Cyclophellitol analogues for profiling of exo- and endo-glycosidases

Schröder, S.P.

Chapter 5

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

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

5.1 Introduction

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

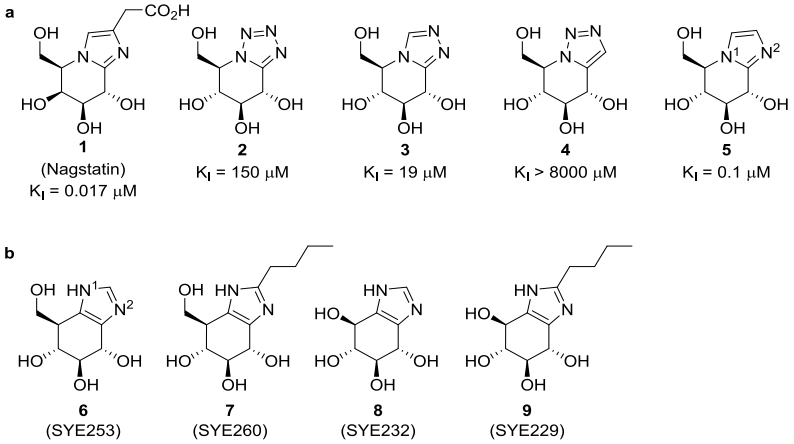
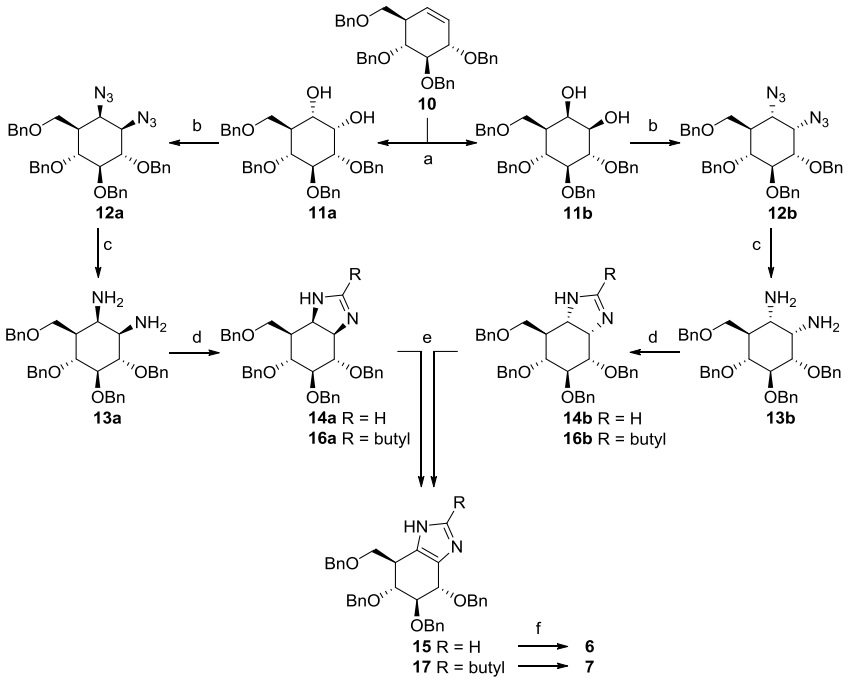


Figure 1. (a) Known cyclitol-azole type glycosidase inhibitors and their inhibitory potencies against porcine NAG (**1**)[3] and sweet almond β-glucosidase (**2**-**5**).[4–7] (b) Gluco-1*H*-imidazoles and respective analogues subject of this study. Numbering of the nitrogen atoms used for describing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

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

5.2 Results and Discussion

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

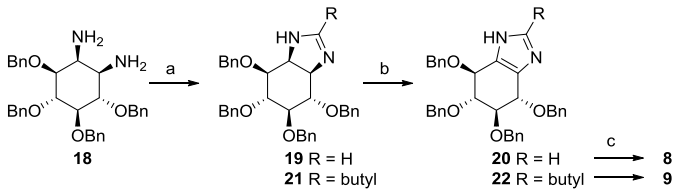


Scheme 1 Synthesis of 1*H*-imidazoles **6** and **7**, starting from cyclohexene **10**.[14] Reagents and conditions: a) $RuCl_3 \cdot 3H_2O$ (7 mol%), $NaIO_4$, EtOAc, MeCN, H_2O, 0 °C, 90 min, yield **11a**: 40%; yield **11b**: 32%; b) MsCl, *N*-methyl-imidazole, Et_3N, $CHCl_3$, rt, 16 h, then NaN_3,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

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

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

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

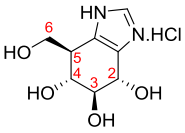


Scheme 2 Synthesis of conduritol B 1*H*-imidazoles **8** and **9**. Reagents and conditions: a) trimethyl orthoformate, HFIP, rt, 16h, yield **19**: 95%; or trimethyl orthovalerate, HFIP, rt, 16 h, yield **21**: 89%; b) For **20**: IBX, DMSO, 45 °C, 16 h, 62%; for **22**: $(COCl)_2$, DMSO, DCM, -60 °C,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

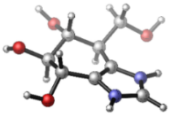
With compounds **6-9** in hand, their lowest energy conformations were modeled by DFT (B3LYP/6-311G(d,p), PCM) calculations on lead compound **6**. The calculated three lowest energy geometries for **6** adopt a $^4H_3$ conformation, differing only in the rotation angle around the C5-C6 axis: *gt*, *gg*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

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

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

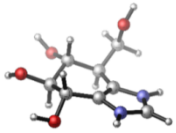


6



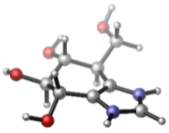
ΔG^T_{aq}= 0.0 kcal/mol

$^4H_3$ *(gt)*



ΔG^T_{aq}= +0.0 kcal/mol

$^4H_3$ *(gg)*



ΔG^T_{aq} = +3.2 kcal/mol

$^4H_3$ *(tg)*

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

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

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

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

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

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

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

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

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

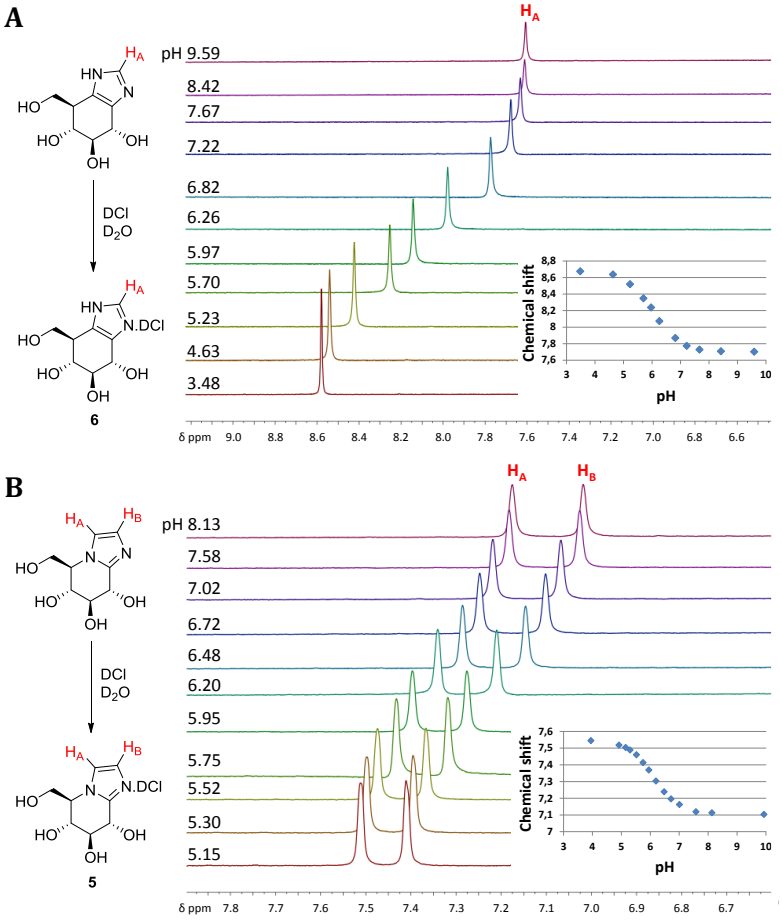


Figure 2 Compound pK_{AH} determination using ^{1}H-NMR. Both imidazoles display similar pK_{AH} values. Experimental details are given in the experimental section. Spectra shown are cropped, showing the azole-ring protons which displayed the highest Δppm shift upon acidification.

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

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

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

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

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

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

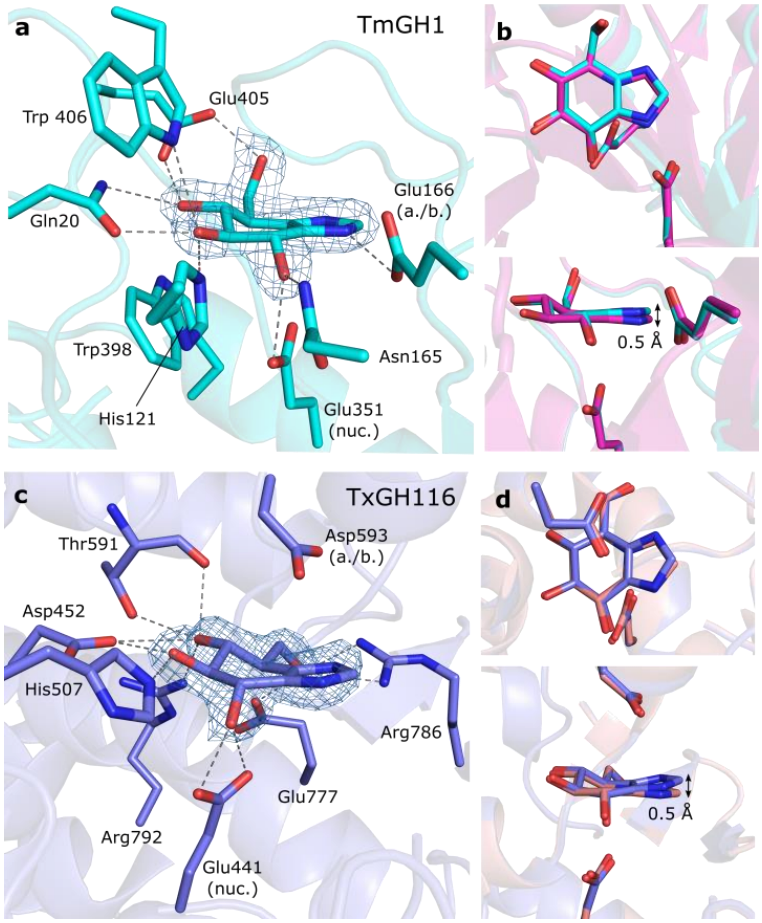


Figure 3 (a) Gluco-*1H*-imidazole **6** in complex with *Tm*GH1, with direct H-bonding interactions shown. (b) Overlay of **6** (cyan) and **5** (pink) within the *Tm*GH1 active site. (c) Gluco-*1H*-imidazole **6** in complex with *Tx*GH116, and direct H-bonding interactions. (d) Overlay of **6** (blue)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

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

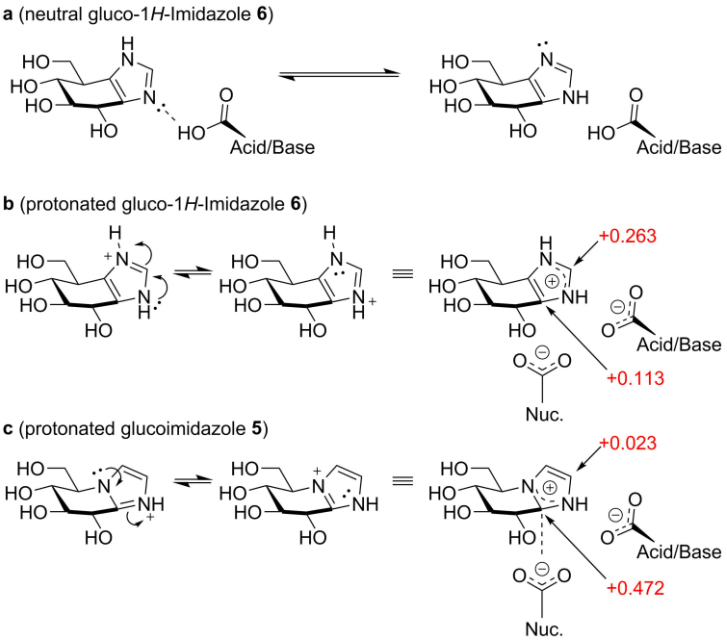


Figure 4 Interactions of gluco-1*H*-imidazole **6** and classical glucoimidazole **5** with the catalytic residues. (a) In the neutral state, gluco-*1H*-imidazoles may undergo prototropic tautomerism, reducing the availability of the N2 lone pair for interaction with a protonated glucosidase acid/base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

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

5.3 Conclusion

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

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

Experimental procedures

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

General procedure 1 (GP1): Bis-azidation

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

General procedure 2 (GP2): Azide reduction

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

General procedure 3 (GP3): Imidazoline formation

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

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

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

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

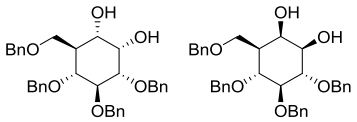
General procedure 5 (GP5): Oxidation to the imidazole (Swern conditions)

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

General procedure 6 (GP6): Hydrogenation

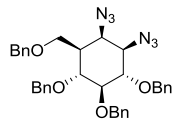
The imidazole starting material was dissolved in MeOH (0.03 M) under N_2 atmosphere, then HCl (1.25M in MeOH, 10 equiv.) and $Pd(OH)_2/C$ (20 wt%) were added and the mixture was purged with H_2 with a balloon. The mixture was stirred vigorously for 16 h, filtered over a small Celite pad and finally concentrated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

Compound 11a and compound 11b



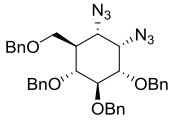
Cyclohexene **10**[35] (1.28 g, 2.46 mmol) was dissolved in EtOAc (15 mL) and MeCN (15 mL) and cooled to 0 °C. A solution of $RuCl_3.3H_2O$ (36 mg, 0.17 mmol) and $NaIO_4$ (789 mg, 3.69 mmol) in H_2O (4.9 mL) was added and the mixture was stirred vigorously at 0 °C for 90 min. The mixture was quenched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

Compound 12a



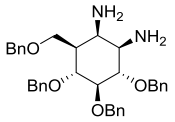
Starting from **11a** (55 mg, 0.1 mmol) and following **GP1**, the product was purified by flash column chromatography (pentane/EtOAc, 15:1) affording compound **12a** as a white solid (45 mg, 74%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.44 – 7.11 (m, 20H), 4.96 – 4.74 (m, 5H), 4.55 – 4.37 (m, 3H), 4.16 (t,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

Compound 12b



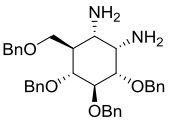
Starting from **11b** (55 mg, 0.1 mmol) and following **GP1**, the product was purified by flash column chromatography (pentane/EtOAc, 15:1) affording compound **12b** as a white solid (35 mg, 58%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.13 (m, 20H), 4.95 – 4.67 (m, 5H), 4.48 (d, *J* = 4.3 Hz, 1H),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

Compound 13a



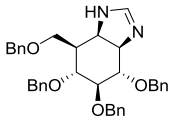
Starting from **12a** (367 mg, 0.61 mmol) and following **GP2**, the product was purified by flash column chromatography (DCM/MeOH, 99:1 → 49:1) affording compound **13a** as a colorless oil (269 mg, 80%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.42 – 7.11 (m, 20H), 4.99 (d, *J* = 11.1 Hz, 1H), 4.96 – 4.83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

Compound 13b



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

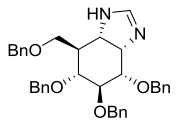
Compound 14a



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

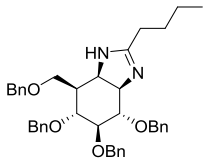
4.64 (m, 5H), 4.51 – 4.42 (m, 3H), 4.01 (dd, *J* = 9.5, 4.4 Hz, 1H), 3.86 (dd, *J* = 9.4, 6.2 Hz, 1H), 3.77 (dd, *J* = 9.2, 4.1 Hz, 1H), 3.68 (t, *J* = 8.8 Hz, 1H), 3.58 – 3.51 (m, 1H), 3.50 – 3.40 (m, 2H), 2.25 (ddt, *J* = 12.3, 8.4, 4.2 Hz, 1H). ^{13}C-NMR (101 MHz, CD_3CN) δ 156.2, 140.0, 139.9, 139.8,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

Compound 14b



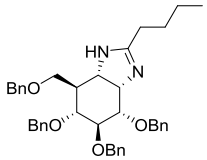
Starting from **13b** (110 mg, 0.2 mmol) and following **GP3** using trimethyl orthoformate, the product was purified by flash column chromatography (DCM/MeOH, 99:1 → 4:1) affording compound **14b** as a colorless oil (98 mg, 87%). ^{1}H-NMR (400 MHz, CD_3CN) δ 7.38 – 7.20 (m, 20H), 7.03 (s, 1H), 4.78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

Compound 16a



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

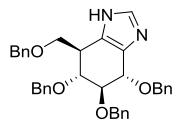
Compound 16b



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

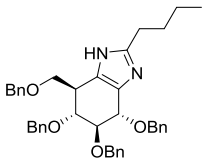
1.73 (t, *J* = 10.4 Hz, 1H), 1.52 (m, 2H), 1.31 (m, 2H), 0.87 (t, *J* = 7.3 Hz, 3H). ^{13}C-NMR (101 MHz, CD_3CN) δ 168.5, 140.1, 140.0, 139.8, 129.3, 129.2, 129.2, 129.1, 128.8, 128.7, 128.7, 128.6, 128.4, 128.4, 128.4, 128.3, 83.6, 80.0, 79.0, 74.6, 74.5, 73.8, 73.3, 69.1, 62.5, 61.0, 46.4, 29.7, 29.5,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

Compound 15



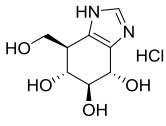
Starting from **14a** (43 mg, 76 µmol) and following **GP4**, the product was purified by flash column chromatography (DCM/MeOH, 99:1 → 67:1) affording compound **15** as a colorless oil (32 mg, 75%). Using the same conditions, product **15** could be obtained from imidazoline **14b** (82 mg, 0.15 mmol)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

Compound 17



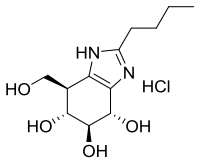
Starting from **16a** (74 mg, 0.12 mmol) and following **GP5**, the product was purified by flash column chromatography (DCM/MeOH, 199:1 → 99:1) affording compound **17** as a colorless oil (56 mg, 76%). Using the same conditions, product **17** could be obtained from imidazoline **16b** (40 mg, 64.7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

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



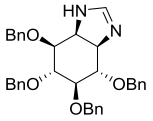
Starting from **15** (26 mg, 46.4 µmol) following **GP6**, the pure product was afforded as a colorless oil (12 mg, quant.). ^{1}H-NMR (400 MHz, D_2O) δ 8.55 (s, 1H), 4.69 (d, *J* = 7.6 Hz, 1H), 4.14 (dd, *J* = 11.4, 2.8 Hz, 1H), 3.93 (dd, *J* = 11.4, 4.8 Hz, 1H), 3.79 (t, *J* = 9.4 Hz, 1H), 3.71 (dd,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

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



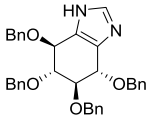
Starting from **17** (48 mg, 77.8 µmol) following **GP6**, the pure product was afforded as a colorless oil (24 mg, quant.). ^{1}H-NMR (400 MHz, MeOD) δ 4.53 (d, *J* = 7.4 Hz, 1H), 4.12 (dd, *J* = 10.8, 3.1 Hz, 1H), 3.89 (dd, *J* = 10.8, 5.0 Hz, 1H), 3.72 (t, *J* = 9.0 Hz, 1H), 3.60 (dd, *J* = 9.3, 7.4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

Compound 19



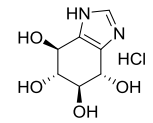
Starting from tetrabenzyl *myo*-diaminocyclitol[15] **18** (400 mg, 0.74 mmol) and following **GP3** using trimethyl orthoformate, the product was purified by flash column chromatography (DCM/MeOH, 99:1 → 7:3) affording the title compound as a colorless oil (387 mg, 95%). ^{1}H-NMR (400 MHz, CD_3CN)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

Compound 20



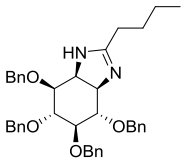
Starting from **19** (55 mg, 0.1 mmol) and using **GP4**, the product was purified by flash column chromatography (DCM/MeOH, 199:1 → 67:1) affording the title compound as a colorless oil (34 mg, 62%). ^{1}H-NMR (400 MHz, CD_3CN) δ 7.57 (s, 1H), 7.44 – 7.14 (m, 20H), 4.92 – 4.83 (m, 4H), 4.80 (d, *J*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

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



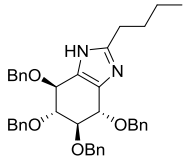
Starting from **20** (18 mg, 32.9 µmol) and following **GP6**, the pure product was afforded as a colorless oil (8.0 mg, quant.). ^{1}H-NMR (400 MHz, D_2O) δ 8.70 (s, 1H), 4.76 (m, 2H), 3.74 – 3.58 (d, *J* = 2.7 Hz, 2H). ^{13}C-NMR (101 MHz, D_2O) δ 135.8. 127.5, 75.9, 66.5. HRMS (ESI-TOF) m/z: $[M+Na]^+$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

Compound 21



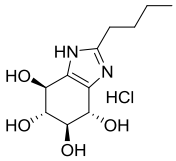
Starting from **18** (53 mg, 0.1 mmol) and following **GP3** using trimethyl orthovalerate, the product was purified by flash column chromatography (DCM/MeOH, 99:1 → 8:2) affording the title compound as a colorless oil (53 mg, 89%). ^{1}H-NMR (400 MHz, CD_3CN) δ 7.42 – 7.18 (m, 20H), 4.77 – 4.56 (m,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

Compound 22



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

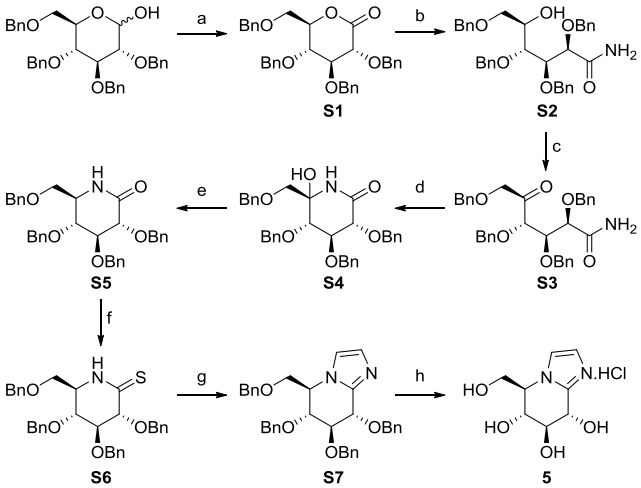
Compound 9 (conduritol B-2-butyl-1*H*-imidazole, SYE229)



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

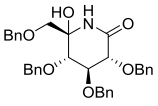
Synthesis of glucoimidazole 5

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



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

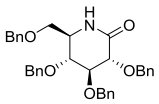
Compound S4



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

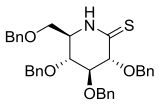
3.32 (q, *J* = 9.6 Hz, 2H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 171.9, 138.3, 137.9, 137.4, 137.1, 128.5, 128.4, 128.4, 128.4, 128.4, 128.3, 128.2, 128.1, 128.1, 128.0, 127.9, 127.8, 127.8, 127.7, 82.0, 79.5, 79.3, 77.3, 75.4, 75.2, 74.8, 73.5, 72.3.

Compound S5



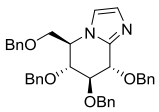
Compound **S4** (6.0 g, 10.8 mmol) was dissolved in MeCN (150 mL), formic acid (40 mL) and $NaCNBH_3$ (1.35 g, 21.6 mmol) was added and the mixture was refluxed for 2 h. Then, the reaction was cooled on ice, quenched with aq. HCl (0.1 M, 80 mL) and stirred for 15 minutes. The reaction mixture was filtered,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

Compound S6



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

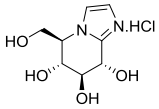
Compound S7



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

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

Compound 5 (glucoimidazole)



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

Determining compound pK_{AH} values

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

Biochemical and Biological Methods

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

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

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

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

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

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

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

Kinetic studies

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

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

Protein expression and crystallography

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

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

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

ITC

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

DFT-Geometry optimisation

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

DFT-Restricted conformational energy surface calculations

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

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Chem., 1986, **261**, 8263–8269.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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