Inhibitors and activity-based probes for β-D-glucuronidases, heparanases and β-L-arabinofuranosidases

Borlandelli, V.

4

Synthesis and Biochemical Evaluation of Activity-Based Probes and Inhibitors of Retaining β-L-Arabinofuranosidases

Published in part in:

Nicholas G. S. McGregor*, Joan Coines*, Valentina Borlandelli, Satoko Amaki, Marta Artola, Alba Nin-Hill, Daniël Linzel, Chihaya Yamada, Takatoshi Arakawa, Akihiro Ishiwata, Yukishige Ito, Gijsbert A. van der Marel, Jeroen D. C. Codée, Shinya Fushinobu, Herman S. Overkleeft, Carme Rovira, Gideon J. Davies. *Angew. Chem. Int. Ed.* **2021**, *60,* 5754–5758.

4.1 Introduction

β-L-arabinofuranosidases are rare catabolic microorganismal biocatalysts processing the hydrolytic cleavage of β-linked L-*arabino*furanosyl monosaccharidic units from the structure of *arabino*-oligosaccharides (AOS) and of complex plant-sourced glycoproteins (like solanaceous lectins)[1-2] and dietary pectic glycans[3-4] such as rhamnogalacturonan II (RG-II, Figure 4.1). Similarly to the majority of furanoside-processing enzymes, β-L-arabinofuranosidases are found in the proteome of the eukaryotic microorganismal domain, as exemplified by the human gut microbiome.[3,5] Distinct bacterial members of the human gut microbiome express β-L-arabinofuranosidases, and more abundantly α-L-arabinofuranosidases,[6-10] for the utilisation of specific plant-derived saccharidic substrates acting as nutrients and putative prebiotics. In particular, β-L-arabinofuranosidases are encoded in AOS utilisation genes by several *Bifidobacteria* species;[5] whereas some *Bacteroides* gut microbial species, exemplified by *Bacteroides thetaiotaomicron*, express these enzymes specifically for the catabolic depolymerisation of pectin.[3]

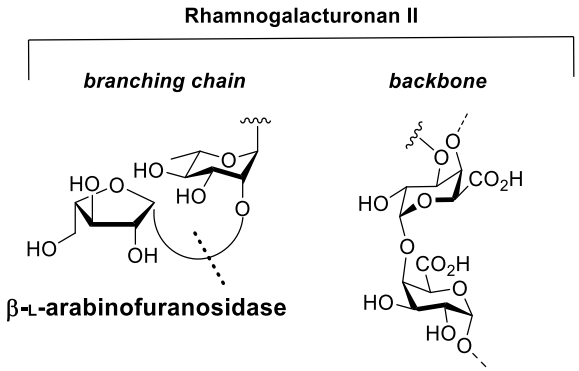


Figure 4.1. Simplified representation of the point of β-L-arabinofuranosidase enzymatic cleavage of β-L-*arabino*furanosidic units from plant-based pectic glycans, as exemplified by rhamnogalacturonan II. The structures of the RG-II polygalacturonic backbone and of a generic branching chain terminating with a β-linked L-*arabino*furanosyl unit are shown. The detailed structure of RG-II branching chains is still not fully established.[11-12]

Crystallographic information became recently available for two distinct microorganismal β-L-arabinofuranosidases belonging to GH families 127 (HypBA1 from *Bifidobacterium longum*)[13-14] and 146 (expressed by gut symbiont *Bacteroides thetaiotaomicron*).[15] Crystal complexes for both enzymes showed an uncommon catalytic active site structure within the repertoire of known GHs: a coordination complex constituted by three cysteines and a glutamate residue centered around a divalent zinc ion in the active site. These structural features potentially qualify HypBA1 and *Bt*GH146 as metal-dependent glycosidases. Although the

occurrence of metalloglycosidases is exceptional, several classes of hydrolases (esterase, peptidase, and β-lactamase) require a metal ion cofactor mechanistically involved in the catalytic reaction. The mechanism of both HypBA1 and *Bt*GH146 was unresolved for many years, and crystallographic[13-14] and site-directed mutagenesis[16] experiments aimed at elucidating the mechanism of action of these enzymes gave both unexpected and contrasting results. Specifically, the anomeric carbon of the β-L-*arabino*furanose substrate in its co-crystal complex with HypBA1 was observed in close proximity to C417,[14] suggesting the hydrolysis of a thioglycosyl-enzyme intermediate (tGEI) during the enzymatic reaction. Strikingly, the L-arabinofuranose substrate in complex with *Bt*GH146 was observed to be adjacent to C416,[15] the corresponding residue of C417 in HypBA1. Based on the crystallographic and mutational information, it was proposed that a cysteine residue (C417 in HypBA1) might act as nucleophile in place of a canonical glutamate or aspartate, with participation of Zn^{2+} cation as cofactor in the mechanism (Figure 4.2A).[14] Furthermore, biochemical and crystallographic studies of HypBA1 established this enzyme as retaining GH, capable of hydrolysing terminal β-L-*arabino*furanosides with retention of configuration at the anomeric center.[16] In line with the range observed in canonical retaining GHs, crystallographic evidence[14] for the distance (4.9 Å) between two carboxylates (E332 and E338) within the active site of HypBA1 was in favour of the hypothesis of a "two-carboxylate" enzymatic mechanism, thus contrasting with the evidences for a thiol-partecipating catalytic mechanism. Together, the collected information at the time was inconclusive for the assignment of the catalytic nucleophile as well as for the identification of the putative thioglycosyl intermediate for β-L-arabinofuranosidases HypBA1 and *Bt*GH146.

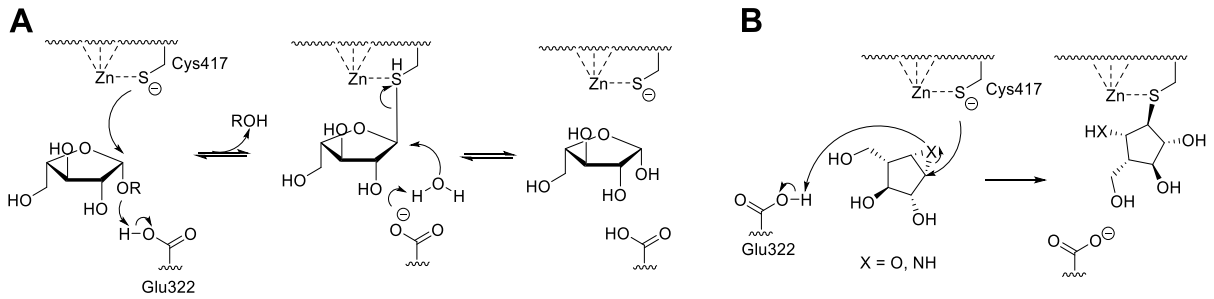


Figure 4.2. (A) Putative mechanism of action for β-L-arabinofuranosidase from GH127 proposed by Ito *et al.*[14] (B) Proposed mechanism of action of cyclophellitol-type β-L-*arabino*furano-configured inhibitors and probes, involving a putative thioglycosyl covalent adduct similarly to the ester-linked glycosidic covalent enzyme-substrate adduct in conventional glycosidases.

Under the assumption that processing of native β-L-*arabino*furanosides by retaining β-L-arabinofuranosidases involves formation of a transient thioglycosyl covalent ligand-enzyme adduct during the catalytic reaction, it was anticipated that β-L-*arabino*furano-configured cyclophellitol-type probing tools would react with retaining β-L-arabinofuranosidases by forming a covalent thioglycosyl enzyme-ligand intermediate (Figure 4.2B), which resembles the ester-linked ligand-enzyme adduct in canonical glycosidases. Globally, the limited understanding regarding the catalytic mechanism and conformational itinerary adopted by these retaining β-L-arabinofuranosidases justifies the need for mechanistic probes which allow visualisation of these selected enzymes in their reacted form. ABPP provides an excellent approach to interrogate functional proteins in complex biological milieu, including biomass-degrading enzymes[17] and human gut microbial systems.[18] Probe design is mostly accomplished by mimicking the transition state conformation and configuration, thus prior knowledge on the mechanistic properties of the enzyme and the conformational properties of its native substrates is important. Five-membered-ring furanosides are characterised by a pronounced and less investigated conformational complexity when compared to the well-established conformational landscapes of six-membered pyranosides.[19-21] Although the conformational landscape has not been determined specifically for β-L-*arabino*furanosides, computational analysis of their stereoisomeric α-L-*arabino*furanoside gave indirect insight into the thermal constraints for furanosides conformers.[22] Among the twenty thermally-accessible ring conformations of furanosides estimated by Cremer-Pople analysis (Figure 4.3A),[23] two envelope-type conformations (that is, ^{3}E and E_3) are exclusively capable of stabilising the oxocarbenium cation[24-25] formed in the transition state (TS) of glycoside hydrolysis catalysed by conventional retaining GHs, similarly to the TS of glycosylations of furanosyl acceptors, indicating that all retaining arabinofuranosidases are expected to process their substrates through an envelope-like ^{3}E or E_3 TS conformation.

Because the planar orientation of atoms O1, C1, C2 and C4 in the envelope-like conformations ^{3}E and E_3 accommodates the geometry of an unsaturated double bond (such as the oxocarbenium bond), it was conceived that *arabino*furano-configured cyclophellitols or cyclophellitol aziridines would adopt an envelope-like conformation in which the three-membered ring would exploit the planar orientation of atoms available in this particular conformations. Thus, β-L-*arabino*furano-configured cyclophellitols or cyclophellitol aziridines would serve as putative conformational β-L-arabinofuranosidase mimetics of the oxocarbenium-like TS. Inspection of the conformational possibilities available to

*arabino*furanosides in general calls forth an additional consideration. Namely, the ^{3}E oxocarbenium-stabilising envelope conformation of α-D-*arabino*furanosyl cyclophellitol-type scaffolds appears to emulate the $^{4}H_{3}$ conformation of six-membered ring 1,6-*epi*-cyclophellitol aziridine,[26] an established covalent inhibitor of retaining α-glucosidases (Figure 4.3B). This implies the potential utility of five-membered strained scaffolds as putative conformational TS mimetics of glyco-pyranosidases intercepting a $^{4}H_{3}$ TS conformation depending on the configurational patterns of the scaffolds,[27] thereby further expanding the potential scope of furanoside-like structures.

With the aim to interrogate the mechanism of action of distinct retaining β-L-arabinofuranosidases and with the long-term aim to profile novel enzymatic entities of the human gut microbiome endowed with retaining β-L-arabinofuranosidase activity, activity-based probes (ABPs) and inhibitors emulating β-L-*arabino*furanosides were developed. This chapter presents the design, synthesis and preliminary biochemical evaluation of a panel of β-L-*arabino*furanosyl cyclophellitol and cyclophellitol aziridine inhibitors and ABPs (Figure 4.4) designed to detect and modulate the hydrolytic activity of retaining β-L-arabinofuranosidases.

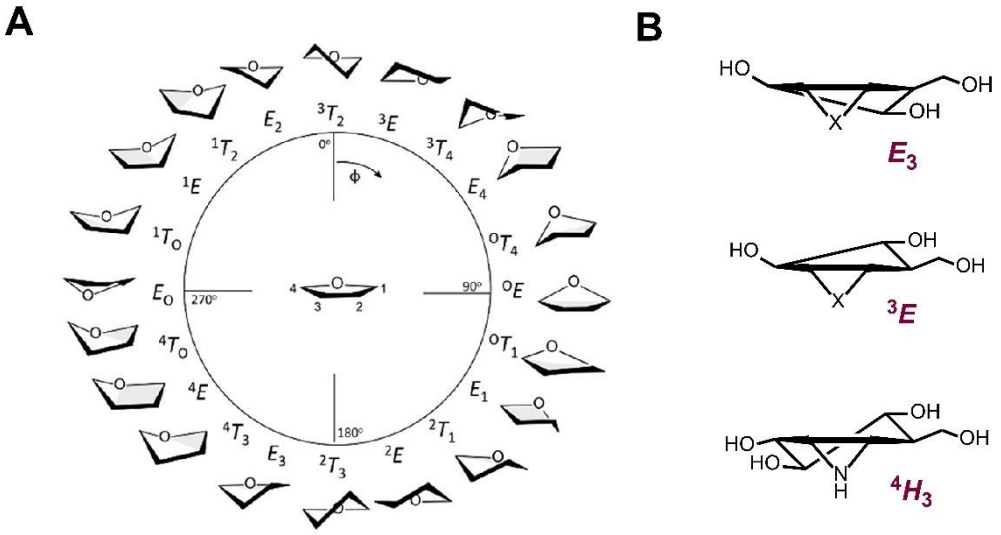


Figure 4.3. (A) Graphical representation of the conformations of a 5-membered furanosyl ring according to the Cremer–Pople angle ϕ.[23] (B) Schematic representation of α-D-*arabino*furanosidic cyclophellitol-type scaffolds in E_3 and ^{3}E conformation, with the ^{3}E conformer resembling the $^{4}H_{3}$ TS conformation of retaining α-D-glucosidases similarly to 1,6-*epi*-cyclophellitol aziridine.

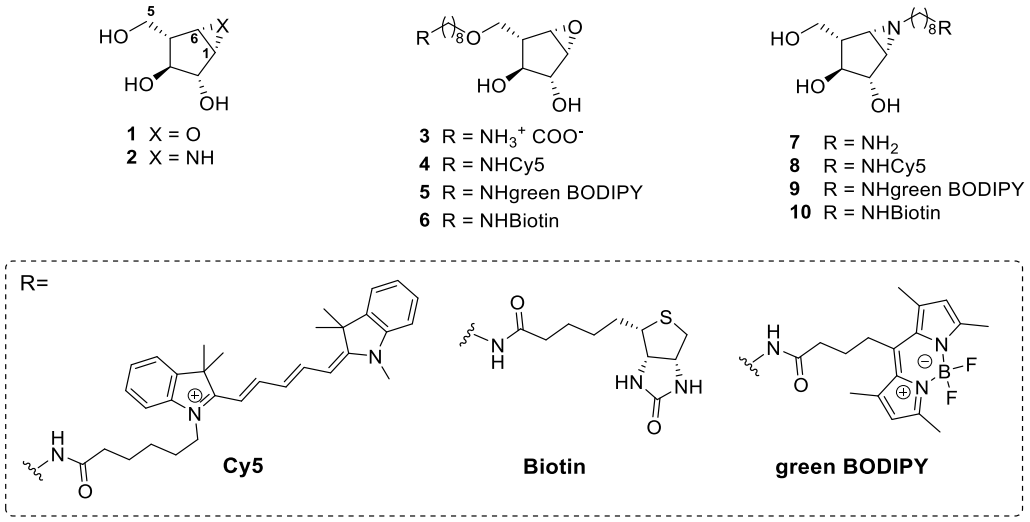


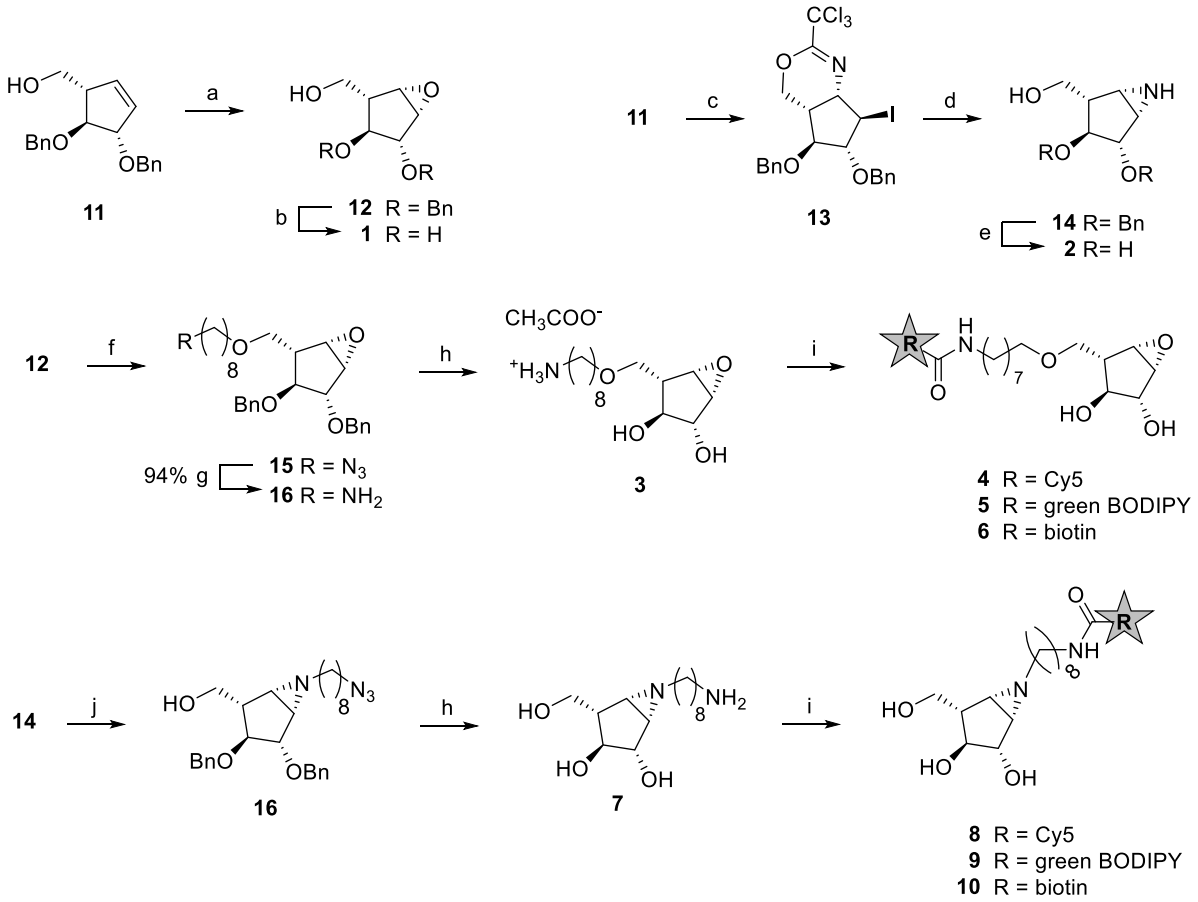
Figure 4.4. Structures of inhibitors and activity-based probes (ABPs) targeting retaining β-L-arabinofuranosidases presented in this chapter.

4.2 Results and discussion

4.2.1 Synthesis of β-L-*arabino*furanosyl inhibitors and activity-based probes

Starting from commercially available α-D-*galacto*pyranoside, bisbenzylated L-*arabino*furanosyl cyclopentene **11** was obtained in nine steps following the synthetic strategy recently utilised for the preparation of a set of α-L-*arabino*furanosyl probes and inhibitors.[22] The synthesis of both epoxide-based and aziridine-based β-L-*arabino*furanosyl scaffolds **3**–**10** involved installation of the electrophilic trap prior incorporation of the linker (Scheme 4.1).

The primary homoallylic hydroxyl group in **11** was employed to impart the desired β-facial selectivity for both sets of compounds. Specifically, epoxidation of **11** with *m*CPBA yielded bisbenzylated β-L-*arabino*furanosyl epoxide **12** with high degree of stereoselectivity when performed at low temperature, following recently reported procedure.[22] Next, the linker was introduced by *O*-alkylation of **12** with freshly prepared azido-octyltriflate in presence of a stoichometric amount of NaH. Step-wise Staudinger azide reduction of *O*-alkylated **15** and global reductive debenzylation of **16** gave prospective inhibitor **3**. Preactivation of a carboxy-reporter (Cy5-acid, greenDBP-acid or biotin) with pentafluorophenyltrifluoroacetate (PFP-TFA) prior to addition to amine **3** afforded probes **4**-**6**. These were purified by reverse-phase HPLC chromatography with exception of green Bodipy-tagged **5**, which was purified by silica gel column chromatography due to its lipophilic nature. Aziridine-based ABPs **8**-**10** together with inhibitors **2** and **7** were accessible from bisbenzylated β-L-*arabino*furanosyl aziridine **14**. The aziridine was introduced using a stereospecific sequence of reactions which previously



Scheme 4.1. Synthesis of β-L-*arabino*furanosyl inhibitors **1**-**3** and **7**, and activity-based probes **8**–**9** and **8**–**10**. Reagents and conditions: (a) *m*-CPBA, NaH_2PO_4/Na_2HPO_4 buffer, DCM, 4 ˚C, 3 days, 74%. (b) H_2, Pd/C, CH_3OH, 18 h, 18%. (c) CCl_3CN, DBU, DCM, 0 ˚C → rt, 3 h; then NIS, $CHCl_3$, 0 ˚C → rt, 17 h, quant. yield. (d) HCl in CH_3OH, DCM/CH_3OH, 6 h; then Amberlite IRA-67, 1 day, 63% over 2 steps. (e) Na(s), NH_3, THF/^{t}BuOH, -60 ˚C, 45 min, 23%; (f) azido-octyltriflate, NaH, THF, -15 ˚C → rt, 28 h, 84%; (g) polymer-bound PPh_3, AcCN/H_2O, 70 ˚C, 5h, 94%; (h) $Na_{(s)}$, NH_3, THF/^{t}BuOH, - 60 ˚C, 45 min, **3**: 84%, **7**: 77%; (i) reporter-COOH, Pfp-TFA, DMF; then **3** or **7**, DIPEA, DMF, 24 h. **4**: 7%; **5**: 41%; **6**: 23%; **8**: 8%; **9**: 45%; **10**: 8%; (j) azido-octyltriflate, DIPEA, DCM, -15 ˚C → rt, 20 h, 44%.

found successful application in the aziridination of pyranose-like substrates[28-30] yet had not previously been put to the test on five-membered-ring substrates. Specifically, unsaturated **11** was subjected to trichloroimidation of the primary hydroxyl group after which base-mediated iodocyclisation yielded iodo-derivative **13** quantitatively and stereospecifically. Acidic hydrolysis of cyclic imidate **13** generated *in situ* the primary ammonium species which upon base treatment yielded aziridine **14** with concomitant iodine displacement. At this stage, the synthesis of aziridine-based probes diverges from the preparation of non-alkylated inhibitor **2**. Global debenzylation of **14** under Birch conditions followed by reverse-phase HPLC chromatography using a HILIC column gave final product **2**. *En route* to probes **8**-**10**, *N*-

alkylation of **14** with freshly prepared azido-octyltriflate followed by one-pot reductive debenzylation and azide reduction under Birch conditions afforded primary amine **7**. Similarly to the preparation of ABPs **4**-**6**, pentafluorophenyl-trifluoroacetate-activated esters of the reporters were generated *in situ* for the ensuing amide coupling with amine **7**, giving, after HPLC purification, target compounds **8**-**10**.

4.2.2 *In vitro* evaluation of β-L-*arabino*furanosyl inhibitors and activity-based probes

As preliminary studies, the inhibitory activity of newly-developed inhibitors **1**, **2**, **7** and probe **4** as prospective β-L-arabinofuranosidase inactivators was studied, and inhibition kinetics were measured for recombinant HypBA1 and *Bt*GH146 at the optimal pH (pH 4.5 for rHypBA1 and pH 7.5 for r*Bt*GH146), using hydrolysis of chromogenic *p*-nitrophenyl-β-L-*arabino*furanoside[31] as read-out of residual enzymatic activity. Remarkably, secondary aziridine **2** displayed the superior inhibitory efficacy against recombinant HypBA1 (k_{inact}/K_I > $4.2 \cdot 10^3$ $M^{-1} \cdot s^{-1}$ as illustrated in Table 4.1) compared with *Bt*GH146 ($k_{inact}/K_I = 62$ $M^{-1} \cdot s^{-1}$, Table 4.1) and the highest inhibitory efficacy towards recombinant β-L-arabinofuranosidases within the set of tested compounds, thus proving its elevated reactivity towards both retaining HypBA1 and *Bt*GH146.

Table 4.1. Inhibition kinetic parameters of β-L-*arabino*furanosyl compounds **1**, **2**, **4** and **7** measured for retaining rHypBA1 and r*Bt*GH146.

	r*Bt*GH146			rHypBA1		
Compound	k_{inact} **(min^{-1})**	K_I **(μM)**	k_{inact}/K_I **(M^{-1}·s^{-1})**	k_{inact} **(min^{-1})**	K_I **(μM)**	k_{inact}/K_I **(M^{-1}·s^{-1})**
1	n.d.	n.d.	< 0.5	n.d.	n.d.	14.7
2	79 ± 6	350 ± 40	62	> 1	n.d.	> 4200
4	3.2 x 10^{-2} ± 0.3 x 10^{-2}	140 ± 10	3.8	n.d.	n.d	< 0.1
7	8.3 ± 0.4	190 ± 20	12	1.3 ± 0.3	280 ± 90	76

Inactivation rates and inhibition constants (k_{inact} and K_I). [a]n.d.: not determined.

Interestingly, the k_{inact}/K_I value measured for the interaction between rHypBA1 and **2** is significantly higher than the one corresponding to the inhibition of fungal α-L-arabinofuranosidases from *Aspergilllus niger* (*An*AbfA, GH51) and from *Aspergillus kawachii* (*Ak*AbfB, GH54) by the stereoisomeric α-L-*arabino*furano-configured cyclophellitol aziridine (that is: 39 $M^{-1} \cdot s^{-1}$ against *An*AbfA and 28 $M^{-1} \cdot s^{-1}$ against *Ak*AbfB).[15,22] Tertiary aziridine **7** and epoxide **1** displayed more than two orders of magnitude lower rHypBA1 k_{inact}/K_I values than **2**, indicating that these compounds are less capable to covalently react with the tested enzymes or

to adopt the competent binding pose for the enzymatic reaction. Within the set of tested scaffolds, Cy5-tagged epoxide ABP **4** showed the lowest k_{inact}/K_I value towards rHypBA1, and moderate k_{inact}/K_I value (3.8 $M^{-1} \cdot s^{-1}$) towards r*Bt*GH146, making it a slow inhibitory and labeling tool for the studied enzymes.

Comparison of k_{inact}/K_I values for **1** and **2** against rHypBA1 reveals that inhibitory efficacy is dependent on the nature of the warhead. This trend holds true for r*Bt*GH146 inhibition as well, though to a lesser extent. Aziridines **2** and **7** and epoxide-type probe **4** displayed r*Bt*GH146 inhibitory potency (K_I values) and kinetics (k_{inact} values) in the same order of magnitude, outperforming epoxide **1**. From this assessment, it can be inferred that *N*-tagged β-L-*arabino*furanosyl aziridines possess superior inhibitory efficacy for both rHypBA1 and r*Bt*GH146 compared to *O*-alkylated β-L-*arabino*furanosyl epoxides, and may be the preferred retaining β-L-arabinofuranosidase ABPs.

To establish the reactivity of epoxides **3**-**6** and aziridines **2** and **7** in relation to the measured kinetic parameters, intact mass spectrometry analysis was conducted on samples of rHypBA1 and r*Bt*GH146 incubated with these compounds for 1 or 16 hours (see Appendix, Figure 4.S2). In accordance with the above-described inhibitory data, the selected compounds displayed reactivities dependent on the nature of the electrophilic trap and of the reporter when treated with the tested enzymes. In particular, covalent adduct formation was observed upon incubation of rHypBA1 with both aziridine **2** and *N*-alkyl aziridine **7** after 1 hour. However, the observed mass difference (+146.2 for **2** and +273.8 for **7**) differs of one unit from the expected values (+145 for **2** and +272 for **7**). The HypBA1 binding cleft differs from *Bt*GH146 for a non-conserved cysteine residue (C415) placed in proximity of the catalytic nucleophile C417. The reported mass difference for rHypBA1 might account for a change in protonation state of this additional non-conserved thiol, which may undergo protonation under the tested conditions. An alternative hypothesis is that one of the non-covalently modified conserved cysteines (C418, C340) may lose coordination to the zinc cation, as a result of protein denaturation during the intact mass spectrometry analysis. More experiments are needed to investigate the reported observation further. Conversely, r*Bt*GH146 exhibited broader reactivity and formed ligand-enzyme adducts with all tested compounds with exception of **3**, **5**, and **6**, which all displayed zero to partial covalent adduct formation after 1 hour. However, reactivity with **5** or **6** could be driven to completion by overnight incubation. Comprehensively, these observations further corroborate the findings derived from the *in vitro* kinetic assessment with chromogenic substrate.

4.2.3 Structural characterisation of enzyme-ligand interactions with 1 and 2

Insight into the mechanism of distinct β-L-arabinofuranosidases from GH families 127 and 146 was next obtained with 3D crystal analysis of complexes for **1** and **2** with rHypBA1 and r*Bt*GH146. Both crystal structures revealed ligand density in close proximity to the $Zn(Cys)_3(Glu)$ coordination complex (Figure 4.5 and 4.6). Analysis of the crystal structures of both enzymes reacted with both epoxide **1** and aziridine **2** revealed a remarkable difference for the mode of action of r*Bt*GH146 enzyme. While the active site cysteine in rHypBA1 (C417) was found at covalent-bond distance to C1 (the carbon emulating the anomeric carbon in natural substrate β-L-arabinofuranosides) of both non-alkylated epoxide **1** (Figure 4.5B) and aziridine **2** (Figure 4.6B) in their respective cocrystals, the situation was different for r*Bt*GH146. Reasons for this can be found by closer inspection of their crystal structures in complex with r*Bt*GH146.

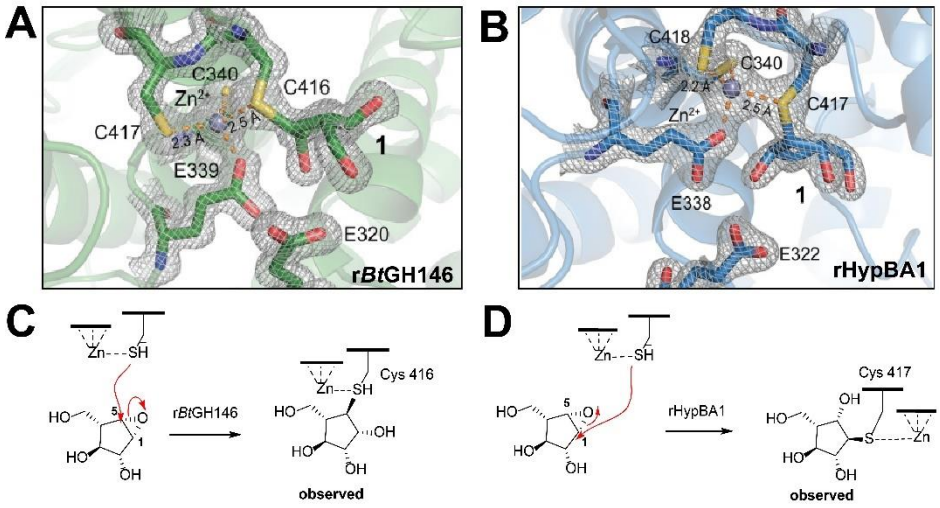


Figure 4.5. (A) Active site structure of r*Bt*GH146 upon reaction with **1** revealing electron density indicative of covalent bond between C6 and C416. The $2F_o$-F_c map is shown contoured to 2σ for key residues, and bonds coordinating the zinc ion are depicted as dashed orange lines. (B) Active site structure of rHypBA1 bound to **1**, with density and coordination shown as in panel A. (C) Schematic representation of the enzymatic processing of **1** by r*Bt*GH146, illustrating nucleophilic attack at C6 of **1**. (D) Schematic representation of the enzymatic processing of **1** by rHypBA1, illustrating nucleophilic attack at C1 of **1**.

When treated with **1**, the active site cysteine nucleophile (C416) in r*Bt*GH146 reacted with the unexpected epoxide carbon, that is, C6 – the carbon replacing the ring-oxygen in natural substrate β-L-arabinofuranosides (Figure 4.5A). However, when treated with aziridine **2**, the cysteine thiol is observed at covalent-bond distance to C1, the pseudo-anomeric carbon (Figure 4.6A). The atypical regiochemical outcome with **1** is accompanied by a change in non-covalent interactions and in spatial orientation of the substrate within the enzymatic active site, as

indicated by comparison of crystal complex r*Bt*GH146:**1** with the crystal structure obtained for native ligand β-L-*arabino*furanose. Specifically, it was found that occupancy of r*Bt*GH146 active site by **1** differs from the binding pose of the native ligand by a 75° net rotation. Together with this, primary hydroxyl O5 of **1** in r*Bt*GH146:**1** was at H-bond distance from binding residue E320 in place of E217 – the residue interacting with the primary hydroxyl O5 atom of both native substrate β-L-*arabino*furanose and aziridine **2** (Figure 4.6C) in their correspondent crystal complexes with r*Bt*GH146.

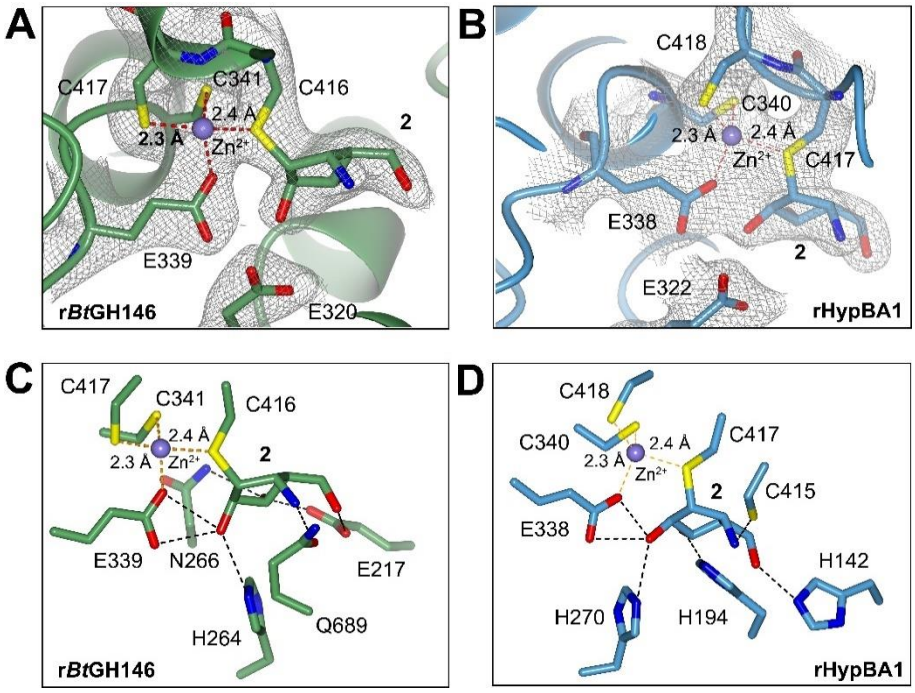


Figure 4.6. (A) Active site structure of r*Bt*GH146 upon reaction with **2** revealing electron density indicative of covalent bond between C1 and C416. The $2F_o$-F_c map is shown contoured to 1σ for key residues, and bonds coordinating the zinc ion are depicted as dashed orange lines. (B) Active site structure of rHypBA1 bound to **2**, with density calculated for **2** and C417, and coordination shown as in panel A. (C) Active site structure of r*Bt*GH146 upon reaction with **2** showing hydrogen bond interactions between **2** and active site of r*Bt*GH146, with hydrogen bonds shown as black dashed lines and coordination to the zinc ion as dashed red lines with distances in Angstroms. (D) Active site structure of rHypBA1 bound to **2**, with hydrogen bonds and coordination shown as in panel C.

Importantly, these crystallographic studies provided structural evidence in favour of a putative mechanism of action involving a thioglycosyl-enzyme intermediate. The mechanistic information inferred from the observed crystallographic results for **1** was further assessed by molecular dynamics (MD) simulations and QM/MM MD calculations informed by obtained X-ray crystal structures. MD simulations revealed that C6 of **1** is in closer proximity of C416 compared to C1 during the entire simulation, priming C6 for nucleophilic attack by C416.

QM/MM MD calculations informed by the obtained X-ray crystal structures pointed towards a $^4E \rightarrow [E_3]^{\ddagger} \rightarrow {}^2T_3$ conformational itinerary for the deglycosylation of putative thioglycosyl-enzyme intermediate (tGEI) in HypBA1. Altogether, intact mass spectrometry measurement of the covalent enzyme-epoxide adduct, enzymatic activity assays, computational studies and 3D crystal structures underscored the occurrence of a covalent enzyme-substrate intermediate adduct in GH127-mediated substrate processing.

For non-alkylated aziridine **2**, analysis of the H-bond interactions with the active sites of both rHypBA1 and r*Bt*GH146 showed similarities in the pattern of non-covalent interactions involving hydroxyl O1 of **2**, a nearby histidine (H264 in r*Bt*GH146, H270 in rHypBA1) and a zinc-coordinated glutamate (E338 in r*Bt*GH146, E339 in rHypBA1). However, hydroxyl O5 and the nitrogen atom of the *trans*-diaxially opened aziridine **2** were found to interact differently with the active site in the two crystal complexes examined. In particular, primary hydroxyl O5 in **2** interacted with different H-bond acceptor residues within the binding pocket – E217 in r*Bt*GH146:**2** (Figure 4.6C), and H142 in HypBA1:**2** (Figure 4.6D). As per the amino group of the opened aziridine in **2**, that was observed to make contact with the thiol group of non-conserved cysteine C415 in rHypBA1:**2**. For r*Bt*GH146, however, no cysteine equivalent to C415 in rHypBA1 is present, and N1 of **2** exhibited exclusively H-bond to Q689 in r*Bt*GH146:**2** complex. The elevated degree of similarity in spatial occupancy between the native ligand and aziridine cyclitol **2** in complex with r*Bt*GH146, together with the "correct" regiochemical outcome at C1 by **2** with r*Bt*GH146, is evidence of the superior binding affinity of the non-alkylated aziridine substrate for r*Bt*GH146 compared to **1**.

Comparison of r*Bt*GH146:**1** and r*Bt*GH146:**2** highlighted a major difference in non-covalent interactions which might have an influence on the spatial occupancy of r*Bt*GH146 binding pocket. In particular, the amino group of reacted aziridine **2** with r*Bt*GH146 was at H-bond with Q689. It was hypothesised that this key interaction is enabled by the H-bond donor capability of the aziridine moiety, and this non-covalent interaction might be arguably at play throughout the enzymatic reaction. To maintain this H-bond, substrate **2** is oriented in a similar position to the one adopted by the native ligand in r*Bt*GH146 enzymatic pocket – with the anomeric carbon in close proximity of C416, and with primary hydroxyl O5 interacting with E217. On the contrary, the absence of H-bond donor ability in the epoxide ring of **1** might be responsible for its alternative spatial orientation – and hence different regiochemical outcome – observed in r*Bt*GH146:**1**, in which the primary hydroxyl group at C5 resembled the amino group of aziridine **2** as H-bond donor for interaction with nearby E320, while C6 was on the E339-Zinc-C416 axis

for nucleophilic attack. Globally, the regiochemical outcomes inferred from the above crystal structures and the observed ligand spatial occupancies within the active sites of both rHypBA1 and r*Bt*GH146 can be likely ascribed to the H-bond acceptor capability and to the superior electrophilic reactivity – when protonated – of the aziridine moiety compared to the epoxidic ring. These two features make **2** readily poised for nucleophilic attack by the catalytic nucleophile of both enzymes. In contrast, epoxide **1** does not react as effectively within the active site of r*Bt*GH146. The reactivity trend derived from the above crystal structures is mirrored by the kinetic parameters estimated against recombinant rHypBA1 and r*Bt*GH146 as shown in Section 4.2.2, and is specifically in accordance with the modest binding affinity of **1** for r*Bt*GH146 (K_I = 350 ± 40 μM) and with the faster inhibition kinetics (k_{inact}/K_I) displayed by **2** with both enzymes. When combined together, these crystallographic results indicate the potential effective broad-spectrum inhibitory ability of cyclophellitol-based aziridines against rHypBA1 and r*Bt*GH146.

While the covalent adduct of r*Bt*GH146 with aziridine **2** is the expected result and provides further evidence supporting annotation of this enzyme, like rHypBA1, as a retaining β-L-arabinofuranosidase utilising a cysteine catalytic site nucleophile, the difference in mode of action towards aziridine **2** and epoxide **1** is remarkable. While a definite explanation for this difference is not available at present, it was noticed that the aziridine inhibits – therefore binds and reacts – much faster than the epoxide, which may indicate that, upon binding, there is limited time for ligand repositioning.

4.2.4 Fluorescent labeling of recombinant β-L-arabinofuranosidases

With the fluorescent probes **4**-**6** and **8**-**10** readily available, their labelling efficacy towards selected recombinant β-L-arabinofuranosidases was investigated next. First, the time-dependent labeling of r*Bt*GH146 and rHypBA1 by epoxide-armed ABP **4** was studied by measuring in-gel fluorescence after ABP incubation at distinct time intervals up to 4 hours (Figure 4.7). r*Bt*GH146 was effectively labelled by **4** after 4 hour ABP incubation, whereas the labeling potency towards rHypBA1 was dramatically reduced, with moderately detectable bands observed after 1 hour at higher contrast. These results are in accordance with the initial *in vitro* kinetic assessment trend, and illustrate that epoxide-armed β-L-*arabino*furanosyl scaffolds are suitable ABPs for r*Bt*GH146 but not rHypBA1.

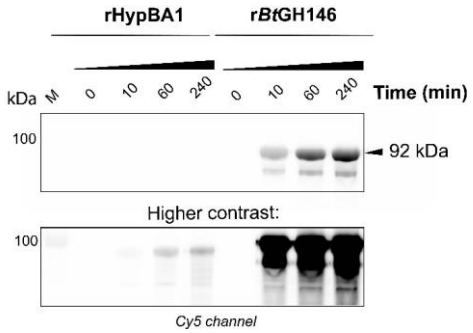


Figure 4.7. Time-dependent fluorescent labeling of rHypBA1 (left) and r*Bt*GH146 (right) by epoxide-armed Cy5-tagged ABP **4** during 4 hour incubation at optimal enzyme pH (rHypBA1: pH 4.5 NaOAc, 1 mM DTT; r*Bt*GH146: pH 7.0 NaPi, 1 mM DTT), showing a faster and more effective labeling of r*Bt*GH146 by **4** compared to rHypBA1 within the tested time window.

Having assessed the labelling efficacy of **4** against two different β-L-arabinofuranosidase isozymes and taking into account the biomedical relevance of *Bt*GH146 in human gut physiology, ensuing in-gel studies were conducted on r*Bt*GH146. Detection limits of r*Bt*GH146 by distinct fluorescent epoxide-type probes **4** and **5** were established by in-gel fluorescence labeling. As illustrated in Figure 4.8, Cy5-equipped and green Bodipy-equipped ABPs **4** and **5** exhibited comparable labeling efficacy against r*Bt*GH146 in the micromolar range after 1 hour ABP incubation. This implies that both probes may be used interchangeably in future labeling of complex biological samples for their ability to label r*Bt*GH146.

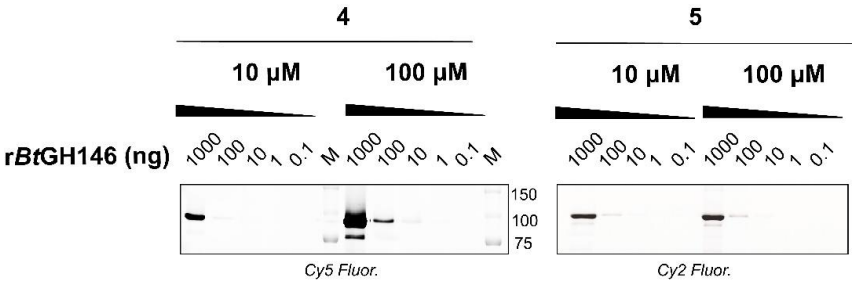


Figure 4.8. Detection sensitivity of Cy5-tagged ABP **4** (left) and green Bodipy-tagged ABP **5** (right) towards r*Bt*GH146 after 1 hour incubation at pH 7.5 (HEPES/NaCl), with 10 and 100 µM final probe concentration.

Within the set of aziridine-type probes, Cy5- and green Bodipy-ABPs **8** and **9** were assessed for their in-gel r*Bt*GH146 labeling efficacy. On the basis of the reported *in vitro* k_{inact} and K_I values measured for **2** and **7**, experiments were conducted with 1 or 10 µM final ABP concentration (Figure 4.9). Interestingly, fluorescent bands were detected after 3 minutes incubation with 10 µM fluorescent ABPs **8** or **9**, whereas no significant labeling was observed under 4 hours with lower concentration of ABP **8** (1 µM).

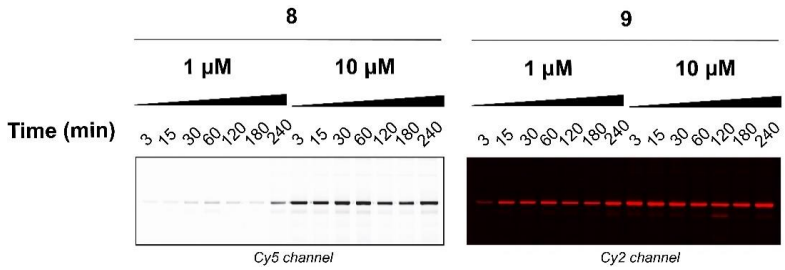


Figure 4.9. Fluorescent labeling of r*Bt*GH146 (100 ng) by Cy5-tagged ABP **8** (left) and green Bodipy-tagged ABP **9** (right) after 1 hour incubation at pH 7.4 (PBS/NaCl) and 1 or 10 μM final probe concentration, with ABP **9** displaying faster and more effective labeling of the tested enzyme than **8** at 1 μM ABP concentration.

To compare the labeling efficacy of green Bodipy-tagged probes in relation to the nature of the electrophilic trap, epoxide-type probe **5** and aziridine-type ABP **9** were tested against r*Bt*GH146 (Figure 4.10). Importantly, a visible fluorescent band was detected after 1 hour incubation of 100 ng r*Bt*GH146 with 1 or 10 μM of aziridine **9**. This amount of enzyme could be detected with epoxide ABP **5** as well, though at higher ABP concentrations (100 μM). It can thus be concluded that green Bodipy-aziridine ABP **9** has greater labeling efficacy compared to the epoxide-type probe. This observation is in agreement with the obtained *in vitro* kinetic parameters for the same compounds and holds true independently of the nature of the fluorescent reporter (Cy5 or green Bodipy).

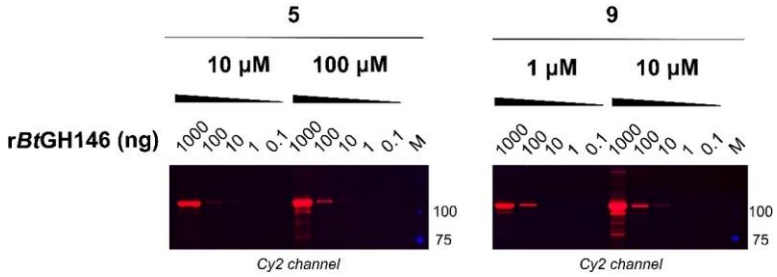


Figure 4.10. Detection sensitivity of green Bodipy-tagged epoxide-type ABP **5** and aziridine-type ABP **9** towards r*Bt*GH146 after 1 hour incubation (with **5**: at pH 7.5 HEPES/NaCl; with **9**: pH 7.4 PBS/NaCl), showing the lower r*Bt*GH146 detection limit of aziridine-type probe **9** with respect to **5** under the tested conditions.

Taken together, the results indicate that fluorescent aziridine-type β-L-*arabino*furanosyl probes can be more effective in labeling recombinant *Bt*GH146. The reactivity trend observed with the epoxide-type and aziridine-type probes is likely due to the site of installation of the linker and of the reporter in the probe design. This trend is analogous to the labeling capabilities of epoxide-type and azidirine-type probes for retaining β-D-glucosidases: namely, epoxide-based ABPs are selective for lysosomal glucocerebrosidase (GBA1),[32] whereas aziridine-based

probing tools label a broad range of mammalian and non-mammalian β-D-glucosidases with high detection sensitivity.[33]

Globally, these findings further verify and complement the *in vitro* inhibition kinetic and crystallographic studies described in the previous sections, and are likely to guide future ABP labeling of recombinant *Bt*GH146. As per HypBA1, its oxidative sensitivity and its lesser gene expression by the human gut microbiota makes it a more challenging target for ABPP. Future in-gel labeling experiments on HypBA1 might inform about the labelling efficacies towards this isozyme.

4.3 Conclusion

This chapter presented the synthesis of a panel of β-L-*arabino*furanosyl-configured epoxides and aziridines as inhibitors and ABPs for retaining β-L-arabinofuranosidases. The synthetic strategy builds on the preparation of a previously reported set of α-L-*arabino*furanosyl cyclophellitol-inspired probes and inhibitors,[22] and diverges from this at the stage of installation of the electrophilic trap onto L-*arabino*-cyclopentene **13**. The compounds were tested *in vitro* for their inhibitory kinetics and labeling efficacies. Crystallographic analysis conducted on complexes of β-L-*arabino*furanosyl cyclophellitol **1** with r*Bt*GH146 and rHypBA1 demonstrated unambiguously the formation of a thioglycosyl intermediate involving a Zn-coordinated cysteine nucleophile present in retaining β-L-arabinofuranosidases from GH families 127 and 146. With the mechanism of action established, the regiochemical outcome of the catalytic reactions as inferred from the crystal structures of **1** with both tested enzymes was investigated further by 3D crystal analysis of complexes of rHypBA1 and r*Bt*GH146 with non-alkylated aziridine **2**, the correspondent aziridine equivalent of **1**. Altogether, the collected results indicate a superior broad-spectrum labeling efficacy of aziridine-type β-L-arabinofuranosyl scaffolds towards different isozymes endowed with β-L-arabinofuranosidase activity. In contrast, epoxide-based probes presented slower labeling of *Bt*GH146 and significantly reduced inhibition efficacy towards rHypBA1, however their labeling abilities could be modulated by extending the ABP incubation time. It is anticipated that aziridine-type tools will be instrumental in broad-spectrum labeling of low-abundance β-L-arabinofuranosidases in complex biological samples, while epoxide-type β-L-*arabino*furanosyl cyclitols might find application because of their specificity as probing tools for discovery of distinct enzyme families.

4.4 Acknowledgements

Nicholas McGregor from the University of York, UK, is kindly acknowledged for the kinetic assessment experiments, and for the crystallographic studies with inhibitor **1** presented in this thesis and his valuable discussions. Wendy Offen, Olga Moroz and Gideon Davies from the University of York, UK, are kindly acknowledged for the crystallographic studies with inhibitor **2** presented in this thesis and their valuable discussions. Lars Binkhorst is acknowledged for his contribution to the synthesis of aziridine inhibitor **2** in the context of his MSc internship. Akihiro Ishiwata from RIKEN, Japan, is kindly acknowledged for providing *p*-nitrophenyl-β-L-arabinofuranoside[31] for kinetic measurements.

4.5 Experimental methods

4.5.1 Biochemical experiments

Materials

All chemicals were purchased from Sigma Aldrich unless otherwise specified. Tested β-L-arabinofuranosyl-configured compounds were synthesised as described in the supplemental synthetic methods and charaterisation section.

Recombinant enzyme production

r*Bt*GH146 – r*Bt*GH146 was synthesised and cloned into pET28a(+) with an *N*-terminal TEV protease-cleavable His6 tag by Genscript (Hong Kong). The enzyme was produced in transformed BL21(DE3) gold *E. coli* (Agilent) grown at 37 °C in autoinduction medium (1% tryptone, 0.5% yeast extract, 25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 0.05% glucose, 0.5% glycerol, 0.2% lactose) overnight. Cells were collected by centrifugation at 4000 *x g* for 15 minutes, resuspended in buffer A (50 mM NaPi, 300 mM NaCl, 20 mM imidazole, pH 7.5) and lysed by a single pass through a French press. Lysate was clarified by centrifugation at 40000 *x g* for 15 minutes then passed through a 5 mL histrap FF crude column (GE Healthcare). The major UV-active fractions eluted during a 10 CV gradient from 20-500 mM imidazole were pooled, concentrated to ~10 mg/mL using a 30 kDa MWCO centrifugal concentrator and purified using a Superdex 200 column (GE Healthcare) equilibrated with 20 mM HEPES pH 7.5 (SEC buffer). Eluted r*Bt*GH146 was pooled, concentrated to ~10 mg/mL and frozen at -80 °C.

rHypBA1 – rHypBA1 was produced and purified as previously described.[14]

Enzyme labelling detected by intact mass spectrometry

r*Bt*GH146 was diluted to 1 mg/mL in SEC buffer. Compounds **1-7 (**0.1 mM final concentration) were added to the enzyme solution and the reactions were incubated at 37 °C for different time periods.

2 μL samples were taken after 1, 6, and 24 hours of incubation for **1,** or 1 and 16 hours for **2-7**. These samples (2 μL) were then diluted with 48 μL of 1% formic acid, 10% acetonitrile, and analysed as described previously.[22]

rHypBA1 was diluted to 0.1 mg/mL in 50 mM NaPi (pH 4.5) and compounds **1-7** (0.1 mM final concentration) were added. 10 uL samples were taken after 10, 60, and 1300 minutes of incubation for **1** or 60 and 960 minutes for **2-7**. These 10 uL samples were diluted with 10 uL of 1% formic acid, 10% acetonitrile and stored at -20 °C until intact MS could be analysed in the same manner as for the r*Bt*GH146 samples.

Enzyme Inhibition Kinetics

rHypBA1 – rHypBA1 (20 mg/mL stored at -80 °C) was freshly thawed and diluted to 1 mg/mL in 50 mM Na-acetate (pH 4.5) supplemented with 1 mM DTT (assay buffer). Inhibitors **1**, **2**, **4** and **7** were dissolved in water at 5 mM concentration and used to prepare a dilution series from 1 mM to 16 μM in the assay buffer alongside a buffer control without inhibitor. A working enzyme solution was prepared at 20 μg/mL in assay buffer. To 35 μL of pre-warmed inhibitor solution was added 35 μL of enzyme solution and the inhibition reactions were incubated at 37 °C. Aliquots (7.5 μL) of these inactivation mixtures were removed at time intervals (5, 10, 20, 30, 40, and 60 minutes of incubation), and diluted with 142.5 μL of pre-warmed 0.25 mM 4-nitrophenyl-β-L-arabinofuranoside[31] in the assay buffer and next incubated at 37 °C. Aliquots (40 μL) of the resulting samples were taken from each substrate hydrolysis reaction at defined time intervals (1, 3, and 8 minutes of incubation) and mixed with 40 μL of stop solution (200 μM Na_2CO_3) in a 384-well plate. Absorbance change at 405 nm wavelength (A405) was read using an Epoch plate reader (Biotek). Hydrolysis rates were determined as the slope of a linear fit of A405 *versus* time. Slope values were converted into rates using a 40 μL calibration series of 4-nitrophenol in assay buffer mixed with 40 μL of stop solution. To account for slow enzyme activity loss in the no inhibitor control, rates were converted into residual activity by dividing each measured rate by the inhibitor-free hydrolysis rate measured at that incubation point. Residual activities were then plotted against incubation time using OriginPro graphing software (OriginLab) for each inhibitor concentration and fitted with exponential decay curves having y offset (y_0) values fixed at 0 (with the exception of the uninhibited rates where $y_0 = 1$). Extracted apparent decay constant (k_{app}) values were then plotted against the concentration of compound **1**, **2**, **4** and **7** in each inhibition reaction with error estimates taken as the standard error from exponential decay fit. Since no inflection was observed in the k_{app} *vs* [I] curve, an error-weighted linear fit was performed to determine k_i/K_I.

r*Bt*GH146 – r*Bt*GH146 (20 mg/mL stored at -80 °C) was freshly thawed and diluted to 1 mg/mL in 50 mM NaPi pH 7.0 with the assay buffer (1 mM DTT, 20 mM HEPES (pH 7.5) supplemented with 200 mM NaCl). Inhibitors **1**, **2**, **4** and **7** were dissolved in water at 5 mM concentration and used to prepare a dilution series from 1 mM to 16 μM in the assay buffer alongside a buffer control without

inhibitor. A working enzyme solution was prepared at 20 μg/mL in assay buffer. Next steps followed the above described procedure for HypBA1.

Fluorescent labeling and SDS-PAGE of recombinant HypBA1 and *Bt*GH146 by ABP 4

ABP **4** (10 mM stock, 100% w/v DMSO) was diluted with Milli-Q water to prepare a 1 mM ABP working solution (10% w/v). The enzyme stock (HypBA1: 20 mg/mL HypBA1 in 20 mM MOPS, pH 7.0, 1 mM DTT; *Bt*GH146: 1 mg/mL in 20 mM MOPS pH 7.5) was thawed on ice. The assay buffer for rHypBA1 was prepared by mixing 500 μL of 1 M NaOAc, 50 μL of 1 M DTT and 9.45 mL of deioinised water. For *Bt*GH146 labeling, the assay buffer (pH 7.0 NaPi, 1 mM DTT) was prepared with deionized water. The enzyme stock (10 μL) was diluted to 0.01 mg/mL enzyme concentration with the assay buffer (990 μL assay buffer). Aliquots (100 μL) of the enzyme working solution was heated to 95 °C for 5 minutes to generated an inactivated enzyme solution. To both the working and inactivated enzyme solutions (90 μL each), ABP working solution (10 μL) was added to afford final 0.1 mM ABP concentration. The samples were incubated at 37 °C. At regular time intervals (recorded contemporaneously: 0, 10, 60, 240 minutes), 15 μL samples were taken and mixed with 5 μL of 4x SDS-PAGE loading buffer (40% glycerol, 4% SDS, 250 mM Tris-HCl, pH 6.8, 10% 2-mercaptoethanol, 0.2 mg/mL bromophenol blue). The samples were immediately heated to 95 °C for 2 minutes, then stored frozen (-20 °C) until ready for analysis. Wet slab-gels were scanned on fluorescence using the Typhoon FLA 9500 (GE Healthcare) at λ_{EX} = 635 nm and λ_{EM} ⩾665 nm for Cy5-tagged ABP **4**.

Detection sensitivity of fluorescent ABPs 4, 5 and 9 towards r*Bt*GH146

To prepare for labeling, recombinant *Bt*GH146 stock (1 mg/mL in 20 mM MOPS pH 7.5) was diluted with assay buffer (for ABPs **4**: 20 mM HEPES 150 mM NaCl, pH 7.5; for ABPs **5** and **9**: PBS 150 mM NaCl, pH 7.4) to varying enzyme concentrations (111, 11, 1, 0.1, 0.01 μg/mL). ABPs **4**, **5** and **9** stocks (10 mM in DMSO) were diluted with assay buffer to 1 mM and 0.1 mM ABP concentration (for **4** and **5**) or to 100 μM and 10 μM (for ABP **9**). For labeling, 9 μL of each of the enzyme working solutions were loaded in separate epps on ice, to yield a total final amount of 1000, 100, 10, 1, 0.1 ng *Bt*GH146 per reaction. The enzyme solution was thus incubated with 1 μL ABP (**4**, **5**, or **9**) at 37 °C for 1 h (Figure 4.4 and 4.6). Samples were then denatured with 2.66 μL sample buffer (5x Laemmli buffer, containing 50% (v/v) 1M Tris-HCl pH 6.8, 50% (v/v) glycerol, 10% (w/v) Dithiothreitol (DTT), 10% (w/v) sodium dodecyl sulphate (SDS), 0.01% bromophenol blue) and heated at 98 °C for 5 minutes. Proteins were resolved by electrophoresis in sodium dodecylsulfate (SDS-PAGE) 10% polyacrylamide gels and wet slab gels were scanned as described above. Wet slab-gels were scanned on fluorescence using the Typhoon FLA 9500 (GE Healthcare) at λ_{EX} = 635 nm and λ_{EM} ⩾665 nm for Cy5-tagged ABP **4** and **9**, and at λ_{EX} = 473 nm and λ_{EM}⩾510 nm for greenBODIPY-tagged ABP **5**.

Time-dependent fluorescent labeling of r*Bt*GH146 by aziridine-type ABP 8

Recombinant *Bt*GH146 stock (1 mg/mL in 20 mM MOPS pH 7.5) was diluted with assay buffer (PBS 150 mM NaCl, pH 7.4) to afford 11 µg/mL working enzyme solution. ABP **8** (10 mM in DMSO) was diluted with the assay buffer to 100 µM and 10 µM ABP concentration. For labeling, 9 µL of enzyme working solution (100 ng total *Bt*GH146 in final reaction mixture) was loaded in separate epps on ice, and 1 µL ABP working solution was added in each epp. Reactions were incubated at 37 °C while shaking. At time intervals (3, 15, 30, 60, 120, 180, 240 minutes), samples were denatured with 5x Laemmli buffer, boiled at 98 °C for 5 minutes and subjected to SDS-PAGE and fluorescence scan as described above.

4.5.2 Crystallographic Collection Data and Refinement Statistics

Crystallisation and Soaking of 1 with r*Bt*GH146 or rHypBA1

r*Bt*GH146 – Crystals of r*Bt*GH146 were grown from a 1.2 µL sitting drop containing a 1:1 mixture of 10 mg/mL *Bt*GH146 in SEC buffer (pre-incubated at 37 °C for 16 hours following supplementation with 200 µM **1**) with 20% PEG3350, 0.2 M ammonium formate, 2.5% (170 mM) arabinose, 0.1 M pH 6.0 MES. Crystals were transferred into mother liquor supplemented with 20% PEG400 prior to cryocooling in liquid N_2. Diffraction data were collected out to 1.41 Å on beamline I03 at Diamond Light Source (Harwell) and processed using the Xia2[34] pipeline with Dials.[35]

rHypBA1 – rHypBA1 protein labelled with **1** was used for crystallisation. After purification by Ni affinity chromatography, the protein sample was incubated with 0.1 mM inhibitor **1** at 37 °C for 1 hour in 50 mM Na-acetate (pH 4.5) buffer. The labelled sample was further purified by size exclusion chromatography using HiLoad 16/60 Superdex 200 pg column in 10 mM HEPES-NaOH (pH 7.5) and 150 mM NaCl. The purified sample was concentrated and buffer-exchanged to 10 mM HEPES-NaOH (pH 7.5) with an ultrafiltration centrifugal membrane unit (Vivaspin Turbo 15, 10 kDa molecular weight cutoff; Sartorius Stedim Biotech, Göttingen, Germany). Crystals were grown at 20 °C for 2 days using the sitting drop vapor diffusion method by mixing 0.5 µL of protein solution containing 40 mg/mL rHypBA1:**1** with equal volume of a reservoir solution, which contained 0.7 M Na-citrate and 0.1 M MES-NaOH, pH 6.5. For cryoprotectant, 15% (w/v) trehalose was used.

Crystallisation and Soaking of 2 with r*Bt*GH146 or rHypBA1

r*Bt*GH146 – Crystals of r*Bt*GH146 was purified according to the protocols priorly described (see section *Recombinant enzyme production*) and buffer-exchanged into 20 mM MES pH 6.0 at 22 mg/ml. It was initially co-crystallised with 2.5% (w/v) arabinose, as crystals grown without arabinose were found to be unsuitable for X-ray analysis, using an Oryx robot (Douglas instruments) with 150 nL protein solution plus 150 nL reservoir solution in 96-well format plates (MRC 2-well crystallization microplate, Swissci, Switzerland) equilibrated against 60 µL reservoir solution which consisted of 19% PEG 3350, MES/Bis-Tris-propane mix with pH 6.0- 6.25, 0.2 M ammonium formate. The crystals

obtained were in the form of clusters of plates; these were used for microseed matrix screening (MMS)[36], which was carried out using the Oryx robot according to published protocols.[37,38] Briefly, the crystals were vortexed for 1 min with a Seed Bead (Hampton Research) with 50 µL well solution added. Crystals suitable for soak experiments were obtained after several rounds of seeding using 150 nL protein solution plus 50 nL seeding stock plus 100 nL well solution, with the latter consisting of 15-17% PEG 3350, MES pH 6.25-6.5, 0.2 M ammonium formate. They were soaked with **2** for 2-4 hours as follows: 2 µL of a 5 mM solution of **2** were diluted with 8 µL mother liquor of the corresponding crystallisation condition giving a final concentration of 1 mM; approximately 0.2-0.5 µL of this solution were added to the drop containing crystals.

rHypBA1 – rHypBA1 crystals were grown by the sitting drop vapour diffusion method, using protein at 37 mg/ml in 10 mM HEPES pH 7.5, in a volume ratio of 0.6:0.5 with the well solution, which consisted of 0.7 M sodium citrate, 0.1 M 2-(*N*-morpholino)ethanesulfonic acid (MES) pH 6.5, 10 mM dithiothreitol. The crystal was soaked with **2** which had been dissolved at 20 mM in water, diluted with well solution to 4 mM, and added to the protein drop to a final concentration of 2 mM. The crystal was fished after 22 hours, via a cryoprotectant solution comprised of the well solution components supplemented with 25% (v/v) glycerol, into liquid nitrogen.

X-ray data collection and structure solution

Diffraction data for rHypBA1:**1** and r*Bt*GH146:**1** complexes were collected at the Photon Factory of the High Energy Accelerator Research Organisation (KEK-PF, Tsukuba, Japan) and processed using XDS.[39] The structure of the r*Bt*GH146:**1** complex was solved by molecular replacement using Phaser[40] with the known complex with L-arabinofuranose (PDBID: 5OPJ) as the search model. The resulting solution showed clear density for the bound ligand within the enzyme active site. Ligand coordinates and dictionaries were generated using jLigand[41] and built into the model using Coot,[42] followed by alternating rounds of manual model building and density refinement using Coot and REFMAC5[43] within the CCP4 suite.[44]

The structure of the HypBA1:**1** complex was determined by molecular replacement using Molrep[45] with the rHypBA1 structure (PDBID: 3WKX) as the search model. Ligand restraint generation and crystallographic refinement were carried out as above.

Diffraction data for rHypBA1:**2** and r*Bt*GH146:**2** complexes were collected at Diamond Light Source (UK) to 2.09 Å and 2.0 Å resolution, respectively, processed with DIALS[46] as incorporated in Xia2[34] pipeline and scaled with *AIMLESS*. For r*Bt*GH146:**2**, the space group was $P2_12_12_1$, with unit cell dimensions of a=93.84, b= 98.549, c= 196.584 Å, and $\alpha=\beta=\gamma=90.0°$. All computations were carried out using programs from the CCP4 suite[47] unless otherwise stated, as incorporated in CCP4Cloud user interface.[48] The structure was solved using MOLREP[45] with PDB entry 6YQH as an initial model and refined with cycles of REFMAC[43] alternating with manual model correction in COOT.[49] The

dictionaries for the ligand and its covalent linkage were created in JLigand[41], and the ligand was built using COOT.

For rHypBA1:**2**, the space group was $P3_2\ 2\ 1$, with unit cell dimensions of 77.71, 77.71, 252.09 Å and $\alpha=\beta=90.0°$ $\gamma=120.0°$. As the structure was sufficiently isomorphous with PDB entry 7DIF it was solved using REFMAC[43] with the protein chain of this model. The ligand was built and dictionary restraints generated using *AceDRG*.[50] The structure was refined by performing cycles of manual building in COOT[49], interspersed with refinement using REFMAC5. The programs were run in the *CCP*4*i*2 suite.[51].

4.5.3 Chemical synthesis

General experimental details

All reagents were purchased from Acros, Sigma Aldrich, Biosolve, VWR, Fluka, Merck and Fisher Scientific and used as received unless stated otherwise. Tetrahydrofuran (THF), dichloromethane (DCM), *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 F254) 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 C18 column (Gemini, 4.6 mm x 50 mm, 5 μm particle size, Phenomenex), coupled to a LCQ Advantage Max (Thermo Finnigan) ion-trap spectrometer (ESI^+). Reversed-phase HPLC purification was performed on an Agilent Technologies 1200 series instrument equipped with a semi-preparative column (Gemini C18, 250 mm x 10 mm, 5 μm particle size, Phenomenex) or with HILIC column (VP 250/10 NUCLEODUR HILIC, 5 μm particle size, Macherey-Nagel). The applied buffers were: (A) aqueous 50 mM NH_4HCO_3, and (B) CH_3CN. For gel filtration, an ÄTKA explorer (GE Healthcare) equipped with Bio-Gel P-2 resin (Bio-Rad) was used, with eluting solution of AcOH (1%) in MilliQ. ^1H NMR and ^{13}C NMR spectra were recorded on a Brüker AV-400 (400 and 101 MHz respectively) Brüker AV-500 (500 MHz and 126 MHz) or a Brüker DMX-600 (600 and 151 MHz respectively) spectrometer in the given solvent. Chemical shifts (δ) are given in ppm relative to tetramethylsilane (TMS) as internal standard (0 ppm in ^1H NMR with $CDCl_3$) or the residual signal of the deuterated solvent. Coupling constants (J) are given in Hz. All given ^{13}C-NMR spectra are proton decoupled. The following abbreviations are used to

describe peak patterns when appropriate: s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), Ar (aromatic), C_q (quarternary carbon). 2D NMR experiments (COSY, HSQC) were carried out to assign protons and carbons of the new structures. High-resolution mass spectrometry (HRMS) analysis was performed with a LTQ Orbitrap mass spectrometer (Thermo Finnigan), equipped with an electronspray ion source in positive mode (source voltage 3.5 kV, sheath gas flow 10 mL/min, capillary temperature 250 °C) with resolution R = 60000 at *m*/*z* 400 (mass range *m*/*z* = 150 – 2000) and dioctyl phthalate (*m*/*z* = 391.28428) as a "lock mass". The high-resolution mass spectrometer was calibrated prior to measurements with a calibration mixture (Thermo Finnigan).

Experimental Procedures and Characterisation Data of Products

Known compounds **13**[22] and 1-azido-8-trifluoromethylsulfonyloctane[29] were synthesised following procedures previously described and their spectroscopic data are in agreement with those previously reported.

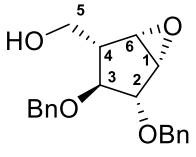
General Procedure A: Birch reaction

Ammonia (5-10 mL) was condensed in an oven-dried flask cooled to -60 °C. Freshly-cut $Na_{(s)}$ wires (22 eq) were added to the liquid ammonia and stirred with a glass-coated stirring bar at -60 °C until complete dissolution (20 minutes). Next, dibenzylated compound **12**, **15** or **17** (20-50 μmol, 1 eq) was dissolved in dry THF (1 mL) and ^{t}BuOH (28 eq). The resulting solution was added drop-wise along the cold walls of the flask containing liquid ammonia at -60 °C. Upon addition, the solution remained dark blue and was stirred for additional 45 min at -60 °C. Next, the reaction was quenched (with 1 mL milli-Q H_2O for aziridine-based compounds **12** and **15**; with 28 eq of 3.1 M NH_4Cl aqueous solution for epoxide-based compound **17**). The reaction mixture was allowed to attain gradually room temperature over 2 h. Crude was co-evaporated thrice with milli-Q water, and purified as described for each distinct product.

General Procedure B: Amide coupling of reporter tag

To a solution of tag-COOH (1 eq) in dry DMF (0.154 – 0.176 M), PFP-TFA (1.5 eq) and DIPEA (3.5 eq) were added under stirring. After 2 h, the reaction was quenched with Milli-Q H_2O (2.1 eq), and the mixture was added at room temperature to a solution of the desired β-L-*arabino*furanosyl amine **3** or **7** (0.77 eq), either dissolved in dry DMF (250 μL) or used as dry solid. Additional DIPEA (2.5 eq) was added to the reaction mixture. Upon full conversion (as monitored by LC-MS), the solution was concentrated to dryness and the crude purified by reverse-phase HPLC or by silica gel chromatography as described. Volatiles were removed by lyophilisation.

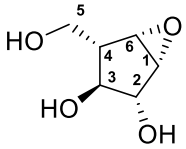
Compound 11



To a buffered solution of NaH_2PO_4/ Na_2HPO_4 (1:1 ratio, 13 mL) under vigorous stirring at 0 ˚C, cyclopentene **13** (200 mg, 0.64 mmol) dissolved in DCM (6 mL) was added. Afterwards, a solution of *m*-CPBA (200 mg, 1.16 mmol) in DCM (6 mL) was

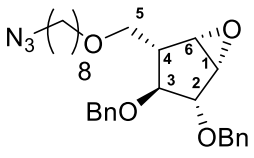
added and the reaction was stirred at 4 ˚C overnight. As TLC indicated the presence of starting material after 19 h, more *m*-CPBA (105 mg, 0.61 mmol) was added. After stirring the reaction mixture for 3 additional days, the layers were separated, and the aqueous phase was washed with EtOAc (2x 100 mL). The combined organic layers were dried over Mg_2SO_4, filtered and concentrated under reduced pressure. Purification of the crude by automatic silica gel column chromatography (25 g silica, 0→70% EtOAc in pentane) afforded the title compound 3 as major product (155 mg, 74%). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.45 – 7.27 (m, 10H, 10CH Ar), 4.79 (d, *J* = 11.9 Hz, 1H, C*H*HPh), 4.69 (d, *J* = 11.3 Hz, 2H, CH_2Ph), 4.55 (d, *J* = 11.6 Hz, 1H, C*H*HPh), 4.11 (dd, *J* = 5.3, 1.3 Hz, 1H, CH-3), 3.91 (dd, *J* = 10.6, 5.4 Hz, 1H, C*H*HOH), 3.84 (dd, *J* = 10.6, 7.5 Hz, 1H, C*H*HOH), 3.56 (m, 3H, H-1, H-6, H-3), 2.23 (tdd, *J* = 7.2, 5.4, 1.4 Hz, 1H, H-4). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 138.2, 136.9 (2C_q Ar), 128.7, 128.6, 128.1, 128.0, 127.9 (10CH Ar), 85.7 (CH, C-2), 81.0 (CH, C-3), 72.7, 71.9 (2CH_2Ph), 62.2 (CH_2OH), 55.1, 54.0 (2CH, C-1, C-2), 46.8 (CH, C-4). HR-MS: *m/z* calcd for $C_{20}H_{22}NaO_4^+$: 349.3812 [*M*+Na]$^+$; found 349.1420.

Compound 1 (DL063)



Compound **11** (200 mg, 640 µmol) was dissolved in MeOH (6 mL) and thoroughly degassed under argon atmosphere. $Pd(OH)_2$ (20% on carbon, 2.6 mg, 18 µmol) was added to the solution and the mixture was further degassed with argon. H_2 gas was bubbled through the reaction mixture with a balloon at rt overnight while stirring. The reaction mixture was then bubbled with argon and filtered over Celite®, and the filter was washed several times with MeOH. Filtrate was concentrated under reduced pressure and purified by silica gel column chromatography (0%→20% MeOH in DCM), and the purified product was freeze-dried to afford final product epoxide **1** as white solid (2.5 mg, 18%). ^{1}H NMR (500 MHz, MeOD) δ = 3.94 (dd, *J* = 6.0, 1.4 Hz, 1H, H-2), 3.81 (dd, *J* = 10.6, 4.7 Hz, 1H, C*H*HOH), 3.71 – 3.63 (m, 1H, C*H*HOH), 3.55 (dd, *J* = 3.2, 1.2 Hz, 1H, H-3), 3.46 (dd, *J* = 3.0, 1.6 Hz, 1H, H epoxide), 3.18 (dd, *J* = 7.5, 6.1 Hz, 1H, H epoxide), 1.97 (dddd, *J* = 9.3, 7.6, 4.7, 1.3 Hz, 1H, CH-4). ^{13}C NMR (126 MHz, MeOD): δ = 80.4 (CH-3), 76.0 (CH-4), 61.7 (CH2), 57.6, 54.9 (CH-1, CH-2), 49.6 (CH-5).

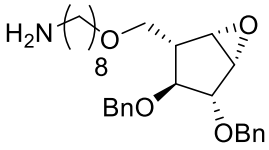
Compound 14



Epoxide **11** (120 mg, 0.369 mmol) was dissolved in dry THF (0.700 mL, 0.21 M) under $N_{2(g)}$ atmosphere, and cooled to 0 °C. NaH (60% wt, 7.73 mg, 0.193 mmol) was added and the solution was stirred for 15 min at 0 °C. Next, the solution was cooled to -20 °C, and freshly prepared azido-octyltriflate[29] solution (0.383 M in THF, 0.500 mL, 0.193 mmol) was added dropwise along the cold flask walls. The solution was gradually warmed up to rt. After 5 h, TLC showed full consumption of the starting material. The reaction was quenched on ice with water, diluted with EtOAc (20 mL) and washed with water, and brine. Organic layer was dried over Na_2SO_4, filtered and volatiles were removed under reduced pressure. Purification by silica

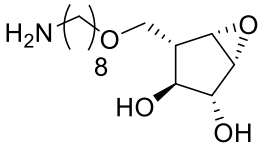
gel column chromatography (pentane/EtOAc 20:1→7:1) afforded titled product as an oil (149 mg, 0.310 mmol, 84%). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.44 – 7.23 (m, 10H, CH Ar), 4.77 (d, *J* = 11.9 Hz, 1H, ½ CH_2Ph), 4.71 – 4.62 (m, 2H, CH_2Ph), 4.54 (d, *J* = 11.6 Hz, 1H, ½ CH_2Ph), 4.09 (dd, *J* = 5.3, 1.4 Hz, 1H, H-2), 3.61 – 3.52 (m, 4H, H-1/H-6/CH_2-5), 3.45 (td, *J* = 6.6, 2.4 Hz, 2H, OCH_2linker), 3.38 (dd, *J* = 7.1, 5.3 Hz, 1H, H-3), 3.24 (t, *J* = 7.0 Hz, 2H, CH_2N_3), 2.29 (dddd, *J* = 9.0, 6.9, 5.4, 1.5 Hz, 1H, H-4), 1.58 (td, *J* = 7.5, 5.4 Hz, 5H, 5 CH*H*linker), 1.42 – 1.21 (m, 11H, 11 CH*H*linker) ppm. ^{13}C NMR (126 MHz, $CDCl_3$): δ = 138.5, 138.2 ($2C_q$), 128.6, 128.5, 128.0, 127.9, 127.8 (6CH Ar), 85.8 (CH, C-2), 81.6 (CH, C-3), 72.6 (OCH_2linker), 71.8, 71.5 ($2CH_2$Ph), 69.6 (CH_2, C-5), 55.3 (CH, epoxide), 54.5 (CH, epoxide), 51.6 (CH_2N_3), 45.2 (CH, C-4), 29.8, 29.4, 29.2, 28.9, 26.8, 26.2 ($6CH_2$, linker) ppm. HR-MS (ESI): *m/z* calcd for $C_{28}H_{41}N_4O_4^+$: 497.31223 [*M*+NH_4]$^+$, found 497.31214; *m/z* calcd for $C_{28}H_{37}N_3O_4Na^+$ 502.26763 [*M*+Na]$^+$, found 502.26752.

Compound 17



Azide **14** (149 mg, 0.311 mmol) was dissolved in dry acetonitrile (6.22 mL, 0.05 M) and transferred into a microwave tube. To this solution, water (56 µl, 3.11 mmol) and polymer-bound PPh_3 (163 mg, 0.622 mmol) were added at rt. The reaction mixture was heated to 70 °C, and stirred for 5 h. After 5 h, TLC showed full conversion; hence the solution was cooled to room temperature and filtered over a sintered glass filter. The filter was washed with acetonitrile (3x 13 mL), volatiles were removed under reduced pressure and the crude was used in the next step without further purification (133 mg, 0.292 mmol, 94%). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.46 – 7.22 (m, 10H, CH Ar), 4.77 (d, *J* = 11.9 Hz, 1H, ½ CH_2Ph), 4.70 – 4.62 (m, 2H, CH_2Ph), 4.54 (d, *J* = 11.6 Hz, 1H, ½ CH_2Ph), 4.09 (dd, *J* = 5.3, 1.4 Hz, 1H, H-2), 3.62 – 3.52 (m, 4H, H-1/H-6/CH_2-5), 3.44 (td, *J* = 6.7, 2.3 Hz, 2H, OCH_2linker), 3.38 (dd, *J* = 7.1, 5.3 Hz, 1H, H-3), 2.72 – 2.60 (m, 2H, C*H*$_2$$NH_2$), 2.29 (dddd, *J* = 9.0, 7.0, 5.4, 1.5 Hz, 1H, H-4), 1.82 (s, 2H, NH_2), 1.63 – 1.51 (m, 2H, CH_2linker), 1.49 – 1.37 (m, 2H, 2CH*H*linker), 1.36 – 1.21 (m, 10H, 8CH*H*linker, grease) ppm. ^{13}C NMR (126 MHz, $CDCl_3$): δ = 138.4, 138.1 ($2C_q$), 128.7, 128.6, 128.4, 127.9, 127.7 (6CH Ar), 85.8 (CH, C-2), 81.6 (CH, C-3), 72.6 (OCH_2linker), 71.8, 71.5 ($2CH_2$Ar), 69.5 (CH_2, C-5), 55.3 (CH, epoxide), 54.5 (CH, epoxide), 45.2 (CH, C-4), 42.2 (CH_2NH_2), 33.6, 29.7, 29.5, 26.9, 26.8, 26.2 ($6CH_2$, linker) ppm. HR-MS (ESI): *m/z* calcd for $C_{28}H_{40}NO_4^+$: 454.29519 [*M*+H]$^+$, found 454.29501.

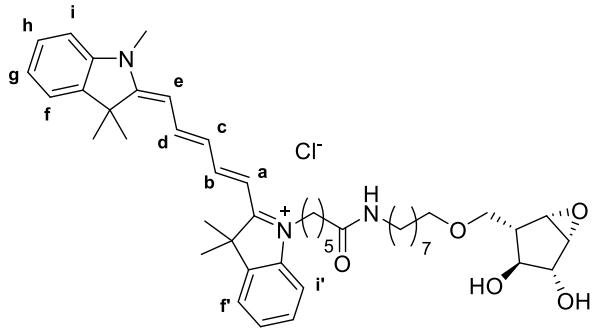
Compound 3 (VB-C049)



Ammonia was condensed in an oven-dried flask at -60 °C, and sodium (189 mg, 8.21 mmol) was freshly cut and added to the liquid ammonia solution at -60 °C. The solution was stirred at this temperature for 40 min. Amine **17** (133 mg, 0.292 mmol) was co-evaporated thrice with distilled toluene and then dissolved in dry THF (5 ml) and ^{t}BuOH (1 ml, 10.46 mmol). The resulting solution was added drop-wise along the cold walls of the

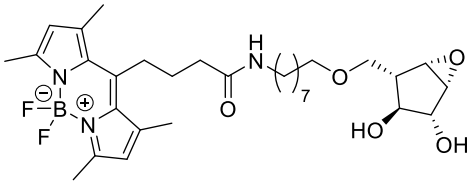
sodium/ammonia solution at -60 °C, and the reaction was stirred at this temperature. After 45 min, the reaction was quenched with an aqueous solution of 3.11 M NH_4Cl (3 ml, 9.35 mmol) and it was allowed to warm up to rt. Crude was co-evaporated thrice with Milli-Q water and desalted by size-exclusion chromatography over Bio-Gel P-2 resin eluting with 1% AcOH in water. The desalted fractions were purified by reverse-phase HPLC chromatography equipped with HILIC column (eluent A: 50 mM NH_4HCO_3 in water; eluent B: CH_3CN). The isolated product was lyophilised upon NMR analysis, to yield titled compound as pale yellow solid (79.5 mg, 0.245 mmol, 84%). ^{1}H NMR (500 MHz, MeOD): δ = 3.95 (dd, *J* = 6.0, 1.5 Hz, 1H, H-2), 3.64 (dd, *J* = 9.3, 4.6 Hz, 1H, C*H*H-5), 3.56 (t, *J* = 9.4 Hz, 1H, CH*H*-5), 3.52 (t, *J* = 6.5 Hz, 1H, OC*H*H_2linker), 3.51 – 3.47 (m, 2H, H-epoxide/ ½ OCH*H*linker), 3.46 (dd, *J* = 3.2, 1.5 Hz, 1H, H-epoxide), 3.19 (dd, *J* = 7.6, 6.0 Hz, 1H, H-3), 2.89 (t, *J* = 7.7 Hz, 2H, C*H*$_2$$NH_2$), 2.06 (dddd, *J* = 9.5, 7.6, 4.6, 1.3 Hz, 1H, H-4), 1.90 (s, 3H, OAc), 1.71 – 1.55 (m, 4H, 2 CH_2linker), 1.49 – 1.35 (m, 8H, 4 CH_2linker), 1.22 (s, 2H) ppm. ^{13}C NMR (126 MHz, MeOD): δ 80.2 (CH), 76.1 (CH), 72.3 (CH_2), 70.5 (CH_2), 57.8 (CH, epoxide), 55.1 (CH, epoxide), 47.4 (CH, C-4), 40.8 (CH_2NH_2), 31.1 (CH), 30.6, 30.3, 30.2, 28.9, 27.4, 27.1 (6 CH_2linker) ppm. HR-MS (ESI): *m/z* calcd for $C_{14}H_{28}NO_4^+$: 274.20128 $[M+H]^+$, found 274.20130.

Compound 4 (VB-C034)



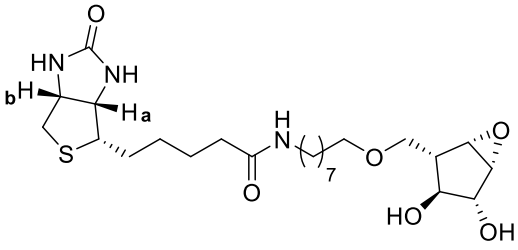
Titled product was obtained as a blue powder (2.2 mg, 2.9 μmol, 7.4% yield) starting from amine **3** (10 mg, 39 μmol) according to general procedure B, upon purification by reverse-phase HPLC chroamtography (A: 50 mM NH_4OAc in H_2O, B: CH_3CN). ^{1}H NMR (850 MHz, MeOD): δ = 8.29 – 8.20 (m, 2H, H-d, H-b), 7.51 – 7.48 (m, 2H, H-i, H-i'), 7.41 (qd, *J* = 7.4, 1.2 Hz, 2H, H-f, H-f'), 7.34 – 7.23 (m, 4H), 6.63 (t, *J* = 12.4 Hz, 1H), 6.28 (dd, *J* = 13.7, 10.1 Hz, 2H), 4.11 (t, *J* = 7.5 Hz, 1H, CH_2N^+), 3.94 (dd, *J* = 6.0, 1.5 Hz, 1H, H-2), 3.64 – 3.61 (m, 4H, CH_3N, C*H*H), 3.55 – 3.52 (m, 1H, CH*H*), 3.52 – 3.43 (m, 4H, CH_2, H-1, H-6), 3.18 (dd, *J* = 7.7, 6.0 Hz, 1H, H-3), 3.12 (t, *J* = 7.2 Hz, 2H, C*H*$_2$NHCO), 2.20 (t, *J* = 7.3 Hz, 2H, CH_2C=O), 2.09 – 2.01 (m, 1H, H-4), 1.83 (p, *J* = 7.8 Hz, 2H, C*H*$_2$$CH_2N^+$), 1.73 (s, 6H, 2$CH_3$), 1.69 (p, *J* = 7.4 Hz, 2H, C*H*$_2$$CH_2$C=O), 1.63 – 1.54 (m, 2H, 2CH*H*linker), 1.49 – 1.42 (m, 4H, 4CH*H* linker), 1.39 – 1.27 (m, 10H, 10CH*H* linker), 1.22 (s, 3H, CH_3), 0.93 – 0.85 (m, 2H, 2CH*H* linker) ppm. ^{13}C NMR (214 MHz, MeOD): δ = 175.7, 175.4, 174.7 (3C_q), 155.53 (CH Ar), 144.3, 143.6, 142.6, 142.5 (4C_q), 129.8, 126.6, 126.3, 126.2, 123.4, 123.3, 112.1, 111.8, 104.4, 104.3 (11CH Ar), 80.2 (CH, C-2), 76.1 (CH, C-3), 72.4 (CH_2), 70.5 (CH_2) 57.8 (CH, C-1/C-6), 55.1 (CH, C-1/C-6), 47.3 (CH, C-4), 44.8 (CH_2N^+), 40.4 (CH_2NH), 36.7 (*C*H_2C=O), 31.1 (CH_3N), 30.7, 30.5, 30.4, 28.2, 28.0 (6CH_2, linker), 27.9, 27.8 (2CH_3), 27.4, 27.2, 26.6 (3CH_2, linker) ppm. HR-MS (ESI): *m/z* calcd for $C_{46}H_{64}N_3O_5^+$: 738.48405 $[M]^+$, found 738.48412.

Compound 5 (VB-C063)



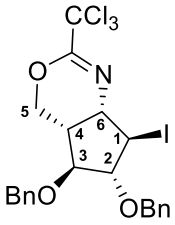
Titled product was obtained as an orange-red powder (14.1 mg, 23.9 μmol, 41% yield) according to general procedure B starting from amine **3** (30 mg, 59 μmol) with purification by silica gel chromatography (DCM/MeOH 99:1→95:5). ^{1}H NMR (500 MHz, MeOD): δ = 6.14 (s, 2H, 2H Ar), 3.93 (d, *J* = 6.0 Hz, 1H, H-2), 3.62 (dd, *J* = 9.4, 4.7 Hz, 1H, C*H*H-5), 3.57 – 3.50 (m, 1H, CH*H*-5), 3.51 – 3.41 (m, 4H, CH_2OC(5), H-1, H-6), 3.18 (q, *J* = 6.9, 6.3 Hz, 3H, CH_2NH, H-3), 3.07 – 2.97 (m, 2H, C$H_2CH_2CH_2C$=O), 2.47 – 2.43 (m, 12H, $4CH_3$), 2.38 (t, *J* = 7.1 Hz, 2H, CH_2C=O), 2.05 (td, *J* = 8.8, 4.5 Hz, 1H, H-4), 1.99 – 1.87 (m, 2H, CH_2CH_2C=O), 1.64 – 1.54 (m, 2H, 2CH*H* linker), 1.51 (t, *J* = 6.9 Hz, 2H, 2CH*H* linker), 1.43 – 1.32 (m, 10H, 8CH*H* linker) ppm. ^{13}C NMR (126 MHz, MeOD): δ = 174.8, 155.2, 147.2, 142.4, 132.6 ($5C_q$), 122.7 (2CH Ar), 80.2 (CH, C-2), 76.1 (CH, C-3), 72.4 (CH_2O linker), 70.5 (CH_2, C-5), 57.8 (CH, C-1/C-6, epoxide), 55.1 (CH, C-1/C-6, epoxide), 47.4 (CH, C-4), 40.5 (CH_2NH), 37.1 (*C*H_2C=O), 30.7, 30.5, 30.4, 29.3, 28.6, 28.0, 27.2 ($7CH_2$ linker), 16.6 ($4CH_3$), 14.4 ppm. HR-MS (ESI): *m/z* calcd for $C_{31}H_{47}BF_2N_4O_4Na^+$: 612.33963 [*M*+Na]$^+$, found 612.33922.

Compound 6 (VB-C079)



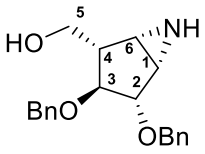
Amine **3** (9.89 mg, 36.3 μmol) was reacted according to general protocol B ensued by reverse-phase HPLC purification (A: 20 mM NH_4OAc in H_2O, B: CH_3CN), to afford the titled product as a white powder (4.27 mg, 8.5 μmol, 23% yield). After NMR analysis, the desired product was lyophilised thrice to remove volatiles. ^{1}H NMR (500 MHz, MeOD): δ = 4.49 (ddd, *J* = 7.9, 5.0, 0.9 Hz, 1H, H-b), 4.30 (dd, *J* = 7.9, 4.5 Hz, 1H, H-a), 3.94 (dd, *J* = 6.0, 1.5 Hz, 1H, H-2), 3.64 (dd, *J* = 9.3, 4.7 Hz, 1H, C*H*H-5), 3.59 – 3.44 (m, 5H, CH*H*-5, CH_2O linker, H-1, H-6), 3.24 – 3.10 (m, 4H, CHS, H-3, CH_2NHC=O), 2.93 (dd, *J* = 12.8, 5.0 Hz, 1H, C*H*HS), 2.71 (d, *J* = 12.7 Hz, 1H, CH*H*S), 2.24 – 2.15 (m, 2H, CH_2C=O), 2.06 (dddd, *J* = 9.7, 7.6, 4.6, 1.4 Hz, 1H, H-4), 1.79 – 1.55 (m, 6H, 6CH*H* linker), 1.53 – 1.31 (m, 12H, 12CH*H* linker) ppm. ^{13}C NMR (126 MHz, MeOD): δ = 176.0, 166.1 ($2C_q$), 80.2 (CH, C-2), 76.1 (CH, C-3), 72.4 (CH_2O, linker), 70.5 (CH_2, C-5), 63.4 (CH, C-a), 61.6 (CH, C-b), 57.8 (CH, C-1), 57.0 (CHS), 55.1 (CH, C-6), 47.4 (CH, C-4), 41.1 (CH_2S), 40.4 (CH_2NH), 36.8 (*C*H_2C=O), 30.7, 30.5, 30.4, 30.4, 29.8, 29.5, 27.9, 27.2, 27.0 ($9CH_2$, linker) ppm. HR-MS (ESI): *m/z* calcd for $C_{24}H_{42}N_3O_6S^+$: 500.27888 [*M*+H]$^+$, found 500.27884.

Compound 16



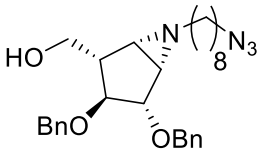
To a solution of cyclopentene **13** (2.73 g, 5.52 mmol) in $CHCl_3$ (138 mL) at 0 °C, NIS (310 g, 13.8 mmol) were added portion-wise under stirring. The solution was allowed to attain gradually rt. After 4 h, full conversion was observed and the solution was quenched with sat. aq. $NaHCO_3$ solution (70 mL), washed thrice with $CHCl_3$ (280 mL), dried over Na_2SO_4, filtered and concentrated to dryness. Crude was co-evaporated thrice with distilled toluene, and run through a short silica plug, yielding the desired product as an oil (3.2 g, 5.52 mmol, quant. yield). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.41 – 7.27 (m, 10H, 10CH Ar), 4.79 (d, *J* = 11.6 Hz, 1H, 1CH*H*Ph), 4.67 – 4.48 (m, 3H, 3CH*H*Ph), 4.46 (ddd, *J* = 4.6, 3.6, 1.1 Hz, 1H, H-1), 4.34 – 4.25 (m, 2H, CH_2-5), 4.22 – 4.17 (m, 2H, H-6, H-2), 3.79 (dd, *J* = 8.4, 4.5 Hz, 1H, H-3), 2.81 – 2.74 (m, 1H, H-4) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ = 128.7, 128.6, 128.2, 128.1 (10CH Ar), 92.7 (CH, C-1), 83.7 (CH, C-3), 72.9, 72.4 (2CH_2Ph), 65.7 (CH_2, C-5), 64.2 (CH, C-2), 39.0 (CH, C-4), 26.6 (CH, C-6) ppm. HR-MS (ESI): *m/z* calcd for $C_{22}H_{22}Cl_3INO_3^+$: 579.97045 [*M*+H]$^+$, found 579.97026.

Compound 12



Cyclic trichloroacetimidate **16** (3.2 g, 5.52 mmol) was dissolved in a solution of MeOH/DCM (1:1 v/v, 110 mL) and the solution was cooled to 0 °C. A solution of HCl (1.25 M in MeOH, 39.7 mL, 49.7 mmol) was added dropwise at 0 °C. Upon addition, the solution was allowed to slowly warm up to rt and was stirred over three days. Upon full conversion, the reaction was quenched with H_2O (0.5 mL). Afterwards, Amberlite IRA-67 was added under vigorous stirring until neutral pH. After 22 h full conversion was observed, and the reaction mixture was filtered over a glass filter and concentrated to dryness. Purification by silica gel chromatography (DCM/MeOH 99:1 + 0.05% TEA→98:2 + 0.05% TEA) afforded the titled product as a pale yellow oil (1.13 g, 3.47 mmol, 63% over 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.47 – 7.23 (m, 10H, 10CH Ar), 4.76 (d, *J* = 11.9 Hz, 1H, CH*H*Ph), 4.66 (q, *J* = 23.5, 11.7 Hz, 2H, 2CH*H*Ph), 4.54 (d, *J* = 11.5 Hz, 1H, CH*H*Ph), 4.12 (dd, *J* = 5.5, 2.6 Hz, 1H, H-2), 3.90 (dd, *J* = 10.7, 4.2 Hz, 1H, C*H*HOH), 3.79 (dd, *J* = 10.7, 6.0 Hz, 1H, CH*H*OH), 3.64 (dd, *J* = 7.5, 5.4 Hz, 1H, H-3), 2.80 – 2.54 (m, 2H, H-1, H-6), 2.24 – 2.16 (m, 1H, H-4) ppm. ^{13}C NMR (126 MHz, $CDCl_3$): δ = 138.5, 138.3 (2C_q), 128.5, 128.4, 128.0, 127.9, 127.8, 127.7 (6CH Ar), 85.9 (CH, C-2), 81.3 (CH, C-3), 72.6, 71.7 (2 CH_2Ph), 62.7 (CH_2, C-5), 46.0 (CH, C-4), 34.7 (CH, C-1/C-6), 33.3 (CH, C-1/C-6) ppm. HR-MS (ESI): *m/z* calcd for $C_{20}H_{24}NO_3^+$: 326.17507 [*M*+H]$^+$, found 326.17481.

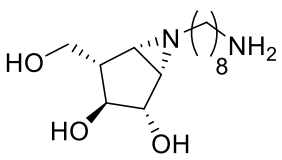
Compound 15



Aziridine **12** (673 mg, 2.13 mmol) was co-evaporated thrice at room temperature with dry distilled toluene. The compound was dissolved in dry DCM (20 ml, 0. M) and DIPEA (531 µL, 2.98 mmol) was added at room temperature. The solution was cooled to -20 °C. A solution of freshly prepared

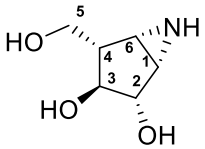
azido-octyltriflate[29] (0.49 M in DCM, 2.98 mmol) was added along the cold walls of the flask. The reaction was allowed to reach 0 °C, and stirred for 20 h. After full consumption of the starting material, the solution was diluted with DCM (150 ml), washed with a saturated $NaHCO_3$ solution (20 ml). The aqueous layer was extracted twice with DCM, and the combined organic layers were washed with brine, dried over Na_2SO_4 and volatiles were removed under reduced pressure. Purification by silica gel column chromatography (pentane/EtOAc 10:1→3:1) afforded the desired product as an oil (452 mg, 0.944 mmol, 44% yield). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.50 – 7.15 (m, 10H, 10CH Ar), 4.81 – 4.63 (m, 3H, 3CH*H*Ph), 4.55 (d, *J* = 11.6 Hz, 1H, CH*H*Ph), 4.03 (dd, *J* = 5.6, 2.8 Hz, 1H, H-2), 3.96 (dd, *J* = 10.5, 4.0 Hz, 1H, C*H*H-6), 3.84 (dd, *J* = 10.5, 5.2 Hz, 1H, CH*H*-6), 3.75 (dd, *J* = 7.4, 5.6 Hz, 1H, H-3), 3.24 (t, *J* = 6.9 Hz, 2H, CH_2N_3), 2.92 – 2.66 (m, 1H, OH), 2.32 – 2.20 (m, 1H, C*H*HN$_{aziridine}$), 2.10 (h, *J* = 4.4 Hz, 1H, H-4), 2.06 (dd, *J* = 5.1, 3.0 Hz, 1H, H-6), 2.02 (dd, *J* = 5.1, 2.9 Hz, 1H, H-1), 2.00 – 1.88 (m, 1H, CH*H*N$_{aziridine}$), 1.62 – 1.46 (m, 4H, 4CH*H* linker), 1.40 – 1.14 (m, 13H, 8CH*H* linker, grease/pentane trace) ppm. ^{13}C NMR (126 MHz, $CDCl_3$): δ = 138.7, 138.6 ($2C_q$), 128.5, 128.0, 127.9, 127.8, 127.7 (6CH Ar), 85.9 (CH, C-2), 82.3 (CH, C-3), 72.7, 71.7 (2 CH_2Ph), 63.3 (CH_2, C-5), 58.8 ($CH_2N_{aziridine}$), 51.6 (CH_2N_3), 45.7 (CH, C-4), 43.1 (CH, C-6), 41.8 (CH, C-1), 29.7, 29.5, 29.2, 28.9, 27.3, 26.8 ($6CH_2$linker) ppm. HR-MS (ESI): *m/z* calcd for $C_{28}H_{39}N_4O_3^+$: 479.30167 [*M*+H]$^+$, found 479.30139.

Compound 7 (VB-C048)



Azide **15** (136 mg, 0.284 mmol) was treated according to general procedure A, to yield titled product (59.5 mg, 0.219 mmol, 77% yield) as white powder upon lyophilisation. ^{1}H NMR (500 MHz, D_2O): δ = 4.05 (dd, *J* = 6.3, 2.2 Hz, 1H, H-2), 3.78 (dd, *J* = 10.7, 4.9 Hz, 1H, C*H*HOH), 3.69 (dd, *J* = 10.7, 9.6 Hz, 1H, CH*H*OH), 3.21 (dd, *J* = 8.2, 6.5 Hz, 1H, H-3), 3.01 – 2.95 (m, 2H, C*H*$_2$NH$_2$), 2.35 – 2.26 (m, 3H, H-1, H-6, C*H*HN$_{aziridine}$), 2.11 – 2.00 (m, 2H, H-4, CH*H*N$_{aziridine}$), 1.64 (q, *J* = 7.5 Hz, 2H, 2CH*H* linker), 1.55 – 1.46 (m, 2H, 2CH*H* linker), 1.41 – 1.26 (m, 8H, 8CH*H* linker) ppm. ^{13}C NMR (126 MHz, D_2O): δ = 78.3 (CH, C-2), 75.9 (CH, C-3), 61.2 (CH_2, C-6), 57.6 ($CH_2N_{aziridine}$), 46.8 (CH, C-4), 44.5, 41.8 (2CH, C-1, C-6) 39.6 (CH_2NH_2), 28.6, 28.4, 28.1, 26.8, 26.3, 25.5 ($6CH_2$ linker) ppm. HR-MS (ESI): *m/z* calcd for $C_{14}H_{29}N_2O_3^+$: 273.21727 [*M*+H]$^+$, found 273.21710.

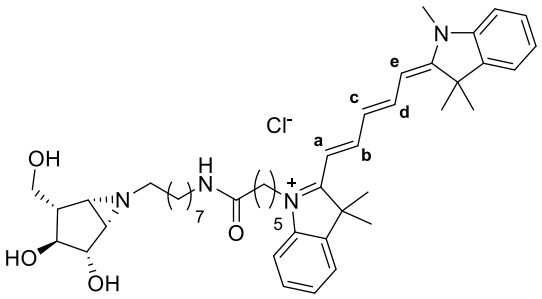
Compound 2 (VB-C075)



Aziridine **12** (31 mg, 95 µmol) was reacted according to general procedure A, yielding titled product (3.2 mg, 22 µmol, 23% yield) after HPLC purification with HILIC column (eluent A: 10 mM NH_4OAc in H_2O; eluent B: CH_3CN). ^{1}H NMR (500 MHz, D_2O): δ = 3.81 (dd, *J* = 6.4, 2.8 Hz, 1H, H-2), 3.54 (dd, *J* = 10.9, 4.5 Hz, 1H, C*H*H-6), 3.40 (t, *J* = 10.2 Hz, 1H, CH*H*-6), 2.90 (dd, *J* = 8.0, 6.4 Hz, 1H, H-3), 2.42 (dd, *J* = 5.5, 2.6 Hz, 2H, H-1, H-6), 1.79 (dddd, *J* = 9.7, 7.4, 4.5, 2.5 Hz, 1H, H-4) ppm. ^{13}C NMR (126 MHz, D_2O): δ = 78.8 (CH, C-2),

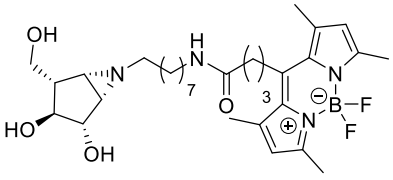
75.2 (CH, C-3), 61.2 (CH_2OH), 47.1 (CH, C-4), 36.4, 32.9 (2 CH, C-1, C-6) ppm. HR-MS (ESI): *m/z* calcd for $C_6H_{12}NO_3^+$: 146.08117 $[M+H]^+$, found 146.08121.

Compound 8 (VB-C085)



Amine **7** (27 mg, 100 µmol) was reacted according to general procedure B followed by reverse-phase HPLC purification with C18 column (A: 50 mM NH_4HCO_3; B: CH_3CN), to afford the titled product as a blue powder (5.8 mg, 7.5 µmol, 7.5% yield). After NMR analysis, the desired product was lyophilised to remove volatiles. ^{1}H NMR (500 MHz, MeOD): δ = 8.25 (t, *J* = 13.0 Hz, 2H, H-b, H-d), 7.50 (d, *J* = 7.4 Hz, 2H, 2CH Ar), 7.42 (dddd, *J* = 8.6, 7.5, 4.4, 1.2 Hz, 2H), 7.35 – 7.24 (m, 4H), 6.63 (t, *J* = 12.4 Hz, 1H), 6.28 (dd, *J* = 13.7, 5.7 Hz, 2H), 4.59 (s, 1H), 4.11 (t, *J* = 7.4 Hz, 2H, CH_2N^+), 3.88 (dd, *J* = 6.3, 3.0 Hz, 1H, H-2), 3.76 (dd, *J* = 10.2, 4.4 Hz, 1H, C*H*H-5), 3.63 (s, 4H, CH_3N, CH*H*-5), 3.17 (dd, *J* = 8.0, 6.3 Hz, 1H, H-3), 3.12 (t, *J* = 7.2 Hz, 2H, C*H*$_2$NHC=O), 2.33 – 2.23 (m, 1H), 2.20 (t, *J* = 7.3 Hz, 2H, CH_2C=O), 2.13 (dd, *J* = 5.2, 2.8 Hz, 1H, H-6), 2.08 (dd, *J* = 5.2, 3.0 Hz, 1H, H-1), 2.01 – 1.93 (m, 1H), 1.91 (s, 4H), 1.82 (q, *J* = 7.6 Hz, 2H, 2CH*H* linker), 1.73 (s, 12H, $4CH_3$), 1.71 – 1.64 (m, 2H, 2CH*H* linker), 1.59 – 1.50 (m, 2H, 2CH*H* linker), 1.49 – 1.40 (m, 4H, 4CH*H* linker), 1.37 – 1.26 (m, 8H, 8CH*H* linker) ppm. ^{13}C NMR (126 MHz, MeOD): δ = 175.7, 175.4, 174.7 ($3C_q$), 155.5 (CH), 144.3, 143.6, 142.7, 142.5, 129.8 ($5C_q$), 129.7, 126.6, 126.3, 126.3, 123.4, 123.3, 112.0, 111.9, 104.4, 104.3 ($10CHsp^2$), 80.2 (CH, C-2), 77.4 (CH, C-3), 62.7 (CH_2, C-5), 59.7 ($CH_2N_{aziridine}$), 46.2 (CH, C-1), 44.8 (CH_2N^+), 43.0 (CH, C-6), 40.4 (*C*H_2NHC=O), 36.7 (*C*H_2C=O), 31.5 (CH_3N), 30.6, 30.4, 30.3, 28.3, 28.2, ($6CH_2$ linker), 27.9 (CH_3), 27.8 (CH_3), 27.4, 26.6, 23.4 ($3CH_2$ linker) ppm. HR-MS (ESI): *m/z* calcd for $C_{46}H_{65}N_4O_4^+$: 737.50003 $[M]^+$, found 737.49978.

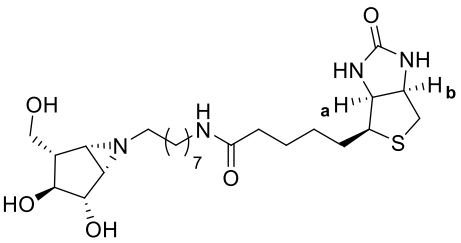
Compound 9 (VB-C066)



Amine **7** (11 mg, 40 µmol) was reacted according to general procedure B followed by silica gel column chromatography (DCM/MeOH 99:1→9:1 + 0.05% TEA), to afford the titled product as a red powder (11 mg, 18 µmol, 45% yield). After NMR analysis, the desired product was lyophilised to remove volatiles. ^{1}H NMR (500 MHz, MeOD): δ = 6.14 (s, 2H, 2CH Ar), 3.87 (dd, *J* = 6.3, 3.0 Hz, 1H, H-2), 3.75 (dd, *J* = 10.2, 4.5 Hz, 1H, C*H*H-5), 3.65 (t, *J* = 9.9 Hz, 1H, CH*H*-5), 3.21 – 3.14 (m, 3H, H-3, C*H*$_2$NHC=O), 3.07 – 2.99 (m, 2H, CH_2), 2.45 (d, *J* = 4.3 Hz, 6H), 2.38 (t, *J* = 7.1 Hz, 2H, CH_2C=O), 2.30 – 2.20 (m, 1H, C*H*$HN_{aziridine}$), 2.11 (dd, *J* = 5.2, 2.8 Hz, 1H, H-6), 2.07 (dd, *J* = 5.2, 3.0 Hz, 1H, H-1), 1.99 – 1.83 (m, 3H, CH*H*$N_{aziridine}$, 2CH*H* linker), 1.53 (q, *J* = 12.3, 9.7 Hz, 4H, 4CH*H* linker), 1.41 – 1.28 (m, 8H, 8CH*H* linker) ppm. ^{13}C NMR (126 MHz, MeOD): δ = 174.8, 155.2, 147.2, 142.4 ($4C_q$), 122.7 ($CHsp^2$), 80.2 (CH, C-2), 77.5 (CH, C-3), 62.7 (CH_2,

C-5), 59.7 ($CH_2N_{aziridine}$), 46.2 (CH, C-1), 43.0 (CH, C-6), 40.5 (CH_2NH), 37.1 (CH_2C=O), 31.1 (CH, C-4), 30.6, 30.4, 30.3, 29.3, 28.6, 28.3, 28.0 ($7CH_2$ linker), 16.6 (CH_3), 14.4 (CH_3) ppm. HR-MS (ESI): *m/z* calcd for $C_{31}H_{48}BF_2N_4O_4^+$: 589.37367 [*M*+H]$^+$, found 589.37328.

Compound 10 (VB-C087)



Amine **7** (29 mg, 108 µmol) was reacted according to general procedure B followed by reverse-phase HPLC purification (A: 50 mM NH_4HCO_3: B: CH_3CN), to afford the titled product as a white powder (4.24 mg, 8.5 µmol, 7.9% yield). After NMR analysis, the desired product was lyophilised to remove volatiles. ^{1}H NMR (850 MHz, D_2O): δ = 4.54 – 4.50 (m, 1H, H-b), 4.34 (dd, *J* = 8.0, 4.5 Hz, 1H, H-a), 3.97 (dd, *J* = 6.5, 2.4 Hz, 1H, H-2), 3.69 (dd, *J* = 10.7, 4.8 Hz, 1H, C*H*H-5), 3.61 (dd, *J* = 10.7, 9.5 Hz, 1H, CH*H*-5), 3.24 (q, *J* = 4.4 Hz, 1H, CHS), 3.14 (dd, *J* = 8.1, 6.5 Hz, 1H, H-3), 3.13 – 3.04 (m, 2H, C*H*$_2$NHC=O), 2.91 (dd, *J* = 13.4, 5.9 Hz, 1H, C*H*HS), 2.70 (d, *J* = 13.0 Hz, 1H, CH*H*S), 2.29 – 2.21 (m, 3H, H-1, H-6, C*H*HN$_{aziridine}$), 2.16 (t, *J* = 7.1 Hz, 2H, CH_2C=O), 2.02 – 1.95 (m, 2H, H-5, CH*H*N$_{aziridine}$), 1.67 – 1.60 (m, 1H, C*H*HCHS), 1.61 – 1.56 (m, 1H, CH*H*CHS), 1.56 – 1.47 (m, 2H, C*H*$_2$CH$_2$C=O), 1.44 – 1.38 (m, 4H, C*H*$_2$CH$_2$NHC=O, C*H*$_2$CH$_2$N$_{aziridine}$), 1.35 – 1.28 (m, 2H, C*H*$_2$CH$_2$CHS), 1.25 – 1.17 (m, 8H, 8CH*H* linker) ppm. ^{13}C NMR (214 MHz, D_2O): δ = 177.5, 166.2 (2C=O), 79.1 (CH, C-2), 76.6 (CH, C-3), 63.0 (CH, C-a), 62.0 (CH_2, C-5), 61.2 (CH, C-b), 58.4 ($CH_2N_{aziridine}$), 56.3 (CHS), 47.6 (CH, C-5), 45.5 (CH, C-6), 42.9 (CH, C-1), 40.6 (CH_2S), 40.2 (*C*H_2NHC=O), 36.5 (*C*H_2C=O), 29.5, 29.2, 28.7, 28.6, 27.27, 26.9, 26.1 ($8CH_2$ linker) ppm. HR-MS (ESI): *m/z* calcd for $C_{24}H_{43}N_4O_5S^+$: 499.29487 [*M*+H]$^+$, found 499.29486.

4.6 References

[1] M. J. Kieliszewski, D. T. Lamport, *Plant J.* **1994**, *5*, 157-172.

[2] M. J. Kieliszewski, A. M. Showalter, J. F. Leykam, *Plant J.* **1994**, *5*, 849-861.

[3] D. Ndeh, A. Rogowski, A. Cartmell, A. S. Luis, A. Baslé, J. Gray, I. Venditto, J. Briggs, X. Zhang, A. Labourel, N. Terrapon, F. Buffetto, S. Nepogodiev, Y. Xiao, R. A. Field, Y. Zhu, M. A. O'Neil, B. R. Urbanowicz, W. S. York, G. J. Davies, D. Wade Abbott, M.-C. Ralet, E. C. Martens, B. Henrissat, H. J. Gilbert, *Nat.* **2017**, *544*, 65-70.

[4] A. J. Whitcombe, M. A. O'Neill, W. Steffan, P. Albersheim, A. G. Darvill, *Carbohydr. Res.* **1995**, *271*, 15–29.

[5] A. A. Arzamasov, D. van Sinderen, D. A. Rodionov, *Front. Microbiol.* **2018**, *9*, 776.

[6] S. Kaneko, M. Sano, I. Kusakabe, *Appl Environ Microbiol.* **1994**, *60*, 3425–3428.

[7] M. Komeno, H. Hayamizu, K. Fujita, H. Ashida, *Appl Environ Microbiol.* **2019**, *85*, e02582-18.

[8] M. A. Schell, M. Karmirantzou, B. Snel, D. Vilanova, B. Berger, G. Pessi, M.-C. Zwahlen, F. Desiere, P. Bork, M. Delley, R. D. Pridmore, F. Arigoni, *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 14422-14427.

[9] A. Margolles, C. G. de los Reyes-Gavilán, *Appl. Environ. Microbiol.* **2003**, *69*, 5096–5103.

[10] M. Komeno, Y. Yoshihara, J. Kawasaki, W. Nabeshima, K. Maeda, Y. Sasaki, K. Fujita, H. Ashida, *Appl. Microbiol. Biotechnol.* **2022**, *106*, 1957-1965.

[11] S. Pérez, M. A. Rodríguez-Carvajal, T. Doco, *Biochimie* **2003**, *85*, 109–121.

[12] M. Pabst, R. M. Fischl, L. Brecker, W. Morelle, A. Fauland, H. Köfeler, F. Altmann, R. Léonard, *Plant. J.* **2013**, *76*, 61-72.

[13] Z. Zhu, M. He, C.-H. Huang, T.-P. Ko, Y.-F. Zeng, Y.-N. Huang, Shiru Jia, F. Lu, J.-R. Liu, R.-T. Guo, *Acta Crystallogr. F. Struct. Biol. Commun.* **2014**, *70*, 636–638.

[14] T. Ito, K. Saikawa, S. Kim, K. Fujita, A. Ishiwata, S. Kaeothip, T. Arakawa, T. Wakagi, G. T. Beckham, Y. Ito, S. Fushinobu, *Biochem. Biophys. Res. Commun.* **2014**, *447*, 32–37.

[15] A. S. Luis, J. Briggs, X. Zhang, B. Farnell, D. Ndeh, A. Labourel, A. Baslé, A. Cartmell, N. Terrapon, K. Stott, E. C. Lowe, R. McLean, K. Shearer, J. Schückel, I. Venditto, M.-C. Ralet, B. Henrissat, E. C. Martens, S. C. Mosimann, D. W. Abbott, H. J. Gilbert, *Nat. Microbiol.* **2018**, *3*, 210–219.

[16] K. Fujita, Y. Takashi, E. Obuchi, K. Kitahara, T. Suganuma, *J. Biol Chem.* **2014**, *289*, 5240 - 5249.

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

[18] P. B. Jariwala, S. J. Pellock, D. Goldfarb, E. W. Cloer, M. Artola, J. B. Simpson, A. P. Bhatt, W. G. Walton, L. R. Roberts, M. B. Major, G. J. Davies, H. S. Overkleeft, M. R. Redinbo, *ACS Chem. Biol.* **2020**, *15*, 217–225.

[19] X. Biarnés, A. Ardèvol, A. Planas, C. Rovira, A. Laio, M. Parrinello, *J. Am. Chem. Soc.* **2007**, *129*, 10686–10693.

[20] A. Ardèvol, X. Biarnés, A. Planas, C. Rovira, *J. Am. Chem. Soc.* **2010**, *132*, 16058–16065.

[21] J. Iglesias-Fernández, L. Raich, A. Ardèvol, C. Rovira, *Chem Sci.* **2015**, *6*, 1167–1177.

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

[23] D. Cremer, J. A. Pople, *J. Am. Chem. Soc.* **1975**, *97*, 1354–1358.

[24] C. H. Larsen, B. H. Ridgway, J. T. Shaw, K. A. Woerpel, *J. Am. Chem. Soc.* **1999**, *121*, 12208–12209.

[25] D. M. Smith, M. B. Tran, K. A. Woerpel, *J. Am. Chem. Soc.* **2003**, *125*, 14149-14152.

[26] 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, J. M. F. G. Aerts, *ACS Cent. Sci.* **2016**, *2*, 351–358.

[27] G. J. Davies, A. Planas, C. Rovira, *Acc. Chem. Res.* **2012**, *45*, 308–316.

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

[29] S. P. Schröder, J. W. van de Sande, W. W. Kallemeijn, C.-L. Kuo, M. Artola, E. J. van Rooden, J. Jiang, T. J. M. Beenakker, B. I. Florea, W. A. Offen, G. J. Davies, A. J. Minnaard, J. M. F. G. Aerts, J. D. C. Codée, G. A. van der Marel, H. S. Overkleeft, *Chem. Commun.* **2017**, *53*, 12528–12531.

[30] J. B. Jiang, W. W. Kallemeijn, D. W. Wright, A. N. 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. Schurmann, D. Mink, R. G. Boot, J. D. C. Codée, G. A. van der Marel, G. J. Davies, J. M. F. G. Aerts, H. S. Overkleeft, *Chem. Sci.* **2015**, *6*, 2782–2789.

[31] S. Kaeothip, A. Ishiwata, T. Ito, S. Fushinobu, K. Fujita, Y. Ito, *Carbohydr. Res.* **2013**, *382*, 95-100.

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

[33] 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, H. S. Overkleeft, *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 12529-12533.

[34] G. Winter, C. M. C. Lobley, S. M. Prince, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69*, 1260–1273.

[35] G. Winter, D. G. Waterman, J. M. Parkhurst, A. S. Brewster, R. J. Gildea, M. Gerstel, L. Fuentes-Montero, M. Vollmar, T. Michels-Clark, I. D. Young, N. K. Sauter, G. Evans, *Acta Crystallogr. Sect. D Struct. Biol.* **2018**, *74*, 85–97.

[36] A. D'Arcy, T. Bergfors, S. W. Cowan-Jacob, M. Marsh, *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **2014**, *70*, 1117-

1126.

[37] P. D. Shaw Stewart, S. A. Kolek, R. A. Briggs, N. E. Chayen, P. F. M. Baldock, *Crystal Growth & Design* **2011**, *11*, 3432-3441.

[38] A. K. Shah, Z.-J. Liu, P. D. Stewart, F. D. Schubot, J. P. Rose, M. G. Newton, B.-C. Wang, *Acta Crystallogr. Sect. D* **2005**, *61*, 123-129.

[39] W. Kabsch, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 125–132.

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

[41] A. A. Lebedev, P. Young, M. N. Isupov, O. V. Moroz, A. A. Vagin, G. N. Murshudov, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2012**, *68*, 431–440.

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

[43] G. N. Murshudov, P. Skubák, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67*, 355–367.

[44] C. Ballard, R. Keegan, E. Krissinel, A. Lebedev, V. Uski, D. Waterman, M. Wojdyr, *Acta Crystallogr. Sect. A Found. Adv.* **2014**, *70*, C1723–C1723.

[45] A. Vagin, A. Teplyakov, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 22–25.

[46] R. J. Gildea, J. Beilsten-Edmands, D. Axford, S. Horrell, P. Aller, J. Sandy, J. Sanchez-Weatherby, C. D. Owen, P. Lukacik, C. Strain-Damerell, R. L. Owen, M. A. Walsh, G. Winter, *Acta Crystallogr. Sect. D Struct. Biol.* **2022**, *78*, 752-769.

[47] 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, K. S. Wilson, *Acta Crystallogr. Sect. D* **2011**, *67*, 235-242.

[48] E. Krissinel, A. A. Lebedev, V. Uski, C. B. Ballard, R. M. Keegan, O. Kovalevskiy, R. A. Nicholls, N. S. Pannu, P. Skubák, J. Berrisford, M. Fando, B. Lohkamp, M. Wojdyr, A. J. Simpkin, J. M. H. Thomas, C. Oliver, C. Vonrhein, G. Chojnowski, A. Basle, A. Purkiss, M. N. Isupov, S. McNicholas, E. Lowe, J. Triviño, K. Cowtan, J. Agirre, D. J. Rigden, I. Uson, V. Lamzin, I. Tews, G. Bricogne, A. G. W. Leslie, D. G. Brown, *Acta Crystallogr. Sect. D Struct. Biol.* **2022**, *78*, 1079-1089.

[49] P. Emsley, K. Cowtan, *Acta Crystallogr Sect. D Biol. Crystallogr.* **2004**, *60*, 2126-2132.

[50] F. Long, R. A. Nicholls, P. Emsley, S. Graæulis, A. Merkys, A. Vaitkus, G. N. Murshudov, *Acta Crystallogr. Sect. D Struct. Biol.* **2017**, *73*, 112-122.

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

APPENDIX

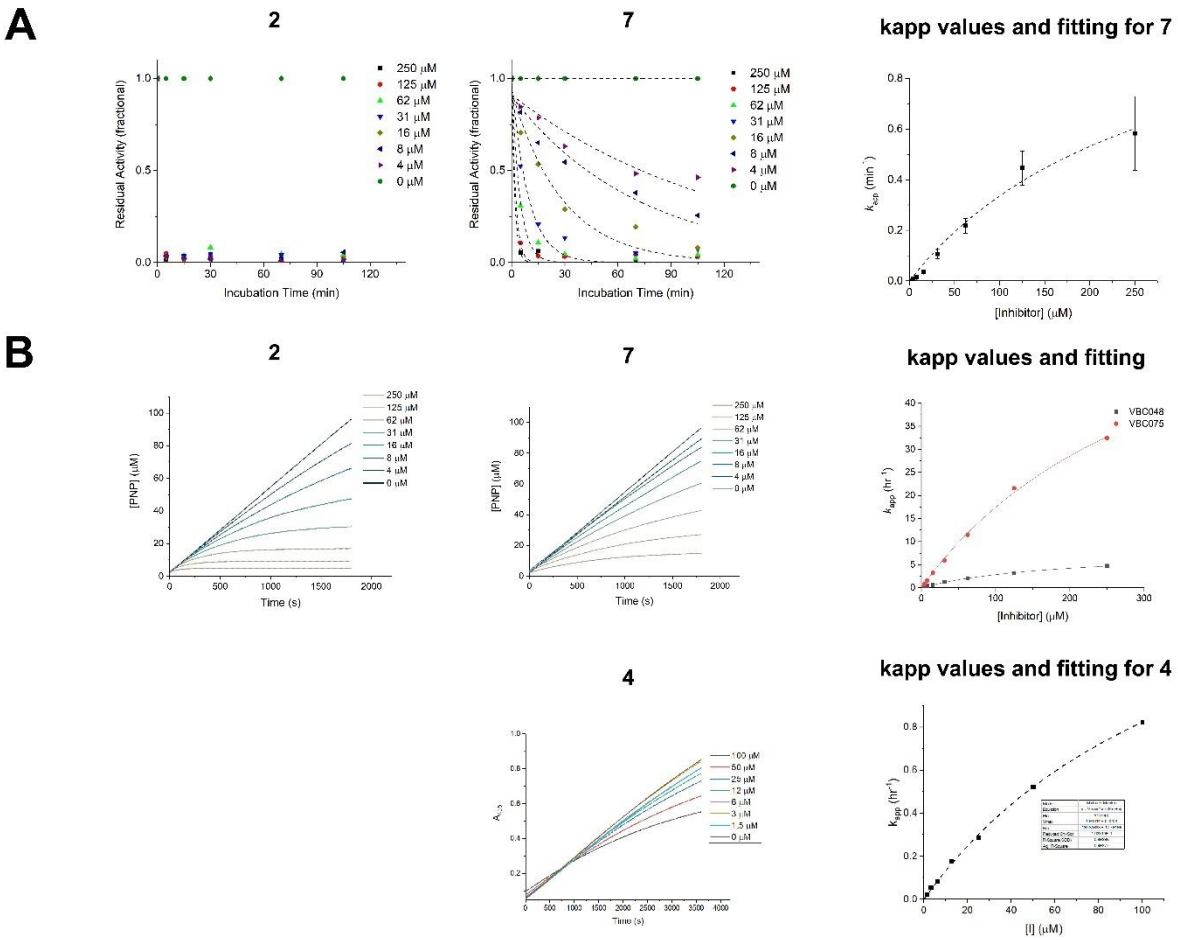


Figure 4.S1. Inhibition kinetics for compounds **2**, **4** and **7** against HypBA1 and *Bt*GH146. (A) Inhibition of HypBA1. Plots of residual enzyme activity *vs* incubation time at different concentration of inhibitors **2** and **7**. Parameters for irreversible inhibition by **7** is modelled by plotting measured inactivation rate constants (k_{app}) *vs* [Inhibitor]. (B) Inhibition of *Bt*GH146. Plots of measured residual absorbance at 405 nm (A405) corresponding to *p*-nitrophenol (PNP) concentration [PNP] *vs* incubation time at different concentration of inhibitors **2**, **4** and **7**. Parameters for irreversible inhibition by **2**, **4** and **7** are modelled by plotting measured inactivation rate constants (k_{app}) *vs* [Inhibitor].

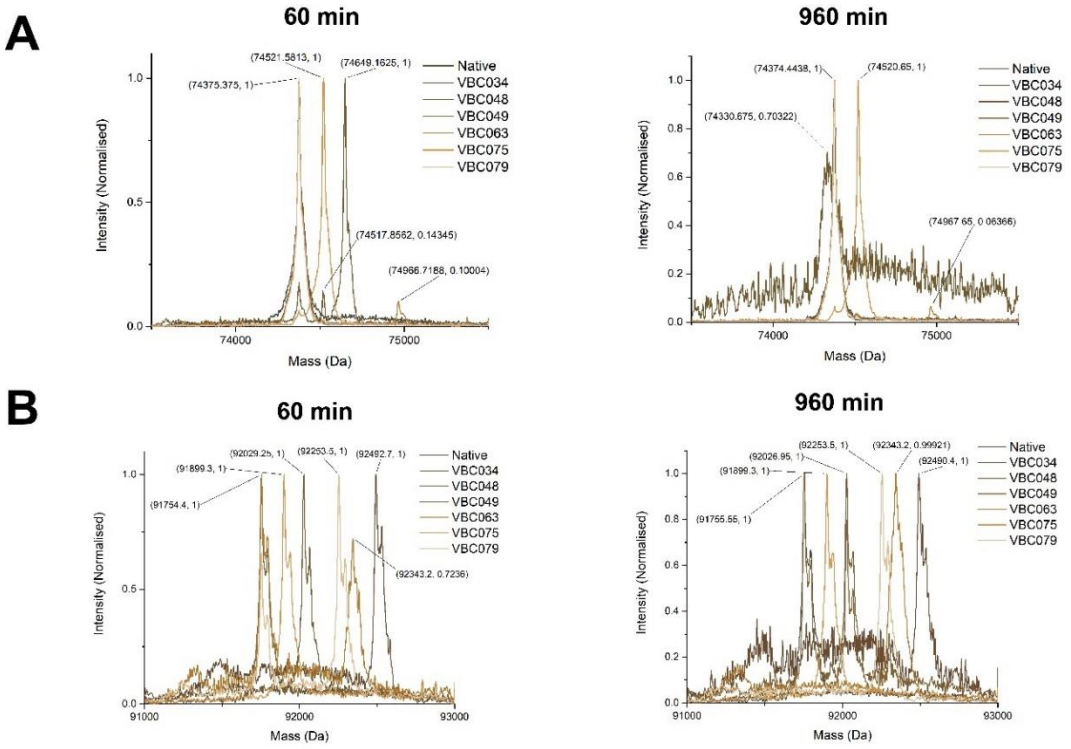


Figure 4.S2. Intact mass spectrometry (MS) of tested compounds **2** – **7** with r*Bt*GH146 and rHypBA1. (A) Intact MS of rHypBA1 following incubation with with 0.1 mM compound in pH 4.5 Na-acetate buffer at 22 °C for different lengths of time (60 or 960 minutes). Expected mass (native): 74376 Da, with **2**: 74521 Da, with **3**: 74649 Da, with **4**: 75113 Da, with **5**: 74988 Da, with **6**: 74875 Da, with **7**: 74648 Da. (B) Intact MS of r*Bt*GH146 following incubation with with 0.1 mM compound in pH 7.5 HEPES buffer at 37 °C for different lengths of time (60 or 960 minutes). Expected mass (native): 91512 Da, with **2**: 91657 Da, with **3**: 91785 Da, with **4**: 92250 Da, with **5**: 92123 Da, with **6**: 92011 Da, with **7**: 91784 Da.