Synthetic peptides as tools in chemical immunology

Doelman, W.

5

Synthesis of pre-organized cyclic ligands for the Cholera toxin B-subunit

Introduction

Cholera is the disease caused by the bacterium *Vibrio Cholerae*. It is probably most famous for being the subject of the first-ever epidemiology study, when John Snow showed the role of the local water supply in the deadly spread of Cholera in London in 1854.[1] However, despite this historical nature and intimate knowledge of the disease, it remains a major health problem today. Cholera is responsible for an estimated 100,000 deaths annually, primarily in developing countries. It often rears its head in crisis situations, where a clean water supply is absent. The disease is characterized by a watery diarrhea that induces death by dehydration unless adequately treated with anitbiotics.[2]

The causative agent of Cholera is a secreted protein toxin named cholera toxin (CT). CT consists of an A subunit, which is non-covalently attached to a ring of five copies of a second subunit, aptly named "subunit B" (CTB). The overall stoichiometry of the toxin is therefore AB_5 (Figure 1B).[3] The A subunit (CTA) contains the catalytic domain of the toxin. Once the complex has entered the gut endothelial cells, CTA disassociates from the pentameric ring of B subunits as two fragments; CTA1 and CTA2. CTA1 acts as an ADP-ribosyltransferase in the cytosol, leading to ADP-ribosylation of the alpha subunit of the heterotrimeric G-protein (Gsα). This modification leads to the overproduction of cAMP by activation of adenylate cyclase, which in turn leads to the phosphorylation of the chloride transporter cystic fibrosis transmembrane conductance regulator (CFTR). This activates the transport of chloride ions out of the cell, leading to the efflux of water due to the resulting change in osmotic pressure.[4]

CTB is responsible for cell entry of CT. It does so by binding carbohydrates on the epithelial cell surface, leading to internalization via an as of yet unknown mechanism. In the 70s, three different research groups found the ganglioside GM1 (Figure 1A) to be the main ligand for CTB.[5–7] This interaction has been shown to have a K_d of about 60 nM, as measured by surface plasmon resonance (SPR), which is unusually strong for a lectin-carbohydrate interaction.[8] The role of GM1 as the receptor for CT is not unambiguous: fucosylated oligosaccharides, such as Lewis X and Lewis Y[8] or the A, B and H blood group antigens[9], were recently found to also be ligands for CTB, binding to a separate, secondary binding site on the protein. Further evidence questioning the importance of GM1 came from a study by Wands *et al.*[10] who were unable to detect GM1 gangliosides on the human colonic epithelial cell line T84, despite this cell line being commonly used to study CTB internalization.[11,12] This was compounded by Yrlid and co-workers who showed that the CT mediated diarrheal response in GM1 deficient mice is no different from that of wildtype mice.[13] All these data support the hypothesis that non GM1-carbohydrates (such as the Lewis-type glycans) are the main ligands for CTB in the human intestine.

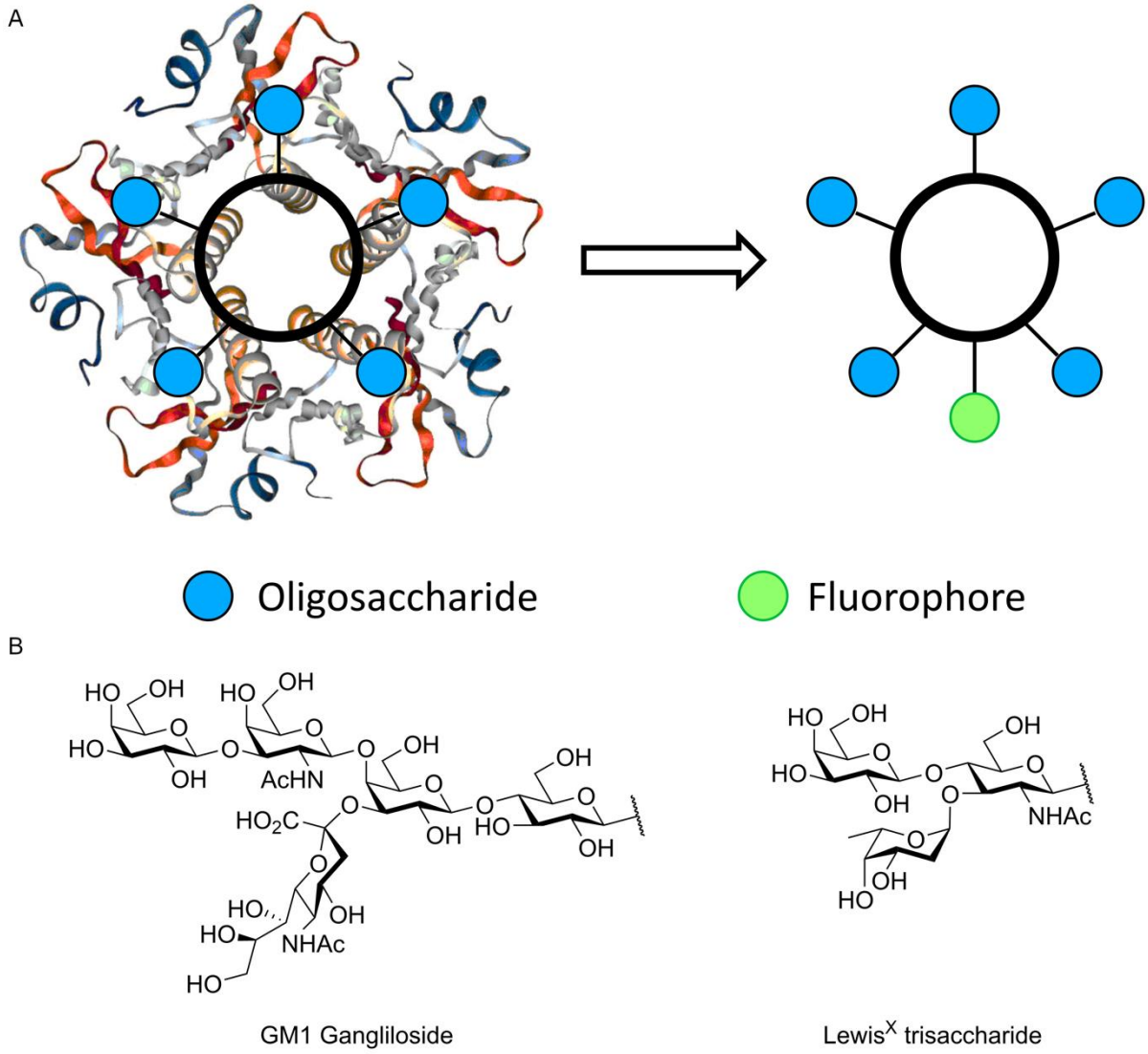


Figure 1. A) Cartoon displaying the general structure and biding mode of pentavalent CTB ligands and the proposed incorporation of a fluorophore for fluorescent detection (protein structure was adapted from PDB: 1CHQ). B) Structure of the GM1 ganglioside and the Lewis X trisaccharide.

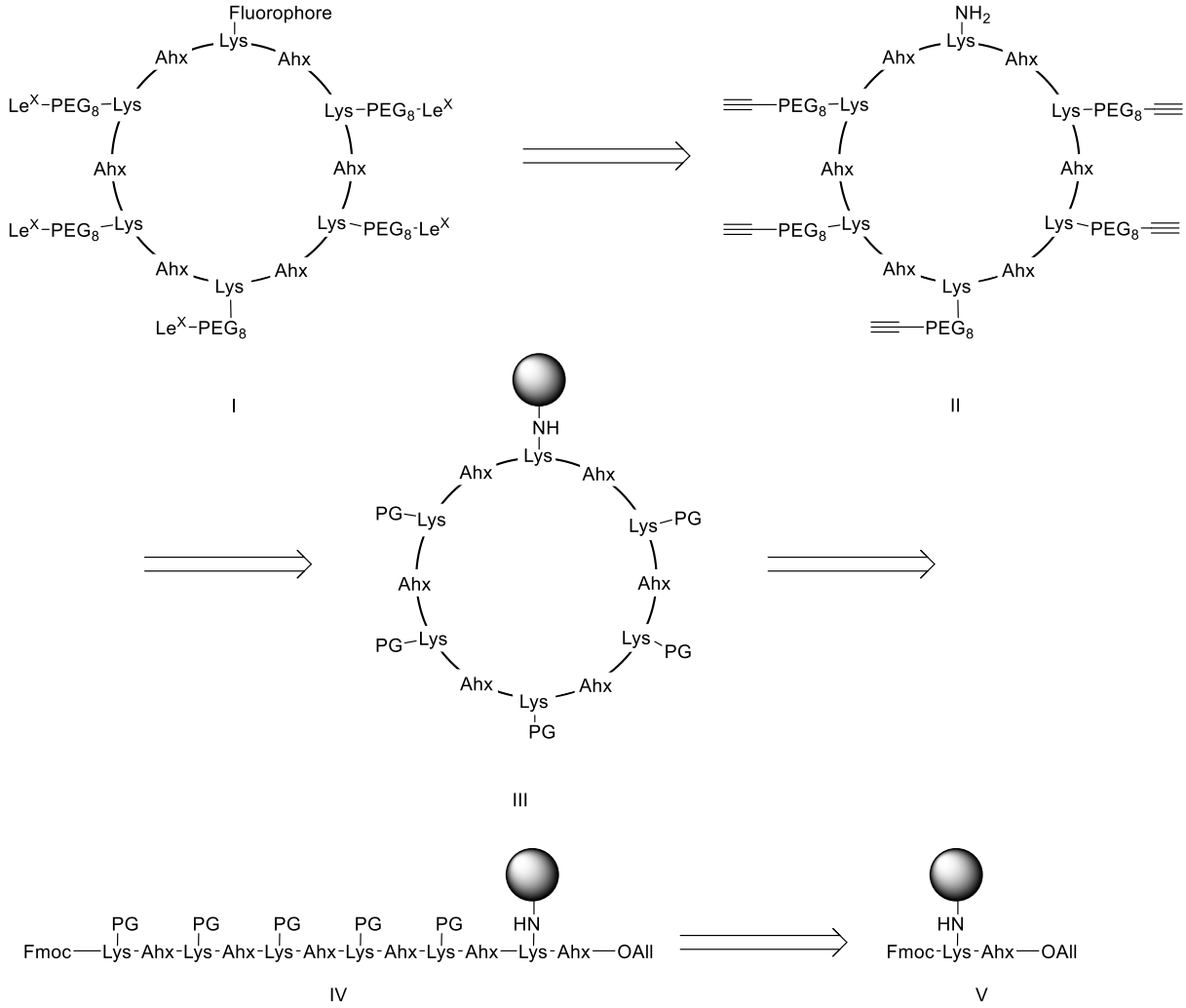
Since carbohydrate-mediated uptake of CT is a key process in the disease progress of cholera, several competitive inhibitors for the carbohydrate binding sites have been developed. While monovalent ligands are also known,[14] most approaches focus on exploiting the pentavalent nature of the protein to increase affinity. Multiple pentavalent scaffolds have been modified with carbohydrates and used as CTB ligands. One of the oldest designs is the pentacyclen based design pioneered by Fan *et al*.[15] Other designs based on calixarene structures[16] or scaffolds based on cyclic peptides have also been studied.[17] In all of these designs, long, polyethylene glycol based linkers are placed between the scaffold core and the carbohydrate ligands, to span the distance between the different binding sites. This is necessary since the diameter of the CTB ring is about 50 Å, with the distance between neighboring binding sites about 35 Å.[18] This makes the protein substantially larger than the diameter of most cyclic cores.

To date, no ligand has been synthesized that allows for the detection of CTB via fluorescent techniques. A ligand like that could be useful for the detection of cholera toxin in waste water[19] or the study of pathogen-host interactions by fluorescent microscopy. To this end, one of the known scaffold designs could be modified to allow for the incorporation of a fluorophore. Figure 1B shows a cartoon displaying this general idea. Of the earlier described designs, the cyclic peptide-based ligands appear the most promising for this application, given the vast amounts of known chemistry for the synthesis and derivatization of peptidic molecules. For the carbohydrate ligand, the Lewis^X trisaccharide was chosen over the more commonly used GM1 pentasaccharide, since recent studies have indicated this as the more relevant ligand for CT.

To allow efficient binding to the multivalent CTB lectin, a pre-organized multivalent carbohydrate ligand was envisaged. By synthesizing a cyclic peptide scaffold containing 5 carbohydrate attachment points, as well as an attachment point for a fluorophore, this pentavalency was deemed achievable (Figure 1B). Linkers of defined length between the trisaccharides and the cyclic peptide were also included in the design, as Fan and coworkers have shown spacer length to be critical for a strong binding interaction.[17] The cyclic peptide would consist of six lysine residues, five of which would be attached to the linker modified carbohydrates and one would serve as the fluorophore attachment point. In between the lysine residues spacer moieties will be added to give the cyclic peptide the correct circumference. By using a combination of amide bond forming reactions and copper catalyzed click chemistry, the core scaffold can be elaborated with different oligosaccharides or fluorophores in a modular fashion. This chapter describes the synthesis of such a peptidic scaffold as well as the synthesis of the full, pentavalent Lewis^X-based fluorescent CTB ligand (Scheme 1).

Results and discussion

The ideal fluorescent, peptide based, CTB ligand would have a cyclic core consisting of lysine residues used for functionalization, spaced out into an optimal ring size by spacer amino acids. The work of Fan and coworkers[17] on cyclic peptide based CTB ligands already showed aminohexanoic acid to be an effective spacer for these structures. From this same work it was deduced that a linker length of 8 ethylene glycol units between the peptide core and the Lewis^X trisaccharides should be optimal. The fluorophore itself can be attached directly to the scaffold, since it should not interact with the protein. All of this taken together gives target structure I, which is retrosynthetically analyzed in Scheme 1.



Scheme 1. Retrosynthetic analysis of the cyclic scaffold presenting pentavalent carbohydrate ligands and a fluorophore.

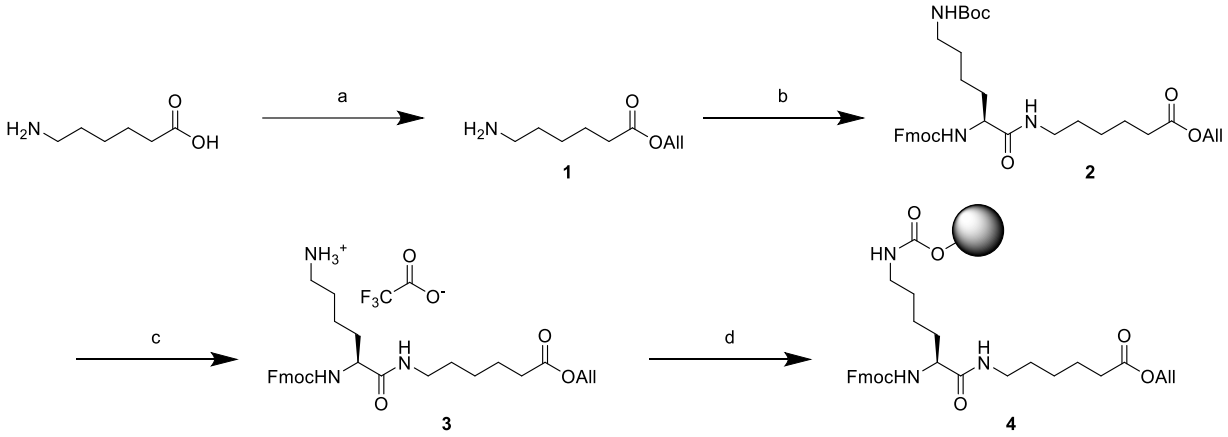
In order to synthesize the desired ligand, a cyclic scaffold containing five carbohydrate attachment points and a single fluorophore attachment site, three things need to be kept in mind: i) introduction of the carbohydrate should happen late-stage using high yielding chemistry, as the sugars are precious, ii) introduction of the fluorophore can most easily be accomplished in the form of an amide forming reaction using a fluorophore NHS-ester, as these reagents are generally commercially available and allow for efficient labeling, and iii) the attachment points used for both oligosaccharides and fluorophore should not interfere with cyclisation of the scaffold.

The first point, efficient late-stage attachment of carbohydrates, can be accomplished using copper(I) catalyzed click chemistry, as previously demonstrated in Chapter 4. A synthesis producing a protected azido-Lewis^X was previously described in Chapter 2, and deprotection of this advanced intermediate should yield a clickable Le^X for the conjugation reaction. To conjugate this oligosaccharide to the scaffold, an alkyne terminated polyethylene glycol molecule could be considered as the linker of choice, as these are often used as biocompatible linkers.[20] Incorporation onto the peptidic scaffold would be accomplished in a facile manner

by amide coupling a carboxylic acid terminated linker to a lysine sidechain. A sixth, unmodified lysine residue would then enable late-stage fluorophore labeling. (Scheme 1, intermediate II) This scaffold itself would be constructed from a properly protected cyclic intermediate (Scheme 1, intermediate III). This leaves the formation of the cyclic peptidic core as the last consideration. On-resin head-to-tail cyclization of a properly orthogonally protected peptide would enable the rapid synthesis of this scaffold using well established chemistry.

To allow for on-resin backbone cyclization, as well as selective coupling of linkers to five of the six lysine sidechains, a strategy of triple orthogonal protected amines is required. (Scheme 1, intermediate IV) To enable a head-to-tail cyclisation approach for the synthesis of the scaffold core, the peptide will be anchored to the solid support via the sidechain of one of the lysine residues, in the form of a carbamate linkage with a resin-bound *para*-hydroxybenzyl functionality, enabling release of the peptide under acidic treatment.[21] The C-terminus of the peptide will remain masked in the form of an allyl ester until the cyclisation reaction is carried out. Using the standard Fmoc protecting group on the N-terminus of the linear peptide immediately creates orthogonality with the acid labile resin anchoring, as well as the allyl ester, while also being synthetically the most straightforward, given the commercial availability of building blocks bearing this protecting group. The remaining lysine sidechains could be protected using highly acid labile trityl based protecting groups. While these types of acid sensitive protected lysine building blocks are commonly used for on-resin sidechain modifications,[22,23] the compatibility with the resin anchoring carbamate linkage needs to be investigated. The synthesis of the linear peptide will start from an immobilized dipeptide bearing a C-terminal allyl ester and an N-terminal Fmoc (Scheme 1, intermediate V).

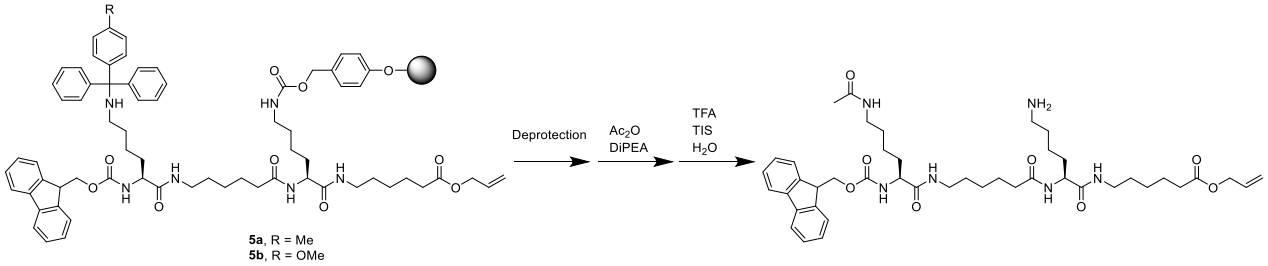
The synthesis of the cyclic peptide was initiated with the synthesis of fully protected dipeptide **2**, bearing a N-terminal Fmoc, a C-terminal allyl ester and a Boc-group protecting the lysine sidechain. This synthesis commenced with the allylation of 6-aminohexanoic acid using $SOCl_2$ as an activator to obtain amine **1** in 94% yield, followed by HCTU mediated coupling to Fmoc-Lys(Boc)-OH, giving the desired dipeptide **2** in 92% yield. To create the carbamate linkage between the resin and the lysine sidechain, the Boc protecting group was removed by treatment with a 1/1 mixture of TFA/DCM (v/v) for 15 minutes, followed by removal of the volatiles *in vacuo*. This afforded crude peptide **3**, which was dissolved in 1 M DiPEA in DCM, in order to neutralize the acid, and *para*-nitrophenolcarbonate-actived Tentagel S NPC resin was added. After stirring for 16 hours, the peptide was immobilized on the resin, affording intermediate **4**.



Scheme 2. Synthesis of fully protected dipeptide **3** followed by resin-anchoring via the lysine sidechain. Reagents and conditions: a) allyl alcohol, $SOCl_2$, 0°C, 94% b) Fmoc-Lys(Boc)-OH, HCTU, DiPEA, DMF, 92% c) TFA, DCM d) Tentagel S NPC, DiPEA, DCM.

To answer the question of protecting group orthogonality between the resin linker and the protecting groups used for the lysine sidechain, model immobilized peptide **5a** was first synthesized by extending **4** with Fmoc-Ahx-OH and Fmoc-Lys(Mtt)-OH. A series of deprotection conditions (Table 1, #1-4) were tested on this immobilized peptide. After deprotection, the resin was treated with Ac_2O and DiPEA followed by resin cleavage with a standard 95% TFA cocktail. By performing LC-MS analysis of the crude peptide, the amount of protecting group liberation could be determined by peak integration, by comparing peak intensity of the acetylated versus non-acetylated peptide. In order to determine whether resin cleavage took place during the trityl deprotection, the flowthrough of the reaction was concentrated *in vacuo,* redissolved in 1:1:1 H_2O/MeCN/tBuOH and subjected to LC-MS analysis as well.

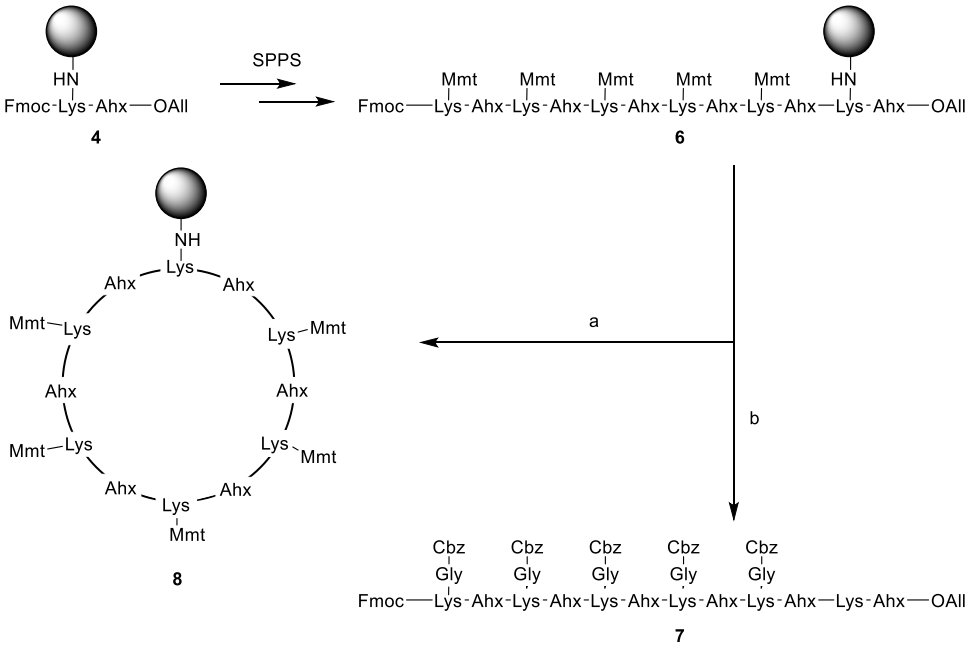
Table 1. Optimization of lysine sidechain deprotection in presence of the acid-labile carbamate resin linkage.



#	Reagent	R =	Time	Remark
1	1% TFA/DCM (v/v)	CH_3	30 min	10% Mtt deprotection
2	1% TFA/DCM (v/v)	CH_3	7 x 2 min	Peptide cleaved from resin
3	1% TFA/DCM + 1% TES (v/v)	CH_3	1 h	50% Mtt deprotection
4	1:2:7 AcOH/TFE/DCM (v/v)	CH_3	1 h	50% Mtt deprotection
5	1% TFA/DCM (v/v)	OCH_3	30 min	Peptide (partially) cleaved from resin
6	**1:2:7 AcOH/TFE/DCM (v/v)**	**OCH_3**	**1 h**	**Full Mmt deprotection**

The typical method for the selective liberation of Lys(Mtt) is treatment with 1% TFA in DCM (v/v) for 30 minutes (entry 1) or multiple very short incubations, repeated until the yellow color, indicative of liberated trityl carbocations, stops appearing (entry 2). While the single acidic treatment was mostly unsuccessful in liberating Mtt, the repeated incubations was able fully deprotect the protecting group, however simultaneously also fully cleaving the solid support linker. Since retritylation of the sidechain could play a role in the poor conversion found for entry 1, this reaction was reattempted with the addition of TES as a cation scavenger (entry 3). While this increased the level of conversion from 10% to 50%, these numbers are nevertheless insufficient for the efficient synthesis of the desired modified peptide. As a final alternative, the resin bound peptide was treated with a mixture of acetic acid and trifluoroethanol in DCM, as described by Aletras *et al.*[24], for one hour. However, this again yielded an unacceptable low conversion to the desired partially deprotected peptide.

Disappointingly, no condition for the full deprotection could be found that was able to fully liberate methyltrityl on lysine while simultaneously being fully orthogonal to the resin linker. Since most conditions led to partial liberation of the trityl without touching the resin linker, a more acid labile protecting group, in the form of 4-methoxytrityl, was incorporated instead, yielding resin bound peptide **5b**. This immobilized peptide was subjected to the same screening approach. When treating this immobilized peptide with 1% TFA in DCM (v/v) for half an hour (entry 5), some peptide liberation was observed, in contrast to what was observed earlier with resin bound peptide **5a** under the same conditions. Gratifyingly, when using the acetic acid/trifluoroethanol mixture (entry 6) complete amine liberation with full retention of support linkage was observed. With an effective orthogonal deprotection strategy in hand, the synthesis of the cyclic peptide itself was investigated next.

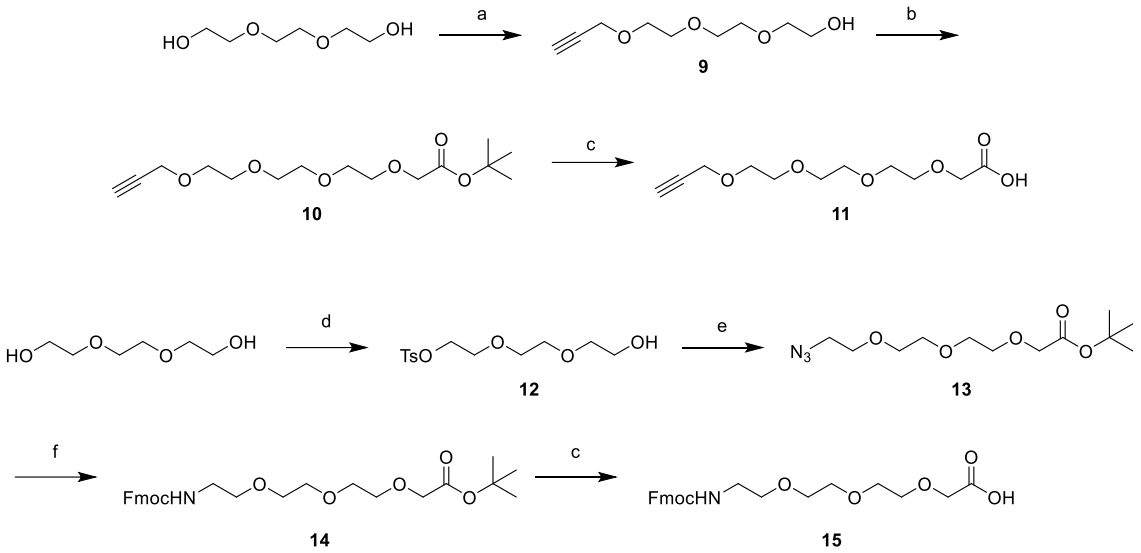


Scheme 3. Synthesis of fully protected cyclic scaffold **8** and exploration of deprotection and coupling of five Lys(Mmt) lysine residues. Reagents and conditions: a) i) $Pd(PPh_3)_4$, $PhSiH_3$, DCM ii) 2% DBU/DMF (v/v) iii) PyAOP, DiPEA, DMF b) i) AcOH, TFE, DCM ii) Cbz-Gly-OH, HCTU, DiPEA, DMF iii) TFA, TIS, DCM.

The immobilized linear peptide **6** was first synthesized using standard Fmoc based SPPS, using only two equivalents to couple the precious Fmoc-Lys(Mmt)-OH building block. Before cyclization could commence, first the C- and N-terminus of the peptide had to be liberated. The C-terminal allyl ester was cleaved using repeated treatment with $Pd(PPh_3)_4$ and phenylsilane. Complete removal of the protecting group was confirmed by a test cleavage of a small portion of resin (~1mg) followed by LC-MS analysis. Initially, the N-terminus was deprotected using 20% piperidine in DMF, as is typical in Fmoc based SPPS. However, this led to the formation of a C-terminal piperidine amide during the cyclization reaction, even when extensive washes were carried out between deprotection and cyclization. Presumably piperidine forms a salt bridge with the deprotected C-terminus, retaining the base on the resin. By deprotecting the N-terminus using 2% DBU in DMF (v/v) instead, this side reaction could be suppressed. Cyclization was attempted by treating the resin with a solution containing PyBOP as the condensation agent and DiPEA as general base for 16 hours, as was described by Fan and co-workers.[17] However, even after repeated 16 hour treatment, some non-cyclized material was still present. By using a more reactive condensation agent, PyAOP,[25] full consumption of linear peptide could be attained after a single 16 hour reaction.

With the fully cyclized peptide in hand, the deprotection and subsequent amide formation of the five lysine residues was explored. It was decided to optimize this reaction on the linear peptide **6**, since cyclisation side-products could conceivably make analysis of incompletely acylated mixtures more cumbersome. To this end, resin bound peptide **6** was first treated with acetic acid and trifluoroethanol in DCM for one hour, followed by DCM and DMF washes. Here, a cheap model carboxylic acid was needed to test the coupling reaction. Cbz-Gly-OH was used, as its lack of steric bulk on the C-α position combined and good UV absorption for LC-MS analysis makes this an ideal test reagent. Therefore, the partially deprotected immobilized peptide was reacted with a mixture of Cbz-Gly-OH, HCTU and DiPEA, which were first preactivated together for 5 minutes before addition to the resin. After 1 hour the resin was thoroughly washed with DMF and DCM and the bound peptide was liberated by treatment with a 95% TFA cocktail (95:2.5:2.5 TFA/H_2O/TIS), followed by LC-MS analysis. Here, besides the desired product **7**, peaks belonging to species where one, two or three lysine residues were acetylated, instead of the desired acylation with Cbz-glycine, were observed. This side reaction presumably happens because of residual acetic acid in the resin matrix, which is activated by HCTU. Some different post-deprotection wash protocols were evaluated, with a thorough washing with 10% TEA in DMF (v/v) for 2 x 15 minutes giving a sufficiently clean result.

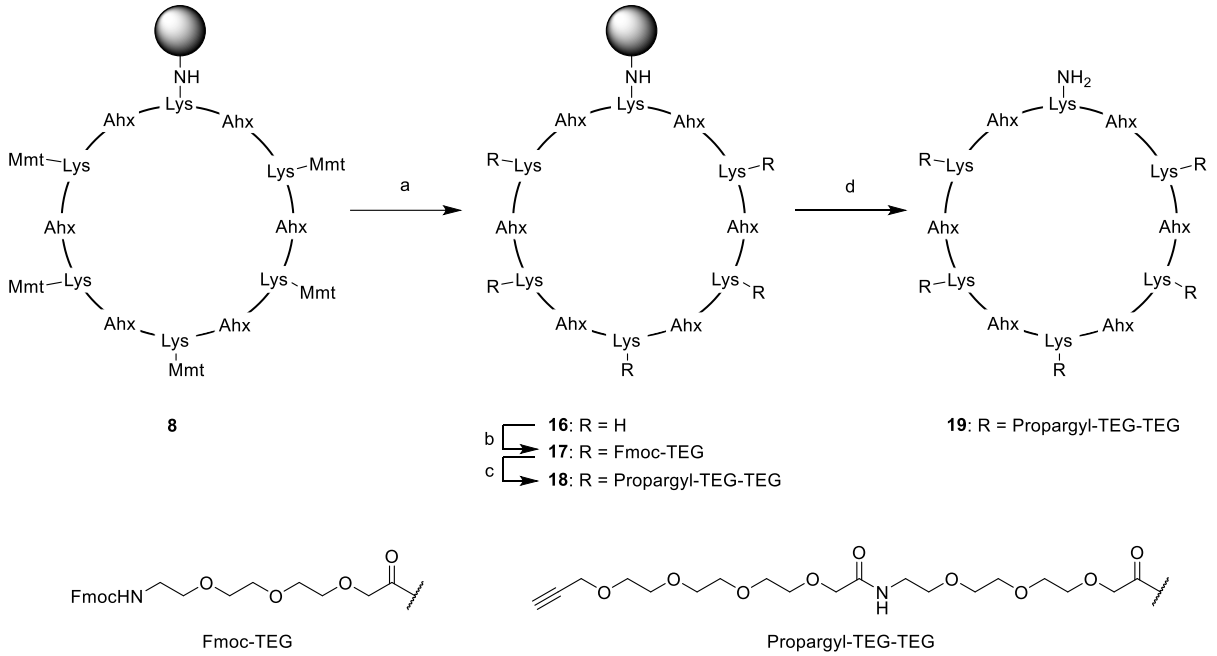
Next, the synthesis of the required long PEG linkers was evaluated. At first, the method of Zhang *et al.*[26], where long, functionalized and monodisperse PEG spacers are synthesized starting from a common building block, was attempted. However, this method turned out to be cumbersome and low yielding, so it was decided to make the required long linkers using on-resin amide bond formation from two shorter linkers **11** and **15**, the synthesis of which is shown in Scheme 4.



Scheme 4. Synthesis of PEG linkers **11** and **15** from triethylene glycol. Reagents and conditions: a) propargyl bromide, NaH, THF, 89% b) *tert*-butyl bromoacetate, NaH, THF, 64% c) TFA, DCM, **11** quant, **15** quant d) TsCl, Et_3N, DCM, 86% e) i) NaN_3, DMF, 70°C ii) *tert*-butyl bromoacetate, NaH, THF, 50% over two steps f) i) PPh_3, H_2O, THF ii) Fmoc-OSu, N-methylmorpholine, DCM, 90%.

Briefly, the propargyl terminated TEG linker **11** was synthesized by selective monoalkylation of triethylene glycol with propargyl bromide, mediated by NaH, obtaining the alkyne **9** in 89% yield, followed by alkylation of the second alcohol with *tert*-butyl bromoacetate, again mediated by NaH, in 64% yield, producing **10**. This protected acid was then treated with TFA to give **11** in quantitative yield. Similarly, the synthesis of Fmoc protected TEG-linker **15** was commenced by monotosylation of triethylene glycol, performed with *para*-toluenesulfonyl chloride in presence of triethylamine, to produce **12** in 86% yield. Next, substitution with sodium azide in DMF at elevated temperature, followed by alkylation with *tert*-butylbromoacetate and NaH in THF, produced intermediate **13** in 50% yield. The azide was converted to an amine by Staudinger reduction with triphenylphosphine and water and immediately protected with an Fmoc group using Fmoc-OSu and *N*-methylmorpholine, yielding fully protected linker **14** in 90% yield. After deprotection of the *tert*-butyl ester with TFA, **15** was obtained in quantitative yield.

In order to obtain scaffold **19**, fully protected cyclic peptide **8** was first treated with the Mmt cleavage cocktail, giving resin bound peptide **16**. After extensive washing, TEG linker **15** was coupled using HCTU as the condensation reagent in presence of DiPEA, producing resin-bound peptide **17**. This was followed by Fmoc deprotection and coupling of linker **11** under the same conditions. This resin bound molecule (**18**) was then treated with a 95% TFA cocktail for one hour, liberating the resin bound scaffold. The obtained crude mixture was purified over RP-HPLC, but the desired molecule was obtained in extremely poor yield (< 1%).



Scheme 5. Synthesis of cyclic scaffold **19**. Reagents and conditions: a) AcOH, TFE, DCM b) **15**, HCTU, DiPEA c) **11**, HCTU, DiPEA d) TFA, TIS, H_2O.

Careful investigation of LC-MS spectra obtained when test-cleaving immobilized peptides **17** and **18** revealed incomplete linker incorporation, especially in the second acylation step (**17** → **18**). By changing the coupling reagent used in these reactions from HCTU to PyAOP, an increased conversion to the desired molecule could be achieved. When the molecule was resynthesized using this altered method, a new complication was observed. When analyzing the mass spectrum (Figure 2A) of the crude molecule, the two main ions observed were found to have m/z = 1768.25 and m/z = 1415.00. While the first mass is readily assigned as the z=2 ion of the expected product **19** (exact mass = 3533.07 Da), the second ion would belong to an unlikely z=2.5 ionization. A more likely explanation would be the formation of the cyclic dimer, where first two linear peptides couple to form a linear dimer, followed by cyclization. This would imply the observed masses are actually the z = 4 and z = 5 ions of the dimer molecule (exact mass = 7066.15 Da). This kind of side reaction has been observed more often when performing on-resin cyclisation reactions.[27] In order to better understand the amount of cyclic monomer versus cyclic dimer that is being formed, an improved way of analyzing cyclic intermediate **8** was envisioned. Since a cyclic polylysine derivative as has been synthesized here has very little retention on reverse phase silica, a method to increase the retention time and separation between monomer and dimer peaks on LC-MS was used, as outlined in Figure 2B. Capitalizing on the ability to selectively liberate the Mmt groups on the lysine sidechains, a hydrophobic moiety could be introduced, increasing retention on reverse phase silica. Benzoyl groups where chosen as the moiety of choice both for their ease of introduction using benzoic anhydride in DCM as well as their added benefit of having good UV absorption properties, aiding the analysis of the mixture by LC-UV.

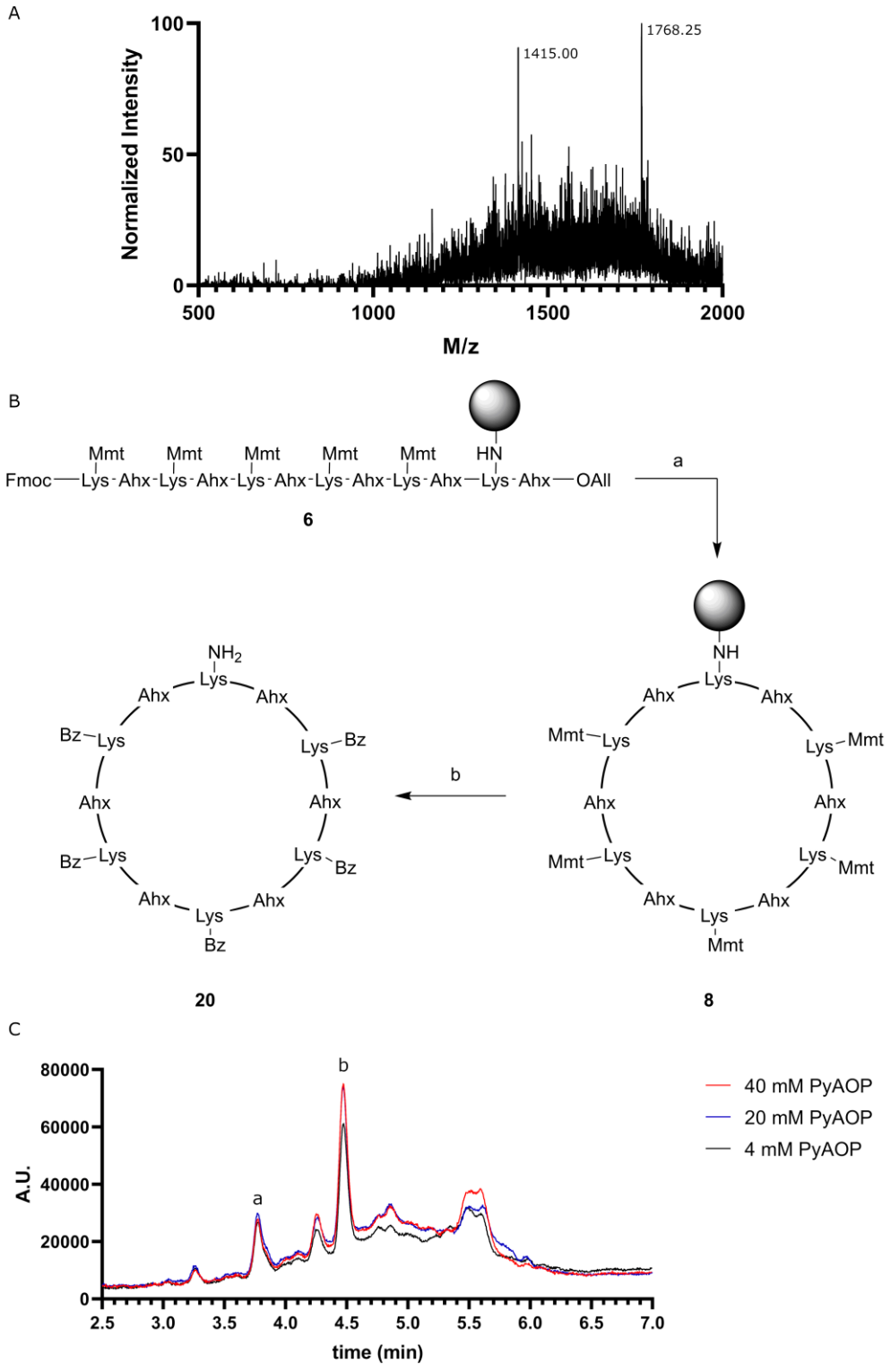
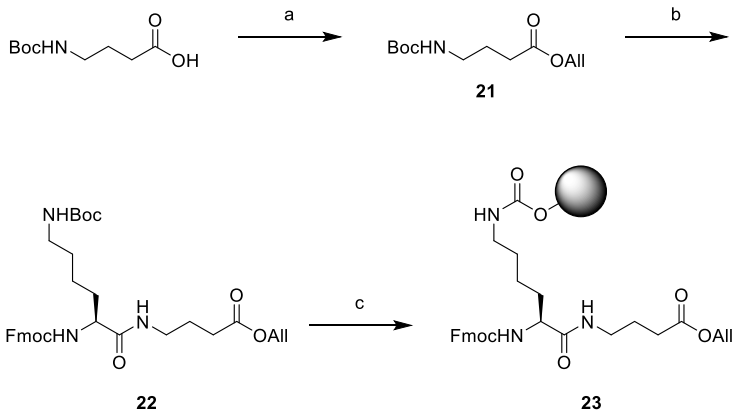


Figure 2. A) Partial mass spectrum obtained when analyzing crude **19**, the two largest peaks are annotated. B) Outline of the methodology used to analyze cyclization products. C) Partial LC-UV chromotagram of the benzoylated derivatives Reagents and conditions: a) i) $Pd(PPh_3)_4$, $PhSiH_3$, DCM ii) DBU, DMF iii) PyAOP, DiPEA, DMF b) i) AcOH, TFE, DCM ii) Bz_2O, DiPEA, DCM iii) TFA, H_2O, TIS C) partial LC-UV analysis (gradient: 30-70% B over 10 minutes) of cyclization reactions carried out using different concentrations of PyAOP. Peak (a) was assigned as the desired monocyclized product and (b) as the undesired dimeric cyclic product.

To evaluate the effect the concentration of the condensation agent PyAOP has on the cyclization, three test cyclization reactions were carried out. The Mmt groups of the immobilized peptides were then removed, followed by benzoylation and resin cleavage. The crude cleavage products were then analyzed by LC-MS, and the traces are shown in Figure 2C. As can be seen in this chromatogram, the majority of observed product was in the form of the cyclic dimer, regardless of concentration of PyAOP. A similar result was obtained when the cyclization was carried out in DMSO instead of DMF. These results suggest that the prevalence to form dimers is inherent to the structure of linear peptide **6**.

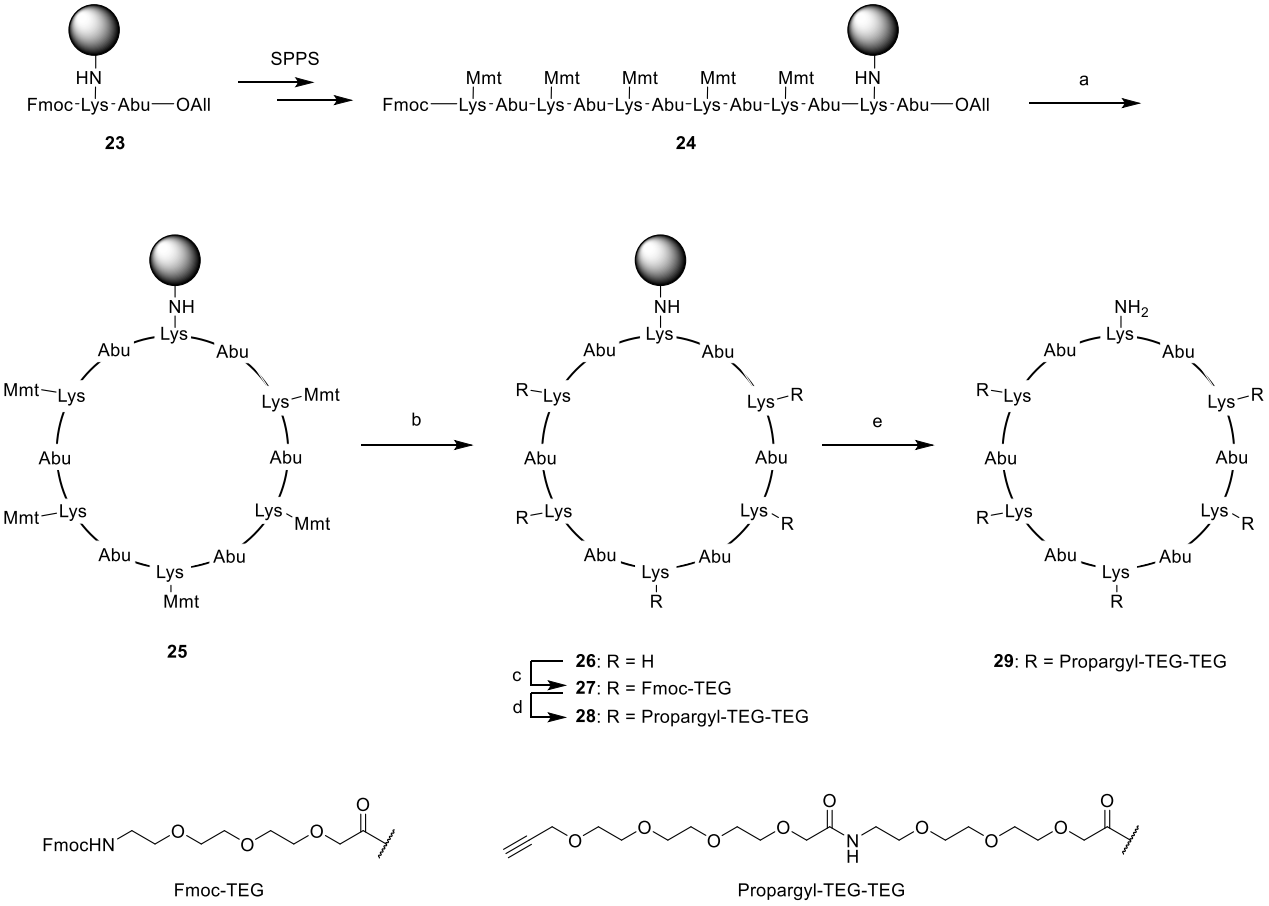
One variable that might be key to the formation of the undesired peptide cyclic peptide is the length of the linear peptide used in the cyclization reaction. Replacing aminohexanoic acid with a shorter spacer, for instance the four carbon long aminobutyric acid, might decrease the peptide's propensity to form cyclic dimers over the desired cyclic monomer. In order to test this, a new immobilized dipeptide **23** was synthesized (Scheme 6). The synthesis started with the formation of an allylester on Boc-aminobutynoic acid, mediated by EDC, in presence of DiPEA and DMAP as nucleophilic catalyst, producing **21** in 75% yield. This was followed by Boc deprotection in 1/1 (v/v) TFA/DCM. After removing the volatiles *in vacuo* the crude amine was immediately coupled to Fmoc-Lys(Boc)-OH, affording dipeptide **22** in 90% yield. This dipeptide was then, after facile removal of the Boc group protecting the lysine side chain, loaded onto Tentagel S NPC resin to afford immobilized intermediate **23**.



Scheme 6. Synthesis of immobilized dipeptide **23**. Reagents and conditions: a) Allyl alcohol, EDC, DMAP, DiPEA, DCM, 16h, 75% b) i) TFA, DCM, 15 min ii) Fmoc-Lys(Boc)-OH, HCTU, DiPEA, DMF, 2h, 90% over two steps c) i) TFA, DCM, 15 min ii) Tentagel S NPC, DiPEA, DCM, 16 h.

With immobilized dipeptide **23** in hand, synthesis of the full scaffold proceeded smoothly, as presented in Scheme 7. The full linear peptide **24** was synthesized as before, and deallylation using $Pd(PPh_3)_4$ and $PhSiH_3$ proceeded smoothly. After removal of the Fmoc using 2% DBU in DMF cyclisation was attempted using PyAOP. After overnight treatment, a small portion of the resin was again subjected to selective Mmt removal followed by benzoylation with benzoic anhydride. After resin liberation with the 95% TFA cocktail, LC-MS analysis showed the desired monocyclic molecule as the main product. The rest of immobilized peptide **25** was then

subjected to the same deprotection condition to liberate the Mmt groups. After vigorous washing to remove remaining acetic acid, Fmoc protected TEG linker **15** was coupled using PyAOP as the condensation agent in presence of DiPEA. The Fmoc group was removed using the conventional method (20% piperidine in DMF (v/v)) and the propargyl terminated TEG **11** was coupled to the linkers, again mediated by PyAOP/DiPEA. The peptide was released from the resin using the standard 95% TFA cocktail and purified by preparative RP-HPLC, yielding **29** in 4.6% yield.

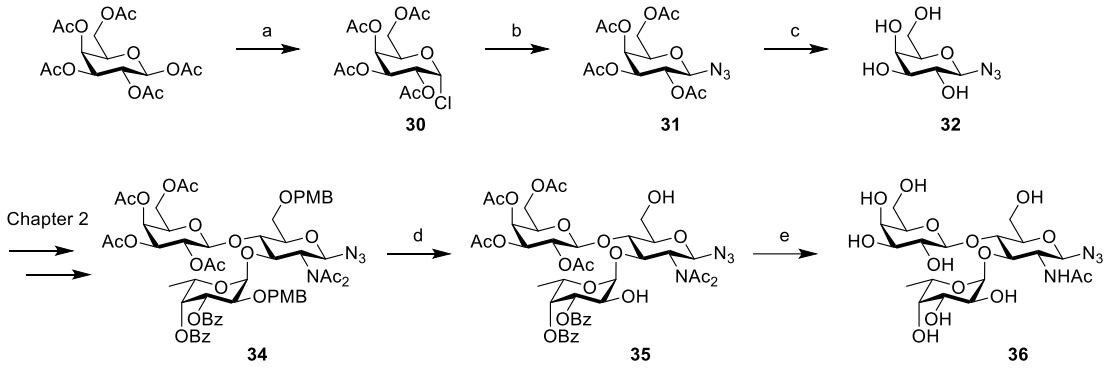


Scheme 7. Synthesis of cyclic Abu containing peptide **29**. Reagents and conditions: a) i) $Pd(PPh_3)_4$, $PhSiH_3$, DCM ii) DBU, DMF iii) PyAOP, DiPEA, DMF b) AcOH, TFE, DCM c) **15**, PyAOP, DiPEA, DMF d) i) piperidine, DMF ii) **11**, PyAOP, DiPEA, DMF e) TFA, TIS, H_2O.

With the peptidic scaffold now in hand, the next step was the synthesis of the clickable sugars. Besides Lewis^X azide, an 1-azido derivative of galactose was also synthesized to explore binding differences caused by interaction with the distinct pockets, since galactose has been shown to interact with the GM1 binding pocket of CTB.[28]

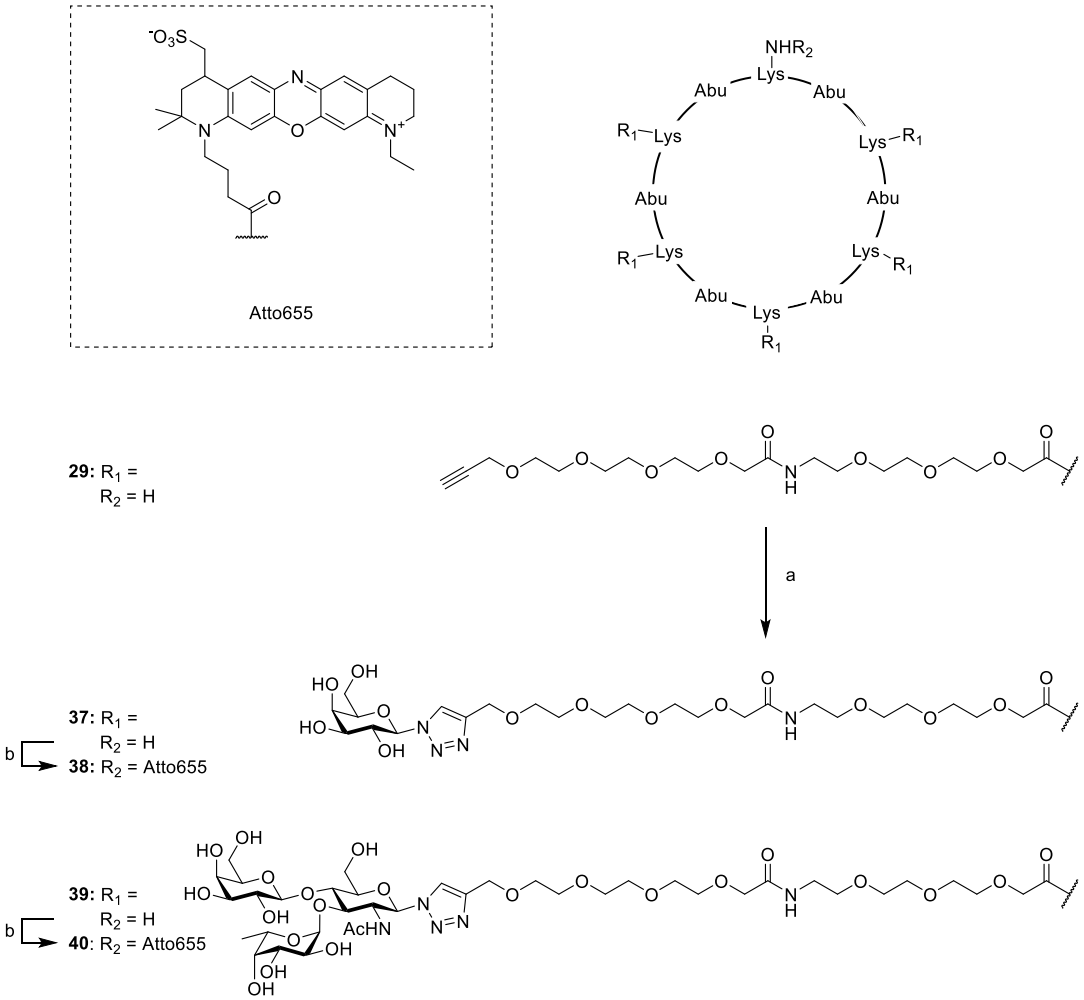
The synthesis of clickable galactoside **32** started from commercially available 1,2,3,4,6-penta-O-acetyl-β-galactopyranoside by introduction of an anomeric chloride using $TiCl_4$ in refluxing DCM, producing α-chloride **30** in 93% yield. A simple S_N2-like substitution with NaN_3 produced the protected azidosugar **31** in 83% yield, which was then deacetylated under Zemplén conditions to produce compound **32** in 92% yield. The synthesis of the azido Lewis X trisaccharide started off with fully protected sugar **34**, the synthesis of which was described

in Chapter 2. The *para*-methoxybenzyl protecting groups were removed oxidatively using DDQ, yielding the partially deprotected compound **35** in 71% yield. To remove the remaining ester protecting groups, Zemplén conditions were again employed, producing compound **36** in quantitative yield.



Scheme 8. Synthesis of galactoside **32** and trisaccharide **36**. Reagents and conditions: a) $TiCl_4$, DCM, 40°C, 1h, 93% b) NaN_3, DMF, 50°C, 3h, 82% c) NaOMe, MeOH, 1.5h, 92% d) DDQ, H_2O, DCM, 71% e) NaOMe, MeOH, 24h, quant.

An attempt was then made to couple these glycosides to the pentameric scaffolds. Scaffold **29** was dissolved in DMSO (50 mM), and glycosyl azides were added as 0.2 M stocks (2 eq per alkyne for **32** and 1.2 eq per alkyne for **36**). A copper click mix was prepared (1:3:2 $CuSO_4$/THPTA/sodium ascorbate) and 15 mol% Cu(I) was added to the reaction mixture. The reactions were gently agitated at 40°C for 48h, after which a small sample was taken out for LC-MS analysis, confirming full conversion to the pentaglycosylated products. Size exclusion purification of **37** and **39**, on HW40 resin using an ammonium acetate buffer, was attempted. While the products were obtained in reasonable purity according to LC-MS analysis, removal of remaining salts proved troublesome. Repeated cycles of dissolution in aqueous acetonitrile, followed by lyophilization were carried out, to no avail. An attempt was made to exchange the acetate ions for hydroxy ions using an anion exchange resin, but this resulted into partial degradation of the material. Instead, an alternative synthetic route was employed, where purification after the click reaction was omitted. To this end, 100 nmol of scaffold **29** was again reacted with sugar **32** in the presence of Cu(I) to give **37**. When LC-MS indicated product formation, DiPEA was added, followed by Atto655-NHS, to label the free lysine. Initial attempts to carry out the reaction at room temperature yielded no conversion, so the reaction was heated to 40°C for 16 hours, creating the desired labeled glycoconjugates. These were again isolated using HW40 size exclusion chromatography. This time, an ammonium bicarbonate buffer was used, since this salt is easier removed by lyophilization and the increase pH (compared to ammonium acetate) is of little concern. Using this method, pentavalent glycoconjugate **38** was isolated in 87% yield and the more complex conjugate **40** was isolated in 42%.



Scheme 9. Synthesis of the final glycoconjugates **38** and **40**. Reagents and conditions: a) **32** or **36**, $CuSO_4$, THPTA, sodium ascorbate, 40°C, 48h b) Atto655-NHS, DiPEA, 40°C, 16h, **38** (87% over 2 steps), **40** (42% over 2 steps).

Conclusion

Since the interactions between carbohydrate binding proteins and their ligands tend to be weak, multivalency is often exploited to increase affinity.[29] As technologies are developed that are able to effectively visualize these kinds of interactions, multivalent scaffolds containing fluorophores for microscopy analysis become more relevant.[30] The combination of the key role the multivalent lectin CTB plays in cholera infection and the unmet medical needs concerning this disease makes CTB an interesting target for further development of multivalent carbohydrate probes. While many different multivalent CTB binding scaffolds have been developed over the years as inhibitors for the cholera toxin itself, modifying one of these to allow for introduction of a fluorescent detector group is non-trivial. Cyclic peptide-based inhibitors seemed like a logical target, as robustness and flexibility of modern solid support peptide synthesis helps the rapid production of these target molecules.

However, the synthesis and late stage functionalization of cyclic peptides is still not without its challenges. As seen in this chapter, cyclization reactions can give unexpected side products, like the aforementioned cyclic dimer. Here, it seems that ring size is still a concern, even while

working under the pseudo-dilution conditions of on-resin cyclization. Fortunately, decreasing the length of the linear peptide, and therefore the circumference of the cyclic molecule, enabled smooth cyclization. Based on previous literature the change from a 6 carbon spacer to a 4 carbon spacer is unlikely to negatively impact the multivalent binding of the molecule to CTB.[17]

Another point of attention is protecting group orthogonality, which in many cases is not fully explored. In this case, conditions where the trityl-based protective groups on the lysine sidechains could be fully removed without touching the resin linker had to be discovered. Here, the extremely acid-sensitive monomethoxytrityl group afforded the desired selectivity and deprotection efficiency. This allowed for the direct incorporation of the polyethylene glycol-based linkers using standard SPPS chemistry, avoiding handling the polyamine intermediate or performing complicated in-solution chemistry after resin liberation. Instead, the completed scaffold molecule was purified by RP-HPLC and derivatized using CuAAC chemistry, following by amine labeling with a commercially available fluorophore NHS ester. This way, cumbersome HPLC purification was avoided for the precious final compound, which could instead be purified over size exclusion chromatography.

In conclusion, this chapter describes the synthesis of a fluorescently labeled pentavalent ligand for CTB bearing either five galactose or five Lewis^X ligands. In the future, this ligand will be applied for the visualization of cholera toxin in cholera infected zebrafish. Furthermore, variants of this scaffold presenting different numbers of specific carbohydrates will be evaluated as ligands for different lectins, for instance DC-SIGN[31] or the mannose-6-phophate[32] receptor.

Experimental

General methods for synthesis and characterization of compounds

Solvents were purchased from Honeywell, VWR or Alfa Aesar. Anhydrous solvents were prepared by drying over 4Å molecular sieves. Reagents purchased from chemical suppliers were used without further purification, unless stated otherwise. All reactions were performed under nitrogen atmosphere and/or under exclusion of H_2O, unless stated otherwise. Reactions were followed by thin layer chromatography which was performed using TLC silica gel 60 F254 on aluminium sheets, supplied by Merck. Compounds were visualized using UV absorption (254 nm) and/or a spray reagent, either permanganate (5 g/L $KMnO_4$, 25 g/L K_2CO_3) or sulfuric acid (10% v/v in EtOH). ^1H and ^{13}C NMR spectra were recorded using a Brüker AV400 (400 /101 MHz) and COSY and HSQC 2D experiments were used to assign peaks. Recorded data was interpreted and analyzed using MestReNova 12 software. Chemical shifts are reported in ppm (δ) in reference to an internal standard (TMS) or the residual solvent peak. High resolution mass spectra were recorded by direct injection (2 µL of a 2 µM solution in H_2O/MeCN 1:1 and 0.1% formic acid) on a mass spectrometer (Thermo Fisher Exactive HF Orbitrap) equipped with an electrospray ion source in positive mode. The high resolution mass spectrometer was calibrated prior to measurements with a calibration mixture (Thermo Finnigan). IR spectra were recorded using a Shimadzu IRSpirit fourier transform infrared spectrometer.

SPPS methods

Loading of Tentalgel NPC resin

Protected dipeptide **2** or **22** (500 µmol) were weighed out into a 50 mL round bottom flask. DCM (2 mL) and TFA (2 mL) were added and the resulting suspension was stirred for 15 minutes, during which the solid fully dissolved. Next, the volatiles were removed *in vacuo* and residual TFA was removed by toluene co-evaporation (3x2 mL). The residue was taken up in 1 M DiPEA in anhydrous DCM (2 mL) and Tentagel S NPC (400 mg, 0.26 mmol/g loading) was added. Another 3 mL of anhydrous DCM were added and the resulting suspension was stirred under an N_2 atmosphere overnight. The next day, the suspension was transferred to a fritted syringe and washed extensively with DCM (3x10 mL) and DMF (3x10 mL).

Manuel synthesis of linear peptides

For the synthesis of linear precursor of the cyclic peptides, manual SPPS protocols were employed on a 100 µmol scale. Liberation of the Fmoc protecting group was accomplished using 20% (v/v) piperidine in DMF (2x 10 mL, 3+7 min), followed by extensive washing with DMF (5x10 mL). The sequence was extended alternatively with either a spacer amino acid (Fmoc-Ahx-OH or Fmoc-Abu-OH) or Fmoc-Lys(Mmt)-OH. For the spacer amino acids, 4 eq to resin loading were used, together with 4 eq HCTU (as 0.4 M solution in DMF) and 8 eq DiPEA and the coupling reaction was carried out for 45 minutes. Fmoc-Lys(Mmt)-OH was coupled using 2 eq of amino acid to resin loading, together with 2 eq HCTU (as a 0.2 M in DMF solution) together with 4 eq of DiPEA, and the reaction was run for 1.5 h. After each coupling, the resin was washed thoroughly with DMF (4x10 mL) before the next round of Fmoc deprotection was started. After synthesis of the full-length linear peptide, the quality of the immobilized peptide was checked by treating a small (~1mg) portion of resin using 200 µL 95% TFA cocktail (95:2.5:2.5 TFA/TIS/H_2O) for 1 hour. After this time the TFA was drained into ~800 µL of ice

cold Et_2O and the formed precipitate was collected by centrifugation. The supernatant was discarded and the pellet redissolved in 200 µL 1:1:1 H_2O/ACN/tBuOH and subjected to LC-MS analysis.

General method for on resin cyclization

This general protocol was used on a scale of 50 µmol of resin bound linear peptide. The C-terminal allyl ester protecting group was removed using $Pd(PPh_3)_4$ and $PhSiH_3$ as scavenger. The resin was first extensively washed with DCM (5x 10 mL). DCM was degassed by bubbling N_2 through ~10mL of DCM for ~1.5 minutes. $Pd(PPh_3)_4$ (11.6 mg, 10 µmol) was weighed in an Eppendorf reaction vessel and dissolved in 1 mL of degassed DCM. $PhSiH_3$ (61 µL, 500 µmol) was added and the resulting solution was added to the drained resin. The resin was then reacted at room temperature under light agitation for 15 minutes. The solution was drained from the resin and the resin was washed once with DCM (5 mL). This procedure was repeated 3 more times to fully liberate the C-terminus of the peptide. Following allyl ester cleavage, the resin was thoroughly washed with DCM (5x 5 mL). Before continuing, the complete removal of the allyl ester was confirmed by the cleavage of a small portion of resin (1 mg) in 95% TFA cocktail followed by LC-MS analysis. The resin was then washed with DMF (3x 5 mL), followed by Fmoc cleavage by 2% DBU (v/v) in DMF (3x5 mL, 5 min each) and extensive DMF washes (3x 5 mL). The resin was then washed 2 times with anhydrous DMF (2x 5 mL) and transferred to a flask, were it was suspended in 20 mL of anhydrous DMF and placed under an N_2 atmosphere. PyAOP (250 µmol, 52 mg, 5 eq) and DiPEA (500 µmol, 35 µL, 10 eq) were added, and the suspension was stirred overnight. The next day, the resin was transferred back to the fritted syringe, drained and washed extensively with DMF (3x5 mL) and DCM (3x5 mL). The drained resin was stored at -20°C.

General method for cyclization analysis

A small portion of resin bearing immobilized cyclic peptide (~1 mg) was treated with 500 µL of a mixture of 1:2:7 AcOH/TFE/DCM for 1 hour followed by extensive washing with DCM (3x2 mL). The deprotected immobilized peptide was then treated with a solution of benzoic anhydride (50 mg) and DiPEA (25 µL) in DCM (450 mL) for 15 minutes. After extensive washing (5x2 mL DCM) 200 µL of a 95% TFA cocktail was added and the resulting suspension was stirred by light agitation for 1 h. The TFA was drained into 800 µL of icecold Et_2O and the formed suspension was collected by centrifugation. The supernatant was discarded and the pellet dissolved in DMSO (200µL) and subjected to LC-MS analysis.

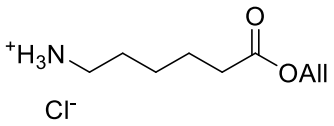
Modification of cyclic peptides with TEG linkers

This general protocol was used on a scale of 50 µmol of resin bound cyclic peptide. The resin was suspended in DCM for 20 minutes to swell the resin before synthesis In order to remove the Mmt protection groups, the resin was suspended in 2 mL of 1:2:7 AcOH/TFE/DCM for 1 hour. This was followed by extensive washing with DCM (3x5 mL) and DMF (3x 5 mL). The resin was then treated with a 10% (v/v) solution of Et_3N in DMF (2x 5 mL, 15 min each) followed by though washing with DMF (5x 5 mL). To a solution of TEG linker **15** (500µmol, 10 eq) in DMF (0.5 M) was added PyAOP (260 mg, 500 µmol, 10 eq) and DiPEA (174 µL, 1 mmol, 20 eq) and the solution was allowed to preactivate for 5 minutes before being added to the drained resin. The suspension was reacted at room temperature under gentle agitation overnight, and the resin was drained and extensively washed with DMF (4x 5 mL). The Fmoc groups were removed as stated before (20% piperidine (v/v) in DMF, 2x 5 mL, 3+7 min) followed by extensive washing with DMF (5x 5 mL). Propargylated TEG **11** (123 mg, 500µmol, 10 eq) and PyAOP (260 mg, 500 µmol, 10 eq) were dissolved in DMF (1 mL) and DiPEA (174 µL, 1 mmol, 20

eq) was added. The suspension was reacted at room temperature under gentle agitation overnight, and the resin was drained and extensively washed with DMF (3x 5 mL) and DCM (3x 5 mL). The resin was then treated with 2 mL of a 95% TFA cocktail (95:2.5:2.5 TFA:TIS:H_2O) for one hour. The TFA solution was filtered through the syringe frit into ice cold Et_2O (10 mL) and the resulting precipitate was collected by centrifugation. The crude scaffold was then subjected to RP-HPLC to obtain the pure compound.

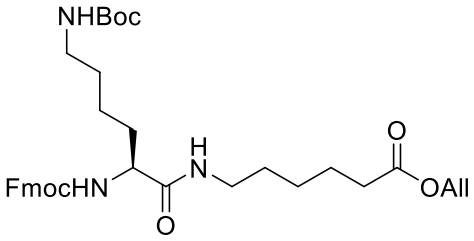
Solution phase synthesis

6-Aminohexanoic acid allyl ester hydrochloride (1)



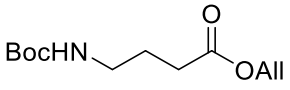
To an icebath cooled flask containing allyl alcohol (10 mL) thionyl chloride (3 mL, 40 mmol, 10 eq) was added dropwise. 6-aminohexanoic acid (524 mg, 4 mmol) was carefully added to the mixture and the reaction was allowed to warm up to room temperature overnight. The volatiles were removed *in vacuo* and the remaining solid was washed with Et_2O to obtain allyl ester **1** in 94% yield (775 mg, 3.74 mmol). This compound was used in the next step without further purification. **^{1}H NMR** (400 MHz, DMSO) δ 8.06 (s, 3H), 5.91 (ddt, J = 17.2, 10.7, 5.5 Hz, 1H), 5.28 (dq, J = 17.2, 1.6 Hz, 1H), 5.20 (dq, J = 10.5, 1.4 Hz, 1H), 4.53 (dt, J = 5.4, 1.4 Hz, 2H), 2.72 (t, J = 7.5 Hz, 2H), 2.33 (t, J = 7.4 Hz, 2H), 1.60 – 1.48 (m, 4H), 1.36 – 1.25 (m, 2H) **^{13}C NMR** (101 MHz, DMSO) δ 172.5, 132.8, 117.8, 64.3, 38.5, 33.2, 26.6, 25.3, 23.9

Fmoc-Lys(Boc)-Ahx-OAll (2)



Fmoc-Lys(Boc)-OH (281 mg, 0.6 mmol, 1.2 eq) was dissolved in dry DMF (2 mL) and HCTU (248 mg, 0.6 mmol, 1.2 eq) and DiPEA (260 µL, 1.5 mmol, 3 eq) were added. The mixture was stirred for 5 minutes before hydrochloride salt **1** (103 mg, 0.5 mmol) was added. The reaction was stirred for 2 hours, diluted with EtOAc (50 mL) and washed successively with 1 M $HCl_{(aq)}$, saturated aqueous $NaHCO_3$ and brine. The organic layer was dried over $MgSO_4$, filtered and concentrated. Silica gel column chromatography (1/1 EtOAc/pentane) yielded dipeptide **2** in 92% yield (287 mg, 0.46 mmol) **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.73 (d, J = 7.5 Hz, 2H, Fmoc-Ar), 7.55 (d, J = 7.1 Hz, 2H, Fmoc-Ar), 7.36 (t, J = 7.2 Hz, 2H, Fmoc-Ar), 7.30 – 7.23 (m, 2H, Fmoc-Ar), 6.71 (s, 1H, NH), 6.01 – 5.79 (m, 2H, NH, -C<u>H</u>=CH_2), 5.28 (d, J = 17.2 Hz, 1H, -CH=C<u>H</u>H), 5.21 (d, J = 10.4 Hz, 1H, -CH=CH<u>H</u>), 4.86 (s, 1H, NH), 4.54 (d, J = 5.6 Hz, 2H, Allyl-CH_2), 4.41 – 4.24 (m, 2H, Fmoc-CH_2), 4.24 – 4.12 (m, 2H, α-Lys), 3.31 – 2.95 (m, 4H, ε-Ahx, ε-Lys), 2.27 (t, J = 7.4 Hz, 2H, α-Ahx), 1.91 – 1.16 (m, 21H, β-Lys, β-Ahx, γ-Lys, γ-Ahx, δ-Lys, δ-Ahx, tBu). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 173.2 (C=O), 171.8 (C=O), 156.4 (C=O), 156.2 (C=O), 143.7 (Fmoc-C_q), 141.2 (Fmoc-C_q), 132.2 (-<u>C</u>H=CH_2), 127.7 (Fmoc-Ar), 127.0 (Fmoc-Ar), 125.0 (Fmoc-Ar), 119.9 (Fmoc-Ar), 118.2 (-CH=<u>C</u>H_2), 79.0 (tBu-C_q), 67.0 (Fmoc-CH_2), 65.0 (Allyl-CH_2), 54.8 (α-Lys), 47.0 (Fmoc-CH), 40.0, 39.2 (ε-Lys, ε-Ahx), 33.9 (α-Ahx), 32.3 (β-Lys), 29.6, 29.0, 28.4 (tBu-CH_3), 26.2, 24.3, 22.5 (β-Lys, β-Ahx, γ-Lys, γ-Ahx, δ-Lys, δ-Ahx). **HRMS** (ESI) m/z $[M + H]^+$ calcd for $C_{35}H_{47}N_3O_7H$ 622.3487 found 622.3487

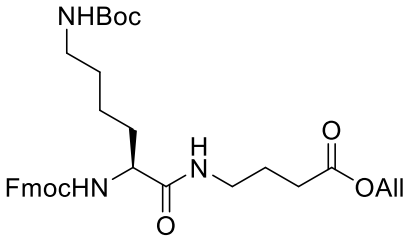
4-(*tert*-Butoxycarbonylamino)butyric acid allyl ester (21)



4-(*tert*-Butoxycarbonylamino)butyric acid (1.22g, 6.0 mmol) was dissolved in DCM and the solution was cooled in an ice bath. Allyl alcohol (2 mL, 30

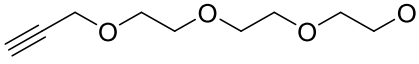
mmol, 5 eq), EDC·HCl (1.38 g, 7.2 mmol, 1.2 eq), DiPEA (5.2 mL, 30 mmol, 5 eq) and DMAP (67 mg, 0.6 mmol, 0.1 eq) were added successively. The reaction was allowed to warm to room temperature overnight after which TLC analysis (40% EtOAc/pentane + 0.1% AcOH) indicated starting material consumption. The reaction mixture was diluted with DCM and washed with 1 M $HCl_{(aq)}$ and sat. aq. $NaHCO_3$, dried over $MgSO_4$ and filtered. The dried organics were concentrated to give allyl ester **21** in 75% yield (1.09 g, 4.5 mmol). This compound was used without further purification. **^{1}H NMR** (300 MHz, $CDCl_3$) δ 5.91 (ddt, *J* = 17.1, 10.4, 5.7 Hz, 1H, -C<u>H</u>=CH_2), 5.31 (dq, J = 17.2, 1.6 Hz, 1H, -CH=C<u>H</u>H), 5.26 – 5.15 (m, 2H, -CH=CH<u>H</u>, NH), 4.58 (dt, J = 5.7, 1.3 Hz, 2H, Allyl-CH_2), 3.16 (q, J = 6.5 Hz, 2H, α-Abu), 2.39 (t, J = 7.4 Hz, 2H, γ-Abu), 1.83 (p, J = 7.1 Hz, 2H, β-Abu), 1.43 (s, 9H, tBu) **^{13}C NMR** (75 MHz, $CDCl_3$) δ 172.6 (C(O)OAll), 155.8 (C(O)OtBu), 131.9 (-<u>C</u>H=CH_2), 117.9 (-CH=<u>C</u>H_2), 78.6, (tBu-C_q), 64.8 (Allyl-CH_2), 39.6 (γ-Abu), 31.2 (α-Abu), 28.1 (tBu-CH_3), 25.0 (β-Abu)

Fmoc-Lys(Boc)-Abu-OAll (22)



Protected amino acid **21** (1.09 g, 4.5 mmol) was dissolved in DCM (5 mL) and TFA (5 mL) was added. The resulting mixture was stirred for 15 minutes before the volatiles were removed *in vacuo*. The crude amine was co-evaporated thrice with toluene (3 x 5 mL) to remove traces of TFA and kept under N_2. In a separate flask, Fmoc-Lys(Boc)-OH (2.53 g, 5.4 mmol, 1.2) and HCTU (2.24 g, 5.4 mmol, 1.2 eq) were dissolved in DMF (18 mL), followed by the addition of DiPEA (3.9 mL, 22.5 mmol, 5 eq). The resulting solution was stirred for 5 minutes before being added to the flask containing the crude amine. The reaction was stirred overnight at room temperature and TLC analysis (1/1 EtOAc/pentane) indicated the formation of a product. The reaction mixture was diluted with EtOAc and washed successively with 1 M $HCl_{(aq)}$, saturated aqueous $NaHCO_3$ and brine. The organic layer was dried over $MgSO_4$, filtered and concentrated. Silica gel column chromatography (1/1 EtOAc/pentane) produced dipeptide **22** in 90% yield (2.56 g, 4.3 mmol). **^{1}H NMR** (400 MHz, DMSO) δ 7.96 – 7.88 (m, 3H, Abu-N$^{\gamma}$H, Fmoc-Ar), 7.74 (dd, J = 7.4, 3.6 Hz, 2H, Fmoc-Ar), 7.49 – 7.39 (m, 3H, Lys-N$^{\alpha}$H), 7.34 (t, J = 7.4 Hz, 2H, Fmoc-Ar), 6.79 (t, J = 5.6 Hz, 1H, Lys-N$^{\varepsilon}$H), 5.90 (ddt, J = 17.2, 10.7, 5.4 Hz, 1H, -C<u>H</u>=CH_2), 5.28 (dq, J = 17.2, 1.6 Hz, 1H, -CH=C<u>H</u>H), 5.19 (dq, J = 10.5, 1.4 Hz, 1H, CH=CH<u>H</u>), 4.53 (dt, J = 5.4, 1.4 Hz, 2H, Allyl-CH_2), 4.30 – 4.18 (m, 3H, Fmoc-CH_2, Fmoc-CH), 3.90 (td, *J* = 8.5, 5.4 Hz, 1H, α-Lys), 3.15 – 3.02 (m, 2H, γ-Abu), 2.97 – 2.83 (m, 2H, ε-Lys), 2.35 (t, J = 7.5 Hz, 2H, α-Abu), 1.66 (p, J = 7.1 Hz, 2H, β-Abu), 1.61 – 1.46 (m, 2H, β-Lys), 1.43 – 1.18 (m, 13H, tBu, γ-Lys, δ-Lys) **^{13}C NMR** (101 MHz, DMSO) δ 172.8 (C=O), 172.4 (C=O), 156.4 (C=O), 156.0 (C=O), 144.4 (Fmoc-C_q), 144.3 (Fmoc-C_q), 141.2 (Fmoc-C_q), 133.2 (-<u>C</u>H=CH_2), 128.1 (Fmoc-Ar), 127.5 (Fmoc-Ar), 125.8 (Fmoc-Ar), 120.6 (Fmoc-Ar), 118.1 (-CH=<u>C</u>H_2), 77.8 (tBu-C_q), 66.0 (Fmoc-CH_2), 64.8 (Allyl-CH_2), 55.2 (α-Lys), 47.1 (Fmoc-CH), 38.2 (γ-Abu), 32.1 (β-Lys), 31.2 (α-Abu), 29.7 (δ-Lys), 28.7 (tBu-CH_3), 24.9 (β-Abu), 23.4 (γ-Lys) **HRMS** (ESI) m/z [M + H]$^+$ calcd for $C_{33}H_{43}N_3O_7H$ 594.3174 found 594.3174

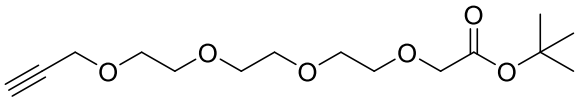
2-(2-(2-(prop-2-yn-1-yloxy)ethoxy)ethoxy)ethan-1-ol (9)



Triethylene glycol (3.0 g, 20 mmol, 2 eq) was dissolved in anhydrous THF (50 mL) and cooled to 0°C in and ice bath. A dispersion of NaH (60% w/w) in mineral oil (0.52 g, 13 mmol, 1.3 eq) was added and the reaction was stirred at 0°C for 30 minutes, followed by the addition of propargyl bromide (1.1 mL, 10 mmol). The reaction was stirred for another 2 hours at 0°C followed by 15 hours at room temperature, at which

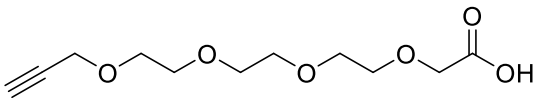
point TLC analysis (1/4 acetone/DCM) indicated complete consumption of starting material. The reaction was quenched by the addition of H_2O (2 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo*. Silica gel column chromatography (5% → 20% acetone/DCM) yielded the desired alcohol **9** as a yellow oil in 89% yield (1.66 g, 8.9 mmol). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 4.20 (d, J = 2.4 Hz, 2H, CH2 propargyl), 3.76 – 3.64 (m, 10H), 3.62 – 3.57 (m, 2H), 3.25 (s, 1H, OH), 2.51 (t, J = 2.4 Hz, 1H, CH propargyl) **^{13}C NMR** (101 MHz, $CDCl_3$) δ 79.4 (C propargyl), 74.7 (CH propargyl), 72.4, 70.4, 70.1, 70.1, 68.8, 61.34 58.2 (CH2 propargyl)

tert-butyl 3,6,9,12-tetraoxapentadec-14-ynoate (10)



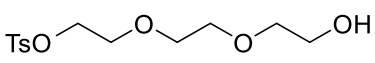
Alcohol **9** (5.37 g, 28.5 mmol) was dissolved in anhydrous THF (140 mL) and cooled to 0°C. A dispersion of NaH (60% w/w) in mineral oil (1.6 g, 40 mmol, 1.4 eq) was added and the suspension was stirred for 30 minutes at 0°C. *tert*-butylbromoacetate (6.4 mL, 43 mmol 1.5 eq) was added to the reaction mixture and the reaction was stirred for an additional 90 minutes at 0°C, followed by the removal of the icebath. The reaction was stirred at room temperature for 20 hours, after which TLC analysis (10% acetone/DCM) indicated consumption of starting material. The reaction mixture was poured in water (700 mL) and the aqueous layer extracted with DCM. The organic layer was dried over $MgSO_4$, filtered and concentrated. Silica gel column chromatography (1% → 10% acetone/DCM) yielded the desired ester **10** as a yellow oil in 64% yield (5.48g, 18.1 mmol). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 4.21 (d, J = 2.4 Hz, 2H, CH2 propargyl), 4.03 (s, 2H, $OCH_2C(O)O$), 3.75 – 3.64 (m, 12H), 2.45 (t, J = 2.4 Hz, 1H, CH propargyl), 1.48 (s, 9H, $C(CH_3)_3$) **^{13}C NMR** (101 MHz, $CDCl_3$) δ 169.6 (C=O), 81.5 ($\underline{C}(CH_3)_3$), 79.6 (C propargyl), 74.6 (CH propargyl), 70.9, 70.7, 70.6, 70.4, 69.1, 69.0 ($O\underline{C}H_2C(O)O$), 58.4 (CH_2 propargyl), 28.1 ($C(\underline{C}H_3)_3$)

3,6,9,12-tetraoxapentadec-14-ynoic acid (11)



tert-butyl ester **10** (1.16g, 3.8 mmol) was dissolved in DCM (7.6 mL) and TFA (2.9 mL, 38 mmol, 10 eq) was added. The reaction mixture was stirred at room temperature for 1 hour, after which TLC analysis (10% acetone/DCM) indicated full consumption of the starting material. The volatiles were removed *in vacuo* and the resulting residue was dissolved in ddH_2O and lyophilized to obtain **11** as a yellow oil in quantitative yield (918 mg, 3.72 mmol). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 9.31 (s, 1H, COOH), 4.21 (d, J = 2.4 Hz, 2H, CH_2 propargyl), 4.19 (s, 2H, $OCH_2C(O)O$), 3.79 – 3.74 (m, J = 5.8, 2.7 Hz, 2H), 3.74 – 3.66 (m, 10H), 2.47 (t, J = 2.3 Hz, 1H, CH propargyl). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 173.4 (C=O), 79.5 (C propargyl), 74.8 (CH propargyl), 71.2, 70.9, 70.6, 70.4, 70.3, 70.3, 69.1, 68.6 ($O\underline{C}H_2C(O)O$), 58.4 (CH_2 propargyl) **ν_{max}/cm^{-1}** 3255 (Alkyne), 1742 (C=O) **HRMS** (ESI) m/z: $[M + Na]^+$ calcd for $C_{11}H_{18}O_6Na$ 269.0996, found 269.0993

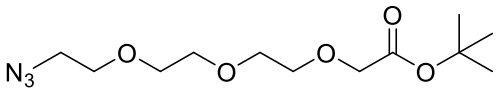
2-(2-(2-hydroxyethoxy)ethoxy)ethyl 4-methylbenzenesulfonate (12)



Triethylene glycol (14 mL, 100 mmol, 10 eq) and Et_3N (2.1 mL, 15 mmol, 1.5 eq) were dissolved together in anhydrous DCM (14 mL). The mixture was cooled to 0°C in an icebath and tosyl chloride (1.9 g, 10 mmol) was added portionwise. The reaction was stirred for 72 hours and TLC analysis (10% acetone/DCM) indicated full consumption of tosyl chloride. The reaction mixture was then diluted with DCM (100 mL), washed with water (2x 50

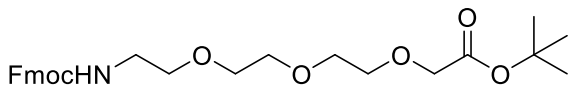
mL) and the combined aqueous layers were back-extracted with DCM (50 mL). The combined organic layers were then washed with 1 M $HCl_{(aq)}$, dried over $MgSO_4$, filtered and concentrated under reduced pressure, giving the desired monosubstituted ethylene glycol **12** in 86% yield (2.60 g, 8.55 mmol). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.79 (d, J = 8.2 Hz, 2H, Ts-Ar), 7.35 (d, J = 8.2 Hz, 2H, Ts-Ar), 4.20 – 4.14 (m, 2H, CH_2OTs), 3.75 – 3.67 (m, 4H, PEG-CH_2), 3.60 (s, 4H, PEG-CH_2), 3.58 – 3.54 (m, 2H, PEG-CH_2), 2.77 (s, 1H, OH), 2.44 (s, 3H, Ts-CH_3). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 144.9, 132.7, 129.8, 127.9 (Ts-Ar), 72.4, 70.6, 70.5, 70.2, 69.2, 68.5, 61.5 (PEG-CH_2), 21.6 (Ts-CH_3)

tert-butyl 2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)acetate (13)



Tosylate **12** (2.60 g, 8.5 mmol) was dissolved in anhydrous DMF (8.5 mL) and NaN_3 (1.1 g, 17 mmol, 2 eq) was added. The reaction was heated to 70°C and stirred at this temperature for 36 hours, after which TLC analysis indicated full consumption of starting material. The reaction was allowed to cool to room temperature, diluted with H_2O (100 mL) and extracted four times with DCM (4x 50 mL). The combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude azide was then dissolved in anhydrous THF (8 mL), together with tert-butyl bromoacetate (3.2 mL, 21 mmol, 2.5 eq). The reaction mixture was then cooled to 0°C in an icebath and a 60% (w/w) dispersion of NaH in mineral oil was added (840 mg, 21 mmol, 2.5 eq) portionwise. After addition was complete, the reaction was allowed to warm to room temperature over night and was quenched by the addition of saturated aqueous $NaHCO_3$ (50 mL). The aqueous layer was extracted twice with DCM (2x 50 mL) and the combined aqueous layers were dried over $MgSO_4$, filtered and concentrated under reduced pressure. Silica gel column chromatography (20% → 40% EtOAc/pentane) gave the desired ethylene glycol derivative **13** in 50% yield (1.23 g, 4.27 mmol). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 4.03 (s, 2H, $OCH_2C(O)O$), 3.76 – 3.65 (m, 10H), 3.39 (t, J = 5.0 Hz, 2H, CH_2N_3), 1.48 (s, 9H, $C(CH_3)_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 169.6 (C=O), 81.5 ($\underline{C}(CH_3)_3$), 70.7, 70.7, 70.6, 70.6, 70.0, 69.0 (PEG-CH_2), 50.7 (CH_2N_3), 28.1 ($C(\underline{C}H_3)_3$). **ν_{max}/cm^{-1}** 2111 (N_3), 1747 (C=O)

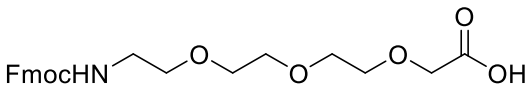
tert-butyl 1-(9H-fluoren-9-yl)-3-oxo-2,7,10,13-tetraoxa-4-azapentadecan-15-oate (14)



Azide **13** (3.76 g, 13 mmol) was dissolved in anhydrous THF (85 mL) and cooled to 0°C in an ice bath. PPh_3 (4.46 g, 17 mmol, 1.3 eq) was added and the reaction was stirred at room temperature for 24 h. When TLC analysis (10% acetone/DCM) indicated full consumption of starting material, H_2O (610 µL 34 mmol, 2.6 eq) was added and the reaction was stirred for an additional 24 h. When TLC analysis (10% MeOH/DCM + 0.1% Et_3N) indicated full consumption of the iminophosphorane intermediate, the reaction mixture was diluted with H_2O (150 mL). The aqueous layer was washed three times with toluene each toluene layers was separately extracted with water. The combined aqueous layers were concentrated to give the crude amine as a yellow oil. This oil was dissolved in DCM (100 mL) and Fmoc-OSu (5.26 g, 15.6 mmol, 1.2 eq) and N-methylmorpholine (1.4 mL, 13 mmol, 1 eq) were added. The reaction mixture was stirred at room temperature for 1 h, when TLC analysis (10% MeOH/DCM indicate full consumption of the amine. The reaction mixture was then washed once with H_2O and twice with 1 M $HCl_{(aq)}$. Each aqueous layer was back-extracted with DCM and the combined organic layers were dried over $MgSO_4$, filtered and concentrated under reduced pressure. Silica gel column chromatography (40% → 80% EtOAc/pentane) gave the desired protected ethylene glycol **14** in 90% yield (5.7 g, 11.7 mmol). **^{1}H NMR** (400 MHz,

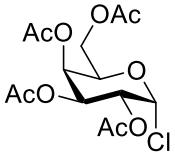
$CDCl_3$) δ 7.75 (d, J = 7.5 Hz, 2H, Fmoc-Ar), 7.60 (d, J = 7.4 Hz, 2H, Fmoc-Ar), 7.38 (t, J = 7.4 Hz, 2H, Fmoc-Ar), 7.30 (td, J = 7.4, 1.0 Hz, 2H, Fmoc-Ar), 5.52 (s, 1H, NH), 4.39 (d, J = 7.0 Hz, 2H, Fmoc-CH_2), 4.21 (t, J = 6.9 Hz, 1H, Fmoc-CH), 3.99 (s, 2H, $OCH_2C(O)O$), 3.73 – 3.59 (m, 8H, PEG-CH_2), 3.56 (t, J = 4.9 Hz, 2H), 3.39 (dd, J = 10.4, 5.2 Hz, 2H, CH_2NH), 1.45 (s, 9H, $C(CH_3)_3$) **^{13}C NMR** (101 MHz, CDCl3) δ 169.6 (C=O), 156.5 (C=O), 144.00, 141.27, 127.63, 127.03, 125.10, 119.93 (Fmoc-Ar), 81.6 ($\underline{C}(CH_3)_3$), 77.5, 70.7, 70.5, 70.3, 70.0 (PEG-CH_2), 68.9 ($O\underline{C}H_2C(O)O$), 66.5 (Fmoc-CH_2), 47.2 (Fmoc-CH), 40.9 (CH_2NH), 28.1 ($C(\underline{C}H_3)_3$)

1-(9H-fluoren-9-yl)-3-oxo-2,7,10,13-tetraoxa-4-azapentadecan-15-oic acid (15)



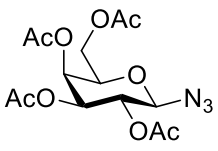
tert-butyl ester **14** (1.46 g, 3 mmol) was dissolved in TFA (15 mL) and the reaction was stirred at room temperature for 2 h, after which TLC analysis (60% EtOAc/pentane) indicated full consumption of starting material. The volatiles were removed *in vacuo* to give building block **15** in quantitative yield (1.47g, 3 mmol). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 9.05 (s, 1H, C(O)OH), 7.76 (d, J = 7.5 Hz, 2H, Fmoc-Ar), 7.60 (d, J = 7.4 Hz, 2H, Fmoc-Ar), 7.39 (t, J = 7.4 Hz, 2H, Fmoc-Ar), 7.31 (t, J = 7.4 Hz, 2H, Fmoc-Ar), 5.57 (s, 1H, NH), 4.53 – 4.34 (m, J = 21.1, 10.1 Hz, 2H, Fmoc-CH_2), 4.22 (t, J = 7.0 Hz, 1H, Fmoc-CH), 4.18 – 4.10 (m, 2H, $OC\underline{H}_2C(O)OH$), 3.78 – 3.51 (m, 10H), 3.40 (s, 2H, $C\underline{H}_2NH$). **^{13}C NMR** (101 MHz, CDCl3) δ 173.1 (C=O), 156.9 (C=O), 144.0, 141.3, 127.7, 127.1, 125.2, 120.0 (Fmoc-Ar), 71.1, 70.5, 70.2, 70.1, 70.0 (PEG-CH_2), 68.5 ($O\underline{C}H_2C(O)OH$), 66.8 (Fmoc-CH_2), 47.2 (Fmoc-CH), 40.9 (CH_2NH). **HRMS** (ESI) m/z $[M + Na]^+$ calcd for $C_{23}H_{27}NO_7Na$ 452.1680 found 452.1678

2,3,4,6-tetra-O-acetyl-α-D-galactopyranosyl chloride (30)



1,2,3,4,6-penta-O-acetyl-β-D-galactopyranoside (3.9 g, 10 mmol) was dissolved in anhydrous DCM (38 mL) and a 1M $TiCl_4$ solution in DCM (12 mL, 12 mmol, 1.2 eq) was added. The reaction was heated to reflux for 1 hour, after which TLC analysis (1/1 EtOAc/pentane) indicated full conversion of the starting material to a higher running spot. The reaction was allowed to cool to room temperature, diluted with DCM and washed with saturated aqueous $NaHCO_3$ and brine. The organic layer was dried over $MgSO_4$, filtered and concentrated. The crude was loaded onto a silica in DCM and eluted in 1/1 EtOAc/pentane to remove residual titanium salts. After removal of the solvent galatosyl chloride **30** was obtained in 93% yield (3.29g, 9.25 mmol, 93%) and used in the next step without further purification. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 6.40 (d, J = 3.8 Hz, 1H, H1), 5.52 (d, J = 2.0 Hz, 1H, H4), 5.40 (dd, J = 10.7, 3.1 Hz, 1H, H3), 5.25 (dd, J = 10.7, 3.9 Hz, 1H, H2), 4.55 (t, J = 6.4 Hz, 1H, H5), 4.19 (dd, J = 11.4, 6.1 Hz, 1H, H6a), 4.11 (dd, J = 11.4, 6.9 Hz, 1H, H6b), 2.16 (s, 3H, Ac), 2.11 (s, 3H, Ac), 2.06 (s, 3H, Ac), 2.01 (s, 3H, Ac) **^{13}C NMR** (101 MHz, $CDCl_3$) δ 169.9 (C=O), 169.7 (C=O), 169.6 (C=O), 169.3 (C=O), 91.1 (C1), 69.2 (C5), 67.5 (C2), 66.9 (C4), 66.7 (C3), 60.8 (C6), 20.3 (Ac), 20.2 (Ac), 20.2 (Ac)

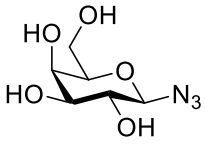
2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl azide (31)



Glycosyl chloride **30** (3.29 g, 9.3 mmol) was dissolved in anhydrous DMF (19 mL) and NaN_3 (1.21 g, 18.6 mmol, 2 eq) was added. The reaction mixture was heated to 50°C for three hours, after which TLC analysis (1/9 EtOAc/DCM) indicated full consumption of starting material. The reaction was allowed to cool to room temperature. The reaction mixture was then diluted with saturated aqueous $NaHCO_3$ (200 mL) and extracted with Et_2O (200 mL). The organic layer was washed with brine (200 mL), dried over $MgSO_4$ and concentrated. Silica gel column chromatography (40% → 50% Et_2O/pentane) yielded azido sugar

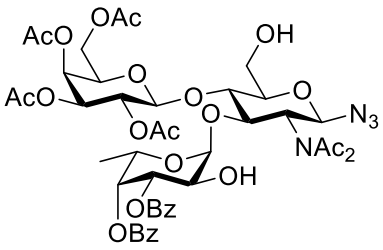
31 in 82% yield (2.82 g, 7.58 mmol). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 5.43 (d, J = 3.2 Hz, 1H, H4), 5.20 – 5.03 (m, 2H, H2, H3), 4.68 (d, J = 8.5 Hz, 1H, H1), 4.18 (d, J = 6.7 Hz, 2H, H6), 4.14 – 4.07 (m, 1H, H5), 2.18 (s, 3H, Ac), 2.10 (s, 3H, Ac), 2.06 (s, 3H, Ac), 1.99 (s, 3H, Ac) **^{13}C NMR** (101 MHz, $CDCl_3$) δ 170.0 (C=O), 169.9 (C=O), 169.6 (C=O), 169.1 (C=O), 87.9 (C1), 72.5 (C5), 70.4 (C3), 67.8 (C2), 66.7 (C4), 61.1 (C6), 20.3 (Ac), 20.3 (Ac), 20.2 (Ac)

β-D-galactopyranosyl azide (32)



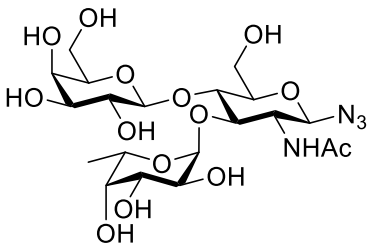
A fresh solution of sodium methoxide was prepared by dissolving a small chuck of sodium metal in methanol (10 mL). The pH of the solution was checked to be 11 and protected glycoside **31** was added. The reaction was stirred at room temperature for 1.5 hours, after which TLC analysis (4/1 EtOAc/pentane) indicated full consumption of starting material. The reaction mixture was neutralized with Amberlite H^+ resin, filtered and concentrated. The obtained material was lyophilized to obtain monosaccharide **32** in 92% yield (189 mg, 0.92 mmol). **^{1}H NMR** (400 MHz, D_2O) δ 4.63 (d, J = 8.7 Hz, 1H, H1), 3.92 (d, J = 3.4 Hz, 1H, H4), 3.80 – 3.67 (m, 3H, H6, H5), 3.64 (dd, J = 9.9, 3.4 Hz, 1H, H3), 3.47 (dd, J = 9.8, 8.8 Hz, 1H, H2) **^{13}C NMR** (101 MHz, D_2O) δ 90.5 (C1), 77.2 (C5), 72.6 (C3), 70.3 (C2), 68.5 (C4), 60.9 (C6)

Azido 2,3,4,6-tetra-O-acetyl-β-D-galactopyranoside(1→4)-[3,4-di-O-benzoyl-α-L-fucopyranoside-(1→3)]-2-deoxy-2-(N-acetylacetamido)-β-D-glucopyranoside (35)



Protected trisaccharide **34** (243 mg, 0.2 mmol, see Chapter 2 for the preparation of this molecule) was dissolved in DCM (1.9 mL) and H_2O (0.1 mL). 2,3-Dichloro-5,6-dicyano-1,4-benzoquinone (182 mg, 0.8 mmol, 4 eq) was added and the resulting suspension was stirred overnight. The next day, TLC analysis (1/1 EtOAc/pentane) indicated full consumption of starting material. The reaction mixture was diluted with DCM and quenched by the addition of saturated aqueous $NaHCO_3$. The layers were separated and the aqueous layer was re-extracted with DCM. The combined organic layers were dried over $MgSO_4$, filtered and concentrated. Silicagel column chromatography (1/1 EtOAc/pentane) yielded partially deprotected trisaccharide **35** in 71% yield (138 mg, 142 µmol). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.04 (dd, J = 8.4, 1.3 Hz, 2H, Bz), 7.82 (dd, J = 8.3, 1.2 Hz, 2H, Bz), 7.60 (tt, J = 7.0, 1.3 Hz, 1H, Bz), 7.51 – 7.42 (m, 3H, Bz), 7.32 – 7.24 (m, 2H, Bz), 5.69 (d, J = 2.7 Hz, 1H, H4''), 5.64 (d, J = 8.8 Hz, 1H, H1), 5.57 (dd, J = 10.5, 3.4 Hz, 1H, H3''), 5.48 (dd, J = 3.3, 0.7 Hz, 1H, H4'), 5.20 – 5.07 (m, 3H, H2', H5'', H3'), 4.91 – 4.83 (m, 3H, H1'', H1', H6'a), 4.79 (t, J = 9.5 Hz, 1H, H3), 4.24 (dd, J = 10.5, 4.0 Hz, 1H, H2''), 4.17 (dd, J = 11.7, 6.1 Hz, 1H, H6'b), 4.03 (d, J = 11.3 Hz, 1H, H6a), 4.00 – 3.91 (m, 2H, H5', H4), 3.87 (d, J = 9.9 Hz, 1H, H6b), 3.68 (t, J = 9.3 Hz, 1H, H2), 3.56 (d, J = 9.9 Hz, 1H, H5), 2.51 (s, 3H, Ac), 2.45 (s, 3H, Ac), 2.28 (s, 3H, Ac), 2.08 (s, 3H, Ac), 2.06 (s, 3H, Ac), 2.00 (s, 3H, Ac), 1.31 (d, J = 6.6 Hz, 3H, H6'') **^{13}C NMR** (101 MHz, $CDCl_3$) δ 175.4 (C=O), 174.5 (C=O), 170.9 (C=O), 170.6 (C=O), 170.2 (C=O), 169.0 (C=O), 166.3 (C=O), 166.2 (C=O), 133.4 (Bz-CH), 133.1 (Bz-CH), 129.9 (Bz-CH), 129.7 (Bz-CH), 129.7 (Bz-C_q), 129.6 (Bz-C_q), 128.6 (Bz-CH), 128.3 (Bz-CH), 100.1 (C1'), 98.3 (C1''), 86.8 (C1), 77.1 (C5), 74.1 (C4), 72.3 (C4''), 72.0 (C5'), 71.8 (C3''), 71.6 (C3), 71.0 (C3'), 69.1 (C2'), 67.5 (C2''), 67.3 (C4'), 65.2 (C5''), 64.1 (C2), 61.7 (C6'), 59.9 (C6), 28.4 (Ac), 25.8 (Ac), 21.0 (Ac), 20.8 (Ac), 20.7 (Ac), 16.1 (C6'')

β-D-galactopyranoside(1→4)-[α-L-fucopyranoside-(1→3)]-2-deoxy-2-acetamido-β-D-glucopyranosyl azide (36)



A fresh solution of sodium methoxide was prepared by dissolving a small chuck of sodium metal in methanol (2 mL). The pH of the solution was checked to be 12 before it was added to a separate flask containing partially protected glycoside **35** (155 mg, 160 µmol). After 24 hours, LC-MS analysis indicated full conversion to the desired product. The reaction mixture was neutralized by addition of amberlite H^+ ion exchange resin and the solution was carefully transferred into a centrifuge tube containing ice-cold Et_2O (20 mL). The resin was washed twice with methanol (2x1 mL) and these washes were also added to the ether. A precipitate was formed that was isolated by centrifugation. The supernatant was discarded and the pellet was washed once with Et_2O (5 mL). The pellet was then transferred to a flask and dried under vacuum, yielding desired oligosaccharide **36** in quantitative (94 mg) yield. **^{1}H NMR** (400 MHz, D_2O) δ 4.99 (d, J = 4.0 Hz, 1H, H1''), 4.76 – 4.73 (m, 1H, H5''), 4.68 (d, J = 8.8 Hz, 1H, H1), 4.34 (d, J = 7.8 Hz, 1H, H1'), 3.93 – 3.51 (m, 13H), 3.48 (dd, J = 7.6, 4.4 Hz, 1H, H5), 3.38 (dd, J = 9.7, 7.9 Hz, 1H, H2), 1.93 (s, 3H, Ac), 1.06 (d, J = 6.6 Hz, 3H, H6'') **^{13}C NMR** (101 MHz, D_2O) δ 174.5 (C=O), 101.8 (C1'), 98.8 (C1''), 88.6 (C1), 77.4, 74.9, 74.9 (C5), 72.9, 72.4, 71.9, 71.0, 69.2, 68.3, 67.7, 66.7 (C5''), 61.5 (C6), 59.6 (C6'), 55.3 (C2), 22.2 (Ac), 15.3 (C6'') **HRMS** (ESI) m/z $[M + Na]^+$ calcd for $C_{20}H_{34}N_4O_{14}Na$ 577.1964 found 577.1966

***Cyclo*(-Lys-Abu-Lys(TEG-TEG-propargyl)-Abu-Lys(TEG-TEG-propargyl)-Abu-Lys(TEG-TEG-propargyl)-Abu-Lys(TEG-TEG-propargyl)-Abu-Lys(TEG-TEG-propargyl)-Abu-) (29)**

Following the general procedures for cyclization and sidechain modification, compound **29** was isolated using RP-HPLC in 4.6% (7.7 mg, 2.3 µmol) yield. **LC-MS** RT = 4.4 min (C18, 10-90% B over 9 minutes) **LRMS** calcd $[M+3H]^{3+}$ = 1122.96, $[M+2H]^{2+}$ = 1683.95; observed M/z = 1123.25, 1684.08 **HRMS** (ESI) m/z $[M + 4H]^{4+}$ calcd for $C_{155}H_{269}N_{23}O_{57}H_4$ 842.2287 found 842.2289

***Cyclo*(-Lys(Atto655)-Abu-Lys(TEG-TEG-triazole-β-Gal)-Abu-Lys(TEG-TEG-triazole-β-Gal)-Abu-Lys(TEG-TEG-triazole-β-Gal)-Abu-Lys(TEG-TEG-triazole-β-Gal)-Abu-Lys(TEG-TEG-triazole-β-Gal)-Abu-) (38)**

To a solution of scaffold **29** (100 nmol, 50 mM in DMSO) was added a solution of glycosyl azide **32** in water (10 eq, 5 µL of 200 mM). A click mix was prepared by mixing 1 µL of 0.1 M $CuSO_4$, 2 µL 0.1 M sodium ascorbate and 3 µL of 0.1 M THPTA, and 1 µL of this mixture was added to the scaffold solution. The solution was gently agitated and heated to 40°C for 48h, after which a 0.5 µL portion was taken out and diluted into 35 µL 1:1:1 H_2O/MeCN/tBuOH and subjected to LC-MS analysis, indicating full conversion to the desired glycoconjugate. The reaction mixture was diluted with 200 µL MilliQ water and lyophilized. The crude glycoconjugate was redissolved in 50 µL DMSO and a solution of DiPEA in DMSO (10 eq, 10 µL of 100 mM) was added, followed by a solution of Atto655-NHS in DMSO (2 eq, 40 µL of 5 mM). The solution was gently agitated and heated to 40°C for 16h, after which a 1 µL portion was taken out and diluted into 39 µL 1:1:1 H_2O/MeCN/tBuOH and subjected to LC-MS analysis, indicating formation of the labeled molecule. The crude was diluted with water and subjected to HW40 size exclusion chromatography in a NH_4OAc buffer containing 10% MeCN. The fractions containing fluorophore labeled product were pooled and lyophilized, producing compound **38** in 87% yield (0.43

mg, 87 nmol) as a blue powder. **LC-MS** RT = 5.7 min (C18, 00-50% B over 9 minutes). **LRMS** calcd $[M+4H]^{4+}$ = 1226.36, $[M+3H]^{3+}$ = 1634.81; observed M/z = 1126.42, 1634.83

Cyclo(-Lys(Atto655)-Abu-Lys(TEG-TEG-triazole-Le^X)-Abu-Lys(TEG-TEG-triazole-Le^X)-Abu-Lys(TEG-TEG-triazole-Le^X)-Abu-Lys(TEG-TEG-triazole-Le^X)-Abu-Lys(TEG-TEG-triazole-Le^X)-Abu-) (40)

To a solution of scaffold **29** (100 nmol, 50 mM in DMSO) was added a solution of glycosyl azide **36** in DMSO (6 eq, 6 µL of 100 mM). A click mix was prepared by mixing 1 µL of 0.1 M $CuSO_4$, 2 µL 0.1 M sodium ascorbate and 3 µL of 0.1 M THPTA, and 1 µL of this mixture was added to the scaffold solution. The solution was gently agitated and heated to 40°C for 48h, after which a 0.5 µL portion was taken out and diluted into 35 µL 1:1:1 H_2O/MeCN/tBuOH and subjected to LC-MS analysis, indicating full conversion to the desired glycoconjugate. Next, a solution of DiPEA in DMSO (10 eq, 10 µL of 100 mM) was added, followed by a solution of Atto655-NHS in DMSO (2 eq, 40 µL of 5 mM). The solution was gently agitated and heated to 40°C for 16h, after which a 1 µL portion was taken out and diluted into 39 µL 1:1:1 H_2O/MeCN/tBuOH and subjected to LC-MS analysis, indicating formation of the labeled molecule. The crude was diluted with water and subjected to HW40 size exclusion chromatography in a NH_4OAc buffer containing 10% MeCN. The fractions containing fluorophore labeled product were pooled and lyophilized, producing compound **40** in 42% yield (0.29 mg, 42 nmol) as a blue powder. **LC-MS** RT = 4.3 min (C18, 10-50% B over 9 minutes). **LRMS** calcd $[M+5H]^{5+}$ = 1330.43, $[M+4H]^{4+}$ = 1662.78; observed M/z = 1330.50, 1662.83

References

(1) CAMERON, D.; JONES, I. G. John Snow, the Broad Street Pump and Modern Epidemiology. *Int. J. Epidemiol.* **1983**, *12* (4), 393–396. https://doi.org/10.1093/IJE/12.4.393.

(2) Clemens, J. D.; Nair, G. B.; Ahmed, T.; Qadri, F.; Holmgren, J. Cholera. *Lancet* **2017**, *390* (10101), 1539–1549. https://doi.org/10.1016/S0140-6736(17)30559-7.

(3) Fan, E.; Merritt, E. A.; Verlinde, C. L. M. J.; Hol, W. G. J. AB5 Toxins: Structures and Inhibitor Design. *Current Opinion in Structural Biology*. Elsevier Current Trends December 1, 2000, pp 680–686. https://doi.org/10.1016/S0959-440X(00)00152-4.

(4) Wernick, N. L. B. B.; Chinnapen, D. J.-F.; Cho, J. A.; Lencer, W. I. Cholera Toxin: An Intracellular Journey into the Cytosol by Way of the Endoplasmic Reticulum. *Toxins (Basel).* **2010**, *2* (3), 310–325. https://doi.org/10.3390/toxins2030310.

(5) Cuatrecasas, P. Interaction of Vibrio Cholerae Enterotoxin with Cell Membranes. *Biochemistry* **1973**, *12* (18), 3547–3558. https://doi.org/10.1021/bi00742a031.

(6) King, C. A.; Heyningen, W. E. V. Deactivation of Cholera Toxin by a Sialidase-Resistant Monosialosylganglioside. *J. Infect. Dis.* **1973**, *127* (6), 639–647. https://doi.org/10.1093/infdis/127.6.639.

(7) Holmgren, J.; Lönnroth, I.; Svennerholm, L. Tissue Receptor for Cholera Exotoxin: Postulated Structure from Studies with GM1 Ganglioside and Related Glycolipids. *Infect. Immun.* **1973**, *8* (2), 208.

(8) Heim, J. B.; Hodnik, V.; Heggelund, J. E.; Anderluh, G.; Krengel, U. Crystal Structures of Cholera Toxin in Complex with Fucosylated Receptors Point to Importance of Secondary Binding Site. *Sci. Rep.* **2019**, *9* (1), 1–14. https://doi.org/10.1038/s41598-019-48579-2.

(9) Heggelund, J. E.; Burschowsky, D.; Bjørnestad, V. A.; Hodnik, V.; Anderluh, G.; Krengel, U. High-Resolution Crystal Structures Elucidate the Molecular Basis of Cholera Blood Group Dependence. *PLoS Pathog.* **2016**, *12* (4), e1005567. https://doi.org/10.1371/journal.ppat.1005567.

(10) Wands, A. M.; Fujita, A.; McCombs, J. E.; Cervin, J.; Dedic, B.; Rodriguez, A. C.; Nischan, N.; Bond, M. R.; Mettlen, M.; Trudgian, D. C.; Lemoff, A.; Quiding-Järbrink, M.; Gustavsson, B.; Steentoft, C.; Clausen, H.; Mirzaei, H.; Teneberg, S.; Yrlid, U.; Kohler, J. J. Fucosylation and Protein Glycosylation Create Functional Receptors for Cholera Toxin. *Elife* **2015**, *4* (OCTOBER2015). https://doi.org/10.7554/eLife.09545.

(11) Lencer, W.; Delp, C.; Neutra, M.; Madara, J. Mechanism of Cholera Toxin Action on a Polarized Human Intestinal Epithelial Cell Line: Role of Vesicular Traffic. *J. Cell Biol.* **1992**, *117* (6), 1197–1209. https://doi.org/10.1083/JCB.117.6.1197.

(12) Lencer, W. I.; Constable, C.; Moe, S.; Rufo, P. A.; Wolf, A.; Jobling, M. G.; Ruston, S. P.; Madara, J. L.; Holmes, R. K.; Hirst, T. R. Proteolytic Activation of Cholera Toxin and Escherichia Coli Labile Toxin by Entry into Host Epithelial Cells: SIGNAL TRANSDUCTION BY A PROTEASE-RESISTANT TOXIN VARIANT *. *J. Biol. Chem.* **1997**, *272* (24), 15562–15568. https://doi.org/10.1074/JBC.272.24.15562.

(13) Cervin, J.; Wands, A. M.; Casselbrant, A.; Wu, H.; Krishnamurthy, S.; Cvjetkovic, A.; Estelius, J.; Dedic, B.; Sethi, A.; Wallom, K.-L. L.; Riise, R.; Bäckström, M.; Wallenius, V.; Platt, F. M.; Lebens, M.; Teneberg, S.; Fändriks, L.; Kohler, J. J.; Yrlid, U. GM1 Ganglioside-Independent Intoxication by Cholera Toxin. *PLoS Pathog.* **2018**, *14* (2), e1006862. https://doi.org/10.1371/journal.ppat.1006862.

(14) Minke, W. E.; Roach, C.; Hol, W. G. J.; Verlinde, C. L. M. J. Structure-Based Exploration of the Ganglioside GM1 Binding Sites of Escherichia Coli Heat-Labile Enterotoxin and Cholera Toxin for the Discovery of Receptor Antagonists. *Biochemistry* **1999**, *38* (18), 5684–5692. https://doi.org/10.1021/bi982649a.

(15) Fan, E.; Zhang, Z.; Minke, W. E.; Hou, Z.; Verlinde, C. L. M. J.; Hol, W. G. J. High-Affinity Pentavalent Ligands of Escherichia Coli Heat-Labile Enterotoxin by Modular Structure-Based Design [10]. *J. Am. Chem. Soc.* **2000**, *122* (11), 2663–2664. https://doi.org/10.1021/ja993388a.

(16) Garcia-Hartjes, J.; Bernardi, S.; Weijers, C. A. G. M.; Wennekes, T.; Gilbert, M.; Sansone, F.; Casnati, A.; Zuilhof, H. Picomolar Inhibition of Cholera Toxin by a Pentavalent Ganglioside GM1os-Calix[5]Arene. *Org. Biomol. Chem.* **2013**, *11* (26), 4340–4349. https://doi.org/10.1039/c3ob40515j.

(17) Zhang, Z.; Liu, J.; Verlinde, C. L. M. J.; Hol, W. G. J.; Fan, E. Large Cyclic Peptides as Cores of Multivalent Ligands: Application to Inhibitors of Receptor Binding by Cholera Toxin. *J. Org. Chem.* **2004**, *69* (22), 7737–7740. https://doi.org/10.1021/jo0489770.

(18) Pickens, J. C.; Mitchell, D. D.; Liu, J.; Tan, X.; Zhang, Z.; Verlinde, C. L. M. J.; Hol, W. G. J.; Fan, E. Nonspanning Bivalent Ligands as Improved Surface Receptor Binding Inhibitors of the Cholera Toxin B Pentamer. *Chem. Biol.* **2004**, *11* (9), 1205–1215. https://doi.org/10.1016/J.CHEMBIOL.2004.06.008.

(19) Achtsnicht, S.; Neuendorf, C.; Faßbender, T.; Nölke, G.; Offenhäusser, A.; Krause, H.-J.; Schröper, F. Sensitive and Rapid Detection of Cholera Toxin Subunit B Using Magnetic Frequency Mixing Detection. *PLoS One* **2019**, *14* (7), e0219356. https://doi.org/10.1371/JOURNAL.PONE.0219356.

(20) Erak, M.; Bellmann-Sickert, K.; Els-Heindl, S.; Beck-Sickinger, A. G. Peptide Chemistry Toolbox – Transforming Natural Peptides into Peptide Therapeutics. *Bioorg. Med. Chem.* **2018**, *26* (10), 2759–2765. https://doi.org/10.1016/J.BMC.2018.01.012.

(21) Alsina, J.; Rabanal, F.; Chiva, C.; Giralt, E.; Albericio, F. Active Carbonate Resins: Application to the Solid-Phase Synthesis of Alcohol, Carbamate and Cyclic Peptides. *Tetrahedron* **1998**, *54* (34), 10125–10152. https://doi.org/10.1016/S0040-4020(98)00606-1.

(22) Reintjens, N. R. M. M.; Tondini, E.; Vis, C.; McGlinn, T.; Meeuwenoord, N. J.; Hogervorst, T. P.; Overkleeft, H. S.; Filippov, D. V.; Marel, G. A. van der; Ossendorp, F.; Codée, J. D. C. C.; van der Marel, G. A.; Ossendorp, F.; Codée, J. D. C. C. Multivalent, Stabilized Mannose-6-Phosphates for the Targeted Delivery of Toll-Like Receptor Ligands and Peptide Antigens. *ChemBioChem* **2021**, *22* (2), 434–440. https://doi.org/10.1002/CBIC.202000538.

(23) Ning, X.; Liu, D.; Liu, S.; Zhang, M.; Shen, H. Modification of Oligopeptides on Aspartic Acid or Lysine Residues by Solid-Phase Synthesis through On-Resin Side-Chain Conjugation. *Synlett* **2018**, *29* (19), 2588–2594. https://doi.org/10.1055/S-0037-1611060.

(24) Aletras, A.; Barlos, K.; Gatos, D.; Koutsogianni, S.; MAMOS, P. Preparation of the Very Acid-sensitive Fmoc-Lys(Mtt)-OH Application in the Synthesis of Side-chain to Side-chain Cyclic Peptides and Oligolysine Cores Suitable for the Solid-phase Assembly of MAPs and TASPs. *Int. J. Pept. Protein Res.* **1995**, *45* (5), 488–496. https://doi.org/10.1111/j.1399-3011.1995.tb01065.x.

(25) Albericio, F.; Cases, M.; Alsina, J.; Triolo, S. A.; Carpino, L. A.; Kates, S. A. On the Use of PyAOP, a Phosphonium Salt Derived from HOAt, in Solid-Phase Peptide Synthesis. *Tetrahedron Lett.* **1997**, *38* (27), 4853–4856. https://doi.org/10.1016/S0040-4039(97)01011-3.

(26) Zhang, H.; Li, X.; Shi, Q.; Li, Y.; Xia, G.; Chen, L.; Yang, Z.; Jiang, Z.-X. Highly Efficient Synthesis of Monodisperse Poly(Ethylene Glycols) and Derivatives through Macrocyclization of Oligo(Ethylene Glycols). *Angew. Chem. Int. Ed.* **2015**, *54* (12), 3763–3767. https://doi.org/10.1002/ANIE.201410309.

(27) Thakkar, A.; Trinh, T. B.; Pei, D. Global Analysis of Peptide Cyclization Efficiency. *ACS Comb. Sci.* **2013**, *15* (2), 120–129. https://doi.org/10.1021/CO300136J.

(28) O'Neal, C. J.; Amaya, E. I.; Jobling, M. G.; Holmes, R. K.; Hol, W. G. J. Crystal Structures of an Intrinsically Active Cholera Toxin Mutant Yield Insight into the Toxin Activation Mechanism. *Biochemistry* **2004**, *43* (13), 3772–3782. https://doi.org/10.1021/bi0360152.

(29) Fehres, C. M.; Kalay, H.; Bruijns, S. C. M.; Musaafir, S. A. M.; Ambrosini, M.; Van Bloois, L.; Van Vliet, S. J.; Storm, G.; Garcia-Vallejo, J. J.; Van Kooyk, Y. Cross-Presentation through Langerin and DC-SIGN Targeting Requires Different Formulations of Glycan-Modified Antigens. *J. Control. Release* **2015**, *203*, 67–76.

(30) Riera, R.; Hogervorst, T. P.; Doelman, W.; Ni, Y.; Pujals, S.; Bolli, E.; Codée, J. D. C.; van Kasteren, S. I.; Albertazzi, L. Single-Molecule Imaging of Glycan–Lectin Interactions on Cells with Glyco-PAINT. *Nat. Chem. Biol.* **2021**, *17* (12), 1281–1288. https://doi.org/10.1038/s41589-021-00896-2.

(31) van Kooyk, Y.; Geijtenbeek, T. B. H. DC-SIGN: Escape Mechanism for Pathogens. *Nat. Rev. Immunol. 2003 39* **2003**, *3* (9), 697–709. https://doi.org/10.1038/nri1182.

(32) Ghosh, P.; Dahms, N. M.; Kornfeld, S. Mannose 6-Phosphate Receptors: New Twists in the Tale. *Nat. Rev. Mol. Cell Biol. 2003 43* **2003**, *4* (3), 202–213. https://doi.org/10.1038/nrm1050.