Engineered Neoglycoproteins as Tools to Study Biologically Relevant Multivalent Interactions

Dissertation

Zur Erlangung des akademischen Grades

D o c t o r r e r u m n a t u r a l i u m (Dr. rer. nat.)

Im Fach Chemie

Spezialisierung: Organische und Bioorganische Chemie

Eingereicht an der Mathematisch-Naturwissenschaftlichen Fakultät

der Humboldt-Universität zu Berlin

von

M.Sc. Simon Klenk

Präsidentin der Humboldt-Universität zu Berlin

Prof. Dr. Sabine Kunst

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät

Prof. Dr. Elmar Kulke

Gutachter:	1. Prof. Dr. Christian Hackenberger 2. Prof. Dr. Andreas Herrmann 3. Dr. Christoph Böttcher
Tag der mündlichen Prüfung:	17.12.2018

Die vorliegende Arbeit wurde im Zeitraum vom 01.11.2013 bis zum 30.06.2018 unter der Leitung von Prof. Dr. Christian P. R. Hackenberger am Institut für Chemie der Humboldt Universität zu Berlin und am Leibniz-Forschungsinstitut für Molekulare Pharmakologie (FMP) angefertigt.

Ich erkläre, dass ich die Dissertation selbständig und nur unter Verwendung der von mir gemäß §7 Abs. 3 der Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät, veröffentlicht im Amtlichen Mitteilungsblatt der Humboldt-Universität zu Berlin Nr. 17/2012 am 27.06.2012 angegebenen Hilfsmittel angefertigt habe.

Ort, Datum Unterschrift

Submitted manuscripts

1.) S. Klenk, D. Lauster, K. Ludwig, S. Nojoumi, S. Behren, S. Saenger, S. Franz, K. Zscheppang, U. Hoffmann, M. Bardua, A. Hamann, T. Wolff, A. C. Hocke, S. Hippenstiel, S. De Carlo, J. Neudecker, N. Budisa, R. P. Netz, C. Böttcher, S. Liese, A. Herrmann, C. P. R. Hackenberger:
Phage capsid nanoparticles with defined ligand arrangement block influenza virus entry - An antiviral strategy

Talks

2.) M. Glanz, S. Klenk, C. P. R. Hackenberger:
Protein Conjugations: A Versatile Tool for Chemical Biology Research
SFB765 "Disziplinforum", October 2016, Berlin

3.) S. Klenk, D. Lauster, A. Herrmann, C. P. R. Hackenberger:
Targeting the Influenza A Virus by Engineered Viral Capsid Glyco-Conjugates with Optimized and Defined Ligand Presentation
SFB765 "Sonderkolloquium", June 2016, Berlin

Posters

1.) M. Glanz, S. Klenk, D. Lauster, S. Nojoumi, S. Behren, K. Ludwig, Ch. Böttcher, R. Haag, N. Budisa, A. Herrmann, C. P. R. Hackenberger:
Multivalent peptide- and protein-based scaffolds for inhibition of influenza virus binding to host cells
3. International Symposium of the SFB 765 „Multivalent Interactions of Pathogens with Biological Surfaces", October 2017, Berlin

2.) S. Klenk, D. Lauster, S. Liese, S. Nojoumi, S. Behren, K. Ludwig, S. Sänger, Th. Wolff, Ch. Böttcher, N. Budisa, A. Herrmann, C. P. R. Hackenberger:
Fighting Viruses with Viruses: A Supramolecular Multivalent Influenza Inhibitor based on the Bacteriophage Qβ Capsid
SFB 1093 Symposium "Supramolecular Chemistry on Proteins", September 2016, Frankfurt
(Poster Price Award)

3.) M. Glanz, S. Klenk, D. Lauster, S. Nojoumi, S. Behren, K. Ludwig, Ch. Böttcher, R. Haag, N. Budisa, A. Herrmann, C. P. R. Hackenberger:
Multivalent peptide- and protein-based scaffolds for inhibition of influenza virus binding to host cells
GDCh Wissenschaftsforum, September 2017, Berlin

4.) S. Klenk, D. Lauster, S. Liese, S. Nojoumi, S. Behren, K. Ludwig, Ch. Böttcher, N. Budisa, A. Herrmann, C. P. R. Hackenberger:
Fighting Viruses with Viruses: A Multivalent Influenza Inhibitor based on the Bacteriophage Qβ Capsid
Biochemistry 2016: "Shaping the Molecules of Life: Chemical Biology of Nucleic Acid and Protein Modifications", July 2016, Frankfurt
(Poster Price Award)

5.) S. Klenk, M. Glanz, S. Nojoumi, D. Lauster, S. Behren, A. Herrmann, N. Budisa, C. P. R. Hackenberger:
Multivalent peptide- and protein-based scaffolds for inhibition of influenza virus binding to host cells
Joint Meeting ECBS & ICBS, October 2015, Berlin

6.) S. Klenk, L. Artner, N. Bohlke, D. Lauster, J. Storm, J. Hütter, A. Herrmann, B. Lepenies, N. Budisa, C. P. R. Hackenberger:
GFP as a Multivalent Scaffold for the Synthesis of Functional Neoglycoproteins
3. International Symposium of the SFB 765 „Multivalency in Chemistry and Biochemistry", October 2014, Berlin

Acknowledgements

I would like to thank Prof. Dr. Christian Hackenberger for giving me the opportunity to pursue my doctoral studies on such a multifarious topic in his group and for being my supervisor. I appreciate the freedom, trust and opportunities I was given during my time in the lab and beyond.

I also want to thank Prof. Dr. Andreas Herrmann for reviewing this thesis, for being a stimulating and humorous collaborator and being part of my SFB765 "Betreuungsteam".

Moreover, I would like to thank my bachelor and master thesis supervisor Prof. Dr. Sabine Laschat and my research assistant supervisor Dr. Claudia Pacholski for their support and advice during my studies in Stuttgart.

Thanks to all of my great collaborators within the SFB765: Dr. Susanne Liese, Prof. Dr. Roland Netz, Dr. Kai Ludwig, Dr. Christoph Böttcher, Saba Nojoumi, Prof. Dr. Nediljko Budisa, Dr. Jens Dernedde, Dr. Sandra Sänger, Marlena Stadtmüller, and Dr. Thorsten Wolff. This thesis would not have been possible without your knowledge, expertise and commitment. Special thanks to Dr. Daniel Lauster for his support, patience and his willingness to always perform another assay. I appreciate our scientific discussions, our lunch/dinner breaks and chats.

I thank my student Sandra Behren for the motivated and productive teamwork and the incredibly honest opinion on mensa food.

I had the indescribable luck and pleasure to work in an awesome group during my doctorate. I am grateful for all the current and past members of the Hackenberger group I got to know during the past five years: Alec Michels, Alice Baumann, Andrew Grimes, Anett Hauser, Anselm Schneider, Dr. Antoine Wallabregue, Dagmar Krause, Dr. Debasish Bhowmick, Dr. Dominik Schumacher, Ines Kretzschmar, Dr. Jordi Bertran-Vicente, Katrin Wittig, Kristin Kemnitz-Hassanin, Kristina Siebertz, Dr. Lukas Artner, Lutz Adam, Marc-André Kasper, Dr. Marcie Jaffee, Maria Glanz, Marianne Dreißigacker, Martin Penkert, Dr. Michaela Mühlberg, Dr. Nicole Nischan, Dr. Olaia Nieto García, Dr. Oliver Reimann, Dr. Philipp Ochtrop, Dr. Robert Vallée, Sergej Schwagerus, Simon Reiske, Dr. Stefan Reinke, Tom Sauer, Dr. Vera Martos, and Dr. Wenyi Li. Thank you, Alec, for proofreading my thesis and hosting me. Thanks to my French teachers Alice, Anselm, Antoine and Kristina. Thank you, Kristin, for your everlasting patience and your organizing talent – cell-culturing could

not have been more efficient and funnier. Thank you so much, dear office crew, for uncountable scientific and political discussions, chats, giggles, games, support, and comfort food. Thank you, Simon, for the numerous taxi rides and our squash matches. Last but not least, Marc, it was a pleasure working with you. Your synthetic advice and experience was so helpful and you always knew what I wanted to hear. Thank you for introducing me to the world of peated whisky and cocktails thereof and for finally hosting me once.

I would like to thank all of my friends that I got to know from many different parts of my life, be it science, music, sports, or GDCh seminars. You shaped so many good experiences and happy moments of mine, that I cannot be happier.

I am grateful for my (grown) family who always supported me during my studies and my decisions I made. Thank you for always having a sympathetic ear for me when I need advice or help in any matters.

My last but foremost thanks is dedicated to my amazing and wonderful wife Sinja. Your constant and unconditional support throughout all situations and stages of our life is incredible and cannot be put into words.

"Je suis de ceux qui pensent que la science est d'une grande beauté. Un savant dans son laboratoire n'est pas seulement un technicien : c'est aussi un enfant placé en face des phénomènes naturels qui l'impressionnent comme un conte de fées."

"I am among those who think that science has great beauty. A scientist in his laboratory is not only a technician: he is also a child placed before natural phenomena which impress him like a fairy tale."

"Ich gehöre zu denen, die glauben, dass die Wissenschaft von großer Schönheit ist. Ein Wissenschaftler in seinem Labor ist nicht nur ein Techniker: Er ist auch ein Kind vor Naturphänomenen, die ihn wie ein Märchen beeindrucken."

Marie Curie, 1933

1. Table of Contents

1. Table of Contents 1
2. Abbreviations 5
3. Abstract 9
4. Kurzzusammenfassung 11
5. Introduction 13
5.1 Motivation 13
5.2 Influenza Viruses 14
5.2.1 Influenza A Virus 14
5.2.2 Influenza B Virus 21
5.2.3 Influenza C Virus 22
5.2.4 Influenza D Virus 22
5.3 Multivalency 23
5.3.1 Theoretical Consideration of Multivalency 23
5.3.2 Examples for Multivalency in Antiviral Strategies 27
5.4 Virus-like Particles 29
5.4.1 VLPs as Vaccines 29
5.4.2 VLPs as Viral Surrogates 32
5.4.3 VLPs as Imaging Agents 32
5.4.4 VLPs as (Targeted) Delivery Vehicles 33
5.5 Copper-catalyzed Alkyne-Azide Cycloaddition as Bioconjugation Reaction 35
5.6 Incorporation of Unnatural Amino Acids 39
5.6.1 Selective Pressure Incorporation 40
5.6.2 Amber Suppression 42
5.7 Sialic Acid 43
5.7.1 General 43
5.7.2 Biosynthesis 45
5.7.3 Chemical Glycosylation 46
6. Objectives 53
7. Results and Discussion 57

7.1 Optimizing the Synthetic Route Towards α-linked Sialic Acid Conjugates 57
7.1.1 Acceptor Synthesis 57
7.1.2 Donor Synthesis 64
7.1.3 Glycosylation Reactions 67
7.2 Synthesis of an *N*-Acetylgalactosamine Derivative as Control 77
7.3 Synthesis of 2-Azido Sialic Acid 79
7.4 Deprotection of Glycosylated Sialic Acid Derivatives 79
7.5 Synthesis of Azidosialyllactose Derivatives 80
7.6 Bioconjugation Reactions with Qβ Bacteriophage Capsids 81
7.7 Biophysical and Biological Evaluation of Sialylated Qβ Constructs 84
7.7.1 Cryo Electron Microscopy 84
7.7.2 Hemagglutination Inhibition Assay 86
7.7.3 Microscale Thermophoresis 90
7.7.4 Infection Inhibition 91
7.8 Modelling of the Multivalent Binding Event 93
7.9 Interior Modification of Qβ Capsids 94
7.10 Assay Setup for Another Structurally Defined Protein Scaffold 101

8. Summary and Outlook 109
9. Materials and Methods 115
9.1 General Information 115
9.1.1 Reagents and Solvents 115
9.1.2 Column Chromatography 115
9.1.3 Preparative HPLC 115
9.1.4 NMR 115
9.1.5 UPLC-MS 116
9.1.6 High-resolution ESI-MS 116
9.1.7 IR Spectroscopy 116
9.1.8 Optical Rotation 116
9.1.9 Protein Concentration Determination 116
9.1.10 Intact Protein MS 117
9.1.11 Tryptic Digest and ESI Orbitrap MS/MS 117

9.1.12 Electron Microscopy 118
9.1.13 Biological and Biophysical Methods 119
9.1.14 Human *ex vivo* Lung Tissue Experiments 119
9.1.15 Modelling 119
9.2 Organic Synthesis 120
9.2.1 Donor Synthesis 120
9.2.2 Acceptor Synthesis 125
9.2.3 Glycosylation Reactions 134
9.2.4 Deprotection of Carbohydrates 155
9.2.5 Azidosialyllactose Derivatives 159
9.2.6 CuAAC Ligand (TEOTA) 160
9.3 Qβ[Hpg] Expression and Purification 161
9.3.1 Analysis of Qβ[Hpg] 162
9.4 Bioconjugation 163
9.4.1 Analysis of Qβ[Gal3] 164
9.4.2 Analysis of Qβ[Sia0] 165
9.4.3 Analysis of Qβ[Sia1] 165
9.4.4 Analysis of Qβ[Sia3] 166
9.4.5 Analysis of Qβ[Sia5] 167
9.4.6 Analysis of Qβ[Sia8α] 168
9.4.7 Analysis of Qβ[Sia8β] 169
9.4.8 Analysis of Qβ[2.6SL] 170
9.4.9 Analysis of Qβ[2.3SL] 171
9.5 Interior Capsid Modification 172
9.6 HepG2 Cell Culture and Uptake Experiments 174
10. References 177

2. Abbreviations

1,2-DCE *1,2-dichloroethane*

4-NBDT ..... *4-nitrobenzenediazonium tetrafluoroborate*

aaRS ..... *aminoacyl-tRNA-synthetase*

Ada ..... *adamantanyl*

Aha ..... *azidohomoalanine*

aq. ..... *aqueous*

ASF ..... *asialofetuin*

ASGPR ..... *asialoglycoprotein receptor*

ATR ..... *attenuated total reflection*

BCA ..... *bicinchoninic acid*

BSA ..... *bovine serum albumin*

BSL ..... *biosafety level*

BTTAA ..... *2-[4-[(bis[(1-tert-butyl-1H-1,2,3-triazol-4-yl)*
methyl]amino)methyl]-1H-1,2,3-triazol-1-yl] acetic acid

CMP ..... *cytosine monophosphate*

CPMV ..... *cowpea mosaic virus*

CuAAC ..... *copper-catalyzed azide-alkyne cycloaddition*

CytD ..... *cytochalasin D*

DIPEA ..... *N,N-diisopropylethylamine*

DMAP *4-dimethylaminopyridine*

DMC ..... *2-chloro-1,3-dimethylimidazolinium chloride*

DMF ..... *dimethylformamide*

DMP ..... *dimethoxypropane*

DMSO ..... *dimethyl sulfoxide*

DMTST ..... *dimethylthiomethyl sulfonium trifluoromethanesulfonate*

DOTA ..... *1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid*

EDTA ..... *ethylenediaminetetraacetic acid*

EM ..... *electron microscopy*

eq. ..... *equivalents*

ESI ..... *electrospray ionization*

FACS ..... *fluorescence-activated cell sorting*
FAM ..... *carboxyfluorescein*
Fet ..... *fetuin*
FPLC ..... *fast protein liquid chromatography*
Gal ..... *galactose*
GalNAc ..... *N-acetylgalactosamine*
GFP ..... *green fluorescent protein*
HA ..... *hemagglutinin*
HAI ..... *hemagglutination inhibition assay*
HBV ..... *hepatitis B virus*
HEF ..... *hemagglutinin-esterase fusion*
HEV ..... *hepatitis E virus*
Hex ..... *hexane*
HHC ..... *human hepatocellular carcinoma*
HPAI ..... *highly pathogenic avian influenza*
Hpg ..... *homopropargylglycine*
HPLC ..... *high performance liquid chromatography*
HPV ..... *human papillomavirus*
IAV ..... *influenza A virus*
IPTG ..... *isopropyl-β-D-thiogalactopyranoside*
ISSD ..... *infantile free sialic acid storage disease*
IUPAC ..... *International Union of Pure and Applied Chemistry*
LC ..... *liquid chromatography*
LPAI ..... *low pathogenic avian influenza*
ManNAc ..... *N-acetylmannosamine*
MD ..... *molecular dynamics*
MeSOTf ..... *methylsulfenyl trifluoromethanesulfonate*
MRI ..... *magnetic resonance imaging*
MS ..... *mass spectrometry*
NA ..... *neuraminidase*
Neu ..... *neuraminic acid*

Neu5Ac ..... *N-acetylneuraminic acid*
NHS ..... *N-hydroxysuccinimide*
NIS ..... *N-iodosuccinimide*
NMR ..... *nuclear magnetic resonance*
NPC ..... *nitrophenylchloroformate*
P ..... *phosphate*
PAGE ..... *polyacrylamide gel electrophoresis*
PEG ..... *polyethylene glycole*
PEP ..... *phosphoenolpyruvate*
PME ..... *particle mesh Ewald*
RawIntDen ..... *raw intensity density*
RNA ..... *ribonucleic acid*
RNP ..... *ribonucleoprotein*
RPMI ..... *Roswell Park Memorial Institute medium*
sat ..... *saturated*
SDS ..... *sodium dodecyl sulfate*
SEC ..... *size-exclusion chromatography*
Sia ..... *sialic acid*
SL ..... *sialyllactose*
SLe^x ..... sialyl $Lewis^x$
S_N ..... *nucleophilic substitution*
SPI ..... *selective pressure incorporation*
TBTA ..... *tris-(benzyltriazolylmethyl)amine*
TCEP ..... *tris(2-carboxyethyl)phosphine*
TEOTA ..... *tris[(1-(2-ethoxy-2-oxoethyl)-1H-1,2,3-triazol-4-yl)methyl]amine*
TFA ..... *trifluoroacetic acid*
TfOH ..... *trifluoromethanesulfonic acid*
THF ..... *tetrahydrofuran*
THPTA ..... *tris[(1-hydroxypropyl-1H-1,2,3-triazol-4-yl)methyl]amine*
TMSOTf ..... *trimethylsilyl trifluoromethanesulfonate*
TMV ..... *tobacco mosaic virus*

TRANCE ..... *TNF-related activation-induced cytokine*

UDP ..... *uridine diphosphate*

UDP-GlcNAc ..... *N-acetylglucosamine*

UPLC ..... *ultra performance liquid chromatography*

VLP ..... *virus-like particle*

WHO ..... *World Health Organization*

3. Abstract

Multivalent ligand–receptor binding events are crucial in nature and responsible for mediating many biological interactions. One particularly prominent example is the cellular uptake of influenza viruses. Viral infection occurs through the recognition and binding of sialic acid-terminated cell-surface glycans by the viral membrane protein hemagglutinin. Consequently, the inhibition of this initial binding event by a suitable hemagglutinin inhibitor could represent a promising strategy to develop new antiviral drugs. Targeting the sialic acid binding site of hemagglutinin is especially desired since this protein domain is highly conserved throughout different influenza strains. An obvious strategy would therefore be to use sialic acid, the natural ligand of hemagglutinin, as an inhibitor, however, this strategy is hindered by the low affinity of monomeric sialic acid towards hemagglutinin. In order to efficiently inhibit hemagglutinin with this strategy, it is possible to also apply the concept of multivalency by presenting numerous sialic acid residues on a multivalent scaffold.

In this thesis, the bacteriophage Qβ capsid served as such a scaffold and facilitated the generation of a monodisperse multivalent system. It was modified with homopropargylglycine as an unnatural amino acid which served as a bioconjugation handle for sialic acid attachment by means of copper-catalyzed alkyne-azide cycloaddition. The corresponding synthesis of compatible azide-carrying sialic acid derivatives was one of the main tasks in this thesis. Stereoselective sialic acid synthesis of the desired α-anomer still remains a major challenge in the field of synthetic organic chemistry due to the special chemical features of sialic acid. To this end, straightforwardly accessible 5-*N*-acetyl adamantanyl thiosialoside was evaluated as glycosylation donor in the α-selective synthesis of sialosides. Efficient activation of this donor was achieved at –78°C under optimized *N*-iodosuccinimide/triflic acid promotion which led to high α-selectivities and overall yields for the desired sialosides. The number of synthetic steps was significantly reduced with this method,

compared to current state-of-the-art literature protocols. Particularly azidoethylene glycol linked sialic acids were synthesized with the optimized protocol, which were used for subsequent bioconjugation reactions to the Qβ capsid.

The synthesized sialic acid modified Qβ capsid particles were then thoroughly characterized by multiple biophysical and biological assays with regard to their ability to bind to hemagglutinin and to inhibit influenza infection. Low nanomolar affinities were measured in these assays and very efficient infection inhibition in a comparable concentration range was observed in *in vitro* cellular, *in vivo* mouse and *ex vivo* human model systems. Furthermore, several pathologically relevant influenza strains could be bound with the strategy presented here, which demonstrated the versatility of this approach. The monodisperse and defined structure of the Qβ scaffold additionally allowed for the establishment of a theoretical model describing the underlying binding modes which is difficult in comparable approaches based on polydisperse polymers or nanoparticle as multivalent scaffolds.

Taken together, the optimized synthetic route towards azide-bearing sialic acids supported the efficient synthesis of high affinity hemagglutinin binders and influenza inhibitors in the form of defined proteinaceous particles which in turn facilitated a better understanding of the involved multivalent binding event.

4. Kurzzusammenfassung

Multivalente Ligand-Rezeptor-Bindungsereignisse sind von entscheidender Bedeutung in der Natur und verantwortlich für die Vermittlung vieler biologischer Wechselwirkungen. Ein besonders prominentes Beispiel ist die zelluläre Aufnahme von Influenzaviren: Die virale Infektion erfolgt durch die Erkennung und Bindung von Sialinsäure-terminierten Zelloberflächenglykanen durch das virale Membranprotein Hämagglutinin. Folglich könnte die Hemmung dieses Bindungsereignisses durch einen geeigneten Hämagglutinin-Inhibitor eine vielversprechende Strategie zur Entwicklung neuer antiviraler Arzneimittel darstellen. Da diese Proteindomäne über verschiedene Influenzastämme hinweg größtenteils konserviert ist, wäre es von besonderer Bedeutung, die Sialinsäure-Bindungsstellen von Hämagglutinin zu adressieren. Eine naheliegende Strategie ist daher, Sialinsäure als natürlichen Liganden von Hämagglutinin ebenso als Inhibitor zu verwenden. Dieses Vorgehen wird allerdings durch die geringe Affinität von monomerer Sialinsäure gegenüber Hämagglutinin erschwert. Um Hämagglutinin auf diese Weise trotzdem effizient zu inhibieren, ist es möglich, ebenfalls das Konzept der Multivalenz anzuwenden und viele Sialinsäurereste auf einem multivalenten Gerüst zu präsentieren.

In der vorliegenden Arbeit diente das Kapsid der Bakteriophage Qβ als ein solches Gerüst und ermöglichte die Bildung eines monodispersen multivalenten Systems, welches mit Homopropargylglycin als unnatürliche Aminosäure modifiziert wurde. Das so eingeführte Alkin diente als funktionelle Gruppe für die kupferkatalysierte Alkin-Azid-Cycloaddition als Biokonjugationsmethode zur Anbindung von Sialinsäuregruppen. Die entsprechende Synthese der hierzu kompatiblen Azid-modifizierten Sialinsäurederivaten war eine der Hauptaufgaben dieser Arbeit. Die stereoselektive Sialinsäuresynthese zum gewünschten α-Anomer ist aufgrund der besonderen chemischen Eigenschaften von Sialinsäure, noch immer eine große Herausforderung auf dem Gebiet der synthetischen organischen Chemie. Zu diesem

Zweck wurde das einfach zugängliche 5-*N*-Acetyladamantanylthiosialosid als Glycosylierungsdonor in der α-selektiven Synthese von Sialosiden evaluiert. Eine effiziente Aktivierung dieses Donors wurde unter optimierten Bedingungen bei –78°C mit *N*-Iodsuccinimid und Trifluormethansulfonsäure erreicht, was zu hohen α-Selektivitäten und Gesamtausbeuten der gewünschten Sialoside führte. Die Anzahl der Syntheseschritte wurde mit dieser Methode im Vergleich zu den aktuellen Literaturprotokollen signifikant reduziert. Mit dem optimierten Protokoll wurden insbesondere Azidoethylenglykol-verknüpfte Sialinsäuren synthetisiert, die für nachfolgende Biokonjugationsreaktionen an das Qβ-Kapsid verwendet wurden.

Die synthetisierten Sialinsäure-modifizierten Qβ-Kapsidpartikel wurden dann eingehend mit Hilfe mehrerer biophysikalischer und biologischer Tests hinsichtlich ihrer Fähigkeit an Hämagglutinin zu binden und eine Influenza-Infektion zu inhibieren charakterisiert. Niedrige nanomolare Affinitäten wurden in diesen Assays gemessen. Eine sehr effiziente Infektionshemmung in einem vergleichbaren Konzentrations-bereich konnte in einem *in vitro* Zell-, sowie in einem *in vivo* Maus- als auch einem menschlichen *ex vivo* Modellsystem beobachtet werden. Darüber hinaus konnten verschiedene pathologisch relevante Influenzastämme über die hier vorgestellte Strategie gebunden werden, was die Vielseitigkeit dieses Ansatzes unterstreicht. Die monodisperse und definierte Struktur des Qβ-Gerüsts erlaubte es außerdem ein theoretisches Modell der zugrundeliegenden Bindungsmodi zu erstellen. In vergleichbaren Ansätzen, in denen polydisperse Polymere oder Nanopartikel als multivalente Gerüste verwendet wurden, ist solch eine Untersuchung nicht möglich.

Zusammengenommen trug der optimierte Syntheseweg zu azidhaltigen Sialinsäuren zur effizienten Synthese hochaffiner Hämagglutinin-Binder und Influenza-Inhibitoren in Form von definierten Protein-Partikeln bei, die wiederum ein besseres Verständnis des involvierten multivalenten Bindungsereignisses ermöglichten.

5. Introduction

5.1 Motivation

With up to 10% of adults and 30% of children worldwide suffering from the flu annually, causing 3 – 5 million cases of severe illness and up to 500,000 fatalities, the threat to the population is unquestioned.[1] Infection prevention by vaccination should be the first line of defense against influenza, but high mutation rates and frequent genetic rearrangements preclude comprehensive protection against seasonal influenza strains.[2] Despite the demand for antiviral drugs, only a few are approved for treating influenza infections. These can be categorized into two classes: 1.) amantadine and rimantadine are targeting the proton channel M2 of influenza A,[3] and 2.) the neuraminidase (NA) inhibitors zanamivir and oseltamivir (known as *Tamiflu*).[4] Blocking the M2 proton channel impairs the release of the viral genome into a host cell and therefore reduces infection. Neuraminidase inhibitors, on the other hand, prevent newly formed viral particles from spreading and thereby stem further infection. Thus, both classes of drugs only act on cells which have already been infected by the virus. Remarkably, there is no approved agent targeting hemagglutinin (HA) – the second large surface protein of an influenza virion besides NA – despite HA being responsible for initial recognition and binding to a host cell, as well as mediating the release of viral genetic material into the cell. Blocking HA binding sites would prevent the attachment of virions to host cell surfaces so that penetrating the host cell cannot occur at all. Therefore, this work focuses on the development, synthesis and evaluation of multivalent HA binding structures that can act as influenza inhibitors.

5.2 Influenza Viruses

5.2.1 Influenza A Virus

Infection with the flu is caused by the influenza virus, with the majority and most severe cases for humans being caused by influenza A viruses (IAV).[5] IAV is, like all other influenza genera, an RNA virus of the family of Orthomyxoviridae and therefore uses negative-sense single-stranded RNA ((-)ssRNA) to infect host cells. The IAV genome consists of eight segments which encode for eleven viral proteins.[6] The main reservoir of IAV are found in the realm of aquatic birds, but many other reservoirs and intermediate hosts, such as human, dog, canine, horse, seal, mink and bat have been identified so far.[7]

Infections with type A viruses are mostly accompanied by fever, coughing and nasal congestion,[8] but chills, sore throat, running nose, muscle pain, fatigue, headaches and pneumonia are also commonly encountered. Influenza type A viruses were responsible for several epidemics and pandemics during the last 100 years. These include 1918's 'Spanish Flu' which affected about 500 million people worldwide (~33% of world's population at that time), leading to 50 – 100 million deaths,[9] as well as the most recent 2009's flu pandemic.

5.2.1.1 Nomenclature

To this day, IAV is divided into 41 subtypes which are characterized by the type of viral surface proteins HA and NA (HxNy, x = 1 – 18, y = 1 – 11).[10] Various subtypes have adapted to different hosts, leading to the development of terms like 'human' and 'avian' influenza, as well as 'swine flu', 'horse flu' and 'dog flu'. Moreover, avian influenza can be further categorized into 'low pathogenic avian influenza (LPAI)' and 'highly pathogenic avian influenza (HPAI)'.

A more detailed nomenclature for isolated influenza strains was accepted and published by the World Health Organisation (WHO).[11] This system can be applied to all influenza genera, since it consists of the following components:

1.) antigenic type (A, B, C, D),

2.) original host (if not from human)

3.) geographical origin

4.) strain number

5.) year of isolation

6.) for IAV, the subtype (HxNy) follows in parenthese

e.g. 'A/mallard/Germany/439/2004(H3N2)' (IAV, isolated from mallard in Germany in 2004 with strain number 439) or 'A/PuertoRico/8/1934 (H1N1)' (IAV from human origin, isolated in Puerto Rico in 1934 with strain number 8).

5.2.1.2 Structure

Influenza viruses consist of a lipid envelope with spherical geometry in the size-range of 100 nm, or more often with filamentous geometry averaging 300 nm in length.[6] Most of the viral surface is covered with HA and NA spike proteins, with HA being the most abundant (4 : 1 HA : NA). The lipid envelope is interspersed with M2 ion channels and underlying is a capsid of the M1 matrix protein. Inside of that capsid the ribonucleoprotein (RNP) complex is found which contains the viral RNA, nucleoprotein and polymerase (Figure 1). Targets for (potential) antiviral therapies are described in more detail in the following section.

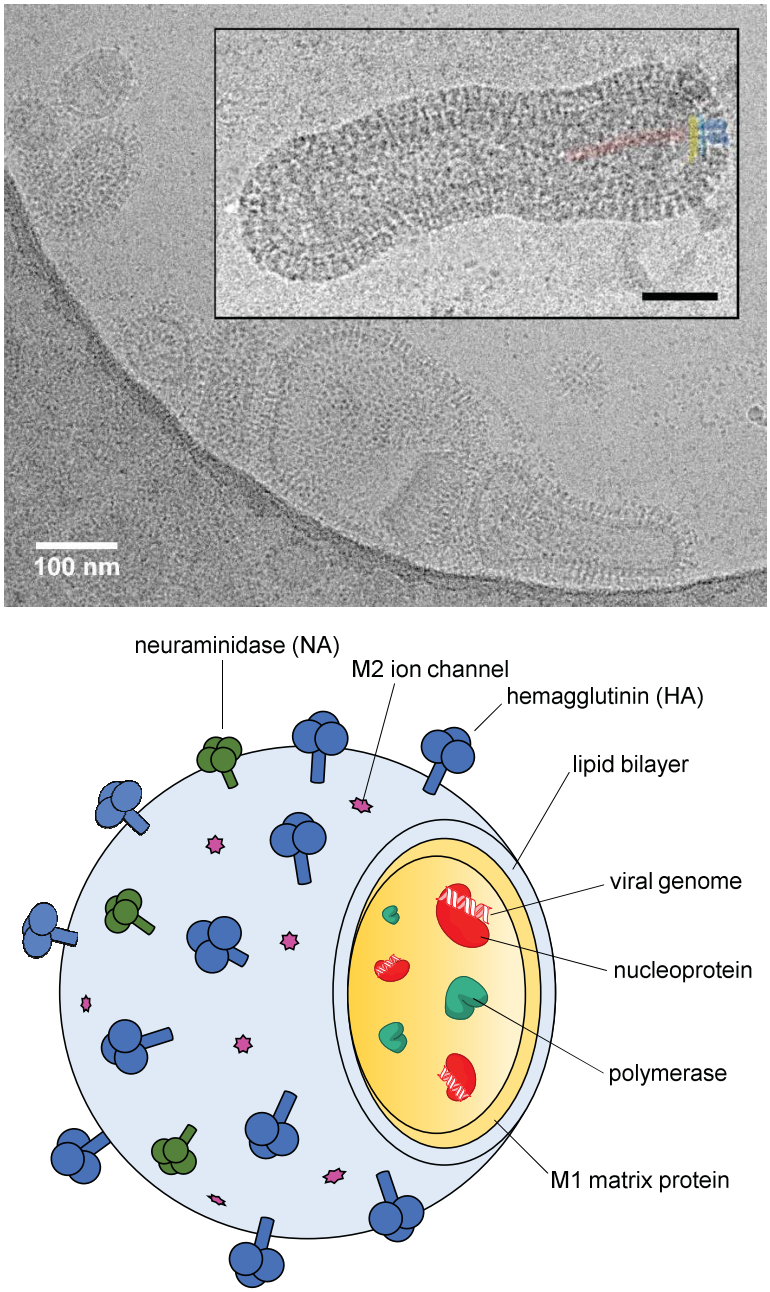


Figure 1: Transmission electron microscopy pictures (kindly provided by *Kai Ludwig*, FU Berlin) of influenza virions (top) which show variances in virus morphology. Inset: close-up image of a virion with colored HA spikes (blue), lipid membrane (cyan), M1 matrix protein layer (yellow), and RNP complex (red); and a schematic representation of an influenza virion (bottom) which additionally depicts NA spikes (green), the M2 ion channel (magenta), and viral polymerase (turquoise).

5.2.1.2.1 Hemagglutinin

HA is a homotrimeric, mushroom-shaped protein with a length of ~13 nm and a top-view diameter of ~5 nm. One monomer possesses a primary and a secondary binding site for sialic acid (Sia), located on the topside of the globular HA1 domain and

sideways on the globular head, respectively (Figure 2).[12] However, the binding affinity of monomeric sialic acid residues such as α–methylsialic acid and sialyllactose is in the low millimolar range (2 - 6 mM).[13] In spite of that, the virus can bind tightly to cell surfaces by the interaction of many HAs on the viral membrane (600 – 1200 per virion) and many sialic acid moieties presented on the cell surface (50 – 200 per nm^2 cell surface).[14] This effect, called multivalency (*vide infra*, section 5.3), is often found in biological systems especially when it comes to intercellular signaling or protein-carbohydrate interactions.[15]

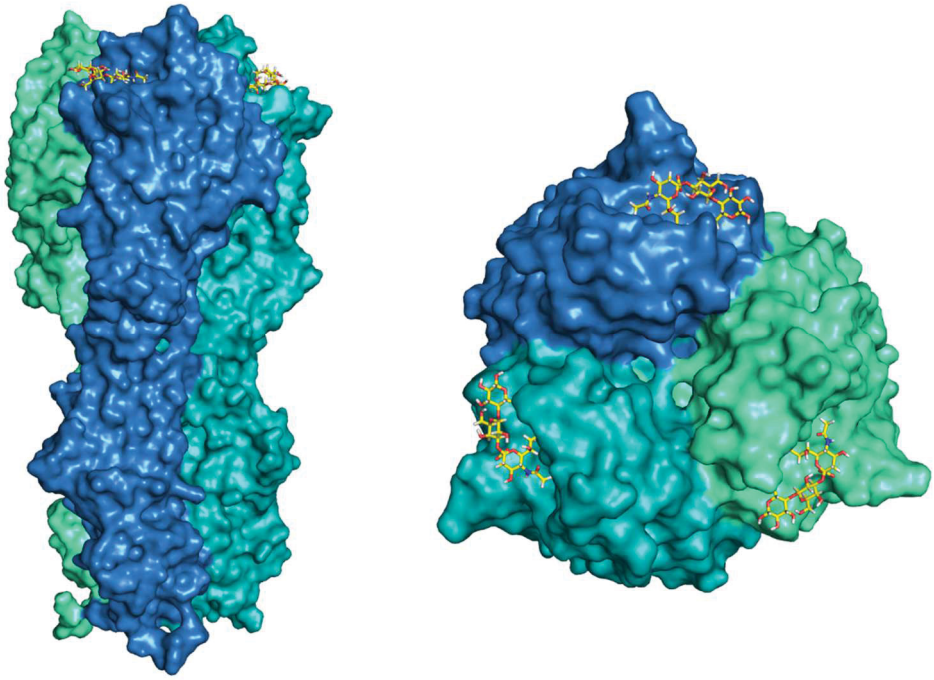


Figure 2: Crystal structure of a hemagglutinin trimer as a side view (left) and top view (right) with sialyllactose (yellow sticks) as bound ligand at the sialic acid binding sites (pdb: 1HGG).[16]

Apart from just binding to a host cell surface, HA is also responsible for host discrimination. Human pathogenic influenza strains predominately bind to α2,6-linked Sia glycans, whereas avian influenza strains bind to α2,3-linked structures.[17] It should be noted, that humans do not exclusively express α2,6-linked structures, but they are most abundant in their upper respiratory tract which is most accessible to inhaled virions.[18]

5.2.1.2.2 Neuraminidase

NA is a homotetrameric protein complex and also mushroom-shaped with a flat 8x8 nm head domain and a 6 – 10 nm long stem region (Figure 3).[19] It is an enzyme that catalyzes the cleavage of terminal sialic acid residues from glycans, with the active sites being located on top of the flat head domain. The removal of host cell sialic acids in order to promote the release of budding virions is often described as the main task of NA, but it has also been shown that NA prevents virions from self-agglutinating[20] and that it facilitates viral entry into host cells.[21]

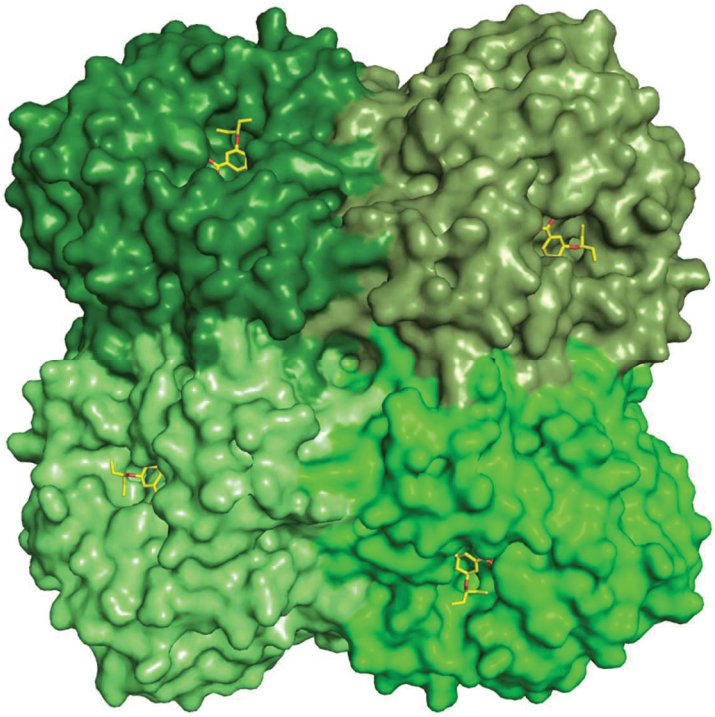


Figure 3: Crystal structure of the head domain of tetrameric neuraminidase with Oseltamivir (yellow sticks) as bound ligand at the sialic acid binding site (pdb: 2HU4).[22]

5.2.1.2.3 M2 proton channel

The matrix protein 2 forms a homotetrameric proton channel which consists of four transmembrane helices stabilized by disulfide bonds (Figure 4).[23] The formed channel is highly selective for protons and pH-regulated.[24] In its closed state at higher pH the channel is closed by a 'tryptophan gate' – tryptophan residues (Trp41) are in close proximity and stabilized by intermolecular contacts with aspartic acid residues

(Asp44).[25] Upon lowering the pH, histidine residues (His37) get protonated which leads to a conformational change and eventually to the opening of the gate. The actual proton current is proposed to be established by a water network located in the channel pore which enables 'proton-hopping' through the channel, the direct diffusion of hydronium ions or by a proton shuttling mechanism through local conformational changes.[26]

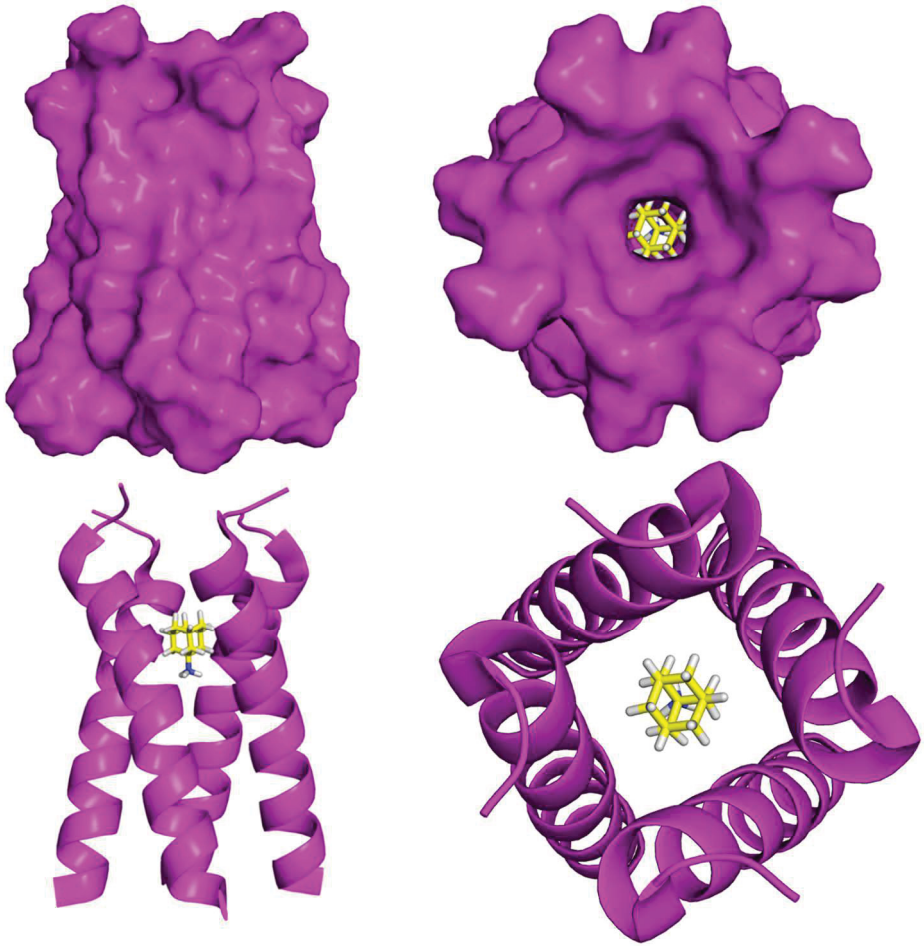


Figure 4: NMR structure of homotetrameric M2 proton channel as a side view (left) and top view (right) with rimantadine (yellow sticks) as bound ligand in the channel pore (pdb: 2RLF).[25]

5.2.1.3 Infection Mechanism

After the virus is inhaled by the host, viral HA binds to sialic acid terminated glycans on epithelial cells in the respiratory tract.[12a, 27] Precursor hemagglutinin is then cleaved by endogenous host proteases which activates the virus and leads to its infectious form.[28] After endocytosis and cellular uptake of the virus particle into

endosomes the pH drops which leads to proton influx through the M2 proton channel which in turn prepares the release of viral genes into the host cell.[29] The proton influx and acidic conditions in the endosome lead to an extensive conformational change of HA which finally causes the fusion of viral and endosomal membranes and releases the viral genome as an RNP complex into the host cell.[30] The viral nucleoprotein carries nuclear import signals which promotes its active transport into the host cell nucleus.[31] Inside the nucleus, complementary positive-sense RNA is produced and the viral genome is replicated and exported to the cytoplasm for translation into viral proteins.[32] To form a new viral particle, RNP complexes are excreted from the cell nucleus and transported to the outer cell membrane where translated HA, NA and M2 proteins are recruited.[33] After budding, NA cleaves off sialic acid residues from the newly formed viral particles and the membrane of the cell which enables the release of the new virion (Figure 5).[34]

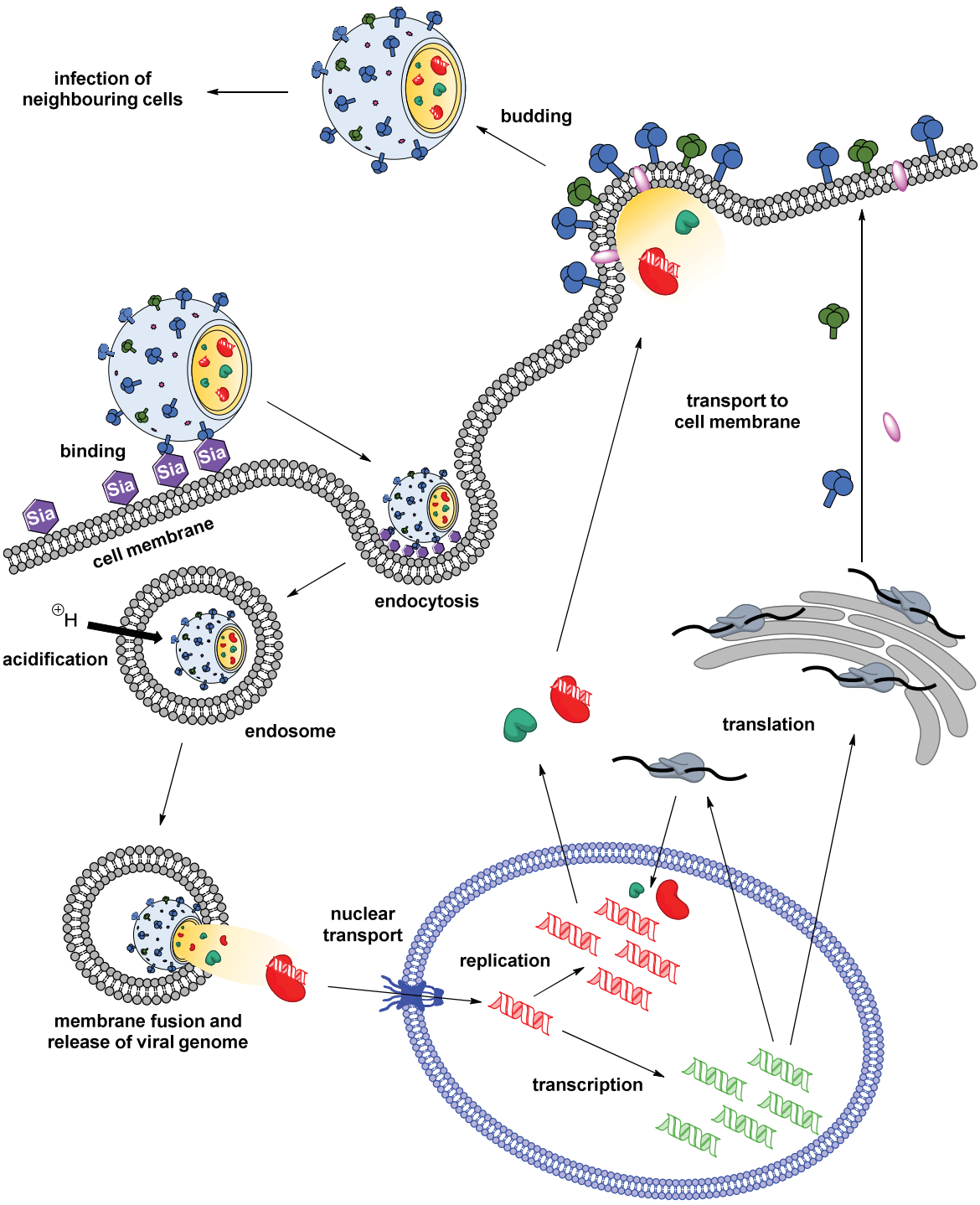


Figure 5: Schematic representation of the Influenza virus infection and reproduction mechanism.

5.2.2 Influenza B Virus

Influenza viruses of type B are closely related to IAV but the only known natural reservoirs of this type are humans and seals.[35] There are two circulating lineages partly responsible for seasonal flu outbreaks: B/Victoria and B/Yamagata.[5] Temporal

and geographical occurrences of these two distinct lineages fluctuate significantly, resulting in complicated global recommendations for seasonal flu vaccine compositions.

The mutation rate of influenza type B viruses is two to six times lower compared to IAV but higher than for influenza C.[36] The low number of natural reservoirs and mutation rate results in very efficient vaccination and the risk of pandemic or even epidemics caused by an influenza Type B virus is very low.

5.2.3 Influenza C Virus

The influenza C type differs significantly from previously described Type A and B. Type C exhibits only seven RNA segments in its genome which encode for nine proteins. Isolation and identification of this type is more difficult than for Type A and B due to its different growth requirements, but studies still suggest its worldwide occurrence.[37] Symptoms caused by an infection with influenza C virus are usually milder compared to the other types.[38]

The most prominent structural difference is seen for the viral surface glycoproteins. Only one is found on the surface of influenza type C: the hemagglutinin-esterase fusion (HEF). The esterase function is comparable to NA of influenza type A and B and responsible for the removal of host cell ligands to enable viral evasion.[39]

Since influenza C undergoes only antigenic drift, combined with a generally slow mutation rate[36a] and the mild symptoms associated with influenza C infection, it is not thought to be capable of causing threatening pandemics or epidemics. This is mirrored in the fact that there is no vaccination against Type C viruses.

5.2.4 Influenza D Virus

The most recently discovered influenza genre, subtype D, was first described in 2011 in the US. It was isolated from pigs which exhibited influenza-like symptoms and was

classified as related to human influenza C viruses.[40] Further investigation of these viral specimens revealed some conserved domains such as the viral entry machinery, but also discovered differences in other key functional sites. Evolutionary studies suggest that the influenza D genes diverged from human influenza C viruses only hundreds to a few thousand years ago.[41] Influenza D is predominantly found in cattle and pigs and so far two distinct subtypes are known.[42]

5.3 Multivalency

Many biologically relevant interactions are mediated by actually very weak forces, e.g. hydrogen bonding in carbohydrate-protein interactions. The key principle enabeling strong and efficient binding while also maintaining reversibility is multivalency. The most vivid picture to describe multivalency is the 'Velcro-principle': Velcro consists of one layer covered with many hooks and another layer which carries numerous loops. If one hook catches one loop, the resulting linkage is very weak and barely noticeable if pulled apart. However, if the two layers are brought together completely, a large number of hooks catch loops resulting in a strong interaction which does not come loose under minor force, but which can still be unfastened if enough force is applied. This macroscopic picture can be transferred to the molecular level: if a large surface covered with carbohydrate ligands (e.g. a cell surface) interacts with a surface which is also densely covered with a proteinaceous receptor (e.g. viral envelop), a tight interaction is enabled and infection of a cell by a virus only then possible (*vide supra*, section 5.2.1.3).

5.3.1 Theoretical Consideration of Multivalency

Multivalency is such an extraordinarily complex phenomenon that the Velcro example mentioned above only helps to visualize the very basic principle. Among others, *Kiessling*, *Whitesides* and co-workers contributed significantly to the understanding of

multivalent interactions by experimental and theoretical studies and important fundamentals are described in the following section.[14, 43]

5.3.1.1 Thermodynamics

A multivalent binding event is usually determined by the free binding enthalpy ΔG_{multi} and the average free energy contribution (ΔG_{avg}) of each individual participating interaction between a ligand and receptor (N) can be determined by:

$$\Delta G_{avg} = \frac{\Delta G_{multi}}{N} \tag{1}$$

An essential aspect of multivalent interactions is cooperativity α. That means that ΔG_{avg} can be greater ($\alpha > 1$), similar ($\alpha = 1$) or less ($\alpha < 1$) than the actual monovalent free binding energy ΔG_{mono}:

$$\Delta G_{avg} = \alpha \cdot \Delta G_{mono} \tag{2}$$

When it comes to multivalency, most of the cases display positive cooperativity ($\alpha > 1$), thus the multivalent binding is preferable with regard to free energy. It has to be noted, however, that multivalent binding with negative cooperativity ($\alpha < 1$) can be still significantly stronger than the monovalent scenario. Because this causal connection between cooperativity and tightness of binding is not given for multivalent systems, *Whitesides* introduced a more vivid measure to describe the enhanced binding of a multivalent system:

$$\beta = \frac{K_{multi}}{K_{mono}} \tag{3}$$

This enhancement factor β describes the ratio between the multivalent association constant (K_{multi}) and the monovalent association constant (K_{mono}), which are both experimentally accessible parameters. Additionally, the exact number of individual interactions (N) doesn't have to be known in order to determine β.

The free binding energy ΔG_{multi} consists of an enthalpy (ΔH_{multi}) and an entropy (ΔS_{multi}) term:

$$\Delta G_{multi} = \Delta H_{multi} - T \cdot \Delta S_{multi} \tag{4}$$

The enthalpic term is basically the sum of all monovalent binding enthalpies (ΔH_{mono}), but ΔH_{multi} can further be influenced:

$$\Delta H_{multi} = N \cdot \Delta H_{mono} \tag{5}$$

Equation (5) is in principle only valid, if a perfect conformational match between receptor and ligand exists. As soon as either the multivalent receptor or ligand has to modify its conformation to enable binding, a preferable or diminished binding occurs which results in a gain or loss in enthalpy, respectively. Since these gains and losses of enthalpy upon binding are not easy to quantitatively describe, it is very difficult to determine ΔH_{multi} exactly.

The entropy contribution to multivalent binding is of great importance for the actual binding scenario and has to be carefully considered in the design process of synthetic multivalent architectures. The entropy term ΔS_{multi} consists of a translational ($\Delta S_{trans,N}$), rotational ($\Delta S_{rot,N}$) and conformational entropy ($\Delta S_{conf,N}$) part, as well as a solvation entropy contribution ($\Delta S_{solv,N}$) of the surrounding solvent molecules:

$$\Delta S_{multi} = \Delta S_{trans,N} + \Delta S_{rot,N} + \Delta S_{conf,N} + \Delta S_{solv,N} \tag{6}$$

The translational and rotational entropies describe the freedom of a system to move or rotate through space or around the three principal axis, respectively. Both contributions are only logarithmically dependent on the mass and size of the corresponding molecules, which results in the fact that there is only little difference between a mono- and a multivalent system. The concentration of two binding partners must also be taken into consideration since the loss of translational entropy upon binding is higher in more dilute scenarios. Conformational entropy $\Delta S_{conf,N}$ can generally be considered to be lost upon binding, since the conformational degree of freedom of a bound ligand is less compared to the free ligand. The remaining $\Delta S_{solv,N}$ contribution comes mainly from bound solvent molecules, presumably water for a biological setting, which is released from the interacting surfaces of multivalent receptors and ligands upon binding, and which therefore increase entropy. Since the interacting surfaces of a multivalent system are approximately similar compared to *N*-

times a monovalent system, $\Delta S_{solv,}$ does not differ very much between those two scenarios and just adds a constant factor to the overall entropy.

Derived from these considerations some conclusions can be drawn: 1.) Using flexible linkers, e.g. widely applied ethylene glycols, to connect monovalent receptors or ligands in order to generate a multivalent scenario are worse compared to rigid linkers because of the significant loss of conformational entropy upon binding. 2.) On the other hand, it is also possible that entropic losses are compensated by enthalpic benefits, e.g. when a flexible linker enables the system to be bound without strain. 3.) The precise quantification of multivalent thermodynamics is an almost impossible issue to solve, because of numerous contributing factors which can hardly comprehensively be considered.

5.3.1.2 Kinetics

Important contributions from kinetics must also be considered. The very first monovalent interaction between multivalent systems can kinetically be treated as an intermolecular process, whereas all subsequent steps are then in principle intramolecular and therefore faster. This results in an overall averaged association rate (k_{on}) which is often faster compared to a monovalent system.[44] However, it was also demonstrated that the increased k_{on} of a multivalent system is not the predominant reason for the enhanced multivalent binding. It turns out that a tremendous decrease of the dissociation rate (k_{off}) in the order of 10^4 contributes largely to the beneficial binding scenario and dominates the kinetic influence of multivalency (Figure 6).[45]

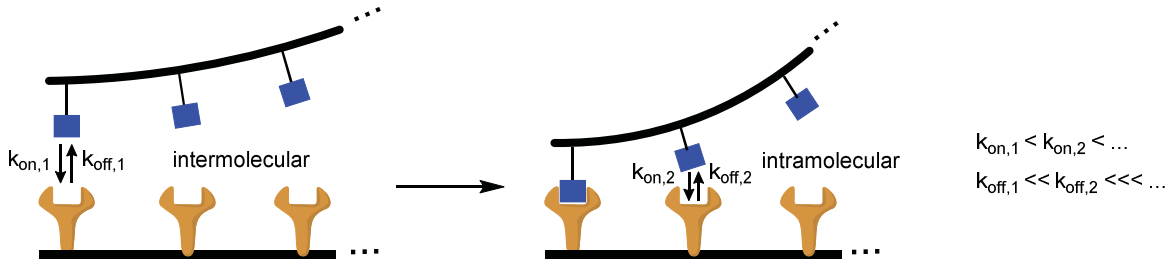


Figure 6: Schematic representation of the influence of multivalency on the binding kinetics. The first binding event is intermolecular and therefore slower than any subsequent binding event in this system. The overall dissociation rate of the multivalent bound system is strongly decreased.

5.3.2 Examples for Multivalency in Antiviral Strategies

Since many biologically relevant processes are mediated by multivalent interactions, it is also possible to intervene by using designed synthetic multivalent structures. A heavily studied field in this context is the infection inhibition of viruses, which relies on the presentation of numerous specific ligands or receptors on a cell surface. The work presented here was focused on the inhibition of influenza virus and therefore already existing strategies in this context are presented in the following section.

By far the most studied multivalent scaffolds for the inhibition of influenza virus infection are polymeric carrier systems.[14] Again *Whitesides* contributed many studies to this field, many of which involved linear polyacrylamide backbones decorated with *C*-linked sialic acid residues.[46] With this system they were able to achieve one of the highest enhancement factors known so far for multivalent influenza inhibition as determined by an erythrocyte agglutination assay (10^8).[47] Apart from linear polymers a lot of globular structures were also studied, e.g. polyglycerol dendrons,[48] liposomes[49] and nanoparticles,[48] and even proteins were modified with sialic residues in order to inhibit influenza infection.[50] All the aforementioned systems share a common trait in that they present many sialic acid residues on a certain surface and therefore mimic and compete with the actual ligand decorated surface of a host cell. However, these architectures are hardly rationally designed and are rather the result of many screens and trial-and-error efforts because of polydispersity, random coil characteristics, dynamics and unselective conjugation chemistries.

Another inhibitor design strategy involves precisely matching smaller di- and trivalent entities which target a certain receptor. Those were for example synthesized from a central phenyl moiety which served as branching point for peptidic or ethylene glycol linkers, terminated by sialic acid residues.[51] Others used modular DNA scaffolds to present two sialic acid residues with varying distances in order to screen for the optimal binding geometry and to achieve high binding affinities.[52] Their DNA scaffold was additionally compared to bivalent ethylene glycol-based systems which displayed sialic acid residues in similar mean distances. Binding affinities for these systems were rather in the range of monovalent interactions, which clearly demonstrates the influence of linkers and their rigidity (compare Section 5.3.1). A trivalent approach with DNA three-way junctions was also described, in which sialic acids were positioned in an optimal spatial arrangement to fit to the HA binding sites.[53] By attaching three sialic acid residues per arm of the three-way junction, a high affinity binder was synthesized, which exhibits a combination of chelating and clustering effects. However, the binding affinity was lower when five sialic acid residues were attached on each arm, most likely because steric hindrance came into play. Not only DNA but also proteins were used for the design of tailored HA binders. Such a protein-based inhibitor was developed by a computational approach in which a monomeric broadly neutralizing antibody-derived protein was designed to fit into one HA binding site.[54] In a second step, this protein was trimerized in order to fit to the trimeric HA geometry, which led to low and even sub-nanomolar binding affinities against several HA strains. These approaches are mainly based on structural information, e.g. crystal structures of respective receptors, and are comparable to classical rationally designed small molecule inhibitors. Therefore, the exact number of ligands of the system is known as well as their spatial arrangement and binding conformation, both of which are crucial to gaining deeper understanding of the underlying binding modes. The overall enhancement factors for these smaller systems are remarkably large, since there are only a small number of ligands which fit exactly

to the target. However, the steric shielding which can be provided by aforementioned larger, polymeric architectures is often a very important or even dominating factor in terms of effective infection inhibition, which cannot be provided by these smaller structures.[55]

5.4 Virus-like Particles

Virus-like particles (VLPs) are supramolecular self-assembled structures based on expressed capsid proteins of viruses.[56] VLPs are non-infectious as they do not contain the actual viral genome. The shapes and sizes are determined by the original virus, with helical rod-like and icosahedral being the most prominent ones. Sources for VLPs have often been plant and animal viruses, as well as bacteriophages.

VLPs offer several advantages as tools in natural and applied sciences: 1.) The inherent monodispersity of expressed proteins and defined symmetry of assembled VLPs is often beneficial in terms of characterization and potential medical applications;[57] 2.) particle dimensions in the nanometer range are comparable to extensively studied inorganic nanoparticles[58] which enables potential similar fields of applications with proteinaceous particles; 3.) the possibility to genetically modify and incorporate bioorthogonal handles enables their precise spatial placement as well as providing knowledge about the number of functional moieties on the exterior and/or interior surfaces of a VLP.[59]

5.4.1 VLPs as Vaccines

VLPs are currently most widely used in the field of vaccination. The major advantage over commonly used inactivated or attenuated viruses as vaccines is the substantially increased security level of VLP-based agents.[60] As mentioned before, VLPs which do not carry the viral genomic material at all can neither replicate nor revert to an enhanced pathogenic phenotype in the vaccinated individual. Furthermore, the production of VLP-based vaccines can be achieved in significantly less time compared

to attenuated viruses, which can be of importance when it comes to sudden outbreaks of infections and prevention of epidemics.

To date, there are five approved VLP-based vaccines against three viruses (Table 1): 1.) Hepatitis B Virus (HBV); 2.) Hepatitis E Virus (HEV); and 3.) Human Papillomavirus (HPV) (Figure 7).[61]

Table 1: Approved VLP-based vaccines

Trade Name	Against	Company	Licensure
Recombivax HB®	HBV	*Merck*	1986
Engerix®-B	HBV	*GlaxoSmithKline*	1988
Hecolin®	HEV	*Xiamen Innovax Biotech*	2011
Gardasil®	HEV	*Merck*	2006
Cervarix®	HPV	*GlaxoSmithKline*	2009

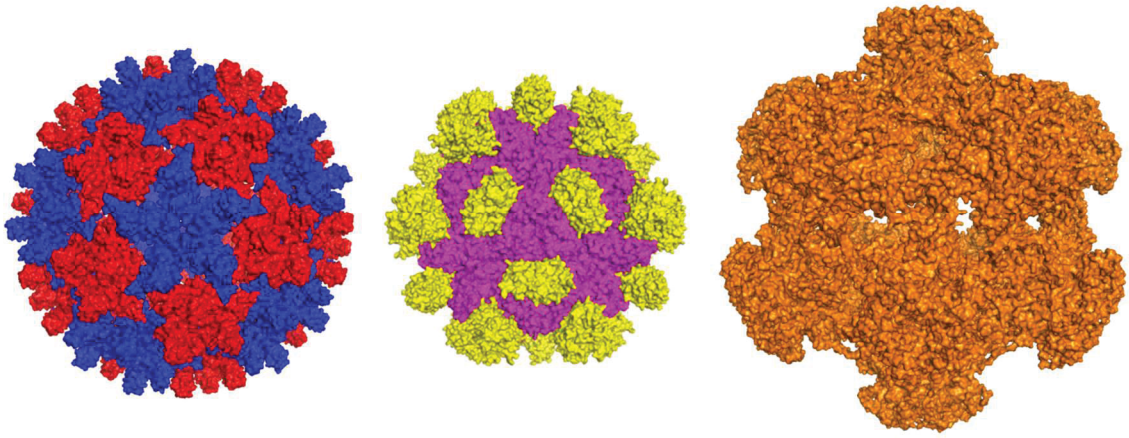


Figure 7: Crystal structures of HBV (left, pdb: 1QGT),[62] HEV (middle, pdb: 3HAG)[63] and HPV (right, pdb: 1DZL)[64] viral capsids. There usual particle sizes are about 40 nm, 35 nm and 50 nm, respectively.

It should be noted that the use of VLP vaccines is not limited to increasing protection against the original virus. Foreign antigenic epitopes can be presented on the surface of a VLP carrier through chemical or genetic fusion. This was for example shown by fusing peptides that represent loops of HIV's surface protein gp120 to the capsid of the RNA bacteriophage MS2, which results in potent immunogenicity of these constructs.[65] The bacteriophage MS2 is a (+)ssRNA virus that naturally infects bacteria of the *Enterobacteriaceae* such as *E.coli*. The genome of MS2 was actually the very first one to be sequenced[66] completely by the group of

Walter Fiers in 1976.[67] VLPs of MS2 consist of 180 copies of the capsid coat protein which assemble into an icosahedral particle of 27.5 nm in diameter (Figure 8, left).[68] A chemical VLP surface modification was used to synthesize a therapeutic vaccine for nicotine addiction.[69] In this study, succinimidyl-activated nicotine was conjugated to solvent-exposed amino acids on the surface of the bacteriophage Qβ. Qβ belongs to the same viral family as MS2 and is therefore structurally very much similar to MS2 (Figure 8, right).[70] Those nicotine-Qβ conjugates were found to induce potent antibody responses in mice and humans.

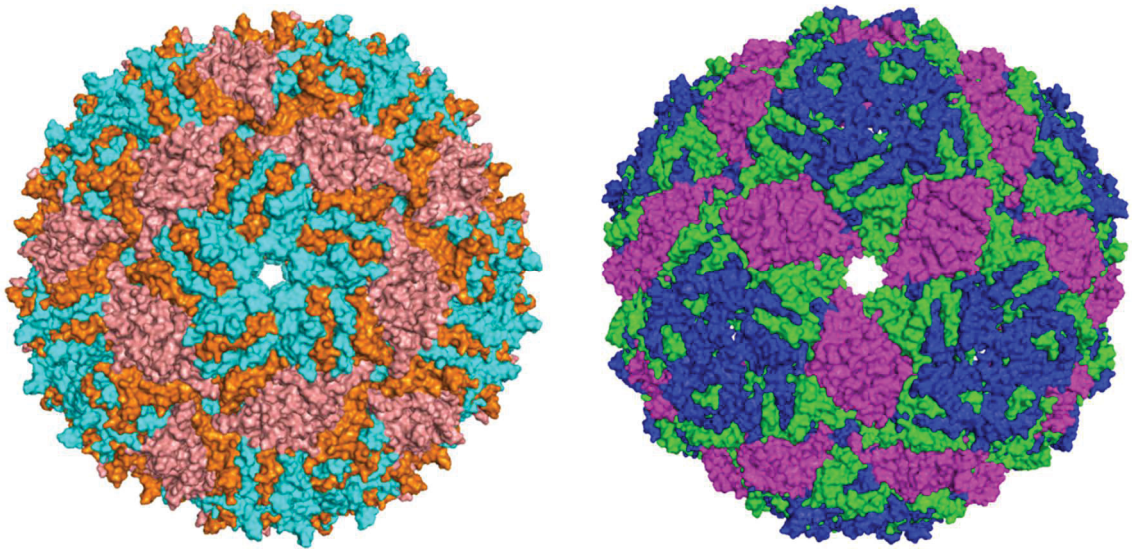


Figure 8: Crystal structures of the two very similar MS2 (left, pdb: 2MS2)[68] and Qβ (right, pdb: 1QBE)[70] bacteriophages.

The aforementioned strong induction of immune responses can also be exploited for the development of therapeutic vaccines against endogenous antigens which play a role in chronic diseases. The formation of antibodies against an organism's own proteins (autoantibodies) is usually hampered by the natural tolerance of the immune system towards these. This was shown to be surmountable by, for example, covalently attaching TNF-related activation-induced cytokine (TRANCE) to the bacteriophage Qβ capsid, resulting in an autoantibody response against TRANCE in mice.[71]

5.4.2 VLPs as Viral Surrogates

Research on highly infective and pathogenic viruses like Ebola and Marburg requires laboratories of biosafety level 4 (BSL-4) and careful handling of samples. Both are members of the *Filoviridae* family and are therefore enveloped, (-)ssRNA viruses[72] which cause severe hemorrhagic fever in humans with exceptionally high mortality. Virions of the *Filoviridae* family are helical filamentous particles often 'U'- or shepherd's crook-shaped with a diameter of 80 nm and usual filament lengths of 0.8-1 µm.[73] It was shown that coexpression of viral matrix proteins and surface glycoproteins is sufficient to trigger the assembly of Ebola and Marburg VLPs of similar morphology compared to authentic virions without infective danger.[74] These VLPs were still able to induce corresponding immune responses, which allows research on those highly pathogenic viruses without the need to work in BSL-4 facility rendering them as useful viral surrogates.

5.4.3 VLPs as Imaging Agents

Another field of application for VLPs is their use as imaging agents. Particularly for magnetic resonance imaging (MRI) a number of studies have been performed in which VLPs were utilized to improve the sensitivity and contrast of this technique. The lower nanometer sized and rigid structures with large rotational correlation times increase water relaxivity rates and therefore sensitivity. Furthermore, potentially hundreds to thousands of contrast agent molecules (commonly Gd^{3+}-complexes) could be attached to a single VLP which significantly increases contrast. Furthermore, additional targeting moieties could be attached to the exterior surface of a VLP in order to locate the contrast agent-loaded particle to its desired location in an organism which further increases contrast.[75] Tobacco mosaic virus (TMV) VLPs were for instance conjugated with Gd(DOTA) complexes at either the exterior or interior surface of the particle by means of copper-catalyzed azide-alkyne cycloaddition (CuAAC) reactions (*vide infra*, section 5.5).[76] TMV is a helical, rod-

shaped (+)ssRNA plant virus with 300 nm in length and about 18 nm in width (Figure 9).[77] CuAAC catalyzed TMV-Gd(DOTA) conjugation results in the loading of up to 2000 or 3500 Gd complexes, respectively, significantly increasing the relaxivity per particle. Recent efforts to use xenon as a contrast agent for MRI resulted in the development of cancer cell targeting MS2-based constructs which carry multiple cryptophane A moieties as Xe-cages.[78] The study demonstrates the high sensitivity that can be achieved by combining specific targeting with high loading of contrast agent.

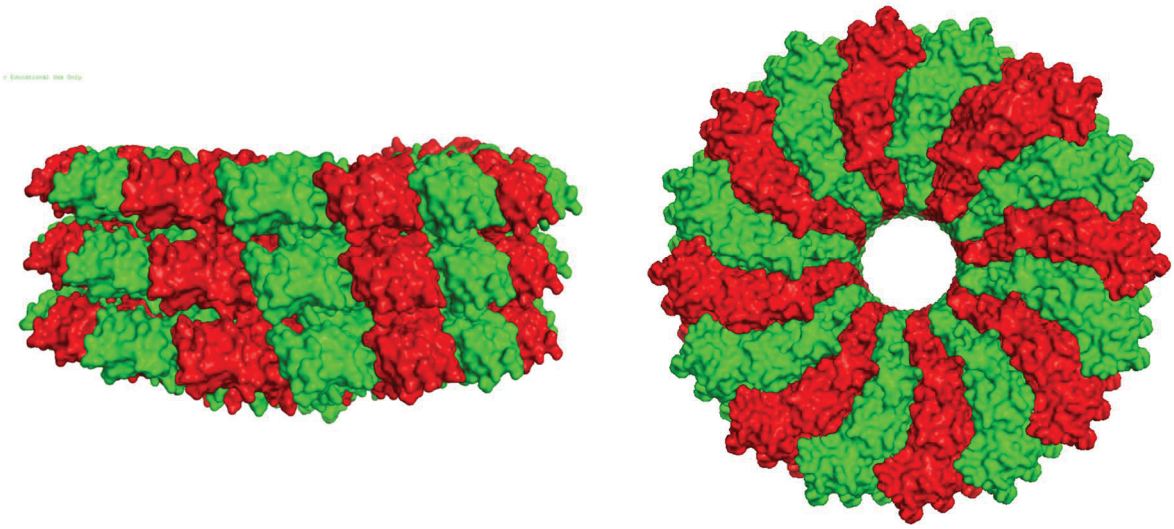


Figure 9: Excerpt of the crystal structure of helical TMV as a side view (left) and a cross-section (right) through the helical filament (pdb: 2TMV)[79].

Apart from MRI applications, it was shown that Qβ VLPs carrying green fluorescent proteins (GFPs) in their interior are useful for optical imaging.[80] Straightforward production together with the enhanced thermal stability and additional targeting possibilities of the encapsulated GFP shows again the versatility and utility of VLP platforms.

5.4.4 VLPs as (Targeted) Delivery Vehicles

VLPs come with a large interior cavity which can potentially serve as a container for certain cargo. The protein shell would furthermore protect cargo against unwanted recognition, degradation and systemic toxicity in addition to functioning as a

biocompatible and soluble packaging which can be located to a specific target. All of these characteristics are important prerequisites for drug delivery vehicles.[60] One example of a VLP-based drug delivery vehicle is again based on MS2 VLPs.[59b] Upon incorporation of an additional cysteine residue on the interior surface of the MS2 capsid and subsequent conjugation with a *Taxol*-maleimide, water soluble MS2-*Taxol* conjugates were synthesized. In cell-based viability assays it was shown that these conjugates exhibit similar toxicity profiles as the free drug at similar concentrations accompanied with the major advantage that no solubilizing agents (such as DMSO or a detergent) were necessary. MS2 was also used as scaffold for the efficient targeting of human hepatocellular carcinoma (HHC) cells.[81] In this study, the capsid surface was modified with the SP94 peptide which binds to HHC cells about 10,000-fold better than hepatocytes for selective targeting of the cancerous cells. The MS2 VLPs were loaded with DNA-conjugated anti-cancer drug cocktail containing doxorubicin, cisplatin and 5-fluorouracil which eventually led to selective cell death of HHC cells at low nanomolar drug concentrations. When they loaded DNA-conjugated ricin toxin A-chain and additionally presented an endosomal escape peptide on the VLP surface, the whole HHC cell population was killed at 100 fM concentration of the toxin.

The symmetrical and often known structure of VLPs also offers the possibility for the attachment of a defined number of ligands by the incorporation and conjugation of unnatural bioorthogonal reporters (*vide infra*, section 5.5 and 5.6), and therefore control over the distribution and valency of synthesized constructs. This strategy was for example followed by *Finn* and co-workers when they used the Qβ capsid which carried a known number of alkynes in the protein sequence at known positions. After copper-catalyzed alkyne-azide cycloaddition reactions (*vide infra*, section 5.5) to attach ligands for the CD22 receptor of B cells to Qβ capsids, the resulting particles were shown to bind tightly to CD22 expressing cells.[82] This Qβ scaffold was later also used by *Davis* and co-workers to efficiently block cellular Ebola infection.[83] in this case, the Qβ capsids was used to present an exact number of mannose

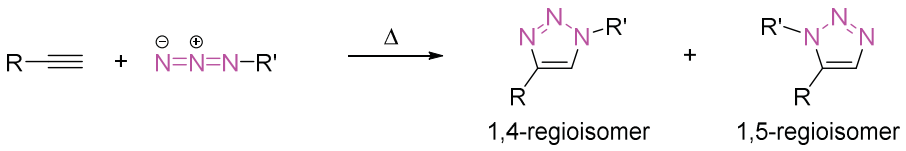
glycodendrons homogeneously distributed over the whole capsid surface. In that way it was possible to address and block DC-SIGN of T-lymphocyte cells, and to protect thereby these cells from Ebola infections. A clear multivalent enhancement of inhibitory properties was observed for the Qβ-glycodendron conjugates, compared to monomeric mannose. The authors hypothesized that the symmetrical presentation of their glycodendrons played an important role for efficient binding, but the structural investigation was not presented in that study.

One important aspect to keep in mind, when materials such as VLPs are used for biomedical purposes, is that these potentially induce an immune response. Although this property is desired when VLPs and conjugates thereof are used as vaccines (see Section 5.4.1), it could in the same way hamper the use of VLPs as drug delivery vehicles.[84]

5.5 Copper-catalyzed Alkyne-Azide Cycloaddition as Bioconjugation Reaction

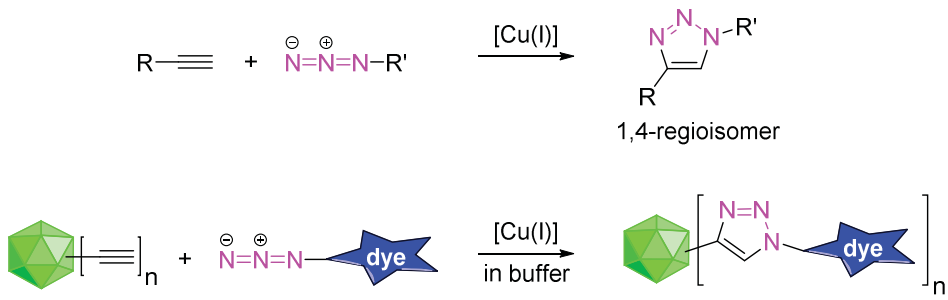
Some of the previously mentioned approaches to present targeting moieties or antigenic epitopes on the surface of VLPs employed protein modification reactions known as bioconjugation reactions. These types of reactions have to fulfill special requirements compared to common reactions in organic chemistry and are often members of the class of bioorthogonal reactions.[85] Bioorthogonality was originally coined as a term for reactions that can occur in living organisms (cells[86] or even higher organisms such as zebrafish[87]) and that meet certain special criteria: 1.) unique functional groups (i.e. bioorthogonal reporters) are necessary in order to achieve selectivity in a biological framework; 2.) these functional groups and every mechanistic step of the reaction need to be compatible with water and large excesses of biologically occurring nucleophiles (thiols, amines, hydroxyl groups); 3.) the reaction has to be fast at low concentrations of the reaction partners and has to proceed at ambient temperatures; 4.) it must be possible to incorporate at least one

of the reaction partners into a biomolecule; 5.) the formed bonds must be stable under biological conditions. Since these conditions are not only a prerequisite for chemical reactions in living organism but also applicable and rather beneficial for *in vitro* conjugations of isolated biomolecules, a plethora of bioorthogonal reactions has been developed, e.g. for protein labeling purposes.[88] One of the most prominent examples is arguably CuAAC.[89] The reaction of 1,3-dipoles such as azides with dipolarophiles such as alkynes to form 1,2,3-triazoles was initially studied by *Huisgen* in the 60s.[90] This original *Huisgen* reaction was promoted solely by heating and delivered a mixture of 1,4- and 1,5-regioisomers of the corresponding triazole (Scheme 1).



Scheme 1: Thermally induced alkyne-azide cycloaddition reaction delivers two regioisomers of the corresponding 1,2,3-triazole product.

In 2002, two research groups reported independently from each other a copper(I)-catalyzed version of an alkyne-azide cycloaddition. The group of *Sharpless* performed a mainly small molecule-based study and screened different alkyne-azide pairs and reactions conditions.[91] Comparable studies were performed in the lab of *Meldal* in which their main focus was in the context of solid-phase peptide synthesis.[92] Both papers reported the exclusive 1,4-selectivity of CuAAC at moderate temperatures and acceptable time scales (Scheme 2, top). It was additionally mentioned in *Sharpless*' paper, that the reaction proceeds evenly well in human plasma, which already showed the potential applicability of this reaction in a biological setting. Only one year later, in 2003, the group of *Finn* reported together with *Sharpless* the bioconjugation reaction between cowpea mosaic virus (CPMV) VLPs and fluorescein labels.[93] Alkyne or azide moieties were introduced at reactive lysine or cysteine residues of CPMV and subsequent CuAAC was performed with the corresponding dye-counterpart (Scheme 2, bottom).



Scheme 2: Copper-catalyzed variant of the alkyne-azide cycloaddition reaction, which delivers exclusively the 1,4-regioisomer as a product (top) and which is compatible with aqueous and ambient conditions (bottom) as demonstrated by the reaction between a viral particle (CPMV, green icosahedron) and a dye (blue star).

It had been already observed at that time that further acceleration of the CuAAC reaction was possible by the addition of tris(triazolyl)amine ligands. In 2004, again one year later, this observation was published by *Fokin* together with *Sharpless* in which they introduced tris-(benzyltriazolylmethyl)amine (TBTA) as a prominent example of rate-accelerating and Cu(I)-stabilizing ligands.[94] Further studies led to the development of many different tris(triazolyl)amine ligands with tuned and optimized properties for certain scenarios, such as tris[(1-(2-ethoxy-2-oxoethyl)-1*H*-1,2,3-triazol-4-yl)methyl]amine (TEOTA),[95] tris[(1-hydroxypropyl-1*H*-1,2,3-triazol-4-yl)methyl]amine (THPTA)[96] and 2-[4-[(bis[(1-tert-butyl-1*H*-1,2,3-triazol-4-yl)methyl]amino)methyl]-1*H*-1,2,3-triazol-1-yl] acetic acid (BTTAA).[97] THPTA and BTTAA are particularly useful for bioconjugation reactions since they are water-soluble and therefore applicable in biological environments. Another important feature of these Cu(I)-chelating ligands is that they intercept copper-catalysis derived reactive oxygen species and therefore reduce cytotoxic effects (Figure 10).[98]

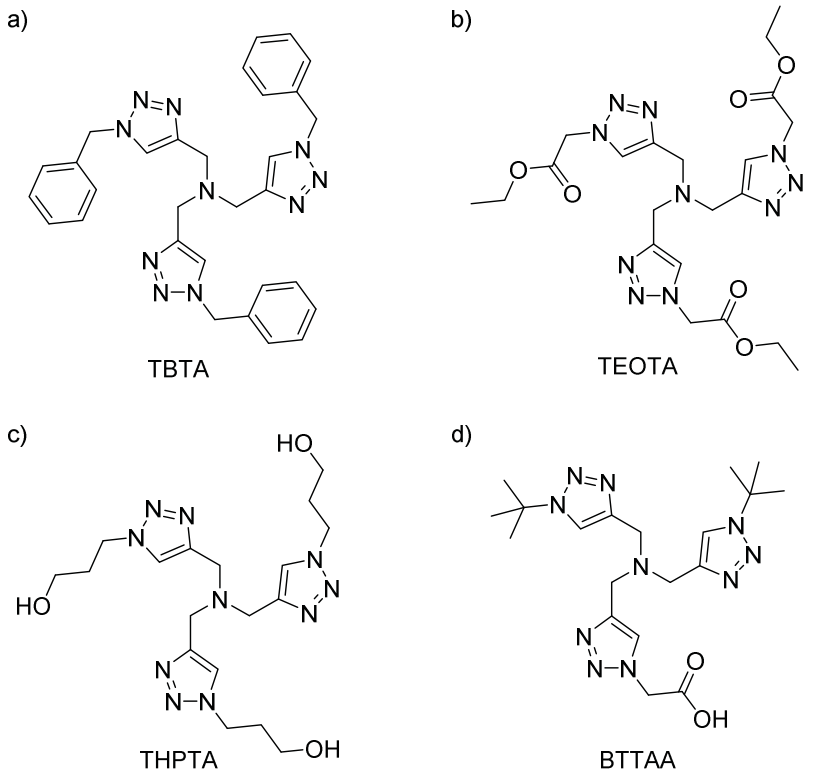
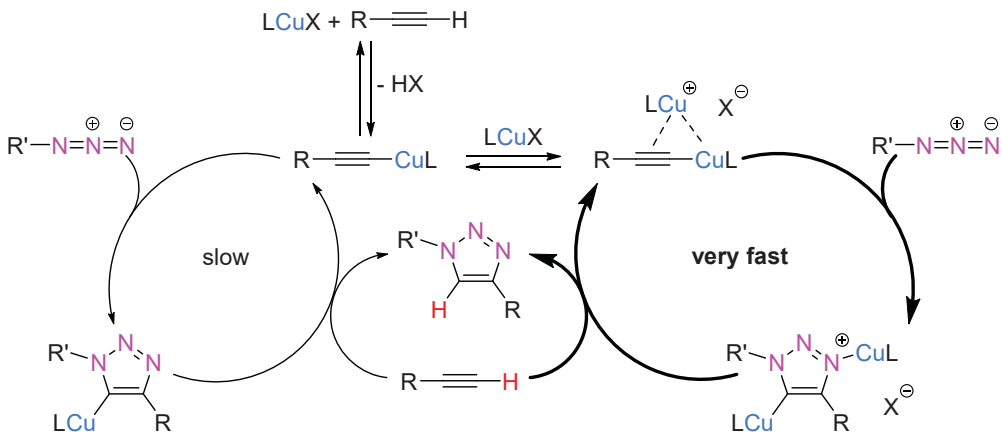


Figure 10: Commonly used CuAAC accelerating ligands: a) tris-(benzyltriazolylmethyl)amine (TBTA); b) tris[(1-(2-ethoxy-2-oxoethyl)-1*H*-1,2,3-triazol-4-yl)methyl]amine (TEOTA); water-soluble c) tris[(1-hydroxypropyl-1*H*-1,2,3-triazol-4-yl)methyl]amine (THPTA); and d) 2-[4-[(bis[(1-tert-butyl-1*H*-1,2,3-triazol-4-yl)methyl]amino)methyl]-1*H*-1,2,3-triazol-1-yl] acetic acid (BTTAA).

Since CuAAC is such a versatile and robust method, a lot of studies were committed to elucidate the corresponding reaction mechanism. Early studies suggested that a mononuclear copper-acetylide serves as key intermediate in the catalytic cycle of the CuAAC.[99] However, increasing evidence was found by experimental and theoretical investigations that a dinuclear copper-acetylide-azide complex is involved in the catalytic process.[100] In a mass spectrometry focused report it was furthermore shown that the use of the above tris-(triazolyl)amine ligands promotes the formation and stabilization of even trinuclear copper-complexes which could play a role in the catalytic cycle.[101] This was interestingly already proposed by *Meldal* in 2008.[102] A very elegant work by *Bertrand* describes for the first time the isolation of key intermediates of the catalytic CuAAC cycle.[103] With these isolated compounds in hand, it was possible to pinpoint that both the mononuclear and dinuclear copper

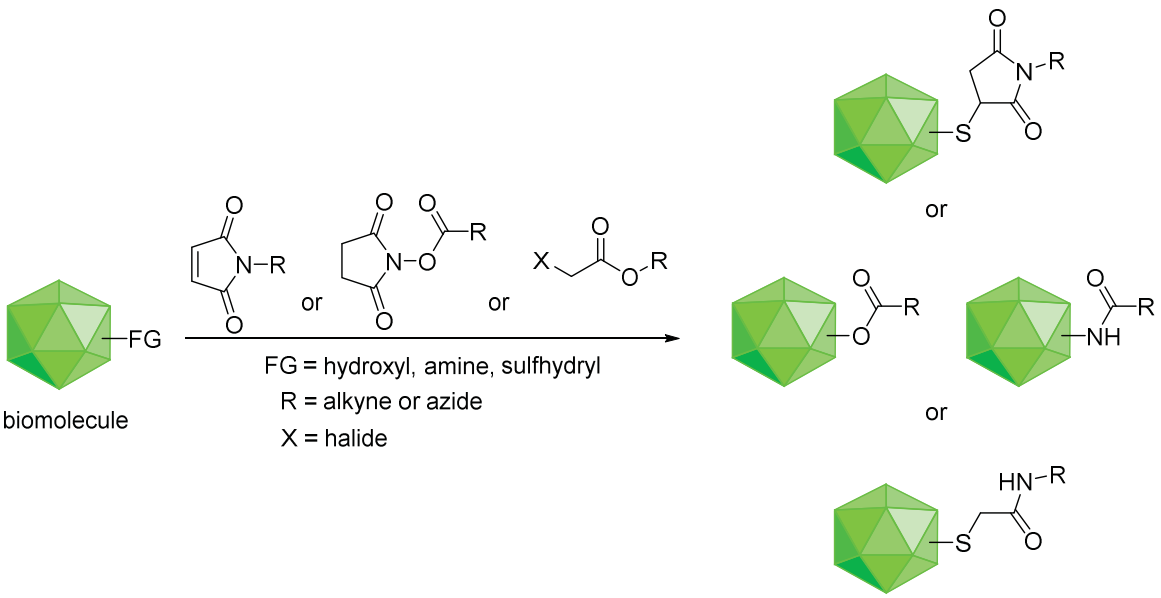
complexes are active intermediates, but the dinuclear pathway is kinetically strongly favored (Scheme 3).



Scheme 3: The two CuAAC pathways which occur during the reaction. Mononuclear pathway (left cycle) exhibits significantly slower reaction kinetics compared to the one with a dinuclear copper species as an intermediate (right cycle).

5.6 Incorporation of Unnatural Amino Acids

One requirement for a bioorthogonal reaction mentioned before was the presence of at least one functional group compatible with the reaction partners that can be incorporated into a biomolecule, e.g. protein. Many options are available when using the CuAAC reaction, since the reaction partners – alkynes and azides – are both very stable and inert under biological conditions.[93] Natural amino acids can be used to attach these functionalities for example by utilizing nucleophilic residues such as the sulfhydryl of cysteine, amine of lysine and hydroxyl of serine and threonine.[104] These residue-specific methods often utilize maleimide, *N*-hydroxysuccinimide (NHS) or alkyl halide chemistry (Scheme 4) which also come – despite their straightforward handling and efficiency – with some drawbacks: 1.) Selectivity is inversely correlated with reactivity and fine-tuning thereof is often not adequately achievable; 2.) polyfunctionalization often occurs because the protein contains multiple copies of the targeted amino acid, which leads to 3.) heterogeneity of the resulting product and therefore difficulties for subsequent analysis and further applications.



Scheme 4: Classic approaches to incorporate unnatural functionalities into a biomolecule, e.g. protein, by reacting natural nucleophiles such as –OH, -SH, and –NH_2 with corresponding electrophiles such as maleimides, NHS-esters and alkyl-halides.

A possible strategy for overcoming these aspects is to directly incorporate unnatural alkyne- or azide-functionalized amino acids into the protein sequence. Different techniques based on genetic code expansion and selective pressure have been developed to this end,[88b] with the latter being the method utilized throughout this work and therefore described in more detail in the following section.

5.6.1 Selective Pressure Incorporation

Selective pressure incorporation (SPI) is a proteome wide, residue-specific method which replaces a natural amino acid by a supplied unnatural surrogate which carries the functionality of choice for later conjugation steps. It was already shown in the late 1950s that methionine could be replaced completely by selenomethionine in the proteome of *E.coli* and bacterial growth is still properly maintained.[105] The *E.coli* strain used for their study was "methionine-requiring" which basically describes auxotrophy. Auxotrophy is defined by IUPAC as the state of an organism unable to produce essential compounds, such as amino acids, by itself.[106] Therefore, these organisms require exogenously supplied sources, e.g. the amino acid pool of the surrounding medium, which allows for the selective replacement of a certain amino

acid by an analogue. If this analogue is sterically and electronically similar enough to the natural amino acid it should replace, competition for the natural aminoacyl-tRNA-synthetase (aaRS) is possible and incorporation by the endogenous translational machinery is feasible.[88b] Additionally, the codon of the natural amino acid which should be replaced has to be present in the protein sequence of choice, and the expression of this protein has to be induced while the unnatural amino acid analogue is present inside the organism.[107] Many studies utilizing methionine-auxotrophic systems have been published,[108] but also other amino acids such as proline, phenylalanine, leucin, isoleucine and tryptophan can be replaced.[88b] Subsequent bioconjugation reactions benefit from the fact that unnatural amino acids with an alkyne (e.g. homopropargylglycine – Hpg)[108b] or azide group (azidohomoalanine – Aha)[108d] at the side chain are well accepted substrates for the methionyl-tRNA-synthetase of *E.coli*. However, their processing by the methionyl-tRNA-synthetase is accompanied by significantly reduced activation kinetics.[109]

The usual experimental procedure for a methionine-auxotrophic system is the following: 1.) Growing the expression host, e.g. *E.coli,* in chemically defined minimal medium which contains the 20 natural amino acids. This initial growing phase is necessary to obtain the required number of bacteria in a healthy state. 2.) At this state, the growing medium is exchanged to one which is depleted of methionine and supplied with the desired unnatural amino acid. 3.) During the subsequent settling phase, residual methionine which is present in the cells from the previous growing phase is consumed by the cells metabolism. 4.) Only after that, the expression of the desired protein is induced by the addition of an appropriate inducer, e.g. isopropyl-β-D-thiogalactopyranoside (IPTG). 5.) Since there is only the unnatural amino acid methionine-analogue present for the cell to utilize, this is loaded onto the corresponding aaRS and incorporated into the protein sequence (Figure 11).

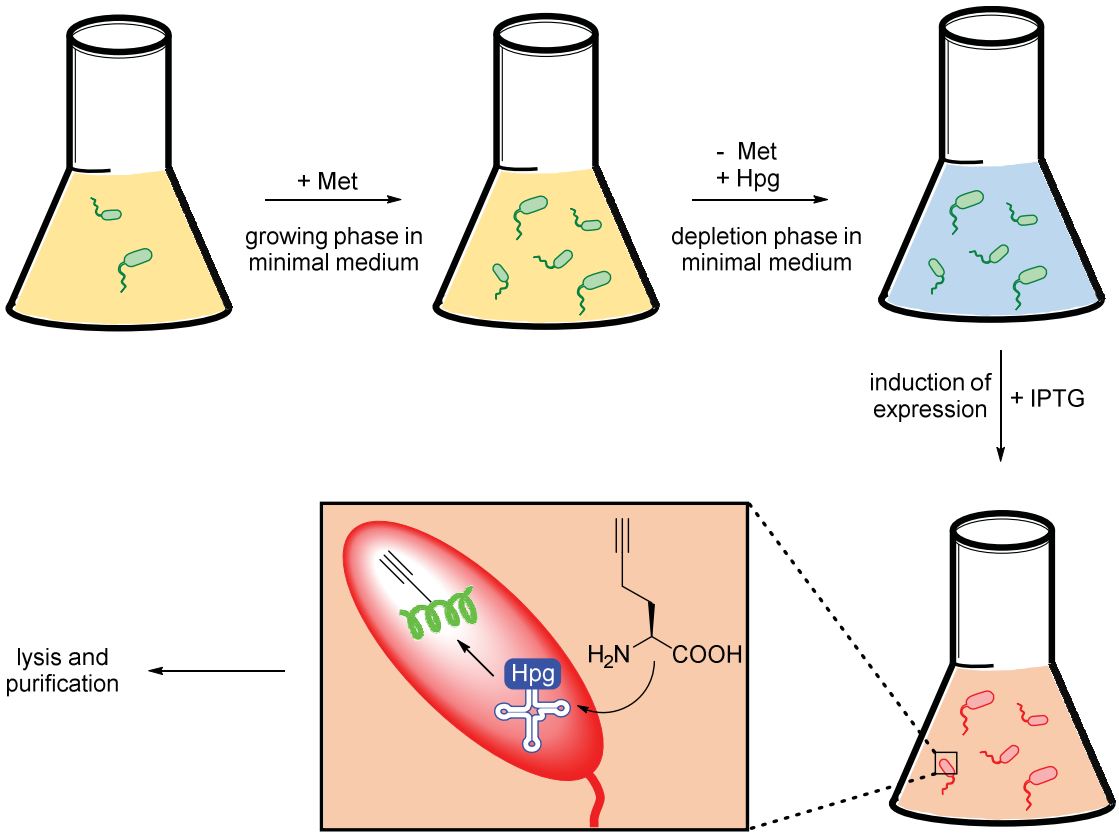


Figure 11: Schematic procedure of a methionine auxotrophic expression setup. Methionine auxotrophic bacteria are grown in minimal medium containing methionine. After reaching a certain bacterial density methionine is depleted from the medium and replaced by an unnatural surrogate, e.g. Hpg. Subsequently, expression is induced and Hpg is incorporated into protein sequences by the endogenous translational machinery of the host.

It has to be mentioned that the above described method alone does not necessarily solve the problem of heterogeneous mixtures in subsequent conjugation reactions, since the amino acid which is to be replaced can occur several times in a protein sequence. However, combined with standard molecular biology techniques, e.g. mutagenesis, it is possible to end up for example with a unique methionine residue which is replaced by an unnatural amino acid and leads eventually to a unique bioorthogonal handle in the resulting protein.

5.6.2 Amber Suppression

The genetic code which is used to translate a gene into a protein contains stop codons as a signal for the ribosome to release the synthesized protein. Three stop codons are available (UAA – ochre, UAG – amber, UGA – opal) which usually do not

encode for an amino acid but bind to release factors. However, two naturally occurring amino acids were identified which are incorporated into a protein by the suppression of such a stop codon and the availability of compatible tRNAs: selenocysteine (UGA) and pyrrolysine (UAG).[110] *Schultz* an co-workers developed and used this stop codon suppression for the incorporation of an unnatural amino acid (*O*-methyl-L-tyrosine) into a protein sequence.[111] The crucial step in this study was to generate a tRNA/aaRS pair which was able to bind to the amber stop codon and to specifically accept the unnatural amino acid, respectively. Unnatural amino acids which are useful for further conjugation reactions, e.g. *p*-azido-L-phenylalanine[112] and *p*-propargyloxyphenylalanine,[113] have also been incorporated by this technique.

The advantages of stop codon suppression are the selective substitution, incorporation and positioning of unnatural amino acids in the protein sequence, comparable to site-directed mutagenesis.[111] Efficiency and fidelity of the technique can be very high and a large variety of unnatural amino acids can be incorporated. By the combination of more than one stop codon suppression and the corresponding tRNA/aaRS pairs it was even possible to incorporate multiple unnatural amino acids into one protein.[114] However, the generation of or search for compatible tRNA/aaRS pairs can be cumbersome. Furthermore, the binding of these stop codons to the actual release factors always represent a competing event which would result in truncated proteins and eventually to heterogenous protein mixtures.

5.7 Sialic Acid

5.7.1 General

Sialic acid is a hypernym for derivatives of the nine-carbon monosaccharide neuraminic acid (Neu). The term “neuraminic” was chosen because it was initially cleaved from brain lipid samples in which a nitrogen-containing organic acid was identified.[115] Unsubstituted Neu is not found in nature but a plethora of different *O*-

and *N*-substitutions and combinations thereof have been discovered so far.[116] More precisely, acetylation, phosphorylation, sulfation and methylation generate the more than 40 different neuraminic acid derivatives currently known (Figure 12). The most prominent example of these is *N*-acetylneuraminic acid (Neu5Ac) which is the predominant variant found in mammals. Neu5Ac is usually the terminal sugar of glycan chains of cell-surface glycoproteins as well as soluble proteins. It fulfills various tasks, including mediating cell-cell interactions, participating in immune response, masking underlying epitopes, being a ligand for pathogen receptors, and many more.[116-117]

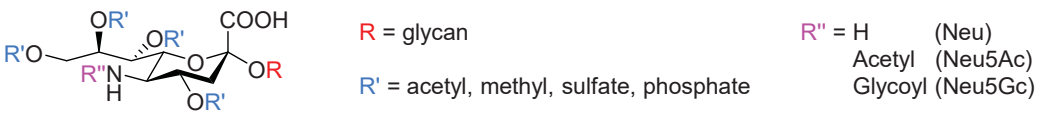
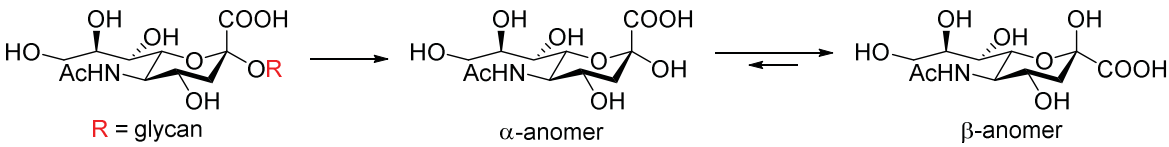


Figure 12: Natural occurring neuraminic acid derivatives with different *O*- and *N*-substituents.

The naturally occurring sialic acid conjugates display exclusively α-linkages to the penultimate sugar unit.[118] This is however different when the sialic acid residues are cleaved off from the actual conjugate which leads to an equilibrium state with 90% of the β-anomer as the predominant form (Scheme 5).

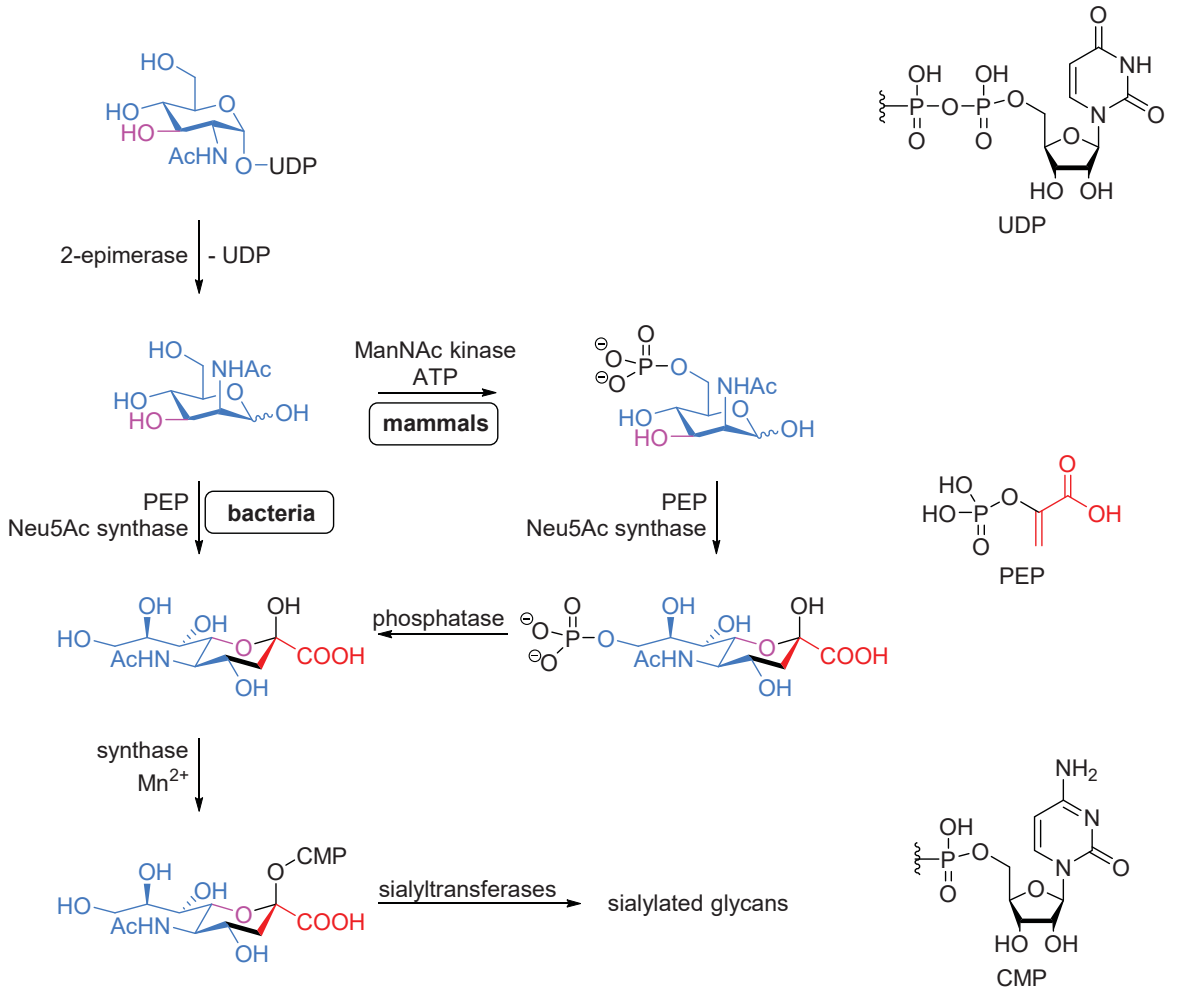


Scheme 5: Anomeric configuration of Neu5Ac in glycans (left) and free in solution (middle and right).

Apart from being essential for influenza viruses to infect host cells (*vide supra*, section 5.2), sialic acids can also be involved in storage diseases which are a direct result of an improper sialic acid metabolism regulation.[119] This leads to neurodegenerative symptoms such as retardation, spasticity and epileptic seizures. The infantile free sialic acid storage disease (ISSD) is the more severe form, leading to massive developmental delays, enlarged livers and hearts and usually very early death of affected children.

5.7.2 Biosynthesis

The biosynthesis of Neu5Ac is quite similar for bacteria and mammals (Scheme 6).[120] It starts with uridine diphosphate-*N*-acetylglucosamine (UDP-GlcNAc) which is transformed to *N*-acetylmannosamine (ManNAc) by a 2-epimerase. An additional phosphorylation step occurs in mammalian cells at this stage, leading to 6'-phosphorylated ManNAc-6-P.



Scheme 6: Biosynthetic pathways of Neu5Ac in bacteria and mammals.

The next step is similar for both species and involves the action of a Neu5Ac synthase which promotes the condensation of phosphoenolpyruvate (PEP) and ManNAc or ManNAc-6P, respectively. A phosphatase in mammals cleaves of the accompanying 9'-phosphate of the resulting Neu5Ac-9-P and generates Neu5Ac, which is the direct product of the Neu5Ac synthase in bacteria. In the last step, a

cytosine monophosphate (CMP) is introduced at the anomeric carbon of Neu5Ac to eventually form CMP-sialic acid. This sialic acid building block is now a substrate for sialyltransferases and can be incorporated into glycan structures by the organism.

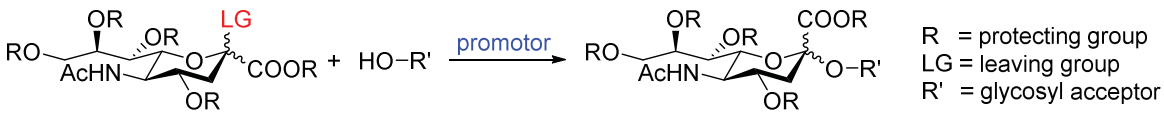
It was reported by *Reutter* that this biosynthetic pathway tolerates minor structural variation of the *N*-substituent, demonstrated with elongated *N*-acetyl groups.[121] This finding was already exploited during the birth of bioorthogonal conjugation reactions (*vide supra*, section 5.5) for the purpose of incorporating ketones and azides as reactive and selectively addressable handles on a cell surface.[122] In these cases, *Bertozzi* and co-workers fed cells with several unnatural ManNAc derivatives. These were transformed by the cells metabolism and ended up as sialic acid residue on the cell surface glycans. It was also shown that not only modified *N*-acetyl groups were accepted in the biosynthetic pathway. By using a 4-azido ManNAc derivative it was possible to generate 7-azido sialic acid derivatives on cell surfaces, which could also be labeled in living zebrafish larvae.[123]

5.7.3 Chemical Glycosylation

The main challenge for the chemical synthesis of sialic acid-containing glycans is to ensure the exclusive formation of α-linkages between the glycan and the sialic acid. Glycosylation methods in this direction are hampered by particular structural features of sialic acid. First, the C-3 methylene group makes sialic acid very susceptible to elimination reactions, especially under harsh reaction conditions leading to the formation of the corresponding glycal. In addition, with the missing substituent at C-3 it is furthermore not possible to exploit neighboring group participation in corresponding glycosylation reactions. Finally, the carboxyl group at the anomeric position destabilizes intermediately formed oxocarbenium ions through its electron withdrawing character.[124]

Chemical glycosylation reactions usually consist of: a glycosyl donor, which carries a suitable leaving group at the anomeric carbon; a glycosyl acceptor, usually carrying a

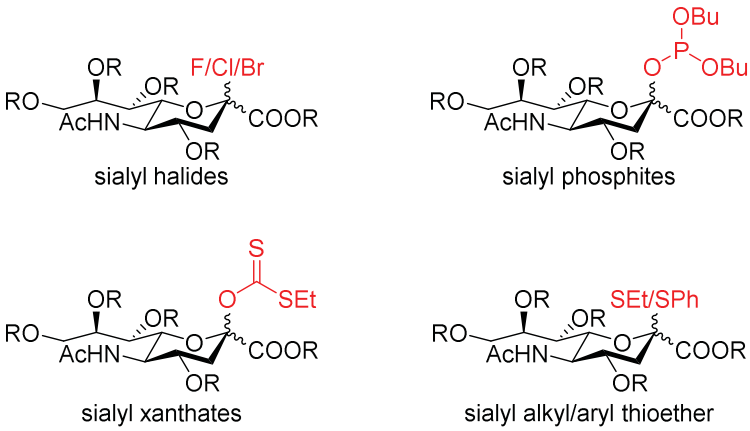
free nucleophilic hydroxyl group; and a promotor which initiates the reaction (Scheme 7). In most cases, the reaction can be classified as an S_N1 or S_N2 nucleophilic substitution, and therefore the anomeric outcome of the glycosylation reaction can be tuned by the nucleophilicity of the acceptor.[125]



Scheme 7: Typical chemical glycosylation reaction which involves a suitable protected sialyl donor with a leaving group, a glycosyl acceptor with a free hydroxyl group and a promotor which drives the desired reaction.

On the donor side, a lot of research was committed towards finding optimal leaving group/promotor combinations in order to efficiently activate a sialic acid donor under preferably mild conditions so as to prevent the above mentioned elimination reaction and to foster the formation of the desired α-linkage. Commonly used leaving groups are halogens, phosphites, xanthates, thioalkyl and thioaryl derivatives (Table 2).[126]

Table 2: Commonly used examples of sialyl donors and their corresponding promotor reagents.



Donor	Promotor
halides	Ag_2CO_3, $HgCN_2/HgBr_2$, $ZnBr_2$
phosphites	TMSOTf
xanthates	DMTST, MeSOTf
alkyl/aryl thioethers	DMTST, NIS/TfOH

Sialic acid halide donors are probably the most easily accessible, with the corresponding chloro derivative being the most frequently used tool for the synthesis of Neu5Ac-containing compounds from the 60s to the 80s.[126] Classical promotors for the activation of chloro donors are heavy metal salts such as silver(I) carbonate[127] and mercury(II) cyanide/mercury(II) bromide.[128] These procedures led only to modest yields and α-selectivities, additionally accompanied with toxicity of the promotors for the environment and health. The situation could be improved by choosing a different promotor. For example, zinc(II) bromide turned out to be especially effective.[129] The corresponding fluoro and bromo donors were also studied but led to predominantly β-linked products[130] or suffered from low stability and therefore competing elimination reactions to the glycal, respectively.[131]

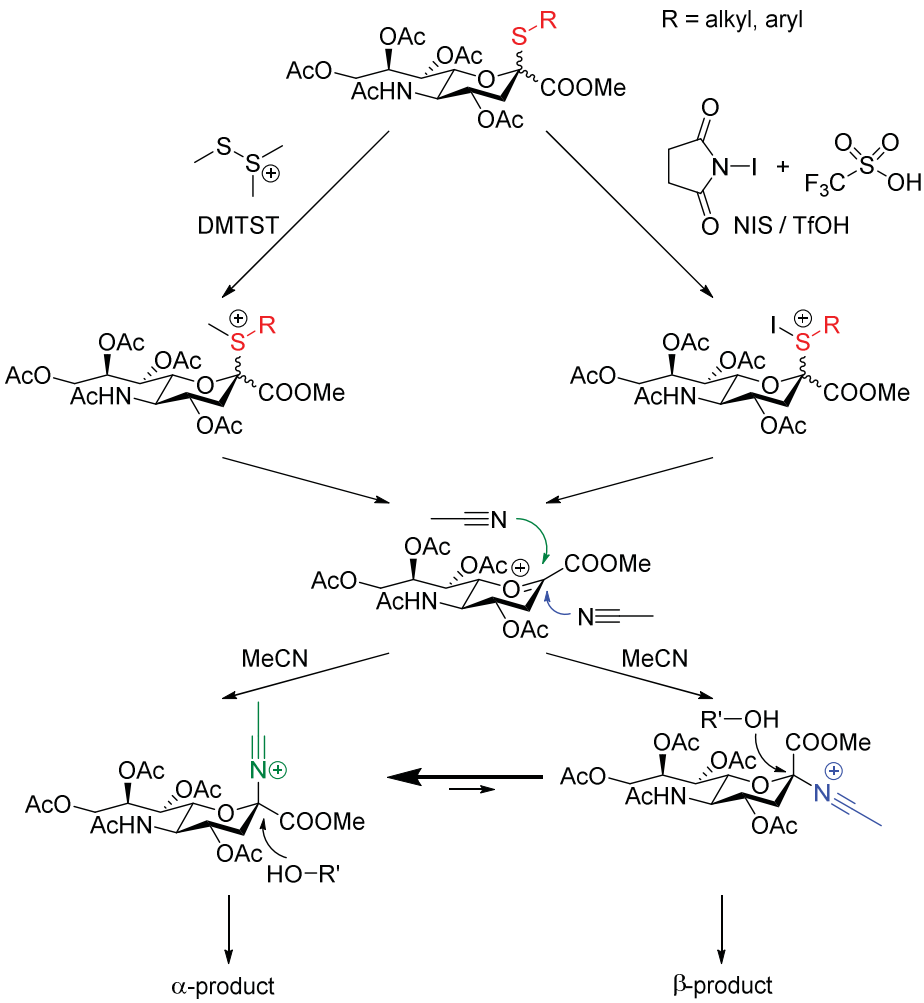
A more modern approach is the use of phosphite donors. A phosphite group can be easily installed at the anomeric position of sialic acid, starting from the free 2-hydroxy group with the corresponding chlorophosphite or phosphoramidite reagents. Two derivatives were devised which delivered good glycosylation yield and also α-selectivities: 1.) diethyl phosphites and 2.) dibenzyl phosphites. Both can be activated by catalytic amounts of trimethylsilyl trifluoromethanesulfonate (TMSOTf), whereas the diethyl derivative gave slightly better yields in the synthesis of the tetrasaccharide sialyl Lewis^{x} (SLe^{x}).

Very frequently used sialic acid donors are thiol containing variants. Xanthates can be activated by dimethylthiomethyl sulfonium trifluoromethanesulfonate (DMTST)[132] or methylsulfenyl trifluoromethanesulfonate (MeSOTf),[133] with the latter being more reactive. The beneficial aspect of this method is that the xanthates can be activated at low temperature, e.g. –70°C, even in the presence of other thiol-containing donor compounds. This allows for the sequential setup of glycans in “one pot”. Such a xanthate donor was for example used for the synthesis of a sialylated threonine

building block for solid-phase peptide synthesis.[134] With that building block it was possible for the first time to synthesize a sialylated Tn antigen on a preparative scale.

Other thiol donors can be either thioalkyl or thioaryl derivaties which can both be activated in principle in the same way as the xanthates (e.g. DMTST).[126] However, significant improvement was achieved after switching to the *N*-iodosuccinimide (NIS)/trifluoromethanesulfonic acid (TfOH) couple as a promotor, especially with regards to the glycosylation of sterically demanding acceptors.[135]

The activation mechanism is the same for both of the promotors mentioned above.[126] The first step is the formation of a sulfonium intermediate, which is a good leaving group. Subsequently, an oxonium ion is formed which is then attacked at the anomeric carbon by the hydroxyl group of an acceptor. It had been shown that acetonitrile as a co-solvent in these reactions significantly increases the yield of the desired α-linked products.[135-136] This can be explained by the reaction of an actual solvent molecule with the oxonium ion resulting in a nitrilium ion which in turn preferentially occurs in the β-configuration. A subsequent nucleophilic attack at the anomeric position leads to an S_N2-type inversion of conformation and therefore predominantly to the α-linkage (Scheme 8).



Scheme 8: Activation mechanism of sialyl donors with an alkyl or aryl thioether leaving group. Intermediately formed sulfonium species are very good leaving groups, leading to an oxocarbenium-species as a common intermediate. Acetonitrile as solvent can support α-selective glycosylation by acting as a directing group.

Major advances in the field of α-selective glycosylation reactions of sialic acid donors were achieved by *Crich* and co-workers. Their thorough studies involved initially *S*-phenyl donor derivatives which carried a special 5-*N*,4-*O*-oxazolidinone protection plus an additional *N*-acetamido protection.[137] They observed very good yields and α-selectivities, even for quite bulky acceptor compounds. The latter could be further improved by switching from the *S*-phenyl to the corresponding *S*-adamantanyl (Ada) sialic acid donor (Figure 13).[138] That change allowed the proper activation of the system even at low temperatures, e.g. –78°C which again significantly improved the α-selectivity. Furthermore, adamantanethiol is a solid compound with less volatility

compared to other low-molecular weight thiols. Therefore, its handling is significantly easier as it can easily be weighed in with less of the pronounced thiol-characteristic unpleasant smell.

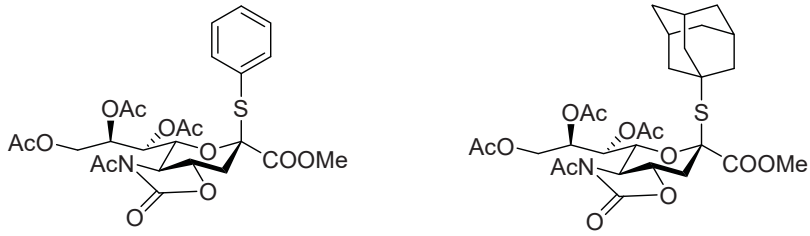
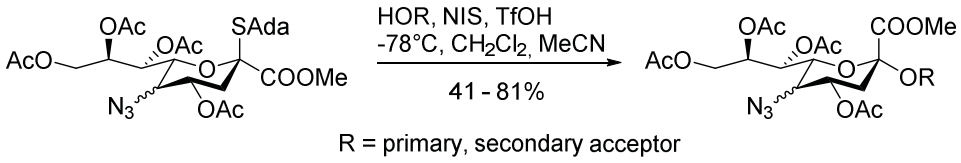


Figure 13: Sialyl donors with a 5-*N*,4-*O*-oxazolidinone and additional *N*-acetamido protecting group as introduced by *Crich*.

The explanation behind these observations is mainly based on two aspects: 1.) The bicyclic character of a 5-*N*,4-*O*-oxazolidinone protected sialic acid leads to a locked conformation which in turn results in 2.) aligned dipole moments of the corresponding carbonyl groups and therefore a pronounced electron withdrawing effect on the anomeric position.[139] This effect destabilizes intermediate oxocarbenium ions and therefore inhibits unwanted elimination pathways.

In a recent work by the same group, it was observed that even without the 5-*N*,4-*O*-oxazolidinone protection high α-selectivities in glycosylation reactions can be achieved.[140] For their synthesis of pseudaminic acid derivatives it was necessary to protect the 5-*N*-acetyl group by an azide and to use this derivative as glycosylation donor (Scheme 9).



Scheme 9: 5-Azidothiosialoside as glycosylation donor led to high yields of desired α-linked products.

They reported very good yields and selectivities for that donor with various primary and secondary alcohols as acceptors. Previously unsatisfactory results for such 5-azido derivatives were also described but the lower activation temperature (–78°C vs.

–40°C) they could apply through the *S*-adamantanyl leaving group was hypothesized as a reason why their donor showed significantly improved results.

6. Objectives

Since the sanitary and economic burdens of influenza infections are as yet unresolved, the search for novel therapeutic and protective strategies carries a great importance. In this regard, it would be of advantage to deepen the understanding of the underlying multivalent interactions between an influenza virus and its host cell (compare section 5.2.1.3 and 5.3). Studies in this manner have so far been hampered by the use of polydisperse influenza binders like polymers that do not allow for a precise quantitative analysis of those multivalent interactions. Scaffold functionalization for well-defined scaffold-virus interaction will be accomplished in this thesis by using the Qβ bacteriophage capsid as a monodisperse proteinaceous scaffold which is decorated at known positions with HA-interacting sialic acid residues. Challenges in this regard are: a.) the provision of sufficient amounts of sialic acid derivatives that can be coupled to the scaffold and which have to be synthesized by means of organic chemistry; b.) the successful conjugation thereof to the Qβ bacteriophage capsid; and c.) the subsequent biophysical and biological evaluation of the resulting constructs and the derivation of a reasonable multivalent binding model based on these results. Therefore, the three main goals of this thesis were defined as:

1.) Optimizing the current state-of-the-art α-selective sialic acid glycosylation strategy by using a synthetically easier accessible sialic acid donor in order to synthesize ethylene glycol-linked sialic acid conjugates. The optimization should result in a reduced number of synthetic steps as well as improved overall yield of the desired α-linked products.
2.) Generation of defined sialic acid-decorated Qβ bacteriophage capsids and analysis thereof.
3.) Evaluation of the inhibitory potential of generated sialic acid-decorated Qβ bacteriophage capsids against influenza viruses and modelling of underlying

binding modes to deepen the understanding of multivalent pathogen interactions.

Glycosylation reactions of sialic acids, though well-studied, still belong in the challenging field of carbohydrate chemistry. The required α-linked products can be obtained almost exclusively by using oxazolidinone protected thioadamantanyl donors (compare section 5.7.3) at the cost of significantly reduced overall yields because of the additional protection steps necessary for the synthesis of the corresponding donor. In the first part of this thesis, the goal was to determine the potential of simple 5-*N*-acetyl adamantanyl thiosialoside **1** as a donor for the actual glycosylation reactions. This compound is an early intermediate in the synthesis of the oxazolidinone protected thioadamantanyl glycosylation donor **2**. The use of an anomeric mixture of **1** would significantly increase the overall yields, since the anomere separation would be performed at a later stage of the synthetic route and more material is used for the actual glycosylation reaction.

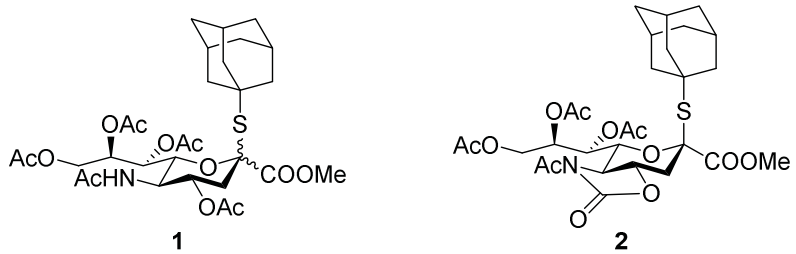
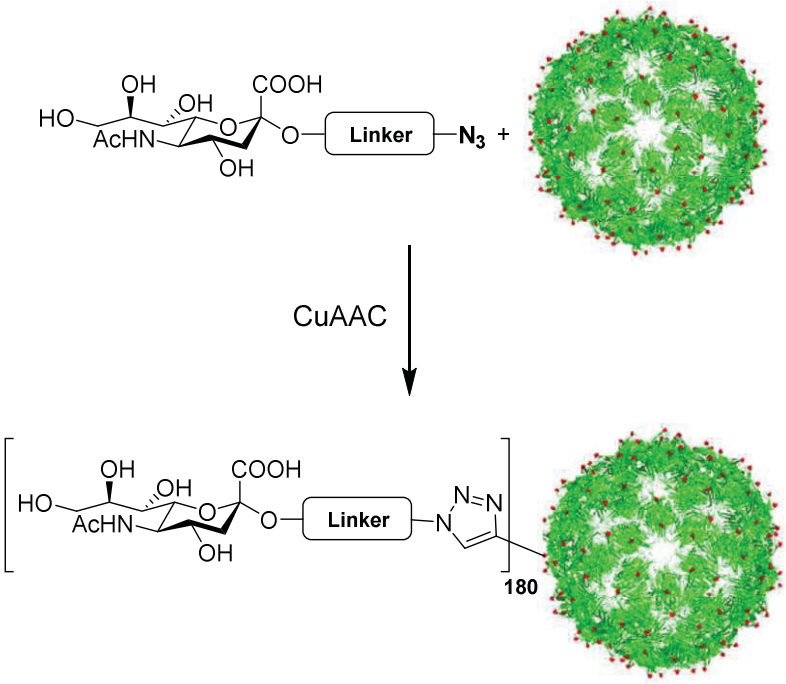


Figure 14: The anomeric mixture of 5-*N*-acetyl adamantanyl thiosialoside **1** (left) and 5-*N*,4-*O*-carbonyl-protected thiosialoside **2** (right) which are compared in this thesis in terms of their potential as efficient donor for the synthesis of α-linked sialosides.

The corresponding donor and reaction conditions will be used to synthesize sufficient quantities of α-linked sialic acid derivatives carrying azidoethylene glycols of different lengths as linker for later protein conjugation reactions. Ethylene glycols were chosen because of the cheap starting materials, readily available varying lengths and thereby varying linker flexibility, and their straightforward modification possibilities. It is also known that ethylene glycols show beneficial properties in drug delivery and biomedical applications, such as good solubility, low toxicity and high stability.[141]

In the second part, these azidoethylene glycol linked sialic acids will be globally deprotected and used to modify alkyne-functionalized Qβ bacteriophage capsids[59a] by means of already published CuAAC conditions.[83] This will generate defined sialic acid-decorated Qβ particles which carry a known number of residues at known positions and which will then be used as HA binders and eventually influenza inhibitors. With the known structural features of the system, it will be possible to deduce a model of the underlying binding modes, to study the influence of linker flexibility, and to get further insights into multivalent virus interactions.



Scheme 10: Anticipated bioconjugation reaction between azidoethylene glycol linked sialic acids and alkyne-functionalized Qβ bacteriophage capsids by means of CuAAC.

In order to use the synthesized sialic acid-Qβ capsids as a diagnostic tool, it would be necessary to additionally attach a reporter molecule, e.g. fluorophore, to the particles. This additional modification should not interfere with the sialic acid-decorated surface of the capsid. Therefore, the interior protein surface of Qβ will be used for an additional conjugation, which firstly requires the removal of encapsidated expression host RNA. Afterwards, an accessible tyrosine residue will be probed to be addressed by a diazonium coupling reaction with 4-nitrobenzenediazonium tetrafluoroborate

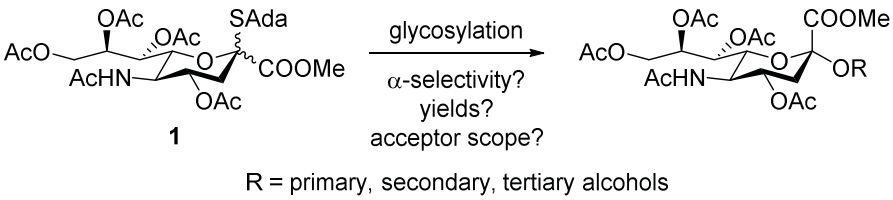
followed by a reductive cleavage with sodium dithionite to generate a reactive handle for a further conjugation reaction.

In the last part of this theses, a quantitative live cell fluorescence microscopy assay should be set up in order to analyze the asialoglycoprotein receptor-mediated uptake of engineered GalNAc-GFP conjugates into hepatic HepG2 cells. For that purpose, positive and negative controls have to be established in order to validate the performance of anticipated fluorescence microscopy assay. Therefore, commercially available fetuin and asialofetuin will be used and labeled by *N*-hydroxysuccinimide-activated fluorescein and subsequently applied to HepG2 cells. After incubation, the cells should be analyzed by fluorescence microscopy and the uptake of fluorescently labeled protein will be quantified by image analysis.

7. Results and Discussion

7.1 Optimizing the Synthetic Route Towards α-linked Sialic Acid Conjugates

The first part of this thesis focuses on optimizing the synthetic route towards α-linked sialic acid glycosides. The α-linkage is the glycosidic configuration which is found in biology and which is necessary for recognition by naturally occurring receptors, such as the HA protein of influenza viruses. The current state-of-the-art synthetic strategy involves a lengthy route with many protecting group manipulations and an early anomer separation for the generation of 5-*N*,4-*O*-oxazolidinone protected sialic acid donors, which significantly hampers the obtainable yields (compare section 5.7.3). Based on recent reports,[140, 142] it was anticipated that the bicyclic oxazolidinone protection is not a prerequisite to obtain high α-selectivities. Therefore, more accessible sialic acid thioadamantanyl glycosides should be tested as donors in glycosylation reactions. To study the acceptor scope which can be glycosylated with these donors, a range of acceptors with different steric demands was required and the corresponding syntheses are described in the following section.



Scheme 11: Planned glycosylation reactions with synthetically easier accessible sialic acid donor **1**.

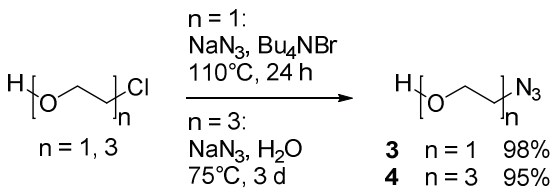
7.1.1 Acceptor Synthesis

7.1.1.1 Azidoethylene Glycols

The goal of this thesis was to modify Qβ phage capsids with sialic acid residues in a defined manner which is fostered by the symmetrical and rigid assembly of the capsid. Since the conjugation reactions will occur at fixed positions on the capsid surface, azidoethylene glycol linkers of varying length (1-5, and 8 ethylene glycol units) were

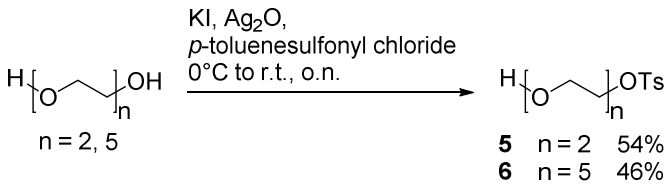
chosen as a first set of glycosyl acceptors, based on preliminary studies by Lukas Artner (FU Berlin).[143] This allows for the screening of distances which can be covered by the respective linker to eventually identify an optimal binding scenario between sialic acid-modified Qβ capsids and HA or whole viruses, respectively.

Attempts were made to synthesize azidoethylene glycols with one to five and eight ethylene glycol units for this purpose. In this process it was possible to obtain azidomonoethylene **3**[144] and azidotriethylene glycol **4**[145] directly after one step, starting from the corresponding commercially available chloroethylene glycols and sodium azide, with 98% and 95% yield, respectively (Scheme 12).



Scheme 12: One-step chloride-azide exchange for the synthesis of azidomonoethylene glycol **3** and azidotriethylene glycol **4**.

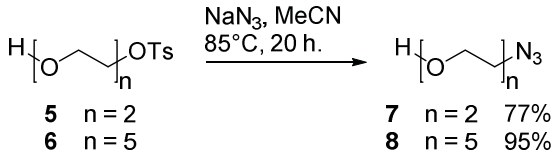
For azidodiethylene glycol **7** and azidopentaethylene glycol **8** the synthetic route started with the corresponding commercially available diols. These were first converted to the corresponding tosylates **5** and **6** with potassium iodide, silver(I) oxide and *p*-toluenesulfonyl chloride in dichloromethane which delivered the desired products in 54% and 46%, respectively (Scheme 13).[146]



Scheme 13: Monotosylation reaction of di- and pentaethylene glycol to install a leaving group for the subsequent azide-exchange.

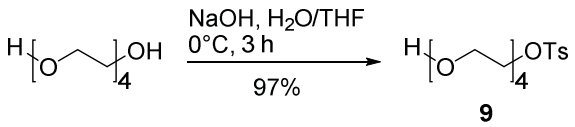
The monotosyl derivatives **5** and **6** were then treated with sodium azide in acetonitrile in order to promote the desired tosyl-azide exchange.[147] With that, 77%

and 95%, respectively, of the azidodiethylene glycol **7** and azidopentaethylene glycol **8** were obtained (Scheme 14).



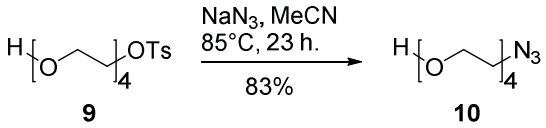
Scheme 14: Tosyl-azide exchange of **5** and **6** in order to obtain desired azidodiethylene glycol **7** and azidopentaethylene glycol **8**.

The remaining tetra- and octaethylene glycol derivatives **10** and **14** were both accessible through the synthetic route towards azidooctaethylene glycol **14**, since it consists of two tetraethylene glycol building blocks. Therefore, commercially available tetraethylene glycol was monotosylated in the first step with sodium hydroxide and *p*-toluenesulfonyl chloride in a water-tetrahydrofuran mixture.[148] With these reaction conditions it was possible to obtain the monotosylated tetraethylene glycol **9** in 97% yield (Scheme 15). This yield is significantly higher compared to the previously applied tosylation conditions for the synthesis of **5** and **6** (*vide supra*) utilizing potassium iodide and silver(I) oxide. The direct nucleophilic reaction pathway promoted by a base therefore proved to be more efficient at ethylene glycol tosylation.



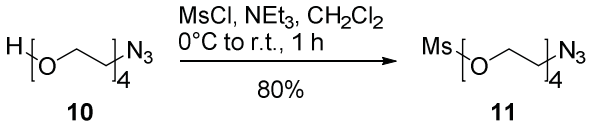
Scheme 15: Monotosylation reaction of tetraethylene glycol on the route towards azidotetra- **10** and azidooctaethylene glycol **14**.

The final step for the synthesis of azidotetraethylene glycole **10** is again the tosyl-azide exchange which was performed as describe above for **7** and **8** with sodium azide in acetonitrile giving **10** in 83% yield (Scheme 16).[147]



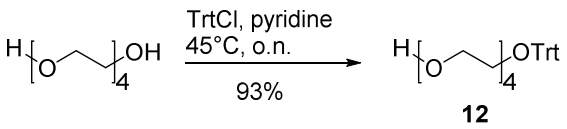
Scheme 16: Final tosyl-azide exchange for the synthesis of azidotetraethylene glycol **10**.

As mentioned above the azidooctaethylene glycol **14** consists of two tetraethylene glycol building blocks. Sandra Behren (HU Berlin) worked on the synthesis of **14** in her master thesis.[142] In her work, the two building blocks were coupled under Williamson ether synthesis conditions, and therefore one fragment had to contain the desired azido group in addition to a mesitylated hydroxyl group as leaving group on the other end of the molecule. This was readily achieved with the above synthesized azidotetraethylene glycol **10** which was mesitylated with triethylamine and methanesulfonyl chloride in dichloromethane. The mesitylated azidotetraethylene glycol **11** was obtained in 80% yield under these conditions (Scheme 17).[149]



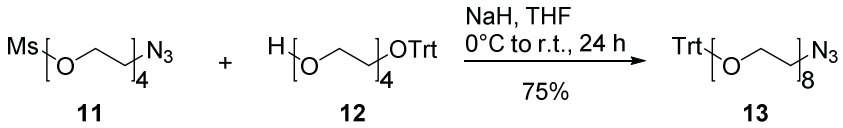
Scheme 17: Mesitylation of **10** for the synthesis of the first building block necessary for the Williamson ether synthesis of azidooctaethylene glycol **14**.

The second tetraethylene glycol building block consequently had to carry a free hydroxyl functionality and a suitable protected alcohol on the other end. The protecting group of choice in this case was a triphenylmethyl (trityl) group, since its incorporation and deprotection are readily achieved and the trityl moiety is compatible with the strongly basic reaction conditions of the Williamson ether synthesis. Therefore, tetraethylene glycol was monotrityl protected using a stoichiometric amount of trityl chloride and pyridine as base, resulting in the monotritylated tetraethylene glycol **12** in 93% yield (Scheme 18).[150]



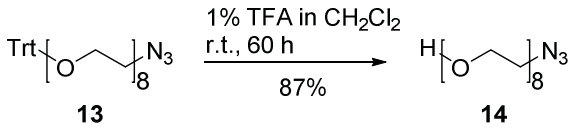
Scheme 18: Monotritylation reaction of tetraethylene glycol to obtain **12** as second tetraethylene building block for the Williamson ether synthesis.

With both building blocks **11** and **12** in hand, Williamson ether synthesis of trityl-protected azidooctaethylene glycol **13** was performed in dry tetrahydrofuran with sodium hydride as the corresponding base (Scheme 19). The desired product was obtained in 75% yield.[142]



Scheme 19: Williamson ether synthesis of trityl-protected azidooctaethylene glycol **13** using the building blocks **11** and **12**.

The final step on this route was the trityl-deprotection which is easily achieved under slightly acidic conditions. Therefore, **13** was treated with 1% trifluoroacetic acid in dichloromethane and the desired azidooctaethylene glycol linker **14** was obtained in 87% yield (Scheme 20).[142]



Scheme 20: Final trityl-deprotection step towards azidooctaethylene glycol **14** under slighty acidic conditions.

7.1.1.2 Additional Acceptors for Glycosylation Reactions

In order to screen sterically varied and more demanding acceptors in the glycosylation studies, other acceptor compounds were necessary. These should cover the simple secondary and tertiary alcohols cyclohexanol and adamantanol, respectively, as well as actual protected sugar acceptors (Figure 15). Cyclohexanol, adamantanol and the methyl-2,3,4-tri-*O*-benzyl-α-D-glucopyranoside are commercially available and didn't

have to be synthesized, whereas the two differently protected galactose derivatives **15** and **19** had to be synthesized.

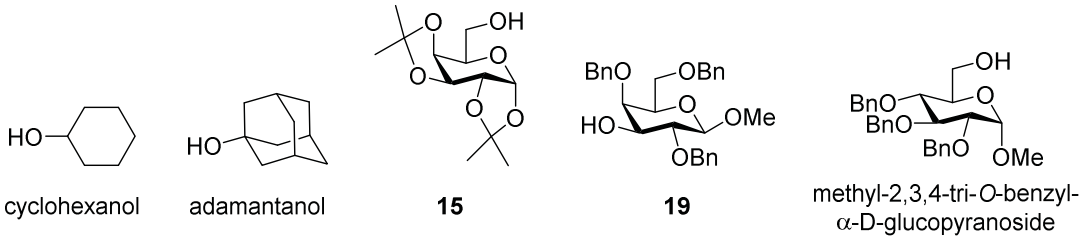
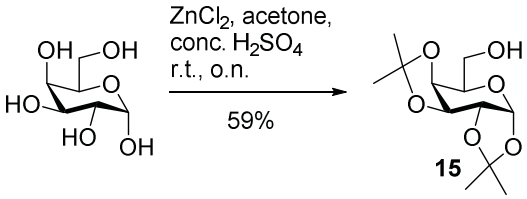


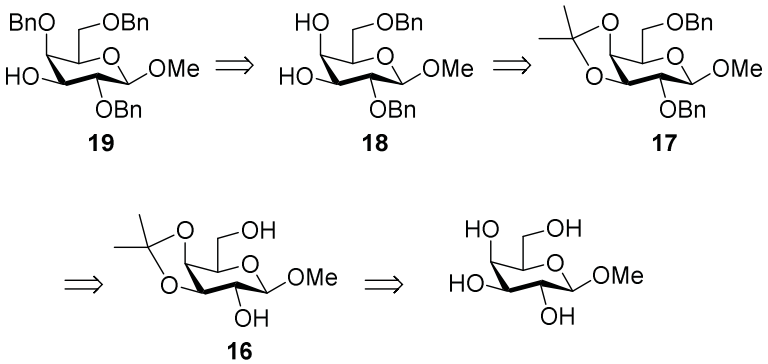
Figure 15: Sterically more demanding glycosylation acceptors which should be tested in the glycosylation reactions.

Isopropylidene protected galactose with a free 6-OH group **15** is accessible through a one-step synthesis, starting from unprotected galactose with zinc(I) chloride in acetone and concentrated sulfuric acid as a catalyst (Scheme 21).[151] With this protocol **15** was obtained in 59% yield.



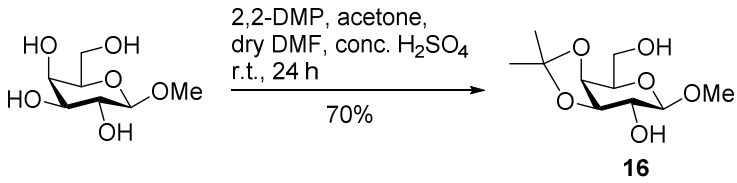
Scheme 21: One-step synthesis of isopropylidene-protected galactose derivative **15** with a free 6-OH group.

A multi-step synthesis was necessary for the benzyl-protected galactose derivative with a free 3-OH group **19** according to retrosynthetic Scheme 22.



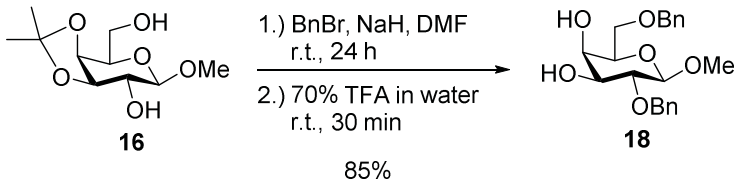
Scheme 22: Retrosynthetic scheme for the synthesis of benzyl-protected galactose derivative with a free 3-OH group **19**.

In the first step, commercially available methyl-β-D-galactopyranoside was isopropylidene-protected at the 3- and 4-OH functionalities in 2,2-dimethoxypropane, acetone and dry *N,N*-dimethylformamide with concentrated sulfuric acid as a catalyst (Scheme 23).[152] The respective product **16** was obtained in 70% yield.



Scheme 23: Isopropylidene protection of the 3- and 4-OH group of methyl-β-D-galactopyranoside as first step in the synthesis of benzyl-protected galactose derivative with a free 3-OH group **19**.

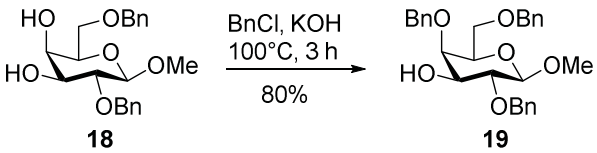
In the next step the remaining 2- and 6-OH groups were benzylated with benzyl bromide and sodium hydride in *N,N*-dimethylformamide and the obtained product **17** after column chromatography was directly subjected to the isopropylidene deprotection step under acidic conditions with 70% trifluoroacetic acid in water.[153] The methyl-2,6-di-*O*-benzyl-β-D-galactopyranoside **18** was obtained in 85% after this two-step procedure (Scheme 24).



Scheme 24: Two-step procedure for the synthesis of **18**: 1.) benzylation of 2- and 6-OH group, directly followed by 2.) the isopropylidene deprotection under acid conditions.

The final benzylation step had to be selective for the 4-OH group. In a study by *Morishima et al.*[154] it was found that the partial benzylation of galactopyranosides is dependent on the base which is used to deprotonate the corresponding hydroxyl groups. They observed in this study, that the use of 1.3 equivalents of potassium hydroxide as a base in benzyl chloride delivers predominantly the benzylated 4-OH group when **18** is used as starting material, but with a rather underwhelming yield (41%). In a first attempt to reproduce this data, the reported reaction conditions

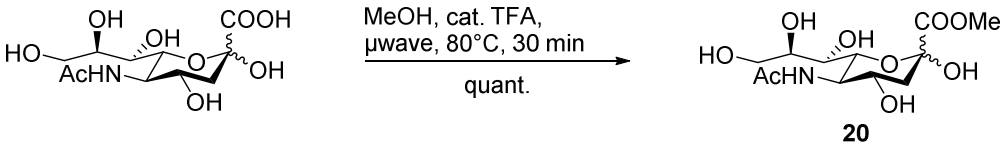
were applied (Scheme 25) and the reaction went well with very good isolated yield of 80% of the desired product **19**.



Scheme 25: Selective benzylation of the 4-OH group of **18** to obtain desired methyl-2,4,6-tri-O-benzyl-β-D-galactopyranoside **19**.

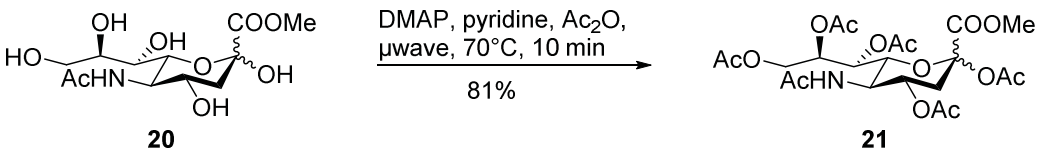
7.1.2 Donor Synthesis

In order to compare more accessible donors such as **1** with the state-of-the-art oxazolidinone-protected donor **2** (Figure 14), it was necessary to synthesize **2** and to perform the anticipated glycosylation reactions with it. The synthetic route towards **2** started with the esterification of the acid-functionality of commercially available Neu5Ac with catalytic amounts of trifluoroacetic acid in methanol and under microwave irradiation (Scheme 26).[155] The use of a microwave as a heating source sped up the reaction tremendously (30 min vs. several hours with classical heating) and delivered the Neu5Ac methyl ester **20** in quantitative yield.



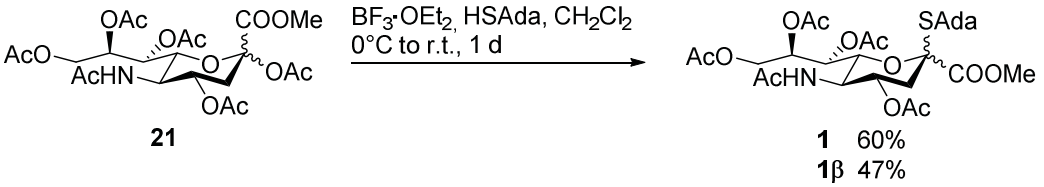
Scheme 26: Microwave-assisted esterification of Neu5Ac with methanol.

The subsequent peracetylation was also performed under microwave irradiation. For that, the methyl ester **20** was mixed with 4-*N*,*N*-dimethylaminopyridine (DMAP) in pyridine and acetic anhydride and after again only 10 minutes the corresponding peracetylated methyl ester **21** was obtained in 81% yield (Scheme 27).[155]



Scheme 27: Microwave-assisted peracetylation of methyl ester **20** to obtain peracetylated methyl ester **21**.

With the following introduction of a thioadamantanyl moiety at the anomeric carbon of **21** two of the anticipated donors were already prepared. The reaction is promoted with boron trifluoride etherate as the activating Lewis acid and performed with 1-adamantanthiol as coupling partner in dry dichloromethane (Scheme 28).[138]



Scheme 28: Introduction of the thioadamantanyl moiety into the peracetylated methyl ester **21** in order to obtain glycosylation donors **1** and **1β**.

The resulting anomeric mixture of the product can be isolated as a mixture to give donor **1** in 60% yield with an α/β ratio of 24:76 according to integration of the 3-H_{eq} proton signals in the NMR spectrum of **1** (Figure 16). For the route towards oxazolidinone protected donor **2** it was necessary to separate the anomers and to isolate the respective β-anomer exclusively. Therefore, a second batch of the reaction was performed and the resulting anomeric mixture was separated by flash column chromatography. This resulted in a slight loss of product material, with only 47% of **1β** being isolated. Unfortunately, it was inexplicably not possible throughout this work to obtain isolated yields in the range of published values (81% of **1β**). The prepared **1β** was used as glycosylation donor in addition to being used for the synthesis of donor **2**.

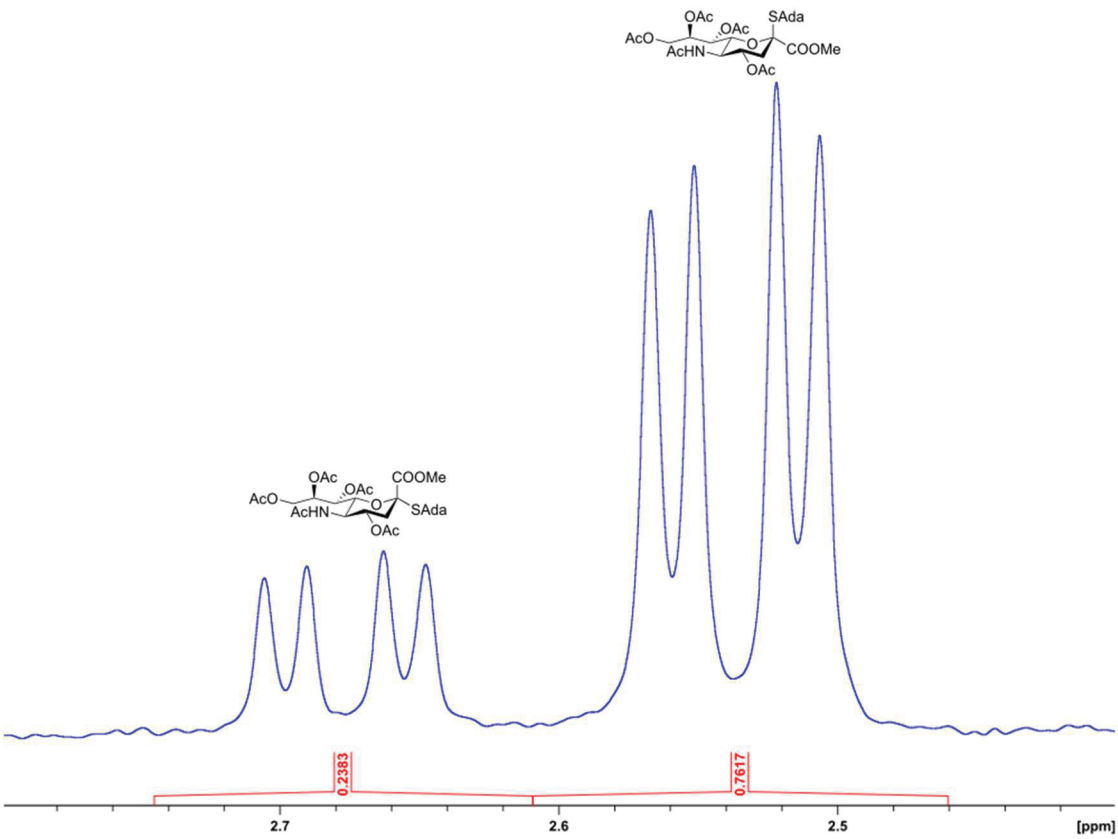
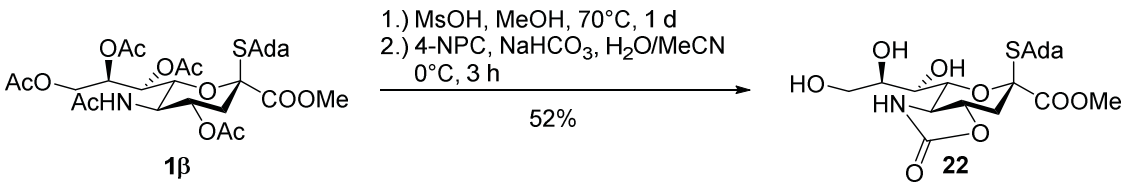


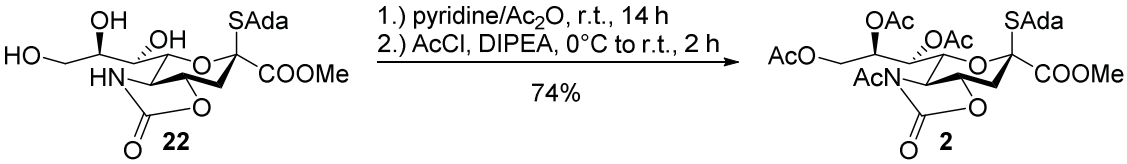
Figure 16: Excerpt of the ^{1}H-NMR spectrum of donor **1**. Shown is the region between 2.4 and 2.8 ppm where the signals for the 3-H_{eq} protons of both anomers in the mixture are located.

To this end, β-adamantanyl thiosialoside **1β** was globally deacetylated with methanesulfonic acid in methanol and after a brief neutralization and washing procedure the oxazolidinone moiety was introduced by the use of 4-nitrophenylchloroformate (4-NPC) and sodium bicarbonate as base in in a water/acetonitrile-mixture (Scheme 29).[138] The obtained yield of 52% is lower compared to published values (63%), which can potentially be attributed to the harsh deprotection conditions with methanesulfonic acid. However, the published protocol involved additional three synthetic steps, which could be avoided here in order to access the desired material faster.



Scheme 29: Introduction of the oxazolidinone protecting group after global deacetylation of thioadamantanyl sialoside **1β**.

In a last step towards donor **2** the hydroxyl groups and the carbamate NH-group had to be reacetylated. This was achieved in a two-step protocol which started with the usual pyridine/acetic anhydride mixture, followed by an additional step in which acetyl chloride was used in conjunction with diisopropylethylamine as the corresponding base (Scheme 30).[138]



Scheme 30: Final two-step reacetylation reaction towards donor **2**.

With that, **2** was obtained in 74% isolated yield and all donor compounds were ready for the following glycosylation studies (Figure 17).

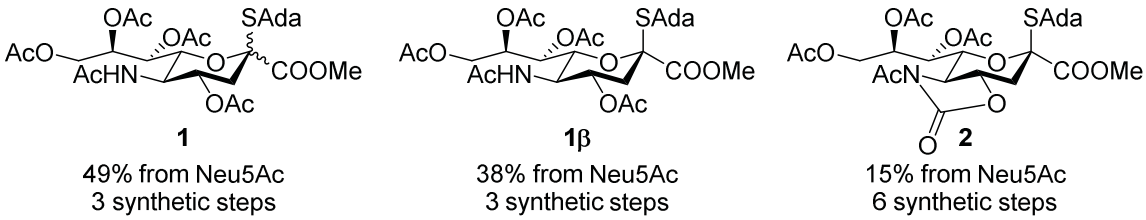


Figure 17: Summary of synthesized sialic acid glycosylation donors: anomeric mixture **1**, pure β-anomer **1β** and oxazolidinone protected **2**.

7.1.3 Glycosylation Reactions

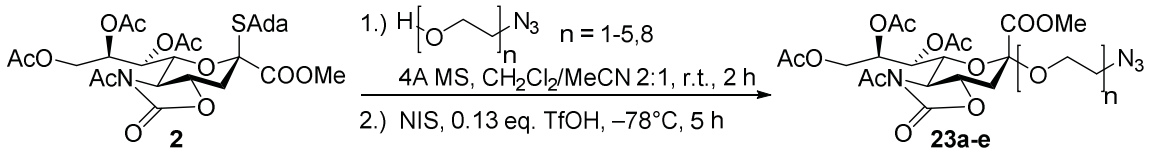
In order to set up a benchmark, the first set of glycosylation reactions was performed with the established oxazolidinone protected donor **2** and the range of azidoethylene glycols **3**, **7**, **4**, **10**, **8**, and **14**. Reaction conditions were applied as reported earlier from the *Crich* group,[140] in which the closely related 5-azido-2-thioadamantanyl sialic

acid derivatives were used as glycosylation donors. These conditions involved the use of acid-washed 4Å molecular sieves, 1.2 equivalents of the corresponding glycosyl acceptor, 2.0 equivalents of *N*-iodosuccinimide with 0.13 equivalents of triflic acid as promoter couple in a dichloromethane/acetonitrile 2:1 mixture (Table 3, top). In order to evaluate the conversion of the corresponding reactions, the reaction crude was analyzed by LC-MS analysis after work-up and conversion was determined by integration and calculated using the following equation:

(‘*product peak area*’ / (‘*starting material peak area*’ + ‘*product peak area*’)) · 100%

Isolation of the products and separation of the anomers was performed by preparative HPLC, in order to have a chromatographic method which can be applied for all anticipated glycosylation reactions, even if some product might not be separable by flash column chromatography. This allowed the direct comparison of the viability of the corresponding donor compounds in glycosylation reactions.

Table 3: Results of glycosylation reaction using donor **2**.

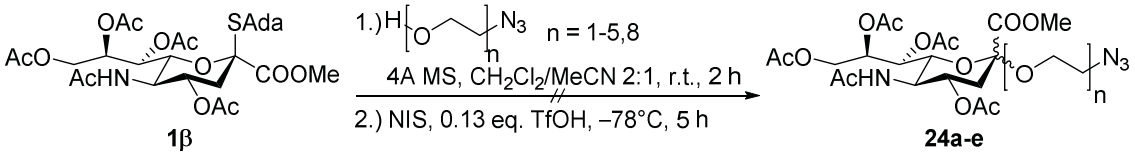


Entry	n =	Conversion	Crude α/β	Isolated Yield α	Isolated Yield β
1	1 (**23a**)	77%	93 : 7	43%	< 5%
2	2 (**23b**)	84%	100 : 0	42%	-
3	3 (**23c**)	77%	100 : 0	48%	-
4	4 (**23d**)	93%	94 : 6	57%	-
5	5 (**23e**)	88%	93 : 7	53%	-
6	8 (**23f**)	93%	97 : 3	58%	-

It can be seen from Table 3 that the reaction proceeded with at least 77% and up to 93% conversion under these conditions. Furthermore, the almost exclusive α-selectivity which was expected from donor **2** was indeed achieved, which showed the reliability of this system when α-linked products are needed. After isolation of the products by HPLC, 42-58% of the desired α-anomers were obtained. It was only

possible for the shortest azidomonoethylene glycol linked product to additionally isolate the corresponding β-anomer **23aβ** which, however, was not completely separated from the corresponding α-anomer. Additionally, it can be seen from this first set of experiments, that the length of the azidoehtylene glycol linker does not influence the outcome of the glycosylation reaction. This might be because the overall difference in steric demand of the flexible ethylene glycol chains did not vary significantly with chain length. Furthermore, it can be expected that the random coils of these rather short oligomers, which are formed in solution, are not large enough to mask the reactive hydroxyl-end of the molecule in it, which would lead to a hampered glycosylation reaction.

The first donor compared with oxazolidinone donor **2** was chosen to be the pure β-anomer **1β**. For that purpose, the exact same reaction conditions stated above were applied together with the same azidoethylene glycols (Scheme 31). In this set of reactions, though, almost no conversion to the desired products was observed by analyzing the reaction crude by LC-MS and the starting donor was still present predominantly.



Scheme 31: Glycosylation reactions with pure β-anomer **1β** as glycosylation donor with initial reactions conditions which did not lead to any conversion of the starting material.

An important observation was that during these reactions the characteristic deep yellow color of a successfully activated reaction mixture was absent. This rendered the hypothesis, that probably the 0.13 equivalents of triflic acid used here were not sufficient to activate glycosyl donor **1β**. To test this, a screen of equivalents of triflic acid was set up and tested with donor **1β** and azidotriethylene glycol **4** (Table 4). All other reactions conditions were kept constant during this set of experiments and the

conversion was again monitored by LC-MS analysis of the crude reactions after work-up.

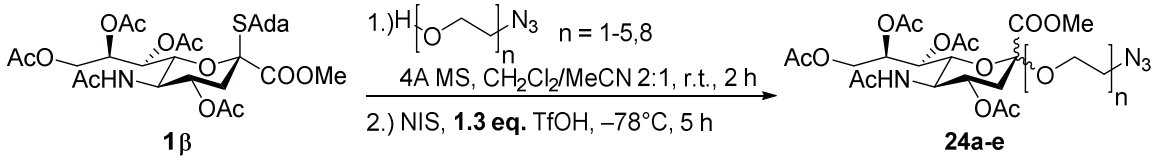
Table 4: Conversion of **1β** dependent on TfOH equivalents.

Entry	TfOH	Conversion
1	1x (0.13 eq.)	17%
2	2x (0.26 eq.)	36%
3	5x (0.65 eq.)	55%
4	10x (1.30 eq.)	97%

From this screen it was clear, that with increasing equivalents of triflic acid the conversion towards the desired products increased and reached basically full conversion with the tenfold amount (1.30 equivalents) of triflic acid compared to the initial conditions (Table 4, entry 4). When these conditions were applied the characteristic yellow color appeared readily as an indicator for successful activation of the reaction.

When the found reaction conditions were transferred to the whole range of azidoethylene glycols, proper activation was achieved in all cases (Table 5).

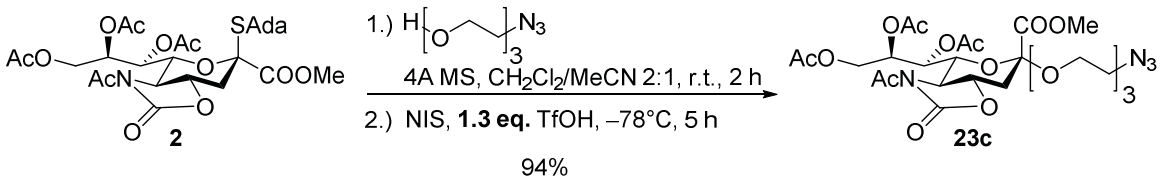
Table 5: Results of glycosylation reaction using donor **1β**.



Entry	n =	Conversion	Crude α/β	Isolated Yield α	Isolated Yield β
1	1 (**24a**)	86%	78 : 22	49%	4%
2	2 (**24b**)	91%	85 : 15	45%	< 11%
3	3 (**24c**)	90%	76 : 24	39%	4%
4	4 (**24d**)	94%	73 : 27	45%	10%
5	5 (**24e**)	97%	74 : 26	46%	5%
6	8 (**24f**)	88%	75 : 25	42%	7%

This set of reactions with donor **1β** and the increased amount of activating triflic acid also achieved good conversion in the range of 86-97%. The proportion of desired α-anomers in the obtained products was expectedly slightly reduced, compared to the reactions with oxazolidinone protected donor **2** (compare Table 3), but with 73-85% still relatively high, indicating that a good amount of α-selectivity is retained in this type of donor. After isolation of the respective anomers by preparative HPLC, the desired α-anomers were obtained in 39-49% yield, which is slightly less compared to oxazolidinone donor **2**, but not far off. Here, it was possible to isolate the corresponding β-anomers in all cases, whereby the azidodiethylene glycol linked product (Entry 2) was contaminated with the corresponding α-anomer.

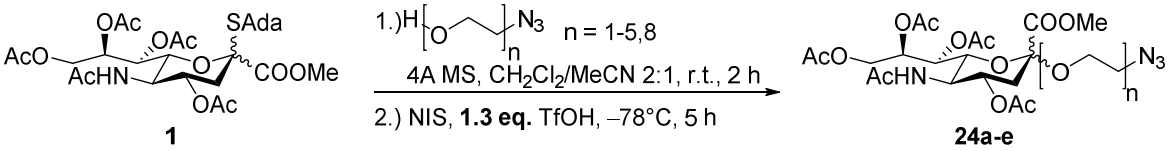
To see if an increased amount of triflic acid would influence the anomeric outcome of the glycosylation reaction with oxazolidinone protected donor **2**, the optimized reaction conditions were applied in a test glycosylation of **2** with azidotriethylene glycol **4** with 1.3 equivalents of triflic acid (Scheme 32). The reaction worked equally well, compared to the initial reaction conditions and the anomeric outcome was not influenced (α/β ratio of 97:3).



Scheme 32: Test glycosylation of oxazolidinone protected donor **2** with azidotriethylene glycol **4** and the increased amount of triflic acid as activator.

These optimized reaction conditions were finally applied for the glycosylation reactions of anomeric mixture **1** as donor (Table 6). In this set of experiments, a slight increase in conversion to the desired products (92-98%) was observed compared to donor **1β**. With portions of the desired α-anomer in the range of 74-80% and preparative HPLC isolated yields thereof in the range of 43-50%, this donor gave very comparable results to the pure β-anomer **1β** donor.

Table 6: Results of glycosylation reaction using donor **1**.



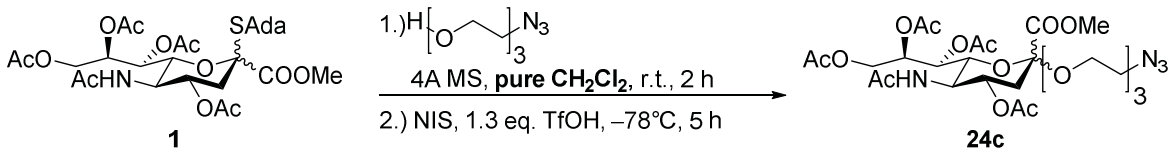
Entry	n =	Conversion	Crude α/β	Isolated Yield α	Isolated Yield β
1	1 (**24a**)	92%	79 : 21	50%	2%
2	2 (**24b**)	92%	80 : 20	48%	< 8%
3	3 (**24c**)	97%	80 : 20	50%	10%
4	4 (**24d**)	97%	74 : 26	44%	11%
5	5 (**24e**)	98%	76 : 24	48%	12%
6	8 (**24f**)	94%	80 : 20	43%	7%

Comparing the results from the three tested donor compounds **1**, **1β** and **2**, it can be stated that the glycosylation reactions towards ethylene glycol-linked sialic acids worked in all the cases. For non-oxazolidinone protected donors **1β** and **1** it was necessary to increase the equivalents of activating acid. In Table 7 the overall isolated yields of the desired α-linked sialic acid-azidoethylene glycols are summarized and compared. It can clearly be seen, that the early anomer separation and additional protecting group manipulation for the synthesis of oxazolidinone protected donor **2**, resulted in a significant loss of obtained material after glycosylation reactions (Entry 1 vs. Entry 3). When the anomerically pure **1β** was used in glycosylation reactions with azidoethylene glycols, overall yields were in between of those from donor **2** and **1** (Entry 2).

Table 7: Summarized results of different sialic acid donors applied in glycosylation reactions.

Entry	Donor	Range of Overall Isolated Yield of Desired α-Products, Starting from Neu5Ac 3
1	**2**	6-8%
2	**1β**	15-19%
3	**1**	21-24%

At this point, it became apparent that anomeric mixture **1** can be used as a glycosylation donor and to further deepen the knowledge about this system, more test reactions were carried out. In section 5.7.3, Scheme 8 the participating effect of nitrile solvents is depicted in the context of sialic acid glycosylation reactions. In order to see if this holds true for the glycosylation reaction with donor **1**, a test reaction in pure dichloromethane, but otherwise unchanged reaction conditions, was carried out (Scheme 33).



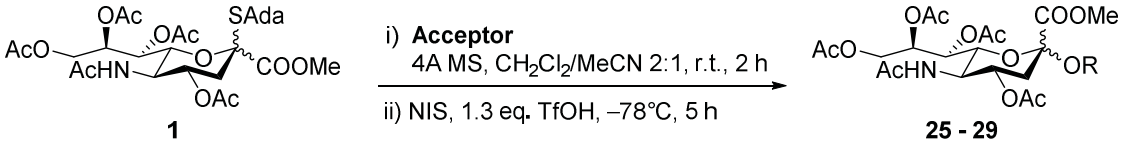
Scheme 33: Test glycosylation with donor **1** performed in pure dichloromethane as solvent.

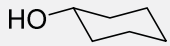
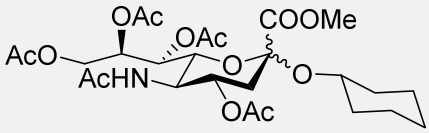
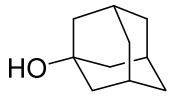
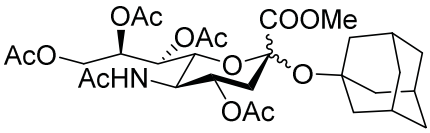
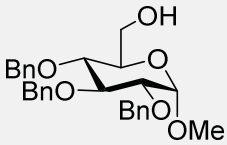
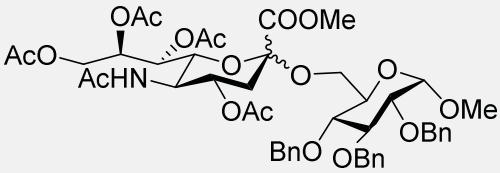
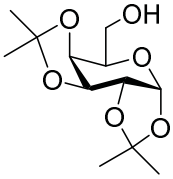
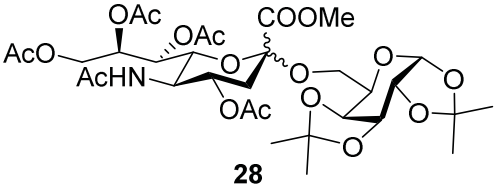
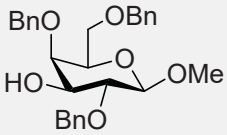
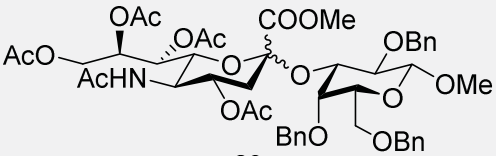
Analysis of the reaction crude revealed that a significantly reduced portion of the desired α-anomer was present (α/β ratio of 62:38) when the reaction was performed in pure dichloromethane compared to a mixture which contained 1/3 of acetonitrile (α/β ratio of 80:20, Table 6, Entry 3). It has to be noted, that higher proportions of acetonitrile (50% and more) are not applicable, since the reaction mixtures freeze due to the higher melting point of acetonitrile.

In order to demonstrate the synthetic versatility of donor **1**, it was also necessary to test different glycosyl acceptors, which were synthesized before (see section 7.1.1.2, Figure 15). From Table 8 it is obvious that many varieties of different glycosyl acceptors can be used to glycosylate sialic acid donor **1** with comparable yields, when the optimized reaction conditions were applied. Again, the corresponding products were isolated by preparative HPLC to obtain comparable results for all acceptors. With that, it was possible to isolate the simple cyclohexyl **25** and adamantanyl glycosylated **26** products in 35% (Entry 1) and 39% (Entry 2) isolated yield of the desired α-anomer, respectively. The use of sterically demanding, protected carbohydrates is of special interest with regard to oligosaccharide synthesis. To this

end, a benzyl protected glucose derivative and the isopropylidene protected derivative of galactose **15** – two commonly employed protecting groups in carbohydrate chemistry – was tested with which 38% (Entry 3) and 40% (Entry 4), respectively, of the corresponding desired α-anomers were obtained.

Table 8: Substrate scope of glycosylation reactions with anomeric mixture **1** as donor.



Entry	Acceptor	Product	Isolated Yield α (β)
1		 **25**	35% (13%)
2		 **26**	39% (7.5%)
3		 **27**	38% (6%) flash column: 58% (4%)
4	 **15**	 **28**	40% (-)
5	 **19**	 **29**	-

In one case, when methyl-2,4,6-tri-*O*-benzyl-β-D-galactopyranoside **19** was used as an acceptor (Entry 5), it was impossible to isolate a pure product even by preparative HPLC. However, the problem in this case was actually that the acceptor **19** is not

compatible with the glycosylation reaction conditions. This hypothesis is supported by the fact that an additional HPLC fraction was isolated, which presumably should contain the starting acceptor **19**. NMR analysis of the isolated material, however, revealed that the respective fraction was in fact methyl-2,**3**,6-tri-*O*-benzyl-β-D-galactopyranoside **30** (Figure 18).

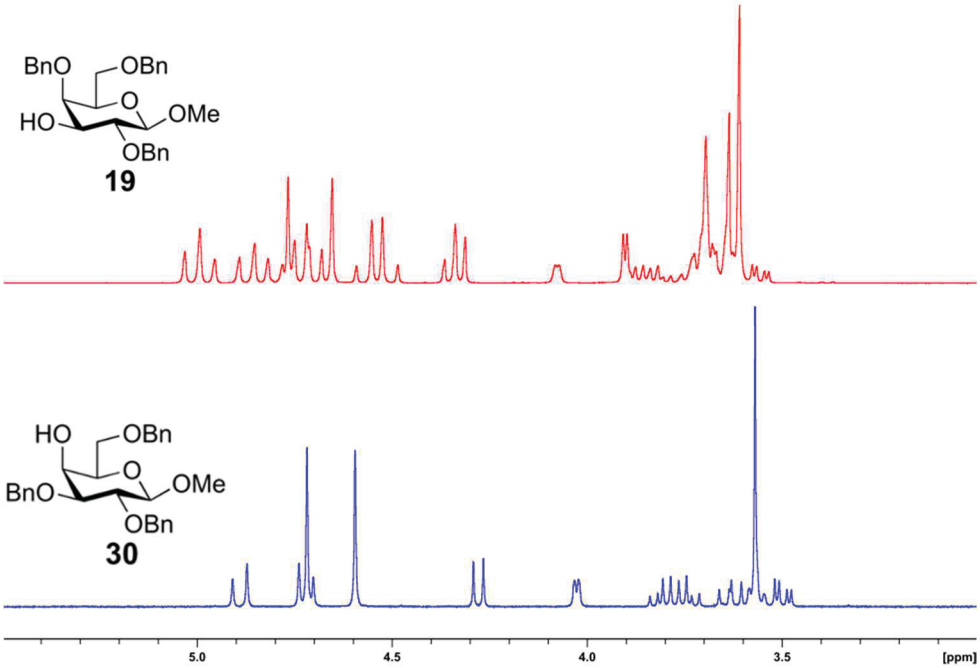


Figure 18: Comparison between the ^{1}H-NMR spectra (3.0-5.5 ppm) of glycosyl acceptor **19** which was used for the reaction (top, red) and isolated compound **30** after the reaction (bottom, blue).

This observation points towards a benzyl migration reaction, which is certainly conceivable under very acidic conditions (Figure 19). Consequently, the use of acid sensitive acceptors has to be carefully evaluated and is in some circumstances not possible, representing a limitation of the method described herein.

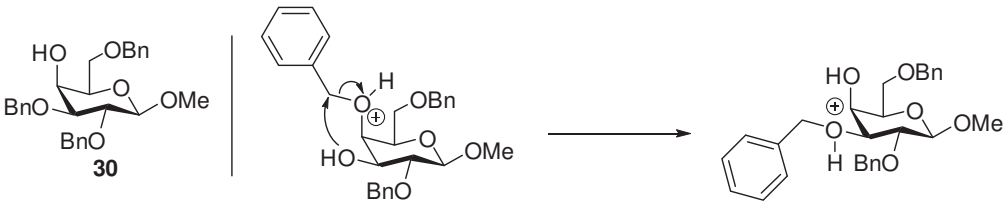


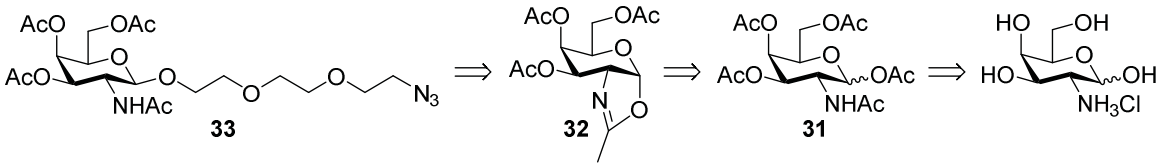
Figure 19: Isolated side product **30** and proposed mechanism of the occurred side reaction.

The relatively low isolated yields reported so far in this chapter can probably be attributed to the HPLC purification, which was applied for the sake of comparability of all performed glycosylation reactions. In an actual synthetic route for a glycosylated sialic acid derivative it would be beneficial to isolate the corresponding products and desired anomers by flash column chromatography. To test if this is possible when the anomeric mixture **1** is used as a glycosyl donor and anomeric mixtures of the corresponding products are generated, two test glycosylations were performed under the optimized reaction conditions described above. In the first experiment, azidotriethylene glycol **4** was used as an acceptor and flash column chromatography on silica gel with a gradient of ethyl acetate in hexanes was tested. Indeed, it was possible to isolate the corresponding product, but separation of the anomers was not achieved. In the second experiment, methyl-2,3,4-tri-*O*-benzyl-α-D-glucopyranoside was used as an acceptor and as a model for a disaccharide synthesis. In this case, the isolation as well as separation of the anomers was readily achieved and with 58% isolated yield of the desired α-anomer **27α** significantly more material was obtained compared to the HPLC purified case (38%, Table 8 , Entry 3). These two experiments show that, indeed, purification and isolation of the desired α-linked glycosylation products is possible in certain cases. However, if HPLC equipment is available in a laboratory, the use of donor **1** in glycosylation reactions can even be advantageous. The necessary HPLC purification of the anomeric product mixture reduces the amount of purifications throughout the synthetic route to one flash column chromatography and one HPLC. When using oxazolidinone protected donor **2** overall four flash columns are necessary to obtain the desired α-linked products. This renders the herein described method as a high yielding and fast synthetic procedure for the synthesis of α-linked sialic acid glycosides if HPLC equipment is available and the desired products have to be accessed quickly. Whenever the resulting anomeric product mixtures are not separable by common flash chromatography and an HPLC

is not available, the oxazolidinone protected donor **2** still represent a state-of-the-art method to synthesize almost exclusively the desired α-linked products.

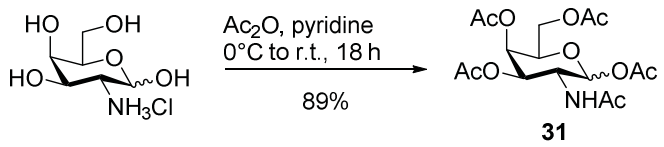
7.2 Synthesis of an *N*-Acetylgalactosamine Derivative as Control

At a later stage of this study, the specific binding of sialylated Qβ bacteriophage capsids to influenza viruses is investigated. In order to exclude any false positive result, an *N*-acetylgalactosamine (GalNAc) derivative was theorized to serve as a control sugar moiety, since GalNAc decorated Qβ capsids should not show any interaction with an influenza virus in the planned binding and infection inhibition assays (see section 7.7). To this end, GalNAc modified with an azidotriethylene glycol linker **34** was chosen. The protected precursor thereof, **33**, is readily accessible from the GalNAx oxazoline **32** which serves as a donor that directs the glycosylation reaction towards the β-linked product. Oxazoline **32** can be obtained by Lewis acid activation of the peracetylated GalNAc precursor **31** (Scheme 34) which is synthesized from the commercially available galactosamine hydrochloride.



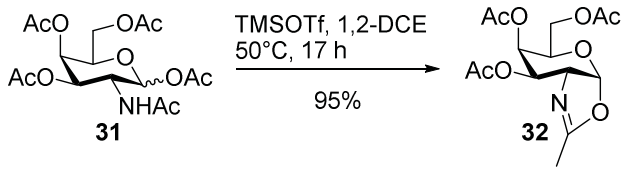
Scheme 34: Retrosynthetic scheme for the synthesis of GalNAc modified with an azidotriethylene glycol linker **34**. Galactosamine hydrochloride is first peracetylated to obtain **31**. Lewis acid activation gives the oxazoline **32** which serves as a donor for the glycosylation with azidotriethylene glycol as acceptor to form **33**.

Starting from commercially available galactosamine hydrochloride, the first step was peracetylation with pyridine and acetic anhydride.[156] The peracetylated GalNAc derivative **31** was obtained from this reaction in 89% after crystallization from methanol (Scheme 35).



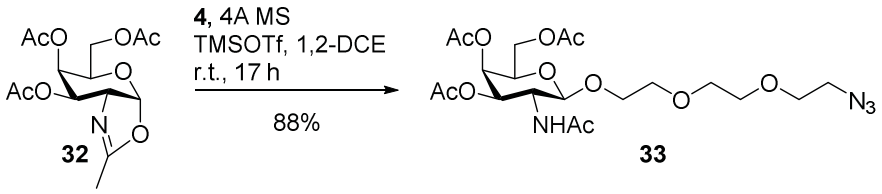
Scheme 35: Synthesis of peracetylated GalNAc **31** starting from commercially available galactosamine hydrochloride with pyridine and acetic anhydride.

The next step is the formation of GalNAc oxazoline **32** as the actual donor compound. Therefore, peracetylated GalNAc **31** was activated by the Lewis acid TMSOTf in 1,2-dichloroethane (1,2-DCE).[157] The desired oxazoline **32** was obtained after column chromatography in 95% yield (Scheme 36).



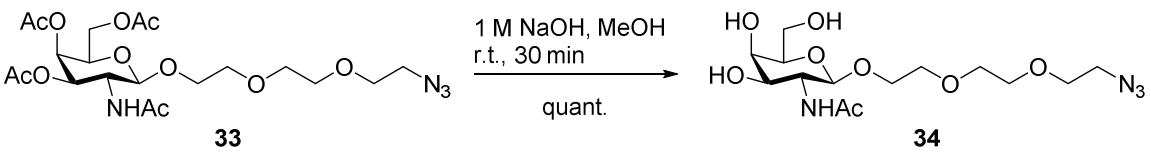
Scheme 36: Synthesis of GalNAc oxazoline **32** starting from peracetylated GalNAc **31** with TMSOTf as activating Lewis acid.

Subsequently, oxazoline **32** was subjected to glycosylation conditions with azidotriethylene glycol **4** as acceptor and TMSOTf as the activating agent in 1,2-DCE.[158] After column chromatography, only the glycosylated β-linked product **33** was obtained in 88% yield (Scheme 37).



Scheme 37: Glycosylation reaction of oxazoline **32** with azidoethylene glycol **4** as acceptor and TMSOTf as activating agent in 1,2-DCE.

In the final deprotection step, **33** was stirred in the presence of aqueous sodium hydroxide solution in methanol to obtain the globally deprotected GalNAc derivative **34** in quantitative yield (Scheme 38).

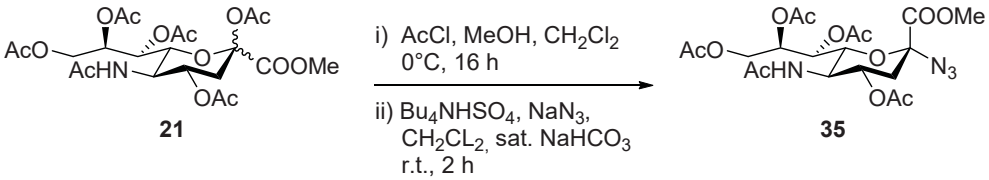


Scheme 38: Global deprotection of glycosylated GalNAc derivative **33** with aqueous sodium hydroxide solution in methanol.

The obtained azidotriethylene glycol GalNAc derivative **34** was then used in the next step for the bioconjugation reaction to the Qβ capsid (see section 7.6)

7.3 Synthesis of 2-Azido Sialic Acid

In order to attach the sialic acid residue directly to the Qβ bacteriophage without any linker it was necessary to incorporate an azide at the anomeric center of sialic acid. For that purpose, peracetylated Neu5Ac methyl ester **21** was transformed to the corresponding 2-chloro derivative by *in situ* generated HCl. This activated intermediate was directly used for the synthesis of 2-azido sialic acid **35** which was achieved by the action of tetrabutylammonium hydrogensulfate and sodium azide (Scheme 39).[159]

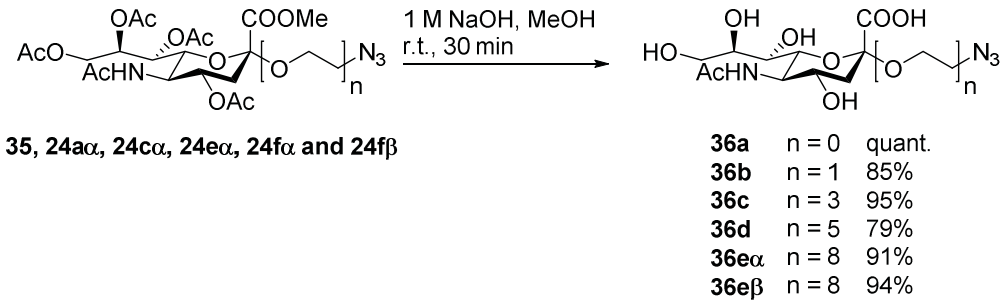


Scheme 39: Synthesis of 2-azido sialic acid **35** through an intermediately formed 2-chloro derivative.

7.4 Deprotection of Glycosylated Sialic Acid Derivatives

In order to reduce the sample size in the planned biophysical and biological assays, only a subset of the previously synthesized azidoethylene glycol sialic acids was chosen to be further processed. Since still the largest possible distance range should be covered, 2-azidosialic acid **35**, the mono- (**24aα**), tri- (**24cα**), penta- (**24eα**) and octaethylene glycol derivatives (**24fα** and **24fβ**) were picked and subjected to alkaline deprotection conditions in aqueous sodium hydroxide solution in methanol,

analogously to the deprotection of GalNAc derivative **33**. The corresponding deprotected azidoethylene glycol sialic acids **36** were obtained after lyophilization in 79% to quantitative yields (Scheme 40) and were ready for the subsequent conjugation reactions (see section 7.6).

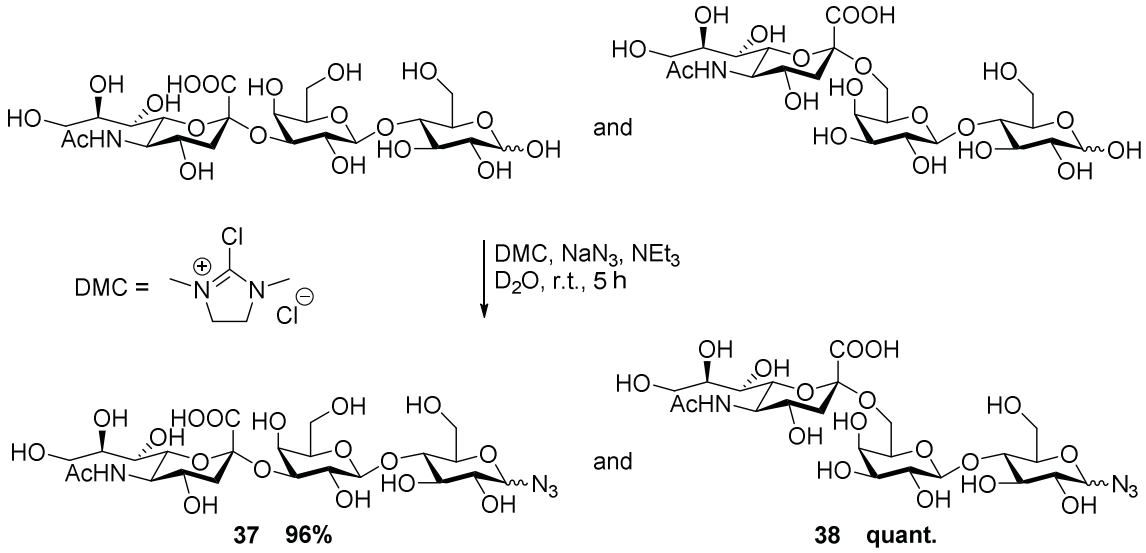


Scheme 40: Global deprotection of glycosylated sialic acid derivatives **35**, **24aα**, **24cα**, **24eα**, **24fα** and **24fβ** with aqueous sodium hydroxide solution in methanol.

7.5 Synthesis of Azidosialyllactose Derivatives

A very important feature of influenza viruses is their selectivity for avian and human hosts (see section 5.2.1.2.1). Since this selectivity is mediated by the respective glycan structures (α2,3- vs. α2,6-linked sialic acid to the penultimate sugar) on the hosts cell surfaces, it would be beneficial to also have these structures conjugated to the Qβ bacteriophage capsids. Therefore, it is necessary to introduce an azide in a di- or oligosaccharide which exhibits above mentioned α2,3- and α2,6-linkages. For that purpose, sialyllactose (SL) was chosen. SL is a trisaccharide which consists of a sialic acid, galactose and glucose moiety and of which both regioisomers (3' and 6') are commercially available in reasonable quantities. In order to avoid cumbersome protecting group manipulations on this complex molecule, a literature known one-step procedure for the introduction of an azide group at the reducing end of the carbohydrate was chosen.[160] According to this protocol, both regioisomers were treated with 2-chloro-1,3-dimethylimidazolinium chloride (DMC), sodium azide and triethylamine in deuterated water (Scheme 41). After a transient acetyl-protection

and reversed-phase purification, the desired azidosialyllactose derivatives **37** and **38** were obtained in 96% and quantitative yield, respectively.



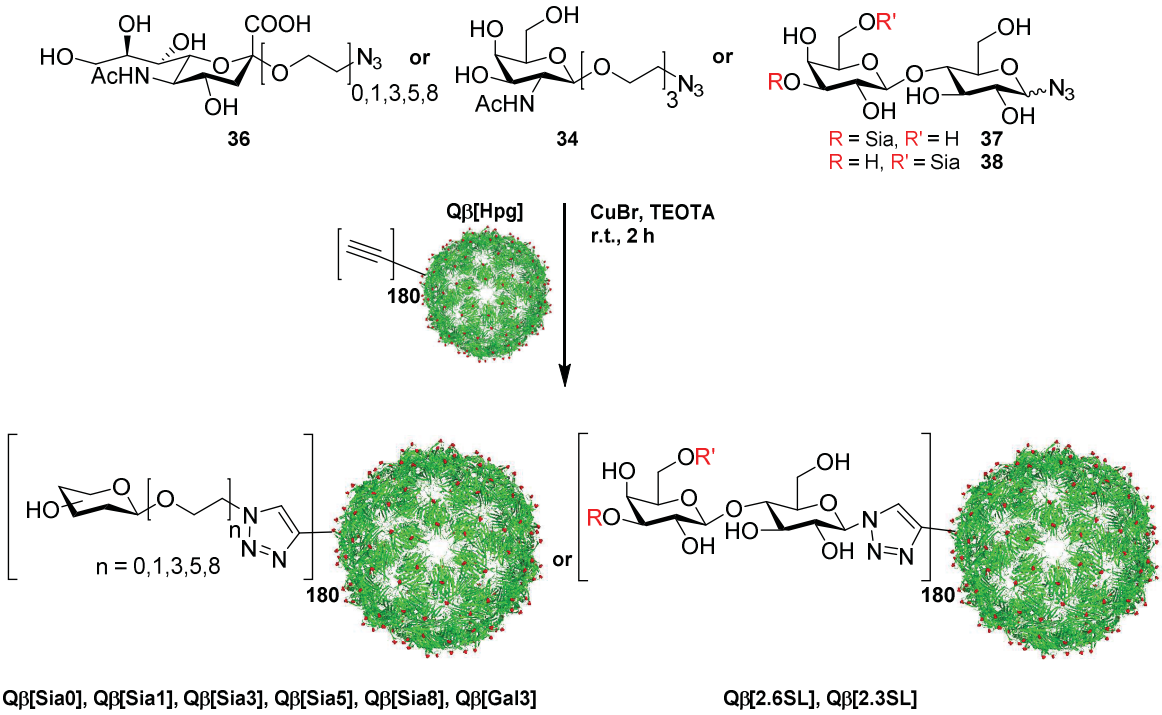
Scheme 41: One-step incorporation of an azido group at the reducing end of 3'- and 6'-sialyllactose.

7.6 Bioconjugation Reactions with Qβ Bacteriophage Capsids

Having all deprotected carbohydrate azides in hand, the next step was the conjugation reaction between these carbohydrates and the Qβ bacteriophage capsid derivative **Qβ[Hpg]**. The protein was auxotrophically expressed by *Saba Nojoumi*, TU Berlin. Isolation and purification of the assembled capsid particles from cell lysate was readily achieved by a PEG precipitation and subsequent ultracentrifugation steps through 10-40% sucrose gradients (section 9.3).[59a, 83] MS analysis of the obtained material confirmed the successful incorporation of desired Hpg as unnatural amino acid at position 16 in the coat protein sequence. Since the Qβ capsid is assembled from 180 coat proteins (see Section 5.4.1), it leads to 180 Hpg residues in a fully assembled Qβ[Hpg] particle.

The conjugation between **Qβ[Hpg]** and the carbohydrate azides was achieved by CuAAC. An already published protocol was chosen for this system, which involved the use of Cu(I)Br as copper source, TEOTA as the Cu(I)-stabilizing ligand, both

solubilized in acetonitrile, together with all the unprotected sugar derivatives which were described above (Scheme 42).[83] The extraordinary stability of **Qβ[Hpg]** towards organic solvents came into play here, since the final acetonitrile content in the reactions was about 33%.



Scheme 42: Bioconjugation reactions by means of CuAAC between Qβ bacteriophage capsid **Qβ[Hpg]** and unprotected carbohydrate azides **34**, **36**-**38**.

After an extensive purification procedure, which involved spin-filtration, dialysis against EDTA buffer, dialysis against pure buffer, and size-exclusion chromatography, the corresponding capsid conjugates were analyzed by intact protein MS and tryptic digest followed by MS/MS experiments. In almost all the reactions, full consumption of the protein starting material and clean formation of the desired conjugates was observed. It has to be noted, however, that additional minor peaks were observed in respective MS spectra. Those could be assigned to protein starting material with an incorporated methionine instead of Hpg and a double CuAAC conjugation product. Methionine can be incorporated through methionine metabolic turnover in the expression host during auxotrophic expression and cannot be ruled out completely.[161] A double CuAAC conjugation product was the result of an additional Hpg present in

the protein sequence. This is the case if the starting methionine is replaced by Hpg during auxotrophic expression and incomplete cleavage of the *N*-terminal amino acid by aminopeptidases occurs.[162] The MS/MS experiments conclusively showed the modification at position 16 of the coat protein and showed no evidence of unselective conjugation to natural amino acid side chains. Only in the case of synthesizing **Qβ[Sia0]** was incomplete conversion and the formation of an unidentified side product observed. **Qβ[Sia0]** was therefore not further considered in any subsequent experiments.

A frequently mentioned problem with CuAAC as a conjugation reaction is the potential risk of residual copper ion retention in the sample, which would result in unwanted toxic effect in a biological testing scenario. To make sure, that no harmful amounts of copper were present in the synthesized Qβ conjugates, ICP-MS measurements were carried out. These experiments revealed that without dialysis as a purification step, micromolar concentrations of copper were indeed retained in a **Qβ[Sia1]** sample (Entry 2, Table 9). After dialysis, the amount of copper was reduced to nanomolar concentrations (Entry 3, Table 9, measured for three independently synthesized **Qβ[Sia1]** samples), which is in the same order of magnitude as a **Qβ[Hpg]** sample prior to CuAAC (Entry 1, Table 9). Having clarified the residual copper concentration post-conjugation, the conjugated Qβ capsids were ready for biological testing, including binding and infection inhibition assays.

Table 9: ICP-MS measurements of the copper content in Qβ samples before CuAAC, after CuAAC without dialysis and after CuAAC with dialysis as additional purification step.

Entry	Sample	Copper content in a 100 nM Qβ capsid sample
1	**Qβ[Hpg]**	0.7 μM
2	**Qβ[Sia1]** no dialysis	15.5 μM
3	**Qβ[Sia1]** dialysis	< 0.5 μM

7.7 Biophysical and Biological Evaluation of Sialylated Qβ Constructs

7.7.1 Cryo Electron Microscopy

Since both interaction partners are in size ranges which are optimally suited for electron microscopy (EM), cryoEM was chosen as a first analytical tool to investigate the interaction between sialic acid conjugated Qβ capsids and influenza viruses. All measurements were performed by *Kai Ludwig* and *Christoph Böttcher*, FU Berlin, and *Sacha de Carlo*, FEI Company. The azidoethylene glycol modified sialic acid conjugates **Qβ[Sia1]**, **Qβ[Sia5]**, as well as **Qβ[Hpg]** and **Qβ[Gal3]** as controls were incubated with A/X31 virus solution, frozen and imaged by cryoEM. As expected, the **Qβ[Hpg]** and **Qβ[Gal3]** did not show any interaction with A/X31 virions (Figure 20A and B) since both samples did not carry any ligand specific for influenza receptors. Thereby, it was shown that there are neither unspecific interactions between the Qβ scaffold itself and influenza virions, nor unspecific hydrophilic interactions when other sugars are present on the Qβ surface. In the case of **Qβ[Sia1]** the picture was completely different. Here, the A/X31 virions were fully wrapped in Qβ particles (Figure 20C). Furthermore, it was observed that the present virions were clumped up and crosslinked by Qβ particles. Having sialic acids attached to Qβ capsid via a longer pentaethylene glycol linker in **Qβ[Sia5]**, however abolished both the previously observed full decoration of X31 virions with Qβ particles, as well as the aggregation of Qβ-crosslinked virions (Figure 20D). These observations already pointed towards a significant difference between Qβ conjugates which carry sialic acid residues with different linker lengths, as well as suggesting that stronger interactions occur with a shorter linker (one ethylene glycol unit vs. five).

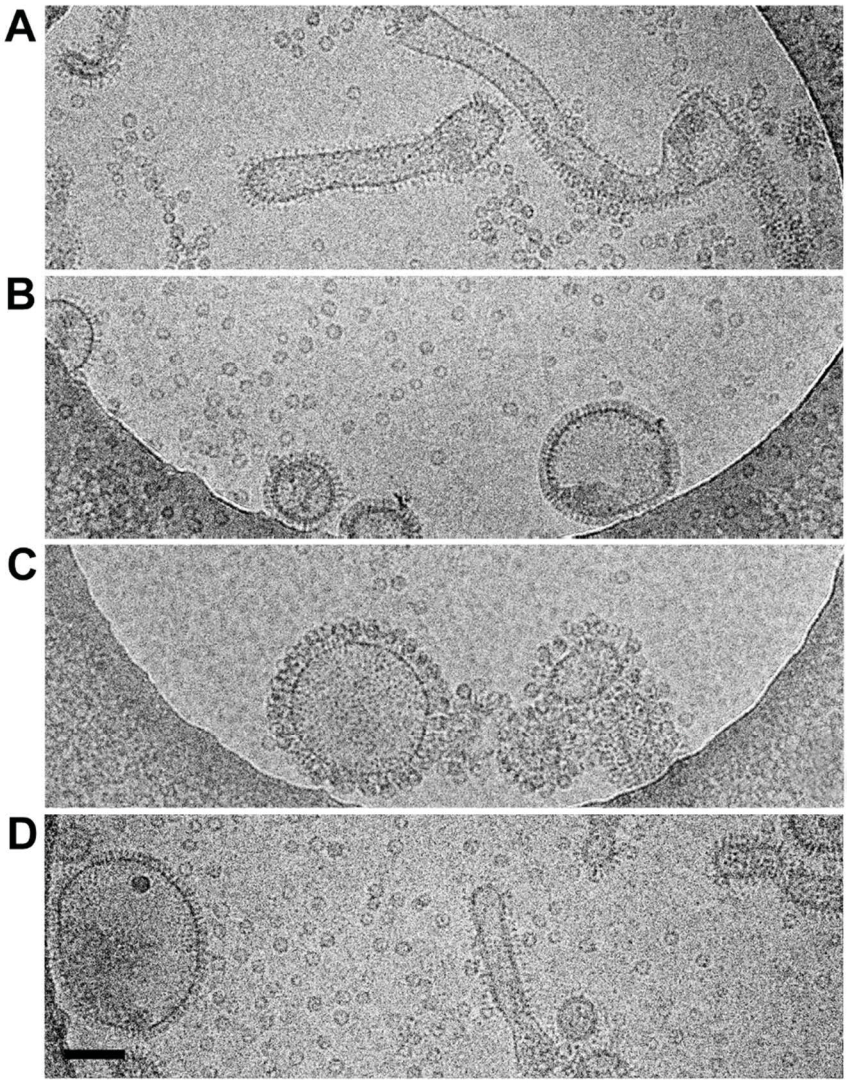


Figure 20: CryoEM images of **A) Qβ[Hpg]**, **B) Qβ[Gal3]**, **C) Qβ[Sia1]**, **D) Qβ[Sia5]** (small homogenous particles) incubated with A/X31 virus (large, irregular shaped particles). Scale bar corresponds to 100 nm.

To obtain a better understanding of the actual binding between **Qβ[Sia1]** particles and X31 virions, cryoTEM tomography was performed (Figure 21). With the resulting three-dimensional representation of a bound influenza virion, it was possible to count exactly 42 **Qβ[Sia1]** particles attached to the viral surface. The distances between the attached **Qβ[Sia1]** particles and the viral surface were found to be in the same range as the known height of HA spikes. A very important observation was the fact that most of the attached Qβ particles were actually found to cover only one HA spike. This points towards a binding scenario in which potentially all three

sialic acid binding sites of an HA spike are occupied by three sialic acid residues of one **Qß[Sia1]** particle.

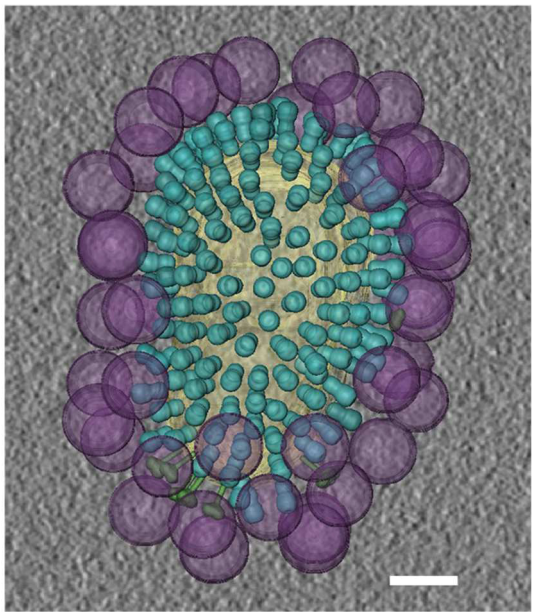


Figure 21: Three-dimensional reconstruction of a cryoEM tomography measurement of an influenza virion covered by Qβ[Sia1] particles. Colors: virus membrane yellow, HA turquois, NA green, and Qβ[Sia1] particles purple. Scale bar corresponds to 25 nm.

This hypothesis can further be supported by the fact that isolated HA spikes (cleaved from the viral surface and free in solution) were still able to bind to **Qß[Sia1]** particles (Figure 22). The series of images in Figure 22 shows that for all observable HA spikes, the HA stem region was located perpendicular to the Qβ surface. If all three sialic acid binding site of the HA head were not occupied, crooked and tangential orientations would also be expected.

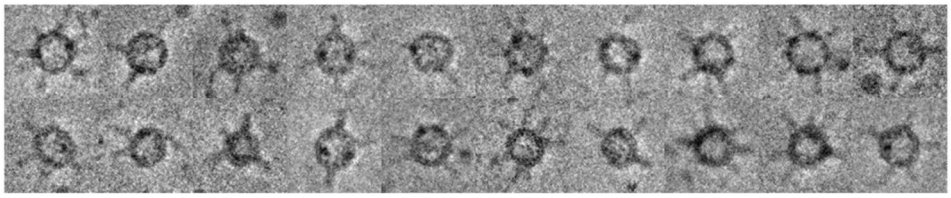


Figure 22: A series of 20 Qβ[Sia1] particles to which cleaved HA spikes are bound.

7.7.2 Hemagglutination Inhibition Assay

Observations from the above described EM measurements were already very insightful, but no quantitative statements could be made. Therefore, the next set of

experiments were hemagglutination inhibition assays (HAI). These experiments were performed by *Daniel Lauster*, HU Berlin, and are part of his doctoral thesis.[163] For that purpose every conjugated Qβ phage capsid was incubated with different influenza viruses and subjected to red blood cells. Upon efficient shielding of the influenza virus with the Qβ phage capsids, the agglutination of red blood cells with the virus is inhibited. From the minimal amount of Qβ needed to inhibit agglutination, the hemagglutination inhibition constant K_{HAI} can be derived. In a first experiment, the ethylene glycol linked Qβ sialic acid conjugates were tested against the A/X31 virus and low K_{HAI} values in the lower nanomolar range were measured (Figure 23). The best performing construct in this set was **Qβ[Sia1]** with a K_{HAI} value of 1.3 nM. This represents an enhancement factor based on the amount of sialic acid residues of 560,000 compared to monovalent 2.6-sialyllactose. These values are remarkable for only 180 sialic acid moieties on the Qβ capsid surface. Other polymeric multivalent sialic acid scaffolds had to carry several hundreds or even thousands of sialic acid residues to achieve at least similar numbers.[46, 164] From the results it also became clear that there is definitely a linker length dependency for K_{HAI} and that shorter linker lengths resulted in lower K_{HAI} values and therefore stronger ability to inhibit agglutination. Such observations are in line with previous observations from cryoEM measurements. The control constructs **Qβ[Hpg]** and **Qβ[Gal3]** did not show any detectable hemagglutination inhibition up to the highest possibly measured concentration, again indicating that there is no unspecific interaction between the scaffold and another carbohydrate. Furthermore, the **Qβ[Sia8β]** construct was tested in this experiment and did not show any hemagglutination inhibition. This clearly demonstrates the specific targeting of the sialic acid binding site of HA by the Qβ capsid which relies on the proper stereochemical configuration (α as in natural occurring glycans) of the sialic acid conjugates.

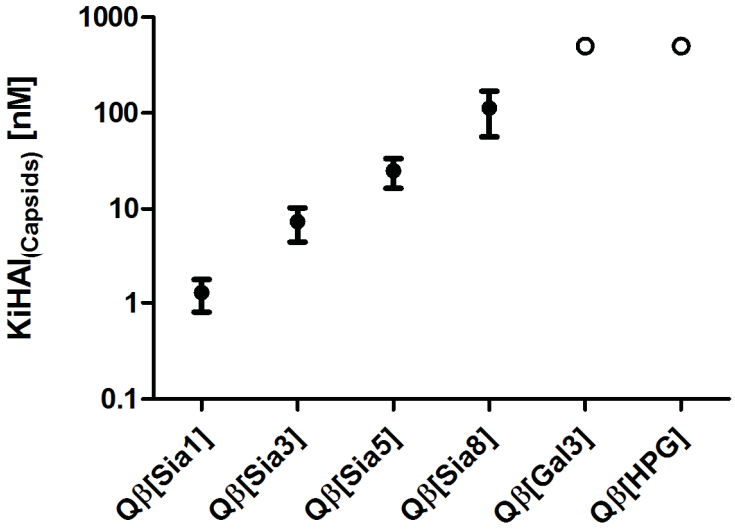


Figure 23. Summarized results of the first set of HAI against A/X31 virus with ethylene glycol linked Qβ sialic acid conjugates as inhibitors. Control constructs **Qβ[Hpg]** and **Qβ[Gal3]** did not show any hemagglutination inhibition ability up to the highest possibly measured concentration (empty circles).

In the next set of HAI assays, multiple different influenza strains were tested against sialylated Qβ particles. The derived K_{HAI} values are summarized in Table 10.

Table 10: Summarized K_{HAI} values against different influenza strains

Virus strain	Qβ[Sia1]	Qβ[Sia3]	Qβ[Sia5]
A/reassortant/X31 (H3N2)	1.3 (±0.5)	7.4 (±2.9)	25.0 (±8.5)
A/Panama/2007/1999 (H3N2)	3.1 (±0.6)	>100	>100
A/mallard/Germany/439/2004 (H3N2)	25.0 (±4.2)	>100	>100
A/Puerto Rico/8/1934 (H1N1)	>100	25.0 (±0.0)	25.0 (±0.0)
A/WSN/1933 (H1N1)	>100	12.5 (±4.2)	12.5 (±0)
A/teal/Germany/2005 (H5N1)	>100	>100	>100
B/Thüringen/2/2006	>100	>100	>100

Table 10 continued

Virus strain	Qβ[Sia8]	Qβ[Gal]	Qβ[Hpg]
A/reassortant/X31 (H3N2)	112.5 (±56.1)	>500	>500
A/Panama/2007/1999 (H3N2)	>100	>100	>100
A/mallard/Germany/439/2004 (H3N2)	>100	>100	>100
A/Puerto Rico/8/1934 (H1N1)	50.0 (±16.7)	>100	>100
A/WSN/1933 (H1N1)	12.5 (±0)	>100	>100
A/teal/Germany/2005 (H5N1)	>100	>100	>100
B/Thüringen/2/2006	>100	>100	>100

The results indicated that for different influenza strains, different linker lengths were optimal for efficient hemagglutination inhibition. For the tested H3N2 strains shorter linker lengths lead to lower K_{HAI} values, whereas agglutination of H1N1 strains was more inhibited with linkers of intermediate length. Agglutination of other pathologically relevant influenza strains such as an avian H5N1 and an influenza B strain could not be inhibited at all with these Qβ constructs. Since the applicability of the simple azidoethylene glycol linked sialic acid residues seemed to be limited, the trisaccharide modified Qβ capsids **Qβ[2.6SL]** and **Qβ[2.3SL]** were tested against the strains which were previously not inhibited (H5N1 and influenza B) in addition to the H3N2 and H1N1 strains which were both inhibited before (Table 11). **Qβ[2.6SL]** was able to inhibit the agglutination of the latter two strains but **Qβ[2.3SL]** did not show any activity. This was the anticipated result since both strains are pathogenic in humans and recognize 2.6-linked sialic residues. Analogously, the avian H5N1 strain was inhibited by **Qβ[2.3SL]** and the human pathogenic influenza B strain was inhibited by **Qβ[2.6SL]**.

Table 11: Summarized K_{HAI} values of **Qβ[2.6SL]** and **Qβ[2.3SL]** against different influenza strains

Virus strain	Qβ[2.6SL]	Qβ[2.3SL]
A/reassortant/X31 (H3N2)	41.7 (±10.4)	>250
A/WSN/1933 (H1N1)	52.1 (±10.4)	>250
A/teal/Germany/2005 (H5N1)	>250	9.1 (±3.4)
B/Thüringen/2/2006	1.7 (±1.1)	>250

7.7.3 Microscale Thermophoresis

To further solidify the biophysical characterization of sialylated Qβ phage capsids and their binding to influenza virions, the binding affinity was measured by microscale thermophoresis (MST). These experiments were performed by *Daniel Lauster*, HU Berlin, and are part of his doctoral thesis.[163] Fluorescently labeled A/X31 viruses were incubated with the azidoethylene glycol linked sialic acid modified Qβ capsids. The resulting apparent dissociation constants $K_{D,app}$ are summarized in Table 12, and are in very good agreement with the K_{HAI} values which were measured before (Section 7.7.2).

Table 12: Summarized apparent dissociation constants $K_{D,app}$ of Qβ capsids measured by MST

	$K_{D,app}$ (nM)
Qβ[Sia1]	1.6 (±0.2)
Qβ[Sia3]	17.3 (±3)
Qβ[Sia5]	32.6 (±8.4)
Qβ[Sia8]	90.6 (±12.3)
Qβ[Gal3]	>500

Both HAI assays and MST measurements, pointed towards a highly affine binding of the synthesized sialic acid carrying Qβ capsids in the lower nanomolar range. The EM images showed that upon binding to an influenza virion, the viral particles were fully covered with Qβ capsids. Furthermore, several different influenza strains could be

bound by these constructs. This suggested that these particles could potentially be used as effective influenza inhibitors, which is topic of the following section.

7.7.4 Infection Inhibition

Infection inhibition experiments were performed by Daniel Lauster, HU Berlin, and are part of his doctoral thesis.[163] A cellular assay was performed first, in order to obtain *in vitro* infection inhibition results. For that, the viral nucleoprotein expression in A549 cells was quantified and related to the quantity of infected cells (Figure 24).

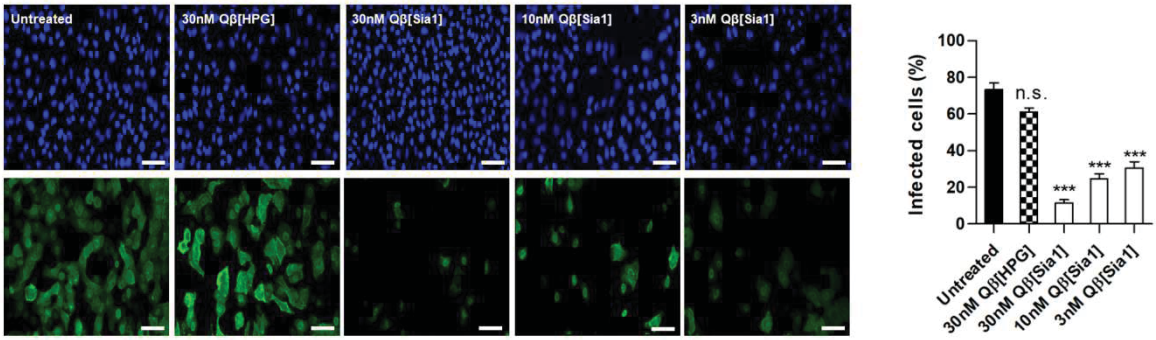


Figure 24: In vitro infection inhibition assay on A549 human lung carcinoma cells against pathogenic A/Panama/2007/1999 virus (blue: cell nuclei DAPI stain, green: viral nucleoprotein; *** = $p<0.005$)

It could be seen that cells which were treated with 3, 10 and 30 nM of **Qβ[Sia1]** showed significantly less viral nucleoprotein expression compared to untreated cells or cells which were treated with unmodified **Qβ[Hpg]** capsids. This result is in good agreement with previously observed binding constants which were all in the lower nanomolar range and showed that the efficient binding of **Qβ[Sia1]** capsids to influenza viruses can be translated into efficient infection inhibition *in vitro*.

The next step was the *in vivo* application of **Qβ[Sia1]** capsids as influenza inhibitors in a mouse model. For that, the A/X31 virus was preincubated with **Qβ[Sia1]** and subsequently administered into the nose of the mouse. Infection progression was followed by monitoring the body weight of the animals (Figure 25). Typically, the buffer treated animals lost the most weight on day four after infection (~12.5%). Upon preincubation with **Qβ[Sia1]** this effect could be significantly reduced to only

~7.5%, which clearly shows that even in an *in vivo* setting infection inhibition can be achieved with the sialylated Qβ capsids presented here.

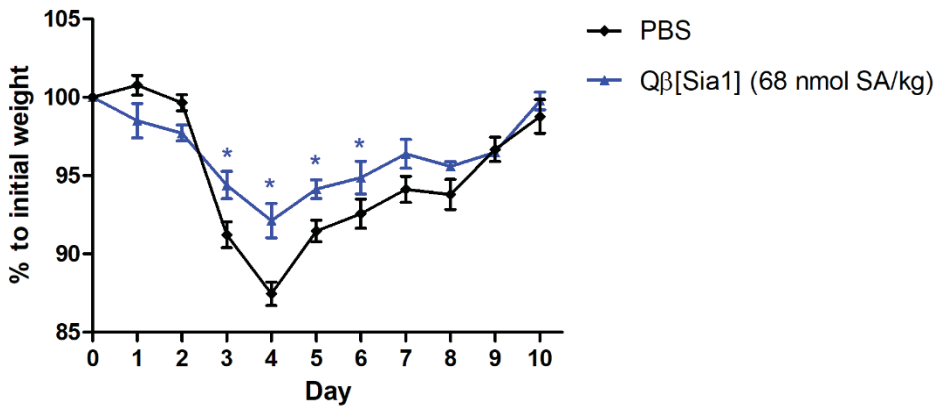


Figure 25: Infection inhibition of A/X31 by **Qβ[Sia1]** in an in vivo mouse model.

In a final set of experiments a human *ex vivo* lung tissue model was used to test infection inhibition in the most relevant tissue for human pathogenic influenza infection. These experiments were performed by *Sandra Sänger*, RKI Berlin. Human lung tissue was infected with A/Panama/2007/99 and subsequently incubated in the presence of 10 nM **Qβ[Sia1]**. The viral load was determined by a plaque assay which showed a reduced viral load after 48 h of incubation compared to tissue which was treated with unmodified **Qβ[Hpg]** as a control (Figure 26). It was possible to further decrease the viral load, when an additional boost of 10 nM **Qβ[Sia1]** was added after 24 h during the incubation (10 + 10 nM).

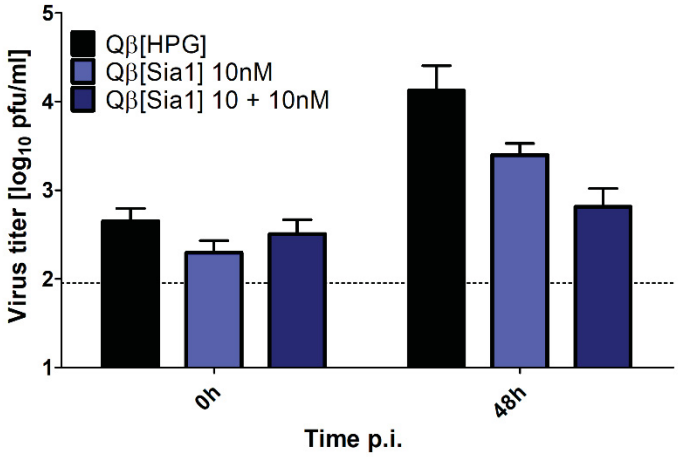


Figure 26: Viral load of infected human lung tissue after infection with influenza and incubation with **Qβ[Sia1]** as inhibitor, determined by a plaque assay.

Taken together these infection inhibition studies suggest that sialylated Qβ capsids could serve as efficient influenza inhibitors in very biologically relevant settings in which only very low nanomolar doses were necessary for efficient infection inhibition.

7.8 Modelling of the Multivalent Binding Event

Based on the defined and monodisperse structure of Qβ phage capsids and the selective conjugation method, it was possible to derive a model of the binding event between ethylene glycol linked sialic acid residues on the surface of Qβ phage capsids and HA. Modelling and computational analyses for that purpose were performed by *Susanne Liese*, FU Berlin, and are part of her doctoral thesis.[165] A mathematical description of the corresponding dissociation constants K_{D} for the above mentioned interaction was found to be

$$K_{\mathrm{D}} = \frac{K_{\mathrm{mono}}^{3}\omega_{\mathrm{sia}}^{3}}{60(C_3(d_1) + C_3(d_2))} \qquad \textbf{Eq. 1}$$

Where K_{mono} describes a single binding affinity of sialic acid to HA, ω_{sia} accounts for the restriction of the rotational degrees of freedom of each Sia unit that arises due to steric repulsion between the Sia unit and the attached ethylene glycol linker, d_1 = 5.0 nm and d_2 =5.9 nm are the edge lengths of two triangles on the capsid

surface with Hpg as the apices (Figure 27A), and C_3 being the cooperativity factor which describes the conditional probability that all three Sia units simultaneously bind to an HA binding site given that one Sia unit is already bound. Inserting all known parameters into equation 1 revealed that the above equation describes observed dissociation constants measured by MST (Section 7.7.3) in good agreement with the experiment (Figure 27B). This fact strongly supported the hypothesis that a single sialic acid-decorated Qβ phage capsids binds to one HA spike by occupying all three sialic acid binding sites of the HA spike.

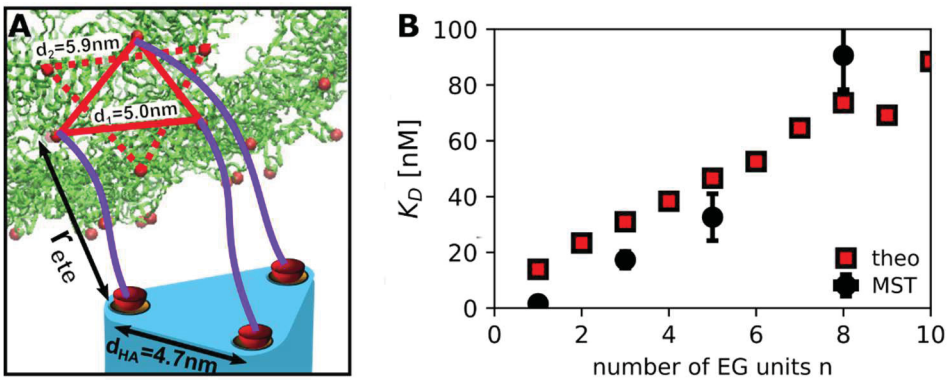


Figure 27: A) Schematic representation of a Qβ capsid bound to HA (blue) in a trimeric binding mode. The triplets on the capsid are indicated by two triangles (red, solid and red, dashed line). The Sia units (red half spheres) are attached to the Qβ capsid via ethylene glycol linkers (purple line), whereby the end-to-end distance of the linker are denoted as r_{ete}. **B)** Comparison between the experimentally and theoretically obtained apparent dissociation constant in dependence of the number of ethylene glycol monomers.

The triangular arrangement of sialic acid residues on the Qβ surface exhibited similar edge lengths as the triangular arrangement of sialic acid binding sites of the HA head (d_1 and d_2 vs. d_{HA}), which turned out to be a crucial component of the overall high binding affinity.

7.9 Interior Modification of Qβ Capsids

Based on the findings described herein, we envisioned to use sialylated Qβ capsids as diagnostic tools for detecting and discriminating between different influenza strains. Such an application, however, would require the attachment of an additional reporter

(e.g. a fluorescent dye) to the capsid constructs to generate a detectable signal. It would be beneficial to attach these additional moieties to the interior surface of the capsid, to not interfere with the sialic acids on the exterior surface. To achieve such an interior modification, several studies report co-expression and encapsidation of fluorescent proteins and packing of nucleic acids,[80, 166] which represents a rather inflexible and limited method with respect to the functional entities that can be utilized. Thus, an approach which would allow for an interior chemical modification of the capsid protein would be preferable. However, such an approach is impeded by the random host RNA which is packed into the capsid upon its expression and assembly and which prevents the access to amino acid side chains for conjugation reactions. The covalent modification of the interior Qβ surface was nonetheless reported in a prior study by the *Finn* group, which required hydrolysis of encapsidated RNA prior to the actual conjugation reaction.[167] This procedure was used here and RNA hydrolysis was induced in the presence of heavy metals at elevated temperatures using 2 mM lead(II) acetate at 37°C for 24 h with which efficient hydrolysis of corresponding encapsidated RNA was achieved. Removal of the packed material could be analyzed by agarose gel electrophoresis and absorption measurements. Upon removal of negatively charged nuclei acids from the Qβ interior, electrophoretic behavior was significantly shifted towards less migration through the gel (Figure 28A and B). Furthermore, the characteristic absorption band of nucleic acids at 260 nm vanished (Figure 28C). The emptied Qβ[Hpg] particles were then ready to undergo bioconjugation reactions on the interior surface.

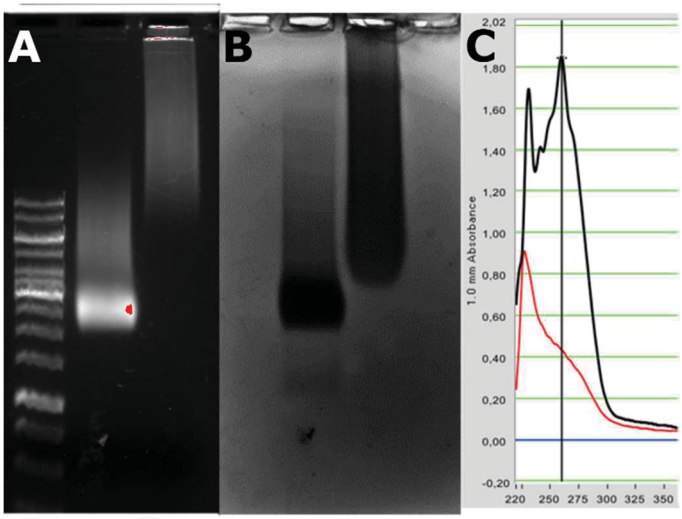
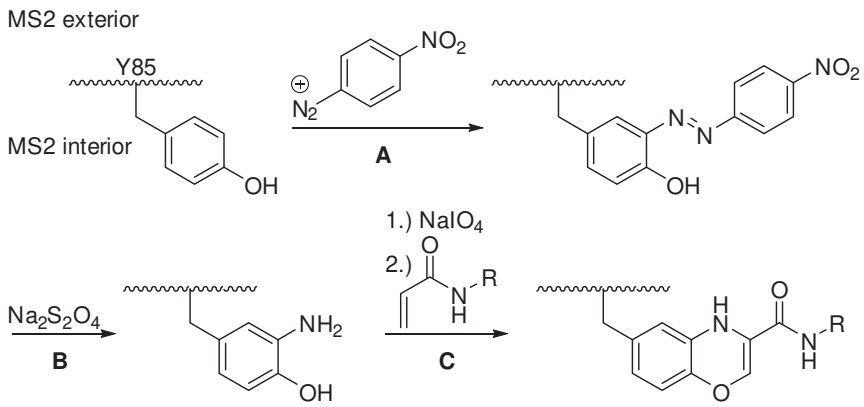


Figure 28: Analysis of Qβ[Hpg] particles which underwent Pb^{2+} induced hydrolysis of encapsidated RNA. **A)** Agarose gel electrophoresis of Qβ[Hpg] before (middle lane) and after hydrolysis (right lane) visualized by the fluorescence signal of ethidium bromide-stained nucleic acids (left lane: marker). **B)** The same gel stained with Coomassie Brilliant Blue to visualize proteins. **C)** NanoDrop absorption measurements before (black line) and after hydrolysis (red line) indicating the vanished nucleic acid band at 260 nm (vertical line).

In search of suitable bioorthogonal reactions it had to be considered that only natural amino acids were available and not all of them were solvent exposed. A comparable issue was reported for the structurally similar MS2 phage by *Francis*.[168] They decided to address a solvent exposed tyrosine at position 85 of the MS2 coat protein with a multi-step protocol which involves a diazonium-coupling reaction with a 4-nitrobenzenediazonium salt (Scheme 43A), obtained from diazotization of 4-nitroaniline. Subsequent reduction by sodium dithionite (Scheme 43B), resulted in an *o*-aminotyrosine derivative which was able to undergo oxidative coupling reactions with electron-poor alkenes (Scheme 43C).



Scheme 43: Interior tyrosine modification of the MS2 capsid as reported by *Francis*.[168] **A)** Diazonium-coupling. **B)** Reductive cleavage by dithionite. **C)** Oxidative coupling with an electron-poor alkene.

Structural similarity between MS2 and Qβ suggests that this procedure could potentially be transferred to the empty Qβ[Hpg] particles, which were prepared earlier. Careful investigation of the Qβ phage capsid crystal structure indeed revealed the presence of interior tyrosine residues at position 89 which could potentially be addressed (Figure 29).

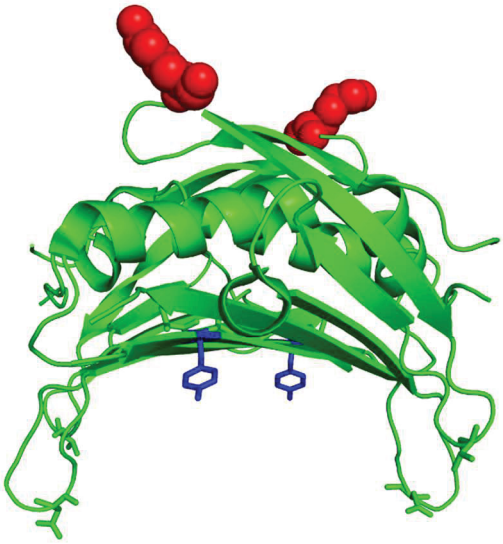
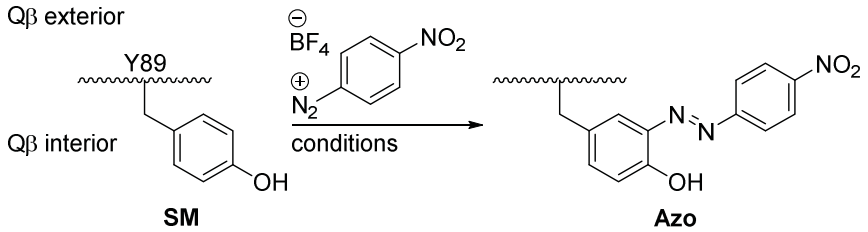


Figure 29: Schematic representation of a Qβ coat protein dimer (pdb: 4L8H) with highlighted position 16 (red spheres) at the exterior surface which was used in this thesis for CuAAC conjugation reactions, and tyrosine 89 (blue sticks) in the capsid interior.

This hypothesis was tested by using commercially available 4-nitrobenzenediazonium tetrafluoroborate (4-NBDT) as diazotization reagent. The use of a solid which can be weighed in accurately was anticipated to be significantly more convenient,

compared to the *in situ* diazotization of the corresponding amine. Reaction conditions were based on the MS2 study,[168] but were screened and optimized for 4-NBDT as the diazonium-coupling agent and Qβ as the protein. For that purpose, a 278 nM (50 µM of coat protein) stock solution of emptied Qβ[Hpg] capsids in potassium phosphate buffer (pH 9.2) was prepared, of which 10 µL was used for each reaction. A 2.5 mM stock solution of 4-NBDT in water was prepared of which 1 µL was added to Qβ[Hpg] under varying conditions (Table 13).

Table 13: Condition screening for the diazonium coupling between 4-NBDT and emptied Qβ[Hpg]. Percentages of starting material (**SM**), mono- and double-diazonium-coupling were derived from MS peak intensities of those three species.



Entry	Equivalents of 4-NBDT	Temperature	Time [min]	SM	1x Addition	2x Addition
1	5	4°C	5	63%	32%	5%
2	5	4°C	10	39%	44%	17%
3	5	4°C	15	36%	48%	16%
4	5	r.t.	10	55%	37%	8%
5	5+5	r.t.	10	41%	42%	17%
6	5+5+5	r.t.	15	4%	43%	53%
7	**5+5**	**r.t. + 4°C**	**15**	**25%**	**51%**	**24%**
8	10	4°C	10	30%	46%	24%
9	5+5	4°C	10	25%	50%	25%

When the emptied Qβ[Hpg] particles were incubated with 5 eq. of 4-NBDT at 4°C for 5 minutes, the mass spectrum showed three distinct species, which corresponded

to the non-reacted protein (**SM**), and the addition of one and two 4-NBDT units (Table 13, Entry 1). It turned out that most of the protein (63%) did not react under these conditions and only 32% of the desired conjugate was formed. 5% of the unwanted double-addition was also observed. Increasing the reaction time to 10 and 15 minutes (Table 13, Entry 2 and 3), led to a better conversion of the starting material, but the desired mono-diazonium-coupling was only slightly increased to 44% and 48%, respectively. Performing the reaction at room temperature for 10 minutes did not improve the outcome, since 55% of starting material was left and only 37% mono-coupling was achieved (Table 13, Entry 4). Building on this, multiple subsequent additions of 4-NBDT were tested. The reaction was again performed at room temperature and another portion of 4-NBDT was added after 5 minutes, which only marginally increased the amount of mono-coupling to 42%, but mainly increased the amount double-diazonium-coupling to 17% (Table 13, Entry 5). Adding even another portion of 4-NBDT after 10 minutes and letting the reaction proceed for another 5 minutes did not increase the mono-addition at all, but rather led to a large amount of double-addition (53%, Table 13, Entry 6). The best result in this screen was obtained by reacting the first portion of 4-NBDT at room temperature for 5 minutes, then cooling the mixture to 4°C and adding another portion of 4-NBDT for 10 minutes at this temperature (Table 13, Entry 7). This led to the formation of 51% of the desired mono-diazonium-coupling product with only 25% of the starting material left. Varying these optimal conditions further by directly using one portion of 10 equivalents of 4-NBDT at 4°C for 10 minutes led to slightly less conversion but still only 46% of the desired product could be obtained (Table 13, Entry 8). Performing the reaction solely at 4°C and reacting the second portion of 4-NBDT only for 5 minutes (Table 13, Entry 9) gave similar results as the optimal conditions in Entry 7.

No conditions were found with which more than 50% of the desired mono-diazonium-coupled product could be obtained. For the desired labeling with a fluorophore,

however, this would still be enough to work with. Thus, the next step, a reductive cleavage promoted by sodium dithionite, was tested with a diazonium-coupled sample according to the published MS2 protocol[168] by adding the solid reagent to the protein and incubating the mixture at room temperature for 2 hours (Figure 30). The reaction mixture was analyzed by MS, which showed unreacted starting material (**SM**) that had not been diazonium-coupled in the previous step, and to a lower extent the desired aminotyrosine (**aY**).

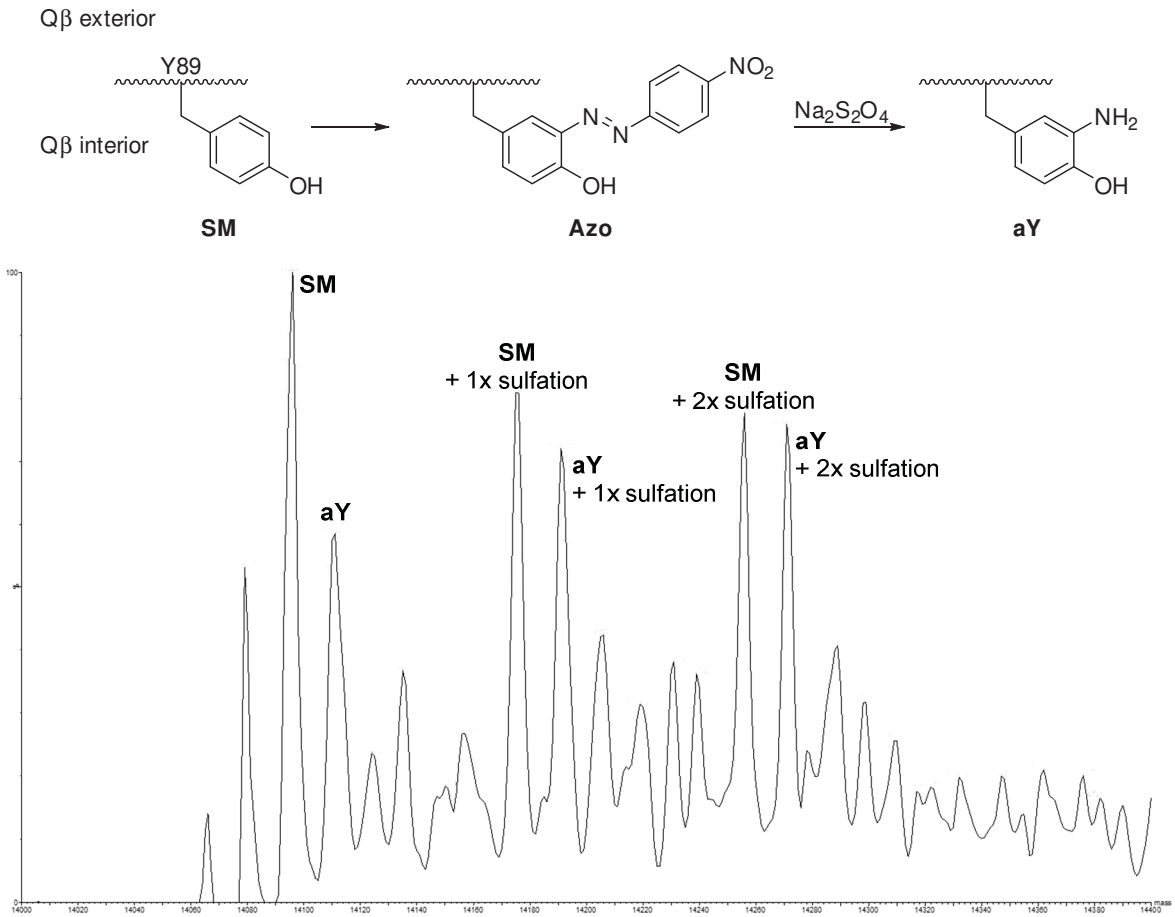


Figure 30: Analysis of the reductive azo-cleavage of diazonium-coupled intermediate (**Azo**) with sodium dithionite towards aminotyrosine (**aY**) by MS.

Hence, the desired product could be obtained. However, a serious issue subsequently arose. Major side products were observed derived from both **SM** and **aY**, with masses of +80 Da from the respective compound. Since **SM** also underwent this side reaction, it could not be a result of the diazonium-coupling, which leaves the sodium

dithionite treatment as a potential source of these side products. In fact, sodium dithionite-induced protein sulfation was observed previously in studies dealing with the dithionite reduction of nitrotyrosine to aminotyrosine,[169] which renders sulfation to be the most probable side reaction here. This always occurred, was not controllable, and the degree of sulfation varied over a large range.

Taken together, this reaction led to a mixture of products which was not desirable for the planned application. Furthermore, it could be conceivable that the sulfation as a side reaction could occur on **aY** itself, which would potentially hamper or prevent a subsequent reaction. Because of these results and the high number of transformation steps which were needed for this strategy, this route was not further pursued.

7.10 Assay Setup for Another Structurally Defined Protein Scaffold

In the PhD thesis of Lukas M. Artner (FU Berlin) from our group another strategy for using structurally defined protein glycoconjugates was studied in order to target the asialoglycoprotein receptor (ASGPR).[143] ASGPR is a membrane protein, which is difficult to crystallize. This membrane protein of hepatocytes is responsible for recognition, binding, internalization and therefore clearing of desialylated glycoproteins from the blood stream. To date only a homology model of ASGPR[170] and a crystal structure of the carbohydrate binding domain[171] are available, which makes it difficult to rationally target this receptor for therapeutic or diagnostic purposes. In the previously mentioned PhD thesis several engineered GFPs were generated. The inherent fluorescence of GFP offered the possibility to use it directly in FACS analysis and to have a fluorescence-based readout for structural integrity of the protein. One of these was designed to carry three Hpg residues at positions which would match the trimeric structural model of the ASGPR. Conjugating ethylene glycol-linked GlaNAc azides by means of CuAAC to these GFP scaffolds led to

fluorescent probes which were recognized and internalized into HepG2 cells by the ASGPR, shown by fluorescence-activated cell sorting (FACS) analysis (Figure 31).

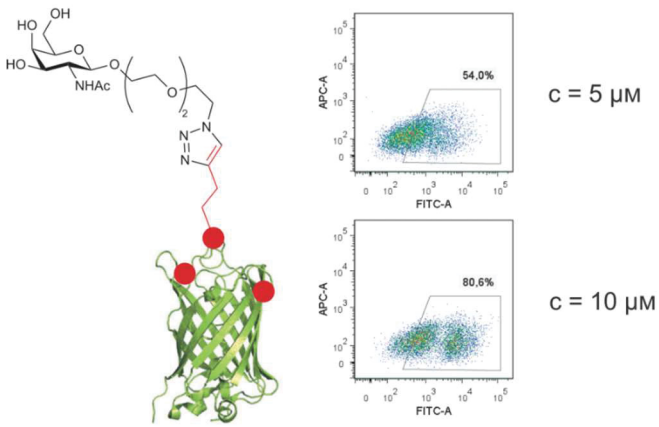


Figure 31: Trivalent GalNAc-GFP conjugate (left) which was internalized by HepG2 cells and analyzed by FACS. Adapted and reassembled from the PhD Thesis of Lukas M. Artner.[143]

It was reported that CuAAC reactions always led to incomplete conversion of the starting GFP[Hpg] and therefore to a mixture of the desired trimeric conjugates, non-reacted starting material and mono- and double-conjugation products.

This conjugation reaction was afterwards optimized by Saba Nojoumi (TU Berlin) and complete conversion of GFP[Hpg] to the desired glycoconjugates was achieved. This work aimed to test these fully modified glycoproteins for their ability to target the ASGPR to be internalized into HepG2 cells. Unfortunately, no or very little uptake was repeatedly observed by FACS analysis in these experiments (data not shown). These were unexpected results and it was not clear whether the problem was due to handling issues or a legitimate different experimental outcome caused by utilizing fully reacted GFP[Hpg]. Therefore, it was planned to perform uptake experiments and to analyze these conjugates by live-cell fluorescent microscopy. Microscopy would additionally allow for the proper tracking of the fluorescent probes, revealing whether they are sticking to the cell membrane or getting internalized, and if so, where they are localizing intracellularly.

To setup these assays and accomplish the analysis by fluorescence microscopy, it was initially necessary to establish suitable positive and negative controls. Commercially available asialofetuin (ASF) and fetuin (Fet) were chosen for that purpose, respectively. Fetuin is a glycosylated protein which occurs in the bloodstream. The protein glycans are terminated by sialic acid, which is why Fet should not be internalized into HepG2 cells by ASGPR mediated uptake. The desialylated variant ASF on the other side, was previously shown to be able to interact with ASGPR and to be taken up into HepG2 cells.[172] Unlike GFP, Fet and ASF had to be labeled fluorescently in order to use them as probes in the anticipated assays. Based on a literature labeling protocol,[172b] Fet and ASF were labeled with NHS-activated carboxyfluorescein (NHS-FAM). The proteins were prepared as a 2 mg/mL solution in basic PBS borate buffer (pH 8) to which NHS-FAM at a concentration of 1 mg/mL dissolved in DMSO was added. After 1.5 h at 37°C, the proteins were purified by size-exclusion chromatography (SEC), analyzed by SDS-PAGE and used for the anticipated cell uptake studies.

In a first experiment, HepG2 cells were cultured on 8-well *ibidi*® slides in RPMI medium which contained phenol red. The cells were incubated with 5 µM of ASF and Fet as a positive and negative control, respectively, at 37°C for 2 h. After washing with fresh medium, the cells were imaged, beginning with cells which were not treated with the corresponding proteins. Significant autofluorescence was detected inside these cells (Figure 32), which would hamper accurate quantification later. It was, however, possible to mostly abolish this interfering autofluorescence by culturing the cells in phenol red-free medium throughout all stages of the experiment.

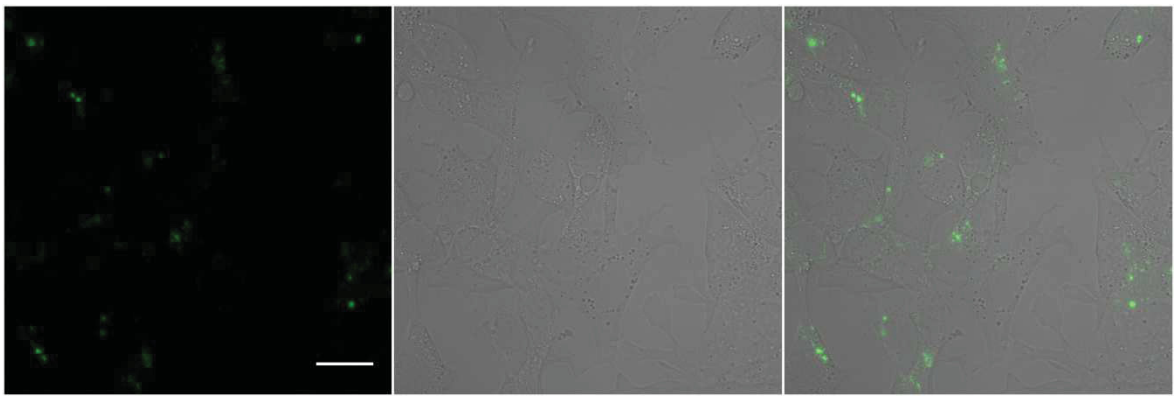


Figure 32: Fluorescence microscopy images of HepG2 cells, which were cultured in phenol red containing medium. Significant fluorescence is detected inside the cells (FAM channel – left, transmission – middle, merge – right, scale bar: 20 µm).

In the next experiment a quantification procedure was tested. Therefore, HepG2 cells were cultured as before with phenol red-free medium and incubated again with 5 µM of the corresponding glycoproteins at 37°C for 2 h. After the washing step, the cells were treated with Hoechst33342 to stain the nuclei for cell counting during image analysis. Next, the raw intensity density (RawIntDen) of the FAM channel was considered and divided by the number of cells, which gave a value that corresponded to "FAM fluorescence per cell". This strategy relies on the assumptions that every cell took up the FAM-labeled proteins to the same extent, and that there is no extracellular background nor autofluorescene in the FAM channel. Since no background fluorescence was detected, autofluorescence was eliminated by phenol red-free medium, and FAM fluorescence was found to be homogeneously distributed over the whole cell-covered area, these assumptions were taken as fulfilled.

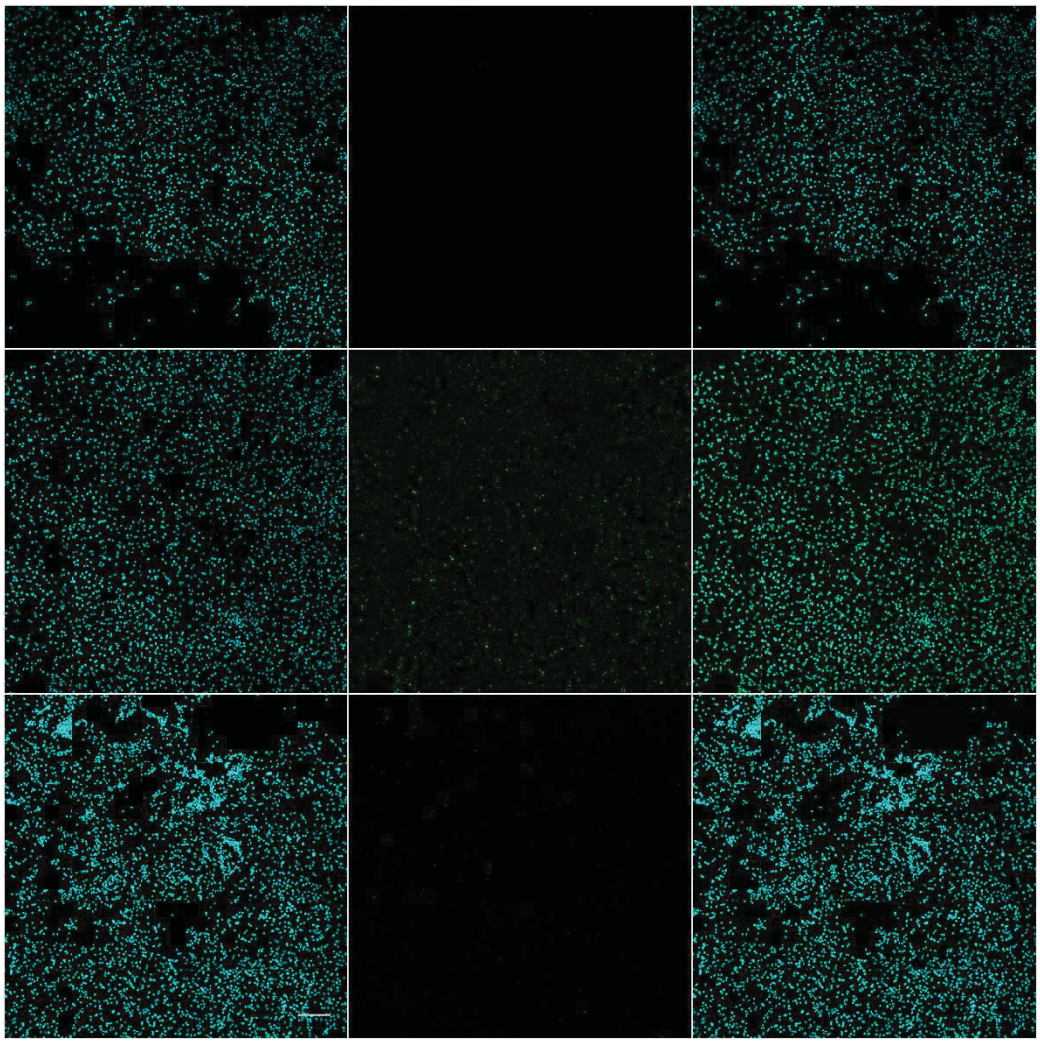


Figure 33: Fluorescence microscopy images of untreated HepG2 cells (upper panel), ASF treated (middle panel), and Fet treated (lower panel) cells. Cell nuclei were stained with Hoechst33342 (left column) and FAM fluorescence (middle column) was detected and merged (right column). Scale bar: 200 µm.

It could already be seen from the microscopy pictures that cells treated with ASF exhibited more FAM fluorescence compared to both the untreated and Fet treated cells (Figure 33). In the case of Fet, there was still some fluorescence detected in the cells, but much less than in the ASF treated cells. These estimates were supported by applying the above describe quantification method (Figure 34).

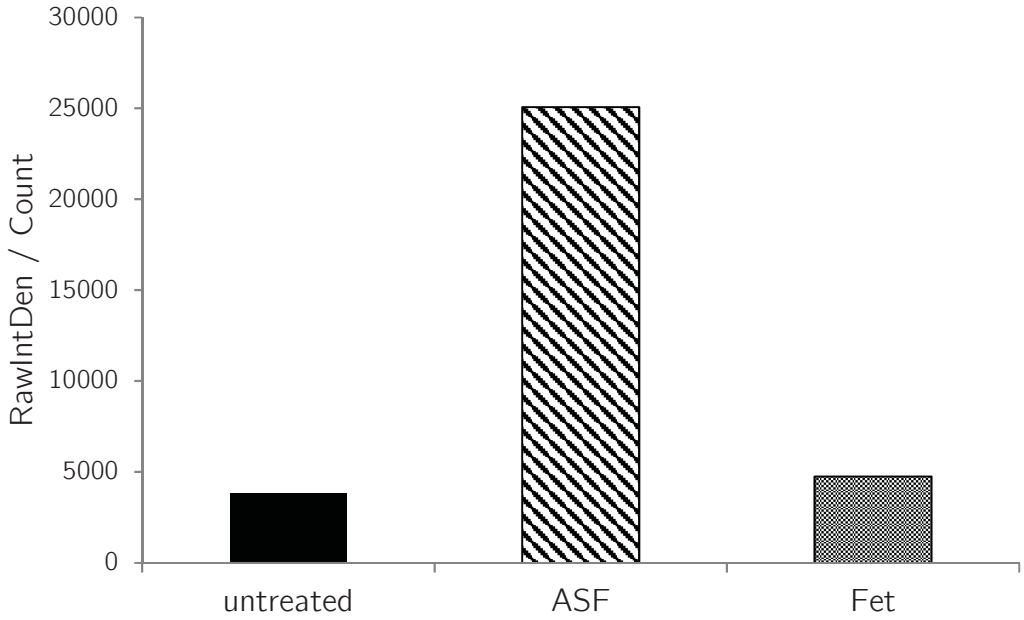


Figure 34: Quantification of labeled protein internalized by HepG2 cells using RawIntDen of the FAM channel divided by the number of cells. 5 µM of ASF and Fet were incubated at 37°C for 2 h.

The quantification showed that ASF resulted in five times more uptake into HepG2 cells than Fet under these conditions and that the Fet fluorescence is only marginally higher compared to the untreated sample. These observations conclusively showed that ASF and Fet could be used as a positive and negative control, respectively, as well as that the proposed quantification method worked and gave reasonable results.

Additional control experiments were performed to verify an ASGPR mediated uptake mechanism. To this end, endocytosis inhibitor cytochalasin D (CytD) was used in order to block receptor-mediated uptake pathways.[173] If labeled ASF got internalized via such a pathway, intracellular fluorescence should be significantly reduced upon coincubation with CytD. Another option to this end would be to block ligand binding sites of ASGPR by applying saturating concentrations of monomeric carbohydrates such as galactose. The influence of CytD is exemplified in Figure 35. HepG2 cells were cultured and treated as before, but additional samples were included which were coincubated with 10 µM of CytD. Again, the untreated cells only showed very little fluorescence and ASF was internalized much more than Fet. Furthermore, Fet uptake was essentially not influenced by the presence of CytD, which is an expected

observation since Fet should not be taken up at all by ASGPR. Samples which were treated with ASF and CytD, however, showed inconclusive results. For some samples, cellular fluorescence was reduced in the presence of CytD, but in other cases the observed fluorescence was even higher (Figure 35, ASF2+CytD) although ASF samples without CytD showed very similar fluorescence (Figure 35, ASF vs. ASF2).

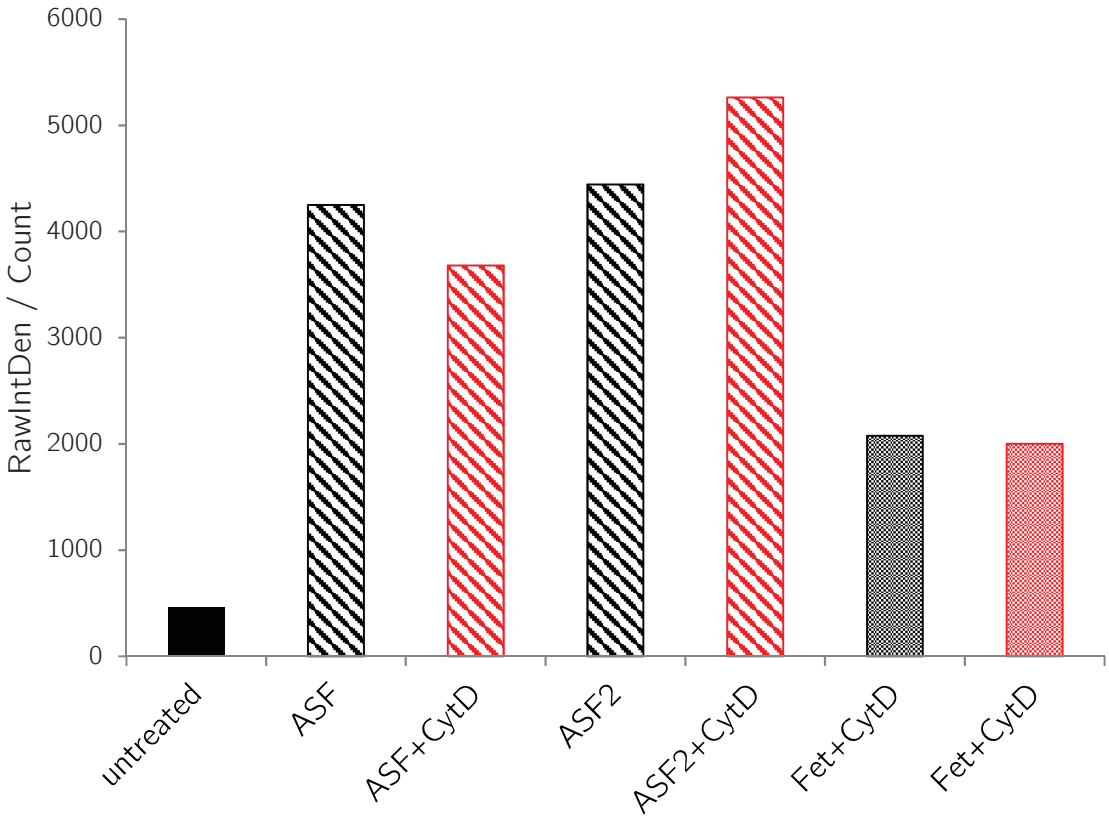


Figure 35: Quantification of labeled protein internalized by HepG2 cells using RawIntDen of the FAM channel divided by the number of cells. Cells were treated with CytD and without CytD as an inhibitor. Two samples of ASF treated cells were quantified (ASF and ASF2).

Using galactose as an inhibitor for ASGPR-mediated uptake was also tested by coincubation with 5 and 10 mM of monomeric galactose (Gal) and otherwise unchanged conditions. These conditions did not decrease the intracellular fluorescence, but actually slightly increased fluorescence was observed for both ASF and Fet treated cells at both applied concentrations of Gal (Figure 36).

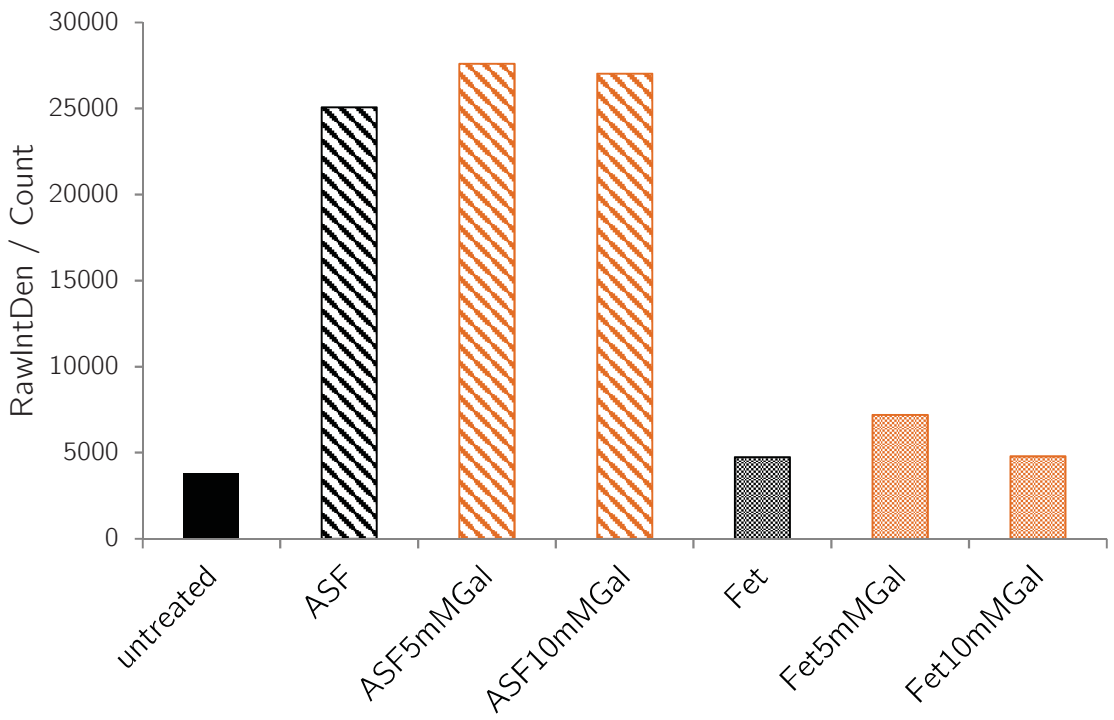


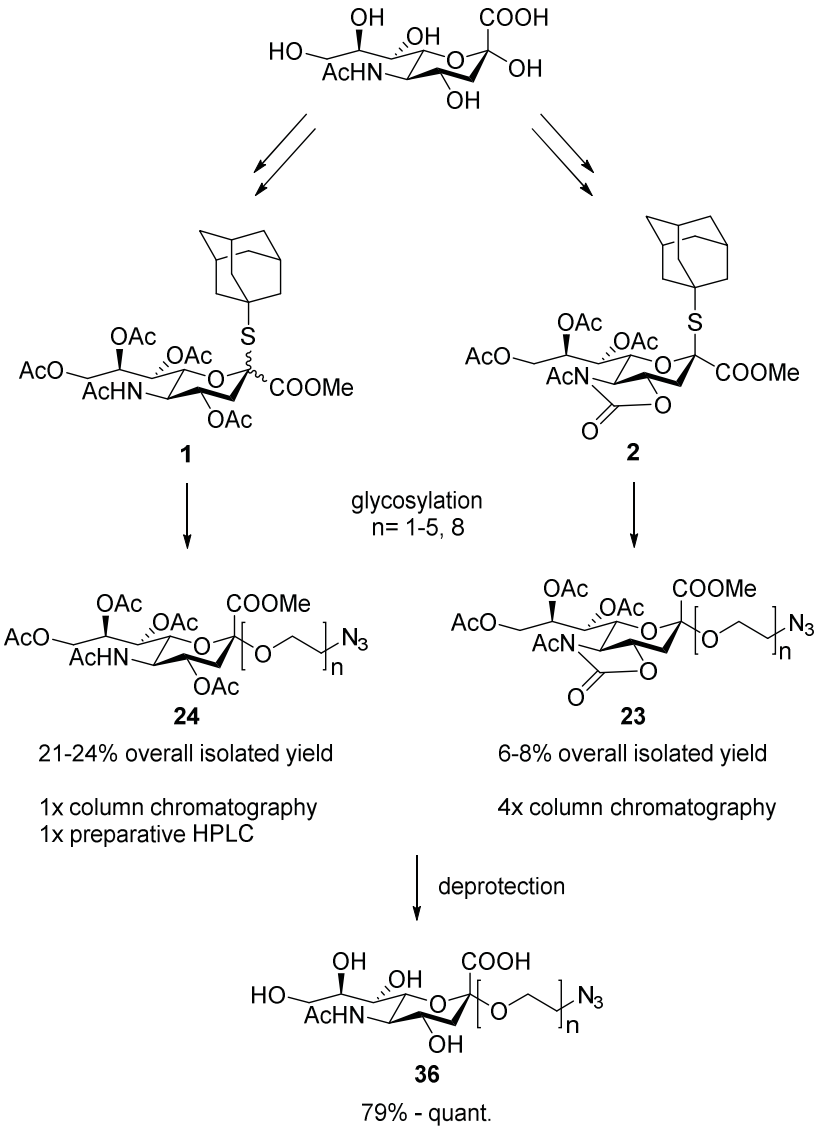
Figure 36: Quantification of labeled protein internalized by HepG2 cells using RawIntDen of the FAM channel divided by the number of cells. Cells were treated with 5 and 10 mM Gal as inhibitor.

The results showed that ASF and Fet could in principle be used as a positive and negative control for ASGPR-mediated uptake studies. An unequivocal verification of this receptor-mediated uptake, however, could not be obtained, since utilized inhibitors showed unexpected and inconclusive results. Especially for the anticipated application, where only small differences between GFP conjugates were to be expected, this setup appears to be not sensitive enough.

8. Summary and Outlook

This work describes the use of engineered proteins as multivalent scaffolds and their use in a biological context. The main focus lay on the synthesis and evaluation of novel influenza entry inhibitors based on sialic acid conjugates of the bacteriophage Qβ capsid. Therefore, sialic acid derivatives had to be accessible in reasonable quantities which is why the first part of this thesis dealt with the evaluation of 5-*N*-acetyl adamantanyl thiosialoside **1** as a glycosylation donor for the synthesis of α-linked sialosides. Naturally occurring sialic acids exist exclusively as α-sialosides and their recognition and function relies on the proper stereochemical configuration. A lot of protecting group manipulations are necessary for the synthesis of current state-of-the-art oxazolidinone protected sialyl donor **2**. However, **2** ensures almost exclusive α-selectivity in glycosylation reactions. In order to facilitate the synthetic accessibility of sialic acid derivatives which can be used in subsequent protein conjugation reactions, faster and higher yielding strategies are desired.

It was possible to synthesize a range of azidoethylene glycol linked sialic acid derivatives with the synthetically easier accessible **1** as sialyl donor in good yields and α-selectivities. To achieve this, more equivalents of activating triflic acid had to be used compared to the activation of oxazolidinone protected donor **2**. In a direct comparison of overall isolated yields during this work, donor **1** performed clearly better with 21-24% compared to donor **2** which gave only 6-8% of the desired α-linked sialosides (Scheme 44). Reported yields in the literature, however, for the preparation of **2** are higher than those achieved in this work. Considering this fact, the method described in this thesis still remains an attractive alternative, since a significant lower number of synthetic steps and purifications were necessary for the synthesis of desired α-linked products.



Scheme 44: Summarized overall isolated yields starting from neuraminic acid of desired protected α-linked sialosides modified with a set of azidoethylene glycols. By deprotection under basic conditions, sialic acid derivatives were obtained for the subsequent CuAAC conjugation to Qβ capsids.

The optimized reaction conditions were successfully transferred to the synthesis of sterically more demanding glycosyl acceptors such as cyclohexanol, adamantanol, and other protected carbohydrates, maintaining the good α-selectivities. Only in one case, with methyl-2,4,6-tri-*O*-benzyl-β-D-galactopyranoside **19** as glycosyl acceptor, the reaction failed because of the acid-lability of the acceptor itself. Since anomeric mixtures are formed as the product of these glycosylation reactions it was necessary to separate these, which was readily achieved by preparative HPLC for all the described cases. Isolation and separation of the corresponding anomers by flash

column chromatography would be also beneficial to further increase the isolated yield. This failed however for the azidotriethylene glycol linked sialic acid derivative. For larger compounds, e.g. when a benzyl protected glucose derivative was used as an acceptor, isolation of the desired α-anomer by column chromatography was easily possible and gave a higher isolated yield compared to the HPLC purified case.

Future aspects which could be addressed along this line are to further extend the substrate scope of the found reactions conditions. It was observed, that some benzyl protected acceptors are not suited for the reported method and therefore other compatibly protected derivatives have to be found in order be able to glycosylate these kinds of substrates. Furthermore, the influence of the solvent which is used for the glycosylation reaction can be further studied. It is known that nitrile solvents are beneficial in terms of α-selectivity, which is why usually acetonitrile is applied in a mixture with dichloromethane. The amount of acetonitrile which can be used is however limited when the reactions are performed at –78°C because of the relatively high melting point (–45°C). Therefore, it could be beneficial to switch to a different solvent system which allows for a higher content of a nitrile solvent or even no co-solvent at all. One possibility could be to exchange dichloromethane for a more polar solvent, e.g. DMF or propylene carbonate. Those solvents should be able to solvate the intermediately formed oxocarbenium ions better than dichloromethane and therefore increase their reactivity. Furthermore, propionitrile as a homologue of acetonitrile exhibits a lower melting point (< –80°C) and the glycosylation reaction could potentially be performed without any co-solvent at –78°C.

In the second part of this thesis, synthesized azidoethylene glycol linked sialic acids were deprotected and conjugated to the Qβ bacteriophage capsids carrying alkynes as reactive handle. The conjugation was readily achieved by copper-catalyzed alkyne-azide cycloadditons and the fully modified Qβ particles were obtained and characterized. No unmodified coat protein was found by MS analysis after the

reaction which points towards a full conversion of the starting protein. In order to characterize the ability of these sialylated Qβ constructs to bind to hemagglutinin, several methods were combined in a comprehensive study performed by collaborators. First, electron microscopy was used to directly visualize the interaction of the synthesized capsids with influenza. It was observed that with short linker lengths, as in **Qβ[Sia1]**, influenza virions were fully wrapped with, as well as clumped and interconnected by the corresponding Qβ particles.

Further quantitative analysis of the actual binding by hemagglutination inhibition assays and microscale thermophoresis measurements showed binding affinities in the lower nanomolar range. The **Qβ[Sia1]** construct as the best binder resulted in a $K_{D,app}$ value of 1.6 nM. This represented an enhancement factor compared to monomeric sialyllactose of ~10^5. It should be kept in mind that this high affinity was obtained with only 180 sialic acid residues presented on the surface of the Qβ capsid. Compared to other polymeric multivalent scaffolds or nanoparticles, the number of sialic acid residues which is necessary to achieve such affine binding is significantly reduced. Based on an established theoretical model of the underlying binding event, it turned out that the way how the sialic acids are spatially arranged on the capsid surface, played a major role in this binding enhancement.

The tight binding together with the electron microscopy pictures led to the hypothesis that sialylated Qβ particles are potential influenza infection inhibitors. To test this, several infection inhibition experiments, ranging from a cellular *in vitro* assay, over an *in vivo* mouse model to an *ex vivo* human lung tissue setup, were carried out. In all the experiments it was possible to inhibit the influenza infection with low nanomlar doses of **Qβ[Sia1]**.

Important next steps for the further development of sialylated Qβ capsids as influenza inhibitors include the synthetic scale up in order to be able to provide

sufficient amounts of the corresponding constructs. A big step towards this was achieved in the first part of this thesis, by significantly streamlining the synthetic accessibility of corresponding sialic acid azides. The production of large quantities of recombinantly expressed Qβ capsids is already established but transferring an auxotrophic expression system to this scale is certainly a special challenge. Furthermore, additional safety issues have to be addressed, namely toxicity profiling has to be performed, as well as studying the immune response of an organism to sialylated Qβ particles.

Apart from the application as actual influenza inhibitor, the here presented system could also be applied as a diagnostic tool for the analysis of influenza strains. It was shown in this thesis that it is possible to discriminate between avian and human pathogenic influenza strains when 2.3- and 2.6-sialyllactose derivatives, respectively, were conjugated to the Qβ scaffold. If these **Qβ[2.3SL]** and **Qβ[2.6SL]** constructs were additionally labeled with two different fluorophores, an easy analytical tool could be established to discriminate newly emerging influenza strains from nature, of which the host specificity is not known, yet. Modifying the interior surface of Qβ capsids towards this end was explored in this thesis, since the exterior targeting surface would remain unaltered. Tyrosine Y89 residue of the Qβ coat protein was chosen as the conjugation target for that purpose. It was necessary to remove randomly encapsulated genetic material from the expression host, which was achieved by incubating Qβ[Hpg] with Pb^{2+} to induce hydrolysis of the corresponding RNA. A subsequent diazonium-coupling reaction with commercially available 4-nitrobenzenediazonium tetrafluoroborate was studied and various reaction conditions were tested. The optimal reaction setup involved two additions of 5 equivalents of the diazonium reagent, the first at room temperature and the second at 4°C, with which 51% of the desired mono-diazonium-coupled product was obtained. To ultimately generate a reactive handle which could be further modified, it was necessary to perform a reductive cleavage of the diazo-bond with sodium dithionite

to generate an aminotyrosine derivative. It turned out that this reaction led primarily to the formation of sulfated side products. Because of these problems and the cumbersome strategy, this was not further pursued. However, an interior modification of the Qβ capsid would still be an attractive tool for labeling or delivery purposes. Another possibility to achieve this goal could be to incorporate an additional cysteine residue on the interior surface of the capsid. This would reduce the effort for interior conjugation only to the hydrolysis of packed RNA and many fluorescent dyes, drugs or other probes are already commercially available as thiol-reactive reagents. Furthermore, expression of the Qβ coat protein as a fluorescent protein-fusion could also be used to generate fluorescent capsid particles without any other additional labeling step.

In the last part of this thesis, the setup of a quantitative assay based on live-cell fluorescence microscopy was attempted. Based on earlier results, it was necessary to establish a positive and negative control for the ASGPR-mediated uptake of fluorescently labeled proteins into hepatic HepG2 cells. Therefore, asialofetuin and fetuin were labeled with NHS-activated carboxyfluorescein and subsequently tested in cellular uptake studies. Asialofetuin was observed to be internalized five times more than the sialylated fetuin. Experiments to verify the receptor-mediated nature of this uptake by using cytochalasin D as endocytosis inhibitor and monomeric galactose to block ligand binding sites, however, were inconclusive. This observation combined with the fact that the anticipated application would result in minimal uptake efficiency variations demonstrated that the microscopy setup appeared to be not viable. To overcome these issues, it would be necessary to screen more inhibitors so as to find one which can inhibit ASGPR-mediated uptake enough to observe clear differences in cellular uptake. With such an inhibitor in hand it could be conceivable to switch again back to FACS for quantification purposes which would give more reliable and comparable results.

9. Materials and Methods

9.1 General Information

9.1.1 Reagents and Solvents

Dry solvents were obtained from *Acros Organics*. All other solvents and reagents were purchased from commercial suppliers as reagent grade and used without further purification. Acidic resin for ion-exchange/neutralization (*DOWEX®-50WX4*) was washed extensively with methanol prior to use. Acid-washed molecular sieves were prepared according to a published protocol by soaking 4 Å molecular sieves with 2 N HCl for 12 h, subsequent removal of HCl under reduced pressure, washing with water and drying at 200°C for 24 h.[140]

9.1.2 Column Chromatography

The silica gel for flash column chromatography was *Acros Silica gel 60 Å, 0.035-0.070 mm*. Mobile phases are given separately (*vide infra*) for the respective compounds.

9.1.3 Preparative HPLC

Preparative HPLC purifications were performed on a *Gilson PLC2020* equipped with a *VP250/32 NUCLEODOUR C18 HTec, 5 µm* column (flow: 30 mL/min; A = H_2O, B = MeCN; gradient: 5% B 0-5 min, 5-90% B 5-60 min).

9.1.4 NMR

1H- and 13C-NMR spectra were recorded on a *Bruker Avance 300* (300 MHz) and *Bruker Avance 600* (600 MHz). Measurements were done at room temperature and chemical shifts are reported in parts per million (ppm) relative to the residual solvent peaks.[174] Coupling constants *J* are reported in Hz and multiplicities are given as s = singlet, d = doublet, t = triplet, q = quadruplet, m = multiplet, b = broad.

9.1.5 UPLC-MS

LC-MS analyses were carried out on a *Waters H-class* instrument, equipped with a *Quaternary Solvent Manager*, a *Waters* autosampler, a *Waters TUV* detector connected to an *Acquity QDA* mass detector using an *Acquity UPLC-BEH C18 1.7 µm 2.1x50mm RP* column (flow: 0.6 mL/min; A = H2O + 0.1% TFA, B = MeCN + 0.1% TFA; short gradient: 5% B 0-0.5 min, 5-95% B 0.5-3 min, 95% B 3-4 min; long gradient: 5% B 0-2 min, 5-95% B 2-13 min, 95% B 13-14 min).

9.1.6 High-resolution ESI-MS

High resolution ESI-MS spectra were recorded on an *Agilent 6220 TOF Accurate Mass* coupled to an *Agilent 1200 LC* and were measured at 35°C between 100–2000 m/z. The column was an *Accucore RP-MS* (30 x 2,1mm; 2.6µm particle size) ran with a flow of 0.8 mL/min and the following gradient (A = water, B = acetonitrile): 95%A + 5%B for 0.2 min, then 95%A + 5%B to 1%A + 99%B until 1.1 min, then 1%A + 99%B until 2.5 min.

9.1.7 IR Spectroscopy

IR spectra were recorded on a *Bruker Tensor27* instrument by the ATR technique at room temperature. Samples were applied as a solution and measurements were done after the solvent had evaporated.

9.1.8 Optical Rotation

Optical rotation was measured on a *Jasco P-2000 Polarimeter* at 589 nm (sodium lamp) and specific rotations are given in $deg \cdot mL \cdot dm^{-1} \cdot g^{-1}$.

9.1.9 Protein Concentration Determination

Concentrations of Qβ containing solutions were measured by commercially available *Pierce™ BCA Protein Assay Kit* (catalogue number 23225) using the BSA calibration curve (0-2000 µg/mL). Evaluation of the assay results was carried out by using the *Mathematica* software (*Wolfram Research, Inc.*, Version 10.2, Champaign, IL).

9.1.10 Intact Protein MS

Intact protein masses were measured by LC-ESI-MS. For that, 8 µL of capsid solution (23 – 177 µM of protein) were treated with 2 µL of 1 M TCEP and 4 µL of this mixture were injected to an *Acquity* UPLC (*Waters Corporation*) equipped with a *Waters Acquity UPLC Protein BEH C4 300 Å* column using water (A) and acetonitrile (B) as eluent which both contained 0.1% formic acid (95% A isocratic for 2.5 min, then gradient to 100% B in 7.5 min). The UPLC was coupled to an *LCT Premier Micromass Technologies* spectrometer operated at 3.2 kV capillary voltage, 60 V cone voltage, 300°C desolvation temperature, 150°C source temperature and 600 L/h desolvation gas. The obtained data were processed by *MassLynx* software (4.1, *Waters Corporation*) and protein spectra were deconvoluted by *Waters MaxEnt1* algorithm. Myoglobin was used as reference for spectra calibration.

9.1.11 Tryptic Digest and ESI Orbitrap MS/MS

Enzymatic in-gel digestion of Qβ conjugates was performed following a standard protocol.[175] Follow-up analysis was carried out on an *Orbitrap Fusion* (*Thermo Scientific*) connected to a *Dionex Ultimate 3000 NCS-3500RS Nano LC* system. Each sample was loaded for 5 min at a flow of 5.6 µL/min (2% acetonitrile, 0.05% TFA) onto an *Acclaim PepMap 100* C_{18} (3 µm, 75µm × 20 mm) precolumn, followed by gradient elution (10% acetonitrile, 0.1% formic acid for 7.8 min, 18% acetonitrile in 1 min, followed by 30% acetonitrile in 20 min and 50% acetonitrile in another 5 min, then 80% acetonitrile in 4 min, flow = 200 nl/min) on a *PikoTip* column (75 µm × 250 mm) equipped *with ReproSil-Pur 120* C_{18}*-AQ* 3 µm material. Mass spectra were obtained in a range of 350–1500 m/z with a resolution of 12000. Subsequent MS/MS data were obtained by electron-transfer dissociation to detect intact carbohydrate modifications.

9.1.12 Electron Microscopy

Electron microscopy was performed by Kai Ludwig, Christoph Böttcher (FU Berlin) and Sacha de Carlo (FEI Company).

9.1.12.1 Cryo-Electron Microscopy

For binding experiments 15 µl phage capsids (concentration 200 nmol/L) were incubated with 15 µL influenza virus A/X31 (~1 mg/mL total protein) with gentle agitation for 30 min at room temperature. Perforated (1 µm hole diameter) carbon film-covered microscopical 200 mesh grids (R1/4 batch of *Quantifoil*, *MicroTools GmbH*, Jena, Germany) were hydrophilized by 60 s glow discharging at 8 W in a *BALTEC MED 020* device. Then 5 µL of the corresponding sample were pipetted to the hydrophilized grid and the supernatant fluid was immediately removed with a piece of filter paper until an ultrathin layer of the sample solution (typical thickness of 100-250 nm) was obtained spanning the holes of the carbon film. The sample was instantly vitrified by plunging the grids into liquid ethane using a guillotine-like apparatus. The vitrified sample was subsequently transferred under liquid nitrogen into a *Tecnai F20 TEM* (*FEI*, Oregon, USA) equipped with field emission gun and operating at 160 kV by the use of a *Gatan* tomography cryo-holder (*Model 914*). Microscopy was carried out at a 94 K sample temperature using the low-dose protocol of the microscope. Micrographs were taken with a *FEI Eagle 4k × 4k* CCD camera using the twofold binning mode. To generate stereo images the same image section was recorded at two different tilt angles (4° and -4°) using the C*ompustage* of the microscope. The resulting micrographs were aligned using the software *StereoPhoto Maker* (*Masuji Suto*, Japan).

9.1.12.2 Cryo-Electron Tomography

Tomograms were acquired on a *Titan Krios™ Transmission Electron Microscope* (*FEI*, Oregon, USA) operating at 200 kV and equipped with a *Volta phase plate*. Single axis tilt series (±64° in 2° steps) were recorded with a *Falcon-II 4k × 4k* direct

electron detector at 29 K primary magnification and a defocus of -17 µm with a total dose less than 70 e^-/Å.

Tomogram reconstruction was done using *Imod* (Kremer J.R., D.N. Mastronarde and J.R. McIntosh (1996)), computer visualization of three-dimensional image data using *IMOD*[176] and *Inspect3D* (*FEI*, Oregon, USA). Manual segmentation and 3D rendering were performed using *Amira, Version 6.0* (*FEI*, Oregon, USA).

9.1.13 Biological and Biophysical Methods

Biological and biophysical experiments were performed by Daniel Lauster (HU Berlin).[163]

9.1.14 Human *ex vivo* Lung Tissue Experiments

Experiments with human lung tissue were performed by Sandra Sänger (RKI Berlin). Lung tissue was infected with $4x10^5$ pfu for 90 min in 500 µL RPMI medium at r.t. After infection lung tissue was washed several times and subsequently incubated for 48 h at 37°C in RPMI medium containing either 10 nM or 100 nM of Qβ[Sia1]. Viral load in the supernatant was determined by plaque titration on MDCK cells.

9.1.15 Modelling

Modelling and theoretical considerations were performed by Susanne Liese (FU Berlin)

9.1.15.1 Molecular Dynamics Simulations

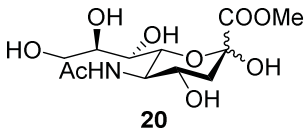
The MD simulations are performed with the *GROMACS* simulation package and the charmm35r forcefield. The following simulation parameters are used: The time step is set to 2 fs. The temperature is set to 300 K. For temperature-/pressure coupling the v-rescale/Parrinello-Rahman algorithm with a relaxation time of 0.1 ps/0.5 ps is chosen. The pressure is set isotropically to 1 bar with a water compressibility of $4.5x10^{-5}$ bar^{-1}. All simulations are performed with periodic boundary conditions in all

three directions. The cut-off of non-bonded interactions is set to 1.0 nm. Long range electrostatic interactions are treated with the PME method. The length of all hydrogen bonds is fixed. The PEG oligomers are solvated in tip3p water. After an energy minimization a 200 ns NPT-simulation is performed. The first 10 ns are discarded in the analysis.

9.2 Organic Synthesis

9.2.1 Donor Synthesis

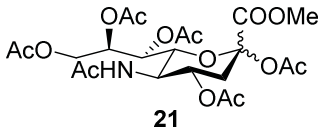
9.2.1.1 Synthesis of Methyl- 5-acetamido-3,5-dideoxy-D-glycero-D-galacto-2-nonulo-pyranosonate 20



According to a literature protocol,[155] Neu5Ac (1.00 g, 3.20 mmol, 1.0 eq.) was suspended in 20 mL methanol microwave tube. TFA (0.10 mL, 1.30 mmol, 0.4 eq.) was added dropwise. Then, the reaction mixture was subjected to microwave irradiation for 30 min at 80°C. The mixture was concentrated under reduced pressure. The product 2 could be obtained as white solid (1.03 g, 3.20 mmol, quant.). Spectral data are in accordance with those reported previously.[155]

^{1}H-NMR (300 MHz, MeOD): δ = 1.84-1.93 (m, 1H, 3-H_{ax}), 2.02 (s, 3H, NAc), 2.24-2.31 (m, 1H, 3-H_{eq}), 3.50-3.61 (m, 2H, 7-H, 9-H_a), 3.70 (m, 1H, 8-H), 3.81 (s, 3H, COOMe), 3.78-3.86 (m, 1H, 9-H_b), 3.91 (d, J = 10.2 Hz, 1H, 5-H), 3.99-4.08 (m, 2H, 4-H, 6-H) ppm; ^{13}C-NMR (75 MHz, MeOD): δ = 21.9, 38.5, 51.9, 53.3, 63.0, 66.5, 68.0, 69.9, 70.1, 95.1, 171.3, 174.6 ppm.

9.2.1.2 Synthesis of Methyl-5-acetamido-2,4,7,8,9-penta-*O*-acetyl-3,5-dideoxy-D-glycero-β-D-galacto-2-nonulopyranosonate 21

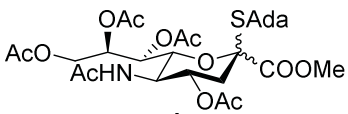


21

According to a literature protocol,[155] DMAP (39.1 mg, 0.32 mmol, 0.05 eq.), pyridine (15.5 mL, 192 mmol, 30.0 eq.) and acetic anhydride (15.1 mL, 160 mmol, 25.0 eq.) were added to Neu5Ac methyl ester **20** (2.17 g, 6.40 mmol, 1.0 eq.) in a microwave tube and the reaction mixture was subjected to microwave irradiation at 70°C for 10 min and was then concentrated under reduced pressure and azeotroped with toluene (3 x 20 mL). The crude product was dissolved in 150 mL DCM and washed with 1 M HCl (3 x 50 mL) and water (3 x 50 mL). The combined organic layers were dried over $MgSO_4$ and the solvent was removed under reduced pressure. Peracetylated Neu5Ac methyl ester **21** was obtained as white solid (2.78 g, 5.21 mmol, 81 %). Spectral data are in accordance with those reported previously.[155]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.88 (s, 3H, NAc), 2.02-2.13 (m, 16H, 5 x OAc, 3-H_{ax}), 2.53 (dd, J = 13.4, 4.9 Hz, 1H, 3-H_{eq}), 3.77 (s, 3H, COOMe), 4.00-4.19 (m, 3H, 6-H, 5-H, 9-H_a), 4.48 (dd, J = 12.4, 2.5 Hz, 1H, 9-H_b), 5.02-5.08 (m, 1H, 8-H), 5.14-5.26 (m, 2H, 4-H, 7-H), 5.34-5.48 (m, 1H, NH) ppm, ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 21.1, 23.4, 36.2, 49.6, 53.3, 62.4, 68.1, 68.6, 71.6, 73.1, 95.7, 97.8, 166.6, 168.5, 168.6, 170.6, 170.6, 170.7, 170.9, 171.3 ppm.

9.2.1.3 Synthesis of Methyl-(2-adamantanyl-5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-2-thio-D-glycero-α/β-D-galacto-non-ulopyranoside)-onate 1

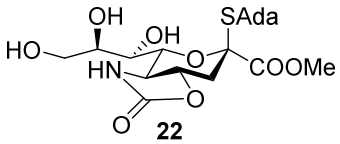


1

According to a literature protocol,[138] peracetylated Neu5Ac methyl ester **21** (6.00 g, 11.3 mmol, 1.0 eq.) was dissolved in 75 mL dry CH_2Cl_2 under inert atmosphere. After the addition of 1-adamantanethiol (2.08 g, 12.4 mmol, 1.1 eq.) at r.t., the solution was cooled to 0°C and $BF_3{\cdot}Et_2O$ (3.42 mL, 27.0 mmol, 2.4 eq.) was added. The reaction was stirred for 1 d at r.t. and was then quenched with 5 mL of triethylamine. The mixture was diluted with CH_2Cl_2, washed with 1 M HCl (3 x 50 mL) and sat. aq. $NaHCO_3$ (3 x 50 mL), dried over $MgSO_4$ and concentrated under reduced pressure. After purification by column chromatography (EtOAc/Hex 1:1 → 9:1), the β-anomer **1β** could be obtained as white solid (3.37 g, 5.25 mmol, 57 %). To obtain the anomeric mixture **1**, both anomers were collected during column chromatography after another batch of the reaction was run. This led to 60% isolated yield of the anomeric mixture **1** with a ratio of α/β = 24:76 according to integration of the 3-H_{eq} proton signals. Spectral data of **1β** are in accordance with those reported previously.[138]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.60-1.66, 1.83-1.89 (mb, 13H, Ada, 3-H_{ax}), 1.96-2.03 (m, 13H, 3 x OAc, Ada), 2.07, 2.13 (2 x s, 6H, 2 x OAc), 2.54 (dd, *J* = 13.5, 4.6 Hz, 1H, 3-H_{eq}), 3.83 (s, 3H, COOMe), 4.02-4.13 (m, 1H, 6-H), 4.21 (dd, *J* = 12.3, 9.1 Hz, 1H, 9-H_a), 5.54 (dd, *J* = 10.5, 2.7 Hz, 1H, 5-H), 5.02 (dd, *J* = 12.3, 1.5 Hz, 1H, 9-H_b), 5.15-5.32 (m, 3H, 8-H, 4-H, 7-H), 5.46-5.47 (m, 1H, NH) ppm; ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.8, 20.9, 21.1, 21.3, 29.9, 36.1, 40.1, 43.7, 49.6, 50.6, 53.0, 63.6, 69.2, 69.5, 72.9, 74.3, 86.3, 170.1, 170.3, 170.5, 171.0, 171.5 ppm.

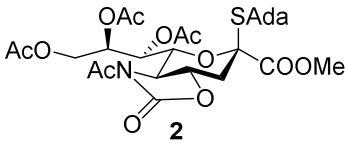
9.2.1.4 Synthesis of Methyl-(2-adamantanyl -5-amino-5-*N*,4-*O*-carbonyl-3,5-dideoxy-2-thio-D-glycero-β-D-galacto-non-2-ulopyranoside)onate 22



According to a literature protocol,[138] the β-2-adamantanyl thiosialoside **1β** (2.50 g, 3.90 mmol, 1.0 eq.) was dissolved in 30 mL dry MeOH under argon. Methanesulfonic acid (0.76 mL, 11.7 mmol, 3.0 eq.) was added and the reaction was stirred at 70°C for 1 d. The reaction was cooled to r.t. and was subsequently quenched with 2 mL of triethylamine and concentrated under reduced pressure. The residue was mixed with $NaHCO_3$ (2.62 g, 31.2 mmol, 8.0 eq.) in 35 mL water and 17.5 mL MeCN. At 0°C 4-nitrophenylchloroformate (1.97 g, 9.75 mmol, 2.5 eq.) dissolved in 17.5 mL MeCN was added dropwise to the stirred solution. The reaction mixture was stirred for 3 h at this temperature and was then extracted with 150 mL EtOAc. The combined organic layers were washed with 1 M HCl (3 x 50 mL) and brine (3 x 50 mL), dried over $MgSO_4$ and the solvent was removed under reduced pressure. After purification by column chromatography with a gradient (EtOAc/MeOH 50:0 → 50:2), **22** could be obtained as white solid (9.92 g, 2.01 mmol, 52 %). Spectral data are in accordance with those reported previously.[138]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.65 (s, 6H, Ada), 1.87-2.04 (m, 9H, Ada), 2.16 (t, J = 12.6 Hz, 1H, 3-H_{ax}), 2.76 (dd, J = 12.6, 3.7 Hz, 1H, 3-H_{eq}), 3.47 (t, J = 10.4 Hz, 1H, 5-H), 3.77-3.94 (m, 6H, 7-H, COOMe, 8-H, 9-H_a), 4.08-4.15 (m, 1H, 9-H_b), 4.47 (d, J = 9.6 Hz, 1H, 6-H), 4.65 (td, J = 11.8, 3.7 Hz, 1H, 4-H) ppm; ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 30.1, 36.3, 43.7, 51.3, 53.6, 60.7, 70.8, 71.3, 71.7, 72.3, 86.2, 160.4, 171.7 ppm.

9.2.1.5 Synthesis of Methyl-(2-adamantanyl 5-acetamido-7,8,9-tri-*O*-acetyl-5-*N*,4-*O*-carbonyl-3,5-dideoxy-2-thio-D-glycero-β-D-galacto-non-2-ulo-pyranoside)onate 2

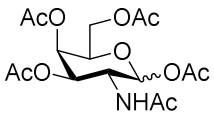


According to a literature protocol,[138] compound **22** (0.81 g, 1.76 mmol, 1.0 eq.) was dissolved in 9.60 mL pyridine and 13.5 mL acetic anhydride and the reaction mixture was stirred at r.t. for 14 h. Then, the reaction was concentrated and azeotroped with toluene (3 x 20 mL). After the addition of 10 mL dry CH_2Cl_2 and DIPEA (3.00 mL, 11.0 mmol, 10.0 eq.), the solution was cooled to 0 °C. Acetylchloride (1.00 mL, 14.1 mmol, 8.0 eq.) was added. The solution was stirred for 1 h at this temperature and another hour at r.t., before pouring it into sat. aq. $NaHCO_3$ solution (50 mL). After extraction with CH_2Cl_2 (3 x 20 mL), the combined organic layers were washed with brine (3 x 20 mL), dried over $MgSO_4$ and solvent was removed under reduced pressure. After purification by column chromatography (EtOAc/Hex 0.5:1.5 → 1:1.5), the product **2** could be obtained as white solid (0.82 g, 1.31 mmol, 74%). Spectral data are in accordance with those reported previously.[138]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.63-1.69 (m, 6H, Ada), 1.82-1.92 (mb, 3H, Ada), 1.94-2.03 (mb, 9H, OAc, Ada), 2.11, 2.12 (2 x s, 6H, OAc), 2.17-2.26 (m, 1H, 3-H_{ax}), 2.48 (s, 3H, NAc), 2.80 (dd, *J* = 12.7, 3.6 Hz, 1H, 3-H_{eq}), 3.66 (dd, *J* = 11.4, 9.4 Hz, 1H, 5-H), 3.83 (s, 3H, CO 2 Me), 4.10-4.18 (m, 1H 9-H_a), 4.65-4.79 (m, 3H, 4-H, 9-H_b, 6-H), 5.28-5.32 (m, 1H, 8-H), 5.70 (t, *J* = 2.7 Hz, 1H, 7-H) ppm; ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 21.0, 21.1, 21.5, 25.5, 30.1, 36.2, 39.0, 43.9, 51.6, 53.3, 60.5, 63.6, 72.7, 73.8, 74.7, 75.4, 85.9, 154.0, 169.6, 170.0, 170.9, 171.4, 172.6 ppm.

9.2.1.6 Synthesis of Galactose Oxazoline 32

9.2.1.6.1 Synthesis of 2,3,4,6-Tetra-*O*-acetyl-D-galactopyranose 31



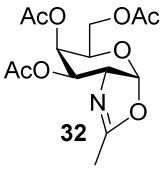
31

D-Galactosamine hydrochloride (2.12 g, 9.80 mmol) was dissolved in 60 mL pyridine and cooled to 0°C. Acetic anhydride (40 mL) was added slowly, the reaction mixture was allowed to warm up to r.t. and was stirred for 18 h at this temperature.

Subsequently, the mixture was azeotroped several times with toluene until a solid residue remains. The product **31** (3.40 g, 8.70 mmol, 89%) was obtained after two rounds of crystallization from methanol. Spectral data are in accordance with those reported previously.[156]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.95 (s, 3H, OAc), 2.02 (s, 3H, OAc), 2.04 (s, 3H, OAc), 2.13 (s, 3H, OAc), 2.17 (s, 3H, OAc), 3.98-4.22 (m, 4H), 5.10 (dd, *J* = 11.3, 3.3 Hz, 1H), 5.37 (d, *J* = 3.0 Hz, 1H), 5.59 (d, *J* = 9.4 Hz, 1H), 5.71 (d, *J* = 8.8 Hz, 1H) ppm.

9.2.1.6.2 Synthesis of Galactose Oxazoline 32

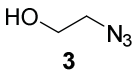


According to a literature protocol,[157] 2,3,4,6-tetra-*O*-acetyl-D-galactopyranose **31** (3.00 g, 7.70 mmol, 1 eq.) was dissolved in 1,2-dichloroethane, and heated to 50°C. TMSOTf (1.65 mL, 9.10 mmol, 1.2 eq.) was added dropwise and the reaction mixture was stirred for 17 h at this temperature. After cooling to r.t., an excess of triethylamine was added and the solvent was removed under reduced pressure. Galactose oxazoline **32** (2.40 g, 7.29 mmol, 95%) was obtained after column chromatography (EtOAc) as a colorless syrup. Spectral data are in accordance with those reported previously.[157]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 2.03-2.08 (m, 9H), 2.11 (s, 3H), 3.95-4.32 (m, 4H), 4.90 (dd, *J* = 7.5, 3.3 Hz, 1H), 5.45 (t, *J* = 2.9 Hz, 1H), 5.99 (d, *J* = 6.8 Hz, 1H) ppm.

9.2.2 Acceptor Synthesis

9.2.2.1 Synthesis of 2-Azidoethanol 3

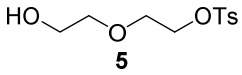


According to a literature protocol,[144] a mixture of 2-chloroethanol (10.0 g, 124 mmol, 1.0 eq.), sodium azide (16.0 g, 373 mmol, 1.8 eq.) and tetra-*n*-butylammonium bromide (4.00 g, 12.4 mmol, 0.1 eq.) was heated at 110°C for 24 h. The mixture was cooled to r.t., diluted with Et_2O and filtered through a plug of celite. The solvent was removed under reduced pressure, the residue was taken up in water and extracted with Et_2O (4 x 100 mL), the organic phase was dried over $MgSO_4$ for 4 h and the solvent was removed under reduced pressure at r.t. The product **3** was obtained as colorless oil (10.6 g, 122 mmol, 98%). Spectral data are in accordance with those reported previously.[144]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 3.28 (t, *J* = 5.1 Hz, 2H), 3.63 (t, *J* = 5.1 Hz, 2H), 3.87 (s, 1H, OH) ppm; ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 53.0, 60.8 ppm.

9.2.2.2 Synthesis of Diethylene Glycol Azide 7

9.2.2.2.1 Synthesis of 2-(2-Hydroxyethoxy)ethyl-*p*-methylbenzenesulfonate 5

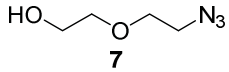


According to a literature protocol,[146] diethylene glycol (5.00 g, 47.1 mmol, 1.0 eq.) was dissolved in 100 mL CH_2Cl_2 and cooled to 0°C. Potassium iodide (1.56 g, 9.42 mmol, 0.2 eq.) and silver(I) oxide (16.4 g, 70.7 mmol, 1.5 eq.) were added and a solution of *p*-toluenesulfonyl chloride (9.88 g, 51.8 mmol, 1.1 eq.) in CH_2Cl_2 was added dropwise to the reaction mixture. The mixture was allowed to warm up to r.t. and was stirred for 17 h, after which it was filtered through a plug of celite with the help of EtOAc. Solvents were removed under reduced pressure and the residue was purified by column chromatography (EtOAc/Hex 1:1 → 100% EtOAc) to obtain **5** (6.65 g, 25.5 mmol, 54%). Spectral data are in accordance with those reported previously.[147]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 2.35 (s, 1H, OH), 2.42 (s, 3H, ArCH_3), 3.46-3.52 (m, 2H), 3.61-3.68 (m, 4H), 4.14-4.19 (m, 2H), 7.32 (d, *J* = 8.2 Hz, 2H, Ar), 7.77

(d, J = 8.3 Hz, 2H, Ar) ppm; ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 21.7, 61.6, 68.6, 69.3, 72.5, 128.0, 129.9, 132.9, 145.0 ppm.

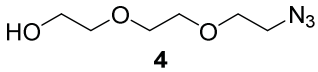
9.2.2.2.2 Synthesis of 2-(2-Azidoethoxy)ethanol 7



According to a literature protocol,[147] tosylated diethylene glycol **5** (6.65 g, 25.5 mmol, 1.0 eq.) was dissolved in 40 mL MeCN and sodium azide (2.49 g, 38.3 mmol, 1.5 eq.) was added. The reaction mixture was equipped with a reflux condenser and heated to 85°C for 20 h. After cooling to r.t., Et_2O was added and the mixture was filtered through a plug of celite. Solvent was removed under reduced pressure to obtain diethylene glycol azide **7** (2.57 g, 19.6 mmol, 77%) as a colorless oil. Spectral data are in accordance with those reported previously.[147]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 2.26 (s, 1H, OH), 3.40 (t, J = 5.0 Hz, 2H, CH_2N_3), 3.58-3.62 (m, 2H), 3.66-3.71 (m, 2H), 3.72-3.76 (m, 2H) ppm.

9.2.2.3 Synthesis of Triethylene Glycol Azide 4

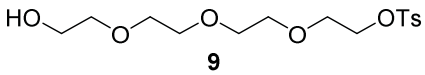


According to a literature protocol,[145] triethylene glycol chloride (3.37 g, 20.0 mmol, 1.0 eq.) was dissolved in 20 mL water. Sodium azide (2.60 g, 40.0 mmol, 1.5 eq.) was added and the reaction was stirred at 75°C for 3 d. Then, the solvent was removed under reduced pressure, the residue was taken up in Et_2O and filtered through a plug of celite. Et_2O was removed under reduced pressure. Triethylene glycole azide **4** was obtained as light yellow oil (3.32 g, 18.9 mmol, 95%) without further purification. Spectral data are in accordance with those reported previously.[145]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 2.79 (br, 1H), 3.32 (t, J = 5.1 Hz, 2H), 3.50-3.55 (m, 2H), 3.57-3.68 (m, 8H) ppm; ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 50.7, 61.7, 70.0, 70.4, 70.7, 72.7 ppm.

9.2.2.4 Synthesis of Tetraethylene Glycol Azide 10

9.2.2.4.1 Synthesis of Tetraethylene Glycol *p*-Methylbenzenesulfonate 9

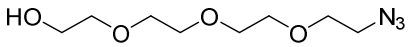


9

According to a literature protocol,[148] NaOH (8.40 g, 210 mmol, 1.5 eq.) dissolved in 35 mL water was added to a solution of tetraethylene glycol (186 mL, 1.40 mol, 10 eq.) in 35 mL THF at 0°C. A solution of *p*-toluenesulfonyl chloride (26.7 g, 140 mmol, 1.0 eq.) in 105 mL THF was added slowly in the cold and the reaction was stirred for 3 h at this temperature. The mixture was then poured into ice water and the aqueous phase was extracted with CH_2Cl_2 (3 x 100 mL). The combined organic layers were washed with water (3 x 100 mL), dried over $MgSO_4$ and concentrated under reduced pressure. The product **9** could be obtained as colorless oil (47.2 g, 135 mmol, 97%). Spectral data are in accordance with those reported previously.[148]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 2.45 (s, 3H), 2.68 (s, 1H), 3.57-3.70 (m, 14H), 4.14 (t, *J* = 4.6 Hz, 2H), 7.31 (d, *J* = 8.1 Hz, 2H), 7.77 (d, *J* = 8.2 Hz) ppm; ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 21.91, 61.96, 68.94, 69.54, 70.58, 70.72, 70.90, 70.98, 72.72, 128.24, 130.11, 133.18, 145.11 ppm.

9.2.2.4.2 Synthesis of Tetraethylene Glycol Azide 10



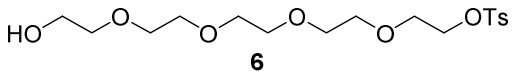
10

According to a literature protocol,[147] tosylated tetraethylene glycol **9** (10.0 g, 28.7 mmol, 1.0 eq.) was dissolved in 60 mL MeCN and sodium azide (2.80 g, 43.1 mmol, 1.5 eq.) was added. The reaction mixture was equipped with a reflux condenser and heated to 85°C for 23 h. After cooling to r.t., water was added and extracted with CH_2Cl_2 (5 x 40 mL). The organic phase was dried over $MgSO_4$ and after removal of the solvent under reduced pressure tetraethylene glycol azide **10** (5.23 g, 23.9 mmol, 83%) as a colorless oil. Spectral data are in accordance with those reported previously.[177]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 3.17-3.24 (m, 2H), 3.31 (s, 1H), 3.36-3.42 (m, 2H), 3.43-3.54 (m, 11H) ppm; ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 50.2, 61.0, 69.5, 69.8, 70.0, 70.1, 70.2, 72.2 ppm.

9.2.2.5 Synthesis of Pentaethylene Glycol Azide 8

9.2.2.5.1 Synthesis of Pentaethylene Glycol *p*-Methylbenzenesulfonate 6

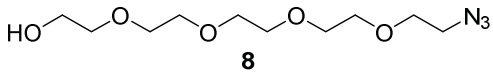


6

According to a literature protocol,[146] pentaethylene glycol (10.0 g, 42. mmol, 1.0 eq.) was dissolved in 200 mL CH_2Cl_2 and cooled to 0°C. Potassium iodide (1.39 g, 8.40 mmol, 0.2 eq.) and silver(I) oxide (14.6 g, 63.0 mmol, 1.5 eq.) were added and a solution of *p*-toluenesulfonyl chloride (8.80 g, 46.2 mmol, 1.1 eq.) in CH_2Cl_2 was added dropwise to the reaction mixture. The mixture was allowed to warm up to r.t. and was stirred for 25 h, after which it was filtered through a plug of celite with the help of EtOAc. Solvents were removed under reduced pressure and the residue was purified by column chromatography (EtOAc) to obtain **6** (7.65 g, 19.5 mmol, 46%). Spectral data are in accordance with those reported previously.[146]

t_R = 2.19 min (short gradient), $C_{17}H_{29}O_8S$ $[M + H]^+$: m/z = 393.16; found: 393.42.

9.2.2.5.2 Synthesis of Pentaethylene Glycol Azide 8



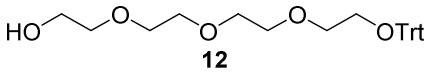
8

According to a literature protocol,[147] tosylated pentaethylene glycol **6** (7.65 g, 19.5 mmol, 1.0 eq.) was dissolved in 50 mL MeCN and sodium azide (1.90 g, 29.3 mmol, 1.5 eq.) was added. The reaction mixture was equipped with a reflux condenser and heated to 85°C for 20 h. After cooling to r.t., water was added and extracted with CH_2Cl_2 (5 x 40 mL). The organic phase was dried over $MgSO_4$ and after removal of the solvent under reduced pressure pentaethylene glycol azide **8** (4.89 g, 18.6 mmol, 95%) as a colorless oil. Spectral data are in accordance with those reported previously.[177]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 3.12 (t, *J* = 5.0 Hz, 2H), 3.31 (t, *J* = 4.6 Hz, 2H), 3.34-3.47 (m, 17H) ppm.

9.2.2.6 Synthesis of Octaethylene Glycol Azide 14

9.2.2.6.1 Synthesis of 2-(2-(2-(2-Trityloxyethoxy)ethoxy)ethoxy)ethanol 12

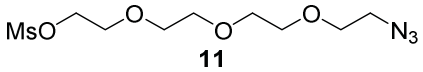


12

According to a literature protocol,[150] tetraethylene glycol (20.0 mL, 150 mmol, 10 eq.), trityl chloride (4.18 g, 15.0 mmol, 1.0 eq.) and pyridine (1.82 mL, 22.5 mmol, 1.5 eq.) were mixed under argon and stirred at 45°C overnight. After cooling to r.t., water (100 mL) was added and the aqueous phase was extracted with toluene (3 x 50 mL). The combined organic layers were washed with 0.3 M HCl (3 x 30 mL) and water (3 x 30 mL), dried over $MgSO_4$ for 1 h, and the solvent was removed under reduced pressure. The product **12** could be obtained as light yellow gel (6.11 g, 14.0 mmol, 93%). Spectral data are in accordance with those reported previously.[150]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 3.26 (t, *J* = 5.4 Hz, 2H), 3.58-3.64 (m, 2H), 3.67-3.73 (m, 12H), 7.22-7.34 (m, 9H), 7.47-7.50 (m, 6H) ppm.

9.2.2.6.2 Synthesis of Azidotetraethylene Glycol Methanesulfonate 11

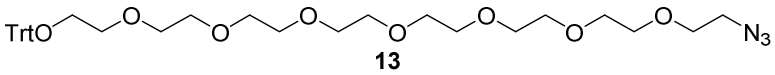


11

According to a literature protocol,[149] to a solution of tetraethylene glycol azide **10** (5.92 g, 27.0 mmol, 1.0 eq.) and triethylamine (7.49 mL, 54.0 mmol, 2.0 eq.) in 100 mL CH_2Cl_2, methanesulfonyl chloride (2.72 mL, 35.1 mmol, 1.3 eq.) was added at 0°C. The reaction was stirred for 10 min in the cold and was then allowed to warm up to r.t. After further stirring for 1 h, the mixture was concentrated under reduced pressure. After purification by column chromatography (EtOAc) the product **11** could be obtained as colorless oil (6.45 g, 21.7 mmol, 80 %). Spectral data are in accordance with those reported previously.[149]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 3.05 (s, 3H), 3.37 (t, *J* = 5.1 Hz, 2H), 3.61-3.68 (m, 10H), 3.72-3.78 (m, 2H), 4.33-4.38 (m, 2H) ppm.

9.2.2.6.3 Synthesis of Trityloctaethylene Glycol Azide 13

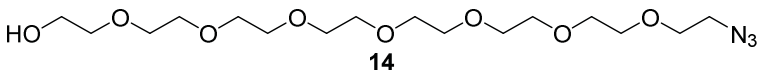


13

Tritylated tetraethylene glycol **12** (1.68 g, 3.85 mmol, 1.1 eq.) was dissolved in 20 mL dry THF and sodium hydride (0.13 g, 5.25 mmol, 1.5 eq.) was added under argon. After stirring for 30 min at r.t., the reaction mixture was cooled 0°C. Azidotetraethylene glycol methanesulfonate **11** (1.04 g, 3.50 mmol, 1.0 eq.) was added in the cold and the reaction was then allowed to warm up to r.t. and stirred for 24 h. Then, the solvent was removed under reduced pressure. After purification by column chromatography (EtOAc), the trityloctaethylene glycol azide **13** was obtained as pale yellow oil (1.68 g, 2.64 mmol, 75%). Spectral data are in accordance with those reported previously.[150]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 3.24 (t, *J* = 5.2 Hz, 2H), 3.39 (t, *J* = 5.2 Hz, 2H), 3.62-3.72 (m, 28H), 7.2.-7.34 (m, 9H), 7.44-7.50 (m, 6H) ppm; ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 51.0, 63.6, 70.4, 70.88, 70.90, 70.93, 70.95, 70.99, 71.00, 71.02, 71.03, 71.1, 86.8, 127.2, 128.7, 129.0, 144.4 ppm.

9.2.2.6.4 Synthesis of Octaethylene Glycol Azide 14



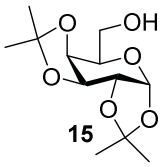
14

Trityloctaethylene glycol azide **13** (6.92 g, 10.9 mmol, 1.0 eq.) was dissolved in 100 mL CH_2Cl_2. TFA (1 mL, 1%) was added and the reaction was stirred at r.t. 60 h. The organic phase was washed with sat. aq. $NaHCO_3$ (3 x 25 mL) and water (3 x 25 mL), dried over $MgSO_4$ and concentrated under reduced pressure. After purification by column chromatography (EtOAc/MeOH 10:0 → 10:1), the product

14 was obtained as pale yellow oil (3.72 g, 9.42 mmol, 87%). Spectral data are in accordance with those reported previously.[149]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 2.77 (s, 1H), 3.36 (t, *J* = 5.2 Hz, 2H), 3.56-3.71 (m, 30H) ppm; ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 8.49, 45.65, 50.64, 61.50, 70.00, 70.23, 70.44, 70.46, 70.48, 70.50, 70.51, 70.57, 70.60, 70.63, 72.55 ppm.

9.2.2.7 Synthesis of 1,2:3,4-Di-*O*-isopropylidene-α-D-galactopyranose 15



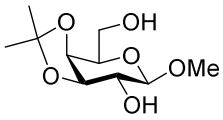
15

According to a literature protocol, [151] zinc(I) chloride (2.84 g, 20.9 mmol, 1.5 eq.) was dissolved in acetone (28 mL, 381 mmol, 27.4 eq.) and galactose (2.50 g, 13.9 mmol, 1 eq.) and 0.1 mL of concentrated sulfuric acid was added to the solution. The reaction was stirred overnight, quenched by the addition of 30 mL sat. aq. $NaHCO_3$, filtered through a plug of celite and acetone was removed under reduced pressure. The residual aqueous phase was extracted with Et_2O (3 x 100 mL), the combined organic phases were dried over $MgSO_4$, and the solvent was removed under reduced pressure. The product **15** (2.14 g, 8.22 mmol, 59%) was obtained as a colorless syrup after column chromatography (EtOAc/Hex 1:3). Spectral data are in accordance with those reported previously.[151]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.32 (s, 6H), 1.44 (s, 3H), 1.52 (s, 3H), 2.30 (sb, 1H), 3.66-3.91 (m, 3H), 4.26 (d, *J* = 7.9 Hz, 1H), 4.29-4.35 (m, 1H), 4.56-4.64 (m, 1H), 5.55 (d, *J* = 5.0 Hz, 1H) ppm.

9.2.2.8 Synthesis of Methyl-2,4,6-tri-*O*-benzyl-β-D-galactopyranosid 19

9.2.2.8.1 Synthesis of Methyl-3,4-di-*O*-isopropylidene-β-D-galactopyranosid 16

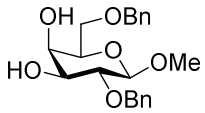


16

According to a literature protocol,[152] methyl-β-D-galactopyranosid (2.00 g, 10.3 mmol, 1.0 eq.) was dissolved in 80 mL dry acetone, 3.67 mL dry DMF, 3.67 mL 2,2-dimethoxypropane (DMP) and 0.19 mL of concentrated sulfuric acid were added. The reaction mixture was stirred for 24 h at r.t. after which it was neutralized by the addition of solid $NaHCO_3$. After filtration the solvent was removed under reduced pressure and the product **16** (1.69 g, 7.21 mmol, 70%) was obtained after column chromatography (EtOAc). Spectral data are in accordance with those reported previously.[178]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.34 (s, 3H), 1.51 (s, 3H), 2.24 (sb, 1H), 2.60 (sb, 1H), 3.49-3.55 (m, 1H), 3.56 (s, 3H, OMe), 3.78-3.92 (m, 2H), 3.95-4.05 (m, 1H), 4.06-4.14 (m, 2H), 4.16 (dd, *J* = 5.6, 1.7 Hz, 1H) ppm.

9.2.2.8.2 Synthesis of Methyl-2,6-di-*O*-benzyl-β-D-galactopyranosid 18



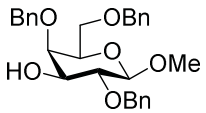
18

According to a literature protocol,[153] **16** (1.00 g, 4.27 mmol, 1.0 eq.) was dissolved in 10 mL dry DMF and cooled to 0°C. benzyl bromide (2.28 mL, 1.44 mmol, 4.5 eq.) and 60% sodium hydride in mineral oil (1.5 g, 34.2 mmol, 8.0 eq.) were added, the reaction was allowed to warm up to r.t. and was stirred for 24 h. The reaction was quenched with an excess of MeOH, poured into ice-cold brine and extracted with EtOAc (3 x 50 mL). The combined organic phases were washed with brine, dried over $MgSO_4$ and the solvent was removed under reduced pressure. The product **17** (1.77 g, 4.27 mmol, quant.) was obtained after column chromatography (Hex/EtOAc 9:1) and deprotected subsequently. Therefore, **17** was dissolved in 70% TFA and stirred at r.t. for 30 min. The solvent was removed by a blowing nitrogen stream and the residue was taken up in EtOAc, washed with sat. aq. $NaHCO_3$ (3 x 50 mL) and brine (3 x 50 mL), dried over $MgSO_4$ and the solvent was removed under reduced pressure. The product **18** (1.42 g, 3.79 mmol, 85%) was obtained

after column chromatography (EtOAc/Hex 1:1). Spectral data are in accordance with those reported previously.[153]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 2.61 (d, *J* = 4.2 Hz, 1H), 2.70 (d, *J* = 3.3 Hz, 1H), 3.49 (dd, *J* = 9.4, 7.5 Hz, 1H), 3.57 (s, 3H), 3.55-3.65 (m, 2H), 3.71-3.85 (m, 2H), 3.96-4.00 (m, 1H), 4.28 (d, *J* = 7.5 Hz, 1H), 4.59 (s, 2H), 4.66 (d, *J* = 11.4 Hz, 1H), 4.94 (d, *J* = 11.5 Hz, 1H), 7.27-7.42 (m, 10H) ppm.

9.2.2.8.3 Synthesis of Methyl-2,4,6-tri-*O*-benzyl-β-D-galactopyranosid 19



19

According to a literature protocol,[154] methyl-2,6-di-*O*-benzyl-β-D-galactopyranosid **18** (375 mg, 1.00 mmol, 1.0 eq.), potassium hydroxide (73.0 mg, 1.30 mmol, 1.3 eq.) and benzyl chloride (5.0 mL, 43.5 mmol, 43.5 eq.) were mixed under argon and heated to 100°C for 3 h. The mixture was cooled to r.t. and taken up in toluene. The solvent was removed under reduced pressure and the product **19** (370 mg, 0.80 mmol, 80%) was obtained after column chromatography (Hex + 2-20% EtOAc in 9% increments). Spectral data are in accordance with those reported previously.[154]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 3.52-4.10 (m, 9H), 4.34 (t, *J* = 7.8 Hz, 1H), 4.54 (q, *J* = 10.6 Hz, 1H), 4.64-4.79 (m, 3H), 4.85 (t, *J* = 11.0 Hz, 1H), 4.99 (t, *J* = 11.5 Hz, 1H), 7.29-7.49 (m, 15H) ppm.

9.2.3 Glycosylation Reactions

9.2.3.1 General Glycosylation Method A (with Donor 2)

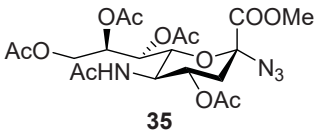
For glycosylation reactions with donor **2** the recent Crich protocol was employed:[140] Briefly, a Schlenk flask was equipped with powdered acid-washed 4 Å molecular sieves, sialic acid donor **2** (94 mg, 0.15 mmol, 1.0 eq.) and the corresponding ethylene glycol azide (0.18 mmol, 1.2 eq.). Dry CH_2Cl_2/MeCN (2:1, 2 mL) was

added, the resulting suspension was stirred for 2 h at r.t. under argon and was then cooled down to –78°C (dry ice/acetone bath) followed by the addition of NIS (42 mg, 0.18 mmol, 2.0 eq.) and TfOH (2 µL, 0.02 mmol, 0.13 eq.). The reaction was stirred for 5 h, quenched with triethylamine (100 µL), filtered through Celite® with the aid of CH_2Cl_2, washed with 20% aq. $Na_2S_2O_3$ solution, dried over $MgSO_4$ and finally concentrated under reduced pressure. The anomeric ratio of the crude product was determined by LC-MS. Purification and separation of anomers was achieved by preparative HPLC.

9.2.3.2 General Glycosylation Method B (with Donors 1 and 1β)

Glycosylations with the anomeric mixture 6 and the pure β-anomer 6β (96 mg, 0.15 mmol, 1.0 eq.) and respective acceptors (0.18 mmol, 1.2 eq.) were performed as described in general glycosylation method A except that the amount of TfOH was increased tenfold (20 µL, 0.2 mmol, 1.3 eq).

9.2.3.3 Synthesis of Methyl-(5-acetamido-2-azido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-glycero-α-D-galacto-non-ulopyranoside)-onate 35



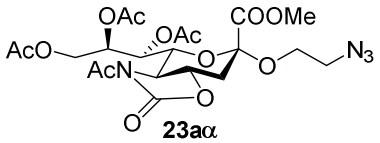
35

According to a literature protocol,[159] peracetylated Neu5Ac methyl ester **21** (750 mg, 1.41 mmol, 1.0 eq.) was dissolved in 9 mL dry CH_2Cl_2 and 1.1 mL acetyl chloride and cooled to 0°C. In a separate flask, 1.1 mL acetyl chloride was cooled to 0°C, dry MeOH was added dropwise and the resulting mixture was transferred to the dissolved peracetylated Neu5Ac methyl ester **21**. The reaction mixture was stirred for 16 h in the cold, volatiles were removed under reduced pressure, the residue was taken up in toluene and evaporated until no smell of HCl was detected anymore. According to another literature protocol,[179] the crude intermediate, tetrabutylammonium hydrogensulfate (479 mg, 1.41 mmol, 1.0 eq.) and sodium azide (458 mg, 7.05 mmol, 5.0 eq.) were dissolved in 7.5 mL CH_2Cl_2 and 7.5 mL sat. aq. $NaHCO_3$

and stirred vigorously at r.t. for 2 h. The reaction mixture was diluted with CH_2Cl_2, layers were separated and the aq. phase was extracted with CH_2Cl_2 (1 x 50 mL). The combined organic phases were washed with sat. aq. $NaHCO_3$, dried over $MgSO_4$ and the solvent was removed under reduced pressure. The product **35** (599 mg, 1.16 mmol, 82% over two steps) was obtained after column chromatography (EtOAc). Spectral data are in accordance with those reported previously.[179]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.79-1.87 (m, 1H, 3-H_{ax}), 1.88 (s, 3H), 2.03 (s, 3H), 2.04 (s, 3H), 2.12 (s, 3H), 2.14 (s, 3H), 2.56 (dd, *J* = 13.0, 4.8 Hz, 1H, 3-H_{eq}), 3.88 (s, 3H, $COOCH_3$), 4.05-4.17 (m, 2H), 4.31-4.40 (m, 1H), 4.98-5.11 (m, 1H), 5.24-5.38 (m, 3H) ppm.

9.2.3.4 Synthesis of Methyl-((2-azido)ethyl-5-acetamido-7,8,9-tri-*O*-acetyl-5-*N*,4-*O*-carbonyl-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyrano-side)-onate 23aα

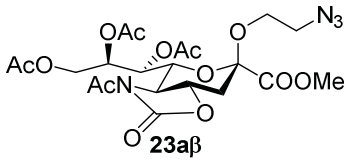


Compound **23aα** was synthesized according to general glycosylation method A and was obtained as an amber syrup (35.0 mg, 64.3 µmol, 43%).

t_R = 6.32 min (long gradient), $[\alpha]_D^{24}$ = -115.3 (c 0.81, $CHCl_3$); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 2.02 (s, 3H, OAc), 2.06-2.17 (m, 7H, 2x OAc, 3-H_{ax}), 2.48 (s, 3H, NAc), 2.87 (dd, *J* = 12.2, 3.5 Hz, 1H, 3-H_{eq}), 3.36 (t, *J* = 4.8 Hz, 2H, CH_2N_3), 3.51-3.60 (m, 1H), 3.74 (dd, *J* = 11.1, 9.4 Hz, 1H), 3.82 (s, 3H, COOMe), 3.96-4.11 (m, 3H), 4.37 (dd, *J* = 12.3, 2.8 Hz, 1H), 4.64 (dd, *J* = 9.4, 1.5 Hz, 1H), 5.38-5.48 (m, 1H), 5.59 (dd, *J* = 8.5, 1.5 Hz, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.9, 21.0, 21.3, 24.8, 36.7, 50.6, 53.2, 59.1, 63.2, 64.9, 68.9, 71.7, 74.9, 75.8, 99.1, 153.7, 168.8, 170.2, 170.4, 170.8, 172.1; ESI-HRMS ($C_{21}H_{28}N_4O_{13}Na$) $[M + Na]^+$:

m/z = 567.1551; found: 567.1554; FTIR-ATR (neat): 2109, 1797, 1744, 1372, 1290, 1267, 1218, 1144, 1126, 1048, 1027 cm^{-1}.

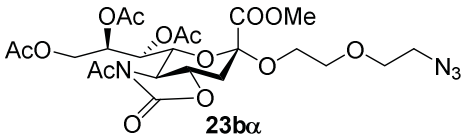
9.2.3.5 Synthesis of Methyl-((2-azido)ethyl-5-acetamido-7,8,9-tri-*O*-acetyl-5-*N*,4-*O*-carbonyl-3,5-dideoxy-D-glycero-β-D-galacto-non-2-ulopyranoside)-onate 23aβ



Compound **23aβ** was synthesized according to general glycosylation method A and was obtained as amber syrup (<4.3 mg, <7.9 µmol, <5%) containing traces of **23aα** as impurity.

t_R = 6.23 min (long gradient); $[\alpha]_D^{24}$ = -10.7 (c 0.36, $CHCl_3$); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 2.03 (s, 3H), 2.04-2.08 (m, 1H, 3-H_{ax}), 2.12 (s, 3H), 2.14 (s, 3H), 2.51 (s, 3H), 2.72 (dd, *J* = 12.1, 3.9 Hz, 1H, 3-H_{eq}), 3.36-3.41 (m, 1H), 3.44-3.59 (m, 2H), 3.64-3.69 (m, 1H), 3.70-3.80 (m, 2H), 3.84 (s, 3H, COOMe), 4.03-4.11 (m, 1H), 4.45 (dd, *J* = 9.4, 2.0 Hz, 1H), 4.57-4.68 (m, 1H), 5.30-5.39 (m, 1H), 5.61-5.66 (m, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.933, 20.937, 21.3, 24.9, 36.3, 50.7, 53.2, 59.5, 62.9, 63.5, 71.5, 73.0, 74.7, 74.75, 98.3, 153.8, 166.7, 170.1, 170.8, 171.2, 172.4; ESI-HRMS ($C_{21}H_{28}N_4O_{13}Na$) $[M + Na]^+$: *m/z* = 567.1551; found: 567.1549; FTIR-ATR (neat): 2109, 1796, 1743, 1458, 1373, 1283, 1217, 1175, 1145, 1125, 1047, 1025 cm^{-1}.

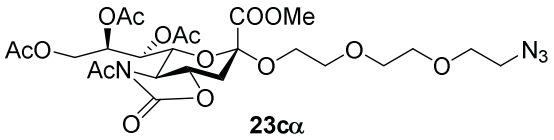
9.2.3.6 Synthesis of Methyl-((2-azidoethoxy)ethyl-5-acetamido-7,8,9-tri-*O*-acetyl-5-*N*,4-*O*-carbonyl-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranoside)-onate 23bα



Compound **23bα** was synthesized according to general glycosylation method A and was obtained as amber syrup (37.4 mg, 63.5 µmol, 42%).

t_R = 6.41 min (long gradient); $[\alpha]_D^{24}$ = +6.2 (c 0.70, $CHCl_3$); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 2.01 (s, 3H, OAc), 2.06-2.18 (m, 7H, 2x OAc, 3-H_{ax}), 2.46 (s, 3H, NAc), 2.87 (dd, *J* = 12.1, 3.5 Hz, 1H, 3-H_{eq}), 3.34-3.40 (m, 2H), 3.42-3.52 (m, 1H), 3.61-3.75 (m, 5H), 3.79 (s, 3H, COOMe), 3.89-4.11 (m, 3H), 4.36 (dd, *J* = 12.3, 2.8 Hz, 1H), 4.59 (dd, *J* = 9.3, 1.7 Hz, 1H), 5.37-5.45 (m, 1H), 5.56 (dd, *J* = 8.1, 1.7 Hz, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.9, 21.0, 21.2, 24.8, 36.5, 50.7, 53.1, 59.1, 63.2, 65.1, 69.0, 70.0, 70.1, 71.9, 75.1, 75.6, 99.3, 153.8, 168.7, 170.1, 170.4, 170.8, 172.1; ESI-HRMS ($C_{23}H_{32}N_4O_{14}Na$) $[M + Na]^+$: *m/z* = 611.1813; found: 611.1813; FTIR-ATR (neat): 2110, 1796, 1742, 1435, 1371, 1290, 1217, 1143, 1125, 1048, 1027 cm^{-1}.

9.2.3.7 Synthesis of Methyl-((2-(2-azidoethoxy)ethoxy)ethyl-5-acetamido-7,8,9-tri-*O*-acetyl-5-*N*,4-*O*-carbonyl-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranoside)-onate 23cα

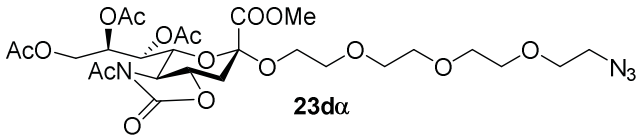


Compound **23cα** was synthesized according to general glycosylation method A and was obtained as amber syrup (45.3 mg, 71.6 µmol, 48%).

t_R = 2.47 min (short gradient); $[\alpha]_D^{24}$ = +4.0 (c 1.08, $CHCl_3$); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 2.01 (s, 3H, OAc), 2.06-2.18 (m, 7H, 2x OAc, 3-H_{ax}), 2.46 (s, 3H, NAc), 2.87 (dd, *J* = 12.1, 3.5 Hz, 1H, 3-H_{eq}), 3.37 (t, *J* = 5.0 Hz, 2H, CH_2N_3), 3.43-3.52 (m, 1H), 3.61-3.74 (m, 9H), 3.78 (s, 3H, COOMe), 3.86-4.07 (m, 3H), 4.34 (dd, *J* = 12.3, 2.9 Hz, 1H), 4.58 (dd, *J* = 9.4, 1.7 Hz, 1H), 5.35-5.43 (m, 1H), 5.55 (dd, *J* = 8.0, 1.6 Hz, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.8, 21.0, 21.2, 24.8, 36.5, 50.7, 53.1, 59.1, 63.1, 65.2, 69.1, 70.1, 70.3, 70.6, 70.7, 71.9, 75.1,

75.6, 99.3, 153.8, 168.8, 170.1, 170.3, 170.7, 172.1; ESI-HRMS ($C_{25}H_{37}N_4O_{15}$) $[M + H]^+$: m/z = 633.2255; found: 633.2259; FTIR-ATR (neat): 2109, 1796, 1743, 1436, 1371, 1289, 1217, 1174, 1143, 1124, 1047, 1027 cm^{-1}.

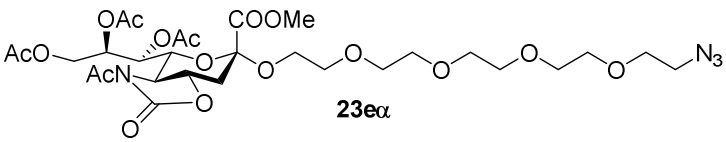
9.2.3.8 Synthesis of Methyl-((2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl-5-acetamido-7,8,9-tri-*O*-acetyl-5-*N*,4-*O*-carbonyl-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranoside)-onate 23dα



Compound **23dα** was synthesized according to general glycosylation method A and was obtained as amber syrup (58.0 mg, 85.7 µmol, 57%).

t_R = 2.44 min (short gradient); $[\alpha]_D^{24}$ = +3.2 (c 0.79, $CHCl_3$); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 2.01 (s, 3H, OAc), 2.06-2.16 (m, 7H, 2x OAc, 3-H_{ax}), 2.46 (s, 3H, NAc), 2.86 (dd, J = 12.1, 3.5 Hz, 1H, 3-H_{eq}), 3.36 (t, J = 5.0 Hz, 2H, CH_2N_3), 3.42-3.51 (m, 1H), 3.60-3.74 (m, 13H), 3.78 (s, 3H, COOMe), 3.85-4.09 (m, 3H), 4.33 (dd, J = 12.3, 2.9 Hz, 1H), 4.58 (dd, J = 9.3, 1.7 Hz, 1H), 5.34-5.43 (m, 1H), 5.55 (dd, J = 8.0, 1.6 Hz, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.8, 21.0, 21.2, 24.8, 35.5, 50.7, 53.1, 59.1, 63.1, 65.2, 69.1, 70.1, 70.14, 70.5, 70.67, 70.69, 70.7, 71.8, 75.1, 75.5, 99.3, 153.8, 168.8, 170.1, 170.3, 170.7, 172.1; ESI-HRMS ($C_{27}H_{41}N_4O_{16}$) $[M + H]^+$: m/z = 677.2518; found: 677.2518; FTIR-ATR (neat): 2108, 1796, 1742, 1641, 1436, 1371, 1289, 1217, 1123, 1047, 1027 cm^{-1}.

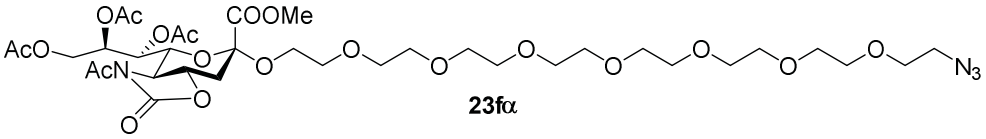
9.2.3.9 Synthesis of Methyl-((2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethoxy)-ethyl-5-acetamido-7,8,9-tri-*O*-acetyl-5-*N*,4-*O*-carbonyl-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranoside)-onate 23eα



Compound **23eα** was synthesized according to general glycosylation method A and was obtained as amber syrup (57.0 mg, 79.1 µmol, 53%).

t_R = 2.45 min (short gradient); $[\alpha]_D^{24}$ = +2.8 (c 0.86, $CHCl_3$); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 2.00 (s, 3H, OAc), 2.05-2.17 (m, 7H, 2x OAc, 3-H_{ax}), 2.45 (s, 3H, NAc), 2.86 (dd, *J* = 12.1, 3.5 Hz, 1H, 3-H_{eq}), 3.36 (t, *J* = 5.0 Hz, 2H, CH_2N_3), 3.41-3.51 (m, 1H), 3.58-3.73 (m, 17H), 3.77 (s, 3H, COOMe), 3.85-4.08 (m, 3H), 4.33 (dd, *J* = 12.3, 2.8 Hz, 1H), 4.57 (dd, *J* = 9.3, 1.7 Hz, 1H), 5.33-5.42 (m, 1H), 5.54 (dd, *J* = 8.0, 1.6 Hz, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.8, 21.0, 21.2, 24.8, 36.5, 50.7, 53.1, 59.1, 63.1, 65.2, 69.1, 70.1, 70.14, 70.5, 70.6, 70.61, 70.7, 70.71, 70.73, 71.8, 75.1, 75.5, 99.3, 153.8, 168.8, 170.1, 170.3, 170.7, 172.1; ESI-HRMS ($C_{29}H_{45}N_4O_{17}$) $[M + H]^+$: *m/z* = 721.2780; found: 721.2788; FTIR-ATR (neat): 2109, 1796, 1743, 1436, 1371, 1289, 1217, 1123, 1047, 1027 cm^{-1}.

9.2.3.10 Synthesis of Methyl-((2-(2-(2-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)-ethoxy)ethoxy)ethoxy)ethoxy)ethyl-5-acetamido-7,8,9-tri-*O*-acetyl-5-*N*,4-*O*-carbonyl-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulo-pyranoside)-onate 23fα

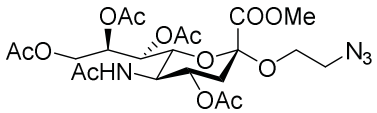


Compound **23fα** was synthesized according to general glycosylation method A and was obtained as amber syrup (74.0 mg, 86.8 µmol, 58%).

t_R = 2.45 min (short gradient); $[\alpha]_D^{24}$ = +4.3 (c 0.77, $CHCl_3$); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.99 (s, 3H, OAc), 2.04-2.15 (m, 7H, 2x OAc, 3-H_{ax}), 2.44 (s, 3H, NAc), 2.85 (dd, *J* = 11.9, 3.6 Hz, 1H, 3-H_{eq}), 3.34 (t, *J* = 5.0 Hz, 2H, CH_2N_3), 3.40-3.49 (m, 1H), 3.57-3.72 (m, 29H), 3.76 (s, 3H, COOMe), 3.83-4.06 (m, 3H), 4.32 (dd, *J* = 12.1, 3.0 Hz, 1H), 4.56 (dd, *J* = 9.3, 1.7 Hz, 1H), 5.32-5.41 (m, 1H), 5.53 (dd, *J* = 8.0, 1.6 Hz, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.8, 21.0, 21.2,

24.7, 36.5, 50.7, 53.0, 59.0, 63.0, 65.1, 69.0, 70.0, 70.1, 70.5, 70.53, 70.55, 70.57, 70.6, 70.65, 70.7, 71.8, 75.0, 75.5, 99.3, 153.7, 168.7, 170.1, 170.2, 170.7, 172.0; ESI-HRMS ($C_{35}H_{57}N_4O_{20}$) $[M + H]^+$: *m/z* = 853.3566; found: 853.3562; FTIR-ATR (neat): 2109, 1796, 1743, 1454, 1372, 1290, 1219, 1099, 1047, 1027 cm^{-1}.

9.2.3.11 Synthesis of Methyl-((2-azido)ethyl-5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranoside)-onate 24aα

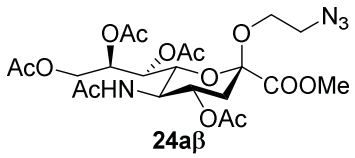


24aα

Compound **24aα** was synthesized according to general glycosylation method B and was obtained as white powder (from donor **1**: 41.7 mg, 74.4 µmol, 50%, from donor **1β**: 41.3 mg, 73.7 µmol, 49%).

t_R = 2.22 min (short gradient); $[\alpha]_D^{22}$ = -26.2 (c 0.84, $CHCl_3$); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.88 (s, 3H), 1.97 (t, *J* = 12.6 Hz, 1H, 3-H_{ax}) 2.02 (s, 3H), 2.03 (s, 3H), 2.13 (s, 3H), 2.14 (s, 3H), 2.62 (dd, *J* = 12.9, 4.6 Hz, 1H, 3-H_{eq}), 3.25-3.33 (m, 1H), 3.34-3.52 (m, 2H), 3.81 (s, 3H, COOMe), 3.93-4.01 (m, 1H), 4.02-4.16 (m, 3H), 4.28 (dd, *J* = 12.4, 2.6 Hz, 1H), 4.81-4.94 (m, 1H), 5.20 (d, *J* = 9.4 Hz, 1H), 5.31 (dd, *J* = 8.8, 1.7 Hz, 1H), 5.36-5.44 (m, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.9, 20.95, 21.0, 21.3, 23.3, 38.1, 49.5, 50.6, 53.0, 62.5, 64.3, 67.3, 68.3, 69.0, 72.6, 98.8, 168.3, 170.2, 170.3, 170.4, 170.8, 171.1; ESI-HRMS ($C_{22}H_{33}N_4O_{13}$) $[M + H]^+$: *m/z* = 561.2044; found: 561.2045; FTIR-ATR (neat): 2108, 1742, 1663, 1544, 1438, 1369, 1304, 1217, 1129, 1037 cm^{-1}.

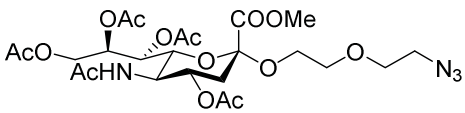
9.2.3.12 Synthesis of Methyl-((2-azido)ethyl-5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-glycero-β-D-galacto-non-2-ulopyranoside)-onate 24aβ



Compound **24aβ** was synthesized according to general glycosylation method B and was obtained as white powder (from donor **1**: 1.6 mg, 2.85 µmol, 2%, from donor **1β**: 3.7 mg, 6.60 µmol, 4%).

t_R = 2.18 min (short gradient); ^{1}H-NMR (600 MHz, $CDCl_3$): δ = 1.90 (s, 3H), 1.91-1.95 (m, 1H, 3-H_{ax}), 2.01 (s, 3H), 2.02 (s, 3H), 2.09 (s, 3H), 2.15 (s, 3H), 2.48 (dd, *J* = 13.0, 4.9 Hz, 1H, 3-H_{eq}), 3.30-3.36 (m, 1H), 3.53-3.63 (m, 2H), 3.66-3.672 (m, 1H), 3.82 (s, 3H, COOMe), 4.02-4.10 (m, 2H), 4.15 (q, *J* = 10.4 Hz, 1H), 4.76 (dd, *J* = 12.4, 2.6 Hz, 1H), 5.24-5.33 (m, 3H), 5.40 (dd, *J* = 4.5, 2.2 Hz, 1H); ^{13}C-NMR (150 MHz, $CDCl_3$): δ = 20.92, 20.94, 21.0, 21.2, 23.2, 37.3, 49.4, 50.7, 53.0, 62.5, 63.3, 68.5, 68.9, 71.7, 71.9, 98.7, 167.2, 170.2, 170.6, 170.8, 170.9, 171.1; ESI-HRMS ($C_{22}H_{33}N_4O_{13}$) $[M + H]^+$: *m/z* = 561.2044; found: 561.2042; FTIR-ATR (neat): 2110, 1743, 1666, 1542, 1439, 1371, 1220, 1167, 1124, 1035 cm^{-1}.

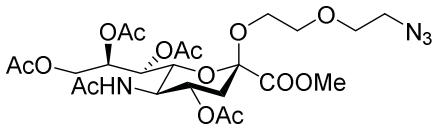
9.2.3.13 Synthesis of Methyl-((2-azidoethoxy)ethyl-5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranoside)-onate 24bα



Compound **24bα** was synthesized according to general glycosylation method B and was obtained as white powder (from donor **1**: 43.2 mg, 71.5 µmol, 48%, from donor **1β**: 67.7 mg, 67.7 µmol, 45%).

t_R = 5.54 min (long gradient); $[\alpha]_D^{22}$ = -20.0 (c 0.82, $CHCl_3$); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.87 (s, 3H), 1.93-2.01 (m, 1H, 3-H_{ax}), 2.02 (s, 3H), 2.03 (s, 3H), 2.13 (s, 3H), 2.14 (s, 3H), 2.61 (dd, J = 12.8, 4.6 Hz, 1H, 3-H_{eq}), 3.33-3.40 (m, 2H), 3.41-3.49 (m, 1H), 3.59-3.73 (m, 4H), 3.79 (s, 3H, COOMe), 3.87-3.97 (m, 1H), 4.01-4.13 (m, 3H), 4.29 (dd, J = 12.4, 2.6 Hz, 1H), 4.77-4.92 (m, 1H), 5.13-5.25 (m, 1H), 5.30 (dd, J = 8.6, 1.4 Hz, 1H), 5.35-5.45 (m, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.9, 20.92, 21.0, 21.2, 23.3, 38.0, 49.5, 50.7, 52.9, 62.5, 64.5, 67.4, 68.5, 69.2, 70.1, 70.2, 72.5, 98.9, 168.3, 170.2, 170.3, 170.4, 170.8, 171.1; ESI-HRMS ($C_{24}H_{37}N_4O_{14}$) $[M + H]^+$: m/z = 605.2306; found: 605.2307; FTIR-ATR (neat): 2111, 1742, 1663, 1546, 1443, 1370, 1303, 1221, 1127, 1038 cm^{-1}.

9.2.3.14 Synthesis of Methyl-((2-azidoethoxy)ethyl-5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-glycero-β-D-galacto-non-2-ulopyranoside)-onate 24bβ



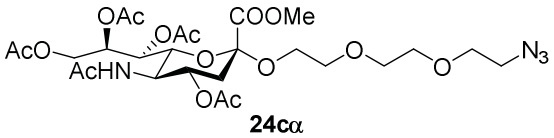
24bβ

Compound **24bβ** was synthesized according to general glycosylation method B and was obtained as amber syrup (from donor **1**: <7.5 mg, <12.4 µmol, <8%, from donor **1β**: <10.0 mg, <16.5 µmol, <11%) containing traces of **24bα** as impurity.

t_R = 5.61 min (long gradient); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.86 (s, 3H), 1.87-1.94 (m, 1H, 3-H_{ax}), 2.01 (s, 3H), 2.03 (s, 3H), 2.06 (s, 3H), 2.14 (s, 3H), 2.44 (dd, J = 12.8, 5.0 Hz, 1H, 3-H_{eq}), 3.33-3.46 (m, 1H), 3.48-3.54 (m, 1H), 3.57-3.77 (m, 5H), 3.80 (s, 3H, COOMe), 3.81-3.88 (m, 1H), 4.04-4.18 (m, 2H), 4.44 (dd, J = 10.6, 2.3 Hz, 1H), 4.93 (dd, J = 12.4, 2.5 Hz, 1H), 5.22-5.36 (m, 2H), 5.40 (t, J = 5.3 Hz, 1H), 5.82 (d, J = 10.4 Hz, 1H); ^{13}C-NMR (150 MHz, $CDCl_3$): δ = 20.9, 21.0, 21.1, 21.15, 23.2, 37.4, 49.1, 51.3, 52.8, 62.8, 63.1, 68.8, 69.3, 69.8, 70.2, 71.2, 72.6, 98.1, 167.6, 170.3, 170.4, 170.8, 170.95, 170.97; ESI-HRMS

($C_{24}H_{37}N_4O_{14}$) $[M + H]^+$: *m/z* = 605.2306; found: 605.2306; FTIR-ATR (neat): 2116, 1743, 1678, 1540, 1458, 1371, 1216, 1126, 1037 cm^{-1}.

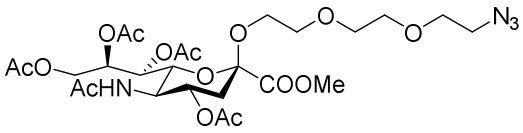
9.2.3.15 Synthesis of Methyl-((2-(2-azidoethoxy)ethoxy)ethyl-5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranoside)-onate 24cα



Compound **24cα** was synthesized according to general glycosylation method B and was obtained as white powder (from donor **1**: 48.2 mg, 74.3 µmol, 50%, from donor **1β**: 37.7 mg, 58.1 µmol, 39%).

t_R = 2.25 min (short gradient); $[\alpha]_D^{23}$ = -22.2 (c 0.82, $CHCl_3$); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.87 (s, 3H), 1.92-2.01 (m, 1H, 3-H_{ax}) 2.02 (s, 3H), 2.03 (s, 3H), 2.12 (s, 3H), 2.13 (s, 3H), 2.60 (dd, *J* = 12.8, 4.6 Hz, 1H, 3-H_{eq}), 3.38 (t, *J* = 5.1 Hz, 2H, CH_2N_3), 3.42-3.49 (m, 1H), 3.60-3.70 (m, 8H), 3.78 (s, 3H, COOMe), 3.85-3.95 (m, 1H), 4.02-4.13 (m, 3H), 4.28 (dd, *J* = 12.4, 2.6 Hz, 1H), 4.78-4.91 (m, 1H), 5.13-5.25 (m, 1H), 5.30 (dd, *J* = 8.5, 1.4 Hz, 1H), 5.34-5.42 (m, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.9, 20.97, 21.0, 21.2, 23.3, 38.1, 49.5, 50.8, 52.9, 62.5, 64.6, 67.4, 68.5, 69.2, 70.1, 70.3, 70.7, 70.8, 72.6, 98.9, 168.4, 170.2, 170.3, 170.4, 170.8, 171.2; ESI-HRMS ($C_{26}H_{41}N_4O_{15}$) $[M + H]^+$: *m/z* = 649.2568; found: 649.2569; FTIR-ATR (neat): 2111, 1742, 1660, 1549, 1444, 1370, 1303, 1222, 1127, 1039 cm^{-1}.

9.2.3.16 Synthesis of Methyl-((2-(2-azidoethoxy)ethoxy)ethyl-5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-glycero-β-D-galacto-non-2-ulopyranoside)-onate 24cβ

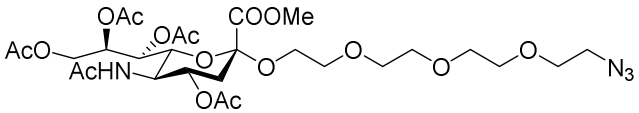


24cβ

Compound **24cβ** was synthesized according to general glycosylation method B and was obtained as amber syrup (from donor **1**: 9.4 mg, 14.5 µmol, 10%, from donor **1β**: 4.2 mg, 6.48 µmol, 4%).

t_R = 2.28 min (short gradient); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.86-1.91 (m, 1H, 3-H_{ax}), 1.89 (s, 3H), 2.01 (s, 3H), 2.03 (s, 3H), 2.05 (s, 3H), 2.14 (s, 3H), 2.44 (dd, *J* = 12.8, 5.0 Hz, 1H, 3-H_{eq}), 3.44-3.50 (m, 2H), 3.51-3.56 (m, 1H), 3.62-3.72 (m, 6H), 3.75-3.83 (m, 5H), 3.84-3.91 (m, 1H), 4.06-4.18 (m, 2H), 4.45 (dd, *J* = 10.6, 2.3 Hz, 1H), 4.89 (dd, *J* = 12.4, 2.5 Hz, 1H), 5.21-5.31 (m, 2H), 5.37-5.43 (m, 1H), 5.94 (d, *J* = 10.3 Hz, 1H); ^{13}C-NMR (150 MHz, $CDCl_3$): δ = 20.95, 20.98, 21.1, 21.2, 23.1, 37.4, 48.9, 50.8, 52.8, 62.8, 63.2, 68.7, 69.1, 69.4, 69.9, 70.2, 70.7, 71.2, 72.4, 98.1, 167.6, 170.4, 170.6, 170.8, 170.9, 171.0; ESI-HRMS ($C_{26}H_{41}N_4O_{15}$) $[M + H]^+$: *m/z* = 649.2568; found: 649.2563; FTIR-ATR (neat): 2106, 1742, 1668, 1542, 1439, 1371, 1223, 1120, 1035 cm^{-1}.

9.2.3.17 Synthesis of Methyl-((2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl-5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranoside)-onate 24dα

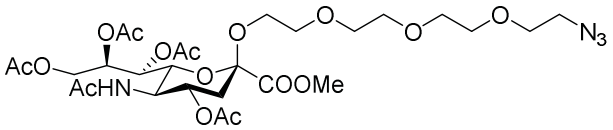


24dα

Compound **24dα** was synthesized according to general glycosylation method B and was obtained as white powder (from donor **1**: 45.4 mg, 65.5 µmol, 44%, from donor **1β**: 46.2 mg, 66.7 µmol, 45%).

t_R = 2.25 min (short gradient); $[\alpha]_D^{23}$ = -18.7 (c 0.83, $CHCl_3$); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.85 (s, 3H), 1.95 (t, *J* = 12.7 Hz, 1H, 3-H_{ax}), 2.00 (s, 3H), 2.02 (s, 3H), 2.11 (s, 3H), 2.12 (s, 3H), 2.60 (dd, *J* = 12.8, 4.6 Hz, 1H, 3-H_{eq}), 3.37 (t, *J* = 5.1 Hz, 2H, CH_2N_3), 3.40-3.47 (m, 1H), 3.56-3.69 (m, 12H), 3.77 (s, 3H, COOMe), 3.83-3.93 (m, 1H), 4.00-4.12 (m, 3H), 4.27 (dd, *J* = 12.4, 2.6 Hz, 1H), 4.75-4.89 (m, 1H), 5.21-5.32 (m, 2H), 5.33-5.40 (m, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.9, 20.93, 21.0, 21.2, 23.3, 38.0, 49.4, 50.7, 52.9, 62.4, 64.6, 67.4, 68.6, 69.2, 70.1, 70.2, 70.6, 70.7, 70.72, 70.8, 72.5, 98.9, 168.4, 170.2, 170.22, 170.4, 170.8, 171.1; ESI-HRMS ($C_{28}H_{45}N_4O_{16}$) $[M + H]^+$: *m/z* = 693.2831; found: 693.2830; FTIR-ATR (neat): 2110, 1744, 1658, 1547, 1447, 1370, 1303, 1222, 1129, 1040 cm^{-1}.

9.2.3.18 Synthesis of Methyl-((2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl-5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-glycero-β-D-galacto-non-2-ulopyranoside)-onate 24dβ



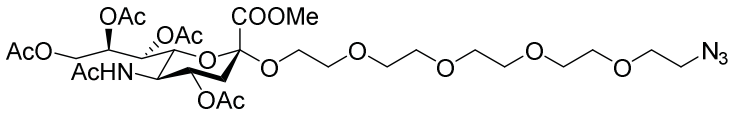
24dβ

Compound **24dβ** was synthesized according to general glycosylation method B and was obtained as amber syrup (from donor **1**: 11.3 mg, 16.3 µmol, 11%, from donor **1β**: 9.9 mg, 14.3 µmol, 10%).

t_R = 2.30 min (short gradient); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.82-1.91 (m, 1H, 3-H_{ax}), 1.87 (s, 3H), 2.00 (s, 3H), 2.02 (s, 3H), 2.04 (s, 3H), 2.14 (s, 3H), 2.42 (dd, *J* = 12.8, 5.0 Hz, 1H, 3-H_{eq}), 3.39 (t, *J* = 5.0 Hz, 2H, CH_2N_3), 3.48-3.56 (m, 1H), 3.61-3.70 (m, 9H), 3.71-3.76 (m, 2H), 3.79 (s, 3H, COOMe), 3.80-3.94 (m, 2H), 4.04-4.19 (m, 2H), 4.58 (dd, *J* = 10.7, 2.2 Hz, 1H), 4.88 (dd, *J* = 12.4, 2.4 Hz, 1H), 5.17-5.34 (m, 2H), 5.34-5.43 (m, 1H), 6.26 (d, *J* = 10.1 Hz, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.96, 21.0, 21.16, 21.17, 23.1, 37.3, 48.6, 50.8, 52.8,

62.8, 63.0, 68.7, 69.3, 69.6, 70.0, 70.1, 70.2, 70.6, 70.7, 70.8, 72.3, 97.9, 167.7, 170.4, 170.6, 170.8, 170.85, 170.9; ESI-HRMS ($C_{28}H_{45}N_4O_{16}$) $[M + H]^+$: m/z = 693.2831; found: 693.2837; FTIR-ATR (neat): 2108, 1740, 1685, 1664, 1543, 1438, 1370, 1222, 1120, 1073, 1033 cm^{-1}.

9.2.3.19 Synthesis of Methyl-((2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethoxy)-ethyl-5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranoside)-onate 24eα

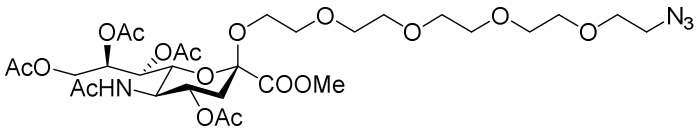


24eα

Compound **24eα** was synthesized according to general glycosylation method B and was obtained as white powder (from donor **1**: 52.5 mg, 71.3 µmol, 48%, from donor **1β**: 50.6 mg, 68.7 µmol, 46%).

t_R = 2.26 min (short gradient); $[\alpha]_D^{23}$ = -18.3 (c 0.83, $CHCl_3$); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.85 (s, 3H), 1.95 (t, J = 12.9 Hz, 1H, 3-H_{ax}), 2.00 (s, 3H), 2.01 (s, 3H), 2.11 (s, 3H), 2.12 (s, 3H), 2.58 (dd, J = 12.8, 4.6 Hz, 1H, 3-H_{eq}), 3.36 (t, J = 5.1 Hz, 2H, CH_2N_3), 3.39-3.47 (m, 1H), 3.54-3.69 (m, 16H), 3.76 (s, 3H, COOMe), 3.83-3.92 (m, 1H), 3.99-4.12 (m, 3H), 4.27 (dd, J = 12.4, 2.5 Hz, 1H), 4.76-4.89 (m, 1H), 5.23-5.31 (m, 2H), 5.32-5.40 (m, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.9, 20.92, 20.94, 21.2, 23.3, 38.0, 49.4, 50.7, 52.8, 62.4, 64.6, 67.3, 68.6, 69.2, 70.1, 70.2, 70.6, 70.61, 60.63, 70.7, 70.73, 70.75, 72.5, 98.9, 168.3, 170.1, 170.2, 170.4, 170.7, 171.1; ESI-HRMS ($C_{30}H_{49}N_4O_{17}$) $[M + H]^+$: m/z = 737.3093; found: 737.3090; FTIR-ATR (neat): 2109, 1743, 1659, 1547, 1449, 1369, 1302, 1220, 1128, 1072, 1039 cm^{-1}.

9.2.3.20 Synthesis of Methyl-((2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethoxy)-ethyl-5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-glycero-β-D-galacto-non-2-ulopyranoside)-onate 24eβ

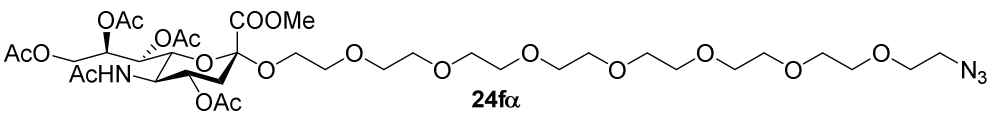


24eβ

Compound **24eβ** was synthesized according to general glycosylation method B and was obtained as amber syrup (from donor **1**: 13.8 mg, 18.7 µmol, 12%, from donor **1β**: 5.6 mg, 7.60 µmol, 5%).

t_R = 2.32 min (short gradient); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.85-1.91 (m, 1H, 3-H_{ax}), 1.87 (s, 3H), 2.00 (s, 3H), 2.02 (s, 3H), 2.04 (s, 3H), 2.14 (s, 3H), 2.43 (dd, *J* = 12.8, 5.0 Hz, 1H, 3-H_{eq}), 3.38 (t, *J* = 5.0 Hz, 2H, CH_2N_3), 3.45-3.57 (m, 1H), 3.60-3.76 (m, 15H), 3.79 (s, 3H, COOMe), 3.80-3.92 (m, 2H), 4.04-4.19 (m, 2H), 4.59 (dd, *J* = 10.6, 2.1 Hz, 1H), 4.88 (dd, *J* = 12.3, 2.4 Hz, 1H), 5.16-5.35 (m, 2H), 5.36-5.41 (m, 1H), 6.35 (d, *J* = 10.3 Hz, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.98, 21.0, 21.16, 21.17, 23.1, 37.3, 48.5, 50.8, 52.8, 62.9, 63.0, 68.8, 69.3, 69.6, 69.7, 70.1, 70.2, 70.4, 70.5, 70.6, 70.8, 70.83, 72.3, 97.9, 167.7, 170.4, 170.7, 170.8, 170.82, 170.84; ESI-HRMS ($C_{30}H_{49}N_4O_{17}$) $[M + H]^+$: *m/z* = 737.3093; found: 737.3093; FTIR-ATR (neat): 2108, 1740, 1665, 1543, 1439, 1370, 1223, 1119, 1073, 1034 cm^{-1}.

9.2.3.21 Synthesis of Methyl-((2-(2-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)-ethoxy)ethoxy)ethoxy)ethoxy)ethyl-5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranoside)-onate 24fα

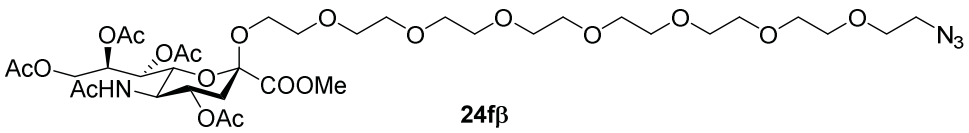


24fα

Compound **24fα** was synthesized according to general glycosylation method B and was obtained as white powder (from donor **1**: 55.7 mg, 64.1 µmol, 43%, from donor **1β**: 54.1 mg, 62.3 µmol, 42%).

t_R = 2.30 min (short gradient); $[\alpha]_D^{23}$ = -17.6 (c 0.89, $CHCl_3$); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.84 (s, 3H), 1.94 (t, J = 12.7 Hz, 1H, 3-H_{ax}), 1.99 (s, 3H), 2.00 (s, 3H), 2.10 (s, 3H), 2.11 (s, 3H), 2.58 (dd, J = 12.8, 4.6 Hz, 1H, 3-H_{eq}), 3.36 (t, J = 5.1 Hz, 2H, CH_2N_3), 3.39-3.47 (m, 1H), 3.51-3.69 (m, 28H), 3.76 (s, 3H, COOMe), 3.81-3.91 (m, 1H), 3.97-4.11 (m, 3H), 4.26 (dd, J = 12.4, 2.5 Hz, 1H), 4.75-4.88 (m, 1H), 5.25-5.31 (m, 2H), 5.32-5.39 (m, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.8, 20.9, 20.94, 21.2, 23.2, 38.0, 49.4, 50.7, 52.8, 62.4, 64.5, 67.3, 68.5, 69.2, 70.1, 70.15, 70.5, 70.6, 70.61, 70.62, 70.67, 70.7, 70.73, 72.5, 98.9, 168.3, 170.1, 170.2, 170.3, 170.7, 171.1; ESI-HRMS ($C_{36}H_{61}N_4O_{20}$) $[M + H]^+$: m/z = 869.3879; found: 869.3880; FTIR-ATR (neat): 2107, 1743, 1665, 1545, 1446, 1369, 1302, 1219, 1121, 1038 cm^{-1}.

9.2.3.22 Synthesis of Methyl-((2-(2-(2-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)-ethoxy)ethoxy)ethoxy)ethoxy)ethyl-5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-glycero-β-D-galacto-non-2-ulopyranoside)-onate 24fβ



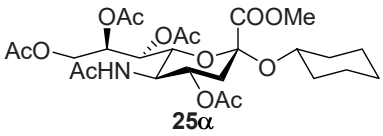
24fβ

Compound **24fβ** was synthesized according to general glycosylation method B and was obtained as amber syrup (from donor **1**: 9.6 mg, 11.0 µmol, 7%, from donor **1β**: 9.0 mg, 10.4 µmol, 7%).

t_R = 2.35 min (short gradient); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.81-1.90 (m, 1H, 3-H_{ax}), 1.87 (s, 3H), 1.99 (s, 3H), 2.02 (s, 3H), 2.04 (s, 3H), 2.14 (s, 3H), 2.43 (dd, J = 12.8, 5.0 Hz, 1H, 3-H_{eq}), 3.38 (t, J = 5.1 Hz, 2H, CH_2N_3), 3.47-3.54 (m,

1H), 3.60-3.72 (m, 27H), 3.79 (s, 3H, COOMe), 3.80-3.92 (m, 2H), 4.04-4.23 (m, 2H), 4.59 (dd, J = 10.7, 2.2 Hz, 1H), 4.88 (dd, J = 12.4, 2.4 Hz, 1H), 5.16-5.33 (m, 2H), 5.35-5.42 (m, 1H), 6.39 (d, J = 10.1 Hz, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.97, 21.0, 21.16, 21.2, 23.1, 37.3, 48.5, 50.8, 52.8, 62.9, 63.0, 68.8, 69.3, 69.6, 69.65, 70.1, 70.2, 70.4, 70.5, 70.51, 70.6, 70.66, 70.67, 70.7, 70.75, 70.78, 70.8, 72.3, 97.9, 167.7, 170.4, 170.8; ESI-HRMS ($C_{36}H_{61}N_4O_{20}$) $[M + H]^+$: m/z = 869.3879; found: 869.3872; FTIR-ATR (neat): 2109, 1741, 1685, 1666, 1544, 1440, 1370, 1223, 1116, 1034 cm^{-1}.

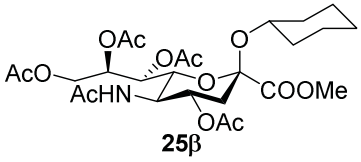
9.2.3.23 Synthesis of Methyl-(2-cyclohexyl-5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-2-D-glycero-α-D-galacto-non-2-ulopyranoside)-onate 25α



Compound **25α** was synthesized according to general glycosylation method B with commercially available cyclohexanol and was obtained as white amorphous solid (from donor **1**: 30.4 mg, 53.0 µmol, 35%). Spectral data are in accordance with those reported previously.[136]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.40-1.77 (m, 9H), 1.87 (s, 3H), 1.90-1.99 (m, 1H, 3-H_{ax}), 2.01 (s, 3H), 2.03 (s, 3H), 2.13 (s, 3H), 2.14 (s, 3H), 2.30-2.44 (m, 1H), 2.58 (dd, J = 12.7, 4.6 Hz, 1H, 3-H_{eq}), 3.64-3.74 (m, 1H), 3.77 (s, 3H), 3.99-4.15 (m, 3H), 4.30 (dd, J = 12.4, 2.4 Hz, 1H), 4.75-4.89 (m, 1H), 5.25-5.40 (m, 3H).

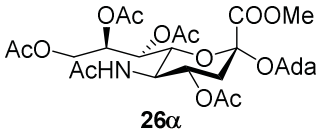
9.2.3.24 Synthesis of Methyl-(2-cyclohexyl-5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-2-D-glycero-β-D-galacto-non-2-ulopyranoside)-onate 25β



Compound **25β** was synthesized according to general glycosylation method B with commercially available cyclohexanol and was obtained as white amorphous solid (from donor **1**: 11.0 mg, 19.2 µmol, 13%). Spectral data are in accordance with those reported previously.[136]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.11-1.75 (m, 9H), 1.87 (s, 3H), 1.95-2.01 (m, 1H, 3-H_{ax}), 2.01 (s, 3H), 2.04 (s, 3H), 2.07 (s, 3H), 2.13 (s, 3H), 2.19-2.29 (m, 1H), 2.52 (dd, *J* = 13.1, 4.9 Hz, 1H, 3-H_{eq}), 3.68-3.75 (m, 1H), 3.77 (s, 3H), 4.02-4.18 (m, 3H), 4.93 (dd, *J* = 12.3, 2.3 Hz, 1H), 5.10-5.17 (m, 1H), 5.18-5.35 (m, 2H), 5.36-5.41 (m, 1H).

9.2.3.25 Synthesis of Methyl-(2-adamantanyl-5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-2-D-glycero-α-D-galacto-non-2-ulopyranoside)-onate 26α



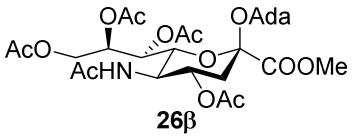
26α

Compound **26α** was synthesized according to general glycosylation method B with commercially available 1-adamantanol and was obtained as white amorphous solid (from donor **1**: 37.0 mg, 59.1 µmol, 39%).

t_R = 2.70 min (short gradient); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.50-1.67 (m, 6H), 1.77-1.85 (m, 6H), 1.88 (s, 3H), 1.93-2.01 (m, 1H, 3-H_{ax}), 1.99 (s, 3H), 2.02 (s, 3H), 2.05-2.10 (m, 3H), 2.12 (s, 3H), 2.15 (s, 3H), 2.51 (dd, *J* = 12.9, 4.5 Hz, 1H, 3-H_{eq}), 3.74 (s, 3H), 4.00 (q, *J* = 10.4 Hz, 1H), 4.10 (dd, *J* = 12.5, 4.7 Hz, 1H), 4.24-4.38 (m, 2H), 4.64-4.82 (m, 1H), 5.22-5.38 (m, 3H); ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.89, 20.94, 21.0, 21.2, 23.4, 31.1, 36.1, 41.0, 43.3, 49.8, 52.5, 62.2, 67.4, 69.1, 72.9, 78.9, 98.0, 170.0, 170.3, 170.4, 170.8, 171.20, 171.22; ESI-HRMS ($C_{30}H_{44}NO_{13}$) $[M + H]^+$: *m/z* = 626.2807; found: 626.2817; FTIR-ATR (neat):

2913, 2856, 2251, 1742, 1665, 1547, 1444, 1369, 1221, 1167, 1126, 1037, 996, 911, 860, 821, 783, 731, 650 cm^{-1}.

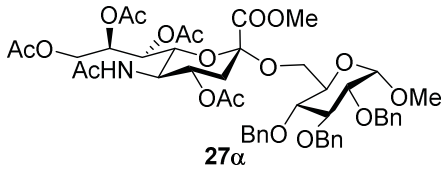
9.2.3.26 Synthesis of Methyl-(2-adamantanyl-5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-2-D-glycero-β-D-galacto-non-2-ulopyranoside)-onate 26β



Compound **26β** was synthesized according to general glycosylation method B with commercially available 1-adamantanol and was obtained as white amorphous solid (from donor **1**: 7.0 mg, 11.2 µmol, 7.5%).

t_R = 2.62 min (short gradient); ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.52-1.65 (m, 6H), 1.66-1.76 (m, 1H, 3-H_{ax}), 1.81-1.89 (m, 6H), 1.87 (s, 3H), 2.00 (s, 3H), 2.01 (s, 3H), 2.07 (s, 3H), 2.08-2.15 (m, 3H), 2.13 (s, 3H), 2.46 (dd, *J* = 12.6, 4.7 Hz, 1H, 3-H_{eq}), 3.79 (s, 3H), 4.06 (q, *J* = 10.3 Hz, 1H), 4.15-4.27 (m, 2H), 4.97 (dd, *J* = 12.3, 1.6 Hz, 1H), 5.16-5.37 (m, 3H), 5.38-5.44 (m, 1H); ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 20.86, 20.96, 21.1, 21.3, 23.3, 31.1, 36.1, 40.6, 43.1, 49.8, 52.8, 53.6, 63.4, 69.1, 69.5, 72.8, 74.0, 78.2, 97.1, 169.9, 170.42, 170.44, 170.6, 171.3, 171.4; ESI-HRMS ($C_{30}H_{44}NO_{13}$) $[M + H]^+$: *m/z* = 626.2807; found: 626.2817; FTIR-ATR (neat): 2912, 2856, 2252, 1743, 1665, 1548, 1437, 1370, 1229, 1153, 1125, 1038, 1010, 947, 913, 734, 650 cm^{-1}.

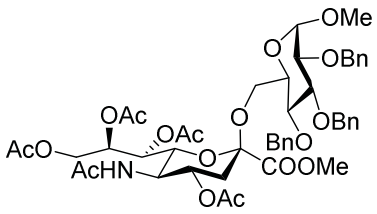
9.2.3.27 Synthesis of Methyl-2,3,4-tri-*O*-benzyl-6-O-[methyl-(5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-α-D-glycero-D-galacto-2-nonulopyranosyl)onate]-α-D-glucopyranosid 27α



Compound **27α** was synthesized according to general glycosylation method B with commercially available methyl-2,3,4-tri-*O*-benzyl-6-O-α-D-glucopyranosid and was obtained as white amorphous solid (from donor **1**: 53.3 mg, 56.8 µmol, 38%). Spectral data are in accordance with those reported previously.[180]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.80 (s, 3H), 1.85 (s, 3H), 1.91-1.99 (m, 1H, 3-H_{ax}), 2.00 (s, 3H), 2.02 (s, 3H), 2.12 (s, 3H), 2.65 (dd, *J* = 12.8, 4.6 Hz, 1H, 3-H_{eq}), 3.36 (s, 3H), 3.38-3.45 (m, 1H), 3.51 (dd, *J* = 9.7, 3.5 Hz, 1H), 3.56-3.64 (m, 1H), 3.71-3.79 (m, 2H), 3.74 (s, 3H), 3.90-4.07 (m, 3H), 4.09 (dd, *J* = 10.8, 1.9 Hz, 1H), 4.23 (dd, *J* = 10.6, 3.8 Hz, 1H), 4.60 (d, *J* = 3.5 Hz, 1H), 4.65 (d, *J* = 12.3 Hz, 1H), 4.72-4.94 (m, 6H), 5.09 (d, *J* = 9.9 Hz, 1H), 5.25 (dd, *J* = 9.4, 1.9 Hz, 1H), 5.29-5.37 (m, 1H), 7.24-7.39 (m, 15H).

9.2.3.28 Synthesis of Methyl-2,3,4-tri-*O*-benzyl-6-O-[methyl-(5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-β-D-glycero-D-galacto-2-nonulopyranosyl)onate]-α-D-glucopyranosid 27β



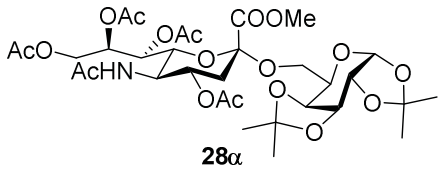
27β

Compound **27β** was synthesized according to general glycosylation method B with commercially available methyl-2,3,4-tri-*O*-benzyl-6-O-α-D-glucopyranosid and was obtained as white amorphous solid (from donor **1**: 8.4 mg, 8.96 µmol, 6%). Spectral data are in accordance with those reported previously.[180]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.84-1.94 (m, 1H, 3-H_{ax}), 1.87 (s, 3H), 1.97 (s, 3H), 2.01 (s, 3H), 2.05 (s, 3H), 2.14 (s, 3H), 2.48 (dd, *J* = 13.0, 5.0 Hz, 1H, 3-H_{eq}), 3.35 (s, 3H), 3.62 (dd, *J* = 9.5, 3.6 Hz, 1H), 3.67-3.86 (m, 4H), 3.72 (s, 3H), 3.93-4.02 (m, 1H), 4.05-4.19 (m, 2H), 4.26 (dd, *J* = 10.5, 2.1 Hz, 1H), 4.67-5.01

(m, 7H), 5.08 (dd, J = 12.3, 2.5 Hz, 1H), 5.11-5.21 (m, 1H), 5.24-5.41 (m, 3H), 7.25-7.43 (m, 15H).

9.2.3.29 Synthesis of 5-Acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-α-D-glycero-D-galacto-2-nonulopyranosylonate-(2→6)-1,2:3,4-di-*O*-isopropylidene-α-D-galactopyranoside 28α

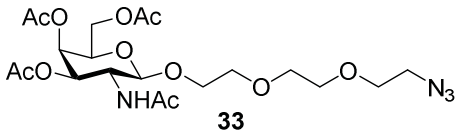


28α

Compound **28α** was synthesized according to general glycosylation method B and was obtained as white amorphous solid (from donor **1**: 44.4 mg, 60.5 µmol, 40%). Spectral data are in accordance with those reported previously.[181]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.32 (s, 3H), 1.32 (s, 3H), 1.41 (s, 3H), 1.53 (s, 3H), 1.88 (s, 3H), 2.02 (s, 3H), 2.04 (s, 3H), 2.12 (s, 3H), 2.13 (s, 3H), 2.61 (dd, J = 12.8, 4.6 Hz, 1H, 3-H_{eq}), 3.56-3.63 (m, 1H), 3.78 (s, 3H), 3.80-3.90 (m, 2H), 3.97-4.39 (m, 7H), 4.59 (dd, J = 7.9, 2.4 Hz, 1H), 4.80-4.94 (m, 1H), 5.13-5.20 (m, 1H), 5.32 (dd, J = 7.9, 1.6 Hz, 1H), 5.36-5.46 (m, 1H), 5.50 (d, J = 5.0 Hz, 1H).

9.2.3.30 Synthesis of 2-(2-(2-Azidoethoxy)ethoxy)ethyl-2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-galactopyranoside 33



33

According to a literature protocol,[158] galactose oxazoline **32** (255 mg, 0.77 mmol, 1.0 eq.) and triethyleneglycol azide **4** (202 mg, 1.16 mmol, 1.5 eq.) were dissolved in 3.5 mL 1,2-dichloroethane under argon and 4Å molecular sieves were added (300 mg). The mixture was stirred at r.t. for 1 h. Subsequently, TMSOTf (70.0 µL, 0.39 mmol, 0.5 eq.) was added dropwise, the reaction mixture was stirred overnight,

afterwards filtered through a plug of celite and the filtrate was washed with sat. aq. $NaHCO_3$ (3 x 20 mL) and water (3 x 20 mL), dried over $MgSO_4$ and the solvent was removed under reduced pressure. The product **33** (686 mg, 1.36 mmol, 88%) was obtained after column chromatography (CH_2Cl_2 + 0.5-6% MeOH in 1% increments) as colorless solid. Spectral data are in accordance with those reported previously.[158]

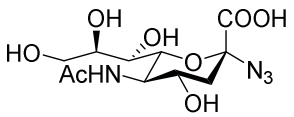
^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.96 (s, 3H), 1.97 (s, 3H), 2.02 (s, 3H), 2.13 (s, 3H), 3.44 (t, *J* = 5.0 Hz, 1H, CH_2N_3), 3.58-3.74 (m, 9H), 3.80-3.93 (m, 3H), 4.07-4.15 (m, 2H), 4.16-4.26 (m, 1H), 4.76 (d, *J* = 8.6 Hz, 1H), 5.04 (dd, *J* = 11.2, 5.0 Hz, 1H), 5.26-5.32 (m, 1H), 6.19 (d, *J* = 9.3 Hz, 1H) ppm.

9.2.4 Deprotection of Carbohydrates

9.2.4.1 General Deprotection Method

Glycosylated sugars were dissolved in MeOH (0.05 mL per mg protected sugar). To the solution was added 1 M aqueous NaOH (0.015 mL per mg protected sugar) and stirred at r.t. for 30 min. Complete conversion was verified by TLC (EtOAc). The mixture was neutralized by the addition of DOWEX®-50WX4 H+ resin, filtered, concentrated under reduced pressure und finally lyophilized to provide the corresponding deprotected azido sugars.

9.2.4.2 Synthesis of [5-Acetamido-3,5-dideoxy-2-azido-D-glycero-α-D-galacto-non-2-ulo-pyranoside]onate 36a



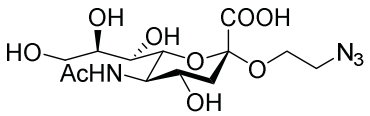
36a

According to the general deprotection method **36a** was obtained as colorless solid (64.0 mg, 0.20 mmol, quant.).

^{1}H-NMR (300 MHz, D_2O): δ = 1.65 (t, *J* = 12.0 Hz, 1H, 3-H_{ax}), 1.98 (s, 3H, NAc), 2.61 (dd, *J* = 12.9, 4.1 Hz, 1H, 3-H_{eq}), 3.50-3.66 (m, 3H), 3.75-3.88 (m, 4H); ^{13}C-

NMR (75 MHz, D_2O): δ = 21.9, 38.1, 51.4, 62.7, 67.3, 68.0, 70.6, 73.7, 89.4, 170.6, 174.9; ESI-HR-MS calculated for $C_{11}H_{19}N_4O_8^+$: *m/z* = 335.1203; found: *m/z* = 335.1197.

9.2.4.3 Synthesis of (2-Azidoethyl)-5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranosidic acid 36b

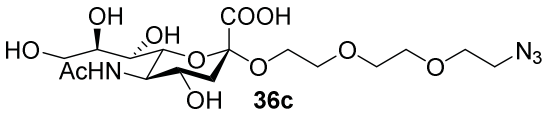


36b

According to the general deprotection method **36b** was obtained as colorless gel (70.5 mg, 0.19 mmol, 85%).

^{1}H-NMR (300 MHz, D_2O): δ = 1.76 (t, *J* = 12.1 Hz, 1H, 3-H_{ax}), 2.00 (s, 3H, NAc), 2.70 (dd, *J* = 12.6, 4.4 Hz, 1H, 3-H_{eq}), 3.46 (t, *J* = 4.9 Hz, 2H, CH_2N_3), 3.61-3.97 (m, 9H); ^{13}C-NMR (75 MHz, D_2O): δ = 21.9, 39.2, 50.2, 51.6, 62.7, 63.0, 67.5, 68.1, 70.9, 72.7, 99.2, 171.5, 174.9; ESI-HR-MS calculated for $C_{13}H_{21}N_4O_9^-$: *m/z* = 377.1314; found: *m/z* = 377.1302.

9.2.4.4 Synthesis of (2-(2-(2-Azidoethoxy)ethoxy)ethyl)-5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranosidic acid 36c



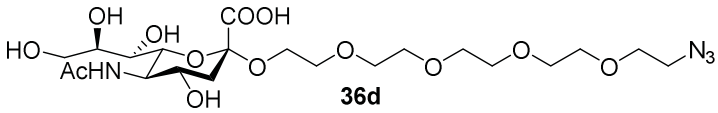
36c

According to the general deprotection method **36c** was obtained as light yellow solid (71.0 mg, 0.15 mmol, 95%).

^{1}H-NMR (300 MHz, D_2O): δ = 1.65 (t, *J* = 12.1 Hz, 1H, 3-H_{ax}), 2.01 (s, 3H, NHAc), 2.71 (dd, *J* = 12.5, 4.5 Hz, 1H, 3-H_{eq}), 3.48 (t, *J* = 4.9 Hz, 2H, CH_2N_3), 3.53 – 3.60 (m, 2H), 3.63 – 3.72 (m, 10H), 3.75-3.92 (m, 5H); ^{13}C-NMR (75 MHz, D_2O): δ = 21.9, 39.3, 50.0, 51.6, 62.7, 63.2, 67.5, 68.1, 69.1, 69.29, 69.32, 69.4,

70.9, 72.6, 99.1, 171.6, 174.8; ESI-HR-MS calculated for $C_{17}H_{29}N_4O_{11}^-$: *m/z* = 465.1838; found: *m/z* = 465.1841.

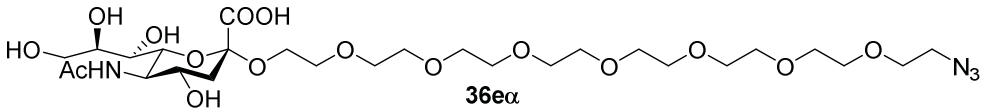
9.2.4.5 Synthesis of (2-(2-(2-(2-(2-Azidoethoxy)ethoxy)ethoxy)ethoxy)ethyl)-5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranosidic acid 36d



According to the general deprotection method **36d** was obtained as yellow oil (123 mg, 0.22 mmol, 79%).

^{1}H-NMR (300 MHz, D_2O): δ = 1.75 (t, *J* = 12.1 Hz, 1H, 3-H_{ax}), 2.00 (s, 3H, NAc), 2.69 (dd, *J* = 12.7, 4.4 Hz, 1H, 3-H_{eq}), 3.48 (t, *J* = 4.9 Hz, 2H, CH_2N_3), 3.56-3.76 (m, 20H), 3.77-3.88 (m, 4H), 3.88-3.97 (m, 1H); ^{13}C-NMR (75 MHz, D_2O): δ = 21.9, 39.3, 50.0, 51.6, 62.7, 63.1, 67.4, 68.1, 69.1, 69.3, 69.38, 69.4, 69.5, 70.9, 71.6, 72.6, 99.1, 171.6, 174.8; ESI-HR-MS calculated for $C_{21}H_{37}N_4O_{13}^-$: *m/z* = 553.2363; found: *m/z* = 553.2358.

9.2.4.6 Synthesis of (2-(2-(2-(2-(2-(2-(2-Azidoethoxy)ethoxy)ethoxy)-ethoxy)ethoxy)ethoxy)ethoxy)ethyl)-5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranosidic acid 36eα

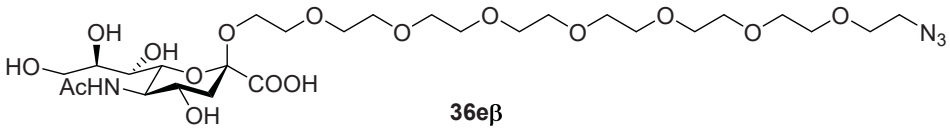


According to the general deprotection method **36eα** was obtained as colorless solid (20.9 mg, 30.4 µmol, 91 %).

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.65 (t, *J* = 12.1 Hz, 1H, 3-H_{ax}), 2.01 (s, 3H, NAc), 2.72 (dd, *J* = 12.4, 4.5 Hz, 1H, 3-H_{eq}), 3.48 (t, *J* =4.8 Hz, 2H, CH_2N_3), 3.52-3.60 (m, 3H), 3.63-3.79 (m, 30H), 3.79-3.94 (m, 4H) ppm; ^{13}C-NMR (75 MHz, $CDCl_3$): δ = 21.9, 40.1, 50.0, 51.7, 62.5, 63.1, 68.1, 69.1, 69.3, 69.37, 69.42, 71.6,

72.4, 100.4, 173.3, 174.9 ppm; ESI-HR-MS calculated for $C_{27}H_{50}N_4O_{16}Na^+$: m/z = 709.3120, found: m/z = 709.3116.

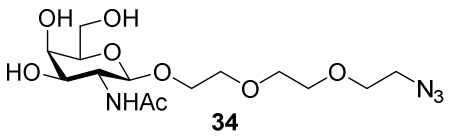
9.2.4.7 Synthesis of (2-(2-(2-(2-(2-(2-(2-(2-Azidoethoxy)ethoxy)ethoxy)-ethoxy)ethoxy)ethoxy)ethoxy)ethyl)-5-acetamido-3,5-dideoxy-D-glycero-β-D-galacto-non-2-ulopyranosidic acid 36eβ



According to the general deprotection method **36eβ** was obtained as colorless solid (10.5 mg, 15.2 µmol, 94 %).

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.74 (t, *J* = 12.2 Hz, 1H, 3-H_{ax}), 2.03 (s, 3H, NAc), 2.41 (dd, *J* = 13.1, 4.9 Hz, 1H, 3-H_{eq}), 3.48 (t, *J* = 4.9 Hz, 2H, CH_2N_3), 3.52-3.58 (m, 1H), 3.64-3.75 (m, 30H), 3.77-3.95 (m, 5H), 4.04-4.14 (m, 1H) ppm; ^{13}C-NMR (150 MHz, $CDCl_3$): δ = 22.1, 39.4, 50.1, 51.8, 62.1, 63.4, 66.5, 68.0, 69.2, 69.47, 69.51, 69.52, 69.56, 69.57, 69.58, 69.59, 69.8, 70.5, 99.1, 172.8, 174.8 ppm; ESI-HR-MS calculated for $C_{27}H_{50}N_4O_{16}H^+$: m/z = 687.3295, found: m/z = 687.3291.

9.2.4.8 Synthesis of 2-[2-(2-Azidoethoxy)ethoxy]ethyl-2-*N*-acetylamido-2-deoxy-β-D-galacto-pyranoside 34



According to the general deprotection method **34** was obtained as pale yellow gel (515 mg, 1.36 mmol, quant.).

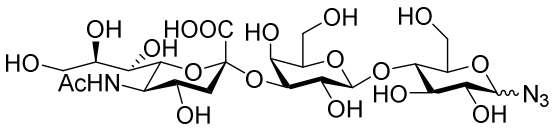
^{1}H-NMR (300 MHz, D_2O): δ = 2.02 (s, 3H, NHAc), 3.48 (t, *J* = 5.1 Hz, 2H, CH_2N_3), 3.62-3.79 (m, 13H), 3.83- 3.93 (m, 2H), 3.95-4.04 (m, 1H), 4.48 (d, *J* = 8.4 Hz, 1H, 1-H); ^{13}C-NMR (75 MHz, D_2O): δ = 22.1, 50.0, 52.2, 60.8, 67.7,

68.7, 69.1, 69.4, 69.5, 69.6, 71.0, 75.0, 101.4, 174.6; ESI-HR-MS calculated for $C_{14}H_{26}N_4O_8Na^+$: *m/z* = 401.1648; found: *m/z* = 401.1649.

9.2.5 Azidosialyllactose Derivatives

According to a literature protocol,[160] commercially available 2,3- and 2,6-sialyllactose (10.0 mg, 15.0 µmol, 1.0 eq.) was dissolved in 250 µL D_2O and 2-chloro-1,3-dimethylimidazolinium chloride (DMC, 25.4 mg, 150 µmol, 10 eq.), NaN_3 (24.8 mg, 381 µmol, 25 eq.) and NEt_3 (31.2 µL, 225 µmol, 15 eq.) were added and the reaction was stirred at r.t. for 5 h and lyophilized afterwards. The solid residue was acetylated with 150 µL acetic anhydride and 450 µL pyridine at r.t. over night. The mixture was concentrated under reduced pressure and the residue was purified by reversed phase *SepPak*® cartridges (*Waters*, 2 g C18 material, wetted with H_2O:MeCN 1:1, equilibrated with pure H_2O, loaded reaction crude, washed with 10 mL H_2O, eluted product with 20 mL MeCN). The MeCN fractions were pooled, concentrated under reduced pressure and lyophilized. To the residue was added LiOH (2 mg), the mixture was dissolved in 200 µL of an H_2O.MeOH 1:1 mixture and the reaction was stirred at r.t. over night. DOWEX®-50WX4 H+ resin was added to neutralize the mixture, filtered off, rinsed with H_2O and the filtrate was lyophilized.

9.2.5.1 2,3-Sialyllactose Azide 37

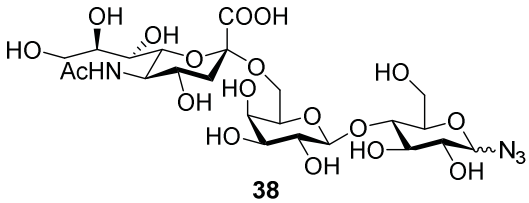


37

According to the general method **37** was obtained as colorless solid (9.50 mg, 14.4 µmol, 96%). Spectral data are in accordance with those reported previously.[160]

ESI-HR-MS calculated for $C_{23}H_{38}N_4O_{18}K^+$: m/z = 697.1818, found: m/z = 697.1817.

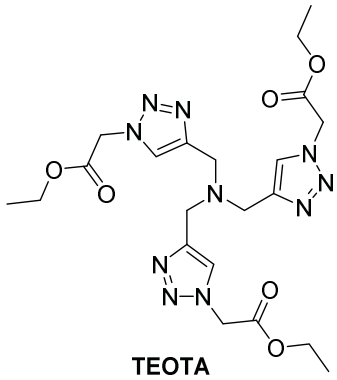
9.2.5.2 2,6-Sialyllactose Azide 38



According to the general method **38** was obtained as colorless solid (9.87 mg, 15.0 µmol, quant.). Spectral data are in accordance with those reported previously.[160]

ESI-HR-MS calculated for $C_{27}H_{38}N_4O_{18}H^+$: m/z = 659.2259, found: m/z = 659.2254.

9.2.6 CuAAC Ligand (TEOTA)



According to a literature protocol,[95] tripropargylamine (0.30 g, 2.29 mmol, 1.0 eq.) and ethyl azidoacetate (0.89 g, 6.86 mmol, 3 eq.) were dissolved in dry MeCN under argon. DIPEA (1.20 mL, 6.86 mmol, 3 eq.) and $[Cu(MeCN)_4]PF_6$ (31.0 mg, 82.4 µmol, 0.036 eq.) were added and the reaction mixture was stirred at r.t. for 4 d. The solvent was removed under reduced pressure, the residue was taken up in CH_2Cl_2 and washed with water. The organic phase was dried over $MgSO_4$ and solvent was removed under reduced pressure. After column chromatography (CH_2Cl_2 + 0-10% MeOH in 1% increments) TEOTA was obtained as pale yellow solid (915 mg, 1.76 mmol, 77%). Spectral data are in accordance with those reported previously.[95]

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.26 (t, *J* = 7.1 Hz, 9H, $OCH_2C\mathit{H_3}$), 3.78 (sb, 6H), 4.22 (q, *J* = 7.1 Hz, 6H, $OC\mathit{H_2}CH_3$), 5.14 (s, 6H), 7.84 (s, 3H) ppm.

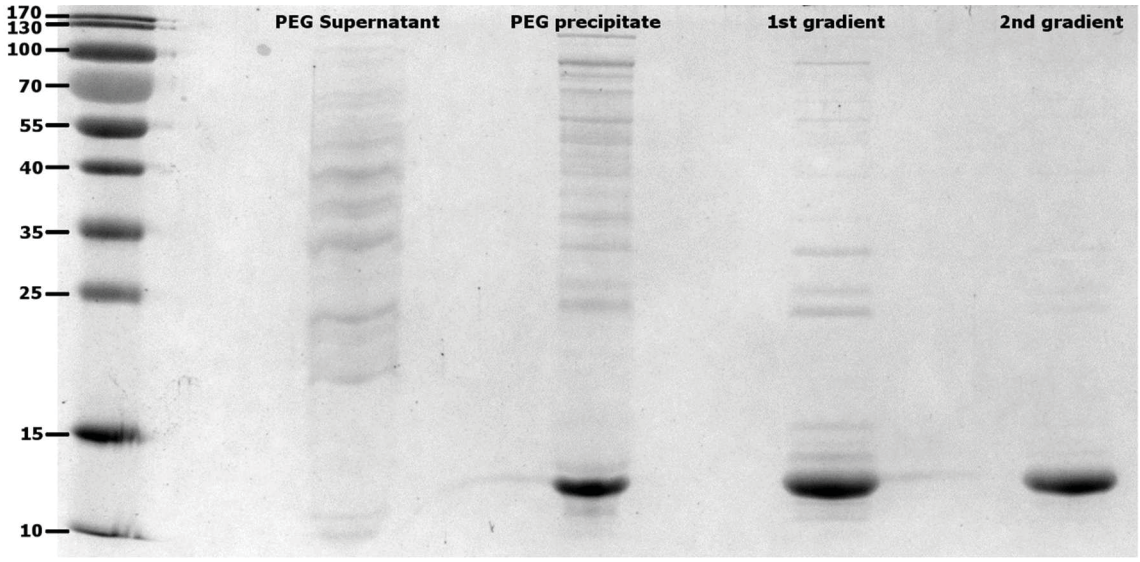
9.3 Qβ[Hpg] Expression and Purification

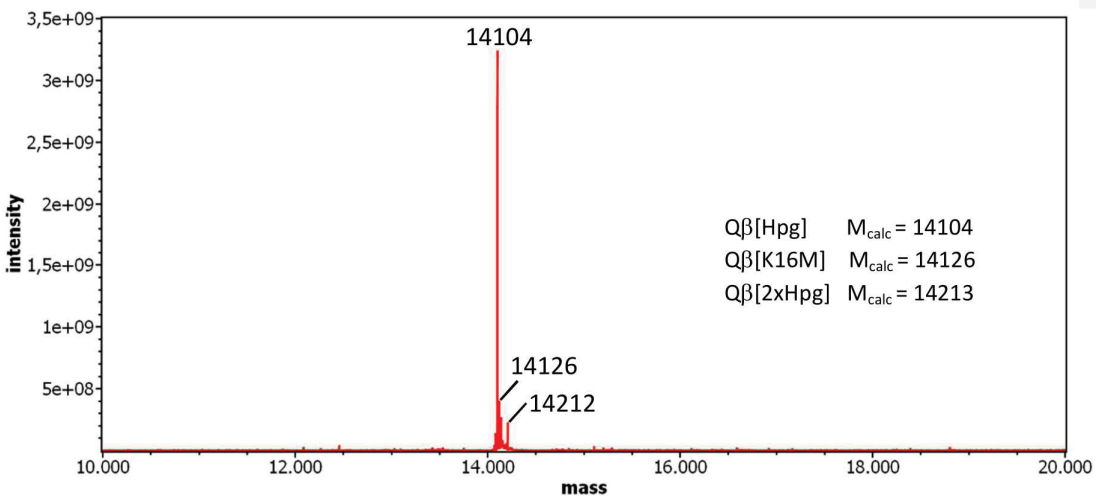
Auxotrophic expression of **Qβ[Hpg]** was performed by Saba Nojoumi (TU Berlin). As expression media *M9 Minimal Medium* and *New Minimal Medium* were used. A preculture of *E. coli* B834(pQE60_QβK16M, pREP4) was prepared in 5 mL minimal media supplemented with Kanamycin50 and Ampicillin100 which was subsequently incubated at 37°C, 200 rpm for 8 h. Minimal medium (625 mL) substituted with Kanamycin50, Ampicillin100 and 3.35 mL methionine (50 mM) was inoculated with 6 mL of the preculture. The resulting main culture was incubated at 30°C, 180 rpm overnight (18 h). Subsequently, the culture was washed twice with 35 mL minimal media supplemented with Kanamycin50 and Ampicillin100 lacking methionine (400 rpm, 4°C, 20 min). Washed culture pellets were resuspended in minimal media supplemented with Kanamycin50, Ampicillin100 and the resulting suspension was brought to 625 mL in preheated minimal medium supplemented with Kanamycin50, Ampicillin100 at 30°C. The unnatural amino acid Hpg was added to the culture (50 mg/L). The expression culture was incubated at 37°C, 180 rpm for 30 min subsequently the culture was incubated at 30°C, 180 rpm for 1 h. Expression of Qβ[K16M] was induced with IPTG (1 mM) and was incubated at 30°C, 180 rpm, overnight (18 h). Cells were harvested (8000 x g, 4°C, 20 min) and stored at -20°C until purification.

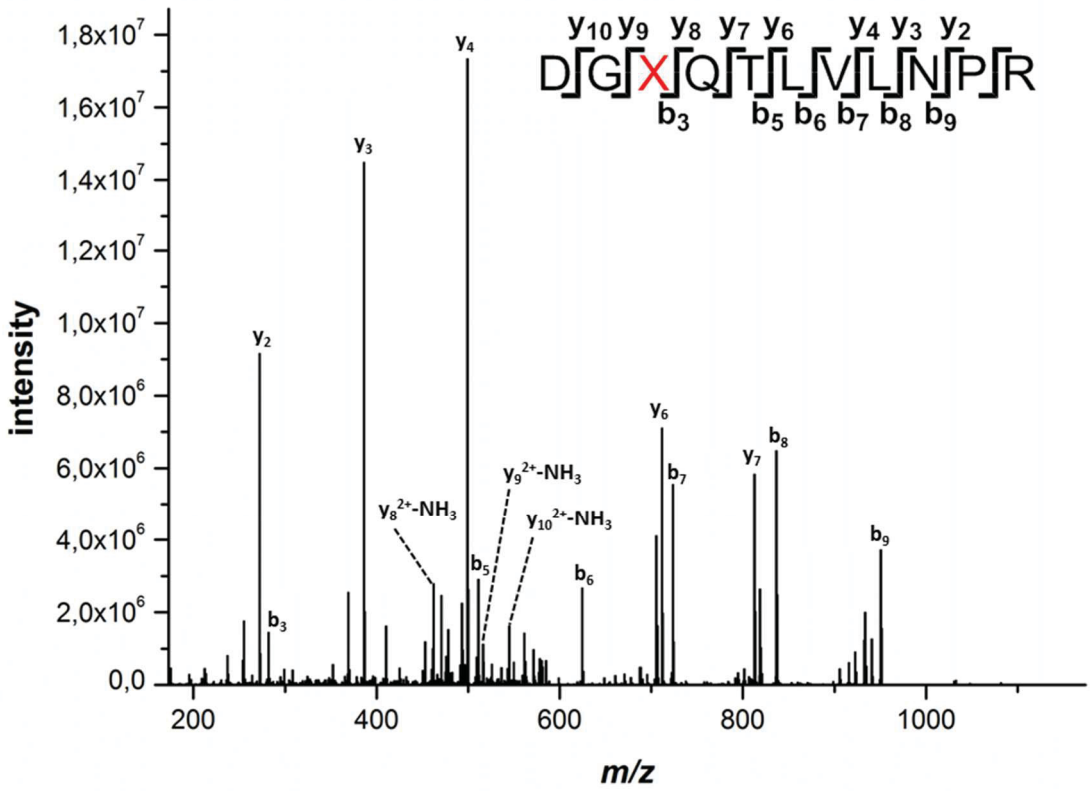
Thawed cell pellets were resuspended in lysis buffer (100 mM potassium phosphate, pH 7.3 ≙ KPi) and lyzed by ten passes through an *LM10 Microfluidizer* (*Microfluidics*, Westwood, Massachusetts) at 14,000 psi. The resulting slurry was incubated with DNaseI in the presence of 10 mM $MgCl_2$ at room temperature for 30 min and subsequently centrifuged to remove cell debris (50,000 x g, 20 min, 4°C). To the supernatant 10% (w/v) of PEG8000 and NaCl (100 mM final concentration) was added and incubated over night at 4°C. The precipitate was collected by centrifugation (10,000 rpm, 15 min, 4°C) and dissolved in KPi buffer. Extraction with $CHCl_3$/nBuOH (1:1) was carried out and phase separation as well as removal of

insoluble material was achieved by centrifugation (10,000 rpm, 15 min, 4°C). The resulting Qβ capsid containing supernatant was stored at 4°C. Further purification was achieved by sucrose gradient (10-40%) ultracentrifugation in an *Optima LE-80K* ultracentrifuge (*Beckman Coulter*) with an *SW 40 Ti* rotor (125,000 x g, 4.5 h, 4°C). Qβ containing fractions were identified by SDS-PAGE, pooled, concentrated and washed in 30 kDa cut-off centrifugal filter units. Analysis of the purification progress and successful incorporation of Hpg at the desired position was carried out by SDS-PAGE, intact protein MS and tryptic digest followed by LC-MS/MS.

9.3.1 Analysis of Qβ[Hpg]





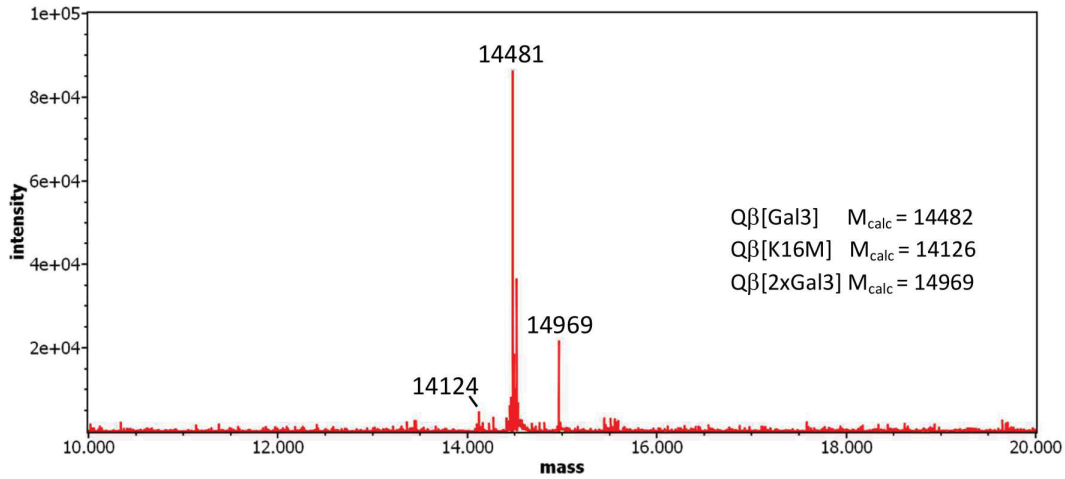


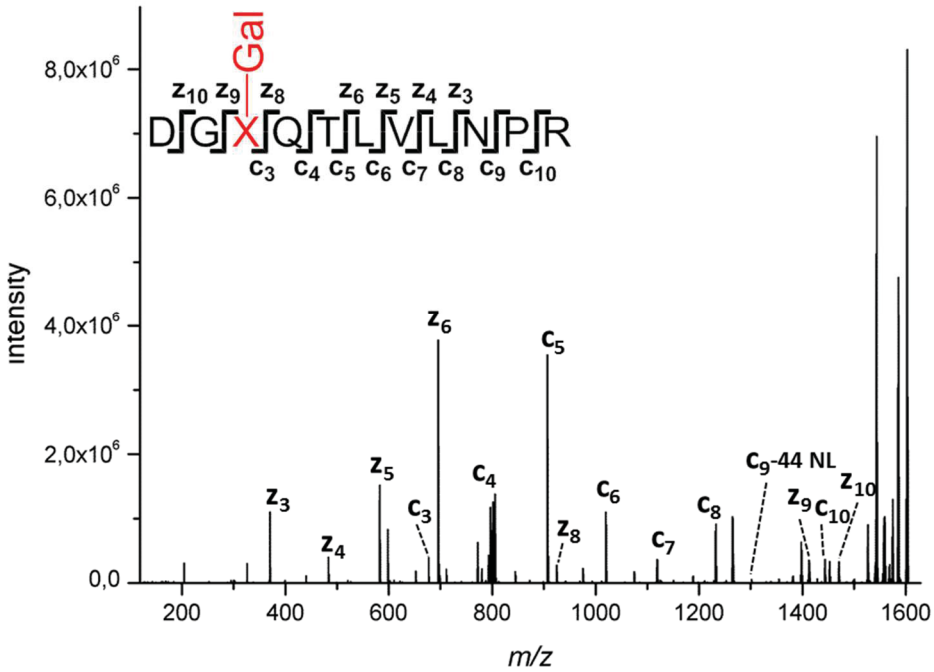
9.4 Bioconjugation

Conjugation reactions were performed according to a published protocol.[83] Briefely, to 500 µL of a 1.2 µM **Qß[Hpg]** capsid solution (213.5 µM coat protein ≙ Hpg concentration) were added 90 eq. of the respective azido sugar, dissolved in 250 µL of KPi. Stock solutions of CuBr (99.999% trace metal basis, 10 mg/mL) and TEOTA (100 mg/mL) in MeCN were prepared. Subsequently, 187.6 µL of the CuBr solution and 72.6 µL of the TEOTA solution were mixed and 200 µL of the resulting Cu/ligand mix were transferred to the protein/sugar mixture. Reaction proceeded at 25°C at 800 rpm on a thermoshaker for 30 min. Another 100 µL portion of the Cu/ligand mix were added to the reaction and after additional 1.5 h the precipitate was removed by centrifugation and large excess of low molecular weight compounds was reduced by spin filtration.

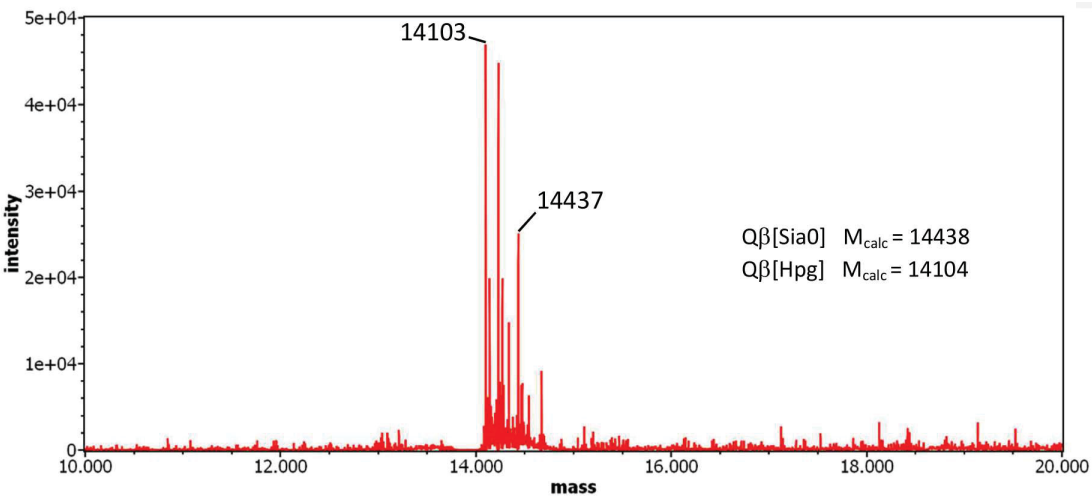
Purification of the reaction crude was again achieved by sucrose gradient ultracentrifugation (125,000 x g, 4.5 h, 4°C in a SW40 Ti rotor, 10-40% gradient). The relevant fractions were identified by SDS-PAGE, pooled, concentrated and washed in 30 kDa cut-off centrifugal filter units. Further polishing steps included dialysis against EDTA-containing potassium phosphate buffer (5 mM EDTA) and SEC on an *Äkta* FPLC system equipped with a *Superose6* column.

9.4.1 Analysis of Qβ[Gal3]

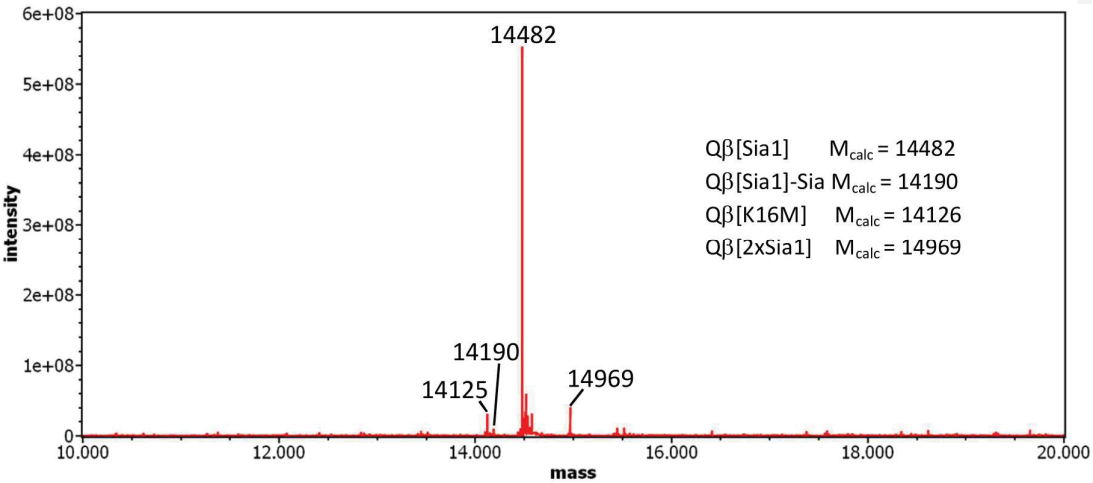


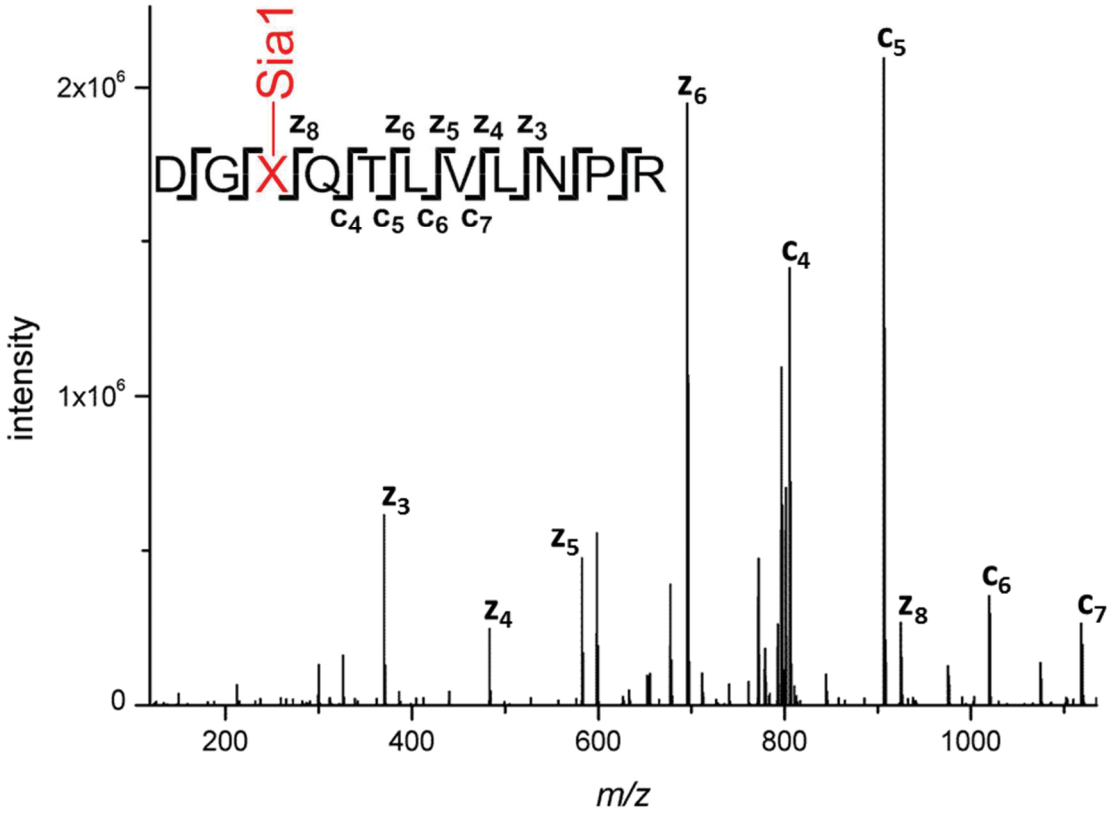


9.4.2 Analysis of Qβ[Sia0]

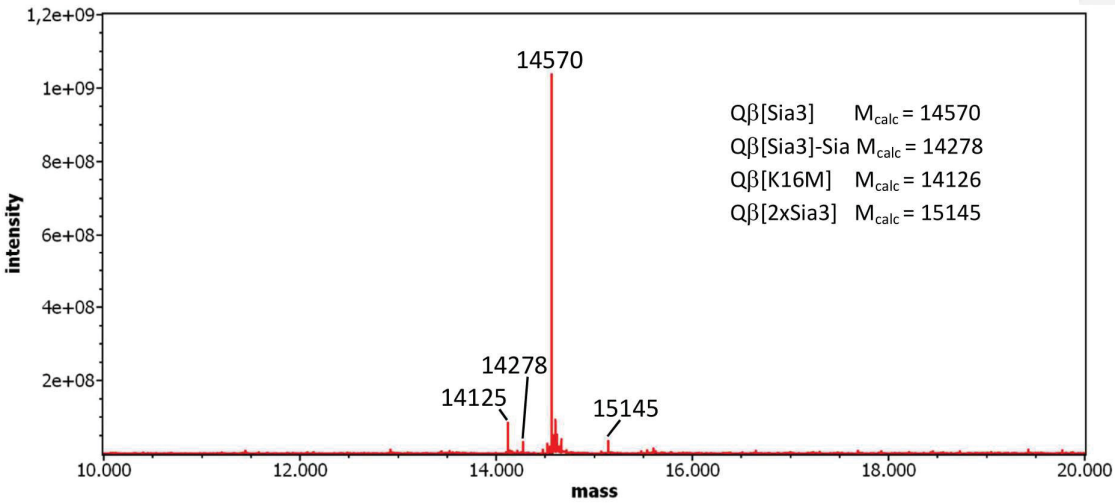


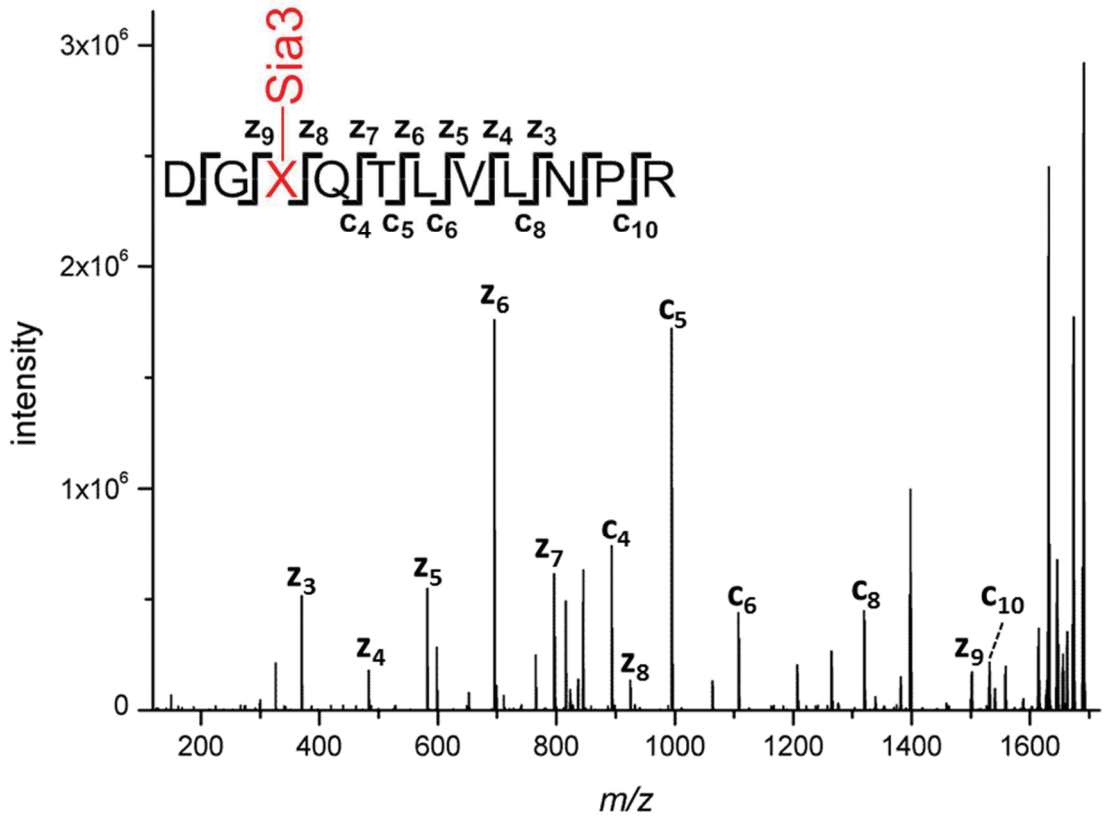
9.4.3 Analysis of Qβ[Sia1]



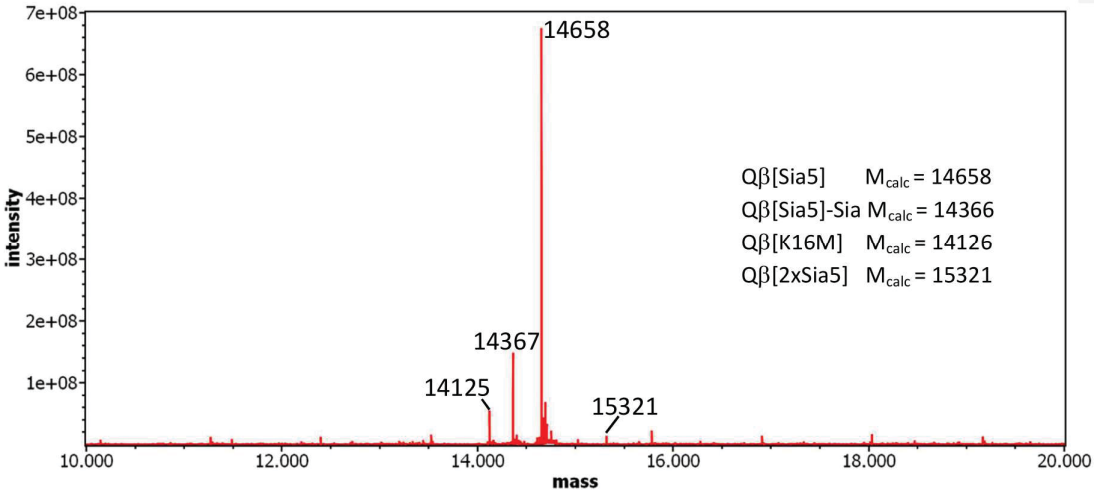


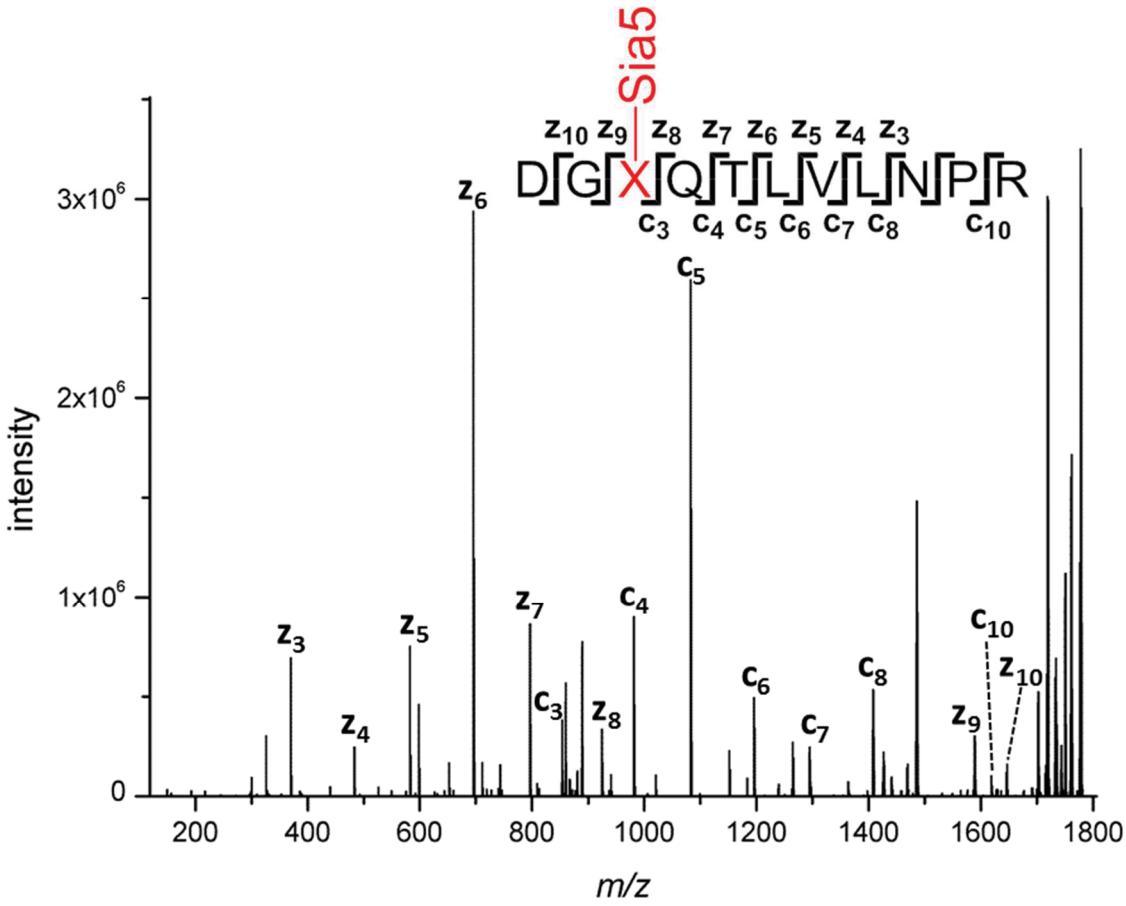
9.4.4 Analysis of Qβ[Sia3]



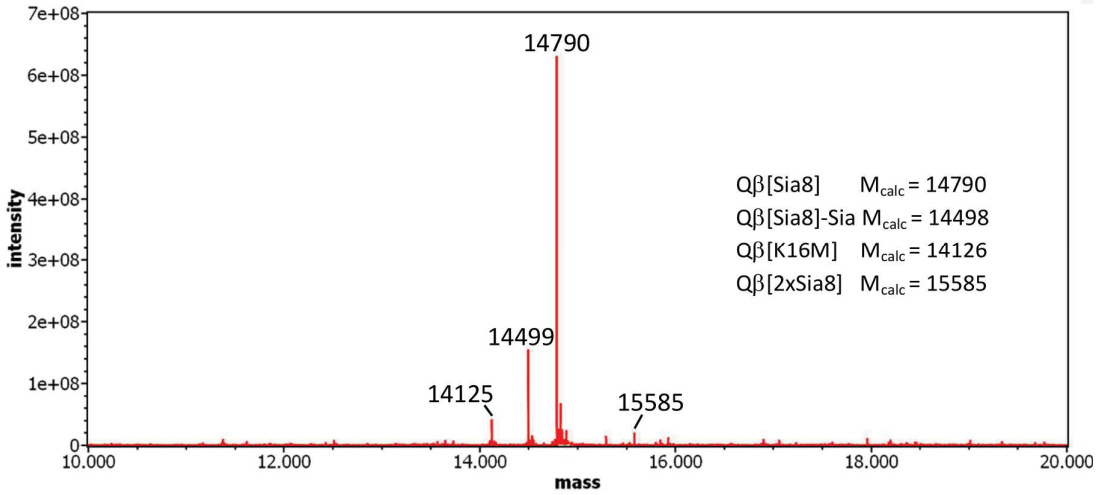


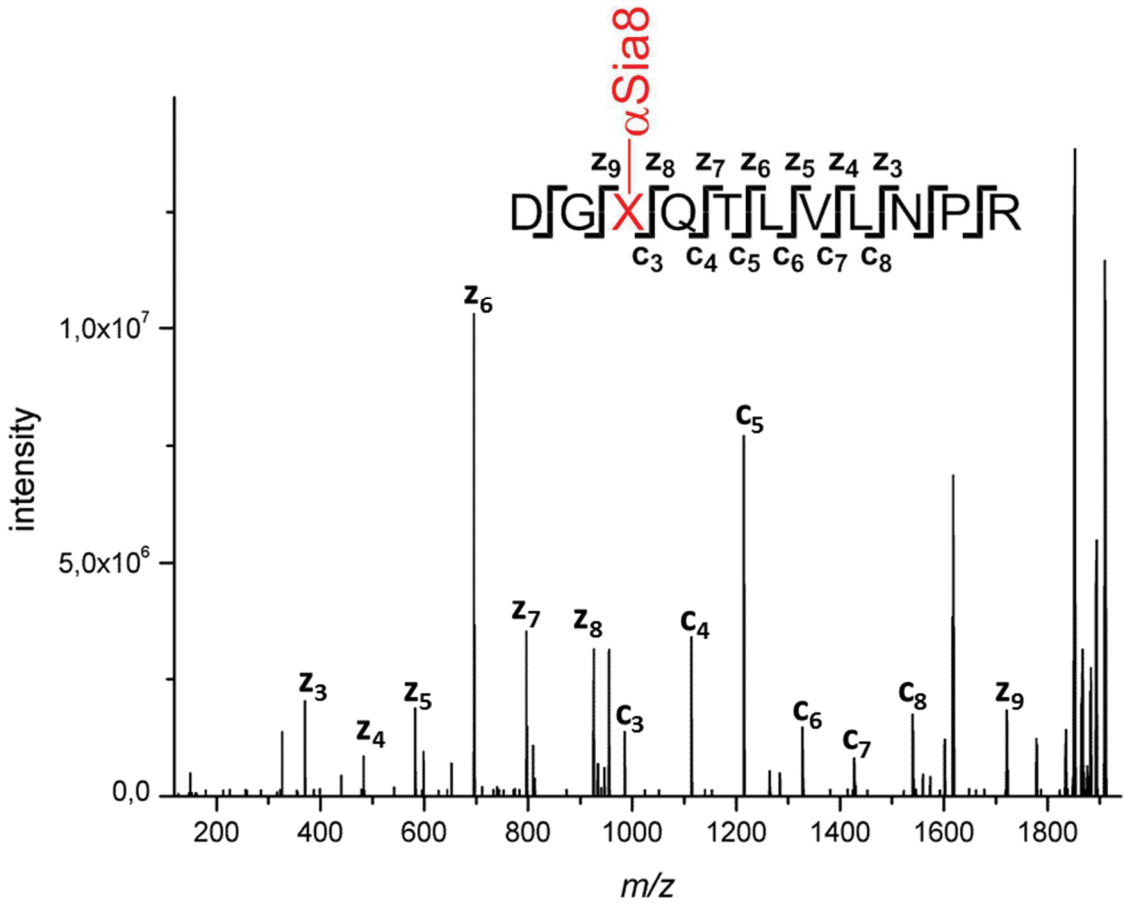
9.4.5 Analysis of Qβ[Sia5]



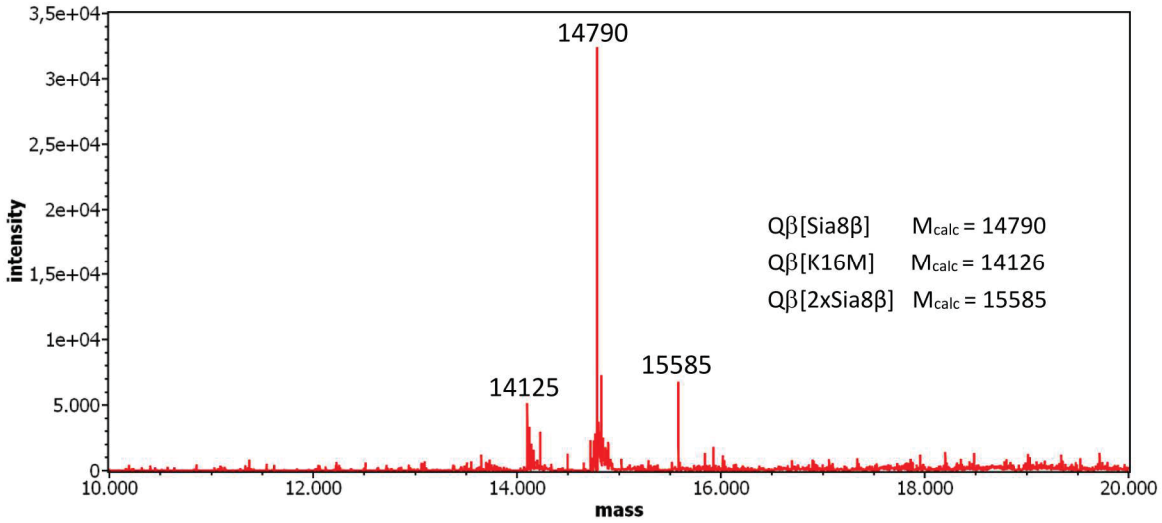


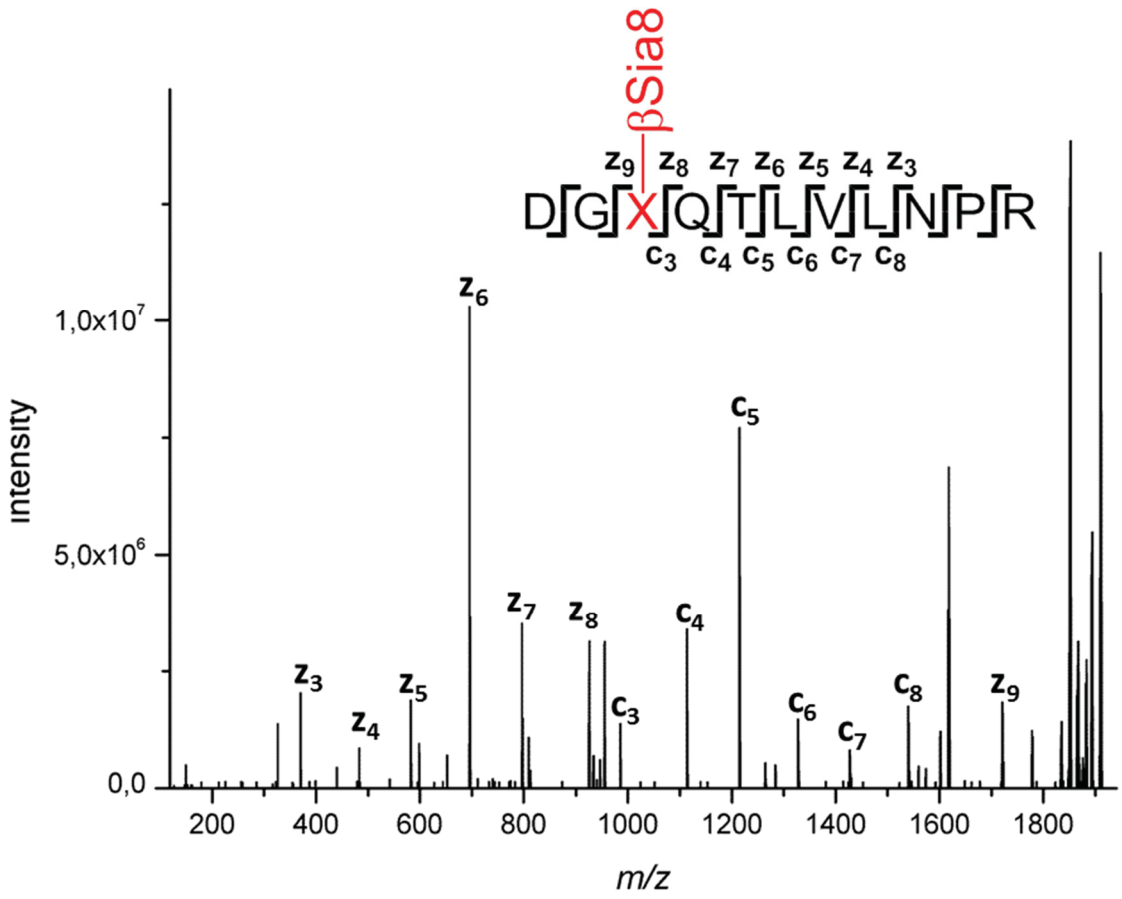
9.4.6 Analysis of Qβ[Sia8α]



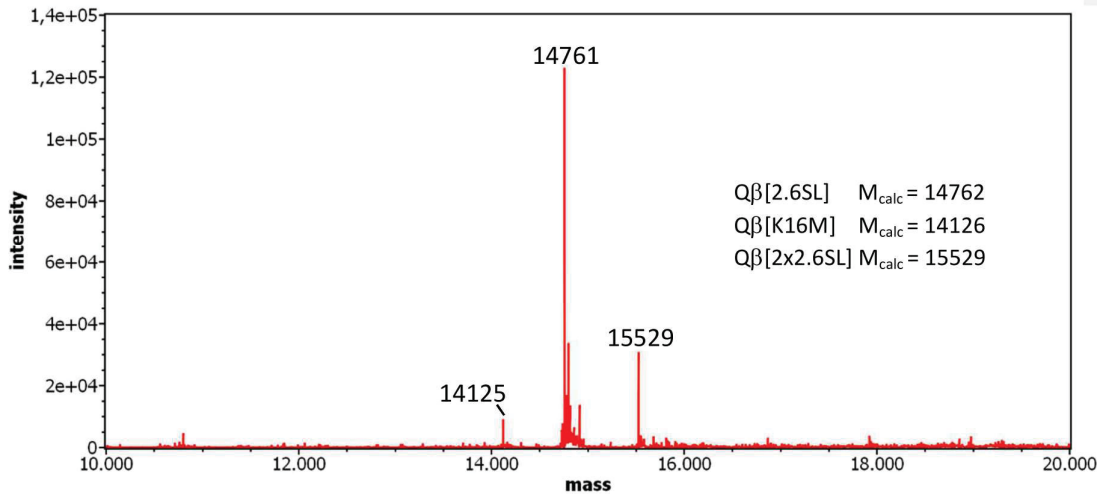


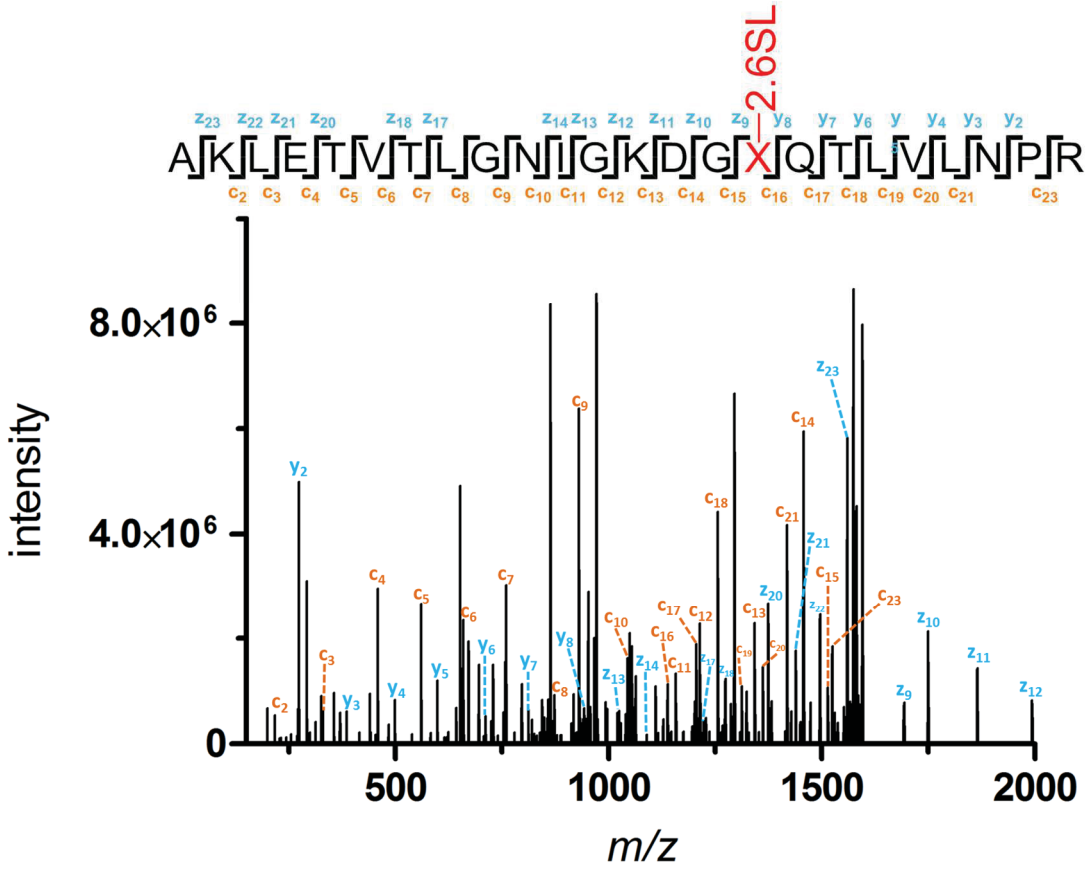
9.4.7 Analysis of Qβ[Sia8β]



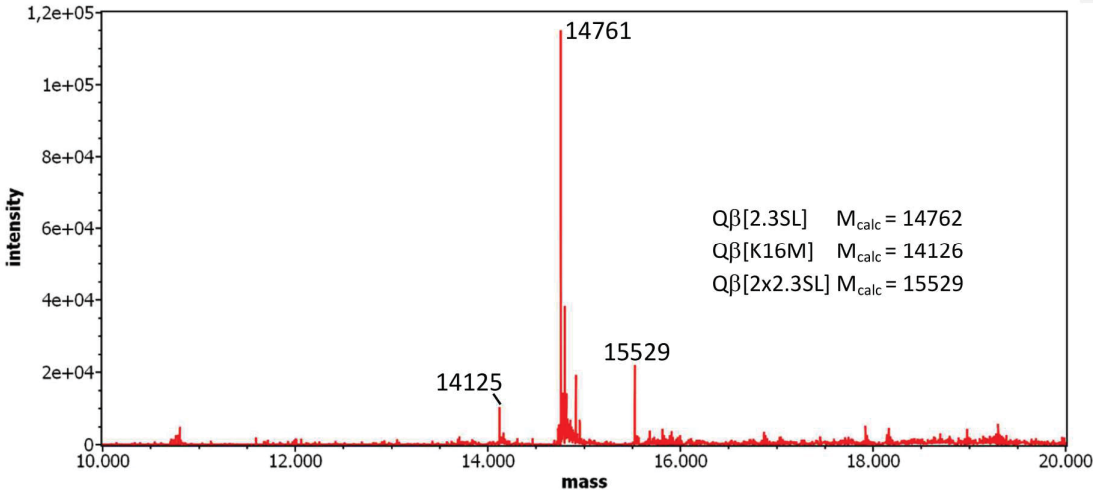


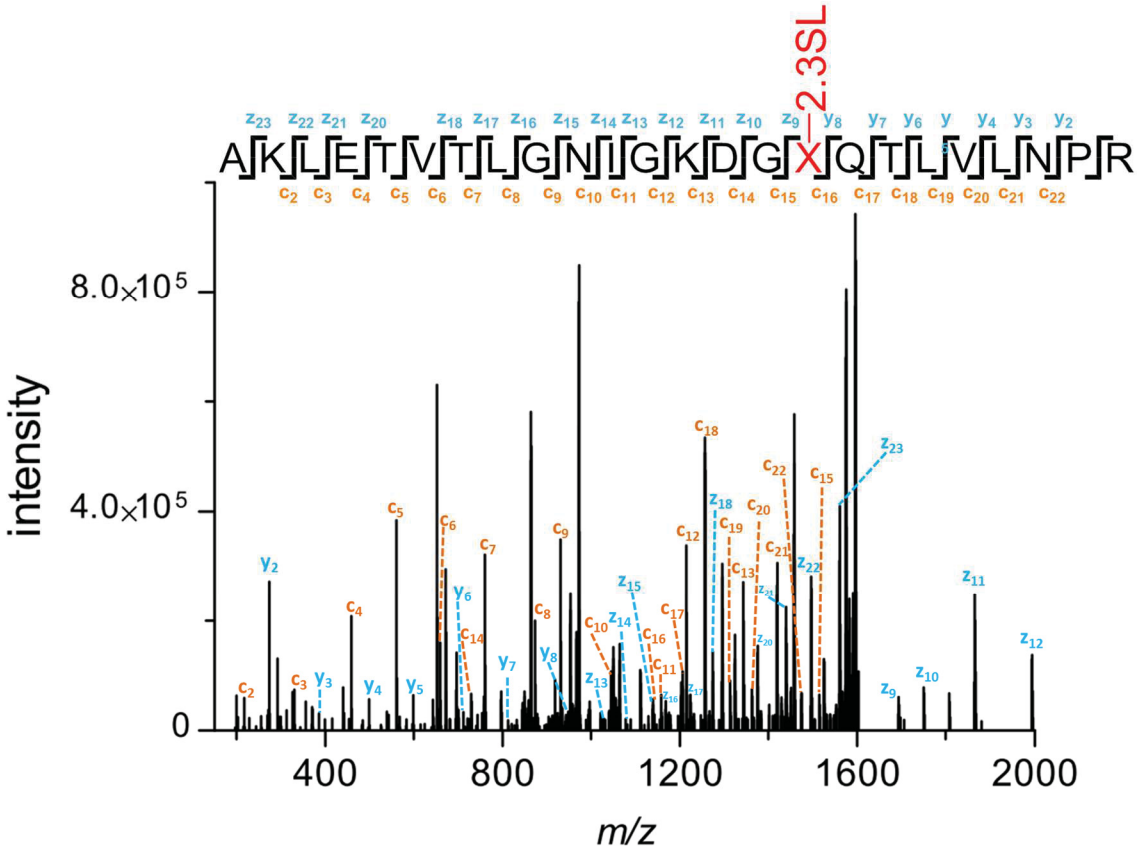
9.4.8 Analysis of Qβ[2.6SL]





9.4.9 Analysis of Qβ[2.3SL]



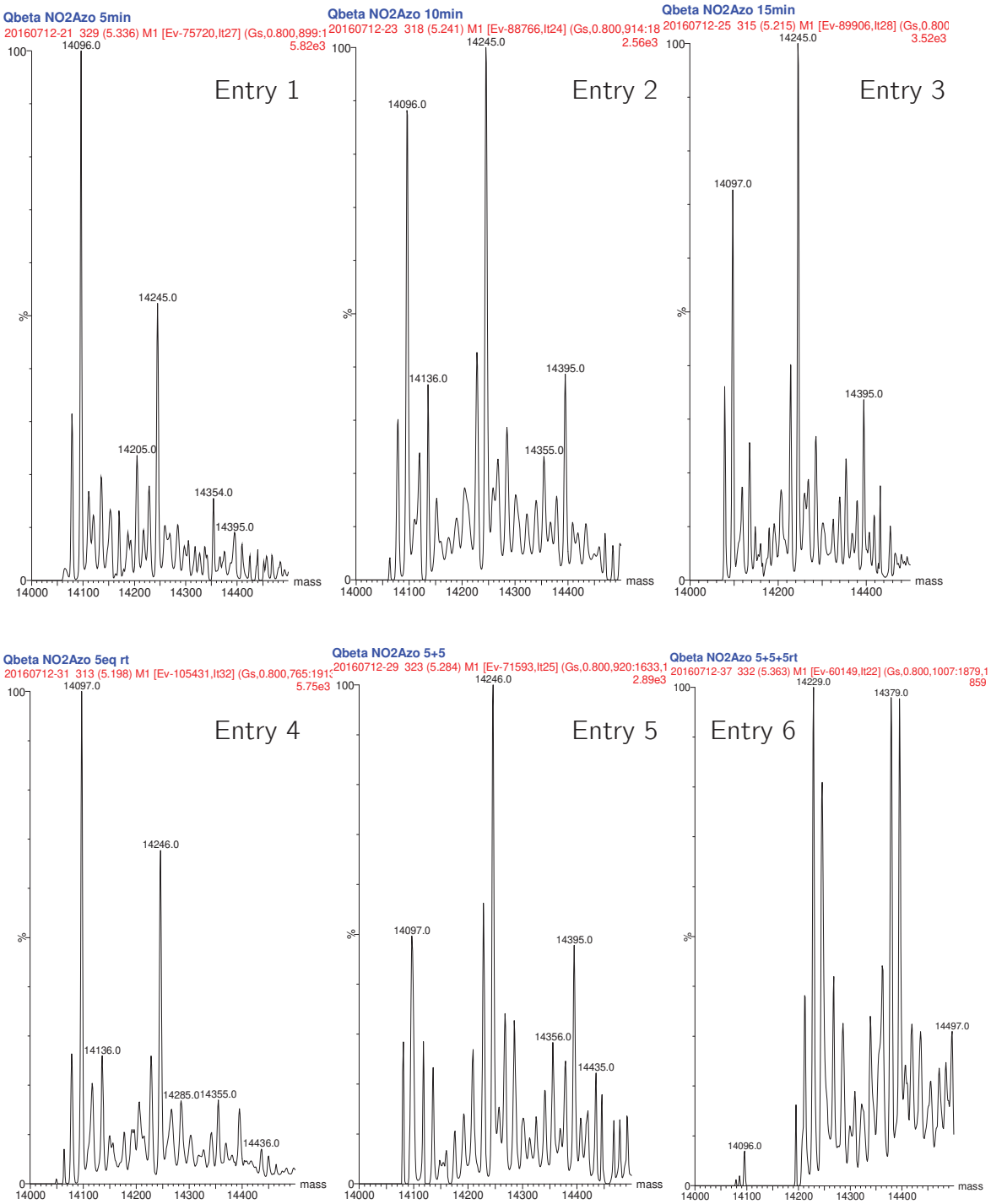


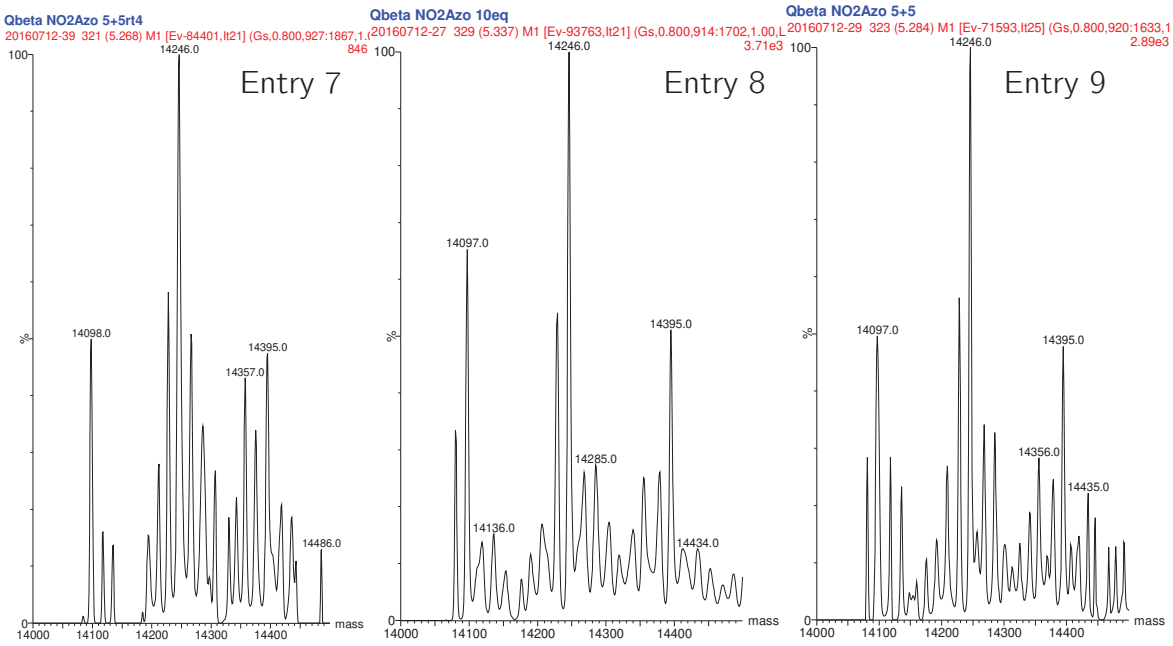
9.5 Interior Capsid Modification

According to a literature protocol,[167] hydrolysis of encapsidated expression host RNA was achieved by adding $PbOAc_2$ up to a final concentration of 2.5 mM to Qβ[Hpg] in 50 mM HEPES buffer (pH 7.4). The mixture was incubated for 24 h at 37°C and purified from heavy metals amd degraded RNA by spin-filtration.

Diazonium-coupling reaction were performed based on a literature procedure,[168] with commercially available 4-nitrobenzenediazonium tetrafluoroborate (4-NBDT) as diazonium reagent. Tested conditions are described in Section 7.9. The optimal conditions which were found are the following: A stock solution of emptied Qβ[Hpg] capsids in potassium phosphate buffer (pH 9.2) (278 nM ≙ 50 µM of coat protein) was prepared, of which 10 µL was used for each reaction. A 2.5 mM stock solution of 4-NBDT in water was prepared of which 1 µL (5 eq.) was added to Qβ[Hpg] at room temperature. After 5 minutes, the mixture was cooled to 4°C and another portion of

4-NBDT (1 µL) was added and incubated at this temperature for additional 10 minutes. The reaction was stopped by the addition of 1 µL of a 25 mM *p*-cresol stock solution in water. After dilution with 400 µL potassium phosphate buffer (pH 9.2), the sample was washed by spin-filtration and subsequently analyzed by MS. The product MS spectra for all conditions in Section 7.9, Table 13 are schown below.



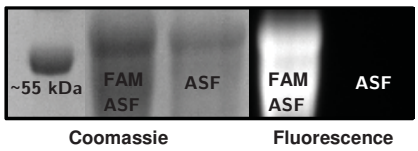


Reductive cleavage of the diazonium-coupled product was tested by adding 0.5 mg of solid sodium dithionite (2.87 µmol) and incubation at room temperature for 2 h. The sample was directly analyzed by MS.

9.6 HepG2 Cell Culture and Uptake Experiments

HepG2 cells were cultured in RPMI 1640 medium (*Biochrom*) containing 10% fetal calf serum, 1% penicillin/streptomycin and phenol red as an indicator (RPMI+) at 37°C under a 5% CO_2 atmosphere. Cells were passaged after detachment with accutase. One day before the experiment, 300 µL of a $2.5 \cdot 10^5$/mL cell suspension were transferred to each well of an 8-well *ibidi*® glass bottom slide which was coated with poly-L-lysine. The cells were allowed to adhere in RPMI+ over night, and were washed with RPMI medium without fetal calf serum, penicillin/streptomycin and phenol red (RPMI-) before the experiment was started.

Asialofetuin and fetuin were FAM-labeled as it is described in the main text. SDS-PAGE showed successful labelling of the protein.



HepG2 cells were incubated in RPMI- with FAM-labeled asialofetuin and fetuin (5 µM final concentration each) for 2 h at 37°C under a 5% CO_2 atmosphere. Cytochalasin D (10 µM) and galactose (5 and 10 mM) were used as uptake inhibitors and coincubated. Subsequently, the cell were washed two times with fresh RPMI- and 200 µL of fresh RPMI- was added for the imaging process. Cell nuclei were stained with Hoechst33342. Images were acquired on *Zeiss LSM710* and *LSM780* confocal microscopes.

10. References

[1] http://www.who.int/immunization/topics/influenza/en/, accessed on 12.01.2018.

[2] http://www.who.int/biologicals/vaccines/influenza/en/, accessed on 12.01.2018.

[3] T. Jefferson, V. Demicheli, C. Di Pietrantonj, D. Rivetti, *Cochrane Database Syst. Rev.* **2006**, CD001169.

[4] M. v. Itzstein, *Nat. Rev. Drug Discov.* **2007**, *6*, 967–974.

[5] A. J. Hay, V. Gregory, A. R. Douglas, Y. P. Lin, *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **2001**, *356*, 1861–1870.

[6] N. M. Bouvier, P. Palese, *Vaccine* **2008**, *26 Suppl 4*, D49-53.

[7] C. R. Parrish, P. R. Murcia, E. C. Holmes, *J. Virol.* **2015**, *89*, 2990-2994.

[8] S. A. Call, M. A. Vollenweider, C. A. Hornung, D. L. Simel, W. McKinney, *J. Am. Med. Assoc.* **2005**, *293*, 987-997.

[9] K. T. Jeffery, M. M. David, *Emerging Infect. Dis.* **2006**, *12*, 15.

[10] https://www.cdc.gov/flu/avianflu/influenza-a-virus-subtypes.htm, accessed on 14.01.2018.

[11] *Bull. W.H.O.* **1980**, *58*, 585-591.

[12] a) W. Weis, J. H. Brown, S. Cusack, J. C. Paulson, J. J. Skehel, D. C. Wiley, *Nature* **1988**, *333*, 426–431; b) N. K. Sauter, G. D. Glick, R. L. Crowther, S. J. Park, M. B. Eisen, J. J. Skehel, J. R. Knowles, D. C. Wiley, *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 324–328.

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

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

[15] a) Y. C. Lee, R. T. Lee, *Acc. Chem. Res.* **1995**, *28*, 321–327; b) J. R. Cochran, D. Aivazian, T. O. Cameron, L. J. Stern, *Trends Biochem. Sci* **2001**, *26*, 304–310; c) J. C. Sacchettini, L. G. Baum, C. F. Brewer, *Biochemistry* **2001**, *40*, 3009–3015; d) C. Brewer, *Curr. Opin. Struct. Biol.* **2002**, *12*, 616–623.

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

[17] a) L. G. Baum, J. C. Paulson, *Acta Histochem. Suppl.* **1990**, *40*, 35-38; b) J. N. Couceiro, J. C. Paulson, L. G. Baum, *Virus Res.* **1993**, *29*, 155-165.

[18] U. Kumlin, S. Olofsson, K. Dimock, N. Arnberg, *Influenza Other Respi. Viruses* **2008**, *2*, 147-154.

[19] Y. A. Shtyrya, L. V. Mochalova, N. V. Bovin, *Acta Naturae* **2009**, *1*, 26-32.

[20] G. M. Air, W. G. Laver, *Proteins* **1989**, *6*, 341-356.

[21] M. N. Matrosovich, T. Y. Matrosovich, T. Gray, N. A. Roberts, H. D. Klenk, *J. Virol.* **2004**, *78*, 12665-12667.

[22] R. J. Russell, L. F. Haire, D. J. Stevens, P. J. Collins, Y. P. Lin, G. M. Blackburn, A. J. Hay, S. J. Gamblin, J. J. Skehel, *Nature* **2006**, *443*, 45.

[23] a) R. J. Sugrue, A. J. Hay, *Virology* **1991**, *180*, 617-624; b) L. J. Holsinger, R. A. Lamb, *Virology* **1991**, *183*, 32-43.

[24] L. H. Pinto, G. R. Dieckmann, C. S. Gandhi, C. G. Papworth, J. Braman, M. A. Shaughnessy, J. D. Lear, R. A. Lamb, W. F. DeGrado, *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 11301-11306.

[25] J. R. Schnell, J. J. Chou, *Nature* **2008**, *451*, 591-595.

[26] R. M. Pielak, J. J. Chou, *Biochim. Biophys. Acta* **2011**, *1808*, 522-529.

[27] J. F. Hers, *Am. Rev. Respir. Dis.* **1966**, *93*, Suppl:162-177.

[28] a) H.-D. Klenk, W. Garten, *Trends Microbiol.* **1994**, *2*, 39–43; b) D. A. Steinhauer, *Virology* **1999**, *258*, 1–20.

[29] A. Helenius, *Cell* **1992**, *69*, 577–578.

[30] a) T. Maeda, S.-i. Ohnishi, *FEBS Letters* **1980**, *122*, 283–287; b) T. Maeda, K. Kawasaki, S. Ohnishi, *Proc. Natl. Acad. Sci. U. S. A.* **1981**, *78*, 4133–4137; c) A. Yoshimura, K. Kuroda, K. Kawasaki, S. Yamashina, T. Maeda, S. Ohnishi, *J. Virol.* **1982**, *43*, 284–293.

[31] J. F. Cros, P. Palese, *Virus Res.* **2003**, *95*, 3-12.

[32] J. C. Kash, A. G. Goodman, M. J. Korth, M. G. Katze, *Virus Res.* **2006**, *119*, 111-120.

[33] D. P. Nayak, E. K. Hui, S. Barman, *Virus Res.* **2004**, *106*, 147-165.

[34] a) M. B. Eisen, S. Sabesan, J. J. Skehel, D. C. Wiley, *Virology* **1997**, *232*, 19–31; b) I.-C. Huang, W. Li, J. Sui, W. Marasco, H. Choe, M. Farzan, *J. Virol.* **2008**, *82*, 4834–4843.

[35] A. D. Osterhaus, *Science* **2000**, *288*, 1051–1053.

[36] a) M. Yamashita, M. Krystal, W. M. Fitch, P. Palese, *Virology* **1988**, *163*, 112–122; b) E. Nobusawa, K. Sato, *J. Virol.* **2006**, *80*, 3675–3678.

[37] Y. Matsuzaki, N. Katsushima, Y. Nagai, M. Shoji, T. Itagaki, M. Sakamoto, S. Kitaoka, K. Mizuta, H. Nishimura, *J. Infect. Dis.* **2006**, *193*, 1229–1235.

[38] a) A. C. Joosting, B. Head, M. L. Bynoe, D. A. Tyrrell, *Br. Med. J.* **1968**, *4*, 153–154; b) S. Katagiri, A. Ohizumi, M. Homma, *J. Infect. Dis.* **1983**, *148*, 51–56.

[39] G. Herrler, I. Dürkop, H. Becht, H. D. Klenk, *J. Gen. Virol.* **1988**, *69 (Pt 4)*, 839–846.

[40] B. M. Hause, M. Ducatez, E. A. Collin, Z. Ran, R. Liu, Z. Sheng, A. Armien, B. Kaplan, S. Chakravarty, A. D. Hoppe, R. J. Webby, R. R. Simonson, F. Li, *PLoS Path.* **2013**, *9*, e1003176.

[41] Z. Sheng, Z. Ran, D. Wang, A. D. Hoppe, R. Simonson, S. Chakravarty, B. M. Hause, F. Li, *Arch. Virol.* **2014**, *159*, 249–255.

[42] E. A. Collin, Z. Sheng, Y. Lang, W. Ma, B. M. Hause, F. Li, *J. Virol.* **2015**, *89*, 1036–1042.

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

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

[45] F. Karush, in *Immunoglobulins* (Eds.: G. W. Litman, R. A. Good), Springer US, Boston, MA, **1978**, pp. 85-116.

[46] M. Mammen, G. Dahmann, G. M. Whitesides, *J. Med. Chem.* **1995**, *38*, 4179–4190.

[47] M. Mammen, K. Helmerson, R. Kishore, S.-K. Choi, W. D. Phillips, G. M. Whitesides, *Chem. Biol.* **1996**, *3*, 757-763.

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

[49] J. E. Kingery-Wood, K. W. Williams, G. B. Sigal, G. M. Whitesides, *J. Am. Chem. Soc.* **1992**, *114*, 7303–7305.

[50] T. J. Pritchett, J. C. Paulson, *J. Biol. Chem.* **1989**, *264*, 9850–9858.

[51] a) G. D. Glick, J. R. Knowles, *J. Am. Chem. Soc.* **1991**, *113*, 4701–4703; b) M. Waldmann, R. Jirmann, K. Hoelscher, M. Wienke, F. C. Niemeyer, D. Rehders, B. Meyer, *J. Am. Chem. Soc.* **2014**, *136*, 783–788.

[52] V. Bandlow, S. Liese, D. Lauster, K. Ludwig, R. R. Netz, A. Herrmann, O. Seitz, *J. Am. Chem. Soc.* **2017**, *139*, 16389-16397.

[53] M. Yamabe, K. Kaihatsu, Y. Ebara, *Bioconjugate Chem.* **2018**, *29*, 1490-1494.

[54] E.-M. Strauch, S. M. Bernard, D. La, A. J. Bohn, P. S. Lee, C. E. Anderson, T. Nieusma, C. A. Holstein, N. K. Garcia, K. A. Hooper, R. Ravichandran, J. W. Nelson, W. Sheffler, J. D. Bloom, K. K. Lee, A. B. Ward, P. Yager, D. H. Fuller, I. A. Wilson, D. Baker, *Nat. Biotechnol.* **2017**, *35*, 667.

[55] J. Vonnemann, S. Liese, C. Kuehne, K. Ludwig, J. Dernedde, C. Böttcher, R. R. Netz, R. Haag, *J. Am. Chem. Soc.* **2015**, *137*, 2572–2579.

[56] G. T. Jennings, M. F. Bachmann, *Biol. Chem.* **2008**, *389*.

[57] H.-C. Tsai, T. Imae, in *Prog. Mol. Biol. Transl. Sci., Vol. 104* (Ed.: A. Villaverde), Academic Press, **2011**, pp. 101-140.

[58] I. Khan, K. Saeed, I. Khan, *Arab. J. Chem.* **2017**.

[59] a) E. Strable, D. E. Prasuhn, A. K. Udit, S. Brown, A. J. Link, J. T. Ngo, G. Lander, J. Quispe, C. S. Potter, B. Carragher, D. A. Tirrell, M. G. Finn,

Bioconjugate Chem. **2008**, *19*, 866–875; b) W. Wu, S. C. Hsiao, Z. M. Carrico, M. B. Francis, *Angew. Chem. Int. Ed.* **2009**, *48*, 9493–9497.

[60] D. Yan, Y. Q. Wei, H. C. Guo, S. Q. Sun, *Appl. Microbiol. Biotechnol.* **2015**, *99*, 10415-10432.

[61] X. Zhang, L. Xin, S. Li, M. Fang, J. Zhang, N. Xia, Q. Zhao, *Hum. Vaccin. Immunother.* **2015**, *11*, 1277-1292.

[62] S. A. Wynne, R. A. Crowther, A. G. W. Leslie, *Mol. Cell* **1999**, *3*, 771-780.

[63] T. S. Y. Guu, Z. Liu, Q. Ye, D. A. Mata, K. Li, C. Yin, J. Zhang, Y. J. Tao, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 12992-12997.

[64] X. S. Chen, R. L. Garcea, I. Goldberg, G. Casini, S. C. Harrison, *Mol. Cell* **2000**, *5*, 557-567.

[65] D. S. Peabody, B. Manifold-Wheeler, A. Medford, S. K. Jordan, J. do Carmo Caldeira, B. Chackerian, *J. Mol. Biol.* **2008**, *380*, 252-263.

[66] E. V. Koonin, M. Y. Galperin, in *Sequence - Evolution - Function: Computational Approaches in Comparative Genomics*, Boston, **2003**.

[67] W. Fiers, R. Contreras, F. Duerinck, G. Haegeman, D. Iserentant, J. Merregaert, W. Min Jou, F. Molemans, A. Raeymaekers, A. Van den Berghe, G. Volckaert, M. Ysebaert, *Nature* **1976**, *260*, 500.

[68] R. Golmohammadi, K. Valegård, K. Fridborg, L. Liljas, *J. Mol. Biol.* **1993**, *234*, 620-639.

[69] M. Patrik, J. G. T., W. Jörg, R. Franziska, L. Ylva, R. Kirsten, R. W. A., M. Philipp, B. M. F., *Eur. J. Immunol.* **2005**, *35*, 2031-2040.

[70] R. Golmohammadi, K. Fridborg, M. Bundule, K. Valegård, L. Liljas, *Structure* **1996**, *4*, 543–554.

[71] G. Spohn, K. Schwarz, P. Maurer, H. Illges, N. Rajasekaran, Y. Choi, G. T. Jennings, M. F. Bachmann, *J. Immunol.* **2005**, *175*, 6211-6218.

[72] J. H. Kuhn, S. Becker, H. Ebihara, T. W. Geisbert, K. M. Johnson, Y. Kawaoka, W. I. Lipkin, A. I. Negredo, S. V. Netesov, S. T. Nichol, G. Palacios, C. J. Peters, A. Tenorio, V. E. Volchkov, P. B. Jahrling, *Arch. Virol.* **2010**, *155*, 2083-2103.

[73] T. W. Geisbert, P. B. Jahrling, *Virus Res.* **1995**, *39*, 129-150.

[74] a) K. L. Warfield, C. M. Bosio, B. C. Welcher, E. M. Deal, M. Mohamadzadeh, A. Schmaljohn, M. J. Aman, S. Bavari, *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 15889-15894; b) K. L. Warfield, D. L. Swenson, G. Demmin, S. Bavari, *Expert Rev. Vaccines* **2005**, *4*, 429-440.

[75] L. Liepold, S. Anderson, D. Willits, L. Oltrogge, J. A. Frank, T. Douglas, M. Young, *Magn. Reson. Med.* **2007**, *58*, 871-879.

[76] M. A. Bruckman, S. Hern, K. Jiang, C. A. Flask, X. Yu, N. F. Steinmetz, *J Mater Chem B* **2013**, *1*, 1482-1490.

[77] a) M. A. Lauffer, *J. Am. Chem. Soc.* **1944**, *66*, 1188-1194; b) A. Klug, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **1999**, *354*, 531-535.

[78] K. Jeong, C. Netirojjanakul, H. K. Munch, J. Sun, J. A. Finbloom, D. E. Wemmer, A. Pines, M. B. Francis, *Bioconjugate Chem.* **2016**.

[79] K. Namba, R. Pattanayek, G. Stubbs, *J. Mol. Biol.* **1989**, *208*, 307-325.

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

[81] C. E. Ashley, E. C. Carnes, G. K. Phillips, P. N. Durfee, M. D. Buley, C. A. Lino, D. P. Padilla, B. Phillips, M. B. Carter, C. L. Willman, C. J. Brinker, J. d. C. Caldeira, B. Chackerian, W. Wharton, D. S. Peabody, *ACS Nano* **2011**, *5*, 5729-5745.

[82] E. Kaltgrad, M. K. O'Reilly, L. Liao, S. Han, J. C. Paulson, M. G. Finn, *J. Am. Chem. Soc.* **2008**, *130*, 4578-4579.

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

[84] M. Zdanowicz, J. Chroboczek, *Acta Biochim. Pol.* **2016**, *63*, 469-473.

[85] E. M. Sletten, C. R. Bertozzi, *Acc. Chem. Res.* **2011**, *44*, 666-676.

[86] H. C. Hang, C. Yu, D. L. Kato, C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 14846-14851.

[87] S. T. Laughlin, J. M. Baskin, S. L. Amacher, C. R. Bertozzi, *Science* **2008**, *320*, 664-667.

[88] a) K. Lang, J. W. Chin, *ACS Chem. Biol.* **2014**, *9*, 16-20; b) K. Lang, J. W. Chin, *Chem. Rev.* **2014**, *114*, 4764-4806.

[89] C. J. Pickens, S. N. Johnson, M. M. Pressnall, M. A. Leon, C. J. Berkland, *Bioconjugate Chem.* **2018**, *29*, 686-701.

[90] R. Huisgen, *Angew. Chem. Int. Ed.* **1963**, *2*, 565-598.

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

[92] C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057-3064.

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

[94] T. R. Chan, R. Hilgraf, K. B. Sharpless, V. V. Fokin, *Org. Lett.* **2004**, *6*, 2853-2855.

[95] Z. Zhou, C. J. Fahrni, *J. Am. Chem. Soc.* **2004**, *126*, 8862-8863.

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

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

[98] V. Hong, N. F. Steinmetz, M. Manchester, M. G. Finn, *Bioconjugate Chem.* **2010**, *21*, 1912-1916.

[99] 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.

[100] a) M. Ahlquist, V. V. Fokin, *Organometallics* **2007**, *26*, 4389-4391; b) V. O. Rodionov, V. V. Fokin, M. G. Finn, *Angew. Chem. Int. Ed.* **2005**, *44*, 2210-2215; c) B. T. Worrell, J. A. Malik, V. V. Fokin, *Science* **2013**, *340*, 457-460.

[101] H. Chen, C. Cai, S. Li, Y. Ma, S. Luozhong, Z. Zhu, *Chem. Eur. J.* **2017**, *23*, 4730-4735.

[102] M. Meldal, C. W. Tornoe, *Chem. Rev.* **2008**, *108*, 2952-3015.

[103] L. Jin, D. R. Tolentino, M. Melaimi, G. Bertrand, *Sci. Adv.* **2015**, *1*.

[104] E. Baslé, N. Joubert, M. Pucheault, *Chem. Biol.* **2010**, *17*, 213-227.

[105] G. N. Cohen, D. B. Cowie, *C. R. Hebd. Seances Acad. Sci.* **1957**, *244*, 680-683.

[106] B. Nagel, H. Dellweg, L. M. Gierasch, in *Pure Appl. Chem., Vol. 64*, **1992**, p. 143.

[107] J. T. Ngo, D. A. Tirrell, *Acc. Chem. Res.* **2011**, *44*, 677-685.

[108] a) J. C. M. van Hest, D. A. Tirrell, *FEBS Letters* **1998**, *428*, 68-70; b) J. C. M. van Hest, K. L. Kiick, D. A. Tirrell, *J. Am. Chem. Soc.* **2000**, *122*, 1282–1288; c) K. L. Kiick, D. A. Tirrell, *Tetrahedron* **2000**, *56*, 9487-9493; d) K. L. Kiick, E. Saxon, D. A. Tirrell, C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 19-24.

[109] K. L. Kiick, J. C. M. van Hest, D. A. Tirrell, *Angew. Chem. Int. Ed.* **2000**, *39*, 2148-2152.

[110] M. Rother, J. A. Krzycki, *Archaea* **2010**, *2010*.

[111] L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001**, *292*, 498-500.

[112] J. W. Chin, S. W. Santoro, A. B. Martin, D. S. King, L. Wang, P. G. Schultz, *J. Am. Chem. Soc.* **2002**, *124*, 9026-9027.

[113] A. Deiters, P. G. Schultz, *Bioorg. Med. Chem. Lett.* **2005**, *15*, 1521-1524.

[114] a) W. Wan, Y. Huang, Z. Wang, W. K. Russell, P. J. Pai, D. H. Russell, W. R. Liu, *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 3211-3214; b) M. G. Hoesl, N. Budisa, *Angew. Chem. Int. Ed.* **2011**, *50*, 2896–2902.

[115] a) G. Blix, *Hoppe-Seyler's Z. Physiol. Chem.* **1936**, *240*, 43; b) E. Klenk, *Hoppe-Seyler's Z. Physiol. Chem.* **1941**, *268*, 50.

[116] R. Schauer, *Glycoconj. J.* **2000**, *17*, 485-499.

[117] B. Wang, J. Brand-Miller, *Eur. J. Clin. Nutr.* **2003**, *57*, 1351.

[118] E. Severi, A. Müller, J. R. Potts, A. Leech, D. Williamson, K. S. Wilson, G. H. Thomas, *J. Biol. Chem.* **2008**, *283*, 4841-4849.

[119] D. Adams, W. A. Gahl, in *GeneReviews((R))* (Eds.: M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. H. Bean, K. Stephens, A. Amemiya), Seattle (WA), **1993**.

[120] M. E. Tanner, *Bioorg. Chem.* **2005**, *33*, 216-228.

[121] a) H. Kayser, R. Zeitler, C. Kannicht, D. Grunow, R. Nuck, W. Reutter, *J. Biol. Chem.* **1992**, *267*, 16934-16938; b) O. T. Keppler, P. Stehling, M. Herrmann, H. Kayser, D. Grunow, W. Reutter, M. Pawlita, *J. Biol. Chem.* **1995**, *270*, 1308-1314.

[122] a) L. K. Mahal, K. J. Yarema, C. R. Bertozzi, *Science* **1997**, *276*, 1125-1128; b) E. Saxon, C. R. Bertozzi, *Science* **2000**, *287*, 2007-2010.

[123] H. Möller, V. Böhrsch, J. Bentrop, J. Bender, S. Hinderlich, C. P. R. Hackenberger, *Angew. Chem. Int. Ed.* **2012**, *51*, 5986-5990.

[124] J. Tiralongo, I. Martinez Duncker, *Sialobiology: Structure, Biosynthesis and Function. Sialic Acid Glycoconjugates in Health and Disease*, Bentham Science Publishers, **2013**.

[125] S. van der Vorm, T. Hansen, H. S. Overkleeft, G. A. van der Marel, J. D. C. Codee, *Chem. Sci.* **2017**, *8*, 1867-1875.

[126] G. J. Boons, A. V. Demchenko, *Chem. Rev.* **2000**, *100*, 4539-4566.

[127] R. Kuhn, P. Lutz, D. L. Macdonald, *Chem. Ber.* **1966**, *99*, 611-617.

[128] H. Paulsen, H. Tietz, *Angew. Chem. Int. Ed.* **1982**, *21*, 927-928.

[129] K. Higashi, S. Miyoshi, S. Nakabayashi, H. Yamada, y. Ito, *Chem. Pharm. Bull. (Tokyo)* **1992**, *40*, 2300-2303.

[130] H. Kunz, H. Waldmann, *J. Chem. Soc., Chem. Commun.* **1985**, 638-640.

[131] H. Paulsen, H. Tietz, *Carbohydr. Res.* **1984**, *125*, 47-64.

[132] A. Marra, P. Sinaÿ, *Carbohydr. Res.* **1990**, *195*, 303-308.

[133] W. Birberg, H. Lönn, *Tetrahedron Lett.* **1991**, *32*, 7453-7456.

[134] L. Beate, K. Horst, *Angew. Chem. Int. Ed.* **1997**, *36*, 618-621.

[135] A. Hasegawa, T. Nagahama, H. Ohki, K. Hotta, H. Ishida, M. Kiso, *J. Carbohyd. Chem.* **1991**, *10*, 493–498.

[136] O. Kanie, M. Kiso, A. Hasegawa, *J. Carbohyd. Chem.* **1988**, *7*, 501-506.

[137] D. Crich, W. Li, *J. Org. Chem.* **2007**, *72*, 2387–2391.

[138] D. Crich, W. Li, *J. Org. Chem.* **2007**, *72*, 7794–7797.

[139] P. K. Kancharla, C. Navuluri, D. Crich, *Angew. Chem. Int. Ed.* **2012**, *51*, 11105–11109.

[140] B. Dhakal, S. Buda, D. Crich, *J. Org. Chem.* **2016**, *81*, 10617–10630.

[141] K. Katrin, H. Richard, F. Dagmar, S. U. S., *Angew. Chem. Int. Ed.* **2010**, *49*, 6288-6308.

[142] S. Behren, *Synthesis of Thiol-Reactive Sialic Acid Derivatives for Protein Bioconjugations*, Humboldt-Universität zu Berlin, **2015**.

[143] L. M. Artner, *Convergent Syntheses of Multivalent Binding Glycoconjugates*, Freie Universität Berlin, **2014**.

[144] H. R. Pfaendler, V. Weimar, *Synthesis* **1996**, *1996*, 1345–1349.

[145] S. P. Amaral, M. Fernandez-Villamarin, J. Correa, R. Riguera, E. Fernandez-Megia, *Org. Lett.* **2011**, *13*, 4522–4525.

[146] J. Murray, D. Nowak, L. Pukenas, R. Azhar, M. Guillorit, C. Wälti, K. Critchley, S. Johnson, R. S. Bon, *J. Mater. Chem. B* **2014**, *2*, 3741–3744.

[147] C. L. Johnson, Z. Guo, *J. Carbohyd. Chem.* **2013**, *32*, 301–323.

[148] M. S. Cubberley, B. L. Iverson, *J. Am. Chem. Soc.* **2001**, *123*, 7560–7563.

[149] S. Svedhem, C.-Å. Hollander, J. Shi, P. Konradsson, B. Liedberg, S. C. T. Svensson, *J. Org. Chem.* **2001**, *66*, 4494–4503.

[150] A. C. French, A. L. Thompson, B. G. Davis, *Angew. Chem. Int. Ed.* **2009**, *48*, 1248-1252.

[151] G. Sennari, T. Hirose, M. Iwatsuki, S. Ōmura, T. Sunazuka, *Chem. Commun.* **2014**, *50*, 8715–8718.

[152] N. Khiar, M. Martinlomas, *J. Org. Chem.* **1995**, *60*, 7017-7021.

[153] C.-y. Li, G.-j. Liu, W. Du, Y. Zhang, G.-w. Xing, *Tetrahedron Lett.* **2017**, *58*, 2109-2112.

[154] N. Morishima, S. Koto, M. Oshima, A. Sugimoto, S. Zen, *Bull. Chem. Soc. Jpn.* **1983**, *56*, 2849–2850.

[155] J. P. Saludes, D. Sahoo, I. A. Monreal, *New J. Chem.* **2014**, *38*, 507.

[156] K. Agarwal, R. Kaul, M. Garg, A. Shajahan, S. K. Jha, S.-G. Sampathkumar, *J. Am. Chem. Soc.* **2013**, *135*, 14189-14197.

[157] H. Saneyoshi, Y. Yamamoto, K. Kondo, Y. Hiyoshi, A. Ono, *J. Org. Chem.* **2017**, *82*, 1796-1802.

[158] P. C. N. Rensen, S. H. van Leeuwen, L. A. J. M. Sliedregt, T. J. C. van Berkel, E. A. L. Biessen, *J. Med. Chem.* **2004**, *47*, 5798–5808.

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

[160] T. Machida, K. Lang, L. Xue, J. W. Chin, N. Winssinger, *Bioconjugate Chem.* **2015**, *26*, 802–806.

[161] H. M. G., A.-R. C. G., N. Sebastian, R. Marina, W. Christina, W. Birgit, B. Nediljko, A. Garabed, *ChemCatChem* **2011**, *3*, 213-221.

[162] W. Birgit, M. Lars, B. Nediljko, *ChemBioChem* **2009**, *10*, 217-220.

[163] D. Lauster, *Entwicklung multivalenter Inhibitoren des Eintritts von Influenzaviren in Wirtszellen*, Humboldt-Universität zu Berlin, **2018**.

[164] a) D. Lauster, M. Glanz, M. Bardua, K. Ludwig, M. Hellmund, U. Hoffmann, A. Hamann, C. Böttcher, R. Haag, C. P. R. Hackenberger, A. Herrmann,

Angew. Chem. Int. Ed. **2017**, *56*, 5931-5936; b) S. Bhatia, D. Lauster, M. Bardua, K. Ludwig, S. Angioletti-Uberti, N. Popp, U. Hoffmann, F. Paulus, M. Budt, M. Stadtmüller, T. Wolff, A. Hamann, C. Böttcher, A. Herrmann, R. Haag, *Biomaterials* **2017**, *138*, 22-34.

[165] S. Liese, *Statistical Mechanics Models for Multivalent Binding*, Freie Universität Berlin, **2017**.

[166] a) J. D. Fiedler, S. D. Brown, J. L. Lau, M. G. Finn, *Angew. Chem. Int. Ed.* **2010**, *49*, 9648-9651; b) J. L. Lau, M. M. Baksh, J. D. Fiedler, S. D. Brown, A. Kussrow, D. J. Bornhop, P. Ordoukhanian, M. G. Finn, *ACS Nano* **2011**, *5*, 7722-7729.

[167] M. L. Hovlid, J. L. Lau, K. Breitenkamp, C. J. Higginson, B. Laufer, M. Manchester, M. G. Finn, *ACS Nano* **2014**, *8*, 8003–8014.

[168] J. M. Hooker, E. W. Kovacs, M. B. Francis, *J. Am. Chem. Soc.* **2004**, *126*, 3718-3719.

[169] a) B. Ghesquiere, N. Colaert, K. Helsens, L. Dejager, C. Vanhaute, K. Verleysen, K. Kas, E. Timmerman, M. Goethals, C. Libert, J. Vandekerckhove, K. Gevaert, *Mol. Cell. Proteomics* **2009**, *8*, 2642–2652; b) B. Ghesquiere, M. Goethals, J. van Damme, A. Staes, E. Timmerman, J. Vandekerckhove, K. Gevaert, *Rapid Commun. Mass Spectrom.* **2006**, *20*, 2885–2893.

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

[171] M. Meier, M. D. Bider, V. N. Malashkevich, M. Spiess, P. Burkhard, *J. Mol. Biol.* **2000**, *300*, 857-865.

[172] a) U. Treichel, K.-H. Meyer zum Büschenfelde, R. J. Stockert, T. Poralla, G. Gerken, *J. Gen. Virol.* **1994**, *75*, 3021-3029; b) Y. Y. T. Seow, M. G. K. Tan, K. T. Woo, *Nephron* **2002**, *91*, 431-438.

[173] F. L. Miya, R. Robyn, H. J. E., S. S. L., *Traffic* **2000**, *1*, 161-171.

[174] G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw, K. I. Goldberg, *Organometallics* **2010**, *29*, 2176–2179.

[175] A. Shevchenko, H. Tomas, J. Havli, J. V. Olsen, M. Mann, *Nat. Protoc.* **2007**, *1*, 2856–2860.

[176] J. R. Kremer, D. N. Mastronarde, J. R. McIntosh, *J. Struct. Biol.* **1996**, *116*, 71-76.

[177] G. Yang, H. m. Ding, Z. Kochovski, R. Hu, Y. Lu, Y. q. Ma, G. Chen, M. Jiang, *Angew. Chem. Int. Ed.* **2017**, *56*, 10691-10695.

[178] P. L. Barili, G. Catelani, G. Fabrizi, D. Lamba, *Carbohydr. Res.* **1993**, *243*, 165-176.

[179] M. Weïwer, C.-C. Chen, M. M. Kemp, R. J. Linhardt, *Eur. J. Org. Chem.* **2009**, *2009*, 2611–2620.

[180] K. Okamoto, T. Kondo, T. Goto, *Tetrahedron* **1987**, *43*, 5909–5918.
[181] H. D. Premathilake, C. P. Gobble, P. Pornsuriyasak, T. Hardimon, A. V. Demchenko, C. De Meo, *Org. Lett.* **2012**, *14*, 1126–1129.