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

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

Chapter 3

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

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

3.1 Introduction

Glycosidases are hydrolytic enzymes that catalyze the hydrolysis of interglycosidic linkages in oligosaccharides and glycoconjugates.[1] Glycosidases are involved in a broad array of human pathologies and are therefore promising therapeutic targets.[2] Activity-based protein profiling (ABPP) provides an effective tool for the analysis of enzyme composition, expression and activity, and can assist in the discovery of enzyme inhibitors.[3] ABPP relies on the addition of a mechanism-based inhibitor to the sample of interest, which reacts with the catalytic machinery of the target enzyme in a covalent and irreversible manner. The inhibitor carries a reporter tag and is termed an activity-based probe (ABP). As the first example on the development of ABPP, Cravatt and co-workers have described an irreversible fluorophosphonate based ABP **1** (Figure 1a), which displays high reactivity towards a wide range of serine hydrolases (SH).[4] ABP **1** was used to reveal active SH composition and expression level in various tissues, and as well in the development of highly specific and potent inhibitors against, for instance, monoacylglycerol lipase in a competitive ABPP setting.[5] Cyclophellitol (**2**, Figure 1b), first isolated from *Phellinus* sp. is a covalent irreversible retaining β-glucosidase inhibitor.[6] The configurational isomers, *galacto*-cyclophellitol (**2**)[7] and *manno*-cyclophellitol (**3**)[8] are inactivators for β-galactosidases and β-mannosidases, respectively, and their selectivity arises from their absolute configurations. Several cyclophellitol based ABPs for selective glycosidase profiling have been published.[9–11] For example, treatment of mouse tissue lysates with acyl ABP **5** (Figure 1c) resulted in tissue-specific fluorescent labelling of the three retaining β-glucosidases, GBA1, 2 and 3 and lactase-phlorizin hydrolase (LPH) without non-specific cross-labelling. Later studies revealed that alkyl ABP **6** also effectively labels these retaining β-glucosidases.[12] While the selectivity of these probes is an attractive value, it also impedes the study of multiple enzyme classes with a single probe, as for

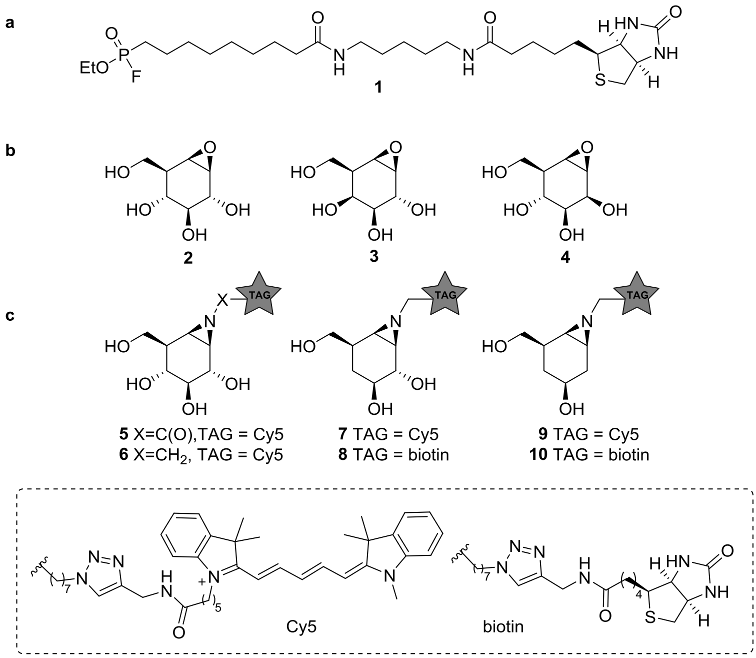


Figure 1 (a) Irreversible fluorophosphonate ABP **1** for serine hydrolases. (b) Structures of cyclophellitols **2**-**4**, selective covalent and irreversible retaining glycosidase inhibitors. (c) Selective β-glucosidase ABPs **5**-**6**, 4-deoxy ABPs **7**-**8** and 2,4-deoxy ABPs **9**-**10**.

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

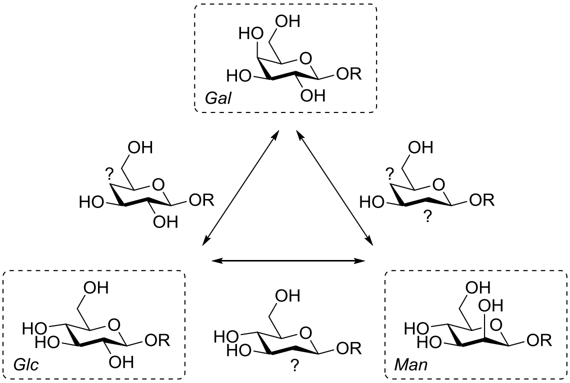
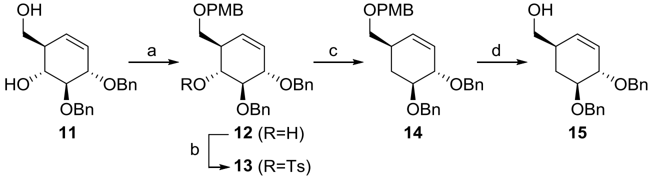


Figure 2 β-Galactosides (Gal), β-glucosides (Glc) and β-mannosides (Man) are distinguished by the configuration of the hydroxyl group(s) at C-2 and/or C-4.

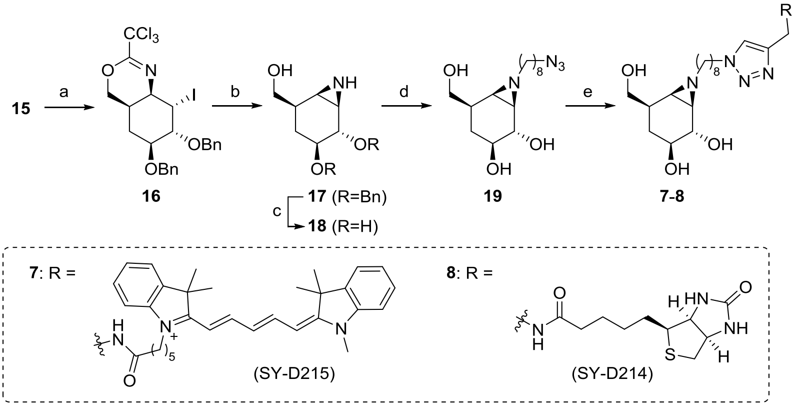
3.2 Results and Discussion

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



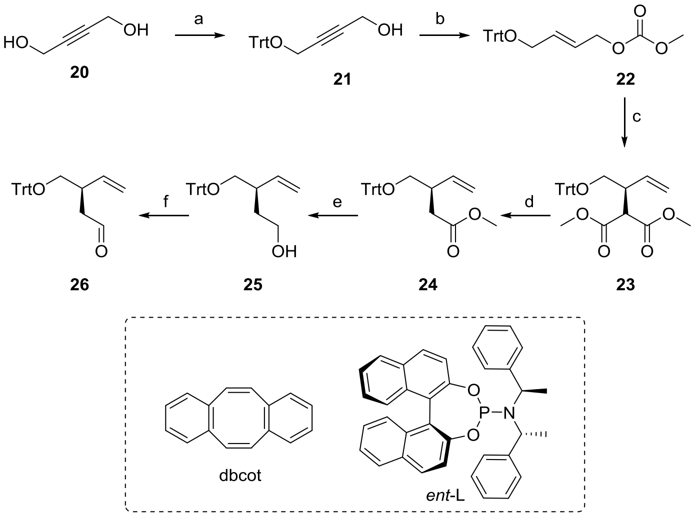
Scheme 1 Deoxygenation of cyclohexene **11**. Reagents and conditions: a) PMB-Cl, 2-aminoethyl diphenylborinate (10 mol%), K_2CO_3, KI, MeCN, 60 °C, 4h, 95%; b) TsCl, pyridine, 60 °C, 16h, 90%; c) $LiAlH_4$, THF, reflux, 16h, 86%; d) HCl (10 mol%), DCM/HFIP, 15 min, quant.

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



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

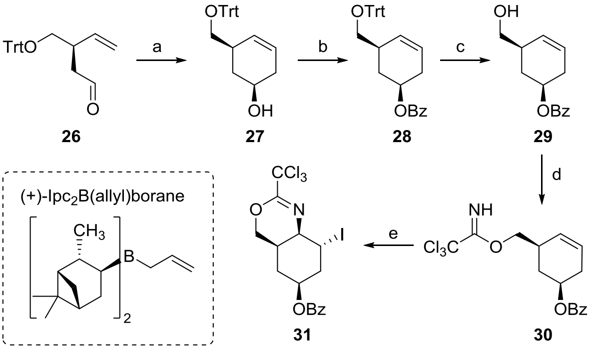
Synthesis of β-2,4-deoxy ABPs - The first steps in the synthesis route were adapted from a paper by Helmchen *et al.*[16] Commercially available diol **20** was tritylated to afford **21** (Scheme 3). The alkyne functionality was reduced by lithium aluminum hydride and the crude allylic alcohol was subsequently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



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

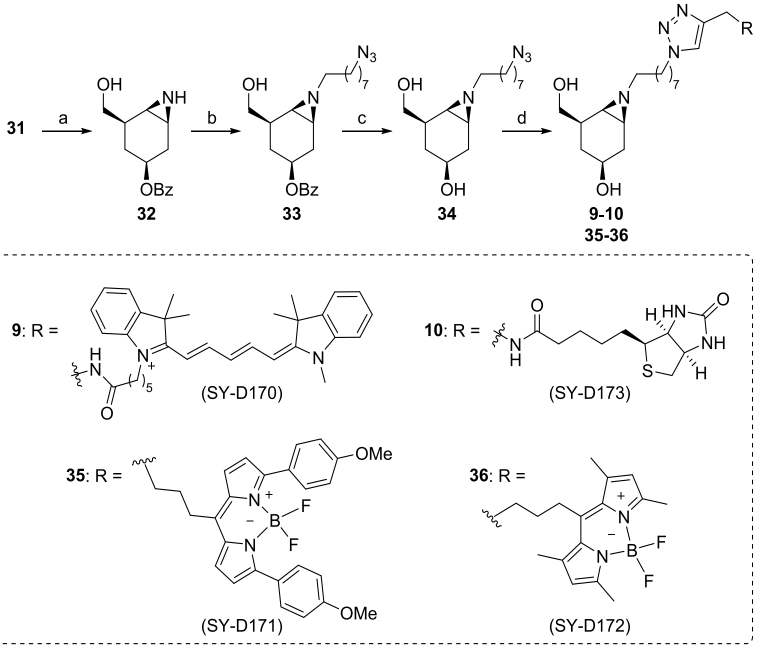
Krapcho decarboxylation. Due to the high reaction temperatures employed, it was found critical to thoroughly degass the reaction mixture beforehand, as well as addition of a radical inhibitor to prevent side-product formation. Reduction of the ester yielded alcohol **25**, which was oxidized to aldehyde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

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



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

Next, cyclic imidate **31** was hydrolyzed under acidic conditions, and subsequent treatment of the resulting ammonium salt with excess base resulted in cyclization to give aziridine **32** (Scheme 5). This compound proved unstable at elevated (~40 °C) temperature, and concentration of the compound after work-up and column purification had to be performed on an ice-bath. The aziridine was alkylated with freshly prepared 8-azidooctyl trifluoromethanesulfonate, affording **33** in moderate yield. Final deprotection under Zémplen conditions resulted in isolation of **34**, which was equipped with various reporter tags via azide/alkyne click ligation to give 2,4-deoxy ABPs **9-10** and **35-36.**



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

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

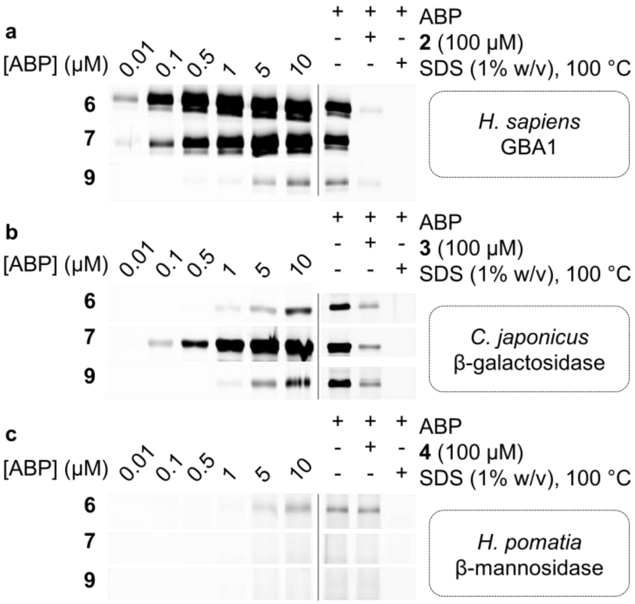


Figure 3 Fluorescent labelling of recombinant *H. sapiens* GBA1, *C. japonicus* β-galactosidase and *H. pomatia* β-mannosidase with ABPs **6**, **7** and **9** (left panel), and competition experiments with selective cyclophellitol inhibitors **2**, **3** or **4** (right panel). (a) GBA1 is potently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

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

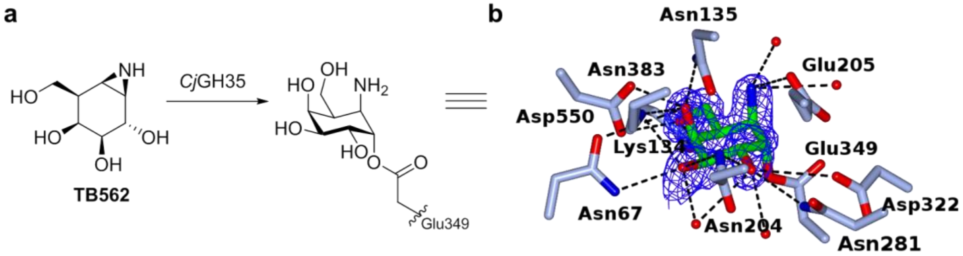


Figure 4 (a) Schematic depiction of g*alacto*-cyclophellitol aziridine **TB562** reacting with the active-site carboxylate of β-galactosidase *Cj*GH35. (b) Structure of **TB562** bound in the *Cj*GH35 active site, obtained by X-ray crystallography (PDB: 5JAW). The catalytic nucleophile (Glu349) is covalently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

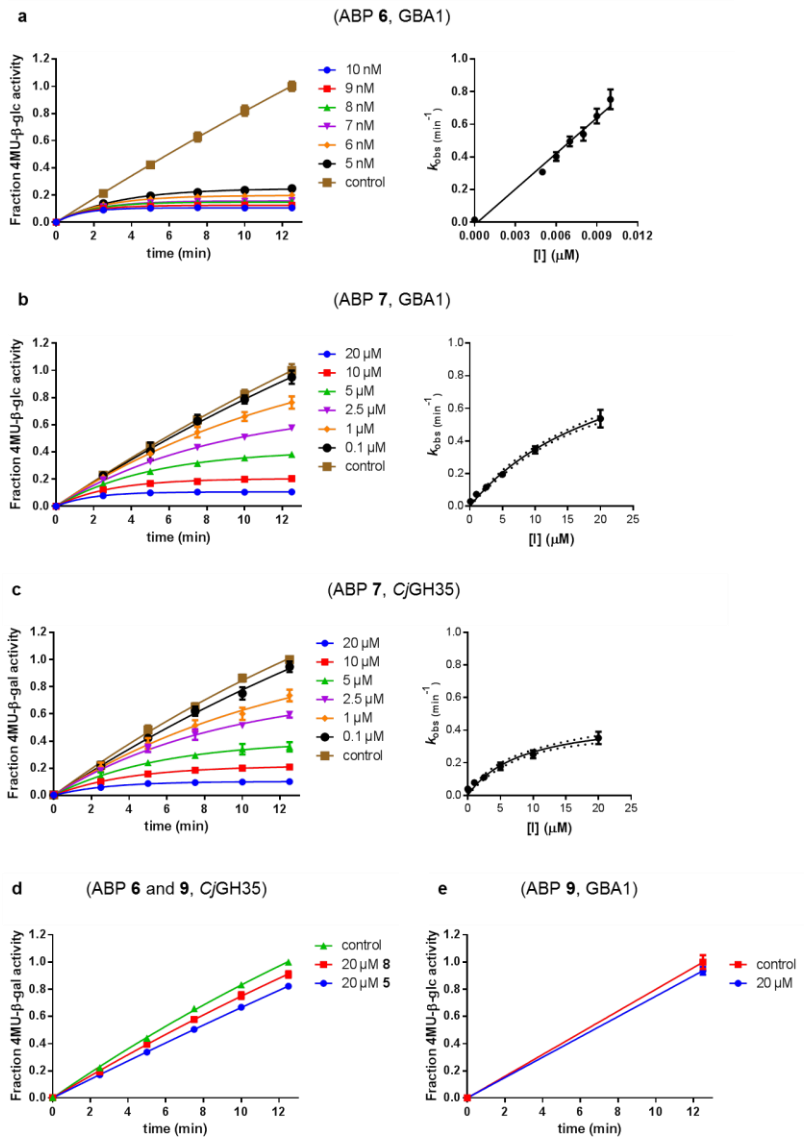


Figure 5 Kinetic plots of GBA1 and *Cj*GH35 inhibition by **6**, **7** and **9** with 4-methylumbelliferyl β-D-glucopyranoside and β-D-galactopyranoside substrates. (a-c) Left: plots of relative substrate activity versus time at a series of different concentrations inhibitor. Right: pseudo first-order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

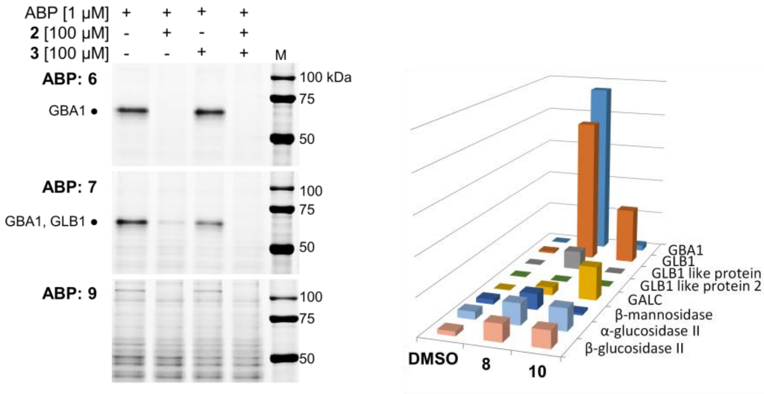
Next, glycosidase profiling by probes **7** and **9** in lysates from mouse (C57bl/6j, Jackson's laboratories) kidney and liver tissue, and human wild-type fibroblast cells (Lonza) was investigated, and compared to that of β-glucosidase selective probe **6**. The samples were incubated with the ABPs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

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

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

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

a Mouse liver



b Mouse kidney

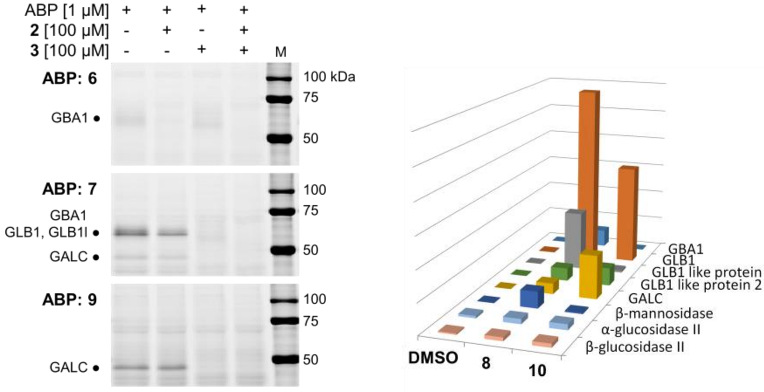


Figure 6 Fluorescent labelling of mouse liver (a) and kidney (b) lysates with probes **6**, **7** and **9**, complemented with competition experiments with β-glucosidase specific covalent inhibitor **2** and β-galactosidase specific **3**. The bar charts display the LC-MS/MS pull-down analyses with the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

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

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

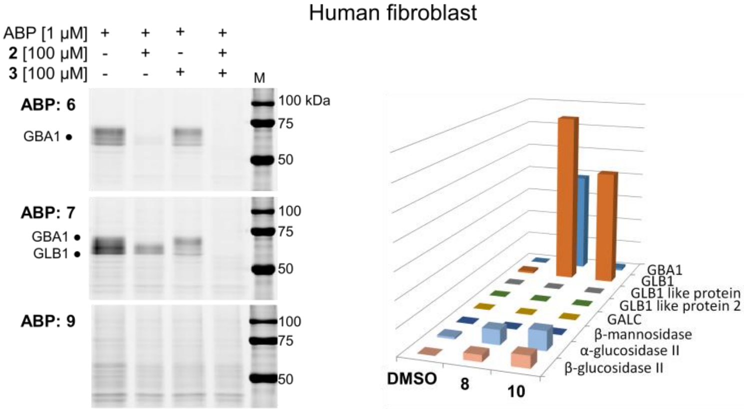


Figure 7 Fluorescent labelling of human fibroblast lysates with probes **6, 7** and **9**, complemented with competition experiments with β-glucosidase specific covalent inhibitor **2** and β-galactosidase specific **3**. The bar charts display the LC-MS/MS pull-down analyses with the biotinylated ABP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

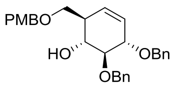
3.3 Conclusion

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

Experimental procedures

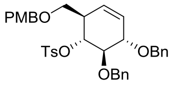
General: Chemicals were purchased from Acros, Sigma Aldrich, Biosolve, VWR, Fluka, Merck and Fisher Scientific and used as received unless stated otherwise. Tetrahydrofuran (THF), *N,N*-dimethylformamide (DMF) and toluene were stored over molecular sieves before use. Traces of water from reagents were removed by co-evaporation with toluene in reactions that required anhydrous conditions. All reactions were performed under an argon atmosphere unless stated otherwise. TLC analysis was conducted using Merck aluminum sheets (Silica gel 60 F_{254}) with detection by UV absorption (254 nm), by spraying with a solution of $(NH_4)_6Mo_7O_{24}\cdot4H_2O$ (25 g/L) and $(NH_4)_4Ce(SO_4)_4\cdot2H_2O$ (10 g/L) in 10% sulfuric acid or a solution of $KMnO_4$ (20 g/L) and K_2CO_3 (10 g/L) in water, followed by charring at ~150 ˚C. Column chromatography was performed using Screening Device b.v. silica gel (particle size of 40 – 63 μm, pore diameter of 60 Å) with the indicated eluents. For reversed-phase HPLC purifications an Agilent Technologies 1200 series instrument equipped with a semi-preparative column (Gemini C18, 250 x 10 mm, 5 μm particle size, Phenomenex) was used. LC/MS analysis was performed on a Surveyor HPLC system (Thermo Finnigan) equipped with a C_{18} column (Gemini, 4.6 mm x 50 mm, 5 μm particle size, Phenomenex), coupled to a LCQ Adventage Max (Thermo Finnigan) ion-trap spectrometer (ESI^+). The applied buffers were H_2O, MeCN and 1% aqueous TFA. ^1H NMR and ^{13}C NMR spectra were recorded on a Brüker AV-400 (400 and 101 MHz respectively) or a Brüker DMX-600 (600 and 151 MHz respectively) spectrometer in the given solvent. Chemical shifts are given in ppm (δ) relative to the residual solvent peak or tetramethylsilane (0 ppm) as internal standard. Coupling constants are given in Hz. High-resolution mass spectrometry (HRMS) analysis was performed with a LTQ Orbitrap mass spectrometer (Thermo Finnigan), equipped with an electronspray ion source in positive mode (source voltage 3.5 kV, sheath gas flow 10 mL/min, capillary temperature 250 °C) with resolution R = 60000 at m/z 400 (mass range m/z = 150 – 2000) and dioctyl phthalate (m/z = 391.28428) as a "lock mass". The high-resolution mass spectrometer was calibrated prior to measurements with a calibration mixture (Thermo Finnigan). *Cj*GH35 enzyme was kindly provided by Prof. dr. Gideon Davies, University of York, U.K.

Compound 12



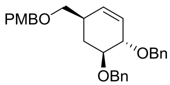
Diol **10** (1.36 g, 4.0 mmol) was dissolved in dry MeCN (20 mL), then PMBCl (0.81 mL, 6.0 mmol), KI (664 mg, 4.0 mmol), K_2CO_3 (608 mg, 4.4 mmol) and 2-aminoethyl diphenylborinate (90 mg, 0.4 mmol) were added and the mixture was stirred at 60 °C for 4 h. Then, the mixture was diluted with EtOAc (200 mL), washed with H_2O (2 x 100 mL) and brine, dried over $MgSO_4$, filtrated and concentrated. Flash purification by silica column chromatography (pentane/EtOAc, 4:1) gave the title compound as a colorless oil (1.75 g, 95%). ^1H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.26 (m, 10H), 7.24 (d, *J* = 8.6 Hz, 2H), 6.87 (d, *J* = 8.6 Hz, 2H), 5.73 (dt, *J* = 10.2, 2.3 Hz, 1H), 5.60 (d, *J* = 10.2 Hz, 1H), 4.99 (d, *J* = 11.3 Hz, 1H), 4.79 (d, *J* = 11.3 Hz, 1H), 4.67 (q, *J* = 11.5 Hz, 2H), 4.46 (s, 2H), 4.21 – 4.17 (m, 1H), 3.79 (s, 3H), 3.74 – 3.61 (m, 2H), 3.61 – 3.51 (m, 2H), 2.95 (brs, OH), 2.57 – 2.48 (m, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 159.3, 138.8, 138.4, 130.3, 129.4, 128.6, 128.6, 128.3, 128.1, 128.0, 127.9, 127.8, 126.8, 113.9, 84.0, 80.3, 75.0, 73.1, 71.7, 71.4, 70.9, 55.4, 44.1 ppm. HRMS (ESI): *m/z* = $[M+Na]^+$ calc for $C_{29}H_{32}O_5$ 483.21420, found 483.21368.

Compound 13



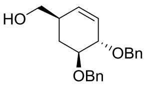
Cyclohexene **12** (2.54 g, 5.51 mmol) was co-evaporated with toluene and dissolved in pyridine (55 mL). Tosyl chloride (10.5 g, 55.1 mmol) was added in portions over 10 minutes, and then the mixture was stirred at 60 °C for 16 h. The mixture was cooled to 0 °C, quenched with H_2O (20 mL) and diluted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

Compound 14



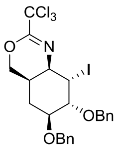
Tosylate **13** (3.07 g, 5.0 mmol) was co-evaporated with toluene (3x) and subsequently dissolved in dry THF (100 mL). The mixture was cooled to 0 °C and $LiAlH_4$ (2.4 M in THF, 15.6 mL, 37.4 mmol) was added and the mixture was refluxed overnight. After cooling to 0 °C, the reaction was quenched by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

Compound 15



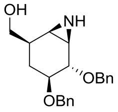
Compound **14** (1.91 g, 4.29 mmol) was dissolved in a mixture of DCM/HFIP (1:1 v/v, 43 mL), and aq. HCl (12 M, 36 µL, 0.43 mmol) was added. The mixture was stirred for 15 minutes and was subsequently quenched by addition of sat. aq. $NaHCO_3$ (10 mL). The mixture was diluted with DCM (200 mL) and washed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

Compound 16



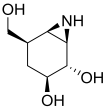
Compound **15** (1.48 g, 4.56 mmol) was dissolved in DCM (23 mL). Trichloroacetonitrile (687 µL, 6.85 mmol) and DBU (68 µL, 0.46 mmol) were added and the mixture was stirred overnight at rt. The mixture was concentrated, and filtered over a small silica plug. Concentration of the eluate afforded an oil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

Compound 17



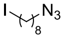
Compound **16** (35 mg, 59 µmol) was dissolved in MeOH/DCM (1:1, 1.2 mL) and HCl (1.25 M in MeOH, 443 µL, 0.55 mmol) was added. The mixture was stirred 24 h, and was subsequently neutralized by addition of Amberlite IRA-67. After stirring overnight, the reaction mixture was filtered and concentrated. Flash purification by silica column chromatography (DCM/MeOH, 49:1 → 19:1) gave the title compound as a colorless oil (16 mg, 80%). ^1H NMR (400 MHz, $CDCl_3$) δ 7.43 – 7.21 (m, 10H), 4.77 (s, 2H), 4.65 (d, J = 4.0 Hz, 2H), 3.76 (dd, J = 10.4, 4.8 Hz, 1H), 3.69 (dd, J = 10.5, 6.1 Hz, 1H), 3.65 (d, J = 8.0 Hz, 1H), 3.43 (ddd, J = 11.9, 8.0, 3.7 Hz, 1H), 2.40 (dd, *J* = 5.7, 2.7 Hz, 1H), 2.28 (d, *J* = 6.0 Hz, 1H), 2.16 – 2.05 (m, 1H), 1.65 (dt, *J* = 12.9, 4.2 Hz, 1H), 1.27 (q, *J* = 12.6 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.9, 138.5, 128.5, 128.4, 127.9, 127.8, 127.7, 127.6, 80.5, 80.1, 72.8, 71.6, 66.0, 36.4, 32.8, 32.3, 24.0 ppm. TLC-MS (ESI) m/z: $[M+H]^+$ calc for $C_{21}H_{26}NO_3$ 340.18, found 340.3.

Compound 18



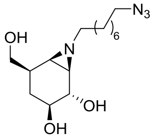
Ammonia (3 mL) was condensed in a flask at -60 °C, and lithium wire (7 mg, 0.94 mmol) was added. The resulting deep-blue solution was stirred for 30 minutes to dissolve all lithium. Aziridine **17** (16 mg, 47 µmol) was taken up in dry THF (1 mL) and added to the reaction mixture. After stirring for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

8-azido-1-iodooctane



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

Compound 19

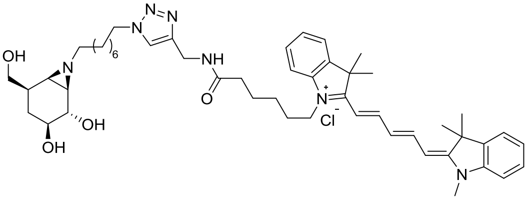


Starting from benzyl protected aziridine **17** (97 mg, 0.29 mmol), the compound was deprotected as described above and purified by elution over Amberlite CG-50 (NH_4^+). After evaporation of the solvent the intermediate was re-dissolved in DMF (1 mL), 1-azido-8-iodooctane (161 mg, 0.57 mmol) and K_2CO_3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

General procedure for click reactions

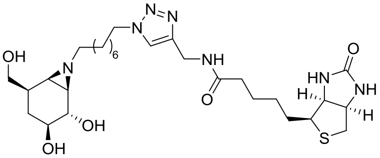
The azido compound (3-5 mg) was dissolved in DMF (0.5 mL), then the alkyne-tag (1.1 eq), $CuSO_4$ (0.2 eq) and sodium ascorbate (0.4 eq) were added and the mixture was stirred for 72 h at rt. The reaction mixture was concentrated and purified by semi-preparative reversed phase HPLC (linear gradient.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

Compound 7 (SY-D215)



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

Compound 8 (SY-D214)

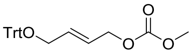


Following the general procedure, reaction of compound **18** (5.7 mg, 18.3 µmol) with biotin-alkyne[30] afforded the title compound as a white solid (5.3 mg, 49%). ^{1}H NMR (600 MHz, D_2O) δ 7.89 (s, 1H), 4.61 (dd, *J* = 7.9, 4.9 Hz, 1H), 4.45 (d, *J* = 6.8 Hz, 2H), 4.43 – 4.35 (m, 3H), 3.66 (dd, *J*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

Compound 21

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

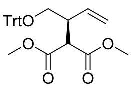
Compound 22



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

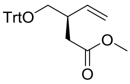
g, 74%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.52 – 7.36 (m, 6H), 7.35 – 7.12 (m, 9H), 5.99 (dtt, *J* = 15.1, 5.8, 1.4 Hz, 1H), 5.88 (dt, *J* = 15.6, 4.5 Hz, 1H), 4.65 (dd, *J* = 6.0, 1.1 Hz, 2H), 3.77 (s, 3H), 3.68 – 3.60 (m, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 155.7, 144.1, 132.4, 128.6, 127.9, 127.1,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

Compound 23



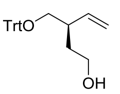
$[Ir(dbcot)Cl]_2$ (86 mg, 0.1 mmol)[17] and *ent*-L (120 mg, 0.2 mmol) were added to a flame-dried flask under argon, then dry THF (5 mL) was added. To this mixture was added triazabicyclodecene (56 mg, 0.4 mmol) and the mixture was stirred for 15 minutes at rt. A dry solution of compound **22** (1.94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

Compound 24



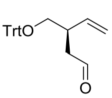
A solution of compound **23** (4.4 g, 9.9 mmol), sodium chloride (867 mg, 14.9 mmol), water (1.1 mL, 61.4 mmol) and BHT (110 mg, 0.5 mmol) in DMSO (50 mL) was put under vacuum and flushed with argon. The mixture was sonicated for 30 minutes under an argon flow and then heated at 150 °C for 22 h. The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

Compound 25



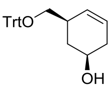
To a solution of compound **24** (18.5 g, 47.8 mmol) in dry THF (240 mL) was added LiBH$_4$ (5.2 g, 239 mmol) in portions at 0 °C and the mixture was stirred for 3 days at rt. The mixture was cooled on ice, quenched with sat. aq. NH$_4$Cl, diluted with water (500 mL) and brine (100 mL), and extracted with EtOAc (3x 150 mL). The combined organic layers were washed with brine, dried over MgSO$_4$, filtrated and concentrated. The product was purified by flash chromatography (pentane/EtOAc; 6:1 → 5:1) to afford the product as a colourless oil (16.7 g, 98%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.48 – 7.40 (m, 6H), 7.33 – 7.19 (m, 9H), 5.73 (ddd, J = 17.2, 10.3, 8.6 Hz, 1H), 5.13 – 5.01 (m, 2H), 3.61 (m, 2H), 3.10 (dd, J = 9.0, 5.7 Hz, 1H), 3.03 (dd, J = 9.0, 6.4 Hz, 1H), 2.44 (m, 1H), 1.78 (m, 1H), 1.65 – 1.51 (m, 2H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 144.3, 140.1, 128.8, 127.9, 127.1, 116.1, 86.7, 67.1, 61.2, 34.8. $[\alpha]_D^{20}$ -15.0 (c 0.4, DCM). IR (ATR, cm^{-1}) 3350, 1489, 1448, 1064. HRMS: [M+H]$^+$ calc for C$_{25}$H$_{26}$O$_2$ 359.20110 found 359.24025.

Compound 26



Compound **25** (2.65 g, 7.4 mmol) was co-evaporated three times with toluene and dissolved in dry DCM (15 mL) under argon. To the mixture was added TEMPO (231 mg, 1.48 mmol) and PhI(OAc)$_2$ (2.4 g, 7.4 mmol), and the mixture was stirred for 5 h. The reaction was quenched with aq. Na$_2$S$_2$O$_3$ (10 mL), diluted with water (300 mL), and extracted with DCM (3 x 150 mL). The combined organic layers were washed with sat. aq. NaHCO$_3$ and water, dried over MgSO$_4$, filtrated and concentrated. The product was purified by flash chromatography (pentane/EtOAc; 50:1 → 20:1) to afford an orange oil (2.3 g, 87%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.70 (s, 1H), 7.41 (m, 6H), 7.36 – 7.14 (m, 9H), 5.83 – 5.64 (m, 1H), 5.13 – 4.97 (m, 2H), 3.19 (dd, J = 9.0, 5.3 Hz, 1H), 3.03 (dd, J = 9.0, 7.2 Hz, 1H), 2.96 – 2.84 (m, 1H), 2.62 (ddd, J = 16.5, 5.8, 2.0 Hz, 1H), 2.43 (ddd, J = 16.5, 7.9, 2.3 Hz, 1H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 202.1, 144.1, 138.0, 128.8, 128.0, 128.0, 127.9, 127.4, 127.2, 116.6, 86.9, 66.2, 45.6, 38.9. $[\alpha]_D^{20}$ -14.3 (c 0.8, DCM). IR (ATR, cm^{-1}) 1724, 1490, 1448, 1068. HRMS: [M+H]$^+$ calc for C$_{25}$H$_{24}$O$_2$ 357.18546 found 357.15177.

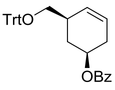
Compound 27



Aldehyde **26** (600 mg, 1.68 mmol) was dissolved in dry THF (20 mL) and cooled to -90 °C. A solution of 1M (+)-ipc$_2$B(allyl)borane solution in pentane was added to dry THF (10 mL), cooled to -90°C, and added dropwise via cannula. The mixture was stirred for 1 h, quenched with MeOH (6 mL) and diluted with sodium-phosphate buffer (pH 7, 0.5M, 12 mL). Hydrogen peroxide (13.2 mL, 30%) was added dropwise and the mixture was warmed to rt and stirred for 1 h. Next, the mixture was diluted with sat. aq. NaHCO$_3$ (100 mL) and extracted with EtOAc (3 x 75 mL). The combined organic layers were washed with water and brine, dried over MgSO$_4$, filtrated and concentrated. The crude product was dissolved in DCM (15 mL) and Grubb's II catalyst (71 mg, 5 mol%) was added. The mixture was refluxed overnight under argon. The product was purified by flash chromatography (pentane/EtOAc; 8:1 → 5:1) to afford the product as colourless oil (518 mg, 77%, dr 92:8). ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.49 – 7.40 (m, 6H), 7.33 – 7.18 (m, 9H), 5.88 –

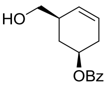
5.52 (m, 2H), 3.94 (dddd, *J* = 11.1, 9.1, 5.5, 3.5 Hz, 1H), 3.12 – 2.94 (m, 2H), 2.67 – 2.52 (m, 1H), 2.37 (d, *J* = 16.8 Hz, 1H), 2.21 – 2.10 (m, 1H), 2.04 – 1.87 (m, 1H), 1.80 (m, OH), 1.45 – 1.12 (m, 1H). ^{13}C NMR (MHz,) δ 144.3, 128.8, 128.6, 127.9, 127.0, 125.3, 86.5, 67.6, 67.5, 37.6, 36.4,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

Compound 28



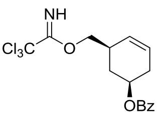
Compound **27** (675 mg, 1.7 mmol) was co-evaporated with toluene, dissolved in dry DCM (17 mL) and cooled to 0°C. Triethyl amine (1.2 mL, 8.5 mmol), benzoyl chloride (0.39 mL, 3.39 mmol) and DMAP (10 mg, 0.09 mmol) were added and the mixture was stirred overnight at rt. The reaction was quenched with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

Compound 29



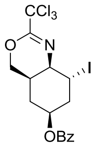
Compound **28** (660 mg, 1.39 mmol) was dissolved in a mixture of DCM (7 mL) and MeOH (7 mL). Then, CSA (16 mg, 0.07 mmol) was added and the mixture was stirred for 5 h at rt. The reaction was quenched with Et_3N (0.15 mL) and concentrated. The product was purified by flash chromatography (pentane/EtOAc;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

Compound 30



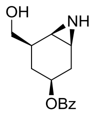
Compound **29** (290 mg, 1.25 mmol) was dissolved in dry DCM (12 mL). Then trichloroacetonitrile (0.25 mL, 2.5 mmol) and DBU (9.3 µL, 0.06 mmol) were added and the mixture was stirred overnight. The mixture was concentrated and the product was purified by flash chromatography (pentane/EtOAc; 20:1) to afford the product as a colourless oil (429 mg, 91%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.30 (s, 1H), 8.05 (d, *J* = 7.5 Hz, 2H), 7.56 (t, *J* = 7.4 Hz, 1H), 7.45 (d, *J* = 7.7 Hz, 2H), 5.78 (ddt, *J* = 10.0, 4.9, 2.5 Hz, 1H), 5.69 (d, *J* = 10.2 Hz, 1H), 5.30 (dddd, *J* = 12.0, 9.2, 5.7, 3.5 Hz, 1H), 4.26 (d, *J* = 6.8 Hz, 2H), 2.95 (m, 1H), 2.67 – 2.50 (m, 1H), 2.35 – 2.16 (m, 2H), 1.67 (q, *J* = 11.3 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.1, 162.9, 133.0, 130.5, 129.7, 128.4, 126.9, 126.0, 72.5, 70.2, 35.6, 31.5, 31.2. $[\alpha]_D^{20}$ +27.3 (c 0.3, DCM). IR (ATR, cm^{-1}) 1714, 1664, 1271, 1070, 794, 709. HRMS: $[M+H]^+$ calc for $C_{16}H_{16}Cl_3NO_3$ 376.02740 found 376.02706.

Compound 31



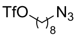
Compound **30** (561 mg, 1.49 mmol) was co-evaporated with toluene, dissolved in dry $CHCl_3$ (15 mL) and cooled to 0°C. Then, NIS (570 mg, 2.23 mmol) was added and the mixture was stirred overnight at rt. The reaction was quenched with aq. $Na_2S_2O_3$ (10 mL) and stirred for 10 minutes. The mixture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

Compound 32



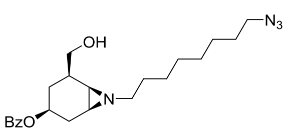
Compound **31** (51 mg, 0.1 mmol) was dissolved in a mixture of DCM (0.5 mL) and MeOH (0.5 mL) and cooled to 0°C. A solution of 1.25M HCl in MeOH (176 µL, 0.22 mmol) was added, and the mixture was stirred 16 h at rt. Then, Amberlite IRA-67 was added until neutral pH and the mixture was stirred for 2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

1-azido-8-trifluoromethylsulfonyloctane



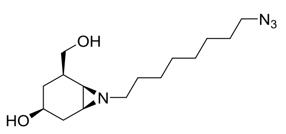
To dry DCM (5.8 mL) was added 8-azidooctan-1-ol (100 mg, 0.58 mmol) and pyridine (57 µL, 0.70 mmol) and the mixture was cooled to -20 °C. Triflic anhydride (118 µL, 0.70 mmol) was added and the mixture was stirred for 15 minutes. Then the mixture was diluted with DCM, and washed with cold water (3 x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

Compound 33



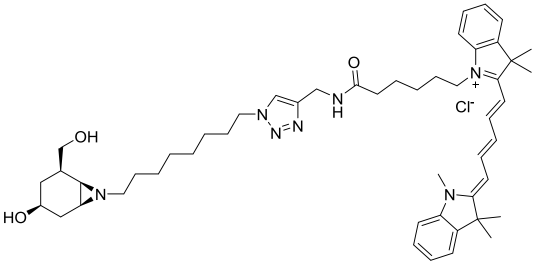
Compound **32** (54 mg, 0.22 mmol) was dissolved in dry THF (2 mL) and cooled to 0°C. DIPEA (42 µL, 0.24 mmol) was added and then 8-azidooctyl trifluoromethanesulfonate (73 mg, 0.24 mmol) in THF (0.5 mL). The mixture was stirred at 0 °C for 2 h and then diluted with water (20 mL) and brine (5 mL). The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

Compound 34



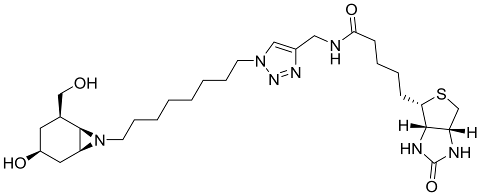
Compound **33** (46 mg, 0.11 mmol) was dissolved in MeOH (1.1 mL), then 0.1M NaOMe in MeOH (690 µL, 0.07 mmol) was added and the mixture was stirred for 24 h at rt. Then the mixture was diluted with sat. aq. $NaHCO_3$ (10 mL), and extracted with EtOAc (5 x 10 mL). The combined organic layers were dried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

Compound 9 (SY-D170)



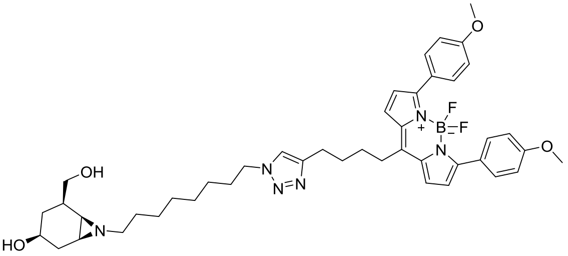
Following the general procedure starting from compound **34** (3.8 mg, 13 μmol), the product was obtained as a blue powder (3.2 mg, 30%). ^{1}H NMR (600 MHz, CD$_3$OD) δ 8.24 (td, *J* = 13.1, 2.9 Hz, 2H), 7.83 (s, 1H), 7.49 (d, *J* = 7.4 Hz, 2H), 7.41 (q, *J* = 7.3 Hz, 2H), 7.28 (m, 4H), 6.62 (t, *J* = 12.4 Hz, 1H), 6.28 (dd, *J* = 14.0, 2.1 Hz, 2H), 4.41 (s, 2H), 4.36 (t, *J* = 7.1 Hz, 2H), 4.09 (t, *J* = 7.5 Hz, 2H), 3.66 – 3.59 (m, 1H), 3.63 (s, 3H), 3.49 (m, 2H), 2.39 – 2.31 (m, 1H), 2.25 (t, *J* = 7.3 Hz, 2H), 2.22 – 2.17 (m, 1H), 2.07 – 2.00 (m, 2H), 1.87 (m, 2H), 1.83 (m, 2H), 1.73 (s, 12H), 1.71 – 1.65 (m, 4H), 1.55 – 1.45 (m, 6H), 1.33 – 1.27 (m, 8H), 0.97 (q, *J* = 11.7 Hz, 1H). ^{13}C NMR (151 MHz, CD$_3$OD) δ 175.7, 175.4, 174.6, 155.5, 155.5, 146.1, 144.2, 143.5, 142.6, 142.5, 136.0, 131.2, 129.8, 129.7, 128.8, 126.6, 126.3, 126.2, 124.1, 123.4, 123.3, 112.0, 111.8, 104.4, 104.2, 68.3, 66.3, 61.9, 51.3, 50.5, 44.8, 41.5, 38.8, 38.6, 36.5, 35.6, 33.8, 32.4, 31.5, 31.3, 31.1, 30.4, 30.2, 30.0, 28.3, 28.1, 27.9, 27.8, 27.4, 27.3, 26.4. HRMS: [M$^+$] calc for C$_{50}$H$_{70}$N$_7$O$_3$ 816.55401 found 816.55291.

Compound 10 (SY-D173)



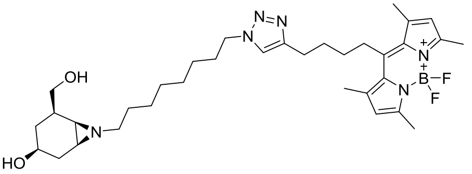
Following the general procedure starting from compound **34** (4.1 mg, 14 μmol), the product was obtained as a white powder (3.5 mg, 44%). ^{1}H NMR (600 MHz, CD$_3$OD) δ 7.84 (s, 1H), 4.49 (dd, *J* = 7.9, 4.9 Hz, 1H), 4.42 (s, 2H), 4.37 (t, *J* = 7.1 Hz, 2H), 4.29 (dd, *J* = 7.9, 4.4 Hz, 1H), 3.63 (dd, *J* = 10.2, 8.0 Hz, 1H), 3.50 (m, 2H), 3.19 (dt, *J* = 9.6, 5.2 Hz, 1H), 2.93 (dd, *J* = 12.7, 5.0 Hz, 1H), 2.71 (d, *J* = 12.7 Hz, 1H), 2.39 (dt, *J* = 11.7, 7.8 Hz, 1H), 2.23 (m, 3H), 2.04 (m, 2H), 1.89 (m, 3H), 1.76 – 1.63 (m, 5H), 1.63 – 1.51 (m, 4H), 1.49 (dd, *J* = 14.1, 10.3 Hz, 1H), 1.42 (m, 2H), 1.39 – 1.25 (m, 8H), 1.18 (t, *J* = 7.0 Hz, 1H), 0.97 (q, *J* = 11.8 Hz, 1H). ^{13}C NMR (151 MHz, CD$_3$OD) δ 176.0, 166.1, 146.2, 124.1, 68.3, 66.2, 63.3, 61.9, 61.6, 57.0, 51.4, 41.5, 41.1, 38.8, 38.7, 36.5, 35.6, 33.8, 32.4, 31.3, 30.4, 30.2, 30.0, 29.7, 29.5, 28.3, 27.4, 26.7. HRMS: [M+H]$^+$ calc for C$_{28}$H$_{47}$N$_7$O$_4$S 578.3483 found 578.3498.

Compound 35 (SY-D171)



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

Compound 36 (SY-D172)



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

Labelling and SDS-PAGE of recombinant enzymes

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

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

Competitive ABPP (recombinant enzyme)

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

In vitro labelling of lysates and SDS-PAGE analysis

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

Crystal structure of TB562 in *Cj*GH35

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

Pull-down and LC-MS/MS analysis

All pull-down experiments were performed in duplicates. Mouse kidney lysates, mouse liver lysates (1.0 mg total protein) or human fibroblasts lysates (250 µg total protein) were incubated with either 0.25% (v/v) DMSO, ABP **8** or **10** (50 µM) for 3 hours at 37 ℃, in a total volume of 200 µL McIlvaine buffer pH 5.0, followed by denaturation by addition of 10% (w/v) SDS 50 µL and boiling for 5 minutes at 100 °C. Samples were then reduced with the use of dithiothreitol (DTT), alkylated by iodoacetamide (IAA) and further prepared for pull-down with DynaBead MyOne Straptavidin Beads C1 as published previously.[34] After pull-down, all samples were used for on-bead digestion. The samples were treated with on-bead digestion buffer (100 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM $CaCl_2$, 2% (v/v) acetonitrile and 10 ng/μL trypsin) and incubated overnight in a shaker at 37 ℃. The supernatant, containing tryptic-digested peptides was then desalted using StageTips. Consequently, the acetonitrile was evaporated using a SpeedVac at 45 ℃ followed by addition of 20 μL of LC-MS sample solution (95:3:0.1, H_2O:acetonitrile:formic acid) for LC-MS analysis. All peptide samples were analysed with a two hour gradient of 5% to 25% acetonitrile on nano-LC, hyphenated to an LTQ-Orbitrap and identified using the Mascot protein search engine.[35] Raw data was calculated using MaxQuant against Uniprot of human (for fibroblasts) or mouse (for mouse kidney and mouse liver) proteome database to obtain an identification list of found proteins. The abundance of the protein hits was quantified as previously described,[25] in unsupervised mode using the default settings of the PLGS (Waters) and IsoQuant software.

Kinetics

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

References

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

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

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

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

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

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

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

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

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

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

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

12 J. Jiang, T. J. M. Beenakker, W. W. Kallemeijn, G. A. van der Marel, H. van den Elst, J. D. C. Codée, J. M. F. G. Aerts and H. S. Overkleeft, *Chem. Eur. J.*, 2015, **21**, 10861–10869.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

29 O. Kaczmarek, H. A. Scheidt, A. Bunge, D. Föse, S. Karsten, A. Arbuzova, D. Huster and J. Liebscher, *Eur. J. Org. Chem.*, 2010, 1579–1586.

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

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

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

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

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

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

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

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