Universiteit
Leiden
The Netherlands

Activity-based protein profiling of glucosidases, fucosidases and glucuronidases

Jiang, J.

Citation
Jiang, J. (2016, June 23). *Activity-based protein profiling of glucosidases, fucosidases and glucuronidases*. Retrieved from https://hdl.handle.net/1887/41279

Version:	Not Applicable (or Unknown)
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/41279

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

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/41279 holds various files of this Leiden University dissertation

Author: Jiang Jianbing
Title: Activity-based protein profiling of glucosidases, fucosidases and glucuronidases
Issue Date: 2016-06-23

5

Detection of active mammalian GH31 α-glucosidases in health and disease using in-class, broad-spectrum activity-based probes

Jianbing Jiang, Chi-Lin Kuo, Liang Wu, Christian Franke, Wouter W. Kallemeijn, Bogdan I. Florea, Eline van Meel, Gijsbert A. van der Marel, Jeroen D. C. Codée, Rolf G. Boot, Gideon J. Davies, Herman S. Overkleeft and Johannes M. F. G. Aerts, *ACS Central Science*, **2016**, in press, DOI: org/10.1021/acscentsci.6b00057.

5.1 Introduction

Lysosomal α-glucosidase (GAA, α-1,4-glucosidase, acid maltase) (E.C.3.2.1.20) is a retaining α-glucosidase, which has been classified into CAZy family GH31.[1-4] Following processing within endoplasmic reticulum the 110 kDa (952 AA) precursor is transported to the lysosomes where it is modified to active 76 and 70 kDa isoforms.[5,6] Within the lysosomes, GAA catalyzes the degradation of glycogen via a general acid/base catalyzed double displacement mechanism,[7,8] releasing a molecule of α-glucose with net retention of stereochemistry at its anomeric center (Figure 1A). Deficiency in GAA leads to the glycogen storage disease type II, known as Pompe disease.[9] In Pompe patients, intra-lysosomal glycogen accumulation causes progressive muscle weakness in heart and skeletal muscles and also affects the liver and nervous system.[10] Different clinical forms of Pompe disease are usually discerned based on age of onset. The infantile-onset form manifests at 4 to 8 months and, when untreated, results in death in the first years of life.[11] Later onset forms generally progress more slowly and are characterized by progressive decrease in muscle strength in the legs followed by smaller muscles in the trunk

and arms and ultimately to fatality through respiratory failure.[12] The severity of the Pompe disease and its progress correlates with the extent of enzyme loss. Pompe disease is currently treated by chronic intravenous administration of recombinant human GAA (rGAA, alglucosidase alfa; Myozyme). Impressive delay of fatal symptoms by enzyme replacement therapy (ERT) is observed in infantile Pompe disease patients,[13] but treatment of late onset disease requires very large amounts of therapeutic enzyme due to poor correction of muscle cell pathology.[14]

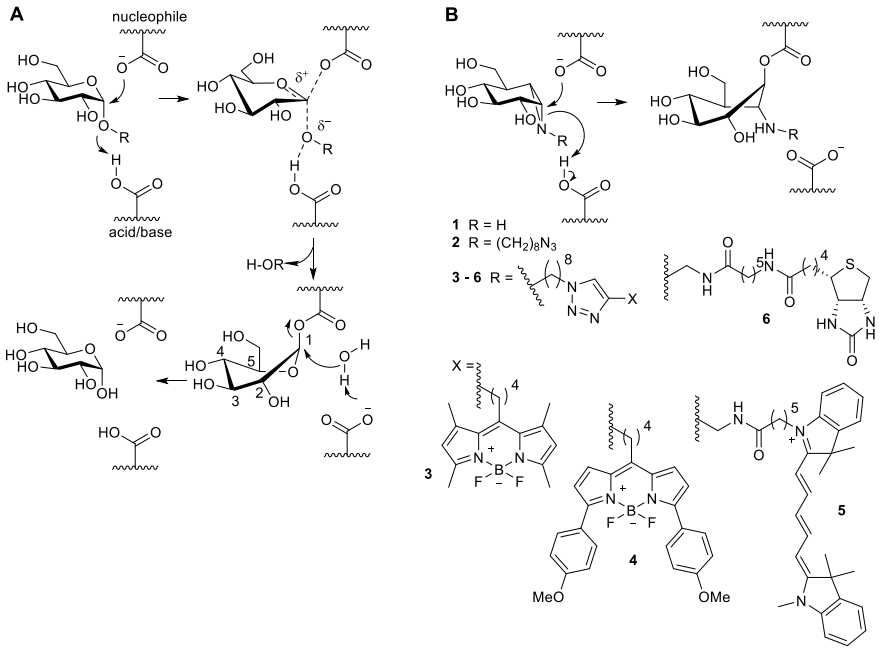


Figure 1. The mechanism of action of retaining α-glucosidases allows the development of activity-based probes. (A) Koshland double-displacement mechanism employed by retaining α-glucosidases. (B) α-Glucose-configured *N*-alkyl cyclophellitol aziridines as mechanism-based irreversible retaining α-glucosidase inhibitors (**1**, **2**) and probes (**3**-**6**).

The identification of GAA activity in living cells or tissues is challenging. It has been previously shown that activity-based protein profiling (ABPP) is an effective methodology to quantify β-glucosidases responsible for the lysosomal storage disorder, Gaucher disease.[15,16] Linking "reporter moieties" to the aziridine nitrogen in the retaining β-glucosidase inhibitor, cyclophellitol aziridine, yielded in-class broad-spectrum retaining β-glucosidase probes. These probes label, besides lysosomal glucosylceramidase (GBA-the enzyme deficient in Gaucher disease) also neutral glucosylceramidase (GBA2), broad-specificity β-glucosidase (GBA3) and lactase phloridzin hydrolase (LPH).[16-21] Based on these findings, it was hypothesized that the corresponding α-glucopyranose configured, nitrogen-substituted cyclophellitol aziridines (Figure 1B), would be effective ABPs for specific labeling of GH31 retaining α-glucosidases,

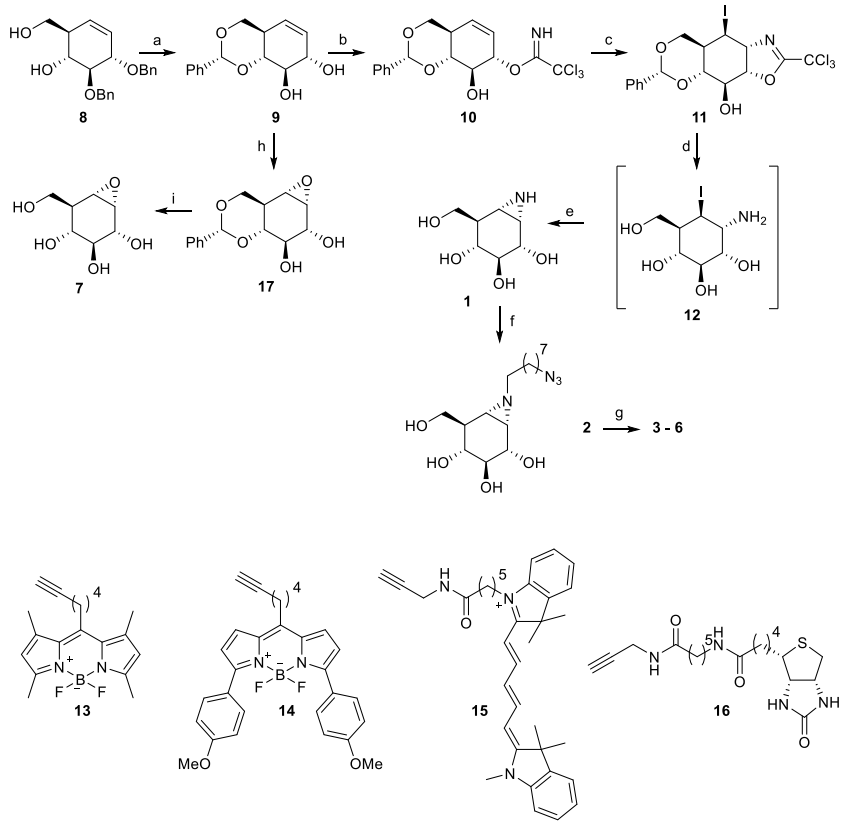
besides GAA also ER α-glucosidase II (GANAB, a crucial enzyme in the quality control of newly formed glycoproteins in the ER) and intestinal α-glucosidases involved in food processing. Here, the successful development of such in-class GH31 α-glucosidase ABPs is reported. Both mass spectrometry and X-ray crystallography demonstrate the covalent labeling of the probe on a representative bacterial GH31 enzyme, whch is homologous to GAA. GH31 α-glucosidases are labeled *in situ* by designed ABPs and that the absence of GAA in Pompe disease cells is readily demonstrated in a diagnostic manner.

5.2 Results

Synthesis of mechanism-based inhibitors and activity-based probes

As the first research objective, to develop an efficient route of synthesis for the preparation of α-configured cyclophellitol aziridine **1** was set out (Scheme 1). The synthesis route for cyclohexene **8**, the precursor in the total synthesis of cyclophellitol (a natural product and a retaining β-glucosidase inhibitor) reported by Madsen and co-workers,[22] has been developed in previous works on retaining β-glucosidase probes development.[15,16] It was reasoned that an iodocyclization scheme, with an appropriate nitrogen nucleophile delivered to the alkene-derived iodonium ion from the allylic alcohol position, would yield an appropriately configured 2-amino-1-iodo species for subsequent intramolecular cyclisation in a stereospecific fashion. For this purpose, the benzyl protective groups in **8** were removed by Birch reduction, after which the 4,6-benzylidene (glucopyranose numbering) was installed to give **9**. The allylic alcohol in **9** proved the most reactive of the two secondary alcohols towards trichloroacetonitrile in the presence of 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) as the catalytic base, providing imidate **10** as the major product in the presence of small amounts of the bis-imidate. Subsequent key iodocyclization afforded in a stereospecific fashion cyclic imidate **11**. Both cyclic imidate and benzylidene acetal were hydrolysed under acidic conditions and the resulting crude primary amine **12** was treated with sodium bicarbonate giving aziridine **1**. The aziridine nitrogen in **1** was alkylated with 1-iodo-7-azidoheptane to yield **2**, onto which and by means of copper (I)-catalyzed azide-alkyne [2+3] cycloaddition either a BODIPY-green dye, a BODIPY-red dye, a Cy5 dye or a biotin was conjugated, to give ABPs **3**, **4**, **5** and **6**, respectively (Scheme 1). 1,6-*epi*-cyclophellitol **7** was prepared by stereoselective epoxidation of **9** with 3-chloroperbenzoic acid (*m*CPBA) and deprotection of benzylidene acetal group of the resulting **17** by Pearlman's catalyst hydrogenation. As reference competitive inhibitors, deoxynojirimycin derivatives **18**-**22** and maltose **23** were used for screening (SI, Figure S1).[23]

Scheme 1. Synthesis of the cyclophellitol aziridine inhibitors **1**, **2**, probes **3**-**6** and 1,6-*ep*i-cyclophellitol **7**.



Reagents and conditions: (a) i) Li, NH_3, THF, -60 °C, 57%; ii) $PhCH(OMe)_2$, CSA, DMF, 61%. (b) CCl_3CN, DBU, DCM, 0 °C. (c) $NaHCO_3$, I_2, H_2O, two steps yield: 41%. (d) 37% HCl aq., dioxane. (e) $NaHCO_3$, MeOH, two steps yield: 63%. (f) 1-azido-8-iodooctane, K_2CO_3, DMF, 80 °C, 39%. (g) **13**, **14**, **15** or **16**, $CuSO_4$ (1.0 M in H_2O), sodium ascorbate (1.0 M in H_2O), DMF, 38% **3**, 11% **4**, 24% **5**, 23% **6**. (h) *m*CPBA, DCM, 40 °C, 44%. (i) $Pd(OH)_2/C$, H_2, MeOH, 68%;

In vitro inhibition and labeling of recombinant human GAA

The inhibitory properties of compounds **1**-**7** (Figure 2) were firstly tested toward rGAA (Myozyme, Genzyme) at pH 4.0 using 4-methylumbelliferyl-α-glucose as substrate (Figure 2A). All compounds proved to be potent rGAA inhibitors with apparent IC_{50} values in the nanomolar range. The most optimized potent inhibitor of the series is cyclophellitol aziridine **1**, with epoxide **7** and *N*-alkyl aziridine **2** within the same range, albeit slightly weaker (Figure 2A). Fluorescent ABPs **3**-**5** all showed about 10-fold increased apparent IC_{50} values compared to **1**, and biotin-conjugated ABP **6** a further 2-fold increase. *In situ* enzyme inhibition within human fibroblasts exhibited similar potency, except for biotin compound **6** (IC_{50} > 10 µM), which apparently is the least cell-permeable of the series. The pH optimum of 5.0 found for labeling of 110 kDa rGAA differed slightly from that of enzymatic activity towards 4MU-α-glucose pH

4.0 (Figure 2B). When rGAA was pre-incubated for 30 min with 10 µM of compounds **1**-**2** and **4**-**7**, followed by labeling with compound **3**, green fluorescent labeling of rGAA by compound **3** was abrogated (Figure 2C). Likewise, the presence of high concentrations of the substrates 4MU-α-glucose and maltose reduced labeling by compound **3** as does prior denaturation of rGAA by incubation with 2% SDS. Inhibitors **1**-**7** reacted too fast for kinetics analysis: for instance, time dependent labeling of rGAA by 100 nM ABP **3** at 4 °C and 37 °C demonstrated most of the enzyme could be labeled within 2 min (SI, Figure S2A).

A

Compounds	IC_{50} (nM)	
	In vitro	*In situ*
1, CF021	37.8 ± 3.4	71.2 ± 9.0
2, CF022	51.7 ± 2.3	75.8 ± 10.9
3, JJB382	458.9 ± 38.4	502.2 ± 4.1
4, JJB347	383.1 ± 22.1	561.6 ± 352.2
5, JJB383	268.7 ± 7.1	42.2 ± 4.9
6, JJB384	767.1 ± 3.5	> 10,000
7, JJB307	54.1 ± 4.9	97.6 ± 14.5

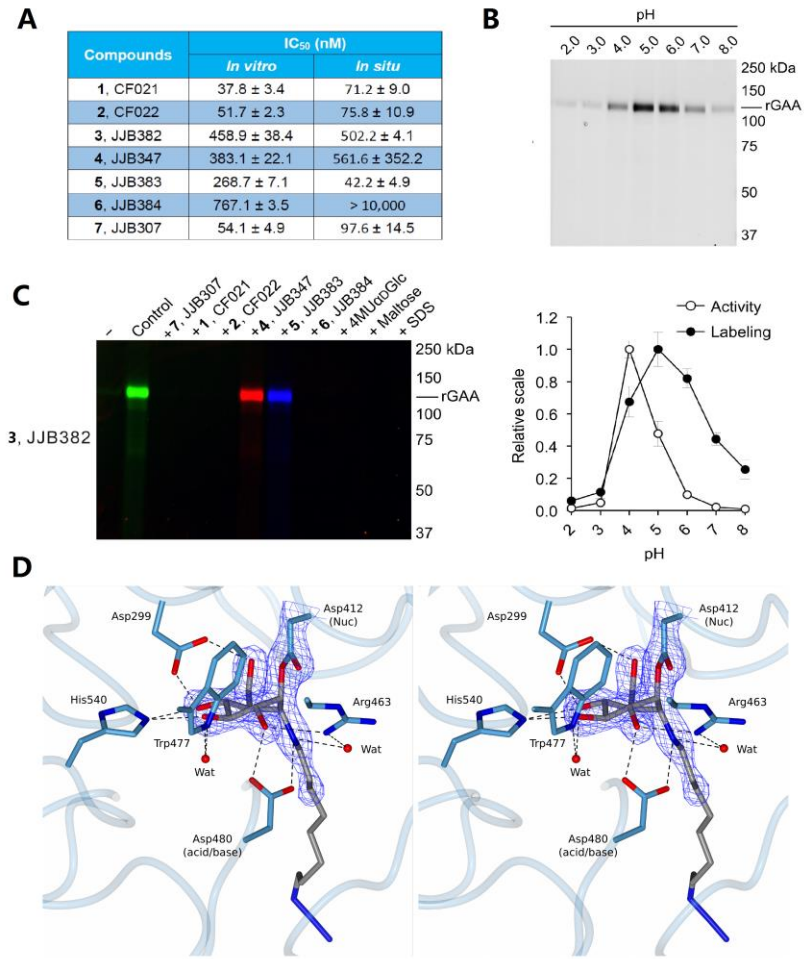


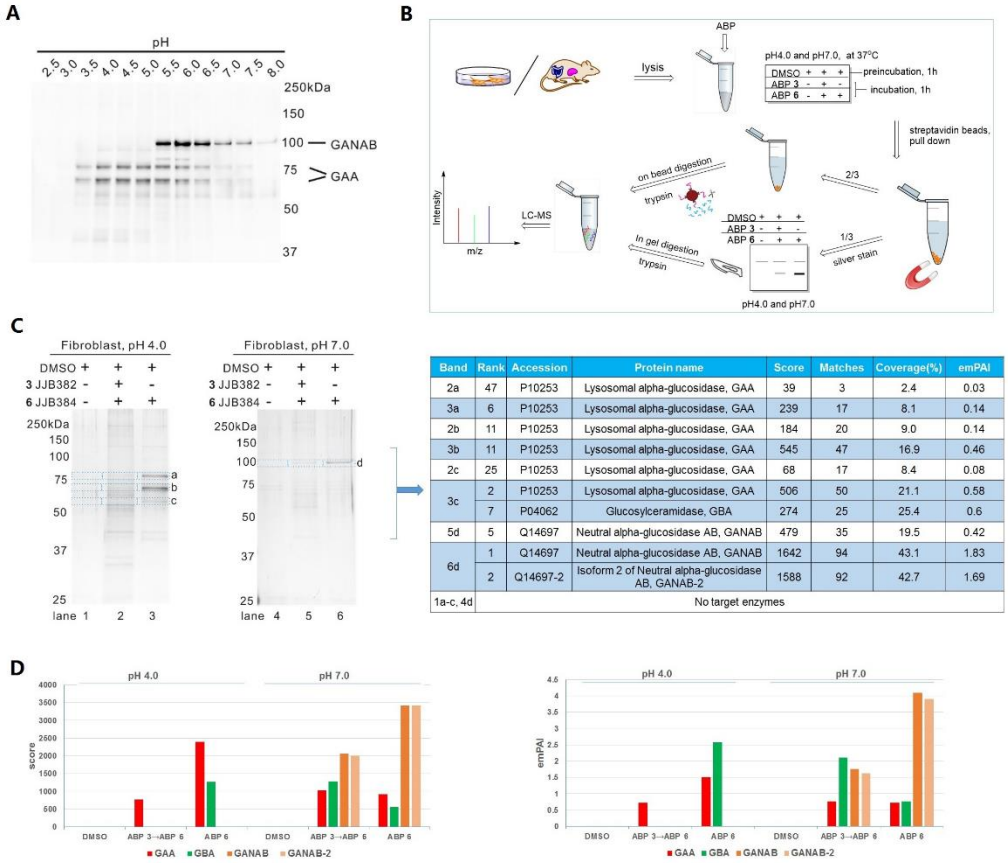
Figure 2. *In vitro* inhibition and labeling of α-glucosidases. A) Inhibition of recombinant GAA. Apparent IC_{50} values (extrapolated with one phase exponential association) are the average values of two separate experiments measured in duplicate, error ranges depict standard deviation B) Labeling of rGAA with compound **3** at various pH as compared to activity towards 4-MU-α-D-glucopyranose. Error bars represent standard deviation. C) ABP labeling of rGAA with **3** competed with compounds **1**-**2** and **4**-**7**. D) Stereoscopic view of the *Cj*Agd31B active site in complex with compound **2**, showing covalent link to C*j*Agd31B enzymatic nucleophile Asp412, and H-bonding interactions to neighboring residues. Electron density is REFMAC maximum-likelihood/σ_A-weighted 2 F_o–F_c synthesis contoured at 0.49 electrons per Å^3.

To establish the mechanism-based mode of action of the probes, a bacterial homologue of GAA from CAZy family GH31, the α-glucosidase from Cellvibrio japonicus, C*j*Agd31B, was used, which is readily amenable to structure studies of GH31 ligand complexes and displays 27% identity to the human GAA enzyme over 615 amino acids including absolute conservation of the active centre "-1" glucosyl site.[24] The 1.95 Å resolution structure of *Cj*Agd31B (Figure 2D) soaked with cyclophellitol aziridine **2** revealed unambiguous electron density showing covalent binding of the ring opened cyclophellitol aziridine β-linked to the active site nucleophile residue (Asp412; equating to Asp518 of GAA and Asp542 of GANAB). Apart from demonstrating irrevocably the mechanism-based mode of action of the probes, the GH31 complex with **2** is consistent with the reaction pathway employed by GH31 in processing α-glucosidic linkages. Substrates are proposed to bind in a $^4C_1$ conformation with catalysis occurring via a $^4H_3$ oxacarbenium ion-like transition state to a covalent adduct in $^1S_3$ skew-boat conformation.[25,26] This $^1S_3$ conformation of the enzyme-inhibitor adduct is well captured in the complex shown in Figure 2D. Similarly, the $^1S_3$ conformation of the enzyme-inhibitor **1** adduct was also well captured in 1.85 Å resolution (SI, Figure S2B).

ABP labeling of multiple α-glucosidases in homogenates of fibroblasts

Having established the expected covalent inactivation, ABP labeling of α-glucosidases was subsequently examined in a more complex and physiologically relevant biological mixture of proteins. For this, homogenates of cultured fibroblasts were prepared and exposed to 1.0 μM ABP **5** at various pH values. Incubation with ABP **5** at pH 4.0 gave clean labeling of what appeared to be the mature lysosomal forms of GAA (two bands at around 70 kDa). At higher pH values an additional protein with a molecular weight of around 100 kDa was observed (Figure 3A). GAA precursor amounts are low in cells and are – as with the mature form – optimally active at low pH. Therefore, it was thought that the protein labeled at 100kDa would unlikely be the GAA pro-form. Rather, it was envisaged that this band would correspond to another α-glucosidase, possibly ER α-glucosidase II. Unambiguously, to establish the nature of the 70 kDa proteins identified at acidic pH as well as the 100 kDa protein seen at neutral pH, affinity purification was performed and chemical proteomics of fibroblast homogenates were treated with biotin-conjugated ABP **6**. A lysate of fibroblasts was incubated at pH 4.0 and pH 7.0 with ABP **6**, in either the presence or absence of pre-treated ABP **3**. Next, biotinylated proteins were enriched by pull down using streptavidin-coated magnetic beads. Loaded streptavidin beads were split for in-gel digestion and on-bead digestion. Captured proteins for subsequent in-gel digestion were released in Laemmli buffer, resolved on SDS-PAGE and visualized by silver staining (Figure 3B). Three distinct bands were obtained that all could be competed for by inclusion of α-cyclophellitol aziridine **3**. Two bands with apparent molecular masses of around 70 kDa are most prominent in the pull-down at pH 4.0 and one band at approximately 100 kDa was the predominant species labeled at pH 7.0. Tryptic digestion was performed on these three protein bands and the resulting tryptic peptides were analyzed by

nanoscale liquid chromatography coupled to tandem mass spectrometry (nano-LCMS/MS). The proteins were identified via matching of the obtained peptide sequences against the Mascot database. In this manner, the ~70 kDa proteins were identified as the two mature forms (70 and 76 kDa) of GAA (Figure 3C). The labeled 100 kDa protein proved to be GANAB and isoform 2 of GANAB; the known retaining ER α-glucosidase II.



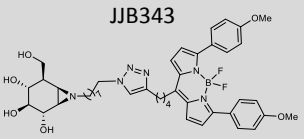
Band	Rank	Accession	Protein name	Score	Matches	Coverage(%)	emPAI
2a	47	P10253	Lysosomal alpha-glucosidase, GAA	39	3	2.4	0.03
3a	6	P10253	Lysosomal alpha-glucosidase, GAA	239	17	8.1	0.14
2b	11	P10253	Lysosomal alpha-glucosidase, GAA	184	20	9.0	0.14
3b	11	P10253	Lysosomal alpha-glucosidase, GAA	545	47	16.9	0.46
2c	25	P10253	Lysosomal alpha-glucosidase, GAA	68	17	8.4	0.08
3c	2	P10253	Lysosomal alpha-glucosidase, GAA	506	50	21.1	0.58
	7	P04062	Glucosylceramidase, GBA	274	25	25.4	0.6
5d	5	Q14697	Neutral alpha-glucosidase AB, GANAB	479	35	19.5	0.42
6d	1	Q14697	Neutral alpha-glucosidase AB, GANAB	1642	94	43.1	1.83
	2	Q14697-2	Isoform 2 of Neutral alpha-glucosidase AB, GANAB-2	1588	92	42.7	1.69
1a-c, 4d	No target enzymes						

Figure 3. Labeling of multiple α-glucosidases and their identification by proteomics. A) Labeling of proteins in fibroblast lysate with compound **3**. B) Work flow of pull down proteomics experiments. C) In-gel digestion silver staining and identification of target proteins modified by biotin-ABP **6**. D) Glycosidases identification with Mascot emPAI and score values after on-bead pull-down and processing of fibroblast lysate with biotin-ABP **6**.

Using on-bead trypsin digestion very similar results were obtained. GAA in the pH 4.0 incubation and GANAB in the pH 7.0 incubation topped the exponentially modified Protein Abundance Index (emPAI), a relative quantitation of the proteins in a mixture based on protein coverage by the peptide matches in a database search result, and protein scores from Mascot search report (Figure 3D and SI, Figure S3). Comparable experiments with mouse liver lysates (SI, Figure S4, S5) resulted in identification of GAA and GANAB as the major targets of biotin-

aziridine ABP **6**, matching the fluorescent bands on SDS-PAGE. Using a final concentration of 10 μM of biotin-aziridine **6** in the affinity purification experiments, enrichment of a protein with apparent molecular mass of 60-65 kDa was also observed (SI, Figure S4). Proteomics analysis identified this protein as GBA, the lysosomal acid β-glucosidase deficient in Gaucher disease. To corroborate this observation, the possible inhibition of enzymatic activity of pure GBA by compounds **1**-**7** was determined. Indeed, GBA is also inhibited by all of the prepared compounds with apparent IC_{50} values in the range of 592.8 nM – 155.3 μM (Table 1). For comparison, the corresponding β-configured cyclophellitol aziridine JJB343[15] was a far more potent inhibitor towards GBA with 500-fold lower apparent IC_{50} values than α-aziridine **4** JJB347 in the same test.

Table 1. Apparent IC_{50} values of α-configured cyclophellitol azridrines **1-6** epoxide **7** JJB307 and on rGBA, determined *in vitro*. JJB343 (cyclophellitol-β-aziridine-BODIPY) was included for comparison. Data were average values of two separate experiments measured in duplicate, error ranges depict standard deviation.

Compounds	*In vitro* IC_{50} (nM)
1 CF021	41,943.5 ± 490.0
2 CF 022	603.2 ± 28.0
3 JJB382	592.8 ± 214.3
4 JJB347	1,076.0 ± 63.6
5 JJB383	815.8 ± 526.4
6 JJB384	2,060.0 ± 219.2
7 JJB307	155,333.0 ± 1,773.4
JJB343 	2.05 ± 0.93

Labeling of α-glucosidases present in murine gastrointestinal tracts

Given the observed high affinity labeling with the ABPs of both GAA and GANAB, labeling of proteins was examined in extracts from intestines, tissue known to contain additional retaining GH31 α-glucosidases. For this purpose, mouse intestine was freshly collected, food remains and bacteria removed by rinsing and the tissue lysed in KPi-Triton buffer using sonication. The lysate was incubated with biotinylated compound **6** and bound proteins were analyzed by gel electrophoresis and chemical proteomics as before (SI, Figure S6, S7). In this way, biotin-ABP **6** could label GAA and GANAB, as well as sucrase-isomaltase (Sis) and maltase-glycoamylase (MGAM), two retaining α-glucosidases expressed specifically in intestinal tissue. In addition, the retaining β-glucosidases, lactase (Lct) and GBA were identified. The findings were recapitulated by labeling lysates of pull-down and supernatant samples with ABPs **3** and

6, protein separation by gel electrophoresis and fluorescence imaging of the wet gel slabs, followed by transferring protein to polyvinylidene fluoride (PVDF) membrane for Western blot analysis with HRP-streptavidin. Fluorescent bands, chemiluminescence bands and silver stain protein bands matched well, and the expected molecular masses for GAA, GANAB, Sis, MGAM, Lct and GBA were observed (SI, Figure S6).

Labeling and concomitant inhibition of α-glucosidases in intact cells.

Prompted by the effective *in vitro* labeling of GH31 α-glucosidases in cell and tissue extracts, the labeling of enzymes in intact cultured fibroblasts was studied with ABPs **3** and **5**. Fibroblasts were exposed to 100 nM ABP **5** for various times (from 10 min to 4 h) by including these in the culture medium. Subsequently, cells were extensively washed and harvested, lysates were prepared at 4 °C and then separated by SDS-PAGE followed by fluorescence scanning of the wet gel slabs (Figure 4A left). Fluorescent labeling of GAA and GANAB was prominent in blue bands. Of note, no GBA was concomitantly labeled *in situ*, illustrating the far lower affinity of the α-glucopyranose ABP to label this β-glucosidase, but GBA was labeled by excess of ABP **3** under *in vitro* conditions (Figure 4A right). During the procedure an excess of ABP **3** was added to exclude artificial labeling of enzymes by non-reacted free ABP **5** during the process of cell lysis. This addition did not change the result, demonstrating that labeling of GAA and GANAB truly occurs *in situ*.

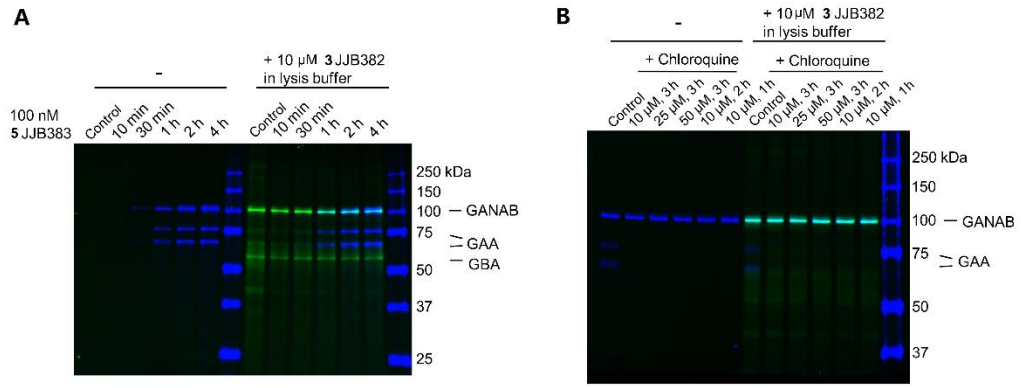


Figure 4. *In situ* labeling of GAA and GANAB in living fibroblasts. A) Time-dependent labeling of GAA and GANAB in fibroblasts by ABP **5** *in situ* and ABP **3** *in vitro*. B) Chloroquine blocks *in situ* GAA labeling.

GAA requires an acid pH for labeling by the α-glucopyranose ABP (Figure 2A). The effect of chloroquine, an agent that raises intralysosomal pH, was therefore investigated.[27] Fibroblasts were incubated with either increasing amount of chloroquine (10, 25 and 50 μM) for three h or increasing treatment time (1, 2 and 3 h) after which ABP **5** was added to the medium. After the time points cells were washed, harvested and analyzed on labeled protein. As shown in Figure 4B, the presence of chloroquine abolished *in situ* labeling of GAA but not GANAB.

Diagnostic application: probing for GAA activity in Pompe disease patient material

Finally, the value of the ABPs for diagnosis of Pompe disease was studied. Fibroblasts of normal individuals and Pompe disease patients suffering from the infantile and adult variants of disease were cultured. Cell lysates were labeled with 0.5 µM of ABP **3** for 30 min at pH 4.0 and pH 7.0. Gel electrophoresis of the denatured protein content and subsequent fluorescence imaging and immunoblots revealed a prominent distinction between material from normal persons and Pompe disease patients (Figure 5). Marked absence of fluorescently labeled mature 70-76 kDa GAA with concomitant normal levels of GANAB was only observed for the patient fibroblasts. This finding demonstrates the potential of α-glucosidase ABPs in laboratory diagnosis of Pompe disease.

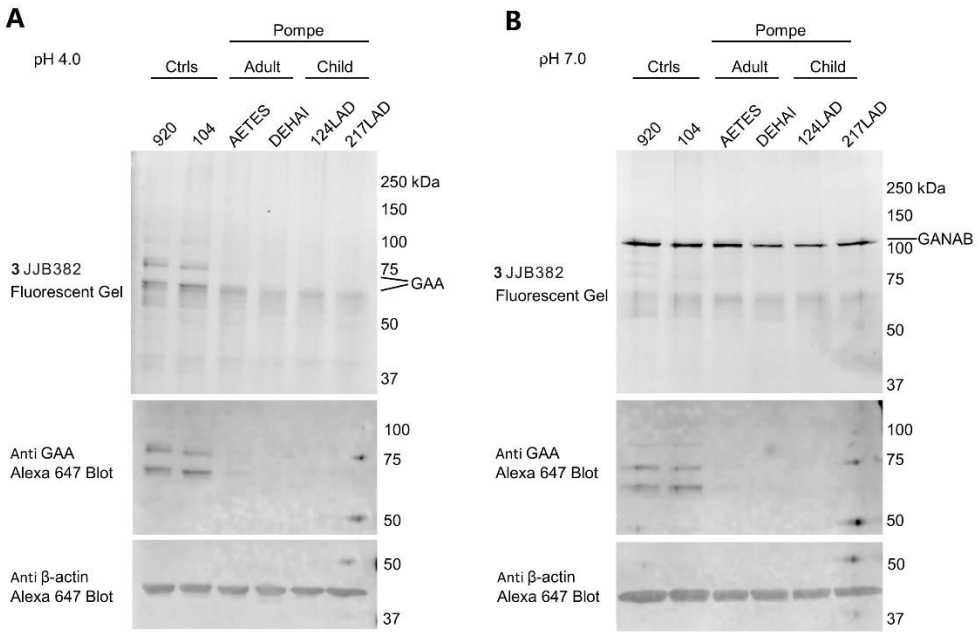


Figure 5. *In vitro* ABP labeling and Western blot detection of α-glucosidases in wild type and Pompe fibroblasts. A) *In vitro* GAA labeling at pH 4.0 with **3**, followed by Western blot detection of GAA in various fibroblast lysates, containing wild type or mutant (Pompe) GAA. B) *In vitro* GANAB labeling at pH 7.0 with **3** followed by Western blot detection of GAA in various fibroblast lysates, containing wild type or mutant (Pompe) GAA**.**

5.3 Discussion

Activity-based protein profiling (ABPP) has emerged in the past fifteen years as a powerful technique to identify enzymes and to study their activity in the context of the physiological processes they partake in, both *in vitro* and *in situ*. In the first instance developed for serine hydrolases and cysteine proteases,[28,29] ABPs can on paper be designed for any enzyme, provided that a covalent enzyme-substrate adduct emerges during enzyme action. Retaining

glycosidases that employ a Koshland two-step double displacement mechanism fulfill this requirement, and it has been reported previously on the versatility of the natural product, cyclophellitol, as a scaffold for activity-based glycosidase probe design. In this work, the generic design principle is adopted by the development of a set of α-configured cyclophellitol aziridines as in-class GH31 retaining α-glucosidase ABPs. The probes label GH31 retaining α-glucosidases in a tissue-dependent fashion and allow detection and identification of these by in-gel fluorescence and by chemical proteomics. They are highly selective towards GH31 retaining α-glucosidases when applied in the appropriate concentration and even at higher concentrations show little to no cross-reactivity, apart from labeling the retaining β-glucosidase, GBA. The ABPs label subcellular α-glucosidases optimally at the pH at which the α-glucosidases display maximal enzymatic activity. Thus, whereas GAA in fibroblasts is maximally labeled at pH 4.0-5.0, GANAB, or ER-α-glucosidase II, is optimally labeled at neutral pH. This finding underscores that the probes report on functional enzymes, rather than protein levels, an observation that is corroborated by the X-ray structure of C*j*Agd31B rGAA, in which ABP precursor **2** has reacted with the active site nucleophile, Asp412.

The tissue-specific and pH-dependent mode of action of the ABPs allow for probing each of the targeted α-glucosidases independently and within their physiological context, as well as in health and disease. The diagnostic value of the probes is demonstrated in the assessment of the lack of GAA activity in infantile and adult Pompe disease tissue (Figure 5). The in-class broad-spectrum GH31 α-glucosidase probes may also find use in the discovery of inhibitors specific for either GAA or GANAB in a competitive ABPP assay. Selective inhibitors for GAA or GANAB would be of interest: GAA inhibitors for pharmacological chaperone discovery in the context of Pompe disease and GANAB inhibitors for antiviral or anticancer drug discovery (GANAB activity being a crucial factor in the quality control of nascent *N*-linked glycoproteins). An initial competitive ABPP was performed in which fibroblasts expressing GAA and GANAB were first treated with varying concentrations of the iminosugars **18**-**22** and maltose **23** followed by incubation at both pH 4.0 and pH 7.0 with ABP **5**. As can be seen in the Figure 6 these compounds efficiently inhibit both enzymes. Indeed no selective inhibitors for either enzyme are known and the ABPs may be of help in identifying such compounds. The ABPs finally also efficiently target and identify intestinal α-glucosidases, which are key players in glucose assimilation and the targets of the anti-diabetic drugs, miglitol and acarbose. The ability to also study these intestinal enzymes in detail in their physiological surroundings holds promise for the future identification of more effective (and selective with respect to GAA/GANAB inhibition) inhibitors, which may come from natural sources and become part of nutraceutical regimes.

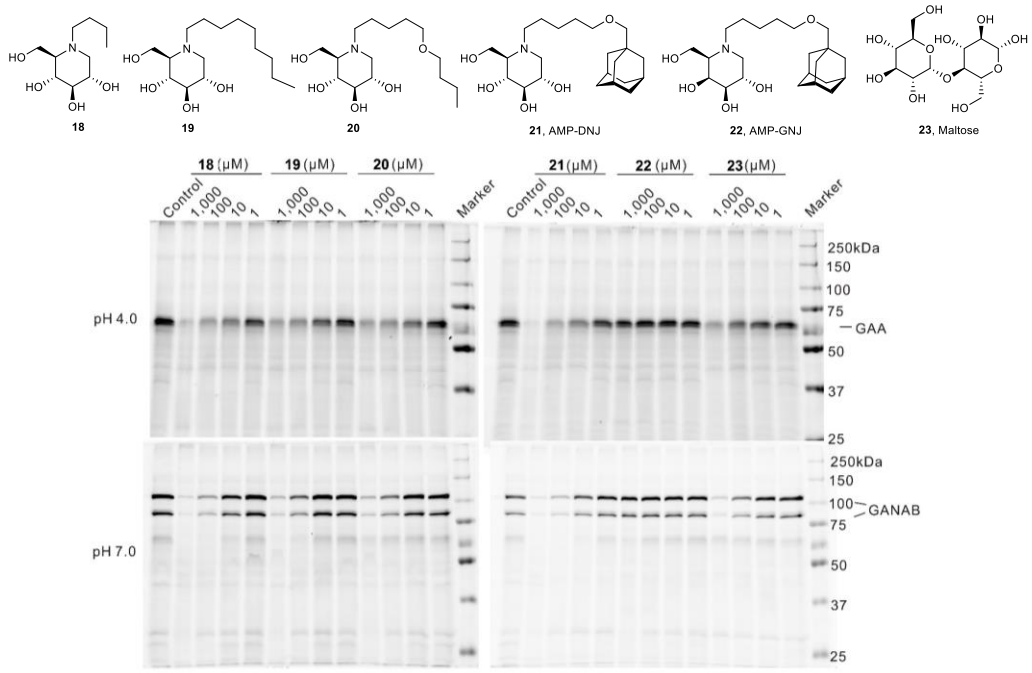


Figure 6. ABP **5** competitive labeling of α-glucosidases in mouse liver extract at pH 4.0 and pH 7.0 with pre-incubation of inhibitors **18-23**.

5.4 Experimental methods

Biological assays:

Materials

Chemicals were obtained from Sigma-Aldrich, if not otherwise indicated. Trypsin and Endoproteinase GluC were commercially available from Promega. Recombinant GAA was obtained from Genzyme (Cambridge, MA, USA). Fibroblasts were obtained with consent from donors. Pompe patients were diagnosed on the basis of reduced GAA activity. Cell lines were cultured in HAMF12-DMEM medium (Invitrogen) supplied with 10% (v/v) FCS. Mouse tissue were isolated according to guidelines approved by the ethical committee of Leiden University (DEC#13191). All the cell or tissue lysates were prepared in potassium phosphate lysis buffer(25 mM in pH 6.5, supplemented with protease inhibitor 1x cocktail (Roche)) via homogenization with silent crusher S equipped with Typ 7 F/S head (30 rpm x 1000, 3 × 7 sec) on ice and lysate concentration was determined with BCA Protein Assay Kit (Pierce™). The protein fractions were stored in small aliquots at -80 °C until use.

Enzyme activity assays and IC_{50} measurements.

The α-D-glucosidase activity of lysosomal α-D-glucosidase GAA was assayed at 37 °C by incubating with 3.0 mM 4-methylumbelliferyl-α-D-glucopyranoside as substrate in 150 mM McIlvaine buffer, pH 5.0, supplemented with 0.1% (w/v) BSA. Activity of rGBA was measured using similar conditions but with 3.8 mM 4-methylumbelliferyl-β-D-glucopyranoside as substrate at pH 5.2, supplemented with 0.1% (v/v) Triton X-100 and 0.2% (w/v) sodium taurocholate. To determine the apparent *in vitro* IC_{50} value, recombinant GAA or rGBA was firstly pre-incubated with a

range of inhibitor dilutions for 30 min at 37 °C, prior to addition of the substrate. To determine the influence of pH on the enzymatic activity, enzyme mixtures were firstly pre-incubated for 30 min on ice with McIlvaine buffers of pH 2.0–8.0 whereafter substrate was added, dissolved in nanopure H_2O. The enzymatic reaction was quenched by adding excess NaOH-glycine (pH 10.3), after which fluorescence of liberated 4-methylumbelliferyl was measured with a fluorimeter LS55 (Perkin Elmer) using λ_{EX} 366 nm and λ_{EM} 445 nm. The *in situ* IC_{50} value was determined by incubating fibroblast cell lines expressing wild-type GAA, grown to confluence, with a range of inhibitor dilutions for 2 h. Hereafter, cells were washed three times with PBS and subsequently harvested by scraping in potassium phosphate buffer (25 mM K_2HPO_4-KH_2PO_4, pH 6.5, supplemented with 0.1% (v/v) Triton X-100 and protease inhibitor cocktail (Roche)). Residual GAA activity was measured using the aforementioned substrate assay. All *in situ* IC_{50} values were determined by replicating each assay twice in duplo in two separate cell lines. Data was corrected for background fluorescence, then normalized to the untreated control condition and finally curve-fitted via one phase exponential decay function (GraphPad Prism 5.0).[15,16]

ABP 7 pull-down and LC-MS/MS analysis

3mg total protein from human fibroblast lysate or 6 mg total protein from mouse intestines lysate was incubated with either 0.1% (v/v) DMSO, 5.0 µM **7** JJB384, or firstly with 5.0 µM **3** JJB382 followed by 5.0 µM **6** JJB384. For mouse liver lysate, 6.0 mg total protein and 10 µM ABPs **3** and **6** were used. Each step taking 30 min at 37 °C, in a total volume of 0.50 mL McIlvaine buffer of pH 4.0 and pH 7.0, subsequently denatured through the addition of 10% (w/v) SDS 125 µL and boiling for 5 min at 100 °C. From here on, samples were prepared for pull-down with streptavidin beads as published earlier[2]. After pull-down procedure, the samples were divided, 2/3 for on-bead digestion and 1/3 for in-gel digestion. On-bead digestion samples including beads were treated by the trypsin digestion buffer (a mixture containing 100 mM Tris-HCl pH 7.8, 100 mM NaCl, 1.0 mM $CaCl_2$, 2% MeCN and 10 ng/µL trypsin) and the bead suspension was incubated in a shaker at 37 °C overnight. The supernatant containing the trypsin-digested peptides was desalted using stage tips, followed by evaporation of MeCN and dilution in 70 µL sample solution (H_2O/MeCN/TFA, 95/3/0.1, v/v/v) for LC-MS analysis. The beads containing active-site peptides were treated with endoproteinase Glu-C digestion buffer (100ng/µL in PBS solution); incubated in a shaker at 37 °C overnight after which the supernatant was desalted using stage tips and for LC-MS. In-gel digestion samples were eluted by boiling the beads at 100 °C in 30 µL of 1 x Laemmli buffer. The eluted proteins were separated on 10% protein gels at 200 V for 1 h, and the protein gels were silver stained using the Invitrogen kit,[30] and visualized by Bio-Rad Chemi-Doc MP Imager using the silver stain channel. Bands were excised with a surgery knife by hand and treated with in-gel digestion buffer (10 mM NH_4HCO_3, 5% ACN, 1mM $CaCl_2$, 10 ng/µL trypsin). The supernatant containing the trypsin-digested peptides was desalted using Stage Tips and prepared for LC-MS. All the peptide samples were analyzed with a 2 h gradient of 5%→25% ACN on nano-LC, hyphenated to an LTQ-Orbitrap and identified via the Mascot protein search engine, and the Raw data was calculated by MaxQuant program against the Uniprot human or mouse proteome database to present the protein identification list.[2] Mascot identifications were manually validated. The identification results were exported as Excel file including protein accession numbers, Mascot peptide scores, mass of the protein, % coverage of the protein by amino acids identified by LC-MS, peptide matches, miss cleavages, C-terminal peptides and protein emPAI values. [30]

SDS-PAGE analysis and fluorescence scanning

For labeling procedures see below. All the labeling samples were pre-incubated in 150 mM McIlvaine buffer on ice for 5 min. Electrophoresis was performed with sodium dodecylsulfate containing 10% polyacrylamide gels. Wet slab-gels were then scanned for ABP-emitted fluorescence using a Bio-Rad ChemiDoc MP imager using green Cy2 (λ_{EX} 470 nm, bandpass 30 nm; λ_{EM} 530 nm, bandpass 28) for **3** JJB382, red Cy3 (λ_{EX} 530 nm, bandpass 28 nm; λ_{EM} 605 nm, bandpass 50) for **4** JJB347, and blue Cy5 (λ_{EX} 625 nm, bandpass 30 nm; λ_{EM} 695 nm, bandpass 55) for **5** JJB383. All samples were denatured with 5× Laemmli buffer (50% (v/v) 1.0 M Tris−HCl, pH 6.8, 50% (v/v) 100% glycerol, 10% (w/v) DTT, 10% (w/v) SDS, 0.01% (w/v) bromophenol blue), boiled for 5 min at 100 °C, and separated by gel electrophoresis on 10% (w/v) SDS-PAGE gels running continuously at 90 V for 30 min and 200 V for 50 min.

***In vitro* labeling of GAA**

The detection limit of each ABP was analyzed by labeling 1 pmol rGAA with 1,000−0.01 fmol **3** JJB382 in 150 mM McIlvaine buffer, pH 5.0, for 30 min at 37 °C. Influence of pH on ABP labeling involved pre-incubation of either 100 fmol rGAA or 10 µg murine liver lysate at pH 2−8 for 30 min on ice, prior to addition of 1.0 µM **3** JJB382, dissolved in Nanopure H_2O and incubating for 30 min at 37 °C. Assessment of 3 JJB382 labeling kinetics involved pre-cooling of 100 fmol rGAA on ice for 15 min, followed by addition of similarly cooled 100 nM **3** JJB382 solution. After mixing, **3** JJB382 labeling was chased for 0−60 min at either 4 °C or 37 °C, whereafter labeling was stopped by denaturation. For competitive ABPP on rGAA, 1 pmol rGAA was pre-incubated with inhibitors (100 µM **7** JJB307, **1** Chris021, **2** Chris022, 10 µM **4** JJB347, **5** JJB383 or **6** JJB384, 10 mM 4MU-α-D-glucopyranoside, 2.5 M maltose) for 30 min at 37 °C, or boiled for 4 min in 2% (w/v) SDS, prior to labeling with 1.0 µM **3** JJB382 for 30 min at 37 °C. Labeling of human fibroblast lysates was performed on 100 µg total protein, using 1.0 µM **3** JJB382 dissolved in either 150 mM McIlvaine buffer, pH 4.0 for labeling of GAA, or pH 7.0 for labeling of GANAB. For pH optimum labeling on lysate of C57Bl6/J mouse liver or human fibrlblast, 50 µg total protein was incubated with 1.0 µM **3** JJB382 in McIlvaine buffers of pH 2.0−8.0 for 30 min at 37 °C. For competitive ABPP on C57Bl6/J mouse liver, 50 µg of total protein was pre-incubated with compounds 18-23 with decreasing concentrations (1000 µM, 100 µM, 10 µM, 1.0 µM) in 150 mM McIlvaine buffer pH 4.0 and pH 7.0 for 30 min at 37 °C, prior to labeling with 1.0 µM **5** JJB383 for 30 min at 37 °C.

Visualizing pulled-down proteins by fluorescent detection and Western Blot.

After pull-down of mouse intestinal enzymes using **6** JJB384, the supernatants and eluents (20 µL from both) of all treatment conditions (DMSO, competition, and **6** JJB384, under pH 4.0 or pH 7.0) were subjected to SDS-PAGE and fluorescent detection (Bio-Rad ChemiDoc MP Imager; Cy2 channel). Proteins were subsequently transferred to PVDF membranes using a Trans-Blot® Turbo system (BioRad), blocked in 4% (w/v) BSA in TBST, and detected for biotin **6** JJB384 binding with Streptavidin-HRP antibody (Dako). The blot was developed in the dark using a 10 mL luminal solution, 100 µL ECL enhancer and 3.0 µL 30% H_2O_2 solution. Chemiluminescence was visualized using the same ChemiDoc imager (BioRad).

***In situ* labeling of GAA**

Fibroblasts were treated with **5** JJB383 (100 nM, for 10 to 240 min) in culture medium. After washing 3 times with PBS, cells were harvested in potassium phosphate buffer (25 mM K_2HPO_4-KH_2PO_4, pH 6.5, supplemented with 0.1% (v/v) Triton X-100 and protease inhibitor cocktail (Roche)), prepared with or without 10 µM **3** JJB382. 20 µg total protein

from each cell homogenate of the lysate was subjected to SDS-PAGE. Labeling were visualized by fluorescent detection using Bio-Rad ChemiDoc MP Imager under the channels Cy5 for **5** JJB383 and Cy2 for **3** JJB382 with the above described settings. For Chloroquine treatment experiment, cultured fibroblasts were incubated with chloroquine (10-50 µM, for up to 3 h) and subsequently with **5** JJB383 (100 nM, for 3 h) in medium. Cells were harvested with or without 10 mM **3** JJB382 with the above describe methods. Homogenates were denatured, resolved on SDS-PAGE and detected for **5** JJB383 labeling by fluorescent scanning.

Pompe GAA detection by Western blot

Following fluorescent scanning of SDS-PAGE, proteins on wet slab gel were transferred to PVDF membrane and blocked as described in the previous section. For GAA detection, the membrane was incubated firstly with mouse polyclonal anti GAA and subsequently with goat anti mouse Alexa647 (Life Technologies). Blot was scanned on a Typhoon FLA 9500 Imager (GE Healthcare) using 633 nm laser and LPR filter, and 100 mm as pixel size. Rabbit anti β-actin (Cell Signaling) and goat anti rabbit Alexa647 (Invitrogen) were used for loading control.

***Agd31B* expression and 3-D crystallography**

*A*gd31B expression and purification was carried out as previously described.[24] Protein crystals were obtained using 1.8 M ammonium sulfate, 0.10 M HEPES (pH = 7.0), 2% PEG 400 at 20 °C by the sitting drop vapor diffusion method. Crystal complexes with **1** CF021 and **2** CF022 were obtained by soaking in mother liquor containing 5.0 mM probe for 2 h, before cyroprotecting in 2.0 M lithium sulfate, 0.10 M HEPES (pH = 7.0), 2% PEG 400, and flash freezing in liquid N_2 for data collection (Table 2).

All data were collected at beamline I04 of the Diamond Light Source, processed using XDS[31] and reduced using Aimless.[32] Complex structures were solved by molecular replacement using MolRep,[33] before subsequent rounds of manual model building and refinement using Coot[34] and REFMAC5[35] respectively. Refinements were carried out using TLS determination of molecular motions.[36] Ligand coordinates were built using jLigand.[37] Crystal structure figures were generated using ccp4mg.[38]

Table 2. Crystal data collection and refinement statistics

	C*j*Agd31B-**2** CF022-complex PDB: 5123	C*j*Agd31B- **1** Chris021-complex PDB: 5124
Data collection		
Space group	P622	P622
Cell dimensions		
a, b, c (Å)	198.0, 198.0, 103.0	197.3, 197.3, 102.9
α, β, γ (°)	90, 90, 120	90, 90, 120
Resolution (Å)	85.74-1.95 (1.99-1.95)	49.83-1.85 (1.88-1.85)
R_{merge}	0.13 (1.64)	0.10 (1.21)
$I/\sigma I$	19.6 (2.3)	26.8 (3.0)
Completeness (%)	100 (100)	100 (100)
Redundancy	20.0 (20.1)	24.5 (25.0)
Refinement		
Resolution (Å)	85.74-1.95	49.83-1.85
No. reflections	82110	95234
R_{work}/ R_{free}	0.17/0.19	0.17/0.19
Protein	6299	6295
Ligand/ion	102	81
Water	616	625
B-factors (TLS refinement)		
Protein	33.8	31.99
Ligand/ion	64.8	54.42
Water	39.7	38.0
R.m.s deviations		
Bond lengths (Å)	0.012	0.013
Bond angles (°)	1.57	1.57

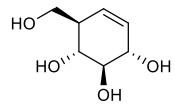
Values in parentheses are for highest-resolution shell.

Synthesis:

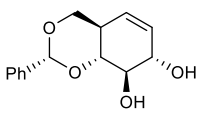
General synthesis

All reagents were of a commercial grade and were used as received unless stated otherwise. Dichloromethane (DCM), tetrahydrofuran (THF) and *N, N*-dimethylformamide (DMF) were stored over 4 Å molecular sieves, which were dried *in vacuo* before use. All reactions were performed under an argon atmosphere unless stated otherwise. Solvents used for flash column chromatography were of pro analysis quality. Reactions were monitored by TLC analysis using Merck aluminum sheets pre-coated with silica gel 60 with detection by UV absorption (254 nm) and by spraying with a solution of $(NH_4)_6Mo_7O_{24}.H_2O$ (25 g/L) and $(NH_4)_4Ce(SO_4)_4.H_2O$ (10 g/L) in 10% sulfuric acid followed by charring at

~150 °C or by spraying with an aq. solution of $KMnO_4$ (7%) and K_2CO_3 (2%) followed by charring at ~150 °C. Column chromatography was performed using either Baker or Screening Device silica gel 60 (0.04 - 0.063 mm) in the indicated solvents. ^{1}H NMR and ^{13}C NMR spectra were recorded on Bruker AV-850 (850/214 MHz), Bruker DMX-600 (600/150 MHz) and Bruker AV-400 (400/100 MHz) spectrometer in the given solvent. Chemical shifts are given in ppm relative to the deuterated chloroform or methanol residual solvent peak or tetramethylsilane (TMS) as internal standard. Coupling constants are given in Hz. All given ^{13}C-NMR spectra are proton decoupled. High-resolution mass spectra were recorded with a LTQ Orbitrap (Thermo Finnigan). Optical rotations were measured on Propol automatic polarimeter (Sodium D-line, λ = 589 nm). LC-MS analysis was performed on an LCQ Adventage Max (Thermo Finnigan) ion-trap spectrometer (ESI^+) coupled to a Surveyor HPLC system (Thermo Finnigan) equipped with a C_{18} column (Gemini, 4.6 mm x 50 mm, 3.0 µm particle size, Phenomenex) equipped with buffers A: H_2O, B: MeCN (MeCN) and C: 1% aq. TFA, For reversed-phase HPLC purifications an Agilent Technologies 1200 series instrument equipped with a semi preparative Gemini C_{18} column (10 x 250 mm) was used. The applied buffers were A: 50 mM NH_4HCO_3 in H_2O, B: MeCN.

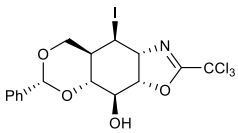


(1R,2R,3S,6R)-6-(Hydroxymethyl)-cyclohex-4-ene-1,2,3-triol The starting material diol **8** was prepared according to previously reported procedure.[22] Diol compound **8** (680 mg, 2.0 mmol, 1.0 eq.) was co-evaporated 3 times with toluene and then dissolved in dry THF (25 mL) and cooled to -60°C under argon atmosphere. Ammonia (20 mL) was condensed at -60°C under argon atmosphere. Lithium (207 mg, 30.0 mmol, 15 eq.) was added and the mixture was stirred until lithium was completely dissolved. To this solution was slowly added the solution of **8** in THF. The reaction mixture was stirred for 30 min at -60 °C and then quenched by adding of H_2O (30 mL). The resulting solution was allowed to come to room temperature and stirred until all ammonia had evolved. Next, the solution was concentrated *in vacuo*, redissolved in H_2O, and neutralized with Amberlite IR-120 H^+. Then, the filtration mixture was concentrated *in vacuo* and purified by silica gel column chromatography (10%→20%, MeOH in DCM) to afford product as colorless oil (180 mg, 1.1 mmol, 57%). TLC: R_f 0.48 (DCM/MeOH, 4/1, v/v); $[\alpha]_D^{20}$ +105.2 (*c* = 1, MeOH); ^{1}H-NMR (400 MHz, CD_3OD): δ ppm 5.66 – 5.57 (m, 2H), 4.06 – 4.03 (m, 1H), 3.81 (dd, *J* = 10.6, 4.1 Hz, 1H), 3.66 - 3.59 (m, 1H), 3.49 - 3.41 (m, 2H), 2.31 - 2.26 (br, 1H); ^{13}C-NMR (100 MHz, CD_3OD): δ ppm 130.97, 128.57, 78.81, 73.62, 72.03, 63.45, 47.65; HRMS: calculated for $C_7H_{12}O_4$ $[M+H^+]$ 161.08084, found: 161.08087.

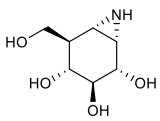


(1R,3R,6R,9S,10S)-9,10-Dihydroxy-3-phenyl-2,4-dioxabicyclo[4.4.0]dec-7-ene (9) (1R,2R,3S,6R)-6-(Hydroxymethyl)-cyclohex-4-ene-1,2,3-triol (180 mg, 1.1 mmol, 1.0 eq.) was dissolved in dry DMF (2.0 mL) and dry MeCN (6.0 mL) in an inert atmosphere. 10-Camphorsulfonic acid (52 mg, 0.25 mmol, 0.20 eq.) was added to the solution, followed by $PhCH(OMe)_2$ (253 µL, 1.7 mmol, 1.5 eq.). After 48 h, the reaction was quenched with Et_3N (31.0 µL, 0.23 mmol, 0.2 eq.) and concentrated *in vacuo*. The reaction mixture was separated out with EtOAc and H_2O and the aq. layer was further extracted with EtOAc. The combined organic layers were washed with brine, dried over Na_2SO_4, filtered and concentrated *in vacuo*. The product was purified by silica gel column chromatography (30%→70%, EtOAc in pentane) to afford **9** (170 mg, 0.69 mmol, 61%). TLC: R_f 0.45 (pentane/EtOAc, 7/3, v/v); $[\alpha]_D^{20}$ +30.8 (10 mg/mL in $CDCl_3$); ^{1}H-NMR (400 MHz, $CDCl_3$): δ ppm 7.52 – 7.47 (m, 2H), 7.37 – 7.31 (m, 3H), 5.57 – 5.52 (m, 2H), 5.28 – 5.25 (m, 1H), 4.24 - 4.18 (m, 2H), 3.86 (dd, *J* = 10.1, 7.5 Hz, 1H), 3.58 – 3.51 (m, 2H), 2.84 (d, *J* = 8.0 Hz,1H), 2.60 – 2.54 (m, 1H); ^{13}C-NMR

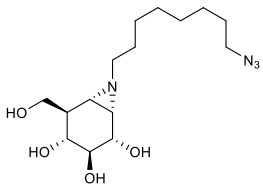
(100 MHz, $CDCl_3$): δ ppm 137.79, 130.49, 130.64, 129.16, 128.34, 126.39, 124.05, 102.16, 80.72, 75.42, 73.74, 69.96, 38.53; HRMS: calculated for $C_{14}H_{16}O_4$ [M+H$^+$] 249.11214, found: 249.11224.



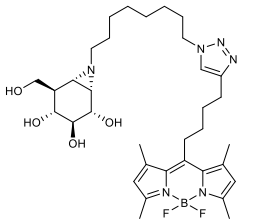
(1R,6R,7R,8R,9S,10R)-10-Hydroxy-7-iodo-3-phenyl-12-trichloromethyl-13-aza-2,4,11-trioxatricyclo[7.4.4.0-8.4.3.0]tridec-12-ene (11) A solution of **9** (535 mg, 2.1 mmol, 1.0 eq.) in dry DCM (60 mL) was made under argon atmosphere and cooled to 0 °C. DBU (64 µL, 0.43 mmol, 0.20 eq) and trichloroacetonitrile (216 µL, 11 mmol, 5 eq.) was added to the solution. After 3 h, more trichloroacetonitrile (108 µL, 5.4 mmol, 2.5 eq.) was added to the reaction mixture. Starting material was fully conversed to imidate **10** with higher running spot on TLC after 21 h stirring at room temperature. Then iodine (1.7 g, 6.7 mmol, 3.1 eq), $NaHCO_3$ (1.8 g, 21 mmol, 10 eq.) and H_2O (3.8 mL) was added to the reaction mixture. After 26 h of the first addition of the iodine and $NaHCO_3$, iodine (1.70 g, 6.68 mmol, 3.1 eq.) and $NaHCO_3$ (1.81 g, 21.5 mmol, 10 eq.) was added to the reaction mixture. After approximately 96 h after the first addition of iodine and $NaHCO_3$, the reaction was quenched by adding $Na_2S_2O_3$ (10% aq. solution) until the purple color had disappeared. The reaction mixture was separated out and the DCM layer was concentrated *in vacuo*, redissolved in EtOAc and washed with H_2O and brine. The initial aq. layer was extracted with EtOAc and the combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. The product was purified by silica gel column chromatography (0%→16%, EtOAc in pentane) to afford **11** (455 mg, 0.88 mmol, 41%) was produced. TLC: R_f 0.57 (PE/EtOAc, 4/1, v/v); ^{1}H-NMR (400 MHz, $CDCl_3$): δ ppm 7.49 – 7.44 (m, 2H), 7.49 – 7.33 (m, 3H), 5.62 (s, 1H), 5.22 (t, *J* = 7.5 Hz, 1H), 4.85 – 4.83 (m, 1H), 4.72 (s, 1H), 4.21 (dd, *J* = 11.3, 4.7 Hz, 1H), 4.03 (t, *J* = 9.8 Hz, 1H), 3.92 (t, *J* = 10.6 Hz, 1H), 3.83 – 3.79 (m, 1H), 3.08 (s, 1H), 1.20 – 1.13 (m, 1H); ^{13}C-NMR (100 MHz, $CDCl_3$): δ ppm 163.72, 137.29, 129.48, 128.51, 126.26, 101.64, 86.69, 77.77, 75.19, 74.93, 72.57, 34.90, 23.69; HRMS: calculated for $C_{16}H_{15}Cl_3NO_4$ [M+H$^+$] 519.91567, found: 519.91504.



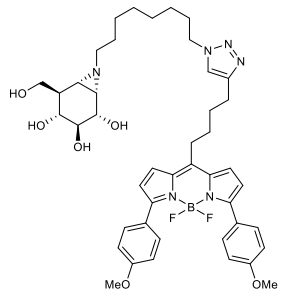
(1S,2R,3S,4R,5R,6R)-5-(Hydroxymethyl)-7-aza-bicyclo-[4.1.0]heptane-2,3,4-triol (1) A solution of **11** (455 mg, 0.88 mmol, 1.0 eq.) was dissolved in 1, 4-dioxane (9.2 mL) and heated to 60 °C. Then aq. HCl (37%, 2.60 mL) was added to the solution. The reaction mixture was stirred at 60 °C overnight. The reaction mixture was concentrated *in vacuo* and then separated out with EtOAc and H_2O. The aq. layer was washed with EtOAc and concentrated *in vacuo* and co-evaporated with toluene. The crude product of the free amine intermediate **12** was dissolved in MeOH (30 mL) and $NaHCO_3$ (2.9 g, 35 mmol, 40 eq.) was added to the solution. The reaction mixture was stirred at room temperature for 4 days. The reaction mixture filtered and then concentrated *in vacuo*. The residue was redissolved in H_2O and filtered over a pad of Amberlite IR-120 H$^+$ resin, washed with H_2O and followed by 1.0 M aq. NH_4OH. The filtrate was concentrated *in vacuo* to afford the aziridine product **1** (97 mg, 0.56 mmol, 63%) as light brown oil. ^{1}H-NMR (400 MHz, D_2O): δ ppm 4.11 – 4.04 (m, 2H), 3.97 – 3.90 (m, 1H), 3.55 – 3.39 (m, 2H), 2.82 – 2.80 (m, 1H), 2.57 (d, *J* = 6.4 Hz, 1H), 2.11 – 2.01 (m, 1H). ^{13}C-NMR (100 MHz, D_2O): δ ppm 73.46, 71.27, 70.07, 61.50, 44.37, 35.66, 31.69; HRMS: calculated for $C_7H_{13}NO4$ [M+H$^+$] 176.09173, found: 176.09163.



(1S,2R,3S,4R,5R,6R)-2,3,4-trihyrdoxy-5-(hydroxymethyl)-7-(8-azidooctyl)-7-aza-bicyclo[4.1.0]heptane (2) A solution of **1** (97 mg, 0.56 mmol, 1.0 eq.) in dry DMF (2.2 mL) was made under argon atmosphere and heated to 80 °C. 1-azido-8-iodooctane (234 mg, 0.83 mmol, 1.5 eq) and K_2CO_3 (345 mg, 2.5 mmol, 4.5 eq.) were added to the solution. After 21 h stirring, the reaction was quenched with MeOH (0.15 mL) and filtered over a pad of celite. The filtrate was concentrated *in vacuo*. The product was purified by silica gel column chromatography (5%→18% MeOH in DCM) to afford product **2** CF022 as yellow oil (72 mg, 0.22 mmol, 39%). TLC: R_f 0.39 (DCM/MeOH, 3/1, v/v); $[\alpha]_D^{21}$ +25.4 (*c* = 1, MeOH); ^{1}H-NMR (850 MHz, CD_3OD): *δ* ppm 3.87 (dd, *J* = 10.7, 4.0 Hz, 1H), 3.66 (dd, *J* = 8.6, 3.7 Hz, 1H), 3.63 (dd, *J* = 10.8, 7.1 Hz, 1H), 3.34 – 3.31 (m, 1H), 3.29 (t, *J* = 6.8 Hz, 2H), 3.05 (t, *J* = 10.0 Hz, 1H), 2.36 – 2.33 (m, 1H), 2.17 – 2.14 (m, 1H), 1.86 – 1.82 (m, 2H), 1.68 (d, *J* = 6.5 Hz, 1H), 1.60 – 1.57 (m, 4H), 1.40 – 1.32 (m, 8H); ^{13}C-NMR (214 MHz, CD_3OD): *δ* ppm 75.76, 73.37, 72.45, 63.52, 62.25, 52.46, 46.89, 46.00, 41.95, 30.58, 30.48, 30.19, 29.90, 28.35, 27.77; LC/MS: R_t 4.42 min, linear gradient 10%→90% B in 12.5 min; ESI-MS: *m/z* = 329.20 $(M+H)^+$; HRMS: calculated for $C_{15}H_{28}N_4O_5$ $[M+H^+]$ 329.21833, found: 329.21828.

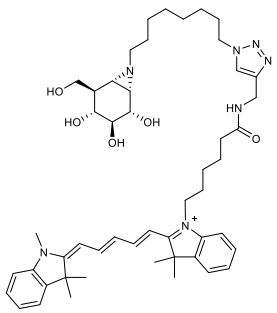


(1S,2S,3S,4R,5R,6S)-7-(8-(4-(4-(5,5-difluoro-1,3,7,9-tetramethyl-5H-4l4,5l4-dipyrrolo[1,2-c:2',1'-f][1,3,2]diazaborinin-10-yl)butyl)-1H-1,2,3-triazol-1-yl)octyl)-5-(hydroxymethyl)-7-azabicyclo[4.1.0]heptane-2,3,4-triol (3) A solution of azide compound **2** (17 mg, 0.052 mmol, 1.0 eq.) in DMF (1.5 mL) was made under argon atmosphere, green BODIPY compound **13** (19 mg, 0.058 mmol, 1.1 eq), $CuSO_4$ (1.0 M in H_2O, 10 µL, 0.010 mmol, 0.20 eq.) and sodium ascorbate (1.0 M in H_2O, 11 µL, 0.011 mmol, 0.21 eq) were added to the solution and the mixture was stirred at room temperature for 12 h. The resulting mixture was concentrated under reduced pressure and the crude product was purified by semi-preparative reversed HPLC (linear gradient: 45%→53% B in A, 3 CV, solutions used A: 50 mM NH_4HCO_3 in H_2O, B: MeCN) and lyophilized to afford product **3** as orange solid (13 mg, 20 µmol, 39%). ^{1}H-NMR (850 MHz, CD_3OD): *δ* ppm 7.73 (s, 1H), 6.11 (s, 2H), 4.34 (t, *J* = 7.0 Hz, 2H), 3.87 (dd, *J* = 8.7, 3.9 Hz, 1H), 3.65 (dd, *J* = 8.5, 3.7 Hz, 1H), 3.61 (dd, *J* = 10.7, 7.2 Hz, 1H), 3.31 – 3.30 (m, 1H), 3.05 (t, *J* = 10.0 Hz, 1H), 3.02 – 2.97 (m, 2H), 2.77 (t, *J* = 7.3 Hz, 2H), 2.43 (s, 6H), 2.37 (s, 6H), 2.34 – 2.26 (m, 1H), 2.15 – 2.11 (m, 1H), 1.91 – 1.81 (m, 6H), 1.67 – 1.62 (m, 3H), 1.57 – 1.53 (m, 2H), 1.36 – 1.22 (m, 8H).; ^{13}C-NMR (214 MHz, CD_3OD) *δ* ppm 154.90, 148.49, 147.88, 142.18, 132.57, 123.36, 122.60, 75.74, 73.34, 72.41, 63.50, 62.19, 51.20, 46.86, 45.99, 41.93, 32.23, 31.24, 30.83, 30.42, 29.88, 29.05, 28.22, 27.32, 25.87, 16.48, 14.44; LC/MS: R_t 5.88 min, linear gradient 10%→90% B in 12.5 min; ESI-MS: *m/z* = 657.33 $(M+H)^+$, 679.53 $[M+Na]^+$; HRMS: calculated for $C_{34}H_{51}BF_2N_6O_4$ $[M+H^+]$ 657.41117, found: 657.41033.

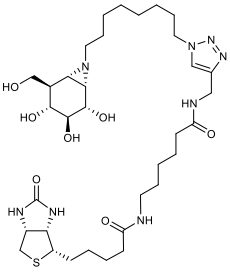


(1S,2S,3S,4R,5R,6S)-7-(8-(4-(4-(5,5-difluoro-3,7-bis(4-methoxyphenyl)-5H-4 ^{4},5 ^{4}-dipyrrolo[1,2-c:2',1'-f][1,3,2]diazaborinin-10-yl)butyl)-1H-1,2,3-triazol-1-yl)octyl)-5-(hydroxymethyl)-7-azabicyclo[4.1.0]heptane-2,3,4-triol (4) A solution of azide compound **2** (18 mg, 0.055 mmol, 1.0eq.) in DMF (1.5 mL) was made under argon atmosphere, red BODIPY compound **14** (29 mg, 0.060 mmol, 1.1 eq), $CuSO_4$ (1.0 M in H_2O, 11 µL, 0.011 mmol, 0.20 eq.) and sodium ascorbate (1.0 M in H_2O, 12 µL, 0.012 mmol, 0.21 eq) were added to the solution and the mixture was stirred at

room temperature for 14 h. The resulting mixture was concentrated under reduced pressure and the crude product was purified by semi-preparative reversed HPLC (linear gradient: 51%→57% B in A, 3 CV, solutions used A: 50 mM NH_4HCO_3 in H_2O, B: MeCN) and lyophilized to afford product **4** as purple solid (5.1 mg, 6.3 µmol, 12%). ^{1}H-NMR (850 MHz, CD_3OD): δ ppm 7.88 – 7.79 (m, 4H), 7.68 (s, 2H), 7.43 (d, *J* = 4.3 Hz, 3H), 7.02 – 6.94 (m, 4H), 6.69 (d, *J* = 4.3 Hz, 2H), 4.32 (t, *J* = 7.1 Hz, 2H), 3.86 – 3.84(m, 7H), 3.64 (dd, *J* = 8.5, 3.7 Hz, 1H), 3.60 (dd, *J* = 10.7, 7.1 Hz, 1H), 3.32 – 3.31 (m, 1H), 3.07 – 3.03 (m, 3H), 2.78 (t, *J* = 6.9 Hz, 2H), 2.29 – 2.26 (m, 1H), 2.12 – 2.09 (m, 1H), 1.88 – 1.79 (m, 6H), 1.64 (d, *J* = 6.5 Hz, 1H), 1.55 – 1.49 (m, 2H), 1.30 – 1.20 (m, 8H); ^{13}C-NMR (214 MHz, CD_3OD): δ ppm 162.20, 158.79, 148.60, 146.78, 137.49, 132.16, 132.14, 132.12, 128.42, 126.51, 123.24, 121.03, 114.85, 114.63, 75.75, 73.36, 72.46, 63.53, 62.18, 55.83, 51.24, 46.86, 45.97, 41.92, 34.13, 31.23, 30.98, 30.40, 30.33, 29.87, 28.20, 27.33, 25.78; LC/MS: R_t 6.82 min, linear gradient 10%→90% B in 12.5 min; ESI-MS: *m/z* = 813.33 $(M+H)^+$; HRMS: calculated for $C_{44}H_{55}BF_2N_6O_6$ $[M+H^+]$ 813.43245, found: 813.43137.

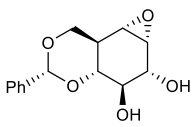


3,3-dimethyl-1-(6-oxo-6-(((1-(8-((1S,2S,3S,4R,5R,6S)-2,3,4-trihydroxy-5-(hydroxymethyl)-7-azabicyclo[4.1.0]heptan-7-yl)octyl)-1H-1,2,3-triazol-4-yl)methyl)amino)hexyl)-2-((1E,3E)-5-((E)-1,3,3-trimethylindolin-2-ylidene)penta-1,3-dien-1-yl)-3H-indol-1-ium (5) A solution of azide compound **2** (17 mg, 0.052 mmol, 1.0 eq.) in DMF (1.5 mL) was made under argon atmosphere, Cy5 compound **15** (29 mg, 0.057 mmol, 1.1 eq), $CuSO_4$ (1.0 M in H_2O, 10 µL, 0.010 mmol, 0.20 eq.) and sodium ascorbate (1.0 M in H_2O, 11 µL, 0.011 mmol, 0.21 eq) were added to the solution and the mixture was stirred at room temperature for 12h. The resulting mixture was concentrated under reduced pressure and the crude product was purified by semi-preparative reversed HPLC (linear gradient: 40%→70% B in A, 3 CV, solutions used A: 50 mM NH_4HCO_3 in H_2O, B: MeCN) and lyophilized to afford product **5** as dark blue powder (10 mg, 12 µmol, 23%). ^{1}H-NMR (600 MHz, CD_3OD): δ ppm 8.32 – 8.16 (m, 2H), 7.84 (s, 1H), 7.55 – 7.48 (m, 2H), 7.41 (m, 2H), 7.35 – 7.21 (m, 4H), 6.62 (t, *J* = 12.4 Hz, 1H), 6.29 – 6.27 (m, 2H), 4.41 (s, 2H), 4.36 (t, *J* = 7.1 Hz, 2H), 4.09 (t, *J* = 7.5 Hz, 2H), 3.86 (dd, *J* = 10.7, 3.9 Hz, 1H), 3.69 – 3.57 (m, 5H), 3.33 – 3.30 (m, 1H), 3.10 – 3.00 (m, 1H), 2.34 – 2.29 (m, 1H), 2.25 (t, *J* = 7.3 Hz, 2H), 2.15 – 2.11 (m, 1H), 1.93 (s, 6H), 1.90 – 1.77 (m, 6H), 1.73 – 1.57 (m, 14H), 1.60 – 1.51 (m, 2H), 1.50 – 1.42 (m, 2H), 1.33 – 1.27 (m, 8H); ^{13}C-NMR (150 MHz, CD_3OD): δ ppm 178.47, 175.73, 175.38, 174.59, 155.54, 155.47, 146.13, 144.23, 143.53, 142.61, 142.51, 129.77, 129.73, 126.60, 126.27, 126.21, 124.15, 123.41, 123.28, 112.01, 111.84, 104.41, 104.23, 75.72, 73.31, 72.39, 63.48, 62.17, 51.33, 50.53, 50.50, 46.87, 46.00, 44.74, 41.92, 36.46, 35.56, 31.50, 31.30, 30.44, 30.41, 29.94, 28.25, 28.12, 27.93, 27.78, 27.37, 27.30, 26.39; LC/MS: R_t 5.18 min, linear gradient 10%→90% B in 12.5 min; ESI-MS: *m/z* = 848.60 $(M)^+$; HRMS: calculated for $C_{50}H_{70}N_7O_5^+$ $[M+H^+]$ 849.54653, found: 849.55112.

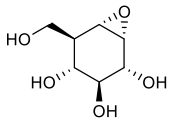


6-(5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamido)-N-((1-(8-((1S,2S,3S,4R,5R,6S)-2,3,4-trihydroxy-5-(hydroxymethyl)-7-azabicyclo[4.1.0]heptan-7-yl)octyl)-1H-1,2,3-triazol-4-yl)methyl)hexanamide (6) A solution of azide compound **2** (15 mg, 0.0457 mmol, 1.0 eq.) in DMF (1.5 mL) was made under argon atmosphere, biotin compound **16** (19.8 mg, 0.0502 mmol, 1.1 eq), $CuSO_4$ (1.0 M in H_2O, 9.1 µL, 0.0091

mmol, 0.20 eq.) and sodium ascorbate (1.0 M in H_2O, 9.6 µL, 0.0096 mmol, 0.21 eq) were added to the solution and the mixture was stirred at room temperature for 12 h. The resulting mixture was concentrated under reduced pressure and the crude product was purified by semi-preparative reversed HPLC (linear gradient: 19%→28% B in A, 12 min, solutions used A: 50 mM NH_4HCO_3 in H_2O, B: MeCN) and lyophilized to afford product **6** as white solid (7.6 mg, 11 µmol, 23%). ^{1}H-NMR (400 MHz, CD_3OD): *δ* ppm 7.85 (s, 1H), 4.49 (dd, *J* = 7.8, 4.8 Hz, 1H), 4.44 – 4.34 (m, 4H), 4.30 (dd, *J* = 7.9, 4.5 Hz, 1H), 3.86 (dd, *J* = 10.7, 3.9 Hz, 1H), 3.69 – 3.58 (m, 2H), 3.34 – 3.30 (m, 1H), 3.25 – 3.10 (m, 3H), 3.05 (t, *J* = 10.0 Hz, 1H), 2.93 (dd, *J* = 12.7, 5.0 Hz, 1H), 2.74 – 2.64 (m, 1H), 2.37 – 2.29 (m, 1H), 2.26 – 2.10 (m, 5H), 1.95 – 1.79 (m, 4H), 1.79 – 1.40 (m, 14H), 1.37 – 1.30 (m, 10H); ^{13}C-NMR (100 MHz, CD_3OD): *δ* ppm 176.01, 166.11, 161.52, 146.22, 124.16, 75.74, 73.34, 72.41, 63.50, 63.38, 62.20, 61.62, 57.02, 51.35, 46.89, 46.01, 41.95, 41.05, 40.18, 36.81, 36.75, 35.58, 31.29, 30.45, 30.12, 29.95, 29.78, 29.50, 28.26, 27.54, 27.37, 26.93, 26.52; LC/MS: R_t 4.37 min, linear gradient 10%→90% B in 12.5 min; ESI-MS: *m/z* = 723.47 $(M+H)^+$; HRMS: calculated for $C_{34}H_{58}N_8O_7S$ $[M+H^+]$ 723.42245, found: 723.42219.



(1aR,2R,3R,3aR,7aS,7bS)-5-phenylhexahydro-2H-oxireno[2',3':3,4]benzo[1,2-d][1,3]dioxine-2,3-diol (17) A solution of compound **9** (100 mg, 0.40 mmol, 1.0 eq.) and 3-chloroperbenzoic acid (*m*CPBA, 75%) (135 mg, 0.60 mmol, 1.5 eq) in dry DCM (10 mL) were made under argon atmosphere and refluxed at 40°C for 24h. The reaction mixture was washed by sat. aq. $NaHCO_3$ and brine. The organic layer was dried over Na_2SO_4, filtered and concentrated *in vacuo*. The product was purified via silica gel column chromatography (50%→100%, EtOAc in pentane) to afford the title compound **17** (47 mg, 0.18 mmol, 44%). TLC: R_f 0.30 (pentane/EtOAc, 3/1, v/v); ^{1}H-NMR (400 MHz, $CDCl_3$): *δ* ppm 7.49 – 7.46 (m, 2H), 7.37 – 7.33 (m, 3H), 5.48 (s, 1H), 4.38 (dd, *J* = 12.0, 4.0 Hz, 1H), 3.97 – 3.95 (m, 1H), 3.74 – 3.68 (m, 1H), 3.64 – 3.60 (m, 1H), 3.32 – 3.25 (m, 2H), 2.78 (d, *J* = 4.0 Hz, 1H), 2.22 – 2.15 (m, 1H); ^{13}C-NMR (100 MHz, $CDCl_3$): *δ* ppm 137.46, 129.38, 128.45, 126.44, 102.14, 79.79, 73.32, 72.76, 68.37, 55.68, 51.71, 38.00; HRMS: calculated for $C_{14}H_{16}O_5$ $[M+H^+]$ 265.10705, found: 265.10720.



(1R,2R,3S,4R,5R,6S)-5-(hydroxymethyl)-7-oxabicyclo[4.1.0]heptane-2,3,4-triol (7) A mixture of product **17** (20 mg, 0.076 mmol, 1.0 eq.) and $Pd(OH)_2/C$ (15 mg, 20% wt. loading(dry basis)) in MeOH (2.0 mL) was stirred at room temperature under hydrogen atmosphere overnight. The catalyst was then filtered off and washed with MeOH. The filtrate and washings were combined and concentrated under reduced pressure. Crude product was purified by semi-preparative reversed phase HPLC (linear gradient: 0%→20%, 3 CV, solutions used A: 50mM NH_4HCO_3 in H_2O, B: MeCN) and lyophilization to afford the title compound *epi*-cyclophellitol **7** (9.0 mg, 0.051 mmol, 68%) as white solid. TLC: R_f 0.26 (MeOH:$CDCl_3$, 1/3, v/v); ^{1}H-NMR (400 MHz, D_2O): *δ* ppm 3.91 – 3.86 (m, 2H), 3.78 – 3.74 (m, 1H), 3.44 – 3.43 (m, 1H), 3.41 - 3.37 (m, 1H), 3.34 – 3.29 (m, 2H), 2.04 – 1.99 (m, 1H); ^{13}C-NMR (100 MHz, D_2O): *δ* ppm 73.05, 71.23, 69.35, 60.32, 57.47, 55.09, 44.12; HRMS-: calculated for $C_7H_{12}O_5$ $[M+H^+]$ 177.07575, found: 177.07571.

5.5 References

[1] L. H. Hoefsloot, M. Hoogeveen-Westerveld, A. J. Reuser and B. A. Oostra, *Biochem. J.* **1990**, 272, 493-497.

[2] S. Chiba, K. Hiromi, N. Minamiura, M. Ohnishi, T. Shimomura, K. Suga, T. Suganuma, A. Tanaka, S. Tomioka and T. Yamamoto, *J. Biochem*. **1979**, 85, 1135-1141.

[3] V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho and B. Henrissat, *Nucleic acids Res.* **2014**, 42, D490-495.

[4] G. J. Davies and S. J. Williams, *Biochem. Soc. Trans.* **201**6, 44, 79-87.

[5] R. J. Moreland, X. Jin, X. K. Zhang, R. W. Decker, K. L. Albee, K. L. Lee, R. D. Cauthron, K. Brewer, T. Edmunds and W. M. Canfield, *J. Biol. Chem.* **2005**, 280, 6780-6791.

[6] H. A. Wisselaar, M. A. Kroos, M. M. Hermans, J. van Beeumen and A. J. Reuser, *J. Biol. Chem.* **1993**, 268, 2223-2231.

[7] S. S. Lee, S. He and S. G. Withers, *Biochem. J*. **2001**, 359, 381-386.

[8] D. E. Koshland, *Biol. Rev.* **1953**, 28, 416-436.

[9] H. G. Hers, *Biochem. J.***1963**, 86, 11-16.

[10] A. J. Reuser, M. A. Kroos, M. M. Hermans, A. G. Bijvoet, M. P. Verbeet, O. P. Van Diggelen, W. J. Kleijer and A. T. Van der Ploeg, *Muscle Nerve. Suppl.* **1995**, 3, S61-69.

[11] P. S. Kishnani, W. L. Hwu, H. Mandel, M. Nicolino, F. Yong and D. Corzo, *J. Pediatr.* **2006**, 148, 671-676.

[12] L. P. Winkel, M. L. Hagemans, P. A. van Doorn, M. C. Loonen, W. J. Hop, A. J. Reuser and A. T. van der Ploeg, *J. Neurol.* **2005**, 252, 875-884.

[13] P. S. Kishnani, D. Corzo, M. Nicolino, B. Byrne, H. Mandel, W. L. Hwu, N. Leslie, J. Levine, C. Spencer, M. McDonald, J. Li, J. Dumontier, M. Halberthal, Y. H. Chien, R. Hopkin, S. Vijayaraghavan, D. Gruskin, D. Bartholomew, A. van der Ploeg, J. P. Clancy, R. Parini, G. Morin, M. Beck, G. S. De la Gastine, M. Jokic, B. Thurberg, S. Richards, D. Bali, M. Davison, M. A. Worden, Y. T. Chen and J. E. Wraith, *Neurology* **2007**, 68, 99-109.

[14] A. T. van der Ploeg, P. R. Clemens, D. Corzo, D. M. Escolar, J. Florence, G. J. Groeneveld, S. Herson, P. S. Kishnani, P. Laforet, S. L. Lake, D. J. Lange, R. T. Leshner, J. E. Mayhew, C. Morgan, K. Nozaki, D. J. Park, A. Pestronk, B. Rosenbloom, A. Skrinar, C. I. van Capelle, N. A. van der Beek, M. Wasserstein and S. A. Zivkovic, *N. Engl. J. Med.* **2010**, 362, 1396-1406.

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

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

[17] J. Jiang, T. J. M. Beenakker, W. W. Kallemeijn, G. A. van der Marel, H. van den Elst, J. D. C. Codee, J. M. F. G. Aerts and H. S. Overkleeft, *Chem. Eur. J.* **2015**, 21, 10861-10869.

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

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

[20] B. T. Adams, S. Niccoli, M. A. Chowdhury, A. N. Esarik, S. J. Lees, B. P. Rempel and C. P. Phenix, *Chem. Commun.* **2015**, 51, 11390-11393.

[21] A. Alcaide, A. Trapero, Y. Perez and A. Llebaria, *Org. Biomol. Chem.* **2015**, 13, 5690-5697.

[22] F. G. Hansen, E.. Bundgaard, R. Madsen, A Short Synthesis of (+)-Cyclophellitol. *J. Org. Chem.* **2005**, 70, 10139-10142..

[23] A. Ghisaidoobe, P. Bikker, A. C. de Bruijn, F. D. Godschalk, E. Rogaar, M. C. Guijt, P. Hagens, J. M. Halma, S. M. Van't Hart, S. B. Luitjens, V. H. van Rixel, M. Wijzenbroek, T. Zweegers, W. E. Donker-Koopman, A. Strijland, R. Boot,. G. van der Marel, H. S. Overkleeft, J. M. Aerts, R. J. van den Berg. Identification of Potent and Selective Glucosylceramide Synthase Inhibitors from a Library of N-Alkylated Iminosugars. *ACS Med. Chem. Lett.* **2011,** 2, 119-123.

[24] J. Larsbrink, A. Izumi, G. R. Hemsworth, G. J. Davies and H. Brumer, *J. Biol. Chem.* **2012**, 287, 43288-43299.

[25] G. J. Davies, A. Planas and C. Rovira, *Acc. Chem. Res.* **2012**, 45, 308-316.

[26] G. Speciale, A. J. Thompson, G. J. Davies and S. J. Williams, *Curr. Opin. Struct. Biol.* **2014**, 28, 1-13.

[27] B. Poole and S. Ohkuma, *J. Cell Biol.* **1981**, 90, 665-669.

[28] Y. Liu, M. P. Patricelli and B. F. Cravatt, *Proc. Natl. Acad. Sci. U S A* **1999**, 96, 14694-14699.

[29] D. Greenbaum, K. F. Medzihradszky, A. Burlingame and M. Bogyo, *Chem. Biol.* **2000**, 7, 569-581.

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

[31] W. Kabsch, *Acta Crystallogr. D* **2010**, 66, 125-132.

[32] P. R. Evans and G. N. Murshudov, *Acta Crystallogr. D* **2013**, 69, 1204-1214.

[33] A. Vagin and A. Teplyakov, MOLREP: an automated program for molecular replacement. *J Appl. Crystallogr.* **1997**, 30, 1022-1025.

[34] P. Emsley and K. Cowtan, *Acta Crystallogr D* **2004**, 60, 2126-2132.

[35] G. N. Murshudov, A. A. Vagin and E. J. Dodson, *Acta Crystallogr. D* **1997**, 53, 240-255.

[36] J. Painter and E. A. Merritt, *Acta Crystallogr. D* **2006**, 62, 439-450.

[37] A. A. Lebedev, P. Young, M. N. Isupov, O. V. Moroz, A. A. Vagin and G. N. Murshudov, J. Painter and E. A. Merritt, *Acta Crystallogr. D* **2012**, 68, 431-440.

[38] S. McNicholas, E. Potterton, K. S. Wilson and M. E. M. Noble, *Acta Crystallogr. D* **2011**, 67, 386-394.

5.6 Supporting Information

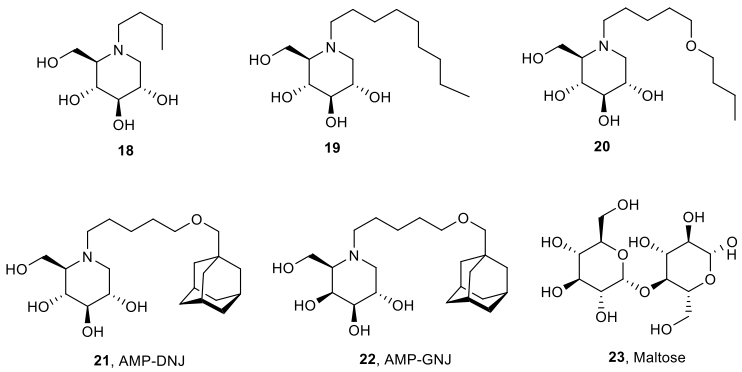


Figure S1. Chemical structures of iminosugar compounds **18**-**22** and maltose **23**.

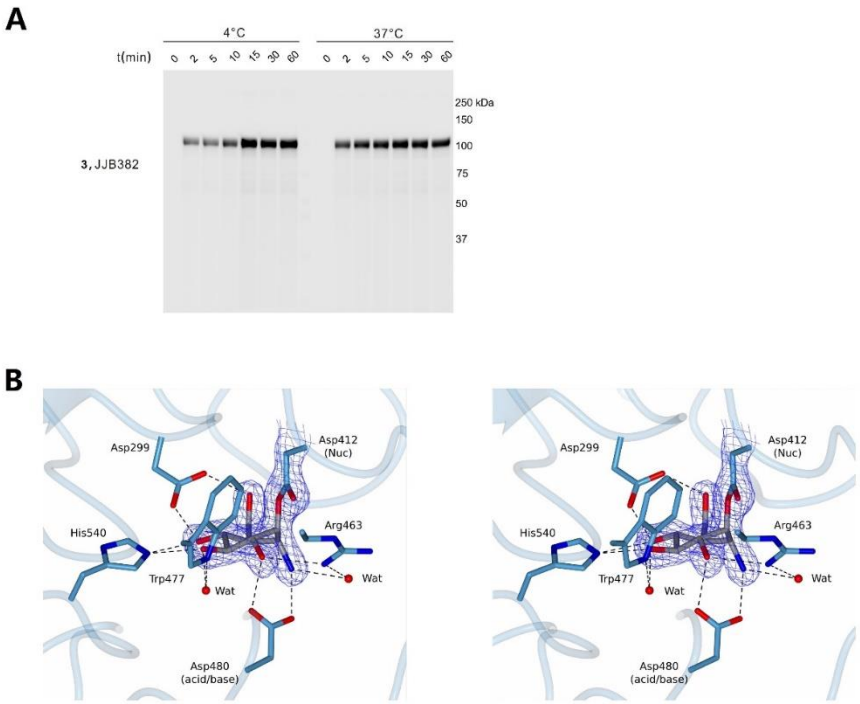


Figure S2. A) Rates determined through direct 4 °C and 37 °C different temperature labeling with **3** JJB382; B) 3-D view of crystal structure of C*j*Agd31B GAA-**1** CF021-complex.

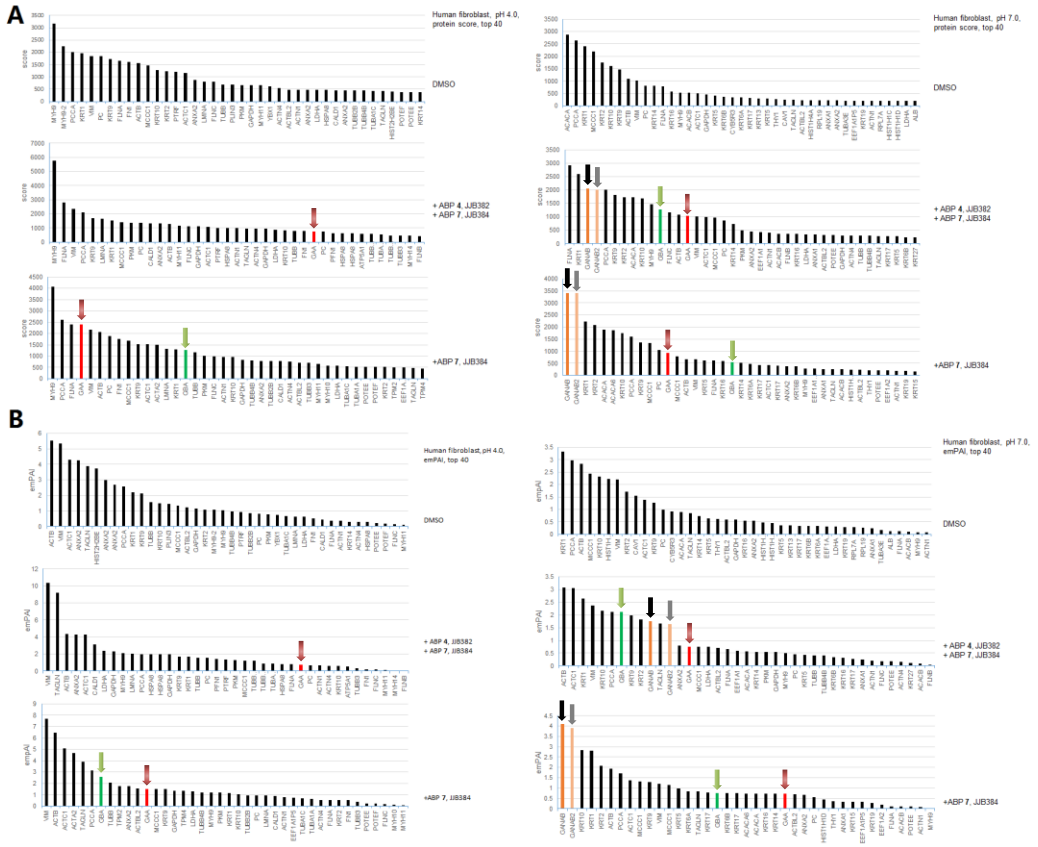
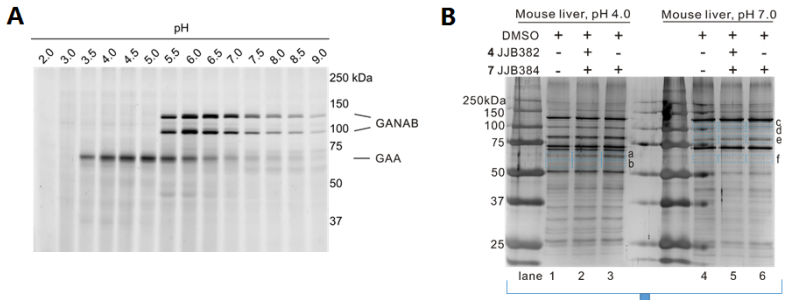


Figure S3. On-bead digest chemical proteomics data of human fibroblast lysate labeling of α-glucosidases. Protein score A) and emPAI value B) of top 40 proteins on- bead digetion analysis using samples of DMSO, ABP **3** competitive inhibition and ABP **6** direct labeling in pH 4.0 and pH 7.0.



Band	Rank	Accession	Protein name	Score	Matches	Coverage(%)	emPAI
2a	2	P70699	Lysosomal alpha-glucosidase, GAA	471	33	18.2	0.44
	4	P17439	Glucosylceramidase, GBA	253	20	19.6	0.47
3a	2	P70699	Lysosomal alpha-glucosidase, GAA	760	40	19.9	0.72
	4	P17439	Glucosylceramidase, GBA	238	21	19.8	0.39
2b	17	P17439	Glucosylceramidase, GBA	139	13	13.4	0.32
3b	13	P17439	Glucosylceramidase, GBA	125	11	15.3	0.18
6c	2	Q8BHN3-2	Isoform 2 of Neutral alpha-glucosidase AB, GANAB-2	909	79	46.8	1.64
6d	2	Q8BHN3	Neutral alpha-glucosidase AB, GANAB	762	62	30.0	1.06
6e	3	Q8BHN3	Neutral alpha-glucosidase AB, GANAB	641	47	23.7	0.82
5f	18	P17439	Glucosylceramidase, GBA	125	11	12.2	0.25
	20	P70699	Lysosomal alpha-glucosidase, GAA	120	8	6.3	0.06
6f	26	P70699	Lysosomal alpha-glucosidase, GAA	67	4	3.8	0.09
1a, 1b, 4c-f, 5c-e	No target enzymes						

Figure S4. A) Mouse liver lysate labeling of α-glucosidases in various pH *in vitro* with ABP **3**. B) In-gel digestion silver staining gel and identification of target proteins modified by biotin-ABP **6** with mass spectrometry and mascot search analysis, comparing with DMSO and ABP **3** competitive inhibition samples.

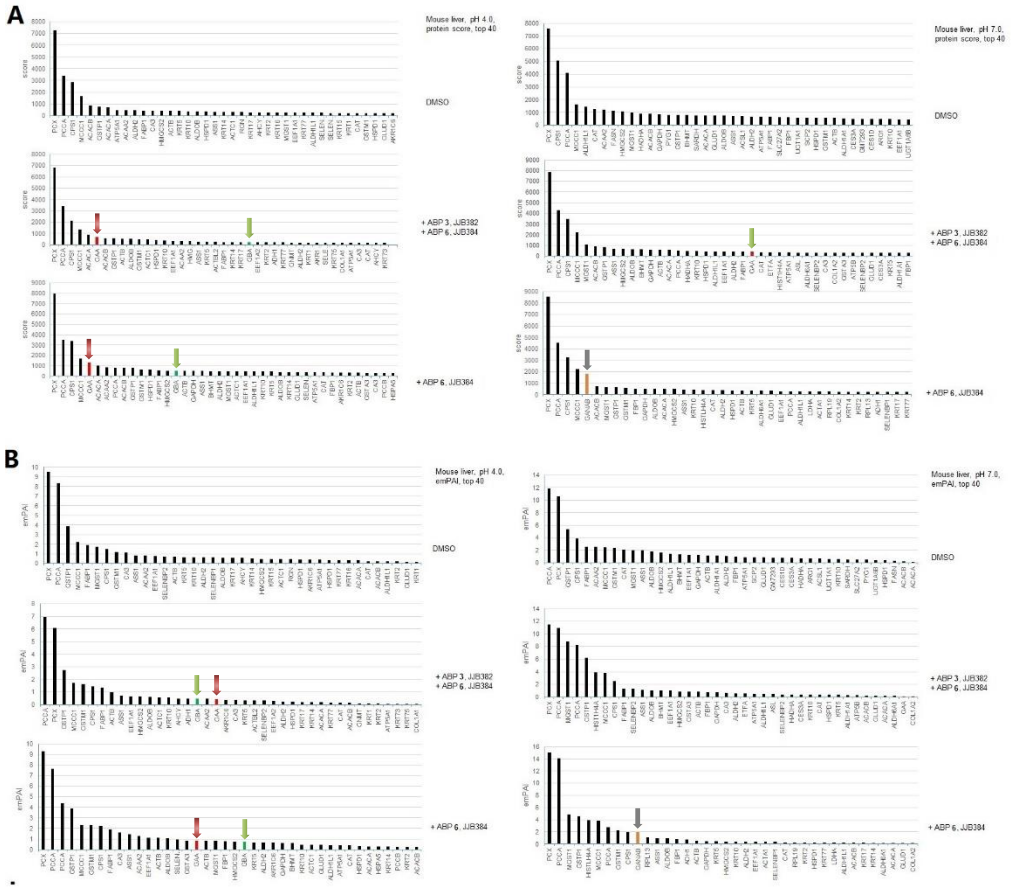
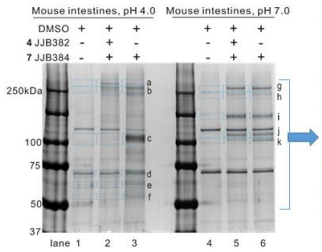


Figure S5. On-bead digest chemical proteomics data of mouse liver lysate labeling of α-glucosidases. Protein score A) and emPAI value B) of top 40 proteins on-bead digestion analysis using samples of DMSO, ABP **3** competitive inhibition and ABP **6** direct labeling in pH 4.0 and pH 7.0.

A



Band	Rank	Accession	Protein name	Score	Matches	Coverage(%)	emPAI
2a	4	B5THE2	Maltase-glucoamylase, MGAM	312	35	12.3	0.18
	30	B5THE3	Sucrase-isomaltase, Sis	43	5	2.5	0.01
3a	19	B5THE2	Maltase-glucoamylase, MGAM	160	12	4.5	0.07
	20	B5THE3	Sucrase-isomaltase, Sis	151	11	4.2	0.06
3b	16	B5THE3	Sucrase-isomaltase, Sis	292	29	12.3	0.13
	31	B5THE2	Maltase-glucoamylase, MGAM	125	11	15.3	0.18
3c	1	B5THE3	Sucrase-isomaltase, Sis	1260	51	14.4	0.37
	33	B5THE2	Maltase-glucoamylase, MGAM	129	9	3.4	0.04
3d	12	B5THE3	Sucrase-isomaltase, Sis	211	19	10.3	0.09
	23	B5THE2	Maltase-glucoamylase, MGAM	124	9	4.9	0.04
	29	A2AFL5	Lysosomal alpha-glucosidase, GAA	55	1	6.2	0.09
3e	12	B5THE3	Sucrase-isomaltase, Sis	309	20	6.5	0.1
	17	B5THE2	Maltase-glucoamylase, MGAM	241	12	4.5	0.07
3f	7	B5THE2	Maltase-glucoamylase, MGAM	505	30	12.8	0.2
	20	B5THE3	Sucrase-isomaltase, Sis	265	19	8	0.1
	72	P17439	Glucosylceramidase, GBA	28	2	2.9	0.05
5g	1	B5THE3	Sucrase-isomaltase, Sis	468	45	13.9	0.28
6g	4	B5THE3	Sucrase-isomaltase, Sis	421	47	15.3	0.3
5h	18	B5THE3	Sucrase-isomaltase, Sis	190	9	3.7	0.06
6h	1	B5THE3	Sucrase-isomaltase, Sis	302	28	10.3	0.18
5i	6	F8VPT3	Protein Lct, Lct	567	54	13.3	0.27
6i	1	F8VPT3	Protein Lct, Lct	549	65	14.1	0.32
5j	11	B5THE3	Sucrase-isomaltase, Sis	553	36	11.8	0.23
	43	B5THE2	Maltase-glucoamylase, MGAM	53	8	2.9	0.03
6j	2	B5THE3	Sucrase-isomaltase, Sis	605	52	16.1	0.3
5k	2	B5THE3	Sucrase-isomaltase, Sis	551	33	11.2	0.24
6k	1	Q8BHN3	Neutral alpha-glucosidase AB, GANAB	683	65	34.4	1.0
	2	B5THE3	Sucrase-isomaltase, Sis	614	39	12.3	0.23
1a-f, 2b-f, 4g-k	No target enzymes						

B

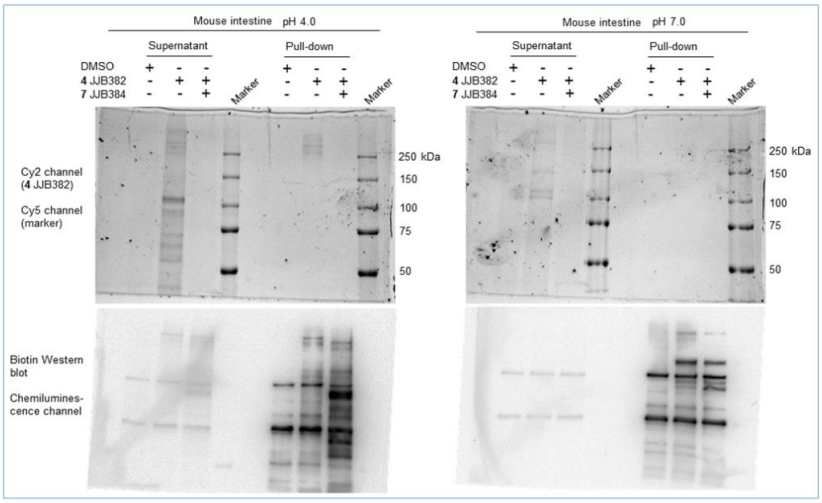


Figure S6. Mouse intestine lysate labeling of α-glucosidases and their identification by proteomics. A) In-gel digestion silver staining gel and identification of target proteins modified by biotin-ABP **7** with mass spectrometry and mascot search analysis, comparing with DMSO and ABP **3** competitive inhibition samples. B) Pull-down and supernatant samples labeling of α-glucosidases in pH 4.0 and pH 7.0 by fluorescent SDS-PAGE scanning for ABP **3** and Western blotting for ABP **6** analysis

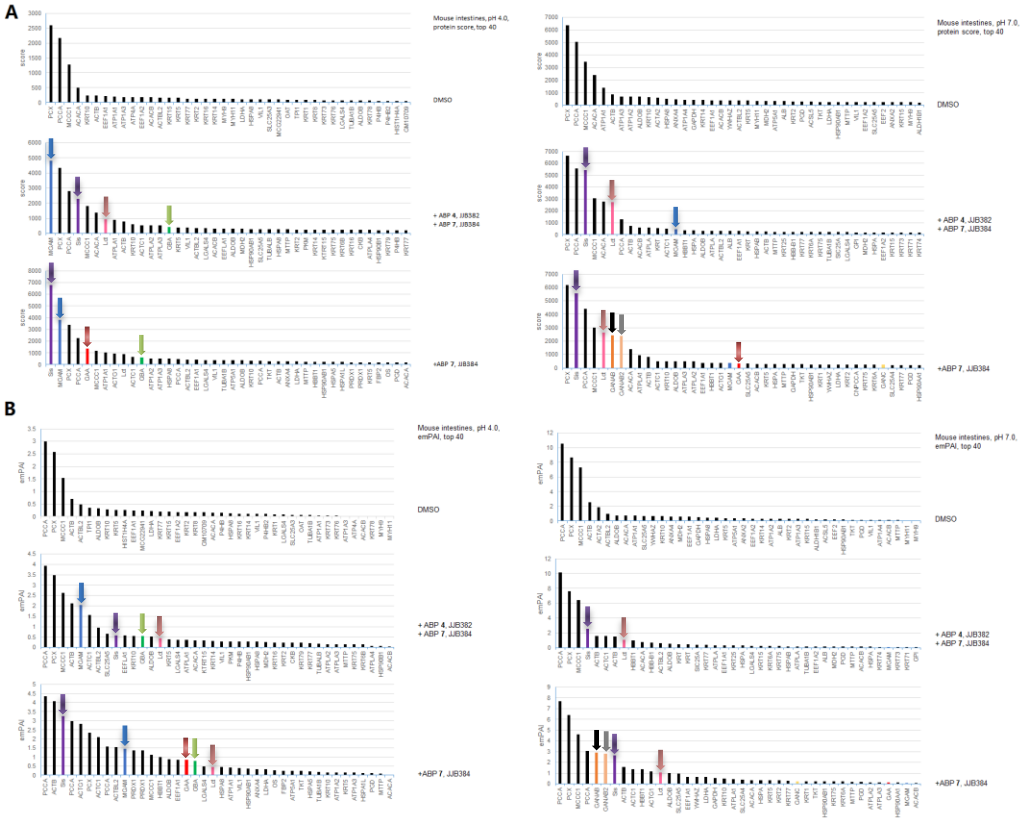


Figure S7. On-bead digest chemical proteomics data of mouse intestine lysate labeling of α-glucosidases. Protein score A) and emPAI value B) of top 40 proteins on-bead digestion analysis using samples of DMSO, ABP **3** competitive inhibition and ABP **6** direct labeling in pH 4.0 and pH 7.0.