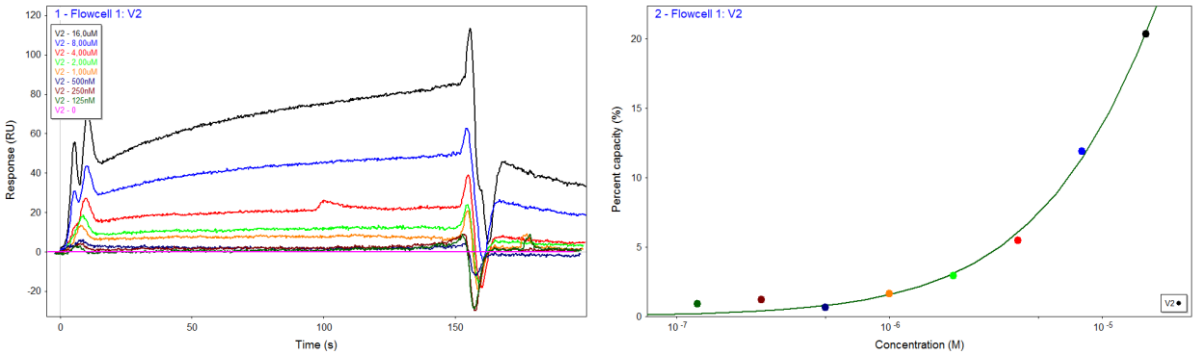


Figure 48 SPR sensorgram (left) and binding capacity (right) of V2 as analyte and di-glycosylated peptide as immobilized binding partner. Affinity from equilibrium binding responses (right).

	Kinetics		
sample	k_{on} [M^{-1} s^{-1}]	k_{off} [s^{-1}]	K_D [nM]
CVN2L0	167	1.5e-2	90 e3
V2	892	3.3e-2	37 e3
V4	755	4.0e-2	53 e3

Table 5 Kinetics data obtained for (CV-N)2L0, V2 and V4 as analytes against immobilized diglycosylated peptide as binding partner. $K_D = k_{off}/k_{on}$

While kinetics data of immobilized HA and (CV-N)2L0 and its variants as analytes presented analyzable sensorgrams, the data for the immobilized peptides was only partially analyzable (Table 5). Since peptides and HA differ in molecular weight, spatial extent, glycosylation richness (mono/di-mannose – high-mannose) as well as linkage (triazole – Asn-glycosylation), no distinct explanation regarding the influence of these properties on binding can be given. On the contrary, (CV-N)2L0 showed stronger binding when an additional mannose unit was introduced to the peptide (**14**), indicating the importance of glycosylation on binding.

The affinity and kinetic studies towards the chip, immobilized with the glycosylated peptides, showed no significant preference among all investigated proteins. The results show that the affinity and kinetics in micromolar concentrations are significantly lower than reported data for fully glycosylated HA, which are shown above. However, when comparing the data from the mono-glyosylated and di-glycosylated peptide, the additional mannose-group shows a notable effect. While the obtained binding curves of protein to mono-glycosylated peptide show only slight signals and therefore no analyzable data (see next paragraph). Measuring the binding of proteins to the di-glycsoylated peptide allowed the calculation of binding affinity and kinetics (Table 5). As a reference, binding of all isolated proteins to immobilized HA was investigated.

4.6.3 Mono-mannosylated peptide immobilized, CV-N and variants as analytes

Binding to the immobilized mono-glycosylated peptide allowed only for qualitative binding data (data partially shown in supplementary information, 7.4.2, Figure 79). Therefore, no kinetics data could be generated since binding has shown to be hardly analyzable (signal to noise ratio insufficient).

5 Conclusions and Outlook

Within this project, binding studies with cyanovirin-N (4 disulfide bonds) and designed mutants that can only form 3 or 2 disulfide bonds were carried out against HA and synthetic glycopeptides. Peptides covering a partial sequence of HA and containing a propargylglycine modification were successfully synthesized using solid-phase peptide synthesis and isolated in good purities and yields. The glycosylation of the modified peptides was achieved using copper-catalyzed azide-alkyne click chemistry. Mannosylation of the peptide was obtained *via* a triazole-linkage, thereby mimicking *N*-glycosylation.

Since binding studies of the protein of interest have shown that CV-N binds specifically to Manα(1-2)Manα units, a biologically relevant model peptide needed to be generated. However, no reported data regarding the synthesis of C_1-N_3 functionalized Manα(1-2)Manα were available. Therefore, we developed a synthesis route, making this species available and successfully isolated this compound (Figure 49, compound **9**). Since the overall yield is very low (3 % after 5 steps), this route shows room for improvement. By starting the synthesis from mannose and generation of **3** using acetic anhydride and hydrobromic acid, followed by direct transformation to compound **4**, undesired hydrolysis of **3** can be limited. In addition, the selection of temperature for the preparation of **5** and further optimizing its purification, would probably allow the isolation in higher yields. With this additional building block in hand, CuAAC reactions were used to synthesize and isolate the second glycopeptide **14**, constituted of azido di-mannose directly linked to the peptide *via* a 1,4-disubstituted triazole ring. Mono- and di-mannosylated peptides were later on studied using SPR.

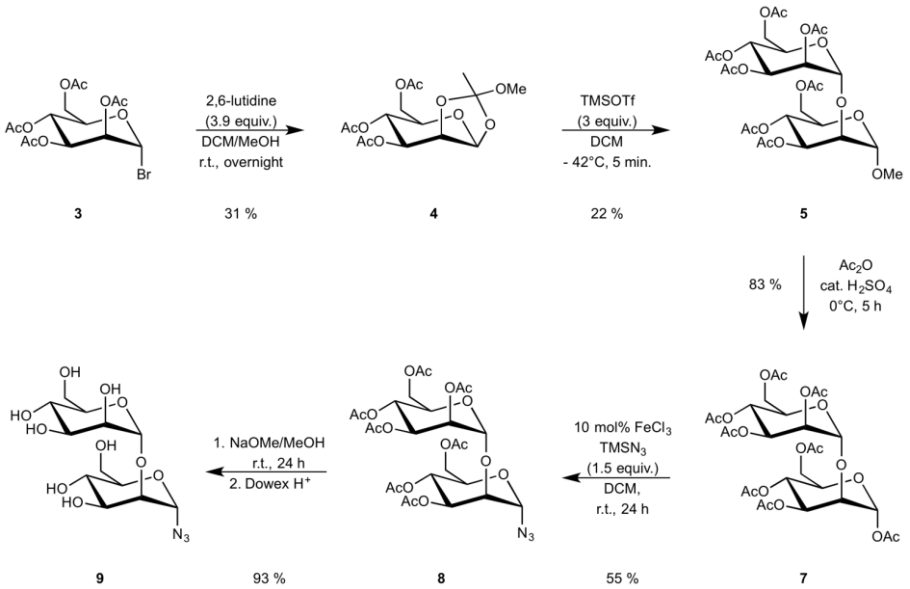


Figure 49 Developed synthesis route for synthesis of building blocks **9** and **8**.

Although the synthesis of our new building block and its application in click chemistry had been shown, the purification of the glycopeptides was unsatisfying due to TBTA

contamination. Therefore, we approached this problem by performing the CuAAC reaction with the peptide still bound to the resin (Figure 50). As a result, the reaction mixture was easily removed after the reaction had been quenched and only one purification step was required to obtain the clean glycopeptides at increased yields.

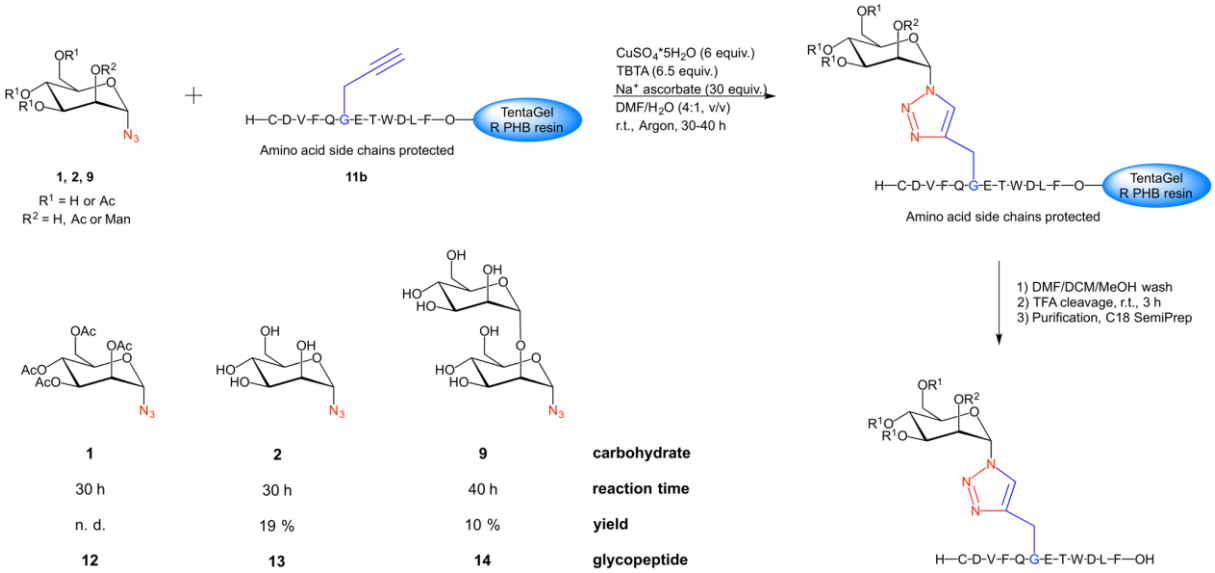


Figure 50 Alternate approach of CuAAC using on-resin conditions.

The expression and isolation of (CV-N)2L0 and its variants had shown to be successful for four out of six designed variants. Purification of the desired proteins was performed from the soluble as well as the insoluble fraction (inclusion bodies), resulting in much higher yields obtained from the insoluble fractions. While one variant (V2) has been identified to be a mixture of two proteins with slightly different masses and has to be analyzed further, (CV-N)2L0 and variants 3-5 were isolated in good purity.

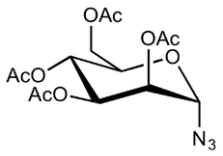
The final SPR studies revealed nanomolar affinities of (CV-N)2L0 and V2 (although not pure) towards immobilized HA, while V4 has shown micromolar and V3 millimolar binding affinity. This observation could be explained by the number of varying disulfide bonds of these proteins ((CV-N)2L0: 4, V2: 3, V3-4: 2 bonds) and could indicate the importance of disulfide bonds for anti-viral activity. Binding studies of our synthesized glycopeptides, which were immobilized, have shown lower affinities compared to HA. Binding of (CV-N)2L0 and its variants to the di-glycosylated peptide was shown in micromolar affinity. Moreover, V4 is binding either di-mannosylated peptide or HA with micromolar binding affinity, too. The immobilized mono-glycosylated peptide and binding of the variants to this peptide produced only qualitative data, where kinetics data were hardly analyzable, but estimated to be in millimolar range for (CV-N)2L0 and V2.

In conclusion, it has been shown that the number of disulfide bonds and the specific glycosylation can have a significant impact on protein binding affinities. These preliminary results will be further investigated in NMR experiments, where the interaction between glycopeptides and proteins will be further studied with much higher resolution.

6 Experimental

6.1 Carbohydrate synthesis

6.1.1 2,3,4,6-Tetra-*O*-acetyl-α-D-mannopyranosyl azide (1)



1

Figure 51 Acetylated mannose azide **1**.

Data for starting material **1** (commercially obtained and analyzed):

^{1}H NMR (600 MHz, $CDCl_3$) δ 5.39 (d, *J* = 1.6 Hz, 1H, H_1), 5.32-5.22 (m, 2H, H_4, H_3), 5.17-5.13 (m, 1H, H_2), 4.31 (dd, *J* = 12.5, 5.6 Hz, 1H, H_{6a}), 4.16 (m, 2H, H_{6b}, H_5), 2.17, 2.12, 2.06, 2.00 (4 x s, 12H, CH_3) ppm.

^{13}C NMR (151 MHz, $CDCl_3$) δ 170.8, 170.1, 169.9, 169.8 (4 x $COCH_3$), 87.6 (C_1), 70.7 (C_5), 69.3 (C_2), 68.4 (C_3), 65.7 (C_4), 62.3 (C_6), 21.0, 20.9, 20.8, 20.8 (4 x CH_3) ppm.

IR spectroscopy: $\tilde{\nu}$ = 2120 (N_3), 1753 (C=O), 1371, 1213, 1045, 957, 687 cm^{-1}.

ESI$^+$-HRMS (m/z) for $C_{14}H_{19}N_3O_9Na$ = 396.1019 $[M+Na]^+$; found: 396.1013 Da.

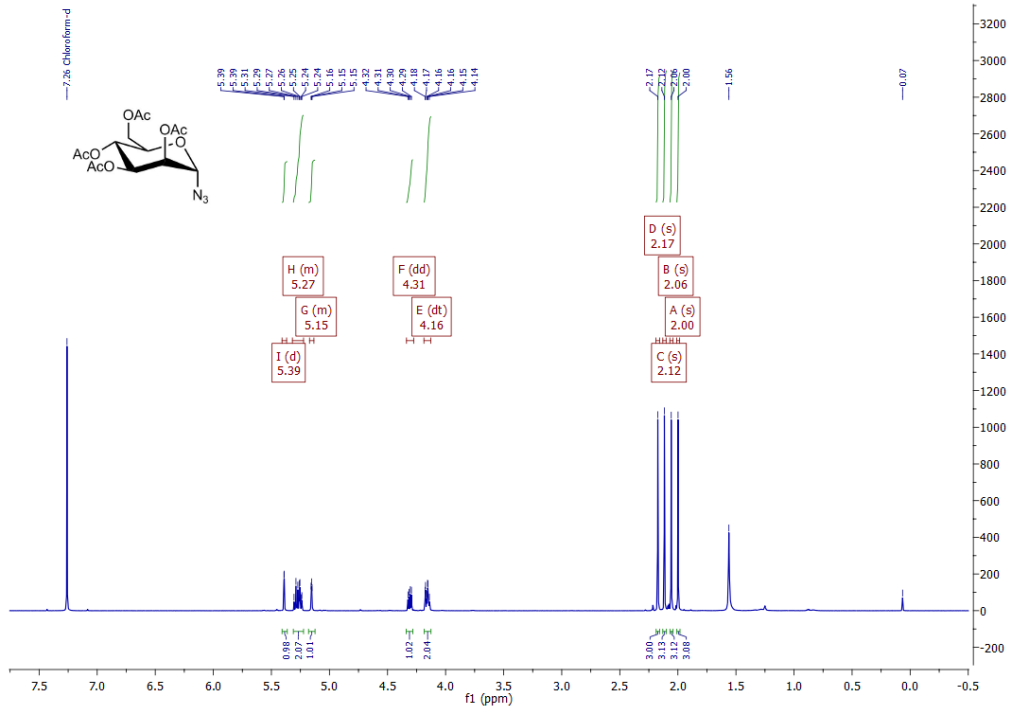


Figure 52 ^{1}H NMR spectrum of sugar **1**.

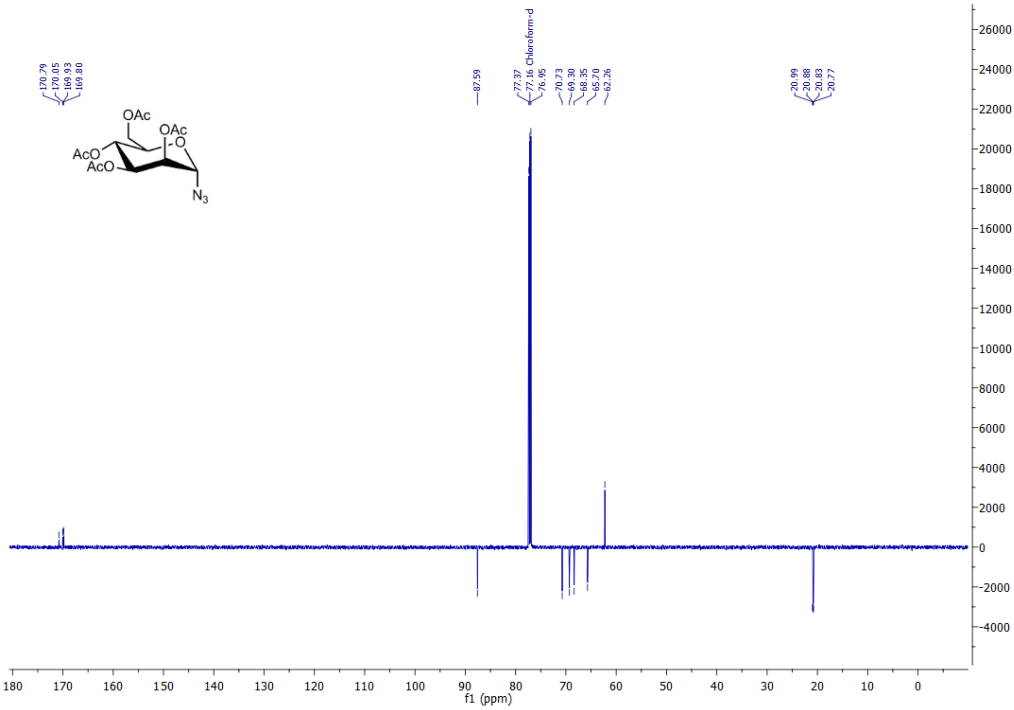
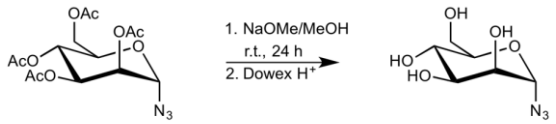


Figure 53 ^{13}C NMR spectrum of sugar **1**.

6.1.2a Synthesis of α-D-mannopyranosyl azide (2) using NaOMe/MeOH[72]



1 2a

Figure 54 Deprotection of acetylated mannose azide **1** – entry a.

In an oven-dried round-bottom flask, tetraacetylated mannopyranosyl azide **1** (56 mg, 0.15 mmol) was dissolved in dry methanol (1 ml) and sodium methoxide (0.5 M in MeOH, 60 µl) was added until pH 9-10. The solution was stirred at room temperature for 24 hours. The reaction mixture was neutralized by the addition of Dowex 50 H^+ until neutral pH was reached, filtered and lyophilized until complete dryness, isolating the deprotected azido-mannose **2** (28.3 mg, 0.14 mmol, 92 %) as a white solid. Obtained data in accordance with literature.[72,75]

^{1}H-NMR (600 MHz, D_2O) δ 5.41 (d, *J* = 1.8 Hz, 1H, H_1), 3.94-3.51 (m, 6H, H_2, H_3, H_4, H_5, H_{6a}, H_{6b}) ppm.

^{13}C NMR (151 MHz, D_2O) δ 89.7 (C_1), 74.5 (C_5), 69.7 (C_3), 69.7 (C_2), 66.3 (C_4), 60.7 (C_6) ppm.

IR spectroscopy: $\tilde{\nu}$ = 3337 (-O-H), 2934 (-C-H), 2122 (N_3), 1248, 1051, 939, 783, 650 cm^{-1}.

ESI$^+$-HRMS (m/z) for $C_6H_{11}N_3O_5Na$ = 228.0596 $[M+Na]^+$; found: 228.0594 Da.

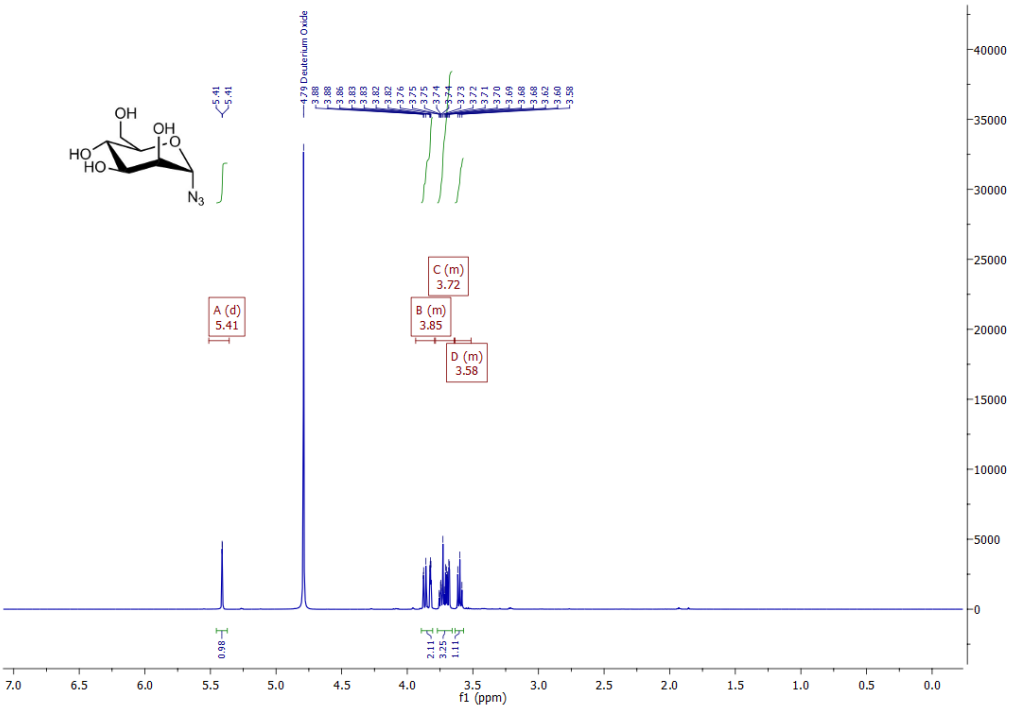


Figure 55 ^{1}H NMR spectrum of sugar **2**.

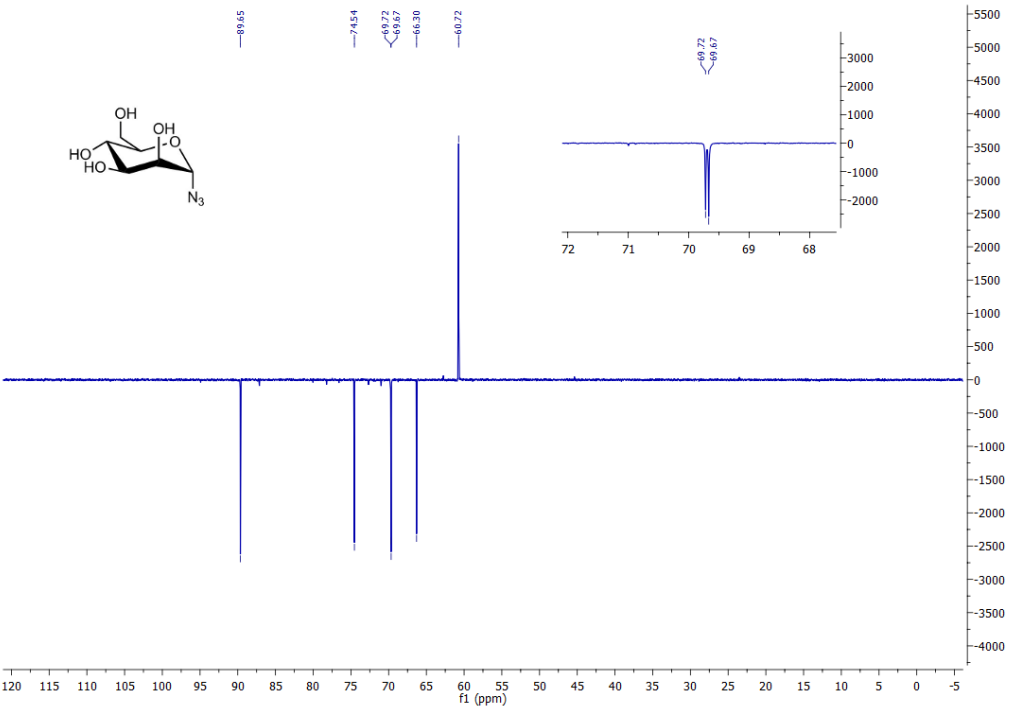
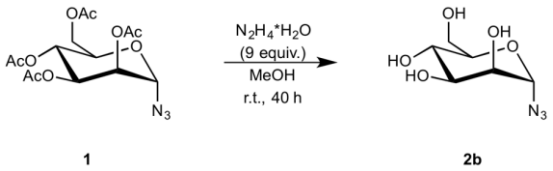


Figure 56 ^{13}C NMR spectrum of sugar **2**.

6.1.2b Synthesis of α-D-mannopyranosyl azide (2) using hydrazine[73]

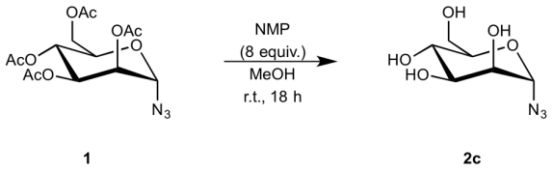


1 2b

Figure 57 Deprotection of acetylated **1** – entry b.

In an Eppendorf tube, 2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranosyl azide **1** (11 mg, 0.03 mmol) was dissolved in dry MeOH (250 µl), followed by the addition of hydrazine monohydrate (6 µl, 0.13 mmol, 4.5 equiv.). The solution was stirred overnight at room temperature, followed by the addition of hydrazine monohydrate (6 µl, 0.13 mmol, 4.5 equiv.) and full conversion after additional 24 hours of stirring. After quenching the reaction with acetic acid (5 µl), the solvent was removed *in vacuo*, the crude product diluted with water and lyophilized over several days to obtain product **2** as a colorless syrup (6 mg, 0.03 mmol, 99 %).

6.1.2c Synthesis of α-D-mannopyranosyl azide (2) using NMP/MeOH[74]

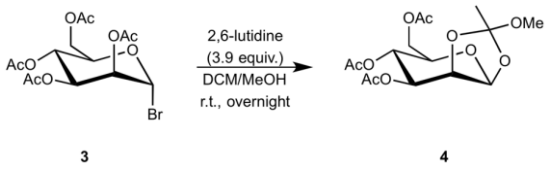


1 2c

Figure 58 Deprotection of acetylated **1** – entry c.

In an oven-dried round-bottom flask, dry MeOH (17 ml) was added to 2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranosyl azide (373 mg, 1 mmol, 1 equiv.), followed by *N*-methylpyrrolidine (0.83 ml, 8 mmol, 5 % in MeOH (v/v)). The solution was stirred at room temperature for 18 hours, followed by removal of the solvent *in vacuo*, dilution with water and final lyophilization overnight, obtaining the deprotected azido-mannose **2** (199 mg, 0.97 mmol, 97 %) as a white solid.

6.1.3 Synthesis of methyl-1,2-orthoacetate mannose (4)[80,81,87]



3 4

Figure 59 Synthesis of methyl-1,2-orthoacetate mannose **4**.

In an oven-dried round-bottom flask, compound **3** (2.5 g, 6.08 mmol, 1 equiv.) was dissolved in dry MeOH and dry DCM (20 ml, 1:1, (v/v)) and 2,6-lutidine (2.73 ml, 23.5 mmol, 3.9 equiv.) was added dropwise. The solution was stirred overnight at room temperature, followed by removal of the solvent *in vacuo*. The residue was co-evaporated with toluene (2 x 10 ml), followed by re-dissolving in $CHCl_3$ (20 ml). The solution was washed with ice-cold 3% aq. $NaHCO_3$ (10 ml), water (10 ml) and brine (10 ml), followed by extraction of the aqueous layer with $CHCl_3$ (3 x 20 ml). The combined organic layers were dried over Na_2SO_4, filtered and concentrated *in vacuo*. The crude product was purified by column chromatography (PE/EtOAc = 3:1 → 2:1) to give 683 mg of pure **4** (1.88 mmol, 31 %) as colorless needles, in accordance with literature.[80,81,87,88]

^{1}H NMR (600 MHz, $CDCl_3$) δ 5.50 (d, *J* = 2.5 Hz, 1H, H_1), 5.30 (t, *J* = 9.8 Hz, 1H, H_4), 5.15 (dd, *J* = 9.9, 4.0 Hz, 1H, H_3), 4.62 (dd, *J* = 4.0, 2.6 Hz, 1H, H_2), 4.24 (dd, *J* = 12.2, 4.9 Hz, 1H, H_{6b}), 4.14 (dd, *J* = 12.2, 2.6 Hz, 1H, H_{6a}), 3.68 (ddd, *J* = 9.6, 4.9, 2.6 Hz, 1H, H_5), 3.28 (s, 3H, OCH_3), 2.13 (s, 3H, OC(O)CH_3), 2.08 (s, 3H, OC(O)CH_3), 2.06 (s, 3H, OC(O)CH_3), 1.75 (s, 3H, OC(O)CH_3) ppm.

^{13}C NMR (151 MHz, $CDCl_3$) δ 170.9 (OC(O)CH_3), 170.6 (OC(O)CH_3), 169.6 (OC(O)CH_3), 124.7 (quat. *C*, orthoester), 97.5 (C_1), 76.7 (C_2), 71.4 (C_5), 70.8 (C_3), 65.5 (C_4), 62.4 (C_6), 50.1 (OCH_3), 24.6 (orthoester, CH_3), 20.9 (OC(O)CH_3), 20.9 (OC(O)CH_3), 20.9 (OC(O)CH_3) ppm.

ESI$^+$-MS (m/z) for $C_{15}H_{22}O_{10}Na$ = 385.11 $[M+Na]^+$; found: 385.03 Da.

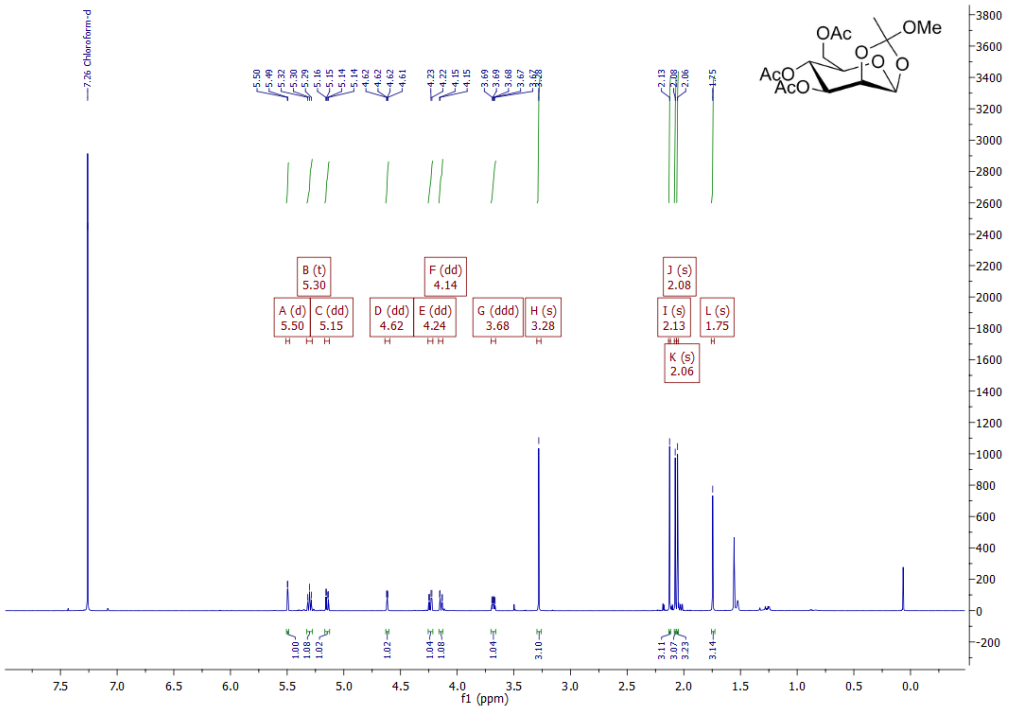


Figure 60 ^{1}H NMR spectrum of sugar **4**.

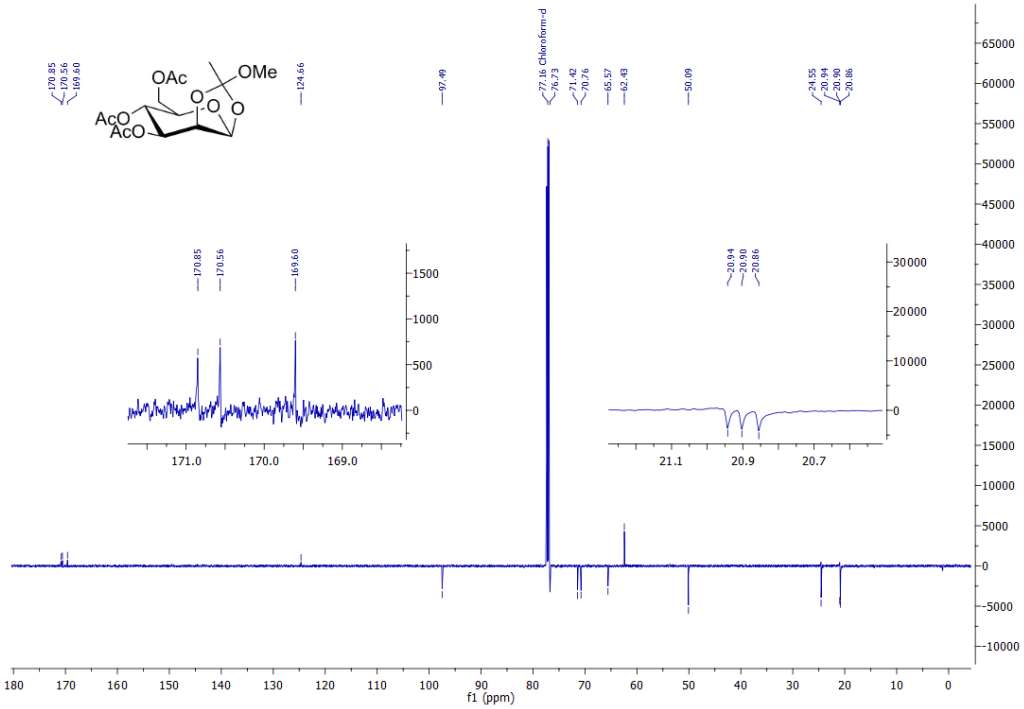


Figure 61 ^{13}C NMR spectrum of sugar **4**.

6.1.4 Synthesis of methyl-di-mannose (5)[80,81]

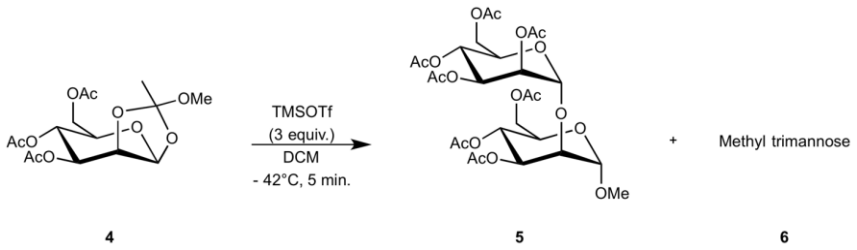


Figure 62 Synthesis of methyl-di-mannose **5**.

In an oven-dried round-bottom flask, compound **4** (515 mg, 1.42 mmol) was dissolved in DCM (20 ml) and cooled to -42°C (dry ice in ACN), followed by the dropwise addition of trimethylsilyl trifluoromethanesulfonate (770 µl, 4.27 mmol, 3 equiv.). After 5 minutes of stirring, the solution was diluted with DCM (20 ml) and quenched with aq. $NaHCO_3$ (10 ml). The phases were separated and the aqueous layer was extracted with DCM (3 x 10 ml). After combining the organic layers, the solution was washed with brine (20 ml) and dried over Na_2SO_4, followed by filtration and concentration *in vacuo*. The crude product was purified by column chromatography (hexane/EtOAc = 2/1 → 1/1), obtaining methyl-di-mannose **5** as a white foam (100 mg, 0.15 mmol, 22 % yield), in accordance with literature.[79,89,90] The formation of methyl-tri-mannose **6** was observed in the ^{1}H NMR spectrum of the crude product. However, due to the challenging purification, this species was not isolated.

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.44-5.22 (m, 5H, *H*$_3$, *H*$_4$, *H*$_{2'}$, *H*$_{3'}$, *H*$_{4'}$), 4.91 and 4.84 (2 d, each 1H, *J* = 1.8 Hz, *H*$_1$, *H*$_{1'}$), 4.26-4.09 (m, 5H, *H*$_{5'}$, *H6*$_{a,b}$, *H*$_{6a',b'}$), 4.03 (dd, *J* = 3.0, 2.0 Hz, 1H, *H*$_2$), 3.90 (ddd, *J* = 9.5, 4.4, 2.5 Hz, 1H, *H*$_5$), 3.41 (s, 3H, OC*H*$_3$), 2.15 (s, 3H, OC(O)*CH*$_3$), 2.14 (s, 3H, OC(O)*CH*$_3$), 2.09 (s, 3H, OC(O)*CH*$_3$), 2.08 (s, 3H, OC(O)*CH*$_3$), 2.04 (s, 3H, OC(O)*CH*$_3$), 2.03 (s, 3H, OC(O)*CH*$_3$), 2.01 (s, 3H, OC(O)*CH*$_3$).

^{13}C NMR (151 MHz, $CDCl_3$) δ 171.1 (*C*OCH_3), 170.7 (*C*OCH_3), 170.6 (*C*OCH_3), 170.1 (*C*OCH_3), 169.9 (*C*OCH_3), 169.7 (*C*OCH_3), 169.6 (*C*OCH_3), 99.4 and 99.3 (2C, *C*$_1$, *C*$_{1'}$), 77.0 (*C*$_2$), 70.4, 69.9, 69.2, 68.6 and 68.5 (5C, *C*$_{2'}$, *C*$_3$, *C*$_{3'}$, *C*$_5$, *C*$_{5'}$), 66.5 and 66.2 (2C, *C*$_4$, *C*$_{4'}$), 62.7 and 62.3 (2C, *C*$_6$, *C*$_{6'}$), 55.4 (*OC*H_3), 20.9, 20.9, 20.9 and 20.8 (7 x CO*C*H_3) ppm.

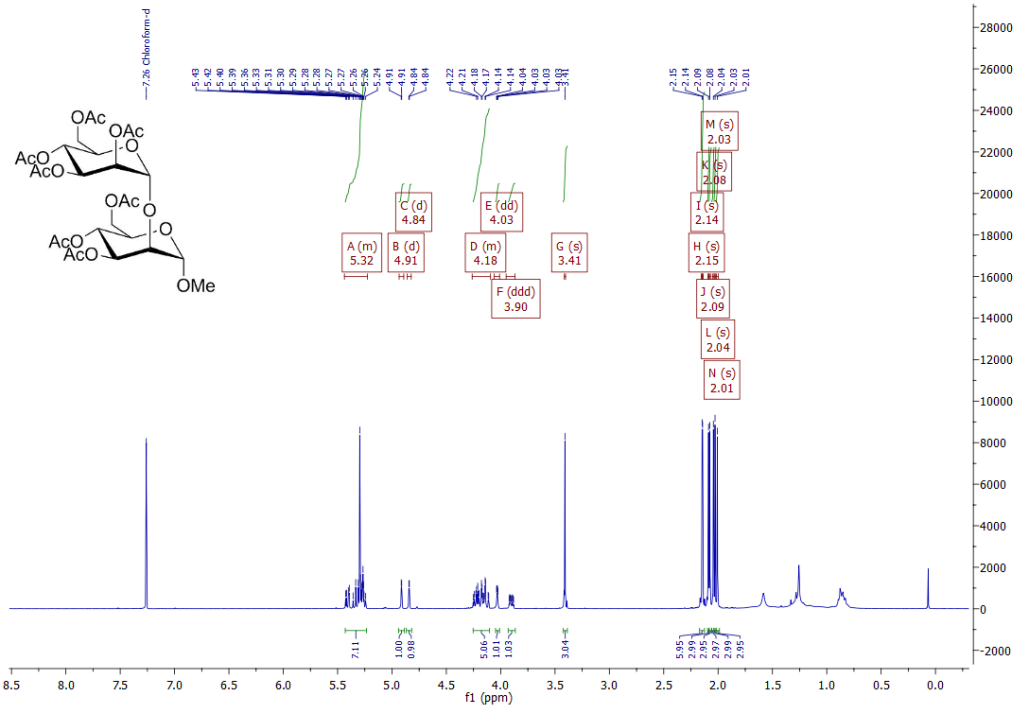


Figure 63 ^{1}H NMR spectrum of sugar **5**.

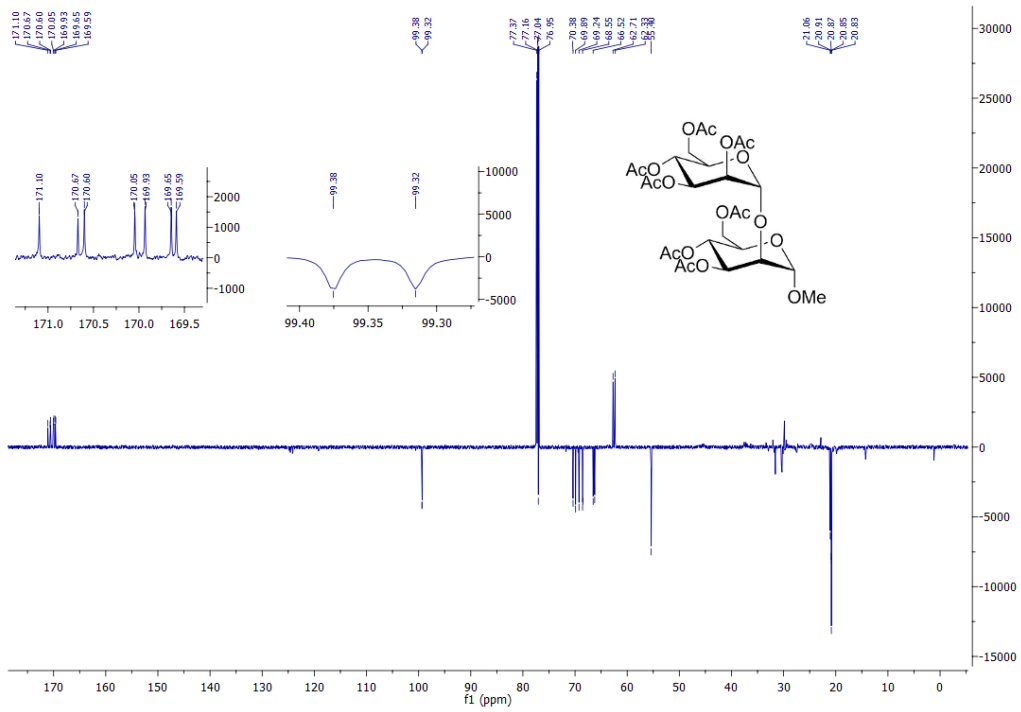
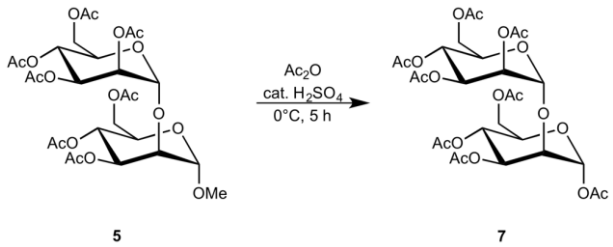


Figure 64 ^{13}C NMR spectrum of sugar **5**.

6.1.5 Acetylation of methyl-di-mannose (5)[79,82]



5 7

Figure 65 Acetylation of methyl-di-mannose **5**.

In a round-bottom flask, methyl-di-mannose **5** (84 mg, 0.13 mmol) was cooled in an ice bath. After the addition of ice-cold acetic anhydride (1 ml), sulfuric acid (4 % in acetic anhydride (v/v), 1 ml), was added dropwise. The reaction mixture was stirred for 5 hours at 0°C. The solution was then diluted with DCM (10 ml), neutralized with ice-cold saturated $NaHCO_3$ solution (2 ml) and washed with water (10 ml). The combined aqueous phase was extracted with dichloromethane (3 x 10 ml), followed by combining the organic phases. After filtration and drying with Na_2SO_4, the solvent was removed *in vacuo*. The crude product was purified by column chromatography (hexane/EtOAc = 1:1). The product **7** was obtained as a white foam (0.11 mmol, 73 mg, 83 % yield), in accordance with literature.[79,82,89–91]

^{1}H NMR (600 MHz, $CDCl_3$) δ 6.24 (d, *J* = 2.0 Hz, 1H, H_1), 5.46-5.38 (m, 2H, $H_{3'}$, $H_{4'}$), 5.31-5.22 (m, 3H, $H_{2'}$, H_3, H_4), 4.94 (d, *J* = 1.6 Hz, 1H, $H_{1'}$), 4.25-4.09 (m, 5H, $H_{5'}$, $H6_{a,b}$, $H_{6a',b'}$), 4.05-4.03 (m, 1H, H_2), 4.01 (ddd, *J* = 10.1, 3.5, 2.6 Hz, 1H, H_5), 2.15 (s, 6H, COCH_3), 2.14 (s, 3H, COCH_3), 2.10 (s, 3H, COCH_3), 2.09 (s, 3H, COCH_3), 2.04 (s, 6H, COCH_3), 2.01 (s, 3H, COCH_3) ppm.

^{13}C NMR (151 MHz, $CDCl_3$) δ 171.1 (*C*OCH_3), 170.9 (*C*OCH_3), 170.7 (*C*OCH_3), 170.1 (*C*OCH_3), 169.9 (*C*OCH_3), 169.6 (*C*OCH_3), 169.3 (*C*OCH_3), 168.4 (*C*OCH_3), 99.4 ($C_{1'}$), 91.5 (C_1), 76.0 (C_2), 70.8, 69.9, 69.8, 69.7, 68.4 (5C, $C_{2'}$, C_3, $C_{3'}$, C_5, $C_{5'}$), 66.3 and 65.6 (2C, C_4, $C_{4'}$), 62.5 and 61.8 (2C, C_6, $C_{6'}$), 21.07 (CO*C*H_3), 21.04 (CO*C*H_3), 20.87 (CO*C*H_3), 20.83 (CO*C*H_3), 20.81 (CO*C*H_3), 20.78 (CO*C*H_3), 20.78 (CO*C*H_3).

ESI$^+$-MS (m/z) for $C_{28}H_{38}O_{19}Na$ = 701.19 $[M+Na]^+$; found: 701.19 Da.

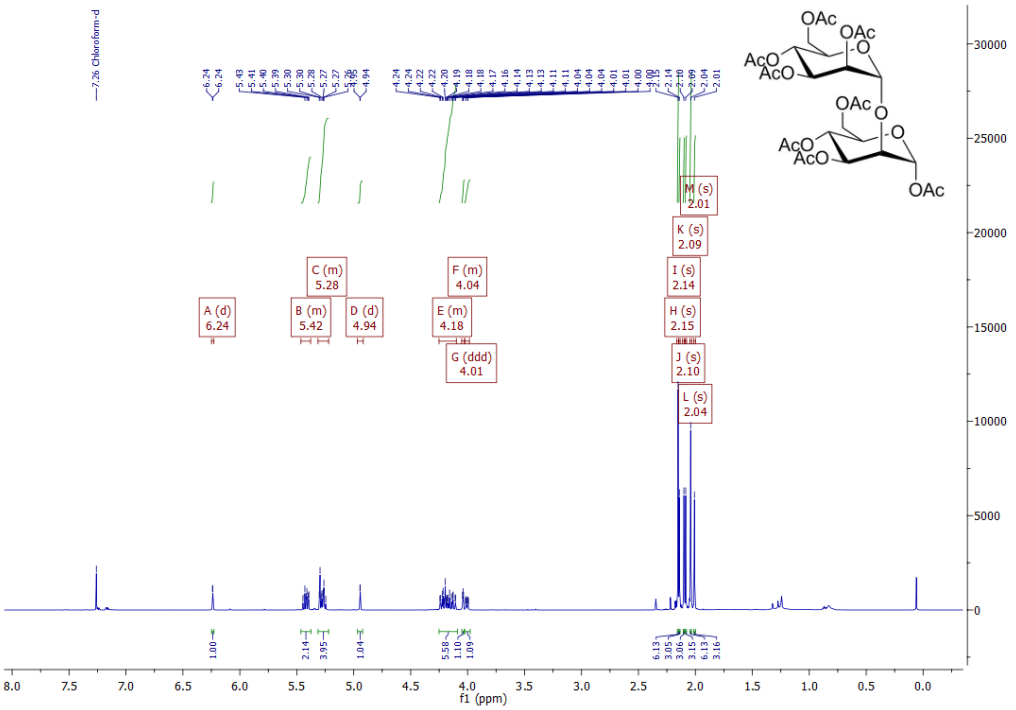


Figure 66 ^{1}H NMR spectrum of sugar **7**.

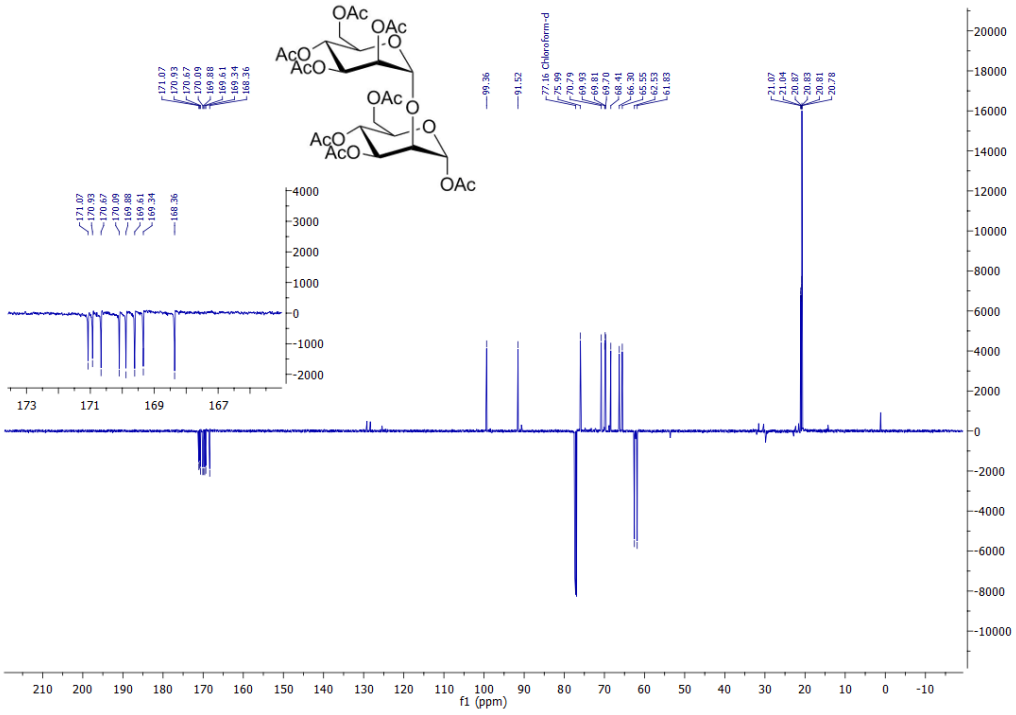


Figure 67 ^{13}C NMR spectrum of sugar **7**.

6.1.6 Azide-modification of acetylated di-mannose (7)[77]

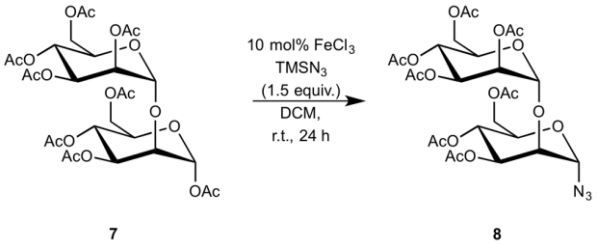


Figure 68 Azide-modification of acetylated di-mannose **7**.

In a round-bottom flask, compound **7** (64 mg, 0.09 mmol, 1 equiv.) was combined with 10 mol% dry iron(III) chloride (1.6 mg, 10 µmol, 0.1 equiv.), followed by the addition of $TMSN_3$ (0.15 mmol, 1.5 equiv., 20 µl) in DCM (2 ml). The reaction mixture was the stirred for 24 hours, when TLC indicated complete consumption of the starting material. The solution was diluted with DCM (8 ml), washed with a sat. $NaHCO_3$ solution (5 ml) and the aqueous layer was extracted with DCM (3 x 5 ml). The organic layers were combined and washed with brine (5 ml), dried over Na_2SO_4, filtered and the solvent was removed in vacuum. The crude product was purified by column chromatography (hexane/EtOAc = 1:1) to give product **8** as a white solid (33 mg, 0.05 mmol, 55 % yield).

^{1}H NMR (600 MHz, $CDCl_3$) δ 5.57 (d, *J* = 2.0 Hz, 1H, *H*$_1$), 5.39-5.30 (m, 2H), 5.26-5.21 (m, 2H), 5.17 (dd, *J* = 9.9, 3.2 Hz, 1H), 4.89 (d, *J* = 1.1 Hz, 1H, *H*$_{1'}$), 4.25 (dd, *J* = 12.4, 4.7 Hz, 1H), 4.20-4.13 (m, 4H), 4.12-4.06 (m, 1H), 3.96 (t, *J* = 3.1 Hz, 1H), 2.14 (s, 6H, C*H*$_3$), 2.09 (s, 3H, C*H*$_3$), 2.07 (s, 3H, C*H*$_3$), 2.04 (s, 3H, C*H*$_3$), 2.03 (s, 3H, C*H*$_3$), 2.00 (s, 3H, C*H*$_3$).

^{13}C NMR (151 MHz, $CDCl_3$) δ 171.0 (*C*OCH$_3$), 170.6 (*C*OCH$_3$), 170.4 (*C*OCH$_3$), 170.0 (*C*OCH$_3$), 169.9 (*C*OCH$_3$), 169.6 (*C*OCH$_3$), 169.4 (*C*OCH$_3$), 99.4 (*C*$_{1'}$), 88.3 (*C*$_1$), 76.8 (*C*$_2$), 70.7, 69.7, 69.6, 69.3 and 68.4 (5C, *C*$_{2'}$, *C*$_3$, *C*$_{3'}$, *C*$_5$, *C*$_{5'}$), 66.4 and 65.6 (2C, *C*$_4$, *C*$_{4'}$), 62.8 and 61.9 (2C, *C*$_6$, *C*$_{6'}$), 21.0 (2 x CO*C*H$_3$), 20.8 (3 x CO*C*H$_3$), 20.8 (2 x CO*C*H$_3$).

ESI$^+$-HRMS (m/z) for $C_{26}H_{35}N_3O_{17}Na$ = 684.1864 $[M+Na]^+$; found: 684.1864 Da.

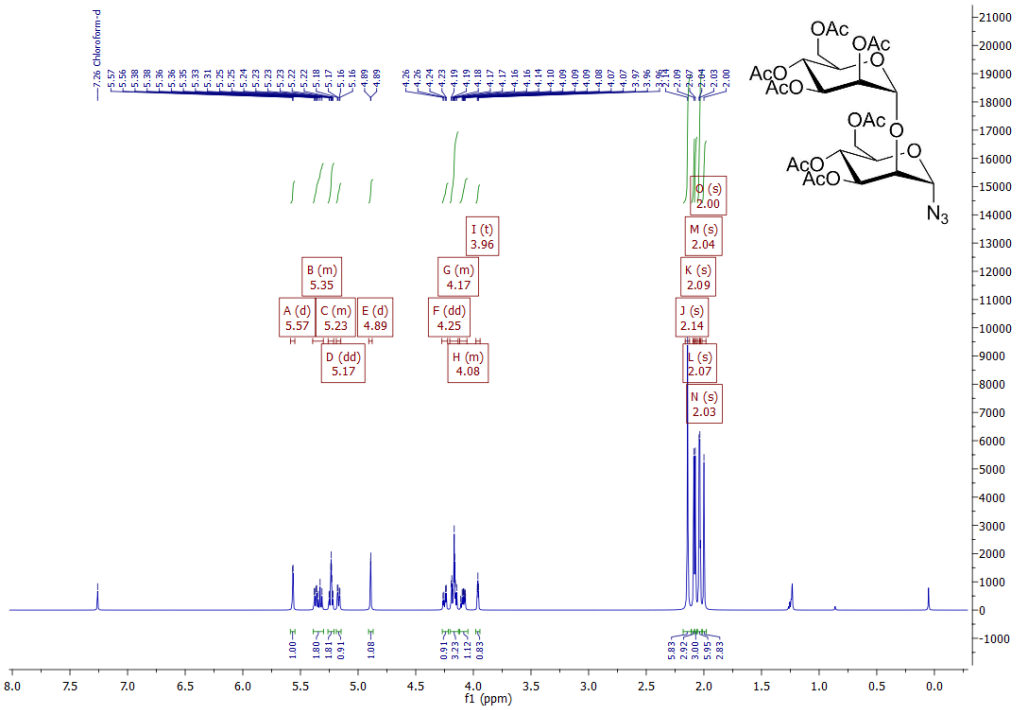


Figure 69 ^{1}H NMR spectrum of sugar **8**.

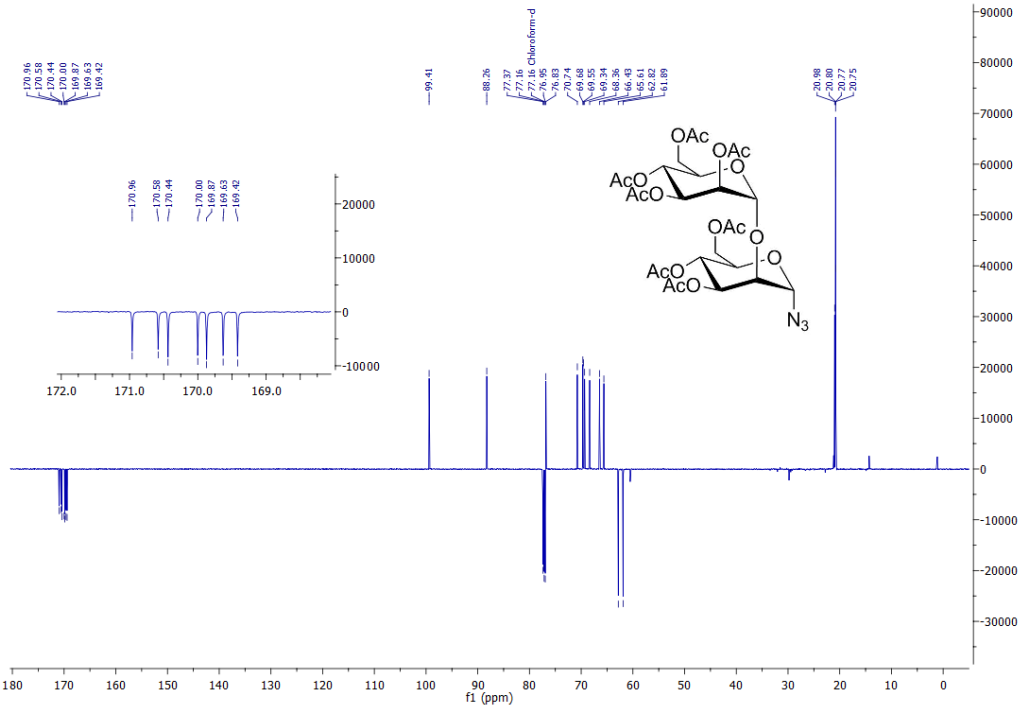
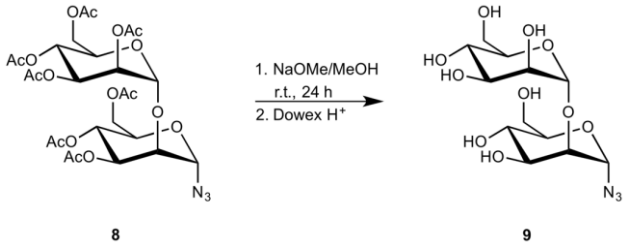


Figure 70 ^{13}C NMR spectrum of sugar **8**.

6.1.7 Synthesis of di-mannopyranosyl azide (9)



8 9

Figure 71 Synthesis of di-mannopyranosyl azide **9**.

In a round-bottom flask, compound **8** (33 mg, 0.05 mmol, 1 equiv.) was dissolved in methanol (1 ml), followed by the addition of catalytic amounts of sodium methoxide (0.5 M in MeOH, 50 µl) with stirring. After 24 hours, TLC indicated a completed reaction. Regenerated Dowex 50 H^+ was added until a neutral pH was reached. The solution was then diluted with water, filtered and lyophilized to give product **9** as a white solid (17 mg, 0.05 mmol, 93 % yield).

^{1}H NMR (400 MHz, D_2O) δ 5.79-5.77 (m, 1H, *H1*), 5.08 (d, *J* = 1.5 Hz, 1H, *H1'*), 4.12-4.09 (m, 1H, skeleton H), 3.99-3.65 (m, 10H, skeleton protons).

^{13}C NMR (151 MHz, D_2O) δ 102.2 (*C1'*), 88.2 (*C1*), 78.4 (*C2*), 74.5, 73.3, 70.2, 69.8 and 69.4 (5C, *C2', C3*, *C3', C5, C5'*), 66.7 and 66.4 (2C, *C4*, *C4'*), 61.0 and 60.7 (2C, *C6*, *C6'*).

IR spectroscopy: $\tilde{\nu}$ = 3273 (-O-H), 2930 (-C-H), 2117 (N_3), 1412, 1243, 1063, 932, 820 cm^{-1}.

ESI$^+$-HRMS (m/z) for $C_{12}H_{21}N_3O_{10}Na$ = 390.1125 $[M+Na]^+$; found: 390.1104 Da.

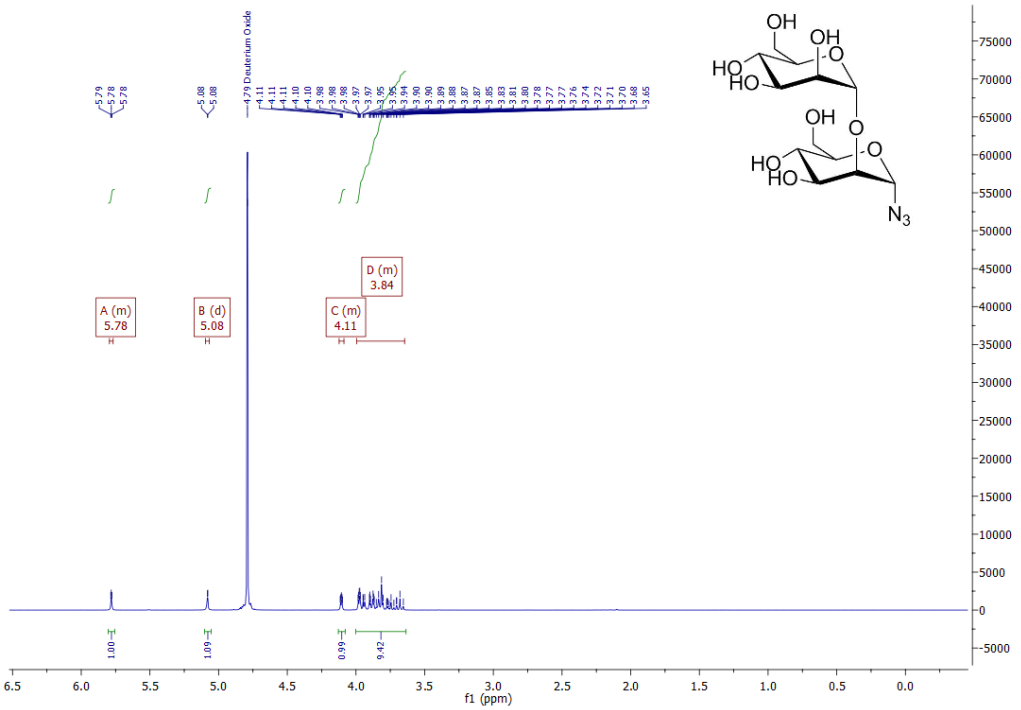


Figure 72 ^{1}H NMR spectrum of sugar **9**.

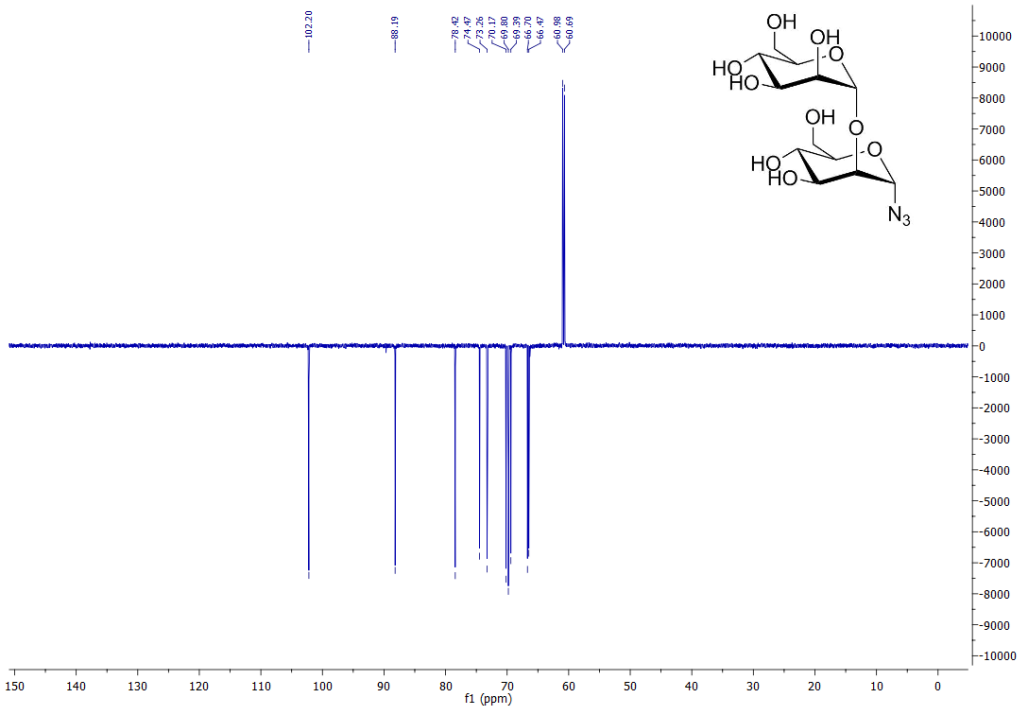


Figure 73 ^{13}C NMR spectrum of sugar **9**.

7 Supplementary information

Molecular graphics were obtained using the UCSF Chimera package[23] and PDB files imported from the Protein Data Bank[92]. Reaction schemes and structural formulas were drawn by using ChemDraw (PerkinElmer, Waltham, USA). NMR analysis was performed by using MestReNova (Mestrelab Research, Santiago de Compostela, Spain). SPR data were analyzed using Scrubber (BioLogic Software, Campbell, Australia).

7.1 Restriction digest

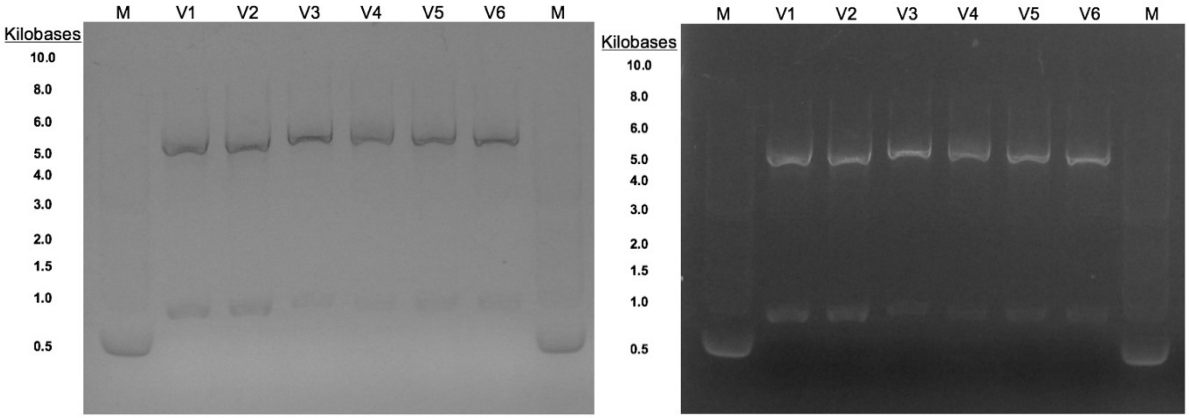


Figure 74 Agarose gel after restriction digest (left and right). Insert: 781 bp, remaining vector: 5280 bp.

7.2 Protein data

M-(CV-N)2L0 methylated cyanovirin (designed dimer), expressed in *E. coli*.

Sequence:

MGHHHHHHIEGRLGKFSQTCYNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDG SLKWQPSNFIETCRNTQLAGSSELAAECKTRAQQFVSTKINLDDHIANIDGTLKYELGKFSQT CYNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKWQPSNFIETCRNTQLAGSSEL AAECKTRAQQFVSTKINLDDHIANIDGTLKYE

Calculated mass (Da): 23474.83 (av., Cys reduced), 23460.26 (mono., Cys reduced); 23466.83 (av., Cys oxidized), 23452.26 (mono., Cys oxidized)

Extinction coefficient (M^{-1} cm^{-1}), calculated using the ProtParam tool from expasy.org, 19940 (all Cys reduced); 20440 (all Cys oxidized)

(CV-N)2L0 sequence from PDB-ID ‘3S3Y’:

Sequence:

GHHHHHHIEGRLGKFSQTCYNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGS
LKWQPSNFIETCRNTQLAGSSELAAECKTRAQQFVSTKINLDDHIANIDGTLKYELGKFSQTC
YNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKWQPSNFIETCRNTQLAGSSELA
AECKTRAQQFVSTKINLDDHIANIDGTLKYE

Calculated mass: 23343.63 (av., Cys reduced), 23329.22 (mono., Cys reduced);
23335.63 (av., Cys oxidized), 23321.22 (mono., Cys oxidized)

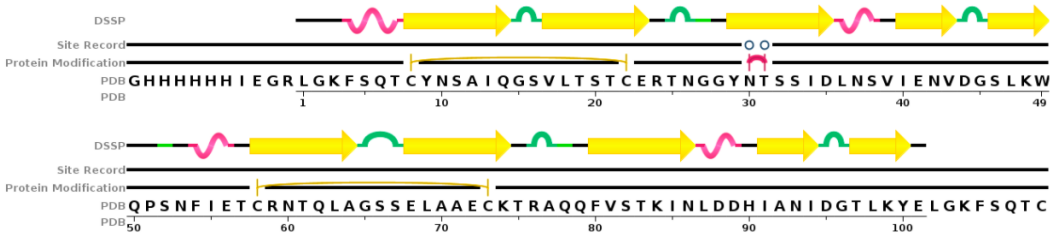


Figure 75 Secondary structure elements of (CV-N)2L0, obtained from PDB (3S3Y).

```
Query  2    GHHHHHHIEGRLGKFSQTCYNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKW  61
            GHHHHHHIEGRLGKFSQTCYNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKW
Sbjct  1    GHHHHHHIEGRLGKFSQTCYNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKW  60

Query  62   QPSNFIETCRNTQLAGSSELAAECKTRAQQFVSTKINLDDHIANIDGTLKYELGKFSQTC  121
            QPSNFIETCRNTQLAGSSELAAECKTRAQQFVSTKINLDDHIANIDGTLKYELGKFSQTC
Sbjct  61   QPSNFIETCRNTQLAGSSELAAECKTRAQQFVSTKINLDDHIANIDGTLKYELGKFSQTC  120

Query  122  YNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKWQPSNFIETCRNTQLAGSSE  181
            YNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKWQPSNFIETCRNTQLAGSSE
Sbjct  121  YNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKWQPSNFIETCRNTQLAGSSE  180

Query  182  LAAECKTRAQQFVSTKINLDDHIANIDGTLKYE  214
            LAAECKTRAQQFVSTKINLDDHIANIDGTLKYE
Sbjct  181  LAAECKTRAQQFVSTKINLDDHIANIDGTLKYE  213
```

Figure 76 Alignment of investigated M-(CV-N)2L0 within this project and PDB-ID 3S3Y.

7.3 SPR sensorgrams

7.3.1 HA immobilized chip

V3

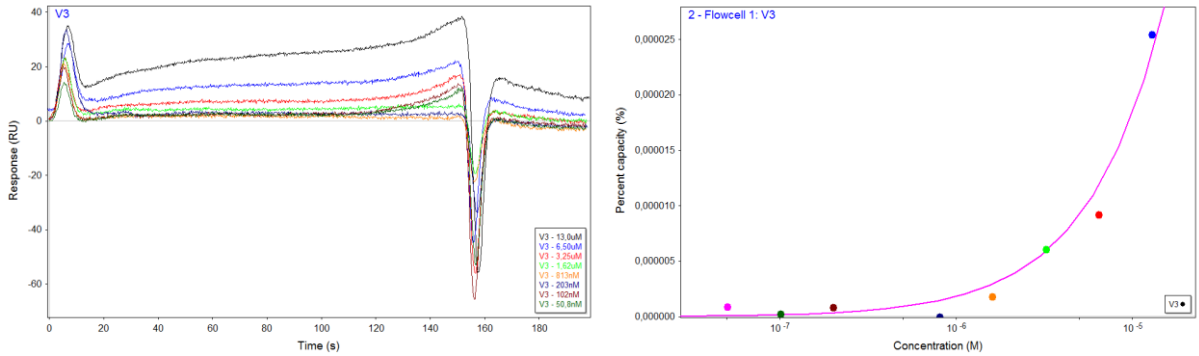


Figure 77 SPR sensorgram (left) and binding capacity (right) of V3 as analyte and HA as immobilized binding partner.

V4

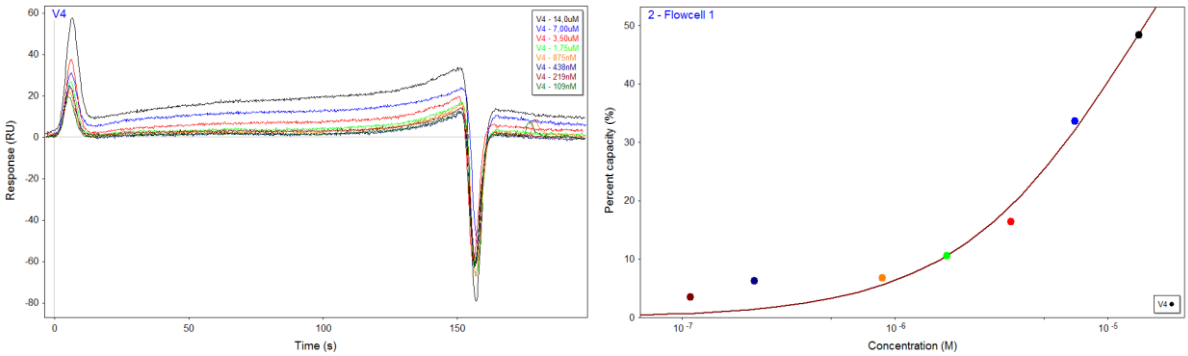


Figure 78 SPR sensorgram (left) and binding capacity (right) of V4 as analyte and HA as immobilized binding partner.

7.3.2 (CV-N)2L0 binding to HA, mono- and di-glycosylated peptide

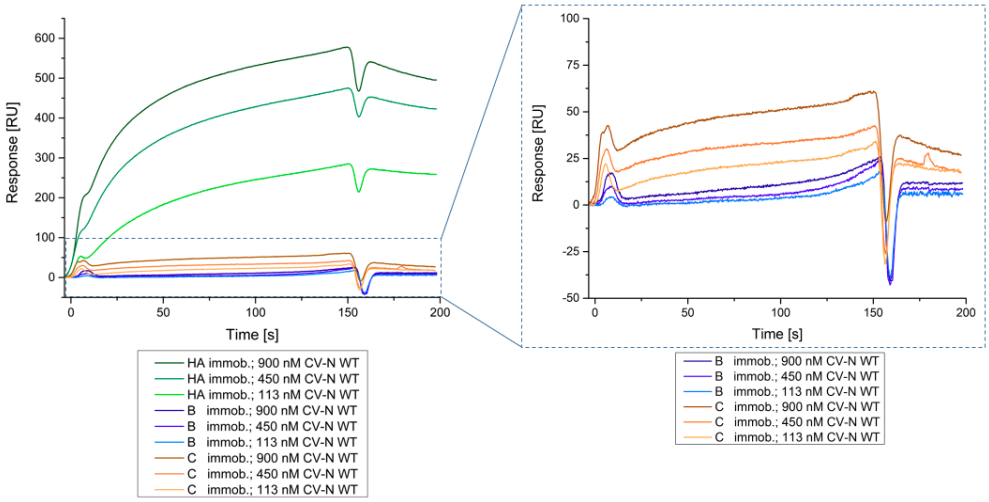


Figure 79 Qualitative comparison of binding of (CV-N)2L0 to immobilized **HA** (green), mono- (**B**, blue) and di-glycosylated (**C**, orange)) peptides with three different concentrations.

8 Acknowledgment

The results of this master thesis were obtained in the research group of Prof. Dr. Christian F. W. Becker (supervisor) and under co-supervision of Dr. Irene Maier, who initiated this project. In addition, protein expression was partially carried out at the University of Natural Resources and Life Sciences (Vienna), under the supervision of Dr. Martin Dragosits. SPR experiments were planed and performed by Dr. Irene Maier, while I provided the material, peptides and expressed and purified proteins.

I would like to express my sincere gratitude to Dr. Irene Maier for the continuous support not only during my experimental work and for discussing the questions that had arisen during this project. Moreover, I would like to thank her for introducing me to surface plasmon resonance and for carefully reviewing this thesis.

Also, I would like to thank Prof. Dr. Christian F. W. Becker, who gave me the opportunity of being part of his research group. The time I spent in his group enhanced my passion for biological chemistry and its techniques for expanding nature's toolbox. I am very thankful that he mediated this interesting, challenging and interdisciplinary project.

Special thanks goes to Dr. Martin Dragosits, who patiently supported me with the transformation, protein expression and purification. He always found time to answer my questions or to provide me with suggestions for overcoming challenges. I am very thankful that I was allowed to work in this group (PI Dr. Ian Wilson) and that I could ask everyone for support.

Furthermore, I would like to thank Anne Conibear for discussing ideas, challenges and scientific advice I received through my work.

I would like to thank all members of the Becker group since everyone contributed to this thesis. Moreover, Julia Schörghuber, Christian Denner, Manuel Gintner and Alexander Preinfalk supported me regarding carbohydrate synthesis. In addition, I would like to thank all people of the NMR- and MS-center core facilities.

First and foremost, I would like to thank my girlfriend who always supports and encourages me to pursue my dreams. I deeply appreciate her cheerful, altruistic and creative character as well as the years and adventures we spent together. Besides, I would like to acknowledge my friends for their valuable personalities and friendship.

Last but not least, I would like to thank my whole family and especially my parents. I am very grateful for their unconditional support through my entire life and the upbringing I enjoyed by them. They taught me to value the important things in life and the endurance for accomplishing goals – not only related to sports.

9 List of figures

10 List of tables

11 References

[1] I. J. Goldstein, R. D. Poretz, Academic Press, **1986**, pp. 33–247.

[2] R. S. Singh, A. K. Tiwary, J. F. Kennedy, *Crit. Rev. Biotechnol.* **1999**, *19*, 145–178.

[3] R. C. F. Cheung, J. H. Wong, W. Pan, Y. S. Chan, C. Yin, X. Dan, T. B. Ng, *Appl. Microbiol. Biotechnol.* **2015**, *99*, 3755–3773.

[4] A. F. . Santos, M. D. C. da Silva, T. H. Napoleão, P. M. G. Paiva, M. T. S. Correia, *Curr. Top. Pept. Protein Res.* **2014**, *15*, 41–66.

[5] J. Chen, D. Huang, W. Chen, C. Guo, B. Wei, C. Wu, Z. Peng, J. Fan, Z. Hou, Y. Fang, et al., *PLoS One* **2014**, *9*, e86455.

[6] C. Wu, W. Chen, J. Chen, B. Han, Z. Peng, F. Ge, B. Wei, M. Liu, M. Zhang, C. Qian, et al., *J. Biochem.* **2015**, *157*, 539–548.

[7] A. R. Garrison, B. G. Giomarelli, C. M. Lear-Rooney, C. J. Saucedo, S. Yellayi, L. R. H. Krumpe, M. Rose, J. Paragas, M. Bray, G. G. Olinger, et al., *Antiviral Res.* **2014**, *112*, 1–7.

[8] K. R. Gustafson, R. C. Sowder, L. E. Henderson, J. H. Cardellina, J. B. McMahon, U. Rajamani, L. K. Pannell, M. R. Boyd, *Biochem. Biophys. Res. Commun.* **1997**, *238*, 223–228.

[9] C. A. Bewley, K. R. Gustafson, M. R. Boyd, D. G. Covell, A. Bax, G. M. Clore, A. M. Gronenborn, *Nat. Struct. Biol.* **1998**, *5*, 571–578.

[10] F. Yang, C. A. Bewley, J. M. Louis, K. R. Gustafson, M. R. Boyd, A. M. Gronenborn, G. M. Clore, A. Wlodawer, *J. Mol. Biol.* **1999**, *288*, 403–412.

[11] L. M. I. Koharudin, L. Liu, A. M. Gronenborn, *Proc. Natl. Acad. Sci.* **2013**, *110*, 7702–7.

[12] L. G. Barrientos, J. M. Louis, I. Botos, T. Mori, Z. Han, B. R. O'Keefe, M. R. Boyd, A. Wlodawer, A. M. Gronenborn, *Structure* **2002**, *10*, 673–686.

[13] Y. K. Fujimoto, D. F. Green, *J. Am. Chem. Soc.* **2012**, *134*, 19639–19651.

[14] E. Matei, J. M. Louis, J. Jee, A. M. Gronenborn, **2012**, *79*, 1538–1549.

[15] E. Matei, R. Basu, W. Furey, J. Shi, C. Calnan, C. Aiken, A. M. Gronenborn, *J. Biol. Chem.* **2016**, *291*, 18967–18976.

[16] C. A. Bewley, S. Otero-Quintero, *J. Am. Chem. Soc.* **2001**, *123*, 3892–3902.

[17] C. a Bewley, *Structure* **2001**, *9*, 931–940.

[18] I. Botos, B. R. O'Keefe, S. R. Shenoy, L. K. Cartner, D. M. Ratner, P. H. Seeberger, M. R. Boyd, A. Wlodawer, *J. Biol. Chem.* **2002**, *277*, 34336–34342.

[19] R. Fromme, Z. Katiliene, P. Fromme, G. Ghirlanda, *Protein Sci.* **2008**, *17*, 939–944.

[20] R. S. Singh, A. K. Walia, J. S. Khattar, D. P. Singh, J. F. Kennedy, *Int. J. Biol. Macromol.* **2017**, *102*, 475–496.

[21] R. Fromme, Z. Katiliene, P. Fromme, G. Ghirlanda, *Protein Sci.* **2008**, *17*, 939–944.

[22] S. R. Shenoy, B. R. O'Keefe, a J. Bolmstedt, L. K. Cartner, M. R. Boyd, *J. Pharmacol. Exp. Ther.* **2001**, *297*, 704–710.

[23] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605–1612.

[24] C. A. Bewley, S. Kiyonaka, I. Hamachi, *J. Mol. Biol.* **2002**, *322*, 881–889.

[25] J. R. Keeffe, P. N. P. Gnanapragasam, S. K. Gillespie, J. Yong, P. J. Bjorkman, S. L. Mayo, *Proc. Natl. Acad. Sci.* **2011**, *108*, 14079–14084.

[26] R. Percudani, B. Montanini, S. Ottonello, *Proteins Struct. Funct. Genet.* **2005**, *60*, 670–678.

[27] R. Apweiler, H. Hermjakob, N. Sharon, *Biochim. Biophys. Acta - Gen. Subj.* **1999**, *1473*, 4–8.

[28] L. X. Wang, M. N. Amin, *Chem. Biol.* **2014**, *21*, 51–66.

[29] C. T. Walsh, S. Garneau-Tsodikova, G. J. Gatto, *Angew. Chemie - Int. Ed.* **2005**, *44*, 7342–7372.

[30] D. F. Zielinska, F. Gnad, J. R. Wiśniewski, M. Mann, *Cell* **2010**, *141*, 897–907.

[31] D. J. Kelleher, R. Gilmore, *Glycobiology* **2006**, *16*, 47R–62R.

[32] H. Freeze, M. Aebi, *Curr Opin Struct Biol* **2005**, *Okt*, 490–8.

[33] K. Ohtsubo, D. Marth, Jamie, *Cell* **2006**, *126*, 855–867.

[34] R. A. Dwek, *Chem. Rev.* **1996**, *96*, 683–720.

[35] G. Walsh, R. Jefferis, *Nat. Biotechnol.* **2006**, *24*, 1241–52.

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

[37] S. T. Anisfeld, P. T. Lansbury, *J. Org. Chem.* **1990**, *55*, 5560–5562.

[38] T. Conroy, K. A. Jolliffe, R. J. Payne, *Org. Biomol. Chem.* **2010**, *8*, 3723.

[39] P. Wang, B. Aussedat, Y. Vohra, S. J. Danishefsky, *Angew. Chemie Int. Ed.* **2012**, *51*, 11571–11575.

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

[41] Z. Guo, N. Shao, *Med. Res. Rev.* **2005**, *25*, 655–678.

[42] C. Unverzagt, Y. Kajihara, *Chem. Soc. Rev.* **2013**, *42*, 4408.

[43] R. G. Spiro, *Glycobiology* **2002**, *12*, 43R–56R.

[44] A. Varki, *Glycobiology* **1993**, *3*, 97–130.

[45] R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.

[46] S. B. H. Kent, *Annu. Rev. Biochem.* **1988**, *57*, 957–989.

[47] M. Ambrosi, N. R. Cameron, B. G. Davis, *Org. Biomol. Chem.* **2005**, *3*, 1593.

[48] M. Amblard, J.-A. Fehrentz, J. Martinez, G. Subra, *Mol. Biotechnol.* **2006**, *33*, 239–254.

[49] R. Huisgen, *Angew. Chemie Int. Ed. English* **1963**, *2*, 565–598.

[50] R. Huisgen, *Angew. Chem. Int. Ed.* **1963**, *2*, 633–645.

[51] E. Haldón, M. C. Nicasio, P. J. Pérez, *Org. Biomol. Chem.* **2015**, *13*, 9528–9550.

[52] C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057–64.

[53] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chemie - Int. Ed.* **2002**, *41*, 2596–2599.

[54] 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.

[55] G. C. Tron, T. Pirali, R. A. Billington, P. L. Canonico, G. Sorba, A. A. Genazzani, *Med. Res. Rev.* **2008**, *28*, 278–308.

[56] H. C. Kolb, B. Sharpless, K. B. Sharpless, *Drug Discov. Today* **2003**, *8*, 1128–1137.

[57] V. D. Bock, H. Hiemstra, J. H. Van Maarseveen, *European J. Org. Chem.* **2006**, *2006*, 51–68.

[58] C. W. Tornøe, S. J. Sanderson, J. C. Mottram, G. H. Coombs, M. Meldal, *J. Comb. Chem.* **2004**, *6*, 312–324.

[59] M. Whiting, J. Muldoon, Y.-C. Lin, S. M. Silverman, W. Lindstrom, A. J. Olson, H. C. Kolb, M. G. Finn, K. B. Sharpless, J. H. Elder, et al., *Angew. Chemie Int. Ed.* **2006**, *45*, 1435–1439.

[60] C. D. Hein, X.-M. Liu, D. Wang, *Pharm. Res.* **2008**, *25*, 2216–2230.

[61] J. Lutz, Z. Zarafshani, *Adv. Drug Deliv. Rev.* **2008**, *60*, 958–970.

[62] B. Voit, *New J. Chem.* **2007**, *31*, 1139–1151.

[63] D. Fournier, R. Hoogenboom, U. S. Schubert, *Chem. Soc. Rev.* **2007**, *36*, 1369.

[64] K. D. Hänni, D. A. Leigh, *Chem. Soc. Rev.* **2010**, *39*, 1240–1251.

[65] H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chemie - Int. Ed.* **2001**, *40*, 2004–2021.

[66] J. E. Hein, V. V Fokin, *Chem Soc Rev.* **2011**, *39*, 1302–1315.

[67] D. G. Drescher, M. J. Drescher, N. A. Ramakrishnan, *Methods Mol. Biol.* **2009**, *493*, 323–343.

[68] S. Victoria, *J. Bacteriol. Parasitol.* **2012**, *3*, DOI 10.4172/2155-9597.1000e110.

[69] M. Ritzefeld, N. Sewald, *J. Amino Acids* **2012**, *2012*, 1–19.

[70] P. Schuck, *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 541–566.

[71] I. E. Valverde, F. Lecaille, G. Lalmanach, V. Aucagne, A. F. Delmas, *Angew. Chemie - Int. Ed.* **2012**, *51*, 718–722.

[72] V. Percec, P. Leowanawat, H. J. Sun, O. Kulikov, C. D. Nusbaum, T. M. Tran, A. Bertin, D. A. Wilson, M. Peterca, S. Zhang, et al., *J. Am. Chem. Soc.* **2013**, *135*, 9055–9077.

[73] J. Li, Y. Wang, *Synth. Commun.* **2004**, *34*, 211–217.

[74] O. Moradei, S. Leit, C. du Mortier, Alicia, F. Cirelli, J. Thiem, *J. Carbohydr. Chem.*

1993, *12*, 13–22.

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

[76] V. Percec, P. Leowanawat, H. J. Sun, O. Kulikov, C. D. Nusbaum, T. M. Tran, A. Bertin, D. A. Wilson, M. Peterca, S. Zhang, et al., *J. Am. Chem. Soc.* **2013**, *135*, 9055–9077.

[77] S. B. Salunke, N. S. Babu, C.-T. Chen, *Chem. Commun.* **2011**, *47*, 10440.

[78] C. Uriel, J. Ventura, A. M. Gómez, J. C. López, B. Fraser-Reid, *J. Org. Chem.* **2012**, *77*, 795–800.

[79] Z. Szurmai, L. Balatoni, A. Lipták, *Carbohydr. Res.* **1994**, *254*, 301–309.

[80] C. Uriel, J. Ventura, A. M. Gómez, J. C. López, B. Fraser-Reid, *J. Org. Chem.* **2012**, *77*, 795–800.

[81] L. Chen, Z. Tan, *Tetrahedron Lett.* **2013**, *54*, 2190–2193.

[82] N. Hamon, C. C. Mouline, M. Travert, *European J. Org. Chem.* **2017**, *2017*, 4803–4819.

[83] A. C. Conibear, K. Farbiarz, R. L. Mayer, M. Matveenko, H. Kählig, C. F. W. Becker, *Org. Biomol. Chem.* **2016**, *14*, 6205–6211.

[84] A. Ben-Bassat, K. Bauer, S. Y. Chang, K. Myambo, A. Boosman, *J. Bacteriol.* **1987**, *169*, 751–757.

[85] P. H. Hirel, M. J. Schmitter, P. Dessen, G. Fayat, S. Blanquet, *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 8247–51.

[86] A. Micsonai, F. Wien, L. Kernya, Y.-H. Lee, Y. Goto, M. Réfrégiers, J. Kardos, *Proc. Natl. Acad. Sci.* **2015**, *112*, E3095–E3103.

[87] J. Beignet, J. Tiernan, C. H. Woo, B. M. Kariuki, L. R. Cox, *J. Org. Chem.* **2004**, *69*, 6341–6356.

[88] A. S. Perlin, *Can. J. Chem.* **1963**, *41*, 399–406.

[89] J. Mares, J. U. Müller, A. Skirgailiene, A. Neumoin, C. A. Bewley, R. R. Schmidt, O. Zerbe, *ChemBioChem* **2006**, *7*, 1764–1773.

[90] P. Taylor, M. T. N-glycoproteins, *Elements* **1998**, *17*, 417–437.

[91] Y. Itoh, S. Tejima, *Chem. Pharm. Bull. (Tokyo).* **1984**, *32*, 957–966.

[92] H. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. Shindyalov, P. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235–242.