# Multivalent Carbohydrate Inhibitors of Bacterial Lectins and Toxins

## Multivalente Koolhydraat Inhibitoren van Bacteriële Lectinen en Toxinen

(met een samenvatting in het Nederlands)

**Proefschrift**

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus prof. dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op dinsdag 28 april 2015 des middags te 4.15 uur

door

**Ou Fu**

geboren op 17 december 1984 te Shanghai, China

Promotor: Prof. dr. R.J. Pieters

This research was financially supported by the Dutch Technology Foundation STW.

## Assessment committee

Prof. dr. R.J.M. Klein Gebbink

Prof. dr. G.J. de Jong

Prof. dr. R.M.J. Liskamp

Prof. dr. H. Zuilhof

Dr. C.A.M. de Haan

ISBN: 978-90-3936-315-7

Cover image: Designed by Ou Fu

Printed and Lay-out by: Proefschriftmaken.nl II Uitgeverij BOXPress

Published by: Uitgeverij BOXPress, ’s-Hertogenbosch

**“Look deep into nature and you will understand everything better”**

*-Albert Einstein*

# Table of contents

# Chapter 1

## General Introduction

## Multivalency Effects in Carbohydrate-based Therapeutic Design

## 1. Carbohydrates & their roles in medicinal chemistry

Carbohydrates are the most abundant class of natural products which consists of mono-, di-, oligo-, and polysaccharides. Especially in oligo- and polysaccharides, a variety of glycosidic linkage types results in highly branched structurally complex and diverse stereo-isomers leading to a huge amount of structural variation. Besides proteins, nucleic acids and lipids, carbohydrates are some of the most important components of living cells. The surface of mammalian cells is covered by a dense carbohydrates-coated complex, designated the glycocalyx, wherein carbohydrates are mainly present as conjugates of proteins (glycoproteins, proteoglycans) and lipids (glycolipids).[1] A selected group of monosaccharide residues is present in glycoproteins and glycolipids e.g. N-acetyl-D-glucosamine (GlcNAc), N-acetyl-D-galactosamine (GalNAc), D-glucose (Glc), D-galactose (Gal), D-mannose (Man), L-fucose (Fuc), D-xylose (Xyl), and N-acetylneuraminic acids (Neu5Ac), *etc*.[2] They are shown in **Fig.1**. Besides their use as essential and fundamental constituents of the cell surface and energy sources, carbohydrates may also participate in molecular recognition events that are associated with a variety of biological process e.g. intercellular recognition in infection, cancer, and the immune response, *etc*.[3] For instance, the important process of recruiting white blood cells to the infectious site in response to an injury or immunologic challenge is mediated by a specific binding between sialic-acid containing glycans and a family of the proteins known as selectins.[4] With respect to their structural diversity and biological importance, the synthesis of different sugar epitopes has attracted a growing attention from carbohydrate chemists over the years. As a result, most recent efforts and new developments of chemically-defined glycoconjugates in glycobiology research have focussed on carbohydrate-mediated recognition at the molecular-level.

As shown in **Fig.1**, on the surface of many microbes e.g. viruses and bacteria there are numerous glycan-binding proteins or so-called “lectins”. This enables the bacteria and viruses to bind or adhere to glycosylated cell surface as a prerequisite for infectious disease.[5] This specific carbohydrate-recognition plays a key role inside the human body where it facilitates the resulting bacterial and viral infections and

constitutes a major problem to human health. Therefore, under pathogenic conditions, cell surface

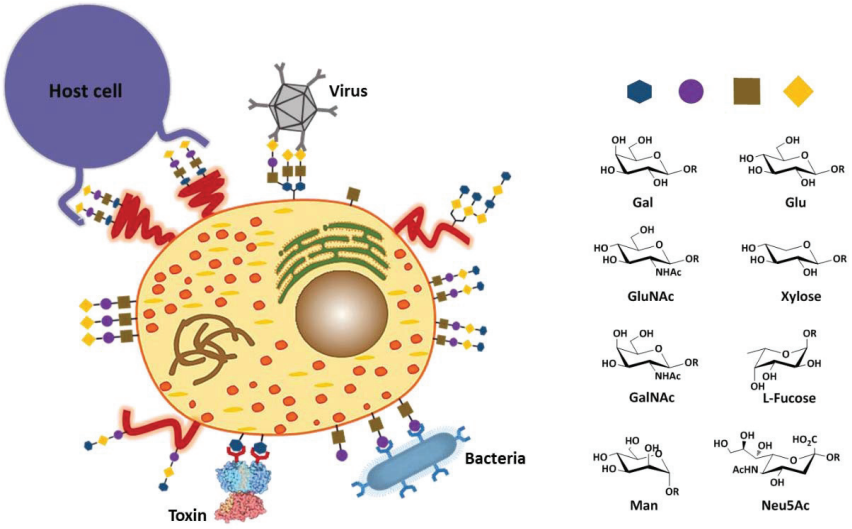



**Figure 1.** Illustration of monosaccharide residues and protein–carbohydrate interactions at the cell surface mediating cell–cell binding, cell–microbe (bacterial, viral) adhesion, cell-toxin adhesion, the red ribbons represent sugar-linked lipids or proteins. Adapted from reference 5a for the figure.

carbohydrate as specific receptors are essential for the subsequent invasive infection, where certain microbes could mimic the host glycans thereby aid their evasions from host immune response.[5c,5d] One infectious disease via carbohydrate-based adhesion is meningitis, which is caused by a zoonotic pathogen named *Streptococcus suis* binding to the Galα1-4Gal sequence on the host surface in pigs.[6] Another common incidence of infection is associated with the uropathogenic *E. coli*. It mainly utilizes two different types of fimbriae with two particular sugar specificities, the Galα1-4Gal specific P-fimbriae (containing the PapG adhesion protein) and the mannose specific type 1 fimbriae (containing the FimH adhesin).[7] Another well-known example of a cell-virus interaction is the influenza infection (common flu), where an initial sialic acid dependent binding is essentially required to pull the trigger of infection.[8] Carbohydrate-protein interactions play pivotal roles in enabling certain pathogens to penetrate or invade through epithelial barriers, whereupon they can distribute through the bloodstream thereafter produce the deep-seated infections. DC-SIGN (dendritic cell-specific intercellular adhesion molecule-3-grabbing nonintegrin) is a CX-type lectin that was reported to be instrumental for pathogenic recognition on

the surface of dendritic cells. It is found that DC-SIGN binds strongly to HIV via its encoded glycoprotein gp120.[9] as The same is true for hepatitis C virus to promote infection.[10, 11] Therefore, the development of carbohydrate-based anti-adhesion therapies is of great value. In spite of their obvious potential, there are significant challenges in the development of carbohydrate-based drugs. The hurdles include the oral availability and the pharmacokinetics. Despite this, several successful carbohydrate-based drugs ('sweet-medicines') have been developed.[12] (**Fig.2**)

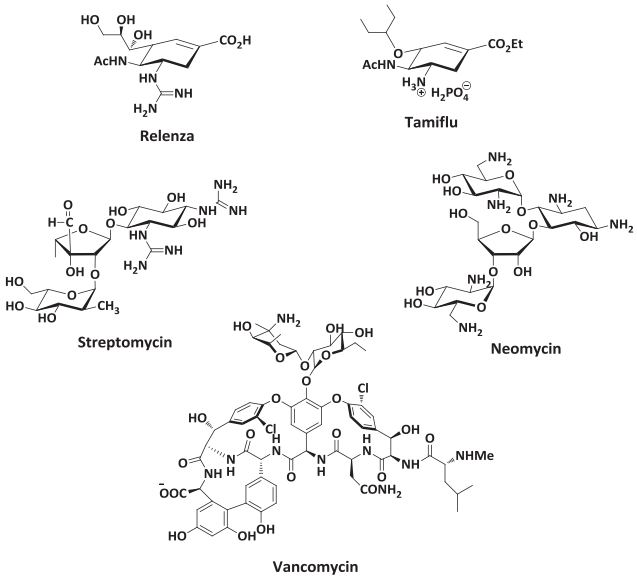



**Figure 2.** Carbohydrate and carbohydrate-derived drugs.[10]

As shown in **Fig.2**, Relenza and Tamiflu[13] are approved therapeutic agents for influenza infection. They could prevent the release of virus and reduce the magnitude and durations of symptoms with high efficacy. However, the usage of these drugs is limited, where they need to be administered within 2 days after symptoms development. And their adverse effects such as bronchospasms and neuropsychiatric side effects also minimize their usage.[14] Moreover, with the massive use of anti-microbial agents, the ever-growing antibiotic resistance has emerged at an alarming rate and given a rising concern for humanity. As reported in 2013, genes expressing resistance to Relenza (and Tamiflu) were found in Chinese patients infected with avian influenza A H7N9.[15] Unfortunately, a neomycin

phosphotransferase II gene also appeared which conferred resistance to this aminoglycoside antibiotic.[16a] In addition, some bacterial strains of *Staphylococcus aureus* and *Enterococcus* have already developed resistance to the glycopeptide antibiotic vancomycin.[16b] Therefore, there is an urgent need for both new classes of antibiotics that are effective against current resistant strains and new classes that possess highly effective anti-infection activity. However, the low affinity of monovalent carbohydrates to adhesion proteins requires a high dose of the sugar for effective inhibition. Inspired by Nature using the multivalent manner of binding epitopes in ligand-receptor interactions to achieve the overall enhanced affinity, application of a multivalency approach in drug discovery offers an opportunity for developing novel anti-infective compounds.[3a] A concerted effort has been made by numerous researchers to investigate the multivalency principle in inhibitor design.

## 2. Definition and mechanism of multivalency effect

Carbohydrate-based molecular interactions have been shown to be generally, but Nature seems to compensate these drawbacks with a multivalent receptor and ligand presentation. Over the past years, advanced by numerous biochemical studies, we have increased our understanding of multivalent interactions being fundamental to the modulation of many critical biological systems. As mentioned, they are prevalent in biology e.g. in the viral and bacterial adhesion to cell surfaces and in the binding of cells to other cells, *etc*. Their unique properties imparted through multivalent binding interactions have stimulated a variety of designed multivalent systems to overcome the intrinsic weak monosaccharide-protein interactions. In a molecular protein-carbohydrate recognition event, binding two or more ligands to one biological entity simultaneously can enhance binding affinity by several orders of magnitude. In the studies by Lee and co-workers, it was first recognized that the carbohydrate-protein interactions were increased with the number of carbohydrate residues involved. This phenomenon is known as the “cluster-glycoside effect”, where a dramatic improvement in binding affinity was exhibited by a cluster of glycosides to the asialoglycoprotein receptor in comparison with that of individual ligand.[17] Besides the collective effects on the affinity, multivalent interactions have a number of characteristics, which are not accessible by the corresponding

monovalent system. In general, multivalent ligands binding to receptors can be categorized into several types (**Fig.3**).

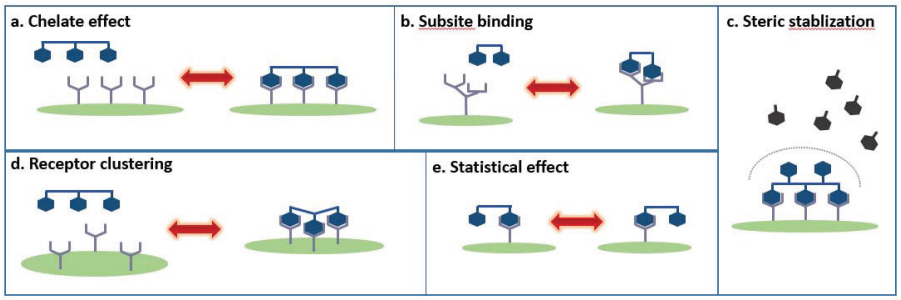



**Figure 3.** General types of multivalent ligand binding. Multivalent ligands possess multiple copies of a recognition element presented from the same scaffold, which allows a variety of additional mechanistic options: (a) Chelate effect. (b) Subsite binding. (c) Steric stabilization. (d) Receptor clustering (e) Statistical effect. Adapted from reference 18 for the figure.

As shown in **Fig.3**, several mechanisms have been proposed to explain the increased affinities of multivalent ligands for their binding partners. The higher binding affinities of multivalent complexes could be from an entropically favoured binding. Since the cost of translational entropy is already paid after the first receptor-ligand contact, all the subsequent binding interactions could thereby proceed with smaller penalties in translational entropy.[19] As a consequence, with a lower entropic cost of multivalent ligands at the binding sites, they eventually impart an overall enhanced binding for the interaction. This is commonly referred to as the chelate effect (**Fig.3a**). In another type of multivalent interaction, the binding energy is mainly gained from contacting a separated secondary site or subsite through extension of the primary binding site. By this assisted-binding site, multivalent ligands are prone to become more effective inhibitors with the ability to occupy the sites simultaneously (**Fig.3b**).[20] The case of favourable nonspecific interactions in proximity to the receptors' surface could also be considered. Likewise, as another variation, multivalent ligands could bind to multiple non-oligomeric receptors. Since the cell membrane is fluid, an increased affinity observed through multiple recognition is primarily contributed by the mobility of diffused receptors. This binding process facilitated by the receptor clustering, differentiates it from the chelation mode. (**Fig.3d**). Even when only one receptor is involved in an interaction, multivalent

ligands can display higher affinities due to their high effective local concentration of ligands which should favour the rebinding mechanism, a phenomenon that is commonly called the statistical rebinding effect (**Fig.3e**). The effect is contributed by the overall slower off-rate of binding, due to the reoccupation of the site that occurs when the first ligand releases, thereby resulting in higher affinity. Another mechanism by which a multivalent ligand can act is through steric stabilization, which can inhibit the interaction of other binding elements (**Fig.3c**). In certain cases of inhibition of binding to viral surface receptors, this additional activity of bound multivalent ligands could shield the host cell from further contact with pathogens.[3c, 21] Nevertheless, in certain cases, multivalent ligands can also bind multivalent receptors in an unexpected fashion. Potentially, instead of an intramolecular association, formation of large aggregates facilitated by an intermolecular binding is possible. Such an effect is dependent on a range of factors including concentrations, affinities, valencies, and ratios of interacting species, etc. When the ligands cannot span multiples sites on a single multivalent receptor, the occurrence of ligand-induced aggregation seems likely.

Not only could the increased strength of multivalent interaction be the result of a synergistic cooperation of numbers of ligand-receptor contacts, but could also be demonstrated by their higher thermodynamic and kinetic stabilities. The thermodynamic profile of multivalent interactions has been extensively investigated by numerous researchers.[3a, 22] The following theoretical equation for calculating the strength of the multivalent association is derived from a review by Whitesides and coworkers.[3a] The free energy change (ΔG) of a ligand-receptor association is determined by enthalpic (ΔH) and entropic (ΔS) components. The equation of a standard relationship of monovalent association ($\Delta G^{mono}$, $\Delta H^{mono}$, $\Delta S^{mono}$) is expressed as follows, where the total entropy change involves the changes from both translational and rotational entropies.

$$\Delta G^{mono} = \Delta H^{mono} - T\Delta S^{mono}$$

$$\Delta S^{mono} \approx \Delta S^{mono}\,(\text{translational}) + \Delta S^{mono}\,(\text{rotational})$$

To elucidate the thermodynamic basis of multivalent interactions, it is easy to start with a simple divalent system. Under certain assumptions e.g. an unstrained and independent manner of divalent association, the free energy of binding ($\Delta G^{di}$) could be expressed in a similar way, where the enthalpy change is twice the enthalpy of monovalent binding. Regarding to the change of entropy, it consists of three components. The first two parts are as same as the monovalent association but the contribution of the change in entropy from the linker's participation is taken into consideration.

$\Delta G^{di} = \Delta H^{di} - T\Delta S^{di}$

$\Delta H^{di} \approx 2\ \Delta H^{mono}$

$\Delta S^{di} \approx \Delta S^{mono}$ (translational) + $\Delta S^{mono}$ (rotational) + $\Delta S^{di}$ (conformational, linker)

In cases of low valency such as a divalent or trivalent ligand, the above equation predicts that the intrinsic affinity of a monovalent system is rather important to the strength of multivalent association. As indicated by the divalent system, the free energy of an advantageous multivalent association is deduced from two ideal systems. One involves n independent pairs of complexes of n ligand copies and n receptor copies. The other one is a multivalent ligand with n sites binding to a multivalent receptor with n sites as well to form one perfectly matched multivalent complex. Thus, the equations are expressed as follows.

$\Delta G^{multi} = n\Delta G^{mono} = n\Delta H^{mono} - nT\Delta S^{mono}$

$\Delta G^{multi(n)} = \Delta H^{multi(n)} - T\Delta S^{multi(n)} \approx n\Delta H^{mono} - T\Delta S^{mono} - T\Delta S^{multi(n)}$ (conformational, linker)

To subtract these above two equations, the change of free energy could be represented as $\Delta\Delta G^{multi} = \Delta G^{multi} - n\Delta G^{mono} = (n-1)T\Delta S^{mono} - T\Delta S^{multi}$. As implied, due largely to the entropic contribution of monovalent binding ($\Delta S^{mono}$ negative), the tighter multivalent association is favoured. This is in accordance with the result of the chelation effect as well because it suffers a relatively smaller entropy loss after first binding event. Even though it is often associated with the conformational entropy penalty of the linker upon receptor-ligand association, the enthalpy term makes a dominant and positive contribution to the free energy to intensify the

multivalent association. In addition, the parameter n also indicates that the multivalent association is strengthened as a function of the valency.

The fundamental concept of multivalent binding has been elucidated in both general binding mechanism and thermodynamic aspects. However, a complete mechanistic description of multivalent binding is still difficult because a combination of several binding events could be involved in the resulting interaction which could be driven by enthalpy or entropy. In principle, from a thermodynamic point of view, it suggests that the binding enhancement could be influenced by several binding parameters e.g. a higher valency number (n), a tighter monovalent association (more negative $\Delta H^{mono}$), a lower conformational entropy cost rigid linker part, *etc*. Nevertheless, there are many possibilities that could complicate this simple picture. The correct spacings of linker molecules and their rigidity/flexibility, the mode of presenting binding epitopes, the conformational restrained architecture, favourable/non-favourable additional interactions, all of them could affect the final outcome in multivalent system.

## 3. Multivalent design of anti-bacterial adhesive agents

Bacteria and protein toxins produced by bacteria are the cause of many life threatening diseases. Antibiotics are prevalently taken to inhibit bacterial growth and work at the first time to combat and eradicate bacterial disease. However, bacteria have evolved resistance to many current antibiotics,[23] which affords the bacteria the capability to remain alive. The emergence and spread of bacterial resistance poses an increasingly significant concern to the healthcare community.[24] To this end, more effective agents are urgently needed to circumvent the bacterial resistance. As mentioned above, carbohydrates are key mediators for bacterial and viral infection, immunological recognition of tumor cells and pathogens, and hormone-cell recognition, *etc*.[25] In a bacterial infection, the first step of the bacterial pathogens is often the adherence to the epithelial surfaces of host tissues. Many of them attach to the host via carbohydrate-protein interactions via their surface proteins interacting with specific carbohydrates present on eukaryotic glycoproteins or glycolipids.[26] Advances in the identification and synthesis of glycans have made carbohydrates attractive potential therapeutic targets. Therefore, bacterial adhesion

and the resulting infectious disease can in principle be prevented by ligands that mimic oligosaccharides binding sites and thus block this recognition or adhesion process.[27] However, the intrinsic interaction between an individual carbohydrate and protein is relatively weak. Multivalency is evidently very important and necessary in protein–glycan interactions between pathogenic microorganisms and mammalian host cells.[3a] Because they present multiple copies of a receptor-binding element, multivalent ligands can bind to receptors with more strongly . These weak interactions can be exceptionally amplified. It has an implication in the design and presentation of carbohydrates in multivalent form to target different pathogens and thereby modulate their interactions (**Fig.4**).[28] Much effort has been made in the preparation of carbohydrate-based multivalent ligands thereby serving as powerful inhibitors. Synthetic scaffolds like proteins, peptides, fullerenes, calixarenes, dendrimers, nanoparticles, *etc.* with different valences, linkers and spatial organizations have been used to multimerize carbohydrates and to enhance the affinity and specificity of the monomers towards their multivalent targets. Definite advancements have been achieved. In the following part, some outstanding examples from the literature are discussed to highlight the multivalent carbohydrates made by various research groups as potential therapeutic agents to inhibit the adhesion of bacteria and their toxins.

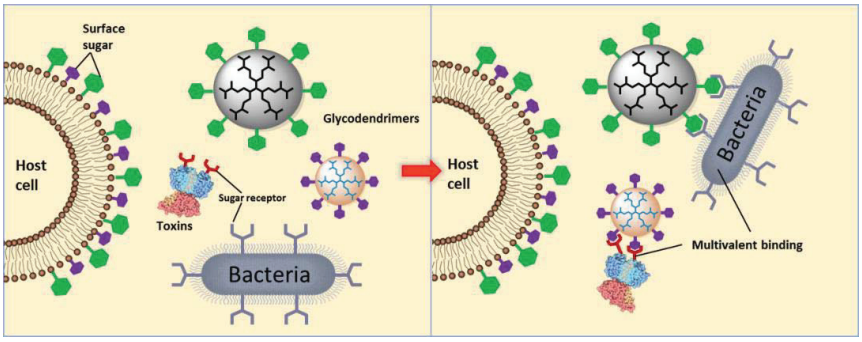



**Figure 4.** Multivalent binding of a bacterium or a bacterial toxin to a glycodendrimer. Adapted from reference 28 for the figure.

## 3.1 Inhibition of *Pseudomonas aeruginosa* (lung pathogens)

*Pseudomonas aeruginosa* is an opportunistic human pathogen. It is the causative agent of a variety of infections and an important morbidity and mortality factor for

immuno-compromised and cystic fibrosis patients.[29a] Recently, it was demonstrated that in some cases this bacterium has become multi-antibiotic resistant owing to its formation of surface attached biofilms. The induced biofilms play an important role in antibiotic resistance and disease progression.[29b] This altered microenvironment provides a barrier for antibiotic access.[30] Therefore, there is an urgent need to develop novel therapeutic agents effective against *P. aeruginosa*. Biofilm formation is partly mediated by two adhesion factors, PA-IL (LecA) and PA-IIL (LecB), which bind selectively to D-galactose and L-fucose as evidenced by studies with deletion mutants.[31] Interfering with these adhesion processes inhibits not only tissue adhesion but also seemingly the biofilm development. Both LecA and LecB inhibition could therefore provide an efficient way to treat or prevent infections of *Pseudomonas aeruginosa*. As shown in **Figure 5**, a crystal structure revealed LecA to be a tetrameric lectin with specificity for terminal galactoside residues with medium-binding affinity ($K_d$ around 50 μM).[32, 34] The low intrinsic affinity of carbohydrate-protein interactions hampers the development of low-molecular-weight inhibitors. Gaining advantage from the cluster effect observed in the binding of multiple carbohydrates to lectins and the structural knowledge of LecA, several groups set out to exploit the multivalency strategy to interfere this pathogen adhesion and efficient ligand preparation using CuAAC.[33]

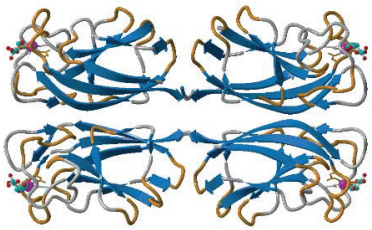


**Figure 5.** Structure of LecA in complex with galactose (PDB code 1OKO).[34]

As indicated by the crystal structure of a LecA-galactose complex, the shortest distance between one pair of neighboring binding sites is ca. 26Å. Guided with this structural knowledge, recently, our group[35] has designed a series of symmetrical divalent ligands presenting galactose residues with varied length of the spacer to fit the a LecA divalent binding mode. A relatively rigid and straight spacer based on the alternation of glucose moieties linked at the 1 and 4 positions by a 1,2,3-triazole unit

via ‘‘click’’ chemistry was designed. By varying the number of glucose building blocks, a variety of spacer distances was achieved. The binding profile of the synthesized compounds was determined by an ELISA-type assay. Among them, compound **1** and compound **2** with a 3 unit spacer (**Fig.6**) were found to contain the optimal spacer to interact with LecA. Clearly, the linker length was proven to be an important factor for the binding potency. As shown in **Fig.6**, both of them were composed of tri-glucose

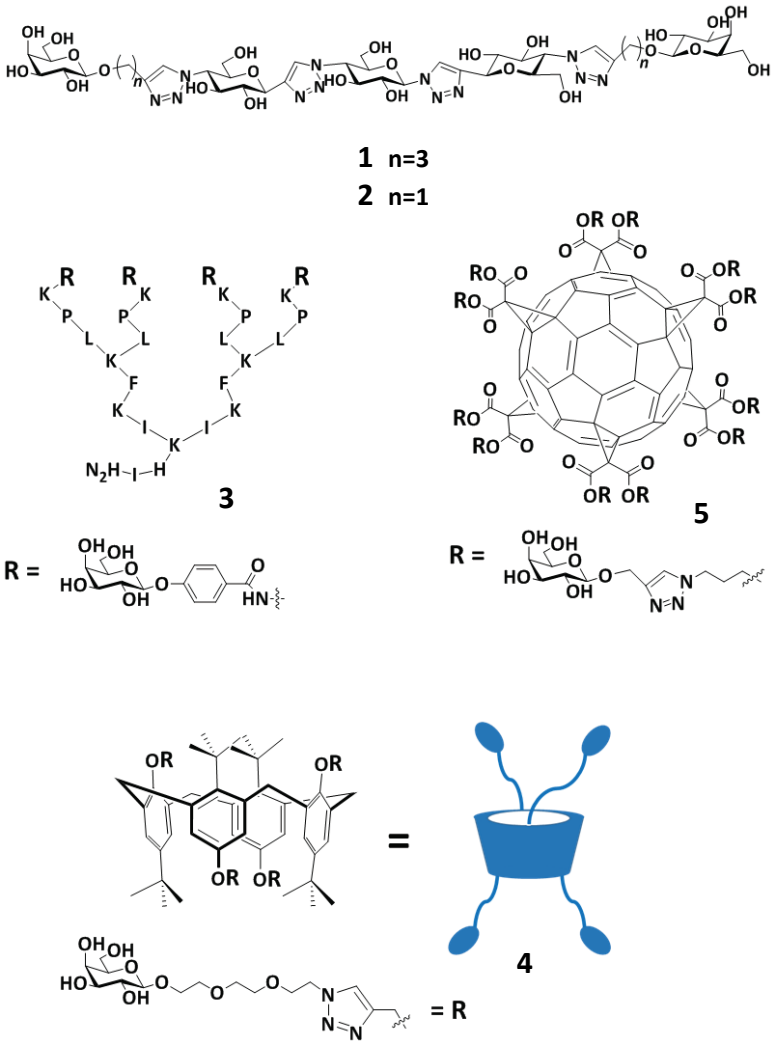



**Figure 6.** Structures of LecA inhibitors in varied multivalent display

spacer appending two galactose units with two aglycone linker lengths (n=1 and n=3) at both ends. With an appropriate spacer-length, it was found that a chelation type

binding between compound **1** and compound **2** was achieved. In a multivalent system, the chelate effect leads to less unfavorable entropy as carbohydrate ligands are bridged by suitable spacers and the ligands simultaneously bind to neighbouring binding sites of the adhesin. The entropic barriers are thought to be reduced after binding of the first ligand. Compound **1** exhibited an $IC_{50}$ value 0.12 µM , which was a 484-fold increase in relative potency per sugar compared to a monovalent reference compound. When shortening the aglycone linker from C3 to C1 in compound **2**, an optimum point was reached. It turned out that a remarkable low nanomolar inhibition was obtained ($IC_{50}$= 2.7 nM), which was over 7500-fold stronger than the relevant reference compound. These results were also in agreement with isothermal calorimetry (ITC) studies. Compound **1** was found as second-best binder with a $K_d$ value of 130 nM, while compound **2** had a much stronger binding to LecA with a $K_d$ of 28 nm, thereby being close to 800-fold more potent than the monovalent probe. This extraordinary finding could be explained by the fact that the dimeric geometry of LecA was perfectly matched by compound **2**.

Reymond *et al*. introduced[36] a glycopeptide dendrimer inhibitor **3** (**Fig.6**), where related β-phenylgalactosides were ligated to four N-terminal lysines of a peptidic dendritic scaffold to probe the multivalency effect on LecA binding. In comparison to the reference D-galactose, compound **3** showed a 4000-fold increased inhibition activity (MIC= 0.78 µM) in a hemagglutination assay and an 875-fold increased binding affinity ($K_d$= 0.1 µM) in an isothermal titration calorimetry (ITC) study. Moreover, as compared to a reference compound it was found that the presence of the aromatic β-phenyl aglycone linker part led to an additional contribution to the binding affinity gain. This was consistent with the thermodynamic parameters obtained from ITC, in which compound **3** had the most favored Gibbs free energy (ΔG). It was also shown by X-ray crystallography that compound **3** contains an aromatic aglycone linker that engages in an unprecedented T-stack interaction with His 50 at the LecA galactose binding site. A molecular dynamics (MD) study revealed that the peptido-arms appeared too short to bridge two galactose binding sites within the same LecA dimer. The chelation effect was unlikely to occur with

compound **3**. The enhanced binding affinity achieved by multivalency effect could result from the secondary interactions with the LecA surface.

The group of Vidal has also developed LecA inhibitors. In this case they were a family of calix[4]arene-based multivalent glycoconjugates .[37] In total seven topologically isomeric calix[4]arene glycoconjugates were prepared through the attachment of galactose azide moieties to alkyne-derivatized calix[4]arene platforms via an optimized CuAAC reaction. The synthesized glycoconjugates comprised one mono-functionalised derivative, two 1,2- or 1,3-divalent regioisomers, one trivalent and three tetravalent isomers with varying topology . As determined by ITC, the binding affinities were enhanced with increased ligand valencies. Compared to a monosaccharide probe, a trivalent conjugate displayed an improvement in the binding of 73- fold (24-fold per ligand), whereas the tetravalent conjugates were considerably more potent. Especially, compound **4** (**Fig.6**), which presents carbohydrates on both faces of the calixarene. This was the most potent compound with a dissociation constant of 176 nM, a 200-fold potency increase per galactose. The surface plasmon resonance (SPR) studies in addition confirmed the findings from ITC. Molecular modelling suggests that the affinity achieved by compound **4** in this alternated 1,3-topology was a chelate-based binding of two target lectins at the same time. This specific topology of compound **4** to present two galactose residues at opposite faces appeared to bridge two binding sites in both intra and inter-protein. The inhibitory potency was shown to be strongly dependent on both the valency and the topology.

A recent study showed that fullerene hexakis-adducts bearing peripheral functional groups were suitable for multivalent ligand preparation. Its unique globular topology allowed each residue on the periphery to be almost equidistant to its neighbors throughout the molecular architecture. Regarding that, the most promising fullerene-based glycocluster **5** was constructed with twelve appending galactose epitopes.[38] In an ELLA (Enzyme Linked Lecin Assay), a dramatic multivalency effect was observed of which compound **5** displayed a significant “glycoside cluster effect” with a 5500-fold increase in binding when comparing a monovalent carbohydrate reference probe. The observed remarkable inhibition

could be due to the multivalent design of the molecule but also to beneficial additional contacts with the lectin binding pocket through the triazole ring. Indicated by LecA crystal structure, a hydrophobic interaction between the aglycone part of the bound galactose and amino residues (Tyr36, Pro38, His50, and Pro51) around the binding pocket was preferred.[39] This was also in agreement with what had been found in compound **3**. A nanomolar ($IC_{50}$= 40 nM) and a submicromolar ($IC_{50}$= 0.36 μM) performance of compound **5** in LecA adhesion inhibition were profiled by HIA and SPR assays, albeit with some differences depending on the analytical technique.

## 3.2. Inhibition of Cholera Toxin ($AB_5$ toxin)

Cholera is caused by the cholera toxin (CT) produced by *Vibrio cholerae* and accounts for around two million deaths annually.[40] Especially, cholera deaths are still a major problem in remote areas where effective treatment is unavailable. Both cholera toxin (CT) and the closely related heat labile enterotoxin (LT) produced by enterotoxigenic *Escherichia coli* (ETEC) are still a major threat to global human health. CT represents a well-defined multivalent protein, which consists of a single disease-causing A-subunit that is surrounded by five lectin-like B-subunits (CTB). The ring-shaped homopentamer B-subunits are responsible for attachment of the toxin to the intestinal surface by recognizing the exposed GM1-oligosaccharide (GM1os) moieties (**Fig.7**) which then facilitates the entry of the toxin into the cell by endocytosis.[41] In the initial course of the disease, antibiotics are given to shorten the illness, lessen the diarrhoeal purging, decrease the need for rehydration fluids and shorten the hospital stay. Unfortunately, prophylactic use of antibiotics greatly increases the risk of the development of resistance. In last two decades, it was reported that strains resistant to antibiotics like tetracycline, ampicillin, kanamycin, streptomycin, sulphonamides, trimethoprim, and gentamicin have appeared in several cholera endemic countries.[42] Therefore, with a growing antibiotic resistance increases the need for more effective, therapeutic and cost-effective agents. Multivalency is a key feature of CT adhesion to the cell-surface and has thus provided the primary focus for creating multivalent CT inhibitors. The goal of taking advantage of the multivalency effect in combination with chemical synthesis lead to the assembly of GM1os in multivalent versions. Multivalent glycoconjugates based on different

carrier platforms like pentacyclen, dendrimers, calixarenes, corannulenes, and even the CTB protein itself have successfully been demonstrated to inhibit CTB binding.[44, 45, 49, 51, 52]

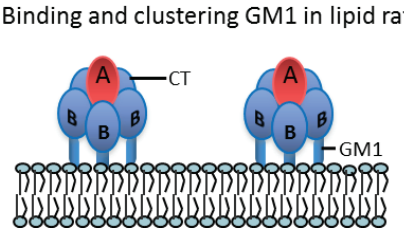



**Figure 7**. Cholera Toxin complex with GM1 pentasaccharide. Adapted from reference 43 for the figure.

Pioneering work on LT and later also CT inhibition by multivalent carbohydrates was done by Fan and co-workers.[44a] In their approach, they used a series of galactosylated multivalent ligands with different linker lengths. Five copies of non-branched monovalent and branched bivalent galactose fingers were attached to a semi-rigid pentacyclen core with extended linear flexible linkers with various repeating units (n = 1 to 4). All synthesized branched decavalent and non-branched pentavalent glycoconjugates were tested for their inhibitory ability in an ELISA assay. In terms of multivalency effect, a million-fold improvement in inhibitory power ($IC_{50}$= 40 nM) was observed for decavalent compound **6** (**Fig.8**) with four repeated linker units (n = 4) in comparison to the monovalent galactose reference, while the non-branched pentavalent showed a ten times lower potency. The compounds with lower number of linker units exhibited less inhibitory potency. Further DLS analysis showed that this decavalent ligand **6** was capable of binding to more than one target protein molecule. It was observed that the highly potent compound **6** formed a 1:2 ligand-CT complex.

In our group,[45] we investigated the inhibition of CT as well. In both pentacyclen and cyclic peptides based multivalent CT inhibitors design, the important concept of effective linker length to match inter-binding site distance was also applied in our case.[44b, 46] The multivalent GM1os-dendrimers with varied valencies (2, 4, and 8) were thereby prepared (**Fig.8**). Evaluated by an ELISA-type assay, the inhibition potency of compounds was increased with valency. The divalent compound **7** was almost 10,000 times more potent than the monovalent reference with an $IC_{50}$ in low

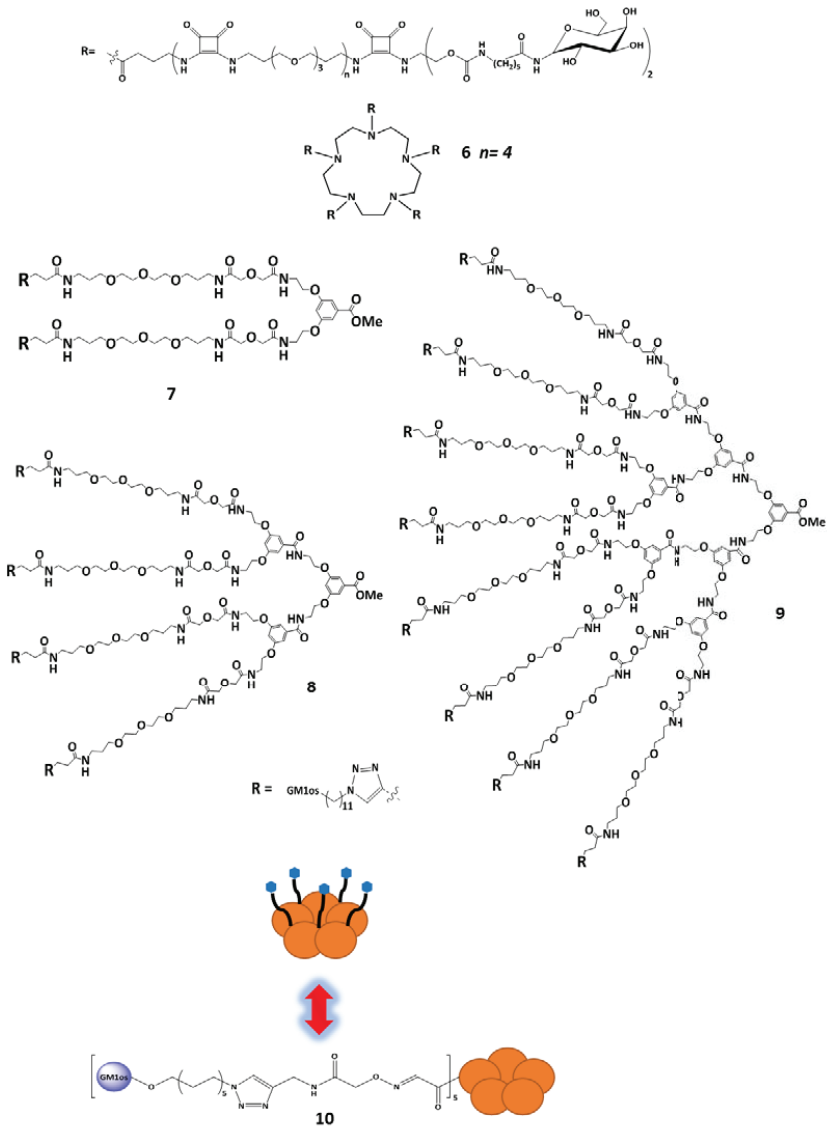



**Figure 8.** Multivalent presentation of CT inhibitors.

nanomolar range ($IC_{50}$= 2 nM). An additional increase was observed in tetravalent system, compound **8** gave an $IC_{50}$ of 0.23 nm. The octavalent compound **9** exhibited the best binding affinity with an $IC_{50}$ value of 50 pM, which is over 5 orders of magnitude more potent (relative inhibitory potency of 47,500 per ligand) compared to a monovalent GM1os derivative. All these results clearly revealed the strong multivalent binding to CTB. Studies by atomic force microscopy (AFM) and analytical

ultracentrifugation (AUC) indicated that a ligand-mediated protein aggregation might account for the unprecedented potency gain of the CTB inhibition.[47] In a follow-up study by our group[48], a simplified galactoside linked to the same multivalent system was probed. The octavalent ligand was shown to exert a 20,000 fold tighter binding when comparing to the weak monovalent binding and could inhibit as effective as a monovalent GM1 derivative due to the multivalency effect.

An attractive strategy to a multivalent inhibitor based on the use of an inactive mutant CTB protein as a scaffold was recently reported (**Fig.8**).[49] As mentioned by Fan et al., precisely matching the size and spacing of the ligands to the binding sites of CTB can optimize the inhibitory potency. The size and configuration of an inactivated toxin with an appended ligand to each subunit is a precise fit that matches the target CTB protein. Based on the work of Jobling and Holmes[50], a nonbinding mutant of CTB (W88E) was designed, where five N-terminal vicinal amino alcohol groups of threonine residues could be selectively oxidized to an aldehyde and then reacted with five GM1os ligands containing a highly reactive oxyamine tail. The oxime ligation resulted in the homopentavalent ligand **10**. An enzyme-linked lectin assay (ELLA) was performed to determine the inhibitory potency of the neoglycoprotein. Pentavalent ligand **10** exhibited an exceptionally low $IC_{50}$ value of 104 pM, which corresponded to a 5100-fold enhancement compared to monovalent GM1os or a 14300-fold enhancement compared to a monovalent GM1 azide derivative. The combined DLS, SV-AUC, and SDS-PAGE studies confirmed that the protein-based pentavalent ligand **10** binds to CTB in a 1:1 ratio forming a hamburger-like protein heterodimer.

Recently, a set of synthesized pentavalent GM1os-presenting multivalent binders were reported by Zuilhof and co-workers[51], based on a 5-fold symmetrical sym-substituted corannulene scaffold with attached PEG-linker units (n = 1, 2, 4). In the inhibitory efficiency study, compound **11** (**Fig.9**) showed the lowest $IC_{50}$ value of 5 nM with an inhibition potency that is nearly 4000 times stronger than that of monomeric GM1os reference. Possibly a lack of accessibility to binding cites with the shorter linker (n=1) and an entropically less favourable linker folding occurred with the longer linker (n =4) inducing an $IC_{50}$-value increase was observed in both

cases. An underlying supramolecular self-aggregation of the corannulene-GM1os derivatives in competition with binding to CTB was suggested for the low nanomolar $IC_{50}$ of compound **11** which was significantly higher than obtained from another study by the same group. In that case a calix[5]arene-based scaffold was made, to which five GM1os units were linked.[52] An improved $IC_{50}$ value (0.45 nM) was determined by the same ELISA inhibition test.

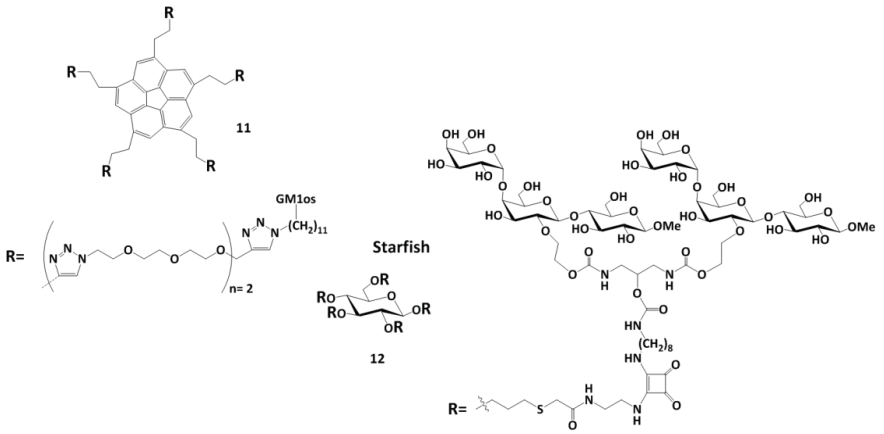



**Figure 9.** Multivalent presentation of CT inhibitors.

A related study, but now aiming to inhibit a different $AB_5$ toxin was performed by the Bundle group[53], where a decavalent glycocluster compound **12**, named "Starfish", was reported (**Fig.9**). Two copies of the natural $P^k$-trisaccharides ligands [54] were linked to the five arms of a central glucose core molecule furnishing a pentameric display. The inhibition assay was carried out with Shiga-like toxins (SLT) belonging to the same family of $AB_5$ toxins as the cholera toxin (CT) and the heat-labile toxin (LT).[55] SLTs can cause serious complications such as acute renal disease and failure in infected patients.[56] Compound **12** exhibited more than a million-fold increase in inhibition over a monovalent reference compound with a subnanomolar activity ($IC_{50}$= 0.4 nM) to SLT1 and a low nanomolar activity to SLT2 ($IC_{50}$=6 nM), respectively. Evidenced by a crystal structure study, each "Starfish" molecule was unexpectedly complexed with two toxin pentamers.

## 3.3 Inhibition of *Streptococcus suis*

*Streptococcus suis* is a Gram-positive bacterium, and an important cause of meningitis, pneumonia and sepsis in pigs. It is also zoonotic and can cause meningitis with possible residual deafness in humans.[57] The massive prophylactic use of antibiotics has contributed to the emergence and spread of antibiotic resistance. The incidence of drug resistance evolved by *S.suis* is clear.[58] Therefore, novel means to prevent and treat infections and to combat the antimicrobial resistance are needed. As *S. suis* requires carbohydrate-specific adhesion prior to the infective process, an anti-adhesion therapy based on the inhibition of bacterial attachment to carbohydrate receptors is attractive. There are two types of carbohydrate-binding *S. suis* strains, $P_N$ and $P_O$. Both of them recognize the terminal Galα(1-4)Gal (galabiose) sequence but with some different specificities.[59, 60] On the basis of this difference, a library of galabiose derivatives had been screened, of which two modified galabiosides were identified as improved *S. suis* monovalent inihibitors.[60] However, the efficiency of the monovalent inhibitor could not compete with a multivalent presentation. Important progress has been made towards multivalent galabiose ligands as potent anti-adhesive agents.

The first investigation was conduct by the Magnusson group.[61] They reported the synthesis of a series of a mono to tetra galabiosylated conjugates and their inhibition potency was determined by HIA assay of *S. suis*. With the exception of compound **13** (**Fig.10**) which showed nanomolar potency, all the other dimeric galabiosides were only slightly more active than the reference monomeric TMSEt galabioside.[62] The largest improvement in inhibitory potency was displayed by the tetravalent analogue with the longest spacer **14** (**Fig.10**) with a several hundred fold enhanced potency against two galabiose-specific strains of *S. suis*, type $P_N$ (628) (MIC= 3 nM) and type $P_O$ (836) (MIC= 2 nM). The results show a clear connection between inhibitory potency and the number of galabiose units present on the inhibitor. The effective anti-adhesion properties of both compound **13** and compound **14** may be connected to the presence of a significant linker. Lower inhibitory potencies were found for inhibitors with shorter linkers.

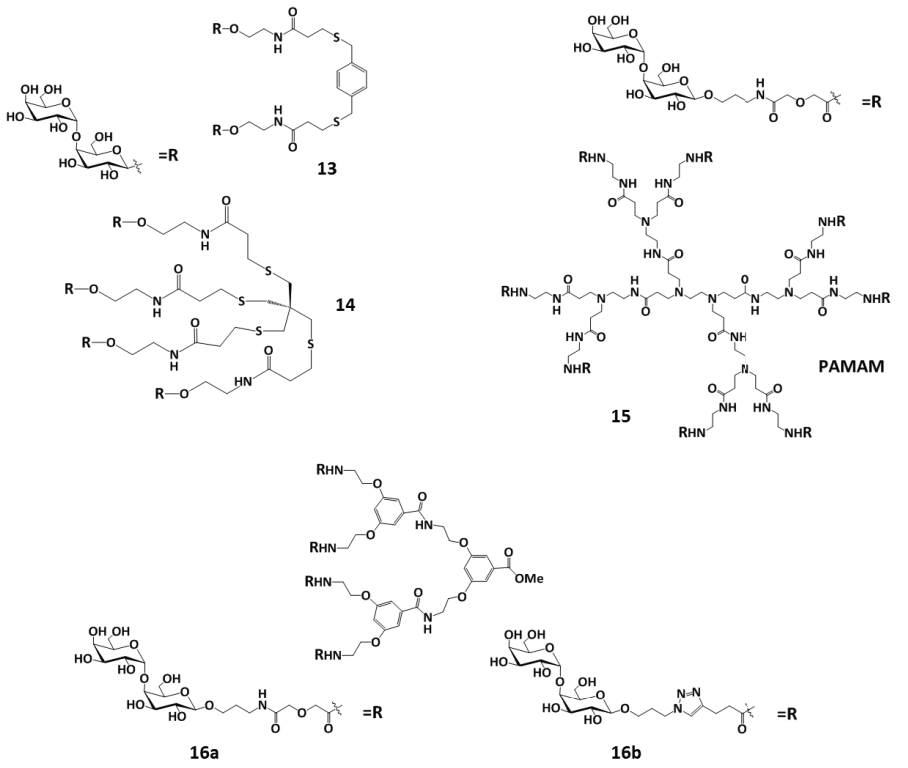



**Figure 10.** Structures of *S.suis* inhibitors.

One of the most striking examples with high biological potential was illustrated by our group.[63] General carboxylic acid-bearing galabiose building blocks were coupled to a poly-(amidoamine) (PAMAM)-based dendritic scaffold. Significant multivalency activitiy of compound **15** (**Fig.10**) was determined by an HIA experiment. It showed a minimal inhibitory concentration in the subnanomolar range (MIC= 0.3 nM) for *S. suis* strain $P_N$ 628 and a low nanomolar MIC of 2 nM for strain $P_O$ 836, which made **15** over 3000-fold and 800-fold more potent than the corresponding monovalent analogue. The results were further confirmed by an SPR assay, in which compound **15** presented the lowest $IC_{50}$ value of 7 nM. The achieved lower MICs for the multivalent compounds was a result of a multivalent or cluster effect.[64] Additionally, two similar studies[63, 65] were performed with the aromatic dendritic scaffolds (**Fig.10**) with a varing number of galabiose units (n= 2 to 4 and 2 to 8) that were conguated via peptide coupling and 1,3-dipolar cycloaddition, respectively. In first set, it was found that inhibiton potency was increased with the increasing valency of

inhibitors. The strongest inhibition power was observed for compound **16a** with an $IC_{50}$ value of 7 nM, which was 256 times more potent than the mono reference. In second case, an optimum in binding affinity was reached by compound **16b** containing four galabiose units. It could efficiently inhibit *S. suis* with an MIC value of 2.4 nM.

Without any doubt, all these unprecedented results stressed the influence of multivalency in bacterial adhesion and bacterial toxin inhibition. As evidenced in all multivalent systems, the synthetic strategies are powerful methods to control the valency and topology of the ligands, the nature and effective length of the linkers, and the architecture of the scaffolds, which in turn impacts the resultant binding affinity of multivalent inhibitors with their binding targets. These synthetic multivalent glycoconjugates have strong inhibitory effects on bacteria and their toxins therefore have great potential to serve as adhesion inhibitors and therapeutic agents. The multivalency strategy is further strongly supported by using multivalent glycoconjugates in other carbohydrate-based therapies e.g. in viral and fungal infections and even for *glycosidases* inhibition. As depicted in **Fig.11**, compound **17**, a set of β-(1,3) glucan oligosaccharide conjugates to the $CRM_{197}$ carrier protein was proven to be a strong inhibitor of fungal adhesion of *Candida albicans* to human epithelial cells. *C. albicans* is a major cause for hospital-acquired systemic infections in immunocompromised patients.[66] Recently, a trivalent ligand **18** was reported as an entry inhibitor that binds specifically to influenza hemagglutinin H5 of avian influenza with high nanomolar affinity.[67] Besides pathogenic adhesins, glycosidases are important therapeutic targets. Inhibition of these enzymes can disrupt the biosynthesis of oligosaccharides and thus interference in carbohydrate processing could has enormous therapeutic potential in many diseases such as diabetes, Gaucher disease, HIV, *etc*.[68] The cyclodextrin(CD)-based iminosugar cluster **19** revealed an impressive inhibitory activity and selectivity for α-mannosidase in the nanomolar range ($K_i$=22 nM).[69] All these intriguing findings indicate the possibility that the multivalency strategy could be employed as a major approach in the inhibition of a given carbohydrate-protein interaction. However, in spite of the roaring success of multivalent glycoconjugates, there is still room for substantial

improvement. With respect to the weak-binding of monovalent ligands, both high-throughput screening and fragment-based design allow the ligand to be further functionalized with other chemical groups to improve its individual binding properties. The important monovalent ligands will then collectively influence the multivalent compounds' performance. What's more, it is also possible to introduce an improved template or scaffold molecule to enhance the binding affinity, where the template molecule could organize the ligands for optimal binding to the target protein.[70] More importantly, with a deeper understanding of the protein-carbohydrate binding domain, more precise multivalent system can be created. The structure-based design of multivalent ligands provides a big opportunity to develop high affinity protein inhibitors in a variety of biological and medical applications.

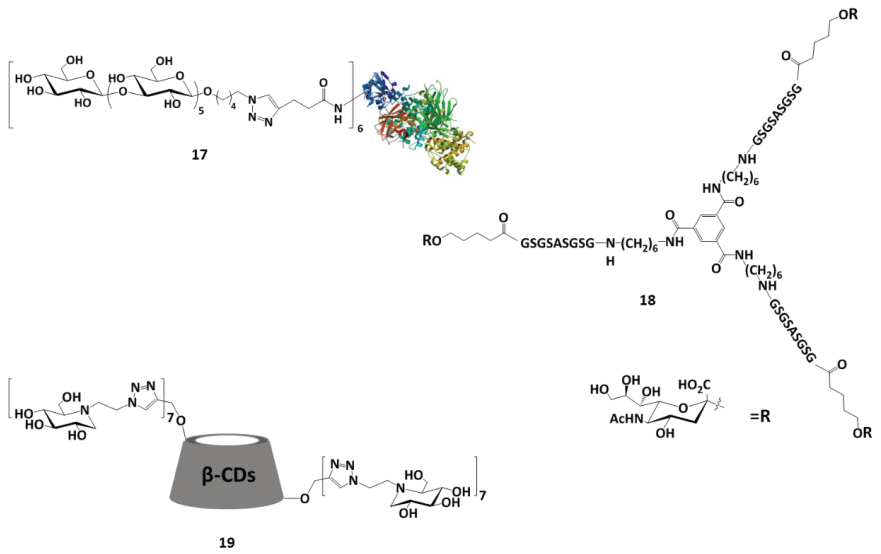



**Figure 11**. Multivalent inhibitors of other therapeutic use.

## 4. Glyconanoparticles and multivalency effect

Nanoparticle (NPs) is a collective term given for any type of objects with sizes within the 1 to 100 nm length scale.[71] As such certain polymers or spherical metal aggregates can also be classified as nanoparticles.[72] Both polymeric and metal nanoparticles have already shown distinct advantages in pharmaceutical research applications such as easy preparation and purification, tunable physiological

properties e.g. water solubility, cytotoxic activity, increased stability during storage and administration, *etc*. Their sizes, their surface properties, their morphologies, and the composites of the materials could be easily optimized to achieve desired biological activities.[73a] In particular, the multifunctionality can be further taken advantage of attaching a second molecule of interest onto the same nanoparticle. Moreover, their chemical versatility can be tailored to be compatible with a wide range of chemical strategies. Regarding all these unique properties together, NPs are widely used for *in vitro* and *in vivo* imaging, drug-delivery, and bio-molecule targeting.[73b] This opens up a wide range of possibilities for incorporating biologically relevant molecules e.g. carbohydrates to nanoparticles to functionalize these platforms. Nanoparticle dimensions can be comparable to those of biomacromolecules. Their globular shape makes them suitable for the presentation of oligosaccharide molecules of high valency to mimic the glycocalyx surface of cells.[74]. Featured with a high density of the surface ligands and large specific surface area, NPs are an ideal multivalent platform for multivalent interactions with suitable target proteins and pathogen surfaces. Glycopolymers have already been reported to exhibit larger multivalency effects in some cases than other multivalent saccharides.[75] The high density of binding sites on nanoparticle increases the rate of clustering.[76] This kind of polyvalent agent in particular has drawn attention for potential medical applications. Recent research on grafting information-carrying ligands such as carbohydrates on nanoparticle surfaces, specific targeting has been achieved.

Following this principle, Haag *et al.*[77] described a polyvalent system consisting of a globular biocompatible polymeric core (hyperbranched polyglycerol, hPG) with a large number of terminal Man (mannose) residues (n= 10 to 60) on their periphery. Due to its hyperbranched architectures and high density of peripheral functional groups, it was envisaged to use hPG as a multivalent tool to explore these mannose-coated glycopolymers as cell-surface glycans mimics. The binding affinity was evaluated by a competitive SPR-based binding assay with Concanavalin A (Con A), a mannose-specific lectin, which is extensively used in model studies with glycoarchitectures, to monitor their binding potency. By increasing the level of Man

functionalization, the $IC_{50}$ value was decreased from micro-to-nanomolar range. A marked increase in inhibitory potencies ($IC_{50}$= 35 nM) was achieved by compound **20** appending the most Man units (n= 60), whereas the corresponding monovalent mannoside (methyl-Man) requires millimolar concentrations for the inhibition.

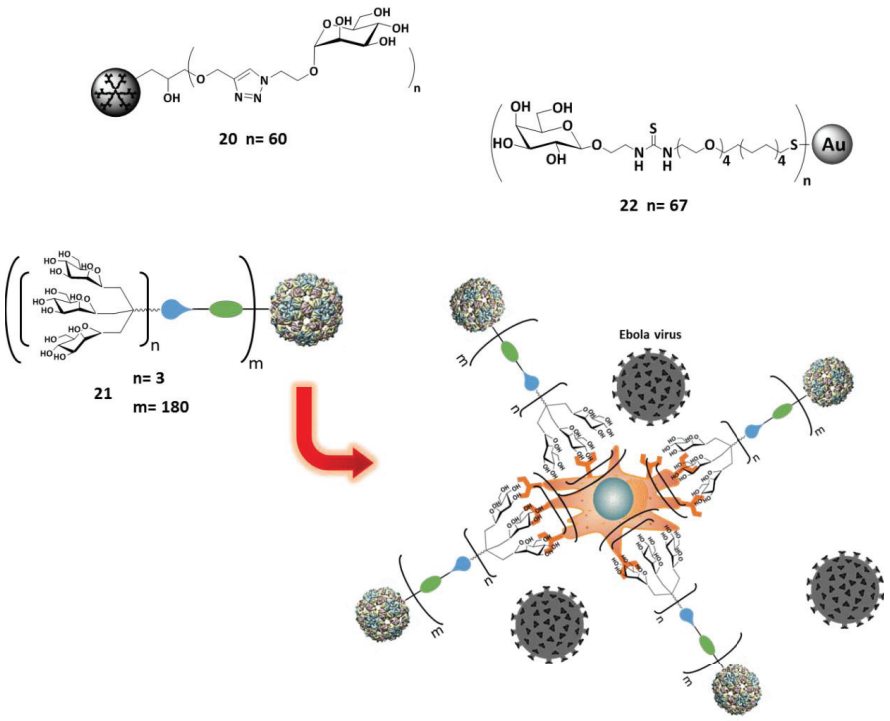



**Figure 12.** Structures of glyconoparticles and schematic infection model of compound **21** with Ebola (shown in black spotted ball) infection (adapted from reference 78 for the figure).

This ligand polyvalency effect was even further increased by Davis *et al*.[78] An elegant design was using a polymeric assembly of multivalent glycan-presenting protein monomers. These synthetic glycol-constructs were coated with a very high number of glycans (1,620 terminal mannosyl moieties in total) yet distributed in a homogeneous manner. The diameters of the resulting virus-like glycodendrinanoparticles **21** were observed up to 32nm. These glycol-assemblies could potentially mimic the virus particles e.g. HIV/SIV in both size and their highly glycosylated surface thereby block the viral infections. DC-SIGN one of the most important pathogen recognition receptors, recognizes mannoses in a multivalent

manner.[79] By using DC-SIGN as an entry point some viruses e.g. HIV-1 and Ebola are capable of escaping from the processing and degradation events carried out by the immune defence machinery at antigen-presenting cells[9] Due to the therapeutic importance of DC-SIGN regarding pathogen invasion. They applied this idea here to the inhibitory immunomodulation of a DC-SIGN-pathogen glycoprotein interaction. Tested in an Ebola viral infection model,[80] compound **21** displayed a significant antiviral activity to block the infection process at a picomolar concentration ($IC_{50s}$= 910 pM). Strong binding enhancements resulting from a high ligand presentation was demonstrated. The prospect of using such ultrahigh valent ligands for therapeutic intervention is promising.

As one of the most applied metal NPs, gold NPs can be easily synthesized with low polydispersities in a range of sizes. They have shown to exhibit good biocompatibility and are intrinsic non-toxic to human cells.[81] Carbohydrate functionalized gold nanoclusters (glyconanoparticles, GNPs) have shown their promising potential to intervene in carbohydrate-mediated biological processes. For this reason, they have been widely produced and explored in biomedical applications. Recently, a systematic study was done by Pérez and co-workers[82] to synthesize a set of galactose functionalised GNPs with varying ligand valencies (n= 12 to 67) and varied ligand presentation densities (11 to 100%). Hybrid galatose-NPs were evaluated by HIA assay as ligands for LecA binding. Multivalency effects were observed when compared to the monovalent ligand. The data from the experiments revealed that the inhibitory potency was amplified with increasing sugar valency and presentation density with decreasing $IC_{50}$ values from the micro- to the nanomolar range. Compound **22** with the highest ligand valency afforded a significant increase in binding affinity with a $K_d$ of 50 nM per sugar, which is nearly a 3000-fold affinity enhancement in comparison to the monovalent reference probe. The preference of PA-IL for high ligand valencies and larger presentation densities could be explained by a combination of structural preferences and statistical rebinding.

All in all, NPs as scaffolds for carbohydrate ligand display have recently emerged and glyconanomaterials with their strong binding properties are great potential as

therapeutics as well as other medical uses e.g. biomedical imaging and detection (signal enhancement).

## 5. Outline of the thesis

In this chapter, a description was given of multivalent glycoconjugates, which have emerged as a family of potential anti-infective therapeutics. Multivalency effects are playing a critical role in most of these advanced developments. With a growing knowledge of the nature of protein–carbohydrate interactions, more and more fine-tuned multivalent glycoconjugates will be further invented in the future. This thesis contributes to research on the fundamental properties of multivalent systems to gain a deeper understanding of multivalency effects in biological systems. It covers improved synthetic methods for the construction of multivalent carbohydrates and detailed studies involving approaches towards activity optimization.

As indicated by the complex of LecA with galactose ligands, the short distance between the two adjacent binding sites is 26 Å. Based on this structural knowledge, initial efforts had been focused on the development of symmetrical divalent galactose inhibitors with rigid spacers to match this geometry. The present study, is based on the presence of two aspartic acid residues (Asp 47) on the path of the spacer of our best divalent ligand. The presence of these two acids creates two negatively charged groups on LecA surface. Therefore, it was attractive to introduce other functional groups to afford additional beneficial interaction with these groups to tighten the binding between the ligand /spacer and the proteins. In **Chapter 2**, the attempts of spacer optimization were carried out to functionalize the 6-OH groups at both spacer glucose ends. Advances in chemical synthesis techniques enabled us to introduce diverse chemical groups on the spacer including amine, pyridinium, phenyl and even carboxylates as negative control. The inhibitors obtained were tested against LecA.

In **Chapter 3**, the synthesis of GM1-based pentavalent glycodendrimers against cholera toxin (CT) inhibition is described. A series of di-, tetra- and octavalent glycodendrimers as CT inhibitors had previously been synthesized. Although extraordinary inhibition potency had been achieved by presenting four or eight

oligosaccharide epitopes, the observed protein aggregation caused by these valency mismatched inhibitors let us wonder if this mismatch was needed for a high affinity gain. To answer the question a matched pentavalent GM1-based ligand was prepared by adding a $5^{th}$ arm to the existing tetravalent scaffold. The multivalent ligand was synthesized in good yield by coupling of alkyn-linked GM1 probes to the azide-bearing dendritic scaffold via CuAAC. An ELISA assay was conducted to establish its potency profile followed by a SV-AUC analysis, which allowed us to observe the aggregation behaviour of the ligands in detail. For comparison, the mismatched tetravalent GM1-dendrimer was also prepared.

Chemoenzymatic synthesis of the GM1-ganglioside is a complicated and time-consuming and expensive process. We therefore also used a simplified monosaccharide, *m*-nitrophenyl-α-D-galactoside (MNPG, previously reported as a monovalent CT inhibitor), where the galactose moiety can serve as an 'anchor' and the nitro group of the hydrophobic aglycone contributes to the binding. Impressively, for such a relatively simple compound, MNPG has a sizeable micromolar affinity for CT, although it is much less potent than the original nanomolar GM1 ligand. To circumvent this drawback, a multivalency strategy was implemented. Hyperbranched polyglycerols (HPGs, >10 nm) form highly branched, narrowly disperse, three-dimensional nanoparticles. Their high density of peripheral functional groups make them ideal scaffolds for the multivalent presentation of saccharides. In **Chapter 4**, we were inspired to build up a polyvalent construct presenting many copies of the MNPG ligand and explore their multivalency effect in CT inhibition. These feasible and cost-effective sugar ball-like molecules might form an avenue for therapeutic intervention.

The inhibition of *Streptococcus suis* with galabiose-containing derivatives had been studied before *S. suis* adhesion had been efficiently inhibited by multivalent galabiose dendrimers at nanomolar concentration. For *S.suis* there are two galabiose-binding strains: $P_O$ and $P_N$. It was found that the adhesin combining site of $P_N$ strains could tolerate larger substitutions at the HO-3' position of the terminal α-galactose in the galabiose structure, of which chemical modification could impart additional binding to *S. suis*. In **Chapter 5**, a new galabiose derivative with a

phenylurea derivatization at C-3' of the terminal galactose is presented as an improved monovalent ligand for *S. suis*. This modified ligand was linked to our effective multivalent scaffold, in order to combine these beneficial properties into a single molecule for optimal inhibition. This novel multivalent glycodendrimer might shed a new light in the development of anti-adhesion compounds for *S. suis*.

## Reference


[1] A. Frey, K. T. Giannasca, R. Weltzin, P. J. Giannasca, H. Reggio, W. I. Lencer, M. R. Neutra, *J. Exp. Med*. **1996**, 184, 1045-1059.
[2] J. Montreuil, *Adv. Carbohydr. Chem. Biochem*. **1980**, 37, 157-223.
[3] a) M. Mammen, S.-K. Choi, G. M. Whitesides, *Angew. Chem*. **1998**, 110, 2908-2953; *Angew. Chem. Int. Ed*. **1998**, 37, 2754-2794; b) J. J. Lundquist, E. J. Toone, *Chem. Rev*. **2002**, 102, 555-578; c) L. L. Kiessling, J. E. Gestwicki, L. E. Strong, *Angew Chem*. **2006**, 118, 2408-2429; *Angew. Chem. Int. Ed*. **2006**, 45, 2348-2368.
[4] a) J. B. Lowe, *Curr. Opin. Cell Biol*. **2003**, 15, 531-538; b) L. A. Lasky, *Annu. Rev. Biochem*. **1995**, 64, 113-140.
[5] a) K.-A. Karlsson, *Adv. Exp. Med. Biol.* **2001**, 491, 431-43; b) J. Holgersson, A. Gustafsson, M. E. Breimer, *Immunol. Cell Biol*. **2005**, 83, 694-708; c) A. V. Karlyshev, J. M. Ketley, B. W. Wren, *FEMS Microbiol. Rev*. **2005**, 29, 377-390; d) F. Lehmann, E. Tiralongo, J. Tiralongo, *Cell. Mol. Life Sci*. **2006**, 63, 1331-1354.
[6] N. P. Pera, A. Kouki, S. Haataja, H. M. Branderhorst, R. M. J. Liskamp, G. M. Visser, J. Finne, R. J. Pieters, *Org. Biomol. Chem.*, **2010**, 8, 2425-2429.
[7] T. Lindhorst, *Topics Curr. Chem*. **2002**, 218, 201-235.
[8] R. G. Webster, W. J. Bean, O. T. Gorman, T. M. Chambers, Y. Kawaoka, *Microbiol. Mol. Biol. Rev*. **1992**, 56, 152-179.
[9] Y. van Kooyk, T. B. H. Geijtenbeek, *Nat. Rev. Immunol.* **2003**, 3, 697-709.
[10] T. B. H. Geijtenbeek, D. S. Kwon, R. Torensma, S. J. van Vliet, G. C. F. van Duijnhoven, J. Middel, Ine L. M. H. A. Cornelissen, Hans S. L. M. Nottet, V. N. KewalRamani, D. R. Littman, C. G. Figdor, Y. van Kooyk, *Cell* **2000**, 100, 587-597.
[11] M. Sanchez-Navarro, J. Rojo, *Drug News Perspect*. **2010**, 23, 557-572.
[12] B. Ernst, J. L. Magnani, *Nature Reviews Drug Discovery*, **2009**, 8, 661-677.
[13] a) M. von Itzstein, W.-Y. Wu, G. B. Kok, M. S. Pegg, J. C. Dyason, B. Jin, T. V. Phan, M. L. Smythe, H. F. White, S. W. Oliver, P. M. Colman, J. N. Varghese, D. M. Ryan, J. M. Wood, R. C. Bethell, V. J. Hotham, J. M. Cameron, C. R. Penn, *Nature.* **1993**, 363, 418-423; b) M. von Itzstein, *Nat. Rev. Drug. Discov.* **2007**, 6, 967-974; c) C. U. Kim, W. Lew, M. A. Williams, H. Liu, L. Zhang, S. Swaminathan, N. Bischofberger, M. S. Chen, D. B. Mendel, C. Y.Tai, W. G. Laver, R. C. Stevens, *J. Am. Chem. Soc.* **1997**, 119, 681-690.
[14] K. Hata, K. Koseki, K. Yamaguchi, S. Moriya, Y. Suzuki, S. Yingsakmongkon, G. Hirai, M. Sodeoka, M. von Itzstein, T. Miyagi, *Antimicrob. Agents Chemother*. **2008**, 52, 3484-3491.
[15] Y. Hu, S. Lu, Z. Song, W. Wang, P. Hao, J. Li, X. Zhang, H.-L. Yen, B. Shi, T. Li, W. Guan, L. Xu, Y. Liu, S. Wang, X. Zhang, D. Tian, Z. Zhu, J. He, K. Huang, H. Chen, L. Zheng, X. Li, J. Ping, B. Kang, X. Xi, L. Zha, Y. Li, Z. Zhang, M. Peiris, Z. Yuan, *Lancet.* **2013**, 381, 2273-2279.

[16] a) R. L. Yenofsky, M. Fine, J. W. Pellow, *Proc. Nat. Acad. Sci. U.S.A.*, **1990**, 87, 3435-3439; b) H. C. Neu, *Science*. **1992**, 257, 1064-1073.
[17] Y. C. Lee, R. R. Townsend, M. R. Hardy, J. Lonngren, J. Arnarp, M. Haraldsson, H. Lonn, *J. Biol. Chem.* **1983**, 258(1), 199-202.
[18] J. E. Gestwicki, C. W. Cairo, L. E. Strong, K. A. Oetjen, L. L. Kiessling, *J. Am. Chem. Soc.* **2002**, 124, 14922-14933.
[19] W. P. Jencks, *Proc. Natl. Acad. Sci. U. S. A.*, **1981**, 78, 4046-4050.
[20] L. L. Kiessling, J. E. Gestwicki, L. E. Strong, *Angew. Chem., Int. Ed.*, **2006**, 45, 2348-2368.
[21] M. Mammen, G. Dahmann, G. M. Whitesides, *J. Med. Chem*. **1995**, 38, 4179-4190.
[22] S.-K. Choi, Synthetic multivalent molecules, John Wiley & Sons, **2004**.
[23] K. Lewis, *Nature Rev. Drug Discov*. **2013**, 12, 371-387.
[24] M. A. Fischbach, C. T. Walsh, *Science*. **2009**, 325, 1089-1093.
[25] a) Y. M. Chabre, R. Roy, *Adv. Carbohydr. Chem. Biochem.* **2010**, 63, 165-393; b) R. Bertozzi, L. L. Kiessling, *Science* **2001**, 291, 2357-2364.
[26] a) K. F. Johnson, *Glycoconjugate J*. **1999**, 16, 141-146; b) H. Remaut, G. Waksman, *Curr Opin Struct Biol*. **2004**, 14, 161-170.
[27] F. Baleux, L. Loureiro-Morais, Y. Hersant, P. Clayette, F. Arenzana-Seisdedos, D. Bonnaffe, H. Lortat-Jacob, *Nat. Chem. Biol*. **2009**, 5, 743-748.
[28] M. A. Mintzer, E. L. Dane, G. A. O'Toole, M. W. Grinstaff, *Mol. Pharm*. **2012**, 9(3), 342-354.
[29] a) P. D. Lister, D. J. Wolter, N. D. Hanson, *Clin. Microbiol. Rev.*, **2009**, 22, 582-610; b) H. C. Flemming, J. Wingender, *Nat. Rev. Microbiol*. **2010**, 8, 623-633.
[30] V. E. Wagner, B. H. Iglewski, *Clin. Rev. Allergy Immunol.*, **2008**, 35, 124-134.
[31] a) R. Loris, D. Tielker, K.-E. Jaeger, L. Wyns, *J. Mol. Biol*. **2003**, 331, 861-870; b) A. Imberty, M. Wimmerová, E. P. Mitchell, N. Gilboa-Garber, *Microbes Infect*. **2004**, 6, 221-228.
[32] M. Smadhi, S. de Bentzmann, A. Imberty, M. Gingras, R. Abderrahim, P.G. Goekjian, *Beilstein J. Org. Chem*. **2014**, 10, 1981-1990.
[33] a) C. W. Tornoe, C. Christensen, M. Meldal, *J. Org. Chem*. **2002**, 67, 3057-3064; b) V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem.* **2002**, 114, 2708-2711; *Angew. Chem. Int. Ed.* **2002**, 41, 2596-2599; c) Q. Wang, T. R. Chan, R. Hilgraf, V. V. Fokin, K. B. Sharpless, M. G. Finn, *J. Am. Chem. Soc*. **2003**, 125, 3192-3193.
[34] A. Nurisso, B. Blanchard, A. Audfray, L. Rydner, S. Oscarson, A. Varrot, A. Imberty, *J. Biol. Chem*. **2010**, 285, 20316-20327.
[35] F. Pertici, N. J. de Mol, J. Kemmink, R. J. Pieters, *Chem. Eur. J*. **2013**, 19, 16923-16927.
[36] R. U. Kadam, M. Bergmann, M. Hurley, D. Garg, M. Cacciarini, M. A. Swiderska, C. Nativi, M. Sattler, A. R. Smyth, P. Williams, M. Cámara, A. Stocker, T. Darbre, J.-L. Reymond, *Angew. Chem. Int. Ed*. **2011**, 50, 10631-10635.
[37] S. Cecioni, R. Lalor, B. Blanchard, J.-P. Praly, A. Imberty, S. E. Matthews, S. Vidal, *Chem. Eur. J.* **2009**, 15, 13232-13240.
[38] S. Cecioni, V. Oerthel, J. Iehl, M. Holler, D. Goyard, J.-P. Praly, A. Imberty, J.-F. Nierengarten, S. Vidal, *Chem. Eur. J*. **2011**, 17, 3252-3261.

[39] G. Cioci, E. P. Mitchell, C. Gautier, M. Wimmerova, D. Sudakevitz, S. Perez, N. Gilboa-Garber, A. Imberty, *FEBS Lett*. **2003**, 555, 297-301.
[40] M. Kosek, C. Bern, R. L. Guerrant, *Bull. W. H. O*. **2003**, 81, 197-204.
[41] E. Fan, E. A. Merritt, C. L. Verlinde, W. G. Hol, *Curr. Opin. Struct. Biol.*, **2000**, 10, 680-686.
[42] D. A. Sack, R. B. Sack, G. B. Nair, A. K. Siddique, *Lancet*, **2004**, 363, 223-233.
[43] V. K. Viswanathan, K. Hodges, G. Hecht. *Nature Reviews Microbiology.* **2009**, 7, 110-119.
[44] Z.-S. Zhang, E. A. Merritt, M. Ahn, C. Roach, Z. Hou, C. L. M. J. Verlinde, W. G. J. Hol, E. Fan, *J. Am. Chem. Soc*. **2002**, 124, 12991-12998; b) E. Fan, Z. Zhang, W. E. Minke, Z. Hou, C. L. M. J. Verlinde, W. G. J. Hol, *J. Am. Chem. Soc*. **2000**, 122, 2663-2664.
[45] A. V. Pukin, H. M. Branderhorst, C. Sisu, C. A. Weijers, M. Gilbert, R. M. Liskamp, G. M. Visser, H. Zuilhof, R. J. Pieters, *ChemBioChem*, **2007**, **8**, 1500-1503.
[46] Z. Zhang, J. Liu, C. L. Verlinde, W. G. Hol, E. Fan, *J. Org. Chem*. **2004**, 69, 7737-7740.
[47] C. Sisu, A. J. Baron, H. M. Branderhorst, S. D. Connell, C. A. Weijers, R. de Vries, E. D. Hayes, A. V. Pukin, M. Gilbert, R. J. Pieters, H. Zuilhof, G. M. Visser, W. B. Turnbull, *ChemBioChem* **2009**, 10, 329-337.
[48] H. M. Branderhorst, R. M. J. Liskamp, G. M. Visser, R. J. Pieters, *Chem. Commun.*, **2007**, 5043-5045.
[49] T. R. Branson, T. E. McAllister, J. Garcia-Hartjes, M. A. Fascione, J. F. Ross, S. L. Warriner, T. Wennekes, H. Zuilhof, W. B. Turnbull, *Angew. Chem. Int. Ed*. **2014**, 53, 8323-8327.
[50] M. G. Jobling, R. K. Holmes, *Mol. Microbiol*. **1991**, 5, 1755-1767.
[51] M. Mattarella, J. Garcia-Hartjes, T. Wennekes, H. Zuilhof, J. S. Siegel, *Org. Biomol. Chem.*, **2013**, 11, 4333-4339.
[52] J. Garcia-Hartjes, S. Bernardi, C. A. G. M. Weijers, T. Wennekes, F. Sansone, A. Casnati, H. Zuilhof *Org. Biomol. Chem.*, **2013**, 11, 4340-4349.
[53] P. I. Kitov, J. M. Sadowska, G. Mulvery, G. D. Armstrong, H. Ling, N. S. Pannu, R. J. Read, D. R. Bundle, *Nature*. **2000**, 403, 669-672.
[54] M. A. Karmali, M. Petric, C. Lim, P. C. Fleming, G. S. Arbus, H. Lior, *J. Infect. Dis*. **1985**, 151, 775-782.
[55] A. D. O'Brien, V. L. Tesh, A. Donohue-Rolfe, M. P. Jackson, S. Olsnes, K. Sandvig, A. A. Lindberg, G. T. Keusch, *Curr. Top. Microbiol. Immunol*. **1992**, 180, 65-94.
[56] H. Karch, P. I. Tarr, M. Blelaszewska, *Int. J. Med. Microbiol.* **2005**, 295, 405-418.
[57] H. F. Wertheim, H. D. Nghia, W. Taylor, C. Schultsz, *Clin. Infect. Dis*. **2009**, 48, 617-625.
[58] C. Palmieri, P. E. Varaldo, B. Facinelli, *Front Microbiol*. **2011**, 2, 235.
[59] S. Haataja, K. Tikkanen, U. Nilsson, G. Magnusson, K.-A. Karlsson, J. Finne, *J. Biol. Chem*. **1994**, 269, 27466-27472.
[60] J. Ohlsson, A. Larsson, S. Haataja, J. Alajääski, P. Stenlund, J. S. Pinkner, S. J. Hultgren, J. Finne, J. Kihlberg, U. J. Nilsson, *Org. Biomol. Chem.*, **2005**, 3, 886-900.
[61] H. C. Hansen, S. Haataja, J. Finne, G. Magnusson, *J. Am. Chem. Soc*. **1997**, 119, 6974-6979.
[62] K. Jansson, S. Ahlfors, T. Frejd, J. Kihlberg, G. Magnusson, J. Dahmén, G. Noori, K. Stenvall, *J. Org. Chem*. **1988**, 53, 5629-5647.

[63] J. A. F. Joosten, V. Loimaranta, C. C. M. Appeldoorn, S. Haataja, F. A. El Maate, R. M. J. Liskamp, J. Finne, R. J. Pieters, *J. Med. Chem*. **2004**, 47, 6499-6508.
[64] J. J. Lundquist, E. J. Toone, *Chem. Rev*. **2002**, 102, 555-578.
[65] H. M. Branderhorst, R. Kooij, A. Salminen, L. H. Jongeneel, C. J. Arnusch, R. M. J. Liskamp, J. Finne, R. J. Pieters, *Org. Biomol. Chem.*, **2008**, 6, 1425-1434.
[66] R. Adamo, Q. Y. Hu, A. Torosantucci, S. Crotti, G. Brogioni, M. Allan, P. Chiani, C. Bromuro, D. Quinn, M. Tontini, F. Berti, *Chem. Sci.*, **2014**, 5, 4302-4311.
[67] M. Waldmann, R. Jirmann, K. Hoelscher, M. Wienke, F. C. Niemeyer, D. Rehders, B. Meyer *J. Am. Chem. Soc*. **2014**, 136, 783-788.
[68] a) T. M. Gloster, G. J. Davies, *Org. Biomol. Chem*. **2010**, 8, 305-320; b) T. Kajimoto, M. Node, *Curr. Top. Med. Chem*. **2009**, 9, 13-33.
[69] C. Decroocq, D. Rodríguez-Lucena, V. Russo, T. M. Barragán, C. O. Mellet, P. Compain, *Chem. Eur. J*. **2011**, 17, 13825-13831.
[70] P. I. Kitov, G. L. Mulvey, T. Griener, T. Lipinski, D. Solomon, E. Paszkiewicz, J. Jacobson, J. M. Sadowska, M. Suzuki, K. Yamamura, G. D. Armstrong, D. R. Bundle, *Proc. Natl. Acad. Sci. U. S. A.*, **2008**, vol. 105, no. 44, 16837-16842.
[71] J. Kreuter (Ed.), Colloidal drug delivery systems, vol. 66 Marcel Dekker, New York (**1994**), 219-342.
[72] S. M. Janib, A. S. Moses, J. A. MacKay, *Adv. Drug Delivery Rev.*, **2010**, 62, 1052-1063.
[73] a) E. Allemann, R. Gurny, E. Doelker, *Eur. J. Pharm. Biopharm*. **1993**, 39, 173; b) J. Jagur-Grodzinski, *React. Funct. Polym*. **1999**, 39, 99; b) R. Sperling, G. P. Rivera, F. Zhang, M. Zanella, W. Parak, *Chem. Soc. Rev*. **2008**, 37, 1896-1908.
[74] A. G. Barrientos, J. M. de la Fuente, T. C. Rojas, A. Fernández, S. Penadés, *Chem. Eur. J*., **2003**, 9, 1909.
[75] Y. Miura, *Polymer J*. **2012**, 44, 679-689.
[76] M. Domurado, D. Domurado, S. Vansteenkiste, A. De Marre, E. Schacht, *J. Controlled Release* **1995**, 33, 115-123.
[77] I. Papp, J. Dernedde, S. Enders, S. B. Riese, T. C. Shiao, R. Roy, R. Haag, *ChemBioChem* **2011**, 12, 1075-1083.
[78] R. Ribeiro-Viana, M. Sánchez-Navarro, J. Luczkowiak, J. R. Koeppe, R. Delgado, J. Rojo, B. G. Davis, *Nat Commun.* **2012**, 3, 1303.
[79] Y. Guo, H. Feinberg, E. Conroy, D. A. Mitchell, R. Alvarez, O. Blixt, M. E. Taylor, W. I. Weis, K. Drickamer, *Nat. Struct. Mol. Biol*. **2004**, 11, 591-598.
[80] C. P. Alvarez, F. Lasala, J. Carrillo, O. Muñiz, A. L. Corbí, R. Delgado, *J. Virol*. **2002**, 76, 6841-6844.
[81] S. Rana, A. Bajaj, R. Mout, V. M. Rotello, *Adv. Drug Delivery Rev.*, **2012**, 64, 200-216.
[82] M. Reynolds, M. Marradi, A. Imberty, S. Penadés, S. Pérez, *Chem. Eur. J*. **2012**, 18, 4264-4273.

# Chapter 2

## Functionalization of a Rigid Divalent LecA Ligand

Part of this chapter has been accepted for publication:

O. Fu, A.V. Pukin, H. C. Quarles van Ufford, J. Kemmink, N. J. de Mol, R. J. Pieters, *Chemistry Open*, **2015**, in press (DOI: 10.1002/open.201402171).

## Introduction

Enhancing the binding potency of carbohydrate inhibitors of protein-carbohydrate interactions is an important goal towards medicinal applications.[1] This is mostly the case because of the relatively weak interactions between proteins and carbohydrate ligands. Multivalency is a proven method to achieve this.[2] LecA is an important virulence factor of *Pseudomonas aeruginosa* involved in bacterial adhesion and biofilm formation.[3] The protein has recently become a popular target protein for the design of multivalent inhibitors.[4] The tetrameric protein promotes the adhesion of the bacteria to the tissue cell surfaces, thus facilitating subsequent steps such as cell invasion and biofilm formation. Inhibiting bacterial adhesion proteins has the potential to become a mild and less resistance-prone method to treat and prevent bacterial infections.[5] Two of the four binding sites, with specificity for galactosides, are relatively close together with a separation of ca. 26 Å.[6] This arrangement has led us to design and synthesize divalent galactoside ligands with well-defined and rigid spacers that should allow a chelation type divalent binding mode.[7] Flexible spacers are commonly used as they are forgiving of imperfect design and will usually yield sizeable potency enhancements in multivalent systems.[7] They are, however, not optimal as there will be a significant entropy loss upon binding and moreover achieving selectivity will be less likely. In our search for an optimal spacer we found compound **1** (**Fig.1a**) to be a highly potent divalent ligand with nanomolar inhibitory potency.[8a] The spacer of this structure contains direct linkages between the glucose moieties and the connected triazoles.[8] This arrangement leads to a relatively rigid structure, in which rotations of the components can take place, but the overall geometry remains mostly linear. Most importantly, good solubility in water was observed. We found that three glucose-triazole units was the optimal length for LecA inhibition, where divalent ligands with 2 and 4 units showed far inferior inhibition. All data were consistent with a chelating binding mode, especially convincing was the stoichiometry derived from ITC binding experiments. Furthermore the short linkage between the galactoside ligand and the triazole of just a single carbon proved to be a major contributing factor to the success of this compound.

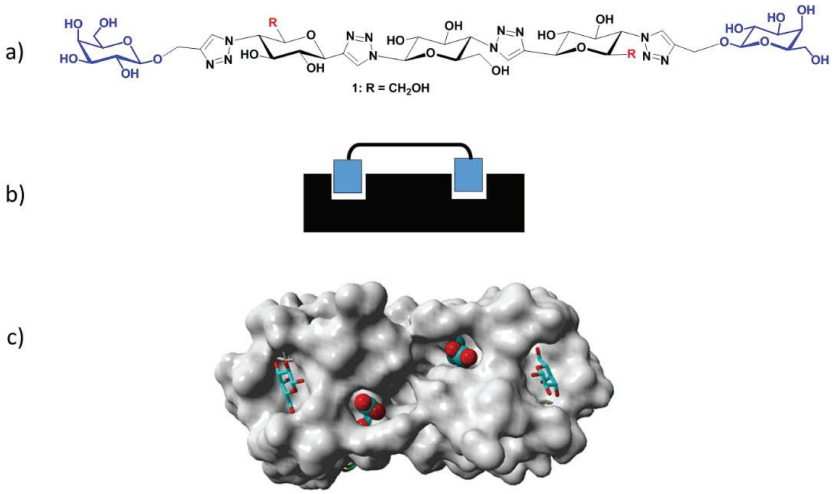



**Figure 1.** a) Structure of potent divalent LecA inhibitor **1** with the relatively rigid glucose-triazole-based spacer; b) Schematic divalent binding mode of a divalent ligand to two LecA subunis; c) X-ray structure of LecA with bound galactose moieties. The two Asp 47 carboxylates in the spacer path are shown explicitly.[6]

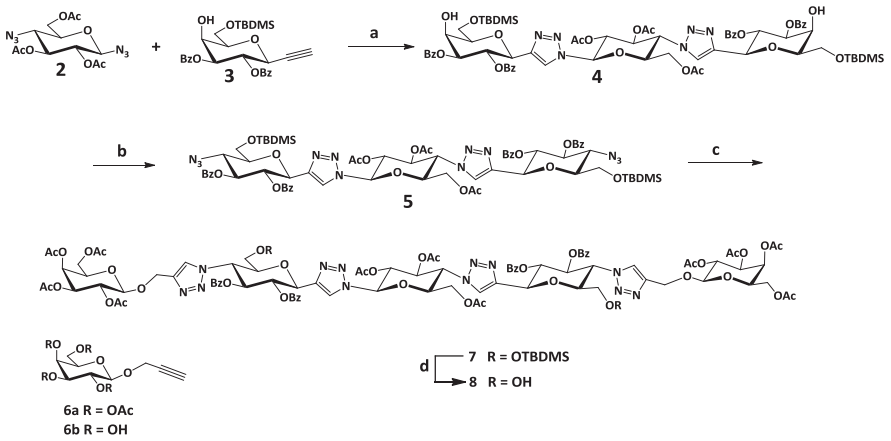
In order to further optimize the potency and to explore the principle of protein-spacer interactions we here report on our functionalization of the spacer of **1** with various functional groups. Looking at the path between the two bound galactosides that the spacer is likely to span on the LecA protein surface when chelating bivalent binding is occurring, the presence of the two carboxylates (Asp 47 in each subunit) is apparent (**Fig.1c**). These two carboxylates are likely to be in close proximity to the 6-OH group on the terminal spacer glucoside units, depending on the rotational state of the molecule. This C(6) positions can modified by using the proper protecting groups in the synthesis.

Molecular modeling was used in order to gauge whether positively charged functional groups at the C(6) positions on the terminal spacer glucoside units would be able to interact with the Asp 47 residues. Firstly, ammonium groups were used (as derived from **12**, *vide infra*). Creating a stabilized conformation with the positive charges in close proximity to the Asp 47 carboxylates was possible, and this orientation was used as the starting point of additional simulations. When running an unrestrained nanosecond MD simulation with explicit water molecules, the stabilized geometry persisted throughout the simulation. Especially, the two hydrogen bonded salt-bridges, between the Asp 47 side chains and the ammonium

groups of the protonated form of **12** remained within 3 Å and were compatible with a geometry that included two fully bound galactoside ligand units. Such a bound geometry was also possible for a compound that included a pyridinium group (as in **13**, but resulted in longer distances between the carboxylate oxygens and the pyridinium nitrogens, due to bigger steric requirements of the pyridinium units. Furthermore, while performing a similar MD simulation this structure was not maintained indeed indicating the larger steric requirements of the pyridinium group when facing the protein surface. These experiments lead to the conclusion that the introduction of ammonium groups would be the most promising. We chose to prepare a series of derivatives with various new substituents in the spacer in the hope to further enhance the LecA inhibitory potency of the molecules. Positively charged substituents were included, as reasoned above. Also negatively charged and more lipophilic groups were included that could possible benefit from interacting with other proximate parts of the protein that would not be obvious beforehand.

## Results and discussion

## Synthesis

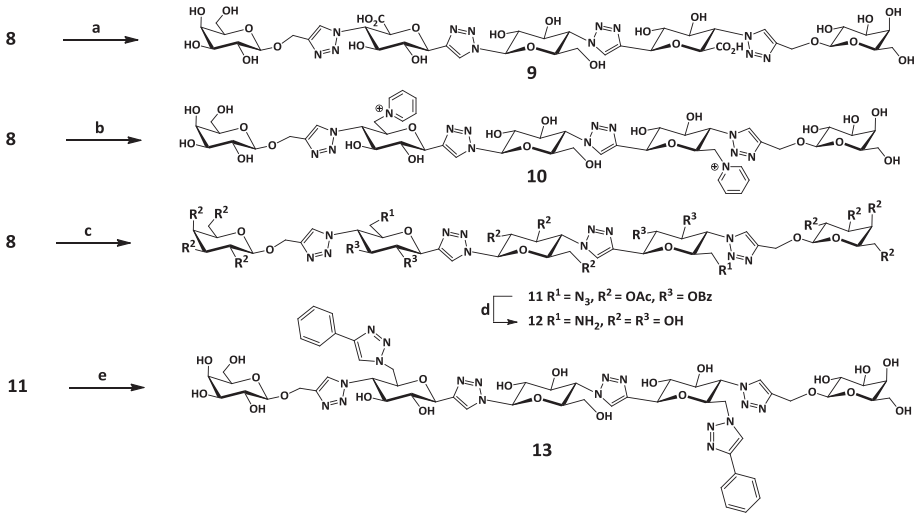



**Scheme 1.** Reagents and conditions: a) $CuSO_4 \cdot 5H_2O$, NaAsc, DMF with 10% $H_2O$, 80°C, 30 min, 85%; b) i) $Tf_2O$, 10% pyridine in $CH_2Cl_2$, 0°C, 3 h ; ii) $NaN_3$, acetone/$H_2O$ (4:1), r.t., o/n, 98%; c) **6a**, $CuSO_4 \cdot 5H_2O$, NaAsc, DMF with 10% $H_2O$, 80°C, 30 min, 64%; d) p-TsOH, MeCN/$H_2O$ (7:1), r.t., 6 h, 83%.

As shown in **Scheme 1**, the synthesis of the modified spacers started with two previously prepared building blocks **2** and **3**.[8] These building blocks were coupled by

a double CuAAC reaction. Next, the two galactosyl axial 4-OH groups of the resulting product **4** were turned into triflates, thus enabling the subsequent substitution by azide with inversion to give **5**. The terminal azides were subsequently linked to the protected galactosyl ligand **6a** by CuAAC yielding **7**. Selective removal of the TBDMS groups then gave **8** which is ready for further functionalization.

The two hydroxymethylene groups of compound **8** were oxidized with TEMPO to carboxylic acids. Removal of the protecting groups gave **9**. Reaction of **8** with triflic anhydride in the presence of pyridine followed by Zemplén deprotection gave the bis pyridinium compound **10**. Tosylation of the primary hydroxyls of **8** followed by reaction with $NaN_3$ gave the intermediate **11**. Zemplén deprotection of **11**, followed by the hydrogenation of the azido groups gave the diamine **12**. CuAAC coupling of phenylacetylene to **11** and subsequent Zemplén deprotection gave the bis-triazole **13**. All final products were purified by preparative HPLC (**Scheme 2**).




**Scheme 2.** Reagents and conditions: a) 1) TEMPO, NaOCl, NaBr, $Bu_4NBr$, $NaHCO_3/Na_2CO_3$ (pH= 9.5), 0°C, 2 h ; 2) NaOMe, MeOH, r.t., o/n, 38%; b) 1) $Tf_2O$, 10% pyridine in $CH_2Cl_2$, 0°C→r.t., 3 h; 2) NaOMe, MeOH, r.t., o/n, 44%; c) 1) TsCl, DABCO, $CH_2Cl_2$, r.t., o/n; 2) $NaN_3$, DMF, 95°C, o/n, 34%; d) 1) NaOMe, MeOH, r.t., o/n; 2) $H_2$, Pd/C, r.t., 60%; e) 1) phenylacetylene, $CuSO_4{\cdot}5H_2O$, NaAsc, DMF with 10% $H_2O$, 80°C, 30 min; 2) NaOMe, MeOH, r.t., o/n, 24%.

## Inhibition study

**Table 1.** Inhibitory potencies ($IC_{50}$) of the divalent inhibitors on LecA binding as determined by a chip-based ELISA type assay,[a] and $K_d$'s, stoichiometry and binding thermodynamic parameters obtained from ITC studies.[b]

| Compound | ELISA | ITC | n[d] | ΔH | −TΔS |
|---|---|---|---|---|---|
| | $IC_{50}$ (nM) | $K_d$ (nM) | | [kJ mol$^{-1}$] | [kJ mol$^{-1}$] |
| **1** | 1.8 | 57 (±7) | 0.5 | -49.1 | 10 |
| **9** | 31 | 68 (±10) | 0.5 | -50 | 9.1 |
| **10** | 5.2 | 89 (±7) | 0.51 | -50.1 | 9.9 |
| **12** | 19 | 56 (±3) | 0.49 | -48.7 | 7.4 |
| **13** | 4.9 | 57[c] (±7) | 0.5 | -44.4 | 3 |

[a] FITC-labelled LecA (5 mg/mL) binding to a galactoside functionalized chip surface.
[b] [LecA] = 20-40 mM; [c] 1% DMSO was present in this experiment. [d] stoichiometry

The compounds were tested in an ELISA-like assay on an array chip as previously reported.[8] Fluorescein labelled LecA was incubated with the inhibitors and exposed to a galactoside functionalized chip surface. Detection of the fluorescence allowed quantification of the binding. Monovalent ligand **6b** was previously determined to have an $IC_{50}$ in this assay of 22 μM.[8a] Clearly all divalent compounds were far more potent and showed major multivalency effects (**Table 1**). The previous best inhibitor **1** still remained the most potent one in the present series with, as before, an $IC_{50}$ in the 2-3 nM range. The pyridinium functionalized **10** and the phenylacetylene derived **13** showed only a minor drop in potency with $IC_{50}$'s in the 5 nM range. Larger potency drops of ca. an order of magnitude were observed for both the negatively charged bis-carboxylate **9** and the positively charged bis-amine **12**. Subsequently isothermal calorimetry (ITC) experiments were conducted, which confirmed the divalent binding mode in all cases with the stoichiometry n values being close to 0.5. As before,[8a] the $K_d$ values were somewhat higher than the $IC_{50}$'s from the chip-based ELISA-type assay. Furthermore the small potency differences of the ELISA-type assay were not seen in the ITC assay. All compounds showed inhibitory potencies

within a narrow range (57-89 nM). Interestingly, when looking at the enthalpic and entropic components of the binding event, the lipophilic and non-charged compound **13**, showed enthalpy-entropy compensation, with a lower beneficial binding enthalpy and at the same time a lowered entropy loss upon binding. The latter is understandable as an example of the hydrophobic effect where water molecules in an ice-like structure are liberated from the lipophilic surfaces. For this reason, some degree of lipophilic association is suggested for **13**.

## Conclusion

New derivatives of a highly potent divalent LecA ligand were prepared. It proved possible to use selectively protected 6-OH groups to build up the ligand in its protected form. Subsequently the selectively deprotectable silyl groups were removed and the resulting primary hydroxyls were converted to carboxylate groups by oxidation, to tosylates and subsequently to azides by substitution, to pyridiniums via their corresponding triflates. The azides allowed CuAAC coupling and also further reduction to the corresponding amino groups. While, it was anticipated that some of these groups would be able to take advantage of additional beneficial interactions with the nearby proteins surface, we did not observe this through enhanced inhibitory or binding potencies. This was even true for the positively charged groups, that could possibly take advantage of the nearby positioned carboxylate groups of Asp 47. Molecular modeling indicated that these interactions were geometrically possible and that they could be energetically favorable. However, a possible alternative scenario where the newly added groups point into the solution and thus away from the protein surface, is apparently more favorable in this case. In future designs, these options should be avoided. Nevertheless, the functionalization of the spacers will also be of importance to fine tune other properties of this type of ligand that are of importance for drug development, such as ADME and toxicity.

## Experimental procedure

## General

Chemicals were obtained from commercial sources and were used without further purification unless noted otherwise. Compounds **2**, **3** and **6** were synthesized

following literature procedures.[7] Solvents were purchased from Biosolve (Valkenswaard, The Netherlands). All moisture-sensitive reactions were performed under a nitrogen atmosphere. Anhydrous solvents were dried over molecular sieves of 4 Å or 3 Å. TLC was performed on Merck pre-coated Silica 60 plates. Spots were visualized by UV light and also by 10% $H_2SO_4$ in MeOH. Microwave reactions were carried out in a Biotage microwave Initiator (Uppsala, Sweden). The microwave power was limited by temperature control once the desired temperature was reached. Sealed vessels of 2-5 mL and 10-20 mL were used. Analytical HPLC runs were performed on a Shimadzu automated HPLC system with a reversed-phase column (Reprospher 100, C4, 5 µm, 250 X 4.6 mm, Dr. Maisch GmbH, Germany) that was equipped with an evaporative light scattering detector (PLELS 1000, Polymer Laboratories, Amherst, MA, USA) and a UV/Vis detector operating at 220 nm and 250 nm. Preparative HPLC runs were performed on an Applied Biosystems workstation. Elution was effected by using a linear gradient of 5% MeCN/0.1% TFA in $H_2O$ to 5% $H_2O$/0.1% TFA in MeCN or by a gradient of H2O to 30% MeCN in $H_2O$. $^1H$ and $^{13}C$ NMR spectroscopy was carried on an Agilent 400-MR spectrometer operating at 400 MHz for $^1H$ and 100 MHz for $^{13}C$. HSQC and TOCSY NMR (500 MHz) were performed with a VARIAN INOVA-500. Electrospray Mass experiments were performed in a Shimadzu LCMS QP-8000. High resolution mass spectrometry (HRMS) analysis was recorded using Bruker ESI-Q-TOF II. The proton numbering scheme of all compounds can be found in the supporting information and was used in the assignments of the signals in the NMR spectra here.

## Isothermal titration microcalorimetry (ITC)

The lectin LecA was obtained from Sigma-Aldrich and it was dissolved in buffer (0.1 M TRIS-HCl, 6 mM $CaCl_2$, pH= 7.5) and degassed. Protein concentration (between 20 and 40 µM depending on the ligand affinity) was checked by measurement of optical density by using a theoretical molar extinction coefficient of 28,000. Carbohydrate ligands were dissolved directly into the same buffer, degassed, and placed in the injection syringe (concentration range: 0.1-0.2 mM). ITC was performed using a MicroCal Auto ITC200. LecA was placed into the 200 µL sample cell, at 25°C. Titration was performed with 2.5 µL injections of carbohydrate ligands every 120 sec. Data

were fitted using the "one-site model" using MicroCal Origin 7 software according to standard procedures. Fitted data yielded the stoichiometry (n), the association constant ($K_a$), the enthalpy (ΔH) and the entropy of binding. The $K_d$ value was calculated as $1/K_a$ and T is 298 K. Each ligand test was duplicated.

## LecA inhibiton assay

Lectin LecA was FITC labelled according to a literature procedure.[9] Microarray experiments were performed by using a PamChip array run on a PamStation 12 instrument (Pam-Gene B.V., 's Hertogenbosch, The Netherlands). Data were obtained by real-time imaging of the fluorescence signal by a CCD camera. Images were analyzed by using BioNavigator 6 software (Pam-Gene). Each array slide contains spots in duplicate. The fluorescence intensities were expressed in arbitrary units and the relative intensities were the average of the two duplicate spots. Aliquots of a solution of FITC-labelled LecA (5 µg $mL^{-1}$ for all tested compounds) in HEPES/PBS buffer (10 mM HEPES, 100 mM NaCl, 0.1% BSA. pH= 7.4), containing different concentrations of the inhibitors were incubated for 1 h at r.t. and subsequently added to the galactoside-functionalized chip. The binding process was monitored for 2 h and the end values of the fluorescence detection were taken for the determination of the $IC_{50}$ by using Prism 5 (Graphpad Software, Inc.).

## Molecular Modelling

All molecular modelling studies were performed using the molecular-graphics, -modeling and -simulation program Yasara version 13.9.8. The bivalent ligands were first constructed in Yasara as isolated molecules. Subsequently, the complex with LecA was built by superposition of one of the galactose units of the ligand with a bound galactose of one of the subunits of the LecA crystal structure (PDB ID: 1OKO). The other galactose unit of the divalent ligand was then pulled into the adjacent galactose binding site of LecA by restraint MD using distance restraints based on the position of the bound galactose present in the X-ray structure with respect to a number of LecA residues. Possible electrostatic interactions between two positively charged amino and pyridinium substitutions in the linker region with two Asp 47s of LecA were investigated in more detail. In order to induce these interactions the subunits comprising the substituted moieties were first rotated and subsequently

the nitrogen atoms of respectively the amino and the pyridinium group were restrained to bring them in close proximity to the carbon atoms of the carboxylic acid moieties of Asp 47. After restrained MD the molecules were subjected to a 1000 ps free MD simulation in water.

## General procedure of micro-wave assisted "click reaction"

To a solution of the azide and alkyne compounds in DMF with 10% $H_2O$, was added $CuSO_4 \cdot 5H_2O$ and NaAsc. The mixture was then heated by microwave irradiation at 80°C for 30 min. When the mixture cooled to r.t., the copper salts were removed by a resin (Cuprisorb) and the solvents were removed under reduced pressure. The residue was purified by silica gel chromatography to afford the corresponding 1,2,3-triazole product.

## Compound 4

A 'click reaction' of a mixture of compound **2** (0.253 g, 0.71 mmol), compound **3** (0.83 g, 1.63 mmol), $CuSO_4 \cdot 5H_2O$ (0.06 g, 0.24 mmol) and NaAsc (0.096 g, 0.48 mmol) in DMF (5 mL) containing 10% $H_2O$ was performed following the general procedure described above to afford compound **4** (0.84 g, 85%). $^1$H NMR (400MHz, $CDCl_3$): δ, ppm 8.03-7.89 (m, 5H, 4 × CH benzoyl, H-1), 7.86-7.76 (m, 5H, 4 × CH benzoyl, H-1), 7.52-7.26 (m, 12H, 12 × CH benzoyl), 6.08 (dd, $J_{4,3}$= $J_{4,5}$= 10 Hz, 1H, H-4), 6.02-5.90 (m, 2H, H-4, H-3'), 5.81 (dd, $J_{5',4'}$= $J_{5',6'}$= 9 Hz, 1H, H-5'), 5.60 (dd, $J_{4',3'}$= $J_{4',5'}$= 9 Hz, 1H, H-4'), 5.51-5.41 (m, 2H, 2 × H-5), 5.02-4.92 (m, 2H, 2 × H-3), 4.81 (dd, $J_{6',5'}$= $J_{6',7'}$= 10 Hz, 1H, H-6'), 4.67-4.54 (m, 3H, H-7', 2 × H-6), 4.16-4.07 (m, 1H,, H-8'a), 4.07-3.93 (m, 2H, 2 × H-8), 3.91-3.82 (m, 2 × H-7), 3.70-3.60 (m, 1H, H-8'b), 3.57-3.51 (m, 1H, 6-OH), 3.35-3.29 (1H, 6-OH), 2.02, 1.65, 1.58 (s, 9H, 3 × $CH_3$, acetyl), 0.90 (s, 18H, 2 × $SiC(CH_3)_3$), 0.07 (2 × s, 12H, 2 × $Si(CH_3)_2$). $^{13}$C NMR (100 MHz, $CDCl_3$): δ, ppm 169.97, 169.10, 168.95 (3 × C=O acetyl), 166.14, 166.05, 165.63, 165.55 (4 × C=O benzoyl), 145.61, 145.58 (2 × C-2), 133.44-133.21 (CH benzoyl), 129.99-128.41 (CH benzoyl), 129.39-129.20 (C benzoyl) 123.43, 122.13 (2 × C-1), 85.55 (C-3'), 77.97 (2 × C-7), 75.75 (C-5), 75.48 (C-5), 74.84 (C-7'), 74.03, 73.77 (2 × C-3), 72.41 (C-5'), 70.85 (C-4'), 69.94, 69.59 (2 × C-4), 68.92, 68.53 (2 × C-6), 63.65, 63.13 (2 × C-8), 61.61 (C-8'), 59.84 (C-6'), 25.99 ($SiC(CH_3)_3$), 20.72, 19.98, 19.89 (3 × $CH_3$ acetyl), 18.44 ($SiC(CH_3)_3$),

-5.31 ($Si(CH_3)_2$). MS (ESI) m/z calcd for $C_{68}H_{84}N_6O_{21}Si_2$ $(M+H)^+$ 1378.59, found 1378.00; HRMS (Q-TOF) m/z calcd for $C_{68}H_{84}N_6O_{21}Si_2$ $(M+H)^+$ 1377.5228, found 1377.5326.

## Compound 5

Compound **4** (0.84 g, 0.61 mmol) in anhydrous $CH_2Cl_2$ (30 mL) containing pyridine (4.22 mL) was treated with triflic anhydride (12.2 mL of a 1 M solution in $CH_2Cl_2$, 12.2 mmol). The mixture was stirred at 0°C for 3 h. after which cold 1N $KHSO_4$ (20 mL) was added. The organic layer was washed twice with cold $H_2O$ (20 mL) and, once with cold brine (20 mL), dried over $Na_2SO_4$, filtered and concentrated. The residue was used for the next step without further purification. The residue was dissolved in an acetone/$H_2O$ mixture (15 mL, 4:1) and sodium azide (0.397 g, 6.1 mmol) was added. The mixture was stirred at r.t. overnight and diluted with a cold $H_2O$/ $CH_2Cl_2$ mixture (50 mL, 4:1). The water layer was separated and extracted with $CH_2Cl_2$ (10 mL). The combined organic layers were dried on $Na_2SO_4$, filtered and concentrated to afford compound **5** (0.86 g, 0.60 mmol, 98%) as a yellowish solid. $^1H$ NMR (400 MHz, $CDCl_3$): δ, ppm 8.00-7.93 (m, 4H, 4 × CH benzoyl), 7.85-7.72 (m, 5H, 4 × CH benzoyl, H-1), 7.65 (s, 1H, H-1), 7.58-7.26 (m, 12H, 12 × CH benzoyl), 5.90 (d, $J_{3',4'}$= 9 Hz, 1H, H-3'), 5.86-5.65 (m, 3H, H-4, 2 × H-5), 5.57 (m, 2H, H-4, H-5'), 5.42 (dd, $J_{4',3'}$= $J_{4'5'}$= 10 Hz, 1H, H-4'), 4.97-4.88 (m, 2H, 2 × H-3), 4.81 (dd, $J_{6',5'}$= $J_{6',7'}$= 10 Hz, 1H, H-6'), 4.61-4.51 (m, 1H, H-7'), 4.17-4.03 (m, 3H, H-8'a, 2 × H-6), 4.03-3.93 (m, 2H, 2 × H-8), 3.68-3.55 (m, 3H, H-8'b, 2 × H-7), 2.01, 1.69, 1.62 (s, 9H, 3 × $CH_3$, acetyl), 0.95 (s, 18H, 2 × $SiC(CH_3)_3$), 0.18-0.04 (m, 12H, 2 × $Si(CH_3)_2$). $^{13}C$ NMR (100 MHz, $CDCl_3$): δ, ppm 169.85, 169.01, 168.96 (3 × C=O acetyl), 165.85, 165.78, 165.56, 165.47 (4 × C=O benzoyl), 145.18 (2 × C-2), 133.57-133.39 (CH benzoyl), 129.96-128.44 (CH benzoyl), 129.39-129.20 (C benzoyl) 123.08, 121.61 (2 × C-1), 85.64 (C-3'), 79.88, 79.86 (2 × C-7), 75.01, 74.96 (2 × C-5), 74.69 (C-7'), 73.60, 73.35 (2 × C-3), 72.60 (C-4'), 72.17, 72.12 (2 × C-4), 70.73 (C-5'), 62.50, 62.40 (2 × C-8), 61.50 (C-8'), 60.43 (2 × C-6), 59.84 (C-6'), 26.04 ($SiC(CH_3)_3$), 20.67, 20.01, 19.94 (3 × CH3 acetyl), 18.57 ($SiC(CH_3)_3$), -4.97-(-5.26) ($Si(CH_3)_2$). MS (ESI) m/z calcd for $C_{68}H_{82}N_{12}O_{19}Si_2$ $(M+H)^+$ 1427.54, found 1427.75; HRMS (Q-TOF) m/z calcd for $C_{68}H_{82}N_{12}O_{19}Si_2$ $(M+H)^+$ 1427.5358, found 1427.5389.

## Compound 7

A 'click reaction' of a mixture of compound **5** (0.88 g, 0.62 mmol), compound **6a** (0.570 g, 1.48 mmol) with $CuSO_4 \cdot 5H_2O$ (28 mg, 0.11 mmol) and NaAsc (0.0443 g, 0.22 mmol) in DMF (15 mL) containing 10% $H_2O$ was was performed following the general procedure described above to afford compound **7** (0.87 g, 64%). $^1$H NMR (400MHz, $CDCl_3$): δ, ppm 7.88 (s, 1H, H-1), 7.82-7.68 (m, 9H, 8 × CH benzoyl, H-1), 7.68-7.59 (2 × s, 2H, 2 × H-1), 7.51-7.39 (m, 4H, 4 × CH benzoyl), 7.36-7.26 (m, 9H, 8 × CH benzoyl), 6.35-6.22 (m, 2H, 2 × H-5), 5.94 (d, $J_{3',4'}$= 9 Hz, 1H, H-3'), 5.88-5.73 (m, 2H, 2 × H-4), 5.67-5.52 (m, 2H, H-5', H-4'), 5.43-5.33 (m, 2H, 2 × H-13), 5.27-5.07 (m, 6H, 2 × H-3, 2 × H-11, 2 × H-6), 5.03-4.71 (m, 7H, 2 × H-12, 2 × H-9a, H-6', 2 × H-9b), 4.63-4.53 (m, 1H, H-7'), 4.48-4.37 (m, 4H, 2 × H-7, 2×H-10), 4.29-4.04 (m, 5H, 2 × H-15a, H-8'a, 2 × H-15b), 3.94-3.85 (m, 2H, 2 × H-14), 3.80 (d, $J_{8a,8b}$= 12 Hz, 2H, 2 × H-8a), 3.65 (dd, $J_{8'b,8'a}$= 12 Hz, $J_{8'b,7'}$= 3 Hz, 1H, H-8'b), 3.45-3.32 (m, 2H, 2 × H-8b), 2.20-1.88 (m, 21H, 7 × $CH_3$, acetyl), 1.77-1.63 (m, 12H, 4 × $CH_3$, acetyl), 0.94-0.84 (m, 18H, 2 × $SiC(CH_3)_3$), 0.05-(-0.07) (m, 12H, 2 × $Si(CH_3)_2$). $^{13}$C NMR (100 MHz, $CDCl_3$): δ, ppm 170.83-168.96 (C=O acetyl), 165.51 (C=O benzoyl), 165.39 (C=O benzoyl), 165.07 (C=O benzoyl), 144.90 (2 × C-2), 143.54 (2 × C-2), 133.64, 129.81-128.53 (CH benzoyl), 128.81 (C benzoyl) 123.66, 123.57, 123.14, 121.67 (4 × C-1), 99.29, 99.26 (2 × C-10), 85.70 (C-3'), 79.40, 79.33 (2 × C-7), 75.02 (C-7'), 74.20, 73.85 (2 × C-5), 73.67, 73.42 (2 × C-3), 72.70 (C-5'), 72.22, 72.15 (2 × C-4), 70.99 (2 × C-12), 70.78 (2 × C-14), 70.73 (C-4'), 68.74 (2 × C-11), 67.25 (2 × C-13), 61.84 (2 × C-9, 2 × C-8), 61.43 (2 × C-15, C-8'), 60.17, 60.12 (2 × C-6), 59.81 (C-6'), 25.96 ($SiC(CH_3)_3$), 20.97-19.92 ($CH_3$ acetyl), 18.52 ($SiC(CH_3)_3$), -5.11-(-5.34) ($Si(CH_3)_2$). MS (ESI) m/z calcd for $C_{102}H_{126}N_{12}O_{39}Si_2$ $(M+2H)^{2+}$ 1100.89, found 1100.60; HRMS (Q-TOF) m/z calcd for $C_{102}H_{126}N_{12}O_{39}Si_2$ $(M+2H)^{2+}$ 1100.3892, found 1100.8980.

## Compound 8

Compound **7** (0.87 g, 0.40 mmol) was dissolved in MeCN/$H_2O$ (7:1, 24 mL), and p-TsOH (0.114 g, 0.60 mmol) was added. The mixture was stirred at r.t. for 6 h then diluted with $CH_2Cl_2$ (50 mL) which was followed by the addition of 10% $NaHCO_3$ (20 mL). The organic layer was washed with brine (20 mL) and dried over $Na_2SO_4$, filtered and concentrated. The residue was purified by silica gel chromatography to afford

compound **8** (0.65 g, 0.33 mmol, 83.3%). $^{1}$H NMR (400MHz, $CDCl_3$): δ ppm 8.22 (s, 1H, H-1), 8.12 (s, 1H, H-1), 7.92 (s, 1H, H-1), 7.87 (s, 1H, H-1), 7.81 (d, J= 8 Hz, 4H, 4 × CH benzoyl), 7.76-7.69 (m, 4H, 4 × CH benzoyl), 7.47-7.35 (m, 4H, 4 × CH benzoyl), 7.31-7.16 (m, 8H, 8 × CH benzoyl), 6.45-6.33 (m, 2H, 2 × H-5), 6.30 (d, $J_{3',4'}$= 8 Hz, 1H, H-3'), 6.01-5.88 (m, 2H, 2 × H-4), 5.81 (dd, $J_{5',4'}$= $J_{5',6'}$= 10 Hz, 1H, H-5'), 5.68 (dd, $J_{4',3'}$= $J_{4',5'}$= 10 Hz 1H, H-4'), 5.53-5.37 (m, 4H, 2 × H-13, 2 × H-6), 5.32 (d, $J_{3,4}$= 12 Hz, 2H, 2 × H-3), 5.22-5.09 (m, 3H, 2 × H-11, H-6'), 5.04-4.80 (m, 6H, 2 × H-12, 2 × H-9a, 2 × H-9b), 4.73-4.65 (m, 1H, H-7'), 4.51-4.40 (m, 4H, 2 × H-7, 2 × H-10), 4.33-4.23 (m, 2H, 2 × H-15a), 4.17-4.07 (m, 2H, 2 × H-15b), 4.04-3.93 (m, 3H, 2 × H-14, H-8'a), 3.80 (d, $J_{8a,8b}$= 12 Hz, 2H, 2 × H-8a), 3.56 (dd, $J_{8'b,8'a}$= 12 Hz, $J_{8'b,7'}$= 4 Hz, 1H, H-8'b), 3.29 (d, $J_{8b,8a}$= 12 Hz, 2H, 2 × H-8b), 2.20-1.93 (m, 21H, 7 × $CH_3$, acetyl), 1.86-1.48 (m, 12H, 4 × $CH_3$, acetyl). $^{13}$C NMR (100 MHz, $CDCl_3$): δ, ppm 170.88-169.10 (C=O acetyl), 165.51 (C=O benzoyl), 165.26 (C=O benzoyl), 165.11 (C=O benzoyl), 144.59, 144.40, 143.75, 143.71 (4 × C-2), 133.60, 129.91-128.44 (CH benzoyl), 128.77, 128.37 (C benzoyl) 123.79 (2 × C-1), 123.63 (2 × C-1), 99.20, 99.10 (2 × C-10), 85.29 (C-3'), 79.35, 79.15 (2 × C-7), 74.85 (C-7'), 73.99 (2 × C-5), 73.68, 73.33 (2 × C-3), 72.72 (C-5'), 72.12 (2 × C-4), 70.94 (2 × C-12), 70.83 (2 × C-14), 70.81 (C-4'), 68.84 (2 × C-11), 67.27 (2 × C-13), 61.64 (2 × C-9), 61.45 (2 × C-15), 61.31 (C-8'), 60.64, 60.56 (2 × C-8), 59.92, 59.88 (2 × C-6), 59.84 (C-6'), 20.96-19.83 ($CH_3$ acetyl). HRMS (Q-TOF) m/z calcd for $C_{90}H_{98}N_{12}O_{39}$ $(M+2H)^{2+}$ 986.3027, found 986.8078.

## Compound 9

A solution of compound **8** (51.6 mg, 0.026 mmol), (2,2,6,6-Tetramethylpiperidin-1-yl)oxyl (0.164 mg, 1.05 µmol) in $CH_2Cl_2$ (1 mL ) was added to a solution of NaBr (1.1 mg, 10.5 µmol), and $Bu_4NBr$ (3.4 mg, 10.5 µmol) in satd. $NaHCO_3$ (1 mL, its pH was adjusted to 9.5 with satd. $Na_2CO_3$). At 0°C, NaOCl (6-14%, 105 µL) was added drop wise. The mixture was stirred vigorously at 0°C for 2 h and then quenched by the addition of $Na_2S_2O_3$ (sat., 0.5 mL), after which $H_2O$ (1.5 mL), $CH_2Cl_2$ (4 mL) were added, and the pH was adjusted to pH=3 with aqueous HCl (6N). The organic phase was washed with brine (1 mL), dried over $Na_2SO_4$, and concentrated *in vacuo*. The resulting oxidized compound was exposed to Zémplén conditions, followed by $H^+$-resin and was concentrated. The residue was subjected to preparative HPLC

purification, which gave compound **9** (11.1 mg, 0.010 mmol, 38%). $^{1}$H NMR (400 MHz, $D_2O$): δ, ppm 8.56 (s, 1H, H-1), 8.43 (s, 1H, H-1), 8.26 (s, 2H, 2 × H-1), 6.10 (d, $J_{3',4'}$= 12 Hz, 1H, H-3'), 5.06 (d, $J_{9a,9b}$= 12Hz, 2H, 2 × H-9a), 5.01-4.91 (m, 5H, 2 × H-9b, H-6', 2 × H-3), 4.89-4.72 (m, 4H, 2 × H-7, 2 × H-6), 4.58-4.41 (m, 6H, H-7', 2 × H-10, H-5', 2 × H-5), 4.30 (dd, $J_{4',3'}$= $J_{4',5'}$= 9.2 Hz, 1H, H-4'), 4.08 (dd, $J_{4,3}$= $J_{4,5}$= 9.2 Hz, 2H, 2 × H-4), 3.97-3.92 (m, 2H, 2 × H-13), 3.88-3.70 (m, 6H, 2 × H-15ab, 2 × H-14), 3.69-3.61 (m, 3H, 2 × H-12, H-8'a), 3.56 (dd, $J_{11,10}$= $J_{11,12}$= 8 Hz, 2H, 2 × H-11), 3.40 (dd, $J_{8'b,8'a}$= 12 Hz, $J_{8'b,7'}$= 4 Hz, 1H, H-8'b). $^{13}$C NMR (100 MHz, $D_2O$): δ, ppm 172.26 (2 × C-8), 145.12 (C-2), 145.01 (C-2), 144.38 (2 × C-2), 126.65 (2 × C-1), 126.51 (C-1), 125.47 (C-1), 102.56 (2 × C-10), 88.28 (C-3'), 78.82 (2 × C-7), 77.71 (C-7'), 75.95 (2 × C-14), 75.04 (C-5), 75.00 (C-5), 74.40 (C-3), 74.38 (C-3), 74.28 (C-5'), 73.64 (C-4), 73.59 (C-4), 73.42 (2 × C-12), 73.30 (C-4'), 71.39 (2 × C-11), 69.33 (2 × C-13), 64.60 (2 × C-6), 62.41 (2 × C-9), 62.33 (C-6'), 61.68 (2 × C-15), 60.32 (C-8'). MS (ESI) m/z calcd for $C_{40}H_{56}N_{12}O_{26}$ $(M+H)^+$ 1121.34, found 1121.05; HRMS (Q-TOF) m/z calcd for $C_{40}H_{56}N_{12}O_{26}$ $(M+H)^+$ 1121.3429, found 1121.3465, $(M+Na)^+$ 1143.3282.

## Compound 10

Compound **8** (49.2 mg, 0.025 mmol) was dissolved in dry $CH_2Cl_2$ (2 mL) with pyridine (200 µL). The mixture was cooled down to 0°C, after which triflic anhydride (203 µL, 0.203 mmol) was added drop wise to the above solution. The reaction was allowed to warm up to r.t. and stirred for 3 h after which 1N $KHSO_4$ (2 mL) and $CH_2Cl_2$ (15 mL) were added. The organic layer was washed with $H_2O$ (5 mL), with brine (5 mL), dried on $Na_2SO_4$, filtered and concentrated. The residue was purified by preparative-HPLC to afford the corresponding pyridinium compound. The resulting material was then treated with 0.5 M NaOMe in MeOH (5 mL), stirred at r.t. overnight after which 1N HCl was added to adjust pH≈6 and the solvents were evaporated *in vacuo*. The residue was purified by preparative HPLC which gave compound **10** (13.3 mg, 0.011 mmol, 44%). $^{1}$H NMR (400 MHz, $D_2O$): δ, ppm 8.69 (d, $J_{16,17}$= 8 Hz, 4H, 4 × H-16), 8.56 (dd, $J_{18,17}$= $J_{18,17}$= 8 Hz, 2H, 2 × H-18), 8.42 (s, 1H, H-1), 8.38 (s, 2H, 2 × H-1), 8.28 (s, 2H, H-1), 8.05-7.98 (m, 4H, 4 × H-17), 6.06 (d, $J_{3',4'}$= 8 Hz, 1H, H-3'), 5.12 (d, $J_{9a,9b}$=12 Hz, 2H, 2 × H-9a), 5.00-4.92 (m, 3H, 2 × H-9b, H-6'), 4.92-4.73 (m, 8H, 2 × H-7, 2 × H-6, 2 × H-3, 2 × H-8a), 4.62-4.42 (m, 6H, 2 × H-10, 2 × H-8b, H-7', H-5'), 4.34-4.23 (m, 3H,

2 × H-5, H-4'), 4.07-3.98 (m, 2H, 2 × H-4), 3.98-3.93 (m, 2H, 2 × H-13), 3.88-3.73 (m, 6H, 2 × H-15ab, 2×H-14), 3.71-3.63 (m, 3H, 2 × H-12, H-8'a), 3.63-3.54 (m, 2H, 2 × H-11), 3.35 (dd, $J_{8'b,8'a}$= 12 Hz, $J_{8'b,7'}$= 4 Hz, 1H, H-8'b). $^{13}$C NMR (100 MHz, $D_2O$): δ, ppm 146.77 (2 × C-18), 145.32 (4 × C-16) 144.63 (2 × C-2), 144.37 (C-2), 144.23 (C-2), 128.21 (4×C-17), 125.96 (2 × C-1), 125.56 (C-1), 124.75 (C-1), 102.42 (2 × C-10), 87.61 (C-3'), 77.08 (C-7'), 75.87 (2 × C-7), 75.49 (2 × C-14), 74.68(2 × C-5), 73.84 (2 × C-3), 73.73 (C-5'), 72.96(2 × C-4), 72.89 (2 × C-12), 72.72 (C-4'), 70.84 (2 × C-11), 68.76 (2 × C-13), 63.61 (2 × C-6), 62.07 (2 × C-9), 61.71 (C-6'), 61.43 (2 × C-8), 61.20 (2 × C-15), 59.76 (C-8'). MS (ESI) m/z calcd for $C_{50}H_{68}N_{14}O_{22}^{2+}$ $M^{2+}$ 608.23, found 608.30; HRMS (Q-TOF) m/z calcd for $C_{50}H_{68}N_{14}O_{22}^{2+}$ $M^{2+}$ 608.2311, found 608.2356.

## Compound 11

Compound **8** (246 mg, 0.125 mmol) was dissolved in dry $CH_2Cl_2$ (5 mL) and tosyl chloride (238.4 mg, 1.25 mmol) and 1,4-diazabicyclo[2.2.2]octane (38 mg, 0.34 mmol) were added. The mixture was stirred at r.t. overnight after which the solvent was removed. The residue was redissolved in $CH_2Cl_2$ (15 mL) and washed with $H_2O$ (5 mL), brine (5 mL), dried on $Na_2SO_4$, filtered and concentrated. The resulting material was purified by silica gel chromatography to give the tosylated compound (220 mg, 77.2%). The residue was then dissolved in dry DMF (8 mL) to which sodium azide (33.6 mg, 0.52 mmol) was added. The mixture was stirred at 95°C overnight after which the solvent was removed. The residue was dissolved in $CH_2Cl_2$ (15 mL), washed with 1N $KHSO_4$ (5 mL), $H_2O$ (5 mL), dried on $Na_2SO_4$, filtered and concentrated. The residue was purified by silica gel chromatography to afford compound **11** (86.8 mg, 0.043 mmol, 34.4%). $^1$H NMR (400MHz, $CDCl_3$): δ 7.95 (s, 1H, H-1), 7.82-7.765 (m, 11H, 4 × CH benzoyl, 3 × H-1), 7.51-7.39 (m, 4H, 4 × CH benzoyl), 7.36-7.22 (m, 8H, 8 × CH benzoyl), 6.36-6.23 (m, 2H, 2 × H-5), 5.98 (d, $J_{3',4'}$= 8 Hz, 1H, H-3'), 5.92-5.76 (m, 2H, 2 × H-4), 5.65 (t, $J_{5',4'}$= $J_{5',6'}$= 8 Hz, 1H, H-5'), 5.57 (t, $J_{4',3'}$= $J_{4',5'}$= 8 Hz, 1H, H-4'), 5.43-5.35 (m, 2H, 2 × H-13), 5.29 (d, $J_{3,4}$= 8 Hz, 2H, 2 × H-3), 5.22-5.09 (m, 4H, 2 × H-11, 2 × H-6), 5.02-4.83 (m, 5H, 2 × H-12, 2 × H-9a, H-6'), 4.77 (d, $J_{9b,9a}$= 12 Hz, 2H, 2 × H-9b), 4.71-4.56 (m, 3H, 2 × H-7, H-7'), 4.44 (d, $J_{10,11}$= 8 Hz, 2H, 2 × H-10), 4.28-4.18 (m, 2H, 2 × H-15a), 4.18-4.05 (m, 3H, 2 × H-15b, H-8'a), 3.94-3.86 (m, 2H, 2 × H-14), 3.72-3.62 (m, 1H, H-8'b), 3.56 (t, $J_{8a,8b}$= 12 Hz, 2H, 2 × H-8a), 2.95, 2.88 (2×dd, $J_{8b,8a}$=

12 Hz, $J_{8b,7}$= 4 Hz, 2H, 2 × H-8b), 2.21-1.92 (21H, 7 × $CH_3$, acetyl), 1.83-1.59 (12H, 4 × $CH_3$, acetyl). $^{13}$C NMR (100 MHz, $CDCl_3$): δ, ppm 170.88-169.06 (C=O acetyl), 165.36 (C=O benzoyl), 165.28 (C=O benzoyl), 164.99 (C=O benzoyl), 144.72, 144.65, 144.25, 144.24 (4 × C-2), 133.75-133.54, 129.85-128.49 (CH benzoyl), 128.72, 128.29, 128.26 (C benzoyl) 123.49, 123.40, 123.30 (3 × C-1), 121.72 (2 × C-1), 99.66, 99.60 (2 × C-10), 85.68 (C-3'), 77.93, 77.72 (2 × C-7), 75.00 (C-7'), 73.76, 73.71 (2 × C-5), 73.50, 73.28 (2 × C-3), 72.66 (C-5'), 72.07, 72.03 (2 × C-4), 70.91 (2 × C-12, C-4'), 70.85 (2 × C-14), 70.81 (C-4'), 68.78 (2 × C-11), 67.23 (2 × C-13), 62.17, 62.13 (2 × C-9), 61.43 (2 × C-15, C-8'), 60.81 (2 × C-6), 59.72 (C-6'), 50.58, 50.37 (2 × C-8), 20.94-19.92 ($CH_3$ acetyl). MS (ESI) m/z calcd for $C_{90}H_{96}N_{18}O_{37}$ $(M+2H)^{2+}$ 1011.91, found 1012.35; HRMS (Q-TOF) m/z calcd for $C_{90}H_{96}N_{18}O_{37}$ $(M+2H)^{2+}$ 1011.3092, found 1011.3091, $(M+H+Na)^{2+}$ 1022.8011, $(M+2Na)^{2+}$ 1033.2923.

## Compound 12

Compound **11** (22.6 mg, 0.020 mmol) was treated with 0.5 M NaOMe in MeOH (2.5 mL) at r.t. overnight and briefly with $H^+$-resin. The residue was concentrated and dissolved in $H_2O$ (2 mL) and Pd/C (15 mg, 10% Pd basis) was added. The pH was adjusted to pH=1 by 6N HCl and the mixture was stirred at r.t. under an $H_2$ atmosphere until the hydrogenation was complete. After that, the reaction mixture was filtered through celite. The filtrate was concentrated under reduced pressure and subjected to preparative HPLC purification, which gave compound **12** (12.9 mg, 0.012 mmol, 60%). $^1$H NMR (400MHz, $D_2O$): δ ppm 8.56 (s, 1H, H-1), 8.43 (s, 1H, H-1), 8.33 (s, 2H, 2 × H-1), 6.13 (d, $J_{3',4'}$= 8 Hz, 1H, H-3'), 5.09 (d, $J_{9a,9b}$= 10 Hz, 2H, 2 × H-9a), 5.03-4.89 (m, 5H, 2 × H-3, H-6', 2 × H-9b), 4.89-4.70 (m, 2H, 2 × H-6), 4.66-4.47 (m, 6H, 2 × H-7, 2 × H-10, H-7', H-5'), 4.41-4.28 (m, 3H, 2 × H-5, H-4'), 4.08-3.99 (m, 2H, 2 × H-4), 3.99-3.93 (m, 2H, 2 × H-13), 3.89-3.73 (m, 6H, 2 × H-15ab, 2 × H-14), 3.71-3.64 (m, 3H, 2 × H-12, H-8'a), 3.57 (dd, $J_{11,10}$= $J_{11,12}$= 8 Hz, 2H, 2 × H-11), 3.40 (dd, $J_{8'b,8'a}$= 12 Hz, $J_{8'b,7'}$= 4 Hz, 1H, H-8'b), 3.24-3.11 (m, 2H, 2 × H-8a), 2.86 (d, $J_{8b,8a}$= 12 Hz, 2H, 2 × H-8b). $^{13}$C NMR (100 MHz, $D_2O$): δ, ppm 144.94, 144.82 (2 × C-2), 144.69 (2 × C-2), 126.18 (2 × C-1), 126.13 (C-1), 125.07 (C-1), 102.63 (2 × C-10), 87.97 (C-3'), 77.46 (C-7'), 75.76 (2 × C-14), 74.90 (2 × C-5), 74.51 (2 × C-7), 74.33 (2 × C-3), 74.07 (C-5'), 73.45 (2 × C-4), 73.18 (2 × C-12), 73.13 (C-4'), 71.14 (2 × C-11), 69.09 (2 × C-13), 64.27

(2 × C-6), 62.39 (2 × C-9), 62.08 (C-6'), 61.48 (2 × C-15), 60.12 (C-8'), 40.59 (2 × C-8). MS (ESI) m/z calcd for $C_{40}H_{62}N_{14}O_{22}$ $(M+2H)^{2+}$ 546.21, found 546.90; HRMS (Q-TOF) m/z calcd for $C_{40}H_{62}N_{14}O_{22}$ $(M+H)^{+}$ 1091.4163, found 1091.4282, $(M+Na)^{+}$ 1113.4132.

## Compound 13

A 'click reaction' of a mixture of compound **11** (61.8 mg, 0.031 mmol), phenylacetylene (11.1 mg, 0.109 mmol), NaAsc (6.5 mg, 0.033 mmol) and $CuSO_4 \cdot 5H_2O$ (4.1 mg, 0.016 mmol) was performed following the general procedure described above to afford the coupling product compound (54 mg, 80%), which was then treated with NaOMe in MeOH (0.5 M, 5 mL). The mixture was stirred at r.t. overnight, treated with $H^{+}$-resin, concentrated and subjected to preparative HPLC purification, which gave compound **13** (9.8 mg, 7.3 µmol, 24%). $^{1}$H NMR (400MHz, $D_2O$ with 30% $CD_3CN$): δ ppm 8.62 (s, 1H, H-1), 8.50, 8.49 (2 × s, 2H, 2 × H-1), 8.46 (s, 1H, H-1), 8.42, 8.41 (2 × s, 2H, 2 × H-1), 8.05-7.98 (m, 4H, 4 × C-17), 7.75 (t, $J_{18,17}$= $J_{18,19}$= 8 Hz, 4H, 4 × C-18), 7.71-7.62 (m, 2H, 2 × C-19), 6.21 (d, $J_{3'4'}$= 9 Hz, 1H, H-3'), 5.27 (d, $J_{9a,9b}$= 12 Hz, 2H, 2 × H-9a), 5.16-4.99 (m, 7H, 2 × H-9b, 2 × H-3, H-6', 2 × H-7), 4.92-4.83 (m, 2H, 2 × H-6), 4.83-4.70 (m, 4H, 2 × H-8a, 2 × H-10), 4.70-4.58 (m, 4H, 2 × H-8b, H-5', H-7'), 4.57-4.48 (m, 2H, 2 × H-5), 4.43 (t, $J_{4',3'}$= $J_{4',5'}$= 8 Hz, 1H, H-4'), 4.21-4.11 (m, 4H, 2 × H-4, 2 × H-13), 4.08-3.90 (m, 6H, 2 × H-15ab, 2 × H-14), 3.89-3.82 (m, 2H, 2 × H-12), 3.82-3.73 (m, 3H, 2 × H-11, H-8'a), 3.46 (dd, $J_{8'b,8'a}$= 13 Hz, $J_{8'b,7'}$= 4 Hz, 1H, H-8'b). $^{13}$C NMR (100 MHz, $D_2O$ with 30% $CD_3CN$): δ, ppm 147.90 (4 × C-2), 145.44 (2 × C-2), 130.82-129.48 (4 × C-18, 2 × C-19), 130.00 (2 × C-16), 126.40 (2 × C-1), 126.21 (4 × C-17), 125.78 (C-1), 124.76 (C-1), 123.66 (2 × C-1), 102.73 (2 × C-10), 88.11 (C-3'), 77.554 (C-7'), 76.27 (2 × C-7), 75.81 (2 × C-14), 75.31 (2 × C-5), 74.50 (2 × C-3), 74.20 (C-5'), 73.62 (2 × C-4), 73.38 (2 × C-12), 73.25 (C-4'), 71.29 (2 × C-11), 69.18 (2 × C-13), 64.29 (2 × C-6), 62.37 (2 × C-9), 62.07 (C-6'), 61.55 (2 × C-15), 60.22 (C-8'), 51.44 (2 × C-8). MS (ESI) m/z calcd for $C_{56}H_{70}N_{18}O_{22}$ $(M+2H)^{2+}$ 674.63, found 674.95; HRMS (Q-TOF) m/z calcd for $C_{56}H_{70}N_{18}O_{22}$ $(M+H)^{+}$ 1348.2624, found 1348.4964.

## Reference

[1] a) J. E. Hudak, C. R. Bertozzi, *Chem. Biol*. **2014**, 21, 16-37; b) B. Ernst, J. L. Magnani, *Nature Rev. Drug. Discov*. **2009**, 8, 661-677.
[2] a) T. K. Dam, T. A. Gerken, C. F. Brewer, *Biochemistry*, **2009**, 48, 3822-3827; b) S. G. Gouin *Chem. Eur. J*. **2014**, 20, 11616-11628; c) N. Jayaraman *Chem. Soc. Rev.* **2009**, 38, 3463-3483; d) R. J. Pieters, *Org. Biomol. Chem*. **2009**, 7, 2013-2025; e) M. Gingras, Y. M. Chabre, M. Roy, R. Roy, *Chem. Soc. Rev*. **2013**, 42, 4823-4841; f) P. M. Levine, T. P. Carberry, J. M. Holub, K. Kirshenbaum, *MedChemComm* **2013**, 4, 493-509; f) L. L. Kiessling, J. C. Grim, *Chem. Soc. Rev*. **2013**, 42, 4476-4491; g) M. C. Galan, P. Dumy, O. Renaudet, *Chem. Soc. Rev*. **2013**, 42, 4599-4612.
[3] a) A. Imberty, M. Wimmerová, E. P. Mitchell, N. Gilboa-Garber, *Microbes Infect.*, **2004**, 6, 221-228; b) G. Cioci, E. P. Mitchell, C. Gautier, M. Wimmerová, D. Sudakevitz, S. Pérez, N. Gilboa-Garber, A. Imberty, *FEBS Letters*, **2003**, 555, 297-301.
[4] a) M. L. Gening, D. V. Titov, S. Cecioni, A. Audfray, A. G. Gerbst, Y. E. Tsvetkov, V. B. Krylov, A. Imberty, N. E. Nifantiev, S. Vidal, *Chem. Eur. J.* **2013**, 19, 9272-9285; b) S. Cecioni, J. Praly, S. E. Matthews, M. Wimmerova, A. Imberty, S. Vidal, *Chem. Eur. J*. **2012**, 18, 6250-6263; c) B. Gerland, A. Goudot, G. Pourceau, A. Meyer, S. Vidal, E. Souteyrand, J. Vasseur, Y. Chevolot, F. Morvan, *J. Org. Chem*. **2012**, 77, 7620-7626; d) S. Cecioni, S. Faure, U. Darbost, I. Bonnamour, H. Parrot-Lopez, O. Roy, C. Taillefumier, M. Wimmerova, J. Praly, A. Imberty, S. Vidal, *Chem. Eur. J*. **2011**, 17, 2146-2159; e) S. Cecioni, V. Oerthel, J. Iehl, M. Holler, D. Goyard, J. Praly, A. Imberty, J. Nierengarten, S. Vidal, *Chem. Eur. J*. **2011**, 17, 3252-3261; f) B. Gerland, A. Goudot, C. Ligeour, G. Pourceau, A. Meyer, S. Vidal, T. Gehin, O. Vidal, E. Souteyrand, J. Vasseur, Y. Chevolot, F. Morvan, *Bioconjug. Chem*. **2014**, 25, 379-392; g) A. Novoa, T. Eierhoff, J. Topin, A. Varrot, S. Barluenga, A. Imberty, W. Römer, N. Winssinger, *Angew. Chem. Int. Ed.* **2014**, 53, 8885-8889; h) R. U. Kadam, M. Bergmann, D. Garg, G. Gabrieli, A. Stocker, T. Darbre, J.-L. Reymond, *Chem. Eur. J*. **2013**, 19, 17054-17063; i) R. U. Kadam, D. Garg, J. Schwartz, R. Visini, M. Sattler, A. Stocker, T. Darbre, J. L. Reymond, *ACS Chem. Biol.* **2013**, 8, 1925-1930.
[5] a) R. J. Pieters, *Med. Res. Rev*. **2007**, 27, 796-816; b) A. Bernardi, J. Jiménez-Barbero, A. Casnati, C. De Castro, T. Darbre, F. Fieschi, J. Finne, H. Funken, K. Jaeger, M. Lahmann, T. K. Lindhorst, M. Marradi, P. Messner, A. Molinaro, P. V. Murphy, C. Nativi, S. Oscarson, S. Penadés, F. Peri, R. J. Pieters, O. Renaudet, J. Reymond, B. Richichi, J. Rojo, F. Sansone, C. Schäffer, W. B. Turnbull, T. Velasco-Torrijos, S. Vidal, S. Vincent, T. Wennekes, H. Zuilhof, A. Imberty, *Chem. Soc. Rev*. **2013**, 42, 4709-4727; c) R. J. Pieters, *Adv. Exp. Med. Biol*. **2011**, 715, 227-240; d) N. P. Pera, R. J. Pieters, *MedChemComm* **2014**, 5, 1027-1035.
[6] Measured between the two anomeric oxygens of the bound galactose units of the X-ray structure with pdb code: 1OKO, see: A. Imberty, M. Wimmerová, E. P. Mitchell, N. Gilboa-Garber, *Microbes Infect.*, **2004**, 6, 221-228.
[7] V. Wittmann, R. J. Pieters, *Chem. Soc. Rev*. **2013**, 42, 4492-4503.
[8] a) F. Pertici, N. J. de Mol, J. Kemmink, R. J. Pieters, *Chem. Eur. J.* **2013**, 19, 16923-16927; b) F. Pertici, R. J. Pieters, *Chem. Commun*. **2012**, 48, 4008-4010.
[9] J. W. Goding, *J. Immunol. Methods*, **1976**, 13, 215-226.

# Chapter 3

## Tetra- vs Pentavalent Inhibitors of Cholera Toxin

Part of this chapter has been accepted for publication:

O. Fu, A. V. Pukin, H. C. Quarles van Ufford, T. R. Branson, D. M. E. Thies-Weesie, W. B. Turnbull, G. M. Visser, R. J. Pieters, *Chemistry Open*, **2015**, in press (DOI: 10.1002/open.201500006)

## Introduction

The cholera disease is a major source of suffering around the world. This diarrhoeal disease is caused by the *Vibrio cholerae* bacteria, but it is the cholera toxin (CT) it produces that is the actual pathogenic species. The toxin attaches itself to the intestinal cell wall where it is subsequently internalized and the A subunit of this $AB_5$ toxin[1] subsequently initiates the disease by raising the cellular cAMP concentration followed by fluid efflux into the intestines.[2] The initial attachment of the toxin to the intestinal cell surface is caused by the 5 B-subunits ($CTB_5$) that surround the A-subunit. While a single B binding site already binds with nanomolar affinities to a GM1-oligosaccharide (GM1os), simultaneous binding of more than one B-subunits of the toxin can greatly enhance its affinity. Blocking the initial attachment of the toxin to the cell surface has the potential to block the disease. Considering the fact that the toxin itself takes advantage of multivalency[3][4] in its binding to the cell surface, it was clear that, in order to interfere effectively, a multivalent ligand system would have to be designed. Several multivalent systems have been designed based on dendrimers,[5][6] polymers[7][8], peptides[9] and also pentavalent scaffolds[10][11][12] and evaluated and have clearly shown the promise of the multivalency approach.[13][14][15][16] In one such approach, we attached the GM1os to dendritic scaffolds of varying valencies. Especially, the effective were the tetra- and octavalent systems, which were able to inhibit $CTB_5$ at subnanomolar concentrations and with orders of magnitude potency enhancements over the corresponding monovalent ligand.[17] Subsequent studies with the close relative of the cholera toxin, the heat labile enterotoxin of *E. coli* (LT) showed that the multivalent ligands, when mixed with the toxin, would lead to aggregates involving many toxin molecules.[18] This was shown by analytical ultracentrifuge experiments as well as by atomic force microscopy. The observed aggregation was attributed to the mismatch in valency between the multivalent ligand (four or eight) and the multisubunit toxin (five). In fact, it was considered a possibility that the enormous potency enhancements observed in the inhibition assay with the cholera toxin could be due to the mismatch and the subsequent aggregation that the multivalent ligands initiated. On the other hand, there were reports in the literature, which described

pentavalent CT or LT ligands that were shown to be potent toxin inhibitors and clearly formed a 1:1 complex with the toxin, as judged by dynamic light scattering (DLS) experiments.[10] Based on the 1:1 design, several pentavalent CT inhibitors were reported, and it was suggested that this design was beneficial to the inhibition,[11][12] also including a modified version of the cholera toxin that can no longer bind GM1 and was outfitted with 5 GM1os ligands.[19] However, so far no experiments were undertaken that compared a matching pentavalent inhibitor with inhibitors of nonmatching valencies based on closely related scaffolds. Therefore, it remains very much unclear which of the two approaches - 1:1 design or mismatch-aggregation - is the best. We here address this question and report on the synthesis and evaluation in the same assay of tetra- and pentavalent GM1-based ligand systems for CT inhibition (**Fig.1**).

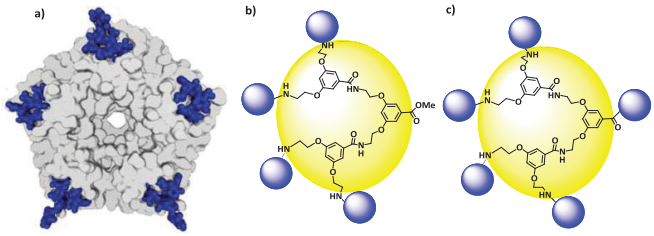
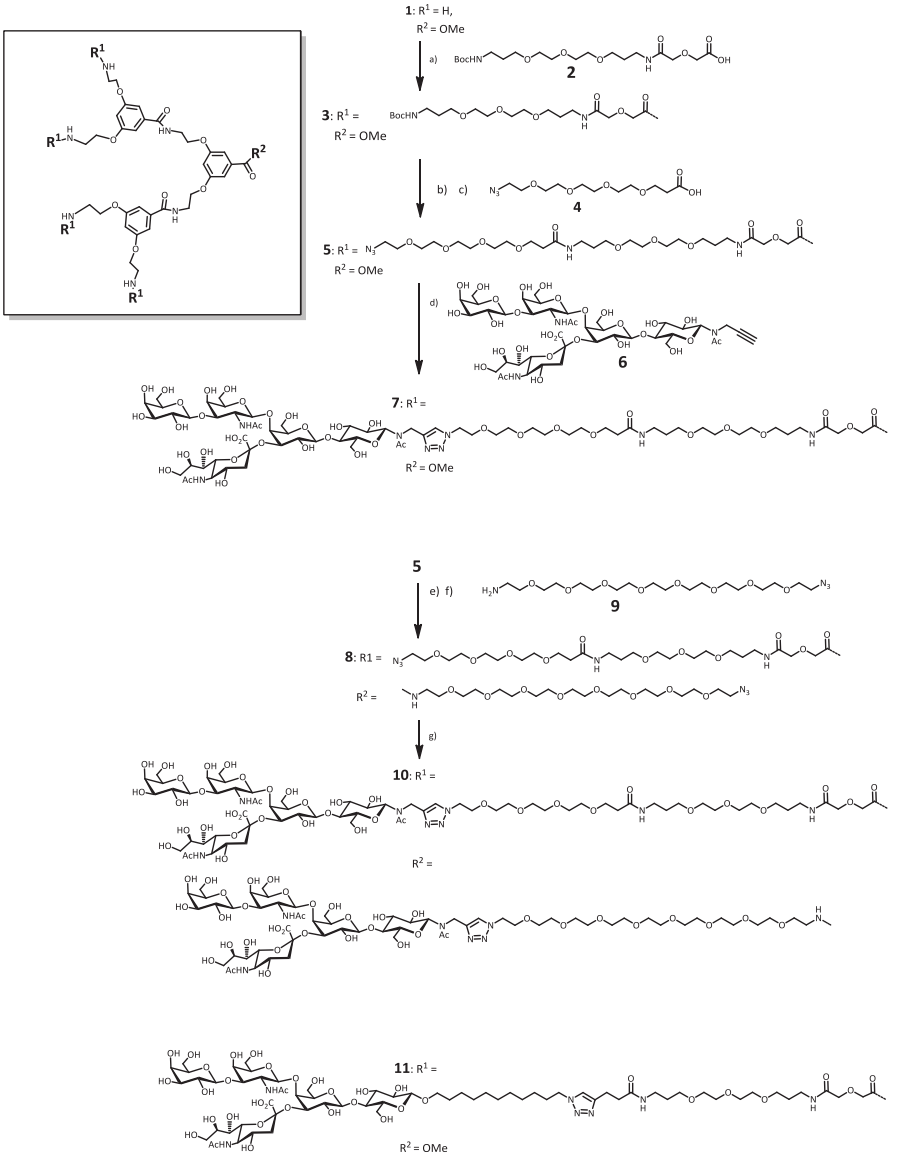



**Figure 1**. a) X-ray structure of the cholera toxin B-subunit bound to GM1os (PDB ID: 3CHB); b) & c) General architecture of the tetravalent (b) and pentavalent (c) ligands described here.

## Results and discussion

### Synthesis

As shown in **Scheme 1**, the synthesis started with the preparation of the scaffold for the tetravalent inhibitor **5**, which was subsequently used for the preparation of the scaffold for the pentavalent inhibitor **8**. The overall design of the tetravalent inhibitor was kept close to the previous version (**11**)[17] although there were differences in the spacer arms due to the availability and use of a different GM1os building block, i.e. **6** in this case. The length of the spacer arm was almost the same as before, with the present one just measuring two atoms longer. Furthermore, the



**Scheme 1.** Reagents and Conditions; a) **2**, BOP, DIPEA, DMF, r.t., o/n, 40%; b) TFA, $CH_2Cl_2$, r.t., 3 h; c) **4**, HATU, DIPEA, DMF, r.t., o/n, 60%; d) $CuSO_4 \cdot 5H_2O$, Na ascorbate, DMF/$H_2O$ 1:1, microwave, 80 °C, 20 min, 30%; e) Tesser's base (30: 9: 1, dioxane/MeOH/4 N NaOH), r.t. until hydrolyzed; f) **9**, BOP, DIPEA, DMF, r.t., o/n, 51%; g) **6**, $CuSO_4 \cdot 5H_2O$, Na ascorbate, DMF/$H_2O$ 1:1, microwave, 80 °C, 20 min, 41%.

previous partly hydrophobic partly hydrophilic spacer arm was now replaced by one consisting almost entirely of hydrophilic ethylene glycol units. The synthesis started

with the elongation of the four arms of **1** as previously described.[5] The spacer **2**[20] was coupled to the dendritic scaffold **1** by the action of BOP and DIPEA, which resulted in **3** in 50% yield. After that, TFA was used to remove the Boc protecting group from the amino groups of **3**, and a coupling reaction between **3** and **4**[21] using HATU and DIPEA afforded the tetrameric full length scaffold **5** in 60% yield over two steps. Microwave-assisted CuAAC was subsequently used to conjugate the GM1os derivative **6** to the scaffold **5**, which efficiently yielded the tetravalent GM1 derivative **7**. The latter was purified by preparative HPLC.

The tetravalent scaffold **5** formed the starting point for the synthesis of the pentavalent version. To this end, the methyl ester of **5** was saponified quantitatively by Tesser's base.[22] The resulting carboxylic acid was coupled to the commercially available spacer **9** using BOP and DIPEA, which successfully gave **8** as the pentavalent scaffold in 51% yield over two steps. Subsequently, a microwave-assisted CuAAC conjugation reaction was employed on **8** and **6** leading to the formation of the pentavalent GM1 derivative **10**, which was purified by preparative HPLC.

## Inhibition

The compounds were evaluated as $CTB_5$ inhibitors in an ELISA type assay as described before.[17] A 96-well ELISA plate was coated by the natural bovine brain GM1 ganglioside. The remaining binding sites were blocked with BSA. Horseradish peroxidase (HRP)-conjugated $CTB_5$ was incubated with varying concentrations of the tested inhibitors for 2 hours at room temperature. After that, the remaining activity of $CTB_5$ was measured upon addition of the solutions to the wells and incubation for 30 min at room temperature to allow for binding of the remaining toxin. After incubation and washing, the amount of bound toxin was quantified by using a chromogenic substrate for HRP. The previously reported[17] tetravalent GM1 compound **11** was used here as a reference in inhibitory potency evaluation. In the present assay it showed an $IC_{50}$ of 190 pM, a value close to the previously reported one (230 pM). (**Table 1**) The new tetravalent GM1 **7** exhibited a very similar inhibitory potency with an $IC_{50}$ of 160 pM. This result shows that a slightly different spacer length and considerably different spacer polarity did not lead to significantly different inhibitory properties. The pentavalent GM1 derivative **10** exhibited an $IC_{50}$

of 260 pM, which is in the same range as the values found for both of the tetravalent ligands. This indicates that in our assay, the potency of the ligand of matching valency, does not essentially differ from the potencies of its non-matching analogues.

**Table 1.** Potency of the inhibitors.[a]

| Compound[a] | Valency | $IC_{50}$ [nM] |
|---|---|---|
| **11** | 4 | 0.19 (±0.02) |
| **7** | 4 | 0.16 (±0.04) |
| **10** | 5 | 0.26 (±0.02) |

[a] Determined in an ELISA with $CTB_5$-HRP (0.43 nM) and wells coated with GM1.

## Sedimentation velocity analytical ultracentrifugation (SV-AUC)

In order to learn whether the pentavalent geometry of **10** leads to a different, possibly less aggregative, binding mode, SV-AUC[23][24] experiments were undertaken. First a sample with just $CTB_5$ was measured. It contained a single species with a sedimentation coefficient of 4.4 S corresponding to a mass of 58 kDa for the protein pentamer. Sisu *et al*.[18] had previously used SV-AUC to test the tetrameric GM1os dendrimer **11** with $LTB_5$ and it was found to strongly aggregate the protein while no discrete oligomers were observed. In the present experiments tetravalent inhibitor **7**, which is structurally similar to **11**, was added to $CTB_5$ at a pentamer concentration of 50 µM. With the addition of 0.2-1.0 equivalents, a dramatic reduction in the overall signal was observed, as had previously been shown for **11** and $LTB_5$, indicating rapid sedimentation of aggregating large particles. However, with inhibitor **7**, the emergence of a peak at 7.2±0.2 S was seen with a predicted mass of approximately 110 kDa, which corresponds to a dimer of CTB pentamers. With increasing amounts of inhibitor up to 10 equivalents, the amount of the dimer species increased and the emergence of some stable CTB pentamer was also observed. Excess and unbound inhibitor was observed as a peak at 0.9±0.1 S corresponding with a mass of 8 kDa.

Pentavalent inhibitor **10** matched the number of ligand groups to the number of binding sites of $CTB_5$ and so it was expected that this inhibitor should form stable 1:1

complexes. However, the AUC results were very similar to those observed for tetravalent ligand **7**. A reduction in signal indicated large scale aggregation and some dimerisation of CTB pentamers was observed. Again, at higher equivalents of inhibitor, some CTB pentamer was seen but with no significant difference to the tetravalent inhibitor.

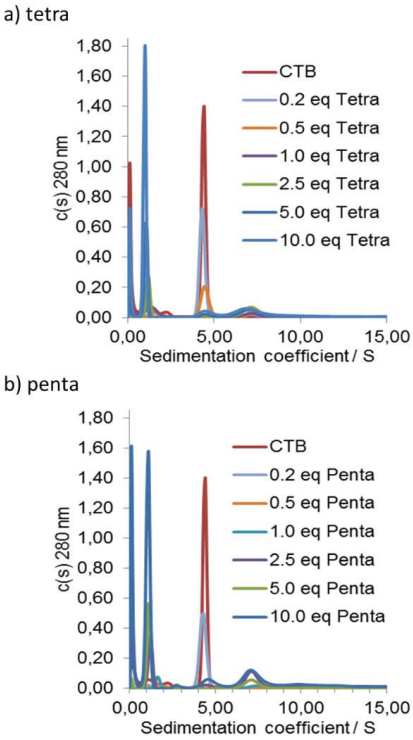



**Figure 2.** Sedimentation velocity analytical ultracentrifugation profiles of tetravalent **7** (a) and pentavalent **10** (b), recorded with increasing amounts of multivalent ligand.

## Conclusion

For the first time a penta and tetravalent cholera toxin inhibitor based on the same scaffold were compared. Cleary, the pentavalent geometry of **10** did not yield major benefits over the tetravalent **7**, in fact it was a little worse. However, it was still a strong inhibitor so major steric clashes did not occur in the binding of **10** to the toxin. Nevertheless one can argue that the potency per arm is significantly reduced by a factor of ca. 2. Both **7** and **10** behaved very similar in the Sedimentation velocity analytical ultracentrifugation (SV-AUC). As noted before for **11**, aggregation occurs in

the binding resulting in higher order structures, while only minor amounts of bound pentamer (or its dimer) could be detected. The arms of the systems described here are designed in agreement with the concept that their 'effective length'[10] should match the distance they should cover. The lengths of their extended conformations are therefore far longer. While the fifth arm is slightly shorter than the other four, it should be kept in mind that it is easily capable of bridging the fifth site and also that is attached to a different site of the scaffold.

It is of interest to compare our results to related pentavalent systems. Even though inhibition assay results cannot be directly compared, it is a fact that the same assay are used in these studies. One pentavalent GM1 system based on a corranulene scaffold, exhibited an $IC_{50}$ of 5 nM, presumably not living up to its full potential due to self-association of the scaffold.[12] A related calix[5]arene based system showed higher potency with an $IC_{50}$ of 450 pM. Neither of these systems showed a compelling argument in favor of a pentavalent system, consistent with our results. $CTB_5$, whose binding sites were disabled, was recently used as a scaffold for the display of 5 GM1 units. This construct showed an $IC_{50}$ of 104 pM and was shown to form a 1:1 complex with the toxin. While impressive, it is not very different from the 160 pM observed here and our previously reported octavalent GM1 structured showed an $IC_{50}$ of 50 pM.[17]

It seems that a pre-organized system can indeed bind in a 1:1 fashion with $CTB_5$,[19] however systems which can adopt more geometries, such as those described here can be equally potent, and this may possibly be due to their ability to form higher order structures or simply due to statistical options for binding. Bundle and Kitov provided theoretical support for the latter to explain the enhancements in the inhibition of $AB_5$ toxins.[25] Their model emphasized the importance of a statistical term that describes in how many ways a multivalent ligand can bind to multiple binding sites. This was called avidity entropy. It was used to explain why an octavalent system was a better Shiga-like toxin inhibitor than the matched pentavalent one.

## Experimental procedure

### General

Unless stated otherwise, chemicals were obtained from commercial sources and used without further purification. Solvents were purchased from Biosolve (Valkenswaard, the Netherlands). Acid spacers **2**[13] and **4**[21] were synthesized following literature procedures. Compound **6** was purchased from Elicityl (France). Microwave reactions were carried out in a dedicated microwave oven: the Biotage Initiator. The microwave power was limited by temperature control once the desired temperature was reached. A sealed vessel of 2-5 mL was used. Analytical HPLC runs were performed on a Shimadzu automated HPLC system with a reversed-phase column (Reprospher 100, C8, 5 μm, 250 X 4.6 mm, Dr. Maisch GmbH, Germany), equipped with an evaporative light-scattering detector (PLELS 1000, Polymer Laboratories) and a UV/VIS detector operating at 220 and 254 nm. Preparative HPLC runs were performed on an Applied Biosystems workstation. Elution was effected using a gradient of 5% MeCN and 0.1% TFA in $H_2O$ to 5% $H_2O$ and 0.1% TFA in MeCN. $^1H$ NMR (400 MHz) and $^{13}C$ NMR (100 MHz) spectra were recorded on an Agilent 400-MR spectrometer. HSQC and TOCSY NMR (500 MHz) measurements were performed on a VARIAN INOVA-500. Electrospray Mass experiments were performed on a Shimadzu LCMS QP-8000. High resolution mass spectrometry (HRMS) analysis was recorded using Bruker ESI-Q-TOF II. The proton numbering scheme of all compounds can be found in the supporting information and was used in the assignments of the signals in the NMR spectra here.

### $CTB_5$ inhibition assay

A 96-well plate was coated with a solution of GM1 (100 μL, 2 μg/mL) in phosphate buffered saline (PBS). Unattached ganglioside was removed by washing with PBS and the remaining binding sites of the surface were blocked with BSA (1%) which was followed by washing with PBS. Samples of toxin-peroxidase conjugate (CTB-HRP; Sigma) and inhibitor in PBS with BSA (0.1%) and Tween-20 (0.05%) were incubated at room temperature for 2 h and were then transferred to the GM1-coated plate. After 30 min of incubation the solution was removed and the wells were washed with BSA

(0.1 %)/Tween-20 (0.05%) in PBS. To identify toxin binding to surface-bound GM1, the wells were treated with a freshly prepared solution of o-phenylenediamine/$H_2O_2$ in citrate buffer (100 mL) for 15 min. After being quenched with $H_2SO_4$, the absorbance in each well was measured at 490 nm.

## Analytical Ultracentrifugation Experiments

Mixtures of $CTB_5$ with various amounts of inhibitors were prepared within 1 h before analysis was carried out. Samples (0.4 mL) were centrifuged in 12 mm pathlength 2-sector aluminium centrepiece cells with sapphire windows in a An60Ti analytical rotor running in an Optima XL-I or Optima XL-A analytical ultracentrifuge (Beckman Instruments, Inc., Palo Alto, California 94304) at 60 krpm and at a temperature of 25 °C. Changes in solute concentration were detected by 300 absorbance scans measured at 280 nm over a period of 5-6 h. Analysis and fitting of the data was performed using the software SedFit.[26] A continuous c (s) distribution model was fitted to the data, taking every 2nd scan. The resolution was set at 200 over a sedimentation coefficient range of 0.0-20.0 S. Parameters were set for the partial specific volume as 0.73654 mL/g, the buffer density of 1.04910 g/mL and the buffer viscosity at 0.00141 Pa s, as calculated using SEDNTERP for phosphate buffered saline 0.1 M. The frictional coefficient, the baseline and the raw data noise were floated in the fitting. The meniscus and bottom of the cell path were also floated after initial estimations from the raw data.

## Compound 3

To a solution of tetra-amine **1**[5] (443 mg, 0.82 mmol) and spacer **2**[20] (1.7 g, 3.90 mmol) in dry DMF (15 mL), BOP (2.56 g, 5.79 mmol) and DIPEA (1.48 g 11.48 mmol) were added. The mixture was stirred at r.t. overnight and then concentrated under reduced pressure. The residue was purified by silica gel chromatography to afford **3** (780 mg, 50%). $^1H$ and $^{13}C$ NMR were consistent with ref.[5] MS (ESI) m/z calcd for $C_{105}H_{174}N_{14}O_{40}$ $(M+2H-2 \times Boc)^{2+}$ 1086.58, found 1086.75.

## Compound 5

Compound **3** (780 mg, 0.33 mmol) was treated with TFA in $CH_2Cl_2$ (1:1, 20 mL) 3 h at r.t., after which the volatiles were removed under reduced pressure and the residue

was dried under high vacuum. Meanwhile, compound **4** was prepared following the literature procedure.[21] The obtained amine TFA salt of **3** and the spacer **4** (670 mg, 2.30 mmol) were dissolved in anhydrous DMF (15 mL), and HATU (875 mg, 2.30 mmol) and DIPEA (892 mg, 6.90 mmol) were added. The mixture was stirred at r.t. overnight and then concentrated under reduced pressure. The residue was purified by silica gel chromatography to afford **5** (600 mg, 60%). $^{1}$H NMR (400MHz, $CDCl_3$): δ, ppm 7.70, 7.58, 7.40, 6.81 (14H, 4 x br t, J= 5 Hz, C(O)NH), 7.15 (2H, s, CH, aryl-2, aryl-6), 6.99 (4H, s, CH, 2 x aryl-2', 2 x aryl-6'), 6.70 (1H, s, CH, aryl-4), 6.54 (2H, s, CH, 2 x aryl-4'), 4.17, 4.06 (12H, 2 x br t, J= 5 Hz, $OCH_2CH_2NH$), 4.02, 3.99 (2 x 8H, 2 x s, $OCH_2C(O)$), 3.86 (3H, s, $C(O)OCH_3$), 3.82–3.75 (4H, m, $OCH_2CH_2NH$), 3.72–3.43 (120H, m, $OCH_2$, $OCH_2CH_2NH$), 3.40–3.32 (16H, m, $OCH_2CH_2N_3$, $CH_2NHC(O)$), 3.32-3.24 (8H, m, $CH_2NHC(O)$), 2.42 (8H, t, J= 5 Hz, $C(O)CH_2CH_2O$), 1.82–1.66 (16H, m, $OCH_2CH_2CH_2NH$). $^{13}$C NMR (100 MHz, $CDCl_3$): δ, ppm 171.16, 169.52, 168.85, 167.51 (C(O)NH), 166.68 ($C(O)OCH_3$), 159.84, 159.70 (C, aryl), 136.69 (C, aryl), 132.14 (C, aryl), 108.31 (CH, aryl-2, aryl-6), 106.76 (CH, aryl-4), 106.41 (CH, aryl-2', aryl-6'), 104.70 (CH, aryl-4'), 71.02 ($OCH_2C(O)$), 70.72–69.35 ($OCH_2$), 67.35 ($OCH_2$), 66.89, 66.62 ($OCH_2CH_2NH$), 52.40 ($C(O)OCH_3$), 50.74 ($OCH_2CH_2N_3$), 39.72, 38.46 ($OCH_2CH_2NH$), 37.22, 37.04 ($CH_2NHC(O)$), 36.81 ($C(O)CH_2CH_2O$), 29.38, 29.14 ($OCH_2CH_2CH_2NH$). MS (ESI) m/z calcd for $C_{134}H_{226}N_{26}O_{54}$ $(M+3H)^{3+}$ 1022.12, found 1022.40; HRMS (Q-TOF) m/z calcd for $C_{134}H_{226}N_{26}O_{54}$ $(M+H)^{+}$ 3064.5738, found 3064.6000, $(M+3H)^{3+}$ 1022.5257, found 1022.5348.

## Compound 7

A solution of tetravalent **5** (7 mg, 2.28 µmol), **6**, (14.8 mg, 13.4 µmol), sodium ascorbate (8.1mg, 41.1µmol) and $CuSO_4 \cdot 5H_2O$ (5.1 mg, 20.6 µmol) in DMF/$H_2O$ (1:1, 2 mL) was prepared and heated under microwave irradiation at 80°C for 20 min. After cooling down to room temperature, the copper salts were removed by a resin (Cuprisorb) and filtered off. The filtrate was then concentrated *in vacuo*, and the residue was purified by preparative HPLC and obtained by freeze-drying as an off-white powder (5 mg, 30%). $^{1}$H NMR (500 MHz, $D_2O$): δ, ppm 8.06, 7.96 (4H, 2 x s, CH, triazole), 7.15 (2H, s, CH, aryl-2, aryl-6), 6.83 (5H, s, CH, 2 x aryl-2', 2 x aryl-6', aryl-4), 6.68 (2H, s, CH, 2 x aryl-4'), 5.64, 5.15 (4H, 2 x d, $J_{1,2}$=8 Hz, $J_{1,2}$=8 Hz, $H_{Glc}$-1), 4.79 (4H,

$H_{GalNAc}$-1), 4.66-4.51 (24H, m, $NCH_2C_{triazole}$, $CH_2N_{triazole}$, $H_{Gal}$-1, $H_{Gal'}$-1), 4.27-4.07 (12H, m, $OCH_2CH_2NH$), 4.10, 4.04 (2 x 8H, 2 x s, $OCH_2C(O)$), 3.86 (3H, s, $C(O)OCH_3$), 3.77–3.65 (12H, m, $OCH_2CH_2NH$), 3.65-3.45 (112H, m, $OCH_2$), 3.41 (4H, t, $J_{2,3}$= $J_{2,4}$= 9 Hz, $H_{Gal}$-2), 3.31-3.17 (16H, m, $CH_2NHC(O)$), 2.69 (4H, dd, $J_{3a,3b}$= 13.5 Hz, $J_{3a,4}$= 4 Hz, $H_{NeuAc}$-3), 2.48 (8H, t, J= 6 Hz, $C(O)CH_2CH_2O$), 2.25 (12H, s, $NC(O)CH_3$), 2.04, 2.02 (2 x 12H, 2 x s, $NHC(O)CH_3$), 1.96 (4H, t, $J_{3b,3a}$= $J_{3b,4}$= 11.5 Hz, $H_{NeuAc}$-3), 1.78–1.67 (16H, m, $OCH_2CH_2CH_2NH$). $^{13}$C NMR (125 MHz, $D_2O$): δ, ppm 175.54, 175.27, 174.64, 174.24, 171.71 (COOH, C(O)NH), 164.71 ($C(O)OCH_3$), 160.05 (C, aryl), 145.38 (C, triazole), 132.49 (C, aryl), 125.67, 125.42 (CH, triazole), 109.43 (CH, aryl-2, aryl-6), 106.94 (CH, aryl-2', aryl-6', aryl-4), 105.55 (CH, aryl-4'), 105.35 ($C_{Gal'}$-1), 103.21 ($C_{Gal}$-1), 103.05 ($C_{GalNAc}$-1), 102.22 ($C_{NeuAc}$-2), 87.49, 82.91 ($C_{Glc}$-1), 80.79 ($C_{GalNAc}$-3), 78.65 ($C_{Glc}$-4), 77.41 ($C_{Glc}$-5), 77.30 ($C_{Gal}$-3), 75.50 ($C_{Gal'}$-5), 75.11 ($C_{Gal}$-4), 75.07 ($C_{GalNAc}$-5), 75.06 ($C_{Gal}$-5), 73.62 ($C_{NeuAc}$-6), 73.07 ($C_{Gal'}$-3), 72.54 ($C_{NeuAc}$-8), 71.28 ($C_{Gal'}$-2), 70.60 ($C_{Gal'}$-4), 70.54 ($OCH_2C(O)$), 70.50 ($C_{NeuAc}$-7, $C_{Gal}$-2), 70.49 ($OCH_2C(O)$), 70.12 ($OCH_2$), 70.02 ($OCH_2$), 69.99 ($OCH_2$), 69.25 ($C_{GalNac}$-4), 69.02 ($C_{NeuAc}$-4), 68.96 ($OCH_2$), 68.49 ($C_{Glc}$-3), 67.49 ($OCH_2CH_2NH$), 67.38 ($C_{Glc}$-2), 67.25 ($OCH_2CH_2NH$), 63.48 ($C_{NeuAc}$-9), 61.55 ($C_{GalNAc}$-6, $C_{Gal'}$-6), 61.05 ($C_{Gal}$-6), 60.66 ($C_{Glc}$-6), 53.39 ($C(O)OCH_3$), 52.15 ($C_{NeuAc}$-5), 51.86 ($C_{GalNAc}$-2), 50.71 ($CH_2N_{triazole}$), 40.36, 39.05 ($OCH_2CH_2NH$), 37.73 ($C_{NeuAc}$-3), 36.93 ($CH_2NHC(O)$), 36.75 ($NCH_2C_{triazole}$), 36.66 ($C(O)CH_2CH_2O$), 28.85 ($OCH_2CH_2CH_2NH$), 23.10 ($C_{GalNAc}$-$NHC(O)CH_3$), 22.61 ($C_{NeuAc}$-$NHC(O)CH_3$), 21.76 ($C_{Glc}$-1-$NC(O)CH_3$). HRMS (Q-TOF) m/z calcd for $C_{302}H_{486}D_8N_{38}O_{170}$ $(M+H)^+$ 7386.3064, found 7386.2000.

## Compound 8

The obtained tetramer **5** (305 mg, 0.10 mmol) was treated with Tesser's base (30: 9: 1, dioxane/MeOH/4 N NaOH, 5 mL). The mixture was stirred at r.t. until the total disappearance of the starting material. After that, the reaction was quenched by adding 1N $KHSO_4$ and the mixture was concentrated under reduced pressure. The residue was redissolved in $CH_2Cl_2$ (20 mL) and washed with 1N $KHSO_4$ (10 mL), $H_2O$ (10 mL) and brine (10 mL), dried on $Na_2SO_4$, and concentrated *in vacuo*. The resulting acid was used for the next step without further purification. To a solution of this acid and amine spacer **9** (O-(2-Aminoethyl)-O'-(2-azidoethyl)heptaethylene glycol, 70 mg,

0.16 mmol, Sigma-Aldrich) in dried DMF (10 mL), BOP (60 mg, 0.13 mmol) and DIPEA (40 mg, 0.31 mmol) were added. The mixture was stirred at r.t. overnight. Afterwards, the reaction was stopped and concentrated. The residue was suspend into $CH_2Cl_2$ (30 mL) and washed with 1N $KHSO_4$ (15 mL), 1N $NaHCO_3$ (15 mL), $H_2O$ (15 mL) and brine (15 mL), respectively. The organic layer was collected, dried on $Na_2SO_4$, and filtered. After concentration, the resulting material was purified by silica gel chromatography to afford **8** (175 mg, 0.05 mmol, 51% over two steps) as a colorless oil. $^1$H NMR (400 MHz, $CDCl_3$): δ, ppm 7.76, 7.67, 7.58, 7.47, 6.87 (15H, 5 x br t, J= 5 Hz, C(O)NH), 7.02 (2H, s, CH, aryl-2, aryl-6), 6.96 (4H, s, CH, 2 x aryl-2', 2 x aryl-6'), 6.63 (1H, s, CH, aryl-4), 6.52 (2H, s, CH, 2 x aryl-4'), 4.17, 4.05 (12H, 2 x b t, J= 5 Hz, OC*H*$_2$CH$_2$NH), 4.02, 3.98 (2 x 8H, 2 x s, $OCH_2C(O)$), 3.80–3.72 (4H, m, $OCH_2CH_2$NH), 3.71–3.44 (152H, m, $OCH_2$, $OCH_2CH_2$NH), 3.40–3.31 (20H, m, $OCH_2CH_2N_3$, C*H*$_2$NHC(O)), 3.31-3.23 (8H, m, C*H*$_2$NHC(O)), 2.42 (8H, t, J= 5 Hz, C(O)C*H*$_2$CH$_2$O), 1.80–1.68 (16H, m, $OCH_2CH_2CH_2$NH). $^{13}$C NMR (100 MHz, $CDCl_3$): δ, ppm 171.69, 169.56, 168.91, 167.56, 167.29 (C(O)NH), 159.97, 159.70 (C, aryl), 136.82 (C, aryl), 136.69 (C, aryl), 106.48 (CH, aryl-2, aryl-6), 106.35 (CH, aryl-2', aryl-6'), 104.81 (CH, aryl-4), 104.72 (CH, aryl-4'), 71.02 ($OCH_2C(O)$), 70.73–69.36 ($OCH_2$), 67.38 ($OCH_2$), 66.75, 66.62 ($OCH_2CH_2$NH), 50.75 ($OCH_2CH_2N_3$), 40.01, 39.75, 38.49 ($OCH_2CH_2$NH), 37.22 (C*H*$_2$NHC(O)), 36.93 (C(O)C*H*$_2$CH$_2$O), 29.39, 29.18 ($OCH_2CH_2CH_2$NH). MS (ESI) m/z calcd for $C_{151}H_{260}N_{30}O_{61}$ $(M+3H)^{3+}$ 1157.94, found 1157.65, $(M+2H)^{2+}$ 1736.91, found 1736.65; HRMS (Q-TOF) m/z calcd for $C_{151}H_{260}N_{30}O_{61}$ $(M+H)^+$ 3470.8165, found 3470.9000.

## Compound 10

A solution of pentavalent scaffold **8** (8.4 mg, 2.42 µmol), **6** (16 mg, 14.55 µmol), sodium ascorbate (6.92 mg, 35 µmol) and $CuSO_4{\cdot}5H_2O$ (4.35 mg, 17.4 µmol) in DMF/$H_2O$ (1:1, 2 mL) was prepared prepared and heated under microwave irradiation at 80°C for 20 min. After cooling down to room temperature, the copper salts were removed by a resin (Cuprisorb) and filtered off. The filtrate was then concentrated *in vacuo*, and the residue was purified by preparative HPLC and obtained by freeze-drying as an off-white powder (8.7 mg, 41%). $^1$H NMR (500 MHz, $D_2O$): δ, ppm 8.07, 7.96 (5H, 2 x s, CH, triazole), 7.02 (2H, s, CH, aryl-2, aryl-6), 6.85,

6.80 (5H, 2 x s, CH, 2 x aryl-2', 2 x aryl-6', aryl-4), 6.70 (2H, s, CH, 2 x aryl-4'), 5.64, 5.15 (5H, 2 x d, $J_{1,2}$=8.5 Hz, $J_{1,2}$=8.5 Hz, $H_{Glc}$-1), 4.78 (5H, $H_{GalNAc}$-1), 4.67-4.52 (30H, m, $NCH_2C_{triazole}$, $CH_2N_{triazole}$, $H_{Gal}$-1, $H_{Gal'}$-1), 4.27-4.09 (12H, m, $OCH_2CH_2NH$), 4.11, 4.05 (2 x 8H, 2 x s, $OCH_2C(O)$), 3.76–3.65 (12H, m, $OCH_2CH_2NH$), 3.68-3.45 (146H, m, $OCH_2$, $CH_2NHC(O)$), 3.40 (5H, t, $J_{2,3}$= $J_{3,4}$= 9 Hz, $H_{Gal}$-2), 3.29-3.19 (16H, m, $CH_2NHC(O)$), 2.69 (5H, dd, $J_{3a,3b}$=13 Hz, $J_{3a,4}$= 4 Hz, $H_{NeuAc}$-3), 2.49 (8H, t, J= 6 Hz, $C(O)CH_2CH_2O$), 2.26 (15H, s, $NC(O)CH_3$), 2.04, 2.02 (2 x 15H, 2 x s, $NHC(O)CH_3$), 1.96 (5H, t, $J_{3b,3a}$= $J_{3b,4}$= 11 Hz, $H_{NeuAc}$-3), 1.80–1.69 (16H, m, $OCH_2CH_2CH_2NH$). $^{13}$C NMR (125 MHz, $D_2O$): δ, ppm 175.54, 175.28, 174.29, 174.14, 172.33, 171.75 (COOH, C(O)NH), 160.07 (C, aryl), 145.47 (C, triazole), 136.36 (C, aryl), 125.41, 125.39 (CH, triazole), 107.40(CH, aryl-2, aryl-6), 106.89 (CH, aryl-2', aryl-6'), 106.01 (CH, aryl-4), 105.51 (CH, aryl-4'), 105.28 ($C_{Gal'}$-1), 103.16 ($C_{Gal}$-1), 103.01 ($C_{GalNAc}$-1), 101.76 ($C_{NeuAc}$-2), 87.38, 82.87 ($C_{Glc}$-1), 80.77 ($C_{GalNAc}$-3), 78.59 ($C_{Glc}$-4), 77.38 ($C_{Gal}$-3), 77.32 ($C_{Glc}$-5), 75.45 ($C_{Gal'}$-5), 75.03 ($C_{GalNAc}$-5), 74.98 ($C_{Gal}$-4), 74.73 ($C_{Gal}$-5), 73.64 ($C_{NeuAc}$-6), 73.04 ($C_{Gal'}$-3), 72.56 ($C_{NeuAc}$-8), 71.16 ($C_{Gal'}$-2), 70.59 ($C_{NeuAc}$-7, $C_{Gal}$-2), 70.44 ($OCH_2C(O)$), 70.37 ($OCH_2C(O)$), 70.16 ($C_{Gal'}$-4), 70.10 ($OCH_2$), 70.09 ($OCH_2$), 70.03 ($OCH_2$), 69.14 ($C_{GalNac}$-4), 68.97 ($C_{NeuAc}$-4), 68.94 ($OCH_2$), 68.43 ($C_{Glc}$-3), 67.41 ($OCH_2CH_2NH$), 67.32 ($C_{Glc}$-2), 67.26 ($OCH_2CH_2NH$), 63.41 ($C_{NeuAc}$-9), 61.49 ($C_{GalNAc}$-6, $C_{Gal'}$-6), 60.88 ($C_{Gal}$-6), 60.61 ($C_{Glc}$-6), 52.10 ($C_{NeuAc}$-5), 51.76 ($C_{GalNAc}$-2), 50.60 ($CH_2N_{triazole}$), 40.19, 40.17, 39.01 ($OCH_2CH_2NH$), 37.61 ($C_{NeuAc}$-3), 36.83 ($CH_2NHC(O)$), 36.70 ($NCH_2C_{triazole}$), 36.59 ($C(O)CH_2CH_2O$), 28.78 ($OCH_2CH_2CH_2NH$), 23.08 ($C_{GalNAc}$-$NHC(O)CH_3$), 22.58 ($C_{NeuAc}$-$NHC(O)CH_3$), 21.68 ($C_{Glc}$-1-$NC(O)CH_3$). HRMS (Q-TOF) m/z calcd for $C_{361}H_{591}D_4N_{45}O_{206}$ $(M-H)^-$ 8862.7567, found 8862.4000.

## Reference


[1] E. A. Merritt, W. G. J. Hol, **1995**, 165-171.
[2] L. De Haan, T. R. Hirst, *Mol. Membr. Biol*. **2006**, 21, 77-92.
[3] 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.
[4] R. J. Pieters, *Org. Biomol. Chem*. **2009**, 7, 2013-2025.
[5] D. Arosio, I. Vrasidas, P. Valentini, R. M. J. Liskamp, R. J. Pieters, A. Bernardi, *Org. Biomol. Chem.* **2004**, 2, 2113-2124.
[6] H. M. Branderhorst, R. M. J. Liskamp, G. M. Visser, R. J. Pieters, *Chem. Commun*. **2007**, 5043-5045.

[7] H.-A. Tran, P. I. Kitov, E. Paszkiewicz, J. M. Sadowska, D. R. Bundle, *Org. Biomol. Chem*. **2011**, 9, 3658-71.
[8] S. J. Richards, M. W. Jones, M. Hunaban, D. M. Haddleton, M. I. Gibson, *Angew. Chem. Int. Ed. Engl.* **2012**, 51, 7812-7816.
[9] S. Liu, K. L. Kiick, *Macromolecules* **2008**, 41, 764-772.
[10] E. K. Fan, Z. S. Zhang, W. E. Minke, Z. Hou, C. L. M. J. Verlinde, W. G. J. Hol, V. Uni, *J. Am. Chem. Soc.* **2000**, 122, 2663-2664.
[11] J. Garcia-Hartjes, S. Bernardi, C. A. G. M. Weijers, T. Wennekes, M. Gilbert, F. Sansone, A. Casnati, H. Zuilhof, *Org. Biomol. Chem*. **2013**, 11, 4340-4349.
[12] M. Mattarella, J. Garcia-Hartjes, T. Wennekes, H. Zuilhof, J. S. Siegel, *Org. Biomol. Chem*. **2013**, 11, 4333-4339.
[13] T. R. Branson, W. B. Turnbull, *Chem. Soc. Rev*. **2013**, 42, 4613-4622.
[14] V. Wittmann, R. J. Pieters, *Chem. Soc. Rev*. **2013**, 42, 4492-4503.
[15] Z. Zhang, E. A. Merritt, M. Ahn, C. Roach, Z. Hou, C. L. M. J. Verlinde, W. G. J. Hol, E. Fan, *J. Am. Chem. Soc.* **2002**, 124, 12991-12998.
[16] S. Cecioni, A. Imberty, S. Vidal, *Chem. Rev.* **2015**, 115, 525-561.
[17] A. V. Pukin, H. M. Branderhorst, C. Sisu, C. A. G. M. Weijers, M. Gilbert, R. M. J. Liskamp, G. M. Visser, H. Zuilhof, R. J. Pieters, *Chembiochem* **2007**, 8, 1500-1503.
[18] C. Sisu, A. J. Baron, H. M. Branderhorst, S. D. Connel, C. A. G. M. Weijers, R. de Vries, E. D. Hayes, A. V. Pukin, M. Gilbert, R. J. Pieters, H. Zuilhof, G. M. Visser, W. B. Turnbull, *ChemBioChem* **2009**, 10, 329-337.
[19] T. R. Branson, T. E. McAllister, J. Garcia-Hartjes, M. A. Fascione, J. F. Ross, S. L. Warriner, T. Wennekes, H. Zuilhof, W. B. Turnbull, *Angew. Chem. Int. Ed. Engl.* **2014**, 53, 8323-8327.
[20] R. Autar, A. S. Khan, M. Schad, J. Hacker, R. M. J. Liskamp, R. J. Pieters, *ChemBioChem* **2003**, 4, 1317-1325.
[21] H. Herzner, H. Kunz, *Carbohydr. Res*. **2007**, 342, 541-557.
[22] G. I. Tesser, I. C. Balvert-Geers, *Int. J. Pept. Protein Res*. **1975**, 7, 295-305.
[23] P. Schuck, M. A. Perugini, N. R. Gonzales, G. J. Hewlett, D. Schubert, *Biophys. J.* **2002**, 82, 1096-1111.
[24] J. Lebowitz, M. S. Lewis, P. Schuck, *Protein. Sci*. **2002**, 2067–2079.
[25] P. I. Kitov, D. R. Bundle, *J. Am. Chem. Soc*. **2003**, 125, 16271-16284.
[26] P. Schuck, *Biophys. J*. **2000**, 78, 1606-1619.

# Chapter 4

## Carbohydrate-linked Hyperbranced Polyglycerol Polymers against Cholera Toxin

## Introduction

Carbohydrate-protein interactions play significant roles in biological processes, such as cell-cell communication, viral and bacterial infection, inflammation, tumor cell metastasis, immune responses, and others.[1] The weak intrinsic binding constants of native carbohydrate epitopes, are often compensated by a multivalent presentation of the ligands leading to potency enhancements, the so-called "glycoside cluster effect".[2] One example where this plays a role is cholera toxin (CT), a pathologically active protein secreted by the *Vibrio cholerae* bacteria.[3] The arrangement of the five identical GM1 binding B-subunits around the central A subunit enables the toxin to strongly bind to the carbohydrate moieties of the GM1 gangliosides on the intestinal cell surface in a simultaneous multivalent manner. Owing to the detrimental effect of the cholera toxin on our health, interfering with its attachment to the cell-surface has attracted a wide interest. The development of strong binding inhibitors of CT would not only be beneficial for the disease treatment but also useful for the toxin detection in infected samples leading to front-line defence.[4] For this reason the development of high-affinity agents for CT is very important and the multivalency effect has been a key principle in their design. As an attractive group of bioactive compounds it is notable that multivalent carbohydrate constructs including carbohydrate clusters, glycopolymers and glycopeptides have already been introduced in the early phase of the drug development process.[5]

To exert its toxic effects, CT needs to bind to at least one copy of its native ligand, ganglioside GM1 [Galβ1-3GalNAcβ1-4(Neu5Acα2-3) Gal-β1-4Glc-Ceramide], displayed on the exterior of intestinal epithelial cells.[6] A reported crystal structure of the $CTB_5$-GM1 complex indicated that the terminal galactose could serve as an 'anchor' point in potential antagonist design.[7] Using the information obtained from crystallographic analysis, an anchor-based approach was exploited by Minke *et al.* to design and optimize small molecule antagonists, which unveiled *meta*-nitrophenyl-α-D-galactopyranoside (MNPG) (**Fig.1**) to be the most effective antagonist among a screen of 35 commercially available O-1 substituted galactose derivatives. It displayed an approximate 100-fold increase in affinity over free galactose.[8] Followed by Fan and Hol *et al.*, a library of antagonists incorporating the MNPG core

was designed. One of their two lead compounds, morpholine derivatized MNPG (**Fig.1A**) showed enhanced affinity in the low micromolar range ($K_d$= 12μM).[9] However, the modest overall affinity of these compounds compared to native GM1os (low micromolar vs. low nanomolar $K_d$[10]) still left much room for improvement. For that reason, it is our aim to incorporate an optimized set of monovalent antagonist into a multivalent presentation.

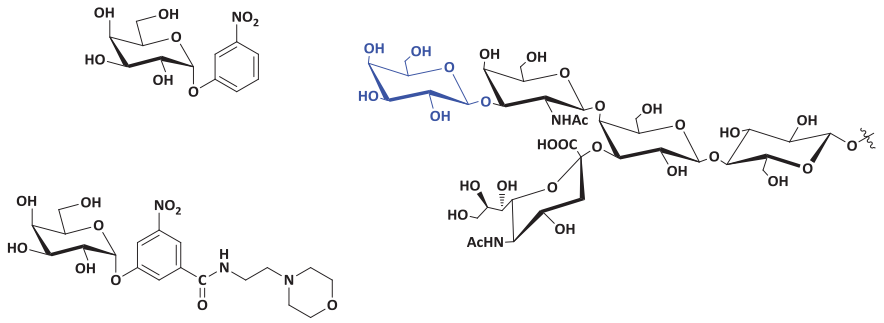


**Fig 1A.** General structure of MNPG and its morpholine derivative and GM1os.[8, 9]

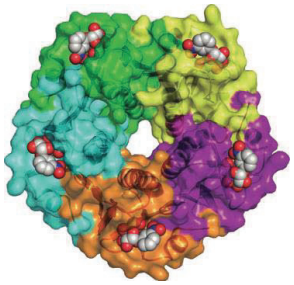


**Fig 1B.** Crystal structure of $CTB_5$ with MNPG bound to each monomer. The image was generated using PyMol (http://www.pymol.org) from co-ordinates deposited in the protein data bank (PDB: 1eei).[8b]

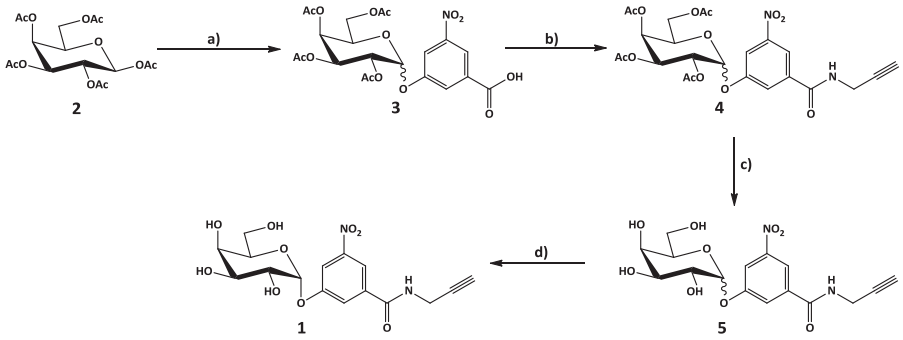
With specific features of being highly biocompatible, narrowly disperse, globular and having a dendritic structure, hyperbranced polyglycerols (hPGs)-based scaffolds have become increasingly important in the field of multivalency applications.[11] By an efficient one-pot synthesis involving epoxide ring-opening multibranching polymerization of glycidol can be achieved,[12] The high density of peripheral hydroxyl end-groups on the resulting hPGs polymer can be accessed for further functionalization via established chemical methods. Recently, it was reported by Haag and co-workers that hPGs were ideal scaffolds for the multivalent presentation

of galactose units. The galactose functionalized hPGs showed strong inhibition in an *in vitro* selectin binding assay with $IC_{50}$'s value in the nanomolar range.[11a] Moreover, in our previous study, we had also shown that a highly active CTB inhibitor can be made by using a multivalent dendritic galactose system. The inhibitory potency had been improved by orders of magnitude.[13] The combination of several of these observations prompted us to make use of the hyperbranched polymer hPGs as a cheap and water-soluble multivalent scaffold and outfit it with the interested and underexplored MNPG ligand to create a cheap and highly potent anti-cholera compound. In other words we aimed to develop a time-saving and cost-effective strategy combining a multivalent design and CuAAC ligation methodology to create a cost-effective anti-cholera polymer.

## Results

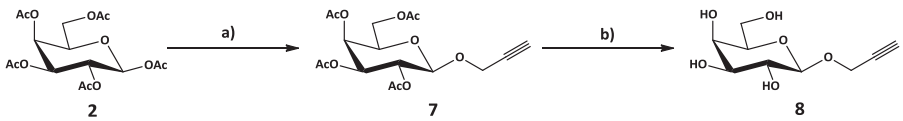
As indicated by a crystal structure study of MNPG in complex with $CTB_5$,[14] two key features of the MNPG binding mode were well-defined: 1) the surprising α-linkage preference for this compound could be explained as the binding mode places the α-linked O-1 aglycone closer to the surface of the binding site affording better shape complementarity; 2) the presented nitrophenyl could displace a water molecule at the canonical water-binding site, where one of the nitro O atoms could be placed at the top position to duplicate the H-bonding.[8] For this particular reason, these very special arrangements should be retained in an appropriate synthesis of the MNPG conjugate. Furthermore, it has recently been shown by Haag *et.al* [11a] that hPGs bearing peripheral carbohydrate adducts could be efficiently prepared by grafting unprotected sugar derivatives onto the polymer core with CuAAC. The specific MNPG ligand **1** was prepared which contains a propargyl moiety to allow its subsequent conjugation to the polymer.

The synthetic route to **1** is shown in **Scheme 1**. The acid-bearing MNPG analogue **3** was synthesized according to a previously reported procedure.[15] Commercially available galactose pentaacetate **2** was used for the glycosylation and 3-hydroxy-5-nitrobenzoic acid with $SnCl_4$ as a Lewis acid, which afforded **3** in a poor yield of 10%.



**Scheme 1.** a) 3-hydroxy-5-nitro benzoic acid, $SnCl_4$, $CH_2Cl_2$, 10%; b) propargylamine, EDCI, DMAP, DMF, 35%; c) NaOMe, MeOH, 89%; e) β-galactosidase, $NH_4HCO_3$ (pH=7.3), 45%.

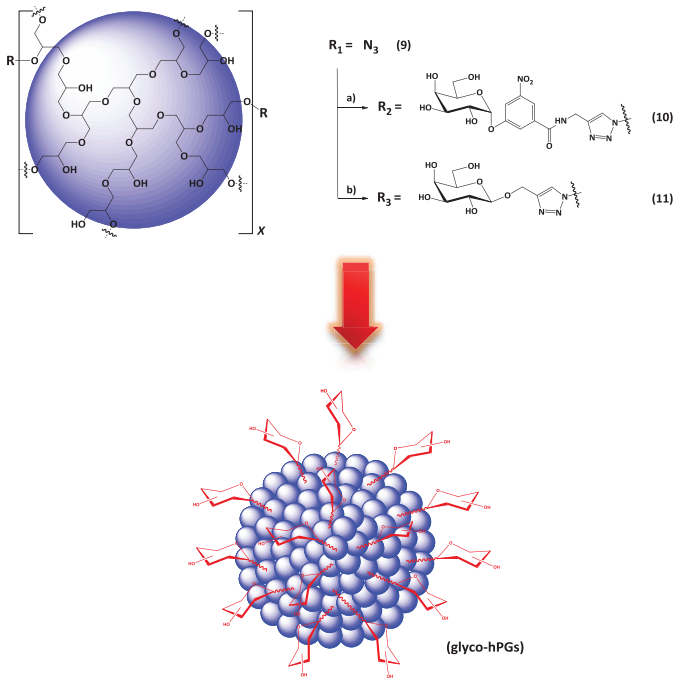
This low-yielding glycosylation might be associated with the carboxylic acid in compound **3**, where the hydroxyl group could be substituted by another oxygen nucleophile on the 3-hydroxy-5-nitrobenzoic acid molecule. To enable the conjugation to hPGs, a propargylamine tether was selected and attached to **3** through amide bond formation using EDCI and DMAP. The resulting alkynylated compound **4** was obtained in a moderate yield of 35%, of which the acetyl protecting groups were selectively removed under Zemplén conditions to give **5** as a mixture of both α and β anomers in 89% yield. Subsequently, the resulting mixture of **5** was subjected to enzymatic digestion using a β-galactosidase in a buffer (pH= 7.3) and produced **1** as the sole α-anomeric form.



**Scheme 2.** a) i) $I_2$, HMDS, $CH_2Cl_2$; ii) propargyl alcohol, $Ag_2CO_3$, $CH_2Cl_2$, 62.5% over two steps; b) NaOMe, MeOH, 91%.

As described before, galactose itself can also bind to CTB with a relatively low affinity, having an approximate $IC_{50}$ in the high milimolar range (≈100 mM) by a CT GD1b Enzyme-Linked Adhesion Assay.[8] This is also identified by our previous work,[13] in which the monovalent galactose exhibited weak inhibitory potency with an $IC_{50}$ of 240 mM. Nevertheless, after enhancement by the multivalency effect, a galactose-

containing dendrimer in its octavalent form showed an $IC_{50}$ which was roughly equal to a GM1os derivative (ca. 20-30 micromolar). With this in mind, galactoside **8** was also prepared for comparison to the MNPG derivative. The synthesis of **8** was shown in **Scheme 2**. It started with the same galactose pentaacetate **2**, where the acetyl group at the anomeric center was replaced by an iodine in the presence of $I_2$ and HMDS. The anomeric halide was subsequently reacted in a Koenings-Knorr glycosylation reaction with propargyl alcohol using $Ag_2CO_3$ as the promotor to afford **7** in 63% yield over two steps. A final deacylation step was performed by reaction of **7** with NaOMe in MeOH to generate the desired compound **8** in 91% yield.




**Scheme 3.** a) **1**, $CuSO_4 \cdot 5H_2O$, Na ascorbate, $THF/H_2O$, r.t.; b) **8**, $CuSO_4 \cdot 5H_2O$, NaAsc $MeOH/H_2O$, r.t. or microwave irradiation.

The hyperbranched polyglycerols (hPGs) provided by the Haag group were prepared via an inverse miniemulsion-templated, acid-catalyzed, ring-opening polymerization of glycerol triglycidyl ether.[16] Unreacted glycidyl epoxides were then opened by an azide nucleophile, resulting in globular polymer **9** with an average molecular weight

of 100 kDa and a degree of azide functionalization of 70% according to the $^{1}H$ NMR spectrum. As shown in **Scheme 3**, the reactions were carried out in THF/$H_2O$/MeOH at room temperature or under microwave irradiation. Employing a developed protocol for CuAAC, azide-containing polyglycerol **9** was reacted with proparglyated MNPG **1** to seemingly furnish the glycopolymers **10**. In parallel, galactose derivative **8** was also linked to the polymer in a similar fashion to afford **11** as a simplified version. The resulting glycopolymers **10** and **11** were both isolated and purified from the reaction mixture by dialysis for further characterization.

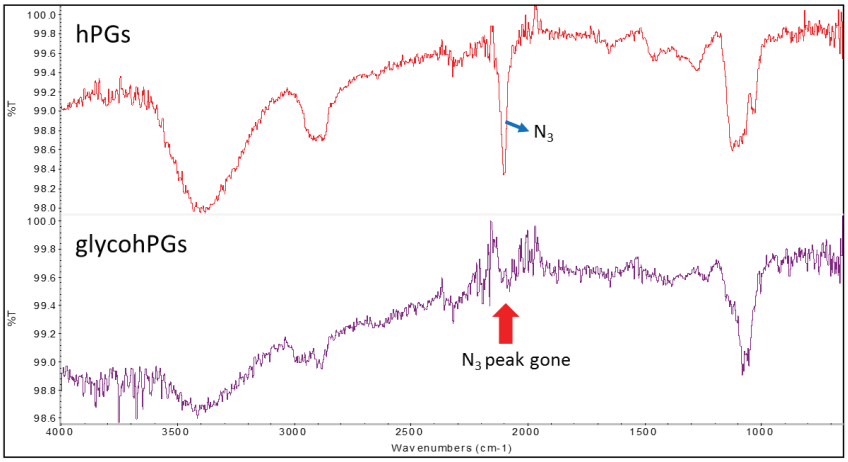
## Characterization

Using click chemistry as a conjugation technique, MNPG and galactose-decorated hPGs **10** and **11** were prepared by grafting terminal alkyne galactoside derivatives with azide-containing polyglycerol polymers, respectively. During the course of this conjugation, IR spectroscopy, as an analytical tool, was particularly helpful to verify the completion of the CuAAC reaction. As shown in **Fig.2A**, compared to the spectrum of azido-hPGs **9**, the IR data of post-glycosylation (compound **11**) showed that the peak corresponding to azide (ca. 2098 $cm^{-1}$) residue was no longer present in the final glycostructures. The strong bands around 3300 $cm^{-1}$ in both IR spectra were derived from the OH groups but not from the alkyne residue (ca. 3289 $cm^{-1}$, C≡H). The excess of alkyne precursors was completely eliminated in the dialysis step. Indeed, one thing is worthy to be clarified. In the case of the glyco-hPGs, no azide stretching peak was observed at 2098 $cm^{-1}$. In order to verify the dispersity, the degree of functionalization of the end glycopolymer adducts were characterized by an enzymatic method using a galactose assay kit (Sigma-Aldrich). It was determined that the galactose-coated hPGs **11** was prepared with 88 sugar units per polymer (Mw=118,530 Da).

Meanwhile, the particle sizes of the resultant glycoconjugates were also quantified by dynamic light scattering (DLS) measurements. For compound **11**, the average diameter was about 35 nm, which was significantly larger than hPGs itself (19 nm) (**Fig.2B**). This observation might be associated with the limited swelling behaviour of the polymer in an aqueous solution. On the other hand, the increase in particle size could also indicate that the galactose units were properly attached to the hydrogels.

Similar analysis on MNPG-coated hPGs analogue was conducted as well. However, the obtained result displayed a much expanded size near 257 nm.

**A)**




**B)**

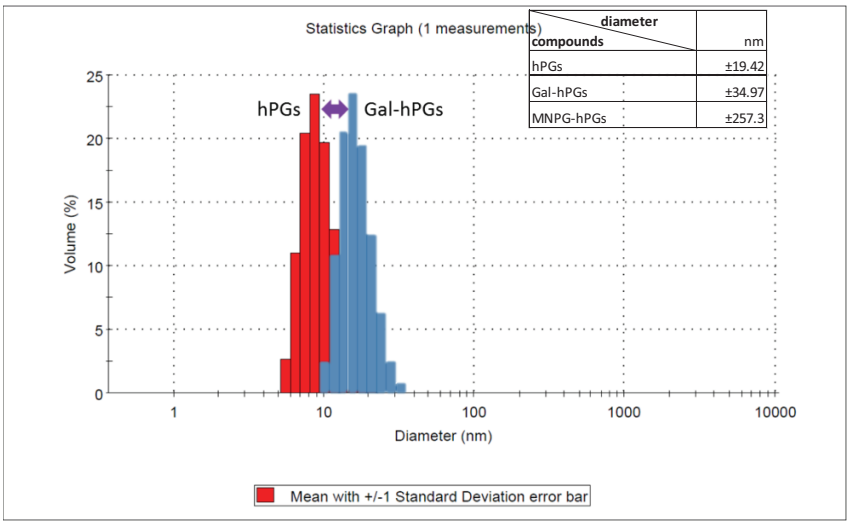



**Figure 2.** A) IR spectra before and after post-glycosylation (hPGs **9** vs **11**); B) DLS experiment of particle size of hPGs (red) and gal-hPGs (blue)

Prior to their biological evaluation, the obtained end-products were dialyzed to totally remove the copper-based impurities and unreacted sugar ligands, and they were then lyophilized. The $^{1}$H NMR spectra of **11** in comparison with the unsubstituted polymer **9** (**Fig.3**) showed a similar broadening of methylene signals indicating that no side reactions had occurred on the polymer backbone. The 1,2,3-

triazole five-membered ring, formed after CuAAC coupling could be observed in the spectrum of Gal-hPGs **11**, at around 8.1 ppm. In addition, the NMR spectrum of **11** also showed extra signals in the 4.00-5.00 ppm region, which were attributed to the saccharide **8** after linkage to the hPGs. According to these data it is clear that the carbohydrate attachment to the polymer had succeeded.

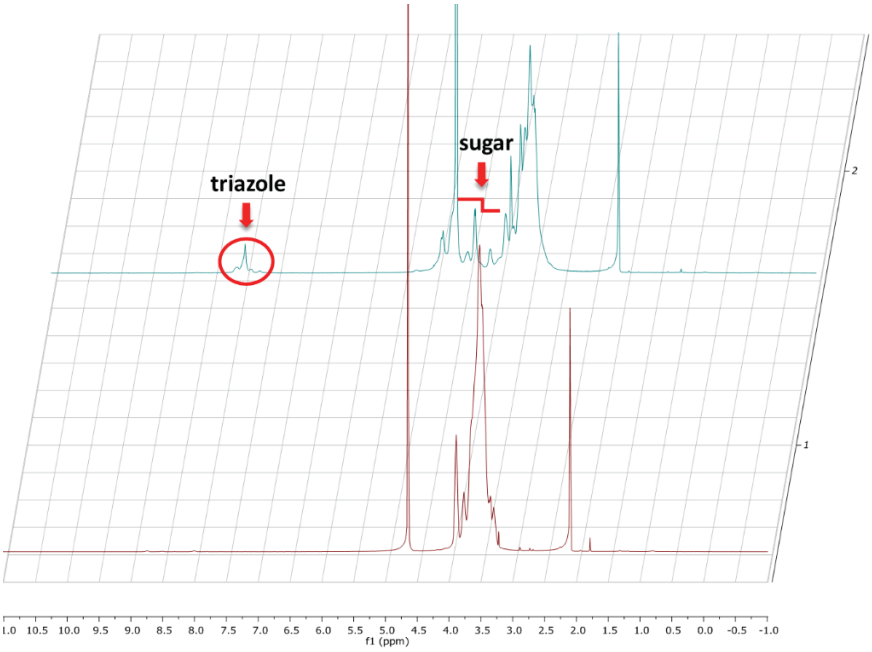



**Figure 3.** NMR spectra of hPGS **9** and Gal-hPGs **11**

Unfortunately, under the same circumstances, the click approach did not seem to work properly for the MNPG conjugation. The analysis of the resultant glycoconstruct was found to be very difficult due to a low degree of ligand conjugation. Unlike the spectrum of Gal-hPGs **11**, all the peaks presented in **Fig.4** were relatively weak, especially the specific signals from 1,2,3-triazole ring and sugar fragment could not be clearly observed in the spectrum. In this case, the overall yield of end-product **10** was unexpectedly less than 10%.

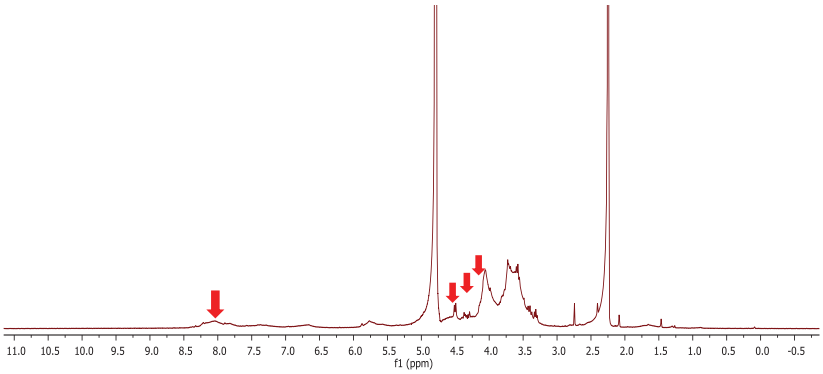



**Figure 4.** NMR spectrum of MNPG-hPGs **10**

The establishment of a multivalent display of MNPG ligands was not very successful. With the click reaction, something unanticipated was taking place. The hPGs used was the same for both the synthesis of **11** and the attempted synthesis of **10**. The only difference was the employment of the galactoside residue instead of its MNPG counterpart. Owing to that fact, we hypothesized that this uncertain issue could be contributed to the MNPG ligand. With a detailed look at the NMR spectra, the origin of the problem became clear. To our surprise, as shown in **Fig.5A**, it was found that compound **5**, before treatment with galactosidase, showed a clear peak corresponding to alkyne hydrogen at δ 2.68 ppm. However, after enzymatic dialysis and prep-HPLC purification, the spectrum of compound **1** showed a much reduced alkyne hydrogen signal**,** which was eventually lost, while the rest of the signals remained essentially unchanged (green vs blue spectra). Similar phenomena were also observed in the APT-$^{13}$C NMR, of which the specific peak of terminal *sp* carbon at δ 75.75 ppm had disappeared. It was suggested that the disappearance of the peak belonging to the end-alkyne group might be associated with the cleavage of β-anomer by the enzyme. Additionally, mass spectra further confirmed the observation. The observed mass of the unknown compound ($(MS+H)^+$= 765.15) was indicative of a dimer since the spectrum of mono MNPG **1** ($(MS+H)^+$= 383.20) (**Fig.5B**).

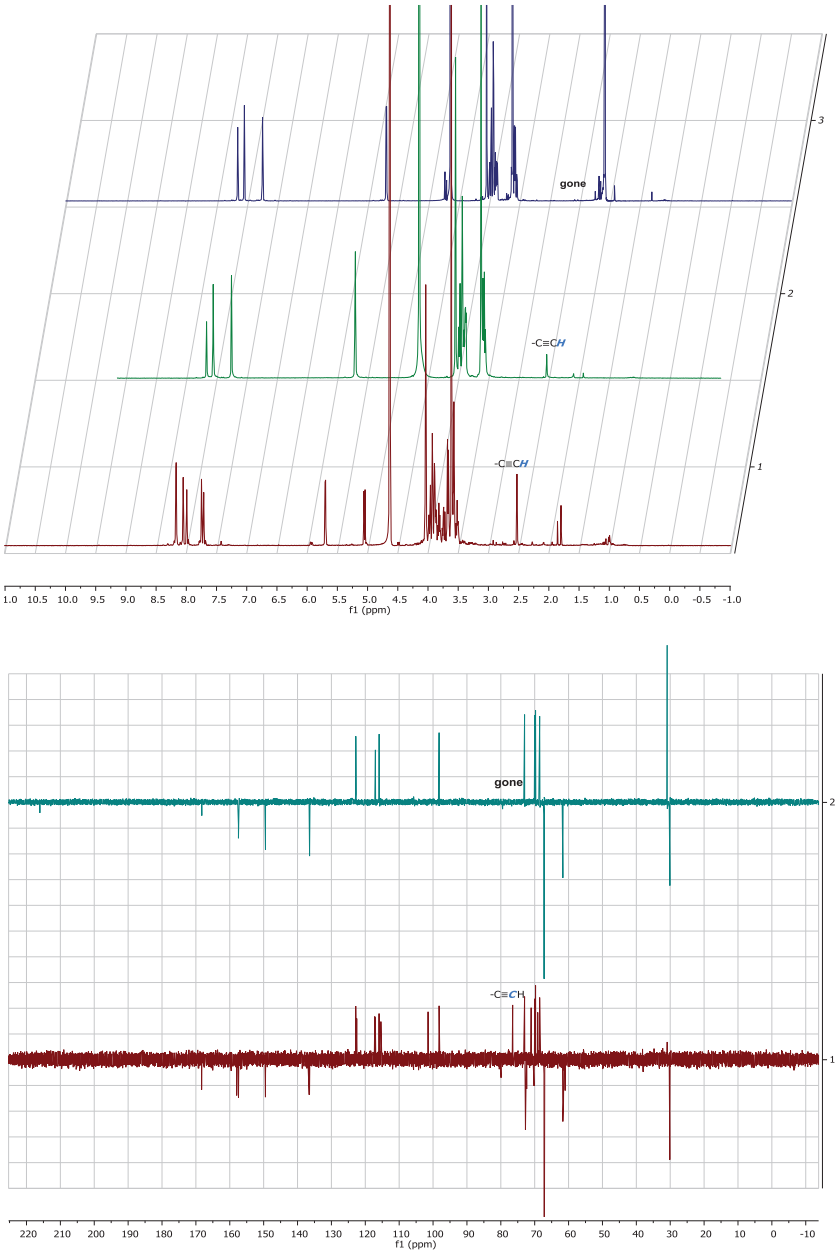



**Figure 5A.** $^{1}H$ and APT-$^{13}C$ NMR of MNPG ligand dimerization, both specific peaks of terminal alkyne group, at δ 2.68 ppm of $^{1}H$ and δ 75.75 ppm of $^{13}C$ were eventually disappeared.

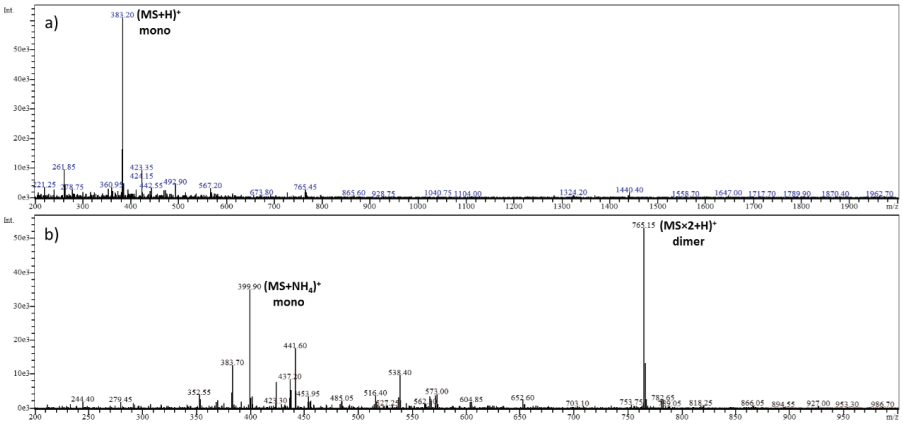



**Figure 5B.** ESI-MS of MNPG ligand dimerization in the course of time. a) freshly prepared solution of MNPG ligand; b) the same solution stored overnight.

All of these results indicated that the missing alkyne signal might be caused by a Glaser/Hay[17]-like oxidative coupling of the two propargyl MNPG ligands, which afforded a symmetrical diyne as depicted in **Fig.6**. The occurrence of this undesired alkyne cross-linking could be caused by: the presence of trace magnesium chloride and Tris salts in the enzyme, the acidity of the terminal alkyne group the attached nitro group (-$NO_2$) on the aromatic ring, or the nitro group could possibly alter the oxidation behaviour of the compound. All of them could trigger the C-C formation under modified oxidative conditions. As a consequence, the resultant apparently dimerized compound made the conjugation between MNPG ligand **1** and hPGs **9** impossible. Surprisingly this behaviour was not seen for **11**.

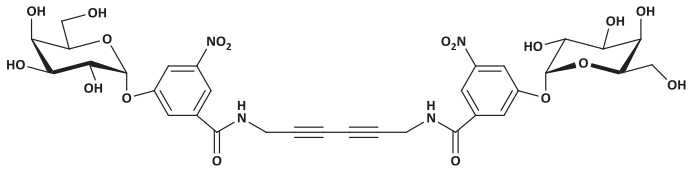


**Figure 6.** 1,3-diyne coupling product of the MNPG ligand

Subsequent attempts are proposed to circumvent this problem. For a longer lasting MNPG ligand, it could be possible to install a protecting group at terminal alkyne e.g. a silyl group, which could be selectively and easily removed at a later stage prior to the click reaction. Secondly, the unfavourable MNPG self-coupling could possibly also be avoided by attaching the mixture **5** to the hPGs firsts. Then afterwards, the

resultant glycoconjugate could be treated with the galactosidase to remove the β-anomers, leading to the polymer bearing the MNPG units only in the desired α configuration. Besides that, the remaining aglycone parts of pre-cleaved β isomers could serve as dummy ligands by this approach. In this manner, it could represent another advantage for ligand density control, an important issue for multivalent inhibitor design.

## Inhibition assay

The inhibition performance of galactose-functionalized hPGs **11** was determined by an ELISA type assay against $CTB_5$[13] was performed by H. C. Quarles van Ufford (Linda) at Utrecht University, where toxin- horseradish peroxidase conjugate ($CTB_5$-HRP) with a varied concentration of glycopolymer inhibitors were incubated and tested on a 96 well GM1 coated plate via absorbance determination. The previously reported $IC_{50}$ value of galactose was used in this study as the monovalent inhibitor reference. The results are shown in **Table 1**. As reported in the prior study, the monovalent reference compound was a weak inhibitor and could only inhibit CT binding with an $IC_{50}$ value in the submolar range ($IC_{50}$= 0.24 M). In comparison, a significant binding performance was observed for gal-hPGs **11** having a low micromolar range inhibition concentration ($IC_{50}$=1.3 μM), where it exhibited more than five orders of magnitude increase in affinity over the mono galactose. On a per sugar basis, the binding efficiency of **11** was 2100-fold more potent than the monovalent counterpart. Clearly, this observed strong CT inhibition effect was a result of the multivalency effect, where the numerous galactose peripherals promoted the rebinding events at a high rate. The data shown in this study indicated that the synthetic glycopolymer **11** could act as a multivalent inhibitor of CT binding.

Surprisingly, even a simplified glycol-construct **11** could afford relatively high inhibition efficiency against CT. Even so, there is still more that needs to be addressed between the carbohydrate peripherals and the toxin with respect to the underlying the multivalent interaction. Indeed, the inhibition strength could also be featured with many factors. As presented by Haag *et al.*[19] using sialic acid-containing hPGs in a flu virus activity study, the binding properties could be varied with varying ligand densities. There could exist a saturation point on polymer

functionalization. After reaching that plateau, the inhibition would not be significantly improved above that point. Apart from the ligand density, it was also mentioned that the particle size could be an important element for refined binding performance when a certain matching particle size was reached.[19] In addition, there are many other factors e.g. an optimum spacing distance in-between ligand and polymer, a beneficial spacer unit connecting the sugar appendages and polymer could further alter the efficiency and specificity of the glycopolymer as a CT inhibitior. Nevertheless, all the results shown above validated the possibility and promise of appending MNPG onto a polymeric scaffold for a multivalency implementation.

**Table 1.** Potency of the inhibitors measured by using an ELISA-type assay.[a]

| Compound | ELISA $IC_{50}$ (M) | rel.pot. (per sugar) |
|---|---|---|
| **Galactose** | $2.4 \times 10^{-1}$[b] | 1 (1) |
| **11** | $1.3 \times 10^{-6}$ | 185,000 (2100) |

[a] Determined in an ELISA with $CTB_5$-HRP (0.43 nM) and wells coated with GM1.

[b] $IC_{50}$ value reported previously[13] using same assay conditions was used in this study

## Conclusion

In summary, a new class of multivalent glycopolymers had been invented as potential inhibitors against CT. In this study, we synthesized two simple carbohydrates such as galactose and MNPG, which were grafted to hPGs' scaffold in a multivalent manner via a "click" reaction. Although the introduction of MNPG to the hPGs was problematic due to the unexpected self-dimerization of the ligand, the galacto-construct **11** as a simplified version was readily made with 88 sugar units on it. Its inhibition strength was investigated by an ELISA type assay against CT. A strong binding performance was observed for **11** ($IC_{50}$= 1.3 µM) with a 2100-fold increased affinity over the mono reference compound. This work is still at an early stage. However, it highlights the potential of an MNPG-containing polymer **10** being potentially a more effective and strong CT inhibitor. Current work is on developing a new protocol for the preparation of MNPG-functionalized hPGs.

## Experimental procedure

### General Information

$^{1}H$ NMR and $^{13}C$ NMR spectra were recorded on an Agilent 400-MR spectrometer (400 and 100 MHz for $^{1}H$ and $^{13}C$, respectively) .The spectra were calibrated using the residual solvent signal as internal standard. FT-IR spectra were recorded on a PerkinElmer UATR TWO FT-IR spectrometer operating from 4000-650 $cm^{-1}$ on the Universal diamond ATR top-plate. ESI-MS spectra samples were measured on Shimadzu LCMS QP-8000. High resolution mass spectrometry (HRMS) analysis was obtained by using the Bruker ESI-Q-TOF II.

### Enzymatic Method for Determining Galactose (Sigma-Aldrich)

This assay protocol is suitable for the colorimetric detection of Galactose and Lactose in cell and tissue culture supernatants, urine, plasma, serum, and other biological samples using the Galactose Assay Kit (MAK012). In this assay kit, galactose is oxidized by galactose oxidase resulting in a colored (570 nm) product, proportional to the galactose present. This kit has a linear detection range of 2–10 nmoles galactose for the colorimetric assay.

### Size characterization

Glycoparticles were suspended in distilled deionized water/methanol and their average size and size distribution were measured by DLS using a Malvern CGS-3 multiangle goniometer (Malvern Ltd., Malvern, U.K.) with a JDS uniphase 22mW He-Ne laser operating at 632nm, an optical fiber-based detector and a digital LV/LSE-5003 generator, measurement angle 90°).

### Materials

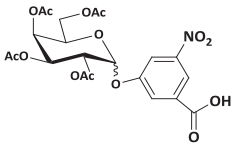
Reactions requiring dry or oxygen-free conditions were carried out under argon (Schlenk conditions). All reagents and solvents were purchased from commercial suppliers and used without further purification. All solvents were reagent grade and used as received. Dry solvents were purchased from Biosolve and kept on 3A molecular sieves. TLC was performed on silica gel plates (60 F254, 20 cm x 20 cm, Merck), detection was effected by charring with 10% sulphuric acid in methanol, followed by heat treatment. Flash chromatography was performed on silica gel 60 Å

(230-400 mesh, particle size 40-63 µm, SiliCycle). Dialysis was performed in a Slide-A-lyzer cassette (Thermo-Scientific, 10,000 MWCO) at 4$^{o}$C. The purification time was between 24 and 48 h against MiliQ water until no yellowish color was present in the buffer and the buffer was changed after every 4 hours.

## $CTB_5$ inhibition assay

A 96-well plate was coated with a solution of GM1 (100 µL, 2 µg/mL) in phosphate buffered saline (PBS). Unattached ganglioside was removed by washing with PBS and the remaining binding sites of the surface were blocked with BSA (1%) which was followed by washing with PBS. Samples of toxin-peroxidase conjugate ($CTB_5$-HRP; Sigma) and inhibitor in PBS with BSA (0.1%) and Tween-20 (0.05%) were incubated at room temperature for 2 h and were then transferred to the GM1-coated plate. After 30 min of incubation the solution was removed and the wells were washed with BSA (0.1 %)/Tween-20 (0.05%) in PBS. To identify toxin binding to surface-bound GM1, the wells were treated with a freshly prepared solution of o-phenylenediamine/$H_2O_2$ in citrate buffer (100 mL) for 15 min. After being quenched with $H_2SO_4$, the absorbance in each well was measured at 490 nm.

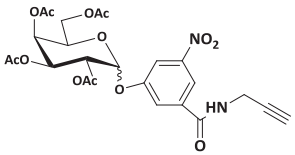
## Compound 3 [14]




1,2,3,4,6-penta-o-acetyl-β-D-galactopyranoside (5 g, 12.82 mmol) in anhydrous $CH_2Cl_2$ (50 mL) was mixed with 3-hydroxy-5-nitrobenzoic acid (2.403g, 13.12 mmol) and stirred under $N_2$. The mixture was heated until reflux and then Tin chloride (IV) neat solution (1.6 mL, 16.03 mmol) was added drop wise via syringe. The reaction was stirred at same condition for 96 h after which the reaction was quenched with the addition of cold $H_2O$ (20 mL). The organic layer was separated and washed with cold $H_2O$ (20 mL) five more times and once with brine (20 mL). The combined organic layer was dried over $Na_2SO_4$, filtered and concentrated. The residue was purified by silica gel chromatography to afford the title compound (360 mg) as an α and β mixture in 1:1 ratio. $^1$H NMR (400MHz, DMSO): δ, ppm 8.36 (s, 2 x CH, aryl), 8.12-

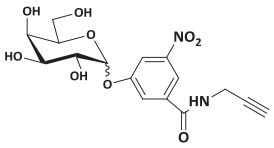
8.02 (m, 2 x CH, aryl), 7.89 (s, 2 x CH, aryl), 6.10 (d, 1H, $J_{1,2}$= 4Hz, H-1$_\alpha$), 5.73 (d, 1H, $J_{1,2}$= 8Hz, H-1$_\beta$), 5.47-5.36 (m, 6H, 2 x H-3, 2 x H-4), 5.30-5.21 (m, 2H, 2 x H-2), 4.19-3.95 (m, 6H, 2 x H-5, 2 x H-6a, 2 x H-6b), 2.20-1.73 (m, 12H, acetyl). $^{13}$C NMR (100MHz, DMSO): δ, ppm 170.07-169.28 (C(O)OH, *C*(O)CH$_3$ ), 156.41, 155.84, 148.38 (C aryl), 123.94, 122.90, 117.86 (CH aryl), 97.43 (C-1$_\beta$), 94.85 (C-1$_\alpha$), 70.96, 69.99 (2 x H-5), 68.11, 67.85 (2 x C-4), 67.54, 67.40 (2 x C-3), 66.98, 66.59 (2 x C-2), 61.84, 61.46 (2 x C-6), 20.49-20.15 (CH$_3$, acetyl). MS (ESI) m/z calcd for $C_{21}H_{23}NO_{14}$ $(M-H+TFA)^-$, found 626.05.

## Compound 4



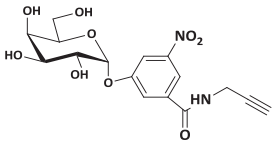

To a solution of **3** (278 mg, 0.54 mmol) and propargyl amine (3 3mg, 0.60 mmol) in anhydrous DMF (10 mL), EDCI ((1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide, 101 mg, 0.65 mmol) and DMAP (79.5 mg, 0.65 mmol) were added. The mixture was stirred at r.t. overnight. Upon the completion of reaction, it was stopped and the solvents was removed *in vacuo*. The residue was purified by silica gel chromatography to afford amide **4** (90.8 mg, 0.17 mmol, 31%) as a yellowish oil. $^1$H NMR (400MHz, $CDCl_3$): δ, ppm 8.31 (s, 2 x CH, aryl), 8.08, 8.0 (2 x s, 2 x CH, aryl), 7.90, 7.84 (2 x s, 2 x CH, aryl), 6.92-6.80 (br s, 1H, C(O)NH), 5.89 (d, 1H, $J_{1,2}$= 3.6Hz, H-1$_\alpha$), 5.57-5.09 (m, 6H, 2 x H-4, 2 x H-2, 2 x H-3), 5.19 (d, 1H, $J_{1,2}$= 8Hz, H-1$_\beta$), 4.31-4.23 (m, 2 x 2H, NH*CH$_2$*C≡CH), 4.22-4.02 (m, 6H, 2 x H-5, 2 x H-6a, 2 x H-6b), 2.32 (s, 1H, C≡*CH*), 2.20-1.73 (m, 12H, acetyl). $^{13}$C NMR (100MHz, $CDCl_3$): δ, ppm 170.94-169.46, 164.19 (C(O)NH, *C*(O)CH$_3$ ), 157.30, 157.00 (C aryl), 149.29, 149.15 (C aryl), 136.81, 136.61 (C aryl), 122.48, 122.26 (CH aryl), 116.36, 116.28, 114.83 (CH aryl), 99.19 (C-1$_\beta$), 95.69 (C-1$_\alpha$), 78.97 (-*C*≡CH), 73.89 (-C≡*CH*), 72.09 (2 x C-5), 70.10 (2 x C-3), 68.35, 68.05 (2 x C-4), 67.48, 67.32 (2 x C-2), 62.07, 61.70 (2 x C-6), 20.88-20.65 (CH$_3$ acetyl). MS (ESI) m/z calcd for $C_{24}H_{26}N_2O_{13}$ $(M+H)^+$ 551.14, found 551.05; HRMS (Q-TOF) m/z calcd for $C_{24}H_{26}N_2O_{13}$ $(M+H)^+$ 551.1435, found 551.1519.

## Compound 5



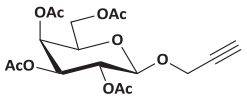
Amide **4** (120 mg, 0.22 mmol) was treated with NaOMe in MeOH (0. 5 M, 5 mL). After deacylation, the reaction was quenched by adding acetic acid to ca. pH 6. The solvent was taken off under reduced pressure and the residue was purified by silica gel chromatography, which gave the MNPG derivative **5** (74 mg, 0.19 mmol, 88%) as a yellowish oil. MS (ESI) m/z calcd for $C_{16}H_{18}N_2O_9$ $(M+H)^+$ 383.10, found 383.20.

## Compound 1



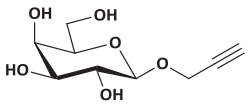
The α and β MNPG mixture **5** (74 mg, 0.19 mmol) was treated with β-Galactosidase (1.8 mg, 1KU, *E.coli*, Grade VIII, lyophilized powder, ≥500 units/mg protein, Sigma-Aldrich) in $NH_4HCO_3$ buffer (10 mL, pH=7.3). The mixture was stirred at r.t. for 2 days. Upon the completion of enzymatic hydrolysis, the solvent was taken off under reduced pressure. The residue was subjected to preparative-HPLC for purification, which afforded the sole α-anomer **1** (37 mg). However, it was eventually found to be mostly **dimerized** according to NMR spectra. $^1$H NMR (400MHz, $D_2O$): δ, ppm 8.31 (s, 1H, aryl), 8.20, 7.90 (2 x s, 2H, aryl), 5.85 (d, 1H, $J_{1,2}$= 3.6Hz, H-1$_\alpha$), 4.19 (s, 2H, -NH*CH$_2$*), 4.15-4.00 (m, 3H, H-3, H-2, H-5), 3.79-3.67 (m, 3H, H-4, H-6a, H-6b), 2.68 (s, 1H, C≡*CH* **then gone**). $^{13}$C NMR (100MHz, $D_2O$): δ, ppm 168.26 (-C(O)NH), 157.44, 149.46, 136.44 (C aryl), 122.80, 117.02, 115.91 (CH aryl), 98.22 (C-1$_\alpha$), 79.47 (-*C*≡CH), 72.99 (C-5), 69.99 (C-3), 69.74 (C-2), 68.53 (C-4), 61.68 (C-6), 30.09 (-$NHCH_2$). HRMS (Q-TOF) m/z calcd for $C_{32}H_{34}N_4O_{18}$ $(M+2H)^{2+}$ 382.3142, found 383.1118.

## Compound 7



1,2,3,4,6-penta-o-acetyl-β-D-galactopyranoside (3g, 7.69mmol) was dissolved in anhydrous $CH_2Cl_2$ (20 mL), which was followed by the addition of iodine (1.17 g, 4.61 mmol) and hexamethyldisilane (0.68 g, 4.65 mmol). The mixture was stirred at r.t. overnight. Upon the complete disappearance of the starting material, the reaction was diluted with $CH_2Cl_2$ (20 mL) and quenched by the addition of 10% $NaS_2O_3$ (20 mL). The organic phase was separated and washed with $NaHCO_3$ (sat., 20 mL). The collected organic layer was dried over $Na_2SO_4$, filtered and concentrated *in vacuo*. The obtained white foam was used directly without further purification. The crude iodo-compound in anhydrous $CH_2Cl_2$ (20 mL) was then treated with propargyl alcohol (0.75 g, 13.43 mmol) and $Ag_2CO_3$ (3.1 g, 11.24 mmol) and stirred at r.t. overnight. The mixture was filtered over celite, the filtrate was concentrated under reduced pressure and the residue was purified by silica gel chromatography, which gave **7** (1.16 g, 62.5% over two steps) as a colorless oil. Spectroscopic data of $^1H$ and $^{13}C$ NMR were consistent with ref [18]. MS (ESI) m/z calcd for $C_{17}H_{22}O_{10}$ $(M+Na)^+$ 409.12, found 408.85. $^1H$ NMR (400MHz, $CDCl_3$): δ, ppm 5.4 (dd, 1H, H-4), 5.22 (dd, 1H, H-2), 5.06 (dd, 1H, H-3), 4.73 (d, 1H, $J_{1,2}$ = 8.0 Hz, H-$1_\beta$), 4.38 (d, 2H, $OCH_2$), 4.21-4.08 (m, 2H, H-6), 3.98-3.89 (m, 1H, H-5), 2.49-2.44 (t, 1H, -C≡CH), 2.18-1.95 (12H acetyl).

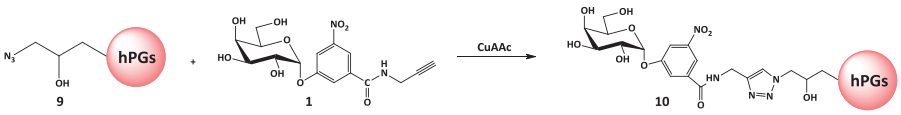
## Compound 8




The propargyl galactoside **7** (200 mg, 0.52 mmol) was treated with NaOMe in MeOH (0.5 M, 5 mL) to remove the acetyl protecting groups. The mixture was stirred at r.t. overnight. Upon completion, the reaction was quenched by adding $H^+$-resin to a pH of ca.6 when the resin was removed by filtration. The solvents were removed under reduced pressure. The residue was purified by silica gel chromatography, which afforded the desired product **8** (102 mg, 91%) MS (ESI) m/z calcd for $C_9H_{14}O_6$ $(M+NH_4)^+$ 236.08, found 236.35. $^1H$ NMR (400MHz, $D_2O$): δ, ppm 4.60 (d, 1H, $J_{1,2}$= 8Hz, H-1), 4.56-4.44 (m, 2H, $OCH_2$), 3.96 (d, 1H, H-4), 3.83-3.65 (m, 4H, H-5, H-3, H-6a, H-6b), 3.55 (dd, 1H, H-2), 2.94 (s, 1H, C≡CH). $^{13}C$ NMR (100MHz, $D_2O$): δ, ppm 101.75

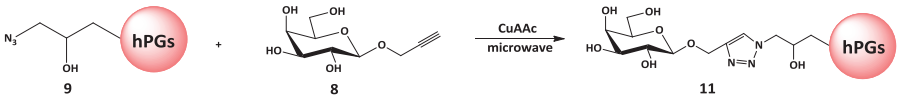
(C-1), 79.10 (-*C*≡CH), 76.85 (C≡*CH*), 75.93 (C-5), 73.39 (C-3), 71.19 (C-2), 69.24 (C-4), 61.59 (C-6), 57.12 ($OCH_2$).

## Compound 10




A solution of azido-polymer **6** (5 mg, 0.02 mmol $N_3$-groups) with compound **1** (11 mg, 0.029 mmol) in THF/$H_2O$ (2:1, 1 mL) was prepared. $CuSO_4{\cdot}5H_2O$ (0.5 mg, 0.002 mmol), NaAsc (7.7 mg, 0.039 mmol) was dissolved in $H_2O$ (0.25 mL) separately in two vessels. Then, the aqueous solution of NaAsc was added to the $CuSO_4{\cdot}5H_2O$ solution. The resultant aqueous solution was added into the reaction mixture in a drop-wise fashion. The reaction mixture was thoroughly degassed using argon. The mixture was stirred at r.t. overnight. Afterwards, the organic solvent was evaporated, and the resulting aqueous mixture was dialyzed against water for 2-3 days to afford the pure product, which afforded compound **10** (<1mg, <10%). (Nearly no reaction under microwave condition as well) The resultant products were characterized by $^1$HNMR, IR and DLS.

## Compound 11




A solution of azido-polymer **6** (3 mg, 0.012mmol $N_3$) with compound **8** (4 mg, 0.018 mmol), in MeOH/$H_2O$ (4:1, 1 mL) was prepared. $CuSO_4{\cdot}5H_2O$ (0.3 mg, 1.2 µmol) NaAsc (0.5 mg, 2.5 µmol) was dissolved in $H_2O$ (minimal amount) separately in two vessels. The resultant aqueous solution was added into the reaction mixture in a drop-wise fashion. The reaction mixture was thoroughly degassed using argon. Then, the reaction was performed using a microwave-assisted CuAAC reaction for 20 min then the copper salts were removed by a resin (Cuprisorb). The organic solvent were evaporated, and the resulting aqueous mixture was dialyzed against water for 2-3 days to afford the pure product, which afforded compound **11** (4 mg, 24.7%). The resultant products were characterized by $^1$HNMR, IR and DLS.

## Reference


[1] a) C. R. Bertozzi, L. L. Kiessling, *Science*. **2001**, 291, 2357-2364. b) A. Varki, *Glycobiology*, **1993**, 3, 97-130. c) A. Varki, R. Cummings, J. Esko, H. Freeze, G. W. Hart, J. Marth, *Essentials of glycobiology*, Cold Spring Harbor Laboratory Press, New York, **1999**.

[2] a) R. T. Lee, Y. Ichikawa, T. Kawasaki, K. Drickamer, Y. C. Lee, *Arch. Biochem. Biophys.* **1992**, 299, 129-136; b) Y. C. Lee, R. T. Lee, *Acc. Chem. Res.* **1995**, 28, 321-327; c) R. T. Lee, Y. C. Lee, *Glycoconjugate J.* **2000**, 17, 543-551; d) J. J. Lundquist, E. J. Toone, *Chem. Rev*. **2002**, 102, 555-578; e) R. J. Pieters, *Med. Res. Rev*. **2007**, 27(6), 796-816.

[3] K. J. Ryan, C. G. Ray (**2004**). *Sherris Medical Microbiology* (4th ed.). McGraw Hill. p. 375. ISBN 0-8385-8529-9.

[4] S. Ahn-Yoon, T. R. DeCory, A. J. Baeumner, R. A. Durst, *Anal. Chem*. **2003**, 75, 2256-2261

[5] K. Ozaki, R. T. Lee, Y. C. Lee and T. Kawasaki, *Glycoconjugate J*. **1995**, 12, 268; J. G. Zhang, O. B. Krajden, R. K. Kainthan, J. N. Kizhakkedathu, I. Constantinescu, D. E. Brooks, M. I. C. Gyongyossy-Issa, *Bioconjugate Chem*. **2008**, 19, 1241.

[6] M. J. S. De Wolf, E. Dams, W. S. H. Dierick, *Biochim. Biophys. Acta* **1994**, 1223, 296.

[7] a) T. K. Sixma, S. E. Pronk, K. H. Kalk, B. A. M. van Zanten, A. M. Berghius, W. G. J. Hol, *Nature*, **1992**, 355, 561. b) W. E. Minke, F. Hong, C. L. M. J. Verlinde, W. G. J. Hol, E. Fan, *J. Biol. Chem*., **1999**, 274 (47), 33469.

[8] a) W. E. Minke, C. Roach, W. G. J. Hol, C. L. M. J. Verlinde, *Biochemistry*, **1999**, 38 (18), 5684-5692; b) E. Fan, E. A. Merritt, Z. Zhang, J. C. Pickens, C. Roach, M. Ahn, W. G. Hol, *Acta crystallogr*. **2001**, 57, 201-212.

[9] D. D. Mitchell, J. C. Pickens, K. Korotkov, E. Fan, W. G. J. Hol, *Bioorg. Med. Chem*. **2004**, 12, 907-920.

[10] W. B. Turnbull, B. L. Precious, S. W. Homans, *J. Am. Chem. Soc.* **2004**, 126, 1047-1054.

[11] a) I. Papp, J. Dernedde, S. Enders, R. Haag, *Chem. Commun.* **2008**, 5851-5853; b) J. Dernedde, A. Rausch, M. Weinhart, S. Enders, R. Tauber, K. Licha, M. Schirner, U. Zugel, A. von Bonin, R. Haag, *Proc. Natl. Acad. Sci. U.S.A.,* 2010, 107, 19679-19684.

[12] a) E. J. Vandenberg, *J. Polym. Sci. Part A*, 1985, 23, 915-949; b) A. Dworak, W. Walach, B. Trzebicka, *Macromol. Chem. Phys.* **1995**, 196, 1963-1970.

[13] H. M. Branderhorst, R. M. J. Liskamp, G. M. Visser, R. J. Pieters, *Chem. Commun.* **2007**, 5043-5045.

[14] E. A. Merritt, S. Sarfaty, I. K. Feil, W. G. J. Hol, *Structure*, 1997, 5, 1485-1499.

[15] E. A. Merritt, Z. Zhang, J. C. Pickens, M. Ahn, W. G. Hol, E. Fan, *J. Am. Chem. Soc*. **2002**, 124, 8818-8824.

[16] A. L. Sisson, D. Steinhilber, T. Rossow, P. Welker, K. Licha, R. Haag, *Angew. Chem.* **2009**, 121, 7676-7681; *Angew. Chem. Int. Ed*. **2009**, 48, 7540-7545.

[17] P. Siemsen, R. C. Livingston, F. Diederich, *Angew. Chem. Int. Ed.* **2000**, 39, 2632.

[18] H. B. Mereyala, S. R. Gurrala, *Carb. Res.* **1998**, 307, 351-354.

[19] I. Papp, C. Sieden, A. L. Sisson, J. Kostka, C. Bottcher, K. Ludwig, A. Herrmann, R. Haag, *ChemBioChem.* **2011**, 12, 887-895.

# Chapter 5

## Multivalent C-3″-Substituted Globotriosides as Inhibitors of *Streptococcus suis* Adhesins

## Introduction

*Streptococcus suis* (*S. suis*) is a major and prevalent porcine pathogen and an important zoonotic agent i.e. also capable of infecting humans.[1] *S. suis* infection could be transmitted to humans by occasional exposure to contaminated pigs or pig meat.[2] In humans, *S. suis* could cause systemic infection e.g. meningitis, septic shock with possible residual deafness.[3] Up to now, there are in total 35 serotypes recognized according to their capsular polysaccharide antigens, of which *S. suis* serotype 2 (SS2) is considered the most virulent in swine and humans.[4] In the past few years, the number of infected humans has been reported to increase. Recent outbreaks of severe human infections by *S. suis* were reported in China in 2005 involving 215 cases and 38 deaths.[5] This pathogen associated with bacterial meningitis remains a big threat to public health in other Asian areas such as Hong Kong, Thailand, Vietnam, *etc.* where intensive pig farming is still common.[6] Capsular polysaccharide, extracellular protein factor (EF), muramidase-released protein (MRP), suilysin, several adhesins, *etc.* have been identified as potential virulence factors of *S. suis* serotype 2.[7] Especially, the capsular polysaccharide was shown as a key factor for bacteria escaping from phagocytosis thereby clearly making an essential contribution to virulence in animal models of infection.[8] Nevertheless, owing to the variety of virulence factors of serotype variants, a complete understanding of how *S. suis* invades the host and finally crosses the blood brain barrier in human has not been achieved.[9] More research is needed to elucidate the virulence of *S. suis* and fully understand the mechanism of pathogenicity in both pigs and humans.

Vaccines and antimicrobial agents are used to control and attenuate *S. suis* infection. But since the pathogenic mechanism to the development of disease is not well-known, the protection of vaccines is incomplete and only partially efficient. An effective vaccine is still under development.[10] Most importantly, antibiotic resistance of *S. suis* is now a frequent occurrence.[11] The massive use of antibiotics has contributed to the emergence. As a result, *S. suis* is now receiving growing attention, not only for its virulence but also for its antibiotics resistance evolution. To circumvent these severe issues, more effective therapeutic agents are urgently needed. As noted, most mammalian cell surfaces are glycosylated e.g. due to the

presence of glycolipids and glycoproteins. Adhesion to these specific glyco-receptor on the cell surface has been shown to be a prerequisite for bacterial infection initiation. Interference of this bacterial adhesion could potentially prevent *S. suis* infectious disease at the very first stage. These featured carbohydrate determinants might offer biochemists appealing targets. Considering that their chemical structure is similar to the innate host ligands, carbohydrate-based anti-adhesives thereby are less likely to contribute to the evolution of bacterial drug resistance and represents a promising approach for the treatment of *S. suis*.[12]

So far, the *S. suis* adhesins involving carbohydrate-mediated adhesion have been characterized by hemagglutination inhibition assays. Some have the preference for sialic acid-containing oligosaccharides in poly-N-acetyllactosaminyl chains.[13] The others e.g. type $P_N$ and type $P_O$ are found to have a binding specificity for the Galα1-4Gal (galabiose) disaccharides unit of the $P_1$ and $P^K$ blood group antigens.[14] Recently, the Streptococcal adhesin P (SadP) was reported that could bind to galabiose residue of host oligosccharides. It contains an N-terminal domain for sugar binding and a C-terminal LPXTG-motif for anchoring.[15] As a crucial receptor for the Streptococcal adhesins, an extended look has been taken at the galabiose unit. There are two types of galabiose-binding strains $P_N$ and $P_O$ that have evolved differences in their galabiose recognition. Hydrogen bonding of HO-4', HO-6', HO-2, and HO-3 of galabiose was shown to occur in the binding to the *S. suis* adhesins. However, unlike the narrow combining site of the $P_O$ strain around the terminal galactose of galabiose, substitution at the terminal galactose is allowed and accepted by the $P_N$ strain.[14] This different mode of intermolecular interaction with the *S. suis* adhesin, in return, provides a good opportunity to make structural alterations aiming at potent galabiose-based monovalent inhibitors. In an initial attempt by Nilsson *et al*.,[16] a library of galabiose-derived inhibitors was evaluated aiming at enhanced inhibitory performance. In this systematic study, they found that the introduction of an aromatic group at the C-3' or an alkyl group at the O-2' position of the terminal galactose unit gave a big improvement in the inhibitory power against *S. suis*. Especially, compound **1** carrying a phenylurea substituent at the C-3' position exhibited the most potent inhibition efficacy in the low nanomolar range ($IC_{50}$= 30

nM) with one order of magnitude improvement in comparison to its parent compound **2** ($IC_{50}$= 310 nM). (**Fig.1**)

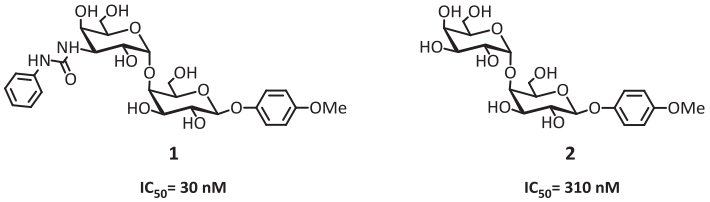



**Figure 1.** Structures of galabiose derivatives.[16]

From the perspective of intrinsic carbohydrate-protein interactions, the relatively low binding affinity of individual carbohydrates is a well-known characteristic, with usually binding in the milli- to micromolar range.[17] Apparently, in biological systems, multiple interactions of carbohydrate residues can simultaneously occur at binding sites to greatly strengthen the avidity towards their recognition partners. This particular type of multivalency effect is the so-called "glycoside cluster effect".[18] Regarding this effect, it has an attractive implication of using a multivalency strategy to develop high affinity ligands as potential therapeutic agents against bacterial infection. With the aid of the multivalency effect, multivalent inhibitors constructed from multiple bioactive compounds e.g. galabiosides linked to a scaffold or core molecule have successfully been introduced against *S. suis* adhesion. Magnusson *et al*. synthesized a series of low valent galabiosides and realized complete inhibition of *S. suis* adhesion in the low nanomolar range.[19] This potential for inhibition was also explored in our group,[20] where a series of di-, tetra-, octavalent galabioside dendrimers were synthesized to address the multivalency effect in *S. suis* inhibition. During a hemagglutination inhibition assay, it was found that the inhibitory power was increased with increasing valency. The best inhibitor was the tetravalent compound **3** (**Fig.2A**) bearing four copies of the galabiose moiety with an MIC of 2.5 nM, which was a total 160 fold increase of binding affinity compared to the monovalent reference compound **4**. However, the related octavalent-version did not provide more inhibitory power than the tetravalent one. The inhibitiory concentration was even slightly increased to 3.9 nM. Furthermore, in an infection peritonitis mouse model, a similar tetravalent galabiose inhibitor 5 (**Fig.2B**) was

tested for its *in vivo* effects.[21] It was shown to have a possible anti-adhesion effect against *S. suis* serotype 2 infection in mice.

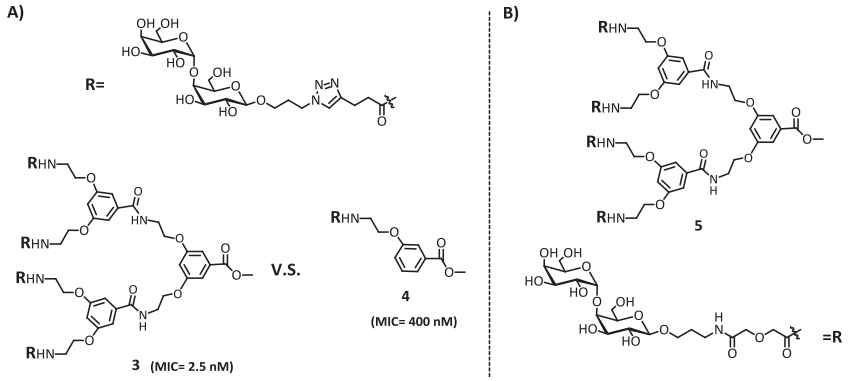



**Figure 2.** Multivalent galabiosides as inhibitor of *S.suis.*[20, 21]

As mentioned above, it is clear that both the advancement in the mono galabioside inhibitor and introduction of multivalency have enhanced the binding affinity in an excellent fashion. Now, an inspiring question has risen that how about employing these two positive impact factors together to bring about an even more effective inhibitor. As a proof of concept, we were intriqued to explore the combination of chemically modified mono-inhibitor with the multivalency effect in this work. The study was based on an earlier determination of the monovalent derivative of the *S. suis* receptor trisaccharide Galα1-4Galβ1-4Glc (globotriose[22]), a better ligand than galabiose even though the glucose moiety is not critical for binding. The suggested globotriose analogue is shown in **Fig.3**, of which the C-3" position is derivatized with a phenylurea group. As part of this research, with respect to our previous finding[20] of the tetravalent scaffold,[23] a synthesis was designed where four azide functionalized globotriose derivatives could be efficiently attached to the scaffold under micro-wave assisted CuAAC conditions (**Fig.3**). The biological activities of the resultant glycoconjugates were profiled towards the relevant *S. suis* adhesin SadP, and is also to be done by hemagglutination inhibiting assay of the whole bacterial strains $P_N$ and $P_O$.

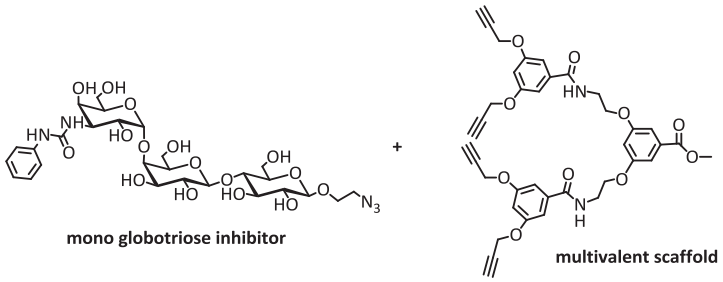
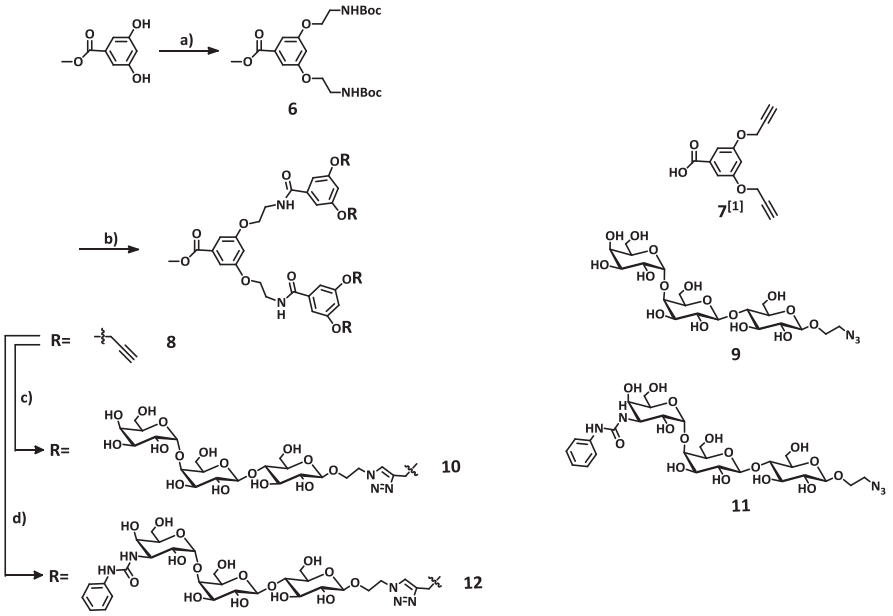



**Figure 3.** mono globotriose inhibitor and multivalent scaffold.

## Results and Discussion

## Synthesis




**Scheme 1.** Reagents and conditions: a) N-Boc-2-bromoethyl-amine, $K_2CO_3$, 18-crown-6, DMF, 40°C, 8 h, 49%; b) i) TFA/$CH_2Cl_2$ (1:1), r.t., 3 h; ii) compound **7**[1], BOP, DIPEA, DMF, r.t., 22%; c) compound **9**, $CuSO_4 \cdot 5H_2O$, Na ascorbate, DMF with 10% $H_2O$, 38%; d) compound **11**, $CuSO_4 \cdot 5H_2O$, Na ascorbate, DMF with 10% $H_2O$, 47%.

The synthesis of the glycosylated dendritic compounds was shown in **Scheme 1**. The commercially available methyl 3,5-dihydroxybenzoate was used as the starting material. By reaction with N-Boc-2-bromoethyl-amine under basic conditions, the di-substituted product **6** was obtained in a moderate yield of 49%. Then, the Boc protecting groups were selectively removed with the treatment of TFA in $CH_2Cl_2$. The

afforded amino-product was then coupled to building block **7**, which was synthesized following a previous report.[23] Coupling of **6** and **7** was carried out in DMF using BOP and DIPEA, which gave the desired alkyne scaffold **8** with a yield of 22% with a sum of incompletely coupled by-products. Subsequently, the globotriose derivative **9**[24] with its azide tail could be efficiently linked to the four alkyne-ends of **8** via CuAAC. Compound **10** was obtained in 38% yield after preparative HPLC purification. The same procedure was also performed for the phenylurea analogue **11**,[24]s which resulted in the glycoconjugate **12** in 47% yield after purification. Compound **10** and compound **12** were both characterized by $^{1}H$, $^{13}C$ NMR and ESI-MS.

## Inhibition study

The examination of inhibition efficiency of the synthetic globotriosides was performed by an ELISA assay using pigeon ovomucoid, displaying galabiose units, on the well surface. An inhibition curve was calculated for each synthesized globotriosides and $IC_{50}$ values were determined (**Fig.4**; **Table 1**). It was shown that the mono reference compound **11** with phenylurea derivatization inhibited the SadP binding with an $IC_{50}$ value in a high nanomolar range ($IC_{50}$= 902 nM). In comparison, a multivalency effect was clearly observed for both tetravalent compounds. The obtained data of **10** and **12** both showed a connection between the valency of globotriose units and resulting inhibition properties. Inhibitor **10** exhibited a 5-fold increase in binding affinity as compared to **11**. However, the improvement of the binding affinity of **10** was limited when corrected for valency. A more effective binding to the *S. suis* adhesin was displayed by **12**, of which the binding affinity was 11 times more potent than the monovalent reference **11**. The reasonable explanation for the difference between tetravalant **10** and **12** was due to the functionality of the substituted phenylurea group at C-3”, which made **12** more efficient than **10**. All these preliminary results demonstrated that both optimized mono ligand and multivalency strategy enable **12** as the most advantageous compound for SadP anti-adhesion. More analysis of bio-tests e.g. with whole bacteria are needed to further investigate the inhibition performance.

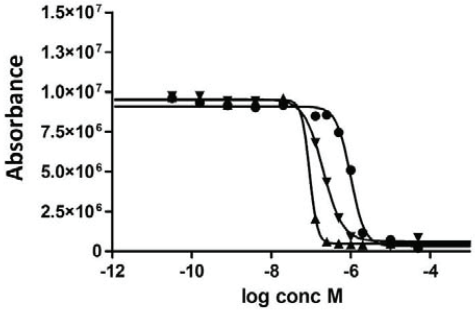



**Figure 4**. Potency of the inhibitors measured by using an ELISA-type assay. Inhibition curves for monovalent **11** (⬮), tetravalent **10** (▼), and tetravalent **12** (▲) galabiose compounds.

**Table 1**. Potency of the inhibitors measured by using an ELISA-type assay

| Inhibitor | S. suis adhesin SadP | |
|---|---|---|
| | $IC_{50}$ [nM] | rel.pot.[a] (per sugar)[b] |
| **11** | 902 | 1 (1) |
| **10** | 204 | 5 (1.2) |
| **12** | 92 | 11 (2.75) |

[a] Relative potency ) $IC_{50}$ (monovalent **11**)/$IC_{50}$ (multivalent compound).

[b] Relative potency per sugar ) relative potency/valency.

## Conclusion

In this study, a novel tetravalent globotrioside as a potential *S. suis adhesin* inhibitor was introduced. Based on the monovalent inhibitor research, four C-3" substituted globotrioside copies were attached to a tetravalent scaffold developed in our prior study via micro-wave assisted CuAAC to give compound **12** in a moderate yield (47%). For comparison, a regular tetravalent globotriosides **10** was also prepared. The synthesized galabiosides were tested for their inhibition of a newly identified *Streptococcus* Adhesin P (SadP) binding to a galabiose-coated surface. Their inhibition concentration was recorded by an ELISA type assay. The mono phenylurea analogue **11** was used as a reference. Both ligand functionalization and ligand valency had a positive effect on the resulting inhibition properties. A big improvement had been achieved for tetravalent **12** with an $IC_{50}$ value of 92 nM representing an 11-fold increase in potency, whereas the improvement was reduced

in the case of **10**. However, more inhibition studies are needed to understand the bacterial adhesin and inhibitor interaction in detail. A hemagglutination inhibition assay of human erythrocytes by *S. suis* bacteria is on the way.

## Experimental procedure

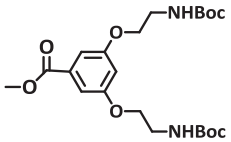
Chemicals were obtained from commercial sources and were used without further purification unless noted otherwise. Compound **7** was synthesized following literature procedures.[23] Compound **9** and compound **11** were provided by Nilsson *et al*. (Lund University, Sweden). Solvents were purchased from Biosolve (Valkenswaard, The Netherlands). All moisture- sensitive reactions were performed under a nitrogen atmosphere. Anhydrous solvents were dried over molecular sieves of 4 Å or 3 Å. TLC was performed on Merck precoated Silica 60 plates. Spots were visualized by UV light and also by 10% $H_2SO_4$ in MeOH, iodine, $KMnO_4$. Microwave reactions were carried out in a Biotage microwave Initiator (Uppsala, Sweden). The microwave power was limited by temperature control once the desired temperature was reached. Sealed vessels of 2-5 mL were used. Analytical HPLC runs were performed on a Shimadzu automated HPLC system with a reversed-phase column (Reprospher 100, C18, 5 µm, 250 X 4.6 mm, Dr.Maisch GmbH, Germany) that was equipped with an evaporative light scattering detector (PLELS 1000, Polymer Laboratories, Amherst, MA, USA) and a UV/Vis detector operating at 220 nm and 250 nm. Preparative HPLC runs were performed on an Applied Biosystems workstation. Elution was effected by using a linear gradient of 5% MeCN/0.1% TFA in $H_2O$ to 5% $H_2O$/0.1% TFA in MeCN. $^1H$ and $^{13}C$ NMR spectroscopy was carried on an Agilent 400-MR spectrometer operating at 400 MHz for $^1H$ and 100 MHz for $^{13}C$. HSQC and TOCSY NMR (500 MHz) were performed with a VARIAN INOVA-500. Electrospray Mass experiments were performed in a Shimadzu LCMS QP-8000. MALDI-TOF-MS were recorded on a Shimadzu Axima-CFR with a-cyano-4-hydroxycinnamic acid or sinapic acid as a matrix. Insulin and adrenocorticotropin fragment 18-39 (Acth) were used for calibration. The ELISA was performed at Turku University by Dr. S. Haataja.

## General procedure of micro-wave assisted CuAAC “click reaction”

To a solution of alkyne compound (1 equiv.) and azide compound (1.2 equiv.) in DMF with 10% $H_2O$, was added $CuSO_4 \cdot 5H_2O$ (1.2 equiv) and NaAsc (2.4 equiv.). The

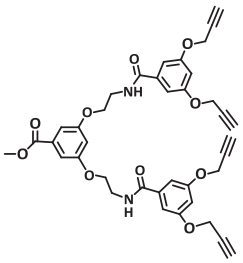
mixture was then irradiated with microwaves and heated to 80°C. The reaction was stirred at these conditions for 30 min. When it reached r.t., the copper salts were removed by a resin (Cuprisorb) and filtrated off. Then the solvents were removed under reduced pressure. The residue was subjected to preparative-HPLC to afford the corresponding 1,2,3-triazole product.

## Compound 6



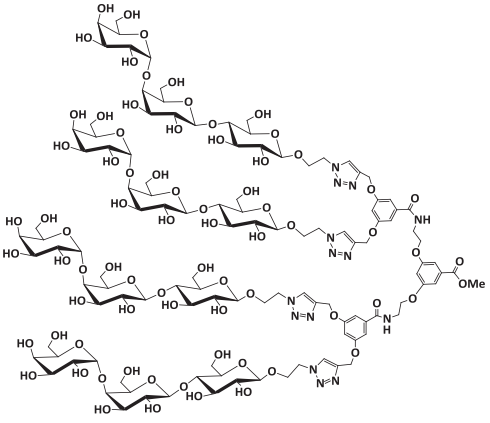

Methyl 3,5-dihydroxybenzoate (4.80 g, 28.57 mmol) and N-Boc-2-bromoethyl-amine (14.37 g, 64.12 mmol) were dissolved in dry DMF (100 mL) which was followed by the addition of 18-crown-6 (0.50 g, 1.89 mmol) and $K_2CO_3$ (17.40 g, 0.126 mol). The mixture was stirred at 40°C for 48 h. Upon completion of the reaction, it was stopped and filtered through Celite. The filtrate was then concentrated under reduced pressure. The residue was extracted with EtOAc (50 mL) and washed with $H_2O$ (25 mL) two times and brine (25 mL) two times. The combined organic layers were dried over $NaSO_4$, filtered and concentrated. The residue was subject to silica gel chromatography and then precipitated in EtOAc/Hex to afford compound **6** as white solid (6.40 g, 14.09 mmol, 49%). $^1$H NMR (400 MHz, $CDCl_3$): δ, ppm 7.16 (d, 2H, J= 2.4 Hz, CH, aryl), 6.62 (dd, 1H, J= 2.4 Hz, CH, aryl), 4.99 (br s, 2H, 2 × NH), 4.02 (t, 4H, J= 5.1Hz, 2 × $OCH_2$), 3.89 (s, 3H, $OCH_3$), 3.52(m, 4H, 2 × C*H*$_2$NH), 1.44 (s, 18H, 2 × $C(CH_3)_3$). $^{13}$C NMR was consistent with ref.[25] MS (ESI) m/z calcd for $C_{22}H_{34}N_2O_8$ $(M+Na)^+$ 477.23, found 477.25.

## Compound 8



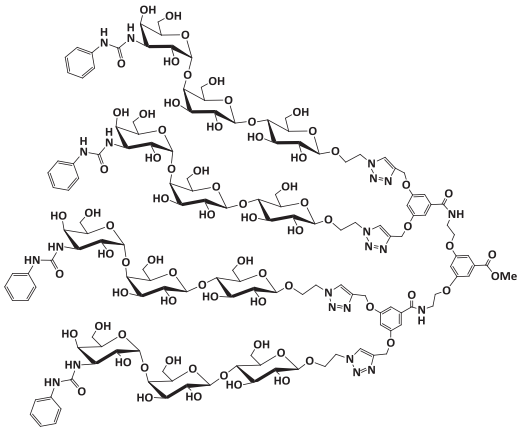
Compound **6** (1.10 g, 2.42 mmol) was treated with TFA/$CH_2Cl_2$ (30 mL, 1:1). The mixture was stirred at r.t. for 3 h. Upon the disappearance of starting material by TLC, the reaction was stopped and the solvents were removed *in vacuo*. The residue was used for the next step without further purification. To a solution of the deprotected compound (0.56 g, 2.20 mmol) and compound **7**[1] (1.27 g, 5.52 mmol) in dried DMF (20 mL), BOP (2.43 g, 5.50 mmol) and DIPEA (1.85 g, 14.31 mmol) were added. The mixture was stirred at r.t. overnight. After that, the reaction was stopped and concentrated under reduced pressure. The residue was redissolved into EtOAc (50 mL) and washed with 1N $KHSO_4$ (20 mL), aqueous $NaHCO_3$ (20 mL), and $H_2O$ (20 mL), respectively. The combined organic layer was dried over $NaSO_4$, filtered and concentrated. The residue was then subjected to flash chromatography, which gave the title compound **8** (0.33 g, 0.49 mmol, 22%). $^1$H NMR (400 MHz, $CDCl_3$): δ, ppm 7.16 (d, 2H, J= 2.4 Hz, CH, aryl), 7.02 (d, 4H, J= 2.4 Hz, CH, aryl), 6.74 (dd, 2H, J= 2.4Hz, CH, aryl), 6.70 (t, 2H, J= 5.6 Hz, N*H*C(O)), 6.64 (dd, 1H, J= 2.4Hz, CH, aryl), 4.69 (d, 8H, J= 2.4 Hz, 4 × $OCH_2$C), 4.13 (t, 4H, J= 4.8 Hz, 2 × $OCH_2$), 3.89 (s, 3H, $OCH_3$), 3.84 (m, 4H, 2 × C*H*$_2$NH), 2.54 (t, 4H, J= 2.4 Hz, 4 × C≡CH). $^{13}$C NMR (100 MHz, $CDCl_3$): δ, ppm 167.45, 166.59 (C(O)NH); 159.58, 158.89 (C, aryl); 136.58, 132.37 (C, aryl); 108.43, 106.88, 106.55, 105.68 (*C*H, aryl), 78.07 (*C*≡CH); 76.24 (C≡*C*H); 67.12 ($OCH_2$); 56.28 (O*C*H$_2$C); 52.49 ($OCH_3$); 39.68 ($CH_2NH$). MS (ESI) m/z calcd for $C_{38}H_{34}N_2O_{10}$ $(M+H)^+$ 679.22, found 679.10.

## Compound 10



A solution of compound **8** (3 mg, 4.42 µmol), compound **9** (12 mg, 0.021 mmol), NaAsc (8.2 mg, 0.041 mmol) and $CuSO_4 \cdot 5H_2O$ (5.31 mg, 0.021 mmol) in DMF/$H_2O$ (2 mL, 9:1) was prepared. The mixture was treated following the general procedure of click reaction as described above to afford compound **10** (5 mg, 1.68 µmol, 38%) as white powder. $^1$H NMR (500 MHz, $D_2O$): δ, ppm 8.06 (s, 4H, 4 × $CH_{triazole}$), 6.87 (s, 2H, 2 × CH, aryl), 6.79 (s, 4H, 4 × CH, aryl), 6.54 (s, 2H, 2 × CH, aryl), 6.51 (s, 1H, CH, aryl), 5.00-4.90 (m, 12H, 4 × H-1", 4 × $OCH_2C_{triazole}$), 4.60 (br s, 8H, 4 × $N_{triazole}CH_2$), 4.41 (d, 8H, J= 8 Hz, 4 × H-1, 4 × H-1'), 4.35 (t, 4H, J= 6.5 Hz, 4 × H-3'), 4.26-4.19 (m, 4H, 2 × $OCH_2$), 4.08-3.97(m, 16H, 4 × $OCH_2CH_2N_{triazole}$), 3.95-3.80 (m, 26H), 3.79-3.76 (m, 30H), 3.76-3.66 (m, 25H, $OCH_3$), 3.66-3.53 (m, 14H, 2 × $CH_2NH$), 3.53-3.45 (m, 10H), 3.25 (t, 4H, J= 9 Hz, 4 × H-2'). $^{13}$C NMR (125 MHz, $D_2O$): 160.31, 159.74 (C, aryl); 143.58 ($C_{triazole}$), 134.07 (C, aryl); 126.94, 126.88 and 125.92 (CH, triazole); 109.29, 107.62, 107.60 and 106.12 (CH, aryl); 104.41, 103.32 (C-1, C-1'); 101.38 (C-1''); 79.85; 78.28; 76.31; 75.65; 75.19; 73.61 (C-2'), 73.09; 71.76, 71.72 (C-3'); 69.83, 68.88 ($OCH_2$, $OCH_2CH_2N_{triazole}$); 67.40; 61.94 ($OCH_2C_{triazole}$); 61.32; 61.06; 53.46 ($OCH_3$); 52.10, 51.14 ($N_{triazole}CH_2$); 40.32 ($CH_2NH$). HRMS (MALDI-TOF) m/z calcd for $C_{118}H_{174}N_{14}O_{74}$ $(M+Na)^+$ 2994.028, found 2994.722.

## Compound 12




A solution of compound **8** (2.0 mg, 3.0 µmol), compound **11** (9.8 mg, 0.021 mmol), NaAsc (5.6 mg, 0.028 mmol) and $CuSO_4 \cdot 5H_2O$ (3.5 mg, 0.014 mmol) in DMF/$H_2O$ (2 mL, 9:1) was prepared. The mixture was treated following the general procedure of the click reaction as described above to afford compound **12** (4.8 mg, 1.40 µmol,

47%) as white powder. $^{1}$H NMR (500 MHz, $D_2O$ with 30% $CD_3CN$): δ, ppm 8.14 (s, 4H, 4 × $CH_{triazole}$), 7.40-7.28 (m, 16H, 16 × CH, PhNH), 7.15 (s, 2H, 2 × CH, aryl), 7.06 (dd, 4H, J= 7 Hz, 4 × CH, PhNH), 7.01 (s, 4H, 4 × CH, aryl), 6.82 (s, 3H, 3 × CH, aryl), 5.18 (s, 8H, 4 × $OCH_2C_{triazole}$), 4.96 (s, 4H, 4 × H-1"), 4.63 (s, 8H, 4 × $N_{triazole}CH_2$), 4.44 (d, 4H, J= 8 Hz, 4 × H-1), 4.41-4.33 (m, 8H, 4 × H-1', 4 × H-3'), 4.27-4.15 (m, 10H, 2 × $OCH_2$), 4.12-3.95(m, 20H, 4 × O*C*$H_2$$CH_2N_{triazole}$), 3.93-3.77 (m, 25H, $OCH_3$), 3.76-3.71 (m, 36H, 2 × *C*$H_2$NH), 3.70-3.59 (m, 22H), 3.59-3.43 (m, 16H), 3.24 (t, 4H, J= 8 Hz, 4 × H-2'). $^{13}$C NMR (125 MHz, $D_2O$ with 30% $CD_3CN$): 161.68, 160.99, and 159.09 (C, aryl); 144.79 ($C_{triazole}$); 140.67 (C, PhNH); 138.05 (C, aryl); 131.13 (CH, PhNH); 127.69, 126.90 (CH, triazole); 125.18 (CH, PhNH); 120.94 (CH, PhNH); 110.38, 108.74 and 107.41 (CH, aryl); 105.31, 104.21 (C-1, C-1'); 101.92 (C-1''); 81.11; 79.34; 77.26; 76.70; 76.34; 74.50 (C-2'); 73.07, 72.92 (C-3'); 69.68, 69.63 ($OCH_2$, O*C*$H_2CH_2N_{triazole}$); 69.16; 62.99 (O*C*$H_2C_{triazole}$); 62.18; 61.97; 54.22 ($OCH_3$); 52.29, 52.02 ($N_{triazole}CH_2$); 41.07 ($CH_2NH$). HRMS (MALDI-TOF) m/z calcd for $C_{146}H_{198}N_{22}O_{74}$ $(M+Na)^+$ 3467.244, found 3467.943.

## Reference


[1] Y. Tang, H. Zhao, W. Wu, D. Wu, X. Li, W. Fang, *Folia Microbiol. (Praha)* **2011**, 56, 541-548.

[2] M. Ip, K. S. Fung, F. Chi, E. S. Cheuk, S. S. Chau, B. W. Wong, S. L. Lui, M. Hui, R. W. Lai, P. K. Chan, *Diagn. Microbiol. Infect. Dis*. 2007, 57, 15-20.

[3] a) H. F. Wertheim, H. D. Nghia, W. Taylor, C. Schultsz, *Clin. Infect. Dis.* **2009**, 48, 617-625; b) C. Palmieri, P. E. Varaldo, B. Facinelli, *Front. Microbiol*., **2011**, 2, 235.

[4] a) J. J. Staats, I. Feder, O. Okwumabua, M. M. Chengappa, *Vet. Res. Commun*. **1997**, 21, 381-407; b) Z. R. Lun, Q. P. Wang, X. G. Chen, A. X. Li, X. Q. Zhu, *Lancet Infect. Dis.* **2007**, 7, 201-209.

[5] C. Chen, J. Tang, W. Dong, C. Wang, Y. Feng, J. Wang, F. Zheng, X. Pan, D. Liu, M. Li, Y. Song, X. Zhu, H. Sun, T. Feng, Z. Guo, A. Ju, J. Ge, Y. Dong, W. Sun, Y. Jiang, J. Wang, J. Yan, H. Yang, X. Wang, G.-F. Gao, R. Yang, J. Wang, J. Yu, *PLoS One*, **2007**, 2, e315

[6] a) A. Kerdsin, S. Dejsirilert, P. Puangpatra, S. Sripakdee, K. Chumla, N. Boonkerd, P. Polwichai, S. Tanimura, D. Takeuchi, T. Nakayama, S. Nakamura, Y. Akeda, M. Gottschalk, P. Sawanpanyalert, K. Oishi, *Emerg. Infect. Dis.*, **2011**, 17, 835-842; b) N. T. Mai, N. T. Hoa, T. V. Nga, L. D. Linh, T. T. Chau, D. X. Sinh, N. H. Phu, L. V. Chuong, T. S. Diep, J. Campbell, H. D. Nghia, T. N. Minh, N. V. Chau, M. D. de Jong, N. T. Chinh, T. T. Hien, J. Farrar, C. Schultsz, *Clin. Infect. Dis*. **2008**, 46, 659-667.

[7] J. Brassard, M. Gottschalk, S. Quessy, *Vet. Microbiol* **2004**, 102, 87-94; b) S. J. King, P. J. Heath, I. Luque, C. Tarradas, C. G. Dowson, A. M. Whatmore, *Infect. Immun*. **2001**, 69, 7572-7582; c) Y. Li, G. Martinez, M. Gottschalk, S. Lacouture, P. Willson, J. D. Dubreuil, M. Jacques, J. Harel, *Infect. Immun*. **2006**, 74, 305-312; d) H. E. Smith, H.

J. Wisselink, N. Stockhofe-Zurwieden, U. Vecht, M. M. Smits, *Adv. Exp. Med. Biol*. **1997**, 418, 651–655.
[8] a) D. Takamatsu, M. Osaki, P. Tharavichitkul, S. Takai, T. Sekizaki, *J Med Microbiol* **2008**, 57, 488-94; b) W. Li, Y. Wan, Z. Tao, H. Chen, R. Zhou, *Vet. Microbiol.* **2013**, 162, 186-194.
[9] M. Gottschalk, M. Segura, J. Xu, *Anim Health Res Rev*. **2007**, 8, 29-45.
[10] R. Higgins & M. Gottschalk, (**1999**). Streptococcal diseases. In Diseases of Swine, 8th edn (Straw, B. E., D'Allaire, S., Mengeling, W. L. & Taylor, D. J., Eds), 563-578. Iowa State University Press, Ames, IA, USA.
[11] M. T. Holden, H. Hauser, M. Sanders, T. H. Ngo, I. Cherevach, A. Cronin, I. Goodhead, K. Mungall, M. A. Quail, C. Price, E. Rabbinowitsch, S. Sharp, N. J. Croucher, T. B. Chieu, N. T. Mai, T. S. Diep, N. T. Chinh, M. Kehoe, J. A. Leigh, P. N. Ward, C. G. Dowson, A. M. Whatmore, N. Chanter, P. Iversen, M. Gottschalk, J. D. Slater, H. E. Smith, B. G. Spratt, J. Xu, C. Ye, S. Bentley, B. G. Barrell, C. Schultsz, D. J. Maskell, J. Parkhill, *PLoS One* **2009**, 4, e6072.
[12] L. Sihvonen, D. N. Kurl, J. Henrichsen, *Acta Vet. Scand*. **1988**, 29, 9-13.
[13] J. Liukkonen, S. Haataja, K. Tikkanen, S. Kelm, J. Finne, *J. Biol. Chem.* **1992**, 267, 21105-21111.
[14] S. Haataja, K. Tikkanen, U. Nilsson, G. Magnusson, K.-A. Karlsson, J. Finne, *J. Biol. Chem.*, **1994**, 269, 27466-27472.
[15] A. Kouki, S. Haataja, V. Loimaranta, A. T. Pulliainen, U. J. Nilsson, J. Finne, *J. Biol. Chem*. **2011**, 286, 38854-38864.
[16] J. Ohlsson, A. Larsson, S. Haataja, J. Alajääski, P. Stenlund, J. S. Pinkner, S. J. Hultgren, J. Finne, J. Kihlberg, U. J. Nilsson, *Org. Biomol. Chem.*, **2005**, 3, 886-900.
[17] H. Lis, N. Sharon, *Chem. Rev*. **1998**, 98, 637-674.
[18] Y. C. Lee, R. T. Lee, *Acc. Chem. Res.*, **1995**, 28, 321-327.
[19] H. C. Hansen, S. Haataja, J. Finne, G. Magnusson, *J. Am. Chem. Soc*. **1997**, 119, 6974-6979.
[20] J. A. F. Joosten, V. Loimaranta, C. C. M. Appeldoorn, S. Haataja, F. A. El Maate, R. M. J. Liskamp, J. Finne, R. J. Pieters, *J. Med. Chem*. **2004**, 47, 6499-6508; b) H. M. Branderhorst, R. Kooij, A. Salminen, L. H. Jongeneel, C. J. Arnusch, R. M. J. Liskamp, J. Finne, R. J. Pieters, *Org. Biomol. Chem.*, **2008**, 6, 1425-1434.
[21] R. J. Pieters, H. C. Slotved, H. M. Mortensen, L. Arler, J. Finne, S. Haataja, J. A. F. Joosten, H. M. Branderhorst, K. A. Krogfelt, *Biology* **2013**, 2, 702-718.
[22] S. Haataja, Z.-Y. Zhang, K. Tikkanen, G. Magnusson, J. Finne, *Glycoconjugate Journal* **1999**, 16, 67-71.
[23] D. T. S. Rijkers, G. W. van Esse, R. Merkx, A. J. Brouwer, H. J. F. Jacobs, R. J. Pieters, R. M. J. Liskamp, *Chem. Commun.*, **2005**, 4581-4583.
[24] Compound **4** and compound **6** were provided by Nilsson *et al*. (Lund University, Sweden)
[25] A. J. Brouwer, S. J. E. Mulders, R. M. J. Liskamp, *Eur. J. Org. Chem*. **2001**, 1903-1915.
[26] H. Senpuku, K. Matin, S. M. D. Abdus, I. Kurauchi, S. Sakurai, M. Kawashima, T. Murata, H. Miyazaki, N. Hanada, *Scand. J. Immunol*. **2001**, 54, 109-116.

# Chapter 6

## Summary, Perspectives & Remarks

## Summary

**Chapter 1** describes carbohydrates contain a remarkable amount of information and complexity within their structure. They are associated with a wide range of pathological recognition events.[1a] Understanding their roles in disease-causing adhesion at the molecular level is crucial for carbohydrate-based therapeutics development. Nature uses carbohydrates and enhances weak binding interactions. Multivalency is a key principle in these interactions and is followed by researchers to improve the treatment of many infectious diseases caused by bacteria and bacterial toxins. The high prevalence of carbohydrates and emerging anti-microbial resistance emphasizes the importance of gaining a full understanding of the modes of action that are central to disease pathogenesis. Being a key principle in nature, a fundamental understanding of the multivalency effect is important in this context. Recent advances in the synthesis of glycoconjugates e.g. glycoclusters, glycodendrimers, glycoproteins and glycopolymers that show a multivalency effect have been highlighted.[1b,1c] It was shown that multivalent glycoarchitectures could be generated including varying shapes, orientations, valencies and ligand densities. In this thesis, synthetic glycomimetics that exhibit multivalency effects have been designed against *P. aeruginosa* LecA, Cholera toxin and *S. suis*. Such efforts and promising results might lead to the next generation of multivalent glycoconjugates with tailored biological activities for therapeutic applications.

We described four modified rigid divalent analogues of an inhibitor of LecA, a virulence factor of *Pseudomonas aeruginosa* in **Chapter 2**. This work was stimulated by our previous study.[2] With the structural knowledge of LecA, the distance between two binding sites is 26 Å (as measured between the anomeric oxygens of the two bound galactoside ligands in an X-ray structure[2c]). A galactosylated divalent LecA inhibitor with a rigid glucose-triazole based spacer spanning that distance was introduced with significant inhibitory benefit. Next, focusing on structural optimization, two aspartic acid residues were spotted on the protein surface next to binding sites. Taking these two efforts together, the aim of this work was to rationally design derivatives of the existing divalent LecA inhibitor that might take advantage of potential beneficial secondary interactions between spacer and protein.

Using the structure of the lead compound[2b] as a starting point, a new series of synthetic analogues was prepared and customized with various elements such as a carboxylate anion, an ammonium cation, an aryl group and a positively charged hydrophobic pyridinium moiety. The selective deprotection of two TBDMS groups on two glucose units of the spacer facilitated the modification of the resulting primary hydroxyl groups at the C-6 positions. Four specifically functionalized inhibitors were obtained through TEMPO oxidation, azide substitution then reduction and cycloaddition, triflation along with pyridine addition, respectively. In a structure-activity relationship determination of the modified compounds against LecA, the binding properties were obtained by an ELISA type assay and ITC measurements. It was expected that the structural changes of the spacer would afford additional binding benefits. Both measurements revealed that the new compounds did not exhibit a higher potency than the original structure they were based on. Molecular dynamics based on structural data indicated the engagement of especially the newly introduced charged ammonium ions with the protein carboxylates. However it seems that the energy gain was limited and other conformations were preferred, possibly pointing the newly introduced functional groups into the solvent. This behaviour must be taken into account in future work and new designs are in progress.

The exceptional pentameric protein structure of the Cholera toxin (CT), an $AB_5$ toxin, makes it an ideal target for the application of multivalent ligands to interfere with the pathological recognition of GM1 gangliosides expressed on host cell surface. Impressive results have been reported in our prior design of di-, tetra- and octavalent ligands as CT binders.[3] It was speculated that, beyond the multivalency effect, protein aggregation associated with a mismatched valency of inhibitor was a major contributor to the observed large affinity enhancement.[4] Accordingly, there was an interest to verify this hypothesis by synthesizing a well-defined pentavalent version and consequently evaluate their inhibitory potency. Hence, in **Chapter 3**, we have developed a convenient strategy to synthesize a pentavalent GM1-based inhibitor that matched the valency of CT. Five GM1 copies were assembled onto an aromatic dendritic core through five PEGylated linkers of the appropriate “effective

length"[5], where a microwave-assisted CuAAC reaction was employed to enable the conjugation. The inhibitory properties were determined by an ELISA assay. The observed CTB inhibitory potency of penta-GM1 glycodendrimer was in the picomolar range with an $IC_{50}$ value of 260 pM. This was a bit worse than that of the tetravalent version which was also prepared. Nevertheless, these data still validated the pentavalent compound as a promising candidates against CT. To probe the influence of the valency in ligand-protein interactions, the analytical ultracentrifuge (SV-AUC)[6] technique was applied to characterize the possible ligand-induced protein aggregation induced by the various inhibitors. Interestingly, the degree of protein aggregation in case of the matched (penta) valency and the mismatched (tetra) valency were very similar.

Nowadays, a number of multivalent CT inhibitors using inherently weak and simple ligands e.g. galactose and lactose have been developed with increased affinity gains.[5,7] One such mono saccharide, *m*-nitrophenyl-α-D-galactoside (MNPG) with an $IC_{50}$ value of 0.72 mM against CT, is of particular interest.[8] As a following study, in **Chapter 4**, we present a synthetic route towards a high molecular weight MNPG-conjugated nanoparticle as a possible enhanced CT antagonist. To maximize the multivalency effect, biocompatible, dendrimer-like hyperbranched polyglycerols (hPGs)[9] were chosen as the core structure, whose high density peripheral groups on the particle surface could be further functionalized. The alkyne-terminated MNPG derivative was successfully synthesized and prepared in an anomerically pure form via a selective enzymatic hydrolysis at the final stage. The conjugation with the azide-functionalized hPGs was conducted under standard click chemistry conditions. Nevertheless, the preferred MNPG-coated glycopolymers were not obtained. In contrast, galactosylated hPGs as an even simpler ligand system was successfully synthesized with 88 galactose units per polymer under the same conditions with no problem. To address this unexpected issue, this synthetic MNPG molecule was investigated in more detail by NMR and ESI-MS measurements. Surprisingly, it was found that a dimerization of the MNPG-alkyne had occurred instead of a CuAAC conjugation to the polymer surface. This probably took place during the enzymatic cleavage of the β-anomer of the MNPG anomeric mixture. A Glaser-like coupling of

two alkyne MNPG molecules was triggered thereby resulting in the failure of aryl triazole ring formation. To explore multivalent glycodendritic polymers as CT inhibitors, a detailed study was focused on the galactose system. A strong CT inhibition activity of the galactose hPGs was observed with an $IC_{50}$ value of 1.3 μM in an ELISA assay, which was more than 2,100 fold lower than the $IC_{50}$ value of monovalent galactose when corrected for valency. This shed light on the possibility of making polymeric MNPG glycoconjugates which should be even more effective agents for CT inhibition. We are currently working on a newly designed route to MNPG units linked to the hPGs.

**Chapter 5** deals with glycodendritic compounds to interfere with galabiose (Galα1-4Gal)-mediated *Streptococcus suis* (*S. suis*) adhesion. Two well-known subtypes of S. suis designated as $P_N$ and $P_O$ along with the recently characterized *Streptococcus* Adhesin P (SadP) are capable of recognizing galabiose-containing oligosaccharides.[10] Type $P_N$ and $P_O$ strains had been claimed to have a somewhat different galabiose-binding specificity. Based on this difference, a systematic study of *S. suis* with galabiose derivatives was carried out.[11] As a result, a potent monovalent inhibitor was found bearing a phenylurea group on C-3' of the terminal galactose unit. In another independent work of our group, a strong inhibition of *S. suis* was observed in the nanomolar range via a multiple display of galabiose ligands on a dendrimer scaffold.[12] Therefore, combing these two studies, a galabiose-containg trisaccharide, globotriose (Galα1-4Galβ1-4Glu, $GbO_3$) and its phenylurea analogue were used in the multivalency concept aiming to enhance *S. suis* affinity. Both globotriose derivatives were "clicked" to the tetravalent dendritic molecules. The inhibition performances of the resultant compounds were tested against SadP. A monovalent globotrioside with a phenylurea group was used as a monovalent reference compound. The best inhibitor was the phenylurea derivatized tetravalent $GbO_3$ with the lowest inhibitory concentration ($IC_{50}$) of 92 nM, which is eleven-fold more efficient than the monovalent reference compound. Meanwhile, the regular tetravalent version exhibited a small increase in affinity towards SadP only by a factor of five. These preliminary results demonstrated that the mono ligand optimization and multivalent mechanism of action are two key factors for the observed inhibition. From this point

of view, their further investigation is being followed up by hemagglutination inhibition assay of human erythrocytes by *S. suis* bacteria.

## Perspectives

As pointed out in the discussion on multivalent system design, the nature of the spacer represents a key factor for the observed improved binding affinity. On the basis of a molecular modelling study, two aspartic acid residues (Asp 47) found at the LecA surface were targeted to serve as two secondary binding sites in the pursuit for additional binding affinity in **Chapter 2**. A small library of modified divalent derivatives was fine-tuned with a variety of functional groups for specific affinity to LecA via spacer-mediated protein interactions. Unfortunately, this functionalization at both C-6 positions of the terminal glucoses of the spacer did not afford the expected binding enhancement over the parent compound. Even so, a molecular dynamic study of LecA with inhibitors indicated that the newly introduced functional groups should be able to interact favourably with the targeted aspartates. Recently, it was revealed from an X-ray structure[13] of the parent compound bound to the LecA protein that the O-2 hydroxyl group of end-glucose of the rigid spacer already makes a water bridged hydrogen bond with Asp 47, and furthermore, the O-3 hydroxyl group forms a hydrogen bond with Tyr 36 directly (**Fig.1A**).

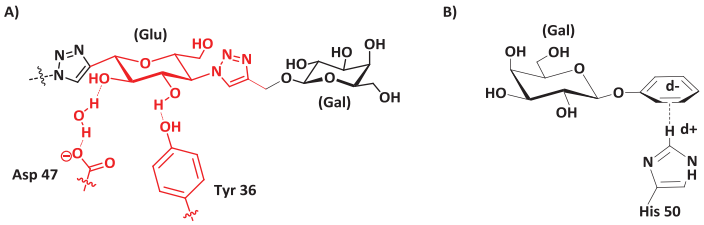



**Figure 1.** Interaction of inhibitor's spacer end unit with LecA.[13, 14]

Possibly the spacer protein interactions already present in the system, makes it less likely to engineer new ones, since the existing ones need to be broken first. To circumvent the lack of binding enhancement arising from the non-participation of the newly introduced functional group, the next attempt for spacer modification should be oriented at the C-2 & C-3. These two positions are more likely to meet the observed LecA binding geometries. Furthermore, it was reported[14] that LecA displayed a preference for ligands with an aglycon containing an aromatic aglycon that can be part of a CH-π T-shape interaction with His 50 (**Fig.1B**). According to that, it seems attractive that, instead of the triazole alycone part of the

galactose unit, a benzene ring group could be introduced. The resulting CH-π T-shape interaction with His 50 thereby might lead to further affinity gain.

Besides that, it was discovered by Varrot *et al.*[15] that there is a secondary LecA sugar binding site as revealed by the crystal study of the LecA-melibiose (Galα1-6Glc) complex. An additional sugar binding site of bound glucose was revealed in proximity to the primary galactose binding site. Even though the affinity of the secondary pocket for glucose is probably weak, this finding could afford an alternative way to enhance the binding affinity by introducing extra glucoses to contact the protein.

With respect to the tetrameric structure of LecA, it is tempting to introduce a tetravalent ligand to match the valency of all of LecA's four binding sites. On the basis of the existing divalent system, a proposed tetravalent pattern is depicted in **Fig.2**. Undoubtedly, this is a big challenge. As indicated by the LecA-galactose crystal structure, the longer distance between one pair of two binding sites is 79 Å.[16] To meet that distance, a long second spacer with the effective length to bridge these two binding sites is needed. However, so far, there is still no complete study in this field. The important factor of rigidity/flexibility has to be taken into account in the creation of the 2nd spacer.

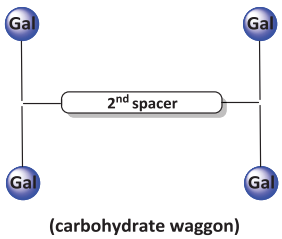



**Figure 2.** Valency-matched LecA inhibitor

As described in both **Chapter 3** and **Chapter 4**, we have described the applications of tailored multivalent carbohydrate ligands as CT inhibitors to probe the influence of valency on a multivalent binding system. Even though the expected 1: 1 manner of complexing valency-matched multivalent ligands with $CTB_5$ was not observed, the significant inhibition effect exhibited by five-fold display of GM1 oligosaccharides makes it still one of the most potent agents for CT inhibition. Instead, it was found that both penta- and a newly synthesized tetra-GM1 ligand behaved similarly leading

to cross-linked complexes. So far, only two studies showed a detectable toxin-ligand complex in a 1:1 ratio: using one inactive CT toxin and another pentacyclen as template structures presenting the five ligands.[17] The cross-linked aggregates in our case underscored that not only the valency of the ligand but the ligand topology are also essential for the high affinity and specific molecular recognition events. Given that, it is worth noting that the five pendent GM1 head groups in our system were not pre-organized in an exact pentagonal pattern as is more so the case for the other two. Their observed toxin-ligand 1:1 complexes might be associated with nearly perfect pre-organization of GM1 ligands prior to binding thereby stabilizing their complexation. A recent study done by Bundle *et al.*[18] also demonstrated and used the idea of pre-organization to develop a $GbO_3$-based multivalent inhibitor against Shiga-like toxin (SLT), which is in the same family of $AB_5$ toxins as cholera and the heat-labile toxin. They had even emphasized that the principle of pre-organization could be a strong factor for binding significance. Taking all of these into consideration, a penta-arylcyclopentadiene (**Fig.3**) draws our attention, which could be easily prepared via a palladium catalysed reaction of aryl bromides with metallocenes.[19] Regarding to its radial topology, it might open up an opportunity to present the GM1 ligands with an increased degree of pre-organization. The ligands are hopefully posed into correct positions to bind their complementary targets. Provided that, the result in turn could help us interpret the multivalent binding mechanism in terms of valency and topology.

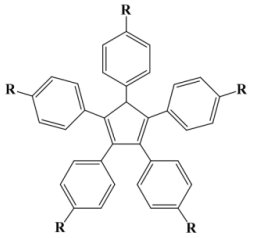



**Figure 3.** penta-aryl cyclopentadiene

Some recent findings have demonstrated the potential of hyperbranched polyglycerols (hPGs) in chemical and biomedical applications.[20] Especially, their peripheral hydroxyl groups enable the attachment of multiple ligands e.g. carbohydrates[21] on their surface. The resulting glycopolymers could perform a

multivalent binding mechanism and lead to an increased binding affinity. From this point of view, a polyvalent display of MNPG epitopes could be a nice approach with enhanced inhibition efficiency of CT. Unfortunately, as stated in **Chapter4**, the self-coupling of MNPG molecules instead of CuAAC coupling with the azide-functionalized hPGs was observed unexpectedly. To circumvent this unanticipated terminal alkyne cross-coupling, an option would be to conduct the enzymatic procedure after the loading of both α and β MNPGs onto the hPGs. Alternatively, a CuAAC reaction could be performed by alkyn hPGs with the corresponding azido-MNPGs. Both methods followed by galactosidase treatment should avoid the occurrence of the problematic 1,3-diynes thereby yield the desired MNPG-hPGs end-products. Apart from the synthetic issue, the inhibitory effect of the resulting glycopolymer depends on several factors e.g. ligand density, degree of functionalization (DF), surface environment, nature of the spacer, and even particle size.[22] It was pointed out that there is a saturation point in the degree of functionalization, above which no additional affinity was achieved.[22] Although a major binding enhancement was observed in the galactoseconjugated polymer, more research is needed. All of them are vital factors underlying multivalent interactions of glycopolymer. This work is still in its infancy.

A study of *S. suis* binding by galabiose (Galα1-4Gal) derivatives has disclosed that the type $P_N$ strains could tolerate large substitution at O-2" and C-3" positions of the terminal galactose unit in oligosaccharides (**Fig.4A**), whereas $P_O$ strains could not.[23] With respect to that, two potential galabiose-based monovalent inhibitors with a phenylurea group at C-3" and methoxymethyl group at O-2" were previously identified and characterized.[11] In **Chapter 5**, the phenylurea derived globotriose (Galα1-4Galβ1-4Glu, $GbO_3$) was used for multivalent display. As an extension, it is proposed to place both the phenylurea group at C-3" and the methoxymethyl group at O-2" of one molecule to afford a novel monovalent compound. Nevertheless, it would be interesting to see whether this combination could afford additive binding affinity or no significant difference or even worse performance. As shown in **Chapter 5**, we had synthesized tetravalent $GbO_3$-based *S. suis* inhibitors. The resultant bioactivities of glycodendrimers validated the multivalency strategy in the *S. suis*

inhibitors design. Thus far, the highest valency of a multivalent inhibitor for *S. suis* is up to 8.[12] With the experiences of the glycopolymer (**Chapter 3**), it is of interest to prepare carbohydrate-coated polymers with a maximized multivalency effect, also for *S. suis*. Besides the galabiose-specific *S. suis* adhesins, some *S. suis* strains are known to recognize a terminal sialic acid motif.[24] Considering this, it is appealing to create polymeric material that accommodates both ligands i.e. galabiose and sialic acid at same time. The resulting hetero glycopolymer could there be multivalent and multifunctional. Glyco-metal-nanoparticles (GNPs) as another multivalent sugar-based material has attract much attention. This concept could also be extended to magnetic glyconanoparticles that could act as diagnostic tools of *S. suis* as previously developed in our group.[25] (**Fig.4B**).

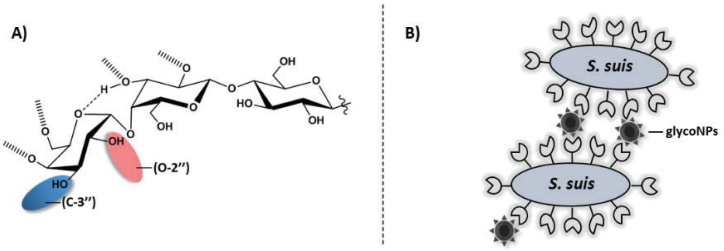



**Figure 4.** A) Hydrogen binding patterns of *S. suis* $P_N$ and indicated modification positions (shown in red and blue ovals, respectively)[23]; B) glycoNPs as biosensor for *S. suis* detection.[25]

## Remarks

Being a key principle in nature, the multivalency effect is increasingly being exploited in the design and synthesis of multivalent carbohydrate analogues to modulate carbohydrate-mediated pathogenicity. In this thesis, even with some challenges e.g. spatial and topological requirements, ligand density control, *etc.*, a number of multivalent inhibitors have been introduced against *P. aeruginosa* LecA, Cholera toxin and *S. suis* with varying degrees of potency enhancements. Unlike the low affinity and poor specificity of mono saccharides, their multiple display on synthetic scaffolds as an implementation of multivalency, clearly enhances these properties. Armed with advancing knowledge of protein-carbohydrate recognition and synthetic chemistry tools, more potent multivalent ligands will be developed with more and more applications for therapeutic and diagnostic purposes.

## Reference


[1] a) M. Fukuda, O. Hindsgaul, Molecular Glycobiology. New York: Oxford University Press; 1994, 1-52; b) A. Bernardi, J. Jiménez-Barbero, A. Casnati, C. De Castro, T. Darbre, F. Fieschi, J. Finne, H. Funken, K.-E. Jaeger, M. Lahmann, T. K. Lindhorst, M. Marradi, P. Messner, A. Molinaro, P. Murphy, C. Nativi, S. Oscarson, S. Penadés, F. Peri, R. J. Pieters, O. Renaudet, J.-L. Reymond, B. Richichi, J. Rojo, F. Sansone, C. Schäffer, W. B. Turnbull, T. Velasco-Torrijos, S. Vidal, S. Vincent, T. Wennekes, H. Zuilhof, A. Imberty, *Chem. Soc. Rev.*, 2013, 42, 4709-4727; c) V. Wittmann, R. J. Pieters, *Chem. Soc. Rev.*, 2013, 42, 4492-4503.

[2] a) F. Pertici, R. J. Pieters, *Chem. Commun*. **2012**, 48, 4008-4010; b) F. Pertici, N. J. de Mol, J. Kemmink, R. J. Pieters, *Chem. Eur. J*. **2013**, 19, 16923-16927; c) A. Imberty, M. Wimmerová, E. P. Mitchell, N. Gilboa-Garber, *Microbes Infect*. **2004**, *6*, 221-228.

[3] A. V. Pukin, H. M. Branderhorst, C. Sisu, C. A. Weijers, M. Gilbert, R. M. Liskamp, G. M. Visser, H. Zuilhof, R. J. Pieters, *ChemBioChem*, **2007**, **8,** 1500-1503.

[4] C. Sisu, A. J. Baron, H. M. Branderhorst, S. D. Connel, C. Weijers, R. de Vries, E. D. Hayes, A. V. Pukin, M. Gilbert, R. J. Pieters, H. Zuilhof, G. M. Visser, W. B. Turnbull, *ChemBioChem* **2009**, 10, 329-337.

[5] E. Fan, Z. Zhang, W. E. Minke, Z. Hou, C. L. M. J. Verlinde, W. G. J. Hol, *J. Am. Chem. Soc*. **2000**, 122, 2663-2664.

[6] J. Lebowitz, M. S. Lewis, P. Schuck, *Protein Sci.* **2002**, 11, 2067-2079.

[7] H. M. Branderhorst, R. M. Liskamp, G. M. Visser, R. J. Pieters, *Chem. Commun.*, **2007**, 5043-5045; c) I. Vrasidas, N. J. de Mol, R. M. J. Liskamp, R. J. Pieters, *Eur. J. Org. Chem*. **2001**, 4685-4692.

[8] W. E. Minke, C. Roach, W. G. J. Hol, C. L. M. J. Verlinde, *Biochemistry* **1999**, 38, 5684-5692.

[9] a) H. Frey, R. Haag, *Rev. Mol. Biotechnol*. **2002**, 90, 257-267; b) H. Türk, A. Shukla, P. C. A. Rodrigues, H. Rehage, R. Haag, *Chem. Eur. J*. **2007**, 13, 4187-4196.

[10] A. Kouki, R. J. Pieters, U. J. Nilsson, V. Loimaranta, J. Finne, S. Haataja, *Biology* **2013**, 2, 918-935.

[11] J. Ohlsson, A. Larsson, S. Haataja, J. Alajääski, P. Stenlund, J. S. Pinkner, S. J. Hultgren, J. Finne, J. Kihlberg, U. J. Nilsson, *Org. Biomol. Chem.*, **2005**, 3, 886-900.

[12] J. A. F. Joosten, V. Loimaranta, C. C. M. Appeldoorn, S. Haataja, F. A. El Maate, R. M. J. Liskamp, J. Finne, R. J. Pieters, *J. Med. Chem*. **2004**, 47, 6499-6508.

[13] Reymond *et al.*, to be published

[14] a) R. U. Kadam, D. Garg J. Schwartz, R. Visini, M. Sattler, A. Stocker, T. Darbre, J.-L. Reymond, *ACS Chem. Biol*. **2013**, 8, 1925-1930; b) R. U. Kadam, M. Bergmann, M. Hurley, D. Garg, M. Cacciarini, M. A. Swiderska, C. Nativi, M. Sattler, A. R. Smyth, P. Williams, M. Cámara, A. Stocker, T. Darbre, J.-L. Reymond, *Angew. Chem. Int. Ed*. **2011**, 50, 10631-10635.

[15] B. Blanchard, A. Imberty, A. Varrot, *Proteins*. **2014**, 821060-1065.

[16] A. Nurisso, B. Blanchard, A. Audfray, L. Rydner, S. Oscarson, A. Varrot, A. Imberty, *J. Biol. Chem*. **2010**, 285, 20316-20327.

[17] a) T. R. Branson, T. E. McAllister, J. Garcia-Hartjes, M. A. Fascione, J. F. Ross, S. L. Warriner, T. Wennekes, H. Zuilhof, W. B. Turnbull, *Angew. Chem. Int. Ed*. **2014**, 53, 8323-8327; b) Z.-S. Zhang, E. A. Merritt, M. Ahn, C. Roach, Z. Hou, C. L. M. J. Verlinde, W. G. J. Hol, E. Fan, *J. Am. Chem. Soc*. **2002**, 124, 12991-12998.

[18] P. I. Kitov, G. L. Mulvey, T. P. Griener, T. Lipinski, D. Solomon, E. Paszkiewicz, J. M. Jacobson, J. M. Sadowska, M. Suzuki, K. I. Yamamura, G. D. Armstrong, D. R. Bundle, *Proc. Natl. Acad. Sci. U. S. A.* **2008**, 105, 16837-16842.
[19] M. Miura, S. Pivsa-Art, T. Satoh, M. Nomura, M. Miura, G. Dyker, J. Heiermann, *Chem. Commun.*, **1998**, 17, 1889-1890.
[20] A. Sisson, D. Steinhilber, T. Rossow, P. Welker, K. Licha, R. Haag, *Angew. Chem. Int. Ed*. **2009**, 48, 7540-7545.
[21] a) I. Papp, J. Dernedde, S. Enders, R. Haag, *Chem. Commun*., **2008**, 5851-5853; b) I. Papp, J. Dernedde, S. Enders, S. B. Riese, T. C. Shiao, R. Roy, R. Haag, *ChemBioChem* **2011**, 12, 1075-1083.
[22] I. Papp, C. Sieben, A. L. Sisson, J. Kostka, C. Böttcher, K. Ludwig, A. Herrmann, R. Haag, *ChemBioChem* **2011**, 12, 887-895.
[23] S. Haataja, K. Tikkanen, U. Nilsson, G. Magnusson, K.-A. Karlsson, J. Finne, *J. Biol. Chem.*, **1994**, 269, 27466-27472.
[24] J. Liukkonen, S. Haataja, K. Tikkanen, S. Kelm, J. Finne, *J. Biol. Chem.* **1992**, 267, 21105-21111.
[25] N. Parera Pera, A. Kouki, S. Haataja, H. M. Branderhorst, R. M. J. Liskamp, G. M. Visser, J. Finne, R. J. Pieters, *Org. Biomol. Chem.*, **2010**, 8, 2425-2429.

# Appendices

## Abbreviation

δ chemical shift

Ac acetyl

AcOH acetic acid

$Ac_2O$ acetic anhydride

AFM atomic force microscopy

AgOTf silver trifluoromethanesulfonate

Boc *tert*-butyloxycarbonyl

$Boc_2O$ di-*tert*-butyl-dicarbonate

BOP (Benzotriazol-1-yloxy)tris(dimethylamino)phosphonium hexafluorophosphate

BSA bovine serum albumin

$Bu_4NBr$ tetra-n-butylammonium bromide

Bz benzoyl

cacld calculated

CuAAC Cu(I) catalysed Azide Alkyne Cycloaddition

CT/Ctx Cholera toxin

CTB Cholera toxin B subunit

d doublet

dd double doublet

D/Da Dalton

DIPEA *N-N'*-diisopropylethylamine

DLS dynamic light scattering

DMAP 4-dimethylaminopyridine

DMF dimethylformamide

DMSO dimethylsulfoxide

EDCI 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide

ELISA enzyme-linked immunosorbent assay

ESI-MS electronspray ionization mass spectrometry

ESI-Q-TOF-MS electrospray ionization quadrupole time-of-flight mass spectrometry

EtOAc ethylacetate

FITC fluorescein isothiocyanate

Gal galactose

$GbO_3$ globotriaosylceramide

GM1 monosialotetrahexosylganglioside

HATU 1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate

HEPES 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HIA heamagglutination inhibition assay

HMDS hexamethyldisilane

HPLC high pressure liquid chromatography

HR-MS high-resolution mass spectrometry

HRP horseradish peroxidase

HSQC heteronuclear single quantum coherence

Hz Hertz

$IC_{50}$ half maximal inhibition concentration

IR infrared spectroscopy

ITC isothermal titration calorimetry

J coupling constant

$K_d$ dissociation constant

$KHSO_4$ potassium hydrogensulfate

LecA *Pseudomonas aeruginosa* Lectin PA-IL

m mulitplet

M molar

MALDI matrix-assisted laser desorption/ionization

MD molecular dynamic

MeCN acetonitrile

MeOH methanol

MIC minimal inhibitory concentration

MNPG m-nitrophenyl-α-D-galactoside

NaAsc sodium ascorbate

$Na_2CO_3$ sodium carbonate

NaBr sodium bromide

$NaHCO_3$ sodium bicarbonate

NaOCl sodium hypochlorite

NaOMe sodium methoxide

NaOH sodium hydroxide

$NEt_3$ triethylamine

NMR nuclear magnetic resonance

NPs nanoparticles

PBS phosphate buffered saline

PDB protein data bank

ppm parts per million

r.t. room temperature

$SnCl_4$ Tin (IV) chloride

*S. suis Streptococcus suis*

SV-AUC sedimentation velocity analytical ultracentrifugation

t triplet

TBDMS *tert*-butyldimethylsilyl

TEMPO (2,2,6,6-tetramethylpiperidin-1-yl)oxy

Tf trifluoromethanesulfonate

TFA trifluoroacetic acid

$Tf_2O$ triflic anhydride

THF tetrahydrofuran

TMSOTf trimethylsilyl trifluoromethanesulfonate

*p*-TsOH *p*-toluenesulfonic acid

TLC thin layer chromatography

TsCl 4-toluenesulfonyl chloride

TOCSY total correlation spectroscopy

## Samenvatting

**Hoofdstuk 1** beschrijft dat koolhydraten een indrukwekkende hoeveelheid informatie in hun structuur bevatten. Veel ziektes beginnen echter door binding aan koolhydraten. Een goed begrip van deze binding en de adhesie van bijvoorbeeld pathogenen is cruciaal voor de ontwikkeling van therapeutica gebaseerd op koolhydraten. De natuur gebruikt koolhydraten en versterkt hun zwakke bindingsinteracties. Multivalentie is het belangrijkste principe hierin en wordt gebruikt door onderzoekers om behandeling van ziekteverwekkende bacteriën en bacteriële toxines te verbeteren. Het feit dat koolhydraten veel voorkomen en ook de toenemende resistentie van bacteriën tegen antibiotica benadrukt het belang van het verkrijgen van een goed begrip van de koolhydraatbiningsmechanismen die tot ziektes leiden. Aangezien multivalentie in de natuur het sleutelprincipe is, is een fundamenteel begrip van het multivalentie effect van groot belang in deze context. Recent is er veel vooruitgang geboekt in de synthese van glycoconjugaten, bijvb. glycoclusters, glycodendrimeren, glyco-eiwitten en glycopolymeren die een multivalentie effect laten zien. Er is getoond dat multivalente architecturen kunnen worden gecreëerd met verschillende vormen, oriëntaties, valenties en ligand dichtheden. In dit proefschrift, zijn synthetische glycomimetica beschreven die een multivalentie effect lieten zien tegen het eiwit LecA van *P. aeruginosa*, het cholera toxine en de bacterie *S. suis*. Deze veelbelovende resultaten zouden tot de volgende generatie multivalente glycoconjugaten met specifieke biologische activiteiten voor therapeutische toepassingen kunnen leiden.

In **hoofdstuk 2** hebben we vier gemodificeerde derivaten beschreven van een starre divalente inhibitor van LecA, een eiwit van *Pseudomonas aeruginosa* betrokken bij het infectieproces. Dit werk was geïnspireerd op de voorafgaande studie. Er was toen gebruik gemaakt van de bekende kristalstructuur van het eiwit LecA met de afstand van 26 Å tussen twee bindingsplaatsen (gemeten tussen de twee anomere zuurstoffen van de twee gebonden galactose moleculen). Een divalent molecuul bestaande uit een rigide glucose-triazool spacer met twee galactose liganden aan beide uiteinden bleek een krachtige remmer. In dit werk is ingezoomd op twee aspartaat residuen op het eiwit oppervlak niet ver van de bindingsplaatsen. Het doel

was beide aspecten te combineren en derivaten te ontwerpen van de oorspronkelijke LecA inhibitor die gebruik maken van hopelijk nuttige interacties van z'n spacer met delen van het eiwit zoals de twee genoemde aspartaten. Met de vorige inhibitor als uitgangssituatie werd een serie analogen gesynthetiseerd met nieuwe groepen zoals carboxylaten, ammonium groepen, aromatische groepen en positief geladen pyridinium groepen. De selectieve ontscherming van twee TBDMS groepen op twee spacer glucose eenheden maakte de modificatie van de twee primaire hydroxy groepen op de C(6) posities mogelijk. Vier nieuw gefunctionaliseerd inhibitoren werden verkregen door TEMPO oxidatie, triflaat vorming gevolgd door azide substitutie en reductie of CuAAC, of triflaat vorming gevolgd door pyridine additie. De structuur-activiteitsrelaties van deze stoffen voor LecA werden bepaald met een soort ELISA assay en met ITC metingen. De verwachting was dat de structuurveranderingen in sommige gevallen tot verhoogde binding aan het eiwit zou leiden. Beiden metingen gaven echter te zien dat de nieuwe verbindingen niet sterker bonden dan de oorspronkelijke structuur waarop ze waren gebaseerd. Moleculaire dynamica gebaseerd op de lecA kristalstructuur liet zien dat met name de ammonium groepen een goede interactie zouden kunnen hebben met de twee carboxylaten van het eiwit. Echter aangezien we dit niet terug zagen in de experimenten lijkt het erop dat andere conformaties van de spacer belangrijker zijn die wellicht de nieuwe groepen richting het oplosmiddel dirigeren i.p.v het eiwit oppervlak. Dit gegeven wordt meegenomen in nieuwe ontwerpen.

De bijzondere pentamere eiwitstructuur van het cholera toxine (CT), een $AB_5$ toxine, maakt het zeer geschikt voor inhibitie of remming door multivalente liganden. Deze kunnen voorkomen dat de toxine kan binden aan de GM1 gangliosiden op het cel oppervlak en tot ziekten kan leiden. Indrukwekkende resultaten zijn door onze groep beschreven van di-tetra- en octavalente liganden als CT inhibitoren. Er is gespeculeerd dat naast het multivalentie (chelaat) effect, eiwit aggregatie een belangrijke bijdrage levert aan de waargenomen inhibitieversterking. Deze aggregatie zou kunnen optreden doordat de valenties van de liganden en de toxine niet gelijk waren. Hierdoor werd het interessant om deze hypothese te verifiëren met de synthese van een pentvalente inhibitor gevolgd door het meten van z'n

inhibitie. Derhalve is in **hoofdstuk 3** een nieuw ontwikkelde strategie beschreven voor de synthese van een pentavalente GM1 inhibitor met dezelfde valentie als CT. Vijf GM1 moleculen werden verbonden aan een aromatisch dendritisch molecuul via vijf PEG-achtige 'linker'moleculen met de geschikte 'effectieve lengte' door gebruik te maken van de CuAAC reactie in de magnetron. De mate van inhibitie werd bepaald met behulp van een ELISA. De waargenomen CTB inhibitie ($IC_{50}$) van het pentavalente GM1 glycodendrimeer was 260 pM. Deze inhibitie was iets minder krachtig dan die van de tetravalente versie die ook gemaakt was. Niettemin, geven deze data aan dat de pentavalente verbinding een veelbelovende kandidaat tegen CT is. Om de rol van de valentie in de binding te bestuderen werden analytische ultracentrifuge (SV-AUC) experimenten uitgevoerd waarmee mogelijke aggregatie, geïnduceerd door het ligand, bepaald kan worden. Het bleek dat de mate van eiwit aggregatie door penta- en tetravalente liganden ongeveer gelijk was.

Er zijn in de literatuur studies verschenen over behoorlijk krachtige multivalente CT inhibitoren die eenvoudige en zwak binden liganden gebruiken zoals galactose en lactose. Een voorbeeld van een monosaccharide derivaat, *m*-nitrophenyl-α-D-galactoside (MNPG) was beschreven met een $IC_{50}$ waarde van 0.72 mM in een CT ELISA assay. In **hoofdstuk 4** hebben we een synthetische route beschreven naar een nanodeeltje waaraan MNPG geconjugeerd wordt als een mogelijke verbeterde CT antagonist. Voor een groot multivalentie effect werd een dendrimeer-achtige hypervertakte polyglycerol (hPGs) gekozen als de drager waaraan met hoge dichtheid groepen kunnen worden gezet. Een MNPG derivaat met een alkyn eraan is succesvol gesynthetiseerd in anomeer zuivere vorm d.m.v. een enzymatische hydrolyse als laatste stap. Het azide-gefunctionaliseerde hPGs polymeer werd geconjugeerd onder standaard 'click' chemie condities. Niettemin was het gewenste MNPG-glycopolymeer niet verkregen. Echter, het galactose gefunctionaliseerde hPGs, een eenvoudiger system, was wel succesvol gesynthetiseerd met 88 galactose eenheden per polymeer, gebruikmakend van dezelfde condities. Om dit onverwachte verschil te verklaren werd MNPG in detail bestudeerd met NMR en ESI-MS. Verrassend was gevonden dat het MNPG-alkyn molecuul gedimeriseerd was zodat het geen CuAAC koppeling met het polymeer kon doen. De dimerisatie vond

plaats tijdens de enzymatische splitsing van het ongewenste β-anomeer in het anomere mengsel. Een Glaser-achtige koppeling van twee MNPG-alkyn moleculen voorkwam dat de gewenste triazool gevormd kon worden. Om toch het multivalentie effect van het polymeer te bestuderen werd het galactose-polymeer bestudeerd als remmer van CT. Een sterke inhibitie werd waargenomen met een $IC_{50}$van 1.3 μM wat meer dan 2100 keer lager is per suiker dan vrij galactose. Dit resultaat geeft een indicatie dat een conjugaat van het betere MNPG ligand nog aanzienlijk krachtiger kan zijn als CT remmer. Er wordt nu gewerkt aan een nieuwe route om MNPG aan het hPGs polymeer te koppelen.

**Hoofdstuk 5** gaat over glycodendrimeren die de adhesie van *Streptococcus suis* (*S. suis*), zoals veroorzaakt door galabiose (Galα1-4Gal) binding, kunnen remmen. Twee goed bekende subtypes van *S. suis*, $P_N$ en $P_O$ binden aan galabiose bevattende koolhydraten, veroorzaakt door het recent gekarakteriseerde eiwit *Streptococcus* Adhesin P (SadP). De twee subtypes hebben een iets verschillende specificiteit voor galabiose derivaten. Dit verschil was in kaart gebracht door een systematische studie. Een krachtige monovalente remmer was gevonden die een fenylureum groep bevatte op de C-3'van de terminale galactose. Uit een studie van onze groep was gebleken dat sterke nanomolaire remming van *S. suis* mogelijk was door meerdere galabiose liganden te verbinden aan een dendrimeer. In dit hoofdstuk zijn de twee studies gecombineerd, door een galabiose bevattende trisaccharide, globotriose (Galα1-4Galβ1-4Glu, $GbO_3$) en z'n fenylureum derivaat te 'clicken' aan een tetravalent dendrimeer. De sterkte van de gemaakte inhibitoren werd getest met SadP. Een monovalente globotrioside met een verbonden fenylureum group werd gebruikt als monovalente referentie verbinding. De beste remmer was het fenylureum derivaat van de tetravalente $GbO_3$ met een $IC_{50}$ van 92 nM, hetgeen 11 keer beter is dan de monovalente referentieverbinding. De ongesubstitueerde tetravalente $GbO_3$ liet een beperkte vijf-voudige versterking zien van de inhibitie van SadP. Deze initiële data laten zien dat zowel het optimaliseren van het monovalente ligand als ook de multivalente presentatie beide belangrijke factoren zijn in de remming. Deze remming zal nog verder worden bestudeerd met behulp van de inhibitie van 'hemagglutination' door *S. suis*.

## Curriculum Vitae

Ou Fu was born on Dec. 17th, 1984 in Shanghai, China. In summer 2003, he finished his high school education and enrolled at East China Normal University (Shanghai) for his undergraduate studies in chemistry. In 2007, he earned his BSc in chemistry with a major in Applied Chemistry and completed a second major in International Economics and Business. In the same year, he moved to the Netherlands and started his master programme of "Drug Innovation" in Utrecht University. During that period, he conducted two academic internships in two research groups affiliated to the Utrecht Institute of Pharmaceutical Science (UIPS). One was in Pharmaceutics where he worked on molecular imprinting under the supervision of Dr. R. van Nostrum in the laboratory of Prof. dr. W. Hennink, the other was in Medicinal Chemistry and Chemical Biology where he specialized in antifungal agent making with the guidance of Prof. dr. R.J. Pieters. After receiving his master's degree, in 2010, he continued his work in Medicinal Chemistry and Chemical Biology as a PhD candidate, under the supervision of Prof. dr. R.J. Pieters. His work was focused on designing and synthesizing multivalent carbohydrate-based compounds as potential anti-bacterial agents. The results of his doctoral work are published in this thesis and have been presented orally as well as posters in several national conferences.

## List of publications and presentations

### Publications

"Functionalization of a Rigid Divalent Ligand for LecA, a Bacterial Adhesion Lectin", by Ou Fu, Aliaksei V. Pukin, H. C. Quarles van Ufford, Johan Kemmink, Nico J. de Mol, Roland J. Pieters, *Chemistry Open*, **2015**, in press (DOI: 10.1002/open.201402171).

"Tetra- versus Pentavalent Inhibitors of Cholera Toxin", by Ou Fu, Aliaksei V. Pukin, H. C. Quarles van Ufford, Thomas R. Branson, Dominique M. E. Thies-Weesie, W. Bruce Turnbull, Gerben M. Visser, Roland J. Pieters, *Chemistry Open*, **2015**, in press (DOI: 10.1002/open.201500006).

"Convenient Stereoselective Synthesis of Substituted Ureido Glycosides Using Stable 4-Chlorophenylcarbamates without the Requirement of Lewis Acids", by Steffen van der Wal, Ou Fu, Stamatia Rontogianni, Roland J. Pieters, Rob M. J. Liskamp, *SYNLETT* **2014**, 25, 205-208.

### Conferences

"Optimizing multivalent carbohydrate ligands to block protein-carbohydrate interactions", by Ou Fu, Roland J. Pieters, 20th Bijvoet Tutorial Symposium, 2014, 14-15 April, Soesterberg, the Netherlands. Oral presentation.

"Functionalizing a rigid spacer to optimize a potent divalent P. aeruginosa lectin inhibitor", by Ou Fu, Roland J. Pieters, 2013, 4-5 November, Lunteren, the Netherlands. Poster presentation.

"Tetra- and Pentavalent Inhibitors of Cholera Toxin", by O. Fu, H. M. Branderhorst, R. J. Pieters, Dutch Organization for Scientific Research-Chemical Science (NWO-CW) study group convention 2012, 22-24 October, Lunteren, the Netherlands. Poster presentation.

"Micro-chip based array of studying multivalency effect in glycobiology", by O. Fu, H. M. Branderhorst, R.Kooij, R. M. J. Liskamp, R. J. Pieters, NWO-CHAINS chemistry conference, 2011, 28-30 November, Maarsen, the Netherlands. Poster presentation.

“Chemoenzymatic Preparation of Galabiose for Studies in Glycobiology”, by O. Fu, H. M. Branderhorst, R.Kooij, R. M. J. Liskamp, R. J. Pieters, Dutch Organization for Scientific Research-Chemical Science (NWO-CW) study group convention, 2010, 25-27 October, Lunteren, the Netherlands. Poster presentation.

## Acknowledgements

Over the course of my long research journey as a PhD candidate in Medicinal Chemistry and Chemical Biology group, it has now finally culminated in the completion of this thesis. It is without doubt achieved by the wonderful supervision, support, and guidance given to me from prof. dr. R.J. Pieters. Your contribution towards its completion was unmatched. Since I started my writing, it had been one of the most significant academic challenges I had never expected. It was you that read my numerous revisions and helped me clarify all the confusions during the entire writing process. Thank you for staying late and spending extra hours for pointing out my typos and and other mistakes. Your intellectual support at crucial times especially reviewing and editing the entire document at final stage were invaluable, which led to the completion of this thesis on time. Besides my supervisor, I convey my profound gratitude to Prof. dr. R.J.M. Klein Gebbink, Prof. dr. G.J. de Jong, Prof. dr. R.M.J. Liskamp, Prof. dr. H. Zuilhof, Dr. C.A.M. de Haan. Thank you for sparing your precious time to evaluate my work and paramount notes. Without your approvals, this thesis would not have been possible.

Now, the time is followed by a special retrospective look back to the very first time when I landed at Schiphol and started my pursuit of my master programme in Utrecht University. I recall at that time Roland welcomed me to join the Medicinal Chemistry and Chemical biology group as an internee. Here, I was introduced to a new perspective on the interface of chemistry and biology at the molecular level. During that period of training, my academic work was wholeheartedly supported and appreciated by you and other colleagues I had been working with. By then, the seed of my interest in synthetic medicinal chemistry was planted and gradually increased. With your inspiration, I took my first and most important step of starting this long PhD journey in MedChem. Over there, not only did you create a highly dynamic working environment necessary to conduct my research, but also provided me the tools and financial means to complete the project. Thank you for giving me the encouragement and time to develop my own ideas. Especially in my first years' research, you helped me set the stage for the later achievements. I trully appreciate

your guidance leading my endeavour to success in the last year and your patience in helping me improve as a researcher.

As I planned and executed my research, I received the help and support from many people in so many ways. I would like to thank Johan Kemmink, who helped me conduct two dimensional NMR spectral analysis and molecular dynamics simulation of my final products. I would like to thank Javier Sastre Torano and Hans W. Hilbers for carrying out all the high resolution mass spectrometry (Q-TOF and MALDI) tests of my synthesized compounds. Many thanks to Linda Quarles van Ufford for helping me perform all the ELISA assays of my inhibitors against *P. aeruginosa* and CT, respectively. Special thanks to Dr. W.B. Turnbull, Thomas Branson and Dominique M.E. Thies-Weesie for helping me set and conduct the SV-AUC measurements of my glycodendrimers. Thanks also goes to Nico de Mol for all the thermodynamic studies and handling a massive amount of open-ended data from the ITC measurements. All your work awarded me a wider field of view in multivalent binding.

I would like to thank Dr. G.M. Visser for the numerous suggestions and insights. It was much fun discussing with you about the field of CT inhibition. I learned a lot from our conversations.

I would like to thank Prof. dr. R. Haag for providing his azido functionalized polymers with me for my sugar-ball study. Thanks to Dr. S. Bhatia, Markus Hellmund for their valuable advice and information of reaction conditions as the guidance for carbohydrate-grafted nanoparticles introduction.

I would like to thank Prof. dr. U. Nilsson and Priya Verma for providing and synthesizing the analogues of globotrioside for my multivalent glyco-ligands against *S. suis* study. I would also like to thank Dr. S. Haataja for his collaboration on testing anti-adhesion potential of my end-products in *in vitro* assays.

To other dedicated members of MedChem: Nathaniel Martin, Dirk Rijkers, Rob Ruijtenbeek, John Kruijtzer, Rob Liskamp, Ed Moret and Jack den Hartog, I highly appreciate all the guidance you have given me. Thank you for your interest and fruitful discussion about my research every time when I presented my project on Friday's group meeting. Your invaluable comments and advice even critique had

been an important factor for my success in my academic period and the improvement of my presentation skills as well. Especially, I wish to express my deep gratitude to Dirk Rijkers. It was very kind of you taking your time on my Dutch samenvatting maken polishing.

Thanks to all the members of “sugar” group, who had directly or indirectly helped me in the development of my work. No doubt, this work could not be a success without the pioneering and superb quality of observational work done specifically by Hilbert and Francesca. Francesca, who worked at the fumehood next to me, was a very “big” and talented chemist. It was always nice talking to you about chemistry, food, culture, arts etc. I remembered that each time I had some issues with chemistry, you were keen to help me out. You still helped me out even after you left for Italy! Special thanks to Aliaksei, thank you for sharing your experience and knowledge about organic synthesis especially on carbohydrate chemistry. I had learned a lot from the critical discussions on numerous parts of my projects with you. Nuría, my gratitude for the time you spent answering my questions. The extension of your enthusiasm and passion from life to work had a great impact on me. Thank you, Raymond as well as Hilde for your suggestive remarks about my project and help in the lab. I truly appreciated all the remote technical assistance I received from you guys to solve my problems throughout my work involving labelling, dialysis, etc. I have to thank Jie Shi and Hao Zhang for being my paranymphs and standing by my sides on my defence day. Without you, the defence would not be complete. Thanks Marjon, Alen, Suhela, Wenjing, Guangyun, Diksha, Arwin and Robb for all the related discussion to improve the quality of my work in the “sugar meeting”. In spite of the enormous work pressures I had been facing, all of you inspired my every effort to make it happen.

Here, I would like to take this opportunity to extend my appreciation to all my colleagues for your help at every stage of my work. My sincere gratitude goes to Rik and Steffen who had helped me wherever and whenever I got stuck, and for being patient with me. Each of you played a very crucial role helping me with the Dutch language and culture. The happy time we spent together for the Chinese hot pot was a very memorable occassion. Jin, my first Chinese fellow, thank you for the helpful

work discussion and I enjoyed playing table tennis with you and Steffen after work. Jeroen, thanks for teaching me the operation of PamStation and PC software for the data readout. Many thanks to Timo, Peter, Paul and Matthijs for all the moments you helped me solve the problems I encountered with NMR machine and HPLC apparatus. Additionally, thank you guys for showing me how good you are not only at execution of chemical reactions but also launches of varied ways to open the beer bottles without a bottle opener. “Impossible is nothing”. Loek, thanks for the help in the lab and organizing all the fishing events. You guys are always helpful as well as so much fun. Thanks also to other Medchemmers including Xin, Laurens, Gosia, Jack, Helmus. I really enjoyed chatting with you about everything. All you guys are more than colleagues for me. Moreover, it’s time to express my thanks to friendly colleagues from other research groups. Thank Yang, Yinan, Dandan in Pharmaceutics group, who helped me with the DLS test of my polymer material and a further thanks to Yinan with assistance in performing the colorimetric tests. My gratitude goes towards Bo for his helpful discussion about how to determine the saccharide units on my glycopolymer backbone and special thanks to Luís for the practical guidance of cassette usage and protocol for dialysis. I would also like to thank Yuxing in Organic Chemistry and Catalysis group for helping me with the infrared measurements which were essential to track the carbohydrate coating on polymer surfaces. Special thanks to Dr. R. van Nostrum (Pharmaceutics) and Prof. dr. R. Gebbink (Organic Chemistry and Catalysis) for providing me the platforms to undergo my tests. Many thanks to Houjiang in Biomolecular Mass Spectrometry and Proteomics group for performing the HR-MS test of my big molecules for emergency management and Daniele, an Erasmus exchange master student for producing some primary data for my project. I am deeply grateful for all the time we worked together. Thank you all.

Also, I owe a big thank you to Iskandar Gandasasmita, Piek Wei Ong from Pamgene Inc., who provided me all the glyco-functionalized and FAEC chips for my glycan arrays.

In addition, I would like to take this opportunity to thank Natatia and Paula for the help of my administration work and organization of annual lab outing and Christmas events. They were lots of fun!

Last but not least, it comes to my personal source of gratitude to my parents. It is very hard to be far away from you. Never anticipated that the frustration and loneliness was along with my write-up and academic research. You are my enduring source of strength and a constant source of support-emotionally, spiritually, and financially. You would always love and stand by me unconditionally. Mom, no matter what accomplishments I achieved is all because of you and Dad. I love you both. This thesis is dedicated to you.

Please allow me to express my infinite gratitude to numerous Asian friends in Utrecht who offered their support and love in varied ways as well. 张浩，施杰，鲁文静，于光允，杨欣，张金强，周厚江和戴思思（夫妇），石洋和孙庆雪（夫妇），黄宇星，罗文豪，钱庆云，高金宝，魏莹，杨鹏飞，冯博，王刚，刘铂洋，刘芳，陈祎楠，李丹丹，娄博，陈伟峦，孙飞龙，关吉斌，伍江波，赵玉珑，孙志祥，刘彦超和沈倩（夫妇），李衡宇，匡迎焕，彭博，赵一鸣，彭茂，李帅，刘金峰，张淼，杨浩然，张帅，李晓倩，李筱，王雨婷，王然，滕菲，徐诗晗，杨帆，李嘉伦，王洪洪和张芳（夫妇），李璋，方锐，张硕伟，李巍，梁旭旭，吴斌，庞垚，宋涌，阮謦，李源，陈挚，陶亮，还有郑尚镐（정상호），Naushad Hossain, Farshad Ramazani, Mazda Radmalekshahi, Teck Yew Low, Jackie Ye, Kelvin Lau, 感谢有幸能够结识你们这些来自五湖四海的朋友，从而点缀了我的荷兰求学生活，使它变得如此多姿多彩，也是你们铸就了今天的自己。虽然我们终将各奔前程，往后鲜有交集。但和你们的相遇是一种缘分，已深深地烙在我心里，借此机会，由衷地谢谢！

In the end, it is a matter of great satisfaction and pleasure to present this thesis “Multivalent carbohydrate inhibitors of bacterial lectins and toxins”. Once more, I would like to extend my sincere and heartfelt gratitude to all the people who had helped me. Without your help, cooperation and encouragement, I would not have come this far. Thanks to all the colleagues of MedChem group for creating a sound work environment and giving me a great time.

Working with you guys has itself been rewarding.

Hartelijk bedankt!!!

非常感谢！

Ou Fu （傅鸥）