Cyclophellitol analogues for profiling of exo- and endo-glycosidases

Schröder, S.P.

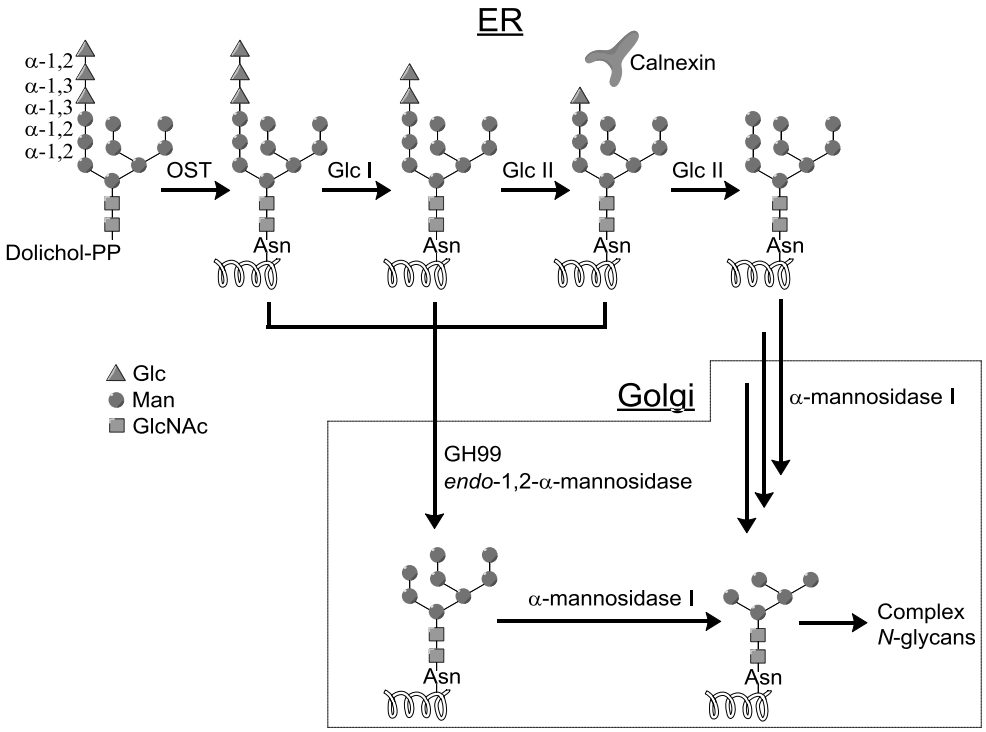
Chapter 6

Synthesis of spiro-epoxyglycosides as potential probes for GH99 *endo*-mannosidases

6.1 Introduction

N-linked glycans are complex oligosaccharides that are linked to asparagine (Asn) residues in proteins in eukaryotic organisms.[1] They play important roles in protein function, stability, folding and targeting and are therefore essential to cellular function.[2] Erroneous *N*-glycan composition is associated with various diseases such as viral infections, Alzheimer disease and metastatic cancer.[3–5] In the endoplasmic reticulum (ER), the 14-mer polysaccharide $Glc_3Man_9GlcNAc_2$-

diphosphodolichol is coupled to the Asn residue of the target protein by the action of the enzyme, oligosaccharyl transferase (OST), after which the glycan undergoes stepwise 'trimming' of the non-reducing end glucoside residues (Scheme 1).[6] The terminal α-1,2 linked glucose residue is cleaved by membrane bound α-glucosidase I in the ER, after which the soluble α-glucosidase II cleaves off the second α-1,3-glucose monomer. The remaining glucose residue is important for quality control as it enables the binding of calnexin and calreticulin; chaperones acting as lectins, inducing the proper folding of the glycoprotein. When the protein is properly folded the last α-1,3-glucose residue is removed by α-glucosidase II, followed by trimming of the terminal mannose residues by ER/Golgi α-mannosidase I. Finally, the core $Man_5GlcNAc_2$ *N*-



Scheme 1 Schematic representation of the assembly and processing of *N*-linked glycans in the ER and Golgi apparatus. Normal processing is performed by Glc I and II by stepwise 'trimming' of the terminal glucose residues in the ER, followed by α-mannosidase I which trims of the first mannose residue. Calnexin recognizes $Glc_1Man_9GlcNAc_2$ and acts as a folding chaperone for quality control. GH99 *endo*-α-1,2-mannosidase is present in the Golgi apparatus and provides an alternative route towards the mature *N*-glycan by cleaving $Glc_{1\text{-}3}Man_9GlcNAc_2$ in an *endo*-fashion, affording $Man_8GlcNAc_2$.

glycan is redecorated with a variety of saccharides. The final structure of the mature glycan depends on the specific glycoprotein. Because α-glucosidases I and II play important roles in the early stages of glycan maturation, these enzymes were marked as therapeutic targets to control diseases involving incorrect *N*-glycosylation.[7–10] However, inhibition of these enzymes did not result in blocking *N*-glycosylation: cells inhibited with α-glucosidase inhibitors as well as mutant cell lines lacking α-glucosidase II retained up to 80% of normal *N*-glycan maturation.[11–13] Spiro and co-workers[14,15] identified an *endo*-glycosidase residing in the Golgi and able to circumvent inhibition of α-glucosidases I and II, namely *endo*-1,2-α-mannosidase (later classified as a member of glucoside hydrolase family 99 (GH99); see *cazypedia.org*). While the sequential deglucosylation steps by α-glucosidase I and II takes place in the ER, *endo*-1,2-α-mannosidase is present in the Golgi apparatus. The enzyme surpasses the stepwise trimming of the terminal glucoside residues by glucosidase I and II by direct *endo* cleavage of the immature glycan $Glc_{1\text{-}3}Man_9GlcNAc_2$, releasing $Glc_{1\text{-}3}Man$. The resulting $Man_8GlcNAc_2$ *N*-glycan subsequently re-enters the normal processing route.

In order to further study the role and function of this enzyme, bacterial orthologues from *B. thetaiotaomicron (Bx)* and *B. xylanisolvens (Bt)* (42% and 41% sequence identity, respectively) were incubated with α-glucopyranosyl-1,3-mannopyranosyl fluoride, which resulted in the release of hydrolyzed product with retention of stereochemistry.[16] Generally, retaining glycosidases follow a classical Koshland catalytic mechanism (Figure 1A).[17] The substrate enters the active site, where the aglycon is protonated by the catalytic acid and is displaced by the catalytic nucleophile, forming a covalent intermediate. The resulting catalytic conjugate base deprotonates a water molecule which then hydrolyzes this intermediate, resulting in the hydrolyzed product with retention of anomeric stereochemistry. In order to investigate the catalytic mechanism of *endo*-α-1,2-mannosidase, *Bx*GH99 was crystallized in complex with the iminosugars, α-glucopyranosyl-1,3-deoxynojirimycin (GlcDMJ) and α-glucopyranosyl-1,3-isofagomine (ManIFG) by Thompson *et al.* to locate the catalytic active site residues.[16] These crystal structures indicate the presence of the acid-base residue coordinated to a water molecule located at the "α-face" of the anomeric position of the substrate. Of particular interest is the apparent absence of a catalytic nucleophile suitable for nucleophilic attack at the anomeric carbon. The closest candidate for nucleophilic attack was found to be Glu-333, situated close (2.6 Å) to 2-OH of the substrate. Because there was no evidence for

conformational flexibility in the active site, a unusual hydrolytic mechanism was proposed in which Glu-333 deprotonates the 2-OH, which then displaces the aglycon via neighboring group participation, forming a 1,2-anhydro epoxide which is then hydrolyzed by a water molecule (Figure 1B). From a conformational point of view, the catalytic itinerary of the substrate during catalysis starts with the Michaelis complex ($^4C_1$), followed by its transition state (^4E) which upon loss of the aglycon forms the 1,2-anhydro sugar intermediate ($^4H_5$) which is subsequently hydrolyzed by water (Figure 1C).[18]

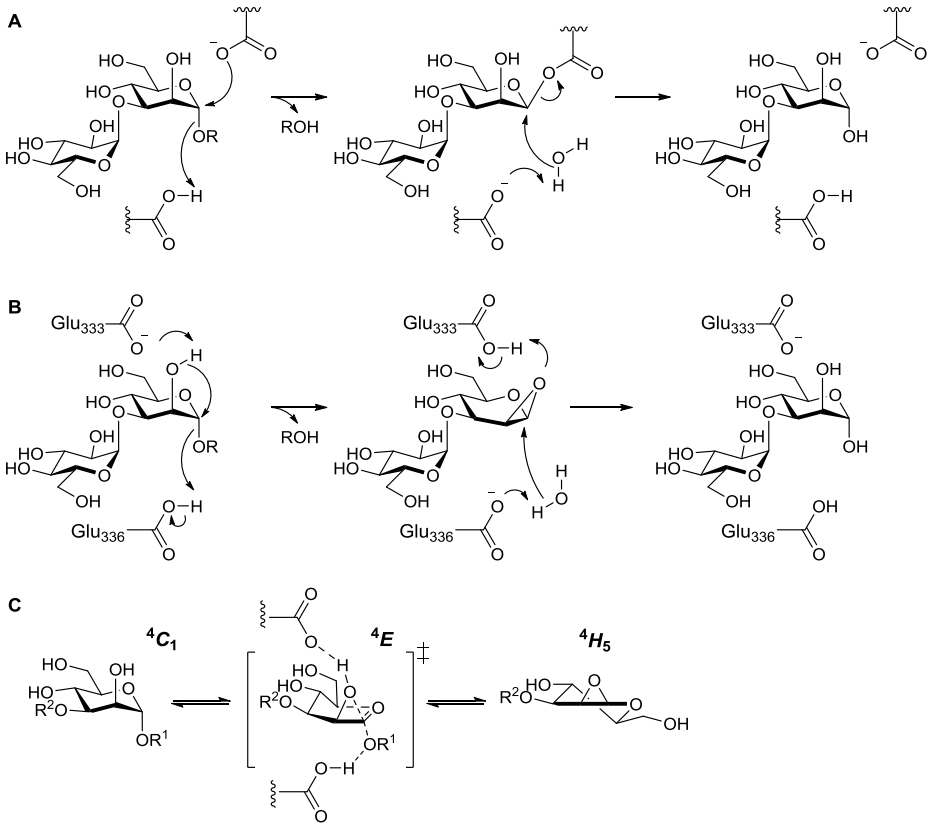


Figure 1 A) Classical Koshland reaction mechanism *via* a covalent enzyme-inhibitor intermediate; B) Reaction mechanism for GH99 *endo*-α-1,2-mannosidase proposed by Thompson *et al;*[16] C) Conformational itinerary for the substrate hydrolysis by GH99 *endo*-α-1,2-mannosidase.

Since *N*-glycosylation only rarely occurs in bacteria and *Bx* and *Bt* are both crucial members of the gut microbiota, the bacterial GH99 *endo*-α-mannosidases might have

originated by horizontal gene transfer from eukaryotic hosts. In this way, *Bx* and *Bt* are able to process highly functionalized yeast cell wall α-mannans present in their host diet, and use it as carbon source. Indeed, *Bx*- and *Bt*GH99 enzymes may be more accurately classified as *endo*-α-mannanases, as they show better affinity for α-mannopyranosyl-1,3-isofagomine (ManIFG, Figure 2) than for α-glucopyranosyl-1,3-isofagomine (GlcIFG).[19] The preference for a mannoside residue at subsite -2 (-4.7 kJ mol^{-1}) was further elucidated by comparing the crystal structures of *Bx*GH99 complexed with GlcIFG and ManIFG, respectively, which revealed repulsion of the 2-OH of the Glc residue by the hydrophobic sidechain of a tryptophan in the active site. In eukaryotes the Trp residue in the enzyme is replaced by a highly conserved tyrosine residue, allowing preferential H-bonding with the 2-OH of Glc in subsite -2. Recently, mannoeuromycin (ManNOE) has been reported as the most potent *endo*-α-1,2-mannanase inhibitor for bacterial GH99 enzymes with K_D values in the low nanomolar range, possibly due to the interaction of O2 with the proposed catalytic base.[20]

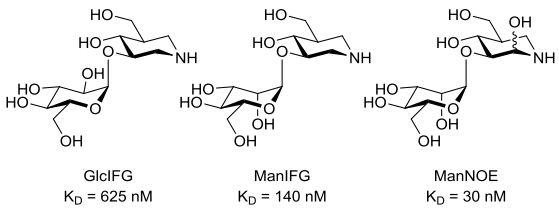


Figure 2 Isofagomine disaccharide analogues α-glucopyranosyl-1,3-isofagomine (GlcIFG), α-mannopyranosyl-1,3-isofagomine (ManIFG) and mannoeuromycin (ManNOE) are potent inhibitors of bacterial GH99 *endo*-α-mannanases. A mannopyranosyl residue at the non-reducing end is favored for *Bt*GH99. Introduction of a hydroxyl group at C2 also increases the potency. K_D values are given for *Bt*GH99.

To date, the synthesis of small-molecules designed for functional and structural investigation of GH99 *endo*-α-1,2-mannosidases and *endo*-α-1,2-mannanases have focused exclusively on reversible inhibitors. In order to study enzyme function in a biological setting, to screen for novel inhibitors, as well as to further elucidate the enzymatic catalytic reaction mechanism, the development of an irreversible ABP would be of interest. This Chapter describes the synthesis of two spiro-epoxyglycosides that contain either a glucose or mannose residue at the non-reducing end, designed to respectively inhibit eukaryotic GH99 *endo*-α-mannosidase and

bacterial *endo*-α-mannanase enzymes (Figure 3). Both compounds contain a spiro-epoxide at position C-2 that may serve as an electrophile to trap the catalytic base residue. Following the proposed unusual catalytic mechanism of the enzyme, the catalytic base is situated close to the methylene group of the spiro-epoxide, and is anticipated to be able to open the ring via nucleophilic attack resulting in a covalent intermediate. The compounds are also equipped with a reporter tag, and following activity-based protein profiling (ABPP) protocols the labeling efficiency of bacterial GH99 *endo*-α-mannanases by these spiro-epoxyglycosides is investigated. Activity-based labeling of these enzymes in the active site would give a strong indication that GH99 enzymes indeed follow the proposed unusual reaction mechanism depicted in Figure 1B. Additionally, it would be of interest to compare the potency of these compounds towards *endo*-α-mannanases, depending on the pyranosyl substituent at the -2 subsite of the enzyme.

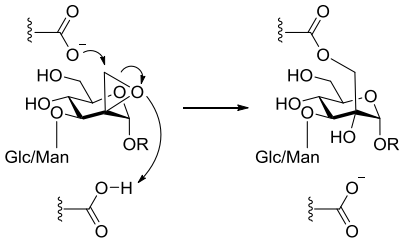
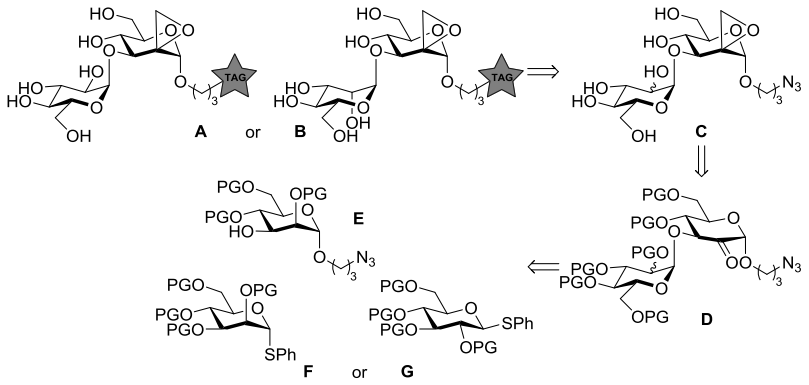


Figure 3 General representation of the target compounds described in this Chapter, including a schematic representation of the covalent binding of the inhibitor to the enzyme based on the proposed reaction mechanism. The spiro-epoxyglycosides contain either an α-glucopyranosyl or an α-mannopyranosyl residue at the non-reducing end and the R-group represents the reporter tag.

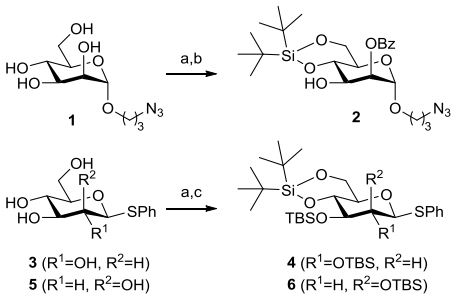
It was envisioned that compounds **A** and **B** could be obtained from matching azides **C** via click-chemistry (Scheme 2). The spiro-epoxide could be installed via a Corey-Chaykovsky epoxidation of ketones **D**, which in turn could be obtained by coupling of acceptor **E** and donor **F** or **G** and subsequent oxidation of the C-2 hydroxyl group. Silyl groups were chosen as global protecting groups since they could ultimately allow deprotection under mild conditions, whilst leaving the azide and acid-labile spiro epoxide intact.



Scheme 2 Retrosynthetic scheme for the assembly of spiro-epoxyglycosides **A** and **B**. Compound **A** contains a glucose residue at the -2 subsite and is therefore anticipated to favor binding by eukaryotic GH99 *endo*-α-1,2-mannosidases. In contrast, compound **B** contains a mannose residue at the -2 subsite and is therefore anticipated to favor binding by bacterial GH99 *endo*-α-1,2-mannanases.

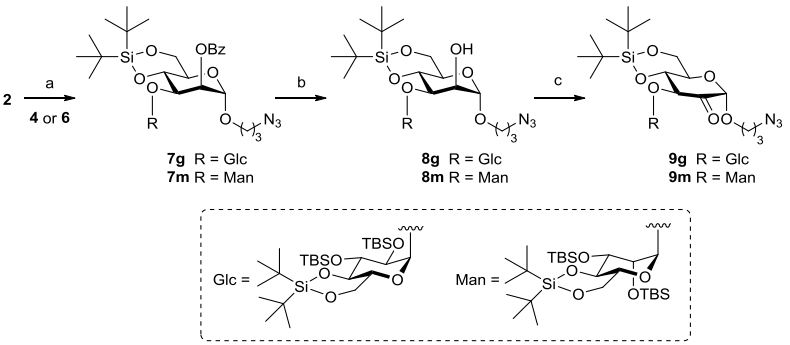
6.2 Results and Discussion

Acceptor **2** was synthesized by 4,6-silylidene protection of compound **1** (prepared by Fischer glycosylation of D-mannose with 1-chloro-propanol and subsequent substitution of the chlorine by azide),[21] followed by formation of the 2,3-orthobenzoate and final treatment with acid (Scheme 3).[22] Glucopyranosyl-configured donor **4** was synthesized from anomeric β-thiophenyl glucopyranoside **3**.[23] While 4,6-silydene protection proceeded smoothly, elevated temperatures were required for complete TBS-protection of the hydroxyl groups on position 2 and 3, presumably due to steric hindrance. Under the same reaction conditions, anomeric α-thiophenyl mannopyranoside **5**[24] was converted to fully protected α-thioglycoside donor **6**.



Scheme 3 Synthesis of acceptor **2** and donors **4** and **6**. Reagents and conditions: a) $^{t}Bu_2Si(OTf)_2$, 2,6-lutidine, DMF, -50 °C; b) $PhCH(OMe)_3$, CSA, 2h, then AcOH, H_2O, 16h, 57% over 2 steps; c) TBSOTf, DMAP, pyridine, 60 °C, 16h, yield **4**: 85% over 2 steps; yield **6**: 81% over 2 steps.

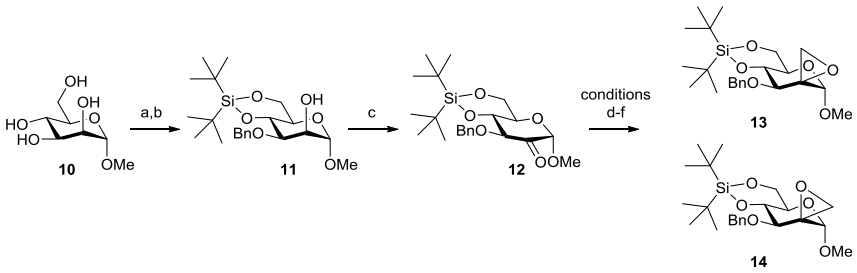
Acceptor **2** was coupled to **4** or **6** in a *N*-iodosuccinimide (NIS)/trimethylsilyl triflate (TMSOTf) mediated coupling at low temperature. Both glycosylations proceeded in excellent yield and stereoselectivity. Pedersen and Bols and co-workers[25] recently reported that silylidene protected mannosyl donors can be used for the stereoselective installation of beta-mannosidic linkages. The contrasting selectivity obtained here is likely the result of the steric buttressing effect of the large silyl ether protecting groups at the C-2- and C-3-hydroxyls, much in line with the steric effects large protecting groups and functionalities have in glycosylations of otherwise beta-selective benzylidene mannosyl donors.[26] Thus, the beta-face of mannosyl donor **6** is effectively shielded for attack by the incoming nucleophile. The stereoselectivity in the glycosylation of glucosyl donor **4** can be rationalized by the reactivity of the used donor. The 'arming' silyl protecting groups allow this donor to readily form an oxocarbenium ion, which will likely take up a $^3H_4$-like conformation, which is preferentially attacked from the alpha-face to provide the 1,2-*cis*-linked product.[27] Next, the benzoyl groups were deprotected under standard Zémplen conditions affording compounds **8g** and **8m**. The alcohols were then oxidized to ketones **9g** and **9m** with Dess-Martin periodinane (DMP), and after work-up the resulting ketones appeared to be in equilibrium with the corresponding hydrates.



Scheme 4 α-Selective glycosylations of acceptor **2** with donors **4** and **6**, followed by transformation into key-intermediate ketones **9g** and **9m**. Reagents and conditions: a) Donor **4** or **6**, NIS, TMSOTf, DCM, 4Å MS, -40 °C, 1h, yield **7g**: 92%; yield **7m**: 88%; b) NaOMe, MeOH, DCM, yield **8g**: 95%; yield **8m**: 86%; c) DMP, DCM, yield **9g**: 98%; yield **9m**: 96%.

Next, the transformation of ketones **9g** and **9m** into their corresponding spiro-epoxides was explored. In the first instance, the projected series of transformations were tested on a model substrate, which was synthesized for this purpose as follows

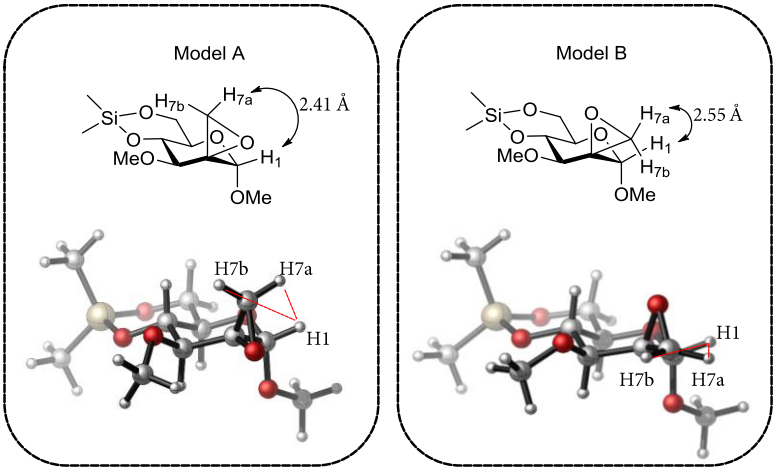
(Scheme 5). α-D-Methyl mannopyranoside was 4,6-silylidene protected and subsequently regioselectively benzylated at position 3 using organotin chemistry,[28] affording compound **11**. The remaining alcohol was oxidized under Swern conditions, yielding ketone **12**. The transformation of ketones into their corresponding spiro-epoxides can be accomplished by methylene insertion, using several reagents including diazomethane[29] and the classical Corey-Chaykovsky epoxidation utilizing (un)stabilized sulphur methylides.[30–32] The ratio of equatorially/axially inserted methylene could be influenced by the choice of methylenating agent. Under optimized conditions, equatorial methylene insertion was favoured using diazomethane in cold ethanol, yielding axial methylene **13** and equatorial methylene **14** in a combined quantitative yield in approximately 1:2 ratio, respectively. In contrast, axial insertion was favoured by employing dimethylsulfoxonium methylide at elevated temperatures, giving rise to **13** and **14** in approximately 3:1 ratio, respectively. Ultimately, axial methylene insertion was highly favoured using dimethylsulfonium methylide, resulting in the formation of **13** and **14** in approximately 9:1 ratio in moderate yield.

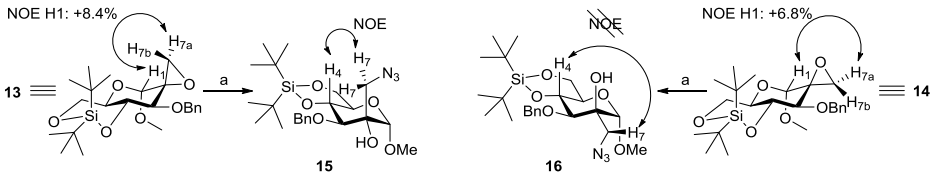


Scheme 5 Synthesis of test substrate **12** and its transformation into its corresponding spiro-epoxides. Reagents and conditions: a) $(^tBu)_2Si(OTf)_2$, 2,6-lutidine, DMF, rt, 30 min; b) Bu_2SnO, toluene, reflux, then BnBr, TBABr, CsF, 84% over 2 steps; c) DMSO, Ac_2O, rt, 61%; d) CH_2N_2, EtOH, 0°C, 10 min, **13**:**14** 1:1.7, 100%; e) $Me_3S(O)I$, *n*-BuLi, THF, 60 °C, 10 min, **13**:**14** 2.7:1, 62%; f) Me_3SI, NaH, DMSO, THF, -5 °C, 30 minutes, **13**:**14** 8.7:1, 52%.

The configuration of **13** and **14** was determined by NMR-NOE experiments. According to DFT calculations on model compounds A and B (Scheme 6), the H1-H7a distance in **13** (model A) is 2.41 Å and 2.55 Å in **14** (model B). Correspondingly, irradiation of H7a in **13** resulted in a NOE-enhancement of 8.4%, whereas irradiation of H7a in **14** resulted in a reduced enhancement of 6.8%. Since NOE-enhancement is linearly correlated to the internuclear distance,[33,34] it was concluded that the compound having the strongest NOE enhancement was **13**. To further validate this theory, spiro-

epoxides **13** and **14** were reacted with sodium azide at elevated temperatures to afford the corresponding azido-alcohols **15** and **16**. Indeed, when protons H7 were irradiated, a NOE enhancement was observed for H4 in **15**, while it was totally absent in **16**, further proving the absolute configuration of spiro-epoxides **13** and **14**.



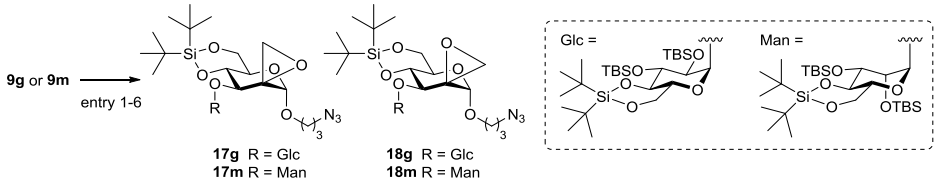


Scheme 6 Atomic distances for H1-H7a in model compounds A and B calculated by DFT (top). Alternative representations of **13** and **14**, and its corresponding NOE enhancements measured by 1D-NOEdif NMR experiments (bottom). After opening the spiro epoxides with sodium azide, a NOE correlation between H4-H7 was seen for **15**, but not for **16**. Reagents and conditions: a) NaN_3, NH_4Cl, H_2O, DMF, 65 °C, 16 h, yield **15**: 75%; yield **16**: 58%.

Having optimized the reaction conditions affording spiro-epoxides from ketone **12**, these conditions were applied to disaccharide analogues **9g** and **9m** (Table 1). Reaction of **9g** with diazomethane as methylenating agent resulted in the formation of the axial- (**17g**) and equatorial (**18g**) methylenes in a 1:1 ratio and in good yields (entry 1). Their absolute configuration was determined by 1D-NOEdiff measurements: upon irradiation of H7a, the signal for H1 was enhanced by 6.4% for **17g**, whereas

irradiation of H7a in **18g** gave an enhancement of the H1 signal of 5.2%. Reaction of **9m** with diazomethane also resulted in a mixture of **17m** and **18m**, in favor of the equatorial methylene group in almost quantitative yield (entry 2). In both cases, the isomers were difficult to separate by standard column chromatography, so other methods in favor of formation of axial methylenes were required. It was anticipated that the classical Corey-Chaykovsky epoxidation methodology using dimethylsulfoxonium methylide would favor the formation of the equatorial methylenes **18g** and **18m**. Indeed, also in these cases both isomers were obtained, however the formation of axial methylenes was still favored in both cases (entries 3-4). Finally, using dimethylsulfonium methylide, only the kinetically favored axial methylenes **17g** and **17m** were formed, albeit in moderate yields (entries 5-6).

Table 1 Transformation of ketones **9g** and **9m** into their corresponding spiro-epoxides under different reaction conditions. Diazomethane favors methylene insertion from the equatorial face of the ketone, while (un)stabilized methylides favor insertion from the axial position.

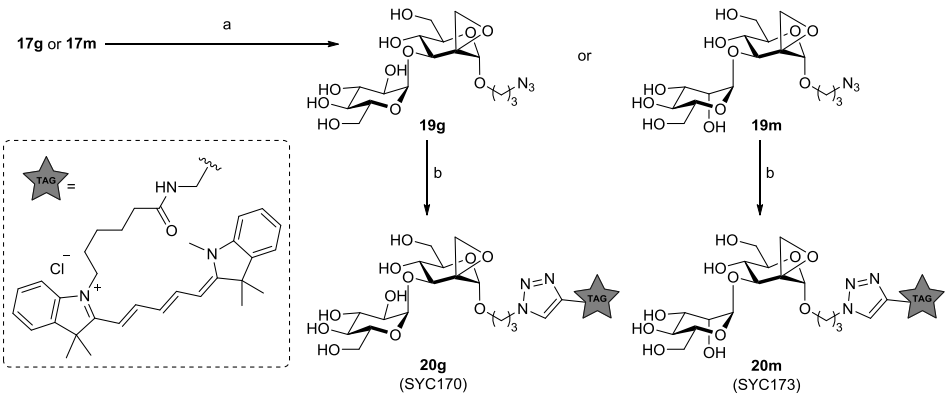


Entry	Starting material	Conditions	**17** : **18**	Yield[a]
1	**9g**	CH_2N_2, EtOH, 0 °C	1 : 1	78%
2	**9m**		1 : 3	97%
3	**9g**	$SOMe_3I$, *n*-BuLi, THF, 60 °C	5 : 1	83%
4	**9m**		2 : 1	85%
5	**9g**	SMe_3I, NaH, DMSO, THF, -10 °C	1 : 0	50%
6	**9m**		1 : 0	53%

[a] combined yield after column chromatography

With spiro-epoxides **17g** and **17m** in hand, global deprotection was accomplished by reaction with tetrabutylammonium fluoride (TBAF) in THF (Scheme 7). It was found that an excess (15 eq.) of reagent and prolonged reaction time (5 days) was needed to accomplish full deprotection. Due to the high polarity of the deprotected polyols as well as the excess of ammonium salts present in the reaction mixture, the purification

of **19g** and **19m** was found to be troublesome using standard column chromatography or size-exclusion. This problem was circumvented by quenching the reaction with potassium hexafluorophosphate, forming ammonium hexafluorophosphate salts which during work-up could be easily extracted into the organic phase whilst leaving the product in the water phase. Finally, a fluorescent Cy5 tag was installed at the azide handle using click-chemistry, which after HPLC purification afforded spiro-epoxyglycosides **20g** and **20m**.



Scheme 7 Global desilylation followed by click ligation of a fluorescent reporter tag affording the target compounds. Reagents and conditions: a) TBAF, THF, 5 days, yield **19g**: 97%; yield **19m**: 74%; b) Cy5-alkyne,[35] $CuSO_4 \cdot 5H_2O$, sodium ascorbate, DMF, rt, 16h, yield **20g**: 32%; yield **20m**: 34%.

Having fluorescent spiro-epoxyglycosides **20g** and **20m** in hand, their labeling efficiency towards recombinant *Bt*- and *Bx*GH99 *endo*-α-1,2-mannanase was evaluated (Figure 4A). The compounds label both enzymes in a concentration-dependent manner, with concentrations as low as 100 nM. Previous work by Hakki and co-workers indicated a preference of a mannosyl residue at the -2 subsite of the enzyme.[19] Interestingly, no significantly increased potency of **20m** over **20g** could be detected. Additionally, the pH dependence of labeling was determined (Figure 4B). Both spiro-epoxyglycosides label the enzymes at an optimum between pH 6-8, corresponding to the optimal pH for GH99 enzymatic activity.[16] Next, the rate of labeling was investigated (Figure 4C). Both *Bx* and *Bt* are labeled by spiro-epoxyglycosides **20g** and **20m** in a time-dependent manner, and substantial labeling is observed within 5 minutes of incubation.

In order to evaluate whether labeling of *Bt* and *Bx* is activity-based, labeling of wild-type *Bx*GH99 was compared to analogous active-site mutants (Figure 5A). While WT enzyme is labeled by spiro-epoxyglycosides **20g** and **20m** within 5 minutes, the catalytic base (here, the nucleophile) mutant E333Q nor catalytic acid mutant E336Q are labeled with these compounds, suggesting that labeling is indeed activity-based. However, prolonged incubation times do lead to labeling of the mutant enzymes, indicating that either the spiro-epoxide is susceptible to ring opening by the mutant catalytic residues, or other residues are available for binding. Nonetheless, denaturation of *Bt* and *Bx* totally abrogates labeling by spiro-epoxyglycosides **20g** and **20m**, clearly indicating that labeling requires native enzyme (Figure 5B).

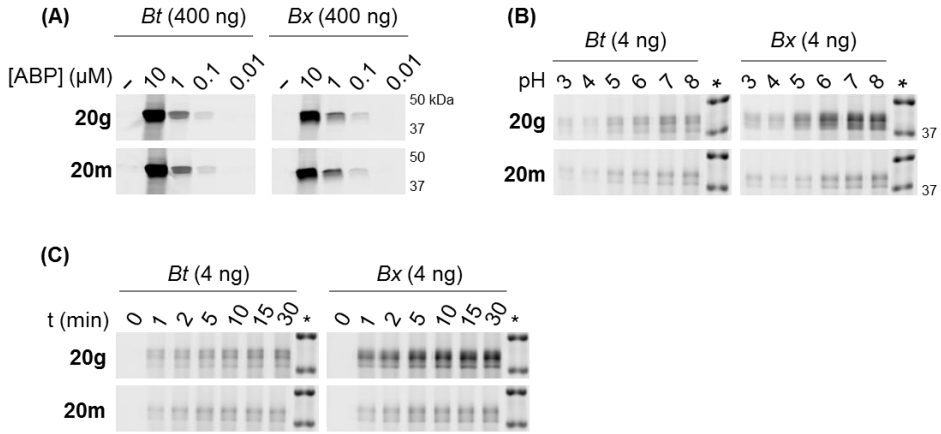


Figure 4 A) Detection limit of *Bt*- and *Bx*GH99 *endo*-α-1,2-mannanases, labeled with fluorescent spiro-epoxyglycosides **20g** or **20m**. B) Effect of pH on labeling of *Bt* and *Bx* enzymes. C) Time-dependent labeling of *Bt* and *Bx*.

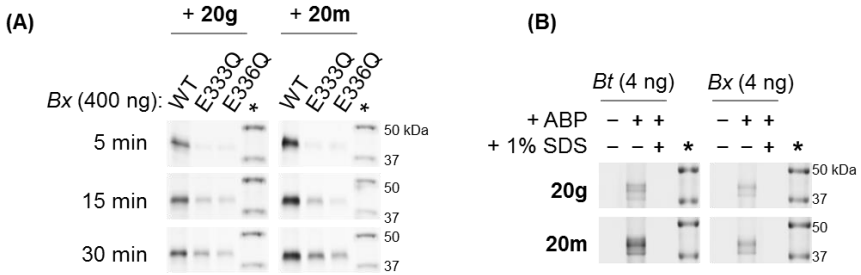


Figure 5 A) Time-dependent labeling of wild-type and mutant *Bx* enzymes with **20g** and **20m**. E333 is the (proposed) catalytic base, E336 is the (proposed) catalytic acid. B) Boiling of *Bt* and *Bx* with 1% SDS prior to incubation with **20g** or **20m** abrogates labeling.

To further evaluate whether the covalent inhibition of *Bt* and *Bx* is truly activity-based, the processing of α-galactosidase A (GLA) by these enzymes was investigated (Figure 6A). GLA contains three *N*-glycosylation sites, of which two are decorated with oligo-mannose structures, and one contains complex oligosaccharides low in mannose content.[36,37] In all experiments, GLA was pre-labeled with TB340[38] to ensure fluorescent detection on gel. Without additives, GLA gives a distinct major band at ~50 kDa (Figure 6B, lane 1). Incubation of GLA with *Bt*GH99 *endo*-α-1,2-mannanase

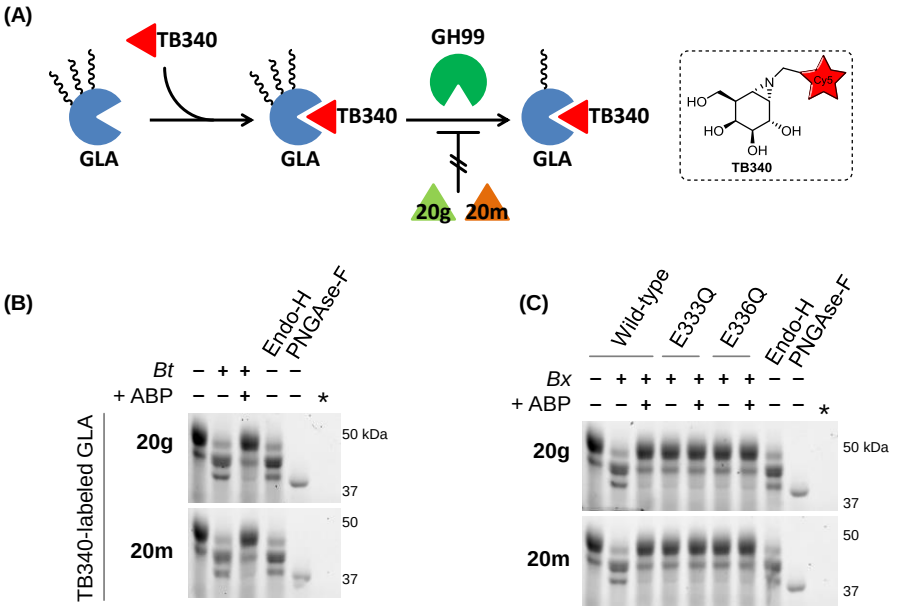


Figure 6 (A) Schematic representation of processing of human alpha-galactosidase GLA by GH99 endomannosidase. GLA is pre-labeled by fluorescent TB340, and contains high-mannose *N*-glycans which can be truncated by endomannosidase, resulting in a decrease in GLA molecular weight. Activity-based labeling of endomannosidase by spiro-epoxyglycosides **4** or **5** (prior to incubation with GLA) blocks its activity, and is therefore unable to process GLA.B) GLA is processed by *Bt endo*-α-1,2-mannanase, resulting in a shift in molecular weight (lane 2). Processing by *Bt* is abrogated by pre-incubation of *Bt* with **20g** or **20m** (lane 3). As control experiments, Endo-H also de-mannosylates (presumably the two high-mannose *N*-glycans of) GLA (lane 4). PNGase-F cleaves off all three glycans, resulting in a lower band migration (lane 5). C) Processing of TB340-labeled GLA by *Bx endo*-α-1,2-mannanase (lane 2). Pre-incubation of *Bx* with **20g** or **20m** abrogates GLA processing (lane 3). Active-site mutants are unable to process GLA (lanes 4-7).

results in demannosylation of the two high mannose *N*-glycans decorating the exterior of GLA, resulting in a shift of the GLA band into lower bands at ~42 kDa (lane 2). This shift in molecular weight positively corresponds to the shift observed when GLA is incubated with Endo-H (lane 4), a known demannosylating enzyme. Treatment of GLA with PNGase-F (lane 5), which fully cleaves *N*-glycans leaving Asn, results in the band migration to a lower molecular weight, probably caused by full deglycosylation of all three *N*-glycans. When *Bt* was pre-incubated with **20g** or **20m**, demannosylation of TB340-labeled GLA was not observed (lane 3), further indicating that binding of **20g** and **20m** occurs in the *Bt*GH99 active-site. An identical experiment was set up for *Bx* enzymes, including the E333Q and E336Q mutant enzymes (Figure 6C). Similar to *Bt*, WT *endo*-α-1,2-mannanase from *Bx* is able to process the *N*-glycans decorating the surface of the enzyme, giving rise to a shift in molecular weight (lane 2) which is similar to processing with Endo-H (lane 8). Pre-incubation of WT *Bx* by **20g** and **20m** prior to consecutive incubation with TB340-labeled GLA resulted in an observed absence of glycan processing (lane 3), indicating that binding of **20g** and **20m** abrogates enzymatic activity. Interestingly, while mutants E333Q and E336Q are labeled by spiro-epoxyglycosides **20g** and **20m** after prolonged reaction times, they are evidently unable to process TB340-labeled GLA (lanes 4-7). Again, labeling of *Bx* mutants at prolonged reaction times could be non-specific, however labeling of the catalytic residues in the active site mutants due to the high reactivity of the spiro-epoxide cannot be excluded.

Finally, the inhibitory potencies of **19g**, **19m**, ManIFG and yeast mannan were tested on *Bt*GH99, using spiro-epoxyglycoside **20m** as fluorescent read out (Figure 7). The enzyme was first pre-incubated with the competitor for 30 min at 37 °C, followed by labeling with 1 μM **20m** for 30 min at 37 °C. Compounds **19g** and **19m** both show a concentration-dependent competition of fluorescent labeling in the range of 1.000-10 μM, although full labeling competition could not be achieved under these conditions. Similarly, ManIFG gave concentration-dependent competition but full competition was not achieved. However pre-incubation by mannan, a substrate for *endo*-α-1,2-mannanase, achieved full competition of the signal, suggesting that processing of spiro-epoxyglycoside **20m** by *Bt*GH99 is specific and activity-based. A similar competition was performed for *Bx*GH99, and it was shown that while pre-incubation with **19g** did not fully abrogate labeling, pre-incubation with the highest concentration of **19m** did accomplish full competition, possibly hinting at a slight preference for a mannosyl residue in subsite -2. Additionally, yeast mannan from *S.*

cerevisiae (an α-1,6 linked mannose backbone branched with α-1,2 and α-1,3 mannoses[39]) showed concentration dependent (partial) competition, and ManIFG was able to fully compete the fluorescent signal at the highest concentration used, suggesting that processing of spiro-epoxyglycoside **20m** by *Bx*GH99 is specific and activity-based.

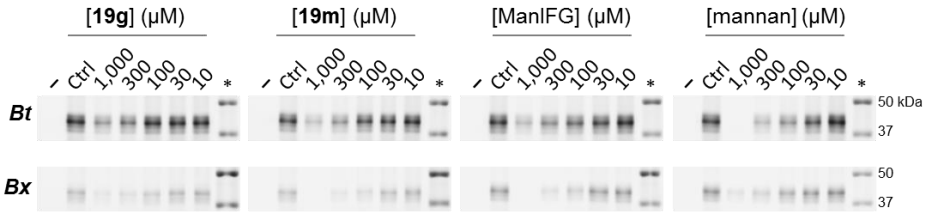


Figure 7 Competitive activity-based protein profiling. *Bt* (top) and *Bx* (bottom) enzymes were incubated with different concentrations of inhibitor followed by labeling with **20m**. For both enzymes, a slight preference for **19m** over **19g** is observed. ManIFG is able to fully compete labeling by **20m** in *Bx*, but not in *Bt*, whereas mannan fully competes labeling by **20m** in *Bt*, but not in *Bx*.

6.3 Conclusion

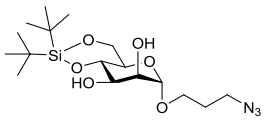
N-glycosylation is an important process in eukaryotic cells that regulates protein folding, targeting and function. The first steps in glycan maturation are performed by stepwise trimming of the terminal glucose residues from the *N*-linked branched oligosaccharide $Glc_3Man_9GlcNAc_3$-Asn by α-glucosidase I and II in the ER. GH99 *endo*-α-1,2-mannosidase is present in the Golgi apparatus and circumvents this pathway by directly cleaving $Glc_{1\text{-}3}Man$ oligosaccharides after which the trimmed *N*-glycan reenters the normal maturation pathway. Orthologues of this enzyme are also found in *Bacteroides thetaiotaomicron (Bt)* and *Bacteroides xylanisolvens (Bx)*, crucial members of the gut microbiota, and these bacterial enzymes are in fact *endo*-α-1,2-mannanases. The catalytic mechanism of these GH99 enzymes is unclear, however an unusual catalytic mechanism involving a 1,2-anhydro-epoxide intermediate was proposed by Thompson *et al.* To further study the enzyme role and function, and simultaneously investigate its reaction mechanism, compounds **20g** and **20m** bearing a spiro-epoxide at C2 as electrophilic trap have been synthesized. These compounds contain either a glucosyl or a mannosyl residue at the non-reducing end. While the former spiro-epoxyglycoside is anticipated to prefer binding by *endo*-α-1,2-mannosidases, the latter could prefer binding by *endo*-α-1,2-mannanases. Based on the proposed catalytic mechanism, the catalytic base is anticipated to covalently and irreversibly bind the spiro-epoxide by nucleophilic substitution. Using SDS-PAGE

labelling studies, these spiro-epoxyglycosides appeared to covalently label *Bx*- and *Bt*GH99 in a concentration- and time-dependent manner. Optimal labelling was achieved at pH 6-8, conform the enzymatic pH optimum for catalysis. Both *Bt* and *Bx* are labelled within 5 minutes of incubation, whilst *Bx* active-site mutants only show labelling after prolonged reaction times. Additionally, denaturing the enzyme prior to probe incubation completely abrogates labelling. Using a Fabrazyme processing assay, it was shown that labelling of *Bt* and *Bx* completely blocks the enzymatic *endo*-mannanase activity, strongly indicating that the spiro-epoxyglycosides bind in the active-site of the enzymes. Finally, using a competitive ABPP strategy it was shown that labelling of *Bt* and *Bx* by spiro-epoxyglycoside **20m** can be competed by the competitive inhibitor ManIFG and natural substrate mannan. Additionally, competition was achieved by the matching non-tagged spiro-epoxyglycosides. While *endo*-α-1,2-mannanase has a preference for mannosyl residues at the -2 binding subsite, the effect of either a mannosyl or glucosyl residue at the non-reducing end of the inhibitor on competition was only minimal. Altogether, these results suggest that the spiro-epoxyglycosides covalently bind the catalytic nucleophile of GH99 *endo*-α-1,2-mannanases, which could ultimately be confirmed by X-ray crystallography or MS-based techniques (further discussed in Chapter 8).

Experimental procedures

General: Chemicals were purchased from Acros, Sigma Aldrich, Biosolve, VWR, Fluka, Merck and Fisher Scientific and used as received unless stated otherwise. Tetrahydrofuran (THF), *N*,*N*-dimethylformamide (DMF) and toluene were stored over molecular sieves before use. Traces of water from reagents were removed by co-evaporation with toluene in reactions that required anhydrous conditions. All reactions were performed under an argon atmosphere unless stated otherwise. TLC analysis was conducted using Merck aluminum sheets (Silica gel 60 F_{254}) with detection by UV absorption (254 nm), by spraying with a solution of $(NH_4)_6Mo_7O_{24}{\cdot}4H_2O$ (25 g/L) and $(NH_4)_4Ce(SO_4)_4{\cdot}2H_2O$ (10 g/L) in 10% sulfuric acid or a solution of $KMnO_4$ (20 g/L) and K_2CO_3 (10 g/L) in water, followed by charring at ~150 ˚C. Column chromatography was performed using Screening Device b.v. silica gel (particle size of 40 – 63 μm, pore diameter of 60 Å) with the indicated eluents. For reversed-phase HPLC purifications an Agilent Technologies 1200 series instrument equipped with a semi-preparative column (Gemini C18, 250 x 10 mm, 5 μm particle size, Phenomenex) was used. LC/MS analysis was performed on a Surveyor HPLC system (Thermo Finnigan) equipped with a C_{18} column (Gemini, 4.6 mm x 50 mm, 5 μm particle size, Phenomenex), coupled to a LCQ Adventage Max (Thermo Finnigan) ion-trap spectrometer (ESI^+). The applied buffers were H_2O, MeCN and 1% aqueous TFA. ^1H NMR and ^{13}C NMR spectra were recorded on a Brüker AV-400 (400 and 101 MHz respectively) or a Brüker DMX-600 (600 and 151 MHz respectively) spectrometer in the given solvent. Chemical shifts are given in ppm (δ) relative to the residual solvent peak or tetramethylsilane (0 ppm) as internal standard. Coupling constants are given in Hz. High-resolution mass spectrometry (HRMS) analysis was performed with a LTQ Orbitrap mass spectrometer (Thermo Finnigan), equipped with an electronspray ion source in positive mode (source voltage 3.5 kV, sheath gas flow 10 mL/min, capillary temperature 250 °C) with resolution R = 60000 at m/z 400 (mass range m/z = 150 – 2000) and dioctyl phthalate (m/z = 391.28428) as a "lock mass". The high-resolution mass spectrometer was calibrated prior to measurements with a calibration mixture (Thermo Finnigan). Yeast mannan from *S. cerevisiae* was purchased from Sigma. ManIFG, as well as *B. thetaiotaomicron* (*Bt*) and *B. xylanisolvens* (*Bx*) enzymes were supplied by Prof. dr. Gideon Davies (University of York, UK). Recombinant α-galactosidase (Fabrazyme) was purchased from Genzyme (Cambridge, MA, USA). The α-galactosidase ABP TB340 was synthesized as described earlier.[38]

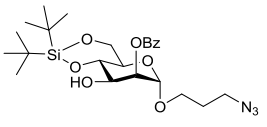
Compound S1



Compound **1**[40] (1.00 g, 3.80 mmol) was co-evaporated with dry toluene and dissolved in dry DMF (38 mL). The resulting solution was cooled to -50 °C and $Si^tBu_2(OTf)_2$ (1.11 mL, 3.42 mmol, 0.9 EQ) and 2,6-lutidine (0.44 mL, 3.80 mmol) were added. The reaction was stirred at -50 °C for 30 minutes and subsequently quenched with brine (400 mL). The aqueous layer was extracted with Et_2O (4x 100 mL). The combined organic layers were washed with 1M aqueous HCl (2 × 100 mL), H_2O (100 mL), and brine and dried over Na_2SO_4. The solvents were removed under reduced pressure and

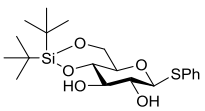
the crude product was purified by gradient column chromatography (EtOAc/pentane, 1:4 to 1:2). The product was obtained as white solid (970 mg, 70%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 4.81 (d, *J* = 1.4 Hz, 1H), 4.11 (dd, *J* = 10.0, 5.0 Hz, 1H), 4.07 – 4.00 (m, 2H), 3.96 (t, *J* = 10.2 Hz, 1H), 3.86 – 3.76 (m, 2H), 3.69 (td, *J* = 10.0, 5.0 Hz, 1H), 3.50 (ddd, *J* = 10.0, 6.3, 5.2 Hz, 1H), 3.40 (td, *J* = 6.6, 2.0 Hz, 2H), 1.94 – 1.82 (m, 2H), 1.06 (s, 9H), 1.00 (s, 9H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 166.0, 133.4, 129.9, 129.7, 128.5, 98.1, 75.2, 72.0, 70.2, 67.4, 66.6, 64.6, 48.2, 28.8, 27.4, 27.0, 22.8, 20.0. IR (neat): υ 3524, 2934, 2886, 2097, 1732, 1717, 1558, 1472, 1267, 1095, 1072, 1026, 885, 826, 710, 654. $[\alpha]_D^{20}$ (c 0.1, DCM): -16. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{24}H_{37}N_3O_7SiNa$ 530.22930, found 530.22907.

Compound 2



Compound **S1** (889 mg, 2.20 mmol) was dissolved in trimethyl orthobenzoate (5.7 mL) and CSA (102 mg, 0.44 mmol) was added. The reaction was stirred for 2 hours at room temperature and cooled to 0 °C. Aqueous AcOH (50%, 20 mL) was added and the mixture was stirred overnight while the cooling bath was allowed to reach room temperature. The solution was poured into saturated aqueous $NaHCO_3$ (50 mL) and the water layer was extracted with CH_2Cl_2 (3 × 50 mL). The combined organic layers were washed with $NaHCO_3$ (50 mL) and dried over $MgSO_4$. The solvents were removed under reduced pressure and the crude product was purified by gradient column chromatography (EtOAc/pentane, 1:99 to 1:10). The product was obtained as colorless oil (922 mg, 82%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.15 – 7.98 (m, 2H), 7.63 – 7.54 (m, 1H), 7.51 – 7.41 (m, 2H), 5.42 (dd, *J* = 3.4, 1.6 Hz, 1H), 4.88 (d, *J* = 1.4 Hz, 1H), 4.23 – 4.06 (m, 3H), 3.99 (t, *J* = 10.2 Hz, 1H), 3.86 – 3.76 (m, 2H), 3.59 – 3.49 (m, 1H), 3.43 (t, *J* = 6.6 Hz, 2H), 1.99 – 1.81 (m, 2H), 1.09 (s, 9H), 1.02 (s, 9H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 166.0, 133.4, 129.9, 129.7, 128.5, 98.1, 75.2, 72.0, 70.2, 67.4, 66.6, 64.6, 48.2, 28.8, 27.4, 27.0, 22.8, 20.0. IR (neat): υ 3524, 2934, 2886, 2097, 1732, 1717, 1558, 1472, 1267, 1095, 1072, 1026, 885, 826, 710, 654. $[\alpha]_D^{20}$ (c 0.1, DCM): -16. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{24}H_{37}N_3O_7SiNa$ 530.22930, found 530.22907.

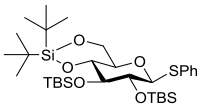
Compound S2



Compound **3**[23] (2.6 g, 9.5 mmol) was dissolved in dry DMF (100 mL) under Ar-atmosphere. The mixture was cooled to -50 °C and 2,6-lutidine (3.3 mL, 28.5 mmol) and $Si^tBu_2(OTf)_2$ (3.4 mL, 10.5 mmol) was added. The reaction was stirred for 2 hours at -50 °C and subsequently quenched with H_2O (100 mL). The water layer was extracted with EtOAc (3 × 100 mL). The organic layers were combined and washed with H_2O (2 × 200 mL) and brine (200 mL) and dried over $MgSO_4$. The solvents were removed under reduced pressure and the crude product was purified by gradient column chromatography (EtOAc/pentane, 1:4 to 1:2). The product was obtained as a white solid (3.58 g, 91%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.57 – 7.46 (m, 2H), 7.38 – 7.28 (m, 3H), 4.60 (d, *J* = 9.7 Hz, 1H), 4.21 (dd, *J* = 10.2, 5.1 Hz, 1H), 3.90 (t, *J* = 10.2 Hz, 1H), 3.68 (t, *J* = 9.0 Hz, 1H), 3.60 (t, *J* = 8.7 Hz, 1H), 3.51 – 3.37 (m, 2H), 2.92 (s, 1H), 2.77 (s, 1H), 1.04 (s, 9H), 0.98 (s, 9H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ

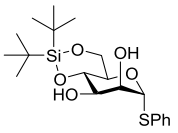
132.9, 131.7, 129.1, 128.3, 88.6, 77.8, 76.4, 74.5, 71.8, 66.1, 27.4, 27.0, 22.7, 19.9. IR (neat): υ 3241, 2932, 2858, 1695, 1471, 1058. $[\alpha]_D^{20}$ (c 0.06, DCM): -57.0. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{20}H_{32}O_5SSiNa$ 435.16319, found 435.16315.

Compound 4



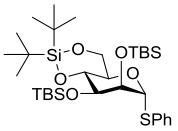
Compound **S2** (1.0 g, 2.42 mmol) was co-evaporated with toluene (3x), dissolved in dry pyridine (5 mL) and cooled to 0°C. DMAP (30 mg, 0.24 mmol) and TBSOTf (3.33 mL, 14.5 mmol) were added and the mixture was heated to 60°C and stirred overnight. The mixture was carefully diluted with water (25 mL) and extracted with DCM (3x 50 mL). The combined organic layers were washed with aq. 1M HCl (3x 25 mL) and brine, dried over Na_2SO_4, filtrated and concentrated. The crude product was purified by gradient column chromatography (pentane/EtOAc, 400:1 to 200:1), affording the product as a white solid (1.44 g, 93%). Analytical data was in accordance with those reported in literature.[41]

Compound S3



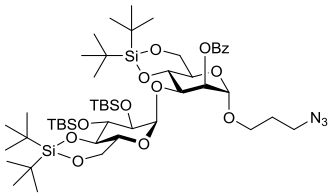
This compound was prepared from **5**[24] (4.9 g, 18 mmol) as described for the preparation of **S2** to afford the product (6.6 g, 89%) as a white solid. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.52 –7.20 (m, 5H), 5.53 (s, 1H), 4.30 (d, *J* = 3.1 Hz, 1H), 4.24 (td, *J* = 10.0, 5.0 Hz, 1H), 4.15 – 4.08 (t, *J* = 9.4 Hz, 1H), 4.05 (dd, *J* = 10.0, 5.0 Hz, 1H), 3.96 (t, *J* = 10.1 Hz, 1H), 3.87 (dd, *J* = 9.1, 3.3 Hz, 1H), 2.67 (brs, 2xOH), 1.05 (s, 9H), 1.03 (s, 9H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 133.9, 131.5, 129.3, 127.7, 87.8, 75.0, 72.4, 72.1, 67.9, 66.2, 27.6, 27.2, 22.8, 20.2. IR (neat): υ 3384, 2932, 2858, 1474, 1064. $[\alpha]_D^{20}$ (c 0.4, DCM): +227. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{20}H_{32}O_5SSiNa$ 435.16319, found 435.16301.

Compound 6



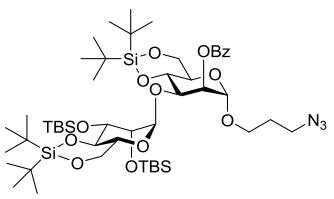
This compound was prepared from **S3** (6.6 g, 16 mmol) as described for the preparation of **4** to afford the product (9.3 g, 91%) as a pale yellow oil which crystallized at -20 °C. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.48 – 7.25 (m, 5H), 5.29 (d, *J* = 1.5 Hz, 1H), 4.28 (t, *J* = 9.0 Hz, 1H), 4.19 (m, 1H), 4.17 – 4.11 (m, 1H), 4.11 – 4.08 (m, 1H), 3.96 (t, *J* = 9.7 Hz, 1H), 3.87 (dd, *J* = 8.9, 2.5 Hz, 1H), 1.09 (s, 9H), 1.07 (s, 9H), 0.99 (s, 9H), 0.92 (s, 9H), 0.21 (s, 3H), 0.18 (s, 3H), 0.14 (s, 3H), 0.07 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 134.9, 131.3, 129.3, 127.4, 89.9, 75.0, 74.6, 73.0, 69.6, 67.1, 27.8, 27.3, 26.3, 25.8, 22.9, 20.2, 18.5, 18.2, -3.9, -4.1, -4.4, -4.4. IR (neat): υ 2931, 2857, 1471, 1250, 1096. $[\alpha]_D^{20}$ (c 1.0, DCM): +91. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{32}H_{61}O_5SSi_3$ 641.35420, found 641.36460.

Compound 7g



Compound **4** (2.00 g, 3.12 mmol) and compound **2** (1.58 g, 3.12 mmol) were combined and co-evaporated with toluene (3x). The mixture was dissolved in dry CH_2Cl_2 (20 mL) and stirred with activated 4A MS for 30 minutes at room temperature. The reaction was cooled to -50 °C and NIS (842 mg, 3.74 mmol) and TMSOTf (68 µL, 0.37 mmol) were added. The reaction mixture was warmed to -40 °C, stirred for 1 hour and subsequently neutralized with NEt_3 (2 mL). The mixture was diluted with CH_2Cl_2 (200 mL) and washed with saturated aqueous Na_2SO_3 (2 × 100 mL), H_2O (100 mL) and subsequently dried over $MgSO_4$. The solvents were removed under reduced pressure and the crude product was purified by gradient column chromatography (EtOAc/pentane, 1:50 to 1:40). The product was obtained as a white foam (2.98 g, 92%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.08 (d, *J* = 7.3 Hz, 2H), 7.59 (t, *J* = 7.4 Hz, 1H), 7.47 (t, *J* = 7.7 Hz, 2H), 5.49 (s, 1H), 5.16 (d, *J* = 2.9 Hz, 1H, H-1 'donor'), 4.86 (s, 1H, H-1 'acceptor'), 4.44 (t, *J* = 9.4 Hz, 1H), 4.22 – 4.15 (m, 2H, H-3), 4.04 (t, *J* = 10.2 Hz, 1H), 3.95 – 3.79 (m, 3H), 3.75-3.66 (m, 2H), 3.61 (t, *J* = 8.4 Hz, 1H), 3.54 (dt, *J* = 10.2, 5.7 Hz, 1H), 3.49 – 3.41 (m, 2H), 1.91 (dq, *J* = 13.5, 6.9 Hz, 1H), 1.13 (s, 9H), 1.04 (s, 9H), 1.03 (s, 9H), 0.97 (s, 9H), 0.91 (s, 9H), 0.79 (s, 9H), 0.17 (s, 3H), 0.05 (s, 3H), 0.03 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 165.4, 133.2, 129.9, 129.6, 128.5, 98.1, 97.9, 78.6, 75.2, 74.3, 73.4, 72.6, 71.1, 67.8, 67.8, 66.9, 66.6, 64.4, 48.1, 28.9, 27.5, 27.1, 27.0, 26.4, 26.2, 22.7, 22.7, 20.0, 20.0, 18.1, 18.0, -3.2, -3.5, -3.6, -4.4. ^{13}C-HMBC-GATED NMR (101 MHz, $CDCl_3$) δ 98.1 ($J_{C1,H1}$ = 170.6 Hz, C1 'donor'), 97.9 ($J_{C1,H1}$ = 172.1 Hz, C1 'acceptor'). IR (neat): υ 2966, 2859, 2093, 1732, 1472, 1260, 1096, 1069, 1045, 827. $[\alpha]_D^{20}$ (c 0.1, DCM): +20. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{50}H_{91}N_3O_{12}Si_4$+Na 1060.55720, found 1060.55694.

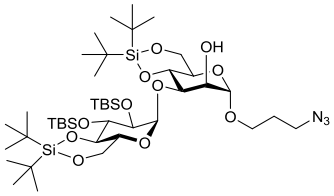
Compound 7m



This compound was prepared from **6** (378 mg, 0.59 mmol) and **2** (299 mg, 0.59 mmol) as described for the preparation of **7g**, to afford the product (538 mg, 88%) as a pale yellow oil. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.02 (m, 2H), 7.62 – 7.53 (m, 1H), 7.45 (t, *J* = 7.7 Hz, 2H), 5.34 (dd, *J* = 3.5, 1.6 Hz, 1H), 4.96 (d, *J* = 1.9 Hz, 1H), 4.82 (d, *J* = 1.4 Hz, 1H), 4.28 (t, *J* = 9.5 Hz, 1H), 4.15 (m, 3H), 4.08 (dd, *J* = 9.4, 3.6 Hz, 1H), 4.02 – 3.93 (t, *J* = 10.3 Hz, 1H), 3.90 – 3.77 (m, 4H), 3.77 – 3.67 (m, 2H), 3.57 – 3.37 (m, 3H), 2.02 – 1.77 (m, 2H), 1.09 (s, 9H), 1.04 – 1.00 (s, 9H), 1.00 (s, 9H), 0.95 – 0.89 (m, 9H), 0.87 – 0.82 (m, 9H), 0.82 – 0.76 (m, 9H), 0.07 (s, 3H), 0.00 (s, 3H), -0.13 (s, 3H), -0.17 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 165.5, 133.4, 130.0, 129.7, 128.6, 103.4, 98.2, 75.2, 75.1, 74.0, 73.6, 72.4, 71.8, 69.5, 67.7, 67.5, 67.1, 64.8, 48.4, 29.0, 27.9, 27.7, 27.2, 27.2, 26.2, 25.8, 22.9, 22.9, 20.1, 19.9, 18.4, 18.2, -4.3, -4.4, -4.4, -4.7. ^{13}C-HMBC-GATED NMR (101 MHz, $CDCl_3$) δ 103.4 ($J_{C1,H1}$ = 172.1 Hz, C1 'donor'), 98.2 ($J_{C1,H1}$ = 172.5 Hz, C1 'acceptor'). IR (neat): υ 2931, 2858, 20998, 1729, 1472, 1226, 1096, 1068.

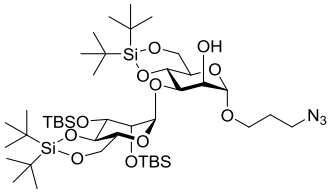
$[\alpha]_D^{20}$ (c 0.4, DCM): +1. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{50}H_{92}N_3O_{12}Si_4$ 1038.57526, found 1038.57587.

Compound 8g



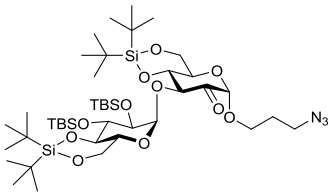
Compound **7g** (610 mg, 0.59 mmol) was co-evaporated with toluene (3x) and dissolved in a mixture of DCM/MeOH (9 mL, 1:1). NaOMe (30 wt%, 560 µL) was added and the reaction mixture was stirred for 24 h. The reaction was neutralized with AcOH and the solvents were removed under reduced pressure. The crude product was purified by gradient column chromatography (EtOAc/pentane, 1:11 to 1:8). The product was obtained as a white foam (519 mg, 95%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 5.34 (d, *J* = 3.1 Hz, 1H), 4.81 (d, *J* = 0.7 Hz, 1H), 4.29 (t, *J* = 9.3 Hz, 1H), 4.10 – 4.02 (m, 2H), 3.98 (t, *J* = 10.3 Hz, 1H), 3.95 (s, 1H), 3.88 (dd, *J* = 9.2, 3.3 Hz, 1H), 3.86 – 3.76 (m, 3H), 3.76 – 3.66 (m, 3H), 3.58 (dd, *J* = 8.2, 3.1 Hz, 1H), 3.54 – 3.47 (m, 1H), 3.38 (td, *J* = 6.5, 1.7 Hz, 2H), 3.00 (s, 1H, OH), 1.94 – 1.78 (m, 2H), 1.05 (s, 9H), 1.04 (s, 9H), 1.00 (s, 9H), 0.98 (s, 9H), 0.93 (s, 9H), 0.92 (s, 9H), 0.14 (s, 3H), 0.13 (s, 3H), 0.11 (s, 3H), 0.09 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 99.7, 97.4, 78.7, 75.1, 74.6, 74.5, 74.3, 71.1, 67.6, 67.4, 67.0, 66.4, 64.4, 48.4, 29.0, 27.6 (3x), 27.5 (3x), 27.2 (3x), 27.1 (3x), 26.4 (3x), 26.4 (3x), 22.9, 22.7, 20.1, 20.1, 18.3, 18.3, -3.1, -3.3, -3.4, -3.9. IR (neat): υ 2931, 2856, 2099, 1472, 1252, 1132, 1095, 1069, 1043, 868, 827, 772, 654. $[\alpha]_D^{20}$ (c 0.1, DCM): +44. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{43}H_{87}N_3O_{11}Si_4$+Na 956.53099, found 956.53097.

Compound 8m



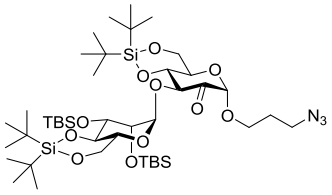
This compound was prepared from **7m** (501 mg, 0.48 mmol) as described for the preparation of **8g** to afford the product (386 mg, 86%) as a colorless oil. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 5.00 (d, *J* = 1.9 Hz, 1H), 4.79 (d, *J* = 1.1 Hz, 1H), 4.17 (t, *J* = 9.2 Hz, 1H), 4.11 (m, 3H), 3.99 – 3.88 (m, 4H), 3.88 – 3.82 (m, 2H), 3.82 – 3.77 (m, 1H), 3.77 – 3.62 (m, 2H), 3.49 (m, 1H), 3.40 (td, *J* = 6.5, 3.1 Hz, 1H), 2.37 – 2.03 (brs, OH), 1.97 – 1.77 (m, 2H), 1.04 (m, 18H), 0.99 (s, 9H), 0.97 (s, 9H), 0.93 (s, 9H), 0.86 (s, 9H), 0.12 (s, 3H), 0.11 (s, 3H), 0.10 (s, 3H), 0.02 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 103.2, 99.6, 77.7, 74.5, 74.1, 73.3, 72.4, 71.5, 69.6, 67.4, 67.33, 66.8, 64.5, 48.4, 28.9, 27.8, 27.6, 27.2, 27.1, 26.3, 25.8, 22.9, 22.7, 20.1, 18.6, 18.2, -3.9, -4.1, -4.3, -4.6. IR (neat): υ 2930, 2858, 2098, 1472, 1250, 1096, 1031. $[\alpha]_D^{20}$ (c 0.4, DCM): +32. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{43}H_{88}N_3O_{11}Si_4$ 934.54904, found 934.54959.

Compound 9g



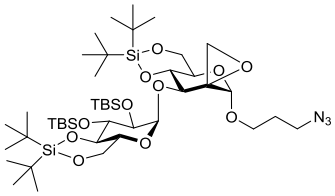
Compound **8g** (2.20 g, 2.36 mmol) was co-evaporated with dry toluene (3x) and dissolved in dry CH_2Cl_2 (65 mL). Dess-Martin periodinane (2.00 g, 4.71 mmol) was added and the mixture was stirred overnight. Celite was added and the solvents were removed under reduced pressure. The product was purified by gradient column chromatography (EtOAc/pentane, 1:70 to 1:4). The product was obtained as a white foam (2.15 g, 98%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 5.18 (d, *J* = 2.8 Hz, 1H), 4.73-4.72 (m, 2H), 4.24 – 4.02 (m, 5H), 4.02 – 3.90 (m, 1H), 3.90 – 3.73 (m, 3H), 3.67 (t, *J* = 8.6 Hz, 1H), 3.62 – 3.54 (m, 2H), 3.40 (t, *J* = 6.6 Hz, 2H), 1.94 – 1.81 (m, 2H), 1.06 (s, 9H), 1.04 (s, 9H), 1.02 (s, 9H), 0.98 (s, 9H), 0.93 (s, 9H), 0.92 (s, 9H), 0.17 (s, 3H), 0.12 (s, 3H), 0.11 (s, 3H), 0.08 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 196.2, 100.2, 98.2, 80.0, 78.9, 78.7, 74.9, 73.9, 67.49, 67.45, 67.0, 66.0, 65.3, 48.0, 27.5, 27.4 (3x), 27.1 (3x), 27.0 (3x), 26.5 (3x), 26.41 (3x), 22.7, 22.6, 20.0, 20.0, 18.3, 18.1, -3.0, -3.5, -3.7, -3.9. IR (neat): υ 2932, 2859, 2099, 1757, 1474, 1387, 1362, 1252, 1161, 1093, 1070, 1043, 866, 827, 775, 652. $[\alpha]_D^{20}$ (c 0.1, DCM): +50. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{43}H_{85}N_3O_{11}Si_4$+Na 954.51534, found 954.51535.

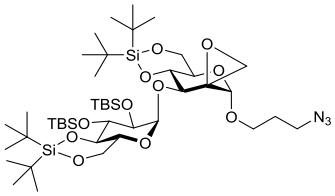
Compound 9m



This compound was prepared from **8m** (355 mg, 0.38 mmol) as described for the preparation of **9g** to afford the product (340 mg, 96%) as a yellow oil. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 4.81 (d, *J* = 1.9 Hz, 1H), 4.75 (s, 1H), 4.45 (d, *J* = 9.1 Hz, 1H), 4.31 – 4.19 (m, 2H), 4.18 – 4.12 (t, *J* = 9.2 Hz, 1H), 4.12 – 3.98 (m, 4H), 3.97 – 3.88 (m, 2H), 3.88 – 3.78 (m, 2H), 3.61 – 3.52 (dt, *J* = 9.9, 5.4 Hz, 1H), 3.41 (t, *J* = 6.5 Hz, 2H), 2.00 – 1.76 (m, 2H), 1.05 (s, 9H), 1.05 (s, 9H), 1.04 (s, 9H), 1.00 (s, 9H), 0.93 (s, 9H), 0.86 (s, 9H), 0.15 (s, 3H), 0.12 (s, 3H), 0.09 (s, 3H), 0.01 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 195.8, 103.4, 100.3, 82.2, 79.5, 74.4, 73.3, 72.1, 68.9, 67.3, 67.1, 66.6, 65.3, 48.0, 28.8, 27.6, 27.4, 27.1, 27.0, 26.2, 25.7, 22.8, 22.7, 20.1, 20.0, 18.4, 18.1, -4.0, -4.2, -4.5, -4.7. IR (neat): υ 2933, 2858, 2087, 1755, 1471, 1254, 1155. $[\alpha]_D^{20}$ (c 0.4, DCM): +47. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{43}H_{86}N_3O_{11}Si_4$ 932.53339, found 932.53363.

Compound 17g and 18g





Method A – *dimethyl sulfonium methylide*

A 1M solution of dimsyl-sodium was prepared from sodium hydride (60 wt%, 200 mg, 5 mmol) in dry DMSO (2.5 mL) and heating this mixture to 70°C for 1 h. The olive green solution was cooled to room

temperature and diluted with dry THF (2.5 mL). A fraction (0.12 mL, 0.19 mmol) of this mixture was added to a dried flask and cooled on an ice-salt bath. Then, a solution of trimethylsulfonium iodide (26.3 mg, 0.129 mmol) in dry DMSO (0.43 mL) and dry THF (0.4 mL) was added drop wise and the mixture was stirred for 5 minutes. Then, compound **9g** (100 mg, 0.11 mmol, co-evaporated with toluene (3x) beforehand) in dry THF (0.64 mL) was added and the mixture was stirred for 30 minutes. The mixture was diluted with water (20 mL) and extracted with Et_2O/pentane (2:1, 4x 15 mL). The combined organic layers were washed with water (20 mL), dried over Na_2SO_4, filtrated and concentrated. The crude product was purified by gradient column chromatography (pentane/EtOAc, 60:1 to 50:1) to afford solely product **17g** (51 mg, 50%) as an oil.

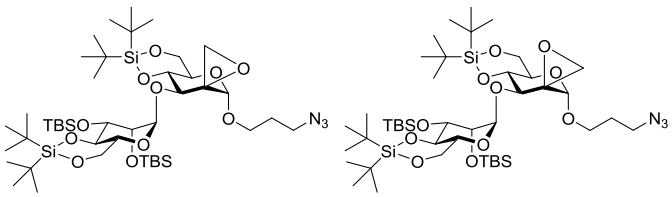
Method B – *dimethyl sulfoxonium methylide*
Trimethylsulfoxonium iodide (37.8 mg, 0.172 mmol) was suspended in dry THF (2 mL) and cooled to 0°C. *n*-Butyllithium (2 M in pentane, 80 µL, 0.16 mmol) was added and the mixture was heated to 60°C. Compound **9g** (100 mg, 0.11 mmol) was co-evaporated with toluene (3x), dissolved in dry THF (1 mL) and added drop wise to the ylide solution. After 10 minutes, the mixture was cooled to room temperature and quenched with MeOH (0.5 mL). The mixture was evaporated and the crude product was purified by gradient column chromatography (pentane/EtOAc, 60:1) to give a mixture of compounds **17g** and **18g** (90 mg, ratio **17g**:**18g** 5:1, total yield 88%) as a colorless oil.

Method C – *diazomethane*
To a glass tube were added aq. KOH (40%, 5 mL) and Et_2O (20 mL) and this mixture was cooled to 0°C. Then, 1-methyl-3-nitro-1-nitrosoguanidine (2.9 g, 10 mmol) was added in portions with swirling. A fraction (2 mL) of the bright yellow ether layer was added drop-wise to a solution of compound **9g** (100 mg, 0.11 mmol) in EtOH (3 mL) at 0°C. After stirring for 10 minutes, acetic acid (glacial) was added drop wise until the yellow mixture turned colorless. The mixture was concentrated and co-evaporated with toluene (3x). The crude products were purified by column chromatography (pentane/acetone, 150:1), affording compound **17g** and **18g** (79 mg, ratio **17g:18g** 1:1, total yield 78%).

Data for compound **17g** (axial methylene) ^{1}H-NMR (400 MHz, $CDCl_3$) δ 5.34 (d, *J* = 3.0 Hz, 1H), 4.40 – 4.22 (m, 2H), 4.09 (dd, *J* = 10.4, 5.1 Hz, 2H), 4.04 – 3.93 (m, 3H), 3.87-3.76 (m, 4H), 3.66 (t, *J* = 8.7 Hz, 1H), 3.56 – 3.39 (m, 4H), 3.23 (d, *J* = 5.6 Hz, 1H), 2.63 (d, J = 5.6 Hz, 1H), 1.89 (q, J = 5.7 Hz, 2H), 1.07 (s, 18H), 1.05 (s, 9H), 1.02 (s, 9H), 0.96 (s, 9H), 0.95 (s, 9H), 0.16 (s, 3H), 0.13 (s, 3H), 0.11 (s, 3H), 0.10 (s, 3H). 13C-NMR (101 MHz, CDCl3) δ 101.2, 96.9, 79.7, 78.2, 74.7, 73.8, 67.7, 67.4, 66.9, 66.8, 66.1, 64.4, 58.3, 48.1, 48.0, 29.0, 27.6 (3x), 27.3 (3x), 27.2 (3x), 27.0 (3x), 26.5 (3x), 26.4 (3x), 22.8, 22.5, 20.1, 20.0, 18.2, 18.1, -3.08, -3.38 (2x), -4.03. IR (neat): υ 2934, 2858, 2320, 2094, 1095, 1043, 827, 773. [α]D20 (c 0.1, DCM): +98 (c 0.1, DCM). HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{44}H_{88}N_3O_{11}Si_4$ 946.54904, found 946.54953. Data for compound **18g** (equatorial methylene): 1H-NMR (400 MHz, CDCl3) δ 5.33 (d, J = 2.9 Hz, 1H), 4.27 – 4.15 (m, 3H), 4.18 (s, 1H), 4.10 – 4.01 (m, 2H), 4.00 – 3.95 (t, J

= 10.1 Hz, 1H), 3.92 – 3.76 (m, 3H), 3.75 – 3.59 (m, 4H), 3.50 – 3.35 (m, 3H), 3.15 (d, J = 5.1 Hz, 1H), 2.61 (d, J = 5.1 Hz, 1H), 1.86 (quintet, J = 6.3 Hz, 2H), 1.04 (s, 9H), 1.04 (s, 9H), 0.98 (s, 18H), 0.91 (s, 9H), 0.90 (s, 9H), 0.13 (s, 3H), 0.10 (s, 3H), 0.09 (s, 3H), 0.08 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 102.5, 97.4, 79.3, 78.5, 76.0, 74.8, 68.7, 67.6, 67.1, 66.3, 64.2, 59.4, 48.3, 46.4, 29.9, 29.1, 27.6, 27.5, 27.2, 27.1, 26.3, 22.9, 22.7, 20.1, 18.4, 18.3, -3.6, -3.6, -4.0. IR (neat): υ 2930, 2858, 2099, 1472, 1252, 1091, 1043. $[\alpha]_D^{20}$ (c 0.1, DCM): +54. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{16}H_{27}N_3O_{11}Na$ 968.53099, found 968.53089.

Compound 17m and 18m



Compounds **17m** and **18m** were prepared from **9m** as described for the preparation of **17g** and **18g**, and could be separated by careful column chromatography.

Method A – *dimethyl sulfonium methylide*

Starting from compound **9m** (100 mg, 0.11 mmol), the product **17m** (54 mg, 53%) was obtained as the single product.

Method B – *Trimethylsulfoxonium iodide*

Starting from compound **9m** (150 mg, 0.16 mmol), product **17m** and **18m** were obtained as a mixture (129 mg, 85%, ratio **17m**:**18m** = 2:1)

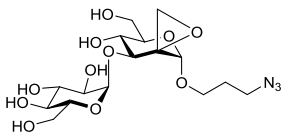
Method C – *Diazomethane*

Starting from compound **9m** (265 mg, 0.28 mmol), product **17m** and **18m** were obtained as a mixture (263 mg, 98%, ratio **17m**:**18m** = 1:2).

Data for compound **17m**: ^{1}H-NMR (400 MHz, $CDCl_3$) δ 4.94 (d, *J* = 2.0 Hz, 1H), 4.28 (s, 1H), 4.25 – 4.11 (m, 5H), 3.98 – 3.85 (m, 6H), 3.85 – 3.75 (m, 3H), 3.48 (m, 3H), 2.98 (d, *J* = 5.3 Hz, 1H), 2.64 (d, *J* = 5.3 Hz, 1H), 2.02 – 1.80 (m, 2H), 1.08 (s, 9H), 1.06 (s, 9H), 1.03 (s, 9H), 1.01 (s, 9H), 0.96 (s, 9H), 0.87 (s, 9H), 0.15 (s, 3H), 0.14 (s, 3H), 0.11 (s, 3H), 0.03 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 102.1, 101.3, 79.0, 74.4, 73.5, 72.6, 68.7, 67.4, 67.2, 66.9, 64.5, 58.3, 48.2, 47.6, 28.9, 27.7, 27.7, 27.2, 27.1, 26.3, 25.7, 22.8, 22.8, 20.2, 20.1, 18.7, 18.2, -4.0, -4.1, -4.3, -4.6. IR (neat): υ 2931, 2098, 1741, 1251, 1159, 1097. $[\alpha]_D^{20}$ (c 0.05, DCM): +44. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{44}H_{88}N_3O_{11}Si_4$ 946.54904, found 946.54933. Data for **18m**: ^{1}H-NMR (400 MHz, $CDCl_3$) δ 5.12 (d, *J* = 2.0 Hz, 1H), 4.28 (d, *J* = 9.4 Hz, 1H), 4.22 – 4.14 (m, 3H), 4.14 – 4.05 (m, 2H), 4.00 – 3.93 (t, *J* = 10.3 Hz, 1H), 3.93 – 3.83 (m, 3H), 3.83 – 3.77 (m, 2H), 3.55 – 3.43 (m, 2H), 3.40 (m, 2H), 3.08 (d, *J* = 4.7 Hz, 1H), 2.67 (d, *J* = 4.7 Hz, 1H), 1.94 – 1.82 (m, 2H), 1.05 (s, 9H), 1.04 (s, 9H), 1.00 (s, 9H), 0.99 (s, 9H), 0.94 – 0.90 (m, 9H), 0.84 (s 9H), 0.14

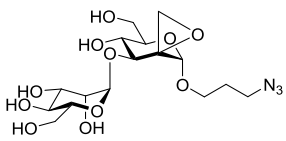
(s, 3H), 0.11 (s, 3H), 0.10 (s, 3H), 0.00 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 102.2, 101.9, 78.2, 74.4, 73.4, 72.1, 70.8, 69.7, 68.2, 67.3, 66.5, 64.2, 58.9, 48.2, 46.6, 28.8, 27.7, 27.5, 27.01, 26.2, 25.6, 22.7, 22.6, 20.0, 19.9, 18.4, 18.1, -4.0, -4.2, -4.2, -4.8. IR (neat): υ 2929, 2098, 1741, 1251, 1161, 1099. $[\alpha]_D^{20}$ (c 0.05, DCM): +38. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{44}H_{88}N_3O_{11}Si_4$ 946.54904, found 946.54940.

Compound 19g



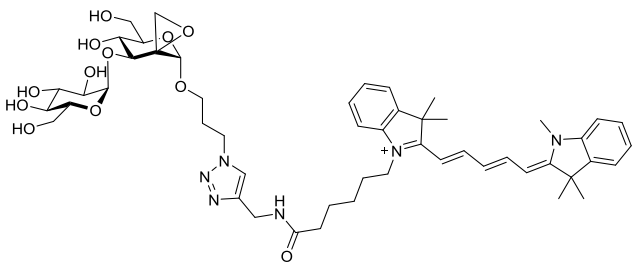
Compound **17g** (145 mg, 0.153 mmol) was co-evaporated with toluene (3x) and dissolved in dry THF (14.5 mL). TBAF (1 M in THF, 2.3 mL, 2.3 mmol) was added and the mixture was stirred overnight at room temperature. The solution was eluted with THF over a small Dowex-50WX4-200-Na^+ packed column, concentrated and purified by gradient column chromatography (EtOAc/MeOH, 19:1 to 9:1). The product was dissolved in water and lyophilized to afford the title compound as a white solid (64.8 mg, 97%). ^{1}H-NMR (400 MHz, D_2O) δ 5.23 (d, *J* = 3.8 Hz, 1H), 4.50 (s, 1H), 4.25 (d, *J* = 9.0 Hz, 1H), 3.95 – 3.68 (m, 8H), 3.63 – 3.53 (m, 2H), 3.53 – 3.43 (m, 3H), 3.41 – 3.34 (t, *J* = 8 Hz, 1H), 3.17 (d, *J* = 4.5 Hz, 1H), 2.87 (d, *J* = 4.6 Hz, 1H), 1.97-1.85 (m, 2H). ^{13}C-NMR (101 MHz, D_2O) δ 100.0, 99.2, 73.0, 72.9, 72.5, 71.7, 71.5, 70.91, 69.1, 64.8, 60.3, 60.2, 58.7, 48.4, 48.1, 27.8. IR (neat): υ 3369, 2927, 2108, 1521, 1026. $[\alpha]_D^{20}$ (c 0.1, DCM): +174. HRMS (ESI) m/z: $[M+NH_4]^+$ calc for $C_{16}H_{31}N_4O_{11}$ 455.19838, found 455.19849.

Compound 19m



This compound was prepared from **17m** (59 mg, 0.623 mmol) as described for the preparation of **19g** to afford the product (20 mg, 74%) as a white solid. ^{1}H-NMR (400 MHz, D_2O) δ 5.05 (d, *J* = 1.6 Hz, 1H), 4.45 (s, 1H), 4.18 (d, *J* = 9.3 Hz, 1H), 3.95 (dd, *J* = 3.2, 1.8 Hz, 1H), 3.77 (m, 6H), 3.64 (m, 3H), 3.57 – 3.48 (m, 1H), 3.41 (t, *J* = 6.5 Hz, 2H), 3.10 (d, *J* = 4.5 Hz, 1H), 2.81 (d, *J* = 4.5 Hz, 1H), 1.95 – 1.78 (m, 2H). ^{13}C-NMR (101 MHz, D_2O) δ 101.0, 100.0, 73.3, 72.8, 72.5, 70.9, 70.4, 69.9, 66.3, 64.7, 60.8, 60.2, 58.7, 48.3, 48.0, 27.8. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{16}H_{27}N_3O_{11}$ 460.1538, found 460.1544.

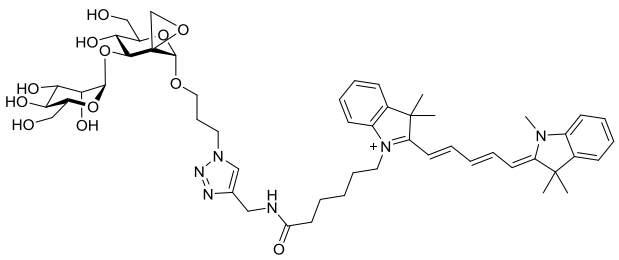
Compound 20g (SYC170)



Compound **19g** (4.83 mg, 11.0 µmol) was dissolved in DMF (0.5 mL) and placed under Argon. Then the Cy5-alkyne[35] (6.1 mg, 11.0 µmol), aq. $CuSO_4$ (0.1 M, 44 µL, 4.4 µmol) and aq. sodium ascorbate (0.1 M, 44 µL, 4.4 µmol) were added and the mixture was stirred overnight at room temperature. The product was purified by HPLC (NH_4CO_3) to

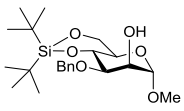
afford the title compound as a blue solid (3.54 mg, 32%). ^{1}H-NMR (400 MHz, MeOD) δ 8.24 (t, *J* = 13.0 Hz, 2H), 7.89 (s, 1H), 7.49 (d, *J* = 7.4 Hz, 2H), 7.44 – 7.38 (m, 2H), 7.32 – 7.23 (m, 4H), 6.62 (t, *J* = 12.4 Hz, 1H), 6.28 (d, *J* = 13.7 Hz, 2H), 5.14 (d, *J* = 3.8 Hz, 1H), 4.85 (s, 1H), 4.53 (t, *J* = 6.8 Hz, 2H), 4.42 (s, 2H), 4.32 (s, 1H), 4.16 (d, *J* = 9.1 Hz, 1H), 4.10 (t, *J* = 7.4 Hz, 2H), 3.84 – 3.64 (m, 9H), 3.63 (s, 3H), 3.56 (t, *J* = 9.3 Hz, 1H), 3.40 (dd, *J* = 9.7, 3.8 Hz, 1H), 3.37 – 3.32 (1, 9H), 3.07 (d, *J* = 5.3 Hz, 1H), 2.70 (d, *J* = 5.4 Hz, 1H), 2.25 (t, *J* = 7.3 Hz, 2H), 2.23 – 2.15 (m, 2H), 1.88 – 1.76 (m, 2H), 1.75 – 1.67 (m, 17H), 1.51 – 1.44 (m, 2H). ^{13}C-NMR (101 MHz, MeOD) δ 180.3, 175.7, 175.4, 174.7, 155.5, 155.5, 146.1, 144.3, 143.6, 142.6, 142.5, 129.8, 129.7, 126.6, 126.3, 126.2, 124.7, 123.4, 123.3, 112.0, 111.9, 104.4, 104.3, 102.4, 101.8, 76.8, 75.1, 74.4, 74.0, 73.9, 73.0, 71.2, 65.3, 62.4, 62.4, 59.8, 50.6, 50.5, 48.4, 44.8, 36.5, 35.6, 31.5, 31.0, 28.1, 27.9, 27.8, 27.3, 26.4. HRMS (ESI) m/z: $[M]^+$ calc for $C_{51}H_{69}N_6O_{12}$ 957.4968, found 957.5005.

Compound 20m (SYC173)



This compound was prepared from **19m** (3.72 mg, 8.5 µmol) as described for the preparation of **20g** to afford the product (2.9 mg, 34%) as a blue solid. ^{1}H-NMR (600 MHz, MeOD) δ 8.24 (t, *J* = 13.0 Hz, 2H), 7.90 (s, 1H), 7.49 (d, *J* = 7.4 Hz, 2H), 7.44 – 7.39 (m, 2H), 7.28 (dt, *J* = 16.4, 7.6 Hz, 4H), 6.62 (t, *J* = 12.4 Hz, 1H), 6.28 (d, *J* = 13.7 Hz, 2H), 5.19 (d, *J* = 1.3 Hz, 1H), 4.85 (s, 1H), 4.54 (t, *J* = 6.8 Hz, 2H), 4.42 (s, 2H), 4.29 (s, 1H), 4.23 (d, *J* = 9.2 Hz, 1H), 4.10 (t, *J* = 7.4 Hz, 3H), 3.90 (dd, *J* = 3.2, 1.7 Hz, 1H), 3.85 – 3.68 (m, 9H), 3.66 (d, *J* = 9.5 Hz, 1H), 3.63 (s, 3H), 3.60 (dd, *J* = 9.5, 3.3 Hz, 1H), 3.57 – 3.52 (m, 1H), 3.36 – 3.32 (m, 1H), 3.01 (d, *J* = 5.3 Hz, 1H), 2.68 (d, *J* = 5.4 Hz, 1H), 2.26 (t, *J* = 7.3 Hz, 2H), 2.20 (dq, *J* = 13.1, 6.7 Hz, 2H), 1.83 (m, 2H), 1.73 (s, 17H), 1.47 (m, 2H). ^{13}C-NMR (150 MHz, MeOD) δ 180.3, 175.7, 175.4, 174.7, 155.5, 155.5, 146.1, 144.3, 143.6, 142.6, 142.5, 129.8, 129.7, 126.6, 126.3, 126.2, 124.8, 123.4, 123.3, 112.1, 111.8, 104.4, 104.3, 102.6, 102.4, 74.7, 74.3, 74.1, 73.3, 72.7, 72.1, 68.2, 65.2, 62.7, 62.3, 59.9, 50.6, 50.5, 48.3, 44.8, 36.5, 35.7, 31.5, 31.1, 31.0, 28.1, 28.0, 27.8, 27.3, 26.4. HRMS (ESI) m/z: $[M]^+$ calc for $C_{51}H_{69}N_6O_{12}$ 957.4968, found 957.4995.

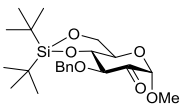
Compound 11



Methyl α-D-mannopyranoside **10** (6.1 g, 31.4 mmol) was dissolved in dry DMF (250 mL) and cooled to -50°C. Then, $^{t}Bu_2Si(OTf)_2$ (10.0 mL, 30.8 mmol, 0.98 EQ) and 2,6-lutidine (11.0 mL, 94 mmol) were added and the mixture was stirred for 30 minutes. Water (250 mL) was added and the mixture was extracted with EtOAc (3x 200 mL). The combined organic layers were washed with aq. 1M HCl (2x 100 mL), water (100 mL) and brine. The organic layer was dried over Na_2SO_4, filtrated and concentrated. The crude was co-evaporated with toluene (3x) and dissolved in toluene (400 mL). Dibutyltin oxide (7.6 g, 30.5

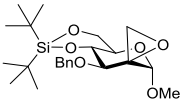
mmol) was added and the mixture was refluxed overnight under argon. The mixture was cooled to room temperature, and tetrabutylammonium bromide (10.3 g, 32.0 mmol), cesium fluoride (4.7 g, 31.1 mmol) and benzyl bromide (3.8 mL, 31.7 mmol) were added. The mixture was refluxed for 3 h and subsequently quenched with aq. $NaHCO_3$ (300 mL) and extracted with EtOAc (3x 200 mL). The combined organic layers were washed with water, dried over Na_2SO_4, filtrated and concentrated. The crude product was purified by gradient column chromatography (pentane/EtOAc, 10:1 to 6:1) to afford the title compound as a colorless oil (10.9 g, 84%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.42 – 7.27 (m, 5H), 4.93 (d, *J* = 12.0 Hz, 1H), 4.78 (d, *J* = 12.0 Hz, 1H), 4.69 (d, *J* = 1.3 Hz, 1H), 4.25 (t, *J* = 9.3 Hz, 1H), 4.13 – 4.09 (m, 1H), 4.03 – 3.95 (m, 2H), 3.73 – 3.66 (td, *J* = 10.0, 5.2 Hz, 1H), 3.64 (dd, *J* = 9.0, 3.5 Hz, 1H), 3.37 (s, 3H), 2.62 (brs, OH), 1.09 (s, 9H), 1.01 (s, 9H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 138.6, 128.5, 127.9, 127.8, 100.9, 78.4, 75.1, 73.5, 70.2, 67.0, 66.9, 55.3, 27.6, 27.2, 22.8, 20.1. IR (neat): υ 3474, 2931, 2858, 1471, 1094, 1066. $[\alpha]_D^{20}$ (c 0.2, DCM): + 39. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{22}H_{36}O_6SiNa$ 447.21734, found 447.21710.

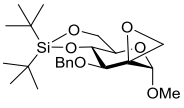
Compound 12



Compound **11** (9.54 g, 22.47 mmol) was co-evaporated with toluene (3x), dissolved in dry DMSO (121 mL) and cooled to 10°C. Acetic anhydride (38 mL, 404 mmol) was added and the mixture was stirred overnight at room temperature. The mixture was quenched with sat. aq. $NaHCO_3$ (400 mL) and water (500 mL) and stirred for 30 minutes. The mixture was extracted with Et_2O (4x 200 mL) and the combined organic layers were washed with sat. aq. $NaHCO_3$ (2x 200 mL), water (100 mL) and brine. The organic layer was dried over Na_2SO_4, filtrated and concentrated. The crude product was purified by gradient column chromatography (pentane/EtOAc, 40:1 to 4:1), affording the title compound as a white solid (5.84 g, 61%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.43 (m, 2H), 7.29 (m, 3H), 4.97 (d, *J* = 12.7 Hz, 1H), 4.76 (d, *J* = 12.7 Hz, 1H), 4.64 (s, 1H), 4.31 (d, *J* = 9.3 Hz, 1H), 4.20 (dd, *J* = 9.8, 4.5 Hz, 1H), 4.13 (t, *J* = 9.3 Hz, 1H), 4.05 (td, *J* = 9.6, 4.5 Hz, 1H), 3.95 (t, *J* = 9.9 Hz, 1H), 3.43 (s, 3H), 1.08 (s, 9H), 1.02 (s, 9H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 197.8, 138.0, 128.4, 127.7, 127.5, 101.4, 82.9, 79.6, 73.4, 67.1, 66.6, 55.9, 27.5, 27.1, 22.81, 20.0. IR (neat): υ 1750, 1472, 1362, 1158, 1094. $[\alpha]_D^{20}$ (c 0.5, DCM): +5. HRMS (ESI) m/z: $[M+NH_4]^+$ calc for $C_{22}H_{38}O_6NSi$ 440.24629, found 440.24532.

Compound 13 and 14





Method A – *dimethylsulfonium methylide*

Compound **12** (317 mg, 0.75 mmol) was subjected to the same reaction conditions as for the preparation of compound **17g**, and compounds **13** (151 mg, 46%) and **14** (19 mg, 6%) were obtained.

Method B – *dimethylsulfoxonium methylide*

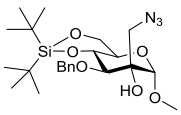
Compound **12** (100 mg, 0.24 mmol) was subjected to the same reaction conditions as for the preparation of compound **17g**, and a mixture of compounds **13** and **14** (64 mg, ratio **13**:**14** 2.7:1, 62%) was obtained.

Method C - *diazomethane*

Compound **12** (100 mg, 0.24 mmol) was subjected to the same reaction conditions as for the preparation of compound **17g**, and a mixture of compounds **13** and **14** (103 mg, ratio **13**:**14** 1:1.66, 100%) was obtained.

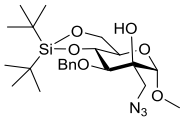
Data for compound **13** (axial methylene): ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.43 – 7.27 (m, 5H), 4.76 (d, *J* = 11.9 Hz, 1H), 4.74 – 4.69 (d, *J* = 11.9 Hz, 1H), 4.22 – 4.19 (s, 1H), 4.14 (m, 1H), 4.05 – 3.98 (m, 1H), 3.98 – 3.85 (m, 3H), 3.41 (s, 3H), 3.24 (d, *J* = 5.4 Hz, 1H), 2.65 (d, *J* = 5.4 Hz, 1H), 1.08 (s, 9H), 1.02 (s, 9H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 139.0, 128.3, 127.7, 127.5, 102.4, 78.7, 76.4, 75.2, 67.3, 66.9, 59.3, 55.5, 48.4, 27.6, 27.2, 22.9, 20.1. IR (neat): υ 2931, 2858, 1471, 1362, 1101, 1050. $[\alpha]_D^{20}$ (c 0.2, DCM): +35. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{24}H_{36}O_6Si$ 437.23539, found 437.23545. Data for compound **14** (equatorial methylene): ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.37 – 7.27 (m, 5H), 4.96 (d, *J* = 11.6 Hz, 1H), 4.72 (d, *J* = 11.7 Hz, 1H), 4.22 (t, *J* = 9.3 Hz, 1H), 4.14 (dd, *J* = 10.1, 4.9 Hz, 1H), 4.09 (s, 1H), 4.05 – 3.98 (m, 2H), 3.87 (td, *J* = 9.9, 4.9 Hz, 1H), 3.36 (s, 3H), 3.11 (d, *J* = 4.8 Hz, 1H), 2.67 (d, *J* = 4.8 Hz, 1H), 1.09 (s, 9H), 1.04 (s, 9H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 138.3, 128.4, 128.1, 127.7, 103.3, 78.3, 74.9, 74.3, 67.9, 66.7, 58.7, 55.0, 46.4, 29.7, 27.5, 27.1, 22.7, 19.9. IR (neat): υ 2931, 2857, 1471, 1362, 1089, 1047. $[\alpha]_D^{20}$ (c 0.2, DCM): +47. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{23}H_{36}O_6SiNa$ 459.21734, found 459.21691.

Compound 15



Compound **13** (50 mg, 0.12 mmol) was dissolved in DMF (1.0 mL). Water (164 µL), ammonium chloride (49 mg, 0.92 mmol) and sodium azide (22 mg, 0.34 mmol) were added and the mixture was heated to 65°C and stirred overnight. Water (25 mL) was added and the mixture was extracted with EtOAc (4x 10 mL). The combined organic layers were washed with water (3x 15 mL), brine and dried over Na_2SO_4, filtrated and concentrated. The crude product was purified by column chromatography (pentane/EtOAc, 50:1), affording the title compound as an oil (41 mg, 75%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.45 – 7.26 (m, 5H), 4.97 – 4.91 (m, 1H), 4.89 – 4.82 (m, 1H), 4.74 (s, 1H), 4.14 – 4.09 (dd, *J* = 10.0, 4.7 Hz, 1H), 3.92 – 3.86 (m, 2H), 3.81 – 3.77 (m, 1H), 3.77 – 3.71 (m, 1H), 3.66 (d, *J* = 9.0 Hz, 1H), 3.46 (s, 3H), 3.23 – 3.15 (d, *J* = 13.2 Hz, 1H), 2.61 – 2.56 (s, OH), 1.06 (s, 9H), 1.01 (s, 9H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 138.9, 128.4, 127.9, 127.7, 99.9, 83.7, 76.7, 76.6, 75.9, 66.9, 66.7, 56.0, 52.1, 27.6, 27.2, 22.8, 20.1. IR (neat): υ 2931, 2857, 2106, 1471, 1069. $[\alpha]_D^{20}$ (c 0.5, DCM): +27. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{24}H_{37}N_3O_6Si$ 480.25244, found 480.25196.

Compound 16



Compound **14** (19 mg, 44 µmol) was co-evaporated with toluene and dissolved in DMF (0.87 mL). Sodium azide (17 mg, 0.26 mmol) was added and the mixture was heated to 100°C and stirred overnight. The mixture was cooled to room temperature, diluted with water (10 mL) and extracted with EtOAc (4x 5 mL). The combined organic layers were washed with water (3x 10 mL), brine and dried over Na2SO4, filtrated and concentrated. The crude product was purified by gradient column chromatography (pentane/EtOAc, 50:1 to 30:1), affording the title compound as an oil (12 mg, 58%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.43 – 7.29 (m, 5H), 5.05 (d, *J* = 11.1, 1H), 4.71 (d, *J* = 11.1, 1H), 4.60 (s, 1H), 4.27 (t, *J* = 9.2, 1H), 4.13 (dd, *J* = 10.1, 5.0, 1H), 4.00 (t, *J* = 10.3, 1H), 3.74 (td, *J* = 10.1, 5.0, 1H), 3.52 (s, 2H), 3.49 (d, *J* = 3.1, 1H), 3.40 (s, 3H), 3.24 – 3.20 (d, *J* = 12.7 Hz, 1H), 2.51 (s, OH), 1.10 – 1.07 (s, 9H), 1.03 (s, 9H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 138.0, 128.6, 128.6, 128.3, 101.0, 79.2, 76.7, 76.4, 75.5, 66.8, 66.7, 55.9, 55.6, 27.6, 27.2, 22.8, 20.1. IR (neat): υ 2933, 2857, 2089, 1471, 1048. $[\alpha]_D^{20}$ (c 0.2, DCM): +52. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{23}H_{37}N_3O_6SiNa$ 502.23438, found 502.23428.

Labeling of *Bt* and *Bx* enzymes. To determine the detection limit, 400 ng recombinant *B. thetaiotaomicron* (*Bt*) and *B. xylanisolvens* (*Bx*) were labeled in 150 mM McIlvaine buffer, pH 7.0 (citric acid–Na_2HPO_4) with 0.0001–10 µM spiro-epoxyglycoside **20g** or **20m** for 1 h at 37 °C. The samples were then denatured with 5× Laemmli buffer (50% (v/v) 1 M Tris–HCl, pH 6.8, 50% (v/v) 100% (v/v) glycerol, 10% (w/v) DTT, 10% (w/v) SDS, 0.01% (w/v) bromophenol blue), boiled for 4 min at 100 °C, and separated by electrophoresis on 10% (w/v) SDS-PAGE gel running continuously at 90 V.[42] Wet slab-gels were scanned on fluorescence using a Typhoon FLA 9500 (GE Healthcare at λ_{EX} 532 nm and λ_{EM} 575 nm for ABP TB340; and at λ_{EX} 635 nm and λ_{EM} 665 nm for **20g** and **20m**. The pH optimum was analyzed using 4 ng enzyme incubated with 1 µM **20g** and **20m** dissolved in McIlvaine buffer, pH 3-8, for 30 min at 37 °C. The time-dependent labeling kinetics were assessed identically for wild-type *Bx* and *Bt* enzyme, with the incubation stopped after 0, 1, 2, 5, 10, 15 or 30 min by denaturation with Laemmli buffer. Time-dependent labeling of *Bx* wild-type, E333Q and E336Q enzymes was assessed by incubating 400 ng for 5, 15, 30 or 60 min with 1 µM **20g** and **20m** dissolved in McIlvaine buffer, pH 7. The effect of denaturation was assessed on 4 ng wild-type *Bt* and *Bx* by boiling for 4 min at 100 °C prior to incubating with 1 µM **20g** and **20m** for 30 min at 37 °C. Activity-based protein profiling utilized 4 ng *Bt* and *Bx* enzyme that was pre-incubated with 10-1,000 µM **19g**, **19m** or ManIFG, or 0.3-30 µg/µL mannan, at pH 7.0 for 30 min at 37 °C, followed by labeling with 1 µM **20g** and **20m** for 30 min at 37 °C.

Functional Fabrazyme assay. Recombinant α-galactosidase was diluted 1:2 in 50 mM McIlvaine buffer, pH 4.6, and pre-labeled with 2 µM TB340 for 1 h at 37 °C. Subsequently, the mixture was diluted to 1:500 in 150 mM McIlvaine buffer, pH 7.0. In parallel, 400 ng *Bx* wild-type, E333Q and E336Q were incubated in the presence or absence of 10 µM **20g** or **20m**, dissolved in 150 mM McIlvaine buffer, pH 7.0, for 1 h at 37 °C. Subsequently, the *Bx* mixture (10 µL) was incubated with 10

µL TB340-labeled Fabrazyme for 8 h at 37 °C. Hereafter, samples were denatured, separated on SDS-PAGE gel and visualized by fluorescence scanning, as described above (vide supra). As control, 10 µL TB340-labeled Fabrazyme was treated by either Endo-H or PNGase-F, following the manufacturer's instructions (New England Biolabs).

References

1 A. Helenius and M. Aebi, *Science*, 2001, **291**, 2364–2370.
2 M. Molinari, *Nat. Chem. Biol.*, 2007, **3**, 313–320.
3 T. Feizi and M. Larkin, *Glycobiology*, 1990, **1**, 17–23.
4 Y.-Y. Zhao, M. Takahashi, J.-G. Gu, E. Miyoshi, A. Matsumoto, S. Kitazume and N. Taniguchi, *Cancer Sci.*, 2008, **99**, 1304–1310.
5 K. Akasaka-Manya, H. Manya, Y. Sakurai, B. S. Wojczyk, Y. Kozutsumi, Y. Saito, N. Taniguchi, S. Murayama, S. L. Spitalnik and T. Endo, *Glycobiology*, 2010, **20**, 99–106.
6 A. Herscovics, *Biochim. Biophys. Acta*, 1999, **1473**, 96–107.
7 Y. T. Pan, H. Hori, R. Saul, B. a Sanford, R. J. Molyneux and a D. Elbein, *Biochemistry*, 1983, **22**, 3975–3984.
8 V. W. Sasak, J. M. Ordovas, a D. Elbein and R. W. Berninger, *Biochem. J.*, 1985, **232**, 759–766.
9 L. Foddy and R. C. Hughes, *Eur. J. Biochem.*, 1988, **175**, 291–299.
10 R. G. Spiro, *J. Biol. Chem.*, 2000, **275**, 35657–35660.
11 S. E. H. Moore and R. G. Spiro, *J. Biol. Chem.*, 1990, **265**, 13104–13112.
12 K. Fujimoto and R. Kornfeld, *J. Biol. Chem.*, 1991, **266**, 3571–3578.
13 C. Völker, C. M. De Praeter, B. Hardt, W. Breuer, B. Kalz-Füller, R. N. Van Coster and E. Bause, *Glycobiology*, 2002, **12**, 473–483.
14 W. A. Lubas and R. G. Spiro, *J. Biol. Chem.*, 1987, **262**, 3775–3781.
15 W. a Lubas and R. G. Spiro, *J. Biol. Chem.*, 1988, **263**, 3990–3998.
16 A. J. Thompson, R. J. Williams, Z. Hakki, D. S. Alonzi, T. Wennekes, T. M. Gloster, K. Songsrirote, J. E. Thomas-Oates, T. M. Wrodnigg, J. Spreitz, A. E. Stutz, T. D. Butters, S. J. Williams and G. J. Davies, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 781–786.
17 D. E. Koshland, *Biol. Rev.*, 1953, **28**, 416–436.
18 G. Speciale, A. J. Thompson, G. J. Davies and S. J. Williams, *Curr. Opin. Struct. Biol.*, 2014, **28**, 1–13.
19 Z. Hakki, A. J. Thompson, S. Bellmaine, G. Speciale, G. J. Davies and S. J. Williams, *Chem. Eur. J.*, 2015, **21**, 1966–1977.
20 M. Petricevic, L. F. Sobala, P. Z. Fernandes, L. Raich, A. J. Thompson, G. Bernardo-Seisdedos, O. Millet, S. Zhu, M. Sollogoub, J. Jiménez-Barbero, C. Rovira, G. J. Davies and S. J. Williams, *J. Am. Chem. Soc.*, 2017, **139**, 1089–1097.
21 Y. Zhang, C. Chen, L. Jin, H. Tan, F. Wang and H. Cao, *Carbohydr. Res.*, 2015, **401**, 109–114.
22 G. Despras, R. Robert, B. Sendid, E. MacHez, D. Poulain and J. M. Mallet, *Bioorg. Med. Chem.*, 2012, **20**, 1817–1831.
23 M. S. Motawia, C. E. Olsen, K. Enevoldsen, J. Marcussen and B. L. Møller, *Carbohydr. Res.*, 1995, **277**, 109–123.
24 M. Martín-Lomas, N. Khiar, S. García, J. L. Koessler, P. M. Nieto and T. W. Rademacher, *Chem. Eur. J.*, 2000, **6**, 3608–3621.
25 M. Heuckendorff, J. Bendix, C. M. Pedersen and M. Bols, *Org. Lett.*, 2014, **16**, 1116–1119.
26 D. Crich and V. Dudkin, *Tetrahedron Lett.*, 2000, **41**, 5643–5646.

27 S. van der Vorm, T. Hansen, H. S. Overkleeft, G. A. van der Marel and J. D. C. Codée, *Chem. Sci.*, 2017, **8**, 1867–1875.
28 J. Castilla, M. I. Matheu and S. Castill, 2011, 9622–9629.
29 K. Sato and J. Yoshimura, *Carbohydr. Res.*, 1979, **73**, 75–84.
30 E. J. Corey and M. Chaykovsky, *J. Am. Chem. Soc.*, 1962, **84**, 3782–3783.
31 E. J. Corey and M. Chaykovsky, *J. Am. Chem. Soc.*, 1965, **87**, 1353–1364.
32 E. J. Corey and M. Chaykovsky, *J. Am. Chem. Soc.*, 1962, **84**, 867–868.
33 D. Neuhaus and M. P. Williamson, *The Nuclear Overhauser Effect in Structural and Conformational Analysis, 2nd ed.*, Wiley: New York, 2000.
34 R. A. Bell and J. K. Saunders, *Can. J. Chem.*, 1970, **48**, 1114–1122.
35 O. Kaczmarek, H. A. Scheidt, A. Bunge, D. Föse, S. Karsten, A. Arbuzova, D. Huster and J. Liebscher, *Eur. J. Org. Chem.*, 2010, 1579–1586.
36 S. C. Garman and D. N. Garboczi, *J. Mol. Biol.*, 2004, **337**, 319–335.
37 Y. Sohn, J. M. Lee, H. R. Park, S. C. Jung, T. H. Park and D. B. Oh, *BMB Rep.*, 2013, **46**, 157–162.
38 J. Jiang, T. J. M. Beenakker, W. W. Kallemeijn, G. A. van der Marel, H. van den Elst, J. D. C. Codée, J. M. F. G. Aerts and H. S. Overkleeft, *Chem. Eur. J.*, 2015, **21**, 10861–10869.
39 T. Nakajima and C. E. Ballou, *J. Biol. Chem.*, 1974, **249**, 7685–7694.
40 V. Ladmiral, G. Mantovani, G. J. Clarkson, S. Cauet, J. L. Irwin and D. M. Haddleton, *J. Am. Chem. Soc.*, 2006, **128**, 4823–4830.
41 C. M. Pedersen, L. U. Nordstrøm and M. Bols, *J. Am. Chem. Soc.*, 2007, **129**, 9222–9235.
42 M. D. Witte, W. W. Kallemeijn, J. Aten, K.-Y. Li, A. Strijland, W. E. Donker-Koopman, A. M. C. H. van den Nieuwendijk, B. Bleijlevens, G. Kramer, B. I. Florea, B. Hooibrink, C. E. M. Hollak, R. Ottenhoff, R. G. Boot, G. A. van der Marel, H. S. Overkleeft and J. M. F. G. Aerts, *Nat. Chem. Biol.*, 2010, **6**, 907–913.