Inhibition of Galectins and O-GlcNAc Transferase with Di- and Multivalent Ligands

Hao Zhang

Supervisor: Prof. Dr. Roland J. Pieters

Cover image: Crystal structure of Gal-3 with Lactose and OGT with a peptide substrate

ISBN: 978-90-393-6973-9

The printing of this thesis was financially supported by:

Utrecht Institute for Pharmaceutical Sciences

Printed by: **ProefschriftMaken||www.proefschriftmaken.nl**

Cover by: **Hao Zhang and Bihao Lan**

Inhibition of Galectins and O-GlcNAc Transferase with Di- and Multivalent Ligands

Inhibitie van Galectines en O-GlcNAc Transferase met Di- en Mulitivalente Liganden

(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 woensdag 18 april 2018 middags te 12.45 uur

door

Hao Zhang

geboren op 13 march 1988 te Shandong Province, China

Promotor: Prof. dr. R. J. Pieters

The research described in this thesis was financially supported by the China Scholarship Council (CSC) under file number of 201306210049.

谨以此书献给我的家人！

不忘初心，方得始终。

——解自《华严经》

“Now I have come to crossroads in my life, I always knew what the right path was. Without exception, I knew, but I never took it, you know why, it was too damn hard.”

-*Scent of a Woman*

Table of Contents

List of Abbreviations

Ac	**acetyl**
Ac_2O	**acetic anhydride**
Alloc	**allyloxycarbonyl**
BSA	**Bovine serum albumin**
Bop	**Benzotriazol1yloxy)tris(dimethylamino)phosphonium hexafluophosphate**
B_{max}	**maximal binding signal**
COSY	**correlation spectroscopy**
CuAAC	**Cu (I)-catalyzed azide-alkyne cycloaddition**
DIPEA	***N,N*-diisopropylethylamine**
DMSO	**dimethylsulfoxide**
DCM	**dichloromethane**
EDC	**1-(3-Dimethylaminopropyl)-3-ethylcarbodiimide**
EtOAc	**ethylacetate**
ESI	**electrospray ionization**
Fmoc	**fluorenylmethyloxycarbonyl**
GTs	**glycosyltransferases**
GlcNAc	***N*-acetyl-D-glucosamine**
Glu	**glutamine**
HBP	**hexosamine biosynthetic pathway**
HSQC	**heteronuclear singular quantum correlation**
Gal-1	**galectin-1**
Gal-3	**galectin-3**
HOBt	***N*-hydroxybenzotriazole**
HPLC	**high performance liquid chromatography**
HRMS	**high resolution mass spectrometry**
IC_{50}	**median inhibition concentration**
K_d	**dissociation constant**
MeOH	**methanol**
MS	**mass spectrometry**
m/z	**mass to charge ratio (mass spectrometry)**
NMR	**nuclear magnetic resonance**
NHS	***N*-hydroxysuccinimidyl**
NOESY	**nuclear overhauser effect spectroscopy**
OGA	***O*-linked-*N*-acetyl-glucosaminase**

OGT	***O*-linked-*N*-acetyl-glucosaminyl transferase**
PEG	**polyethylene glycol**
Ph	**phenyl**
SDS-PAGE	**sodium dodecyl sulfate polyacrylamide gel electrophoresis**
SPPS	**solid phase peptide synthesis**
tBu	***tert*-butyl**
TDG	**thiodigalactoside**
TBAF	**tetra-*n*-butylammonium fluoride**
TFA	**trifluoroacetic acid**
TSTU	**2-Succinimido-1,1,3,3-tetra-methyluronium tetrafluoroborate**
THF	**tetrahydrofuran**
TLC	**thin layer chromatography**
TPR	**tetratricopeptide repeats**
UDP	**uridine diphosphate**

Amino acids

Ala	**A**	**Alanine**
Arg	**R**	**Arginine**
Asn	**N**	**Asparagine**
Asp	**D**	**Aspartic acid**
Cys	**C**	**Cysteine**
Glu	**E**	**Glutamic acid**
Gln	**Q**	**Glutamine**
Gly	**G**	**Glycine**
His	**H**	**Histidine**
Ile	**I**	**Isoleucine**
Leu	**L**	**Leucine**
Lys	**K**	**Lysine**
Met	**M**	**Methionine**
Phe	**F**	**Phenylalanine**
Pro	**P**	**Proline**
Ser	**S**	**Serine**
Thr	**T**	**Threonine**
Trp	**W**	**Tryptophan**
Tyr	**Y**	**Tyrosine**
Val	**V**	**Valine**

Chapter 1

General introduction I

1.1 Carbohydrates and lectins in molecular recognition

Carbohydrates are the most abundant organic compounds found on the planet, and they can be divided into four major classes: mono-, di-, oligo- and polysaccharides. These classes have numerous roles but it is becoming more and more clear that molecular recognition is an important one. The surface of mammalian cells is covered by a dense layer of carbohydrates designated the glycocalyx, in which, carbohydrates are conjugated to proteins (glycoproteins, proteoglycans) and to lipids (glycolipids). This way the carbohydrates are well presented to take part in molecular recognition processes. Numerous proteins are able to interact with the carbohydrates. Among such proteins, lectins binding to mono- and oligosaccharides specifically have come into the forefront of biological research.

Lectins are carbohydrate-binding proteins and can thus recognize particular glycans among the vast array expressed in biological systems. Most animal lectins can be classified into four distinct families: 1) C-type lectins (including the selectins); 2) P-type lectins; 3) pentraxins; 4) galectins, formerly known as S-type or S-Lac lectins.[1] During the past decades, there has been remarkable progress in elucidating the functions of the lectins related to their carbohydrate binding properties.

1.2 Galectins and biological activities

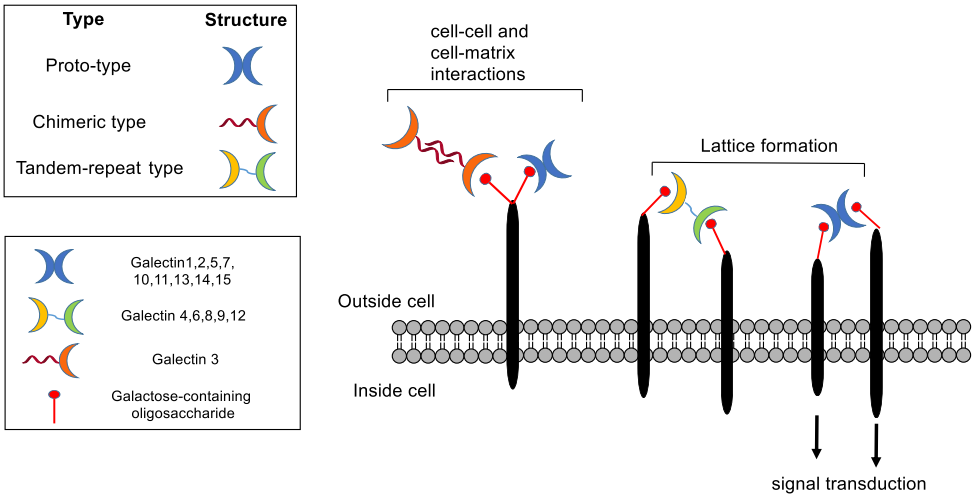


Figure 1. Subfamilies and functions of galectins.

The galectins were first discovered by the Barondes group in the early 90s while analyzing tissue extracts for their ability to bind immobilized β-galactosides. Galectins as a subfamily of lectins, can recognize β-galactosides specifically and contain conserved

carbohydrate-recognition domains (CRDs) consisting of about 130 amino acids, which are responsible for carbohydrate binding. They display an intriguing combination of intra- and extracellular activities by interacting with various glycoproteins both inside and outside of cell. To date, there are 15 members, and are present in organisms ranging from nematodes to mammals.

Galectins can be subdivided into three types based on their structure: proto-type (galectin-1, 2, 5, 7, 10, 11, 13, 14 and 15); tandem type containing two distinct CRDs connected through a polypeptide linker of up to 70 amino acids (galectin-4, 6, 8, 9 and 12); and finally chimeric type (Gal-3) is in a class of its own because it is composed of an N-terminal collagen-like domain with several repeats of a peptide sequence rich in proline, glycine, and tyrosine residues followed by a C-terminal domain.[2]

Galectins are found both inside and outside cells (Figure 1). Extracellularly, galectins can interact with cell-surface glycoconjugates and glycoproteins containing suitable saccharides and consequently crosslink many glycoconjugates. These interactions can produce a cascade of signal events and may lead to distinct cellular responses. Through this mechanism, they modulate mitosis, apoptosis and cell-cycle progression. Intracellularly, galectins have been identified to be essential for pre-mRNA splicing in the nucleus and the regulation of cell growth, apoptosis and cell-cycle progression.[3,4] Besides, Gal-3 was found to shuttle between the nucleus and the cytoplasm[5]. However, the exact mechanisms in these processes are often not clear and protein-protein interactions are involved in most of the intracellular cases, rather than lectin-carbohydrate interactions.[4]

1.3 Carbohydrate-recognition domains (CRDs) structure

The galectin CRD is of a β-sandwich type with two antiparallel β-sheets, in which the carbohydrate binding site is a long groove consisting of five subsites (A-E) as described by Leffler *et al* (Figure 2).[6] This site is long enough to contain a tetrasaccharide and the subsite C with its highly conserved residues can specifically recognize the β-galactoside.[7,8] The second most conserved region, subsite D, commonly binds the glucose in case of lactose. The subsite A and B structures are more varied than subsite C and D, which leads to different selectivities for the various galectins. Finally, subsite E doesn't bind to endogenous saccharides but still contributes to the binding affinity of galectins to certain carbohydrates even if it is less well defined in the structure.

1.4 Galectin inhibitors

As was mentioned in the previous pages, galectins can modulate various biological functions and play profound roles in many pathological processes. To decipher the precise action mechanisms of different galectins, selective inhibitors are an urgent need. In addition, the therapeutic potential of these inhibitors is increasing as more and more observations are pointing to the relevance of the galectins in many conditions and

diseases like inflammation, cancer and fibrosis. Several galectin inhibitors have been developed with an emphasis on Gal-1 and -3. These inhibitors can be divided into small molecule carbohydrate-based inhibitors, non-carbohydrate-based inhibitors, and carbohydrate-based multivalent inhibitors.

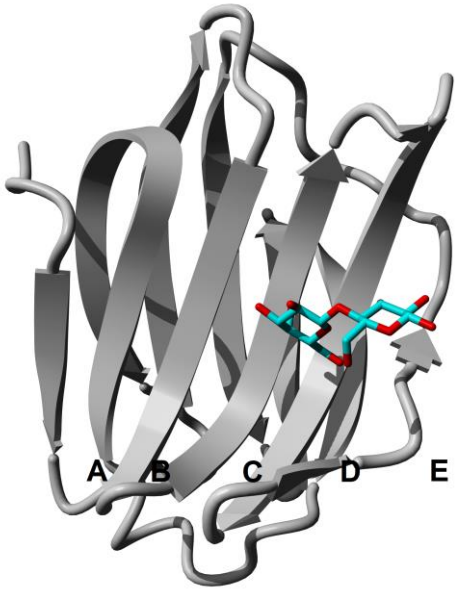


Figure 2. Visualization of a high-resolution X-ray structure of galectin-3 CRD in complex with lactose (PDB id 3ZSJ). The figure was generated with YASARA.

1.4.1 Small molecule carbohydrate-based inhibitors

As galectins recognize β-galactosides, and as a group, share a significant number of conserved amino acids in their CRD's, the prevailing majority of inhibitor discovery efforts have focused on the synthetic modification of lactose (Compound **1**: K_d = 220 µM for Gal-3, K_d = 190 µM for Gal-1) and *N*-acetyllactosamine (Compound **2**: K_d = 67 µM for Gal-3) (Figure 3A).[9,10,11] An obvious increase in binding affinity for Gal-3 was obtained by the introduction of an aromatic moiety at C3 of galactose.[12,13,14] The resulting molecule had a strong interaction with an arginine residue (Arg-144) through arene-guanidinium interactions. The phenomenon was first found by Nilsson and co-workers, and led to the further development of optimized inhibitors with enhanced affinity (Compound **3**: K_d = 0.32 µM for Gal-3).[11] In the meantime, a new scaffold, thiodigalactoside (TDG, compound **4**: K_d = 49 µM for Gal-3, K_d = 24 µM for Gal-1), was found to produce more advanced Gal-3 inhibitors because of its enhanced glycolytic stability and similar binding mode compared

to lactose and LacNAc.[15,13] Further research showed not only that a similar derivatization with aromatic amides at C3 of the TDG galactose could increase the affinity for Gal-3 through interacting with arginine, but the same modification at the second galactose C3 further enhanced the binding. Moreover, a triazole linkage between the carbohydrate and the aromatic moiety instead of an ether, ester, or amide linkage was found to increase the affinity for Gal-1 and -3 even further (Compound **5**: K_d = 0.66 μM for Gal-3), as the synthesis of carbohydrate based 1,2,3-triazole analogs by Cu(I)-assisted 1,3-dipolar azide-alkyne cycloaddition (CuAAC) was applied.[16,17] Van Hattum and co-workers introduced 4-phenyl-1H-1,2,3-triazol-1-yl substituents at the TDG C3 introduced by CuAAC, which increased the affinity for Gal-3 to 44 nM but yielded little selectivity between Gal-1 and -3 (Compound **6**). The bulkier 4-(4-phenoxyphenyl)-1H- 1,2,3-triazol-1-yl substituent, however, increased the preference for Gal-3 over Gal-1 to more than 200-fold (Compound **7**: K_d = 0.36 μM for Gal-3, K_d = 84 μM for Gal-1). In line with these finding, structure-based drug design was utilized to develop selective galectin-3 inhibitors. A series of patents have been filed about TDG-based inhibitors for Gal-3, including the TD139 (Compound **8**: K_d = 0.87 μM for Gal-3, K_d = 84 μM for Gal-1), which have exhibited promising effects in *in vivo* experiments attenuating the late-stage progression of lung fibrosis after bleomycin.[18–20,21]

Compared with disaccharides, the monosaccharides (Figure 3B) have an innately weaker affinity.[10] However, monosaccharide-based inhibitors can offer higher ligand efficiency and glycolytic stability, and are more easily modified to make them suitable as a therapeutic agent. The C1 and C3 hydroxyls of galactose were modified to improve affinity and selectivity for Gal-1 and -3. Most of these molecules have a higher affinity than galactose but can't reach the affinity of unmodified lactose.[10,12,22] A few monosaccharides showed a better potency than lactose. For example, compound **9** showed a K_d of 313 μM for Gal-1 (2.5 fold higher than lactose in a same assay).[23] Oxime ether **10**, a galactose derivative (K_d = 330 μM for Gal-3), showed an enhanced affinity for Gal-3 compared to galactose and another 30-fold affinity enhancement was observed by attachment of a triazole group at C3 (**11**: K_d = 11 μM).[10] In addition, the O2 of galactose can be derivatized to obtain potential galectin inhibitors but this influences the binding less because the hydroxy group on C2 is pointing away from the protein into the solvent. A new scaffold, talose (the C2 epimer of galactose), was used to obtain taloside-based inhibitors of galectins. Some O2-substitued talosides were shown to achieve affinity enhancements by additional ligand-protein interactions of substituents like in compound **12** and **13**.[24,25] These inhibitors still need to be optimized but talosides may offer desirable hydrolytic stability in talose-based drug candidates as they are usually "unknown" substrates for carbohydrate-processing enzyme in mammalians.

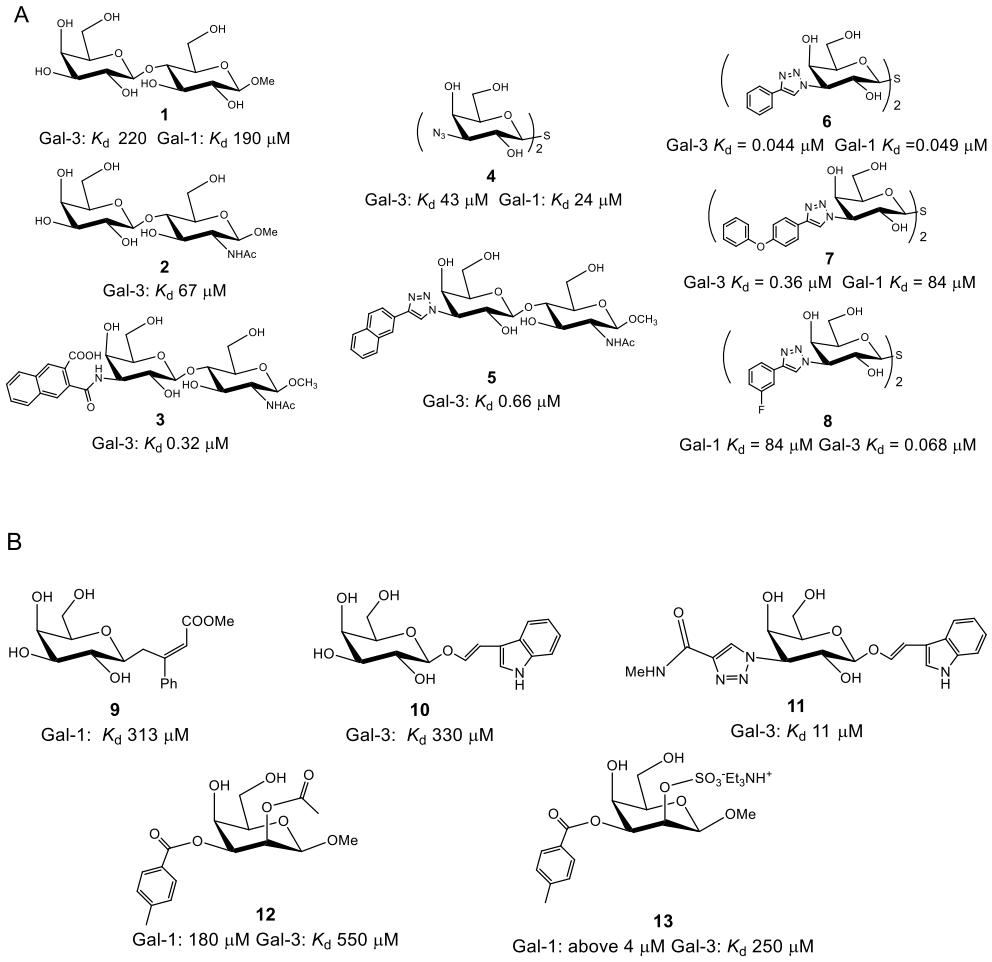
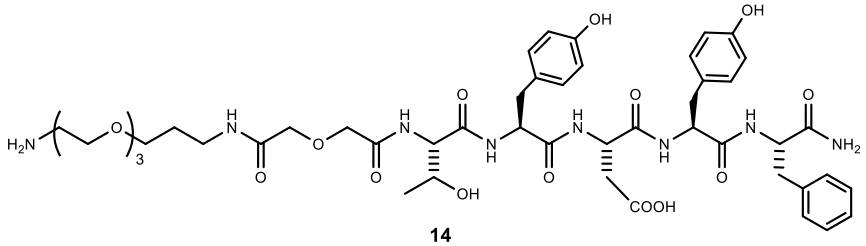


Figure 3. Small molecule carbohydrate-based inhibitors. (A) disaccharide galectin inhibitors; (B) monosaccharide galectin inhibitors.

1.4.2 Non-carbohydrate-based inhibitors

Currently, non-carbohydrate-based inhibitors for galectins are mainly peptides and peptidomimetics (Figure 4). Compared with carbohydrates, peptides are easier to synthesize but can exhibit strong immunogenicity.[26] In addition, some peptides called "carbohydrate-mimetic peptides (CMPs)" have been identified and these compounds are recognized by carbohydrate-binding proteins.[27,28] Short pentapeptides like compound **14** containing Tyr-X-Tyr motif were identified that interfere with galectin binding to its natural carbohydrate.[29,30] Their inhibition was confirmed with surface plasmon resonance (SPR) by using a solubilizing tail on the peptide. The resulting effect of peptides was explained by

the nature of a multivalent ligand which can cause galectin aggregation through cross-linking several lectin molecules. However, the targeting site of these peptides and binding mode still need to be elucidated.[28,30,30] Another case of synthetic peptides as galectin inhibitors was reported by Zou and co-workers, in which peptides (**15**, **16** and **17**) bound with high affinity to recombinant Gal-3 (K_d ≈ 5-80 nM). By blocking the Gal-3 CRD, the peptide significantly inhibited Gal-3 on the cell surface and hence reduced metastasis-associated carcinoma cell adhesion, but the inhibiting mechanism of it is still not clear.[32]



14
Gal-1 IC_{50} 5.2 μM

SMEPALPDWWWKMFK

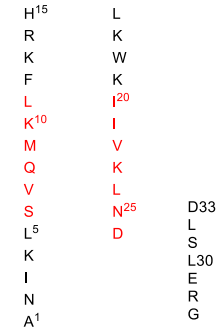
15
Gal-3 K_d 17 nM

AYTKCSRQWRTCMTTH

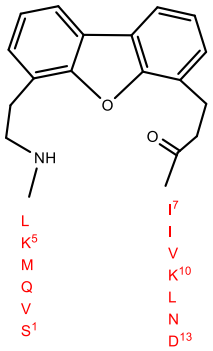
16
Gal-3 K_d 5.7 nM

ANTPCGPYTHDCPVKR

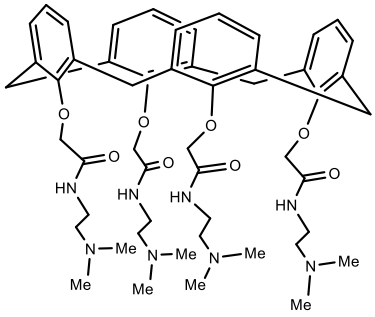
17
Gal-3 K_d 72 nM



18
Gal-1 K_d 90 nM



19
Gal-1 K_d 5-100 μM



20
Gal-1 K_d 30 μM

The colored sequences in **18** are preserved in **19**

Figure 4. Non-carbohydrate-based inhibitors

The most well-studied peptide-based galectin inhibitor is compound **18**, anginex, a potent antiangiogenic and antitumor peptide, that has been shown to bind Gal-1, -2, -7, -8N and -9N, but not Gal-3, -4N, -4C and -9C.[33] Anginex is a synthetic peptide (ANIKLSVQMKLFKRHLKW KIIVKLNDGRELSLD) derived from the β-strand regions of anti-

angiogenic proteins PF4, IL-8 and BPI.[34,35,36] Surface plasmon resonance showed that anginex bound Gal-1 with a K_d of about 90 nM.[37] To reduce anginex's molecular size and peptidic nature, a structure-activity relationship (SAR) study was conducted, which identified some key amino acid residues and indicated that the conformation of the peptide (anti-parallel β-sheet) was essential for its bioactivity.[38,39] With this knowledge, a partial peptide mimetic of anginex, compound **19** (6DBF7), was designed and it contains six amino acid residues at the N-terminus and seven at the C-terminus linked by a dibenzofuran (DBF) moiety (SVQMKL-[DBF]-IIVKLND).[39] The novel anti-angiogenic compound was shown to be more effective *in vivo* than parent anginex. In addition, NMR studies showed that compound **19** and its analogs bound to Gal-1 on one side of the β-sandwich opposite from the lectin's carbohydrate binding site (K_d = 5-100 μM), which indicates that the peptidomimetic is a noncompetitive inhibitor that targets Gal-1 with a different binding mode compared to most of the carbohydrate galectin inhibitors.[40] With the success of the partial peptidomimetics, a series of topomimetics were synthesized based on a calix[4]arene backbone, which allowed chemical substituents to approximate the molecular dimensions and amphipathic features of compound **18** and **19**. Among these molecules, compound **20** (calixarene 0118) was shown to be a potent angiogenesis inhibitor *in vitro* and highly effective at inhibiting tumor growth in murine tumor models.[41] Considering some similarity to compounds **18** and **19**, Gal-1 was assumed to be the target of compound **20**. NMR spectroscopy (^{15}N-^{1}H HSQC) showed that compound **20** bound to Gal-1 also at a site away from the lectin's carbohydrate binding site with a K_d of 30 μM.[42]

1.4.3 Carbohydrate-based multivalent inhibitors

The concept of multivalency has gained interest in biological systems since cell-cell recognition and cell signaling often rely on the formation of multiple receptor-ligand complexes. Diverse carbohydrates decorating the cell surface or in their free state, can make contact with a protein and then produce the corresponding effect. The individual protein-carbohydrate interactions are usually weak but multivalent ligands can compensate for this drawback through binding two or more ligands to one biological entity simultaneously. Currently, two typical mechanisms can explain the increased affinities of multivalent ligands for their binding partners: the chelation effect and the statistical rebinding effect. The chelation effect describes the intramolecular interaction between ligands and receptors with multiple binding sites. When the ligands are linked by a suitable spacer, the higher binding affinity of multivalent ligands can be achieved because of an entropically favored binding.[43] The statistical rebinding effect is intermolecular. A multivalent ligand can display higher affinities in the binding of only one receptor site. The effect is obtained by an overall slower off-rate binding as the receptor site is rapidly reoccupied by a nearby (sub)ligand when the first one leaves.[44]

Galectins are often found as dimers (covalent or noncovalent) and sometimes even oligomers can be part of multivalent interactions. Significant efforts have been put into the design of multivalent galectin inhibitors that contain carbohydrate ligands like lactose and LacNAc, to achieve multivalency effects (Figure 5). Divalent lactose-bearing compounds showed increased potency to galectins compared to lactose, but the relative potency increases on a per lactose basis were very small (compound **21** and **22**).[45,23] To improve it, different scaffolds were introduced leading to larger enhancements as were clearly seen in the case of compound **23**, **24** and **25.**[46,47,48] Results suggest a stronger interaction between these compound and galectins, especially for Gal-3 where the relative potency per lactose ligand unit as compared to free lactose is increased. Further research showed that the effects on Gal-3 likely resulted from the aggregation behavior of the chimeric Gal-3.[48,47] Apart from synthetic scaffolds, proteins like bovine serum albumin (BSA) were also used as a multivalent scaffolds and different glycoproteins were prepared by conjugation of the protein to glycans. In studies reported by Elling and co-workers, neo-glycoproteins containing BSA as a scaffold and a tetrasaccharide (LacNAc-LacNAc: Galβ1,4GlcNAcβ1, 3Galβ1,4GlcNAc or LacDiNAc-LacNAc: GalNAcβ1,4GlcNAcβ1,3Galβ1,4GlcNAc) for multivalent binding were prepared. The binding experiments by an ELISA-type assay showed the K_d value of neo-glycoproteins with different numbers of attached glycans (compound **26a-k** and **27a-k**) reached the nanomolar range and the nature of the terminal sugar of the tetrasaccharides could increase selectivity between Gal-1 and Gal-3.[49]

Complex polysaccharides extracted and modified from natural sources such as egg, apple and citrus exhibited inhibiting effects on galectins.[50,51,52] Mayo and co-workers have identified some galactomannans (GMs) with α-1,6-galactose branched form a β-1,4-mannose backbone, that bind Gal-1 and -3 in the micromolar range (from 80 to 10 μM) though seemingly at a site different from the conventional carbohydrate binding domain.[53,54,55] Among these GMs, GM-CT-01 (Davanat, compound **28**) was discovered to target Gal-1 and -3 (Gal-1: 10 μM, Gal-3: 2.8 μM) and was patented as a cancer therapeutic that can decrease the negative side effects of co-therapeutics such as 5-fluorouracil (5-FU) and Adriamycin.[56,50,57,58] It has a weight-average molecular weight of 59 kDa and an average repeating unit of 17 β-D-Man residues and 10 α-D-Gal residues, with an average polymeric molecule containing approximately 12 such repeating units. As a modified version of GM-CT-01, GR-MD-02 known as a Gal-3 inhibitor (Gal-1: 8 μM, Gal-3: 2.9 μM), was found and it is a complex polysaccharide with rhamnogalacturonate backbone and branches terminating with galactose and arabinose at varying intervals.[57,58] In the phase II trial against non-alcoholic steatohepatitis with cirrhosis (NASH-CX), a significant biological activity of GR-MD-02 was demonstrated in patients.[45] However, it has to be noted that both GM-CT-01 and GR-MD-02 have no obvious selectivity between Gal-1 and Gal-3 and may bind to other targets as well.[59,60]

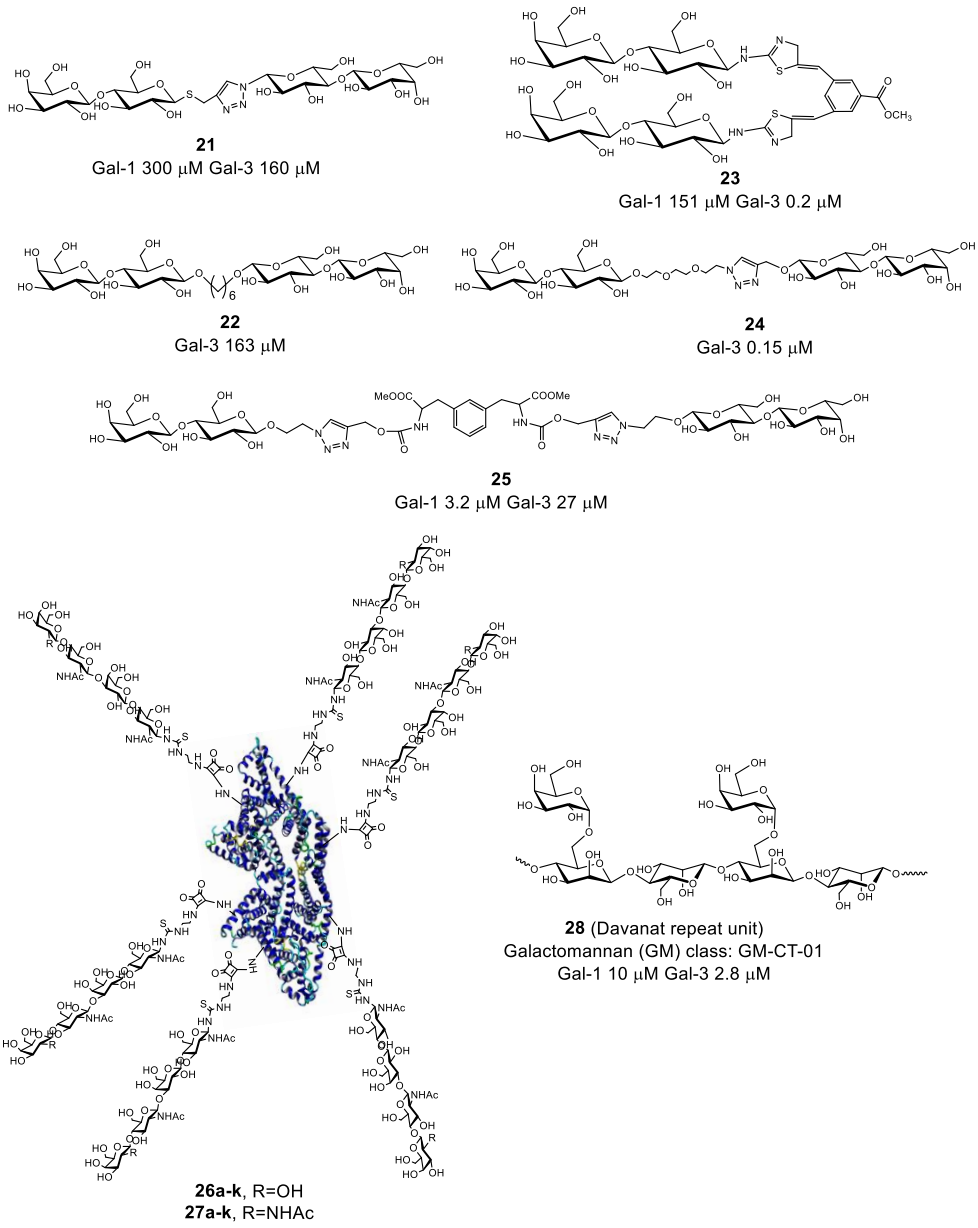


Figure 5. Carbohydrate-based multivalent inhibitors

1.5 Project aim

Galectins have diverse functions and are involved in many biological processes like cell growth, adhesion and signaling. In addition, there is increasing evidence that galectins can be therapeutic targets for cancer and fibrosis. However, the functional mechanisms of

galectins remain unclear in many cases. This is true in part because they are present both inside and outside the cell and display an intriguing combination of intra- and extracellular activities. Furthermore, they can modulate the action of different glycoproteins through recognizing specific glycan motifs of the glycoproteins, with the added complexicity of multivalency and aggregation. Therefore, selective and potent inhibitors for galectins would be valuable tools to investigate the biological functions of these proteins.

In this thesis, we describe the design and synthesis of specific galectin inhibitors through a strategy of multivalency. Chapter 3 describes a synthesis of multivalent ligands inhibiting Gal-3, in which multiple lysine residues of albumin were chemically conjugated with “thiodigalactoside” epitopes to make neo-glycoproteins. The neo-glycoproteins were finally evaluated in binding studies with human Gal-1 and -3 to determine the binding properties.

Chapter 4 describes the synthesis of galectin inhibitors with a potential “chelate effect”, which are designed to bind to the two different binding sites on galectins simultaneously. In this chapter, a series of asymmetric “hybrid” compounds were prepared, that combine two galectin ligands, i.e., thiodigalactoside derivatives and calixarene 0118. Moreover, NMR spectroscopy was used to evaluate the interactions of these compounds with Gal-1 or -3. In addition, cellular experiments were conducted to compare the cytotoxic effects of the hybrids with those of calixarene 0118.

1.6 References

(1) Kilpatrick, D. C. Animal Lectins: A Historical Introduction and Overview. *Biochim. Biophys. Acta* **2002**, *1572*, 187–197.

(2) Cooper, N. W. D. Galectinomics: Finding Themes in Complexity. *Biochim. Biophys. Acta - Gen. Subj.* **2002**, *1572* (2–3), 209–231.

(3) Dagher, S. F.; Wangt, J. L.; Patfrersont, R. J. Identification of Galectin-3 as a Factor in Pre-mRNA Splicing. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 1213–1217.

(4) Wang, J. L.; Gray, R. M.; Haudek, K. C.; Patterson, R. J. Nucleocytoplasmic Lectins. *Biochim. Biophys. Acta - Gen. Subj.* **2004**, *1673*, 75–93.

(5) Davidson, P. J.; Davis, M. J.; Patterson, R. J.; Ripoche, M.; Poirier, F.; Wang, J. L. Shuttling of Galectin-3 between the Nucleus and Cytoplasm. *Glycobiology* **2002**, *12* (5), 329–337.

(6) Leffler, H.; Carlsson, S.; Hedlund, M.; Qian, Y.; Poirier, F. Introduction to Galectins. *Glycoconj. J.* **2002**, *19* (7–9), 433–440.

(7) Knibbs, R. N.; Agrwal, N.; Wang, J. L.; Goldstein, I. J. Carbohydrate-Binding Protein 35: II. Analysis of the Interaction of the Recombinant Polypeptide with Saccharides. *J. Biol. Chem.* **1993**, *268* (20), 14940–14947.

(8) Salomonsson, E.; Carlsson, M. C.; Osla, V.; Hendus-Altenburger, R.; Kahl-Knutson, B.; Öberg, C. T.; Sundin, A.; Nilsson, R.; Nordberg-Karlsson, E.; Nilsson, U. J.; Karlsson, A.; Rini, J. M.; Leffler, H. Mutational Tuning of Galectin-3 Specificity and Biological Function. *J. Biol. Chem.* **2010**, *285* (45), 35079–35091.

(9) Seetharaman, J.; Kanigsberg, A.; Slaaby, R.; Leffler, H.; Barondes, S. H.; Rini, J. M. X-Ray Crystal Structure of the Human Galectin-3 Carbohydrate Recognition Domain at 2.1-Angstrom Resolution. *J. Biol. Chem.* **1998**, *273* (21), 13047–13052.

(10) Tejler, J.; Salameh, B.; Leffler, H.; Nilsson, U. J. Fragment-Based Development of Triazole-Substituted O-Galactosyl Aldoximes with Fragment-Induced Affinity and Selectivity for Galectin-3. *Org. Biomol. Chem.* **2009**, *7* (19), 3982–3990.

(11) Sörme, P.; Arnoux, P.; Kahl-Knutsson, B.; Leffler, H.; Rini, J. M.; Nilsson, U. J. Structural and Thermodynamic Studies on Cation-II Interactions in Lectin-Ligand Complexes: High-Affinity Galectin-3 Inhibitors through Fine-Tuning of an Arginine-Arene Interaction. *J. Am. Chem. Soc.* **2005**, *127* (6), 1737–1743.

(12) Giguère, D.; Patnam, R.; Bellefleur, M.-A.; St-Pierre, C.; Sato, S.; Roy, R. Carbohydrate Triazoles and Isoxazoles as Inhibitors of Galectins-1 and -3. *Chem. Commun.* **2006**, No. 22, 2379–2381.

(13) Cumpstey, I.; Sundin, A.; Leffler, H.; Nilsson, U. J. C2-Symmetrical Thiodigalactoside Bis-Benzamido Derivatives as High-Affinity Inhibitors of Galectin-3: Efficient Lectin Inhibition through Double Arginine-Arene Interactions. *Angew. Chemie - Int. Ed.* **2005**, *44* (32), 5110–5112.

(14) Sörme, P.; Qian, Y.; Nyholm, P. G.; Leffler, H.; Nilsson, U. J. Low Micromolar Inhibitors of Galectin-3 Based on 3'-Derivatization of N-Acetyllactosamine. *ChemBioChem* **2002**, *3* (2–3), 183–189.

(15) Leffler, H.; Barondes, S. H. Specificity of Binding of Three Soluble Rat Lung Lectins to Substituted and Unsubstituted Mammalian β-Galactosides. *J. Biol. Chem.* **1986**, *261* (22), 10119–10126.

(16) Campo, V. L.; Marchiori, M. F.; Rodrigues, L. C.; Dias-Baruffi, M. Synthetic Glycoconjugates Inhibitors of Tumor-Related Galectin-3: An Update. *Glycoconj. J.* **2016**, *33* (6), 853–876.

(17) Salameh, B. A.; Cumpstey, I.; Sundin, A.; Leffler, H.; Nilsson, U. J. 1H-1,2,3-Triazol-1-Yl Thiodigalactoside Derivatives as High Affinity Galectin-3 Inhibitors. *Bioorganic Med. Chem.* **2010**, *18* (14), 5367–5378.

(18) Leffler, H.; Salameh, B. A.; Nilsson, U. 3-Triazolyl-Galatoside Inhibitors of Galectins. *U.S. Pat.* **2010**, *7,700,763*.

(19) Nilsson, U.; Leffler, H.; Mukhopadhyay, B.; Bengal, K. W.; Rajput Vishal. Novel Galectoside Inhibitors of Galectins. *U.S. Pat.* **2014**, *0,336,146*.

(20) Nilsson, U.; Leffler, H. Galactoside Inhibitor of Galectin-3 and Its Use for Treating Pulmonary Fibrosis. *C.A. Pat.* **2015**, *104755088*.

(21) Hsieh, T.; Lin, H.; Tu, Z.; Lin, T.; Wu, S. Dual Thio-Digalactoside-Binding Modes of Human Galectins as the Structural Basis for the Design of Potent and Selective Inhibitors. *Sci. Rep.* **2016**, No. July, 1–9.

(22) Hattum, H. Van; Branderhorst, H. M.; Moret, E. E.; Nilsson, U. J.; Leffler, H.; Pieters, R. J. Tuning the Preference of Thiodigalactoside- and Lactosamine-Based Ligands to Galectin-3 over Galectin-1. *J. Med. Chem.* **2013**, *56* (3), 1350–1354.

(23) Giguère, D.; Bonin, M. A.; Cloutier, P.; Patnam, R.; St-Pierre, C.; Sato, S.; Roy, R. Synthesis of Stable and Selective Inhibitors of Human Galectins-1 and -3. *Bioorganic Med. Chem.* **2008**, *16* (16), 7811–7823.

(24) Collins, P. M.; Öberg, C. T.; Leffler, H.; Nilsson, U. J.; Blanchard, H. Taloside Inhibitors of Galectin-1 and Galectin-3. *Chem. Biol. Drug Des.* **2012**, *79* (3), 339–346.

(25) Öberg, C. T.; Blanchard, H.; Leffler, H.; Nilsson, U. J. Protein Subtype-Targeting through Ligand Epimerization: Talose-Selectivity of Galectin-4 and Galectin-8. *Bioorganic Med. Chem. Lett.* **2008**, *18* (13), 3691–3694.

(26) Slovin, S. F.; Keding, S. J.; Ragupathi, G. Carbohydrate Vaccines as Immunotherapy for Cancer. *Immunol. Cell Biol.* **2005**, *83*, 418–428.

(27) Simon-haldi, M.; Mantei, N.; Franke, J.; Voshol, H.; Schachner, M. Identification of a Peptide Mimic of the L2 / HNK-1 Carbohydrate Epitope. *J. Neurochem.* **2002**, *83* (6), 1380–1388.

(28) Matsubara, T. Potential of Peptides as Inhibitors and Mimotopes: Selection of Carbohydrate-Mimetic Peptides from Phage Display Libraries. *J. Nucleic Acids* **2012**, *12* (ID740982), 15.

(29) Valentini, P.; Lensch, M.; Arnusch, C. J.; Andre, S.; Russwurm, R.; Siebert, H.; Fischer, M. J. E.; Gabius, H.; Pieters, R. J. Interference of the Galactose-Dependent Binding of Lectins by Novel Pentapeptide Ligands. *Bioorg. Med. Chem.* **2004**, *14*, 1437–1440.

(30) Arnusch, C. J.; Kuwabara, I.; Russwurm, R.; Kaltner, H.; Gabius, H.; Pieters, R. J. Identification of Peptide Ligands for Malignancy- and Growth-Regulating Galectins Using Random Phage-Display and Designed Combinatorial Peptide Libraries. *Bioorg. Med. Chem.* **2005**, *13*, 563–573.

(31) Burke, S. D.; Zhao, Q.; Schuster, M. C.; Kiessling, L. L.; January, R. V. Synergistic Formation of Soluble Lectin Clusters by a Templated Multivalent Saccharide Ligand. *J. Am. Chem. Soc.* **2000**, *122* (27), 4518–4519.

(32) Zou, J.; Glinsky, V. V; Landon, L. A.; Matthews, L.; Deutscher, S. L. Peptides Specific to the Galectin-3 Carbohydrate Recognition Domain Inhibit Metastasis-Associated Cancer Cell Adhesion. *Carcinogenesis* **2005**, *26* (2), 309–318.

(33) Salomonsson, E.; Thijssen, V. L.; Griffioen, A. W.; Nilsson, U. J.; Leffler, H. The Anti-Angiogenic Peptide Anginex Greatly Enhances Galectin-1 Binding Affinity for Glycoproteins. *J. Biol. Chem.* **2011**, *286* (16), 13801–13804.

(34) Gray, B. H.; Haseman, J. R.; Mayo, K. H. B/PI-Derived Synthetic Peptides: Synergistic Effects in Tethered Bactericidal and Endotoxin Neutralizing Peptides. *Biochim. Biophys. Acta* **1995**, *1244*, 185–190.

(35) Mayo, K. H.; Haseman, J.; Ilyina, E.; Gray, B. Designed β-Sheet-Forming Peptide 33mers with Potent Human Bactericidal/permeability Increasing Protein-like Bactericidal and Endotoxin Neutralizing Activities. *Biochim. Biophys. Acta - Gen. Subj.* **1998**, *1425* (1), 81–92.

(36) Mayo, K. H.; van der Schaft, D. W.; Griffioen, a W. Designed Beta-Sheet Peptides That Inhibit Proliferation and Induce Apoptosis in Endothelial Cells. *Angiogenesis* **2001**, *4* (1), 45–51.

(37) Thijssen, V. L. J. L.; Postel, R.; Brandwijk, R. J. M. G. E.; Dings, R. P. M.; Nesmelova, I.; Satijn, S.; Verhofstad, N.; Nakabeppu, Y.; Baum, L. G.; Bakkers, J.; Mayo, K. H.; Poirier, F.; Griffioen, A. W. Galectin-1 Is Essential in Tumor Angiogenesis and Is a Target for Antiangiogenesis Therapy. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (43), 15975–15980.

(38) Dings, R. P. M.; Mayo, K. H. A Journey in Structure-Based Drug Discovery: From Designed Peptides to Protein Surface Topomimetics as Antibiotic and Antiangiogenic Agents. *Acc. Chem. Res.* **2007**, *40* (10), 1057–1065.

(39) Mayo, K. H.; Dings, R. P. M.; Flader, C.; Nesmelova, I.; Hargittai, B.; Van Der Schaft, D. W. J.; Van Eijk, L. I.; Walek, D.; Haseman, J.; Hoye, T. R.; Griffioen, A. W. Design of a Partial Peptide Mimetic of Anginex with Antiangiogenic and Anticancer Activity. *J. Biol. Chem.* **2003**, *278* (46), 45746–45752.

(40) Dings, R. P. M.; Kumar, N.; Miller, M. C.; Loren, M.; Rangwala, H.; Hoye, T. R.; Mayo, K. H. Structure-Based Optimization of Angiostatic Agent 6DBF7, an Allosteric Antagonist of Galectin-1. *J. Pharmacol. Exp. Ther.* **2013**, *344* (3), 589–599.

(41) Dings, R. P. M.; Chen, X.; Hellebrekers, D. M. E. I.; van Eijk, L. I.; Zhang, Y.; Hoye, T. R.; Griffioen, A. W.; Mayo, K. H. Design of Nonpeptidic Topomimetics of Antiangiogenic Proteins with Antitumor Activities. *J. Natl. Cancer Inst.* **2006**, *98* (13), 932–936.

(42) Dings, R. P. M.; Miller, M. C.; Nesmelova, I.; Astorgues-Xerri, L.; Kumar, N.; Serova, M.; Chen, X.; Raymond, E.; Hoye, T. R.; Mayo, K. H. Antitumor Agent Calixarene 0118 Targets Human Galectin-1 as an Allosteric Inhibitor of Carbohydrate Binding. *J. Med. Chem.* **2012**, *55* (11), 5121–5129.

(43) Page, M. I.; Jencks, W. P. Entropic Contributions to Rate Accelerations in Enzymic and Intramolecular Reactions and the Chelate Effect. *Proc. Natl. Acad. Sci. U. S. A.* **1971**, *68* (8), 1678–1683.

(44) Mammen, M.; Choi, S.; Whitesides, G. M. Polyvalent Interactions in Biological Systems : Implications for Design and Use of Multivalent Ligands and Inhibitors. *Angew. Chem. Int. Ed.* **1998**, *37*, 2754–2794.

(45) André, S.; Liu, B.; Gabius, H.; Roy, R. First Demonstration of Differential Inhibition of Lectin Binding by Synthetic Tri- and Tetravalent Glycoclusters from Cross-Coupling of Rigidified 2-Propynyl Lactoside. *Org. Biomol. Chem.* **2003**, *1* (22), 3909–3916.

(46) Tejler, J.; Tullberg, E.; Nilsson, U. J. Synthesis of Multivalent Lactose Derivatives by 1,3-Dipolar Cycloadditions: Selective Galectin-1 Inhibition. *Carbohydr. Res.* **2006**, *341*, 1353–1362.

(47) Marchiori, M. F.; Souto, D. E. P.; Bortot, L. O.; Pereira, J. F.; Kubota, L. T.; Cummings, R. D.; Dias-baruffi, M.; Carvalho, I.; Leiria, V. L. Synthetic 1,2,3-Triazole-Linked Glycoconjugates Bind with High Affinity to Human Galectin-3. *Bioorg. Med. Chem.* **2015**, *23* (13), 3414–3425.

(48) Vrasidas, I.; André, S.; Valentini, P.; Böck, C.; Lensch, M.; Kaltner, H.; Liskamp, R. M. J.; Gabius, H.; Pieters, R. J. Rigidified Multivalent Lactose Molecules and Their Interactions with Mammalian Galectins : A Route to Selective Inhibitors. *Org. Biomol. Chem.* **2003**, *1*, 803–810.

(49) Böcker, S.; Laaf, D.; Elling, L. Galectin Binding to Neo-Glycoproteins: LacDiNAc Conjugated BSA as Ligand

for Human Galectin-3. *Biomolecules* **2015**, No. 5, 1671–1696.

(50) Platt, D.; Klyosov, A. Selectively Depolymerized Galactomannan Polysaccharide. *U.S. Pat.* **2011**, *0,077,217*.

(51) Wang, H.; Huang, W.; Orwenyo, J.; Banerjee, A.; Vasta, G. R.; Wang, L. Design and Synthesis of Glycoprotein-Based Multivalent Glyco-Ligands for Infuenza Hemagglutinin and Human Galectin-3. *Bioorganic Med. Chem.* **2013**, *21* (7), 2037–2044.

(52) Seko, A.; Koketsu, M.; Nishizono, M.; Enoki, Y.; Ibrahim, H. R.; Juneja, L. R.; Kim, M.; Yamamoto, T. Occurrence of a Sialylglycopeptide and Free Sialylglycans in Hen's Egg Yolk. *Biochim. Biophys. Acta* **1997**, *1335*, 23–32.

(53) Miller, M. C.; Klyosov, A.; Mayo, K. H. The α-Galactomannan Davanat Binds Galectin-1 at a Site Different from the Conventional Galectin Carbohydrate Binding Domain. *Glycobiology* **2009**, *19* (9), 1034–1045.

(54) Miller, M. C.; Ippel, H.; Suylen, D.; Klyosov, A. A.; Traber, P. G.; Hackeng, T.; Mayo, K. H. Binding of Polysaccharides to Human Galectin-3 at a Noncanonical Site in Its Carbohydrate Recognition Domain. *Glycobiology* **2016**, *26* (1), 88–99.

(55) Miller, M. C.; Klyosov, A. A.; Mayo, K. H. Structural Features for α-Galactomannan Binding to Galectin-1. *Glycobiology* **2012**, *22* (4), 543–551.

(56) Klyosov, A.; Platt, D. Co-Administration of a Polysaccharide with a Chemotherapeutic Agent for the Treatment of Cancer. *U.S. Pat.* **2006**, *7,012,068*.

(57) Traber, P. G.; Chou, H.; Zomer, E.; Hong, F.; Klyosov, A.; Fiel, M.; Friedman, S. L. Regression of Fibrosis and Reversal of Cirrhosis in Rats by Galectin Inhibitors in Thioacetamide-Induced Liver Disease. *PLoS One* **2013**, *8* (10), e75361.

(58) Traber, P. G.; Zomer, E. Therapy of Experimental NASH and Fibrosis with Galectin Inhibitors. *PLoS One* **2013**, *8* (12), e83481.

(59) Miller, M. C.; Ribeiro, J. P.; Roldós, V.; Martín-santamaría, S.; Cañada, F. J.; Nesmelova, A.; André, S.; Pang, M.; Klyosov, A.; Baum, L. G.; Jiménez-barbero, J. Structural Aspects of Binding of α-Linked Digalactosides to Human Galectin-1. *Glycobiology* **2011**, *21* (12), 1627–1641.

(60) Miller, M. C.; Nesmelova, I. V; Platt, D.; Klyosov, A.; Mayo, K. H. The Carbohydrate-Binding Domain on Galectin-1 Is More Extensive for a Complex Glycan than for Simple Saccharides : Implications for Galectin-Glycan Interactions at the Cell Surface. *Biochem. J.* **2009**, *421*, 211–221.

Chapter 2

General introduction II

2.1 Carbohydrate and protein glycosylation

Glycosylation is the process by which a carbohydrate, i.e. a glycosyl donor, is covalently linked to a target macromolecule (a glycosyl acceptor). Protein glycosylation in biology mainly refers to an enzyme-directed process that attaches glycans to proteins, and the resulting modified proteins can exhibit different functions and subsequently be involved in numerous biological processes. One of the most fascinating and yet frustrating features of protein glycosylation is the variety of the attached glycans, which is the so called “microheterogeneity”. Thus, glycosylation can turn a single protein into distinct “glycoforms”. Mechanistically, this “microheterogeneity” phenomenon might be the result of the complex microenvironment, i.e. the endoplasmic reticulum (ER) and Golgi apparatus, in which multiple, sequential, and partially competitive glycosylation and deglycosylation processes go simultaneously.

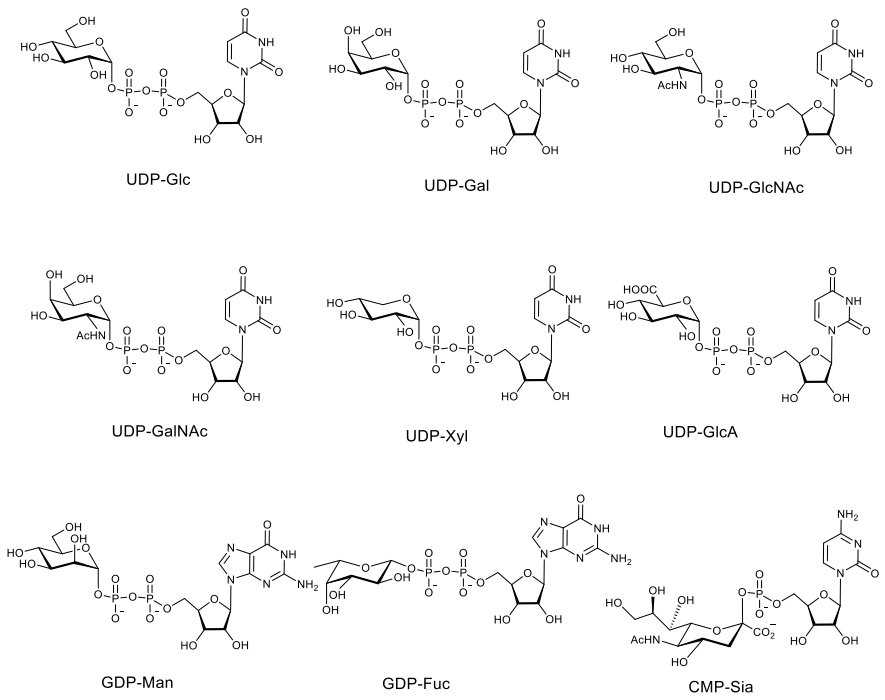


Figure 1. Sugar nucleotide donors used by mammals.

The protein glycosylation process itself has several common features. First of all, the glycosylation usually uses an activated form of the carbohydrate (nucleotide sugars) as a sugar donor for the reaction and the conversions are catalyzed by glycosyltransferases (GTs). These sugar donors are synthesized within the cytosolic or nuclear compartment

from monosaccharide precursors of endogenous or exogenous origin. Among an extensive range of sugar donors present in nature, only 9 sugar nucleotide donors are used by mammals (Figure 1). Secondly, protein glycosylation is a site-specific modification, which means the glycans can be linked to the protein in specific manners providing *N*-, *O*- or *C*-linked glycans.

2.2 Carbohydrate and glycosyltransferase

Carbohydrate moieties on a glycosylated protein act as words of a cellular language and their chemical diversity governs the corresponding functions of the resulting protein. Considering the stereochemical complexity of carbohydrates, glycosyltransferases need to transfer the sugar in a regiospecific and stereospecific way, which can involve inversion or retention of the configuration at the anomeric center (Figure 2).

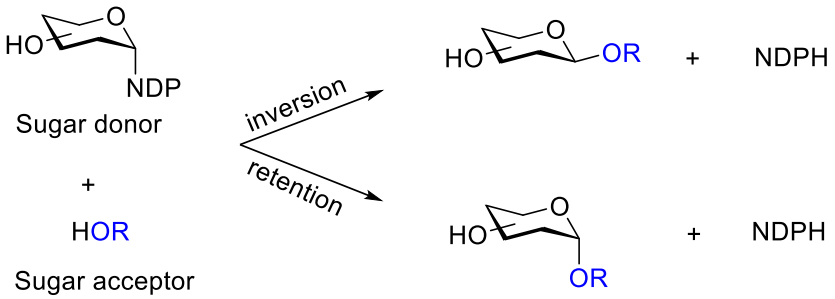


Figure 2. Enzymatic reaction with either an inverting or a retaining glycosyltransferase.

The mechanism of inverting glycosyltransferases is suggested to be a direct S_N2-like displacement via an oxocarbenium-ion transition state assisted by a catalytic base. This is exemplified in the research about O-linked β-N-acetylglucosamine transferase (OGT) undertaken by Walker and co-workers. Their crystal structure of human OGT containing sugar donor UDP-N-acetylglucosamine (UDP-GlcNAc) and a 14-residue peptide substrate as the sugar acceptor, showed OGT catalyzes the transfer with inversion of anomeric stereochemistry. [1]

The mechanism for retaining GTs is controversial but there are two main mechanism proposed: double-displacement ($2\times S_N2$) and internal return (S_Ni). Double displacement mechanism involves formation of a covalent glycosyl-enzyme intermediate and two S_N2 reactions in total. The first nucleophilic attack produces the glycosyl-enzyme intermediate with inversion of configuration on the anomeric center of the sugar donor. In the second step, the intermediate is attacked by a hydroxyl group of sugar acceptor, which is assisted by a catalytic base and results in an overall retention of configuration. For the S_Ni

mechanism, the GT interacts with the sugar donor and produces an oxocarbenium ion to shield on one side of the anomeric center. On the opposite side, the nearby acceptor is deprotonated by the phosphate group, then attach to the anomeric center after the phosphate leaves. This results in retention of configuration.[2]

2.3 Glycosyltransferase families

GTs are classified into families based on amino acid sequence similarities. Since Campbell and co-workers first started the classification of GTs in 1997, over 90 distinct GTs families have been found. [2,3] To date, X-ray crystal structures have been published for over 100 GTs in 38 GT families, which have been classified into three folds called GT-A, GT-B and GT-C. The GTs protein topology in GT-A fold consists of two associated β/α/β Rossmann domains and was first described for the SpsA from *Bacillus subtilis*.[4] Like the GT-A fold, the architecture of GT-B enzymes consists of two β/α/β Rossmann domains, but these are associated flexibly and face "each other". The first GT-B three dimensional structure was reported in 1994 for a bacteriophage T4 β-glucosyltransferase, which can transfer a glucose from uridine diphosphoglucose to hydroxymethyl groups of modified cytosine bases in DNA.[4,5] The third fold, GT-C, was predicted on the basis of iterative sequence searches but the first three dimensional structure of a GT-C enzyme wasn't published until 2008. This was STT3 from *Pyrococcus furiosus*, which can catalyze the transfer of a heptasaccharide onto a peptide in an Asn-X-Thr/Ser-motif-dependent manner.[6] The STT3 crystal structure consists of four structural domains, the mainly α-helical domain of which is located at the center and surrounded by the other three β-sheet rich domains. It should be noted that not all the enzymes adopting GT-A or GT-B folds are glycosyltransferases and there is no relation between the overall fold of an enzyme and its biological function.

2.4 O-linked N-acetylglucosamine transferase

Post-translational modifications, such as phosphorylation, acetylation or methylation, regulate protein function in cellular processes. O-GlcNAcylation as a widespread form of post-translational modification has been found on hundreds of cytosolic and nuclear proteins. Remarkably, the O-GlcNAc modification is regulated by only two opposing enzymes. The enzyme catalyzing the attachment of N-acetylglucosamine onto serines/threonines of the protein is O-linked N-acetylglucosamine transferase (OGT), and the enzyme removing the O-GlcNAc from the modified protein is O-GlcNAcase (OGA). More than 1,000 substrate proteins that are O-GlcNAcylated by OGT have been identified to date and therefore OGT is involved in numerous biological processes such as transcription, cell cycle progression, the stress response and nutrient sensing.

Structurally OGT contains the GT-B fold and it comprises two distinct regions: an N-terminal region consisting of tetratricopeptide (TPR) repeat units and a catalytic region containing three domains: the amino (N)-terminal domain (N-Cat), the carboxy (C)-terminal domain (C-Cat), and the intervening domain (Int-D) (Figure 3A).[8,9] The UDP

moiety lies in a pocket which is in the C-Cat domain near a cleft between the C-Cat and N-Cat domain. Besides, the crystal structure of a complex containing UDP and a peptide substrate has been published, which provides insight into the catalytical mechanism of OGT.[8] The OGT-UDP-peptide complex is shown in Figure 3B. The structure and a competitive inhibition experiment indicate an action mechanism of OGT, in which the donor substrate UDP-GlcNAc binds to the active site of the enzyme first, followed by the acceptor peptide of the protein, and then OGT transfers GlcNAc from the donor onto the peptide (Figure 3D). In this transfer, several residues around the substrate are critical for catalytic activity (Figure 3C). For example, Lys 842 stabilizes the UDP-GlcNAc by contacts with the phosphate of UDP, and His 498 locating between the reactive serine hydroxyl and the GlcNAc binding pocket was proposed to catalyze the nucleophilic attack as a base.[10,11]

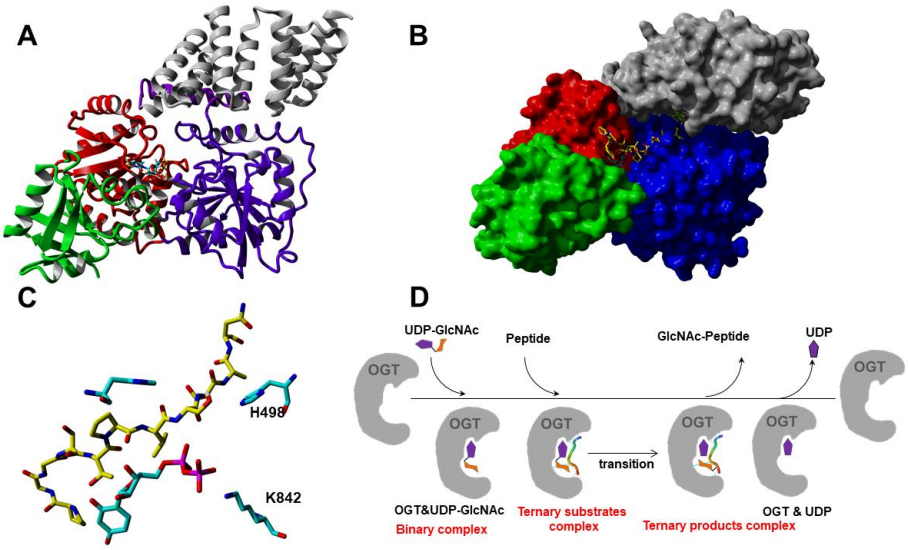


Figure 3. (A) Schematic of the OGT domain architecture with the TPR units shown in grey, the transitional helix (H3) in purple, the N-Cat domain in blue, the Int-D domain in green, and the C-Cat domain in red; (B) OGT-UDP-peptide complex; (C) View of UDP and key residues (carbon atoms shown in cyan) and part of the CKII peptide (carbon atoms shown in yellow); (D) Proposed mechanism of OGT.

2.5 O-GlcNAc transferase inhibitors

Much research has revealed that OGT appears to be a therapeutic target in several human pathologies. This is especially the case in diabetes and cancer because of their strong connections with the metabolic status of the cell (Figure 4). Like kinases that use ATP as a substrate in the phosphorylation, OGT uses UDP-GlcNAc as the sugar donor substrate, which is a product of hexosamine biosynthetic pathway (HBP). Nutrient

variations change the flux through the HBP either increasing or decreasing UDP-GlcNAc levels, affecting the O-GlcNAcylation of many proteins, since OGT is highly sensitive to UDP-GlcNAc levels across a very broad range of concentrations. Through this mechanism, OGT overexpression along with excessive nutritional intake (glucose and glutamine), an important signal of diabetes and cancer, and the resulting O-GlcNAcylated proteins are involved in different diseases. Besides, the N-terminal domain consisting of tetratricopeptide (TPR) units can mediate the recognition of proteins, which implies a role for substrate selection via protein-protein interactions. With the catalytic and non-catalytic activities, OGT is emerging as a topic of current interest, and selective inhibitors targeting OGT are needed to identify all of its functions in biological processes and validate OGT as a therapeutic target. Based on their different structural features, inhibitors can be classified into three types: sugar donor mimic inhibitors, bi-substrate inhibitors and small molecule inhibitors identified from screening (Figure 5).

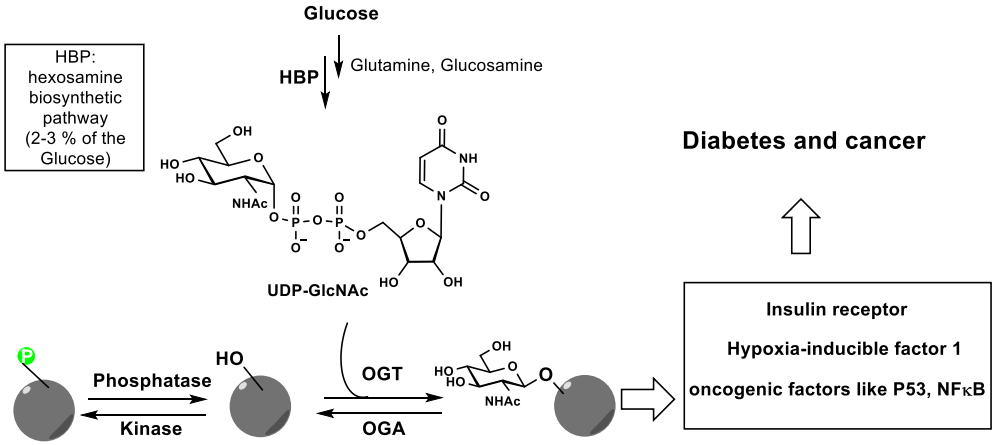


Figure 4. OGT as a therapeutic target linking glucose metabolism to protein O-GlcNAcylation in different diseases.

2.5.1 Sugar donor mimic inhibitors

UDP-GlcNAc plays a role as a sugar donor in the process of O-GlcNAcylation. Interestingly, UDP itself can inhibit OGT *in vitro* with an IC_{50} of 1.8 μM, making it the most potent OGT inhibitor to date (Figure 5). With this interesting finding, some OGT inhibitors have been identified from UDP analogues. For example, C-UDP exhibited a high potency like UDP but the two compounds are not suitable for use in cell biology or medicine because of poor cell permeability and selectivity issues. Besides, various small molecules derived from moieties of the sugar donor UDP-GlcNAc are reported to perturb

O-GlcNAcylation in cells, including alloxan and benzyl 2-acetamido-2-deoxy-α-D-galactopyranoside (BAGDP). The uracil analogue alloxan was the first OGT inhibitor reported and has been used in cell biology research (IC_{50} = 18 μM).[12] However, its use is restricted by many off-target effects and cellular toxicity due to its oxidative capacity.[13,14] BADGP, as an *N*-acetylgalactosamine (GalNAc) mimic, is capable of changing the O-glycosylation in cells but is not selective for OGT.[15]

Modification of UDP-GlcNAc produced several OGT inhibitors that are less potent than UDP.[16] Ac-5S-GlcNAc as a precursor can be administered to a cell to hijack the hexosamine pathway towards the production of UDP-5S-GlcNAc. After enzymatic processing, this compound decreased global *O*-GlcNAcylation within cells with an EC_{50} of 5 μM and has been used in a number of biological studies.[17,18] Nevertheless, UDP-5S-GlcNAc converted from Ac-5S-GlcNAc, is an isostere of UDP-GlcNAc, and may inhibit not only OGT but also other UDP-GlcNAc-dependent enzymes in the cell. UDP-5S-GlcNAc can e.g. be epimerized to UDP-5S-GalNAc and affect the corresponding enzyme that uses UDP-GalNAc.[17,19] To summarize, although these inhibitors derived from the sugar donor show moderate potency and some of them have been used in research, the drawback is obvious due to their limited potency and nonnegligible off-target effect in cell biology experiments. Not only does the charged nature of UDP make it cell-impermeant, but also it also binds to a wide range of other enzymes.

2.5.2 Bisubstrate inhibitors

A selective OGT inhibitor based on just the sugar donor is difficult to obtain because GTs usually use a common glycosyl donor. On the other hand, acceptor analogues have the potential to be selective inhibitors because of their more specific binding to the target, albeit with a lower affinity indicated by typical K_m values in the millimolar or high micromolar range. In comparison with the acceptor, donor affinities are usually in the low micromolar range.[20,21] Bisubstrate analogues, in which donor and acceptor analogues are covalently attached to each other, have the potential to offer both potent and selective inhibition.[22]

OGT bisubstrate inhibitors derived from both the OGT donor and acceptor substrate were published in 2014 for the first time, which are shown in Figure 5. These two compounds (Goblin 1 and Goblin 2) consist of the UDP part of the sugar donor and in addition a sugar acceptor peptide, connected via a short linker.[23] It should be noted here that the OGT inhibition of Goblin 1 and 2 is almost entirely the result of the UDP part and the peptide part had no detectable independent binding potency to OGT. In other words, the donor and acceptor moiety in the Goblins do not synergistically contribute to the OGT inhibition as expected from a proper bisubstrate inhibitor. Besides, the specificity of Goblin towards OGT needs to be validated and poor cell permeability restricted their applications in cell biology experiments.

2.5.3 Small molecule inhibitors from high-throughput screening (HTS)

High-throughput screening (HTS) approaches have also been used to identify small molecule OGT inhibitors. Several OGT inhibitors with moderate potency have been reported.[24,25] A fluorescent UDP-GlcNAc displacement assay was designed to screen library compounds, and then several positive hits were evaluated for OGT inhibition with a radiometric assay. IC_{50} values and the mode of inhibition were determined for two of the best: ST045849 (IC_{50} = 53 μM) and BZX (IC_{50} = 10 μM); both were competitive with the UDP-GlcNAc.[25] A study showed that ST045849 perturbed pancreatic development to reduce the insulin production, and this finding indicated the important and specific role of OGT in β-cell development.[26] Another recent report showed ST045849 as an OGT inhibitor blocking O-GlcNAcylated Notch1 and Notch1 protein expression in the plasma membrane.[27] BZX contains a dicarbamate moiety that connects two aromatic groups and it was used to treat breast cancer cells by Caldwell and co-workers. In their research, the resulting reduced O-GlcNAcylation by OGT inhibition caused shrinkage and an anti-invasion effect of the cancer cells, which suggested that OGT could represent novel therapeutic targets for breast cancer.[28] One unanticipated finding about BZX is that it can covalently bind OGT to cause an irreversible loss of enzyme activity. Mass spectrometry and X-ray crystallography have shown the five heteroatom dicarbamate core in BZX cross-links the active site of OGT (Lys 842 and Cys 917) through a double-displacement mechanism. While the inhibiting mechanism of BZX provides new insight into the unique architecture of the OGT active site, it may produce potential toxic and off-target effect.[29]

Another cell-permeable small molecule OGT inhibitor, OSMI-1, has been identified through a combination of HTS hits and follow-up chemistry. This compound inhibits OGT with an IC_{50} of 2.7 μM and has been tested in several mammalian cell lines.[19] Compared with other OGT inhibitors, OSMI-1 is noncompetitive with respect to UDP-GlcNAc, which means it binds to a different site of OGT and likely adopts a different mode of action. Similar inhibitors have been discovered via “tethering in situ click chemistry (TISCC)” in very recent research. APNT and APBT inhibit OGT with IC_{50} of 66 μM and 139 μM and also inhibit O-GlcNAcylation in cells without competing with donor or acceptor substrates of OGT.[30] More research is necessary to understand how these compounds inhibit OGT and such newly discovered information can then be used to design novel OGT inhibitors and better understand the regulatory mechanisms and roles of OGT.

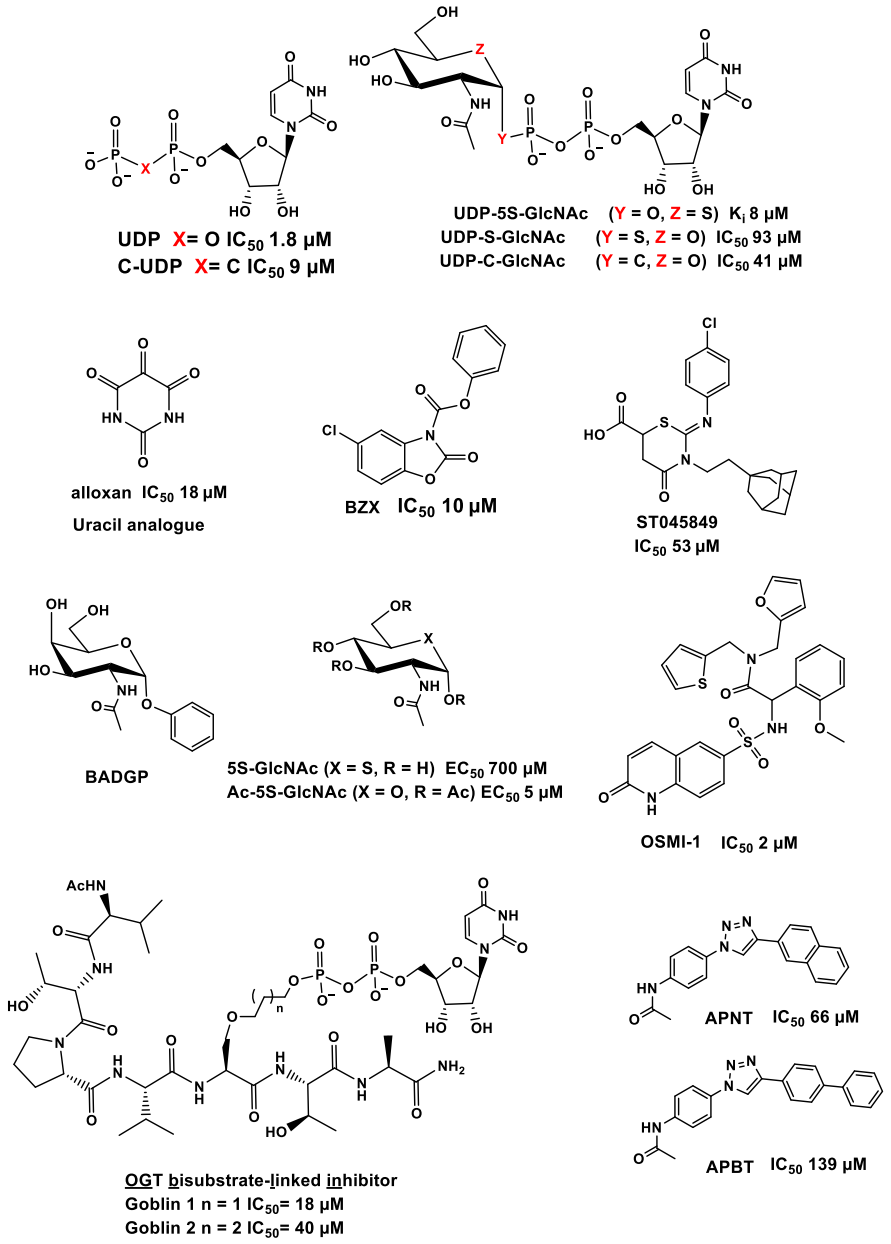


Figure 5. OGT inhibitors discussed in the introduction.

2.6 Project aim

O-GlcNAc transferase (OGT) catalyzes the glycosylation of cytosolic and nuclear proteins and several compounds that manipulate OGT activity have been discovered to be used as tools to better understand its function. Further research on some OGT inhibitor

designs is limited by their diphosphate moiety that strongly contributes to the binding but results in poor uptake properties and/or selectivity. We recently identified two peptides (RBL-2 and ZO-3) as OGT sugar acceptor substrates through a peptide microarray approach. By making peptide variants, we successfully identified their O-GlcNAc sites and optimized the sequence of the peptides. Inspired by these findings, we obtained two potent OGT peptide inhibitors and thus started to design and synthesize potential bisubstrate inhibitors. Furthermore, one approach of conjugating a peptide derived from RBL-2 to uridine was performed as described in Chapter 5 of this thesis. In parallel, compounds obtained from a virtual screening were introduced as a replacement of uridine to make a novel OGT bisubstrate inhibitor.

2.7 References

(1) Pathak, S.; Alonso, J.; Schimpl, M.; Rafie, K.; Blair, D. E.; Borodkin, V. S.; Schüttelkopf, A. W.; Albarbarawi, O.; van Aalten, D. M. F. The Active Site of O-GlcNAc Transferase Imposes Constraints on Substrate Sequence. *Nat. Struct. Mol. Biol.* **2015**, *22* (9), 744–750.

(2) Breton, C.; Fournel-gigleux, S.; Palcic, M. M. Recent Structures , Evolution and Mechanisms of Glycosyltransferases. *Curr. Opin. Struct. Biol.* **2012**, *22* (5), 540–549.

(3) Campbell, J.; Davies, D.; Bulone, V.; Henrissat, B. A Classification of Nucleotide-Diphospho-Sugar Glycosyltransferases Based on Amino Acid Sequence Similarities. *Biochem. J.* **1997**, *326*, 929–942.

(4) Charnock, S. J.; Davies, G. J. Structure of the Nucleotide-Diphospho-Sugar Transferase, SpsA from Bacillus Subtilis,in Native and Nucleotide-Complexed Forms. *Biochemistry* **1999**, *38* (20), 6380–6385.

(5) Vrielinkl, A.; Ruger, W.; Driessen, H. P. C.; Freemont, P. S. Crystal Structure of the DNA Modifying Enzyme: β-Glucosyltransferase in the Presence and Absence of the Substrate Uridine Diphosphoglucose. *EMBO J.* **1994**, *13* (15), 3413–3422.

(6) Roychoudhury, R.; Pohl, N. L. B. New Structures, Chemical Functions, and Inhibitors for Glycosyltransferases. *Curr. Opin. Chem. Biol.* **2010**, *14* (2), 168–173.

(7) Birch, H. L.; Alderwick, L. J.; Rittmann, D.; Krumbach, K.; Etterich, H.; Grzegorzewicz, A.; Mcneil, M. R.; Eggeling, L.; Besra, G. S.; Al, B. E. T.; Acteriol, J. B. Identification of a Terminal Rhamnopyranosyltransferase (RptA) Involved in Corynebacterium Glutamicum Cell Wall Biosynthesis. *J. Bacteriol.* **2009**, *191* (15), 4879–4887.

(8) Lazarus, M. B.; Nam, Y.; Jiang, J.; Sliz, P.; Walker, S. Structure of Human O-GlcNAc Transferase and Its Complex with a Peptide Substrate. *Nature* **2011**, *469* (7331), 564–567.

(9) Jínek, M.; Rehwinkel, J.; Lazarus, B. D.; Izaurralde, E.; Hanover, J. A.; Conti, E. The Superhelical TPR-Repeat Domain of O-Linked GlcNAc Transferase Exhibits Structural Similarities to Importin α. *Nat. Struct. Mol. Biol.* **2004**, *11* (10), 1001–1007.

(10) Martinez-fleites, C.; Macauley, M. S.; He, Y.; Shen, D. L.; Vocadlo, D. J.; Davies, G. J. Structure of an O-GlcNAc Transferase Homolog Provides Insight into Intracellular Glycosylation. *Nat. Struct. Mol. Biol.* **2008**, *15* (7), 764–765.

(11) Lazarus, M. B.; Jiang, J.; Gloster, T. M.; Zandberg, W. F.; Whitworth, G. E.; Vocadlo, D. J.; Walker, S. Structural Snapshots of the Reaction Coordinate for O-GlcNAc Transferase. *Nat. Chem. Biol.* **2012**, *8* (12), 966–968.

(12) Konrad, R. J.; Zhang, F.; Hale, J. E.; Knierman, M. D.; Becker, G. W.; Kudlow, J. E. Alloxan Is an Inhibitor of the Enzyme O-Linked N -Acetylglucosamine Transferase. *Biochem. Biophys. Res. Commun.* **2002**, *293*, 207–212.

(13) Elsner, M.; Tiedge, M.; Guldbakke, B.; Munday, R.; Lenzen, S. Importance of the GLUT2 Glucose Transporter for Pancreatic Beta Cell Toxicity of Alloxan. *Diabetologia* **2002**, *45*, 1542–1549.

(14) Zhang, H.; Gao, G.; Brunk, U. T. Extracellular Reduction of Alloxan Results in Oxygen radicaLmediated Attack on Plasma and Lysosomal Membranes. *Acta Pathol. Microbiol. Immunol. Scand.* **1992**, *100*, 317–325.

(15) Hennebicq-reig, S.; Lesuffleur, T.; Capon, C.; Bolos, C. D. E.; Kim, I.; Moreau, O.; Recchi, M.; Mae, E.; Richet, C.; He, B.; Aubert, J.; Real, F. X.; Zweibaum, A.; Delannoy, P.; Degand, P.; Huet, G. Permanent Exposure of Mucin-Secreting HT-29 Cells to Benzyl-N-Acetyl-α-D- Galactosaminide Induces Abnormal O-Glycosylation of Mucins and Inhibits Constitutive and Stimulated MUC5AC Secretion. *Biochem. J.*

1998, *334*, 283–295.

(16) Dorfmueller, H. C.; Borodkin, V. S.; Blair, D. E.; Pathak, S.; Navratilova, I.; Aalten, D. M. F. Van. Substrate and Product Analogues as Human O-GlcNAc Transferase Inhibitors. *Amino Acids* **2011**, *40* (3), 781–792.

(17) Gloster, T. M.; Zandberg, W. F.; Heinonen, J. E.; Shen, D. L.; Vocadlo, D. J. Europe PMC Funders Group Hijacking a Biosynthetic Pathway Yields a Glycosyltransferase Inhibitor within Cells. *Nat. Chem. Biol.* **2011**, *7* (3), 174–181.

(18) Upregulation, M. H. V.; Donovan, K.; Alekseev, O.; Qi, X.; Cho, W.; Azizkhan-clifford, J. O-GlcNAc Modification of Transcription Factor Sp1 Mediates Hyperglycemia-Induced VEGF-A Upregulation in Retinal Cells. *Biochem. Mol. Biol.* **2014**, *55* (12), 7862–7873.

(19) Ortiz-meoz, R. F.; Jiang, J.; Lazarus, M. B.; Orman, M.; Janetzko, J.; Fan, C.; Duveau, D. Y.; Tan, Z.; Thomas, C. J.; Walker, S. A Small Molecule That Inhibits OGT Activity in Cells. *ACS Chem. Biol.* **2015**, No. 10, 1392–1397.

(20) Okazaki, K.; Nishigaki, S.; Ishizuka, F.; Ogawa, S. Potent and Specific Sialyltransferase Inhibitors : Imino-Linked 5a'-carbadisaccharides. *Org. Biomol. Chem.* **2003**, No. 1, 2229–2230.

(21) Galan, M. C.; Venot, A. P.; Phillips, S.; Boons, G. The Design and Synthesis of a Selective Inhibitor of Fucosyltransferase VI. *Org. Biomol. Chem.* **2004**, No. 2, 1376–1380.

(22) Izumi, M.; Yuasa, H.; Hashimoto, H. Bisubstrate Analogues as Glycosyltransferase Inhibitors. *Curr. Top. Med. Chem.* **2009**, *9*, 87–105.

(23) Borodkin, V. S.; Schimpl, M.; Gundogdu, M.; Rafie, K.; Dorfmueller, H. C.; Robinson, D. a; Aalten, D. M. F. Van. Bisubstrate UDP-Peptide Conjugates as Human O-GlcNAc Transferase Inhibitors. *Biochem. J.* **2014**, *457* (3), 497–502.

(24) Gross, B. J.; Swoboda, J. G.; Walker, S. A Strategy to Discover Inhibitors of O-Linked Glycosylation. *J. Am. Chem. Soc.* **2008**, *130* (2), 440–441.

(25) Gross, B. J.; Kraybill, B. C.; Walker, S. Discovery of O-GlcNAc Transferase Inhibitors. *J. Am. Chem. Soc.* **2005**, *127* (42), 14588–14589.

(26) Filhoulaud, G.; Guillemain, G.; Scharfmann, R. The Hexosamine Biosynthesis Pathway Is Essential for Pancreatic Beta Cell Development. *J. Biol. Chem.* **2009**, *284* (36), 24583–24594.

(27) Jeon, J. H.; Suh, H. N.; Kim, M. O.; Ryu, J. M.; Han, H. J. Glucosamine-Induced OGT Activation Mediates Glucose Coordinately Contributed to the Regulation of Maintenance of Self-Renewal in Mouse Embryonic Stem Cells. *Stem Cells Dev.* **2014**, *23* (17), 2067–2079.

(28) Caldwell, S. A.; Jackson, S. R.; Shahriari, K. S.; Lynch, T. P.; Sethi, G.; Walker, S.; Vosseller, K.; Reginato, M. J. Nutrient Sensor O-GlcNAc Transferase Regulates Breast Cancer Tumorigenesis through Targeting of the Oncogenic Transcription Factor FoxM1. *Oncogene* **2010**, *29* (19), 2831–2842.

(29) Jiang, J.; Lazarus, M. B.; Pasquina, L.; Sliz, P.; Walker, S. A Neutral Diphosphate Mimic Crosslinks the Active Site of Human O-GlcNAc Transferase. *Nat. Chem. Biol.* **2012**, *8* (1), 72–77.

(30) Wang, Y.; Zhu, J.; Zhang, L. Discovery of Cell-Permeable O-GlcNAc Transferase Inhibitors via Tethering in Situ Click Chemistry. *J. Med. Chem.* **2017**, *60* (1), 263–272.

Chapter 3

Thiodigalactoside−bovine serum albumin conjugates as high-potency inhibitors of galectin-3: an outstanding example of multivalent presentation of small molecule inhibitors

This chapter have been accepted for publication:

Zhang, H.; Laaf, D.; Elling, L.; Pieters, R. J. Thiodigalactoside–Bovine Serum Albumin Conjugates as High-Potency Inhibitors of Galectin-3: An Outstanding Example of Multivalent Presentation of Small Molecule Inhibitors. *Bioconjug. Chem*. 2018, *in press.*

Abstract:

Galectin inhibitors are urgently needed to understand the mode of action and druggability of different galectins, but potent and selective agents are still evading researchers even now. Small-sized inhibitors based on thiodigalactoside (TDG) have shown their potential while modifications at their C3 position have indicated a strategy to improve selectivity and potency. Considering the role of galectins as a glycoprotein traffic police, involved in multivalent bridging interactions, we aim to create multivalent versions of the potent TDG inhibitors. Here, we present for the first time the multivalent attachment of a TDG derivative using bovine serum albumin (BSA) as the multivalent scaffold. An efficient synthetic method is presented to obtain a novel type of neo-glycosylated proteins loaded with different numbers of TDG moieties. A polyethylene glycol (PEG)-spacer is introduced between the TDG and the protein scaffold maintaining an appropriate accessibility for an adequate galectin interaction. These novel conjugates were evaluated in galectin binding and inhibition studies *in vitro*. The compound with a moderate density of 19 conjugated TDGs has been identified as one of the most potent multivalent Gal-3 inhibitors so far, with a clear demonstration of the benefit of a multivalent ligand presentation. The described method may facilitate the development of specific galectin inhibitors and their application in biomedical research.

3.1 Introduction

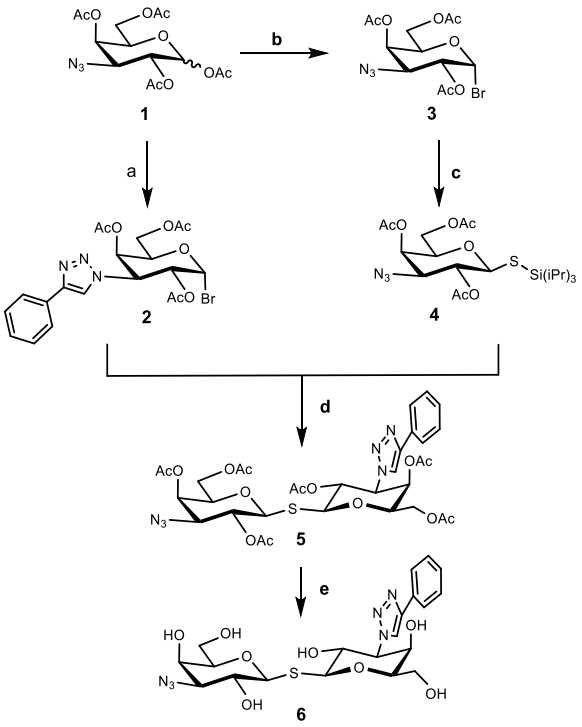
A dense layer of carbohydrates is found on mammalian cells and the variety of the attached glycans results in specific profiles for molecular recognition. This recognition involving the so-called 'sugar code' is operational by reversible interaction of carbohydrate-binding proteins. Members of this protein class are described as lectins and fulfil a variety of effector functions in terms of cellular communication.[1] Galectins as one subtype of lectins that can specifically recognize β-galactosides and are found in fungi, invertebrates and vertebrates.[2] Fifteen different galectins have been identified in humans until now and they play crucial roles in the organization of receptor-lectin complexes (lattices) and regulation of immune responses.[3,4,5] Their concave shaped groove for oligosaccharide binding, namely the carbohydrate recognition domain (CRD), is highly preserved. Among all the known galectins, galectin-1 (Gal-1) and galectin-3 (Gal-3) are the most thoroughly studied due to their involvement in angiogenesis, tumor progression and metastasis.[6,7,8] The participation in malignant processes makes them promising targets for anti-cancer therapy. In this regard, immense efforts have been spent on the synthesis of potent and specific ligands. The prevailing majority of drug discovery efforts have been focused on the derivatives of lactose and N-acetyllactosamine (methyl β-lactoside, K_d = 220 μM for Gal-3, K_d = 190 μM for Gal-1) (methyl β-LacNAc, K_d = 67 μM for Gal-3).[9,10] Thiodigalactoside (TDG, K_d = 49 μM for Gal-3, K_d = 24 μM for Gal-1) has been identified as a more potent Gal-3 inhibitor with additional advantages such as its enhanced glycolytic stability while maintaining a similar binding mode compared to lactose and LacNAc.[11,12]

Apart from designing small molecule inhibitors, a multivalent-based strategy has been adopted to promote the interactions between ligands and protein, as this more closely mimics the natural way in which galectins interact with glycoproteins. So far different scaffolds have been reported to carry multiple galectin ligands.[13,14,15,16,17] In our previous work, bovine serum albumin (BSA) was used as plain protein carrier for neo-glycosylation.[18,19] However, the conjugation technique adopted in those studies was complicated which restricted extensive applications. In addition, the squaric acid functional group used in the synthesis of the spacer may cause allergic and/or immunogenic reactions when applied *in vivo*.[20] *N*-hydroxysuccinimidyl (NHS) esters, the most commonly used amine-reactive reagents, enables a larger variety of amine-reactive ligands for labeling proteins.[21,22,23] Benefiting from this prior information, we successfully approached a type of neo-glycoproteins through conjugating NHS functionalized-TDG to lysine residues of BSA. The most obvious finding to emerge from the present study is that the TDG-conjugates exhibit outstanding high inhibitory potencies despite a low or moderate number of attached ligands. By combining a highly potent monovalent ligand with a beneficial multivalent presentation resulted in some of the most effective Gal-3 inhibitors. Besides, the multivalent TDG-conjugates represent the first example of decorating a non-glycosylated carrier with TDG derivatives.

3.2 Results and Discussion

3.2.1 Synthesis of unsymmetrical TDG precursor

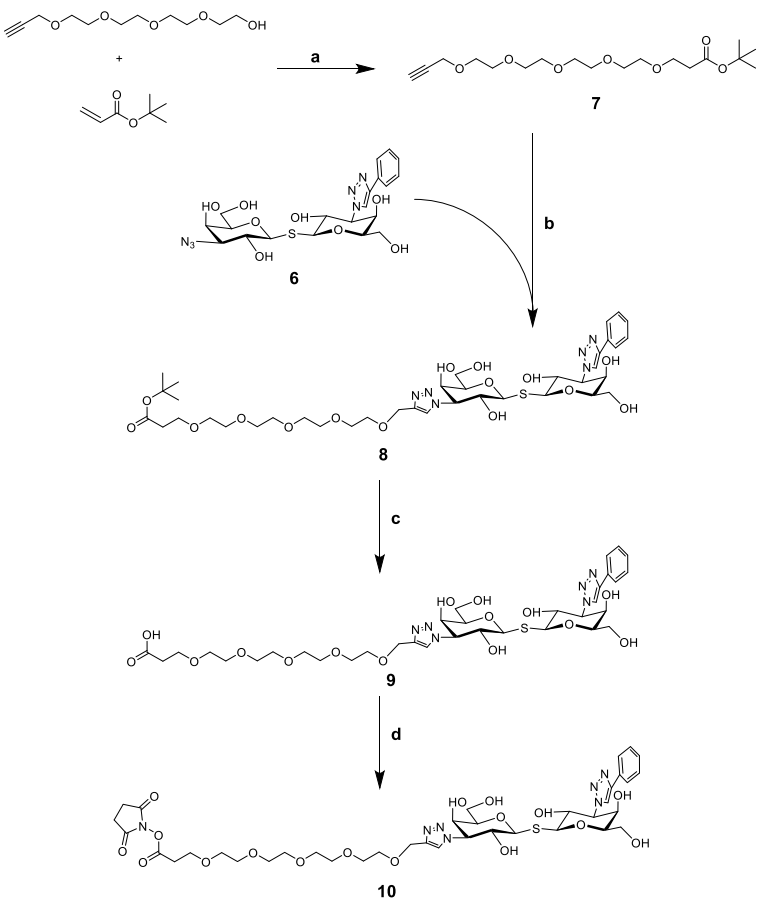
The synthesis of carboxy-functionalized TDG started as shown in Scheme 1. The key building block was the unsymmetrical TDG derivative **6** carrying a phenyltriazole and an azide at the 3- and 3'-positions. To construct compound **6**, two building blocks, namely tri-isopropylsilyl thioglycoside (compound **4**) and glycosyl halide (compound **2**), were prepared using a published method. The coupling involved a one-pot desilylation and glycosyl thiol alkylation with the glycosyl halide.[24] Hence, compound **1**, prepared from commercially available 1,2,5,6-diacetone-α-D-glucofuranoside through a known four-step reaction,[25] was brominated ($TiBr_4$, 67% yield) to afford compound **3**, which was converted to thio-glycoside **4** stereoselectively with tri-isopropylsilylthiol (TIPSSH) in the presence of K_2CO_3. Meanwhile, a copper-catalyzed azide-alkyne cycloaddition of compound **1** with phenylacetylene provided the corresponding triazole crude, which reacted with HBr to obtain the glycosyl halide **2** (79% yield in two steps). Desilylating and activating with TBAF turned compound **4** into a thiol nucleophile which subsequently replaced the anomeric bromide of compound **2** through an SN_2 reaction. Purification by silica column chromatography gave the resulting compound **5** in 54% yield. After removing the acetyl protecting group, the resulting crude **6** was used for the next step without further purification (Scheme 1).



Scheme 1. Synthesis of compound **6**. Reagents and conditions: (a) (i) Phenylacetylene, $CuSO_4$, sodium ascorbate, DMF/H_2O, 80°C, microwave, (ii) HBr, CH_2Cl_2, 25 °C, 79% yield in two steps; (b) $TiBr_4$, CH_2Cl_2/EtOAc, 25 °C, 67%; (c) TIPSSH, K_2CO_3, CH_3CN, 25 °C, 30%; (d) TBAF, CH_3CN, 25 °C, 62%; (e) NaOMe, CH_3OH, 25 °C.

3.2.2 Synthesis of polyethylene glycol (PEG)-spacer and carboxy-functionalized TDG

The synthesis of polyethylene glycol (PEG)-spacer (compound **7**) started from tetraethylene glycol. As previously reported, the reaction of tetraethylene glycol with an equal amount of propargyl bromide in the presence of NaH in THF at room temperature gave the monoalkyne terminated PEG_4 .[26] Then the Michael addition of resulting compound with tert-butyl acrylate in the presence of catalytic sodium metal gave compound **7** in 70% yield.[27] Having assembled the important intermediates (compound **6** and **7**), the next objective was their CuAAC (Copper-catalyzed azide–alkyne cycloaddition) conjugation assisted by copper iodide. After purification with size-exclusion chromatography on Bio-Gel, compound **8** was obtained as a white fluffy solid. Removal of the tert-butyl group from compound **8** (TFA: DCM = 1: 1) gave carboxyl compound **9** (90% yield), which was transferred to the corresponding NHS-ester **10** through coupling with TSTU in DMF (Scheme 2). Compound **10** is prone to hydrolysis (e.g. during purification), thus it was directly used for further reaction.

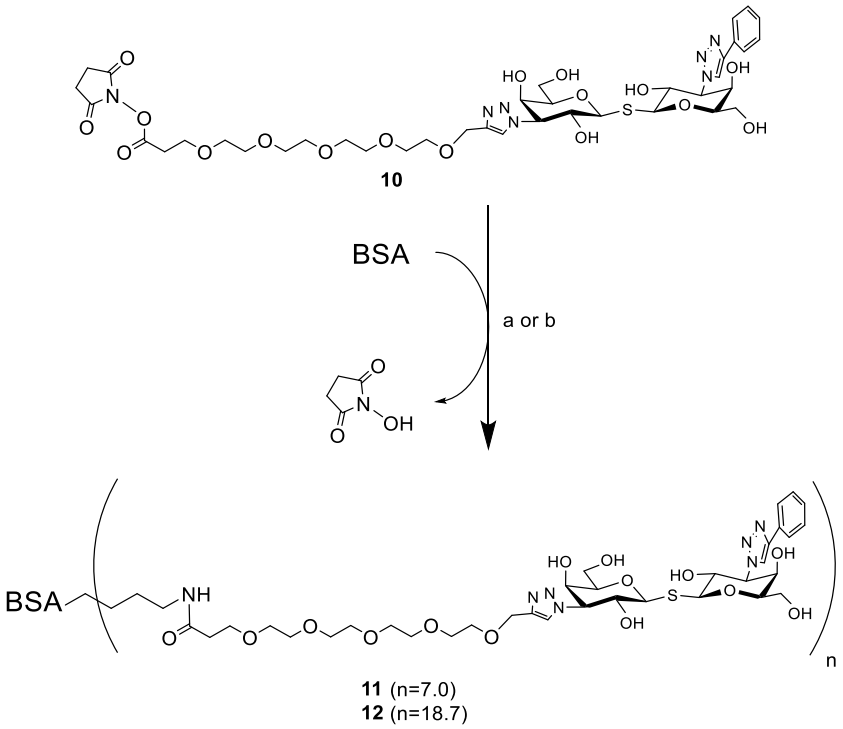


Scheme 2. Preparation of carboxy- (**9**) and NHS-functionalized (**10**) TDG derivatives. Reagents and conditions: (a) Na, THF, 0 to 25 °C, 39%; (b) CuI, CH_3OH, 25 °C, 64%; (c) TFA/DCM, 25 °C, 98%; (d) TSTU, DiPEA, DMF, 25 °C.

3.2.3 Neo-glycoprotein synthesis and analysis

Compared to the click-chemistry and squarate linker chemistry, NHS-mediated coupling is a straightforward and convenient coupling strategy for the modification of protein carriers. [28,29,30,31,19,21,23] The lysine residues of bovine serum albumin (BSA) react with the NHS ester moiety of the TDG derivative, but a crucial factor that needs to be taken into consideration is the distance between the TDG and BSA. Benefiting from previous work on chitooligomer spacers in neo-glycoproteins, we concluded that a suitable spacer with a certain significant length needed to be used for maintaining an appropriate ligand accessibility and proper galectin interaction of the final products.[31] A PEG spacer with a similar length as used the mentioned study, was incorporated into conjugation agent **9**. Furthermore, the modification of carbohydrates or derivatives thereof with PEG

as a biocompatible molecule is a commonly used technique.[32] The NHS functionalized-TDG **10** readily reacted with amino groups of BSA using a reaction buffer that contained 35 mM HEPES (pH = 7.0). The total amount of compound **10** was divided into three portions and added batch-wise after every 24h. As a result, we obtained compound **11** that was verified by the TNBSA-assay to carry 7.0 ± 1.0 TDG moieties per BSA molecule (Scheme 3) and the coupling efficiency was found to be 8 %. Shifting the pH to a slightly higher value (pH 8.0 - 9.0) by the addition of triethylamine (TEA) gave arise to compound **12**. The TNBSA-assay confirmed that the number of attached TDGs was now 18.7±1.6 corresponding to a coupling efficiency of 20 %. The elevation of the pH may deprotonate the amino groups of the lysine residues to a degree sufficiently high for fast and efficient coupling. In accordance with our previous findings, reducing sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) confirmed the attachment of the TDGs.



11 (n=7.0)
12 (n=18.7)

Scheme 3. Loading of BSA protein carriers with NHS-functionalized compound **10** resulting in compounds **11** and **12**. Reagents and conditions: (a) compound **11**: BSA (0.06mM, 150 µL) in HEPES buffer (35 mM, pH = 7.0), compound **10** in DMF (54 mM, 3 × 5µL), 72 h, 4 °C, (b) compound **12**: BSA (0.06 mM, 50 µL) in HEPES buffer (pH = 8.0 - 9.0, adjusted with TEA), compound **10** in DMF (54 mM, 3 × 1.67 µL), 72h, 4°C.

3.2.4 Evaluation of multivalent BSA-conjugates as galectin ligands

In collaboration with the biomaterials group of Lothar Elling (RWTH Aachen University), multivalent TDG-conjugates **11** and **12** were utilized as immobilized ligands in solid-phase binding assays with recombinant human His_6-tagged Gal-1 and Gal-3 (Figure 1). Except for half-maximal binding signal corresponding to the dissociation constant (K_d) and maximal binding signal (B_{max}), binding efficiency (B_{max}/K_d) was adopted to evaluating these two compounds (Table 1). The binding studies showed that Gal-1 bound **11** in a three-fold more efficient manner than Gal-3 did. Gal-1 showed higher affinity for conjugate **11** with a 6-fold reduced apparent dissociation constant K_d in comparison with Gal-3 with p<0.001 (student t-test). However, Gal-3 showed an elevated capacity when binding to conjugate **11**, indicated by the higher B_{max} value (Table 1). A possible explanation is the glycoside cluster effect, which may result in galectin oligomerization and thus to increase the binding signal.[33,34] This effect might be more pronounced for Gal-3, which is generally considered to form higher oligomers when binding to multivalent ligands. Gal-1 aggregation is likely limited to the formation of dimers. [35]

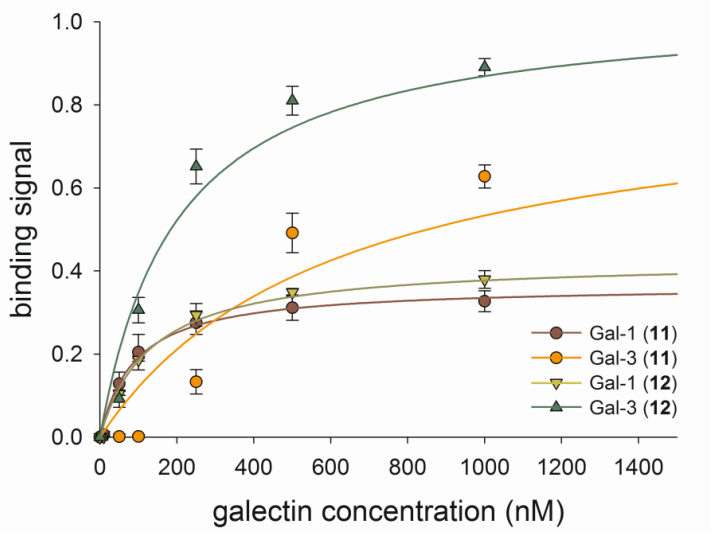


Figure 1. Behavior of recombinant human Gal-1 and Gal-3 for binding to immobilized neo-glycoconjugates **11** and **12.** The subtracted blank value (no Gal-1 and/or Gal-3) was 0.047±0.003.

Conjugate **12** contains a higher number of TDG derivatives, and is therefore bound to Gal-3 with a higher affinity. In addition, the highest binding efficiency was observed in the assay for Gal-3 toward conjugate **12** with a lower K_d value and raised B_{max} value (Table 1, Figure 2). The enhancement of Gal-3 binding efficiency for compound **12** was significant with p<0.001 (student t-test). In contrast, Gal-1 was not affected much by conjugate **12,**

the capacity (B_{max}) of which was slightly increased while the affinity was decreased. Thus, the binding efficiencies for Gal-1 toward conjugates **11** and **12** are almost the same within experimental error (Figure 2, Table 1).

Table 1. Binding behavior of Gal-1 and Gal-3 using multivalent TDG conjugates as immobilized ligands.

Galectin	ligand	K_d (µM)[a]	B_{max}(-)[a]	Galectin binding efficiency [μM^{-1}][b]
Gal-1	**11**	0.090±0.012	0.37±0.01	4.1±0.8
Gal-1	**12**	0.131±0.017	0.43±0.01	3.3±0.6
Gal-3	**11**	0.616±0.246	0.86±0.12	1.4±0.5
Gal-3	**12**	0.199±0.446	1.04±0.06	5.2±0.3

[a] determined in ELISA; [b] ratio of B_{max} and K_d

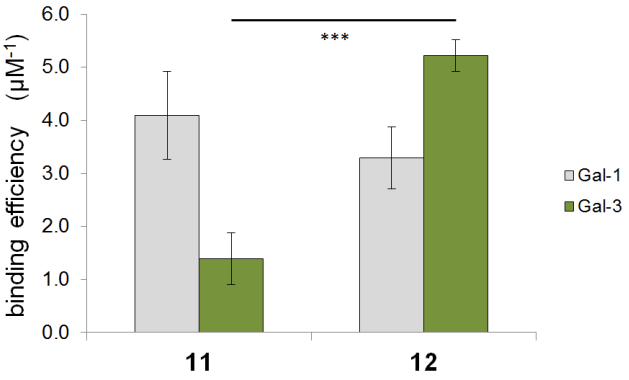


Figure 2. Binding efficiencies of recombinant human Gal-1 and Gal-3 for binding to immobilized neo-glyconjugates **11** and **12**. Triple asterisks corresponds to a significant difference with a confidence interval of $p<0.001$.

To sum up, Gal-1 and Gal-3 bind conjugates **11** and **12** with high affinity through recognizing the conjugated TDGs. These findings are consistent with our previous studies, in which galectins interacted with the corresponding monovalent TDG compounds in a fluorescence polarization assay.[36] However, we did not detect any specificity differences as seen for previously synthesized neo-glycoproteins carrying poly-*N*-acetyllactosamine (poly-LacNAc) derivatives.[18] A possible reason is that the TDG ligands primarily interact with a very conserved region in the galectin's CRD and thus produce little variations in

binding affinities.[37,38,39] The larger aromatic substituents (e.g. 4-phenoxyphenyl) on the TDG scaffold could possibly offer a selectivity among galectins, which could eventually promote the discovery of selective galectin inhibitors based on neo-glycoproteins.[36]

3.2.5 Evaluation of multivalent BSA-conjugates as Gal-3 inhibitors

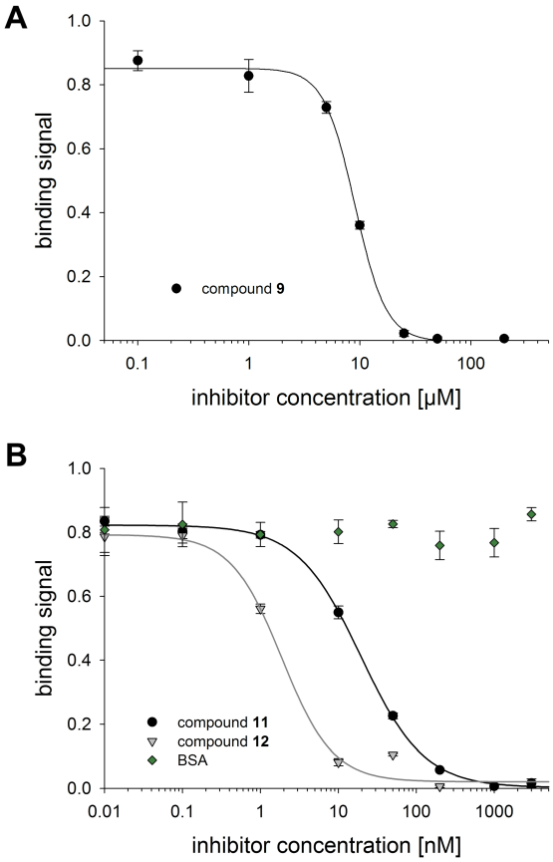


Figure 3. Competitive inhibition of Gal-3 (5.56 µM) binding to ASF using monovalent carboxy-functionalized compound **9** (A) and multivalent TDG-conjugates **11** and **12** at indicated concentration (B). Refer to Scheme 2 and Scheme 3 for compound structures. The subtracted blank value (no Gal-3) was 0.091±0.005.

The univalent TDG derivative **9** and multivalent glycoconjugates **11** and **12** were also evaluated in terms of their capability to prevent Gal-3 from binding to immobilized asialofetuin (ASF). The ASF glycoprotein was considered as a standard galectin ligand as reported before.[18,19,31,40,41] In our case, Gal-3 was incubated together with increasing amounts of inhibitors **9**, **11** or **12** until the ASF-galectin interaction was blocked completely.

Non-modified BSA was utilized as a negative control that can't perturb the binding between Gal-3 and ASF. As Figure 3 indicates, a complete inhibition of Gal-3 binding to ASF was reached when inhibitor **9** (>25 µM) was used. The same inhibition effect also occurred for conjugates **11** and **12** (>200 nM). The resulting sigmoidal inhibition curves were the basis for the calculation of the IC_{50} value, defined as the inhibitor concentration at which the half-maximal inhibition was reached (Table 2). The inhibition strength of compound 9 was in the low micromolar range and fits the range of *N',N''*-diacetyl-lactosamine (LacdiNAc)-LacNAc tetrasaccharide, which was identified as specific ligand of Gal-3.[19] However, we assume that the NHS-/PEG-modification of one C3 atom (asymmetric character) of **9** may reduce its affinity when applied as non-conjugated Gal-3 inhibitor. When comparing previous relevant studies, it becomes clear that Gal-3 has a much higher affinity for the symmetrically modified TDG. When both TDG's galactose moeities carried the same substituents such as C(3)-benzamides,[38,42] or C(3)-triazoles,[36,43] the Gal-3 affinities (K_d) were between 22 and 360 nM in a fluorescence polarization assay.

When the asymmetric TDG derivative **9** is conjugated to a non-glycosylated serum protein scaffold to give the conjugates **11** and **12**, it's binding potential is fully exhibited. As shown in Table 2, the conjugates **11** and **12** have extraordinary high inhibitory potencies and low IC_{50} values. Even though the conjugate **11** only carried a quantity of seven TDG derivatives, the inhibition strength was raised 465-fold compared to the monovalent compound **9**. This corresponds to an improvement factor of 66 per loaded TDG. The impact of multivalent ligand presentation was even more pronounced for glycoconjugate **12** carrying a higher number (n=18.7) of TDG derivatives. Here, the inhibition strength was raised more than 4800-fold compared to the compound **9**, representing an improvement factor per TDG of 256.

Our findings suggest that multivalent conjugates **11** and **12**, but not monovalent compound **9**, inactivate more Gal-3 molecules than the amount of presented TDG derivatives, as seen before.[17] On the one hand, both multivalent inhibitors may induce the formation of Gal-3 complexes, cross-linked by their N-termini.[44] On the other hand, type-C Gal-3 self-association is most likely. Here, the non-occupied CRD of Gal-3 molecules interact with already TDG-bound Gal-3 leading to an oligomerization and stacking as reported before.[45] To the best of our knowledge, the tremendously diminished IC_{50} value makes multivalent glycoconjugate **12** one of the most effective Gal-3 inhibitors. The multivalent design promotes the cluster glycoside effect resulting in a highly efficient entrapment of Gal-3.[1,33,34].Similar neo-glycoproteins with a cargo of different poly-LacNAc derivatives were recently synthesized and used as Gal-3 inhibitors.[18,19,29] Herein, we may use them as a reference to evaluate the presented results. We take those BSA neo-glycoconjugates bearing the LacNAc-LacNAc (n = 7.5) or LacdiNAc-LacNAc (n = 7.4) glycans as the ideal benchmarks because they have an equal modification density

compared to the conjugate **11**. In that case only moderate inhibition strengths were observed, with IC_{50} values of 850 nM ([LacNAc-LacNAc]$_{n= 7.4}$ - BSA) and 1100 nM ([LacdiNAc-LacNAc]$_{n= 7.4}$ - BSA), respectively.[19] Clearly the in TDG ligand has a potency advantage but the implementation of the PEG-spacer may also be a favorable feature in terms of ligand accessibility and flexibility.[32,46] To evaluate conjugate **12** similarly, previously synthesized neo-glycoproteins may also serve as ideal references. Conjugates with LacNAc-LacNAc (n = 17.8), LacdiNAc-LacNAc (n = 18.0) or derivatized poly-LacNAc hexasaccharides of equal modification density (n = 16 - 19) were prepared and thoroughly studied in terms of galectin interaction.[19,18] The respective inhibition constants ranged between 60-90 nM and 37-76 nM. With an outstanding low IC_{50} (1.88 nM), conjugate **12** has at least more than 20-fold elevated potency compared to those reference neo-glycoproteins.

Table 2. Inhibition constants and inhibitory potencies of TDG derivative **9** and multivalent TDG-conjugates **11** and **12**.

Inhibitor	IC_{50} (nM)[a]	number of loaded TDG moieties (-)[b]	relative inhibitory potency	relative inhibitory potency per glycan
9	9030±27	1.0±0.0	1.0±0.0	1.0±0.0
11	19.40±1.09	7.0±1.0	465.5±27.5	66.5±13.4
12	1.88±0.38	18.7±1.6	4803.2±985.2	256.9±74.7

[a] ELISA; [b] TNBSA-assay

3.3 Conclusion and outlook

TDG derivatives have been used as galectin inhibitors for research. The aromatic groups on the C3 and C3' position of TDG produce selectivity and affinity toward galectins. In the present study, we report on the synthesis of an asymmetrical TDG structure that can be used to yield multivalent compounds through conjugating to a protein scaffold. To obtain the key precursor, a straightforward approach was used to lead to the NHS functionalized-TDG derivative. Subsequent reaction with BSA gave multivalent TDG-glycoconjugates. Weak alkaline pH, adjusted by TEA, was crucial for an effective conjugation. To the best of our knowledge, this is the first example of conjugating a TDG derivative to a carrier. The multivalent presentation on conjugates **11** and **12** unlocks TDG's full potential. Extraordinary high multivalency factors were observed that resulted in one of the most effective inhibition of Gal-3 *in vitro* until now. The result is clearly a

combination of the binding properties of the monovalent ligand and the multivalent display by the BSA. As previously noted, potent galectin inhibition cannot be achieved with very weak or non-binding ligands, conjugated to BSA.[31] Furthermore we note that, while a multivalent scaffold can enhance existing binding potency, the specificity at the multivalent level remains the same.[47] In other systems, very strong multivalence effects have been reported leading to picomolar inhibition, usually involving the simultaneous binding of ligands to nearby binding sites.[34] This chelation type mechanism is less likely to contribute to the present system, due to the monovalent nature of the protein. Considering this, other modes of action such as statistical rebinding or aggregation, usually lead to smaller effects,[34] which makes the present results more notable. In the present system, the PEG-spacer likely helps to make the TDGs accessible for the inter-acting galectins. The multivalent TDG-modified conjugates (**11**, **12**) can provide ideal properties for biomedical applications with its structural features. Both the serum protein and PEG are biocompatible and they have been used in different biological contexts and therapeutic fields. Therefore, cell culture *in vitro* experiments are planned to elucidate the power of the synthesized conjugate and *in vivo* applications may follow. The TDGs with different functional groups on C3 and C3' position (e.g. 4-phenoxyphenyl) will be attached onto BSA to produce selective galectin inhibitors on a multivalent level.

3.4 Experimental section

3.4.1 Reagents and general information

All reagents employed were of American Chemical Society (ACS) grade or finer and were used without further purification unless otherwise stated. TLC analysis was performed on Merck precoated silica gel 60 F-254 plates. Spots were visualized with UV-light, ninhydrin stain (1.5 g ninhydrin and 3.0 mL acetic acid in 100 mL n-butanol), Potassium permanganate (1.5 g $KMnO_4$, 10 g K_2CO_3, and 1.25 mL 10% NaOH in 200 mL H_2O) and sulfuric acid (10 % sulfuric acid in methanol). Column chromatography was performed using Silica-P Flash silica gel (60 Å, particle size 40 – 63 µm) from Silicycle (Canada). Microwave reactions were carried out in a Biotage Initiator (300 W) reactor. Lyophilization was performed on a Christ Alpha 1-2 apparatus. ^1H and ^{13}C NMR spectroscopy was carried on an Agilent 400-MR spectrometer operating at 400 MHz for ^1H and 101 MHz for ^{13}C. HSQC, TOCSY and NOESY (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. Analytical LC-MS (electrospray ionization) was performed on Thermo-Finnigan LCQ Deca XP Max using same buffers and protocol as described for analytical HPLC.

Analytical HPLC was performed on a Shimadzu-10AVP (Class VP) system using a Phenomenex Gemini C18 column (110Å, 5 µm, 250×4.60 mm) at a flow rate of 1 mL min. The used buffers were 0.1 % trifluroacetic acid in CH_3CN: H_2O= 5: 95 (buffer A) and 0.1 %

trifluroacetic acid in CH_3CN: H_2O= 95: 5 (buffer B). UV-absorption was measured at 214 and 254 nm.

3.4.2 Synthetic procedures and compound characterization

2,4,6-tri-*O*-acetyl-3-deoxy-3-(4-phenyl-1H-1,2,3-triazol)-α-D-galactopyranosyl bromide (1)

Compound **1** (300 mg, 0.80 mmol), sodium L-ascorbate (237.6 mg, 1.2 mmol), $CuSO_4.5H_2O$ (100 mg, 0.40 mmol), Tris(3-hydroxypropyltriazolylmethyl)amine (THPTA, 4.2 mg, 0.0096 mmol), and phenylacetylene (176.4 µl, 1.6 mmol) were dissolved in DMF (13.5 mL) and H_2O (1.5 mL). The reaction was performed under microwave irradiation at 80 °C for 40 mins. Subsequently, the solvent was evaporated and the residue was dissolved in CH_2Cl_2 (100 mL), washed with H_2O (1 × 100 mL) and brine (1 × 100 mL), dried over Na_2SO_4, filtered and concentrated. The crude product was obtained as a white solid (310 mg) and used for next reaction directly. To a solution of crude (310 mg) in dry CH_2Cl_2 (50 mL) the HBr (33% HBr in acetic acid, 2.0 mL) was dropwise added under N_2 atmosphere. The solution was sealed and stirred overnight at room temperature. A saturated $NaHCO_3$ solution (50 mL) was added to quench the reaction and then the organic layer was washed with H_2O (1 × 50 mL) and brine (1 × 50 mL), dried with Na_2SO_4 and filtered. The residue was purified by silica chromatography (hexanes: EtOAc = 1: 1) and gave the product **2** as a light yellow solid (312 mg, two steps, 79% overall yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.85 – 7.70 (m, 3H, ar, triazole), δ 7.41 (dd, J = 8.3, 6.8 Hz, 2H, ar), 7.37 – 7.28 (m, 1H, ar), 6.86 (d, J = 3.8 Hz, 1H, H1), 5.79 (dd, J = 11.4, 3.8 Hz, 1H, H2), 5.63 (dd, J = 3.1, 1.3 Hz, 1H, H4), 5.32 (d, J = 3.0 Hz, 1H, H3), 4.67 – 4.58 (m, 1H, H5), 4.22 (dd, J = 11.6, 6.3 Hz, 1H, H6a), 4.11 (dd, J = 11.6, 6.3 Hz, 1H, H6b), 2.05, 2.04, 1.93 (3s, each 3H, OCH_3).

^{13}C NMR (101 MHz, $CDCl_3$) δ 170.25, 169.55, 168.94, 128.89, 128.49, 125.70 , 119.33 , 88.52 , 77.17 , 71.34 , 67.76 , 66.85 , 60.86 , 58.66 , 31.40 , 29.67 , 20.58 , 20.45 , 20.34 .

HRMS (EI, m/z) calculated for $C_{20}H_{22}BrN_3O_7H+$ ($[M+H]^+$): 496.0714, found 496.0707.

2,4,6-tri-O-acetyl-3-azido-α-D-galactopyranosyl bromide (3)

Compound **1** (1.2 g, 3.2 mmol) was dissolved in CH_2Cl_2 (50 mL) and EtOAc (5.0 mL) and then titanium tetrabromide ($TiBr_4$, 2.4 g, 6.4 mmol) was added slowly. The reaction mixture was stirred under sealed conditions overnight at room temperature. NaOAc (2.0 g, 24 mmol) was added to quench the reaction and washed with H_2O (3 × 50 mL). The organic layer was dried over Na_2SO_4, filtered and concentrated in vacuo. Silica chromatography (Hexanes: EtOAc = 3:1) yielded **7** as a clear oil (855 mg, 67.4%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.68 (d, J = 3.9 Hz, 1H, H1), 5.48 (dd, J = 3.4, 1.4 Hz, 1H, H4), 4.98 – 4.89 (m, 1H, H2), 4.39 (td, J = 6.8, 6.1, 1.5, 0.7 Hz, 1H, H5), 4.17 (dd, J = 11.5, 6.1 Hz, 1H, H6a), 4.11 (dd, J = 10.6, 3.3 Hz, 1H, H3), 4.03 (dd, J = 11.5, 6.8 Hz, 1H, H6b), 2.16 (s, 3H, $C(O)CH_3$), 2.14 (s, 3H, $C(O)CH_3$), 2.05 (s, 3H, $C(O)CH_3$).

^{13}C NMR (101 MHz, $CDCl_3$) δ 170.27, 169.75 and 169.58 ($C(O)CH_3$), 88.19 (C-1), 71.29 (C-5), 69.47 (C-2), 67.05 (C-4), 60.95 (C-6), 58.32 (C-3), 20.69, 20.60, 20.50 ($C(O)CH_3$).

Tri-isopropylsilyl 3-azido-2,4,6-tri-*O*-acetyl-1-thio-β-D-galactopyranoside (4)

To a solution of **3** (770 mg, 1.95 mmol) in dry CH_3CN (10 mL) N_2 gas was purged for 10 mins, then K_2CO_3 (809 mg, 5.86 mmol) was added followed by tri-isopropylsilylthiol (TIPSSH, 628 μL, 2.93 mmol), and the reaction was stirred for 3 h at room temperature. After complete conversion of the starting material according to TLC monitoring, the solvent was evaporated and the residue was dissolved in CH_2Cl_2 (20 mL), washed with H_2O (2 x 20 mL). The organic layer was dried over Na_2SO_4, filtered and concentrated in vacuo. Silica chromatography (hexanes: EtOAc = 4: 1) yielded **4** as a white solid (300 mg, 30%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.42 (dd, J = 3.5, 1.0 Hz, 1H, H4), 5.20 (t, J = 9.8 Hz, 1H, H2), 4.60 (d, J = 9.5 Hz, 1H, H1), 4.11 (dd, J = 11.5, 5.6 Hz, 1H, H6a), 4.00 (dd, J = 11.6, 7.1 Hz, 1H, H6b), 3.79 (ddd, J = 6.9, 5.5, 1.1 Hz, 1H, H5), 3.53 (dd, J = 10.1, 3.4 Hz, 1H, H3), 2.15, 2.12, 2.02 (3S, 9H, 3 $COCH_3$), 1.25 (m, 3H, -SiC_3H_3), 1.14 – 1.07 (m, 18H, -$SiC_3H_3C_6H_{18}$).

^{13}C NMR (101 MHz, $CDCl_3$) δ 170.57, 170.25, 169.39 (3 $COCH_3$), 80.37 (C-1), 75.47 (C-5), 62.87 (C-3), 72.17 (C-2), 68.12 (C-4), 62.13 (C-6), 20.84, 20.76, 20.51 (3$COCH_3$), 18.52, 18.22 (6 –SiCHCH_3), 12.76 (3 –SiCH).

HRMS (EI, m/z): calculated for $C_{21}H_{37}N_3O_7SSiNa^+$ ([M+Na]$^+$): 526.2014, found 526.2011.

3-azido-3'-phenyl-2,2',4,4',6,6'-hexa-O-acetyl β-D-thiodigalactoside (5)

The solution of **2** (190 mg, 0.38 mmol) in dry CH_3CN (10 mL) was added by **4** (193 mg, 0.38 mmol), N_2 gas was purged for 10 mins through the solution, and tetra-n-butylammoniumfluoride (TBAF, 1M in THF, 460 μL) was added. Following complete conversion of the starting material after 5 mins according to TLC analysis, the solvent was evaporated and silica chromatography (hexanes:EtOAc = 1: 1 → 1: 2) gave compound **5** (160 mg, 54 %) as white solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ = 7.80 (s, 1H, triazole), 7.77-7.71 (m, 2H, ar), 7.43-7.28 (m, 3H, ar), 5.75-5.68 (m, 1H, H-2'), 5.61 (d, 1H, J =3.2 Hz, H-4'), 5.47 (d,1H, J4,3 =3.4 Hz, J4,5 = 1.1 Hz, H-4), 5.21 (dd, 1H, J3',2'= 8.6 Hz, J3',4'= 2.5 Hz, H-3'), 5.17 (m, 1H, H-2), 4.98 (1H, d, J1',2'= 9.8 Hz, H-1'), 4.84 (1H, J1,2 = 10.0 Hz, H-1), 4.11 (m, 5H, H-5', H-6ab, H-6a'b'), 3.89

(1H, td, J5,4 = 1.2 Hz, J5,6ab = 6.4 Hz, H-5), 3.67 (1H, J3,2 = 10.1 Hz, J3,4 = 3.4 Hz, H-3), 2.15, 2.13, 2.08, 2.05, 2.04, and 2.02 (6s, total 18H, C(O)CH_3).

^{13}C NMR (101 MHz, $CDCl_3$) δ 170.34, 170.25, 169.86, 169.54, 169.37, 168.70 , 147.92 , 129.96 , 128.90 , 128.47 , 125.68 , 118.26 , 82.11 , 81.51 , 77.34 , 77.02 , 76.70 , 75.53, 68.71 , 68.43 , 67.66 , 66.33 , 62.94 , 62.78 , 61.56 , 61.38 , 20.76 , 20.68 , 20.62 , 20.61 , 20.47 , 20.38 .

HRMS (EI, m/z) calculated for $C_{32}H_{38}N_6O_{14}SH^+$ ($[M+H]^+$): 763.2239, found 763.2277.

3-(4-phenyl-1H-1,2,3-triazol)-3'-azido-thiodigalactoside (6)

NaOMe (40 mg, 2.5 mmol) was added in the solution of compound **5** (120 mg, 0.16 mmol) in CH_3OH (5.0 mL) and the mixture was stirred for 6 h at room temperature. The solution was neutralized with DOWEX-H^+ resin, filtered, and evaporated. Crude **6** was obtained as a white solid and used in next step without further purification.

Tert-butyl 4,7,10,13,16-pentaoxanonadec-18-ynoate (7)

To a solution 3,6,9,12-tetraoxapentadec-14-yn-1-ol of (200 mg, 0.86 mmol) in 5 mL of THF was added sodium (0.6 mg, 0.025 mmol). When the sodium was dissolved, tert-butyl acrylate (0.125 mL, 0.86 mmol) was added. The solution was stirred for 20h at room temperature and H_2O (1 mL) was added to quench the reaction. After removal of the solvent, the residue was suspended in brine and extracted three times with ethyl acetate. The combined organic layers were dried over Na_2SO_4 before the solvent was removed. The resulting oil was purified by silica chromatography (hexanes: EtOAc = 1: 1 → 1: 2) to give compound **7** (120 mg, 39 %) as a colorless oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 4.17 (d, J = 2.4 Hz, 2H), 3.73 – 3.53 (m, 18H), 2.46 (t, J = 6.6 Hz, 2H), 2.40 (t, J = 2.4 Hz, 1H), 1.41 (s, 9H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 170.82, 80.42, 74.43, 70.57, 70.56, 70.53, 70.46, 70.36, 70.33, 69.07, 69.07, 66.85, 58.34, 36.2 , 28.08, 28.05.

Preparation of compound 8

The compound **7** (33 mg, 0.092 mmol) and compound **6** crude (48 mg) were dissolved into CH_3CN (2.0 mL) and then CuI (18 mg, 0.093 mmol) was added into the solution. The resulting mixture was heated under microwave irradiation to 80 °C for 90 mins. After complete conversion of the starting material according to TLC monitoring, the mixture was concentrated in vacuo, and then H_2O (1.0 mL) was added. A clear solution was obtained after centrifuge, which was purified by size-exclusion chromatography (Bio-Gel P2 fine;

column 2.5 cm × 120 cm; flow rate 0.3 ml/min; elution with H_2O/*n*-Butanol = 95/ 5). The fractions containing the product were pooled and freeze-dried to give compound **8** (52 mg, 0.060 mmol, 64 %) as a white fluffy solid.

^{1}H NMR (500 MHz, D_2O) δ 8.55 (s, 1H), 8.26 (s, 1H), 7.87 (d, J = 7.7 Hz, 2H), 7.55 (t, J = 7.6 Hz, 2H), 7.48 (t, J = 7.5 Hz, 1H), 6.81 – 6.69 (m, 2H), 5.14 (dd, J = 9.8, 7.4 Hz, 2H), 5.02 (td, J = 13.7, 10.5, 2.9 Hz, 2H), 4.80 – 4.73 (m, 2H), 4.74 (s, 2H), 4.46 (dt, J = 27.5, 10.3 Hz, 3H), 4.26 (dd, J = 22.4, 3.0 Hz, 2H), 4.06 (td, J = 8.5, 4.4 Hz, 3H), 3.89 – 3.76 (m, 4H), 3.75 (t, J = 2.4 Hz, 2H), 3.66 (dq, J = 7.8, 4.3, 3.5 Hz, 16H), 3.40 – 3.26 (m, 2H), 3.00 – 2.88 (m, 4H), 2.56 (t, J = 6.0 Hz, 2H), 1.45 (s, 9H).

^{13}C NMR (126 MHz, D_2O, extracted from HSQC) δ 121.43, 124.35, 125.65, 126.36, 125.65, 129.20, 129.23, 128.75, 84.13, 84.12, 66.86, 66.97, 63.11, 66.85, 67.89, 79.54, 79.55, 61.03, 66.44, 61.05, 61.06, 69.53, 69.52, 69.50, 70.21, 42.99, 42.99, 35.06, 27.18, 27.17, 27.18.

HRMS (EI, m/z) calculated for $C_{38}H_{58}N_6O_{15}SNa^+$ ($[M+Na]^+$): 893.3579, found 893.3595.

Preparation of compound 9

Compound **8** (52 mg, 0.060 mmol) was added into TFA/CH_2Cl_2 (10 mL, 1:1) and the solution was stirred for 2 h at room temperature. After being fully evaporated, the residue was purified by size-exclusion chromatography (Bio-Gel P2 fine; column 2.5 cm × 120 cm; flow rate 0.3 ml/min; elution with H_2O/ *n*-Butanol = 95/ 5). The fractions containing the product were pooled and freeze-dried to give compound **9** (48 mg, 0.059 mmol, 98 %) as a white solid.

^{1}H NMR (400 MHz, D_2O) δ 8.40 (s, 1H), 8.11 (s, 1H), 7.75 – 7.68 (m, 2H), 7.44 – 7.37 (m, 2H), 7.36 – 7.28 (m, 1H), 4.99 (dd, J = 9.8, 6.1 Hz, 2H), 4.87 (ddd, J = 11.9, 10.7, 3.0 Hz, 2H), 4.59 (s, 2H), 4.31 (dt, J = 20.7, 10.2 Hz, 2H), 4.16 – 4.06 (m, 2H), 3.91 (td, J = 7.6, 4.4 Hz, 2H), 3.74 – 3.60 (m, 5H), 3.60 – 3.48 (m, 17H), 2.31 (t, J = 6.7 Hz, 2H).

^{13}C NMR (101 MHz, D_2O) δ 180.11, 163.09, 162.74, 147.43, 143.78 , 129.19 , 128.82 , 125.65 , 124.36 , 121.38 , 117.70 , 114.80 , 84.10 , 79.48 , 69.48 , 69.45 , 69.43 , 69.17 , 68.85 , 67.98 , 67.93 , 67.84 , 66.94 , 66.83 , 66.77 , 66.74 , 63.04 , 61.00 , 37.52 .

HRMS (EI, m/z): calculated for $[C_{34}H_{50}N_6O_{15}SH]^+$ ($[M+H]^+$): 815.3128, found 815.3135.

Preparation of compound 10

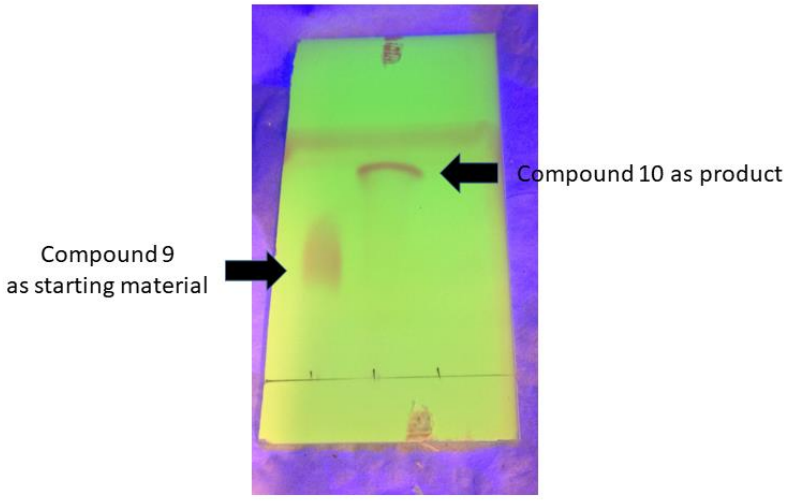
Compound **9** (22 mg, 0.027 mmol) was dissolved in anhydrous DMF (0.5 mL) and DiPEA (4.69 μl, 0.027 mmol) was added, followed by addition of TSTU (8.1 mg, 0.027

mmol). The resulting mixture was stirred for 30 mins, then TLC and HPLC showed that the starting material was fully converted (depicted in 3.4.4). To avoid hydrolysis, the crude **10** was used for labeling the protein directly.

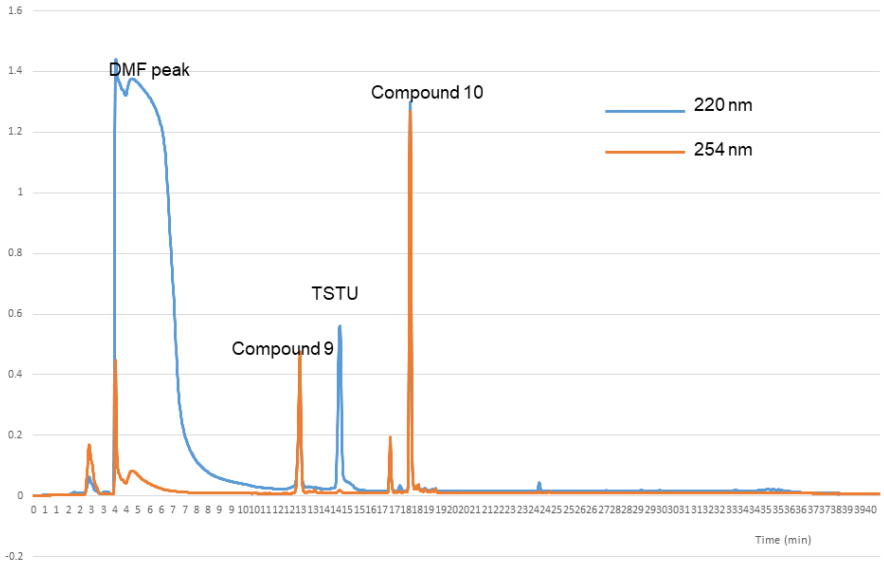
Preparation of multivalent TDG-conjugates 11 and 12.

A volume of 150 µL BSA (60 µM in 35 mM HEPES buffer, pH 7.0) was mixed with 5 µL of the coupling agent (compound **10** crude, 54 mM in DMF) was added into the solution and the reaction mixture was incubated at 4 °C. After 24 h and 48 h, additional volumes of 5 µL compound **10** crude were added. As to synthesis of conjugate **12**, the pH of the BSA solution was elevated to pH 9.0 using TEA before adding the conjugation agent. Conjugates **11** and **12** were isolated and rinsed with H_2O using VivaSpin 500 centrifugal concentrators (Sartorius Stedim Biotech, Goettingen, Germany) with an MWCO of 10 kDa. The protein concentration was determined by Bradford reagent (Carl Roth, Karlsruhe, Germany) according to manufacturer's instruction.

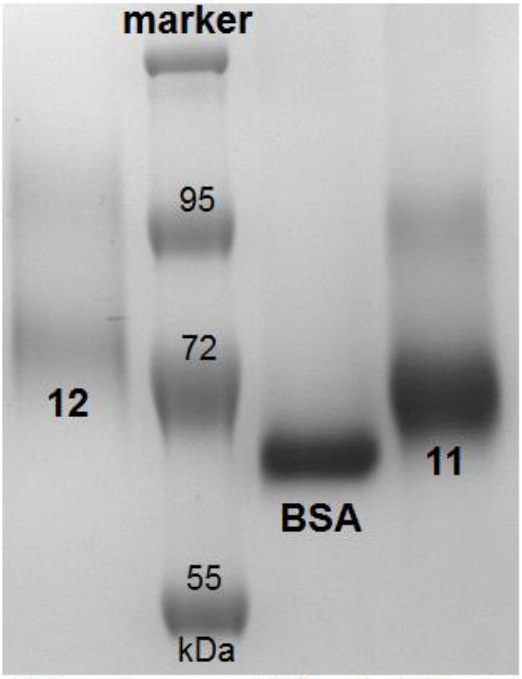
3.4.3 TLC and HPLC monitor the conversion of compound 9 into 10



EtOAc: CH_3OH: H_2O = 5: 2: 1



3.4.4 TDG-conjugates 11 and 12 confirmed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)



10% reducing gel (25 mA, 120 min)

TDG-conjugates **11** and **12**, separated on a reducing gel by SDS-PAGE. The utilized marker was PageRuler Prestained Protein Ladder (ThermoFisher Scienitific).

3.5 Reference

(1) Cecioni, S.; Imberty, A.; Vidal, S. Glycomimetics versus Multivalent Glycoconjugates for the Design of High A Ffi Nity Lectin Ligands. *Chem. Rev.* **2015**, *115* (1), 525–561.

(2) Barondes, S. H.; Cooper, D. N. W.; Gitt, M. A.; Leffler, H. Galectins. Structure and Function of a Large Family of Animal Lectins. *J. Biol. Chem.* **1994**, *269* (33), 20807–20810.

(3) Elola, M. T.; Blidner, A. G.; Ferragut, F.; Bracalente, C.; Rabinovich, G. A. Assembly, Organization and Regulation of Cell-Surface Receptors by Lectin-Glycan Complexes. *Biochem. J.* **2015**, *469* (1), 1–16.

(4) Nabi, I. R.; Shankar, J.; Dennis, J. W. The Galectin Lattice at a Glance. *J. Cell Sci.* **2015**, *128*, 2213–2219.

(5) Chaudhari, A. D.; Gude, R. P.; Kalraiya, R. D.; Chiplunkar, S. V. Endogenous Galectin-3 Expression Levels Modulate Immune Responses in Galectin-3 Transgenic Mice. *Mol. Immunol.* **2015**, *68* (2), 300–311.

(6) Maris, C.; Adanja, I.; Mercier, M. Le; Haene, N. D.; Decaestecker, C.; Baum, L.; Salmon, I. VEGFR1 and VEGFR2 Involvement in Extracellular Galectin-1- and Galectin-3-Induced Angiogenesis. *PLoS One* **2013**, *8* (6), e67029.

(7) Fortuna-costa, A.; Gomes, A. M.; Kozlowski, E. O.; Stelling, Mariana, P.; Pavão, M. S. Extracellular Galectin-3 in Tumor Progression and Metastasis. *Front. Oncol.* **2014**, *4*, Article 138.

(8) Funasaka, T.; Raz, A.; Nangia-Makker, P. Galectin-3 in Angiogenesis and Metastasis. *Glycobiology* **2017**, *24* (10), 886–891.

(9) Seetharaman, J.; Kanigsberg, A.; Slaaby, R.; Leffler, H.; Barondes, S. H.; Rini, J. M. X-Ray Crystal Structure of the Human Galectin-3 Carbohydrate Recognition Domain at 2.1-Å Resolution. *J. Biol. Chem.* **1998**, *273* (21), 13047–13052.

(10) Tejler, J.; Salameh, B.; Leffler, H.; Nilsson, U. J. Fragment-Based Development of Triazole-Substituted O-Galactosyl Aldoximes with Fragment-Induced Affinity and Selectivity for Galectin-3. *Org. Biomol. Chem.* **2009**, *7*, 3982–3990.

(11) Cumpstey, I.; Carlsson, S.; Leffler, H.; Nilsson, U. J. Synthesis of a Phenyl Thio-β-D-Galactopyranoside Library from 1,5-Difluoro-2,4-Dinitrobenzene: Discovery of Efficient and Selective Monosaccharide Inhibitors of Galectin-7. *Org. Biomol. Chem.* **2005**, *3*, 1922–1932.

(12) Leffler, H.; Barondes, S. H. Specificity of Binding of Three Soluble Rat Lung Lectins to Substituted and Unsubstituted Mammalian β-Galactosides. *J. Biol. Chem.* **1986**, *261* (22), 10119–10126.

(13) Andre, S.; Kaltner, H.; Furuike, T.; Nishimura, S.; Gabius, H. J. Persubstituted Cyclodextrin-Based Glycoclusters as Inhibitors of Protein-Carbohydrate Recognition Using Purified Plant and Mammalian Lectins and Wild-Type and Lectin-Gene-Transfected Tumor Cells as Targets. *Bioconjugate Chem.* **2004**, *15* (1), 87–98.

(14) Pieters, R. J.; Vrasidas, I.; Kaltner, H.; Kuwabara, I.; Liu, F.; Liskamp, R. M. J.; Gabius, H. Wedgelike Glycodendrimers as Inhibitors of Binding of Mammalian Galectins to Glycoproteins, Lactose Maxiclusters, and Cell Surface Glycoconjugates. *ChemBioChem* **2001**, *2*, 822–830.

(15) Marchiori, M. F.; Souto, D. E. P.; Bortot, L. O.; Pereira, J. F.; Kubota, L. T.; Cummings, R. D.; Dias-baruffi, M.; Carvalho, I.; Leiria, V. L. Synthetic 1,2,3-Triazole-Linked Glycoconjugates Bind with High Affinity to Human Galectin-3. *Bioorg. Med. Chem.* **2015**, *23* (13), 3414–3425.

(16) Tejler, J.; Tullberg, E.; Nilsson, U. J. Synthesis of Multivalent Lactose Derivatives by 1,3-Dipolar Cycloadditions: Selective Galectin-1 Inhibition. *Carbohydr. Res.* **2006**, *341*, 1353–1362.

(17) Vrasidas, I.; André, S.; Valentini, P.; Böck, C.; Lensch, M.; Kaltner, H.; Liskamp, R. M. J.; Gabius, H.; Pieters, R. J. Rigidified Multivalent Lactose Molecules and Their Interactions with Mammalian Galectins: A Route to Selective Inhibitors. *Org. Biomol. Chem.* **2003**, *1* (5), 803–810.

(18) Laaf, D.; Bojarova, P.; Pelantova, H.; Kren, V.; Elling, L.; Laaf, D.; Bojarová, P.; Pelantová, H.; Křen, V.; Elling, L. Tailored Multivalent Neo-Glycoproteins: Synthesis , Evaluation and Application of a Library of Galectin-3-Binding Glycan Ligands Tailored Multivalent Neo-Glycoproteins : Synthesis , Evaluation and Application of a Library of Galectin-3-Binding Glycan Liga. *Bioconjug. Chem.* **2017**, *28* (11), 2832–2840.

(19) Böcker, S.; Laaf, D.; Elling, L. Galectin Binding to Neo-Glycoproteins: LacDiNAc Conjugated BSA as Ligand for Human Galectin-3. *Biomolecules* **2015**, No. 5, 1671–1696.

(20) Mawas, F.; Niggemann, J.; Jones, C.; Corbel, M. J.; Kamerling, J. P.; Vliegenthart, J. F. G. Immunogenicity in a Mouse Model of a Conjugate Vaccine Made with a Synthetic Single Repeating Unit of Type 14 Pneumococcal Polysaccharide Coupled to CRM197. *Infect. Immun.* **2002**, *70* (9), 5107–5114.

(21) Cheng, K.; El-boubbou, K.; Landry, C. C. Binding of HIV-1 gp120 Glycoprotein to Silica Nanoparticles Modified with CD4 Glycoprotein and CD4 Peptide Fragments. *Appl. Mater. Interfaces* **2012**, *4*, 235–243.

(22) Jin, X.; Newton, J. R.; Montgomery-smith, S.; Smith, G. P. A Generalized Kinetic Model for Amine Modification of Proteins with Application to Phage Display. *Biotechniques* **2009**, *46* (3), 175–182.

(23) Patil, U. S.; Qu, H.; Caruntu, D.; O'Connor, C. J.; Sharma, A.; Cai, Y.; Tarr, M. A. Labeling Primary Amine Groups in Peptides and Proteins with N-Hydroxysuccinimidyl Ester Modified Fe3O4@SiO2 Nanoparticles Containing Cleavable Disulfide-Bond Linkers. *Bioconjug. Chem.* **2013**, *24* (9), 1562–1569.

(24) Mandal, S.; Nilsson, U. J. Tri-Isopropylsilyl Thioglycosides as Masked Glycosyl Thiol Nucleophiles for the Synthesis of S-Linked Glycosides and Glyco-Conjugates. *Org. Biomol. Chem.* **2014**, *12*, 4816–4819.

(25) Lowary, T. L.; Hindsgaul, O. Recognition of Synthetic O-Methyl,epimeric,and Amino Analogues of the Acceptor α-L-Fuc P-(1→2)-β-D-Galp-OR the Blood-Group A and B Gene-Specified Glycosyltransferases. *Carbohydr. Res.* **1994**, *251*, 33–67.

(26) Sung, S. R.; Han, S. C.; Jin, S.; Lee, J. W. Convergent Synthesis and Characterization of Dumbbell Type Dendritic Materials by Click Chemistry. *Bull. Korean Chem. Soc.* **2011**, *32* (11), 3933–3940.

(27) Seitz, O.; Kunz, H. HYCRON, an Allylic Anchor for High-Efficiency Solid Phase Synthesis of Protected Peptides and Glycopeptides. *J. Org. Chem.* **1997**, *62* (4), 813–826.

(28) Wang, H.; Huang, W.; Orwenyo, J.; Banerjee, A.; Vasta, G. R.; Wang, L. Design and Synthesis of Glycoprotein-Based Multivalent Glyco-Ligands for Influenza Hemagglutinin and Human Galectin-3. *Bioorg. Med. Chem.* **2013**, *21* (7), 2037–2044.

(29) Fischöder, T.; Laaf, D.; Dey, C.; Elling, L. Enzymatic Synthesis of N-Acetyllactosamine (LacNAc) Type 1 Oligomers and Characterization as Multivalent Galectin Ligands. *Molecules* **2017**, *22*, 1320–1335.

(30) Tietze, L. F.; Schroter, J. C.; Gabius, S.; Brinck, U.; Goerlach-graw, A.; Gabius, H. Conjugation of P-Aminophenyl Glycosides with Squaric Acid Diester to a Carrier Protein and the Use of Neoglycoprotein in the Histochemical Detection of Lectins. *Bioconjug. Chem.* **1991**, *2* (3), 148–153.

(31) Laaf, D.; Bojarová, P.; Mikulová, B.; Pelantová, H.; Křen, V.; Elling, L. Two-Step Enzymatic Synthesis of β-D-N-Acetylgalactosamine-(1→4)-D-N-Acetylglucosamine (LacdiNAc) Chitooligomers for Deciphering Galectin Binding Behavior. *Adv. Syntheis Catal.* **2017**, *359* (12), 2101–2108.

(32) Giorgi, M. E.; Agusti, R.; Lederkremer, R. M. De. Carbohydrate PEGylation, an Approach to Improve Pharmacological Potency. *Beilstein Joural Org. Chem.* **2014**, *10*, 1433–1444.

(33) Lundquist, J. J.; Toone, E. J. The Cluster Glycoside Effect. *Chem. Rev.* **2002**, *102* (2), 555–578.

(34) Pieters, R. J. Maximising Multivalency Effects in Protein-Carbohydrate Interactions. *Org. Biomol. Chem.* **2009**, *7*, 2013–2025.

(35) Leffler, H.; Carlsson, S.; Hedlund, M.; Qian, Y.; Poirier, F. Introduction to Galectins. *Glycoconj. J.* **2002**, *19* (7–9), 433–440.

(36) Hattum, H. Van; Branderhorst, H. M.; Moret, E. E.; Nilsson, U. J.; Leffler, H.; Pieters, R. J. Tuning the Preference of Thiodigalactoside- and Lactosamine-Based Ligands to Galectin-3 over Galectin-1. *J. Med. Chem.* **2013**, *56* (3), 1350–1354.

(37) Blanchard, H.; Bum-Erdene, K.; Matthew W Hugo. Inhibitors of Galectins and Implications for Structure-Based Design of Galectin-Specific Therapeutics. *Aust. J. Chem.* **2014**, *67* (12), 1763–1779.

(38) Öberg, C. T.; Leffler, H.; Nilsson, U. J. Inhibition of Galectins with Small Molecules. *Chimia (Aarau).* **2011**, *65* (1), 18–23.

(39) Mackinnon, A.; Chen, W.-S.; Leffler, H.; Panjwani, N.; Schambye, H.; Sethi, T.; Nilsson, U. J. Design, Synthesis, and Applications of Galectin Modulators in Human Health. In *Top med chem*; 2014; Vol. 12, pp 95–121.

(40) Kupper, C. E.; Böcker, S.; Liu, H.; Adamzyk, C.; Kamp, J. Van De; Lethaus, B.; Jahnen-dechent, W.; Neuss, S.; Müller-newen, G.; Elling, L. Fluorescent SNAP-Tag Galectin Fusion Proteins as Novel Tools in Glycobiology. *Curr. Pharm. Des.* **2013**, *19*, 5457–5467.

(41) Laaf, D.; Steffens, H.; Pelantova, H.; Kren, V.; Elling, L. Chemo-Enzymatic Synthesis of Branched N-Acetyllactosamine Glycan Oligomers for Galectin-3 Inhibition. *Adv. Syntheis Catal.* **2017**, *359* (22), 4015–4024.

(42) Salameh, B. A.; Leffler, H.; Nilsson, U. J. 3-(1,2,3-Triazol-1-Yl)-1-Thio-Galactosides as Small, Efficien, and Hydrolytically Stable Inhibitors of Galectin-3. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3344–3346.

(43) Delaine, T.; Collins, P.; MacKinnon, A.; Sharma, G.; Stegmayr, J.; Rajput, V. K.; Mandal, S.; Cumpstey, I.; Larumbe, A.; Salameh, B. A.; Kahl-knutsson, B.; Hattum, H. Van; Scherpenzeel, M. Van; Pieters, R. J.; Sethi, T.; Schambye, H.; Oredsson, S.; Leffler, H.; Blanchard, H.; Nilsson, U. J. Galectin-3-Binding Glycomimetics That Strongly Reduce Bleomycin-Induced Lung Fibrosis and Modulate Intracellular Glycan Recognition. *ChemBioChem* **2016**, *17* (18), 1759–1770.

(44) Ahmad, N.; Gabius, H. J.; André, S.; Kaltner, H.; Sabesan, S.; Roy, R.; Liu, B.; Macaluso, F.; Brewer, C. F. Galectin-3 Precipitates as a Pentamer with Synthetic Multivalent Carbohydrates and Forms Heterogeneous Cross-Linked Complexes. *J. Biol. Chem.* **2004**, *279* (12), 10841–10847.

(45) Lepur, A.; Salomonsson, E.; Nilsson, U. J.; Leffler, H. Ligand Induced Galectin-3 Protein Self-Association. *J. Biol. Chem.* **2012**, *287* (26), 21751–21756.

(46) Liese, S.; Netz, R. R. Influence of Length and Flexibility of Spacers on the Binding Affinity of Divalent Ligands. *Beilstein Joural Org. Chem.* **2015**, *11*, 804–816.

(47) Parera Pera, N.; Branderhorst, H. M.; Kooij, R.; Maierhofer, C.; Van Der Kaaden, M.; Liskamp, R. M. J.; Wittmann, V.; Ruijtenbeek, R.; Pieters, R. J. Rapid Screening of Lectins for Multivalency Effects with a Glycodendrimer Microarray. *ChemBioChem* **2010**, *11* (13), 1896–1904.

Chapter 4

Synthesis and evaluation of “hybrid” galectin inhibitors designed to simultaneously occupy two different binding sites

Abstract

Galectins have diverse functions and are involved in many biological processes because of their complex intra- and extracellular activities. Selective and potent inhibitors for galectins will be valuable tools to investigate the biological functions of these proteins. Therefore, we described the synthesis of galectin inhibitors with a potential "chelate effect". These compounds are designed to bind the two different binding sites on galectins simultaneously. In this chapter, a series of asymmetric "hybrid" compounds were prepared, which combined two galectin ligands 1) a substituted thiodigalactoside derivative and 2) a calixarene derivative. NMR spectroscopy was used to evaluate the interactions of these compounds with Galectin-1 and -3. In addition, cellular experiments were conducted to compare the cytotoxic effects of the hybrids with those of the calixarene derivative.

4.1 Introduction

As a subfamily of lectins, galectins can specifically recognize β-galactosides. They contain conserved carbohydrate-recognition domains (CRDs) consisting of about 130 amino acids that are responsible for the carbohydrate binding.[1] They also display an intriguing combination of intra- and extracellular activities and are involved in many biological process like cell growth, cell adhesion and signaling.[2,3] In addition, galectins also play a profound role in many diseases like inflammation, cancer and fibrosis.[4,5,6,7] To understand the action mechanism of galectins and their relevant therapeutic potential, there is a clear need for selective and potent inhibitors targeting different galectins.

The last two decades have witnessed a growing number of galectin inhibitors derived from carbohydrates and peptide.[8,9,10] The majority of the galectin inhibitors that were carbohydrate-based included modifications of lactose, a naturally galectin-binding ligand. Typically, those inhibitors competitively occupy the carbohydrate binding site. Furthermore, an obvious increase in binding affinity for certain galectins like galectin-3 (Gal-3) was found when an aromatic group was introduced at the C3 position of the galactose moiety.[11–13] In the meantime, a new scaffold, thiodigalactoside (TDG, K_d = 49 μM for Gal-3, K_d = 24 μM for Gal-1), produced more potent inhibitors in addition to enhancing glycolytic stability while maintaining a similar binding mode compared to lactose.[11,14]

Compared to carbohydrates, peptides are easier to synthesize. For carbohydrate-binding proteins, some peptides may prevent carbohydrate binding either by directly occupying the carbohydrate-binding site ("carbohydrate mimetic") or by simply occluding access to the carbohydrate to the binding site.[15–19] For peptide galectin inhibitors, both of the action modes were suggested but the exact binding modes were not always elucidated.[17,20] The best-studied peptide-based galectin inhibitor is anginex (βpep-33), a potent anti-angiogenic and anti-tumor peptide, which was shown to bind Gal-1, -2, -7, -8N and -9N, but not Gal-3, -4N, -4C and -9C.[5,21] In attempts to make

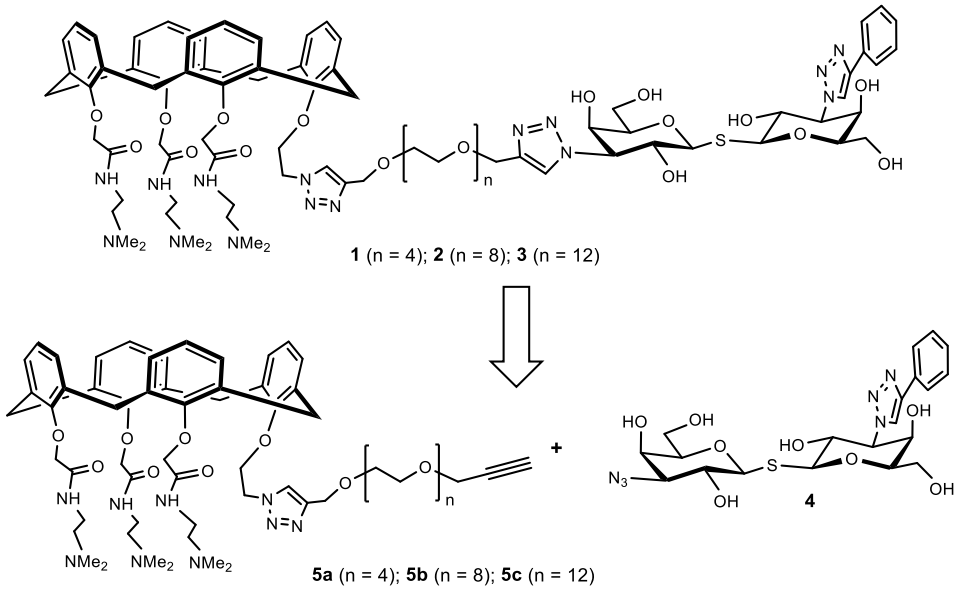
simplified relatives certain calix[4]arene derivatives such as calixarene 0118 were found to have an angiogenesis effect like anginex.[22,23] Considering the functional similarity, Gal-1 was assumed to be the target of 0118. Subsequent studies using ^{15}N-^{1}H HSQC NMR spectroscopy confirmed the hypothesis. It was shown that 0118 binds to Gal-1 at a site away from the galectin's carbohydrate binding site and the K_d is around 30 μM.[24]

In our group, a study conducted by Hilde van Hattum indicated that aromatic substituents at the TDG C3 by CuAAC (Copper-catalyzed azide–alkyne cycloaddition) strongly increased the affinity for certain galectins. In addition, among the symmetrical TDG derivatives the use of different aromatic groups showed that selectivity between Gal-1 and Gal-3 could be created.[25] So far, however, there has been little discussion about non-symmetrical TDG compounds for galectin inhibition. The recent published work on 0118 and the available data on TDG-based inhibitors lead us to consider the creation of hybrid or hetero-bivalent galectin inhibitors containing both a calix[4]arene ligand and a TDG ligand connected via a tether, to make a series of nonsymmetrical galectin inhibitors. These can potentially interact with the two different sites on the galectins simultaneously.[26] In this chapter, we set out to make a series of such non-symmetrical "hybrid" compounds, that combine the two core structures of TDG and 0118. Moreover, NMR spectroscopy was used to demonstrate the interactions between these compounds with Gal-1 and Gal-3. All compounds were characterized in terms of cytotoxicity and anti-proliferative effects on HUVEC cells (Human umbilical vein endothelial cells), HT-29 cells (Human colorectal adenocarcinoma grade II) and on Caki-2 cells (Human Renal Cancer Cell Line), along with TDG control compounds and calixarene control compounds.

4.2 Results and Discussion

4.2.1 Design and synthesis of hybrid galectin ligands

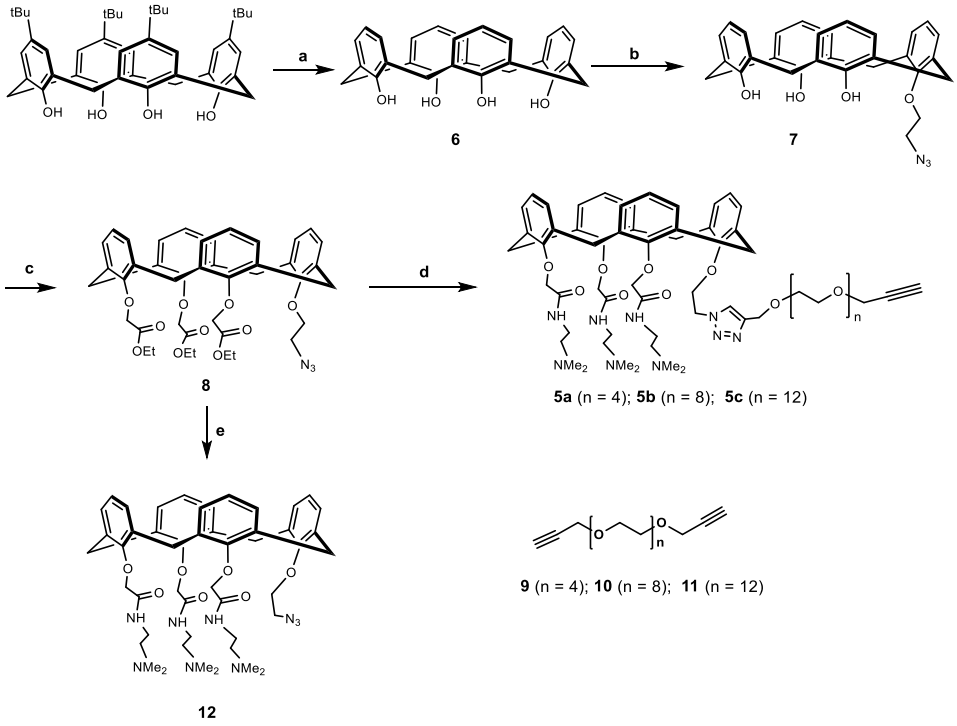
The synthesis of hybrid galectin ligands **1-3** was designed on the basis of the retrosynthetic analysis shown in Scheme 1. To obtain the target molecule, the precursors based on the calixarenes **5a-c** and the precursor based on thiodigalactoside (**4**) needed to be synthesized. The route towards **5a-5c** involved a direct CuAAC reaction of the calixarene analog with dipropargylated polyethylene glycol.



Scheme 1. Structure and retrosynthetic analysis of hybrid galectin ligands.

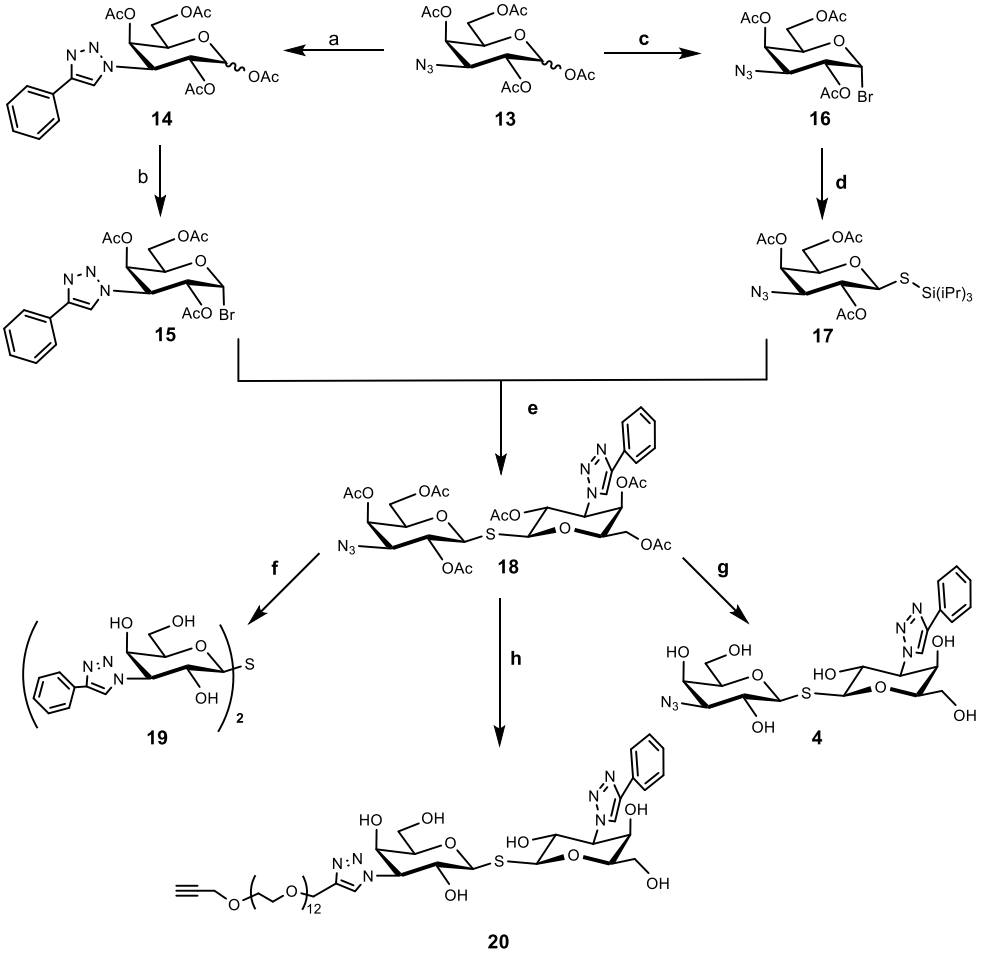
The synthesis of the calixarene precursor proceeded as shown in Scheme 2. Based on a published procedure, compound **7** was obtained through mono-alkylating **6** with 1-azido-2-iodoethane.[26] Then, the treatment with an excess of ethyl bromoacetate in the presence of K_2CO_3 gave compound **8**. It should be noted that the compound **8** was a mixture of isomers as revealed by the NMR spectrum with respect to the rotation of the various benzene rings in the calixarene. It was used in the next reaction without further purification. Previous research has revealed the synthesis of compounds **9-11**, and they were separately reacted with the compound **8.** Subsequently the resulting products were reacted with *N,N*-dimethylethylenediamine to yield **5a-c.**[27,28] These target compounds were obtained pure after preparative HPLC. The NMR and Mass spectrum showed they were in the "cone" conformation and they were used for further reactions.

To construct the required TDG derivatives, the tri-isopropylsilyl thio-glycoside **17** and the glycosyl halide **15** were needed. The TDG synthesis was based on a published method that is a one-pot desilylation and glycosyl thiol alkylation with a glycosyl halide (Scheme 3).[29] The synthesis benefited from work described in Chapter 3, in which the compound **18** was obtained. After removing the acetyl protecting group of **18**, the resulting crude **4** was used for "click" with **5a-c** without further purification. Besides this, CuAAC modification of compound **18** with phenylacetylene and compound **11** gave the corresponding products, that yielded control compounds **19** and **20** after deprotection.

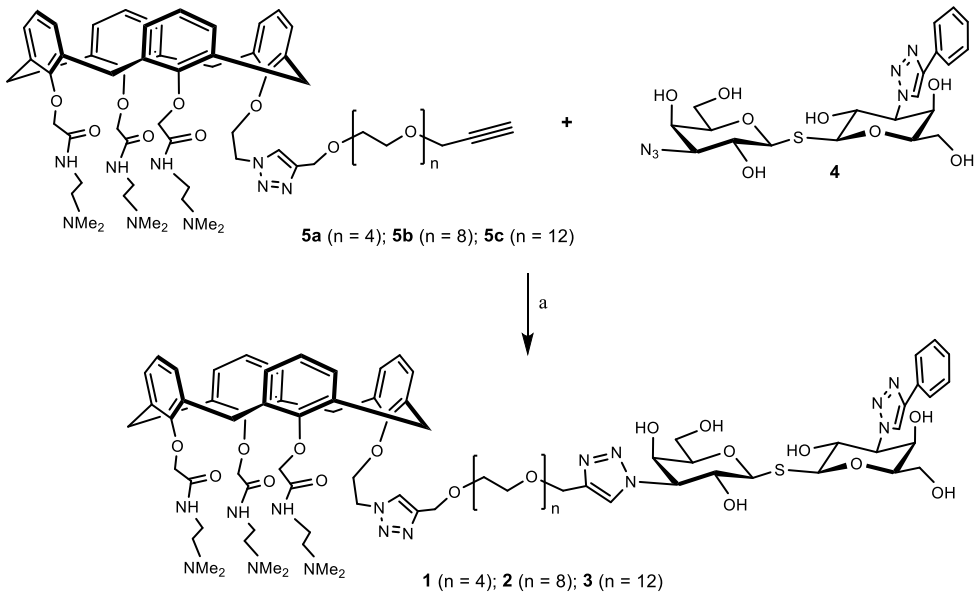


Scheme 2. Synthesis of targets **5a-5c.** Reaction Conditions: (a) $AlCl_3$, phenol, toluene, 98%; (b) 1-azido-2-iodoethane, NaOMe, CH_3CN, 85°C, 67%; (c) ethyl bromoacetate, K_2CO_3, CH_3CN, 85°C; 65%; (d) (i) compound **9-11**, CuI, CH_3CN, 80 °C, microwave; (ii) *N,N*-dimethylethylenediamine, 50°C, 10% in two steps (**5a**), 23% in two steps (**5b**), 27% in two steps (**5c**); (e) *N,N*-dimethylethylenediamine, 50 °C, 39 %.

Having assembled the important intermediates, calixarene parts **5a-5c** and TDG part **4**, the next objective became their CuAAC conjugation and purification of the resulting product to give the final hybrid compound (Scheme 4). Taking hybrid **3** as an example, compound **5c** and **4** were coupled using CuAAC under microwave irradiation to give the target compound (12 mg, in 22% yield) after purification by the preparative HPLC. Through analysis of the NMR spectrum, the calixarene part of hybrid **3** was confirmed to be in the “cone” conformation (i.e. identical to 0118)[5,28] and was used for the biological testing and binding studies.



Scheme 3. Synthesis of TDG component. Reaction conditions: (a) Phenylacetylene, $CuSO_4$, sodium ascorbate, DMF/H_2O, 80°C, microwave; (b) HBr, CH_2Cl_2, 79% for the two steps; (c) $TiBr_4$, CH_2Cl_2/EtOAc, 67%; (d) TIPSSH, K_2CO_3, CH_3CN, 30%; (e) TBAF, CH_3CN, 62%; (f) (i) Phenylacetylene, $CuSO_4$, sodium ascorbate, DMF/H_2O, 80 °C, microwave; (ii) NaOMe, CH_3OH, 23% in two steps; (g): NaOMe, CH_3OH; (h) (i) compound **11**, CuI, CH_3OH, 80 °C, microwave; (ii) NaOMe, CH_3OH, 5.0% in two steps.



Scheme 4. Synthesis of hybrid **1-3.** Hybrid **1-3** were synthesized in a similar fashion starting from the corresponding precursor. Reaction Conditions: CuI, CH_3OH, 80 °C, microwave, 90 min.

4.2.2 Conformation analysis of hybrid galectin ligands

Calix[4]arenes are not planar and can adopt four different stable conformations: cone, partial cone (paco), 1,2-alternate (1,2-alt), and 1,3-alternate (1,3-alt). As the phenolic hydroxyls in the lower rim are replaced by bulky groups, rotation of the aromatic rings will be prevented and stable conformations will be produced. A method to determine the conformation has been published,[30,31] which shows the ^{13}C-NMR shift of bridged methylene carbons connecting each pair of aromatic rings appears at 30 - 31 ppm when neighboring aromatic rings are oriented to the same side and at 37- 40 ppm when they point to the opposite side (Figure 1A). In the NMR assignments of the calixarene analogues we synthesized, the bridged methylene carbon can be observed and the chemical shift is at 30- 31 ppm, which implies that these compounds adopted the “cone” conformation. In addition, the NOESY spectrum was determined, with information about the spatial proximities of protons (Figure 1B). In the survey of all the cross-peaks encountered from the aromatic protons (displayed in red) of the calixarene moiety in hybrids **1-3**, NOE’s only appear from the bridged methylene protons, and none are observed protons from substituent group of the lower rim which are more remote in the cone conformation.

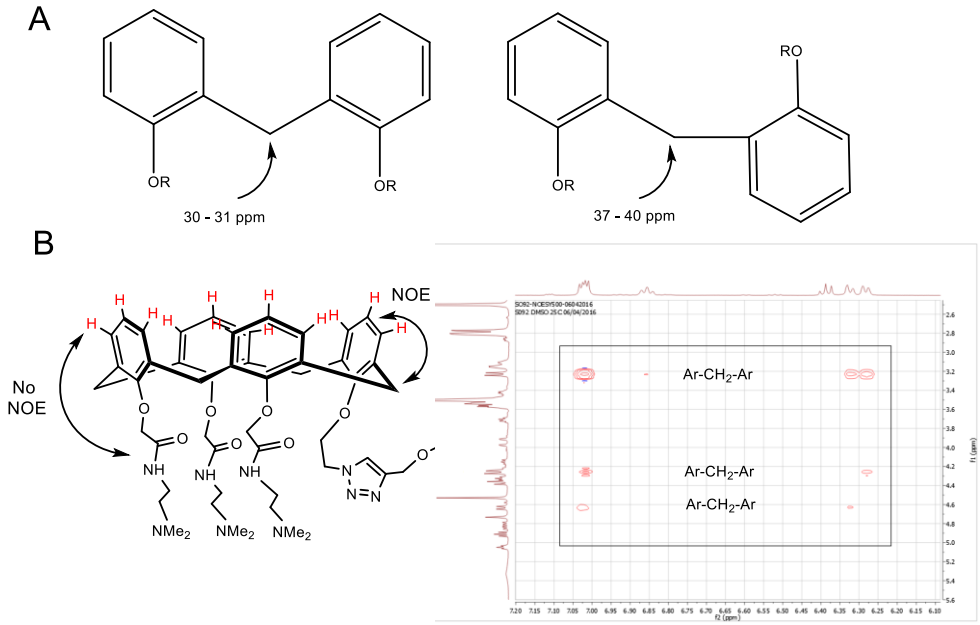


Figure 1. A: ^{13}C-NMR rule for the assignment of the orientation of neighboring aromatic rings in calixarenes; B: cross peaks encountered by aromatic protons (in red) of calixarene in NOESY. For clarity, other protons are not shown.

4.2.3 NMR spectroscopy to demonstrate the interaction between ligands with Gal-1 and Gal-3

Hybrid **1-3** and control compounds (**12** and **20**) were utilized as Gal-1 and -3 ligands in the NMR perturbation experiment by Dr. Hans Ippel (University of Maastricht) in collaboration with Professor Kevin Mayo (University of Minnesota). Figure 2 shows a ^{15}N-filtered NOESY spectrum of ^{15}N-labeled Gal-3 in the presence of hybrid **2** (protein: ligand = 1: 1.2), only showing ligand signals and aromatic protein signals and a ^{1}H spectra of the free hybrid **2**. We clearly see that the ligand occurs in a slow exchange between the free and the bound state, as shown by the corresponding 1D proton spectra at different ligand: protein ratios. In addition, the chemical shift changes of the TDG moiety, especially for its phenyl group proton signals ε and ζ (around 7.5 ppm), indicate that the TDG is involved in the interaction. The same phenomena also occur in the case of Gal-1, and in both cases the system is in slow exchange at least with respect to the TDG part. This result is in according with our previous work on symmetric TDG derivatives like compound **19** binding Gal-1 and Gal-3 in the nanomolar range. Here, the NMR results indicate a preference for one particular mode of binding of the asymmetric TDG compounds to Gal-3, because only one set of signals of bound ligand is observed in slow to intermediate exchange (exchange rate < 150 ms).

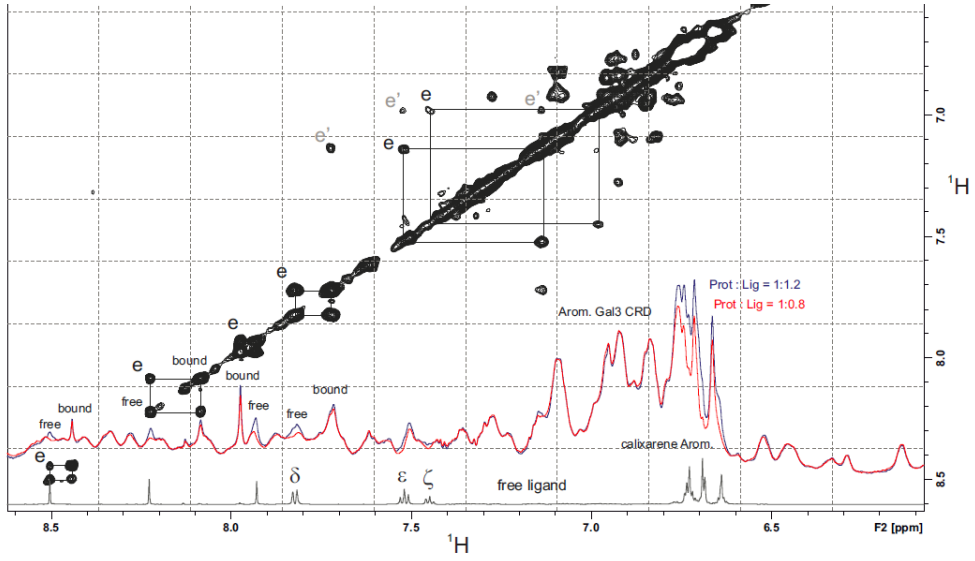


Figure 2. One-dimensional ^{1}H spectra and ^{15}N-filtered 2D NOESY spectra of Gal-3 CRD (50 μM) titrated with hybrid **2** at increasing molar ratios (red, protein: ligand = 1: 0.8; blue, protein: ligand = 1: 1.2; 150 ms mixing time). Cross peaks labelled *e* indicate exchange peaks between free and bound thiogalactoside resonances, whereas peaks labelled *e'* indicate NOE-mediated exchange peaks within the coupled ^{1}H spin system of the aromatic Phe ring in **2**. Assignments of proton resonances of **2** in the apo state are given in the bottom 1D spectrum, recorded at 50 μM of free ligand **2**.

To illuminate the perturbations of Gal-1 and -3 caused by the hybrid compounds, chemical shift perturbation degree maps are made, in which resonance positions of every amide backbone amino acid in the protein complex is shown relative to the resonance position of the corresponding apo state. A perturbation plot of a protein during a ligand titration provides a relatively good insight into which part of the protein the ligand interacts with. Take Gal-3 as an example, the titration experiments were conducted with hybrid **2** and **3** separately (protein: ligand = 1: 1.2) and the corresponding root-mean-square (RMS) delta values of backbone amide ^{1}H and ^{15}N chemical shifts upon complexation are shown in Figures 3A and 3B. It is interesting to note that in both cases the ligand can bind the protein, but the amplitudes and pattern of chemical shift perturbation pattern over the amino acid sequence of Gal-3 remain basically the same. This finding is unexpected and suggests that the PEG linker between the TDG and the calixarene moiety in the two ligands doesn’t alter the structural binding motif. A possible explanation for this might be that only one moiety (TDG or calixarene) interacts with the

protein which then leads to a same effect. To verify the hypothesis, we tested TDG control compound **20** and calixarene control compound **12** under the same conditions and the result is shown in Figure 3C and 3D.

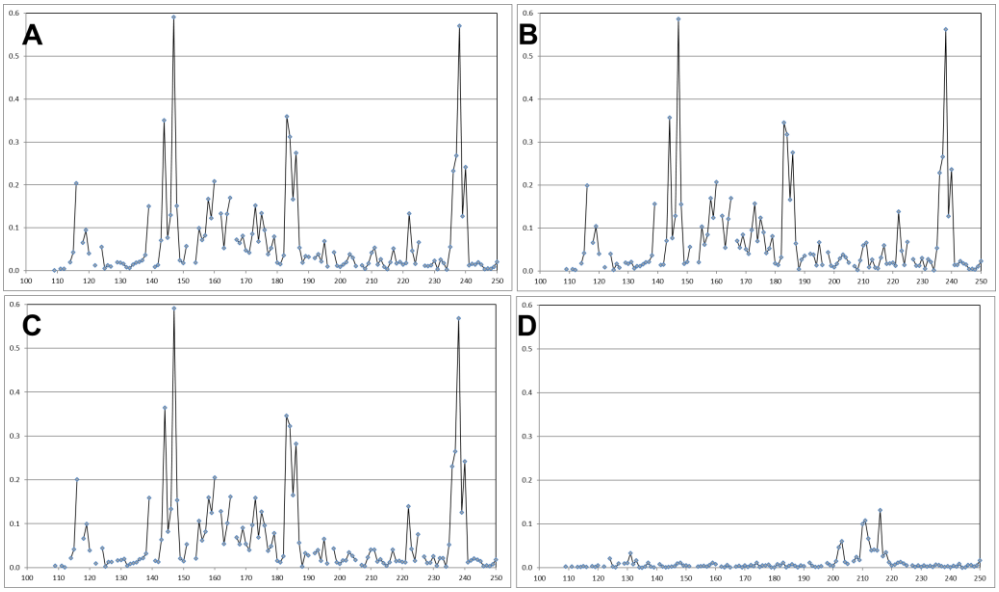


Figure 3. Perturbation degree changes observed for Gal-3 CRD (108-250, 50 μM) in the presence of ligands are shown vs the residue number of Gal-3. A: Gal-3 CRD in complex with hybrid **2** (Protein: ligand = 1: 1.2). B: Gal-3 CRD in complex with hybrid **3** (Protein: ligand = 1: 1.2). C: Gal-3 CRD in complex with compound **20** (Protein: ligand = 1: 1.2). D: Gal-3 CRD in complex with compound **12** (Protein: ligand = 1: 15). High perturbation values of delta chemical shifts reflect large structural changes in the protein after binding to the ligand.

In Figure 3C, compound **20** contains TDG and linker and it exhibits a similar interaction with Gal-3 compared to hybrids **2** and **3**, not only with respect to the residue number but also with respect to the perturbation degree. In contrast, compound **12,** containing the calixarene, only produced little perturbation on Gal-3 even when the ligand concentration was increased by more than 10 times (Figure 3D). Besides, the residue numbers involved in the interaction between compound **12** with Gal-3 were different from the TDG compound which is consistent with the reported notion that calixarenes bind at a different site of the galectins.[5] The same results were seen in the case of Gal-1. All data support the idea that the hybrid compounds interact with the galectins using mostly a single moiety, i.e. TDG. Moreover, compound **12** shows very little affinity for Gal-1 which is a surprising considering the reported data of its parent compound, calixarene 0118.

To summarize these results, hybrid **2** and **3** bind to Gal-1 and Gal-3 with high affinity, which corroborates our previous research.[25,32] However, we didn’t detect an expected synergy effect of TDG and calixarene, which means the binding mode of the hybrid compounds is not a chelate mode, at least not to a large extent. It was shown that the TDG is the most important contributor for binding. In comparison, the calixarene moiety, both as a separate molecule or as a part of hybrid compound showed less strong binding. However, NMR results still show specific binding of the calixarene part at the backside of Gal-3 (residues 200-220) although with limited affinity, as shown by the need for up to 80-fold excess substrate to reach saturated binding and by the simultaneous presence of NMR shifts in a fast exchange rate scale.

4.2.4 Cytotoxicity studies

Usually calixarene analogues can inhibit the cell proliferation effectively but there is little relevant research about TDG analogues. In this regard, evaluation of the hybrid compounds in terms of cytotoxicity and inhibition of cell proliferation is of interest. In collaboration with the medical oncology group of Professor Arjan Griffioen (VU University Medical Center), all the compounds were assessed by employing a colorimetric cell viability assay. Dose-response curves obtained after incubation of Caki-2, HT-29 and HUVECs cells with increasing concentrations of the new compounds are expressed as the percentage of cell survival relative to control with vehicle only.

As showed in Figure 4, all of compounds carrying a calixarene unit (compound **12**, hybrid **2** and hybrid **3**) were found to inhibit cell survival. They were at least equipotent to parent 0118, while one of them (compound **12**) did inhibit cell survival more potently. As for the TDG derivatives (compounds **19** and **20**), they did not exhibit any effect on cell viability. Another important finding is that hybrid **2** and **3** exhibited minimal killing effect on HUVECs (Figure 4A), a non-cancerous cell line, whereas the viability of cancer cells (Caki-2 and HT-29) was inhibited greatly by them (Figure 4B and 4C). Taken together, these results clearly show that the hybrid compounds contain a fully functional calixarene moiety that exhibits its effects on cell viability. Interestingly, it seems that the linkage to the TDG unit leads to a reduced toxicity effect on the non-cancerous cells.

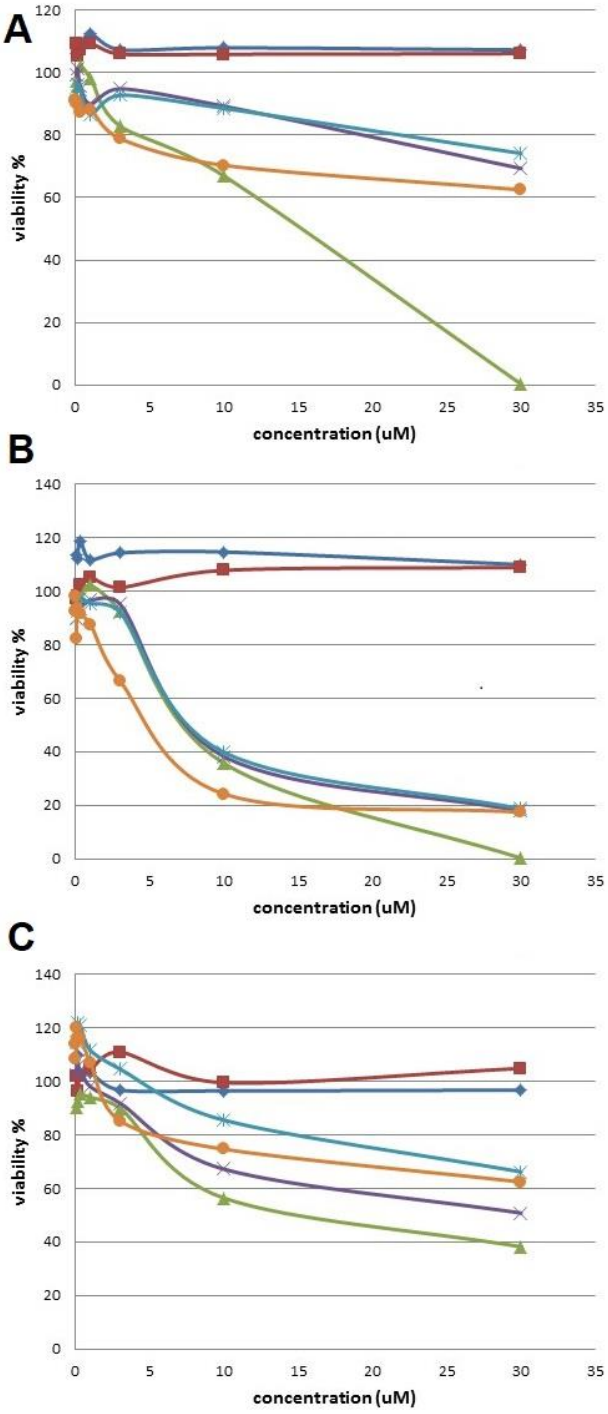

A
viability %
concentration (uM)
B
viability %
concentration (uM)
C
viability %
concentration (uM)

Figure 4. Dose-response curves obtained after incubation of HUVECs (A), HT-29 (B) and Caki-2 (C) with increasing concentrations of the new compounds. Blue curve: compound **20**, Red curve: compound **19**, Green curve: compound **12**, Purple curve: hybrid **2**, Light blue curve: hybrid **3**, Orange curve: calixarene 0118.

4.3 Conclusion

We here describe the design and synthesis of a series of galectin inhibitors designed to simultaneously occupy two different binding sites. TDG derivatives have been validated to be valuable inhibitors for galectin research, and they target the carbohydrate recognition domain. By comparison, the calixarene 0118 whose discovery was inspired by the peptide-based galectin inhibitor (anginex), is reported to interact with galectins at a binding site that is distinct from that of TDG. We thus report on the synthesis of a non-symmetrical TDG conjugated to a calixarene that yielded potentially hetero bivalent or hybrid ligands for galectins. To obtain the desired compounds, TDG and calixarene containing precursors were synthesized separately and linked through CuAAC. Using our compounds, NMR spectroscopy was used to study the interaction between the hybrid ligands and Gal-1 and Gal-3. The result clearly showed that compounds strongly bound to the galectins. It was also clear that the galectin binding mainly resulted from the TDG moiety and the desired “chelating” enhanced binding effect involving the calixarene moiety was not achieved. In addition, cytotoxicity studies were done to evaluate proliferation inhibition of the hybrid ligands. By comparing with parents TDG and calixarene 0118, we found that the hybrid compounds affected cell viability greatly and this time clearly through the calixarene moiety. Interestingly, the killing effect on HUVECs was less than for the other two cell lines when using compounds carrying a TDG moiety, suggesting that combining a TDG and a calixarene provides selectivity for the killing of cancerous instead of non-cancerous cells. Overall it is clear that both components of the hybrid compounds are fully functional as evidenced by galectin binding and viability studies. On the binding side, the modified calixarene did not help the binding, but the effects on biological systems are intriguing and may involve other proteins in addition to galectins which may warrant further study.

4.4 Experimental section

4.4.1 Reagents and general methods

TLC analysis was performed on Merck precoated silica gel 60 F-254 plates. Spots were visualized with UV-light, ninhydrin stain (1.5 g ninhydrin and 3.0 mL acetic acid in 100 mL n-butanol), Potassium permanganate (1.5 g $KMnO_4$, 10 g K_2CO_3, and 1.25 mL 10% NaOH in 200 mL H_2O) and sulfuric acid (10 % sulfuric acid in methanol). Column chromatography was performed using Silica-P Flash silica gel (60 Å, particle size 40 – 63 µm) from Silicycle (Canada). Microwave reactions were carried out in a Biotage Initiator (300 W) reactor. Lyophilization was performed on a Christ Alpha 1-2 apparatus. ^{1}H and ^{13}C NMR

spectroscopy was carried on an Agilent 400-MR spectrometer operating at 400 MHz for H^1 and 101 MHz for ^{13}C. HSQC, TOCSY and NOESY (500 MHz) were performed with a VARIAN INOVA-500. Complexes between uniformly Gal3-[15N] CRD and various TDG and Calixarene ligands were studied on a Bruker Avance III HD 700 MHz spectrometer, equipped with a TCI cryoprobe. Amide proton assignments of apo Gal-3 CRD were taken from a previous research.[33] Assignments of protein signals in the 1: 1 complex between hybrid **3** and Gal-3 were confirmed by means of a 3D ^{15}N-edited NOESY spectrum (150 ms mixing time). Chemical shifts of apo ligand resonances were derived from a combination of 2D correlation spectra, whereas signals of bound ligand were assigned through its exchange peaks in a ^{15}N-filtered NOESY spectrum (150 ms mixing time) at 30 ^oC.

Electrospray mass spectrometry were performed on a Shimadzu LCMS QP-8000. High resolution mass spectrometry (HRMS) analysis was recorded using a Bruker ESI-Q-TOF II. Analytical LC-MS (electrospray ionization) was performed on Thermo-Finnigan LCQ Deca XP Max using same buffers and protocol as described for analytical HPLC.

Analytical HPLC was performed on a Shimadzu-10AVP (Class VP) system using a Phenomenex Gemini C18 column (110 Å, 5 µm, 250 × 4.60 mm) at a flow rate of 1 mL min. The used buffers were 0.1 % trifluroacetic acid in CH_3CN: H_2O = 5: 95 (buffer A) and 0.1 % trifluroacetic acid in CH_3CN: H_2O = 95: 5 (buffer B). Runs were performed using a standard protocol: 100% buffer A for 2min, then a linear gradient of buffer B (0-100% in 38 min) and UV-absorption was measured at 214 and 254 nm. Purification using preparative HPLC was performed on an Applied Biosystems workstation with a Phenomenex Gemini C18 column (10 µm, 110 Å, 250 × 21.2 mm) at a flow rate of 6.25 mL min^{-1}. Runs were performed by a standard protocol: buffer A for 5 min followed by a linear gradient of buffer B (0 – 100% in 70 min) with the same buffer as described for the analytical HPLC.

4.4.2 Synthetic procedures and compound characterization

Preparation of crude 4: 3-(4-phenyl-1H-1,2,3-triazol)-3'-azido-thiodigalactoside

NaOMe (40 mg, 2.5 mmol) was added to a solution of **18** (120 mg, 0.16 mmol) in CH_3OH (5.0 mL) and the mixture was stirred for 6 h at room temperature. The solution was neutralized with DOWEX-H^+ resin, filtered, and evaporated. Crude **4** (85 mg) was obtained as a white solid and used in the next step without further purification.

25,26,27,28-tetrahydroxycalix[4]arene (6)

P-tert-Butylcalix[4]arene (5.0 g, 7.70 mmol) was suspended in toluene (50 mL), then $AlCl_3$ (8.2 g, 61.60 mmol) and phenol (5.8 g, 61.60 mmol) was added. The reaction mixture was stirred overnight, and then 1 N HCl (50 mL) was added slowly. After 10 min the reaction mixture was extracted by CH_2Cl_2 (3 × 100 mL), and the combined organic phase

was concentrated. CH_3OH (50.0 mL) was added to the residues to make a slurry. After filtration, compound **6** (3.2 g, 7.60 mmol, 98 %) was collected as a white solid and used for next step without further purification.

^{1}H NMR (400 MHz, $CDCl_3$) δ 10.17 (d, *J* = 1.2 Hz, 4H, OH), 7.04 (dd, J = 7.5, 1.2 Hz, 8H, Ar-H), 6.71 (td, J = 7.6, 1.2 Hz, 4H, Ar-H), 4.23 (m, 4H, CH_2), 3.51 (m, 4H, CH_2).

^{13}C NMR (101 MHz, $CDCl_3$) δ 148.77, 128.98, 128.24, 122.29, 122.24, 31.88, 31.71, 31.46, 31.31.

HRMS (EI, m/z): calculated for $C_{28}H_{24}O_4H^+$([M+H]$^+$): 425.1747, found 425.1729.

25,26,27-Trihydroxy-28-(2'-azidoethoxy) calix[4]arene (7)

To **6** (2.8 g, 6.70 mmol) dissolved in CH_3CN (50 mL) NaOMe (434 mg, 8.03 mmol) was added. The mixture was refluxed for 30 min and 1-azido-2-iodoethane (1.3 g, 6.68 mmol) was added. The mixture was refluxed overnight and the reaction was monitored by TLC. An additional amount of NaOMe (200 mg, 3.70 mmol) was added and the mixture was refluxed for 1 day. Then the mixture was worked up by evaporating the solvent. Then CH_2Cl_2 (100 mL) was added to the residue and the mixture was washed with water (3 × 50 mL). After drying with Na_2SO_4, the organic solvent was evaporated and the residue was purified by silica column chromatography using hexanes: EtOAc = 5: 1. The target compound **7** was obtained as a white solid (1.3 g, 2.60 mmol, 39%) and it was used in the next step. Unreacted starting material (1.2 g, 2.80 mmol, 42%) was also retrieved.

^{1}H NMR (400 MHz, $CDCl_3$) δ 9.64 (s, 1H, OH), 9.14 (s, 2H, OH), 7.10 (d, J = 7.7 Hz, 2H, Ar-H), 7.09 – 7.00 (m, 6H, Ar-H), 6.90 (dd, J = 7.9, 7.3 Hz, 1H, Ar-H), 6.70 (tt, J = 7.5, 1.2 Hz, 3H, Ar-H), 4.40 (d, J = 13.0 Hz, 2H, Ar_2CH_2), 4.30 (d, J = 13.8 Hz, 2H, Ar_2CH_2), 4.25 (dd, J = 5.6, 4.3 Hz, 2H, OCH_2), 4.05 (dd, J = 5.6, 4.3 Hz, 2H, CH_2N_3), 3.51 (d, J = 2.5 Hz, 2H, Ar_2CH_2), 3.48 (d, J = 3.2 Hz, 2H, Ar_2CH_2).

^{13}C NMR (101 MHz, $CDCl_3$) δ 151.04, 150.92, 149.28, 134.15, 129.66, 129.06, 128.91, 128.80, 128.54, 128.49, 128.24, 126.62, 122.08, 121.07, 74.42, 51.38, 32.01, 31.45.

HRMS (EI, m/z): calculated for $C_{30}H_{27}N_3O_4Na^+$([M+Na]$^+$): 516.1894, found 516.1851.

25,26,27-Tri[(ethoxycarbonyl)methoxy]-28-(2'-azidoethoxy) calix[4]arene (8)

To a solution of **7** (1.3 g, 2.60 mmol) in CH_3CN (50 mL) K_2CO_3 (1.8 g, 13.04 mmol) was added. The mixture was stirred for 2 h and then an excess of ethyl bromoacetate (2.6 g, 15.82 mmol) was added. The mixture was heated to 85 °C and refluxed overnight while monitoring the reaction by TLC. After removing the solvent, the residue was taken up in

CH_2Cl_2 (100 mL) and then the organic part was washed with H_2O (3 × 100 mL). After separation and drying with Na_2SO_4, the organic layer was evaporated and the impure compound was purified by column chromatography (PE: EtOAc = 5:1 → 4:1). The product (1.7 g, 2.2 mmol) was obtained as colorless oil. According to ^{1}H NMR, COSY, HSQC and HMBC, this sample contains approx. 75% of the compound **8** (corresponding to 1.275 g, 1.7 mmol, 65 %), which is the "cone" conformation and 25% of other isomers. The solid was used in the next step without further purification.

^{1}H NMR (400 MHz, $CDCl_3$, extracted from HSQC) δ 6.84 (dd, J = 6.9, 1.5 Hz, 4H), 6.75 (ddd, J = 8.2, 6.6, 4.0 Hz, 2H), 6.53 – 6.44 (m, 6H), 6.30 – 6.22 (m, 1H), 4.82 (s, 1H), 4.80 (d, J = 8.7 Hz, 3H), 4.69 (d, J = 16.0 Hz, 2H), 4.63 (d, J = 13.5 Hz, 2H), 4.52 (d, J = 16.1 Hz, 2H), 4.33 – 4.19 (m, 7H), 4.15 (t, J = 6.4 Hz, 2H), 3.96 (t, J = 6.4 Hz, 2H), 3.25 (d, J = 13.0 Hz, 3H), 3.20 (s, 1H), 1.31 (td, J = 7.1, 5.2 Hz, 9H).

HRMS (EI,m/z): calculated for $C_{42}H_{45}N_3O_{10}Na^+$ ([M+Na]$^+$):774.2997, found 774.2963.

25,26,27-Tris-*N*-(*N,N*-dimethyl-2-aminoethyl)carbamoylmethoxy-28-(2'-azidoethoxy) calix[4]arene (12)

To the compound **8** (350 mg) was added under nitrogen *N,N*-dimethylethylenediamine (7.0 mL). The mixture was stirred for 24 h at 50 °C. The excess of *N,N*-dimethylethylenediamine was removed by evaporation under reduced pressure. The residue was dissolved in CH_3CN (20 mg/mL) and purified by preparative HPLC using the standard protocol. Fractions containing the product (t_R = 35 min, broad peak) were pooled and the water was removed by freeze-drying to obtain the pure compound **12** as a clear oil (160 mg, 0.18mmol, 39%).

^{1}H NMR (500 MHz, DMSO-d_6) δ 6.94 (dd, J = 7.6, 4.3 Hz, 4H), 9.99 – 9.95 (m, 3H), 8.52 (q, J = 7.1, 6.6 Hz, 3H), 6.80 (t, J = 7.5 Hz, 2H), 6.40 (t, J = 7.5 Hz, 2H), 6.36 – 6.28 (m, 3H), 4.66 (d, 2H), 4.59 (d, J = 13.5 Hz, 2H), 4.40 (d, J = 14.0 Hz, 2H), 4.32 (d, J = 13.4 Hz, 2H), 4.28 (d, J = 14.0 Hz, 2H), 4.08 (d, J = 6.2 Hz, 2H), 3.89 (t, J = 6.2 Hz, 2H), 3.55 (dt, J = 13.0, 6.4 Hz, 6H), 3.23 (d, J = 5.7 Hz, 2H), 3.20 (d, J = 6.5 Hz, 8H), 2.81 (d, J = 6.9 Hz, 18H), .

^{13}C NMR (126 MHz, DMSO-d_6, extracted from HSQC) δ 129.30, 123.07, 123.03, 128.39, 74.23, 30.90, 74.03, 30.91, 74.03, 71.75, 51.18, 34.42, 30.91, 55.87, 42.50, 39.65

HRMS (EI,m/z): calculated for $C_{48}H_{63}N_9O_7H^+$ ([M+H]$^+$): 878.4923, found 878.4934.

Tetra(ethyleneglycol)-di(2-propinyl) (9)

To a solution of tetraethylene glycol (8.65 mL, 0.050 mol) in THF (50 mL) was added NaH (60% oil dispersion, 2.40 g, 0.060 mol) portionwise at 0 °C, and the reaction was

stirred at 0 °C for 1 h. Propargyl bromide (7.14 g, 0.060 mol) was added to the reaction solution, and the reaction was stirred at room temperature overnight. The reaction was quenched by adding ice cold water, and the mixture was extracted with CH_2Cl_2 (3 × 100 mL), washed with brine, dried over Na_2SO_4. After filtration, the solvent was removed *in vacuo*, the residue was purified by column chromatography (hexanes: acetone = 4: 1→ 3: 1) to afford dipropargylated compound **9** as a colorless oil (2.2 g, 0.008 mmol, 16%). In the meantime, monoproparylated tetraethylene glycol (5.3g, 0.023 mmol, 46%) as a byproduct was isolated for the synthesis of other likers.

^{1}H NMR (400 MHz, $CDCl_3$) δ 4.19 (d, J = 2.4, 4H), 3.75 – 3.57 (m, 16H), 2.41 (t, 1H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 79.64, 74.44, 70.50, 70.56, 70.38, 69.09 , 58.37.

HRMS (EI, m/z): calculated for $C_{14}H_{22}O_5Na^+$ ($[M+Na]^+$): 293.1359, found 293.1349.

Monopropargyl(tetraethylene glycol) tosylate

Monopropargyl-tetraethylene glycol (2.1 g, 0.009 mmol) obtained from the preparation of compound **9** was dissolved in CH_2Cl_2 (15 mL) and then p-toluenesulfonyl chloride (2.2 g, 11.5 mmol) and trimethylamine (2.5 mL, 0.018 mmol) were added. The reaction was stirred at r.t. for 16 h. The solvent was removed *in vacuo*, and the residue was purified by column chromatography (hexanes: EtOAc = 1: 1) to afford glycol tosylate as a yellow oil (2.9 g, 0.0075 mmol, 83 % yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.80(d, 2H), 7.36 – 7.30 (m, 2H), 4.19 (d, 2H), 4.18 – 4.14 (m, 2H), 3.70 – 3.51 (m, 14H), 2.45 (s, 3H), 2.40 (t, 1H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 144.91, 133.19, 129.95, 128.13, 77.36 , 74.65 , 70.91 , 70.75 , 70.70 , 70.56 , 69.39 , 69.27 , 68.84 , 58.55 , 21.79 .

Octa(ethyleneglycol)-di(2-propinyl) (10)

NaH (60% oil dispersion, 140 mg, 3.50 mmol) was added to a solution of monoproparyl-tetraethylene glycol (270 mg, 1.16 mmol) in THF (10.0 mL) and stirred for 1 h, followed by adding a solution of monopropargyl(tetraethylene glycol) tosylate (448 mg, 1.16 mmol) in THF (2.0 mL). The resulting mixture was stirred at r.t. for 24 h, and quenched by adding ice cold water, and the mixture was extracted with CH_2Cl_2 (3 × 100 mL), washed with brine, dried over Na_2SO_4. After filtration, the solvent was removed *in vacuo*, the residue was purified by column chromatography (EtOAc: CH_3OH = 10: 1) to afford dipropargylated compound **10** as a colorless oil (462 mg, 1.03 mmol, 89 %).

^{1}H NMR (400 MHz, $CDCl_3$) δ 3.72 – 3.55 (m, 32H), 4.17 (d, J = 2.3 Hz, 4H), 2.41 (t, 1H).

^{13}C NMR (101 MHz, CDCl3) δ 79.71, 74.64, 70.64, 70.61, 70.60, 70.43, 69.15, 58.44.

HRMS (EI, m/z): calculated for $C_{22}H_{38}O_9Na^+$ ([M+Na]$^+$): 469.2408, found: 469.2420.

4,7,10,13,16,19,22,25,28,31,34,37,40-tridecaoxatritetraconta-1,42-diyne (11)

NaH (60% oil dispersion,340 mg, 8.50 mmol) was treated to a solution of tetraethylene glycol (276 mg, 1.42 mmol) in THF (10.0 mL) and stirred for 1h, followed by adding a solution of monopropargyl(tetraethylene glycol) tosylate (1.2 g, 3.11 mmol) in THF (2.0 mL). The resulting mixture was stirred at r.t. for 24 h, and quenched by adding ice cold water, and the mixture was extracted with CH_2Cl_2 (3 × 100 mL), washed with brine, dried over Na_2SO_4. After filtration, the solvent was removed *in vacuo*, the residue was purified by column chromatography (EtOAc: CH_3OH = 10: 1) to afford dipropargylated compound **11** as a colorless oil (620 mg, 1.00 mmol, 70%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 4.18 (d, J = 2.4 Hz, 4H), 3.69 – 3.60 (m, 48H), 2.41 (s, 1H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 79.63, 74.50, 70.57, 70.55, 70.53, 70.37, 69.07, 58.36.

HRMS (EI, m/z): calculated for $C_{30}H_{54}O_{13}H^+$ ([M+H]$^+$): 623.3637, found: 623.3691.

25,26,27-Tris-*N*-(*N*,*N*-dimethyl-2-aminoethyl)carbamoylmethoxy-28-{2'-[4-(2,5,8,11,14-pentaoxaheptadec-16-yn-1-yl)-1H-1,2,3-triazol]ethoxy}calix[4]arene (5a)

Compound **9** (99 mg, 0.37 mmol) and **8** (175 mg, 0.37 mmol) were dissolved in CH_3CN (5.0 mL) and then CuI (44.3 mg, 0.23 mmol) was added to the solution. The resulting mixture was heated under microwave irradiation at 80 °C for 90 min. The mixture was concentrated *in vacuo*, and then CH_2Cl_2 (5.0 mL) was added. A clear solution was obtained after centrifugation, which was concentrated *in vacuo* to afford the crude triazole (260 mg). To the crude (260 mg) *N*,*N*-dimethylethylenediamine (7.0 mL) was added under nitrogen. The mixture was stirred for 24 h at 50 °C. The excess of *N*,*N*-dimethylethylenediamine was removed by evaporation under reduced pressure. The residue was dissolved in CH_3CN (20 mg/mL) and purified by preparative HPLC using the standard protocol. Fractions containing the product (t_R = 40 min) were pooled and the water was removed by freeze-drying to obtain the pure compound **5a** as a clear oil (45 mg, 0.039mmol, yield 10 % in two steps).

^{1}H NMR (400 MHz, DMSO-d_6) δ 10.00 (d, J = 35.1 Hz, 3H), 8.60 – 8.36 (m, 3H), 8.17 (s, 1H), 7.05 – 6.91 (m, 4H), 6.89 – 6.78 (m, 2H), 6.35 (m, 6H),4.41 – 4.26 (m, 9H), 4.24 (s, 2H), 4.13 (d, J = 2.4 Hz, 1H), 3.26 – 3.19 (m, 6H), 3.19 – 3.07 (m, 5H), 2.81 (s, 13H), 2.77 (s, 6H), 5.05 (q, J = 6.9 Hz, 2H), 4.73 (d, J = 4.1 Hz, 2H), 4.64 (d, J = 13.4 Hz, 2H), 4.53 (d, J = 4.4 Hz, 2H), 3.60 – 3.39 (m, 16H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 170.69 , 169.56 , 159.15 , 158.82 , 157.24 , 156.28 , 154.70 , 154.65 , 144.19 , 144.15 , 135.90 , 135.80 , 135.39 , 135.34 , 133.57 , 133.54 , 133.37 , 133.30 , 129.49 , 129.27 , 129.23 , 128.35 , 124.87 , 123.44 , 123.20 , 123.09 , 77.53 , 74.36 , 74.14 , 71.89 , 70.16 , 70.13 , 70.11 , 70.07 , 70.03 , 69.88 , 69.45 , 69.37 , 68.92 , 63.88 , 63.84 , 57.90 , 56.16 , 56.13 , 55.90 , 49.37 , 49.32 , 42.85 , 42.82 , 40.25 , 40.05 , 34.30 , 34.26 , 30.68 .

HRMS (EI, m/z): calculated for $C_{62}H_{85}N_9O_{12}H^+$ ([M+H]$^+$): 1148.6390, found 1148.6360.

Preparation of hybrid 1

Compound **5a** (45 mg, 0.039 mmol) and **4** (20 mg, 0.039 mmol) were dissolved in CH_3OH (3.0 mL) and then CuI (7.5 mg, 0.039 mmol) was added into the solution. The resulting mixture was heated under microwave irradiation at 80 °C for 90 min. The mixture was concentrated *in vacuo*, and then CH_2Cl_2 (5.0 mL) was added. A clear solution was obtained after centrifugation, which was concentrated *in vacuo* to afford the crude compound. The crude was dissolved in CH_3CN (20 mg/mL) and purified by preparative HPLC using the standard protocol. Fractions containing the product (t_R = 35 min) were pooled and the buffer was removed by freeze-drying to obtain the pure hybrid **1** as an off-white foam (4.0 mg, 0.0024 mmol, 6.2 %).

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.16 (s, 1H), 8.53 (s, 1H), 8.47 – 8.35 (m, 3H), 8.01 (s, 1H), 7.86 (d, *J* = 7.7 Hz, 2H), 7.44 (t, *J* = 7.6 Hz, 2H), 7.32 (t, *J* = 7.4 Hz, 1H), 7.01 (dd, *J* = 7.9, 4.5 Hz, 5H), 6.85 (t, *J* = 7.5 Hz, 2H), 6.39 (t, *J* = 7.5 Hz, 2H), 5.05 (t, *J* = 7.6 Hz, 3H), 4.91 (t, *J* = 10.0 Hz, 2H), 4.82 (ddd, *J* = 21.8, 10.6, 2.9 Hz, 2H), 4.73 (s, 3H), 4.63 (d, *J* = 13.4 Hz, 3H), 4.53 (s, 5H), 4.39 (s, 1H), 4.36 (s, 2H), 4.32 (d, *J* = 7.4 Hz, 2H), 4.27 (d, *J* = 3.7 Hz, 3H), 4.25 – 4.22 (m, 4H), 3.71 (q, *J* = 7.0 Hz, 2H), 4.15 (t, *J* = 9.7 Hz, 1H), 4.00 – 3.95 (m, 1H), 3.91 (s, 1H), 3.56 (dd, *J* = 11.7, 6.7 Hz, 10H), 3.52 (s, 25H), 3.50 (s, 15H), 3.48 (s, 2H), 3.39 (d, *J* = 4.6 Hz, 27H), 3.23 (d, *J* = 13.5 Hz, 6H), 2.80 (s, -N(CH_3)$_2$), 3.18 – 3.09 (m, 10H), 2.77 (s, -N(CH_3)$_2$).

^{13}C NMR (126 MHz, DMSO-d_6, extracted from HSQC) δ 121.27, 124.92, 124.87, 123.67, 125.47, 129.37, 128.11, 129.50, 123.36, 123.13, 123.74, 128.36, 49.01, 49.61, 84.19, 67.35, 67.27, 67.08, 67.66, 74.41, 30.76, 64.42, 63.81, 62.94, 74.17, 71.83, 74.15, 74.75, 30.77, 67.20, 68.50, 67.92, 67.36, 68.04, 80.42, 79.84, 70.03, 69.94, 69.35, 60.48, 61.08, 72.27, 70.70, 70.10, 73.04, 73.44, 66.83, 68.40, 67.62, 69.18, 71.48, 73.84, 71.89, 68.00, 69.52, 68.77, 34.40, 33.81, 76.95, 70.11, 30.77, 57.45, 56.61, 56.01, 55.19, 55.78, 56.39, 40.28, 40.16, 40.42, 39.79, 40.46, 40.30, 39.75, 41.80, 42.54, 43.98, 43.15, 44.84, 45.63, 40.74, 40.44, 39.83, 40.57, 40.04, 43.14, 42.53, 40.45, 41.24, 40.33, 41.54, 40.01, 41.95.

HRMS (EI, m/z): calculated for $C_{82}H_{111}N_{15}O_{20}SH^+$ ([M+H]$^+$):1658.7923, found 1658.7974.

25,26,27-Tris-*N*-(*N*,*N*-dimethyl-2-aminoethyl)carbamoylmethoxy-28-{2'-[4-(2,5,8,11,14,17,20,23,26- nonaoxanonacos-28-yn-1-yl)-1H-1,2,3-triazol]ethoxy}calix[4]arene (5b)

Compound **10** (89 mg, 0.20 mmol) and **8** (75 mg, 0.10 mmol) were dissolved in to CH_3CN (5.0 mL) and then CuI (19 mg, 0.10 mmol) was added into the solution. The resulting mixture was heated under microwave irradiation to 80 °C for 90 min. After complete conversion of the starting material according to TLC monitoring, the mixture was concentrated *in vacuo*, and then CH_2Cl_2 (5.0mL) was added. A clear solution was obtained after centrifugation, which was concentrated *in vacuo*, and the resulting compound (70 mg) was obtained after silica chromatography (CH_2Cl_2: CH_3OH = 10: 1). To the compound (70 mg) *N*,*N*-dimethylethylenediamine (5.0 mL) was added under nitrogen. The mixture was stirred for 24 h at 50 °C. The excess of *N*,*N*-dimethylethylenediamine was removed by evaporation under reduced pressure. The mixture was dissolved in CH_3CN (20 mg/mL) and purified by preparative HPLC using the standard protocol. Fractions containing the product (t_R = 40 min) were pooled and the buffer was removed by freeze-drying to obtain the pure compound **5b** as a clear oil (30 mg, 0.023 mmol, 23% for the two steps).

^{1}H NMR (500 MHz, DMSO-d_6) δ 9.59 – 9.55 (m, 3H, -NHMe$_2$), 8.50 – 8.34 (m, 3H, -C(O)NH), 8.17 (s, 1H, CH triazole), 7.02 (t, *J* = 6.0 Hz, 4H, Ar-H), 6.86 (t, *J* = 7.6 Hz, 2H, Ar-H), 5.05 (t, *J* = 7.7 Hz, 2H,-CH$_2$N$_{triazole}$), 4.73 (s, 2H, Ar-OCH$_2$C(O)), 6.39 (t, *J* = 7.6 Hz, 2H, Ar-H), 6.30 (dd, *J* = 20.4, 7.5 Hz, 4H, Ar-H), 4.63 (d, *J* = 13.4 Hz, 2H, Ar$_2$CH$_2$), 4.53 (s, 2H, -CH$_2$O-), 4.38 (d, *J* = 14.0 Hz, 2H, Ar-OCH$_2$C(O)), 4.32 (t, *J* = 7.6 Hz, 2H, Ar-OCH$_2$CH$_2$), 4.28 – 4.23 (m, 3H, Ar-OCH$_2$C(O) and Ar$_2$CH$_2$), 4.13 (d, *J* = 2.2 Hz, 2H, -CH$_2$C≡CH), 3.59 – 3.52 (m, 6H, -NHCH$_2$), 3.43 – 3.33 (m, -OCH$_2$CH$_2$O-), 3.23 (d, *J* = 13.5 Hz, 5H, Ar$_2$CH$_2$), 3.14 (s, 6H, -CH$_2$NMe$_2$), 2.80 (s, -N(CH$_3$)$_2$), 2.77 (s, -N(CH$_3$)$_2$).

^{13}C NMR (126 MHz, DMSO-d_6, extracted from HSQC) δ 123.87, 128.60, 128.45, 128.30, 122.37, 122.11, 127.40, 127.33, 127.38, 48.66, 74.01, 29.86, 62.77, 73.76, 73.20, 29.79, 57.80, 70.45, 68.51, 32.86, 33.44, 33.44, 69.71, 68.20, 68.24, 29.25, 55.51, 55.05, 56.48, 54.98, 42.77, 39.64, 40.78, 42.19, 43.86, 38.99

HRMS (EI, m/z): calculated for $C_{70}H_{101}N_9O_{16}H^+$ ([M+H]$^+$): 1324.7439, found 1324.7462.

Preparation of hybrid 2

5b (30 mg, 0.023 mmol) and **4** (18.6 mg, 0.036 mmol) were dissolved in to CH_3OH (3 mL) and then CuI (4.6 mg, 0.024 mmol) was added to the solution. The resulting mixture was heated under microwave irradiation at 80 °C for 90 min. The mixture was concentrated *in vacuo*, and then CH_2Cl_2 (5.0 mL) was added. A clear solution was obtained after centrifugation, which was concentrated *in vacuo* to afford the crude compound. The crude was dissolved in CH_3CN (20 mg/mL) and purified by preparative HPLC on the

standard protocol. Fractions containing the product (t_R = 35 min) were pooled and the buffer was removed by freeze-drying to obtain the pure hybrid **2** as an off-white foam (12 mg, 0.0065 mmol, 28%).

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.53 (s, 1H), 9.66 – 9.62 (m, 3H), 8.48 – 8.36 (m, 3H), 8.17 (s, 1H), 8.01 (s, 1H), 7.90 – 7.84 (m, 2H), 7.44 (t, *J* = 7.7 Hz, 2H), 7.32 (t, *J* = 7.4 Hz, 1H), 7.02 (dd, *J* = 7.6, 4.1 Hz, 4H), 6.86 (td, *J* = 7.5, 1.8 Hz, 2H), 6.39 (t, *J* = 7.5 Hz, 2H), 6.35 – 6.26 (m, 4H), 5.32 (s, 4H), 5.05 (t, *J* = 7.6 Hz, 2H), 4.91 (dd, *J* = 10.9, 9.5 Hz, 2H), 4.82 (ddd, *J* = 21.6, 10.5, 3.0 Hz, 2H), 4.73 (s, 2H), 4.63 (d, *J* = 13.4 Hz, 2H), 4.53 (s, 5H), 4.36 (s, 1H), 4.33 (d, *J* = 7.7 Hz, 2H), 4.27 (s, 2H), 4.25 (s, 1H), 4.24 (d, *J* = 2.9 Hz, 1H), 4.17 – 4.10 (m, 1H), 3.98 (d, *J* = 3.1 Hz, 1H), 3.91 (d, *J* = 3.2 Hz, 1H), 3.71 (dt, *J* = 9.0, 6.3 Hz, 2H), 3.61 – 3.50 (m, 30H), 3.50 (s, 22H), 3.23 (d, *J* = 13.5 Hz, 5H), 3.15 (d, *J* = 5.2 Hz, 6H), 2.90 – 2.81 (m, 4H), 2.81 – 2.75 (m, 21H).

^{13}C NMR (126 MHz, DMSO-d_6, extracted from HSQC)δ 121.92, 124.86, 124.21, 125.42, 125.41, 129.37, 128.81, 128.08, 129.52, 129.60, 129.52, 123.38, 123.33, 123.10, 123.12, 128.37, 128.39, 128.40, 49.60, 84.20, 67.29, 67.71, 73.52, 31.36, 63.76, 63.84, 70.13, 74.15, 63.82, 70.90, 71.80, 74.17, 30.78, 67.83, 67.19, 69.15, 67.99, 68.62, 79.26, 69.37, 60.47, 69.37, 69.54, 34.42, 34.41, 69.23, 30.78, 58.59, 55.94, 56.47, 43.16, 43.16, 43.16, 41.03

HRMS (EI,m/z): calculated for $C_{90}H_{127}N_{15}O_{24}SH^+$, $[M+H]^+$: 1834.8972, found 1834.8973.

25,26,27-Tris-*N*-(*N*,*N*-dimethyl-2-aminoethyl)carbamoylmethoxy-28-{2'-[4-(2,5,8,11,14,17,20,23,26,29,32,35,38-tridecaoxahentetracont-40-yn-1-yl)-1H-1,2,3-triazol]ethoxy}calix[4]arene (5c)

The compound **10** (128 mg, 0.20 mmol) and **8** (78 mg, 0.10 mmol) were dissolved into CH_3CN (5.0 mL) and then CuI (19 mg, 0.10 mmol) was added to the solution. The resulting mixture was heated under microwave irradiation at 80 °C for 90 min. After complete conversion of the starting material according to TLC monitoring, the mixture was concentrated *in vacuo*, and then CH_2Cl_2 (5.0 mL) was added. A clear solution was obtained after centrifugation, which was concentrated in vacuo, and the resulting compound (120 mg) was achieved through silica chromatography (CH_2Cl_2: CH_3OH = 10: 1). To the compound (120 mg) *N,N*-dimethylethylenediamine (5.0 mL) was added under nitrogen. The mixture was stirred for 24 h at 50 °C. The excess of *N,N*-dimethylethylenediamine was removed by evaporation under reduced pressure. The mixture was dissolved in CH_3CN (20 mg/mL) and purified by preparative HPLC using the standard protocol. Fractions containing the product (t_R = 40 min, broad peak) were pooled and the water was removed by freeze-drying to obtain the compound **5c** as a clear oil (40 mg, 0.027 mmol, 27% for the two steps).

^{1}H NMR (500 MHz, DMSO-d_6) δ 9.75 – 9.59 (m, 3H, -NHMe$_2$), 8.49 – 8.38 (m, 3H, -C(O)NH), 8.18 (s, 1H, CH triazole), 7.01 (dd, *J* = 7.6, 4.0 Hz, 4H, Ar-H), 6.85 (t, *J* = 7.5 Hz, 2H, Ar-H), 6.39 (t, *J* = 7.5 Hz, 2H, Ar-H), 6.31 (dd, *J* = 20.1, 7.5 Hz, 4H, Ar-H), 5.05 (t, *J* = 7.6 Hz, 2H,-CH$_2$N$_{triazole}$), 4.74 (s, 2H, Ar-OCH$_2$C(O)), 4.64 (d, *J* = 13.4 Hz, 2H, Ar$_2$CH$_2$), 4.53 (s, 2H, -CH$_2$O-), 4.38 (d, *J* = 14.0 Hz, 2H, Ar-OCH$_2$C(O)), 4.33 (t, *J* = 7.6 Hz, 2H, Ar-OCH$_2$CH$_2$), 4.30 – 4.25 (m, 3H, , Ar-OCH$_2$C(O) and Ar$_2$CH$_2$), 4.14 (d, *J* = 2.3 Hz, 2H, -CH$_2$C≡CH), 3.56- 3.52 (m, 6H, -C(O)NHCH$_2$), 3.46 – 3.27 (m, -OCH$_2$CH$_2$O-), 3.23 (d, *J* = 13.6 Hz, 5H, Ar$_2$CH$_2$), 3.17 – 3.10 (m, 6H, -CH$_2$N(CH$_3$)$_2$), 2.80 (s, -N(CH$_3$)$_2$), 2.77 (s, -N(CH$_3$)$_2$).

^{13}C NMR (126 MHz, DMSO-d_6, extracted from HSQC) δ 124.87, 129.23, 129.48, 123.97, 123.33, 123.11, 128.38, 128.46, 49.02, 74.50, 30.26, 63.79, 70.14, 74.17, 71.64, 74.19, 30.28, 56.98, 66.46, 72.30, 70.73, 70.13, 30.21, 43.14, 44.01.

HRMS (EI, m/z): calculated for $C_{78}H_{117}N_9O_{20}H^+$([M+H]$^+$): 1500.8488, found 1500.8480.

Preparation of hybrid 3

Compound **5c** (40 mg, 0.027 mmol) and **4** (17 mg, 0.033 mmol) were dissolved in CH_3OH (3.0mL), then CuI (5.1 mg, 0.027 mmol) was added into the solution. The resulting mixture was heated under microwave irradiation at 80 °C for 90 min. The mixture was concentrated *in vacuo*, and then CH_2Cl_2 (5 mL) was added. A clear solution was obtained after centrifugation, which was concentrated *in vacuo* to afford the crude compound. The crude was dissolved in CH_3CN (20 mg/mL) and purified by preparative HPLC using the standard protocol. Fractions containing the product (t_R = 35 min, broad peak) were pooled and the buffer was removed by freeze-drying to obtain the pure hybrid **3** as an off-white foam (12 mg, 0.0060 mmol, 22%).

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.53 (s, 1H), 8.17 (s, 1H), 9.68 – 9.64 (m, 3H), 8.50 – 8.34 (m, 3H), 8.01 (s, 1H), 7.87 (d, *J* = 7.7 Hz, 1H), 7.44 (t, *J* = 7.7 Hz, 2H), 7.31 (d, *J* = 7.5 Hz, 1H), 7.02 (dd, *J* = 7.6, 4.2 Hz, 3H), 6.89 – 6.82 (m, 2H), 6.39 (t, *J* = 7.5 Hz, 2H), 6.30 (dd, *J* = 20.3, 7.4 Hz, 3H), 5.32 (s, 4H), 5.05 (t, *J* = 7.5 Hz, 2H), 4.91 (dd, *J* = 11.2, 9.5 Hz, 2H), 4.82 (ddd, *J* = 21.3, 10.6, 3.0 Hz, 2H), 4.73 (s, 2H), 4.64 (d, *J* = 13.4 Hz, 2H), 4.53 (s, 4H), 4.39 (s, 1H), 4.36 (s, 1H), 4.33 (t, *J* = 7.6 Hz, 2H), 4.27 (s, 1H), 4.26 (d, *J* = 11.3 Hz, 2H), 4.23 (d, *J* = 10.8 Hz, 2H), 3.97 (d, *J* = 3.1 Hz, 1H), 3.91 (d, *J* = 3.2 Hz, 1H), 3.71 (dt, *J* = 9.4, 6.3 Hz, 2H), 3.61 – 3.44 (m, 56H), 3.42 (t, *J* = 6.2 Hz, 2H), 3.23 (d, *J* = 13.6 Hz, 4H), 3.17 – 3.09 (m, 7H), 2.84 (d, *J* = 22.3 Hz, 2H), 2.82 – 2.75 (m, 16H).

^{13}C NMR (126 MHz, DMSO-d_6, extracted from HSQC)δ 121.87, 125.80, 123.66, 126.06, 126.06, 129.36, 129.36, 127.90, 129.51, 128.87, 129.48, 123.34, 123.35, 123.11, 123.11, 128.37, 128.37, 128.42, 129.40, 50.47, 84.17, 67.93, 67.76, 74.41, 30.77, 74.16, 70.13, 74.17, 70.11, 74.16, 30.79, 67.82, 74.15, 70.71, 69.64, 67.20, 67.82, 67.82, 70.73, 70.13,

67.37, 68.02, 78.90, 79.25, 79.86, 70.14, 70.15, 68.75, 69.19, 68.38, 34.40, 70.13, 33.81, 30.79, 30.77, 70.13, 56.56, 70.13, 54.38, 56.53, 56.61, 54.22, 39.97, 43.14, 43.21, 41.74

HRMS (EI, m/z) calculated for $C_{98}H_{143}N_{15}O_{28}SH^+$, $[M+H]^+$: 2011.0020, found: 2011.0092.

Bis-{3-deoxy-3[4-(phenyl)-1H-1,2,3-triazol-1-yl]-β-D-galactopyranosyl} sulfane (19)

Compound **18** (10 mg, 0.013 mmol) and phenylacetylene (1.43µl, 0.013 mmol) were dissolved in CH_3CN (2.0 mL) and then CuI (2.50 mg, 0.013 mmol) was added to the solution. The resulting mixture was heated under microwave irradiation at 80 °C for 90 min. After complete conversion of the starting material according to TLC monitoring, the mixture was concentrated *in vacuo*, and then CH_2Cl_2 (5.0 mL) was added. A clear solution was obtained after centrifugation, which was concentrated *in vacuo.* The residue (12 mg) was dissolved in CH_3OH (5 mL) and NaOMe (40 mg, 2.5 mmol) was added. The mixture was stirred for 6 h at room temperature and was neutralized with DOWEX-H^+ resin, filtered, and evaporated. The crude was purified by preparative HPLC using the standard protocol. Fractions containing the product (t_R = 30 min) were pooled and the buffer was removed by freeze-drying to obtain compound **19** as a white solid (2.0 mg, 0.003 mmol, 23 %).

^{1}H NMR (500 MHz, D_2O) δ 8.55 (s, 2H, CH triazole), 7.87 (d, J = 7.8 Hz, 4H, Ar-H), 7.55 (t, J = 7.5 Hz, 4H, Ar-H), 7.48 (t, J = 7.5 Hz, 2H, Ar-H), 5.17 (d, J = 9.7 Hz, 2H, H1), 5.04 (d, J = 10.9 Hz, 2H, H3), 4.49 (t, J = 10.2 Hz, 2H, H2), 4.29 (m, 2H, H4), 4.08 (t, J = 6.2 Hz, 2H, H5), 3.87 (dd, J = 12.0, 7.7 Hz, 2H, H6a), 3.79 (dd, J = 12.0, 4.4 Hz, 2H, H6b).

^{13}C NMR (126 MHz, D_2O, extracted from HSQC) δ 121.43, 126.38, 129.92, 129.53, 84.35, 66.97, 67.38, 68.03, 79.66, 60.41, 60.89

HRMS (EI, m/z) calculated for $C_{28}H_{32}N_6O_8SNa^+$ ($[M+Na]^+$): 635.1895, found 635.1898.

Preparation of compound 20

The compound **18** (30 mg, 0.039 mmol) and compound **11** (42 mg, 0.067 mmol) were dissolved to CH_3CN (2.0 mL) and then CuI (7.50 mg, 0.039 mmol) was added to the solution. The resulting mixture was heated under microwave irradiation at 80 °C for 90 min. After complete conversion of the starting material according to TLC monitoring, the mixture was concentrated *in vacuo*, and then CH_2Cl_2 (5.0 mL) was added. A clear solution was obtained after centrifugation, which was concentrated *in vacuo*. The corresponding compound was obtained by silica chromatography (CH_2Cl_2: CH_3OH = 10: 1→ 5: 1), which was dissolved in CH_3OH (5 mL) and NaOMe (40 mg, 2.5 mmol) was added. The mixture was stirred for 6 h at room temperature. The solution was neutralized with DOWEX-H^+ resin, filtered, and evaporated. The crude was purified by preparative HPLC using the standard protocol. Fractions containing the product (t_R = 30 min) were pooled and the

buffer was removed by freeze-drying to obtain compound **20** as a clear oil (3.0 mg, 0.002 mmol, 5.0 %).

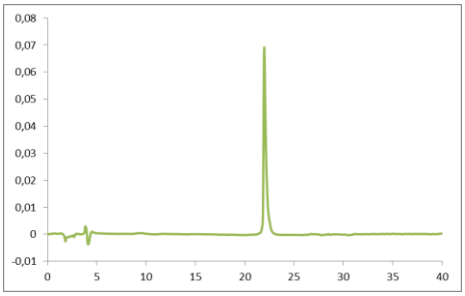
^{1}H NMR (500 MHz, D2O) δ 8.55 (s, 1H), 8.26 (s, 1H), 7.91 – 7.84 (m, 2H), 7.56 (t, J = 7.6 Hz, 2H), 7.49 (d, J = 7.4 Hz, 1H), 5.14 (dd, J = 9.7, 6.9 Hz, 2H), 5.02 (ddd, J = 13.7, 10.6, 3.0 Hz, 2H), 4.74 (s, 2H), 4.55 – 4.39 (m, 2H), 4.30 – 4.21 (m, 4H), 4.06 (td, J = 8.5, 4.3 Hz, 2H), 3.86 (ddd, J = 12.6, 7.8, 5.2 Hz, 2H), 3.81 – 3.62 (m, 50H).

^{13}C NMR (126 MHz, D_2O, extracted from HSQC) δ 121.44, 124.43, 125.77, 125.77, 129.35, 129.35, 128.97, 84.27, 84.29, 66.92, 66.90, 63.05, 69.52, 66.80, 68.02, 57.65, 67.96, 79.66, 60.91, 60.93, 68.90, 69.53.

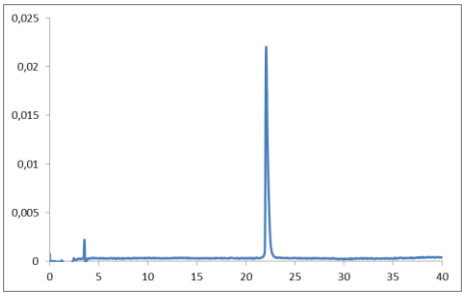
HRMS (EI, m/z) calculated for $C_{50}H_{80}N_6O_{21}SH^+$ ($[M+H]^+$): 1133.5170, found 1133.5183.

4.4.3 Analytical HPLC for new compounds

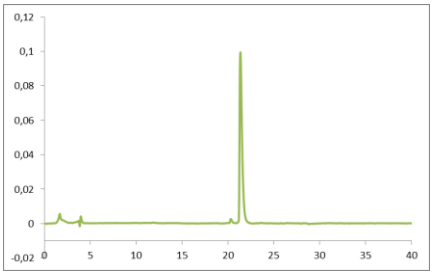
Compound **5a**



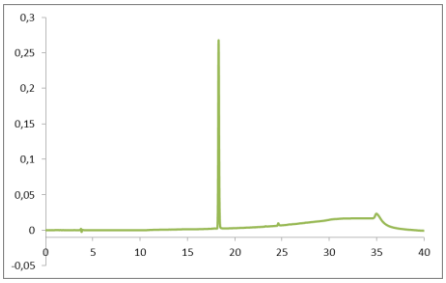
Compound **5b**



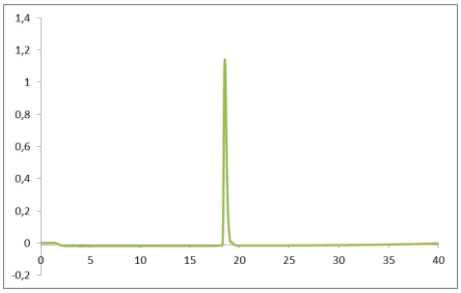
Compound **5c**



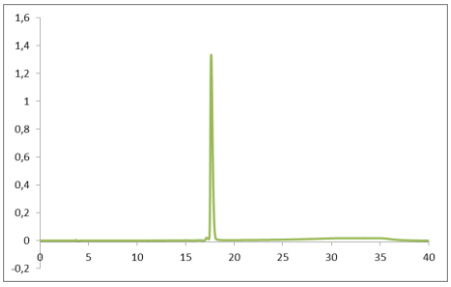
Compound **12**



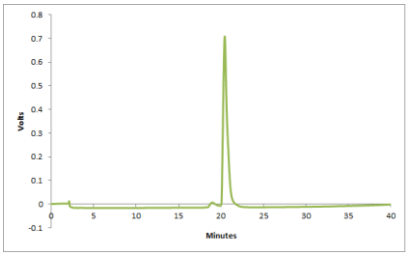
Compound **19**



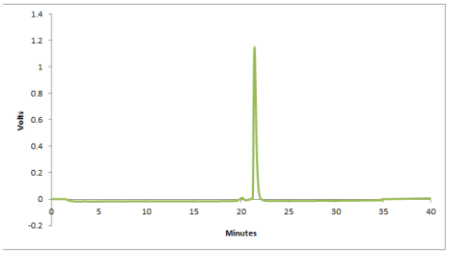
Compound **20**



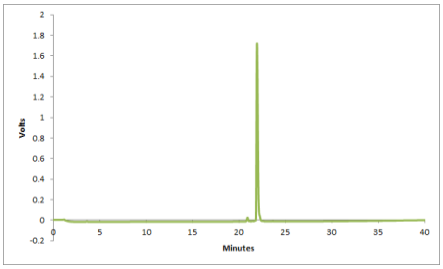
hybrid **1**



hybrid **2**



HPLC of hybrid **3**



4.5 References:

(1) Cooper, N. W. D. Galectinomics: Finding Themes in Complexity. *Biochim. Biophys. Acta - Gen. Subj.* **2002**, *1572* (2–3), 209–231.

(2) Liu, F.; Patterson, R. J.; Wang, J. L. Intracellular Functions of Galectins. *Biochim. Biophys. Acta - Gen. Subj.* **2002**, *1572*, 263–273.

(3) Perillo, N. L.; Baum, L. G. Galectins : Versatile Modulators of Cell Adhesion, Cell Proliferation, and Cell Death. *J. Mol. Med.* **1998**, *76*, 402–412.

(4) Liu, F.-T.; Rabinovich, G. A. Galectins as Modulators of Tumour Progression. *Nat. Rev.* **2005**, *5* (1), 29–41.

(5) Thijssen, V. L. J. L.; Postel, R.; Brandwijk, R. J. M. G. E.; Dings, R. P. M.; Nesmelova, I.; Satijn, S.; Verhofstad, N.; Nakabeppu, Y.; Baum, L. G.; Bakkers, J.; Mayo, K. H.; Poirier, F.; Griffioen, A. W. Galectin-1 Is Essential in Tumor Angiogenesis and Is a Target for Antiangiogenesis Therapy. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (43), 15975–15980.

(6) Traber, P. G.; Zomer, E. Therapy of Experimental NASH and Fibrosis with Galectin Inhibitors. *PLoS One* **2013**, *8* (12), e83481.

(7) Liu, F.; Rabinovich, G. A. Galectins : Regulators of Acute and Chronic Inflammation. *Ann. N. Y. Acad. Sci.* **2010**, *1183*, 158–182.

(8) Blanchard, H.; Yu, X.; Collins, P. M.; Bum-Erdene, K. Galectin-3 Inhibitors: A Patent Review (2008–present). *Expert Opin. Ther. Pat.* **2014**, *24* (10), 1053–1065.

(9) Mackinnon, A.; Chen, W.-S.; Leffler, H.; Panjwani, N.; Schambye, H.; Sethi, T.; Nilsson, U. J. Design, Synthesis, and Applications of Galectin Modulators in Human Health. In *Top med chem*; 2014; Vol. 12, pp 95–121.

(10) Blanchard, H.; Bum-Erdene, K.; Bohari, M. H.; Yu, X. Galectin-1 Inhibitors and Their Potential Therapeutic Applications: A Patent Review. *Expert Opin. Ther. Pat.* **2016**, *26* (5), 537–554.

(11) Cumpstey, I.; Sundin, A.; Leffler, H.; Nilsson, U. J. C2-Symmetrical Thiodigalactoside Bis-Benzamido Derivatives as High-Affinity Inhibitors of Galectin-3: Efficient Lectin Inhibition through Double Arginine-Arene Interactions. *Angew. Chemie - Int. Ed.* **2005**, *44* (32), 5110–5112.

(12) Sörme, P.; Qian, Y.; Nyholm, P. G.; Leffler, H.; Nilsson, U. J. Low Micromolar Inhibitors of Galectin-3 Based on 3'-Derivatization of N-Acetyllactosamine. *ChemBioChem* **2002**, *3* (2–3), 183–189.

(13) Giguère, D.; Patnam, R.; Bellefleur, M.-A.; St-Pierre, C.; Sato, S.; Roy, R. Carbohydrate Triazoles and Isoxazoles as Inhibitors of Galectins-1 and -3. *Chem. Commun.* **2006**, No. 22, 2379–2381.

(14) Leffler, H.; Barondes, S. H. Specificity of Binding of Three Soluble Rat Lung Lectins to Substituted and Unsubstituted Mammalian β-Galactosides. *J. Biol. Chem.* **1986**, *261* (22), 10119–10126.

(15) Matsubara, T. Potential of Peptides as Inhibitors and Mimotopes: Selection of Carbohydrate-Mimetic Peptides from Phage Display Libraries. *J. Nucleic Acids* **2012**, *12* (ID740982), 15.

(16) Simon-haldi, M.; Mantei, N.; Franke, J.; Voshol, H.; Schachner, M. Identification of a Peptide Mimic of the L2 / HNK-1 Carbohydrate Epitope. *J. Neurochem.* **2002**, *83* (6), 1380–1388.

(17) Zou, J.; Glinsky, V. V; Landon, L. A.; Matthews, L.; Deutscher, S. L. Peptides Specific to the Galectin-3 Carbohydrate Recognition Domain Inhibit Metastasis-Associated Cancer Cell Adhesion. *Carcinogenesis* **2005**, *26* (2), 309–318.

(18) Landon, L. A.; Deutscher, S. L. Combinatorial Discovery of Tumor Targeting Peptides Using Phage Display. *J. Cell. Biochem.* **2003**, *90*, 509–517.

(19) Burke, S. D.; Zhao, Q.; Schuster, M. C.; Kiessling, L. L.; January, R. V. Synergistic Formation of Soluble Lectin Clusters by a Templated Multivalent Saccharide Ligand. *J. Am. Chem. Soc.* **2000**, *122* (27), 4518–4519.

(20) van der Schaft, D. W. J.; Dings, R. P. M.; de Lussanet, Q. G.; van Eijk, L. I.; Nap, A. W.; Beets-Tan, R. G. H.; Bouma-Ter Steege, J. C. A.; Wagstaff, J.; Mayo, K. H.; Griffioen, A. W. The Designer Anti-Angiogenic Peptide Anginex Targets Tumor Endothelial Cells and Inhibits Tumor Growth in Animal Models. *FASEB J.* **2002**, *16* (14), 1991–1993.

(21) Mayo, K. H.; Haseman, J.; Ilyina, E.; Gray, B. Designed β-Sheet-Forming Peptide 33mers with Potent Human Bactericidal/permeability Increasing Protein-like Bactericidal and Endotoxin Neutralizing Activities. *Biochim. Biophys. Acta - Gen. Subj.* **1998**, *1425* (1), 81–92.

(22) Dings, R. P. M.; Mayo, K. H. A Journey in Structure-Based Drug Discovery: From Designed Peptides to Protein Surface Topomimetics as Antibiotic and Antiangiogenic Agents. *Acc. Chem. Res.* **2007**, *40* (10), 1057–1065.

(23) Dings, R. P. M.; Chen, X.; Hellebrekers, D. M. E. I.; van Eijk, L. I.; Zhang, Y.; Hoye, T. R.; Griffioen, A. W.; Mayo, K. H. Design of Nonpeptidic Topomimetics of Antiangiogenic Proteins with Antitumor Activities. *J. Natl. Cancer Inst.* **2006**, *98* (13), 932–936.

(24) Dings, R. P. M.; Miller, M. C.; Nesmelova, I.; Astorgues-Xerri, L.; Kumar, N.; Serova, M.; Chen, X.; Raymond, E.; Hoye, T. R.; Mayo, K. H. Antitumor Agent Calixarene 0118 Targets Human Galectin-1 as an Allosteric Inhibitor of Carbohydrate Binding. *J. Med. Chem.* **2012**, *55* (11), 5121–5129.

(25) Hattum, H. Van; Branderhorst, H. M.; Moret, E. E.; Nilsson, U. J.; Leffler, H.; Pieters, R. J. Tuning the Preference of Thiodigalactoside- and Lactosamine-Based Ligands to Galectin-3 over Galectin-1. *J. Med. Chem.* **2013**, *56* (3), 1350–1354.

(26) Läppchen, T.; Dings, R. P. M.; Rossin, R.; Simon, J. F.; Visser, T. J.; Bakker, M.; Walhe, P.; van Mourik, T.; Donato, K.; van Beijnum, J. R.; Griffioen, A. W.; Lub, J.; Robillard, M. S.; Mayo, K. H.; Grüll, H. Novel Analogs of Antitumor Agent Calixarene 0118: Synthesis, Cytotoxicity, Click Labeling with 2-[18F]fluoroethylazide, and in Vivo Evaluation. *Eur. J. Med. Chem.* **2015**, *89*, 279–295.

(27) Sung, S. R.; Han, S. C.; Jin, S.; Lee, J. W. Convergent Synthesis and Characterization of Dumbbell Type Dendritic Materials by Click Chemistry. *Bull. Korean Chem. Soc.* **2011**, *32* (11), 3933–3940.

(28) Bryant, L. H.; Yordanov, A. T.; Linnoila, J. J.; Brechbiel, M. W.; Frank, J. a. First Noncovalently Bound Calix [4] Arene ± Gd III ± Albumin Complex. *Angew. Chem. Int. Ed.* **2000**, *39*, 1641–1643.

(29) Mandal, S.; Nilsson, U. J. Tri-Isopropylsilyl Thioglycosides as Masked Glycosyl Thiol Nucleophiles for the Synthesis of S-Linked Glycosides and Glyco-Conjugates. *Org. Biomol. Chem.* **2014**, *12* (27), 4816–4819.

(30) Jaime, C.; De Mendoza, J.; Prados, P.; Nieto, P. M.; Sanchez, C. Carbon-13 NMR Chemical Shifts. A Single Rule to Determine the Conformation of calix[4]arenes. *Sect. Title Phys. Org. Chem.* **1991**, *56* (10), 3372–3376.

(31) Magrans, J. O.; Mendoza, J. De. Are 1,3-Di- O -benzoylcalix[4]arenes an Exception to the 13 C-NMR Rule for Conformational Determination? *J. Org. Chem.* **1997**, *3263* (7), 4518–4520.

(32) Williamson, M. P. Using Chemical Shift Perturbation to Characterise Ligand Binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16.

(33) Ippel, H.; Miller, M. C.; Berbís, M. A.; André, S.; Hackeng. F, T. M.; Caňada, J.; Weber, C.; Gabius, H.;

Jiménez-Barbero, J.; Mayo, K. H. $^{1}H,^{13}C$,and ^{15}N Backbone and Side-Chain Chemical Shift Assignments for the 36 Proline-Containing, Full Length 29 kDa Human Chimera-Type Galectin-3. *Biomol. NMR Assign.* **2015**, *9* (1), 59–63.

Chapter 5

Inhibition of O-GlcNAc transferase by peptidic hybrids

Abstract

O-GlcNAc transferase (OGT) attaches a GlcNAc on specific substrate proteins using UDP-GlcNAc as the sugar donor. This modification can alter protein function directly by regulating cellular signaling and transcription pathways in response to altered nutrient availability and stress. Specific inhibitors of OGT would be valuable tools for biological studies and lead structures for therapeutics. The existing OGT inhibitors are mainly derived from sugar donor substrates, but poor cell permeability and off-target effects limit their use. Here, we describe our progress on OGT inhibition based on substrate peptides identified by array screening. Subsequently, bisubstrate inhibitors were prepared by conjugating these peptides to uridine in various ways. In parallel, an *in silico* fragment screening was conducted to obtain small molecules targeting the UDP binding pocket. After evaluation of the initial hits, one of these small molecules was introduced into a novel OGT hybrid inhibitor, as the replacement of uridine. The novel compounds inhibit OGT activity with IC_{50} values in the micromolar range.

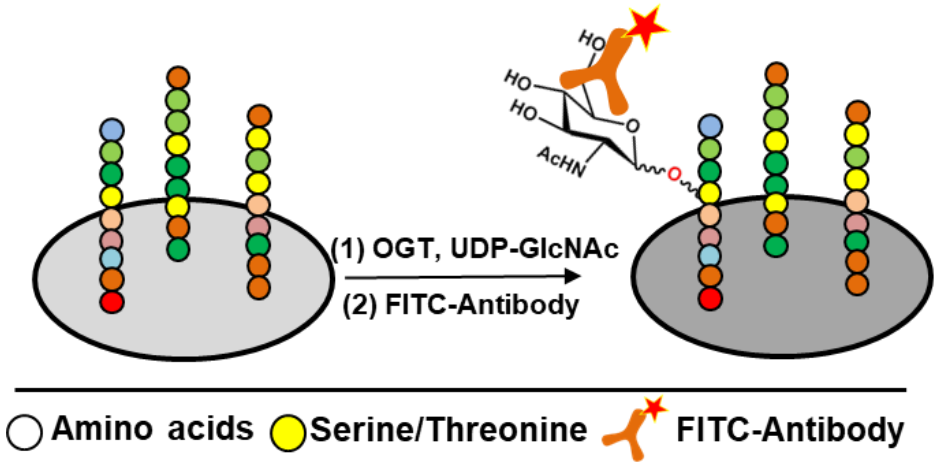
5.1 Introduction

O-GlcNAc transferase (OGT) is an essential mammalian enzyme involved in the dynamic O-GlcNAcylation of cytosolic and nuclear proteins. Through catalyzing the attachment of N-acetylglucosamine to specific serines and threonines of protein, OGT is associated with numerous biological processes such as transcription, cell cycle progression, stress response and nutrient sensing.[1–5] Prior studies have noted the potential of OGT as a therapeutic target as it is linked to diseases such as Alzheimer's disease and cancer.[6–8] OGT is one of only two enzymes modulating O-GlcNAcylation and its expression is correlated to the metabolic status of cell directly.[7,9–11] In metabolic diseases like cancer and diabetes, increase of various metabolic inputs like glucose into the cell alters the production of UDP-GlcNAc through the hexosamine biosynthetic pathway. This promotes O-GlcNAcylation since OGT is highly sensitive to intracellular UDP-GlcNAc levels. Elevated O-GlcNAcylation of key signaling molecules and transcription factors can be involved in the development and pathology of diseases.[8,10,12,13] For this reason, OGT is emerging as a therapeutic target of current interest. A selective and cell-permeable OGT inhibitor is strongly needed to further decipher its function in biological process and explore its potential as a therapeutic target.

In recent years, biochemical and crystallographic studies have clearly revealed the mechanistic aspects of OGT's catalytic process. The sugar donor substrate UDP-GlcNAc binds to the active site of the enzyme first, followed by the sugar acceptor substrate.[14,15] The ternary substrate approaches the transition state for glycosyl transfer of *N*-acetylglucosamine from the donor onto the specific serine or threonine of the acceptor. To discover OGT inhibitors, a few OGT sugar donor analogues have been identified that show a moderate inhibitory effect on OGT activity.[16–18] However, these compounds cannot

permeate into cells, which restricts their use in research. In addition, most of these compounds have a severe off-target effect because of their similarity to commonly used UDP-GlcNAc.[17] Other types of inhibitors have been described that are non-substrate inhibitors but some off-target or toxicity effects were still noted.[19]

Here, we describe our approach to obtaining OGT inhibitors targeting acceptor and donor sites using a dynamic peptide microarray and a fragment-based method, respectively. In collaboration with Jie Shi within our group, two selective OGT peptide substrates were identified by using peptide microarrays (Scheme 1). First their potential for inhibition was studied by substituting the serine in the O-GlcNAc modification site with an alanine. Inspired by his work, a further optimization of these peptides resulted in new inhibitors with improved inhibitory effect on OGT activity. Moreover, a study on bisubstrate OGT inhibitors was performed by conjugating part of the sugar donor (uridine) to the substrate peptides with neutral linkers, which showed a variation in the magnitude of the synergy effects. Besides, the process of searching for fragments targeting the UDP binding pocket of OGT, provided us with four compounds inhibiting OGT activity *in vitro*, one of which as a replacement of uridine was incorporated into a novel OGT bisubstrate inhibitor.



Scheme 1. The schematic depiction of the peptide microarray approach to afford peptides as substrate of OGT. Briefly, the microarray was blocked with BSA, and followed by the addition of OGT, UDP-GlcNAc, and a FITC conjugated antibody. The enzymatic reaction was run by pumping the mixture up and down through the porous Al_2O_3 chip material.

5.2 Results and Discussion

5.2.1 The optimization of peptide inhibitors derived from OGT acceptor substrate

Using a peptide microarray approach, we recently identified the peptide RBL-2_410-422 as an OGT substrate (Scheme 1). This peptide is derived from the RBL-2 protein, which is a key regulator of entry into cell division and may function as a tumor suppressor. The serine that is modified by O-GlcNAc was identified in this peptide.[20,21] Interestingly, replacing this serine with an alanine in peptide RBL-2_410-422_420A resulted in a competitive OGT inhibitor (Figure 1). This finding suggests a significant interaction between the peptide substrate and OGT that could form the basis for OGT inhibitor development. Another specific substrate peptide was identified among tyrosine kinase substrates and was derived from the tight junctions ZO-3 protein: ZO-3_357_371.[22] Again replacing the serine at the modification site with an alanine yielded an OGT inhibitor. To identify the minimum structural requirements for OGT inhibition, two series of shorter peptides based on the original sequences were synthesized and they were tested as OGT inhibitors (Figure 2). In the series of truncated RBL-2_410-422_S420A peptides, surprisingly, an actual potency improvement was observed for a shorter peptide, i.e. octapeptide **Pep6** (Figure 2A). Additional experiments showed that it exhibited an IC_{50} value of 385 μM.

Similar results were obtained for the ZO-3_357-371_S369A derivatives with the optimal peptide being heptapeptide **Pep13** (Figure 2B), that showed an IC_{50} value of 193 μM. Besides providing more support for OGT being sequence specific, our data also motivated us to improve potency and selectivity. To this end, we proceeded to develop bisubstrate inhibitors by conjugating parts of mimics of the sugar donor analogues with these identified peptide inhibitors.

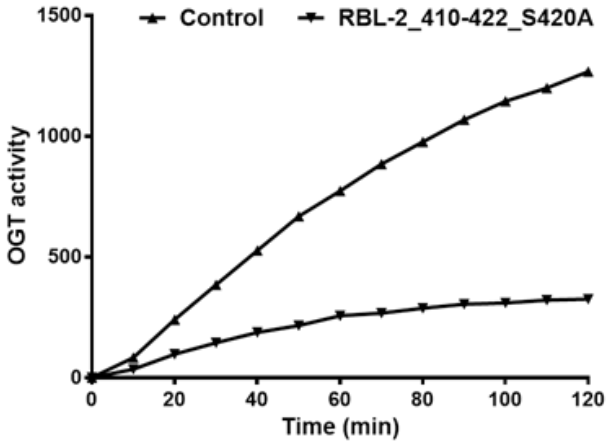
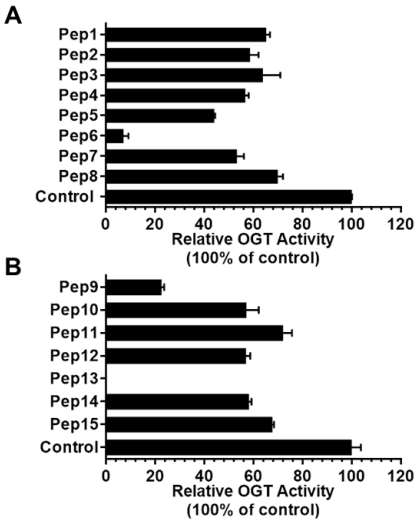


Figure 1. The RBL-2_410-422_S420A inhibited OGT activity competitively. The effect of RBL-2_410-422_S420A (0.5 mM) on OGT activity was determined using a peptide microarray assay on which RBL-2_410-422 was immobilized. A parallel assay was performed using water as control. Real time OGT activity was reflected by the fluorescent antibody bound to O-GlcNAcylated RBL-2_410-422.



Peptide	Sequence
Pep1	KENSPCVTPVATA
Pep2	ENSPCVTPVATA
Pep3	NSPCVTPVATA
Pep4	SPCVTPVATA
Pep5	PCVTPVATA
Pep6	CVTPVATA
Pep7	VTPVATA
Pep8	TPVATA

Peptide	Sequence
Pep9	CGRESSYDIYRVPSAQS
Pep10	DIYRVPSAQS
Pep11	IYRVPSAQS
Pep12	YRVPSAQS
Pep13	RVPSAQS
Pep14	VPSAQS
Pep15	PSAQS

Figure 2. (A) A series of peptides derived from RBL-2_410-422_S420A (**Pep1**) were produced with sequences shown in the right table and their inhibitory effect on OGT activity were determined at concentration of 1 mM using UDP-Glo assay. **Pep6** is obtained as the best OGT inhibitor among these peptides. (B) A series of peptides derived from ZO-3_357-371_S369A (**Pep9**) were produced with sequences shown in the right table and their inhibitory effect on OGT activity were determined at concentration of 1 mM using UDP-Glo assay. **Pep13** is obtained as the best OGT inhibitor among these peptides. All peptides were synthesized by standard Fmoc SPPS protocols described in Experimental Section.

5.2.2 Design and synthesis bisubstrate inhibitors by conjugating peptides to uridine with neutral linkers

As revealed by privious research, the K_d of sugar donor (UDP-GlcNAc) for OGT is even higher than the one of UDP, which points that the GlcNAc moiety even hampers the binding.[23] Accordingly, we have focused on the UDP alone and its binding site to boost the peptides' affinity. UDP itself as a product of O-GlcNAcylation inhibits OGT *in vitro* with an IC_{50} of 1.8 μM, making it the most potent OGT inhibitor to date.[23,18] Despite its strong affinity for OGT the charged phosphates preclude cell permeability. Aiming to combine the peptide selectivity with UDP's potency but removing its diphosphate, we prepared uridine-peptide conjugates (compounds **1-5** shown in Figure 3) with neutral linkers of different length.

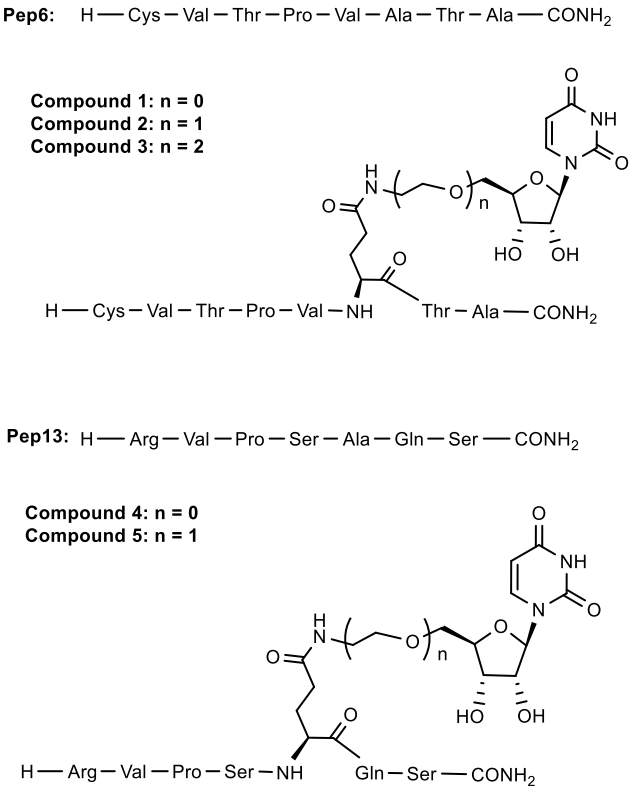


Figure 3. Structure of **Pep6**, **Pep13** and bisubstrate derivatives. All the compounds were prepared by standard Fmoc SPPS protocols. The modified "Glu" building blocks were prepared as described in Scheme 2. Incorporation of the modified "Glu" building block were performed by using corresponding building block (2.0 eq), BOP reagent (4.0 equiv), and DiPEA (8 eq). Upon completion of SPPS, peptides were cleaved from the resin to yield the crude product. Purified compounds were obtained by preparative HPLC and compounds identity was confirmed by analytical HPLC and high-resolution mass spectrometry.

The ability of these compounds to inhibit OGT activity was again determined using UDP-Glo Glycosyltransferase assay and was further compared with **Pep6** and **Pep13**. As shown in Table 1, an IC_{50} of 297 μM was determined for **1** whereas no obvious inhibition was shown for **2** and **3**, indicating the importance of spacer length. One unanticipated finding was that similar modifications on **Pep13** would render the compounds inactive, indicating differences in the binding modes of **Pep13** and **Pep6**. Overall the enhancement of the binding due to the introduction of UDP is modest at best. For this reason, we started a search for more potent fragments that could more productively replace uridine in the context of a bisubstrate or hybrid inhibitor.

Table 1. OGT inhibition values obtained for **Pep6**, **Pep13** and bisubstrate derivatives.[a]

Compound	IC_{50}[μM]	Compound	IC_{50}[μM]
Pep6	385	**Pep13**	193
1	297	**4**	>1000
2	>1000	**5**	>1000
3	493		

[a]IC_{50} values are reported in μM and are averages obtained from triple independent duplicate analysis of each compound.

5.2.3 Discovery of small molecules targeting UDP binding pocket from a fragment screening and a novel OGT bisubstarte inhibitor

In collaboration with Professor Marko Anderluh (University of Ljubljana), a fragment library containing more than 216,000 fragments to identify compounds binding to the UDP site of OGT, was screened with a docking protocol in FRED software was screened. The library contained fragments from the Asinex, ChemBridge, ChemDiv, Enamine, Life chemicals, Key Organics, Maybridge and Otava fragment libraries. Analysis of interactions of UDP in the OGT binding site revealed that uracil moiety of UDP forms two hydrogen bonds with Ala896 backbone carbonyl and amine groups. To mimic this hydrogen bonding network Ala896 was used as a constraint in virtual screening (docking) experiment and hence only compounds containing proximal hydrogen bond donor and acceptor groups were retrieved as potential hits.

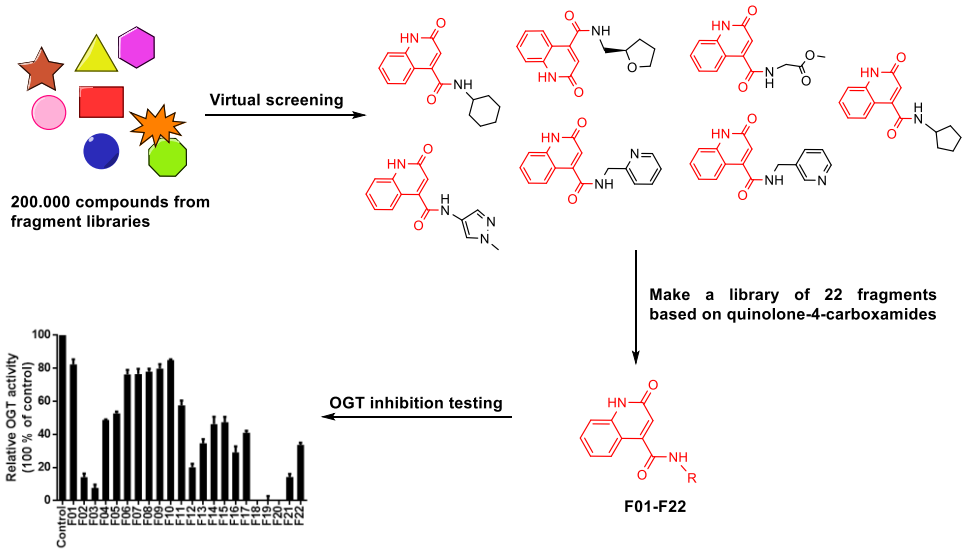


Figure 4. The schematic depiction of fragment screening and testing. Fragments screening yielded seven hits with same heterocyclic scaffold that were synthetically modified and tested for OGT inhibition at 1 mM using a UDP-Glo assay.

Among these virtual hits, seven compounds contained the same scaffolds as they were all quinolone-4-carboxamides (Figure 4). A common feature of these molecules is that the quinolone ring is anchored in the uridine binding site of OGT and different carboxamides point to the diphosphate group binding site. To capitalize on this finding, a series of 22 fragments (**F01-F22**) carrying diverse carboxamides (listed in Table 3) was prepared. This library was constructed by coupling 2-hydroxyquinoline-4-carboxylic acid and various amines with EDC/HOBt as a coupling agent. The inhibitory effect of these compounds on OGT activity was initially measured at concentration of 1 mM using a UDP-Glo assay and several fragments were found to inhibit OGT activity by 90%. The most potent fragments are shown in Figure 5A, and their inhibition profile at 200 μM. Docking pose of OGT inhibitor F20 predicted two hydrogen bonds with Ala896, which mimic the binding mode of the uracil moiety of UDP. Additional hydrogen bonds are predicted between Thr921 side chain and inhibitor carboxylate group and Thr922 side chain and carbonyl group of the carboxamide at position 4 of the quinolone-2-one ring (Figure 6). Moreover, comparing the similar **F19** and **F20** activities suggests that a free carboxylate group is not a prerequisite for binding, thus making this a suitable site for conjugation to the peptide.

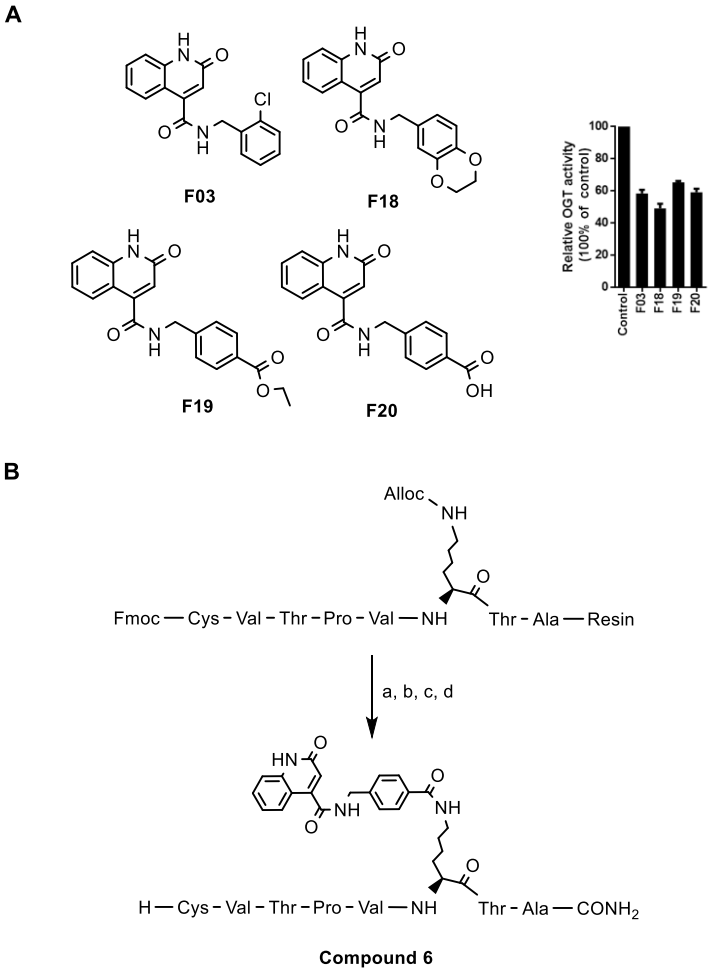


Figure 5. (A) Top 4 fragments structures and OGT inhibition validation. **F03**, **F18**, **F19** and **F20** were tested for OGT activity inhibition at 0.2 mM using UDP-Glo assay. (B) On resin synthesis of compound **6** as a Pep6-F20 conjugate: (a) $Pd(PPh_3)_4$, phenylsilane, DCM, RT, 1h. (b) **F20**, Bop, DiPEA, DMF, 2h. (c) Piperidine/DMF, 1h. (d) TFA/TIPS/EDT/H_2O, RT, 3h.

Considering that the shortest linker was the most effective in **1**-**3**, a lysine was selected for conjugating the fragment, placing the fragment at a similar distance than the uridine group in **1**. This choice also allowed a convenient on-resin synthesis, in which **F20** can be covalently linked to the amine group of the Lys side chain. After cleavage and purification, compound **6** was obtained and tested for its ability to inhibit OGT in our assay. An IC_{50} of 117 μM was determined, which indicated that an improved synergy effect of the two components hybrid inhibitor.

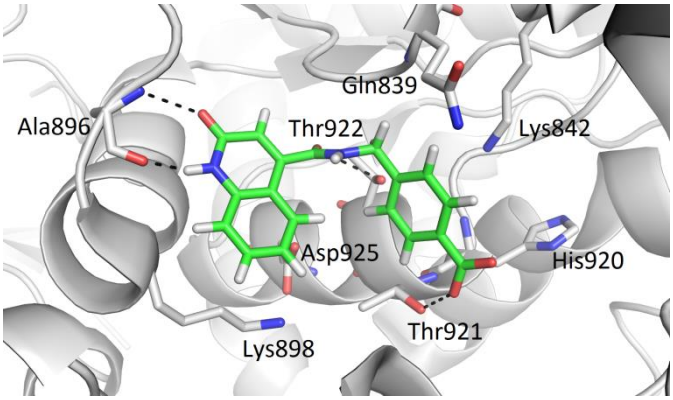


Figure 6. Predicted binding mode of OGT inhibitor **F20** (in green sticks) in OGT binding site (in grey cartoon, PDB entry: 4N39). The ligand and the neighbouring protein side-chains are shown as stick models, coloured according to the chemical atom type (blue, N; red, O; orange, S; green, Cl). Hydrogen bonds are indicated by black dotted lines.

5.3 Conclusion

In conclusion, we successfully designed and synthesized several bi-substrate or hybrid OGT inhibitors. Starting from peptide substrates and replacing the modification site serine with an alanine showed inhibition. Trimming the peptides to 7-8 amino acids actually yielded the more potent inhibitors **Pep6** and **Pep13**. Such an enhancement is unusual as trimming usually leads to reduced potency. Linking uridine to **Pep6** in various ways lead to a modest synergy effect but not for **Pep13**. In addition, fragments were identified to block the sugar donor site of OGT through a combination of virtual screening and rational redesign. Importantly, we have demonstrated that removing diphosphate moiety is not a prerequisite when designing hybrid inhibitors. Linking a fragment to **Pep6** displayed enhanced inhibition showing the promise of this approach for making potent and selective OGT inhibitors.

5.4 Experimental Section

5.4.1 Reagents and general method

Unless stated otherwise, chemicals were obtained from commercial sources and used without further purification. All solvents were purchased from Biosolve (Valkenswaard, The Netherlands) and were stored on molecular sieves (4 Å). Fmoc-L-Glu-OtBu and Fmoc-L-Lys(Alloc)-OH were obtained from Iris Biotech GMBH (Marktredwitz, Germany) and other standard Fmoc amino acids were obtained from GL Biochem Ltd (Shanghai, China).

All reactions and fractions from column chromatography were monitored by thin layer chromatography (TLC) using plates with a UV fluorescent indicator (normal SiO_2, Merck 60 F254). One or more of the following methods were used for visualization: 10% H_2SO_4 in MeOH, molybdenum blue, I_2 or ninhydrine followed by warming until spots could be visible. Flash chromatography was performed using Merck type 60, 230–400 mesh silica gel. Removal of solvent was performed under reduced pressure using a rotary evaporator.

^1H and ^{13}C NMR spectroscopy was carried on an Agilent 400-MR spectrometer operating at 400 MHz for ^1H and 101 MHz for ^{13}C. HSQC, TOCSY and NOESY (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. Analytical LC-MS (electrospray ionization) was performed on Thermo-Finnigan LCQ Deca XP Max using same buffers and protocol as described for analytical HPLC.

Analytical HPLC was performed on a Shimadzu-10AVP (Class VP) system using a Phenomenex Gemini C18 column (110Å, 5 µm, 250×4.60 mm) at a flow rate of 1 mL min. The used buffers were 0.1 % trifluroacetic acid in CH_3CN: H_2O= 5: 95 (buffer A) and 0.1 % trifluroacetic acid in CH_3CN: H_2O = 95: 5 (buffer B). Runs were performed using standard protocol (a) or (b), (a): 100% buffer A for 2 min, then a linear gradient of buffer B (0-100% in 38 min) and UV-absorption was measured at 214 and 254 nm; (b) 100% buffer A for 2 min, then a linear gradient of buffer B (0-100% in 48 min) and UV-absorption was measured at 214 and 254 nm. Purification using preparative HPLC was performed on an Applied Biosystems workstation with a Phenomenex Gemini C18 column (10 µm, 110 Å, 250×21.2 mm) at a flow rate of 6.25 mL/min. Runs were performed by a standard protocol: buffer A for 5 min followed by a linear gradient of buffer B (0 – 100% in 70 min) with the same buffer as described for the analytical HPLC.

5.4.2 Molecular Modelling

Compound library. Fragment libraries from Asinex, ChemBridge, ChemDiv, Enamine, Life chemicals, Key Organics, Maybridge and Otava were downloaded in SDF format. These libraries were merged and duplicates removed, which resulted in a library containing 216,472 compounds. For these compounds a library of conformers was generated using OMEGA software (Release 2.5.1.4, OpenEye Scientific Software, Inc., Santa Fe, NM, USA; www.eyesopen.com)[24] using default settings, which resulted in a maximum of 200 conformers per ligand. The average number of conformers in the library was 51.

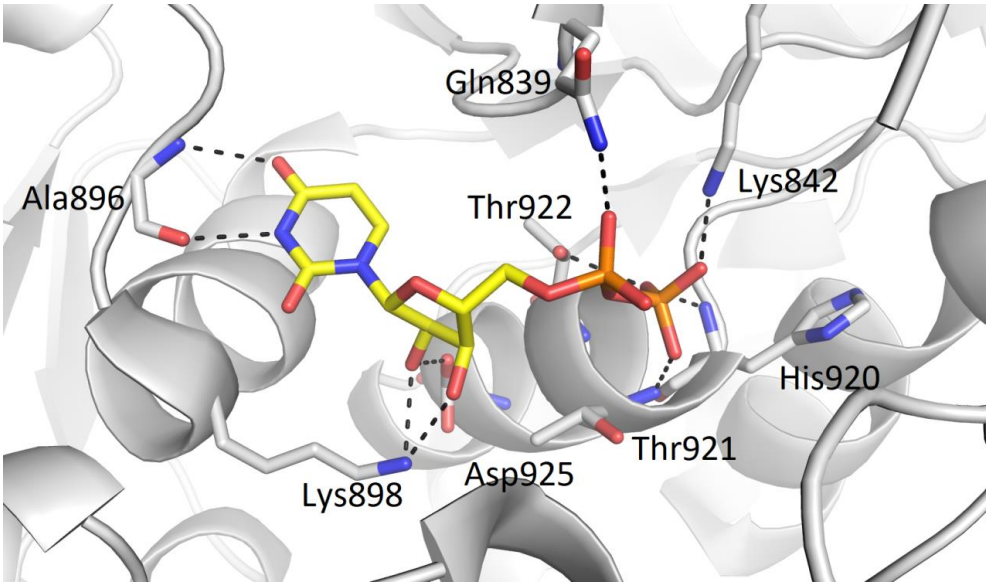


Figure 7. Binding mode of UDP (in yellow sticks) in OGT binding site (in grey cartoon, PDB entry: 4N39). The ligand and the neighbouring protein side-chains are shown as stick models, coloured according to the chemical atom type (blue, N; red, O; orange, S; green, Cl). Hydrogen bonds are indicated by black dotted lines.

Structure-based virtual screening and docking. For docking with FRED software (Release 3.2.0.2, OpenEye Scientific Software, Inc., Santa Fe, NM, USA; www.eyesopen.com), OGT binding site (PDB entry: 4N39)[25] was prepared using MAKE RECEPTOR (Release 3.2.0.2, OpenEye Scientific Software, Inc., Santa Fe, NM, USA; www.eyesopen.com). The grid box around the UDP bound in the OGT crystal structure was generated automatically and was not adjusted. This resulted in a box with the following dimensions: 16.67 Å *15.00 Å * 20.33 Å and the volume of 5083 $Å^3$. For “Cavity detection” slow and effective “Molecular” method was used for detection of binding sites. Inner and outer contours of the grid box were also calculated automatically using “Balanced”

settings for “Site Shape Potential” calculation. The inner contours were disabled. Ala896 was defined as hydrogen bond donor and acceptor constraint for the docking calculations.

Fragment library, prepared by OMEGA, was then docked to the prepared UDP-binding site of OGT (Figure 7, PDB entry: 4N39) using FRED (Release 3.2.0.2. OpenEye Scientific Software, Inc., Santa Fe, NM, USA; www.eyesopen.com).[26,27,28] Docking resolution was set to high, other settings were set as default. A hit list of top 500 ranked molecules was retrieved and the best ranked FRED-calculated pose for each compound was inspected visually and used for analysis and representation.

5.4.3 Chemistry

Synthesis of all the peptides was achieved by following a standard Fmoc SPPS strategy on a Symphony Multiple Peptide Synthesizer and starting from Rink amide resin. The modified “Glu” building blocks **10-12** were prepared and an on-resin approach for preparing compound **1-5** was established as described in Scheme 2.

Peptides were assembled on Rink amide tentagel S resin (227 mg with a resin loading of 0.22 mmol/g). With the exception of the modified building block, peptide couplings were performed with protected Fmoc amino acid (4.0 eq), BOP reagent (4.0 eq), and DIPEA (8.0 eq) in DMF (total volume 5 mL) at ambient temperature for 1 hour.

Incorporation of the modified Glu was performed by using corresponding building block (2.0 eq), BOP reagent (4.0 equiv), and DiPEA (8 eq). Upon completion of SPPS, peptides were cleaved from the resin and deprotected by using a mixture of TFA/EDT/TIPS/H_2O (9:0.25:0.25:0.5, V/V/V/V) followed by Et_2O precipitation to yield the crude peptides. **Pep1-15** were confirmed by LC/MS analysis, in each case providing the expected mass (Table 2). **1-5** were purified by preparative HPLC as described in general method and their identities were confirmed by analytical HPLC and high-resolution mass spectrometry.

Table 2. Analytical data of **Pep1-Pep15**

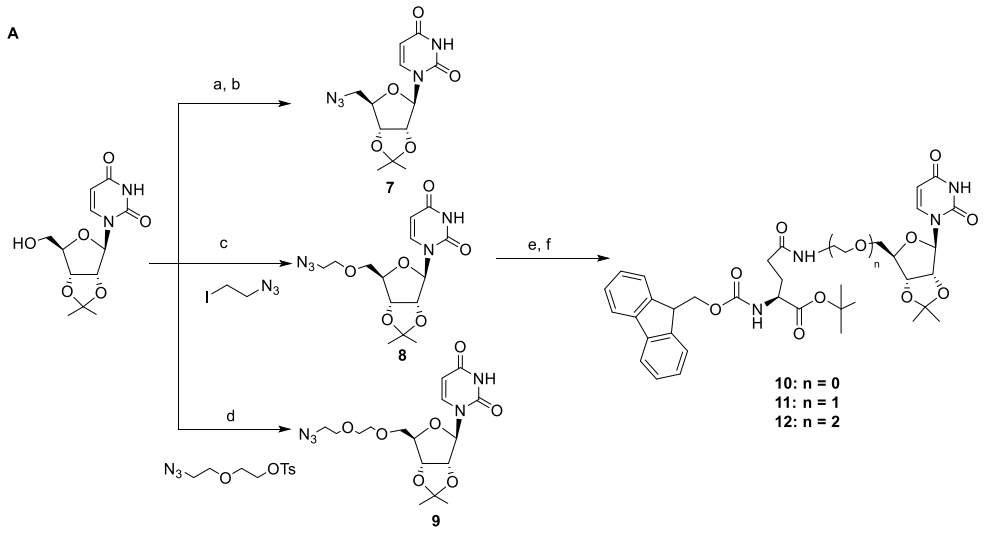
Compound	R_t (min)[a]	Molecular formula	Exact Mass	Measured Mass[b]
Pep1	8.94	$C_{55}H_{94}N_{16}O_{19}S$	1314.66	1313.95
Pep2	11.49	$C_{49}H_{82}N_{14}O_{18}S$	1186.57	1186.71

Pep3	9.20	$C_{44}H_{75}N_{13}O_{15}S$	1057.52	1057.29
Pep4	9.16	$C_{40}H_{69}N_{11}O_{13}S$	943.48	943.86
Pep5	9.58	$C_{37}H_{64}N_{10}O_{11}S$	856.45	856.80
Pep6	8.22	$C_{32}H_{57}N_{9}O_{10}S$	759.39	759.50
Pep7	7.64	$C_{29}H_{52}N_{8}O_{9}$	656.39	656.41
Pep8	6.05	$C_{24}H_{43}N_{7}O_{8}$	557.32	557.48
Pep9	11.00	$C_{80}H_{125}N_{25}O_{28}S$	1915.88	1915.80
Pep10	23.50	$C_{49}H_{79}N_{15}O_{16}$	1133.58	1133.65
Pep11	21.00	$C_{45}H_{74}N_{14}O_{13}$	1018.56	1018.74
Pep12	18.09	$C_{39}H_{63}N_{13}O_{12}$	905.47	905.55
Pep13	10.53	$C_{30}H_{54}N_{12}O_{10}$	742.41	742.51
Pep14	8.24	$C_{24}H_{42}N_{8}O_{9}$	586.31	586.94
Pep15	4.57	$C_{19}H_{33}N_{7}O_{8}$	487.24	487.73

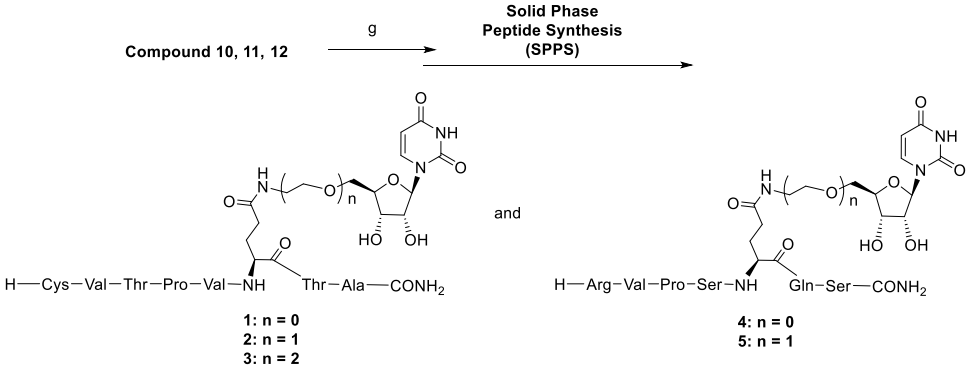
[a]Retention time measured on a a Prosphere C18 column (250 × 4.6 mm, 300 Å, 10 µm) and an acetonitrile gradient (5-95%, 0.1% of TFA) in 40 minutes; flow rate of 1 ml/min. [b] ESI-LRMS.

For preparing **1-5**, modified "Glu" building blocks **10-12** were first synthesized following a method described in Scheme 2 (A). Briefly, a 2',3'-O-isopropylidene uridine precursor was converted into **7-9** which bear azido groups. For **7**, the 5'-hydroxy group was first converted into a *p*-toluenesulfonyl ester with *p*-toluenesulfonylchloride and pyridine as the catalyst/base. The following nucleophilic substitution with sodium azide gave **7** in

high yield. The ensuing catalytic hydrogenation with Lindlar catalyst proceeded to completion within 3 h, and the desired amine derivative was found to be pure by LCMS and used without purification in the next reaction step due to its inherent instability. Treatment of with Fmoc-Glu-OtBu in dry dichloromethane, with DiPEA as the base, gave the **10** in relatively high yield. In the meantime, synthesis of the **8** and **9** began with the synthesis of azido linkers via a modified literature method.[29,30,31] Via the route, two linkers with hundred milligram quantities were readily prepared. These linkers were reacted with the 5'-hydroxy group of 2',3'-O-isopropylidene uridine precursor in CH_3CN with NaOMe under reflux, which produced corresponding **8** and **9**. These intermediates were subsequently transformed into corresponding amine under reaction conditions similar to those used for **7**. The resulting compound confirmed by LCMS were reacted to Fmoc-L-Glu-OtBu to give **11** and **12** in same condition as **10** used. The **10**, **11** and **12** were deprotected by the treatment with TFA/DCM, then were incorporated into **Pep6** and **Pep13** to replace the GlcNAc site residue Ala separately by standard automated SPPS strategy as described in Scheme 2 (B). After deprotection and cleavage from the resin, crude peptides were precipitated in cold Et_2O, which was purified by preparative HPLC. The purity and identity were confirmed via analytical HPLC and high-resolution mass spectrometry.



B



Scheme 2 (A) Synthesis of **1-5**. Reagents and conditions: a) TsCl, pyridine, 0 ^{o}C→ RT, 12h; b) NaN_3, DMF, RT, 12h; c) $ICH_2CH_2N_3$, NaOMe, CH_3CN, 85 ^{o}C, 12h; d) $TsOCH_2CH_2OCH_2CH_2N_3$, NaOMe, CH_3CN, 85 ^{o}C, 12h; e) Lindlar catalyst, H_2, RT, 3h; f) Fmoc-L-Glu-OtBu, Bop, DiPEA, CH_2Cl_2, RT, 20h; (B) Reagents and conditions: g) TFA/CH_2Cl_2, RT, 2h.

5.4.4 Synthetic procedures and compound characterization

The 2',3'-O-isopropylidene uridine and **7** were prepared according to established literature procedures and analytical data were in agreement with published values.[32,33] In addition, 1-azido-2-iodoethane was prepared based on the established procedure and cude was used in next reaction without further purification because of its potential explosive character.[29,34] Similar method was used to produce 2-(2-azidoethoxy)ethyl 4-methylbenzenesulfonate and analytical data was in agreement with published values.[31]

Tert-butylN2-(((9H-fluoren-9-yl)methoxy)carbonyl)-N5-(((3aR,4R,6R,6aR)-6-(2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)-L-glutaminate (10)

7 (600 mg, 1.9 mmol) was dissolved in methanol (150 mL) and argon was bubbled through the solution for 10 min. Lindlar catalyst (200 mg) was added and the suspension stirred under hydrogen. After 3 h, TLC indicated the disappearance of the starting azide and the catalyst was filtered off. The solvent was evaporated under reduced pressure at no more than 40 °C and the colourless oil (550 mg) was achieved. LCMS indicated a high degree of purity so the product was used without purification for further synthesis. The crude (500mg) was dissolved in DCM (50 ml), and N-alpha-(9-Fluorenyl methyloxycarbonyl)-L-glutamic-acid alpha t-butyl ester (Fmoc-L-Glu-OtBu) (901 mg, 2.12 mmol), BOP (0.94 g, 2.12 mmol), and DiPEA (768 µl, 4.42 mmol) were added. The mixture was stirred for 20 hours. The solvent was evaporated and residue was dissolved in EtOAc (100 ml) and washed with 1 M $KHSO_4$ (3 × 100 ml), saturated $NaHCO_3$ (3 × 100ml). The

organic layer was dried over Na_2SO_4 and filtered and the solvent was removed by evaporation. Silica gel chromatography (eluent: DCM/CH_3OH = 98.5/1.5 → 98/2) afforded **10** as colourless oil (700 mg, 1.01 mmol, 57%). R_f = 0.6 (10% CH_3OH in DCM).

^{1}H NMR (400 MHz, DMSO-d_6) δ 8.01 (t, *J* = 5.9 Hz, 1H), 7.86 (dt, *J* = 7.6, 0.9 Hz, 2H), 7.69 (d, *J* = 5.9 Hz, 1H), 7.67 (d, *J* = 6.4 Hz, 1H), 7.63 (d, *J* = 7.9 Hz, 1H), 7.38 (ddd, *J* = 8.3, 7.4, 1.0 Hz, 2H), 7.29 (tt, *J* = 7.4, 1.3 Hz, 2H), 5.72 (s, 1H), 5.71 (d, *J* = 2.1 Hz, 1H), 5.59 (d, *J* = 8.0 Hz, 1H), 4.99 (dd, *J* = 6.5, 2.3 Hz, 1H), 4.65 (dd, *J* = 6.6, 4.3 Hz, 1H), 4.26 (ddd, *J* = 16.4, 9.6, 6.6 Hz, 2H), 4.19 (d, *J* = 6.6 Hz, 1H), 3.98 – 3.90 (m, 1H), 3.90 – 3.80 (m, 1H), 3.37 (dt, *J* = 13.7, 5.7 Hz, 1H), 3.22 (dt, *J* = 13.2, 6.2 Hz, 1H), 2.15 (t, *J* = 7.7 Hz, 2H), 1.98 – 1.80 (m, 1H), 1.81 – 1.68 (m, 1H), 1.43 (s, 3H), 1.35 (s, 9H), 1.24 (s, 3H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 171.99, 171.80, 170.75, 163.65, 156.49, 150.66, 144.21, 143.54, 141.14, 128.06, 127.49, 125.70, 125.67, 120.54, 113.73, 102.21, 92.53, 85.31, 83.86, 81.87, 80.97, 66.05, 60.18, 55.33, 54.56, 47.07, 41.10, 31.94, 28.07, 27.41, 27.09, 25.61.

HRMS (EI) m/z, calculated for $C_{36}H_{42}N_4O_{10}H^+$ [M+H$^+$]: 691.2974, found 691.2983.

The deprotection of **10** was conducted using 10 ml of TFA/DCM (1: 1). After evaporating the solvent fully, the residue was used for SPPS directly without purification.

1-((3aR,4R,6R,6aR)-6-((2-azidoethoxy)methyl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)pyrimidine-2,4(1H,3H)-dione (8)

2',3'-O-isopropylidene uridine (0.50 g, 1.8 mmol) was dissolved in CH_3CN (50 ml). After the addition of 1-azido-2-iodoethane (0.35 g, 1.8 mmol) and NaOMe (0.28 g, 5.2 mmol) at room temperature, the solution was refluxed and stirred overnight. After judging by TLC (CH_2Cl_2: MeOH, 10:1), the solution was concentrated in vacuo, then the crude mixture was dissolved in 200 ml EtOAc and washed with brine (1 × 200 ml) and water (2 × 200 ml). The combined organic layer was dried over Na_2SO_4, filtered, and concentrated in vacuo. The resulting yellow oil was purified by silica gel chromatography eluting with a gradient of CH_2Cl_2: MeOH (100:1 → 50:1) to yield **8** (340 mg, 0.96 mmol, 53%). In the meantime, 90 mg starting material was collected back.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.45 (s, 1H), 5.75 (d, *J* = 8.1 Hz, 1H), 5.63 (d, *J* = 2.7 Hz, 1H), 4.95 (dd, *J* = 6.4, 2.7 Hz, 1H), 4.92 (dd, *J* = 6.4, 3.0 Hz, 1H), 4.34 – 4.27 (m, 1H), 4.14 (t, *J* = 6.2 Hz, 2H), 3.90 (dd, *J* = 12.0, 2.6 Hz, 1H), 3.78 (dd, *J* = 12.0, 3.4 Hz, 1H), 3.51 (t, *J* = 6.2 Hz, 2H), 2.97 (d, *J* = 24.5 Hz, 1H), 1.56 (s, 3H), 1.34 (s, 3H), .

^{13}C NMR (101 MHz, $CDCl_3$) δ 162.58, 151.00, 140.81, 114.39, 101.87, 96.26, 87.07, 84.28, 80.50, 62.71, 48.30, 39.74, 27.33, 25.34.

HRMS (EI, m/z): calculated for $C_{14}H_{19}N_5O_6Na^+$ ([M+Na]$^+$): 376.1228, found: 376.1254.

Tert-butyl N2-(((9H-fluoren-9-yl)methoxy)carbonyl)-N5-(2-(((3aR,4R,6R,6aR)-6-(2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methoxy)ethyl)-L-glutaminate (11)

8 (340 mg, 0.96 mmol) was dissolved in CH_3OH (150 mL) and argon was bubbled through the solution for 10 min. Lindlar catalyst (300 mg) was added and the suspension stirred under hydrogen. After 3 h, TLC indicated the disappearance of the starting azide and the catalyst was filtered off. The solvent was evaporated under reduced pressure at not more than 40 °C and the colourless oil (315 mg) dried was achieved. LCMS indicated a high degree of purity so the product was used without purification for further synthesis. The residue was dissolved in 50 ml dry DCM, and Fmoc-L-Glu-OtBu (491 mg, 1.15 mmol), BOP (512 mg, 1.15 mmol), and DiPEA (419 µl, 2.41 mmol) were added. The mixture was stirred for 20 hours. The solvent was evaporated and residue was dissolved in EtOAc (100 ml) and washed with 1 M $KHSO_4$ (3 × 100 ml), saturated $NaHCO_3$ (3 × 100ml). The organic layer was dried over Na_2SO_4 and filtered and the solvent was removed by evaporation. Silica gel chromatography (eluent: DCM/CH_3OH = 98 /2 → 30/1) afforded product **11** as colourless oil (420 mg, 0.57 mmol, 59% yield in two steps).

^{1}H NMR (400 MHz, DMSO-d_6) δ 7.85 (dt, *J* = 7.6, 0.9 Hz, 2H), 7.80 (d, *J* = 8.1 Hz, 2H), 7.69 (d, *J* = 7.4 Hz, 2H), 7.62 (d, *J* = 7.8 Hz, 1H), 7.38 (td, *J* = 7.5, 1.1 Hz, 2H), 7.29 (tt, *J* = 7.5, 1.2 Hz, 2H), 5.83 (d, *J* = 2.5 Hz, 1H), 5.72 (d, *J* = 8.1 Hz, 1H), 5.10 – 5.02 (m, 1H), 4.86 (dd, *J* = 6.3, 2.6 Hz, 1H), 4.72 (dd, *J* = 6.3, 3.4 Hz, 1H), 4.33 – 4.14 (m, 3H), 4.08 (q, *J* = 4.2 Hz, 1H), 3.93 (t, *J* = 6.3 Hz, 2H), 3.90 – 3.81 (m, 2H), 3.56 (tt, *J* = 9.8, 4.7 Hz, 2H), 3.49 (t, *J* = 6.3 Hz, 2H), 3.38 (t, *J* = 5.8 Hz, 2H), 3.13 (q, *J* = 5.8 Hz, 2H), 2.14 (t, *J* = 7.7 Hz, 2H), 1.97 – 1.83 (m, 1H), 1.73 (m, 1H), 1.45 (s, 3H), 1.35 (s, 9H), 1.25 (s, 3H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 171.86, 171.77, 162.53, 156.48, 150.96, 144.22 , 141.13 , 140.74 , 129.34 , 128.07 , 127.71 , 127.49 , 125.70 , 121.81 , 120.54, 120.45, 113.31 , 101.29 , 92.64 , 87.24 , 84.35 , 80.95 , 80.87 , 66.07 , 61.64 , 54.63 , 47.07 , 36.50 , 32.09 , 28.14 , 28.08 , 27.45 , 27.02 , 25.58.

HRMS (EI) m/z, calculated for $C_{38}H_{46}N_4O_{11}H^+$ [M+H$^+$]: 735.3236, found 735.3244.

The deprotection of **11** was conducted using 10 ml of TFA/DCM (1: 1). After evaporating the solvent fully, the residue was used for SPPS directly without purification.

1-((3aR,4R,6R,6aR)-6-((2-(2-azidoethoxy)ethoxy)methyl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)pyrimidine-2,4(1H,3H)-dione (9)

2',3'-O-isopropylidene uridine (0.60 g, 2.1 mmol) was dissolved in CH_3CN (50 ml). After addition of 2-(2-azidoethoxy)ethyl 4-methylbenzenesulfonate (0.72 g, 2.5 mmol) and NaOMe (0.34 g, 6.3 mmol)at room temperature, the solution was refluxed and stirred for 2 days. After judging by TLC (CH_2Cl_2: MeOH = 5: 1), the solution was concentrated in vacuo, then the crude mixture was dissolved in 200 ml EtOAc and washed with brine (1 × 200 ml) and water (2 × 200 ml). The combined organic layers were dried over Na_2SO_4, filtered, and concentrated in vacuo. The resulting yellow oil was purified by silica gel chromatography eluting with a gradient of CH_2Cl_2: MeOH (50: 1) to yield the compound **9** (620 mg, 1.56 mmol, 74%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.32 (d, *J* = 8.1 Hz, 1H), 5.75 (d, *J* = 8.0 Hz, 1H), 5.54 (d, *J* = 2.8 Hz, 1H), 5.04 (dd, *J* = 6.5, 2.8 Hz, 1H), 4.96 (dd, *J* = 6.5, 3.5 Hz, 1H), 4.29 (td, *J* = 3.5, 2.5 Hz, 1H), 4.23 – 4.11 (m, 2H), 3.91 (dd, *J* = 12.1, 2.5 Hz, 1H), 3.80 (d, *J* = 11.9 Hz, 1H), 3.75 – 3.71 (m, 2H), 3.66 (t, *J* = 5.0 Hz, 2H), 3.33 (td, *J* = 4.7, 1.2 Hz, 2H), 2.83 (s, 1H), 1.57 (s, 3H), 1.36 (d, *J* = 0.9 Hz, 3H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 140.93, 114.43 , 102.20 , 97.14 , 87.25 , 84.03 , 80.41 , 69.64 , 67.58 , 62.86 , 50.86 , 39.98 , 27.40 , 25.40 .

HRMS (EI, m/z): calculated for $C_{16}H_{23}N_5O_7Na^+$ ([M+Na]$^+$): 420.1490, found: 420.1493.

Tert-butyl N2-(((9H-fluoren-9-yl)methoxy)carbonyl)-N5-(2-(2-(((3aR,4R,6R,6aR)-6-(2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methoxy)ethoxy)ethyl)-L-glutaminate (12)

9 (360 mg, 0.91 mmol) was dissolved in methanol (150 mL) and argon was bubbled through the solution for 10 min. Lindlar catalyst (200 mg) was added and the suspension stirred under hydrogen. After 3 h, TLC indicated the disappearance of the starting azide and the catalyst was filtered off. The solvent was evaporated under reduced pressure at not more than 40 °C and the colourless oil (350 mg) dried was achieved. LCMS indicated a high degree of purity so the product was used without purification for further synthesis. The residue was dissolved in 50 ml dry DCM, and Fmoc-L-Glu-OtBu (441 mg, 1.04 mmol), BOP (502 m g, 1.13 mmol), and DiPEA (410 µl, 2.35 mmol) were added. The mixture was stirred for 20 hours. The solvent was evaporated and residue was dissolved in EtOAc (100 ml) and washed with 1 M $KHSO_4$ (3 × 100 ml), saturated $NaHCO_3$ (3 × 100ml). The organic layer was dried over Na_2SO_4 and filtered and the solvent was removed by evaporation. Silica gel chromatography (eluent: DCM/ CH_3OH = 98.5/ 1.5 → 95/ 5) afforded product **12** as colourless oil (510 mg, 0.65 mmol, 70 % yield in two steps).

^{1}H NMR (400 MHz, DMSO-d_6) δ 7.85 (dt, *J* = 7.6, 0.9 Hz, 2H), 7.80 (d, *J* = 8.1 Hz, 2H), 7.69 (d, *J* = 7.4 Hz, 2H), 7.62 (d, *J* = 7.8 Hz, 1H), 7.38 (td, *J* = 7.5, 1.1 Hz, 2H), 7.29 (tt, *J* = 7.5, 1.2 Hz,

2H), 5.83 (d, *J* = 2.5 Hz, 1H), 5.72 (d, *J* = 8.1 Hz, 1H), 5.10 – 5.02 (m, 1H), 4.86 (dd, *J* = 6.3, 2.6 Hz, 1H), 4.72 (dd, *J* = 6.3, 3.4 Hz, 1H), 4.33 – 4.14 (m, 3H), 4.08 (q, *J* = 4.2 Hz, 1H), 3.93 (t, *J* = 6.3 Hz, 2H), 3.90 – 3.81 (m, 2H), 3.56 (tt, *J* = 9.8, 4.7 Hz, 2H), 3.49 (t, *J* = 6.3 Hz, 2H), 3.38 (t, *J* = 5.8 Hz, 2H), 3.13 (q, *J* = 5.8 Hz, 2H), 2.14 (t, *J* = 7.7 Hz, 2H), 1.97 – 1.83 (m, 1H), 1.73 (ddd, *J* = 16.2, 11.6, 7.7 Hz, 1H), 1.45 (s, 3H), 1.35 (s, 9H), 1.25 (s, 3H).

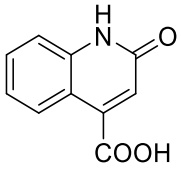
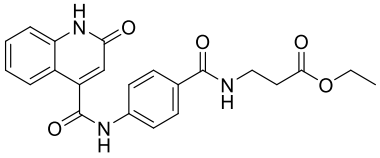
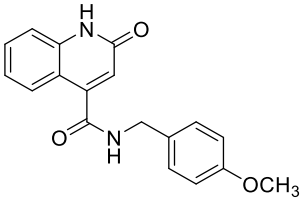
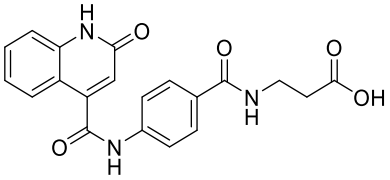
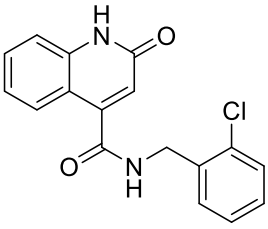
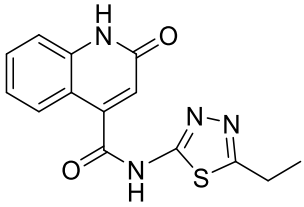
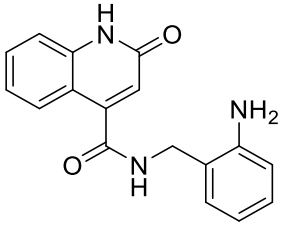
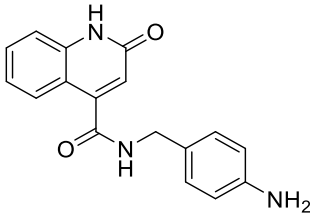
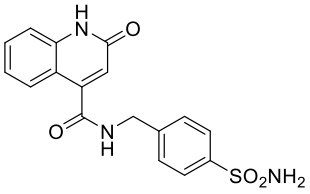
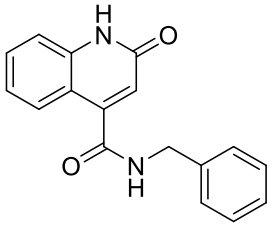
^{13}C NMR (101 MHz, DMSO-d_6) δ 171.81, 171.71, 156.48, 150.85, 144.22, 141.14 , 140.85 , 128.05 , 127.48 , 125.69 , 125.67 , 120.52 , 113.31 , 101.24 , 92.66 , 87.26 , 84.34 , 80.95 , 80.90 , 69.13 , 66.78 , 66.06 , 61.66 , 54.60 , 47.08 , 45.50, 38.99 , 31.96 , 28.07 , 27.45 , 27.13 , 25.59 .

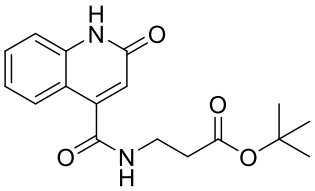
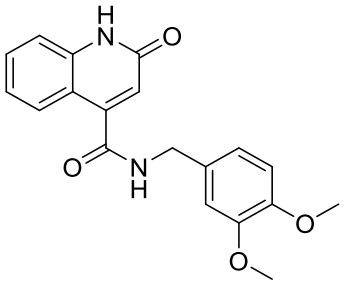
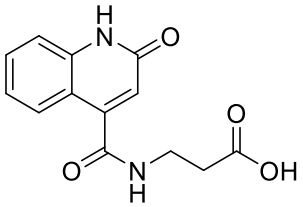
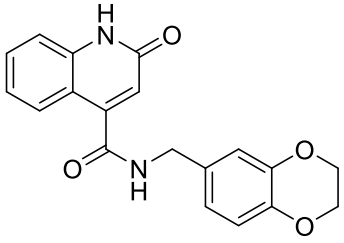
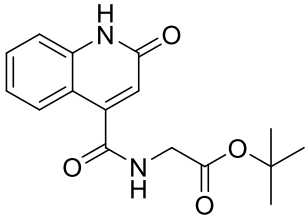
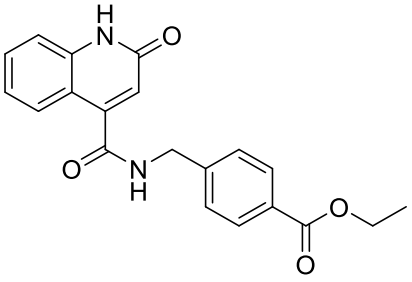
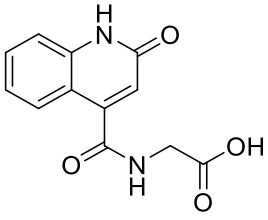
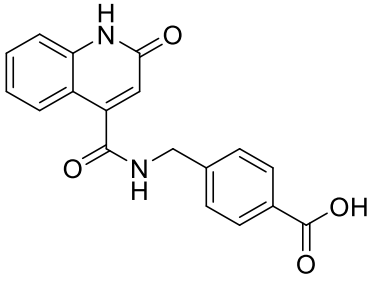
HRMS (EI, m/z): calculated for $C_{40}H_{50}N_4O_{12}Na^+$ ([M+Na]$^+$): 801.3317, found: 801.3325.

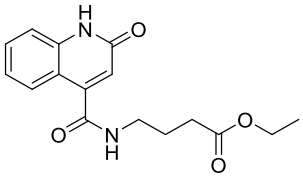
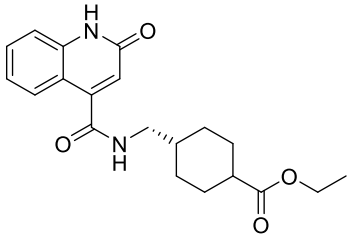
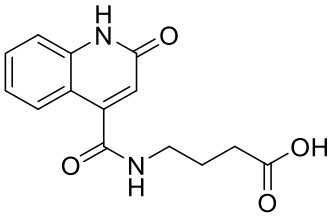
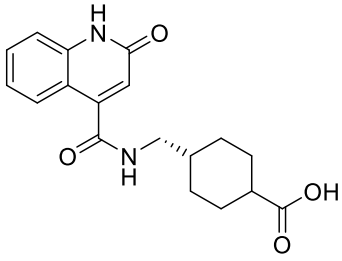
The deprotection of **12** was conducted using 10 ml of TFA/DCM (1:1). After evaporating the solvent fully, the residue was used for SPPS directly without purification.

5.4.5 Structures of prepared fragments F01-F22

Table 3. Structures of prepared fragments **F01-F22**.

Cpd.		Cpd.	
F01		F12	
F02		F13	
F03		F14	
F04		F15	
F05		F16	

F06		F17	
F07		F18	
F08		F19	
F09		F20	

F10		F21	
F11		F22	

5.4.6 Analytical HRMS and HPLC traces for Pep6, Pep13 and compounds 1-6

Pep6: HRMS (EI, m/z): calculated for $C_{32}H_{57}N_9O_{10}SH^+$ ([M+H]$^+$): 760.4022, found: 760.4026.

Pe13: HRMS (EI, m/z): calculated for $C_{30}H_{54}N_{12}O_{10}H^+$ ([M+H]$^+$): 743.4164, found: 743.4144.

1: HRMS (EI, m/z): calculated for $C_{43}H_{70}N_{12}O_{16}SH^+$ ([M+H]$^+$): 1043.4826, found: 1043.4816.

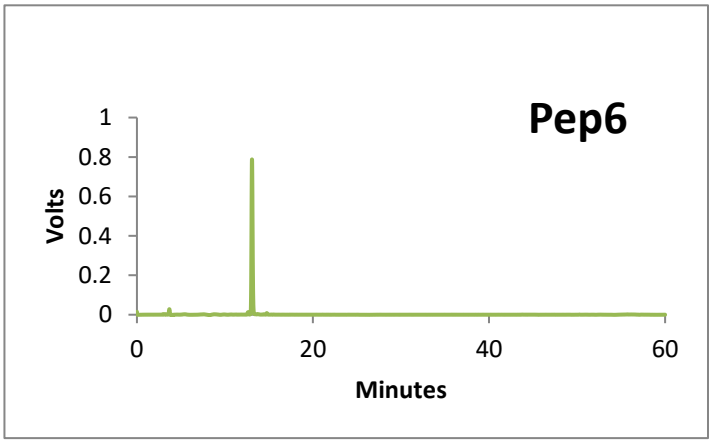
2: HRMS (EI, m/z): calculated for $C_{45}H_{74}N_{12}O_{17}SH^+$ ([M+H]$^+$): 1087.5088, found: 1087.5096.

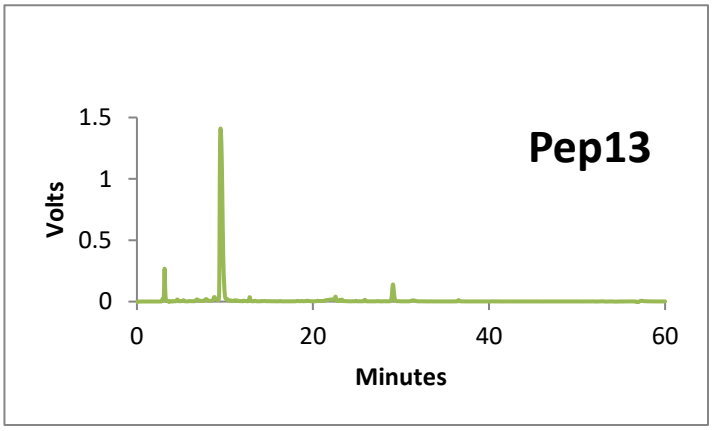
3: HRMS (EI, m/z): calculated for $C_{47}H_{78}N_{12}O_{18}SH^+$ ([M+H]$^+$): 1131.5351, found: 1131.5306.

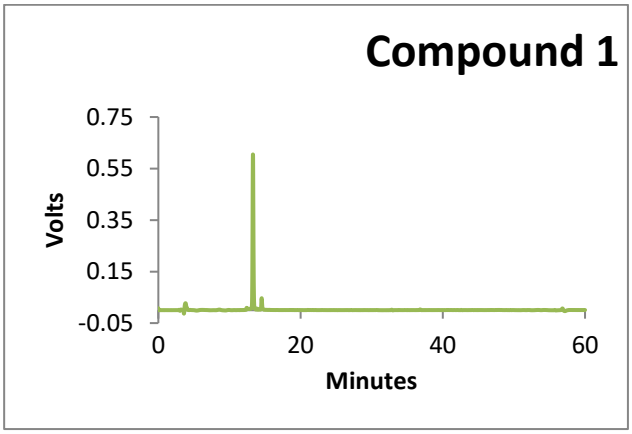
4: HRMS (EI, m/z): calculated for $C_{41}H_{67}N_{15}O_{16}H^+$ ([M+H]$^+$): 1026.4963, found: 1026.4960.

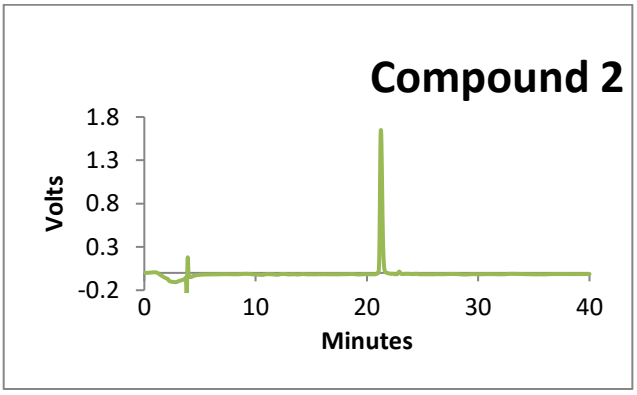
5: HRMS (EI, m/z): calculated for $C_{43}H_{71}N_{15}O_{17}H^+$ ([M+2H]$^{2+}$/2): 535.7649, found: 535.7656.

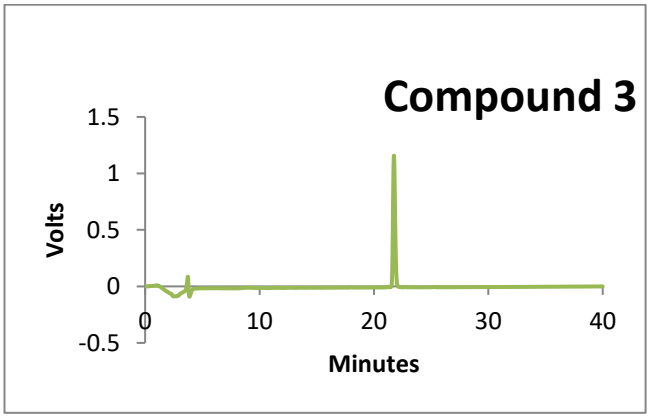
6: HRMS (EI, m/z): calculated for $C_{53}H_{76}N_{12}O_{13}SH^+$ ([M+H]$^+$): 1121.5454, found: 1121.5455

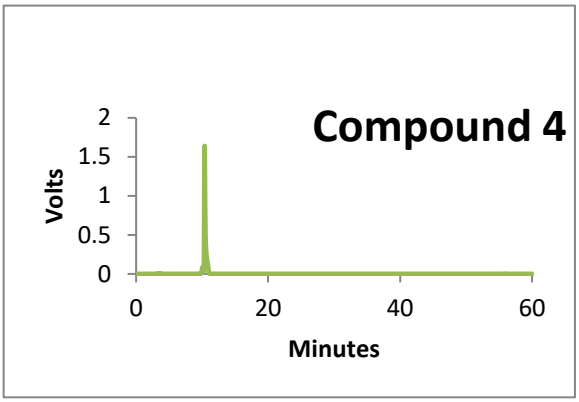


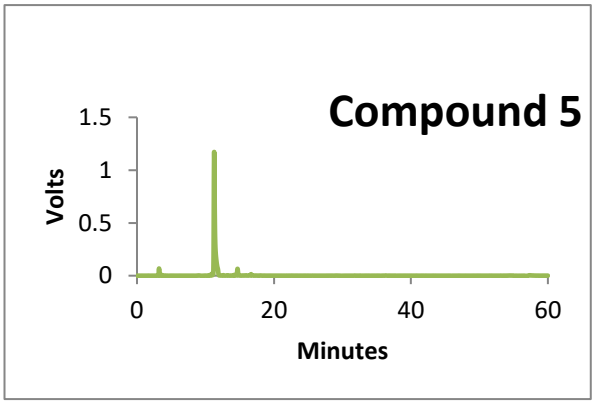


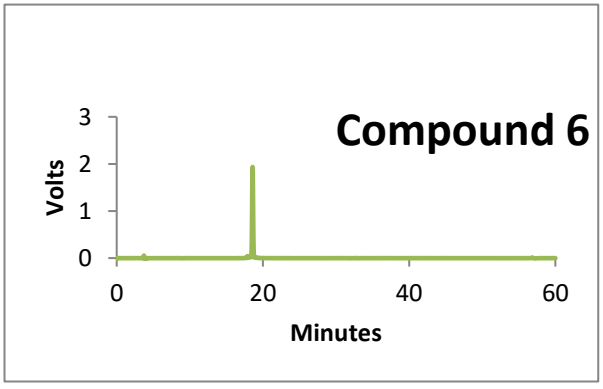

Compound 1
0.75
0.55
0.35
0.15
-0.05
Volts
0
20
40
60
Minutes


Compound 2
1.8
1.3
0.8
0.3
-0.2
Volts
0
10
20
30
40
Minutes

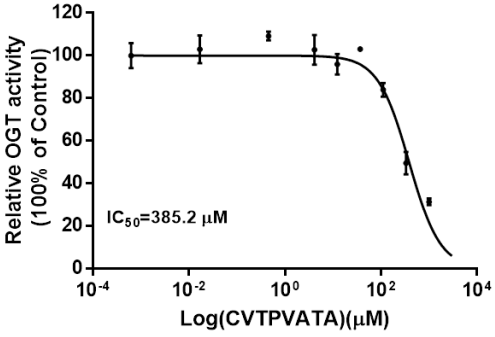

Compound 3
1.5
1
0.5
0
-0.5
Volts
0
10
20
30
40
Minutes


Compound 4
Volts
2
1.5
1
0.5
0
0
20
40
60
Minutes

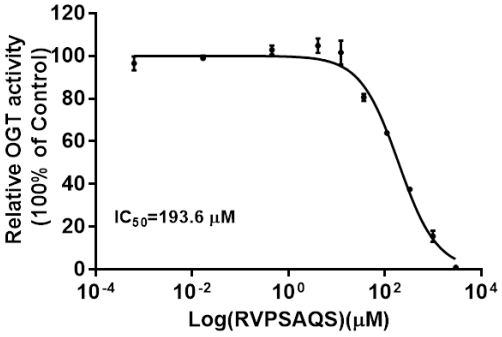

Compound 5
Volts
1.5
1
0.5
0
0
20
40
60
Minutes


Compound 6
Volts
3
2
1
0
0
20
40
60
Minutes

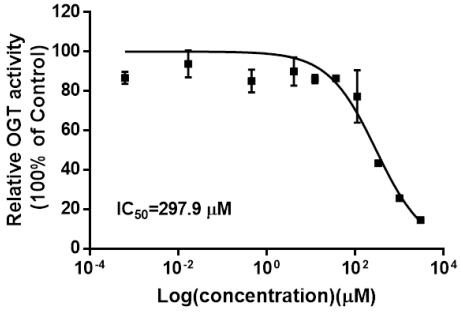
5.4.7 Inhibition effect of compounds in UDP-Glo assay



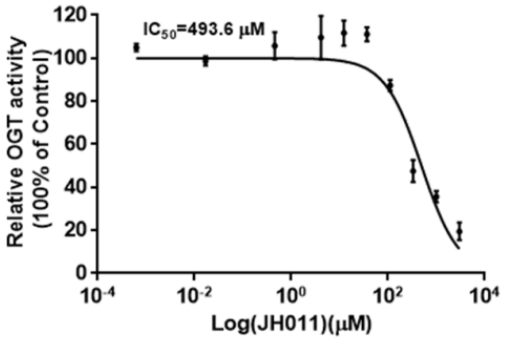
Inhibition of OGT for **Pep6**



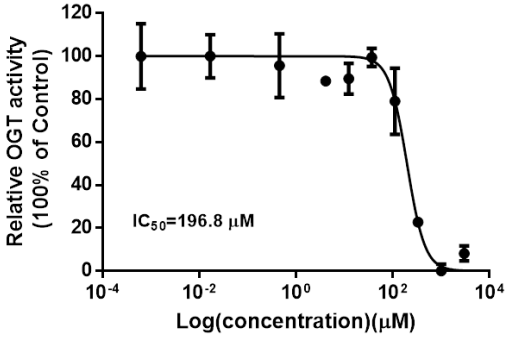
Inhibition of OGT for **Pep13**



Inhibition of OGT for **Compound 1**



Inhibition of OGT for **Compound 3**



Inhibition of OGT for **Compound 6**

5.5 Reference

(1) Hanover, J. A.; Krause, M. W.; Love, D. C. Bittersweet Memories: Linking Metabolism to Epigenetics through O-GlcNAcylation. *Nat. Rev.* **2012**, *13* (5), 312–321.

(2) Daou, S.; Mashtalir, N.; Hammond-martel, I.; Pak, H.; Yu, H.; Sui, G. Crosstalk between O-GlcNAcylation and Proteolytic Cleavage Regulates the Host Cell Factor-1 Maturation Pathway. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (7), 2747–2752.

(3) Butkinaree, C.; Park, K.; Hart, G. W. Biochimica et Biophysica Acta O-Linked β-N-Acetylglucosamine (O-GlcNAc): Extensive Crosstalk with Phosphorylation to Regulate Signaling and Transcription in Response to Nutrients and Stress. *Biochim. Biophys. Acta* **2010**, *1800* (2), 96–106.

(4) Groves, J. A.; Lee, A.; Yildirir, G.; Zachara, N. E. Dynamic O-GlcNAcylation and Its Roles in the Cellular Stress Response and Homeostasis. *Cell Stress Chaperones* **2013**, *18*, 535–558.

(5) Hart, G. W.; Slawson, C.; Ramirez-correa, G.; Lagerlof, O. Cross Talk Between O-GlcNAcylation and Phosphorylation: Roles in Signaling, Transcription, and Chronic Disease. *Annu.Rev.Biochem* **2011**, No. 80, 825–858.

(6) Dias, W. B.; Hart, G. W. O-GlcNAc Modification in Diabetes and Alzheimer's Disease. *Mol. Biosyst.* **2007**, *3*, 766–772.

(7) Slawson, C.; Hart, G. W. O-GlcNAc Signalling: Implications for Cancer Cell Biology. *Nat. Rev.* **2011**, *11* (9), 678–684.

(8) Queiroz, R. M. De; Carvalho, É.; Dias, W. B. O-GlcNAcylation: The Sweet Side of the Cancer. *Front. Oncol.* **2014**, *4* (132), 1–10.

(9) Slawson, C.; Copeland, R. J.; Hart, G. W. O-GlcNAc Signaling: A Metabolic Link between Diabetes and Cancer? *Trends Biochem. Sci.* **2010**, *35* (10), 547–555.

(10) Fardini, Y.; Dehennaut, V.; Lefebvre, T.; Issad, T. O-GlcNAcylation: A New Cancer Hallmark? *Front. Endocrinol. (Lausanne).* **2013**, *4* (99), 1–14.

(11) Harwood, K. R.; Hanover, J. A. Nutrient-Driven O-GlcNAc Cycling-Think Globally but Act Locally. *J. Cell Sci.* **2014**, *127*, 1857–1867.

(12) Hart, G. W.; Housley, M. P.; Slawson, C. Cycling of O-Linked β-N-Acetylglucosamine on Nucleocytoplasmic Proteins. *Nature* **2007**, *446* (26), 1017–1022.

(13) Ferrer, C. M.; Lynch, T. P.; Sodi, V. L.; Falcone, J. N.; Schwab, L. P.; Peacock, D. L.; Vocadlo, D. J.; Seagroves, T. N.; Reginato, M. J. Article O-GlcNAcylation Regulates Cancer Metabolism and Survival Stress Signaling via Regulation of the HIF-1 Pathway. *Mol. Cell* **2014**, *54* (5), 820–831.

(14) Lazarus, M. B.; Jiang, J.; Gloster, T. M.; Zandberg, W. F.; Whitworth, G. E.; Vocadlo, D. J.; Walker, S. Structural Snapshots of the Reaction Coordinate for O-GlcNAc Transferase. *Nat. Chem. Biol.* **2012**, *8* (12), 966–968.

(15) Lazarus, M. B.; Nam, Y.; Jiang, J.; Sliz, P.; Walker, S. Structure of Human O-GlcNAc Transferase and Its Complex with a Peptide Substrate. *Nature* **2011**, *469* (7331), 564–567.

(16) Borodkin, V. S.; Schimpl, M.; Gundogdu, M.; Rafie, K.; Dorfmueller, H. C.; Robinson, D. a; Aalten, D. M. F. Van. Bisubstrate UDP-Peptide Conjugates as Human O-GlcNAc Transferase Inhibitors. *Biochem. J.* **2014**, *457* (3), 497–502.

(17) Trapannone, R.; Rafie, K.; Aalten, D. M. F. Van. O-GlcNAc Transferase Inhibitors: Current Tools and

Future Challenges. *Biochem. Soc. Trans.* **2016**, *44* (part 1), 88–93.

(18) Dorfmueller, H. C.; Borodkin, V. S.; Blair, D. E.; Pathak, S.; Navratilova, I.; Aalten, D. M. F. Van. Substrate and Product Analogues as Human O-GlcNAc Transferase Inhibitors. *Amino Acids* **2011**, *40* (3), 781–792.

(19) Ortiz-meoz, R. F.; Jiang, J.; Lazarus, M. B.; Orman, M.; Janetzko, J.; Fan, C.; Duveau, D. Y.; Tan, Z.; Thomas, C. J.; Walker, S. A Small Molecule That Inhibits OGT Activity in Cells. *ACS Chem. Biol.* **2015**, No. 10, 1392–1397.

(20) Shi, J.; Sharif, S.; Ruijtenbeek, R.; Pieters, R. J. Activity Based High-Throughput Screening for Novel O-GlcNAc Transferase Substrates Using a Dynamic Peptide Microarray. *PLoS One* **2016**, *11* (3), e0151085.

(21) Pathak, S.; Alonso, J.; Schimpl, M.; Rafie, K.; Blair, D. E.; Borodkin, V. S.; Schüttelkopf, A. W.; Albarbarawi, O.; van Aalten, D. M. F. The Active Site of O-GlcNAc Transferase Imposes Constraints on Substrate Sequence. *Nat. Struct. Mol. Biol.* **2015**, *22* (9), 744–750.

(22) Shi, J.; Tomašič, T.; Sharif, S.; Brouwer, A. J.; Anderluh, M.; Ruijtenbeek, R. Peptide Microarray Analysis of the Cross-Talk Between O-GlcNAcylation and Tyrosine Phosphorylation. *FEBS Lett.* **2017**, *591* (13), 1872–1883.

(23) Schimpl, M.; Zheng, X.; Borodkin, V. S.; Blair, D. E.; Ferenbach, A. T.; Schüttelkopf, A. W.; Navratilova, I.; Aristotelous, T.; Albarbarawi, O.; Robinson, D. A.; Macnaughtan, M. A.; Aalten, D. M. F. Van. O-GlcNAc Transferase Invokes Nucleotide Sugar Pyrophosphate Participation in Catalysis. *Nat. Chem. Biol.* **2012**, *8* (12), 969–74.

(24) Hawkins, P. C. D.; Skillman, A. G.; Warren, G. L.; Ellingson, B. A.; Stahl, M. T. Conformer Generation with OMEGA: Algorithm and Validation Using High Quality Structures from the Protein Databank and Cambridge Structural Database. *J. Chem. Inf. Model.* **2010**, *50* (4), 572–584.

(25) Lazarus, M. B.; Jiang, J.; Kapuria, V.; Bhuiyan, T.; Janetzko, J.; Zandberg, W. F.; Vocadlo, D. J.; Herr, W.; Walker, S. HCF-1 Is Cleaved in the Active Site of O-GlcNAc Transferase. *Science (80-.).* **2013**, *342* (6), 1235–1238.

(26) Mcgann, M. FRED Pose Prediction and Virtual Screening Accuracy. *J. Chem. Inf. Model.* **2011**, *51* (1), 578–596.

(27) Mcgaughey, G. B.; Sheridan, R. P.; Bayly, C. I.; Culberson, J. C.; Kreatsoulas, C.; Lindsley, S.; Maiorov, V.; Truchon, J.; Cornell, W. D. Comparison of Topological, Shape, and Docking Methods in Virtual Screening. *J. Chem. Inf. Model.* **2007**, *47* (4), 1504–1519.

(28) Mcgann, M. R.; Almond, H. R.; Nicholls, A.; Grant, J. A.; Brown, F. K. Gaussian Docking Functions. *Biopolymers* **2003**, *68*, 76–90.

(29) Popr, M.; Hybelbauerová, S.; Jindřich, J. A Complete Series of 6-Deoxy-Monosubstituted Tetraalkylammonium Derivatives of α-, β-, and γ -Cyclodextrin with 1,2,and 3 Permanent Positive Charges. *Beilstein Joural Org. Chem.* **2014**, *10*, 1390–1396.

(30) Demko, Z. P.; Sharpless, K. B. An Intramolecular [2+3] Cycloaddition Route to Fused 5-Heterosubsittuted Tetrazoles. *Org. Lett.* **2001**, *3* (25), 4091–4094.

(31) Abellán-flos, M.; Tanç, M.; Supuran, C. T.; Vincent, S. P. Biomolecular Chemistry Exploring Carbonic Anhydrase Inhibition with Multimeric Coumarins Displayed on a Fullerene Scaffold. *Org. Biomol. Chem.* **2015**, *13* (27), 7445–7451.

(32) Gobec, S.; Gravier-pelletier, C.; Le, Y.; Pec, S.; Babic, A. Synthesis of 1-C-Linked Diphosphate Analogues of UDP- N -Ac-Glucosamine and UDP- N -Ac-Muramic Acid. **2008**, *64*, 9093–9100.

(33) Fengjuan, S.; Xiaoliu, L.; Xiaoyuan, Z.; Zhanbin, Q.; Qingmei, Y.; Hua, C.; Jinchao, Z. Synthesis and Biological Activities of Novel S-Triazine Bridged Dinucleoside Analogs. *Chinese J. Chem.* **2011**, *29*, 1205–1210.

(34) Amides, B.; Lang, S.; Murphy, J. A. The 2-(2-Azidoethyl) Cycloalkanone Strategy for Bridged Amides and Medium-Sized Cyclic Amine Derivatives in the Aubé-Schmidt Reaction. *SYNLETT* **2010**, *No.4*, 0529–0534.

Chapter 6

Summary

Samenvatting

6.1 Summary

The work described in this thesis focuses on the development of inhibitors for two distinct types of proteins that act on carbohydrates: the galectins and O-GlcNAc transferase (OGT). These two types of protein can bind or convert certain carbohydrates to exhibit different functions in a wide variety of biological processes.

In **chapter 1**, a general introduction is given on the galectins, a group of proteins that recognize β-galactosides specifically and contain conserved carbohydrate-recognition domains (CRDs) consisting of about 130 amino acids. To date, we know there are 15 members in the galectin family, and they are present in organisms ranging from nematodes to mammals. Galectins display an intriguing combination of intra- and extracellular activities. To decipher the precise action mechanisms of different galectins and explore their therapeutic potential, selective inhibitors are an urgent need. Currently several potent galectin inhibitors have been developed with an emphasis on galectin-1 and -3. Based on their structural features, these inhibitors can be divided into three types: carbohydrate-based small molecule inhibitors, non-carbohydrate-based inhibitors, and carbohydrate-based multivalent inhibitors. Among these inhibitors, small molecule inhibitors based on thiodigalactoside (TDG) have shown their potential while aryl modifications at their C3 position improved selectivity and potency. In addition, non-carbohydrate-based inhibitors such as calixarene 0118 were discovered, that can bind to Gal-1 on the opposite side of the lectin's carbohydrate binding site. The third type of galectin inhibitors, i.e the multivalent carbohydrate based ones, usually make use of a suitable scaffolds to bear multiple ligands so as to achieve a multivalency effect that enhances the interactions between ligands and proteins.

Chapter 2 introduces O-GlcNAcylation as an example of protein glycosylation. Protein glycosylation is the process by which a carbohydrate, i.e. the glycosyl donor, is covalently linked to a target protein (the glycosyl acceptor). The resulting modified proteins can exhibit different functions and subsequently be involved in numerous biological processes. While most glycosylated proteins are present on the cell surface, O-GlcNAc transferase (OGT) acts entirely on proteins that function intracellularly. It can catalyze the attachment of O-GlcNAc to serines and threonines of certain proteins. OGT is emerging as a topic of interest, and selective OGT inhibitors are strongly needed to further decipher its roles in biological processes and explore its potential as a therapeutic target. Based on their structural features, inhibitors can be classified into three types: sugar donor mimic inhibitors, bi-substrate inhibitors and small molecule inhibitors identified from high throughput screening. However, despite significant progress, most of these compounds can't permeate into cells or are not selective enough and show off-target effects, which restricts their use in research and potential applications.

Chapter 3 describes the synthesis of neo-glycoproteins as multivalent galectin inhibitors. The approach involves conjugating NHS-functionalized thiodigalactoside (TDG)-based inhibitors to lysine residues of the BSA serum protein. To obtain the NHS functionalized-TDG, a synthetic approach was developed to create a non-symmetrical TDG derivative linked to a PEG-based spacer terminating in an NHS ester. Subsequent reaction with BSA gave multivalent TDG-glycoconjugates. Slightly alkaline pH was found to be crucial for an effective conjugation. To the best of our knowledge, this is the first example of the conjugation of a TDG derivative to a carrier. In collaboration with the biomaterials group of Professor Lothar Elling (RWTH Aachen University), inhibitory potencies of these conjugates were studied *in vitro*, which showed that the multivalent presentation of the conjugates unlocked the non-symmetrical TDG's full potential. Large multivalency effects were observed that resulted in one of the most effective inhibition of Gal-3 *in vitro* until now. In the future, TDGs with different functional groups on the C3 and C3' position (e.g. 4-phenoxyphenyl) will be involved in new conjugates, that potentially provide a higher galectin selectivity on a multivalent level.

Chapter 4 describes the synthesis of hetero-bivalent or hybrid galectin inhibitors. This was approached through the conjugation of a TDG derivative to a calixarene 0118 derivative. The molecules were designed to bind to the two different binding sites on the galectins simultaneously. To obtain these hybrid compounds, TDG and calixarene components were synthesized separately after which they were linked by a CuAAC "click" reaction. Using our compounds, NMR spectroscopy was used to study the interactions between ligands with Gal-1 and Gal-3, which was conducted by Dr. Hans Ippel (University of Maastricht) in collaboration with Professor Kevin Mayo (University of Minnesota). The data showed that the new hybrid inhibitors indeed strongly bound to galectins-1 and -3 but that the binding mainly resulted from the TDG moiety and the desired "chelating" binding effect for hybrid compounds was not observed. In addition, cytotoxicity studies were done to evaluate proliferation inhibition of these compounds on cancer cells in collaboration with the medical oncology group of Professor Arjan Griffioen (VU University Medical Center). By comparison with the parent TDG and calixarene 0118 compounds, we found that the hybrid compounds affected cell viability greatly, but the effects were all due to the calixarene moiety. Interestingly, the killing effect on non-cancerous HUVECs was less than that of the two cancer cell lines, suggesting that the TDG may possibly enhance selectivity.

Chapter 5 describes an approach to obtain OGT inhibitors targeting both the donor and acceptor sites. The approach used a combination of peptide optimization and a fragment-based method. Two substrate peptides were previously identified by Jie Shi within our group, and their potential to be developed into OGT inhibitors was evaluated by substituting the O-GlcNAc site serine with an alanine. Through subtracting amino acids

from the found peptides, we identified two peptide inhibitors targeting the OGT acceptor substrate site. Furthermore, we designed and synthesized bisubstrate or hybrid inhibitors by conjugating these peptides to uridine, a moiety of the sugar donor. Using the right linker was important to observe a synergy effect between the two components of the hybrid inhibitor. In parallel, an *in silico* fragment screening was conducted to obtain small molecules blocking the sugar donor site of OGT. This was performed in collaboration with Professor Marko Anderluh (University of Ljubljana). The evaluation of the initial hits provided us with four quinolone-4-carboxamide analogues that significantly inhibited OGT *in vitro*. The common unit was used as a replacement of uridine in novel hybrid inhibitors. These compounds inhibit OGT activity with IC_{50} values in the micromolar range and structural data could facilitate the further development and ultimately their application in biomedical research.

6.2 Samenvatting in het Nederlands

Het werk beschreven in dit proefschrift gaat over de ontwikkeling van remmers van twee verschillende typen eiwiten die beide werken op koolhydraten: de galectines en O-GlcNAc transferase (OGT). Beide typen eiwitten kunnen binden aan koolhydraten of kunnen deze omzetten en hierdoor invloed hebben op een variëteit aan biologische processen.

In **hoofdstuk 1** wordt een algemene inleiding over de galectines gegeven, een groep eiwitten die specifiek β-galactosides binden en die een geconserveerd koolhydraat bindend domain (CRD) van ca. 130 aminozuren bevatten. Tot nu toe weten we dat er 15 leden van de galectine familie bestaan en dat ze aanwezig zijn in vele organismes variërend van nematoden tot zoogdieren. Galectins hebben een intrigerende hoeveelheid intra- en intercellulaire biologische activiteiten. Om al deze activiteiten in detail verder op te helderen en uiteindelijk gebruik te maken van hun therapeutische potentiaal zijn selectieve remmers dringend nodig. Momenteel zijn er meerder krachtige galectine remmers ontwikkeld met een nadruk op galectines 1 en 3. Op basis van hun structuren zijn deze remmers in te delen in drie groepen: kleine koolhydraten en hun derivaten, niet-koolhydraten en multivalente koolhydraten. Onder deze remmers zijn de thiodigalactosiden (TDG) veelbelovend vooral als de C3 positie aromatische substituenten bevat want dan wordt de remming aanzienlijk versterkt en de selectiviteit vergroot. Ook zijn niet-suiker remmers beschreven zoals het calixareen 0118 dat galectine-1 kan binden aan de andere kant dan de CRD. Het derde type galectine remmers, de multivalente koolhydraten, maken doorgaans gebruik van een geschikt kernmolecuul waaraan meerdere remmers gekoppeld kunnen worden zodat de interacties met de eiwitten versterkt kunnen worden door het multivalentie principe.

In **hoofdstuk 2** wordt O-GlcNAcylering geintroduceerd als een voorbeeld van eiwit glycosylering. Eiwit glycosylering is een proces waarin een koolhydraat (de donor) wordt gekoppeld aan een bepaald eiwit (de acceptor). Eiwit glycosylering in de biologie refereert vooral naar het enzym gedreven proces waarin koolhydraatketens aan eiwitten gekoppeld worden en de gemodificeerde eiwitten andere functies krijgen en als zodanig invloed op bepaalde biologische processen kunnen hebben. Hoewel de meeste geglycosyleerde eiwitten op het celoppervlak zitten, werkt de O-GlcNAc transferase alleen op eiwitten die binnen de cel hun werk doen. Het kan de koppeling van N-acetyl glucosamine (GlcNAc) aan bepaalde serines en threonines van bepaalde eiwitten katalyseren. OGT is een onderwerp met een groeiende belangstelling en selectieve remmers van OGT zijn urgent nodig om de verschillende rollen in biologische processen verder op te helderen en verder te onderzoeken of remmers ervan geneesmiddelen kunnen worden. Op basis van hun structuren zijn er drie typen remmers van OGT: suiker-donor mimetica, bi-substraat remmers en kleine moleculen verkregen van “high throughput screening”. Ondanks flinke

vooruitgang op dit gebied gaan de meeste remmers niet de cel in of zijn ze onvoldoende selectief en binden ook andere eiwitten, waardoor hun gebruik in biologisch onderzoek en richting toepassingen nu nog beperkt is.

In **hoofdstuk 3** beschrijft de synthese van nieuwe suiker-eiwit conjugaten, ofwel neoglyco-eiwitten, als multivalente galectine remmers. De aanpak betreft de koppeling van NHS-gefunctionaliseerde TDG remmers aan de lysine aminozuren van het serum eiwit BSA. Om de benodigde NHS-gefunctionaliseerde TDG remmers te krijgen, werd een synthetische route ontwikkeld naar een niet-symmetrische TDG gekoppeld aan één kant aan een PEG staart met een NHS ester aan het eind. Vervolgens gaf de koppeling de gewenste TDG-BSA conjugaten. De koppeling verliep alleen goed onder licht basische omstandigheden. Voorzover wij weten is dit het eerste voorbeeld van een koppeling van een TDG derivaat aan een eiwit. In samenwerking met de groep van Professor Lothar Elling (RWTH Universiteit van Aken) werden de remmingen van de nieuwe conjugaten bestudeerd met *in vitro* assays. De resultaten lieten zien dat de effecten van de multivalente presentatie van de TDG remmers groot waren. Grote multivalentie effecten zijn gemeten en het resultaat is één van de meest krachtige galectine-3 remmers ooit. Op basis hiervan kunnen in de toekomst TDG's met verschillende groepen op de C3 en de C3'posities (zoals 4-phenoxyphenyl) geïntroduceerd worden die wellicht een betere selectiviteit te zien zullen geven.

Hoofdstuk 4 beschrijft de synthese van hetero-bivalente of hybride galectine remmers. De aanpak betrof de koppeling van een TDG molecuul aan derivaat van calixareen 0118. De remmers zijn ontworpen om tegelijkertijd de twee verschillende bindingsplaatsen op de galectines te binden. Om deze hybrides te verkrijgen, werden de TDG en calixareen componenten apart gesynthetiseerd waarna ze middels een CuAAC "click" reactie aan elkaar gekoppeld zijn. NMR spectroscopie werd gebruikt om de interacties tussen de nieuwe verbindingen en Gal-1 en Gal-3 te bestuderen, en dat is gedaan door Dr. Hans Ippel (Universiteit van Maastricht) in samenwerking met Professor Kevin Mayo (Universiteit van Minnesota). De data van de studie lieten zien dat de nieuwe hybrides inderdaad sterk bonden aan de galectines-1 en -3, maar dat de binding vrijwel volledig kon worden toegeschreven aan de TDG component en dat de beoogde chelaat binding niet optrad. Verder zijn er ook cytotoxiciteitsstudies uitgevoerd met de verbindingen op kankercellen in samenwerking met de groep van Professor Arjan Griffioen (VU Medisch Centrum). Door de verbindingen te vergelijken met de onderdelen TDG en het calixareen 0118 derivaat, bleek dat de hybrides een groot effect hadden op de levensvatbaarheid van de cellen, maar dat de effecten volledig waren toe te schrijven aan het calixareen deel van de moleculen. Een interessante observatie hierbij was dat de niet-kanker cellijn HUVEC aanzienlijk minder last had van de hybrides dan de calixareen

referentieverbindingen hetgeen suggereert dat het TDG deel wellicht de selectiviteit vergroot.

In **Hoofdstuk 5** wordt een studie beschreven naar remmers van OGT die zowel elementen van de donor als van de acceptor bevatten. De aanpak bestond uit een combinatie van peptide optimalisatie en een 'fragment-based' methode. Gebruik makende van twee eerder gevonden peptidesubstraten werd bestudeerd of deze peptides de basis konden vormen van nieuwe remmers door de serine waar de GlcNAc op komt te vervangen door een alanine. Door verder de lengte van de peptides in te korten werden geschikte remmers gevonden met affiniteit voor de acceptor plaats van OGT. Op basis hiervan zijn bisubstraat of hybride remmers gemaakt waarin de peptides aan uridine, een belangrijk deel van de donor, zijn gekoppeld. De lengte van het verbindingsstuk tussen beide delen bleek erg belangrijk, maar met de juiste versie was het mogelijk een synergie-effect tussen beide delen van het molecuul waar te nemen. Naast deze aanpak werd ook een "*in silico* fragment screening" uitgevoerd met als doel om kleine moleculen te vinden met affiniteit voor de donor locatie van OGT. Dit deel was uitgevoerd in samenwerking met Professor Marko Anderluh (Universiteit van Ljubljana). De modificatie van de eerste hits van de screening gaf ons vier quinolone-4-carboxamides die OGT significant konden remmen. Deze eenheid werd door ons ingebouwd in nieuwe hybride remmers ter vervanging van het uridine deel. De verbindingen remden OGT met bemoedigende micromolaire IC_{50} waarden. In de toekomst kunnen structurele data helpen bij het verbeteren van de remming en uiteindelijke toepassingen in biomedisch onderzoek.

Appendices

Curriculum Vitae

List of Publications

Acknowledgements

Curriculum Vitae

Hao Zhang was born in Shandong, China on March 13th, 1988. In summer 2006, he graduated from No.6 High School of Zibo City and enrolled at Minzu University of China (Beijing) for his bachelor study in chemistry. In 2010, he earned his BSc and then moved to Peking Union Medical College for his master study. During this period, he worked on computer-aided drug design and organic synthesis in the group led by Prof. Peixun Liu. In 2013, he obtained Master of Science degree and was awarded with outstanding graduate of Beijing. In the same year, he moved to the Netherlands and started his PhD research project in Utrecht University. He joined the group of Medicinal Chemistry and Chemical Biology under supervision of Prof. Dr. Roland J. Pieters. His work focused on designing and synthesizing di- and multivalent ligands targeting proteins that act on carbohydrates, the results of which are described in this thesis.

List of Publications

Publications from this thesis

● **Zhang, H.; Laaf, D.; Elling, L.; Pieters, R. J. Thiodigalactoside–Bovine Serum Albumin Conjugates as High-Potency Inhibitors of Galectin-3: An Outstanding Example of Multivalent Presentation of Small Molecule Inhibitors. *Bioconjug. Chem.* 2018, *in press.***

● **Zhang, H.; Tomašič, T.; Shi, J.; Weiss, M.; Ruijtenbeek, R.; Anderluh, M.; Pieters, R. J. Inhibition of O-GlcNAc transferase (OGT) by peptidic hybrids. *Manuscript in preparation.***

● **Zhang, H.; Ippel, H.; Miller, M. C.; Griffioen, A. W.; Pieters, R. J.; Mayo, K. H. Design of hybrid galectin inhibitors targeting dual binding sites with antitumor activities. *Manuscript in preparation.***

Other publications

● **Brittoli, A.; Fallarini, S.; Zhang, H.; Pieters, R. J.; Lombardi, G. "*In vitro*" studies on galectin-3 in human natural killer cells. *Immunology Letters.* 2018, 194, *February*, 4-12.**

● **Zhang, H.; Zan, J.; Yu G.; Jiang, M.; Liu, P. A combination of 3D-QSAR, molecular docking and molecular dynamics simulation studies of benzimidazole-quinolinone derivatives as iNOS inhibitors. *Int. J. Mol. Sci.* 2012, 13, 11210-11227.**

Acknowledgements

After the past four years of my research and social life in Utrecht University, I finally get the chance to look back on it and express my thanks to the people I have met. It is without doubt that the first big thanks should be given to my supervisor, **Prof. dr. Roland J. Pieters**. Since I first met you, I have realized that you are such a "treasure". I am so lucky to have you, who is always generous and kind, lead me into the world of chemistry and guide me to know the carbohydrate, the peptide and above all, the wonderful molecular world. In the last four years, you answered my every question at every moment and always encouraged me to try my immature thoughts. Your effort towards my PhD project is unmatched and this thesis only gives a partial view. Except for the intellectual support, you shared me a lot of experience about how to sell points, which made me aware of the importance of "selling" itself. I will always keep your words "be kind to your audience" in mind, and so let's go next one in case someone is looking for his/her name.

I would like to give my big thanks to **Nathaniel**. I am so grateful to you because you built an awesome working environment in the east lab. Without your effort, I can't easily get skill and help from the other members and can't get through many difficult moments especially in the early stage of my PhD. **Johan**, thanks for all the help with the spectra and patient guide about instrument operation. Even if we didn't get succeed in the project of STD-NMR, the process of learning from you gave me a lot of confidence. Keep running in Groningen! **Javier**, thanks for the measurement of all the high resolution mass spectrometry and timely maintenance of those instruments in 5th floor. **Rob**, I felt a bit hesitated at the beginning of our OGT project until the first meeting with you. Your timely and effective communication made the whole project easier and I learned a lot from those discussions in our regular meetings. **John**, thanks for your help on HPLC and great suggestion on peptide synthesis. You have "warned" me not to use the freeze chemical agent with a hug, which is a great relief for a poor researcher who have searched for "a precious solvent" DMF all day in the lab. **Dirk**, thanks for your great suggestions about peptide synthesis. Sometimes I even felt a bit jealous especially when Xin told me your tolerance towards him (I am sure Roland won't see here). **Ed**, thanks for your guide on the modeling and your trust that involved me into the board of the UIPS PhD committee. Thank you for everything you did for Chinese students and you'd better change your office because I already told "few" Chinese students to turn to you if they have any problem about the salary and house. **Arwin**, thanks for your daily maintenance in the lab. I am sorry I yelled at you once, but it is great to help you know how to handle a crazy Chinese, isn't it?

Prof. dr. Lothar Elling, thanks for your great effort in our collaboration and scientific suggestions to my thesis. To my delight you can join my defense in Utrecht, which is a

great honor of mine. **Dr. Hans Ippel**, thanks for your contribution to our collaboration. Your technic in the biomacromolecule NMR is terrific and I hope we can publish one same terrific article in the near future. **Dominic**, my unseen friend, we two made a great collaboration and I can feel your positive attitude through your every email. I truly hope you can enjoy the late stage of your PhD and have a good life in Aachen.

Xin, a great roommate and friend. Over the four years together I already got used to your accompany and considerate arrangement for our daily life. After departing from Utrecht, I felt like something was missing especially at the moment when I have to cook for others. Please set up your restaurant as soon as possible and get yourself ready because I won't pay the meal. **Jie**, my private consultant. You provided your knowledge for my research and helped me out of those different problems. You always made the work get easier, and I wish you good luck in Spain. **Matthijs**, my lab guard, challenger of Bruce Lee and my buddy. It is so difficult to find suitable words to describe my gratitude for what you have done to me. You told me everything about the country and its culture like a gracious host. You gave me advice and encouragement like an old brother when I felt down or puzzled. You helped me out from different troubles even though I didn't mention them at all. Those daily coffee talks, bar drinking, hangover at chains we did together bring us closer and finally turn into a wisp of "wired" smoke soaking into my memory and soul. Thanks a lot, my friend, best wishes to you and your family and I hope you everything goes well. My another smoke maker, **Alen**, a great teller and my friend. Although you left "nothing" to me in your book, I knew we had a great relationship that nothing could surpass. Thanks for those nice talks that greatly helped me learn about how to behave in a foreign society. You inspired me in many ways, which included but not limited to sausage cooking, techno music and combining closestool with sink. We still have business to discuss and so come to China, I can make Peking Duck, Xiaomi electronics and Jack Ma get ready.

Peter, thank you for all the help especially in my early stage of PhD. You played a great leading role when you were here, in and outside of work. Even now I still can recall your iconic laughter. Good luck with **Nikki** in Germany. **Marjon**, a great listener and friend, For many times you have been patient to listen to my talk even some of them were quite negative. I haven't showed you too much Chinese architecture so maybe that is worth to arrange a visit to China. **Laurens**, I have enjoyed many conversations with you and thanks for your great help in the lab. By the way, at least three Chinese girls told me you looked handsome and I told them you were a papa already. So, stop dreaming! **Timo**, there was time when I felt something wrong after you changed your seat. Until one day I hear your funny talk outside of your new office, I recognized your contribution to our group is dramatically underestimated and the people making you move might make a mistake. Now you are kind of boss for yourself and please remember your funny character can earn

money. **Paul**, I never forget the first time you took me to the "Primus" and told me to choose one beer from a list of funny names (one is "korea wolf" if I remember right). Since then I have stepped into a beer world. I will always drink one more for you. **Kamal**, forgive me I have taken a wrong way for you and your wife on the beach but you could learn about my good intention to help your wife know and admire you more. Sometimes I miss the good days when we played football together, you kicked the ball and the ball kicked me. Take **Vida** to China one day and maybe we can better the Chinese football together. **Frederik**, my man, your coming to my office territory ended a "messy but organized" age of the desk (Yes, that is created by you, **Matthijs**). Unfortunately, I can't stay long with you otherwise sooner or later, I will make your desk like mine. Your humble and gentle character impressed me a lot and thank you for those nice talks. I wish you everything goes well in Utrecht. **Tom**, the current "DJ" of east lab, thanks for the energy you have brought to the group and I felt enjoyable to work with you in lab and committee, even if I couldn't get your every joke. Keep the music on! **Charlotte**, thanks for your kindness to me and it was a great time to play with you in Dota. Don't be too surprised when you see my game name in future WTF, which only means I am trying to deliver happiness to you. And don't be too nice to Tom, he doesn't deserve that. **Tim**, your seat is only three steps away from mine but I still believe it is more quickly to find you through Steam. May you "WINNER WINNER, CHICKEN DINNER" every day. **Jack**, **Apoorva**, **Ingrid**, **Pieter**, **Yongzhi**, **Yvette**, **Gaël**, **Mehman**, **Ivan**, **Rosanne**, **Hanna**, **Núria**, **Victor**, **Seino** and **Tom W.**, all of you built the dynamic team and I always feel energized with your positive input. Thank you all and I wish every happiness will always be with you. **Steffen** and **Gosia**, you were in the group at the beginning of my PhD, but you gave me a huge help that is precious especially for a beginner. Good luck for you!

I would like to thank all the members of "sugar" group, who have helped me in different ways. Special thanks to **Ou**, the first Chinese I met in Utrecht. At first glance you look quite cool but you got "hot" after just two hours. I will never forget the days when you took me to tour the city, try comparatively authentic Chinese food, and play basketball. You became a Chinese DJ in the empty lab and pumped your arms after day's work, which showed me your wild soul under your suntanned skin. Don't try to hide it anymore because that is your nature. I hope we can see each other again in China and best wishes to you and your parents. **Aliaksei**, master of organic chemistry and language. Thank you for sharing your experience and knowledge about organic synthesis especially on carbohydrate chemistry. You have showed me a lot of your tricks on language and I guess only the Chinese can stop your "show off". Good luck with you career. **Barbara**, thank God Roland brought you into our group and let you work at the fume hood next to me. I can finally have an excellent workmate especially in torturous weekend, that turned into not that bad because of you and those talks about your worse experimental results. Without you I can't go through those trouble days and I hope you have a bright future. **Reshmi**, I

didn't meet any psychic tutor until I talked with you. From then on, I started to bother you when I feel confused. Not only that, you can also answer my question about research, and maybe that is why you didn't choose to be a real psychic tutor. Thanks for all the help and I hope you and your family have a good time in the Netherlands. **Suhela**, I am sorry I didn't help you out from your purification problem but you made it by yourself in the end, which remind us not to ask a "problem" to solve your problem even if you know the "problem" can solve his problem. Thank you for being part of OGT/OGA project! **Merel**, my first and only student here, thanks for the help during my PhD and I am honored to be a witness of your route to academic. Good luck in Leuven. Thank **Wenjing**, **Guangyun**, **Diksha** and **Xuan** for all the talks and teamwork. Without your help I can't make my project go well.

Here, I would like to take this opportunity to thank all my Chinese friend. 杨欣，施杰，我们三个人漂洋过海跑到这里相识，一起开始又一起结束，一起度过了这确实不怎么容易的四年，我真是三生有幸能跟你俩并肩作战，千言万语换一句好生保重，后会有期兄弟。鸥，作为你的师弟我只能说科比万岁，虽然你知道我想说 Beat LA。 一起打球打游戏吃吃喝喝是那时候最快乐的日子，祝你早日找到你的 soulmate，也顺祝叔叔阿姨身体健康！阿泰，乌特勒支认识你的人那么多，一半不知道你的大名罗文豪，看在你已经献身祖国科研事业，我也勉为其难叫你一声罗老师，祝你早日晋升，我期待在未来院士名单上看到你的大名。鲁文静，你现在最需要的应该就是鼓励，坚持住，读博士不比减肥难哈哈。光允，好好做自己，有事儿就说话。小岛，看到你媳妇我就知道你早晚会是一特成功人士，好好干，祝你俩荷兰生活开心美满，哥们北京等你。谢谢 2013 年一起到乌特的兄弟姐妹们，每次想到你们都让我心头一暖。谢谢戴美玲，祝你在国外一切顺利。谢谢郭勇，跟你一起损杨欣简直不要太爽。谢谢娄博，如果不嫌弃希望再次合作啊。谢谢刘芳，热心肠的好姑娘。谢谢张玉茂，乌特留学百事通，祝你跟茂嫂生活幸福美满。还有王朝文，温仕成，杨浩然，孙飞龙，郭洪波，赵玉珑，王宏凯，祝你们心想事成！谢谢邓慧师姐，李文涛师兄，黄宇星师兄，石洋师兄，祝你们工作顺利。谢谢 UIPS 的师弟师妹们，韦萱，李静，陈建明，高永智，刘明龙，孙丽凤，陈度伸，张良伟，张宇睿，朝乐梦，姜明，祝你们留学生活顺利！

借此机会我想感谢乌特勒支篮球队的兄弟们。张帅，再过几十年我预感你会是篮球场最烦的那种老大爷，又难缠又能说，祝你跟嫂子王莹天天开心。刘金峰，扎实的后卫，祝你早日回国任教。郑尚镐（Sangho），MVP of UTRECHT, best wishes to you and your future wife. I hope we can play basketball together again and welcome to China！邓博超，谢谢你带我吃遍了乌特和阿姆，哥们等你和 PP 来北京。谢谢范玉，虽然你每逢大赛就失踪，但我还得说你打的不错。谢谢陶崴，你投三分的时候有安东尼和科比一起打铁的样子。还有关吉斌，王梦麟，崔伯龙，熊武，叶诚沛，钱律，杨为晔，郑旭，高宇，李安，姜嘉炜，祝你们在生活和事业里都如球场上那么牛逼。

我想谢谢我的家人，没有你们的无私付出我不可能走到今天。谢谢我的爸爸妈妈，儿子久未尽孝于前，愧疚于心，成此书也不抵万一，唯愿父母身体安康。岳父岳母，蒙您青眼，允我为婿又维护有加，可我于妻于您都照顾未周，今得回国，自当尽心竭力。姐姐姐夫，我不在的日子里为我承责太多，弟弟无以为报，只能好好工作，不负所期。最后，给我的妻子陈妹，你用了最美的年华来等我，也请让我用余生来与你相伴。

再次由衷的谢谢！

张 浩

2018 年 3 月 30 日于北京