Exploring chemical space in covalent and competitive glycosidase inhibitor design

Chen, Y.

2

Activity-Based Protein Profiling of Retaining α-Amylases in Complex Biological Samples

Manuscript published as:

Yurong Chen,[*] Zachary Armstrong,[*] Marta Artola, Bogdan I. Florea, Chi-Lin Kuo, Casper de Boer, Mikkel S. Rasmussen, Maher Abou Hachem, Gijsbert A. van der Marel, Jeroen D. C. Codée, Johannes M. F. G. Aerts, Gideon J. Davies, and Herman S. Overkleeft, *J. Am. Chem. Soc.* **2021**, *143*, 2423-2432.

2.1 Introduction

Retaining *α*-amylases, belonging to the glycoside hydrolase (GH) 13 family (www.cazy.org),[1] are starch-processing enzymes present in plants, animals and microorganisms.[2-4] They catalyze the hydrolysis of internal α-1,4-glycosidic linkages in starch and related polysaccharides through a two-step Koshland double displacement mechanism (Figure 2.1). In this mechanism, the aglycon is protonated by the general acid/base residue, concomitant with an S_N2 nucleophilic attack by the catalytic nucleophile at the anomeric center to form a covalent enzyme-substrate intermediate with inversion of anomeric configuration. The covalent enzyme-substrate adduct is subsequently hydrolyzed to yield oligosaccharide products (for instance linear and alpha-1,6-branched *malto*-oligosaccharides) with net retention of α-anomeric configuration.[5]

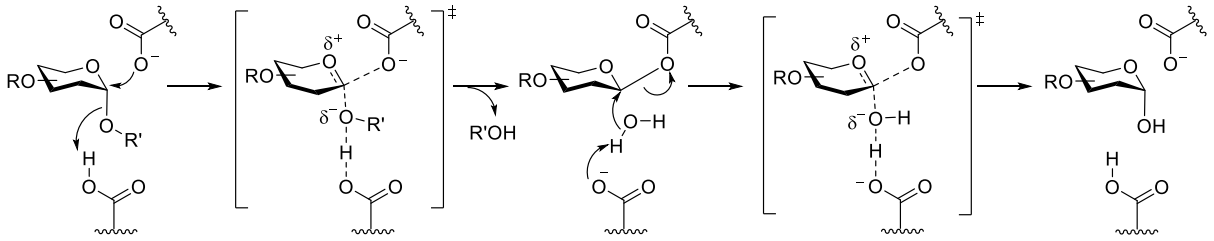


Figure 2.1. General Koshland double displacement mechanism employed by retaining α-amylases.

Starch is the main product of major human and animal feedstock crops including wheat, rice, maize, tapioca and potato. Traditionally, large-scale starch processing, an industry established over a hundred years ago, relied on the degradation of starch by acid hydrolysis.[6,7] Amylases have gained considerable traction as alternatives to conventional chemical starch hydrolysis, as they are more environmentally benign, are less energy intensive and moreover may produce more defined oligosaccharide products from starch polymers.[8,9] Today, a large number of α-amylases from microbial sources are commercially available and are widely used for the production of high glucose and fructose syrups.[10,11] Another major application of α-amylases is found in the detergent industry, where they act at low temperatures and alkaline pH to remove tough starch stains from clothes and porcelains under mild conditions.[12] α-Amylases are furthermore extensively employed in textile, food, paper and biofuel industries. The α-amylases currently applied in industrial biotechnology originate from a limited number of fungal and bacterial strains, including *Aspergillus oryzae* and *Bacillus licheniformis*.[11,13] The microbial world however may harbor many additional α-amylases including ones that differ functionally in terms of activity, substrate specificity and resistance to (harsh) biotechnological conditions.

In addition to their potential in biotechnological applications, α-amylases have received increased interest in the past decades in the context of human health. In mammals, digestion of starch by salivary and pancreatic α-amylases leads to linear and branched *malto*-oligosaccharides. These are largely hydrolyzed to glucose by α-glucosidases located in the intestinal mucosa, which then enters the blood stream by means of facilitated diffusion. Control of postprandial glucose levels, which has proven to be a key factor for the treatment of diabetes and obesity,[14] might be achieved by activity modulation of starch-processing enzymes. Anti-diabetic drugs currently used in the clinic include the α-glucosidase inhibitors,[15,16] miglitol, voglibose and acarbose, the latter of which also inhibits human pancreatic α-amylase (HPA).[17,18] However, their effectiveness is offset by undesirable side effects, including diarrhea, flatulence, and abdominal pain, secondary effects that arise predominantly from inhibition of intestinal α-glucosidases. Hence, selective inhibitors targeting exclusively HPA, the first enzyme in the digestion sequence, would be of great interest.[19-24]

Activity-based protein profiling (ABPP) has emerged as a powerful strategy to obtain qualitative and quantitative activity information on enzymes of interest in complex mixtures.[25,26] ABPP is especially useful for retaining glycosidases that form a covalent enzyme–substrate intermediate during hydrolysis. It relies on the availability of suitable activity-based probes (ABPs) that are generally generated from covalent and irreversible inhibitors by equipping these with a reporter entity (a fluorophore or biotin) or a bio-orthogonal tag. By means of ABPP, enzyme activities can be rapidly identified and quantified (comparative ABPP) and the effect of inhibitors on target enzymes studied in cell extracts and living cells (competitive ABPP). Previous work on retaining glycosidase ABPP is rooted in the known mechanism-based *β*-glucosidase inhibitor cyclophellitol, isolated from *Phellinus sp.*[27] In the first instance cyclophellitol-derived ABPs were applied to the study of retaining *exo*-acting *α*- and *β*-glycosidases.[28-32] More recently, this methodology was expanded towards retaining endoglycosidases in work on xylobiose-configured ABPs to investigate *endo*-acting xylanases in *Aspergillus* secretomes.[33,34] This chapter addresses the efficacy of retaining glycosidase ABPP by the design and chemical synthesis of a panel of maltobiose *epi*-cyclophellitol derived retaining α-amylase inhibitors and ABPs, the kinetic and structural analysis of their interactions with microbial α-amylases, and the application of these probes for rapid detection and identification of α-amylases in fungal secretomes, mouse tissue and human saliva. The described work complements a recent study by Withers and colleagues,[24] who reported a chemo-enzymatic synthesis of maltobiose *epi*-cyclophellitol, to which the chemically derived

material and as well the structural work on α-amylase-inhibitor complexes could be matched. The results described in this chapter moreover illustrate the potential of retaining α-amylase ABPP in the annotation of α-amylases in complex biological samples.

2.2 Results and discussion

2.2.1 Design and synthesis of α-amylase inhibitors and activity-based probes

1,6-*Epi*-cyclophellitol, as well as its aziridine and cyclic sulfate analogues (Figure 2.2A) are potent mechanism-based α-glucosidase inhibitors, and it has recently become clear that their tagged versions are effective activity-based retaining α-exoglucosidase probes.[30,35] These monosaccharidic cyclitols however do not inhibit *endo*-acting α-amylases, which arguably is because of the absence of productive binding with the larger *endo*-glycosidase active sites (Chapter 1). In line with this consideration, Withers and co-workers subsequently reported that the maltose disaccharide mimetic α-1,4-glucopyranosyl *epi*-cyclophellitol (compound **1a**, Figure 2.2B), which they prepared by enzymatic α-glucosylation of 1,6-*epi*-cyclophellitol, is an effective mechanism-based, covalent and irreversible HPA inhibitor.[24] Crystallographic analysis of an HPA-inhibitor complex points to the O4' of the nonreducing end sugar,[24] which may be alkylated to block potential exoglycosidase action, as a promising site for orthogonal functionalization (as in **1b**) to allow introduction of a reporter entity (a fluorophore or biotin, as in **1c** and **1d**). Alternatively, and in line with previous[33] work on xylanase ABPs maltose-configured cyclophellitol aziridines can be designed bearing the reporter tag on the aziridine nitrogen, either with or without an O4' cap (compounds **2a**-**2c**, Figure 2.2B). Finally, and in line with previous results on 1,6-*epi*-cyclophellitol cyclosulfate being a potent and selective retaining α-exoglucosidase inhibitor,[30,35] it was decided to synthesize the corresponding maltose cyclosulfate (compound **3a**, Figure 2.2B) as well, with the aim to investigate whether this Michaelis complex emulating inhibitor would also act on amylases.

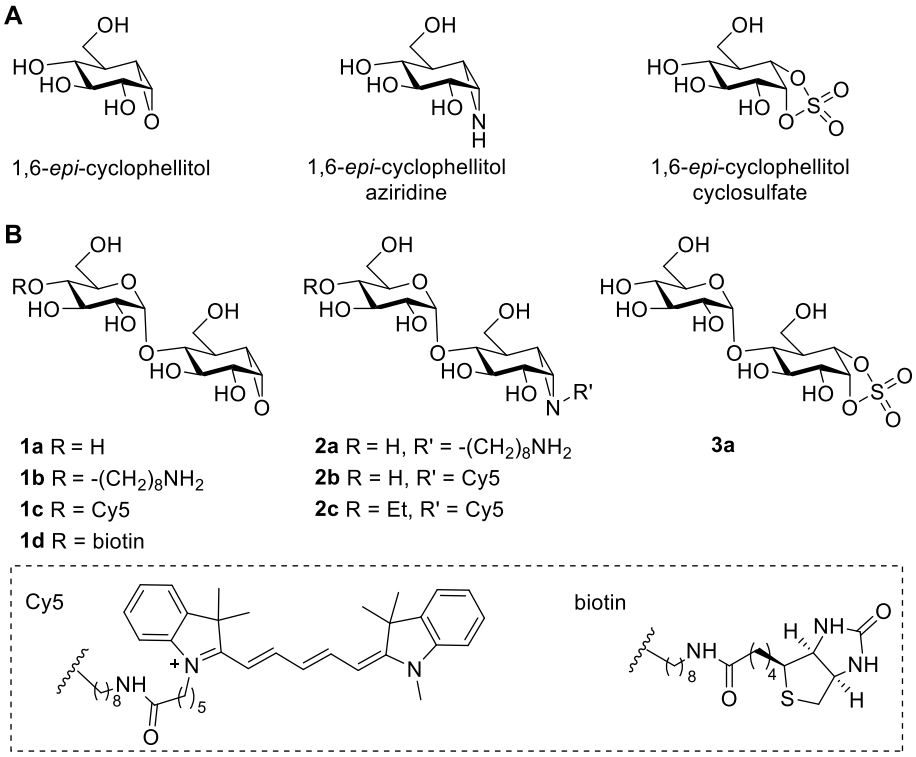
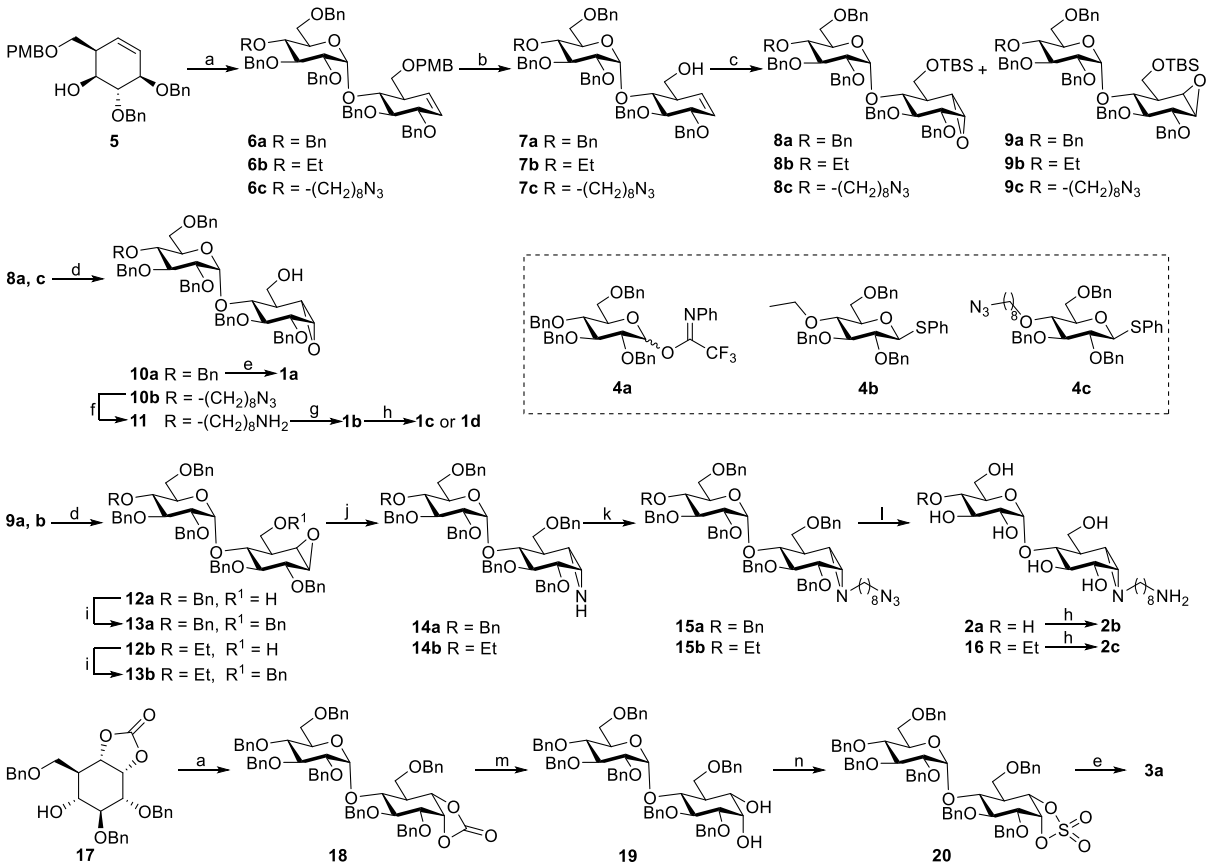


Figure 2.2. Design of mechanism-based retaining *α*-glucosidase inhibitors and ABPs. A) Structures of retaining α-exoglucosidase inhibitors; B) Structures of retaining α-endoglucosidase inhibitors and ABPs, which are the subject of the described studies in this chapter.

The recently reported synthesis of xylobiose-cyclophellitols was accomplished via the direct glycosylation of xylo-cyclophellitol (aziridine) acceptors.[33] This route however appeared not feasible for the construction of compounds **1a-d** since the epoxide functionality in partially protected 1,6-*epi*-cyclophellitols proved unstable towards stoichiometric amounts of Lewis acid under the conditions required to effect formation of α-1,4-glycosidic bonds using contemporary glycosylation conditions. Alternatively, cyclohexene **5** was found to readily undergo glycosylation with imidate donor **4a** under *N*,*N*-dimethyl formamide (DMF) and trifluoromethanesulfonic acid (TfOH) activating conditions, affording pseudodisaccharide **6a** in good yield.[36] Selective deprotection of the 4-methoxybenzyl group in **6a** followed by epoxidation afforded an inseparable mixture of *α/β*- epoxides in a 1:3 ratio, which were silylated and separated by silica gel column chromatography to yield pure **8a** and **9a**. Desilylation and global debenzylation of **8a** by Pearlman's catalyst hydrogenation gave maltobiose *epi*-cyclophellitol **1a**, of which all analytical and spectroscopical data matched those of the chemo-enzymatically produced compound.[24] The epoxide in **9a** was converted to aziridine **14a** by treatment with sodium azide followed by a Staudinger-type ring closure. Alkylation and global

deprotection of **14a** under Birch conditions afforded alkyl-aziridine **2a**, which was equipped with a Cy5 fluorophore to produce ABP **2b** (Scheme 2.1).



Scheme 2.1. Synthesis of the inhibitors and ABPs subject of the here-described studies. Reagents and conditions: a) for **6a** and **18**: **4a**, TfOH (1.5 eq), DMF, DCM, 4 Å MS, -20 °C to 0 °C, **6a** 96%, **18** 57%; for **6b**: **4b**, NIS, TfOH (1.5 eq), DMF, DCM, 4 Å MS, 0 °C, 88%; for **6c**: **4c**, NIS, TfOH (1.5 eq), DMF, DCM, 4 Å MS, 0 °C, 87%; b) DDQ, DCM/H_2O (19/1), rt, **7a** 80%, **7b** 74%, **7c** 71%; c) *i*) *m*CPBA, DCM, 0 °C to rt; *ii*) TBSCl, DMAP, imidazole, DCM, rt, **8a** 20%, **9a** 58%, **8b** 21%, **9b** 65%, **8c** 22%, **9c** 63%; d) TBAF, THF, rt, **10a** 90%, **10b** 87%, **12a** 91%, **12b** 88%; e) $Pd(OH)_2$/C, H_2, MeOH/H_2O/dioxane (2/1/2), rt, **1a** 100%, **3a** 89%; f) polymer-bound PPh_3, MeCN, H_2O, 70 °C, 79%; g) $Pd(OH)_2$/C, H_2, HOAc, ^{t}BuOH/H_2O/dioxane (1/2/1), rt, 91%; h) Cy5-OSu or biotin-OSu, DIPEA, DMF, rt, **1c** 33%, **1d** 16%, **2b** 10%, **2c** 17%; i) BnBr, NaH, TBAI, DMF, 0 °C to rt, **13a** 82%, **13b** 75%; j) *i*) NaN_3, $LiClO_4$, DMF, 100 °C; *ii*) polymer-bound PPh_3, MeCN, 60 °C, **14a** 48%, **14b** 42%; k) 8-azidooctyltrifluoromethanesulfonate, DIPEA, DCM, 0 °C to rt, **15a** 90%, **15b** 88%; l) Na, NH_3 (liq.), ^{t}BuOH, THF, -60 °C, **2a** 79%, **16** 94%; m) NaOMe, DCM/MeOH (1/1), rt, 93%; n) *i*) $SOCl_2$, TEA, DCM, 0 °C; *ii*) $RuCl_3.3H_2O$, $NaIO_4$, EtOAc/ACN/H_2O (2/2/1), 0 °C, 58%.

Previous studies on *endo*-xylanases revealed that the *xylobiose*-cyclophellitol aziridine probe was susceptible to hydrolysis by *exo*-xylosidases present in *Aspergillus* secretomes.[33] Thus, it was hypothesized that "capping" the O4' position of maltobiose *epi*-cyclophellitol azirines with

an ethyl group would prevent cleavage by *exo*-acting α-glucosidases in complex biological samples. This was achieved by using O4 ethyl capped thiophenyl donor **4b** for glycosylation with acceptor **5** under *N*-iodosuccinimide (NIS)/DMF/TfOH catalysis, giving 4'-ethyl pseudodisaccharide **6b** in good yield. Following the same route as for **2a**, 4'-ethyl capped maltobiose *epi*-cyclophellitol aziridine **16** could be readily obtained from intermediate **9b**, which was equipped with a Cy5 fluorophore to produce ABP **2c**. We constructed epoxide-based ABPs by using thiophenyl donor **4c**, where the azidooctyl group was introduced in an early stage for later installation of reporter entities and also to prevent *exo*-glycosidase processing. Following the established route, epoxide **1b** was obtained after desilylation, azide reduction and debenzylation of **8c**. Following coupling with a Cy5 fluorophore or biotin moiety, ABPs **1c** and **1d** were obtained respectively.

The same strategy was also applied to construct maltobiose *epi*-cyclophellitol cyclosulfate **3a**. Carbonate acceptor **17** was readily glycosylated with donor **4a** under TfOH/DMF conditions, giving pseudodisaccharide **18** in moderate yield. Treatment of **18** with sodium methoxide gave diol **19**. Generation of the cyclosulfite using thionyl chloride and subsequent oxidation gave perbenzylated cyclosulfate **20**, the benzyl ethers of which were removed by Pearlman's catalyst hydrogenation to afford inhibitor **3a**.

2.2.2 Kinetic studies for time-dependent inhibition of *Aspergillus oryzae* α-amylase by maltobiose inhibitors

To assess the potency of the synthesized maltobiose *epi*-cyclophellitols, the time-dependent inhibition of *Aspergillus oryzae* α-amylase (hereafter termed as Taka-amylase) was monitored. Taka-amylase, one of the earliest amylases to be characterized and used commercially,[37,38] is a retaining α-amylase belonging to GH13 subfamily 1. This enzyme uses a pair of catalytic active site residues – Asp206 as a nucleophile and Glu230 as the acid/base – to cleave amylose chains and requires several subsites to be occupied to catalyse hydrolysis at meaningful rates.[39]

To assess time-dependent inhibition, the hydrolysis of 2-chloro-4-nitrophenyl maltotrioside (CNP-M3) by Taka-amylase was measured after pre-incubation with 2 mM of inhibitors **1a**, **1b, 2a** and **3a**. All four inhibitors displayed a time-dependent inhibition of Taka-amylase, however the rates of inhibition varied greatly (Figure 2.3). Both the *epi*-cyclophellitol epoxide **1a** and aziridine **2a** had slow rates of inactivation, with rate constants of 0.0061 ± 0.0002 min^{-1} and 0.005 ± 0.002 min^{-1} respectively. The rate constant for the inactivation by **1a** is somewhat higher than was previously observed for its inhibition of HPA, determined to be 0.001 min^{-1} at

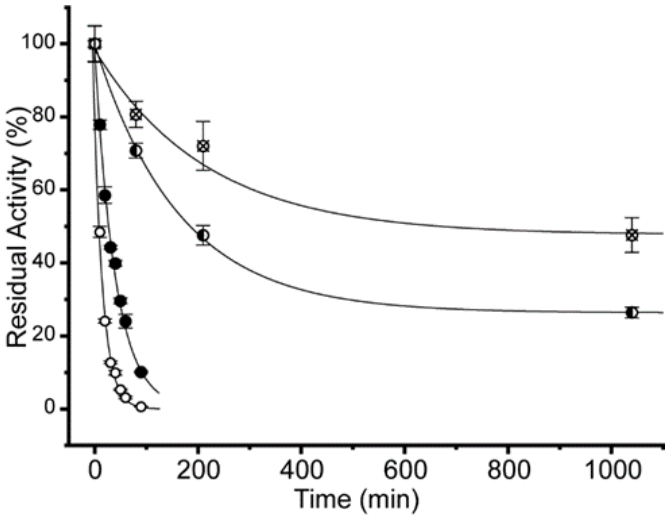


Figure 2.3. Time-dependent inactivation of Taka-amylase. Residual activity of Taka-amylase incubated with 2 mM of either **1a** (◐), **1b** (○), **2a** (⊗) or **3a** (●) was assayed over time and fit to an exponential decay curve.

a higher concentration of inhibitor (3.6 mM).[24] It is somewhat surprising that the aziridine had a rate of inhibition slightly worse than that of the epoxide, as previous work has demonstrated aziridine based inhibitors to be more potent for α-glucosidases.[30] It was speculated that *N*-alkylation of the aziridine may reduce the initial binding of inhibitor **2a**, consequently reducing inactivation rates. Inhibition by the cyclosulfate inhibitor **3a** was 4-5 fold faster (rate constant = 0.0250 ± 0.0009 min^{-1}) when compared to **1a** or **2a**. This observation agrees well with our previous work, which showed that the monosaccharide *epi*-cyclophellitol cyclosulfate was indeed the most potent cyclophellitol type inhibitor of recombinant lysosomal acid α-glucosidase.[35] Surprisingly, 4'-alkylamine **1b** proved to be the most effective inhibitor tested, with an inhibition rate constant that was nearly 10 times that of the untagged **1a** (rate constant = 0.057 ± 0.003 min^{-1}). This suggests that initial binding of **1b** is aided by interactions between the -3 subsite and the 4'-alkylamine present, which in turn increases the rate of inhibition. Previous kinetic analysis of Taka-amylase catalytic rates has shown dramatic rate increases for the hydrolysis of a maltotetrasaccharide over a maltotrioside,[39] highlighting the importance of binding in the -3 position with a natural substrate.

2.2.3 Structural analysis of inhibitor-enzyme complexes

X-ray crystallography studies were conducted next to further understand the binding of inhibitors **1a**, **1b**, **2a** and **3a** to the active site of Taka-amylase. All four inhibitors formed covalent complexes with Taka-amylase, solved with resolution between 1.67 and 1.35 Å, and were covalently linked to the enzyme through the catalytic nucleophile Asp206 (Figure 2.4).

The slow reaction rates necessitated prolonged incubation of the inhibitors with pre-formed crystals. The *epi*-cyclophellitol epoxides were soaked for 3 days. At this stage **1a** reached 80% active site occupancy, while **1b** had complete occupancy, reflective of its increased rate constant of inhibition when compared to **1a**. After a week-long soak with **2a** Taka-amylase crystals had approximately 70% inhibitor occupancy, thus soak length for this compound was further extended to 1 month after which complete occupancy was observed. The soak with **3a** was also extended to 1 month to allow for complete active site occupancy.

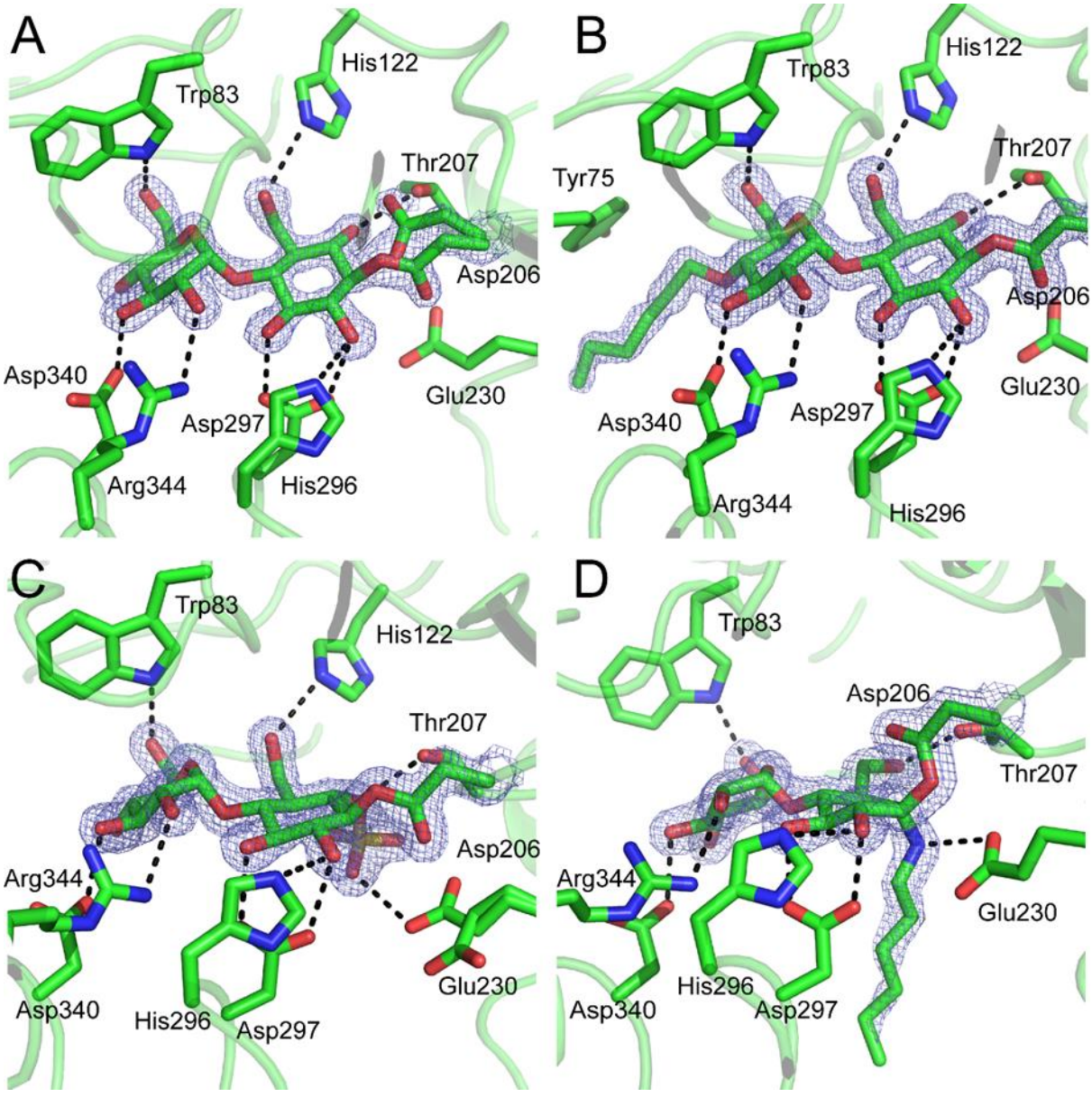


Figure 2.4. Covalent inhibition of Taka-amylase with glycosylated cyclophellitols. Crystal structures are shown between Taka-amylase and each of **1a** (A), **1b** (B), **3a** (C), and **2a** (D). Electron density ($2F_o–F_c$) is shown for both the ligand and the catalytic nucleophile (Asp206) as a blue mesh contoured at 1.5σ (**1a** = 0.59 e^-/ Å^3, **1b** = 0.61 e^-/ Å^3, **2a** = 0.46 e^-/ Å^3, **3a** = 0.54 e^-/ Å^3). The polypeptide is shown

in cartoon form with active site residues shown as sticks. Apparent hydrogen bonding interactions are shown as dotted black lines. Active-site residue Arg204 is omitted for clarity.

The crystal structure of **1a** bound to the active site nucleophile (Asp206) showed the cyclitol moiety to adopt a $^4C_1$ chair conformation, which is also observed for **1b** and **3a**. This conformation matches that observed for the -1 sugar by Caner *et al.*[24] for the binding of enzymatically synthesized **1a** to HPA and the conformation of an *in-situ* generated 5-fluoro-idose trisaccharide, also to HPA.[23] Although a detailed computational analysis of the conformational itineraries for a GH13 α-amylase has yet to be done, a detailed analysis of a homologous GH13 amylosucrase from *Neisseria polysaccharea* has been published.[40] This analysis proposes a $^4C_1 \rightarrow [^4H_3]^{\ddagger} \rightarrow {}^4C_1$ trajectory for the glycosylation step, consistent with the observed conformation of the inhibitor bound in the -1 position. The nucleophile, Asp206, in the **1a** complex structure is present in two conformations, one minor unbound conformation which aligns with the uncomplexed form (PDB: 6YQ7), and the other rotated approximately 120 degrees from the uncomplexed orientation that we ascribe to the conformation present in the covalently inhibited enzyme (Figure 2.4, panel A, PDB: 6YQ9). This second conformation is also seen for the **1b** complex which is at full occupancy. Bound **1a** also forms 9 direct hydrogen bonds with Taka-amylase – with His122, Arg204, Thr207, His296 and Asp297 in the -1 subsite and Trp83, Asp340 and Arg344 in the -2 subsite – replicating those seen in the structure of a complex between Taka-amylase and acarbose (PDB: 7TAA), with the exception of the hydrogen bond between the 6-OH of **1a** and Thr207, and the lack of an H-bond with the catalytic nucleophile that is observed for acarbose.[41]

The structure of 4'-alkylamine *epi*-cyclophellitol **1b** bound in the active site (Figure 2.4, panel B, PDB: 6YQC) maintains an identical hydrogen bonding network to that seen in the complex with **1a**, the only difference being the complete occupancy of the inhibitor, resulting in the absence of an unbound Asp206 conformer, and the presence of an alkyl tail. The 4' tail of **1b** can only be modelled to the sixth carbon (away from the maltose core) indicating that the terminal amine may form multiple interactions in the -3 subsite. The portion of the alkyl tail that could be modelled with confidence is appropriately positioned to form a C-H/π hydrogen bond through the second carbon which is 3.9 Å away from Tyr75. This interaction may in part be responsible for better binding of **1b** to Taka-amylase and the resulting increased rate of inhibition.

The structure of the *epi*-cyclophellitol cyclosulfate **3a** complex has many similarities with complexes of the epoxide inhibitors **1a** and **1b** (Figure 2.4, panel C, PDB: 6YQB). The cyclitol

is covalently attached to Asp206, again rotated approximately 120° from the unreacted enzyme, and is in a $^4C_1$ conformation. The -2 sugar is positioned exactly as for **1a**, however the presence of an exocyclic sulfur at the 6 position has resulted in Thr207 being positioned within hydrogen bonding distance to the O6 that forms the sulfate ester. The sulfate group is also positioned within hydrogen bonding distance to the acid base residue, Glu230, however this residue appears in two different conformations in the structure.

Binding of **2a** to Taka-amylase is strikingly dissimilar to the other three inhibitor complexes (Figure 2.4, panel D, PDB: 6YQA). The inhibitor is again covalently bound to Asp206, however the nucleophile is rotated nearly 180° when compared to the other three complexes, such that the non-bonded δO is positioned above the "β" face of the bound cyclitol ring rather than below it. The hydrogen bonds between the inhibitor and the enzyme -1 subsite are also altered, with the cyclitol methoxy alcohol rotated to form a hydrogen bond with Thr207 – instead of His122 – and the C6 nitrogen forms a hydrogen bond with the acid/base residue (Glu230). Furthermore, the reacted cyclitol, rather than $^4C_1$, is in an unprecedented E_3 conformation, the emergence of which we attribute to the presence of the alkyl-chain. Previous observations underscore the notion that cyclophellitol aziridines with *N*-alkylation are generally well accepted in both *exo*-acting enzymes as these often have open + subsites[28-32] and *endo*-acting enzymes that target *β*-linkages,[33] which typically contain long, tunnel or groove-like active sites. The active site of Taka-amylase, conversely, is a V-shaped surface depression with a severe angle between the -1 and +1 sugar binding sites, this in turn constrains the conformation of the aziridine alkylation and the conformation of the cyclitol ring.

2.2.4 *In vitro* labeling of recombinant human saliva α-amylase with ABPs 1c, 2b and 2c

With the fluorescent ABPs **1c**, **2b** and **2c** in hand, their labeling efficiency and specificity towards recombinant human saliva α-amylase (Type XIII-A, HSA) was investigated next. Kinetic analysis for inactivation of Taka-amylase revealed the parent inhibitors **1a**, **1b** and **2a** as slow enzyme binders. Thus, recombinant α-amylase was incubated with ABPs **1c**, **2b** or **2c** at two quite high concentrations (100 μM or 10 μM, 1 h, 37 °C, pH 7). As shown in Figure 2.5A, HSA was effectively labeled by epoxide-based ABP **1c**, while the labeling potency of aziridine-based ABPs **2b** and **2c** was dramatically reduced, and only very weak bands were observed even at 100 μM probe concentrations. In accordance with the observations from the kinetic studies, epoxide inhibitor **1b** is much more potent than aziridine inhibitor **2a**. This observation is supported by the X-ray structure analysis of inhibitor-enzyme complexes, which indicates that interactions between the -3 subsite in α-amylase active sites and the O4'

substituent may contribute to the initial binding of **1b** and the resulting increased inhibition potency while the alkylation of aziridine warhead in **2a** is not preferred during enzyme binding.

2.2.5 ABPP of retaining α-amylases in complex biological samples

As the next research objective, the visualization of α-amylases was attempted in concentrated human saliva (untreated human saliva and saliva supernatant were also tested in an initial screening, and the results were shown in Figure 2.S2) and lysates from mouse pancreas and salivary gland with ABP **1c**. As shown in Figure 2.5B, ABP **1c** labeled a distinct band at ~57 kDa, corresponding to the molecular weight of HSA.[42] The same band was also observed in mouse pancreas lysates where significant labeling appeared at 25 μM (2 h, 37 °C, pH 7), while at higher probe concentrations background fluorescence was observed. Interestingly, labeling in mouse salivary gland lysates with **1c** showed two clear bands, not only the one at ~57 kDa, but also a slightly stronger band at ~70 kDa. Next, the pH dependence of labeling was determined (Figure 2.5C). ABP **1c** exhibits a similar pH-dependent activity towards α-amylase (at the time unknown ~57 kDa band) in these three biological samples, and the optimal pH for labeling is determined at 5.0, different from the reported neutral pH optimum in HPA[21] and HSA[42] activity. Of note, in mouse salivary gland lysates, the band at ~70 kDa only appeared following labeling at pH ranging from 6.0-8.0 with optimal labeling observed at pH 7.5. Additionally, the efficiency and the rate of labeling of α-amylase (at the time unknown ~57 kDa band) were further investigated under optimal pH 5.0 (Figure 2.5D and 2.5E). In these three samples, α-amylase (the ~57 kDa band) was effectively labeled by **1c** in a concentration- and time-dependent manner, consistent with the irreversible inhibition mechanism one would expect for the maltose cyclophellitols.

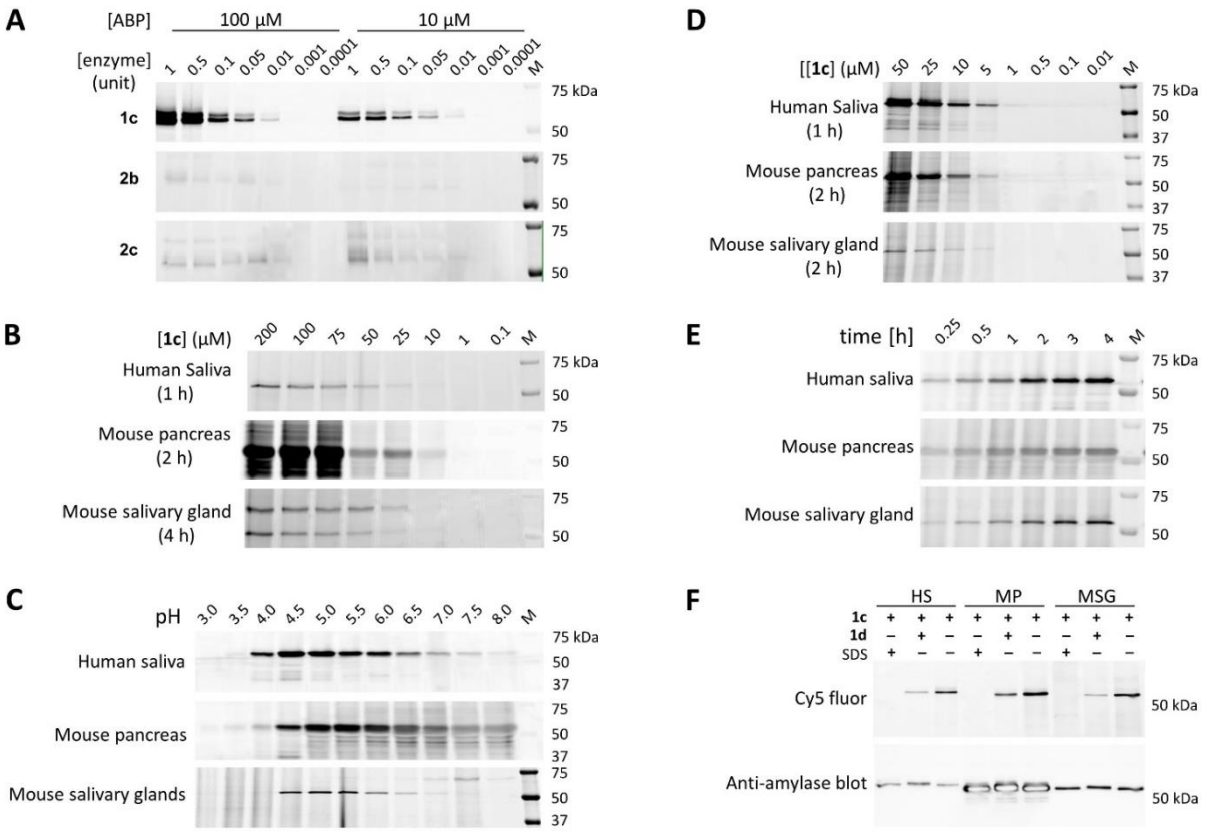


Figure 2.5. A) Fluorescent labeling of recombinant human saliva α-amylase (Type XIII-A) with ABPs **1c**, **2b** and **2c** at pH 7 after 1 h incubation at 37 °C; B) Fluorescent labeling of concentrated human saliva, mouse pancreas and salivary gland lysates (55 µg total protein) with different concentrations of ABP **1c** (pH 7.0, 37 °C); C) pH-dependent labeling with ABP **1c** in complex biological samples. The optimal pH is approximately 5.0 in these three samples; D) Detection limit of α-amylase in complex biological samples labeled with ABP **1c** (pH 5.0, 37 °C); E) Time-dependent labeling of α-amylase in complex biological samples with ABP **1c** (pH 5.0, 37 °C); F) Competitive ABPP on amylases in extracts from human saliva (HS), mouse pancreas (MP) and mouse salivary gland (MSG), offset against Western blot detection of amylase abundance using an anti-amylase antibody.

Having established that amylase activities could be visualized in extracts from human saliva, murine pancreas and murine salivary glands, competitive ABPP (cABPP) experiments were performed to see whether these activities could be negated by inclusion of a non-tagged inhibitor. Pre-incubation of biological samples with inhibitor **1d** followed by labeling with Cy5 ABP **1c** resulted in decrease of fluorescent signals similar to denaturation of the samples with 2% SDS (Figure 2.5F). After fluorescent imaging, proteins were transferred to a polyvinylidene fluoride (PVDF) membrane for Western blot analysis with an anti-amylase antibody. The observed fluorescent bands and chemiluminescence bands match well where no inhibitor is included, with the chemiluminescence bands remaining equal in intensity where the ABP signal

decreases in an inhibitor-dependent fashion, demonstrating that the labeled bands by ABP **1c** can be α-amylase. cABPP assays were also performed with inhibitors **1a** and **1b** in concentrated human saliva (Figure 2.S4). Compared with **1d**, the inhibition potency of these two was reduced and only minor competition was observed with "untagged" **1a** at the highest concentration applied (100 μM), further indicating that alkylation of the O4' position of the epoxide-based inhibitor may improve its inhibition potency.

2.2.6 ABPP of fungal secretomes

The probes developed herein can also be applied to monitor the production of amylases by biotechnologically relevant fungi. Secretomes from the saprophytic ascomycete *Aspergillus nidulans* grown on various starches have previously been shown to contain a complex mixture of amylases, glucosidases and oxidative enzymes.[43] ABPP was performed on secretomes from a time course of *A. nidulans* strain FGSC A4 grown in minimal media supplemented with 1% wheat starch as a carbon source. Secretome samples were taken daily over a 5-day period and the fluorescent Cy5 ABP **1c** was incubated with these secretome samples to reveal a complex pattern of fluorescent bands that could first be detected starting on day 3. The intensity of amylase bands increased after day 3 with a maximum intensity at day 5 (Figure 2.6C).

Conventional industrial starch hydrolysis relies on enzymes that are both thermostable and acid tolerant.[44,45] Both of these metrics can be measured directly from the secretome using ABPP, capturing these parameters in the native enzyme environment. Labeling of the day 5 secretome in the presence of McIlvaine buffers at pH between 3.0 and 8.0 revealed pH dependent labeling patterns (Figure 2.6A). The most intense band – observed at ~66 kDa – was labeled between pH 4.0–7.0 while the bands appearing at ~52 and ~48 kDa were constrained to labeling between 4.0-5.5 and 5.0-7.0 respectively. The temperature-dependent labeling of the day 5 *A. nidulans* secretome at pH 5.0 and at temperatures between 25 and 55 °C was assessed next (Figure 2.6B). Within this range the bands at ~66, ~52 and ~48 kDa all showed temperature dependent labeling, with labeling efficiency decreased above 37 °C, presumably due to denaturation of the corresponding proteins at this temperature. Conversely the bands at ~58 kDa and ~44 kDa remained present at the highest temperature tested.

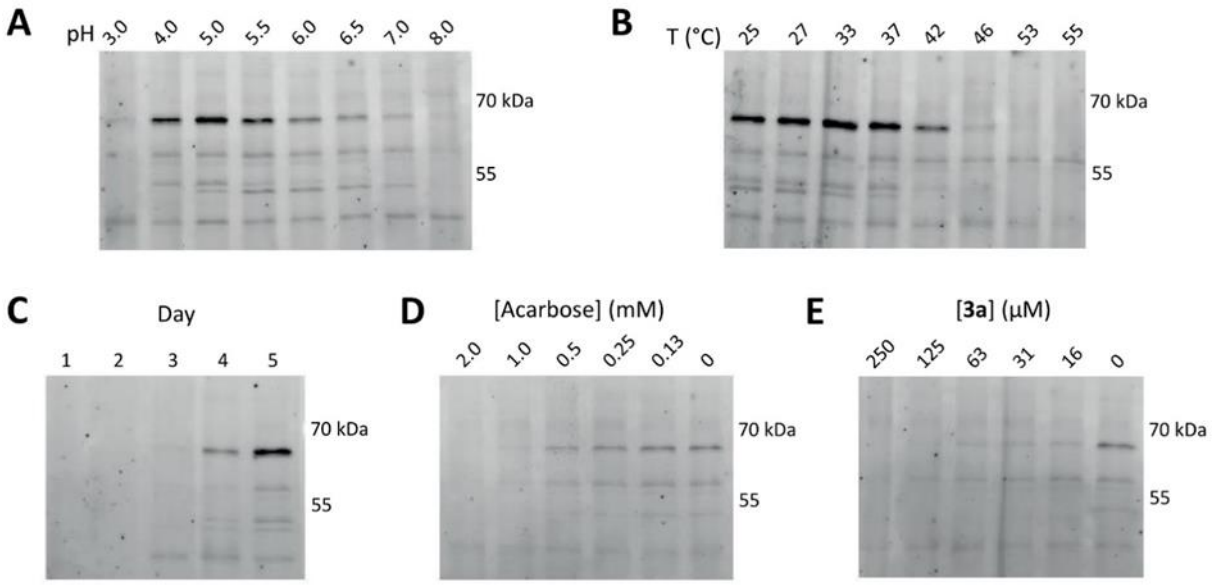


Figure 2.6. A) Fluorescent labeling of day 5 *A. nidulans* secretome with ABP **1c** at pH 3-8 after incubation for 1 h at 37 °C; B) Temperature dependent labeling of day 5 *A. nidulans* secretome with ABP **1c** after incubation at pH 5.0 for 1 h at 25-55 °C; C) Fluorescent labeling of day 1-5 *A. nidulans* secretomes with ABP **1c** at pH 5.0 after incubation for 1 h at 37 °C; D) Competitive ABPP of day 5 *A. nidulans* secretome with acarbose. The secretome was pre-incubated with acarbose at varying concentrations (30 min, pH 5.0, 37 °C) prior to labeling with ABP **1c** (67.5 μM, 1 h, pH 5.0, 37 °C); E) Competitive ABPP of day 5 *A. nidulans* secretome with **3a**. The secretome was pre-incubated with **3a** at varying concentrations (30 min, pH 5.0, 37 °C) prior to labeling with ABP **1c** (67.5 μM, 1 h, pH 5.0, 37 °C).

cABPP was applied next to show the sensitivity of enzyme labeling to both competitive and covalent inhibitors (Figure 2.6D and 2.6E). As mentioned above, cABPP allows for the identification of active enzyme inhibitors, without the need for purification and even if the identity of the active enzyme is unknown.[46,47] Pre-incubation of secretomes for 30 minutes with either acarbose, or **3a** prior to labeling with ABP **1c** resulted in decreased labeling of specific bands in a concentration dependent manner. Enzyme bands at ~66, ~58, ~52 and ~48 kDa were all inhibited by either acarbose or **3a**, while the band at 44 kDa appeared to be insensitive to these α-amylase inhibitors.

2.2.7 Identification of ABP labeled bands

To further determine the specificity of the developed ABPs, ABP-labeled proteins in human saliva samples were analyzed. For this purpose, a human saliva sample was incubated with biotinylated ABP **1d**, a no probe control (DMSO only), or a competitor-inhibited control (first treatment with **1b**). ABP **1d**-labeled proteins were enriched by using magnetic streptavidin

beads. After on-bead tryptic digestion, the resulting peptides were analyzed by LC-MS/MS and matched against the human UniProt database, using the Mascot search engine as previously described.[48] As expected, AMY1A (P04745) was clearly abundant in the ABP **1d** pulldown experiment (Table 2.S2). Other proteins with high scores were background proteins such as enolase standard and keratin contaminations. Pretreatment of the saliva sample with inhibitor **1b** caused a significant loss of AMY1A peptide abundance, and almost no AMY1A was found in the untreated control. These results demonstrate AMY1A in human saliva samples can be selectively labeled and identified via biotinylated ABP **1d**.

Based on the known content of *A. nidulans* secretome,[43] the most prominent ~66 kDa band labeled with **1c** was tentatively assigned as the mature form of α-amylase AmyB (ANIA_03402), see Table 2.S4 for list of mature α-amylase molecular masses. The two minor bands at ~52 and ~48 kDa were also tentatively assigned as AmyA (ANIA_03388) and AmyF (ANIA_0201) α-amylases, respectively. As the wheat starch grown secretome of *A. nidulans* only contains three amylases,[43] the final two minor bands could not be confidently assigned. It can however be speculated that these are proteins that are highly abundant in the secretome.

To identify the labeled enzymes in the *A. nidulans* secretome an activity-based protein pulldown was performed as well. The day 5 secretome was treated with either the biotinylated ABP **1d** or an equivalent volume of DMSO, then a pulldown was performed using streptavidin beads, followed by an on bead tryptic digest and LC-MS/MS analysis. All three amylases from the *A. nidulans* secretome (AmyA, AmyF and AmyB) were detected at elevated levels in the probe **1d**-treated samples relative to the negative control and total secretome, confirming their initial assignment of the major and two minor bands observed by in-gel fluorescence (Table 2.S3). In addition, ChiB, an endochitanase and AgdB, an α-glucosidase were also detected, both of which are highly abundant in the secretome (Table 2.S3). The molecular weights of the detected amylases match well with the labeled bands seen from ABPP with probe **1c**. Furthermore, these results suggest the identities of the minor bands seen at 44 and 58 kDa may belong to ChiB and AgdB respectively – the latter, which is proteolytically processed from its 108 kDa form to a heterodimer with the catalytic protomer having a 58 kDa molecular weight.[49]

2.3 Conclusion

This Chapter describes the synthesis of a panel of *maltobiose*-configured *epi*-cyclophellitol derivatives as mechanism-based covalent inhibitors and ABPs for retaining α-amylases. Kinetic studies of the inhibition of Taka-amylase revealed that cyclosulfate warhead (**3a**) is the most

potent, compared with its epoxide and aziridine counterparts (**1a** and **2a**). As well, alkylation of O4' of the nonreducing end sugar in **1a** led to compound **1b** that inhibits α-amylase up to 10-fold more potently than the parent compound. Compound conformations can strongly influence the inhibitory potencies and selectivities of cyclophellitol analogues. X-ray studies have shown that both epoxide and cyclosulfate-based inhibitors bind to the active site nucleophilic residue of Taka-amylase in a $^4C_1$ chair conformation, whilst the structure of aziridine **2a** complex was found in an unprecedented E_3 conformation, likely due to the presence of the aziridine alkyl-chain, which might account for its lower potency. The inability of fluorescent ABPs **2b** and **2c** to modify recombinant human saliva α-amylase further indicated that substituted aziridines are not accepted by the tested amylases. In contrast, epoxide-based ABP **1c** effectively labels α-amylases in human saliva and mice tissue in a concentration- and time-dependent manner, with optimum labeling at pH 5.0. In addition, it can also detect α-amylases in fungal secretomes and allow easy read out of pH and temperature dependence and susceptibility to inhibitors. Looking ahead, we speculate that installation of reporter entities at O4' of the nonreducing end sugar of cyclosulfate-based inhibitor **3a** would yield more potent ABPs based on kinetic and crystallographic studies. In conclusion, the described maltobiose *epi*-cyclophellitol ABPs allow activity-based profiling of α-amylases in complex biological samples, and may find use in the discovery of new and effective human amylase inhibitors that may have potential as therapeutic reagents for the treatment of type 2 diabetes, and in the discovery of new microbial amylases with potentially advantageous properties for biotechnological application.

2.4 Acknowledgements

Zachary Armstrong and Gideon Davies from University of York, UK are kindly acknowledged for the kinetic, X-ray and fungal secretome labeling experiments, and valuable discussion. Mikkel Rasmussen and Maher Abou Hachem from Technical University of Denmark are kindly acknowledged for providing the *Aspergillus nidulans* secretomes.

2.5 Experimental methods

2.5.1 Biochemical experiments

Materials

Recombinant human saliva α-amylase (Type XIII-A) was obtained as a lyophilized power from Sigma-Aldrich (Product Number: A1031). Taka-amylase was purchased as a solid power from Sigma-Aldrich (Product Number: 86247). Anti-Pancreatic alpha amylase antibody was purchased as a liquid from Abcam (Product Number: ab199132). Concanavalin A beads was purchased from Sigma-Aldrich

(Product Number: C9017). Mouse tissue were isolated according to guidelines approved by the ethical committee of Leiden University (DEC#13191). All the tissue lysates were prepared in potassium phosphate lysis buffer (25 mM in pH 6.5, supplemented with 0.1% (v/v) Triton X-100 and protease inhibitor 1x cocktail (Roche)) via homogenization with silent crusher S equipped with Typ 7 F/S head (30 rpm x 1000, 3 × 7 sec) on ice and lysate concentration was determined with Bicinchoninic acid (BCA) Protein Assay Kit (PierceTM).[50] The protein fractions were stored in small aliquots at -80 °C until use. BCA protein assay kit was acquired from Thermo Fisher Scientific. Trypsin was commercially available from Promega. Dynabeads™ MyOne™ Streptavidin T1 was obtained from Invitrogen. Dithiothreitol (DTT) was obtained from Biochemica and Iodoacetamide (IAA) was obtained from Sigma Aldrich. All other chemicals were obtained from standard commercial sources.

Preparation of human saliva sample

Unstimulated whole saliva sample (20 mL) from a normal subject was expectorated into a 50-mL polypropylene tube. The tube was centrifuged at 4 °C for 10 min at 4000 rpm. Supernatant was collected into a new, ice cold 50-mL tube and stored at -20 °C. The pellets were resuspended into 1 mL ice-cold KPi lysis buffer (25 mM phosphate buffer pH 6.5 supplemented with 0.1% (v/v) Triton X-100 and protease inhibitor 1x cocktail (Roche)). To homogenize, the suspension was sonicated for 60 seconds at 60% amplitude strength on ice. Obtained lysates were stored in aliquots at -20 °C. Protein concentration of lysate was determined by using the BCA method.[50]

Purification of Taka-amylase

To purify Taka-amylase for crystallization, 10 g of powder was dissolved in 200 mL of buffer A (20 mM Tris pH 7.5, 5 mM $CaCl_2$). This suspension was concentrated and washed extensively with buffer A using 10 kDa cutoff Vivaspin centrifugal filter columns (Sigma-Aldrich). Concentrated protein was further purified using a gel-filtration column (HiLoad 16/600 Superdex 75 pg; GE Healthcare), which had been pre-equilibrated with buffer B (20 mM Tris pH 7.5, 5 mM $CaCl_2$, 200 mM NaCl). Peak fractions were pooled, concentrated and washed with buffer A. Protein stocks were diluted to 30 mg/ml and flash frozen in liquid nitrogen.

Inhibition Kinetics

To assess the time dependent inactivation of the synthesized inhibitors samples, Taka-amylase (60 µM) was incubated with 2 mM of either **1a**, **1b**, **2a**, **3a** or buffer alone in assay buffer (20 mM MES, 20 mM NaCl, 5 mM $CaCl_2$ pH 5.0) at 30 °C. Aliquots of these inactivation mixtures were removed at time intervals and diluted 100-fold into assay cells containing assay buffer and 2-chloro-4-nitrophenyl α-maltotrioside (5 mM) pre-incubated at 30 °C, and absorbance change was monitored at 400 nm. Measured residual rates as a function of time for each inactivator were fit to an equation describing first order decay with offset in OriginPro 2019 (OriginLab).

Crystallization of Taka-amylase

Taka-amylase at 30 mg/mL was tested against a range of commercial crystallization screens. Large split crystals were found in 0.1 M sodium citrate pH 5.6, 20% (v/v) 2-Propanol, 20% (w/v) PEG 4,000, a condition that was used for further optimization. The optimized crystals were grown in MRC maxi 48-well plates (Swissci) using the sitting-drop vapor-diffusion method at 20 °C using a well-solution composed of: 0.1 M sodium citrate pH 5.6, 18% (v/v) 2-propanol. Droplets contained a mixture of 500 nL protein stock and 500 nL of well-solution. Crystals appeared within one week. Crystals were soaked in 0.1 M sodium citrate pH 5.6, 25 % (v/v) ethylene glycol, 20 % (w/v) PEG 4,000 prior to flash cooling in liquid nitrogen. Inhibitor complexes for **1a**, **1b**, **2a** and **3a** were obtained by soaking crystals in well solution containing 10 mM of the inhibitor for several days (3 days for **1a** and **1b** and 1 month for **2a** and **3a**) before flash cooling in liquid nitrogen.

All diffraction data were collected at Diamond beamlines (un-liganded and **1a** complex: I03, complexes with **1b**, **2a** and **3a**: I04), and data were processed with the CCP4i2 suite.[51] All structures of Taka-amylase were solved by molecular replacement using the PDB structure 7TAA as the search model in Phaser.[52,53] Refinement was performed using cycles of maximum-likelihood refinement using REFMAC5[54] were interspersed with manual corrections of the models using COOT.[55] Structural figures were drawn with PyMol (DeLano Scientific LLC, http://pymol.sourceforge.net/).

Labeling and SDS-PAGE of recombinant human saliva α-amylase

To prepare for labeling, recombinant human saliva α-amylase stock (1 unit/µL) was diluted with assay buffer (150 mM McIlvaine buffer supplemented with 10 mM $CaCl_2$ and 10 mM NaCl, pH 7.0, if not otherwise specified) to final concentrations of 1, 0.5, 0.1, 0.05, 0.01, 0.001, 0.0001 unit in 10 µL volume. ABPs **1c**, **2b** or **2c** stock (10 mM in DMSO) was diluted with assay buffer (2% (v/v) DMSO) to 3.5x of intended assay concentrations. For labeling, 10 µL enzyme dilution was incubated with 4 µL ABP (**1c**, **2b** or **2c**) dilution at 37 °C for 1 h (Figure 2.5A). To determine the detection limit, 0.2 unit pure enzyme in 10 µL assay buffer (pH 5.0) was incubated with 4 µL ABP **1c** at varying concentrations at pH 5.0 for 4 h, at 37 °C (Figure 2.S5A). The time-dependent labeling was assessed by incubating 0.2 unit pure enzyme in 10 µL assay buffer (pH 5.0) with 4 µL of 175 µM ABP **1c** dissolved in assay buffer pH 5.0 at 37 °C for 10, 30, 60, 120, 180, 240 or 360 min (Figure 2.S5B). Samples were then denatured with 4 µL sample buffer (5x Laemmli buffer, containing 50% (v/v) 1M Tris-HCl pH 6.8, 50% (v/v) glycerol, 10% (w/v) Dithiothreitol (DTT), 10% (w/v) sodium dodecyl sulphate (SDS), 0.01% bromophenol blue) and heated at 98 °C 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 140V 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 and images were

processed using ImageLab 5.2.1 (BioRad). Gels were subsequently extensively stained with Coomassie Brilliant Blue solution for assessing total protein amounts in each lane of sample.

Labeling of untreated human saliva, saliva supernatant and concentrated human saliva

To prepare for labeling, concentrated human saliva pellet (16.9 μg/μL in KPi buffer) was diluted with assay buffer pH 7.0 (2% (v/v) DMSO) to final concentration of 55 μg in 24 μL volume. Untreated saliva and saliva supernatant were thaw on ice and 24 μL of each was directly used for labeling. ABPs **1c** (10 mM in DMSO) was diluted with assay buffer pH 7.0 (2% (v/v) DMSO) to 5x of intended assay concentrations. For labeling, 6 μL of ABP **1c** dilution was incubated with 24 μL enzyme dilution at 37 °C for 1 h. Samples were then denatured with 8 μL sample buffer (5x Laemmli buffer, containing 50% (v/v) 1M Tris-HCl pH 6.8, 50% (v/v) glycerol, 10% (w/v) Dithiothreitol (DTT), 10% (w/v) sodium dodecyl sulphate (SDS), 0.01% bromophenol blue) and heated at 98 °C for 5 minutes, and subjected to SDS-PAGE and fluorescence scan as described above. Results were shown in Figure 2.S2.

Labeling of lysates and SDS-PAGE analysis

Concentrated human saliva pellet (55 μg total protein per sample), mouse pancreas lysates (55 μg total protein per sample), or mouse salivary gland lysates (55 μg total protein per sample) were diluted in assay buffer (pH 7.0, if not otherwise specified) and incubated with ABP **1c** at intended concentrations (2% (v/v) DMSO) at 37 °C for the intended time periods (Figure 2.5B). Influence of pH on ABP labeling was analyzed using 55 μg total protein incubated with 25 μM **1c** (for concentrated human saliva and mouse pancreas) or 50 μM **1c** (for mouse salivary gland), dissolved in assay buffer (2% (v/v) DMSO), pH 3.0−8.0, for 4 h at 37 °C (Figure 2.5C). To determine the detection limit under optimal pH, 55 μg protein was labeled with 0.01−50 μM ABP **1c** in assay buffer pH 5.0 (2% (v/v) DMSO) at 37 °C for the intended time periods (Figure 2.5D). The time-dependent labeling under optimal pH was assessed by incubating 55 μg protein with 10 μM **1c** (for concentrated human saliva and mouse pancreas) or 50 μM **1c** (for mouse salivary gland), dissolved in assay buffer pH 5.0 (2% (v/v) DMSO) at 37 °C for 0.25, 0.5, 1, 2, 3 or 4 h (Figure 2.5E). Samples were then denatured with sample buffer (5x Laemmli buffer, containing 50% (v/v) 1M Tris-HCl pH 6.8, 50% (v/v) glycerol, 10% (w/v) Dithiothreitol (DTT), 10% (w/v) sodium dodecyl sulphate (SDS), 0.01% bromophenol blue) and heated at 98 °C for 5 minutes and subjected to SDS-PAGE and fluorescence scan as described above.

DMSO control experiment

To prepare for labeling, assay buffer pH 5.0 was supplemented with varying concentrations of DMSO (0%, 0.5%, 1% and 2% (v/v)). Concentrated human saliva (16.9 μg/μL in KPi buffer) was diluted with assay buffer containing 0–2% (v/v) DMSO to a final concentration of 55 μg protein in 24 μL volume. ABPs **1c** (10 mM stock in DMSO) was diluted with assay buffer containing 0–2% (v/v) DMSO to 5x of intended assay concentrations. For labeling, 6 μL of ABP **1c** dilution was incubated with 24 μL

enzyme dilution at 37 °C for 1 h. Samples were then denatured with 8 µL Laemmli (5x) sample buffer and heated at 98 °C for 5 minutes, and subjected to SDS-PAGE and fluorescence scan as described above. Results were shown in Figure 2.S3.

Labeling of *A. nidulans* secretome and SDS-PAGE analysis

A. nidulans strain FGSC A4 was grown in minimal medium containing 1 % (w/v) wheat starch (Sigma-Aldrich, S5127) as described previously.[56] Secretome samples were removed daily over a period of five days, flash frozen in liquid nitrogen and stored at -20 ˚C until use in labeling experiments. Secretome samples (12 µL) were diluted with 150 mM McIlvaine buffer of appropriate pH to a final volume of 15 µL, before addition of 3 µL ABP **1c** to a final concentration of 100 µM. Labeling reactions were carried out for 1 h at 37 °C with shaking at 400 rpm, then stopped by addition of 6 µL Laemmli (4X) sample buffer and boiling at 95°C for 5 minutes. 10 µL of each reaction was separated on a 10% SDS-PAGE gel prior to imaging using the Cy5 laser/filter settings on a Typhoon 5 scanner (GE Healthcare). PageRuler Plus Prestained protein ladder (Thermo Fisher Scientific) was used as marker.

Competitive ABPP experiments

Concentrated human saliva (5 µg total protein), mouse pancreas (5 µg total protein) and mouse salivary gland (25 µg total protein) were diluted with assay buffer (pH 5.0, 2% (v/v) DMSO) to a final 10 µL volume, pre-incubated with 2.5 µL **1d** (final 10 µM for human saliva sample; final 50 µM for mouse pancreas and salivary gland samples) at 37 °C for 4 h. Then, 2.5 µL of ABP **1c** was added to a final concentration of 10 µM (for human saliva and mouse pancreas) or 50 µM (for mouse salivary gland). The labeling reaction was incubated at 37 °C for 2 h. Additionally a positive control was performed: the protein was pre-incubated with 2.5 µL assay buffer (pH 5.0, 2% (v/v) DMSO) for 4 h at 37 °C, and subsequently labelled with ABP **1c** at the concentrations described above. A negative control was also performed: 2.5 µL SDS (10 % (w/v)) was added to the 10 µL protein dilution, boiling at 98 °C for 5 min, and incubated with ABP **1c** at 37 °C for 2 h. Samples were then denatured with 4 µL Laemmli (5x) sample buffer and heated at 98 °C for 5 minutes. Proteins were resolved by electrophoresis in 10% SDS-PAGE gels, running at a constant of 90V for 30 minutes followed by 140V 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 and images were processed using ImageLab 5.2.1 (BioRad).

Pre-incubation of secretomes with inhibitors was carried out in 18 µL reactions which contained, 12 µL of secretome, 3 µL McIlvaine buffer pH 5.0 and 3 µL of inhibitor at varying concentrations. After pre-incubation for 30 min at 37°C, ABP **1c** was added to a final concentration of 67.5 µM. This labeling reaction was incubated at 37°C for 1 h with shaking at 400 rpm, then stopped by addition of 5 µL Laemmli (4X) sample buffer and boiling at 95°C for 5 minutes. 10 µL of each reaction was separated on a 10% SDS-PAGE gel prior to imaging using the Cy5 laser/filter settings on a Typhoon 5 scanner

(GE Healthcare). PageRuler Plus Prestained protein ladder (Thermo Fisher Scientific) was used as marker.

Western blotting

Proteins resolved by SDS–PAGE were transferred to a PVDF membrane using a Trans-Blot Turbo system (Bio-Rad). Membrane was blocked in 10 mL of 5% (w/v) powder milk in TBST (50 mM Tris (pH 7.5), 150 mM NaCl, 0.1% Tween 20) for 1 h at room temperature, then incubated with rabbit polyclonal anti-Pancreatic alpha amylase (Abcam; ab199132) at 1:2,000 dilution in TBST (5% (w/v) powder milk) at 4 °C overnight. Membrane was washed 3× with TBST and blocked in 5% (w/v) powder milk in TBST for 30 min at room temperature, then incubated with HRP-conjugated mouse anti-rabbit IgG at 1:5,000 dilution in 5% (w/v) powder milk in TBST for 1 h at room temperature. Membrane was washed again 3× with TBST, 3× with TBS and 2× with Demi-H_2O and blots visualized using Amersham prime ECL reagent (GE Healthcare) and recorded using a Bio-Rad ChemiDoc system.

Detection sensitivity of ABP 1c towards α-amylase in human plasma

To prepare for labeling, recombinant human saliva α-amylase stock (1 unit/μL) was diluted with assay buffer pH 5.0 to final concentrations of 0, 0.0005, 0.005, 0.025, 0.05, 0.5 unit/μL. ABPs **1c** stock (10 mM in DMSO) was diluted with assay buffer pH 5.0 to two different concentrations: 200 μM and 50 μM.

Pre-wash of Concanavalin A (ConA) beads: 940 μL of ConA bead slurry was put in a 15 mL centrifuge tube and washed with 4700 μL of ConA buffer (0.1 M NaOAc, 0.1 M NaCl, 1 mM $MgCl_2$, 1 mM $CaCl_2$, 1 mM $MnCl_2$, pH 6.0. Keep on ice all the time when using). Centrifuging at 4 °C for 1 min and the supernatant (4700 μL) was carefully pipette out. After washing 4x more, 564 μL of ConA buffer was added to the bead slurry and 100 μL of the bead slurry was pipette into a series of 1.5 mL Eppendorf tubes.

For labeling:

Set **A**: 2 μL of enzyme dilution + 22 μL of assay buffer + 8 μL of 200 μM ABP **1c** dilution was incubated at 37 °C for 4 h. Samples were then denatured with 8 μL Laemmli (5x) sample buffer and heated at 98 °C for 5 minutes, and subjected to SDS-PAGE as described above.

Set **B**: 2 μL of enzyme dilution + 22 μL of assay buffer + 8 μL of 200 μM ABP **1c** dilution was incubated at 37 °C for 4 h. During this period, 25 μL of human plasma was added to each of the Eppendorf tubes that contain 100 μL pre-washed ConA bead slurry, binding at 4 °C on a roller for 1 h. After binding, the beads were washed with ConA buffer (5 x 1000 μL) and all the supernatant was carefully pipette out upon the fifth wash. Samples were then denatured with 8 μL Laemmli (5x) sample buffer, vortexed slightly and the final 40 μL mixture was immediately added to the ConA purified

plasma and heated at 98 °C for 5 minutes. After short spin down, the supernatants were collected in new Eppendorf tubes and subjected to SDS-PAGE as described above.

Set **C**: 2 µL of enzyme dilution was spiked into 25 µL of human plasma, which was added to each of the Eppendorf tubes that contain 100 µL pre-washed ConA bead slurry, incubating at 4 °C on a roller for 1 h. After washing, 32 µL of 50 µM ABP **1c** dilution was immediately added to the ConA purified plasma and the mixture was incubated at 37 °C for 4 h. Samples were then denatured with 8 µL Laemmli (5x) sample buffer and heated at 98 °C for 5 minutes. After short spin down, the supernatants were collected in new Eppendorf tubes and subjected to SDS-PAGE as described above. Results were shown in Figure 2.S1.

Pull-down and LC-MS/MS analysis

Proteomics experiment was performed with human saliva in triplicate for the DMSO control, **1b**-inhibited control and **1d** pulldown. Human saliva sample (300 µg total protein) was incubated with either 0.5% (v/v) DMSO or ABP **1d** (5 µM), or firstly with 25 µM inhibitor **1b** followed by 5 µM of ABP **1d**, each step taking 4 h at 37 °C, in a total volume of 200 µL McIlvaine buffer pH 5.0 supplemented with 10 mM $CaCl_2$ and 10 mM NaCl, followed by denaturation by addition of 10% (w/v) SDS 20 µL and heating for 5 minutes at 98 °C. The solutions were centrifuged at 10,000 rpm for 10 min at room temperature and clear supernatants were transferred carefully into new 2-mL Eppendorf tubes, followed by methanol-chloroform precipitation: for each sample, 800 µL methanol was added into the 220 µL protein solution followed by 200 µL chloroform and vortexed vigorously upon each addition. Then 600 µL Milli-Q was added and the suspension was vortexed vigorously prior to centrifuging at 10,000 rpm for 10 minutes at room temperature. The aqueous top layer was removed carefully and 600 µL methanol was added, then the sample was vortexed slightly and centrifuged as above. The supernatant was then removed, and the pellet was left to air-day for 5 min.

The protein pellet was resuspended in 100 µL PBS buffer with 2% (w/v) SDS, vortexed vigorously and heated to 45 °C for 20 minutes with shaking at 600 rpm. 14 µL of 50 mM DTT was added and the samples were incubated at 65 °C for 30 minutes to reduce disulfide bonds, then 200 µL PBS buffer was added and the suspensions were sonicated for 30 seconds at 20% amplitude at 0 °C. Then, 14 µL of freshly-prepared 150 mM iodoacetamide (IAA) was added and the samples were incubated in the dark for 30 minutes at room temperature. The excessive IAA was subsequently quenched by adding 7 µL of 50 mM DTT, incubating for 30 minutes at room temperature. The 335 µL protein suspensions were stepwise diluted with PBS buffer (200 µL/1000 µL/500 µL, vortex vigorously upon each addition) to a final 2-mL volume. Then the resulting 2-mL suspensions were centrifuged at 10,000 rpm for 10 minutes at room temperature and supernatants were transferred carefully into new 2-mL Eppendorf tubes.

The samples were then ready for pull-down with prewashed Dynabeads™ MyOne™ Streptavidin T1 (30 µL) as published previously.[57] After pull-down, all samples were used for on-bead digestion.

The samples were treated with 100 µL on-bead digestion buffer (50 mM NH_4HCO_3 (pH 8), 2% ACN, 1 mM $CaCl_2$, 2.5 ng/µL of trypsin) and incubated overnight at 37 °C. The supernatants, containing tryptic-digested peptides were then transferred to new 2-mL low-binding tubes and desalted using StageTips. Consequently, the acetonitrile was evaporated using a SpeedVac at 45 °C followed by addition of 50 µL of LC-MS sample solution (97:3:0.1, H_2O:acetonitrile:formic acid) for LC-MS analysis. All peptide samples were analyzed 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.[58] Raw data was calculated using MaxQuant against UniProt of human saliva proteome database to obtain an identification list of found proteins. The abundance of the protein hits was quantified as previously described,[48] in unsupervised mode using the default settings of the PLGS (Waters) and IsoQuant software. Raw and quantified data can be found in Table 2.S2.

Fungal Secretome Proteomics

Proteomic analysis was performed in triplicate for the DMSO control, **1d** pulldown and total secretome. *A. nidulans* day 5 wheat secretome was concentrated 12.5-fold using a Vivaspin centrifugal concentrator (3 kDa cut-off). Concentrated secretome (40 µl) was buffered with 30 mM McIlvane buffer pH 5.0 and transferred to a Lo-bind 0.5 ml tube (Eppendorf). The buffered secretome was then treated with either 300 µM **1d** or DMSO control. Reactions were incubated 4 hours at 30 °C with shaking at 450 rpm. Labeling was terminated by the addition of DTT and SDS to a final concentration of 4 mM and 0.2 % (v/v) respectively, and heating to 95 °C for 5 minutes. The denatured proteins were then cooled to 25 °C, IAA was added to a final concentration of 24 mM then samples were incubated in the dark for 30 minutes at 25°C. IAA, DTT, SDS, buffer and excess probe were removed by acetone precipitation; 4 volume equivalents of acetone (-20 °C) was added to the sample and the mixture was vortexed. The sample was then placed at -20 °C for one hour, then proteins were pelleted by centrifugation at 14,000 x *g* for 1 minute. The supernatant was removed, then the pellet was washed with the same volume of -20 °C acetone then centrifuged again at 14,000 x *g* for 1 minute. The supernatant was then removed, and the pellet was left to briefly air dry.

Samples were then resuspended 50 µL of 4 M urea, then diluted with 450 µL of 0.05% SDS in phosphate buffered saline. Strep Mag Sepharose beads (GE healthcare; 20 µL; washed twice with phosphate buffered saline (PBS) were added to each sample, followed by incubation at 25 °C with constant agitation to prevent bead settling for 3 hours. Beads were collected using a magnetic rack and the supernatant was discarded. Beads were washed with 1 mL of 2 % SDS at 25 °C then with 1 mL of 2 % SDS at 65 °C to eliminate non-specific binding, followed by 1 ml of 2 M urea for 5 minutes to remove SDS and then 1 mL of water two times to remove residual SDS and urea. The washed beads were then resuspended in 20 µL of 0.05 M triethylamonium bicarbonate and 1 µL of 0.5 µg/µL trypsin (Promega V5280) in 50 mM acetic acid was added. The on-bead digest was incubated overnight at 37 °C. Following the tryptic digest, the beads were immobilized and released peptides were removed. The total

secretome samples were treated as above, except the acetone pellet was resuspended directly in 0.05 M triethylamonium bicarbonate and digested with trypsin in solution.

The samples were desalted via a C18 ZipTip then resolubilised with aqueous 0.1% TFA, before injection onto a 50 cm EasyNano C18 PepMap column. Peptides were eluted at a flow rate of 300 nL/min over 1 hour acquisitions onto an Orbitrap Fusion Tribrid mass spectrometer operated in DDA mode. Tandem mass spectra were searched using both Mascot and X!Tandem search programs against the *Emericella nidulans* strain FGSC A4 subset of the Uniprot database and typical contaminants. Carboxymethylation of Cysteine was set as a fixed modification and oxidation of Met was set as variable. Resulting spectral matches were combined in Scaffold and filtered to require a global protein and PSM false discovery rate of 1%, and a minimum of two unique peptides per protein group identification. Label-free quantification was performed using Progenesis QI (Waters). Following chromatographic alignment, peaks were integrated and assigned. Protein abundance was estimated using the integrated intensity of non-conflicting peptides. Results of this analysis for all identified proteins can be found in Table 2.S3.

2.5.2 Chemical synthesis

General experimental details

All reagents were of experimental grade and were used without further purification unless stated otherwise. Dichloromethane (DCM) and tetrahydrofuran (THF) were stored over 3 Å molecular sieves and *N*,*N*-dimethylformamide (DMF) was stored over 4 Å molecular sieves, which were dried *in vacuo* before use. All reactions were performed under an Argon or N_2 atmosphere unless stated otherwise. Reactions were monitored by analytical thin layer chromatography (TLC) using Merck aluminum sheets pre-coated with silica gel 60 with detection by UV-absorption (254 nm) and by spraying with a solution of $(NH_4)_6Mo_7O_{24}\cdot H_2O$ (25 g/L) and $(NH_4)_4Ce(SO_4)_4\cdot H_2O$ (10 g/mL) in 10% sulfuric acid followed by charring at ~150 °C or by spraying with an aqueous solution of $KMnO_4$ (7%) and K_2CO_3 (2%) followed by charring at ~150 °C. Column chromatography was performed manually using Screening Device b.v. silica gel 60 (0.04-0.063 mm) in the indicated solvents. LC-MS analysis was performed on a LCQ Advantage Max (Thermo Finnigan) ion-trap spectrometer (ESI^+) coupled to a Surveyor HPLC system (Thermo Finnigan) equipped with a C18 column (Gemini, 4.6 mm x 50 mm, 5 µM particle size, Phenomenex). The applied buffers were H_2O, acetonitrile (MeCN) and 1% aqueous trifluoroacetic acid (TFA). ^{1}H-NMR and ^{13}C-NMR spectra were recorded on Bruker AV-400 (400/101 MHz), Bruker AV-500 (500/126 MHz), and Bruker AV-850 (850/214 MHz) spectrometers in the given solvent. Chemical shifts (δ) are given in ppm relative to tetramethylsilane (TMS) as internal standard (^{1}H NMR in $CDCl_3$) or the residual signal of the deuterated solvent. Coupling constants (*J*) are given in Hz. All given ^{13}C-NMR spectra are proton decoupled. The following abbreviations are used to describe peak patterns when appropriate: s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), Ar (aromatic), C_q (quarternary

carbon). 2D NMR experiments (COSY, HSQC) were carried out to assign protons and carbons of the new structures. 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).

Experimental Procedures and Characterization Data of Products

Known compounds **4a**[36], **5**[33], **S1**[36], **S2**[33], Cy5-OSu[59], biotin-OSu[60] were synthesized following procedures previously described and their spectroscopic data are in agreement with those previously reported.

Standard procedure A: glycosylation with imidate donor (4a)

The donor (1.5 equiv.) was co-evaporated with toluene (3 x) and dissolved in dry DCM (0.05 M) under nitrogen and stirred over fresh flame-dried molecular sieves 3 Å, after which DMF (16 equiv. of donor) was added to the solution. The solution was cooled to -20 °C and TfOH (1.5 equiv.) was added. After 1 h, the pre-activation was complete as indicated by TLC-analysis. Acceptor (1.0 equiv.) was added to the solution and the mixture was placed in an ice bath. The reaction was stirred at 0 °C until TLC-analysis showed complete conversion of the acceptor. The reaction was quenched with Et_3N, filtered and concentrated *in vacuo*. The products were purified by size exclusion (eluent MeOH/DCM: 1/1) and silica gel column chromatography (See experimental description below for eluent system). *Note: for the glycosylation of donor **4a** with acceptor **5**, the reaction was first kept at -20 °C for 24 h then warmed to 0 °C and stirred until TLC-analysis showed complete conversion of the acceptor.*

Standard procedure B: glycosylation with thio-donors (4b and 4c)

The donor (1.5 equiv.) was co-evaporated with toluene (3 x) and dissolved in dry DCM (0.05 M) under nitrogen and stirred over fresh flame-dried molecular sieves 3 Å, after which DMF (16 equiv. of donor) was added to the solution. The solution was cooled to -20 °C, NIS (1.5 equiv.) and TfOH (1.5 equiv.) were added successively. The mixture was pre-activated at -20 °C for 2 h. Then acceptor (1.0 equiv.) was added to the solution and the mixture was placed in an ice bath. The reaction was stirred at 0 °C until TLC-analysis showed complete conversion of the acceptor. The reaction mixture was diluted with DCM and was quenched with sat. aq. $Na_2S_2O_3$. The organic phase was washed with water and brine, dried with anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The products were purified by size exclusion (eluent MeOH/DCM: 1/1) and silica gel column chromatography (See experimental description below for eluent system).

General procedure C: Deprotection of PMB protecting group

The starting material (1.0 equiv.) was dissolved in DCM/H_2O (19/1, 0.05 M). DDQ (1.2 equiv.) was added to the mixture. The reaction was stirred at rt until TLC-analysis indicated full consumption of the starting material (± 4 h). Then the mixture was diluted with DCM and washed with sat. aq. $NaHCO_3$ (2 x), water (1 x) and brine, dried with anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The product was purified by silica gel column chromatography (See experimental description below for eluent system).

General procedure D: Epoxidation with *m*-CPBA

The starting material (1.0 equiv.) was dissolved in dry DCM (0.05 M) and cooled 0 °C, *m*-CPBA (6.0 equiv.) was added and the reaction was stirred at rt overnight. The mixture was diluted with DCM and washed with sat. aq. $Na_2S_2O_3$ (1 x), sat. aq. $NaHCO_3$ (1 x), water and brine successively, dried with anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The crude was directly used for next step (silylation) without further purification.

General procedure E: Protection with TBS group

The starting material (1.0 equiv.) was dissolved in dry DCM (0.1 M) and cooled 0 °C, imidazole (3.0 equiv.), TBSCl (2.0 equiv.) and DMAP (0.1 equiv.) were added successively and the reaction was stirred at rt overnight. The mixture was diluted with DCM, washed with water (3 x) and brine, dried with anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The product was purified by silica gel column chromatography (See experimental description below for eluent system).

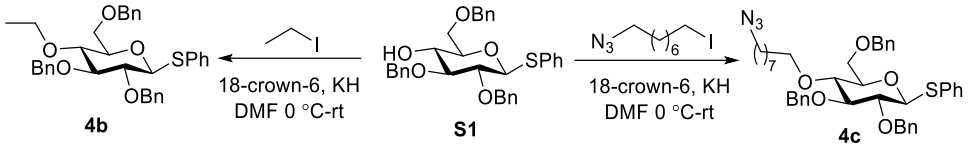
General procedure F: Deprotection of TBS group

The starting material (1.0 equiv.) was dissolved in dry THF (0.1 M), TBAF (1.0 M in THF, 6.0 equiv.) was added. The reaction was stirred at rt until TLC-analysis indicated full consumption of the starting material (± 2h). The reaction was quenched sat. aq. NH_4Cl, extracted with DCM (2 x), the combined organic layers were washed with sat. aq. $NaHCO_3$, water and brine, dried with anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The product was purified by silica gel column chromatography (See experimental description below for eluent system).

General procedure G: Protection with Bn group

The starting material (1.0 equiv.) was dissolved in dry DMF (0.1 M) and cooled to 0 °C. To the mixture was added NaH (60% in mineral, 2.0 equiv.) and stirred at 0 °C for 30 min. Then BnBr (1.5 equiv.) and TBAI (0.05 equiv.) were added successively and the mixture was warmed to rt and stirred overnight. The reaction was quenched with H_2O at 0 °C, extracted with EtOAc (2 x), the combined organic layers were washed with water (2 x) and brine, dried with anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The product was purified by silica gel column chromatography (See experimental description below for eluent system).

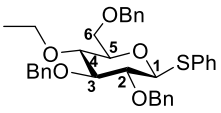
Scheme 2.2. Preparation of donor **4b** and **4c**



For **4b**: KH (30 wt%, 0.36 g, 2.7 mmol) was dispersed in dry DMF (18 mL), purged with N_2 and cooled to 0 °C. Compound **S1** (1.0 g, 1.8 mmol) was added to the mixture in several portions over a period of 15 min, then the reaction was warmed to rt and stirred for 30 min. After which the mixture was cooled to 0 °C again, 18-crown-6 (97 mg, 0.37 mmol) and iodoethane (0.44 mL, 5.5 mmol) were added. After stirring at rt overnight, the reaction was carefully quenched with H_2O at 0 °C, diluted with water (80 mL), extracted with EtOAc (2 x 100 mL), the combined organic layers were washed with water (2 x 150 mL), brine, dried over Na_2SO_4, filtered and concentrated *in vacuo*. The product was purified by flash column chromatography (pentane/EtOAc 20:1→12:1) affording compound **4b** (906 mg, 1.59 mmol, 86%) as a white solid.

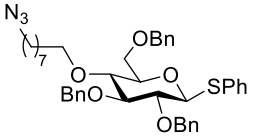
For **4c**: Starting from compound **S1** (1.50 g, 2.76 mmol) and 1-azido-8-iodooctane (1.17 g, 4.15 mmol), the reaction was carried out following the same procedure described for **4b**. The product was purified by flash column chromatography (pentane/EtOAc 20:1→15:1) affording compound **4c** (1.27 g, 1.82 mmol, 66%) as a yellow oil.

Compound 4b



^{1}H NMR (400 MHz, $CDCl_3$) δ 7.63 – 7.53 (m, 2H, CH Ar), 7.41 – 7.24 (m, 15H, CH Ar), 7.23 – 7.17 (m, 3H, CH Ar), 4.90 – 4.81 (m, 3H, CH*H* Bn), 4.71 (d, *J* = 10.3 Hz, 1H, CH*H* Bn), 4.65 (d, *J* = 9.6 Hz, 1H, H1), 4.56 (dd, *J* = 29.6, 11.9 Hz, 2H, CH*H* Bn), 3.86 – 3.76 (m, 2H, H6a and CH_3CH*H*O), 3.75 – 3.69 (m, 1H, H6b), 3.60 (m, 2H, H4 and CH_3CH*H*O), 3.49 – 3.41 (m, 3H, H2, H3 and H5), 1.13 (t, *J* = 7.0 Hz, 3H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.5, 138.4, 138.2, 133.9 ($4C_q$ Ar), 132.0, 129.0, 128.6, 128.5, 128.4, 128.3, 128.0, 127.8, 127.7, 127.6, 127.5 (20CH Ar), 87.5 (C1), 86.7 (C4), 80.7 (C2), 79.3 (C3), 78.1 (C5), 75.9, 75.5, 73.5 ($3CH_2$ Bn), 69.1 (C6), 68.5 (O*C*H_2CH_3), 15.9 (CH_3). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{35}H_{38}O_5SNa$ 593.2332, found 593.2330.

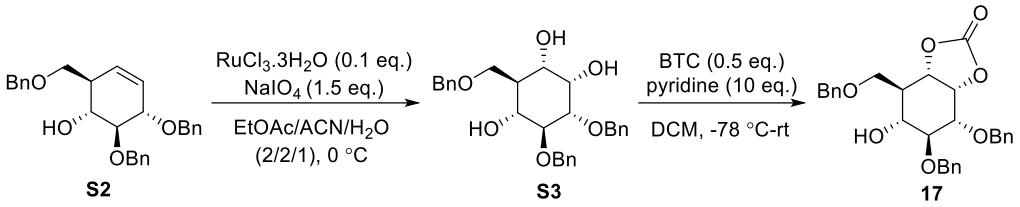
Compound 4c



^{1}H NMR (400 MHz, $CDCl_3$) δ 7.58 (dd, *J* = 6.6, 3.0 Hz, 2H, CH Ar), 7.44 – 7.11 (m, 18H, CH Ar), 4.90 – 4.78 (m, 3H, CH*H* Bn), 4.70 (d, *J* = 10.3 Hz, 1H, CH*H* Bn), 4.65 (d, *J* = 9.6 Hz, 1H, H1), 4.56 (dd, *J* = 29.6, 11.9 Hz, 2H, CH*H* Bn), 3.82 – 3.66 (m, 3H, H6a, H6b and O*CH*H linker), 3.61 (t, *J* = 8.4 Hz, 1H, H4), 3.55 – 3.37 (m, 4H, O*CH*H linker, H2, H3 and H5), 3.22 (t, *J* = 7.0 Hz, 2H, CH_2N_3), 1.66 – 1.40 (m, 4H, $2CH_2$ linker), 1.39 – 1.12 (m, 8H, $4CH_2$ linker). ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.6, 138.5,

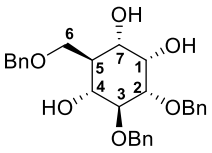
138.2, 134.0 (4C_q Ar), 132.0, 129.0, 128.5, 128.4, 128.3, 128.0, 127.9, 127.8, 127.7, 127.6, 127.5 (20CH Ar), 87.5 (C1), 86.8 (C4), 80.8 (C2), 79.4 (C3), 78.1 (C5), 75.9, 75.5, 73.5 (3CH_2 Bn), 73.3 (OCH_2 linker), 69.2 (C6), 51.5 (CH_2N_3), 30.4, 29.5, 29.2, 28.9, 26.8, 26.2 (6CH_2 linker). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{41}H_{49}N_3O_5SNa$ 718.3285, found 718.3281.

Scheme 2.3. Preparation of acceptor **17**



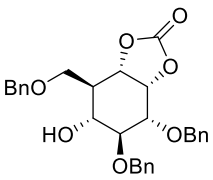
Cyclohexane **S2** (1.94 g, 4.51 mmol) was dissolved in EtOAc (25 mL) and MeCN (25 mL) and cooled to 0 °C. A solution of $RuCl_3.3H_2O$ (94 mg, 0.45 mmol) and $NaIO_4$ (1.45 g, 6.78 mmol) in H_2O (9 mL) was added and the mixture was stirred at 0 °C vigorously for 2 h. The reaction was quenched by addition of sat. aq. $Na_2S_2O_3$ (10 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 80 mL). The combined organic layers were washed with brine, dried over Na_2SO_4, filtered and concentrated *in vacuo*. The crude was purified by flash column chromatography (pentane/EtOAc, 5:1 → 2:1) affording compound **S3** (750 mg, 1.61 mmol, 36%) as a clean oil. Compound **S3** (565 mg, 1.21 mmol) was dissolved in dry DCM (5 mL) and pyridine (1.0 mL, 12 mmol) was added. The mixture was purged with N_2 and cooled to -78 °C. A solution of triphosgene (0.18 g, 0.61 mmol) in dry DCM (2 mL) was added to the mixture at -78 °C. Then the reaction was stirred at rt for 1 h and treated with sat. aq. NH_4Cl (60 mL), extracted with DCM (2 x 80 mL). The combined organic layers were washed successively with water, sat. aq. $NaHCO_3$, water and brine, dried over Na_2SO_4, filtered and concentrated *in vacuo*. The crude was purified by flash column chromatography (pentane/acetone 11:1→7:1) affording compound **17** (0.46 g, 0.94 mmol, 78%) as a yellow oil.

Compound S3



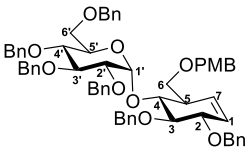
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.49 – 7.11 (m, 15H, CH Ar), 4.99 (d, J = 11.4 Hz, 1H, CH*H* Bn), 4.75 – 4.62 (m, 3H, CH*H* Bn), 4.56 – 4.45 (m, 2H, CH*H* Bn), 4.13 (t, J = 2.7 Hz, 1H, H1), 3.88 (ddd, J = 9.2, 4.0, 1.2 Hz, 1H, H6a), 3.80 – 3.67 (m, 2H, H3 and H6b), 3.53 (dt, J = 10.8, 1.7 Hz, 1H, H7), 3.42 – 3.29 (m, 2H, H2 and H4), 2.89 – 2.63 (br s, 2H, 2OH), 2.16 (tdd, J = 10.5, 6.1, 3.9 Hz, 1H, H5). ^{13}C NMR (126 MHz, $CDCl_3$) δ 138.8, 137.9, 137.9 (3C_q Ar), 128.6, 128.6, 128.5, 128.0, 128.0, 127.9, 127.9, 127.7 (15CH Ar), 82.4 (C3), 79.9 (C2), 75.5, 73.7, 72.1 (3CH_2 Bn), 70.3 (C4), 70.2 (C1), 70.17 (C7), 70.0 (C6), 42.5 (C5). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{28}H_{32}O_6Na$ 487.2091, found 487.2091.

Compound 17



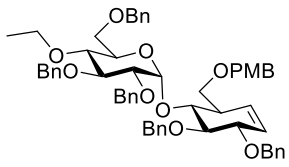
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.20 (m, 15H, CH Ar), 4.82 – 4.72 (m, 2H, H1 and H7), 4.70 – 4.54 (m, 4H, CH*H* Bn), 4.52 (s, 2H, CH*H* Bn), 3.89 – 3.78 (m, 2H, H6a and H2), 3.73 – 3.64 (m, 2H, H3 and H4), 3.61 (dd, *J* = 9.5, 3.9 Hz, 1H, H6b), 2.78 (s, 1H, OH), 2.44 – 2.30 (m, 1H, H5). ^{13}C NMR (101 MHz, $CDCl_3$) δ 154.6 (C=O), 137.7, 137.7, 137.1 ($3C_q$ Ar), 128.7, 128.6, 128.6, 128.3, 128.1, 128.0, 128.0, 127.8, 127.8 (15CH Ar), 81.3 (C3), 76.0 (C2), 75.0 (C1), 74.4 (C7), 73.6, 73.4, 73.2 ($3CH_2$ Bn), 69.6 (C4), 66.7 (C6), 43.2 (C5). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{29}H_{30}O_7Na$ 513.1884, found 513.1887.

Compound 6a



Starting from compound **5** (0.46 g, 1.0 mmol), using donor **4a** (1.1 g, 1.5 mmol) and following **Standard procedure A**. The crude was divided into two portions and purified by size exclusion (DCM/MeOH = 1/1) giving the target product as a single isomer (970 mg, α/β > 20/1). Then the compound was further purified by flash column chromatography (pentane/EtOAc 13:1→9:1) to obtain compound **6a** (0.95 g, 0.96 mmol, 96%) as a clean oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.03 (m, 32H, CH Ar), 6.83 – 6.78 (m, 2H, CH Ar), 5.74 (tt, *J* = 4.8, 2.4 Hz, 2H), 5.63 (dt, *J* = 10.1, 2.1 Hz, 1H), 5.00 (d, *J* = 12.1 Hz, 1H), 4.95 – 4.87 (m, 2H), 4.86 – 4.74 (m, 3H), 4.72 – 4.22 (m, 12H), 4.10 (t, *J* = 9.1 Hz, 1H), 4.03 – 3.86 (m, 2H), 3.79 (d, *J* = 16.4 Hz, 5H), 3.71 – 3.65 (m, 1H), 3.65 – 3.58 (m, 2H), 3.56 – 3.45 (m, 3H), 2.67 (ddt, *J* = 8.2, 5.3, 2.7 Hz, 1H, H5). ^{13}C NMR (101 MHz, $CDCl_3$) δ 159.2, 139.3, 139.0, 138.6, 138.4, 138.2, 138.0, 130.5 ($8C_q$ Ar), 129.8 (C7), 129.2, 129.0, 128.4, 128.4, 128.4, 128.3, 128.3, 128.3, 128.2, 128.1, 128.0, 128.0, 127.9, 127.9, 127.9, 127.8, 127.8, 127.8, 127.7, 127.7, 127.6, 127.6, 127.3, 127.1, 127.0, 126.8 (32CH Ar), 126.6 (C1), 113.9, 113.8 (2CH Ar), 97.0 (C1'), 85.0, 82.0, 80.8, 79.6, 77.9, 75.6, 75.1, 74.1, 73.6, 73.2, 72.8, 72.5, 71.8, 70.9, 69.6 (C6), 68.4 (C6'), 55.4 (CH_3 PMB), 43.9 (C5). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{63}H_{66}O_{10}Na$ 1005.4548, found 1005.4550.

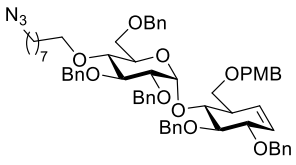
Compound 6b



Starting from compound **5** (0.46 g, 1.0 mmol), using donor **4b** (0.86 g, 1.5 mmol) and following **Standard procedure B**. The crude was divided into two portions and purified by size exclusion (DCM/MeOH = 1/1) giving the target product as a single isomer (845 mg, α/β > 20/1). Then the compound was further purified by flash column chromatography (pentane/EtOAc 13:1→9:1) to obtain compound **6b** (0.81 g, 0.88 mmol, 88%) as a clean oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37 – 7.06 (m, 27H, CH Ar), 6.86 (d, *J* = 8.3 Hz, 2H, CH Ar), 5.78 – 5.67 (m, 2H, H1 and H1'), 5.66 – 5.60 (m, 1H, H7), 4.98 (d, *J* = 12.0 Hz, 1H, CH*H* Bn), 4.92 – 4.84 (m, 2H, CH*H* Bn), 4.77 (d, *J* = 10.9 Hz, 1H, CH*H* Bn), 4.59 (td, *J* = 12.7, 12.1, 10.5 Hz, 3H, CH*H* Bn), 4.49 – 4.28 (m, 5H, CH*H* Bn), 4.28 – 4.22 (m, 1H, H2), 4.08 (t, *J* = 9.1 Hz, 1H, H4), 3.94 – 3.85 (m, 2H, H3 and H3'), 3.78 (s, 5H, 1CH*H* Et, CH_3 PMB and H5'), 3.66 –

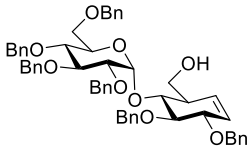
3.39 (m, 7H, H6ab, H6'ab, H4', H2' and 1CH*H* Et), 2.67 (dq, *J* = 6.8, 3.5 Hz, 1H, H5), 1.07 (q, *J* = 6.7 Hz, 3H, CH_3 Et). ^{13}C NMR (101 MHz, $CDCl_3$) δ 159.2, 139.4, 139.0, 138.4, 138.3, 138.1, 130.5 (7C_q Ar), 129.8 (C7), 129.1, 128.4, 128.4, 128.4, 128.3, 128.3, 128.2, 128.2, 128.18, 128.0, 128.0, 128.0, 127.9, 127.9, 127.8, 127.8, 127.7, 127.7, 127.6, 127.5, 127.5, 127.1, 126.8 (27CH Ar), 126.6 (C1), 113.8 (2CH Ar), 97.1 (C1'), 85.0 (C3), 81.9 (C3'), 80.8 (C2), 79.4, 77.8, 75.6 , 74.1, 73.6, 73.4 (C4), 72.9, 72.5, 71.8, 71.0 (C5'), 69.7 (C6), 68.4 (CH_2 Et), 68.3 (C6'), 55.4 (CH_3 PMB), 43.9 (C5), 16.0 (CH_3 Et). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{58}H_{64}O_{10}Na$ 943.4392, found 943.4385.

Compound 6c



Starting from compound **5** (560 mg, 1.20 mmol), using donor **4c** (1.27 g, 1.80 mmol) and following **Standard procedure B**. The crude was divided into two portions and purified by size exclusion (DCM/MeOH = 1/1) giving the target product as a single isomer (1.21 g, α/β > 20/1). Then the compound was further purified by flash column chromatography (pentane/EtOAc 13:1→9:1) to obtain compound **6c** (1.10 g, 1.05 mmol, 87%) as a clean oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.40 – 7.02 (m, 27H, CH Ar), 6.90 – 6.81 (m, 2H, CH Ar), 5.77 – 5.68 (m, 2H, H1 and H1'), 5.63 (dt, *J* = 10.1, 2.1 Hz, 1H, H7), 4.98 (d, *J* = 12.0 Hz, 1H, CH*H* Bn), 4.94 – 4.82 (m, 2H, CH*H* Bn), 4.75 (d, *J* = 10.9 Hz, 1H, CH*H* Bn), 4.65 – 4.28 (m, 8H, CH*H* Bn), 4.25 (dt, *J* = 7.4, 2.4 Hz, 1H, H2), 4.08 (t, *J* = 9.1 Hz, 1H, H4)), 3.95 – 3.83 (m, 2H, H3 and H3'), 3.81 – 3.67 (m, 5H, 1CH*H*O linker, CH_3 PMB and H5'), 3.66 – 3.29 (m, 7H, H6ab, H6'ab, H4', H2', 1CH*H*O linker), 3.20 (t, *J* = 7.0 Hz, 2H, CH_2N_3), 2.67 (dp, *J* = 8.3, 2.8 Hz, 1H, H5), 1.54 (p, *J* = 7.0 Hz, 2H, 2CH*H* linker), 1.43 (h, *J* = 7.8, 6.9 Hz, 2H, 2CH*H* linker), 1.26 (ddd, *J* = 18.1, 12.9, 5.5 Hz, 8H, 8CH*H* linker). ^{13}C NMR (101 MHz, $CDCl_3$) δ 159.2, 139.3, 139.1, 138.4, 138.2, 138.1, 130.5 (7C_q Ar), 129.8 (C7), 129.1, 128.4, 128.4, 128.4, 128.3, 128.2, 128.2, 128.1, 128.0, 127.9, 127.8, 127.8, 127.7, 127.7, 127.6, 127.5, 127.5, 127.1, 126.8 (27CH Ar), 126.6 (C1), 113.8 (2CH Ar), 97.1 (C1'), 85.0 (C3), 81.9 (C3'), 80.8 (C2), 79.4, 77.8, 75.6, 74.1, 73.6, 73.3 (C4), 73.2 (CH_2O linker), 72.9, 72.5, 71.8, 71.0 (C5'), 69.6 (C6), 68.4 (C6'), 55.4 (CH_3 PMB), 51.5 (CH_2N_3), 43.9 (C5), 30.5, 29.5, 29.2, 28.9, 26.8, 26.2 (6CH_2 linker). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{64}H_{75}N_3O_{10}Na$ 1068.5345, found 1068.5342.

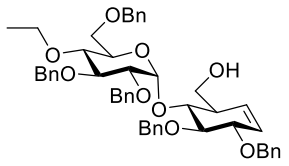
Compound 7a



Starting from compound **6a** (1.04 g, 1.06 mmol), and following **General procedure C**, the reaction was purified by flash column chromatography (pentane/EtOAc 5:1→2:1) to obtain compound **7a** (0.73 g, 0.85 mmol, 80%) as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.41 – 7.12 (m, 26H, CH Ar), 7.11 – 6.98 (m, 4H, CH Ar), 5.86 – 5.75 (m, 2H, H1 and H1'), 5.65 (dt, *J* = 10.1, 2.1 Hz, 1H, H7), 5.05 (d, *J* = 12.2 Hz, 1H, CH*H* Bn), 4.88 (dd, *J* = 11.6, 8.1 Hz, 2H, CH*H* Bn), 4.75 (dd, *J* = 10.9, 3.1 Hz, 2H, CH*H* Bn), 4.65 – 4.43 (m, 6H, CH*H* Bn), 4.38 (d, *J* = 10.8 Hz, 1H, CH*H* Bn), 4.27 (ddt, *J* = 7.3, 3.8, 1.9 Hz,

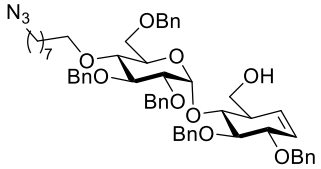
1H, H2), 4.18 (t, *J* = 9.2 Hz, 1H, H4), 4.03 – 3.87 (m, 4H, H3, H3', H5' and H6a), 3.71 – 3.61 (m, 2H, H6'a and H6b), 3.47 (dd, *J* = 9.8, 4.0 Hz, 1H, H2'), 3.40 (dd, *J* = 9.8, 7.1 Hz, 1H, H6'b), 3.29 (dd, *J* = 10.2, 8.7 Hz, 1H, H4'), 3.10 (br s, 1H, OH), 2.57 (dp, *J* = 8.8, 3.0 Hz, 1H, H5). ^{13}C NMR (101 MHz, $CDCl_3$) *δ* 139.4, 138.7, 138.3, 138.1, 137.8, 137.3 ($6C_q$ Ar), 129.8 (C7), 128.6, 128.6, 128.5, 128.5, 128.3, 128.3, 128.3, 128.3, 128.1, 128.1, 128.0, 127.9, 127.7, 127.7, 127.6, 127.5 (C1), 127.0, 126.5, 97.3 (C1'), 84.5 (C3), 82.0 (C3'), 80.9 (C2), 79.2 (C2'), 78.4 (C4'), 75.6, 75.2, 73.8, 73.6, 73.1(C4), 73.0, 71.7, 71.5 (C5'), 69.1 (C6'), 61.9 (C6), 45.8 (C5). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{55}H_{58}O_9Na$ 885.3973, found 885.3971.

Compound 7b



Starting from compound **6b** (0.78 g, 0.85 mmol), and following **General procedure C**, the reaction was purified by flash column chromatography (pentane/EtOAc 7:1→3:1) to obtain compound **7b** (0.50 g, 0.62 mmol, 74%) as a colorless oil. ^{1}H NMR (500 MHz, $CDCl_3$) *δ* 7.45 – 7.02 (m, 25H, CH Ar), 5.81 (dt, *J* = 10.2, 2.4 Hz, 1H, H1), 5.75 (d, *J* = 4.0 Hz, 1H, H1'), 5.65 (dt, *J* = 10.2, 2.1 Hz, 1H, H7), 5.02 (d, *J* = 12.1 Hz, 1H, CH*H* Bn), 4.94 – 4.78 (m, 2H, CH*H* Bn), 4.73 (d, *J* = 10.9 Hz, 1H, CH*H* Bn), 4.63 – 4.58 (m, 3H, CH*H* Bn), 4.54 (d, *J* = 11.5 Hz, 1H, CH*H* Bn), 4.46 (s, 2H, CH*H* Bn), 4.26 (ddt, *J* = 7.3, 3.1, 2.0 Hz, 1H, H2), 4.16 (dd, *J* = 9.6, 8.7 Hz, 1H, H4), 4.02 – 3.81 (m, 4H, H6a, H3, H5' and H3'), 3.76 (dq, *J* = 9.2, 7.0 Hz, 1H, CH*H* Et), 3.67 (td, *J* = 11.1, 10.5, 2.3 Hz, 2H, H6'a and H6b), 3.50 (dd, *J* = 9.9, 6.9 Hz, 1H, H6'b), 3.45 – 3.36 (m, 2H, H2' and CH*H* Et), 3.10 (dd, *J* = 10.1, 8.7 Hz, 1H, H4'), 3.02 (t, *J* = 6.3 Hz, 1H, OH), 2.58 (dh, *J* = 8.8, 2.9 Hz, 1H, H5), 1.04 (t, *J* = 7.0 Hz, 3H, CH_3 Et). ^{13}C NMR (126 MHz, $CDCl_3$) *δ* 139.5, 138.8, 138.3, 138.2, 137.4 ($5C_q$ Ar), 129.8 (C7), 128.6, 128.5, 128.4, 128.3, 128.2, 128.2, 128.1, 128.0, 127.9, 127.7, 127.7, 127.6, 127.6, 127.5, 127.0 (C1), 126.6, 97.4 (C1'), 84.5 (C3), 81.8 (C3'), 80.8 (C2), 79.1 (C2'), 78.8 (C4'), 75.6, 73.9, 73.7 ($3CH_2$ Bn), 73.3 (C4), 73.1, 71.7 ($2CH_2$ Bn), 71.6 (C5'), 69.3 (C6'), 68.6 (CH_2 Et), 62.1 (C6), 45.8 (C5), 15.8 (CH_3 Et). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{50}H_{56}O_9Na$ 823.3817, found 823.3821.

Compound 7c

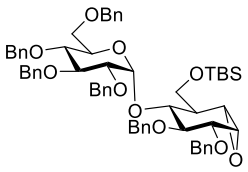


Starting from compound **6c** (1.10 g, 1.05 mmol) and following **General procedure C**, The reaction was purified by flash column chromatography (pentane/EtOAc 7:1→3:1) to obtain compound **7c** (0.70 g, 0.75 mmol, 71%) as a yellow oil. ^{1}H NMR (400 MHz, $CDCl_3$) *δ* 7.38 – 7.11 (m, 23H, CH Ar), 7.10 – 7.02 (m, 2H, CH Ar), 5.81 (dt, *J* = 10.2, 2.4 Hz, 1H, H1), 5.76 (d, *J* = 4.0 Hz, 1H, H1'), 5.65 (dt, *J* = 10.1, 2.1 Hz, 1H, H7), 5.03 (d, *J* = 12.1 Hz, 1H, CH*H* Bn), 4.92 – 4.81 (m, 2H, CH*H* Bn), 4.71 (d, *J* = 10.9 Hz, 1H, CH*H* Bn), 4.64 – 4.50 (m, 4H, CH*H* Bn), 4.46 (s, 2H, CH*H* Bn), 4.26 (ddt, *J* = 7.3, 3.4, 2.0 Hz, 1H, H2), 4.17 (t, *J* = 9.1 Hz, 1H, H4), 4.01 – 3.81 (m, 4H, H6a, H3, H5' and H3'), 3.69 (tdd, *J* = 10.4, 7.6, 2.6 Hz, 3H, CH*H*O linker, H6'a and H6b), 3.50 (dd, *J* = 9.9, 6.8 Hz, 1H, H6'b),

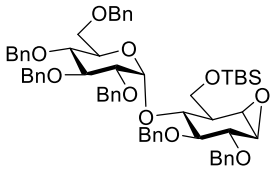
3.43 (dd, *J* = 9.8, 4.0 Hz, 1H, H2'), 3.31 (dt, *J* = 9.0, 6.8 Hz, 1H, CH*H*O linker), 3.21 (t, *J* = 6.9 Hz, 2H, CH_2N_3), 3.10 (dd, *J* = 10.1, 8.7 Hz, 1H, H4'), 2.58 (dh, *J* = 8.7, 2.8 Hz, 1H, H5), 1.54 (p, *J* = 7.0 Hz, 2H, 2CH*H* linker), 1.40 (dq, *J* = 13.4, 6.7 Hz, 2H, 2CH*H* linker), 1.33 – 1.05 (m, 8H, 8CH*H* linker). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.3, 138.7, 138.2, 138.1, 137.3 ($5C_q$ Ar), 129.7 (C7), 128.5, 128.4, 128.3, 128.2, 128.1, 128.1, 127.9, 127.9, 127.7, 127.6, 127.6, 127.5, 127.5, 127.4, 126.9 (C1), 126.4, 97.3 (C1'), 84.4 (C3), 81.7 (C3'), 80.7 (C2), 79.0 (C2'), 78.8 (C4'), 75.4, 73.8, 73.6 ($3CH_2$ Bn), 73.3 (CH_2O linker), 73.2 (C4), 72.9, 71.6 ($2CH_2$ Bn), 71.5 (C5'), 69.2 (C6'), 62.0 (C6), 51.4 (CH_2N_3), 45.7 (C5), 30.2, 29.3, 29.0, 28.8, 26.7, 26.0 ($6CH_2$ linker). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{56}H_{67}N_3O_9Na$ 948.4770, found 948.4769.

Compound 8a and 9a

Starting from compound **7a** (0.70 mg, 0.81 mmol) and following **General procedure D and E**, the product was purified by flash column chromatography (pentane/Et_2O 13:1→9:1) to obtain compound **8a** (0.16 g, 0.16 mmol, 20%) as a clean oil and compound **9a** (0.47 g, 0.47 mmol, 58%) as a clean oil.



8a: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.38 – 7.10 (m, 30H, CH Ar), 5.34 (d, *J* = 3.6 Hz, 1H, H1'), 4.96 (dd, *J* = 27.9, 11.2 Hz, 2H, CH*H* Bn), 4.89 – 4.79 (m, 3H, CH*H* Bn), 4.78 – 4.67 (m, 2H, CH*H* Bn), 4.59 (dd, *J* = 17.7, 12.2 Hz, 2H, CH*H* Bn), 4.51 – 4.43 (m, 3H, CH*H* Bn), 4.02 (t, *J* = 9.4 Hz, 1H, H3'), 3.95 (t, *J* = 9.1 Hz, 1H, H3), 3.88 (dt, *J* = 9.8, 3.0 Hz, 1H, H5'), 3.85 – 3.75 (m, 3H, H2, H4 and H6a), 3.75 – 3.70 (m, 2H, H6b and H6'a), 3.66 (dd, *J* = 10.0, 9.0 Hz, 1H, H4'), 3.58 (dd, *J* = 10.4, 2.0 Hz, 1H, H6'b), 3.48 (dd, *J* = 9.9, 3.6 Hz, 1H, H2'), 3.31 (d, *J* = 4.2 Hz, 1H, epoxide), 3.08 (d, *J* = 4.2 Hz, 1H, epoxide), 2.39 (dq, *J* = 8.4, 4.8, 4.1 Hz, 1H, H5), 0.87 (s, 9H, $(CH_3)_3CSi$), 0.02 (d, *J* = 5.2 Hz, 6H, $2CH_3Si$). ^{13}C NMR (126 MHz, $CDCl_3$) δ 139.6, 138.8, 138.5, 138.4, 138.4, 138.0 ($6C_q$ Ar), 128.5, 128.5, 128.3, 128.2, 128.1, 128.1, 127.9, 127.9, 127.8, 127.7, 127.6, 127.5, 127.0, 127.0, 98.7 (C1'), 81.9 (C3'), 80.9 (C3), 79.6 (C2'), 79.0 (C2), 78.6 (C4), 78.0 (C4'), 75.7, 75.1, 74.7, 73.7, 73.1, 72.9 ($6CH_2$ Bn), 71.1 (C5'), 68.6 (C6'), 63.0 (C6), 55.7 (epoxide), 55.2 (epoxide), 44.1 (C5), 26.0 ($(CH_3)_3CSi$), 18.2 ($(CH_3)_3CSi$), -5.2 (CH_3Si), -5.4 (CH_3Si). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{61}H_{72}O_{10}SiNa$ 1015.4787, found 1015.4792.

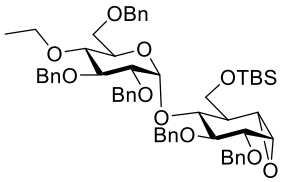


9a:^{1}H NMR (400 MHz, $CDCl_3$) δ 7.55 – 6.83 (m, 30H, CH Ar), 5.70 (d, *J* = 3.6 Hz, 1H, H1'), 4.93 – 4.67 (m, 6H, CH*H* Bn), 4.64 – 4.54 (m, 2H, CH*H* Bn), 4.49 – 4.39 (m, 4H, CH*H* Bn), 4.15 (dd, *J* = 9.1, 4.3 Hz, 1H, H6a), 3.94 (dd, *J* = 9.9, 8.7 Hz, 1H, H3'), 3.90 – 3.85 (m, 1H, H3), 3.76 (dt, *J* = 10.1, 2.6 Hz, 1H, H5'), 3.74 – 3.59 (m, 6H, H6'ab, H6b, H2, H4 and H4'), 3.53 (dt, *J* = 4.1, 1.0 Hz, 1H, epoxide), 3.45 (dd, *J* = 9.9, 3.7 Hz, 1H, H2'), 3.22 (d, *J* = 3.8 Hz, 1H, epoxide), 2.36 (dtd, *J* = 10.5, 5.4, 2.5 Hz, 1H, H5), 0.87 (s, 9H, $(CH_3)_3CSi$), 0.04 (d, *J* = 2.1 Hz, 6H, $2CH_3Si$). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.1, 138.8, 138.7, 138.2, 138.0, 137.6 ($6C_q$ Ar), 128.6, 128.5, 128.4, 128.3, 128.3, 128.1, 128.1, 128.0, 128.0, 127.8, 127.7, 127.6, 127.6, 127.5, 127.1, 126.5, 97.0 (C1'), 85.0, 81.9 (C3'), 80.3 (C3), 79.3 (C2'),

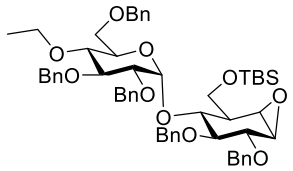
77.7 (C4'), 75.7, 74.8, 73.8, 73.7, 73.0, 72.9 (6CH_2 Bn), 71.5 (C5'), 70.0, 68.4 (C6'), 62.8 (C6), 55.0 (epoxide), 53.8 (epoxide), 43.8 (C5), 26.1 ((*C*H_3)$_3$CSi), 18.5 ((CH$_3$)$_3$*C*Si), -5.3 (CH_3Si), -5.4 (CH_3Si). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{61}H_{72}O_{10}SiNa$ 1015.4787, found 1015.4789.

Compound 8b and 9b

Starting from compound **7b** (447 mg, 0.558 mmol) and following **General procedure D and E**, the product was purified by flash column chromatography (pentane/Et_2O 11:1→5:1) to obtain compound **8b** (108 mg, 0.116 mmol, 21%) as a light yellow oil and compound **9b** (338 mg, 0.363 mmol, 65%) as a light yellow oil.



8b:^{1}H NMR (400 MHz, $CDCl_3$) δ 7.51 – 7.01 (m, 25H, CH Ar), 5.34 (d, *J* = 3.6 Hz, 1H, H1'), 5.02 (d, *J* = 11.6 Hz, 1H, CH*H* Bn), 4.94 – 4.81 (m, 3H, CH*H* Bn), 4.80 – 4.70 (m, 2H, CH*H* Bn), 4.67 (d, *J* = 12.0 Hz, 1H, CH*H* Bn), 4.60 (d, *J* = 12.3 Hz, 1H, CH*H* Bn), 4.49 (dd, *J* = 14.5, 12.1 Hz, 2H, CH*H* Bn), 4.01 – 3.91 (m, 2H), 3.88 – 3.65 (m, 7H), 3.61 (dd, *J* = 10.3, 2.1 Hz, 1H), 3.57 – 3.41 (m, 3H), 3.36 – 3.31 (m, 1H, H1), 3.10 (dd, *J* = 4.3, 0.8 Hz, 1H, H7), 2.43 (dt, *J* = 7.1, 3.4 Hz, 1H, H5), 1.11 (t, *J* = 7.0 Hz, 3H, CH_3 Et), 0.90 (s, 9H, (CH$_3$)$_3$CSi), 0.05 (d, *J* = 4.2 Hz, 6H, 2CH_3Si). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.6, 138.9, 138.4, 138.4, 138.0 (5C_q Ar), 128.5, 128.4, 128.3, 128.2, 128.1, 127.8, 127.8, 127.6, 127.6, 127.4, 127.0, 127.0, 98.7 (C1'), 81.8, 80.9, 79.4, 79.0, 78.5, 78.0, 75.7 (CH_2 Bn), 74.8 (CH_2 Bn), 73.7 (CH_2 Bn), 73.1 (CH_2 Bn), 72.9 (CH_2 Bn), 71.21 (C5'), 68.6 (C6'), 68.5 (CH_2 Et), 63.0 (C6), 55.7 (C7), 55.2 (C1), 44.1 (C5), 26.0 ((*C*H_3)$_3$CSi), 18.2 ((CH$_3$)$_3$*C*Si), 15.9 (CH_3 Et), -5.2 (CH_3Si), -5.4 (CH_3Si). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{56}H_{70}O_{10}SiNa$ 953.4631, found 953.4635.

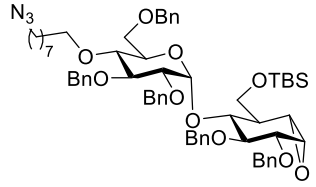


9b:^{1}H NMR (400 MHz, $CDCl_3$) δ 7.43 – 6.93 (m, 25H, CH Ar), 5.67 (d, *J* = 3.6 Hz, 1H, H1'), 4.98 – 4.36 (m, 10H, CH*H* Bn), 4.14 (dd, *J* = 9.2, 4.3 Hz, 1H, H6a), 3.90 – 3.82 (m, 2H), 3.80 – 3.56 (m, 7H), 3.52 (dt, *J* = 3.8, 1.2 Hz, 1H, epoxide), 3.49 – 3.35 (m, 3H), 3.21 (d, *J* = 3.8 Hz, 1H, epoxide), 2.35 (tdd, *J* = 9.4, 4.3, 1.5 Hz, 1H, H5), 1.03 (t, *J* = 7.0 Hz, 3H, CH_3 Et), 0.91 (s, 9H, (CH$_3$)$_3$CSi), 0.07 (d, *J* = 1.9 Hz, 6H, 2CH_3Si). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.1, 139.0, 138.3, 138.1, 137.6 (5C_q Ar), 128.6, 128.4, 128.4, 128.3, 128.2, 128.1, 128.0, 127.9, 127.8, 127.6, 127.6, 127.4, 127.1, 126.5, 97.0 (C1'), 85.0, 81.8, 80.3, 79.2, 77.7, 75.6 (CH_2 Bn), 73.8 (CH_2 Bn), 73.7 (CH_2 Bn), 73.1 (CH_2 Bn), 72.8 (CH_2 Bn), 71.6, 70.0, 68.4 (C6'), 68.2 (CH_2 Et), 62.8 (C6), 55.0 (epoxide), 53.7 (epoxide), 43.7 (C5), 26.1 ((*C*H_3)$_3$CSi), 18.5 ((CH$_3$)$_3$*C*Si), 15.9 (CH_3 Et), -5.3 (CH_3Si), -5.4 (CH_3Si). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{56}H_{70}O_{10}SiNa$ 953.4631, found 953.4632.

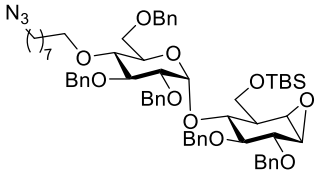
Compound 8c and 9c

Starting from compound **7c** (666 mg, 0.719 mmol) and following **General procedure D and E**, the product was purified by flash column chromatography (pentane/Et_2O 11:1→7:1) to obtain compound

8c (165 mg, 0.156 mmol, 22%) as a yellow oil and compound **9c** (478 mg, 0.453 mmol, 63%)as a yellow oil.

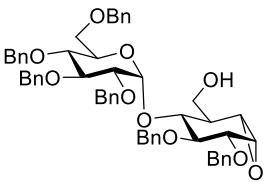


8c:^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.13 (m, 25H, CH Ar), 5.33 (d, *J* = 3.6 Hz, 1H, H1'), 5.12 – 4.34 (m, 10H, CH*H* Bn), 4.02 – 3.89 (m, 2H), 3.86 – 3.65 (m, 7H), 3.60 (dd, *J* = 10.3, 2.1 Hz, 1H), 3.50 – 3.37 (m, 3H), 3.36 – 3.31 (m, 1H, epoxide), 3.28 (t, *J* = 6.9 Hz, 2H, CH_2N_3), 3.11 (dd, *J* = 4.2, 0.9 Hz, 1H, epoxide), 2.43 (dt, *J* = 6.7, 3.4 Hz, 1H, H5), 1.69 – 1.55 (m, 2H, 2CH*H* linker), 1.48 (dt, *J* = 8.7, 2.8 Hz, 2H, 2CH*H* linker), 1.39 – 1.24 (m, 8H, 8CH*H* linker), 0.90 (s, 9H, $(CH_3)_3CSi$), 0.09 – 0.03 (m, 6H, $2CH_3Si$). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.6, 138.9, 138.4, 138.4, 138.0 ($5C_q$ Ar), 128.4, 128.4, 128.3, 128.2, 128.1, 128.0, 128.0, 127.8, 127.7, 127.6, 127.4, 127.0, 98.7 (C1'), 81.8, 80.9, 79.4, 79.0, 78.6, 78.0, 75.6 (CH_2 Bn), 74.8 (CH_2 Bn), 73.7 (CH_2 Bn), 73.2 (CH_2O linker), 73.1 (CH_2 Bn), 72.9 (CH_2 Bn), 71.2, 68.7 (C6'), 63.0 (C6), 55.7 (epoxide), 55.2 (epoxide), 51.6 (CH_2N_3), 44.1 (C5), 30.5, 29.5, 29.2, 28.9, 26.8, 26.2 ($6CH_2$ linker), 26.0 ((CH_3)$_3$CSi), 18.1 ((CH_3)$_3$*C*Si), -5.2 (CH_3Si), -5.4 (CH_3Si). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{62}H_{81}N_3O_{10}SiNa$ 1078.5583, found 1078.5581.



9c:^{1}H NMR (400 MHz, $CDCl_3$) δ 7.45 – 6.88 (m, 25H, CH Ar), 5.69 (d, *J* = 3.7 Hz, 1H, H1'), 5.00 – 4.28 (m, 10H, CH*H* Bn), 4.15 (dd, *J* = 9.1, 4.3 Hz, 1H, H6a), 3.90 – 3.80 (m, 2H), 3.78 – 3.56 (m, 7H), 3.56 – 3.50 (m, 1H, epoxide), 3.46 – 3.29 (m, 3H), 3.25 – 3.17 (m, 3H, 1H epoxide and CH_2N_3), 2.36 (tdd, *J* = 9.4, 4.4, 2.4 Hz, 1H, H5), 1.55 (dq, *J* = 8.3, 6.8 Hz, 2H, 2CH*H* linker), 1.46 – 1.33 (m, 2H, 2CH*H* linker), 1.32 – 1.08 (m, 8H, 8CH*H* linker), 0.91 (s, 9H, $(CH_3)_3CSi$), 0.08 (d, *J* = 2.3 Hz, 6H, $2CH_3Si$). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.1, 138.9, 138.2, 138.0, 137.6 ($5C_q$ Ar), 128.5, 128.4, 128.3, 128.2, 128.2, 128.0, 128.0, 127.8, 127.7, 127.6, 127.5, 127.4, 127.0, 126.4, 96.9 (C1'), 85.1, 81.8, 80.2, 79.1, 77.7, 75.5 (CH_2 Bn), 73.7 (CH_2 Bn), 73.6 (CH_2 Bn), 73.0 (CH_2 Bn), 72.9 (CH_2O linker), 72.7 (CH_2 Bn), 71.6, 69.7, 68.4 (C6'), 62.8 (C6), 54.9 (epoxide), 53.7 (epoxide), 51.5 (CH_2N_3), 43.7 (C5), 30.5, 29.5, 29.2, 28.9, 26.8, 26.2 ($6CH_2$ linker), 26.1 ((CH_3)$_3$CSi), 18.4 ((CH_3)$_3$*C*Si), -5.3 (CH_3Si), -5.4 (CH_3Si). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{62}H_{81}N_3O_{10}SiNa$ 1078.5583, found 1078.5581.

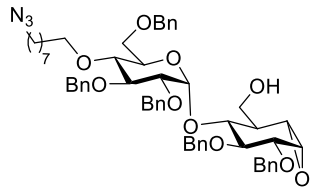
Compound 10a



Starting from compound **8a** (0.10 g, 0.10 mmol) and following **General procedure F**, the product was purified by flash column chromatography (pentane/EtOAc 4:1→2:1) to obtain compound **10a** (80 mg, 0.09 mmol, 90%) as a clean oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.43 – 6.99 (m, 30H, CH Ar), 5.58 (d, *J* = 4.0 Hz, 1H, H1'), 4.97 – 4.83 (m, 3H, CH*H* Bn), 4.76 – 4.63 (m, 4H, CH*H* Bn), 4.58 (d, *J* = 12.3 Hz, 1H, CH*H* Bn), 4.55 – 4.45 (m, 2H, CH*H* Bn), 4.39 (dd, *J* = 11.5, 7.7 Hz, 2H, CH*H* Bn), 4.01 – 3.75 (m, 7H, H3', H5', H6ab, H2, H3 and H4), 3.66 (dd, *J* = 9.9, 1.7 Hz, 1H, H6'a), 3.47 – 3.31 (m,

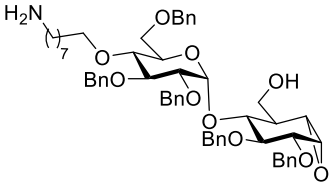
3H, H2', H6'b and epoxide), 3.29 – 3.08 (m, 2H, H4' and epoxide), 2.34 – 2.25 (m, 1H, H5). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.5, 138.5, 138.2, 138.1, 137.7, 137.1 (6C_q Ar), 128.7, 128.6, 128.5, 128.4, 128.3, 128.3, 128.2, 128.1, 128.1, 128.0, 127.8, 127.7, 127.5, 127.0, 126.5, 97.8 (C1'), 81.8, 81.4, 79.8, 79.0 (C2'), 78.5 (C4'), 75.6 (CH_2 Bn), 75.1 (CH_2 Bn), 74.6 (CH_2 Bn), 74.5, 73.7 (CH_2 Bn), 73.2 (CH_2 Bn), 72.8 (CH_2 Bn), 71.5 (C5'), 69.5 (C6'), 61.6 (C6), 55.5 (epoxide), 54.9 (epoxide), 44.2 (C5). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{55}H_{58}O_{10}Na$ 901.3922, found 901.3920.

Compound 10b



Starting from compound **8c** (0.15 g, 0.14 mmol) and following **General procedure F**, the product was purified by flash column chromatography (pentane/EtOAc 3:1→1:1) to obtain compound **10c** (117 mg, 0.124 mmol, 86%) as a colorless oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.41 – 7.00 (m, 25H, CH Ar), 5.53 (d, *J* = 3.9 Hz, 1H, H1'), 4.91 (s, 2H, CH*H* Bn), 4.82 (d, *J* = 10.9 Hz, 1H, CH*H* Bn), 4.71 (d, *J* = 7.7 Hz, 3H, CH*H* Bn), 4.64 – 4.53 (m, 2H, CH*H* Bn), 4.50 (d, *J* = 12.1 Hz, 1H, CH*H* Bn), 4.39 (d, *J* = 12.1 Hz, 1H, CH*H* Bn), 3.97 – 3.76 (m, 7H, H3', H5', H6ab, H2, H3 and H4), 3.76 – 3.66 (m, 2H, H6'a and CH*H*O linker), 3.49 (dd, *J* = 9.8, 7.6 Hz, 1H, H6'b), 3.43 – 3.26 (m, 3H, H2', epoxide and CH*H*O linker), 3.26 – 3.19 (m, 3H, epoxide and CH_2N_3), 3.03 (dd, *J* = 10.1, 8.5 Hz, 1H, H4'), 2.31 (ddd, *J* = 7.7, 4.4, 2.6 Hz, 1H, H5), 1.56 (p, *J* = 7.0 Hz, 2H, 2CH*H* linker), 1.41 (p, *J* = 6.9, 6.5 Hz, 2H, 2CH*H* linker), 1.36 – 1.11 (m, 8H, 8CH*H* linker). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.5, 138.6, 138.2, 138.1, 137.2 (5C_q Ar), 128.6, 128.4, 128.4, 128.2, 128.2, 128.2, 128.1, 128.1, 127.8, 127.8, 127.6, 127.5, 127.4, 126.9, 126.5, 97.9 (C1'), 81.5, 81.4, 79.7, 79.0 (C4'), 78.9 (C2'), 75.4 (CH_2 Bn), 74.9, 74.6 (CH_2 Bn), 73.7 (CH_2 Bn), 73.3 (CH_2O linker), 73.2 (CH_2 Bn), 72.7 (CH_2 Bn), 71.7, 69.7 (C6'), 61.6 (C6), 55.4 (epoxide), 54.8 (epoxide), 51.5 (CH_2N_3), 44.2 (C5), 30.3, 29.4, 29.1, 28.9, 26.7, 26.0 (6CH_2 linker). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{56}H_{67}N_3O_{10}Na$ 964.4719, found 964.4717.

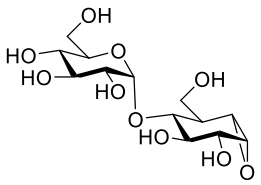
Compound 11



Compound **10b** (110 mg, 0.117 mmol) was dissolved in MeCN (2.4 mL), polymer-bound triphenylphosphine (3 mmol/g loading, 973 mg, 0.292 mmol) and H_2O (21.1 µL, 1.17 mmol) were added and the mixture was stirred overnight at 70 °C. TLC-analysis indicated total consumption of the starting material, and additional H_2O (500 µL) was added and the mixture was stirred for 5 h at 70 °C. The mixture was cooled to rt, filtered, washed with MeCN (3 x 5 mL). The combined filtrate was diluted with MeCN (50 mL), dried over Na_2SO_4, filtered and concentrated *in vacuo*. The crude was purified by flash column chromatography (DCM/MeOH 19:1→15:1) affording compound **11** (84 mg, 92 µmol, 79%) as a clean oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37 – 6.99 (m, 25H, CH Ar), 5.55 (d, *J* = 3.9 Hz, 1H, H1'), 4.90 (s, 2H, CH*H* Bn), 4.82 (d, *J* = 11.0 Hz, 1H, CH*H* Bn), 4.69 (q, *J* = 6.3, 5.6 Hz, 3H, CH*H*

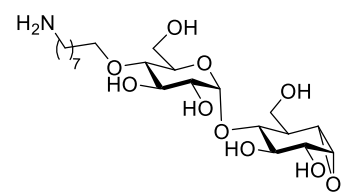
Bn), 4.57 (s, 2H, CH*H* Bn), 4.49 (d, *J* = 12.2 Hz, 1H, CH*H* Bn), 4.39 (d, *J* = 12.1 Hz, 1H, CH*H* Bn), 3.85 (ddt, *J* = 22.1, 12.8, 9.4 Hz, 7H, H3', H5', H6ab, H2, H3 and H4), 3.75 – 3.65 (m, 2H, H6'a and CH*H*O linker), 3.54 – 3.43 (m, 1H, H6'b), 3.40 – 3.24 (m, 3H, H2', epoxide and CH*H*O linker), 3.22 (d, *J* = 4.1 Hz, 1H, epoxide), 2.86 (t, *J* = 7.7 Hz, 2H, C*H*$_2$NH$_2$), 2.31 (dt, *J* = 7.7, 3.4 Hz, 1H, H5), 1.66 (dq, *J* = 15.4, 7.7, 7.0 Hz, 2H, 2CH*H* linker), 1.45 – 1.35 (m, 2H, 2CH*H* linker), 1.36 – 1.10 (m, 8H, 8CH*H* linker). ^{13}C NMR (101 MHz, $CDCl_3$) *δ* 139.5, 138.7, 138.2, 138.1, 137.3 (5C$_q$ Ar), 128.6, 128.5, 128.4, 128.4, 128.2, 128.2, 128.2, 128.1, 127.9, 127.8, 127.7, 127.5, 127.5, 127.0, 126.5, 97.9 (C1'), 81.6, 81.5, 79.8, 79.0 (C4'), 78.9 (C2'), 75.4, 74.8, 74.6, 73.8, 73.4 (CH_2O linker), 73.2, 72.8, 71.7, 69.7 (C6'), 61.7 (C6), 55.4 (epoxide), 54.9 (epoxide), 44.3 (C5), 40.4 (CH_2NH_2), 30.3, 29.4, 29.0, 28.8, 26.6, 26.1 (6CH_2 linker). HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{56}H_{70}NO_{10}$ 916.4994, found 916.4993.

Compound 1a



Compound **10a** (21 mg, 24 μmol) was dissolved in a mixture of MeOH/H_2O/dioxane (2/1/2, 1 mL) under Argon and $Pd(OH)_2$/C (20 wt%, 25 mg, 36 μmol) was added. While stirring vigorously, the mixture was flushed with a H_2 balloon. After stirring for 3 h under H_2 atmosphere, the mixture was filtered over a small celite pad and evaporate to afford the product in high purity as a white powder (7.8 mg, quant) after lyophilization. ^{1}H NMR (400 MHz, MeOD) *δ* 5.02 (d, *J* = 3.9 Hz, 1H, H1'), 3.90 (dd, *J* = 11.1, 3.0 Hz, 1H, H6a), 3.85 – 3.76 (m, 3H, H6b, H6'a and H4'), 3.74 – 3.66 (m, 2H, H6'b and H2), 3.66 – 3.54 (m, 2H, H3 and H3'), 3.44 (dd, *J* = 9.7, 3.9 Hz, 1H, H2'), 3.41 – 3.33 (m, 1H, H4), 3.30 – 3.25 (m, 2H, epoxide and H5'), 3.21 (d, *J* = 4.0 Hz, 1H, epoxide), 2.05 (ddd, *J* = 9.0, 5.6, 3.0 Hz, 1H, H5). ^{13}C NMR (101 MHz, MeOD) *δ* 101.1 (C1'), 80.8 (C4), 72.9 (C3'), 72.7 (C3), 72.3 (C2), 71.8 (C2'), 70.5 (C4'), 69.3 (C5'), 60.5 (C6'), 59.5 (C6), 55.8 (epoxide), 53.4 (epoxide), 43.1 (C5). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{13}H_{22}O_{10}Na$ 361.1105, found 361.1102.

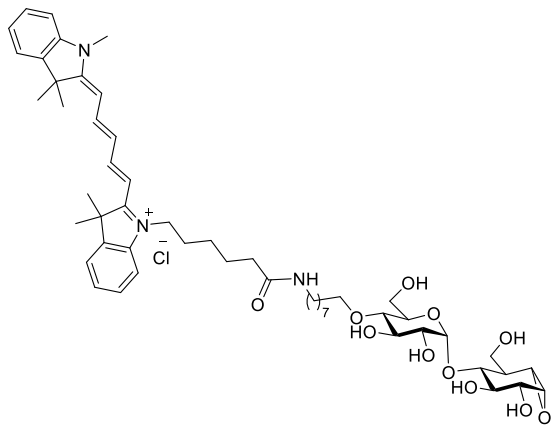
Compound 1b



Compound **11** (30 mg, 33 μmol) was dissolved in a mixture of ^{t}BuOH/H_2O/dioxane (1/2/1, 1.6 mL) under Argon, then HOAc (19 μL, 0.33 mmol) and $Pd(OH)_2$/C (20 wt%, 35 mg, 50 μmol) were added. While stirring vigorously, the mixture was flushed with a H_2 balloon. After stirring for 6.5 h under H_2 atmosphere, the mixture was filtered over a small celite pad. The filtrate was concentrated and purified by semi-preparative reversed phase HPLC (linear gradient. Solution used: A: 50 mM NH_4HCO_3 in H_2O, B: MeCN). The product was obtained as a white powder (13.9 mg, 29.9 μmol, 91%) after lyophilization. ^{1}H NMR (850 MHz, D_2O) *δ* 5.11 (d, *J* = 4.0 Hz, 1H, H1'), 3.92 (dd, *J* = 8.9, 1.8 Hz, 1H, H2), 3.87 (dd, *J* = 11.4, 2.9 Hz, 1H, H6a), 3.82 – 3.77 (m, 2H, CH*H*O linker and H6'a), 3.76 – 3.67 (m, 4H, H3', H5', H6b and H6'b), 3.65 – 3.59 (m, 2H, H3 and CH*H*O linker), 3.54 (ddd, *J* = 9.9, 4.5 Hz, 1H, H2'), 3.46 – 3.42 (m, 2H, H4 and epoxide),

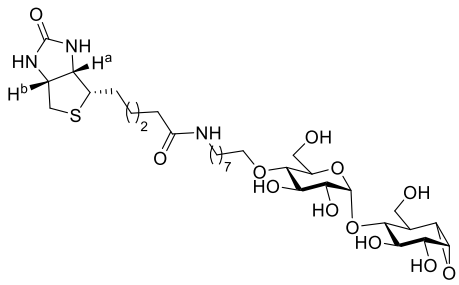
3.32 (d, *J* = 4.2 Hz, 1H, epoxide), 3.24 (t, *J* = 9.6 Hz, 1H, H4'), 2.94 (t, *J* = 7.6 Hz, 2H, C*H*$_2$NH$_2$), 2.19 (ddd, *J* = 9.0, 5.7, 2.9 Hz, 1H, H5), 1.64 – 1.58 (m, 2H, 2CH*H* linker), 1.58 – 1.51 (m, 2H, 2CH*H* linker), 1.35 – 1.26 (m, 8H, 8CH*H* linker). ^{13}C NMR (214 MHz, D$_2$O) *δ* 100.6 (C1'), 79.9 (C4), 77.6 (C4'), 73.4 (C3), 73.4 (CH$_2$O linker), 72.9 (C3'), 71.8 (C2'), 71.6 (C5'), 70.6 (C2), 60.3 (C6), 60.2 (C6'), 57.3 (epoxide), 55.2 (epoxide), 42.8 (C5), 39.5 (CH$_2$NH$_2$), 29.2, 28.2, 28.1, 26.7, 25.5, 25.1 (6CH$_2$ linker). HRMS (ESI) m/z: [M+H]$^+$ calc for $C_{21}H_{40}NO_{10}$ 466.2647, found 466.2644.

Compound 1c



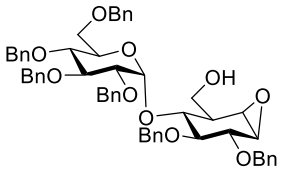
Compound **1b** (5.4 mg, 12 μmol) was dissolved in dry DMF (0.5 mL), then DIPEA (6.0 μL, 34 μmol) and Cy5-OSu (7.9 mg, 13 μmol) were added and the mixture was stirred at rt for 40 h. Full conversion was observed by LC-MS and the product was purified by semi-preparative reversed phase HPLC (linear gradient. Solution used: A: 50 mM NH_4HCO_3 in H_2O, B: MeCN). The fractions were concentrated under reduced pressure, co-evaporated with Milli-Q/MeCN (1/1, 3 x), dissolved in Milli-Q/^{t}BuOH (1/1) again and lyophilized to obtain the title compound (3.7 mg, 3.9 μmol, 33%) as a blue powder. ^{1}H NMR (500 MHz, MeOD) *δ* 8.30 – 8.19 (m, 2H), 7.49 (d, *J* = 7.4 Hz, 2H), 7.41 (tdd, *J* = 7.7, 3.8, 1.2 Hz, 2H), 7.33 – 7.23 (m, 4H), 6.63 (t, *J* = 12.4 Hz, 1H), 6.28 (dd, *J* = 13.7, 5.9 Hz, 2H), 5.00 (d, *J* = 3.9 Hz, 1H, H1'), 4.11 (t, *J* = 7.4 Hz, 2H, CH$_2$N$^+$), 3.92 – 3.83 (m, 2H, H6a and CH*H*O linker), 3.82 – 3.72 (m, 3H, H2, H6b and H6'a), 3.72 – 3.65 (m, 3H, H3', H5' and H6'b), 3.63 (s, 3H, CH$_3$N), 3.62 – 3.52 (m, 2H, H3 and CH*H*O linker), 3.43 (dd, *J* = 9.7, 3.9 Hz, 1H, H2'), 3.36 – 3.33 (m, 1H, H4), 3.27 (dd, *J* = 4.0, 1.8 Hz, 1H, epoxide), 3.22 – 3.18 (m, 1H, epoxide), 3.19 – 3.07 (m, 3H, H4' and C*H*$_2$NH), 2.20 (t, *J* = 7.3 Hz, 2H, CH$_2$C=O), 2.04 (ddt, *J* = 11.5, 6.0, 2.6 Hz, 1H, H5), 1.83 (dt, *J* = 15.0, 7.4 Hz, 2H, 2CH*H* linker), 1.73 (s, 14H, 4CH$_3$ and 2CH*H* linker), 1.63 – 1.50 (m, 2H, 2CH*H* linker), 1.50 – 1.41 (m, 4H, 4CH*H* linker), 1.32 (dq, *J* = 5.6, 2.9, 2.3 Hz, 8H, 8CH*H* linker). ^{13}C NMR (126 MHz, MeOD) *δ* 175.7, 175.4, 174.7, 155.5, 155.5, 144.3, 143.6, 142.7, 142.5, 129.7, 129.7, 126.7, 126.2, 126.2, 123.4, 123.3, 112.1, 111.8, 104.4, 104.3, 103.2 (C1'), 83.1 (C4), 79.2 (C4'), 75.3 (C3'), 74.9 (C3), 74.1 (C2'), 74.0 (CH$_2$O linker), 73.7 (C5'), 72.7 (C2), 62.3 (C6'), 61.6 (C6), 58.0 (epoxide), 55.5 (epoxide), 50.6 (C$_q$), 50.5 (C$_q$), 45.2 (C5), 44.8 (CH$_2$N$^+$), 40.4 (CH$_2$NH), 36.7 (CH$_2$C=O), 31.5 (CH$_3$N), 31.4, 30.5, 30.4, 30.3, 28.2 (5CH$_2$ linker), 28.0 (2CH$_3$), 27.9 (CH$_2$ linker), 27.8 (2CH$_3$), 27.4, 27.2, 26.6 (3CH$_2$ linker). HRMS (ESI) m/z: [M]$^+$ calc for $C_{53}H_{76}N_3O_{11}$ 930.5474, found 930.5475.

Compound 1d



Compound **1b** (5.7 mg, 12 μmol) was dissolved in dry DMF (0.5 mL), then DIPEA (4.3 μL, 24 μmol) and biotin-OSu (4.6 mg, 13 μmol) were added and the mixture was stirred at rt for 22 h. Full conversion was observed by LC-MS and the product was purified by semi-preparative reversed phase HPLC (linear gradient. Solution used: A: 50 mM NH_4HCO_3 in H_2O, B: MeCN). The fractions were concentrated under reduced pressure, co-evaporated with Milli-Q/MeCN (1/1, 3 x), dissolved in Milli-Q/^{t}BuOH (1/1) again and lyophilized to obtain the title compound (1.4 mg, 2.0 μmol, 16%) as a white solid. ^{1}H NMR (850 MHz, MeOD) δ 5.00 (d, J = 3.9 Hz, 1H, H1'), 4.51 – 4.48 (m, 1H, H^b), 4.31 (dd, J = 7.9, 4.5 Hz, 1H, H^a), 3.90 – 3.85 (m, 2H, H6a and CH*H*O linker), 3.82 – 3.75 (m, 3H, H2, H6b and H6'a), 3.72 – 3.65 (m, 3H, H3', H5' and H6'b), 3.61 – 3.55 (m, 2H, H3 and CH*H*O linker), 3.43 (dd, J = 9.7, 3.9 Hz, 1H, H2'), 3.33 (d, J = 9.6 Hz, 1H, H4), 3.27 (dd, J = 4.1, 1.8 Hz, 1H, epoxide), 3.23 – 3.19 (m, 2H, epoxide and CHS), 3.18 – 3.13 (m, 3H, H4' and CH_2NH), 2.93 (dd, J = 12.8, 5.0 Hz, 1H, CH*H*S), 2.71 (d, J = 12.7 Hz, 1H, CH*H*S), 2.23 – 2.14 (m, 2H, CH_2C=O), 2.05 (ddd, J = 9.2, 5.8, 3.0 Hz, 1H, H5), 1.78 – 1.53 (m, 6H, 6CH*H* linker), 1.50 (t, J = 7.0 Hz, 2H, 2CH*H* linker), 1.45 (q, J = 7.6 Hz, 2H, 2CH*H* linker), 1.41 – 1.30 (m, 8H, 8CH*H* linker). ^{13}C NMR (214 MHz, MeOD) δ 176.0, 166.1, 103.2 (C1'), 83.0 (C4), 79.3 (C4'), 75.3 (C3'), 74.9 (C3), 74.1 (C2'), 74.0 (CH_2O linker), 73.7 (C5'), 72.7 (C2), 63.4 (CH^a), 62.3 (C6'), 61.7 (C6), 61.6 (CH^b), 57.9 (epoxide), 57.0 (CHS), 55.5 (epoxide), 45.2 (C5), 41.1 (CH_2S), 40.4 (CH_2NH), 36.9 (CH_2C=O), 31.4, 30.5, 30.4, 30.3, 29.8, 29.5, 27.9, 27.1, 27.0 (9CH_2 linker). HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{31}H_{54}N_3O_{12}S$ 692.3422, found 692.3412.

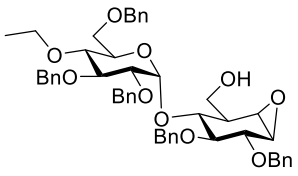
Compound 12a



Starting from compound **9a** (322 mg, 0.324 mmol) and following **General procedure F**, the product was purified by flash column chromatography (pentane/EtOAc 4:1→2:1) to obtain compound **12a** (0.26 g, 0.29 mmol, 91%) as a clean oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.47 – 6.91 (m, 30H, CH Ar), 5.66 (d, J = 3.9 Hz, 1H, H1'), 4.87 (t, J = 11.5 Hz, 2H, CH*H* Bn), 4.75 (ddd, J = 26.1, 19.7, 11.7 Hz, 4H, CH*H* Bn), 4.61 – 4.55 (m, 2H, CH*H* Bn), 4.55 – 4.38 (m, 4H, CH*H* Bn), 4.03 (dd, J = 11.3, 5.0 Hz, 1H, H6a), 4.01 – 3.85 (m, 5H, H6b, H3, H3' and H5'), 3.73 – 3.63 (m, 2H, H6'a), 3.58 (dd, J = 10.3, 5.0 Hz, 1H, H6'b), 3.50 – 3.41 (m, 2H, H2'), 3.39 (dd, J = 4.1, 1.7 Hz, 1H, epoxide), 3.16 (d, J = 3.8 Hz, 1H, epoxide), 2.28 – 2.20 (m, 1H, H5). ^{13}C NMR (126 MHz, $CDCl_3$) δ 139.2, 138.8, 138.2, 138.2, 137.8, 137.6 (6C_q Ar), 128.6, 128.6, 128.5, 128.5, 128.3, 128.3, 128.3, 128.2, 128.1, 128.1, 128.0, 127.9, 127.7, 127.7, 127.5, 127.1, 126.5, 97.4 (C1'), 84.8, 81.9, 80.2 (C3'), 79.3 (C2'), 78.1, 75.5 (CH_2 Bn), 75.1 (CH_2 Bn), 73.8 (CH_2 Bn), 73.7 (CH_2 Bn), 73.2 (CH_2 Bn), 72.9 (CH_2 Bn), 71.5 (C5'), 70.3, 68.9 (C6'), 61.9

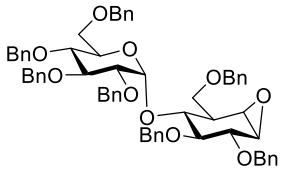
(C6), 56.5 (epoxide), 52.5 (epoxide), 43.3 (C5). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{55}H_{58}O_{10}Na$ 901.3922, found 901.3921.

Compound 12b



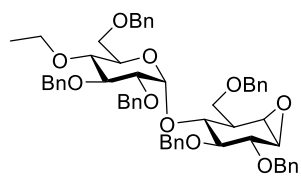
Starting from compound **9b** (338 mg, 0.36 mmol) and following **General procedure F**, the product was purified by flash column chromatography (pentane/EtOAc 3:1→2:1) to obtain compound **12b** (261 mg, 0.32 mmol, 88%) as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.42 – 6.99 (m, 25H, CH Ar), 5.66 (d, *J* = 3.9 Hz, 1H, H1'), 4.95 – 4.35 (m, 10H, CH*H* Bn), 4.08 – 3.91 (m, 3H), 3.91 – 3.64 (m, 6H), 3.59 (dd, *J* = 10.3, 4.9 Hz, 1H, H6'b), 3.46 – 3.37 (m, 3H, H2', CH*H* Et and epoxide), 3.26 (dd, *J* = 10.1, 8.8 Hz, 1H), 3.17 (d, *J* = 3.8 Hz, 1H, epoxide), 2.31 – 2.20 (m, 1H, H5), 1.04 (t, *J* = 7.0 Hz, 3H, CH_3 Et). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.2, 138.8, 138.2, 137.8, 137.5 (5C_q Ar), 128.6, 128.5, 128.4, 128.3, 128.2, 128.2, 128.1, 128.1, 127.9, 127.9, 127.7, 127.6, 127.5, 127.1, 126.5 (25CH Ar), 97.4 (C1'), 84.9, 81.8, 80.1, 79.1, 78.3, 75.5, 73.8, 73.7, 73.2, 72.9, 71.6, 70.0, 68.8 (C6'), 68.5 (CH_2 Et), 62.0 (C6), 56.5 (epoxide), 52.5 (epoxide), 43.3 (C5), 15.8 (CH_3 Et). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{50}H_{56}O_{10}Na$ 839.3766, found 839.3771.

Compound 13a



Starting from compound **12a** (0.26 g, 0.30 mmol) and following **General procedure G**, the product was purified by flash column chromatography (pentane/EtOAc 11:1→7:1) to obtain compound **13a** (234 mg, 0.24 mmol, 81%) as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37 – 6.96 (m, 35H, CH Ar), 5.71 (d, *J* = 3.7 Hz, 1H, H1'), 4.88 (dd, *J* = 13.8, 11.4 Hz, 2H, CH*H* Bn), 4.82 – 4.66 (m, 4H, CH*H* Bn), 4.61 – 4.31 (m, 8H, CH*H* Bn), 3.94 – 3.79 (m, 4H), 3.73 – 3.58 (m, 4H), 3.55 (dd, *J* = 10.6, 2.8 Hz, 1H, H6'a), 3.51 – 3.40 (m, 3H, H6'b, H2' and epoxide), 3.19 (d, *J* = 3.8 Hz, 1H, epoxide), 2.43 (dddd, *J* = 9.5, 7.7, 3.8, 1.6 Hz, 1H, H5). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.0, 138.87, 138.6, 138.3, 138.1, 138.0, 137.5 (7C_q Ar), 128.6, 128.5, 128.4, 128.4, 128.3, 128.3, 128.2, 128.1, 128.1, 127.9, 127.9, 127.8, 127.7, 127.7, 127.6, 127.5, 127.5, 127.1, 126.5, 96.9 (C1'), 85.2, 82.0, 80.2 (C3'), 79.3 (C2'), 77.7, 75.6, 75.1, 73.8, 73.6, 73.1, 73.0, 72.8, 71.1, 69.5, 69.1 (C6), 68.2 (C6'), 55.7 (epoxide), 53.2 (epoxide), 41.7 (C5). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{62}H_{64}O_{10}Na$ 991.4392, found 991.4394.

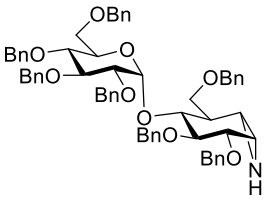
Compound 13b



Starting from compound **12b** (255 mg, 0.312 mmol) and following **General procedure G**, the product was purified by flash column chromatography (pentane/EtOAc 11:1→7:1) to obtain compound **13b** (213 mg, 0.235 mmol, 75%) as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.41 – 6.98 (m, 30H, CH Ar), 5.67 (d, *J* = 3.7 Hz, 1H, H1'), 4.92 – 4.66 (m, 5H, CH*H* Bn), 4.66 – 4.26 (m, 7H, CH*H* Bn), 3.92 – 3.58 (m, 8H), 3.54 – 3.36 (m, 6H), 3.19 (d, *J* = 3.8 Hz, 1H, epoxide), 2.44 (dddd,

J = 9.6, 7.9, 3.8, 1.6 Hz, 1H, H5), 1.06 (t, *J* = 7.0 Hz, 3H, CH_3 Et). ^{13}C NMR (101 MHz, $CDCl_3$) *δ* 139.1, 139.0, 138.3, 138.2, 138.0, 137.6 ($7C_q$ Ar), 128.6, 128.5, 128.4, 128.4, 128.3, 128.2, 128.2, 128.1, 128.0, 127.9, 127.7, 127.7, 127.7, 127.6, 127.5, 127.5, 127.1, 126.5, 97.0 (C1'), 85.1, 81.9, 80.2, 79.2, 77.8, 75.5, 73.7, 73.6, 73.2, 73.1, 72.8, 71.3, 69.8, 69.2 (C6), 68.4 (CH_2 Et), 68.2 (C6'), 55.6 (epoxide), 53.3 (epoxide), 41.7 (C5), 16.0 (CH_3 Et). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{57}H_{62}O_{10}Na$ 929.4235, found 929.4234.

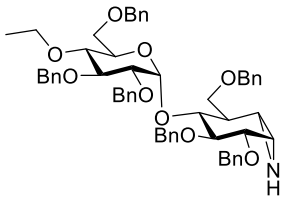
Compound 14a



Compound **13a** (234 mg, 0.241 mmol) was dissolved in dry DMF (2.4 mL) and purged with N_2. Then NaN_3 (158 mg, 2.41 mmol) and $LiClO_4$ (514 mg, 4.82 mmol) were added and the reaction was stirred at 100 ^{0}C for 18 h under inert atmosphere. The reaction mixture was diluted with H_2O (50 mL) and extracted with EtOAc (2 x 40 mL), the combined organic layers were washed with water (2 x 50 mL) and brine, dried with anhydrous Na_2SO_4, filtered and concentrated *in vacuo* to afford a mixture of azido-alcohols (0.21 g). The crude (0.21 mmol) was dissolved in dry MeCN (2 mL) and polymer-bound PPh_3 (1.6 mmol/g loading, 0.26 g, 0.42 mmol) was added to the solution. The reaction was stirred at 60 ^{0}C for 16 h under inert atmosphere. Then the beads were removed by filtration, the organic fitrate was concentrated *in vacuo* and purified by silica gel column chromatography (pentane/acetone 7:1→4:1) to obtain **14a** (96 mg, 99 µmol, 48%) as a clean oil. ^{1}H NMR (400 MHz, $CDCl_3$) *δ* 7.46 – 6.97 (m, 35H, CH Ar), 5.48 (s, 1H, H1'), 4.94 – 4.73 (m, 5H, CH*H* Bn), 4.69 (s, 2H, CH*H* Bn), 4.60 – 4.31 (m, 7H, CH*H* Bn), 3.97 (t, *J* = 9.3 Hz, 1H, H3'), 3.93 – 3.72 (m, 3H, H3, H2 and H6a), 3.61 (dddd, *J* = 14.3, 10.2, 6.3, 3.3 Hz, 5H, H6b, H4, H4', H5' and H6'a), 3.49 – 3.41 (m, 2H, H6'b and H2'), 2.49 (dd, *J* = 6.3, 2.8 Hz, 1H, aziridine), 2.37 (d, *J* = 7.2 Hz, 2H, H5), 2.30 (d, *J* = 6.2 Hz, 1H, aziridine). ^{13}C NMR (101 MHz, $CDCl_3$) *δ* 138.9, 138.6, 138.4, 138.3, 138.0 ($7C_q$ Ar), 128.8, 128.5, 128.5, 128.5, 128.4, 128.4, 128.3, 128.25, 128.2, 128.2, 128.1, 128.0, 128.0, 127.9, 127.8, 127.8, 127.7, 127.7, 127.6, 127.6, 127.5, 127.4, 127.1, 127.0, 126.8 (35CH Ar), 82.0, 82.0, 79.5, 79.5, 77.8, 77.8, 75.7 (CH_2 Bn), 75.1, 73.7, 73.6, 73.0, 72.8 ($5CH_2$ Bn), 72.2 (CH_2 Bn, assigned by HSQC), 70.9, 70.7 (C6), 68.4 (C6'), 43.4 (C5, assigned by HSQC), 33.8 (CH aziridine, assigned by HSQC), 33.1 (CH aziridine, assigned by HSQC). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{62}H_{65}NO_9Na$ 990.4552, found 990.4548.

Compound 14b

compound **13b** (0.21 g, 0.23 mmol) was dissolved in dry DMF (2.4 mL) and purged with N_2. Then NaN_3 (0.15 g, 2.3 mmol) and $LiClO_4$ (0.49 g, 4.6 mmol) were added and the reaction was stirred at 100 ^{0}C for 18 h under inert atmosphere. After cooling to rt, the reaction mixture was diluted with H_2O (50 mL) and extracted with EtOAc (2 x 40 mL). The combined organic layers were washed with water (2 x 50 mL) and brine, dried with anhydrous Na_2SO_4, filtered and concentrated *in vacuo* to afford a mixture of azido-

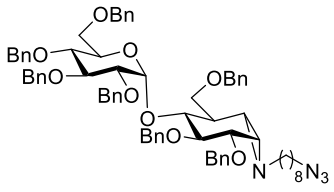


alcohols (204 mg). Starting from the azido-alcohol mixture (0.21 mmol), the reaction was carried out following the same procedure described for **14a**. The product was purified by flash column chromatography (pentane/EtOAc 7:1→1:1) to obtain **14b** (81 mg, 89 µmol, 42%) as a clean oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38 – 7.02 (m, 30H, CH Ar), 5.48 – 5.36 (m, 1H, H1'), 4.90 – 4.35 (m, 12H, CH*H* Bn), 3.94 – 3.68 (m, 5H), 3.67 – 3.54 (m, 3H), 3.50 – 3.35 (m, 5H), 2.48 (dd, *J* = 6.3, 2.6 Hz, 1H, aziridine), 2.38 (q, *J* = 6.0, 5.5 Hz, 1H, H5), 2.30 (dd, *J* = 6.1, 1.1 Hz, 1H, aziridine), 1.05 (t, *J* = 7.0 Hz, 3H, CH_3 Et). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.4, 138.9, 138.7, 138.4, 138.4, 138.0 ($6C_q$ Ar), 128.5, 128.4, 128.2, 128.2, 128.1, 128.1, 128.0, 128.0, 127.7, 127.7, 127.6, 127.6, 127.6, 127.4, 127.0, 126.8 (30CH Ar), 97.4 (C1'), 81.8, 81.8, 79.4, 79.4, 77.8, 77.8, 75.6 (CH_2 Bn), 74.1, 73.6, 73.0, 72.9, 72.1 ($5CH_2$ Bn), 71.1, 70.8 (C6), 68.4 (CH_2 Et), 68.4 (C6'), 42.8 (C5, assigned by HSQC), 33.0 (C1 and C7, overlap), 15.9 (CH_3 Et). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{57}H_{63}NO_9Na$ 928.4395, found 928.4391.

Synthesis of 1-azido-8-trifluoromethylsulfonyloctane

To a dry DCM (0.1 M, 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 ^{0}C. Triflic anhydride (118 µL, 0.70 mmol) was added and the mixture was stirred for 15 min. Then the mixture was diluted with DCM (50 mL), washed with cold water (3 x 30 mL). The organic layer was dried with anhydrous Na_2SO_4, filtered and concentrated *in vacuo* at rt. The crude was used directly for the alkylation of the aziridine.

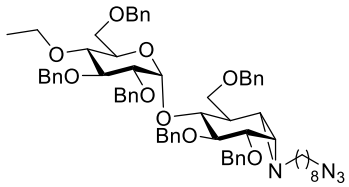
Compound 15a



Aziridine **14a** (70 mg, 72 µmol) was dissolved in dry $CHCl_3$ (0.8 mL) and cooled 0 ^{0}C. Then DIPEA (15 µL, 87 µmol) and freshly-made 8-azidooctyl trifluoromethanesulfonate (1.0 M in $CHCl_3$, 216 µL) were added and the mixture was stirred at rt for 2 h until TLC-analysis showed complete conversion of **14a**. Then another portion of DIPEA (38 µL, 216 µmol) and MeOH (220 µL) were added and the mixture was stirred at rt overnight. The reaction mixture was diluted with EtOAc (60 mL), washed with sat. aq. $NaHCO_3$ (30 mL), water and brine, dried with anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The product was purified by silica gel column chromatography (pentane/acetone 20:1→18:1) to obtain **16a** (73 mg, 65 µmol, 90%) as a clean oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.43 – 6.96 (m, 35H, CH Ar), 5.56 (d, *J* = 3.6 Hz, 1H, H1'), 4.97 – 4.85 (m, 3H, CH*H* Bn), 4.78 (t, *J* = 10.8 Hz, 2H, CH*H* Bn), 4.74 – 4.64 (m, 2H, CH*H* Bn), 4.59 – 4.28 (m, 7H, CH*H* Bn), 3.98 (t, *J* = 9.3 Hz, 1H, H3'), 3.88 – 3.72 (m, 3H, H3, H2 and H6a), 3.67 – 3.52 (m, 5H, H6b, H4, H4', H5' and H6'a), 3.44 (dt, *J* = 10.4, 2.8 Hz, 2H, H6'b and H2'), 3.23 (t, *J* = 7.0 Hz, 2H, CH_2N_3), 2.31 (td, *J* = 7.0, 6.6, 3.3 Hz, 2H, H5 and CH*H*N), 2.07 (dt, *J* = 10.5, 6.9 Hz, 1H, CH*H*N), 1.80 (dd, *J* = 6.5, 3.0 Hz, 1H, H1), 1.56 (dq, *J* = 14.9, 7.6, 6.9 Hz, 5H, H7 and 4CH*H* linker), 1.39 – 1.20 (m,

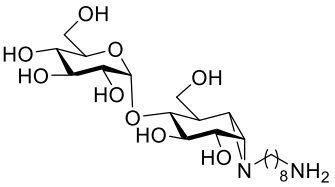
8H, 8CH*H* linker). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.8, 139.0, 138.9, 138.6, 138.5, 138.3, 138.0 (7C_q Ar), 128.4, 128.4, 128.3, 128.2, 128.2, 128.1, 128.0, 128.0, 127.9, 127.8, 127.7, 127.6, 127.5, 127.4, 127.3, 126.8, 126.6 (35CH Ar), 97.7 (C1'), 83.1 (C3), 82.0 (C3'), 80.3 (C2), 79.5 (C2'), 77.9 (C4'), 75.9 (C4), 75.6, 75.1, 74.5, 73.6, 72.9, 72.9, 72.2 (7CH_2 Bn), 70.9 (C5'), 70.8 (C6), 68.3 (C6'), 61.2 (CH_2N), 51.6 (CH_2N_3), 43.4 (C5), 42.1 (C1), 41.6 (C7), 29.9, 29.6, 29.2, 28.9, 27.4, 26.8 (6CH_2 linker). HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{70}H_{81}N_4O_9$ 1121.5998, found 1121.5994.

Compound 15b



Starting from aziridine **14b** (63 mg, 70 µmol), the reaction was carried out following the same procedure described for **15a**. The product was purified by flash column chromatography (pentane/EtOAc 13:1→11:1) to obtain **15b** (65 mg, 61 µmol, 88%) as a clean oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.49 – 6.95 (m, 30H, CH Ar), 5.51 (d, J = 3.6 Hz, 1H, H1'), 4.89 (d, J = 1.4 Hz, 2H, CH*H* Bn), 4.84 (d, J = 10.8 Hz, 1H, CH*H* Bn), 4.79 – 4.62 (m, 3H, CH*H* Bn), 4.57 (d, J = 12.1 Hz, 1H, CH*H* Bn), 4.52 – 4.43 (m, 3H, CH*H* Bn), 4.43 – 4.32 (m, 2H, CH*H* Bn), 3.92 – 3.68 (m, 5H), 3.67 – 3.58 (m, 3H), 3.53 (dd, J = 10.5, 3.1 Hz, 1H), 3.48 – 3.34 (m, 4H), 3.23 (t, J = 7.0 Hz, 2H, CH_2N_3), 2.36 – 2.25 (m, 2H, H5 and CH*H*N), 2.09 (dt, J = 11.5, 7.4 Hz, 1H, CH*H*N), 1.80 (dd, J = 6.5, 3.0 Hz, 1H, H1), 1.62 – 1.48 (m, 5H, H7 and 4CH*H* linker), 1.38 – 1.24 (m, 8H, 8CH*H* linker), 1.05 (t, J = 7.0 Hz, 3H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.8, 139.0, 139.0, 138.5, 138.4, 138.0 (6C_q Ar), 128.4, 128.4, 128.3, 128.2, 128.1, 128.0, 127.9, 127.7, 127.6, 127.5, 127.5, 127.4, 127.3, 126.7, 126.6 (30CH Ar), 97.8 (C1'), 83.0, 81.8, 80.3, 79.4, 77.9, 76.4, 75.5 (CH_2 Bn), 74.4, 73.6, 72.9, 72.9, 72.2 (5CH_2 Bn), 71.0, 71.0 (C6), 68.4 (CH_2 Et), 68.3 (C6'), 61.2 (CH_2N), 51.6 (CH_2N_3), 43.3 (C5), 42.1 (C1), 41.6 (C7), 29.9, 29.6, 29.2, 28.9, 27.4, 26.8 (6CH_2 linker), 15.9 (CH_3). HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{65}H_{79}N_4O_9$ 1059.5842, found 1059.5838.

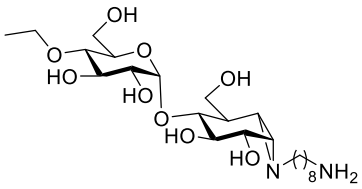
Compound 2a



Ammonium (3.0 mL) was condensed in a dry flask at -60 ^{0}C, and sodium (34 mg, 1.4 mmol) was added. The resulting deep-blue solution was stirred for 15 min to dissolve all sodium. Aziridine **15a** (23 mg, 20 µmol) and ^{t}BuOH (20 µL, 200 µmol) were taken up in dry THF (1.0 mL) and slowly added to the reaction mixture. After stirring for 1 h, the reaction was carefully 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 (MH_4^+) with 0.5 M NH_4OH as eluent, concentrated *in vacuo*, affording the title compound (7.5 mg, 16 µmol, 79%) as a light-yellow powder after lyophilization. ^{1}H NMR (400 MHz, D_2O) δ 5.11 (d, J = 4.0 Hz, 1H, H1'), 3.94 – 3.64 (m, 7H), 3.60 – 3.47 (m, 2H), 3.35 (dt, J = 26.7, 9.7 Hz, 2H), 2.94 (t, J = 7.5 Hz, 2H, C*H*$_2$NH$_2$), 2.27 (t, J = 7.5 Hz, 2H, CH_2N), 2.11 – 1.93 (m, 2H, H5 and H1), 1.82 (d, J = 6.6 Hz, 1H, H7), 1.66 – 1.57 (m, 2H,

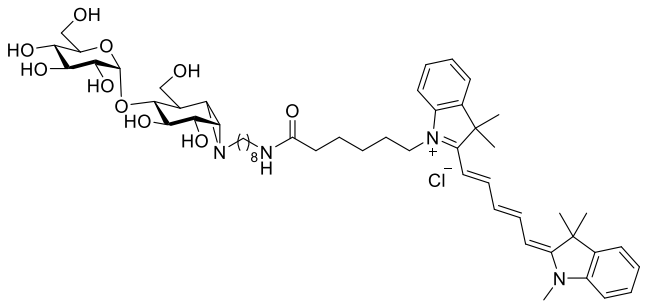
2CH*H* linker), 1.51 (s, 2H, 2CH*H* linker), 1.36 – 1.22 (m, 8H, 8CH*H* linker). ^{13}C NMR (101 MHz, D_2O) δ 101.0 (C1'), 81.3, 74.9, 73.3, 72.9, 72.1, 71.3, 69.7, 61.9 (C6'), 60.8 (C6), 60.4 (CH_2N), 44.5 (C1), 43.6 (C5), 41.1 (C7), 39.9 (CH_2NH_2), 29.0, 28.8, 28.4, 27.2, 26.7, 25.9 (6CH_2 linker). HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{21}H_{41}N_2O_9$ 465.2807, found 465.2806.

Compound 16



Ammonium (8.0 mL) was condensed in a dry flask at -60 ^{0}C, and sodium (94 mg, 4.1 mmol) was added. The resulting deep-blue solution was stirred for 15 min to dissolve all sodium. Aziridine **15b** (62 mg, 59 µmol) and ^{t}BuOH (56 µL, 0.59 mmol) were taken up in dry THF (1.5 mL) and slowly added to the reaction mixture. After stirring for 1 h, the reaction 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 (MH_4^+) with 0.5 M NH_4OH as eluent, concentrated *in vacuo*, affording the title compound (27 mg, 55 µmol, 94%) as a white powder after lyophilization. ^{1}H NMR (400 MHz, D_2O) δ 5.09 (d, J = 4.0 Hz, 1H, H1'), 3.89 – 3.64 (m, 9H), 3.61 – 3.48 (m, 2H), 3.34 – 3.21 (m, 2H), 2.93 (dt, J = 15.2, 7.2 Hz, 2H, C*H*$_2$NH$_2$), 2.26 (t, J = 7.3 Hz, 2H, CH_2N), 2.06 – 1.94 (m, 2H, H1 and H5), 1.81 (d, J = 6.6 Hz, 1H, H7), 1.66 – 1.41 (m, 4H, 4CH*H* linker), 1.29 (s, 8H, 8CH*H* linker), 1.17 (t, J = 7.0 Hz, 3H, CH_3). ^{13}C NMR (101 MHz, D_2O) δ 100.5 (C1'), 80.9, 77.5, 74.6, 72.8, 71.7, 71.6, 70.9, 68.9 (CH_2 Et), 61.5 (C6), 60.1 (C6'), 60.0 (CH_2N), 44.2 (C1), 43.1 (C5), 40.7 (C7), 39.6 (CH_2NH_2), 28.6, 28.4, 28.1, 27.2, 26.4, 25.5 (6CH_2 linker), 14.7 (CH_3). HRMS (ESI) m/z: $[M+H]^+$ calc for $C_{23}H_{45}N_2O_9$ 493.3120, found 493.3120.

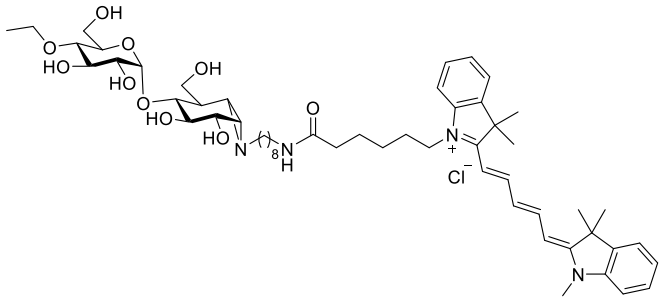
Compound 2b



Compound **2a** (7.0 mg, 15 µmol) was dissolved in dry DMF (0.5 mL), then DIPEA (5.7 µL, 33 µmol) and Cy5-OSu (10 mg, 16 µmol) were added and the mixture was stirred at rt for 16 h. Full conversion was observed by LC-MS and the product was purified by semi-preparative reversed phase HPLC (linear gradient. Solution used: A: 50 mM NH_4HCO_3 in H_2O, B: MeCN). The fractions were concentrated under reduced pressure, co-evaporated with Milli-Q/MeCN (1/1, 3 x), dissolved in Milli-Q/^{t}BuOH (1/1) again and lyophilized to obtain the title compound (1.4 mg, 1.5 µmol, 10%) as a blue powder. ^{1}H NMR (600 MHz, MeOD) δ 8.29 – 8.20 (m, 2H), 7.50 (dt, J = 7.4, 1.4 Hz, 2H), 7.42 (tdd, J = 7.6, 4.9, 1.2 Hz, 2H), 7.33 – 7.23 (m, 4H), 6.63 (t, J = 12.5 Hz, 1H), 6.28 (dd, J = 13.7, 7.4 Hz, 2H), 4.98 (d, J = 3.9 Hz, 1H, H1'), 4.11 (t, J = 7.4 Hz, 2H, CH_2N^+), 3.88 (dd, J = 11.0, 3.1 Hz, 1H, H6a), 3.85 – 3.78 (m, 1H, H6a'), 3.76 – 3.66 (m, 4H, H2, H5', H6b and H6'b), 3.62 (d, J = 15.3 Hz, 4H, CH_3N and H3'), 3.55 (dd, J = 9.9,

8.8 Hz, 1H, H3), 3.45 – 3.40 (m, 1H, H2'), 3.27 (d, *J* = 9.2 Hz, 1H, H4'), 3.21 – 3.15 (m, 1H, H4), 3.12 (t, *J* = 7.2 Hz, 2H, C*H*$_2$NH), 2.31 (dt, *J* = 11.6, 7.1 Hz, 1H, CH*H*N aziridine), 2.20 (t, *J* = 7.3 Hz, 2H, CH$_2$C=O), 2.13 (t, *J* = 5.8 Hz, 1H, CH*H*N aziridine), 1.95 – 1.88 (m, 1H, H5), 1.87 – 1.79 (m, 3H, H1 and 2CH*H* linker), 1.73 (s, 15H, H7, 2CH*H* linker and 4CH$_3$), 1.56 (q, *J* = 7.5 Hz, 2H, 2CH*H* linker), 1.50 – 1.42 (m, 4H, 4CH*H* linker), 1.36 – 1.27 (m, 8H, 8CH*H* linker). ^{13}C NMR (214 MHz, MeOD) *δ* 175.7, 175.4, 174.7, 155.5, 155.5, 144.3, 143.6, 142.7, 142.5, 129.8, 129.8, 126.6, 126.3, 126.3, 123.4, 123.3, 112.1, 111.9, 104.4, 104.3, 103.2 (C1'), 84.1 (C4), 75.9 (C3), 75.0 (C3'), 74.4, 74.0 (C2'), 72.7, 71.4 (C4'), 62.9 (C6), 62.7 (C6'), 62.2 (CH$_2$N aziridine), 50.6 (C$_q$), 50.5 (C$_q$), 45.9 (C1), 45.8 (C5), 44.8 (CH$_2$N$^+$), 42.1 (C7), 40.4 (CH$_2$NH), 36.7 (*C*H$_2$C=O), 31.5 (CH$_3$N), 30.7, 30.5, 30.4, 30.3, 28.4, 28.2, 28.0 (7CH$_2$ linker), 27.9 (2CH$_3$), 27.8 (2CH$_3$), 27.4, 26.6 (2CH$_2$ linker). HRMS (ESI) m/z: [M]$^+$ calc for C$_{53}$H$_{77}$N$_4$O$_{10}$ 929.5634, found 929.5627.

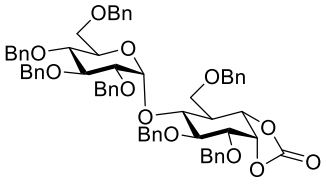
Compound 2c



Compound **16** (7.0 mg, 14 µmol) was dissolved in dry DMF (0.5 mL), then DIPEA (5.5 µL, 31 µmol) and Cy5-OSu (10 mg, 17 µmol) were added and the mixture was stirred at rt overnight. Full conversion was observed by LC-MS and the product was purified by semi-preparative reversed phase HPLC (linear gradient. Solution used: A: 50 mM NH$_4$HCO$_3$ in H$_2$O, B: MeCN). The fractions were concentrated under reduced pressure, co-evaporated with Milli-Q/MeCN (1/1, 3 x), dissolved in Milli-Q/^{t}BuOH (1/1) again and lyophilized to obtain the title compound (2.4 mg, 2.4 µmol, 17%) as a blue powder. ^{1}H NMR (850 MHz, MeOD) *δ* 8.30 – 8.21 (m, 2H), 7.50 (dt, *J* = 7.3, 1.5 Hz, 2H), 7.44 – 7.39 (m, 2H), 7.32 – 7.25 (m, 4H), 6.63 (t, *J* = 12.4 Hz, 1H), 6.28 (dd, *J* = 13.7, 11.2 Hz, 2H), 4.97 (d, *J* = 3.9 Hz, 1H, H1'), 4.11 (t, *J* = 7.5 Hz, 2H, CH$_2$N$^+$), 3.92 – 3.84 (m, 2H, H6a and CH*H*O Et), 3.77 (dd, *J* = 11.7, 2.1 Hz, 1H, H6'a), 3.74 – 3.61 (m, 9H, H2, H3', H5', H6b, H6'b, CH*H*O Et and CH$_3$N), 3.56 – 3.52 (m, 1H, H3), 3.47 – 3.40 (m, 1H, H2'), 3.20 – 3.09 (m, 4H, H4, H4' and C*H*$_2$NH), 2.31 (ddt, *J* = 14.7, 7.7, 3.6 Hz, 1H, CH*H*N aziridine), 2.20 (t, *J* = 7.2 Hz, 2H, CH$_2$C=O), 2.15 (ddd, *J* = 11.7, 8.7, 7.2 Hz, 1H, CH*H*N aziridine), 1.90 (ddd, *J* = 9.9, 6.7, 3.1 Hz, 1H, H5), 1.87 – 1.75 (m, 3H, H1 and 2CH*H* linker), 1.73 (s, 15H, H7, 2CH*H* linker and 4CH$_3$), 1.64 – 1.52 (m, 2H, 2CH*H* linker), 1.46 (dq, *J* = 14.2, 6.9, 6.4 Hz, 4H, 4CH*H* linker), 1.39 – 1.27 (m, 8H, 8CH*H* linker), 1.20 – 1.15 (m, 3H, CH$_3$ Et). ^{13}C NMR (214 MHz, MeOD) *δ* 175.7, 175.4, 174.7, 155.5, 155.5, 144.3, 143.6, 142.7, 142.5, 129.78, 129.8, 126.7, 126.3, 126.3, 123.4, 123.3, 112.1, 111.9, 104.4, 104.3, 103.1 (C1'), 84.1, 79.2, 75.8 (C3), 75.2, 74.1 (C2'), 73.6, 72.7, 69.3 (CH$_2$ Et), 62.9 (C6), 62.1 (C6'), 62.1 (CH$_2$N aziridine), 50.6 (C$_q$), 50.5 (C$_q$), 45.9 (C1), 45.8 (C5), 44.8 (CH$_2$N$^+$), 42.2 (C7), 40.4 (CH$_2$NH), 36.7 (*C*H$_2$C=O), 31.51 (CH$_3$N), 30.7,

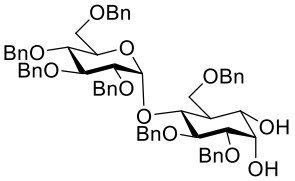
30.4, 30.4, 30.3, 28.4, 28.2, 28.0 ($7CH_2$ linker), 28.0 ($2CH_3$), 27.8 ($2CH_3$), 27.4, 26.6 ($2CH_2$ linker), 16.0 (CH_3 Et). HRMS (ESI) m/z: $[M]^+$ calc for $C_{55}H_{81}N_4O_{10}$ 957.5947, found 957.5945.

Compound 18



Starting from compound **17** (98 mg, 0.20 mmol), using donor **4a** (214 mg, 0.30 mmol) and following **Standard procedure A**. The reaction mixture was purified by size exclusion (DCM/MeOH = 1/1) to give the product as a mixture of two isomers (130 mg, α/β = 13/1, determined by ^{1}H NMR). Then the mixture was further purified by flash column chromatography (pentane/acetone 13:1→9:1) to obtain compound **18** (115 mg, 114 µmol, 57%) as a clean oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.39 – 7.03 (m, 35H, CH Ar), 5.14 (d, *J* = 3.5 Hz, 1H, H1'), 5.03 (t, *J* = 9.0 Hz, 1H, H7), 4.92 (d, *J* = 10.9 Hz, 1H, CH*H* Bn), 4.82 (dd, *J* = 10.9, 6.3 Hz, 2H, CH*H* Bn), 4.79 – 4.72 (m, 2H, 1CH*H* Bn and H1), 4.61 (s, 2H, CH*H* Bn), 4.55 (d, *J* = 11.9 Hz, 1H, CH*H* Bn), 4.44 (td, *J* = 21.8, 20.8, 12.0 Hz, 4H, CH*H* Bn), 4.25 (d, *J* = 12.2 Hz, 2H, CH*H* Bn), 4.17 (d, *J* = 11.5 Hz, 1H, CH*H* Bn), 3.99 – 3.83 (m, 5H, H2, H3', H3, H4 and H6a), 3.76 (dt, *J* = 10.1, 2.4 Hz, 1H, H5'), 3.73 – 3.66 (m, 1H, H4'), 3.62 (dd, *J* = 9.7, 2.4 Hz, 1H, H6b), 3.57 (dd, *J* = 9.8, 3.5 Hz, 1H, H2'), 3.42 (dd, *J* = 10.9, 2.8 Hz, 1H, H6'a), 3.29 (dd, *J* = 10.9, 1.9 Hz, 1H, H6'b), 2.80 (ddd, *J* = 11.9, 9.4, 2.3 Hz, 1H, H5). ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.8 (C=O), 138.8, 138.6, 138.4, 138.1, 138.0, 137.2, 137.1 ($7C_q$ Ar), 128.7, 128.6, 128.5, 128.4, 128.4, 128.3, 128.2, 128.1, 128.1, 128.0, 128.0, 128.0, 127.9, 127.8, 127.7, 127.6, 127.6, 127.3 (35CH Ar), 95.1 (C1'), 82.3, 81.0, 79.8 (C2'), 77.7 (C4'), 75.7, 75.1 ($2CH_2$ Bn), 74.1, 74.2 (C1), 73.9 (C7), 73.8, 73.5, 73.3, 73.0 ($4CH_2$ Bn), 71.9, 71.6 (CH_2 Bn), 71.0 (C5'), 68.1 (C6'), 65.6 (C6), 41.3 (C5). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{63}H_{64}O_{12}Na$ 1035.4290, found 1035.4286.

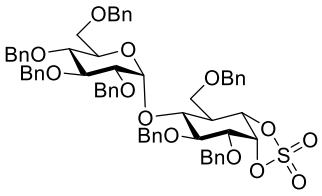
Compound 19



Compound **18** (90 mg, 89 µmol) was dissolved in DCM/MeOH (1/1, 2.0 mL), then 30 wt% NaOMe in MeOH solution (20 µL, 0.11 mmol) was added and the reaction mixture was stirred at rt for 2 h. After which the mixture was diluted with DCM (2.0 mL) and neutralized with washed Amberlite IR-120 H^+ resin until pH ≈ 7, filtered, washed with DCM (3 x 3 mL) and concentrated *in vacuo*. The product was purified by flash column chromatography (pentane/acetone 7:1→3:1) to obtain compound **19** (82 mg, 83 µmol, 93%) as a clean oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.04 (m, 35H, CH Ar), 5.81 (d, *J* = 3.7 Hz, 1H, H1'), 4.99 (d, *J* = 11.9 Hz, 1H, CH*H* Bn), 4.89 (d, *J* = 10.9 Hz, 1H, CH*H* Bn), 4.83 – 4.72 (m, 3H, CH*H* Bn), 4.65 – 4.56 (m, 3H, CH*H* Bn), 4.56 – 4.48 (m, 2H, CH*H* Bn), 4.47 – 4.36 (m, 3H, CH*H* Bn), 4.27 (d, *J* = 12.2 Hz, 1H, CH*H* Bn), 4.15 (t, *J* = 2.8 Hz, 1H), 4.05 (t, *J* = 9.0 Hz, 1H), 3.97 (dd, *J* = 9.9, 8.8 Hz, 1H), 3.91 – 3.68 (m, 5H), 3.65 (dd, *J* = 10.1, 8.8 Hz, 1H), 3.54 – 3.41 (m, 3H), 3.38 (dd, *J* = 10.7, 1.9 Hz, 1H), 2.49 (brs, 3H), 2.30 (tt, *J* = 10.6, 3.8 Hz, 1H). ^{13}C NMR

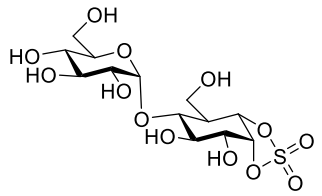
(101 MHz, $CDCl_3$) δ 139.0, 138.9, 138.6, 138.2, 138.1, 138.0, 137.8 ($7C_q$ Ar), 128.6, 128.4, 128.4, 128.3, 128.2, 128.1, 128.0, 127.9, 127.8, 127.8, 127.7, 127.6, 127.6, 127.3, 127.1, 126.5, 96.4 (C1'), 82.8, 82.1, 80.5, 79.5, 77.8, 75.6 (CH_2 Bn), 75.0, 74.0, 73.6, 73.4, 72.9, 72.5 ($6CH_2$ Bn), 71.0, 69.8, 69.0, 68.3 (C6'), 67.6 (C6), 42.9 (C5). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{62}H_{66}O_{11}Na$ 1009.4497, found 1009.4495.

Compound 20



Compound **19** (82 mg, 83 µmol) was dissolved in dry DCM (1.5 mL) and purged with N_2. After cooling to 0 ^{0}C, TEA (92 µL, 0.66 mmol) and thionyl chloride (24 µL, 0.33 mmol) were added successively. After stirring at 0 ^{0}C for 15 min, the reaction was diluted with DCM (50 mL), washed with sat. aq. $NaHCO_3$ (30 mL), H_2O (2 x 30 mL) and brine, dried over Na_2SO_4, filtered and concentrated *in vacuo* to give the crude sulfite as a pale-yellow oil. The crude was directly dissolved in a mixture of EtOAc and MeCN (1/1, 2.0 mL) and cooled to 0 ^{0}C. A solution of $RuCl_3.3H_2O$ (2 mg, 8.3 µmol) and $NaIO_4$ (35 mg, 0.16 mmol) in H_2O (0.4 mL) was added and the mixture was stirred vigorously at 0 ^{0}C for 2 h. The reaction was quenched by addition of sat. aq. $Na_2S_2O_3$ (1 mL) and the mixture was stirred vigorously for 15 min. Then the mixture was diluted with H_2O (30 mL) and extracted with EtOAc (2 x 40 mL). The combined organic layers were washed with brine, dried over Na_2SO_4, filtered and concentrated *in vacuo*. The crude was purified by flash column chromatography (pentane/acetone 13:1→11:1) affording compound **20** (50 mg, 48 µmol, 58%) as a clean oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.45 – 6.99 (m, 35H, CH Ar), 5.28 (d, *J* = 3.7 Hz, 1H, H1'), 5.22 (dd, *J* = 9.8, 6.5 Hz, 1H, H7), 4.99 (dd, *J* = 6.9, 3.4 Hz, 1H, H1), 4.90 (d, *J* = 10.4 Hz, 1H, CH*H* Bn), 4.81 (dd, *J* = 10.9, 4.0 Hz, 2H, CH*H* Bn), 4.76 – 4.58 (m, 3H, CH*H* Bn), 4.56 – 4.20 (m, 8H, CH*H* Bn), 3.99 – 3.85 (m, 5H, H2, H3', H3, H4 and H6a), 3.78 – 3.63 (m, 2H, H5' and H4'), 3.62 – 3.51 (m, 2H, H6b and H2'), 3.49 – 3.40 (m, 1H, H6'a), 3.31 (d, *J* = 10.8 Hz, 1H, H6'b), 2.87 (t, *J* = 10.1 Hz, 1H, H5). ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.7, 138.5, 138.2, 137.9, 137.8, 137.3, 137.2 ($7C_q$ Ar), 128.7, 128.6, 128.6, 128.5, 128.4, 128.4, 128.3, 128.1, 128.0, 128.0, 127.9, 127.8, 127.8, 127.7, 127.3 (35CH Ar), 95.8 (C1'), 82.1, 81.1, 79.7 (C2'), 79.4 (C1), 78.3 (C7), 77.7 (C4'), 75.7, 75.1 ($2CH_2$ Bn), 74.3, 73.7 (CH_2 Bn), 73.6, 73.5, 73.0, 72.4 ($4CH_2$ Bn), 71.2, 71.1 (C5'), 68.1 (C6'), 64.8 (C6), 41.8 (C5). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{62}H_{64}O_{13}SNa$ 1071.3960, found 1071.3965.

Compound 3a



Compound **20** (20 mg, 19 µmol) was dissolved in a mixture of MeOH/H_2O/dioxane (2/1/2, 2 mL) under Argon and $Pd(OH)_2$/C (20 wt%, 20 mg, 29 µmol) was added. While stirring vigorously, the mixture was flushed with a H_2 balloon. After stirring for 3 h under H_2 atmosphere, the mixture was filtered over a small celite pad and evaporate to afford the product **3a** in high purity as a white solid (7.1 mg, 17 µmol, 89%) after lyophilization. ^{1}H NMR (400

MHz, D_2O) δ 5.47 (dd, J = 4.6, 3.4 Hz, 1H, H1), 5.35 – 5.23 (m, 2H, H1' and H7), 4.02 – 3.87 (m, 3H, H2, H3 and H6a), 3.86 – 3.55 (m, 7H), 3.40 (t, J = 9.2 Hz, 1H), 2.43 – 2.32 (m, 1H, H5). ^{13}C NMR (101 MHz, D_2O) δ 100.3 (C1'), 84.4 (C1), 82.1 (C7), 76.2, 73.2, 72.9, 72.8, 71.7, 69.2, 68.0, 60.4 (C6'), 56.5 (C6), 43.7 (C5). HRMS (ESI) m/z: $[M+Na]^+$ calc for $C_{13}H_{22}O_{13}SNa$ 441.0673, found 441.0670.

2.6 References

[1] Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P. M.; Henrissat, B., *Nucleic Acids Res.* **2014**, *42* (D1), D490-D495.

[2] Huang, N.; Stebbins, G. L.; Rodriguez, R. L., *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 7526-7530.

[3] Brayer, G. D.; Luo, Y.; Withers, S. G., *Protein Sci.* **1995**, *4*, 1730-1742.

[4] Gupta, R.; Gigras, P.; Mohapatra, H.; Goswami, V. K.; Chauhan, B., *Process Biochem.* **2003**, *38* (11), 1599-1616.

[5] Zechel, D. L.; Withers, S. G., *Acc. Chem. Res.* **2000**, *33* (1), 11-18.

[6] Nakazawa, Y.; Wang, Y. J., *Carbohydr. Res.* **2003**, *338* (24), 2871-2882.

[7] Wang, S.; Copeland, L., *Crit. Rev. Food Sci. Nutr.* **2015**, *55* (8), 1081-1097.

[8] Kirk, O.; Borchert, T. V.; Fuglsang, C. C., *Curr. Opin. Biotech.* **2002**, *13* (4), 345-351.

[9] van der Maarel, M. J. E. C.; van der Veen, B.; Uitdehaag, J. C. M.; Leemhuis, H.; Dijkhuizen, L., *J. Biotech.* **2001**, *94*, 137-155.

[10] Souza, P. M. d.; Magalhães, P. d. O. e., *Braz. J. Microbiol.* **2010**, *41*, 850-861.

[11] Sundarram, A.; Murthy, T. P. K., *J. Appl. Environ. Microbiol.* **2014**, *2*, 166-175.

[12] Hmidet, N.; El-Hadj Ali, N.; Haddar, A.; Kanoun, S.; Alya, S.-K.; Nasri, M., *Biochem. Eng. J.* **2009**, *47* (1-3), 71-79.

[13] Rana, N.; Walia, A.; Gaur, A., *Natl. Acad. Sci. Lett.* **2013**, *36* (1), 9-17.

[14] Mooradian, A. D.; Thurman, J. E., *Drugs* **1999**, *57* (1), 19-29.

[15] Scott, L. J.; Spencer, C. M., *Drugs* **2000**, *59*, 521-549.

[16] Chiasson, J.-L.; Josse, R. G.; Gomis, R.; Hanefeld, M.; Karasik, A.; Laakso, M., *The Lancet* **2002**, *359* (9323), 2072-2077.

[17] Li, C.; Begum, A.; Numao, S.; Park, K. H.; Withers, S. G.; Brayer, G. D., *Biochemistry* **2005**, *44*, 3347-3357.

[18] Nahoum, V.; Roux, G.; Anton, V.; RougÉ, P.; Puigserver, A.; Bischoff, H.; Henrissat, B.; Payan, F., *Biochem. J.* **2000**, *346* (1), 210-208.

[19] Tarling, C. A.; Woods, K.; Zhang, R.; Brastianos, H. C.; Brayer, G. D.; Andersen, R. J.; Withers, S. G., *Chembiochem* **2008**, *9* (3), 433-438.

[20] Williams, L. K.; Zhang, X.; Caner, S.; Tysoe, C.; Nguyen, N. T.; Wicki, J.; Williams, D. E.; Coleman, J.; McNeill, J. H.; Yuen, V.; Andersen, R. J.; Withers, S. G.; Brayer, G. D., *Nat. Chem. Biol.* **2015**, *11* (9), 691-696.

[21] Braun, C.; Brayer, G. D.; Withers, S. G., *J. Biol. Chem.* **1995**, *270*, 26778-26781.

[22] Zhang, R.; McCarter, J. D.; Braun, C.; Yeung, W.; Brayer, G. D.; Withers, S. G., *J. Org. Chem.* **2007**, *73*, 3070-3077.

[23] Zhang, R.; Li, C.; Williams, L. K.; Rempel, B. P.; Brayer, G. D.; Withers, S. G., *Biochemistry* **2009**, *48* (45), 10752-10764.

[24] Caner, S.; Zhang, X.; Jiang, J.; Chen, H. M.; Nguyen, N. T.; Overkleeft, H.; Brayer, G. D.; Withers, S. G., *FEBS Lett.* **2016**, *590* (8), 1143-1151.

[25] Cravatt, B. F.; Wright, A. T.; Kozarich, J. W., *Annu. Rev. Biochem.* **2008**, *77*, 383-414.

[26] Willems, L. I.; Overkleeft, H. S.; van Kasteren, S. I., *Bioconjugate Chem.* **2014**, *25* (7), 1181-1191.

[27] Atsumi, S.; Umezawa, K.; Iinuma, H.; Naganawa, H.; Nakamura, H.; Iitaka, Y.; Takeuchi, T., *J. Antibiot.* **1990**, *43* (1), 49-53.

[28] Jiang, J.; Kallemeijn, W. W.; Wright, D. W.; van den Nieuwendijk, A.; Rohde, V. C.; Folch, E. C.; van den Elst, H.; Florea, B. I.; Scheij, S.; Donker-Koopman, W. E.; Verhoek, M.; Li, N.; Schurmann, M.; Mink, D.; Boot, R. G.; Codée,

J. D. C.; van der Marel, G. A.; Davies, G. J.; Aerts, J.; Overkleeft, H. S., *Chem. Sci.* **2015**, *6* (5), 2782-2789.

[29] Willems, L. I.; Beenakker, T. J.; Murray, B.; Scheij, S.; Kallemeijn, W. W.; Boot, R. G.; Verhoek, M.; Donker-Koopman, W. E.; Ferraz, M. J.; van Rijssel, E. R.; Florea, B. I.; Codée, J. D.; van der Marel, G. A.; Aerts, J. M. F. G.; Overkleeft, H. S., *J. Am. Chem. Soc.* **2014**, *136* (33), 11622-11625.

[30] Jiang, J.; Kuo, C. L.; Wu, L.; Franke, C.; Kallemeijn, W. W.; Florea, B. I.; van Meel, E.; van der Marel, G. A.; Codée, J. D.; Boot, R. G.; Davies, G. J.; Overkleeft, H. S.; Aerts, J. M., *ACS Cent. Sci.* **2016**, *2* (5), 351-358.

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

[32] McGregor, N. G. S.; Artola, M.; Nin-Hill, A.; Linzel, D.; Haon, M.; Reijngoud, J.; Ram, A.; Rosso, M. N.; van der Marel, G. A.; Codée, J. D. C.; van Wezel, G. P.; Berrin, J. G.; Rovira, C.; Overkleeft, H. S.; Davies, G. J., *J. Am. Chem. Soc.* **2020**, *142* (10), 4648-4662.

[33] Schröder, S. P.; de Boer, C.; McGregor, N. G. S.; Rowland, R. J.; Moroz, O.; Blagova, E.; Reijngoud, J.; Arentshorst, M.; Osborn, D.; Morant, M. D.; Abbate, E.; Stringer, M. A.; Krogh, K.; Raich, L.; Rovira, C.; Berrin, J. G.; van Wezel, G. P.; Ram, A. F. J.; Florea, B. I.; van der Marel, G. A.; Codée, J. D. C.; Wilson, K. S.; Wu, L.; Davies, G. J.; Overkleeft, H. S., *ACS Cent. Sci.* **2019**, *5* (6), 1067-1078.

[34] de Boer, C.; McGregor, N. G. S.; Peterse, E.; Schröder, S. P.; Florea, B. I.; Jiang, J.; Reijngoud, J.; Ram, A. F. J.; van Wezel, G. P.; van der Marel, G. A.; Codée, J. D. C.; Overkleeft, H. S.; Davies, G. J., *RSC Chemical Biology* **2020**, *1* (3), 148-155.

[35] Artola, M.; Wu, L.; Ferraz, M. J.; Kuo, C. L.; Raich, L.; Breen, I. Z.; Offen, W. A.; Codée, J. D. C.; van der Marel, G. A.; Rovira, C.; Aerts, J.; Davies, G. J.; Overkleeft, H. S., *ACS Cent. Sci.* **2017**, *3* (7), 784-793.

[36] Wang, L.; Overkleeft, H. S.; van der Marel, G. A.; Codée, J. D. C., *J. Am. Chem. Soc.* **2018**, *140* (13), 4632-4638.

[37] Sherman, H. C.; Kendall, E. C.; Clark, E. D., *J. Am. Chem. Soc.* **1910**, *32* (9), 1073-1086.

[38] Takamine, J., *J. Soc. Chem. Ind.* **1898**, *17* (2), 118-120.

[39] Kato, S.; Shimizu-Ibuka, A.; Mura, K.; Takeuchi, A.; Tokue, C.; Arai, S., *Bioscience, Biotechnology, and Biochemistry* **2014**, *71* (12), 3007-3013.

[40] Alonso-Gil, S.; Coines, J.; André, I.; Rovira, C., *Front. Chem.* **2019**, *7*, 1-10.

[41] Brzozowski, A. M.; Davies, G. J., *Biochemistry* **1997**, *36* (36), 10837-10845.

[42] Fisher, S. Z.; Govindasamy, L.; Tu, C.; Agbandje-McKenna, M.; Silverman, D. N.; Rajaniemi, H. J.; McKenna, R., *Sect. F: Struct. Biol. Crystallogr. Commun.* **2006**, *62* (2), 88-93.

[43] Nekiunaite, L.; Arntzen, M. O.; Svensson, B.; Vaaje-Kolstad, G.; Abou Hachem, M., *Biotechnol. Biofuels* **2016**, *9* (1), 187-202.

[44] Parashar, D.; Satyanarayana, T., *Front. Bioeng. Biotechnol.* **2018**, *6*, 1-14.

[45] Crabb, W.D., Mitchinson, C., *Trends Biotechnol.* **1997**, *15*, 349-352.

[46] Wu, L.; Armstrong, Z.; Schröder, S. P.; de Boer, C.; Artola, M.; Aerts, J. M.; Overkleeft, H. S.; Davies, G. J., *Curr. Opin. Chem. Biol.* **2019**, *53*, 25-36.

[47] Armstrong, Z.; Kuo, C. L.; Lahav, D.; Liu, B.; Johnson, R.; Beenakker, T. J. M.; de Boer, C.; Wong, C. S.; van Rijssel, E. R.; Debets, M. F.; Florea, B. I.; Hissink, C.; Boot, R. G.; Geurink, P. P.; Ovaa, H.; van der Stelt, M.; van der Marel, G. M.; Codée, J. D. C.; Aerts, J.; Wu, L.; Overkleeft, H. S.; Davies, G. J., *J. Am. Chem. Soc.* **2020,** *142* (30), 13021-13029.

[48] Kuharev, J.; Navarro, P.; Distler, U.; Jahn, O.; Tenzer, S., *Proteomics* **2015**, *15* (18), 3140-3151.

[49] Kato, N.; Suyama, S.; Shirokane, M.; Kato, M.; Kobayashi, T.; Tsukagoshi, N., *Appl. Environ. Microbiol.* **2002**, *68* (3), 1250-1256.

[50] Smith, P. K.; Krohn, R. I.; Hermanson, G. T.; Mallia, A. K.; Gartner, F. H.; Provenzano, M. D.; Fujimoto, E. K.; Goeke, N. M.; Olson, B. J.; Klenk, D. C., *Anal. Biochem.* **1985**, *150* (1), 76-85.

[51] Potterton, L.; Agirre, J.; Ballard, C.; Cowtan, K.; Dodson, E.; Evans, P. R.; Jenkins, H. T.; Keegan, R.; Krissinel, E.; Stevenson, K.; Lebedev, A.; McNicholas, S. J.; Nicholls, R. A.; Noble, M.; Pannu, N. S.; Roth, C.; Sheldrick, G.; Skubak, P.; Turkenburg, J.; Uski, V.; von Delft, F.; Waterman, D.; Wilson, K.; Winn, M.; Wojdyr, M., *Acta. Crystallogr., Sect. D: Struct. Biol.* **2018**, *74* (2), 68-84.

[52] Skubák, P.; Pannu, N. S., *Nat. Commun.* **2013,** *4* (1), 1-6.

[53] McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., *J. Appl. Crystallogr.* **2007,** *40* (4), 658-674.

[54] Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A., *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011,** *67* (4), 355-367.

[55] Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., *Acta. Crystallogr., Sect. D: Biol. Crystallogr.* **2010,** *66* (4), 486-501.

[56] Nekiunaite, L.; Arntzen, M. O.; Svensson, B.; Vaaje-Kolstad, G.; Abou Hachem, M., *Biotechnol. Biofuels* **2016**, *9* (1), 187-202.

[57] Li, N.; Kuo, C.-L.; Paniagua, G.; van den Elst, H.; Verdoes, M.; Willems, L. I.; van der Linden, W. A.; Ruben, M.; van Genderen, E.; Gubbens, J.; van Wezel, G. P.; Overkleeft, H. S.; Florea, B. I., *Nat. Protoc.* **2013**, *8* (6), 1155-1168.

[58] Ishihama, Y., *Mol. Cell. Proteomics* **2005**, *4*, 1265-1272.

[59] Kvach, M. V.; Ustinov, A. V.; Stepanova, I. A.; Malakhov, A. D.; Skorobogatyi, M. V.; Shmanai, V. V.; Korshun, V. A., *Eur. J. Org. Chem.* **2008**, *2008* (12), 2107-2117.

[60] Susumu, K.; Uyeda, H. T.; Medintz, I. L.; Pons, T.; Delehanty, J. B.; Mattoussi, H., *J. Am. Chem. Soc.* **2007**, *129*, 13987-13996.

APPENDIX

2.S1 Supplemental Results and Discussion

2.S1.1 Detection sensitivity of ABP 1c towards α-amylase in human plasma

From the medical perspective, it would be very attractive to measure α-amylase in human samples that allow to discriminate between disease and healthy patients. For instance, serum amylase level is often used as a biomarker of acute pancreatitis.[1-3] The normal range of serum amylase level for adults is usually between 15 international units (IU)/L to 110 IU/L (may vary between laboratories).[4] However, patients with acute pancreatitis present elevated α-amylase level in plasma, which can rise up to 4 to 6 times the upper limit of normal range. Here, we explored a potential application of our ABP in biomedical sector by analysing the detection of α-amylase in human plasma.

To determine the sensitivity of detection, the amount of enzyme detectable with the Cy5 ABP **1c** was first explored by spiking recombinant human salivary amylase at varying concentrations (0–1 unit) into 25 μL of normal human plasma (Figure 2.S1, set C), followed by Concanavalin A (ConA) beads purification to clean the sample. Then the purified glycoproteins were labeled with 50 μM ABP **1c** at pH 5.0 for 4 h. Unfortunately, recombinant human salivary amylase can only be detected at the maximum concentration applied here (1 unit/25 μL plasma), which is much higher than the upper limit of amylase level in normal (0.003 unit/25 μL) or pancreatitis (0.017 unit/25 μL) plasma. Besides, all other purified glycoproteins in plasma were also labeled under this condition, which can further decrease the sensitivity of detection. In addition, two positive control experiments were included as well by first incubating pure enzyme with 50 μM ABP **1c** at pH 5.0 for 4 h, with (Figure 2.S1, set B) or without (Figure 2.S1, set A) adding the final reaction mixture into ConA purified plasma. The detection limit of pure recombinant amylase with ABP **1c** (no plasma control, set A) is about 0.01 unit, which is just above the level of 0.017 unit/25 μL plasma of acute pancreatitis patients, but still too insensitive to the level in normal person.

In conclusion, the nonspecific labeling of all other purified glycoproteins together with the low probe potency points to a detection limit of 1 unit/25 μL plasma, meaning that these biochemical probes are not optimal for the detection of α-amylase in human plasma of patients suffering from pancreatitis.

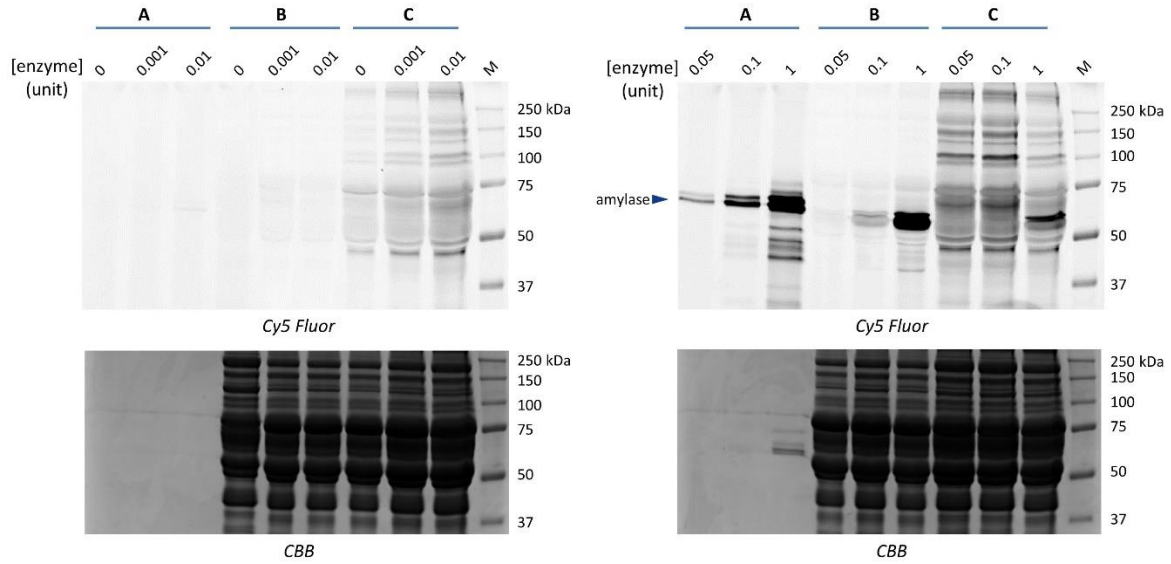


Figure 2.S1. Detection limit of α-amylase in human plasma with ABP **1c**. Set **A**: no plasma control, recombinant human saliva α-amylase (0–1 unit) was incubated with 50 μM ABP **1c** at pH 5.0 for 4 h; Set **B**: recombinant human saliva α-amylase (0–1 unit) was incubated with 50 μM ABP **1c** at pH 5.0 for 4 h, then the reaction mixture was added to ConA purified plasma; Set **C**: recombinant human saliva α-amylase (0–1 unit) was first spiked into 25 μL normal human plasma, followed by ConA purification. The purified glycoproteins were then incubated with 50 μM ABP **1c** at pH 5.0 for 4 h.

2.S1.2 Fluorescent labeling of different saliva samples with ABP 1c

In initial studies on human saliva, labeling of different saliva samples (untreated saliva, saliva supernatant and concentrated saliva pellet) with ABP **1c** were investigated. For labeling, untreated saliva (24 µL per sample), saliva supernatant (24 µL per sample), or concentrated saliva pellet (55 µg total protein per sample, diluted with assay buffer pH 7.0 to a final 24 µL volume) was incubated with 6 µL of ABP **1c** at intended concentrations in assay buffer pH 7.0 at 37 °C for 1 h. As is shown in Figure 2.S2, the detection limit of α-amylase in all these three samples is very similar (10 µM of **1c**) although slightly stronger labeling was observed in untreated saliva and saliva supernatant. However, the protein concentration of human saliva varies among different individuals, to facilitate the quantification of the protein amount needed, and to study the optimal labeling pH, the concentrated saliva pellet (16.9 µg/µL in Kpi buffer) was considered for further studies.

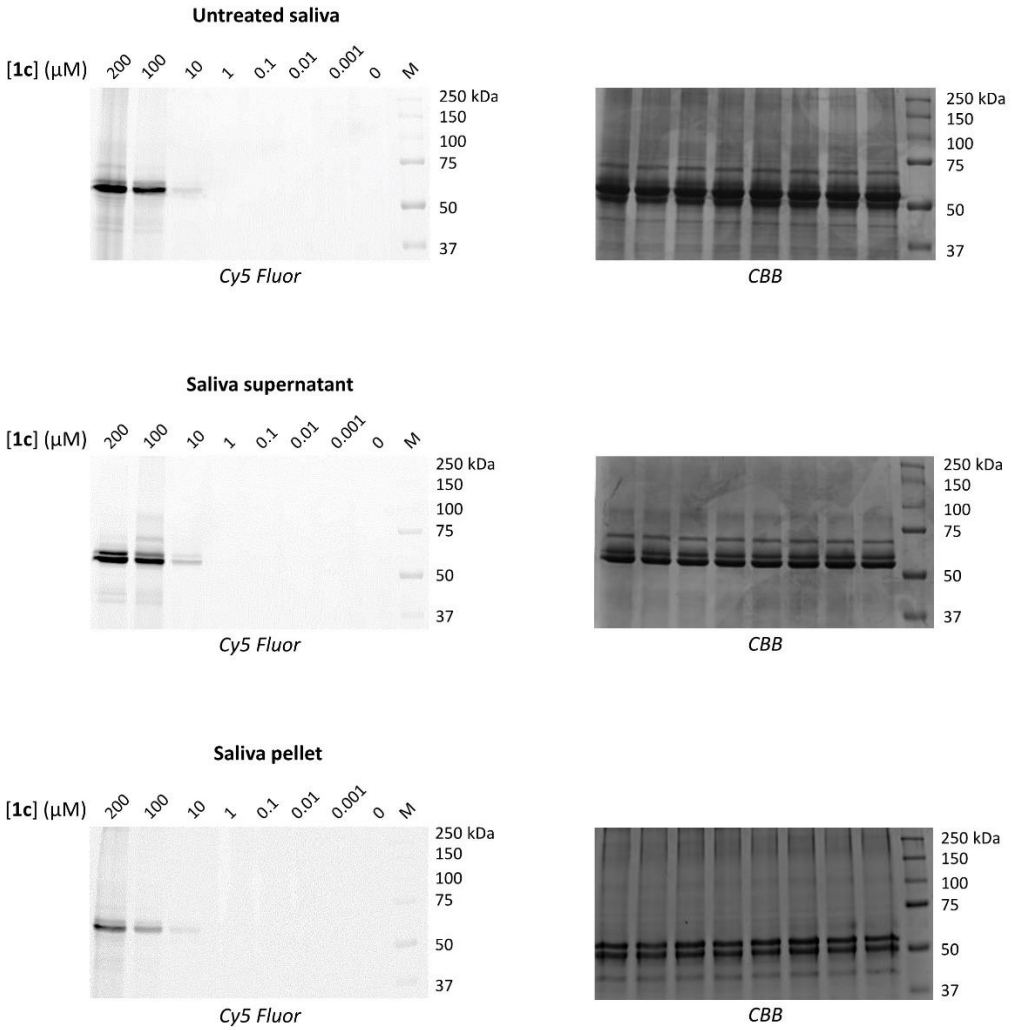


Figure 2.S2. Fluorescent labeling of untreated human saliva (24 µL), saliva supernatant (24 µL) and concentrated human saliva pellet (55 µg total protein) with different concentrations of ABP **1c** (1 h, 37 °C, pH 7.0).

2.S2 Supplemental Figures and Tables

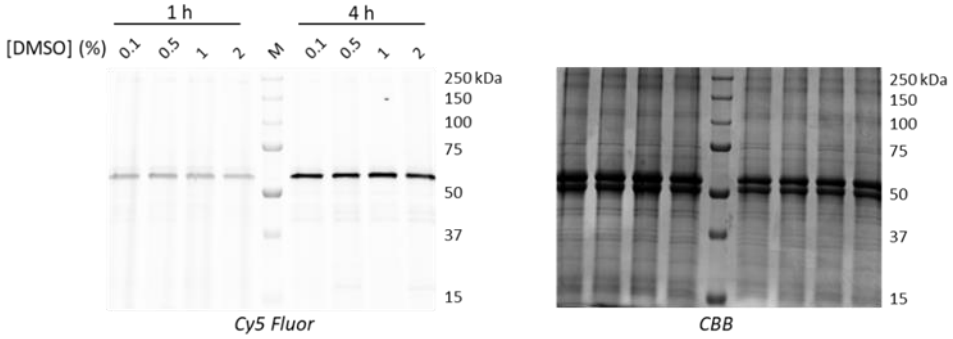


Figure 2.S3. Fluorescent labeling of concentrated human saliva (55 µg total protein) with ABP **1c** (10 µM, 37 °C, pH 5.0, 1 h or 4 h) in the presence of varying concentrations of DMSO (v/v).

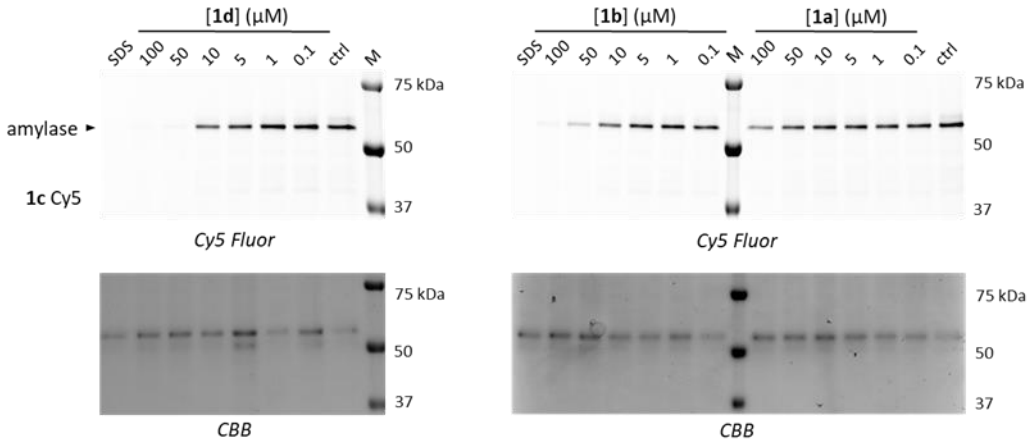


Figure 2.S4. Competitive assay of concentrated human saliva. The sample (5 µg total protein/10 µL) was pre-incubated with 2.5 µL **1a**, **1b** or **1d** at different concentrations (4 h, 37 °C, pH 5.0), followed by labeling with 2.5 µL of 10 µM ABP **1c** (2 h, 37 °C, pH 5.0).

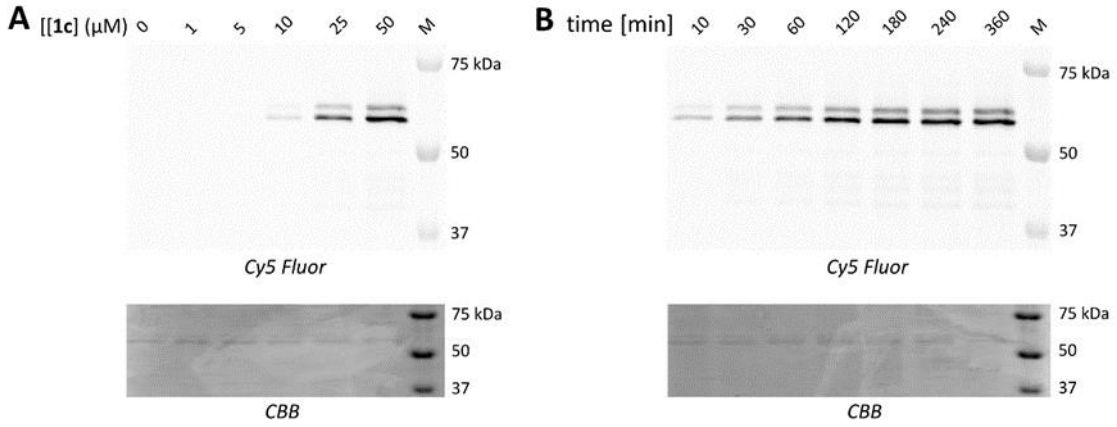


Figure 2.S5. A) Concentration-dependent labeling of recombinant human saliva α-amylase (0.2 unit) with ABP **1c** under optimal pH 5.0 after incubation 4 h at 37 °C; B) Time-dependent labeling of recombinant human saliva α-amylase with 50 μM of ABP **1c** at pH 5.0 and 37 °C.

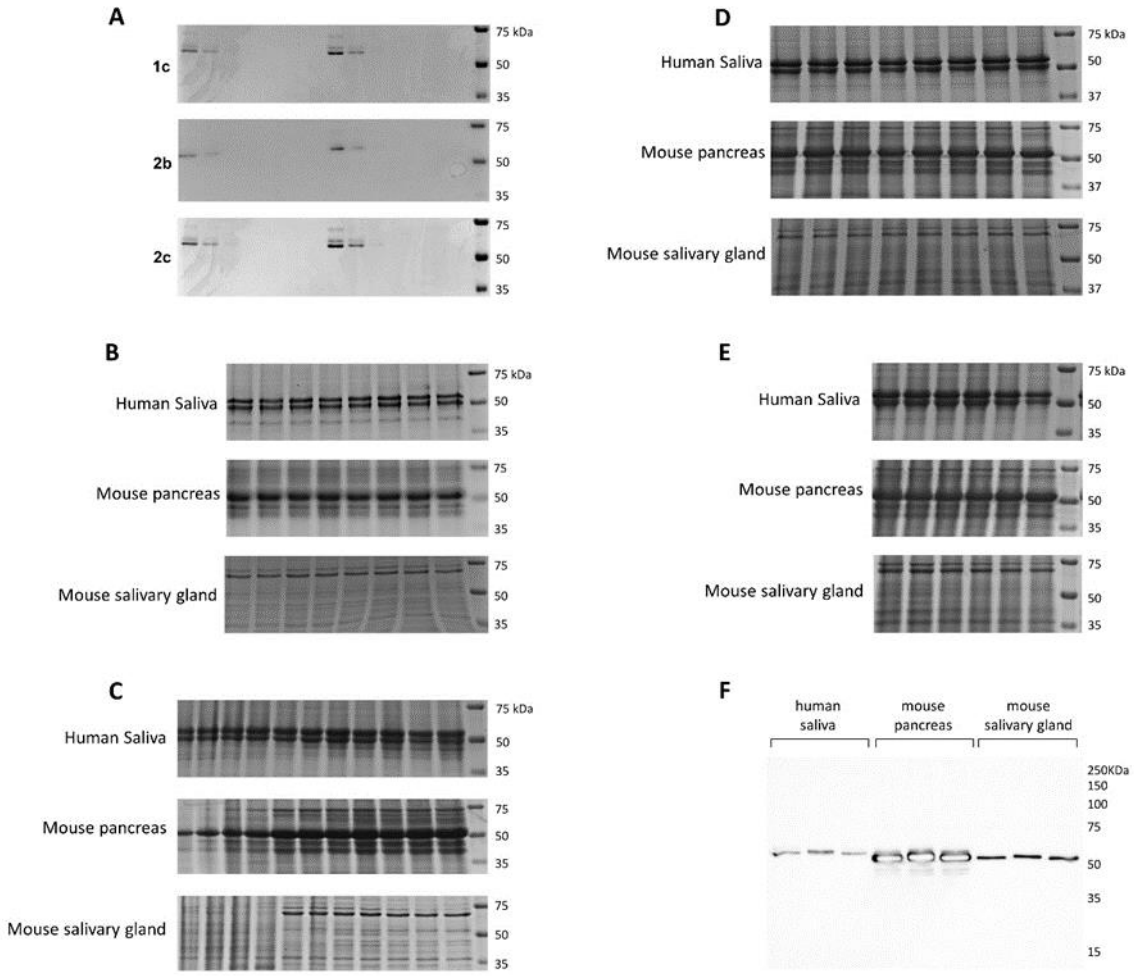


Figure 2.S6. Coomassie Brilliant Blue staining and full-length image of western blot of gels presented in main text Figure 2.5.

Table 2.S1. Data collection and refinement statistics (molecular replacement)

	Taka-amylase	**Taka-amylase + 1a**	**Taka-amylase + 2a**	**Taka-amylase + 3a**	**Taka-amylase + 1b**
PDB ID	6YQ7	6YQ9	6YQA	6YQB	6YQC
Data collection					
Space group	$P12_11$	$P12_11$	$P12_11$	$P12_11$	$P12_11$
Cell dimensions					
a, b, c (Å)	65.2, 103.1, 75.2	65.2, 103.1, 75.4	65.7, 103.3, 75.5	65.6, 102.9, 75.6	64.8, 99.0, 75.5
α, β, γ (°)	90.0, 103.6, 90.0	90.0, 103.6, 90.0	90.0, 103.8, 90.0	90.0, 103.8, 90.0	90.0, 103.6, 90.0
Resolution (Å)	59.65-1.58 (1.61-1.58)	103.10-1.55 (1.58-1.55)	103.32-1.67 (1.70-1.67)	59.78 -1.50 (1.53-1.50)	73.33-1.35 (1.37-1.35)
$CC_{1/2}$	0.99 (0.73)	0.99 (0.87)	0.99 (0.72)	0.99 (0.92)	0.99 (0.76)
I / σI	5.7 (1.7)	11.8 (2.1)	9.6 (1.1)	10.3 (1.6)	7.5 (1.0)
Completeness (%)	98.4 (97.0)	99.9 (99.8)	100 (100)	98.7 (97.0)	95.5 (73.0)
Redundancy	4.2 (4.2)	4.0 (3.8)	4.2 (4.2)	6.8 (6.8)	3.9 (2.7)
Refinement					
Resolution (Å)	1.58	1.55	1.67	1.50	1.35
No. reflections	129,847	139,928	113,215	153,221	192,746
R_{work} / R_{free}	0.16/0.19	0.15/0.17	0.17/0.20	0.16 / 0.19	0.18/0.20
No. atoms					
Protein	7,441	7,480	7,404	7,464	7,468
Ligand/ion	40/2	90/2	105/2	106/2	151/2
Water	518	815	519	785	687
B-factors					
Protein	21	18	28	22	18
Ligand/ion	36/16	23/13	34/23	27/17	22/14
Water	28	28	34	31	28
R.m.s. deviations					
Bond lengths (Å)	0.01	0.01	0.01	0.01	0.01
Bond angles (°)	1.7	1.8	1.7	1.7	1.7
*ligands in the uncomplexed structure are derived from ethylene glycol/Calcium					

Table 2.S2. Proteins identified in concentrated human saliva after biotinylated probe **1d**-treatment, competitor-inhibited control, and no probe control pull down experiments

					Intensity		
Accession	**Protein description**	**Abbreviation**	**cover [%]**	**Score**	**no probe**	**COMP (1b+1d)**	**Probe (1d)**
P60709	Actin, cytoplasmic 1	ACTB	23.7	62.07	3.37E+07	7.06E+07	6.05E+07
P13647	Keratin, type II cytoskeletal 5	KRT5	21	56.28	6.44E+07	4.16E+07	5.04E+07
P04745	Alpha-amylase 1	AMY1A	56.4	323.31	8.55E+06	2.12E+09	1.05E+10
P02769	BOVIN Serum albumin	ALB	42.7	320.79	2.37E+10	1.99E+10	3.18E+10
P04264	Keratin, type II cytoskeletal 1	KRT1	41	283.71	2.48E+09	1.37E+09	2.26E+09
P13645	Keratin, type I cytoskeletal 10	KRT10	34.1	251.8	1.14E+09	6.89E+08	1.14E+09
P00761	Trypsin	TRYP	25.1	231.57	4.89E+09	4.09E+09	4.28E+09
P22629	Streptavidin	SAV	41.7	185.2	2.57E+10	2.42E+10	2.84E+10
P00924	Enolase 1	ENO1	37.1	159.13	2.90E+09	3.04E+09	3.56E+09
P35527	Keratin, type I cytoskeletal 9	KRT9	33.4	154.79	5.52E+08	2.96E+08	3.79E+08
P04083	Annexin A1	ANXA1	40.2	151.37	2.93E+08	2.84E+06	1.25E+07
P35908	Keratin, type II cytoskeletal 2 epidermal	KRT2	25.5	118.83	2.20E+08	1.19E+08	1.62E+08
P07477	Trypsin-1	PRSS1	8.1	11.06	9.14E+07	9.35E+05	6.30E+06
P48668	Keratin, type II cytoskeletal 6C	KRT6C	26.8	106.72	1.05E+08	3.04E+07	4.65E+07
P01591	Immunoglobulin J chain	IGJ	7.5	94,959	0.00E+00	3.77E+06	3.09E+06
P19013	Keratin, type II cytoskeletal 4	KRT4	26.5	94,148	6.13E+07	2.78E+05	1.19E+06
P02533	Keratin, type I cytoskeletal 14	KRT14	28.8	93,395	1.78E+08	6.80E+07	9.36E+07
P31025	Lipocalin-1	LCN1	6.2	89,124	1.68E+06	4.39E+05	2.73E+05
P04792	Heat shock protein beta-1	HSPB1	8.3	87,612	2.20E+06	0.00E+00	7.36E+04
Q96DA0	Zymogen granule protein 16 homolog B	ZG16B	7.7	83,808	1.82E+06	1.43E+06	7.95E+06
P04080	Cystatin-B	CSTB	12.2	79,422	6.11E+06	1.43E+06	4.49E+06
P15924	Desmoplakin	DSP	4.2	78,899	3.28E+07	3.52E+06	5.99E+06
Q92817	Envoplakin	EVPL	5.9	74,045	1.44E+07	5.74E+05	3.02E+05
P25311	Zinc-alpha-2-glycoprotein	AZGP1	3.4	71,511	1.40E+06	2.16E+06	4.82E+06
Q96FQ6	Protein S100-A16	S100A16	11.7	71,024	6.61E+06	3.85E+05	3.86E+05
P07355	Annexin A2	ANXA2	2.9	70,778	3.52E+06	8.60E+05	1.51E+06
P19957	Elafin	PI3	8.5	66,777	1.84E+07	2.14E+07	2.42E+07
Q13835	Plakophilin-1	PKP1	5	36,197	1.57E+07	1.02E+06	1.39E+06
P14923	Junction plakoglobin	JUP	7	35,763	1.13E+07	7.10E+05	1.43E+06
P13646	Keratin, type I cytoskeletal 13	KRT13	19.2	32,911	3.34E+07	9.50E+05	2.61E+06
P08779	Keratin, type I cytoskeletal 16	KRT16	23.9	32,806	1.50E+07	7.95E+06	9.48E+06
P02788	Lactotransferrin	LTF	6.2	31,411	2.70E+06	7.97E+06	2.03E+07
P00450	Ceruloplasmin	CP	3.8	25,213	1.37E+07	8.43E+06	1.53E+07
P01876	Ig alpha-1 chain C region	IGHA1	11	25,078	2.44E+06	1.34E+07	2.22E+07
Q9UBC9	Small proline-rich protein 3	SPRR3	43.2	25,045	3.32E+07	4.00E+05	1.20E+06
Q9UBG3	Cornulin	CRNN	17.8	23,478	2.75E+07	1.18E+06	1.31E+06
P05109	Protein S100-A8	S100A8	23.7	21,429	1.73E+07	1.18E+06	5.84E+06
P62805	Histone H4	HIST1H4A	33	20,959	2.71E+07	3.14E+06	7.53E+06
P10412	Histone H1.4	HIST1H1E	15.5	19,715	1.05E+07	5.54E+05	3.31E+06
P35325	Small proline-rich protein 2B	SPRR2B	36.1	19,356	2.88E+07	5.94E+05	5.64E+05
Q9Y646	Carboxypeptidase Q	CPQ	4.7	18,965	2.16E+07	1.61E+07	2.35E+07
Q08188	Protein-glutamine gamma-glutamyltransferase	TGM3	3.5	14,731	4.26E+06	3.20E+06	5.58E+06
Q06830	Peroxiredoxin-1;Peroxiredoxin-2	PRDX1;PRDX2	10.6	14,399	4.28E+06	2.38E+04	8.18E+05
A8K2U0	Alpha-2-macroglobulin-like protein 1	A2ML1	1.7	14,234	9.20E+05	4.67E+06	6.00E+06
P01036	Cystatin-S;Cystatin-SN;Cystatin-SA	CST4;CST1;CST2	20.6	13,111	8.36E+05	2.72E+06	1.72E+07
P01834	Ig kappa chain C region	IGKC	34.6	12,985	2.55E+06	5.83E+06	1.38E+07
P68104	Elongation factor 1-alpha 1	EEF1A1	5	12,822	5.09E+06	3.79E+05	1.16E+05
P01857	Ig gamma-1 chain C region	IGHG1	6.7	12,503	1.02E+06	7.25E+05	4.58E+06
P09211	Glutathione S-transferase P	GSTP1	10	11,362	1.17E+06	1.50E+06	3.73E+06
P12273	Prolactin-inducible protein	PIP	17.8	10,784	6.18E+05	4.43E+05	4.67E+06
Q9Y577	E3 ubiquitin-protein ligase TRIM17	TRIM17	1.7	-2	1.66E+05	1.93E+05	1.40E+05
Q96NH3	Protein broad-minded	TBC1D32	0.8	-2	5.81E+06	4.98E+06	8.47E+06
Q9UL36	Zinc finger protein 236	ZNF236	0.5	-2	1.46E+06	5.61E+06	8.81E+06

Table 2.S3. Analysis of proteins identified in the *Aspergillus nidulans* secretomes after biotinylated probe **1d**-treatment, no probe control, and total secretome pull down experiments

					Raw abundance								
					Probe	Probe	Probe	No_Probe	No_Probe	No_Probe	Secretome	Secretome	Secretome
Accession	Abbreviation	Predicted Function	Unique peptides	Mass Da	D412_1	D412_2	D412_3	D412_4	D412_5	D412_6	D412_7	D412_8	D412_9
G5EAT0	AmyB	Amylase, GH13	14	69125	3.07E+07	5.08E+07	5.85E+07	1.69E+05	1.01E+05	1.30E+05	2.55E+06	2.57E+06	5.12E+06
Q5B7U2	AmyF	Amylase, GH13	11	50626	2.39E+06	4.75E+06	5.77E+06	9.93E+04	6.43E+04	5.48E+04	9.55E+05	1.09E+06	1.99E+06
G5EAZ3	ChiB	Endochitinase, GH18	24	44293	1.02E+06	1.42E+06	1.85E+06	1.48E+06	1.26E+06	9.78E+05	1.03E+08	1.29E+08	2.52E+08
G5EB45	AmyA	Amylase, GH13	5	54729	4.44E+05	8.52E+05	1.09E+06	1.34E+04	1.14E+04	3.66E+03	4.57E+04	3.33E+04	1.42E+05
G5EB11	AgdB	alpha-glucosidase, GH31	23	108005	4.27E+05	6.12E+05	8.62E+05	2.62E+05	1.70E+05	1.16E+05	1.21E+07	1.42E+07	2.72E+07
Q5B3Q5	-	Glycoside Hydrolase Family 55	25	97260	1.43E+05	2.61E+05	2.77E+05	1.58E+05	8.70E+04	6.12E+04	8.09E+06	1.18E+07	2.14E+07
Q5ATD5	-	Peptide hydrolase	8	54200	3.52E+04	5.16E+04	9.64E+04	3.95E+04	2.84E+04	2.24E+04	3.87E+06	5.30E+06	1.27E+07
Q5AUM3	abnC	Arabinan endo-1,5-alpha-L-arabinosidase, GH43	4	34580	3.65E+04	3.36E+04	5.03E+04	1.62E+04	2.81E+03	2.49E+04	1.63E+06	4.48E+06	3.50E+06
Q5AUI4	-	Extracellular lipase	2	31613	4.97E+03	3.86E+03	3.76E+03	7.22E+03	1.77E+03	1.29E+03	9.26E+04	1.44E+05	2.09E+05

					Spectral counts								
					Probe	Probe	Probe	No_Probe	No_Probe	No_Probe	Secretome	Secretome	Secretome
Accession	Abbreviation	Predicted Function	Unique peptides	Mass Da	D412_1_SC	D412_2_SC	D412_3_SC	D412_4_SC	D412_5_SC	D412_6_SC	D412_7_SC	D412_8_SC	D412_9_SC
G5EAT0	AmyB	Amylase, GH13	14	69125	34	40	41	0	0	0	16	13	17
Q5B7U2	AmyF	Amylase, GH13	11	50626	10	18	21	0	0	0	6	8	9
G5EAZ3	ChiB	Endochitinase, GH18	24	44293	1	1	4	0	1	0	71	87	102
G5EB45	AmyA	Amylase, GH13	5	54729	5	4	5	0	0	0	3	2	5
G5EB11	AgdB	alpha-glucosidase, GH31	23	108005	9	9	9	0	0	0	42	65	63
Q5B3Q5	-	Glycoside Hydrolase Family 55	25	97260	2	3	3	0	0	0	27	42	56
Q5ATD5	-	Peptide hydrolase	8	54200	2	2	1	0	0	0	10	17	21
Q5AUM3	abnC	Arabinan endo-1,5-alpha-L-arabinosidase, GH43	4	34580	2	1	2	0	0	0	5	8	5
Q5AUI4	-	Extracellular lipase	2	31613	1	1	2	0	0	0	4	2	4

Table 2.S4. α-Amylases previously identified in the starch secretomes of *Aspergillus nidulans*

Protein (name)	**CBM**	**Protein family**	**Uniprot**	**Molecular mass (kDa)***
α-amylase (AmyA)		GH13_1	G5EB45	52.6
α-amylase (AmyB)	CBM20	GH13_1	G5EAT0	66.7
α-amylase (AmyF)		GH13_1	Q5B7U2	48.3
*The mass is of the mature protein lacking the predicted signal peptide using Signal P (http://www.cbs.dtu.dk/services/SignalP/)				

2.S3 Supplemental References

[1] Moridani, M. Y.; Bromberg, I. L., *Clin. Biochem.* **2003,** *36* (1), 31-33.

[2] Ismail, O. Z.; Bhayana, V., *Clin. Biochem.* **2017,** *50* (18), 1275-1280.

[3] Rompianesi, G.; Hann, A.; Komolafe, O.; Pereira, S. P.; Davidson, B. R.; Gurusamy, K. S., *Cochrane Database Syst. Rev.* **2017,** *4*, CD012010.

[4] Ali Esmaili, H., *Biomed. & Pharmacol. J.* **2017,** *10* (1), 389-394.