Universiteit
Leiden
The Netherlands

Cyclophellitol analogues for profiling of exo- and endo-glycosidases

Schröder, S.P.

Citation
Schröder, S. P. (2018, May 17). *Cyclophellitol analogues for profiling of exo- and endo-glycosidases*. Retrieved from https://hdl.handle.net/1887/62362

Version:	Not Applicable (or Unknown)
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/62362

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/62362 holds various files of this Leiden University dissertation

Author: Schröder, Sybrin P.
Title: Cyclophellitol analogues for profiling of exo- and endo-glycosidases
Date: 2018-05-17

Chapter 7

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

7.1 Introduction

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

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

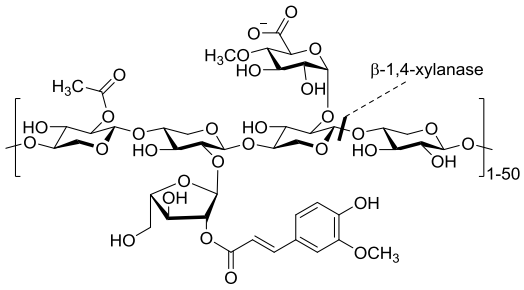


Figure 1 Chemical structure of hemicellulosic xylan. The polysaccharide contains up to 200 carbohydrates and is comprised of a β-1,4-xylan backbone, decorated with α-D-glucuronic acids and α-L-arabinofuranosyl residues carrying acetate, ferulate or coumarate esters. *Endo*-1,4-β-xylanases hydrolyze the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

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

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

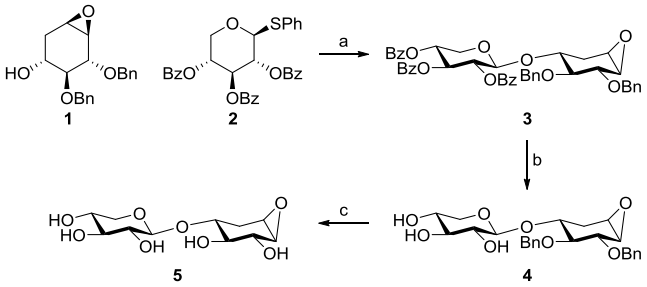
Endo-β-1,4-xylanases serve as catalysts in many industrial applications including food processing (fruit juicing, brewing, baking), textile production, the production of detergents, antioxidants and surfactants.[17] Additionally, a major application of these enzymes lies in the paper and pulp industry,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

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

7.2 Results and Discussion

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

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

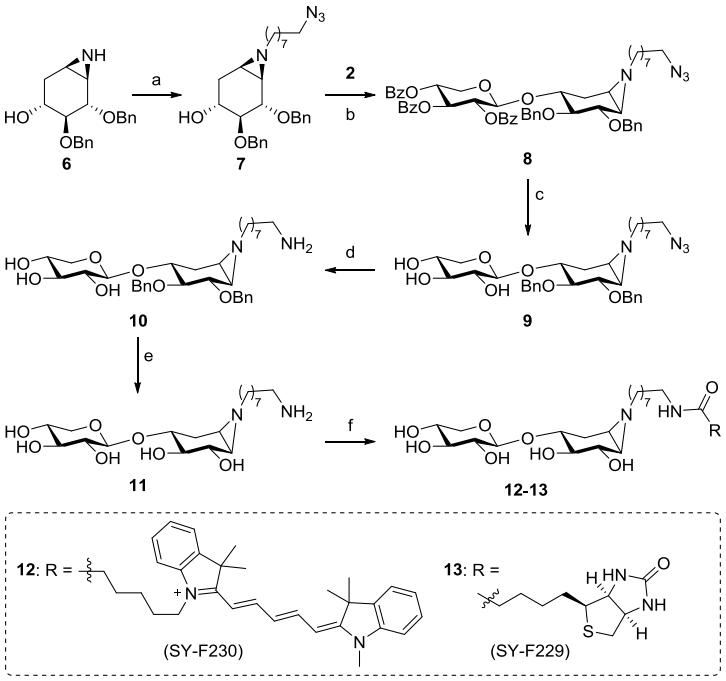


Scheme 1 Synthesis of xylobiose epoxide **5** by glycosylation of epoxide acceptor **1**. Reagents and conditions: a) NIS, TMSOTf (10 mol%), 4Å MS, DCM, -40 °C, 1 h, 75%; b) NaOMe, MeOH, DCM, 16 h, rt, 87%; c) $Pd(OH)_2/C$, H_2, MeOH, H_2O, dioxane, 4 h, rt, quant.

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

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

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



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

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

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

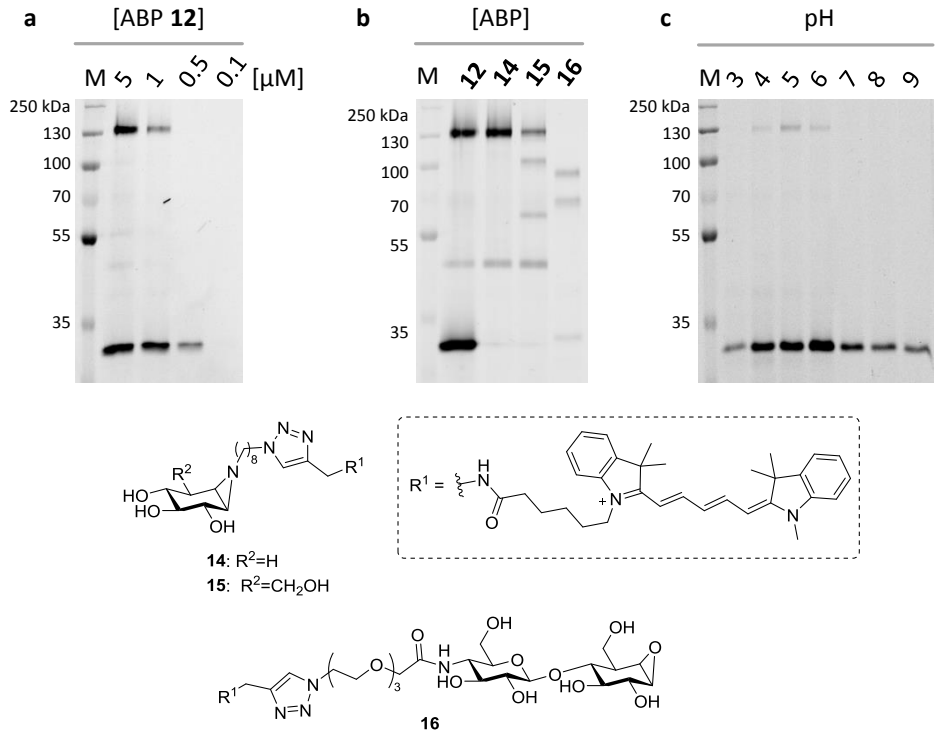
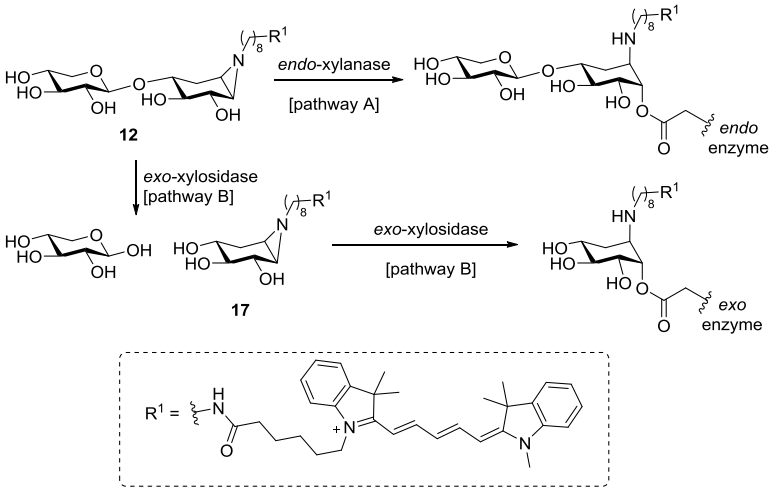


Figure 2 (a) Fluorescent labeling of *A. niger* secretome grown on xylan (1%, 8 days) with different concentrations of *xylobiose*-cyclophellitol ABP **12**. (b) Labeling of this secretome with **14** (5 µM) suggests that the bands at ~130 kDa and ~45 kDa are possibly *exo*-xylosidases. These bands are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

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

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

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



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

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

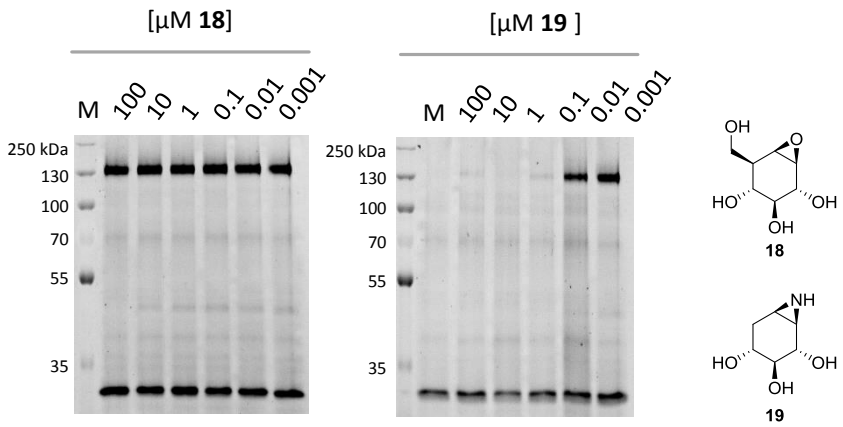


Figure 3 Competition assay of *A. niger* secretome. The sample was pre-incubated with **18** or **19** at different concentrations (30 min, 37 °C, pH 6) and then labeled with **12** (5 μM, 30 min, 37 °C). Pre-incubation with **18** did not decrease labeling with **12** at any concentration, whereas **19**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

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

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

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

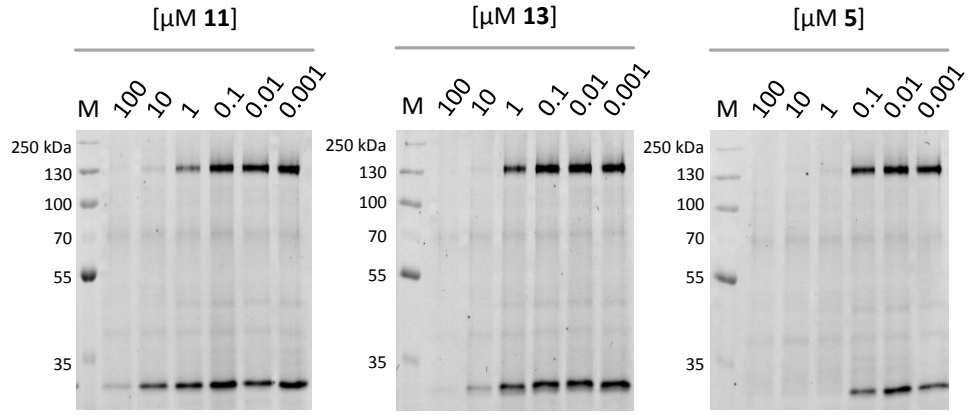


Figure 4 Competition assay of *A. niger* secretome. The sample was pre-incubated with **11**, **13** and **5** at different concentrations (30 min, 37 °C) and then labelled with **12** (5 μM, 30 min, 37 °C, pH 6). Both **11** and **13** are weak inactivators of *endo*-xylanase, moreover the *exo*-xylosidase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

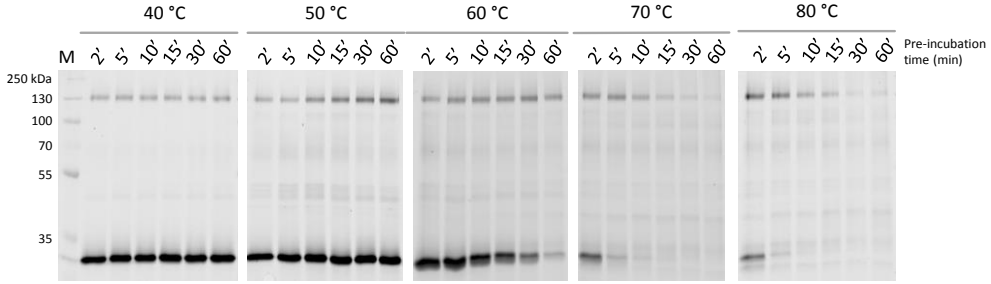


Figure 5 Monitoring of *endo*-xylanase activity over time at different temperatures. The enzyme retains full activity after incubation at 50 °C for 60 minutes. At 60 °C, the enzyme half-life is approximately 15 minutes and nearly full deactivation is reached after 60 minutes. At 70-80°C, nearly all enzyme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

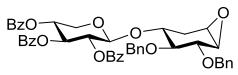
7.3 Conclusion

Up to date, all cyclophellitol (aziridine) based glycosidase ABPs reported in literature are monomeric in structure and therefore could only be utilized for the monitoring of *exo*-glycosidases. While Withers and co-workers have published the chemoenzymatic synthesis of α-1,4-glucosyl *epi*-cyclophellitol as covalent irreversible α-amylase inhibitor,[45] this approach was certainly not trivial and enzymatic tolerance of substrates equipped with a reporter tag might pose a major obstacle for the chemoenzymatic synthesis of *endo*-glycosidase probes. In this Chapter, a new straightforward synthetic route towards glycosylated *xylo*-cyclophellitols is described, which relies on the direct glycosylation of *xylo*-cyclophellitol (aziridine) acceptors with an appropriate thioglycoside donor under NIS/TMSOTf activating conditions. The epoxide and aziridine functionalities were tolerant towards the Lewis acidic reaction conditions, and glycosylation products could be obtained in good yields. Labeling of the secretome of *Aspergillus niger* (grown on xylan) by the *xylobiose*-cyclophellitol aziridine fluorescent probe identified a ~30 kDa band on gel, which is presumed to be the *endo*-xylanase XlnC. The optimal buffer pH for labeling this enzyme was pH 6.0, and the enzyme appeared stable up to 50 °C for 60 minutes; at higher temperatures the activity was abolished after this time period. The probe also identified the presence of a ~130 kDa protein in the secretome, and by competition assays with appropriate *exo*-glycosidase inhibitors, this band is presumed to be an *exo*-xylosidase. It is hypothesized that following incubation with the *xylobiose*-cyclophellitol ABP, the *exo*-xylosidase is labeled by an *exo*-xylosidase probe generated *in situ* by *exo*-hydrolysis of the *xylobiose*-cyclophellitol ABP. Competition assays indicated that *xylobiose*-cyclophellitol alkyl aziridines are weak inactivators of the *endo*-xylanase, and due to the inherent instability of the *endo*-xylanase probe towards hydrolysis, that these inhibitors are processed by the *exo*-xylosidase at higher rate. *Xylobiose*-cyclophellitol (epoxide) is a more potent inactivator of the *endo*-xylanase compared to the alkyl aziridine analogues, in contrast to observations for *exo*-xylosidase inhibitors. Identification of the labeled enzymes may be achieved using *xylobiose*-cyclophellitol aziridine equipped with a biotin tag, via protein enrichment techniques followed by LC-MS/MS identification. Lastly, the instability of *endo*-cyclophellitol probes towards *exo*-hydrolysis complicates the selective profiling of *endo*-glycosidases in biologically complex samples such as presented here, therefore the synthesis of stabilized *endo*-probes would be of interest (see Chapter 8).

Experimental procedures

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

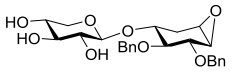
Compound 3



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

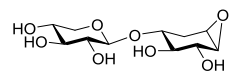
3.50 (dd, *J* = 12.0, 8.5 Hz, 1H), 3.42 (dd, *J* = 10.0, 7.7 Hz, 1H), 3.16 (br s, 1H), 3.09 (d, *J* = 3.6 Hz, 1H), 2.49 (ddd, *J* = 14.5, 5.0, 2.1 Hz, 1H), 1.72 – 1.63 (ddd, *J* = 14.5, 10.5, 1.0 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): 165.8, 165.6, 165.2, 138.8, 137.8, 133.5, 133.5, 133.4, 130.3,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

Compound 4



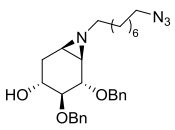
Compound **3** (115 mg, 0.15 mmol) was dissolved in a mixture of dry DCM (0.75 mL) and MeOH (1.5 mL). Then, NaOMe in MeOH (5.4 M, 11 µL, 40 mol %) was added and the mixture was stirred 16h at rt. The mixture was neutralized with $Et_3N.HCl$ (10 mg, 0.075 mmol) and concentrated. Flash column chromatography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

Compound 5



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

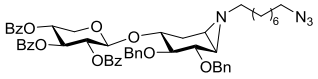
Compound 7



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

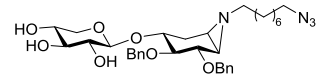
concentrated. Flash purification by silica column chromatography (DCM/MeOH, 99:1) gave the title compound as an oil (204 mg, 86%). ^{1}H NMR (400 MHz, $CDCl_3$): 7.42 – 7.22 (m, 10H), 4.96 (d, *J* = 11.4 Hz, 1H), 4.77 (d, *J* = 11.5 Hz, 1H), 4.65 (dd, *J* = 11.4, 7.3 Hz, 2H), 3.74 (d, *J* = 8.0 Hz, 1H),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

Compound 8



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

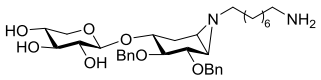
Compound 9



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

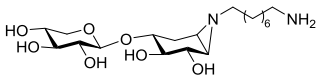
11.4, 7.2 Hz, 1H), 2.00 (dt, *J* = 11.6, 7.1 Hz, 1H), 1.71 (d, *J* = 5.6 Hz, 1H), 1.67 (d, *J* = 13.4 Hz, 1H), 1.58 (dt, *J* = 14.4, 6.8 Hz, 2H), 1.52 (d, *J* = 6.0 Hz, 1H), 1.50 – 1.41 (m, 2H), 1.32 (d, *J* = 23.0 Hz, 8H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): 138.5, 138.1, 128.6, 128.4, 128.1, 128.0,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

Compound 10



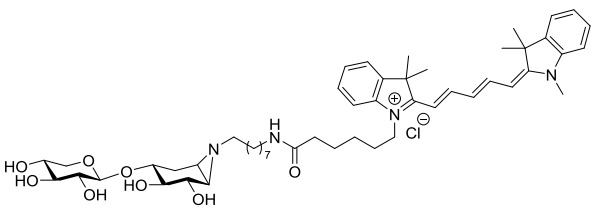
Compound **9** (111 mg, 0.19 mmol) was dissolved in MeCN (3.8 mL), polymer-bound triphenylphosphine (3 mmol/g loading, 126 mg, 0.38 mmol) and H_2O (34 µL, 1.9 mmol) were added and the mixture was stirred overnight at 70 °C. TLC indicated total consumption of the starting material, and additional H_2O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

Compound 11



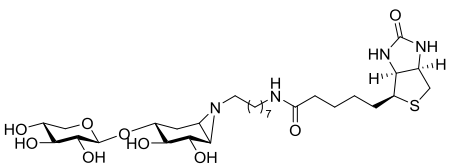
Ammonia (1 mL) was condensed in a flask at -60 °C, and lithium wire (13 mg, 1.86 mmol) was added. The resulting deep-blue solution was stirred for 30 minutes to dissolve all lithium. Aziridine **10** (30 mg, 49 µmol) was taken up in dry THF (1 mL) and added to the reaction mixture. After stirring for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

Compound 12 (SY-F230)



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

Compound 13 (SY-F229)



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

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

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

References

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

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

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

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

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

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

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

26 M. D. Witte, W. W. Kallemeijn, J. Aten, K.-Y. Li, A. Strijland, W. E. Donker-Koopman, A. M. C. H. van den Nieuwendijk, B. Bleijlevens, G. Kramer, B. I. Florea, B. Hooibrink, C. E. M. Hollak, R. Ottenhoff, R. G. Boot, G. A. van der Marel, H. S. Overkleeft and J. M. F. G. Aerts, *Nat. Chem. Biol.*, 2010, **6**, 907–913.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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