Cyclophellitol analogues for profiling of exo- and endo-glycosidases

Schröder, S.P.

Cyclophellitol analogues for profiling of *exo*- and *endo*-glycosidases

PROEFSCHRIFT

ter verkrijging van

de graad van Doctor aan de Universiteit Leiden,

op gezag van Rector Magnificus prof. mr. C. J. J. M. Stolker,

volgends besluit van het College voor Promoties

te verdedigen op donderdag 17 mei 2018

klokke 15:00 uur

door

Sybrin Paul Schröder

Geboren te Nijmegen in 1989

Promotiecommissie

Promotores:	Prof. dr. H. S. Overkleeft
	Prof. dr. G. A. van der Marel
Co-promotor:	Dr. J. D. C. Codée
Overige leden:	Prof. dr. M. van der Stelt
	Prof. dr. J. M. F. G. Aerts
	Dr. M. Artola
	Prof. dr. ir. A. J. Minnaard (Rijksuniversiteit Groningen)
	Prof. dr. G. J. Davies (University of York, Verenigd Koninkrijk)
	Prof. dr. S. J. Williams (University of Melbourne, Australië)

Printed by Ridderprint B.V.

ISBN: 978-94-6299-964-0

Voor O.B.

With your feet on the air and your head on the ground

try this trick and spin it

your head will collapse, but there's nothing in it

and you'll ask yourself

where is my mind?

- The pixies

Table of contents

Chapter 1

Introduction and outline

Glycosidic bonds are one of the most stable linkages within natural biopolymers, with half-lives of starch and cellulose reaching approximately 5 million years.[1] Glycosidases are enzymes that catalyze the hydrolysis of glycosidic bonds in oligosaccharides[2], glycoproteins[3] and glycolipids[4], by increasing the hydrolysis rate by 10^{17}-fold.[1] Correct processing of glycans is essential to cellular function as glycans are involved in energy-storage, signaling pathways, folding, stability, activity and functioning of proteins. Glycosidases are generally highly specific towards the absolute configuration of their glycosidic substrate, and due to the structural diversity of glycans,[5] many distinct glycosidase families have evolved. To date, glycosidases are categorized in over 130 glycoside hydrolase (GH) families, based on structural similarities and amino acid sequence in their catalytic domain (see: *cazypedia.org*).[6] While research on glycosidases has successfully unraveled their function in various complex biological systems, the role and catalytic mechanism of many glycosidases remains unclear. Glycosidases can be studied using specific (covalent) inhibitors, which allow interrogation of their catalytic machinery, mechanism and itinerary from

X-ray crystallographic snapshots of (covalent) enzyme-inhibitor complexes. Additionally, covalent activity based probes (ABPs) allow profiling of glycosidases in biological settings, thereby enabling the study of glycosidase involvement in health and disease. This Thesis describes the synthesis and biochemical evaluation of such (covalent) glycosidase inhibitors and ABPs. Since glycosidases are highly specific towards their substrates, the design of such inhibitors requires knowledge of their function and mechanistic aspects. Some of these aspects are introduced in this Chapter.

1.1 Glycosidases in biomedicine and biotechnology

Aberrant glycosidase function often results in accumulation of its substrates, which is at the basis of a number of human diseases.[7] For example, impaired activity of lysosomal β-glucocerebrosidase (glucosylceramidase, GBA1), caused by inherited mutations in the gene encoding for this enzyme, results in accumulation of its substrate, glucosylceramide, in macrophages, inducing spleen and liver enlargement, skeletal disorders, anemia and other symptoms.[8] This lysosomal storage disorder is called Gaucher disease and current treatments are based on intravenous administration of recombinant GBA1 (enzyme replacement therapy[9]) or reducing glucosylceramide biosynthesis by inhibiting glucosylceramide synthase (substrate reduction therapy[10]). Recently, competitive GBA inhibitors have been evaluated as alternatives for the treatment of Gaucher disease.[11] This strategy, termed pharmacological chaperone therapy (PCT), relies on the administration of such competitive GBA inhibitors to assist folding of the nascent GBA polypeptide into its active conformation in the endoplasmic reticulum. Other lysosomal storage disorders include Pompe disease (α-glucosidase deficiency),[12] Fabry disease and Krabbe disease (deficiency in α- and β-galactocerebrosidase, respectively),[13,14] mannosidosis (deficiency in α- and β-mannosidases)[15,16] and mucopolysaccharidosis type III-B (deficiency in α-*N*-acetyl-glucosaminidase).[17]

Glycosidases are widely applied enzymes in biotechnology. For example, they are used for enzymatic “biobleaching” in the paper and pulp industries.[18] Glycosidases are also used in food production, as ingredients of detergents, in the production of textiles,[19] and in the production of biofuels (such as ethanol) from organic (waste) material.[20] In the latter application, polymeric plant material is hydrolysed by glycosidases to afford monomeric saccharides, which are subsequently fermented to produce combustible materials.

1.2 Mechanistic aspects of glycosidases

Glycosidases can be subdivided into two distinct classes, namely *exo*- and *endo*-glycosidases (Figure 1). *Exo*-glycosidases hydrolyse polysaccharides at the terminal non-reducing (in some cases, also reducing) end of the carbohydrate chain. Typically, their active site is shaped as a pocket, providing space to accommodate a monosaccharide.[21] In contrast, e*ndo*-glycosidases hydrolyse polysaccharide chains at internal positions. Their active site is either cleft- or tunnel shaped, and can accommodate multiple monosaccharide residues.[21] Glycosidase active sites can be dissected in multiple subsites (-*n* to +*n*) which interact with the polysaccharide substrate.[22] The -*n* subsite(s) accommodate the glycon moiety, whereas the aglycon is positioned in the +*n* subsite(s). The hydrolytic apparatus is situated at subsite -1. Generally, e*xo*-glycosidases have strong interactions with the monosaccharide occupying the -1 position, but less so with the aglycon at +1. The minimal substrate size of e*ndo*-glycosidases is a trisaccharide, hence *endo*-glycosidases have strong interactions at subsites -2 to +1, however, several subsites are generally involved in substrate binding. For example, the active site of cellobiohydrolase I from *Trichoderma reesei* spans up to 11 binding sites (-7 to +4).[23]

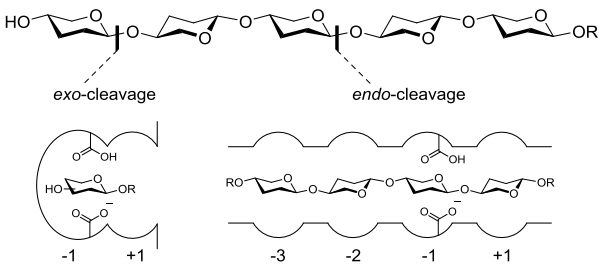


Figure 1 Glycosidases can be classified into *endo*- and *exo*-acting hydrolases. *Exo*-glycosidases generally cleave their glycan substrate at the non-reducing end of the polysaccharide chain. Their active site is pocket-shaped and contains only two subsites (-1 and +1). *Endo*-glycosidases cleave polysaccharides at internal positions, and their cleft or tunnel shaped active sites can be divided into multiple subsites.

Glycosidases can also be divided in two major classes based on the stereochemical outcome of the hydrolytic reaction. These classes employ distinct catalytic mechanisms as was first outlined by Daniel E. Koshland, Jr in 1953.[24] Inverting glycosidases hydrolyse their substrates with inversion of stereochemistry at the anomeric position (Figure 2). During this one-step, single displacement mechanism,

the reaction itinerary proceeds through an oxocarbenium-like transition state. The catalytic apparatus consists of a base and acid residue, most often a carboxylate and a carboxylic acid, thus aspartic acid and/or glutamic acid (Asp/Glu) residues. The catalytic residues are located 6-12 Å apart, providing space for the substrate and a water molecule. When the substrate is correctly coordinated in the active site, the catalytic base deprotonates the water molecule, which subsequently displaces the substrate's aglycon moiety. The aglycon is displaced with protic assistance of the catalytic acid/base residue.[25–27]

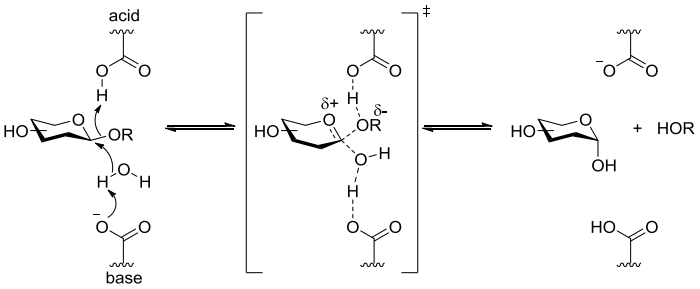


Figure 2 Catalytic reaction mechanism of the hydrolysis of a β-glycoside employed by inverting glycosidases, using a one-step single displacement mechanism.

Retaining glycosidases employ a two-step, double displacement mechanism to hydrolyse their substrates (Figure 3). The active site contains a catalytic nucleophile (Asp/Glu) and an acid/base residue (Asp/Glu), situated approximately 5.5 Å apart. During the first step (glycosylation), the substrate undergoes nucleophilic attack at the anomeric centre, while the aglycon is protonated by the catalytic acid/base. An oxocarbenium-like transition state emerges, the aglycon is expelled and a covalent intermediate is formed between the substrate and the enzyme. During the second step (deglycosylation), the catalytic base deprotonates a water molecule, which then attacks the anomeric centre of the enzyme-substrate complex. Following a second transition state, the covalent enzyme-substrate bond is cleaved and the hydrolysed substrate is released, with retention of stereochemistry at the anomeric position.

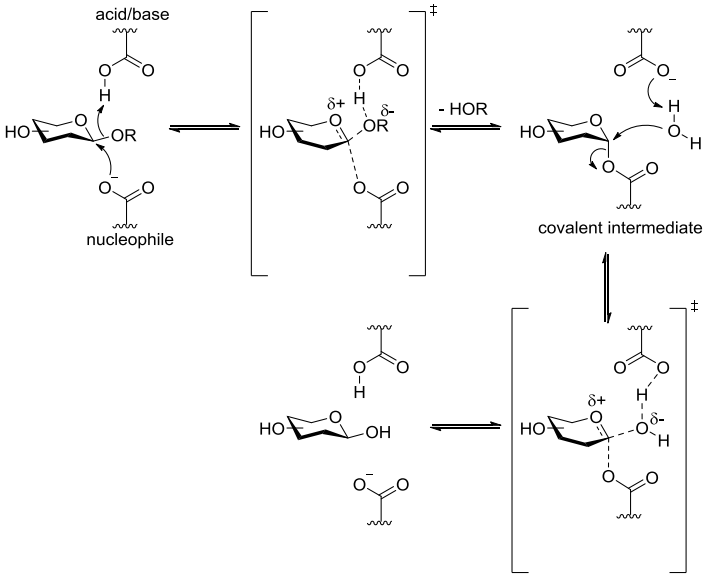


Figure 3 Catalytic reaction mechanism of the hydrolysis of a β-glycoside employed by retaining glycosidases, using a two-step double displacement mechanism. The reaction pathway involves a covalent enzyme-substrate intermediate, which is subsequently hydrolyzed.

Most glycosidases can be categorized in these two classes, however there are exceptions. For instance, 2-*N*-acetyl substrates are hydrolysed by GH families 18, 20, 25, 56, 84 and 85 via neighbouring group participation.[28–31] These enzymes lack a catalytic nucleophile; instead, the 2-acetamido group acts as such and intramolecularly displaces the aglycon. Therefore, while the enzymatic itinerary follows a two-step double displacement mechanism, no covalent intermediate is formed with the substrate. Other enzyme classes include: sialidases,[32,33] which utilize a tyrosine residue as catalytic nucleophile; myrosinases,[34] which lack a catalytic acid and employ an external base; and NAD-dependent glycosidases,[35,36] which employ an oxidation-elimination-reduction pathway. Recently, it was proposed that GH99 *endo*-mannosidases follow a two-step, double displacement mechanism with neighbouring group participation of OH-2, forming an 1,2-anhydro-epoxide which is subsequently hydrolyzed.[37–39]

The classical Koshland hydrolytic reaction mechanism is widely appreciated by the scientific community; typically schematically represented by placement of the catalytic residues at the top and bottom face of the substrate anomeric centre. In

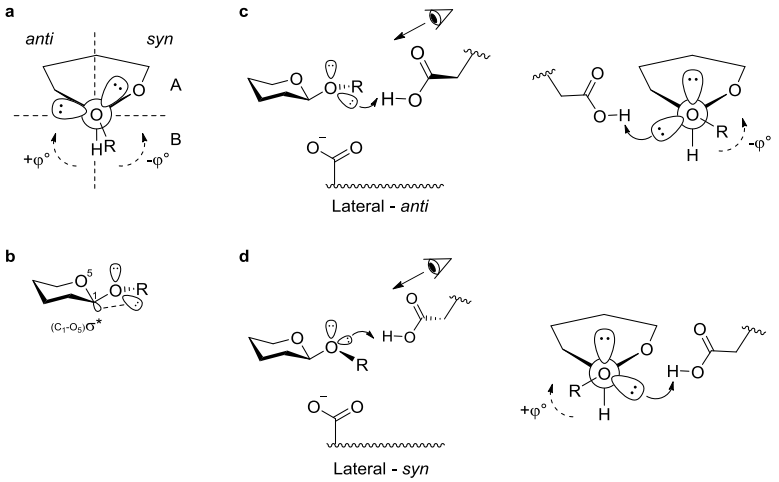


Figure 4 Semi-lateral protonation of the aglycon by the catalytic acid/base residue. (a) The spacial environment of a glycoside can be divided into two half-spaces (*syn* and *anti*) which are separated by a plane defined by the anomeric C-H bond. The catalytic acid/base in the majority of glycosidases resides in the *anti* half-space. A plane perpendicular to the anomeric C-H bond defines the A or B half-spaces. For β-glycosidases the catalytic acid/base resides in the A half-space, whereas for α-glycosidases it is situated in the B half-space. (b) The exo-anomeric effect in β-glycosides. (c) Substrate protonation by an *anti*-protonator. (d) Substrate protonation by a *syn*-protonator.

reality, while the catalytic nucleophile residue is indeed placed below or above the anomeric centre, this does not hold true for the catalytic acid/base. As first postulated by Vasella and co-workers[40] and later confirmed by X-ray crystallographic snapshots of many enzyme-inhibitor complexes,[41] the catalytic acid/base is positioned semi-lateral ('half-space' A for β-glycosidases, 'half-space' B for α-glycosidases) from the aglycon oxygen, and is situated in the *syn* (near the endocyclic oxygen) or *anti* (opposite the endocyclic oxygen) half-space (Figure 4a).

The *syn* or *anti* positioning of the catalytic acid/base is conserved within GH families, and roughly 70% of all families are *anti*-protonators. The lowest energy conformation in β-glycosides involves the glycosidic anomeric bond rotated in a minus-synclinal angle (-ϕ°) to allow hyperconjugative overlap of its antiperiplanar oxygen lone-pair orbital with the C1-O5 antibonding (σ*) orbital (Figure 4b), a phenomenon called the exo-anomeric effect.[42] In *anti*-protonating β-glycosidases, the catalytic acid/base is situated in the *anti*-A quadrant and protonates this antiperiplanar glycosidic oxygen lone-pair orbital (Figure 4c). This cancels out the stabilizing exo-anomeric effect,

thereby lowering the activation barrier for hydrolysis. *Syn*-protonating β-glycosidases display their catalytic acid/base in the *syn*-A quadrant (Figure 4d). Protonation of the aglycon requires a minus-anticlinal (+ϕ°) rotation of the glycosidic bond, thereby cancelling out the exo-anomeric effect as well.

As depicted in Figure 2 and 3, the mechanistic reaction itineraries for inverting and retaining glycosidases follow different steps involving substrate distortions towards the transition state(s). Such conformational distortions were initially proposed for hen egg white lysozyme,[43] and eventually found to be glycosidase-specific.[44,45] For example, upon substrate entrance into the active site, most retaining β-glucosidases distort their substrates into a $^{1}S_{3}$ conformation (Michaelis complex), followed by a $^{4}H_{3}$ transition state conformation (Figure 5a).[46–48] Following substitution of the aglycon, the resulting covalent intermediate occupies a relaxed $^{4}C_{1}$ conformation, which is then further hydrolysed. While many reaction itineraries of other glycosidase classes are yet unknown, this same $^{1}S_{3} \rightarrow [^{4}H_{3}]^{\ddagger} \rightarrow {}^{4}C_{1}$ itinerary is also proposed for other glycosidase families including GH2 β-galactosidases[49,50], GH10 *endo*-xylanases[51,52] and GH39 xylosidases.[53] In contrast, many α-glucosidases[54,55] and α-galactosidases[56] follow a 'reversed' $^{4}C_{1} \rightarrow [^{4}H_{3}]^{\ddagger} \rightarrow {}^{1}S_{3}$ conformational pathway (Figure 5b).

Hydrolysis of α-mannopyranosides by retaining α-mannosidases involves substitution of the aglycon to furnish a covalent β-mannoside-enzyme complex. The construction of β-glycosidic linkages from α-mannosides is a well-known challenge for synthetic organic chemists.[57] Substitution of α-mannoside donors towards β-mannopyranosides is hampered by the steric hindrance of O2 with the incoming nucleophile. Furthermore, α-mannosides are stabilized by the anomeric effect and the opposing [C1,O1 – C2,O2] dipoles so that conformational distortion is disfavoured. Lastly, introduction of a glycosidic β-linkage in mannopyranosides induces the Δ2 effect[58]; a destabilizing alignment of the [C1,O(1,5) – C2,O2] dipoles. Nature has provided an effective way of introducing glycosidic β-linkages in mannopyranosides. By forming a Michaelis complex with the substrate in the $^{0}S_{2}$ conformation, the aglycon is arranged in a pseudo axial position, and inline attack of the carboxylate is facilitated (Figure 5c).[59,60] Following a $B_{2,5}$ transition state, the covalent intermediate is formed in the $^{1}S_{5}$ conformation which reduces the destabilizing Δ2 effect. In analogy, β-mannosidases follow the 'reversed' itinerary, employing $^{1}S_{5} \rightarrow [B_{2,5}]^{\ddagger} \rightarrow {}^{0}S_{2}$ substrate conformations (Figure 5d).[61–63]

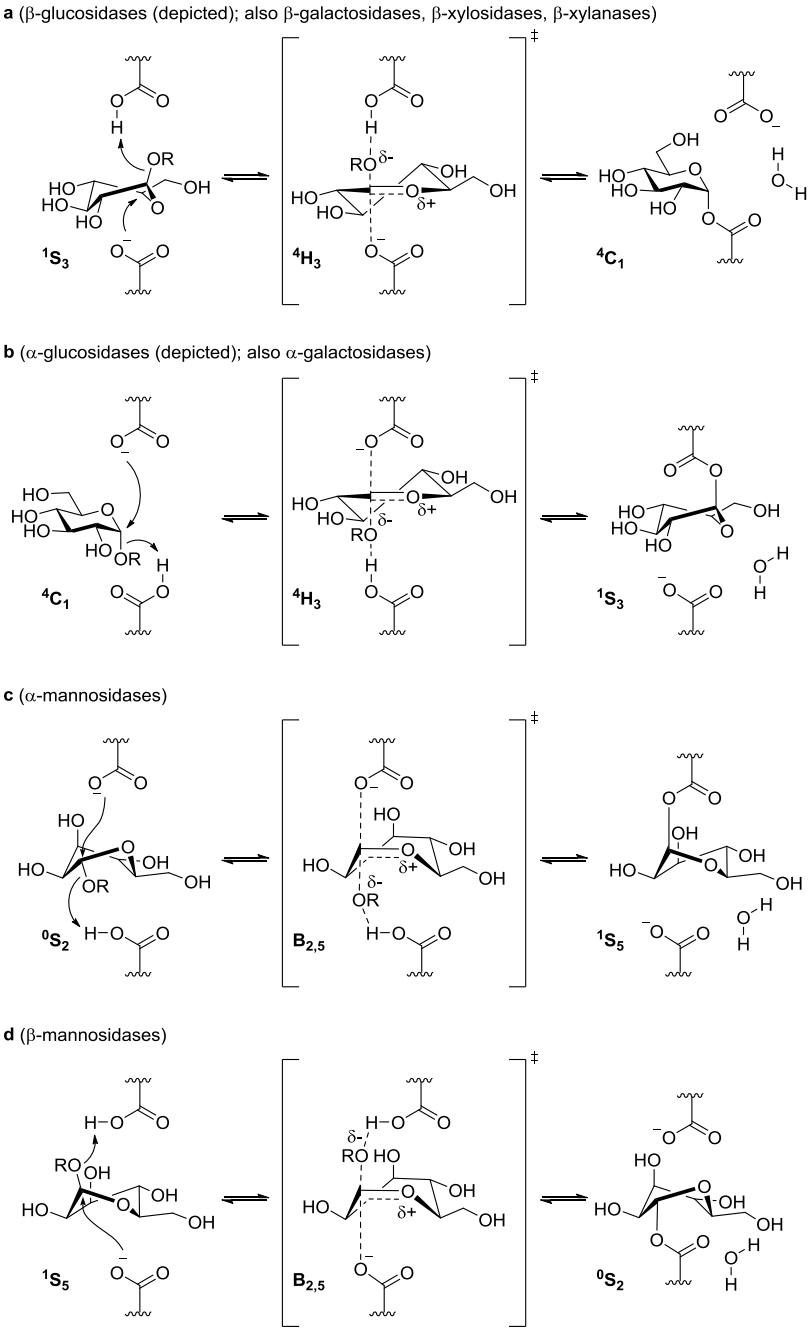


Figure 5 Different conformational itineraries of the substrate during processing by retaining α- and β-glycosidases. The hydrolysis of the covalent intermediates is omitted for clarity.

1.3 Activity-based protein profiling

Enzymes that form a covalent bond with their substrate during catalysis are amenable to activity-based protein profiling (ABPP).[64,65] ABPP commences with incubation of a relevant biological sample with an activity-based probe (ABP), which reacts selectively, covalently and irreversibly with the enzyme (class) of interest. With this technique, generally only active enzymes are targeted. An ABP generally consists of a recognition element, which ensures recognition by the target enzyme selectively, and an electrophile (warhead), which reacts covalently and irreversibly with the catalytic machinery of functional enzymes. Furthermore, the ABP is equipped with a reporter (fluorescent or affinity) tag, which enables quantitative and qualitative read-out of the biochemical assay (Figure 6a).

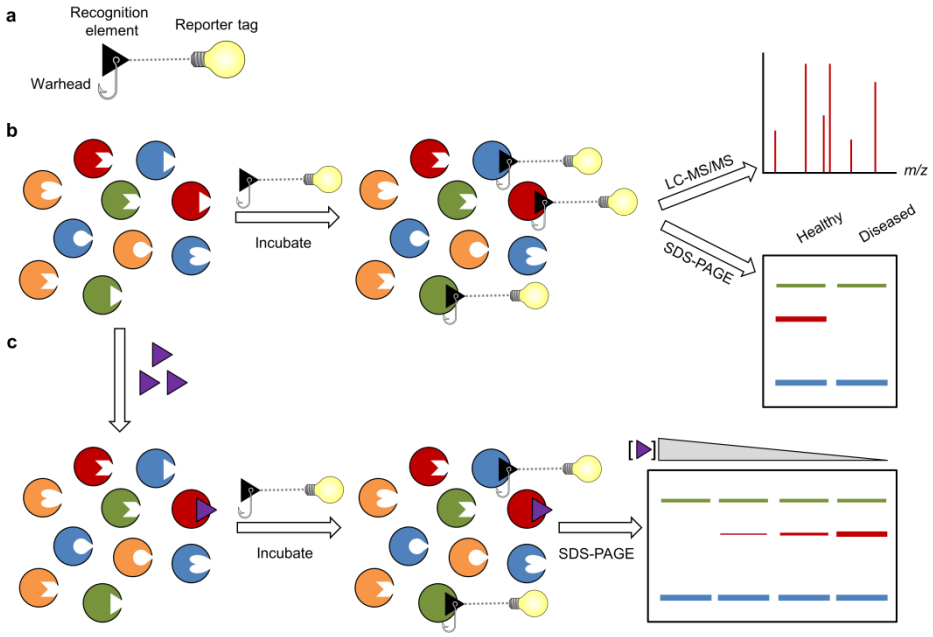


Figure 6 Schematic overview of activity-based protein profiling (ABPP) techniques. (a) The activity-based probe consists of a recognition element containing an electrophilic trap for irreversible reaction with the target enzyme, and a reporter tag, which enables analysis of the experiment. (b) Incubation of a complex biological sample with the probe results in binding with the target enzymes, which can then be identified by proteomics or analyzed by SDS-PAGE in a comparative ABPP study. (c) In a competitive ABPP assay, the complex biological sample is incubated with a range of inhibitor concentrations, and then post-labeled by the probe. Only the fraction of non-inhibited target enzymes will bind to the probe.

For example, probes that are equipped with a biotin affinity tag are used to enrich the enzyme(s) that reacted with the ABP, via affinity purification using streptavidin immobilized on beads. The enriched protein(s) are subsequently removed from the beads using proteolysis (usually trypsinolysis) and the resulting peptides analysed by LC-MS/MS and matched against protein sequence databases to identify the target protein (Figure 6b).[66] Alternatively, active enzymes of interest can be visualized by probes containing a fluorescent reporter tag, for instance a BODIPY, cyanine or TAMRA fluorophore. Following sample incubation and irreversible inactivation of the target enzyme(s) with the probe, the complex protein mixture is denatured and then analysed by SDS-PAGE. Using fluorescent imaging, the protein band(s) with corresponding molecular weight correlating to the targeted enzyme are visualized. Precursor enzymes, inactive or malfunctioning enzymes are generally not targeted by this technique and the band intensity can be linked to enzyme expression levels. Thus, ABPs can be used to correlate enzyme concentration to tissues, or as diagnostic tools for investigating active enzyme levels in patients *versus* healthy individuals (comparative ABPP[67–69]). Furthermore, fluorescent ABPs are used in competition assays, where the enzyme of interest is pre-incubated with different inhibitor concentrations followed by labelling with the ABP (competitive ABPP,[70,71] Figure 6c). With such assays, inhibitory potency is visualized which corresponds to the decrease in fluorescent signal. When the ABP used targets multiple enzyme classes, off-target activity of the inhibitor can be visualized and identified. Enzymes that are currently studied by ABPP include serine hydrolases,[72–74] cysteine hydrolases,[75–77] proteasomes,[78–80] kinases[81–83] and also glycosidases.[84–86]

1.4 Covalent, irreversible, cyclophellitol-based glycosidase probes

An important class of covalent retaining glycosidase inhibitors are the cyclitol-epoxides.[87] For example, conduritol B epoxide (CBE, **1**, Figure 7a), first described by Legler,[88] is a potent inactivator of lysosomal β-glucocerebrosidase (GBA1)[89] and other glucosidases.[90] The cyclitol core partially mimics the glucopyranoside substrate and is recognized as such. Following coordination in the active site, the catalytic nucleophile performs a nucleophilic attack on the oxirane at the anomeric position, and the nucleophilic displacement is aided by protonation of the ring oxygen by the catalytic acid/base, resulting in a covalent, irreversible complex (Figure 7b). However, CBE is not β-glucosidase selective and also inhibits α-glucosidases due to its C_2-symmetry, albeit with lower potency.[91] Cyclophellitol (**2**, Figure 7a), isolated as a natural product from *Phellinus* sp.[92] and available via various synthetic routes,[93,94] is also a covalent

irreversible β-glucosidase inactivator. It is structurally related to CBE, but contains a hydroxymethylene functionality at C5 (instead of hydroxyl for CBE), is selective for β-glucosidases over α-glucosidases and displays higher potency due to its higher natural substrate resemblance.[95] Subsequently it was found that the close structural analogue cyclophellitol aziridine[96] (**3**, Figure 7a) displayed even higher potency towards β-glucosidases.[97]

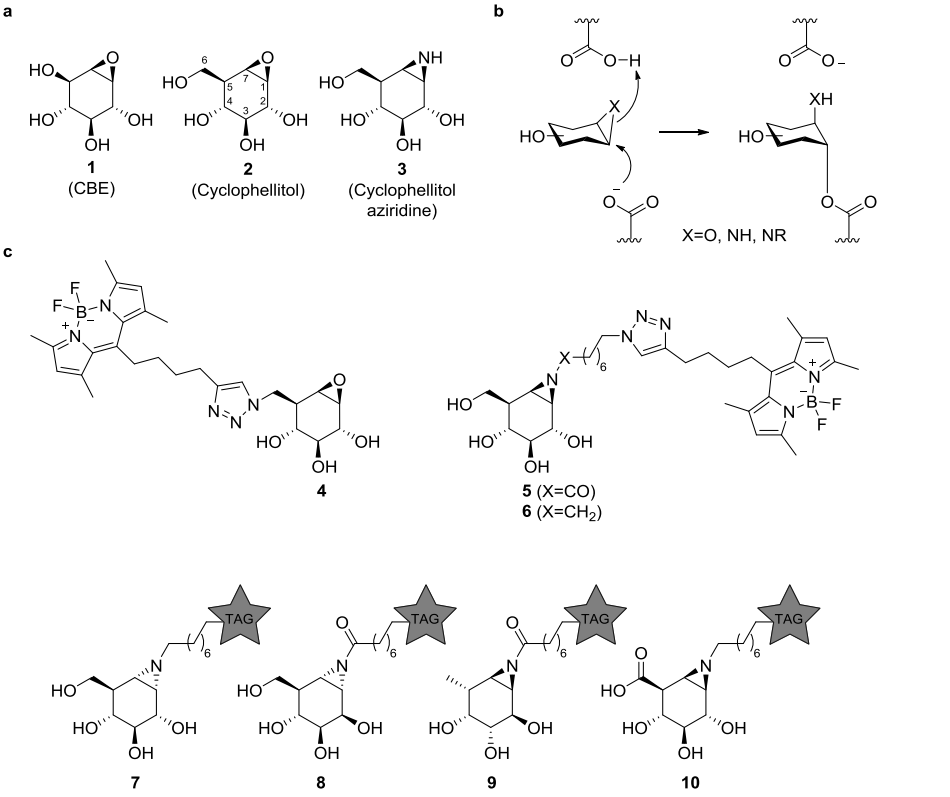


Figure 7 (a) Covalent irreversible inhibitors of β-glucosidases: conduritol B epoxide (**1**), cyclophellitol (**2**) and cyclophellitol aziridine (**3**). (b) Mechanism of glycosidase inactivation by cyclitol epoxides and aziridines. After opening of the electrophilic trap, a covalent adduct is formed which is stable towards hydrolysis. (c) Cyclophellitol equipped with a fluorescent tag at C6 (**4**) is a potent probe for GBA1. Cyclophellitol aziridine acylated (**5**) or alkylated (**6**) with a fluorescent tag are potent probes for the mammalian retaining β-glucosidases, GBA1, GBA2, GBA3 and LPH. Specific ABPs for α-glucosidases (**7**),[98] α-galactosidases (**8**),[99] α-fucosidases (**9**)[68] and β-glucoronidases (**10**)[100] have also been developed.

Due to the covalent and irreversible mode of action that cyclophellitol employs towards β-glucosidases, its core structure was used for the generation of activity based probes. Endowing position C6 in cyclophellitol (**2**) with a fluorescent BODIPY tag afforded ABP **4** (Figure 7c), which proved to be a highly potent and selective probe for human GBA1 over the other three human β-glucosidases GBA2, GBA3 and lactase-phlorizin hydrolase (LPH).[101] In a subsequent study, the nitrogen in cyclophellitol aziridine (**3**) was equipped with a fluorescent tag affording *N*-acyl ABP **5**.[102] In this case, the reporter moiety occupies the space normally taken up by the natural substrate aglycon, and it was found that this probe labels all β-glucosidases GBA1, GBA2, GBA3 and LPH in mice with high potency. However, *N*-acyl ABP **5** was found to be somewhat unstable during synthesis, purification and handling. Later it was found that its more accessible *N*-alkylated counterpart (**6**) was equally effective in β-glucosidase labeling.[103] In analogy, specific ABPs based on the cyclophellitol template have been synthesized and evaluated for labeling of α-glucosidases (**7**),[98] α-galactosidases (**8**),[99] α-fucosidases (**9**)[68] and β-glucoronidases (**10**).[100]

1.5 Outline of this Thesis

To this day, all cyclophellitol-based inhibitors and ABPs have been close analogues of their natural substrate counterparts. As a result, these probes showed high selectivity towards their target glycosidases. While such probes are of high value for studying these specific enzyme classes, they impede the simultaneous profiling of glycosidases that process different substrate configurations with a single probe. The first Chapters in this Thesis focus on the structural derivatization of cyclophellitol-based probes with the aim of enabling inter-class labelling of glycosidases. **Chapter 2** describes the synthesis of cyclophellitol-based inhibitors and probes, which lack the hydroxymethylene functionality at C5, and their labelling is biochemically evaluated on complex biological samples. The synthesis of cyclophellitol-based probes which are lacking hydroxyl substituents at positions C2 and C4 are described in **Chapter 3**, as well as the biochemical evaluation of these probes towards inter-class labelling of α- and β-glucosidases, mannosidases and galactosidases in cell extracts. **Chapter 4** reports on the synthesis and biochemical evaluation of D-*lyxo* and D-*arabino*furanoside analogues of cyclophellitol aziridine. The synthesis of a novel reversible inhibitor, gluco-1*H*-imidazole, is described in Chapter **5** and its inhibitory activity and binding properties are investigated. To date, all cyclitol-epoxide/aziridine based probes are monomeric in structure and therefore only target *exo*-glycosidases. In the last Chapters of this Thesis, synthetic methodologies for the construction of *endo*-

glycosidase probes are described. **Chapter 6** reports on the introduction of a spiro-epoxide warhead onto a disaccharide moiety for the purpose of constructing probes for GH99 *endo*-α-mannosidases, and their labelling potency is evaluated. The construction of x*ylobiose*-cyclophellitol probes from *xylo*-cyclophellitol acceptors via direct glycosylation is described in **Chapter 7**, and the xylanase labelling potency is evaluated in the secretome of *Aspergillus niger*. Finally, **Chapter 8** provides a summary of this Thesis, followed by future prospects including the synthesis of multimeric irreversible probes and inhibitors for GH76 *endo*-α-mannosidase.

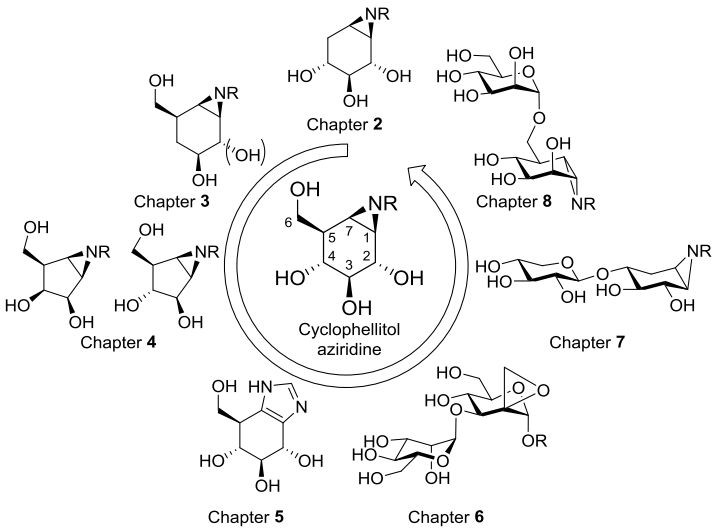


Figure 8 Global overview of the inhibitors and activity-based probes (based on the cyclophellitol aziridine scaffold, depicted in the middle) discussed in this Thesis.

References

1 R. Wolfenden, X. Lu and G. Young, *J. Am. Chem. Soc.*, 1998, **120**, 6814–6815.

2 W. D. Comper and T. C. Laurent, *Physiol. Rev.*, 1978, **58**, 255–315.

3 A. Herscovics, *Biochim. Biophys. Acta*, 1999, **1473**, 96–107.

4 W. Curatolo, *Biochim. Biophys. Acta*, 1987, **906**, 111–136.

5 R. D. Cummings, *Mol. Biosyst.*, 2009, **5**, 1087–1104.

6 V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho and B. Henrissat, *Nucleic Acids Res.*, 2014, **42**, 490–495.

7 P. J. Meikle, J. J. Hopwood, A. E. Clague and W. F. Carey, *JAMA*, 1999, **281**, 249–254.

8 E. Sidransky, *Mol. Genet. Metab.*, 2004, **83**, 6–15.

9 N. W. Barton, R. O. Brady, J. M. Dambrosia, A. M. Di Bisceglie, S. H. Doppelt, S. C. Hill, H. J. Mankin, G. J. Murray, R. I. Parker, C. E. Argoff, R. P. Grewal and K.-T. Yu, *N. Engl. J. Med.*, 1991, **324**, 1464–1470.

10 T. Cox, R. Lachmann, C. Hollak, J. Aerts, S. van Weely, M. Hrebícek, F. Platt, T. Butters, R. Dwek, C. Moyses, I. Gow, D. Elstein and A. Zimran, *Lancet*, 2000, **355**, 1481–1485.
11 A. R. Sawkar, W.-C. Cheng, E. Beutler, C.-H. Wong, W. E. Balch and J. W. Kelly, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 15428–15433.
12 A. T. van der Ploeg and A. J. Reuser, *Lancet*, 2008, **372**, 1342–1353.
13 K. Suzuki, *J. Child Neurol.*, 2003, **18**, 595–603.
14 Y. A. Zarate and R. J. Hopkin, *Lancet*, 2008, **372**, 1427–1435.
15 D. Malm and Ø. Nilssen, *Orphanet J. Rare Dis.*, 2008, **3**, 21.
16 F. Percheron, M. J. Foglietti, M. Bernard and B. Ricard, *Biochimie*, 1992, **74**, 5–11.
17 M. J. Valstar, G. J. G. Ruijter, O. P. van Diggelen, B. J. Poorthuis and F. A. Wijburg, *J. Inherit. Metab. Dis.*, 2008, **31**, 240–252.
18 V. Kumar, J. Marín-Navarro and P. Shukla, *World J. Microbiol. Biotechnol.*, 2016, **32**, 1–10.
19 M. K. Bhat, *Biotechnol. Adv.*, 2000, **18**, 355–383.
20 M. Balat, *Energy Convers. Manag.*, 2011, **52**, 858–875.
21 G. Davies and B. Henrissat, *Structure*, 1995, **3**, 853–859.
22 G. J. Davies, K. S. Wilson and B. Henrissat, *Biochem. J.*, 1997, **321**, 557–559.
23 C. Divne, J. Ståhlberg, T. T. Teeri and T. A. Jones, *J. Mol. Biol.*, 1998, **275**, 309–325.
24 D. E. Koshland, *Biol. Rev.*, 1953, **28**, 416–436.
25 D. L. Zechel and S. G. Withers, *Acc. Chem. Res.*, 2000, **33**, 11–18.
26 J. D. McCarter and G. Stephen Withers, *Curr. Opin. Struct. Biol.*, 1994, **4**, 885–892.
27 V. L. Y. Yip and S. G. Withers, *Org. Biomol. Chem.*, 2004, **2**, 2707–2713.
28 D. J. Vocadlo and S. G. Withers, *Biochemistry*, 2005, **44**, 12809–12818.
29 A. C. Terwisscha van Scheltinga, S. Armand, K. H. Kalk, A. Isogai, B. Henrissat and B. W. Dijkstra, *Biochemistry*, 1995, **34**, 15619–15623.
30 B. L. Mark, D. J. Vocadlo, S. Knapp, B. L. Triggs-Raine, S. G. Withers and M. N. G. James, *J. Biol. Chem.*, 2001, **276**, 10330–10337.
31 S. Knapp, D. Vocadlo, Z. Gao, B. Kirk, J. Lou and S. G. Withers, *J. Am. Chem. Soc.*, 1996, **118**, 6804–6805.
32 A. G. Watts, I. Damager, M. L. Amaya, A. Buschiazzo, P. Alzari, A. C. Frasch and S. G. Withers, *J. Am. Chem. Soc.*, 2003, **125**, 7532–7533.
33 M. F. Amaya, A. G. Watts, I. Damager, A. Wehenkel, T. Nguyen, A. Buschiazzo, G. Paris, A. C. Frasch, S. G. Withers and P. M. Alzari, *Structure*, 2004, **12**, 775–784.
34 W. P. Burmeister, S. Cottaz, P. Rollin, A. Vasella and B. Henrissat, *J. Biol. Chem.*, 2000, **275**, 39385–39393.
35 V. L. Y. Yip, A. Varrot, G. J. Davies, S. S. Rajan, X. Yang, J. Thompson, W. F. Anderson and S. G. Withers, *J. Am. Chem. Soc.*, 2004, **126**, 8354–8355.
36 S. S. Rajan, X. Yang, F. Collart, V. L. Y. Yip, S. G. Withers, A. Varrot, J. Thompson, G. J. Davies and W. F. Anderson, *Structure*, 2004, **12**, 1619–1629.
37 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.
38 Z. Hakki, A. J. Thompson, S. Bellmaine, G. Speciale, G. J. Davies and S. J. Williams, *Chem. Eur. J.*, 2015, **21**, 1966–1977.
39 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.
40 T. D. Heightman and A. T. Vasella, *Angew. Chem. Int. Ed.*, 1999, **38**, 750–770.
41 W. Nerinckx, T. Desmet, K. Piens and M. Claeyssens, *FEBS Lett.*, 2005, **579**, 302–312.
42 M. Wu, W. Nerinckx, K. Piens, T. Ishida, H. Hansson, M. Sandgren and J. Ståhlberg, *FEBS J.*, 2013, **280**, 184–198.
43 C. C. F. Blake, D. F. Koenig, G. A. Mair, A. C. T. North, D. C. Phillips and V. R. Sarma, *Nature*, 1965, **206**, 757–761.
44 G. Speciale, A. J. Thompson, G. J. Davies and S. J. Williams, *Curr. Opin. Struct. Biol.*, 2014, **28**, 1–13.
45 G. J. Davies, A. Planas and C. Rovira, *Acc. Chem. Res.*, 2012, **45**, 308–316.
46 G. Sulzenbacher, H. Driguez, B. Henrissat, M. Schülein and G. J. Davies, *Biochemistry*, 1996, **35**, 15280–15287.
47 G. J. Davies, L. Mackenzie, A. Varrot, M. Dauter, A. M. Brzozowski, M. Schülein and S. G. Withers, *Biochemistry*, 1998, **37**, 11707–11713.
48 X. Biarnés, J. Nieto, A. Planas and C. Rovira, *J. Biol. Chem.*, 2006, **281**, 1432–1441.
49 R. W. Wheatley, J. C. Kappelhoff, J. N. Hahn, M. L. Dugdale, M. J. Dutkoski, S. D. Tamman, M. E. Fraser and R. E. Huber, *Arch. Biochem. Biophys.*, 2012, **521**, 51–61.
50 D. H. Juers, T. D. Heightman, A. Vasella, J. D. McCarter, L. Mackenzie, S. G. Withers and B. W. Matthews, *Biochemistry*, 2001, **40**, 14781–14794.
51 R. Suzuki, Z. Fujimoto, S. Ito, S. I. Kawahara, S. Kaneko, K. Taira, T. Hasegawa and A. Kuno, *J. Biochem.*, 2009, **146**, 61–70.
52 V. Notenboom, S. J. Williams, R. Hoos, S. G. Withers and D. R. Rose, *Biochemistry*, 2000, **39**, 11553–11563.
53 M. Czjzek, A. Ben David, T. Bravman, G. Shoham, B. Henrissat and Y. Shoham, *J. Mol. Biol.*, 2005, **353**, 838–846.
54 A. L. Lovering, S. L. Seung, Y. W. Kim, S. G. Withers and N. C. J. Strynadka, *J. Biol. Chem.*, 2005, **280**, 2105–2115.
55 Y. W. Kim, A. L. Lovering, H. Chen, T. Kantner, L. P. McIntosh, N. C. J. Strynadka and S. G. Withers, *J. Am. Chem. Soc.*, 2006, **128**, 2202–2203.
56 A. I. Guce, N. E. Clark, E. N. Salgado, D. R. Ivanen, A. A. Kulminskaya, H. Brumer and S. C. Garman, *J. Biol. Chem.*, 2010, **285**, 3625–3632.
57 D. Crich, *Acc. Chem. Res.*, 2010, **43**, 1144–1153.
58 R. E. Reeves, *J. Am. Chem. Soc.*, 1950, **72**, 1499–1506.
59 S. Numao, D. A. Kuntz, S. G. Withers and D. R. Rose, *J. Biol. Chem.*, 2003, **278**, 48074–48083.
60 A. J. Thompson, G. Speciale, J. Iglesias-Fernández, Z. Hakki, T. Belz, A. Cartmell, R. J. Spears, E. Chandler, M. J. Temple, J. Stepper, H. J. Gilbert, C. Rovira, S. J. Williams and G. J. Davies, *Angew. Chem. Int. Ed.*, 2015, **54**, 5378–5382.
61 R. J. Williams, J. Iglesias-Fernández, J. Stepper, A. Jackson, A. J. Thompson, E. C. Lowe, J. M. White, H. J. Gilbert, C. Rovira, G. J. Davies and S. J. Williams, *Angew. Chem. Int. Ed.*, 2014, **53**, 1087–1091.
62 L. E. Tailford, W. a Offen, N. L. Smith, C. Dumon, C. Morland, J. Gratien, M.-P. Heck, R. V Stick, Y. Blériot, A. Vasella, H. J. Gilbert and G. J. Davies, *Nat. Chem. Biol.*, 2008, **4**, 306–312.
63 V. M. a. Ducros, D. L. Zechel, G. N. Murshudov, H. J. Gilbert, L. Szabó, D. Stoll, S. J. Withers and G. J. Davies, *Angew. Chem. Int. Ed.*, 2002, **114**, 2948–2951.
64 B. F. Cravatt, A. T. Wright and J. W. Kozarich, *Annu. Rev. Biochem.*, 2008, **77**, 383–414.
65 M. J. Evans and B. F. Cravatt, *Chem. Rev.*, 2006, **106**, 3279–3301.
66 K. V. M. Huber and G. Superti-Furga, *Methods Mol. Biol.*, 2016, **1394**, 211–218.

67 Y. Yang, X. Yang and S. H. L. Verhelst, *Molecules*, 2013, **18**, 12599–12608.

68 J. Jiang, W. W. Kallemeijn, D. W. Wright, A. M. C. H. van den Nieuwendijk, V. C. Rohde, E. C. Folch, H. van den Elst, B. I. Florea, S. Scheij, W. E. Donker-Koopman, M. Verhoek, N. Li, M. Schürmann, D. Mink, R. G. Boot, J. D. C. Codée, G. A. van der Marel, G. J. Davies, J. M. F. G. Aerts and H. S. Overkleeft, *Chem. Sci.*, 2015, **6**, 2782–2789.

69 M. Shahiduzzaman and K. M. Coombs, *Front. Microbiol.*, 2012, **3**, 1–5.

70 M. J. Niphakis and B. F. Cravatt, *Annu. Rev. Biochem.*, 2014, **83**, 341–377.

71 E. Deu, Z. Yang, F. Wang, M. Klemba and M. Bogyo, *PLoS One*, 2010, **5**, e11985.

72 G. M. Simon and B. F. Cravatt, *J. Biol. Chem.*, 2010, **285**, 11051–11055.

73 Y. Liu, M. P. Patricelli and B. F. Cravatt, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 14694–14699.

74 D. Kidd, Y. Liu and B. F. Cravatt, *Biochemistry*, 2001, **40**, 4005–4015.

75 G. Wang, U. Mahesh, G. Y. J. Chen and S. Q. Yao, *Org. Lett.*, 2003, **5**, 737–740.

76 A. M. Sadaghiani, S. H. L. Verhelst, V. Gocheva, K. Hill, E. Majerova, S. Stinson, J. A. Joyce and M. Bogyo, *Chem. Biol.*, 2007, **14**, 499–511.

77 D. Kato, K. M. Boatright, A. B. Berger, T. Nazif, G. Blum, C. Ryan, K. A. H. Chehade, G. S. Salvesen and M. Bogyo, *Nat. Chem. Biol.*, 2005, **1**, 33–38.

78 N. Li, C.-L. Kuo, G. Paniagua, H. van den Elst, M. Verdoes, L. I. Willems, W. A. van der Linden, M. Ruben, E. van Genderen, J. Gubbens, G. P. van Wezel, H. S. Overkleeft and B. I. Florea, *Nat. Protoc.*, 2013, **8**, 1155–1168.

79 D. S. Hewings, J. A. Flygare, I. E. Wertz and M. Bogyo, *FEBS J.*, 2017, **284**, 1540–1554.

80 G. de Bruin, B. T. Xin, M. Kraus, M. van der Stelt, G. A. van der Marel, A. F. Kisselev, C. Driessen, B. I. Florea and H. S. Overkleeft, *Angew. Chem. Int. Ed.*, 2016, **55**, 4199–4203.

81 Q. Zhao, X. Ouyang, X. Wan, K. S. Gajiwala, J. C. Kath, L. H. Jones, A. L. Burlingame and J. Taunton, *J. Am. Chem. Soc.*, 2017, **139**, 680–685.

82 B. R. Lanning, L. R. Whitby, M. M. Dix, J. Douhan, A. M. Gilbert, E. C. Hett, T. O. Johnson, C. Joslyn, J. C. Kath, S. Niessen, L. R. Roberts, M. E. Schnute, C. Wang, J. J. Hulce, B. Wei, L. O. Whiteley, M. M. Hayward and B. F. Cravatt, *Nat. Chem. Biol.*, 2014, **10**, 760–767.

83 D. Kitagawa, K. Yokota, M. Gouda, Y. Narumi, H. Ohmoto, E. Nishiwaki, K. Akita and Y. Kirii, *Genes to Cells*, 2013, **18**, 110–122.

84 M. D. Witte, M. T. C. Walvoort, K. Y. Li, W. W. Kallemeijn, W. E. Donker-Koopman, R. G. Boot, J. M. F. G. Aerts, J. D. C. Codée, G. A. van der Marel and H. S. Overkleeft, *ChemBioChem*, 2011, **12**, 1263–1269.

85 L. I. Willems, J. Jiang, K. Y. Li, M. D. Witte, W. W. Kallemeijn, T. J. N. Beenakker, S. P. Schröder, J. M. F. G. Aerts, G. A. van der Marel, J. D. C. Codée and H. S. Overkleeft, *Chem. Eur. J.*, 2014, **20**, 10864–10872.

86 K. A. Stubbs and D. J. Vocadlo, *Methods Enzymol.*, 2006, **415**, 253–268.

87 B. P. Rempel and S. G. Withers, *Glycobiology*, 2008, **18**, 570–586.

88 G. Legler, *Hoppe. Seylers. Z. Physiol. Chem.*, 1966, **345**, 197–214.

89 L. Premkumar, A. R. Sawkar, S. Boldin-Adamsky, L. Toker, I. Silman, J. W. Kelly, A. H. Futerman and J. L. Sussman, *J. Biol. Chem.*, 2005, **280**, 23815–23819.

90 G. Legler, *Methods Enzymol.*, 1977, **46**, 368–381.

91 M. M. P. Hermans, M. a Krooss, J. Van Beeurnens, B. a Oostras and A. J. J. Reusersll, *Biochemistry*, 1991, **266**, 13507–13512.

92 S. Atsumi, K. Umezawa, H. Iinuma, H. Naganawa, H. Nakamura, Y. Iitaka and T. Takeuchi, *J. Antibiot.*, 1989, **43**, 49–53.

93 J. Jiang, M. Artola, T. J. M. Beenakker, S. P. Schröder, R. Petracca, C. de Boer, J. M. F. G. Aerts, G. A. van der Marel, J. D. C. Codée and H. S. Overkleeft, *Eur. J. Org. Chem.*, 2016, **2016**, 3671–3678.

94 F. G. Hansen, E. Bundgaard and R. Madsen, *J. Org. Chem.*, 2005, **70**, 10139–10142.

95 S. G. Withers and K. Umezawa, *Biochem. Biophys. Res. Commun.*, 1991, **177**, 532–537.

96 M. Nakata, C. Chong and Y. Niwata, *J. Antibiot.*, 1993, **46**, 2–5.

97 K. Tatsuta, *Pure Appl. Chem.*, 1996, **68**, 1341–1346.

98 J. Jiang, C. L. Kuo, L. Wu, C. Franke, W. W. Kallemeijn, B. I. Florea, E. van Meel, G. A. van der Marel, J. D. C. Codée, R. G. Boot, G. J. Davies, H. S. Overkleeft and J. M. F. G. Aerts, *ACS Cent. Sci.*, 2016, **2**, 351–358.

99 L. I. Willems, T. J. M. Beenakker, B. Murray, S. Scheij, W. W. Kallemeijn, R. G. Boot, M. Verhoek, W. E. Donker-Koopman, M. J. Ferraz, E. R. Van Rijssel, B. I. Florea, J. D. C. Codée, G. A. Van Der Marel, J. M. F. G. Aerts and H. S. Overkleeft, *J. Am. Chem. Soc.*, 2014, **136**, 11622–11625.

100 L. Wu, J. Jiang, Y. Jin, W. W. Kallemeijn, C.-L. Kuo, M. Artola, W. Dai, C. van Elk, M. van Eijk, G. A. van der Marel, J. D. C. Codée, B. I. Florea, J. M. F. G. Aerts, H. S. Overkleeft and G. J. Davies, *Nat. Chem. Biol.*, 2017, **13**, 867–873.

101 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.

102 W. W. Kallemeijn, K. Y. Li, M. D. Witte, A. R. A. Marques, J. Aten, S. Scheij, J. Jiang, L. I. Willems, T. M. Voorn-Brouwer, C. P. A. A. Van Roomen, R. Ottenhoff, R. G. Boot, H. Van Den Elst, M. T. C. Walvoort, B. I. Florea, J. D. C. Codée, G. A. Van Der Marel, J. M. F. G. Aerts and H. S. Overkleeft, *Angew. Chem. Int. Ed.*, 2012, **51**, 12529–12533.

103 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.

Chapter 2

Synthesis and biochemical evaluation of D-*xylo*-cyclophellitols

Parts of this chapter have been published:
S.P. Schröder *et al.*, A Divergent Synthesis of L-*arabino*- and D-*xylo*-Configured Cyclophellitol Epoxides and Aziridines, *Eur. J. Org. Chem.* **2016**, 4787-4794

2.1 Introduction

Glycosidases are responsible for the cleavage of glycosidic linkages and are key in the turnover of glycans and glycoconjugates – a large and highly diverse class of biopolymers including polysaccharides, glycoproteins and glycolipids.[1,2] The importance of glycosidases in biological and biomedical research is reflected in the numerous studies on the discovery and development of glycosidase inhibitors. These studies are often inspired by nature, and many natural products have been identified as glycosidase inhibitors. One such natural product glycosidase inhibitor is cyclophellitol (Figure 1), a cyclitol epoxide discovered in 1990 by Atsumi and co-workers. They found that this compound, produced by the fungal strain *Phellinus sp*, inhibits retaining β-glucosidase activities by mimicking the natural substrate and

reacting with the active site nucleophile to yield a stable covalent and irreversible adduct.[3–5] Cyclophellitol is selective against retaining β-glucosidases over inverting β-glucosidases, enzymes that do not form a covalent enzyme-substrate adduct and instead directly deliver water to the activated (protonated) glycosidic linkage. This feature, which sets cyclophellitol apart from the competitive glucosidase inhibitor, deoxynojirimycin,[6] has led to a recent increased interest in the compound and its functional and configurational analogues, both in structural and chemical glycobiology studies.

In one early study on cyclophellitol analogues, substitution of the epoxide oxygen for nitrogen led to cyclophellitol aziridine, an equally strong retaining β-glucosidase inhibitor with the additional feature of a secondary amine amenable for chemical modification.[7] Another close analogue comprises conduritol B epoxide (CBE, Figure 1). CBE lacks the C5 methylene compared to cyclophellitol (cyclophellitol numbering is depicted in Figure 1) and is a synthetic retaining glucosidase inhibitor derived from the natural product, conduritol. CBE was first described by Legler and co-workers and has been the glycosidase inhibitor of choice for many years.[8,9] CBE inhibits retaining β-glucosidases considerably less effectively than cyclophellitol, and besides inhibits retaining α-glucosidases as well. Conduritol B aziridine was identified as a mechanism-based glucosidase inactivator by Withers and co-workers[10] and recently used as a starting point for the synthesis of *N*-alkylated glucosidase inhibitors.[11]

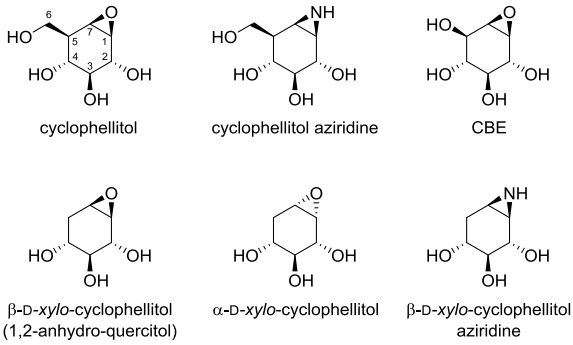
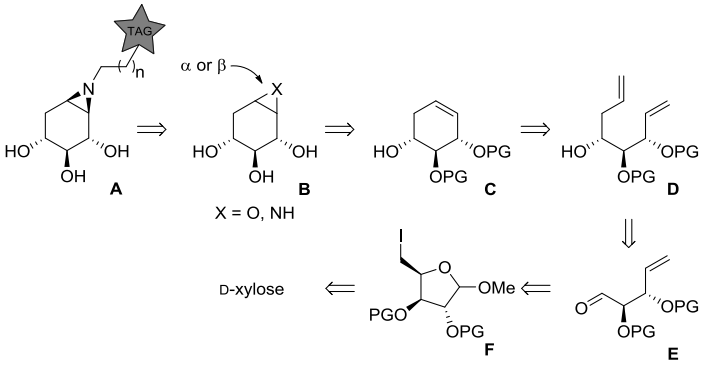


Figure 1 Chemical structures of cyclophellitol and its functional and structural analogues.

Following the discovery of cyclophellitol and CBE, quite a number of structural and configurational analogues have appeared in literature.[5,7,12,13] However, studies on cyclitol epoxides and cyclitol aziridines emulating pentopyranosides (instead of

hexopyranosides) are scarce. Ogawa and co-workers reported the synthesis of β-D-*xylo*-cyclophellitol, referred to as 1,2-anhydro-quercitol (Figure 1).[14] Configurational and functional analogues of β-D-*xylo*-cyclophellitol have not appeared in the literature. In this Chapter, a newly developed synthesis route towards β-D-*xylo*-cyclophellitol as well as the epimeric α-D-*xylo*-cyclophellitol is described. Additionally this route provides access to the aziridine analogues, and the β-aziridine is subsequently functionalized with reporter tags to afford multiple β-D-*xylo*-ABPs. Because these inhibitors and probes are structural analogues of cyclophellitol (lacking the hydroxymethylene moiety at C5), the inhibitory potency and selectivity towards human β-glucosidases GBA1 and GBA2 is investigated. Furthermore, the β-D-*xylo*-ABPs could be used as a tool to monitor β-xylosidase activity in the context of biotechnology research, for example related to food processing, paper production or biofuel development.[15,16,17] For these reasons, the efficiency of the β-D-*xylo*-ABPs to tag and visualize a β-xylosidase was studied and the results of these studies are presented as well.

It was envisioned that β-D-*xylo*-cyclophellitol ABPs **A** (Scheme 1) would be available from the set of epoxide- and aziridine inhibitors **B**. These inhibitors could be synthesized from cyclic alkene **C**, which in turn would emerge from ring-closing metathesis (RCM) of diene **D**. Several studies on asymmetric allylation of aldehyde **E** have appeared in the literature in recent years which include variations in the protective group scheme, the nature of the nucleophile and the stereochemical outcome of the newly introduced chiral center. Aldehyde **E** is readily available by Vasella fragmentation[18] of protected methyl iodofuranoside **F**, which is obtained from D-xylose in 3 steps according to literature.[19]

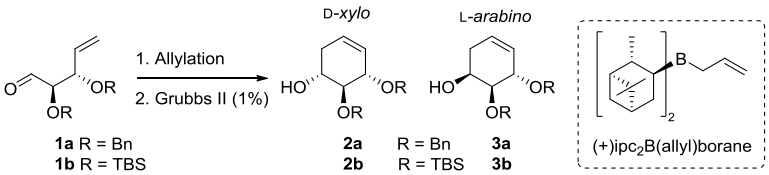


Scheme 1 Retrosynthesis of ABPs **A** and inhibitors **B**.

2.2 Results and Discussion

As the first research objective, optimal conditions were searched for the preparation enantiopure homo-allylic alcohol **C** in pure diastereomeric form and with the optimal protective group pattern for ensuing elaboration to the target epoxides and aziridines. For this, attention was focused on literature reports on nucleophilic allylation of aldehydes **E**. Madsen and co-workers thoroughly studied domino elimination-alkylation reactions of methyl iodofuranoside **F** affording products **D** directly.[20–22] The diastereomeric reaction outcome could be directed by using either zinc or indium to promote alkylation with allyl bromide, affording a diastereomeric mixture of homoallylic alcohol **D** in 1:3 and 7:2 ratios ('down'/'up'), respectively. Apart from zinc- or indium mediated Barbier additions to aldehyde **E**,[23,24] examples of Grignard-like allylations on the carbonyl group are scarce. Isopropylidene-protected aldehyde **E** has been allylated using catalytic asymmetric Keck conditions[25] or by using allyl-zinc reagents,[26] but the isopropylidene protective group would implicate acidic deprotection in the final stages of the synthesis, conditions that may be incompatible with the targeted epoxides and aziridines. It was therefore decided to investigate the efficacy of benzyl-protected alkenal **1a**[19] and *tert*-butyldimethylsilyl-protected alkenal **1b**[27], readily available according to the literature procedures, as the starting materials in the allylation studies. For practical reasons the resulting dienes were not isolated,

Table 1 Stereo-controlled allylation of aldehydes **1a** and **1b**



1a R = Bn
1b R = TBS

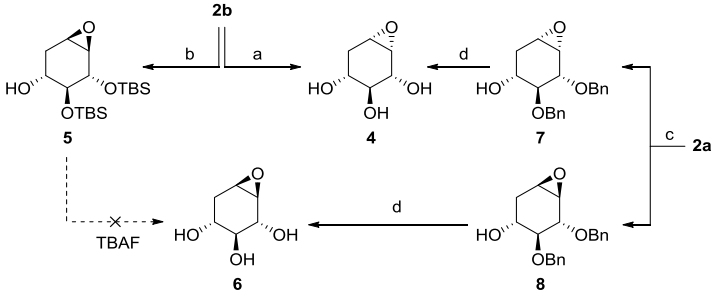
2a R = Bn **3a**
2b R = TBS **3b**

Entry	Compound	Allylation conditions	Ratio **2**:**3**[a]	Yield
1	**1a**	AllylMgBr, THF, 0 °C	1.2:1	90%[b]
2	**1a**	AllylBr, Zn, THF/H_2O, 40 °C	3:1	68%[b]
3	**1a**	(-)-ipc$_2$B(allyl)borane, THF, -90 °C	0:1	71%[c]
4	**1a**	(+)-ipc$_2$B(allyl)borane, THF, -90 °C	9:1	76%[c]
5	**1b**	(+)-ipc$_2$B(allyl)borane, THF, -90 °C	1:0	71%[c]

[a] Crude product ratio determined by ^{1}H-NMR. [b] Combined isolated yield. [c] Isolated yield of major isomer.

instead the crude dienes were directly submitted to RCM to furnish matching cyclohexenes **2** and/or **3**. Reaction of **1a** with allyl magnesium bromide resulted in a nearly 1:1 mixture of diastereomers (Table 1, entry 1). Reacting **1a** with allyl bromide and zinc under sonication (entry 2) resulted in a 3:1 preference for the formation of alkene **2a**. Ultimately it was found that with the use of the (-)-enantiomer of Brown's allylborane reagent, (-)-ipc$_2$B(allyl)borane (entry 3),[28] cyclohexene **3a** was formed in 71% yield, without the formation of observable amounts of **2a**, findings that align with the stereoselective γ-silylallylboration of aldehyde **1a** reported by Heo *et al.*[29] Similarly, by using (+)-ipc$_2$B(allyl)borane (entry 4), cyclohexene **2a** was obtained in 9:1 ratio and 76% yield. Additionally, under these same reaction conditions the target cyclic alkene **2b** could be obtained as a single isomer, when TBDMS groups were used as protecting groups (entry 5).

As the next objective, the transformation of cyclohexenes **2a** and **2b** into the corresponding epoxides was studied (Scheme 2). The α-epoxide **4** could be obtained selectively by employing a two-step procedure involving the complete deprotection of cyclohexene **2b**, followed by *m*-CPBA oxidation directed by the allylic alcohol. Next, a method to selectively afford β-epoxide **6** was investigated. Whereas reaction of **2b** with *m*-CPBA resulted in poor conversion, the cyclohexene could be easily epoxidized with dimethyldioxirane (DMDO) to afford β-epoxide **5** as a single isomer. Although the

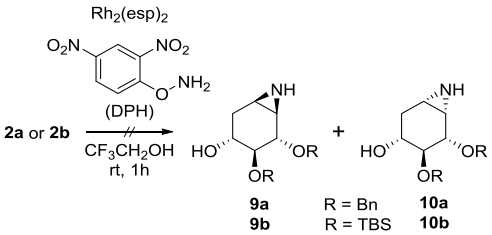


Scheme 2 Synthesis of epoxides **4** and **6**. Reagents and conditions: a) HCl, MeOH, 16h, then *m*-CPBA, $NaHCO_3$, MeOH, 16h, 64%; b) dimethyldioxirane, acetone, 2h, 77%; c) oxone, CH_3COCF_3, $NaHCO_3$, EDTA, MeCN, H_2O, 0 °C, 1h, **7**: 23%, **8**: 57%; d) $Pd(OH)_2/C$, H_2, MeOH, H_2O, dioxane, **4**: 62%, **6**: 74%.

product could be smoothly deprotected by TBAF, it proved to be troublesome to afford β-epoxide **6** in pure state due to contamination by highly polar ammonium salts. As an alternative, it was found that when benzyl protected cyclohexene **2a** was

epoxidized by methyl(trifluoromethyl)dioxirane (generated *in situ* from 1,1,1-trifluoroacetone and oxone[30]), epoxides **7** and **8** were obtained as a separable mixture. Subsequent deprotection of the benzyl groups in epoxides **7** and **8** with Pearlman's catalyst under H_2 atmosphere resulted in D-*xylo*-cyclophellitols **4** and **6**, respectively.

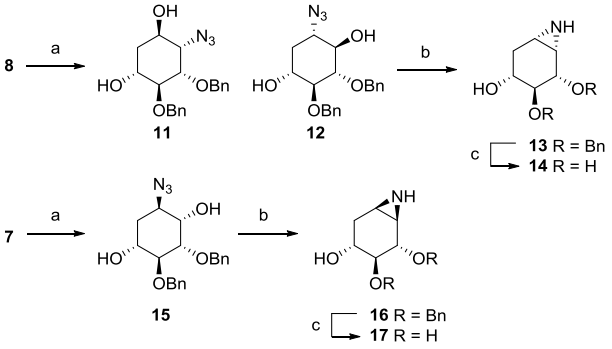
Next, the versatility of cyclohexenes **2a** and **2b** as starting materials for the synthesis of the target aziridines by means of direct aziridination was investigated. A few years ago, Jat *et al.* described the direct stereospecific synthesis of unprotected aziridines from alkenes, using $Rh_2(esp)_2$ and *O*-(2,4-dinitrophenyl)hydroxylamine (DPH) in trifluoroethanol.[31] Besides linear olefins, also cyclic alkenes were readily converted to the aziridines, and in the case of cholesterol, a directing effect of an adjacent hydroxyl group on the stereospecific delivery of the aziridine was suggested. Unfortunately, cyclohexene **2a** proved to be quite unreactive under these conditions, and it was found that high loading of catalyst (25 mol %) and aminating agent (4 equivalents) was required to effectuate full consumption of the starting material (Scheme 3). Under these conditions, aziridines **9a** and **10a** were produced as a mixture in low yield and purity. Unfortunately, similar results were observed when cyclohexene **2b** was subjected to the same reaction conditions. Therefore, further optimization of the reaction conditions was not pursued.



Scheme 3 Direct aziridination of cyclohexenes **2a** or **2b** with *O*-(2,4-dinitrophenyl)hydroxylamine (DPH) under rhodium catalysis was found unsuccessful.

In an alternative attempt to obtain the target aziridines, their synthesis from the corresponding epimeric epoxides was investigated (Scheme 4). For this purpose, the benzyl-protected epoxides were chosen as starting compounds to prevent protecting group migrations during the subsequent reactions, as well to ensure facile deprotection towards the final compounds. Epoxide **8** was reacted with sodium azide in the presence of lithium perchlorate at elevated temperature to afford a mixture of

azido-alcohols **11** and **12**. Interestingly, in contrast to the Fürst-Plattner rule, compound **12** was formed in favor of **11**. While the exact reason for this outcome is not known, it could be caused by a Lewis-acidic chelation effect of lithium due to the excess lithium perchlorate employed. A similar behavior was found with azidolysis of benzyl-protected CBE under the presence of $LiClO_4$.[32] The mixture of azido-alcohols was subsequently ring-closed under anhydrous Staudinger conditions using polymer-bound triphenylphosphine,[33] resulting in aziridine **13**, the benzyl ethers of which were removed under Birch conditions to afford D-*xylo*-cyclophellitol aziridine **14**. Similarly, epoxide **7** was treated with sodium azide and in this case the Fürst-Plattner rule was obeyed and thus only a single regioisomer **15** was formed. Ring-closure of this azido-alcohol resulted in aziridine **16**, and the benzyl groups were removed under Birch conditions resulting in D-*xylo*-cyclophellitol aziridine **17**.

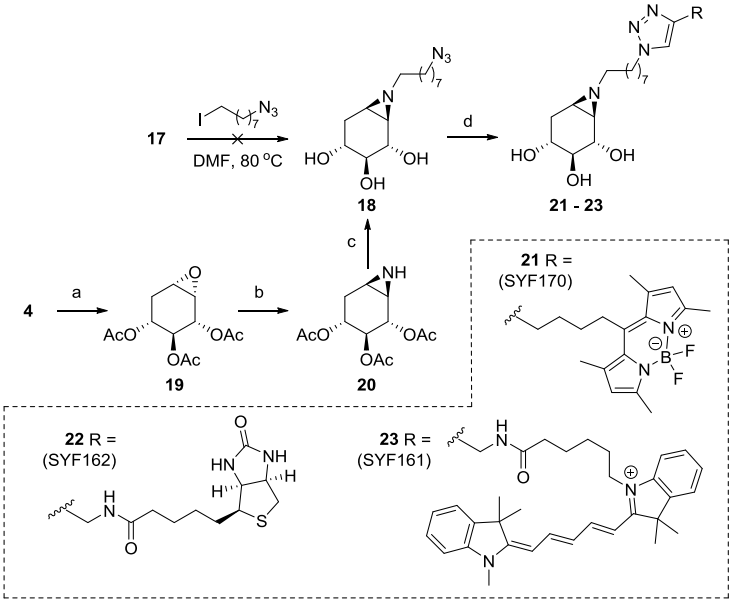


Scheme 4 Synthesis of aziridines **14** and **17** from their parent epoxides. Reagents and conditions: a) NaN_3, $LiClO_4$, DMF, 100 °C, 16h, yield **11**: 33%, yield **12**: 50%, yield **15**: 80%; b) polymer-bound triphenylphosphine, MeCN, 60 °C, 16h, yield **13**: 55%, yield **16**: 77%; c) Li, NH_3, THF, -60 °C, 1h, yield **14**: 96%, yield **17**: 73%.

Having obtained the target aziridine inhibitors, the synthesis route towards β-D-xylo-cyclophellitol aziridine ABPs was investigated (Scheme 5). Direct alkylation of the unprotected aziridine with the appropriate 1-azido-8-iodooctane linker resulted in a complex reaction mixture. An alternative synthetic route towards **18** was then demanded allowing facile alkylation of the aziridine. It was found that protected aziridines could be easily alkylated with alkyl-triflates in chloroform, in high yields without occurrence of side reactions. However, direct alkylation of aziridine **16** would necessitate the removal of the benzyl protecting groups in the presence of the azide linker in the final stage of the synthesis. As azides are prone to undergo reduction

under the standard reductive conditions of benzyl deprotections (i.e. Birch conditions, palladium catalyzed hydrogenation), other protective groups for the secondary alcohols were chosen. Therefore, α-epoxide **4** was protected with acetyl groups and the epoxide was opened with sodium azide under elevated temperatures to give a mixture of azido-alcohols, which was then subjected to Staudinger-type ring closure to give the target β-aziridine **20** in low yield. This aziridine could then be smoothly alkylated by the appropriate alkyl-triflate followed by deprotection with sodium methanolate affording alkylated aziridine **18**. The azide handle could then be functionalized by click-chemistry with different fluorogenic tags and a biotin moiety to afford β-D-xylo-cyclophellitol aziridine ABPs **21**-**23** after HPLC purification.

With β-D-*xylo*-cyclophellitol ABPs **21**-**23** in hand, their labeling efficiency towards recombinant lysosomal β-glucocerebrosidase (GBA1) was investigated, and compared to known *gluco*-configured cyclophellitol aziridine probes. The enzyme was incubated



Scheme 5 Synthesis of β-D-*xylo*-cyclophellitol aziridine ABPs **21**-**23**. Reagents and conditions: a) Ac_2O, DMAP, pyridine, rt, 16h, 88%; b) 1. NaN_3, $Et_3N.HCl$, 80 °C, 16h; 2. polymer-bound triphenylphosphine, 60 °C, 24h, 13% over 2 steps; c) 1. 8-azidooctyl trifluoromethanesulfonate, DIPEA, DCM, rt, 24h; 2. NaOMe, MeOH, rt, 16h, 73% over 2 steps; d) tag-alkyne, $CuSO_4$, sodium ascorbate, DMF/H_2O, rt, 16h, **21**: 58%; **22**: 82%, **23**: 58%.

in McIlvaine buffer pH 5.2 with different concentrations of **21** or **23** (Figure 2a) for 30 minutes at 37 °C and subsequently the remaining unreacted enzyme was post-labeled by 10 nM β-D-*gluco*-cyclophellitol ABPs JJB367[34] or JJB70[35] (Figure 2d), respectively. From these experiments, it is apparent that **21** and **23** label GBA1 in a concentration dependent manner, and that the labeling is activity-based. Indeed, it is known that GBA1 displays activity towards 4-methylumbelliferyl β-D-xylopyranoside,[36] and the reduced potency (compared to the β-D-*gluco*-cyclophellitol ABPs) presumably arises from the absence of the C5-hydroxymethylene group which may be important for the stabilization of the enzyme-inhibitor transition state complex. Interestingly, while the optimal probe concentration for labeling GBA1 with BODIPY-green tagged **21** is approximately 1.0 μM, the Cy5 analogue **23** does not accomplish full enzyme labeling at this concentration, and the optimum lies between 1 and 10 μM.

In addition to these studies, a competitive assay was executed with the complementary D-*xylo*-cyclophellitol inhibitors **4, 6**, **14, 17**, **18**, **21** and **22** (Figure 2b). Thus, recombinant GBA1 was pre-incubated with different concentrations of inhibitor, followed by labeling of the residual active enzyme with 2 μM of **23**. Nearly complete competition was achieved by pre-incubation with 100 μM β-epoxide **6**. The β-aziridine **17** was approximately 10 times more potent, as full inhibition was achieved around 10 μM. Alkylation of the aziridine (**18**) further increased its potency, however it was found that click ligation of the azidoalkyl linker with a BODIPY-green (**21**) or biotin (**22**) reporter tag reduced its inhibitory activity. Finally, in line with the enzymatic catalytic mechanism, the α-epoxide **4** and -aziridine **14** did not react with the enzyme. The IC_{50} values for inhibition of GBA1 were also determined and are in line with the competition experiments (Figure 2c).

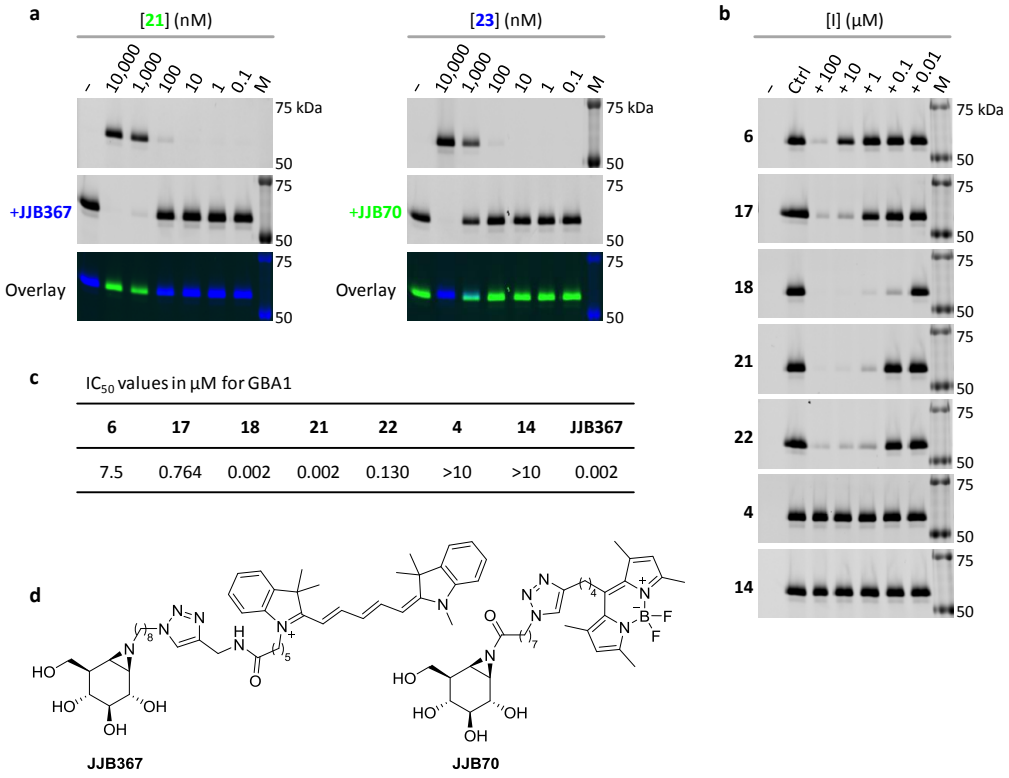


Figure 2 a) determination of the concentration optimum for labeling recombinant GBA1 with ABP **21** (left) and **23** (right). rGBA (1 pmol) was incubated with different concentrations of *xylo*-ABP at pH 5.2 for 30 min at 37 °C and subsequently incubated with 10 nM JJB367 or JJB70. b) Competition assay with various α/β-D-xylo-cyclophellitol epoxides and aziridines. c) IC_{50} values of these inhibitors for GBA1. Values reported are averages from two measurement using β-4-MU-glucopyranoside as substrate at pH 5.2. d) Chemical structures of β-glucosidase selective ABPs JJB367 and JJB70.

Next, the labeling of **23** in mouse liver lysate (C57bl/6j, Jackson's laboratories) was investigated (Figure 3). In this experiment, lysate was pre-incubated with different concentrations of β-epoxide (Figure 3a), β-aziridine (Figure 3b), β-alkyl aziridine (Figure 3c) or BODIPY-aziridine (Figure 3d). These competitors contained either the D-*gluco* (top row) or D-*xylo* (bottom row) configuration. After pre-incubation, the remainder of active enzymes was post-labeled by incubation with either 2 μM JJB367

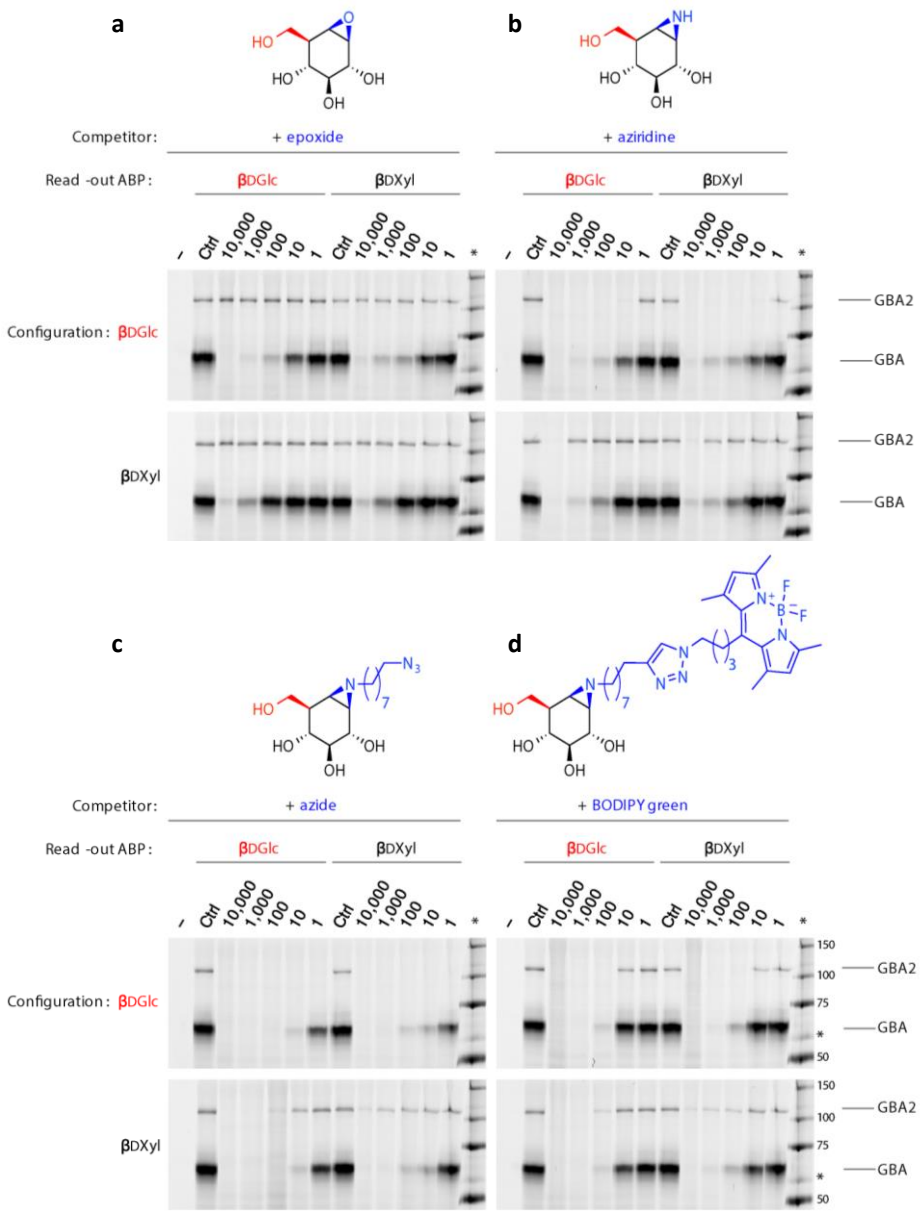


Figure 3 Competition assay of covalent D-*xylo* or D-*gluco* configured cyclophellitols in mouse liver lysate. The lysate was pre-incubated with the cyclitol inhibitor depicted above the gel (top gel: D-*gluco* configuration, bottom gel: D-*xylo* configuration) with different concentrations, and subsequently labeled with D-*gluco* (left) or D-*xylo* configured cyclophellitol aziridine ABP (right).

(left panel) or **23** (right panel). Incubation with **23** resulted in the labeling of GBA1, as well as GBA2 (control lanes). Moreover, the labeling intensity of both enzymes appeared comparable to that of JJB367 at this concentration. Both D-*gluco* and D-*xylo* derivatives selectively inhibit GBA1 over GBA2 (Figure 3a), with the D-*gluco* compound being the most potent of the two. The β-aziridines display similar potency towards GBA1 (Figure 3b). However, the D-*gluco* configured aziridine simultaneously inhibits GBA2, whereas the D-*xylo* aziridine only inhibits this enzyme at the highest concentration tested (10 μM). Subsequently, functionalization of the aziridines with an azidoalkyl spacer results in an increased potency towards GBA1 and GBA2 for both configurations (Figure 3c). This trend is retained for the aziridines functionalized with a BODIPY fluorophore, while the inhibitory potency appears to be slightly reduced for both compounds (Figure 3d).

Lastly, mouse liver, brain and duodenum tissue lysates (C57bl/6j, Jackson's laboratories) were labeled with the biotinylated D-*xylo* ABP and the hits were analyzed by proteomics. For this purpose, the lysates were incubated with **22** with or without pre-incubation with **23** (competition), the resulting biotinylated proteins were enriched with magnetic streptavidin beads, and following tryptic digestion of the

Table 2 Pull-down analysis by LC-MS/MS proteomics of mouse brain, duodenum and liver tissue lysates labeled by biotinylated ABP **22**.

	Probe	Protein	Accession code	Mass (kDa)	Protein score	Unique peptides	Sequence coverage
Brain	DMSO	-	-	-	-	-	-
	23 → 22	-	-	-	-	-	-
	22	GBA1	P17439	58	1.536	6	12 %
Duodenum	DMSO	-	-	-	-	-	-
	23 → 22	-	-	-	-	-	-
	22	GBA1	P17439	58	25.884	9	24 %
Liver	DMSO	GANAB	Q8BHN3	107	1.697	12	15 %
		Glucosidase II	O08795	60	1.206	4	5 %
	23 → 22	-	-	-	-	-	-
	22	GBA1	P17439	58	75.141	21	66 %
		GBA2	Q69ZF3	105	2.409	9	15 %

proteins the peptide fragments were analyzed by nano LC-MS/MS (Table 2). Pull-down analysis of **22** in mouse brain lysates clearly identified the labeling of GBA1, which was positively absent in the competition experiment and negative control. Likewise, labeling with **22** in mouse duodenum identified GBA1 as single glycosidase hit. Finally, a pull-down was performed on mouse liver lysate. This resulted in the identification of GBA1 as well as GBA2, corroborating with the bands labeled by **23** (Figure 3).

The ability of D-*xylo* ABP **23** to label recombinant GH52 β-xylosidase from *Opitutus terrae* (PB90-1, ~81 kDa) was investigated next by incubation of the enzyme with the probe at pH 6.8, 37 °C for 30 minutes with increasing concentrations of **23** (Figure 4a). Optimal labeling was achieved at 500 nM. Furthermore the optimal pH for labeling was determined by incubating the enzyme with the probe at different pH values (Figure 4b), and the most profound labeling was observed at pH 6. A similar experiment was performed, varying the incubation temperature, and optimal labeling occurred at 50 °C (Figure 4c). Interestingly, in contrast with this temperature optimum, *O. terrae* strains are unable to grow at temperatures above 37 °C.[37] Lastly, a competition assay was performed. Pre-incubation of the enzyme with different concentrations of covalent inhibitor followed by labeling with 500 nM **23** shows that β-epoxide **6** is unable to fully inactivate the enzyme at the highest concentration (100 μM) employed (Figure 4d). The β-aziridine **17** proved to be a superior competitor, and alkylation of the aziridine (**18**) further improved its potency. Similarly to the labeling of GBA1 (Figure 2b), the potency was slightly reduced when the aziridine was equipped with a biotin scaffold (**22**). Minor competition was observed by α-epoxide **4** at the highest concentration applied, while cyclophellitol was totally inactive, in line with the enzymatic activity.

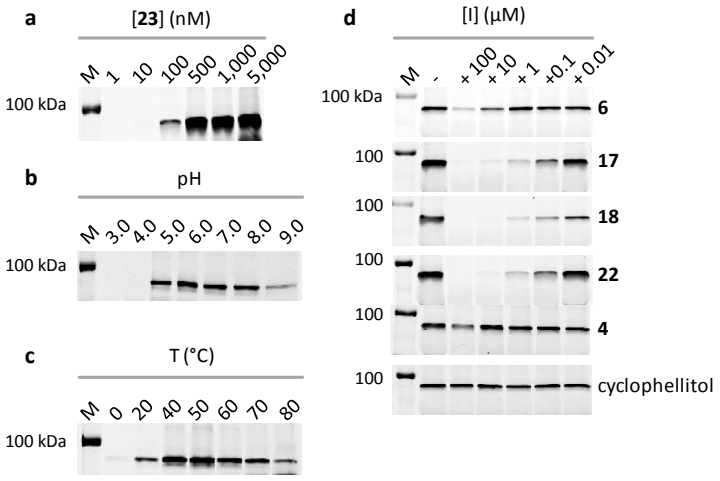


Figure 4 Activity-based labeling of β-xylosidase from *Opitutus terrae* (PB90-1, 50 ng). a) Concentration optimum (50 mM phosphate buffer pH 6.8 at 37 °C, 30 minutes). Optimal probe concentration of 500 nM was used for further screening. b) pH optimum. c) Temperature screening. d) Competitive ABPP assay.

2.3 Conclusion

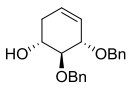
In summary, this Chapter describes the synthesis of D-*xylo*-cyclophellitol, a pentopyranoside-configured analogue of cyclophellitol. The key synthetic step in the synthesis involves the asymmetric allylation of alkenal **1a** towards D-*xylo*-configured cyclohexene **2a**, which is under full diastereomeric control by choice of allylating agent. Additionally, the synthesis of structural analogues of D-*xylo*-cyclophellitol is described bearing an aziridine as electrophilic warhead, and the aziridine is further equipped with several reporter tags to afford D-*xylo*-cyclophellitol activity-based probes. While these probes only partially resemble the natural substrate (lacking the C5 hydroxymethylene group), they are activity-based inactivators of GBA1, albeit with reduced potency in comparison with cyclophellitol and its analogues. Probe **23** selectively labels GBA1 and GBA2 in mouse liver lysate, and through a competition assay it was found that the D-*xylo*-cyclophellitols display similar selectivity towards GBA1 and GBA2 compared to the cyclophellitols, except for D-*xylo*-cyclophellitol aziridine **17** which, in contrast to cyclophellitol aziridine, shows higher activity towards GBA1 over GBA2. Pull-down analysis with biotinylated ABP **22** positively identified GBA1 and GBA2 as unique glycosidase hits in this lysate. Lastly, it was shown that D-*xylo*-cyclophellitols are activity-based covalent inhibitors of β-xylosidase from *O. terrae*, and using ABP **23** the temperature and pH dependence of labeling by the enzyme was visualized, and the potency of these D-*xylo*-cyclophellitols was determined by a competition assay.

The D-*xylo*-cyclophellitol aziridine ABPs could be used for labeling of GBA1 and GBA2 in biological samples, however the configurationally matching D-*gluco*-cyclophellitol aziridine ABP JJB367 proved to be superior for this purpose. Instead, D-*xylo* ABPs **21**-**23** may be used for identification of retaining *exo*-β-xylosidases in biologically relevant samples. Additionally, these ABPs could be used to screen for β-xylosidases with extreme pH and/or temperature tolerance, thereby providing a potential useful tool for the food, paper and biofuel industries.[15,16,17]

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).

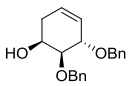
Compound 2a



To a stirred solution of aldehyde **1a**[19] (960 mg, 3.25 mmol) in dry THF (33 mL) was added (+)-ipc$_2$B(allyl)borane (1M in pentane, 4.9 mL) slowly at -90 °C under inert atmosphere. The reaction mixture was stirred for 1 h at -90 °C. After allowing the solution to reach rt, MeOH (6 mL) was added followed by the slow addition of 30% H_2O_2 (20 mL) and sodium-phosphate buffer (0.1 M, pH 7, 20 mL). The biphasic solution was separated, and the aqueous layer was extracted with EtOAc (3 x 75 mL). The combined organic phases were then washed with aq. 10% $Na_2S_2O_3$, water and brine, dried over $MgSO_4$, filtered and evaporated. The crude product was dissolved in dry degassed DCM (13 mL) under argon, and Grubbs 2nd generation catalyst (1% mol) was added. The reaction was refluxed (40 °C) in the dark for 15 h. The solvent was evaporated and flash purification by silica column chromatography (pentane/EtOAc, 9:1) gave product **2a** as colorless oil (760 mg, 76% over 2 steps). ^1H NMR (400 MHz, $CDCl_3$): δ 7.61 – 7.11 (m, 10H), 5.70 (s,

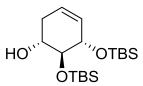
2H), 4.99 (d, *J* = 11.4 Hz, 1H), 4.80 – 4.52 (m, 3H), 4.17 (d, *J* = 6.6 Hz, 1H), 3.80 (dd, *J* = 15.3, 9.0 Hz, 1H), 3.65 – 3.50 (t, *J* = 8.2 Hz, 1H), 2.85 (s, 1H), 2.47 (dd, *J* = 16.2, 4.8 Hz, 1H), 2.23 – 2.04 (m, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 138.6, 138.3, 128.6, 128.5, 128.4, 128.0, 127.9, 127.9, 127.8, 127.7, 126.4, 126.0, 83.5, 80.2, 74.6, 71.4, 68.8, 32.5 ppm. IR: (neat) ν 3350, 3030, 1453, 1352, 1060, 734 cm^{-1}. $[\alpha]D_{20}$ (c = 0.4, DCM): +50. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{20}H_{22}O_3$ 311.16417, found 311.16429.

Compound 3a



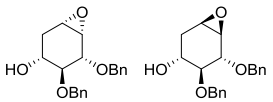
The reaction was carried out following the same procedure described for **2a**, employing aldehyde **1a**[19] (600 mg, 2.02 mmol) and (-)-ipc$_2$B(allyl)borane to give the product as colorless oil (444 mg, 71% over 2 steps). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.40 – 7.26 (m, 10H), 5.73 (s, 2H), 4.80 – 4.53 (m, 4H), 4.24 – 4.13 (m, 2H), 3.71 (dd, *J* = 5.8, 2.5 Hz, 1H), 2.48 – 2.20 (m, 2H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 138.6, 138.4, 128.6, 128.5, 127.9, 127.9, 127.9, 127.8, 126.9, 125.3, 80.4, 75.6, 72.3, 71.9, 66.8, 31.40 ppm. IR: (neat) υ 3439, 3030, 2870, 1496, 1452, 1066, 1026, 732, 694 cm^{-1}. $[\alpha]D_{20}$ (c = 0.1, $CHCl_3$): +84. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{20}H_{22}O_3$ 311.16417, found 311.16435.

Compound 2b



The reaction was carried out following the same procedure described for **2a**, employing aldehyde **1b**[27] (2.22 g, 6.44 mmol) and (+)-ipc$_2$B(allyl)borane to give the product as colorless oil (1.62 g, 71% over 2 steps). ^{1}H NMR (400 MHz, $CDCl_3$): δ 5.72 (dt, *J* = 10.1, 3.5 Hz, 1H), 5.67 – 5.60 (m, 1H), 3.98 (s, 1H), 3.83 (s, 2H), 3.52 (d, *J* = 7.1 Hz, 1H), 2.45 – 2.36 (m, 1H), 2.18 – 2.10 (m, 1H), 0.90 (s, 9H), 0.89 (s, 9H), 0.12 (s, 3H), 0.11 (s, 3H), 0.11 (s, 3H), 0.10 (s, 3H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 126.4, 126.2, 72.8, 70.4, 68.9, 30.5, 25.8, 25.7, 18.0, 17.9, -4.4, -4.6, -4.7 ppm. IR: (neat) υ 3512, 2927, 1471, 1251, 1076, 833 cm^{-1}. $[\alpha]D_{20}$ (c = 0.3, DCM): +71. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{18}H_{38}O_3Si_2$ 381.2252, found 381.2260.

Compound 7 and 8

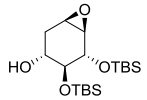


7 8

To a stirred solution of cyclohexene **2a** (1.0 g, 3.22 mmol) in MeCN/0.4 mM aq. EDTA (2:1 v/v, 32 mL), 1,1,1-trifluoroacetone (3.2 mL, 35.4 mmol) was added at 0 °C. Subsequently a solid mixture of $NaHCO_3$ (2.2 g, 25.8 mmol) and oxone (10.9 g, 35.44 mmol) was added in 6 portions over 60 min and the reaction was left to stir at 0 °C for 1 h. Then, water (300 mL) was added and the crude products were extracted with EtOAc (3 x 150 mL). The combined organic phases were washed with brine, dried over $MgSO_4$, filtered and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 4:1) afforded pure **7** and **8** as colorless oils (**7**: 244 mg, 23%; **8**: 601 mg, 57%). <u>For α-isomer **7**:</u> ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.52 – 7.15 (m, 10H), 4.91 (d, *J* = 11.3 Hz, 1H), 4.83 (d, *J* = 11.9 Hz, 1H), 4.77 (d, *J* = 11.9 Hz, 1H), 4.61 (d, *J* = 11.3 Hz, 1H), 3.88 (dd, *J* = 6.7, 2.3 Hz, 1H), 3.67 – 3.47 (m, 2H), 3.36 (dd, *J* = 3.7, 2.3 Hz, 1H), 3.26 (t, *J* = 4.0 Hz, 1H), 2.68 (br s, OH), 2.35 (ddd, *J* = 15.4, 6.2, 4.6 Hz, 1H), 2.01 (dd, *J* = 15.3, 8.6 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 138.3,

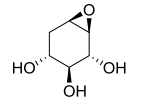
138.1, 128.7, 128.6, 128.1, 128.0, 80.7, 78.8, 75.0, 72.0, 68.4, 54.1, 51.6, 29.9 ppm. IR: (neat) υ 3300, 1103, 1094, 1073 cm^{-1}. $[\alpha]D_{20}$ (c = 0.5, DCM): +24. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{20}H_{22}O_4$ 349.14103, found 349.14087. For β-isomer **8**: ^1H NMR (400 MHz, $CDCl_3$): δ 7.53 – 7.01 (m, 10H), 4.96 (d, *J* = 11.3 Hz, 1H), 4.82 (d, *J* = 11.3 Hz, 1H), 4.73 – 4.60 (m, 2H), 3.83 (d, *J* = 8.0 Hz, 1H), 3.58 (td, *J* = 10.3, 5.4 Hz, 1H), 3.35 – 3.24 (m, 2H), 3.16 (d, *J* = 3.6 Hz, 1H), 2.60 (ddd, *J* = 14.6, 5.3, 1.8 Hz, 1H), 2.50 (br s, OH), 1.81 (ddd, *J* = 14.6, 10.6, 1.6 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 138.4, 137.6, 128.8, 128.7, 128.2, 128.1, 128.1, 128.0, 84.3, 80.0, 75.0, 72.7, 65.2, 54.0, 53.6, 31.3 ppm. IR: (neat) υ 3350, 3031, 1110, 1070, 1048 cm^{-1}. $[\alpha]D_{20}$ (c = 0.4, DCM): +23. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{20}H_{22}O_4$ 349.14103, found 349.14111.

Compound 5



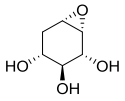
Dimethyldioxirane (DMDO) was freshly prepared according to the procedure by Murray and Singh.[38] Cyclohexene **2b** (72 mg, 0.2 mmol) was dissolved in acetone (3.8 mL) and cooled to 0 °C. Then, a solution of dimethyldioxirane (0.08 M in acetone, 3.8 mL) was added dropwise, and the mixture was stirred for 3 h at rt. Then the mixture was concentrated, and flash purification by silica column chromatography (pentane/EtOAc 15:1) afforded the product as a colorless oil (58 mg, 77%). ^1H NMR (400 MHz, $CDCl_3$): δ 3.81 (d, *J* = 6.7 Hz, 1H), 3.50 – 3.43 (m, 1H), 3.39 (dd, *J* = 8.4, 6.7 Hz, 1H), 3.21 (s, 1H), 3.00 (d, *J* = 3.6 Hz, 1H), 2.47 (dd, *J* = 14.9, 4.1 Hz, 1H), 2.43 (d, *J* = 3.4 Hz, 1H), 1.86 (ddd, *J* = 14.9, 8.8, 2.8 Hz, 1H), 0.94 (s, 9H), 0.90 (s, 9H), 0.18 (s, 3H), 0.14 (s, 3H), 0.11 (s, 3H), 0.09 (s, 3H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 77.0, 72.5, 66.7, 56.5, 52.2, 30.9, 26.0, 25.9, 18.1, 18.1, -3.5, -4.1, -4.4, -4.6 ppm. IR: (neat) υ 3439, 2929, 1766, 1249, 1093, 835 cm^{-1}. $[\alpha]D_{20}$ (c = 1.0, DCM): +11. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{18}H_{38}O_4Si_2$ 397.2201, found 397.2205.

Compound 6[14]



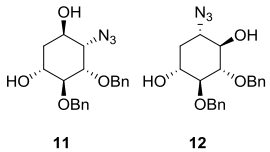
Compound **8** (20 mg, 0.06 mmol) was dissolved in a mixture of MeOH:H_2O:1,4-dioxane (1:1:1, 1.2 mL) and $Pd(OH)_2/C$ (10 mol%) was added under argon. The solution was flushed with H_2 and stirred under H_2 (1 atm) for 5 h at RT. The mixture was filtered over a Celite pad and the solvent was evaporated. The crude deprotected epoxide was absorbed on silica gel and purified by silica column chromatography (DCM/MeOH, 9:1), affording product **6** as a colorless oil (6.6 mg, 74%). ^1H NMR (400 MHz, $CDCl_3$): δ 3.78 (d, *J* = 8.3 Hz, 1H), 3.53 – 3.40 (m, 2H), 3.28 (dd, *J* = 10.2, 8.3 Hz, 1H), 3.21 (d, *J* = 3.6 Hz, 1H), 2.56 (ddd, *J* = 14.9, 5.3, 1.9 Hz, 1H), 1.85 (ddd, *J* = 14.9, 10.6, 1.7 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 76.6, 71.2, 65.5, 56.6, 54.9, 31.2 ppm. IR: (neat) υ 3378, 3320 cm^{-1}. $[\alpha]D_{20}$ (c = 0.1, MeOH): -20. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_6H_{10}O_4$ 147.06519, found 147.06521.

Compound 4[14]



Starting from epoxide **7** (40 mg, 0.12 mmol), the reaction was carried out following the procedure described for **6** to afford the product as a colorless oil (11 mg, 62%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 3.89 (dd, *J* = 8.6, 1.9 Hz, 1H), 3.53 (td, *J* = 10.3, 7.4 Hz, 1H), 3.41 – 3.37 (m, 1H), 3.37 – 3.27 (m, 2H), 2.43 (ddd, *J* = 15.4, 7.2, 5.4 Hz, 1H), 1.78 (dd, *J* = 15.5, 10.4 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 73.7, 71.6, 68.5, 57.5, 52.2, 30.8 ppm. IR: (neat) υ 3469, 3362, 3209 cm^{-1}. [α]D_{20} (c = 0.2, MeOH): +12. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_6H_{10}O_4$ 147.06519, found 147.06519.

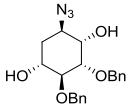
Compound 11 and 12



11 **12**

To a solution of epoxide **8** (200 mg, 0.62 mmol) in dry DMF (6.2 mL), NaN_3 (200 mg, 3 mmol) and $LiClO_4$ (124 mg, 1.2 mmol) were added and the reaction mixture was heated at 100 °C for 18 h under inert atmosphere. H_2O (60 mL) was added, the crude product was extracted with EtOAc (3 x 30 mL) and then the combined organic fractions were washed with H_2O and brine. The organic layer was dried over $MgSO_4$, filtered and concentrated. Flash purification by silica column chromatography (pentane/EtOAc 8:2) afforded the desired azido-alcohols (**11**: 79 mg, 33%; **12**: 111 mg, 50%) as yellow oils. For isomer **11**: ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.40 – 7.21 (m, 10H), 4.80 – 4.51 (m, 4H), 4.06 (td, *J* = 7.2, 4.0 Hz, 1H), 3.98 – 3.86 (m, 2H), 3.67 (m, 2H), 3.05 (br s, OH), 2.40 (br s, OH), 1.96 (ddd, *J* = 10.9, 6.9, 3.9 Hz, 1H), 1.87 (ddd, *J* = 13.7, 7.4, 4.1 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 138.1, 137.2, 128.7, 128.7, 128.3, 128.3, 128.1, 127.9, 79.8, 78.9 (broad signal, assigned by HSQC), 73.9, 73.8, 68.9, 66.4, 64.4, 35.4 ppm. IR: (neat) υ 3380, 2102 cm^{-1}. [α]D_{20} (c = 0.5, DCM): -4. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{20}H_{23}N_3O_4$ 329.15808, found 392.15802. For isomer **12**: ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.42 – 7.26 (m, 10H), 4.97 (d, *J* = 11.3 Hz, 1H), 4.93 (d, *J* = 11.2 Hz, 1H), 4.78 (d, *J* = 11.2 Hz, 1H), 4.70 (d, *J* = 11.3 Hz, 1H), 3.56 (ddd, *J* = 11.9, 8.7, 4.6 Hz, 1H), 3.49 (t, *J* = 9.1 Hz, 1H), 3.41 – 3.20 (m, 3H), 2.63 (br s, OH), 2.36 (br s, OH), 2.17 (dt, *J* = 12.9, 4.6 Hz, 1H), 1.39 (dd, *J* = 24.8, 12.4 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 138.3, 138.2, 128.9, 128.8, 128.2, 128.2, 128.0, 85.5, 83.6, 76.8, 75.7, 75.6, 69.2, 60.0, 33.7 ppm. IR: (neat) υ 3380, 2109 cm^{-1}. [α]D_{20} (c = 0.5, DCM): -51. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{20}H_{23}N_3O_4$ 392.15808, found 392.15797.

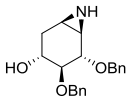
Compound 15



Starting from epoxide **7** (244 mg, 0.75 mmol), the reaction was carried out following the same procedure described for **8** to afford the desired azido-alcohol as a single isomer (222 mg, 80%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.41 – 7.23 (m, 10H), 4.77 – 4.56 (m, 4H), 3.88 (m, 3H), 3.81 (dd, *J* = 6.4, 1.5 Hz, 1H), 3.67 (t, *J* = 6.2 Hz, 1H), 2.91 (br s, OH), 2.47 (br s, OH), 2.14 – 1.99 (m, 1H), 1.92 (ddd, *J* = 13.8, 8.0, 3.9 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 138.1, 137.2, 128.9, 128.7, 128.5, 128.3, 128.1, 127.9, 80.2, 78.9 (broad signal, assigned by HSQC), 74.0, 73.8,

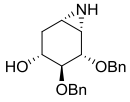
70.6, 68.5, 58.8, 31.7 ppm. IR: (neat) υ 3437, 2916, 2100, 1257, 1074 cm^{-1}. [α]D_{20} (c = 0.4, DCM): -18. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{20}H_{23}N_3O_4$ 392.15808, found 392.15802.

Compound 16



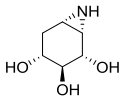
Compound **15** (506 mg, 1.37 mmol) was dissolved in dry CH_3CN (14 mL) and polymer-bound PPh_3 (3 mmol/g loading, 913 mg, 2.74 mmol) was subsequently added to the solution. The reaction was left to stir for 15 h at 60 °C under inert atmosphere. Then the beads were removed by filtration, the organic solvent was evaporated and flash purification by silica column chromatography (DCM/MeOH 49:1) afforded the product as a colorless oil (361 mg, 81%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.43 – 7.22 (m, 10H), 4.94 (d, *J* = 11.4 Hz, 1H), 4.78 (d, *J* = 11.4 Hz, 1H), 4.66 (d, *J* = 5.5 Hz, 1H), 4.64 (d, *J* = 5.4 Hz, 1H), 3.79 (d, *J* = 7.7 Hz, 1H), 3.59 (td, *J* = 10.1, 5.3 Hz, 1H), 3.29 (dd, *J* = 9.7, 7.7 Hz, 1H), 2.47 – 2.32 (m, 2H), 2.27 (d, *J* = 5.9 Hz, 1H), 1.73 (ddd, *J* = 13.4, 10.5, 3.0 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 138.6, 138.0, 128.6, 128.6, 128.0, 128.0, 127.9, 127.9, 84.9 (broad), 81.6 (broad), 74.7, 72.3, 65.6, 33.0, 31.5, 30.9 ppm. IR: (neat) υ 3300, 2862, 1454, 1062 cm^{-1}. [α]D_{20} (c = 0.15, DCM): +15. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{20}H_{23}NO_3$ 326.17507, found 326.17493.

Compound 13



Starting from a mixture of azido-alcohols **11** and **12** (40 mg, 0.11 mmol), the reaction was carried out following the procedure described for **15** to afford the desired product as a colorless oil (20 mg, 55%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.56 – 7.02 (m, 10H), 4.77 – 4.71 (m, 3H), 4.57 (d, *J* = 11.5 Hz, 1H), 3.89 (t, *J* = 4.5 Hz, 1H), 3.70 (dd, *J* = 12.0, 6.3 Hz, 1H), 3.56 (dd, *J* = 6.4, 4.9 Hz, 1H), 2.54 (dd, *J* = 6.0, 4.4 Hz, 1H), 2.44 (t, *J* = 4.6 Hz, 1H), 2.18 (dt, *J* = 14.4, 4.8 Hz, 1H), 1.94 (ddd, *J* = 14.4, 6.5, 1.2 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 138.6, 138.5, 128.6, 128.5, 128.0, 127.9, 127.8, 80.5, 77.8, 73.5, 71.2, 68.8, 32.2, 29.5, 28.2 ppm. IR: (neat) υ 3300, 1111, 1070, 1057, 697 cm^{-1}. [α]D_{20} (c = 0.3, DCM): -45. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{20}H_{23}NO_3$ 326.17507, found 326.17494.

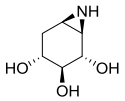
Compound 14



Ammonia (3 mL) was condensed at -60°C under inert atmosphere. Lithium (8 mg, 1.1 mmol) was added and the mixture was stirred until all lithium was completely dissolved (30 min). To the resulting dark-blue mixture, a solution of aziridine **13** (18 mg, 0.06 mmol) dissolved in dry THF (1 mL) was subsequently added. The reaction mixture was stirred for 1 h at -60 °C before being quenched with MilliQ-H_2O. The solution was allowed to come to RT and stirred in a warm water bath (40 °C) until all ammonia had evolved. After solvent evaporation under reduced pressure the crude was re-dissolved in 0.5N NH_4OH and purified over Amberlite CG-50 (NH_4^+). The product was eluted with 0.5N NH_4OH solution and obtained as a colorless oil (5.9 mg, 73%). ^{1}H NMR (400 MHz, D_2O): δ 3.83 (dd, J = 8.6, 3.7 Hz, 1H), 3.44 (m, 1H), 3.21 (dd, J = 10.2, 8.7 Hz, 1H), 2.49 (dd, J = 5.7, 3.8 Hz, 1H), 2.42 – 2.27 (m, 2H), 1.56 (m, 1H) ppm. ^{13}C NMR (101 MHz, D_2O): δ

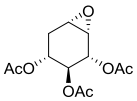
74.3, 71.9, 69.4, 35.4, 31.0, 27.8 ppm. IR: (neat) υ 3306, 1670, 1199, 1136 cm^{-1}. $[\alpha]D_{20}$ (c = 0.1, MeOH): -26. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_6H_{11}NO_3$ 146.08117, found 146.08132.

Compound 17



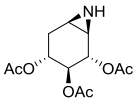
Starting from aziridine **16** (40 mg, 0.12 mmol), the reaction was carried out following the same procedure described for **13** to afford the product as a colorless oil (17 mg, 96%). ^{1}H NMR (400 MHz, D_2O): δ 3.64 (d, *J* = 8.3 Hz, 1H), 3.37 (td, *J* = 10.5, 5.4 Hz, 1H), 3.18 (dd, *J* = 8.6, 1.0 Hz, 1H), 2.46 (br s, 1H), 2.34 (dd, *J* = 13.8, 5.2 Hz, 1H), 2.20 (d, *J* = 6.0 Hz, 1H), 1.79 – 1.58 (m, 1H) ppm. ^{13}C NMR (101 MHz, D_2O): δ 77.5, 72.7, 66.2, 34.57, 31.3, 30.9 ppm. IR: (neat) υ 3365, 1653, 1463 cm^{-1}. $[\alpha]D_{20}$ (c = 0.3, MeOH): -24. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_6H_{11}NO_3$ 146.08117, found 146.08115.

Compound 19



Epoxide **4** (226 mg, 1.55 mmol) was dissolved in dry pyridine (5 mL) and cooled to 0 °C. Then, acetic anhydride (2.2 mL, 23.2 mmol) and a catalytic amount of DMAP were added and the mixture was stirred overnight at rt. The mixture was diluted with EtOAc (50 mL), quenched with sat. aq. $NaHCO_3$ (30 mL) and stirred vigorously for 3 h. The water layer was discarded, the organic phase was diluted with EtOAc (200 mL), washed with sat. aq. $NaHCO_3$, H_2O and brine, dried over $MgSO_4$, filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc 2:1) afforded the product as a yellow solid (371 mg, 88%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 5.34 – 5.22 (m, 2H), 4.93 – 4.80 (m, 1H), 3.44 (d, *J* = 3.9 Hz, 1H), 3.29 – 3.23 (dd, *J* = 4.9, 4.2 Hz, 1H), 2.60 (ddd, *J* = 15.3, 7.7, 5.0 Hz, 1H), 2.12 (s, 3H), 2.04 – 2.01 (m, 1H), 2.02 (s, 3H), 2.02 (s, 3H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ 170.6, 170.1, 169.8, 71.9, 70.2, 68.8, 54.3, 50.5, 29.0, 20.9, 20.9, 20.8 ppm. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{12}H_{16}O_7$ 295.0788, found 295.0795.

Compound 20



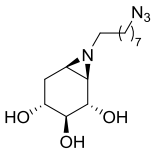
Epoxide **19** (204 mg, 0.75 mmol) was co-evaporated with toluene (3x) and dissolved in DMF under argon. Sodium azide (487 mg, 7.49 mmol) and triethylamine hydrochloride (113 mg, 0.82 mmol) were added and the mixture was stirred overnight at 80 °C. The mixture was diluted with 1N HCl (300 mL) and extracted with EtOAc (3 x 100 mL). The combined organic layers were washed with H_2O, brine, dried over $MgSO_4$, filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc 2:1) afforded the azido-alcohols as a mixture (169 mg, 72%) which was used directly in the next step without further characterization. The mixture of azido-alcohols (160 mg, 0.507 mmol) was co-evaporated with toluene, dissolved in dry THF, and polymer-bound triphenylphosphine (~3 mmol/g loading, 338 mg, 1.02 mmol) was added. After stirring for 24 h at 60 °C, the mixture was filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc 1:2) afforded the product as an oil (18 mg, 13%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 5.15 – 4.94 (m, 3H), 2.56 – 2.47 (m, 2H), 2.21 (d, *J* = 5.8 Hz, 1H), 2.09 (s, 3H), 2.02 (s, 3H), 1.99 (s, 3H), 1.83 (dddd, *J* = 10.5, 8.1, 6.1, 3.4 Hz, 1H) ppm. ^{13}C NMR

(101 MHz, $CDCl_3$): δ 170.4, 170.2, 169.9, 73.6, 73.5, 67.3, 32.4, 31.0, 29.6, 21.1, 21.0, 20.9 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{12}H_{17}NO_6$ 272.1129, found 272.1129.

1-azido-8-trifluoromethylsulfonyloctane

To dry DCM (5.8 mL) was added 8-azidooctan-1-ol (100 mg, 0.58 mmol) and pyridine (57 µL, 0.70 mmol) and the mixture was cooled to -20 °C. Triflic anhydride (118 µL, 0.70 mmol) was added and the mixture was stirred for 15 minutes. Then the mixture was diluted with DCM, and washed with cold water (3 x 10 mL). The organic layer was dried over $MgSO_4$, filtrated and concentrated at rt. The crude product was used directly for the alkylation of the aziridine. ^{1}H NMR (400 MHz, $CDCl_3$) δ 4.55 (t, *J* = 6.5 Hz, 2H), 3.27 (t, *J* = 6.9 Hz, 2H), 1.90 – 1.75 (m, 2H), 1.65 – 1.55 (m, 2H), 1.49 – 1.30 (m, 8H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 77.8, 51.6, 29.4, 29.0, 28.9, 28.9, 26.7, 25.2 ppm.

Compound 18

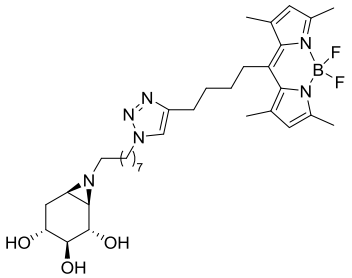


Aziridine **20** (18 mg, 66 µmol) was dissolved in dry DCM (700 µL) and cooled to 0 °C. Then, DIPEA (13 µL, 132 µmol) and 8-azidooctyl trifluoromethanesulfonate (0.5 M, 264 µL) were added and the mixture was stirred 24 h at rt. The reaction was quenched by adding sat. aq. $NaHCO_3$ and the mixture was extracted with EtOAc (3x). The combined organics were washed with brine, dried over $MgSO_4$, filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc 5:1) afforded the product as an oil (21 mg), which was taken up in methanol (500 µL) and subsequently NaOMe (0.1 M, 98 µL) was added. After stirring 16 h at rt, the mixture was quenched by adding triethylamine hydrochloride until the pH of the mixture was neutral. The mixture was concentrated, and flash purification by silica column chromatography (DCM/MeOH 93:7) afforded the product as an oil (14 mg, 73% over two steps). ^{1}H NMR (400 MHz, CD_3OD): δ 3.60 (d, *J* = 8.0 Hz, 1H), 3.39 (m, 1H), 3.28 (t, *J* = 6.8 Hz, 2H), 3.05 (dd, *J* = 9.8, 8.1 Hz, 1H), 2.32 (dd, *J* = 13.8, 5.4 Hz, 1H), 2.24 (m, 1H), 1.85 – 1.80 (m, 1H), 1.67 – 1.62 (m, 1H), 1.62 – 1.34 (m, 12H) ppm. ^{13}C NMR (101 MHz, CD_3OD): δ 79.1, 74.0, 68.1, 61.7, 52.4, 45.3, 41.1, 33.4, 30.6, 30.5, 30.2, 29.9, 28.3, 27.8 ppm. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{14}H_{26}N_4O_3$ 321.1897, found 321.1899.

General procedure for click reactions

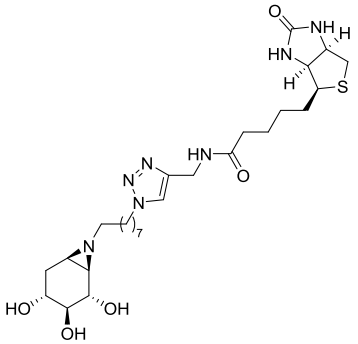
The azido compound (3-5 mg) was dissolved in degassed DMF (0.2 mL), then the alkyne-tag (1.1 eq), $CuSO_4$ (0.2 eq) and sodium ascorbate (0.4 eq) were added and the mixture was stirred for 16 h at rt. The reaction mixture was concentrated and purified by semi-preparative reversed phase HPLC (linear gradient. Solutions used: A: 50 mM NH_4HCO_3 in H_2O, B: acetonitrile).

Compound 21 (SYF170)



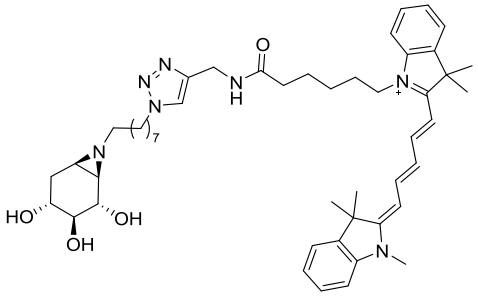
Following the general procedure starting from compound 18 (4.4 mg, 14.7 μmol), the product was obtained as an orange powder (5.4 mg, 58%). ^{1}H NMR (500 MHz, CD_3OD): δ 7.73 (s, 1H), 6.11 (s, 2H), 4.35 (t, *J* = 6.9 Hz, 2H), 3.60 (d, *J* = 8.0 Hz, 1H), 3.38 (m, 1H), 3.05 (dd, *J* = 9.8, 8.0 Hz, 1H), 3.03 – 2.98 (m, 2H), 2.78 (t, *J* = 7.2 Hz, 2H), 2.44 (s, 6H), 2.38 (s, 6H), 2.30 (dd, *J* = 13.1, 5.5 Hz, 1H), 2.20 (t, *J* = 7.3 Hz, 2H), 1.92 – 1.82 (m, 5H), 1.81 – 1.77 (m, 1H), 1.68 – 1.58 (m, 3H), 1.56 (d, *J* = 6.3 Hz, 1H), 1.54 – 1.48 (m, 2H), 1.35 – 1.20 (m, 7H) ppm. ^{13}C NMR (125 MHz, CD_3OD): δ 153.6, 147.2, 146.6, 140.8, 131.2, 122.0, 121.3, 77.8, 72.7, 66.8, 60.3, 49.9, 48.5, 44.0, 39.7, 32.1, 30.9, 29.9, 29.5, 29.1, 29.0, 28.5, 27.8, 26.8, 25.9, 24.6, 15.2, 13.1 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{33}H_{49}BF_2N_6O_3$ 627.4000, found 627.4029.

Compound 22 (SYF162)



Following the general procedure starting from compound **18** (4.3 mg, 14.5 μmol), the product was obtained as a white powder (6.9 mg, 82%). ^{1}H NMR (500 MHz, CD_3OD): δ 7.87 (s, 1H), 4.53 – 4.49 (m, 1H), 4.44 (s, 2H), 4.40 (m, 2H), 4.31 (dd, *J* = 7.9, 4.5 Hz, 1H), 3.61 (d, *J* = 8.0 Hz, 1H), 3.40 (td, *J* = 10.1, 5.5 Hz, 1H), 3.25 – 3.18 (m, 1H), 3.07 (dd, *J* = 9.9, 8.1 Hz, 1H), 2.95 (dd, *J* = 12.7, 5.0 Hz, 1H), 2.72 (d, *J* = 12.7 Hz, 1H), 2.33 (dd, *J* = 13.8, 5.5 Hz, 1H), 2.26 (m, 4H), 1.91 (m, 2H), 1.86 (dd, *J* = 9.6, 3.5 Hz, 1H), 1.79 – 1.51 (m, 7H), 1.48 – 1.27 (m, 11H) ppm. ^{13}C NMR (125 MHz, CD_3OD): δ 174.6, 164.8, 144.9, 122.8, 77.8, 72.7, 66.8, 62.0, 60.3, 60.3, 55.7, 50.0, 44.1, 39.8, 39.7, 35.2, 34.2, 32.1, 29.9, 29.1, 29.1, 28.6, 28.4, 28.1, 26.9, 26.0, 25.4 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{27}H_{45}N_7O_5S$ 580.3276, found 580.3279.

Compound 23 (SYF161)



Following the general procedure starting from compound 18 (4.6 mg, 15.3 μmol), the product was obtained as a blue powder (7.6 mg, 58%). ^{1}H NMR (500 MHz, CD_3OD): δ 8.28 (t, *J* = 13.1 Hz, 2H), 7.88 (s, 1H), 7.51 (d, *J* = 7.4 Hz, 2H), 7.45 – 7.39 (m, 2H), 7.33 – 7.30 (m, 2H), 7.30 – 7.25 (m, 2H), 6.65 (t, *J* = 12.4 Hz, 1H), 6.31 (d, *J* = 13.7 Hz, 2H), 4.42 (s, 2H), 4.40 – 4.36 (m, 2H), 4.13 – 4.09 (m, 2H), 3.65 (s, 3H), 3.60 (d, *J* = 8.0 Hz, 1H), 3.40 (m, 1H), 3.07 (dd, *J* = 9.8, 8.0 Hz, 1H), 2.32 (dd, *J* = 13.9, 5.4 Hz, 1H), 2.26 (m, 4H), 1.93 – 1.79 (m, 5H), 1.76 – 1.67 (m, 2H), 1.74 (s, 12H), 1.68 – 1.61 (m,

1H), 1.59 (d, *J* = 6.3 Hz, 1H), 1.50 (m, 3H), 1.43 – 1.23 (m, 9H) ppm. ^{13}C NMR (125 MHz, CD_3OD): δ 174.4, 174.0, 173.3, 154.2, 154.2, 142.9, 142.2, 141.3, 141.2, 128.4, 128.4, 125.4, 124.9, 124.9, 122.9, 122.1, 122.0, 110.7, 110.5, 103.1, 103.0, 77.8, 72.7, 66.8, 60.3, 50.0, 44.0, 43.4, 39.8, 35.2, 34.3, 32.1, 30.2, 30.0, 29.1, 29.1, 28.6, 26.9, 26.8, 26.6, 26.5, 26.0, 26.0, 25.1 ppm. HRMS (ESI) m/z: $[M]^+$ calc for $C_{49}H_{68}N_7O_4$ 819.5406, found 819.5336.

General procedure for SDS-PAGE experiments

Recombinant enzyme (rGBA: 1 pmol, PB90-1 β-xylosidase: 50 ng) or mouse liver lysate (Jackson's laboratories, C57Bl6/J, 40 µg total protein per sample) was diluted in 150 mM phosphate buffer with appropriate pH. A solution of the ABP with appropriate concentration was added and the mixture was incubated for 30 minutes at 37 °C. In the case of competition experiments, the enzyme solution was pre-incubated with appropriate inhibitor concentrations for 30 minutes at 37 °C. Then, laemmli (4X) was added and the mixture was heated at 100 °C for 5 minutes. The samples were loaded on 10% acrylamide SDS-PAGE gels and the gels were ran at a constant 90V. Wet slab gels were scanned on fluorescence using a Typhoon FLA9500 Imager (GE Healthcare) using λ_{EX} 635 nm; λ_{EM} > 665 nm. Images were acquired, processed and quantified with Image Quant (GE Healthcare).

Determination of IC_{50} values

The activity of recombinant GBA (Cerezyme, Genzyme) was measured at 37 °C with 4-methylumbelliferyl β-D-glucopyranoside as substrate as reported previously.[39] To determine the half-maximal inhibitory concentration (IC_{50}) value, the inhibitors were pre-incubated for 30 min with the enzyme before addition of the substrate mixture. The incubation mixture contained 3 mM β-4-MU-glucopyranoside, 0.2% (w/v) sodium taurocholate, 0.1% (v/v) Triton X-100 and 0.1% (w/v) BSA in 150 mM McIlvaine buffer, pH 5.2. After stopping the incubation with excess NaOH-glycine (pH 10.6), fluorescence of liberated 4-methylumbelliferone was measured with a fluorimeter LS 55 (Perkin Elmer) using λex 366 nm and λem 445 nm. The IC50 values were estimated by non-linear regression analysis of blank-corrected datapoints, using a one-phase exponential decay function (GraphPad Prism 5.0).

References

1 V. Gieselmann, *Biochim. Biophys. Acta*, 1995, **1270**, 103–136.
2 K. A. Stubbs, *Carbohydr. Res.*, 2014, **390**, 9–19.
3 S. Atsumi, H. Iinuma, C. Nosaka and K. Umezawa, *J. Antibiot.*, 1990, **43**, 49–53.
4 J. Marco-Contelles, *Eur. J. Org. Chem.*, 2001, 1607–1618.
5 B. P. Rempel and S. G. Withers, *Glycobiology*, 2008, **18**, 570–586.
6 H. Paulsen, I. Sangster and K. Heyns, *Chem. Ber.*, 1967, **100**, 802–815.
7 K. Tatsuta, *Pure Appl. Chem.*, 1996, **68**, 1341–1346.
8 G. Legler, *Hoppe. Seylers. Z. Physiol. Chem.*, 1966, **345**, 197–214.
9 G. Legler, *Hoppe. Seylers. Z. Physiol. Chem.*, 1968, **349**, 767–774.
10 G. Caron and S. G. Withers, *Biochem. Biophys. Res. Commun.*, 1989, **163**, 495–499.

11 B. T. Adams, S. Niccoli, M. A. Chowdhury, A. N. K. Esarik, S. J. Lees, B. P. Rempel and C. P. Phenix, *Chem. Commun.*, 2015, **51**, 11390–11393.

12 E. Borges de Melo, A. da Silveira Gomes and I. Carvalho, *Tetrahedron*, 2006, **62**, 10277–10302.

13 V. W. Tai, P. Fung, Y. Wong and T. K. M. Shing, *Tetrahedron: Asymm.*, 1994, **5**, 1353–1362.

14 S. Ogawa, S. Uetsuki, Y. Tezuka, T. Morikawa, A. Takahashi and K. Sato, *Bioorg. Med. Chem. Lett.*, 1999, **9**, 1493–1498.

15 P. Biely, *Trends Biotechnol.*, 1985, **3**, 286–290.

16 A. Sunna and G. Antranikian, *Crit. Rev. Biotechnol.*, 1997, **17**, 39–67.

17 D. B. Jordan and K. Wagschal, *Appl. Microbiol. Biotechnol.*, 2010, **86**, 1647–1658.

18 B. Bernet and A. Vasella, *Helv. Chim. Acta*, 1979, **62**, 1990–2016.

19 F. G. Hansen, E. Bundgaard and R. Madsen, *J. Org. Chem.*, 2005, **70**, 10139–10142.

20 C. S. Poulsen and R. Madsen, *J. Org. Chem.*, 2002, **67**, 4441–4449.

21 L. Hyldtoft, C. S. Poulsen and R. Madsen, *Chem. Commun.*, 1999, 2101–2102.

22 L. Hyldtoft and R. Madsen, *J. Am. Chem. Soc.*, 2000, **122**, 8444–8452.

23 L. Keinicke and R. Madsen, *Org. Biomol. Chem.*, 2005, **3**, 4124–4128.

24 J. C. Grim, K. C. A. Garber and L. L. Kiessling, *Org. Lett.*, 2011, **13**, 3790–3793.

25 G. Sabitha, K. Shankaraiah and J. S. Yadav, *Eur. J. Org. Chem.*, 2013, 4870–4878.

26 V. R. Doddi, A. Kumar and Y. D. Vankar, *Tetrahedron*, 2008, **64**, 9117–9122.

27 J. Uenishi and H. Ohmiya, *Tetrahedron*, 2003, **59**, 7011–7022.

28 P. K. Brown, H C; Jadhav, *J. Am. Chem. Soc.*, 1983, **105**, 2092–2093.

29 J. N. Heo, E. B. Holson and W. R. Roush, *Org. Lett.*, 2003, **5**, 1697–1700.

30 D. Yang, M.-K. Wong and Y.-C. Yip, *J. Org. Chem.*, 1996, **60**, 3887–3889.

31 J. L. Jat, M. P. Paudyal, H. Gao, Q.-L. Xu, M. Yousufuddin, D. Devarajan, D. H. Ess, L. Kurti and J. R. Falck, *Science*, 2014, **343**, 61–65.

32 P. Serrano, A. Llebaria and A. Delgado, *J. Org. Chem.*, 2005, **70**, 7829–7840.

33 P. Pöchlauer, E. P. Müller and P. Peringer, *Helv. Chim. Acta*, 1984, **67**, 1238–1247.

34 J. Jiang, Thesis: Activity-based protein profiling of glucosidases, fucosidases and glucoronidases; Leiden University, 2016.

35 K. Y. Li, J. Jiang, M. D. Witte, W. W. Kallemeijn, H. van den Elst, C. S. Wong, S. D. Chander, S. Hoogendoorn, T. J. M. Beenakker, J. D. C. Codée, J. M. F. G. Aerts, G. A. van der Marel and H. S. Overkleeft, *Eur. J. Org. Chem.*, 2014, **2014**, 6030–6043.

36 S. van Weely, M. Brandsma, A. Strijland, J. M. Tager and J. M. F. G. Aerts, *BBA - Mol. Basis Dis.*, 1993, **1181**, 55–62.

37 K. J. Chin, W. Liesack and P. H. Janssen, *Int. J. Syst. Evol. Microbiol.*, 2001, **51**, 1965–1968.

38 R. W. Murray and M. Singh, *Org. Synth.*, 1997, **74**, 91–97.

39 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.

Chapter 3

Inter-class β-glycosidase profiling by deoxygenated activity-based cyclophellitol probes

Parts of this chapter have been published:
S.P. Schröder *et al.,* Towards broad spectrum activity-based glycosidase probes: synthesis and evaluation of deoxygenated cyclophellitol aziridines
Chem. Commun., **2017**, 53, 12528-12531

3.1 Introduction

Glycosidases are hydrolytic enzymes that catalyze the hydrolysis of interglycosidic linkages in oligosaccharides and glycoconjugates.[1] Glycosidases are involved in a broad array of human pathologies and are therefore promising therapeutic targets.[2] Activity-based protein profiling (ABPP) provides an effective tool for the analysis of enzyme composition, expression and activity, and can assist in the discovery of enzyme inhibitors.[3] ABPP relies on the addition of a mechanism-based inhibitor to the sample of interest, which reacts with the catalytic machinery of the target enzyme in a covalent and irreversible manner. The inhibitor carries a reporter tag and is termed an activity-based probe (ABP). As the first example on the development of ABPP,

Cravatt and co-workers have described an irreversible fluorophosphonate based ABP **1** (Figure 1a), which displays high reactivity towards a wide range of serine hydrolases (SH).[4] ABP **1** was used to reveal active SH composition and expression level in various tissues, and as well in the development of highly specific and potent inhibitors against, for instance, monoacylglycerol lipase in a competitive ABPP setting.[5] Cyclophellitol (**2**, Figure 1b), first isolated from *Phellinus* sp. is a covalent irreversible retaining β-glucosidase inhibitor.[6] The configurational isomers, *galacto*-cyclophellitol (**2**)[7] and *manno*-cyclophellitol (**3**)[8] are inactivators for β-galactosidases and β-mannosidases, respectively, and their selectivity arises from their absolute configurations. Several cyclophellitol based ABPs for selective glycosidase profiling have been published.[9–11] For example, treatment of mouse tissue lysates with acyl ABP **5** (Figure 1c) resulted in tissue-specific fluorescent labelling of the three retaining β-glucosidases, GBA1, 2 and 3 and lactase-phlorizin hydrolase (LPH) without non-specific cross-labelling. Later studies revealed that alkyl ABP **6** also effectively labels these retaining β-glucosidases.[12] While the selectivity of these probes is an attractive value, it also impedes the study of multiple enzyme classes with a single probe, as for

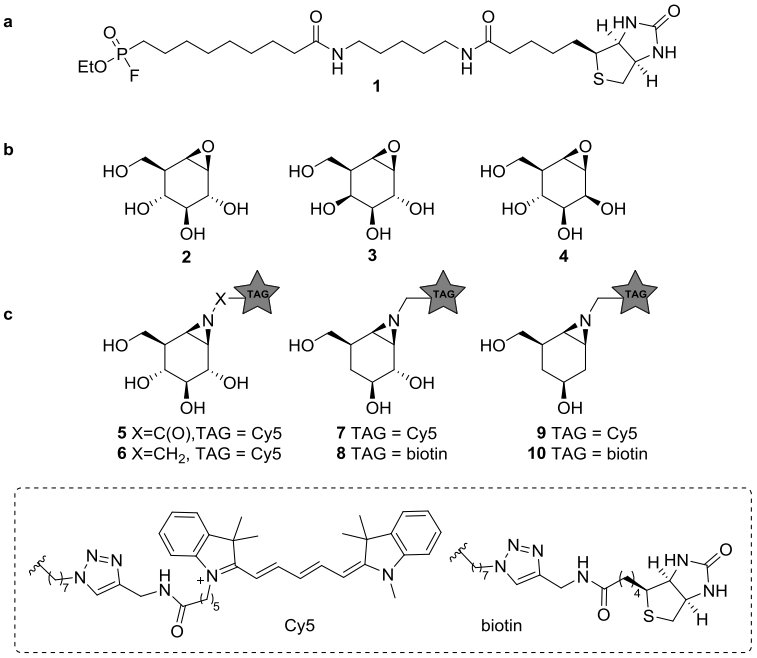


Figure 1 (a) Irreversible fluorophosphonate ABP **1** for serine hydrolases. (b) Structures of cyclophellitols **2**-**4**, selective covalent and irreversible retaining glycosidase inhibitors. (c) Selective β-glucosidase ABPs **5**-**6**, 4-deoxy ABPs **7**-**8** and 2,4-deoxy ABPs **9**-**10**.

example can be done by the broad-spectrum SH ABPs developed by Cravatt and co-workers.[4] Therefore, the development of broad spectrum glycosidase ABPs would be of interest. Since the substrate specificity of glucosidases, mannosidases and galactosidases is determined by the configuration at C2 and C4 of the substrate glycosides (axial or equatorial, see Figure 2), removal of the alcohol functionalities at these positions may abolish active site preference. In this Chapter, the synthesis of four deoxygenated ABPs **7**-**10** and their efficacy and specificity in activity-based glycosidase profiling is presented.

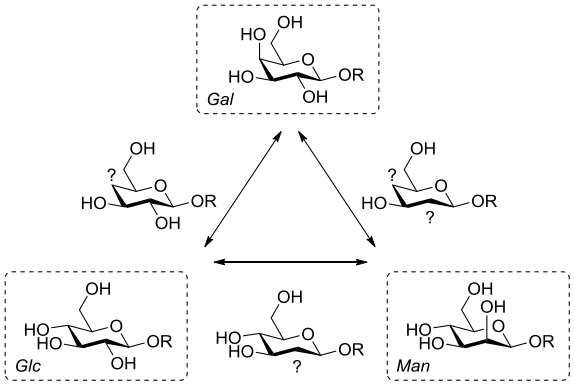
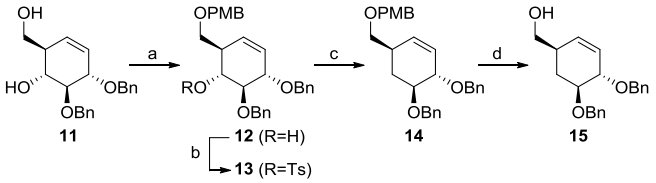


Figure 2 β-Galactosides (Gal), β-glucosides (Glc) and β-mannosides (Man) are distinguished by the configuration of the hydroxyl group(s) at C-2 and/or C-4.

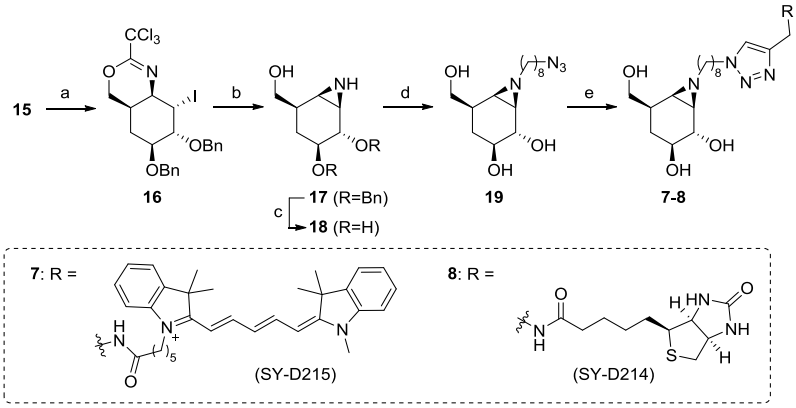
3.2 Results and Discussion

Synthesis of β-4-deoxy ABPs - The synthesis commenced with published diol **11**[13] (Scheme 1), which was selectively protected at the primary alcohol with 4-methoxybenzyl chloride using 2-aminoethyl diphenylborinate as catalyst,[14] resulting in **12**. The remaining secondary alcohol appeared unreactive towards standard tosylation conditions at room temperature; therefore tosylation was performed with an excess of tosyl chloride at elevated temperature overnight, affording **13** in high yield. The tosylate was reduced with an excess of lithium aluminum hydride in refluxing THF overnight, yielding 4-deoxy intermediate **14**. Deprotection of the PMB group at the primary alcohol could be performed by oxidative cleavage by DDQ, however yields varied greatly using this reagent. Alternatively, the primary alcohol could be deprotected effectively when acidic conditions were employed: treatment of **14** with catalytic HCl in HFIP[15] afforded allylic alcohol **15** in quantitative yield.



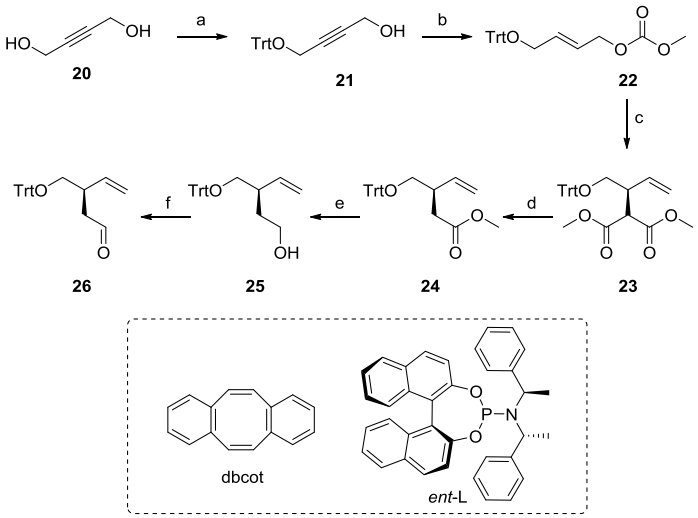
Scheme 1 Deoxygenation of cyclohexene **11**. Reagents and conditions: a) PMB-Cl, 2-aminoethyl diphenylborinate (10 mol%), K_2CO_3, KI, MeCN, 60 °C, 4h, 95%; b) TsCl, pyridine, 60 °C, 16h, 90%; c) $LiAlH_4$, THF, reflux, 16h, 86%; d) HCl (10 mol%), DCM/HFIP, 15 min, quant.

The aziridine warhead was obtained from cyclohexene **15** via a two-step procedure. First, the primary alcohol was reacted with trichloroacetonitrile. After work-up, the resulting crude trichloroacetimidate was reacted with *N*-iodosuccinimide (NIS), which resulted in consecutive cyclization into **16** (Scheme 2). After hydrolysis of the cyclic imidate by methanolic HCl, the resulting ammonium salt treated with basic Amberlite resin which resulted in intramolecular iodine displacement, affording aziridine **17**. The benzyl groups were subsequently removed by Birch debenzylation conditions to afford 4-deoxy-β-aziridine **18**. Prolonged heating of neat aziridine **18** at 40 °C led to decomposition, however alkylation with 1-azido-8-iodooctane at 100 °C in DMF successfully afforded **19** in moderate yield. Azide/alkyne click ligation with Cy5- or biotin-alkyne and subsequent HPLC purification led to the isolation of ABPs **7** and **8**.



Scheme 2 Synthesis of β-4-deoxy ABPs **7** and **8** from cycloxene **15**. a) 1. Cl_3CCN, DBU (10 mol%), DCM, rt, 16h; 2. NIS, $CHCl_3$, 0 °C to rt, 16h, 77%; b) HCl, MeOH/DCM, 24h, then Amberlite IRA-67, 16h, 80%; c) Li, NH_3, THF, -60 °C, 1h, quant.; d) 1-azido-8-iodooctane, K_2CO_3, DMF, 100 °C, 16h, 53%; e) Cy5-alkyne or biotin-alkyne, $CuSO_4$, sodium-ascorbate, DMF, H_2O, rt, 72h, yield **7**: 42%, yield **8**: 49%.

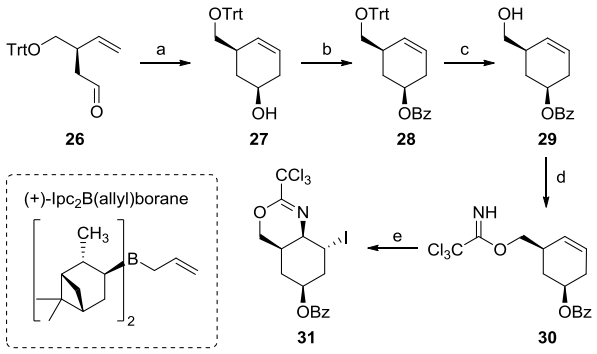
Synthesis of β-2,4-deoxy ABPs - The first steps in the synthesis route were adapted from a paper by Helmchen *et al*.[16] Commercially available diol **20** was tritylated to afford **21** (Scheme 3). The alkyne functionality was reduced by lithium aluminum hydride and the crude allylic alcohol was subsequently reacted with methyl chloroformate to form allylic methyl carbonate **22**. Next, the malonate moiety was installed at the branched allylic terminus via an iridium catalyzed decarboxylative allylic alkylation. The iridium catalyst $[Ir(dbcot)Cl]_2$ was prepared from commercially available $[Ir(cod)Cl]_2$ by ligand exchange with dibenzo[*a,e*]cyclooctene (dbcot). In turn, dbcot is prepared by radical dimerization of α,α'-dibromo-*o*-xylene, double benzylic bromination and subsequent elimination.[17] The phosphoramidite ligand is commercially available, and the enantiomeric outcome of the reaction is directed by the chirality of this ligand. Thus, by employing the iridium catalyzed allylic alkylation reaction in presence of the chiral ligand, *ent*-L, the malonate moiety was installed at the branched allylic terminus (branched:linear product 19:1) in high yield and enantioselectivity (97% *ee*). Malonate **23** is transformed into methyl ester **24** by



Scheme 3 Asymmetric installation of the malonate moiety to linear alkene **22** by iridium catalyzed allylic alkylation. Decarboxylation and subsequent reduction followed by oxidation yielded aldehyde **26**. Reagents and conditions: a) TrtCl, Et_3N, DMAP, DCM, rt, 18h, 65%; b) 1. $LiAlH_4$, THF, 0 °C to rt, 72h; 2. methyl chloroformate, pyridine, DCM, 0 °C, 30 min, 74%; c) $[Ir(dbcot)Cl]_2$ (2 mol%), *ent*-L (4 mol%), triazabicyclodecene (8 mol%), dimethylmalonate, THF, 50 °C, 5h, 92%, branched:linear 19:1, *ee* 97%; d) NaCl, BHT, H_2O, DMSO, 150 °C, 22h, 78%; e) $LiBH_4$, THF, rt, 72h, 98%; f) $PhI(OAc)_2$, TEMPO, DCM, rt, 5h, 87%.

Krapcho decarboxylation. Due to the high reaction temperatures employed, it was found critical to thoroughly degass the reaction mixture beforehand, as well as addition of a radical inhibitor to prevent side-product formation. Reduction of the ester yielded alcohol **25**, which was oxidized to aldehyde **26** using TEMPO/BAIB.

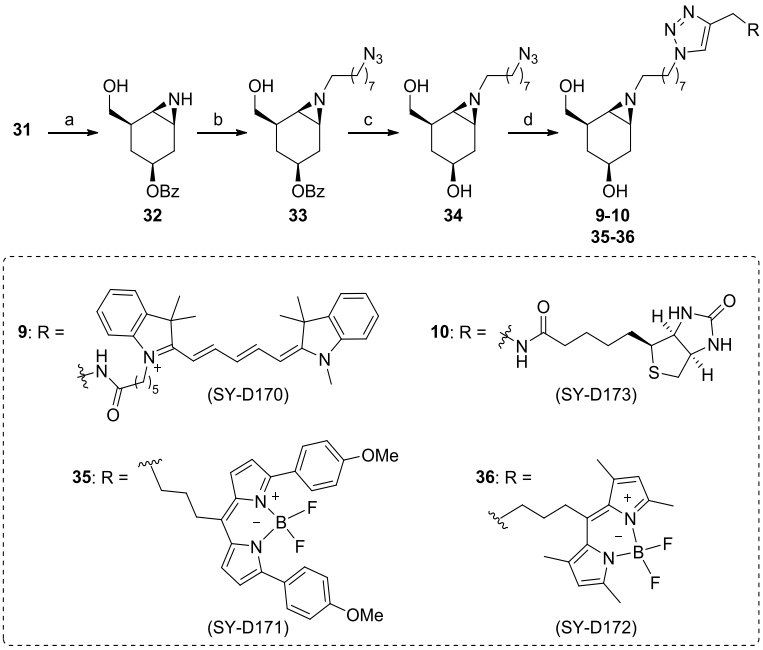
Aldehyde **26** was transformed into cyclic alkene **27** via a two-step procedure (Scheme 4). First, asymmetric allylation was performed using Brown's reagent,[18] followed by ring-closing metathesis by Grubb's II catalyst (*dr* 92:8). The secondary alcohol was benzoylated and **28** could be isolated as a single diastereoisomer. The trityl group was removed under acidic conditions, and the resulting compound **29** was reacted with trichloroacetonitrile to afford imidate **30**. Intramolecular cyclization was achieved by addition of NIS, resulting in cyclic imidate **31**.



Scheme 4 Asymmetric allylation of aldehyde **26** followed by reaction to the primary trichloroacetimidate and subsequent intramolecular cyclization affording cyclic imidate **31**. Reagents and conditions: a) 1. (+)-ipc$_2$B(allyl)borane, THF, -90 °C, 1h; 2. Grubb's II catalyst (5 mol%), DCM, reflux, 16h, 77%, *dr* 92:8; b) BzCl, Et$_3$N, DMAP, DCM, rt, 16h, 84%; c) CSA, MeOH, DCM, rt, 5h, 94%; d) Cl$_3$CCN, DBU, DCM, rt, 16h, 91%; e) NIS, CHCl$_3$, 0 °C to rt, 16h, 96%.

Next, cyclic imidate **31** was hydrolyzed under acidic conditions, and subsequent treatment of the resulting ammonium salt with excess base resulted in cyclization to give aziridine **32** (Scheme 5). This compound proved unstable at elevated (~40 °C) temperature, and concentration of the compound after work-up and column purification had to be performed on an ice-bath. The aziridine was alkylated with freshly prepared 8-azidooctyl trifluoromethanesulfonate, affording **33** in moderate yield. Final deprotection under Zémplen conditions resulted in isolation of **34**, which

was equipped with various reporter tags via azide/alkyne click ligation to give 2,4-deoxy ABPs **9**-**10** and **35**-**36**.



Scheme 5 Synthesis of β-2,4-deoxy ABPs **9**-**10** and **35**-**36** from cyclic imidate **31**. Reagents and conditions: a) HCl, MeOH, DCM, rt, 16h, then Amberlite IRA-67, 2h, 93%; b) 8-azidooctyl trifluoromethanesulfonate, DIPEA, THF, 0 °C, 2h, 53%; c) NaOMe, MeOH, rt, 24h, 82%; d) tag-alkyne, $CuSO_4$, sodium-ascorbate, DMF, H_2O, rt, 72h, yield **9**: 30%, yield **10**: 44%, yield **35**: 27%, yield **36**: 61%.

The labelling efficiency and specificity towards recombinant β-glucosidases, β-galactosidases and β-mannosidases was investigated for fluorescent ABPs **7** and **9**, and compared to that of β-glucosidase selective probe **6**. Enzyme incubation with a range of ABP concentrations (30 minutes, 37 °C, optimal enzyme pH) revealed that recombinant *Homo sapiens* GBA1 is highly efficiently labelled by ABP **6**, with protein bands visualized at probe concentrations as low as 10 nM (Figure 3a). Additionally, 4-deoxy ABP **7** also labelled GBA1 efficiently, although the potency was slightly reduced, likely due to lacking hydrogen-bonding interactions within the active site residue(s) as the result of the absence of OH-4. The labelling potency of 2,4-deoxy ABP **9** proved to be drastically reduced. All bands could be competed with selective β-glucosidase inactivator **2**. Also, denatured enzyme (1% SDS, 100 °C, 5 min) is not labelled by any of the probes indicating that all labelling is activity-based.

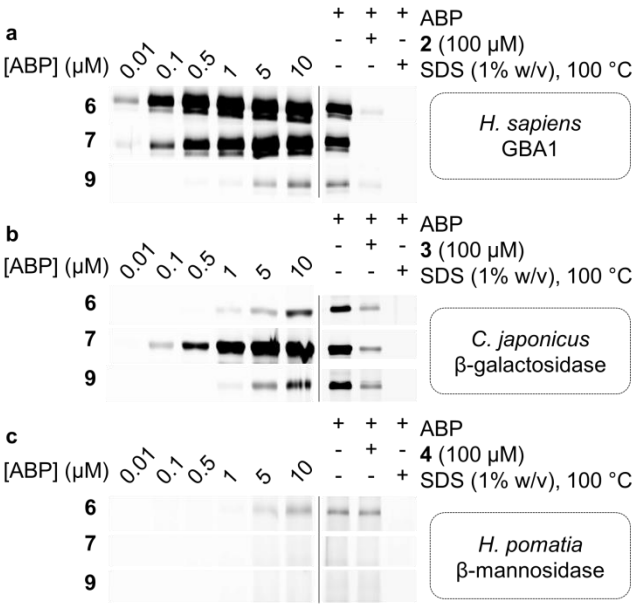


Figure 3 Fluorescent labelling of recombinant *H. sapiens* GBA1, *C. japonicus* β-galactosidase and *H. pomatia* β-mannosidase with ABPs **6**, **7** and **9** (left panel), and competition experiments with selective cyclophellitol inhibitors **2**, **3** or **4** (right panel). (a) GBA1 is potently labelled by **6** and **7**, but not **9**. All labelling was activity-based, as indicated by competition experiments with GBA1 selective inhibitor **2**. (b) β-galactosidase is only potently labelled by **7**, in an activity-based manner. (c) Only ABP **6** weakly labels β-mannosidase, but a competition assay with specific inhibitor **4** indicates that labelling is not activity-based.

As expected, binding kinetics studies revealed that **6** is a far more potent GBA1 inactivator (k_{inact}/K_I = 27.51 ± 0.85 $\mu M^{-1}min^{-1}$) than **7** (k_{inact}/K_I = 0.14 ± 0.08 $\mu M^{-1}min^{-1}$). Kinetic parameters for **9** could not be determined under the applied assay conditions reflecting its low potency (Figure 5). Labelling of *Cellvibrio japonicus* β-galactosidase from glycoside hydrolase family 35 (*Cj*GH35)[19] could be detected at micromolar concentrations of **6** and **9**. Interestingly, 4-deoxy ABP **7** was significantly more potent and labelled the enzyme at concentrations down to 100 nM, similar to GBA1 labelling; which is surprising given the extensive interactions of OH-4, from a *galacto*-cyclophellitol aziridine (TB562), with the enzyme (Figure 4, PDB: 5JAW). Indeed, kinetic parameters for **7** in *Cj*GH35 (k_{inact}/K_I = 0.16 ± 0.07 $\mu M^{-1}min^{-1}$) are comparable to that in GBA1 (kinetic parameters of **6** and **9** could not be determined under our assay conditions, due to low potency). Competition experiments indicated the labelling was activity-based, although full competition was not achieved which could indicate a fraction of non-specific binding. Lastly, it was found that none of the ABPs showed significant labelling of GH2 β-mannosidase from *Helix pomatia* (Figure 3c). Indeed, it is known that OH-2 is highly important for coordination to the glycosidase catalytic nucleophile and stabilizing the transition state (TS).[20] Moreover, probes **6**, **7** and **9** mimic preferentially adopt a $^4H_3$ configuration, which is characteristic for β-glucosidase and β-galactosidase transition states, whereas β-mannosidases often favour a $B_{2,5}$ TS.[21]

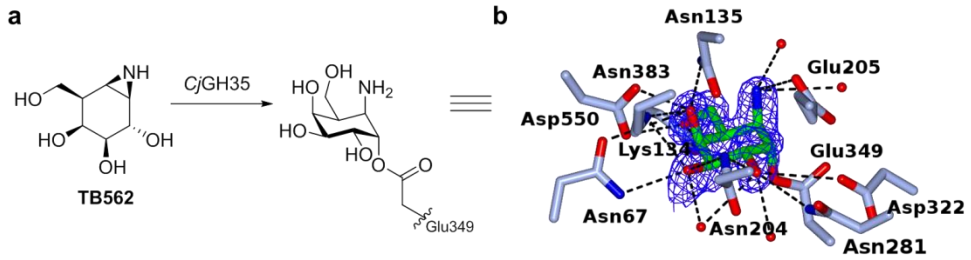


Figure 4 (a) Schematic depiction of g*alacto*-cyclophellitol aziridine **TB562** reacting with the active-site carboxylate of β-galactosidase *Cj*GH35. (b) Structure of **TB562** bound in the *Cj*GH35 active site, obtained by X-ray crystallography (PDB: 5JAW). The catalytic nucleophile (Glu349) is covalently bound to the anomeric centre of the inhibitor, while the catalytic acid/base (Glu205) interacts with the amine functionality. Of note are the multiple active site residues interacting with OH-4 of the substrate.

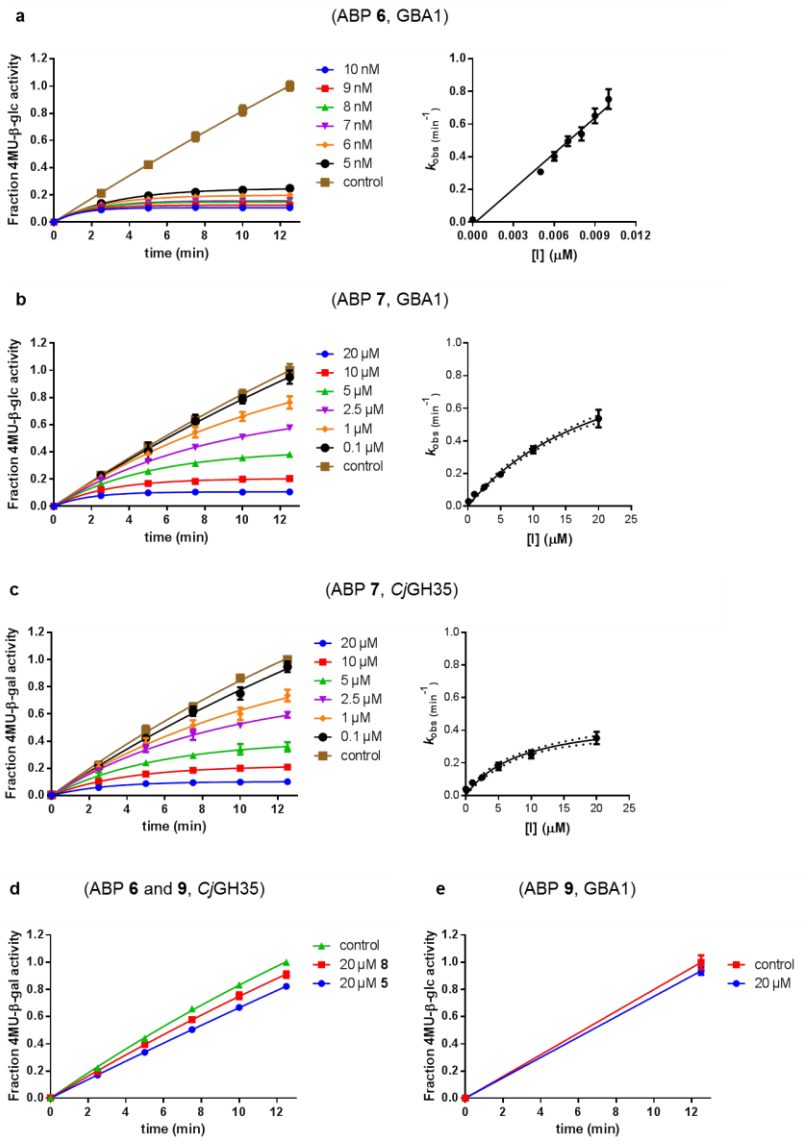


Figure 5 Kinetic plots of GBA1 and *Cj*GH35 inhibition by **6**, **7** and **9** with 4-methylumbelliferyl β-D-glucopyranoside and β-D-galactopyranoside substrates. (a-c) Left: plots of relative substrate activity versus time at a series of different concentrations inhibitor. Right: pseudo first-order rate constant plots. (d) ABPs **6** and **9** do not display significant inhibition of *Cj*GH35 at the highest inhibitor concentration in this assay. (e) ABP **9** does not display significant inhibition of GBA1 at the highest inhibitor concentration in this assay.

Next, glycosidase profiling by probes **7** and **9** in lysates from mouse (C57bl/6j, Jackson's laboratories) kidney and liver tissue, and human wild-type fibroblast cells (Lonza) was investigated, and compared to that of β-glucosidase selective probe **6**. The samples were incubated with the ABPs under identical conditions (1 μM ABP, pH 5.0, 37 °C, 30 min).

In mouse liver lysates, **6** labels a distinct band at ~65 kDa (Figure 6a), corresponding to the molecular weight of GBA1 (58 - 66 kDa[22]), and this band is fully abolished when the sample was pre-incubated with GBA1 inhibitor **2**.[23] Similarly, incubation with **7** resulted in a band at the same height. This band could only be partially competed with **2**, and pre-incubation by specific β-galactosidase inactivator **3** resulted in partial competition as well. Ultimately, full competition was achieved by pre-incubation with both cyclophellitols **2** and **3**. Thus, in contrast to labelling with **6**, incubation with **7** resulted in the concomitant labelling of GBA1 and a β-galactosidase, presumably acid β-galactosidase (GLB1, ~63 kDa).[24]

To identify the bands labelled on gel, the proteins were enriched by incubation of the lysate with biotin probe **8** and subsequent pull-down with streptavidin beads. After tryptic digestion, the peptide fragments were identified by LC-MS/MS and the abundance of the protein hits was quantified as previously described,[25] in unsupervised mode using the default settings of the PLGS (Waters) and IsoQuant software. Indeed, incubation with **8** resulted in enrichment of GBA1 and GLB1, corresponding to the mixed band at ~65 kDa on gel. In contrast, labelling with ABP **9** did not result in labelling of GBA1 or GLB1. Multiple bands were visible on gel, but no competition was observed except for one band at ~115 kDa of an unknown protein. It was therefore concluded that the other bands were a result of non-specific binding. Regardless, the lysate was incubated with **10** and the labelled proteins were identified by proteomic analysis. Whereas no enrichment of GBA1 could be detected, the probe did show increased levels of GLB1 labelling. Additionally, LC-MS/MS analysis detected enrichment of β-galactosylceramidase (GALC), an 80 kDa protein which is processed in the lysosomes into two catalytically active subunits of ~30 and ~50 kDa,[26] and is responsible for the hydrolysis of galactosylceramide. Interestingly, fluorescent labelling with **9** did not identify GALC on gel. However, it is known that expression levels of GALC in the liver are low.[27]

In mouse kidney lysates (Figure 6b), **6** only marginally labels GBA1. In contrast, bands at ~65 kDa and ~45 kDa were observed with probe **7** which could be fully competed out by pre-incubation with **3**, but not **2**. This indicated the labelling of mainly β-galactosidases in this sample, presumably GLB1 and GALC, which was subsequently confirmed by proteomics experiments as described above. Enrichment of GBA1 was

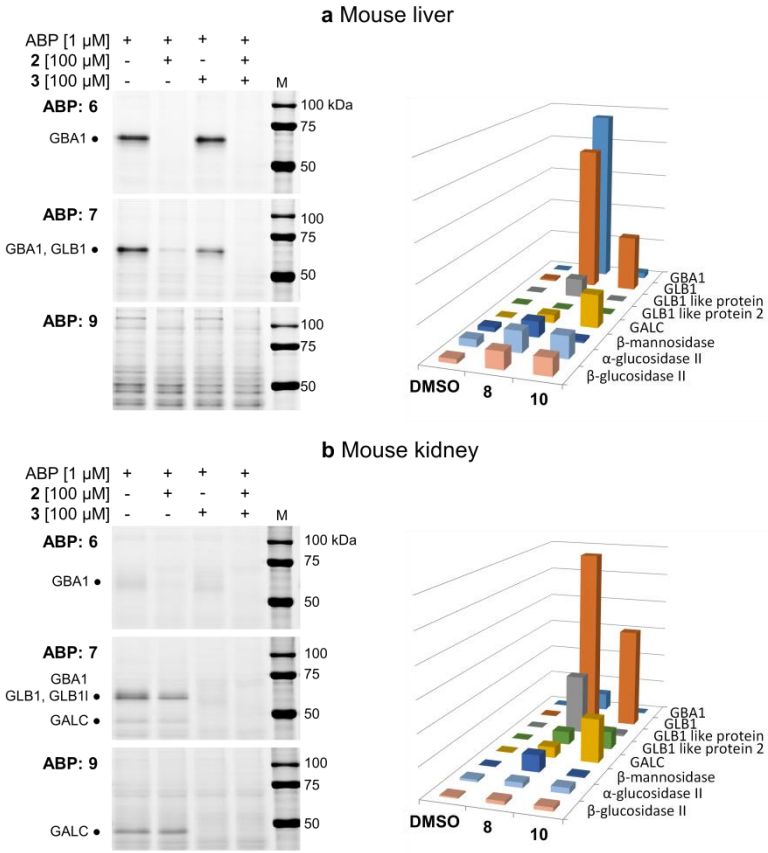


Figure 6 Fluorescent labelling of mouse liver (a) and kidney (b) lysates with probes **6**, **7** and **9**, complemented with competition experiments with β-glucosidase specific covalent inhibitor **2** and β-galactosidase specific **3**. The bar charts display the LC-MS/MS pull-down analyses with the biotinylated ABP analogues corresponding to the gels on the left, compared to the negative control (DMSO). (a) Probe **6** only labels GBA1, while **7** labels GBA1 and GLB1, and **9** only shows non-specific labelling of proteins in gel, although pull-down analysis shows enrichment of GLB1 and GALC with **10**. (b) Only low amounts of GBA1 could be detected by specific probe **6**. 4-Deoxy probe **7** labels GBA1, GLB1 (like) protein and GALC. ABP **9** shows specific fluorescent labelling of GALC on gel.

also detected but less pronounced, in line with the fluorescent assay. Labelling with probe **7** was also observed for GLB1-like protein 1 (GLB1l, ~68 kDa) and 2 (GLB1l2, processed mass unknown), which both possess the GLB1 active site motif.[28] While 2,4-deoxy ABP **9** did not significantly label GBA1 or GLB1, proteomics identified GLB1 and GLB1l2 by pull-down with **10**. Interestingly, a strong band at ~45 kDa was labelled, which could be competed with **3**. Indeed, proteomics revealed increased labelling of GALC by this probe and it was therefore concluded that **9** is a GALC specific ABP.

Lastly, the labelling of human fibroblast lysates was investigated (Figure 7c). A pattern of bands between 55-70 kDa is labelled by **6** which can subsequently be fully competed for with **2**, indicating the labelling of different isoforms[22] of GBA1 in this lysate. Similarly, **7** labels bands between 55-70 kDa, however with increased signal intensity. Approximately 50% competition was achieved by pre-incubation with either **2** or **3**, whereas pre-incubation with both inhibitors completely abrogated labelling. Indeed, pull-down analysis identified both GBA1 and GLB1 enrichment by biotinylated probe **8**. Finally, **9** did not show significant labelling of proteins in the sample, whereas GLB1 was enriched by biotinylated ABP **10** according to LC-MS/MS analysis.

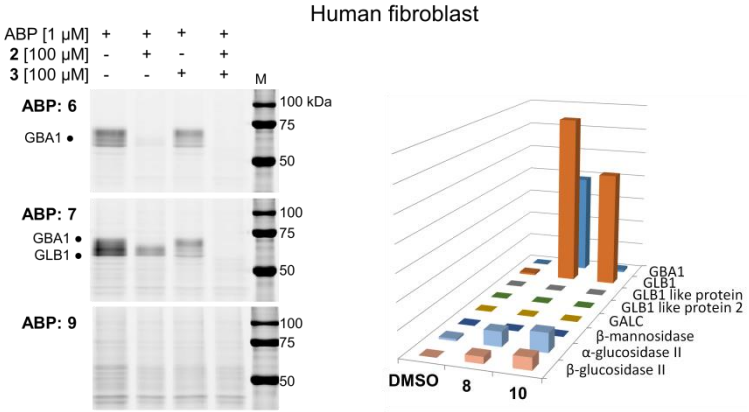


Figure 7 Fluorescent labelling of human fibroblast lysates with probes **6, 7** and **9**, complemented with competition experiments with β-glucosidase specific covalent inhibitor **2** and β-galactosidase specific **3**. The bar charts display the LC-MS/MS pull-down analyses with the biotinylated ABP analogues corresponding to the gels on the left, compared to the negative control (DMSO). ABP **6** labels GBA1 only, while **7** labels both GBA1 and GLB1. ABP **9** does not significantly label proteins, however pull-down analysis shows enrichment of GLB1.

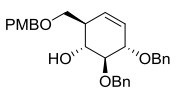
3.3 Conclusion

In summary, this Chapter describes the syntheses towards a set of deoxygenated probes, which were modelled on cyclophellitol aziridine **6**. None of the probes display activity towards β-mannosidases. Deoxygenation at C2 and C4 (ABPs **9**-**10**) enables labelling of purified β-glucosidases and β-galactosidases, albeit with low potency. Nevertheless, it was found that **9** displayed unexpected specificity towards β-galactosylceramidase in mouse kidney lysate where the enzyme is abundant.[27] Deoxygenation at C4 (ABPs **7**-**8**) affects the binding kinetics for GBA1, however *in vitro* labelling proceeded with high efficiency and moreover these probes now identify retaining β-galactosidases as well. Based on the results presented herein it is concluded that, indeed, probes designed for simultaneously profiling retaining glycosidases processing differently configured glycosides are within reach. Earlier, it was observed that probes designed for a specific class of retaining glycosidases may possess cross-reactivity towards other glycosidases,[10] however these results could not be predicted. Here, it is shown that by substitution of the hydroxyl group that distinguishes glucose from galactose for hydrogen on an otherwise unaltered cyclophellitol aziridine yields probes labelling both glucosidases and galactosidases. The strategy is however not general, since additionally stripping off OH-2 yields a largely inactive probe. Future research on differently configured, OH-deleted cyclophellitol aziridines is required to establish if and to what extent this strategy is viable to deliver a set of broad-spectrum glycosidase ABPs, for use in combination with the specific *exo*-glycosidase probes already in hand for dissecting large numbers of retaining glycosidases from various species.

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). *Cj*GH35 enzyme was kindly provided by Prof. dr. Gideon Davies, University of York, U.K.

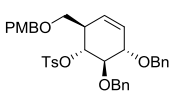
Compound 12



Diol **10** (1.36 g, 4.0 mmol) was dissolved in dry MeCN (20 mL), then PMBCl (0.81 mL, 6.0 mmol), KI (664 mg, 4.0 mmol), K_2CO_3 (608 mg, 4.4 mmol) and 2-aminoethyl diphenylborinate (90 mg, 0.4 mmol) were added and the mixture was stirred at 60 °C for 4 h. Then, the mixture was diluted with EtOAc (200 mL), washed with H_2O (2 x 100 mL) and brine, dried over $MgSO_4$, filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 4:1) gave the title compound as a colorless oil (1.75 g, 95%). ^1H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.26 (m, 10H), 7.24 (d, *J* = 8.6 Hz, 2H), 6.87 (d, *J* = 8.6 Hz, 2H), 5.73 (dt, *J* = 10.2, 2.3 Hz, 1H), 5.60 (d, *J* = 10.2 Hz, 1H), 4.99 (d, *J* = 11.3 Hz, 1H), 4.79 (d, *J* = 11.3 Hz, 1H), 4.67 (q, *J* = 11.5 Hz, 2H), 4.46 (s, 2H), 4.21 – 4.17 (m, 1H), 3.79 (s, 3H), 3.74 – 3.61 (m, 2H), 3.61 – 3.51 (m, 2H), 2.95 (brs, OH), 2.57 – 2.48 (m, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 159.3, 138.8, 138.4, 130.3,

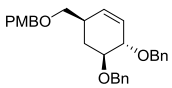
129.4, 128.6, 128.6, 128.3, 128.1, 128.0, 127.9, 127.8, 126.8, 113.9, 84.0, 80.3, 75.0, 73.1, 71.7, 71.4, 70.9, 55.4, 44.1 ppm. HRMS (ESI): *m/z* = $[M+Na]^+$ calc for $C_{29}H_{32}O_5$ 483.21420, found 483.21368.

Compound 13



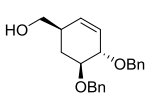
Cyclohexene **12** (2.54 g, 5.51 mmol) was co-evaporated with toluene and dissolved in pyridine (55 mL). Tosyl chloride (10.5 g, 55.1 mmol) was added in portions over 10 minutes, and then the mixture was stirred at 60 °C for 16 h. The mixture was cooled to 0 °C, quenched with H_2O (20 mL) and diluted with EtOAc (200 mL). The organic phase was washed with aq. 1M HCl (3 x 100 mL), H_2O (100 mL), sat. aq. $NaHCO_3$ (100 mL) and brine, dried over $MgSO_4$, filtered and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 7:1 → 5:1) gave the title compound as a colorless oil (3.07 g, 90%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.73 (d, *J* = 8.3 Hz, 2H), 7.31 – 7.18 (m, 10H), 7.16 (d, *J* = 5.6 Hz, 2H), 7.07 (d, *J* = 8.0 Hz, 2H), 6.90 – 6.84 (m, 2H), 5.70 (dt, *J* = 10.2, 2.4 Hz, 1H), 5.61 (dt, *J* = 10.2, 1.8 Hz, 1H), 5.00 – 4.92 (m, 1H), 4.65 (d, *J* = 11.4 Hz, 1H), 4.56 – 4.48 (m, 3H), 4.44 (d, *J* = 11.5 Hz, 1H), 4.32 (d, *J* = 11.5 Hz, 1H), 4.23 – 4.17 (m, 1H), 3.80 (s, 3H), 3.73 (dd, *J* = 10.1, 7.5 Hz, 1H), 3.52 (dd, *J* = 9.3, 3.4 Hz, 1H), 3.41 (dd, *J* = 9.3, 5.4 Hz, 1H), 2.71 – 2.63 (m, 1H), 2.29 (s, 3H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 159.2, 144.3, 138.5, 138.1, 134.7, 130.4, 129.5, 129.4, 128.5, 128.1, 128.1, 128.0, 127.8, 127.8, 127.5, 127.3, 127.0, 113.8, 81.4, 80.4, 80.3, 74.6, 72.9, 72.3, 68.8, 55.4, 43.6, 21.7 ppm. HRMS (ESI): *m/z* = $[M+Na]^+$ calc for $C_{36}H_{38}O_7S$ 637.22305, found 637.22318.

Compound 14



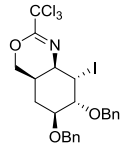
Tosylate **13** (3.07 g, 5.0 mmol) was co-evaporated with toluene (3x) and subsequently dissolved in dry THF (100 mL). The mixture was cooled to 0 °C and $LiAlH_4$ (2.4 M in THF, 15.6 mL, 37.4 mmol) was added and the mixture was refluxed overnight. After cooling to 0 °C, the reaction was quenched by slow addition of EtOAc (100 mL). An aqueous solution of Rochelle's salt (33 wt%, 100 mL) was added and the mixture was stirred vigorously for 1 h. Then, the organic phase was separated and the water phase was extracted with EtOAc (2x 100 mL). The combined organic layers were washed with brine, dried over $MgSO_4$, filtered and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 8:1) gave the title compound as a colorless oil (1.91 g, 86%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.42 – 7.18 (m, 12H), 6.88 (d, *J* = 7.9 Hz, 2H), 5.69 (s, 2H), 4.78 – 4.64 (m, 4H), 4.44 (s, 2H), 4.12 (d, *J* = 7.2 Hz, 1H), 3.80 (s, 3H), 3.69 (td, *J* = 9.1, 7.2, 3.2 Hz, 1H), 3.38 – 326 (m, 2H), 2.56 (s, 1H), 2.20 (d, *J* = 12.8 Hz, 1H), 1.38 (q, *J* = 11.8 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 159.3, 138.9, 130.6, 130.4, 129.4, 128.5, 128.5, 127.9, 127.8, 127.8, 127.6, 113.9, 79.7, 79.4, 73.9, 72.9, 72.2, 71.8, 55.4, 37.1, 30.9 ppm. HRMS (ESI): *m/z* = $[M+H]^+$ calc for $C_{29}H_{32}O_4$ 445.23734, found 445.23729.

Compound 15



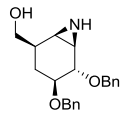
Compound **14** (1.91 g, 4.29 mmol) was dissolved in a mixture of DCM/HFIP (1:1 v/v, 43 mL), and aq. HCl (12 M, 36 µL, 0.43 mmol) was added. The mixture was stirred for 15 minutes and was subsequently quenched by addition of sat. aq. $NaHCO_3$ (10 mL). The mixture was diluted with DCM (200 mL) and washed with brine. The organic phase was dried over $MgSO_4$, filtered and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 4:1 → 2:1) gave the title compound as a colorless oil (1.48 g, quant.). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.24 (m, 10H), 5.81 (dt, *J* = 10.1, 2.5 Hz, 1H), 5.74 (d, *J* = 10.2 Hz, 1H), 4.76 – 4.63 (m, 4H), 4.06 (dq, *J* = 6.5, 2.5 Hz, 1H), 3.73 (ddd, *J* = 10.1, 6.4, 3.6 Hz, 1H), 3.62 – 3.52 (m, 2H), 2.51 – 2.42 (m, 1H), 2.20 (dt, *J* = 13.1, 4.6 Hz, 1H), 1.95 (brs, OH), 1.55 – 1.46 (m, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.8, 138.6, 130.7, 128.5, 128.5, 128.1, 127.9, 127.8, 127.8, 127.7, 78.1, 77.7, 72.0, 71.8, 66.6, 38.5, 29.9 ppm. HRMS (ESI): *m/z* = $[M+H]^+$ calc for $C_{21}H_{24}O_3$ 325.17982, found 325.17987.

Compound 16



Compound **15** (1.48 g, 4.56 mmol) was dissolved in DCM (23 mL). Trichloroacetonitrile (687 µL, 6.85 mmol) and DBU (68 µL, 0.46 mmol) were added and the mixture was stirred overnight at rt. The mixture was concentrated, and filtered over a small silica plug. Concentration of the eluate afforded an oil (1.9 g) which was directly taken up in $CHCl_3$ (40 mL) and cooled to 0 °C. Then *N*-iodosuccinimide (1.37 g, 6.08 mmol) was added and the mixture was stirred overnight while the ice-bath was allowed to slowly reach rt. The mixture was diluted with $CHCl_3$ (150 mL), washed with aq. 10% $Na_2S_2O_3$ (100 mL) and brine. Flash purification by silica column chromatography (pentane/EtOAc, 20:1) gave the title compound as a colorless oil (2.08 g, 77% over 2 steps). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.44 – 7.22 (m, 10H), 4.80 (s, 1H), 4.75 (dd, *J* = 11.5, 8.2 Hz, 2H), 4.64 (dd, *J* = 11.5, 4.0 Hz, 2H), 4.38 (dd, *J* = 11.1, 2.9 Hz, 1H), 4.30 (d, *J* = 11.2 Hz, 1H), 3.99 (t, *J* = 3.9 Hz, 1H), 3.82 (td, *J* = 10.5, 4.9 Hz, 1H), 2.91 – 2.63 (m, 2H), 1.97 (dt, *J* = 13.1, 4.4 Hz, 1H), 1.59 – 1.48 (m, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 152.8, 138.6, 137.9, 128.5, 128.2, 127.9, 127.8, 77.9 (broad, 2 x C), 73.4, 73.4, 72.2, 58.7, 35.7 (broad, assigned by HSQC), 29.6 (broad, assigned by HSQC), 27.0 (broad, assigned by HSQC) ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{23}H_{23}Cl_3INO_3$ 592.97882, found 593.9883.

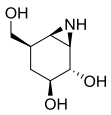
Compound 17



Compound **16** (35 mg, 59 µmol) was dissolved in MeOH/DCM (1:1, 1.2 mL) and HCl (1.25 M in MeOH, 443 µL, 0.55 mmol) was added. The mixture was stirred 24 h, and was subsequently neutralized by addition of Amberlite IRA-67. After stirring overnight, the reaction mixture was filtered and concentrated. Flash purification by silica column chromatography (DCM/MeOH, 49:1 → 19:1) gave the title compound as a colorless oil (16 mg, 80%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.43 – 7.21 (m, 10H), 4.77 (s, 2H), 4.65 (d, *J* = 4.0 Hz, 2H), 3.76 (dd, *J* = 10.4, 4.8 Hz, 1H), 3.69 (dd, *J* = 10.5, 6.1 Hz, 1H), 3.65 (d, *J* = 8.0 Hz, 1H), 3.43 (ddd, *J* = 11.9,

8.0, 3.7 Hz, 1H), 2.40 (dd, *J* = 5.7, 2.7 Hz, 1H), 2.28 (d, *J* = 6.0 Hz, 1H), 2.16 – 2.05 (m, 1H), 1.65 (dt, *J* = 12.9, 4.2 Hz, 1H), 1.27 (q, *J* = 12.6 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.9, 138.5, 128.5, 128.4, 127.9, 127.8, 127.7, 127.6, 80.5, 80.1, 72.8, 71.6, 66.0, 36.4, 32.8, 32.3, 24.0 ppm. TLC-MS (ESI) m/z: $[M+H]^+$ calc for $C_{21}H_{26}NO_3$ 340.18, found 340.3.

Compound 18

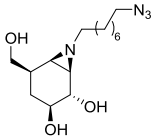


Ammonia (3 mL) was condensed in a flask at -60 °C, and lithium wire (7 mg, 0.94 mmol) was added. The resulting deep-blue solution was stirred for 30 minutes to dissolve all lithium. Aziridine **17** (16 mg, 47 µmol) was taken up in dry THF (1 mL) and added to the reaction mixture. After stirring for 1 h, the mixture was quenched with H_2O. The mixture was slowly warmed to rt and evaporated. The crude was dissolved in H_2O and eluted over a column packed with Amberlite CG-50 (NH_4^+) with 0.5M NH_4OH as eluent, affording the title compound as an oil (8 mg, quant.). ^{1}H NMR (400 MHz, D_2O) δ 3.42 (dd, *J* = 10.8, 7.9 Hz, 1H), 3.37 – 3.30 (m, 2H), 3.21 (ddd, *J* = 12.1, 8.4, 3.5 Hz, 1H), 2.24 (dd, *J* = 5.9, 2.9 Hz, 1H), 2.10 – 1.96 (m, 2H), 1.35 (dt, *J* = 12.7, 4.2 Hz, 1H), 0.68 (q, *J* = 12.4 Hz, 1H) ppm. ^{13}C NMR (101 MHz, D_2O) δ 73.1, 73.0, 64.2, 36.3, 35.0, 31.8, 26.8 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_7H_{14}NO_3$ 160.09682, found 160.09690.

8-azido-1-iodooctane

8-chloro-1-octanol (10.5 g, 64 mmol) was dissolved in DMSO (16 mL), NaN_3 (6.2 g, 96 mmol) was added and the mixture was stirred overnight at 80 °C. The mixture was diluted with EtOAc (100 mL) and washed with H_2O (10 x 60 mL). The organic phase was dried over $MgSO_4$, filtrated and concentrated. The crude intermediate was taken up in DCM (200 mL), Et_3N (14.2 mL, 102 mmol) was added and the mixture was cooled to 0 °C. Then, MsCl (7.4 mL, 96 mmol) was added dropwise and the mixture was allowed to reach rt. After 1 h at rt, the mixture was quenched with H_2O (50 mL) and diluted with DCM (100 mL). The organic phase was washed with 1N HCl (3 x 100 mL), sat. aq. $NaHCO_3$ (1 x 100 mL) and brine, dried over $MgSO_4$, filtrated and concentrated. The crude intermediate was taken up in DMF (640 mL), KI (15.9 g, 96 mmol) was added and the mixture was stirred overnight at 70 °C. The solvent was evaporated, then the crude was diluted with H_2O (1 L) and extracted with Et_2O (3 x 100 mL). The combined organic fractions were washed with brine, dried over $MgSO_4$, filtrated, concentrated and flash purification by silica column chromatography (Et_2O/pentane, 1:99) gave the title compound as a pale yellow oil (13.8 g, 77%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 3.26 (t, *J* = 6.9 Hz, 2H), 3.19 (t, *J* = 7.0 Hz, 2H), 1.88 – 1.73 (m, 2H), 1.66 – 1.50 (m, 2H), 1.46 – 1.26 (m, 8H).

Compound 19

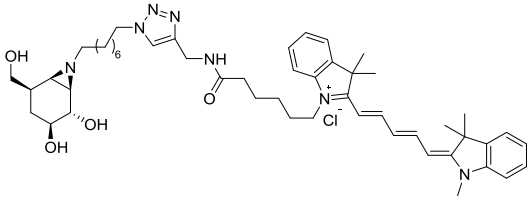


Starting from benzyl protected aziridine **17** (97 mg, 0.29 mmol), the compound was deprotected as described above and purified by elution over Amberlite CG-50 (NH_4^+). After evaporation of the solvent the intermediate was re-dissolved in DMF (1 mL), 1-azido-8-iodooctane (161 mg, 0.57 mmol) and K_2CO_3 (47 mg, 0.34 mmol) were added and the mixture was stirred overnight at 100 °C. The mixture was concentrated at 60 °C, and flash purification by silica column chromatography (DCM/MeOH, 19:1 → 9:1) gave the title compound as colorless oil (47 mg, 53% over two steps). ^{1}H NMR (400 MHz, MeOD) δ 3.60 (dd, *J* = 10.2, 7.9 Hz, 1H), 3.52 – 3.44 (m, 2H), 3.28 (t, *J* = 6.8 Hz, 2H), 3.26 – 3.22 (m, 1H), 2.42 (dt, *J* = 11.5, 7.6 Hz, 1H), 2.15 – 2.01 (m, 2H), 1.84 (dd, *J* = 6.1, 3.2 Hz, 1H), 1.63 (d, *J* = 6.2 Hz, 1H), 1.62 – 1.53 (m, 4H), 1.50 (dt, *J* = 12.7, 4.5 Hz, 1H), 1.36 (m, 8H), 0.97 (q, *J* = 12.1 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 74.5, 74.4, 65.8, 61.9, 52.4, 45.7, 42.5, 38.6, 30.5, 30.3, 30.2, 29.9, 29.4, 28.3, 27.8 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{15}H_{29}N_4O_3$ 313.22342, found 313.22365.

General procedure for click reactions

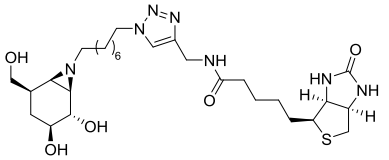
The azido compound (3-5 mg) was dissolved in DMF (0.5 mL), then the alkyne-tag (1.1 eq), $CuSO_4$ (0.2 eq) and sodium ascorbate (0.4 eq) were added and the mixture was stirred for 72 h at rt. The reaction mixture was concentrated and purified by semi-preparative reversed phase HPLC (linear gradient. Solutions used: A: 50 mM NH_4HCO_3 in H_2O, B: acetonitrile.

Compound 7 (SY-D215)



Following the general procedure, reaction of compound **18** (5.7 mg, 18.3 µmol) with Cy5-alkyne[29] afforded the title compound as a blue solid (6.6 mg, 42%). ^{1}H NMR (600 MHz, D_2O) δ 7.81 (s, 1H), 7.81 – 7.73 (m, 2H), 7.33 (dd, *J* = 12.8, 7.3 Hz, 2H), 7.26 (t, *J* = 7.5 Hz, 1H), 7.21 (t, *J* = 9.0 Hz, 2H), 7.14 – 7.07 (m, 2H), 7.05 (t, *J* = 7.2 Hz, 1H), 6.33 (t, *J* = 12.1 Hz, 1H), 6.03 (d, *J* = 13.6 Hz, 1H), 5.97 (d, *J* = 13.4 Hz, 1H), 4.38 (s, 2H), 4.23 (t, *J* = 6.6 Hz, 2H), 3.92 (s, 2H), 3.56 – 3.50 (m, 4H), 3.44 (dd, *J* = 10.5, 6.9 Hz, 1H), 3.34 (ddd, *J* = 12.2, 8.6, 3.5 Hz, 1H), 2.25 (q, *J* = 8.4, 6.5 Hz, 3H), 2.13 (dd, *J* = 6.6, 3.1 Hz, 1H), 1.92 (s, 3H), 1.85 (td, *J* = 11.0, 5.0 Hz, 1H), 1.72 (dd, *J* = 5.8, 3.0 Hz, 1H), 1.70 – 1.63 (m, 4H), 1.63 – 1.55 (m, 4H), 1.40 (d, *J* = 13.9 Hz, 12H), 1.32 (dd, *J* = 14.2, 7.5 Hz, 4H), 1.03 (s, 8H), 0.84 (q, *J* = 12.5 Hz, 1H) ppm. ^{13}C NMR (125 MHz, D_2O) δ 182.4, 176.7, 174.6, 173.5, 154.1, 153.7, 145.7, 143.5, 142.8, 142.0, 129.5, 126.2, 125.9, 125.3, 124.5, 123.3, 123.2, 111.7, 111.6, 104.2, 103.5, 73.7, 73.5, 65.2, 60.8, 51.2, 49.9, 49.8, 45.0, 44.5, 42.0, 37.3, 36.3, 35.2, 31.8, 30.4, 29.5, 29.4, 29.1, 28.9, 27.9, 27.8, 27.6, 27.5, 26.6, 26.4, 26.1 ppm. HRMS (ESI) m/z: $[M]^+$ calc for $C_{50}H_{70}N_7O_4^+$ 832.54838, found 832.54865.

Compound 8 (SY-D214)

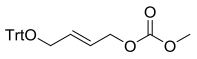


Following the general procedure, reaction of compound **18** (5.7 mg, 18.3 µmol) with biotin-alkyne[30] afforded the title compound as a white solid (5.3 mg, 49%). ^{1}H NMR (600 MHz, D_2O) δ 7.89 (s, 1H), 4.61 (dd, *J* = 7.9, 4.9 Hz, 1H), 4.45 (d, *J* = 6.8 Hz, 2H), 4.43 – 4.35 (m, 3H), 3.66 (dd, *J* = 10.6, 7.6 Hz, 1H), 3.56 (d, *J* = 8.6 Hz, 1H), 3.51 (dd, *J* = 10.6, 6.7 Hz, 1H), 3.39 (ddd, *J* = 12.3, 8.6, 3.7 Hz, 1H), 3.28 (dt, *J* = 10.0, 5.2 Hz, 1H), 2.99 (dd, *J* = 13.1, 5.0 Hz, 1H), 2.78 (d, *J* = 13.0 Hz, 1H), 2.44 (ddd, *J* = 11.4, 9.7, 6.5 Hz, 1H), 2.30 (t, *J* = 7.1 Hz, 2H), 2.25 – 2.17 (m, 1H), 2.13 (ddd, *J* = 11.7, 9.6, 5.3 Hz, 1H), 1.94 (dd, *J* = 6.2, 3.2 Hz, 1H), 1.92 (d, *J* = 1.1 Hz, 1H), 1.89 (p, *J* = 6.9 Hz, 2H), 1.76 (d, *J* = 6.3 Hz, 1H), 1.68 (dtd, *J* = 18.8, 9.8, 4.9 Hz, 2H), 1.64 – 1.58 (m, 2H), 1.58 – 1.43 (m, 3H), 1.36 – 1.16 (m, 10H), 0.88 (q, *J* = 12.5 Hz, 1H) ppm. ^{13}C NMR (125 MHz, D_2O) δ 177.6, 166.3, 145.7, 124.8, 73.7, 73.6, 65.3, 63.0, 61.2, 60.7, 56.4, 51.4, 45.0, 41.9, 40.7, 37.3, 36.2, 35.3, 30.2, 29.4, 29.3, 28.9, 28.8, 28.6, 28.6, 27.3, 26.4, 26.1 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{28}H_{48}N_7O_5S$ 594.34321, found 594.34302.

Compound 21

To a solution of 1,4-butynediol (100 g, 1.16 mol) in dry DCM (465 mL) was added triethylamine (18.0 mL, 128 mmol), trityl chloride (116 mmol, 32.4 g) and DMAP (700 mg, 5.8 mmol). The mixture was stirred for 18 h at rt and subsequently concentrated. The crude mixture was diluted with water (1 L) and extracted with EtOAc (3 x 150 mL). The organic layers were washed with water and brine, dried over $MgSO_4$, filtrated and concentrated. The product was purified by flash chromatography (pentane/EtOAc; 6:1 → 3:1) to afford the product as a yellow oil (24.9 g, 65%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.47 – 7.13 (m, 15 H), 4.23 (s, 2H), 3.82 (s, 2H), 1.78 (s, OH). ^{13}C NMR (101 MHz, $CDCl_3$) δ 143.5, 129.1, 128.7, 128.3, 128.1, 127.3, 87.6, 83.8, 82.6, 53.2, 51.3. IR (ATR, cm^{-1}) 3200, 1489, 1444, 1026. HRMS: $[M+H]^+$ calc for $C_{23}H_{20}O_2$ 329.15416 found 329.00516.

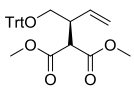
Compound 22



A solution of 2M $LiAlH_4$ in THF (151.6 mmol, 75.5 mL) was added to dry THF (375 mL) at 0°C. An anhydrous solution of compound **21** (24.9 g, 75.8 mmol) in THF (30 mL) was added dropwise and the mixture was stirred over weekend. The mixture was then cooled on ice and quenched with aq. 1N NaOH (75 mL). The resulting cake was filtered over celite and concentrated to 100 mL. The mixture was diluted with water (1 L), extracted with EtOAc (3 x 200 mL) and the combined organic fractions were washed with water and brine, dried over $MgSO_4$, filtrated and concentrated. The crude product was co-evaporated with toluene and dissolved in dry DCM (133 mL). Pyridine (8.1 mL, 100 mmol) was added and the mixture was cooled to 0 °C. Methyl chloroformate (7.7 mL, 100 mmol) was added dropwise and the mixture was stirred for 30 minutes. The reaction was quenched with sat. aq. $NaHCO_3$ (300 mL), extracted with DCM (200 mL) and the organic layer was washed with water, dried over $MgSO_4$, filtrated and concentrated. The product was purified by flash chromatography (pentane/EtOAc; 20:1) to afford the product as a white solid (21.8

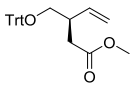
g, 74%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.52 – 7.36 (m, 6H), 7.35 – 7.12 (m, 9H), 5.99 (dtt, *J* = 15.1, 5.8, 1.4 Hz, 1H), 5.88 (dt, *J* = 15.6, 4.5 Hz, 1H), 4.65 (dd, *J* = 6.0, 1.1 Hz, 2H), 3.77 (s, 3H), 3.68 – 3.60 (m, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 155.7, 144.1, 132.4, 128.6, 127.9, 127.1, 124.0, 87.0, 68.0, 63.7, 54.9. IR (ATR, cm^{-1}) 1745, 1447, 1256, 940. HRMS: $[M+Na]^+$ calc for $C_{25}H_{24}O_4$ 411.15723 found 411.15655.

Compound 23



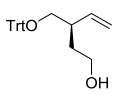
$[Ir(dbcot)Cl]_2$ (86 mg, 0.1 mmol)[17] and *ent*-L (120 mg, 0.2 mmol) were added to a flame-dried flask under argon, then dry THF (5 mL) was added. To this mixture was added triazabicyclodecene (56 mg, 0.4 mmol) and the mixture was stirred for 15 minutes at rt. A dry solution of compound **22** (1.94 g, 5 mmol) in THF (10 mL) was added dropwise. Next, dimethyl malonate (0.74 mL, 6.5 mmol) was added and the mixture was heated to 50 °C. After 5 h, the reaction mixture was concentrated and the product was purified by flash chromatography (pentane/EtOAc; 25:1) to afford the product as a colourless oil (2.04 g, 92%, branched:linear; 95:5). The *ee* (97%) was determined with a ChiralCel OD (Hex/isopropanol; 98:2), 1 mL/min, UV254 nm. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.46 – 7.34 (m, 6H), 7.33 – 7.17 (m, 9H), 5.91 (ddd, *J* = 17.2, 10.2, 9.1 Hz, 1H), 5.24 – 4.99 (m, 2H), 3.81 (d, *J* = 8.5 Hz, 1H), 3.63 (s, 3H), 3.58 (s, 3H), 3.23 – 3.13 (m, 2H), 3.09 (tt, *J* = 9.0, 5.4 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 168.7, 168.7, 143.9, 128.8, 127.9, 127.1, 118.2, 86.8, 64.4, 53.1, 52.5, 52.4, 44.6. $[\alpha]_D^{20}$ -22.3 (c 0.6, DCM). IR (ATR, cm^{-1}) 1734, 1437, 1151. HRMS: $[M+Na]^+$ calc for $C_{28}H_{28}O_5$ 467.18345 found 467.18221.

Compound 24



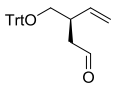
A solution of compound **23** (4.4 g, 9.9 mmol), sodium chloride (867 mg, 14.9 mmol), water (1.1 mL, 61.4 mmol) and BHT (110 mg, 0.5 mmol) in DMSO (50 mL) was put under vacuum and flushed with argon. The mixture was sonicated for 30 minutes under an argon flow and then heated at 150 °C for 22 h. The mixture was cooled to rt, diluted with water (375 mL) and brine (75 mL), and extracted with EtOAc (3 x 200 mL). The combined organic layers were washed with brine, dried over $MgSO_4$, filtrated and concentrated. The product was purified by flash chromatography (pentane/Et_2O; 50:1 → 30:1) to afford the product as a colourless oil (2.99 g, 78%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.48 – 7.36 (m, 6H), 7.34 – 7.17 (m, 9H), 5.75 (ddd, *J* = 17.6, 10.4, 7.7 Hz, 1H), 5.14 – 4.98 (m, 2H), 3.61 (s, 3H), 3.13 (dd, *J* = 8.9, 5.3 Hz, 1H), 3.00 (dd, *J* = 8.8, 6.9 Hz, 1H), 2.87 (m, 1H), 2.63 (dd, *J* = 15.3, 5.9 Hz, 1H), 2.35 (dd, *J* = 15.3, 8.4 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 173.1, 144.2, 138.4, 128.8, 127.9, 127.1, 116.2, 86.6, 66.1, 51.6, 40.8, 36.7. $[\alpha]_D^{20}$ -9.0 (c 0.4, DCM). IR (ATR, cm^{-1}) 1737, 1448, 1170, 1074. HRMS: $[M+Na]^+$ calc for $C_{26}H_{26}O_3$ 409.17797 found 409.17748.

Compound 25



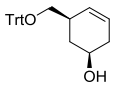
To a solution of compound **24** (18.5 g, 47.8 mmol) in dry THF (240 mL) was added $LiBH_4$ (5.2 g, 239 mmol) in portions at 0 °C and the mixture was stirred for 3 days at rt. The mixture was cooled on ice, quenched with sat. aq. NH_4Cl, diluted with water (500 mL) and brine (100 mL), and extracted with EtOAc (3x 150 mL). The combined organic layers were washed with brine, dried over $MgSO_4$, filtrated and concentrated. The product was purified by flash chromatography (pentane/EtOAc; 6:1 → 5:1) to afford the product as a colourless oil (16.7 g, 98%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.48 – 7.40 (m, 6H), 7.33 – 7.19 (m, 9H), 5.73 (ddd, *J* = 17.2, 10.3, 8.6 Hz, 1H), 5.13 – 5.01 (m, 2H), 3.61 (m, 2H), 3.10 (dd, *J* = 9.0, 5.7 Hz, 1H), 3.03 (dd, *J* = 9.0, 6.4 Hz, 1H), 2.44 (m, 1H), 1.78 (m, 1H), 1.65 – 1.51 (m, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 144.3, 140.1, 128.8, 127.9, 127.1, 116.1, 86.7, 67.1, 61.2, 34.8. $[\alpha]_D^{20}$ -15.0 (c 0.4, DCM). IR (ATR, cm^{-1}) 3350, 1489, 1448, 1064. HRMS: $[M+H]^+$ calc for $C_{25}H_{26}O_2$ 359.20110 found 359.24025.

Compound 26



Compound **25** (2.65 g, 7.4 mmol) was co-evaporated three times with toluene and dissolved in dry DCM (15 mL) under argon. To the mixture was added TEMPO (231 mg, 1.48 mmol) and $PhI(OAc)_2$ (2.4 g, 7.4 mmol), and the mixture was stirred for 5 h. The reaction was quenched with aq. $Na_2S_2O_3$ (10 mL), diluted with water (300 mL), and extracted with DCM (3 x 150 mL). The combined organic layers were washed with sat. aq. $NaHCO_3$ and water, dried over $MgSO_4$, filtrated and concentrated. The product was purified by flash chromatography (pentane/EtOAc; 50:1 → 20:1) to afford an orange oil (2.3 g, 87%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.70 (s, 1H), 7.41 (m, 6H), 7.36 – 7.14 (m, 9H), 5.83 – 5.64 (m, 1H), 5.13 – 4.97 (m, 2H), 3.19 (dd, *J* = 9.0, 5.3 Hz, 1H), 3.03 (dd, *J* = 9.0, 7.2 Hz, 1H), 2.96 – 2.84 (m, 1H), 2.62 (ddd, *J* = 16.5, 5.8, 2.0 Hz, 1H), 2.43 (ddd, *J* = 16.5, 7.9, 2.3 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 202.1, 144.1, 138.0, 128.8, 128.0, 128.0, 127.9, 127.4, 127.2, 116.6, 86.9, 66.2, 45.6, 38.9. $[\alpha]_D^{20}$ -14.3 (c 0.8, DCM). IR (ATR, cm^{-1}) 1724, 1490, 1448, 1068. HRMS: $[M+H]^+$ calc for $C_{25}H_{24}O_2$ 357.18546 found 357.15177.

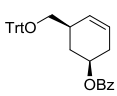
Compound 27



Aldehyde **26** (600 mg, 1.68 mmol) was dissolved in dry THF (20 mL) and cooled to -90 °C. A solution of 1M (+)-ipc$_2$B(allyl)borane solution in pentane was added to dry THF (10 mL), cooled to -90°C, and added dropwise via cannula. The mixture was stirred for 1 h, quenched with MeOH (6 mL) and diluted with sodium-phosphate buffer (pH 7, 0.5M, 12 mL). Hydrogen peroxide (13.2 mL, 30%) was added dropwise and the mixture was warmed to rt and stirred for 1 h. Next, the mixture was diluted with sat. aq. $NaHCO_3$ (100 mL) and extracted with EtOAc (3 x 75 mL). The combined organic layers were washed with water and brine, dried over $MgSO_4$, filtrated and concentrated. The crude product was dissolved in DCM (15 mL) and Grubb's II catalyst (71 mg, 5 mol%) was added. The mixture was refluxed overnight under argon. The product was purified by flash chromatography (pentane/EtOAc; 8:1 → 5:1) to afford the product as colourless oil (518 mg, 77%, dr 92:8). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.49 – 7.40 (m, 6H), 7.33 – 7.18 (m, 9H), 5.88 –

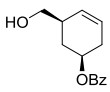
5.52 (m, 2H), 3.94 (dddd, *J* = 11.1, 9.1, 5.5, 3.5 Hz, 1H), 3.12 – 2.94 (m, 2H), 2.67 – 2.52 (m, 1H), 2.37 (d, *J* = 16.8 Hz, 1H), 2.21 – 2.10 (m, 1H), 2.04 – 1.87 (m, 1H), 1.80 (m, OH), 1.45 – 1.12 (m, 1H). ^{13}C NMR (MHz,) δ 144.3, 128.8, 128.6, 127.9, 127.0, 125.3, 86.5, 67.6, 67.5, 37.6, 36.4, 35.0. $[\alpha]_D^{20}$ +33.0 (c 0.2, DCM). IR (ATR, cm^{-1}) 3350, 1489, 1448, 1066. HRMS: [M+H]$^+$ calc for $C_{26}H_{26}O_2$ 371.20110 found 371.31553.

Compound 28



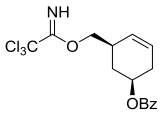
Compound **27** (675 mg, 1.7 mmol) was co-evaporated with toluene, dissolved in dry DCM (17 mL) and cooled to 0°C. Triethyl amine (1.2 mL, 8.5 mmol), benzoyl chloride (0.39 mL, 3.39 mmol) and DMAP (10 mg, 0.09 mmol) were added and the mixture was stirred overnight at rt. The reaction was quenched with sat. aq. $NaHCO_3$ (20 mL), stirred for 30 minutes and then extracted with DCM (3 x 20 mL). The combined organic layers were dried over $MgSO_4$, filtrated and concentrated. The product was purified by flash chromatography (pentane/EtOAc; 50:1) to afford the product as a single diastereoisomer as a colourless oil (672 mg, 84%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.12 – 7.93 (m, 2H), 7.60 – 7.50 (m, 1H), 7.49 – 7.37 (m, 7H), 7.33 – 7.17 (m, 10H), 5.68 (s, 2H), 5.33 – 5.20 (m, 1H), 3.10 (dd, *J* = 8.6, 6.5 Hz, 1H), 3.00 (dd, *J* = 8.5, 6.8 Hz, 1H), 2.74 (s, 1H), 2.62 – 2.48 (m, 1H), 2.32 – 2.23 (m, 1H), 2.24 – 2.14 (m, 1H), 1.55 (q, *J* = 11.6 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.2, 144.3, 132.9, 130.8, 129.7, 128.9, 128.8, 128.4, 127.9, 127.0, 124.7, 86.4, 70.9, 67.4, 37.2, 32.4, 31.5. $[\alpha]_D^{20}$ +36.0 (c 0.2, DCM). IR (ATR, cm^{-1}) 1714, 1448, 1273, 1070. HRMS: [M+Na]$^+$ calc for $C_{33}H_{30}O_3$ 497.20926 found 497.20825.

Compound 29



Compound **28** (660 mg, 1.39 mmol) was dissolved in a mixture of DCM (7 mL) and MeOH (7 mL). Then, CSA (16 mg, 0.07 mmol) was added and the mixture was stirred for 5 h at rt. The reaction was quenched with Et_3N (0.15 mL) and concentrated. The product was purified by flash chromatography (pentane/EtOAc; 4:1) to afford the product as a colourless oil (304 mg, 94%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.04 (d, *J* = 7.4 Hz, 2H), 7.55 (t, *J* = 7.4 Hz, 1H), 7.43 (t, *J* = 7.7 Hz, 2H), 5.76 (m, 1H), 5.66 (d, *J* = 10.2 Hz, 1H), 5.35 – 5.20 (m, 1H), 3.71 – 3.55 (m, 2H), 2.69 – 2.50 (m, 2H), 2.23 (m, 2H), 2.14 – 1.93 (m, 1H), 1.64 (q, *J* = 11.4 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.2, 133.0, 130.6, 129.6, 128.4, 128.0, 125.9, 70.8, 66.6, 39.0, 31.4, 31.3. $[\alpha]_D^{20}$ +17.3 (c 0.3, DCM). IR (ATR, cm^{-1}) 3400, 1712, 1273, 1114. HRMS: [M+Na]$^+$ calc for $C_{14}H_{16}O_3$ 255.0992 found 255.1001.

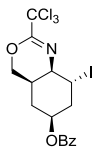
Compound 30



Compound **29** (290 mg, 1.25 mmol) was dissolved in dry DCM (12 mL). Then trichloroacetonitrile (0.25 mL, 2.5 mmol) and DBU (9.3 µL, 0.06 mmol) were added and the mixture was stirred overnight. The mixture was concentrated and the product was purified by flash chromatography (pentane/EtOAc; 20:1) to afford the product as a colourless oil (429 mg, 91%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.30 (s, 1H), 8.05 (d,

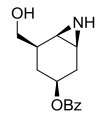
J = 7.5 Hz, 2H), 7.56 (t, *J* = 7.4 Hz, 1H), 7.45 (d, *J* = 7.7 Hz, 2H), 5.78 (ddt, *J* = 10.0, 4.9, 2.5 Hz, 1H), 5.69 (d, *J* = 10.2 Hz, 1H), 5.30 (dddd, *J* = 12.0, 9.2, 5.7, 3.5 Hz, 1H), 4.26 (d, *J* = 6.8 Hz, 2H), 2.95 (m, 1H), 2.67 – 2.50 (m, 1H), 2.35 – 2.16 (m, 2H), 1.67 (q, *J* = 11.3 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.1, 162.9, 133.0, 130.5, 129.7, 128.4, 126.9, 126.0, 72.5, 70.2, 35.6, 31.5, 31.2. $[\alpha]_D^{20}$ +27.3 (c 0.3, DCM). IR (ATR, cm^{-1}) 1714, 1664, 1271, 1070, 794, 709. HRMS: $[M+H]^+$ calc for $C_{16}H_{16}Cl_3NO_3$ 376.02740 found 376.02706.

Compound 31



Compound **30** (561 mg, 1.49 mmol) was co-evaporated with toluene, dissolved in dry $CHCl_3$ (15 mL) and cooled to 0°C. Then, NIS (570 mg, 2.23 mmol) was added and the mixture was stirred overnight at rt. The reaction was quenched with aq. $Na_2S_2O_3$ (10 mL) and stirred for 10 minutes. The mixture was diluted with sat. aq. $NaHCO_3$ (40 mL) and extracted with $CHCl_3$ (3 x 30 mL). The combined organic layers were dried over $MgSO_4$, filtrated and concentrated. The product was purified by flash chromatography (pentane/Et_2O; 10:1) to afford the product as a white foam (714 mg, 96%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.07 – 7.95 (m, 2H), 7.63 – 7.53 (m, 1H), 7.44 (dd, *J* = 8.4, 7.1 Hz, 2H), 5.49 (tt, *J* = 11.0, 4.2 Hz, 1H), 4.85 (q, *J* = 3.3 Hz, 1H), 4.47 (dd, *J* = 11.1, 3.1 Hz, 1H), 4.34 (dd, *J* = 11.2, 1.8 Hz, 1H), 3.92 (t, *J* = 3.6 Hz, 1H), 2.92 – 2.76 (m, 1H), 2.30 (dd, *J* = 14.1, 3.0 Hz, 1H), 2.24 – 2.11 (m, 1H), 1.92 (ddd, *J* = 14.4, 11.0, 3.7 Hz, 1H), 1.55 (q, *J* = 12.8 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 165.9, 152.9, 133.2, 130.1, 129.7, 128.5, 71.8, 70.4, 57.4, 34.8, 29.6, 28.5, 27.0. $[\alpha]_D^{20}$ +6.0 (c 0.3, DCM). IR (ATR, cm^{-1}) 1714, 1670, 1273, 1111, 821, 711. HRMS: $[M+H]^+$ calc for $C_{16}H_{15}Cl_3INO_3$ 501.92406 found 501.92348.

Compound 32

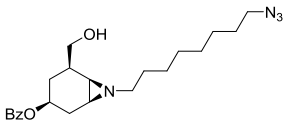


Compound **31** (51 mg, 0.1 mmol) was dissolved in a mixture of DCM (0.5 mL) and MeOH (0.5 mL) and cooled to 0°C. A solution of 1.25M HCl in MeOH (176 µL, 0.22 mmol) was added, and the mixture was stirred 16 h at rt. Then, Amberlite IRA-67 was added until neutral pH and the mixture was stirred for 2 h. The suspension was filtrated and the resin was washed with MeOH (3x). The organic phase was evaporated on an ice-bath. The product was purified by flash chromatography (DCM/MeOH; 16:1) using neutralized SiO_2, and the fractions were evaporated on an ice bath to afford the product as a colourless oil (23 mg, 93%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.01 (m, 2H), 7.55 (m, 1H), 7.48 – 7.38 (m, 3H), 4.97 (dddd, *J* = 12.1, 10.5, 6.7, 3.6 Hz, 1H), 3.85 (dd, *J* = 10.5, 4.4 Hz, 1H), 3.75 (dd, *J* = 10.5, 5.9 Hz, 1H), 2.48 (dtd, *J* = 13.9, 7.0, 1.9 Hz, 1H), 2.40 (dd, *J* = 6.1, 3.1 Hz, 1H), 2.32 (t, *J* = 6.5 Hz, 1H), 2.31 – 2.23 (m, 1H), 1.80 – 1.69 (m, 2H), 1.55 (q, *J* = 12.3 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.1, 133.0, 129.7, 128.5, 128.4, 70.7, 66.2, 36.4, 31.5, 30.1, 27.0, 27.0. $[\alpha]_D^{25}$ +21.6 (c 0.5, DCM). IR (ATR, cm^{-1}) 3300, 1708, 1273, 1112. HRMS:$[M+H]^+$ calc for $C_{14}H_{17}NO_3$ 248.12867 found 248.12802. *NOTE: This compound is unstable at elevated temperatures, and must be concentrated on an ice-bath to minimize decomposition.*

1-azido-8-trifluoromethylsulfonyloctane

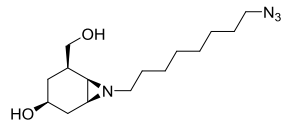
To dry DCM (5.8 mL) was added 8-azidooctan-1-ol (100 mg, 0.58 mmol) and pyridine (57 µL, 0.70 mmol) and the mixture was cooled to -20 °C. Triflic anhydride (118 µL, 0.70 mmol) was added and the mixture was stirred for 15 minutes. Then the mixture was diluted with DCM, and washed with cold water (3 x 10 mL). The organic layer was dried over $MgSO_4$, filtrated and concentrated at rt. The crude product was used directly for the alkylation of the aziridine. ^{1}H NMR (400 MHz, $CDCl_3$) δ 4.55 (t, *J* = 6.5 Hz, 2H), 3.27 (t, *J* = 6.9 Hz, 2H), 1.90 – 1.75 (m, 2H), 1.65 – 1.55 (m, 2H), 1.49 – 1.30 (m, 8H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 77.8, 51.6, 29.4, 29.0, 28.9, 28.9, 26.7, 25.2 ppm.

Compound 33



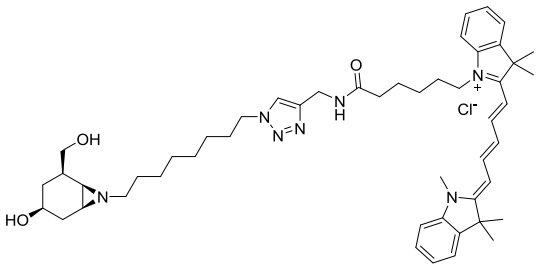
Compound **32** (54 mg, 0.22 mmol) was dissolved in dry THF (2 mL) and cooled to 0°C. DIPEA (42 µL, 0.24 mmol) was added and then 8-azidooctyl trifluoromethanesulfonate (73 mg, 0.24 mmol) in THF (0.5 mL). The mixture was stirred at 0 °C for 2 h and then diluted with water (20 mL) and brine (5 mL). The mixture was extracted with EtOAc (3 x 15 mL) and the combined organic layers were washed with brine, dried over $MgSO_4$, filtrated and concentrated. The product was purified by flash chromatography (pentane/EtOAc; 3:1) to afford the product as a colourless oil (46 mg, 53%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.00 (dd, *J* = 8.3, 1.3 Hz, 2H), 7.62 – 7.48 (m, 1H), 7.42 (t, *J* = 7.7 Hz, 1H), 4.91 (dddd, *J* = 11.9, 10.5, 6.5, 3.5 Hz, 1H), 3.87 (dd, *J* = 10.3, 4.2 Hz, 1H), 3.79 – 3.68 (m, 1H), 3.26 (t, *J* = 6.9 Hz, 2H), 2.47 (m, 2H), 2.39 – 2.30 (m, 1H), 2.23 – 2.10 (m, 2H), 1.78 – 1.65 (m, 3H), 1.64 – 1.51 (m, 8H), 1.34 (m, 7H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.0, 132.9, 130.5, 129.5, 128.3, 70.7, 66.2, 61.0, 51.4, 40.0, 36.1, 35.9, 29.8, 29.5, 29.4, 29.1, 28.8, 27.7, 27.3, 26.6. $[\alpha]_D^{25}$ +18.0 (c 0.4, DCM). IR (ATR, cm^{-1}) 3400, 2927, 2094, 1741, 1274, 1112. HRMS: $[M+H]^+$ calc for $C_{22}H_{32}N_4O_3$ 401.25527 found 401.25370.

Compound 34



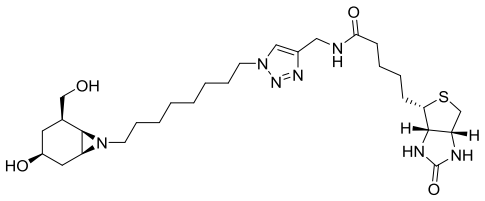
Compound **33** (46 mg, 0.11 mmol) was dissolved in MeOH (1.1 mL), then 0.1M NaOMe in MeOH (690 µL, 0.07 mmol) was added and the mixture was stirred for 24 h at rt. Then the mixture was diluted with sat. aq. $NaHCO_3$ (10 mL), and extracted with EtOAc (5 x 10 mL). The combined organic layers were dried over $MgSO_4$, filtrated and concentrated. The product was purified by flash chromatography using neutralized SiO_2 (DCM/MeOH; 32:1 → 24:1) to afford the product as a colourless oil (28 mg, 82%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.84 (dd, *J* = 10.4, 7.3 Hz, 1H), 3.79 (m, 1H), 3.74 (dd, *J* = 10.3, 5.8 Hz, 1H), 3.26 (t, *J* = 6.9 Hz, 2H), 2.24 (m, 3H), 1.98 (dd, *J* = 14.2, 5.1 Hz, 1H), 1.85 (dt, *J* = 14.1, 3.7 Hz, 1H), 1.75 (m, 2H), 1.66 – 1.42 (m, 6H), 1.44 – 1.20 (m, 8H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 66.6, 66.6, 60.6, 51.4, 39.7, 37.9, 33.7, 30.9, 30.2, 29.4, 29.3, 29.0, 28.8, 27.2, 26.6. $[\alpha]_D^{25}$ +32.7 (c 0.3, MeOH). IR (ATR, cm^{-1}) 3350, 2927, 2091, 1245. HRMS: $[M+H_3O]^+$ calc for $C_{15}H_{28}N_4O_2$ 315.2391 found 315.2399.

Compound 9 (SY-D170)



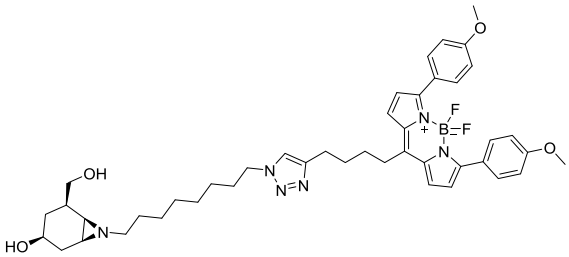
Following the general procedure starting from compound **34** (3.8 mg, 13 μmol), the product was obtained as a blue powder (3.2 mg, 30%). ^{1}H NMR (600 MHz, CD_3OD) δ 8.24 (td, *J* = 13.1, 2.9 Hz, 2H), 7.83 (s, 1H), 7.49 (d, *J* = 7.4 Hz, 2H), 7.41 (q, *J* = 7.3 Hz, 2H), 7.28 (m, 4H), 6.62 (t, *J* = 12.4 Hz, 1H), 6.28 (dd, *J* = 14.0, 2.1 Hz, 2H), 4.41 (s, 2H), 4.36 (t, *J* = 7.1 Hz, 2H), 4.09 (t, *J* = 7.5 Hz, 2H), 3.66 – 3.59 (m, 1H), 3.63 (s, 3H), 3.49 (m, 2H), 2.39 – 2.31 (m, 1H), 2.25 (t, *J* = 7.3 Hz, 2H), 2.22 – 2.17 (m, 1H), 2.07 – 2.00 (m, 2H), 1.87 (m, 2H), 1.83 (m, 2H), 1.73 (s, 12H), 1.71 – 1.65 (m, 4H), 1.55 – 1.45 (m, 6H), 1.33 – 1.27 (m, 8H), 0.97 (q, *J* = 11.7 Hz, 1H). ^{13}C NMR (151 MHz, CD_3OD) δ 175.7, 175.4, 174.6, 155.5, 155.5, 146.1, 144.2, 143.5, 142.6, 142.5, 136.0, 131.2, 129.8, 129.7, 128.8, 126.6, 126.3, 126.2, 124.1, 123.4, 123.3, 112.0, 111.8, 104.4, 104.2, 68.3, 66.3, 61.9, 51.3, 50.5, 44.8, 41.5, 38.8, 38.6, 36.5, 35.6, 33.8, 32.4, 31.5, 31.3, 31.1, 30.4, 30.2, 30.0, 28.3, 28.1, 27.9, 27.8, 27.4, 27.3, 26.4. HRMS: [M$^+$] calc for $C_{50}H_{70}N_7O_3$ 816.55401 found 816.55291.

Compound 10 (SY-D173)



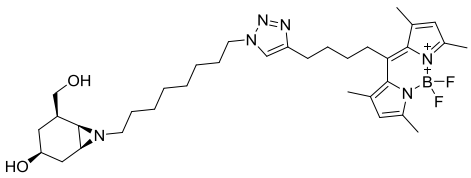
Following the general procedure starting from compound **34** (4.1 mg, 14 μmol), the product was obtained as a white powder (3.5 mg, 44%). ^{1}H NMR (600 MHz, CD_3OD) δ 7.84 (s, 1H), 4.49 (dd, *J* = 7.9, 4.9 Hz, 1H), 4.42 (s, 2H), 4.37 (t, *J* = 7.1 Hz, 2H), 4.29 (dd, *J* = 7.9, 4.4 Hz, 1H), 3.63 (dd, *J* = 10.2, 8.0 Hz, 1H), 3.50 (m, 2H), 3.19 (dt, *J* = 9.6, 5.2 Hz, 1H), 2.93 (dd, *J* = 12.7, 5.0 Hz, 1H), 2.71 (d, *J* = 12.7 Hz, 1H), 2.39 (dt, *J* = 11.7, 7.8 Hz, 1H), 2.23 (m, 3H), 2.04 (m, 2H), 1.89 (m, 3H), 1.76 – 1.63 (m, 5H), 1.63 – 1.51 (m, 4H), 1.49 (dd, *J* = 14.1, 10.3 Hz, 1H), 1.42 (m, 2H), 1.39 – 1.25 (m, 8H), 1.18 (t, *J* = 7.0 Hz, 1H), 0.97 (q, *J* = 11.8 Hz, 1H). ^{13}C NMR (151 MHz, CD_3OD) δ 176.0, 166.1, 146.2, 124.1, 68.3, 66.2, 63.3, 61.9, 61.6, 57.0, 51.4, 41.5, 41.1, 38.8, 38.7, 36.5, 35.6, 33.8, 32.4, 31.3, 30.4, 30.2, 30.0, 29.7, 29.5, 28.3, 27.4, 26.7. HRMS: [M+H]$^+$ calc for $C_{28}H_{47}N_7O_4S$ 578.3483 found 578.3498.

Compound 35 (SY-D171)



Following the general procedure starting from compound **34** (4.3 mg, 14 μmol), the product was obtained as a purple powder (3.0 mg, 27%). ^{1}H NMR (600 MHz, CD_3OD) δ 7.88 – 7.80 (d, *J* = 8.9 Hz, 4H), 7.68 (s, 1H), 7.42 (d, *J* = 4.3 Hz, 2H), 7.00 – 6.93 (d, *J* = 8.9 Hz, 4H), 6.69 (d, *J* = 4.3 Hz, 2H), 4.32 (t, *J* = 7.0 Hz, 2H), 3.84 (s, 6H), 3.60 (dd, *J* = 10.2, 7.8 Hz, 1H), 3.47 (m, 2H), 3.05 (t, *J* = 7.4 Hz, 2H), 2.78 (t, *J* = 6.9 Hz, 2H), 2.30 (dt, *J* = 11.6, 7.7 Hz, 1H), 2.19 (m, 1H), 2.02 (m, 1H), 1.95 (dt, *J* = 11.7, 7.3 Hz, 1H), 1.85 (m, 6H), 1.66 – 1.60 (m, 2H), 1.56 – 1.42 (m, 4H), 1.33 – 1.11 (m, 8H), 0.95 (q, *J* = 11.8 Hz, 1H). ^{13}C NMR (151 MHz, CD_3OD) δ 162.2, 158.8, 146.7, 137.5, 132.2, 128.4, 126.5, 123.3, 121.0, 114.6, 68.3, 66.2, 61.84, 55.8, 51.2, 41. 5, 38.8, 38.6, 34.1, 33.8, 32.4, 31.2, 31.0, 30.4, 30.3, 30.2, 29.9, 28.2, 27.3, 25.7. HRMS: $[M+H]^+$ calc for $C_{44}H_{55}BF_2N_6O_4$ 781.4419 found 781.4428.

Compound 36 (SY-D172)



Following the general procedure starting from compound **34** (4.4 mg, 15 μmol), the product was obtained as a yellow powder (5.6 mg, 61%). ^{1}H NMR (600 MHz, CD_3OD) δ 7.73 (s, 1H), 6.11 (s, 2H), 4.35 (t, *J* = 6.9 Hz, 2H), 3.62 (dd, *J* = 10.2, 7.9 Hz, 1H), 3.49 (m, 2H), 3.07 – 2.95 (m, 2H), 2.78 (t, *J* = 7.2 Hz, 2H), 2.44 (s, 6H), 2.37 (s, 6H), 2.36 – 2.31 (m, 1H), 2.22 (dt, *J* = 13.0, 6.1 Hz, 1H), 2.05 (m, 1H), 2.00 (m, 1H), 1.88 (m, 4H), 1.71 – 1.60 (m, 4H), 1.53 (m, 3H), 1.48 (dd, *J* = 14.0, 10.2 Hz, 1H), 1.36 – 1.18 (m, 8H), 0.97 (q, *J* = 11.7 Hz, 1H). ^{13}C NMR (151 MHz, CD_3OD) δ 154.9, 148.5, 147.9, 142.2, 132.6, 123.4, 122.6, 68.3, 66.2, 61.9, 51.2, 41.5, 38.8, 38.7, 33.8, 32.4, 32.2, 31.2, 30.8, 30.4, 30.2, 29.9, 29.1, 28.2, 27.3, 25.9, 16.5, 14.5. HRMS: $[M+H]^+$ calc for $C_{34}H_{51}BF_2N_6O_2$ 625.4207 found 625.4226.

Labelling and SDS-PAGE of recombinant enzymes

Recombinant glycosidases were labeled with ABPs **6**, **7** and **9** at 37 ℃ for 30 minutes with the optimized conditions for each individual enzyme; 2.5 pmol GBA1 (Cerezyme) in 150 mM McIlvaine buffer pH 5.2 supplemented with 0.1% Triton X-100, 0.2% (w/v) sodium taurocholate, 13 pmol *Cj*GH35 (β-galactosidase) in 150 mM McIlvaine buffer pH 4.5 and 10 pmol β-mannosidase (*Helix pomatia*) in 150 mM McIlvaine buffer pH 5.0 supplemented with 0.1% (w/v) bovine serum albumin (BSA) for stabilization of the recombinant proteins. Samples were denatured with sample buffer (4x Laemmli buffer, containing 50% (v/v) 1M Tris-HCl pH 6.8, 50% (v/v) 100% glycerol, 10% (w/v) Dithiothreitol (DTT), 10% (w/v) sodium dodecyl sulphate (SDS), 0.01% bromophenol blue) and heated at 100 ℃ for 5 minutes. Proteins were resolved by electrophoresis in sodium dodecylsulfate

(SDS-PAGE) 10% polyacrylamide gels, running at a constant of 90V for 30 minutes followed by 120V for approximately 60 minutes. Wet slab gels were scanned on fluorescence using a Typhoon FLA9500 Imager (GE Healthcare) using λ_{EX} 635 nm; λ_{EM} > 665 nm. Images were acquired, processed and quantified with Image Quant (GE Healthcare).

Competitive ABPP (recombinant enzyme)

Recombinant glycosidases were dissolved in the appropriate buffer as described above, and incubated with 100 μM of matching cyclophellitol epoxides **2**, **3** or **4** at 37 ℃ for 1 hour. Then, the enzymes were labelled with the matching ABP **6** (100 nM), **7** (500 nM) or **9** (10 μM) at 37 ℃ for 30 minutes. Additionally a positive control was performed; the enzyme was incubated in buffer supplemented with 1% DMSO for 1 h at 37 °C, and subsequently labelled with the matching ABP with the concentration described above. A negative control was also performed, which was acquired by heat-inactivation of proteins (adding 4x Laemmli buffer and heating at 100 °C for 5 minutes) prior to ABP labelling. The samples were denatured and analysed by SDS-PAGE as described above.

***In vitro* labelling of lysates and SDS-PAGE analysis**

Mouse kidney lysates (50 μg total protein per sample), mouse liver lysates (40 μg total protein per sample) or human fibroblasts lysates (10 μg total protein per sample) were dissolved in McIlvaine buffer pH 5.0 for 5 minutes on ice. The samples were incubated with epoxide inhibitors **2** and/or **3** (100 μM), or DMSO (1% v/v) for 1 hour at 37 ℃, followed by incubation with 1 μM ABP **6**, **7** or **9** for 30 minutes at 37 ℃. The samples were denatured and analysed by SDS-PAGE as described above.

Crystal structure of TB562 in *Cj*GH35

*Cj*GH35 was purified and crystallized as described previously.[19] A crystal was soaked in the presence of a speck of **TB562** powder for 70 hours. The crystal was fished directly into liquid nitrogen without the need for additional cryoprotectant. Data were collected on beamline IO2 at the Diamond Light Source at wavelength 0.97950 Å, and were processed using *DIALS*[31] and scaled with *AIMLESS*[32] to 1.6 Å. In contrast to previous *Cj*GH35 structures, the space group was P1 and the unit cell dimensions, 98.9, 115.8, 116.0 Å, and angles, 90.2, 90.2, 90.4° and was twinned. The structure was solved using programs from the *CCP4* suite;[33] molecular replacement was performed the native coordinates, PDB entry 4D1I, as the model and with refinement using twinned intensities. Full details of data quality and refinement statistics are given in the header information to PDB 5JAW.

Pull-down and LC-MS/MS analysis

All pull-down experiments were performed in duplicates. Mouse kidney lysates, mouse liver lysates (1.0 mg total protein) or human fibroblasts lysates (250 μg total protein) were incubated with either 0.25% (v/v) DMSO, ABP **8** or **10** (50 μM) for 3 hours at 37 ℃, in a total volume of 200 μL McIlvaine buffer pH 5.0, followed by denaturation by addition of 10% (w/v) SDS 50 μL and boiling for 5 minutes at 100 °C. Samples were then reduced with the use of dithiothreitol (DTT), alkylated by

iodoacetamide (IAA) and further prepared for pull-down with DynaBead MyOne Straptavidin Beads C1 as published previously.[34] After pull-down, all samples were used for on-bead digestion. The samples were treated with on-bead digestion buffer (100 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM $CaCl_2$, 2% (v/v) acetonitrile and 10 ng/µL trypsin) and incubated overnight in a shaker at 37 ℃. The supernatant, containing tryptic-digested peptides was then desalted using StageTips. Consequently, the acetonitrile was evaporated using a SpeedVac at 45 ℃ followed by addition of 20 µL of LC-MS sample solution (95:3:0.1, H_2O:acetonitrile:formic acid) for LC-MS analysis. All peptide samples were analysed with a two hour gradient of 5% to 25% acetonitrile on nano-LC, hyphenated to an LTQ-Orbitrap and identified using the Mascot protein search engine.[35] Raw data was calculated using MaxQuant against Uniprot of human (for fibroblasts) or mouse (for mouse kidney and mouse liver) proteome database to obtain an identification list of found proteins. The abundance of the protein hits was quantified as previously described,[25] in unsupervised mode using the default settings of the PLGS (Waters) and IsoQuant software.

Kinetics

Due to the high potency of **6** towards recombinant GBA1, kinetics of all probes were determined following the method described by Wu *et al.*[36] Recombinant enzymes GBA1 (Cerezyme, β-glucosidase) and *Cj*GH35 (overexpressed bacterial β-galactosidase) were used for kinetic experiments. For GBA1 (3.6 nM), 150 mM McIlvaine buffer pH 5.2, supplemented with 0.2% Sodium Taurocholate, 0.1% Triton X-100 and 0.1% BSA was used. For *Cj*GH35 (4.8 nM), 150 mM McIlvaine buffer pH 4.5 supplemented with 0.1% BSA was used. Kinetics were measured by adding a series of concentrations of ABP **6**, **7** or **9** (162.5 µL, maximal reaction concentration was 20 µM due to limited availability of the ABPs) to β-4MU-Glc (1300 µL, reaction [S] 2345 µM) and β-4MU-Gal (1300 µL, reaction [S] 837 µM) at 37 ℃. The t=0 samples were prepared by taking 112.5 µL from the [ABP + S] to a 96-well plate in duplo, after which stop buffer (200 µL, glycine/MeOH 1M, pH 10) and lastly the enzyme (12.5 µL) was added. Then, enzyme (137.5 µL) was added (t=0) to [ABP + S]. At t=2.5, 5.0, 7.5, 10.0 and 12.5 min, 125 µL of the reaction mixture was taken and added to stop buffer (200 µL, glycine/MeOH 1M, pH 10) in the 96-well plate. After 12.5 minutes, the plate was measured with a PerkinElmer Fluorescence Spectrometer LS-55 using BL Studio with excitation at 366 nm and emission at 445 nm. Kinetic experiments were performed in duplo or triplo with duplo measurements each run. Obtained values were plotted as a one phase exponential association and transferred to k_{obs} values. Which, according to Michaelis-Menten kinetics can be plotted against the concentration to determine the K_I, k_{inact} and their ratio.[36] The K_m of each enzyme was determined by incubating the same enzyme concentration as the kinetic experiments with a series of substrate concentrations for 12.5 minutes at 37 °C. Then, the reaction was stopped by addition to stop buffer and measured on fluorescence. The results were then plotted and K_m was determined according to Michaelis-Menten $K_m = ½V_{max}$.[37] K_m measurement was performed in a single measurement with a triplo of each sample. GBA1 K_m=1423 µM; *Cj*GH35 K_m=499 µM. Results were processed and analysed using GraphPad Prism 6.0.

References

1 G. Davies and B. Henrissat, *Structure*, 1995, **3**, 853–859.

2 K. Ohtsubo and J. D. Marth, *Cell*, 2006, **126**, 855–867.

3 M. J. Evans and B. F. Cravatt, *Chem. Rev.*, 2006, **106**, 3279–3301.

4 Y. Liu, M. P. Patricelli and B. F. Cravatt, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 14694–14699.

5 J. Z. Long, W. Li, L. Booker, J. J. Burston, S. G. Kinsey, J. E. Schlosburg, F. J. Pavón, A. M. Serrano, D. E. Selley, L. H. Parsons, A. H. Lichtman and B. F. Cravatt, *Nat. Chem. Biol.*, 2008, **5**, 37–44.

6 S. Atsumi, K. Umezawa, H. Iinuma, H. Naganawa, H. Nakamura, Y. Iitaka and T. Takeuchi, *J. Antibiot.*, 1989, **43**, 49–53.

7 L. I. Willems, T. J. M. Beenakker, B. Murray, B. Gagestein, H. Van Den Elst, E. R. Van Rijssel, J. D. C. Codée, W. W. Kallemeijn, J. M. F. G. Aerts, G. A. Van Der Marel and H. S. Overkleeft, *Eur. J. Org. Chem.*, 2014, 6044–6056.

8 V. W. F. T. Tony, K. M. Shing, *J. Chem. Soc. Perkin Trans. 1*, 1994, 2017–2025.

9 W. W. Kallemeijn, K. Y. Li, M. D. Witte, A. R. A. Marques, J. Aten, S. Scheij, J. Jiang, L. I. Willems, T. M. Voorn-Brouwer, C. P. A. A. Van Roomen, R. Ottenhoff, R. G. Boot, H. Van Den Elst, M. T. C. Walvoort, B. I. Florea, J. D. C. Codée, G. A. Van Der Marel, J. M. F. G. Aerts and H. S. Overkleeft, *Angew. Chem. Int. Ed.*, 2012, **51**, 12529–12533.

10 J. Jiang, C. L. Kuo, L. Wu, C. Franke, W. W. Kallemeijn, B. I. Florea, E. van Meel, G. A. van der Marel, J. D. C. Codée, R. G. Boot, G. J. Davies, H. S. Overkleeft and J. M. F. G. Aerts, *ACS Cent. Sci.*, 2016, **2**, 351–358.

11 L. I. Willems, T. J. M. Beenakker, B. Murray, S. Scheij, W. W. Kallemeijn, R. G. Boot, M. Verhoek, W. E. Donker-Koopman, M. J. Ferraz, E. R. Van Rijssel, B. I. Florea, J. D. C. Codée, G. A. Van Der Marel, J. M. F. G. Aerts and H. S. Overkleeft, *J. Am. Chem. Soc.*, 2014, **136**, 11622–11625.

12 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.

13 F. G. Hansen, E. Bundgaard and R. Madsen, *J. Org. Chem.*, 2005, **70**, 10139–10142.

14 D. Lee, C. L. Williamson, L. Chan and M. S. Taylor, *J. Am. Chem. Soc.*, 2012, **134**, 8260–8267.

15 A. G. Volbeda, H. A. V. Kistemaker, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov and J. D. C. Codée, *J. Org. Chem.*, 2015, **80**, 8796–8806.

16 G. Franck, K. Brödner and G. Helmchen, *Org. Lett.*, 2010, **12**, 3886–3889.

17 G. Franck, M. Brill and G. Helmchen, *Org. Synth.*, 2014, 55–65.

18 P. K. Brown and H. C. Jadhav, *J. Am. Chem. Soc.*, 1983, **105**, 2092–2093.

19 J. Larsbrink, A. J. Thompson, M. Lundqvist, J. G. Gardner, G. J. Davies and H. Brumer, *Mol. Microbiol.*, 2014, **94**, 418–433.

20 S. G. Withers, *Carbohydr. Polym.*, 2001, **44**, 325–337.

21 G. J. Davies, A. Planas and C. Rovira, *Acc. Chem. Res.*, 2012, **45**, 308–316.

22 S. van Weely, J. M. F. G. Aerts, M. B. van Leeuwen, J. C. Heikoop, W. E. Donker-Koopman, J. A. Barranger, J. M. Tager and A. W. Schram, *Eur. J. Biochem.*, 1990, **191**, 669–677.

23 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.

24 S. Tomino and M. Meisler, *J. Biol. Chem.*, 1975, **250**, 7752–7758.

25 J. Kuharev, P. Navarro, U. Distler, O. Jahn and S. Tenzer, *Proteomics*, 2015, **15**, 3140–3151.

26 S. Nagano, T. Yamada, N. Shinnoh, H. Furuya, T. Taniwaki and J. I. Kira, *Clin. Chim. Acta*, 1998, **276**, 53–61.

27 P. Luzi, M. A. Rafi, M. Zaka, M. Curtis, M. T. Vanier and D. A. Wenger, *Mol. Genet. Metab.*, 2001, **73**, 211–223.

28 J. Le Carré, D. F. Schorderet and S. Cottet, *Mol. Vis.*, 2011, **17**, 1287–1297.

29 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.

30 F. M. Cordero, P. Bonanno, M. Chioccioli, P. Gratteri, I. Robina, A. J. Moreno Vargas and A. Brandi, *Tetrahedron*, 2011, **67**, 9555–9564.

31 D. G. Waterman, G. Winter, R. J. Gildea, J. M. Parkhurst, A. S. Brewster, N. K. Sauter and G. Evans, *Acta Crystallogr. Sect. D, Struct. Biol.*, 2016, **72**, 558–575.

32 P. R. Evans and G. N. Murshudov, *Acta Crystallogr. Sect. D, Biol. Crystallogr.*, 2013, **69**, 1204–1214.

33 M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2011, **67**, 235–242.

34 N. Li, C.-L. Kuo, G. Paniagua, H. van den Elst, M. Verdoes, L. I. Willems, W. A. van der Linden, M. Ruben, E. van Genderen, J. Gubbens, G. P. van Wezel, H. S. Overkleeft and B. I. Florea, *Nat. Protoc.*, 2013, **8**, 1155–1168.

35 Y. Ishihama, Y. Oda, T. Tabata, T. Sato, T. Nagasu, J. Rappsilber, M. Mann, *Mol. Cell. Proteomics*, 2005, **4**, 1265–1272.

36 L. Wu, J. Jiang, Y. Jin, W. W. Kallemeijn, C.-L. Kuo, M. Artola, W. Dai, C. van Elk, M. van Eijk, G. A. van der Marel, J. D. C. Codée, B. I. Florea, J. M. F. G. Aerts, H. S. Overkleeft and G. J. Davies, *Nat. Chem. Biol.*, 2017, **13**, 867–873.

37 K. A. Johnson and R. S. Goody, *Biochemistry*, 2011, **50**, 8264–8269.

Chapter 4

D-*arabino*- and D-*lyxo*furanosyl cyclitol aziridines for unbiased glycosidase profiling

4.1 Introduction

Activity-based protein profiling (ABPP) has proven to be a powerful method for the identification of several classes of hydrolytic enzymes, including glycosidases.[1] Recent research has shown that carbasugars mimicking the natural substrate configuration are potent irreversible glycosidase inhibitors. For example, the configurational β-glucopyranose mimic ABP **1** (Figure 1) effectively labels the endogenous β-glucosidases, GBA1, GBA2, GBA3 and LPH in various mouse tissue lysates without off-target labeling of other proteins present in the complex biological sample.[2] Similarly, other configurational pyranoside probes such as **2** and **3** have been employed to selectively identify α-D-galactosidases and α-L-fucosidases, respectively (Figure 1).[3,4] With these specific ABPs in hand, the glycosidase composition of various biological samples could be studied. Moreover, this technique allows the study on the effect of small-molecule inhibitors on the activity of the target enzymes by competitive ABPP experiments.[5] However, the scope of glycosidases studied in such experiments is limited by the high specificity of these ABPs. Therefore it would be of interest to develop non-specific ABPs which display activity towards a wide variety of

glycosidases. With such probes, the overall glycosidase composition in biologically relevant settings as well as its response to external factors could be studied.

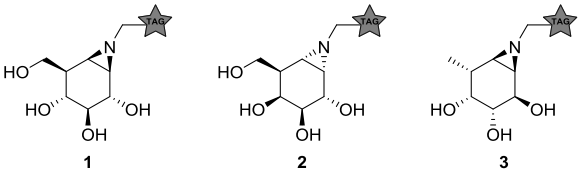


Figure 1 Activity-based probes **1**-**3** used for the selective identification of β-glucosidases, α-galactosidases and α-L-fucosidases, respectively.

The selective glycosidase ABPs discussed above were pyranoside carbasugar mimics. Probes emulating furanosides have not been investigated, while glycosidases acting on furanoside substrates are present in GH families 32, 43, 62, 68, 117 and 130.[6] Bols and co-workers have published the synthesis of D-*arabino*furanosyl aziridine **4** (Figure 2),[7] and shown that it adopts a $^{1,4}B$ conformation (carbohydrate numbering) in solution. While the five-membered ring resembles a furanoside, one could argue that this conformation also mimics pyranoside carbohydrates in a $^{1,4}B$ conformation, which is the Michaelis complex conformation found during the catalytic itinerary of several (GH 1, 5, 7, 12, 18, 20 and 26) retaining glycosidases acting on pyranosides.[8] Therefore, D-*arabino*furanosyl aziridines could potentially serve as covalent inhibitors of retaining glycofuranosidases, as well as glycopyranosidases. While aziridine **4** displayed zero activity towards α-glucosidase (yeast) or β-glucosidase (almonds), alkylation of the aziridine with a methyl (**5**) or benzyl group (**6**) positively induced competitive inhibition towards these enzymes. Although inhibition was moderate, it was anticipated that the incorporation of a large aliphatic reporter tag could further increase its potency (a phenomenon also observed for **1**).[9]

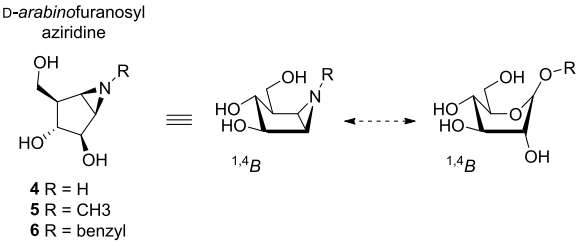
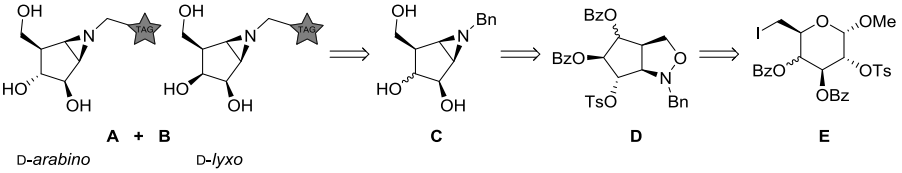


Figure 2 D-*arabino*furanosyl aziridines **4**-**6** synthesized by Bols and co-workers adopt a $^{1,4}B$ conformation in solution.

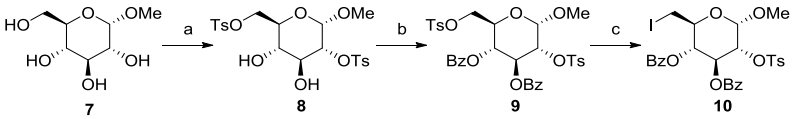
In this Chapter, the synthesis of furanosyl aziridine inhibitors equipped with a reporter tag is described. The synthesis route of Bols and co-workers is adapted and further optimized to afford D-*arabino*furanosyl aziridine derivatives, and the synthesis strategy is then translated to a similar starting material to afford novel D-*lyxo*furanosyl aziridines. It was envisioned that D-*arabino*furanosyl aziridines **A** as well as D-*lyxo*furanosyl aziridines **B** would be available from configurationally matching **C**, which would be accessible by N-O reduction of cyclopentaisoxazolidine **D**. This cis-fused bicyclic system would be available *via* fragmentation of **E** into the corresponding aldehyde, followed by nitrone formation and subsequent [3+2] cycloaddition. The biochemical evaluation of these compounds focused on labeling of glycosidases in an unbiased fashion, as analyzed by SDS-PAGE and proteomics.



Scheme 1 Retrosynthetic analysis of the target D-*arabino*- and D-*lyxo*furanosyl aziridines.

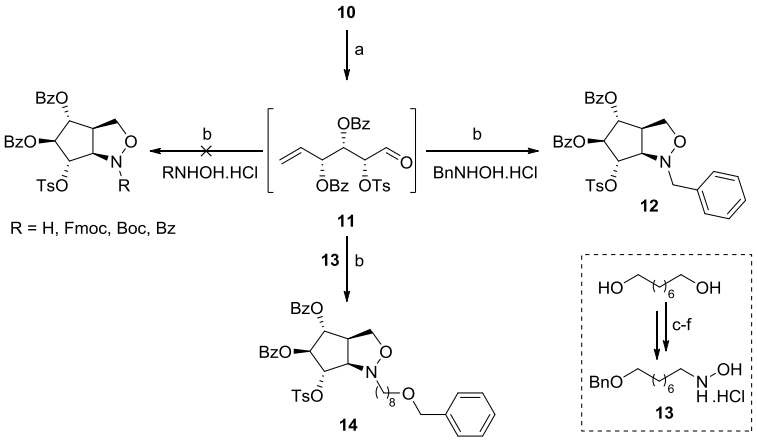
4.2 Results and Discussion

The synthesis of D-*arabino*furanosyl aziridine **4** commenced with the preparation of synthon **10** (Scheme 2). A previously described procedure by Ferrier[10] to prepare intermediate **8** from **7** proved to be low yielding, therefore an alternative route was employed. Methyl α-glucopyranoside **7** was selectively tosylated at position O2 and O6 using a catalytic amount of dibutyltin dichloride, affording bis-tosylate **8**.[11,12] Subsequent benzoylation of the remaining hydroxyls furnished compound **9** which was then regioselectively iodinated at C-6 to afford key-intermediate **10**.



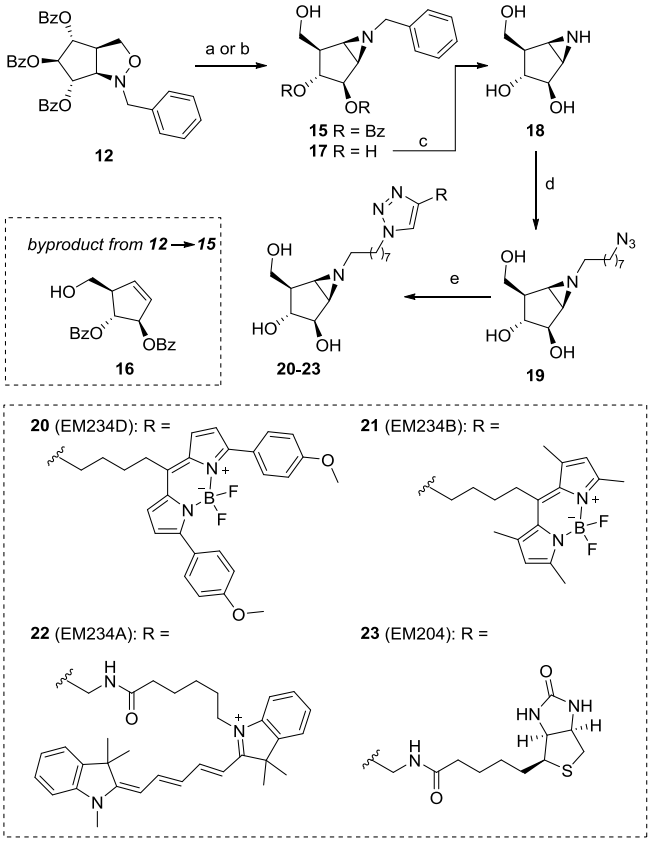
Scheme 2 Synthesis of iodopyranoside **10**. Reagents and conditions: a) TsCl, Bu_2SnCl_2 (40 mol%), Et_3N, MeCN, rt, 24h, 97%; b) Bz_2O, DMAP, pyridine, rt, 16h, 90%; c) NaI, Ac_2O, reflux, 4h, quant.

Iodide **10** was then subjected to a Vasella fragmentation[13] using zinc dust in refluxing ethanol to afford intermediate aldehyde **11**, which was found to be unstable during isolation and was therefore directly used in the next step (Scheme 3). The aldehyde was reacted with *N*-benzyl hydroxylamine, affording the corresponding nitrone intermediate which underwent smooth intramolecular [3+2] cycloaddition with the terminal alkene to afford *N*-benzyl isoxazolidine **12** as a single isomer. Additionally, it would be of interest to investigate the incorporation of other functionalities/protecting groups on the isoxazolidine nitrogen. Therefore, condensation of aldehyde **11** with other hydroxylamines was studied. While the commercially available unprotected hydroxylamine as well as Fmoc-, Boc- and Bz-protected hydroxylamines were found to give complex mixtures upon reaction with aldehyde **11**, *N*-(8-(benzyloxy)octyl)hydroxylamine **13** (obtained by mono-benzylation of 1,8-octanediol followed by oxidation, reductive amination and reduction, respectively) reacted with the aldehyde and underwent subsequent [3+2] cycloaddition to give bicyclic **14**. Although this method could ensure preliminary installation of the alkyl linker in the synthesis route, the [3+2] cycloaddition with hydroxylamine **13** was low yielding and therefore not further pursued. Instead, the synthesis was continued with bicyclic *N*-benzylated **12**.



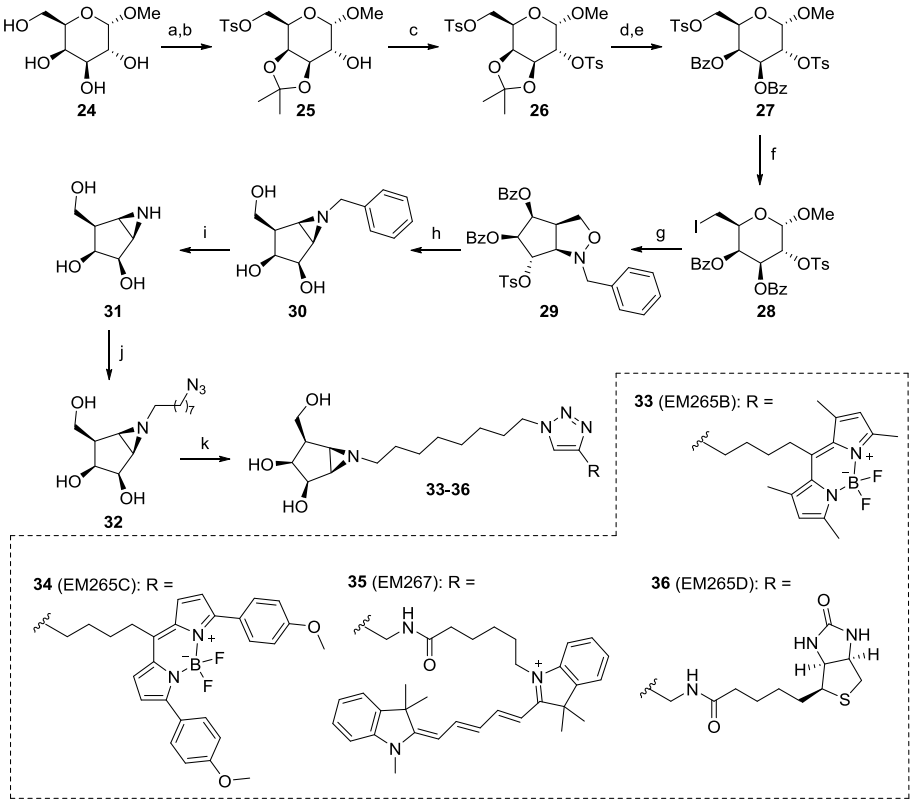
Scheme 3 Synthesis of the isoxazolidines. Reagents and conditions: a) Zn, EtOH, reflux, 2h, work-up; b) RNHOH.HCl, Na_2CO_3, toluene, 50 °C, 2h, for **12**: 61% over 2 steps; for **14**: 24% over 2 steps; c) BnBr, NaH, TBAI, THF, 16h; d) $(COCl)_2$, DMSO, DCM, -78 °C, 1h; e) $NH_2OH.HCl$, NaOH, EtOH, reflux, 1h; f) $NaBH_3CN$, HCl, MeOH, rt, 5h, 17% overall yield over 4 steps.

As reported by Bols et al.,[7] reductive N-O bond cleavage can be achieved using Raney nickel under hydrogen pressure. However the reported yield was moderate (48%), possibly due to concomitant *N*-debenzylation. The use of Raney nickel under hydrogen atmosphere indeed resulted in a moderate yield (52%) after prolonged reaction times (72 h). Alternatively, zinc in acetic acid is reported to reductively cleave N-O bonds.[14] While the reaction of isoxazolidine **12** with zinc in acetic acid gave no reaction at room temperature, heating the reaction to 55 °C resulted in the isolation of product **15** (50%) as well as byproduct **16** (25%, Scheme 4). Lowering the temperature to 35 °C and subsequent work-up followed by basification with Et_3N resulted in the isolation of



Scheme 4 Synthesis of *N*-subsituted D-*arabino*furanosyl aziridines. Reagents and conditions: a) Zn, AcOH, 35 °C, 2h, work-up, then Et_3N, MeOH, 97% **15**; b) Zn, AcOH, 35 °C, 2h, work-up, then NaOMe, MeOH, quant. **17**; c) $Pd(OH)_2/C$, H_2, H_2O, 1h, 45%; d) 1-azido-8-iodooctane, K_2CO_3, DMF, 80 °C, 16h, 40%; e) tag-alkyne, $CuSO_4$, sodium ascorbate, DMF/H_2O, rt, 16h, **20** 37%; **21** 57%; **22** 22%; **23** 40%.

15 in 96% yield. Furthermore, basification of the crude amine with sodium methoxide led to the cyclization to the aziridine followed by global debenzoylation, leading to *N*-benzyl cyclitol **17** in quantitative yield. The aziridine was debenzylated using Pearlman's catalyst in water under hydrogen atmosphere. To avoid reductive opening of the aziridine ring the reaction times were minimized and aziridine **18** could be isolated in 45% yield. Finally, the aziridine was alkylated with 1-azido-8-iodooctane to afford *N*-octylazido aziridine **19** which was equipped with various reporter tags via a Huisgen azide/alkyne 1,3-dipolar cycloaddition[15] to afford *N*-substituted D-arabinofuranosyl aziridines **20**-**23** after HPLC purification.



Scheme 5 Synthesis of D-*lyxo*furanosyl activity based probes. Reagents and conditions: a) TsCl, Et_3N, MeCN, rt, 16h; b) *p*-TsOH, 2,2-dimethoxypropane, DMF, 63% over 2 steps; c; TsCl, Et_3N, MeCN, 40 °C, 16h, 50%; d) HCl, MeOH, 16h, quant.; e) Bz_2O, DMAP, pyridine, rt, 16h, 70%; f) NaI, Ac_2O, reflux, 4h, 88%; g) Zn, EtOH, reflux, 6h, work-up, then BnNHOH.HCl, Na_2CO_3, toluene, 50 °C, 16h, 95%; h) Zn, AcOH, 40 °C, work-up, then NaOMe, MeOH, rt, 3h, quant.; i) $Pd(OH)_2/C$, H_2, H_2O, 30 min, 46%; j) 1-azido-8-iodooctane, K_2CO_3, DMF, 80 °C, 16h, 46%; k) tag-alkyne, $CuSO_4$, sodium ascorbate, DMF/H_2O, rt, 16h, **33** 35%; **34** 29%; **35** 24%; **36** 29%.

The synthesis of D-*lyxo*furanosyl aziridine follows a similar strategy (Scheme 5). Starting from commercially available methyl α-D-galactopyranoside **24**, the primary alcohol was tosylated and the resulting product was protected with a 3,4-isopropylidene acetal resulting in compound **25** in 63% over two steps. Tosylation of the remaining secondary alcohol required gentle heating and product **26** could be isolated in moderate yield. Quantitative deprotection of the isopropylidene acetal was achieved using hydrochloric acid in methanol, which after benzoylation yielded product **27**. S_n2 displacement of the tosyl group by iodine afforded iodopyranoside **28**. Vasella fragmentation followed by work-up, nitrone formation and *in situ* [3+2] dipolar cycloaddition resulted in isoxazolidine **29**. Next, reductive cleavage of the N-O bond and subsequent ring-closure proceeded smoothly, and the crude product was directly deprotected under Zemplén conditions affording *N*-benzyl aziridine **30** in excellent yield over two steps. Debenzylation using Pearlman's catalyst under hydrogen atmosphere afforded aziridine cyclitol **31**, which was alkylated yielding *N*-azidoalkyl aziridine **32**. Several reporter tags were attached to the linker moiety *via* a Huisgen azide/alkyne 1,3-dipolar cycloaddition to afford *N*-alkylated D-*lyxo*furanosyl aziridines **33**-**36**.

With fluorescent D-*arabino*- (**22**) and D-*lyxo*furanosyl (**35**) aziridines in hand, their ability to label various glycosidases in human c920 fibroblast lysates was evaluated at lysosomal (4.0) and cytosolic (7.0) pH (Figure 3A and 3B, respectively). Both aziridines **22** and **35** showed significant labeling of bands at 10 μM; at higher concentrations background fluorescence was observed. At this concentration, fibroblast lysate was incubated with the aziridines at different pH's (Figure 3C and 3D). Several bands appeared following labeling at pH ranging from 4-9 indicating the high reactivity of the aziridine warhead over a wide pH-range, consistent with labeling patterns of analogous cyclitol aziridines.[16] These initial results suggested that for both **22** and **35**, the optimal pH for labeling in fibroblast lysate was ~5.0. Both fluorescent aziridines displayed a similar labeling pattern in fibroblast lysate, with a strong band at ~42 kDa. Moreover, for both aziridines a band at the molecular weight of GBA1 (MW ~56 kDa, annotated with an asterisk) was visualized indicating a possible covalent interaction with this enzyme. Additionally, the unbiased labeling of glycosidases in WT mouse liver lysate (C57bl/6j, Jackson's laboratories) was probed with **21** and **33** (Figure 3E). Again, a band at ~42 kDa was labeled by **21**. However, while the presence of GBA1 could clearly be indicated by the GBA1 selective probe MDW933,[17] the enzyme did not appear to be labeled with 10 μM **21**. In contrast, while

labeling with **33** also showed a band at ~42 kDa, a strong band at ~56 kDa appeared as well, possibly caused by labeling of GBA1.

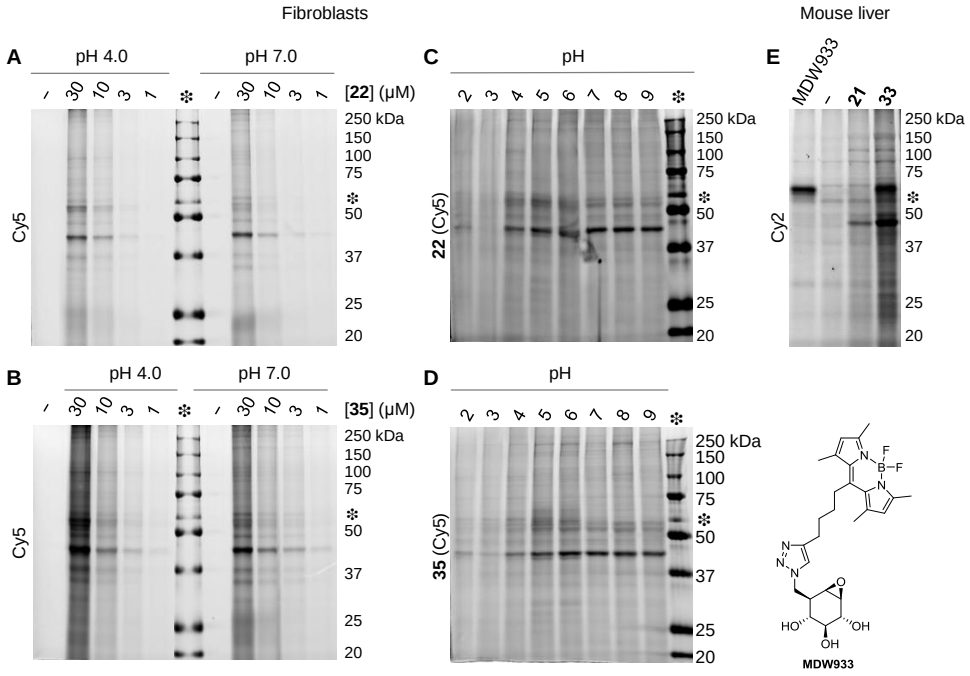


Figure 3 Biochemical evaluation of fluorescent D-*arabino*- (**21** and **22**) and D-*lyxo*furanosyl aziridines (**33** and **35**) in c920 fibroblast lysates; A) concentration range for D-*arabino*-aziridine **22** at cytosolic and lysosomal pH; B) concentration range for D-*lyxo*furanosyl aziridine **35** at cytosolic and lysosomal pH; C) pH range for D-*arabino*-aziridine **22**; D) pH range for D-*lyxo*furanosyl aziridine **35**; E) labeling of **21** and **33** in C57bl/6j mouse liver lysates; MDW933 is a GBA1 selective probe; (-) represents a control experiment without addition of probe; the band for GBA1 is annotated with an asterisk (❊).

Next, competitive ABPP experiments were performed (Figure 4). Fibroblast lysates were pre-incubated with different concentrations (1, 10 or 100 μM) of fluorescent aziridines **21** or **33** (Figure 4A and 4B, respectively) and subsequently labeled with matching concentrations of biotin-aziridines (**23** or **36**, respectively). The biotin-labeled proteins were then visualized by a streptavidin-Cy5 blot and the bands were compared to the matching non-competed lanes. Unfortunately, no significant competition of the bands at ~42 kDa, ~56 kDa or other bands was observed for either **23** or **36** at any of the concentrations used. A similar competition experiment was executed in mouse liver lysates at pH 5.0 (Figure 4C). The samples were treated with

10 µM biotin-aziridine (**23** or **36**, respectively) with or without competition of 10 µM fluorescent aziridine (**21** or **33**, respectively). Again, no significant competition of biotin-labeled proteins could be observed, although detection might have been hampered by the high amount of endogenous biotinylated proteins present. Ultimately, streptavidin pull-down experiments were performed with **23** and **36** on various lysates (human fibroblast, mouse kidney, liver, duodenum and brain) and the labeled proteins were analyzed after tryptic digestion followed by LC-MS/MS proteomics. No glycosidase hits could be identified in any of the lysates, therefore further optimization of labeling conditions was abandoned.

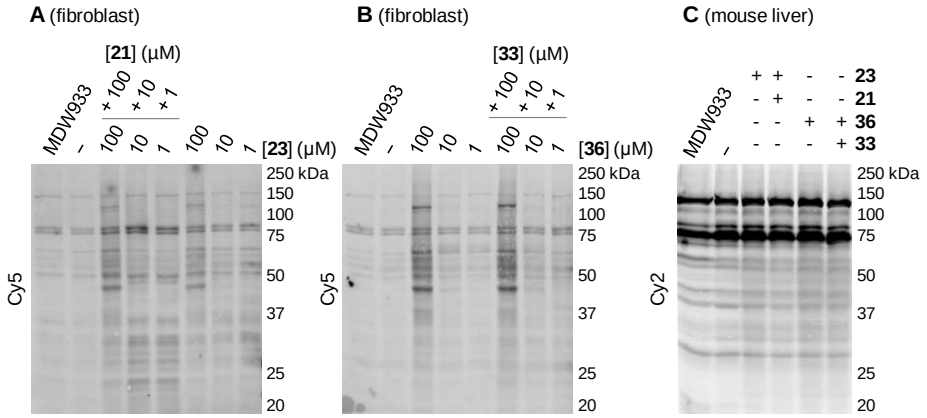


Figure 4 Competition experiments matching biotin and Cy2 tagged aziridines; A) competition of **23** with **21** in human c920 fibroblast lysates; B) competition of **36** with **33** in c920 fibroblast lysates; C) competition of **23** with **21** and **36** with **33** in mouse liver lysate.

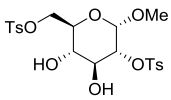
4.3 Conclusion

In the past, activity-based glycosidase profiling studies employed covalent pyranose-configured inhibitors to mimic the substrate of the specific enzyme of interest. In this Chapter, this method is expanded to furanosyl configured aziridines, with the aim to label glycosidases in an unbiased fashion. Fluorescent cyclitol aziridines with D-*arabino* (**20**-**23**) and D-*lyxo* (**33**-**36**) configurations were synthesized and their labeling efficiency at different concentrations and pH values was evaluated in human fibroblast- and mouse kidney lysates. While for both configurations different fluorescent bands could be observed in complex biological samples, competitive ABPP with the matching biotinylated aziridines did not demonstrate significant competition of these signals. Additionally, pull-down experiments were performed on various biological samples but no glycosidase hits could be identified. Ultimately, it was concluded that there were no enzymes present in human fibroblast and mouse kidney lysates which were capable of recognizing D-*arabino*- or D-*lyxo*furanose configured cyclitols as their substrates, although it would be of interest to study the viability of other furanose-based cyclitols (see future prospects, 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).

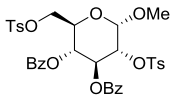
Compound 8



To a stirred solution of dry Et_3N (0.97 mL, 7 mmol) in dry MeCN (30 mL) was added methyl-α-D-glucopyranoside (485 mg, 2.5 mmol), Bu_2SnCl_2 (304 mg, 1.0 mmol) and tosyl chloride (1.14 g, 6.0 mmol) at rt under argon atmosphere and stirred for 24 hours. The mixture was quenched with water and extracted with EtOAc (300 mL). The extract was washed successively with water, aqueous sodium hydrogen carbonate and again with water, followed by drying over $MgSO_4$ and evaporation. Purification of the crude product by column chromatography (pentane/EtOAc, 1:1) yielded product **8** (1.22 g, 97%). ^1H-NMR (400 MHz, $CDCl_3$) δ 7.80 (m, 4H), 7.36 (dd, J = 7.8, 6.6 Hz, 4H), 4.64 (d, J = 3.7 Hz, 1H), 4.29 (dd, J = 11.1, 4.6 Hz, 1H), 4.25 – 4.17 (m, 2H), 3.91 (t, J = 9.3 Hz, 1H), 3.72 (ddd, J = 10.0, 4.5, 2.0 Hz, 1H), 3.47 (t, J = 9.4 Hz, 1H), 3.25 (s, 3H), 2.78 (s, 1OH), 2.63 (s, OH), 2.46 (s, 6H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 145.4, 130.0, 129.9, 128.1, 127.9, 97.3, 78.8, 77.3, 77.0, 76.7, 71.0, 69.7, 68.7, 68.3, 55.6, 29.7, 21.7. IR (neat, cm^{-1}): υ 3510, 2929,

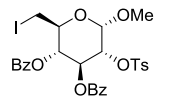
1597, 1355, 1172, 1033, 970, 929, 810, 665. $[\alpha]_D^{20}$ (c 1.0, CH_2Cl_2): +47. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{21}H_{26}O_{10}S_2$ 503.10538 found 503.10437.

Compound 9



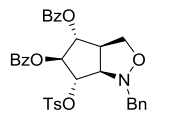
Compound **8** (5.7 g, 11.4 mmol) was dissolved in dry pyridine (300 mL). Benzoyl anhydride (10.3 g, 45 mmol) and DMAP (0.28 g, 3.3 mmol) were added and stirred 16 h at rt. The mixture was quenched with sat. aq. $NaHCO_3$ and extracted with EtOAc. The extract was washed successively with water, sat. aq. $NaHCO_3$ and water, dried over $MgSO_4$ and evaporated. Purification of the crude product by column chromatography (pentane/EtOAc, 2:1) yielded product **9** (7.3 g, 10.3 mmol, 90%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.79 – 7.67 (m, 2H), 7.64 – 7.54 (m, 2H), 7.48 (td, *J* = 7.4, 1.3 Hz, 1H), 7.36 – 7.24 (m, 2H), 7.20 (d, *J* = 8.0 Hz, 1H), 6.93 (d, *J* = 8.0 Hz, 1H), 5.82 (t, *J* = 9.7 Hz, 0H), 5.02 (d, *J* = 3.6 Hz, 0H), 4.53 (dd, *J* = 10.0, 3.6 Hz, 1H), 4.23 – 4.13 (m, 1H), 4.06 (dd, *J* = 11.5, 5.9 Hz, 1H), 3.46 (s, 1H), 2.33 (s, 1H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 165.0, 149.0, 145.1, 145.0, 137.0, 133.7, 133.2, 133.1, 132.7, 132.3, 130.1, 130.0, 129.9, 129.9, 129.8, 129.8, 128.8, 128.6, 128.5, 128.4, 128.2, 128.1, 128.0, 127.7, 97.8, 76.4, 69.5, 68.8, 67.8, 67.4, 56.3, 21.7. IR (neat, cm^{-1}): υ 2943, 1720, 1357, 1273, 1174, 1093, 1066, 1020, 975, 812, 657. $[\alpha]_D^{20}$ (c 0.6, CH_2Cl_2): +15. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{35}H_{34}O_{12}S_2$ 711.15765, found 711.15680.

Compound 10



Sodium iodide (5.0 g, 34 mmol) was added to a mixture of **9** (19.9 g, 28 mmol) in acetic anhydride (50 mL) and refluxed for 4 hours. After cooling to rt, the mixture was filtered over celite, diluted with dichloromethane (200 mL), washed with aqueous sodium thiosulphate, water, dried over $MgSO_4$, filtered, and concentrated. The solid residue was recrystallized from methanol to yield compound **10** (18.6 g, 28 mmol, quant.) as a white solid. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.86 (d, *J* = 7.9 Hz, 2H), 7.62 (dd, *J* = 17.9, 8.1 Hz, 4H), 7.49 (q, *J* = 7.2 Hz, 2H), 7.34 (t, *J* = 7.7 Hz, 2H), 7.29 (d, *J* = 7.7 Hz, 2H), 6.96 (d, *J* = 8.2 Hz, 2H), 5.89 (t, *J* = 9.7 Hz, 1H), 5.20 (t, *J* = 9.6 Hz, 1H), 5.09 (d, *J* = 3.6 Hz, 1H), 4.60 (dd, *J* = 10.0, 3.6 Hz, 1H), 4.05 – 3.95 (m, 1H), 3.57 (s, 3H), 3.34 (dd, *J* = 11.0, 2.5 Hz, 1H), 3.19 (dd, *J* = 11.0, 8.5 Hz, 1H), 2.21 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 165.4, 165.0, 145.0, 133.8, 133.2, 130.3, 130.0, 129.9, 129.8, 128.6, 128.5, 128.2, 127.8, 97.9, 76.8, 72.7, 69.2, 69.1, 56.4, 3.6. IR (neat, cm^{-1}): υ 1681, 1452, 1323, 1178, 931, 802, 667, 634, 617. $[\alpha]_D^{20}$ (c 0.7, CH_2Cl_2): +33. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{28}H_{27}IO_9S$ 689.03183, found 689.03107.

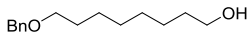
Compound 12



To a suspension of **10** (1.0 g, 1.50 mmol) in EtOH (20 mL) was added Zn dust (1.0 g,15.30 mmol). The mixture was refluxed for 2 h and then filtered through Celite and concentrated to give a yellow oil that was dissolved in CH_2Cl_2 (20 mL), washed with water (2 × 10 mL), dried over $MgSO_4$, filtered and concentrated to give unstable aldehyde **11** as an oil. To a solution of the aldehyde in toluene (8 mL) was added *N*-benzylhydroxylamine hydrochloride (359 mg, 2.25 mmol) and Na_2CO_3 (225 mg, 2.25 mmol), and the

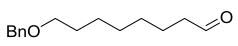
reaction was stirred at 50 ^{0}C for 2 h. The mixture was then concentrated and the residue was dissolved in CH_2Cl_2 (15 mL), washed with water (2 × 10 mL), dried over $MgSO_4$, filtered and concentrated. After crystallization from methanol, product **12** was obtained (0.56 g, 0.92 mmol, 61%) as a white solid. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.00 – 7.93 (m, 2H), 7.88 (dd, *J* = 8.3, 1.2 Hz, 2H), 7.72 (d, *J* = 8.3 Hz, 2H), 7.55 (t, *J* = 7.4 Hz, 2H), 7.41 (td, *J* = 7.6, 5.5 Hz, 4H), 7.36 – 7.22 (m, 6H), 7.07 (d, *J* = 8.1 Hz, 2H), 5.91 (t, *J* = 8.2 Hz, 1H), 5.18 (dd, *J* = 8.1, 6.3 Hz, 1H), 5.09 (dd, *J* = 8.2, 5.6 Hz, 1H), 4.28 – 4.17 (m, 2H), 3.95 (d, *J* = 13.5 Hz, 1H), 3.91 – 3.80 (m, 2H), 3.25 (dtd, *J* = 10.0, 6.5, 3.7 Hz, 1H), 2.20 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 166.2, 165.0, 145.0, 136.7, 133.7, 133.5, 133.4, 130.1, 130.0, 129.9, 129.1, 128.9, 128.7, 128.5, 128.4, 128.0, 127.7, 84.1, 79.1, 77.4, 70.8, 70.1, 59.4, 50.2, 21.7. IR (neat, cm^{-1}): υ 3061, 2879, 1714, 1598, 1355, 1261, 1180, 1109, 1093, 1068, 962, 844, 688. $[\alpha]_D^{20}$ (c 0.8, CH_2Cl_2): +15. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{34}H_{31}NO_8S$ 614.20931, found 614.18433.

8-(benzyloxy)octan-1-ol



To a stirred suspension of NaH (3.2 g, 60 wt%, 141 mmol) in dry THF (150 mL) was added a solution of octane-1,8-diol (8.9 g, 61 mmol) in dry THF (100 mL) at 0 ^{0}C, and stirring was continued for 30 min. Then, BnBr (6.9 mL, 58 mmol) in dry THF (20 mL) was added slowly. The mixture was warmed to rt, then slowly heated with a heatgun to 50 ^{0}C and then cooled to rt. TBAI (11.1 g, 30 mmol) was added and the stirring was continued for 16 h. The reaction was cooled to 0 ^{0}C and slowly quenched with sat. NH_4Cl followed by extraction with EtOAc (3x150 mL). The organic extracts were combined and washed with H_2O and brine, and then dried over $MgSO_4$. The solvent was evaporated, followed by purification of the crude product by column chromatography (pentane/EtOAc, 5:1), to afford the product (7.51 g, 32 mmol, 52%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.12 (m, 5H), 4.50 (s, 2H), 3.63 (t, *J* = 6.6 Hz, 2H), 3.46 (t, *J* = 6.6 Hz, 2H), 1.65 – 1.50 (m, 4H), 1.42 – 1.20 (m, 8H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 138.8, 128.5, 127.8, 127.6, 73.0, 70.6, 63.2, 32.9, 29.9, 29.6, 29.5, 26.3, 25.8. IR (neat, cm^{-1}): υ 3360, 2927, 2854, 1454, 1361, 1097, 731, 694. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{15}H_{24}O_2$ 237.18645, found 237.18486.

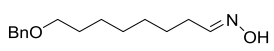
8-(benzyloxy)octanal



To a stirred solution of $(COCl)_2$ (0.51 mL, 6 mmol) in dry CH_2Cl_2 (3,6 mL) was added a pre-mixed solution of dry DMSO (0.85 mL, 12 mmol) and dry CH_2Cl_2 (5.4 mL) at -78 ^{0}C. After 30 min, a solution of 8-(benzyloxy)octan-1-ol (0.717 g, 3 mmol) in CH_2Cl_2 (3.6 mL) was added slowly at the same temperature. The resulting mixture was stirred for 1 h at -78 ^{0}C, and then Et_3N (2.11 mL, 24 mmol) was added. The mixture was warmed to rt, and stirring was continued for 1 h. A saturated solution of NH_4Cl (2 mL) was added slowly to quench the reaction followed by addition of brine, and then the mixture was extracted with CH_2Cl_2. The combined organic extracts were washed with brine, dried over $MgSO_4$, filtered and concentrated. The resulting aldehyde was purified by column chromatography (pentane/EtOAc 4:1), to afford the product (0.589 g, 2.5 mmol, 84%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 9.75 (t, *J* = 1.8 Hz, 1H), 7.31 (dd, *J* = 19.3, 15.5 Hz, 8H), 4.49 (s, 3H), 3.45 (t, *J* = 6.6 Hz, 3H), 2.41 (td, *J* = 7.4, 1.8 Hz, 3H), 1.66 – 1.52 (m, *J* = 12.8, 6.9 Hz, 6H), 1.40 –

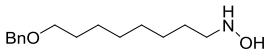
1.27 (m, *J* = 6.0, 3.4 Hz, 9H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 203.0, 138.8, 128.5, 127.8, 127.6, 73.0, 70.5, 44.0, 29.8, 29.3, 29.2, 26.1, 22.2. IR (neat, cm^{-1}): υ 2929, 2854, 1722, 1454, 1361, 1097, 1028, 731, 696. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{15}H_{22}O_2$ 235.17080 found 235.16914.

8-(benzyloxy)octanal oxime



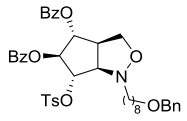
To a solution of the aldehyde (0.354 g, 1.5 mmol) and hydroxylamine hydrochloride (0.37 g, 5.3 mmol) in EtOH (6 mL) were added NaOH pallets (0.54 g, 13.5 mmol) in small portions. The mixture was allowed to stir at rt for 1 h and then refluxed for another 1 h. The reaction mixture was then cooled to rt and 1M HCl was added until pH ≈ 7 and the solution became clear. The mixture was extracted with CH_2Cl_2 (4x10 mL), washed with brine, dried over $MgSO_4$ and concentrated to afford the product (0.373 g, 1.5 mmol, quant.). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.26 (m, 10H), 4.50 (s, 4H), 3.46 (dd, *J* = 7.4, 5.8 Hz, 4H), 2.18 (dt, 2H), 1.66 – 1.55 (m, *J* = 10.4, 4.0 Hz, 5H), 1.51 – 1.39 (m, *J* = 7.4 Hz, 4H), 1.40 – 1.26 (m, *J* = 9.4, 8.9, 4.4 Hz, 13H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 128.5, 127.8, 127.6, 73.0, 70.5, 29.8, 29.6, 29.4, 29.3, 29.2, 26.6, 26.2, 25.0. IR (neat, cm^{-1}): υ 3250, 2927, 2854, 1454, 1361, 1095, 1028, 923, 904, 732, 696. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{15}H_{25}NO_2$ 250.18176, found 250.18009.

Compound 13



The oxime (0.317 g, 1.27 mmol) was dissolved in MeOH (4 mL) and cooled to 0 ^{0}C. $NaBH_3CN$ (0.168 g, 2.67 mmol) was added and 12 M HCl (2.1 mL, 2.54 mmol) was added drop wise. After addition the reaction mixture was allowed to stir at rt for 5 h before adding 1M NaOH until pH ≈ 9. The reaction mixture was concentrated and the product was extracted with CH_2Cl_2 (4 x 10ml), washed with brine, dried over $MgSO_4$, filtered, and concentrated. The crude product was recrystallized from pentane/EtOAc (4:1) to afford product **13** (0.12 g, 0.5 mmol, 40%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.44 – 7.27 (m, 5H), 4.50 (s, 2H), 3.46 (t, *J* = 6.6 Hz, 2H), 2.96 – 2.89 (m, 2H), 1.68 – 1.56 (m, 2H), 1.56 – 1.42 (m, 2H), 1.42 – 1.25 (m, 8H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 128.5, 127.8, 127.6, 73.0, 70.6, 54.2, 29.9, 29.6, 29.5, 27.2, 26.3. IR (neat, cm^{-1}): υ 3255, 2922, 2484, 1359, 1105, 1060, 740, 694. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{15}H_{25}NO_2$ 252.19741 found 252.19576.

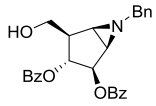
Compound 14



To a suspension of compound **10** (0.5 g, 0.75 mmol) in aqueous EtOH (10 mL) was added Zn dust (0.5 g, 8.05 mmol). The mixture was refluxed for 2 h and then filtered through Celite and concentrated to give a yellow oil that was dissolved in CH_2Cl_2 (20 mL), washed with water (2 × 10 mL), dried over $MgSO_4$, filtered and concentrated to give aldehyde **11** as an oil. To a solution of the aldehyde in toluene (8 mL) was added **13** (420 mg, 1.5 mmol) and Na_2CO_3 (150 mg, 1.5 mmol), and the reaction was heated at 50 ^{0}C for 2 h. The mixture was then concentrated, and the residue was dissolved in CH_2Cl_2 (15 mL), washed with water (2 × 10 mL), dried over $MgSO_4$, filtered and concentrated. The product was

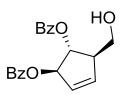
crystallized from methanol to yield compound **14** (220 mg, 0.18 mmol, 24%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.01 – 7.94 (m, 2H), 7.92 – 7.86 (m, 2H), 7.74 (d, *J* = 8.3 Hz, 2H), 7.58 – 7.50 (m, 2H), 7.44 – 7.36 (m, 4H), 7.34 (d, *J* = 4.4 Hz, 5H), 7.10 (d, *J* = 8.0 Hz, 2H), 5.88 (t, *J* = 8.2 Hz, 1H), 5.15 (dd, *J* = 8.0, 6.3 Hz, 1H), 5.01 (dd, *J* = 8.1, 5.7 Hz, 1H), 4.51 (s, 2H), 4.20 (dd, *J* = 9.4, 2.8 Hz, 1H), 4.07 (dd, *J* = 9.4, 7.3 Hz, 1H), 3.71 (dd, *J* = 9.6, 5.5 Hz, 1H), 3.47 (t, *J* = 6.6 Hz, 2H), 3.22 – 3.11 (m, 1H), 2.75 – 2.65 (m, 1H), 2.55 – 2.44 (m, 1H), 1.69 – 1.54 (m, 2H), 1.54 – 1.42 (m, 2H), 1.42 – 1.25 (m, 8H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 144.9, 133.6, 133.5, 130.1, 130.0, 129.8, 129.2, 128.7, 128.5, 128.4, 128.1, 127.8, 127.6, 79.2, 77.4, 73.0, 70.7, 50.3, 29.9, 29.6, 27.9, 27.2, 26.3. IR (neat, cm^{-1}): υ 2916, 2852, 1726, 1705, 1597, 1450, 1267, 1178, 1111, 970, 846, 813, 705, 665. $[\alpha]_D^{20}$ (c 0.2, CH_2Cl_2): -21. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{42}H_{47}NO_9S$ 742.30617, found 742.30479.

Compound 15



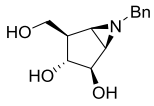
Compound **12** (2.0 g, 0.50 mmol) was dissolved in acetic acid (40 mL). Zinc powder (5 g, 20 mmol) was added and the mixture was stirred in a preheated oil bath at 35^{0}C. After 2 hours the mixture was filtered over Celite and rinsed with methanol. The solvents were evaporated and the solid residue was redissolved in methanol followed by addition of Et_3N until pH ≈ 7, and the reaction was stirred overnight. The mixture was then concentrated, and the residue was dissolved in CH_2Cl_2 (15 mL), washed with water (2 × 10 mL), dried over $MgSO_4$, filtered and concentrated. After recrystallization from methanol compound **15** (0.21 g, 0.49 mmol, 97%) was obtained. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.10 – 8.03 (m, 3H), 7.57 – 7.52 (m, 2H), 7.47 – 7.36 (m, 7H), 7.16 – 7.06 (m, 3H), 5.59 (dd, *J* = 6.1, 3.1 Hz, 1H), 5.45 (dd, *J* = 7.2, 6.2 Hz, 1H), 3.98 (dd, *J* = 11.1, 5.4 Hz, 1H), 3.93 – 3.84 (m, 1H), 3.60 (d, *J* = 13.7 Hz, 1H), 3.25 (d, *J* = 13.7 Hz, 1H), 2.66 (dd, *J* = 4.9, 3.1 Hz, 1H), 2.50 (dd, *J* = 4.9, 3.0 Hz, 1H), 2.48 – 2.39 (m, 1H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 166.7, 166.5, 138.6, 136.0, 133.5, 133.4, 133.3, 133.2, 130.00, 129.9, 129.8, 129.7, 129.6, 129.5, 128.6, 128.5, 128.4, 128.4, 128.3, 127.6, 127.2, 84.1, 81.8, 80.0, 76.8, 64.3, 62.5, 61.3, 55.1, 46.4, 43.1, 42.1. IR (neat, cm^{-1}): υ 3514, 2922, 1724, 1701, 1450, 1261, 1099, 1070, 1026, 705, 686. $[\alpha]_D^{20}$ (c 0.5, CH_2Cl_2): -87. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{27}H_{25}NO_5$ 444.18098, found 444.18027.

Compound 16



Byproduct obtained when the reaction of **12** with zinc in acetic acid was performed at 55 °C, yield 25%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.13 – 7.95 (m, 5H), 7.65 – 7.51 (m, 2H), 7.50 – 7.42 (m, 3H), 6.17 (s, 1H), 6.08 – 5.97 (m, 2H), 5.49 (t, *J* = 2.8 Hz, 1H), 3.90 (dd, *J* = 11.1, 4.4 Hz, 1H), 3.78 (dd, *J* = 11.1, 7.4 Hz, 1H), 3.06 – 2.96 (m, 1H). TLC-MS m/z: $[M+Na]^+$ calc for $C_{20}H_{18}O_5Na$ 361.1, found 361.1. These data are in accordance with literature reports.[18]

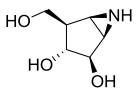
Compound 17



Compound **12** (80 mg, 0.13 mmol) was dissolved in acetic acid (10 mL). Zinc powder (1.2 g, 5 mmol) was added and the mixture was transferred to a preheated oil bath at 35^{0}C. After 2 hours the mixture was filtered over Celite and

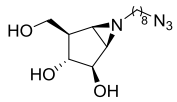
rinsed with methanol. The solvents were evaporated and co-evaporated with toluene. The solid residue was dissolved in methanol (10 mL) and methanolic sodium methoxide was added until pH ≈ 10 and stirred for 3 h. Subsequently, acetic acid was added until pH ≈ 7 and the solvent was evaporated followed by purification of the crude product by column chromatography (DCM/MeOH, 9:1) to afford compound **17** (30 mg, 0.13 mmol, quant.) as a white solid. ^{1}H-NMR (400 MHz, D_2O) δ 7.43 – 7.26 (m, 5H), 4.09 (dd, *J* = 6.5, 2.8 Hz, 1H), 3.71 (dd, *J* = 10.7, 4.9 Hz, 1H), 3.66 – 3.52 (m, 1H), 3.47 (d, *J* = 13.8 Hz, 1H), 3.40 – 3.30 (m, 1H), 3.25 (dd, *J* = 8.2, 6.5 Hz, 1H), 2.60 – 2.48 (m, 2H), 2.08 (dddt, *J* = 9.5, 7.7, 4.9, 2.5 Hz, 1H). ^{13}C-NMR (101 MHz, D_2O) δ 138.3, 128.6, 128.0, 127.5, 78.2, 75.6, 61.0, 60.5, 46.7, 45.0, 42.4. IR (neat, cm^{-1}): υ 2887, 1724, 1450, 1365, 1274, 1259, 1174, 1093, 995, 844, 812, 659. $[\alpha]_D^{20}$ (c 0.8, H_2O): +39. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{13}H_{17}NO_3$ 236.12940, found 236.12816.

Compound 18



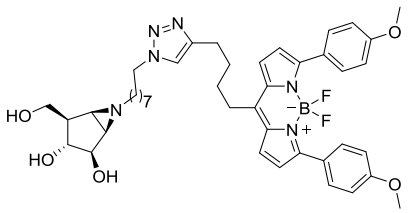
To a solution of **17** (290 mg, 1.23 mmol) in water (24 mL) under argon was added $Pd(OH)_2/C$ (10%wt, 280 mg, 2 mmol) and the mixture was purged with H_2 gas with a balloon. After vigorously stirring for 1 h, the mixture was filtered through Celite and concentrated followed by purification of the crude product by column chromatography (DCM/MeOH, 9:1), to afford product **18** (80 mg, 0.55 mmol, 45%) as a white solid. ^{1}H-NMR (400 MHz, D_2O) δ 4.04 (d, *J* = 6.1 Hz, 1H), 3.73 (dd, *J* = 10.9, 4.8 Hz, 1H), 3.67 – 3.56 (m, 1H), 3.16 (dd, *J* = 7.2 Hz, 1H), 2.66 (s, 2H), 2.08 – 1.96 (m, *J* = 4.6 Hz, 1H). ^{13}C-NMR (101 MHz, D_2O) δ 78.0, 74.7, 60.8, 48.8, 46.2, 44.1, 35.6, 32.6. IR (neat, cm^{-1}): υ 3387, 3197, 2947, 2509, 2385, 1382, 1091, 1020, 877, 678. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_6H_{11}NO_3$ 146.08199, found 146.08104.

Compound 19



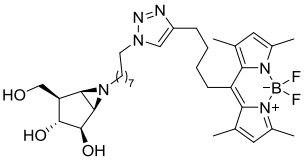
Aziridine **18** (82 mg, 0.57 mmol) was dissolved in dry DMF (2.5 mL). 1-azido-8-iodooctane (0.32 g, 1.13 mmol) and potassium carbonate (0.23 g, 1.7 mmol) were added and the mixture was heated to 80^{0}C and stirred overnight. The mixture was then concentrated and the product was purified by column chromatography on neutralized silica gel (DCM/MeOH, 10:1). After freeze drying product **19** was obtained (69 mg, 0.23 mmol, 40%) as a fluffy powder. ^{1}H-NMR (400 MHz, MeOD) δ 3.92 (dd, *J* = 6.3, 3.0 Hz, 1H), 3.79 (dd, *J* = 10.2, 4.4 Hz, 1H), 3.68 (t, *J* = 10.0 Hz, 1H), 3.30 (t, *J* = 6.8 Hz, 2H), 3.20 (dd, *J* = 8.0, 6.3 Hz, 1H), 2.32 (dt, *J* = 11.6, 7.5 Hz, 1H), 2.18 (dd, *J* = 5.2, 2.7 Hz, 1H), 2.14 (dd, *J* = 5.1, 3.0 Hz, 1H), 2.02 (dt, *J* = 11.7, 7.3 Hz, 1H), 1.98 – 1.90 (m, 1H), 1.65 – 1.51 (m, 4H), 1.42 – 1.33 (m, 8H). ^{13}C-NMR (101 MHz, MeOD) δ 80.2, 77.3, 62.6, 59.6, 52.4, 49.1, 46.2, 43.0, 30.5, 30.5, 30.2, 29.9, 28.2, 27.7. IR (neat, cm^{-1}): υ 3344, 2929, 2856, 2090, 1728, 1359, 1271, 1174, 1091, 1020, 812, 705, 659. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{14}H_{26}N_4O_3$ 299.20940, found 299.20779.

Compound 20 (EM234D)



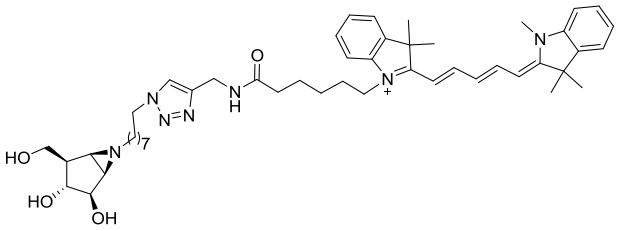
Compound **19** (5.1 mg, 17 µmol) was dissolved in DMF (0.6 mL). BODIPY-red-alkyne (16.2 mg, 34 mmol), $CuSO_4$ (33 µL of 100 mM solution in H_2O) and sodium ascorbate (50 µL of 100 mM solution in H_2O) were added and the solution was stirred overnight at room temperature. The solvent was evaporated and the product was purified by semi-preparative reversed phase HPLC (linear gradient: 49%→55%, solutions used: A: 50 mM NH_4HCO_3 in H_2O, B: acetonitrile). Freeze drying yielded product **20** as a dark blue powder (4.9 mg, 6.3 µmol, 37%). ^{1}H-NMR (600 MHz, MeOD) δ 7.84 (d, *J* = 8.9 Hz, 3H), 7.69 (s, 1H), 7.43 (d, *J* = 4.3 Hz, 2H), 6.97 (d, *J* = 8.9 Hz, 3H), 6.69 (d, *J* = 4.3 Hz, 2H), 4.32 (t, *J* = 7.0 Hz, 2H), 3.87 (dd, *J* = 6.2, 3.0 Hz, 1H), 3.85 (s, 5H), 3.75 (dd, *J* = 10.2, 4.4 Hz, 1H), 3.64 (t, *J* = 10.0 Hz, 1H), 3.16 (dd, *J* = 7.9, 6.3 Hz, 1H), 3.06 (t, *J* = 7.2 Hz, 2H), 2.78 (t, *J* = 6.8 Hz, 2H), 2.22 (dt, *J* = 11.6, 7.4 Hz, 1H), 2.10 (dd, *J* = 5.1, 2.8 Hz, 1H), 2.06 (dd, *J* = 5.1, 3.1 Hz, 1H), 1.96 – 1.77 (m, 9H), 1.54 – 1.41 (m, 2H), 1.33 – 1.15 (m, 9H). ^{13}C-NMR (151 MHz, MeOD) δ 162.2, 158.8, 148.6, 146.8, 137.5, 132.2, 132.1, 132.1, 128.4, 126.5, 123.3, 121.0, 114.6, 80.2, 77.4, 62.6, 59.6, 55.8, 51.2, 49.8, 49.6, 49.4, 49.3, 49.1, 49.0, 48.9, 48.7, 48.6, 46.2, 43.0, 34.1, 31.2, 31.1, 31.0, 30.5, 30.4, 30.3, 29.9, 28.1, 27.3, 25.8. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{33}H_{49}BF_2N_6O_3$ 783.42093, found 783.42211.

Compound 21(EM234B)



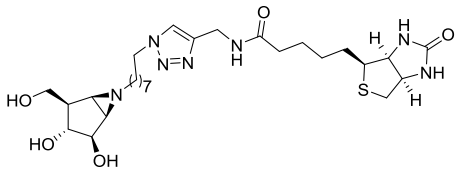
Compound **19** (5.1 mg, 17 µmol) was dissolved in DMF (0.6 mL). BODIPY-green-alkyne (11 mg, 34 mmol), $CuSO_4$ (33 µL of 100 mM solution in H_2O) and sodium ascorbate (50 µL of 100 mM solution in H_2O) were added and the solution was stirred overnight at room temperature. The solvent was evaporated and the product was purified by semi-preparative reversed phase HPLC (linear gradient: 42%→48%, solutions used: A: 50 mM NH_4HCO_3 in H_2O, B: acetonitrile). Freeze drying yielded compound **21** as a yellow powder (6.0 mg, 9.6 µmol, 57%). ^{1}H-NMR (600 MHz, MeOD) δ 7.74 (s, 1H), 6.12 (s, 2H), 4.35 (t, *J* = 7.0 Hz, 2H), 3.88 (dd, *J* = 6.3, 3.0 Hz, 1H), 3.76 (dd, *J* = 10.2, 4.4 Hz, 1H), 3.65 (t, *J* = 10.0 Hz, 1H), 3.17 (dd, *J* = 8.0, 6.3 Hz, 1H), 3.06 – 2.97 (m, 2H), 2.79 (t, *J* = 7.2 Hz, 2H), 2.44 (s, 7H), 2.39 (s, 7H), 2.25 (dt, *J* = 11.6, 7.4 Hz, 1H), 2.13 (dd, *J* = 5.1, 2.8 Hz, 1H), 2.08 (dd, *J* = 5.1, 3.1 Hz, 1H), 1.98 – 1.83 (m, 7H), 1.69 – 1.61 (m, 2H), 1.56 – 1.47 (m, 2H), 1.36 – 1.22 (m, 11H). ^{13}C-NMR (151 MHz, MeOD) δ 154.9, 148.5, 147.9, 142.2, 132.6, 131.0, 123.4, 122.6, 80.2, 77.4, 62.7, 59.7, 51.2, 49.6, 46.2, 43.0, 32.3, 31.2, 31.1, 30.8, 30.5, 30.4, 29.9, 29.1, 28.1, 27.3, 25.9, 16.5 14.4. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{33}H_{49}BF_2N_6O_3$ 627.40062, found 627.40088.

Compound 22 (EM234A)



Compound **19** (5.1 mg, 17 µmol) was dissolved in DMF (0.6 mL). Cy5-alkyne (18.6 mg, 34 mmol), $CuSO_4$ (33 µL of 100 mM solution in H_2O) and sodium ascorbate (50 µL of 100 mM solution in H_2O) were added and the solution was stirred overnight at room temperature. The solvent was evaporated and the product was purified by semi-preparative reversed phase HPLC (linear gradient: 45%→51%, solutions used: A: 50 mM NH_4HCO_3 in H_2O, B: acetonitrile). Freezedrying yielded product **22** as a blue powder (3.2 mg, 3.7 µmol, 22%). ^{1}H-NMR (600 MHz, MeOD) δ 8.30 – 8.17 (m, 3H), 7.84 (s, 1H), 7.50 (d, *J* = 7.3 Hz, 2H), 7.46 – 7.37 (m, 2H), 7.33 – 7.21 (m, 5H), 6.62 (t, *J* = 12.4 Hz, 1H), 6.28 (dd, *J* = 13.7, 2.2 Hz, 3H), 4.40 (s, 2H), 4.36 (t, *J* = 7.1 Hz, 2H), 4.09 (t, *J* = 7.5 Hz, 2H), 3.88 (dd, *J* = 6.3, 3.0 Hz, 1H), 3.76 (dd, *J* = 10.2, 4.4 Hz, 1H), 3.65 (d, *J* = 9.9 Hz, 1H), 3.63 (s, 3H), 3.16 (dd, *J* = 8.0, 6.3 Hz, 1H), 2.29 – 2.21 (m, 3H), 2.13 (dd, *J* = 5.1, 2.8 Hz, 1H), 2.09 (dd, *J* = 5.1, 3.1 Hz, 1H), 1.96 (dt, *J* = 11.7, 7.1 Hz, 1H), 1.93 – 1.85 (m, 8H), 1.82 (dt, *J* = 15.3, 7.9 Hz, 2H), 1.73 (s, 12H), 1.56 – 1.43 (m, 4H), 1.38 – 1.25 (m, 8H). ^{13}C-NMR (151 MHz, MeOD) δ 175.8, 175.4, 174.6, 155.6, 155.5, 146.1, 144.2, 143.5, 142.6, 142.5, 129.8, 129.8, 126.6, 126.3, 126.2, 124.2, 123.4, 123.3, 112.0, 111.9, 104.4, 104.2, 80.2, 77.4, 62.7, 59.7, 56.1, 51.3, 50.5, 50.5, 49.9, 49.6, 46.2, 44.7, 43.0, 36.5, 35.6, 31.5, 31.3, 31.1, 30.5, 30.4, 29.9, 28.2, 28.1, 27.9, 27.8, 27.4, 27.3, 26.4. HRMS (ESI) m/z: $[M]^+$ calc for $C_{49}H_{68}N_7O_4$ 818.5327, found 818.55349.

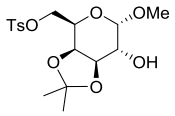
Compound 23 (EM204)



Compound **19** (6.2 mg, 21 µmol) was dissolved in DMF (0.9 mL). Biotin-alkyne (6.5 mg, 23 mmol), $CuSO_4$ (17 µL of 100 mM solution in H_2O) and sodium ascorbate (18 µL of 100 mM solution in H_2O) were added and the solution was stirred overnight at room temperature. The solvent was evaporated and the product was purified by semi-preparative reversed phase HPLC (linear gradient: 19%→25%, solutions used: A: 50 mM NH_4HCO_3 in H_2O, B: acetonitrile). Freeze drying yielded compound **23** as a white powder (4.8 mg, 3.7 µmol, 40%). ^{1}H-NMR (600 MHz, MeOD) δ 7.84 (s, 1H), 4.50 (dd, *J* = 7.8, 4.6 Hz, 1H), 4.42 (s, 2H), 4.37 (t, *J* = 7.1 Hz, 2H), 4.29 (dd, *J* = 7.9, 4.5 Hz, 1H), 3.89 (dd, *J* = 6.3, 3.0 Hz, 1H), 3.76 (dd, *J* = 10.2, 4.4 Hz, 1H), 3.65 (t, *J* = 10.0 Hz, 1H), 3.24 – 3.13 (m, *J* = 10.3, 8.9, 6.0 Hz, 2H), 2.93 (dd, *J* = 12.8, 5.0 Hz, 1H), 2.71 (d, *J* = 12.7 Hz, 1H), 2.28 (dd, *J* = 7.9, 4.3 Hz, 1H), 2.24 (t, *J* = 7.6 Hz, 2H), 2.15 (dd, *J* = 5.1, 2.8 Hz, 1H), 2.11 (dd, *J* = 5.1, 3.1 Hz, 1H), 2.02 – 1.94 (m, *J* = 7.5, 4.6 Hz, 1H), 1.94 – 1.83 (m, 4H), 1.79 – 1.48 (m, 6H), 1.43 (dd, *J* = 15.4, 7.6 Hz, 2H), 1.39 – 1.24 (m, 8H). ^{13}C-NMR (151 MHz, MeOD) δ 176.0, 166.1, 146.2, 124.2, 80.2, 77.4, 63.3, 62.6, 61.6, 59.6, 57.0, 51.4,

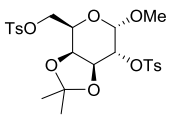
49.8, 49.6, 46.2, 43.0, 41.1, 36.5, 35.6, 31.3, 30.5, 30.4, 29.9, 29.7, 29.4, 28.2, 27.4, 26.7. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{27}H_{45}N_7O_5S$ 580.32826, found 580.32788.

Compound 25



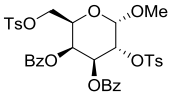
To a stirred solution of dry acetonitrile (200 mL) was added methyl-α-D-galactopyranoside (10.1 g, 51 mmol), triethylamine (12.5 mL, 65 mmol) and tosylchloride (13.7 g, 65 mmol) and stirred overnight. The white precipitate was filtered and washed with pentane. The solid was dissolved in toluene, evaporated and dissolved in dry DMF (200 mL). *p*-Toluenesulfonic acid (0.78 g, 0.15 mmol) and 2,2-dimethoxypropane (102 mmol) were added and the mixture was stirred over night at rt. Then it was quenched with triethylamine until pH ≈ 7. The solvent was evaporated followed by purification of the crude product by column chromatography (DCM/MeOH, 20:1) to afford product **25** (12.4 g, 32 mmol, 63% over 2 steps). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.81 (d, *J* = 8.3 Hz, 2H), 7.35 (d, *J* = 8.0 Hz, 2H), 4.70 (d, *J* = 3.9 Hz, 1H), 4.29 – 4.17 (m, 4H), 4.14 (d, *J* = 6.2 Hz, 1H), 3.79 (td, *J* = 6.0, 3.9 Hz, 1H), 3.42 (s, 3H), 2.48 (d, *J* = 5.8 Hz, 1H), 2.46 (s, 3H), 1.42 (s, 3H), 1.28 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 145.0, 129.9, 128.1, 110.1, 98.0, 75.7, 72.6, 69.3, 68.8, 66.6, 55.6, 27.5, 25.8, 21.8. IR (neat, cm^{-1}): υ 3221, 2941, 1720, 1336, 1253, 1172, 1095, 806, 702. $[\alpha]_D^{20}$ (c 0.8, H_2O): +56. HRMS (ESI) m/z: $[M+H]^+$ calc $C_{17}H_{30}O_8S$ 389.12989, found 389.12976.

Compound 26



Compound **25** (12 g, 32 mmol) was dissolved in dry acetonitrile (150 mL), then triethylamine (15 mL, 80 mmol) and tosylchloride (18 g, 80 mmol) were added and the mixture was stirred overnight at 40 °C. The reaction was quenched with water and extracted with EtOAc (3x). The combined extracts were washed successively with water (3x), aqueous $NaHCO_3$ (3x), brine and dried over $MgSO_4$. The solution was concentrated followed by purification of the crude product by column chromatography (pentane/EtOAc, 3:1) affording product **26** (8.0 g, 16 mmol, 50%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.80 (dd, *J* = 8.3, 5.4 Hz, 4H), 7.34 (t, *J* = 8.6 Hz, 4H), 4.81 (d, *J* = 3.1 Hz, 1H), 4.27 – 4.14 (m, 5H), 4.11 (dd, *J* = 4.9, 2.1 Hz, 1H), 3.32 (s, 3H), 2.46 (s, 3H), 2.43 (s, 3H), 1.18 (s, 3H), 1.05 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 145.1, 145.1, 133.3, 133.0, 130.4, 130.0, 129.8, 128.4, 128.1, 110.1, 97.7, 78.4, 73.4, 73.2, 68.7, 65.5, 56.1, 27.5, 26.3, 21.8, 21.8. IR (neat, cm^{-1}): υ 1724, 1452, 1352, 1172, 1091, 974, 810, 704, 677. $[\alpha]_D^{20}$ (c 0.2, CH_2Cl_2): +60. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{21}H_{26}O_{10}S_2$ 543.13703, found 543.13521.

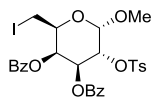
Compound 27



Compound **26** (2.8 g, 5.1 mmol) was dissolved in methanol (80 mL) and HCl (12 M, 0.2 mL) was added. The mixture was stirred overnight followed by addition of Na_2CO_3 until pH ≈ 7 and filtered. The solvent was evaporated and purification of the crude product by column chromatography (DCM/MeOH, 50:1) yielded the 3,4-diol (2.6 g, 5.1

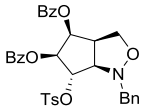
mmol, quant.). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.80 (dd, *J* = 14.0, 8.3 Hz, 4H), 7.35 (dd, 4H), 4.61 (d, *J* = 3.7 Hz, 1H), 4.55 (dd, *J* = 9.2, 3.7 Hz, 1H), 4.22 (dd, *J* = 10.6, 5.4 Hz, 1H), 4.14 (dd, *J* = 10.6, 7.0 Hz, 1H), 4.07 – 3.92 (m, 3H), 3.26 (s, 3H), 2.76 (d, *J* = 3.5 Hz, 1H), 2.58 (s, 1H), 2.45 (d, *J* = 3.1 Hz, 6H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 145.6, 145.3, 132.7, 130.1, 130.1, 128.1, 128.1, 97.4, 77.8, 77.5, 77.2, 76.8, 69.1, 68.5, 67.6, 67.3, 55.8. IR (neat, cm^{-1}): υ 3537, 3464, 2920, 1597, 1352, 1190, 1170, 1002, 974, 837, 812, 769, 678. $[\alpha]_D^{20}$ (c 1.0, CH_2Cl_2): +80. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{21}H_{26}O_{10}S_2$ 503.10538, found 503.10428. Subsequently, the 3,4-diol (2.55 g, 5.1 mmol) was dissolved in dry pyridine (50 mL). Benzoyl anhydride (4.6 g, 20.4 mmol) and DMAP (30 mg, 0.3 mmol) were added and the mixture was stirred overnight. The reaction was quenched by addition of aqueous $NaHCO_3$ and extracted with EtOAc (3 x 100 mL). The combined extracts were washed with water, aqueous $NaHCO_3$, and brine, dried over $MgSO_4$, and concentrated. Purification of the crude product by column chromatography (pentane/EtOAc, 2:1) yielded compound **27** (2.56 g, 3.6 mmol, 70%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.89 – 7.83 (m, 2H), 7.72 – 7.63 (m, 3H), 7.60 (d, *J* = 8.3 Hz, 2H), 7.52 – 7.44 (m, 5H), 7.25 – 7.18 (m, 4H), 6.96 (d, *J* = 8.1 Hz, 2H), 5.70 (d, *J* = 2.6 Hz, 1H), 5.63 (dd, *J* = 10.5, 3.4 Hz, 1H), 5.09 (d, *J* = 3.5 Hz, 1H), 4.93 (dd, *J* = 10.4, 3.5 Hz, 1H), 4.38 (t, *J* = 6.3 Hz, 1H), 4.17 (dd, *J* = 10.3, 7.1 Hz, 1H), 4.01 (dd, *J* = 10.3, 5.6 Hz, 1H), 3.44 (s, 3H), 2.32 (s, 3H), 2.22 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 165.1, 164.9, 145.2, 144.9, 133.8, 133.3, 130.0, 129.9, 129.9, 129.8, 129.0, 128.9, 128.7, 128.2, 128.0, 127.8, 98.2, 74.3, 69.1, 67.5, 67.1, 66.7, 56.3, 21.7. IR (neat, cm^{-1}): υ 1720, 1357, 1271, 1064, 1020, 977, 813, 705, 659. $[\alpha]_D^{20}$ (c 0.2, CH_2Cl_2): +59. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{35}H_{34}O_{12}S_2$ 711.15765, found 711.15689.

Compound 28



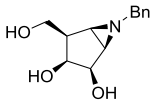
Sodium iodide (612 mg, 4.1 mmol) was added to a mixture of compound **27** (2.44 g, 3.4 mmol) in acetic anhydride (15 mL) and refluxed for 4 hours. After cooling to rt the mixture was filtered over Celite to remove the solids. The residue was dissolved in DCM, washed with aqueous sodium thiosulphate, water, dried over $MgSO_4$ and filtered. The solvent was evaporated followed by crystallization from methanol to yield product **28** (2.0 g, 88%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.06 – 7.88 (m, 2H), 7.66 (dd, *J* = 14.8, 7.9 Hz, 3H), 7.56 – 7.42 (m, 5H), 7.23 (dd, *J* = 8.1, 7.6 Hz, 2H), 6.99 (d, *J* = 8.0 Hz, 2H), 5.86 (dd, 1H), 5.69 (dd, *J* = 10.5, 3.4 Hz, 1H), 5.13 (d, *J* = 3.6 Hz, 1H), 5.00 (dd, *J* = 10.5, 3.6 Hz, 1H), 4.32 (dt, *J* = 7.0 Hz, 1H), 3.54 (s, 3H), 3.20 (d, *J* = 7.1 Hz, 2H), 2.24 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 165.4, 165.0, 144.9, 133.8, 133.3, 130.0, 129.9, 129.8, 129.0, 129.0, 128.8, 128.2, 127.8, 98.3, 74.4, 70.6, 69.9, 67.7, 56.4, 21.7, 0.5. IR (neat, cm^{-1}): υ 1720, 1371, 1273, 1247, 1093, 1020, 842, 817, 705, 655. $[\alpha]_D^{20}$ (c 0.3, CH_2Cl_2): +202. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{28}H_{27}IO_9S$ 689.03183, found 689.03131.

Compound 29



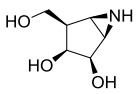
To a suspension of compound **28** (2.3 g, 6.8 mmol) in EtOH (60 mL) was added Zn dust (4 g, 60 mmol). The mixture was refluxed for 6 h and then filtered through Celite and concentrated to give a yellow oil that was dissolved in CH_2Cl_2 (200 mL), washed with water (2 × 100 mL), dried over $MgSO_4$, filtered and concentrated to give the aldehyde as an oil. To a solution of the aldehyde in toluene (30 mL) was added *N*-benzylhydroxylamine hydrochloride (2.4 g, 14 mmol) and Na_2CO_3 (1.6 g , 14 mmol), and the reaction was stirred at 50 ^{0}C for 16 h. The mixture was then concentrated, and the residue was dissolved in CH_2Cl_2 (150 mL), washed with water (2 × 100 mL), dried over $MgSO_4$, filtered and concentrated. Purification of the crude product by column chromatography (pentane/EtOAc, 3:1) yielded product **29** (4.0 g, 6.5 mmol, 95%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.92 (dd, *J* = 8.3, 1.2 Hz, 2H), 7.80 (dd, *J* = 8.3, 1.2 Hz, 2H), 7.74 (d, *J* = 8.3 Hz, 2H), 7.59 – 7.44 (m, 3H), 7.43 – 7.27 (m, 10H), 7.11 (d, *J* = 8.0 Hz, 2H), 5.71 (dd, *J* = 6.7, 4.9 Hz, 1H), 5.50 (dd, *J* = 7.7, 4.8 Hz, 1H), 5.26 (dd, *J* = 7.7, 4.9 Hz, 1H), 4.07 – 3.95 (m, 3H), 3.89 – 3.79 (m, 2H), 3.68 – 3.58 (m, 1H), 2.24 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 165.5, 165.1, 145.0, 136.8, 133.7, 133.6, 133.4, 130.0, 129.9, 129.1, 129.0, 128.6, 128.6, 128.4, 127.9, 127.7, 85.2, 75.4, 71.5, 70.3, 64.9, 59.9, 47.0, 21.7. IR (neat, cm^{-1}): υ 1722, 1598, 1450, 1363, 1274, 1174, 1024, 844, 812, 707, 696. $[\alpha]_D^{20}$ (c 0.3, CH_2Cl_2): +17. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{34}H_{31}NO_8S$ 614.18431, found 614.18427.

Compound 30



Compound **29** (480 mg, 0.8 mmol) was dissolved in acetic acid (15 mL). Zinc powder (1.0 g, 6 mmol) was added and the mixture was stirred in a preheated oil bath at 40 ^{0}C. After 7 h the mixture was filtered over Celite and rinsed with methanol. The solvents were evaporated and the crude was co-evaporated with toluene. The solid residue was dissolved in methanol (10 mL) and methanolic sodium methoxide was added until pH ≈ 10 and stirred for 3 h. Acetic acid was added until pH ≈ 7 and the solvent was evaporated, followed by purification of the crude product by column chromatography (DCM/MeOH, 9:1), to afford product **30** (188 mg, 0.8 mmol, quant.). ^{1}H-NMR (400 MHz, MeOD) δ 7.42 – 7.28 (m, 5H), 4.12 (dd, *J* = 5.2, 2.2 Hz, 1H), 3.84 – 3.72 (m, 2H), 3.65 – 3.55 (m, 2H), 3.17 (d, *J* = 13.5 Hz, 1H), 2.46 (s, 1H), 2.41 (s, 1H), 2.23 – 2.13 (m, 1H). ^{13}C-NMR (400 MHz, MeOD) δ 140.3, 129.4, 128.9, 128.2, 75.0, 71.3, 61.9, 60.2, 47.5, 46.7, 44.0. IR (neat, cm^{-1}): υ 2885, 1724, 1598, 1450, 1274, 1259, 1174, 1093, 995, 908 844, 812, 659. $[\alpha]_D^{20}$ (c 0.6, MeOH): +24. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{13}H_{17}NO_3$ 236.12940, found 236.12809.

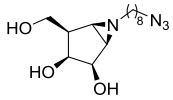
Compound 31



To a solution of **30** (500 mg, 2.13 mmol) in water (50 mL) under argon was added $Pd(OH)_2/C$ (10%wt, 500 mg) and the mixture was purged with H_2 gas with a balloon. After vigorously stirring for 30 minutes, the mixture was filtered through Celite and concentrated followed by purification of the crude product by column chromatography (DCM/MeOH, 9:1), to afford product **31** (144 mg, 0.99 mmol, 46%). ^{1}H-NMR (400 MHz, MeOD) δ 4.24 (dd, *J* = 5.5, 2.2 Hz, 1H), 3.91 – 3.84 (m, 2H), 3.75 (dd, *J* = 10.8, 7.6 Hz, 1H), 2.59 – 2.55 (m, 1H), 2.54 (dd, *J* = 3.6, 2.0

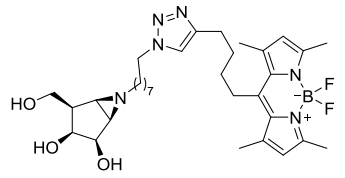
Hz, 1H), 2.33 (qd, *J* = 7.4, 1.9 Hz, 1H). ^{13}C-NMR (101 MHz, MeOD) δ 75.2, 69.7, 60.4, 46.8, 39.0, 35.8. IR (neat, cm^{-1}): υ 3259, 2922, 1730, 1408, 1375, 1101, 1085, 1051, 1031, 867, 682. $[\alpha]_D^{20}$ (c 0.8, MeOH): +8. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_6H_{11}NO_3$ 146.0812, found 146.0815.

Compound 32



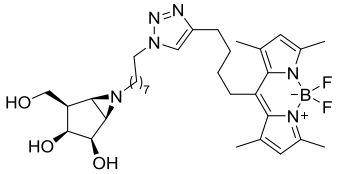
Aziridine **31** (39 mg, 0.27 mmol) was dissolved in dry DMF (1.3 mL). 1-azido-8-iodooctane (0.15 g, 0.53 mmol) and K_2CO_3 (0.11 g, 0.8 mmol) were added and the mixture was heated to 80 ^{0}C and stirred overnight. The mixture was then concentrated and purified by column chromatography on neutralized silica gel (DCM/MeOH, 10:1) and by semi-preparative reversed phase HPLC (linear gradient: 36%→42%, solutions used: A: 50 mM NH_4HCO_3 in H_2O, B: acetonitrile). After freeze drying compound **32** was obtained (37 mg, 0.12 mmol, 46%).^{1}H-NMR (600 MHz, MeOD) δ 4.04 (dd, *J* = 5.2, 2.3 Hz, 1H), 3.87 (dd, *J* = 10.7, 6.9 Hz, 1H), 3.73 (dd, *J* = 10.7, 7.8 Hz, 1H), 3.69 (t, *J* = 5.1 Hz, 1H), 3.28 (t, *J* = 6.9 Hz, 2H), 2.37 (dt, *J* = 11.8, 7.2 Hz, 1H), 2.23 – 2.20 (m, 1H), 2.20 – 2.17 (m, 1H), 2.16 – 2.11 (m, 1H), 1.98 (dt, *J* = 11.8, 6.8 Hz, 1H), 1.63 – 1.56 (m, 2H), 1.56 – 1.49 (m, 2H), 1.44 – 1.32 (m, 9H). ^{13}C-NMR (151 MHz, MeOD) δ 74.9, 71.4, 60.4, 58.4, 52.5, 47.0, 46.8, 43.5, 30.8, 30.5, 30.2, 29.9, 28.3, 27.8. IR (neat, cm^{-1}): υ 3323, 2927, 2854, 2090, 1463, 1253, 1093, 1024, 906, 723, 623. $[\alpha]_D^{20}$ (c 1.0, MeOH): +12. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{14}H_{26}N_4O_3$ 321.1897, found 321.1900.

Compound 33 (EM265B)



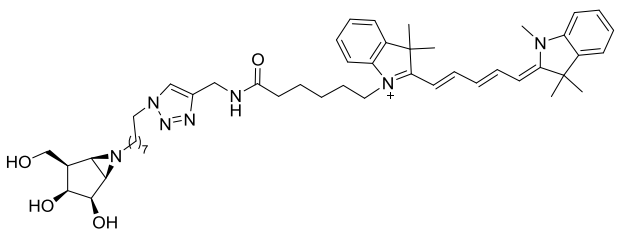
Compound **32** (5.0 mg, 17 μmol) was dissolved in DMF (0.6 mL). BODIPY-green-alkyne (11 mg, 34 mmol), $CuSO_4$ (33 μL of 100 mM solution in H_2O) and sodium ascorbate (50 μL of 100 mM solution in H_2O) were added and the solution was stirred overnight at room temperature. The solvent was evaporated and the product was purified by by semi-preparative reversed phase HPLC (linear gradient: 50%→53%, solutions used: A: 50 mM NH_4HCO_3 in H_2O, B: acetonitrile). After freezedrying product **33** was obtained as a yellow powder (3.8 mg, 6.0 μmol, 35%). ^{1}H-NMR (500 MHz, MeOD) δ 7.74 (s, 1H), 6.12 (s, 2H), 4.35 (t, *J* = 6.9 Hz, 2H), 4.03 (dd, *J* = 5.1, 2.1 Hz, 1H), 3.85 (dd, *J* = 10.7, 6.9 Hz, 1H), 3.74 – 3.66 (m, 2H), 3.06 – 2.98 (m, 2H), 2.78 (t, *J* = 7.2 Hz, 2H), 2.43 (s, 6H), 2.38 (s, 6H), 2.35 – 2.28 (m, 1H), 2.22 – 2.17 (m, 1H), 2.17 – 2.14 (m, 1H), 2.14 – 2.10 (m, 1H), 1.89 – 1.82 (m, 3H), 1.69 – 1.59 (m, 2H), 1.52 – 1.43 (m, 2H), 1.40 – 1.19 (m, 9H). ^{13}C-NMR (125 MHz, MeOD) δ 154.9, 142.2, 122.6, 74.9, 71.4, 60.4, 58.3, 51.2, 47.0, 46.8, 43.5, 43.5, 40.4, 32.2, 31.2, 30.8, 30.7, 30.3, 29.9, 29.1, 28.1, 27.3, 25.9, 16.5, 14.4. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{33}H_{49}BF_2N_6O_3$ 627.4000, found 627.4024.

Compound 34 (EM265C)



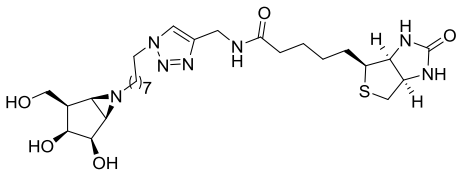
Compound **32** (5.0 mg, 17 μmol) was dissolved in DMF (0.6 mL). BODIPY-red-alkyne (16.2 mg, 34 mmol), $CuSO_4$ (33 μL of 100 mM solution in H_2O) and sodium ascorbate (50 μL of 100 mM solution in H_2O) were added and the solution was stirred overnight at room temperature. The solvent was evaporated and the product was purified by semi-preparative reversed phase HPLC (linear gradient: 55%→61%, solutions used: A: 50 mM NH_4HCO_3 in H_2O, B: acetonitrile). After freezedrying product **34** was obtained as a dark blue powder (3.8 mg, 4.9 μmol, 29%). ^{1}H-NMR (500 MHz, MeOD) δ 7.84 (d, *J* = 8.9 Hz, 4H), 7.69 (s, 1H), 7.43 (d, *J* = 4.3 Hz, 2H), 6.97 (d, *J* = 8.9 Hz, 4H), 6.69 (d, *J* = 4.3 Hz, 2H), 4.33 (t, *J* = 7.0 Hz, 2H), 4.01 (dd, *J* = 5.1, 2.2 Hz, 1H), 3.84 (s, 7H), 3.72 – 3.64 (m, 2H), 3.05 (t, *J* = 7.2 Hz, 2H), 2.78 (t, *J* = 6.7 Hz, 2H), 2.28 (dt, *J* = 11.8, 7.0 Hz, 1H), 2.15 (s, 1H), 2.14 – 2.07 (m, 2H), 1.91 – 1.78 (m, 8H), 1.49 – 1.37 (m, 2H), 1.37 – 1.17 (m, 9H). ^{13}C-NMR (125 MHz, MeOD) δ 132.1, 128.4, 123.3, 121.0, 114.6, 74.9, 71.4, 60.3, 58.3, 55.8, 51.2, 46.9, 46.8, 43.5, 34.1, 30.3, 27.3. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{43}H_{53}BF_2N_6O_5$ 783.4211, found 783.4241.

Compound 35(EM267)



Compound **32** (5.0 mg, 17 μmol) was dissolved in DMF (0.6 mL). Cy5-alkyne (18.6 mg, 34 mmol), $CuSO_4$ (33 μL of 100 mM solution in H_2O) and sodium ascorbate (50 μL of 100 mM solution in H_2O) were added and the solution was stirred overnight at room temperature. The solvent was evaporated and the product was purified by semi-preparative reversed phase HPLC (linear gradient: 55%→61%, solutions used: A: 50 mM NH_4HCO_3 in H_2O, B: acetonitrile). After freezedrying product **35** was obtained as a blue powder (3.5 mg, 4.1 μmol, 24%). ^{1}H-NMR (500 MHz, MeOD) δ 8.27 (t, *J* = 13.0 Hz, 2H), 7.87 (s, 1H), 7.52 (d, *J* = 7.4 Hz, 3H), 7.48 – 7.39 (m, 3H), 7.37 – 7.24 (m, 5H), 6.64 (t, *J* = 12.4 Hz, 1H), 6.30 (d, *J* = 13.5 Hz, 2H), 4.43 (s, 2H), 4.39 (t, *J* = 7.0 Hz, 3H), 4.11 (t, *J* = 7.4 Hz, 2H), 4.07 – 4.03 (m, 1H), 3.87 (dd, *J* = 10.6, 6.9 Hz, 1H), 3.76 – 3.68 (m, 2H), 3.65 (s, 3H), 2.27 (t, *J* = 7.3 Hz, 2H), 2.23 (s, 1H), 2.21 – 2.18 (m, 1H), 2.18 – 2.12 (m, 1H), 1.93 – 1.80 (m, 7H), 1.55 – 1.44 (m, 5H), 1.44 – 1.27 (m, 13H). ^{13}C-NMR (125 MHz, MeOD) δ 174.6, 155.5, 129.8, 126.6, 126.2, 124.1, 123.4, 123.3, 112.0, 111.8, 104.4, 104.2, 74.9, 71.4, 60.4, 58.3, 51.3, 50.5, 47.0, 46.8, 44.8, 43.5, 36.5, 35.6, 31.5, 31.3, 30.7, 30.3, 29.9, 28.1, 27.9, 27.8, 27.3, 26.4. HRMS (ESI) m/z: $[M]^+$ calc for $C_{49}H_{68}N_7O_4$ 818.5327, found 818.5339.

Compound 36 (EM265D)



Compound **32** (5.0 mg, 17 µmol) was dissolved in DMF (0.6 mL). Biotin-alkyne (9.3 mg, 34 mmol), $CuSO_4$ (33 µL of 100 mM solution in H_2O) and sodium ascorbate (50 µL of 100 mM solution in H_2O) were added and the solution was stirred overnight at room temperature. The solvent was evaporated and the product was purified by by semi-preparative reversed phase HPLC (linear gradient: 20%→26%, solutions used: A: 50 mM NH_4HCO_3 in H_2O, B: acetonitrile). After freezedrying product **36** was obtained as a white powder (2.9 mg, 5.0 µmol, 29%). ^{1}H-NMR (500 MHz, MeOD) δ 7.85 (s, 2H), 4.49 (dd, *J* = 7.8, 4.9 Hz, 2H), 4.42 (s, 4H), 4.38 (t, *J* = 7.0 Hz, 4H), 4.29 (dd, *J* = 7.9, 4.5 Hz, 2H), 4.04 (dd, *J* = 5.1, 2.2 Hz, 1H), 3.86 (dd, *J* = 10.7, 6.9 Hz, 1H), 3.74 – 3.67 (m, 2H), 3.53 (t, *J* = 6.6 Hz, 1H), 3.19 (dt, *J* = 10.0, 5.6 Hz, 2H), 2.93 (dd, *J* = 12.7, 5.0 Hz, 2H), 2.70 (d, *J* = 12.7 Hz, 2H), 2.36 (dt, *J* = 11.8, 7.1 Hz, 1H), 2.26 – 2.20 (m, 5H), 2.20 – 2.17 (m, 1H), 2.16 – 2.10 (m, 1H), 1.97 (dt, *J* = 12.1, 6.7 Hz, 1H), 1.78 – 1.55 (m, 8H), 1.55 – 1.46 (m, 4H), 1.46 – 1.27 (m, 19H). ^{13}C-NMR (125 MHz, MeOD) δ 74.9, 71.4, 63.3, 62.9, 61.6, 60.4, 58.3, 57.0, 51.4, 47.0, 46.8, 43.5, 41.1, 36.5, 35.6, 31.3, 30.8, 30.4, 30.0, 29.7, 29.5, 28.2, 27.4, 26.7. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{27}H_{45}N_7O_5S$ 580.3276, found 580.3289.

General procedure for SDS-PAGE experiments

Mouse liver lysate (Jackson's laboratories, C57Bl6/J, 40 µg total protein per sample) or human fibroblast lysate (product code CC-2511, lot no. C92030, Lonza, 10 µg total protein per sample) was diluted in 150 mM phosphate buffer with appropriate pH. A solution of the ABP with appropriate concentration was added and the mixture was incubated for 30 minutes at 37 °C. In the case of competition experiments, the enzyme solution was pre-incubated with appropriate inhibitor concentrations for 30 minutes at 37 °C. Then, laemmli (4X) was added and the mixture was heated at 100 °C for 5 minutes. The samples were loaded on 10% acrylamide SDS-PAGE gels and the gels were ran at a constant 90V. Wet slab gels were scanned on fluorescence using a Typhoon FLA9500 Imager (GE Healthcare) using λ_{EX} 635 nm; λ_{EM} > 665 nm. Images were acquired, processed and quantified with Image Quant (GE Healthcare).

References

1 B. F. Cravatt, A. T. Wright and J. W. Kozarich, *Annu. Rev. Biochem.*, 2008, **77**, 383–414.

2 W. W. Kallemeijn, K. Y. Li, M. D. Witte, A. R. A. Marques, J. Aten, S. Scheij, J. Jiang, L. I. Willems, T. M. Voorn-Brouwer, C. P. A. A. van Roomen, R. Ottenhoff, R. G. Boot, H. van den Elst, M. T. C. Walvoort, B. I. Florea, J. D. C. Codée, G. A. van der Marel, J. M. F. G. Aerts and H. S. Overkleeft, *Angew. Chem. Int. Ed.*, 2012, **51**, 12529–12533.

3 L. I. Willems, T. J. M. Beenakker, B. Murray, S. Scheij, W. W. Kallemeijn, R. G. Boot, M. Verhoek, W. E. Donker-Koopman, M. J. Ferraz, E. R. Van Rijssel, B. I. Florea, J. D. C. Codée, G. A. van der Marel, J. M. F. G. Aerts and H. S. Overkleeft, *J. Am. Chem. Soc.*, 2014, **136**, 11622–11625.

4 J. Jiang, W. W. Kallemeijn, D. W. Wright, A. M. C. H. van den Nieuwendijk, V. C. Rohde, E. C. Folch, H.

van den Elst, B. I. Florea, S. Scheij, W. E. Donker-Koopman, M. Verhoek, N. Li, M. Schürmann, D. Mink, R. G. Boot, J. D. C. Codée, G. A. van der Marel, G. J. Davies, J. M. F. G. Aerts and H. S. Overkleeft, *Chem. Sci.*, 2015, **6**, 2782–2789.

5 D. Leung, C. Hardouin, D. L. Boger and B. F. Cravatt, *Nat. Biotechnol.*, 2003, **21**, 687–691.

6 D. G. Naumoff, *Mol. Biol.*, 2012, **46**, 322–327.

7 O. Lopez Lopez, J. G. Fernández-Bolaños, V. H. Lillelund and M. Bols, *Org. Biomol. Chem.*, 2003, **1**, 478–482.

8 G. Speciale, A. J. Thompson, G. J. Davies and S. J. Williams, *Curr. Opin. Struct. Biol.*, 2014, **28**, 1–13.

9 K.-Y. Li, J. Jiang, M. D. Witte, W. W. Kallemeijn, W. E. Donker-Koopman, R. G. Boot, J. M. F. G. Aerts, J. D. C. Codée, G. A. van der Marel and H. S. Overkleeft, *Org. Biomol. Chem.*, 2014, **12**, 7786–7791.

10 R. J. Ferrier, *J. Chem. Soc., Perkin. Trans. 1*, 1979, 1455–1458.

11 M. Kawana, M. Tsujimoto and S. Takahashi, *J. Carbohydr. Chem.*, 2000, **19**, 67–78.

12 M. J. Martinelli, R. Vaidyanathan, J. M. Pawlak, N. K. Nayyar, U. P. Dhokte, C. W. Doecke, L. M. H. Zollars, E. D. Moher, V. Van Khau and B. Kosmrlj, *J. Am. Chem. Soc.*, 2002, **124**, 3578–3585.

13 B. Bernet and A. T. Vasella, *Helv. Chim. Acta*, 1984, **67**, 1328–1347.

14 H. G. Aurich and F. Biesemeier, *Synthesis*, 1995, 1171–1178.

15 V. D. Bock, H. Hiemstra and J. H. Van Maarseveen, *Eur. J. Org. Chem.*, 2006, 51–68.

16 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.

17 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.

18 R. J. Ferrier, P. Prasit and G. J. Gainsford, *J. Chem. Soc., Perkin. Trans. 1*, 1983, 1629–1634.

Chapter 5

Gluco-1*H*-imidazole: a new class of azole-type β-glucosidase inhibitor

Parts of this chapter have been published:
S.P. Schröder *et al*., Gluco-1*H*-imidazole: a new class of azole-type β-glucosidase inhibitor
J. Am. Chem. Soc. **2018**, http://dx.doi.org/10.1021/jacs.8b02399

5.1 Introduction

Glycosidases catalyze the hydrolysis of oligosaccharides, polysaccharides, glycolipids and glycoproteins, and glycosidase inhibitors are widely regarded as promising therapeutic entities.[1] Azole-containing glycopyranoside mimics are a major class of competitive glycosidase inhibitors.[2] The natural product, nagstatin (**1**), is a potent *O*-linked *N*-acetyl-β-D-glucosaminidase (NAG) inhibitor.[3] Glucopyranose mimics bearing tetrazole (**2**), triazole (**3**, **4**) and imidazole (**5**) moieties have been conceived as β-glucosidase inhibitors by the groups of Tatsuta and Vasella, with inhibitory potencies for sweet almond β-glucosidase ranging from low nanomolar (**5**) to high micromolar (**4**, Figure 1a).[4–7]

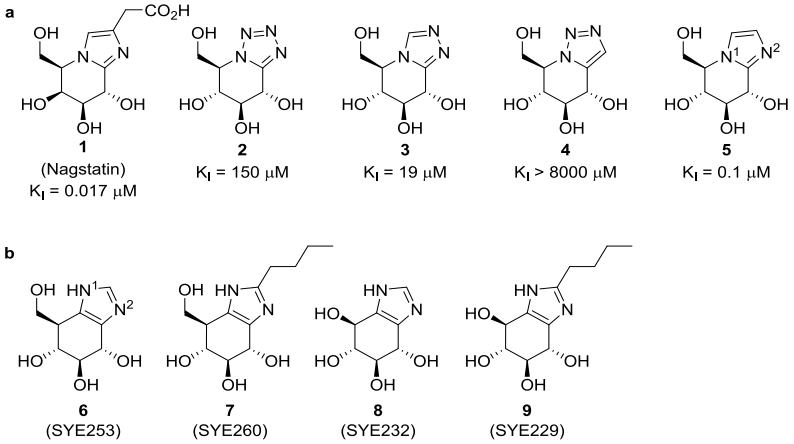
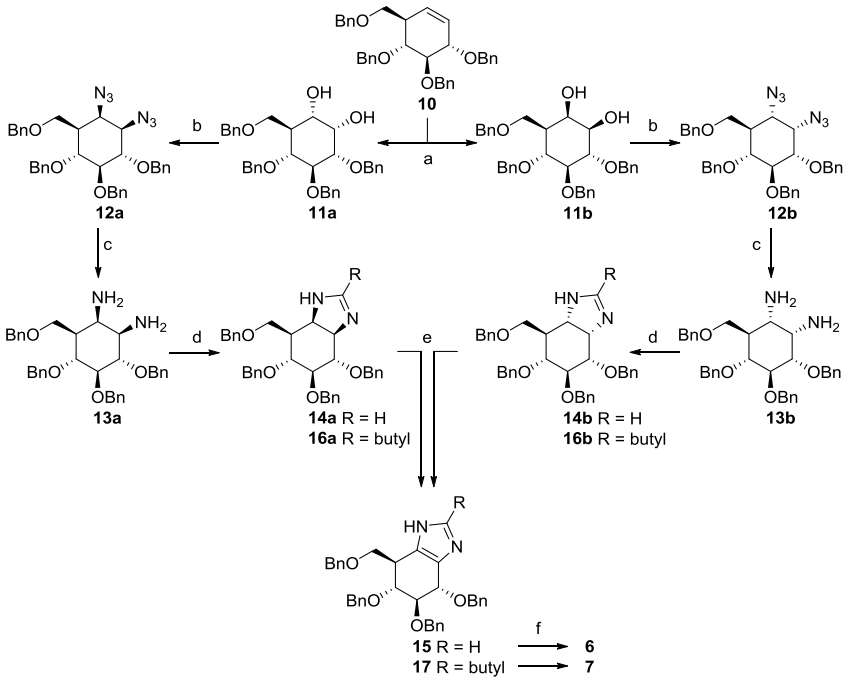


Figure 1. (a) Known cyclitol-azole type glycosidase inhibitors and their inhibitory potencies against porcine NAG (**1**)[3] and sweet almond β-glucosidase (**2**-**5**).[4–7] (b) Gluco-1*H*-imidazoles and respective analogues subject of this study. Numbering of the nitrogen atoms used for describing 1*H*-imidazoles in the text is as depicted in **5** and **6**.

The enzymatic reaction itinerary followed by β-glucosidases normally proceeds via a $^4H_3$ transition state, the thermodynamically most favored conformation of gluco-azoles (see Chapter 1).[8] This conformational transition state mimicry contributes considerably to the inhibitory potency of gluco-azoles. Vasella and co-workers concluded, from extensive studies on gluco-azoles (including compounds **2**-**5**), that the presence of an azole nitrogen at the position taken up by the exocyclic oxygen in a β-glucoside natural substrate is essential for inhibitory potency. They postulated that this 'exocyclic' nitrogen participates in lateral coordination with the catalytic acid/base of "anti-protonating" glucosidases.[9–13] Remarkably, none of the azole-type inhibitors synthesized to date are protic with respect to the azole ring (such as 1*H*-imidazoles), and it was reasoned that such positional isomers would be attractive alternatives to the reported azole glycosidase inhibitors. This Chapter reports on the synthesis and biochemical evaluation of gluco-1*H*-imidazole **6** and its close analogues **7**-**9** as conceptually new gluco-azole type retaining β-glucosidase inhibitors (Figure 1b).

5.2 Results and Discussion

The synthesis of gluco-1*H*-imidazole **6** commenced with cyclohexene **10**,[14] which was transformed into the vicinal diamine following a modified procedure from Llebaria and co-workers for the construction of related 1,2-*cis*-diamines (Scheme 1).[15] It was

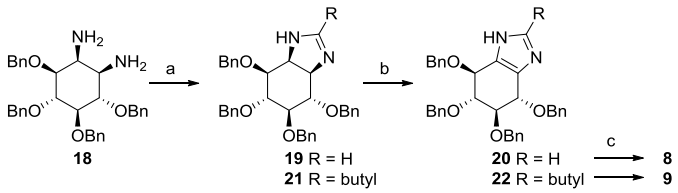


Scheme 1 Synthesis of 1*H*-imidazoles **6** and **7**, starting from cyclohexene **10**.[14] Reagents and conditions: a) $RuCl_3 \cdot 3H_2O$ (7 mol%), $NaIO_4$, EtOAc, MeCN, H_2O, 0 °C, 90 min, yield **11a**: 40%; yield **11b**: 32%; b) MsCl, *N*-methyl-imidazole, Et_3N, $CHCl_3$, rt, 16 h, then NaN_3, DMF, 100 °C, 16 h, yield **12a**: 74%; yield **12b**: 58%; c) PtO_2, H_2, THF, rt, 16 h, yield **13a**: 80%; yield **13b**: 96%; d) trimethyl orthoformate, HFIP, rt, 16h, yield **14a**: 76%; yield **14b**: 87%; or trimethyl orthovalerate, HFIP, rt, 16 h, yield **16a**: 74%; yield **16b**: 78%; e) For **15**: IBX, DMSO, 45 °C, 16 h, 75% from **14a**, 71% from **14b**; for **17**: $(COCl)_2$, DMSO, DCM, -60 °C, 1h, 76% from **16a**; 70% from **16b**; f) $Pd(OH)_2/C$, H_2, HCl, MeOH, quant. **6**; quant. **7**.

found that dihydroxylation using OsO_4/NMO required prolonged reaction times (>7 days) to reach completion. In contrast, rapid conversion of **10** was accomplished using $RuCl_3/NaIO_4$, affording a separable mixture of 1,2-*cis*-dihydroxy isomers **11a** (*S*,*S*) and **11b** (*R*,*R*) in near equimolar quantities.[16] Bismesylation and subsequent azide substitution afforded diazido cyclitols **12a**,**b** with inversion of stereochemistry. Platinum oxide catalyzed hydrogenolysis afforded diamino cyclitols **13a**,**b** in high yield, which were condensed with trimethyl orthoformate in hexafluoroisopropanol (HFIP) to afford β-imidazolines **14a**,**b**.[17] Protected 1*H*-imidazole **15** was obtained by oxidation of either isomer of **14** with IBX/DMSO,[18] and final hydrogenolysis afforded

gluco-1*H*-imidazole **6** in quantitative yield. Derivatization of the imidazole ring of gluco-1*H*-imidazoles can be readily achieved by variation of the trimethyl orthoester employed during imidazoline formation. Thus, 2-butyl imidazolines **16a**,**b** were successfully obtained by condensation of **13a**,**b** with trimethyl orthovalerate. In this case, oxidation to imidazole **17** proved to be low yielding (16% from **16a**) using the IBX/DMSO system, however the yield was significantly improved using oxidation under Swern conditions. Ultimately, hydrogenation afforded gluco-1*H*-2-butyl-imidazole **7** in quantitative yield.

Utilizing tetrabenzyl *myo*-diaminocyclitol[15] **18** as starting material, conduritol B-1*H*-imidazoles **8** and **9** were obtained via essentially the same procedure (Scheme 2). Condensation of the cis-diamine with trimethyl orthoformate in HFIP afforded imidazoline **19**, which was oxidized to **20** using the IBX/DMSO system in 62% yield (Swern oxidation gave a lower yield of 33%). Final deprotection by palladium catalyzed hydrogenolysis afforded **8**. Diamine **18** was also successfully condensed with trimethyl orthovalerate to generate 2-butyl imidazoline **21**, which was oxidized to 1*H*-imidazole **22** using Swern conditions in 74% yield (IBX/DMSO oxidation resulted in a lower yield of 43%). Final deprotection by hydrogenation afforded conduritol-B 2-butyl-1*H*-imidazole **9** in quantitative yield.

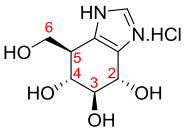


Scheme 2 Synthesis of conduritol B 1*H*-imidazoles **8** and **9**. Reagents and conditions: a) trimethyl orthoformate, HFIP, rt, 16h, yield **19**: 95%; or trimethyl orthovalerate, HFIP, rt, 16 h, yield **21**: 89%; b) For **20**: IBX, DMSO, 45 °C, 16 h, 62%; for **22**: $(COCl)_2$, DMSO, DCM, -60 °C, 1h, 74%; c) $Pd(OH)_2/C$, H_2, HCl, MeOH, quant. **8**; quant. **9**.

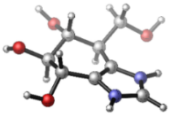
With compounds **6-9** in hand, their lowest energy conformations were modeled by DFT (B3LYP/6-311G(d,p), PCM) calculations on lead compound **6**. The calculated three lowest energy geometries for **6** adopt a $^4H_3$ conformation, differing only in the rotation angle around the C5-C6 axis: *gt*, *gg* and *tg* respectively (Table 1). Based on these optimized structures the spin-spin coupling constants were calculated[19] with

the use of 6-311g(d,p) u+1s as basis set and PCM(H_2O) as solvent model. The calculated total nuclear spin-spin coupling terms were used as calculated spin-spin coupling constants. The calculated $^3J_{(H,H)}$ coupling constants for these low energy $^4H_3$ rotamers of **6** matched well with experimental $^3J_{(H,H)}$ coupling constants, suggesting that **6** adopts a $^4H_3$ conformation in solution.

Table 1 Calculated lowest energy (ΔG^T_{aq}) conformations for gluco-1*H*-imidazole **6**. The calculated $^3J_{(H,H)}$ couplings corresponding to these conformations match well with experimental coupling constants. All coupling constants are in Hertz (Hz).

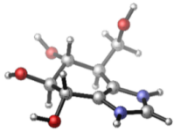


6



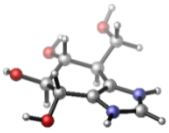
ΔG^T_{aq}= 0.0 kcal/mol

$^4H_3$ *(gt)*



ΔG^T_{aq}= +0.0 kcal/mol

$^4H_3$ *(gg)*



ΔG^T_{aq} = +3.2 kcal/mol

$^4H_3$ *(tg)*

Coupling	Exp. $^3J_{(H,H)}$	Calc. $^3J_{(H,H)}$	Calc. $^3J_{(H,H)}$	Calc. $^3J_{(H,H)}$
H2-H3	7.6	8.0	8.0	7.9
H3-H4	9.8	9.9	10.2	10.0
H4-H5	9.4	10.0	9.8	7.8
H5-H6a	2.8	4.9	3.2	12.5
H5-H6b	4.8	11.7	2.2	4.8
H6a-H6b	11.4	8.8	10.4	8.8

The competitive kinetic constants were determined for 1*H*-imidazoles **6–9** as compared with the canonical glucoimidazole **5** for inhibition of several retaining glucosidases, namely: *Tm*GH1 from *Thermotoga maritima*,[20] *Tx*GH116 from *Thermoanaero-bacterium xylanolyticum*,[21] sweet almond β-glucosidase (GH1), human lysosomal acid β-glucosylceramidase (GBA1, GH30), and human acid α-glucosidase (GAA, GH31). Gluco-1*H*-imidazole **6** displayed micromolar inhibitory activity against all β-glucosidases tested, with K_i values ranging from ~3.9 μM against GBA1 to ~69 μM against *Tx*GH116 (Table 2). Remarkably, given their apparent structural similarity, 1*H*-imidazole **6** proved to be a weaker β-glucosidase inhibitor than classical glucoimidazole **5**, which inhibits β-glucosidases at nanomolar concentrations. Attachment of a butyl moiety, as in **7**, increased the inhibitory potency of the gluco-1*H*-imidazole scaffold; whilst this increase in potency was modest for *Tm*GH1 and *Tx*GH116, **7** inhibited sweet almond β-glucosidase and GBA1 with nanomolar potency.

Notably, the biological role of GBA1 is breakdown of amphiphilic glucosylceramide.[22] Conduritol-1*H*-imidazole **8** did not effectively inhibit any of the glucosidases tested except for GBA1, consistent with a loss of interactions at the O6 position, but also with the fact that GBA1 is effectively inhibited by conduritol B epoxide (CBE).[23] As with the gluco-1*H*-imidazoles, addition of a butyl moiety to the conduritol-1*H*-imidazole scaffold increased activity against sweet almond β-glucosidase and GBA1, rendering **9** a nanomolar inhibitor of these enzymes. Interestingly, none of the compounds displayed any inhibitory activity against cytosolic retaining β-glucosidase GBA2 (GH116) or glucosylceramide synthase (GCS). Finally, none of the compounds were inhibitors of the human α-glucosidase GAA.

Table 2 Inhibition constant (K_i) values in μM for compounds **5**-**9**.

	*Tm*GH1[a]	*Tx*GH116[a]	sweet almond GH1[a]	GBA1[b]	GBA2[c]	GCS[d]	GAA[e]
6	58 ± 1	69 ± 12	23 ± 1	3.9 ± 2	>50[f]	>50[f]	>100[f]
7	39 ± 14	46 ± 5	0.039 ± 0.006	0.133 ± 0.040	>50[f]	>50[f]	>100[f]
8	>100[f]	>100[f]	>100[f]	13.7 ± 6	>50[f]	>50[f]	>100[f]
9	>100[f]	>100[f]	0.221 ± 0.004	0.079 ± 0.001	>50[f]	>50[f]	>100[f]
5	0.026 ± 0.001	0.165 ± 0.006	0.067 ± 0.004[g]	0.070 ± 0.005	>50[f]	>50[f]	>100[f]

Kinetic values are mean ± SD (triplo). [a] Assay measured at pH 6.8 using β-*p*-NPG substrate; [b] Assay measured at pH 5.2 using β-2,4-DNPG substrate; [c] Assay measured *in vitro* at pH 5.2 using 4-MU-β-Glc substrate; [d] Assay measured *in situ* using C6-NBD-ceramide substrate; [e] Assay measured at pH 4.8 using α-*p*-DNPG substrate; [f] Apparent IC_{50}; [g] Literature K_i = 0.100 μM.[6]

As is apparent from the inhibition data depicted in Table 2, 1*H*-imidazoles represent a promising new structural class of glycosidase inhibitors. Competitive GBA1 inhibitors are considered as attractive starting points for the development of pharmacological chaperones: compounds that stabilize the fold of mutant, partially malfunctioning GBA1 in Gaucher patients. The selectivity of compounds **7** and **9** for GBA1 over GBA2 and GCS makes them attractive starting points for this purpose.[24,25] Table 1 also

reveals that 1*H*-imidazole **6** is a comparatively much weaker retaining β-glucosidase inhibitor than the canonical gluco-azole **5**. Given the structures, at first sight, are closely related, this rather striking result was investigated in more depth.

Literature studies on gluco-azoles **2**-**5** and related compounds have shown that the inherent basicity of the azole ring qualitatively correlates to the inhibitory potency.[26] Therefore, the pK_{AH} values of **5** and **6** were determined by ^{1}H-NMR titration experiments (Figure 2). The difference in inhibitory potency of these compounds could not be ascribed to differences in pK_{AH}, since the experimental pK_{AH} of **5** and **6** are almost identical (**5**: pK_{AH} 6.2, lit. 6.1[26]; **6**: pK_{AH} 6.0).

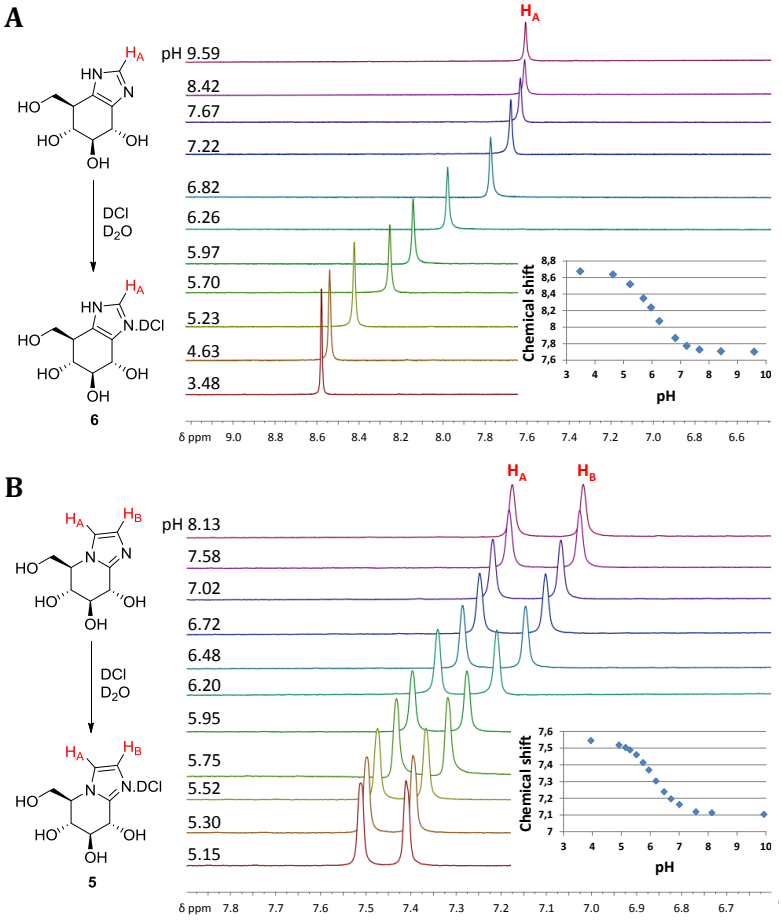


Figure 2 Compound pK_{AH} determination using ^{1}H-NMR. Both imidazoles display similar pK_{AH} values. Experimental details are given in the experimental section. Spectra shown are cropped, showing the azole-ring protons which displayed the highest Δppm shift upon acidification.

The interaction of gluco-1*H*-imidazoles with *Tm*GH1 and *Tx*GH116 was also explored by isothermal titration calorimetry (ITC) at pH 6.8 (at which pH compound **5** shows strongest inhibition in literature studies[10,21]). In line with the K_i data (Table 2), ITC titrations showed nanomolar binding affinities for **5**, and micromolar affinities for **6** and **7** (Table 3). The enthalpic contributions in active site binding for 1*H*-azoles are smaller than those observed for **5**, and only partially compensated for by increased entropic contributions to the binding energy. In line with what has been reported for **5**,[11] the presence of a hydrophobic moiety in **7** reduced its enthalpy of binding compared to **6**, which was outweighed by an increase in favorable entropic contributions to binding, leading to an overall increased ligand affinity in all instances. ITC measurements were also performed at pH 5.8 (enzyme activity optimum), and show a similar trend (data not shown).

Table 3 ITC calculated parameters of binding for **5**-**7** with *Tm*GH1 or *Tx*GH116 at pH 6.8.

	***Tm*GH1**			***Tx*GH116**		
	6	**7**	**5**	**6**	**7**	**5**
N (sites)	0.98 ± 0.06	0.96 ± 0.2	0.96 ± 0.03	0.98 ± 0.09	0.96 ± 0.02	0.98 ± 0.1
K_D (μM)	28.3 ± 6.9	5.02 ± 1.0	0.14 ± 0.02	27.3 ± 0.4	19.6 ± 2.6	0.075 ± 0.01
ΔH[a]	−18.4 ± 1.4	14.2 ± 0.2	−52.7 ± 1.2	−22.5 ± 0.3	−7.4 ± 0.1	−41.8 ± 2.3
−ΔS[a]	−7.6 ± 2.0	−44.6 ± 0.7	13.6 ± 1.3	−3.6 ± 0.2	−19.5 ± 0.4	1.04 ± 2.4
ΔG[a]	−26.0 ± 0.6	−30.3 ± 0.6	−39.1 ± 0.34	−26.1 ± 0.06	−26.9 ± 0.4	−40.7 ± 0.47

All reported values are the mean ± standard deviation from three (ligands **6**, **7**) or four (ligand **5**) technical replicates. [a] Values are in kJmol^{-1}.

Next, the X-ray structure of **6** in complex with *Tm*GH1 was determined to 1.7 Å (PDB: 5OSS), and compared to that of the reported[10] structure of the same enzyme complexed with **5** (PDB: 2CES). Both structures revealed the binding mode of both compounds to be very similar. A single molecule of **6** was observed in the active site after ligand soaking, adopting a ^4E conformation very similar to that observed in the complex of **5** and *Tm*GH1 (Figure 3a).[11] H-bonding interactions made by **6** to active

site residues were identical to those previously observed with **5**, albeit with a slight 'upwards' tilt for **6** compared to **5** (~0.4 or 0.5 Å 'upward' shifts at the apical imidazole carbon, compared with ligands in chains A or B of 2CES respectively; Figure 3b). Crystal structures in *Tx*GH116 at 2.1 Å resolution (Figure 3c,d) also reveal similar binding modes[21] and a ~0.5 Å 'upwards' tilt for **6** compared to **5** (PDB: 5OST and 5BX4, respectively).

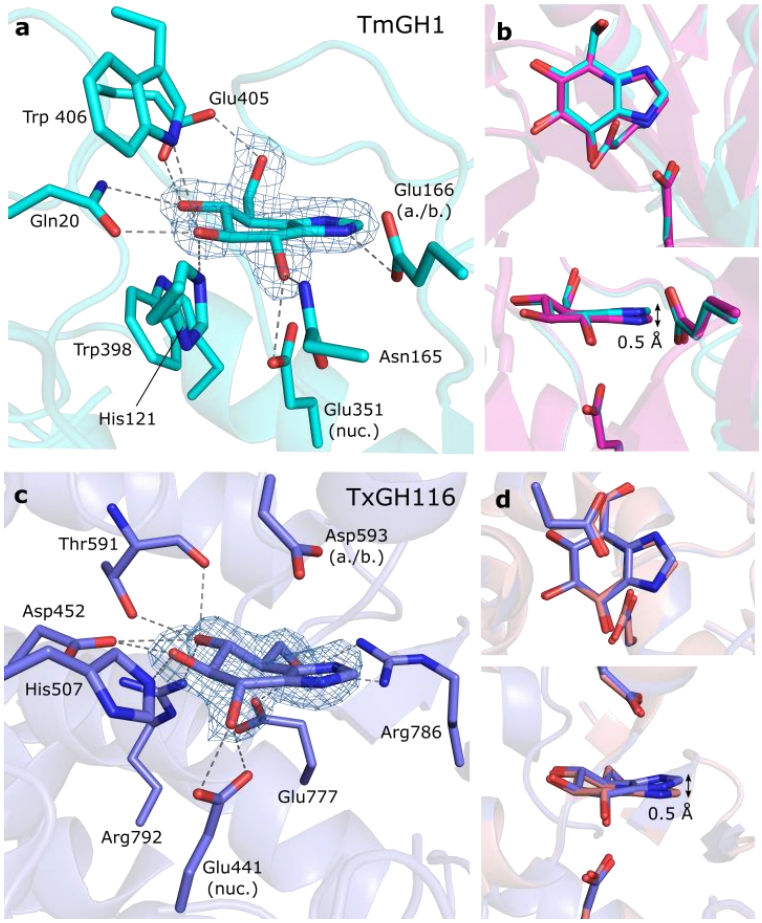


Figure 3 (a) Gluco-*1H*-imidazole **6** in complex with *Tm*GH1, with direct H-bonding interactions shown. (b) Overlay of **6** (cyan) and **5** (pink) within the *Tm*GH1 active site. (c) Gluco-*1H*-imidazole **6** in complex with *Tx*GH116, and direct H-bonding interactions. (d) Overlay of **6** (blue) and **5** (salmon) within the *Tx*GH116 active site. For both *Tm*GH1 and *Tx*GH116, small 'upwards' shifts can be seen for **6** compared to classical glucoimidazole **5**.

The underlying cause for the reduced potency of gluco-1*H*-imidazole **6** compared to **5** is most likely the combination of a number of factors. Repositioning of the N1 atom (from the bridgehead position in **5** to the position in **6**) may have two major consequences that together reduce the binding affinity of **6** compared to **5**. Firstly, considering the situation where the imidazole is in a neutral state[28]: the free lone pair of the N2 atom in **5** likely laterally coordinates to the acid/base residue of the bound glucosidase.[9–12] This positioning of N2 is conserved in **6**, as observed in *Tm*GH1 (Figure 3a). However, and in contrast to **5**, 1*H*-imidazole **6** may undergo prototropic tautomerism (Figure 4a). Thus, whilst the overall pK_{AH} values of **5** and **6** are similar, the N2 lone pair of **6** may be less available for interaction with the glucosidase acid/base, reducing the binding affinity of **6** compared to **5**.

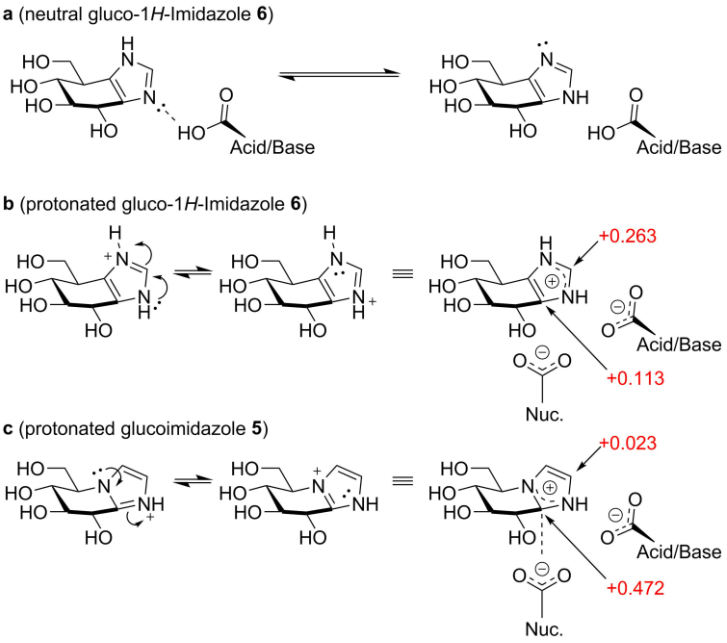


Figure 4 Interactions of gluco-1*H*-imidazole **6** and classical glucoimidazole **5** with the catalytic residues. (a) In the neutral state, gluco-*1H*-imidazoles may undergo prototropic tautomerism, reducing the availability of the N2 lone pair for interaction with a protonated glucosidase acid/base residue. (b) In the protonated state, positive charge is delocalized onto the 'apical' carbon in gluco-1*H*-imidazole, rendering it relatively poorly placed for charge-charge interaction with the anionic glucosidase nucleophile. (c) For classical glucoimidazole **5**, positive charge is delocalized onto the anomeric equivalent carbon, ideally located for charge-charge interaction with the nucleophile residue. Mulliken charges on the 'anomeric' and 'apical' carbon atoms were calculated by DFT and annotated in red.

Protonation of the imidazole in turn (either in solution or by proton abstraction from a catalytic site residue)[28] results in positive charge delocalization. Resulting charge-charge interactions with enzyme active site carboxylates are thought to contribute substantially to enzyme binding energy of azole-type inhibitors.[29] The Mulliken charge on all atoms for protonated **5** and **6** was determined by DFT. Protonation of the azole ring in **5** produces a δ+ charge on the 'anomeric' carbon, which is ideally located for a charge-charge interaction with a retaining glucosidase active site nucleophile (Figure 4c). Conversely, protonation of **6** leads to a δ+ charge largely delocalized onto the 'apical' carbon atom of the imidazole, with the overall δ+ charge also being less pronounced (Figure 4b). This apical δ+ charge is located distal from the catalytic nucleophile and thus relatively poorly positioned for charge-charge interactions, which may explain the reduced binding enthalpy observed in ITC for gluco-1*H*-imidazoles **6** compared to **5**. In addition, the small 'upwards' shift observed in crystal structure complexes of **6** compared to **5** is consistent with a weaker charge-charge interaction of **6** with the catalytic nucleophile. Interestingly, in contrast to neutral **6**, imidazole **5** also contains a significant δ+ character (+0.306 Mulliken charge) on the 'anomeric' carbon in its neutral state.

5.3 Conclusion

To sum up, this Chapter describes the synthesis and in-depth structural and functional analysis of a new class of competitive β-glucosidase inhibitors: the 1*H*-gluco-azoles. Arguably, this class of compounds has no literature precedent and since no obvious route of synthesis existed for related compounds (including the canonical gluco-azole, **5**) new chemistry was designed and developed to effectively access the target compounds. The current route is flexible with respect to the substitution pattern, as shown, and can likely be transferred to configurational isomers targeting other GH family glycosidases, this by applying the route of synthesis on configurational isomers of cyclohexene **10**. The compounds have some interesting structural analogy to those reported by Li and Byers[30], and Field *et al*[31] who have shown that simple 1*H*-imidazoles inhibit several glucosidases effectively. In contrast to these compounds, however, the 1*H*-imidazoles described in this Chapter are not able to form a proton complex between the glucosidase active site residues and the 1*H*-imidazole, this because of tight complexation of the cyclitol moiety within the enzyme active site. 1*H*-imidazole **6** appeared a substantially poorer inhibitor than known gluco-azole **5**. Based on quantum mechanical calculations, it is hypothesized that the reduced

potency (in respect of **5**) is caused by delocalization of the lone pair on the 'glycosidic' nitrogen atom due to tautomerism and/or impaired δ+ charge development at the anomeric center of the inhibitor. As demonstrated by X-ray crystallography, a slight 'upward' tilt of the imidazolium ring from the catalytic nucleophile, possibly due to reduced electrostatic interaction, might be a result of the latter effect. Finally, introduction of an alkyl substituent on the 1*H*-imidazole ring (as in **7**) as well as 'deletion' of the methylene carbon (as in **9**) yielded potent competitive inhibitors of the human lysosomal glucosylceramidase, GBA. These and related compounds may be developed into pharmacological chaperone candidates for Gaucher disease and (by preparing configurational isomers) possibly also for related lysosomal storage disorders characterized by genetic deficiency in other GH family lysosomal glycosidases.

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).

General procedure 1 (GP1): Bis-azidation

The diol starting material was dissolved in dry $CHCl_3$ (0.2 M), then Et_3N (3 equiv.) and *N*-methyl imidazole (10 equiv.) were added and the mixture was cooled to 0 °C. MsCl (4 equiv.) was added and the mixture was stirred 16 h at rt. The mixture was quenched with water at 0 °C, diluted with EtOAc, washed with aq. 1M HCl (2 x), H_2O and brine. The organic layer was dried over $MgSO_4$, filtered and concentrated. After co-evaporation with toluene (2 x), the crude intermediate product was dissolved in dry DMF (0.1 M). NaN_3 (10 equiv.) was added and the mixture was stirred 16 h at 100 °C. Then, the mixture was diluted with H_2O and extracted with Et_2O (3 x). The combined organic layers were washed with H_2O and brine, dried over $MgSO_4$, filtered and concentrated. The product was purified by flash column chromatography using the indicated eluent.

General procedure 2 (GP2): Azide reduction

The bisazido starting material was dissolved in THF (0.05 M) under N_2 atmosphere. PtO_2 (30 mol%) was added, the reaction mixture was purged with H_2 with a balloon, and the mixture was stirred vigorously for 16 h. Then, the mixture was filtered over a small Celite pad and concentrated. The product was purified by flash column chromatography using the indicated eluent.

General procedure 3 (GP3): Imidazoline formation

The diamino starting material was dissolved in HFIP (0.1 M), the appropriate trimethyl orthoester (3 equiv.) was added and the mixture was stirred for 16 h at rt. The mixture was diluted with Et_2O and washed with aq. 1M NaOH (3 x), H_2O and brine, dried over $MgSO_4$, filtered and concentrated. The product was purified by flash column chromatography using the indicated eluent.

It should be noted that for both glucose and conduritol configurations, oxidation of the 2-butyl-imidazolines to the 2-butyl-imidazoles proceeded in only moderate yields when IBX/DMSO was employed. In contrast, we found that oxidation proceeded more smoothly under Swern conditions.[32,33]

General procedure 4 (GP4): Oxidation to the imidazole (IBX, DMSO)

The imidazoline starting material was dissolved in DMSO (0.1 M), IBX[34] (1.5 equiv.) was added and the mixture was stirred 16 h at 45 °C. Next, the mixture was cooled to rt, quenched with aq. 10% $Na_2S_2O_3$ and aq. 1M NaOH. The mixture was stirred for 15 min, diluted with Et_2O, washed with H_2O (3 x) and brine, dried over $MgSO_4$, filtered and concentrated. The product was purified by flash column chromatography using the indicated eluent.

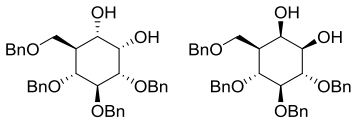
General procedure 5 (GP5): Oxidation to the imidazole (Swern conditions)

To dry DCM (0.1 M based on starting material) was added DMSO (7 equiv.) and the mixture was cooled to -60 °C. Then, oxalyl chloride (5 equiv.) was added slowly and the mixture was stirred for 30 min. The imidazoline starting material was co-evaporated with toluene (2 x), dissolved in dry DCM (1 mL) and added dropwise. The mixture was stirred for 1 h at -60 °C and subsequently quenched with Et_3N (7 equiv.). The cooling bath was removed and the mixture was allowed to reach rt. After stirring 1 h at rt, the mixture was diluted with EtOAc, washed with H_2O (3 x) and brine. The organic layer was dried with $MgSO_4$, filtered and concentrated. The product was purified by flash column chromatography using the indicated eluent.

General procedure 6 (GP6): Hydrogenation

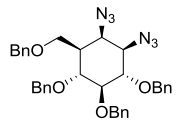
The imidazole starting material was dissolved in MeOH (0.03 M) under N_2 atmosphere, then HCl (1.25M in MeOH, 10 equiv.) and $Pd(OH)_2/C$ (20 wt%) were added and the mixture was purged with H_2 with a balloon. The mixture was stirred vigorously for 16 h, filtered over a small Celite pad and finally concentrated which afforded the pure product.

Compound 11a and compound 11b



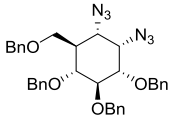
Cyclohexene **10**[35] (1.28 g, 2.46 mmol) was dissolved in EtOAc (15 mL) and MeCN (15 mL) and cooled to 0 °C. A solution of $RuCl_3.3H_2O$ (36 mg, 0.17 mmol) and $NaIO_4$ (789 mg, 3.69 mmol) in H_2O (4.9 mL) was added and the mixture was stirred vigorously at 0 °C for 90 min. The mixture was quenched by addition of aq. 10% $Na_2S_2O_3$ (20 mL) and the mixture was stirred for 15 min. Then the mixture was diluted with H_2O (100 mL) and extracted with EtOAc (3 x 60 mL). The combined organic layers were washed with brine, dried over $MgSO_4$, filtrated and concentrated. The product was purified by flash column chromatography (pentane/EtOAc, 4:1 → 2:1) affording compound **11a** (539 mg, 40%) and **11b** (432 mg, 32%) as white solids. *Analytical data for **11a**:* ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.38 – 7.15 (m, 20H), 4.94 (d, *J* = 10.8 Hz, 1H), 4.87 (d, *J* = 10.8 Hz, 1H), 4.82 (d, *J* = 10.8 Hz, 1H), 4.71 (s, 2H), 4.55 – 4.38 (m, 3H), 4.14 (s, 1H), 3.96 (t, *J* = 9.4 Hz, 1H), 3.84 (dd, *J* = 9.0, 2.5 Hz, 1H), 3.68 – 3.65 (m, 2H), 3.46 – 3.30 (m, 2H), 3.05 (d, *J* = 6.2 Hz, OH), 2.62 (s, OH), 2.18 (tdd, *J* = 10.9, 5.0, 2.5 Hz, 1H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 138.8, 138.5, 138.0, 138.0, 128.5, 128.5, 128.4, 128.4, 128.0, 128.0, 127.9, 127.7, 127.7, 127.7, 127.6, 82.9, 80.0, 77.7, 75.7, 75.3, 73.3, 72.5, 70.3, 69.3, 67.7, 43.2. IR (neat, cm^{-1}): ν 3441, 2868, 1452, 1064. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{35}H_{39}O_6$ 555.27412, found 555.27374. These data are in agreement with those previously reported.[16] *Analytical data for **11b**:* ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.69 – 6.77 (m, 20H), 5.00 – 4.85 (m, 4H), 4.79 (d, *J* = 11.1 Hz, 1H), 4.59 – 4.39 (m, 3H), 4.25 (s, 1H), 3.89 (m, 3H), 3.73 (dd, *J* = 8.9, 2.9 Hz, 1H), 3.61 – 3.45 (m, 2H), 3.34 (s, OH), 2.41 (d, *J* = 4.8 Hz, OH), 1.74 (dq, *J* = 8.0, 2.3 Hz, 1H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 138.7, 138.7, 138.4, 137.6, 128.7, 128.7, 128.6, 128.6, 128.1, 128.1, 127.9, 127.9, 127.9, 127.8, 127.7, 86.7, 82.4, 77.4, 75.8, 75.7, 75.6, 74.6, 73.7, 71.0, 68.9, 43.5. IR (neat, cm^{-1}): ν 3441, 2866, 1452, 1058. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{35}H_{39}O_6$ 555.27412, found 555.27411.

Compound 12a



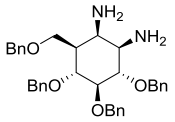
Starting from **11a** (55 mg, 0.1 mmol) and following **GP1**, the product was purified by flash column chromatography (pentane/EtOAc, 15:1) affording compound **12a** as a white solid (45 mg, 74%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.44 – 7.11 (m, 20H), 4.96 – 4.74 (m, 5H), 4.55 – 4.37 (m, 3H), 4.16 (t, *J* = 2.9 Hz, 1H), 3.86 – 3.75 (t, *J* = 9.6 Hz, 1H), 3.73 (dd, *J* = 8.9, 4.3 Hz, 1H), 3.58 – 3.49 (m, 2H), 3.49 – 3.41 (m, 1H), 3.38 (dd, *J* = 10.2, 9.2 Hz, 1H), 2.08 – 1.95 (m, 1H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 138.4, 137.9, 137.8, 137.8, 128.6, 128.6, 128.6, 128.4, 128.2, 128.1, 128.1, 128.0, 127.8, 127.6, 87.0, 81.1, 78.0, 76.1, 75.9, 75.6, 73.6, 67.5, 65.9, 61.1, 43.8. IR (neat, cm^{-1}): ν 2858, 2102, 1359, 1066. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{35}H_{37}N_6O_4$ 605.28708, found 605.33734.

Compound 12b



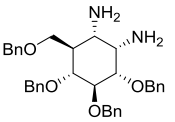
Starting from **11b** (55 mg, 0.1 mmol) and following **GP1**, the product was purified by flash column chromatography (pentane/EtOAc, 15:1) affording compound **12b** as a white solid (35 mg, 58%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.13 (m, 20H), 4.95 – 4.67 (m, 5H), 4.48 (d, *J* = 4.3 Hz, 1H), 4.45 (d, *J* = 3.5 Hz, 1H), 4.34 (d, *J* = 11.5 Hz, 1H), 4.05 (t, *J* = 2.9 Hz, 1H), 3.86 (t, *J* = 9.5 Hz, 1H), 3.82 (d, *J* = 9.2 Hz, 1H), 3.56 (ddd, *J* = 9.6, 6.5, 4.1 Hz, 2H), 3.51 (dd, *J* = 10.6, 2.4 Hz, 2H), 2.02 (t, *J* = 11.2 Hz, 1H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 138.6, 138.4, 138.0, 137.6, 128.7, 128.5, 128.2, 128.1, 128.1, 127.9, 127.8, 127.8, 83.1, 80.4, 77.7, 76.0, 75.7, 73.3, 73.3, 65.0, 63.7, 57.7, 42.5. IR (neat, cm^{-1}): ν 2858, 2098, 1359, 1082. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{35}H_{36}N_6O_4Na$ 627.26902, found 627.26849.

Compound 13a



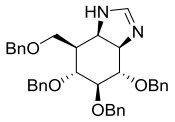
Starting from **12a** (367 mg, 0.61 mmol) and following **GP2**, the product was purified by flash column chromatography (DCM/MeOH, 99:1 → 49:1) affording compound **13a** as a colorless oil (269 mg, 80%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.42 – 7.11 (m, 20H), 4.99 (d, *J* = 11.1 Hz, 1H), 4.96 – 4.83 (m, 3H), 4.66 (d, *J* = 11.1 Hz, 1H), 4.55 – 4.41 (m, 3H), 3.91 (dd, *J* = 11.0, 9.3 Hz, 1H), 3.77 – 3.63 (m, 3H), 3.59 (t, *J* = 9.2 Hz, 1H), 3.38 (t, *J* = 3.0 Hz, 1H), 2.80 (dd, *J* = 10.0, 3.4 Hz, 1H), 1.87 (ddt, *J* = 11.0, 7.4, 3.2 Hz, 1H), 1.53 (s, 4H, 2 x NH_2). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 138.9, 138.7, 138.6, 138.2, 128.6, 128.5, 128.5, 128.1, 128.0, 127.8, 127.8, 127.7, 127.5, 88.3, 82.3, 78.5, 75.8, 75.4, 75.3, 73.3, 68.8, 56.6, 51.7, 44.8. IR (neat, cm^{-1}): ν 2856, 1361, 1066. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{35}H_{41}N_2O_4$ 553.30608, found 553.30585.

Compound 13b



Starting from **12b** (858 mg, 1.42 mmol) and following **GP2**, the product was purified by flash column chromatography (DCM/MeOH, 99:1 → 7:3) affording compound **13b** as a colorless oil (750 mg, 96%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.27 (m, 20H), 4.91 (t, *J* = 11.9 Hz, 2H), 4.79 (d, *J* = 10.7 Hz, 1H), 4.69 (d, *J* = 11.7 Hz, 1H), 4.64 (d, *J* = 11.6 Hz, 1H), 4.46 (d, *J* = 27.6 Hz, 3H), 3.90 (t, *J* = 9.2 Hz, 1H), 3.77 (d, *J* = 7.6 Hz, 1H), 3.64 (d, *J* = 7.8 Hz, 1H), 3.50 (t, *J* = 9.6 Hz, 1H), 3.49 - 3.43 (m, 2H), 2.91 (d, *J* = 10.9 Hz, 1H), 1.97 (t, *J* = 10.9 Hz, 1H), 1.87 (s, 4H, 2 x NH_2). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 139.1, 138.9, 138.5, 128.6, 128.5, 128.1, 128.0, 127.9, 127.9, 127.8, 127.6, 127.6, 83.1, 81.6, 78.9, 75.7, 75.4, 73.2, 72.2, 66.1, 53.3, 49.6, 43.4. IR (neat, cm^{-1}): ν 2860, 1602, 1496, 1452, 1359, 1066. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{35}H_{40}N_2O_4Na$ 575.28803, found 575.28741.

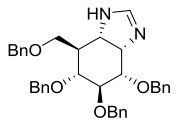
Compound 14a



Starting from **13a** (55 mg, 0.1 mmol) and following **GP3** using trimethyl orthoformate, the product was purified by flash column chromatography (DCM/MeOH, 99:1 → 7:3) affording compound **14a** as a colorless oil (43 mg, 76%). ^{1}H-NMR (400 MHz, CD_3CN) δ 7.41 – 7.18 (m, 20H), 7.07 (s, 1H), 4.84 –

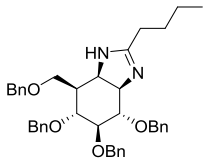
4.64 (m, 5H), 4.51 – 4.42 (m, 3H), 4.01 (dd, *J* = 9.5, 4.4 Hz, 1H), 3.86 (dd, *J* = 9.4, 6.2 Hz, 1H), 3.77 (dd, *J* = 9.2, 4.1 Hz, 1H), 3.68 (t, *J* = 8.8 Hz, 1H), 3.58 – 3.51 (m, 1H), 3.50 – 3.40 (m, 2H), 2.25 (ddt, *J* = 12.3, 8.4, 4.2 Hz, 1H). ^{13}C-NMR (101 MHz, CD_3CN) δ 156.2, 140.0, 139.9, 139.8, 139.7, 129.3, 129.2, 129.2, 128.9, 128.8, 128.7, 128.5, 128.4, 128.4, 85.1, 83.5, 78.9, 74.7, 74.5, 74.1, 73.7, 69.9, 65.6, 60.7, 41.7. IR (neat, cm^{-1}): ν 3278, 3030, 2862, 1654, 1543, 1359, 1066. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{36}H_{38}N_2O_4$ 563.29043, found 563.29022.

Compound 14b



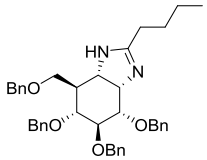
Starting from **13b** (110 mg, 0.2 mmol) and following **GP3** using trimethyl orthoformate, the product was purified by flash column chromatography (DCM/MeOH, 99:1 → 4:1) affording compound **14b** as a colorless oil (98 mg, 87%). ^{1}H-NMR (400 MHz, CD_3CN) δ 7.38 – 7.20 (m, 20H), 7.03 (s, 1H), 4.78 – 4.62 (m, 5H), 4.53 – 4.44 (m, 3H), 3.97 (dd, *J* = 9.3, 4.3 Hz, 1H), 3.86 – 3.81 (m, 1H), 3.81 – 3.78 (m, 1H), 3.70 (t, *J* = 7.0 Hz, 1H), 3.67 – 3.63 (m, 2H), 3.38 (dd, *J* = 11.3, 7.2 Hz, 1H), 1.74 – 1.66 (m, 1H). ^{13}C-NMR (101 MHz, CD_3CN) δ 155.7, 140.1, 140.0, 140.0, 139.8, 129.3, 129.2, 129.2, 128.9, 128.8, 128.7, 128.7, 128.5, 128.4, 128.4, 83.4, 79.6, 79.1, 74.5, 74.4, 73.8, 73.3, 69.2, 61.7, 60.5, 46.2. IR (neat, cm^{-1}): ν 3030, 2868, 1681, 1595, 1454, 1087. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{36}H_{39}N_2O_4$ 563.29043, found 563.29010.

Compound 16a



Starting from **13a** (119 mg, 0.21 mmol) and following **GP3** using trimethyl orthovalerate, the product was purified by flash column chromatography (DCM/MeOH, 99:1 → 8:2) affording compound **16a** as a colorless oil (99 mg, 74%). ^{1}H-NMR (400 MHz, CD_3CN) δ 7.39 – 7.18 (m, 20H), 4.73 (m, 5H), 4.51 – 4.43 (m, 3H), 3.99 (dd, *J* = 9.0, 4.4 Hz, 1H), 3.77 (dt, *J* = 9.2, 4.6 Hz, 2H), 3.68 (t, *J* = 8.8 Hz, 1H), 3.53 (t, *J* = 7.5 Hz, 1H), 3.46 – 3.39 (m, 2H), 2.19 (dq, *J* = 12.8, 4.3 Hz, 1H), 2.07 (td, *J* = 7.4, 3.6 Hz, 2H), 1.51 – 1.42 (m, 2H), 1.31 (dq, *J* = 14.2, 7.2 Hz, 2H), 0.88 (t, *J* = 7.3 Hz, 3H). ^{13}C-NMR (101 MHz, CD_3CN) δ 168.4, 140.2, 140.0, 139.9, 139.8, 129.3, 129.3, 129.2, 129.2, 128.9, 128.8, 128.7, 128.7, 128.4, 128.4, 128.3, 85.5, 83.7, 79.1, 74.8, 74.6, 73.9, 73.5, 70.0, 66.8, 61.6, 42.2, 29.7, 29.4, 23.1, 14.1. IR (neat, cm^{-1}): ν 2868, 1600, 1454, 1363, 1066. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{40}H_{47}N_2O_4$ 619.35303, found 619.35266.

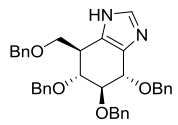
Compound 16b



Starting from **13b** (110 mg, 0.2 mmol) and following **GP3** using trimethyl orthovalerate, the product was purified by flash column chromatography (DCM/MeOH, 99:1 → 4:1) affording compound **16b** as a colorless oil (96 mg, 78%). ^{1}H-NMR (400 MHz, CD_3CN) δ 7.40 – 7.19 (m, 20H), 4.76 (dd, *J* = 11.2, 6.3 Hz, 2H), 4.71 – 4.62 (m, 3H), 4.53 – 4.43 (m, 3H), 4.00 (dd, *J* = 9.0, 3.5 Hz, 1H), 3.83 (t, *J* = 9.4 Hz, 1H), 3.77 – 3.63 (m, 4H), 3.39 (dd, *J* = 11.2, 6.5 Hz, 1H), 2.18 (t, *J* = 7.6 Hz, 2H),

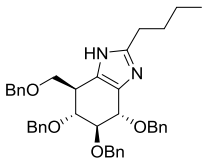
1.73 (t, *J* = 10.4 Hz, 1H), 1.52 (m, 2H), 1.31 (m, 2H), 0.87 (t, *J* = 7.3 Hz, 3H). ^{13}C-NMR (101 MHz, CD_3CN) δ 168.5, 140.1, 140.0, 139.8, 129.3, 129.2, 129.2, 129.1, 128.8, 128.7, 128.7, 128.6, 128.4, 128.4, 128.4, 128.3, 83.6, 80.0, 79.0, 74.6, 74.5, 73.8, 73.3, 69.1, 62.5, 61.0, 46.4, 29.7, 29.5, 23.1, 14.2. IR (neat, cm^{-1}): ν 2862, 1608, 1454, 1359, 1091. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{40}H_{47}N_2O_4$ 619.35303, found 619.35260.

Compound 15



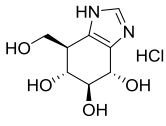
Starting from **14a** (43 mg, 76 µmol) and following **GP4**, the product was purified by flash column chromatography (DCM/MeOH, 99:1 → 67:1) affording compound **15** as a colorless oil (32 mg, 75%). Using the same conditions, product **15** could be obtained from imidazoline **14b** (82 mg, 0.15 mmol) in 71% yield (58 mg). ^{1}H-NMR (500 MHz, CD_3CN) δ 7.51 (s, 1H), 7.42 – 7.18 (m, 20H), 5.05 (d, *J* = 11.5 Hz, 1H), 4.88 – 4.81 (m, 4H), 4.67 – 4.63 (m, 1H), 4.54 – 4.44 (m, 3H), 3.96 (dd, *J* = 9.0, 6.4 Hz, 1H), 3.86 (dd, *J* = 9.0, 3.9 Hz, 1H), 3.76 (t, *J* = 8.1 Hz, 1H), 3.59 (t, *J* = 7.5 Hz, 1H), 3.09 (m, 1H). ^{13}C-NMR (125 MHz, CD_3CN) δ 140.3, 140.0, 139.8, 139.4, 136.8, 129.3, 129.2, 129.2, 129.0, 128.8, 128.8, 128.6, 128.5, 128.4, 128.3, 85.6, 79.5, 77.7, 75.5, 75.3, 73.7, 72.7, 70.0, 40.9. IR (neat, cm^{-1}): ν 3028, 2862, 1496, 1454, 1359, 1087. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{36}H_{37}N_2O_4$ 561.27478, found 561.27454.

Compound 17



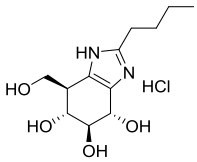
Starting from **16a** (74 mg, 0.12 mmol) and following **GP5**, the product was purified by flash column chromatography (DCM/MeOH, 199:1 → 99:1) affording compound **17** as a colorless oil (56 mg, 76%). Using the same conditions, product **17** could be obtained from imidazoline **16b** (40 mg, 64.7 µmol) in 70% yield (28 mg). ^{1}H-NMR (400 MHz, CD_3CN) δ 7.41 – 7.19 (m, 20H), 4.98 (d, *J* = 11.7 Hz, 1H), 4.87 – 4.76 (m, 4H), 4.63 (dd, *J* = 6.1, 1.5 Hz, 1H), 4.54 – 4.40 (m, 3H), 3.95 (dd, *J* = 8.8, 6.2 Hz, 1H), 3.84 (dd, *J* = 9.0, 4.1 Hz, 1H), 3.81 – 3.72 (m, 1H), 3.66 – 3.55 (m, 1H), 3.13 – 2.96 (m, 1H), 2.63 (d, *J* = 7.6 Hz, 2H), 1.68 – 1.59 (m, 2H), 1.35 (m, 2H), 0.92 (t, *J* = 7.4 Hz, 3H). ^{13}C-NMR (101 MHz, CD_3CN) δ 150.4, 140.4, 140.0, 139.8, 139.5, 129.3, 129.2, 129.2, 129.1, 128.9, 128.8, 128.7, 128.5, 128.4, 128.4, 128.2, 85.4, 79.3, 77.7, 75.3, 75.2, 73.7, 72.5, 69.8, 41.3, 31.6, 29.0, 23.1, 14.1. IR (neat, cm^{-1}): ν 2862, 1454, 1359, 1089. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{40}H_{45}N_2O_4$ 617.33738, found 617.33710.

Compound 6 (gluco-1*H*-imidazole, SYE253)



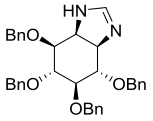
Starting from **15** (26 mg, 46.4 µmol) following **GP6**, the pure product was afforded as a colorless oil (12 mg, quant.). ^{1}H-NMR (400 MHz, D_2O) δ 8.55 (s, 1H), 4.69 (d, *J* = 7.6 Hz, 1H), 4.14 (dd, *J* = 11.4, 2.8 Hz, 1H), 3.93 (dd, *J* = 11.4, 4.8 Hz, 1H), 3.79 (t, *J* = 9.4 Hz, 1H), 3.71 (dd, *J* = 9.8, 7.4 Hz, 1H), 3.01 (s, 1H). ^{13}C-NMR (101 MHz, D_2O) δ 135.2, 128.2, 126.6, 76.9, 69.3, 66.6, 58.8, 40.9 HRMS (ESI-TOF) m/z: $[M+Na]^+$ calc for $C_8H_{12}N_2O_4$ 223.0689, found 223.0702.

Compound 7 (gluco-2-butyl-1*H*-imidazole, SYE260)



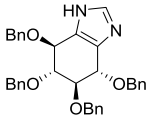
Starting from **17** (48 mg, 77.8 µmol) following **GP6**, the pure product was afforded as a colorless oil (24 mg, quant.). ^{1}H-NMR (400 MHz, MeOD) δ 4.53 (d, *J* = 7.4 Hz, 1H), 4.12 (dd, *J* = 10.8, 3.1 Hz, 1H), 3.89 (dd, *J* = 10.8, 5.0 Hz, 1H), 3.72 (t, *J* = 9.0 Hz, 1H), 3.60 (dd, *J* = 9.3, 7.4 Hz, 1H), 2.94 (t, *J* = 7.6 Hz, 2H), 2.87 (s, 1H), 1.69 – 1.58 (m, 2H), 1.29 (m, 2H), 0.86 (t, *J* = 7.3 Hz, 3H). ^{13}C-NMR (101 MHz, MeOD) δ 150.5, 129.7, 127.6, 78.9, 71.1, 68.2, 60.5, 43.0, 31.0, 26.6, 23.1, 13.8. HRMS (ESI-TOF) m/z: $[M+H]^+$ calc for $C_{12}H_{20}N_2O_4$ 257.1496, found 257.1510.

Compound 19



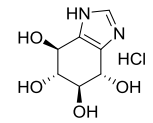
Starting from tetrabenzyl *myo*-diaminocyclitol[15] **18** (400 mg, 0.74 mmol) and following **GP3** using trimethyl orthoformate, the product was purified by flash column chromatography (DCM/MeOH, 99:1 → 7:3) affording the title compound as a colorless oil (387 mg, 95%). ^{1}H-NMR (400 MHz, CD_3CN) δ 7.34 (m, 20H), 7.14 (s, 1H), 4.76 (s, 2H), 4.75 – 4.63 (m, 6H), 4.16 (dd, *J* = 10.7, 4.2 Hz, 1H), 3.97 – 3.87 (m, 2H), 3.82 – 3.73 (m, 2H), 3.57 (dd, *J* = 7.7, 5.6 Hz, 1H). ^{13}C-NMR (101 MHz, CD_3CN) δ 155.2, 139.1, 138.9, 138.8, 138.8, 128.3, 128.3, 128.3, 128.2, 127.9, 127.9, 127.8, 127.8, 127.6, 127.6, 127.5, 127.4, 83.2, 81.3, 80.5, 77.0, 73.0, 72.9, 72.9, 72.4, 64.0, 59.5. IR (neat, cm^{-1}): ν 3030, 2866, 1670, 1452, 1066. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{35}H_{37}N_2O_4$ 549.27478, found 549.27526.

Compound 20



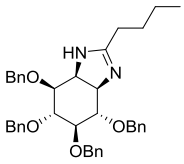
Starting from **19** (55 mg, 0.1 mmol) and using **GP4**, the product was purified by flash column chromatography (DCM/MeOH, 199:1 → 67:1) affording the title compound as a colorless oil (34 mg, 62%). ^{1}H-NMR (400 MHz, CD_3CN) δ 7.57 (s, 1H), 7.44 – 7.14 (m, 20H), 4.92 – 4.83 (m, 4H), 4.80 (d, *J* = 11.2 Hz, 2H), 4.74 – 4.70 (m, 2H), 3.91 (dd, *J* = 4.4, 2.2 Hz, 2H). ^{13}C-NMR (101 MHz, CD_3CN) δ 139.9, 137.9, 129.2, 129.2, 128.8, 128.4, 128.4, 84.5, 78.2 (broad, assigned with HSQC), 75.7, 73.2. IR (neat, cm^{-1}): ν 3030, 2866, 1585, 1496, 1452, 1344, 1053. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{35}H_{35}N_2O_4$ 547.25913, found 547.25897.

Compound 8 (conduritol B-1*H*-imidazole, SYE232)



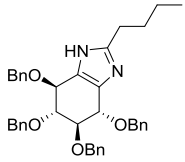
Starting from **20** (18 mg, 32.9 µmol) and following **GP6**, the pure product was afforded as a colorless oil (8.0 mg, quant.). ^{1}H-NMR (400 MHz, D_2O) δ 8.70 (s, 1H), 4.76 (m, 2H), 3.74 – 3.58 (d, *J* = 2.7 Hz, 2H). ^{13}C-NMR (101 MHz, D_2O) δ 135.8. 127.5, 75.9, 66.5. HRMS (ESI-TOF) m/z: $[M+Na]^+$ calc for $C_7H_{10}N_2O_4$ 209.0533, found 209.0542.

Compound 21



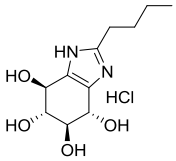
Starting from **18** (53 mg, 0.1 mmol) and following **GP3** using trimethyl orthovalerate, the product was purified by flash column chromatography (DCM/MeOH, 99:1 → 8:2) affording the title compound as a colorless oil (53 mg, 89%). ^{1}H-NMR (400 MHz, CD_3CN) δ 7.42 – 7.18 (m, 20H), 4.77 – 4.56 (m, 8H), 4.11 (dd, *J* = 10.2, 4.1 Hz, 1H), 3.89 – 3.77 (m, 2H), 3.75 – 3.66 (m, 2H), 3.51 (dd, *J* = 7.8, 5.7 Hz, 1H), 2.19 – 2.06 (t, *J* = 7.6 Hz, 2H), 1.55 – 1.42 (m, 2H), 1.29 (m, 2H), 0.84 (t, *J* = 7.3 Hz, 3H). ^{13}C-NMR (101 MHz, CD_3CN) δ 168.9, 140.2, 139.9, 139.8, 129.2, 129.2, 128.9, 128.8, 128.6, 128.5, 128.4, 84.2, 82.3, 81.8, 78.5, 73.9, 73.8, 73.3, 65.6, 61.5, 29.5, 29.4, 23.1, 14.1. IR (neat, cm^{-1}): ν 2870, 1606, 1454, 1357, 1064. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{39}H_{45}N_2O_4$ 605.33738, found 605.33722.

Compound 22



Starting from **21** (53 mg, 0.088 mmol) and following **GP5**, the product was purified by flash column chromatography (DCM/MeOH, 99:1) affording the title compound as a colorless oil (39 mg, 74%). ^{1}H-NMR (400 MHz, CD_3CN) δ 7.49 – 7.17 (m, 20H), 4.87 (m, 8H), 4.70 (dd, *J* = 4.3, 2.1 Hz, 2H), 3.90 (dd, *J* = 4.3, 2.1 Hz, 2H), 2.63 (dd, *J* = 8.2, 7.4 Hz, 2H), 1.64 (m, 2H), 1.35 (m, 2H), 0.91 (t, *J* = 7.4 Hz, 3H). ^{13}C-NMR (101 MHz, CD_3CN) δ 151.6, 140.1, 139.9, 129.2, 128.8, 128.7, 128.4, 128.3, 84.3, 77.0, 75.6, 73.1, 31.5, 29.0, 23.1, 14.1. IR (neat, cm^{-1}): ν 3030, 2870, 1454, 1355, 1058. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{39}H_{43}N_2O_4$ 603.32173, found 603.32178.

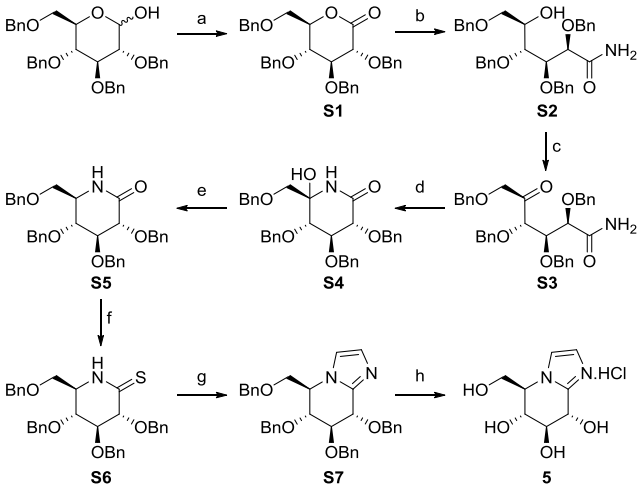
Compound 9 (conduritol B-2-butyl-1*H*-imidazole, SYE229)



Starting from **22** (46 mg, 76.3 μmol) and following **GP6**, the pure product was afforded as a white solid (23 mg, quant.). ^{1}H-NMR (400 MHz, MeOD) δ 4.48 (s, 2H), 3.45 (s, 2H), 2.82 (t, *J* = 7.3 Hz, 2H), 1.70 – 1.54 (m, 2H), 1.27 (q, *J* = 7.1 Hz, 2H), 0.86 (t, *J* = 7.1 Hz, 3H). ^{13}C-NMR (101 MHz, MeOD) δ 151.3, 129.1, 78.1, 68.4, 30.8, 26.6, 23.0, 13.8. HRMS (ESI-TOF) m/z: $[M+Na]^+$ calc for $C_{11}H_{18}N_2O_4$ 265.1159, found 265.1172.

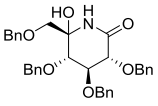
Synthesis of glucoimidazole 5

Glucoimidazole **5** was prepared according to the procedure by Vasella *et al*.[6] from D-gluconolactone (Scheme S1).[36] NMR spectra of this compound are in agreement with those previously reported.



Scheme S1 Synthesis scheme of glucoimidazole **5**. Reagents and conditions: a) Ac_2O, DMSO, rt, 16h; b) NH_3, MeOH, rt, 2h, 57% over two steps; c) Ac_2O, DMSO, rt, 16h; d) NH_3, MeOH, quant. yield over two steps; e) HCOOH, $NaCNBH_3$, MeCN, reflux, 2h, 59%; f) Lawesson's reagent, toluene, rt, 24h, 98%; g) 1. $Hg(OAc)_2$, amino acetaldehyde dimethylacetal, THF, 0 °C, 2h; 2. pTsOH, toluene, 70 °C, 16h, 45% over two steps; h) $Pd(OH)_2/C$, H_2, HCl, MeOH, 6 h, quant.

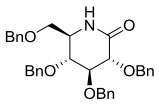
Compound S4



2,3,4,6-O-tetrabenzylglucose (10.0 g, 18.5 mmol) was dissolved in DMSO (100 mL), acetic anhydride (31.5 mL, 333 mmol) was added and the mixture was stirred overnight at rt. The reaction was cooled on ice, quenched with water (900 mL), extracted with Et_2O (3 x 200 mL) and the combined organic fractions were washed with water (3 x 200 mL) and brine. The crude was dissolved in methanolic ammonia (7 M) and stirred overnight at rt. The mixture was concentrated and purified by flash chromatography (pentane/EtOAc, 1:1) to afford an oil (5.8 g, 10.5 mmol, 57%). The oil was taken up in DMSO (35 mL), acetic anhydride (23 mL, 242 mmol) was added and the mixture was stirred overnight at rt. The mixture was quenched with water (300 mL), extracted with Et_2O (3 x 100 mL) and the combined organic fractions were washed with water (100 mL), sat. aq. $NaHCO_3$ (3 x 100 mL) and brine, and concentrated. The crude was dissolved in methanolic ammonia (7 M) and stirred 2h at rt. The mixture was concentrated and purified by flash column chromatography (pentane/EtOAc, 2:1 → 1:1) to give the title compound as an oil (6.0 g, quant.). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.44 – 7.12 (m, 20H), 6.93 (s, 1H), 5.17 (d, *J* = 11.2 Hz, 1H), 4.87 (dd, *J* = 17.3, 11.1 Hz, 2H), 4.75 (t, *J* = 11.8 Hz, 2H), 4.56 – 4.49 (m, 1H), 4.46 (d, *J* = 11.9 Hz, 1H), 4.36 (d, *J* = 11.9 Hz, 1H), 4.24 (t, *J* = 9.2 Hz, 1H), 4.01 (d, *J* = 8.6 Hz, 1H), 3.76 (d, *J* = 7.6 Hz, 2H),

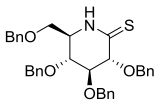
3.32 (q, *J* = 9.6 Hz, 2H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 171.9, 138.3, 137.9, 137.4, 137.1, 128.5, 128.4, 128.4, 128.4, 128.4, 128.3, 128.2, 128.1, 128.1, 128.0, 127.9, 127.8, 127.8, 127.7, 82.0, 79.5, 79.3, 77.3, 75.4, 75.2, 74.8, 73.5, 72.3.

Compound S5



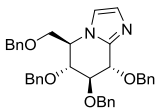
Compound **S4** (6.0 g, 10.8 mmol) was dissolved in MeCN (150 mL), formic acid (40 mL) and $NaCNBH_3$ (1.35 g, 21.6 mmol) was added and the mixture was refluxed for 2 h. Then, the reaction was cooled on ice, quenched with aq. HCl (0.1 M, 80 mL) and stirred for 15 minutes. The reaction mixture was filtered, the resulting crystals were taken up in EtOAc (150 mL), washed with sat. aq. $NaHCO_3$ (3 x 50 mL) and brine, and concentrated to give the title compound as a white solid (3.4 g, 59%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.44 – 7.13 (m, 20H), 6.16 (s, 1H), 5.17 (d, *J* = 11.2 Hz, 1H), 4.85 (d, *J* = 8.3 Hz, 1H), 4.83 (d, *J* = 8.4 Hz, 1H), 4.77 (d, *J* = 11.2 Hz, 1H), 4.72 (d, *J* = 11.1 Hz, 1H), 4.51 – 4.39 (m, 3H), 4.00 (d, *J* = 8.1 Hz, 1H), 3.90 (t, *J* = 8.1 Hz, 1H), 3.65 – 3.50 (m, 3H), 3.34 – 3.19 (m, 1H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 170.6, 138.1, 137.9, 137.7, 137.4, 128.6, 128.5, 128.5, 128.4, 128.4, 128.2, 128.1, 127.9, 127.9, 82.4, 78.9, 77.1, 74.8, 74.7, 73.4, 70.0, 53.9.

Compound S6



Compound **S5** (3.4 g, 6.3 mmol) was co-evaporated with toluene, dissolved in dry toluene (63 mL) and Lawesson's reagent (1.9 g, 4.8 mmol) was added. After stirring overnight at rt, the mixture was stirred over celite and evaporated. The compound was purified with flash column chromatography (pentane/EtOAc, 9:1) to afford the title product as a white solid (3.4 g, 98%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.19 (s, 1H), 7.45 – 7.06 (m, 20H), 5.02 (d, *J* = 11.5 Hz, 1H), 4.74 (d, *J* = 11.5 Hz, 1H), 4.66 (d, *J* = 11.5 Hz, 1H), 4.58 (d, *J* = 11.5 Hz, 1H), 4.50 – 4.41 (m, 4H), 4.35 (d, *J* = 11.5 Hz, 1H), 3.89 (dt, *J* = 11.5, 5.5 Hz, 2H), 3.63 (dd, *J* = 9.8, 3.2 Hz, 1H), 3.57 (dd, *J* = 9.3, 4.7 Hz, 1H), 3.38 (dd, *J* = 9.8, 7.4 Hz, 1H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 200.5, 137.6, 137.4, 137.2, 128.7, 128.6, 128.5, 128.5, 128.4, 128.3, 128.2, 128.1, 128.0, 82.5, 81.4, 78.4, 73.5, 72.9, 72.7, 72.6, 68.4, 56.0.

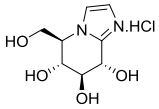
Compound S7



Compound **S6** (101 mg, 0.18 mmol) was dissolved in THF (1 mL) and cooled to 0 °C. Aminoacetaldehyde dimethyl acetal (100 µL, 0.93 mmol) and $Hg(OAc)_2$ (82 mg, 0.26 mmol) were added and the mixture was stirred for 2 h at 0 °C. Celite was added to the mixture and the resulting suspension was filtered over celite. The filtrate was diluted with water (50 mL) and extracted with Et_2O (2 x 25 mL). The combined organic fractions were washed with brine, dried over $MgSO_4$, filtrated and concentrated. The crude was dissolved in toluene (5 mL), pTsOH (95 mg, 0.5 mmol) was added and the mixture was stirred overnight at 70 °C. The mixture was quenched with water (30 mL), extracted with Et_2O (3 x 15 mL) and the combined organic fractions were washed with sat. aq. $NaHCO_3$ (3 x 10 mL) and brine, dried

over $MgSO_4$, filtrated and concentrated. The compound was purified with flash column chromatography (pentane/EtOAc, 2:1) to give the title product as an oil (45 mg, 45% over two steps). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.49 – 7.13 (m, 20H), 7.13 – 7.07 (m, 1H), 7.06 – 7.00 (m, 1H), 5.18 (d, *J* = 11.6 Hz, 1H), 4.85 (dd, *J* = 22.6, 11.1 Hz, 3H), 4.74 (d, *J* = 5.7 Hz, 1H), 4.69 (d, *J* = 11.3 Hz, 1H), 4.50 (d, *J* = 11.1 Hz, 1H), 4.45 (d, *J* = 4.0 Hz, 2H), 4.17 (ddd, *J* = 7.9, 5.2, 2.9 Hz, 1H), 4.08 (dd, *J* = 7.6, 5.8 Hz, 1H), 3.90 – 3.79 (m, 2H), 3.74 (dd, *J* = 10.4, 5.3 Hz, 1H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 144.1, 138.3, 138.0, 137.7, 137.4, 129.5, 128.6, 128.6, 128.5, 128.4, 128.3, 128.2, 128.1, 128.1, 128.0, 127.9, 127.7, 117.4, 82.2, 76.1, 74.4, 74.4, 74.2, 73.4, 72.8, 68.5, 58.2.

Compound 5 (glucoimidazole)



Compound **S7** (45 mg, 80 µmol) was dissolved in methanol (2 mL) under N_2 atmosphere and methanolic HCl (1.25 M, 642 µL, 10 eq.) and $Pd(OH)_2$/C (20 wt%, 34 mg) were added. The flask was purged with H_2 gas and the mixture was stirred vigorously for 6 h. The flask was purged with N_2 gas, the reaction mixture filtered over celite and concentrated to give the title compound (16 mg, quant.). ^{1}H-NMR (400 MHz, D_2O) δ 7.64 (s, 1H), 7.54 (s, 1H), 4.86 (d, *J* = 9.0 Hz, 2H), 4.29 (d, *J* = 13.0 Hz, 1H), 4.24 (d, *J* = 8.6 Hz, 1H), 4.11 (dd, *J* = 12.9, 2.7 Hz, 1H), 4.03 (t, *J* = 9.4 Hz, 1H), 3.91 (t, *J* = 9.5 Hz, 1H). ^{13}C-NMR (101 MHz, D_2O) δ 145.2, 120.6, 119.7, 73.2, 66.9, 66.3, 62.2, 58.3.

Determining compound pK_{AH} values

The pK_{AH} values of the imidazoles were determined with the method described by Gift *et al.*[37] using a Metrohm 691 pH-meter and Hamilton spintrode. The compound was dissolved in D_2O (0.6 mL) and basified with NaOD (0.1 M in D_2O) to pH > 8. Then, the mixture was acidified by stepwise addition of DCl (0.1 M in D_2O) and a ^{1}H-NMR spectrum (Brüker DMX-300) was recorded after each addition. A correction for determination in D_2O instead of H_2O was applied according to Krężel *et al.*[38]

Biochemical and Biological Methods

Enzyme preparations used for IC_{50} and kinetics measurements were as follows: Recombinant human β-glucosidase GBA1 (Cerezyme) and α-glucosidase recombinant human GAA (Myozyme) were obtained from Genzyme, USA. Bacterial β-glucosidase enzymes *Tm*GH1[39] and *Tx*GH116[40] were expressed as previously described. β-Glucosidase from almonds was purchased from Sigma Aldrich as lyophilized powder (7.9 U/mg solid). Cellular homogenates of a stable HEK293 over-expressing GBA2 cell line were obtained as previously described[41] and were pre-incubated for 30 min with 1mM CBE. Proteins were stored in small aliquots at -80 °C until use. *p*-nitrophenyl-β-D-glucopyranoside was purchased from Sigma Aldrich, 4-MU-β-D-glucopyranoside was purchased from Glycosynth, and C6-NBD-ceramide (6-[*N*-methyl-*N*-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)aminododecanoyl] sphingosine) from Molecular probes. GBA1 inhibitor Conduritol-β-Epoxide (CBE) was purchased from Enzo. 2,4-dinitrophenyl-β-D-glucopyranoside[42] and 2,4-dinitrophenyl-α-D-glucopyranoside[43]

were synthesized following synthetic procedures previously described and their spectroscopic data are in agreement with those previously reported.

***In vitro* apparent IC_{50} measurements**

To determine *in vitro* apparent IC_{50} values, 25 µL of enzyme solution was pre-incubated with 25 µL of a range of 6 inhibitor dilutions for 30 min in a 96 well plate, using the following buffers: GBA1 in 150 mM McIlvaine buffer pH 5.2, 0.2% taurocholate (w/v), 0.1% Triton X-100 (v/v) and 0.1% bovine serum albumin (BSA) (w/v); GAA in 150 mM McIlvaine buffer pH 4.8 and 0.1% BSA (w/v); *Tm*GH1, *Tx*GH116 and β-glucosidase from sweet almonds in 50 mM $NaHPO_4$ pH 6.8 and 0.1% BSA (w/v).
After 30 min of pre-incubation, 50 µL of substrate solution in the same buffer was added to this E (25 µL) + I (25 µL) mixture (total reaction volume 100 µL). GBA1 residual activity was measured using final 24 nM concentration of enzyme (Cerezyme) and 200 µM of 2,4-dinitrophenyl-β-D-glucopyranoside substrate, incubated for 30 min at 37 °C. GAA activity was measured using final concentrations of 156 nM and 200 µM of 2,4-dinitrophenyl-α-D-glucopyranoside substrate, for 30 min at 37 °C. *Tm*GH1, *Tx*GH116 and β-glucosidase from sweet almonds residual activity was measured using final concentrations of 37 nM, 82 nM and 0.125 U/mL respectively and 400 µM of *p*-nitrophenyl-β-D-glucopyranoside, for 30 min at 37 °C. Finally, all enzyme reactions were monitored for 10 minutes and the release of 2,4-dinitrophenolate or *p*-nitrophenolate and UV-absorbance was measured at 420 nm in a Tecan GENios Microplate Reader. Values plotted for [I] are those in the final reaction mixture, containing E + I + S. Data was corrected for background absorbance, then normalized to the untreated control condition and finally curve-fitted via one phase exponential decay function (GraphPad Prism 5.0). Apparent i*n vitro* IC_{50} values were determined in technical triplicates.

For GBA2, 12.5 µL of lysate was pre-incubated with 12.5 µL of a range of 7 inhibitor dilutions for 30 min at 37°C. Afterwards, 100 µL of 3.7 mM 4-MU-β-D-glucopyranoside in 150 mM McIlvaine buffer pH 5.8 and 0.1% BSA (w/v) were added and incubated for 1h at 37°C. After stopping the substrate reaction with 200 µL 1M NaOH-Glycine (pH 10.3), liberated 4-MU fluorescence was measured with a fluorimeter LS55 (Perkin Elmer) using λ_{Ex} 366 nm and λ_{Em} 445 nm. All IC_{50} values were determined in duplicate.

In situ apparent IC50 values for GCS were determined with NBD-ceramide as substrate as previously described[44]. RAW 264.7 (American Type culture collection) were cultured in RPMI medium (Gibco) supplemented with 10% FCS, 1 mM GlutaMAX™ and 100 units/mL penicillin/streptomycin (Gibco) at 37°C and 5% CO2. The RAW 264.7 cells were grown to confluence in 12-well plates and pre-incubated for 1h with 300 µM CBE, followed by 1h incubation at 37°C in the presence of a range of 6 inhibitor concentrations and with 1 nmol C6-NBD-ceramide. The cells were washed 3x with PBS and harvested by scraping. After lipid extraction[45], the C6-NBD lipids were separated and detected by

HPLC (λ_{Ex} 470 nm and λ_{Em} 530 nm). IC_{50} values were determined in duplicate from the titration curves of observed formed C6-NBD-glucosylceramide.

Kinetic studies

The kinetic studies of reversible imidazole inhibitors in *Tm*GH1, *Tx*GH116 and β-glucosidase from sweet almonds were performed by monitoring the UV-absorbance of *p*-nitrophenolate released from *p*-nitrophenyl β-D-glucopyranoside. *Tm*GH1, *Tx*GH116 and β-glucosidase from sweet almonds (25 µL) at 37 nM, 82 nM and 0.125 U/mL respectively in 50 mM phosphate buffer (pH 6.8) and 0.1% BSA (w/v) were pre-incubated with a range of inhibitor dilutions (25 µL) for 30 min at 37 °C in a 96 well plate. The reaction was then started by adding 50 µL of different *p*-nitrophenyl β-D-glucopyranoside substrate concentrations (0.05, 0.1, 0.25, 0.5, 0.75, 1.0, 2.5 and 5 mM) in 50 mM phosphate buffer (pH 6.8) to the 50 µL enzyme-inhibitor mixture. For kinetic studies in human recombinant β-glucosidase, 25 µL of 24 nM Cerezyme in 150 mM McIlvaine buffer pH 5.2 supplemented with 0.2% taurocholate (w/v), 0.1% Triton X-100 (v/v) and 0.1% bovine serum albumin (BSA) (w/v), was incubated with a range of inhibitor dilutions (25 µL) for 30 min at 37 °C in a 96 well plate. The reaction was then started by adding 50 µL of different 2,4-dinitrophenyl-β-D-glucopyranoside substrate concentrations (0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6 and 0.7 mM) in the previously described 150 mM McIlvaine buffer (pH 5.2) to the 50 µL enzyme-inhibitor mixture.

The release of *p*-nitrophenolate or 2,4-dinitrophenolate was monitored by absorbance at 420 nm for 10 min (at 25 °C for *Tm*GH1, *Tx*GH116 and β-Glucosidase from almonds or 37 °C for human Cerezyme and Myozyme) in a Tecan GENios Microplate Reader to determine the hydrolysis rate. The K_i values of reversible competitive or linear mixed inhibition were determined by Michaelis-Menten model using standard nonlinear regression (GraphPad Prism 5.0). K_i values were determined in technical triplicates.

Protein expression and crystallography

*Tm*GH1 was produced by expression of the construct pET-28a-*Tm*GH1-His_6 and purified as described by Zechel *et al.*[46] TmGH1 was crystallized by sitting drop vapour diffusion, with the protein at 10 mg/ml in 50 mM imidazole pH 7.0 and the well solution comprised of 11 % polyethylene glycol (PEG) 4000, 0.1 M imidazole pH 7.0, 50 mM calcium acetate, 100 mM trimethylamine *N*-oxide. The protein drop was seeded with a seed stock grown under similar conditions. To generate the ligand complex, a crystal of *Tm*GH1 was soaked with 10 mM gluco-1*H*-imidazole **6** for 4 days, and fished into liquid nitrogen via a cryoprotectant solution comprised of the well solution supplemented with 25 % (v/v) ethylene glycol. Data were collected at Diamond beamline I03, processed using *DIALS*[47] and scaled using *AIMLESS*[48] to a resolution of 1.7 Å. The structure was solved using 1OD0 without the water molecules as the starting model for *REFMAC*[49], and refined by manual rebuilding in *Coot*[50] combined with further cycles of refinement using *REFMAC*. Crystal structure figures were generated using Pymol.

There are two molecules in the asymmetric unit of the *Tm*GH1 crystal structure. **6** is modelled in the active site of chain B only at an occupancy of 0.8, whilst the equivalent site in chain A has been modelled with ethylene glycol in two alternative conformations and two water molecules. The authors have observed that crystal structures of ligand complexes obtained with *Tm*GH1 crystals sometimes yield ligand in only one out of two molecules in the asymmetric unit. It may be that some of the active sites are blocked by N-terminal residues on adjacent chains, as observed for 1OD0.pdb (where 5 residues at the start of chain B extend into the active site of mol A), but for this complex it has not been possible to definitively model N-terminal residues before Val3.

*Tx*GH116 was produced by expression of construct pET30a-*Tx*GH116 Δ1-18 with a C-terminal His_6 tag and purified as described by Charoenwattanasatien *et al*.[21] *Tx*GH116 was crystallized by the sitting drop vapour diffusion method, with a well solution of 0.2M ammonium sulfate, 20 % (v/v) PEG 3350, 0.1 M Bis-Tris pH 6.8. To generate ligand complexes, crystals of *Tx*GH116 were soaked with 10 mM gluco-1*H*-imidazole **6** for 20 hours, before fishing via a cryoprotectant solution with 25 % (v/v) ethylene glycol. Data were collected at Diamond beamline I03, processed using *DIALS* and scaled using *AIMLESS* to a resolution of 2.1 Å. The structure was solved using *MOLREP*[51], with 5BVU as the model, and the solved structure refined by cycles of manual rebuilding in *Coot* and refinement using *REFMAC*. Crystal structure figures were generated using Pymol.

ITC

ITC experiments were carried out using a MicroCal AutoITC200 (Malvern Instruments, formerly GE Healthcare). All titrations were run at 25 °C in 50 mM Sodium Phosphate, pH 5.8 or 6.8. Proteins were buffer exchanged into ITC buffer via at least 3 rounds of dilution/concentration using an Amicon Ultra spin concentrator (Millipore), and further degassed under vacuum prior to use. Cell concentrations of 100 μM (protein) and syringe concentrations of 2 mM (ligand) were used for titrations using gluco-1*H*-Imidazoles **6** and **7**. Cell concentrations of 50 μM and syringe concentrations of 500 μM were used for titrations using **5**. Analyses were carried out using the MicroCal PEAQ-ITC analysis software (Malvern Instruments).

DFT-Geometry optimisation

All calculations were performed with DFT as level of theory in a combination of the B3LYP functional. A conformer distribution search option included in the Spartan 04 program,[52] in gas phase with the use of 6-31G(d) as basis set, was used as starting point for the geometry optimisation. All generated structures were optimized with Gaussian 03[53] at 6-311G(d,p), their zero-point energy corrections calculated, and further optimised with incorporated polarizable continuum model (PCM) to correct for solvation in water. Visualisation of the conformations of interest was done with CYLview.[53]

DFT-Restricted conformational energy surface calculations

The geometry with the lowest, ZPE corrected, solvated free energy was selected as the starting point for the partial conformational energy surface calculation. A survey of the possible neighbouring conformational space was made by scanning two dihedral angles, including the C1-C2-C3-C4 (D1), C3-C4-C5-O (D3) ranging from −60° to −20°. The C5-O-C1-C2 (D5) was fixed at 0° since this is highly favoured. The resolution of this survey is determined by the step size which was set to 5° per puckering parameter. These structures were calculated with Gaussian 03 with a 6-311G(d,p) as basis set. Furthermore, solvation effects of H_2O were taken into account with a polarisable continuum model (PCM) function.

References

1 N. Asano, *Glycobiology*, 2003, **13**, 93–104.

2 V. H. Lillelund, H. H. Jensen, X. Liang and M. Bols, *Chem. Rev.*, 2002, **102**, 515–553.

3 T. Aoyagi, H. Suda, K. Uotani, F. Kojima, T. Aoyama, K. Horiguchi, M. Hamada and T. Takeuchi, *J. Antibiot.* , 1992, **45**, 1404–1408.

4 K. Tatsuta, S. Miura, S. Ohta and H. Gunji, *Tetrahedron Lett.*, 1995, **36**, 1085–1088.

5 P. Ermert and A. Vasella, *Helv. Chim. Acta*, 1991, **74**, 2043–2053.

6 T. Granier, N. Panday and A. Vasella, *Helv. Chim. Acta*, 1997, **80**, 979–987.

7 T. D. Heightman, M. Locatelli and A. Vasella, *Helv. Chim. Acta*, 1996, **79**, 2190–2200.

8 G. J. Davies, A. Planas and C. Rovira, *Acc. Chem. Res.*, 2012, **45**, 308–316.

9 A. Varrot, M. Schülein, M. Pipelier, A. Vasella and G. J. Davies, *J. Am. Chem. Soc.*, 1999, **121**, 2621–2622.

10 T. M. Gloster, S. Roberts, G. Perugino, M. Rossi, M. Moracci, N. Panday, M. Terinek, A. Vasella and G. J. Davies, *Biochemistry*, 2006, **45**, 11879–11884.

11 V. Notenboom, S. J. Williams, R. Hoos, S. G. Withers and D. R. Rose, *Biochemistry*, 2000, **39**, 11553–11563.

12 M. Hrmova, V. A. Streltsov, B. J. Smith, A. Vasella, J. N. Varghese and G. B. Fincher, *Biochemistry*, 2005, **44**, 16529–16539.

13 T. D. Heightman and A. T. Vasella, *Angew. Chem. Int. Ed.*, 1999, **38**, 750–770.

14 F. G. Hansen, E. Bundgaard and R. Madsen, *J. Org. Chem.*, 2005, **70**, 10139–10142.

15 A. Trapero and A. Llebaria, *ACS Med. Chem. Lett.*, 2011, **2**, 614–619.

16 M. Artola, L. Wu, M. J. Ferraz, C. L. Kuo, L. Raich, I. Z. Breen, W. A. Offen, J. D. C. Codée, G. A. van der Marel, C. Rovira, J. M. F. G. Aerts, G. J. Davies and H. S. Overkleeft, *ACS Cent. Sci.*, 2017, **3**, 784–793.

17 S. Khaksar, A. Heydari, M. Tajbakhsh and S. M. Vahdat, *J. Fluor. Chem.*, 2010, **131**, 1377–1381.

18 K. C. Nicolaou, C. J. N. Mathison and T. Montagnon, *J. Am. Chem. Soc.*, 2004, **126**, 5192–5201.

19 T. Bally and P. R. Rablen, *J. Org. Chem.*, 2011, **76**, 4818–4830.

20 L. Liu, Z. Zeng, G. Zeng, M. Chen, Y. Zhang, J. Zhang, X. Fang, M. Jiang and L. Lu, *Bioorg. Med. Chem. Lett.*, 2012, **22**, 837–843.

21 R. Charoenwattanasatien, S. Pengthaisong, I. Breen, R. Mutoh, S. Sansenya, Y. Hua, A. Tankrathok, L. Wu, C. Songsiriritthigul, H. Tanaka, S. J. Williams, G. J. Davies, G. Kurisu and J. R. K. Cairns, *ACS Chem. Biol.*, 2016, **11**, 1891–1900.

22 M. Abdul-Hammed, B. Breiden, G. Schwarzmann and K. Sandhoff, *J. Lipid Res.*, 2017, **58**, 563–577.

23 G. A. Grabowski, K. Osiecki-Newman, T. Dinur, D. Fabbro, G. Legler, S. Gatt and R. J. Desnick, *J. Biol.*

Chem., 1986, **261**, 8263–8269.

24 A. R. Sawkar, W.-C. Cheng, E. Beutler, C.-H. Wong, W. E. Balch and J. W. Kelly, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 15428–15433.

25 R. A. Steet, S. Chung, B. Wustman, A. Powe, H. Do and S. A. Kornfeld, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 13813–13818.

26 N. Panday and A. Vasella, *Synthesis*, 1999, 1459–1468.

27 N. Panday, Y. Canac and A. Vasella, *Helv. Chim. Acta*, 2000, **83**, 58–79.

28 B. Wang, J. I. Olsen, B. W. Laursen, J. C. Navarro Poulsen and M. Bols, *Chem. Sci.*, 2017, **8**, 7383–7393.

29 T. D. Heightman, A. Vasella, K. E. Tsitsanou, S. E. Zographos, V. T. Skamnaki and N. G. Oikonomakos, *Helv. Chim. Acta*, 1998, **81**, 853–864.

30 Y.-K. Li and L. D. Byers, *Biochim. Biophys. Acta*, 1989, **999**, 227–232.

31 R. A. Field, A. H. Haines, E. J. T. Chrystal and M. C. Luszniak, *Biochem. J.*, 1991, **274**, 885–889.

32 D. Keirs and K. Overton, *J. Chem. Soc. Chem. Commun.*, 1987, 1660–1661.

33 G. M. Dubowchik, L. Padilla, K. Edinger and R. a. Firestone, *J. Org. Chem.*, 1996, **61**, 4676–4684.

34 M. Frigerio, M. Santagostino and S. Sputore, *J. Org. Chem.*, 1999, **64**, 4537–4538.

35 T. J. M. Beenakker, D. P. A. Wander, W. A. Offen, M. Artola, L. Raich, M. J. Ferraz, K. Y. Li, J. H. P. M. Houben, E. R. van Rijssel, T. Hansen, G. A. van der Marel, J. D. C. Codée, J. M. F. G. Aerts, C. Rovira, G. J. Davies and H. S. Overkleeft, *J. Am. Chem. Soc.*, 2017, **139**, 6534–6537.

36 H. S. Overkleeft, J. van Wiltenburg and U. K. Pandit, *Tetrahedron*, 1994, **50**, 4215–4224.

37 A. D. Gift, S. M. Stewart and P. K. Bokashanga, *J. Chem. Educ.*, 2012, **89**, 1458–1460.

38 A. Krezel and W. Bal, *J. Inorg. Biochem.*, 2004, **98**, 161–166.

39 T. M. Gloster, P. Meloncelli, R. V. Stick, D. Zechel, A. Vasella and G. J. Davies, *J. Am. Chem. Soc.*, 2007, **129**, 2345–2354.

40 R. Charoenwattanasatien, S. Pengthaisong, I. Breen, R. Mutoh, S. Sansenya, Y. Hua, A. Tankrathok, L. Wu, C. Songsiriritthigul, H. Tanaka, S. J. Williams, G. J. Davies, G. Kurisu and J. R. K. Cairns, *ACS Chem. Biol.*, 2016, **11**, 1891–1900.

41 D. Lahav, B. Liu, R. J. B. H. N. van den Berg, A. M. C. H. van den Nieuwendijk, T. Wennekes, A. T. Ghisaidoobe, I. Breen, M. J. Ferraz, C. L. Kuo, L. Wu, P. P. Geurink, H. Ovaa, G. A. van der Marel, M. van der Stelt, R. G. Boot, G. J. Davies, J. M. F. G. Aerts and H. S. Overkleeft, *J. Am. Chem. Soc.*, 2017, **139**, 14192–14197.

42 H. M. Chen and S. G. Withers, *ChemBioChem*, 2007, **8**, 719–722.

43 S. K. Sharma, G. Corrales and S. Penadés, *Tetrahedron Lett.*, 1995, **36**, 5627–5630.

44 T. Wennekes, R. J. B. H. N. van den Berg, W. Donker, G. A. van der Marel, A. Strijland, J. M. F. G. Aerts and H. S. Overkleeft, *J. Org. Chem.*, 2007, **72**, 1088–1097.

45 E. G. Bligh and W. J. Dyer, *Can. J. Biochem. Physiol.*, 1959, **37**, 911–917.

46 D. L. Zechel, A. B. Boraston, T. Gloster, C. M. Boraston, J. M. Macdonald, D. M. G. Tilbrook, R. V. Stick and G. J. Davies, *J. Am. Chem. Soc.*, 2003, **125**, 14313–14323.

47 D. G. Waterman, G. Winter, R. J. Gildea, J. M. Parkhurst, A. S. Brewster, N. K. Sauter and G. Evans, *Acta Crystallogr. Sect. D, Struct. Biol.*, 2016, **72**, 558–575.

48 P. R. Evans and G. N. Murshudov, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2013, **69**, 1204–1214.

49 A. A. Vagin, R. A. Steiner, A. A. Lebedev, L. Potterton, S. McNicholas, F. Long and G. N. Murshudov, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2004, **60**, 2184–2195.

50 P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2010, **66**, 486–501.

51 A. Vagin and A. Teplyakov, *J. Appl. Crystallogr.*, 1997, **30**, 1022–1025.

52 J. Kong, C. A. White, A. I. Krylov, D. Sherrill, R. D. Adamson, T. R. Furlani, M. S. Lee, A. M. Lee, S. R. Gwaltney, T. R. Adams, C. Ochsenfeld, A. T. B. Gilbert, G. S. Kedziora, V. A. Rassolov, D. R. Maurice, N. Nair, Y. Shao, N. A. Besley, P. E. Maslen, J. P. Dombroski, H. Daschel, W. Zhang, P. P. Korambath, J. Baker, E. F. C. Byrd, T. Van Voorhis, M. Oumi, S. Hirata, C. Hsu, N. Ishikawa, J. Florian, A. Warshel, B. G. Johnson, P. M. W. Gill, M. Head-Gordon and J. A. Pople, *J. Comput. Chem.*, 2000, **21**, 1532–1548.

53 C. Y. Legault, *CYLview, 1.0b, Univ. Sherbrooke, 2009 (http//www.cylview.org).*

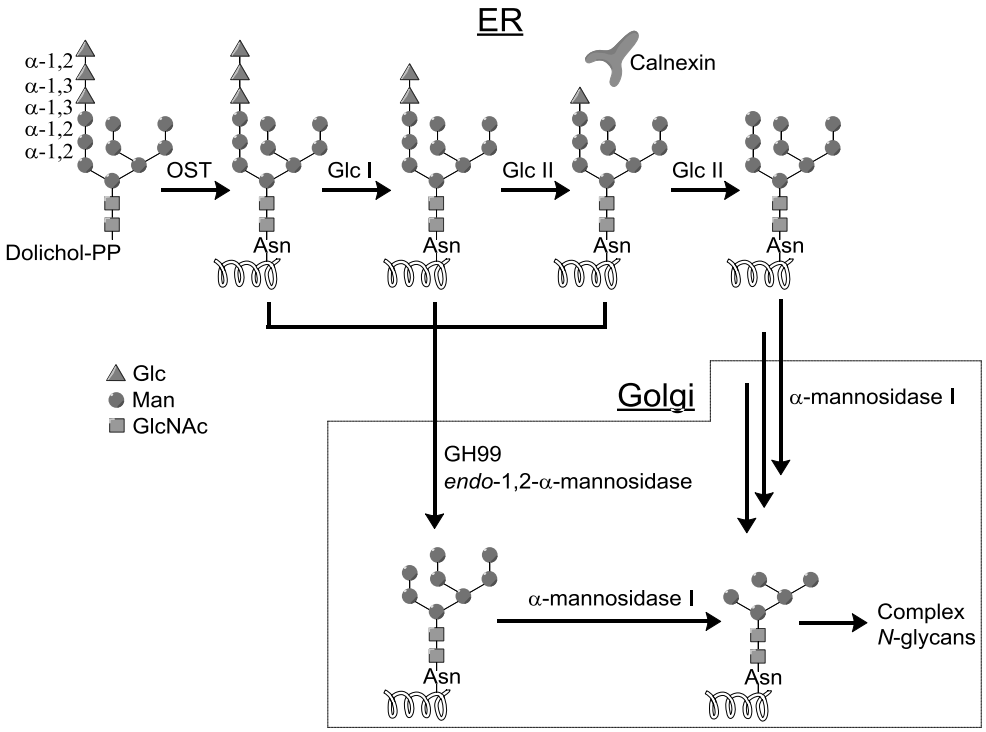
Chapter 6

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]

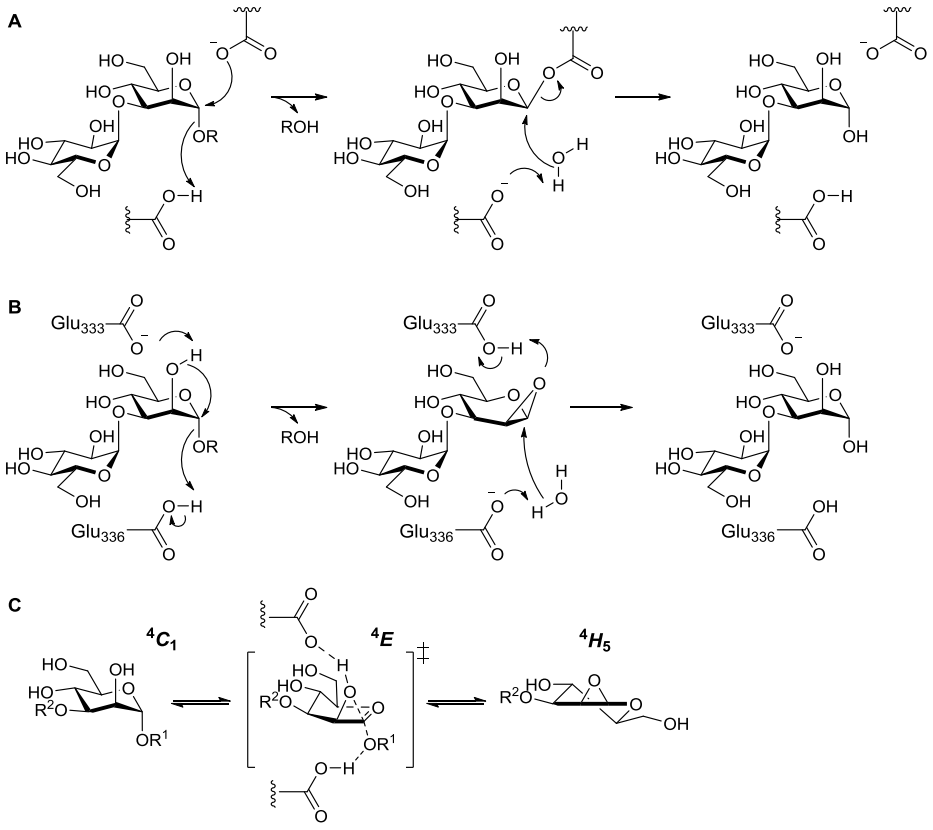


Figure 1 A) Classical Koshland reaction mechanism *via* a covalent enzyme-inhibitor intermediate; B) Reaction mechanism for GH99 *endo*-α-1,2-mannosidase proposed by Thompson *et al;*[16] C) Conformational itinerary for the substrate hydrolysis by GH99 *endo*-α-1,2-mannosidase.

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]

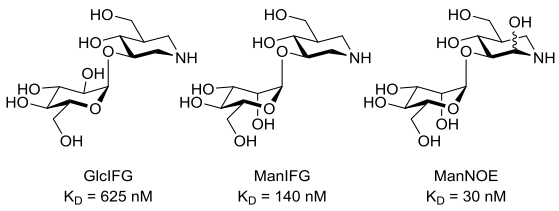


Figure 2 Isofagomine disaccharide analogues α-glucopyranosyl-1,3-isofagomine (GlcIFG), α-mannopyranosyl-1,3-isofagomine (ManIFG) and mannoeuromycin (ManNOE) are potent inhibitors of bacterial GH99 *endo*-α-mannanases. A mannopyranosyl residue at the non-reducing end is favored for *Bt*GH99. Introduction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

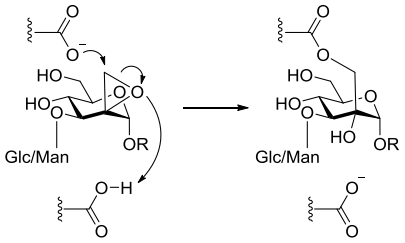
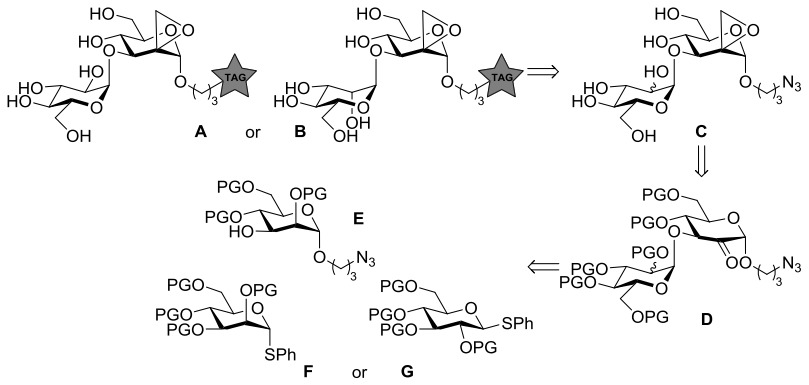


Figure 3 General representation of the target compounds described in this Chapter, including a schematic representation of the covalent binding of the inhibitor to the enzyme based on the proposed reaction mechanism. The spiro-epoxyglycosides contain either an α-glucopyranosyl or an α-mannopyranosyl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

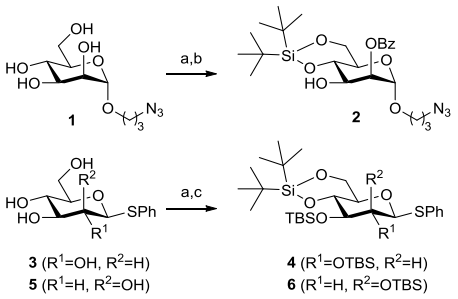
It was envisioned that compounds **A** and **B** could be obtained from matching azides **C** via click-chemistry (Scheme 2). The spiro-epoxide could be installed via a Corey-Chaykovsky epoxidation of ketones **D**, which in turn could be obtained by coupling of acceptor **E** and donor **F** or **G** 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.

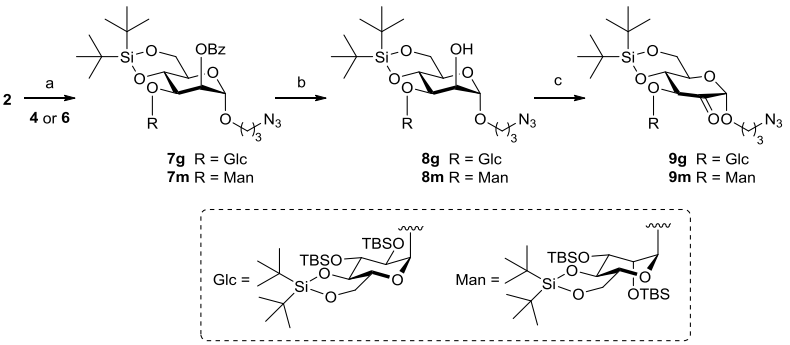
6.2 Results and Discussion

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.

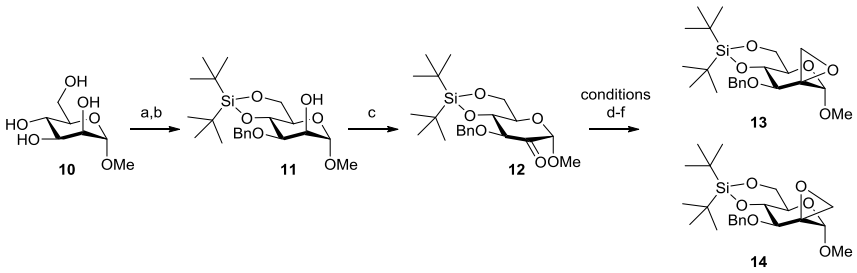
Acceptor **2** was coupled to **4** or **6** in a *N*-iodosuccinimide (NIS)/trimethylsilyl triflate (TMSOTf) mediated coupling at low temperature. Both glycosylations proceeded in excellent yield and stereoselectivity. Pedersen and Bols and co-workers[25] recently reported that silylidene protected mannosyl 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

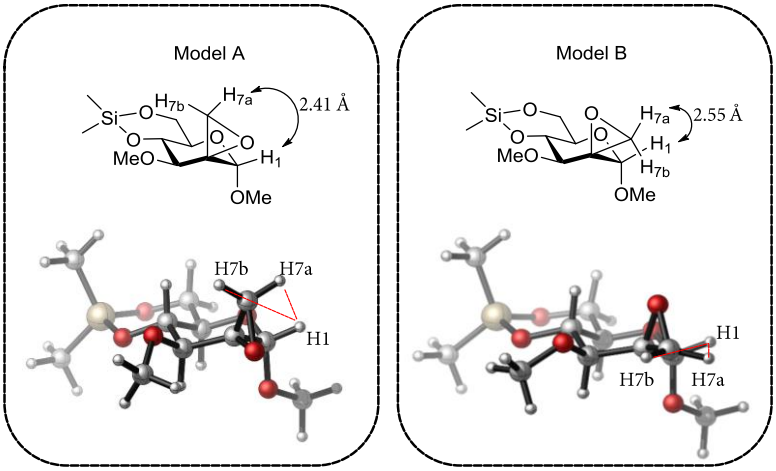
(Scheme 5). α-D-Methyl mannopyranoside was 4,6-silylidene protected and subsequently regioselectively benzylated at position 3 using organotin chemistry,[28] affording compound **11**. The remaining alcohol was oxidized under Swern conditions, yielding ketone **12**. The transformation of ketones into 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.

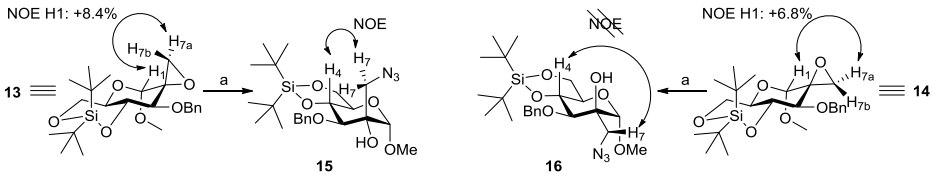


Scheme 5 Synthesis of test substrate **12** and its transformation into its corresponding spiro-epoxides. Reagents and conditions: a) $(^tBu)_2Si(OTf)_2$, 2,6-lutidine, DMF, rt, 30 min; b) Bu_2SnO, toluene, reflux, then BnBr, TBABr, CsF, 84% over 2 steps; c) DMSO, Ac_2O, rt, 61%; d) CH_2N_2, EtOH, 0°C, 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**.



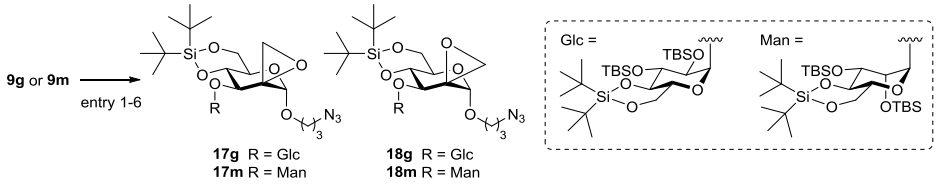


Scheme 6 Atomic distances for H1-H7a in model compounds A and B calculated by DFT (top). Alternative representations of **13** and **14**, and its corresponding NOE enhancements measured by 1D-NOEdif NMR experiments (bottom). After opening the spiro epoxides with sodium azide, a NOE correlation between 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.

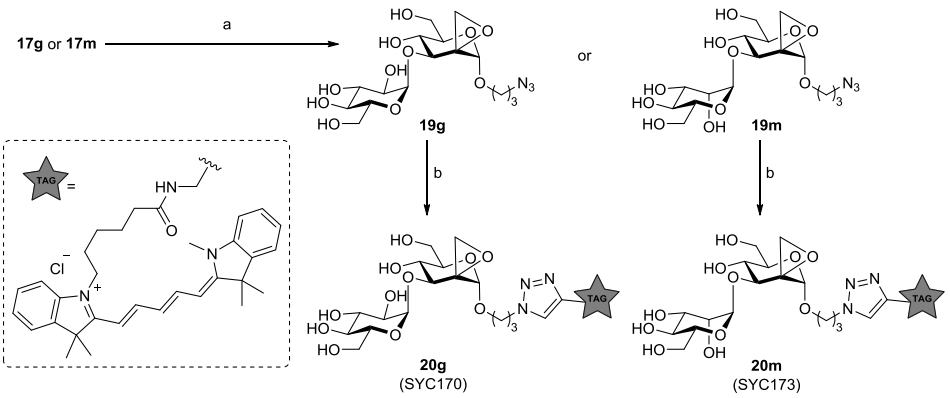


Entry	Starting material	Conditions	**17** : **18**	Yield[a]
1	**9g**	CH_2N_2, EtOH, 0 °C	1 : 1	78%
2	**9m**		1 : 3	97%
3	**9g**	$SOMe_3I$, *n*-BuLi, THF, 60 °C	5 : 1	83%
4	**9m**		2 : 1	85%
5	**9g**	SMe_3I, NaH, DMSO, THF, -10 °C	1 : 0	50%
6	**9m**		1 : 0	53%

[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).

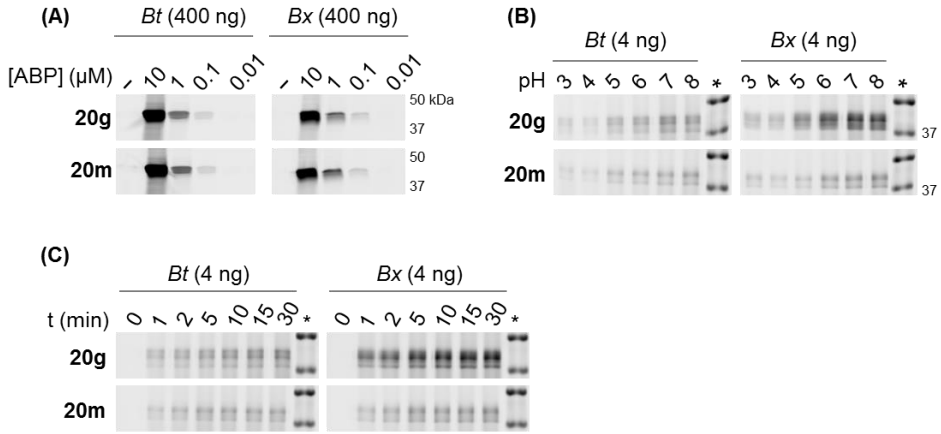


Figure 4 A) Detection limit of *Bt*- and *Bx*GH99 *endo*-α-1,2-mannanases, labeled with fluorescent spiro-epoxyglycosides **20g** or **20m**. B) Effect of pH on labeling of *Bt* and *Bx* enzymes. C) Time-dependent labeling of *Bt* and *Bx*.

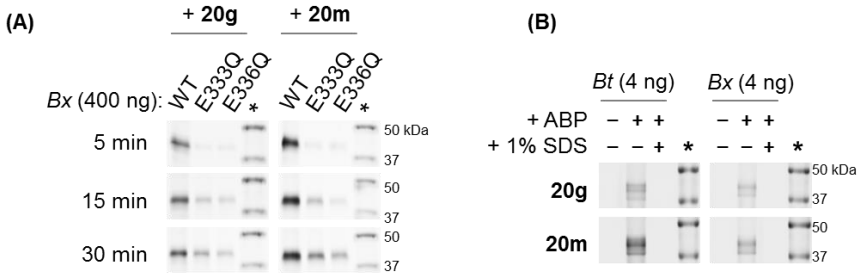


Figure 5 A) Time-dependent labeling of wild-type and mutant *Bx* enzymes with **20g** and **20m**. E333 is the (proposed) catalytic base, E336 is the (proposed) catalytic acid. B) Boiling of *Bt* and *Bx* with 1% SDS prior to incubation with **20g** or **20m** abrogates labeling.

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

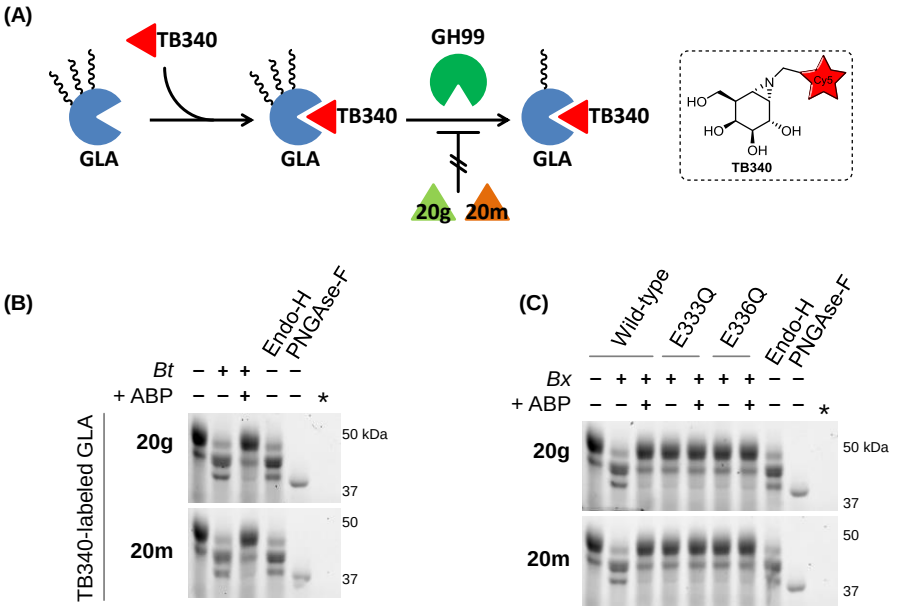


Figure 6 (A) Schematic representation of processing of human alpha-galactosidase GLA by GH99 endomannosidase. GLA is pre-labeled by fluorescent TB340, and contains high-mannose *N*-glycans which can be truncated by endomannosidase, resulting in a decrease in GLA molecular weight. Activity-based labeling 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.

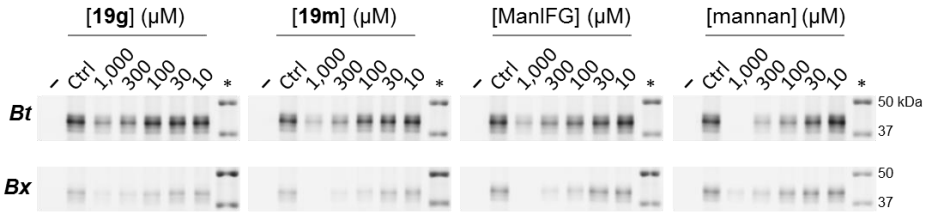


Figure 7 Competitive activity-based protein profiling. *Bt* (top) and *Bx* (bottom) enzymes were incubated with different concentrations of inhibitor followed by labeling with **20m**. For both enzymes, a slight preference for **19m** over **19g** is observed. ManIFG is able to fully compete labeling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

6.3 Conclusion

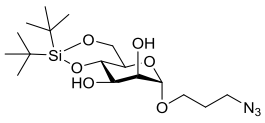
N-glycosylation is an important process in eukaryotic cells that regulates protein folding, targeting and function. The first steps in glycan maturation are performed by stepwise trimming of the terminal glucose residues from the *N*-linked branched oligosaccharide $Glc_3Man_9GlcNAc_3$-Asn by α-glucosidase 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]

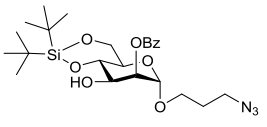
Compound S1



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

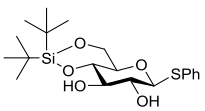
the crude product was purified by gradient column chromatography (EtOAc/pentane, 1:4 to 1:2). The product was obtained as white solid (970 mg, 70%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 4.81 (d, *J* = 1.4 Hz, 1H), 4.11 (dd, *J* = 10.0, 5.0 Hz, 1H), 4.07 – 4.00 (m, 2H), 3.96 (t, *J* = 10.2 Hz, 1H), 3.86 – 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.

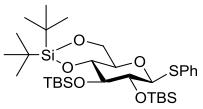
Compound S2



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

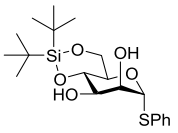
132.9, 131.7, 129.1, 128.3, 88.6, 77.8, 76.4, 74.5, 71.8, 66.1, 27.4, 27.0, 22.7, 19.9. IR (neat): υ 3241, 2932, 2858, 1695, 1471, 1058. $[\alpha]_D^{20}$ (c 0.06, DCM): -57.0. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{20}H_{32}O_5SSiNa$ 435.16319, found 435.16315.

Compound 4



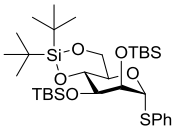
Compound **S2** (1.0 g, 2.42 mmol) was co-evaporated with toluene (3x), dissolved in dry pyridine (5 mL) and cooled to 0°C. DMAP (30 mg, 0.24 mmol) and TBSOTf (3.33 mL, 14.5 mmol) were added and the mixture was heated to 60°C and stirred overnight. The mixture was carefully diluted with water (25 mL) 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



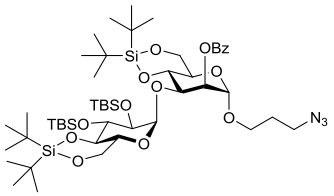
This compound was prepared from **5**[24] (4.9 g, 18 mmol) as described for the preparation of **S2** to afford the product (6.6 g, 89%) as a white solid. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.52 –7.20 (m, 5H), 5.53 (s, 1H), 4.30 (d, *J* = 3.1 Hz, 1H), 4.24 (td, *J* = 10.0, 5.0 Hz, 1H), 4.15 – 4.08 (t, *J* = 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



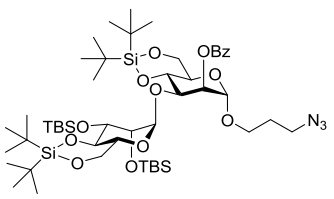
This compound was prepared from **S3** (6.6 g, 16 mmol) as described for the preparation of **4** to afford the product (9.3 g, 91%) as a pale yellow oil which crystallized at -20 °C. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.48 – 7.25 (m, 5H), 5.29 (d, *J* = 1.5 Hz, 1H), 4.28 (t, *J* = 9.0 Hz, 1H), 4.19 (m, 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.

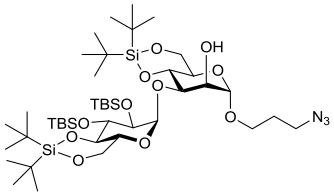
Compound 7m



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.

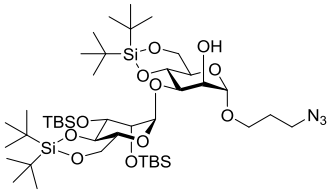
$[\alpha]_D^{20}$ (c 0.4, DCM): +1. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{50}H_{92}N_3O_{12}Si_4$ 1038.57526, found 1038.57587.

Compound 8g



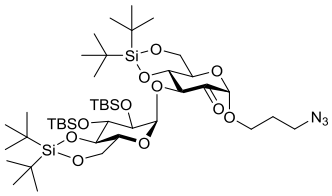
Compound **7g** (610 mg, 0.59 mmol) was co-evaporated with toluene (3x) and dissolved in a mixture of DCM/MeOH (9 mL, 1:1). NaOMe (30 wt%, 560 µL) was added and the reaction mixture was stirred for 24 h. The reaction was neutralized with AcOH and the solvents were removed under reduced pressure. The 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



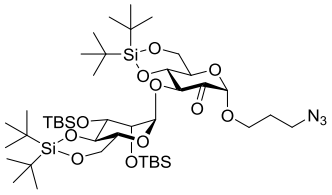
This compound was prepared from **7m** (501 mg, 0.48 mmol) as described for the preparation of **8g** to afford the product (386 mg, 86%) as a colorless oil. ^{1}H-NMR (400 MHz, $CDCl_3$) δ 5.00 (d, *J* = 1.9 Hz, 1H), 4.79 (d, *J* = 1.1 Hz, 1H), 4.17 (t, *J* = 9.2 Hz, 1H), 4.11 (m, 3H), 3.99 – 3.88 (m, 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



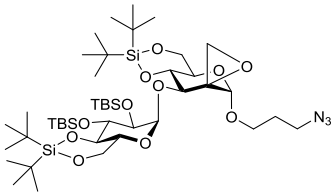
Compound **8g** (2.20 g, 2.36 mmol) was co-evaporated with dry toluene (3x) and dissolved in dry CH_2Cl_2 (65 mL). Dess-Martin periodinane (2.00 g, 4.71 mmol) was added and the mixture was stirred overnight. Celite was added and the solvents were removed under reduced pressure. The product was purified 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.

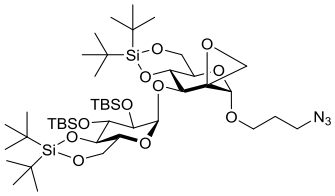
Compound 9m



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.

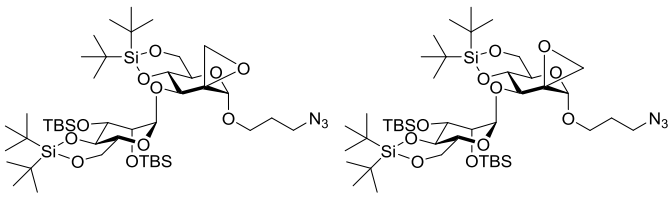
Method B – *dimethyl sulfoxonium methylide*
Trimethylsulfoxonium iodide (37.8 mg, 0.172 mmol) was suspended in dry THF (2 mL) and cooled to 0°C. *n*-Butyllithium (2 M in pentane, 80 µL, 0.16 mmol) was added and the mixture was heated to 60°C. Compound **9g** (100 mg, 0.11 mmol) was co-evaporated with 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)

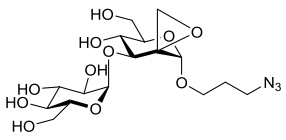
Method C – *Diazomethane*

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

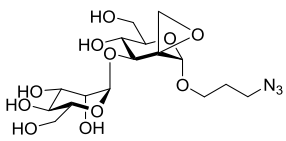
(s, 3H), 0.11 (s, 3H), 0.10 (s, 3H), 0.00 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 102.2, 101.9, 78.2, 74.4, 73.4, 72.1, 70.8, 69.7, 68.2, 67.3, 66.5, 64.2, 58.9, 48.2, 46.6, 28.8, 27.7, 27.5, 27.01, 26.2, 25.6, 22.7, 22.6, 20.0, 19.9, 18.4, 18.1, -4.0, -4.2, -4.2, -4.8. IR (neat): υ 2929, 2098, 1741, 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



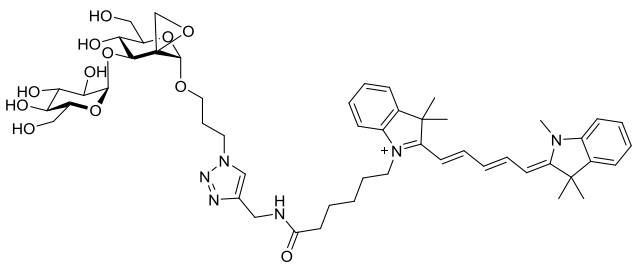
Compound **17g** (145 mg, 0.153 mmol) was co-evaporated with toluene (3x) and dissolved in dry THF (14.5 mL). TBAF (1 M in THF, 2.3 mL, 2.3 mmol) was added and the mixture was stirred overnight at room temperature. The solution was eluted with THF over a small Dowex-50WX4-200-Na^+ packed column, concentrated 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.

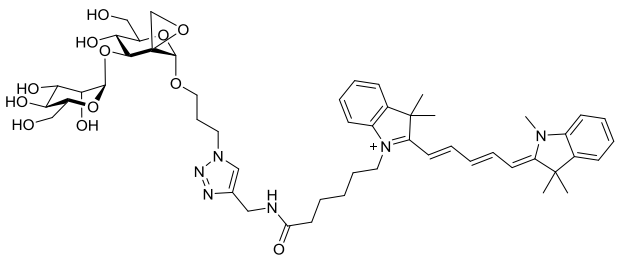
Compound 20g (SYC170)



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

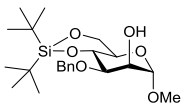
afford the title compound as a blue solid (3.54 mg, 32%). ^{1}H-NMR (400 MHz, MeOD) δ 8.24 (t, *J* = 13.0 Hz, 2H), 7.89 (s, 1H), 7.49 (d, *J* = 7.4 Hz, 2H), 7.44 – 7.38 (m, 2H), 7.32 – 7.23 (m, 4H), 6.62 (t, *J* = 12.4 Hz, 1H), 6.28 (d, *J* = 13.7 Hz, 2H), 5.14 (d, *J* = 3.8 Hz, 1H), 4.85 (s, 1H), 4.53 (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.

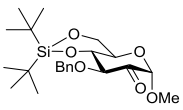
Compound 11



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

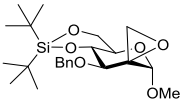
mmol) was added and the mixture was refluxed overnight under argon. The mixture was cooled to room temperature, and tetrabutylammonium bromide (10.3 g, 32.0 mmol), cesium fluoride (4.7 g, 31.1 mmol) and benzyl bromide (3.8 mL, 31.7 mmol) were added. The mixture was refluxed for 3 h and subsequently quenched 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.

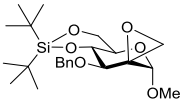
Compound 12



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*

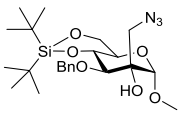
Compound **12** (100 mg, 0.24 mmol) was subjected to the same reaction conditions as for the preparation of compound **17g**, and a mixture of compounds **13** and **14** (64 mg, ratio **13**:**14** 2.7:1, 62%) was obtained.

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.

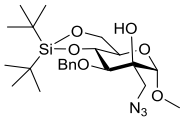
Data for compound **13** (axial methylene): ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.43 – 7.27 (m, 5H), 4.76 (d, *J* = 11.9 Hz, 1H), 4.74 – 4.69 (d, *J* = 11.9 Hz, 1H), 4.22 – 4.19 (s, 1H), 4.14 (m, 1H), 4.05 – 3.98 (m, 1H), 3.98 – 3.85 (m, 3H), 3.41 (s, 3H), 3.24 (d, *J* = 5.4 Hz, 1H), 2.65 (d, *J* = 5.4 Hz, 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.

Chapter 7

Synthesis of *xylobiose*-cyclophellitols for activity-based protein profiling of *Aspergillus niger* secretome

7.1 Introduction

With the ever-growing world population, pollution, global warming and depletion of fossil fuels, the development of technologies to provide natural, sustainable and renewable sources of food stock, energy and other utilities has become of major importance to the scientific community and industry.[1–4] Plant biomass is the most abundant carbon source on earth and holds great potential for the production of renewable resources.[5] Anually, 10^{11} tons of plant biomass is degraded by microbes, corresponding to the energy equivalent of 640 billion barrels of crude oil.[6,7] In perspective, the global crude oil consumption in 2017 was estimated at 36 billion

barrels of crude oil.[8] Plant biomass mainly consists of lignocellulose, complemented by starch, chitin, lipids and proteins.[9] Lignocellulose is the main constituent of plant cell walls and consists of cellulose, hemicellulose and lignin. While lignin is an heterogeneous aromatic biopolymer, cellulose and hemicellulose are polysaccharides and are therefore of high interest for the production of biofuels, such as ethanol, via sugar fermentation.[9] Cellulose consists mostly of linear β-1,4-linked glucose residues which can span up to 15.000 linked monosaccharides. In contrast, hemicellulose polymers are shorter, spanning up to 200 monosaccharides, and are highly branched. Hemicellulose is a heterogeneous polymer, and monosaccharide compositions vary between plant species and classes, notably between softwoods and hardwoods.[10] Overall, xylan is the most abundant polymer in hemicellulose, and is estimated to represent one third of all renewable carbon on earth.[11] Xylan is a β-1,4-xylose polymer which is branched with α-D-glucuronopyranose, 4-*O*-methyl-α-D-glucuronopyranose and α-L-arabinofuranose residues which in turn may contain acetate, ferulate and coumarate esters (Figure 1). Xylan is biosynthetically degraded by hemicellulases[12]; a cooperative set of enzymes including α-D-glucuronidases, α-L-arabinofuranosidases, (acetylxylan-, feruloyl-, coumaroyl)esterases, *exo*-β-xylosidases and *endo*-β-xylanases. Microorganisms that produce hemicellulases include yeasts, fungi and bacteria. Enzymatic expression levels are generally the highest for filamentous fungi, such as *Aspergillus niger* and *Trichoderma reesei*.[13]

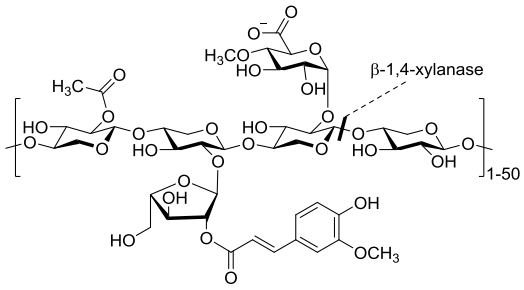


Figure 1 Chemical structure of hemicellulosic xylan. The polysaccharide contains up to 200 carbohydrates and is comprised of a β-1,4-xylan backbone, decorated with α-D-glucuronic acids and α-L-arabinofuranosyl residues carrying acetate, ferulate or coumarate esters. *Endo*-1,4-β-xylanases hydrolyze the xylan backbone at internal positions.

The most important hemicellulases for the biodegradation of xylan are *endo*-β-1,4-xylanases, which cleave the xylan backbone internally, affording two shorter xylan polymers. *Endo*-β-1,4-xylanases are classified in glycoside hydrolase (GH) families,

and members are present in families 5, 7, 8, 10, 11, 30, 43, 51 and 98.[14] GH 10 and 11 glycohydrolases have been most extensively studied, and apart from their similar catalytic activity also share dissimilarities[15]; GH10 members are capable of hydrolyzing branched xylan polymers, have four to five binding subsites (-1 to -5) and also display cellulase activity. In contrast, GH11 members are only active towards non-substituted xylan polymers, contain up to 7 binding subsites and are unreactive towards cellulose. Additionally, while GH10 xylanases are *anti*-protonators, the catalytic acid/base in GH11 members is positioned *syn* with respect to their substrate (see Chapter 1).[16]

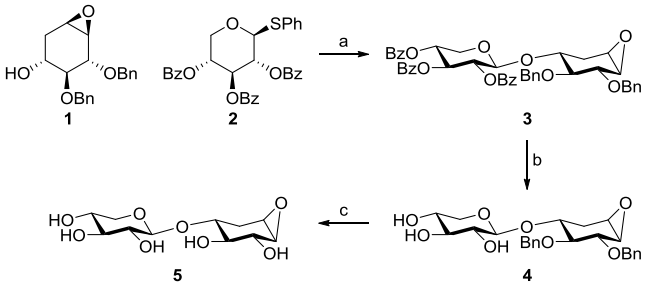
Endo-β-1,4-xylanases serve as catalysts in many industrial applications including food processing (fruit juicing, brewing, baking), textile production, the production of detergents, antioxidants and surfactants.[17] Additionally, a major application of these enzymes lies in the paper and pulp industry, where they are used as biobleaching agents.[18] In this process organic pulp is traditionally heated to elevated temperatures of around 170 °C and subsequently treated with chlorine to remove lignin. This process is costly and produces hazardous chlorinated waste products. *Endo*-1,4-β-xylanases offer a more economically feasible and ecofriendly process to remove lignin, as they break down the hemicellulose crosslinks in lignocellulose, increasing the effectiveness of the chlorination process and thereby decreasing the amount of chemicals required. Also, xylan is branched with 4-*O*-methyl-α-D-glucuronopyranosyl monosaccharides, which are converted to hexenuronic acid during heat treatment. The formation of hexenuronic acid results in undesired paper 'yellowing'. Removal of hemicellulose by xylanases thus enables biobleaching. Xylanases also have a great potential in the production of biofuels, where they are used in the saccharification (hydrolysis of polymeric material to monomeric carbohydrates) of lignocellulose with the aid of *exo*-glycosidases.[17] The resulting monomeric sugar mass is then fermented to produce ethanol, which can be distilled and used as combustible. A major obstacle in the use of enzymes in this process is the toxicity of ethanol towards the microbiota that produce the enzymes.[19] Also, many industrial processes for which the use of xylanases would be desirable require acidic or alkaline pH, or elevated reaction temperatures for optimal production efficiency.[15] Additionally, processes in the food industry sometimes require efficient enzymatic activity at low temperatures to prevent microbial contaminations and food spoiling. Thus, the availability of thermophilic, psychrophilic (cold-adapted), acidophilic and alkaliphilic enzymes is of high importance to the bioindustry.

Many industrially relevant extremophilic xylanases and other enzymes have been identified and isolated from microorganisms that have adapted to extreme environmental conditions.[20] Microorganisms that produce these enzymes could be identified by screening (meta)genomic libraries. Additionally, enzyme properties could be altered by protein engineering. Factors that improve enzymatic activity and/or prevent enzymatic activity loss due to denaturation by extreme conditions have been studied extensively, and include increased hydrogen bonding, disulfide bridges, hydrophobic interactions, proline content and N- or C- terminal stabilization.[18] Indeed, (site directed) mutagenesis has successfully increased the activity and stability of multiple industrially relevant xylanases.[21–24] Fluorogenic assays are traditionally used to determine enzymatic activity/stability, but readouts can easily be misinterpreted by contaminants (such as other enzyme traces) present in the assay. Activity-based protein profiling would potentially represent an alternative way to probe and visualize xylanase activity. So far, several activity-based probes (ABPs) based on the cyclophellitol scaffold have been developed.[25–30] However, due to the monomeric structure of these probes, they are generally specific towards *exo*-acting glycosidases. For the study of *endo*-acting xylanases, a probe that mimics the polymeric xylan backbone is required. In this Chapter, the synthesis of such an ABP is disclosed. The probe is based on the monomeric D-*xylo*-cyclophellitol aziridine scaffold (Chapter 2), which is elongated with a xylose residue at the non-reducing terminus to ensure substrate recognition by the enzyme. Additionally, the fluorescent labeling of this ABP is studied on the industrially relevant secretome of *Aspergillus niger*.

7.2 Results and Discussion

The most straightforward route towards the assembly of multimeric glycosylated cyclophellitol derivatives would be the direct glycosylation of a cyclophellitol (epoxide or aziridine) acceptor with a (poly)saccharide donor. Thus, for the synthesis of the β-1,4-*xylobiose*-cyclophellitol inhibitor, D-*xylo*-cyclophellitol **1** (Chapter 2) containing the electrophilic epoxide warhead was chosen as glycosylation acceptor (Scheme 1). To ensure selective β-xylosylation, thiophenyl donor **2**[31] was selected which bears a participating β-directing benzoyl group at O-2. Whereas it is known that epoxides are prone to undergo Lewis acid-mediated ring opening by trimethylsilyl trifluoromethanesulfonate (TMSOTf),[32] glycosylation of acceptor **1** with donor **2** employing a catalytic amount of TMSOTf resulted in the formation of 'disaccharide' **3** in good yield. Subsequently, the benzoyl groups were removed by Zemplén

debenzoylation to afford **4**. Lastly, the benzyl groups were removed by Pearlman's catalyst under hydrogen atmosphere to afford **5** in quantitative yield. Of note, oxiranes are also known to undergo (palladium catalyzed) reductive opening of the strained three-membered ring,[33] however the epoxide functionality remained unaffected employing an excess of Pearlman's catalyst and short (1-4 h) reaction times.

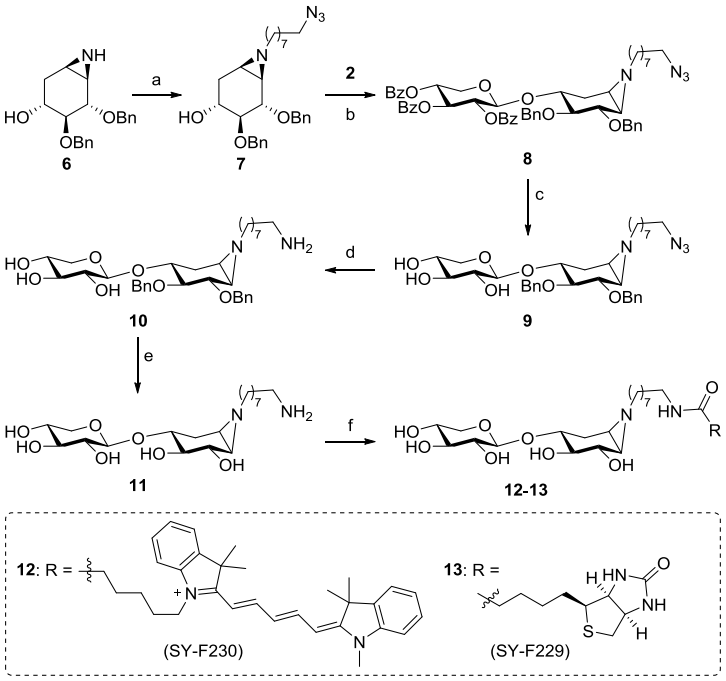


Scheme 1 Synthesis of xylobiose epoxide **5** by glycosylation of epoxide acceptor **1**. Reagents and conditions: a) NIS, TMSOTf (10 mol%), 4Å MS, DCM, -40 °C, 1 h, 75%; b) NaOMe, MeOH, DCM, 16 h, rt, 87%; c) $Pd(OH)_2/C$, H_2, MeOH, H_2O, dioxane, 4 h, rt, quant.

Based on these results, it was investigated whether glycosylation of a donor containing an aziridine functionality was feasible as well. Therefore, aziridine **6** (Chapter 2) was 'protected' by selective *N*-alkylation with 8-azidooctyl trifluoromethanesulfonate resulting in *N*-alkyl-aziridine acceptor **7** (Scheme 2). Indeed, Lewis-acid catalyzed glycosylation with donor **2**[31] resulted in the formation of 'disaccharide' **8** in good yield. Remarkably, a slight excess of TMSOTf appeared to be required to initiate the glycosylation. Subsequently, the benzoyl groups were removed under Zemplén conditions, affording **9**. The azide group was reduced by Staudinger reduction resulting in **10**, which was deprotected with Birch debenzylation conditions, affording **11**. Amide coupling of the primary amine with the corresponding succinate esters (Cy5-OSu[34] and biotin-OSu[35]) followed by HPLC purification resulted in the isolation of *xylobiose*-cyclophellitol ABPs **12** and **13**.

With the *xylobiose*-cylophellitols in hand, the labeling potency of fluorescent aziridine **12** was investigated in a complex biological setting. *Aspergillus niger* is an industrially relevant fungus[36] of which several enzymes have been genetically engineered to provide thermostable enzymes with high turnover.[24,37] Bacterial and fungal strains including *A. niger* are able to excrete a palette of (degradative) enzymes, termed the secretome, into their extracellular environment.[38,39] These proteins are distinguished

by an *N*-terminal signal peptide and are transported across the cell membrane via the ER/Golgi pathway.[40] The presence of signal peptides can be used to predict the variety of excreted proteins through genomics, however it has been shown that such estimations may only cover ~50% of the actual secreted proteins.[41] Additionally, the consistency of secreted enzymes depends on the extracellular medium on which the microbe grows.



Scheme 2 Chemical glycosylation of aziridine acceptor **6** with donor **7** resulted in 'disaccharide' **8** which was further elaborated into ABPs **12**-**13**. Reagents and conditions: a) 8-azidooctyl trifluoromethanesulfonate, DIPEA, DCM, rt, 16 h, 86%; b) NIS, TMSOTf (1.4 eq.), 4Å MS, DCM, -40 °C, 4 h, 77%. c) NaOMe, MeOH, DCM, rt, 16 h, 87%; d) polymer-bound triphenylphosphine, H_2O, MeCN, 70 °C, 16 h, 93%; e) Li, NH_3, THF, 1 h, 85%; f) biotin-OSu or Cy5-OSu, DIPEA, DMF, 16 h, rt, yield **12**: 22%, yield **13**: 25%.

For this study, *A. niger* was grown on Beechwood xylan (1%) as the sole carbon source, and after 8 days the secretome was harvested and concentrated (2.5x). Following incubation of the sample (37 °C, 30 min, pH 5) with a high concentration (5 µM) of ABP **12**, two strong bands (~135 kDa and ~30 kDa) and one weak band (~45 kDa) were visualized on gel (Figure 2a). Lowering the concentration of ABP to 1 µM

abrogated labeling of the ~45 kDa band and reduced labeling of the ~130 kDa band, while at 0.5 µM, only the ~30 kDa band was observed.

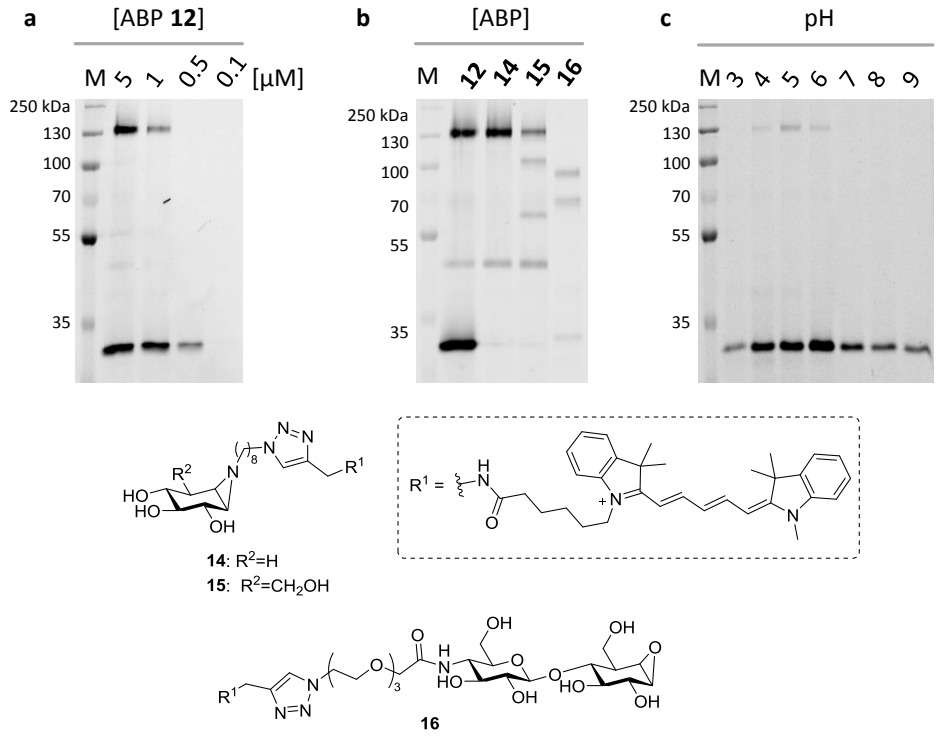
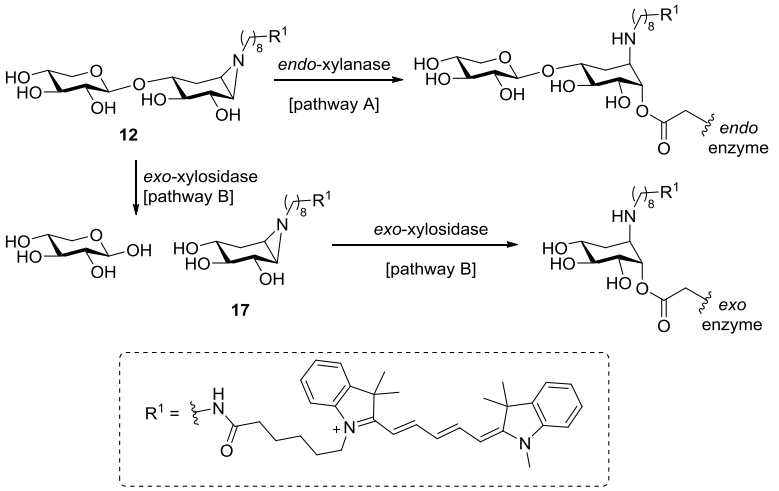


Figure 2 (a) Fluorescent labeling of *A. niger* secretome grown on xylan (1%, 8 days) with different concentrations of *xylobiose*-cyclophellitol ABP **12**. (b) Labeling of this secretome with **14** (5 µM) suggests that the bands at ~130 kDa and ~45 kDa are possibly *exo*-xylosidases. These bands are also labeled with **15** (5 µM), but with reduced efficiency. Probe **16** possibly identified low levels of three different cellulases in this secretome. (c) The optimal pH for labeling the ~30 kDa xylanase was determined by labeling the secretome at different buffer pH's, and is approximately pH 6.

The glycoside hydrolase composition of this secretome was further examined by labeling with *exo*-xylosidase ABP **14**, (Chapter 2), *exo*-glucosidase ABP **15**[42], and *endo*-β-1,4-glucanase ABP **16** (unpublished). At 5 µM probe the bands at ~130 kDa and ~45 kDa (but not ~30 kDa) are labeled by *exo*-xylosidase ABP **14** as well (Figure 2b), suggesting that these bands are in fact *exo*-xylosidases. These bands are also labeled with **15** in reduced potency. Probe **16** visualized weak bands at ~90, 70 and 33 kDa, suggesting that the secretome contains low levels of cellulases as well. The band at ~30 kDa was only labeled with *endo*-probe **12**, indicating that this band corresponds

to a low-molecular weight *endo*-xylanase. Indeed, it is known that growing *A. niger* on xylan activates the transcriptional regulator XlnR[43] which induces secretion of GH10 xylanase XlnC (~34 kDa)[44]. Interestingly, XlnR activation also induces GH11 xylanase XlnB (~24 kDa)[44] secretion, however a band at this molecular weight was not found. Next, the pH influence on labeling of the ~30 kDa xylanase band by **12** was investigated by pre-incubation of the secretome at different buffer pH's followed by labeling (1 μM), and pH 6.0 appeared optimal (Figure 2c).

Labeling of all three bands with *endo*-ABP **12** could be explained by its inherent instability of the glycosidic linkage towards glycosidic hydrolysis by *exo*-acting enzymes. It is hypothesized that during incubation, **12** partially undergoes *exo*-hydrolysis affording D-xylose and *exo*-xylose ABP **17** (which is structurally highly similar to **14**), which concomitantly labels the *exo*-xylosidase (Scheme 3, pathway B).



Scheme 3 Labeling of the secretome with *xylobiose*-cyclophellitol aziridine **12** results in the activity-based labeling of *endo*-xylanase (pathway A). However, the glycosidic bond in **12** may be cleaved *in situ* by *exo*-xylosidases resulting in D-xylose and **17**. Due to the monomeric structure of **17**, this compound is a probe for *exo*-xylosidases (pathway B).

To further examine the bands labeled by **12**, competitive ABPP experiments were performed (Figure 3). Pre-incubation with the β-glucosidase inactivator cyclophellitol **18** followed by labeling with **12** (5 μM) did not affect the labeling of both bands at any concentration, whereas pre-incubation with **19** potently competed with the band at ~130 kDa, further supporting that this band is indeed an *exo*-xylosidase.

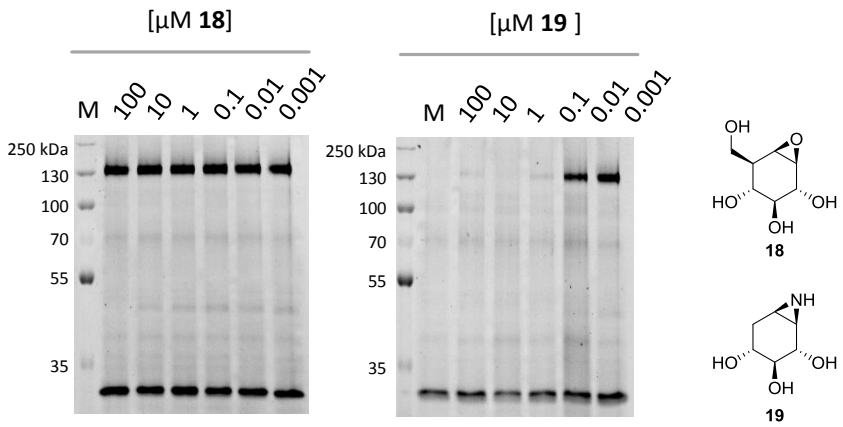


Figure 3 Competition assay of *A. niger* secretome. The sample was pre-incubated with **18** or **19** at different concentrations (30 min, 37 °C, pH 6) and then labeled with **12** (5 μM, 30 min, 37 °C). Pre-incubation with **18** did not decrease labeling with **12** at any concentration, whereas **19** fully competed the ~130 kDa band at concentrations up to 1 μM, indicating that this band is an *exo*-xylosidase.

Competition assays were also performed with *endo*-xylanase inactivators **5**, **11** and **13** (Figure 4). It was found that **11** and **13** are both inactivators of the ~30 kDa xylanase, but high concentration of inhibitor was needed to accomplish full inhibition. Interestingly, the ~130 kDa *exo*-xylosidase was more potently inhibited, indicating that for **11** and **13**, processing by the *exo*-xylosidase (pathway B, Scheme 3) is preferred over *endo*-xylanase inactivation (pathway A, Scheme 3). In contrast, for epoxide **5** *endo*-xylanase deactivation occurred at approximately the same rate as *exo*-xylosidase processing, moreover **5** deactivated *endo*-xylanase (and *exo*-xylosidase) activity with higher potency. This observation is in sharp contrast with the *exo*-xylosidase inactivators described in Chapter 2, for which all aziridine based inhibitors exhibited superior potency over the epoxide.

The applicability of ABP **12** to monitor enzymatic activity over time at different temperatures in biological context is exemplified in Figure 5. For this experiment, *A. niger* secretome was pre-incubated with buffer (pH 6.0) at different temperatures for different time periods (2-60 min), after which the remaining active enzyme was labeled with **12** (5 μM). It was observed that the ~30 kDa *endo*-xylanase retained full activity after pre-incubation of the sample for 60 minutes up to 50 °C. At 60 °C, the enzyme half-life is approximately 15 minutes and nearly all enzymatic activity is lost

after 60 minutes. At 70-80 °C, enzymatic activity is abrogated within 5 minutes. Interestingly, the ~130 kDa *exo*-xylosidase appears to be more stable towards elevated temperatures.

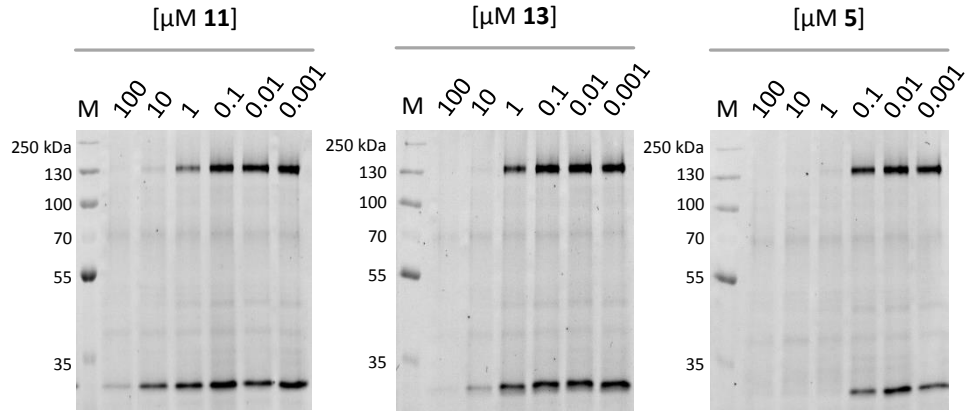


Figure 4 Competition assay of *A. niger* secretome. The sample was pre-incubated with **11**, **13** and **5** at different concentrations (30 min, 37 °C) and then labelled with **12** (5 μM, 30 min, 37 °C, pH 6). Both **11** and **13** are weak inactivators of *endo*-xylanase, moreover the *exo*-xylosidase was inhibited with higher efficiency. Pre-incubation with **5** resulted in inactivation of *endo*-xylanase (and *exo*-xylosidase) with higher potency.

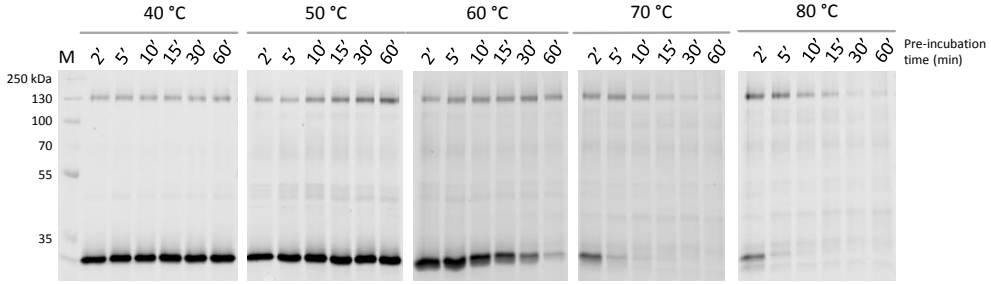


Figure 5 Monitoring of *endo*-xylanase activity over time at different temperatures. The enzyme retains full activity after incubation at 50 °C for 60 minutes. At 60 °C, the enzyme half-life is approximately 15 minutes and nearly full deactivation is reached after 60 minutes. At 70-80°C, nearly all enzyme activity is abrogated within 5 minutes. The ~130 kDa *exo*-xylosidase appears to be more stable towards elevated temperatures.

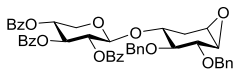
7.3 Conclusion

Up to date, all cyclophellitol (aziridine) based glycosidase ABPs reported in literature are monomeric in structure and therefore could only be utilized for the monitoring of *exo*-glycosidases. While Withers and co-workers have published the chemoenzymatic synthesis of α-1,4-glucosyl *epi*-cyclophellitol as covalent irreversible α-amylase inhibitor,[45] this approach was certainly not trivial and enzymatic tolerance of substrates equipped with a reporter tag might pose a major obstacle for the chemoenzymatic synthesis of *endo*-glycosidase probes. In this Chapter, a new straightforward synthetic route towards glycosylated *xylo*-cyclophellitols is described, which relies on the direct glycosylation of *xylo*-cyclophellitol (aziridine) acceptors with an appropriate thioglycoside donor under NIS/TMSOTf activating conditions. The epoxide and aziridine functionalities were tolerant towards the Lewis acidic reaction conditions, and glycosylation products could be obtained in good yields. Labeling of the secretome of *Aspergillus niger* (grown on xylan) by the *xylobiose*-cyclophellitol aziridine fluorescent probe identified a ~30 kDa band on gel, which is presumed to be the *endo*-xylanase XlnC. The optimal buffer pH for labeling this enzyme was pH 6.0, and the enzyme appeared stable up to 50 °C for 60 minutes; at higher temperatures the activity was abolished after this time period. The probe also identified the presence of a ~130 kDa protein in the secretome, and by competition assays with appropriate *exo*-glycosidase inhibitors, this band is presumed to be an *exo*-xylosidase. It is hypothesized that following incubation with the *xylobiose*-cyclophellitol ABP, the *exo*-xylosidase is labeled by an *exo*-xylosidase probe generated *in situ* by *exo*-hydrolysis of the *xylobiose*-cyclophellitol ABP. Competition assays indicated that *xylobiose*-cyclophellitol alkyl aziridines are weak inactivators of the *endo*-xylanase, and due to the inherent instability of the *endo*-xylanase probe towards hydrolysis, that these inhibitors are processed by the *exo*-xylosidase at higher rate. *Xylobiose*-cyclophellitol (epoxide) is a more potent inactivator of the *endo*-xylanase compared to the alkyl aziridine analogues, in contrast to observations for *exo*-xylosidase inhibitors. Identification of the labeled enzymes may be achieved using *xylobiose*-cyclophellitol aziridine equipped with a biotin tag, via protein enrichment techniques followed by LC-MS/MS identification. Lastly, the instability of *endo*-cyclophellitol probes towards *exo*-hydrolysis complicates the selective profiling of *endo*-glycosidases in biologically complex samples such as presented here, therefore the synthesis of stabilized *endo*-probes would be of interest (see 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).

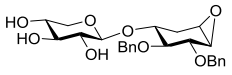
Compound 3



A mixture of **1** (65 mg, 0.2 mmol) and **2**[31] (111 mg, 0.2 mmol) was co-evaporated with toluene (3x) and dissolved in dry DCM (1.3 mL). Activated 4 Å molecular sieves were added and the mixture was stirred for 30 minutes at RT. Then the mixture was cooled to -40 °C, and NIS (54 mg, 0.24 mmol) and TMSOTf (3.6 μL, 10 mol %) were added. After stirring for 1 h, the mixture was quenched with Et_3N, diluted with DCM (30 mL) and aq. 10% $Na_2S_2O_3$ (20 mL) was added. The mixture was warmed to RT and the layers were separated. The organic phase was washed with aq. 10% $Na_2S_2O_3$ (2 x 20 mL) and brine, dried over $MgSO_4$, filtrated and concentrated. Flash column chromatography (pentane:EtOAc 3:1) afforded the product as a white foam (115 mg, 75%). ^1H NMR (400 MHz, $CDCl_3$): 8.01 – 7.90 (m, 6H), 7.57 – 7.24 (m, 19H), 5.77 (t, *J* = 8.3 Hz, 1H), 5.42 (dd, *J* = 8.4, 6.5 Hz, 1H), 5.37 – 5.28 (m, 1H), 4.98 (d, *J* = 10.9 Hz, 1H), 4.85 (d, *J* = 6.5 Hz, 1H), 4.72 (m, 3H), 4.38 (dd, *J* = 12.0, 4.9 Hz, 1H), 3.91 – 3.78 (m, 2H),

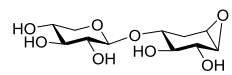
3.50 (dd, *J* = 12.0, 8.5 Hz, 1H), 3.42 (dd, *J* = 10.0, 7.7 Hz, 1H), 3.16 (br s, 1H), 3.09 (d, *J* = 3.6 Hz, 1H), 2.49 (ddd, *J* = 14.5, 5.0, 2.1 Hz, 1H), 1.72 – 1.63 (ddd, *J* = 14.5, 10.5, 1.0 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): 165.8, 165.6, 165.2, 138.8, 137.8, 133.5, 133.5, 133.4, 130.3, 129.9, 129.9, 129.3, 129.2, 129.2, 128.6, 128.6, 128.5, 128.4, 128.3, 128.0, 128.0, 127.7, 99.6, 82.5, 79.6, 75.4, 74.6, 73.5, 71.6, 71.6, 69.7, 62.3, 53.9, 53.5, 30.1 ppm. IR: (neat) 2933, 1722, 1452, 1250, 1087 cm^{-1}. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{46}H_{42}O_{11}$ 771.27999, found 771.28107.

Compound 4



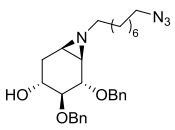
Compound **3** (115 mg, 0.15 mmol) was dissolved in a mixture of dry DCM (0.75 mL) and MeOH (1.5 mL). Then, NaOMe in MeOH (5.4 M, 11 µL, 40 mol %) was added and the mixture was stirred 16h at rt. The mixture was neutralized with Et_3N.HCl (10 mg, 0.075 mmol) and concentrated. Flash column chromatography (DCM:MeOH 20:1) afforded the product as a colorless oil (60 mg, 87%). ^{1}H NMR (400 MHz, $CDCl_3$): 7.36 – 7.17 (m, 10H), 4.77 (d, *J* = 11.0 Hz, 1H), 4.71 (d, *J* = 11.5 Hz, 2H), 4.64 (d, *J* = 11.4 Hz, 1H), 4.56 (br s, OH), 4.33 (d, *J* = 6.8 Hz, 1H), 4.00 (br s, 2x OH), 3.88 (dd, *J* = 11.7, 4.7 Hz, 1H), 3.85 – 3.75 (m, 2H), 3.53 (dd, *J* = 13.4, 8.4 Hz, 1H), 3.43 (m, 2H), 3.35 (t, *J* = 7.6 Hz, 1H), 3.18 (br s, 1H), 3.16 – 3.05 (m, 2H), 2.57 (dd, *J* = 14.3, 2.8 Hz, 1H), 1.78 (dd, *J* = 13.7, 10.7 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): 138.2, 137.6, 128.6, 128.4, 128.3, 128.0, 127.8, 100.9, 81.9, 79.5, 75.4, 74.6, 73.1, 71.9, 71.0, 69.6, 65.1, 53.6, 53.5, 30.3 ppm. IR: (neat) 3396, 2922, 1367, 1037 cm^{-1}. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{25}H_{30}O_8$ 459.20134, found 459.20135.

Compound 5



Compound **4** (45 mg, 0.1 mmol) was dissolved in a mixture of MeOH:H_2O:dioxane (1:1:1, 2.1 mL) under argon and $Pd(OH)_2$/C (20 wt%, 42 mg, 0.3 mmol) was added. While stirring vigorously, the mixture was flushed with a H_2 balloon. After stirring for 4 h under H_2 atmosphere, the mixture was filtered over Celite and evaporated to afford the product in high purity as a colorless oil (28 mg, 100%). ^{1}H NMR (400 MHz, D_2O): 4.39 (d, *J* = 7.9 Hz, 1H), 3.90 (dd, *J* = 11.6, 5.5 Hz, 1H), 3.78 (d, *J* = 8.2 Hz, 1H), 3.63 – 3.54 (m, 2H), 3.48 – 3.43 (m, 1H), 3.42 – 3.35 (m, 2H), 3.27 – 3.19 (m, 2H), 3.17 (d, *J* = 3.7 Hz, 1H), 2.64 (ddd, *J* = 14.8, 5.2, 2.1 Hz, 1H), 1.85 (ddd, *J* = 14.8, 10.4, 1.6 Hz, 1H) ppm. ^{13}C NMR (101 MHz, D_2O): 100.8, 75.6, 74.4, 73.6, 72.7, 71.1, 69.1, 65.1, 56.1, 54.7, 28.7 ppm. IR: (neat) 3259, 2160, 1977, 1379, 1311, 1072, 1024 cm^{-1}. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{11}H_{18}O_8$ 279.10744, found 279.10776.

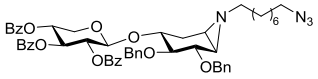
Compound 7



Aziridine **6** (162 mg, 0.50 mmol) was co-evaporated with toluene and dissolved in dry DCM (2 mL). DIPEA (87 µL, 0.50 mmol) was added and subsequently a solution of 1-azido-8-trifluoromethanesulfonyloctane (1M in DCM, 0.55 mL, 0.55 mmol, see Chapter 2 for preparation). After stirring the solution overnight at rt, the reaction was quenched with MeOH (5 mL), the mixture was stirred for additional 2 h and then

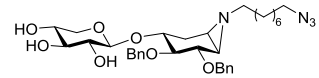
concentrated. Flash purification by silica column chromatography (DCM/MeOH, 99:1) gave the title compound as an oil (204 mg, 86%). ^{1}H NMR (400 MHz, $CDCl_3$): 7.42 – 7.22 (m, 10H), 4.96 (d, *J* = 11.4 Hz, 1H), 4.77 (d, *J* = 11.5 Hz, 1H), 4.65 (dd, *J* = 11.4, 7.3 Hz, 2H), 3.74 (d, *J* = 8.0 Hz, 1H), 3.54 (td, *J* = 10.3, 5.4 Hz, 1H), 3.24 (t, *J* = 6.8 Hz, 2H), 3.22 – 3.17 (m, 1H), 2.54 (s, OH), 2.42 (dd, *J* = 12.9, 5.4 Hz, 1H), 2.27 – 2.13 (m, 2H), 1.71 (dd, *J* = 5.6, 2.5 Hz, 1H), 1.68 – 1.62 (m, 1H), 1.62 – 1.54 (m, 2H), 1.54 – 1.44 (m, 3H), 1.40 – 1.28 (m, 8H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): 138.8, 138.1, 128.6, 128.6, 128.1, 128.0, 127.9, 127.8, 85.3, 81.9, 74.8, 72.2, 66.5, 61.0, 51.6, 41.4, 39.9, 31.8, 29.7, 29.6, 29.2, 28.9, 27.5, 26.8 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{28}H_{38}N_4O_3$ 479.3017, found 479.3022.

Compound 8



Donor **2**[31] (235 mg, 0.42 mmol) and acceptor **7** (145 mg, 0.30 mmol) were combined in a flask, co-evaporated with toluene (3x) and dissolved in dry DCM (3 mL). 4Å molecular sieves were added and the mixture was stirred overnight. The mixture was cooled to -40 °C and NIS (82 mg, 0.36 mmol) and TMSOTf (74 µL, 0.41 mmol) were added. After 4 h the reaction was quenched with Et_3N (200 µL) and 10% aq. $Na_2S_2O_3$ (2 mL) was added. After warming to rt, the mixture was diluted with DCM (100 mL), washed with 10% aq. $Na_2S_2O_3$ (2 x 50 mL) and brine, dried over $MgSO_4$, filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 6:1 → 4:1) gave the title compound as an oil (216 mg, 77%). ^{1}H NMR (400 MHz, $CDCl_3$): 8.00 – 7.92 (m, 6H), 7.57 – 7.44 (m, 3H), 7.42 – 7.22 (m, 16H), 5.74 (t, *J* = 8.2 Hz, 1H), 5.39 (dd, *J* = 8.3, 6.4 Hz, 1H), 5.31 (td, *J* = 8.2, 4.9 Hz, 1H), 4.95 (d, *J* = 11.0 Hz, 1H), 4.86 (d, *J* = 6.3 Hz, 1H), 4.72 (d, *J* = 11.0 Hz, 1H), 4.69 (s, 2H), 4.36 (dd, *J* = 12.0, 4.8 Hz, 1H), 3.82 (td, *J* = 10.3, 5.2 Hz, 1H), 3.74 (d, *J* = 7.6 Hz, 1H), 3.48 (dd, *J* = 12.0, 8.3 Hz, 1H), 3.31 (dd, *J* = 9.9, 7.7 Hz, 1H), 3.25 (t, *J* = 6.9 Hz, 2H), 2.34 – 2.22 (m, 2H), 1.94 (dt, *J* = 11.8, 6.8 Hz, 1H), 1.63 – 1.55 (m, 4H), 1.54 – 1.48 (m, 1H), 1.47 – 1.17 (m, 11H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): 165.8, 165.6, 165.3, 139.2, 138.4, 133.5, 133.4, 133.4, 130.0, 130.0, 129.9, 129.5, 129.3, 129.3, 128.6, 128.5, 128.5, 128.3, 128.2, 128.0, 127.8, 127.5, 99.5, 83.5, 81.5, 76.2, 75.2, 72.9, 71.7, 71.6, 69.7, 62.1, 60.7, 51.6, 41.4, 30.5, 29.7, 29.6, 29.3, 28.9, 27.4, 26.8 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{54}H_{58}N_4O_{10}$ 923.4226, found 923.4243.

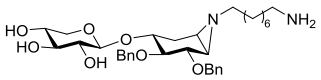
Compound 9



Compound **8** (59 mg, 64 µmol) was dissolved in a mixture of DCM/MeOH (1:2 v/v, 1.2 mL) and NaOMe (5.4 M in MeOH, 6 µL, 32 µmol) was added. After stirring overnight the mixture was neutralized by addition of Et_3N.HCl (5 mg, 35 µmol) and concentrated. Flash purification by silica column chromatography (DCM/MeOH, 97:3 → 94:6) gave the title compound as an oil (34 mg, 87%). ^{1}H NMR (400 MHz, $CDCl_3$): 7.37 – 7.20 (m, 10H), 4.77 (s, 2H), 4.72 (d, *J* = 11.6 Hz, 1H), 4.66 (d, *J* = 11.5 Hz, 1H), 4.45 (d, *J* = 6.3 Hz, 1H), 3.96 (dd, *J* = 11.8, 4.4 Hz, 1H), 3.83 (dd, *J* = 10.2, 5.3 Hz, 1H), 3.79 (d, *J* = 7.6 Hz, 1H), 3.62 – 3.54 (m, 1H), 3.48 (t, *J* = 7.6 Hz, 1H), 3.45 – 3.40 (m, 1H), 3.35 (dd, *J* = 10.0, 7.8 Hz, 1H), 3.24 (t, *J* = 6.9 Hz, 2H), 3.19 (dd, *J* = 11.8, 8.5 Hz, 1H), 2.46 (dd, *J* = 13.1, 5.0 Hz, 1H), 2.33 (dt, *J* =

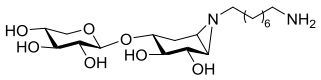
11.4, 7.2 Hz, 1H), 2.00 (dt, *J* = 11.6, 7.1 Hz, 1H), 1.71 (d, *J* = 5.6 Hz, 1H), 1.67 (d, *J* = 13.4 Hz, 1H), 1.58 (dt, *J* = 14.4, 6.8 Hz, 2H), 1.52 (d, *J* = 6.0 Hz, 1H), 1.50 – 1.41 (m, 2H), 1.32 (d, *J* = 23.0 Hz, 8H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): 138.5, 138.1, 128.6, 128.4, 128.1, 128.0, 127.7, 100.4, 82.9, 81.6, 74.8, 74.7, 72.7, 71.8, 71.3, 69.7, 64.7, 60.9, 51.6, 41.3, 39.8, 30.9, 29.6, 29.6, 29.3, 28.9, 27.5, 26.8 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{33}H_{46}N_4O_7$ 611.3439, found 611.3447.

Compound 10



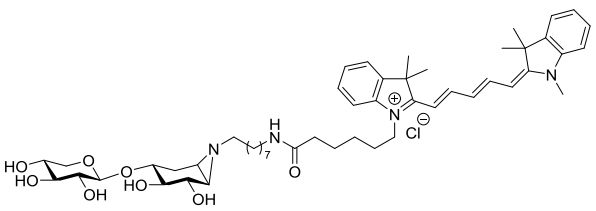
Compound **9** (111 mg, 0.19 mmol) was dissolved in MeCN (3.8 mL), polymer-bound triphenylphosphine (3 mmol/g loading, 126 mg, 0.38 mmol) and H_2O (34 µL, 1.9 mmol) were added and the mixture was stirred overnight at 70 °C. TLC indicated total consumption of the starting material, and additional H_2O (500 µL) was added and the mixture was stirred for 4 h at 70 °C. The mixture was cooled to rt, filtrated, diluted with MeCN (100 mL), dried over $MgSO_4$, filtrated and concentrated, affording the title compound (103 mg, 93%) as an oil. ^{1}H NMR (400 MHz, $CDCl_3$): 7.41 – 7.16 (m, 10H), 4.77 (d, *J* = 3.8 Hz, 2H), 4.72 (d, *J* = 11.7 Hz, 1H), 4.66 (d, *J* = 11.5 Hz, 1H), 4.48 (d, *J* = 6.0 Hz, 1H), 3.99 (dd, *J* = 11.8, 4.1 Hz, 1H), 3.85 (dd, *J* = 10.1, 5.2 Hz, 1H), 3.81 (d, *J* = 7.5 Hz, 1H), 3.58 (q, *J* = 7.5 Hz, 1H), 3.49 (t, *J* = 7.3 Hz, 1H), 3.43 (t, *J* = 6.7 Hz, 1H), 3.34 (dd, *J* = 10.0, 7.6 Hz, 1H), 3.21 (dd, *J* = 11.8, 8.1 Hz, 1H), 3.02 (brs, NH2), 2.65 (t, *J* = 7.0 Hz, 2H), 2.45 (dd, *J* = 13.0, 4.6 Hz, 1H), 2.19 (ddt, *J* = 23.9, 11.5, 5.8 Hz, 2H), 1.69 (d, *J* = 5.3 Hz, 1H), 1.63 (t, *J* = 10.7 Hz, 1H), 1.52 (d, *J* = 5.9 Hz, 1H), 1.49 – 1.38 (m, 4H), 1.30 (s, 8H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): 138.5, 138.2, 128.6, 128.4, 128.1, 128.0, 127.7, 100.6, 83.0, 82.0, 74.7, 74.4, 72.6, 71.6, 71.1, 69.7, 64.7, 60.6, 42.1, 41.0, 39.9, 33.5, 31.1, 29.6, 29.5, 27.4, 26.8 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{33}H_{48}N_2O_7$ 585.3534, found 585.3542.

Compound 11



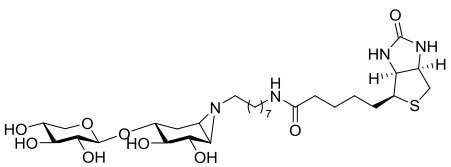
Ammonia (1 mL) was condensed in a flask at -60 °C, and lithium wire (13 mg, 1.86 mmol) was added. The resulting deep-blue solution was stirred for 30 minutes to dissolve all lithium. Aziridine **10** (30 mg, 49 µmol) was taken up in dry THF (1 mL) and added to the reaction mixture. After stirring for 1 h, the mixture was quenched with H_2O. The mixture was slowly warmed to rt and evaporated. The crude was dissolved in H_2O and eluted over a column packed with Amberlite CG-50 (NH_4^+) with 0.5M NH_4OH as eluent, concentrated, and re-purified by HW40 (NH_4HCO_3) affording the title compound as an oil (17.0 mg, 85%). ^{1}H NMR (400 MHz, D_2O): 4.37 (d, *J* = 7.8 Hz, 1H), 3.90 (dd, *J* = 11.6, 5.4 Hz, 1H), 3.66 (d, *J* = 8.5 Hz, 1H), 3.55 (dtt, *J* = 25.5, 10.4, 5.6 Hz, 2H), 3.39 (t, *J* = 9.2 Hz, 1H), 3.31 – 3.18 (m, 3H), 2.94 (t, *J* = 7.6 Hz, 2H), 2.50 (dd, *J* = 13.9, 5.5 Hz, 1H), 2.34 – 2.23 (m, 2H), 2.07 – 2.02 (m, 1H), 1.77 – 1.67 (m, 2H), 1.66 – 1.56 (m, 2H), 1.55 – 1.45 (m, 2H), 1.39 – 1.23 (m, 8H) ppm. ^{13}C NMR (101 MHz, D_2O): 100.9, 75.6, 75.0, 74.9, 72.7, 72.1, 69.1, 65.1, 59.3, 43.1, 39.8, 39.4, 28.7, 28.4, 28.3, 28.0, 26.6, 26.3, 25.4 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{19}H_{37}N_2O_7$ 405.25953, found 405.25943.

Compound 12 (SY-F230)



Compound **11** (4.1 mg, 10 µmol) was dissolved in DMF (0.5 mL), then DIPEA (3.8 µL, 22 µmol) and Cy5-OSu[34] (3.8 mg, 11 µmol) was added and the mixture was stirred overnight at rt. Full conversion was observed by LC-MS and the product was purified by semi-preparative reversed phase HPLC (linear gradient. Solutions used: A: 50 mM NH4HCO3 in H2O, B: acetonitrile), affording the product as a blue solid (2.09 mg, 22%) after lyophilization. 1H NMR (500 MHz, D2O): 7.87 (t, J = 12.8 Hz, 2H), 7.45 (d, J = 5.1 Hz, 2H), 7.35 (q, J = 7.5 Hz, 2H), 7.24 (t, J = 7.4 Hz, 2H), 7.22 – 7.17 (m, 2H), 6.42 (t, J = 12.5 Hz, 1H), 6.11 (t, J = 13.0 Hz, 2H), 4.34 (d, J = 7.8 Hz, 1H), 4.06 (t, J = 6.2 Hz, 2H), 3.89 (dd, J = 11.5, 5.4 Hz, 1H), 3.62 (d, J = 8.4 Hz, 1H), 3.59 (dd, J = 9.3, 5.0 Hz, 1H), 3.56 (s, 3H), 3.48 (td, J = 10.4, 5.5 Hz, 1H), 3.39 (t, J = 9.2 Hz, 1H), 3.28 – 3.19 (m, 3H), 3.00 (t, J = 7.0 Hz, 2H), 2.40 (dd, J = 13.8, 5.4 Hz, 1H), 2.18 (t, J = 6.7 Hz, 2H), 2.12 – 2.04 (m, 1H), 2.03 – 1.96 (m, 1H), 1.84 – 1.75 (m, 3H), 1.69 – 1.57 (m, 4H), 1.54 (s, 6H), 1.53 (s, 6H), 1.36 – 1.25 (m, 6H), 1.17 – 1.05 (m, 8H) ppm. 13C NMR (150 MHz, D2O): 177.0, 174.6, 174.3, 154.1, 143.7, 143.0, 142.2, 142.1, 129.5, 129.5, 126.0, 125.3, 123.3, 123.2, 112.0, 111.7, 104.0, 103.9, 101.9, 76.7, 76.1, 76.0, 73.7, 73.3, 70.2, 66.2, 60.6, 50.0, 49.9, 44.6, 44.0, 40.6, 40.3, 36.5, 31.7, 29.9, 29.8, 29.6, 29.4, 29.3, 27.9, 27.9, 27.8, 27.7, 27.6, 27.2, 26.4, 26.0 ppm. HRMS (ESI) m/z: [M]+ calc for C51H73N4O8+ 869.54229, found 869.54248.

Compound 13 (SY-F229)



Compound **11** (4.1 mg, 10 µmol) was dissolved in DMF (0.5 mL), then DIPEA (3.8 µL, 22 µmol) and biotin-OSu[35] (3.8 mg, 11 µmol) was added and the mixture was stirred overnight at rt. Full conversion was observed by LC-MS and the product was purified by semi-preparative reversed phase HPLC (linear gradient. Solutions used: A: 50 mM NH_4HCO_3 in H_2O, B: acetonitrile), affording the product as a white solid (1.59 mg, 25%) after lyophilization. ^{1}H NMR (500 MHz, D_2O): 4.51 (dd, *J* = 8.0, 4.6 Hz, 1H), 4.34 – 4.29 (m, 2H), 3.84 (dd, *J* = 11.6, 5.5 Hz, 1H), 3.59 (d, *J* = 8.5 Hz, 1H), 3.51 (ddd, *J* = 10.5, 9.1, 5.5 Hz, 1H), 3.44 (td, *J* = 10.4, 5.5 Hz, 1H), 3.33 (t, *J* = 9.2 Hz, 1H), 3.24 (dd, *J* = 8.7, 5.1 Hz, 1H), 3.21 (dd, *J* = 8.3, 1.9 Hz, 1H), 3.20 – 3.17 (m, 1H), 3.15 (dd, *J* = 9.4, 7.9 Hz, 1H), 3.12 – 3.05 (m, 2H), 2.90 (dd, *J* = 13.1, 5.0 Hz, 1H), 2.68 (d, *J* = 13.0 Hz, 1H), 2.43 (dd, *J* = 13.3, 5.6 Hz, 1H), 2.25 – 2.17 (m, 2H), 2.15 (t, *J* = 7.1 Hz, 2H), 1.95 – 1.91 (m, 1H), 1.68 – 1.63 (m, 1H), 1.61 (d, *J* = 6.3 Hz, 2H), 1.60 – 1.47 (m, 3H), 1.47 – 1.36 (m, 4H), 1.35 – 1.26 (m, 2H), 1.25 – 1.15 (m, 8H) ppm. ^{13}C NMR (150 MHz, D_2O): 177.0, 174.6, 174.3, 154.1, 143.7, 143.0, 142.2, 142.1, 129.5, 129.5, 126.0, 125.3, 123.3, 123.2, 112.0, 111.7, 104.0, 103.9, 101.9, 76.7, 76.1, 76.0, 73.7, 73.3, 70.2, 66.2, 60.6, 50.0, 49.9, 44.6, 44.0, 40.6, 40.3, 36.5,

31.7, 29.9, 29.8, 29.6, 29.4, 29.3, 27.9, 27.8, 27.7, 27.6, 27.2, 26.4, 26.0 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{29}H_{51}N_4O_9S$ 631.33713, found 631.33729.

Aspergillus niger (strain NRRL 3) was cultured for 8 days at 30 °C at 250 rpm with Beechwood xylan (>90% xylose content, Sigma Aldrich) as the sole carbon source (1%). The fungus was removed by filtration, and the secretome was concentrated (5x) by ultrafiltration over a Macrosep Advance spin filter (3K cutoff). After concentration, the total protein content (~2.3 μg/μL) was determined by the BCA assay[46] using the BCA kit from Pierce Chemical Company (Rockford, IL). For each SDS-PAGE experiment, the secretome (0.5 μL) was diluted with McIlvaine buffer (9.5 μL, 150 mM) of appropriate pH. Then, the ABP dissolved in buffer (5 μL, <1% DMSO) was added and the sample was incubated at 37 °C for 30 minutes. Subsequently, sample buffer (4X, Laemmli, 5 μL) was added and the sample was denatured at 100 °C for 5 minutes. A portion of this sample (10 μL) was loaded on the gel and 90-180V was applied. Page Ruler Plus was used as marker ladder. The gel was analyzed by fluorescent scanning with a BioRad ChemiDoc system.

References

1 Y. Man, H. Xiao, W. Cai and S. Yang, *Sci. Total Environ.*, 2017, **599–600**, 863–872.
2 K. Moustafa, *Sci. Total Environ.*, 2017, **598**, 639–646.
3 H. L. van Soest, H. S. de Boer, M. Roelfsema, M. G. J. den Elzen, A. Admiraal, D. P. van Vuuren, A. F. Hof, M. van den Berg, M. J. H. M. Harmsen, D. E. H. J. Gernaat and N. Forsell, *Clim. Change*, 2017, **144**, 165–179.
4 M. den Elzen, A. Admiraal, M. Roelfsema, H. van Soest, A. F. Hof and N. Forsell, *Clim. Change*, 2016, **137**, 655–665.
5 J. Marcel, R. Gallo and M. A. Trapp, *J. Braz. Chem. Soc*, 2017, **28**, 1586–1607.
6 A. J. Ragauskas, C. K. Williams, B. H. Davison, G. Britovsek, J. Cairney, C. A. Eckert, W. J. Frederick, J. P. Hallett, D. J. Leak, C. L. Liotta, J. R. Mielenz, R. Murphy, R. Templer and T. Tschaplinski, *Science*, 2006, **311**, 484–9.
7 L. P. Wackett, *Curr. Opin. Chem. Biol.*, 2008, **12**, 187–193.
8 US Energy Information Administration, *Short-Term Energy Outlook (STEO)*, 2017.
9 C. Escamilla-Alvarado, J. A. Pérez-Pimienta, T. Ponce-Noyola and H. M. Poggi-Varaldo, *J. Chem. Technol. Biotechnol.*, 2017, **92**, 906–924.
10 V. Menon and M. Rao, *Prog. Energy Combust. Sci.*, 2012, **38**, 522–550.
11 R. A. Prade, *Biotechnol. Genet. Eng. Rev.*, 1996, **13**, 101–131.
12 D. Shallom and Y. Shoham, *Curr. Opin. Microbiol.*, 2003, **6**, 219–228.
13 N. Kulkarni, A. Shendye and M. Rao, *FEMS Microbiol. Rev.*, 1999, **23**, 411–456.
14 V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho and B. Henrissat, *Nucleic Acids Res.*, 2014, **42**, 490–495.
15 T. Collins, C. Gerday and G. Feller, *FEMS Microbiol. Rev.*, 2005, **29**, 3–23.
16 V. Notenboom, S. J. Williams, R. Hoos, S. G. Withers and D. R. Rose, *Biochemistry*, 2000, **39**, 11553–11563.
17 S. Garg, *Curr. Metabolomics*, 2016, **4**, 23–37.
18 V. Kumar, J. Marín-Navarro and P. Shukla, *World J. Microbiol. Biotechnol.*, 2016, **32**, 1–10.
19 S. Dagley and C. N. Hinshelwood, *J. Chem. Soc.*, 1938, 1942–1948.

20 F. Niehaus, F. Niehaus, C. Bertoldo, C. Bertoldo, M. Kähler, M. Kähler, G. Antranikian and G. Antranikian, *Appl. Microbiol. Biotechnol.*, 1999, **51**, 711–729.

21 D. E. Stephens, K. Rumbold, K. Permaul, B. A. Prior and S. Singh, *J. Biotechnol.*, 2007, **127**, 348–354.

22 C. You, Q. Huang, H. Xue, Y. Xu and H. Lu, *Biotechnol. Bioeng.*, 2010, **105**, 861–870.

23 K. Miyazaki, M. Takenouchi, H. Kondo, N. Noro, M. Suzuki and S. Tsuda, *J. Biol. Chem.*, 2006, **281**, 10236–10242.

24 L. Song, A. Tsang and M. Sylvestre, *Biotechnol. Bioeng.*, 2015, **112**, 1081–1091.

25 W. W. Kallemeijn, K. Y. Li, M. D. Witte, A. R. A. Marques, J. Aten, S. Scheij, J. Jiang, L. I. Willems, T. M. Voorn-Brouwer, C. P. A. A. van Roomen, R. Ottenhoff, R. G. Boot, H. van den Elst, M. T. C. Walvoort, B. I. Florea, J. D. C. Codée, G. A. van der Marel, J. M. F. G. Aerts and H. S. Overkleeft, *Angew. Chem. Int. Ed.*, 2012, **51**, 12529–12533.

26 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.

27 J. Jiang, C. L. Kuo, L. Wu, C. Franke, W. W. Kallemeijn, B. I. Florea, E. van Meel, G. A. van der Marel, J. D. C. Codée, R. G. Boot, G. J. Davies, H. S. Overkleeft and J. M. F. G. Aerts, *ACS Cent. Sci.*, 2016, **2**, 351–358.

28 L. I. Willems, T. J. M. Beenakker, B. Murray, S. Scheij, W. W. Kallemeijn, R. G. Boot, M. Verhoek, W. E. Donker-Koopman, M. J. Ferraz, E. R. Van Rijssel, B. I. Florea, J. D. C. Codée, G. A. van der Marel, J. M. F. G. Aerts and H. S. Overkleeft, *J. Am. Chem. Soc.*, 2014, **136**, 11622–11625.

29 J. Jiang, W. W. Kallemeijn, D. W. Wright, A. M. C. H. van den Nieuwendijk, V. C. Rohde, E. C. Folch, H. van den Elst, B. I. Florea, S. Scheij, W. E. Donker-Koopman, M. Verhoek, N. Li, M. Schürmann, D. Mink, R. G. Boot, J. D. C. Codée, G. A. van der Marel, G. J. Davies, J. M. F. G. Aerts and H. S. Overkleeft, *Chem. Sci.*, 2015, **6**, 2782–2789.

30 L. Wu, J. Jiang, Y. Jin, W. W. Kallemeijn, C.-L. Kuo, M. Artola, W. Dai, C. van Elk, M. van Eijk, G. A. van der Marel, J. D. C. Codée, B. I. Florea, J. M. F. G. Aerts, H. S. Overkleeft and G. J. Davies, *Nat. Chem. Biol.*, 2017, **13**, 867–873.

31 D. Crich and Z. Dai, *Tetrahedron*, 1999, **55**, 1569–1580.

32 S. Murata, M. Suzuki and R. Noyori, *Bull. Chem. Soc. Jpn.*, 1982, **55**, 247–254.

33 E. Thiery, J. Le Bras and J. Muzart, *Eur. J. Org. Chem.*, 2009, 961–985.

34 M. V. Kvach, A. V. Ustinov, I. A. Stepanova, A. D. Malakhov, M. V. Skorobogatyi, V. V. Shmanai and V. A. Korshun, *Eur. J. Org. Chem.*, 2008, 2107–2117.

35 K. Susumu, H. T. Uyeda, I. L. Medintz, T. Pons, J. B. Delehanty and H. Mattoussi, *J. Am. Chem. Soc.*, 2007, **129**, 13987–13996.

36 E. Schuster, N. Dunn-Coleman, J. Frisvad and P. Van Dijck, *Appl. Microbiol. Biotechnol.*, 2002, **59**, 426–435.

37 F. Li, J. Xie, X. Zhang and L. Zhao, *J. Microbiol. Biotechnol.*, 2015, **25**, 11–17.

38 V. Girard, C. Dieryckx, C. Job and D. Job, *Proteomics*, 2013, **13**, 597–608.

39 A. Economou, *Mol. Membr. Biol.*, 2002, **19**, 159–69.

40 G. Blobel and B. Dobberstein, *J. Cell Biol.*, 1975, **67**, 835–851.

41 H. Antelmann, H. Tjalsma, B. Voigt, S. Ohlmeier, S. Bron, J. M. van Dijl and M. Hecker, *Genome Res.*, 2001, **11**, 1484–1502.

42 J. Jiang, Thesis: Activity-based protein profiling of glucosidases, fucosidases and glucoronidases; Leiden University, 2016.

43 A. R. Stricker, R. L. Mach and L. H. De Graaff, *Appl. Microbiol. Biotechnol.*, 2008, **78**, 211–220.

44 A. MacCabe, M. Fernández-Espinar, L. H. De Graaff, J. Visser and D. Ramón, *Gene*, 1996, **175**, 29–33.

45 S. Caner, X. Zhang, J. Jiang, H. M. Chen, N. T. Nguyen, H. Overkleeft, G. D. Brayer and S. G. Withers, *FEBS Lett.*, 2016, **590**, 1143–1151.

46 P. K. Smith, R. I. Krohn, G. T. Hermanson, A. K. Mallia, F. H. Gartner, M. D. Provenzano, E. K. Fujimoto, N. M. Goeke, B. J. Olson and D. C. Klenk, *Anal. Biochem.*, 1985, **150**, 76–85.

Chapter 8

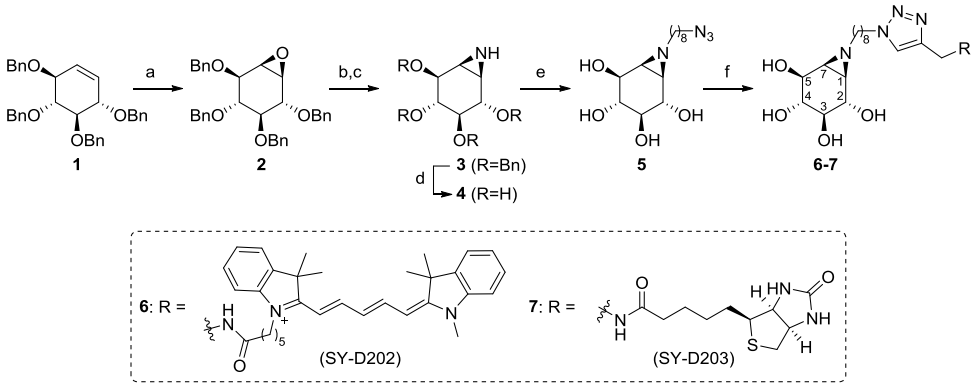
Conclusions and future prospects

Cyclophellitol and cyclophellitol aziridine are potent, selective and irreversible mechanism-based inhibitors of retaining β-glucosidases.[1,2] Cyclophellitol aziridine equipped with a reporter tag allows facile visualization and identification of active β-glucosidases in complex biological samples with high sensitivity.[3] Configurational isomers of cyclophellitol aziridine show high selectivity towards their target glycosidases, which normally processes substrates with matching configurations.[4–7] The work described in this Thesis entails modification of the cyclophellitol core in such a way that glycosidases processing different configurational sugar substrates are targeted simultaneously – thus to create broad-spectrum activity-based glycosidase probes. Additionally, the first synthetic examples of multimeric cyclophellitol aziridine activity-based probes are described. These tools enable the study of *endo*-glycosidases; a field yet unexplored by cyclophellitol-based probes.

1| **Chapter 1** provides a global overview of glycosidase functions in organisms, as well as their mechanistic aspects of catalytic glycoside hydrolysis. The basic principles of activity-based protein profiling are described, as well as the mechanism of inactivation of glycosidases by cyclitol epoxides and aziridines.

2| A new route towards D-*xylo*-cyclophellitols is described in **Chapter 2**. The key step in this synthesis route involves an asymmetric Brown allylation reaction, which following a series of chemical transformations afforded α- and β-configured D-*xylo*-cyclophellitol epoxides and aziridines. The β-aziridine was equipped with different fluorescent tags, and upon incubation with mouse liver lysates, the probe labeled β-glucosylcerebrosidase (GBA) 1 and 2. Pull-down analysis with the biotinylated probe also identified GBA1 and GBA2 in mouse liver, as well as GBA1 in mouse brain and duodenum. In addition, the D-*xylo*-cyclophellitol aziridine probes proved highly potent activity-based inactivators of GH52 β-xylosidase from *Opitutus terrae*.

It would be of interest to further investigate the influence of the substituent at C5 of cyclitols on the inhibition of glycosidases. For example, D-*xylo*-cyclophellitol aziridine equipped with a fluorescent tag was shown to label GBA1 and GBA2 in mouse liver lysate, in line with cyclophellitol aziridine probes reported previously.[3] In contrast,

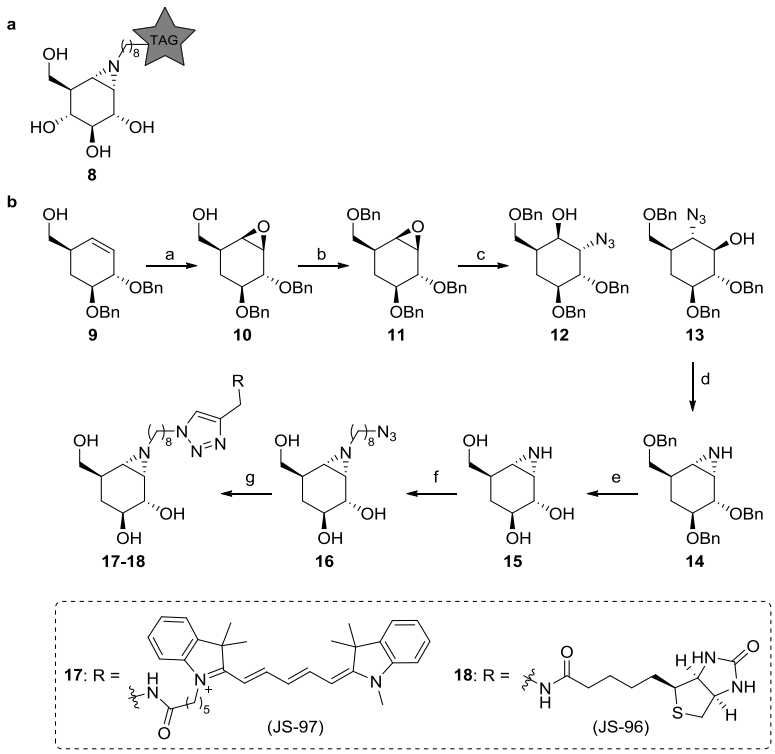


Scheme 1 Synthesis of conduritol B aziridine activity-based probes **6-7**. Reagents and conditions: a) oxone, CF_3COCH_3, $NaHCO_3$, EDTA, H_2O, MeCN, 0 °C, 87%; b) NaN_3, $LiClO_4$, DMF, 100 °C, 16h; c) polymer-bound PPh_3, MeCN, 60 °C, 33% over two steps; d) Li, NH_3, THF, -60 °C, 1h, 99%; e) 1-azido-8-iodooctane, K_2CO_3, DMF, 80 °C, 16h, 40%; f) tag-alkyne, $CuSO_4$, Na-ascorbate, DMF, yield **6**: 40%; yield **7**: 37%.

conduritol B aziridine (which possesses an equatorial hydroxyl group at C5), *N*-alkylated with various alkyl moieties, appeared highly specific for GBA1 over GBA2.[8] To further investigate the influence of the substituent at C5 in cyclitols towards the labeling of glycosidases, conduritol B aziridines **6**-**7** equipped with a reporter tag were synthesized (Scheme 1). The synthesis commenced with epoxidation of cyclohexene **1**[9] to afford **2**, which was converted into aziridine **3** by azidolysis followed by Staudinger-type ring closure. The benzyl protecting groups were removed under Birch conditions to afford conduritol B aziridine **4**, which was alkylated with an azidooctyl spacer to give **5**. Click ligation with the appropriate alkynes afforded conduritol B aziridine ABPs **6** and **7** (see the experimental section for synthetic procedures and compound characterization). Due to its C_2-symmetry axis, conduritol B epoxide is a potent inactivator of both α- and β-glucosidases.[10] Therefore, conduritol B aziridine ABPs **6** and **7** could potentially also serve as broad-spectrum α/β-glucosidase probes.

Chapter 3 describes the synthesis of two deoxygenated analogues (4-deoxy and 2,4-deoxy) of cyclophellitol aziridine. The 4-deoxygenated analogue was synthesized by deoxygenation of a common precursor in the synthesis of cyclophellitol aziridine. The synthesis of the 2,4-deoxy analogue required a different synthetic route, starting from an achiral synthon which was subjected to an iridium catalyzed asymmetric allylic alkylation followed by an asymmetric Brown allylation. It was envisioned that deoxygenation would reduce selectivity of the probe towards glycosidase active-sites, thereby enabling broad-spectrum glycosidase profiling. More specifically, the substrate configuration of OH-4 distinguishes its specificity in recognition by glucosidases and galactosidases, and OH-2 between glucosidases and mannosidases. Indeed, the 4-deoxygenated cyclophellitol aziridine probe enabled inter-class fluorescent labeling of β-glucosidases and β-galactosidases, as well as their identification by proteomics. Concomitant deoxygenation at C2 and C4 severely reduced the potency of the probe towards glycosidase labeling, however it appeared that the probe is selective for β-galactocerebrosidase.

In some initial studies aimed to establish whether a similar broad-spectrum activity could be obtained with deoxygenated analogues of *epi*-cyclophellitol aziridine **8**[4] (Scheme 2a), 4-deoxy *epi*-cyclophellitol aziridine probes were synthesized. Cyclohexene **9** (Chapter 3) was diastereoselectively epoxidized with *m*-CPBA to afford epoxide **10** (Scheme 2b). Benzylation of the primary alcohol gave **11**, which was then



Scheme 2 Synthesis of 4-deoxy *epi*-cyclophellitol aziridine probes. Reagents and conditions: a) *m*-CPBA, DCM, 0 °C, 91%; b) BnBr, NaH, TBAI, DMF, 79%; c) NaN_3, $LiClO_4$, DMF, 80 °C, 48h, ratio **12**:**13** 1:1.3, 83%; d) polymer-bound PPh_3, MeCN, 90 °C, 73%; e) Li, THF, NH_3, -60 °C, quant.; f) 8-azido-1-iodooctane, K_2CO_3, DMF, 100 °C, 55%; g) tag-alkyne, $CuSO_4$, Na-ascorbate, DMF, yield **17**: 44%; yield **18**: 56%.

reacted with sodium azide to provide a mixture of azidoalcohols **12** and **13**. Compound **13** was then subjected to Staudinger-type ring closure to afford aziridine **14**. Debenzylation under Birch conditions (**15**), followed by alkylation gave alkyl aziridine **16** which could be click-ligated with different reporter tags to afford 4-deoxy α-ABPs **17**-**18** (see the experimental section for synthetic procedures and compound characterization). While specific α-glucosidase ABP **8** positively identified the lysosomal acid α-glucosidase (GAA) in mouse kidney, liver and human fibroblast lysates,[4] deoxygenation at C4 totally abolished labeling of glycosidases (Figure 1).

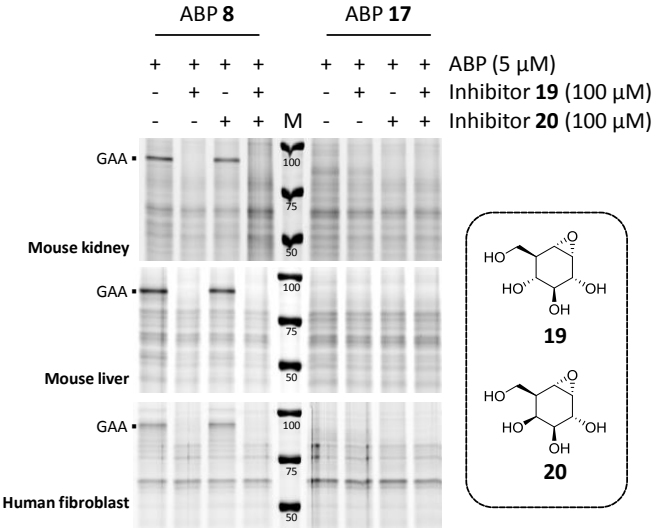
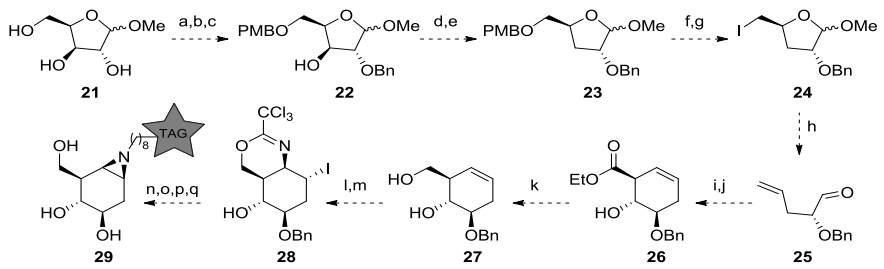


Figure 1 Labelling of mouse kidney, liver and human fibroblast lysates with *epi*-cyclophellitol aziridine probe **8** (left panel) and 4-deoxy *epi*-cyclophellitol aziridine **17** (right panel). Selective probe **8** labels GAA in all lysates and labelling can be competed with selective inhibitor **19** (and not **20**), indicating that labelling is activity-based. The 4-deoxy compound **17** does not significantly label any bands in these lysates in activity-based manner.

Alternatively, it would be of interest to develop a probe that is capable of simultaneously labeling β-glucosidases and β-mannosidases. For this purpose, C2 deoxygenated cyclophellitol aziridine probes could be synthesized (Scheme 3). Starting from methyl D-xylofuranose **21**, the 4,6-diol could be regioselectively protected as the *p*-anisylidene acetal. The remaining alcohol could be benzylated, and reductive acetal cleavage using borane and Bu_2BOTf at low temperature[11] could afford the 6-*O*-*p*-methoxybenzyl (PMB) ether **22**. The hydroxyl group could then be removed by tosylation and hydride reduction to give **23**. Acid catalyzed PMB cleavage using HCl in hexafluoroisopropanol (HFIP)[12] followed by an Appel iodination could afford compound **24**. Zinc-mediated Vasella fragmentation could give aldehyde **25**, which could be subjected to an indium catalyzed Barbier allylation[13] followed by Grubbs metathesis to afford ester **26**. Reduction of the ester could afford **27**, which is then reacted with trichloroacetonitrile and subsequently treated with *N*-iodosuccinimide to afford cyclic imidate **28**. Acidic hydrolysis followed by basification could then afford the aziridine, which is globally debenzylated under Birch conditions followed by



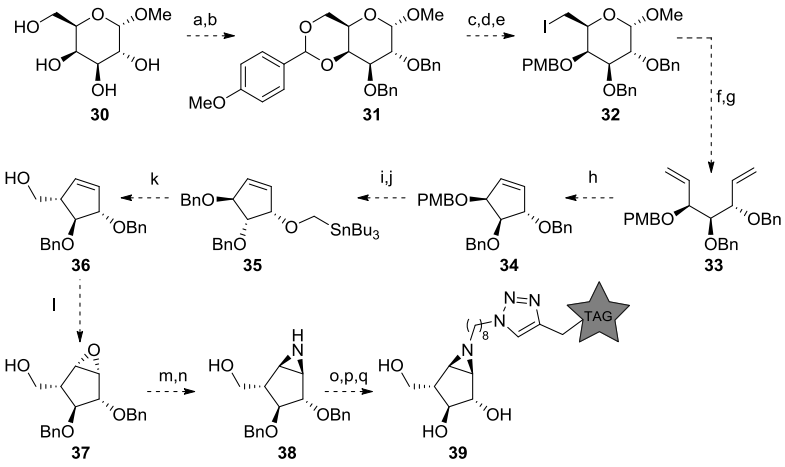
Scheme 3 Proposed synthetic scheme towards C-2 deoxy-ABPs **29**. Reagents and conditions: a) anisaldehyde dimethyl acetal, *p*-TsOH, DMF, 60 °C; b) BnBr, NaH, TBAI, DMF; c) BH_3, Bu_2BOTf, THF, DCM, -78 °C; d) TsCl, pyridine; e) $LiAlH_4$, THF, reflux; f) HCl, HFIP, DCM; g) PPh_3, I_2, THF, reflux; h) Zn, THF, H_2O, sonication; i) 4-ethylbromocrotonate, In, $La(OTf)_3$, H_2O; j) Grubbs II, DCM, reflux; k) DIBAL-H, $NaBH_4$, THF; l) CCl_3CN, DBU, DCM; m) NIS, $CHCl_3$; n) HCl, MeOH, then Et_3N; o) Li, NH_3; p) 1-azido-8-iodooctane, K_2CO_3, DMF, 100 °C; q) tag-alkyne, $CuSO_4$, Na-ascorbate, DMF.

alkylation with a click handle and click-ligation with reporter tags to finally afford 2-deoxy aziridines **29**.

Chapter 4 describes the synthesis of β-D-*lyxofuranosyl* and β-D-*arabinofuranosyl*-cyclophellitol aziridine probes. The ability of these probes to fluorescently label glycosidases in unbiased fashion was evaluated in human fibroblast and mouse liver lysates. Activity-based fluorescent labeling of glycosidases could not be detected for both probes, and pull-down analysis with the biotinylated probes did not identify glycosidases in various biological samples. It was therefore concluded that none of the biological samples (human or mouse) contained glycosidases that accepted the D-furanosyl configured probes as substrate.

In contrast to β-D-arabinofuranosides, α-L-arabinofuranosides are common components in hemicellulosic biomass. The β-1,4-xylan backbone in hemicellulose is randomly substituted at O2, O3 or both by α-L-arabinofuranosides. These side-groups impede rapid processing of hemicellulose by *endo*-xylanases and *exo*-xylosidases, retarding the hydrolysis reaction in industrial applications (such as the production of D-xylose for the production of xylitol or ethanol). A probe that would rapidly detect α-L-arabinofuranosidase activity with high sensitivity could be of interest for the identification and development of α-L-arabinofuranosidases with high substrate turnover rate. For this purpose, an α-L-*arabino*-configured cyclophellitol aziridine probe would be of interest. The synthesis of this probe may start from commercially

available methyl α-D-galactopyranosidase **30** (Scheme 4), which is then regioselectively 4,6-*O* protected as the *p*-anisylidene acetal, followed by benzylation of the remaining alcohols to afford globally protected **31**. The acetal may then be reductively cleaved using borane and Bu_2BOTf to the 4-*O*-PMB ether,[11] followed by tosylation and nucleophilic displacement with iodine at O-6 to afford orthogonally protected iodopyranoside **32**. Zinc-mediated Vasella fragmentation followed by Wittig olefination using triphenylphosphonium methylide would afford diene **33**, which is subsequently ring-closed using Grubbs catalyst to give **34**. The PMB-group may be selectively deprotected using HCl in HFIP,[12] and the liberated alcohol condensated with tributyl-iodomethylstannane to afford **35**. Hydroxymethylation via a [2,3]-Wittig-Still rearrangement[14,15] would subsequently afford alcohol **36**. Stereoselective α-epoxidation using *m*-CPBA would then afford **37**, which is subsequently transformed into aziridine **38** by azidolysis followed by Staudinger-type ring closure. Global deprotection under Birch conditions followed by alkylation with 1-azido-8-iodooctane and subsequent click ligation with an appropriate reporter tag may finally afford α-L-*arabino*-cyclophellitol aziridine probes **39**.

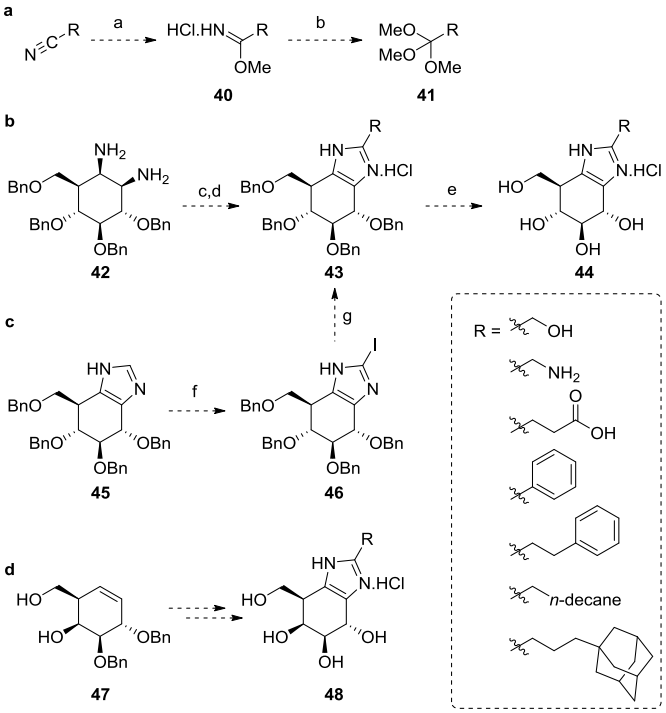


Scheme 4 Proposed synthetic scheme for the preparation of α-L-*arabino*-cyclophellitol aziridine probes **39**. Reagents and conditions: a) anisaldehyde dimethyl acetal, *p*-TsOH, DMF, 60 °C; b) BnBr, NaH, TBAI, DMF; c) BH_3, Bu_2BOTf, THF, DCM, 0 °C; d) TsCl, pyridine; e) NaI, DMF, 80 °C; f) Zn, EtOH, reflux; g) CH_3PPh_3Br, *n*-BuLi, THF, -40 °C; h) Grubbs II, DCM, reflux; i) HCl, HFIP, DCM; j) Bu_3SnCH_2I, KH, dibenzo-18-crown-6, THF, 0 °C; k) *n*-BuLi, THF, -78 °C; l) *m*-CPBA, DCM; m) NaN_3, $LiClO_4$, DMF, 100 °C; n) polymer-bound triphenylphosphine, MeCN, 80 °C; o) Li, NH_3, THF, -60 °C; p) 8-azido-1-iodooctane, K_2CO_3, DMF, 100 °C; q) alkyne-tag, $CuSO_4$, Na-ascorbate, DMF.

5| Gluco-azoles competitively inhibit glucosidases by transition-state mimicry and their ability to interact with catalytic acid residues in glucosidase active sites.[16] However, none of such azole-type inhibitors described to date possess a protic nitrogen characteristic for 1*H*-imidazoles. **Chapter 5** describes the synthesis and biochemical evaluation of gluco-1*H*-imidazole, a gluco-azole bearing a 1*H*-imidazole fused to a glucopyranose-configured cyclitol core, as well as three close analogues as novel glycosidase inhibitors. While both compounds exhibit high structural similarity and binding modes to β-glucosidases from *Thermotoga maritima* and *Thermoanaerobacterium xylanolyticum* (as shown by X-ray crystallography), it was found that unsubstituted gluco-1*H*-imidazole displayed reduced activity towards β-glucosidases, in comparison to the classical glucoimidazole. It can be postulated that the activity of gluco-1*H*-imidazole is lower due to reduced interaction of the 'glycosidic' nitrogen with the glucosidase catalytic acid, caused by prototropic tautomerization of the imidazole proton. Additionally, DFT calculations showed that protonation of glucoimidazole causes δ+ charge development at the 'anomeric' centre, ideally located for interaction with the glucosidase catalytic nucleophile. In gluco-1*H*-imidazole, protonation of the azole causes δ+ character on the 'apical' carbon of the azole ring, poorly placed for interaction with the catalytic nucleophile. In contrast, gluco-1*H*-imidazoles equipped with an *n*-butyl moiety are effective inhibitors of human GBA1, inhibiting this enzyme (deficient in Gaucher disease) in the nanomolar range. Indeed, this positive effect of a hydrophobic moiety at the 'aglycon' subsite was also observed for the classical glucoimidazoles by Vasella *et al*.[17] For example, glucoimidazole substituted with hydroxymethylene or ethylphenyl moieties at position C2 of the imidazole displayed activity in the low nanomolar range.

It would be of interest to derivatize gluco-1*H*-imidazole with different aliphatic moieties (such as hydromethyl, ethylphenyl, adamantyl or *n*-decyl) on the imidazole ring and study its effect on inhibitory activity. For this purpose, diamine **42** (Chapter 5) may be condensed with appropriate trimethyl orthoesters **41** (synthesized in two steps from the matching nitrile using methanolic HCl;[18] Scheme 5a) to afford substituted imidazoles **43**, which would afford inhibitors **44** after palladium-catalyzed debenzylation (Scheme 5b). Alternatively, 1*H*-imidazole **45** may be treated with *N*-iodosuccinimide to afford iodoazole **46**, which may be coupled with suitable organotin moieties under palladium catalysis[19,20] to afford substituted imidazoles **43**.

Moreover, the synthetic procedure described in Chapter 5 for the construction of gluco-1*H*-imidazoles could be adapted to other relevant cyclohexene scaffolds (Scheme 5 d). For example, *galacto*- configured cyclohexene **47**[21] could be utilized for the synthesis of galacto-1*H*-imidazoles **48**, which are potential inhibitors of β-galactosidases. Deficiency of active acid β-galactosidase (GLB1) causes two different lysosomal storage diseases (LSDs), namely G_{M1}-gangliosidosis[22] and Morquio B disease.[23] Chaperone therapy with small molecule inhibitors is reported to be promising treatment of these LSDs,[24] and galacto-1*H*-imidazoles **48** could be investigated for this purpose as well.



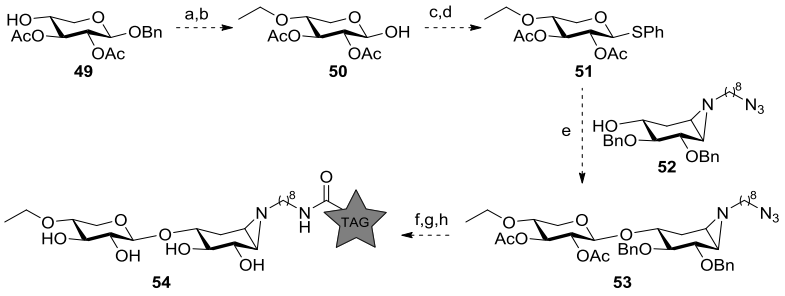
Scheme 5 (a) General synthesis of trimethyl orthoesters from terminal nitriles. (b) Proposed synthetic scheme towards substituted gluco-1*H*-imidazoles by condensation with trimethyl orthoesters. (c) Proposed synthetic scheme towards substituted gluco-1*H*-imidazoles by azole iodination and palladium catalyzed organotin coupling. Reagents and conditions: a) HCl, MeOH, (^{i}Pr)$_2$O, -20 °C; b) MeOH, Et$_2$O, reflux; c) (MeO)$_3$R, HFIP; d) (COCl)$_2$, DMSO, DCM, -60 °C; e) Pd(OH)$_2$/C, H$_2$, HCl, dioxane, MeOH, H$_2$O; f) NIS, CHCl$_3$; g) SnR$_4$, Pd(PPh$_3$)$_4$, DMF, 95 °C. (d) Galacto-1*H*-imidazoles **48** could be prepared from cyclohexene **47**[21] following the synthetic procedures reported in Chapter 5.

6| **Chapter 6** describes the synthesis of two spiro-epoxyglycosides designed to irreversibly inhibit GH99 *endo*-α-mannosidases (containing either a gluco- or mannopyranosyl residue at the non-reducing end). These enzymes are postulated to employ an unusual reaction trajectory, following through deprotonation of OH-2 to form an 1,2-anhydro epoxide transition state which is subsequently hydrolyzed by water.[25] The spiro-epoxyglycosides described contain a spiro-epoxide warhead at C-2, which was installed by a Corey-Chaykovsky spiro-epoxidation reaction. It was shown that the spiro-epoxyglycosides are able to label recombinant GH99 *endo*-α-mannosidases in concentration-, pH- and time-dependent fashion, and that labeling could be competed with the matching non-tagged inhibitors, as well as the enzyme natural substrate. Additionally, labeling of the enzyme totally abrogated the enzyme's native de-mannosylating ability. In order to unequivocally prove the postulated unusual reaction mechanism of GH99 *endo*-α-mannosidases, it would be of interest to obtain a crystal structure of the spiro-epoxyglycosides bound in the active site. Nucleophilic opening of the spiro-epoxide warhead by the putative nucleophile, together with hydrogen bonding of the putative acid/base with the aglycon would prove the suggested mechanism. Alternatively, incubation of the GH99 enzyme with the spiro-epoxyglycosides, followed by proteolysis and LC-MS/MS analysis could identify the labeled peptide fragment, and further fragmentation of this peptide could unveil the labeled nucleophile. Previously, this technique successfully identified the catalytic nucleophile in several glycosidases.[26–29]

7| **Chapter 7** describes the synthesis of *xylobiose*-cyclophellitol aziridine probes. The results described in this Chapter provide the first example that monomeric *xylo*-cyclophellitol epoxides and aziridines can readily be glycosylated with an appropriate thioglycoside donor under NIS/TMSOTf catalysis to afford 'disaccharide' cyclitol epoxides and aziridines in good yields. This finding invites the synthesis of multimeric cyclitol probes for the profiling of *endo*-glycosidases. This is exemplified by the synthesized *xylobiose*-cyclophellitol aziridine probe, which was able to visualize *endo*-xylanase activity in the secretome of industrially relevant fungus *Aspergillus niger* and profile the pH and temperature tolerance of the enzyme.

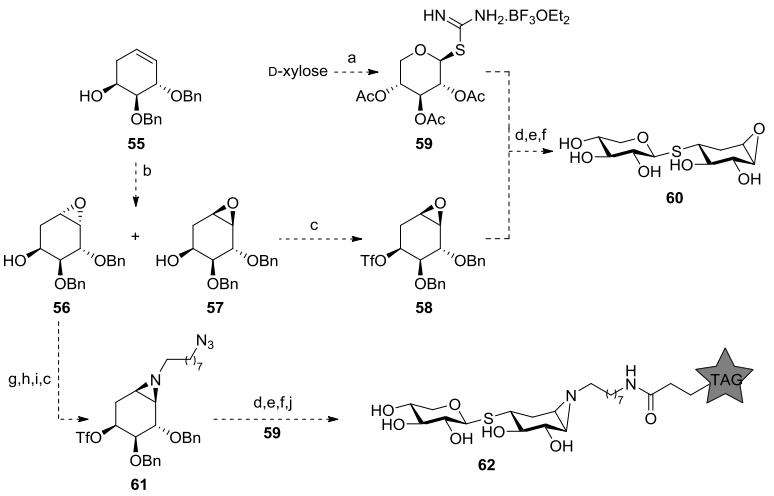
A drawback of the glycosylated *endo*-xylanase probe is that it is susceptible to hydrolysis by *exo*-xylosidases present in the complex biological sample. However, it was shown that this can be prevented by pre-incubation of the sample with a specific irreversible *exo*-xylosidase inhibitor. Alternatively, probes and covalent inhibitors that

are resistant towards *exo*-xylosidase cleavage would be of interest. Fur such purpose, 'capping' O-4 of the non-reducing pyranoside with an ether (e.g. ethoxy) group could prevent *exo*-glycosidase processing (Scheme 6). For example, compound **49**[30] could be alkylated using ethyl triflate, followed by palladium catalyzed debenzylation to afford **50**. Acetylation and reaction with thiophenol under Lewis acidic conditions affords donor **51**, which may be coupled with aziridine acceptor **52** to afford **53**. Deprotection and amide coupling then affords O-4 capped *xylobiose*-cyclophellitol aziridine **54**.



Scheme 6 Proposed synthetic scheme towards O-4 'capped' *xylobiose*-cyclophellitol aziridine **54**, which is anticipated to be resistant to hydrolysis by *exo*-glycosidases. Reagents and conditions: a) ethyl triflate, Et_3N, DCM; b) Pd/C, H_2, THF; c) Ac_2O, pyridine; d) PhSH, $BF_3{\cdot}OEt_2$, DCM; e) NIS, TMSOTf, DCM, -40 °C; f) NaOMe, MeOH, DCM; g) Na, ^tBuOH, NH_3, -60 °C; h) tag-OSu, DIPEA, DMF.

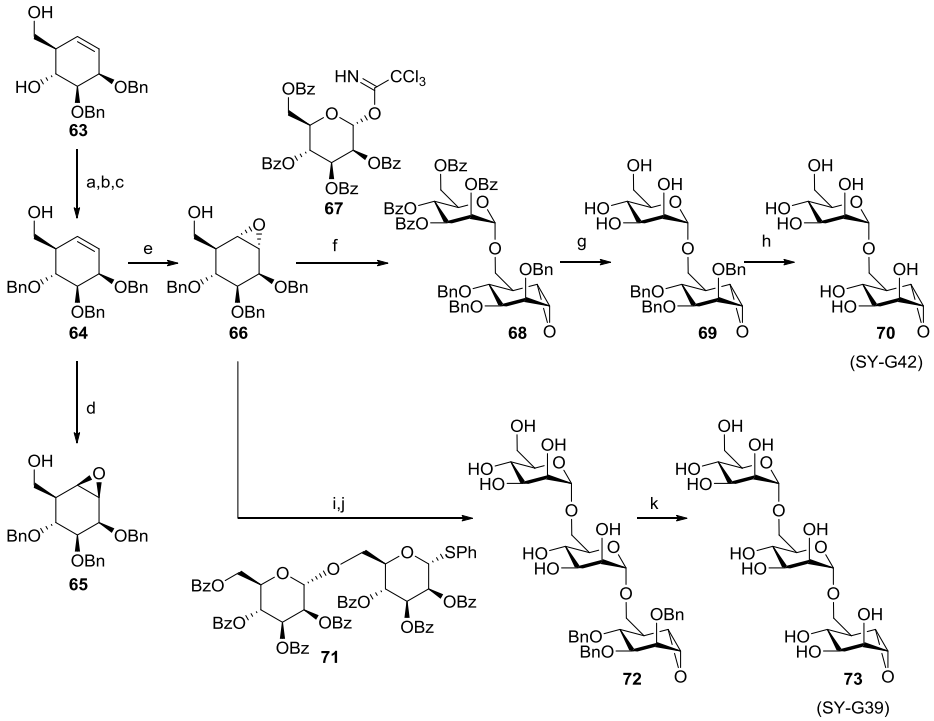
An alternative strategy would be the replacement of the hydrolysable glycosidic linkage by a thioglycosidic linkage to yield stabilized multimeric *endo*-probes. Thioglycosidic linkages are generally resistant to glycosidic processing due to the lower basicity of sulphur compared to oxygen, thereby hampering protonation of the aglycon by the enzyme acid/base residue during catalysis.[31,32] For example, 4-deoxy-4-thio-*xylobiose*-cyclophellitol **60** might be available from the coupling of triflate **58** and isothiouronium donor **59** (Scheme 7).[33,34] Donor **59** is available from D-xylose in two steps by acetylation and substitution by thiourea.[35] Triflate acceptor **58** may be obtained by epoxidation of cyclohexene **55** (Chapter 2), followed by triflation of the alcohol in β-epoxide **57**. The thioglycosidic linkage between **58** and **59** may then be installed by treatment with base, and Zemplén deacetylation followed by debenzylation under Birch conditions would afford thioglycosidic *xylobiose*-cyclophellitol **60**. The synthesis of thioglycosidic *xylobiose*-cyclophellitol aziridine **62** would follow the same strategy, by coupling triflate **61** with isothiouronium donor **59** followed by deprotection and amide coupling with appropriate succinimidyl esters.



Scheme 7 Proposed synthetic scheme towards stabilized thioglycosidic *xylobiose*-cyclophellitol **60** and fluorescent aziridine **62**. Reagents and conditions: a) 1. Ac_2O, $BF_3.OEt_2$; 2. thiourea, MeCN, 70 °C; b) Oxone, CH_3COCF_3, $NaHCO_3$, EDTA, MeCN, H_2O; c) Tf_2O, pyridine, DCM, -20 °C; d) Et_3N, MeCN; e) NaOMe, MeOH; f) Li, NH_3; g) NaN_3, $LiClO_4$, DMF, 100 °C; h) polymer-bound PPh_3, MeCN, 80 °C; i) 8-azido-1-trifluoromethanesulfonyloctane, DIPEA, $CHCl_3$; j) tag-OSu, DIPEA, DMF.

Glycoside hydrolase family GH76 *endo*-α-1,6-mannosidases are found in bacteria and fungi. In bacteria such as *Bacteroides thetaiotaomicron*, a member of the human gut microbiota, these enzymes facilitate the breakdown of large 'high mannose' glycoproteins that make up the outer cell wall of yeasts and filamentous fungi.[36,37] In contrast, GH76 *endo*-α-1,6-mannosidases in fungi display transglycosylation activity, building up the extracellular α-mannan layer.[38] GH76 *endo*-α-1,6-mannosidase gene knockouts in certain fungi impair cell growth and can induce cell death.[39,40] Therefore, the development of (irreversible) GH76 *endo*-α-1,6-mannosidase inhibitors are of interest. For this purpose, two multimeric α-1,6-*manno-epi*-cyclophellitols have been synthesized (Scheme 8). Starting from diol **63**[41], the secondary alcohol was protected by a three-step procedure involving protection of the primary alcohol with TBS, benzylation of the secondary alcohol and desilylation to afford **64**. The olefin could be stereoselectively epoxidized by choice of reagent; using *m*-CPBA β-epoxide **65** was selectively obtained, and in turn α-epoxide **66** was obtained by treatment with *in situ* generated (trifluoromethyl)methyldioxirane. Epoxide **66** was subsequently coupled with trichloroimidate donor **67**[42] to afford 'disaccharide' **68**. Zemplén deacetylation afforded **69**, which was then debenzylated using Pearlman's catalyst to afford α-1,6-

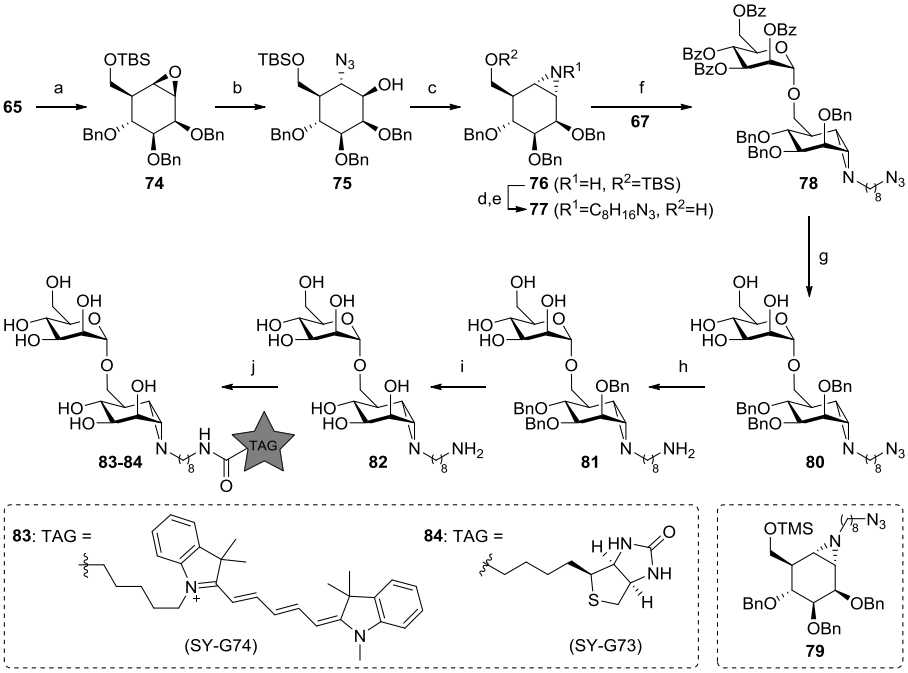
mannobiose-epi-cyclophellitol **70**. Similarly, epoxide **66** was coupled to disaccharide donor **71**[43] followed by deacetylation to afford **72** in low yield, which gave α-1,6-*mannotriose-epi*-cyclophellitol **73** after debenzylation (see the experimental section for synthetic procedures and compound characterization).



Scheme 8 Synthesis of α-1,6-*mannotriose-epi*-cyclophellitols **70** and **73**. Reagents and conditions: a) TBSCl, imidazole, DMF; b) BnBr, NaH, TBAI, DMF; c) TBAF, THF, 89% over three steps; d) *m*-CPBA, DCM , 0 °C, 84%; e) Oxone, CH_3COCF_3, $NaHCO_3$, EDTA, H_2O, MeCN, 0 °C, 89%; f) TMSOTf, DCM, -40 °C, 65%; g) NaOMe, MeOH, DCM, 81%; h) $Pd(OH)_2/C$, H_2, dioxane, MeOH, H_2O, 2h, quant; i) NIS, TMSOTf, DCM, -40 °C; j) NaOMe, MeOH, DCM, 14% over two steps; k) $Pd(OH)_2/C$, H_2, dioxane, MeOH, H_2O, 2h, quant.

In order to enable further studies on the GH76 *endo*-α-1,6-mannosidase, a set of tagged aziridines was synthesized (Scheme 9). Protection of the primary alcohol in epoxide **65** (Scheme 8) gave **74**, which was subjected to azidolysis to afford azidoalcohol **75**. Staudinger-type ring closure afforded aziridine **76**, which was alkylated and desilylated to afford acceptor **77**. Coupling with donor **67**[42] under Lewis

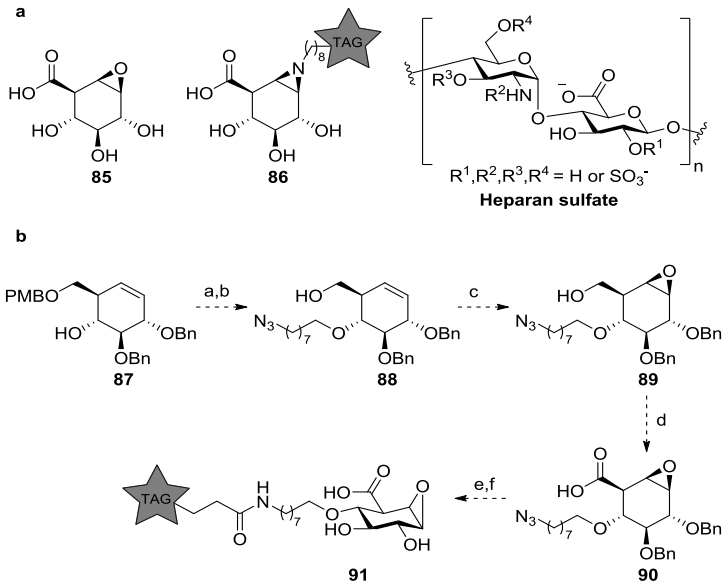
acidic conditions, using a small excess of triflic anhydride afforded 'disaccharide' **78** (of note: when TMSOTf was used as Lewis acid, only silylated acceptor **79** was formed in 78% isolated yield). Deacetylation (**80**) followed by Staudinger reduction (**81**) and debenzylation under Birch conditions afforded **82**, which was ligated with different reporter tags to afford α-*mannobiose-epi*-cyclophellitol aziridines **83** and **84** (see the experimental section for synthetic procedures and compound characterization).



Scheme 9 Synthesis of α-1,6-*mannobiose-epi*-cyclophellitol aziridines **83** and **84**. Reagents and conditions: a) TBSCl, imidazole, THF, rt, 4h, 93%; b) NaN_3, $LiClO_4$, DMF, 80 °C, 16h, 49%; c) polymer-bound PPh_3, MeCN, 90 °C, 16h, 85%; d) 8-azidooctyl trifluoromethanesulfonate, DIPEA, $CHCl_3$, 3h; e) TBAF, THF, 88% over two steps; f) imidate donor **67** (Scheme 8), TfOH, DCM, -40 °C, 74%; g) NaOMe, MeOH, DCM, 79%; h) polymer-bound PPh_3, H_2O, MeCN, 70 °C, 16h, 96%; i) Li, THF, NH_3, -60 °C, 1h, 81%; j) tag-OSu, DIPEA, DMF, 80: 29%; 81: 52%.

Recently, the development of a set of β-glucuronic acid configured cyclophellitols was reported (Scheme 10a).[6] Epoxide **85** was shown to be a potent inactivator of exo-acting β-glucuronidase (GUSB) from *Acidobacterium capsulatum* and its aziridine analogues displayed even higher potencies. Probe **86** equipped with a Cy5 fluorescent tag was shown to effectively label GUSB in human spleen lysates. Additionally, probe **86** also targets endo-acting heparanase (HPSE), as well as its inactive pro-enzyme proHPSE. HPSE is responsible for the intra- and extracellular degradation of heparan

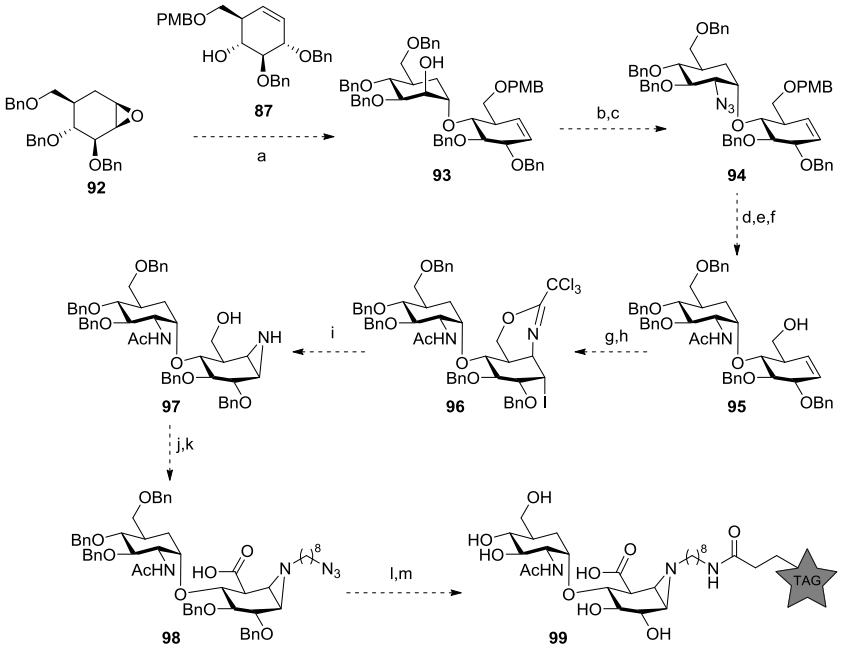
sulfate (HS), and its activity is strictly regulated. Increased extracellular HPSE activity in urine samples is a biomarker for inflammatory kidney disease, acute kidney injury and renal injury in diabetic nephropathy.[44] An activity-based probe that selectively targets HPSE could be used as a diagnostic tool for the quantification of excreted urinary HPSE levels in patients. It is envisioned that installment of a reporter tag onto C4 of epoxide **85** would block processing by exo-acting GUSB, thereby enabling selective HPSE labeling. For this purpose, cyclohexene **87** (Chapter 3) could be alkylated with an 8-azidooctyl linker, followed by PMB deprotection using DDQ to give **88** (Scheme 10b). Using *m*-CPBA, the liberated primary alcohol could then stereoselectively direct the epoxidation to β-epoxide **89**, which is then oxidized to carboxylic acid **90** using the TEMPO/BAIB system. Debenzylation under Birch conditions followed by amide coupling with different reporter tags could then afford a set of potential HPSE selective activity-based probes **91**.



Scheme 10 (a) Chemical structures of β-glucuronidase inhibitor **85**, activity-based probe **86** and the generalized structure of heparan sulfate. (b) Proposed synthetic scheme towards potential heparanase probe **91** which is resistant to processing by *exo*-glucuronidases. Reagents and conditions: a) KHMDS, 8-azido-1-iodooctane, THF; b) DDQ, THF; c) *m*-CPBA, DCM; d) TEMPO, BAIB, DCM, ^{t}BuOH, H_2O; e) Na, ^{t}BuOH, THF, NH_3; f) tag-OSu, DIPEA, DMF.

Alternatively, a stabilized heparanase 'disaccharide' probe may be constructed by introduction of a 'carbasugar' at the non-reducing end via an ether linkage (Scheme

11). For this purpose, epoxide **92** (available in multiple steps from D-glucal)[45] would be reacted with cyclohexene **87** (Chapter 3) using $Cu(OTf)_2$ as Lewis acid catalyst to regioselectively[45] afford 'disaccharide' **93** through trans-diaxial opening of the epoxide ring. The axial hydroxyl group may then be triflated and subsequently displaced by azide to afford equatorial azide **94**. Staudinger reduction, *N*-acetylation and selective cleavage of the *p*-methoxybenzyl ether using HCl in HFIP would afford cyclohexene **95**. The homoallylic alcohol may be reacted with trichloroacetonitrile under DBU catalysis, and the resulting trichloroimidate would be cyclized upon the double bond to afford **96**. Acidic hydrolysis, followed by basification could lead to aziridine **97**, which could be *N*-alkylated with 8-azidooctyl trifluoromethanesulfonate and the hydroxyl group could then be oxidized with the TEMPO/BAIB system to give **98**. Global deprotection using Birch conditions, followed by amide coupling with appropriate succinimidyl esters would finally afford stabilized 'disaccharide' heparanase ABPs **99**.

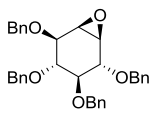


Scheme 11 Proposed synthetic scheme towards stabilized 'disaccharide' heparanase probes **99**. Reagents and conditions: a) $Cu(OTf)_2$, DCM , rt; b) Tf_2O, pyridine, DCM, -20 °C; c) NaN_3, DMF; d) PPh_3, H_2O, MeCN; e) Ac_2O, pyridine; f) HCl, HFIP, DCM; g) CCl_3CN, DBU, DCM; h) NIS, $CHCl_3$; i) HCl, MeOH, then Et_3N; j) 8-azidooctyl trifluoromethanesulfonate, DIPEA, DCM; k) TEMPO, BAIB, H_2O, ^tBuOH, DCM; l) Na, ^tBuOH, THF, NH_3; m) tag-OSu, DIPEA, DMF.

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).

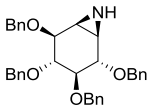
Compound 2



Cyclohexene **1**[9] (1.84 g, 3.63 mmol) was dissolved in a mixture of MeCN (24 mL), dioxane (12 mL) and aq. EDTA buffer (0.4 mM, 12 mL) and the mixture was cooled to 0 °C. 1,1,1-trifluoroacetone (4.9 mL, 54.4 mmol) was added via a pre-cooled needle, and subsequently a mixture of Oxone (11.1 g, 18.1 mmol) and $NaHCO_3$ (2.1 g, 25.4 mmol) was added in 6 portions over 1 h. The mixture was stirred for an additional 2.5 h, then diluted with H_2O (200 mL), extracted with EtOAc (3 x 100 mL) and the combined organic fractions were washed with brine, dried over $MgSO_4$, filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 25:1 → 10:1) afforded the title compound as a white solid (1.65 g, 87%). ^1H NMR (400 MHz, $CDCl_3$) δ 7.41 – 7.20 (m, 20H), 4.88 – 4.65 (m, 8H), 3.97 – 3.84 (m, 3H), 3.71 – 3.57 (m, 1H), 3.47 (dd, *J* = 10.4, 7.9 Hz, 1H), 3.31 (d, *J* = 3.2 Hz,

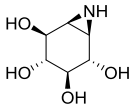
1H), 3.19 (d, *J* = 3.8 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.5, 138.1, 137.5, 128.4, 128.4, 128.2, 128.0, 127.9, 127.8, 127.7, 127.5, 127.50, 83.3, 79.2, 79.1, 78.9, 75.8, 75.4, 73.2, 73.0, 55.2, 53.9 ppm.

Compound 3



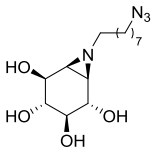
Epoxide **2** (1.55 g, 2.97 mmol) was dissolved in DMF (30 mL), sodium azide (1.93 g, 29.7 mmol) and $LiClO_4$ (6.32 g, 59.4 mmol) were added and the mixture was stirred overnight at 100 °C. The mixture was diluted with H_2O (300 mL) and extracted with Et_2O (3 x 150 mL). The combined organic fractions were washed with H_2O (200 mL) and brine, dried over $MgSO_4$, filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 5:1) gave the azido-alcohols (1.36 g) as an inseparable mixture. This mixture was co-evaporated with toluene (3x) and dissolved in dry MeCN (12 mL) under argon. Polymer-bound triphenylphosphine (~3 mmol/g, 1.6 g, 4.8 mmol) was added and the mixture was stirred overnight at 60 °C. The reaction mixture was filtered and evaporated. Flash purification by silica column chromatography (pentane/EtOAc, 3:1) afforded the title compound as a white solid (508 mg, 33% over 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.43 – 7.26 (m, 20H), 4.91 – 4.69 (m, 8H), 3.87 (d, *J* = 7.7 Hz, 2H), 3.66 (t, *J* = 9.4 Hz, 1H), 3.46 (dd, *J* = 10.3, 7.8 Hz, 1H), 2.53 (dd, *J* = 5.1, 2.8 Hz, 1H), 2.37 (d, *J* = 5.8 Hz, 1H), 0.75 (brs, NH) ppm. ^{13}C NMR (125 MHz, $CDCl_3$) δ 138.9, 138.2, 128.5, 128.4, 128.4, 128.3, 128.2, 128.0, 127.9, 127.9, 127.9, 127.7, 127.6, 127.5, 84.4 (broad, assigned by HSQC), 79.9, 79.7, 75.9, 75.4, 73.0, 34.4, 29.8 (broad, assigned by HSQC) ppm.

Compound 4



Ammonia (10 mL) was condensed in a flask at -60 °C, and lithium wire (56 mg, 8.0 mmol) was added. The resulting deep-blue solution was stirred for 30 minutes to dissolve all lithium. Aziridine **3** (104 mg, 0.2 mmol) was taken up in dry THF (1 mL) and added to the reaction mixture. After stirring for 1 h, the mixture was quenched with H_2O. The mixture was slowly warmed to rt and evaporated. The crude was dissolved in H_2O and eluted over a column packed with Amberlite CG-50 (NH_4^+) with 0.5M NH_4OH as eluent, affording the title compound as an oil (32 mg, 99%). ^{1}H NMR (400 MHz, D_2O) δ 3.87 (dt, *J* = 6.0, 3.3 Hz, 1H), 3.68 (dd, *J* = 5.6, 2.5 Hz, 1H), 3.20 (dd, *J* = 5.8, 2.6 Hz, 2H), 2.61 (dd, *J* = 6.1, 3.6 Hz, 1H), 2.34 (d, *J* = 6.2 Hz, 1H) ppm. ^{13}C NMR (101 MHz, D_2O) δ 75.6, 72.0, 71.5, 70.9, 35.4, 34.9 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_6H_{12}O_9$ 162.07608, found 162.07613. This analytical data is in accordance with the literature.[2]

Compound 5



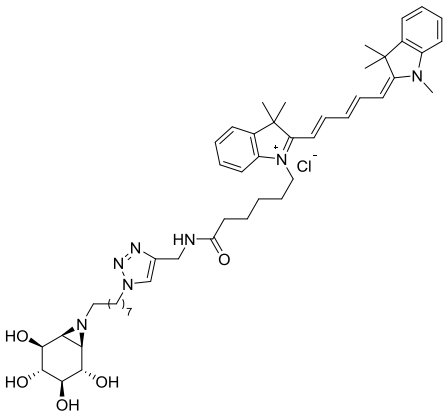
Aziridine **4** (32 mg, 0.2 mmol) was dissolved in DMF (1 mL). Potassium carbonate (38 mg, 0.28 mmol) and 1-azido-8-iodooctane (83 mg, 0.3 mmol) were added and the mixture was stirred overnight at 80 °C. H-NMR analysis of the crude reaction mixture showed mainly starting material. Therefore, additional potassium carbonate (27 mg, 0.2 mmol) and 1-azido-8-iodooctane (110 mg, 0.4 mmol) were

added and the mixture was stirred 4 h at 120 °C and subsequently concentrated. Flash purification by silica column chromatography (DCM/MeOH, 9:1) afforded the title product as an oil (25 mg, 40%). ^{1}H NMR (400 MHz, D_2O) δ 3.61 (dd, *J* = 8.4, 3.6 Hz, 1H), 3.54 (d, *J* = 8.1 Hz, 1H), 3.19 (t, *J* = 6.8 Hz, 2H), 3.14 (dd, *J* = 10.4, 8.5 Hz, 1H), 2.97 (dd, *J* = 10.4, 8.2 Hz, 1H), 2.28 (dt, *J* = 11.6, 7.3 Hz, 1H), 2.07 (dt, *J* = 11.6, 7.4 Hz, 1H), 1.84 (dd, *J* = 6.2, 3.7 Hz, 1H), 1.54 (d, *J* = 6.2 Hz, 1H), 1.53 - 1.43 (m, 4H), 1.35 - 1.22 (d, *J* = 19.0 Hz, 8H) ppm. ^{13}C NMR (101 MHz, D_2O) δ 77.8, 74.0, 73.4, 73.1, 62.0, 52.4, 45.8, 45.4, 30.6, 30.5, 30.2, 29.9, 28.3, 27.8 ppm.

General procedure for click reactions

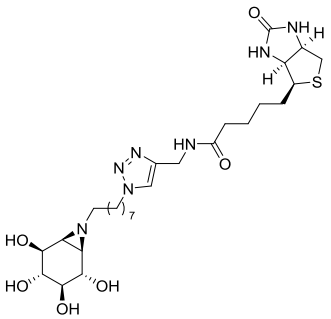
The azido compound (3-5 mg) was dissolved in DMF (0.5 mL), then the alkyne-tag (1.1 eq), $CuSO_4$ (0.2 eq) and sodium ascorbate (0.4 eq) were added and the mixture was stirred for 72 h at rt. The reaction mixture was concentrated and purified by semi-preparative reversed phase HPLC (linear gradient. Solutions used: A: 50 mM NH_4HCO_3 in H_2O, B: acetonitrile).

Compound 6 (SY-D202)



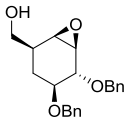
Following the general procedure starting from compound **5** (5.7 mg, 18.1 µmol), the product was obtained as a blue powder (6.3 mg, 40%). ^{1}H NMR (500 MHz, MeOD) δ 8.33 – 8.18 (m, 2H), 7.86 (s, 1H), 7.51 (d, *J* = 7.4 Hz, 2H), 7.47 – 7.38 (m, 2H), 7.36 – 7.25 (m, 4H), 6.64 (t, *J* = 12.4 Hz, 1H), 6.30 (dd, *J* = 13.8, 1.8 Hz, 2H), 4.43 (s, 2H), 4.39 (t, *J* = 7.1 Hz, 2H), 4.11 (t, *J* = 7.5 Hz, 2H), 3.71 (dd, *J* = 8.4, 3.6 Hz, 1H), 3.65 (s, 3H), 3.64 (d, *J* = 8.2 Hz, 1H), 3.24 (dd, *J* = 10.4, 8.4 Hz, 1H), 3.07 (dd, *J* = 10.4, 8.1 Hz, 1H), 2.36 (dt, *J* = 11.6, 7.3 Hz, 1H), 2.27 (t, *J* = 7.3 Hz, 2H), 2.19 – 2.11 (m, 1H), 1.93 – 1.80 (m, 5H), 1.75 (s, 12H), 1.72 (dd, *J* = 8.8, 6.4 Hz, 2H), 1.63 (d, *J* = 6.3 Hz, 1H), 1.56 (q, *J* = 7.4, 7.0 Hz, 2H), 1.49 (p, *J* = 7.7, 7.2 Hz, 2H), 1.40 – 1.28 (m, 8H) ppm. ^{13}C NMR (125 MHz, MeOD) δ 174.3, 174.0, 173.2, 154.1, 142.8, 142.1, 141.2, 141.1, 128.4, 128.3, 125.2, 124.9, 124.8, 122.7, 122.0, 121.9, 110.6, 110.4, 103.0, 102.8, 76.5, 72.7, 72.1, 71.8, 60.6, 49.9, 49.1, 44.4, 44.0, 43.4, 35.1, 34.2, 30.1, 29.9, 29.1, 29.0, 28.5, 26.8, 26.7, 26.5, 26.4, 26.0, 25.9, 25.0 ppm.

Compound 7 (SY-D203)



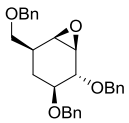
Following the general procedure starting from compound **5** (6.4 mg, 20.3 µmol), the product was obtained as a white powder (4.51 mg, 37%). ^{1}H NMR (500 MHz, MeOD) δ 7.86 (s, 1H), 4.52 (dd, *J* = 7.9, 4.2 Hz, 1H), 4.45 (s, 2H), 4.40 (t, *J* = 7.1 Hz, 2H), 4.31 (dd, *J* = 7.9, 4.5 Hz, 1H), 3.71 (dd, *J* = 8.4, 3.6 Hz, 1H), 3.65 (d, *J* = 8.1 Hz, 1H), 3.27 – 3.18 (m, 2H), 3.08 (dd, *J* = 10.4, 8.1 Hz, 1H), 2.95 (dd, *J* = 12.9, 5.0 Hz, 1H), 2.73 (d, *J* = 12.7 Hz, 1H), 2.38 (dt, *J* = 11.6, 7.2 Hz, 1H), 2.26 (t, *J* = 7.3 Hz, 2H), 2.17 (dt, *J* = 11.6, 7.3 Hz, 1H), 1.95 – 1.87 (m, 3H), 1.79 – 1.66 (m, 3H), 1.65 (d, *J* = 6.2 Hz, 1H), 1.63 – 1.55 (m, 3H), 1.48 – 1.41 (m, 2H), 1.41 – 1.27 (m, 8H) ppm. ^{13}C NMR (125 MHz, MeOD) δ 174.6, 144.9, 122.7, 76.5, 72.7, 72.1, 71.8, 61.9, 60.6, 60.2, 55.6, 49.9, 44.4, 44.0, 39.7, 35.1, 34.2, 29.9, 29.1, 29.0, 28.5, 28.3, 28.0, 26.7, 25.9, 25.3 ppm.

Compound 10



Cyclohexene **9** (3.11 mmol, 1.0 g) was dissolved in DCM. At 0 °C, *m*-CPBA (1.74 g, 7.79 mmol) was added. After stirring overnight, the reaction mixture was diluted with DCM before washing with sat. aq. $NaHCO_3$/10% $Na_2S_2O_3$ (1:1 v/v) and brine. The organic layer was dried, concentrated and purified by column chromatography (pentane/EtOAc, 2:1) to yield the title compound (970 mg, 91%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37 – 7.23 (m, 10H), 4.79 (dd, *J* = 2.5 Hz, 2H), 4.71 – 4.59 (m, 2H), 3.78 – 3.62 (m, 3H), 3.44 (ddd, *J* = 12.1, 8.0, 3.8 Hz, 1H), 3.27 (d, *J* = 3.4 Hz, 1H), 3.19 (d, *J* = 3.7 Hz, 1H), 2.15 (dt, *J* = 10.8, 5.5 Hz, 1H), 1.72 (dt, *J* = 12.9, 4.3 Hz, 1H), 1.16 (q, *J* = 12.4 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.7, 138.1, 128.6, 128.5, 128.0, 128.0, 127.8, 127.7, 79.3, 78.9, 73.2, 71.8, 64.9, 54.4, 54,0, 37.5, 23.8. IR: (neat) υ 2928, 2872, 1454, 1092, 1059, 735 cm^{-1}. HRMS (ESI): *m/z* = $[M+Na]^+$ calcd. for $C_{21}H_{24}O_4$ 363.15668, found 363.15666.

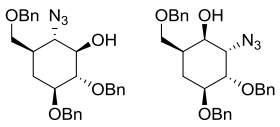
Compound 11



Epoxide **9** (2.8 mmol, 0.95 g) was dissolved in DMF under argon. At 0 °C NaH (60 wt%, 146 mg, 3.64 mmol, 1.3 equiv.) and TBAI (104 mg, 0.28 mmol) were added. After 10 minutes BnBr (466 µL, 3.92 mmol, 1.4 equiv.) was added. The reaction mixture stirred at room temperature. After reaction completion, it was quenched with water. The mixture was extracted with ether, washed with water (3x), brine and dried over $MgSO_4$. The concentrated crude product was then purified by column chromatography (pentane/EtOAc, 94:6) to obtain the title compound (0.95 g, 79%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.42 – 7.22 (m, 15H), 4.84 – 4.73 (m, 2H), 4.69 – 4.50 (m, 4H), 3.72 (d, *J* = 8.0 Hz, 1H), 3.61 – 3.53 (m, 1H), 3.47 – 3.38 (m, 2H), 3.29 (d, *J* = 3.5 Hz, 1H), 3.18 (d, *J* = 3.7 Hz, 1H), 2.30 (dq, *J* = 12.3, 6.3 Hz, 1H), 1.76 (dt, *J* = 12.9, 4.2 Hz, 1H), 1.11 (q, *J* = 12.4 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.7, 138.3, 138.2, 128.6, 128.6, 128.5, 128.0, 127.9, 127.8, 127.8, 127.8, 127.7, 79.3, 79.0, 73.5, 73.2, 72.2, 71.8, 54.6, 54.1, 35.8, 24.3. IR:

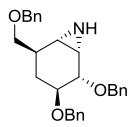
(neat) υ 2860, 1452, 1092, 1028, 733, 696 cm^{-1}. HRMS (ESI): *m/z* = $[M+Na]^+$ calcd. for $C_{28}H_{30}O_4$ 453.20363, found 453.20294.

Compound 12 and 13



Epoxide **11** (933 mg, 2.2 mmol) was dissolved in DMF under argon. NaN_3 (21.7 mmol, 1.4 g, 10 equiv.) and $LiClO_4$ (43.4 mmol, 4.6 g, 20 equiv.) were added. The reaction mixture was stirred at 80 °C for two days. The completed reaction was quenched with water, extracted with ether and washed with water and brine. The crude product, after drying over $MgSO_4$ and concentrating, was purified by column chromatography (pentane/EtOAc, 1:19 → 1:9) resulting in a mixture of the two *trans*-azido alcohols **12**-**13** in a ratio of 1:1.3 (853 mg, 83%). **12:** ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.27 (m, 15H), 5.04 (d, *J* = 11.3 Hz, 1H), 4.70 (dd, *J* = 11.4, 5.2 Hz, 2H), 4.60 (d, *J* = 11.5 Hz, 1H), 4.52 (s, 2H), 3.57 (dd, *J* = 9.1, 4.5 Hz, 1H), 3.53 (d, *J* = 2.2 Hz, 1H), 3.50 (dd, *J* = 4.6, 1.8 Hz, 1H), 3.47 (s, 1H), 3.45 (d, *J* = 2.0 Hz, 1H), 3.35 (q, *J* = 9.3 Hz, 2H), 2.62 (d, *J* = 2.0 Hz, 1H), 2.20 (dd, *J* = 9.9, 4.6 Hz, 1H), 1.51 (s, 1H), 1.49 (d, *J* = 3.3 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.6, 138.4, 138.3, 128.8, 128.6, 128.6, 128.2, 128.1, 127.8, 127.8, 127.7, 85.1, 79.7, 76.4, 75.6, 73.4, 72.0, 70.4, 64.3, 37.8, 31.2. IR: (neat) υ 2924, 2859, 2099, 1452, 1086, 1070, 733, 696 cm^{-1}. HRMS (ESI): *m/z* = $[M+H]^+$ calcd. for $C_{28}H_{31}N_3O_4$ 474.23873, found 474.23836. **13:** ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.41 – 7.26 (m, 14H), 4.81 (d, *J* = 11.7 Hz, 1H), 4.77 – 4.62 (m, 3H), 4.56 – 4.43 (m, 2H), 4.02 (s, 1H), 3.98 – 3.94 (m, 2H), 3.82 – 3.73 (m, 1H), 3.70 (dd, *J* = 9.3, 2.5 Hz, 1H), 3.58 (dd, *J* = 9.3, 3.5 Hz, 1H), 3.45 (d, *J* = 1.3 Hz, 1H), 1.94 (s, 1H), 1.90 (d, *J* = 22.9 Hz, 1H), 1.78 (dd, *J* = 8.9, 4.9 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.1, 138.6, 137.5, 128.7, 128.5, 128.5, 128.1, 127.9, 127.7, 127.7, 127.6, 80.5, 77.6, 73.8, 73.5, 73.4, 72.7, 72.2, 63.9, 35.0, 28.3. IR: (neat) υ 2924, 2864, 2102, 1454, 1090, 1076, 731, 696 cm^{-1}. HRMS (ESI): *m/z* = $[M+H]^+$ calcd. for $C_{28}H_{31}N_3O_4$ 474.23873, found 474.23849.

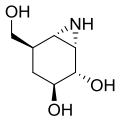
Compound 14



Trans-azido alcohol **13** (0.54 mmol, 254 mg) was co-evaporated with toluene (3x) and dissolved in dry acetonitrile under argon. Polymer-bound triphenyl phosphine (717 mg, 2.2 mmol, 4 equiv.) was added to a dried closed tube. The resin was rinsed with anhydrous acetonitrile (3x) before the starting material was transferred to the closed tube. The reaction mixture was stirred at 90 °C overnight. The mixture was then allowed to cool to room temperature, filtrated and concentrated at room temperature. The crude product was purified by column chromatography (DCM/MeOH, 200:1 → 50:1) to afford the title product (168 mg, 73%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.44 – 7.23 (m, 16H), 4.82 (p, *J* = 12.4 Hz, 2H), 4.71 – 4.59 (m, 2H), 4.59 – 4.44 (m, 2H), 3.73 (dd, *J* = 8.0, 3.5 Hz, 1H), 3.65 (ddd, *J* = 11.4, 8.0, 3.3 Hz, 1H), 3.50 – 3.37 (m, 2H), 2.48 (dd, *J* = 6.0, 3.5 Hz, 1H), 2.30 (d, *J* = 6.1 Hz, 1H), 2.21 (dt, *J* = 12.2, 6.1 Hz, 1H), 1.84 (ddd, *J* = 12.8, 5.3, 3.4 Hz, 1H), 1.14 (q, *J* = 12.0 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.3, 139.1, 138.4, 128.6, 128.4, 128.4, 127.9, 127.8, 127.8, 127.6, 127.5, 80.4, 77.6, 73.4, 73.3, 72.6, 72.4, 36.3, 35.0, 34.7, 31.8.

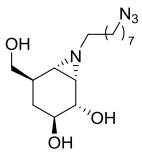
IR: (neat) υ 2857, 1450, 1094, 1074, 735, 696 cm^{-1}. HRMS (ESI): *m/z* = $[M+H]^+$ calcd. for $C_{28}H_{31}NO_3$ 430.23767, found 430.23746.

Compound 15



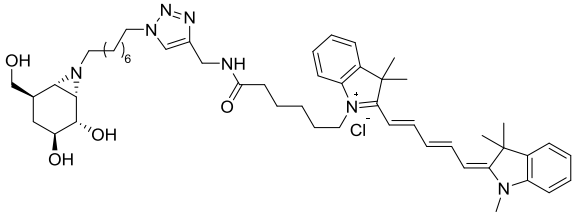
Ammonia was condensed at -60 °C and lithium (1.5 mmol, 11 mg, 30 equiv.) was added. The mixture was stirred for 30 minutes. Then, the aziridine **14** (22 mg, 50 µmol) in THF (1 mL) was added and stirred at -60 °C for 1 hour. The reaction was quenched with D_2O and concentrated. The crude product was purified by column chromatography using Amberlite CG-50 (NH_4^+) with 0.5*N* NH_4OH, affording the title compound as an oil (9 mg, quant.). ^1H NMR (500 MHz, D_2O) δ 3.76 (dd, *J* = 8.7, 3.6 Hz, 1H), 3.64 (dd, *J* = 11.0, 5.7 Hz, 1H), 3.58 (dd, *J* = 11.0, 7.0 Hz, 1H), 3.52 (ddd, *J* = 12.2, 8.7, 3.5 Hz, 1H), 2.58 (dd, *J* = 6.4, 3.7 Hz, 1H), 2.35 (d, *J* = 6.4 Hz, 1H), 2.05 (dq, *J* = 12.2, 6.1 Hz, 1H), 1.67 (dt, *J* = 12.4, 4.4 Hz, 1H), 1.14 (q, *J* = 12.4 Hz, 1H). ^{13}C NMR (126 MHz, D_2O) δ 73.2, 69.5, 64.5, 37.1, 36.5, 34.4, 33.2. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_7H_{14}NO_3$ 160.09682, found 160.09662.

Compound 16



Aziridine **14** (71 mg, 0.166 mmol) was debenzylated and purified as described above, and subsequently dissolved in DMF-d_7 (1.6 mL). K_2CO_3 (0.2 mmol, 27 mg, 1.2 equiv.) and 1-azido-8-iodooctane (0.33 mmol, 93 mg, 2 equiv.) were added. The reaction mixture was stirred overnight at 100 °C. Then, the mixture was concentrated and purified by column chromatography (DCM/MeOH, 20:1 → 9:1) affording the title compound as an oil (29 mg, 55% over 2 steps). ^1H NMR (400 MHz, MeOD) δ 3.60 – 3.57 (m, 1H), 3.57 – 3.50 (m, 2H), 3.47 (dt, *J* = 10.8, 5.5 Hz, 1H), 3.28 (t, *J* = 6.8 Hz, 2H), 2.33 (dt, *J* = 11.7, 7.5 Hz, 1H), 2.16 (dt, *J* = 11.6, 7.5 Hz, 1H), 2.00 (dt, *J* = 12.5, 5.7 Hz, 1H), 1.87 (dd, *J* = 6.4, 3.6 Hz, 1H), 1.70 (d, *J* = 6.4 Hz, 1H), 1.59 (dq, *J* = 13.9, 6.9, 6.1 Hz, 5H), 1.35 (s, 8H), 1.00 (q, *J* = 12.2 Hz, 1H). ^{13}C NMR (101 MHz, MeOD) δ 75.1, 71.3, 66.1, 62.3, 52.4, 46.9, 44.8, 39.6, 35.3, 30.6, 30.4, 30.2, 29.9, 28.3, 27.8. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{15}H_{29}N_4O_3$ 313.22342, found 313.22342.

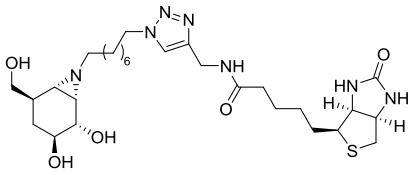
Compound 17 (JS-97)



Following the general procedure starting from compound **16** (4.8 mg, 15 µmol) the title compound was obtained as a blue solid (6.7 mg, 44%). ^1H NMR (500 MHz, D_2O) δ 7.79 (s, 2H), 7.34 (t, *J* = 8.1 Hz, 2H), 7.23 (dt, *J* = 25.5, 8.0 Hz, 3H), 7.09 (dt, *J* = 19.1, 7.5 Hz, 3H), 6.32 (t, *J* = 12.2 Hz, 1H), 6.00 (dd, *J* = 32.4, 13.7 Hz, 2H), 4.35 (s, 2H), 4.23 (t, *J* = 6.9 Hz, 2H), 3.92 (s, 1H), 3.68 (dd, *J* = 8.7, 3.7 Hz, 1H), 3.56 (dd, *J* = 10.9, 5.5 Hz, 1H), 3.52 (s, 2H), 3.51 – 3.43 (m, 3H), 2.22 (t, *J* = 7.0 Hz, 2H), 2.15 (t, *J* = 7.4 Hz, 2H), 2.05 (s, 8H), 2.02 – 1.93 (m, 2H), 1.92 (s, 5H), 1.77 (d, *J* = 6.8 Hz, 1H), 1.67 (s, 3H), 1.63 – 1.55 (m, 3H), 1.41 (d, *J* = 10.3 Hz, 12H),

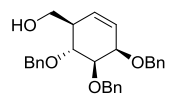
1.06 (d, *J* = 12.7 Hz, 3H), 1.03 (s, 5H). ^{13}C NMR (126 MHz, D_2O) δ 176.0, 173.8, 172.7, 153.2, 142.6, 141.9, 141.1, 128.5, 125.3, 125.0, 124.3, 122.2, 110.7, 103.2, 102.6, 72.7, 70.0, 64.3, 59.7, 50.3, 49.0, 48.9, 45.8, 44.1, 43.5, 36.9, 35.4, 34.3, 33.1, 30.9, 29.5, 28.5, 28.4, 28.1, 27.0, 26.8, 26.6, 26.4, 25.6, 25.5, 25.1, 0.9. HRMS (ESI) m/z: $[M]^+$ calc for $C_{50}H_{70}N_7O_4^+$ 832.54838, found 832.54852.

Compound 18 (JS-96)



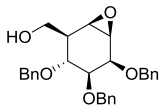
Following the general procedure starting from compound **16** (4.8 mg, 15 µmol), the title compound was obtained as a white solid (5.1 mg, 56%). ^{1}H NMR (400 MHz, MeOD) δ 7.86 (s, 1H), 4.54 (dd, *J* = 7.9, 4.5 Hz, 2H), 4.43 (s, 2H), 4.38 (t, *J* = 7.0 Hz, 2H), 4.33 (dd, *J* = 7.9, 4.5 Hz, 1H), 3.63 (dd, *J* = 8.5, 3.8 Hz, 1H), 3.58 (dd, *J* = 10.9, 5.6 Hz, 1H), 3.55 – 3.45 (m, 2H), 3.22 (dt, *J* = 9.8, 5.4 Hz, 1H), 2.95 (dd, *J* = 12.9, 5.0 Hz, 1H), 2.72 (d, *J* = 12.9 Hz, 1H), 2.24 (dt, *J* = 14.2, 7.4 Hz, 4H), 2.01 (dq, *J* = 12.4, 5.8 Hz, 1H), 1.93 (dd, *J* = 6.5, 3.9 Hz, 1H), 1.88 (d, *J* = 8.2 Hz, 4H), 1.73 (d, *J* = 6.5 Hz, 1H), 1.65 (td, *J* = 13.8, 6.3 Hz, 5H), 1.55 (dt, *J* = 15.5, 8.1 Hz, 5H), 1.39 (p, *J* = 8.5, 7.9 Hz, 3H), 1.29 (s, 5H), 1.03 (q, *J* = 12.4 Hz, 1H). ^{13}C NMR (101 MHz, MeOD) δ 176.6, 166.1, 146.1, 124.4, 74.8, 71.2, 65.7, 63.3, 61.9, 61.5, 56.8, 51.4, 46.8, 44.8, 41.0, 39.0, 36.5, 35.4, 34.9, 31.0, 30.2, 30.1, 29.7, 29.4, 29.2, 28.0, 27.1, 26.5. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{28}H_{48}N_7O_5S$ 594.34321, found 594.34332.

Compound 64



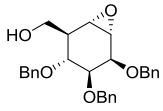
Diol **63**[13] (400 mg, 1.18 mmol) was dissolved in dry DMF (6 mL), imidazole (200 mg, 2.94 mmol) was added and the mixture was cooled to -15 °C (ice/EtOH). Then, TBSCl (186 mg, 1.23 mmol) was added and the mixture was stirred for 2 h while the cooling bath was allowed to warm to rt. The mixture was quenched and diluted with H_2O (120 mL) and extracted with Et_2O (3 x 50 mL). The combined organic layers were washed with H_2O (100 mL) and brine, dried over $MgSO_4$, filtrated and concentrated. The crude product was co-evaporated with toluene, dissolved in dry DMF (6 mL) and cooled to 0 °C. Then, TBAI (43 mg, 0.12 mmol), BnBr (280 µL, 2.35 mmol) and NaH (60 wt%, 89 mg, 2.23 mmol) were added and the mixture was stirred overnight at rt. The reaction was quenched and diluted with H_2O (60 mL) at 0 °C, and extracted with Et_2O (3 x 30 mL). The combined organic layers were washed with H_2O (30 mL) and brine, dried over $MgSO_4$, filtrated and concentrated. The crude product was co-evaporated with toluene, dissolved in THF (6 mL) and TBAF (1M in THF, 3.5 mL, 3.5 mmol) was added. After 2 h, the mixture was concentrated, and flash purification by silica column chromatography (pentane/EtOAc, 3:1 → 2:1) afforded the title compound as an oil (451 mg, 89%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.42 – 7.22 (m, 15H), 5.85 (ddd, *J* = 9.9, 4.6, 2.7 Hz, 1H), 5.66 (dd, *J* = 10.0, 2.5 Hz, 1H), 4.96 (d, *J* = 11.2 Hz, 1H), 4.80 – 4.62 (m, 5H), 4.13 (t, *J* = 4.0 Hz, 1H), 3.97 (dd, *J* = 8.9, 7.3 Hz, 1H), 3.76 – 3.63 (m, 3H), 2.46 – 2.36 (m, 1H), 2.09 (brs, OH) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.9, 138.6, 138.5, 130.5, 128.6, 128.5, 128.5, 128.3, 128.0, 127.9, 127.8, 127.7, 126.9, 80.7, 77.0, 74.5, 72.7, 72.0, 71.8, 64.5, 46.5 ppm. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{28}H_{30}O_4$ 453.2036, found 453.2049.

Compound 65



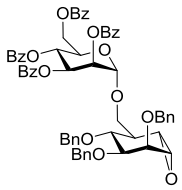
Cyclohexene **64** (451 mg, 1.05 mmol) was dissolved in DCM (10 mL) and cooled to 0 °C on a large ice-bath. Then, mCPBA (<77% wt, 470 mg, 2.1 mmol) was added and the mixture was stirred overnight while the cooling bath was slowly allowed to reach rt. The mixture was diluted with DCM (100 mL) and washed with a mixture of sat. aq. $NaHCO_3$ and aq. 10% $Na_2S_2O_3$ (1:1 v/v, 3 x 50 mL) and brine, dried over $MgSO_4$, filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 2:1 → 1:1) afforded the title compound as a white solid (393 mg, 84%). Additionally, α-epoxide **66** was obtained as an oil (36 mg, 8%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.48 – 7.23 (m, 15H), 4.88 (d, *J* = 11.0 Hz, 1H), 4.83 (d, *J* = 12.4 Hz, 1H), 4.72 (d, *J* = 12.4 Hz, 1H), 4.66 – 4.58 (m, 2H), 4.52 (d, *J* = 11.0 Hz, 1H), 4.03 (t, *J* = 4.6 Hz, 1H), 3.94 (dd, *J* = 10.7, 5.2 Hz, 1H), 3.88 (dd, *J* = 10.7, 6.2 Hz, 1H), 3.75 (t, *J* = 8.5 Hz, 1H), 3.45 (dd, *J* = 9.1, 4.9 Hz, 1H), 3.32 (dd, *J* = 3.7, 2.5 Hz, 1H), 3.20 (t, *J* = 4.1 Hz, 1H), 2.58 – 2.29 (brs, OH), 2.14 (ddt, *J* = 8.5, 5.4, 3.1 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.3, 138.2, 138.1, 128.5, 128.5, 128.4, 128.2, 128.2, 128.0, 127.9, 127.8, 80.0, 74.7, 74.2, 72.6, 71.5, 70.5, 63.0, 54.9, 51.0, 44.3 ppm. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{28}H_{30}O_5$ 469.1985, found 469.1990.

Compound 66



Cyclohexene **64** (326 mg, 0.76 mmol) was dissolved in a mixture of MeCN (7.6 mL) and aq. EDTA (0.4 mM, 3.8 mL) and cooled to 0 °C. 1,1,1-trifluoroacetone (1.0 mL, 11.4 mmol) was added via a pre-cooled needle and subsequently a mixture of Oxone (2.33 g, 3.79 mmol) and $NaHCO_3$ (445 mg, 5.3 mmol) was added in 6 portions over 1 h. After stirring an additional hour, the mixture was diluted with H_2O (100 mL) and extracted with EtOAc (3 x 50 mL). The combined organic fractions were washed with brine, dried over $MgSO_4$, filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 4:1 → 2:1) afforded the title compound as an oil (302 mg, 89%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39 – 16 (m, 15H), 4.90 (dd, *J* = 18.0, 11.7 Hz, 2H), 4.73 – 4.53 (m, 4H), 4.22 (t, *J* = 2.6 Hz, 1H), 3.81 – 3.75 (m, 1H), 3.75 – 3.66 (m, 2H), 3.61 (dd, *J* = 10.8, 5.9 Hz, 1H), 3.19 (t, *J* = 3.1 Hz, 1H), 3.11 (d, *J* = 3.7 Hz, 1H), 2.27 (brs, OH), 2.19 (dt, *J* = 7.8, 5.7 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.5, 138.4, 138.3, 128.5, 128.5, 128.4, 128.2, 127.9, 127.9, 127.8, 127.7, 127.7, 80.2, 74.8, 74.7, 74.0, 73.5, 73.1, 62.6, 54.3, 54.2, 44.3 ppm. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{28}H_{30}O_5$ 469.1985, found 469.1996.

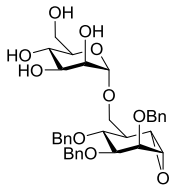
Compound 68



Acceptor **66** (45 mg, 0.1 mmol) and trichloroimidate donor **67**[42] (89 mg, 0.12 mmol) were combined in a flask, co-evaporated with toluene (3x) and dissolved in dry DCM (1 mL). The mixture was cooled to -40 °C and TMSOTf (5.4 µL, 30 µmol) was added. After stirring for 1 h, the reaction was quenched with Et_3N (50 µL) and warmed to rt. The mixture was diluted with H_2O (30 mL) and extracted with DCM (3 x 15 mL). The combined organic fractions were washed with brine,

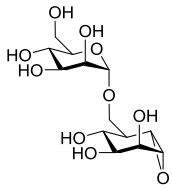
dried over $MgSO_4$, filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 5:1) afforded the title compound as an oil (67 mg, 65%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.07 (ddd, *J* = 12.9, 8.3, 1.2 Hz, 4H), 7.92 (dd, *J* = 8.3, 1.2 Hz, 2H), 7.82 (dd, *J* = 8.3, 1.2 Hz, 2H), 7.62 – 7.57 (m, 1H), 7.56 – 7.51 (m, 1H), 7.50 – 7.46 (m, 1H), 7.45 – 7.22 (m, 26H), 6.13 (t, *J* = 10.0 Hz, 1H), 5.90 (dd, *J* = 10.2, 3.3 Hz, 1H), 5.73 (dd, *J* = 3.2, 1.8 Hz, 1H), 5.08 (d, *J* = 1.7 Hz, 1H), 4.97 (dd, *J* = 11.9, 4.1 Hz, 2H), 4.77 (d, *J* = 12.3 Hz, 1H), 4.72 (d, *J* = 11.7 Hz, 1H), 4.63 (dd, *J* = 21.5, 11.6 Hz, 3H), 4.44 – 4.36 (m, 2H), 4.28 (t, *J* = 2.6 Hz, 1H), 3.91 (dd, *J* = 9.9, 7.5 Hz, 1H), 3.85 – 3.80 (m, 1H), 3.77 (dd, *J* = 9.6, 3.0 Hz, 2H), 3.29 (t, *J* = 3.1 Hz, 1H), 3.27 (d, *J* = 3.5 Hz, 1H), 2.44 (td, *J* = 7.8, 3.9 Hz, 1H) ppm. ^{13}C NMR (125 MHz, $CDCl_3$) δ 166.2, 165.6, 165.5, 165.5, 138.7, 138.6, 133.6, 133.5, 133.3, 133.2, 130.0, 129.9, 129.9, 129.4, 129.2, 129.1, 128.7, 128.6, 128.6, 128.4, 128.1, 128.0, 127.9, 127.9, 127.8, 98.0, 80.4, 75.0, 74.6, 74.0, 73.3, 73.2, 70.5, 70.1, 69.4, 68.3, 66.9, 62.8, 54.5, 54.2, 42.4 ppm. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{62}H_{56}O_{14}$ 1047.3562, found 1047.3627.

Compound 69



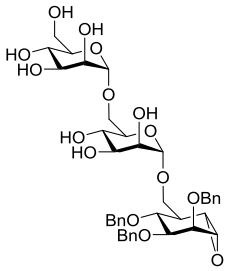
Compound **68** (56 mg, 55 µmol) was dissolved in a mixture of MeOH (0.5 mL) and DCM (0.5 mL), NaOMe (5.4 M in MeOH, 5.1 µL, 27 µmol) was added and the mixture was stirred overnight at rt. The mixture was concentrated and flash purification by silica column chromatography (DCM/MeOH, 19:1) afforded the title compound as an oil (27 mg, 81%). ^{1}H NMR (500 MHz, $CDCl_3$ + MeOD) δ 7.41 – 7.24 (m, 15H), 4.89 (d, *J* = 11.5 Hz, 2H), 4.80 (m, under solvent peak, assigned by HSQC, 1H), 4.72 (d, *J* = 12.0 Hz, 1H), 4.63 (d, *J* = 2.8 Hz, 2H), 4.61 – 4.56 (m, 1H), 4.28 (t, *J* = 2.7 Hz, 1H), 3.88 – 3.81 (m, 2H), 3.81 – 3.72 (m, 4H), 3.69 (ddd, *J* = 9.5, 6.0, 3.1 Hz, 2H), 3.59 (dd, *J* = 9.8, 3.9 Hz, 1H), 3.54 (ddd, *J* = 9.5, 4.5, 2.7 Hz, 1H), 3.23 (t, *J* = 3.1 Hz, 1H), 3.16 (d, *J* = 3.6 Hz, 1H), 2.27 (td, *J* = 7.8, 3.9 Hz, 1H) ppm. ^{13}C NMR (125 MHz, $CDCl_3$ + MeOD) δ 137.9, 137.8, 137.8, 127.7, 127.6, 127.6, 127.3, 127.3, 127.1, 127.1, 126.9, 126.9, 99.8, 79.5, 74.2, 74.0, 73.1, 72.8, 72.5, 72.1, 70.8, 70.2, 66.5, 66.1, 60.8, 53.5 (2 C), 41.8 ppm. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{34}H_{40}O_{10}$ 631.2514, found 631.2531.

Compound 70 (SY-G42)



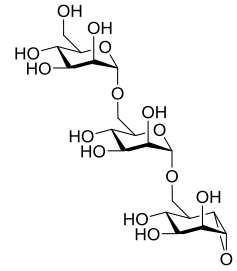
Compound **69** (20 mg, 33 µmol) was dissolved in a mixture of dioxane/MeOH/H_2O (1:1:1, 1.5 mL) under N_2. $Pd(OH)_2$/C (20 wt%, 12 mg, 16 µmol) was added and the mixture was purged with H_2 gas with a balloon. After stirring vigorously for 2 h, the mixture was filtrated over Celite and evaporated, affording the title compound as an oil (12 mg, quant.). ^{1}H NMR (400 MHz, D_2O) δ 4.89 (d, *J* = 1.6 Hz, 1H), 4.41 (t, *J* = 2.9 Hz, 1H), 3.97 – 3.90 (m, 2H), 3.89 – 3.78 (m, 2H), 3.76 – 3.68 (m, 2H), 3.65 – 3.57 (m, 3H), 3.51 (dd, *J* = 10.4, 3.5 Hz, 1H), 3.40 – 3.36 (m, 1H), 3.31 (d, *J* = 3.5 Hz, 1H), 2.14 (ddd, *J* = 9.4, 6.3, 3.3 Hz, 1H) ppm. ^{13}C NMR (101 MHz, D_2O) δ 99.7, 72.9, 70.5, 70.3, 70.0, 67.1, 66.7, 66.1, 65.6, 60.9, 55.8, 55.2, 42.3 ppm. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{13}H_{22}O_{10}$ 361.1105, found 361.1118.

Compound 72



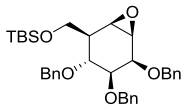
Acceptor **66** (62 mg, 0.14 mmol) and thioglycoside donor **71**[43] (194 mg, 0.17 mmol) were combined in a flask, co-evaporated with toluene (3x) and dissolved in dry DCM (1.4 mL). Molecular sieves (4 Å) were added and the mixture was stirred for 1 h at rt, and subsequently cooled to -40 °C. Then, NIS (36 mg, 0.17 mmol) and TMSOTf (7.5 µL, 42 µmol) were added. After stirring for 1 h, the reaction was quenched with Et_3N (50 µL) and warmed to rt. The mixture was diluted DCM (50 mL), washed with 10% aq. $Na_2S_2O_3$ (2 x 20 mL) and brine, dried over $MgSO_4$, filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 2:1) afforded the target trisaccharide which was highly contaminated with inseparable byproducts. The product was taken up in a mixture of MeOH (0.5 mL) and DCM (0.5 mL) and 5 drops of NaOMe (5.4 M in MeOH) were added. After stirring the reaction overnight, the reaction was neutralized by addition of Amberlite CG-50 (H^+), filtrated and concentrated. Flash purification by silica column chromatography (DCM/MeOH, 19:1 → 15:85) afforded the title compound as an oil (15 mg, 14%). ^{1}H NMR (400 MHz, MeOD) δ 7.44 – 7.18 (m, 15H), 4.87 (d, *J* = 2.4 Hz, 1H), 4.79 (d, *J* = 1.5 Hz, 1H), 4.78 – 4.70 (m, 2H), 4.70 – 4.55 (m, 4H), 4.30 (t, *J* = 2.8 Hz, 1H), 3.90 (dd, *J* = 11.2, 5.4 Hz, 1H), 3.88 – 3.79 (m, 4H), 3.78 (d, *J* = 8.5 Hz, 1H), 3.75 – 3.68 (m, 4H), 3.68 – 3.61 (m, 5H), 3.57 (dd, *J* = 9.9, 3.7 Hz, 1H), 3.24 (t, *J* = 3.1 Hz, 1H), 3.15 (d, *J* = 3.6 Hz, 1H), 2.22 (td, *J* = 7.6, 3.8 Hz, 1H) ppm. ^{13}C NMR (101 MHz, MeOD) δ 140.0, 139.9, 139.8, 129.5, 129.4, 129.4, 129.2, 129.1, 129.0, 128.8, 128.7, 128.7, 101.9, 101.4, 81.4, 76.0, 75.8, 74.7, 74.4, 74.3, 73.7, 73.5, 72.8, 72.6, 72.0, 72.0, 68.6, 68.3, 67.9, 67.0, 62.8, 55.4, 55.2, 43.9 ppm. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{40}H_{50}O_{15}$ 793.3042, found 793.3082.

Compound 73 (SY-G39)



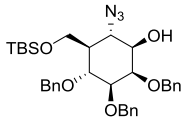
Compound **72** (10 mg, 13 µmol) was dissolved in a mixture of dioxane/MeOH/H_2O (1:1:1, 1.0 mL) under N_2. $Pd(OH)_2$/C (20 wt%, 10 mg, 14 µmol) was added and the mixture was purged with H_2 gas with a balloon. After stirring vigorously for 2 h, the mixture was filtrated over celite and evaporated, affording the title compound as an oil (6.9 mg, quant.). ^{1}H NMR (500 MHz, D_2O) δ 4.91 (s, 2H), 4.44 (t, *J* = 2.9 Hz, 1H), 4.02 – 3.91 (m, 4H), 3.89 (d, *J* = 10.8 Hz, 1H), 3.85 (t, *J* = 3.4 Hz, 1H), 3.83 (t, *J* = 3.1 Hz, 1H), 3.80 – 3.64 (m, 6H), 3.68 – 3.60 (m, 2H), 3.53 (dd, *J* = 10.4, 3.4 Hz, 1H), 3.41 (t, *J* = 3.0 Hz, 1H), 3.34 (d, *J* = 3.6 Hz, 1H), 2.17 (ddd, *J* = 9.2, 6.2, 3.2 Hz, 1H) ppm. ^{13}C NMR (125 MHz, D_2O) δ 99.9, 99.5, 72.8, 71.1, 70.9, 70.6, 70.4, 70.0, 67.2, 66.8, 66.6, 66.2, 65.7, 65.6, 61.0, 55.9, 55.3, 42.3 ppm. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{19}H_{32}O_{15}$ 523.1633, found 523.1649.

Compound 74



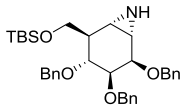
Epoxide **65** (393 mg, 0.88 mmol) was dissolved in dry THF (9 mL), cooled to 0 °C and then imidazole (120 mg, 1.76 mmol) and TBSCl (160 mg, 1.06 mmol) were added and the mixture was stirred for 4 h at rt. The reaction was quenched by addition of H_2O (100 mL) and extracted with Et_2O (3 x 50 mL). The combined organic phases were washed with brine, dried over $MgSO_4$, filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 10:1) afforded the title compound as a colorless oil (459 mg, 93%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.47 – 7.21 (m, 15H), 4.84 (dd, *J* = 11.8, 9.4 Hz, 2H), 4.73 (d, *J* = 12.5 Hz, 1H), 4.63 (s, 2H), 4.46 (d, *J* = 11.3 Hz, 1H), 4.03 (t, *J* = 4.6 Hz, 1H), 3.90 (dd, *J* = 9.5, 4.9 Hz, 1H), 3.73 (t, *J* = 9.4 Hz, 1H), 3.57 (t, *J* = 8.4 Hz, 1H), 3.47 – 3.40 (m, 2H), 3.26 (t, *J* = 4.0 Hz, 1H), 2.14 (dddd, *J* = 9.4, 7.7, 4.9, 2.5 Hz, 1H), 0.88 (s, 9H), 0.04 (s, 6H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.6, 138.5, 138.3, 128.5, 128.5, 128.4, 128.3, 128.2, 128.0, 127.9, 127.8, 127.7, 80.0, 74.5, 73.4, 72.6, 71.5, 70.8, 62.2, 54.5, 51.9, 44.9, 26.1, 18.4, -5.3 ppm. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{34}H_{44}O_5Si$ 583.2850, found 583.2861.

Compound 75



Epoxide **74** (441 mg, 0.79 mmol) was dissolved in DMF (8 mL). Sodium azide (510 mg, 7.9 mmol) and $LiClO_4$ (1.67 g, 15.7 mmol) were added and the mixture was stirred overnight at 80 °C. The mixture was diluted with H_2O (100 mL) and extracted with Et_2O (3 x 50 mL). The combined organic phases were washed with brine, dried over $MgSO_4$, filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 15:1) afforded the title compound as a colorless oil (234 mg, 49%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.43 – 7.12 (m, 15H), 4.69 – 4.41 (m, 6H), 4.15 – 4.04 (m, 1H), 3.96 – 3.87 (m, 2H), 3.77 – 3.67 (m, 4H), 3.66 (dd, *J* = 8.5, 2.6 Hz, 1H), 0.87 (s, 9H), 0.04 (s, 3 H), 0.03 (s, 3 H) ppm. ^{13}C NMR (100 MHz, $CDCl_3$) δ 138.0, 137.8, 128.5, 128.4, 128.1, 127.9, 127.8, 127.6, 78.6, 76.7 (under solvent peak, assigned by HSQC), 74.8, 73.2, 73.1, 72.5 (broad), 70.8, 63.9 (broad), 61.2, 44.6, 25.9, 18.1, -5.5, -5.5 ppm. HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{34}H_{45}N_3O_5Si$ 626.3021, found 626.3023.

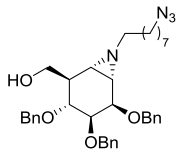
Compound 76



Polymer-bound triphenylphosphine (~3 mmol/g loading, 242 mg, 0.73 mmol) was added to a flame dried microwave tube and rinsed with dry MeCN (3x). Azido-alcohol **75** (219 mg, 0.36 mmol) was co-evaporated with toluene (3x), dissolved in dry MeCN (3.6 mL) and transferred to the microwave tube under nitrogen atmosphere. The tube was fitted with a cap and the reaction was stirred overnight at 90 °C. The mixture was then filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 3:1) afforded the title compound as a colorless oil (172 mg, 85%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.43 – 7.20 (m, 15H), 4.93 (d, *J* = 12.3 Hz, 1H), 4.87 (d, *J* = 11.4 Hz, 1H), 4.71 (d, *J* = 13.0 Hz, 2H), 4.64 (d, *J* = 11.8 Hz, 1H), 4.55 (d, *J* = 11.4 Hz, 1H), 4.20 (t, *J* = 2.3 Hz, 1H), 3.78 – 3.73 (m, 2H), 3.72 – 3.66 (m, 1H), 3.61 (dd, *J* = 9.7, 8.3 Hz, 1H), 2.42 (dd, *J* = 5.7, 2.2 Hz, 1H), 2.32 (d, *J* = 5.8 Hz, 1H),

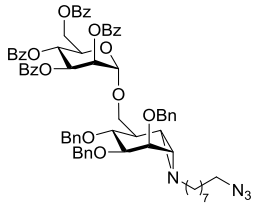
2.11 (td, *J* = 8.0, 4.1 Hz, 1H), 0.88 (s, 9H), 0.04 (s, 3H), 0.03 (s, 3H) ppm. ^{13}C NMR (100 MHz, $CDCl_3$) δ 139.3, 139.2, 128.4, 128.4, 128.1, 127.8, 127.7, 127.5, 127.5, 75.6, 75.2, 74.4, 73.3, 73.0, 63.7, 45.6, 34.3, 31.7, 26.0, 18.4, -5.2, -5.2 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{34}H_{45}NO_4Si$ 560.3191, found 560.3203.

Compound 77



Freshly prepared 8-azidooctyl trifluoromethanesulfonate (0.5 M in $CHCl_3$, 1.1 mL, 0.57 mmol) was added to a solution of aziridine **76** (159 mg, 0.28 mmol) and DIPEA (99 µL, 0.57 mmol) in $CHCl_3$ (2.8 mL) and the mixture was stirred overnight at rt. The reaction was quenched by addition of MeOH (3 mL) and the mixture was stirred for 3 h at rt. The mixture was concentrated, dissolved in EtOAc (150 mL) and washed with H_2O (3 x 75 mL) and brine, dried over $MgSO_4$, filtrated and concentrated. The crude was dissolved in dry THF (2.8 mL), TBAF (1M in THF, 0.85 mL, 0.85 mmol) was added and the mixture was stirred overnight at rt. The mixture was diluted with H_2O (200 mL) and extracted with EtOAc (3 x 50 mL). The combined organic phases were washed with H_2O (50 mL) and brine, dried over $MgSO_4$, filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 4:1) afforded the title compound as a colorless oil (146 mg, 88%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.41 – 7.20 (m, 15H), 4.89 (dd, *J* = 11.7, 9.3 Hz, 2H), 4.73 (d, *J* = 12.2 Hz, 1H), 4.70 (d, *J* = 11.7 Hz, 1H), 4.64 (d, *J* = 11.7 Hz, 1H), 4.57 (d, *J* = 11.3 Hz, 1H), 4.18 (t, *J* = 2.5 Hz, 1H), 3.79 – 3.75 (m, 2H), 3.74 – 3.70 (m, 1H), 3.63 (dt, *J* = 10.3, 5.0 Hz, 1H), 3.23 (t, *J* = 6.9 Hz, 2H), 2.24 – 2.10 (m, 4H), 1.76 (dd, *J* = 6.0, 2.4 Hz, 1H), 1.61 – 1.54 (m, 2H), 1.50 (d, *J* = 6.0 Hz, 1H), 1.48 – 1.40 (m, 2H), 1.39 – 1.22 (m, 8H) ppm. ^{13}C NMR (125 MHz, $CDCl_3$) δ 138.9, 138.8, 128.4, 128.4, 128.2, 127.8, 127.6, 127.6, 127.6, 127.5, 81.7, 76.4, 74.7, 74.5, 73.2, 72.9, 64.7, 60.7, 51.5, 45.3, 42.7, 40.3, 29.7, 29.5, 29.1, 28.9, 27.3, 26.7 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{36}H_{46}N_4O_4$ 599.3592, found 599.3605.

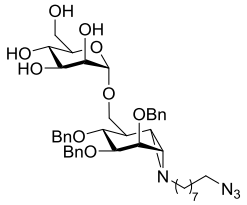
Compound 78



Donor **67**[42] (207 mg, 0.28 mmol) and acceptor **77** (139 mg, 0.23 mmol) were combined in a flask, co-evaporated with toluene (3x) and dissolved in dry DCM (2.3 mL). 4 Å molecular sieves were added, the mixture was stirred for 2 h and then cooled to -40 °C. TfOH (20.5 µL, 0.23 mmol) was added in three portions over 1 h and after stirring for additional 15 minutes the mixture was quenched by addition of Et_3N (100 µL). The mixture was diluted with DCM (50 mL), washed with sat. aq. $NaHCO_3$ (2 x 20 mL) and brine. The organic layer was dried over $MgSO_4$, filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 5:1) afforded the title compound as a colorless oil (201 mg, 74%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.08 (t, *J* = 8.6 Hz, 4H), 7.91 (d, *J* = 7.5 Hz, 2H), 7.83 (d, *J* = 7.5 Hz, 2H), 7.63 – 7.15 (m, 27H), 6.14 (t, *J* = 9.8 Hz, 1H), 5.92 (dd, *J* = 10.1, 3.2 Hz, 1H), 5.75 (d, *J* = 1.7 Hz, 1H), 5.17 – 5.07 (m, 1H), 4.95 (dd, *J* = 11.9, 2.6 Hz, 2H), 4.79 (d, *J* = 12.4 Hz, 1H), 4.75 – 4.60 (m, 3H), 4.57 (d, *J* = 11.5 Hz, 1H), 4.48 – 4.37 (m, 2H), 4.23 (s, 1H), 3.92 – 3.65 (m, 4H), 3.16 (t, *J* = 6.9 Hz, 2H), 2.42 – 2.28

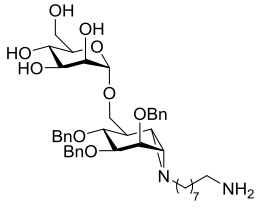
(m, 2H), 2.19 (dt, *J* = 11.7, 6.9 Hz, 1H), 1.90 – 1.80 (m, 1H), 1.68 (d, *J* = 5.9 Hz, 1H), 1.52 (d, *J* = 6.3 Hz, 4H), 1.40 – 1.20 (m, 8H) ppm. ^{13}C NMR (75 MHz, $CDCl_3$) δ 166.2, 165.6, 165.4, 139.1, 139.0, 138.9, 133.6, 133.5, 133.3, 133.2, 130.0, 129.9, 129.8, 129.8, 129.4, 129.2, 129.1, 128.7, 128.6, 128.5, 128.5, 127.9, 127.9, 127.7, 127.6, 127.6, 97.6, 81.7, 75.4, 74.8, 74.5, 73.1, 73.0, 70.6, 70.3, 69.3, 69.2, 67.0, 62.8, 60.8, 51.5, 43.1, 42.9, 40.2, 29.8, 29.6, 29.2, 28.9, 27.4, 26.8 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{70}H_{72}N_4O_{13}$ 1177.5169, found 1177.5215.

Compound 80



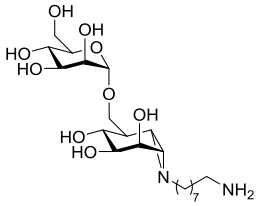
Compound **78** (198 mg, 0.17 mmol) was dissolved in a mixture of DCM/MeOH (1:1 v/v, 1.6 mL), NaOMe (5.4 M in MeOH, 16 µL, 84 µmol) was added and the mixture was stirred overnight at rt. The mixture was concentrated and flash purification by silica column chromatography (DCM/MeOH, 40:1 → 40:3) afforded the title compound as a colorless oil (101 mg, 79%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.46 – 7.07 (m, 15H), 4.91 – 4.84 (m, 2H), 4.82 (s, 1H), 4.70 (d, *J* = 12.3 Hz, 1H), 4.66 (d, *J* = 11.7 Hz, 1H), 4.60 (d, *J* = 11.7 Hz, 1H), 4.46 (d, *J* = 11.4 Hz, 1H), 4.14 (s, 1H), 3.96 (t, *J* = 9.6 Hz, 1H), 3.93 – 3.83 (m, 2H), 3.80 (d, *J* = 9.4 Hz, 1H), 3.73 (dd, *J* = 9.8, 2.6 Hz, 1H), 3.71 – 3.63 (m, 2H), 3.58 (t, *J* = 9.2 Hz, 1H), 3.52 (dd, *J* = 9.5, 3.5 Hz, 1H), 3.47 (d, *J* = 9.7 Hz, 1H), 3.22 (t, *J* = 7.0 Hz, 2H), 2.31 – 2.14 (m, 2H), 2.14 – 2.04 (m, 1H), 1.71 (dd, *J* = 5.7, 1.9 Hz, 1H), 1.57 (p, *J* = 6.9 Hz, 2H), 1.47 (d, *J* = 5.9 Hz, 1H), 1.45 – 1.35 (m, 2H), 1.36 – 1.17 (m, 8H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.0, 138.9, 128.5, 128.4, 127.9, 127.7, 127.6, 127.6, 100.2, 81.6, 74.8, 74.5, 73.1, 72.9, 72.6, 71.9, 71.2, 68.6, 66.2, 61.0, 60.6, 51.6, 43.1, 42.8, 40.3, 29.7, 29.6, 29.2, 28.9, 27.3, 26.8 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{42}H_{56}N_4O_9$ 761.4120, found 761.4150.

Compound 81



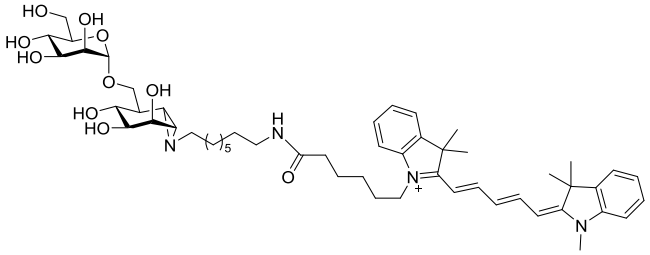
Compound **80** (98 mg, 0.13 mmol) was dissolved in MeCN (2.6 mL), polymer-bound triphenylphosphine (~3 mmol/g loading, 86 mg, 0.26 mmol) and H_2O (23 µL, 1.3 mmol) were added and the mixture was stirred overnight at 70 °C. Then, more H_2O (500 µL) was added and the mixture was stirred for 4 h at 70 °C. The mixture was filtrated and concentrated to afford the product in high purity as an oil (91 mg, 96%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.42 – 7.16 (m, 15H), 4.89 (d, *J* = 11.8 Hz, 2H), 4.84 (s, 1H), 4.72 (d, *J* = 12.3 Hz, 1H), 4.65 (q, *J* = 11.8 Hz, 2H), 4.59 – 4.41 (m, 1H + NH2), 4.20 (s, 1H), 3.97 (dd, *J* = 22.0, 11.5 Hz, 2H), 3.87 (s, 1H), 3.81 – 3.66 (m, 3H), 3.63 – 3.43 (m, 4H), 2.64 (t, *J* = 7.1 Hz, 2H), 2.57 (dd, *J* = 11.6, 5.9 Hz, 1H), 2.21 (td, *J* = 9.3, 4.3 Hz, 1H), 1.76 (d, *J* = 5.5 Hz, 2H), 1.49 (d, *J* = 5.8 Hz, 1H), 1.48 – 1.37 (m, 4H), 1.35 – 1.20 (m, 8H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.0, 139.0, 138.8, 128.4, 128.4, 128.4, 128.0, 127.8, 127.6, 127.5, 127.5, 100.4, 81.9, 75.6, 74.9, 74.5, 73.2, 72.9, 72.8, 71.9, 71.2, 68.5, 66.2, 60.9, 60.8, 43.3, 42.7, 41.4, 40.4, 32.1, 29.8, 29.4, 28.9, 27.2, 26.6 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{42}H_{58}N_2O_9$ 735.4215, found 735.4224.

Compound 82



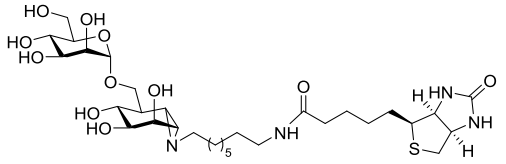
Ammonia (5 mL) was condensed at -60 °C, lithium (21 mg, 3.06 mmol) was added and the mixture was stirred for 30 minutes resulting in a deep blue solution. Compound **81** (45 mg, 61 µmol) dissolved in THF (1 mL) was added dropwise to this solution, the mixture was stirred for 1 hour at -60 °C and subsequently quenched by addition of water (1 mL). The mixture was evaporated, redissolved in H_2O (1 mL) and purified by elution over Amberlite CG-50 (NH_4^+) with NH_4OH (0.5 M) as eluent. The product fractions were combined and evaporated, then the product was redissolved in MeOH, filtrated over Celite and evaporated, affording the product as an oil (23 mg, 81%). ^{1}H NMR (400 MHz, D_2O) δ 4.88 (s, 1H), 4.31 (s, 1H), 3.94 (s, 1H), 3.89 – 3.76 (m, 3H), 3.76 – 3.66 (m, 2H), 3.66 – 3.56 (m, 2H), 3.47 – 3.31 (m, 2H), 2.96 – 2.85 (m, 2H), 2.48 – 2.30 (m, 1H), 2.26 – 2.09 (m, 1H), 1.99 – 1.85 (m, 2H), 1.77 (d, *J* = 5.9 Hz, 1H), 1.63 – 1.55 (m, 2H), 1.55 –1.45 (m, 2H), 1.38 –1.18 (m, 8H) ppm. ^{13}C NMR (101 MHz, D_2O) δ 99.6, 72.8, 71.1, 70.6, 70.0, 67.6, 67.5, 66.7, 66.0, 60.8, 59.7, 43.9, 42.5, 40.4, 39.5, 28.5, 28.0, 27.1, 26.4, 25.5 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{21}H_{40}N_2O_9$ 465.2807, found 465.2810.

Compound 83 (SY-G74)



Compound **82** (8.5 mg, 18.2 µmol) was dissolved in DMF (0.5 mL), then DIPEA (6.3 µL, 36.4 µmol) and Cy5-OSu[46] (12.3 mg, 20.0 µmol) was added and the mixture was stirred overnight. The mixture was diluted with H_2O (0.5 mL) and the resulting crude was purified by HPLC (NH_4HCO_3) affording the product as a blue solid (5.2 mg, 29%). ^{1}H NMR (500 MHz, D_2O) δ 7.81 – 7.72 (m, 2H), 7.37 (t, *J* = 6.4 Hz, 2H), 7.26 (q, *J* = 7.1 Hz, 2H), 7.19 (dd, *J* = 10.8, 8.2 Hz, 2H), 7.11 (q, *J* = 7.3 Hz, 2H), 6.34 (t, *J* = 12.5 Hz, 1H), 6.03 (dd, *J* = 18.9, 13.5 Hz, 2H), 4.83 (d, *J* = 1.4 Hz, 1H), 4.25 (dd, *J* = 3.2, 1.7 Hz, 1H), 4.01 (br s, 2H), 3.89 (dd, *J* = 3.3, 1.7 Hz, 1H), 3.84 (dd, *J* = 12.2, 2.1 Hz, 1H), 3.82 – 3.70 (m, 3H), 3.67 – 3.61 (m, 2H), 3.61 – 3.56 (m, 1H), 3.52 (s, 3H), 3.43 (t, *J* = 10.0 Hz, 1H), 3.36 (dd, *J* = 10.5, 3.5 Hz, 1H), 2.97 (t, *J* = 7.0 Hz, 2H), 2.22 – 2.10 (m, 3H), 2.01 – 1.94 (m, 1H), 1.88 – 1.84 (m, 1H), 1.82 (d, *J* = 4.6 Hz, 1H), 1.80 – 1.75 (m, 2H), 1.65 (d, *J* = 6.1 Hz, 1H), 1.63 – 1.54 (m, 2H), 1.44 (s, 6H), 1.41 (s, 6H), 1.38 – 1.21 (m, 8H), 1.15 – 1.02 (m, 8H) ppm. ^{13}C NMR (125 MHz, D_2O) δ 175.9, 173.5, 173.1, 153.1, 142.7, 142.0, 141.1, 141.0, 128.5, 125.1, 124.3, 122.3, 110.9, 110.7, 103.1, 102.9, 99.6, 72.9, 71.1, 70.8, 70.2, 67.6, 67.4, 66.8, 66.1, 60.9, 59.7, 49.0, 48.9, 44.1, 43.6, 42.5, 40.6, 39.4, 35.5, 30.8, 29.0, 28.6, 28.5, 28.4, 27.0, 26.9, 26.8, 26.6, 26.3, 25.5, 25.1, 22.9 ppm. HRMS (ESI) m/z: $[M]^+$ calc for $C_{53}H_{77}N_4O_{10}$ 930.5712, found 930.5668.

Compound 84 (SY-G73)



Compound **82** (6.6 mg, 14.3 μmol) was dissolved in DMF (0.5 mL), then DIPEA (5 μL, 28.5 μmol) and biotin-OSu[47] (5.4 mg, 15.7 μmol) was added and the mixture was stirred overnight. The mixture was diluted with H_2O (0.5 mL) and the resulting crude was purified by HPLC (NH_4HCO_3) affording the product as a white solid (5.1 mg, 52%). ^{1}H NMR (500 MHz, D_2O) δ 4.90 (d, *J* = 1.4 Hz, 1H), 4.58 (dd, *J* = 7.9, 4.8 Hz, 1H), 4.39 (dd, *J* = 7.9, 4.5 Hz, 1H), 4.35 – 4.28 (m, 1H), 3.96 (dd, *J* = 3.3, 1.7 Hz, 1H), 3.89 – 3.78 (m, 3H), 3.76 – 3.69 (m, 2H), 3.67 – 3.58 (m, 2H), 3.44 (dd, *J* = 19.3, 9.7 Hz, 1H), 3.41 – 3.37 (m, 1H), 3.30 (dt, *J* = 9.8, 5.4 Hz, 1H), 3.21 – 3.10 (m, 2H), 2.97 (dd, *J* = 13.0, 5.0 Hz, 1H), 2.75 (d, *J* = 13.0 Hz, 1H), 2.41 (dt, *J* = 13.4, 6.8 Hz, 1H), 2.25 – 2.18 (m, 3H), 2.01 – 1.93 (m, 2H), 1.81 (d, *J* = 6.1 Hz, 1H), 1.74 – 1.55 (m, 4H), 1.55 – 1.43 (m, 4H), 1.42 – 1.33 (m, 2H), 1.33 – 1.23 (m, 8H) ppm. ^{13}C NMR (125 MHz, D_2O) δ 176.6, 99.7, 73.0, 71.2, 70.7, 70.1, 67.7, 67.6, 66.8, 66.2, 62.1, 61.0, 60.3, 59.8, 55.5, 44.1, 42.6, 40.6, 39.8, 39.3, 35.6, 28.8, 28.7, 28.3, 28.3, 27.9, 27.7, 26.5, 26.0, 25.3 ppm. HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{31}H_{54}N_4O_{11}S$ 691.3583, found 691.3608.

References

1 S. Atsumi, K. Umezawa, H. Iinuma, H. Naganawa, H. Nakamura, Y. Iitaka and T. Takeuchi, *J. Antibiot.*, 1989, **43**, 49–53.

2 G. Caron and S. G. Withers, *Biochem. Biophys. Res. Commun.*, 1989, **163**, 495–499.

3 W. W. Kallemeijn, K. Y. Li, M. D. Witte, A. R. A. Marques, J. Aten, S. Scheij, J. Jiang, L. I. Willems, T. M. Voorn-Brouwer, C. P. A. A. van Roomen, R. Ottenhoff, R. G. Boot, H. van den Elst, M. T. C. Walvoort, B. I. Florea, J. D. C. Codée, G. A. van der Marel, J. M. F. G. Aerts and H. S. Overkleeft, *Angew. Chem. Int. Ed.*, 2012, **51**, 12529–12533.

4 J. Jiang, C. L. Kuo, L. Wu, C. Franke, W. W. Kallemeijn, B. I. Florea, E. van Meel, G. A. van der Marel, J. D. C. Codée, R. G. Boot, G. J. Davies, H. S. Overkleeft and J. M. F. G. Aerts, *ACS Cent. Sci.*, 2016, **2**, 351–358.

5 J. Jiang, W. W. Kallemeijn, D. W. Wright, A. M. C. H. van den Nieuwendijk, V. C. Rohde, E. C. Folch, H. van den Elst, B. I. Florea, S. Scheij, W. E. Donker-Koopman, M. Verhoek, N. Li, M. Schürmann, D. Mink, R. G. Boot, J. D. C. Codée, G. A. van der Marel, G. J. Davies, J. M. F. G. Aerts and H. S. Overkleeft, *Chem. Sci.*, 2015, **6**, 2782–2789.

6 L. Wu, J. Jiang, Y. Jin, W. W. Kallemeijn, C.-L. Kuo, M. Artola, W. Dai, C. van Elk, M. van Eijk, G. A. van der Marel, J. D. C. Codée, B. I. Florea, J. M. F. G. Aerts, H. S. Overkleeft and G. J. Davies, *Nat. Chem. Biol.*, 2017, **13**, 867–873.

7 L. I. Willems, T. J. M. Beenakker, B. Murray, S. Scheij, W. W. Kallemeijn, R. G. Boot, M. Verhoek, W. E. Donker-Koopman, M. J. Ferraz, E. R. Van Rijssel, B. I. Florea, J. D. C. Codée, G. A. Van Der Marel, J. M. F. G. Aerts and H. S. Overkleeft, *J. Am. Chem. Soc.*, 2014, **136**, 11622–11625.

8 B. T. Adams, S. Niccoli, M. A. Chowdhury, A. N. K. Esarik, S. J. Lees, B. P. Rempel and C. P. Phenix, *Chem. Commun.*, 2015, **51**, 11390–11393.

9 G. Luchetti, K. Ding, M. D'Alarcao and A. Kornienko, *Synthesis*, 2008, 3142–3147.

10 B. P. Rempel and S. G. Withers, *Glycobiology*, 2008, **18**, 570–586.

11 J. M. Hernández-Torres, J. Achkar and A. Wei, *J. Org. Chem.*, 2004, **69**, 7206–7211.

12 A. G. Volbeda, H. A. V. Kistemaker, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov and J. D. C. Codée, *J. Org. Chem.*, 2015, **80**, 8796–8806.

13 F. G. Hansen, E. Bundgaard and R. Madsen, *J. Org. Chem.*, 2005, **70**, 10139–10142.

14 Z. You, *Tetrahedron Lett.*, 1993, **34**, 2597–2600.

15 H. Ovaa, B. Lastdrager, J. D. C. Codeé, G. A. van der Marel, H. S. Overkleeft and J. H. van Boom, *J. Chem. Soc. Perkin Trans. 1*, 2002, 2370–2377.

16 T. M. Gloster, P. Meloncelli, R. V. Stick, D. Zechel, A. Vasella and G. J. Davies, *J. Am. Chem. Soc.*, 2007, **129**, 2345–2354.

17 N. Panday, Y. Canac and A. Vasella, *Helv. Chim. Acta*, 2000, **83**, 58–79.

18 S. M. McElvain, R. E. Kent and C. L. Stevens, *J. Am. Chem. Soc.*, 1946, **68**, 1922–1925.

19 P. Lang, G. Magnin, G. Mathis, A. Burger and J. F. Biellmann, *J. Org. Chem.*, 2000, **65**, 7825–7832.

20 A. A. Van Aerschot, P. Mamos, N. J. Weyns, S. Ikeda, E. De Clercq and P. A. Herdewijn, *J. Med. Chem.*, 1993, **36**, 2938–2942.

21 Y. Harrak, C. M. Barra, A. Delgado, A. R. Castaño and A. Llebaria, *J. Am. Chem. Soc.*, 2011, **133**, 12079–12084.

22 N. Brunetti-Pierri and F. Scaglia, *Mol. Genet. Metab.*, 2008, **94**, 391–396.

23 A. Caciotti, S. C. Garman, Y. Rivera-Colón, E. Procopio, S. Catarzi, L. Ferri, C. Guido, P. Martelli, R. Parini, D. Antuzzi, R. Battini, M. Sibilio, A. Simonati, E. Fontana, A. Salviati, G. Akinci, C. Cereda, C. Dionisi-Vici, F. Deodato, A. d'Amico, A. d'Azzo, E. Bertini, M. Filocamo, M. Scarpa, M. di Rocco, C. J. Tifft, F. Ciani, S. Gasperini, E. Pasquini, R. Guerrini, M. A. Donati and A. Morrone, *Biochim. Biophys. Acta - Mol. Basis Dis.*, 2011, **1812**, 782–790.

24 K. Higaki, L. Li, U. Bahrudin, S. Okuzawa, A. Takamuram, K. Yamamoto, K. Adachi, R. C. Paraguison, T. Takai, H. Ikehata, L. Tominaga, I. Hisatome, M. Iida, S. Ogawa, J. Matsuda, H. Ninomiya, Y. Sakakibara, K. Ohno, Y. Suzuki and E. Nanba, *Hum. Mutat.*, 2011, **32**, 843–852.

25 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.

26 C. Braun, T. Lindhorst, N. B. Madsen and S. G. Withers, *Biochemistry*, 1996, **35**, 5458–5463.

27 H. D. Ly, S. Howard, K. Shum, S. He, A. Zhu and S. G. Withers, *Carbohydr. Res.*, 2000, **329**, 539–547.

28 J. C. Díaz Arribas, A. G. Herrero, M. Martín-Lomas, F. Javier Cañada, S. He and S. G. Withers, *Eur. J. Biochem.*, 2000, **267**, 6996–7005.

29 S. S. Lee, S. He and S. G. Withers, *Biochem. J.*, 2001, **359**, 381–386.

30 L. Ziser and S. G. Withers, *Carbohydr. Res.*, 1994, **265**, 9–17.

31 T. H. Fife and T. J. Przystas, *J. Am. Chem. Soc.*, 1980, **102**, 292–299.

32 V. L. Y. Yip and S. G. Withers, *Angew. Chem. Int. Ed.*, 2006, **45**, 6179–6182.

33 F. M. Ibatullin, S. I. Selivanov and A. G. Shavva, *Synthesis*, 2001, **2001**, 419–422.

34 T. Belz, Y. Jin, J. Coines, C. Rovira, G. J. Davies and S. J. Williams, *Chem. Commun.*, 2017, **53**, 9238–9241.

35 S. Ghosh, P. Tiwari, S. Pandey, A. K. Misra, V. Chaturvedi, A. Gaikwad, S. Bhatnagar and S. Sinha, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 4002–4005.

36 A. J. Thompson, F. Cuskin, R. J. Spears, J. Dabin, J. P. Turkenburg, H. J. Gilbert and G. J. Davies, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2015, **71**, 408–415.

37 F. Cuskin, E. C. Lowe, M. J. Temple, Y. Zhu, E. A. Cameron, N. A. Pudlo, N. T. Porter, K. Urs, A. J. Thompson, A. Cartmell, A. Rogowski, B. S. Hamilton, R. Chen, T. J. Tolbert, K. Piens, D. Bracke, W.

Vervecken, Z. Hakki, G. Speciale, J. L. Munōz-Munōz, A. Day, M. J. Peña, R. McLean, M. D. Suits, A. B. Boraston, T. Atherly, C. J. Ziemer, S. J. Williams, G. J. Davies, D. W. Abbott, E. C. Martens and H. J. Gilbert, *Nature*, 2015, **520**, 388–388.

38 J. Ao, J. L. Chinnici, A. Maddi and S. J. Free, *Eukaryot. Cell*, 2015, **14**, 792–803.

39 A. Maddi, C. Fu and S. J. Free, *PLoS One*, 2012, **7**, e38872.

40 H. Kitagaki, H. Wu, H. Shimoi and K. Ito, *Mol. Microbiol.*, 2002, **46**, 1011–1022.

41 C. S. Wong, Thesis: The Synthesis of Mannose-derived Bioconjugates and Enzyme Inhibitors; Leiden University, 2015.

42 B. Becker, R. H. Furneaux, F. Reck and O. a. Zubkov, *Carbohydr. Res.*, 1999, **315**, 148–158.

43 H. Franzyk, M. Meldal, H. Paulsen and K. Bock, *J. Chem. Soc. Perkin Trans. 1*, 1995, 2883–2898.

44 T. J. Rabelink, B. M. van den Berg, M. Garsen, G. Wang, M. Elkin and J. van der Vlag, *Nat. Rev. Nephrol.*, 2017, **13**, 201–212.

45 V. Bordoni, V. Porkolab, S. Sattin, M. Thepaut, I. Frau, L. Favero, P. Crotti, A. Bernardi, F. Fieschi and V. Di Bussolo, *RSC Adv.*, 2016, **6**, 89578–89584.

46 M. V. Kvach, A. V. Ustinov, I. A. Stepanova, A. D. Malakhov, M. V. Skorobogatyi, V. V. Shmanai and V. A. Korshun, *Eur. J. Org. Chem.*, 2008, 2107–2117.

47 K. Susumu, H. T. Uyeda, I. L. Medintz, T. Pons, J. B. Delehanty and H. Mattoussi, *J. Am. Chem. Soc.*, 2007, **129**, 13987–13996.

Nederlandse samenvatting

Glycosidases zijn enzymen die voorkomen in alle vormen van leven. Glycosidases zijn essentieel voor de afbraak van (complexe) koolhydraten, zoals tafelsuiker, maar ook glycolipiden en glycoproteïnes. Koolhydraten zijn lange ketens van monosacchariden die aan elkaar geketend zijn via acetaal bindingen. Deze bindingen zijn van nature erg stabiel in waterig milieu en zullen dus niet eenvoudig verbreken. Om een gezonde stofwisseling te behouden (in bijvoorbeeld het menselijk lichaam) is de afbraak van koolhydraten van essentieel belang voor energietoevoer, maar ook om correcte functie van eiwitten en cellulaire processen te behouden. Glycosidases katalyseren ('versnellen') de afbraak van koolhydraten door ze eigenhandig in stukken te knippen door middel van een hydrolyse reactie met water. Waneer een glycosidase enzym niet goed werkt (door bijvoorbeeld een genetische afwijking) kunnen de suikers die zij normaliter afbreekt ophopen, hetgeen zorgt voor een verstoring in belangrijke cellulaire processen. Omdat er vele verschillende typen glycosidases bestaan, staan deze enzymatische defecten aan de grondslag van verschillende ziektes, zoals de lysosomale stapelingsziektes van Gaucher en Fabry. Tevens wordt verstoorde glycosidase functie onder andere in verband gebracht met kanker, Parkinson en Alzheimer.

De functie en activiteit van glycosidases kan bestudeerd worden met een recentelijk ontwikkelde techniek: activity-based protein profiling (ABPP). Deze techniek berust op het toedienen van een reactieve verbinding (probe) aan een biologische monster waarbij de probe reageert met het enzym van interesse en daarbij andere enzymen ongemoeid laat. De probe is gebaseerd op de natuurlijke verbinding cyclophellitol, dat sterk lijkt op een koolhydraat en ook als zodanig herkend zal worden door het glycosidase enzym. Het enzym zal proberen de probe in stukken te knippen, maar gaat daarbij een niet-reversibele reactie aan met de probe waardoor het enzym niet meer los kan laten. Hierbij word tevens de enzymatische functie uitgeschakeld. De probe bevat een fluorescerende kleurstof en kleurt daarbij het enzym indirect aan, waardoor het met speciale scan-apparatuur gevisualizeerd kan worden. Met dit soort probes kan dus de aan- of afwezigheid van bepaalde glycosidases in biologische monsters worden bepaald (diagnostiek).

Omdat er vele verschillende soorten glycosidases bestaan, zijn er veel verschillende probes nodig om elk soort specifiek te kunnen bestuderen. Hoewel er al meerdere

probes gesynthetiseerd zijn om bijvoorbeeld α- en β-glucosidases, mannosidases, galactosidases en fucosidases aan te kunnen kleuren, dekken deze niet de lading van het grote scala aan glycosidases. Het onderzoek dat staat beschreven in dit proefschrift omvat de synthese en biochemische evaluatie van nieuwe glycosidase probes, die gebruikt kunnen worden om niet eerder verkende soorten glycosidases te kunnen bestuderen in biologische context.

Hoofdstuk 1 geeft een algemene introductie over glycosidases, hun functie en de reactiepaden die zij doorlopen bij het hydrolyseren van een suikerketen. Tevens wordt de ABPP techniek toegelicht, en worden de probes die voorheen gepubliceerd zijn in de literatuur besproken.

Hoofdstuk 2 beschrijft de synthese van een glycosidase probe die bestaat uit de xylose configuratie. De belangrijkste stap in deze synthese bestaat uit een asymmetrische allyleringsreactie met een chiraal allyl-boraan reagens. De probe blijkt uitstekend te reageren met β-xylosidases en kan dus gebruikt worden om deze in biologische context beter te bestuderen. Omdat de xylose configuratie sterk lijkt op de glucose configuratie (met als verschil de hydroxymethyleen groep op positie C-5), is onderzocht of deze probe tevens met glucosidases reageert. Dat blijkt mogelijk; er is dus enige kruisreactiviteit van β-glucosidases met xylose geconfigureerde probes.

Hoofdstuk 3 borduurt voort op bovenstaande bevinding. Ditmaal worden probes gesynthetiseerd die hydroxyl groepen missen op positie C-4 en C-2. De alcoholgroepen op deze posities maken het onderscheid tussen de glucose-, mannose- en galactose configuraties. Door deze groepen weg te laten zou de probe in theorie tegelijk glucosidases, mannosidases en galactosidases kunnen binden waardoor simultaan glycosidases van verschillende klassen bestudeerd kunnen worden. De werkelijkheid blijkt subtieler; door het weghalen van hydroxylgroepen op C-2 en C-4 verliest de probe zijn potentie, maar het verwijderen van enkel de alcohol groep op C-4 geeft een probe die selectief reageert met glucosidases en galactosidases.

Tot nu toe bestaan alle glycosidase probes gebaseerd op het natuurproduct cyclophellitol uit een 6-ring en lijken daarmee sterk op het natuurlijke suiker substraat. In **Hoofdstuk 4** wordt de synthese beschreven van glycosidase probes bestaande uit een 5-ring. Er is onderzocht of deze verbindingen tevens in staat zijn om glycosidases te binden, maar dit bleek niet het geval.

In de jaren ’90 hebben de groepen van Tatsuta en Vasella uitgebreid onderzoek gedaan aan gluco-azolen; azool ringen gefuseerd aan een glucose geconfigureerde piperidine als niet-covalente (reversibele) remmers voor β-glucosidases. Van deze reeks gluco-azolen bleek glucoimidazool (met een exocyclisch stikstof atoom op de plek van het aglycon en een stikstof op de plek van het endocyclisch zuurstof atoom in glucose) de sterkste remmer te zijn. Echter er ontbrak één structuur in deze reeks; de imidazool-gefuseerde verbinding waarbij beiden stikstof atomen exocyclisch zitten ten opzichte van de glucose ring. **Hoofdstuk 5** beschrijft de synthese van deze nieuwe klasse van remmers, genaamd de gluco-1*H*-imidazolen. De synthese vereiste een andere synthetische route dan gebruikt voor de andere gluco-azolen, en wordt gemaakt vanuit een bouwsteen voor cyclophellitol. Hoewel gluco-1*H*-imidazool een reversibele remmer is voor β-glucosidases, is zij minder potent dan glucoimidazool. De reden hiervoor is achterhaald met behulp van kristallografie, isothermale titratie calorimetrie en DFT calculaties, en valt toe te wijzen aan een gebrek aan basiciteit van het exo-cyclisch stikstof atoom, alsmede een verminderde electrostatische interactie met het katalytisch nucleofiel van het enzym na protonering van de azool ring. Desalniettemin valt via de ontwikkelde syntheseroute de ‘suiker’ ring evenals de subsituent op de imidazool ring gemakkelijk aan te passen, hetgeen gezorgd heeft voor reversibele inhibitors met nanomolaire activiteit tegen lysosomale glucocerebrosidase (GBA1), met hoge selectiviteit over GBA2 en glucocerebrosidase synthase. Deze remmers zouden dus gebruikt kunnen worden als medicijnen tegen de ziekte van Gaucher via chaperone therapie.

Hoofdstuk 6 beschrijft de synthese van twee potentiële covalente remmers van *endo*-α-1,2-mannanases van glycosidase familie GH99. Doordat een van de katalytische aminozuren zich op een ongebruikelijke positie bevindt in het katalytische centrum van het enzym, is er gespeculeerd dat e*ndo*-α-1,2-mannanases een ander reactiepad gebruiken om hun substraat om te zetten (via het 1,2-anhydro epoxide). De gesynthetiseerde spiro-epoxyglycosides bestaan uit disaccharides gefunctionaliseerd met een spiro-epoxide op positie C-2, dat dient als het electrofiel voor irreversiebele reactie met het enzym. Met behulp van ABPP technieken wordt aangetoond dat deze verbindingen inderdaad labelen met recombinant GH99 enzym, en de labeling lijkt plaats te vinden in het katalytisch centrum.

In **Hoofdstuk 7** wordt een andere klasse van *endo*-glycosidases bestudeerd; de *endo*-xylanases. Tot noch toe waren alle cyclophellitol-gebaseerde probes monomeer van

structuur, en labelden daardoor voornamelijk glycosidases die enkel een monomeer in hun katalytisch centrum toelaten (*exo*-glycosidases). Dit hoofdstuk beschrijft een methode om gemakkelijk een *xylo*-cyclophellitol probe te verlengen met een xylose suiker residu, via directe chemische glycosylering. Hierdoor ontstaat een disaccharide probe die in staat is om *endo*-xylanases te labelen. Dit wordt gedemonstreerd met het labelen van het secretoom van de industrieel relevante schimmel *Aspergillus niger*. Met deze probes kan het pH optimum, alsmede de temperatuur stabiliteit van xylanases in het secretoom gemakkelijk bestudeerd worden.

Ten slotte wordt in **Hoofdstuk 8** een meer gedetailleerde samenvatting gegeven van de bevindingen in dit Proefschrift, verwoven met hierop gebaseerde voorstellen voor vervolgonderzoek.

List of publications

Catalytic Decarboxylative Alkenylation of Enolates
S. P. Schröder, N. J. Taylor, P. Jackson and V. Franckevičius
Org. Lett., **2013**, *15*, 3778–3781

From covalent glycosidase inhibitors to activity-based glycosidase probes
L. I. Willems, J. Jiang, K.-Y. Li, M. D. Witte, W. W. Kallemeijn, T. J. M. Beenakker, S. P. Schröder, J. M. F. G. Aerts, G. A. van der Marel, J. D. C. Codée and H. S. Overkleeft
Chem. Eur. J., **2014**, *20*, 10864-10872

The synthesis of cyclophellitol-aziridine and its configurational and functional isomers
J. Jiang, M. Artola, T. J. M. Beenakker, S. P. Schröder, R. Petracca, C. de Boer, J. M. F. G. Aerts, G. A. van der Marel, J. D. C. Codée and H. S. Overkleeft
Eur. J. Org. Chem. **2016**, *2016*, 3671-3678

A divergent synthesis of L-*arabino*- and D-*xylo*-configured cyclophellitol epoxides and aziridines
S. P. Schröder, R. Petracca, H. Minnee, M. Artola, J. M. F. G. Aerts, J. D. C. Codée, G. A. van der Marel and H. S. Overkleeft
Eur. J. Org. Chem. **2016**, *2016*, 4787-4794

Towards broad spectrum activity-based glycosidase probes: synthesis and evaluation of deoxygenated cyclophellitol aziridines
S. P. Schröder, J. W. van de Sande, W. W. Kallemeijn, C.-L. Kuo, M. Artola, E. J. van Rooden, J. Jiang, T. J. M. Beenakker, B. I. Florea, W. A. Offen, G. J. Davies, A. J. Minnaard, J. M. F. G. Aerts, J. D. C. Codée, G. A. van der Marel and H. S. Overkleeft
Chem. Commun. **2017**, *53*, 12528-12531

Palladium-Catalysed Construction of All-Carbon Quaternary Centres with Propargylic Electrophiles: Challenges in the Simultaneous Control of Regio-, Chemo- and Enantioselectivity

M. Kenny, S. P. Schröder, N. J. Taylor, P. Jackson, D. J. Kitson,V. Franckevičius
Synthesis **2018**, DOI: 10.1055/s-0036-1591957

Gluco-1*H*-imidazole: a new class of azole-type β-glucosidase inhibitor

S. P. Schröder, L. Wu, M. Artola, T. Hansen, W. A. Offen, M. J. Ferraz, K.-Y. Li, J. M. F. G. Aerts, G. A. van der Marel, J. D. C. Codée, G. J. Davies and H. S. Overkleeft
J. Am. Chem. Soc. **2018**, http://dx.doi.org/10.1021/jacs.8b02399

Curriculum vitae

Sybrin Schröder was born on the 20th of July 1989 in Nijmegen. After finishing his secondary education (VWO) at the Nijmeegse Scholengemeenschap Groenewoud in 2007, he commenced his studies in Chemistry at the Radboud Universiteit Nijmegen. He received his Bachelors degree in 2010 after successfully completing his internship in the group of Prof. dr. Alan E. Rowan. Then, he obtained his Masters degree in 2012 after completing his internship in the group of Prof. dr. Floris P. J. T. Rutjes entailing the synthesis of fluorinated analogues of 1-deoxynojirimycin, followed by an internship abroad at the University of York in the group of Prof. dr. Richard J. K. Taylor focused on palladium catalyzed decarboxylative propargyl transfer reactions. After his Masters degree, he worked as process chemist at ChemConnection in Oss for a period of six months. In 2013, he started his PhD studies at Leiden University in the bio-organic synthesis group under supervision of Prof. dr. Hermen S. Overkleeft. Parts of the research described in this Thesis were presented on posters at Molecules: Synthesis and Properties (Lunteren, 2013), CHAINS (Veldhoven, 2014, 2016), EMBL Chemical Biology (Heidelberg, 2014), COST (Oxford, 2015), Reedijk Symposium (Leiden, 2016) and Eurocarb (2017, Barcelona). Oral presentations about his research were presented at the annual ABPP conferences in Leiden (2017) and Oxford (2018). Sybrin is currently continuing his research in the group of Prof. dr. Overkleeft as a postdoctoral fellow.